Substitutions in the *Escherichia coli* RNA polymerase inhibitor T7 Gp2 that allow inhibition of transcription when the primary interaction interface between Gp2 and RNA polymerase becomes compromised

Andrey Shadrin,1 Carol Sheppard,1 Konstantin Severinov,2,3,4 Steve Matthews1 and Sivaramesh Wigneshweraraj1

1Section of Microbiology and MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London SW7 2AZ, UK
2Waksman Institute for Microbiology and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA
3Institute of Molecular Genetics, Russian Academy of Sciences, Moscow 123182, Russia
4Institute of Gene Biology, Russian Academy of Sciences, Moscow 119334, Russia

The *Escherichia coli*-infecting bacteriophage T7 encodes a 7 kDa protein, called Gp2, which is a potent inhibitor of the host RNA polymerase (RNAp). Gp2 is essential for T7 phage development. The interaction site for Gp2 on the *E. coli* RNAp is the β\(^9\) jaw domain, which is part of the DNA binding channel. The binding of Gp2 to the β\(^9\) jaw antagonizes several steps associated with interactions between the RNAp and promoter DNA, leading to inhibition of transcription at the open promoter complex formation step. In the structure of the complex formed between Gp2 and a fragment of the β\(^9\) jaw, amino acid residues in the β\(^3\) strand of Gp2 contribute to the primary interaction interface with the β\(^9\) jaw. The 7009 *E. coli* strain is resistant to T7 because it carries a charge reversal point mutation in the β\(^9\) jaw that prevents Gp2 binding. However, a T7 phage encoding a mutant form of Gp2, called Gp2\(^b\), which carries triple amino acid substitutions E24K, F27Y and R56C, can productively infect this strain. By studying the molecular basis of inhibition of RNAp from the 7009 strain by Gp2\(^b\), we provide several lines of evidence that the E24K and F27Y substitutions facilitate an interaction with RNAp when the primary interaction interface with the β\(^9\) jaw is compromised. The proposed additional interaction interface between RNAp and Gp2 may contribute to the multi-pronged mechanism of transcription inhibition by Gp2.

INTRODUCTION

Central to the regulation of bacterial transcription is the RNA polymerase (RNAp), the enzyme responsible for the synthesis of all RNA in bacteria. The bacterial RNAp is composed of a multisubunit catalytic core (subunit composition \(\alpha_2\beta\beta'\omega\); abbreviated E) and a sigma (\(\sigma\)) factor subunit that determines the specificity of the RNAp-promoter interaction. In *Escherichia coli*, the \(\sigma^{70}\)-containing RNAp (\(\text{E}\sigma^{70}\)) is responsible for the transcription of most genes during exponential growth (Haugen et al., 2008). Transcription by \(\text{E}\sigma^{70}\) begins with the formation of the initial \(\text{E}\sigma^{70}\)-promoter complex known as the closed promoter complex (RPC). The RPC is often very unstable and isomerizes, via several intermediates, to the transcriptionally proficient open promoter complex (RPO). In the RPO, \(\sim 15\) base pairs of promoter DNA around the transcription start site are melted to form the ‘transcription bubble’ and the template DNA strand is positioned at the active centre of the RNAp. The catalytic β and β\(^\prime\) subunits of the RNAp define the main DNA binding channel (DBC), and the active centre where RNA synthesis takes place is located deep within the DBC. The double-stranded DNA immediately downstream of the active centre (dwDNA) interacts with a segment of the DBC called the downstream DBC (dwDBC), and this interaction is essential for the formation and stability of the RPO (Saecker et al., 2011).

Abbreviations: DBC, DNA binding channel; Dt, doubling time; dwDBC, downstream DBC; dwDNA, DNA immediately downstream of the active centre; HRP, horseradish peroxidase; RNAp, RNA polymerase; RPO, open promoter complex.

A supplementary table is available with the online version of this paper.
The RPo formation also represents an important regulatory step during transcription initiation, and thus the activity of the bacterial RNAp is controlled by an array of transcriptional factors, which repress, stimulate or modulate its ability to form the RPo (Brown et al., 2004). Most bacteriophages (phages) rely on the RNAp of their host bacteria for expression of their genes during the infection process. Therefore, RPo formation by the bacterial RNAp is also subjected to regulation by phage-encoded proteins through covalent RNAp modifications and modulation by RNAp-binding proteins (Nechaev & Severinov, 2006). The biological role of T7 Gp2 is to ensure efficient and coordinated transcription of phage genes by the T7 RNAp without interference from bacterial RNAp (Savalia et al., 2010). In the solution structure of the complex of Gp2 with a fragment of the β′ jaw domain (E. coli RNAp β′ residues 1153–1213), two invariant arginine residues (R56 and R58; R56 and R58 are invariant in 25 and 23, respectively, out of 25 Gp2-like proteins in phages CR8, CR44b, K1F and EcoDS1, is also underlined in bold type. (c) Alignment of the β′ jaw domain amino acid residues 1153–1213 from a representative set of host RNAPs, with those infected by T7-like phages encoding Gp2-like proteins indicated by asterisks. Amino acid residues corresponding to E1158 and E1188 of the E. coli RNAp are shown in bold type. (d) As in (a), except that the structural models of Gp2-like proteins CR44b_13 (orange) and CR8_16 (pink) are superimposed and amino acid residue R15 is indicated. The structural models of the Gp2 homologues were calculated with SWISS-MODEL using NMR structures of T7 Gp2 as a template (Arnold et al., 2006; Camara et al., 2010).

**Methods**

**Bacterial strains and plasmids.** Table S1 (available with the online version of this paper) lists the bacterial strains and plasmids used in this study. Plasmid pCS1:Gp2 and derivatives were constructed by digesting pSW33:Gp2 (and derivatives) with XbaI and HindIII, and the 0.3 kb fragment containing gene 2 was cloned into the same sites in pBAD18 (Invitrogen). The coding sequences of Gp2 homologues CR44b protein 13 (CR44b_13) and CR8 protein 16 (CR8_16) were PCR-amplified from phage genomic DNA (kindly provided by Ana Louisa Toribio and Gordan Dougan, The Wellcome Trust Sanger Institute, Cambridge, UK) with primers containing restriction sites for NdeI and BamHI and cloned into the same sites in plasmid pET33b+ (Novagen) to generate pAS33:CR44b_13 and pAS33:CR8_16, respectively. Similarly, plasmids pAS1:CR44b_13 and pAS1:CR8_16 were constructed using the procedure described for pCS1:Gp2. Mutant forms of T7 Gp2 and CR44b_13 were constructed using the QuikChange Site-Directed PCR Mutagenesis kit (Stratagene) using pCS1:Gp2, pSW33:Gp2 or pAS1:CR44b_13 as a template. T7 Gp2 was overexpressed and purified as described previously: CR44b_13 and CR8_16 were overexpressed and purified using the AttE system (Takara).

The molecular basis by which E24K, F27Y and R56C substitutions in Gp2 suppress the E1188K mutation in the β′ jaw domain, which prevents wild-type Gp2 from binding to the RNAp, is not known. In the context of the Gp2–β′ jaw domain fragment structure, the E24K and F27Y substitutions are located in the middle and close to the end, respectively, of the loop connecting the β1 and β2 strands in Gp2, i.e. at the opposite side to the β3 strand, which forms the interface with the β′ jaw domain and carries R56 (Fig. 1a). In this study, we established a simple in vivo bacterial growth attenuation assay to assess the activity of recombinant Gp2 on E. coli RNAp in the absence of T7 infection, to study the molecular basis by which Gp2 inhibits wild-type, 7009 and BR3 E. coli RNAp. The results strongly suggest that the E24K and F27Y mutations facilitate the interaction between Gp2 and RNAp when the primary interaction interface between Gp2 and RNAp is compromised. Experiments addressing the loop interconnecting β1 and β2 strands in Gp2 homologues from two T7-like phages that infect Citrobacter rodentium further corroborate our results.
The expression of Gp2 was induced at OD600 of ~0.2–0.25 by adding multiwell plate reader (BMG Labtech). The seed cultures were diluted also performed in 96-well microtitre plates in a POLARstar Omega replicates were performed for each growth curve.

Bacterial growth attenuation assays. Seed cultures were grown at 37 °C, shaking at 700 r.p.m. for 6–7 h in a THERMOrstar (BMG Labtech) plate incubator by directly inoculating a colony from a freshly transformed Luria agar plate into 200 μl of Luria broth (LB) medium (Difco) containing 100 μg ampicillin ml⁻¹ and 0.5 % (w/v) glucose (to prevent leaky expression of Gp2 from the P_BAD promoter) into a 96-well microtitre plate (Sterilin). The experimental growth curves were also performed in 96-well microtitre plates in a POLARstar Omega multwell plate reader (BMG Labtech). The seed cultures were diluted 1:100 in a final volume of 200 μl of fresh LB medium containing 100 μg ampicillin ml⁻¹ and incubated at 30 °C, shaking at 500 r.p.m. The expression of Gp2 was induced at OD600 of ~0.2–0.25 by adding 0.04 % (w/v) arabinose for 7009 cultures. The doubling time (Dt) was calculated from the gradient of the growth curve typically between the OD600 corresponding to induction time and OD600 0.7–0.8. The R² values in each case were >0.98. At least three biological and technical replicates were performed for each growth curve.

Quantitative Western blotting. E. coli MG1655 cultures were grown in conditions as described above. At 2 h post induction (see text for details), cell samples were taken and the total amount of Gp2 molecules per cell was determined by Western blotting using standard protocols. E. coli whole-cell extract (corresponding to 2.5 × 10⁷ cells) was analysed by SDS-PAGE on gels that were calibrated with known amounts of purified Gp2. Proteins were transferred to a Hybond-ECL nitrocellulose membrane using a Trans-Blot Semi-Dry transfer cell (Bio-Rad). To detect Gp2, anti-His monoclonal antibodies conjugated with horseradish peroxidase (HRP) (Sigma) in combination with the ECL SuperSignal West Femto Chemiluminescent Substrate kit (Pierce) were used. Digital images of the blots were obtained using an LAS-3000 Fuji Imager, and signal quantification and calculations were performed exactly as described by Piper et al. (2009). Briefly, the amount of Gp2 (ng) (Fig. 2d, table, column 2) at each time point post induction was estimated from the calibration curve. The number of molecules of Gp2 per cell (Fig. 2d, table, column 3) was determined by first calculating the total number of moles of Gp2 [mass (g/l)/Mw (10,000 Da)], multiplying by Avogadro’s constant (6.022 × 10²³ mol⁻¹) and then by dividing the total number of Gp2 molecules by the number of cells in the sample (i.e. 2.5 × 10⁷ cells).

In vitro transcription assays. These were performed exactly as previously described (Câmara et al., 2010).

Efficiency of plaque formation assays. These were performed exactly as previously described (Câmara et al., 2010; Rokytta et al., 2002).

RESULTS

Insights from multiple protein sequence alignment of T7 Gp2 homologues

To determine how amino acids at positions corresponding to T7 Gp2 E24, F27 and R56 compare with corresponding residues in homologous proteins encoded by related phages, we conducted a BLAST search using standard search parameters and T7 Gp2 as a query sequence. The multiple protein sequence alignment of known Gp2 homologues (in the EBI database, July 2012) was done using COBAt (Fig. 1b) (Papadopoulos & Agarwala, 2007). We also built an alignment of the T7 Gp2 target (the β’ jaw fragment region, E. coli β’ amino acids 1153–1213) with corresponding sequences from several bacterial species (Fig. 1c). As expected, the amino acid residues R56 and R58 in T7 Gp2 that are essential for the binding to the E. coli β’ jaw are identical in 25 and 23, respectively, out of 25 known Gp2 homologues (Fig. 1b). Similarly, amino acid residues E1158 and E1188 display a high degree of conservation only in bacteria infected by phages encoding Gp2 homologues (Fig. 1c). Overall, this observation is consistent with previous structural and biochemical analyses which show that the primary interaction interface between T7 Gp2 and the E. coli RNAp involves R56 and R58 in strand β3 of Gp2 and E1158 and E1188 in the β’ jaw domain.

A phenylalanine residue at a position corresponding to F27 of T7 Gp2 is conserved in 18 out of 24 Gp2 homologues, whereas amino acids at a position corresponding to E24 in T7 Gp2 display a significant degree of variation (Fig. 1b). Interestingly, we note that phages CR8, CR44b, K1F and EcoDS1 encode Gp2-like proteins that contain a positively charged amino acid (R15) at the position that corresponds to T7 Gp2 S23, a residue immediately adjacent to E24 that is changed for lysine in Gp2β (Fig. 1b, underlined). From structural models generated for Gp2 homologues from phages CR8 and CR44b and based on the solution structure of the T7 Gp2–β’ jaw domain fragment complex, it is evident that R15 in CR44b_13 and CR8_16 (and by extension also in Gp2 homologues encoded by K1F and EcoDS1) is located in the loop interconnecting β1 and β2 strands of the proteins (Fig. 1d). Thus, it is possible that Gp2 homologues from phages CR8, CR44b, K1F and EcoDS1, like T7 Gp2β, can also inhibit 7009 and/or BR3 RNAp. If true, this would indicate that Gp2 homologues encoded by wild-type CR8, CR44b, K1F and EcoDS1 phages interact with RNAp like the mutant Gp2 encoded by T7β.

A simple in vivo assay measures the activity of Gp2 in the absence of T7 phage infection

To experimentally address the observations derived from multiple sequence alignment and to understand the molecular basis by which the triple amino acid substitutions in Gp2β suppress the E1188K mutation in the β’ jaw domain, we developed a simple in vivo assay to determine and compare the ability of wild-type and mutant Gp2 proteins to inhibit E. coli RNAp in the absence of T7 infection. In this assay, gene 2 (which encodes Gp2) is placed under the control of an arabinose-inducible araBAD promoter (P_BAD) in plasmid pCS1:Gp2 and introduced...
Fig. 2. Establishing a simple in vivo assay to measure the activity of Gp2 in the absence of T7 phage infection. Growth curves of (a) E. coli MG1655 and (b) 7009 cells expressing Gp2 upon induction with arabinose at t=0. The solid red, dotted and solid black lines represent cells harbouring pCS1 : Gp2WT, pBAD18 (empty vector) and pCS1 : Gp2R56E, respectively; the dashed line represents cells harbouring pCS1 : Gp2WT to which ampicillin was not added to the growth medium. (c) As in (a), except that at t=2 h, the cells were pelleted and resuspended in fresh LB to remove inducer (arabinose) prior to continuation of growth. The dashed red line represents cells that were induced for the second time (see text for details) at t=6 h. (d) Quantification of Gp2 molecules per E. coli MG1655 cell at t=0.5, 1 and 2 h post induction of Gp2. Top panel: digitalized image of the Western blot probed with HRP-conjugated anti-His monoclonal antibodies containing purified Gp2 (as a reference standard) and whole-cell protein extracts corresponding to 2.5×10^7 cells. Bottom-left panel: calibration curve relating signal intensity to the amount of purified Gp2 (in ng). Bottom-right panel: table showing the calculated number of molecules of Gp2 per cell for each time point sampled post induction.
Table 1. Dt values for E. coli MG1655, 7009 and BR3 cells expressing wild-type Gp2 and mutant Gp2 containing different combinations of point mutations based on Gp2β

| Gp2                  | Ec MG1655 Dt* (min) ± SD† | FI‡ | Ec 7009 Dt (min) ± SD | FI§ | Ec BR3 Dt (min) ± SD | FI‖ |
|----------------------|---------------------------|-----|----------------------|-----|---------------------|-----|
| Empty vector         | 53 ± 5                    | 1.0 | 54 ± 3               | ND  | 66 ± 1              | ND  |
| Wild-type            | ND                        | ND  | 58 ± 1               | 1.0 | 117 ± 6             | 1.0 |
| E24K                 | ND                        | ND  | 57 ± 4               | 1.0 | 179 ± 6             | 1.5 |
| F27Y                 | ND                        | ND  | 54 ± 3               | 1.0 | 176 ± 10            | 1.5 |
| E24K/F27Y            | ND                        | ND  | 57 ± 2               | 1.0 | 249 ± 7             | 2.1 |
| R56C                 | 64 ± 1                    | 1.2 | 65 ± 2               | 1.2 | 102 ± 4             | 0.9 |
| E24K/R56C            | 221 ± 8                   | 4.1 | 85 ± 3               | 1.6 | 166 ± 9             | 1.4 |
| F27Y/R56C            | 93 ± 11                   | 1.8 | 83 ± 5               | 1.6 | 151 ± 2             | 1.3 |
| E24K/F27Y/R56C       | 280 ± 48                  | 5.3 | 142 ± 5              | 2.6 | 262 ± 1             | 2.2 |
| Gp2β + R58E          | 52 ± 1                    | 1.0 | 59 ± 8               | 1.1 | 73 ± 5              | 0.6 |

*Dt (doubling time) in LB at 30 °C in microtitre plate.
†SD is based on at least three independent measurements for each strain/plasmid combination.
‡Fold increase in Dt with respect to MG1655 E. coli cells containing the empty vector.
§Fold increase in Dt with respect to 7009 E. coli cells expressing wild-type Gp2.
‖Fold increase in Dt with respect to BR3 E. coli cells expressing wild-type Gp2.

into E. coli strain MG1655. As shown in Fig. 2(a), induction of Gp2 expression upon addition of 0.04 % (w/v) arabinose (at t=0) results in efficient attenuation of bacterial growth. The attenuation of bacterial growth is specific, since it is not observed in wild-type cells transformed with the pCS1::Gp2βR56E plasmid, which encodes a functionally defective Gp2 mutant (Câmara et al., 2010) (Fig. 2a), or in E. coli strain 7009 transformed with pCS1::Gp2 (Fig. 2b). Control experiments established that attenuation of bacterial growth was not caused by a loss of selection due to the inhibition of transcription of the β-lactamase gene from pCS1, since attenuation of bacterial growth upon arabinose induction was observed in the absence of ampicillin (Fig. 2a, b). Additional control experiments, shown in Fig. 2(c), established that recombinant Gp2, when expressed in E. coli, acts like a bacteriostatic agent: growth-attenuated E. coli MG1655 cells obtained 2 h (i.e. at t=2) after Gp2 induction resumed growth when inoculated into fresh growth medium after removal of the inducer, and stopped growing upon reinduction with arabinose at t=6 h (Fig. 2c).

Further analysis of whole-cell extracts from cells obtained at t=2 h after Gp2 induction revealed that the total number of Gp2 molecules per E. coli cell is ~5000 molecules (Fig. 2d). Since the total number of RNAp molecules in an exponentially growing E. coli cell is estimated to be ~2000–2500 molecules (Ishihama, 1999), it seems that the total number of Gp2 molecules per E. coli cell exceeds that of the total number of RNAp molecules per E. coli cell by at least twofold under conditions where attenuation of bacterial growth is observed. In summary, the in vivo bacterial growth attenuation assay established here can be used to determine and compare the activity of Gp2 mutants to inhibit the E. coli RNAp in vivo in the absence of T7 phage infection.

**E24K and F27Y substitutions facilitate Gp2 interaction with RNAp when the primary interaction interface is compromised**

It is not known which one of the three changes (or a combination thereof) restores the binding of Gp2β to the 7009 RNAp. Therefore, we constructed single, double and triple mutant versions of Gp2 based on Gp2β and tested their ability to inhibit the wild-type, 7009 and BR3 E. coli RNAp using appropriate strains and the in vivo bacterial growth attenuation assay described above (Fig. 2). Consistent with previous results from alanine-scanning mutagenesis (Câmara et al., 2010), single or double amino acid substitutions at E24 (to K) and/or F27 (to Y) attenuated the growth of E. coli MG1655 to the same degree, as did the wild-type Gp2. On the other hand, expression of Gp2R56C did not significantly affect the growth rate (expressed as Dt in Table 1) as compared with cells harbouring the empty plasmid vector (pBAD18), indicating that the R56C substitution strongly affected the interaction with the wild-type RNAp. However, the Dt of cells expressing double mutant Gp2 variants containing the R56C mutation in combination with either the E24K or the F27Y mutation (i.e. Gp2(E24K/R56C) and Gp2(F27Y/R56C)) was detectably increased (by ~4.1- and ~1.8-fold, respectively) compared with the growth rate of cells harbouring the empty plasmid vector (Fig. 3a, Table 1). Although compared with wild-type Gp2, Gp2_E24K, Gp2_F27Y and Gp2_E24K/F27Y mutants, the Gp2(E24K/R56C) and Gp2(F27Y/R56C) mutants attenuated the growth of MG1655 E. coli cells relatively weakly (Fig. 3a,
Inhibition of *E. coli* RNA polymerase by T7 Gp2

Fig. 3. Ability of single, double and triple mutant versions of Gp2 based on Gp2<sup>b</sup> to attenuate bacterial growth upon induction. Growth curves of (a) *E. coli* MG1655, (b) 7009 and (c) BR3 expressing wild-type Gp2 and Gp2 mutants containing various combinations of mutations based on Gp2<sup>b</sup>. Gp2 expression was induced with arabinose at *t*=0. In each graph, the solid red, dotted and dashed black lines represent cells harbouring pCS1 : Gp2<sup>WT</sup>, pBAD18 (empty vector) and pCS1 : Gp2 encoding the indicated mutant form of Gp2, respectively. In panels (a–c), the same relevant traces for cells harbouring pCS1 : Gp2<sup>WT</sup> and pBAD18 (empty vector) were used.
Table 1), the effect was highly reproducible. When present together, the E24K and F27Y mutations further improved the ability of Gp2b\(^{R56C}\) (i.e. Gp2\(^{b}\)) to attenuate wild-type E. coli growth (Fig. 3a), and the Dt of cells expressing Gp2\(^{b}\) was increased ~5.3-fold compared with that of cells harbouring the empty plasmid vector (Table 1). In summary, it appears that mutations in amino acids residues in and surrounding the loop interconnecting \(\beta_1\) and \(\beta_2\) strands in Gp2 (i.e. E24K and F27Y) can compensate for mutations in Gp2 (i.e. R56C) that compromise (but not abolish; see below) the interaction interface between Gp2 and the \(\beta_9\) jaw domain. Based on Dt values calculated for MG1655 E. coli cells after induction of Gp2\(^{E24K/R56C}\), Gp2\(^{E24K/F27Y}\) and Gp2\(^{b}\) (Table 1), it seems that the charge reversal substitution at E24 (i.e. E24K) in the loop interconnecting the \(\beta_1\) and \(\beta_2\) strands in Gp2 rather than the F27Y mutation at the beginning of the \(\beta_2\) strand is more important in enabling the interaction between Gp2 and the RNAp when the primary interaction interface in the \(\beta_3\) strand is compromised.

T7 Gp2 amino acids 14–59 are sufficient to inhibit the E. coli RNAp

It is evident from the alignment (Fig. 1b) that the highest degree of sequence conservation occurs between residues

## Table 2. Dt values for E. coli strain MG1655 expressing wild-type and mutant versions of Gp2\(^{14–59}\)

| Gp2          | Ec MG1655 Dt* (min) ± SD† | FI‡ |
|--------------|---------------------------|-----|
| Empty vector | 49 ± 1                    | 1.0 |
| Wild-type    | ND                        | ND  |
| Gp2\(^{14–59}\) | 239 ± 16                 | 4.9 |
| Gp2\(^{14–59}\) (R56C) | 57 ± 2                  | 1.2 |
| Gp2\(^{14–59}\) (E24K/F27Y) | 338 ± 25                | 6.9 |

*Dt (doubling time) in LB at 30 °C in microtitre plates.
†SD is based on at least three independent measurements for each strain/plasmid combination.
‡Fold increase in Dt with respect to MG1655 E. coli cells containing the empty vector.
corresponding to T7 Gp2 amino acids 14–59. Therefore, it is possible that amino acids outside this region are dispensable for activity. We used the in vivo bacterial growth attenuation assay to determine whether the truncated form of Gp2 (Gp2 14–59) is functionally active. Results shown in Fig. 4(a) and Table 2 reveal that expression of Gp2 14–59 detectably attenuates MG1655 E. coli cell growth, though not as efficiently as full-length Gp2 expression. The Dt of MG1655 cells expressing Gp2 14–59 is increased ~4.9-fold when compared with MG1655 E. coli cells harbouring the empty plasmid vector (Table 2). Since further N- and C-terminal truncations (e.g. Gp2 16–59, Gp2 14–58 and Gp2 14–55) resulted in proteins unable to affect cell growth rate (data not shown), the results suggest that a minimal functional region of T7 Gp2 comprises amino acids 14–59. The attenuation of cell growth by Gp2 14–59 is specific, since no growth attenuation is detected in the context of the R56C mutation (Fig. 4b). However, the E24K and F27Y mutations clearly improve the efficiency by which Gp2 14–59 attenuates MG1655 cell growth (Fig. 4c). The Dt of MG1655 E. coli cells expressing Gp2 14–59 carrying the E24K/F27Y double substitution is increased ~6.9-fold compared with that of cells harbouring an empty plasmid vector (Table 2). This result indicates that E24K and F27Y substitutions improve the binding of Gp2 to the RNAp even if the primary interaction interface with the β’ jaw domain is intact but the binding affinity is attenuated by N- and C-terminal deletion of evolvable variable segments.

Gp2 homologues from phages that infect C. rodentium inhibit E. coli RNAp even if the interaction with the β’ jaw is compromised

To obtain further experimental evidence that E24K and F27Y substitutions in Gp2 facilitate the interaction of Gp2 with RNAp when the primary interaction interface with the β’ jaw domain is compromised, we conducted experiments with Gp2 homologues, proteins 13 and 16 from phages CR44b and CR8 (CR44b_13 and CR8_16), which infect C. rodentium, a model animal bacterial pathogen used to study human infections by enteropathogenic and enterohaemorrhagic E. coli. Initially, we wanted to determine whether recombinant CR44b_13 and CR8_16 can complement a T7 mutant phage (T7dam64) harbouring a mutation in gene 2 (Burck & Miller, 1978). T7dam64 infections of wild-type E. coli BL21 cells are not productive, since Gp2 function is essential for phage development. However, E. coli BL21 cells containing pSW33 : Gp2 WT (which allows expression of Gp2 from a T7 RNAp-dependent promoter) are successfully infected by T7dam64, as judged by the efficiency of plaque formation, which gave an efficiency of plating (EOP) value of 1.0 ± 0.18. EOP was measured using the following equation: EOP = (number of T7dam64 plaques on strain + test plasmid)/(number of T7dam64 plaques on strain + pSW33 : Gp2 WT). In contrast, the EOP was significantly reduced (to 0 ± 0.01) on lawns of E. coli BL21 cells containing pSW33 : Gp2 R56E, which encodes a functionally defective Gp2 mutant. When plasmids encoding recombinant CR44b_13 (pAS33 : CR44b_13) were transformed into E. coli BL21, T7dam64 formed plaques, albeit at 40% reduced efficiency compared with E. coli BL21 cells harbouring pSW33 : Gp2 WT (EOP of 0.61 ± 0.02). However, T7dam64 was able to form plaques equally well on E. coli BL21 cells harbouring a plasmid encoding CR8_16 (pAS33 : CR8_16) (EOP of 1.01 ± 0.21) and wild-type Gp2 (pSW33 : Gp2 WT). Thus, we conclude that CR44b protein 13 and CR8 protein 16 can partially and fully, respectively, substitute for T7 Gp2 during T7 infection.

Next, we conducted a growth attenuation assay to establish whether CR44b_13 and CR8_16, like T7 Gp2, bind to the β’ jaw domain of the E. coli RNAp. As shown in Fig. 5(a), induction of CR44b_13 and CR8_16 expression with 0.04% (w/v) arabinose from pAS1 : CR44b_13 and pAS1 : CR8_16, respectively, in wild-type E. coli MG1655, resulted in efficient growth attenuation. However, no growth attenuation of E. coli strain JE1134, which encodes a β’ subunit with residues 1149–1190 deleted (Fig. 5b), was observed. We next determined whether CR44b_13 and CR8_16 can inhibit E. coli RNAp harbouring either the E1158K (i.e. found in BR3 E. coli strain) or E1188K (i.e. found in 7009 E. coli strain) mutations, as suggested by sequence analysis (above). As shown in Fig. 5(c), whereas the growth of E. coli strain 7009 was not detectably attenuated by either CR44b_13 or CR8_16, the growth of E. coli strain BR3 was efficiently attenuated by both phage proteins (Fig. 5d). In vitro transcription assays, which report the ability of T7 Gp2 CR44b_13 and CR8_16 proteins to inhibit the synthesis of the tetrancleotide RNA product (ApApUpU) from the UV5 promoter by the RNAp containing σ 70, with purified wild-type, BR3 and 7009 RNAp and T7 Gp2, confirmed the in vivo results (Fig. 5e).

If a charge reversal substitution (E24K) in T7 Gp2 β’ indeed facilitates binding to the RNAp when the primary contact with the β’ jaw domain is weakened (Fig. 3), then R15 of CR44b_13 should be important for attenuation of BR3 E. coli growth. Results shown in Fig. 5(f) confirm that this is indeed the case: the mutant variant of CR44b_13 harbouring the R15E mutation fails to attenuate the growth of E. coli strain BR3 upon induction. However, the R15E mutation has no detectable effect on the ability of CR44b_13 Gp2 to attenuate the growth of wild-type E. coli. Overall, results with Gp2-like proteins encoded by C. rodentium phages provide further support for the view that the molecular basis by which Gp2 β’ is able to inhibit the 7009 RNAp is that E24K and F27Y substitutions facilitate an interaction between Gp2 and RNAp when the primary interaction interface is compromised.

**DISCUSSION**

T7 Gp2 binds tightly to and potently inhibits E. coli RNAp. Previous biochemical and structural studies have shown that amino acid residues in the β3 strand of Gp2 contribute to the primary interaction interface with the β’ jaw domain...
of the RNAp. Results from the current study show that substitutions in amino acid residues of Gp2 located in the region surrounding and including the loop interconnecting the $\beta_1$ and $\beta_2$ strands, and, therefore, located on the opposite side to the $\beta_3$ strand, can compensate for amino acid substitutions in Gp2 and/or the $\beta_9$ jaw domain that compromise the primary interaction interface between Gp2 and the RNAp. This seems to constitute the molecular basis by which T7 $\beta$ (encoding Gp2$\beta$ harbouring the E24K, F27Y and R56C substitutions) can successfully infect the 7009 E. coli strain (harbouring the E1188K mutation) that is resistant to infection by wild-type T7 phage.

Fig. 5. Experiments with Gp2-like proteins from C. rodentium phages CR44b and CR8. Growth curves of (a) E. coli MG1655, (b) JE1134, (c) 7009 and (d) BR3 cells in which the expression of T7 Gp2 and Gp2-like proteins CR44b_13 and CR8_16 was induced at $t=0$ (solid red lines). In each graph, the dotted line represents cells harbouring pBAD18 (empty vector). (e) Autoradiograph of a 20% (w/v) denaturing gel showing synthesis of the transcript ApApUpU (the $^{32}$P-labelled nucleotides underlined) from the lacUV5 promoter by the $\sigma^{70}$-containing RNAp in the presence of Gp2, CR44b_13 and CR8_16. The molar ratio of RNAP to Gp2/Gp2-like proteins is indicated above each lane. The percentage of ApApUpU transcript synthesized (%A) by the RNAp is calculated relative to the reaction without Gp2 and is indicated at the bottom of each lane. (f) Growth curves of E. coli MG1655 (left) and BR3 (right) cells expressing CR44b_13 and harbouring the R15E amino acid substitution. Induction with arabinose was performed at $t=0$. 

A. Shadrin and others
Since the E24K and F27Y substitutions allow (1) the functionally defective Gp2 mutant with the R56C substitution effectively to attenuate wild-type *E. coli* growth (Fig. 3a, Table 1) and (2) detectably improve the ability of the truncated Gp2 form, Gp214–59, to attenuate wild-type *E. coli* growth (Fig. 4, Table 2), it is possible that the region of Gp2 surrounding and including the loop interconnecting the β1 and β2 strands contributes to an ‘auxiliary’ interaction interface between Gp2 and RNAP and contributes to the mechanism by which Gp2 efficiently inhibits RPo formation by the RNAP: the overall architecture of RNAP is reminiscent of a ‘crab claw’. The β and β′ subunits form the two ‘pincers’ of the claw and define the DBC and dwDBC (see Introduction). During RPo formation access to the DBC and dwDBC is controlled by large-scale movements of the β′ subunit (Mekler et al., 2002). Thus, with respect to accessibility of the DBC, RNAP can exist in a so-called ‘closed state’ (in which the width of the DBC is insufficient to allow access of double-stranded DNA) or in an ‘open state’ (in which the DBC is sufficiently wide to allow access of double-stranded DNA) (Chakraborty et al., 2012). The open state is required for RPo formation, and after loading and unwinding of DNA to form the transcription bubble, the RNAP converts into the closed state (Chakraborty et al., 2012; Rutherford et al., 2009). However, in terms of interaction with the RNAP, Gp2 seems to be analogous to the transcription anti-termination factor RfAH, which binds to β′ clamp helices and a β gate loop, which are located directly across from each other on the two sides of the DBC and thereby lock the RNAP in the closed state to ensure transcription processivity, reducing pausing and termination (Sevostyanova et al., 2011). Experiments to verify the model in which Gp2 simultaneously interacts with the β and β′ pincers to inhibit RPo formation are currently in progress.

**ACKNOWLEDGEMENTS**

This work was supported by grants from the BBSRC and Wellcome Trust to S.W., who is also a recipient of a BBSRC David Phillips Fellowship (BB/E023703). We are grateful to Ana Luisa Toribio and Gordon Dougan for sharing unpublished data on *C. rodentium* replication of bacteriophage T7 DNA. We thank Dr Ellen James for technical advice and comments on the manuscript.

**REFERENCES**

Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. (2006). The swiss-model workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22, 195–201.

Belogurov, G. A., Vassylyeva, M. N., Svetlov, V., Klyuyev, S., Grishin, N. V., Vassylyev, D. G. & Artsimovitch, I. (2007). Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol Cell* 26, 117–129.

Browning, D. F. & Busby, S. J. (2004). The regulation of bacterial transcription initiation. *Nat Rev Microbiol* 2, 57–65.

Burck, K. B. & Miller, R. C., Jr (1978). Marker rescue and partial replication of bacteriophage T7 DNA. *Proc Natl Acad Sci U S A* 75, 6144–6148.

Cámara, B., Liu, M., Reynolds, J., Shadrin, A., Liu, B., Kwok, K., Simpson, P., Weinzierl, R., Severinov, K. & other authors (2010). T7 phage protein Gp2 inhibits the *Escherichia coli* RNA polymerase by antagonizing stable DNA strand separation near the transcription start site. *Proc Natl Acad Sci U S A* 107, 2247–2252.
Chakraborty, A., Wang, D., Ebright, Y. W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H. & other authors (2012). Opening and closing of the bacterial RNA polymerase clamp. Science 337, 591–595.

Chamberlin, M. (1974). Isolation and characterization of prototrophic mutants of Escherichia coli unable to support the intracellular growth of T7. J Virol 14, 509–516.

Ederth, J., Artsimovitch, I., Isaksson, L. A. & Landick, R. (2002). The downstream DNA jaw of bacterial RNA polymerase facilitates both transcriptional initiation and pausing. J Biol Chem 277, 37456–37463.

Ederth, J., Mooney, R. A., Isaksson, L. A. & Landick, R. (2006). Functional interplay between the jaw domain of bacterial RNA polymerase and allele-specific residues in the product RNA-binding pocket. J Mol Biol 356, 1163–1179.

Haugen, S. P., Ross, W. & Gourse, R. L. (2008). Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Microbiol 6, 507–519.

Ishihama, A. (1999). Modulation of the nucleoid, the transcription apparatus, and the translation machinery in bacteria for stationary phase survival. Genes Cells 4, 135–143.

James, E., Liu, M., Sheppard, C., Mekler, V., Câmara, B., Liu, B., Simpson, P., Cota, E., Severinov, K. & other authors (2012). Structural and mechanistic basis for the inhibition of Escherichia coli RNA polymerase by T7 Gp2. Mol Cell 47, 755–766.

Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Reyvakin, A., Kapanidis, A. N., Niu, W., Ebright, Y. W., Levy, R. & Ebright, R. H. (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. Cell 108, 599–614.

Mekler, V., Minakhin, L., Sheppard, C., Wigneshweraraj, S. & Severinov, K. (2011). Molecular mechanism of transcription inhibition by phage T7 gp2 protein. J Mol Biol 413, 1016–1027.

Nechaev, S. & Severinov, K. (1999). Inhibition of Escherichia coli RNA polymerase by bacteriophage T7 gene 2 protein. J Mol Biol 289, 815–826.

Nechaev, S. & Severinov, K. (2003). Bacteriophage-induced modifications of host RNA polymerase. Annu Rev Microbiol 57, 301–322.

Papadopoulos, J. S. & Agarwala, R. (2007). COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics 23, 1073–1079.

Piper, S. E., Mitchell, J. E., Lee, D. J. & Busby, S. J. (2009). A global view of Escherichia coli Rsd protein and its interactions. Mol Bioys 5, 1943–1947.

Rokyta, D., Badgett, M. R., Molineux, I. J. & Bull, J. J. (2002). Experimental genomic evolution: extensive compensation for loss of DNA ligase activity in a virus. Mol Biol Evol 19, 230–238.

Rutherford, S. T., Villers, C. L., Lee, J. H., Ross, W. & Gourse, R. L. (2009). Allosteric control of Escherichia coli rRNA promoter complexes by DksA. Genes Dev 23, 236–248.

Saecker, R. M., Record, M. T., Jr & Dehaseth, P. L. (2011). Mechanism of bacterial transcription initiation: RNA polymerase – promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. J Mol Biol 412, 754–771.

Savalia, D., Robins, W., Nechaev, S., Molineux, I. & Severinov, K. (2010). The role of the T7 Gp2 inhibitor of host RNA polymerase in phage development. J Mol Biol 402, 118–126.

Schmitt, M. P., Beek, P. J., Kearney, C. A., Spence, J. L., DiGiovanni, D., Condreay, J. P. & Molineux, I. J. (1987). Sequence of a conditionally essential region of bacteriophage T3, including the primary origin of DNA replication. J Mol Biol 193, 479–495.

Sevostyanova, A., Belogurov, G. A., Mooney, R. A., Landick, R. & Artsimovitch, I. (2011). The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. Mol Cell 43, 253–262.

Edited by: D. Grainger