Mol Genet Genomics (2002) 266: 873–881
DOI 10.1007/s00438-001-0610-0

ORIGINAL PAPER

T. Atlung · H.V. Nielsen · F.G. Hansen

Characterisation of the allelic variation in the *rpoS* gene in thirteen K12 and six other non-pathogenic *Escherichia coli* strains

Received: 6 July 2001 / Accepted: 18 October 2001 / Published online: 21 November 2001
© Springer-Verlag 2001

Abstract The nucleotide sequence of *rpoS,* the gene for the stress sigma factor, was determined in 13 different K12 strains of *Escherichia coli*. The results indicate that the original K12 isolate carried an amber mutation at codon 33, which in 50% of the derivatives is mutated by a single base substitution to a coding triplet, in most cases to CAG encoding glutamine. The six non-K12 strains examined here had GAG, encoding glutamate, in position 33. The two most divergent strains had three and seven neutral substitutions in *rpoS* and carried insertions of 2100 and 2900 bp, respectively, just downstream of the gene. The genetic variations in *rpoS* were compared with the variation in RpoS-related phenotypes, by measuring catalase (KatE) activity, glycogen accumulation and acid phosphatase levels, and a *katEp*-*gfp* fusion was used to visualise *katE* gene transcription. The RpoS phenotypes of the six *rpoS*(33E) strains varied significantly more than that of the K12 *rpoS*(33Q) strains, especially with respect to acid phosphatase levels. This was due to the absence of the gene for the transcriptional activator AppY from four of the *rpoS*(33E) strains, while all the K12 derivatives carried this gene. When cloned into a LacI-controlled vector and compared in a *rpoS*::Tn*10* background, the RpoS(33Q) and RpoS(33E) variants showed the same activity.

Keywords Amber mutation · Gene insertions/deletions · *gfp* fusion · Evolution

Communicated by W. Goebel

T. Atlung (✉) · H.V. Nielsen
Department of Life Sciences and Chemistry,
Roskilde University, DK-4000 Roskilde, Denmark
E-mail: atlung@ruc.dk
Tel.: +45-46-742402
Fax: +45-46-743011

F.G. Hansen
BioCentrum-DTU, Section of Molecular Microbiology,
Building 301, Technical University of Denmark,
DK-2800 Lyngby, Denmark

Introduction

The normal habitat of the facultative anaerobic bacterium *Escherichia coli* is the gastrointestinal tract of animals. Extensive reprogramming of the pattern of gene expression enables *E. coli* to survive the passage through faeces and to stay alive in marine and freshwater environments. Perhaps the most important factor in this adaptation is the RNA polymerase initiation factor σ^S, encoded by the gene *rpoS*, which is required for efficient transcription of numerous genes involved in adaptation to, and survival under, non-optimal conditions (Hengge-Aronis 1996). *rpoS* mutants exhibit a pleiotropic phenotype, with two classical features being sensitivity to H_2O_2 due to low expression of the *katE* gene encoding catalase HPII, and low acid phosphatase (AppA) levels. These phenotypes gave rise to the previous gene designations *katF* and *appR* (Touati et al. 1986, 1991; Mulvey and Loewen 1989). Other phenotypes of *rpoS* mutants include deficiencies in the accumulation of glycogen due to reduced expression of the regulatory gene *glgS* (Hengge-Aronis and Fischer 1992) and altered cell size and shape in stationary phase due to an effect on the regulatory morphogene *bolA* (Lange and Hengge-Aronis 1991a).

The *rpoS* gene is required for prolonged survival in stationary phase both in rich and minimal medium (Lange and Hengge-Aronis 1991b), for stationary phase-induced acid resistance (Small et al. 1994), for thermotolerance (Hengge-Aronis et al. 1991), and for adaptation to growth in medium of high osmolarity (Hengge-Aronis et al. 1993). Recently, it was shown that RpoS is needed for long-term survival of *E. coli* in illuminated seawater, mainly due to the effect of RpoS on expression of catalase (Gourmelon et al. 1997).

The cellular concentration of RpoS is subject to large variations as a function of growth conditions. Regulation of RpoS levels occurs primarily at the post-transcriptional level, and is partly mediated through the rate of degradation (Muffler et al. 1996a). Degradation of the protein is controlled by a complex regulatory mechanism involving

a His-Asp phophostransfer response-regulator, which might also directly regulate the activity of RpoS (Muffler et al. 1996b; Pratt and Silhavy 1996; Becker et al. 1999).

While there is general agreement about the main phenotypic differences between wild type and *rpoS* mutants of *E. coli*, it is not clear what the wild-type *rpoS* sequence in *E. coli* K12 actually is. Three different amino acid sequences have been reported for different purportedly wild-type K12 strains in five publications (Mulvey and Loewen 1989; Ivanova et al. 1992; Tanaka et al. 1993; Blattner et al. 1997; Jishage and Ishihama 1997), four of which are available in GenBank (Accession Nos. X16400, Z14969, D13548, and U00096); three other sequenced K12 strains have amber mutations in the gene, primarily at codon 33 (Ivanova et al. 1992; Jishage and Ishihama 1997; Visick and Clarke 1997). Furthermore, *rpoS* amber mutations may be present in many other K12 strains, in view of the finding that trehalose accumulation in these strains depends on the presence of amber suppressors (Rod et al. 1988) and trehalose accumulation requires expression of the RpoS-controlled *ots* genes (Hengge-Aronis et al. 1991).

We undertook the present study with the aim of finding the true wild-type *rpoS* sequence of *E. coli* in general, and K12 in particular, and, furthermore, of extending the studies of *rpoS* mutations in K12 strains by including strains used in investigations of the AppR phenotype (Touati et al. 1986).

Materials and methods

Bacterial strains, growth media, and enzyme measurements

Strains used in this study are listed in Table 1. Cultures were grown in Luria-Bertani (LB) medium (Miller 1972) for genetic experiments and for determination of acid phosphatase activity as described previously (Atlung et al. 1989). Catalase activity was estimated from the rapidity and degree of bubbling when a drop of H_2O_2 was applied to bacteria grown overnight on LB plates at 37°C. Glycogen accumulation was determined on MOPS minimal medium (Neidhardt et al. 1974) plates containing limiting NH_4Cl (0.95 mM), supplemented with 1% glucose, 1 μg/ml thiamine, 10 μg/ml uracil, 100 μg/ml thymidine, 10 μg/ml adenine, and a synthetic amino acid mixture (FN20) at half the concentrations suggested by Neidhardt et al. (1974). Glycogen was visualised by evaporation of I_2 onto the plate. AB minimal medium (Clark and Maaløe 1967) supplemented with 1% casamino acids (Difco), 0.2% glucose and 1 μg/ml thiamine was used for growth of bacterial strains in experiments in which synthesis of RpoS was induced with IPTG (Apollo Scientific Limited) and *bolAp1-lacZ* expression was measured. Cell density was monitored and β-galactosidase activity was determined as described previously (Brøndsted and Atlung 1994). Amounts of RpoS protein were determined in cultures induced with 0.1 and 0.2 mM IPTG in both exponentially growing cultures and in early stationary phase cultures. Sample preparation and immunoblot analysis was carried out as described previously

Table 1 *E. coli* strains

Strain	Type	Genotype[a]	Source/reference
MC4100	K12	*araD139*, Δ(*argF-lac*)*U169*, *deoC1*, *flb5301*, *relA1*, *rpsL150*, *ptsF25*, *rbsR*	Silhavy et al. (1984)
RH90	K12	MC4100 *rpoS359*::Tn*10*	Lange et al. (1991b)
KL99	K12	Hfr, *thi-1*, *relA1*, *spoT1*, *lac-42*	Cold Spring Harbor Laboratory
SBS1028	K12	*thi1*, *lysA22*, *argA21*, *cysC43*, *pheA97*, *thyA61*, *relA*, *rpsE339*, *srl*::Tn*10*	P. L. Boquet
SBS1029	K12	SBS1028 *rpoS*(33Y ΔC1239)	P. L. Boquet
MT102	K12	*thi*, Δ(*ara-leu*)*7679*, *araD139*, *lac*Δ*X74*, *galU*, *galK*, *rpsL*, *hsdRK12*[b]	M. T. Hansen
KL16	K12	Hfr, *thi-1*, *relA1*, *spoT1*	B. Bachmann
LJ24	K12	*thi-1*, *leu-6*, *lacY1*, *lacI-Z*Δ(*Mlu*), *tonA21*, *rpsL*, *rfbD1*, *supE44*, *rpoS*(33am)[c]	Rasmussen et al. (1991)
FH1218	K12	$\lambda(c^+)$ *thi-1*, *trp-3*, *his-4*, *pyrB*::Tn*5*, *galK2*, *lacY1*, *mtl-1*, *ara-9*, *tsx-3*, *ton-1*, *rpsL8* or *9*, *supE44*, *rpoS*(33am)	Løbner-Olesen et al. (1989)
EMG2	K12	F^+, $\lambda(c^+)$, *rpoS*(33am)	Bachmann (1972), via A. Wright
CSH61	K12	HfrC, *trp*, *thi*, *rpoS*(33am)	Miller (1972)
UTH1038	K12	*rpoB369*(Fts) *rpoS*(33am)[d]	Goldsmith et al. (1970)
CSH57	K12	*thi*, *purE*, *leu*, *trp*, *his*, *argG*, *ilv*, *metA* (or *metB*), *rpsL*, *ara*, *lacY*, *gal*, *malA*, *xyl*, *mtl*, *rpoS*(33am R81S)	Miller (1972)
Hfr6	K12	Hfr, *metB1*, *relA*, *mut2*, *mtl8*, *malB20*, *rpoS*(33S)	Cold Spring Harbor Laboratory
Cp14	B/r	*rpoS*(33E)	Zeuthen and Paro (1971)
WM301	B/r	*leu-19*, *pro-19*, *trp-25*, *his-47*, *arg-28*, *metB55*, *thyA59*, *deoB23*, *lac-11*, *gal-11*, *rpsL56*, *sul1*, *hsdSK12*, *rpoS*(33E)	Zahn and Messer (1979)
JG151	$15T^-$	*thyA* *deoB* *rpoS*(33E)	Donachie (1968)
BR21	BJ4	*rpsL*(Str^R) *rpoB*(Rif^R) *rpoS*(33E)	Krogfelt et al. (1993)
F18		*rpoS*(33E)	Wadolkowski et al. (1988)
C-1005	C	F^+, *thr*, *leu*, *xan*, *his*, *rpsL*, *rpoS*(33E, Δ[643–653])	From G. Bertani (Calendar et al. 1970)
RH95	K12	MC4100, λMAV103 *bolAp1*::*lacZYA* Km^R	Lange et al. (1991a)
TC4392	K12	MT102 *rpoS359*::Tn*10*	P1(RH90) × MT102
TC4545	K12	MT102 *rpoS359*::Tn*10*, λMAV103 *bolAp1*::*lacZ* Km^R	P1(RH95) × TC4392

[a]For genetic symbols, see Bachmann (1990)
[b]*hsdRK12* derivative of MC1000 (Casadaban and Cohen 1980)
[c]Thr^+, *lacI* – *Z*Δ(*Mlu*) derivative of C600 (Bachmann 1972)
[d]GC2553 was obtained from R. D'Ari (see Vinella and D'Ari 1994 for the history of this strain)

Table 2 List of oligonucleotide primers used in this work

Primer	Sequence (5′→3′)
rpoS1	GCTGCGTTATTTGCCGCAGCG
rpoS2	GTTATTCACCACTGTTAACGGCCG
rpoS4	GCTGGACATCCTGGCCGATG
rpoS5	GGGCATCGGACCTTTTATTGTGCAC
rpoS7	**GGCGCGCC**TTACTCRCGRAAYARNGCYTCRAT[a]
rpoS8	ACGGAGTTGAGGTTTTTGACG
rpoS9	ACGATTGAACGGGCGATTATG
rpoS10	GCTGTTTGGCGTTCAGCTCG
rpoS11	CGAGC**TCTAGA**GCTGCGTTATTTGCCGCAGCG
rpoS12	CATGA**GGATCC**GGGCATCGGACCTTTTATTGTGCAC
rpoS18	GTGAACGTGTTCGCCAGATTC
rpoS21	AGCGTTCCATCAGTTACGACAGC
rpoS22	CGGACAACCGTGGTCACTACAGC
rpoS23	GACTCAGGGTTCTGGATTGTGACC
rpoS24	GCTGATGCCAGTTCCAAAGC
rpoS25	GCCGTATCGGATACGTTGG
katE1	CATGA**GGATCC**ATTACTGGCTTCACTAAACGC
katE2	CGAGC**TCTAGA**GCGACATTGAACTCGTCTCC
pCB2	AGCGCAGCAGAGGCGGAT
Y8	GATGGATTATGTTTGCTCCG
Y12	TTTAAATGCCATTGCCGCTC

[a]Restriction enzyme recognition sites incorporated into primer tails are indicated in *bold face*

(Hansen et al. 1991), except that the bands were visualised with ECF substrate RPN5785 (Amersham Pharmacia Biotech) using a Storm 840 phospho-fluoroimager (Amersham Pharmacia Biotech) and quantified using Imagemaster TotalLab. The RpoS antibody was kindly donated by Regine Hengge-Aronis. When appropriate, antibiotics were added at 50 μg/ml (kanamycin), 20 μg/ml (tetracycline), and 100 μg/ml (ampicillin).

DNA manipulations and nucleotide sequence determination

Plasmid DNA was prepared by alkaline lysis (Birnboim and Doly 1979). Restriction enzymes were used as recommended by the supplier (Fermentas). PCR amplifications were performed in a Personal Cycler (Biometra) using Thermoprime Plus (Advanced Biotechnologies) or Expand High Fidelity (Böehringer Mannheim) polymerase in the buffers provided by the suppliers. PCR fragments for determination of the *rpoS* sequence from the different strains were generated with Thermoprime using 1 μl of overnight culture in LB or a colony from a LB plate as template and the primers rpoS1 and rpoS5, except for strains BJ4 and F18, where primer rpoS7 was used instead of rpoS5. PCR fragments spanning the insertion downstream of *rpoS* in strains BJ4 and F18 were generated using the primers rpoS18 and rpoS23 and the Expand Long Template PCR system from Roche Diagnostics. Sequencing of the *rpoS* gene was performed with the primers rpoS1, 2, 3, 4, 8 and 9 and the Thermo Sequenase cycle sequencing kit from Amersham, using the internal primer labelling protocol with [^{35}S]dATP. The sequences of the ends of the PCR products spanning the *rpoS*-o454 intergenic regions from F18 and BJ4 were determined using the primers rpoS18 and rpoS21-25 and the Big Dye Terminator cycle sequencing Ready Reaction kit (Applied Biosystems) and the reactions were analysed on the ABI Prism 310 Genetic Analyzer. All primer sequences used are listed in Table 2.

Construction of plasmids

The plasmids pTAC4550 and pTAC4589 for IPTG-controlled production of the RpoS(33Q) and RpoS(33E) variants, respectively, were constructed as follows: the entire *rpoS* gene was amplified from MC4100 and Cp14 using the primers rpoS11 and rpoS12 (Fig. 1A and Table 2) and High Fidelity polymerase; the PCR fragments were then digested with *Xba*I and *Bam*HI and ligated to plasmid pFH2102 (to be published elsewhere) digested with the same enzymes. To verify that the correct sequence was present, the insert in plasmids pTAC4550 and pTAC4589 was sequenced using the primers rpoS1, 2, 4, 8 and 9. Plasmid pTAC4598 carries the *katE* promoter inserted upstream of a *gfp* gene with codon usage optimised for *E. coli* (F. G. Hansen, to be published elsewhere). The *katE* promoter was amplified using the primers katE1 and katE2, producing a fragment extending from –153 to + 60 relative to the transcript start determined by Tanaka et al. (1997). The *katE* PCR fragment was digested with *Xba*I and ligated to plasmid pFHC2207 digested with *Sma*I and *Xba*I, giving rise to pTAC4598. The entire insert in pTAC4598 was sequenced using the primer pCB2, which anneals just upstream of the multiple cloning site of pFHC2207.

Southern analysis

Chromosomal DNA was prepared, restricted with different enzymes, and Southern analysis was carried out as described previously (Atlung and Hansen 1993). Hybridisation probes were prepared by labelling PCR fragments (purified on Qiaquick spin columns) with [^{35}S]dATP, using DNA polymerase I Klenow fragment and random priming (Megaprime kit; Amersham). The PCR fragment of *rpoS* was prepared using the primers rpoS1 and rpoS5 and the *appY* fragment was obtained using primers Y8 and Y12.

Results

Nucleotide sequence of the *rpoS* genes

The genotypes and sources of the strains for which we have determined the *rpoS* gene sequence are shown in Table 1. The K12 strains we selected include the MC4100 strain, which is normally used as an *rpoS* wild type; five strains originally characterised as low in AppA, and thus candidates for *rpoS* mutants; two characterised as having high AppA levels, and thus supposedly *rpoS*$^{+}$ (Touati et al. 1986); two original K12 isolates (UTH1038 and EMG2) obtained from two different sources, and finally three strains that we have frequently used in physiological experiments. We also

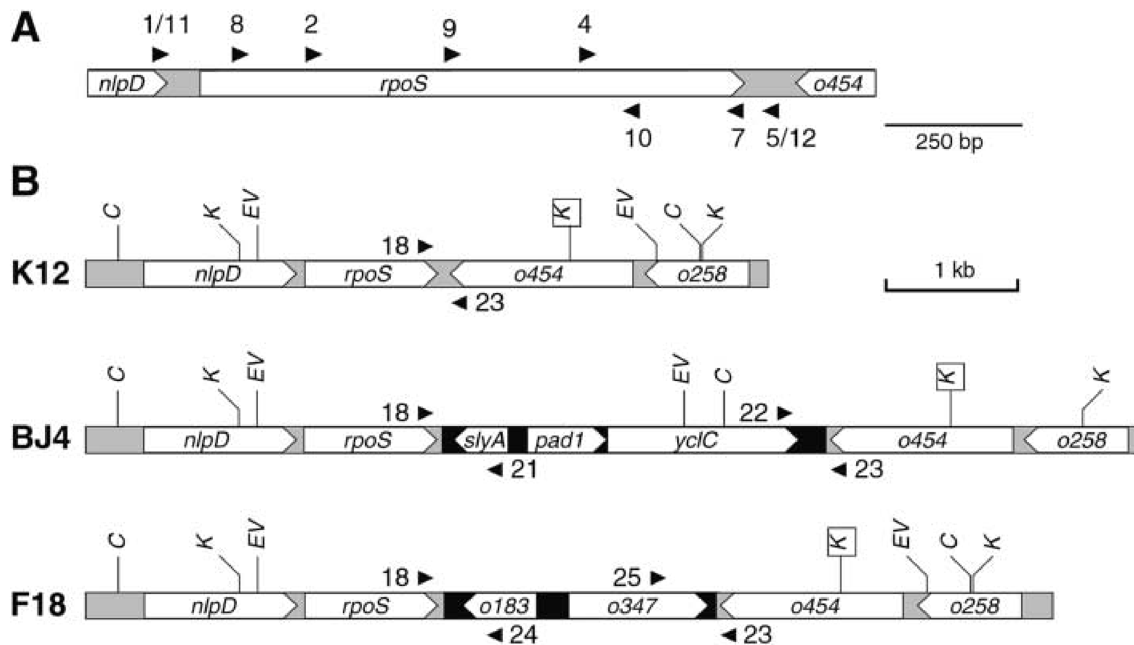

Fig. 1A, B Genetic and restriction enzyme maps of the *rpoS* region of various *E. coli* strains. The positions of the primers used are also indicated. The DNA sequence is shown as a *grey bar* with coding sequences shown as *white arrows* indicating the transcriptional orientation. Primers are denoted as *small black arrows*, those *above* the maps are forward primers while those *below* are reverse primers. The numbers refer to the primer names given in Table 2, where the sequences of the primers are given. **A** Positions of primers used for sequencing of the *rpoS* genes. **B** Maps of the *rpoS* region in the different *E. coli* strains. At the *top* is the map of the K12 strain based on the genomic sequence (GenBank Accession No. U00096). The restriction maps of the BJ4 and F18 strains are based on the results of Southern hybridisation analysis using *rpoS* as probe and assuming that the positions of restriction sites upstream of the *rpoS* gene are the same as in K12. The *black areas* in the maps indicate insertions relative to the K12 map. Coding sequences/genes are identified from the annotation of the genomic K12 sequence (GenBank Accession No. U00096), and for the BJ4 insert from the sequences of the EPEC/EHEC strains (GenBank Accession Nos. AF242208–242211) and for F18 from the sequence of the uropathogenic strain CFTO73 (GenBank Accession No. AF270497). Restriction enzyme symbols: C, *Cla*I; EV, *Eco*RV; K, *Kpn*I. the *boxed* K indicates a partially digested *Kpn*I site

chose six other non-pathogenic *E. coli* strains – four old laboratory strains (two B/r strains, a $15T^-$ and a C strain), and two more recent isolates with no history of spending time in stabs: F18, isolated from a healthy human (Wadolkowski et al. 1988), and BJ4, isolated from a healthy rat. Both these strains are efficient colonisers of the rodent intestine (Krogfelt et al. 1993).

The *rpoS* gene was amplified directly from the chromosome and its nucleotide sequence determined using the primers indicated in Fig. 1A. For all strains except F18 and BJ4, PCR products were obtained using the rpoS primers 1 and 5. We only obtained PCR products from F18 and BJ4 by using rpoS7, in place of rpoS5, as the downstream primer; rpoS7 overlaps the stop codon.

The results of the sequencing analyses are summarised in Table 3. We found that five of the K12 strains had precisely the same *rpoS* sequence as MG1655, used for determination of the *E. coli* genome sequence by the Blattner group (GenBank Accession No. U00096). Another six of the K12 strains, including the two original K12 isolates, deviated from this sequence only in having a TAG triplet at codon 33, while the last two K12 strains had two other substitutions at position 33. All six non-K12 strains had GAG at codon 33, like all the *rpoS* genes sequenced from *E. coli* samples obtained from food or from clinical isolates (Ferreira et al. 1999). All of the 18 strains that we sequenced had TTT as codon 25 and thus a phenylalanine at this position, not leucine as found in two of the previously published K12 sequences (Ivanova et al. 1992; Jishage and Ishihama 1997).

The *rpoS* gene of the $15T^-$ strain showed no other differences from K12, while the two B/r strains had a single neutral base substitution. The two recent intestinal isolates, BJ4 and F18, were the most divergent, having three and seven neutral substitutions relative to K12. The nucleotide sequence found in F18 was identical to that of strain O18:K1:H7 (GenBank Accession No. AF083844) except for the C1154A mutation that creates an amber codon in the latter strain, and to that of the uropathogenic strain CFTO73 (AF270497) except for two differences in neutral substitutions (C756T in CFTO73 and G1482A in F18). Four of the neutral substitutions common to these three strains are also found in the *rpoS* sequences of EPEC1 strains (AF242209 and AF242211). The neutral substitutions found in BJ4 were different from those found in any of the *E. coli rpoS* sequences that have been deposited in GenBank.

Genetic alterations downstream of *rpoS* in F18 and BJ4

As mentioned above we failed to generate a PCR product from strains F18 and BJ4 using the primer

Table 3 Sequence variations in the *rpoS* gene

Strain	Codon 33 variants		Other significant changes	Neutral substitutions
	Codon	Amino acid		
MC4100	CAG	Q		
KL99	CAG	Q		
SBS1028	CAG	Q		
MT102	CAG	Q		
KL16	CAG	Q		
LJ24	TAG	Amber		
FH1218	TAG	Amber		
EMG2	TAG	Amber		
CSH61	TAG	Amber		
UTH1038	TAG	Amber		
CSH57	TAG	Amber	C733A (= R81S)	
Hfr6	TCG	S		
SBS1029	TAT	Y	ΔC1239	
C	GAG	E	Δ(643–653)	
Cp14	GAG	E		C1434T
WM301	GAG	E		C1434T
JG151	GAG	E		
BJ4(BR21)	GAG	E		A921G, C1251T, G1443A
F18	GAG	E		T655C, T849G, T954C, 1065C, C1224T, G1311A, G1482A

[a]Base numbering is as in the GenBank sequence ECOSIGMA38 (Accession No. D13548) and differences relative to this sequence are given. For all strains, except F18 and BJ4, the nucleotide sequence was determined between positions 450 and 1555 (the coding sequence extends from 493 to 1483). For F18 and BJ4 the sequence was determined from position 450 to 450 nt beyond the termination codon (GenBank Accession Nos. AF410776 and AF410777, respectively)

rpoS5, which is located in the intergenic region between *rpoS* and the downstream K12 gene o454. We therefore performed a Southern analysis using a *rpoS* gene probe (data not shown) to construct a restriction map of the *rpoS* region from these strains. For the K12 (LJ24) and C strains the fragment sizes obtained using the restriction enzymes indicated in Fig. 1B were identical to those deduced from the K12 genomic sequence. The results for BJ4 indicated that this strain carries an insertion of approximately 3000 bp between *rpoS* and o454 (see Fig. 1B). The restriction fragment lengths obtained for F18 were compatible with an insertion of approximately 2000 bp in this strain. That both strains carry insertions, and not DNA substitutions, was confirmed by PCR analysis using the primers rpoS18, located near the end of the *rpoS* gene (see Fig. 1B), and rpoS23, located a little into the coding sequence of o454. The K12, B/r, $15T^-$ and C strains yielded PCR products of approximately 330 bp as expected from the K12 genome sequence, while BJ4 and F18 gave fragments of 3200 bp and 2400 bp, respectively, as expected if these strains carry insertions between *rpoS* and o454. The size of the insertions, as well as the results of the Southern analysis, suggests that BJ4 carries the same insertion as the EPEC1, EPEC2 and EHEC2 strains (Herbelin et al. 2000) and that F18 carries the same insertion as the uropathogenic strain CFTO73 and other strains in group B2 (Culham and Wood 2000). These hypotheses were supported by PCR analysis using primers internal to these insertion sequences in combination with the rpoS and o454 primers (see Fig. 1B for the locations of the primers used), and further confirmed by determination of approximately 500 bp of sequence at each end of the two PCR products obtained with the primers rpoS18 and rpoS23 primers. The sequences from F18 showed 99.5% identity to those of CTFO73, while the sequences from BJ4 showed between 98.1% and 94.8% identity to those of the EPEC/EHEC strains in the *slyA* region and between 99.6% and 96.5% identity in the *yclC* region.

Correlation of *rpoS* genotypes and phenotypes

All the strains used for sequencing of *rpoS* in this study were tested for three RpoS-related phenotypes (Fig. 2). Expression of the *katE* gene was judged from catalase activity in a plate test, and directly visualised by introducing plasmid pTAC4598 carrying the *katE* promoter fused to *gfp,* which encodes green fluorescent protein. The acid phosphatase activity, encoded by *appA*, was measured in overnight cultures, and glycogen accumulation was tested on plates.

The two strains carrying small deletions that cause frameshift mutations in *rpoS*, and three of the six K12 *rpoS*(33am) strains (EMG2, CSH61 and UTH1038), had phenotypes very similar to that of the *rpoS*::Tn*10* strain. The other three K12 *rpoS*(33am) strains displayed *katE* expression levels similar to those of the *rpoS*(33Q) strains but were deficient in glycogen accumulation. Two of these strains (LJ24 and FH1218) are known to carry amber suppressors and strain CSH57 must also carry one. The lack of glycogen accumulation could be due to inefficient amber suppression under the special growth conditions used for the glycogen test.

All the five *rpoS*(33Q) strains and the single *rpoS*(33S) strain were phenotypically RpoS^{+}, although there were some variations in the phenotypes amongst the strains (see Fig. 2).

The strains carrying the *rpoS*(33E) allele displayed a greater variability in the RpoS-related phenotypes than was observed for the K12 *rpoS*(33Q) strains. The B/r strain WM301 must carry a *katE* mutation since it had very low catalase activity but was proficient for expression of the *katEp-gfp* fusion, and had high levels of glycogen accumulation. The other B/r strain (Cp14) had normal catalase activity but low glycogen accumulation. Strain F18 displayed a RpoS mutant phenotype with respect to catalase and AppA levels.

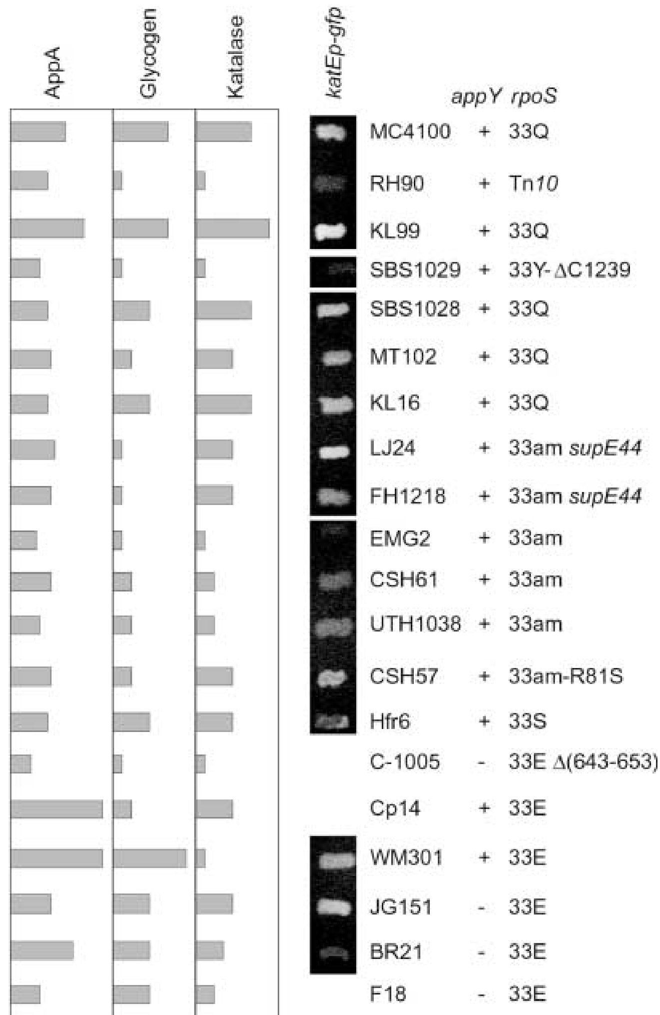

Fig. 2 Comparison of RpoS-regulated phenotypes and *katE* gene expression in the strains studied. AppA levels, glycogen accumulation and catalase activity were determined as described in Materials and methods. The values for strain MC4100 were set at 3 for the graphic representation. The indicated strains carrying plasmid pTAC4598 (*katEp-gfp* fusion) were patched onto a LB plate supplemented with ampicillin and thymidine, and grown overnight at 37°C. The plate was photographed under blue light (> 520 nm) with a video camera equipped with a green filter (approximately 470 nm). There are no *katE-gfp* data for the C strain because it was ampicillin resistant, and for Cp14 and F18 we were unable to obtain transformants. The presence of the *appY* gene was determined by PCR analysis using the primers Y8 and Y12, the absence of *appY* from strains BJ4, F18 and C was verified by Southern analysis using the *appY* fragment from LJ24 as probe and LJ24 as positive control

The variability in *appA* expression amongst the *rpoS*(33E) strains prompted us to test all the strains for the presence of the *appY* gene, since this gene codes for a transcriptional activator that is more important for *appA* expression than RpoS (Atlung et al. 1997). The *appY* gene is carried by a lambdoid prophage located between the *argU* and *envY* genes at 13 min on the *E. coli* K12 map (Bachmann 1990), and it has previously been shown that this phage, and thus the *appY* gene, was only present in 75% of 21 pathogenic *E. coli* strains tested (Nakata et al. 1993). Our genetic analysis showed that four of the *rpoS*(33E) strains lacked the *appY* gene, while all the K12 strains and the two B/r strains carried the gene (Fig. 2).

Comparison of the activity of the RpoS(33Q) and RpoS(33E) proteins

Due to the rather large variability in the RpoS-related phenotypes displayed by the different *rpoS*(33E) strains we could not draw any firm conclusions regarding the relative activities of the RpoS(33Q) and RpoS(33E) proteins. Attempts to introduce the *rpoS*(33E) allele from B/r into K12 by P1 transduction were unsuccessful, even when WM301 *hsdR*(K12) was used as donor. We therefore cloned the *rpoS*(33Q) and *rpoS*(33E) genes into the IPTG-controlled expression vector pFHC2102, and introduced the resulting plasmids pTAC4550 and pTAC4589 into TC4545, a *rpoS*::Tn*10* derivative carrying a *bolAp1-lacZ* fusion, since this is one of the most strongly RpoS-dependent promoters and can be used as a reporter gene to measure relative levels of RpoS activity (Lange and Hengge-Aronis 1991a; Bohannon 1991). Both RpoS variants showed the same activity at all IPTG concentrations (0.03–0.3 mM) when tested in the semi-quantitative plate tests for catalase function and glycogen accumulation. Induction of RpoS synthesis during exponential growth also stimulated the *bolAp1* promoter to the same extent (Fig. 3). The two induced strains contained the same amount of RpoS protein per unit cell mass as determined by immunoblot analysis (data not shown). Stimulation of *bolA* promoter activity increased in parallel with increasing IPTG concentration and was nearly maximal at 0.075 mM IPTG (data not shown). No further stimulation of *bolAp1* was observed above 0.1 mM IPTG, although expression from the vector promoter increases up to 0.5 mM (T. Atlung, unpublished data) and a small decrease in growth rate was observed at the higher IPTG concentrations. The degree of stimulation of the *bolA* promoter upon induction of saturating levels of RpoS during exponential growth was very similar to that observed when comparing *rpoS*$^{+}$ and *rpoS*::Tn*10* strains in stationary phase (Lange and Hengge-Aronis 1991a).

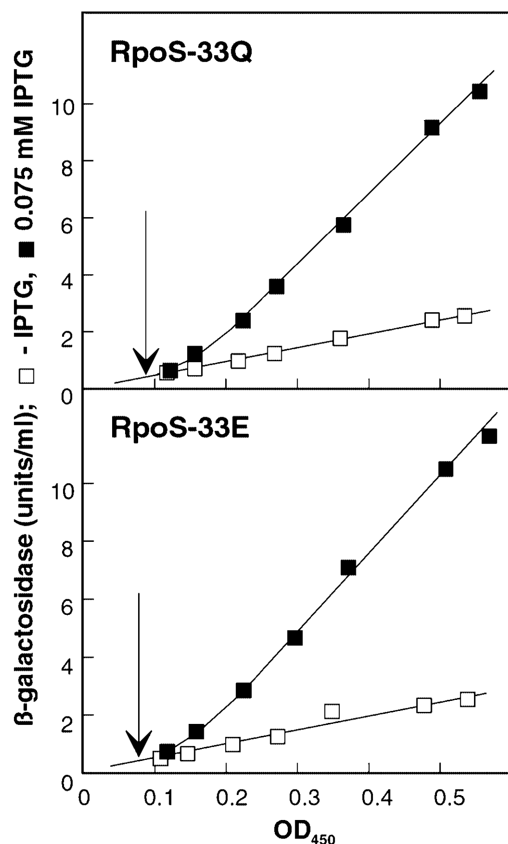

Fig. 3 Comparison of the transcriptional activation of *bolA*p1 by the RpoS-33Q and RpoS-33E variants. Cultures of strain TC4545 carrying plasmid pTAC4550 (RpoS-33Q) or plasmid pTAC4589 (RpoS-33E) were grown at 37°C in AB minimal glucose medium supplemented with casamino acids. At the time indicated by the *arrow* the cultures were divided and IPTG was added to one part (to a final concentration of 0.075 mM), and growth and *β*-galactosidase synthesis was followed. Both strains grew with the same generation time in the absence and presence of IPTG

Discussion

In this work we found that nearly 50% of the K12 strains examined carried an amber mutation in codon 33 of the *rpoS* gene, and that the majority of the remaining K12 strains had a CAG codon instead of the GAG that was present at this position in the six non-K12 *E. coli* strains tested in this work, and in the *rpoS* genes from food and clinical isolates of *E. coli* (Ferreira et al. 1999) as well as in all the other non-K12 *E. coli rpoS* sequences in GenBank. Thus all *E. coli* strains, except K12, have glutamate, not glutamine, at position 33, i.e. the amino acid that is found at this position in nearly all the other sequences deduced for enterobacterial RpoS proteins (GenBank Accession Nos. U66542, Y13230, AJ222716, U35777, X81641, U00119, U05011, U16152, Y13230). The complementation studies showed that the RpoS (33Q) variant had the same activity as the RpoS(33E) protein with respect to the few promoters tested so far. Preliminary experiments, however, suggest that the two proteins might differ in transcriptional activity at one or more operons involved in the development of acid resistance (T. Atlung, unpublished results).

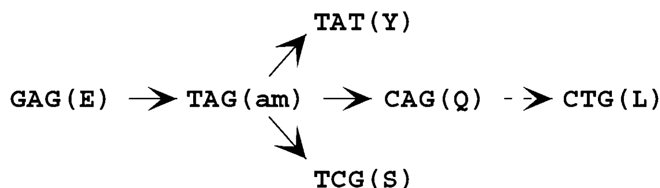

Fig. 4 Scheme depicting the probable evolution of codon 33 in *E. coli* K12 strains

The findings in this work, and in earlier studies of *rpoS* sequences in K12, strongly suggest the following evolutionary relationships (depicted in Fig. 4): the common ancestor of all *E. coli* strains had GAG at position 33, which mutated to TAG in the original K12 lineage, and then further mutations occurred during the evolution of the laboratory derivatives to help the cells to survive in stationary phase, primarily those that restored a functional RpoS protein either by pseudoreversion of the original amber or mediated by an amber suppressor. We did not find any strains with the C-terminal mutation described previously: an insertion after codon 326 leading to a frameshift (Ivanova et al. 1992) which was also found in a screen for mutants that take over stationary phase cultures (Zambrano et al. 1993).

Compared to K12 and the other *E. coli* laboratory strains, the two more recent isolates, F18 and BJ4, both have an insertion downstream of *rpoS*. Strain BJ4 has the same genetic structure in the *rpoS*-o454 intergenic region as the enteropathogenic strains of groups EPEC1 and 2 and EHEC2 (Herbelin et al. 2000), although the neutral substitutions in the BJ4 *rpoS* gene are completely different from those found in these strains. The sequence found in the enteropathogenic strains is also present in *Shigella dysenteria* and in *Salmonella enterica* (Herbelin et al. 2000). Thus, their presence in BJ4 supports the view that the genes *slyA*, *padl* and *yclC* were present in the common ancestor of the enteric bacteria. Strain F18 carries the insertion found in uropathogenic strains (Culham and Wood 2000). The F18 strain displayed a partially mutant phenotype for RpoS, although the *rpoS* gene encoded a fully functional protein. The mutant phenotype might be caused by the insertion downstream of *rpoS,* which is located just at the 3′ end of the proposed terminator and thus might affect the mRNA stability. Alternatively, the RpoS deficiency could be due to mutation in one of the many RpoS regulatory genes.

A recent study showed that an *rpoS* mutant of *E. coli* BJ4 has a growth advantage over the $rpoS^+$ strain in competitive mouse colonisation experiments (Krogfelt et al. 2000). Since the mutant was constructed by transduction of *rpoS*::Tn*10* from K12, our finding that BJ4 carries an insertion downstream of *rpoS*, however, raises the question whether the effect on colonisation was in fact due to inactivation of *rpoS* or was caused by the deletion of the *slyA*, *padl* and *yclC* genes.

Acknowledgements We thank Kirsten Olesen for her expert technical assistance, Regine Hengge-Aronis for donation of RpoS antibody, and Regine Hengge-Aronis and Karen Krogfelt for comments on the manuscript.

References

Atlung T, Hansen FG (1993) Three distinct chromosome replication states are induced by increasing concentrations of DnaA protein in *Escherichia coli*. J Bacteriol 175:6537–6545

Atlung T, Nielsen A, Hansen FG (1989) Isolation, characterization, and nucleotide sequence of *appY*, a regulatory gene for growth-phase-dependent gene expression in *Escherichia coli*. J Bacteriol 171:1683–1691

Atlung T, Knudsen K, Heerfordt L, Brøndsted L (1997) Effects of σ^S and the transcriptional activator AppY on induction of the *Escherichia coli hya* and *cbdAB-appA* operons in response to carbon and phosphate starvation. J Bacteriol 179:2141–2146

Bachmann BJ (1972) Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol Rev 36:525–557

Bachmann BJ (1990) Linkage map of *Escherichia coli* K-12, edition 8. Microbiol Rev 54:130–197

Becker G, Klauck E, Hengge-Aronis R (1999) Regulation of RpoS proteolysis in *Escherichia coli*: the response regulator RssB is a recognition factor that interacts with the turnover element in RpoS. Proc Natl Acad Sci USA 96:6439–6444

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Blattner FR, et al (1997) The complete genome sequence of *Escherichia coli* K-12. Science 277:1453–1474

Bohannon DE (1991) Stationary-phase-inducible "Gearbox" promoters: differential effects of *katF* mutations and role of σ^{70}. J Bacteriol 173:4482–4492

Brøndsted L, Atlung T (1994) Anaerobic regulation of the hydrogenase 1 (*hya*) operon of *Escherichia coli*. J Bacteriol 176:5423–5428

Calendar R, Lindquist B, Sironi G, Clark AJ (1970) Characterization of REP^- mutants and their interaction with P2 phage. Virology 40:72–83

Casadaban MJ, Cohen S (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J Mol Biol 138:179–207

Clark DJ, Maaløe O (1967) DNA replication and the division cycle in *Escherichia coli*. J Mol Biol 23:99–112

Culham DE, Wood JM (2000) An *Escherichia coli* reference collection group B2- and uropathogen-associated polymorphism in the *rpoS-mutS* intergenic region of the *E. coli* chromosome. J Bacteriol 182:6272–6276

Donachie WD (1968) Relationship between cell size and time of initiation of DNA replication. Nature 219:1077–1079

Ferreira A, Rendano L, Wiedmann M, Boor KJ (1999) Characterization of *rpoS* alleles in *Escherichia coli* O157:H7 and in other *E. coli* serotypes. J Appl Microbiol 86:295–301

Goldsmith EP, Cater MS, Butler MA, Greene A (1970) Genetic analysis of the histidine operon in *Escherichia coli* K12. Genetics 66:219–229

Gourmelon M, Touati D, Pommepuy M, Cormier M (1997) Survival of *Escherichia coli* exposed to visible light in seawater: analysis of *rpoS*-dependent effects. Can J Microbiol 43:1036–1043

Hansen FG, Atlung T, Braun RE, Wright A, Hughes P, Kohiyama M (1991) Initiator (DnaA) protein concentration as a function of growth rate in *Escherichia coli* and *Salmonella typhimurium* . J Bacteriol 173:5194–5199

Hengge-Aronis R (1996) Regulation of gene expression during entry into stationary phase. In: Neidhardt FC, Curtiss R III, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE (eds) *Escherichia coli* and *Salmonella:* cellular and molecular biology (2nd edn). American Society for Microbiology, Washington D.C., pp 1497–1512

Hengge-Aronis R, Fischer D (1992) Identification and molecular analysis of *glgS*, a novel growth-phase-regulated and *rpoS*-dependent gene involved in glycogen synthesis in *Escherichia coli*. Mol Microbiol 6:1877–1886

Hengge-Aronis R, Klein W, Lange R, Rimmele M, Boos W (1991) Trehalose synthesis genes are controlled by the putative sigma factor encoded by *rpoS* and are involved in stationary-phase thermotolerance in *Escherichia coli*. J Bacteriol 173:7918–7924

Hengge-Aronis R, Lange R, Henneberg N, Fischer D (1993) Osmotic regulation of *rpoS*-dependent genes in *Escherichia coli*. J Bacteriol 175:259–265

Herbelin CJ, Chirillo SC, Melnick KA, Whittam TS (2000) Gene conservation and loss in the *mutS-rpoS* genomic region of pathogenic *Escherichia coli*. J Bacteriol 182:5381–5390

Ivanova A, Renshaw M, Guntaka RV, Eisenstark A (1992) DNA base sequence variability in *katF* (putative sigma factor) gene of *Escherichia coli*. Nucleic Acids Res 20:5479–5480

Jishage M, Ishihama A (1997) Variation in RNA polymerase sigma subunit composition witin different stocks of *Escherichia coli* W3110. J Bacteriol 179:959–963

Krogfelt KA, Kongsbak Poulsen L, Molin S (1993) Identification of coccoid *Escherichia coli* BJ4 cells in the large intestine of streptomycin-treated mice. Infect Immun 61:5029–5034

Krogfelt KA, Hjulgaard M, Sørensen K, Cohen PS, Givskov M (2000) *rpoS* gene function is a disadvantage for *Escherichia coli* BJ4 during competitive colonization of the mouse large intestine. Infect Immun 68:2518–2524

Lange R, Hengge-Aronis R (1991a) Growth phase-regulated expression of *bolA* and morphology of stationary-phase *Escherichia coli* cells are controlled by the novel sigma factor σ^S. J Bacteriol 173:4474–4481

Lange R, Hengge-Aronis R (1991b) Identification of a central regulator of stationary phase gene expression in *Escherichia coli*. Mol Microbiol 5:49–59

Løbner-Olesen A, Skarstad K, Hansen FG, von Meyenburg K, Boye E (1989) The DnaA protein determines the initiation mass of *Escherichia coli* K-12. Cell 57:881–889

Miller JH (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Muffler A, Traulsen DD, Lange R, Hengge-Aronis R (1996a) Posttranscriptional osmotic regulation of the σ^S subunit of RNA polymerase in *Escherichia coli*. J Bacteriol 178:1607–1613

Muffler A, Fischer D, Altuvia S, Storz G, Hengge-Aronis R (1996b) The response regulator RssB controls stability of the σ^S subunit of RNA polymerase in *Escherichia coli*. EMBO J 15:1333–1339

Mulvey MR, Loewen PC (1989) Nucleotide sequence of *katF* of *Escherichia coli* suggests *KatF* protein is a novel transcription factor. Nucleic Acids Res 17:9979–9991

Nakata N, Tobe T, Fukuda I, Suzuki T, Komatsu K, Yoshikawa M, Sasakawa C (1993) The absence of a surface protease, OmpT, determines the intercellular spreading ability of *Shigella*: the relationship between the *ompT* and *kcpA* loci. Mol Microbiol 9:459–468

Neidhardt FC, Bloch PL, Smith DF (1974) Culture medium for enterobacteria. J Bacteriol 119:736–747

Pratt LA, Silhavy TJ (1996) The response regulator SprE controls the stability of RpoS. Proc Natl Acad Sci USA 93:2488–2492

Rasmussen LJ, Møller PL, Atlung T (1991) Carbon metabolism regulates expression of the *pfl* (pyruvate formate-lyase) gene in *Escherichia coli*. J Bacteriol 173:6390–6397

Rod ML, Alam KY, Cunningham PR, Clark DP (1988) Accumulation of trehalose by *Escherichia coli* K-12 at high osmotic pressure depends on the presence of amber suppressors. J Bacteriol 170:3601–3610

Silhavy TJ, Berman ML, Enquist LW (1984) Experiments with gene fusions. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Small P, Blankenhorn D, Welty D, Zinser E, Slonczewski JL (1994) Acid and base resistance in *Escherichia coli* and *Shigella flexneri*: Role of *rpoS* and growth pH. J Bacteriol 176:1729–1737

Tanaka K, Takayanagi Y, Fujita N, Ishihama A, Takahashi H (1993) Heterogeneity of the principal σ factor in *Escherchia coli*: the *rpoS* gene product, σ^{38}, is a second principal σ factor of RNA polymerase in stationary-phase *Escherichia coli*. Proc Natl Acad Sci USA 90:3511–3515

Tanaka K, Handel K, Loewen PC, Takahashi H (1997) Identification and analysis of the *rpoS*-dependent promoter of *katE*, encoding catalase HPII in *Escherichia coli*. Biochim Biophys Acta 1352:161–166

Touati E, Dassa E, Boquet PL (1986) Pleiotropic mutations in *appR* reduce pH 2.5 acid phosphatase expression and restore succinate utilisation in CRP-deficient strains of *Escherichia coli*. Mol Gen Genet 202:257–264

Touati E, Dassa E, Dassa J, Boquet PL, Touati D (1991) Are *appR* and *katF* the same *Escherichia coli* gene encoding a new sigma transcription initiation factor. Res Microbiol 142:29–36

Vinella D, D'Ari R (1994) Thermoinducible filamentation in *Escherichia coli* due to an altered RNA polymerase β subunit is suppressed by high levels of ppGpp. J Bacteriol 176:966–972

Visick JE, Clarke S (1997) RpoS- and OxyR-independent induction of HPI catalase at stationary phase in *Escherichia coli* and identification of *rpoS* mutations in common laboratory strains. J Bacteriol 179:4158–4163

Wadolkowski EA, Laux DC, Cohen PS (1988) Colonization of the streptomycin-treated mouse large intestine by a human fecal *Escherichia coli* strain. Infect Immun 56:1030–1035

Zahn G, Messer W (1979) Control of the initiation of DNA replication in *E. coli*. II. Function of the *dnaA* product. Mol Gen Genet 168:197–209

Zambrano MM, Siegele DA, Almiron M, Tormo A, Kolter R (1993) Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. Science 259:1757–1760

Zeuthen J, Pato ML (1971) Replication of the F'*lac* sex factor in the cell cycle of *Escherichia coli*. Mol Gen Genet 111:242–255