Molecular Microbiology (2012) 85(1), 142–151 ■
doi:10.1111/j.1365-2958.2012.08099.x
First published online 31 May 2012

Fitness-compensatory mutations in rifampicin-resistant RNA polymerase

Gerrit Brandis,[1] Marie Wrande,[2] Lars Liljas[2] and Diarmaid Hughes[1]*
[1]*Department of Medical Biochemistry and Microbiology, Uppsala University, Biomedical Center, Box 582, SE-75123 Uppsala, Sweden.*
[2]*Department of Cell and Molecular Biology, Uppsala University, Biomedical Center, Box 596, SE-75124 Uppsala, Sweden.*

Summary

Mutations in *rpoB* (RNA polymerase β-subunit) can cause high-level resistance to rifampicin, an important first-line drug against tuberculosis. Most rifampicin-resistant (Rif^R) mutants selected *in vitro* have reduced fitness, and resistant clinical isolates of *M. tuberculosis* frequently carry multiple mutations in RNA polymerase genes. This supports a role for compensatory evolution in global epidemics of drug-resistant tuberculosis but the significance of secondary mutations outside *rpoB* has not been demonstrated or quantified. Using *Salmonella* as a model organism, and a previously characterized Rif^R mutation (*rpoB* R529C) as a starting point, independent lineages were evolved with selection for improved growth in the presence and absence of rifampicin. Compensatory mutations were identified in every lineage and were distributed between *rpoA*, *rpoB* and *rpoC*. Resistance was maintained in all strains showing that increased fitness by compensatory mutation was more likely than reversion. Genetic reconstructions demonstrated that the secondary mutations were responsible for increasing growth rate. Many of the compensatory mutations in *rpoA* and *rpoC* individually caused small but significant reductions in susceptibility to rifampicin, and some compensatory mutations in *rpoB* individually caused high-level resistance. These findings show that mutations in different components of RNA polymerase are responsible for fitness compensation of a Rif^R mutant.

Accepted 11 May, 2012. *For correspondence. E-mail diarmaid.hughes@imbim.uu.se; Tel. (+46) 18 4714507; Fax (+46) 18 471 4673.

Introduction

Rifampicin (RIF) is an important first-line antibiotic for the treatment of tuberculosis (WHO, 2010; Conde and Lapa, 2011). RIF is a bactericidal antibiotic that inhibits the bacterial RNA polymerase (Jindani *et al.*, 2003). RNA polymerase is a large (~400 kDa) enzyme consisting of five subunits (Zhang *et al.*, 1999; Murakami and Darst, 2003; Werner, 2008). The $\alpha_2\beta\beta'\omega$ pentameric core forms a crab-claw-like structure where the α subunits are important for assembly and the ββ′ heterodimer forms the catalytic centre. The role of the ω subunit is not clear. Single-stranded DNA enters the enzyme through the primary channel formed by the ββ′ heterodimer. Nucleotides enter via the secondary channel and the polymerized RNA leaves through the exit channel (Zhang *et al.*, 1999). All channels have a highly positive surface and stabilize the nucleic acids by van der Waals and polar interactions between amino acid residues of the polymerase and the phosphate backbone of the DNA or RNA (Vassylyev *et al.*, 2007). Unlike eukaryotic cells, which encode three distinct RNA polymerases, prokaryotes use only one RNA polymerase to produce rRNAs, mRNAs and tRNAs (Archambault and Friesen, 1993). This makes the bacterial RNA polymerase a fitness-determining enzyme since bacterial growth rate is directly related to the rate of production of rRNAs (Condon *et al.*, 1995).

Structural analysis indicates that RIF binds to the β subunit of bacterial RNA polymerase within the DNA/RNA channel and directly blocks the path of the elongating RNA when the transcript becomes 2–3 nt in length (Campbell *et al.*, 2001). High-level resistance to RIF can be caused by single mutations in *rpoB*, the gene coding for the β subunit of RNA polymerase (Telenti *et al.*, 1993). Most of these mutations change amino acids in or surrounding the RIF-binding pocket and probably cause resistance by interfering with RIF binding to RNA polymerase (Campbell *et al.*, 2001). All of the high-level Rif^R mutations identified in *M. tuberculosis* or *E. coli* are located in one of four regions of the β subunit protein (Campbell *et al.*, 2001; Garibyan *et al.*, 2003). Using *E. coli* numbering for the β subunit amino acids, these regions are the N-terminal cluster (amino acids 146–148), cluster I (amino acids 507–534), cluster II (amino acids 563–574) and cluster III (amino acid 687) (Campbell *et al.*, 2001; Garibyan *et al.*, 2003). In

M. tuberculosis at least 95% of Rif^R mutations occur in cluster I (Ramaswamy and Musser, 1998; Heep *et al.*, 2001). However, DNA sequence analysis of Rif^R *M. tuberculosis* has shown that many clinical isolates carry multiple mutations in the same regions of *rpoB* (Kapur *et al.*, 1994; Musser, 1995; Sekiguchi *et al.*, 2007). Because individual Rif^R mutations have been experimentally shown to cause high-level resistance while reducing growth rate (Bjorkman *et al.*, 1998; Billington *et al.*, 1999; Reynolds, 2000; Sander *et al.*, 2002; Mariam *et al.*, 2004; Gagneux *et al.*, 2006), the finding of multiple mutations in *rpoB* in clinical isolates suggests that the secondary mutations may have been selected for growth rate compensation. This hypothesis is in accordance with an experimental study in which four different *E. coli* Rif^R mutants were evolved with selection for growth compensation, and compensatory mutations were identified as second-site mutations in *rpoB* (Reynolds, 2000). The compensatory mutations improved growth rate without significantly altering the level of resistance. However, in almost half of the compensated mutants examined in that study (9/20), no secondary mutation was identified in the first 2226 nucleotides of *rpoB* that was sequenced (Reynolds, 2000). One possible but untested explanation for the remaining unidentified compensatory mutations is that they map in the final, approximately 1800 nucleotides, of *rpoB*.

Because RNA polymerase is a complex multicomponent enzyme it is not obvious that the compensatory mutations need be restricted to *rpoB*. Indeed, very recently, whole-genome sequencing of paired clinical RIF-resistant and corresponding RIF-susceptible *M. tuberculosis* isolates, and *in vitro* evolution of susceptible strains in the absence of RIF, identified additional mutations in 38 genes and 10 intergenic regions (Comas *et al.*, 2012). The authors suggested that mutations in the RNA polymerase subunit genes *rpoA* and *rpoC* (α and β' subunits respectively) that occurred exclusively in RIF-resistant genomes with *rpoB* mutations were likely to be compensatory (Comas *et al.*, 2012). This hypothesis was strengthened by analysis of several hundred multidrug-resistant (MDR) and RIF-susceptible *M. tuberculosis* clinical isolates, showing that 27.1% of MDR strains carried a non-synonymous amino acid substitution mutation in *rpoA* or *rpoC* (Comas *et al.*, 2012). In an other study, whole-genome sequences of 34 RIF-resistant clinical *M. tuberculosis* isolates from Russia showed that 14 out of 24 strains that carry the *rpoB* mutation S450L carried an additional mutation in *rpoA* or *rpoC* that was not found in drug-sensitive isolates, and it was suggested that these were possibly fitness compensatory mutations (Casali *et al.*, 2012). However, the hypothesis in these papers that individual mutations in *rpoA* or *rpoC* compensate for fitness costs associated with primary resistance mutations in *rpoB* is based on a statistical association and was not directly tested, and the effect of the putative compensatory mutations, on growth and susceptibility, remains to be determined and quantified.

In this study, using *Salmonella enterica* as a model organism, and a previously well-characterized Rif^R mutation, *rpoB* R529C (Jin and Turnbough, 1994), as a starting point, we evolved 20 independent lineages under selection for improved growth competitiveness in the presence or in the absence of RIF. Our aims were (i) to identify the locations of fitness-compensatory mutations; (ii) test whether the putative compensatory mutations were sufficient to explain the compensated phenotype; and (iii) to quantify the effects of the compensatory mutations in terms of improved growth rate and changes in RIF susceptibility. We report the identification of compensatory mutations in *rpoA*, *rpoB* or *rpoC* in each of the evolved strains and their interpretation in terms of the structural information available on bacterial RNA polymerase.

Results

Phenotypes of rpoB *R529C*

We have previously isolated a set of Rif^R mutants carrying individual mutations in *rpoB* in *S. enterica* ATCC 14028s (Wrande *et al.*, 2008). The exponential growth rates of these strains were measured in Luria–Bertani (LB) medium to identify a suitable candidate for compensatory evolution (Table S1). We chose *rpoB* R529C, a mutation that caused slow growth in rich medium (LB generation time: wild-type, 19.55 min; mutant, 54.28 min) and a high level of resistance to RIF [minimum inhibitory concentration (MIC): wild-type, 12 μg ml^{-1}; mutant, 3000 μg ml^{-1}]. The R529C mutation has previously been shown to cause an increased K_M for UTP with *E. coli* RNA polymerase, resulting in an increase in abortive transcription initiation from the *E. coli pyrB1* promoter (Jin and Turnbough, 1994), thus explaining its slow growth rate. Using real-time RT-PCR (*Experimental procedures*) we showed that the R529C mutation in *S. enterica* also reduced transcription efficiency relative to the wild-type (Fig. S1).

*Selection of compensatory mutations in TH7884 (*rpoB *R529C)*

Twenty independent lineages of TH7884 (*rpoB* R529C, Rif^R) were evolved by serial passage (*Experimental procedures*): 10 in LB medium and 10 in LB + RIF. After 60 generations of growth each culture was diluted, plated on LB agar, and relative colony growth rate assessed visually after overnight incubation at 37°C. The slow growth of the original mutant strain on LB agar made it easy to distinguish colonies with faster growth rates. Two fast-growing colonies were picked from each of the 20 lineages. The

Table 1. Compensatory mutations selected in *rpoA*, *rpoB* and *rpoC*.

Mutations in *rpoA, B* or *C*					Selection[d]	
Original	Compensatory	MIC[a]	Dt (min) ± SD[b]	*N*[c]	LB	LB + RIF
Wild-type	Wild-type	12	19.55 ± 0.01	–	–	–
rpoB R529C	Wild-type	3000	54.28 ± 1.08	–	–	–
rpoB R529C	*rpoA* R191C	3000	26.74 ± 0.07	2	–	+ (2)
rpoB R529C	*rpoA* R191S	3000	29.73 ± 0.73	1	–	+
rpoB R529C	*rpoA* Q194P	3000	30.50 ± 0.32	1	+	–
rpoB R529C	*rpoB* D516G	>3000	22.24 ± 0.28	6	+ (3)	+ (3)
rpoB R529C	*rpoB* P560L	3000	26.41 ± 0.34	1	+	–
rpoB R529C	*rpoB* P564S	1500	34.18 ± 0.21	1	+	–
rpoB R529C	*rpoB* E565A	1500	23.96 ± 0.30	1	–	+
rpoB R529C	*rpoB* R637C	3000	31.96 ± 0.42	1	–	+
rpoB R529C	*rpoB* H673Y	3000	29.22 ± 0.36	1	+	–
rpoB R529C	*rpoC* P64L	3000	27.86 ± 0.68	1	+	–
rpoB R529C	*rpoC* L770P	3000	26.79 ± 0.09	1	–	+
rpoB R529C	*rpoC* R1075P	3000	26.17 ± 1.02	1	–	+
rpoB R529C	*rpoC* R1075H	3000	28.30 ± 0.30	1	–	+
rpoB R529C	*rpoC* G1136A	3000	27.96 ± 1.04	1	+	–
rpoB R529C	*rpoC* R1140H	3000	29.83 ± 0.93	1	+	–
rpoB R529C	*rpoC* V1198E	3000	26.13 ± 0.83	1	+	–

a. MIC RIF in μg ml^{-1}.
b. Dt is doubling time (generation time) in LB at 37°C, ± standard deviation (SD) based on three independent measurements for each strain.
c. *N* is the number of independent isolates (from independent lineages) of each mutation.
d. Growth medium in which growth-compensatory mutants were selected. LB or LB with rifampicin at 100 μg ml^{-1}. Numbers in brackets are number of independent isolates of each mutation.

rpoB gene was sequenced in each of these 40 strains. In every strain the original R529C mutation was still present, but 6 different secondary mutations were also identified in *rpoB*, distributed across 11 of the 20 lineages (Table 1). From 9 of the lineages identical *rpoB* mutations were identified in both isolates, while in the remaining 2 lineages a secondary *rpoB* mutation was found in only one of the two isolates sequenced. A P22 phage lysate was also made on each of the 40 strains and used to determine whether the compensatory mutations mapped in the *rpoBC* region (*Experimental procedures*). Positive transduction results (fast growth associated with the *rpoBC* region) were obtained for 33/40 strains, including each of the 20 strains in which a secondary mutation had been identified in *rpoB* by DNA sequencing. We concluded that 13 strains must have a growth compensatory mutation outside of, but closely linked to *rpoB*, while 7 strains must have a compensatory mutation elsewhere in the genome. The gene *rpoA* was regarded as a good candidate for the location of compensatory mutations outside of *rpoBC*. The *rpoA* and *rpoC* genes were PCR amplified from each of the 40 strains and sequenced. The results (Table 1) confirmed the mutation distribution predicted from the transduction data and showed that 13 strains carried a secondary mutation in *rpoC* while 7 strains carried a secondary mutation in *rpoA*. Thus each of the 40 faster-growing strains isolated after 60 generations of growth in LB or LB + RIF had retained the original *rpoB* R529C mutation and acquired an additional single mutation in *rpoA*, *rpoB* or *rpoC*. In all, 22 independent secondary mutations were identified in these 20 lineages, with 3 different mutations in *rpoA*, 6 different mutations in *rpoB*, and 7 different mutations in *rpoC* (Table 1).

The secondary mutations are necessary and sufficient for growth compensation

We tested whether the secondary mutations found in *rpoA*, *rpoB* and *rpoC* after 60 generations of TH7884 growth were indeed responsible for the observed growth compensation. Isogenic strains were constructed in which the identified mutations were transferred by P22 phage-mediated transduction into the wild-type TH6878 (*rpoB* double mutations and *rpoBC* double mutations) or into the slow-growing mutant TH8146 (*rpoA* mutations). In each case transductants were found to inherit the predicted faster-growth phenotype. As expected, transductants that did not inherit the *rpoA* or *rpoC* mutations were also obtained from each cross and they grew as slow as the parental strain TH7884. The genotypes of selected transductants of each phenotype and from each cross were tested and confirmed by DNA sequencing of *rpoA*, *rpoB* and *rpoC*. The growth rates of the reconstructed mutant strains were also measured and found to be indistinguishable from the values of the original selected mutants (Table S2). We concluded that the secondary mutations selected in *rpoA*, *rpoB*, and *rpoC* were necessary and sufficient to generate the growth-compensated phenotypes.

Rifampicin MICs of the compensated mutants

The RIF MIC was measured for each of the selected and reconstructed compensated mutants (Table 1, Table S2). For most of the compensated mutants MIC was unchanged compared with the parental strain (RIF MIC 3000 μg ml^{-1}). The only deviations observed were a reduction in MIC of 1 step for two strains and an increased MIC for one strain (growth at 3000 μg ml^{-1}, the highest concentration tested). Interestingly, the two *rpoB* mutations found in the double mutant with the increased RIF MIC (*rpoB* R529C and D516G) have each previously been shown to individually cause a Rif^R phenotype (Jin and Gross, 1988; Wrande *et al*., 2008). We concluded that in general the growth-compensatory mutations, whether selected in the presence or absence of RIF in the growth medium, and regardless of whether the compensating mutation affected *rpoA*, *rpoB* or *rpoC*, did not significantly reduce RIF MIC of the primary Rif^R mutant. We note that also in other respects the selection of compensatory mutants in the absence and in the presence of RIF did not have any significant influence on the spectrum of compensatory mutations obtained (with the caveat that the sample size is small). Thus, the same number (11) of independent compensatory mutations was obtained from each set of 20 colonies examined after the different selections; the distribution of compensatory mutations between the different RNA polymerase subunits did not differ significantly between the two selections (chi-squared test, *P*-value 0.39); the growth rates of the independent compensated mutants did not differ between the two data sets (Mann–Whitney test, two-tailed *P*-value 0.57).

Rifampicin MICs of compensatory mutations in rpoA *and* rpoC

We constructed strains carrying the *rpoA* and *rpoC* compensatory mutations alone and measured their RIF MICs and growth rates (Table 2). Interestingly, several of the mutations in *rpoA* and *rpoC* caused small but significant increases in the RIF MIC. One of the mutations, *rpoA* R191S, increased RIF MIC from 12 to 32 μg ml^{-1} with no significant effect on growth rate (Table 2). Another of the mutations, *rpoC* L770P, increased RIF MIC from 12 to > 32 μg ml^{-1} but this increase was associated with a significant decrease in growth rate. To our knowledge this is the first time that mutations in *rpoA* and *rpoC* have been shown to cause reduced susceptibility to RIF.

Positions of mutated residues in terms of RNA polymerase structure

The position and role of the fitness-compensatory mutants can be deduced from crystal structures of RNA polymerase (Fig. 1) from *T. thermophilus* and *T. aquaticus* (Campbell *et al*., 2005; Vassylyev *et al*., 2007). The amino acid sequence of the *Salmonella* and *Thermus* proteins differ in some regions but are mostly similar. Table 3 summarizes the information on the position and role of each of the mutated residues. The mutants are mostly found in the central part of the enzyme, where the sequence similarity is high. One of the mutated residues (*rpoC* G1136A) differs in sequence between these bacteria, and another

Table 2. Phenotypes of *rpoA* and *rpoC* compensatory mutations.

Mutation	RIF MIC[a]	Dt (min) ± SD[b]
Wild-type	12	19.55 ± 0.01
rpoA R191C	16	20.72 ± 1.65
rpoA R191S	**32**	19.96 ± 0.22
rpoA Q194P	**24**	20.00 ± 0.61
rpoC P64L	**24**	21.39 ± 0.31
rpoC L770P	**>32**	27.32 ± 0.60
rpoC R1075P	**24**	22.78 ± 1.53
rpoC R1075H	**24**	21.92 ± 0.71
rpoC G1136A	12	21.23 ± 0.40
rpoC R1140H	**32**	24.67 ± 0.52
rpoC V1198E	12	19.89 ± 0.67

a. MIC RIF in μg ml^{-1} based on at least three independent measurements for each strain. Values in **bold** are statistically significantly higher than the MIC of the wild-type (≥ 2 MIC steps increase relative to the wild-type MIC).
b. Dt is doubling time (generation time) in LB with shaking aeration at 37°C, ± standard deviation (S.D.), based on three independent measurements for each strain.

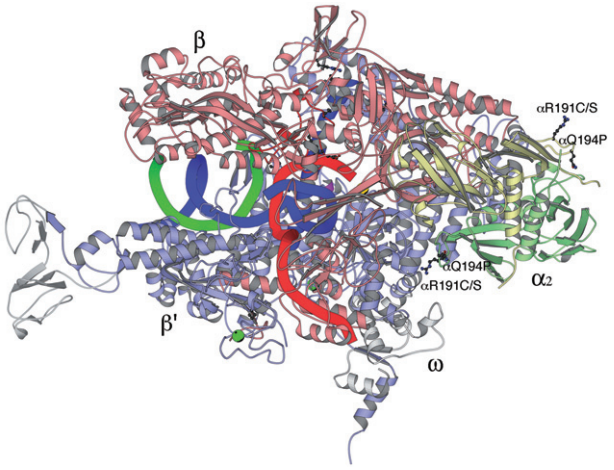

Fig. 1. RNA polymerase complete complex showing mutations in α subunits.
The elongation complex of *T. thermophilus*, PDB code 2O5I. The β subunit is in pale red, the β′ subunit is in pale blue (one segment to the left, not present in the *Salmonella* enzyme, in grey), the α subunits in pale yellow and green and the ω subunit in grey. The complex contains segments of nucleic acid with the template strand of DNA in blue, the non-template DNA strand in green, and RNA in red. Metal ions are marked by gold (Mg^{2+}) and green (Zn^{2+}) spheres. The bridge helix passing between the upper and lower part of the molecule is in dark blue. The mutants found in the α subunit are labelled.

Table 3. Location and role of residues mutated.

Mutant[a]	Tt[b]	Location	Role of side-chain[c]
rpoB R529C	409	Hybrid-binding channel	Binds to phosphate in RNA (RNA-DNA hybrid), binds to rifampicin in complex
rpoA R191C/S	185	α subunit	Binds to β′ E692 in one subunit, exposed with no contacts in other
rpoA Q194P	188	α subunit	Half exposed – no contacts in any of the subunits
rpoB D516G	396	Hybrid-binding channel	Side-chain is 4.8 Å from Arg405, 3.1 Å from rifampicin molecule in complex
rpoB P560L	440	Fork loop 2. Close to upper part of bridge helix	Close to β′ F1071, βF540, βV430
rpoB P564S	444	Fork loop 2. Hybrid-binding channel	Close to RNA in DNA/RNA hybrid, close to Arg409 (4 Å), binds to the same phosphate. Close to rifampicin binding site
rpoB E565A	445	Fork loop 2. Hybrid-binding channel	Points away from RNA towards βN556, 11 Å from active site Mg^{2+} ion
rpoB R637C	516	Close to upper part of bridge helix	Stacked between β′ L1068 (beginning of bridge helix) and βV514 in elongating complex. Pointing away from bridge helix in rifampicin complex
rpoB H673Y	552	Close to upper part of bridge helix	Hydrogen bond to main chain of β′ 1061. No steric problem for mutation
rpoC P64L	52	Zn finger	Proline close to two carbonyl oxygens. Some rearrangement needed in mutant
rpoC L770P	1068	In upper part of bridge helix	See β516 above
rpoC R1075P/H	-	Long insert in trigger loop of *S. emterica*	
rpoC G1136A	1254	In trigger loop	Q in *Thermus*. Disordered in elongating complex
rpoC R1140H	1258	In trigger loop	Close to β′ E1261 (3.9 Å in elongating complex, 2.6 Å in rifampicin complex)
rpoC V1198E	1313	DNA clamp	Hydrophobic region, internal (L1325, V1319, V1281)

a. Position in *S. enterica* α, β, β′ subunits.
b. Corresponding position in *T. thermophilus* α, β and β′ subunits.
c. Based on the *T. thermophilus* elongation complex, PDB code 2O5I (Vassylyev *et al.*, 2007) and the *T. thermophilus* rifampicin complex, PDB code 1YNN (Campbell *et al.*, 2005).

(*rpoC* R1075P/H) falls in a region in *S. enterica* that does not have an obvious counterpart in *Thermus*. The possible mechanisms of compensation, as deduced from the available crystal structures, are discussed below.

Discussion

In this study we evolved Rif^R *S. enterica*, selecting for improved growth fitness, and used genetic mapping and strain reconstructions to identify and characterize fitness-compensatory mutations. Using *S. enterica* is clearly not directly relevant to the evolution of drug resistance in tuberculosis patients. However, an advantage of using a genetically amenable model organism is that we can experimentally test and quantify the phenotypic significance of the putative fitness compensatory mutations isolated in the evolution experiments. Our data are in good agreement with the results from the recent genomic sequence analysis of Rif^R *M. tuberculosis* (Comas *et al.*, 2012) in that we also find a high frequency of fitness-compensatory mutations in *rpoA* and *rpoC* when we evolve a Rif^R mutant with selection for improved growth fitness (Table 1). In addition, we found that compensatory mutations in *rpoB* are also frequently selected. Of the 22 independent fitness-compensatory mutations isolated from 20 lineages, the distribution was: *rpoA*, 4 (18%); *rpoB*, 11 (50%); *rpoC*, 7 (32%). These results are also in agreement with earlier data from *E. coli* where approximately 50% (9/20) of growth compensatory mutations to primary Rif^R mutations in *rpoB* were found in the *rpoB* gene, although in that study the remaining mutations were not identified (Reynolds, 2000).

The genetic reconstructions made here showed that in every case the selected secondary mutation identified in *rpoA*, *rpoB* or *rpoC* was sufficient to account for the measured fitness compensation (Table S2). In addition, it is interesting to note that several of the mutations isolated in *rpoA* and *rpoC* cause a small but significant reduction in susceptibility to RIF, with or without causing a large reduction in growth fitness (Table 2). We do not know the MICs of equivalent mutations in *M. tuberculosis* and so cannot speculate on their clinical significance. However, this finding raises the possibility that the evolution of resistance to RIF could in principle begin with the selection of mutations outside of the classic resistance-determining-region in *rpoB*, and might occur for example, in environments where bacteria are exposed to low levels of antibiotic (Baquero *et al.*, 1998; Gullberg *et al.*, 2011).

Ten of the 20 Rif^R lineages were evolved in the presence of RIF and the remaining 10 in an antibiotic-free environment. Although reversion to the wild-type sequence would have provided the greatest increase in

fitness, all of the isolates tested after selection in the antibiotic-free environment had acquired second-site mutations in *rpoA*, *rpoB* or *rpoC* (Table 1). This bias may reflect the lower probability of acquiring a specific reversion mutation versus acquiring any one of the many different second-site mutations that confer a partial reversion of the growth phenotype (Levin *et al.*, 2000). However, it may also reflect differences in specific mutation rates because one particular mutation, *rpoB* D516G, that causes a small reduction in growth rate compared with the wild-type (Table S1), was selected independently six times, and with equal frequency in the presence and in the absence of RIF. Note that each evolution experiment was inoculated from an independent single colony showing that the compensatory mutations arose independently during the course of the evolution. This mutation is also the compensatory mutation that causes the greatest improvement in growth rate (Table 1). Both mutational events, reversion of R529C and occurrence of D516G, involve the same transition, A:T to G:C. Unless mutation rates differ in each sequence context it is not obvious why D516G (GAC to GGC) should be selected at a higher frequency than reversion of R529C to the wild-type sequence (TGT to CGT). The *rpoB* D516G mutation is also interesting because not only is it a frequent compensatory mutation but it is also a Rif^R mutation in *S. enterica* (RIF MIC 250 μg ml^{-1}) as well as in *E. coli* and *M. tuberculosis* (Campbell *et al.*, 2001; Garibyan *et al.*, 2003). The strain carrying the combination *rpoB* R529C and D516G has a slightly higher RIF MIC than strains carrying *rpoB* R529C alone. Thus, *rpoB* D516G is unusual in that it occurs more frequently that reversion of the original mutation, it compensates for the loss of fitness caused by *rpoB* R529C, and it contributes to a further reduction in the susceptibility of the double mutant to RIF.

RNA polymerase is a highly conserved enzyme among bacteria and a crystal structure of the core RNA polymerase alone (Zhang *et al.*, 1999) and in complex with RIF has been solved from *Thermus aquaticus* (Campbell *et al.*, 2001; 2005). Based on analysis of these structures (Fig. 1, Table 3), and given the high level of amino acid identity with RNA polymerase from *S. enterica*, we suggest that the fitness-compensatory mutations found in this study can be divided into four groups: (i) mutations that restore the properties of the RIF binding site; (ii) mutations that affect interactions between the enzyme and its nucleotide substrates; (iii) mutations that affect the interactions between the enzyme and its RNA product; and (iv) mutations that affect the interactions between enzyme subunits. Each of these putative classes is discussed below.

The primary *rpoB* R529C mutation is located in the channel of the RNA polymerase where the DNA/RNA hybrid of the elongating complex is bound (Fig. 2). The

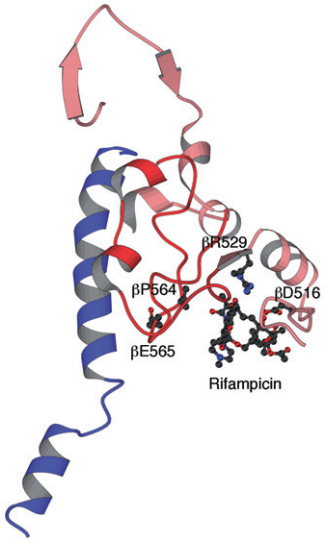

Fig. 2. The central part of the rifampicin complex.
The central part of the rifampicin complex showing the locations of nearby mutations. PDB code 1YNN.

role of the completely conserved arginine residue is probably to stabilize the phosphate backbone of the newly synthesized RNA three and four bases away from the catalytic centre (Vassylyev *et al.*, 2007). Therefore, it is reasonable to believe that an amino acid change to a cysteine at this position alters the properties of the exit channel and the stabilization of the RNA product. This mutation reduces transcription rate and growth rate of the bacteria (Jin and Turnbough, 1994). The first group of compensatory mutations (*rpoB* D516G, *rpoB* P560L, *rpoB* P564S and *rpoB* E565A) are all located at the exit channel close to the original *rpoB* R529C mutation around bases two to four of the synthesized RNA (Figs 2 and 3). This suggests that each of these mutations alters the properties of the exit channel in a way that reduces the effect of the original mutation and thereby increases the growth rate of the bacteria. Two of the mutations remove negative charges, and these may reduce the effect of the loss of the positive charge of the arginine in the primary mutation. One of these mutants, *rpoB* D516G, by itself leads to RIF resistance, and it has been proposed that this mutation in isolation may lead to weaker binding of RIF due to an increased positive charge at the binding site (Campbell *et al.*, 2001).

The second group of compensatory mutations can be divided into two subgroups depending on their location in the RNA polymerase: (i) *rpoB* R637C, *rpoB* H673Y and *rpoC* L770P; and (ii) *rpoC* R1075H, *rpoC* R1075P, *rpoC* G1136A, *rpoC* R1140H and *rpoC* V1198E (Fig. 3, Table 3). The mutations in the first subgroup (*rpoB* R637C, *rpoB* H673Y and *rpoC* L770P) are located very close to each other in the assembled enzyme (Fig. 3). The

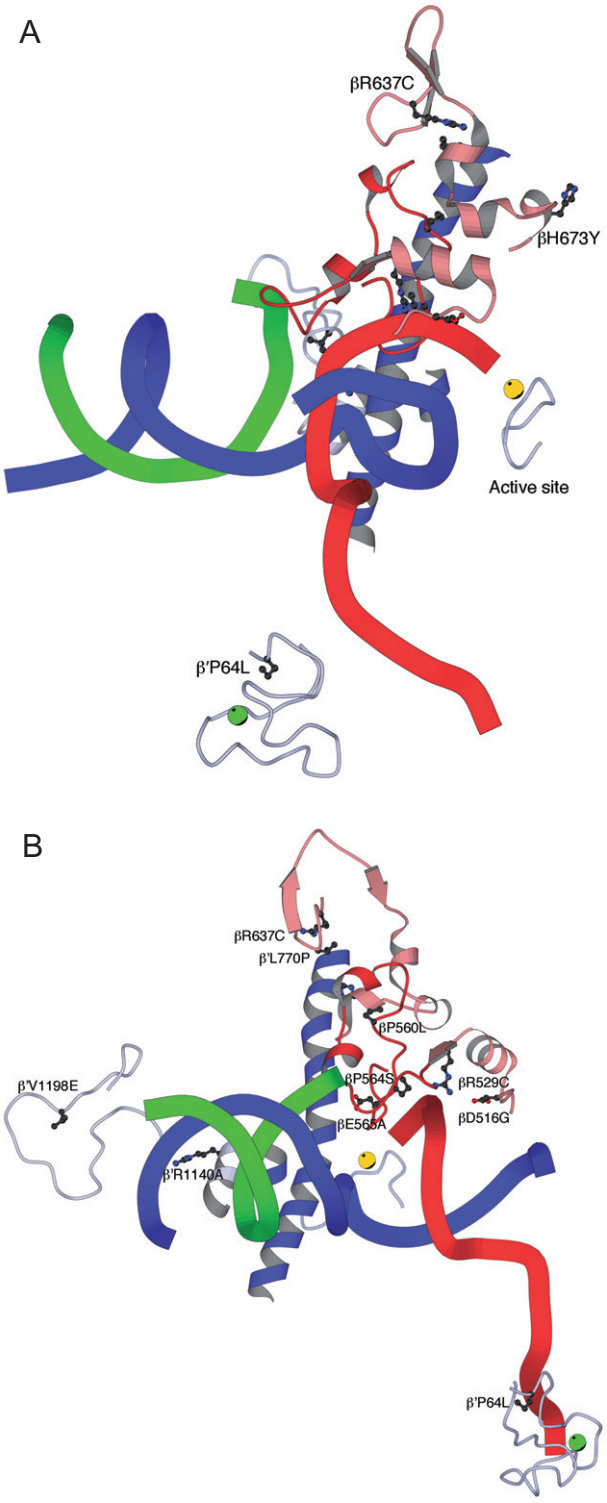

Fig. 3. The central part of the elongating complex.
The central part of the elongating complex with the bridge helix (blue), fork loop 2 (red), the Zn finger (pale blue). All side-chains in the β′ and the β subunits that are mutated are shown and labelled except β′ R1075P/H which is located in a long insert in the trigger loop of *S. enterica* and β′ G1136 that is disordered in this structure (PDB code 2O5I). Its position is at the top of the helix containing the R1140A mutation. (A) In the same orientation as in Fig. 1 (the DNA and RNA molecules can be used as cues for the position in the whole complex), and (B) rotated 90° around a vertical axis so that the mutations are more easily seen.

mutation *rpoC* L770P is found in the bridge helix of the β′ subunit, an important and very flexible structure which interacts with the DNA template strand. The bridge helix is thought to move the template strand through the polymerase and position it in the active centre (Kaplan and Kornberg, 2008; Hein and Landick, 2010) and it has been shown that amino acid substitutions within the bridge helix can increase the elongation rate (Weilbaecher *et al*., 1994; Tan *et al*., 2008). The side-chain of R637 is directly interacting with the bridge helix (Table 3), and all these mutants may influence the dynamics properties of this helix.

The compensatory mutations in the second subgroup (*rpoC* R1075H, *rpoC* R1075P, *rpoC* G1136A, *rpoC* R1140H and *rpoC* V1198E) are located at the other side of the cleft connected by the bridge helix. Two of them, *rpoC* G1136A (Q in *Thermus*) and *rpoC* R1140H, are found in the so-called trigger loop. The trigger loop is a conserved structure that is highly flexible and refolds to assist nucleotide loading and catalysis (Toulokhonov *et al*., 2007). The mutations *rpoC* R1075H and *rpoC* R1075P are found in a long insertion in the trigger loop that is not present in the *Thermus* enzyme. The conformation of this insertion is known from a structural study (Chlenov *et al*., 2005). The position of this insert in the complete structure of the polymerase has been modelled (Opalka *et al*., 2010), but it is still unclear how the insert may influence the refolding of the trigger loop. In previous studies it was shown that the mutation *rpoC* R1075C in the trigger loop can reduce the probability of transcription termination (Weilbaecher *et al*., 1994), while the mutation *rpoC* G1136S and other mutations in this area reduce termination and increase elongation rate (Bar-Nahum *et al*., 2005; Tan *et al*., 2008). The mutant *rpoC* V1198E is found in the small domain adjacent to the trigger loop (Fig. 3). Since this side-chain is packed in the hydrophobic centre of this small domain, the replacement of the non-polar valine by the charged glutamic acid may influence the conformation of the domain. It is thus possible that the compensatory mutations in this group all lead to an altered interaction between the trigger loop and the incoming nucleotides, and thereby increase the elongation rate thus compensating for the growth defect caused by the primary *rpoB* R529C mutation.

The third group consists of one compensatory mutation, *rpoC* P64L, located close to a zinc binding pocket in the β′ subunit of RNA polymerase (Fig. 3). The zinc-binding pocket interacts with the RNA product of polymerization

(King *et al*., 2004). The *rpoC* P64L mutation at this position may change the conformation and affect the cysteine residues inside the binding pocket that are required to chelate the zinc ion, altering interactions between the zinc-binding pocket and the RNA. Since the primary *rpoB* R529C mutation destabilizes the binding between the polymerase and the RNA (Jin and Turnbough, 1994), a change in the binding between the enzyme and the RNA at another position might be the mechanism of compensation.

The fourth and final group of compensatory mutations is found in the α subunit of RNA polymerase (Fig. 1). These mutations (*rpoA* R191C, *rpoA* R191S and *rpoA* Q194P) are all located in an area that has previously been identified to play a crucial role in α subunit dimerization and binding of the dimer to the β′ subunit (Kimura and Ishihama, 1995). In addition, the mutation *rpoA* R191C has been shown to cause a thermo-sensitive phenotype and to reduce transcriptional fidelity (Ishihama *et al*., 1980; Kawakami and Ishihama, 1980) consistent with alterations in the structure and function of the mutated RNA polymerase. The arginine is in one of the α subunits forming a salt link to a glutamate in the β′ subunit. The mutations involving this arginine are likely to influence the stability or the assembly of the complex, but it is difficult to explain from the structure how these mutations in the α subunit influence the properties of the enzyme. However, the selection of these mutations is not just a peculiarity of *Salmonella* because in the recent whole-genome sequencing analysis of *M. tuberculosis* clinical isolates (Comas *et al*., 2012), two mutations were identified at similar positions in the α subunit (R182L, equivalent to position R191 in *Salmonella*; R186C, equivalent to R195 in *Salmonella*). The experimental data show that each of the three selected α subunit mutations substantially increases the growth rate of the original Rif^R R529C mutant (Table 1), and two of them (R191S and Q194P) individually cause a significant reduction in susceptibility to RIF (Table 2).

In conclusion, our results support the recent findings from whole-genome sequence analysis of clinical *M. tuberculosis* strains, showing that the evolution of Rif^R mutants involves the acquisition of multiple mutations in the *rpoA*, *rpoB* and *rpoC* genes of RNA polymerase (Comas *et al*., 2012). Importantly, our study shows by genetic reconstruction that the secondary mutations acquired in each of these three genes are responsible for improving the fitness of a primary Rif^R mutation in *rpoB*. In addition, our findings show that individual mutations in *rpoA* or *rpoC* can reduce susceptibility to RIF. This suggests that exposure of bacteria to low levels of the antibiotic may increase the risk of resistance development, because of the greater number of potential resistance mutations.

Experimental procedures

Bacterial strains and growth conditions

All strains are derived from the virulent *Salmonella enterica* serovar Typhimurium strain ATCC 14028s (Jarvik *et al*., 2010) and are described in Table S3. Transduction mapping and strain constructions were done using P22 HT phage (Schmieger and Backhaus, 1973). Bacteria were grown at 37°C in LB broth and on plates of LB medium solidified with 1.5% agar (Oxoid), 0.2% glucose and 3 mM $CaCl_2$ (LA plates). RIF was added to media as required at 100 μg ml^{-1}. Growth rates were measured using a Bioscreen C machine (Oy Growth Curves Ab Ltd). Overnight cultures were diluted 3000-fold in LB and 300 μl of the diluted culture were incubated at 37°C with continuous shaking in honeycomb microtitre plates, with optical density readings at 5 min intervals and the doubling time was calculated from the increase in optical density at 600 nm.

Evolution by serial passage

Independent lineages of TH7884 were grown in 15 ml tubes with shaking at 37°C overnight in LB and LB + RIF respectively. The lineages were serially passaged after each cycle of growth by transferring 2 μl of overnight culture into 2 ml of fresh media to initiate the next cycle of growth. Each lineage was evolved for six cycles (60 generations), after which dilutions were plated on LA, examined after overnight incubation at 37°C, and the two largest colonies per lineage were picked for further analysis.

PCR amplification and DNA sequencing

To prepare DNA, a single colony was suspended in 100 μl sterile water and boiled for 5 min. The PCR was performed using Ready-To-Go PCR beads (GE Healthcare) in a 25 μl volume containing 0.4 μM forward and reverse Primer and 1 μl DNA sample. PCR conditions were as follows: denaturation at 95°C for 5 min, then 30 cycles of 95°C (30 s), 55°C (30 s) and 72°C (1 min) followed by a final elongation at 72°C for 10 min. Primers for amplification and sequencing of *rpoA*, *rpoB* and *rpoC* are described in Table S4. DNA sequencing was made at Macrogen (Seoul, Korea) using purified PCR product.

Transcription efficiency assay

Overnight cultures were diluted 10^{-3} in LB and grown at 37°C to an OD_{600} of 0.3. Transcription of the *spoU* gene cloned in a pBAD TOPO TA vector (Invitrogen) was induced by adding L-Arabinose to a final concentration of 0.2%. Samples were taken at 0, 0.5, 1, 1.5, 2, 5 and 10 min after induction and stabilized in RNAprotect Bacteria Reagent (Qiagen). Total RNA was prepared using RNeasy Mini Kit (Qiagen) and 5 μg total RNA was DNase treated using Turbo DNA-free kit (Ambion). cDNA was prepared from 500 ng DNA-free RNA using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Quantification of cDNA was performed on a 7300 Real Time PCR System (Applied Biosystems) using Fast SYBR Green Master Mix (Applied Biosystems) in a 20 μl

volume containing 0.375 μM forward and reverse Primer and 5 μl fivefold diluted cDNA sample. PCR conditions were as follows: 50°C for 2 min, then 95°C for 10 min, followed by 40 cycles of 95°C (15 s) and 60°C (1 min). Primers for amplification of *spoU* and *ssrA* (TmRNA) are listed in Table S4. The change in the relative quantity of *spoU* mRNA and TmRNA in every individual sample was calculated using an established mathematical model for relative quantification in real-time PCR (Pfaffl, 2001).

Minimum inhibitory concentration measurements

The RIF MIC values up to 32 μg ml^{-1} were determined using E-test (BioMérieux, France) on Mueller-Hinton agar plates. RIF MIC values in the range 50–3000 μg ml^{-1} were measured by broth microdilution. Overnight cultures were diluted 10^{-3} in LB then 10 μl volumes of diluted culture were inoculated into microtitre plate wells with 190 μl LB containing RIF to give final concentrations of 50, 100, 150, 250, 500, 750, 1500 and 3000 μg ml^{-1}. Cultures were incubated 16–18 h at 37°C before reading.

Acknowledgements

This work was supported by grants from the Swedish Research Council (Vetenskapsrådet), Knut and Alice Wallenberg Foundation (RiboCore Project), EU 7th Framework Program (project PAR), Swedish Strategic Research Foundation (SSF) and the Swedish Governmental Agency for Innovation Systems (Vinnova).

References

Archambault, J., and Friesen, J.D. (1993) Genetics of eukaryotic RNA polymerases I, II, and III. *Microbiol Rev* **57:** 703–724.

Baquero, F., Negri, M.C., Morosini, M.I., and Blazquez, J. (1998) Antibiotic-selective environments. *Clin Infect Dis* **27** (Suppl. 1)**:** S5–11.

Bar-Nahum, G., Epshtein, V., Ruckenstein, A.E., Rafikov, R., Mustaev, A., and Nudler, E. (2005) A ratchet mechanism of transcription elongation and its control. *Cell* **120:** 183–193.

Billington, O.J., McHugh, T.D., and Gillespie, S.H. (1999) Physiological cost of rifampin resistance induced *in vitro* in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **43:** 1866–1869.

Bjorkman, J., Hughes, D., and Andersson, D.I. (1998) Virulence of antibiotic-resistant *Salmonella typhimurium*. *Proc Natl Acad Sci USA* **95:** 3949–3953.

Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S.A. (2001) Structural mechanism for rifampicin inhibition of bacterial *rna* polymerase. *Cell* **104:** 901–912.

Campbell, E.A., Pavlova, O., Zenkin, N., Leon, F., Irschik, H., Jansen, R., *et al*. (2005) Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *EMBO J* **24:** 674–682.

Casali, N., Nikolayevskyy, V., Balabanova, Y., Ignatyeva, O., Kontsevaya, I., Harris, S.R., *et al*. (2012) Microevolution of extensively drug-resistant tuberculosis in Russia. *Genome Res* **22:** 735–745.

Chlenov, M., Masuda, S., Murakami, K.S., Nikiforov, V., Darst, S.A., and Mustaev, A. (2005) Structure and function of lineage-specific sequence insertions in the bacterial RNA polymerase beta′ subunit. *J Mol Biol* **353:** 138–154.

Comas, I., Borrell, S., Roetzer, A., Rose, G., Malla, B., Kato-Maeda, M., *et al*. (2012) Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet* **44:** 106–110.

Conde, M.B., and Lapa, E.S.J.R. (2011) New regimens for reducing the duration of the treatment of drug-susceptible pulmonary tuberculosis. *Drug Dev Res* **72:** 501–508.

Condon, C., Squires, C., and Squires, C.L. (1995) Control of rRNA transcription in *Escherichia coli*. *Microbiol Rev* **59:** 623–645.

Gagneux, S., Long, C.D., Small, P.M., Van, T., Schoolnik, G.K., and Bohannan, B.J. (2006) The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. *Science* **312:** 1944–1946.

Garibyan, L., Huang, T., Kim, M., Wolff, E., Nguyen, A., Nguyen, T., *et al*. (2003) Use of the rpoB gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair (Amst)* **2:** 593–608.

Gullberg, E., Cao, S., Berg, O.G., Ilback, C., Sandegren, L., Hughes, D., and Andersson, D.I. (2011) Selection of resistant bacteria at very low antibiotic concentrations. *PLoS Pathog* **7:** e1002158.

Heep, M., Brandstatter, B., Rieger, U., Lehn, N., Richter, E., Rusch-Gerdes, S., and Niemann, S. (2001) Frequency of rpoB mutations inside and outside the cluster I region in rifampin-resistant clinical *Mycobacterium tuberculosis* isolates. *J Clin Microbiol* **39:** 107–110.

Hein, P.P., and Landick, R. (2010) The bridge helix coordinates movements of modules in RNA polymerase. *BMC Biol* **8:** 141.

Ishihama, A., Shimamoto, N., Aiba, H., Kawakami, K., Nashimoto, H., Tsugawa, A., and Uchida, H. (1980) Temperature-sensitive mutations in the alpha subunit gene of *Escherichia coli* RNA polymerase. *J Mol Biol* **137:** 137–150.

Jarvik, T., Smillie, C., Groisman, E.A., and Ochman, H. (2010) Short-term signatures of evolutionary change in the Salmonella enterica serovar typhimurium 14028 genome. *J Bacteriol* **192:** 560–567.

Jin, D.J., and Gross, C.A. (1988) Mapping and sequencing of mutations in the *Escherichia coli* rpoB gene that lead to rifampicin resistance. *J Mol Biol* **202:** 45–58.

Jin, D.J., and Turnbough, C.L. (1994) An *Escherichia coli* RNA polymerase defective in transcription due to its overproduction of abortive initiation products. *J Mol Biol* **236:** 72–80.

Jindani, A., Dore, C.J., and Mitchison, D.A. (2003) Bactericidal and sterilizing activities of antituberculosis drugs during the first 14 days. *Am J Respir Crit Care Med* **167:** 1348–1354.

Kaplan, C.D., and Kornberg, R.D. (2008) A bridge to transcription by RNA polymerase. *J Biol* **7:** 39.

Kapur, V., Whittam, T.S., and Musser, J.M. (1994) Is *Mycobacterium tuberculosis* 15,000 years old? *J Infect Dis* **170:** 1348–1349.

Kawakami, K., and Ishihama, A. (1980) Defective assembly of ribonucleic acid polymerase subunits in a temperature-sensitive alpha-subunit mutant of *Escherichia coli*. *Biochemistry* **19:** 3491–3495.

Kimura, M., and Ishihama, A. (1995) Functional map of the alpha subunit of *Escherichia coli* RNA polymerase: insertion analysis of the amino-terminal assembly domain. *J Mol Biol* **248:** 756–767.

King, R.A., Markov, D., Sen, R., Severinov, K., and Weisberg, R.A. (2004) A conserved zinc binding domain in the largest subunit of DNA-dependent RNA polymerase modulates intrinsic transcription termination and antitermination but does not stabilize the elongation complex. *J Mol Biol* **342:** 1143–1154.

Levin, B.R., Perrot, V., and Walker, N. (2000) Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics* **154:** 985–997.

Mariam, D.H., Mengistu, Y., Hoffner, S.E., and Andersson, D.I. (2004) Effect of rpoB mutations conferring rifampin resistance on fitness of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **48:** 1289–1294.

Murakami, K.S., and Darst, S.A. (2003) Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol* **13:** 31–39.

Musser, J.M. (1995) Antimicrobial agent resistance in mycobacteria: molecular genetic insights. *Clin Microbiol Rev* **8:** 496–514.

Opalka, N., Brown, J., Lane, W.J., Twist, K.A., Landick, R., Asturias, F.J., and Darst, S.A. (2010) Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. *PLoS Biol* **8:** e1000483.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29:** e45.

Ramaswamy, S., and Musser, J.M. (1998) Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tuber Lung Dis* **79:** 3–29.

Reynolds, M.G. (2000) Compensatory evolution in rifampin-resistant *Escherichia coli*. *Genetics* **156:** 1471–1481.

Sander, P., Springer, B., Prammananan, T., Sturmfels, A., Kappler, M., Pletschette, M., and Bottger, E.C. (2002) Fitness cost of chromosomal drug resistance-conferring mutations. *Antimicrob Agents Chemother* **46:** 1204–1211.

Schmieger, H., and Backhaus, H. (1973) The origin of DNA in transducing particles in P22-mutants with increased transduction-frequencies (HT-mutants). *Mol Gen Genet* **120:** 181–190.

Sekiguchi, J., Miyoshi-Akiyama, T., Augustynowicz-Kopec, E., Zwolska, Z., Kirikae, F., Toyota, E., *et al*. (2007) Detection of multidrug resistance in *Mycobacterium tuberculosis*. *J Clin Microbiol* **45:** 179–192.

Tan, L., Wiesler, S., Trzaska, D., Carney, H.C., and Weinzierl, R.O. (2008) Bridge helix and trigger loop perturbations generate superactive RNA polymerases. *J Biol* **7:** 40.

Telenti, A., Imboden, P., Marchesi, F., Lowrie, D., Cole, S., Colston, M.J., *et al*. (1993) Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *Lancet* **341:** 647–650.

Toulokhonov, I., Zhang, J., Palangat, M., and Landick, R. (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol Cell* **27:** 406–419.

Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H., and Artsimovitch, I. (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448:** 157–162.

Weilbaecher, R., Hebron, C., Feng, G., and Landick, R. (1994) Termination-altering amino acid substitutions in the beta′ subunit of *Escherichia coli* RNA polymerase identify regions involved in RNA chain elongation. *Genes Dev* **8:** 2913–2927.

Werner, F. (2008) Structural evolution of multisubunit RNA polymerases. *Trends Microbiol* **16:** 247–250.

WHO (2010) *Treatment of Tuberculosis: Guidelines*, 4th edn. Geneva: WHO.

Wrande, M., Roth, J.R., and Hughes, D. (2008) Accumulation of mutants in 'aging' bacterial colonies is due to growth under selection, not stress-induced mutagenesis. *Proc Natl Acad Sci USA* **105:** 11863–11868.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999) Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution. *Cell* **98:** 811–824.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.