

**Characterisation of
bacteriophage-encoded
inhibitors of the bacterial
RNA polymerase**

Carol Maria Sheppard

SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ABSTRACT

RNA polymerase (RNAP) is an essential enzyme which catalyses transcription; a highly regulated process. Bacteriophage are viruses which infect bacteria and as a result have evolved a diverse range of mechanisms to regulate the bacterial RNAP to serve the needs of the virus. T7 Gp2 and Xp10 P7 are two bacteriophage-encoded transcription factors that inhibit the activity of the bacterial RNAP. The aim of this study is to investigate the molecular mechanisms of action of Gp2 and P7.

Fluorescence anisotropy experiments proved Gp2 to bind to RNAP, independently of the σ -factor, with a 1:1 stoichiometry and a low nanomolar affinity. *In vitro* transcription assays demonstrated that a negatively charged strip in Gp2 is the major determinant for its inhibitory activity. Furthermore, it was shown that efficient Gp2-mediated inhibition of RNAP also depends upon the highly negatively charged and flexible σ^{70} specific domain, R1.1. Gp2 and R1.1 both bind in the downstream-DNA binding channel and exert long-range antagonistic effects on RNAP-promoter DNA interactions around the transcription start site.

A systematic mutagenesis screen was used to identify residues in P7 necessary for binding to the RNAP; results were interpreted in the context of a newly resolved NMR structure of P7. Electrophoretic mobility shift assays revealed that P7 'traps' a RNAP-promoter DNA complex *en route* to the transcriptionally-competent complex. Preliminary results from a fluorescence based RNAP-DNA interaction assay suggest that P7 may target RNAP interactions with the -35 promoter element and the 'discriminator region'.

This study has contributed to our understanding of how non-bacterial transcriptional factors can influence bacterial gene expression by modulating RNAP activity. This study has also uncovered vulnerabilities in RNAP, which have the potential to be exploited therapeutically. To this end, these structure-function studies of Gp2 and P7 have provided the basis for the rational design of novel anti-bacterial compounds.

DECLARATION OF AUTHORSHIP

I certify that this thesis entitled “Characterisation of bacteriophage-encoded inhibitors of the bacterial RNA polymerase” is written entirely by myself and that the research to which it refers to is my own. I confirm that all main sources of help have been acknowledged and that any ideas or quotations from the work of others have been clearly referenced.

Carol Maria Sheppard

ACKNOWLEDGMENTS

I am grateful to so many people who have made the time during my PhD such an enjoyable and unforgettable experience.

Above all, I would like to thank my supervisor Ramesh who has given me constant guidance, advice and motivation during my PhD. I also appreciate the many opportunities he has provided for me and the wealth of knowledge and skills he has passed onto me, which I'm sure have set me in good stead for my future scientific career. I will miss the banter, verbal abuse and general rudeness to which I am now accustomed. I could not have wished for a better supervisor, thank you for everything!

To my fellow 'Wigs', both past and present, I would not have learnt half of what I know now without you all. A massive thanks to Jon, Bea, Trouble (Zhen), Sophie, Ellen, (An)drey and Dan who together have created such a fun and stimulating environment. You have helped me in so many countless ways; I look forward to our future 'collaborations'. To my other friends and colleagues in CMBI3, thank you for all the chats and laughs and the kindness you have shown me over the past four years. I will miss Friday Cakes and Thirsty Thursdays which have quite literally fuelled me through the bad times and helped me celebrate the good times.

I am, as ever, grateful to my parents who have always been encouraging and supportive towards my education. They have taught me to be ambitious, hard-working and more than anything to be appreciative and curious of the world around me.

Last but in no means least, I would like to thank my fiancé, Alex. Without your love and understanding I could not have done this. Thank you for being 'Roo' and always being there for me.

CONTENTS

ABSTRACT.....	1
DECLARATION OF AUTHORSHIP.....	2
ACKNOWLEDGMENTS.....	3
LIST OF FIGURES.....	9
LIST OF TABLES.....	11
CHAPTER 1	15
INTRODUCTION.....	15
1.1. Bacterial transcription.....	16
1.1.1. Bacterial transcription overview	16
1.2. Structure-function of bacterial RNAP	18
1.2.1. Bacterial RNAP structures	18
1.2.2. Architecture and channels of RNAP	19
1.2.3. Subunits of <i>E. coli</i> RNAP	20
1.2.4. RNAP ‘mobile modules’	22
1.2.5. σ -factors	26
1.2.6. Organisation and function of σ^{70} -like factor domains	27
1.3. The isomerisation from RPc to RPo	30
1.4. Regulation of transcription initiation.....	33
1.5. Bacteriophage.....	37
1.6. Phage-encoded transcription factors of bacterial RNAP	38
1.7. T7 phage-encoded Gp2.....	40
1.8. Xp10 phage-encoded P7	44
1.9. Research objectives	45
1.10. Why is this research important?	46
CHAPTER 2	48
MATERIALS AND METHODS	48

2.1.	Solutions and buffers	49
2.2.	Bacterial strains and culture conditions	51
2.2.1.	Media preparation	51
2.2.2.	Bacterial strains.....	52
2.2.3.	Bacterial plasmids	53
2.3.	Molecular cloning	54
2.3.1.	CaCl ₂ competent cell preparation	54
2.3.2.	Transformation by heat shock	54
2.3.3.	Purification of plasmid DNA.....	54
2.3.4.	Polymerase chain reaction (PCR).....	55
2.3.5.	Restriction digest	55
2.3.6.	Agarose gel electrophoresis.....	56
2.3.7.	Gel extraction and purification	56
2.3.8.	Phenol-chloroform extraction and ethanol precipitation.....	56
2.3.9.	Ligation.....	57
2.3.10.	Colony PCR	57
2.3.11.	Site directed mutagenesis (SDM).....	58
2.3.12.	Sequencing.....	59
2.4.	Protein expression and purification.....	60
2.4.1.	pET system of expression.....	60
2.4.2.	Temperature sensitive system of expression	60
2.4.3.	SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)	60
2.4.4.	Nickel affinity purification (FPLC).....	61
2.4.5.	High-throughput nickel affinity purification	62
2.4.6.	Heparin affinity purification (FPLC).....	62
2.4.7.	Polyethyleneimine and ammonium sulphate precipitation	63
2.4.8.	Bradford assay.....	63
2.4.9.	Activator dependent transcriptional bacterial two-hybrid β -galactosidase assay.	64

2.5.	Immunoblotting	64
2.5.1.	Polyclonal antibody production	64
2.5.2.	Affinity purification of antibody serum.....	65
2.5.3.	Antibody purification against cell background	65
2.5.4.	Western blotting	65
2.6.	<i>In vitro</i> biochemical assays.....	66
2.6.1.	³² P-labelling of DNA probes	66
2.6.2.	Electrophoretic mobility shift assay (EMSA)	66
2.6.3.	³² P-labelling of proteins	67
2.6.4.	Protein-protein native gel binding assay	67
2.6.5.	Abortive transcription assay	67
2.7.	Fluorescent binding assays	68
2.7.1.	Alexa-488 labelling of Gp2	68
2.7.2.	Fluorescence anisotropy	68
2.7.3.	TMR labelling of σ^{70}	69
2.7.4.	RNAP beacon assay	69
CHAPTER 3		70
ESTABLISHING A SOLUTION-BASED ASSAY TO DETECT THE BINDING OF GP2 TO <i>E. COLI</i> RNAP		70
3.1.	Introduction	71
3.2.	Objectives.....	75
3.3.	Over-expression and purification of recombinant proteins	76
3.4.	Fluorescent labelling of Gp2	79
3.5.	FA experimental setup.....	80
3.6.	Effect of the σ -factor on Gp2 binding to E.....	81
3.7.	Important determinants for Gp2-E/E σ binding	82
3.8.	Conclusions	84
CHAPTER 4		86

INHIBITION OF <i>E. COLI</i> RNAP BY GP2: THE ROLE OF THE NEGATIVELY CHARGED STRIP OF AMINO ACID RESIDUES IN GP2.....	86
4.1. Introduction	87
4.2. Objectives.....	89
4.3. Production of Gp2 NCS mutant proteins	90
4.4. 1-D NMR spectra of Gp2 NCS mutant proteins	91
4.5. Binding of NCS mutants to σ^{70}	92
4.6. Inhibition of transcription by NCS mutant proteins.....	94
4.7. Is Gp2 ^{Mut7} retained in the RPo?	95
4.8. Conclusions	97
CHAPTER 5	103
THE ROLE OF σ^{70} REGION 1.1 DOMAIN DURING GP2-MEDIATED INHIBITION OF RNAP	103
5.1. Introduction	104
5.2. Objectives.....	106
5.3. The effect of the σ^{70} N-terminal on Gp2 inhibitory activity.....	107
5.4. The contribution of σ^{70} R1.1 domain to inhibition of σ^{70} by Gp2.....	109
5.5. The effect of Gp2 on promoter-less transcription	114
5.6. The effect of Gp2 and R1.1 on DNA entry into the active site	116
5.7. The effect of Gp2 on accessibility to single-stranded DNA.....	119
5.8. Conclusions	121
CHAPTER 6	125
A PRELIMINARY CHARACTERISATION OF THE XP10 PHAGE-ENCODED, RNAP INHIBITOR, P7	125
6.1. Introduction	126
6.2. Objectives.....	127
6.3. Determination of P7 residues important for binding.	128
6.4. Structural determination of P7	134
6.5. Mechanism of inhibition of transcription initiation by P7	138
6.6. Discussion.....	150

CHAPTER 7	153
OVERVIEW.....	153
7.1. Work leading to this study	154
7.2. Advances made	156
7.3. Future perspectives	159
REFERENCES.....	161

- Sheppard,C., Camara,B., Shadrin,A., Akulenko,N., Liu,M., Baldwin,G., Severinov,K., Cota,E., Matthews,S., and Wigneshweraraj,S. (2011) Inhibition of *Escherichia coli* RNAP by T7 Gp2 protein: role of negatively charged strip of amino acid residues in Gp2. *J Mol Biol* **407**: 623-632.
- Mekler,V., Minakhin,L., Sheppard,C., Wigneshweraraj,S., and Severinov,K. (2011) Molecular mechanism of transcription inhibition by phage T7 Gp2 protein. *J Mol Biol* **413**: 1016-1027.
- James,E., Liu,M., Sheppard,C., Mekler,V., Camara,B., Liu,B., Simpson,P., Cota,E., Severinov,K., Matthews,S., and Wigneshweraraj,S. (2012) Structural and mechanistic basis for the inhibition of *Escherichia coli* RNA polymerase by T7 Gp2. *Mol Cell*. **47**: 755-66.
- Shadrin,A., Sheppard,C., Severinov,K., Matthews,S., and Wigneshweraraj,S. (2012) 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. *Microbiology* **158**: 2753-64.
- Drennan,A., Kraemer,M., Capp,M., Gries,T., Ruff,E., Sheppard,C., Wigneshweraraj,S., Artsimovitch,I., and Record,M.T.Jr. (2012) Key roles of the downstream mobile jaw of *Escherichia coli* RNA Polymerase in transcription initiation. *Biochemistry* **51**: 9447-59.
- Shadrin,A., Sheppard,C., Savalia,D., Severinov,K., Wigneshweraraj,S. (2012) Overexpression of *Escherichia coli* udk mimics the absence of T7 Gp2 function and thereby abrogates successful infection by T7 phage. *Microbiology*.

LIST OF FIGURES

Figure 1.1.1 Phosphodiester bond formation during nucleotidyl transfer.....	17
Figure 1.2.1 Structure of RNAP.	19
Figure 1.2.2 RNAP channels.	20
Figure 1.2.3 RNAP β and β' sequence insertion elements.....	22
Figure 1.2.4 RNAP mobile modules.	25
Figure 1.2.5 σ^{70} consensus promoter sequences.....	26
Figure 1.2.6 The RNAP holoenzyme.....	28
Figure 1.2.7 Conserved domains and functions of σ^{70} type factors.....	30
Figure 1.3.1 Transition between R _{Pc} and R _{Po}	32
Figure 1.4.1 Mechanisms of simple activation.	34
Figure 1.4.2 Mechanisms of simple repression.	36
Figure 1.7.1 Alanine mutagenesis screen of Gp2.....	42
Figure 1.7.2 Charge separation of Gp2 structure.	43
Figure 1.9.1 Gp2 and P7 inhibit R _{Po} formation.	46
Figure 3.1.1 Schematic of the relationship between binding and anisotropy.....	72
Figure 3.1.2 A diagram displaying the orientation of the emitted and excited polarized light.....	73
Figure 3.1.3 Sequence alignment of Gp2 orthologs.	74
Figure 3.3.1 Nickel affinity chromatography of E ^{WT}	77
Figure 3.3.2 Heparin affinity chromatography of E ^{WT}	77
Figure 3.3.3 Nickel affinity chromatography of σ^{70}	78
Figure 3.3.4 Heparin affinity chromatography of σ^{70}	78
Figure 3.3.5 Nickel affinity chromatography of Gp2.....	79
Figure 3.3.6 Lane profile of an SDS-PAGE gel from the nickel affinity chromatography of Gp2.	79
Figure 3.4.1 Native gel binding assay demonstrating the binding of Gp2/Gp2 ^{Alexa488} to E σ^{70}	80
Figure 3.6.1 Graph displaying the binding of Gp2 to E, E σ^{70} and E σ^{54} as measured by FA.....	82
Figure 3.7.1 Graph displaying binding of Gp2 to E σ^{70} and E ^{E118K} σ^{70} as measured by FA.	83
Figure 3.7.2 Graph displaying binding of Gp2 to E σ^{70} and Gp2 ^{R56E} to E σ^{70} as measured by FA.	83
Figure 4.1.1 Charge separation of Gp2 protein.	87
Figure 4.1.2 Sequence alignment of Gp2 orthologs.	88
Figure 4.1.3 Bar chart displaying the Inhibitory activity of Gp2 NCS point mutants.	89
Figure 4.3.1 Structure of Gp2 with NCS mutations highlighted.....	90
Figure 4.4.1 1-D NMR spectra of Gp2 and NCS mutants.	92

Figure 4.5.1 Graph displaying the binding of Gp2 ^{WT} * and NCS mutant versions to Eσ ⁷⁰ measured by FA.	93
Figure 4.6.1 Inhibitory activity of Gp2 NCS mutants.....	95
Figure 4.7.1 Native gel of Gp2 ^{WT} and Gp2 ^{Mut7} binding to Eσ ⁷⁰	96
Figure 4.8.1 Structure of the Gp2-β' jaw fragment with binding determinants.	99
Figure 4.8.2 Structure of Gp2-β' jaw fragment with NCS.	100
Figure 4.8.3 Structural composite model of <i>E. coli</i> RNAP core and Gp2.	101
Figure 5.1.1 Conserved domains and functions of σ ⁷⁰ type factors.....	104
Figure 5.3.1 Graph displaying the binding of Gp2* to Eσ ⁷⁰ and Eσ ^{70ΔR1} as measured by FA.	107
Figure 5.3.2 Inhibitory activity of Gp2 on Eσ ⁷⁰ and Eσ ^{70ΔR1} transcription.....	108
Figure 5.4.1 Inhibitory activity of Gp2 on Eσ ^{70WT} and Eσ ^{70ΔR1.1} transcription.	109
Figure 5.4.2 The <i>in trans</i> effect of σ ^{70R1.1} on Gp2-mediated inhibition of Eσ ^{70ΔR1.1} transcription.	110
Figure 5.4.3 The effect of the order of addition of <i>in trans</i> R1.1 on Gp2-mediated inhibition of transcription.....	111
Figure 5.4.4 Inhibitory activity of Gp2 on Eσ ^{70WT} , Eσ ^{70ΔR1.1} and Eσ ³⁸ transcription.....	112
Figure 5.4.5 Graph displaying the binding of Gp2* to variant holoenzymes as measured by FA.	113
Figure 5.5.1 Sequence of the MS.	114
Figure 5.5.2 Inhibitory effect of Gp2 on Eσ ^{70WT} and Eσ ^{70ΔR1.1} transcription from the MS.	115
Figure 5.5.3 The <i>in trans</i> effect of R1.1 on Gp2-mediated inhibition of Eσ ^{70ΔR1.1} transcription.....	116
Figure 5.6.1 The effect of Gp2 on <i>lacUV5</i> binding to Eσ ⁷⁰	117
Figure 5.6.2 EMSA of ³² P- <i>lacUV5</i> binding to Eσ ⁷⁰ in the presence of Gp2 ^{WT} , Gp2 ^{R56E} and Gp2 ^{Mut7}	118
Figure 5.6.3 EMSA of ³² P- <i>lacUV5</i> binding to Eσ ⁷⁰ and Eσ ^{70ΔR1.1} in the presence of Gp2 and/or R1.1..	119
Figure 5.7.1 Effect of Gp2 and R1.1 on binding and productive engagement of ss-DNA at the active site.....	120
Figure 5.8.1 A model of the effect of Gp2 on R1.1 positioning in Eσ ⁷⁰	124
Figure 6.1.1 Structural context of P7 binding.	127
Figure 6.3.1 Activator dependent transcriptional bacterial two-hybrid system.	130
Figure 6.3.2 Sequence alignment of P7 orthologs.	132
Figure 6.3.3 Bacterial two-hybrid β-galactosidase screen of P7.....	133
Figure 6.4.1 Secondary structure prediction of P7.	134
Figure 6.4.2 Nickel affinity chromatography of P7.	135
Figure 6.4.3 NMR structure of P7.	136
Figure 6.4.4 NMR structure of P7 with the secondary structure prediction overlay.	137
Figure 6.4.5 NMR structure of P7 with mapping of the class IV amino acid residues.	138

Figure 6.5.1 The order of addition effect of P7 on transcription initiation.	140
Figure 6.5.2 Effect of P7 on upstream $^{32}\text{P-lacUV5}$ binding to $\text{E}^{\text{WT}}/\text{NN}\sigma^{70}$	141
Figure 6.5.3 Effect of P7 on $^{32}\text{P-lacUV5}$ binding to $\text{E}^{\text{WT}}/\text{NN}\sigma^{70}$	142
Figure 6.5.4 The order of addition effect of P7 on $^{32}\text{P-lacUV5}$ binding to $\text{E}^{\text{NN}}\sigma^{70}$	143
Figure 6.5.5 RNAP beacon assay.	144
Figure 6.5.6 The effect of P7 on $\text{E}^{\text{NN}}\sigma^{70}$ interactions with the full length T5N25 promoter.	145
Figure 6.5.7 The effect of P7 on $\text{E}^{\text{NN}}\sigma^{70}$ interactions with the T5N25 -10 promoter element.	146
Figure 6.5.8 The effect of P7 on $\text{E}^{\text{NN}}\sigma^{70}$ interactions with the T5N25 discriminator region.	147
Figure 6.5.9 The effect of P7 on $\text{E}^{\text{NN}}\sigma^{70}$ interactions with the T5N25 extended -10 promoter element.	148
Figure 6.5.10 The effect of P7 on $\text{E}^{\text{NN}}\sigma^{70}$ interactions with the T5N25 promoter -35 promoter.	149
Figure 6.5.11 Summary of potential $\text{E}^{\text{NN}}\sigma^{70}$ -promoter DNA interactions targeted by P7.	149
Figure 7.1.1 How does Gp2 inhibit RPo formation?	155
Figure 7.1.2 How does P7 inhibit RPo formation?	156
Figure 7.2.1 Overall mechanism of inhibition of $\text{E}\sigma^{70}$ by Gp2.	158
Figure 7.2.2 Insights into the mechanism of inhibition of $\text{E}^{\text{NN}}\sigma^{70}$ by P7.	159

LIST OF TABLES

Table 1.6.1 Known phage-encoded proteins that bind to and modulate bacterial RNAP.	39
Table 2.2.1 Antibiotics used in this study.	51
Table 2.2.2 Bacterial strains used in this study.	52
Table 2.2.3 Bacterial plasmids used in this study.	53
Table 2.4.1 Composition of SDS polyacrylamide gels.	61
Table 2.6.1 Composition of a native 4.5% polyacrylamide gel.	67
Table 2.6.2 Composition of a urea-denaturing 20% polyacrylamide gel.	68
Table 4.5.1 K_d values of NCS mutants.	94
Table 5.4.1 Affinity of Gp2* binding to variants of E.	114
Table 6.3.1 Classification of P7 mutant proteins.	131

ABBREVIATIONS

1-D	1 dimensional
3-D	3 dimensional
³²P	Radioactive isotope of phosphorous
A	Adenine
α	Alpha subunit of RNAP (encoded by <i>rpoA</i>)
α-CTD	α C-terminal domain
α-NTD	α N-terminal domain
Å	Angstroms
ADP	Adenosine diphosphate
a.a.	Amino acid
Amp	Ampicillin
ATP	Adenosine triphosphate
β	Beta subunit of RNAP (encoded by <i>rpoB</i>)
β'	Beta prime subunit of RNAP (encoded by <i>rpoC</i>)
bp	Base pair
BSA	Bovine serum albumin
C	Cytosine
Ci	Curie
cAMP	Cyclic adenosine monophosphate
Cm	Chloramphenicol
CRP	cAMP receptor protein
Δ	Delta
Da	Daltons
dH₂O	Distilled water
ddH₂O	Double distilled water
DMF	N,N-dimethylformamide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
ds-DNA	Double stranded DNA
DTT	Dithiothreitol
dw-DBC	Downstream-DNA binding channel
E	RNAP core
<i>E. coli</i>	<i>Escherichia coli</i>
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
Eσ	RNAP holoenzyme
FA	Fluorescence anisotropy
FPLC	Fast protein liquid chromatography
G	Guanine
g	Grams
<i>g</i>	Relative centrifugal force
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
Kan	Kanamycin
Kb	Kilo base
KDa	Kilo Daltons
L	Litres
LA	Lysogeny agar

LB	Lysogeny broth
M	Molar
MCS	Multiple cloning site
mg	Milligrams
ml	Millilitres
mM	Millimolar
mRNA	Messenger RNA
MU	Miller units
MW	Molecular weight
NCS	Negatively charged strip
ng	Nanograms
nM	Nanomolar
nm	Nanometre
NT	Nucleotide
NTP	Nucleoside triphosphate
ω	Omega subunit of RNAP (encoded by <i>rpoZ</i>)
OD	Optical density
ONPG	Ortho-nitrophenyl-β-D-galactopyranoside
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEI	Polyethyleneimine
PFU	<i>Pyrococcus furiosus</i> (DNA polymerase)
PKA	Protein kinase A
PNK	Polynucleotide kinase
PVDF	Polyvinylidene difluoride
ρ	Rho protein
RBS	Ribosome binding site
RNA	Ribonucleic acid
RNAP	RNA polymerase
RPc	RNAP promoter DNA closed complex
RPi	RNAP promoter DNA intermediate complex
rpm	Revolutions per minute
RPO	RNAP promoter DNA open complex
rRNA	Ribosomal RNA
RT	Room temperature
SAP	Shrimp alkaline phosphatase
SDM	Site directed mutagenesis
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SE	Standard error
ss-DNA	Single stranded DNA
σ	Sigma subunit of RNAP
σ¹⁹	Sigma subunit of RNAP (encoded by <i>fecI</i>)
σ²⁴	Sigma subunit of RNAP (encoded by <i>rpoE</i>)
σ²⁸	Sigma subunit of RNAP (encoded by <i>rpoF</i>)
σ³²	Sigma subunit of RNAP (encoded by <i>rpoH</i>)
σ³⁸	Sigma subunit of RNAP (encoded by <i>rpoS</i>)
σ⁵⁴	Sigma subunit of RNAP (encoded by <i>rpoN</i>)
σ⁷⁰	Sigma subunit of RNAP (encoded by <i>rpoD</i>)
T	Thymine

<i>T. aquaticus</i>	<i>Thermus aquaticus</i>
TBE	Tris borate EDTA
TBS	Tris buffered saline
TBSTT	Tris buffered saline tween 20
TCEP	Tris-(2-carboxyethyl)phosphine
TEC	Transcription elongation complex
TEMED	N,N,N',N' -tetramethyl-ethylene-diamine
TF	Transcription factor
TFBS	Transcription factor binding site
TG	Tris glycine
TGS	Tris glycine sodium dodecyl sulphate
Tm	Melting temperature
TMR	Tetramethylrhodamine-5-maleimide
tRNA	Transfer RNA
<i>T. thermophilus</i>	<i>Thermus thermophilus</i>
U	Uridine
UP-elements	Upstream elements
UTP	Uridine triphosphate
µg	Micrograms
µl	Microlitres
µM	Micromolar
V	Volts
v/v	Volume per volume
w/v	Weight per volume
WT	Wild type
λ	Wavelength
V	Volts
v/v	Volume per volume
<i>X. oryzae</i>	<i>Xanthomonas oryzae</i>

Amino acid	Three letter codon	One letter codon
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid (Aspartate)	Asp	D
Cysteine	Cys	C
Glutamic acid (Glutamate)	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

CHAPTER 1

INTRODUCTION

1.1. Bacterial transcription

1.1.1. Bacterial transcription overview

In bacteria, RNAP is the key enzyme that catalyses transcription, the first and most highly regulated step of gene expression. RNAP facilitates the process of transcription by which DNA is used as a template for the *de novo* synthesis of messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA) from nucleoside triphosphate (NTP) substrates. A mRNA transcript can then be subsequently translated into a polypeptide chain. Transcription can be divided into three distinct but successive stages; initiation, elongation and termination. To initiate transcription the RNAP associates with a σ -factor to form a complex, which is able to specifically interact with the promoter, a region upstream of the gene. The initial engagement of RNAP- σ with the promoter DNA is referred to as the RNAP-promoter DNA closed complex (RPc), in reference to the conformation of the DNA. The RNAP then unwinds a section of the DNA to form the 'transcription bubble' and positions the transcription start site near the catalytic centre. This transcriptionally competent state is referred to as the RNAP-promoter DNA open complex (RPo). The separation of the DNA strands forms a region of single-stranded DNA that serves as a template for RNA synthesis. The incoming NTPs are then able to base pair to their now exposed deoxyribonucleotide counterparts. In the RNAP catalytic centre, two Mg^{2+} ions are required for coordinating the formation of the phosphodiester bond and consequently the release of a pyrophosphate side product. One Mg^{2+} ion is permanently chelated to three absolutely conserved aspartic acid residues within a NADFDGD motif (Zaychikov *et al.*, 1996) of the RNAP, whereas the other is associated with the β and γ phosphates of the entering NTP and is recruited for each act of catalysis. The arrangement of these ions facilitates the nucleophilic attack of the entering NTP α -phosphate by the 3' hydroxyl at the terminus of the nascent RNA chain and stabilises the pentavalent transition state of the target phosphate (Sosunov *et al.*, 2003; Sosunov *et al.*, 2005).

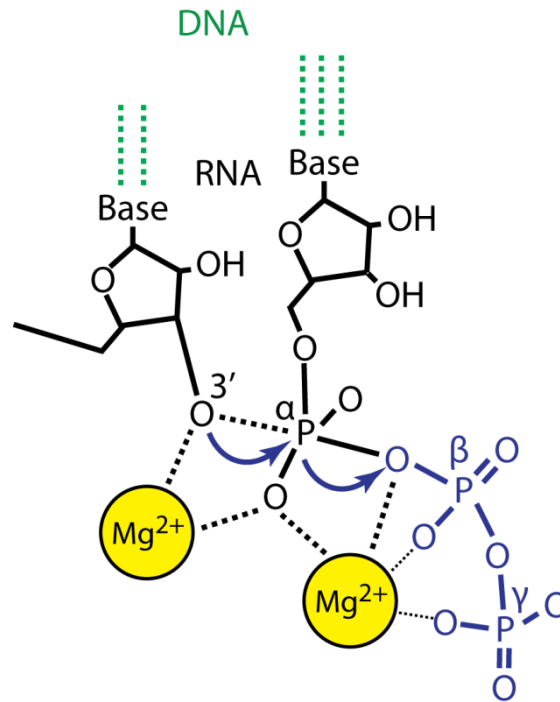


Figure 1.1.1 Phosphodiester bond formation during nucleotidyl transfer.

A schematic showing the nucleophilic attack of the α -phosphate of the incoming NTP by the 3' hydroxyl at the terminus of the nascent RNA chain. The pentavalent transition state of the α -phosphate is coordinated by two Mg^{2+} ions (yellow circles). One of the Mg^{2+} ions is associated with the β and γ -phosphates of the incoming NTP. Phosphodiester bond formation results in the release of a pyrophosphate product (blue). Inspired by (Zhang *et al.*, 2010).

Initially, as the substrate NTPs bind to the growing DNA/RNA hybrid, the template DNA is pulled into the main channel of the RNAP. Meanwhile the promoter region of the DNA remains bound to the σ -factor. The outcome of these counterintuitive events is 'DNA scrunching'. This results in an accumulation of free energy, which initially causes cycles of the premature breaking of the DNA/RNA hybrid and release of short abortive products. It has been hypothesised that these short RNA transcripts may serve to prime initiation and therefore regulate gene expression (Goldman *et al.*, 2009). Once the RNA chain reaches more than 8-11 nucleotides the free energy exceeds that of promoter DNA binding, allowing the RNAP to dissociate from the σ -factor and clear the promoter (Borukhov and Nudler, 2008). Promoter clearance triggers the dissociation of the σ -factor from the RNAP to form a stable transcription elongation complex (TEC). However, in some cases the dissociation of the σ -factor from the RNAP is not compulsory; the σ -factor can remain bound throughout the transcription cycle (Bar-Nahum and Nudler, 2001). In the TEC, the 'transcription bubble' and DNA/RNA hybrid are maintained as the RNAP productively transcribes the DNA from 5' – 3' (with respect to the non-template strand) at an average rate of 30-100 nt/s *in vivo* (Mooney *et*

al., 1998). The conservation of these nucleic acid structures means that the RNAP translocates the DNA template one nucleotide at a time to match the incorporation of individual substrate NTPs with the displacement of one nucleotide of the nascent RNA from the DNA/RNA hybrid. The growth of the RNA chain occurs in parallel with the re-annealing of the upstream duplex DNA and the DNA strand separation at the downstream edge of the 'transcription bubble' (Vassilyev *et al.*, 2007). Transcription can be terminated by two different mechanisms, which are both dependent on RNAP pausing followed by destabilisation of the TEC. Intrinsic termination occurs when the newly synthesised RNA forms a hairpin loop caused by intramolecular base pairing of a G-C rich region, this is subsequently followed by a poly-U tract. These structures cause the melting of the DNA/RNA hybrid and collapse of the 'transcription bubble' (Gusarov and Nudler, 1999). Alternatively, termination can also be induced by the mono-hexameric helicase protein ρ , which threads the nascent RNA through the centre of its ring and translocates towards the RNAP to unwind the DNA/RNA hybrid (Ciampi, 2006). After the release of the nascent RNA and the DNA template, the RNAP becomes available for σ -factor binding and new cycle of transcription.

1.2. Structure-function of bacterial RNAP

1.2.1. Bacterial RNAP structures

RNAP is a large, multi-subunit enzyme which is evolutionary conserved across all three kingdoms of life. Eukaryotic organisms possess at least three distinct nuclear RNAPs; RNA Pol I, RNA Pol II and RNA Pol III for synthesising rRNA, mRNA and tRNA respectively. Conversely, all prokaryotic (bacterial and archaeal) transcription is performed by a single RNAP, analogous to the eukaryotic RNA Polymerase II. The study of prokaryotic transcription, particularly in the bacterial model organism *Escherichia coli* has provided fundamental information on the structure and function of all RNAPs. Due to the large size (~390 kDa) and multi-subunit composition of bacterial RNAP a combination of biochemical, biophysical, structural and genetic approaches is needed to fully understand the mechanism and regulation of transcription. Although the *E. coli* RNAP is the best studied enzyme of its type to date, the only high resolution, x-ray crystallography structures of the RNAP belong to the thermophilic bacterium *Thermus thermophilus* (2.6 Å and 2.5 Å) (Vassilyev *et al.*, 2002; Vassilyev *et al.*, 2007) and *Thermus aquaticus* (3.3 Å and 4 Å) (Zhang *et al.*, 1999; Murakami *et al.*, 2002b) enzymes. However the bacterial RNAP proteins are functionally similar and share a high level of sequence homology. Furthermore, the overall shape and size of a low resolution (19 Å)

structure of the *E. coli* RNAP (Darst *et al.*, 1998) obtained by cryo-electron microscopy is reminiscent of the high resolution structures of *T. thermophilus* and *T. aquaticus* RNAP.

1.2.2. Architecture and channels of RNAP

The bacterial RNAP is a large enzyme, comprising of ~150 Å (long), 115 Å (tall) and 110 Å (wide) dimensions (Zhang *et al.*, 1999). RNAP has an overall negative charge but with an asymmetric charge distribution; the exterior is uniformly negative, whereas the surfaces that contact the nucleic acids are predominantly positive (Murakami *et al.*, 2002b). The architecture of the RNAP, like the eukaryotic (Darst *et al.*, 1991; Gnatt *et al.*, 2001) enzyme reveals a characteristic crab-claw like shape (Figure 1.2.1), in which the major catalytic subunits comprise the smaller and larger ‘pincers’ respectively.



Figure 1.2.1 Structure of RNAP.

A structure of *T. aquaticus* RNAP (Zhang *et al.*, 1999) shown in the ribbon representation. The overall architecture of the multi-subunit RNAP resembles that of a ‘crab claw’.

The ‘pincers’ of the claw form a positively charged (27 Å wide) channel referred to as the main channel or cleft which is often subdivided into two sections: the active site channel where catalysis occurs and the DNA/RNA hybrid resides and the downstream-DNA binding channel (dw-DBC) which holds the duplex DNA, still to be transcribed, downstream of the active site. The main channel branches at the upstream of the RNAP into the RNA exit channel and forms a path through which the nascent RNA transcript can exit the complex. A secondary channel (~10-12 Å wide) at the downstream face of the enzyme leads to the catalytic centre and is responsible for the delivery of substrate NTPs as well as providing access for specific transcriptional regulators (Zhang *et al.*, 1999). The arrangement of the RNAP channels is summarised in Figure 1.2.2.

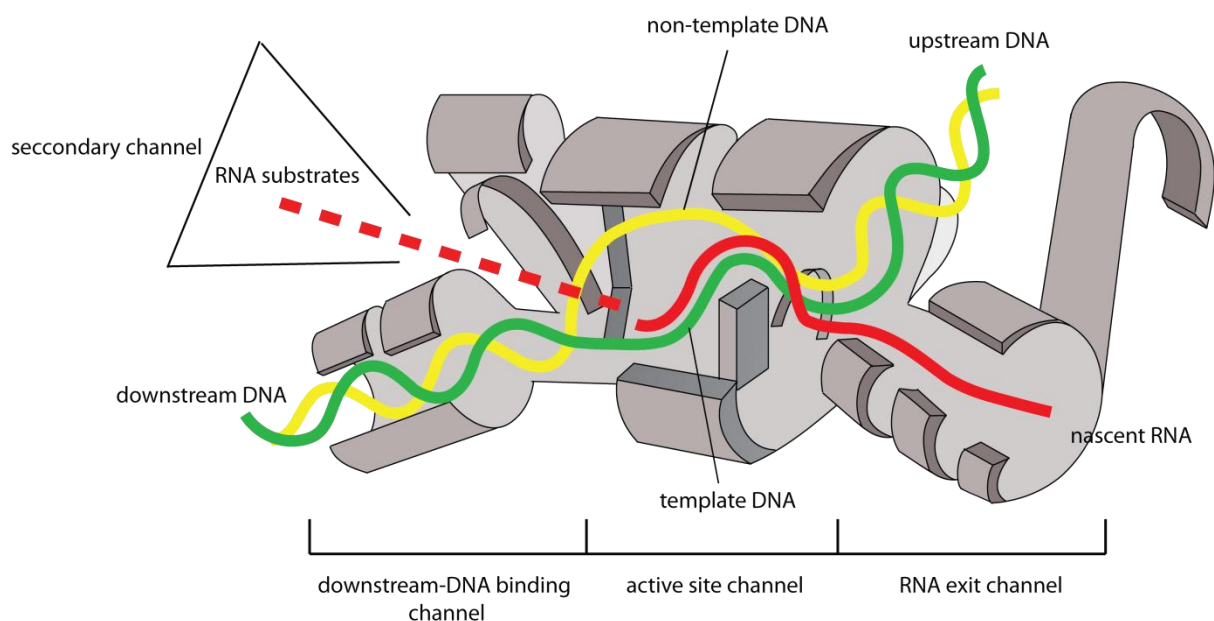


Figure 1.2.2 RNAP channels.

A schematic representation of the RNAP channels and their accommodation of nucleic acids during transcription. The RNA substrates (red dashed line) enter the active site via the secondary channel. The downstream duplex DNA is held in the downstream-DNA binding channel (non-template strand in yellow, template strand in green), catalysis of nucleotidyl transfer occurs in the active site channel and the nascent RNA (red line) exits the RNAP through the RNA exit channel. Inspired by (Nudler, 2009).

1.2.3. Subunits of *E. coli* RNAP

The *E. coli* RNAP core consists of five subunits; two α (36.5 kDa), β (150.6 kDa), β' (155.2 kDa) and ω (10.2 kDa) which are collectively referred to as E. The α -subunits consist of two domains connected by an unstructured flexible linker (Blatter *et al.*, 1994). The N-terminal domains (α -NTD)

of the α_I and α_{II} subunits contact the β and β' subunits respectively (Zhang *et al.*, 1999). The α -NTDs are responsible for RNAP subunit assembly, during which the two α subunits dimerise before association with the β , β' and ω subunits sequentially (Wang *et al.*, 1997; Ishihama, 1981). The C-terminal (α -CTD) domains are connected to the rest of the RNAP via the linkers which permit the differential positioning of the α -CTDs at different promoters. The α -CTDs form sequence specific interactions with regions of DNA or with DNA-bound transcriptional regulatory factors to influence transcriptional activation, repression, termination and anti-termination (Liu *et al.*, 1996; Browning and Busby, 2004). In addition, the length of the linkers has also been shown to effect transcription in a promoter-dependent manner (Meng *et al.*, 2000).

The ω -subunit is unique to bacteria and is also proposed to have a structural role in enzyme assembly, specifically acting as a chaperone to the β' -subunit and in the recruitment of the σ -factor (Mathew and Chatterji, 2006). More recently it has also been suggested that ω may influence promoter specificity by altering the affinities of the σ -factor for the RNAP (Geertz *et al.*, 2011). However its precise function during RNAP transcription remains unclear.

The larger β and β' subunits comprise the major catalytic components of the enzyme which form the 'pincers'. The β and β' subunits were historically subdivided into evolutionary conserved regions from prokaryotic, archaeal and eukaryotic sequences, classified A-I for β (Sweetser *et al.*, 1987; Zhang *et al.*, 1999) and A-H for β' (Jokerst *et al.*, 1989; Zhang *et al.*, 1999). More recently the conserved sequence regions have been updated and redefined; amongst bacterial RNAP, 16 shared regions for the β subunit and 20 for the β' subunit have been identified (Lane and Darst, 2010) (Figure 1.2.3). This comparison revealed that the highly conserved regions are also interspersed with species specific additions referred to as insertion elements.

Relative to *E. coli* the *Thermus* enzymes contain two insertions; β_{i12} (Opalka *et al.*, 2010) and β'_{i2} (β' non-conserved domain) (Chlenov *et al.*, 2005). Compared to the *Thermus* enzymes, *E. coli* RNAP contains four sequence insertion elements; β_{i4} (dispensable region I), β_{i9} (dispensable region II), β'_{i6} (G non-conserved domain) (Chlenov *et al.*, 2005) and β_{i11} (Lane and Darst, 2010) (Figure 1.2.3). These *E. coli* insertion elements have been structurally resolved and the *Thermus* specific regions removed to present a complete structural model of the *E. coli* RNAP (Opalka *et al.*, 2010).

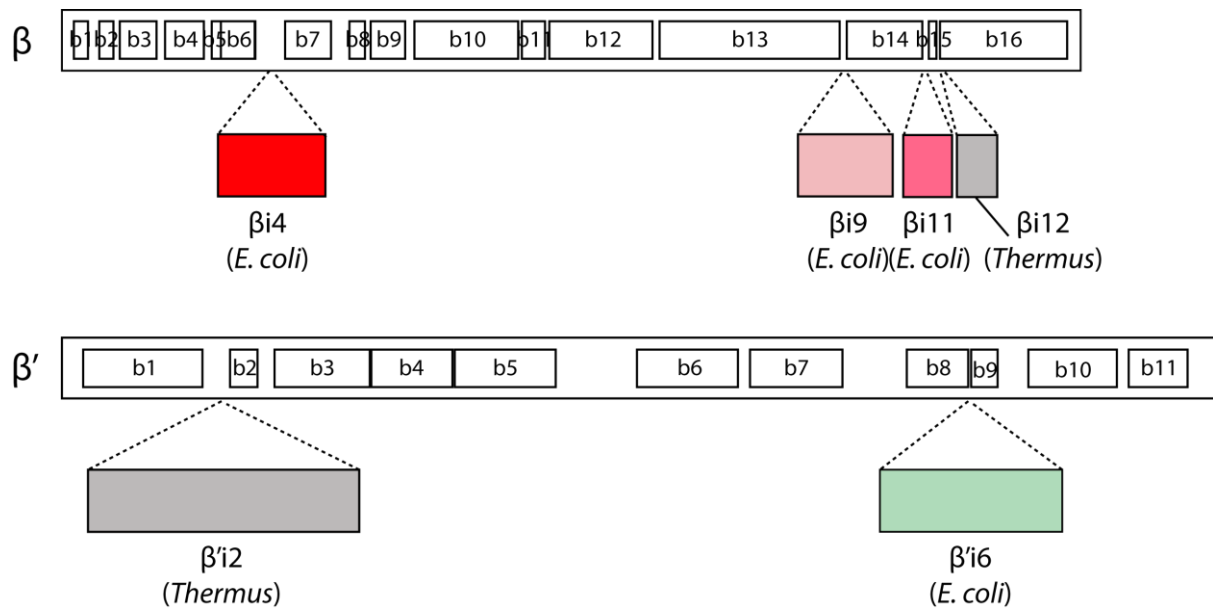


Figure 1.2.3 RNAP β and β' sequence insertion elements.

Representation of the evolutionary conserved regions of the bacterial β and β' subunits. The white boxes indicate conserved regions between all bacteria, whereas the coloured boxes indicate lineage-specific insertion elements of *E. coli* and the *Thermus* RNAPs. Adapted from (Opalka *et al.*, 2010).

1.2.4. RNAP ‘mobile modules’

A more useful and descriptive nomenclature is more commonly used to subdivide the β and β' catalytic domains into structurally conserved and highly flexible regions referred to as ‘mobile modules’. These ‘mobile modules’ possess varied roles in RNAP- σ assembly, nucleic acid interactions, enzyme movement (translocation), catalytic activity and transcription regulation. The β flap, β' Zn-binding element, β' zipper, β' lid and β' coiled-coil are located at the upstream face of the RNAP and primarily form the RNA exit channel. The β' rudder, β' bridge-helix and the β' trigger-loop are located in the active site channel and the β dw-lobe, β' clamp and the β' jaw form the dw-DBC (Figure 1.2.4).

The most substantial conformational changes of the RNAP are a result of the flexibility of the ‘pincers’ which enclose the main channel. The β' pincer is formed predominantly by the β' clamp (*Thermus* 3-619) (which also encompasses the β' zipper, β' Zn-binding element, β' lid, β' coiled-coil and β' rudder). Swinging motions of the β' clamp and the β clamp (*Thermus* 1058-1116) around a hinge referred to as the switch region, can alter the width of the main channel by nearly 25 Å (Darst *et al.*, 2002). The ‘pincers’ can exist in three conformational states; an open clamp state with a

channel width of ~ 20 Å which is sufficient for the accommodation of double stranded (ds)-DNA, a closed state of ~ 12 Å which can accommodate only single stranded (ss)-DNA and a collapsed state with a width of ~ 8 Å which is insufficient for even ss-DNA. The width of the main channel in the open state allows for DNA loading and strand-separation, whereas clamp closure stabilises the RPo and elongation complexes (Chakraborty *et al.*, 2012).

The downstream portion of the β' clamp together with the β dw-lobe (*Thermus* 142-330), β' jaw (*E. coli* 1149-1190) and $\beta'i6$ (*E. coli* 944-1129) contribute to the dw-DBC which accommodates the duplex DNA. The dw-DBC interactions with the dw-DNA can be subdivided into active site proximal and active site distal contacts. Active site proximal contacts with the +5 to +8 DNA is made by the β' clamp and β' dw-lobe. Whereas active site distal, +10 to +20 DNA, is also contacted by the β' clamp as well as the β' jaw and $\beta'i6$. The β' jaw domain contributes to all three stages of transcription by increasing the longevity of the RPo and facilitating pausing during elongation and termination (Ederth *et al.*, 2002a). In particular, the β' jaw stabilises DNA strand-separation together with the β' clamp and β dw-lobe (Nechaev *et al.*, 2000;Wigneshweraraj *et al.*, 2002;Wigneshweraraj *et al.*, 2003;Wigneshweraraj *et al.*, 2005;Wigneshweraraj *et al.*, 2006). In addition, $\beta'i6$ also plays a role in *E. coli* RNAP pausing and termination (Artsimovitch *et al.*, 2003), most likely because the sequence of this insertion element intercepts a critical mobile module known as the β' trigger loop (Opalka *et al.*, 2010).

The highly conserved β' trigger loop (*Thermus* 1238-1250) and β' bridge helix (*Thermus* 1067-1102) are both situated in the active centre and are positioned to separate the main channel from the secondary channel as well as marking the boundary between the active site channel and the dw-DBC (Zhang *et al.*, 1999;Vassilyev *et al.*, 2002;Cramer *et al.*, 2001). The flexible β' trigger loop together with the β' bridge helix function in concert to catalyse nucleotidyl transfer. The β' trigger loop, in its folded conformation is proposed to be directly responsible for the coupling of substrate NTP recognition with the catalysis of phosphodiester bond formation (Wang *et al.*, 2006). In addition the β' trigger loop may also influence the conformation of the β' bridge helix. The state of the β' bridge helix seems to differ in the available structures, consequently it is proposed to adopt at least two conformations, straight and kinked (Vassilyev *et al.*, 2002;Gnatt *et al.*, 2001). These conformations interchange depending on the stage of catalysis and correlate with the translocation of the DNA and RNA through the RNAP (Epshtein *et al.*, 2002;Bar-Nahum *et al.*, 2005).

Several mobile modules are required for the stability of the DNA/RNA hybrid and the processivity of the nascent RNA chain. The β flap (*Thermus* 703-830), β' zipper (*Thermus* 26-47), β' Zn-binding element (*Thermus* 51-83), and β' lid (*Thermus* 525-538), β' rudder (*Thermus* 584-602) and

β' coiled-coil (*Thermus* 540-581) all comprise or are located in close proximity to the RNA exit channel. The β' flap is a major component of the RNA exit channel, acting as a cover for the nascent RNA (Zhang *et al.*, 1999). The β' flap contributes to σ -factor binding and (-35) promoter element recognition and is likely to be required for σ -release during promoter escape (Vassilyev *et al.*, 2002;Kuznedelov *et al.*, 2002b). The flap is also responsible for the recognition and stabilisation of hairpin structures in the nascent RNA (Toulokhonov and Landick, 2003) together with the β' zipper and β' Zn-binding element (Epshtein *et al.*, 2007a;King *et al.*, 2004).

The β' Zn-binding element comprises 4 conserved cysteine residues which may play a critical role in β' folding (Zhang *et al.*, 1999). The positioning of the β' Zn-binding element suggests that in addition to termination it may also be involved in recognition of (-35) promoter DNA elements (Murakami *et al.*, 2002b;Vassilyev *et al.*, 2002). The β' zipper is involved in forming interactions with the promoter DNA (spacer region) to form a stable R_Pc and at some promoters it also facilitates R_Po formation (Yuzenkova *et al.*, 2011).

The β' lid forms contacts with the σ -factor (linker region) as is required for efficient transcription initiation (Toulokhonov and Landick, 2006). The β' lid stacks onto the upstream edge of the DNA/RNA hybrid, thus mimicking the base-stacking interactions within the nucleic acid duplex and sterically blocking its extension (Vassilyev *et al.*, 2007). Therefore, the β' lid is also well positioned to facilitate the dissociation of the RNA chain from the DNA template strand of the DNA/RNA hybrid (Naryshkina *et al.*, 2006;Toulokhonov and Landick, 2006).

The primary contact site of the σ -factor with RNAP is with the β' coiled-coil which extends out from the main channel (Murakami *et al.*, 2002b). The β' coiled-coil induces the σ -factor (R 2.2) to recognise the DNA non-template strand (-10 promoter element) (Young *et al.*, 2001). The β' coiled-coil also supports another loop-like structure called the β' rudder (Zhang *et al.*, 1999). The β' rudder stabilises the RNA at the upstream edge of the DNA/RNA hybrid in the initiation and elongation complexes (Kuznedelov *et al.*, 2002a).

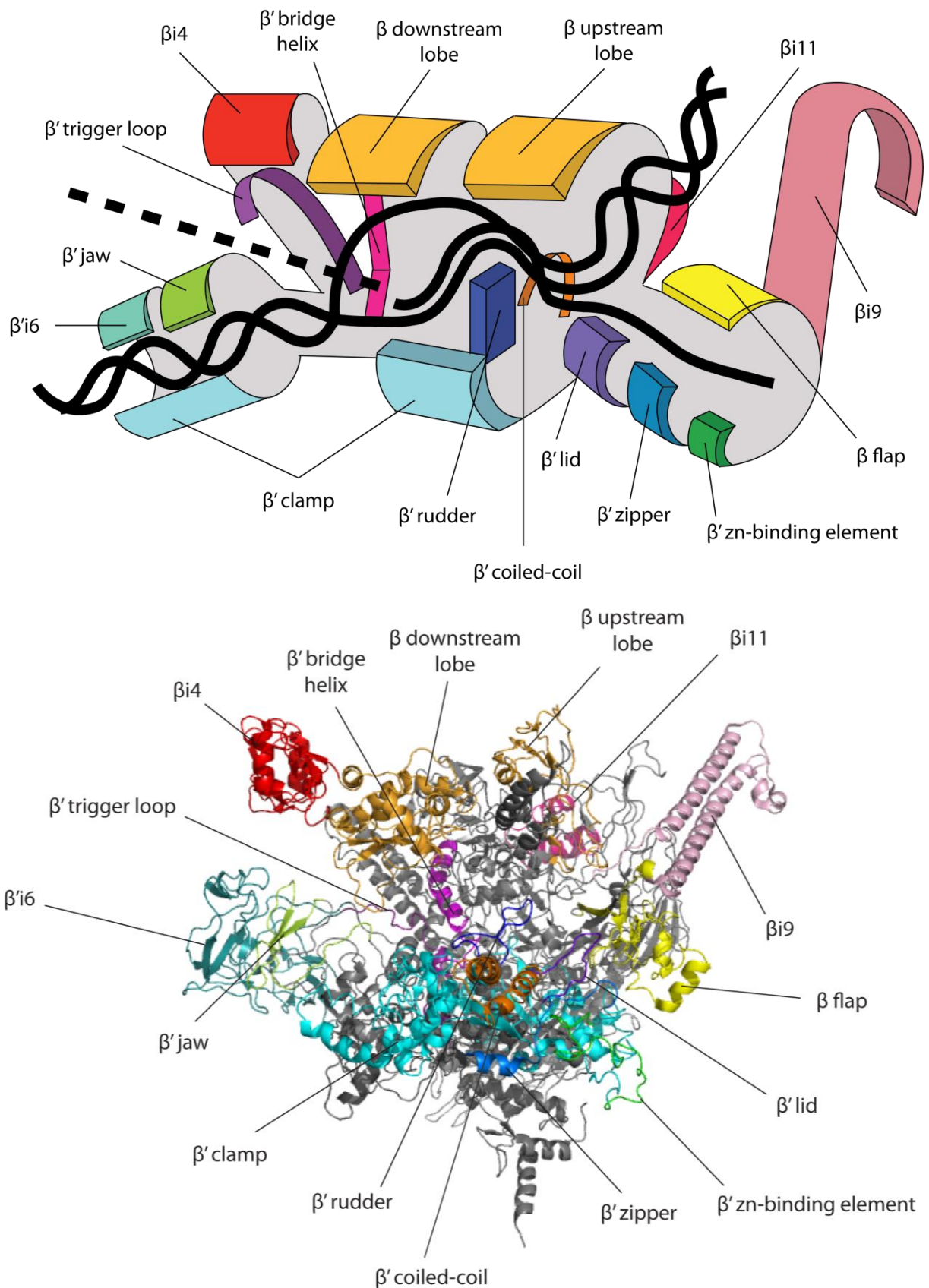


Figure 1.2.4 RNAP mobile modules.

Top: schematic of RNAP with the mobile modules highlighted. Inspired by (Nudler, 2009). Bottom: structural model of *E. coli* RNAP with mobile modules highlighted as above (Opalka *et al.*, 2010).

1.2.5. σ -factors

The core RNAP (E) is catalytically proficient but is unable to initiate promoter specific transcription. E forms extensive contacts with a dissociable sixth subunit, σ , to form a fully functional holoenzyme ($E\sigma$). Upon formation of the holoenzyme and only then, the σ -factor is able to recognise and interact with specific promoter elements, which contain conserved hexameric motifs. The number of σ -factors encoded in a bacterial genome varies considerably from 1 in *Mycoplasma genitalium* to 63 in *Streptomyces coelicolor*, and generally differs in accordance with the lifestyle of the species (Gruber and Gross, 2003). In response to environmental cues, bacterial cells are able to select from their repertoire of σ -factors in order to direct gene expression towards various adaptive responses. Each of these σ -factors accumulate in response to different stress cues and compete for binding to the free E, enabling transcription from a specific subset of promoters therefore altering the global pattern of gene expression.

E. coli has seven different species of σ -factor; σ^{70} (*rpoD*), σ^{54} (*rpoN*), σ^{38} (*rpoS*), σ^{32} (*rpoH*), σ^{28} (*rpoF*), σ^{24} (*rpoE*) and σ^{19} (*fecI*), named according to their respective molecular weights. Even though the σ -factors contact E in a similar manner they all confer differential regulatory properties upon the holoenzyme. The principal and prototypical *E. coli* σ -factor, σ^{70} is responsible for expression of housekeeping genes predominantly during exponential growth. Five of the other *E. coli* σ -factors (σ^{38} , σ^{32} , σ^{28} , σ^{24} and σ^{19}) are closely related to σ^{70} in sequence and structure and therefore belong to the σ^{70} -like family. σ -factors of this type recognise and initiate transcription from promoters containing the non-template consensus -35 ($^{-35}$ TTGACA $^{-30}$) and -10 ($^{-12}$ TATAAT $^{-7}$) sequences (Lisser and Margalit, 1993) (Figure 1.2.5). The length of the spacer region between the -10 and -35 promoter elements is optimally 17 bps (Harley and Reynolds, 1987) and is important for $E\sigma$ -promoter DNA interactions and promoter activity (Mulligan *et al.*, 1985). σ^{70} -like factors can also recognise a TG motif just upstream of the -10 element referred to as the extended -10 promoter element which is present in a subset of promoters that often lack the -35 element (Barne *et al.*, 1997) as well as residues in the discriminator region located between the -10 element and the transcription start site (+1) (Haugen *et al.*, 2006).

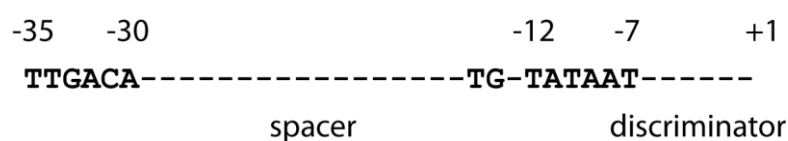


Figure 1.2.5 σ^{70} consensus promoter sequences.

Sequence and position of the non-template consensus promoter elements and regions recognised by σ^{70} .

The σ^{70} -like factors can be broadly divided into four phylogenetic groups (Gruber and Gross, 2003; Paget and Helmann, 2003). Group 1 includes the primary housekeeping σ -factor represented by *E. coli* σ^{70} . Group 2 σ -factors are closely related to group 1 but differ in the fact that they are dispensable for growth. *E. coli* σ^{38} is a group 2 σ -factor which drives the transcription of genes required during the stationary phase, making it the master regulator of the general stress response (Hengge-Aronis, 2002). Group 3 σ -factors are more distantly related and also tend to initiate transcription from specific promoters in response to more precise cues. For instance *E. coli* σ^{32} responds to aberrant protein folding usually as a result of heat shock (Arsene *et al.*, 2000) whereas σ^{28} is responsible for the transcription of late genes in flagellum assembly required for rapid motility (Chilcott and Hughes, 2000). Group 4 includes the most numerous and diverse σ -factors which are also termed ECF σ -factors due to their role in responding to signals from the extracytoplasmic environment. σ^{24} responds to accumulation of aberrant proteins in the cell envelope in order to restore the integrity of the cell membrane (Alba and Gross, 2004), whereas σ^{19} directs transcription of the operon responsible for iron uptake (Visca *et al.*, 2002).

Although most strains of bacteria have multiple σ^{70} -like representatives they tend to have only one major alternative type of σ -factor. The σ^{54} class of σ -factors are associated with the transcription of genes required for specific adaptive changes such as growth under nitrogen limiting conditions and changes to membrane integrity. σ^{54} is the single *E. coli* representative of the σ^{54} class of σ -factors. Unlike σ^{70} , σ^{54} characteristically directs E to promoters with conserved regions at -12 (GC) and -24 (GG) (Barrios *et al.*, 1999). Furthermore, $E\sigma^{54}$ requires an enhancer protein (AAA+ ATPase) to initiate this energy (ATP hydrolysis) dependent form of transcription (Wigneshweraraj *et al.*, 2008; Buck *et al.*, 2000).

1.2.6. Organisation and function of σ^{70} -like factor domains

σ^{70} type factors consist of up to four domains; 1.1, 2, 3 and 4 joined by flexible linkers (Campbell *et al.*, 2002). These domains contain the highly conserved regions (R) 1.1, 1.2-2.4, 3.0-3.1 and 4.1-4.2 as well as a large non-conserved region (σ_{NCR}) between regions 1.2 and 2.1. The interdomain distances in free σ are short, resulting in an overall compact structure (Mekler *et al.*, 2002; Schwartz *et al.*, 2008; Callaci *et al.*, 1999). In the holoenzyme each of the σ -domains and the linkers makes extensive contacts with the upstream face of E, (Murakami *et al.*, 2002b; Vassilyev *et al.*, 2002) whilst the crucial promoter binding elements, σ domains 2 and 4, are solvent exposed to engage with the DNA (Figure 1.2.6).

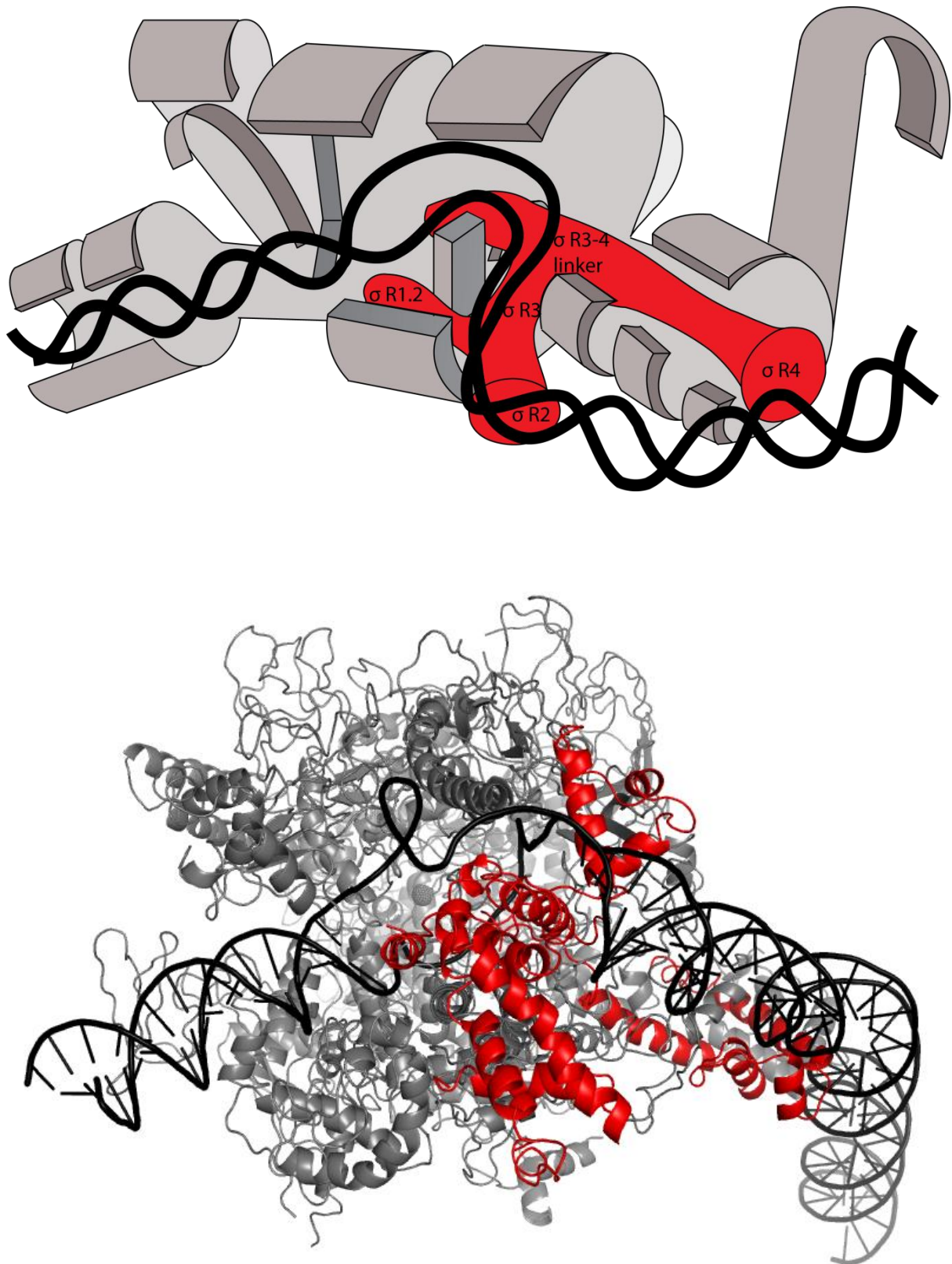


Figure 1.2.6 The RNAP holoenzyme

Top: schematic of RNAP holoenzyme in the RPo formation with the σ -factor coloured red. Inspired by (Nudler, 2009). Bottom: structural model of *T. aquaticus* RPo with the σ -factor highlighted in red (Opalka *et al.*, 2010).

R1.1 is limited to only group 1 σ^{70} -like factors, and although its sequence is poorly conserved its high negativity is always preserved. To date, there is limited structural information available for R1.1 in either the free σ or holoenzyme, with only the full length version of σ^A from *Thermotoga maritima* resolved (Schwartz *et al.*, 2008). Thus far, R1.1 has been assigned with two major functions. In the free σ , R1.1 is believed to function as an auto-inhibitory domain by positioning itself to putatively interact with R4.2 and mask the promoter binding elements (Dombroski *et al.*, 1993a; Dombroski *et al.*, 1993b; Camarero *et al.*, 2002; Schwartz *et al.*, 2008). Binding to the RNAP core induces large scale conformational changes in σ including the movement of R1.1 which relieves the masking and allows for spatial rearrangements of the promoter DNA binding elements to a distance which is compatible for engagement with the -10 and -35 sequences (Callaci *et al.*, 1999). Consequently, most functions of the σ^{70} -like factors are only able to manifest themselves in the context of the holoenzyme. With respect to the holoenzyme, it was proposed that because R1.1 is highly flexible its positioning may alter during the transition between RPc to RPo (Murakami *et al.*, 2002b; Nagai and Shimamoto, 1997). In fact, in the holoenzyme the negative nature of R1.1 allows it to bind deep within the dw-DBC (which may sufficiently widen the channel to facilitate the entry of ds-DNA). However, upon RPo formation R1.1 is displaced from the inside of the channel and is re-positioned to interact with the tip of the β dw-lobe. Therefore R1.1 also acts to promote transcription by serving as a molecular placeholder for the dw-DNA (Mekler *et al.*, 2002).

σ domain 2 is comprised of five conserved regions. R1.2 directly contacts the discriminator region (Haugen *et al.*, 2006). Regions 2.1 and 2.2 are important for forming interactions with the RNAP core, predominantly with the β' rudder, β lobes (Sharp *et al.*, 1999) and the β' coiled-coil (Murakami *et al.*, 2002b). Regions 2.3 and 2.4 have specific roles in promoter binding and DNA strand-separation; R2.4 contains the promoter recognition elements for interaction with the -10 promoter element (Siegele *et al.*, 1989) whereas residues in region 2.3 are responsible for the nucleation of DNA strand-separation by the flipping of a highly conserved adenine at position -11 of the non-template strand (Lim *et al.*, 2001; Tomsic *et al.*, 2001; deHaseth and Helmann, 1995).

σ domain 3 consists of regions 3.0 (Campbell *et al.*, 2002) (originally known as 2.5) and 3.1. R3.0 recognises the extended -10 promoter element (Barne *et al.*, 1997; Campbell *et al.*, 2002) whereas R3.1 serves as another RNAP core binding interface (Sharp *et al.*, 1999; Murakami *et al.*, 2002b).

The linker region between σ domains 3 and 4 constitutes R3.2. The linker contains nine negatively charged residues which are highly conserved in group 1 σ -factors (Gruber and Bryant, 1997) suggesting it has an important function. It seems that part of the linker forms a loop that

protrudes into the active site channel between the β' lid and β' rudder, whereas the rest of the linker is located within the RNA exit channel and therefore occupies the same space as the nascent RNA (Murakami *et al.*, 2002b; Vassylyev *et al.*, 2002). Since part of the loop is negatively charged it may function to serve as a molecular mimic or placeholder for RNA (Murakami *et al.*, 2002b). Consequently it has been hypothesised that the linker region must be displaced upon synthesis of an RNA product ≥ 9 –11 nt, thus extension of the RNA chain during abortive initiation displaces the linker and triggers σ -release and consequently promoter escape (Murakami *et al.*, 2002b; Mekler *et al.*, 2002).

The $\sigma 4$ domain, consisting of R4.1 and R4.2, forms RNAP core interactions with the β flap. R4.2 specifically, determines interactions with the -35 promoter element (Gardella *et al.*, 1989), and induces DNA bending to facilitate interactions between the α -CTDs and specific upstream DNA sequences (UP-elements) (Campbell *et al.*, 2002; Ross *et al.*, 1993). However, the presence of the extended -10 promoters makes region 4.2 dispensable for promoter binding (Kumar *et al.*, 1993).

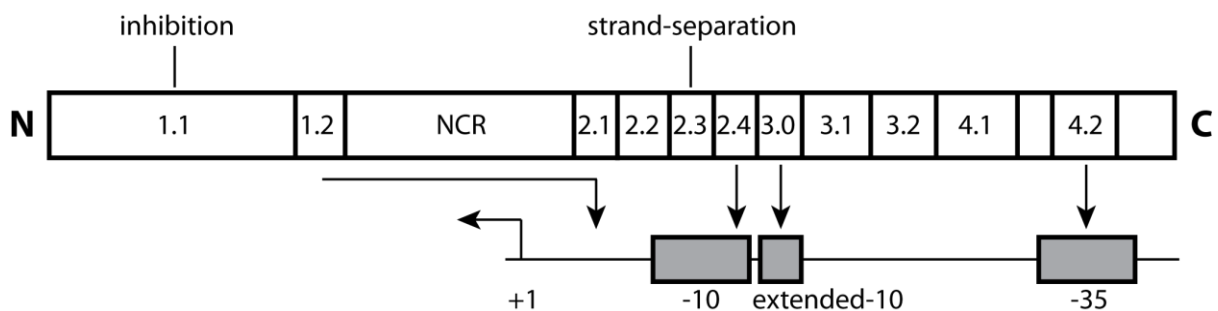


Figure 1.2.7 Conserved domains and functions of σ^{70} type factors.

The conserved regions and non-conserved region (NCR) of the σ^{70} type factors are labelled. Arrows point to the regions of contact between the σ -factor and the discriminator region, -10, extended -10 and -35 promoter elements. Regions 1.1 and 2.3 are required for inhibitory activity and initiation of DNA strand-separation respectively. Adapted from (Gruber and Gross, 2003).

1.3. The isomerisation from RPc to RPo

In addition to the RNAP core and holoenzyme, structures of the *T. aquaticus* and *T. thermophilus* holoenzyme bound to nucleic acids have also been resolved (Murakami *et al.*,

2002a; Vassilyev *et al.*, 2007). These structures, together with extensive biochemical and biophysical data have provided valuable insights into the extent of the holoenzyme-promoter DNA interactions and the dynamics of the conformational changes, which occur during transcription initiation and elongation. The most intricate stage of transcription initiation is the conversion of the R_{Pc} to the R_{Po} to generate the 'transcription bubble', a structure, which remains throughout elongation. The transition between the R_{Pc} and R_{Po} is a multi-step and reversible process, which involves the isomerisation of the DNA in conjunction with large scale conformational changes within the holoenzyme (Figure 1.3.1).

At most σ^{70} promoters, the R_{Pc} is typically formed by the initial engagement of the -10 and -35 promoter DNA elements with σ R2.4 and 4.2 of the holoenzyme. In the R_{Pc} the promoter DNA remains characteristically double stranded and resides outside of the main channel. The DNA in this complex demonstrates only upstream protection between $\sim$ -55 to -20 which primarily results from the contacts with the α -CTDs and σ R4.2 (Sclavi *et al.*, 2005). The lack of downstream-DNA contacts renders the R_{Pc} sensitive to competitors such as heparin and salt.

DNA strand-separation is facilitated by the motion of 'DNA breathing' which can be influenced by a variety of factors, including DNA sequence, supercoiling and wrapping (deHaseth and Helmann, 1995) as well as the chemical environment and the temperature. The nucleation of DNA strand-separation is triggered by aromatic residues in σ R2.3 which are responsible for the flipping of the highly conserved adenine at position -11 of the non-template strand (Heyduk *et al.*, 2006). Propagation of DNA strand-separation then occurs sequentially from the -10 promoter element. As the DNA strands separate the distance between the -10 base pair and the downstream edge of the bubble actually decreases. This implies that the DNA may adopt a scrunched conformation in the active site channel, not unlike the scrunching that occurs during abortive initiation (Chen *et al.*, 2010). The final 'transcription bubble' comprises $\sim$ 12 separated base pairs typically between the -10 to +3 sites. The single-stranded non-template DNA is positioned in a groove between the β lobes and interacts with conserved aromatic residues of σ R2.3. Whereas the single-stranded template DNA is enclosed in a tunnel formed from σ domains 2, 3, the β' lid and β' rudder, which position the +1 site next to the Mg^{2+} ions in the catalytic centre (Murakami *et al.*, 2002a).

The DNA at the downstream edge of the 'transcription bubble' bends or kinks sharply by 90° to position the duplex dw-DNA in line with the dw-DBC (Vassilyev *et al.*, 2007). The entry of the 22 Å ds-DNA in the 10-19 Å wide dw-DBC (Murakami *et al.*, 2002b; Vassilyev *et al.*, 2002) is facilitated by the opening of the clamp domains of the β and β' pincers (Chakraborty *et al.*, 2012) and appears to coincide with the displacement of σ R1.1 (Mekler *et al.*, 2002). The protection of the dw-DNA to +20,

arises from potential interactions with the β' jaw (Wigneshweraraj *et al.*, 2005), β' i6 (Artsimovitch *et al.*, 2003) and β clamp (Wigneshweraraj *et al.*, 2006) which are likely to contribute to the tight binding of the duplex dw-DNA (Kontur *et al.*, 2010), resulting in a stable, transcriptionally competent and competitor resistant RPo.

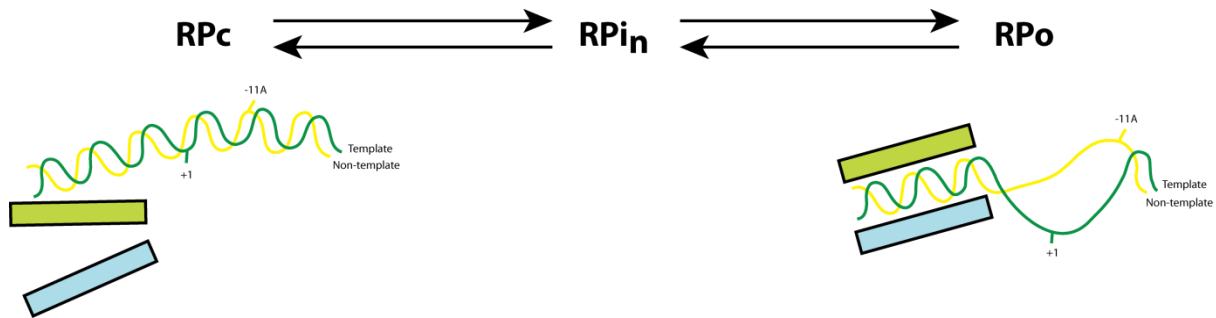


Figure 1.3.1 Transition between RPc and RPo.

A schematic illustrating the characteristic differences between the RPc and RPo. In the RPc the DNA remains in its duplex conformation outside of the main channel of the RNAP. In the RPo the DNA strands are locally separated and the dw-DNA is held by the β' jaw (green rectangle) and β' clamp (blue rectangle). Inspired by (Saecker *et al.*, 2011).

These sequences of events that lead to the formation of the RPo take place through relatively short-lived intermediates which are too transient to capture for structural analysis. However intermediate complexes at the *lacUV5*, T7A1 and λP_R promoters have been characterised using time-resolved footprinting methods (Buc and McClure, 1985; Sclavi *et al.*, 2005; Saecker *et al.*, 2011). The precise order and kinetics of events governing RPo formation are debatable and are likely to differ from promoter to promoter. The major conflicting hypothesis in RPo formation is centred around whether DNA bending and entry of the dw-DNA into the dw-DBC occurs before strand-separation as described by footprinting experiments using the λP_R promoter (Gries *et al.*, 2010; Kontur *et al.*, 2010), or whether DNA strand-separation facilitates DNA bending and dw-DNA loading as suggested from structural analysis and footprinting experiments on the T7A1 promoter (Chen *et al.*, 2010; Rogozina *et al.*, 2009).

1.4. Regulation of transcription initiation

Within a bacterial cell the supply of available RNAP is limited, therefore it is beneficial for the cell to function economically by controlling when and how much of each gene is expressed. The majority of RNAP is utilised for the transcription of genes required for translation and other housekeeping tasks, consequently, bacteria employ a range of methods to fine tune gene expression in response to environmental cues and stresses. The general distribution of RNAP can be affected by various mechanisms including the architecture of the DNA, the strength of promoter DNA sequences and the availability of specific σ -factors or nucleotides. Transcription initiation is also tightly controlled by a network of proteins known as transcription factors (TFs) which function to modify the activity of the holoenzyme at specific promoters. TFs can either up-regulate (activate) or down-regulate (repress/inhibit) transcription from specific or subsets of genes.

Many TFs are DNA binding proteins; in *E. coli* 50% of regulated genes are controlled by seven 'global' protein DNA binding TFs: CRP, FNR, IHF, Fis, ArcA, NarL and Lrp, conversely ~60 TFs can only regulate transcription from one promoter (Martinez-Antonio and Collado-Vides, 2003). DNA binding TFs regulate transcription initiation by binding to specific sites adjacent to promoters, and also often interact with the α -CTDs or σ domain 4. The location and sequence of transcription factor binding sites (TFBSs) typically determines whether a TF acts to repress or activate transcription of a certain gene. TFBSs are usually 12-30 bps in length and contain palindromic sequences or direct repeats as most TFs function as homodimers (van Hijum *et al.*, 2009). Activators increase the affinity of the RNAP for promoter DNA; three mechanisms have been described for simple activation. Class I activators target subunits of the RNAP. The cyclic AMP receptor protein (CRP) is a class I (and class II) activator which responds to elevated cAMP levels by binding to a position upstream of the *lac* promoter. CRP induces a bend in the DNA enabling it to make contact with the α -CTD of RNAP which is favourable for RPc formation, and with the α -NTD which facilitates isomerisation to the RPo (Busby and Ebright, 1999). The bacteriophage λ CI protein is an example of a class II activator which binds close to the -35 (bacteriophage λ PRM) promoter and makes contacts with domain 4 of σ^{70} (Nickels *et al.*, 2002). Class III activators such as members of the MerR family bind in the spacer region and twist the DNA, re-orientating the -10 and -35 promoters to an arrangement more optimal for RNAP binding (Browning and Busby, 2004). In addition, activators can also function to negatively modulate repressors of transcription (van Hijum *et al.*, 2009).

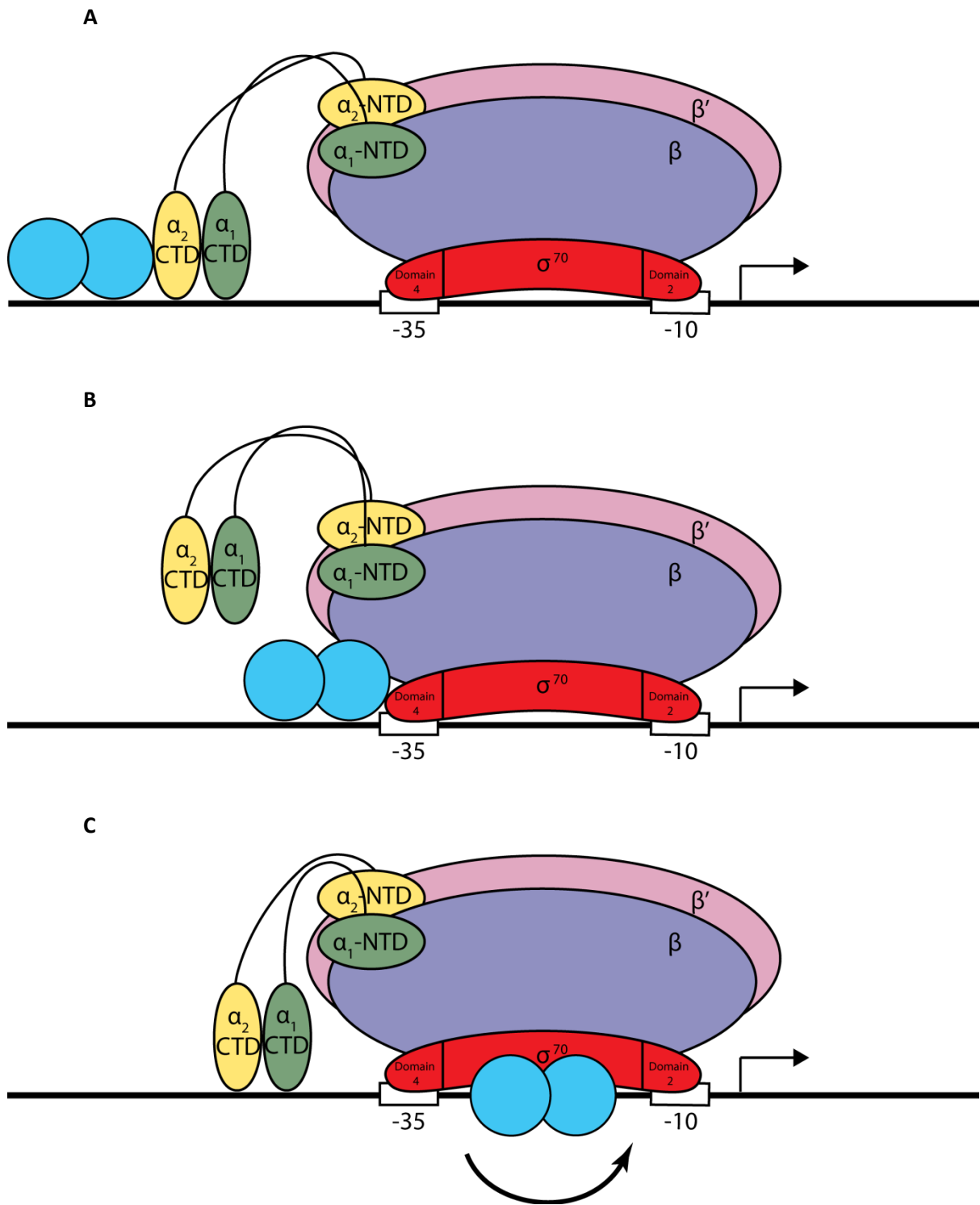
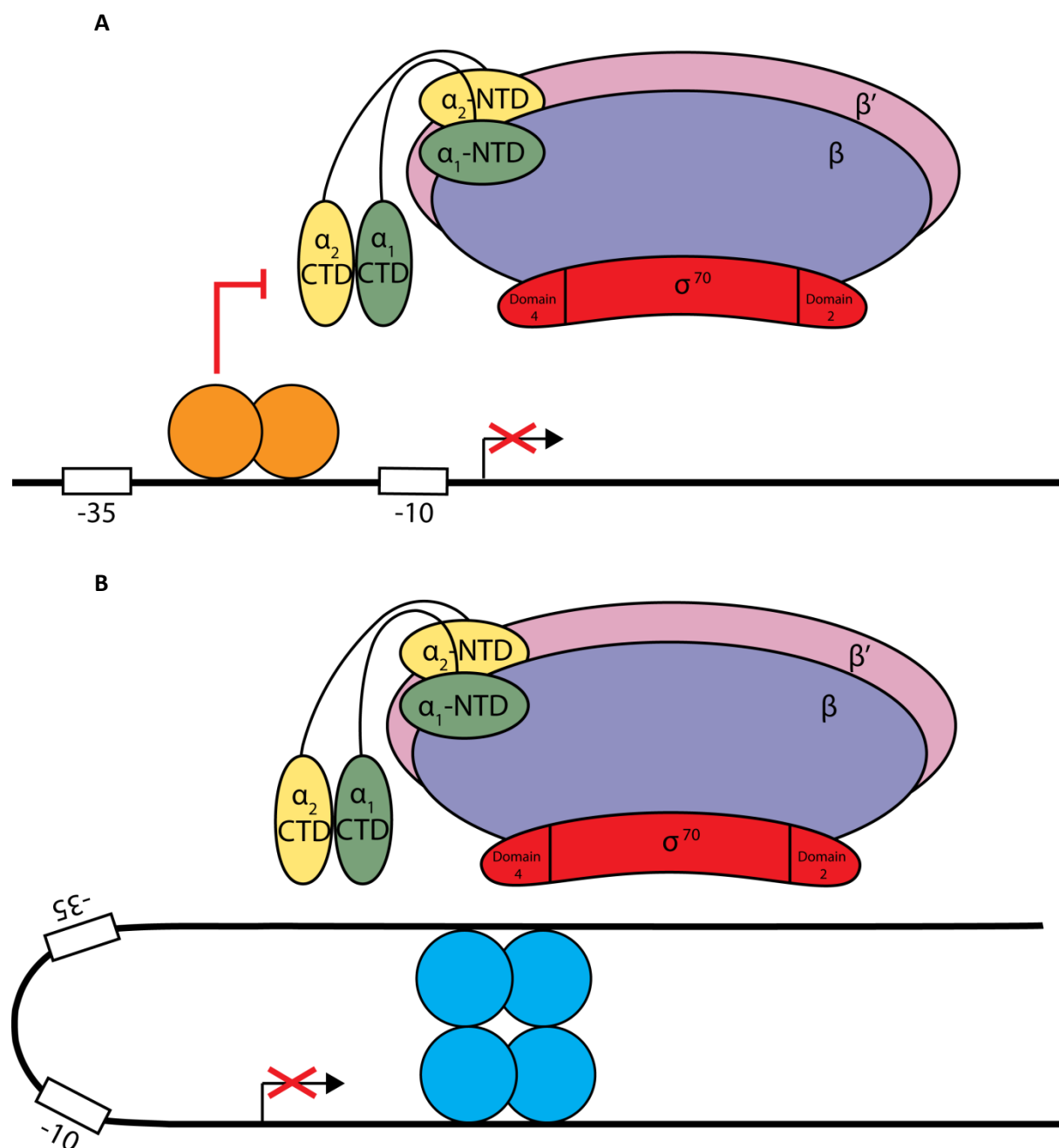


Figure 1.4.1 Mechanisms of simple activation.

Schematics illustrating the organisation of $E\sigma^{70}$ and activators (cyan) during simple activation. A) Class I activation. The activator is bound to an upstream site and contacts the α -CTD of RNAP, recruiting it to the promoter. B) Class II activation. The activator binds adjacent to the -35 promoter element and contacts σ^{70} domain 4. C) Class III activation. The activator binds at or near the -10 and -35 promoter elements and realigns them for optimal $E\sigma^{70}$ binding. Adapted from (Browning and Busby, 2004).

There are also four major modes by which TFs modulate transcription by simple repression. Class I repression can occur by steric hindrance either by directly binding to, or close to the promoter elements such as the *lac* repressor at the *lac* operon. Class II repression is brought about by the interaction of TFs such as GalR which bind upstream and downstream of the *gal* promoter to form a DNA loop, thereby preventing access to this region of DNA (Choy and Adhya, 1992). Class III repressors function as anti-activators. For instance CytR inhibits transcription initiation by contacting the promoter DNA bound CRP dimer and sterically blocking its interactions with the α -CTDs of the RNAP (Busby and Ebright, 1999).



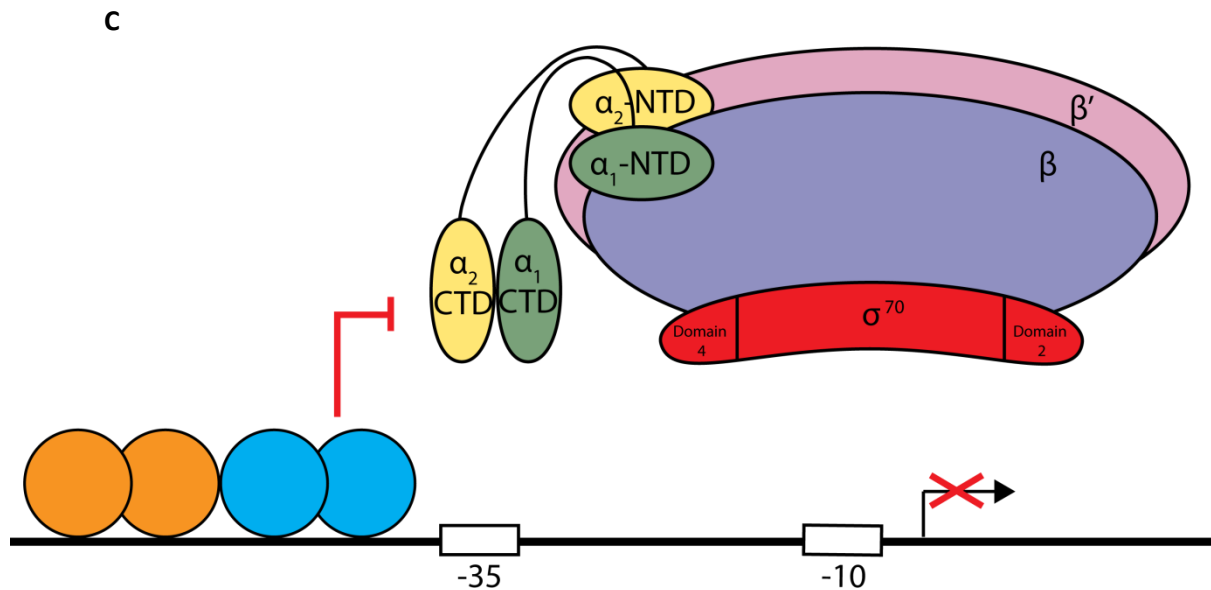


Figure 1.4.2 Mechanisms of simple repression.

Schematics illustrating the organisation of $E\sigma^{70}$, activators (cyan) and repressors (orange) during simple repression. A) Class I repression. The repressors blocks recognition of core promoter elements by $E\sigma^{70}$. B) Class II repression. Repressors bind to distal sites and interact to facilitate DNA looping. C) Class III repression. The repressors bind to activators to prevent them from functioning. Adapted from (Browning and Busby, 2004).

Furthermore, transcription is also subject to regulation by non-DNA binding factors (either proteins or nucleotides) that bind directly to subunits of the holoenzyme, irrespective of the promoter, to modulate the enzymes activity. For instance, alarmones are small modified nucleotides which can be synthesised and degraded quickly in response to the environment. The best studied alarmones are guanine tetraphosphate (ppGpp) and guanine pentaphosphate (pppGpp) which accumulate rapidly in amino acid starved *E. coli* cells in a process known as the stringent response. During amino acid starvation the binding of uncharged tRNAs to the ribosome stalls protein synthesis. This event triggers the synthesis of (p)ppGpp from GTP/GDP by the RelA protein. Other stress conditions such as depletion of phosphorus, carbon and iron also trigger the synthesis of (p)ppGpp, catalysed by SpoT or RelA homologs (Srivatsan and Wang, 2008). (p)ppGpp functions alongside its protein cofactor, DskA. (p)ppGpp is proposed to bind near to the RNAP active centre (Artsimovitch *et al.*, 2004) whereas DskA is structurally similar to the TFs GreA/B and therefore is likely to bind in the secondary channel (Symersky *et al.*, 2006). Together, their effects specifically alter gene expression from the rRNA and amino acid biosynthesis promoters. The rRNA promoters have a GC-rich discriminator region which is suboptimal for interacting with σ R1.2 (Haugen *et al.*, 2006). Therefore RPos from these promoters are intrinsically unstable (p)ppGpp and DskA act to lower the stability of these RPos even further, effectively preventing transcription initiation (Paul *et*

al., 2004). In contrast (p)ppGpp and DskA activate transcription from the amino acid biosynthesis promoters (which have AT-rich discriminator sequences) by increasing the rate of RPc to RPo isomerisation (Paul *et al.*, 2005). In addition (p)ppGpp and DskA have a global effect on gene expression by releasing RNAP from σ^{70} dependent genes, shifting the equilibrium towards the use of alternative σ -factors (Durfee *et al.*, 2008).

An alternative method by which bacterial transcription can be regulated is by non-native transcription factors which are produced by viruses that infect bacteria. To successfully propagate, the bacterial viruses synthesise low-MW, often non-DNA binding proteins, which specifically interact with and modulate the activity of the bacterial RNAP in favour of the limited transcription of host genes required for phage development and/or the efficient transcription of phage genes.

1.5. Bacteriophage

Bacterial viruses referred to as bacteriophages or simply phages are probably the most ubiquitous entities on the planet. Phage were first identified by Felix d'Herelle and Frederick Twort independently as bacterial lytic agents in the early 20th century (D'Herelle, 2007; Twort, 1936). Phages typically adopt one of two types of life cycles; lytic or lysogenic. Lytic phage cause immediate lysis of the bacterial host cell to release the phage progeny. Lysogenic or temperate phage integrate their phage DNA into the bacterial host genome to generate a prophage which multiplies upon host cell division. The prophage remains dormant until host cell conditions deteriorate which stimulates excision from the bacterial genome and then initiates the process of lysis. Thus both lytic and lysogenic phage are dependent on their bacterial hosts for proliferation and consequently have developed a variety of mechanisms to modify cellular components of the host to benefit the survival of the phage.

Regulation of the phage life cycle is achieved by sequential expression of the phage early middle and late genes. In view of the fact that the bacterial RNAP is a key target for regulation, it is not surprising that phages have evolved mechanisms to 'hijack' the host RNAP for expression of at least their early genes if not all of their genes. Thus the host RNAP may be prevented from transcribing its own genes as well as directed to the correct phage genes during the appropriate time during infection.

1.6. Phage-encoded transcription factors of bacterial RNAP

Phages employ two methods of host RNAP modulation: covalent modifications such as ADP-ribosylation or phosphorylation and regulation by non-DNA binding, low-MW proteins which bind directly to the host RNAP to control and/or direct its transcriptional activity (Nechaev and Severinov, 2003). The *E. coli* infecting, ds-DNA, lytic phage, T4, is a good example of how both types of modification can be utilised to divert transcription by the bacterial σ^{70} away from the host genes and towards the phage early and middle genes.

The T4 enzymes Alt and Mod covalently modify the bacterial RNAP early in infection via ADP-ribosylation. Alt acts with low efficiency to ADP-ribosylate the Arg²⁶⁵ residue of the α -CTD of only one of the α -subunits as well as other sites on α , β , β' , and σ^{70} . Whereas Mod specifically ADP-ribosylates the Arg²⁵⁶ residue of both α -subunits (Goff, 1984). ADP-ribosylated σ^{70} is unable to transcribe from UP-element-dependent and CRP-dependent promoters. As UP-element-dependent transcription from rRNA promoters accounts for a large proportion of RNA synthesis during *E. coli* exponential growth, it is thought that this covalent modification functions to increase the number of available RNAP for transcription of T4 genes (Nechaev and Severinov, 2003).

A product of T4 middle-stage expression is AsiA, a low-MW protein (Orsini *et al.*, 1993). AsiA exists as a homodimer which dissociates to bind to σ^{70} as a heterodimer with a 1:1 stoichiometry (Adelman *et al.*, 1997). AsiA specifically interacts with σ R4.2 and prevents its interaction with the -35 promoter element which thereby inhibits transcription from *E. coli* -10/-35 promoters and early phage promoters (Colland *et al.*, 1998; Minakhin *et al.*, 2001; Severinova *et al.*, 1998). In support of this view, extended -10 promoters and pre-formed RPos are not inhibited by AsiA (Adelman *et al.*, 1997). AsiA also cooperates with MotA, an early-stage T4 protein, to activate transcription of *E. coli* σ^{70} dependent middle-stage T4 genes which are predominantly under the control of the extended -10 class of promoters (Ouhammouch *et al.*, 1995). These middle-stage genes also contain a MotA binding site (Guild *et al.*, 1988) which is positioned to allow MotA to bind to σ^{70} R4.2 and therefore induce AsiA to interact with an identical sequence of amino acids at σ^{70} R4.1 (Minakhin *et al.*, 2003). The interaction of AsiA with σ^{70} R4.1 stimulates transcription by increasing the efficiency of RPo formation and promoter escape (Adelman *et al.*, 1998). Thus AsiA functions as a switch for the transition between T4 early-stage and middle-stage transcription; AsiA inhibits transcription from host and early T4 genes but together with MotA, activates transcription from T4 middle genes. T4 late gene expression also requires the host bacterial RNAP but σ^{70} is replaced by the phage encoded promoter specificity factors Gp55 and the activator Gp45 (Nechaev and Severinov, 2003).

Protein	Size (a.a.)	Phage	Host	Targets which stage(s) of transcription?	Effect on transcription
Alc	167	T4	<i>Escherichia coli</i>	Elongation/termination	Proposed to bind to the N-terminal of the β' subunit and inhibit transcription of unmodified cytosine-containing DNA.
AsiA	90	T4	<i>Escherichia coli</i>	Initiation	Interacts with σ R4.2 and prevents its interaction with the -35 promoter.
E3	237	SPO1	<i>Bacillus subtilis</i>	UNK	Proposed to bind to the N-terminal of the β' subunit to inhibit transcription.
Gp2	64	T7	<i>Escherichia coli</i>	Initiation	Binds to the β' jaw domain and antagonises RPo formation at -10/-35 promoters.
Gp24	131	Dp-1	<i>Streptococcus pneumoniae</i>	Initiation	Putative σ -factor which may direct host RNAP to phage genes.
Gp28	225	SPO1	<i>Bacillus subtilis</i>	Initiation	σ -like factor which binds to host RNAP core and directs it to middle promoters.
Gp36	214	PhiEco32	<i>Escherichia coli</i>	Initiation	σ -like factor which binds to host RNAP core, likely to direct transcription to middle or late phage genes.
Gp39	147	P23-45	<i>Thermus thermophilus</i>	Initiation, elongation, termination	Binds to the β flap and inhibits transcription from -10/-35 promoters as well as decreasing pausing during elongation and preventing termination.
Gp55	185	T4	<i>Escherichia coli</i>	Initiation	σ -like factor which binds to host RNAP core to transcribe late phage genes.
Gp69	204	Dp-1	<i>Streptococcus pneumoniae</i>	Initiation	Putative σ -factor which may direct host RNAP to phage genes.
Gp76	53	P23-45	<i>Thermus thermophilus</i>	Initiation	Inhibits transcription from -10/-35 promoters.
MotA	211	T4	<i>Escherichia coli</i>	Initiation	Binds to an upstream DNA site and σ R4.2 which induces AsiA to interact with σ R4.1, stimulating RPC formation and promoter escape.
N	107	λ	<i>Escherichia coli</i>	Elongation/termination	RNA binding protein that interacts with elongating RNAP to prevent termination at <i>nut</i> sites.
Nun	114	HK022	<i>Escherichia coli</i>	Termination	Interacts with RNA <i>nut</i> site and RNAP to abolish N-dependent anti-termination.
ORF67	198	G1	<i>Staphylococcus aureus</i>	Initiation	Binds to σ domain 4 to prevent DNA binding.
P7	73	Xp10	<i>Xanthomonas oryzae</i>	Initiation, termination	Binds to the N-terminal of the β' subunit to inhibit RPo formation at -10/-35 promoters and prevent termination.
Q protein	144	λ	<i>Escherichia coli</i>	Elongation, termination	Binds between the -10/-35 promoter elements and RNAP core to suppress promoter-proximal pausing and prevent termination.
ss-DNA binding (SSB)	265	N4	<i>Escherichia coli</i>	Initiation	Proposed to bind to the C-terminal of the β' subunit and activate Eo ⁷⁰ transcription from late phage promoters.

Table 1.6.1 Known phage-encoded proteins that bind to and modulate bacterial RNAP.

The T4 proteins MotA and AsiA are representatives of a novel group of non-bacterial TFs which bind to and modulate the bacterial RNAP. To date, over 5000 different phages have been isolated, of which ~750 have had their complete genomes sequenced (Hatfull and Hendrix, 2011). Considering the abundance and genetic diversity of the phage, only a very small proportion of the phage-encoded TFs have been studied, from a limited number of hosts. A list of the known and partly characterised phage-encoded proteins which bind to and modulate the bacterial RNAP during different stages of transcription is detailed in Table 1.6.1. This study aims to further contribute to the understanding of the regulation of bacterial RNAP during transcription initiation by investigating the molecular mechanisms of action of two phage-encoded proteins; T7 Gp2 and Xp10 P7.

1.7. T7 phage-encoded Gp2

T7 is also a ds-DNA lytic phage which is dependent on the *E. coli* host RNAP as well as its own RNAP for gene expression. T7 phage relies upon the host RNAP to transcribe early genes including *gene 1* which encodes the T7 single subunit RNAP (Studier, 1972). T7 RNAP is then responsible for the transcription of middle and late genes which are required for host transcription ‘shut-off’ and the production of viral progeny. Therefore the host transcription ‘shut-off’ regulates the switch from early to late phage gene expression. Host transcription ‘shut-off’ by T7 is implicated by two phage proteins, Gp0.7 an early gene product and Gp2 a middle gene product.

Gp0.7 is a serine/threonine protein kinase which acts to phosphorylate various host proteins including RNAP, the endonucleases RNase III and RNase E and the translational components initiation factor 1 (IF1), IF2, IF3, elongation factor G, ribosomal proteins S1 and S6, in addition to itself. Gp0.7 phosphorylation of RNAP occurs specifically at the amino acid residue T1068 of β' i6 and causes p -dependent transcription termination. Gp0.7 is dispensable for T7 growth under normal laboratory conditions but it becomes essential when Gp2 function is attenuated. This suggests that the functional roles of Gp0.7 and Gp2 in host transcription ‘shut-off’ may be overlapping (Severinova and Severinov, 2006).

Gp2 was first assigned a role in the host transcription ‘shut-off’ because it was discovered bound to the inactive bacterial RNAP isolated from T7 infected cells. Moreover, removal of Gp2 from RNAP with high salt levels, recovered transcription activity (Hesselbach and Nakada, 1977a; Hesselbach and Nakada, 1975). Additionally, T7 phage that are defective in *gene protein 2*, which encodes Gp2, can be rescued by the addition of rifampicin (the only clinically used antibiotic

that targets RNAP), further suggesting the involvement of Gp2 as an inhibitor of RNA synthesis (Hesselbach and Nakada, 1975;LeClerc and Richardson, 1979a). It seems that the precise role of Gp2 in the switch from early-late phage gene expression is to prevent anti-terminated transcription from the A3 promoter by the bacterial RNAP, into areas of the phage genome normally transcribed by the T7 RNAP (Savalia *et al.*, 2010). The bacterial RNAP translocates the DNA much slower than the T7 RNAP, which results in a roadblock and causes the later to pause. The paused T7 RNAP may then recruit DNA packaging machinery which introduces a premature break in the DNA resulting in the formation of short concatamers. Thus, Gp2 deficient T7 phage are able to infect and lyse the host cell efficiently, but do not produce viable progeny (Chamberlin, 1974;Center, 1975;LeClerc and Richardson, 1979b;DeWyngaert and Hinkle, 1979;DeWyngaert and Hinkle, 1980;Qimron *et al.*, 2008).

An initial investigation into the interaction between RNAP and Gp2 was based on the *E. coli* strains tsnB, BR3, and Y49. These strains were selected as they are non-permissive for T7 growth because Gp2 is no longer able to bind to and therefore inhibit the bacterial RNAP (Chamberlin, 1974;Studier, 1973;DeWyngaert and Hinkle, 1979;DeWyngaert and Hinkle, 1980;Shanblatt and Nakada, 1982). Sequencing of tsnB, BR3 and Y49, identified two individual charge reversal mutations in the β' jaw of RNAP; E1158K and E1188K, suggesting that these amino acid residues are important determinants for Gp2 binding to the RNAP (Nechaev and Severinov, 1999).

Early work from this laboratory also focused on studying the interaction between Gp2 and RNAP. A comprehensive alanine mutagenesis screen of Gp2 revealed that two amino acid residues, R56 and R58, which are highly conserved in Gp2, seem to be important for inhibitory activity (Figure 1.7.1). Native-PAGE assays showed that the R56A and R58A mutants of Gp2 were significantly impaired in binding to $E\sigma^{70}$. Moreover, the contribution of the charge of Gp2 amino acid residues 56 and 58 was assessed *in vivo*: charge reversal substitutions (R to E) significantly reduced the efficiency of T7 plating more than charge conservation (R to K) or charge removal (R to A). In view of the fact that Gp2-RNAP binding can be disrupted with high salt levels and that the charge reversal mutations, E1158K and E1188K in the β' jaw also abolish Gp2 binding, it would appear that the interaction between Gp2 and the β' jaw is likely to include a significant electrostatic component (Camara *et al.*, 2010). Although the major amino acid residues responsible for the Gp2-RNAP binding have been established, the affinity of the interaction between Gp2 and RNAP remains unknown. Therefore the first objective of this study was to develop and apply a solution based assay to quantitatively measure the affinity between Gp2 and RNAP.

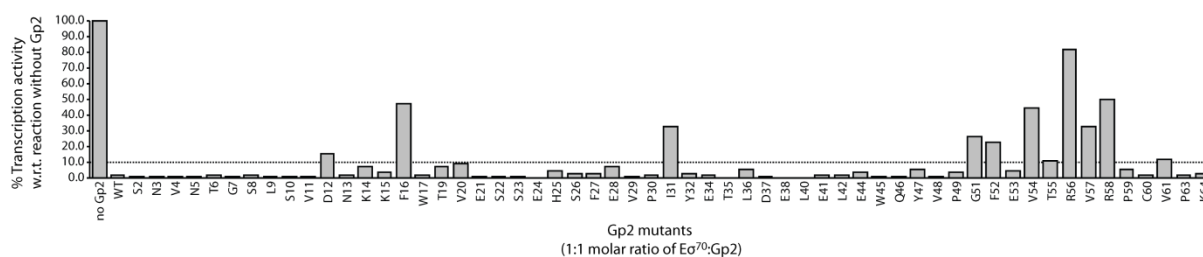


Figure 1.7.1 Alanine mutagenesis screen of Gp2.

A graph showing the effect of single alanine substitutions in Gp2. The inhibitory activity of each single alanine mutant version of Gp2 was assessed using an *in vitro* transcription assay: The ability of $E\sigma^{70}$ to synthesise the ApApUpU RNA transcript from the *lacUV5* promoter template in the presence of each Gp2 single-alanine mutant was normalised to a reaction conducted in the absence of Gp2 (first bar). Gp2 and $E\sigma^{70}$ were present at equimolar amounts. The dotted line indicates the cut-off for $E\sigma^{70}$ activity, below which a Gp2 mutant was considered to have a WT level of inhibitory activity.

Work from this laboratory also led to the determination of a high resolution 3-D, NMR structure of Gp2, which showed Gp2 to adopt a compact $\beta_1\beta_2\alpha_1\beta_3$ topology. The structure was initially used to further consider the results of the alanine mutagenesis screen. The structure of Gp2 revealed that the R56 and R58 amino acid residues were surface exposed at the C-terminal of the protein and positioned to project outwards to form an interaction with the β' jaw, whereas the majority of the remaining amino acid residues important for function (at residues F16, I31, G51, F52, V54 and V57) were located in positions proposed to stabilise the hydrophobic core of the protein or necessary for correct protein folding. Furthermore, a notable feature of the Gp2 structure is the separation of charges created by the R56 and R58 amino acid residues located at in the β_3 strand and a surface exposed strip of aspartic and glutamic acid residues (E34, D37, E38, E41, E44, E53 and E21) which run down the length of the α -helix to the $\beta_1\beta_2$ loop (Figure 1.7.2). The role of this negatively charged strip (NCS) of amino acid residues in Gp2 inhibition of RNAP is unknown. The second objective of this study is to investigate the role of the NCS in the Gp2-mediated inhibition of $E\sigma^{70}$.

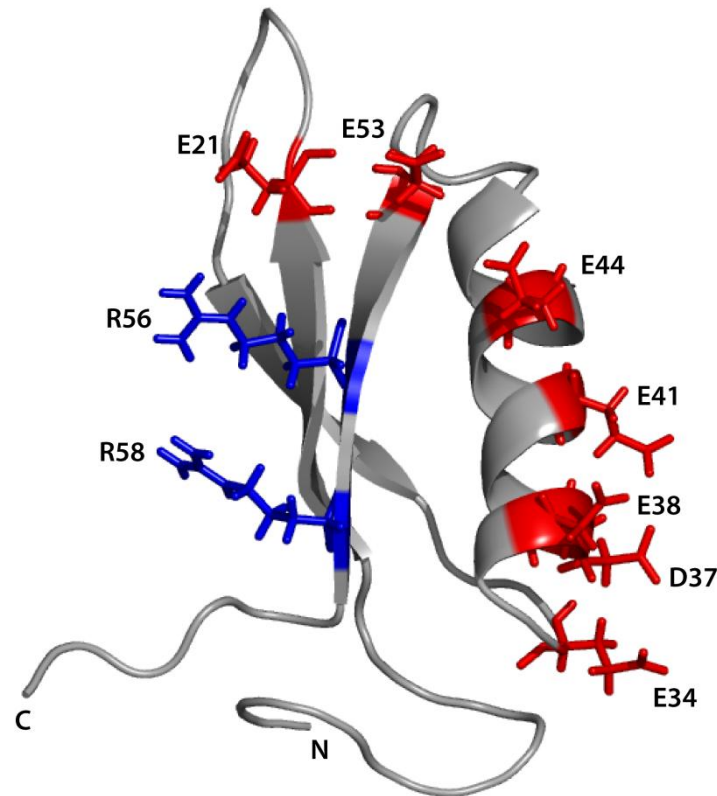


Figure 1.7.2 Charge separation of Gp2 structure.

Ribbon representation of the NMR structure of Gp2 with the positively charged binding determinants, R56 and R58 coloured blue and the aspartic and glutamic acid residues of the NCS (E34, D37, E38, E41, E44, E53 and E21) coloured red.

Gp2 inhibits the stage of transcription initiation, but can only exert its effects prior to the formation of the RPo, meaning that Gp2 has no effect on promoter escape or elongation (Hesselbach and Nakada, 1977b; Nechaev and Severinov, 1999). Gp2-mediated inhibition of transcription initiation appears to be both σ -factor and promoter specific; Gp2 is a potent inhibitor of σ^{70} transcription at -10/-35 promoters such as *lacUV5* and T7A1 but has significantly less effect at extended -10 promoters such as *galP1* (Nechaev and Severinov, 1999). Furthermore, transcription from promoters recognised by the activator and ATP-hydrolysis dependent σ^{54} is also resistant to Gp2 inhibition (Wigneshweraraj *et al.*, 2004). It was originally proposed that Gp2 inhibited transcription initiation by preventing RPo formation as judged by a reduction in promoter DNA binding (Hesselbach and Nakada, 1977b; Nechaev and Severinov, 1999). However, recently the mechanism of inhibition of transcription by Gp2 has been studied in more detail by this laboratory. Results from *in vitro* transcription assays showed that Gp2 does not affect RPo formation but a stage afterwards, leading to the RPo; Gp2 can successfully inhibit transcription initiation from homoduplex DNA templates but has markedly less effect on heteroduplex DNA templates (containing mismatches

between the -10/+3) which mimic the 'transcription bubble' in the RPo. Further *in vitro* transcription assays were conducted with different heteroduplex DNA templates to determine the length and position and of the mismatched segment required to overcome Gp2 inhibition. Results showed that inhibition was attenuated from DNA templates containing +1 proximal heteroduplex segments and that Gp2 is displaced upon $E\sigma^{70}$ -promoter DNA complex formation. This suggests that Gp2 affects a late stage *en route* to RPo formation associated with DNA strand-separation around the +1 site (Camara *et al.*, 2010). The third objective of this study is to further unravel the overall mechanism of inhibition of $E\sigma^{70}$ by Gp2. Since, σ R1.1, like Gp2 is highly negatively charged, also resides in the dw-DBC and has been proposed to play an important role in the progression to RPo formation, the third objective will focus on the potential contribution of σ R1.1 in the Gp2-mediated inhibition of $E\sigma^{70}$.

1.8. Xp10 phage-encoded P7

Xp10 is a ds-DNA lytic phage which infects the phytopathogenic bacterium *Xanthomonas oryzae* a major cause of rice blight. The genome organisation of Xp10 is different to T7 phage and more reminiscent of temperate phages related to λ . The Xp10 genome is split into two groups of genes which are expressed divergently. The leftward genes are involved in genome replication and expression and are expressed early in infection by the host RNAP, whereas rightward genes are required for the formation of phage progeny and host cell lysis and are transcribed by the phage RNAP in addition to the host RNAP (Yuzenkova *et al.*, 2003).

Xp10 was originally reported to employ an anti- σ activity to 'shut-off' host transcription as only the bacterial RNAP core was isolated from Xp10 infected *X. oryzae* cells compared to the isolation of the holoenzyme from un-infected *X. oryzae* cells (Liao and Kuo, 1986). However, later research proved this finding to be untrue. Instead it was shown that a low-MW phage protein referred to as P7 binds to the *X. oryzae* holoenzyme and is responsible for its inactivity in promoter-dependent transcription. Electrophoretic mobility shift assays (EMSAs) and *in vitro* transcription assays demonstrated that P7 prevents RPo formation as well as displaying characteristics of an anti-terminator. Additionally, like AsiA and Gp2, pre-formed RPos and transcription initiation from extended -10 promoters seems to be resistant to P7 inhibition (Nechaev *et al.*, 2002). Therefore it was proposed that P7 may play a dual role in Xp10 development; inhibiting transcription initiation of host RNAP to induce a switch between early and late phage genes and also ensuring that promoter-

distal phage genes (predominantly of the extended -10 type) are expressed, perhaps utilising an anti-termination mechanism (Yuzenkova *et al.*, 2003).

It was hypothesised that P7 may prevent RPo formation by interfering with conformational changes in the σ required for E σ -promoter complex formation. In support of this view, luminescence resonance energy transfer (LRET) experiments demonstrated that in the presence of P7 the interdomain distances between σ R2.4 and 4.2 were reduced to resemble that of free σ . Since the σ R2.4 and R4.2 are required for interaction with the -10 and -35 promoter elements, this result suggests that P7 may act to prevent R_{Pc} formation (Nechaev *et al.*, 2002). The precise mechanism of inhibition of transcription initiation by P7 remains unknown.

P7 appears to have no effect on transcription by *E. coli* E σ^{70} suggesting that P7 binds to an evolutionary variable region of the RNAP. The site of P7 binding to the *X. oryzae* RNAP was determined by co-purification, far-Western blotting and bacterial two-hybrid based interaction analyses. The results indicated that the precise site of P7 binding mapped to the N⁶L⁷F⁸N⁹ amino acid residues of the β' subunit, located at the upstream face of RNAP (Yuzenkova *et al.*, 2008). The fourth and fifth objectives of this study are to further characterise P7 activity by determining which of its amino acid residues are necessary for binding to the RNAP and to investigate the mechanism by which P7 inhibits transcription initiation.

1.9. Research objectives

T7 Gp2 and Xp10 P7 inhibit transcription initiation by RNAP at a stage(s) *en route* to RPo formation (Figure 1.9.1). This thesis concerns a detailed investigation into the molecular mechanism of action of Gp2 and also a preliminary characterisation of P7 with the following objectives:

- Establish and apply a quantitative assay for measuring the binding affinity between Gp2 and RNAP and variant versions thereof.
- Explore the role of the NCS of Gp2 in the inhibition of E σ^{70} .
- Investigate the contribution of σ R1.1 in the overall molecular mechanism of inhibition of E σ^{70} by Gp2.
- Determine the amino acid residues in P7 important for binding to RNAP.
- Investigate the mechanism by which P7 inhibits transcription initiation.

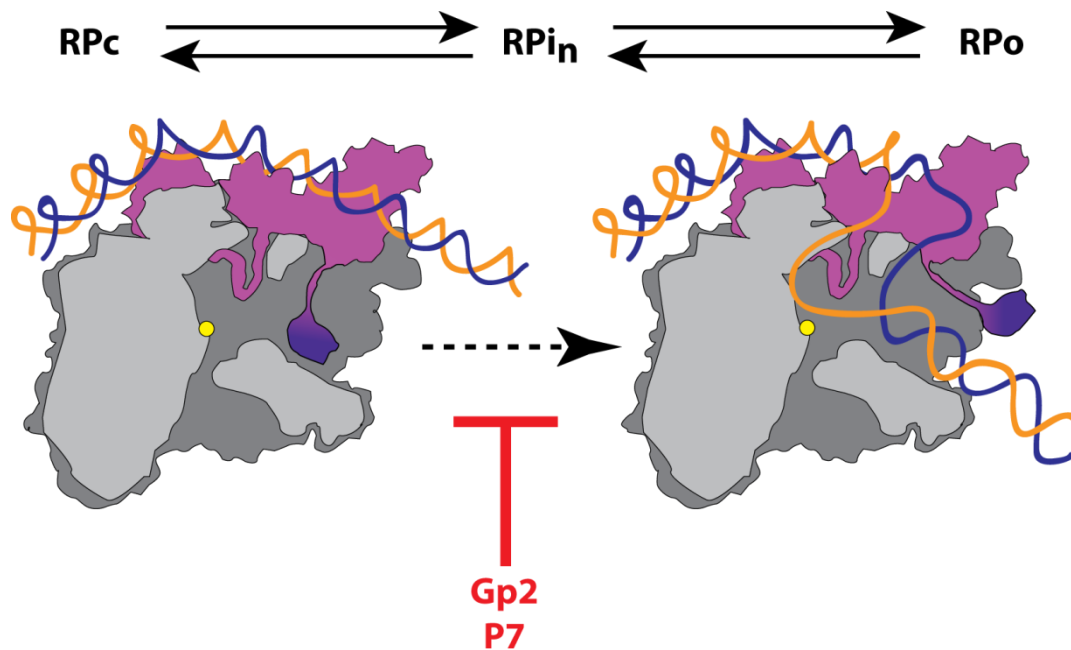


Figure 1.9.1 Gp2 and P7 inhibit RPo formation.

A cartoon depicting the major conformational changes between the RPc and RPo of $E\sigma^{70}$ (E is grey, σ^{70} is purple). The active site is represented by a yellow dot whereas the DNA template strand is shown in orange and the non-template strand in dark blue. Gp2 and P7 inhibit a stage(s) *en route* to RPo formation.

1.10. Why is this research important?

Transcription is a fundamental process in all living organisms. Bacterial transcription is catalysed by the multi-subunit enzyme RNAP. The regulation of bacterial transcription predominantly occurs at the initiation step where the activity of RNAP is controlled by a plethora of transcription regulatory factors which predominantly exert their effects by binding to the promoter DNA. This study focuses on furthering our current understanding of two, non-DNA binding transcription regulatory factors; the phage-encoded proteins, T7 Gp2 and Xp10 P7. These non-bacterial proteins specifically bind to RNAP to inhibit its activity and therefore represent a novel group of regulators of bacterial transcription initiation which are relatively poorly studied compared to their bacterial counterparts. Here, the regulation of RNAP by Gp2 and P7 has been investigated from a detailed biochemical and structural perspective to provide insights into their molecular mechanisms of inhibition. Investigating how such proteins inhibit bacterial RNAP activity is integral to furthering our understanding of the essential process of transcription and the complexity of its regulation.

The overuse and misuse of antibiotics has accelerated the evolution of drug-resistant bacteria, this effect has been exacerbated by the under development of new anti-bacterials. Thus, there is a significant, currently un-met need for the development of novel anti-bacterials. Bacterial RNAP is an essential enzyme and also a proven antibiotic target. Many of the available anti-bacterial compounds were originally isolated from other bacteria. However bacteriophage provide an alternative approach as they have evolved specific mechanisms to inhibit its hosts activity. The study of Gp2 and P7 not only uncovers novel vulnerabilities in the bacterial RNAP which can be exploited therapeutically but also provides a framework for the rational design of anti-bacterial compounds.

CHAPTER 2

MATERIALS AND METHODS

2.1. Solutions and buffers

Lysogeny broth (LB) (Merck)	10 g peptone, 5 g yeast extract, 10 g NaCl, dH ₂ O to 1L.
Lysogeny agar (LA) (Merck)	10 g peptone, 5 g yeast extract, 10 g NaCl, 20 g agar, dH ₂ O to 1L.
SOB (BD)	20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 2.4 g MgSO ₄ , 0.186 g KCl.
SOC	SOB + 20 mM glucose.
DNA loading dye (6x) (Fermentas)	10 mM Tris-HCl (pH 7.6), 0.03% (w/v) bromophenol blue, 0.03% (w/v) xylene, 60% (v/v) glycerol, 60 mM EDTA.
Nickel buffer A	25mM NaH ₂ PO ₄ (pH 7.0), 0.5 M NaCl, 5% (v/v) glycerol.
Lysis buffer	100 ml nickel buffer A + 1 protease inhibitor tablet (Roche).
Nickel buffer B	Nickel buffer A + 1 M imidazole.
Heparin dialysis buffer	10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 5% (v/v) glycerol, 0.1 mM EDTA, 1 mM DTT.
Heparin buffer A	Heparin dialysis buffer + 10 mM MgCl ₂ .
Heparin buffer B	Heparin buffer A + 1 M NaCl.
High-glycerol storage buffer	10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 50% (v/v) glycerol, 0.1 mM EDTA, 1 mM DTT.
High-throughput nickel buffer A	Nickel buffer A + 50 mM imidazole, 0.05% (v/v) Triton X-100.
High-throughput nickel buffer B	Nickel buffer A + 1 M imidazole, 0.05% (v/v) Triton X-100.
Conjugation buffer	10 mM Tris-HCl (pH 7.0), 300 mM NaCl, 0.1 mM EDTA, 5% (v/v) glycerol, 0.5mM TCEP.
Low -glycerol storage buffer	10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 50% (v/v) glycerol.
Acrylamide proto gel (30%) (solution I) (National Diagnostics)	30% (w/v) acrylamide, 0.8 % (w/v) methylene bis-acrylamide.

Acrylamide solution II	1.5 M Tris-HCl, pH 8.8, 0.3% (w/v) SDS.
Acrylamide solution III	0.5 M Tris-HCl, pH 6.8, 0.4% (w/v) SDS.
Laemmli buffer (2x) (Sigma-Aldrich)	4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) 2-mercaptoethanol, 0.004% (w/v) bromophenol blue, 0.125 M Tris-HCl, pH 6.8.
Tris glycine SDS (TGS) (10x) (National Diagnostics)	0.25 M Tris, 1.92 M glycine, 1% (w/v) SDS.
Phosphate buffered saline (PBS)	8 g NaCl, 0.2 g KCl, 1.44 g Na ₂ HPO ₄ , 0.24 g KH ₂ PO ₄ , dH ₂ O to 1 L, final pH 7.4.
Transblot buffer (20x)	24.2 g Tris, 150 g glycine, dH ₂ O to 1 L.
Transblot buffer (1x)	20 ml 20x transblot buffer, 40 ml methanol, dH ₂ O to 400 ml.
Tris buffered saline (TBS) (10x)	24.2 g Tris, 80 g NaCl, dH ₂ O to 1 L, final pH 7.6.
Tris buffered saline Tween 20 (TBSTT)	1x TBS, 0.1% (v/v) Tween 20.
Z-buffer	16.1 g Na ₂ HPO ₄ , 5.5 g NaH ₂ PO ₄ , 0.75 g KCl, 0.246 g MgSO ₄ , 2.7 ml 2-mercaptoethanol, dH ₂ O to 1 L, final pH 7.0.
TM buffer (10x)	10 mM Tris-HCl (pH 8.0), 10 mM MgCl ₂ .
PKA buffer (2x)	80 mM Tris-HCl (pH 6.4), 400 mM NaCl, 50 mM MgCl ₂ , 4 mM DTT, 50% (v/v) glycerol.
Reaction buffer (5x)	200 mM Tris-HCl (pH 8.0), 500 mM NaCl, 50 mM MgCl ₂ , 5 mM DTT.
Footprinting reaction buffer (5x)	200 mM Tris-HCl (pH 8.0), 200 mM NaCl, 50 mM MgCl ₂ .
Native loading dye (6x)	40 mM Tris-HCl, (pH 8.0), 100 mM NaCl, 10 mM MgCl ₂ , 1 mM DTT, 50% (v/v) glycerol, 0.03% (w/v) bromophenol blue.
Transcription stop buffer (3x)	0.03% (w/v) bromophenol blue, 0.03% (w/v) xylene, 20 mM EDTA in deionized formamide.
Tris glycine (TG) (10x) (National Diagnostics)	0.25 M Tris, 1.92 M glycine.
Acrylamide concentrate (National Diagnostics)	237.5 g acrylamide, 12.5 g methylene bis-acrylamide, 7.5 M urea, dH ₂ O to 1 L.

Acrylamide diluent (National Diagnostics)	7.5 M urea.
Tris borate EDTA (TBE) (10x) (Lonza)	0.89 M Tris borate, 0.02 M EDTA, final pH 8.3.
Fluorescence reaction buffer	40 mM Tris-HCl (pH 8.0), 50 mM NaCl, 10 mM MgCl ₂ , 5% (v/v) glycerol, 0.02% (v/v) Tween 20.

2.2. Bacterial strains and culture conditions

2.2.1. Media preparation

All media for bacterial growth were autoclaved at 121°C and 2 bars for 15 min or filter sterilised through a 0.2 µm cellulose acetate syringe filter (VWR). Where necessary, media were supplemented with antibiotics to select for only bacterial strains or plasmids conferring this resistance (Table 2.2.1). Cells grown in liquid culture were incubated in a water bath at 37°C, shaking at 145 rpm unless otherwise stated. The optical density of liquid cultures was measured using a Biochrom Libra S11 spectrophotometer at λ_{600} .

Antibiotic	Abbreviation	Stock concentration	Final concentration	Storage
Ampicillin	Amp	100 mg/ml in ethanol	100 µg/ml	-20°C
Kanamycin	Kan	50 mg/ml in ddH ₂ O	50 µg/ml	-20°C
Chloramphenicol	Cm	25 mg/ml in ethanol	25 µg/ml	-20°C

Table 2.2.1 Antibiotics used in this study.

The bacterial strains and plasmids used in this study are detailed in Table 2.2.2 and Table 2.2.3.

2.2.2. Bacterial strains

Strain	Description	Genotype	Antibiotic resistance	References
<i>E. coli</i> MG1655	A laboratory wild-type strain of <i>E. coli</i> K-12 with minimal genetic manipulations.	F ⁻ λ ⁻ ilvG ⁻ rfb-50 rph-1		(Blattner <i>et al.</i> , 1997)
<i>E. coli</i> DH5α	An <i>E. coli</i> K-12 derived transformable strain. This strain does not produce endonuclease I therefore avoids non-specific digestion resulting in high quality plasmid DNA. DH5α also lacks the alpha portion of the <i>lacZ</i> gene and thus is suitable for blue-white screening.	F ⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(r _K ⁻ m _K ⁺), λ-		<i>Life technologies</i>
<i>E. coli</i> BL21 (DE3)	BL21 (DE3) competent cells allow high-efficiency protein expression of target genes under the control of a T7 promoter. BL21 (DE3) is lysogenic for λ-DE3 which encodes the T7 RNA polymerase under the control of the <i>lacUV5</i> promoter inducible by IPTG.	F ⁻ ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])		(Studier and Moffatt, 1986)
<i>E. coli</i> NovaBlue (DE3)	A K-12 strain ideally suited as a cloning host due to its high transformation efficiency and <i>recA</i> , <i>endA</i> mutations, which result in high yields of unaltered plasmid DNA. This DE3 lysogen is useful for stringent expression due to the presence of the <i>lacI</i> ^q repressor.	endA1 hsdR17 (r _{K12} ⁻ m _{K12} ⁺) supE44 thi-1 recA1 gyrA96 relA1 lac F'[proA ⁺ B ⁺ lacI ^q Z ΔM15::Tn10] (Tet ^R)	Tet	<i>Novagen</i>
<i>E. coli</i> 397c	<i>E. coli</i> 397c contains a 16 bp deletion, resulting in a frameshift that leads to truncation of the β' subunit at the carboxyl terminus. This strain is temperature sensitive for growth at 43.5°C and unable to plate bacteriophage P2 at 33°C.	P90A5 λcl ₈₅₇ h ₈₀ S _{t68} dlac ⁺ rpoC Δ 4065-4080		(Christie <i>et al.</i> , 1996)
<i>E. coli</i> JE1134 (DE3)	MG1655 strain of <i>E. coli</i> lacking the 1149-1190 residues of the RNAP β' subunit, conferring resistance to Gp2. These cells are also lysogenic for λ-DE3.	MG1655 rpoC Δ 1149-1190		(Ederth <i>et al.</i> , 2002c)
<i>E. coli</i> KS1	MC1000 derivative containing a <i>lac</i> promoter bearing a single λ operator centred 62bp upstream of the transcription start site. Used for α-λCl activated transcription measurable in β-galactosidase assays.	MC1000 F' <i>lacI</i> ^q placO _R 2-62	Kan	(Dove <i>et al.</i> , 1997)

Table 2.2.2 Bacterial strains used in this study.

2.2.3. Bacterial plasmids

Plasmid	Description	Antibiotic Resistance	Reference
pET28b ⁺	Cloning and expression of recombinant proteins with an N-terminal His ₆ -tag and optional C-terminal His ₆ -tag. Expression of target genes is under the control of the T7 promoter.	Kan	Novagen
pET33b ⁺	Derived from pET28b ⁺ but carries an additional 15 bp sequence encoding the protein kinase A (PKA) site RRASV for ^{32/33} P-γ-ATP labelling.	Kan	Novagen
pBAD18	Expression of target genes is tightly controlled by the L-arabinose operon. Protein expression is positively and negatively regulated with the addition of arabinose or glucose respectively.	Amp	(Guzman <i>et al.</i> , 1995)
pVS10	Expression of <i>E. coli</i> RNAP under the control of the T7 gene 10 promoter. RNAP contains a His ₆ -tag at the C-terminal of the β subunit.	Amp	(Belogurov <i>et al.</i> , 2007)
pRL663	Expression of <i>E. coli</i> RNAP β' subunit with C-terminal His ₆ -tag under the control of the <i>trc</i> promoter.	Amp	(Wang <i>et al.</i> , 1995)
pBRα	'Prey' plasmid encodes the α-NTD which can be used to generate a fusion protein under the control of the tandem <i>lpp</i> and <i>lacUV5</i> promoters for two-hybrid protein-protein experiments.	Amp	(Dove and Hochschild, 2004)
pACλCI	'Bait' plasmid encodes λCI which can be used to generate a fusion protein under the control of the <i>lacUV5</i> promoter for two-hybrid protein-protein experiments.	Cm	(Dove and Hochschild, 2004)

Table 2.2.3 Bacterial plasmids used in this study.

2.3. Molecular cloning

2.3.1. CaCl₂ competent cell preparation

The *E. coli* strain of interest was streaked out onto an LA plate and incubated at 37°C overnight. A seed culture was prepared by inoculating 5 ml of LB with a single bacterial colony and incubating at 37°C overnight. 200 ml of LB was inoculated with 200 µl of seed culture and cells were grown at 37°C until mid exponential phase, $OD_{\lambda 600} = 0.4-0.45$. The culture was then centrifuged at $3200 \times g$ for 10 min at 4°C. The cell pellet was resuspended in 40 ml of pre-chilled 100 mM CaCl₂ and 15% (v/v) glycerol solution. Cells were then incubated on ice, in a 4°C room overnight or for 2 hours. The cells were centrifuged again at $3200 \times g$ for 10 min at 4°C. The pellet was then resuspended in 6 ml of pre-chilled 100 mM CaCl₂ and 15% (v/v) glycerol solution. Cells were either used immediately or aliquoted into 50 µl volumes, snap frozen in liquid nitrogen and stored at -80°C.

Note: *E. coli* 397c cells are temperature sensitive therefore need to be cultured at 30°C.

2.3.2. Transformation by heat shock

50 µl aliquots of competent cells (refer to section 2.3.1) were thawed on ice for 10 min. Approximately 100 ng of plasmid DNA was added to the cells and then incubated on ice for 30 min. The cells were heat shocked in a 42°C water bath for 1 min before immediately returning to ice for 2 min. Cells were recovered by the addition of 1 ml pre-warmed SOC media and incubation at 37°C, shaking for 1-2 hours. Cells were then centrifuged at $17,000 \times g$ for 2 min and resuspended in 50 µl of the original SOC media. The cell solution was then plated onto an appropriate LA plate and incubated at 37°C overnight.

2.3.3. Purification of plasmid DNA

A 5 ml overnight culture of cells harbouring the desired plasmid was centrifuged at $3200 \times g$ for 10 min and the supernatant discarded. The plasmid DNA was then purified from the cells using the QIAprep spin mini-prep kit (*Qiagen*) as per the manufacturer's instructions. The DNA concentration was measured using a NanoVue Plus spectrophotometer (*GE Healthcare*).

2.3.4. Polymerase chain reaction (PCR)

Reactions were conducted in 0.2 ml thin-walled PCR tubes.

Reaction component	μl per reaction
ddH ₂ O	40.5
10x PFU buffer (<i>Stratagene</i>)	5.0
25 mM dNTP (<i>Sigma</i>) mix	1.0
DNA template (100 ng/μl)	1.0
Forward Primer (10 μM)	1.0
Reverse Primer (10 μM)	1.0
PfuUltra II Fusion HS DNA Polymerase (<i>Stratagene</i>)	0.5
Total	50.0

94°C – 5 min		
94°C – 30 s	denaturing	30 cycles
53°C – 30 s	annealing*	
72°C – 15 s per 1 kb	extension	
72°C – 5 min		
10°C – ∞	hold	

* The annealing temperature was at least 5°C less than the lowest T_m of the primers.

PCR products were purified using a QIAquick PCR purification kit (*Qiagen*) as per the manufacturer's instructions. DNA concentration was measured using a NanoVue Plus spectrophotometer.

2.3.5. Restriction digest

Digestion reactions were performed using NEB enzymes and buffers. A suitable buffer for both enzymes was first selected or the reactions were performed sequentially following a DNA purification step. A typical reaction included 1 μl of 10x NEB buffer (1/2/3/4), 1 μl of each restriction enzyme, 10 μg/ml BSA if required and ddH₂O to 10 μl. Reactions were incubated at 37°C for 2 hours and then heat inactivated at 65°C for 20 min. To prevent re-ligation, the plasmid DNA was dephosphorylated at its 5' ends by adding 2.5 μl of shrimp alkaline phosphatase (SAP) and 2.5 μl of SAP

10x buffer (*Promega*) and incubating at 37 °C for 30 min. The enzyme was then heat inactivated by incubating at 75°C for 15 min. The digestions were purified using the QIAquick PCR purification kit as per the manufacturer's instructions. DNA concentration was measured using a NanoVue Plus spectrophotometer.

2.3.6. Agarose gel electrophoresis

Typically gels were prepared at either 1% (w/v) agarose (DNA > 200bp) or 2% (w/v) agarose (DNA < 200bp) by heating in 1x TBE. SYBR Safe DNA gel stain (*Invitrogen*) was added to the agarose solution at a 1:10,000 dilution and then poured into a level casting tray with an appropriate gel comb and allowed to set. 6x DNA loading dye (*Fermentas*) was added to the samples before loading into the gel alongside an O'Generuler 1 kb or 100 bp DNA marker (*Fermentas*). The gel was run at 100 V in 1x TBE for approximately 30-60 min before being removed and visualised using a Safe Imager Blue Light Transilluminator (*Invitrogen*). Gel images were captured using a Bio-rad GelDoc EZ imager.

2.3.7. Gel extraction and purification

DNA was resolved using an appropriate percentage agarose gel (refer to section 2.3.6) until the desired band(s) was distinguishable from any other DNA bands. The gel was transferred onto a clear glass plate and placed onto a blue light transilluminator safe imager. The band of interest was then excised from the gel using a scalpel and placed on a petri dish before freezing at -80°C. Once frozen, the gel slice was then cut into thin slices and transferred into a microcentrifuge tube. Sufficient ddH₂O was added to just cover the slices which were then vortexed until the agarose had dissolved. The DNA was purified from the agarose using the QIAquick PCR purification kit as per the manufacturer's instructions or by phenol-chloroform and ethanol precipitation (refer to section 2.3.8). The DNA concentration was measured using a NanoVue Plus spectrophotometer.

2.3.8. Phenol-chloroform extraction and ethanol precipitation

An equal volume of phenol-chloroform was added to the DNA containing solution, vortexed and centrifuged at 17,000 x *g* for 2 min. The upper-phase was then extracted being careful not to disrupt the interface, and transferred to a fresh microcentrifuge tube. 2.5 volumes of 100% ethanol and 0.25 volumes of 3M NaOAc pH 5.2 were added to the solution. The DNA was then precipitated by either incubating at -80°C for 1 hour or by freezing in liquid nitrogen. The solution was then centrifuged at 17,000 x *g* for 10 min and the supernatant discarded. The pellet was washed with 70%

pre-chilled ethanol and left to air dry. The DNA pellet was then resuspended in ddH₂O or appropriate buffer.

2.3.9. Ligation

Reactions were conducted using the purified digested plasmid DNA and insert DNA (refer to section 2.3.5) at a 3:1 ratio (refer to equation below). For example, 100 ng of plasmid DNA was combined with 33 ng of insert DNA, 0.5 µl of T4 ligase (5 U/µl) (*Fermentas*), 1 µl of 10x T4 ligase buffer (*Fermentas*) and ddH₂O to total 20 µl. The ligation reactions were incubated at 22°C for 10-60 min. 5 µl of ligation reaction was used per transformation (refer to section 2.3.2).

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng of insert}$$

2.3.10. Colony PCR

Primers were either designed to anneal to regions closely flanking the multiple cloning site (MCS) or alternatively one primer was complementary to the inserted gene and the second primer to a region within the plasmid.

Reactions were conducted in 0.2 ml thin-walled PCR tubes.

Reaction component	µl per reaction
ddH ₂ O	38.25
10x Taq buffer (<i>NEB</i>)	5.00
25 mM dNTP mix	1.25
Forward Primer (10 µM)	2.50
Reverse Primer (10 µM)	2.50
Taq DNA Polymerase (5 U/µl) (<i>NEB</i>)	0.50
Total	50.00

Each PCR reaction was inoculated with a single colony using a pipette tip. This tip was then dabbed onto a marked LA plate and incubated overnight at 37°C as a reference.

94°C – 5 min		
94°C – 30 s	denaturing	30 cycles
53°C – 30 s	annealing*	
72°C – 15 s per 1 kb	extension	
72°C – 5 min		
10°C – ∞	hold	

* The annealing temperature was at least 5°C less than the lowest T_m of the primers.

Positive clones were identified by resolving 5 µl of PCR product on a 1% (DNA > 200bp) or 2% (DNA < 200bp) (w/v) agarose gel (refer to section 2.3.6). Clones of the correct insert size were then sequenced using the appropriate primers (refer to section 2.3.12).

2.3.11. Site directed mutagenesis (SDM)

Primers were designed to anneal to the same sequence on opposite strands of the plasmid and with the desired mutation approximately in the middle of each primer.

Reactions were conducted in 0.2 ml thin-walled PCR tubes.

Reaction component	µl per reaction
ddH ₂ O	40.5
10x PFU buffer (<i>Stratagene</i>)	5.0
25 mM dNTP mix	1.0
DNA template (100 ng/µl)	1.0
Forward Primer (10 µM)	1.0
Reverse Primer (10 µM)	1.0
PfuUltra II Fusion HS DNA Polymerase (<i>Stratagene</i>)	0.5
Total	50.0

95°C – 30 s		
95°C – 30 s	denaturing	15 cycles
55°C – 1 min	annealing*	
68°C – 15 s per 1 kb	extension	
10°C – ∞	hold	

* The annealing temperature was at least 5°C less than the lowest T_m of the primers and often required optimisation for SDM.

The template DNA was digested by adding 6 µl buffer 4 (*NEB*), 3 µl ddH₂O, 1 µl DpnI (20 U/µl) (*NEB*) to the PCR reaction and incubating at 37°C for 1 hour before spiking with 1 µl DpnI and incubating for a further hour. The DNA was then purified using a QIAgen PCR purification kit as per the manufacturer's instructions. The plasmid was transformed into an *E. coli* cloning strain and incubated overnight at 37°C (refer to section 2.3.2). A single colony was then used to inoculate 5 ml of LB, which was incubated overnight, shaking at 37°C. The plasmid DNA was then purified from this culture using a QIAgen mini-prep kit as per the manufacturer's instructions. DNA concentration was measured using a NanoVue Plus spectrophotometer and plasmids were then sequenced using appropriate primers (refer to section 2.3.12).

2.3.12. Sequencing

DNA sequences were determined using the Imperial College Core Sequencing Service at the MRC genomics laboratory. Sequencing reactions contained 3.2 pmoles of primer, 150 - 300 ng of plasmid DNA per 3 kb or 500 - 600 ng for > 3kb plasmid DNA in 10 µl total ddH₂O. Samples were cycle sequenced using BigDye v 3.1 (*Applied Biosystems*) as follows:

94°C – 1 min	
94°C – 10 s	30 cycles
55°C – 15 s	
60°C – 4 min	
4°C – ∞	

Products were then purified by EDTA-ethanol precipitation, resuspended in highly deionised formamide and run on 3730xl DNA Analyser (*Applied Biosystems*). All sequences were then analysed using Chromas and Clone Manager.

2.4. Protein expression and purification

2.4.1. pET system of expression

Expression plasmids containing genes under the control of the T7 promoter were transformed (refer to section 2.3.2) into DE3 *E. coli* strains of cells, plated onto LA and incubated at 37°C overnight. 1 L of LB was inoculated with half of the transformants and grown, shaking at 37°C. When the OD_{λ600} reached 0.4-0.5 the culture was temperature shifted to 25°C for 30 min. Expression was then induced with the addition of a final concentration of 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). The cultures were grown for a further 3 hours, shaking at 25°C. 1 ml samples were taken before and after induction. These were centrifuged at 17,000 x *g* and the pellet was resuspended in 100 µl of ddH₂O for future SDS-PAGE analysis (refer to section 2.4.3). Cells were harvested by centrifugation at 2700 x *g*, 4°C and the pellet frozen at -20°C overnight.

2.4.2. Temperature sensitive system of expression

E. coli 397c cells were transformed with pRL663:rpoC^{WT/NLFN}, plated onto LA and incubated at 30°C overnight. 1 L of pre-warmed (42.5°C) LB was inoculated with half of the transformants together with 0.5 mM IPTG and grown shaking at 42.5°C for 8 hours. Cells were harvested by centrifugation at 2700 x *g*, 4°C and the pellet frozen at -20°C overnight.

2.4.3. SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE gels were prepared at an appropriate percentage according to the molecular weight of the protein of interest (the percentage acrylamide is inversely proportional to protein size). Solutions were prepared as described in Table 2.4.1 with the addition of APS and TEMED immediately prior to pouring. The separating solution was quickly poured into 1 mm thick Mini-PROTEAN glass plates (*Bio-Rad*) to 1 cm below the well level or half this amount if making a split percentage gel. 100 µl of isopropanol was pipetted on top to level the gel solution. Once the gel was set, the isopropanol was decanted and the top of the gel was washed with ddH₂O. Residual ddH₂O was removed by absorption with Whatman paper. The stacking gel was then prepared and poured on top of the separating gel and either a 10 or 15 well comb inserted. Once set, the gels could either be used immediately or for future use; wrapped in wet paper towel, then clingfilm and stored at 4°C.

	Separating				Stacking
	10%	12.5%	15%	20%	
Acrylamide (30%)	3.25ml	4.1ml	5.0ml	6.7ml	0.6ml
Acrylamide Solution II	2.5ml	2.5ml	2.5ml	2.5ml	-
Acrylamide Solution III	-	-	-	-	1ml
ddH ₂ O	4.2ml	3.4ml	2.5ml	0.8ml	2.4ml
10% (w/v) APS	100μl	100μl	100μl	100μl	40μl
TEMED	10μl	10μl	10μl	10μl	4μl

Table 2.4.1 Composition of SDS polyacrylamide gels.

The gels were assembled in a Mini-PROTEAN Tetra Electrophoresis gel tank (*Bio-Rad*) with 1x TGS buffer and pre-warmed by running at 200 V. Samples were prepared by mixing 10 μl of protein sample with 2x Laemmli buffer (*Sigma*), incubated at 95°C for 5 min and then briefly centrifuged. The samples were then loaded into the pre-warmed gel alongside 5 μl of PageRuler Prestained protein ladder (*Fermentas*) and the gel run at 200 V for approximately 50 min or until the dye front reached the bottom of the gel. The gel was then removed from the glass plates, washed three times in dH₂O and then stained with 20 ml of Bio-safe Coomassie stain (*Bio-rad*) for 30-60 min. The gel was then destained by rinsing twice in dH₂O. The gel was viewed on a light box and an image captured using the Biorad GelDoc EZ imager.

2.4.4. Nickel affinity purification (FPLC)

Recombinant His₆-tagged proteins were purified using nickel affinity chromatography. Cells were harvested as per section 2.4.1 and 2.4.2 then thawed and resuspended in 25 ml of pre-chilled lysis buffer and 25 mg of lysozyme. The cell solution was then transferred to a high-speed centrifuge tube and placed in a beaker of ice cold water. The cells were lysed by sonicating for 10 min (10 second pulses) with 40% amplitude using a Sonics Vibra-Cell sonicator or by passing through a French press twice until the cell lysate was transparent. The cell lysate was then centrifuged at 31400 x *g* for 30 min at 4°C. The supernatant (soluble fraction) was decanted into a pre-chilled universal tube and the pellet (insoluble fraction) roughly resuspended in an equivalent amount of nickel buffer A to the soluble fraction. A HisTrap column (*GE Healthcare*) was prepared by washing with 5 x column volumes of nickel buffer A, 0.1 M EDTA, nickel buffer A, ddH₂O, 0.1 M NiCl₂ (1.5 column volumes), ddH₂O, nickel buffer A, nickel buffer B, nickel buffer A. The supernatant of the cell lysate was then loaded into a glass super loop and injected into the AKTA prime (*GE Healthcare*) via

the HisTrap column. Typically a 3-7% nickel buffer B wash was performed before a 40 ml elution gradient. Fractions of washes and peaks were collected. 10 µl samples of each collected fraction were resolved by SDS-PAGE. Pure fractions were dialysed into storage buffer overnight at 4°C using dialysis membrane (*Spectrum laboratories*) with either 3500 or 6-8000 molecular weight (MW) cut off.

2.4.5. High-throughput nickel affinity purification

For small scale preparations of recombinant His₆-tagged proteins a high-throughput method of expression and purification was used. Expression plasmids containing genes under the control of the T7 promoter were transformed into *E. coli* DE3 cells. These transformants were used to inoculate 3 ml of LB in a 24 deep-well plate. The cultures were grown, shaking at 37°C, overnight. 100 µl of Cell lysis reagent (*Sigma*) and 2 µl of Lysonase (*Novagen*) was added to each well and incubated, shaking at RT for 15 min. 20 µl of Magne-His bead suspension (*Promega*) was added to each well before a further incubation, shaking at room temperature (RT) for 30 min. Using a 24-point magnetic block the lysates were washed three times with wash buffer. The proteins were then eluted from the magne-his beads with 200 µl of elution buffer. The eluate was then dialysed into suitable storage buffer using mini-dialysis units (*Thermo Scientific*) with a 3500 MW cut off.

2.4.6. Heparin affinity purification (FPLC)

Heparin is a highly negatively charged glycosaminoglycan with similar properties to RNA and DNA, therefore it can be immobilised for the affinity purification of nucleic acid binding proteins. RNAP and σ -factor containing eluate fractions that still contained contaminants after nickel column purification were pooled and dialysed into heparin dialysis buffer at 4°C overnight. A HiTrap column (*GE Healthcare*) was prepared by passing through 5 x column volumes of ddH₂O, 1 M NaCl, ddH₂O, heparin buffer A, heparin buffer B, heparin buffer A. 10 mM MgCl₂ was then added to the protein sample before loading into a glass super loop and injected into the HiTrap column attached to an AKTA prime (*GE Healthcare*). Typically a 20 ml elution gradient was performed. Fractions of washes and peaks were collected and 10 µl samples of each collected fraction were resolved by SDS-PAGE. Pure fractions were dialysed into storage buffer overnight at 4°C using dialysis membrane with a 6-8000 MW cut off.

2.4.7. Polyethyleneimine and ammonium sulphate precipitation

RNAP expressed from the *E. coli* 397c cells was first purified by polyethyleneimine (PEI) and ammonium sulphate precipitation followed by nickel affinity purification. PEI is used to precipitate nucleic acids and their bound proteins, whereas ammonium sulphate is used to precipitate proteins by 'salting out'. After protein expression (refer to section 2.4.2) the cells were harvested by centrifugation at 2700 x *g* for 30 min at 4°C. The cell pellets were then resuspended in 100 ml of RNAP lysis buffer A before re-centrifugation and storage of the cell pellet -20°C overnight. The cell pellet was thawed on ice and resuspended in 50 ml of RNAP lysis buffer B and 50 mg of lysozyme. The cells were then lysed by sonicating for 10 min (10 second pulses) with 40% amplitude using a Sonics Vibra-Cell sonicator. The cell lysates were centrifuged at 31,400 x *g* for 30 min at 4°C and the supernatant collected. All further steps were carried out at 4°C where possible. The nucleic acids and RNAP were precipitated by the gradual addition and mixing of 0.6% (v/v) PEI (*Sigma*) from a 10% stock (pH 7.9 using HCl), over a 15 min period before a further 30 min of mixing. The PEI precipitate was pelleted in a high speed centrifuge tube at 31,400 x *g* for 20 min and the pellet thoroughly resuspended in 10 ml RNAP wash buffer. The solution was re-centrifuged at 31,400 x *g* for 20 min and resuspended in 10 ml RNAP elution buffer, before further centrifugation at 31,400 x *g* for 20 min. The supernatant was collected and gently mixed with 0.4 g (NH₄)SO₄ (*Sigma*)/ ml supernatant, gradually added over 15 min, followed by a further 30 min of mixing. The precipitate was pelleted at 31,400 x *g* for 10 min, resuspended in nickel buffer A and dialysed into nickel buffer A overnight at 4°C.

2.4.8. Bradford assay

The Bradford assay was used to determine the concentrations of purified protein preparations. The absorbance of the bound form of the Coomassie Brilliant Blue G-250 dye is proportional to the concentration of protein in the sample. A protein standard curve was prepared by pipetting 5 µl of bovine serum albumin (BSA) (*NEB*) into individual wells of a clear, flat-bottomed 96 well microtiter plate. Typically the standard curve consisted of a serial dilution of 2.0, 1.75, 1.5, 1.25, 1.0, 0.5, 0.25, 0.125, 0.05 mg/ml BSA in ddH₂O and also a ddH₂O blank. 5 µl of the protein of interest was also pipetted, in triplicate into the plate. 250 µl of Quickstart Bradford dye reagent (*Bio-Rad*) was added to every protein containing well and then absorbance at 595 nm was measured in a BMG labtech POLARstar Omega plate reader.

2.4.9. Activator dependent transcriptional bacterial two-hybrid β -galactosidase assay.

Bacterial two-hybrid assays were conducted as described in (Dove and Hochschild, 2004); pBR α and pAC λ CI or derivatives were transformed into KS-1 cells and plated onto LA containing Kan, Amp and Cm antibiotics and incubated at 37°C overnight. A single colony of each transformation was used to inoculate 3 ml of LB containing antibiotics and 20 μ M IPTG, these were cultured for 16 hours at 30°C. 0.5 ml of the overnight cultures was diluted in 2 ml of LB containing antibiotics and IPTG. 1 ml of diluted culture was used to take an absorbance reading at 600 nm. The cells in the remaining diluted cultures were lysed by addition of 20 μ l choloform and 20 μ l of 0.1% SDS followed by vortexing for 10 s and shaking at 37°C for 40 min. 10 μ l of the lysed cell solutions were mixed in 990 μ l of Z-buffer and then 10 μ l was transferred to a 96 well plate. The plate was incubated at 30°C for 5 min before the addition of 25 μ l of pre-warmed 4 mg/ml ONPG. The plate was incubated for a further 30 min at 30°C and then the reactions stopped with 50 μ l of 1 M Na₂CO₃. The OD at λ_{600} , λ_{420} and λ_{550} was then measured. The enzymatic activity in miller units (MU) was calculated as follows:

$$MU = 1000 \times [OD_{\lambda_{420}} - (1.75 \times OD_{\lambda_{550}})] \div (t \times v \times OD_{\lambda_{600}})$$

OD $_{\lambda_{600}}$ = cell density.

OD $_{\lambda_{550}}$ = scatter from cell debris, multiplication by 1.75 approximates for scatter at 420 nm.

OD $_{\lambda_{420}}$ = absorbance of o-nitrophenol.

t = time of reaction in min

v = volume of cells in ml

2.5. Immunoblotting

2.5.1. Polyclonal antibody production

Peptide sequences were designed by Eurogentec for rabbit immunisation. Immunisation was conducted using two different peptide sequences per rabbit. Bleeds were taken pre-immunisation, 21 days after immunisation and after termination (final bleed). Anti-peptide ELISA analysis was conducted for the final bleed prior to shipping.

2.5.2. Affinity purification of antibody serum

50 µg of the protein of interest was resolved by SDS-PAGE and transferred to PVDF membrane as per the Western blotting protocol (refer to section 2.5.4). The membrane was then stained with Ponceau S for 5 min and rinsed with dH₂O to visualise the protein. The protein containing sections of the membrane were excised and placed in a 50 ml tube. The membrane was blocked with 25 ml of 5% fat-free milk powder (in PBS) for 1 hour and then incubated with 25 ml 5% fat-free milk powder + 5 ml final bleed antibody serum at 4°C overnight. The membrane was rinsed twice in PBS for a few seconds each time and then the antibody was eluted with 4 ml 100 mM glycine pH 2.7 for exactly 2 min. The eluate was transferred into a 15 ml tube containing 300 µl 1.5 M Tris-HCl pH 8.8 to neutralise the acid. The eluate was then dialysed against 500 ml PBS twice over 8 hrs. The antibody was then aliquoted and frozen at -20°C.

2.5.3. Antibody purification against cell background

To remove cross-reactivity and non-specific background the antibodies were purified against the relevant cell extract. An overnight 200 ml culture of cell strain without the protein of interest was centrifuged at 3200 x *g* and resuspended in 2 ml of PBS. Cells were lysed by sonication as described in section 2.4.4. 100 µl of antibody was then added to 1 ml of cell extract and incubated at 37°C for 1 hour. The cell extract solution was then centrifuged at 31,400 x *g* for 30 min at 4°C. The supernatant was then transferred to the second 1 ml of cell extract and the incubation and centrifugation step was repeated. The antibody containing supernatant was stored at 4°C or in aliquots at -20°C.

2.5.4. Western blotting

Protein samples were resolved using the appropriate percentage SDS-PAGE (refer to section 2.4.3). The gel was then blotted in 1x transblot buffer for 10 min. A gel-sized sheet of PVDF membrane was soaked in methanol for 10 s, in water for 5 min and then equilibrated in 1x transblot buffer for 10 min. Meanwhile, eight gel-sized sheets of Whatman paper were soaked in 1x transblot buffer. The blot was assembled onto a Bio-rad trans-blot semi-dry transfer cell; with the gel on top of the membrane and 4 pieces of Whatman paper either side. Any bubbles were then removed by rolling over the blot with a stripette. The blot was transferred to the membrane by running at 10 V for 1 hour. The membrane was then blocked in 5% fat-free milk powder (in TBS) for 1 hour at RT or overnight at 4°C. The membrane was washed twice for 10 min each in TBSTT and for a further 10 min in TBS. The membrane was then incubated for 1 hour with the primary antibody, diluted in 5%

fat-free milk powder. After incubation the membrane was washed five times for 10 min each in TBSTT. On parafilm, Supersignal west femto/pico ECL solution (*Pierce*) was prepared as per the manufacturer's instructions. The blot was then placed face down into the ECL mixture for 30 s (femto) or 5 min (pico). The membrane was then lightly dried on blue towel, covered with clingfilm and placed into a cassette. The image was either visualised using an LAS-3000 Fuji Imager.

2.6. In vitro biochemical assays

2.6.1. ^{32}P -labelling of DNA probes

Oligonucleotides were custom made and PAGE purified (*Sigma-Aldrich*). Either the non-template (NT) or template (T) strand was labelled with ^{32}P by incubating 20 pmoles of oligonucleotide, 2 μl of 10x polynucleotide (PNK) buffer A, 3 μl of ^{32}P - γ -ATP (250 Ci/mmol) (*Perkin Elmer*) and 1 μl polynucleotide kinase (PNK) (10 U/ μl) (*Fermentas*) in a total volume of 20 μl of ddH₂O at 37°C for 30 min. The enzyme was inactivated by incubating the reaction at 75°C for 10 min. 40 pmoles of the complementary oligonucleotide was added to the reaction along with 4 μl of 10x TM buffer and made up to 40 μl with ddH₂O. The oligonucleotides were denatured at 95°C for 5 min and then allowed to anneal by slowly cooling to RT. The final concentration of labelled duplexed DNA was estimated to be 500 nM.

2.6.2. Electrophoretic mobility shift assay (EMSA)

Typically, EMSAs were conducted in 10 μl reactions; 100 nM RNAP was reconstituted with 400 nM σ -factor in 2 μl 5x reaction buffer and ddH₂O for 5 min. RNAP-promoter complexes were formed by adding 20 nM ^{32}P -labelled promoter DNA followed by a further 5 min incubation. Reactions were either conducted at 4 or 37°C to maximise the percentage of either closed or open complexes. In some instances, transcription factors such as Gp2 or P7 were added in varying molar ratios at different stages of the reaction. Additionally, the stability of RNAP-DNA complexes was assessed by further incubation with 100 $\mu\text{g}/\text{ml}$ heparin. 2 μl of native loading dye was added to reactions before loading onto a 4.5% native polyacrylamide gel (Table 2.6.1) run at 100 V in TG buffer at either 4°C/RT/37°C for 30-90 min depending on temperature and length of DNA. Gels were transferred to Whatman paper, dried and visualised using a FLA-5000 Fujifilm phosphorimager.

Native PAGE (4.5%)	
Acrylamide (30%)	3 ml
10x TG buffer	2 ml
ddH ₂ O	15 ml
10% (w/v) APS	200 µl
TEMED	20 µl

Table 2.6.1 Composition of a native 4.5% polyacrylamide gel.

2.6.3. ³²P-labelling of proteins

Recombinant proteins were ³²P-labelled at the serine residue within the RRASV sequence. Proteins were labelled in reactions containing, 50 µl 2x protein kinase A (PKA) reaction buffer, 500 ng of protein, 1 µl of ³²P-γ-ATP (250 Ci/mmol-Perkin Elmer) and 5 U of PKA made up to 100 µl with ddH₂O. Reactions were incubated at 37°C for 30 min.

2.6.4. Protein-protein native gel binding assay

Protein-protein native gel binding experiments were conducted at 37°C in 10 µl reactions. Typically, increasing concentrations of RNAP were reconstituted with σ-factor using a 1:4 molar ratio and incubated with 2 µl 5x reaction buffer, 1 µg α-lactalbumin and ddH₂O for 5 min. RNAP-Gp2 complexes were formed by adding 50 nM ³²P-Gp2 (refer to section 2.6.3) followed by a further 5 min incubation. 2 µl of native loading dye was added to reactions before loading onto a 4.5% native polyacrylamide gel (Table 2.6.1) run at 100 V in TG buffer for 80 min at 37°C. Gels were transferred to Whatman paper, dried and visualised using a FLA-5000 Fujifilm phosphorimager.

2.6.5. Abortive transcription assay

Transcription reactions were performed in 10 µl volumes at 37°C. Typically 100 nM RNAP was reconstituted with 400 nM σ-factor and incubated for 5 min. 20 nM DNA promoter template was added and incubated for a further 5 min. Then, an elongation mixture contained 0.25 mM ApA dinucleotide primer, 0.05 mM UTP, 0.2 µl of ³²P-α-UTP (250 Ci/mmol) and 100 µg/ml heparin was added and the reaction allowed to proceed for 5 min. Reactions were stopped with 4 µl of stop dye and the samples were resolved on a 20% urea denaturing acrylamide gel (Table 2.6.2) for 45 min at

100 V. The gels were exposed to an imaging plate and visualised by FLA-5000 Fujifilm phosphorimager.

Urea-denaturing polyacrylamide gel (20%)	
Acrylamide concentrate	20 ml
Acrylamide diluents	2.5 ml
10x TBE buffer	2.5 ml
10% (w/v) APS	150 µl
TEMED	20 µl

Table 2.6.2 Composition of a urea-denaturing 20% polyacrylamide gel.

2.7. Fluorescent binding assays

2.7.1. Alexa-488 labelling of Gp2

Gp2 and derivatives were purified by nickel affinity chromatography, dialysed into conjugation buffer and then the protein concentration was determined by Bradford assay (refer to sections 2.4.4, 2.4.5 and 2.4.8). A stock solution of 10 mM Alexa-488 (*Invitrogen*) was prepared by reconstitution in dimethylformamide (DMF). Gp2 was combined with 5 times excess Alexa-488 and incubated at RT for 2 hours. The reaction was quenched with 10 mM 2-mercaptoethanol. The conjugated protein was then dialysed into nickel buffer A at 4°C overnight. The labelled protein was re-purified from excess Alexa-488 by nickel affinity chromatography. The peak fractions were dialysed (3500 MW cut off) at 4°C, overnight into low-glycerol storage buffer and the concentration of labelled protein was measured via the Bradford assay. Labelled protein was snap frozen in liquid nitrogen and stored at -80°C. The conjugation efficiency was calculated using $[\text{Alexa-488}]/[\text{protein}]$. This equation assumes that all Alexa-488 is bound to the protein with a 1:1 ratio.

2.7.2. Fluorescence anisotropy

50 nM Gp2^{Alexa488} was diluted in 400 µl of low-glycerol storage buffer within a 5 mm glass cuvette. The degree of polarisation was measured using a Fluoromax-3 spectrofluorometer (*Horiba*

Scientific) fitted with automated polarisation filters. Measurements were recorded using an excitation and emission wavelengths of 495 nm and 519 nm respectively with a G-factor of 0.63 and slit size of 0.5 nm. Binding curves were generated at 37°C by titrating in RNAP or reconstituted holoenzyme (E:σ at a 1:4 molar ratio) and measuring the anisotropy 5 times in immediate succession. Data were fitted using the tight binding equation in Grafit 5 (Erithacus Software) as show below.

$$A = A_D + (A_{DE} - A_D) \times \left(\frac{\{(K_d + [E]_0 + [Gp2]_T) - \sqrt{(-K_d - [E]_0 - [Gp2]_T)^2 - 4[E]_0[Gp2]_T}\}}{2[Gp2]_T} \right)$$

Where A is the observed anisotropy, A_D is the anisotropy of free Gp2^{Alexa488}, A_{DE} is the anisotropy of RNAP-bound Gp2^{Alexa488}, K_d is the equilibrium dissociation constant, $[Gp2]_T$ is the total concentration of Gp2^{Alexa488} and $[E]_0$ is the RNAP concentration.

2.7.3. TMR labelling of σ⁷⁰

Single cysteine σ⁷⁰ and derivatives were labelled at position 211 as described in (Mekler *et al.*, 2002). Labelling reactions were conducted in 1 ml volumes and contained 20 μM single cys σ⁷⁰ subjected to solid-phase reduction on Reduce-Imm (*Pierce*) as per the manufacturer's instructions, 200 μM tetramethylrhodamine-5-maleimide (TMR) (*Molecular Probes*), 100 mM sodium phosphate (pH 8.0) and 1 mM EDTA. Following a 1 hr incubation at 0°C, products were purified by gel-filtration chromatography on Bio-Gel P6DG (*Bio-Rad*) and stored in high-glycerol storage buffer at -20°C.

2.7.4. RNAP beacon assay

Fluorescence measurements were performed using a QuantaMaster QM4 spectrofluorometer with an excitation wavelength of 550 nm and an emission of 578 nm. Reactions were performed at 25°C in a plastic cuvette with a total volume of 800 μl fluorescence reaction buffer. Reactions typically contained 1 nM labelled Eσ⁷⁰, DNA probes at various concentrations (according to their binding affinities) and inhibitors were added at various stages of the reaction. Time dependent changes in fluorescence intensity were measured after a manual mixing time of 15 s.

CHAPTER 3

ESTABLISHING A SOLUTION-BASED ASSAY TO DETECT THE BINDING OF GP2 TO *E. COLI* RNAP

3.1. Introduction

There are a wide variety of methods to detect and analyse protein-protein interactions. The choice of method ultimately depends on the circumstance of the interaction and the extent of characterisation required. Biochemical techniques such as co-immunoprecipitation, crosslinking, electrophoretic mobility shift assays (EMSAs) as well as phage display and bacterial/yeast two-hybrid assays are typically used for the screening of protein-protein interactions but are limited in terms of specificity, therefore proposed interactions often require further confirmation and a more detailed analysis. Biophysical techniques such as fluorescence anisotropy, fluorescence resonance energy transfer, surface plasmon resonance, dual polarisation interferometry and isothermal titration calorimetry are able to accurately measure binding affinities and kinetics of interactions as well as informing on stoichiometry and conformational changes. However, some of these techniques require more elaborate and specialised equipment which is perhaps beyond the scope of many biological demands.

Previous work in this laboratory has used conventional gel-based methods such as EMSAs to assess the binding between wild-type (WT) and variant versions of Gp2 and *E. coli* RNAP (Camara *et al.*, 2010). However such techniques are neither quantitative nor sensitive enough to accurately measure binding affinities, determine stoichiometries nor to detect subtle differences in binding. In addition, the experimental conditions used in gel-based assays are somewhat artificial, thus often disrupt the binding equilibrium of the reaction which is particularly biased against transient interactions. The nature of solution-based assays is not only more native but also allows for the easy manipulation of pH and/or salt concentration in order to investigate the type of the interaction. This study aims to establish a reliable, solution based method which can be used to accurately detect the binding and measure the affinity of Gp2 to RNAP and variants thereof.

Fluorescent anisotropy (FA) is a solution based technique that is frequently used to measure the binding constants and kinetics of reactions, as it only requires limited modification to one of the binding partners and a spectrofluorometer for obtaining measurements. This technique depends on a significant size difference between the two interacting molecules of which the smaller molecule is fluorescently labelled. When the fluorescently-labelled molecule is in solution it rotates freely, which upon excitation with polarized light has a scrambling effect, generating a depolarized emission (Figure 3.1.1, left panel). The ratio of polarized light relative to the total light intensity of the emission is expressed in terms of anisotropy. The anisotropy measurements describe the average angular displacement of the fluorophore that occurs between the absorption and subsequent

emission of a photon. This angular displacement of the fluorescent molecule is dependent upon the rate and extent of rotation and consequently upon the size and shape of the rotating molecule as well as the viscosity of the solution and local environment of the fluorophore (Joseph R.Lakowicz, 2012). In general, the smaller the molecule the more rapidly it rotates and the more the emitted light is depolarized, resulting in lower anisotropy. Therefore upon binding to a larger molecule a more stable complex is formed which rotates slower in the solution and therefore reduces the degree of scrambling, producing a more polarized emission and higher anisotropy (Figure 3.1.1, right panel). An increase in the concentration of the larger molecule will result in an increase in anisotropy until a state of saturation is reached.

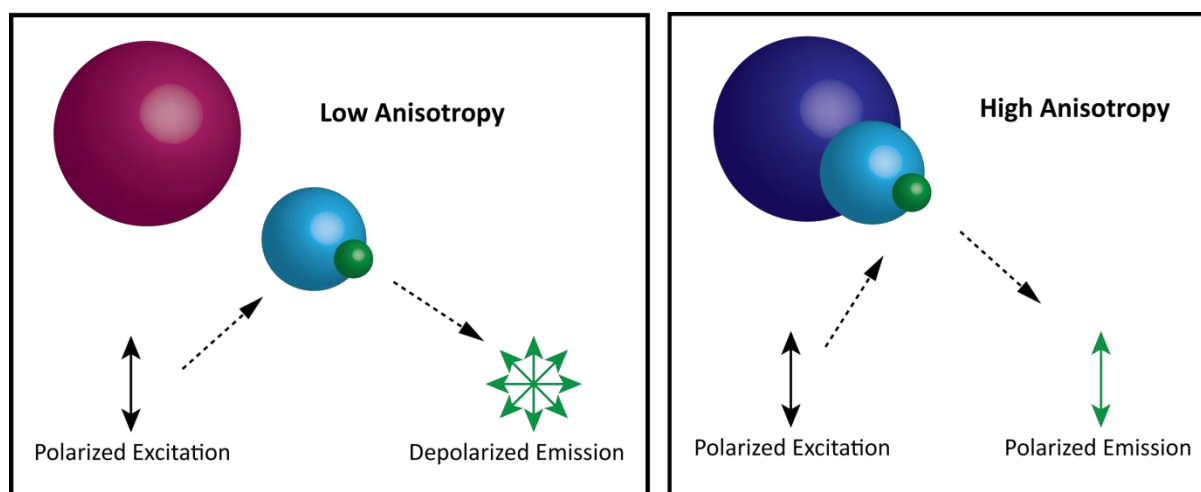


Figure 3.1.1 Schematic of the relationship between binding and anisotropy.

Left panel: upon excitation with polarized light a small (green), fluorescently labelled molecule (light blue) which does not bind to other available molecules (purple) produces a depolarized emission referred to as low anisotropy. Right panel: by binding to another molecule (dark blue) and forming a significantly larger complex, the light emitted after excitation remains comparatively polarized, referred to as high anisotropy.

To obtain measurements during a binding experiment a sample is excited with vertically polarized light, orientated parallel to the z-axis. The intensity of the emission is then detected via polarizers on two different planes. I_v describes the intensity of the emission orientated parallel to the direction of the polarized excitation (z-axis). I_h denotes the intensity of the emission perpendicular to the excitation (x-axis) (Figure 3.1.2). These values are used in the following equation to calculate the anisotropy (r).

$$r = \frac{I_v - I_h}{I_v + 2I_h}$$

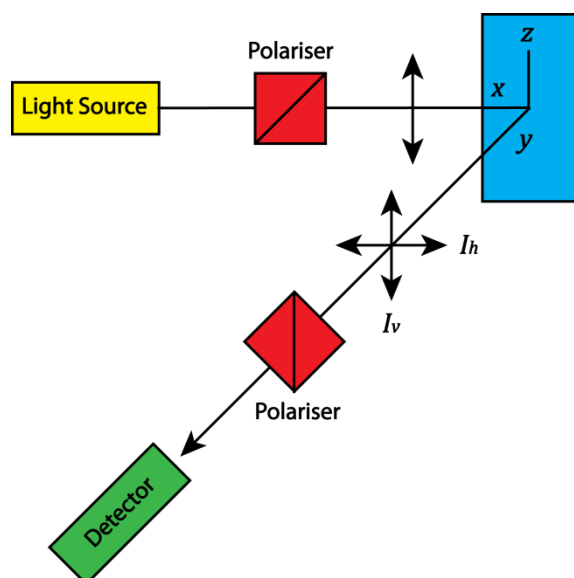


Figure 3.1.2 A diagram displaying the orientation of the emitted and excited polarized light.

To obtain FA measurements the sample is excited with polarized light orientated parallel to the z-axis. Light emitted from the sample is detected parallel to the z-axis (I_v) and perpendicular to the z-axis (I_h).

When measuring the interaction between two molecules a binding curve can be generated by maintaining the concentration of the smaller, fluorescently labelled molecule and titrating in the larger binding partner until maximal binding has been reached. The resulting curve can then be used to calculate the strength of the interaction in terms of binding affinity such as the association or dissociation constants, K_a and K_d respectively. The dissociation constant (K_d) is commonly used to describe the affinity between two molecules and in the case of FA is a reference to the concentration of the fluorescently labelled molecule, expressed in units of moles. K_d defines the equilibrium between bound and unbound states with respect to the reverse reaction i.e. the separation of the molecules. The smaller the K_d the stronger the binding, for instance the antibiotic rifampicin binds to *E. coli* RNAP with a very high affinity; $K_d = 130$ nM (Xu *et al.*, 2005). The measurement of the K_d provides a valuable reference for the comparison of similar interactions. Similarly to rifampicin, Gp2 is also a potent inhibitor of *E. coli* RNAP; however the affinity of Gp2 to RNAP has never been quantitatively assessed. Therefore this chapter aims to establish an FA assay to accurately measure the strength of binding between Gp2 and RNAP as well as inform on the stoichiometry of the interaction. Furthermore, this FA assay and the resulting K_d value can then be used to compare the binding of variant and mutant forms of the two proteins in future studies.

In early attempts to understand T7 infection and identify the processes which involve host-phage component interactions three *E. coli* mutant strains (TsnB, BR3 and Y49) were discovered that were non-permissive for T7 intracellular growth (Shanblatt and Nakada, 1982; Chamberlin,

1974;Studier, 1973). When RNAP was isolated from these strains it was found to be specifically resistant to T7 Gp2 inhibition, attributed to a lack of interaction (DeWyngaert and Hinkle, 1979;Shanblatt and Nakada, 1982). Complementation assays localised these mutations and therefore the proposed site of Gp2 binding to the C-terminus of the largest subunit of RNAP, now known as the β' jaw domain (residues 1149-1190) (Buchstein and Hinkle, 1982). These mutant *rpoC* genes were then sequenced, revealing that the resistant strains possess the charge reversal point mutations, E1158K (BR3 and Y49) and E1188K (TsnB) (Nechaev and Severinov, 1999), suggesting that E1158 and E1188 amino acid residues are necessary determinants for Gp2 to bind to RNAP. In fact, recent structural analysis has confirmed that the E1158 and E1188 amino acid residues reside within the interface of a complex formed between Gp2 and a fragment of the β' jaw (James *et al.*, 2012) (refer to section 4.8).

Previous work from this laboratory included a functional alanine mutagenesis screen of Gp2, which identified two closely located and surface exposed amino acid residues, R56 and R58 as being important for Gp2 inhibition (Camara *et al.*, 2010). In addition, sequence alignments revealed that R56 is conserved in all known Gp2 orthologs, whereas R58 is conserved in all but one case, where it is substituted for the also positively charged lysine (Figure 3.1.3).

Enterobacteria phage T7a	-----MSN-VNTGSLVDNKKFWATVESSEHSFEVPIYAETLDEALELAEWQYVPAGFEVTRVPCVAPK	64
Enterobacteria phage T7b	-----MSN-VNTGSLVDNKKFWATVESSEHSFEVPIYAETLDEALELAEWQYVPAGFEITRVRPCVAPK	64
Enterobacteria phage 13a	-----MSN-VNTGSLVDNKKFWATVESSEHSFEVPIYAETLDEALELAEWQYVPAGFEVTRVPCVAPK	64
Enterobacteria phage T7c	-----MSN-VNTGSLVDNKKFWATVESSEHSFEVPIYAETLDEALELQEWQYVPAGFEVTRVPCVAPK	64
Enterobacteria phage T7.1	-----MSN-VNTGSLVDNKKFWATVESSEHSFEVPIYAETLDEALELAEWQYVPAGFEVTRVPCVAPK	64
Yersinia phage YpP-Y	-----MSN-ANTGSLVDNKKFWATVESSEHSFEVPIYAETLDEALELAEWQYVPAGFEVTRVPCVAPK	64
Yersinia phage YpsP-G	-----MSN-ANTGSLVDNKKFWATVESSEHSFEVPIYAETLDEALELAEWQYVPAGFEVTRVPCVAPK	64
Yersinia pestis phage phiA1122	-----MSN-ANTGSLVDNKKFWATVESSEHSFEVPIYAETLDEALELAEWQYVPAGFEVTRVPCVAPK	64
Salmonella phage Vi06	--MVYPRATTWSFTINGYREVTDMLNVGT---DSKKFWATVESSEHSFEVPIYAETLDEALELAEWQYVQAGFEVTRVPCVTPK	80
Kluyvera phage Kvp1	-----MSN-LNKGSVLVNGDKKFFATVEGQGSFEVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	63
Yersinia phage Berlin	-----MSI-LNKGSVLVNGDKKFFATVEGQGSFEVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	63
Enterobacteria phage BA14	-----MST-LNKGSVLVNGDKKFFATVEGQGSFEVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	63
Yersinia phage Yep-phi	-----MSI-LNKGSVLVNGDKKFFATVEGQGSFEVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	63
Erwinia phage vB_EamP-L1	-----MKRFIAEVEVTQTHAFVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	51
Morganella phage MmP1	-----MSVTNDVNAIKE--NSGTLTETNKKFIVHLQGSSEIEVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	69
Klebsiella phage K11	-----MQKFVVTVETANASYELPVGAGSLDEALELAEWQYVPAGFAVTRIRPDRKV-	49
Enterobacteria phage K30	-----MQKFVVTVETANASYELPVGAGSLDEALELAEWQYVPAGFAVTRIRPDRKV-	49
Enterobacteria phage T3	-----MEQREIKYLLTMEGNTSFEVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	54
Salmonella phage phiSG-JL2	-----MEQREIKYLLTMEGNTSFEVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	54
Yersinia phage phiYeO3-12	MS-----QCNNGCQYAKDVGMENHSCANRCMYSEQKYLITIEGNTSFEVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	78
Enterobacteria phage EcoDS1	-----MSDV-----EKKYIIELEGRVKSFEVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	52
Enterobacteria phage K1F	-----MTEV-----EKKYIIELEGRVKSFEVPIYAETLDEALELAEWQYVPAGFAVTRIRPDRKV-	52
Pseudomonas phage phiIBB-PF7A	-----MKLFKAQVGSSEWVPIYAKDLDEALELAEWQYVPAGFAVTRIRPDRKV-	45
	: : : : * * : : * : : *	

Figure 3.1.3 Sequence alignment of Gp2 orthologs.

Alignment of the amino acid sequences of all known Gp2 orthologs in the single-letter amino acid codon. The length of the proteins is noted at the C-terminus in terms of amino acid residues. A * represents an absolutely conserved amino acid residue, : indicates conservation between groups of strongly similar properties and . indicates conservation between groups of weakly similar properties. The highly conserved R56 and R58 residues are highlighted in blue.

EMSA confirmed that the function of these two residues could be attributed to the binding to RNAP. Considering the nature of these arginines and the glutamic acid mutations in the β' jaw, *in*

vivo experiments were conducted to evaluate the contribution of the charge of the R56 and R58 residues in RNAP binding. Results revealed that charge reversal mutations (R to E) attenuated Gp2 function significantly more than charge removal (R to A) or charge conservation (R to K) (Camara *et al.*, 2010). This data implied that the Gp2-RNAP interaction involves, in part, an electrostatic component, likely to involve the E1158 and E1188 residues in the β' jaw. Consistent with this view, the interaction between Gp2 and RNAP has previously been shown to be salt sensitive (Hesselbach and Nakada, 1975).

In this study the point, charge reversal mutations in the RNAP β' subunit (E1158K and E1188K) and in Gp2 (R56E and R58E) were used as reciprocal negative controls in the FA assay as well as in downstream biochemical experiments.

3.2. Objectives

This chapter aims to establish an FA assay from which the binding affinity and the stoichiometry of the interaction between Gp2 and RNAP can be determined. The large size difference between Gp2 (7 kDa) and RNAP core (390 kDa) maximises the difference in signal between the unbound (depolarized light) and bound (polarized light) which is ideal for distinguishing between subtle variations of different interactions. Thus, this assay will also serve as a reliable method for assessing the binding between mutant forms of Gp2 and RNAP core, hereafter referred to as E, as well as alternative versions of the RNAP holoenzyme (E σ). The charge reversal mutations at β' subunit E1188K and Gp2 R56E will serve as negative controls for the setup of this assay and future experiments described in the following chapters.

The overall objective for this study was to establish a quantitative assay from which the binding affinity of E-Gp2 and variants thereof can be determined. This included the following aims:

- Over-express and purify the essential core components necessary for initial binding experiments including; recombinant *E. coli* E^{WT}, E^{E1188K}, σ^{70} , σ^{54} , Gp2^{WT} and Gp2^{R56E}.
- Fluorescently label Gp2^{WT} and Gp2^{R56E}.
- Establish assay conditions for conducting titration curves.
- Determine the binding affinity and the stoichiometry of the interaction between Gp2^{WT} and E^{WT}, E σ^{70} and E σ^{54} .

3.3. Over-expression and purification of recombinant proteins

To apply this technique to the study of Gp2 and E/ σ binding the relevant proteins first needed to be over-expressed and purified; these included recombinant *E. coli* E^{WT}, E^{E1188K}, σ ⁷⁰, σ ⁵⁴, Gp2^{WT} and Gp2^{R56E}. Preliminary findings had suggested that only a single mutation at either E1158/E1188 of the RNAP β' subunit or R56/R58 of Gp2 was required to entirely disrupt the interaction, therefore it was only necessary to construct one mutant version of each protein. The σ -factors and Gp2 proteins were expressed using the pET system which allows for the controlled, over-expression of recombinant proteins (refer to section 2.4.1). The pET28b⁺ and pET33b⁺ plasmids encode proteins fused to an N-terminal His₆-tag. In addition, the pET33b⁺ plasmid also includes the RRASV sequence upstream of the His₆-tag which enables ³²P-labelling by protein kinase A. Gp2^{WT} and variants were specifically encoded under the pET33b⁺ plasmid so the proteins could be ³²P-labelled and used as probes in gel-based protein-protein interaction experiments as well as the FA assay. The expression of the encoded proteins from these two plasmids is under the control of the T7 promoter and *lac* operator. The BL21/DE3 host cell strains harbour a single chromosomal copy of the T7 gene 1 which encodes T7 RNAP under the control of the *lac* promoter and *lac* operator. Addition of IPTG results in the displacement of the LacI repressor from both the *lac* promoter on the chromosome and the T7 promoter on the pET plasmid. Only then is the *E. coli* RNAP able to transcribe the T7 gene 1 from the chromosome which in turn can express the target gene encoded on the pET plasmid. The pVS10 plasmid encodes all five subunits of *E. coli* E with a C-terminal His₆-tag on the β' subunit under the control of a single T7 gene 10 promoter. Expression of E and mutant versions was performed using the Novablue (DE3) cell strain which also harbours the T7 gene 1 and is inducible with IPTG.

His₆-tags allow for the purification of the recombinant proteins by immobilised metal affinity chromatography (IMAC). The amino acid histidine (His) possesses an imidazole ring side-chain with a lone pair of electrons, therefore the His₆-tag binds strongly to divalent metal ions such as Ni²⁺ which can be immobilised on an agarose column. The recombinant protein can be eluted using an imidazole gradient which competes for the binding to Ni²⁺. A further purification step was used to separate E from its associated σ -factors. Heparin is a highly negatively charged glycosaminoglycan with a similar structure to DNA and RNA. These properties mean it can be exploited for the specific purification of nucleic acid-binding proteins. Therefore E and σ -factor proteins were expressed and purified using a two step fast protein liquid chromatography (FPLC); nickel (Figure 3.3.1 and Figure 3.3.3) and heparin affinity chromatography (Figure 3.3.2 and Figure 3.3.4) (refer to sections 2.4.4

and 2.4.6). Whereas Gp2 was purified using only nickel affinity chromatography (Figure 3.3.5). In some cases where multiple variants of Gp2 were required in the same set of assays the high-throughput method for expression and purification was used (refer to section 2.4.5). Examples of the SDS-PAGE analysis of the expression and purification of the recombinant proteins are shown in Figure 3.3.1 to Figure 3.3.5.

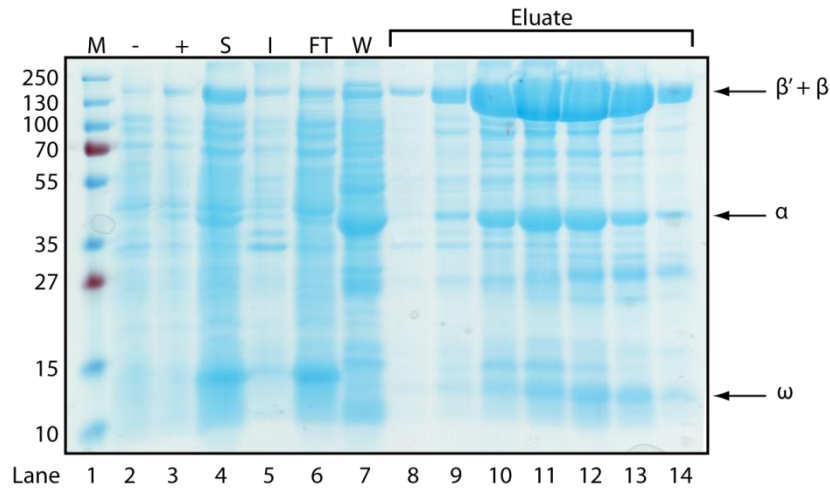


Figure 3.3.1 Nickel affinity chromatography of E^{WT} .

Image of a 10% (w/v) SDS-PAGE gel from the nickel affinity purification of recombinant *E. coli* E^{WT} . M = Marker, - = without IPTG induction, + = with IPTG induction, S = soluble fraction, I = insoluble fraction, FT = flow-through and W = low imidazole wash. The molecular weight of the protein markers is shown in kDa and the migratory positions of the subunits of E, β' , β , α , and ω are indicated with arrows.

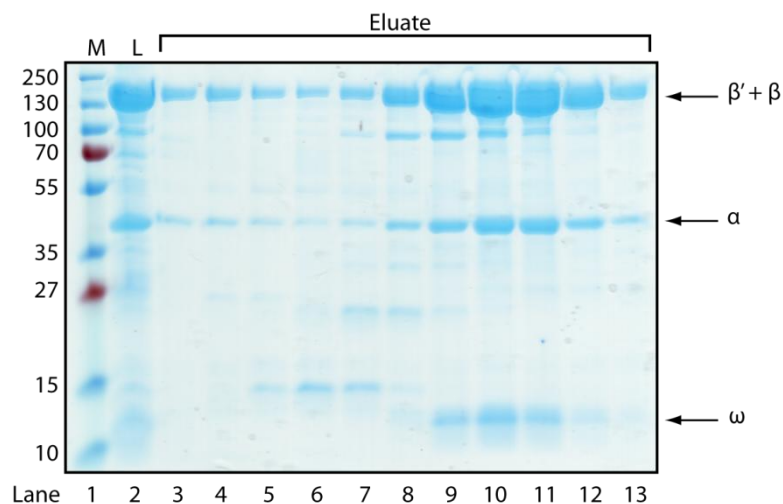


Figure 3.3.2 Heparin affinity chromatography of E^{WT} .

Image of a 10% SDS-PAGE gel from the heparin affinity purification of recombinant *E. coli* E^{WT} . M = Marker and L = load. The molecular weight of the protein markers is shown in kDa and the migratory positions of the subunits of E, β' , β , α , and ω are indicated with arrows.

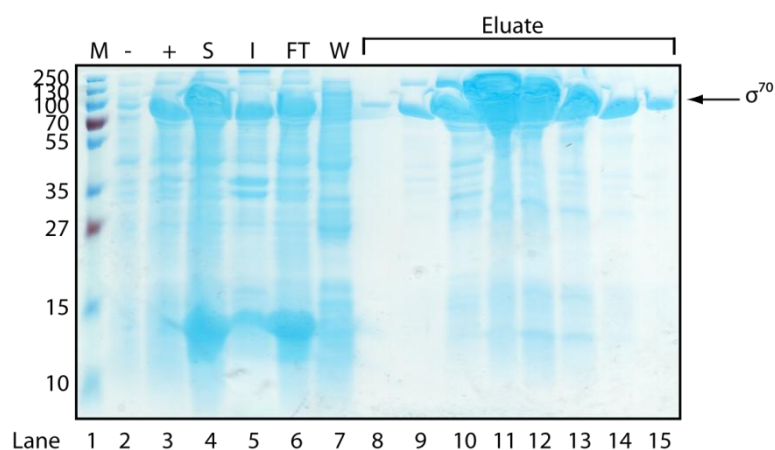


Figure 3.3.3 Nickel affinity chromatography of σ^{70} .

Image of a 12.5% (w/v) SDS-PAGE gel from the nickel affinity purification of recombinant *E. coli* σ^{70} . M = Marker, - = without IPTG induction, + = with IPTG induction, S = soluble fraction, I = insoluble fraction, FT = flow-through and W = low imidazole wash. The molecular weight of the protein markers is shown in kDa and the migratory position of σ^{70} is indicated with an arrow. Note, *E. coli* σ^{70} has an apparent MW of 87 kDa.

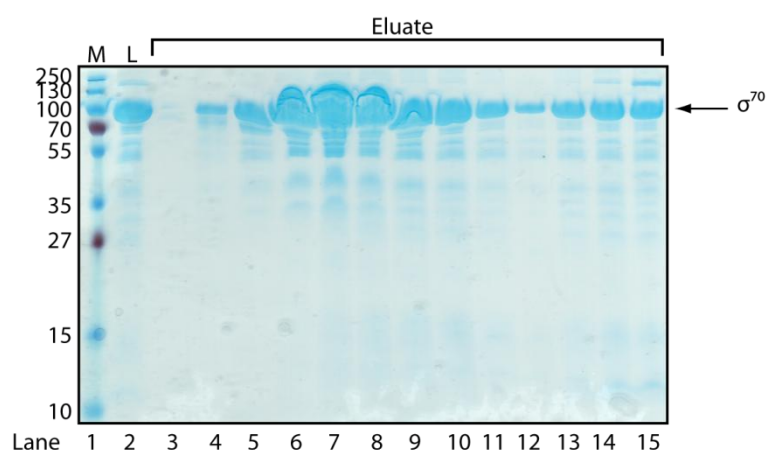


Figure 3.3.4 Heparin affinity chromatography of σ^{70} .

Image of a 12.5% (w/v) SDS-PAGE gel from the heparin affinity purification of recombinant *E. coli* σ^{70} . M = Marker, L = load. The molecular weight of the protein markers is shown in kDa and the migratory position of σ^{70} is indicated with an arrow. Note, *E. coli* σ^{70} has an apparent MW of 87 kDa.

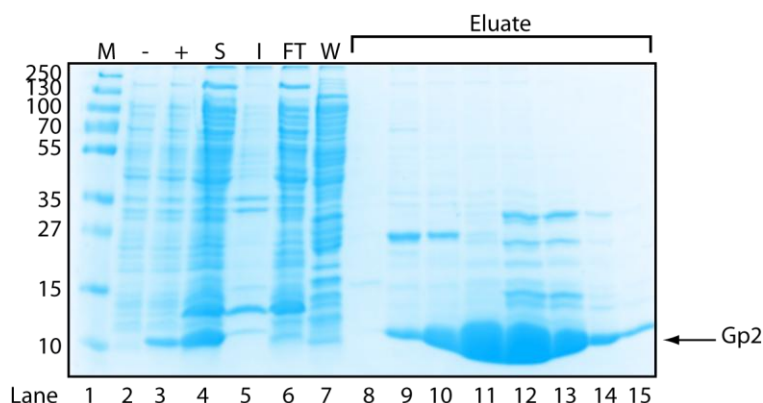


Figure 3.3.5 Nickel affinity chromatography of Gp2.

Image of a 12.5% + 15% (w/v) SDS-PAGE gel from the nickel affinity purification of recombinant T7 Gp2. M = Marker, - = without IPTG induction, + = with IPTG induction, S = soluble fraction, I = insoluble fraction, FT = flow-through, W = low imidazole wash. The molecular weight of the protein markers is shown in kDa and the migratory position of Gp2 is indicated with an arrow.

Eluate fractions containing proteins of >90% homogeneity as judged by densitometry (Figure 3.3.6) were dialysed into low-glycerol storage buffer. The protein concentration was then determined using the Bradford assay (refer to section 2.4.8).

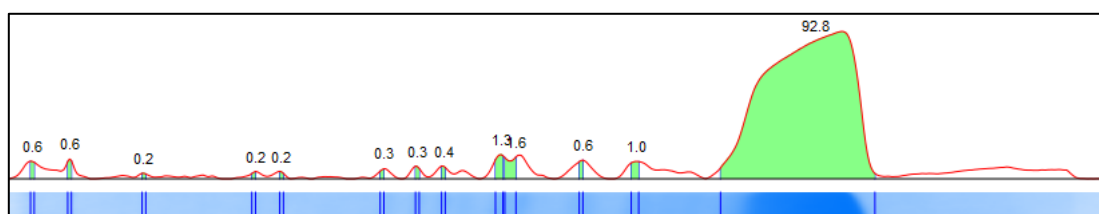


Figure 3.3.6 Lane profile of an SDS-PAGE gel from the nickel affinity chromatography of Gp2.

Graph showing the profile of lane 11 from the SDS-PAGE gel in Figure 3.3.5. Each peak represents a band in the gel, with the corresponding percentage of densitometry shown above.

3.4. Fluorescent labelling of Gp2

To fluorescently label Gp2 a suitable fluorophore needed to be selected. Alexa-488 is only 700 daltons in size, has strong absorption, excellent photostability, is pH insensitive between pH 4 and 10 and has an exceptional fluorescence output per protein conjugate (Panchuk-Voloshina et al., 1999). As a result of these qualities it is frequently used in the study of protein-protein interactions by FA. Alexa-488 is a thiol or sulphydryl reactive probe and therefore labels proteins at their cysteine

residues. Fortunately Gp2 only contains one native cysteine at position 60 which is surface exposed at the C-terminus. Alanine mutagenesis at C60 showed this residue to be relatively unimportant for function (Camara *et al.*, 2010), making it an ideal position for modification. Gp2 was conjugated with Alexa-488 as described in section 2.7.1 and re-purified by nickel affinity chromatography to remove excess un-conjugated fluorophore. Gp2 and Gp2^{Alexa488} were then ³²P-labelled and used in native gel binding assays to establish whether the modification affects the ability of Gp2 to bind to Eσ⁷⁰ relative to an unlabelled version; compare Figure 3.4.1 lanes 2, 3 and 4 to lanes 6, 7 and 8. The Alexa-488 modified version of ³²P-Gp2 bound Eσ⁷⁰ as well as the non-fluorescently modified ³²P-Gp2, as judged by the relative proportions of complex formation. Therefore, the Alexa-488 modification of Gp2 does not detectably affect its ability to bind to Eσ⁷⁰.

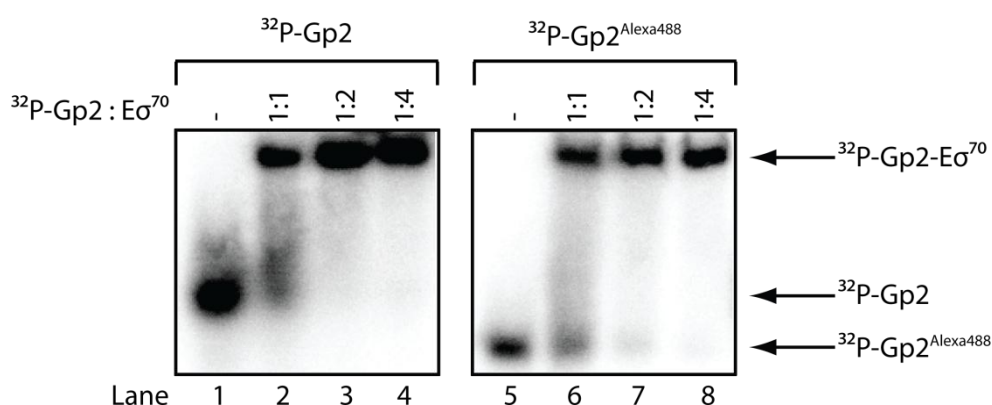


Figure 3.4.1 Native gel binding assay demonstrating the binding of Gp2/Gp2^{Alexa488} to Eσ⁷⁰.

An autoradiograph of a 4.5% (w/v) polyacrylamide native gel showing the binding of ³²P-Gp2 (left panel) and ³²P-Gp2^{Alexa488} (right panel) to Eσ⁷⁰. The molar ratios of Eσ⁷⁰ with respect to Gp2 are indicated at the top of each lane. The migratory positions of free versions of ³²P-Gp2 (lane 1 and 5) and ³²P-Gp2-Eσ⁷⁰ complexes (lanes 2-4 and 6-8) are indicated with arrows.

3.5. FA experimental setup

FA experiments were conducted using a Fluoromax-3 spectrofluorometer as described in section 2.7.2. To begin, the ideal concentration of Gp2^{Alexa488} was established; it was determined that 50 nM Gp2^{Alexa488} gave a sufficient fluorescence signal above the background. The Gp2^{Alexa488} was diluted in 400 μl of low glycerol storage buffer as this is considered a sufficient volume from which to take accurate measurements as well as allowing for an increase in volume upon the addition of

protein. Low glycerol storage buffer was used as it only contains 5% glycerol which does not significantly affect the viscosity of the solution and therefore impair the movement of the proteins. All experiments were conducted at 37°C as this was consistent with *in vitro* transcription assays. Measurements were recorded using a slit size of 0.5 nm and excitation and emission wavelengths of 495 nm and 519 nm respectively as per the requirements for Alexa-488. The G-factor is a constant which defines the ratio of the sensitivity of detection of fluorescence intensities between the parallel and perpendicular components of the emitted light. The G-factor for Gp2^{Alexa488} under these experimental conditions was initially measured and determined to be 0.63; this value was maintained for all future experiments. Binding curves were generated by titrating in either E or reconstituted Eσ formed with a 1:4 molar excess of σ-factor. The anisotropy was measured 5 times in immediate succession. The resulting curve was fitted and the K_d calculated from the tight binding equation shown below:

$$A = A_D + (A_{DE} - A_D) \times \left(\frac{([E]_0 + [Gp2]_T) - \sqrt{([E]_0 + [Gp2]_T)^2 - 4[E]_0[Gp2]_T}}{2[Gp2]_T} \right)$$

Where A is the observed anisotropy, A_D is the anisotropy of free Gp2^{Alexa488}, A_{DE} is the anisotropy of bound Gp2^{Alexa488}, K_d is the equilibrium dissociation constant, [Gp2]_T is the total concentration of Gp2^{Alexa488} and [E]₀ is the concentration of RNAP core (Bellamy and Baldwin, 2001).

3.6. Effect of the σ-factor on Gp2 binding to E

An initial titration curve was generated to assess the binding between Gp2^{Alexa488} and E^{WT} (Figure 3.6.1). An equimolar amount of E was required to saturate the 50 nM of Gp2^{Alexa488} signal which represents a 1:1 binding stoichiometry. The K_d in this instance was calculated to be 0.53 nM which represents a very tight binding interaction. It was previously described that Gp2 has no effect on the activity of E suggesting that Gp2 can only inhibit in the context of Eσ (DeWyngaert and Hinkle, 1979) (also refer to chapters 4 and 5). To investigate the potential requirement for the σ-factor in Gp2 function, the binding of Gp2^{Alexa488} to Eσ⁷⁰ was measured to test whether this strengthened the binding interaction. Interestingly, Gp2^{Alexa488} bound to Eσ⁷⁰ using the same 1:1 molar stoichiometry and with the same affinity as E, K_d = 0.59 nM (Figure 3.6.1). As Gp2 is not an anti-σ, i.e. does not simply disrupt holoenzyme formation, this result implies that Gp2 does not require the σ-factor directly or indirectly, for initial binding to E. Gp2 is known to have no significant effect on transcription from the major alternative holoenzyme, Eσ⁵⁴ (Wigneshweraraj *et al.*, 2004). As there is a low sequence and structural similarity between σ⁷⁰ and σ⁵⁴ it is possible that the

different E- σ interaction may interfere with Gp2 binding. Therefore this lack of inhibition may simply be due to a relatively weak interaction between Gp2 and E σ^{54} compared to that of E σ^{70} . Surprisingly Gp2^{Alexa488} actually bound to E σ^{54} with a 1:1 molar stoichiometry and with a K_d of 0.78 nM which is comparable to that of E and E σ^{70} (Figure 3.6.1). Thus the initial binding of Gp2 to E occurs completely independently of the σ -factor.

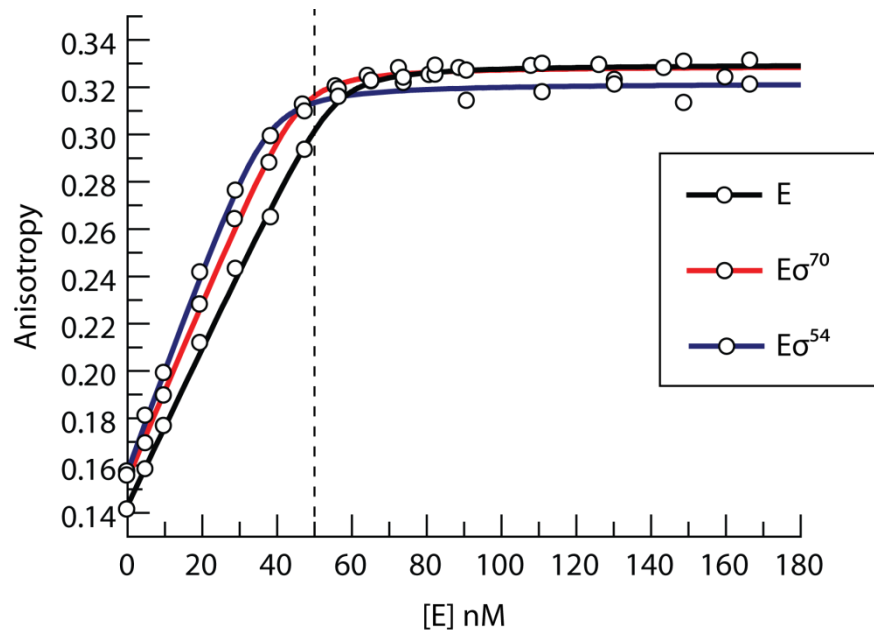


Figure 3.6.1 Graph displaying the binding of Gp2 to E, E σ^{70} and E σ^{54} as measured by FA. Gp2^{Alexa488} (50 nM) binding to *E. coli* E (black line), E σ^{70} (red line) and E σ^{54} (blue) is represented by anisotropy which is displayed as a function of the concentration of E. The dotted line indicates the concentration (50 nM) at which binding becomes saturated.

3.7. Important determinants for Gp2-E/E σ binding

To test whether the Gp2 interactions observed with E, E σ^{70} and E σ^{54} are specific, the same binding curves were performed with the reciprocal controls; E^{E118K} and Gp2^{R56E-Alexa488} (Figure 3.7.1 and Figure 3.7.2). As expected, Gp2 displayed no detectable binding to E, in either the context of Gp2^{Alexa488} + E^{E118K} σ^{70} or Gp2^{R56E-Alexa488} + E σ^{70} . These results confirm that E1188 of the β' subunit of E and R56 of Gp2 are essential determinants for the initial binding of Gp2 to E. Therefore the defective mutants E^{E118K} and Gp2^{R56E} can serve as faithful negative controls in further functional studies.

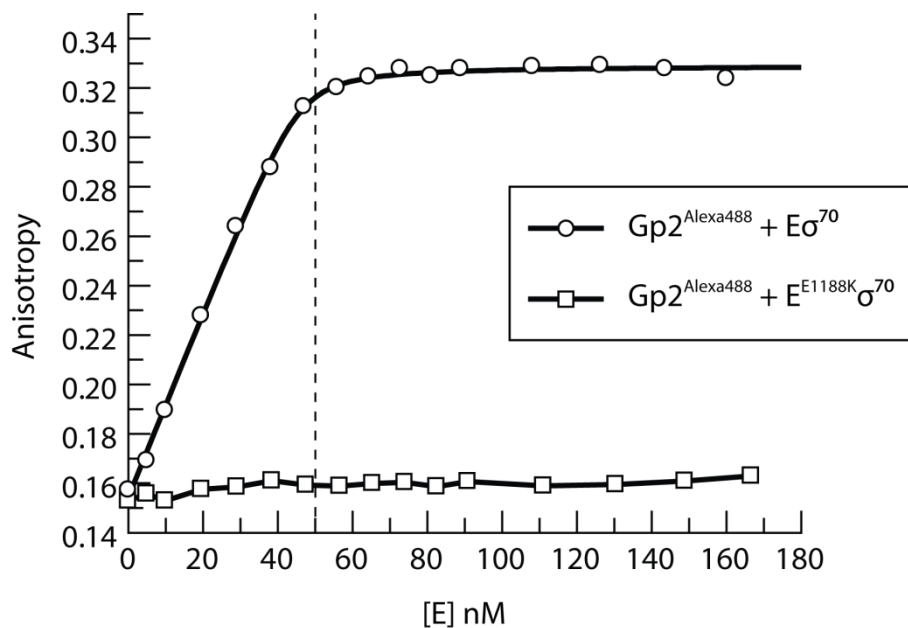


Figure 3.7.1 Graph displaying binding of Gp2 to Eσ⁷⁰ and E^{E1188K}σ⁷⁰ as measured by FA.

Gp2^{Alexa488} (50 nM) binding to *E. coli* Eσ⁷⁰ (circles) and E^{E1188K}σ⁷⁰ (squares) is represented by anisotropy which is displayed as a function of the concentration of E. The dotted line indicates the concentration (50 nM) at which binding becomes saturated.

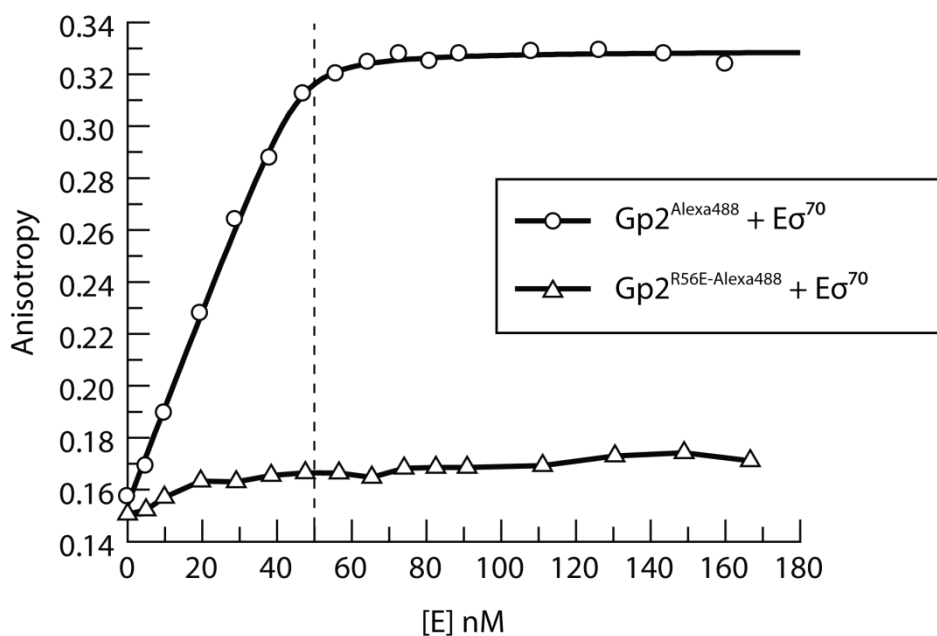


Figure 3.7.2 Graph displaying binding of Gp2 to Eσ⁷⁰ and Gp2^{R56E} to Eσ⁷⁰ as measured by FA.

Gp2^{Alexa488} (50 nM) binding to Eσ⁷⁰ (circles) and Gp2^{R56E}-Alexa488 binding to Eσ⁷⁰ (triangles) is represented by anisotropy which is displayed as a function of the concentration of E. The dotted line indicates the concentration (50 nM) at which binding becomes saturated.

3.8. Conclusions

The main objective of this study was to establish a reliable and quantitative, solution based protein-protein binding assay for detecting the interaction between Gp2 and RNAP. In this case, fluorescence anisotropy proved to be an ideal technique as Gp2 could be easily modified without detectably affecting its binding capability. In addition, the large size difference between Gp2 and RNAP provides a broad range in the change of anisotropy which would be advantageous for future analysis of mutant interactions (refer to chapters 4 and 5).

Gp2 binds to the RNAP core, E, using a 1:1 molar stoichiometry with a sub-nanomolar affinity; a K_d of 0.53 nM is indicative of a very tight binding inhibitor. Interestingly, the affinity of Gp2 to E is independent of the presence of the σ -factor; Gp2 binds to $E\sigma^{70}$ with a K_d of 0.59 nM. It remains possible that the overall interface between Gp2 and E could be altered upon σ -factor binding but any additional contacts are likely to be weak and may not detectably affect the K_d . Furthermore, these findings do not exclude the possibility that the affinity between Gp2 and E is affected when the holoenzyme engages with the promoter DNA to form the RPo or at any subsequent steps in attempted RPo formation. It is conceivable that as the $E\sigma$ -promoter complex progresses through a series of conformational changes the affinity of Gp2 for E may change; for instance it may weaken as the DNA loads into the dw-DBC or it might tighten as the flexible mobile domains of the E are repositioned. The effect of the promoter DNA on Gp2 binding may be experimentally addressed by introducing variant duplexed oligonucleotides into the $Gp2^{Alexa488}$ - $E\sigma^{70}$ complex which mimic the entry and conformational change of the DNA during RPo formation. An even more surprising result is that despite Gp2 being unable to inhibit transcription by $E\sigma^{54}$ (Wigneshweraraj *et al.*, 2004), it seems it is still able to bind to this holoenzyme with a K_d of 0.78 which is not significantly different to that of E or $E\sigma^{70}$. This suggests that the additional ATP consuming, activation steps required during RPo formation in $E\sigma^{54}$ -dependent transcription may either actively displace Gp2 from its binding site or render it functionally inactive by some other unknown mechanism.

It is also possible that as the strength of binding between Gp2 and E is already very tight, a potential further increase in affinity upon $E\sigma$ formation may be undetectable under these experimental conditions. In fact, the binding affinity can only be accurately quantified when the K_d is calculated to be greater than the starting concentration of the fluorescently labelled protein. The K_d values calculated from the experiments performed in this study were not greater than the concentration of $Gp2^{Alexa488}$ which is a major weakness of this assay. Ideally, a significantly lower concentration of $Gp2^{Alexa488}$ should be used, however considering the range of the current estimated

K_d , the fluorescent signal required would be undetectable. One way to counteract this problem of tight binding proteins is to conduct competition assays to produce a dissociation curve which can be fitted to determine the binding affinity more accurately. Alternatively the binding reactions could be performed in higher salt conditions which may compromise the interactions and increase the K_d to a broader and more detectable range.

FA experiments performed with the point mutants E^{E1188K} and Gp2^{R56E-Alexa488} demonstrated that these charge reversal mutations render these proteins completely defective for binding. Thus it was confirmed that the E1188 and R56 residues are indeed essential determinants for binding and by extension the E1158 and R58 residues are also likely to have the same role. It is surprising and perhaps unlikely that the electrostatic interaction between just two pairs of amino acid residues could be responsible for such a tight binding affinity. Therefore, although the initial binding of Gp2 to E is very tight and entirely dependent on the glutamic acid and arginine residues of E and Gp2, these FA measurements do not provide any direct information on the extent of the surface interface (refer to chapter 4). It is possible that after the initial docking, via E1158, E1188, R56 and R58 the proteins form a more extensive binding interface involving other amino acid residues and perhaps even induces conformational rearrangements in the E/E σ which could further secure Gp2 in a tightly bound position. In fact, recent work in this laboratory has explored the possibility that the interface between Gp2 and RNAP may be more extensive than previously described. It seems that substitutions at amino acid positions E24 and F27 in the loop interconnecting the β_1 and β_2 strands may form an auxiliary interaction between Gp2 and RNAP when the primary interaction is compromised. Thus, this additional contact site may contribute to the high binding affinity between Gp2 and RNAP (Shadrin *et al.*, 2012).

In this chapter it has been shown how FA has been established for the accurate and reliable quantification of the binding affinity of Gp2 to RNAP and variant versions thereof. Such an approach could also have implications for pharmaceutical and synthetic biology applications. For instance, rifampicin is currently the only clinically used RNAP targeting antibiotic and is the first line treatment of tuberculosis. The K_d of rifampicin binding to RNAP has been calculated to be 130 nM. Therefore Gp2 actually represents an inhibitor of RNAP with potentially a ~250 fold tighter binding affinity than the available therapeutic alternative.

CHAPTER 4

INHIBITION OF *E. COLI* RNAP BY GP2: THE ROLE OF THE NEGATIVELY CHARGED STRIP OF AMINO ACID RESIDUES IN GP2

4.1. Introduction

Previously an atomic resolution solution structure of Gp2 was determined by NMR. This structure revealed that the protein adopts a $\beta_1\beta_2\alpha_1\beta_3$ topology comprised of three β strands packed against an α helix of which the central β_2 is flanked by the anti-parallel β_1 and β_3 strands (Camara *et al.*, 2010). One prominent feature of this structure of Gp2 is the charge separation of amino acids possessing negative side chains at opposing sides of the protein to the positively charged binding determinants R56 and R58 (Figure 4.1.1).

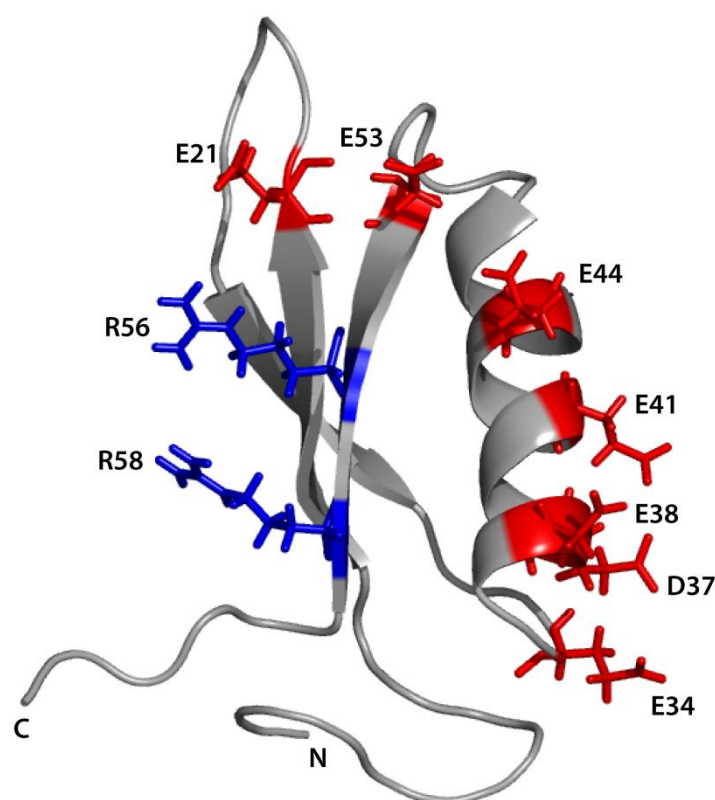


Figure 4.1.1 Charge separation of Gp2 protein.

Ribbon representative of the NMR structure of Gp2 with the positively charged binding determinants, R56 and R58 coloured blue and the negatively charged aspartic and glutamic acid residues (E21, E34, D37, E38, E41, E44 and E53) coloured red.

The negatively charged aspartic and glutamic acid residues E21, E34, D37, E38, E41, E44 and E53 form a contiguous negatively charged strip of surface exposed residues running from the β_1 - β_2

loop down the α -helix. The negative charge of side chains of these seven amino acid residues is partially conserved in orthologs of Gp2 (Figure 4.1.2) suggesting that this negatively charged strip (NCS) is important for function. The role of this NCS in Gp2 function is unknown and therefore the subject of this chapter.

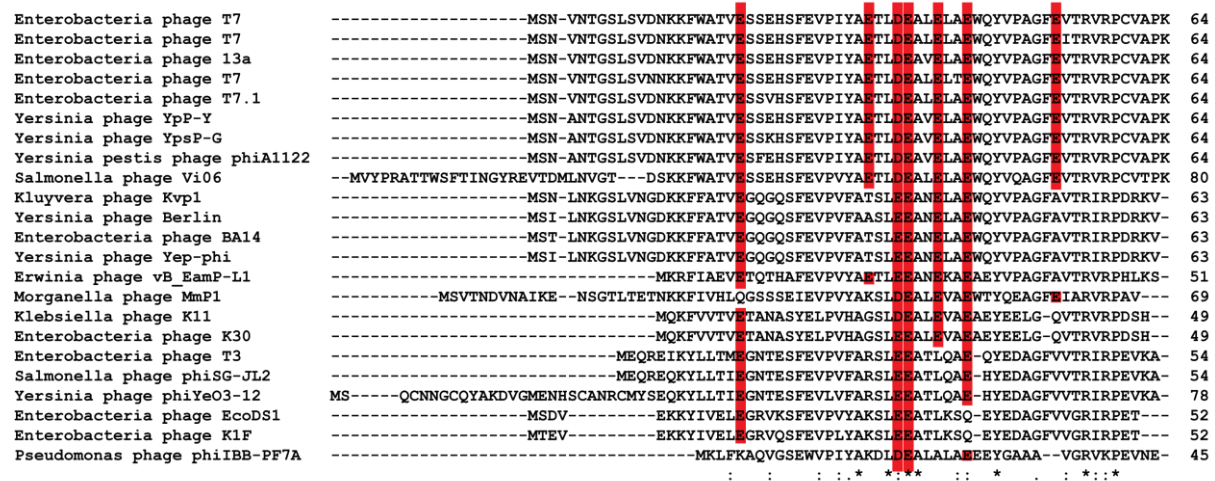


Figure 4.1.2 Sequence alignment of Gp2 orthologs.

Alignment of the amino acid sequences all known Gp2 orthologs in the single-letter amino acid codon. The length of the proteins is noted at the C-terminus in terms of amino acid residues. A * represents an absolutely conserved amino acid residue, : indicates conservation between groups of strongly similar properties and . indicates conservation between groups of weakly similar properties. The NCS residues are highlighted in red.

An initial *in vitro* transcription experiment was conducted to test the inhibitory activity of single alanine mutations within the NCS. The results revealed that the individual point mutations at each of the seven residues had no detectable effect on Gp2 inhibitory activity (Figure 4.1.3) (Sheppard *et al.*, 2011). This data suggests that either the negative charge of the residues in the NCS does not contribute significantly to Gp2 function or because of their high level of conservation as a whole it is more feasible that the negatively charged residues within the NCS may act collectively. As the NCS is positioned at the opposite side of the protein to the R56 and R58 residues which are required for binding to the β' jaw, it is conceivable that the NCS could project into the downstream DNA-binding channel (dw-DBC) of RNAP to inhibit its activity by an unknown mechanism. Investigating the role of the NCS will provide crucial insights into the overall mechanism of action of Gp2.

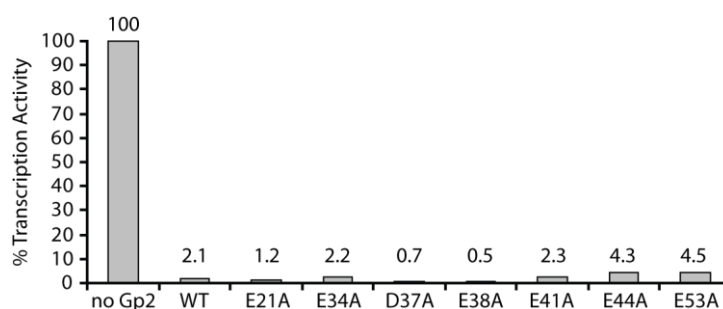


Figure 4.1.3 Bar chart displaying the Inhibitory activity of Gp2 NCS point mutants.

Results from an *in vitro* transcription assay showing the activity of *E. coli* $\text{E}\sigma^{70}$ from the *lacUV5* promoter in the absence of Gp2 and in the presence of Gp2^{WT} or mutant versions of Gp2 with single alanine substitutions in the NCS. Transcription activity is shown in percentage above each bar (Sheppard *et al.*, 2011).

4.2. Objectives

The overall objective of this chapter is to explore the role that the NCS of Gp2 may play in the inhibition of RNAP. This objective was approached by investigating how the removal of the NCS may affect Gp2 function. To this end, four mutant versions of Gp2 were created by site-directed mutagenesis, containing a series of alanine and serine substitutions within the NCS. In collaboration with Prof. Steve Matthews at the centre of structural biology, Imperial College London, the ability of each mutant protein to fold correctly was first tested, then their binding affinities for $\text{E}\sigma^{70}$ were quantified and their inhibitory activities were determined both *in vitro* and *in vivo*. This data provides further insights into the overall molecular mechanism employed by Gp2 during inhibition of RNAP. This study was addressed with the following specific aims:

- Over-express and purify mutant versions of Gp2 containing two, five, six and seven alanine/serine substitutions within the NCS.
- Determine whether the mutations in the NCS of Gp2 compromise the structural integrity of the protein by 1-D NMR.
- Establish if and with what affinity the NCS mutant proteins can bind to $\text{E}\sigma^{70}$ using the previously established FA assay (refer to chapter 3).
- Determine the activity of each NCS mutant protein.
- Explore the mechanism by which Gp2 inhibits the RNAP and what role the NCS may play.

4.3. Production of Gp2 NCS mutant proteins

Following the single alanine mutagenesis of the NCS, an additional mutant version of Gp2 was constructed possessing two mutations (D37A and E38S) within the NCS (Gp2^{Mut2}) at the most highly conserved residues in the Gp2 orthologs (Figure 4.1.2). To further perturb the charge of the NCS three further mutant versions of Gp2 were created in which five (Gp2^{Mut5}), six (Gp2^{Mut6}) and all seven (Gp2^{Mut7}) residues in the NCS were substituted for alternating alanine and serine residues (Figure 4.3.1). These amino acids were chosen as substitutes due to the relatively small size of their side chains, which are less likely to distort the structure of the α -helix.

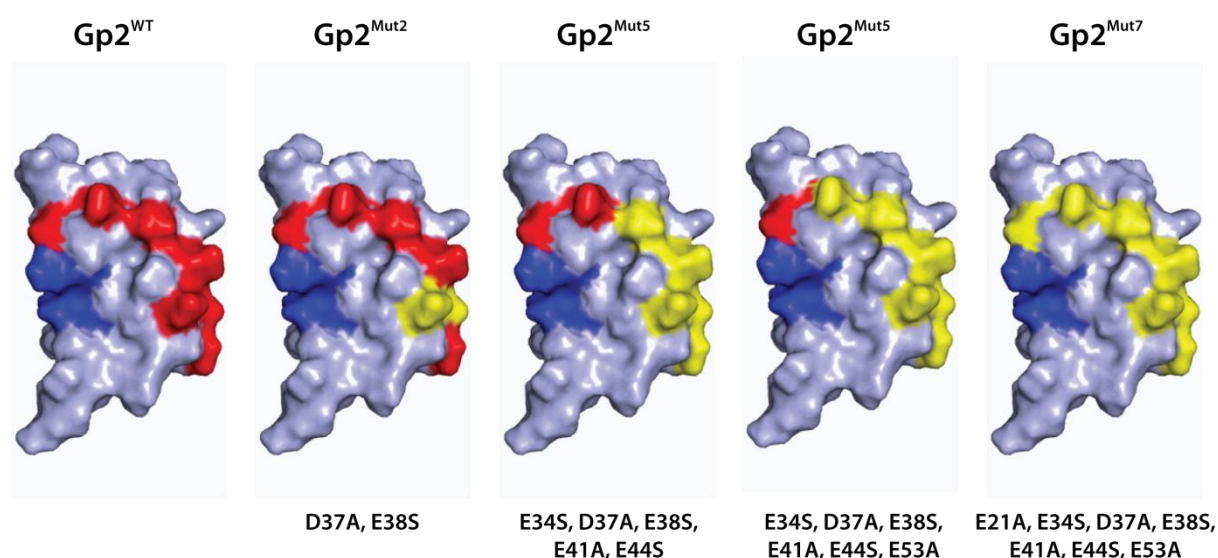


Figure 4.3.1 Structure of Gp2 with NCS mutations highlighted.

Space filled representation of Gp2 with R56 and R58 residues coloured blue and NCS residues coloured red. The NCS alanine and serine substitutions are coloured yellow and their positions described beneath each structure.

The four mutant proteins were over-expressed from the pET28b⁺ and pET33b⁺ plasmids and purified to >90% homogeneity by nickel affinity chromatography (refer to sections 2.4.4 and 3.3).

4.4. 1-D NMR spectra of Gp2 NCS mutant proteins

Since the NCS mutant proteins are the product of a significant degree of modification (>10% for Gp2^{Mut7}) it was a major concern that the extent of the substitutions may compromise the structural integrity of Gp2. To check whether the mutant proteins had retained their overall structure the 1-D NMR spectra of the mutants was measured and compared to that of Gp2^{WT} (Figure 4.4.1). These experiments were conducted in collaboration with Prof. Steve Matthews and colleagues at the Centre of Structural biology, Imperial College London. Resonances beneath 0.5 ppm represent methyl and methylene groups buried in the hydrophobic core of the protein. In the Gp2 spectra the dispersion of resonances at -0.126 ppm to 0.045 ppm (Figure 4.4.1, enlarged box) specifically correspond to the H^{γ1} and H^{γ2} protons in the methyl groups of V20. Residue V20 forms part of the hydrophobic core of the protein which if disrupted significantly effects the folding of Gp2 (Camara *et al.*, 2010). The frequencies of the resonances from these protons can be shifted from the electronic ring currents of neighbouring aromatic side chains (ring current shift). Therefore a change in the local environment of V20 will be represented by this spectrum. For instance F16 is buried within the hydrophobic core of the protein, therefore a substitution to alanine results in a loss of resonances below 0.5 ppm, signifying a disruption to the structure, which results in a significant reduction in inhibitory activity (Camara *et al.*, 2010). However, the frequencies of these resonances are retained in the spectra of all the NCS mutant proteins. Similarly the amide and aromatic side chains of the NCS mutants display a dispersion of resonances between 6.5 ppm to 9 ppm which are comparable to that of Gp2^{WT}.

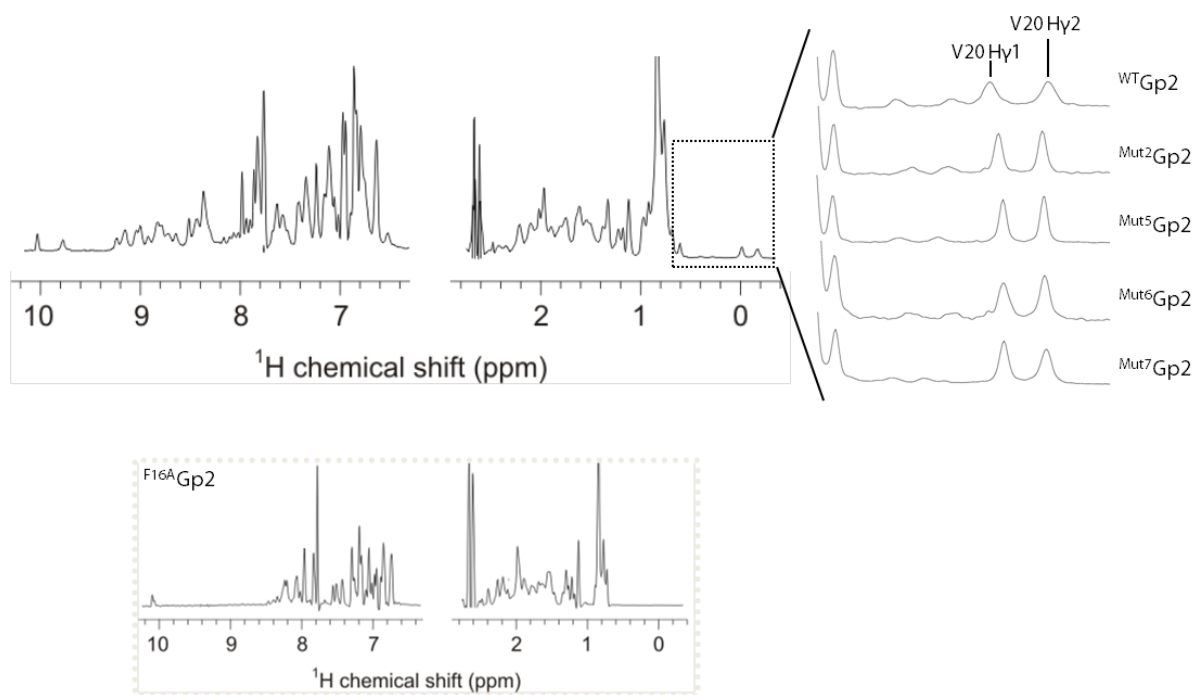


Figure 4.4.1 1-D NMR spectra of Gp2 and NCS mutants.

Detail of proton (^1H)NMR spectra from Gp2^{WT} and NCS mutants. The chemical shift from the amide and aromatic protons are shown on the left side whereas the methyl and methylene groups are shown on the right. The enlarged region shows the ring current-shifted methyl groups from residue V20. The inserted box is an (^1H)NMR spectra from Gp2^{F16A} in which the structure is significantly disrupted due to mutation (Sheppard *et al.*, 2011).

This data demonstrates that the structural features of Gp2^{WT} are preserved in all the mutant proteins, implying that the substitutions within the NCS do not significantly perturb the correct folding of the protein. Consequently the mutant Gp2 proteins could be confidently used in assays to investigate Gp2 binding to and inhibition of RNAP.

4.5. Binding of NCS mutants to E σ ⁷⁰

To test whether the residues in the NCS contributed to the ability of Gp2 to bind to E σ ⁷⁰ the strength of the interaction was quantified by FA. Each Gp2 mutant protein along with Gp2^{WT} was labelled with Alexa-488, hereafter referred to as Gp2* and used in a titration assay with E σ ⁷⁰ (refer to sections 2.7.2 and 3.5). As expected the anisotropy of Gp2^{WT} increased in a dose dependent manner upon increase in concentration of E σ ⁷⁰, whereas no binding was detected for Gp2^{R56E*} + E σ ⁷⁰ or Gp2^{WT*} + E^{E1188K} σ ⁷⁰. In this experiment the K_d of Gp2^{WT*} was calculated to be 0.35 nM. The same

titration assays were then performed for each of the NCS mutants (Figure 4.5.1) and the K_d values were calculated from the fit of each of these binding curves (Table 4.5.1).

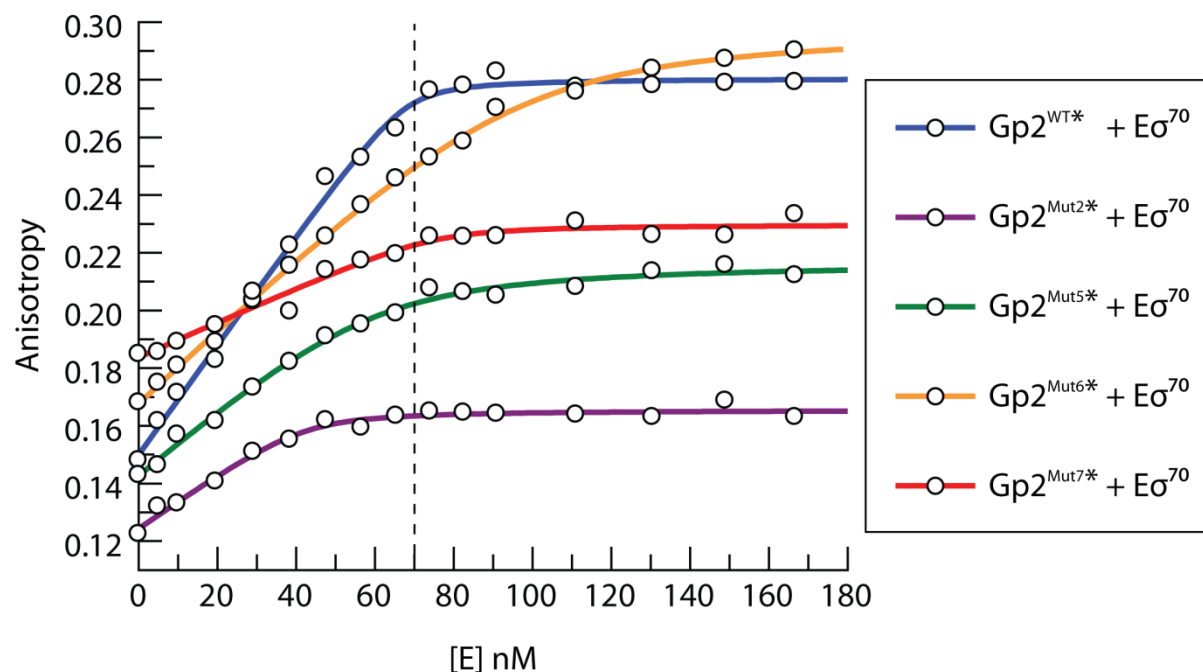


Figure 4.5.1 Graph displaying the binding of Gp2^{WT*} and NCS mutant versions to Eσ⁷⁰ measured by FA. Gp2^{WT*} (blue), Gp2^{Mut2*} (purple), Gp2^{Mut5*} (green), Gp2^{Mut6*} (orange) and Gp2^{Mut7*} (red) binding to Eσ⁷⁰ is represented by anisotropy which is displayed as a function of the concentration of E. The dotted line indicates the concentration at which Gp2^{WT*} binding becomes saturated.

The shape of the binding curves notably differs between Gp2^{WT} and each of the NCS mutants. In particular there is a significant difference in amplitude resulting in lower bound anisotropy values (especially in Gp2^{Mut2*}, Gp2^{Mut5*} and Gp2^{Mut7*}). This difference in amplitude relates to a change in the rotational diffusion which can be caused by a change in the lifetime of the fluorophore or a change to its local environment. This implies that either the individual context of the Alexa-488 fluorophore in each of the Gp2 proteins changes due to the direct effect of the mutations or that the mutations effect the overall binding interface between each of the Gp2 proteins and Eσ⁷⁰ which ultimately changes the Alexa-488 surroundings.

Gp2	K _d (nM)	SE
WT	0.35	0.09
Mut2	2.20	0.54
Mut5	3.20	1.93
Mut6	3.95	1.43
Mut7	2.90	1.71

Table 4.5.1 K_d values of NCS mutants.

The average calculated K_d value representing the binding affinity of each Gp2 protein to Eσ⁷⁰ was determined by FA.

The binding affinities, represented by the K_d values of each of the mutant proteins, are approximately 6-11 fold lower than that of Gp2^{WT}. Although the variation of these K_d values and the shape of the binding curves suggests a slight change in the strength of the interaction, these binding affinities are all in the low nanomolar range and therefore still indicative of tight binding proteins. Furthermore, the saturation of NCS mutant protein binding to Eσ⁷⁰ occurs at approximately the same concentration as that of Gp2^{WT} (represented by the dashed line in Figure 4.5.1), indicative of a 1:1 binding stoichiometry. From these results it was concluded that the mutations within the NCS have only a marginal effect on the ability of the protein to bind to Eσ⁷⁰. This is in marked contrast to a point mutation at R56, which under the same experimental conditions completely abolished the binding ability of Gp2 (refer to Figure 3.7.2) and therefore along with R58 remains the main determinants for the initial binding to RNAP.

4.6. Inhibition of transcription by NCS mutant proteins

Since the primary function of the NCS does not seem to be for binding to Eσ⁷⁰, the inhibitory activity of the mutant proteins was assessed in an *in vitro* transcription assay. The NCS mutant proteins were incubated with Eσ⁷⁰ (refer to section 2.6.5) at 1:1, 1:2 and 1:4 ratios at which Gp2^{WT} is known to almost completely abolish transcription (Figure 4.6.1 lanes 2-4). Interestingly, in contrast to Gp2^{WT} the NCS mutants displayed gradually reduced inhibitory activity, correlating with the progressive removal of the negative charge of the side chains. At equimolar ratios of Gp2 to Eσ⁷⁰, transcription activity was reduced to 1.2%, 53% and 97% for Gp2^{Mut2}, Gp2^{Mut5} and Gp2^{Mut6}

respectively. Whereas Gp2^{Mut7} has no effect on transcription activity at this molar ratio; compare lanes 2, 5, 8, 11 and 14. However at a 1:4 molar ratio of Gp2^{Mut7} over Eσ⁷⁰ there is a notable decrease in transcription (14%) (Figure 4.6.1, lane 16) suggesting that if present at high enough concentrations there still exists slight inhibitory activity. This inhibitory activity may be attributed, in part to steric hindrance; however such high ratios especially in the artificial environment of the experiment are unlikely to be representative of native conditions.

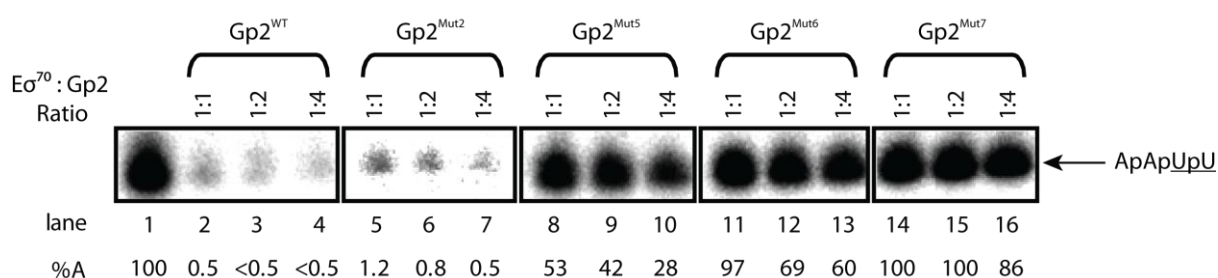


Figure 4.6.1 Inhibitory activity of Gp2 NCS mutants.

An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing the synthesis of the ApApUpU transcript as indicated by the arrow whereby the underlined nucleotides are ³²P-labelled. Transcription was performed by Eσ⁷⁰ from the *lacUV5* template in the presence of 1:1, 1:2 or 1:4 molar ratios of Gp2^{WT} or NCS mutant. The percentage activity of Eσ⁷⁰ in the presence of the Gp2 proteins was calculated with respect to a reaction conducted in the absence of Gp2 and are indicated beneath each lane.

This data strongly suggests that the NCS residues specifically contribute to the inhibitory function of Gp2, and that the potency of inhibition is determined by the number of the negatively charged side chains within the NCS. In support of this view, *in vivo* complementation data conducted by Dr. Andrey Shadrin (a researcher in the Wigneshweraraj laboratory) also showed that T7 phage harbouring an amber mutation in *gene 2* were progressively less effective in infecting BL21 cells when complemented with pET28b⁺:Gp2^{Mut2}/Gp2^{Mut5}/Gp2^{Mut6}/Gp2^{Mut7} than with pET28b⁺:Gp2^{WT} (Sheppard *et al.*, 2011).

4.7. Is Gp2^{Mut7} retained in the RPo?

Gp2^{WT} has no effect on transcription from a pre-formed RPo because it is unable to bind to Eσ⁷⁰ (Camara *et al.*, 2010), hence it was proposed that Gp2 and the dw-DNA compete for binding to

the β' jaw. However, in the presence of the functionally defective mutant $Gp2^{Mut7}$ the RPo is still able to form and therefore transcription can occur, but it is not known whether $Gp2^{Mut7}$ remains bound or is displaced from this complex. To determine whether $Gp2^{Mut7}$ is retained in the RPo that forms in its presence, EMSAs were conducted using Alexa-488 labelled $Gp2^{WT}$ and $Gp2^{Mut7}$ as well as a ^{32}P -labelled *lacUV5* promoter probe which allows for the visualisation of protein and DNA components by fluorescence imaging (FI) and autoradiography (A) respectively. Therefore $Gp2^* - E\sigma^{70}$ and $E\sigma^{70} - ^{32}P$ -*lacUV5* binary complexes as well as the hypothetical $Gp2^* - E\sigma^{70} - ^{32}P$ -*lacUV5* ternary complex could be detected. The migration of free $Gp2^{WT*}$ and $Gp2^{WT*}$ bound to $E\sigma^{70}$ are indicated in Figure 4.7.1 lanes 1 and 2 whereas free *lacUV5* and $E\sigma^{70} - ^{32}P$ -*lacUV5* are indicated in lanes 5' and 6'. As expected, when ^{32}P -*lacUV5* was added to $E\sigma^{70}$ pre-incubated with $Gp2^{WT*}$ the predominate complex continued to be $Gp2^{WT*} - E\sigma^{70}$ (lane 7) with only a small proportion of $E\sigma^{70}$ escaping $Gp2^{WT*}$ inhibition to form the RPo (lane 7'). In contrast and as expected, $Gp2^{Mut7*}$ had no noticeable effect on the formation of the RPo; compare lane 6' to 8' and $Gp2^{Mut7*}$ was undetectable at the migratory position of the RPo upon the addition of ^{32}P -*lacUV5*, compare lane 8 to 8'.

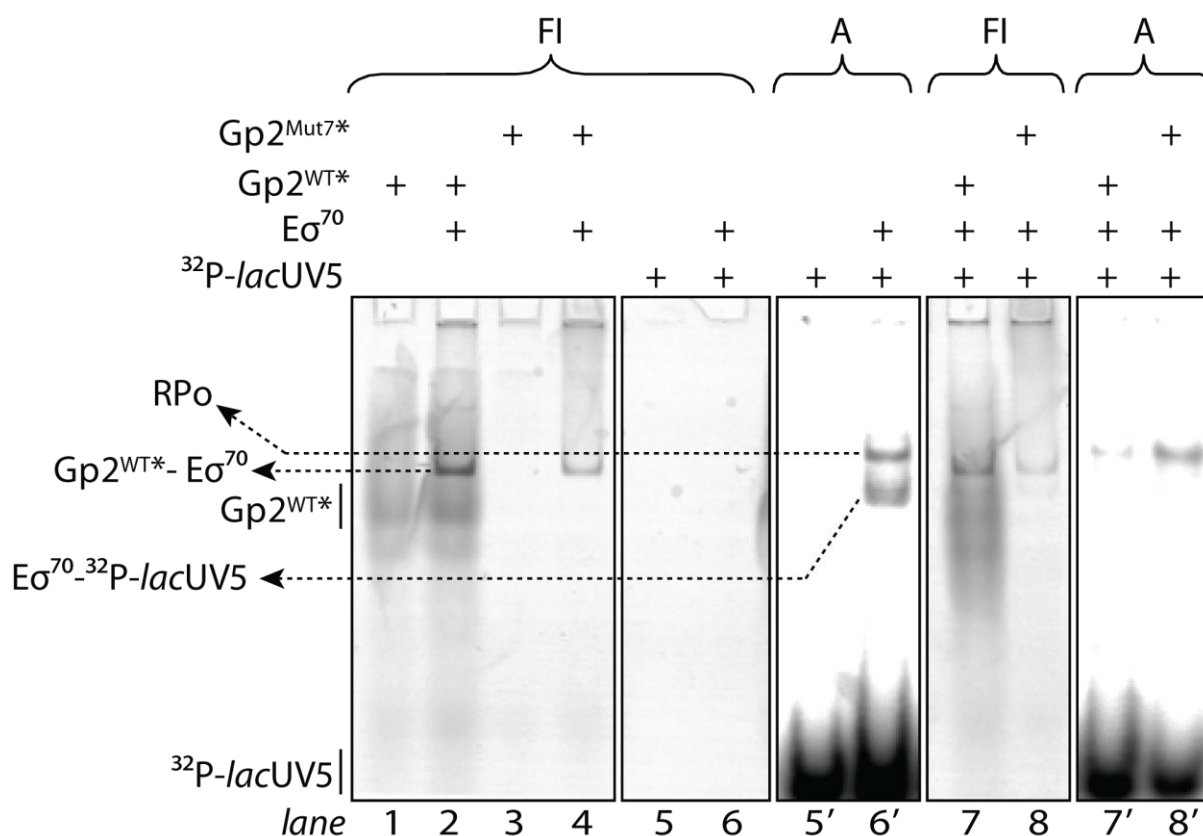


Figure 4.7.1 Native gel of $Gp2^{WT}$ and $Gp2^{Mut7}$ binding to $E\sigma^{70}$.

A fluorescence image (FI) (lanes 1-6 and 7-8) and autoradiograph (A) (lanes 5'-8') of a 4.5% (w/v) polyacrylamide native gel determining whether $Gp2^{Mut7}$ is retained in the RPo. The migration positions of unbound ^{32}P -*lacUV5*, $E\sigma^{70} - ^{32}P$ -*lacUV5*, $Gp2^{WT*}$, $Gp2^{WT*} - E\sigma^{70}$ and RPo are as indicated. Reaction components are marked at the top of each lane.

This result indicates that Gp2^{Mut7*} is displaced from E σ ⁷⁰ during the formation of the transcriptionally competent RPo which is able to form in its presence. In contrast, Gp2^{WT} remains bound leading to efficient prevention of RPo formation and consequently inhibition of transcription. Therefore, like Gp2^{WT}, formation of the RPo complex does not allow for Gp2^{Mut7} binding. However, whereas Gp2^{WT} is able to compete with the dw-DBC for binding to the β' jaw (refer to section 4.8) and consequently antagonise RPo formation in an order of addition dependent manner, the complete removal of the NCS seems to compromise this ability of Gp2^{Mut7}.

4.8. Conclusions

The results in this chapter have described how the NCS of amino acid side chains plays a major role in the inhibitory activity of Gp2. It was confirmed by NMR that the substitution of amino acids within the NCS does not detectably compromise the structural integrity of the Gp2 protein. FA measurements demonstrated that the binding affinities of each of the NCS mutant proteins to E σ ⁷⁰ were in the low nanomolar range and did not differ significantly from that of Gp2^{WT}. These results imply that the NCS is not essential for the initial binding of Gp2 to the E σ ⁷⁰. However, the shape of the binding curves does differ considerably, which actually may be informative to the specifics of the interaction. The difference in amplitudes of the anisotropy may suggest a change to the local environment of the Alexa-488 fluorophore. As the 1-D NMR spectra showed that the mutant proteins are all folded correctly the discrepancies in the binding curves may be due to an overall change in the surface interaction between Gp2 and the dw-DBC of RNAP. Therefore although R56 and R58 are the major determinants for the initial binding or docking of Gp2 to the β' jaw, the NCS, and indeed other residues, may contribute to the overall interaction or could be required for final positioning of Gp2. This is especially plausible considering the potentially strong intermolecular forces between the NCS and the positively charged 'walls' of the dw-DBC. These additional interactions may involve amino acid residues in other domains of the dw-DBC and could help explain how Gp2 has such a tight binding affinity. Alternatively, it has been shown that the oligonucleotide sequence in protein-DNA and amino acid residues in protein-protein interaction studies can affect the properties of an attached fluorophore, potentially quenching the intensity of the emission (Anderson *et al.*, 2008). Therefore the substitution of the amino acid residues in the NCS may directly influence the behaviour of Alexa-488. Despite any discrepancies in binding, these subtle differences cannot account for such drastic changes in Gp2 inhibitory activity.

The results from the *in vitro* transcription assay clearly show that there is a gradual decrease in the inhibitory activity of Gp2 which correlates with the progressive removal of the NCS. Complete removal of the NCS, entirely abolishes Gp2 activity, therefore the NCS is the major determinant for inhibitory function of Gp2. This finding is supported by an *in vivo* assay which showed that amino acid substitutions in the NCS significantly impaired the ability of T7 to infect *E. coli* (Sheppard *et al.*, 2011). The absolute removal of the NCS as judged by Gp2^{Mut7} results in a complete loss of inhibitory activity which implies that Gp2 function is entirely dependent on the conservation of this feature. Therefore Gp2 is unlikely to antagonise RPo formation simply by steric hindrance or mechanisms that involve the presence of the protein *per se*.

As Gp2 is a potent inhibitor of Eo⁷⁰ it is likely that Gp2 has evolved to function by disrupting a number of protein-DNA as well as protein-protein interactions which are necessary for RPo formation. To further explore the role of the NCS in the context of the molecular mechanism of inhibition of RNAP by Gp2 an NMR structure of Gp2 bound to a β' jaw fragment (residues 1153-1213) was determined (James *et al.*, 2012). The complex with Gp2 extends the β -sheet of the β' jaw fragment to an arrangement consisting of seven β -strands running in an anti-parallel conformation. As predicted, the major Gp2 binding determinants R56 and R58 are located in close proximity to one of the RNAP binding determinants, residue E1188 of the β' jaw (Figure 4.8.1) which generates significant electrostatic interactions, aligning the β_3 and β_2 strands of the β' jaw fragment with the β_3 and α_1 of Gp2. Additionally it seems that this positioning, extends the residues of the NCS of Gp2 by residues E1158, D1181, D1184, E1187 and E1188 of the β' jaw (Figure 4.8.2), thus supporting the view that the NCS is an important feature in Gp2 function.

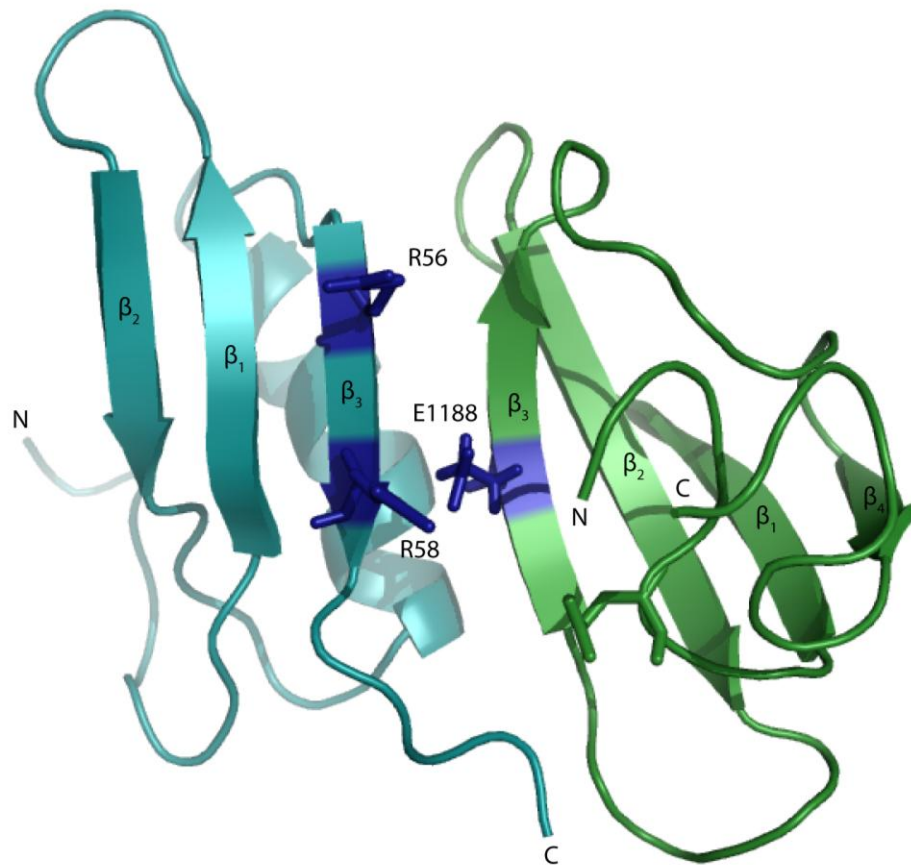


Figure 4.8.1 Structure of the Gp2-β' jaw fragment with binding determinants.

Ribbon representation of the Gp2-β' jaw fragment complex. Gp2 is shown in cyan and the β' jaw fragment in green. The R56, R58 and E1188 residues are shown in stick representation and coloured dark blue. Adapted from (James *et al.*, 2012).

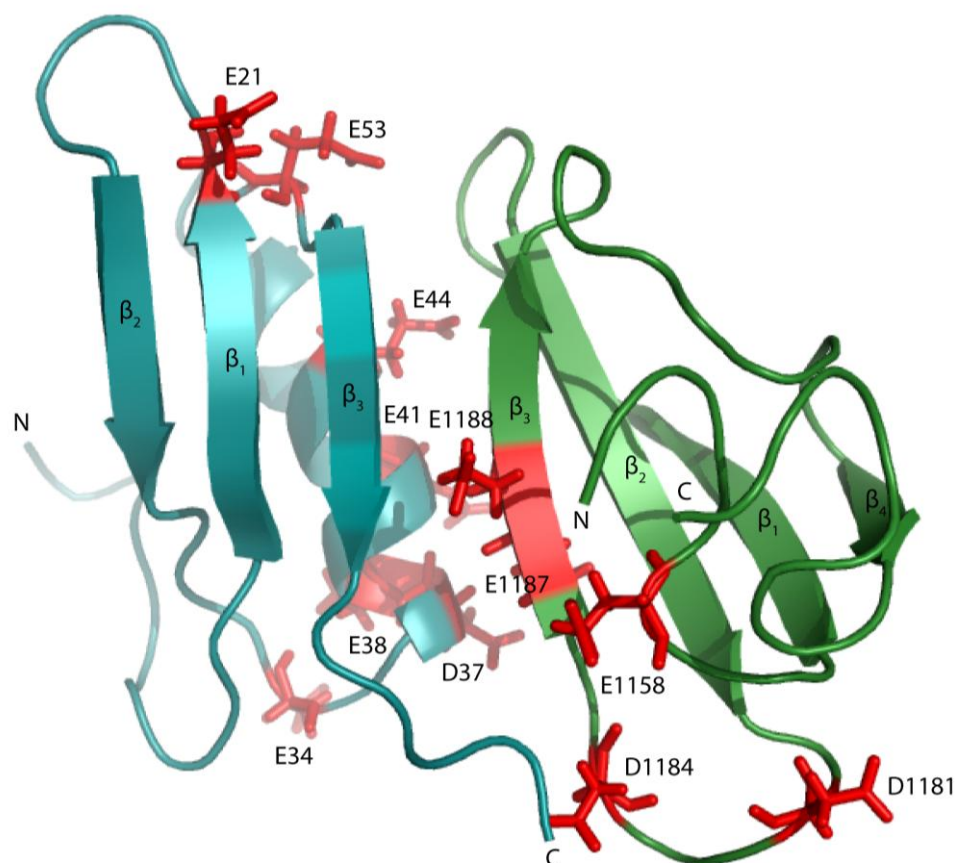


Figure 4.8.2 Structure of Gp2-β' jaw fragment with NCS.

Ribbon representation of the Gp2-β' jaw fragment complex. Gp2 is shown in cyan and the β' jaw fragment in green. The NCS residues (E21, E34, D37, E38, E41, E44 and E53) and the β' jaw NCS extension residues (E1158, D1181, D1184, E1187, E1188) are shown in stick representation and coloured red. Adapted from (James *et al.*, 2012).

To provide further insights into the precise mechanism of inhibition by Gp2, this structural complex was used together with the *E. coli* RNAP core (Opalka *et al.*, 2010) and the transcription initiating *T. aquaticus* (Murakami *et al.*, 2002b) structures to generate a composite structural model of the Gp2-RNAP complex (James *et al.*, 2012). Previous results have implied that Gp2 may inhibit Eσ⁷⁰ activity by directly competing for binding to the dw-DBC with the dw-DNA. The structural model showed that the surface of Gp2 binding does indeed face into the dw-DBC, where the dw-DNA is predicted to be positioned. Moreover, NMR titration experiments with the β' jaw fragment and ds-DNA fragment provided the first experimental evidence that the β' jaw fragment does specifically interact with the dw-DNA. Therefore these findings strengthen the proposal that Gp2 may function to antagonise the interactions between the β' jaw and the dw-DNA. To experimentally test this, NMR spectra were obtained from competition experiments whereby Gp2 was titrated into a β' jaw

fragment-ds-DNA complex. In this context, Gp2 was able to successfully displace the ds-DNA from the β' jaw fragment. Additionally the presence of Gp2 prevented the crosslinking of a ds-DNA segment to the β' jaw fragment (James *et al.*, 2012). These results strongly support the view that Gp2 antagonises interactions between the dw-DNA interactions and the dw-DBC which are necessary for RPo formation.

Due to the negative nature of the NCS and the extension of this charge by the β' jaw it was proposed that Gp2 may antagonise the interactions between the dw-DNA and the dw-DBC by projecting the NCS into the dw-DBC and electrostatically repelling the negatively charged dw-DNA. However in the current structural model the negatively charged residues of the NCS are predominantly facing towards the β' jaw and the $\beta'i6$ domains rather than projecting out into the active site cleft of the dw-DBC (Figure 4.8.3). Therefore this current structural model orientates Gp2 in a position that makes electrostatic repulsion of the dw-DNA unlikely.

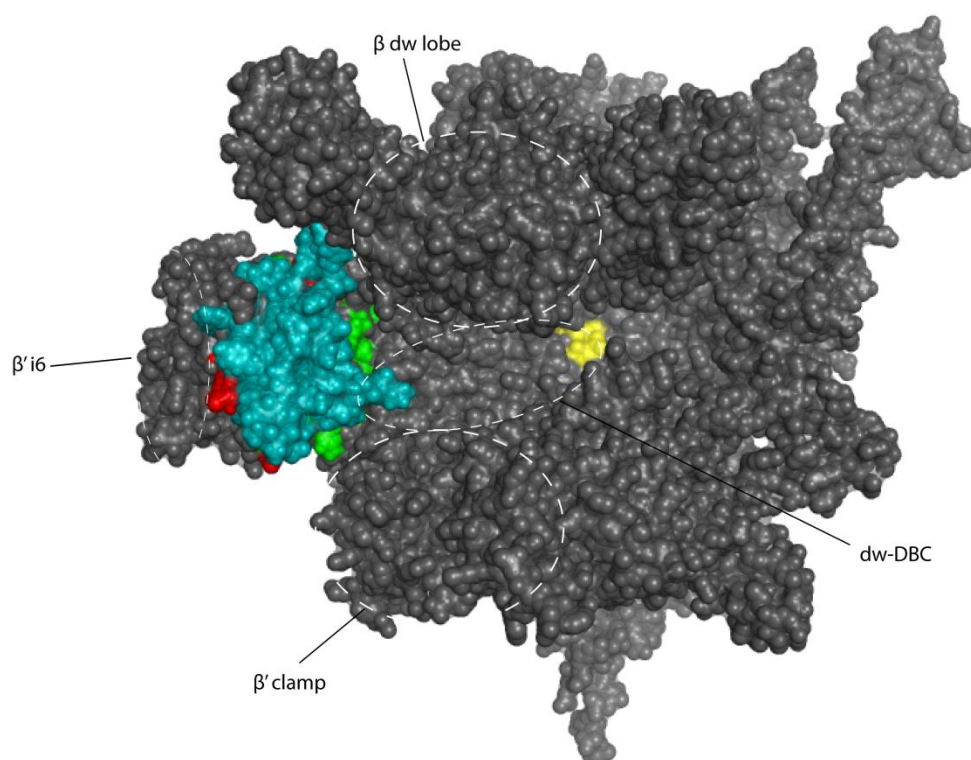


Figure 4.8.3 Structural composite model of *E. coli* RNAP core and Gp2.

Surface representation of Gp2 bound to the *E. coli* RNAP core. Gp2 is shown in cyan and the β' jaw fragment in green. The NCS residues are coloured red but are predominately hidden in this model. The active site is coloured yellow. The dw-DBC and the β dw-lobe, β' clamp and $\beta'i6$ mobile modules are circled and labelled. Adapted from (James *et al.*, 2012).

Alternatively and/or additionally the NCS may be better orientated to interact with the positively charged domains of the dw-DBC and 'lock' σ^{70} in a conformation unfavourable for RPo formation. Structural analysis has revealed that the two 'pincers' of RNAP can exist in numerous conformational states which can vary the diameter of the active-centre cleft between $> 20 \text{ \AA}$ when open and $\sim 8\text{-}12 \text{ \AA}$ when closed (Chakraborty *et al.*, 2012). It is conceivable that the diameter of the active cleft may dictate the entry and the conformational state of the DNA. During early stages of transcription initiation the β' clamp and the β dw-lobe which comprise the 'pincers' of RNAP (Figure 4.8.3) must move apart to permit the ds-DNA to enter the active-centre cleft but later in initiation the β' clamp must close to form a stable RPo. Therefore Gp2 could potentially target the movement of the β' clamp and β dw-lobe to either restrict access of the ds-DNA or to prevent the closure of the 'pincers'. Another possible mechanism by which Gp2 can effect interdomain-protein interactions or conformational changes involves the flexible and highly negatively charged σ^{70} R1.1 domain which also resides in the dw-DBC. The potential role of σ^{70} R1.1 in Gp2-mediated inhibition is the subject of chapter 5.

CHAPTER 5

THE ROLE OF σ^{70} REGION 1.1 DOMAIN DURING GP2-MEDIATED INHIBITION OF RNAP

5.1. Introduction

In order to initiate transcription, the RNAP core must interact with a dissociable promoter specificity factor known as a σ -factor. The σ -factor is responsible for the recognition of the consensus promoter elements as well as steps in DNA strand-separation. *E. coli* possess seven different σ -factors (refer to section 1.2.5) of which the RNAP core predominately associates with σ^{70} which is responsible for initiating the transcription of the housekeeping genes, particularly during exponential growth. σ^{70} -like factors are comprised of four domains (Figure 5.1.1), separated by flexible linkers, which form contacts either with the β and β' subunits of RNAP or the promoter DNA. The major determinants for promoter recognition are located in regions 2.4 and 4.2, which contact the -10 and -35 promoter elements respectively whereas domain 3.0 is responsible for interacting with the extended -10 promoter motif (Barne *et al.*, 1997).

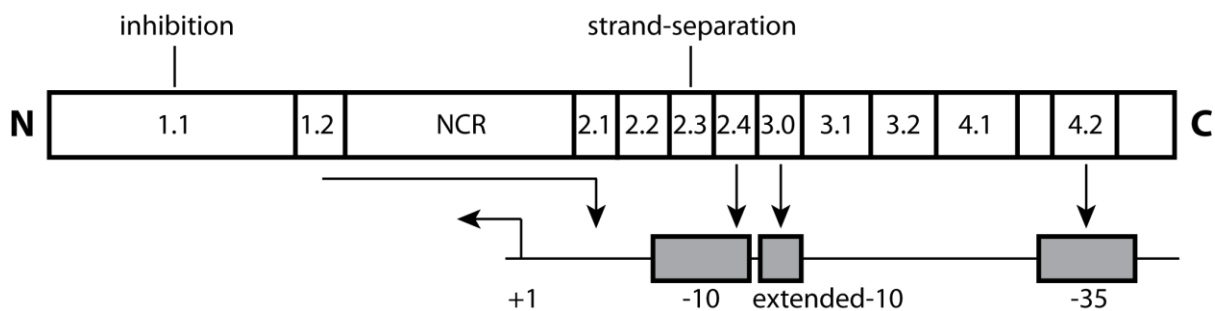


Figure 5.1.1 Conserved domains and functions of σ^{70} type factors.

The conserved regions and non-conserved region (NCR) of the σ^{70} type factors are indicated. Arrows point to the regions of contact between the σ -factor and the discriminator region, the -10, extended -10 and -35 promoter elements of the DNA. The subdomains, regions 1.1 and 2.3 are required for auto-inhibition and DNA strand-separation respectively. Adapted from (Gruber and Gross, 2003).

Domain 1 is only present in a specific subset of σ^{70} -like factors and although it tends to vary in sequence, its overall negative charge is always conserved (Gruber and Gross, 2003). High resolution structures of fragments of free σ (Malhotra *et al.*, 1996; Schwartz *et al.*, 2008; Campbell *et al.*, 2002) and of the *T. aquaticus* and *T. thermophilus* holoenzymes (Murakami *et al.*, 2002b; Vassilyev *et al.*, 2002) have provided valuable information on the extensive E- σ contacts and transcription initiation transition steps however, to date R1.1 remains unresolved in these structures, most likely due to its flexible nature as seen by DNA footprinting (Nagai and Shimamoto, 1997). However, biochemical experiments have shown R1.1 to function as an auto-inhibitory

regulator by concealing the DNA binding determinants in the free protein, presumably to prevent binding of the σ -factor to the promoter DNA in the absence of RNAP core (Dombroski *et al.*, 1992). This view is supported by LRET, FRET and crosslinking data which compared the interdomain distances between free σ and the holoenzyme; (Schwartz *et al.*, 2008; Mekler *et al.*, 2002; Callaci *et al.*, 1999), implying that the free σ is much more compact in structure and therefore must undergo large scale conformational changes to form its extensive interactions upon binding to RNAP core. Further to these observations, it has also been suggested that the positioning of R1.1 is critical for steps in transcription initiation. It was proposed from structural analyses that in the context of $E\sigma^{70}$ the location of R1.1 corresponds to the main channel; which could make it responsible for the widening of the 'pincers' to facilitate entry of the ds-DNA (Murakami *et al.*, 2002b). The location of R1.1 in the downstream portion of the main channel was confirmed by FRET experiments. Moreover, it was also demonstrated that during RPo formation, R1.1 is repositioned from the inside of the positively charged channel to make alternative interactions with the also positively charged β -pincer tip formed by the dw-lobe. From this observation it was proposed that the highly negatively charged R1.1 acts as a molecular mimic of the dw-DNA and therefore serves as a placeholder for the DNA until proficient RPo can occur (Mekler *et al.*, 2002).

Some of the very early work on Gp2 noted that its inhibitory effect on DNA binding to RNAP is 2-fold more effective in the context of $E\sigma^{70}$ than the RNAP core (DeWyngaert and Hinkle, 1979). This is surprising; recall that in chapter 3 it was demonstrated that Gp2 binds to the RNAP core with the same affinity as $E\sigma^{70}$. Therefore this early observation could imply that although σ^{70} may not influence the binding affinity of Gp2 it could, in part, contribute to the inhibitory activity of Gp2. Gp2, like R1.1 has regions of high negative charge and is also located in the same vicinity of the dw-DBC of RNAP. Therefore, the role, if any, of this specific domain of σ^{70} in Gp2 inhibition of $E\sigma^{70}$ was explored. It was hypothesised that as Gp2 targets a stage *en route* to RPo formation, it may be possible that Gp2 could exploit the properties and essential function of R1.1 during this process.

Recent experiments by Mekler *et al.*, which used Gp2 purified from this study, explored the mechanism of action of Gp2; specifically focusing on how the inhibitor influences the binding of promoter DNA to $E\sigma^{70}$. Assays were performed using an engineered T5N25 promoter which includes consensus -10 and -35 elements, an optimised UP element as well as an extended -10 motif. This strong promoter probe referred to as T5N25CONS enables the formation of a stable ternary Gp2- $E\sigma^{70}$ -DNA complex which displays some of the hallmarks of the RPo but crucially is transcriptionally inactive. Thus, this ternary Gp2- $E\sigma^{70}$ -DNA complex can be considered as an intermediate complex that forms during RPo formation. DNA footprints of these complexes show that Gp2 reduces the

DNaseI dw-DNA protection from the +20 to the +8 site which supports the overall view that Gp2 reduces the affinity of the dw-DNA for the dw-DBC. KMnO₄ DNA footprinting revealed that the upstream edge of the transcription bubble (up to -6) was resistant to Gp2 inhibition and therefore single-stranded, whereas the downstream portion (-5 to +2) remained duplexed. These observations strongly supports previous *in vitro* transcription experiments from our laboratory which implied that Gp2 effects a late stage of RPo formation associated with DNA strand separation around the + 1 site (Camara *et al.*, 2010). Quantitative binding assays were then used to dissect the effect of Gp2 on the binding of E σ^{70} to different portions of the T5N25CONS promoter. Results showed that Gp2 has no significant effect on the binding of DNA upstream of the -10 site to E σ^{70} but does inhibit the binding of single-stranded downstream extensions of the template strand around the active site. These findings suggest that the inhibitory effects of Gp2 are not confined solely to its site of binding at the β' jaw. Hence, this study not only investigated the possible contribution of R1.1 in Gp2 inhibition but also how it may directly/indirectly influence the binding of different segments of promoter DNA to E σ^{70} .

5.2. Objectives

The primary focus of this chapter is to explore the potential contribution that σ^{70} R1.1 may have in the inhibition of E σ^{70} by Gp2. This study will involve the *in vitro* analysis of R1.1 deleted versions of σ^{70} in the context of the Gp2 inhibition of E σ^{70} . The main objective is to determine whether R1.1 is required for Gp2 inhibition of E σ^{70} and to investigate the role that R1.1 may play in the mechanism of inhibition. In addition, this study will also further investigate the extent to which Gp2 restricts the entry of the promoter DNA and whether R1.1 may be involved in this antagonistic effect. This data will provide further insights into the overall molecular mechanism employed by Gp2 during the inhibition of transcription initiation. This study includes the following individual aims:

- Over-express and purify $\sigma^{70\Delta R1}$ and $\sigma^{70\Delta R1.1}$ as well as the isolated R1.1 peptide.
- Determine the binding affinity of Gp2 to RNAP reconstituted with the mutant versions of σ^{70} .
- Investigate the inhibitory activity of Gp2 on E σ^{70} transcription using $\sigma^{70\Delta R1}$ and $\sigma^{70\Delta R1.1}$ and with R1.1 added *in trans*.
- Explore the effect of Gp2 on the interactions between E σ^{70} and various segments of the promoter DNA and the potential contribution of R1.1.
- Unravel the mechanism by which R1.1 may contribute to the inhibition of E σ^{70} by Gp2.

5.3. The effect of the σ^{70} N-terminal on Gp2 inhibitory activity

Gp2 binds to the β' jaw domain which forms part of the dw-DBC of RNAP. The N-terminal domain of σ^{70} is also believed to localise within this channel upon $E\sigma^{70}$ formation but at some stage prior to RPo formation it is repositioned to the tip of the β dw-lobe at the exterior of the channel. On the basis of this expected close proximity, preliminary experiments were performed to ascertain if the N-terminus of σ^{70} influences Gp2 binding and activity. FA binding experiments were conducted with both the full length $E\sigma^{70}$ and a version of the $E\sigma^{70}$ lacking σ^{70} amino acid residues 1-130 referred to as $E\sigma^{70\Delta R1}$ (Figure 5.3.1). It was predicted that as Gp2 binds to RNAP core (E) and $E\sigma^{70}$ with equal affinities, the removal of the σ^{70} N-terminal domain was unlikely to affect Gp2-E binding affinity.

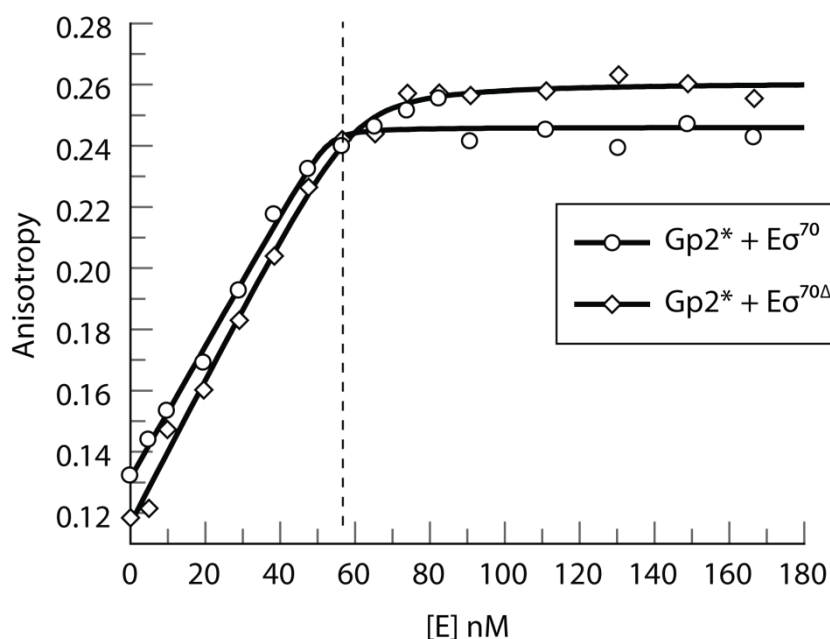


Figure 5.3.1 Graph displaying the binding of Gp2* to $E\sigma^{70}$ and $E\sigma^{70\Delta R1}$ as measured by FA.

Gp2* (50 nM) binding to $E\sigma^{70}$ (circles) and $E\sigma^{70\Delta R1}$ (diamonds) is represented by anisotropy which is displayed as a function of the concentration of E. The dotted line indicates the concentration at which Gp2* binding becomes saturated.

As expected Gp2* binds to $E\sigma^{70\Delta R1}$ with the same equimolar stoichiometry and equivalent affinity as to the full length $E\sigma^{70}$ with K_d values of 0.83 nM and 0.40 nM respectively. This confirms that at least in the context of $E\sigma^{70}$, σ^{70} domains R1.1 and R1.2 have no detectable effect on Gp2 binding.

Next, it was considered whether R1 of σ^{70} could contribute to the inhibitory activity of Gp2. To investigate this potential contribution, the effect of the removal of the σ^{70} N-terminal on Gp2-mediated inhibition of transcription was tested. The inhibitory activity of Gp2 was assessed in an *in vitro* transcription assay; $E\sigma^{70WT}$ or $E\sigma^{70\Delta R1}$ was pre-incubated with either Gp2^{WT} or the defective mutant Gp2^{R56E} prior to the addition of the *lacUV5* DNA template (Figure 5.3.2).

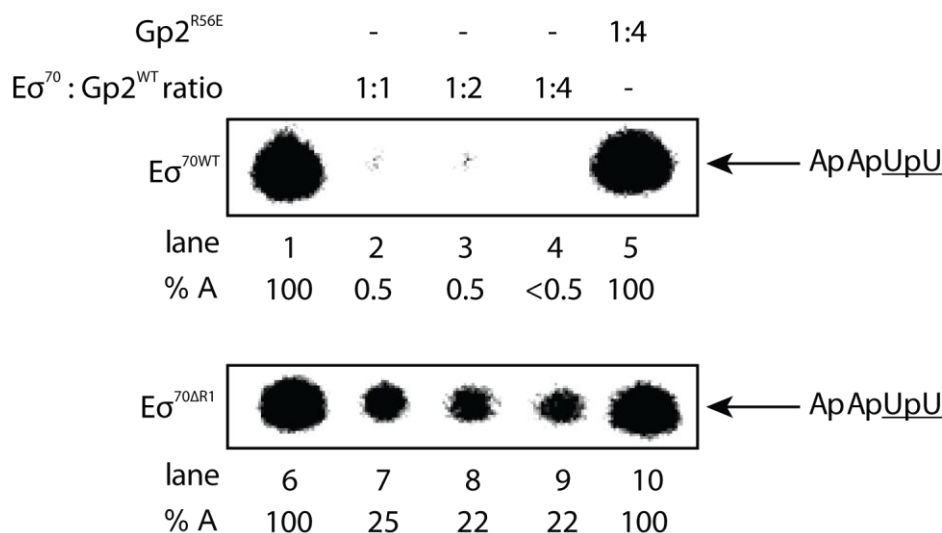


Figure 5.3.2 Inhibitory activity of Gp2 on $E\sigma^{70}$ and $E\sigma^{70\Delta R1}$ transcription.

An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing the synthesis of the ApApUpU transcript as indicated by the arrow whereby the underlined nucleotides are ³²P-labelled. Transcription was performed by $E\sigma^{70}$ (top panel) and $E\sigma^{70\Delta R1}$ (bottom panel) from the *lacUV5* template in the presence of 1:1, 1:2 or 1:4 molar ratios of Gp2^{WT} or 1:4 molar excess of Gp2^{R56E}. The percentage activity (% A) of $E\sigma^{70}$ in the presence of the Gp2 proteins was calculated with respect to reactions conducted in the absence of Gp2 (lanes 1 and 6) and are indicated beneath each lane.

Removal of the σ^{70} N-terminal reduced transcription inhibition by up to ~25% providing the first piece of experimental evidence to show that the N-terminal of σ^{70} is required for the efficient inhibition of $E\sigma^{70}$ transcription by Gp2. Even increasing the molar ratio of $E\sigma^{70}$:Gp2 from 1:1 to 1:4 (lanes 7-9) did not detectably recover Gp2 inhibition, implying that the potential role of σ^{70} N-terminal may not be attributed simply to increasing Gp2 binding or activity but as an extension of the overall inhibitory mechanism. Control reactions were performed with the point mutant Gp2^{R56E} which is defective for binding and therefore has no effect on transcription. Transcription reactions with $E\sigma^{70\Delta R1}$ and 1:4 molar excess of Gp2^{R56E} displayed no change in levels of the ApApUpU transcript (compare lanes 6 and 10), demonstrating that the inhibition shown in lanes 2-4 and 7-9 is specific to Gp2^{WT}.

5.4. The contribution of σ^{70} R1.1 domain to inhibition of $E\sigma^{70}$ by Gp2

Thus far, experiments were performed using a σ^{70} N-terminal deletion mutant deficient in regions 1.1 and 1.2. However, it is explicitly R1.1 that is believed to act as DNA mimic and thus serve as a placeholder for the dw-DNA, whereas R1.2 is known to make protein-DNA contacts with the discriminator region (Haugen *et al.*, 2006). Although mutations in R1.2 have been shown to effect various stages of transcription initiation and elongation (Baldwin and Dombroski, 2001), therefore it was considered whether if it is solely R1.1 (and not R1.2) of σ^{70} which is responsible for the Gp2-mediated inhibition of transcription. Henceforth, a further mutant version of σ^{70} was constructed which lacked only the first 100 amino acid residues comprising exclusively region 1.1, referred to as $E\sigma^{70\Delta R1.1}$. The ability of Gp2 to inhibit $E\sigma^{70}$ reconstituted with either the full length σ^{70} or $\sigma^{70\Delta R1.1}$ was again determined using an *in vitro* transcription assay (Figure 5.4.1). The incubation of 2-fold molar excess of Gp2 with $E\sigma^{70WT}$ before the addition of the *lacUV5* promoter template completely abolished transcription (lane 2) as observed previously. Consistent with the preliminary $\sigma^{70\Delta R1}$ experiment, $E\sigma^{70\Delta R1.1}$ was inhibited far less efficiently (~55% inhibition) than $E\sigma^{70WT}$. Using comparable experimental conditions the affinities of Gp2* for $E\sigma^{70}$ and $E\sigma^{70\Delta R1.1}$ were measured by FA and were found to be equivalent (Figure 5.4.5). Therefore, it seems that full inhibition of $E\sigma^{70}$ by Gp2 specifically requires R1.1 of σ^{70} .

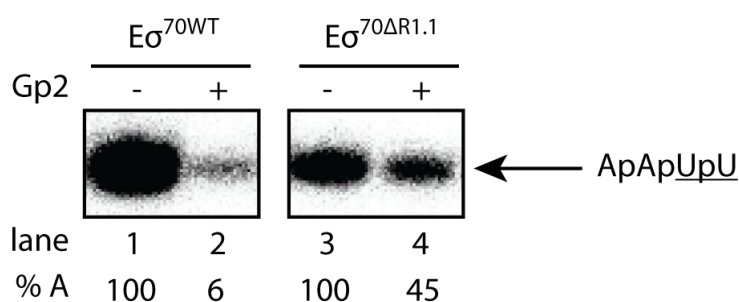


Figure 5.4.1 Inhibitory activity of Gp2 on $E\sigma^{70WT}$ and $E\sigma^{70\Delta R1.1}$ transcription.

An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing the synthesis of the ApApUpU transcript as indicated by the arrow whereby the underlined nucleotides are ^{32}P -labelled. Transcription was performed by $E\sigma^{70WT}$ (lanes 1 and 2) and $E\sigma^{70\Delta R1.1}$ (lanes 3 and 4) from the *lacUV5* template in the presence of 1:2 molar excess of Gp2^{WT}. The percentage activity (% A) of $E\sigma^{70}$ in the presence of Gp2 (lanes 2 and 4) was calculated with respect to a reaction conducted in the absence of Gp2 (lanes 1 and 3) and is indicated beneath each lane.

To confirm that it is the presence and presumably the localisation of R1.1 that is responsible for the inhibitory effect of Gp2 rather than a structural side-effect of the deletion, R1.1 (σ^{70} amino acids residues 1-100) was isolated and its *in trans* effect on $E\sigma^{70\Delta R1.1}$ transcription assessed in the presence and absence of Gp2 to test whether it allows for the full inhibition of $E\sigma^{70}$ by Gp2 (Figure 5.4.2). The addition of R1.1 almost fully complemented for the reduced inhibitory activity of Gp2 on $E\sigma^{70\Delta R1.1}$ (lanes 4-7). Transcription was reduced to that seen with $E\sigma^{70}$, with full inhibition occurring at a 1:4 molar excess of R1.1 over $\sigma^{70\Delta R1.1}$ (lane 6). Thus, the presence of the isolated R1.1 domain *in trans* with $E\sigma^{70\Delta R1.1}$ allows Gp2 to efficiently inhibit transcription initiation.

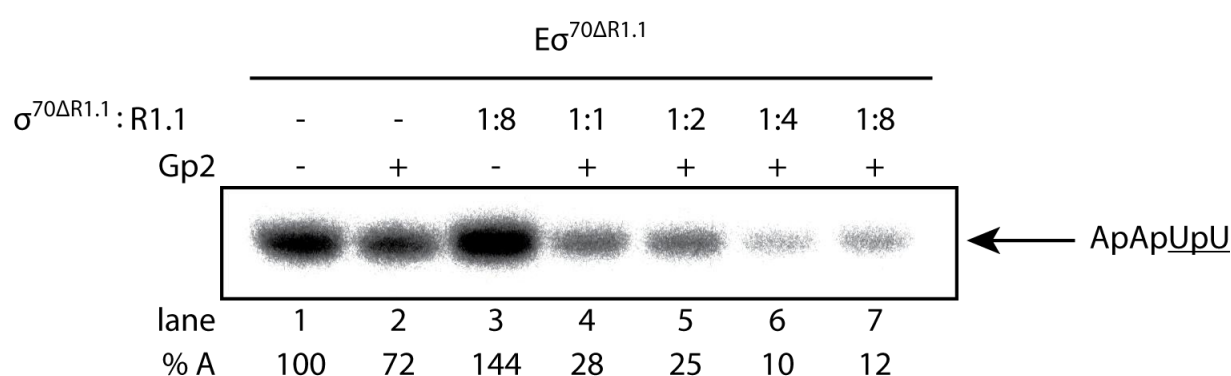


Figure 5.4.2 The *in trans* effect of $\sigma^{70\Delta R1.1}$ on Gp2-mediated inhibition of $E\sigma^{70\Delta R1.1}$ transcription.

An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing the synthesis of the ApApUpU transcript as indicated by the arrow whereby the underlined nucleotides are ^{32}P -labelled. Transcription was performed by $E\sigma^{70\Delta R1.1}$ from the *lacUV5* template in the absence (lane 1) and presence of 1:2 molar excess of Gp2^{WT} (lanes 2-5). R1.1 was added *in trans* to the reactions at molar ratios indicated above each lane. The percentage activity (% A) of $E\sigma^{70}$ in the presence of Gp2 and/or R1.1 was calculated with respect to a reaction conducted in the absence of Gp2 and are indicated beneath each lane.

Furthermore the *in trans* effect of R1.1 proved not to be order of addition dependent (Figure 5.4.3) which suggests that under these experimental conditions Gp2 and R1.1 are able to bind to the dw-DBC independently of each other, supporting the view that R1.1 does not influence the initial binding of Gp2 to the $E\sigma^{70}$. Control reactions established that the *in trans* presence of the isolated R1.1 domain in $E\sigma^{70\Delta R1.1}$, in the absence of Gp2, did not antagonise the ability of $E\sigma^{70\Delta R1.1}$ to transcribe (compare Figure 5.4.2 lanes 1 and 3). In fact, the *in trans* addition of R1.1 in the absence of Gp2 notably increased transcription by ~44%. This effect may be due to the auxiliary role that R1.1 has in support of transcription; prior to its displacement from the dw-DBC, R1.1 is believed to hold open

the ‘pincers’ of RNAP facilitating the entry of the ds-DNA into the channel (Murakami *et al.*, 2002b). Therefore in the absence of Gp2, R1.1 promotes the formation of the RPo.

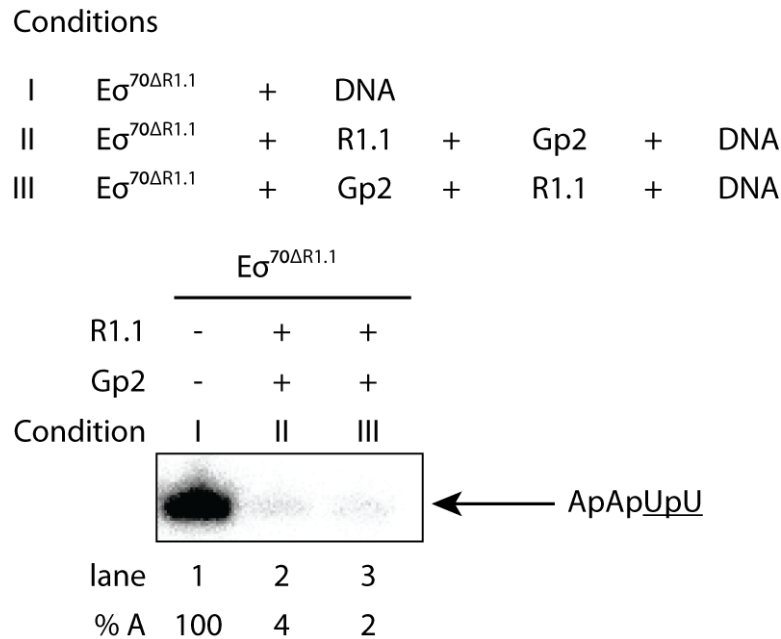


Figure 5.4.3 The effect of the order of addition of *in trans* R1.1 on Gp2-mediated inhibition of transcription.

An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing the synthesis of the ApApUpU transcript as indicated by the arrow whereby the underlined nucleotides are ^{32}P -labelled. Transcription was performed by $E\sigma^{70\Delta R1.1}$ from the *lacUV5* template in the absence (lane 1) and presence of 1:2 molar excess of Gp2 and 1:8 molar excess of R1.1 (lanes 2 and 3). The order of addition of the major components of each reaction is illustrated above. The percentage activity (% A) of $E\sigma^{70}$ in the presence of the Gp2 and R1.1 was calculated with respect to the reaction conducted in the absence of Gp2 and R1.1 (lane 1) and are indicated beneath each lane.

It seems that the R1.1 domain of σ^{70} is required for the efficient inhibition of $E\sigma^{70}$ by Gp2. To further corroborate this finding the effect of Gp2 on transcription from holoenzymes formed from alternative σ -factors was explored. The *E. coli* σ -factor, σ^{38} is responsible for the transcription of genes during the stationary phase. The promoter specificity of σ^{38} overlaps with that of σ^{70} (Gruber and Gross, 2003) but crucially it does not contain a R1.1 domain. Therefore it can serve as a valuable tool for evaluating the role of R1.1 in Gp2-mediated inhibition, in the context of a holoenzyme that natively lacks R1.1. To this end, the ability of Gp2 to inhibit transcription from E σ reconstituted with either $\sigma^{70\text{WT}}$, $\sigma^{70\Delta R1.1}$ or σ^{38} was compared using the *E. coli osmE* promoter (Bordes *et al.*, 2000) (Figure 5.4.4).

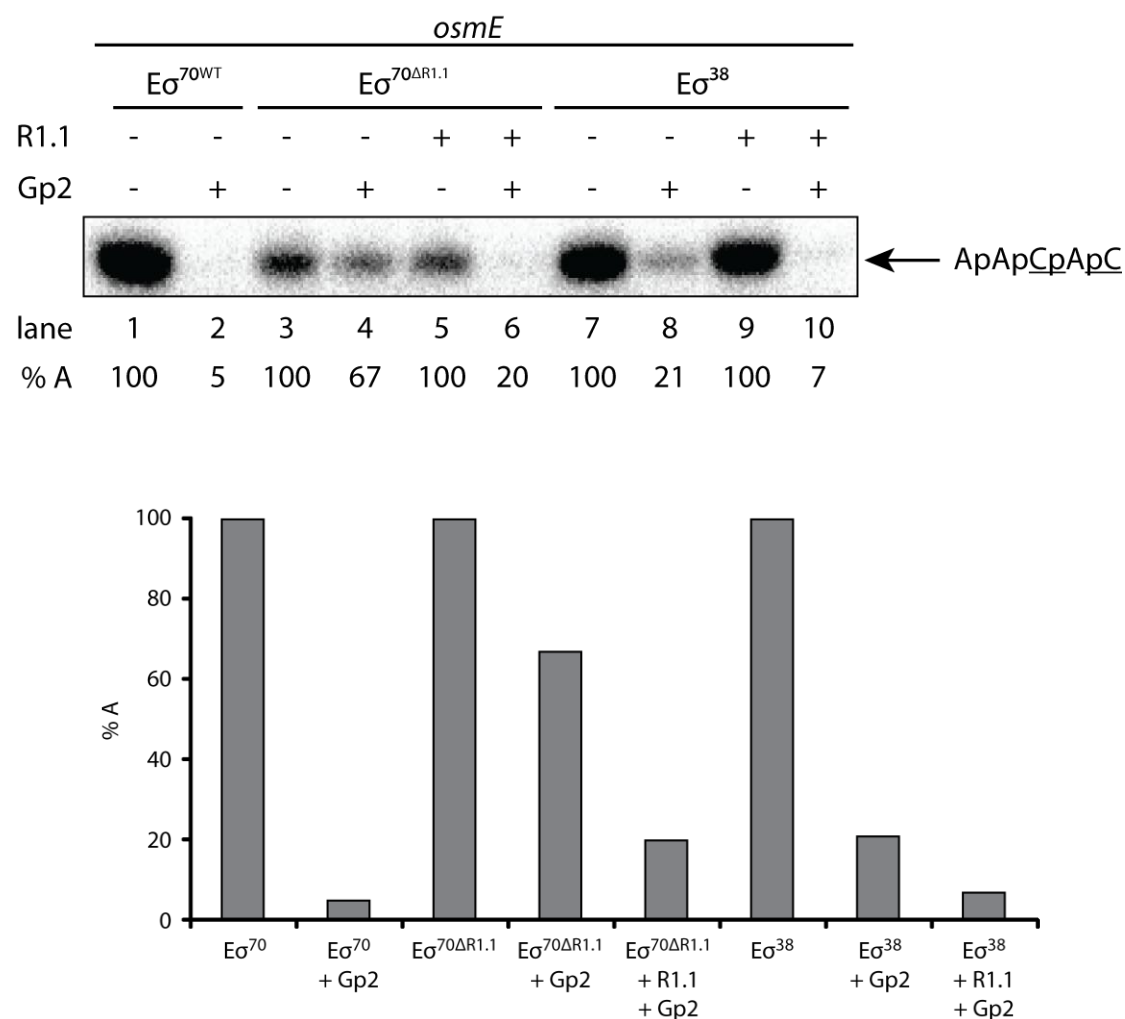


Figure 5.4.4 Inhibitory activity of Gp2 on $E\sigma^{70WT}$, $E\sigma^{70\Delta R1.1}$ and $E\sigma^{38}$ transcription.

Top: An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing the synthesis of the ApApCpApC transcript as indicated by the arrow whereby the underlined nucleotides are ^{32}P -labelled. Transcription was performed by $E\sigma^{70WT}$, $E\sigma^{70\Delta R1.1}$ and $E\sigma^{38}$ from the *osmE* template in the absence (lanes 1, 3, 5, 7 and 9) and presence of 2-fold molar excess of Gp2 (lanes 2, 4, 6, 8 and 10). R1.1 was added *in trans* using a 8-fold molar excess to reactions with and without Gp2 (lanes 5, 6, 9 and 10). The percentage activity (% A) of each holoenzyme in the presence of Gp2 and/or R1.1 was calculated with respect to a reaction conducted in the absence of Gp2 (lanes 1, 3, 5, 7 and 9) and are indicated beneath each lane. Bottom: Graphical representation of lanes 1, 2, 3, 4, 6, 7, 8 and 10.

A 4-fold molar excess of Gp2 over σ^{70WT} almost completely abolished the synthesis of the ApApCpA *osmE* transcript (lane 2) as seen previously on the *lacUV5* promoter template. In contrast, under the same experimental conditions, $E\sigma^{38}$ -dependent transcription was only inhibited by ~80% (lane 8). As expected, full inhibition of transcription by $E\sigma^{70}$ on the *osmE* promoter was R1.1-dependent (compare lanes 3-6). Remarkably, the *in trans* addition of R1.1 to $E\sigma^{38}$ allowed for the further inhibition of transcription by Gp2 to levels comparable to that of $E\sigma^{70WT}$ (lanes 1-2 and 7-10), suggesting that R1.1 actively contributes to the full inhibition of RNAP by Gp2.

The affinity of Gp2 for $E\sigma^{70WT}$, $E\sigma^{70\Delta R1.1}$, $E\sigma^{38}$ and the major alternative holoenzyme $E\sigma^{54}$ was measured by FA (Figure 5.4.5). Recall that Gp2 has no significant effect on $E\sigma^{54}$ -dependent transcription (Wigneshweraraj *et al.*, 2004), however unlike σ^{38} , σ^{54} does contain a region 1 domain which functions to regulate isomerisation to the RPo in a comparable manner to σ^{70} R1.1 (Cannon *et al.*, 1999). The affinity of Gp2 for each of these holoenzymes was found to be equivalent, with K_d values in the same range (Table 5.4.1). Thus, the β' jaw, not R1.1 constitutes as the major docking site of Gp2, however an interaction between R1.1 and Gp2 cannot be excluded. Furthermore, this interesting result signifies that Gp2 is able to selectively inhibit each of the holoenzymes with differing degrees, even though the initial binding of Gp2 is comparable. These results strongly suggest that R1.1 of σ^{70} forms part of the mechanism which contributes to the full and specific inhibition of $E\sigma^{70}$ transcription by Gp2.

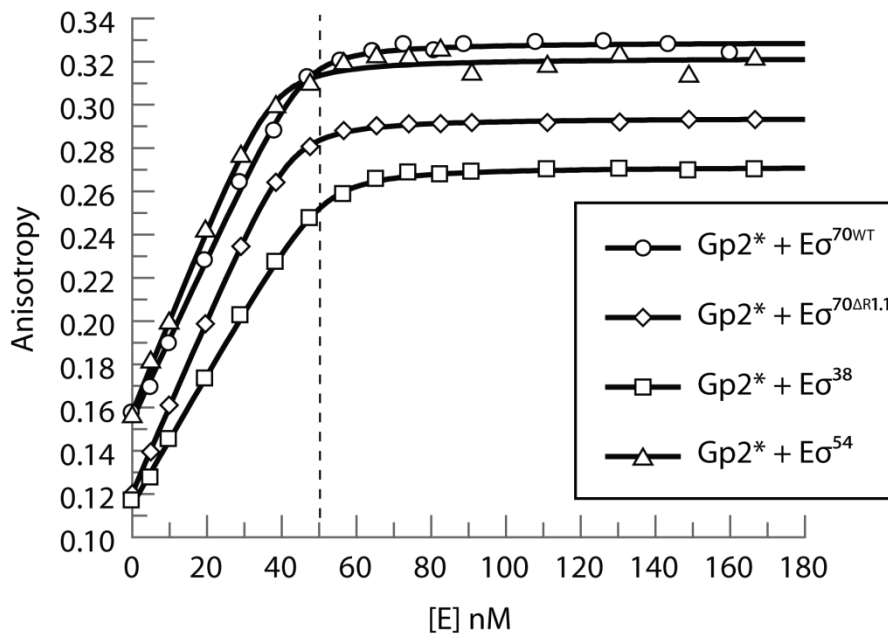


Figure 5.4.5 Graph displaying the binding of Gp2* to variant holoenzymes as measured by FA.

Gp2* binding to $E\sigma^{70WT}$ (circles), $E\sigma^{70\Delta R1.1}$ (diamonds), $E\sigma^{38}$ (squares) and $E\sigma^{54}$ (triangles) is represented by anisotropy which is displayed as a function of the concentration of E. The dotted line indicates the concentration at which Gp2^{WT}* binding to $E\sigma^{70}$ becomes saturated.

E	K _d (nM)	S.E.
E	0.53	0.27
E σ^{70}	0.44	0.34
E $\sigma^{70\Delta R1.1}$	0.66	0.06
E σ^{38}	0.77	0.14
E σ^{54}	0.78	0.40

Table 5.4.1 Affinity of Gp2* binding to variants of E.

K_d values were calculated from FA assays shown in Figure 5.4.5.

5.5. The effect of Gp2 on promoter-less transcription

To further corroborate the role of R1.1 in the inhibition of RNAP by Gp2, experiments were conducted with a promoter-less minimal scaffold, here after referred to as the MS. The MS template consists of an 18 nucleotide long DNA duplex (lacking the -10 and -35 consensus promoter elements) and an 8 nucleotide long RNA-DNA heteroduplex separated by two unpaired DNA bases (Figure 5.5.1). Crucially, transcription from the MS can occur independently of the σ -factor. The addition of $\alpha^{32}\text{P}$ -UTP to the RNAP-MS probe complex results in the synthesis of a 9 nucleotide-long $\alpha^{32}\text{P}$ -UTP-labelled RNA product, hereafter called RNA-U (Kulbachinskiy *et al.*, 2004) independent of the σ -factor (Figure 5.5.2 lane 1).



Figure 5.5.1 Sequence of the MS.

Sequences and assembly of the minimal nucleic acid scaffold formed from an RNA-8mer (red) which is bound to the template strand of the DNA (blue), which in turn is annealed to the non-template strand of the DNA (black). Transcription from this MS results in the addition of a uridine at the 3' end of the RNA.

Gp2 inhibits the synthesis of RNA-U by RNAP from the MS only in the presence of σ^{70} (Figure 5.5.2 lane 5); σ -independent transcription (lanes 1-3) seems to be almost completely unaffected by Gp2, implying that Gp2 does rely significantly on σ^{70} for full inhibition of the RNAP.

Conditions

I	E/E σ^{70WT} /E $\sigma^{70\Delta R1.1}$	+	MS
II	E/E σ^{70WT} /E $\sigma^{70\Delta R1.1}$	+	Gp2 + MS
III	E/E σ^{70WT} /E $\sigma^{70\Delta R1.1}$	+	MS + Gp2

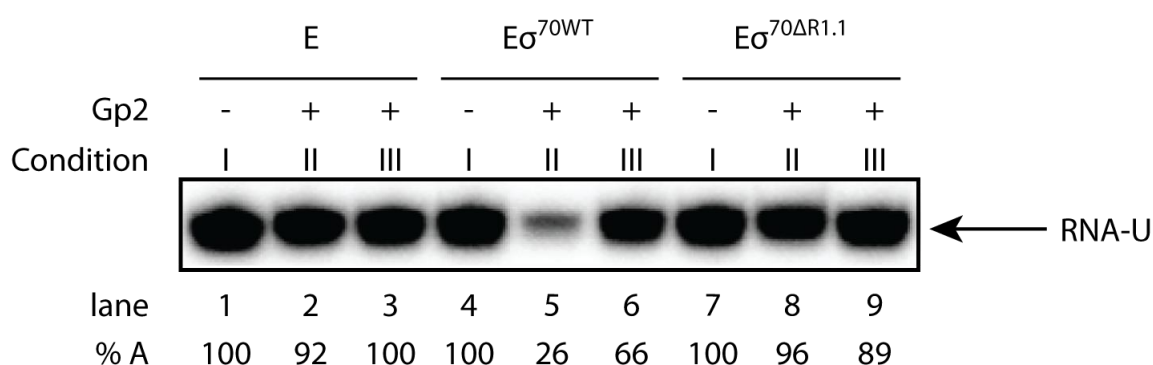


Figure 5.5.2 Inhibitory effect of Gp2 on E σ^{70WT} and E $\sigma^{70\Delta R1.1}$ transcription from the MS.

An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing the synthesis of the RNA-U transcript. Transcription was performed by E σ^{70WT} (lanes 1, 2 and 3) and E $\sigma^{70\Delta R1.1}$ (lanes 4, 5 and 6) from the MS template in the absence (lanes 1 and 4) and presence (lanes 2, 3, 5 and 6) of 1:2 molar excess of Gp2. Order of addition reactions were conducted as shown in the schematic above. The percentage activity (% A) of E σ^{70} in the presence of the Gp2 proteins was calculated with respect to the reactions conducted in the absence of Gp2 and are indicated beneath each lane.

Further to and consistent with previous results (Figure 5.4.1), E $\sigma^{70\Delta R1.1}$ -dependent transcription also displayed no significant inhibition of RNA-U synthesis by Gp2 (lanes 7-9) suggesting that specifically R1.1 of σ^{70} is required for the full inhibition of RNAP by Gp2. In support of this view, when the isolated R1.1 domain was added *in trans* to E $\sigma^{70\Delta R1.1}$ it conferred a significant degree of sensitivity to Gp2 (Figure 5.5.3 lane 4). Results from these MS assays confirm that Gp2 inhibition of transcription by E σ^{70} is dependent on σ^{70} with a specific requirement for R1.1.

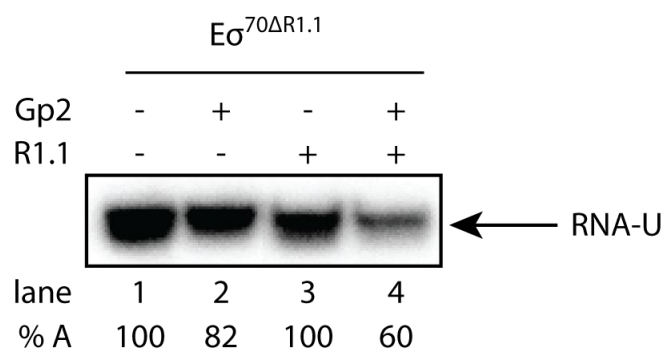


Figure 5.5.3 The *in trans* effect of R1.1 on Gp2-mediated inhibition of $E\sigma^{70\Delta R1.1}$ transcription.

An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing the synthesis of the RNA-U transcript. Transcription was performed by $E\sigma^{70\Delta R1.1}$ from the MS template in the absence (lane 1 and 3) and presence (lanes 2 and 4) of 1:2 molar excess of Gp2. R1.1 was added *in trans* to the specific reactions (lanes 3 and 4) at a 1:8 molar ratio. The percentage activity of $E\sigma^{70}$ in the presence of Gp2 and/or R1.1 was calculated with respect to the reactions conducted in the absence of Gp2 and are indicated beneath each lane.

5.6. The effect of Gp2 and R1.1 on DNA entry into the active site

Experiments investigating the effect of Gp2 on transcription initiation have so far shown that the promoter DNA is able to engage with $E\sigma^{70}$ to form the RPC and initiate DNA strand-separation but is not able to form a fully stable, transcriptionally competent RPo complex (Camara *et al.*, 2010). Recently, fluorescent RNAP beacon assays were performed to quantitatively assess how and to what extent the template DNA strands can access the main channel of RNAP. In these assays the T5N25 promoter was engineered to allow for the formation of a transcriptionally inactive, yet stable Gp2- $E\sigma^{70}$ -DNA quaternary complex. As expected, Gp2 decreased the overall affinity of the dw-DNA to the dw-DBC, however the effect of Gp2 was not limited to just the dw-DNA. Interestingly, Gp2 also decreased the binding affinities of the template and non-template strands of the promoter DNA around the active site (Mekler *et al.*, 2011b). With this in mind, this study intended to determine the point at which Gp2 prevents entry of the promoter DNA into the main channel. To this end, the minimal length of promoter DNA required for Gp2 inhibition to occur was established by assessing the effect Gp2 may have on the binding of DNA fragments to $E\sigma^{70}$ as judged by $E\sigma^{70}$ -DNA complex formation in EMSAs.

An initial EMSA was conducted using a ^{32}P -labelled version of the *lacUV5* (-45/+20)(-45/+20) probe as this is consistent with the template used in *in vitro* transcription assays which represents the minimal footprint of $E\sigma^{70}$. As expected the binding of $E\sigma^{70}$ to this probe is almost completely

abolished by Gp2, with <5% $E\sigma^{70}$ -DNA complexes surviving (Figure 5.6.1). In contrast, Gp2 has no effect on the binding of the (-45/-7)(-45/-7) probe (100% $E\sigma^{70}$ -DNA complex survival). This result confirms that Gp2 does not interfere with the interactions between $E\sigma^{70}$ and the consensus promoter DNA required for R_Pc formation. Therefore further EMSAs were performed with a series of *lacUV5* templates in which the downstream endpoints were extended in 1 bp increments. Results showed that there was only a slight effect on the -6, -5, -4 and -3 truncated probes (>95% $E\sigma^{70}$ -DNA complex survival). However, EMSAs conducted with probes truncated at positions -2, -1 and +1 showed approximately 50%, 85% and 95% reduction in $E\sigma^{70}$ -DNA complex formation in the presence of Gp2.

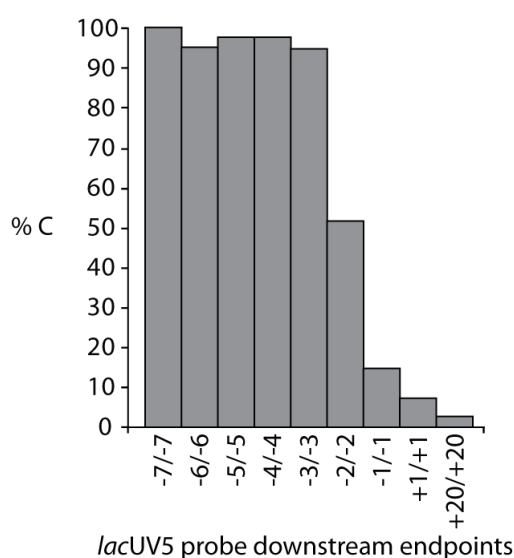


Figure 5.6.1 The effect of Gp2 on *lacUV5* binding to $E\sigma^{70}$.

Graph showing the percentage complex (% C) formed from DNA binding to $E\sigma^{70}$ in the presence of 1:2 molar excess of Gp2. The double stranded *lacUV5* probes with different downstream endpoints are indicated on the x-axis.

These results imply that $E\sigma^{70}$ sensitivity to Gp2 increases as the DNA approaches the active centre (around positions -2 and -1). Gp2 specifically prevents the entry of the DNA into the $E\sigma^{70}$ at the active centre presumably to prevent the positioning of the transcription start site base (+1) at the Mg^{2+} ions. Furthermore, unlike the interaction with the full length DNA probe (-45/+20)(-45/+20), the effect of Gp2 on the entrance of DNA around the active site occurs in an order of addition independent manner (data not shown). Presumably the $E\sigma^{70}$ -DNA contacts of the shorter probes are not extensive enough to successfully compete with the effect of Gp2. Therefore the

inhibitory effect of Gp2 extends beyond its site of binding at the dw-DBC and crucially attenuates the entrance of double-stranded DNA beyond the active site.

The prevention of DNA binding is unlikely to be a simple case of steric hindrance as the Gp2 binding site at the β' jaw is located quite distal from the active centre. In support of this view, the removal of the NCS as in Gp2^{Mut7} completely eliminates the antagonistic effect of Gp2 (Figure 5.6.2 lane 4) to the same extent as the non-binding mutant Gp2^{R56E}, which demonstrates that Gp2 binding *per se* does not prevent DNA entry around the active site.

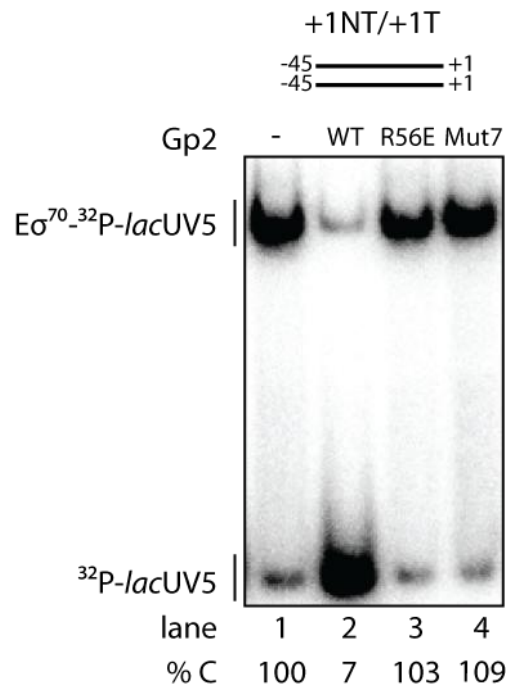


Figure 5.6.2 EMSA of 32 P-*lacUV5* binding to $E\sigma^{70}$ in the presence of Gp2^{WT}, Gp2^{R56E} and Gp2^{Mut7}.

An autoradiograph of a 4.5% (w/v) polyacrylamide native gel showing the effect of Gp2^{WT} (lane 2), Gp2^{R56E} (lane 3) and Gp2^{Mut7} (lane 4) on the binding of the (-45/+1)(-45/+1) *lacUV5* probe to $E\sigma^{70}$. The migration positions of unbound 32 P-*lacUV5* and $E\sigma^{70}$ - 32 P-*lacUV5* are as indicated. The percentage of $E\sigma^{70}$ - 32 P-*lacUV5* complex formation (% C) is calculated relative to a reaction conducted in the absence of Gp2 (lane 1) and is shown beneath each lane.

To investigate whether R1.1 also contributes to this inhibitory effect of Gp2, further EMSAs with the (-45/+1)(-45/+1) probe were conducted using $E\sigma^{70\Delta R1.1}$ (Figure 5.6.3). The effect of Gp2 was markedly reduced in the absence of R1.1 (compare lanes 2 and 4) and, as expected, partially restored when the isolated R1.1 domain was added *in trans* (lane 6). Therefore this Gp2-mediated attenuation of DNA entry into the active centre requires the NCS of Gp2 as well as the presence of the negatively charged R1.1 domain of σ^{70} . These results imply that the prevention of DNA

entry/positioning around the active centre is a result of the indirect and long-range antagonistic effects of Gp2.

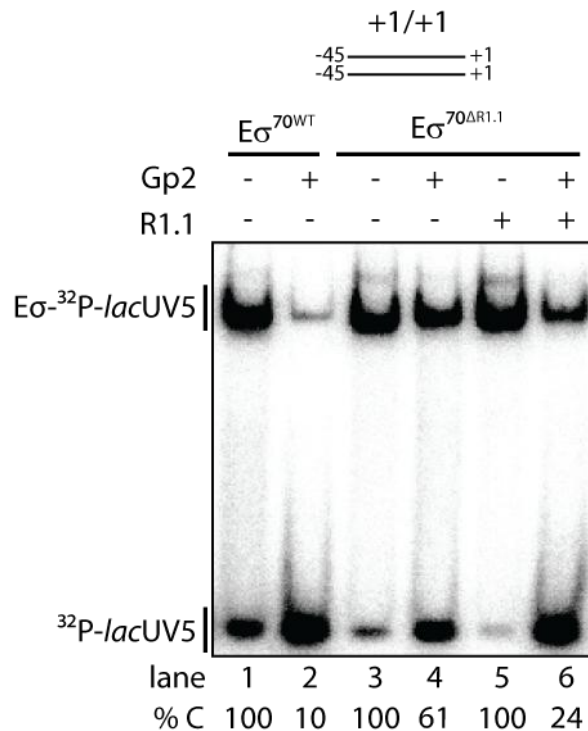


Figure 5.6.3 EMSA of ³²P-lacUV5 binding to Eσ⁷⁰ and Eσ⁷⁰ΔR1.1 in the presence of Gp2 and/or R1.1.

An autoradiograph of a 4.5% (w/v) polyacrylamide native gel showing the effect of Gp2 and/or R1.1 on the binding of the (-45/+1)(-45/+1) *lacUV5* probe to Eσ⁷⁰. The absence or presence of Gp2 and/or R1.1 in the reaction is indicated above each lane. The migration positions of unbound ³²P-lacUV5 and Eσ⁷⁰-³²P-lacUV5 are as indicated. The percentage of Eσ⁷⁰-³²P-lacUV5 complex formation (% C) is calculated relative to a reaction conducted in the absence of Gp2 and R1.1 (lane 1) and is shown beneath each lane.

5.7. The effect of Gp2 on accessibility to single-stranded DNA

Thus far it seems that Gp2 cooperates with R1.1 to prevent entry of ds-DNA around the active site. Next it was investigated whether Gp2 may also restrict access to single-stranded (ss) DNA. Additional EMSA experiments were conducted using fork junction probes with either non-template or template downstream extensions; (-45/+20)(-45/-7) and (-45/-7)(-45/+20) respectively. Recall that the (-45/+20)(-45/+20) promoter probe is sensitive to Gp2 inhibition whereas the (-45/-7)(-45/-7) probe is resistant to Gp2. EMSA experiments revealed that Gp2 had no detectable effect on the binding of Eσ⁷⁰ to either of the fork junction probes (Figure 5.7.1.A).

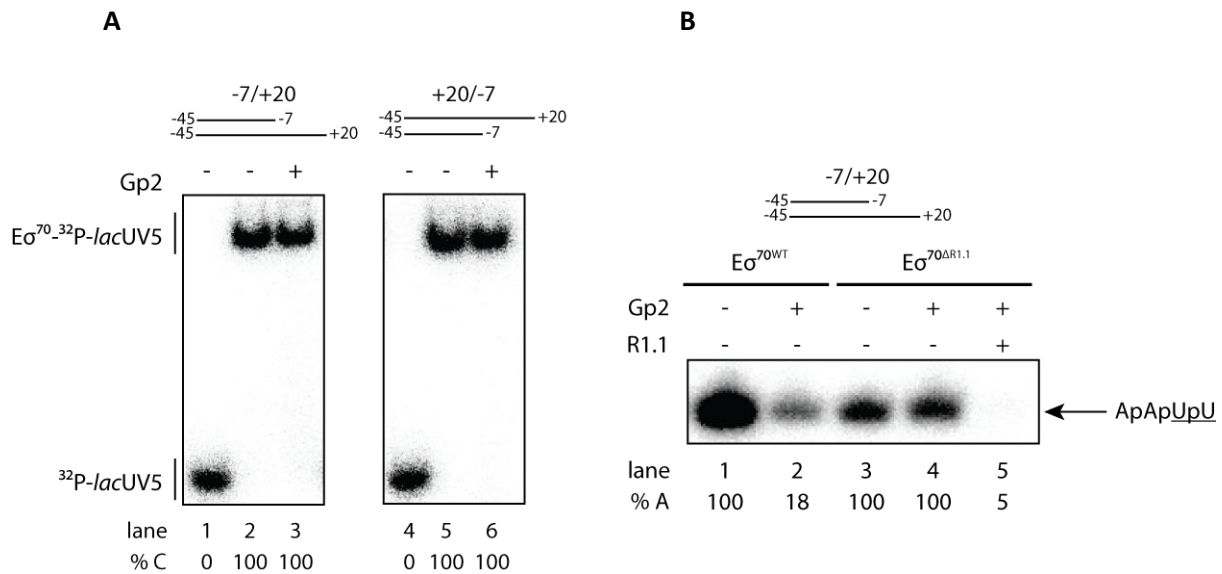


Figure 5.7.1 Effect of Gp2 and R1.1 on binding and productive engagement of ss-DNA at the active site.

A) An autoradiograph of a 4.5% (w/v) polyacrylamide native gel showing the effect of Gp2 on the binding of the $(-45/-7)(-45/+20)$ or $(-45/+20)(-45/-7)$ *lacUV5* probes to Eo⁷⁰. The binding reactions were conducted in the absence (lanes 2 and 5) or presence (lanes 3 and 6) of Gp2. The migration positions of unbound ³²P-*lacUV5* and Eo⁷⁰-³²P-*lacUV5* are as indicated. The percentage of Eo⁷⁰-³²P-*lacUV5* complex formation is calculated relative to a reaction conducted in the absence of Gp2 and R1.1 (lanes 2 and 5) and is shown beneath each lane. B) An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing the synthesis of the ApApUpU transcript as indicated by the arrow whereby the underlined nucleotides are ³²P-labelled. Transcription was performed by either Eo^{70WT} or Eo^{70ΔR1.1} from the *lacUV5* template in the absence (lanes 1 and 3) and presence of 1:4 molar excess of Gp2 (lanes 2, 4 and 5). R1.1 was added *in trans* using a 1:8 molar excess (lane 5). The percentage activity of each holoenzyme in the presence of Gp2 and/or R1.1 was calculated with respect to a reaction conducted in the absence of Gp2 (lanes 1 and 3) and are indicated beneath each lane.

Although, in these experimental conditions Gp2 does not affect fork junction binding, it does not necessarily mean that the ss-DNA section of the fork junction is correctly positioned within the RNAP. Therefore to test whether Eo⁷⁰ can productively engage with the ss-DNA, the $(-45/-7)(-45/+20)$ fork junction DNA, which retains the +1 transcription start site, was used as a template in an *in vitro* transcription assay. Results showed that synthesis of the ApApUpU transcript was effectively abolished in the presence of Gp2 (Figure 5.7.1.B lane 2). Therefore although the fork junction DNA is able to bind to Eo⁷⁰, it seems the ss-DNA extension of the template strand is not permitted to access the active site of RNAP. Remarkably this inhibitory effect of Gp2 on transcription from the $(-45/-7)(-45/+20)$ probe also seems to be entirely R1.1-dependent (compare Figure 5.7.1.B lanes 2, 4 and 5) as removal of R1.1 results in a complete loss of Gp2 inhibition but can be fully recovered when added back *in trans*. This result implies that potentially the primary role of R1.1 in Gp2-mediated inhibition of transcription is to prevent Eo⁷⁰ from productively engaging with the template strand.

5.8. Conclusions

The work in this chapter has further explored the effect of Gp2 on various aspects of transcription initiation. In particular, it has been described how the highly flexible and negatively charged R1.1 domain of σ^{70} contributes to the Gp2-mediated inhibition of $E\sigma^{70}$ -dependent transcription. Under the experimental conditions described here, Gp2 inhibition was reduced significantly in the absence of R1.1. The specificity of this result was tested by the *in trans* addition of R1.1, which notably restored the full Gp2 inhibitory effect. FA binding experiments shown in this chapter and previous chapters have illustrated that Gp2 can bind to the RNAP core independently of the associated σ -factor. In accordance with this observation, FA experiments also demonstrated that R1.1 is not required for the initial binding of Gp2 to RNAP. However, it remains to be known how the affinity of Gp2 for $E\sigma^{70}$ may change upon DNA binding. Since, Gp2 and the dw-DNA compete for binding to the β' jaw a potential role of R1.1 may be to sustain the high affinity that Gp2 has for RNAP upon the addition of DNA. This is a premise that may be experimentally explored in the future by fluorometric means.

It has been noted previously that despite its tight binding affinity for RNAP, Gp2 has no effect on transcription performed using the major alternative promoter specificity factor, σ^{54} . The reasons for the resistance of $E\sigma^{54}$ -dependent transcription to Gp2 remain unknown but presumably involve the dependency of this process on activation and energy consumption (Wigneshweraraj *et al.*, 2004). In this study, it was reported for the first time that Gp2 is able to partially inhibit $E\sigma^{38}$ which is responsible for the transcription of the stationary phase genes. Transcription by $E\sigma^{38}$ was notably less affected by Gp2 compared to $E\sigma^{70}$ but remarkably, complete inhibition was restored upon the *in trans* addition of the σ^{70} R1.1 domain. Therefore it seems that R1.1 provides a mechanism for the selective inhibition of σ^{70} -dependent transcription independent of specific promoter sequence interactions. This observation suggests that Gp2 may have evolved to predominantly target the expression of the major housekeeping genes which ultimately results in the most detrimental effect on the *E. coli* cell. The effect of Gp2 on the other five remaining holoenzyme variants remains to be determined but work is in progress to assess the global effect of Gp2 on *E. coli* transcription using an *in vivo* transcriptomics approach.

In vitro transcription experiments were conducted using the promoter-less MS in an attempt to investigate the dependence of Gp2 inhibition on σ^{70} interactions. Notably, Gp2-mediated inhibition of transcription from the MS was less effective than with the duplex DNA promoter template. This may be because the DNA-RNA hybrid represents a more stable conformation which

presumably can compete more effectively against the effects of Gp2. Nonetheless, results from the MS assay further confirmed that the dependency of Gp2 inhibition on σ^{70} was attributed predominantly to R1.1.

Following on from a recent study which investigated the effect of Gp2 on $E\sigma^{70}$ -promoter interactions it was further explored how Gp2 restricts DNA entry into $E\sigma^{70}$. Gp2 strictly prevents the entry of ds-DNA at position +1 which represents the active site of the enzyme. This antagonistic effect is completely dependent on the NCS of Gp2 which supports the view that this feature is a necessary determinant for inhibitory function (Sheppard *et al.*, 2011). Consistent with the previous *in vitro* transcription assays, this antagonistic effect of Gp2 is also partially dependent on R1.1.

Experiments conducted using fork junction probes revealed that in contrast with ds-DNA probes of equivalent length, Gp2 did not affect their binding to the $E\sigma^{70}$. However Gp2 was still able to attenuate proficient transcription from the extended template version of this probe demonstrating that Gp2 attenuates the correct positioning of the ss-DNA extensions. This effect of Gp2 may be as a result of conformational changes within the RNAP which include the restricted movement of the β' clamp. FRET experiments have shown that in the presence of Gp2, the conformation of the β' clamp only exists in a 'closed' to 'collapsed' state in which the width of the channel is not sufficient to accommodate either ds or ss-DNA (Chakraborty *et al.*, 2012).

In comparison to *in vitro* transcription assays with ds-DNA templates the effect of Gp2 on the fork junction template was entirely dependent on the presence of R1.1. This implies that the prevention of dw-DNA binding to the dw-DBC is mainly caused by Gp2 specific effects which are only partially influenced by the presence of R1.1. In contrast, attenuation of the positioning of ss-DNA, such as that of the 'transcription bubble' is exclusively a result of the effects of the presence of R1.1 during Gp2 inhibition. Due to the nature of this assay the positioning of the non-template strand could not be studied in the same context. Results from these fork junction experiments described here are in support of fluorometric binding assays which demonstrated that Gp2 reduced the binding affinity of fork junctions with template strand over-hangs and ss non-template segments close to the active site (Mekler *et al.*, 2011b). The overall finding that Gp2 prevents DNA entry and/or positioning around the active site is remarkable considering that both Gp2 and R1.1 bind at downstream locations quite distinct from this region. This observation indicates that Gp2 acts via R1.1 using long-range antagonistic effects.

During transcription initiation R1.1 acts as a DNA mimic to serve as a placeholder for the dw-DNA, which is ultimately displaced during RPo formation. Results thus far suggest that during Gp2-

mediated inhibition of transcription, R1.1 is required to prevent the entry and positioning of the DNA around the active site. Therefore it was proposed that Gp2 may be exploiting the function of R1.1 by prolonging its role as a DNA placeholder. To test whether Gp2 also functions to prevent the movement of R1.1, FRET assays were conducted to compare the position of R1.1 in the presence and absence of Gp2. The position of a fluorescein probe incorporated in R1.1 was measured relative to rifampicin that binds to RNAP between the upstream and downstream lobes of the β subunit. The degree of quenching of the fluorescein probe by rifampicin is inversely proportional to distance. As predicted, the quenching efficiency in the RPo was much lower than with $E\sigma^{70}$ and $E\sigma^{70}$ -Gp2, indicating that Gp2 antagonises the obligatory displacement of R1.1 so that its positioning is more reminiscent to that of the RPc (Figure 5.8.1) (James *et al.*, 2012). It remains to be known whether Gp2 restricts R1.1 movement by directly contacting the domain, this may be explored using crosslinking methods. The findings described here confirm that T7 Gp2 has cleverly evolved a mechanism to exploit the native function of R1.1 to prevent a stable RPo from forming.

As yet R1.1 has eluded structural biologists due to its highly acidic and flexible nature. However, by using Gp2 as a tool to secure the position of R1.1 in the dw-DBC may enable the successful determination of its structure in a context not too dissimilar to that of the native holoenzyme. Such a structure may also clarify the positioning of Gp2 and the orientation of its NCS which would provide invaluable insights into the overall molecular mechanism of inhibition.

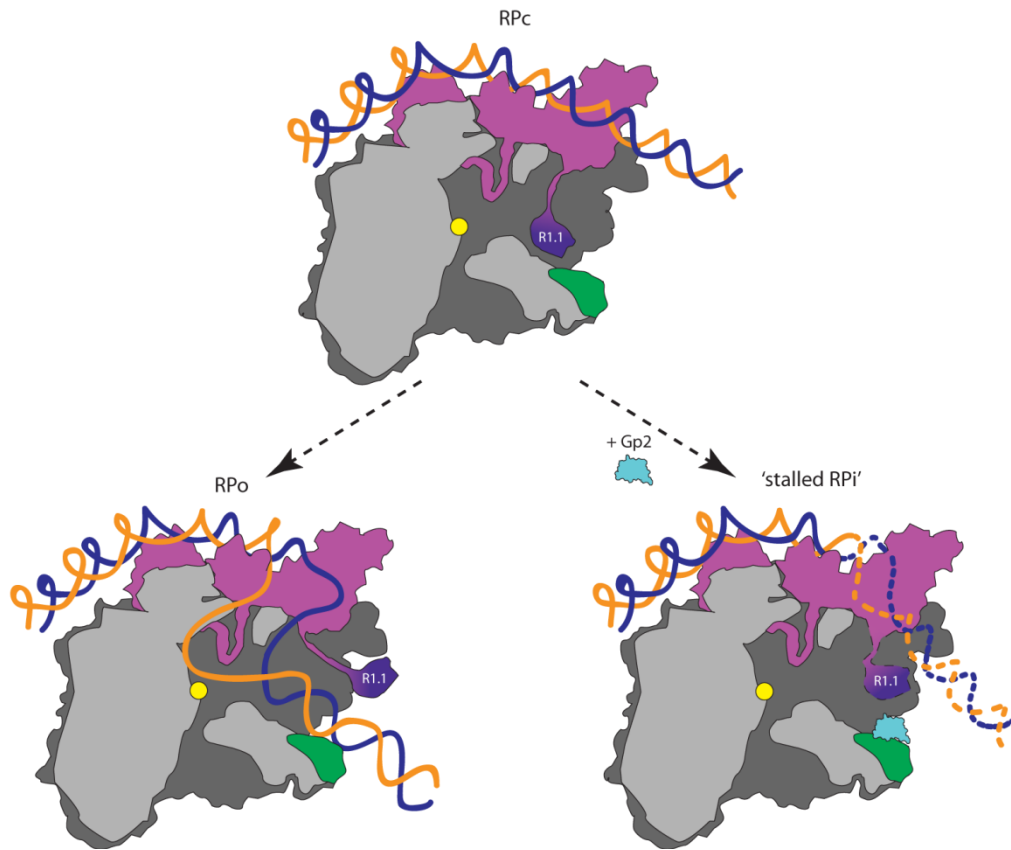


Figure 5.8.1 A model of the effect of Gp2 on R1.1 positioning in $E\sigma^{70}$.

A cartoon depicting how the binding of Gp2 (light blue) to the β' jaw (green) of $E\sigma^{70}$ (E is grey, σ^{70} is purple) antagonises the correct positioning of DNA (non-template is dark blue, template is orange) around the active site (yellow dot) and in the dw-DBC as well as prevents the obligatory movement of R1.1 required to form the RPo. Adapted from (James *et al.*, 2012).

CHAPTER 6

A PRELIMINARY CHARACTERISATION OF THE XP10 PHAGE-ENCODED, RNAP INHIBITOR, P7

6.1. Introduction

P7 was originally isolated from Xp10 bacteriophage-infected *Xanthomonas oryzae* whole cells extracts. When assessed for transcriptional activity, the Xp10-infected cell extracts were able to attenuate RNA synthesis relative to extracts from uninfected cells (Liao and Kuo, 1986; Nechaev *et al.*, 2002). This inhibitory activity was attributed to a 7 kDa protein, P7, which co-purifies with the *X. oryzae* RNAP by binding to the β' subunit and therefore is likely to be responsible for the host transcription shut-off during Xp10 infection (Nechaev *et al.*, 2002). Recombinant P7 was found to inhibit *in vitro* transcription by attenuating RPo formation, as it has no effect on pre-formed RPo complexes. In addition, P7 seems to operate as a bi-functional inhibitor of RNAP as it also displays characteristics of an anti-terminator (Nechaev *et al.*, 2002).

The inhibitory effects of P7 seem to be limited to *X. oryzae* as it has no effect on transcription from *E. coli* E σ^{70} . This specificity of inhibition suggested that P7 may bind to an evolutionary variable region of the β' subunit. The location of P7 binding to the *X. oryzae* β' subunit was determined by far-Western blotting and bacterial two-hybrid based interaction analyses. The results indicated that the precise site of P7 binding to the *X. oryzae* β' subunit mapped to amino acid residues 6-9 which comprise the N⁶L⁷F⁸N⁹ sequence. The same positions in *E. coli* possess a K⁶F⁷L⁸K⁹ sequence and therefore are likely to be responsible for conferring resistance to P7. Moreover, a mutant form of *E. coli* RNAP with β' amino acid residues 6-9 corresponding to that of the *X. oryzae* RNAP is sensitive to P7 inhibitory activity and therefore serves as a useful tool in *in vitro* studies. The β' N-terminal residues are missing in the *Thermus aquaticus* and *Thermus thermophilus* RNAP structures and the *E. coli* insertion element of this region has, as yet, not been resolved. Therefore it is not currently possible to pinpoint the precise location of P7 binding in the context of the currently available structures of bacterial RNAP. However, the general region of P7 binding is deduced to be in the vicinity of three major functional mobile modules of RNAP; the β' Zn-binding element, the β' zipper and the β' coiled-coil as well as σ domains 2 and 3 (Figure 6.1.1). The β' Zn-binding element and the β' zipper are parts of the RNA exit channel, which forms interactions with the nascent RNA. These two elements are required for efficient termination by recognition of hairpin loop structures in the RNA (Epshtein *et al.*, 2007b; King *et al.*, 2004). Therefore P7 may act as an anti-terminator by influencing either or both of these elements and their interactions with the RNA to decrease termination efficiency. Additionally the β' zipper also interacts with the promoter spacer region to facilitate the formation of a stable RPc complex (Yuzenkova *et al.*, 2011) and the β' Zn-binding element is positioned to perhaps recognise the -35 promoter element (Murakami *et al.*,

2002b; Vassilyev *et al.*, 2002). Whereas the β' coiled-coil represents a major σ -factor docking site and specifically induces σ^{70} R2.2 to contact the non-template strand around the -10 promoter element region (Young *et al.*, 2001). Therefore it is also conceivable that P7 could influence the function of one of these three mobile modules or regions of the σ -factor to antagonise holoenzyme-promoter DNA interactions required for proficient transcription initiation.

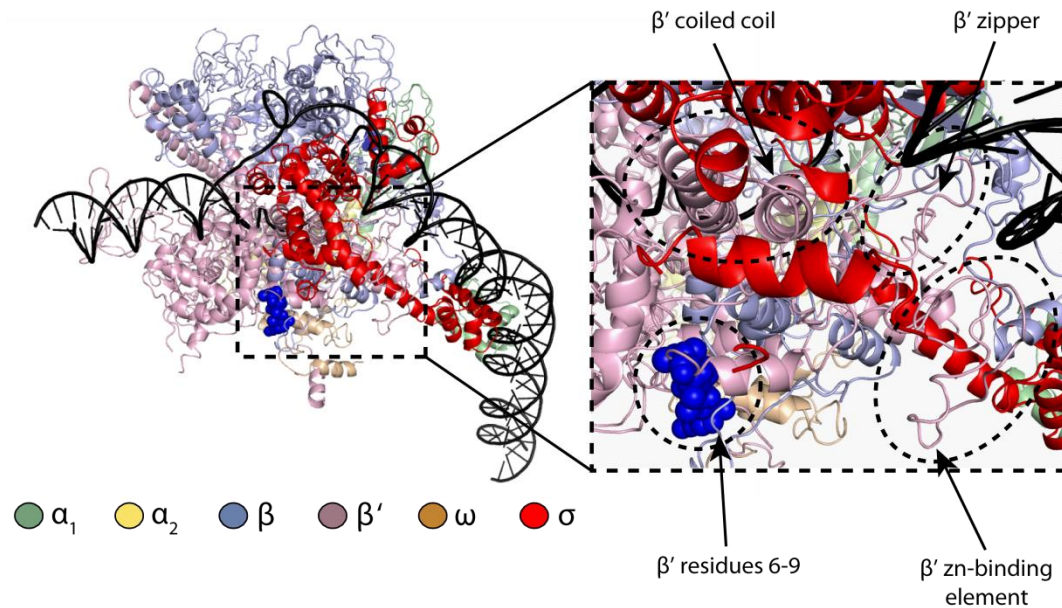


Figure 6.1.1 Structural context of P7 binding.

Ribbon representation of *Thermus aquaticus* E σ^A with the β' (pink), β (purple), α_1 (green), α_2 (yellow), ω (brown) and the σ -factor (red) individually coloured. The predicted location of P7 binding at β' amino acid residues 6-9 is highlighted in blue. The β' coiled coil, zipper and Zn-binding element mobile modules are indicated with arrows.

Although the site of P7 binding to the *X. oryzae* RNAP has been determined, it is not known which amino acid residues in P7 contribute to the binding to RNAP since the structure of P7 is currently unresolved.

6.2. Objectives

The overall purpose of this study is to provide an initial characterisation of the Xp10 P7 protein, with a particular focus on the mechanism of inhibition of transcription initiation. This study

will include a comprehensive mutagenesis screen of the P7 protein to determine which amino acid residues are important for binding function and how their function might be influenced by their structural context. Biochemical assays will be employed to investigate the mechanism by which P7 inhibits transcription initiation. The results from this study will provide further insights into the overall molecular mechanism employed by P7 during the inhibition of transcription initiation and form the basis for a future in depth analysis of P7 structure-function. This initial characterisation of the mechanism of action of P7 will include the following aims:

- Establish and optimise a screen for assessing the binding of a single alanine mutant library of P7 to *X. oryzae* RNAP.
- Determine the structure of P7 in collaboration with Prof. Steve Matthews from the CSB, Imperial College London.
- Assess the relative binding ability of the single alanine mutants in the context of the structure.
- Investigate the mechanism by which P7 inhibits transcription initiation.

6.3. Determination of P7 residues important for binding.

A previous study mapped the P7 interaction site on *X. oryzae* RNAP to the N⁶L⁷F⁸N⁹ amino acid residues at the N-terminal of the β' subunit (Yuzenkova *et al.*, 2008). The aim in this study was to investigate this interaction further, by determining which residues in P7 are important for binding to RNAP. To assess the contribution of each individual amino acid residue of P7 in binding to *X. oryzae* RNAP it was necessary to generate a mutant library of P7 proteins which could be used in a comprehensive interaction screen. The substitution of amino acid residues with alanine is a common approach to assessing whether the properties of an amino acid residue contribute to function as alanine itself possesses one of the most simplistic side chains. To conduct *in vitro* experiments a library of 70 recombinant P7 proteins (omitting the 3 native alanines and the methionine) would need to be over-expressed and purified. However, purified recombinant P7^{WT} can be unstable as it has a tendency to precipitate out of solution, therefore it needs to be carefully handled and stored to prevent inconsistencies between experiments. For that reason, the purification of a library of P7 mutant proteins all with potentially differing degrees of solubility and stability may yield variable results. Instead, it was decided that an *in vivo* approach would be more manageable and consequently produce more reliable and reproducible results. Thus a bacterial two-hybrid assay was used to measure the interaction between the alanine mutant P7 library and RNAP.

Conventional bacterial two-hybrid systems such as the adenylate cyclase system rely on the chance interaction of the protein binding partners in order to reconstitute the adenylate cyclase enzyme and synthesise cyclic adenosine monophosphate (cAMP). cAMP subsequently binds to the cAMP receptor protein (CRP) to activate transcription of the *lacZ* gene. Here, an activator dependent, transcriptional bacterial two-hybrid system was employed; which directly couples the binding of the encoded proteins to the transcriptional output of the *lacZ* gene. This system makes use of the well defined properties of the RNAP α -CTD and the bacteriophage λ CI protein (λ CI) (Dove and Hochschild, 2004). Contacts between the λ operator (O_R2) bound λ CI protein and the α -CTD of the *E. coli* RNAP are required for efficient activation of transcription (Kedzierska *et al.*, 2007; Dove *et al.*, 1997)(Figure 6.3.1.A). This system requires two expression vectors; the prey plasmid, pBR α in which the P7 gene replaces the *E. coli* RNAP α -CTD sequence by fusion to the linker of the NTD of the *rpoA* gene and the bait plasmid, pAC λ CI, which contains an N-terminal fragment of the *X. oryzae* *rpoC* gene fused to the C-terminus of the λ CI gene. Upon IPTG induced expression, in the KS1 reporter strain (bearing the *placO_R2-62* promoter), an interaction between the P7 protein and the *X. oryzae* RNAP β' subunit fragment, generates an α - λ CI chimaera which is bound at the λO_R2 operator located 62 bps upstream of the transcription start site and can successfully activate transcription of the *lacZ* gene (Figure 6.3.1.B). *LacZ* encodes for the enzyme, β -galactosidase which can hydrolyse the lactose analogue, ortho-nitrophenyl- β -D-galactopyranoside (ONPG) into galactose and ortho-nitrophenol of which the later is yellow in colour and therefore provides a quantitative read-out by spectrophotometer analysis at 420 nm (refer to section 2.4.9). The degree of binding is directly proportional to the intensity of the yellow colour, therefore provides a means for not only identifying a binding interaction but also the relative strength of the interaction compared to other binding partners. The single alanine mutant library of P7 proteins was constructed by site-directed mutagenesis using the pBR α -P7^{WT} plasmid as a template (refer to section 2.3.11). Although it had been previously established that only residues N⁶L⁷F⁸N⁹ of the *X. oryzae* RNAP β' -subunit are sufficient for P7 binding, a more substantially sized peptide is required for such interaction assays. The previous P7 interaction study observed that P7 binding to the first 85 amino acids of *X. oryzae* RNAP produced one of the highest β -galactosidase read-outs, therefore this fragment would be ideal for generating a wide range over which to assess weaker binding P7 mutants. In order to compare the binding abilities of each P7 mutant under the exact same experimental conditions a 96-well plate, high-throughput format was established and optimised for the simultaneous quantification of the interaction between the *X. oryzae* RNAP β' N-terminal fragment and the P7 mutant library.

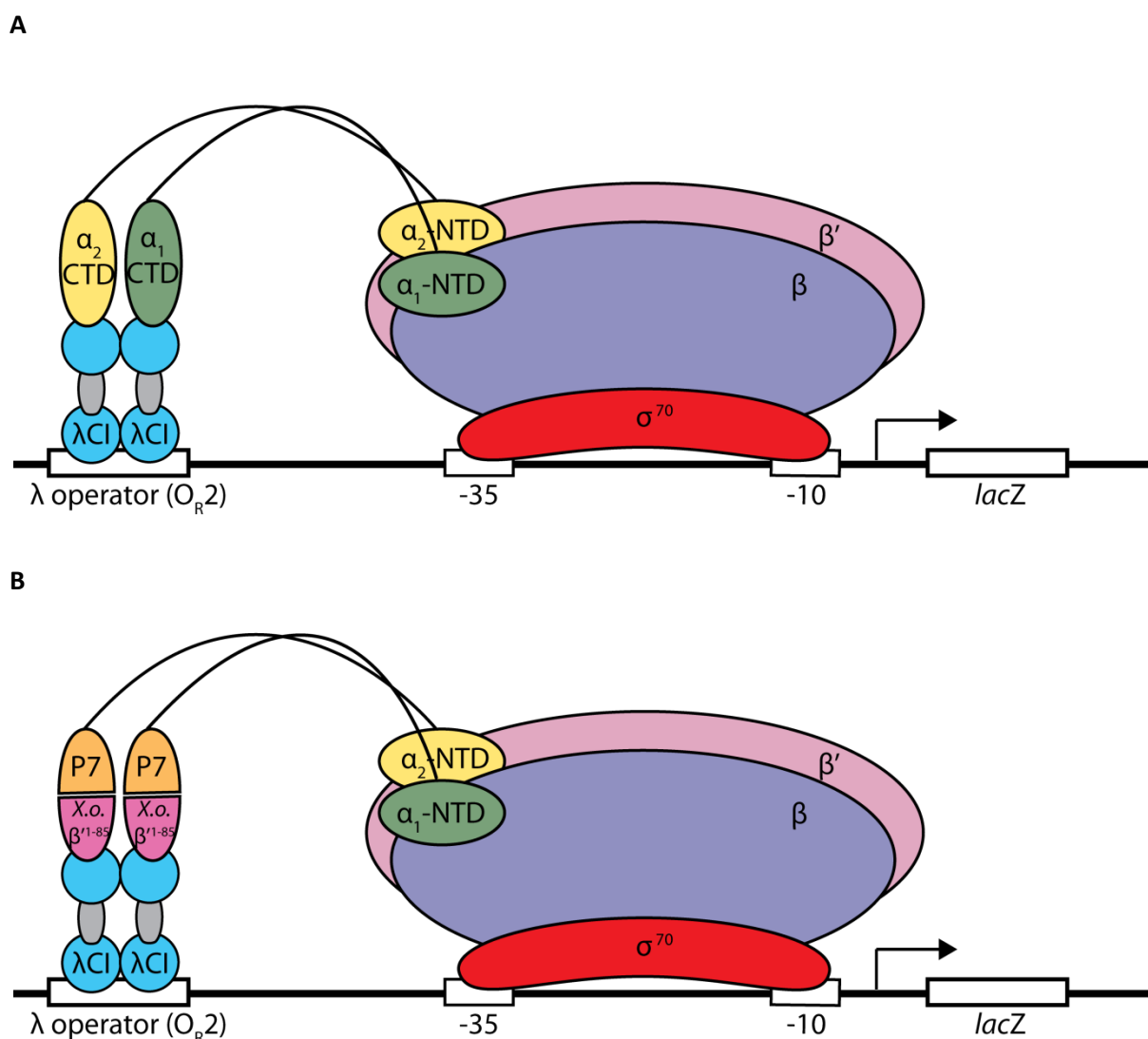


Figure 6.3.1 Activator dependent transcriptional bacterial two-hybrid system.

(A) A schematic illustrating the arrangement of the transcription machinery at the λ operator (O_{R2}) of the KS1 reporter strain. *E. coli* RNAP binds to the -35 and -10 promoter elements via the σ^{70} subunit. The λ CI proteins bind as a dimer to the λ operator (O_{R2}) positioned 62 bp upstream from the transcription start site of the *lacZ* gene. The λ CI proteins activate transcription from the *lacZ* promoter by contacting the α -CTDs. (B) The P7 protein and the *X. oryzae* β' ¹⁻⁸⁵ fragment are fused to the α linker-CTD and the λ CI-CTD respectively. Contact between P7 and *X. oryzae* β' ¹⁻⁸⁵ activates transcription of the *lacZ* gene.

The binding ability of each P7 alanine mutant to the *X. oryzae* RNAP β' fragment can be represented in terms of Miller Units and compared to that of P7^{WT} (Figure 6.3.3). From this screen, each amino acid residue of P7 was categorised according to the binding ability of its respective alanine mutant (Table 6.3.1). Class I mutants possess greater than 75% β -galactosidase activity, which represents binding in a similar range to that of P7^{WT}. This relatively high binding suggests that the respective amino acid residues of these point mutants are less likely to be important for P7 binding. Class II and class III mutants exhibit 51-74% and 26-50% β -galactosidase activity

respectively, therefore represent amino acid residues which may partially contribute to function but ultimately are non-essential for overall binding. Whereas, the most interesting residues are the class IV mutants which have retained less than 25% β -galactosidase activity, indicative of very weak binding. This implies that the class IV amino acid residues are the most important for P7 binding to RNAP. The overall analysis of the mutagenesis screen revealed that P7 is fairly sensitive to mutation as almost 40% of the tested proteins are classified as class IV mutants, whereas only 23% of the proteins are class I mutants. The class IV mutants are also well distributed across the amino acid sequence which suggests that the surface interface, between P7 and *X. oryzae* RNAP may be quite extensive. However, there are two clusters of class IV residues within the amino acid sequence; the GSVLTVKV residues at positions 19-26 and the FDR residues at positions 39-41. It is possible that these residues function together to form crucial interactions with *X. oryzae* RNAP.

Class	% β -galactosidase Activity	Number of P7 Mutant Proteins
I	>75%	16
II	51-74%	17
III	26-50%	9
IV	<25%	27

Table 6.3.1 Classification of P7 mutant proteins.

One limitation to this type of assay is that low β -galactosidase activity is not necessarily a result of binding ability *per se*. It may for instance, be a consequence of low expression levels of the mutant P7 protein. Reliable results from this screen depend on equal expression levels and protein turnover throughout the P7 mutant library. To experimentally address whether the class IV mutants were all present at the same levels as the other mutant proteins, Western blot analysis was performed from cell lysates extracted from the bacterial two-hybrid screen. To this end, polyclonal antibodies were raised against peptides of the P7 sequence (VEEDFDRADYGSDPEF and SYLKRNGGIKDLTKV) and subsequently affinity purified using the recombinant P7 protein and then purified against a non-P7 expressing cell background to eliminate cross-reactivity (refer to sections 2.5.1, 2.5.2, and 2.5.3). All cell lysates tested did exhibit detectable amounts of the α -NTD-P7 mutant fusion proteins (Figure 6.3.3 bottom panel). However the intensity of some of the class IV mutant bands does differ significantly relative to other class IV mutants as well as a class III mutant control and P7^{WT}. Since the P7 mutant plasmids were all created from the same parental plasmid using the same technique it is unlikely that the difference in expression is due to the transcriptional or

translational differences, but perhaps more the stability of the mutant proteins. It is conceivable that the observed decreases in β -galactosidase activity may be caused by aberrant folding of the mutant proteins which could result in protein degradation (as judged by Western blotting). In addition, a loss in interaction could also be due to the formation of an anomalous protein structure whereby the surface interface may either be masked or no longer formed correctly. Therefore the results from the screen may not necessarily reflect directly on the contribution of specific residues for binding to *X. oryzae* RNAP but potentially on their requirement for the correct folding or maintenance of the structural integrity of the protein which consequently may be needed to successfully interact with the *X. oryzae* RNAP. Therefore, from the Western blot analysis it seems that at least four residues; V²⁶, F³⁹, D⁴⁰ and R⁴¹ are necessary for maintaining structural integrity whereas F⁴, I⁷, G⁹, Y¹⁰, N¹², F¹⁴, S²⁰ and V²⁴ exhibit levels most comparable to P7^{WT} and are therefore more likely to be *bona fide* determinants for binding.

In addition to the two-hybrid screen, a sequence alignment of known P7 orthologs was conducted to inform upon which amino acid residues were evolutionary conserved and therefore functionally important. However, Xp10 P7 only shares sequence homology with two other putative proteins, from the OP1 and Xop411 *Xanthomonas* phage (Figure 6.3.2). The alignment demonstrates that the overall sequence of P7 is highly conserved which explains why P7 is quite sensitive to mutation but provides little insight into which residues may be more important for function.

```

Xp10      MNEFTQISGYVNAFGSQRGSVLTVKVENDEGWTILVEEDFDRADYGSDPEFVAEVSSYLKRNNGGIKDLTKVLTR 73
OP1       MNEFTQISGFINAFGSQRGSVLTVKVENDAGWALIEEDFDRAEYGS DPEFVAEVSSFIKRNGGIKDLTMVISR 73
Xop411    MNEFSQIPGFVNAFGLQDGSKLVVKVEDSRGWNVAEEEF SREDYNSDPEFVAEVSSFIKRNGGIKDLTKVVS R 73
          ****:*:.*:;**** * ** *.****;. ** : **:*.* :*.*****;:***** *;:*

```

Figure 6.3.2 Sequence alignment of P7 orthologs.

Alignment of the amino acid sequences of all known P7 orthologs in the single-letter amino acid codon. A * represents an absolutely conserved amino acid residue, : indicates conservation between groups of strongly similar properties and . indicates conservation between groups of weakly similar properties. P7 class I residues are coloured green, class II blue, class III orange and class IV red.

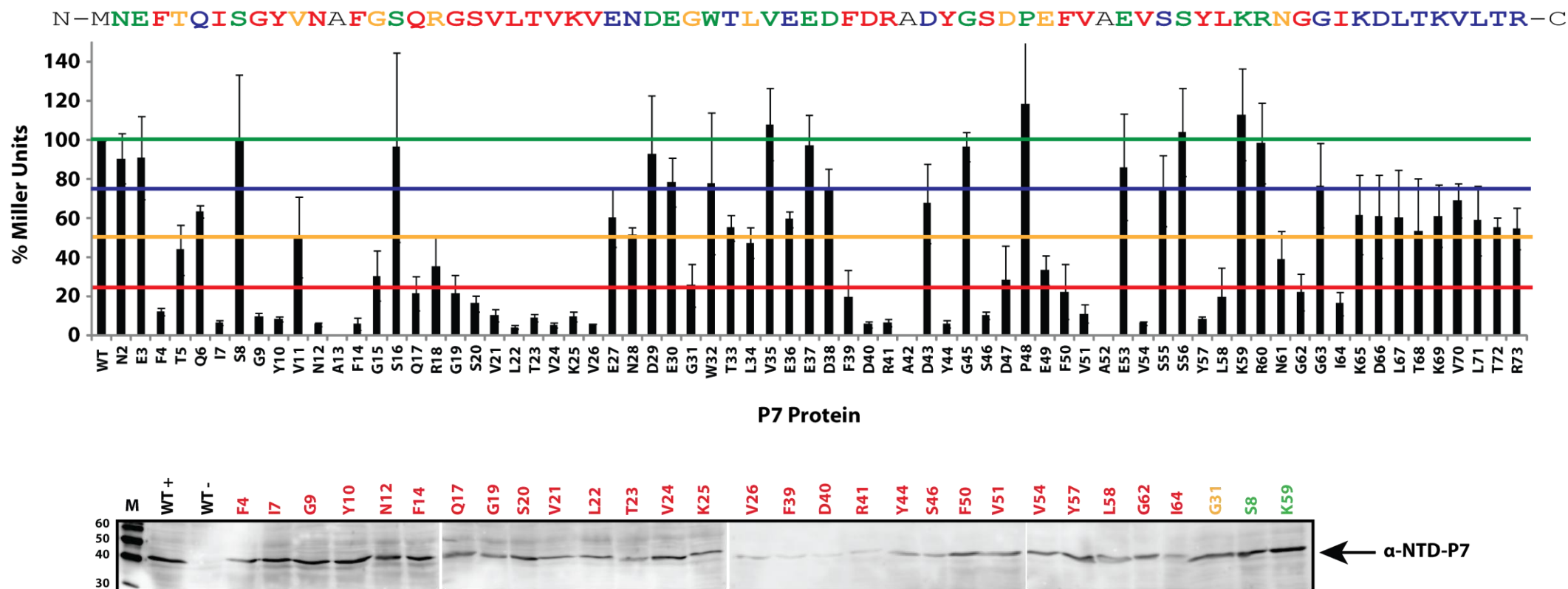


Figure 6.3.3 Bacterial two-hybrid β -galactosidase screen of P7.

Top panel: Single alanine mutant versions of P7 (x-axis) were used in a bacterial two-hybrid β -galactosidase assay. The binding of P7^{WT}-*E. coli* RNAP α -NTD fusion to an *X. oryzae* RNAP β' fragment-bacteriophage λ CI fusion successfully activates transcription of the *lacZ* gene encoding β -galactosidase. The percentage Miller Units (y-axis) represent β -galactosidase activity as a result of mutant P7 fusion proteins binding to the *X. oryzae* RNAP β' fragment fusions compared to that of P7^{WT}. The P7 mutant proteins are classified into four categories; class I $\geq 75\%$ Miller Units (green), class II = 51-74% Miller Units (blue), class III = 26-50% Miller Units (orange) and class IV < 25% Miller Units (red). The single codon amino acid sequence of P7 is shown above and colour-coded according to class. Bottom panel: Western blot analysis of the expression of the class IV mutant proteins. The molecular weights of the protein marker (M) are indicated in kDa. WT+ represents Western blot analysis of P7^{WT} under IPTG inducing conditions, whereas WT- is Western blot detection conducted under non-inducing conditions. The class IV amino acid residues are displayed in red, whereas a borderline class IV/III residue (G31) is in orange and class I controls in green. The migration of *E. coli* RNAP α -NTD-P7 fusions is indicated by an arrow.

6.4. Structural determination of P7

To provide further insights into the potential functional contribution of specific class IV residues the structure of P7 was determined. The results of the mutagenesis screen were initially interpreted in the secondary structure of P7, which was determined using two types of prediction software (Figure 6.4.1). The Chou & Fasman Secondary Structure Prediction (CFSSP) was used to predict which residues are likely to form turns in the protein. A total of 11 turns were predicted, which included a turn at the beginning of the $G^{19}S^{20}V^{21}L^{22}T^{23}V^{24}K^{25}V^{26}$ class IV cluster and a turn at the end of the $F^{39}D^{40}R^{41}$ cluster. Jpred3 was used to predict which residues are likely to form secondary structure features such as α -helices and β -strands. At least 2 α -helices were predicted, comprising residues 50-59 and residues 65-71. A β -strand was predicted to comprise residues 21-26 which encompasses most of the $G^{19}S^{20}V^{21}L^{22}T^{23}V^{24}K^{25}V^{26}$ cluster and a second shorter β -strand formed by residues 33-36.



Figure 6.4.1 Secondary structure prediction of P7.

A summary of the secondary structure prediction of the P7 protein was created using CFSSP and Jpred3. The original P7 single codon amino acid sequence of P7 is shown at the top. CFSSP was used to predict the turns in the protein (T). Jhmm and Jpssm represent a Jnet hmm profile and a Jnet PSIBLAST pssm profile prediction respectively. Jnet provides a final secondary structure prediction of the α -helices (H) and β -strands (E). Lupas predicts coiled coil elements with 14, 21 and 28 window sizes. Jnet was also used to predict buried residues (B) with 0, 5 or 25% solvent exposure. The reliability of the prediction ranges from 0 to 9, with 9 being the highest accuracy.

No coiled coil elements were predicted as it is likely that the P7 protein is too small for such composite features. Amino acid residues were also predicted to be buried in the hydrophobic core of the protein with varying degrees of exposure, for instance many of the hydrophobic and aliphatic residues including isoleucine, leucine and valine are predicted to have only 5% solvent exposure whereas V54 is likely to be deep within the hydrophobic core as it is predicted to have 0% solvent exposure.

To obtain a tertiary structure of P7, the recombinant protein was over-expressed from pET28b⁺ and Ni²⁺ affinity purified (refer to sections 2.4.1 and 2.4.4) (Figure 6.4.2) and the structure was resolved by Prof. Steve Mathews and Dr. Bing Liu using NMR spectroscopy.

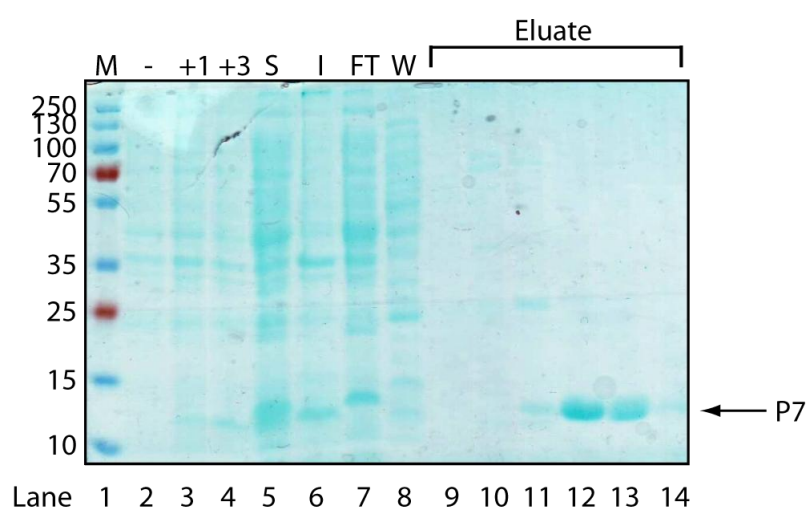


Figure 6.4.2 Nickel affinity chromatography of P7.

Image of a 12.5% + 15% (w/v) SDS-PAGE gel from the nickel affinity purification of recombinant Xp10 P7. M = Marker, - = without IPTG induction, +1 = 1 hr after IPTG induction, +3 = 3 hrs after IPTG induction, S = soluble fraction, I = insoluble fraction, FT = flow-through, W = low imidazole wash. The molecular weight of the protein markers is shown in kDa and the migratory position of P7 is indicated with an arrow.

The high resolution, 3-D NMR structure of P7 revealed the protein to comprise 3 β -strands, of which the central strand runs anti-parallel, followed by a turn and an α -helix (Figure 6.4.3).

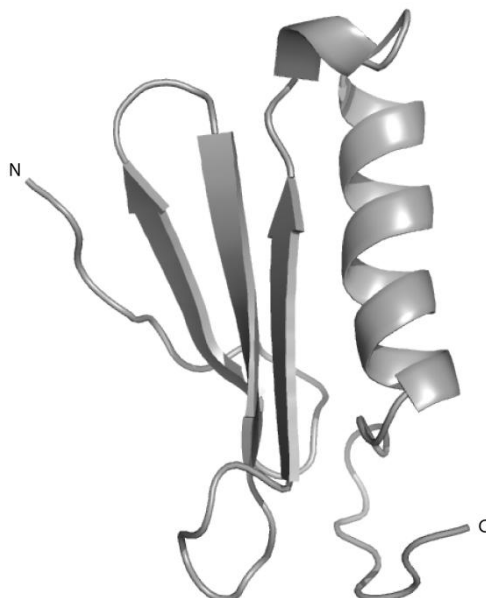


Figure 6.4.3 NMR structure of P7.

An NMR structure of P7 shown in the ribbon representation, possessing a $\beta_1\beta_2\beta_3\alpha_1$ topology whereby the central β -strand runs anti-parallel.

The secondary structure predictions were overlaid onto the NMR structure to validate the accuracy of the predictions (Figure 6.4.4). The overlay showed that the CFSSP software correctly predicted 7/11 of the turns in the protein including the turn in the F³⁹D⁴⁰R⁴¹ cluster, whereas Jpred3 accurately predicted the α -helix and the central β -strand which is comprised by the V²¹L²²T²³V²⁴K²⁵V²⁶ portion of the larger class IV cluster (Figure 6.4.4, left). In addition, almost all of the aliphatic and hydrophobic amino acid residues predicted to possess $\leq 5\%$ solvent exposure are in fact located in the central region of the protein and are positioned to project their hydrophobic side chains into the core (Figure 6.4.4, right).

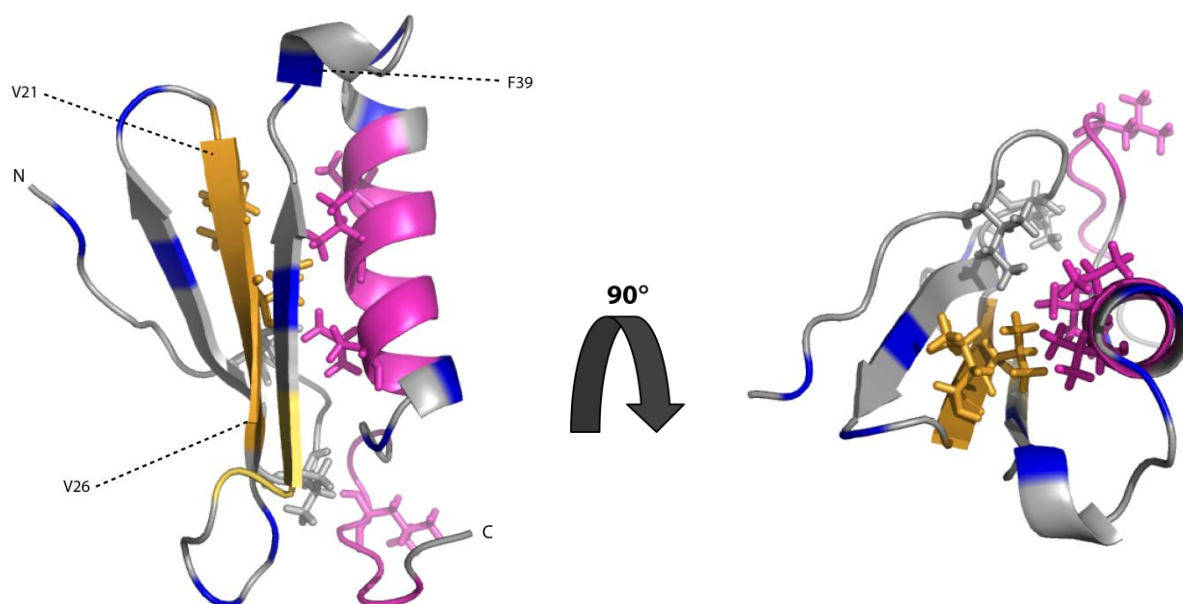


Figure 6.4.4 NMR structure of P7 with the secondary structure prediction overlay.

Left: An NMR structure of P7 as in Figure 6.4.3 but with the secondary structure prediction overlay; α -helices (pink), β -strands (orange), turns (blue) and the amino acid residues with $\leq 5\%$ solvent exposure are shown in the stick representation. The F39 and V21-V26 residues which form part of the class IV clusters are indicated. Right: Top view orientation of the P7 structure displaying the location and projection of the hydrophobic residues with $\leq 5\%$ predicted solvent exposure (coloured as in the left figure).

The availability of a high resolution 3-D structure of P7 allowed for the positioning of the class IV residues in a structural context with the aim of deciphering the potential roles of the individual residues as well as the clusters. As in the primary structure the class IV residues are dispersed throughout the secondary structure features (Figure 6.4.5, left). It seems that many of the class IV residues may be necessary for structural integrity. For instance, the $G^{19}S^{20}V^{21}L^{22}T^{23}V^{24}K^{25}V^{26}$ cluster which forms the β_1 - β_2 loop and β_2 strand and is centrally positioned in the protein along with many of the side chains of the hydrophobic residues of the β_2 -strand appear to project into the core of the protein rather than projecting outwards to form a potential interaction with RNAP (Figure 6.4.5 right). The class IV amino acid residues F50, V51, V54 and Y57 which reside in the α -helix are also hydrophobic and project into the core of the protein (Figure 6.4.5, right). Similarly, the smaller $F^{39}D^{40}R^{41}$ cluster forms a linker region between the β_3 -strand and α -helix which may be required for correct protein folding.

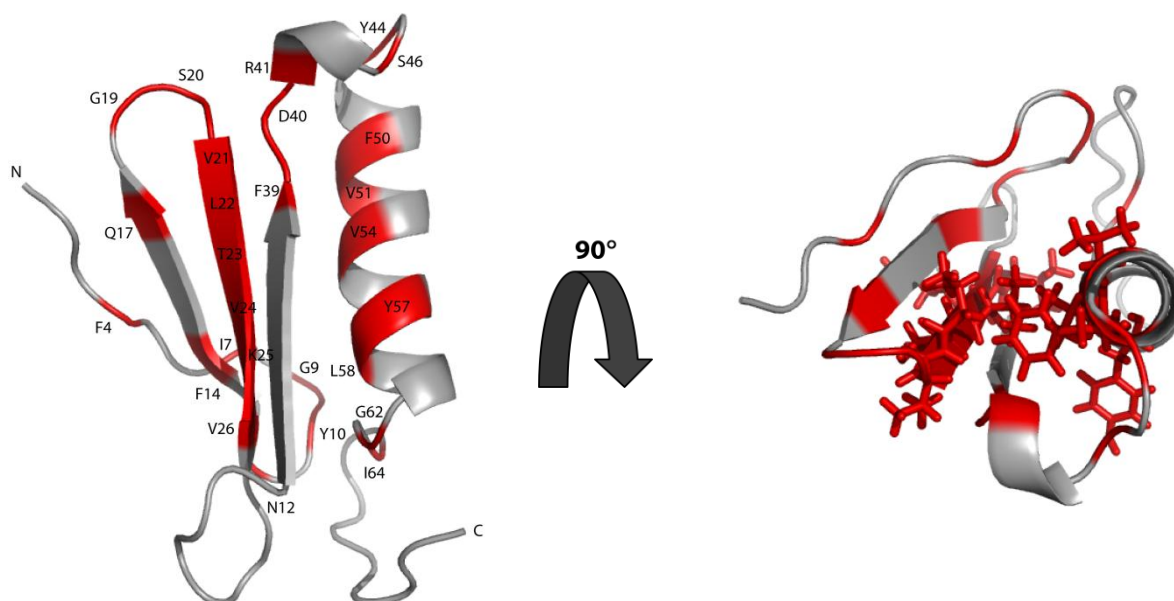


Figure 6.4.5 NMR structure of P7 with mapping of the class IV amino acid residues.

Left: An NMR structure of P7 as in Figure 6.4.3 but with the class IV residues labelled and highlighted in red. Right: Top view orientation of the P7 structure displaying the location and projection of the side chains of the hydrophobic class IV residues within the β_2 strand (L²²–T²³–V²⁴–K²⁵–V²⁶) and the α -helix (F⁵⁰, V⁵¹, V⁵⁴ and Y⁵⁷).

Therefore it seems that ~12 of the class IV residues either form part of the hydrophobic core of the protein and/or are required for preserving vital structural features. This means that ~15 class IV residues remain, which are more likely to be major determinants for binding to RNAP or forming part of the overall surface interface. These include amino acid residues, F⁴, I⁷, G⁹, Y¹⁰, N¹², F¹⁴, Q¹⁷, G¹⁹, S²⁰, V²¹, Y⁴⁴, S⁴⁶, L⁵⁸, G⁶² and I⁶⁴.

6.5. Mechanism of inhibition of transcription initiation by P7

A previous preliminary biochemical analysis of P7 established that it inhibits transcription initiation from -10/+35 promoters by preventing RPo formation (Nechaev *et al.*, 2002). To continue the investigation into the molecular mechanism of inhibition of transcription initiation by P7, a mutant form of the *E. coli* RNAP was created for *in vitro* biochemical experiments as this is the best studied bacterial RNAP. This mutant *E. coli* RNAP contains amino acid substitutions at β' subunit residues 6-9 (K⁶F⁷L⁸K⁹) so that it resembles that of the *X. oryzae* RNAP (N⁶L⁷F⁸N⁹) and therefore

confers P7 sensitivity (Yuzenkova *et al.*, 2008). The *in vitro* activity of P7 was tested using the P7 sensitive version of the *E. coli* RNAP, which is referred to as E^{NN}.

To determine the stage of transcription which is targeted by P7, order of addition *in vitro* transcription assays were performed using a homoduplex *lacUV5* DNA template and a heteroduplex *lacUV5* DNA template with mismatches between the -10 and +3 sites, which mimics the conformation of the DNA in the RPo. Order of addition experiments were conducted with each template by adding P7 before or after the DNA. Additionally, as P7 was previously shown to restrict the σ^{70} interdomain movements of σ R2.4 and R4.2 (Nechaev *et al.*, 2002), order of addition experiments were also performed to test if P7 had an antagonistic effect on E^{NN} σ^{70} formation or stability. Results from experiments using the homoduplex probe revealed that P7 successfully inhibits transcription when added to the reaction prior to the DNA but not after (compare Figure 6.5.1 lanes 2 and 3 to 4). This suggests that P7 is unable to inhibit transcription initiation if the RPo has been pre-formed, which confirms that P7 must target a stage prior to RPo formation as has been suggested previously (Nechaev *et al.*, 2002). Parallel reactions using the heteroduplex template showed that P7 can also still inhibit pre-opened DNA by up to 70% (compare Figure 6.5.1 lane 5 to 6 and 7), therefore it is unlikely to antagonise the process of DNA strand-separation *per se* but may still partially effect conformational changes in E^{NN} σ^{70} or interactions with the DNA which are associated with DNA strand-separation. In addition, P7 inhibition does not seem to be dependent on the order of addition of the σ -factor (compare Figure 6.5.1 lanes 2 to 3 and 6 to 7). This observation corroborates the view that P7 does not function simply to prevent holoenzyme formation.

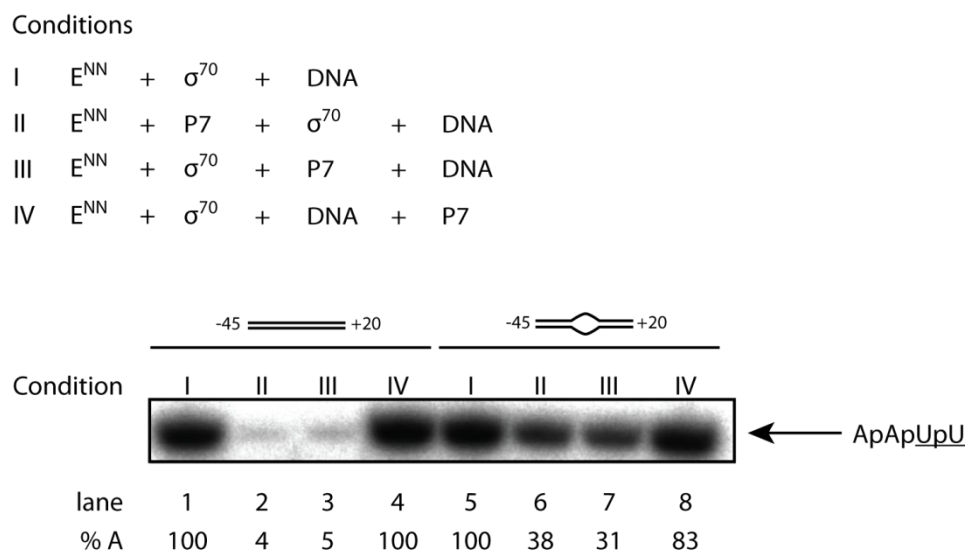


Figure 6.5.1 The order of addition effect of P7 on transcription initiation.

An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing the synthesis of the ApApUpU transcript as indicated by the arrow whereby the underlined nucleotides are ^{32}P -labelled. Transcription was performed by $E^{NN}\sigma^{70}$ from a homoduplex *lacUV5* template (lanes 1-4) and heteroduplex template (lanes 5-8) in the absence (lanes 1 and 5) and presence of 1:2 molar excess of P7 (lanes 2-4 and 6-8). The order of addition of the major components of each reaction is illustrated above. The percentage activity (% A) of $E^{NN}\sigma^{70}$ in the presence of P7 was calculated with respect to the reaction conducted in the absence of P7 with the same *lacUV5* template (lanes 1 and 5), and are indicated beneath each lane.

To test whether P7 may antagonise R_{Pc} formation, EMSA experiments were conducted using a short version of the *lacUV5* probe which contains only the upstream portion of the DNA (-45 to -7), which is sufficient for R_{Pc} formation but is inadequate for isomerisation to the R_{Po}. Results revealed that P7 has no effect on the binding of the (-45/-7)(-45/-7) DNA to $E^{NN}\sigma^{70}$ compared to reactions conducted with either storage buffer (SB) or $E^{WT}\sigma^{70}$ (Figure 6.5.2). Therefore, it seems that P7 does not inhibit transcription initiation by preventing the engagement of the DNA with $E^{NN}\sigma^{70}$. Additional EMSAs were performed using a longer *lacUV5* probe (-45/+20)(-45/+20) and at $\leq 4^\circ\text{C}$ whereby the predominant complex is the R_{Pc}. This experiment also showed that P7 had no effect on R_{Pc} formation (data not shown).

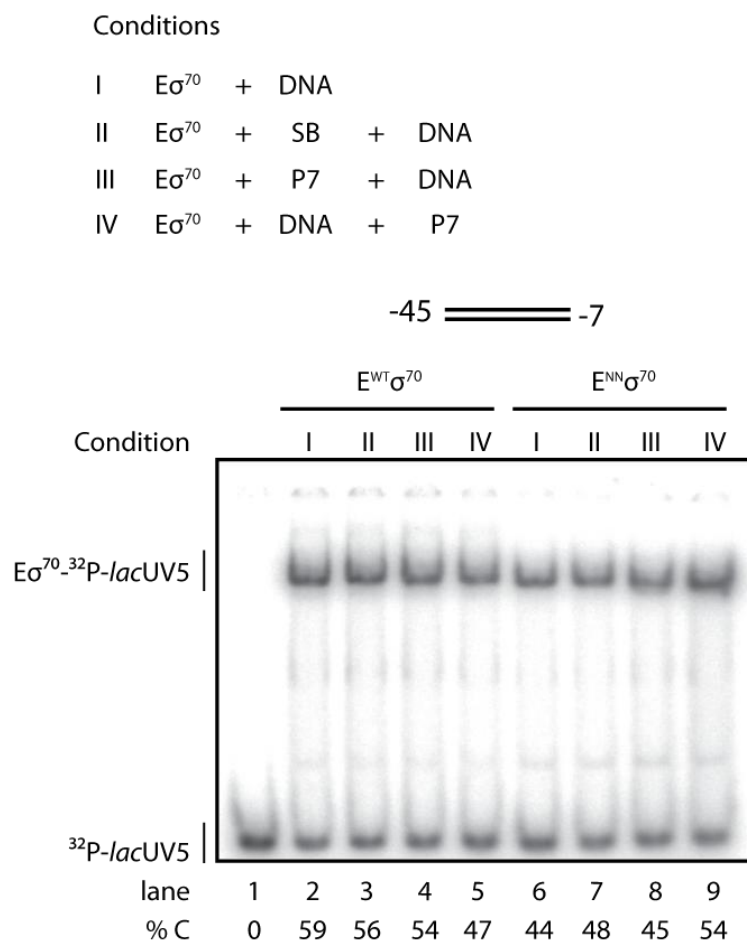


Figure 6.5.2 Effect of P7 on upstream ^{32}P -*lacUV5* binding to $E^{WT}/^{NN}\sigma^{70}$.

An autoradiograph of a 4.5% (w/v) polyacrylamide native gel showing the effect of P7 on the binding of the (-45/-7)(-45/-7) ^{32}P -*lacUV5* probe to $E^{WT}\sigma^{70}$ or $E^{NN}\sigma^{70}$. The order of addition of the major components of each reaction (including SB – storage buffer) is illustrated above. The migration positions of unbound ^{32}P -*lacUV5* and $E\sigma^{70}$ - ^{32}P -*lacUV5* are as indicated. The percentage of $E\sigma^{70}$ - ^{32}P -*lacUV5* complex formation (% C) is calculated relative to a reaction conducted in the absence of P7 and $E\sigma^{70}$ (lane 1) and is shown beneath each lane.

To investigate how P7 may affect events after R_{Pc} leading to R_{Po} formation, similar EMSA experiments were conducted with the (-45/+20)(-45/+20) *lacUV5* templates at 37°C where the most predominant complex is the R_{Po} (Figure 6.5.3). Complexes were challenged with heparin which abolishes the R_{Pc} and intermediate $E\sigma$ -promoter DNA complexes that have not acquired full R_{Po} status. P7 had no detectable effect on complexes formed by $E^{WT}\sigma^{70}$. Reactions with $E^{NN}\sigma^{70}$ formed comparable R_{Po} complexes to $E^{WT}\sigma^{70}$ (compare lanes 2 and 6), however interestingly in the presence of P7 a faster migrating complex is formed (lane 8). This complex is almost completely disrupted in the presence of heparin (lane 9) which suggests it has not yet formed a stable R_{Po}. This result suggests that P7 prevents full isomerisation of the R_{Pc} to R_{Po} by most likely stabilising or ‘trapping’ a normally transient intermediate $E\sigma^{70}$ -promoter complex. The formation of this P7 ‘trapped’ complex

was also observed when EMSAs were conducted using the T7A1, λ PR3, T5N25 and T5N25CONS promoter probes (data not shown).

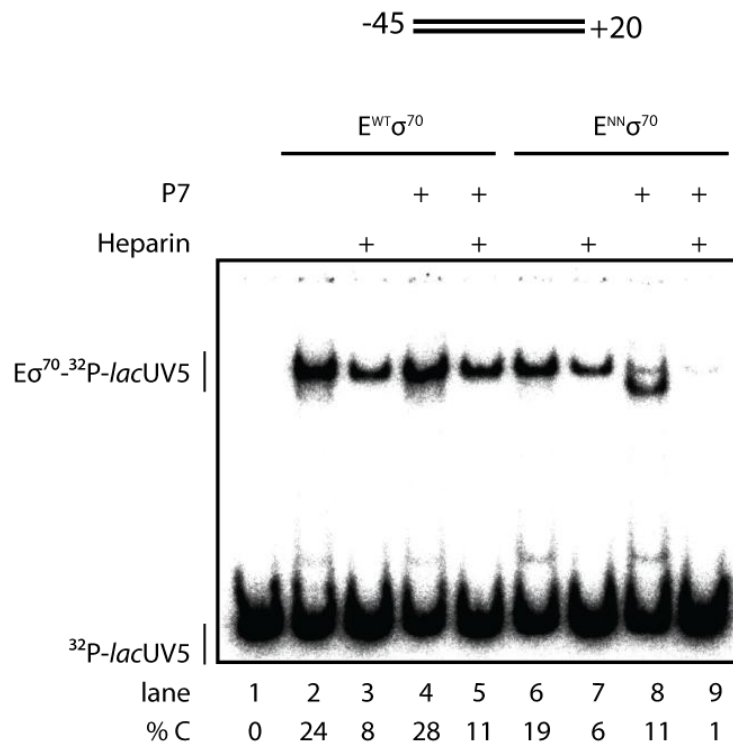


Figure 6.5.3 Effect of P7 on $^{32}\text{P-lacUV5}$ binding to $E^{WT}/E^{NN}\sigma^{70}$.

An autoradiograph of a 4.5% (w/v) polyacrylamide native gel showing the effect of P7 on the binding of the (-45/+20)(-45/+20) $^{32}\text{P-lacUV5}$ probe to $E^{WT}\sigma^{70}$ or $E^{NN}\sigma^{70}$. The order of addition of the major components of each reaction is illustrated above. The migration positions of unbound $^{32}\text{P-lacUV5}$ and $E\sigma^{70-32}\text{P-lacUV5}$ are as indicated. The percentage of $E\sigma^{70-32}\text{P-lacUV5}$ complex formation (% C) is calculated relative to a reaction conducted in the absence of P7 and $E\sigma^{70}$ (lane 1) and is shown beneath each lane.

Next, order of addition experiments were performed to investigate how P7 may affect the equilibrium between heparin sensitive (RPc) complexes and heparin stable (RPo) complexes. When P7 was added prior to the DNA, the same complexes were formed as in Figure 6.5.3 lane 8 with 70% of the complex forming a faster migrating and heparin sensitive P7-trapped complex and 30% forming a slower migrating, heparin stable RPo complex (Figure 6.5.4 lane 6). If P7 is added after the pre-incubation of $E^{NN}\sigma^{70}$ with the DNA, i.e. after RPo formation then the equilibrium between P7-trapped complex and RPo is shifted (lane 8) to a 55:45% equilibrium in favour of the RPo complex.

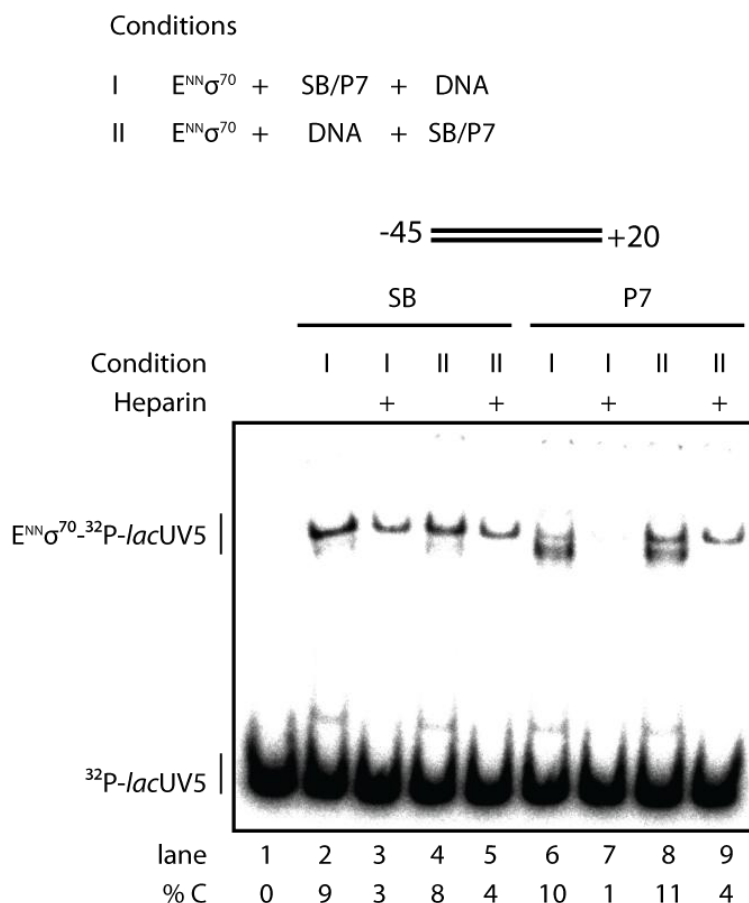


Figure 6.5.4 The order of addition effect of P7 on $^{32}\text{P-lacUV5}$ binding to $E^{NN}\sigma^{70}$.

An autoradiograph of a 4.5% (w/v) polyacrylamide native gel showing the effect of P7 on the binding of the (-45/+20)(-45/+20) *lacUV5* probe to $E^{NN}\sigma^{70}$. The order of addition of the major components of each reaction is illustrated above and the addition of heparin is indicated above each lane. The migration positions of unbound $^{32}\text{P-lacUV5}$ and $E\sigma^{70-32}\text{P-lacUV5}$ are as indicated. The percentage of $E\sigma^{70-32}\text{P-lacUV5}$ complex formation (% C) is calculated relative to a reaction conducted in the absence of P7 and $E\sigma^{70}$ (lane 1) and is shown beneath each lane.

The extent of the engagement of promoter DNA with $E^{NN}\sigma^{70}$ in the P7 ‘trapped’ complex is unknown. To explore how P7 affects $E^{NN}\sigma^{70}$ -promoter DNA interactions a fluorescent binding assay, referred to as the RNAP beacon assay was employed to quantitatively assess the binding between the $E^{NN}\sigma^{70}$ and various segments of the promoter in the presence of P7 (refer to section 2.7.4). In the RNAP beacon assay (Figure 6.5.5) the σ^{70} amino acid residue C211 is labelled with the fluorophore tetramethylrhodamine-5-maleimide (TMR). C211 is located in σ R2.3, in the vicinity of a highly conserved cluster of aromatic tryptophan and tyrosine amino acid residues which effectively quench the fluorescence of TMR by photoinduced electron transfer. Interaction of σ R2.3 with the non-template strand of the -10 element disrupts the contacts between C211-TMR and the aromatic

residues of σ R2.3 which reduces the quenching and consequently augments the fluorescent signal. This allows the binding of $E\sigma^{70}$ -TMR to DNA templates containing the -10 promoter element to be measured fluorescently and as a function of time.

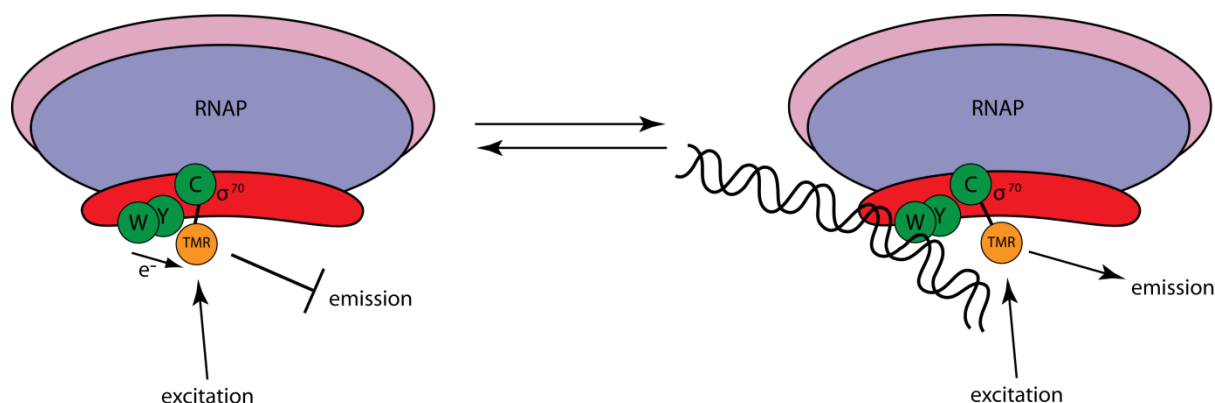


Figure 6.5.5 RNAP beacon assay.

A schematic representation of the RNAP beacon assay. The orange circle represents the TMR fluorophore conjugated to σ^{70} amino acid residue C211. The aromatic amino acid residues tyrosine (Y) and tryptophan (W) effectively quench the fluorescence by photoinduced electron transfer. Binding of DNA to that region disrupts the contacts between the aromatic residues and TMR, resulting in increased fluorescence emission. Inspired by (Mekler *et al.*, 2011c).

RNAP beacon assays were conducted together with Dr. Vladimir Mekler at the Waksman Institute of Microbiology, State University of New Jersey to investigate how P7 may affect $E^{NN}\sigma^{70}$ affinity for specific segments of promoter DNA. Experiments were either conducted using the T5N25 promoter DNA or a derivative containing a consensus -35 element, an optimised UP element and the extended -10 element referred to as T5N25CONS. The binding kinetics of these promoter DNA fragments to $E\sigma^{70}$ -TMR has been well characterised and therefore provides a basis for this inhibitory study (Mekler *et al.*, 2011c; Mekler *et al.*, 2011a; Mekler *et al.*, 2011b).

To begin, experiments were conducted using the T5N25 (-65/+35)(-65/+35) DNA fragment as this possesses a middle range affinity which is convenient for a preliminary investigation. The addition of this DNA fragment to $E^{NN}\sigma^{70}$ -TMR rapidly increased the fluorescent signal from ~27,000 to ~84,000 counts/s, which is indicative of a high binding affinity (Figure 6.5.6, blue line). However, when $E^{NN}\sigma^{70}$ -TMR is pre-incubated with saturating amounts of P7 the addition of the DNA fragment yields a much lower fluorescent signal of ~43,000 counts/s which demonstrates that P7 has reduced the binding affinity of this DNA fragment to $E^{NN}\sigma^{70}$ (Figure 6.5.6, red line). In contrast, and as expected if P7 is added after the formation of the $E^{NN}\sigma^{70}$ -promoter complex then it has only a

negligible effect and reduces fluorescence from ~88,000 to ~87,000 counts/s over the timecourse (Figure 6.5.6, green line). This result is consistent with the *in vitro* transcription and EMSA data which suggests that P7 has a much greater effect prior to the addition of the DNA. This data, together with the EMSA results also implies that P7 reduces the affinity of $E^{NN}\sigma^{70}$ for the promoter DNA but importantly does not prevent complex formation between $E^{NN}\sigma^{70}$ and the promoter DNA.

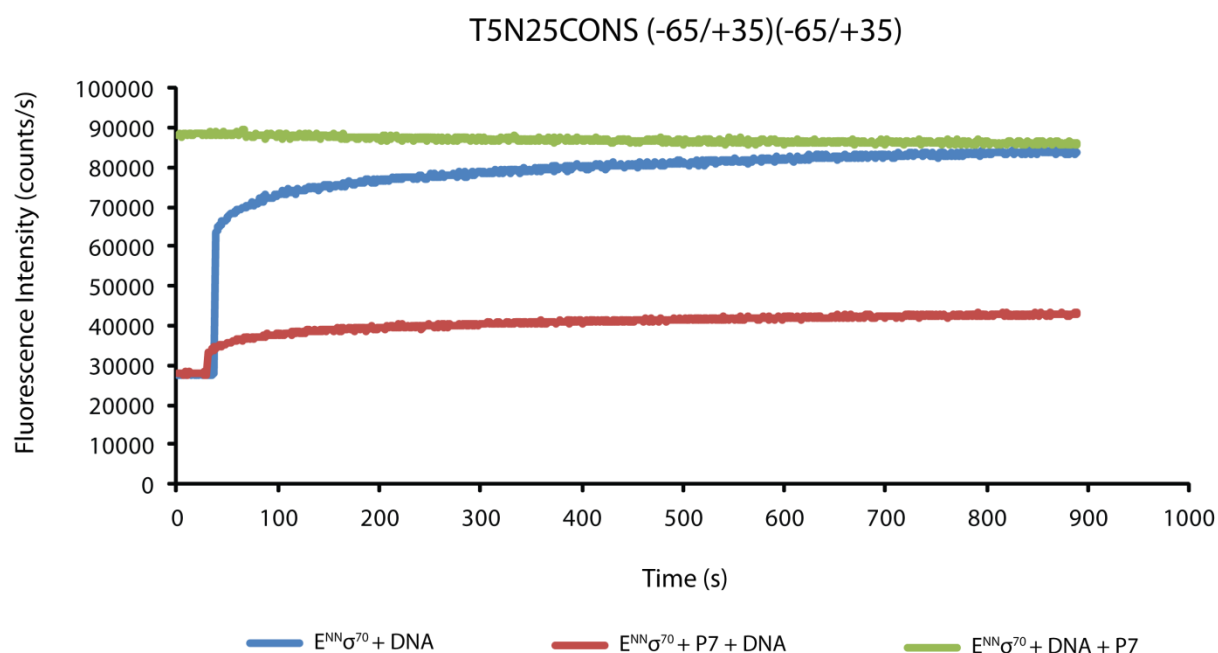


Figure 6.5.6 The effect of P7 on $E^{NN}\sigma^{70}$ interactions with the full length T5N25 promoter.

A time dependent increase in fluorescence upon $E^{NN}\sigma^{70}$ -TMR binding to the T5N25 (-65/+35)(-65/+35) DNA fragment is shown in blue. The pre-incubation of $E^{NN}\sigma^{70}$ -TMR with P7 before addition of the DNA fragment is shown in red, whereas the addition of P7 after $E^{NN}\sigma^{70}$ -TMR-DNA fragment binding is shown in green.

Additional RNAP beacon assays were conducted to identify specific $E^{NN}\sigma^{70}$ -promoter DNA interactions that may be targeted by P7 and therefore responsible for the prevention of RPo formation. To test whether P7 may antagonise interactions with the -10 promoter element a single-stranded T5N25 (-12/-6) fragment was used in the RNAP beacon assay. This oligonucleotide comprises the consensus $^{-12}\text{TATAAT}^{-6}$ sequence of the -10 promoter element on the non-template strand. The -10 promoter element is contacted by σ^{70} R2.4 in the RPo and the non-template strand of this region is specifically recognised by $E\sigma^{70}$ to stabilise the complex during DNA strand-separation (Marr and Roberts, 1997). Because this oligonucleotide has a very low affinity for $E\sigma^{70}$ it may not be possible to detect any change in signal if $E^{NN}\sigma^{70}$ is pre-incubated with P7 which would be quite

ambiguous to interpret. Therefore, considering this oligonucleotide binds weakly and P7 is likely to still be able to compete (as seen in Figure 6.5.4, lane 8), the addition of P7 and DNA was conducted in the reverse order, i.e. $E^{NN}\sigma^{70}$ was incubated with T5N25 (-12/-6) first and then P7 was added. Here it was shown that P7 does not antagonise the interactions between $E^{NN}\sigma^{70}$ -TMR and the T5N25 (12/-6) fragment, in fact a slight increase in fluorescence from ~49,000 to ~52,000 counts/s was actually detected (Figure 6.5.7). Therefore, it seems that P7 is unlikely to antagonise the interactions with the -10 promoter element in the RPo but may instead increase the affinity between $E\sigma^{70}$ (σ R2.4) and this promoter DNA fragment. However this interaction needs to be investigated further using ds-DNA fragments.

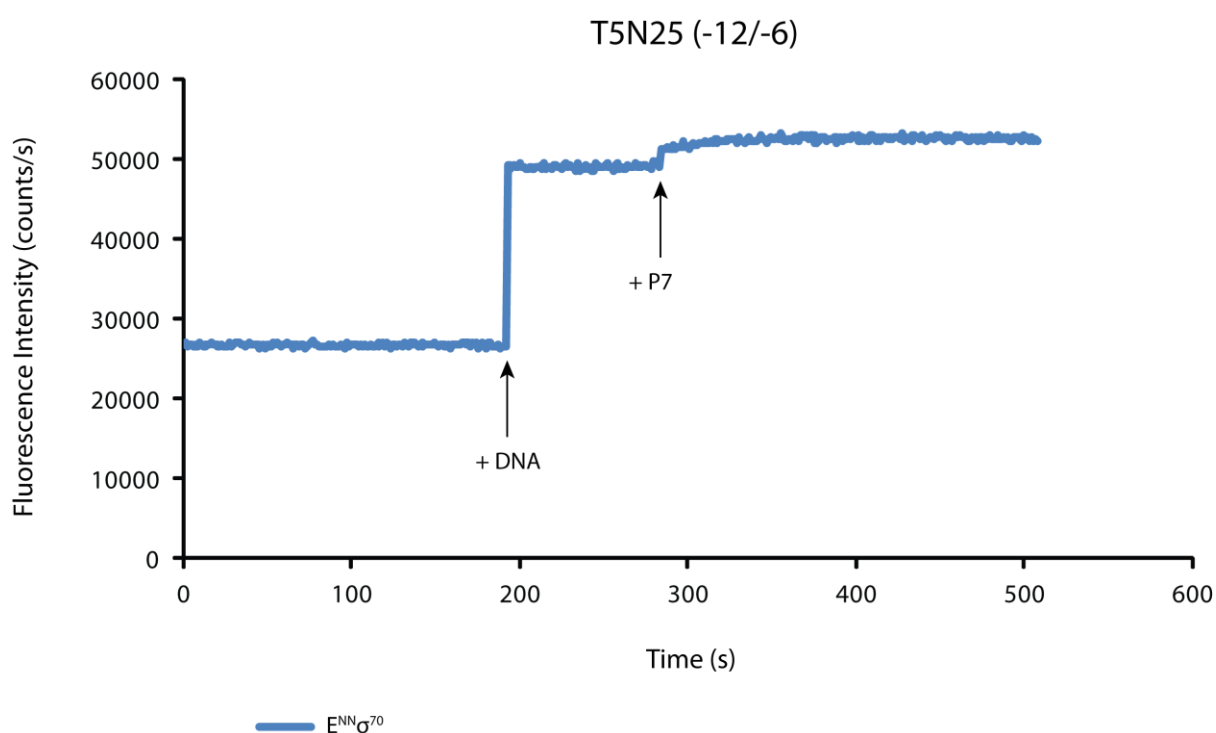


Figure 6.5.7 The effect of P7 on $E^{NN}\sigma^{70}$ interactions with the T5N25 -10 promoter element.

A time dependent measurement of fluorescence of $E^{NN}\sigma^{70}$ -TMR is shown in blue, the addition of the single-stranded T5N25 (-12/-6) DNA fragment and P7 are indicated with arrows.

The DNA promoter fragment was extended downstream to include the discriminator region and the transcription start site; $^{-12}\text{TATATTAGATTCAT}^{+2}$. The discriminator region refers to the section of DNA between the -10 promoter element and the transcription start site, of which the non-template strand is known to be contacted by σ R1.2 (Haugen *et al.*, 2006). An identical binding assay with the single-stranded T5N25 (-12/+2) DNA fragment showed contrasting results to the (-12/-6) oligonucleotide; the addition of P7 to the $E^{NN}\sigma^{70}$ -TMR- T5N25 (-12/+2) complex significantly reduced the fluorescent signal from ~49,000 to ~36,000 counts/s (Figure 6.5.8). This implies that P7 may

target, directly or indirectly interactions between $E^{NN}\sigma^{70}$, specifically σ R1.2 and the discriminator region. This finding needs to be further investigated using alternative sequences, as the sequence of the discriminator region has been shown to be crucial for the regulation of RNAP at certain promoters (Haugen *et al.*, 2006).

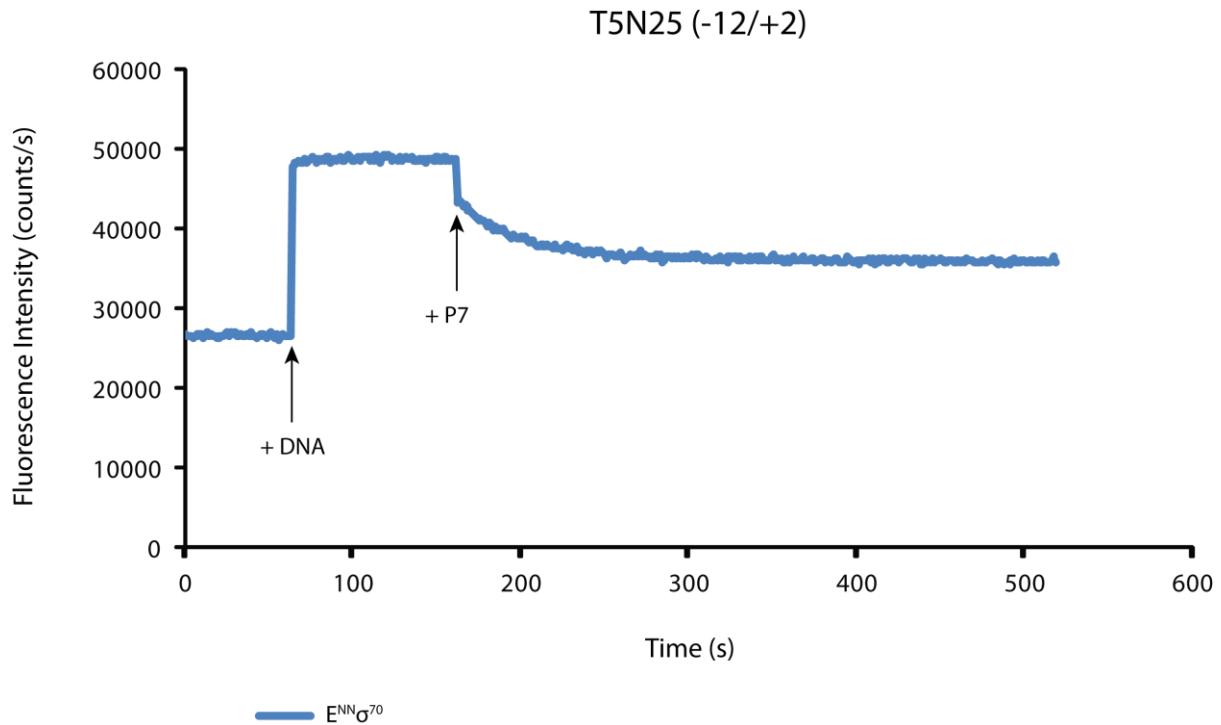


Figure 6.5.8 The effect of P7 on $E^{NN}\sigma^{70}$ interactions with the T5N25 discriminator region.

A time dependent measurement of fluorescence of $E^{NN}\sigma^{70}$ -TMR is shown in blue, the addition of the single-stranded T5N25 (-12/+2) DNA fragment and P7 are indicated with arrows.

Since transcription initiation from the extended -10 type promoters appears to be resistant to P7 (Nechaev *et al.*, 2002), the effect of P7 on $E^{NN}\sigma^{70}$ interactions with a promoter fragment containing this element was tested (Figure 6.5.9). The T5N25CONS (-26/-6)(-26/-12) upstream fork junction does not contain a -35 promoter element but does still have an extended -10 promoter element and the -10 promoter element which due to the nature of the assay is necessary for fluorescence quenching, however, recall that binding of the -10 element does not seem to be effected by P7 (Figure 6.5.7). Results showed that the pre-incubation of $E^{NN}\sigma^{70}$ -TMR with P7 has no significant effect on T5N25CONS (-26/-6)(-26/-12) binding which suggests that P7 does not target this promoter element. This result is consistent with the observation that P7 displays significantly less inhibitory activity on transcription initiation from extended -10 promoters (Nechaev *et al.*, 2002).

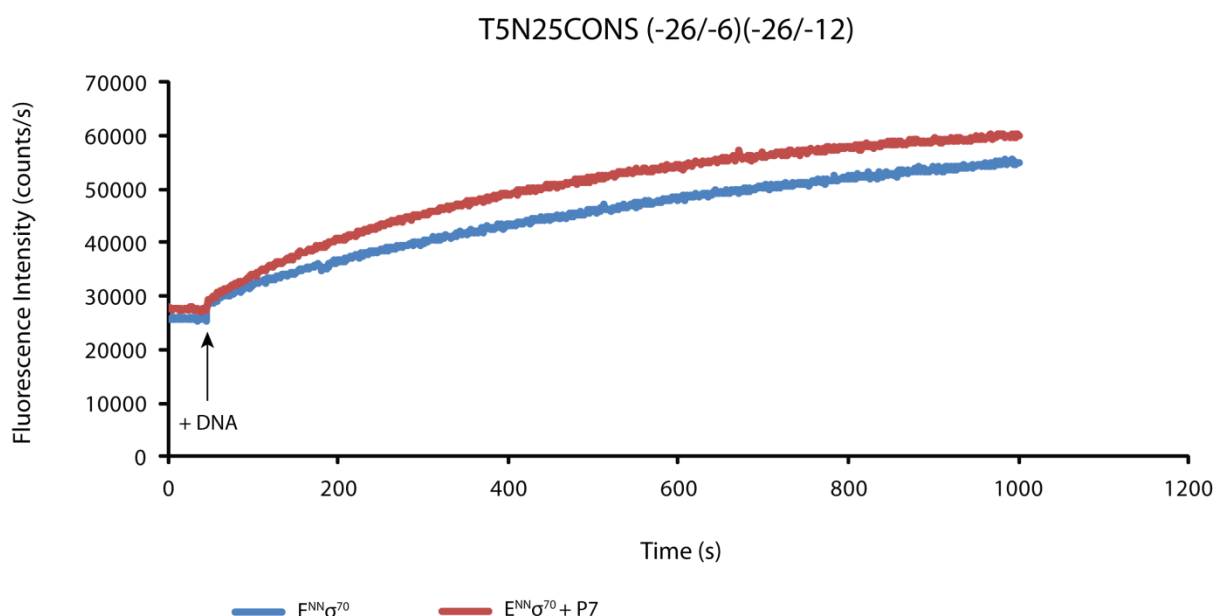


Figure 6.5.9 The effect of P7 on $E^{NN}\sigma^{70}$ interactions with the T5N25 extended -10 promoter element.
A time dependent measurement of fluorescence of $E^{NN}\sigma^{70}$ -TMR binding to T5N25CONS (-26/-6)(-26/-12) is shown in blue whereas $E^{NN}\sigma^{70}$ -TMR-P7 binding to T5N25CONS (-26/-6)(-26/-12) is shown in red.

Results thus far suggest that P7 may antagonise interactions between the $E^{NN}\sigma^{70}$ and the discriminator (specifically the non-template -6 to +2 segment) as only promoter DNA fragments containing this region display a reduced binding affinity. However upstream segments of the promoter DNA (-65 to -26) have not been directly tested. This portion of the promoter DNA encompasses the -35 element which is a potential target for P7 inhibition for two reasons. Firstly, P7 seems to affect the interdomain distance between σ R2.4 and 4.2 which contact the -10 and -35 promoters. Secondly, $E^{NN}\sigma^{70}$ transcription initiation from extended -10 promoters is resistant to P7, recall that $E^{NN}\sigma^{70}$ contacts with an extended -10 promoter make the -35 promoter element dispensable; therefore it seems plausible that P7 may target interactions with the -35 promoter element. To experimentally test whether interactions with the -35 promoter element are a target of P7 inhibition, further fluorescent binding assays were performed with the T5N25CONS (-38/-11)(-38/-12) upstream fork junction. This upstream fork junction encompasses the -35 promoter element but not the discriminator region which is already proposed to be involved in P7 inhibition. Addition of P7 to the $E^{NN}\sigma^{70}$ -TMR- T5N25CONS (-38/-11)(-38/-12) complex decreases the fluorescent signal from ~46,000 to ~38,000 counts/s (Figure 6.5.10). This implies that P7 inhibition may also be dependent on the -35 element; recall that P7 has no effect on $E^{NN}\sigma^{70}$ -TMR interactions with the -26 to -6 segment (Figure 6.5.9), therefore P7 specifically antagonises binding of the -38 to -26 section of promoter DNA.

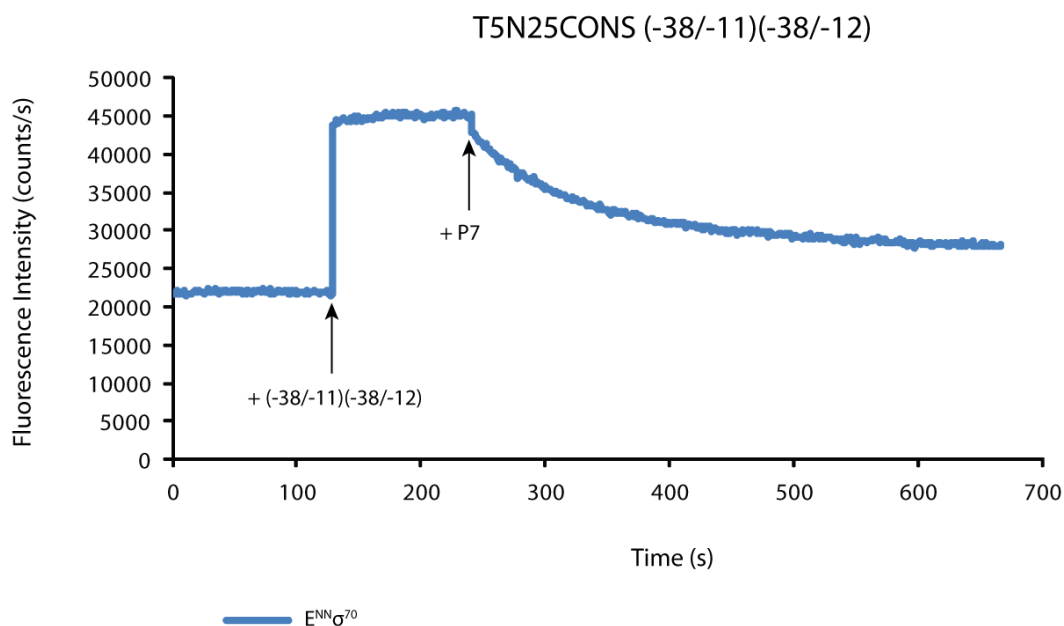


Figure 6.5.10 The effect of P7 on $E^{NN}\sigma^{70}$ interactions with the T5N25 promoter -35 promoter.

A time dependent measurement of fluorescence of $E^{NN}\sigma^{70}$ -TMR binding to T5N25 (-38/-11)(-38/-12). The addition of the (-38/-11)(-38/-12) fragment and P7 are indicated with arrows.

These RNAP beacon experiments have so far demonstrated that P7 may form a ‘trapped’ complex (as seen in Figure 6.5.3 and Figure 6.5.4) by directly or indirectly reducing the affinity of $E^{NN}\sigma^{70}$ to the -35 promoter element and/or the discriminator region (Figure 6.5.11), but not interactions with the -10 promoter element in the RPo. Additionally it seems that contacts with the extended -10 promoter allow $E^{NN}\sigma^{70}$ to escape inhibition by P7.

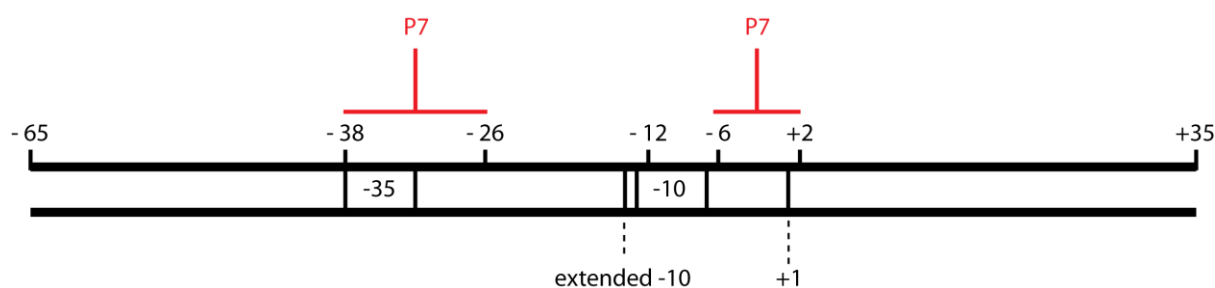


Figure 6.5.11 Summary of potential $E^{NN}\sigma^{70}$ -promoter DNA interactions targeted by P7.

A schematic illustrating the location of the $E^{NN}\sigma^{70}$ -promoter DNA interactions which are proposed to be antagonised by P7. The -35, -10 and extended -10 promoter elements are labelled along with the transcription start site. The positions of the DNA fragment end points used in the RNAP beacon assays are indicated on the top strand.

6.6. Discussion

A bacterial two-hybrid approach was used to map the residues in P7 which are important for binding to the *X. oryzae* RNAP. The mutagenesis screen revealed that P7 is quite sensitive to mutation, which implies that the interaction interface could potentially be quite extensive. The P7 residues were categorised into four classes according to the ability of their respective alanine mutant counterparts to bind to the *X. oryzae* RNAP β' -subunit fragment and induce expression of β -galactosidase. Using this categorisation 27 class IV residues were identified as being important for P7 binding including two clusters: G¹⁹S²⁰V²¹L²²⁻²³V²⁴K²⁵V²⁶ and F³⁹D⁴⁰R⁴¹. Although Western blot analysis of the cell lysates revealed that the class IV mutants were all expressed, there was notable variation in the apparent levels of expression. Secondary structure prediction software proceeded by NMR spectroscopy was used to resolve a high resolution, 3-D representation of P7. When the class IV residues were mapped in the context of the tertiary structure it became apparent that some of these residues may not be directly responsible for binding *per se*. In particular, many of the class IV residues seemed to be positioned to form the hydrophobic core of the protein, including residues within the central β -strand formed by the majority of the G¹⁹S²⁰V²¹L²²⁻²³V²⁴K²⁵V²⁶ cluster and residues in the α -helix. Moreover, the F³⁹D⁴⁰R⁴¹ cluster comprises an apparent turn in the protein between the β_3 -strand and the α -helix. Therefore ~12 of the class IV residues are potentially required for maintaining structural integrity and/or the correct folding of the protein. Unstable or aberrant proteins are more likely to be degraded, which may explain the discrepancies in the apparent expression levels of some of the class IV mutants. However, this does not eliminate the possibility that these residues may play an additional role in forming an interaction with the β' subunit as well as maintaining structural features. The screen has highlighted ~15 further class IV residues (F⁴, I⁷, G⁹, Y¹⁰, N¹², F¹⁴, Q¹⁷, G¹⁹, S²⁰, V²¹, Y⁴⁴, S⁴⁶, L⁵⁸, G⁶² and I⁶⁴), although as yet, none of these have been directly confirmed as important for binding to the RNAP.

The binding of P7^{WT} to E^{NN} and variants could be further studied by fluorescence anisotropy (refer to chapter 3). This would allow confirmation of the P7 residues important for binding to E^{NN}o⁷⁰. For site specific labelling, the fluorophore is typically conjugated via a sulphhydryl group but as P7 does not contain a native cysteine, this amino acid residue would need to be introduced. The mutagenesis screen and structure could be used to select suitable class I and surface exposed amino acid residues as candidates such as S⁸, S¹⁶, P⁴⁸, K⁵⁹, R⁶⁰. Alternatively as P7 is encoded in pET28b⁺ and pET33b⁺, a cysteine residue could be introduced within the N-terminal tag of the recombinant protein.

In addition, a structure of P7 bound to a *X. oryzae* β' N-terminal fragment could be used to form a model of a $E\sigma^{70}$ -P7 complex which would provide a more informative perspective on the orientation of the binding partners and therefore the prospective binding determinants. A structural complex would also help decipher which mobile modules of RNAP (β' Zn-binding element, β' zipper or β' coiled-coil) may be influenced by P7 binding. Ultimately the mutagenesis screen could be implemented together with structural and biochemical data to decipher which residues are important for the function of P7 as an inhibitor of transcription initiation and as an anti-terminator.

Biochemical studies set out to determine the steps in transcription initiation which are targeted by P7. *In vitro* transcription assays and EMSAs confirmed that P7 has no detectable effect on the initial engagement of the DNA with $E^{NN}\sigma^{70}$ and cannot inhibit transcription from a pre-formed $E^{NN}\sigma^{70}$ -promoter DNA complex. Thus P7 may not prevent RPo formation but does seem to inhibit a stage prior to RPo formation. Furthermore, EMSAs also revealed that P7 appears to trap an $E^{NN}\sigma^{70}$ -promoter DNA complex which is faster migrating and heparin sensitive. As this P7 'trapped' complex migrates differentially to a fully competent RPo complex it possibly represents a markedly distinct conformation. Heparin stability is commonly used as a hallmark of the RPo complex or late intermediates; therefore P7 seems to trap an early intermediate whereby the 'transcription bubble' has not yet formed and hence the DNA is unable to compete with heparin. $KMnO_4$ experiments could be conducted to assess the extent, if any, of the DNA strand-separation in the P7 'trapped' complex.

Results from order of addition binding experiments show that P7 disrupts the equilibrium between a heparin unstable and heparin stable complex. Therefore P7 may antagonise $E^{NN}\sigma^{70}$ -promoter DNA interactions or conformational changes in $E^{NN}\sigma^{70}$ required for DNA isomerisation. To investigate if P7 may affect the interactions of $E^{NN}\sigma^{70}$ with promoter DNA, binding experiments were conducted using the RNAP beacon assay. By conducting preliminary fluorescence based binding experiments with specific segments of a P7 sensitive promoter DNA fragment it was shown that P7 may affect the interactions between $E^{NN}\sigma^{70}$ (σ R1.2) and the discriminator region as well as (σ R4.2) and the -35 promoter element.

During transcription initiation the discriminator region is contacted by σ R1.2. The sequence of the discriminator region can influence the strength of this interaction and consequently the lifetime of the $E\sigma^{70}$ -promoter DNA complex. Furthermore this interaction is known to be subject to regulation by transcription factors (Haugen *et al.*, 2006; Haugen *et al.*, 2008). Thus, it is conceivable that P7 may weaken the interaction between the discriminator and σ R1.2 which may shorten the lifetime of the RPo, shifting the equilibrium to the left and thus 'trapping' an intermediate complex.

Since transcription from extended -10 promoters does not rely on the -35 promoter element the results from the RNAP beacon assay, which show a P7 dependency on the -35 promoter element region, may explain why transcription initiation from extended -10 promoters is resistant to P7 inhibition. Recently it has been shown that at certain promoters the β' zipper can facilitate RPo formation by interacting with the spacer region to compensate for σ^{70} interactions with the -35 promoter element, in a similar manner to the extended -10 promoter (Yuzenkova *et al.*, 2011). Therefore it would be interesting to investigate whether transcription initiation from this novel type of promoter is also resistant to P7 inhibition.

It should be emphasised that all the RNAP beacon assay results shown here are preliminary experiments and need repeating with suitable controls (i.e. in the absence of P7 and/or presence of a P7 mutant defective in binding) and with different orders of addition. Furthermore, the effect of P7 on $E^{NN}\sigma^{70}$ interactions with downstream DNA has not yet been assessed; these interactions are crucial for forming a stable RPo and therefore may also be targeted by P7. The P7 'trapped' complex may also be further characterised using DNase I and ExoIII footprinting to determine how the $E\sigma^{70}$ is occupied by the promoter DNA.

Currently, the characterisation of the mechanism of inhibition of transcription by P7 has been conducted using a mutant version of the *E. coli* RNAP as this represents the best studied bacterial RNAP for which composite structures and a plethora of biochemical resources are available. However *X. oryzae* itself is a major rice pathogen and therefore the study of its transcriptional components has relevance to the agricultural industry. Since there appears to be significant variation between the *E. coli* RNAP β' subunit and the *X. oryzae* counterpart (with only 72% sequence homology at the amino acid level) and it has been shown that the choice of DNA templates can influence the potency of an inhibitor (Hauptenthal *et al.*, 2012), it may be more accurate to also investigate the effect of P7 using a more native and relevant system. This approach may also have implications for the continued study of P7 as an anti-terminator of transcription.

CHAPTER 7

OVERVIEW

7.1. Work leading to this study

This study aimed to explore the molecular mechanisms of action of two phage-encoded proteins; T7 Gp2 and Xp10 P7, which bind directly to the bacterial RNAP and inhibit its activity during transcription initiation. Work leading to this study had shown that Gp2 and P7 target steps which prevent the formation of a transcriptionally competent RPo; the precise details of which remained to be elucidated (Figure 7.1.1 and Figure 7.1.2).

Previously, it was established that Gp2 binding to RNAP was dependent on the conservation of the E1158 and E1188 amino acid residues of the β' jaw (Nechaev and Severinov, 1999). The β' jaw forms part of the dw-DBC and amongst other functions is required for stable DNA strand-separation (Ederth *et al.*, 2002b). During the course of this study it was demonstrated by NMR titration experiments that the dw-DNA directly contacts the β' jaw and that this interaction can be disrupted by Gp2, which competes for binding to an overlapping site (James *et al.*, 2012). Early research from this laboratory included a structure-function analysis of Gp2, which revealed that amino acid residues R56 and R58, which project out from the β_3 strand were also required for binding to the β' jaw (Camara *et al.*, 2010). Despite the identification of the major determinants for Gp2-RNAP binding, the affinity of the interaction had never been measured (Figure 7.1.1.①).

An initial investigation into the mechanism of inhibition of RNAP by Gp2 had demonstrated that Gp2 antagonises a late stage of RPo formation associated with DNA strand-separation around the transcription start site (Camara *et al.*, 2010). This idea was later supported by DNA footprinting methods which showed by generating a Gp2-Eo⁷⁰-promoter DNA quaternary complex that only the upstream portion of the 'transcription bubble' is able to form whereas the DNA around the transcription start site remains in its duplex conformation (Mekler *et al.*, 2011b). However, the overall molecular mechanism of action employed by Gp2 to antagonise the interactions between RNAP and the promoter DNA was yet to be unravelled. Analysis of the structure of Gp2 had revealed a continuous negatively charged strip (NCS) of aspartic and glutamic acid residues running from the β_1 - β_2 loop down the α -helix (Camara *et al.*, 2010). The potential role of the NCS in the Gp2-mediated inhibition of RNAP had not been investigated (Figure 7.1.1.②). Furthermore, like Gp2, σ R1.1 is also negatively charged and known to reside in the dw-DBC, proximal to the Gp2 binding site. Considering that the displacement of R1.1 from the dw-DBC is a pre-requisite for RPo formation it was proposed that R1.1 may also play a role in the mechanism by which RNAP is inhibited by Gp2 (Figure 7.1.1.③).

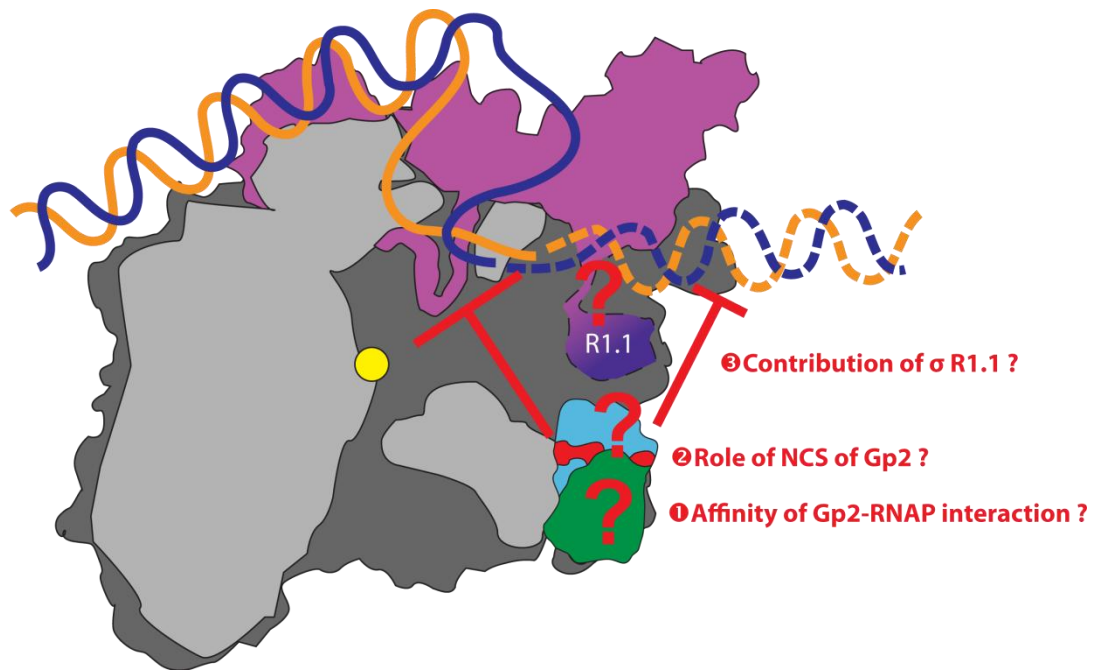


Figure 7.1.1 How does Gp2 inhibit RPo formation?

A cartoon depicting the effects of Gp2 on RPo formation by $E\sigma^{70}$ (grey and purple respectively). The active site is represented by a yellow dot whereas the DNA template strand is shown in orange and the non-template strand in dark blue. Gp2 (light blue) binds to the β' jaw (green), of which the affinity of the interaction has not been previously measured (①). Gp2 prevents the binding of the dw-DNA to the β' jaw and antagonises DNA strand-separation around the transcription start site (illustrated by red bars) by an undefined mechanism. The role of the NCS (red patches) of Gp2 has not been investigated (②) and the potential contribution of σ R1.1 has not been determined (③).

In contrast to Gp2, P7 represents a relatively uncharacterised phage-encoded protein. Prior to this study, it was established that P7 binds to a markedly distinct region of the β' subunit, located more at the upstream face of RNAP. Interaction studies mapped the specific binding site to the $N^6L^7F^8N^9$ amino acid residues (Yuzenkova *et al.*, 2008). However, the amino acid residues in P7 which are important for binding remain uncharacterised (Figure 7.1.2.①). A preliminary investigation into the effect of P7 on RNAP activity had shown that it may prevent RPc formation by restricting the movement of σ R2.4 and 4.2 which are required for promoter recognition (Nechaev *et al.*, 2002). However, the precise mechanism by which P7 antagonises $E\sigma^{70}$ -promoter DNA interactions has not been investigated (Figure 7.1.2.②).

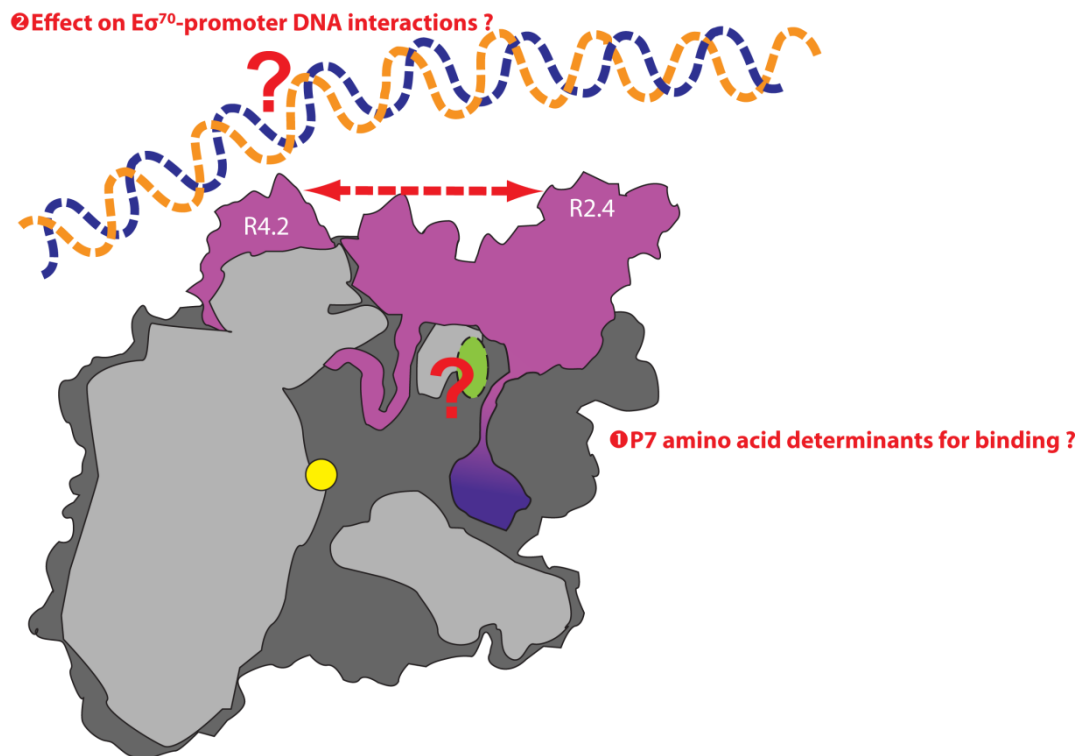


Figure 7.1.2 How does P7 inhibit RPo formation?

A cartoon depicting the effects of P7 on RPo formation by Eσ⁷⁰ (grey and purple respectively). The active site is represented by a yellow dot whereas the DNA template strand is shown in orange and the non-template strand in dark blue. P7 (light green) binds at the upstream face of RNAP; P7 amino acid residues required for binding are not known (①). P7 restricts the interdomain distance between σ R2.4 and 4.2 (illustrated by red dotted arrows). The effect of P7 on Eσ⁷⁰-promoter DNA interactions has not been investigated (②).

7.2. Advances made

From this study we have now proved Gp2 to be a tight binding and potent inhibitor of bacterial RNAP. Results from FA experiments have established that Gp2 binds to the RNAP core with a 1:1 stoichiometry and a low nanomolar affinity and that this initial binding of Gp2 to RNAP occurs independently of the σ-factor. It seems that the number of negatively charged amino acid residues within the NCS determines the potency of Gp2; making this the major feature for the inhibitory function of Gp2 (Figure 7.2.1.①). Using a structural complex of Gp2 and a fragment of the β' jaw we showed that the arrangement of the interaction, extends the residues of the NCS of Gp2 by aspartic and glutamic acid residues of the β' jaw (James *et al.*, 2012), thus supporting the view that the NCS is an important feature in Gp2 function. Since, R56 and R58 as well as the NCS amino acid residues are

highly conserved in Gp2 orthologs the phage encoding these proteins are also likely to employ the same or at least very similar mechanisms of bacterial RNAP inhibition.

The complete removal of the NCS also compromises the ability for Gp2 to remain bound to RNAP upon addition of the promoter DNA. This result supports the view that Gp2 and the dw-DNA compete for binding to the β' jaw (James *et al.*, 2012) (Figure 7.2.1.②). By using the current structural model of Gp2 bound RNAP we have shown that Gp2 is orientated so that the NCS is positioned to face into the positively charged walls of the β' jaw and the $\beta'i6$ of the dw-DBC. Therefore we propose that the role of the NCS of Gp2 may not solely be to electrostatically repel the DNA away from the dw-DNA but may instead be to 'lock' the RNAP in a conformation which is unfavourable for DNA entry. This is supported by recent single-molecule FRET experiments which demonstrated that Gp2 restricts the movement of the β' clamp in favour of a closed and/or collapsed state, narrowing the channel width to ~8- 12 Å (Chakraborty *et al.*, 2012) (Figure 7.2.1.③). This locking mechanism does not occur if the NCS is removed (unpublished data).

We have also described how the highly flexible and negatively charged R1.1 domain of σ^{70} is required for the full Gp2-mediated inhibition of $E\sigma^{70}$ -dependent transcription. During the course of this study FRET experiments were performed to demonstrate that the presence of Gp2 prevents the movement of σ R1.1 from the dw-DBC (Mekler *et al.*, 2011b) (Figure 7.2.1.④). Together Gp2 and R1.1 exert long range antagonistic effects to attenuate the entry of the DNA around the active centre and the correct placement of the single-stranded DNA template strand at the active site (Figure 7.2.1.⑤). Ultimately these direct and indirect effects of Gp2 result in the formation of a partially open and transcriptionally inactive 'transcription bubble' (Mekler *et al.*, 2011b).

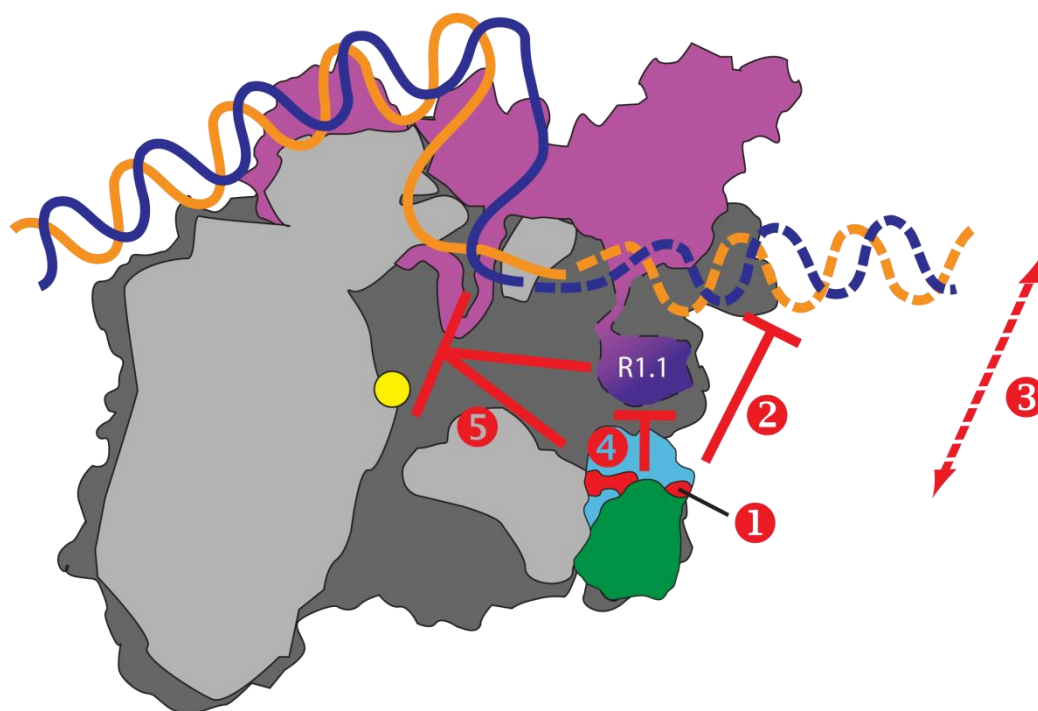


Figure 7.2.1 Overall mechanism of inhibition of $E\sigma^{70}$ by Gp2.

A cartoon depicting the effect of Gp2 on RPO formation by $E\sigma^{70}$ (grey and purple respectively). The active site is represented by a yellow dot, whereas the DNA template strand is shown in orange and the non-template strand in dark blue. Gp2 (light blue) contains a NCS (red patches) which forms the major feature for inhibitory activity (①). Gp2 competes with the dw-DNA for binding to the β' jaw (green) (②). The binding of Gp2 narrows the width of the main channel (illustrated by the red dashed arrows) (③). Gp2 restricts the movement of σ R1.1 from the dw-DBC (④) and together they antagonise the binding of promoter DNA around the active centre (⑤). These effects of Gp2 result in the formation of a partially open and transcriptionally inactive 'transcription bubble'.

We used a structure-function analysis of P7 to inform upon which amino acid residues in P7 may be required for binding to RNAP. The results of a two-hybrid screen were assessed in the context of the newly resolved NMR structure of P7 (Figure 7.2.2.①). We identified approximately 15 amino acid residues that may be important determinants for binding and ~13 additional amino acid residues which are likely to be required for maintenance of the structural integrity of the protein.

We employed both biochemical and biophysical methods to investigate the mechanism of inhibition of transcription initiation by P7. Previously it was shown that P7 reduces the distance between σ R2.4 and 4.2 (Figure 7.2.2.②), suggesting that P7 may prevent R_{Pc} formation. Our results from *in vitro* transcription assays suggested that P7 may not prevent R_{Pc} formation but does antagonise a stage *en route* to the formation of the RPO. In addition, EMSA experiments revealed that P7 traps a markedly distinct $E^{NN}\sigma^{70}$ -promoter DNA complex which is both fast migrating and heparin sensitive. We also conducted preliminary RNAP beacon assay experiments which suggested

that P7 may specifically disrupt $E^{NN}\sigma^{70}$ interactions with the -35 promoter element and the discriminator region (Figure 7.2.2.③). These findings will be subject to further investigations.

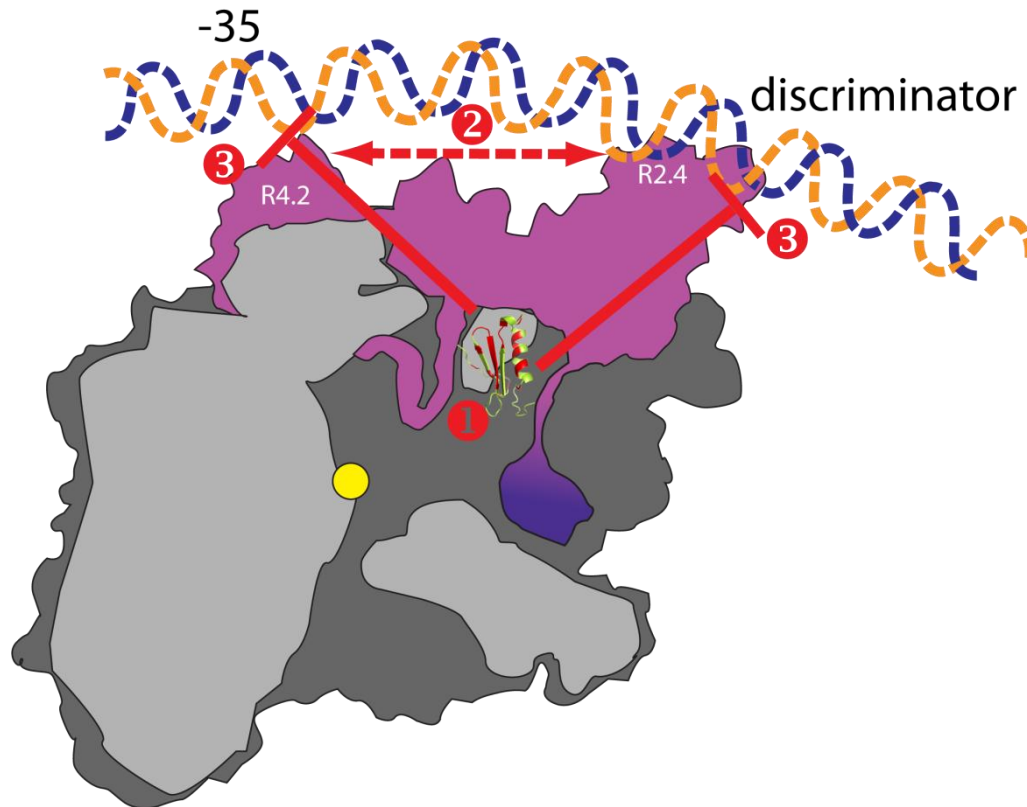


Figure 7.2.2 Insights into the mechanism of inhibition of $E^{NN}\sigma^{70}$ by P7.

A cartoon depicting the effect of P7 on RPo formation by $E\sigma^{70}$ (grey and purple respectively). The active site is represented by a yellow dot, whereas the DNA template strand is shown in orange and the non-template strand in dark blue. The structure of P7 (light green) has been resolved and ~27 amino acid residues (red) have been identified to be either directly or indirectly required for binding (①). P7 restricts the interdomain distance between σ R2.4 and 4.2 (illustrated by red dotted arrows) (②) and may specifically antagonise $E^{NN}\sigma^{70}$ interactions with the -35 promoter element and the discriminator region (③).

7.3. Future perspectives

In this study we have answered many of the objectives originally proposed, although in doing so many more questions have arisen; in particular, two major lines of enquiry may be explored in the immediate future to develop the existing findings.

The work described here and in previous studies has demonstrated Gp2 to be both σ -factor and promoter specific. The current mechanism of inhibition of RNAP by Gp2 has suggested that T7

has evolved to target the activity of the major housekeeping holoenzyme, $E\sigma^{70}$ from the -10/-35 type promoters. In contrast, transcription from extended -10 promoters and initiation directed by σ^{54} both appear to be resistant to Gp2 (Nechaev and Severinov, 1999; Wigneshweraraj *et al.*, 2004), whereas transcription by $E\sigma^{38}$ is only partially inhibited by Gp2. However, the effect of Gp2 has only been studied *in vitro* from a limited number of promoters. The effect of Gp2 on transcription may be assessed in a more biologically relevant manner using an *in vivo* transcriptomics approach. The growth effect of the over-expression of recombinant Gp2 in *E. coli* cells may be monitored over a time-course to determine the point at which growth becomes attenuated. At this point RNA can be extracted and used for microarray analysis. Results will demonstrate the global effect of Gp2 on the *E. coli* transcriptome; detailing which specific promoters are differentially expressed in the presence of Gp2. This will establish which holoenzymes and promoter types are sensitive or resistant to Gp2 inhibition.

The preliminary investigation into the effect of P7 on $E^{NN}\sigma^{70}$ -promoter DNA interactions identified a markedly distinct complex. The P7 'trapped' complex is both faster migrating and heparin sensitive compared to its RPO counterpart and possibly represents a normally transient intermediate complex. This P7 'trapped' complex may be further characterised using DNA footprinting methods. DNaseI and or ExoIII footprinting could be used to determine how the occupation of the DNA by $E^{NN}\sigma^{70}$ changes in the presence of P7. Whereas $KMnO_4$ footprinting would establish if and to what extent the DNA strands are separated. A combination of these methods will inform upon the possible effect of P7 on promoter DNA interactions which will provide a basis for continued RNAP beacon experiments.

Overall this research has provided valuable insights into the molecular mechanisms of action of two novel regulators of bacterial RNAP. This work has contributed to our understanding of the regulation of transcription initiation as well as extending our appreciation for the complexity of the essential enzyme, RNAP.

REFERENCES

- Adelman,K., Brody,E.N., and Buckle,M. (1998) Stimulation of bacteriophage T4 middle transcription by the T4 proteins MotA and AsiA occurs at two distinct steps in the transcription cycle. *Proc Natl Acad Sci U S A* **95**: 15247-15252.
- Adelman,K., Orsini,G., Kolb,A., Graziani,L., and Brody,E.N. (1997) The interaction between the AsiA protein of bacteriophage T4 and the sigma70 subunit of Escherichia coli RNA polymerase. *J Biol Chem* **272**: 27435-27443.
- Alba,B.M., and Gross,C.A. (2004) Regulation of the Escherichia coli sigma-dependent envelope stress response. *Mol Microbiol* **52**: 613-619.
- Anderson,B.J., Larkin,C., Guja,K., and Schildbach,J.F. (2008) Using fluorophore-labeled oligonucleotides to measure affinities of protein-DNA interactions. *Methods Enzymol* **450**: 253-272.
- Arsene,F., Tomoyasu,T., and Bukau,B. (2000) The heat shock response of Escherichia coli. *Int J Food Microbiol* **55**: 3-9.
- Artsimovitch,I., Patlan,V., Sekine,S., Vassilyeva,M.N., Hosaka,T., Ochi,K. *et al.* (2004) Structural basis for transcription regulation by alarmone ppGpp. *Cell* **117**: 299-310.
- Artsimovitch,I., Svetlov,V., Murakami,K.S., and Landick,R. (2003) Co-overexpression of Escherichia coli RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. *J Biol Chem* **278**: 12344-12355.
- Baldwin,N.E., and Dombroski,A.J. (2001) Isolation and characterization of mutations in region 1.2 of Escherichia coli sigma70. *Mol Microbiol* **42**: 427-437.
- 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.
- Bar-Nahum,G., and Nudler,E. (2001) Isolation and characterization of sigma(70)-retaining transcription elongation complexes from Escherichia coli. *Cell* **106**: 443-451.
- Barne,K.A., Bown,J.A., Busby,S.J., and Minchin,S.D. (1997) Region 2.5 of the Escherichia coli RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J* **16**: 4034-4040.
- Barrios,H., Valderrama,B., and Morett,E. (1999) Compilation and analysis of sigma(54)-dependent promoter sequences. *Nucleic Acids Res* **27**: 4305-4313.
- Bellamy,S.R., and Baldwin,G.S. (2001) A kinetic analysis of substrate recognition by uracil-DNA glycosylase from herpes simplex virus type 1. *Nucleic Acids Res* **29**: 3857-3863.
- Blatter,E.E., Ross,W., Tang,H., Gourse,R.L., and Ebright,R.H. (1994) Domain organization of RNA polymerase alpha subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell* **78**: 889-896.

- Borukhov,S., and Nudler,E. (2008) RNA polymerase: the vehicle of transcription. *Trends Microbiol* **16**: 126-134.
- Browning,D.F., and Busby,S.J. (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**: 57-65.
- Buc,H., and McClure,W.R. (1985) Kinetics of open complex formation between Escherichia coli RNA polymerase and the lac UV5 promoter. Evidence for a sequential mechanism involving three steps. *Biochemistry* **24**: 2712-2723.
- Buchstein,S.R., and Hinkle,D.C. (1982) Genetic analysis of two bacterial RNA polymerase mutants that inhibit the growth of bacteriophage T7. *Mol Gen Genet* **188**: 211-218.
- Buck,M., Gallegos,M.T., Studholme,D.J., Guo,Y., and Gralla,J.D. (2000) The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *J Bacteriol* **182**: 4129-4136.
- Busby,S., and Ebright,R.H. (1999) Transcription activation by catabolite activator protein (CAP). *J Mol Biol* **293**: 199-213.
- Callaci,S., Heyduk,E., and Heyduk,T. (1999) Core RNA polymerase from E. coli induces a major change in the domain arrangement of the sigma 70 subunit. *Mol Cell* **3**: 229-238.
- Camara,B., Liu,M., Reynolds,J., Shadrin,A., Liu,B., Kwok,K. *et al.* (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.
- Camarero,J.A., Shekhtman,A., Campbell,E.A., Chlenov,M., Gruber,T.M., Bryant,D.A. *et al.* (2002) Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. *Proc Natl Acad Sci U S A* **99**: 8536-8541.
- Campbell,E.A., Muzzin,O., Chlenov,M., Sun,J.L., Olson,C.A., Weinman,O. *et al.* (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell* **9**: 527-539.
- Cannon,W., Gallegos,M.T., Casaz,P., and Buck,M. (1999) Amino-terminal sequences of sigmaN (sigma54) inhibit RNA polymerase isomerization. *Genes Dev* **13**: 357-370.
- Center,M.S. (1975) Role of gene 2 in bacteriophage T7 DNA synthesis. *J Virol* **16**: 94-100.
- Chakraborty,A., Wang,D., Ebright,Y.W., Korlann,Y., Kortkhonjia,E., Kim,T. *et al.* (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.
- Chen,J., Darst,S.A., and Thirumalai,D. (2010) Promoter melting triggered by bacterial RNA polymerase occurs in three steps. *Proc Natl Acad Sci U S A* **107**: 12523-12528.
- Chilcott,G.S., and Hughes,K.T. (2000) Coupling of flagellar gene expression to flagellar assembly in Salmonella enterica serovar typhimurium and Escherichia coli. *Microbiol Mol Biol Rev* **64**: 694-708.
- 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.

- Choy,H.E., and Adhya,S. (1992) Control of gal transcription through DNA looping: inhibition of the initial transcribing complex. *Proc Natl Acad Sci U S A* **89**: 11264-11268.
- Christie,G.E., Cale,S.B., Isaksson,L.A., Jin,D.J., Xu,M., Sauer,B., and Calendar,R. (1996) Escherichia coli rpoC397 encodes a temperature-sensitive C-terminal frameshift in the beta' subunit of RNA polymerase that blocks growth of bacteriophage P2. *J Bacteriol* **178**: 6991-6993.
- Ciampi,M.S. (2006) Rho-dependent terminators and transcription termination. *Microbiology* **152**: 2515-2528.
- Colland,F., Orsini,G., Brody,E.N., Buc,H., and Kolb,A. (1998) The bacteriophage T4 AsiA protein: a molecular switch for sigma 70-dependent promoters. *Mol Microbiol* **27**: 819-829.
- Cramer,P., Bushnell,D.A., and Kornberg,R.D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**: 1863-1876.
- D'Herelle,F. (2007) On an invisible microbe antagonistic toward dysenteric bacilli: brief note by Mr. F. D'Herelle, presented by Mr. Roux. 1917. *Res Microbiol* **158**: 553-554.
- Darst,S.A., Edwards,A.M., Kubalek,E.W., and Kornberg,R.D. (1991) Three-dimensional structure of yeast RNA polymerase II at 16 A resolution. *Cell* **66**: 121-128.
- Darst,S.A., Opalka,N., Chacon,P., Polyakov,A., Richter,C., Zhang,G., and Wriggers,W. (2002) Conformational flexibility of bacterial RNA polymerase. *Proc Natl Acad Sci U S A* **99**: 4296-4301.
- Darst,S.A., Polyakov,A., Richter,C., and Zhang,G. (1998) Insights into Escherichia coli RNA polymerase structure from a combination of x-ray and electron crystallography. *J Struct Biol* **124**: 115-122.
- deHaseth,P.L., and Helmann,J.D. (1995) Open complex formation by Escherichia coli RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA. *Mol Microbiol* **16**: 817-824.
- DeWynngaert,M.A., and Hinkle,D.C. (1979) Bacterial mutants affecting phage T7 DNA replication produce RNA polymerase resistant to inhibition by the T7 gene 2 protein. *J Biol Chem* **254**: 11247-11253.
- DeWynngaert,M.A., and Hinkle,D.C. (1980) Characterization of the defects in bacteriophage T7 DNA synthesis during growth in the Escherichia coli mutant tsnB. *J Virol* **33**: 780-788.
- Dombroski,A.J., Walter,W.A., and Gross,C.A. (1993a) Amino-terminal amino acids modulate sigma-factor DNA-binding activity. *Genes Dev* **7**: 2446-2455.
- Dombroski,A.J., Walter,W.A., and Gross,C.A. (1993b) The role of the sigma subunit in promoter recognition by RNA polymerase. *Cell Mol Biol Res* **39**: 311-317.
- Dombroski,A.J., Walter,W.A., Record,M.T., Jr., Siegele,D.A., and Gross,C.A. (1992) Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell* **70**: 501-512.
- Dove,S.L., and Hochschild,A. (2004) A bacterial two-hybrid system based on transcription activation. *Methods Mol Biol* **261**: 231-246.

- Dove, S.L., Joung, J.K., and Hochschild, A. (1997) Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature* **386**: 627-630.
- Durfee, T., Hansen, A.M., Zhi, H., Blattner, F.R., and Jin, D.J. (2008) Transcription profiling of the stringent response in *Escherichia coli*. *J Bacteriol* **190**: 1084-1096.
- Ederth, J., Artsimovitch, I., Isaksson, L.A., and Landick, R. (2002a) The downstream DNA jaw of bacterial RNA polymerase facilitates both transcriptional initiation and pausing. *J Biol Chem* **277**: 37456-37463.
- Ederth, J., Artsimovitch, I., Isaksson, L.A., and Landick, R. (2002b) The downstream DNA jaw of bacterial RNA polymerase facilitates both transcriptional initiation and pausing. *J Biol Chem* **277**: 37456-37463.
- Ederth, J., Isaksson, L.A., and Abdulkarim, F. (2002c) Origin-specific reduction of ColE1 plasmid copy number due to mutations in a distinct region of the *Escherichia coli* RNA polymerase. *Mol Genet Genomics* **267**: 587-592.
- Epshtein, V., Cardinale, C.J., Ruckenstein, A.E., Borukhov, S., and Nudler, E. (2007a) An allosteric path to transcription termination. *Mol Cell* **28**: 991-1001.
- Epshtein, V., Cardinale, C.J., Ruckenstein, A.E., Borukhov, S., and Nudler, E. (2007b) An allosteric path to transcription termination. *Mol Cell* **28**: 991-1001.
- Epshtein, V., Mustaev, A., Markovtsov, V., Bereshchenko, O., Nikiforov, V., and Goldfarb, A. (2002) Swing-gate model of nucleotide entry into the RNA polymerase active center. *Mol Cell* **10**: 623-634.
- Gardella, T., Moyle, H., and Susskind, M.M. (1989) A mutant *Escherichia coli* sigma 70 subunit of RNA polymerase with altered promoter specificity. *J Mol Biol* **206**: 579-590.
- Geertz, M., Travers, A., Mehandziska, S., Sobetzko, P., Chandra-Janga, S., Shimamoto, N., and Muskhelishvili, G. (2011) Structural coupling between RNA polymerase composition and DNA supercoiling in coordinating transcription: a global role for the omega subunit? *MBio* **2**.
- Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**: 1876-1882.
- Goff, C.G. (1984) Coliphage-induced ADP-ribosylation of *Escherichia coli* RNA polymerase. *Methods Enzymol* **106**: 418-429.
- Goldman, S.R., Ebright, R.H., and Nickels, B.E. (2009) Direct detection of abortive RNA transcripts in vivo. *Science* **324**: 927-928.
- Gries, T.J., Kontur, W.S., Capp, M.W., Saecker, R.M., and Record, M.T., Jr. (2010) One-step DNA melting in the RNA polymerase cleft opens the initiation bubble to form an unstable open complex. *Proc Natl Acad Sci U S A* **107**: 10418-10423.
- Gruber, T.M., and Bryant, D.A. (1997) Molecular systematic studies of eubacteria, using sigma70-type sigma factors of group 1 and group 2. *J Bacteriol* **179**: 1734-1747.
- Gruber, T.M., and Gross, C.A. (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* **57**: 441-466.

- Guild,N., Gayle,M., Sweeney,R., Hollingsworth,T., Modeer,T., and Gold,L. (1988) Transcriptional activation of bacteriophage T4 middle promoters by the motA protein. *J Mol Biol* **199**: 241-258.
- Gusarov,I., and Nudler,E. (1999) The mechanism of intrinsic transcription termination. *Mol Cell* **3**: 495-504.
- Harley,C.B., and Reynolds,R.P. (1987) Analysis of E. coli promoter sequences. *Nucleic Acids Res* **15**: 2343-2361.
- Hatfull,G.F., and Hendrix,R.W. (2011) Bacteriophages and their genomes. *Curr Opin Virol* **1**: 298-303.
- Haugen,S.P., Berkmen,M.B., Ross,W., Gaal,T., Ward,C., and Gourse,R.L. (2006) rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. *Cell* **125**: 1069-1082.
- Haugen,S.P., Ross,W., Manrique,M., and Gourse,R.L. (2008) Fine structure of the promoter-sigma region 1.2 interaction. *Proc Natl Acad Sci U S A* **105**: 3292-3297.
- Hauptenthal,J., Husecken,K., Negri,M., Maurer,C.K., and Hartmann,R.W. (2012) Influence of DNA Template Choice on Transcription and Inhibition of Escherichia coli RNA Polymerase. *Antimicrob Agents Chemother* **56**: 4536-4539.
- Hengge-Aronis,R. (2002) Stationary phase gene regulation: what makes an Escherichia coli promoter sigmaS-selective? *Curr Opin Microbiol* **5**: 591-595.
- Hesselbach,B.A., and Nakada,D. (1975) Inactive complex formation between E. coli RNA polymerase and inhibitor protein purified from T7 phage infected cells. *Nature* **258**: 354-357.
- Hesselbach,B.A., and Nakada,D. (1977a) "Host shutoff" function of bacteriophage T7: involvement of T7 gene 2 and gene 0.7 in the inactivation of Escherichia coli RNA polymerase. *J Virol* **24**: 736-745.
- Hesselbach,B.A., and Nakada,D. (1977b) I protein: bacteriophage T7-coded inhibitor of Escherichia coli RNA polymerase. *J Virol* **24**: 746-760.
- Heyduk,E., Kuznedelov,K., Severinov,K., and Heyduk,T. (2006) A consensus adenine at position -11 of the nontemplate strand of bacterial promoter is important for nucleation of promoter melting. *J Biol Chem* **281**: 12362-12369.
- Ishihama,A. (1981) Subunit of assembly of Escherichia coli RNA polymerase. *Adv Biophys* **14**: 1-35.
- James,E., Liu,M., Sheppard,C., Mekler,V., Camara,B., Liu,B. *et al.* (2012) Structural and Mechanistic Basis for the Inhibition of Escherichia coli RNA Polymerase by T7 Gp2. *Mol Cell*.
- Jokerst,R.S., Weeks,J.R., Zehring,W.A., and Greenleaf,A.L. (1989) Analysis of the gene encoding the largest subunit of RNA polymerase II in Drosophila. *Mol Gen Genet* **215**: 266-275.
- Joseph R.Lakowicz (2012) Fluorescence Anisotropy. In Principles of Fluorescence Spectroscopy., pp. 291-319.
- Kedzierska,B., Szambowska,A., Herman-Antosiewicz,A., Lee,D.J., Busby,S.J., Wegrzyn,G., and Thomas,M.S. (2007) The C-terminal domain of the Escherichia coli RNA polymerase alpha subunit plays a role in the CI-dependent activation of the bacteriophage lambda pM promoter. *Nucleic Acids Res* **35**: 2311-2320.

- 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.
- Kontur, W.S., Capp, M.W., Gries, T.J., Saecker, R.M., and Record, M.T., Jr. (2010) Probing DNA binding, DNA opening, and assembly of a downstream clamp/jaw in Escherichia coli RNA polymerase-lambdaP(R) promoter complexes using salt and the physiological anion glutamate. *Biochemistry* **49**: 4361-4373.
- Kulbachinskiy, A., Feklistov, A., Krashennnikov, I., Goldfarb, A., and Nikiforov, V. (2004) Aptamers to Escherichia coli core RNA polymerase that sense its interaction with rifampicin, sigma-subunit and GreB. *Eur J Biochem* **271**: 4921-4931.
- Kumar, A., Malloch, R.A., Fujita, N., Smillie, D.A., Ishihama, A., and Hayward, R.S. (1993) The minus 35-recognition region of Escherichia coli sigma 70 is inessential for initiation of transcription at an "extended minus 10" promoter. *J Mol Biol* **232**: 406-418.
- Kuznedelov, K., Korzheva, N., Mustaev, A., and Severinov, K. (2002a) Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. *EMBO J* **21**: 1369-1378.
- Kuznedelov, K., Minakhin, L., Niedziela-Majka, A., Dove, S.L., Rogulja, D., Nickels, B.E. *et al.* (2002b) A role for interaction of the RNA polymerase flap domain with the sigma subunit in promoter recognition. *Science* **295**: 855-857.
- Lane, W.J., and Darst, S.A. (2010) Molecular evolution of multisubunit RNA polymerases: sequence analysis. *J Mol Biol* **395**: 671-685.
- LeClerc, J.E., and Richardson, C.C. (1979a) Gene 2 protein of bacteriophage T7: purification and requirement for packaging of T7 DNA in vitro. *Proc Natl Acad Sci U S A* **76**: 4852-4856.
- LeClerc, J.E., and Richardson, C.C. (1979b) Gene 2 protein of bacteriophage T7: purification and requirement for packaging of T7 DNA in vitro. *Proc Natl Acad Sci U S A* **76**: 4852-4856.
- Liao, Y.D., and Kuo, T.T. (1986) Loss of sigma-factor of RNA polymerase of Xanthomonas campestris pv. oryzae during phage Xp10 infection. *J Biol Chem* **261**: 13714-13719.
- Lim, H.M., Lee, H.J., Roy, S., and Adhya, S. (2001) A "master" in base unpairing during isomerization of a promoter upon RNA polymerase binding. *Proc Natl Acad Sci U S A* **98**: 14849-14852.
- Lisser, S., and Margalit, H. (1993) Compilation of E. coli mRNA promoter sequences. *Nucleic Acids Res* **21**: 1507-1516.
- Liu, K., Zhang, Y., Severinov, K., Das, A., and Hanna, M.M. (1996) Role of Escherichia coli RNA polymerase alpha subunit in modulation of pausing, termination and anti-termination by the transcription elongation factor NusA. *EMBO J* **15**: 150-161.
- Malhotra, A., Severinova, E., and Darst, S.A. (1996) Crystal structure of a sigma 70 subunit fragment from E. coli RNA polymerase. *Cell* **87**: 127-136.
- Marr, M.T., and Roberts, J.W. (1997) Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science* **276**: 1258-1260.

- Martinez-Antonio,A., and Collado-Vides,J. (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* **6**: 482-489.
- Mathew,R., and Chatterji,D. (2006) The evolving story of the omega subunit of bacterial RNA polymerase. *Trends Microbiol* **14**: 450-455.
- Mekler,V., Kortkhonjia,E., Mukhopadhyay,J., Knight,J., Revyakin,A., Kapanidis,A.N. *et al.* (2002) Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* **108**: 599-614.
- Mekler,V., Minakhin,L., and Severinov,K. (2011a) A critical role of downstream RNA polymerase-promoter interactions in the formation of initiation complex. *J Biol Chem* **286**: 22600-22608.
- Mekler,V., Minakhin,L., Sheppard,C., Wigneshweraraj,S., and Severinov,K. (2011b) Molecular mechanism of transcription inhibition by phage T7 gp2 protein. *J Mol Biol* **413**: 1016-1027.
- Mekler,V., Pavlova,O., and Severinov,K. (2011c) Interaction of Escherichia coli RNA polymerase sigma70 subunit with promoter elements in the context of free sigma70, RNA polymerase holoenzyme, and the beta'-sigma70 complex. *J Biol Chem* **286**: 270-279.
- Meng,W., Savery,N.J., Busby,S.J., and Thomas,M.S. (2000) The Escherichia coli RNA polymerase alpha subunit linker: length requirements for transcription activation at CRP-dependent promoters. *EMBO J* **19**: 1555-1566.
- Minakhin,L., Camarero,J.A., Holford,M., Parker,C., Muir,T.W., and Severinov,K. (2001) Mapping the molecular interface between the sigma(70) subunit of E. coli RNA polymerase and T4 AsiA. *J Mol Biol* **306**: 631-642.
- Minakhin,L., Niedziela-Majka,A., Kuznedelov,K., Adelman,K., Urbauer,J.L., Heyduk,T., and Severinov,K. (2003) Interaction of T4 AsiA with its target sites in the RNA polymerase sigma70 subunit leads to distinct and opposite effects on transcription. *J Mol Biol* **326**: 679-690.
- Mooney,R.A., Artsimovitch,I., and Landick,R. (1998) Information processing by RNA polymerase: recognition of regulatory signals during RNA chain elongation. *J Bacteriol* **180**: 3265-3275.
- Mulligan,M.E., Brosius,J., and McClure,W.R. (1985) Characterization in vitro of the effect of spacer length on the activity of Escherichia coli RNA polymerase at the TAC promoter. *J Biol Chem* **260**: 3529-3538.
- Murakami,K.S., Masuda,S., Campbell,E.A., Muzzin,O., and Darst,S.A. (2002a) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**: 1285-1290.
- Murakami,K.S., Masuda,S., and Darst,S.A. (2002b) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* **296**: 1280-1284.
- Nagai,H., and Shimamoto,N. (1997) Regions of the Escherichia coli primary sigma factor sigma70 that are involved in interaction with RNA polymerase core enzyme. *Genes Cells* **2**: 725-734.
- Naryshkina,T., Kuznedelov,K., and Severinov,K. (2006) The role of the largest RNA polymerase subunit lid element in preventing the formation of extended RNA-DNA hybrid. *J Mol Biol* **361**: 634-643.

- Nechaev,S., Chlenov,M., and Severinov,K. (2000) Dissection of two hallmarks of the open promoter complex by mutation in an RNA polymerase core subunit. *J Biol Chem* **275**: 25516-25522.
- Nechaev,S., and Severinov,K. (1999) Inhibition of Escherichia coli RNA polymerase by bacteriophage T7 gene 2 protein. *J Mol Biol* **289**: 815-826.
- Nechaev,S., and Severinov,K. (2003) Bacteriophage-induced modifications of host RNA polymerase. *Annu Rev Microbiol* **57**: 301-322.
- Nechaev,S., Yuzenkova,Y., Niedziela-Majka,A., Heyduk,T., and Severinov,K. (2002) A novel bacteriophage-encoded RNA polymerase binding protein inhibits transcription initiation and abolishes transcription termination by host RNA polymerase. *J Mol Biol* **320**: 11-22.
- Nickels,B.E., Dove,S.L., Murakami,K.S., Darst,S.A., and Hochschild,A. (2002) Protein-protein and protein-DNA interactions of sigma70 region 4 involved in transcription activation by lambdacl. *J Mol Biol* **324**: 17-34.
- Nudler,E. (2009) RNA polymerase active center: the molecular engine of transcription. *Annu Rev Biochem* **78**: 335-361.
- 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**.
- Orsini,G., Ouhammouch,M., Le Caer,J.P., and Brody,E.N. (1993) The asiA gene of bacteriophage T4 codes for the anti-sigma 70 protein. *J Bacteriol* **175**: 85-93.
- Ouhammouch,M., Adelman,K., Harvey,S.R., Orsini,G., and Brody,E.N. (1995) Bacteriophage T4 MotA and AsiA proteins suffice to direct Escherichia coli RNA polymerase to initiate transcription at T4 middle promoters. *Proc Natl Acad Sci U S A* **92**: 1451-1455.
- Paget,M.S., and Helmann,J.D. (2003) The sigma70 family of sigma factors. *Genome Biol* **4**: 203.
- Panchuk-Voloshina,N., Haugland,R.P., Bishop-Stewart,J., Bhalgat,M.K., Millard,P.J., Mao,F. *et al.* (1999) Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. *J Histochem Cytochem* **47**: 1179-1188.
- Paul,B.J., Barker,M.M., Ross,W., Schneider,D.A., Webb,C., Foster,J.W., and Gourse,R.L. (2004) DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**: 311-322.
- Paul,B.J., Berkmen,M.B., and Gourse,R.L. (2005) DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc Natl Acad Sci U S A* **102**: 7823-7828.
- Qimron,U., Kulczyk,A.W., Hamdan,S.M., Tabor,S., and Richardson,C.C. (2008) Inadequate inhibition of host RNA polymerase restricts T7 bacteriophage growth on hosts overexpressing udk. *Mol Microbiol* **67**: 448-457.
- Rogozina,A., Zaychikov,E., Buckle,M., Heumann,H., and Sclavi,B. (2009) DNA melting by RNA polymerase at the T7A1 promoter precedes the rate-limiting step at 37 degrees C and results in the accumulation of an off-pathway intermediate. *Nucleic Acids Res* **37**: 5390-5404.

- Ross,W., Gosink,K.K., Salomon,J., Igarashi,K., Zou,C., Ishihama,A. *et al.* (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**: 1407-1413.
- Saecker,R.M., Record,M.T., Jr., and 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., and Severinov,K. (2010) The role of the T7 Gp2 inhibitor of host RNA polymerase in phage development. *J Mol Biol* **402**: 118-126.
- Schwartz,E.C., Shekhtman,A., Dutta,K., Pratt,M.R., Cowburn,D., Darst,S., and Muir,T.W. (2008) A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. *Chem Biol* **15**: 1091-1103.
- Sclavi,B., Zaychikov,E., Rogozina,A., Walther,F., Buckle,M., and Heumann,H. (2005) Real-time characterization of intermediates in the pathway to open complex formation by Escherichia coli RNA polymerase at the T7A1 promoter. *Proc Natl Acad Sci U S A* **102**: 4706-4711.
- Severinova,E., and Severinov,K. (2006) Localization of the Escherichia coli RNA polymerase beta' subunit residue phosphorylated by bacteriophage T7 kinase Gp0.7. *J Bacteriol* **188**: 3470-3476.
- Severinova,E., Severinov,K., and Darst,S.A. (1998) Inhibition of Escherichia coli RNA polymerase by bacteriophage T4 AsiA. *J Mol Biol* **279**: 9-18.
- Shadrin,A., Sheppard,C., Severinov,K., Matthews,S., and Wigneshweraraj,S. (2012) 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. *Microbiology*.
- Shanblatt,S.H., and Nakada,D. (1982) Escherichia coli mutant which restricts T7 bacteriophage has an altered RNA polymerase. *J Virol* **42**: 1123-1126.
- Sharp,M.M., Chan,C.L., Lu,C.Z., Marr,M.T., Nechaev,S., Merritt,E.W. *et al.* (1999) The interface of sigma with core RNA polymerase is extensive, conserved, and functionally specialized. *Genes Dev* **13**: 3015-3026.
- Sheppard,C., Camara,B., Shadrin,A., Akulenko,N., Liu,M., Baldwin,G. *et al.* (2011) Inhibition of Escherichia coli RNAP by T7 Gp2 protein: role of negatively charged strip of amino acid residues in Gp2. *J Mol Biol* **407**: 623-632.
- Siegele,D.A., Hu,J.C., Walter,W.A., and Gross,C.A. (1989) Altered promoter recognition by mutant forms of the sigma 70 subunit of Escherichia coli RNA polymerase. *J Mol Biol* **206**: 591-603.
- Sosunov,V., Sosunova,E., Mustaev,A., Bass,I., Nikiforov,V., and Goldfarb,A. (2003) Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J* **22**: 2234-2244.
- Sosunov,V., Zorov,S., Sosunova,E., Nikolaev,A., Zakeyeva,I., Bass,I. *et al.* (2005) The involvement of the aspartate triad of the active center in all catalytic activities of multisubunit RNA polymerase. *Nucleic Acids Res* **33**: 4202-4211.
- Srivatsan,A., and Wang,J.D. (2008) Control of bacterial transcription, translation and replication by (p)ppGpp. *Curr Opin Microbiol* **11**: 100-105.

- Studier,F.W. (1972) Bacteriophage T7. *Science* **176**: 367-376.
- Studier,F.W. (1973) Genetic analysis of non-essential bacteriophage T7 genes. *J Mol Biol* **79**: 227-236.
- Studier,F.W., and Moffatt,B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**: 113-130.
- Sweetser,D., Nonet,M., and Young,R.A. (1987) Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. *Proc Natl Acad Sci U S A* **84**: 1192-1196.
- Symersky,J., Perederina,A., Vassilyeva,M.N., Svetlov,V., Artsimovitch,I., and Vassilyev,D.G. (2006) Regulation through the RNA polymerase secondary channel. Structural and functional variability of the coiled-coil transcription factors. *J Biol Chem* **281**: 1309-1312.
- Tomsic,M., Tsujikawa,L., Panaghie,G., Wang,Y., Azok,J., and deHaseth,P.L. (2001) Different roles for basic and aromatic amino acids in conserved region 2 of Escherichia coli sigma(70) in the nucleation and maintenance of the single-stranded DNA bubble in open RNA polymerase-promoter complexes. *J Biol Chem* **276**: 31891-31896.
- Touloukhonov,I., and Landick,R. (2003) The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. *Mol Cell* **12**: 1125-1136.
- Touloukhonov,I., and Landick,R. (2006) The role of the lid element in transcription by E. coli RNA polymerase. *J Mol Biol* **361**: 644-658.
- Twort,F.W. (1936) Further Investigations on the Nature of Ultra-Microscopic Viruses and their Cultivation. *J Hyg (Lond)* **36**: 204-235.
- van Hijum,S.A., Medema,M.H., and Kuipers,O.P. (2009) Mechanisms and evolution of control logic in prokaryotic transcriptional regulation. *Microbiol Mol Biol Rev* **73**: 481-509, Table.
- Vassilyev,D.G., Sekine,S., Laptenko,O., Lee,J., Vassilyeva,M.N., Borukhov,S., and Yokoyama,S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**: 712-719.
- Vassilyev,D.G., Vassilyeva,M.N., Perederina,A., Tahirov,T.H., and Artsimovitch,I. (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**: 157-162.
- Visca,P., Leoni,L., Wilson,M.J., and Lamont,I.L. (2002) Iron transport and regulation, cell signalling and genomics: lessons from Escherichia coli and Pseudomonas. *Mol Microbiol* **45**: 1177-1190.
- Wang,D., Bushnell,D.A., Westover,K.D., Kaplan,C.D., and Kornberg,R.D. (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**: 941-954.
- Wang,D., Meier,T.I., Chan,C.L., Feng,G., Lee,D.N., and Landick,R. (1995) Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex. *Cell* **81**: 341-350.
- Wang,Y., Severinov,K., Loizos,N., Fenyo,D., Heyduk,E., Heyduk,T. *et al.* (1997) Determinants for Escherichia coli RNA polymerase assembly within the beta subunit. *J Mol Biol* **270**: 648-662.

- Wigneshweraraj,S., Bose,D., Burrows,P.C., Joly,N., Schumacher,J., Rappas,M. *et al.* (2008) Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor. *Mol Microbiol* **68**: 538-546.
- Wigneshweraraj,S.R., Burrows,P.C., Nechaev,S., Zenkin,N., Severinov,K., and Buck,M. (2004) Regulated communication between the upstream face of RNA polymerase and the beta' subunit jaw domain. *EMBO J* **23**: 4264-4274.
- Wigneshweraraj,S.R., Burrows,P.C., Severinov,K., and Buck,M. (2005) Stable DNA opening within open promoter complexes is mediated by the RNA polymerase beta'-jaw domain. *J Biol Chem* **280**: 36176-36184.
- Wigneshweraraj,S.R., Nechaev,S., Bordes,P., Jones,S., Cannon,W., Severinov,K., and Buck,M. (2003) Enhancer-dependent transcription by bacterial RNA polymerase: the beta subunit downstream lobe is used by sigma 54 during open promoter complex formation. *Methods Enzymol* **370**: 646-657.
- Wigneshweraraj,S.R., Nechaev,S., Severinov,K., and Buck,M. (2002) Beta subunit residues 186-433 and 436-445 are commonly used by Esigma54 and Esigma70 RNA polymerase for open promoter complex formation. *J Mol Biol* **319**: 1067-1083.
- Wigneshweraraj,S.R., Savalia,D., Severinov,K., and Buck,M. (2006) Interplay between the beta' clamp and the beta' jaw domains during DNA opening by the bacterial RNA polymerase at sigma54-dependent promoters. *J Mol Biol* **359**: 1182-1195.
- Xu,M., Zhou,Y.N., Goldstein,B.P., and Jin,D.J. (2005) Cross-resistance of Escherichia coli RNA polymerases conferring rifampin resistance to different antibiotics. *J Bacteriol* **187**: 2783-2792.
- Young,B.A., Anthony,L.C., Gruber,T.M., Arthur,T.M., Heyduk,E., Lu,C.Z. *et al.* (2001) A coiled-coil from the RNA polymerase beta' subunit allosterically induces selective nontemplate strand binding by sigma(70). *Cell* **105**: 935-944.
- Yuzenkova,J., Nechaev,S., Berlin,J., Rogulja,D., Kuznedelov,K., Inman,R. *et al.* (2003) Genome of Xanthomonas oryzae bacteriophage Xp10: an odd T-odd phage. *J Mol Biol* **330**: 735-748.
- Yuzenkova,Y., Tadigotla,V.R., Severinov,K., and Zenkin,N. (2011) A new basal promoter element recognized by RNA polymerase core enzyme. *EMBO J* **30**: 3766-3775.
- Yuzenkova,Y., Zenkin,N., and Severinov,K. (2008) Mapping of RNA polymerase residues that interact with bacteriophage Xp10 transcription antitermination factor p7. *J Mol Biol* **375**: 29-35.
- Zaychikov,E., Martin,E., Denissova,L., Kozlov,M., Markovtsov,V., Kashlev,M. *et al.* (1996) Mapping of catalytic residues in the RNA polymerase active center. *Science* **273**: 107-109.
- 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 Å resolution. *Cell* **98**: 811-824.
- Zhang,J., Palangat,M., and Landick,R. (2010) Role of the RNA polymerase trigger loop in catalysis and pausing. *Nat Struct Mol Biol* **17**: 99-104.



COMMUNICATION

Inhibition of *Escherichia coli* RNAP by T7 Gp2 Protein: Role of Negatively Charged Strip of Amino Acid Residues in Gp2

Carol Sheppard^{1†}, Beatriz Cámara^{1†}, Andrey Shadrin^{2,3,4},
Natalia Akulenko^{3,4}, Minhao Liu⁵, Geoff Baldwin⁶,
Konstantin Severinov^{3,4}, Ernesto Cota⁵,
Steve Matthews⁵ and Siva R Wigneshweraraj^{1*}

¹Section of Microbiology, Faculty of Medicine & Centre for Molecular Microbiology and Infection, Imperial College London, SW7 2AZ, UK

²Skryabin Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Science Pushchino, Russia

³Waksman Institute for Microbiology and Department of Molecular Biology and Biochemistry, Rutgers, State University of New Jersey, Piscataway, NJ 08854, USA

⁴Institutes of Molecular Genetics and Gene Biology, Russian Academy of Sciences, Moscow, Russia

⁵Division of Molecular Biosciences, Imperial College London, SW7 2AZ, UK

⁶Centre for Structural Biology, Imperial College London, SW7 2AZ, UK

Received 19 November 2010;
received in revised form
3 February 2011;
accepted 4 February 2011
Available online
18 February 2011

Edited by: R. Ebricht

Keywords:

bacteriophage T7;
RNA polymerase;
RNA polymerase inhibitor;
T7 Gp2;
transcription regulation

Gp2, a 7 kDa protein encoded by T7 bacteriophage, is a potent inhibitor of *Escherichia coli* RNA polymerase (RNAP), the enzyme responsible for transcription of all bacterial genes and early viral genes. A prominent feature in the structure of Gp2 is a contiguous strip of seven negatively charged amino acid residues (negatively charged strip or NCS), located along one side of the molecule. The role of the NCS in Gp2 function is not known. Here, the *in vivo* and *in vitro* properties of altered forms of Gp2 with amino acid substitutions in the NCS are described. While mutations in the NCS do not compromise the folding or the ability of Gp2 to bind to the RNAP β' subunit, disruption of the NCS significantly attenuates Gp2 function *in vivo* and its ability to inhibit RNAP *in vitro*. Efficient inhibition of the RNAP by Gp2 also involves the amino terminal region 1 domain of the RNAP promoter specificity subunit σ^{70} , located in the vicinity of the primary Gp2 binding site in β' . The results are discussed in the context of hypothetical molecular mechanisms of RNAP inhibition by Gp2.

© 2011 Elsevier Ltd. All rights reserved.

*Corresponding author. E-mail address:
s.r.wig@imperial.ac.uk.

† Contributed equally to this work.

Abbreviations used: RNAP, RNA polymerase; R_{PC}, closed promoter complex; R_{PO}, open promoter complex; NCS, negatively charged strip; DBC, DNA-binding channel; FP, fluorescence polarization.

Transcription of DNA is catalysed by the DNA-dependent RNA polymerase (RNAP). Bacterial RNAP is a multisubunit enzyme comprised of a five-subunit catalytic core ($\alpha_2\beta\beta'\omega$; abbreviated E) and one of several σ subunits, each conferring upon the catalytic core an ability to initiate transcription from specific DNA sites called promoters.¹ Most bacterial promoters with conserved sequences near

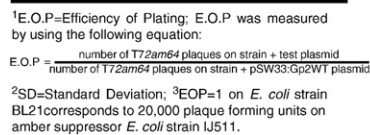


Fig. 1 (*legend on next page*)

positions -35 and -10, with respect to the transcription start site (the +1 site), depend on the RNAP containing a housekeeping σ factor (σ^{70} in *Escherichia coli*). Under certain stress conditions, a relatively small subset of *E. coli* promoters is used by the RNAP containing a major variant σ factor, σ^{54} .² The initiation of transcription by $E\sigma^{70}$ is a multistep process and represents a major regulatory step of gene expression. The σ^{70} factor directed binding of the RNAP to the promoter results in the formation of a competitor-sensitive closed promoter complex (RPC). The transition of the RPC to transcription-initiation competent open promoter complex (RPO) occurs through several intermediates and is accompanied by large-scale conformational changes in $E\sigma^{70}$ and promoter DNA.³⁻⁵ In RPO, the strands of promoter DNA are locally melted to form the transcription bubble, the +1 site of the template DNA strand is placed at the RNAP catalytic site and the duplex DNA downstream of the transcription bubble is tightly held in a trough, called the downstream DNA binding channel (downstream-DBC), formed by the structurally flexible RNAP domains β' jaw, β' clamp, β' I6 and β downstream lobe.^{4,6} Interactions between promoter DNA and the downstream-DBC are responsible for several of the discrete kinetic steps observed *en route* to the RPO as well as for stable maintenance of the transcription bubble once the RPO formation is complete.^{3,7}

Bacterial RNAP is a major target for numerous regulatory proteins and small ligands that fine-tune gene expression levels to the changing needs of the cell. Several bacteriophages (phages) encode small molecular mass regulatory proteins that modify the activity of host RNAP to favour the expression of phage genes.⁸ The *E. coli* bacteriophage T7 provides a paradigm of how a phage-encoded regulatory protein modifies the activity of host bacterium RNAP. The transcription of the T7 genome occurs in three stages. Early T7 genes are transcribed by $E\sigma^{70}$, whereas middle and late T7 genes are transcribed by phage-encoded RNAP. The product of the middle T7 gene 2, a 7 kDa protein called Gp2,

binds to the β' jaw domain in the downstream-DBC of the host RNAP and prevents RPO formation by $E\sigma^{70}$.^{9,10} A contiguous strip of seven negatively charged amino acids (E21, E34, D37, E38, E41, E44 and E53) at one side of the molecule and two invariant and functionally indispensable positively charged arginine residues (R56 and R58) on the opposite side are notable features in the high-resolution solution structure of Gp2 (Fig. 1a).¹⁰ The role of the negatively charged strip (NCS) in T7 Gp2 is not known. Because R56 and R58 are essential for Gp2 interaction with the β' jaw domain, NCS, which should be located on the side opposite the RNAP-binding side of Gp2, could be projected into downstream-DBC and interfere with the positioning of DNA in the channel and/or with σ^{70} region 1, which is thought to occupy the downstream-DBC in free $E\sigma^{70}$.⁵ Here, we describe the *in vivo* and *in vitro* properties of four altered forms of Gp2, which carry alanine or serine substitutions of two, five, six, or all seven amino acid residues of the NCS. Our results show that removal of the negatively charged side chains from the NCS does not affect the structural integrity of Gp2 or compromise its ability to bind $E\sigma^{70}$. However, the mutant Gp2 forms are significantly attenuated for transcription inhibition, with progressive decrease in inhibition efficiency following removal of negatively charged side chains from the NCS. We demonstrate that efficient inhibition of $E\sigma^{70}$ by Gp2 requires region 1 of σ^{70} . The results thus indicate an important role of the NCS in the inhibition of the RNAP by Gp2 and support a model in which the primary Gp2-RNAP interaction serves to position the NCS in the downstream-DBC.

Mutagenesis of the NCS in T7 Gp2

The NMR structure of Gp2 shows that the protein adopts a compact globular fold comprising a three-stranded β -sheet packed against an α -helix in a $\beta_1\beta_2\alpha_1\beta_3$ topology with the central β_2 strand flanked by anti-parallel β_1 and β_3 strands. Residues E21, E34, D37, E38, E41, E44 and E53, which form the NCS, are located in or are proximal to the α_1

Fig. 1. The negatively charged strip of amino acids in Gp2. (a) Ribbon representation of Gp2 highlighting (in red) the side chains of the negatively charged amino acid residues that contribute to the NCS. The functionally essential R56 and R58 are indicated. (b) Effect of single alanine substitutions on inhibition of $E\sigma^{70}$. The graph indicates results from an *in vitro* transcription inhibition assay and reports the activity of $E\sigma^{70}$ in synthesizing the ApApUpU transcript from *lacUV5* in the presence of equimolar amounts of each Gp2 single-alanine mutant. The results are normalized to the activity of $E\sigma^{70}$ in the absence of Gp2. In independent experiments, data obtained fall within 10% of the values shown. The *in vitro* transcription inhibition assay was done as described¹⁰ and the schematic at the top indicates the concentration of reaction components, time added and incubation time. (c) The structure of Gp2 shown as a space filled representation and in the same orientation and colouring as in (a). The consecutive alanine or serine substitutions introduced into Gp2 to generate the Gp2 mutants are depicted in yellow. The lower panel shows an alignment of the amino acid sequences of Gp2-like proteins from known phage genomes in the single-letter amino acid code. The localization of β strands and the α helix in the structure of Gp2 are indicated. The symbols in the sequence alignment are: (*), identical residues; (:), conserved residues; (.) semi-conserved residues. The amino acid residues targeted for mutagenesis are highlighted in red. (e) Plating efficiency of T72am64 phage on *E. coli* strain BL21 transformed with pET33:Gp2 encoding mutant Gp2 proteins.

helix (E34, D37, E38, E41 and E44), the β_1 (E21), and β_3 strands (E53) (Fig. 1a). Initially, we made single alanine substitutions at each NCS residue and tested the ability of Gp2 point mutants to inhibit *E. coli* E σ^{70} in an *in vitro* transcription assay. In this assay, the incubation of E σ^{70} with equimolar amounts of ^{WT}Gp2 before the addition of the *lacUV5* promoter effectively abolishes the synthesis of the ApApUpU transcript by E σ^{70} (Fig. 1b).¹⁰ The results revealed that single alanine substitutions of NCS residues had no detectable effect on the ability of Gp2 to inhibit E σ^{70} in an *in vitro* transcription assay, indicating that the NCS either does not contribute to the function of Gp2 or that the negatively charged side chains of the NCS act synergistically. Therefore, we made four additional mutants in which two (^{Mut2}Gp2), five (^{Mut5}Gp2), six (^{Mut6}Gp2) or all seven (^{Mut7}Gp2) NCS residues were substituted with alanine or serine (Fig. 1c). ^{Mut2}Gp2 contains an alanine (D37A) and a serine (E38S) in two positions that contain negatively charged side chains in all 13 known Gp2 homologs. ^{Mut5}Gp2 contains alanine (D37 and E41) and serine (E34, E38 and E44) instead of negatively charged residues in or close to the α_1 helix (Fig. 1a). The choice of residues introduced instead of the native negatively charged amino acids was structure-based and dictated by the desire to maintain the structural integrity of α_1 . In ^{Mut6}Gp2 and ^{Mut7}Gp2, the remaining NCS residues (E53, or E53 and E21, respectively) are changed to alanine.

Amino acid substitutions in NCS impair the ability of T7 to infect *E. coli*

Initially, the activities of all Gp2 mutants were tested in an *in vivo* complementation assay.¹⁰ To this end, *E. coli* BL21 cells containing the pET-based plasmid expressing wild type gene 2 or gene 2 mutants were infected with T7 phage harbouring an *am64* amber mutation in gene 2. T7^{*am64*} infections of wild-type *E. coli* BL21 cells are not productive, since Gp2 function is essential for phage development.¹¹ However, *E. coli* BL21 cells containing the plasmid expressing ^{WT}Gp2 are productively infected by T7^{*am64*}, as judged by the efficiency of plaque formation, calculated as a ratio of numbers of plaques observed on non-suppressing to suppressing hosts. In contrast, the efficiency of plaque formation is much reduced (>10-fold) on lawns of *E. coli* BL21 cells harbouring the functionally defective ^{R56E}Gp2 mutant that is defective for E σ^{70} binding (Fig. 1d).¹⁰ Plaque formation by T7^{*am64*} on lawns formed by *E. coli* BL21 cells harbouring plasmids expressing the NCS mutants was also affected but to a lesser extent: ^{Mut5}Gp2, ^{Mut6}Gp2 and ^{Mut7}Gp2-producing cells allowed ~58%, ~62% and ~68%, respectively, less plaques to form than those formed by cells expressing ^{WT}Gp2. Plaque formation

on lawns of cells containing plasmid expressing ^{Mut2}Gp2 was reduced by only ~27%. Overall, the results of the *in vivo* complementation assay indicate that the NCS contributes to Gp2 function *in vivo*, though the effects of NCS substitutions, including the mutation that removes all the negatively charged amino acids from the NCS (^{Mut7}Gp2), is less than the effect of the point substitution R56E affecting the Gp2–RNAP interaction. Attempts to determine the relative *in vivo* expression levels of wild-type and mutant Gp2 by Western blotting were unsuccessful because the expression level of plasmid-borne Gp2 appeared to be below the minimum detection limit of the anti-Gp2 antibodies available to us (data not shown).

Multiple amino acid substitutions in NCS do not affect the structure of the protein

To evaluate the properties of Gp2 mutants at a molecular level, all Gp2 mutants, as well as the ^{WT}Gp2 control, were overproduced as amino-terminal His₆-tagged fusion proteins and purified to >90% homogeneity by nickel affinity chromatography. The level of expression of all Gp2 mutants was comparable to that of ^{WT}Gp2 (data not shown). Initially, the (¹H)NMR spectra of all mutants were measured and compared to that of ^{WT}Gp2. The (¹H)NMR spectrum of ^{WT}Gp2 is characteristic of a small, folded domain, as judged by the dispersion of resonances below ~0.5 ppm (corresponding to H ^{γ 1} and H ^{γ 2} protons from V20 at –0.126 ppm and 0.046 ppm, respectively; Fig. 2a).¹⁰ These resonances typically represent buried protons of methyl or methylene groups in hydrophobic cores. Their frequencies are shifted by electronic ring currents from the plane of a neighbouring aromatic side chain. Similarly, there is good dispersion of resonances in the amide and aromatic regions (6.5–9 ppm; Fig. 2a).¹⁰ The results indicate that these features are retained in the spectra of ^{Mut2}Gp2, ^{Mut5}Gp2, ^{Mut6}Gp2 and ^{Mut7}Gp2 (Fig. 2a and data not shown), indicating that removal of the negatively charged side chains at positions 21, 34, 37, 38, 41, 44 and 53 does not affect the native structure of Gp2. In Fig. 2b, the (¹H)NMR spectrum of a Gp2 in which the structure is significantly disrupted by a mutation (F16A) is shown for comparison.¹⁰

Multiple amino acid substitutions in the NCS have little effect on the affinity of Gp2 for *E. coli* RNAP

We developed a fluorescence polarization (FP) assay to quantify the strength of the interaction between Gp2 and E σ^{70} to measure the effect of NCS substitutions on RNAP binding. Gp2 was modified with a thiol-reactive Alexa488-

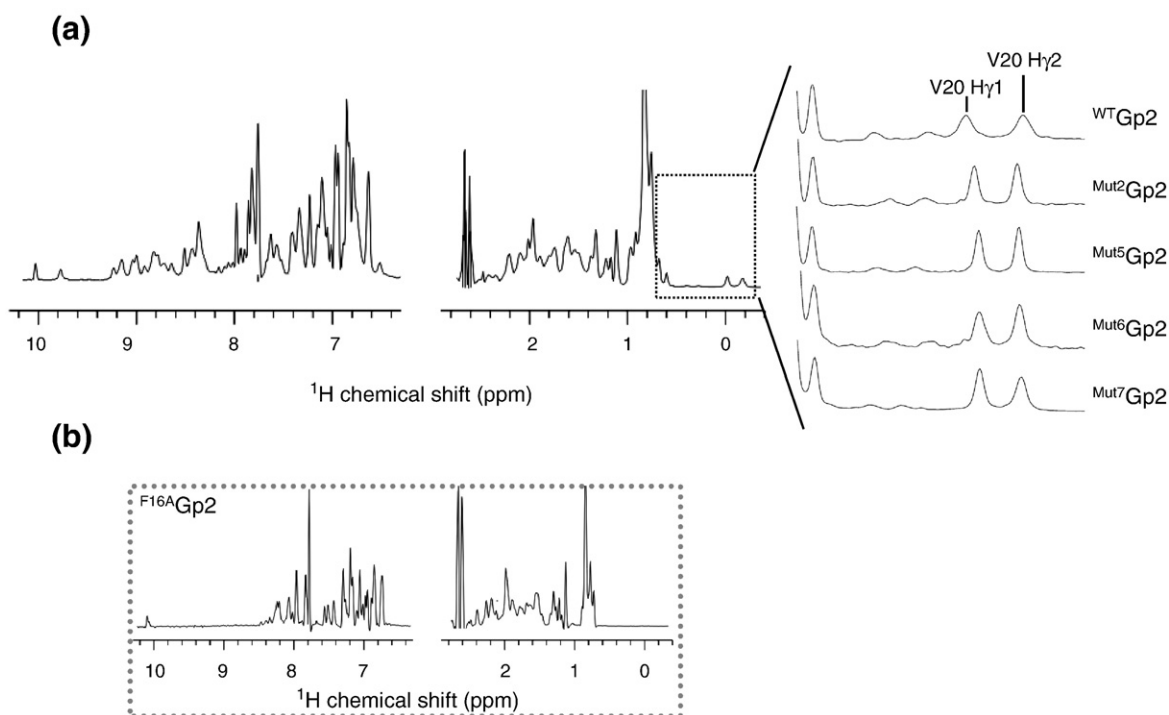


Fig. 2. The (^1H)NMR spectra of Gp2. (a) Detail of proton (^1H)NMR spectra from $^{\text{WT}}$ Gp2 obtained as described,¹⁰ showing the amide/aromatic protons (left) and the methyl/methylene groups (right) in the 0 ppm region for $^{\text{WT}}$ Gp2. The boxed region within the 0 ppm region is enlarged to show the ring current-shifted methyl groups from V20 for WTGp2 and mutants (for details, see the text). (b) A (^1H)NMR spectrum of an altered form of Gp2 ($^{\text{F16A}}$ Gp2)¹⁰ in which the structure is significantly disrupted by the mutation.

fluorophore at a single, surface-exposed, non-essential cysteine residue at position 60 (C60).¹⁰ The ^{32}P -labelled version of Alexa488-modified $^{\text{WT}}$ Gp2 ($^{\text{WT*}}$ Gp2) bound $\text{E}\sigma^{70}$ as well as ^{32}P -labelled non-modified Gp2 in a native gel assay (developed to monitor complex formation between [^{32}P]Gp2 and $\text{E}\sigma^{70}$ (Fig. 3a),¹⁰) indicating that the presence of the Alexa488-fluorophore at C60 does not detectably affect the binding of Gp2 to $\text{E}\sigma^{70}$ or the inhibition of $\text{E}\sigma^{70}$ by Gp2 (data not shown). Incubation of $^{\text{WT*}}$ Gp2 with increasing concentrations of wild type $\text{E}\sigma^{70}$ led to a large change in anisotropy of $^{\text{WT*}}$ Gp2 (Fig. 3b). Control reactions with the Alexa488-modified version of $^{\text{R56E}}$ Gp2 confirmed that the increase in anisotropy is due to specific binding of Gp2 to $\text{E}\sigma^{70}$ (Fig. 3b). As expected, we failed to detect an interaction between $^{\text{WT*}}$ Gp2 and $\text{E}\sigma^{70}$ reconstituted with a mutant form of RNAP core, which contains a charge reversal substitution (E1188K)¹² in the β' jaw domain and is unable to bind to Gp2 (Fig. 3c). The anisotropy of $^{\text{WT*}}$ Gp2 continued to increase with increasing amounts of $\text{E}\sigma^{70}$ in a dose-dependent manner and achieved saturation at about ~ 50 nM $\text{E}\sigma^{70}$, indicating that Gp2 binds to $\text{E}\sigma^{70}$ with a 1:1 binding stoichiometry and with

sub-nanomolar affinity (K_d 0.4 ± 0.02 nM). We conclude that the FP assay allows sensitive and quantitative detection of RNAP interaction with Gp2.

We used the FP assay to measure the strength of the interaction between Alexa488-fluorophore-labelled versions of Gp2 mutants harbouring substitutions in the NCS. As shown in Fig. 3d, the affinity of Alexa488-fluorophore-labelled versions of $^{\text{Mut2}}$ Gp2, $^{\text{Mut5}}$ Gp2, $^{\text{Mut6}}$ Gp2 and $^{\text{Mut7}}$ Gp2 for $\text{E}\sigma^{70}$ was reduced ~ 3 – 12 -fold compared to the $^{\text{WT*}}$ Gp2– $\text{E}\sigma^{70}$ interaction. Nevertheless, the K_d values determined for interaction between $^{\text{Mut2*}}$ Gp2, $^{\text{Mut5*}}$ Gp2, $^{\text{Mut6*}}$ Gp2 and $^{\text{Mut7*}}$ Gp2 and $\text{E}\sigma^{70}$ still reflect a very tight binding affinity and all Gp2 mutants bound to $\text{E}\sigma^{70}$ with a 1:1 molar ratio under our experimental conditions (Fig. 3d and data not shown). We conclude that mutations in the NCS have only a small effect on the affinity of Gp2 for host RNAP and stoichiometric binary complexes containing $\text{E}\sigma^{70}$ and Gp2 can be obtained readily with each mutant. The results are consistent with the earlier finding that the two surface-exposed invariant arginine residues (R56 and R58) in Gp2 are the major determinants for binding the host RNAP.¹⁰

Multiple amino acid substitutions in the NCS attenuate the ability of Gp2 to inhibit the *E. coli* RNAP

The ability of Gp2 mutants to inhibit $E\sigma^{70}$ was tested in an *in vitro* transcription assay. The $Mut2$ Gp2 mutant present in equimolar amounts with $E\sigma^{70}$ inhibits ApApUpU transcript synthesis as well as WT Gp2 (Fig. 3e, lane 5). In contrast, under identical conditions $Mut5$ Gp2 and $Mut6$ Gp2 display a significantly reduced (by ~53% and ~97%, respectively) ability to inhibit ApApUpU synthesis by $E\sigma^{70}$ (Fig. 3e, lanes 8 and 11, respectively). We observe an increased inhibition of transcription activity of $E\sigma^{70}$ *in vitro* at higher concentrations of $Mut5$ Gp2 and $Mut6$ Gp2 (Fig. 3e, lanes 8–13). However, $Mut7$ Gp2, even when present at ~4-fold molar excess over $E\sigma^{70}$, has only a marginal inhibitory effect on $E\sigma^{70}$ transcription. Thus, we propose that the NCS could be involved in inhibitory interactions that Gp2 makes *en route* to RPo formation but is not required for the initial binding of Gp2 to the RNAP (Fig. 3d). Overall, the results from the *in vitro* transcription assay are consistent with the results from the efficiency of plaque formation assays (Fig. 1d), suggesting that amino acid substitutions in NCS impair the ability of Gp2 to effectively inhibit the *E. coli* RNAP.

To determine whether $Mut7$ Gp2 is retained in the RPo that evidently forms in its presence, we conducted native gel mobility-shift assays using Alexa488-fluorophore-labelled Gp2 and the [32 P] *lacUV5* promoter probe. Gp2 containing promoter complexes can be visualised by the presence of a fluorescence signal whereas promoter DNA containing promoter complexes can be visualised by the presence of a radioactive signal on the native gel. Double labelling reports on the formation of promoter complexes containing Gp2. As shown in Fig. 3f, the WT Gp2- $E\sigma^{70}$ and $Mut7$ Gp2- $E\sigma^{70}$ binary complexes migrate detectably faster than the RPo on the native gel (compare lanes 2, 4 and 6'). The addition of the [32 P] *lacUV5* promoter probe to the WT Gp2- $E\sigma^{70}$ complex, as expected, did not lead to the RPo formation (Fig. 3f, lane 7') and only the fluorescence signal originating from the WT Gp2- $E\sigma^{70}$ complex was present in the corresponding lane (Fig. 3f, compare lanes 7 and 7'). Consistent with earlier result (Fig. 3e), the addition of [32 P] *lacUV5* promoter probe to the $Mut7$ Gp2- $E\sigma^{70}$ complex had only a marginal effect on RPo formation (Fig. 3f, lane 8'); no fluorescence signal was detectable in RPo formed in the presence of $Mut7$ Gp2, indicating that $Mut7$ Gp2 dissociates from $E\sigma^{70}$ upon formation of RPo (Fig. 3f, compare lanes 8 and 8'). In support of this view, the fluorescence signal corresponding to the $Mut7$ Gp2- $E\sigma^{70}$ complex in the lane containing the [32 P] *lacUV5* probe is diminished by >50% (Fig. 3f, compare lanes 4 and 8), thus reflecting formation of the RPo by the

$Mut7$ Gp2-bound $E\sigma^{70}$. It appears that in the absence of the NCS, Gp2 is efficiently displaced from the RNAP at some point *en route* to the RPo. In contrast, no such displacement of WT Gp2 occurs, leading to efficient inhibition of RPo formation by $E\sigma^{70}$. Clearly, the NCS seems to contribute to the ability of Gp2 to inhibit the RNAP efficiently, which is compromised in its absence.

Efficient inhibition of *E. coli* $E\sigma^{70}$ by Gp2 requires region 1 σ^{70} region 1

The results described above indicate that the NCS in Gp2 is important for efficient inhibition of $E\sigma^{70}$. In an earlier study, Mekler *et al.* showed that region 1.1 (residues 1–100), the negatively charged multifunctional amino-terminal domain found in most primary σ factors, such as *E. coli* σ^{70} , is located in the vicinity of the downstream-DBC when promoter DNA is not present.⁵ The σ^{70} region 1.1 has been suggested to serve as a “molecular placeholder” for downstream duplex DNA.⁵ During RPo formation, the σ^{70} region 1.1 is repositioned from the downstream-DBC, a step that is necessary for efficient and stable formation of RPo. On the basis of the expected proximity of the NCS in RNAP-bound Gp2 and the negatively charged σ^{70} region 1.1 in $E\sigma^{70}$ in the absence of promoter DNA (Fig. 4a), we considered the possibility that Gp2 inhibition of $E\sigma^{70}$ could involve the σ^{70} region 1.1. To this end, we conducted *in vitro* transcription assays with $E\sigma^{70}$ reconstituted with σ^{70} deleted for region 1 (residues 1–130; here called σ^{70}_{2-4}). Initially, we used the FP assay to establish that Gp2 binds to the $E\sigma^{70}$ reconstituted with σ^{70}_{2-4} with an affinity similar to that of the wild-type $E\sigma^{70}$ (Fig. 4b). However, the ability of Gp2 to inhibit $E\sigma^{70}_{2-4}$ is detectably impaired when compared to reactions containing $E\sigma^{70}$ (Fig. 4c). Overall, the results indicate that efficient inhibition of $E\sigma^{70}$ by Gp2 requires the σ^{70} region 1.

Perspectives: insights into the mechanisms of inhibition of $E\sigma^{70}$ by Gp2

In this study we demonstrated that the NCS has a major role in the ability of Gp2 to function as a potent inhibitor of *E. coli* $E\sigma^{70}$. The results show that the NCS plays a relatively small role in the initial binding of Gp2 to $E\sigma^{70}$ and that multiple substitutions in the NCS do not compromise the overall structural integrity of the protein. There appears to be a clear progressive decrease in efficiency of $E\sigma^{70}$ inhibition with the progressive decrease in the number of negatively charged residues in the NCS. In the restraint-driven structural model of the Gp2-bacterial RNAP

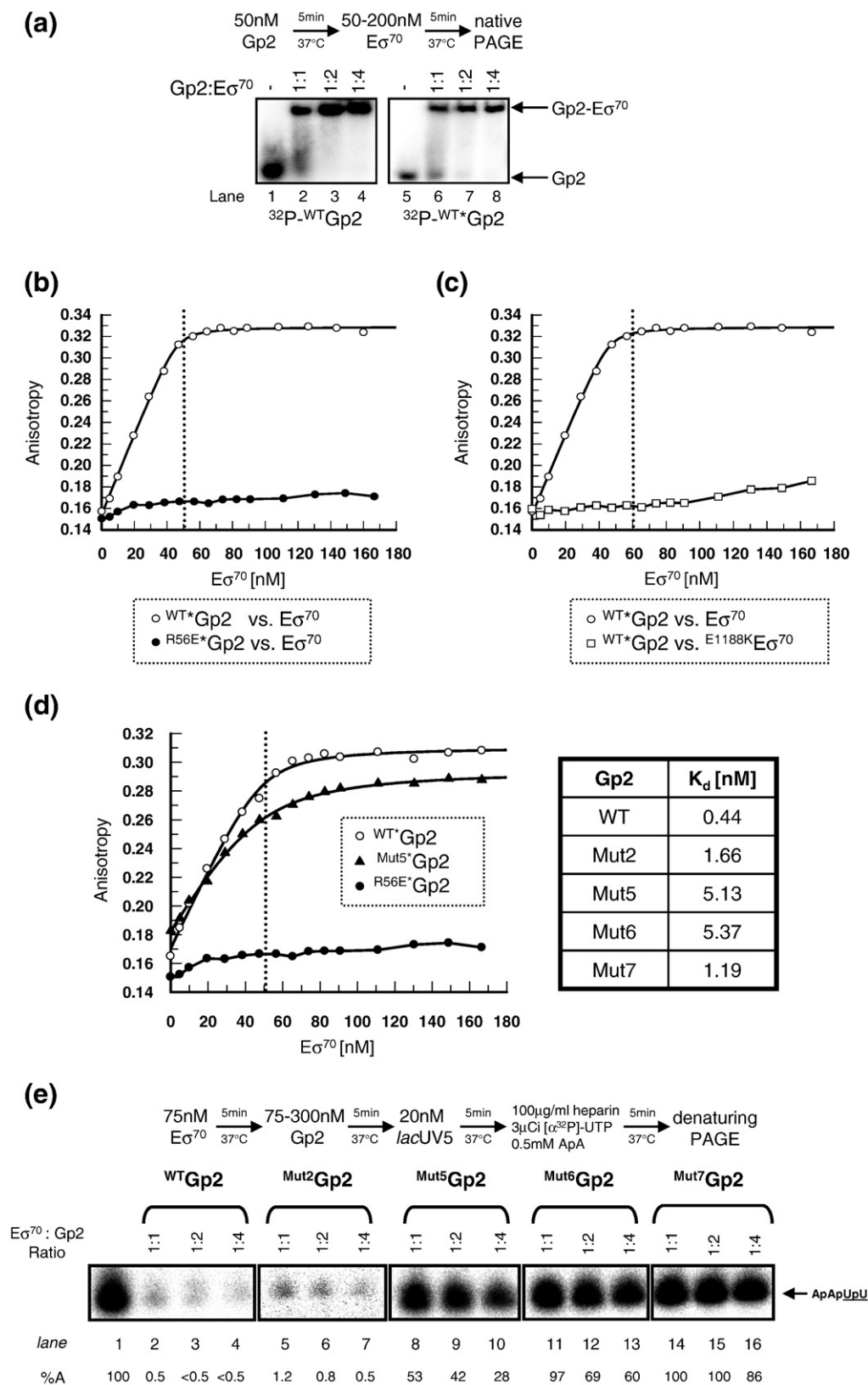


Fig. 3 (legend on next page)

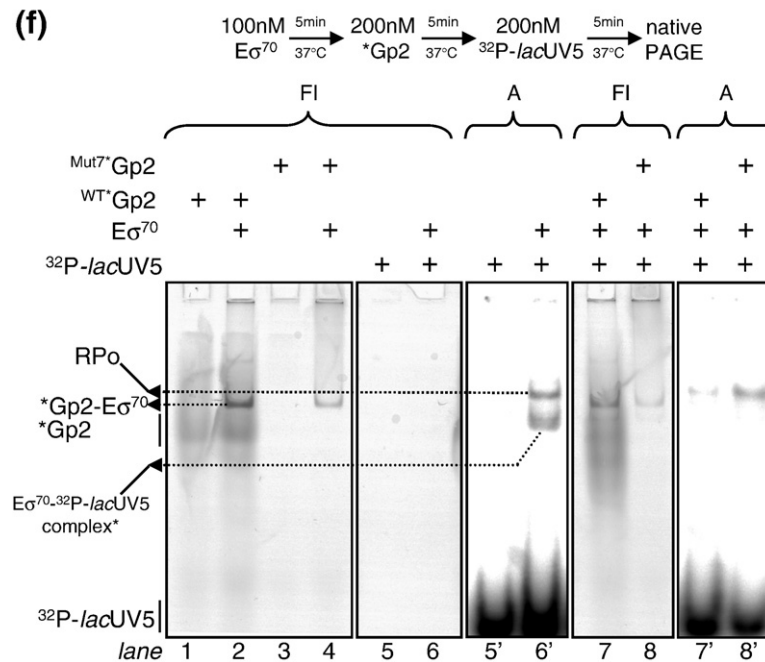


Fig. 3. The binding of Gp2 mutants to *E. coli* $E\sigma^{70}$. (a) An autoradiograph of a 4.5% (w/v) polyacrylamide native gel showing the binding of [32 P] WT Gp2 (left), [32 P] WT *Gp2 (right) to $E\sigma^{70}$. The migration positions of [32 P]Gp2/Gp2* (lanes 1 and 5) and the $E\sigma^{70}$ -Gp2 complex (lanes 2–4 and 6–8) are indicated. The molar ratio of $E\sigma^{70}$ with respect to Gp2 in each lane is shown at the top. The binding assay was done as described¹⁰ in RNAP-binding buffer (10 mM Tris-HCl, pH 8, 100 mM NaCl, 0.1 mM EDTA, 5% (v/v) glycerol) and the schematic at the top indicates the concentration of reaction components, time added and incubation time. (b) Equilibrium protein binding assays with Alexa488-fluorophore-labelled Gp2 and $E\sigma^{70}$. A Fluoromax-3 spectrofluorimeter fitted with automated polarization filters was used to measure fluorescence anisotropy. Data were recorded using an excitation wavelength of 495 nm and an emission wavelength of 519 nm. The binding assays were done at 37 °C in a reaction volume of 400 μ l in RNAP binding buffer. Small amounts of the $E\sigma^{70}$ (reconstituted *in situ* with ~4-fold molar excess of σ^{70} over E) were titrated into 50 nM Alexa488-fluorophore-labelled Gp2. Five measurements of anisotropy were averaged and each protein titration was done in at least duplicate. Data were fit using the tight binding equation in the Grafit 5 software (Erithacus Software) as described.¹³ The graph shows the increase in anisotropy of wild type WT *Gp2 (open circles) and R56E *Gp2 (filled circles) as a function of the concentration of $E\sigma^{70}$, where the concentration at saturation is indicated by the dotted line. (c) As in (b) but reactions were done with WT *Gp2 and either $E\sigma^{70}$ (open circles) or $E^{E1188K}\sigma^{70}$ (open boxes). (d) A graph showing the increase in anisotropy of wild type WT *Gp2 (open circles), R56E *Gp2 (filled circles) and Mut5 *Gp2 (filled triangles) as a function of the concentration of $E\sigma^{70}$. The K_d values calculated for the interaction of WT *Gp2 and *Gp2 mutants with $E\sigma^{70}$ are given in the table. (b)–(d) At least two independent experiments were done to obtain the K_d values; all data were within 5–10% of the K_d values shown. (e) An autoradiograph of a 20% (w/v) denaturing gel showing synthesis of the transcript ApApUpU (indicated by the arrow, where the underlined nucleotides are 32 P-labelled) from the *lacUV5* promoter by $E\sigma^{70}$ in the presence of WT Gp2 (lanes 2–4), Mut2 Gp2 (lanes 5–7), Mut5 Gp2 (lanes 8–10), Mut6 Gp2 (lanes 11–13) and Mut7 Gp2 (lanes 14–16). The percentage of ApApUpU synthesised (%A) by $E\sigma^{70}$ in the presence of Gp2 with respect to reactions in the absence of Gp2 are given at the bottom of the gels. In independent experiments, all data were within 15% of the %A value shown. The *in vitro* transcription assay was done as described¹⁰ and the schematic at the top indicates the concentration of reaction components, time added and incubation time. (g) Experiments to determine whether Mut7 Gp2 is retained in the RPo (for details, see the text). A fluorescence image (FI; lanes 1–8) and autoradiograph (A; lanes 5'–8') of a 4.5% (w/v) polyacrylamide native gel showing the migration positions of Gp2, Gp2- $E\sigma^{70}$, unbound [32 P]*lacUV5* promoter probe and the RPo. The reaction components present in each lane are shown at the top and the schematic indicates the concentration of reaction components, time added and incubation time. The faster migrating complex marked with an asterisk (*) in lane 6' is a heparin-sensitive $E\sigma^{70}$ -[32 P]*lacUV5* promoter complex that has not acquired the RPo status (data not shown).

complex,¹⁰ the NCS projects into the downstream-DBC (Fig. 4a) and is surrounded by structurally flexible domains of the *E. coli* RNAP (β' jaw, Gp2-binding site; β' I6, β' downstream lobe; β' , downstream clamp) that form the downstream DBC and play a role in RPo formation and

stability.^{4,7} Our results support the view that the NCS in Gp2 antagonises protein–protein and/or protein–DNA interactions that occur *en route* to the RPo. We propose three, not necessarily mutually exclusive, mechanisms by which Gp2 might inhibit RPo formation by $E\sigma^{70}$.

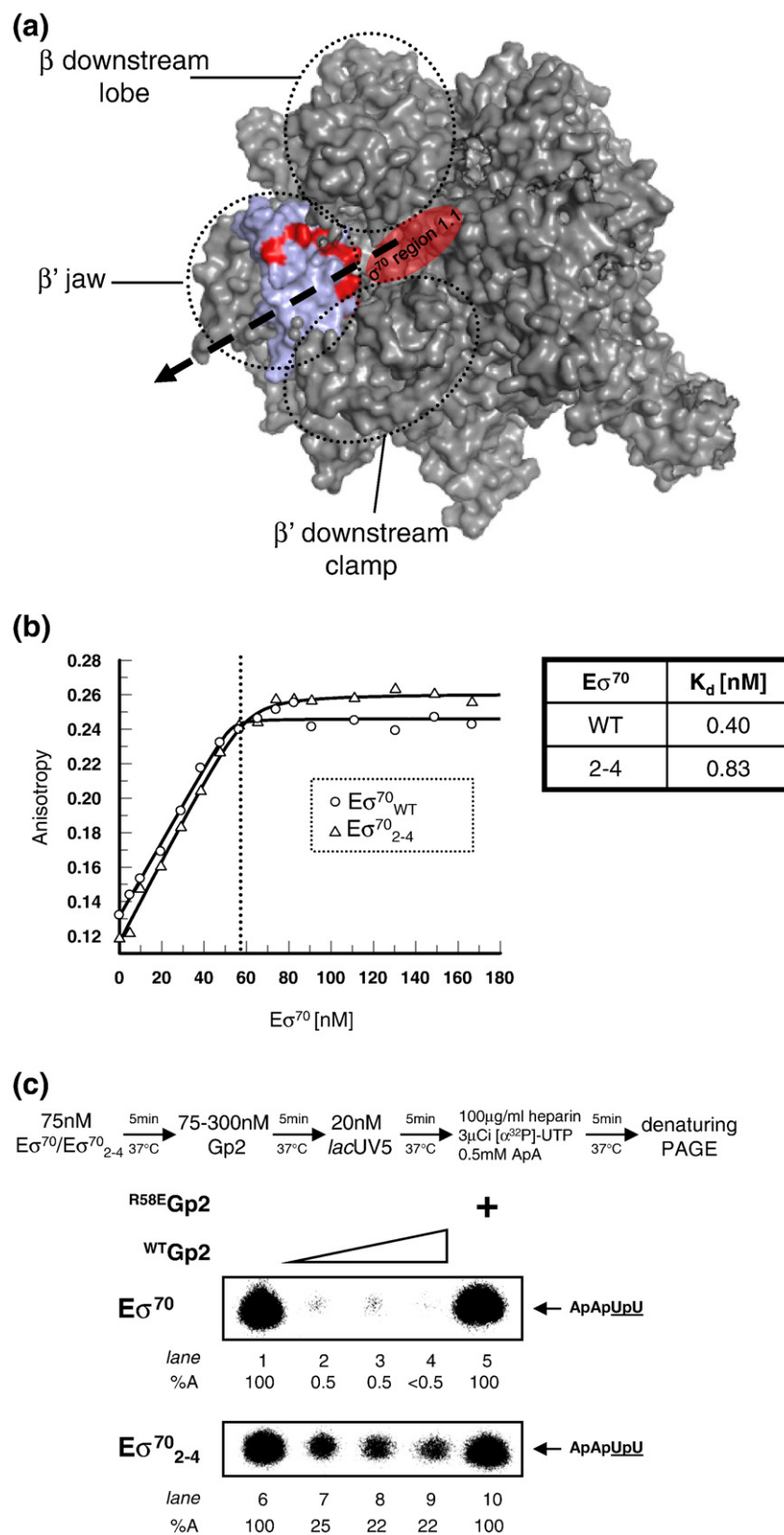


Fig. 4. (a) Restraint-driven model of Gp2 bound to the *Thermus aquaticus* $E\sigma^A$ (σ^A is the *T. aquaticus* equivalent of *E. coli* σ^{70}) shown as a surface representation. The subunits of *T. aquaticus* $E\sigma^A$ are coloured grey and Gp2 is shown in light blue with the NCS indicated in red. The path of the downstream duplex DNA in the downstream-DBC is indicated by the dotted arrow. Dotted circles indicate the structurally flexible domains of the bacterial RNAP that contribute to the downstream-DBC; *T. aquaticus* RNAP does not have the β' I6 domain. The approximate location of σ^{70} region 1.1 is indicated by the red oval. (b) A graph showing the increase in anisotropy of wild type $^{WT}Gp2$ as a function of $E\sigma^{70}$ (open circles) and $E\sigma^{70}_{2-4}$ concentration (open triangles). *E. coli* σ^{70}_{2-4} was constructed and purified exactly as described.¹⁴ $E\sigma^{70}$ and $E\sigma^{70}_{2-4}$ were reconstituted *in situ* using ~4-fold molar excess of E over σ^{70} and σ^{70}_{2-4} , respectively. The K_d values calculated for the interaction of $^{WT}Gp2$ with $E\sigma^{70}$ and $E\sigma^{70}_{2-4}$ are shown in the table. At least two independent experiments were done to obtain the K_d values. The data obtained were all within 5~10% of the K_d values shown. (c) An autoradiograph of a 20% (w/v) polyacrylamide denaturing gel showing synthesis of the transcript ApApUpU (indicated by the arrow, where the underlined nucleotides are ^{32}P -labelled) from the *lacUV5* promoter by $E\sigma^{70}$ and $E\sigma^{70}_{2-4}$ in the presence of increasing amounts (~1, 2 and ~4-fold molar excess over $E\sigma^{70}/E\sigma^{70}_{2-4}$) of $^{WT}Gp2$ (lanes 2–4 and 7–9, respectively). Reactions containing an ~4-fold molar excess of $R58EGp2$ over $E\sigma^{70}/E\sigma^{70}_{2-4}$ are shown as a control (lanes 5 and 10). The percentage of ApApUpU synthesised (%A) by $E\sigma^{70}$ and $E\sigma^{70}_{2-4}$ in the presence of Gp2 with respect to reactions in the absence of with Gp2 (lanes 1 and 6) are given at the bottom of the gels. All data obtained in independent experiments were within 15% of the %A value shown. The *in vitro* transcription assay was done as described¹⁰ and the schematic at the top indicates the concentration of reaction components, time added and incubation time.

Mechanism 1

Electrostatic interactions between the NCS and positively charged surfaces of the β' downstream lobe and the β' downstream clamp domains could "lock" RNAP in a conformation unfavourable for RPo formation.

Mechanism 2

The NCS could electrostatically antagonise interactions between the downstream-DBC and the duplex downstream DNA, which is an essential step for RPo formation and stability.

Mechanism 3

The NCS could electrostatically reposition region 1 of σ^{70} to a location unfavourable for efficient RPo formation to occur. Therefore, in the absence of region 1 of σ^{70} region 1, Gp2 is unable to inhibit $E\sigma^{70}$ efficiently. In support of this view, *E. coli* RNAP reconstituted with σ factors naturally lacking region 1.1 (e.g. $E\sigma^{38}$) is markedly less sensitive than $E\sigma^{70}$ to Gp2.⁹

The availability of Gp2 mutants, such as Mut^7 Gp2, that bind tightly but fail to inhibit $E\sigma^{70}$, opens ways for addressing experimentally the mechanisms by which Gp2 might inhibit $E\sigma^{70}$. Such endeavours are underway in our laboratories.

Acknowledgements

This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) grant (to S.W, E.C and S.M.) and a UK Biochemical Society travel grant (to S.W.). S.W. is a recipient of a David Phillips Fellowship (BB/E023703). Work in K.S. laboratories was supported by NIH grant GM64530, Russian Academy Presidium Molecular and Cell Biology program grant, and Federal Program "Scientific and scientific-pedagogical personnel of innovative Russia 2009-2013", state contract 02.740.11.0771. N.A. was partially supported by a Charles and Johanna Busch Postdoctoral Fellowship from the Waksman Institute, Russian Foundation for Basic Research grant 08-04-00968-a, and Russian Ministry of Education and Science State contract P1166.

References

1. 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.
2. Wigneshweraraj, S., Bose, D., Burrows, P. C., Joly, N., Schumacher, J., Rappas, M. *et al.* (2008). Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor. *Mol. Microbiol.* **68**, 538–546.
3. Davis, C. A., Bingman, C. A., Landick, R., Record, M. T., Jr & Saecker, R. M. (2007). Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase. *Proc. Natl Acad. Sci. USA*, **104**, 7833–7838.
4. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. & Darst, S. A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*, **296**, 1285–1290.
5. Mekler, V., Kortkhonja, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A. N. *et al.* (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell*, **108**, 599–614.
6. Opalka, N., Brown, J., Lane, W. J., Twist, K. A., Landick, R., Asturias, F. J. & Darst, S. A. (2010). Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. *PLoS Biol.* **8**, e1000483.
7. 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.
8. Nechaev, S. & Severinov, K. (2003). Bacteriophage-induced modifications of host RNA polymerase. *Annu. Rev. Microbiol.* **57**, 301–322.
9. Nechaev, S. & Severinov, K. (1999). Inhibition of *Escherichia coli* RNA polymerase by bacteriophage T7 gene 2 protein. *J. Mol. Biol.* **289**, 815–826.
10. Camara, B., Liu, M., Reynolds, J., Shadrin, A., Liu, B., Kwok, K. *et al.* (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. USA*, **107**, 2247–2252.
11. 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.
12. Wigneshweraraj, S. R., Burrows, P. C., Nechaev, S., Zenkin, N., Severinov, K. & Buck, M. (2004). Regulated communication between the upstream face of RNA polymerase and the beta' subunit jaw domain. *EMBO J.* **23**, 4264–4274.
13. Bellamy, S. R. & Baldwin, G. S. (2001). A kinetic analysis of substrate recognition by uracil-DNA glycosylase from herpes simplex virus type 1. *Nucleic Acids Res.* **29**, 3857–3863.
14. Wilson, C. & Dombroski, A. J. (1997). Region 1 of sigma70 is required for efficient isomerization and initiation of transcription by *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **267**, 60–74.



Molecular Mechanism of Transcription Inhibition by Phage T7 gp2 Protein

Vladimir Mekler^{1*}, Leonid Minakhin¹, Carol Sheppard², Sivaramesh Wigneshweraraj² and Konstantin Severinov^{1,3*}

¹Department of Molecular Biology and Biochemistry, Waksman Institute for Microbiology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

²Section of Microbiology, Faculty of Medicine and Centre for Molecular Microbiology and Infection, Imperial College London, SW7 2AZ, UK

³Institutes of Molecular Genetics and Gene Biology, Russian Academy of Sciences, Moscow 119991, Russia

Received 23 July 2011;
received in revised form
14 September 2011;
accepted 15 September 2011
Available online
21 September 2011

Edited by R. Ebricht

Keywords:

RNA polymerase;
transcription inhibition;
T7 bacteriophage;
open promoter complex;
protein beacon assay

Escherichia coli T7 bacteriophage gp2 protein is a potent inhibitor of host RNA polymerase (RNAP). gp2 inhibits formation of open promoter complex by binding to the β' jaw, an RNAP domain that interacts with downstream promoter DNA. Here, we used an engineered promoter with an optimized sequence to obtain and characterize a specific promoter complex containing RNAP and gp2. In this complex, localized melting of promoter DNA is initiated but does not propagate to include the point of the transcription start. As a result, the complex is transcriptionally inactive. Using a highly sensitive RNAP beacon assay, we performed quantitative real-time measurements of specific binding of the RNAP–gp2 complex to promoter DNA and various promoter fragments. In this way, the effect of gp2 on RNAP interaction with promoters was dissected. As expected, gp2 greatly decreased RNAP affinity to downstream promoter duplex. However, gp2 also inhibited RNAP binding to promoter fragments that lacked downstream promoter DNA that interacts with the β' jaw. The inhibition was caused by gp2-mediated decrease of the RNAP binding affinity to template and non-template strand segments of the transcription bubble downstream of the -10 promoter element. The inhibition of RNAP interactions with single-stranded segments of the transcription bubble by gp2 is a novel effect, which may occur via allosteric mechanism that is set in motion by the gp2 binding to the β' jaw.

© 2011 Elsevier Ltd. All rights reserved.

Introduction

RNA polymerase (RNAP) is among the most highly regulated enzymes in the bacterial cell.

*Corresponding authors. K. Severinov is to be contacted at Waksman Institute, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA. E-mail addresses: mekler@waksman.rutgers.edu; severik@waksman.rutgers.edu.

Abbreviations used: RNAP, RNA polymerase; TMR, tetramethylrhodamine.

Bacteriophages have developed ingenious mechanisms to modify host RNAP to make it serve viral needs. Interactions between phage-encoded factors and host RNAPs have long served as a rich source of paradigms for transcriptional regulation (reviewed in Refs. 1 and 2). One such example is provided by *Escherichia coli* bacteriophage T7. The transcription of T7 phage genome proceeds in three stages. Early T7 genes are transcribed from several very strong promoters by *E. coli* RNAP containing primary σ^{70} subunit ($E\sigma^{70}$), whereas middle and late T7 genes are transcribed by phage-encoded RNAP.³ The product of the middle T7 gene 2, a 7-kDa protein called gp2,

binds to the β' jaw domain of the host RNAP and inhibits it.^{4,5} Only one gp2 molecule participates in tight interaction with an RNAP molecule,⁶ and point mutations in, as well as, deletions of the β' jaw abolish this binding.⁵ gp2 was originally thought to simply switch phage gene transcription from host to viral RNAP. However, its essential biological role is to prevent transcription by host RNAP into those parts of T7 genome that are normally transcribed by viral RNAP. The slow-moving *E. coli* RNAP slows down the fast-moving T7 RNAP, causing the latter enzyme to pause, which initiates aberrant processing of T7 replicative intermediates, leading to accumulation of improperly packaged, defective phage particles.^{7,8}

The RNAP site of gp2 binding, the β' jaw, is part of a trough that accepts double-stranded DNA downstream of the transcription start point.⁹ An RNAP mutant lacking the β' jaw (RNAP- $\beta'\Delta$ jaw) is active but forms unstable transcription initiation complexes.⁹ Binding of gp2 to the β' jaw prevents formation of detectable promoter complexes by $E\sigma^{70}$.^{5,10} gp2 has no effect on (and does not bind to) preformed open promoter complexes (RP_o) or transcription elongation complexes.⁵ Thus, the β' jaw interactions with downstream DNA and gp2 appear to be mutually exclusive, suggesting that the mechanism of inhibition is based on steric blockade of the β' jaw interactions with DNA by bound gp2.

It is commonly accepted that RP_o formation is initiated when $E\sigma^{70}$ contacts with DNA upstream of the transcription start point are established. Since gp2 binds the β' jaw, it is expected to inhibit a late stage of RP_o formation. Thus, gp2 is predicted to trap an intermediate complex *en route* to RP_o.¹⁰ However, promoter complexes containing gp2, if they exist, are unstable, and standard methods such as footprinting failed to reveal the interaction of the preformed complex of $E\sigma^{70}$ and gp2 ($E\sigma^{70}$ -gp2) with promoter.¹⁰ Here, we used an engineered promoter with optimized sequence to obtain and partially characterize a specific transcriptionally inactive promoter complex containing $E\sigma^{70}$ and gp2. DNase I footprinting and KMnO₄ probing of this complex reveal that gp2 prevents establishment of downstream DNA contacts and interferes with DNA melting near the transcription start site but has no effect on RNAP interactions with DNA at and upstream the -10 promoter element. Using a recently developed fluorescence RNAP beacon assay,¹¹ we performed quantitative real-time measurements of specific binding of the $E\sigma^{70}$ -gp2 complex to promoter DNA and various promoter fragments. In this way, the effect of gp2 on RNAP interaction with promoters was dissected. The data show that gp2 not only greatly decreases RNAP affinity to downstream promoter duplex but also reduces RNAP affinity to non-template strand segment of the transcription bubble located between

the -10 element and the transcription start site (called "discriminator" on some promoters). Further, gp2 appears to impede the entry of the template strand into the active center cavity. The inhibition of RNAP-discriminator interactions by gp2 is a novel effect, which may be either direct or indirect, that is, due to allosteric changes caused by gp2 binding to the β' jaw. We propose that promoter complexes trapped by gp2 correspond to an intermediate of promoter complex formation. Availability of such complexes, which are very transient in the absence of gp2, opens way for their functional and structural characterization.

Results

Experimental setup: Use of the RNAP beacon assay to monitor the inhibition of *E. coli* $E\sigma^{70}$ interactions with promoter DNA by T7 gp2

T7 gp2 binding to the jaw domain of the *E. coli* RNAP β' subunit decreases $E\sigma^{70}$ affinity to promoter DNA and prevents formation of catalytically competent RP_o.^{5,10} $E\sigma^{70}$ binding to promoter DNA and various promoter fragments can be measured using fluorescence protein beacon assay.¹¹ The beacon assay relies on the detection of fluorescence signal from RNAP holoenzyme containing the σ^{70} subunit with fluorescent label site-specifically incorporated in proximity to region 2.3, the part of σ that recognizes the -10 promoter element. The ground level fluorescence of such RNAP beacon is low because aromatic σ amino acids quench the fluorescence of the label. When RNAP beacon binds promoter DNA or a promoter fragment, specific interactions between σ region 2.3 and DNA are established. Since aromatic amino acids that quench the label participate in interactions with promoter DNA, the fluorescent probe becomes "unquenched", leading to increased fluorescence.

As can be seen from Fig. 2a, the fluorescence signal generated upon the addition of T5 N25 promoter-containing DNA fragment (shown in Fig. 1) to the RNAP beacon preincubated with excess gp2 is considerably lower than that observed in the absence of gp2. The gp2 effect was insignificant when gp2 was added to preformed complex of RNAP beacon and T5 N25 promoter DNA (Fig. 2a). The results are consistent with standard abortive transcript synthesis assay (Fig. 2b): transcription from T5 N25 was strongly inhibited when gp2 was added to RNAP before the addition of promoter DNA (lane 2), but no effect was observed when gp2 was added to preformed open complexes (lane 3). Thus, the beacon assay allows one to follow the inhibitory action of T7 gp2 on RNAP interactions with promoter.

Double-stranded DNA probes			
-65	-30	-10	+35
TCCTTGCTCAAGAATCATAAAAATTATTTGCTTTCAGGAAAATTTTCTGTATAATAGATTCA----			T5 N25
AGAAACGAGTTTCTTAGTATTTTTTAAATAAACGAAAGTCCTTTTAAAAAGACATATTATCTAAGT-----			
TCCTTGCTCAAAAATTTTTTTTAAAGTATTTGACATCAGGAAAATTTTGGTATAATAGATTCA----			N25cons
AGAAACGCGATTTTAAAAAAATTTTCATAAACTGTAGTCCTTTTAAAAAACCATATTATCTAAGT-----			
TCCTTGCTCAAGAATCATAAAAATTATCCACTTTCAGGAAAATTTTCTGCGCGGCAGATTCA----			Non-pro-
AGAAACGAGTTTCTTAGTATTTTTTAAATAGGTGAAAGTCCTTTTAAAAAGACGCGCGTCTAAGT-----			moter DNA
-59	-12		
CGCTAAAATTTTTTTTAAAGTATTTGACATCAGGAAAATTTTCTGT			[-59/-12]
GCGATTTTAAAAAAATTTTCATAAACTGTAGTCCTTTTAAAAAGACA			
-50		* * *+14	
TCATAAAAAATTATTTGACATCAGGAAAATTTTGGTATAATAGATTCAAAATTTGAGAGA			[-50/+14]
AGTATTTTAAATAAACTGTAGTCCTTTTAAAAAACCATATTATCTAAGTATTTAACTCTCT			
Downstream fork junctions			
-12	+20		
TATAATAGATTCAAAATTTGAGAGAGGAGTT	[-12/+20] [+2/+20] (+2)		
ATTTAACTCTCTCCTCAA			
TATAATAGATTCAAAATTTGAGAGAGGAGTT	[-12/+20] [+3/+20] (+3)		
TTTAACTCTCTCCTCAA			
TATAATAGATTCAAAATTTGAGAGAGGAGTT	[-12/+20] [+4/+20] (+4)		
TTAACTCTCTCCTCAA			
TATAATAGATTCAAAATTTGAGAGAGGAGTT	[-12/+20] [+5/+20] (+5)		
TAACTCTCTCCTCAA			
TATATTAGATTCAAAATTTGAGAGAGGAGTT	NC (+2/+3)		
TTAACTCTCTCCTCAA			
GC			
TATATTAGATTCAAAATTTGAGAGAGGAGTT	NC (-1/+3)		
TTAACTCTCTCCTCAA			
TCCGC			
TATATTAGATTCAAAATTTGAGAGAGGAGTT	NC (-4/+3)		
TTAACTCTCTCCTCAA			
CGGTCCGC			
Upstream fork junctions		Single-stranded probes	
-26	+2	-12	+2
ATGCTTCCGGCTCGTATAATGTGTGGA	[-26/+2] [-26/-12]	TATATTAGATTCAT	[-12/+2]
TACGAAGGCCGAGCA			
-26	-3	TATATTAGAT	[-12/-3]
ATGCTTCCGGCTCGTATAATGTGT	[-26/-3] [-26/-12]		
TACGAAGGCCGAGCA			
-26	-7	TATATT	[-12/-6]
GGAAAATTTTTGGTATAAT	[-26/-7] [-26/-12]		
CCTTTTAAAAAACCA			
-38	-11		
TATTTGCTTTCAGGAAAATTTTCTGT	[-38/-11] [-38/-12]		
ATAAACGAAAGTCCTTTTAAAAAGACA			

Hairpin downstream fork junctions

[-12/+9] [+4/+9] SP

-12 +4 +9
TATAATAGATTCATAAATTTG
TTAAAC

[-12/+9] [+4/+9; NC-4/+3] SP

-12 +4 +9
TATAATAGATTCATAAATTTG
TTAAAC
CGGTCCGC

Fig. 1. Structures of DNA probes used. The probe names used in the text are in red. In fork junction probe names, numbers in left and right parentheses correspond to borders of upper and bottom strands of fork junctions, with respect to the transcription start position located at +1 (underlined). "SP" stands for the SP18 linker that joins non-template strand oligos with template strand oligos in short hairpin fork junctions (shown as a vertical line in corresponding probe structures). Asterisks above the [-50/+14] probe indicate positions where the fragments used in experiment shown in Fig. 8 were truncated.

A stable ternary complex containing $E\sigma^{70}$, gp2, and promoter DNA can form when RNAP interactions with promoter are optimized

gp2 also decreased the binding of RNAP beacon to control non-promoter DNA fragment (Fig. 1) that contained multiple substitutions in the -35 and -10 regions and lacked transcription activity *in vitro*. Curiously, the signal observed upon the addition of T5 N25 DNA to $E\sigma^{70}$ -gp2 was higher than that observed upon the addition of non-promoter DNA (Fig. 2a). We interpreted this result by hypothesizing that some specific interactions between $E\sigma^{70}$ -gp2 complex and promoter are established. However, we could not detect specific binding of T5 N25 DNA to RNAP in the presence of gp2 using $KMnO_4$ probing or DNase I footprinting techniques (data not shown). We hypothesized that this negative result may reflect insufficient stability of ternary (RNAP-gp2-T5 N25) complexes. To

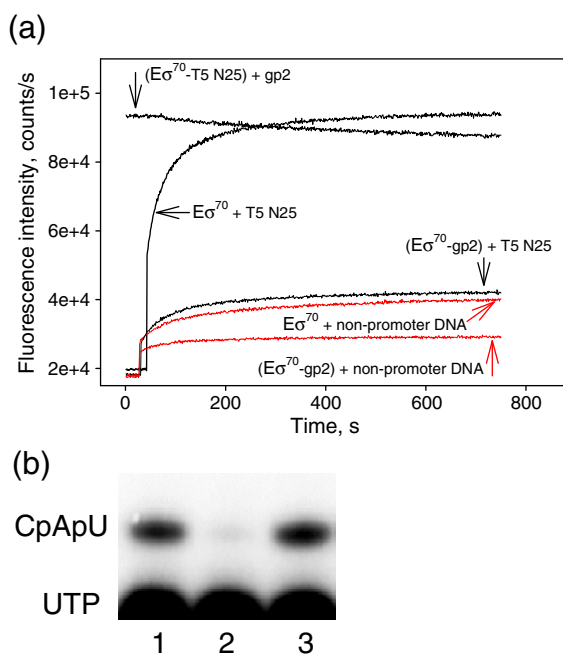


Fig. 2. T7 gp2 inhibits RNAP interaction with the T5 N25 promoter. (a) Time dependence of the change in fluorescence upon mixing 1 nM [211Cys-TMR] σ^{70} RNAP holoenzyme with 2 nM $-65/+35$ T5 N25 (black) or a non-promoter DNA fragment (red) in the presence or in the absence of gp2. The upper curve [labeled as "(RNAP + T5 N25) + gp2"] was obtained in an experiment where gp2 was added to the RNAP beacon preincubated with 2 nM T5 N25 promoter fragment for 30 min. In other experiments, gp2 was added to preformed RNAP-T5 N25 or RNAP-non-promoter DNA complexes. (b) Inhibition of abortive transcript synthesis from T5 N25 promoter fragment by gp2. Reactions contained $E\sigma^{70}$ alone (lane 1) or $E\sigma^{70}$ with gp2 added before (lane 2) or after (lane 3) open complex formation. An autoradiograph of denaturing polyacrylamide gel is presented.

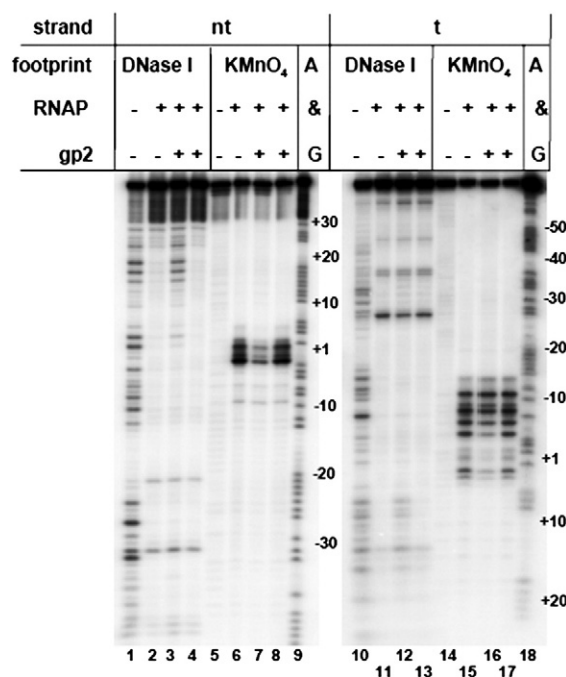


Fig. 3. $E\sigma^{70}$ forms complexes on the N25cons promoter in the presence of gp2. Reactions containing indicated proteins and N25cons promoter DNA fragment terminally labeled at the template ("t") or non-template ("nt") strands were footprinted with DNase I or probed with $KMnO_4$ as described in Materials and Methods. gp2 was added prior to (lanes 3, 7, 12, and 16) or after (lanes 4, 8, 13, and 17) open complex formation. Reaction products were resolved by denaturing PAGE and revealed using PhosphorImager.

improve the putative ternary complex stability, we prepared a derivative of -65 to $+35$ T5 N25 DNA fragment containing the consensus -35 element, an optimized UP element,¹² and the extended -10 element (Fig. 1). We expected that RNAP binding to this promoter, which we call N25cons, would be strengthened due to improved interactions upstream of the -10 promoter element. As expected, the DNA fragment containing N25cons generated significant fluorescence signal when combined with RNAP beacon in the presence of gp2 (Fig. S1a), indicating that specific complex between $E\sigma^{70}$ -gp2 and N25cons forms on this promoter. In the absence of gp2, $E\sigma^{70}$ was highly active on N25cons in the abortive transcription initiation assay (Fig. S1b, lane 1), yet despite the considerable fluorescence signal observed in the beacon assay, transcription from N25cons was highly sensitive to the addition of gp2 prior to promoter complex formation (Fig. S1b, lane 2). gp2 had no effect on preformed open complexes on N25cons (Fig. S1b, lane 3).

$E\sigma^{70}$ -gp2 formed complexes with the N25cons promoter that were readily detectable and clearly distinct from complexes formed by $E\sigma^{70}$ alone or

when gp2 was added after promoter complex formation (Fig. 3). First, DNase I footprinting demonstrated that, while $E\sigma^{70}$ strongly protected DNA up to +18 (with weaker protection extending approximately up to position +25), no protection downstream of position +8 was observed in complexes formed by $E\sigma^{70}$ -gp2 on either DNA strand (Fig. 3, compare lanes 2 and 3 and lanes 11 and 12). In contrast, the upstream portion of the footprint was identical in all complexes. Second, ExoIII treatment of promoter complexes revealed that the downstream boundary of $E\sigma^{70}$ promoter complex, located at position +18 in the absence of

gp2, was mostly shifted to around position +3 in the case of $E\sigma^{70}$ -gp2 complex (Fig. S2). The upstream boundary, located approximately at position -42, was slightly shifted downstream in the presence of gp2, whether added before or after DNA (Fig. S2, compare lanes 6, 7, and 8). While this observation suggests that gp2 may modify RNAP interactions upstream of the -35 promoter element, this effect is weak, as experiment shown in Fig. 7 (see below) demonstrates. $KMnO_4$ probing revealed that, in the absence of gp2, thymine bases at positions from -11 to +3 on the template strand are highly accessible to permanganate, with thymine at position +4 and cytidine at position -14 exhibiting weak sensitivity to the modifying reagent (Fig. 3). In complexes formed by $E\sigma^{70}$ -gp2, the reactivity of bases located from the upstream edge of the transcription bubble up to -6 was altered insignificantly, while intensity of $KMnO_4$ -sensitive bands further downstream was decreased (Fig. 3, compare lanes 15 and 16). On the non-template strand, strong reactivity at and around the transcription start point was observed in the absence of gp2 (Fig. 3, lane 6). This reactivity was strongly decreased in $E\sigma^{70}$ -gp2 complexes (Fig. 3, lane 7). We explain the residual $KMnO_4$ reactivity at and around the transcription start site seen in reactions containing $E\sigma^{70}$ -gp2 complexes by the presence of a small fraction of $E\sigma^{70}$ -promoter complexes in the reactions. The presence of such complexes is also evident from ExoIII footprinting (a minor fraction of exonuclease stops at position +18; Fig. S2, lane 3) and residual abortive initiation activity (Fig. S1b, lane 5). We therefore suspect that, in $E\sigma^{70}$ -gp2 complexes *per se*, the transcription start site is closed, explaining why these complexes are transcriptionally inactive. The $KMnO_4$ reactivity decreased insignificantly upon a 2-min challenge of the $E\sigma^{70}$ -gp2-N25cons complex with 50 μ g/ml of DNA competitor heparin (data not shown).

Overall, the results indicate that when interactions with the upstream promoter part are strong, RNAP forms a specific, stable, yet transcriptionally inactive promoter complex in the presence of gp2. The complex is only partially open, and formation of contacts in the downstream portion of the transcription bubble, as well as contacts with double-stranded DNA downstream of the transcription initiation start point, appears to be compromised. In the ternary complex, the DNA trajectory downstream from the -10 element must differ considerably from that in RP_o . A similar complex may form on the wild-type T5 N25 promoter; however, it is too transient to be detected by footprinting.

Quantitative dissection of the gp2 inhibitory effect using model promoter fragments

$E\sigma^{70}$ specifically binds upstream and downstream fork junctions—partially double-stranded DNA

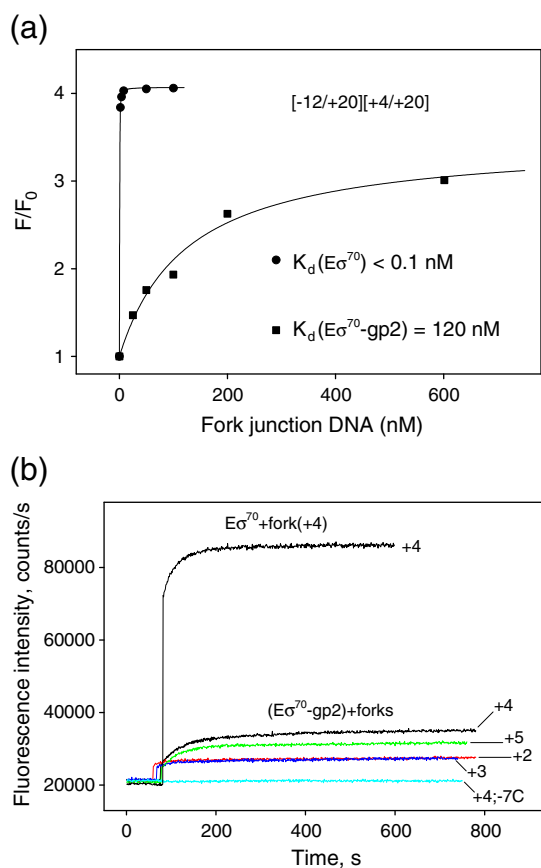


Fig. 4. Binding of downstream fork junction probes to the $E\sigma^{70}$ -gp2 complex. (a) Titration of $E\sigma^{70}$ and $E\sigma^{70}$ -gp2 with [-12/+20][+4/+20] downstream fork junction. The K_d values were determined by fitting the dependence of relative fluorescence signal amplitude (F/F_0) on probe concentration to a chemical equilibrium equation.¹¹ (b) The effect of fork junction point position on $E\sigma^{70}$ -gp2 interaction with downstream fork junction probes. Time dependence of the increase in fluorescence upon mixing the (RNAP beacon-gp2) complex with 50 nM [-12/+20][+2/+20], [-12/+20][+3/+20], [-12/+20][+4/+20], [-12/+20][+5/+20], and [-12/+20;-7C][+4/+20] probes abbreviated as +2, +3, +4, +5, and +4;-7C, correspondingly. The upper curve corresponds to RNAP beacon binding to 50 nM [-12/+20][+4/+20] probe in the absence of gp2.

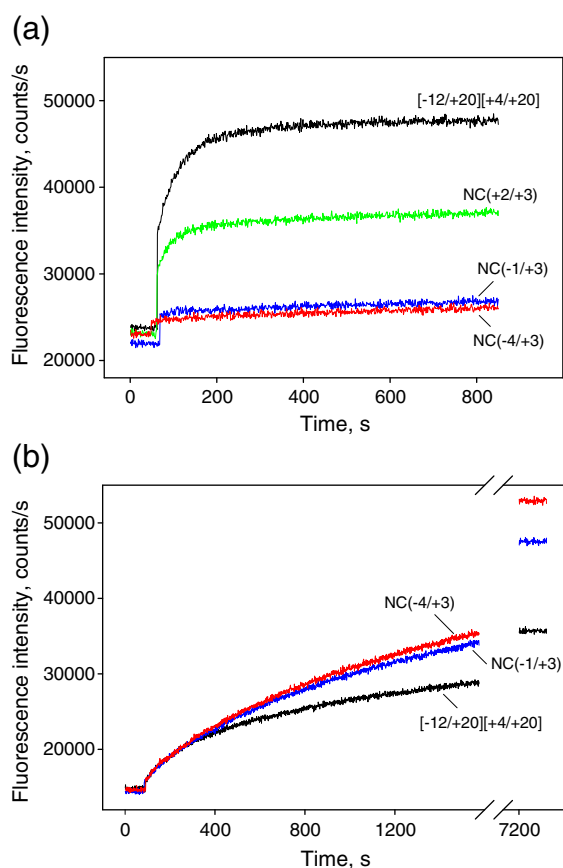


Fig. 5. Binding of downstream fork junction probes with template strand overhangs to the Eσ⁷⁰-gp2 complex and Δβ'(1149-1190)Eσ⁷⁰. (a) Time dependence of the fluorescence intensity in samples containing 1 nM [211Cys-TMR] σ⁷⁰ holoenzyme, 200 nM gp2, and 50 nM indicated downstream fork junction probes. (b) Time dependence of the fluorescence intensity upon the addition of 3 nM indicated downstream fork junction probes to 1 nM [211Cys-TMR] σ⁷⁰ holoenzyme bearing the Δβ'(1149-1190) deletion.

containing short non-template strand overhangs with the -10 promoter element sequence—and this interaction can be monitored by the beacon method.^{11,13,14} By “upstream” and “downstream”, we here mean the location of double-stranded segments relative to the non-template strand overhangs. gp2 inhibits Eσ⁷⁰ binding to downstream fork junction probes, whereas RNAP binding to upstream fork junction probe [-38/-11][-38/-12] (shown in Fig. 1) containing a short single-stranded extension corresponding to the -11A base is not affected.¹⁴ These results indicate that downstream fork junctions contain promoter segments whose recognition by RNAP is targeted by gp2, and therefore, these probes can be used to study the gp2 inhibition mechanism in more detail. We decided to quantitatively measure Eσ⁷⁰ binding to

various downstream fork junctions and some other probes (shown in Fig. 1) in the presence of gp2. To reach binding saturation, much higher concentrations of downstream fork junction probes were required in the presence of gp2 than in its absence (Fig. 4a; see also Ref. 14). However, the binding observed in the presence of gp2 was specific, since control probes bearing a C-to-T substitution at the highly conserved -7 position did not generate any noticeable signal (see, e.g., Fig. 4b). The equilibrium dissociation constant (K_d) for Eσ⁷⁰-gp2 binding to [-12/+20][+4/+20] downstream fork junction probe (with junction point positioned at +4) was calculated by fitting the experimental dependence of the fluorescence signal amplitude on probe concentration to a chemical equilibrium equation (e.g., by titration assay¹¹) (Fig. 4a). The resulting value was 120 nM. Free Eσ⁷⁰ bound the same probe with a K_d of less than 0.1 nM (Fig. 4a). Thus, gp2 decreases RNAP affinity to the downstream fork junction promoter fragments by more than 1000-fold.

As shown in Fig. 4b, probes with junction points at +2 and +3 bound the Eσ⁷⁰-gp2 less efficiently than the [-12/+20][+4/+20] probe. The apparent dependence of downstream fork junction affinities to Eσ⁷⁰-gp2 on the junction position is somewhat surprising, since negligible difference in affinity of these probes to Eσ⁷⁰ was detected.¹⁴

The effect of gp2 on RNAP interaction with a segment of template strand adjacent to the transcription start point was estimated by comparing the binding of [-12/+20][+4/+20] or [-12/+9][+4/+9] SP, a short hairpin downstream fork junction, and their derivatives bearing 2- to 7-nt-long noncomplementary extensions of template strand. Fluorescence signal changes observed upon Eσ⁷⁰-gp2 interaction with these probes are shown in Fig. 5a and Fig. S4. The kinetic curves in Fig. 5a indicate that much higher fluorescence levels are observed when Eσ⁷⁰-gp2 is mixed with parent [-12/+20][+4/+20] probe than with derivative probes containing template strand overhangs. We determined the K_d value for RNAP-gp2 binding to one such probe, NC(-3/+4). Low signal level made it impossible to determine the K_d value for this probe by titration assay. Therefore, an estimate of RNAP-gp2 binding to NC(-3/+4) ($K_d > 0.6$ μM) was obtained from a competition binding assay¹¹ by measuring RNAP-gp2 binding to -12/+3 oligo with and without 0.2 μM NC(-3/+4) (Fig. S3). In contrast, the mutant β'Δjaw RNAP has higher affinity to downstream fork junction probes bearing template strand extensions than to parent probes (Fig. 5b and Table 1). Thus, the inhibitory effect of single-stranded template strand extension on binding of the [-12/+20][+4/+20] derivatives is specifically caused by the gp2 binding to RNAP.

Upon incubation with Eσ⁷⁰-gp2, [-12/+9][+4/+9]SP-based downstream fork junction with

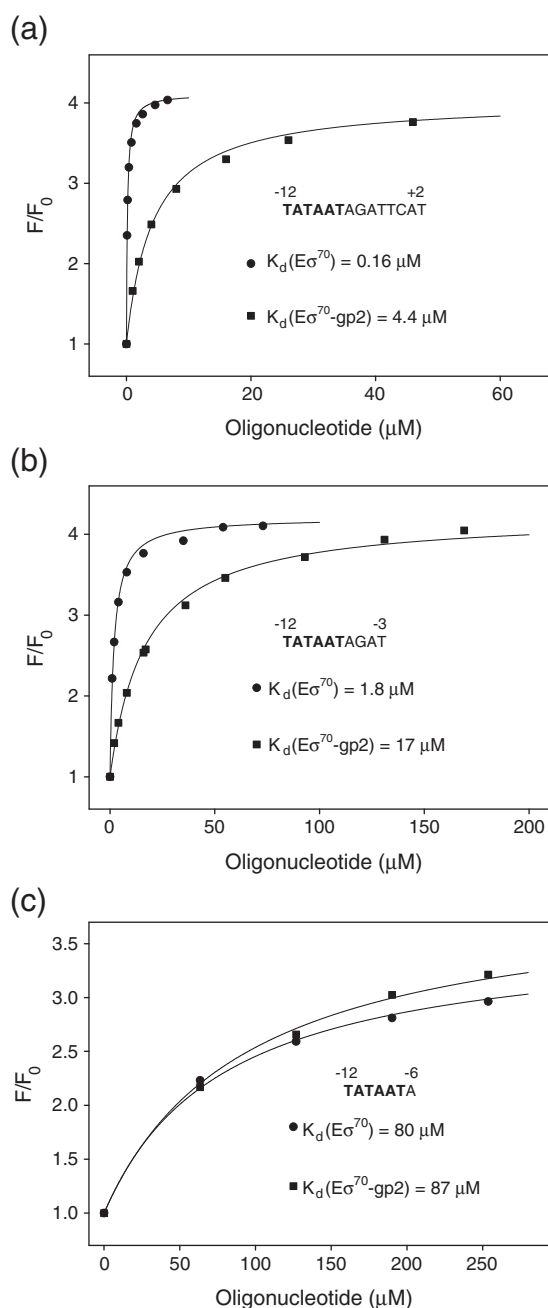


Fig. 6. Binding of non-template oligos containing the -10 promoter element consensus sequence to free $E\sigma^{70}$ and the $E\sigma^{70}$ -gp2 complex. Titration of [211Cys-TMR] σ^{70} holoenzyme and the gp2-[211Cys-TMR] σ^{70} holoenzyme complex with oligonucleotide probes. Continuous lines correspond to nonlinear regression fit of the data.

template strand extension generated saturated signal similar to that observed with parental $[-12/+9][+4/+9]$ SP; however, the kinetics of signal increase was significantly slower with the probe bearing the extension (Fig. S4). These data are in

Table 1. Dissociation constants for the binding of promoter fragments to RNAP beacon, RNAP beacon complex with gp2, and $\Delta\beta'(1149-1190)$ RNAP beacon

DNA probe	K_d RNAP (μ M)	K_d (RNAP-gp2) (μ M)	K_d (RNAP- β' Δ jaw) (μ M)
$[-12/+20][+4/+20]$	$<10^{-4}$	0.12	3.4×10^{-3}
NC($-4/+3$)	$<10^{-4}$	>0.6	1.0×10^{-3}
$-12/+2$	0.16	4.4	0.46
$-12/-3$	1.8	17	
$-12/-6$	80	87	

The structures of DNA probes are shown in Fig. 1. The K_d values presented are averages from two to three individual experiments; the error is $\pm 15\%$.

marked contrast with the situation observed in the absence of gp2, when an overhang from $+3$ to -4 actually improves the affinity of short fork junctions to $E\sigma^{70}$ by ~ 40 -fold (Fig. S5). These results indicate that the interactions of template strand overhangs with RNAP are affected by gp2 in the $E\sigma^{70}$ -gp2 complexes with downstream fork junctions.

If gp2 inhibited transcription solely by targeting interactions with the double-stranded DNA downstream of the transcription start point, then no effects on the interaction with single-stranded DNA in the transcription bubble should be expected. To determine if gp2 effect of transcription bubble formation is direct or indirect, we studied the effect of gp2 on the interaction of RNAP with oligonucleotides containing sequences matching the -10 promoter element consensus sequence and corresponding to single-stranded non-template extensions of downstream fork junction probes. $E\sigma^{70}$ specifically binds such oligonucleotide probes,¹⁵ and the interaction can be measured by the RNAP beacon assay.¹¹ We found that gp2 inhibits binding of $-12/+2$ oligo 27-fold (Fig. 6a and Table 1). The effect was specific since (i) only very small decrease in RNAP binding to $-12/+2$ was observed in the presence of R58E gp2 mutant having impaired affinity to RNAP,¹⁰ and (ii) no inhibition of $-12/+2$ oligo binding was observed upon gp2 addition to RNAP beacon containing RNAP- β' Δ jaw¹⁴ (data not shown).

$E\sigma^{70}$ binding to $-12/-3$ oligo was inhibited by gp2 ~ 9 -fold; the binding of $-12/-6$ oligo was not affected (Fig. 6b and c). The results suggest that gp2 hinders $E\sigma^{70}$ interactions with the $-5/+2$ segment of the non-template strand. Consistent with this conclusion, gp2 also decreased binding of upstream fork junction probes bearing longer ($-11/+2$ and $-11/-3$) non-template strand overhangs, whereas the binding of an upstream fork bearing a short $-11/-7$ overhang was not affected (Fig. S6).

Comparison of K_d values of $E\sigma^{70}$ -gp2 binding to $-12/+2$ oligo and $[-12/+20][+4/+20]$ fork junction shows that the $+3/+20$ segment increases affinity 37-fold. In contrast, the $+3/+20$ segment increased $E\sigma^{70}$

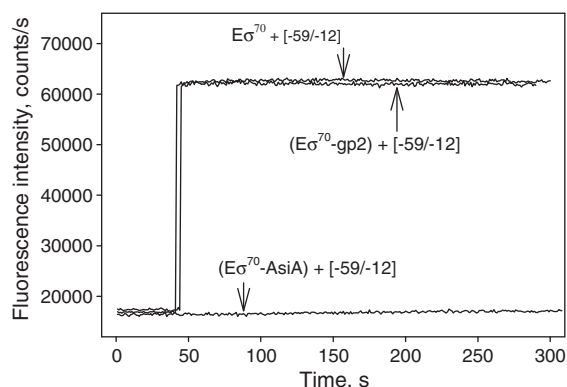


Fig. 7. The effect of T7 gp2 and T4 AsiA on $E\sigma^{70}$ binding to $[-59/-12]$ probe. $[-59/-12]$ probe (2 nM) was combined with RNAP beacon (1 nM) preincubated for 15 min with or without 0.2 μ M T7 gp2 or T4 AsiA, and the increase in fluorescence was monitored.

affinity 9000-fold.¹⁴ Thus, gp2 greatly decreased RNAP affinity to downstream promoter duplex. We also determined K_d values for binding of RNAP beacon containing RNAP- β' Δ jaw to $-12/+2$ oligo and $[-12/+20][+4/+20]$ probe (Table 1). In agreement with previous results, RNAP- β' Δ jaw binds the downstream fork junction probe much weaker than wild-type RNAP;¹⁴ however, the mutant RNAP affinity is ~ 30 -fold higher than that of $E\sigma^{70}$ -gp2. The data also show that, unlike the addition of gp2, deletion of the β' jaw has only a slight effect on the $-12/+2$ oligo binding (Table 1).

As noted above, the ExIII footprinting data suggested that gp2 might affect RNAP-promoter interactions at around position -42 . Moreover, this effect must be distinct from the “normal” function of gp2 as it is observed whether gp2 is added before or after promoter complex formation (Fig. S2). To determine the significance of this new effect, we measured the consequences of gp2 addition on RNAP binding to $[-59/-12]$ probe. A control, T4 AsiA protein, which is known to inhibit RNAP interaction with the -35 element,¹⁶ strongly inhibited RNAP binding to this probe (Fig. 7). In contrast, gp2 had very little or no effect on the binding of this probe to $E\sigma^{70}$ (Fig. 7). We conclude that gp2 does not noticeably affect RNAP-promoter interactions upstream of the -10 promoter element.

Effect of gp2 on RNAP interaction with double-stranded promoter fragments truncated downstream of the transcription start point

The DNase I footprinting of the ternary $E\sigma^{70}$ -gp2-N25cons complex suggests that gp2 blocks RNAP-

promoter contacts downstream from approximately $+8$ position. A set of double-stranded probes was used to assess the influence of promoter segments downstream of the $+6$ position on promoter interactions with the RNAP-gp2 complex. The parent probe was a derivative of the $[-50/+14]$ fragment of the T5 N25 promoter containing the consensus -35 element and the extended -10 element (Fig. 1). The binding of $[-50/+14]$ and three other derivative probes truncated at $+6$, $+8$, and $+12$ was assessed by measuring RNAP beacon signals upon the addition of (i) gp2 to preformed $E\sigma^{70}$ complexes with each probe or (ii) DNA probes to the $E\sigma^{70}$ -gp2 complex. Upon reaching equilibrium binding, signal amplitudes should be the same independent of the order of DNA and gp2 addition. As can be seen in Fig. 8a, the binding curves were strikingly different when gp2 was added to preformed $E\sigma^{70}$ -DNA complexes. The fluorescence intensity decrease observed with the $[-50/+14]$ and $[-50/+12]$ probes was very slow. In

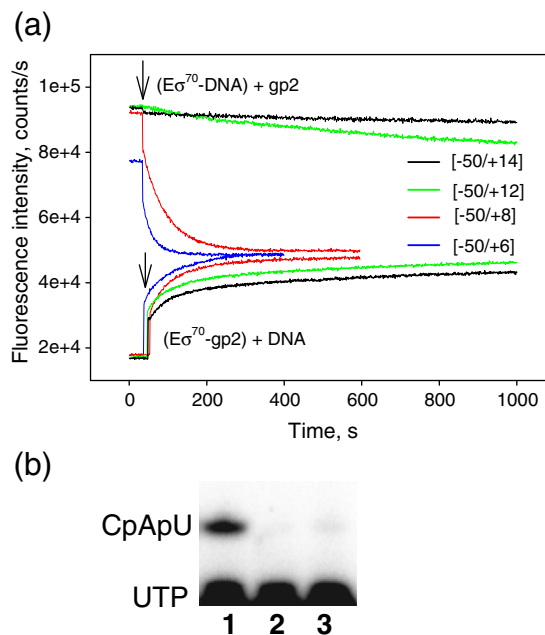


Fig. 8. The effect of gp2 on $E\sigma^{70}$ interaction with double-stranded promoter fragments truncated at various positions downstream of the transcription start point. (a) Time dependence of fluorescence intensity in samples containing 1 nM [211Cys-TMR] σ^{70} holoenzyme, 200 nM gp2, and 2 nM either of $[-50/+14]$, $[-50/+12]$, $[-50/+8]$, or $[-50/+6]$ fragments. Upper curves correspond to experiments in which gp2 was added to the RNAP beacon preincubated with a DNA probe for 30 min; bottom curves show fluorescence intensity changes upon addition of the DNA probes to the RNAP beacon preincubated with gp2. The addition of gp2 or DNA probe is indicated by vertical arrows. (b) Inhibition of abortive transcript synthesis from the $[-50/+6]$ probe by gp2. Reactions contained $E\sigma^{70}$ alone (lane 1) or RNAP with gp2 added before (lane 2) or after (lane 3) open complex formation.

contrast, signal amplitudes rapidly reached equilibrium values upon the addition of gp2 to $E\sigma^{70}$ complexed with $[-50/+6]$ or $[-50/+8]$ probes. Moreover, abortive transcript synthesis from $[-50/+6]$ was inhibited when gp2 was added either prior or after this probe (Fig. 8b, compare lanes 2 and 3). Combined with DNase I footprinting and structural modeling¹⁰ data, these results suggest that $E\sigma^{70}$ interactions with duplex DNA downstream of ca position +9 and with gp2 are mutually exclusive. On the other hand, $E\sigma^{70}$ -DNA contacts upstream of ca position +9 are compatible with gp2 binding; however, the interaction with gp2 leads to formation of an inactive $E\sigma^{70}$ -gp2-promoter complex.

Discussion

The T7 gp2 is a powerful inhibitor of promoter complex formation on most promoters. RNAP interactions with gp2 and promoter were thought to be mutually exclusive, and promoters that are resistant to gp2, such as the *galP1* promoter,⁵ escape inhibition by displacing gp2 from RNAP via optimized interactions of downstream DNA with the enzyme.⁹ The principal result of this work is the demonstration that, under conditions when RNAP contacts with promoter elements are strong, a complex between RNAP σ^{70} holoenzyme, T7 gp2, and promoter can form. This complex is transcriptionally inactive; however, based on DNase I and ExoIII footprints and $KMnO_4$ probing, it resembles "normal" RP_o at and upstream of the -10 promoter element. Downstream of the -10 element, the complex containing gp2 differs from RP_o : the transcription start point is not melted, and contacts with DNA downstream of position +8 are lacking (in RP_o , strong downstream protection extends to position +18). Thus, it appears that, in the presence of gp2, melting of promoter is initiated normally but does not proceed to the transcription initiation start point. The inhibition of promoter melting near the transcription start site is consistent with a proposal based on a structural model of the gp2-RNAP complex.¹⁰ While the promoter complex we describe here was obtained on a specially designed promoter, we surmise that similar complexes may be formed on other promoters. However, these complexes are too transient to be detected by standard biochemical assays.

It could have been expected that gp2 prevents RP_o formation simply by sterically blocking loading of downstream DNA into the enzyme trough formed by the β' jaw and several other RNAP domains.^{9,10} However, the fact that gp2 inhibits transcription from open complexes formed on a model template, which lacks DNA downstream from the +6 position (Fig. 8b), suggests that the mechanism of gp2 inhibition is more complex. In

fact, the absence of promoter opening at and immediately upstream of transcription start point may be caused by gp2 binding *per se*, rather than by the absence of downstream DNA loading. Indeed, gp2 inhibits RNAP binding to various promoter fragments lacking distal part of the downstream promoter duplex segment, including oligonucleotides and upstream fork junctions probes that bear no DNA bases downstream from the transcription start site. Since these effects occur in the absence of β' jaw-DNA contacts, they, therefore, indicate that gp2 inhibits RNAP interaction with downstream part of the transcription bubble between positions -6 and +2. Since the site of the binding of gp2, a small protein, is located away from RNAP sites interacting with single-stranded promoter DNA, we believe that this inhibition may be of allosteric nature. For example, the binding of gp2 to the β' jaw may interfere with movements of the RNAP β' clamp domain that may be required for establishment of interactions with single-stranded DNA segments.^{17,18} In the RNAP holoenzyme, σ^{70} subunit region 1.1 is positioned inside the RNAP-active-site channel from where it is displaced by the downstream promoter segment upon RP_o formation.^{19,20} Recent results suggest that efficient inhibition of $E\sigma^{70}$ by gp2 requires the σ^{70} region 1.1.⁶ Thus, the inhibition of RNAP interactions with single-stranded DNA by gp2 may also involve the σ^{70} region 1.1. Additional biophysical experiments will be required to test these ideas.

The high affinity for RNAP and the two-pronged mechanism of inhibition of open promoter complex formation pathway, first, by affecting promoter opening, and second, by interfering with the β' jaw contacts with downstream DNA, make gp2 a very efficient inhibitor. This is a biologically relevant property, since any leaky transcription from early phage promoters, particularly the A3 promoter,⁸ leads to transcriptional interference between the slow-moving host RNAP and the fast-moving T7 RNAP, causing abnormal progeny phage genome processing and packaging.

Melting of the transcription start site is an obviously critical step during the formation of the catalytically competent open promoter complex. Apparently, different factors and/or changes in the RNAP structure can affect promoter melting near the start site, without having an effect on the nucleation of promoter melting further upstream. The complex formed in the presence of gp2 appears to be similar to previously described complex formed by RNAP lacking the β lobe.²¹ However, the mutant RNAP complexes establish proper downstream contacts in the presence of transcription substrates and are transcriptionally active. The antibiotic myxopyronin binds to a pocket inside the RNAP β clamp domain and blocks promoter melting propagation toward the active site downstream from the -3 position,^{18,22}

again resulting in a complex that may be similar to the gp2-containing complex. Some promoters having GC-rich discriminator sequences are known to form unstable open complexes, a property that is relevant for their specific inhibition during amino acid starvation.^{2,23,24} It is possible that the stringent control-dependent destabilization of complexes formed on these promoters is due to the same mechanism that is used by gp2. Future work will have to establish how such partially opened promoter complexes are related to intermediate complexes on the pathway to RP_o formation.^{25–27}

Our work demonstrates that protein beacon assay is a powerful approach to uncover mechanistic details of transcription factor action. The high sensitivity and the ability to probe RNAP interactions with upstream, downstream, and central (corresponding to the transcription bubble) regions of a promoter allow the use of the beacon assay to study even such potent transcription initiation inhibitors as T7 gp2. In addition, because the beacon assay signal relies on unquenching of σ^{70} region 2.3 aromatic residues with promoter DNA, it allows one to study the local environment of the region 2.3 aromatic cluster in response to different inputs. Other transcription regulators and inhibitors affecting RNAP–promoter interactions, including numerous factors encoded by bacteriophages, can be informatively analyzed by this technique.

Materials and Methods

Proteins

RNAP holoenzyme containing σ^{70} derivative labeled at position 211 with fluorescent label 5-tetramethylrhodamine (TMR) ([211Cys-TMR] σ^{70}) was prepared as in Ref. 11. T7 phage gp2 wild type and R58E mutant, *E. coli* core RNAP containing a deletion (1149–1190) in the β' subunit, and T4 phage AsiA protein were prepared as described previously.^{5,10,16}

DNA probes

DNA oligonucleotides and modified oligonucleotides containing the SP18 spacer were synthesized by Integrated DNA Technologies. Fork junction and double-stranded DNA probes were prepared as in Refs. 11 and 14. A control non-promoter DNA fragment consisted of mutated –65/+35 T5 N25 strands in which –10 and –35 promoter element sequences were changed to CGCGGC and CCACTT, respectively.

Fluorometric assays

Fluorescence measurements were performed using a QuantaMaster QM4 spectrofluorometer (PTI) in transcription buffer [40 mM Tris–HCl (pH 8.0), 100 mM NaCl, 5%

glycerol, 1 mM DTT, and 10 mM MgCl₂] containing 0.02% Tween 20 at 25 °C. Final assay mixtures (800 μ l) contained 1 nM labeled RNAP holoenzyme and DNA probes at various concentrations. The TMR fluorescence intensities were recorded with an excitation wavelength of 550 nm and an emission wavelength of 578 nm. Time-dependent fluorescence changes were monitored after manual mixing of RNAP beacon (800 μ l) and a DNA probe (<20 μ l) in a cuvette; the mixing dead-time was 15 s.

To obtain equilibrium dissociation constants (K_d) of the RNAP beacon with oligonucleotides and fork junctions, we fitted the experimental dependence of the fluorescent signal amplitude on DNA probe concentration to a chemical equilibrium equation (i.e., titration assay) as described in Ref. 11, unless otherwise noted.

The complex of $E\sigma^{70}$ with gp2 ($E\sigma^{70}$ –gp2) was prepared by mixing 1 nM $E\sigma^{70}$ beacon with 200 nM gp2 followed by 10-min incubation at room temperature. Increasing gp2 concentration up to 600 nM did not influence binding of DNA probes to $E\sigma^{70}$ –gp2 under conditions of the measurements.

In vitro abortive initiation assay

Abortive transcription reactions were performed in a final volume of 10 μ l and contained 200 nM *E. coli* σ^{70} holoenzyme (Epicentre), 500 nM gp2 (where indicated), and 50 nM various DNA templates prepared as described above in transcription buffer. Reactions were mixed as indicated and incubated for 10 min at 37 °C, followed by the addition of CpA RNA dinucleotide primer (200 μ M), cold UTP (20 μ M), and [α -³²P]UTP (3000 Ci/mmol). The reactions were incubated for a further 10 min at 37 °C and then terminated by the addition of an equal volume of urea-formamide loading buffer. The reaction products were resolved on a 20% (w/v) polyacrylamide denaturing gel and visualized using a PhosphorImager.

Footprints

For *in vitro* footprinting experiments, the double-stranded DNA promoter fragments were prepared as in Ref. 11 using radioactively 5' end-labeled with [³²P] γ ATP oligonucleotides corresponding to either non-template or template strand. Next, the labeled promoter fragments were purified on Micro Bio-Spin 6 Columns (Bio-Rad) and used for the assays.

DNase I footprinting

Promoter complexes were formed in a final volume of 10 μ l and contained 200 nM *E. coli* σ^{70} holoenzyme (Epicentre), 2 μ M gp2 (where indicated), and 50 nM DNA template in transcription buffer. Reactions were mixed with the order of addition of particular components as indicated, incubated for 10 min at 37 °C, followed by the addition of 0.2 units of DNase I (New England Biolabs). The reactions were proceeded for 30 s at 37 °C and were terminated by the addition of 20 μ l of stop buffer (10 μ g of calf thymus DNA and water) followed by ethanol precipitation. Samples were dissolved in 8 μ l of urea-formamide loading buffer and resolved in 7% polyacrylamide denaturing gel.

KMnO₄ probing

Reactions were set up as described for DNase I footprinting. Promoter complexes were treated with 1 mM KMnO₄ at room temperature for 30 s. Reactions were terminated by the addition of 300 mM β -mercaptoethanol followed by ethanol precipitation and 20-min treatment with 10% piperidine at 95 °C. Reaction products were next treated with chloroform to remove piperidine, ethanol precipitated, dissolved in 8 μ l of urea-formamide loading buffer, and resolved in 7% polyacrylamide denaturing gel.

ExoIII footprinting

Reactions were set up as described above. Promoter complexes were treated with 0.5 units of ExoIII (New England Biolabs) for 2 min at 37 °C. Reactions were terminated by the addition of 20 μ l of stop buffer (10 μ g of calf thymus DNA and water) and precipitated with ethanol. Samples were dissolved in 8 μ l of urea-formamide loading buffer and resolved in 7% polyacrylamide denaturing gel.

Supplementary materials related to this article can be found online at [doi:10.1016/j.jmb.2011.09.029](https://doi.org/10.1016/j.jmb.2011.09.029)

Acknowledgements

This work was supported by the National Institutes of Health grants GM59295 and GM64530 and the Russian Academy of Sciences Presidium Molecular and Cellular Biology Program grant to K.S. S.W. is the recipient of a Biotechnology and Biological Sciences Research Council David Phillips Fellowship (BB/E023703), and C.S. is the recipient of a Biotechnology and Biological Sciences Research Council Doctoral Training Grant studentship.

References

1. Nechaev, S. & Severinov, K. (2003). Bacteriophage-induced modifications of host RNA polymerase. *Annu. Rev. Microbiol.* **57**, 301–322.
2. 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.
3. Molineux, I. J. (2005). The T7 group. In *The Bacteriophages, Second Edition* (Calendar, R. L., ed.), pp. 277–301, Oxford University Press, New York, NY.
4. Hessellbach, B. & Nakada, D. (1975). Inactive complex formation between *E. coli* RNA polymerase and inhibitor protein purified from T7 phage infected cells. *Nature*, **258**, 354–357.
5. Nechaev, S. & Severinov, K. (1999). Inhibition of *Escherichia coli* RNA polymerase by bacteriophage T7 gene 2 protein. *J. Mol. Biol.* **289**, 815–826.
6. Sheppard, C., Cámara, B., Shadrin, A., Akulenko, N., Liu, M., Baldwin, G. *et al.* (2011). Inhibition of *Escherichia coli* RNAP by T7 Gp2 protein: role of negatively charged strip of amino acid residues in Gp2. *J. Mol. Biol.* **407**, 623–632.
7. Qimron, U., Kulczyk, A., Hamdan, S., Tabor, S. & Richardson, C. (2007). Inadequate inhibition of host RNA polymerase restricts T7 bacteriophage growth on hosts overexpressing *udk*. *Mol. Microbiol.* **67**, 448–457.
8. 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.
9. 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.
10. Camara, B., Liu, M., Reynolds, J., Shadrin, A., Liu, B., Kwok, K. *et al.* (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. USA*, **107**, 2247–2252.
11. Mekler, V., Pavlova, O. & Severinov, K. (2011). The interaction of *Escherichia coli* RNA polymerase σ^{70} subunit with promoter elements in the context of free σ^{70} , RNA polymerase holoenzyme, and the β' - σ^{70} complex. *J. Biol. Chem.* **286**, 270–279.
12. Estrem, S. T., Gaal, T., Ross, W. & Gourse, R. L. (1998). Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl Acad. Sci. USA*, **95**, 9761–9766.
13. Guo, J. & Gralla, J. D. (1998). Promoter opening via a DNA fork junction binding activity. *Proc. Natl Acad. Sci. USA*, **95**, 11655–11660.
14. Mekler, V., Minakhin, L. & Severinov, K. (2011). A critical role of downstream RNA polymerase–promoter interactions in the formation of initiation complex. *J. Biol. Chem.* **286**, 22600–22608.
15. Marr, M. T. & Roberts, J. W. (1997). Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science*, **276**, 1258–1260.
16. Severinova, E., Severinov, K., Fenyö, D., Marr, M., Brody, E. N., Roberts, J. W. *et al.* (1996). Domain organization of the *Escherichia coli* RNA polymerase σ^{70} subunit. *J. Mol. Biol.* **263**, 637–647.
17. Darst, S. A., Opalka, N., Chacon, P., Polyakov, A., Richter, C., Zhang, G. & Wriggers, W. (2002). Conformational flexibility of bacterial RNA polymerase. *Proc. Natl Acad. Sci. USA*, **99**, 4296–4301.
18. Mukhopadhyay, J., Das, K., Ismail, S., Koppstein, D., Jang, M., Hudson, B. *et al.* (2008). The RNA polymerase “switch region” is a target for inhibitors. *Cell*, **135**, 295–307.
19. Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A. N. *et al.* (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase promoter open complex. *Cell*, **108**, 599–614.
20. Murakami, K. S., Masuda, S. & Darst, S. A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science*, **296**, 1280–1284.

21. Severinov, K. & Darst, S. A. (1997). A mutant RNA polymerase that forms unusual open promoter complexes. *Proc. Natl Acad. Sci. USA*, **94**, 13481–13486.
22. Belogurov, G. A., Vassilyeva, M. N., Sevostyanova, A., Appleman, J. R., Xiang, A. X., Lira, R. *et al.* (2009). Transcription inactivation through local refolding of the RNA polymerase structure. *Nature*, **457**, 332–335.
23. Paul, B. J., Ross, W., Gaal, T. & Gourse, R. L. (2004). rRNA transcription in *Escherichia coli*. *Annu. Rev. Genet.* **38**, 749–770.
24. 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.
25. Saecker, R. M., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr, Capp, M. W. & Record, M. T., Jr (2002). Kinetic studies and structural models of the association of *E. coli* σ^{70} RNA polymerase with the λP_R promoter: large scale conformational changes in forming the kinetically significant intermediates. *J. Mol. Biol.* **319**, 649–671.
26. Slavi, B., Zaychikov, E., Rogozina, A., Walther, F., Buckle, M. & Heumann, H. (2005). Real-time characterization of intermediates in the pathway to open complex formation by *Escherichia coli* RNA polymerase at the T7A1 promoter. *Proc. Natl Acad. Sci. USA*, **102**, 4706–4711.
27. 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.

Structural and Mechanistic Basis for the Inhibition of *Escherichia coli* RNA Polymerase by T7 Gp2

Ellen James,^{1,6} Minhao Liu,^{1,6} Carol Sheppard,^{1,6} Vladimir Mekler,² Beatriz Cámara,³ Bing Liu,¹ Pete Simpson,¹ Ernesto Cota,¹ Konstantin Severinov,^{2,4,5} Steve Matthews,^{1,*} and Sivaramesh Wigneshweraraj^{1,*}

¹MRC 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

³Laboratorio de Microbiología Molecular y Biotecnología Ambiental, Departamento de Química, Universidad Técnica Federico Santa María, Valparaíso, Chile

⁴Institute of Molecular Genetics, Russian Academy of Science, Moscow 123182, Russia

⁵Institute of Gene Biology, Russian Academy of Science, Moscow 11991, Russia

⁶These authors contributed equally to this work

*Correspondence: s.j.matthews@imperial.ac.uk (S.M.), s.r.wig@imperial.ac.uk (S.W.)

<http://dx.doi.org/10.1016/j.molcel.2012.06.013>

SUMMARY

The T7 phage-encoded small protein Gp2 is a non-DNA-binding transcription factor that interacts with the jaw domain of the *Escherichia coli* (Ec) RNA polymerase (RNAP) β' subunit and inhibits transcriptionally proficient promoter-complex (RPo) formation. Here, we describe the high-resolution solution structure of the Gp2-Ec β' jaw domain complex and show that Gp2 and DNA compete for binding to the β' jaw domain. We reveal that efficient inhibition of RPo formation by Gp2 requires the amino-terminal σ^{70} domain region 1.1 (R1.1), and that Gp2 antagonizes the obligatory movement of R1.1 during RPo formation. We demonstrate that Gp2 inhibits RPo formation not just by steric occlusion of the RNAP-DNA interaction but also through long-range antagonistic effects on RNAP-promoter interactions around the RNAP active center that likely occur due to repositioning of R1.1 by Gp2. The inhibition of Ec RNAP by Gp2 thus defines a previously uncharacterized mechanism by which bacterial transcription is regulated by a viral factor.

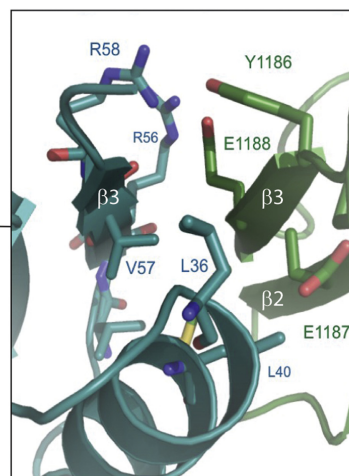
INTRODUCTION

Transcription of DNA is a major focal point of regulation of gene expression in all organisms. In bacteria, transcription is catalyzed by a multisubunit RNA polymerase (RNAP) with subunit composition $\alpha_2\beta\beta'\omega\sigma$ (abbreviated as E σ). The catalytic determinants of the bacterial RNAP are contained in the largest and second-largest subunits (β' and β , respectively). The overall architecture of RNAP is reminiscent of a crab claw (Murakami et al., 2002b). The two pincers of the claw form a positively charged DNA binding channel (DBC). The active center where

RNA synthesis occurs is located deep within the DBC (Figure 1A). A number of flexible domains from the β' and β subunits (β' jaw, β' insertion 6, β' downstream clamp, and β downstream and upstream lobe domains) surround the DBC and contribute to stable binding of DNA (Opalka et al., 2010).

A dissociable RNAP binding factor called the sigma (σ) subunit confers promoter specificity upon the RNAP by recognizing the -35 and -10 (with respect to the transcription start site at $+1$, hereafter called the $+1$ site) elements present in most bacterial promoters. Of the seven σ factors encoded by the *Escherichia coli* (Ec) genome, σ^{70} is responsible for transcription of housekeeping genes and is therefore a major σ factor essential for growth of the cell (Haugen et al., 2008). The six remaining “alternative” σ factors contribute to transcription of genes under specific stress conditions, growth transitions, and/or morphological changes (Gruber and Gross, 2003). Other bacteria also have one σ^{70} -like major σ factor that is responsible for transcription of housekeeping genes, and a variable number of alternative σ factors. All alternative σ factors, with the exception of σ^{54} , are related to σ^{70} .

In Ec, transcription at most σ^{70} -dependent promoters begins with σ^{70} -directed reversible binding of E σ^{70} to the promoter, which results in the formation of the initial closed promoter complex (RPc). At most promoters, the RPc is unstable and either dissociates or isomerizes via several intermediates to the transcription-initiation-competent open promoter complex (RPo). In the RPo, the promoter DNA strands are locally melted and form a transcription bubble spanning positions -12 to $+3$ of the promoter and the $+1$ position of the template strand placed at the RNAP active center (Figure 1A) (Murakami et al., 2002a; Saecker et al., 2011). The double-stranded DNA (dsDNA) immediately downstream of the active center (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 (Murakami et al., 2002a; Saecker et al., 2011). The β' jaw, β' insertion 6, and β' downstream clamp domains contribute to the dwDBC (Murakami et al., 2002a). The dwDNA interaction with the dwDBC can be divided into active



center-proximal (+5 to +8) and -distal (+10 to +20) sets of interactions. The active center-distal interactions mainly involve the β' jaw, β' insertion 6, and β clamp domains. In the structural model of the RPo, the β' jaw domain aligns with the path of the dwDNA and is located closest to the dwDNA. The deletion of the β' jaw domain (amino acid residues 1149–1190) in the *Ec* RNAP dramatically reduces the stability of the RPo (Ederth et al., 2002). Thus, it is very likely, but as yet experimentally unproven, that the β' jaw domain makes sequence-nonspecific contacts with the dwDNA during RPo formation.

Within the σ^{70} family, only proteins that function as major σ factors contain an extended (~100 amino-acid-long) amino-terminal domain, called region 1.1 (R1.1). Sequences of R1.1 from major σ factors from various sources are variable but tend to be acidic. At σ^{70} -dependent promoters, R1.1 plays an important regulatory role during transcription initiation (see below). Although the solution structure of the isolated domain of R1.1 (amino acid residues 1–100) of *Thermotoga maritima* σ^A (a counterpart of *Ec* σ^{70}) is available (Schwartz et al., 2008), the R1.1 domain is not resolved in the crystal structures of the σ^A -containing RNAP from *Thermus aquaticus* and *T. thermophilus* (Murakami et al., 2002a, 2002b; Vassilyev et al., 2002), implying that it is disordered and/or can adopt multiple conformations. Fluorescence resonance energy transfer (FRET) analyses indicate that in the absence of promoter DNA, the position of R1.1 in $E\sigma^{70}$ corresponds to that of the dwDNA in the RPo (Mekler et al., 2002). In other words, in $E\sigma^{70}$, R1.1 is located deep within the active-center cleft, just above the floor of the dwDBC channel, and is positioned to interact with the floor and walls of the dwDBC (Figure 1A) (Mekler et al., 2002). In the RPo, R1.1 is displaced to the tip of the β -subunit pincer (Mekler et al., 2002). Hence, it has been proposed that in free $E\sigma^{70}$, R1.1 acts as a mimic of dwDNA and must be displaced for dwDNA to enter the dwDBC and allow the RNAP active center to access the DNA (Mekler et al., 2002).

Gp2 is a 7 kDa, T7 bacteriophage-encoded, non-DNA-binding transcription factor. The essential biological function of Gp2 is to coordinate transcription of the phage genome by the host and viral RNAPs (Savalia et al., 2010). Gp2 binds tightly to the β' jaw domain of *Ec* RNAP (Cámara et al., 2010; Nechaev and Severinov, 1999) and strongly inhibits transcription from σ^{70} -dependent bacterial promoters (Cámara et al., 2010; Nechaev and Severinov, 1999). Gp2 inhibits RPo formation but does not bind to (and therefore has no effect on) preformed RPo (Nechaev and Severinov, 1999). Recently, using an optimized variant of the σ^{70} -dependent T5 N25 promoter (hereafter called N25cons), we trapped and characterized a ternary $E\sigma^{70}$ -Gp2-promoter complex that had normal contacts with promoter DNA upstream of the +1 site and a partially open transcription bubble, but was transcriptionally inactive and lacked the interaction with the dwDNA (Mekler et al., 2011b). Here, we describe the near-atomic resolution solution structure of Gp2 bound to a fragment

of the β' jaw domain (*Ec* β' subunit residues 1153–1213). We show by nuclear magnetic resonance (NMR) chemical shift perturbation analyses that the β' jaw domain interacts directly with dsDNA, and that the interaction surface for dsDNA overlaps with the Gp2 binding surface. By using RNAP containing mutant σ^{70} lacking R1.1 (amino acid residues 1–100) and determining the effect of the *in trans* addition of the isolated R1.1 domain on the mutant RNAP activity in the presence of Gp2, we demonstrate that Gp2 inhibits RPo formation by $E\sigma^{70}$ not only by sterically preventing dwDNA binding to the dwDBC but also by exerting R1.1-dependent and long-range antagonistic effects on RNAP-promoter DNA interactions near the RNAP active center. Results from biophysical analyses demonstrate that Gp2 changes the microenvironment of R1.1 in $E\sigma^{70}$ and restricts the obligatory displacement of R1.1 from the dwDBC during RPo formation. A combination of direct (by competing for dwDNA binding) and indirect (mediated by R1.1) effects makes Gp2 a highly efficient inhibitor of RPo formation by $E\sigma^{70}$.

RESULTS

Structure of the Complex between Gp2 and the β' Jaw Domain

A lysine substitution at residue E1158 or E1188 in the β' jaw domain prevents Gp2 from binding to the *Ec* RNAP (Figure S1A available online) (Cámara et al., 2010; Nechaev and Severinov, 2003). Therefore, the major Gp2 interacting surface on the β' jaw domain probably includes residues 1158–1188. In agreement with this view, deletion of residues 1149–1190 in the *Ec* RNAP confers resistance to inhibition by Gp2 (Nechaev and Severinov, 1999). We determined the solution structure of the complex between Gp2 and a fragment of the β' jaw domain (representing *Ec* β' jaw domain residues 1153–1213; hereafter referred to as the β' jaw fragment) using multidimensional NMR spectroscopy applied to hybrid-labeled complexes. Backbone C^α , C^β , C' , N , and H^N assignments for each of the labeled binding partner were obtained from HNCACB/CBCA(CO)NH and HN(CA)CO/HNCO spectra, and side-chain assignments were completed with the use of HCCH total correlation spectroscopy spectra. Broadening of some residues in the binding interface indicates the presence of conformational exchange in the isolated complex, perhaps due to the β' jaw fragment not fully representing the complete RNAP interaction. Despite this, the interface was initially characterized based on the manual unambiguous assignment of a few intermolecular nuclear Overhauser effect (NOEs) from $^{13}C/^{15}N$ -filtered NOE spectroscopy-heteronuclear single quantum coherence (HSQC) spectra. Automated NOE assignment methods using the ARIA program were used to complete the NOE assignment of the complex and calculate a family of 10 structures (Figure S1B and Table 1). The solution structure of the *Ec* β' jaw fragment consists primarily of a four-stranded antiparallel β sheet (Figures 1B and 1C) and shows

Figure 1. RPo and RPo Formation and the Structure of the Gp2- β' Jaw Fragment Complex

(A) Cartoon depiction of RPo and RPo formation at σ^{70} -dependent bacterial promoters (the inspiration for the cartoon was taken from Murakami and Darst [2003]). (B) Ribbon representation of the Gp2- β' jaw fragment complex. (C) The same as (B) but rotated by 90° along the horizontal plane. In (B) and (C), the interface region is enlarged in the insets, and the residues located at the interaction interface are shown as sticks and labeled correspondingly. See also Figure S1.

Table 1. Gp2-Jaw Complex

Number of experimental restraints	1,700
Total NOE-derived	1,532
Ambiguous	476
Unambiguous	1,056
Intraresidue	468
Sequential	200
Medium-range ($ i - j \leq 4$)	73
Long-range ($ i - j > 4$)	265
Intermolecular	50
TALOS, ϕ/ψ	168
RMSD from experimental restraints	
Distance, Å	0.023 ± 0.002
Dihedral angle, degrees	0.6 ± 0.1
RMSD from idealized covalent geometry	
Bonds, Å	0.0036 ± 0.0001
Improper angles, degrees	1.35 ± 0.06
Angles, degrees	0.54 ± 0.01
Coordinate RMSD, Å	
Backbone atoms in secondary structure	1.21 ± 0.15
Heavy atoms in secondary structure	1.47 ± 0.15
Ramachandran plot	
Residues in most favored regions, %	75.7
Residues in allowed regions, %	24.2
Residues in disallowed regions, %	0.1
RMSD, root mean-squared deviation.	

high structural similarity to the corresponding region of the β' subunit in the crystal structures of *T. aquaticus* and *T. thermophilus* RNAP (Figure S1C). In the complex with Gp2, the β sheet is extended to a seven-stranded β sheet in an antiparallel arrangement (Figure 1B). The primary interface region is localized to $\beta 3$ and $\beta 2$ of the β' jaw fragment and $\beta 3$ and $\alpha 1$ of Gp2. The two invariant arginine residues in Gp2 that are important for binding to the RNAP, R56 and R58 (Cámara et al., 2010), are located in the interface region in close proximity to E1188, providing a significant ionic interaction across the interface (Figure 1B, inset). Additional interfacial residues include L36, L40, T55, and V57 from Gp2, and V1176, Y1186, E1187, E1188, and M1189 from the β' jaw domain (Figures 1A and 1B, insets). Residues L36, L40, and V57 of Gp2 form the major hydrophobic contacts with V1176 and M1189 of the β' jaw domain at the binding interface (Figures 1A and 1B, insets). Gp2 contains a contiguous strip of seven negatively charged amino acids (E21, E34, D37, E38, E41, E44, and E53; hereafter referred to as the negatively charged strip [NCS]) on the side of the molecule opposing R56 and R58 (Sheppard et al., 2011). Analyses of the role of the NCS by mutagenesis reveal that the NCS is not important for the binding of Gp2 to RNAP, but the disruption of the NCS significantly attenuates the ability of Gp2 to inhibit RPo formation (Sheppard et al., 2011). An examination of the surface electrostatic properties of the Gp2- β' jaw fragment complex reveals that the NCS in Gp2 is extended by residues E1158, D1181, D1184, E1187, and E1188 of the β' jaw domain (Figure S1D).

Interaction of the β' Jaw Domain with dsDNA

To better understand how Gp2 inhibits RPo formation by $E\sigma^{70}$, we derived a composite structural model of the Gp2-RNAP complex using our solution structure of the Gp2- β' jaw fragment complex together with structural models of the *Ec* core RNAP (Opalka et al., 2010) and RPo based on the structure of the *T. aquaticus* RNAP (Murakami et al., 2002a). In the composite structural model, the Gp2-binding surface of the β' jaw domain is facing the dwDBC and toward where the dwDNA would likely lie in the RPo (Figure 2A, insets *i* and *ii*). Because Gp2 antagonizes the interaction between the dwDNA and dwDBC during RPo formation (Mekler et al., 2011a, 2011b), we hypothesized that this region of the β' jaw domain would make sequence-nonspecific direct contacts with dsDNA, and the binding of Gp2 could block or modulate its interaction with dsDNA. To test whether an interaction exists between the β' jaw domain and dsDNA during RPo formation, we conducted an NMR titration experiment with a randomly generated 14 bp dsDNA fragment and ^{15}N -labeled β' jaw fragment. We recorded 2D ^1H - ^{15}N HSQC spectra to monitor the backbone amide chemical shift changes in the β' jaw fragment in the presence of DNA. The NMR spectrum exhibited several specific chemical shift changes, which were in fast exchange on the NMR timescale indicative of a binding constant in the micromolar to millimolar range (Figure 2B). The major perturbed residues (T1169, R1174, and M1189) map to the exposed surface of the $\beta 3$ sheet of the β' jaw fragment (Figure 2A, inset *ii*, and Figure S1A), suggesting that these residues are involved in interaction with dsDNA. Consistent with this view, results from formaldehyde crosslinking experiments showed that whereas the wild-type β' jaw fragment could be crosslinked to the ^{32}P -labeled 14 bp dsDNA, a mutant β' jaw fragment containing an alanine substitution at R1174 could not be detectably crosslinked to the dsDNA (Figure S1E, compare lanes 2 and 4). Further, the R1174A mutation in the context of $E\sigma^{70}$ formed a significantly reduced number of RPo compared with the wild-type $E\sigma^{70}$ (Figure S1F). However, once the RPo were formed, the stabilities of the mutant and wild-type RPo upon challenge with heparin were indistinguishable (Figure S1F), which is not surprising considering that the interface between $E\sigma^{70}$ and DNA is extensive in the RPo, and the effect of a single point mutation in the β' jaw domain on overall DNA binding by $E\sigma^{70}$ would be difficult to detect. Thus, in the composite model, we redefined the potential path of the dsDNA in the RPo with respect to the β' jaw domain (Figure 2A, inset *iii*). A comparison of the DNA-interacting surface in the β' jaw fragment with the Gp2-binding interface from our solution structure of the Gp2- β' -fragment complex reveals significant overlap and suggests that Gp2 and dsDNA may compete for overlapping interaction surfaces on the β' jaw domain in the RNAP (Figure 2A, inset *iii* and *iv*). To test this hypothesis, we performed a competition NMR experiment by titrating Gp2 into a saturated complex of the β' jaw-fragment-dsDNA complex and recorded the changes in the NMR spectrum. As shown in Figure 2C, the characteristic NMR spectrum of the Gp2- β' jaw-fragment complex was regained after the addition of Gp2, thus confirming that in the context of the isolated β' jaw fragment, Gp2 is able to displace dsDNA efficiently. Consistent with this view, results from formaldehyde crosslinking experiments showed that in

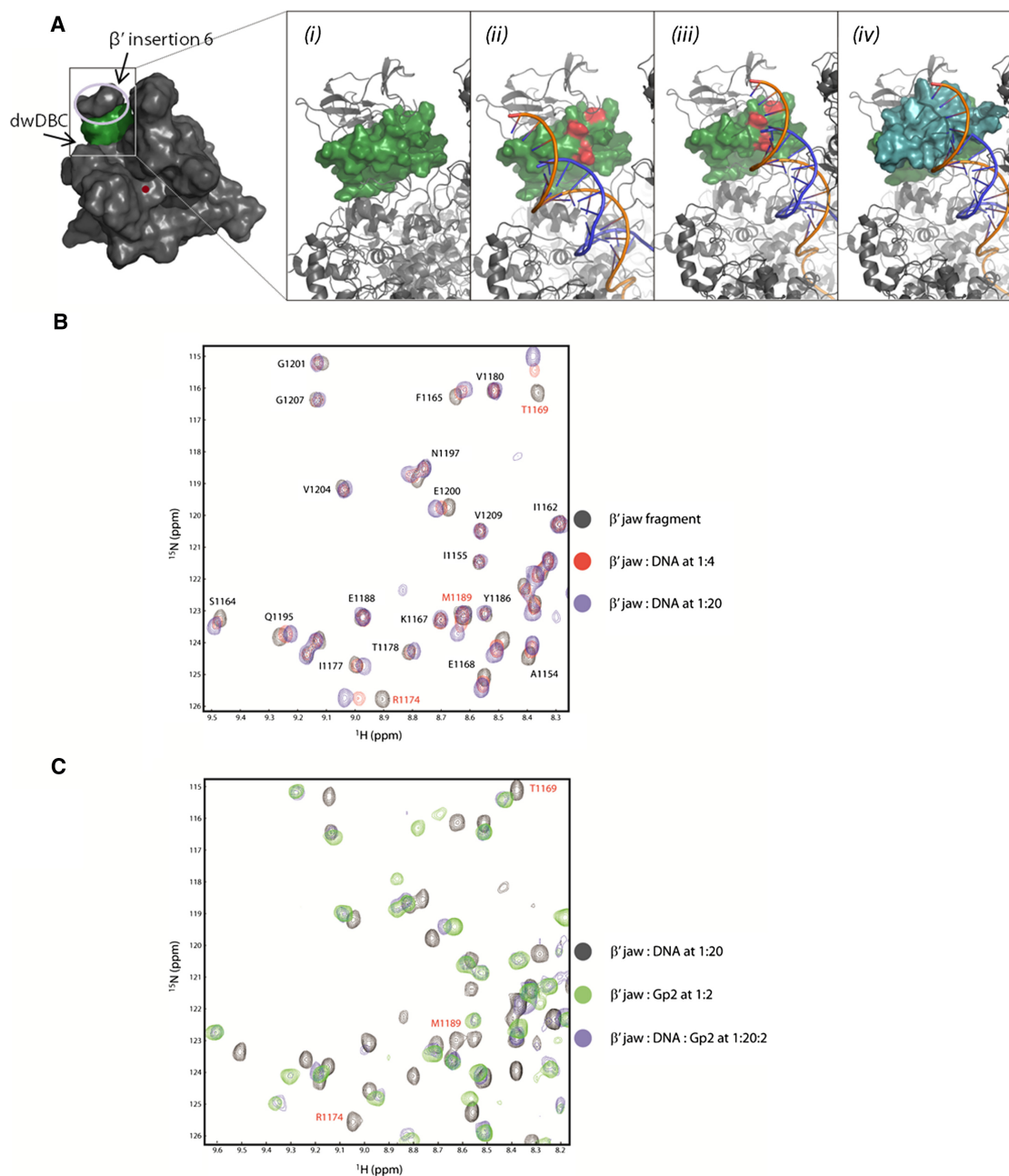


Figure 2. Interaction of the β' Jaw Domain with dsDNA

(A) Surface representation of the *Ec* core RNAP model (Opalka et al., 2010) color-coded as in Figure 1A. The boxed region is enlarged and looks at the DNA-binding surface (shown in ribbon representation in insets i–iv). The β' jaw domain is shown in green as a surface representation and forms part of the DNA-binding face (i). Inset (ii) is as in (i), but showing the path of the dwDNA from the current model of the RPo (Opalka et al., 2010). Highlighted in red are residues T1169, R1174, and M1189, which undergo significant chemical shift changes in β' jaw fragment in the presence of dsDNA (B) and Gp2 (C). Inset (iii) is as in (ii), but showing the redefined path of the dwDNA in the dwDBC. Inset (iv) is as in (iii), but with the surface representation of Gp2 shown in cyan. Note the lack of steric clash between Gp2 and the β' insertion 6 domain, which provides further support for our composite model.

(B) Overlay of 2D ^1H - ^{15}N HSQC spectra of the β' jaw fragment with and without dsDNA recorded at pH 6.5, 303 K (see key for details). Peaks with significant chemical shift differences are indicated in red with their residue numbers (T1169, R1174, and M1189).

(C) As in B, but showing the 2D ^1H - ^{15}N HSQC spectra of the β' jaw fragment with dsDNA (i.e., the β' jaw fragment is ^{15}N labeled) with or without unlabeled Gp2 (see key for details).

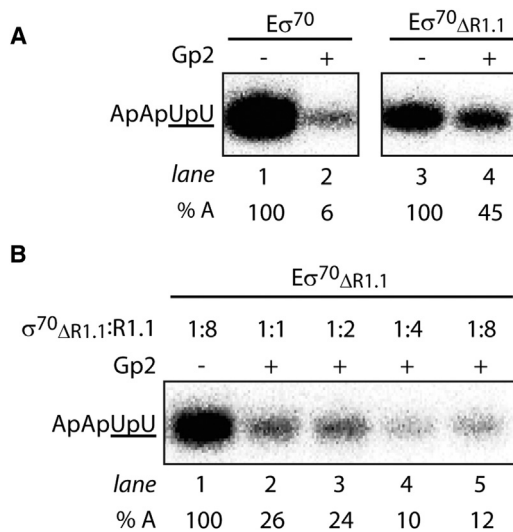


Figure 3. Gp2 Requires R1.1 to Efficiently Inhibit RPo Formation by Eσ⁷⁰

(A) Autoradiograph of 20% (v/v) denaturing urea gels showing the synthesis of the ApApUpU transcript (underlined nucleotides are α³²P labeled) from the *lacUV5* promoter by Eσ⁷⁰ (lanes 1 and 2) and Eσ⁷⁰_{ΔR1.1} (lanes 3 and 4) in the absence and presence of Gp2. The percentage of ApApUpU transcript synthesized (% A) in the reactions with Gp2 with respect to reactions with no Gp2 is given at the bottom of the gel for each reaction.

(B) As above, but showing the synthesis of the ApApUpU transcript in the absence (lane 1) and presence (lanes 2–5) of Gp2 under conditions in which Eσ⁷⁰_{ΔR1.1} was preincubated with increasing amounts of isolated R1.1 domain added in *trans* to the reaction (shown as the ratio of σ⁷⁰_{ΔR1.1} to R1.1). For (A) and (B), all data obtained in at least three independent experiments fell within 5% of the % A value shown. See also Figure S2.

the presence of Gp2, the β' jaw fragment could not be efficiently crosslinked to the ³²P-labeled 14 bp dsDNA (Figure S1E, compare lanes 2 and 3). Overall, the results strongly suggest that Gp2 and DNA compete for overlapping binding sites on the β' jaw domain in the RNAP, and provide a structural basis for and further insights (see below) into the mechanism by which Gp2 inhibits RPo formation by Eσ⁷⁰.

Gp2 Requires R1.1 to Efficiently Inhibit RPo Formation by Eσ⁷⁰

The location of Gp2 in our composite structural model of the Gp2-RNAP complex places Gp2 proximal to the location of R1.1 of σ⁷⁰ in Eσ⁷⁰ inferred from biophysical studies (Mekler et al., 2002). Thus, it is possible that Gp2 could affect R1.1 function during RPo formation, and therefore Gp2 could inhibit RPo formation by a mechanism involving R1.1 of σ⁷⁰. To test this hypothesis, we explored the role of R1.1 in the mechanism by which Gp2 inhibits RPo formation by Eσ⁷⁰. Initially, we determined the ability of Gp2 to inhibit Eσ⁷⁰ reconstituted with either wild-type σ⁷⁰ or σ⁷⁰_{ΔR1.1} using an in vitro transcription assay. Incubation of an ~2-fold molar excess of Gp2 with Eσ⁷⁰ before the addition of a DNA fragment containing the *lacUV5* promoter abolished the synthesis of *lacUV5*-specific ApApUpU transcript (Figure 3A, lanes 1 and 2) (Cámara et al., 2010). In contrast, under identical conditions, Eσ⁷⁰_{ΔR1.1} was inhibited far less efficiently

(~55% inhibition; Figure 3A, lanes 3 and 4) even though the affinity of Gp2 for Eσ⁷⁰_{ΔR1.1} and Eσ⁷⁰ did not differ detectably among the conditions under which the in vitro transcription assays were performed (Figures S2A and S2B). Therefore, it seems that full inhibition of Eσ⁷⁰ RPo formation by Gp2 requires R1.1 of σ⁷⁰. Consistent with this view, the in *trans* addition of the isolated domain of σ⁷⁰ R1.1 (encompassing σ⁷⁰ amino acids 1–100) to Eσ⁷⁰_{ΔR1.1} increased the efficiency of Eσ⁷⁰_{ΔR1.1} transcription inhibition by Gp2, bringing it to the same level as in the case of Eσ⁷⁰ (Figure 3B, lanes 2–5). Thus, the presence of the isolated R1.1 domain in *trans* in Eσ⁷⁰_{ΔR1.1} allows Gp2 to efficiently inhibit RNAP. Full inhibition of Eσ⁷⁰_{ΔR1.1} by Gp2 occurred when the isolated R1.1 domain was added in *trans* to the Eσ⁷⁰_{ΔR1.1} either before or after Gp2 binding (Figure S2C). Control reactions established that the in *trans* presence of the isolated R1.1 domain in Eσ⁷⁰_{ΔR1.1} (in the absence of Gp2) did not antagonize the ability of Eσ⁷⁰_{ΔR1.1} to synthesize the ApApUpU transcript (Figure S2D).

An alternative *Ec* σ factor, σ³⁸, does not contain R1.1 but is able to recognize some σ⁷⁰-dependent promoters (Gruber and Gross, 2003). We compared the ability of Gp2 to inhibit RPo formation by Eσ³⁸ and Eσ⁷⁰ on one such promoter, the *Ec osmE* promoter (Bordes et al., 2000). Although the affinity of Gp2 for Eσ³⁸ and Eσ⁷⁰ did not differ detectably among the conditions under which the in vitro transcription assays were performed (Figures S2A and S2B), Gp2 inhibited Eσ³⁸-dependent synthesis of the ApApCpA *osmE* transcript by only ~80% even under conditions in which the amount of Gp2 exceeded that of Eσ³⁸ by ~4-fold (Figure S2E). In contrast, transcription initiation by Eσ⁷⁰ from this promoter was barely detectable under the same conditions (Figure S2E). As expected, full inhibition of RPo formation by Eσ⁷⁰ on the *osmE* promoter was R1.1 dependent (Figure S2F, lanes 3–6). Moreover, in the presence of the R1.1 domain of σ⁷⁰ added in *trans*, Gp2 fully inhibited RPo formation by Eσ³⁸ on the *osmE* promoter (Figure S2F, lanes 7–10). In summary, even though no detectable differences in the affinity of Gp2 for Eσ⁷⁰, Eσ⁷⁰_{ΔR1.1}, and Eσ³⁸ were observed in two independent experiments (Figures S2A and S2B), we cannot exclude the possibility that the absence of R1.1 (as in the case of Eσ⁷⁰_{ΔR1.1} and Eσ³⁸) can affect the affinity of Gp2 for RNAP. However, the results strongly suggest that R1.1 of σ⁷⁰ is part of the mechanism by which Gp2 inhibits RPo formation by Eσ⁷⁰: Gp2 alone antagonizes dwDNA binding to the dwDBC, leading to partial inhibition of RPo formation, and full inhibition requires R1.1 of σ⁷⁰.

Gp2 Requires R1.1 of σ⁷⁰, but Not the Consensus Promoter DNA Sequences, to Fully Inhibit RPo Formation by Eσ⁷⁰

The σ⁷⁰ factor makes extensive contacts with the consensus promoter DNA sequences (i.e., the –35 and –10 promoter elements) in the RPC and during RPo formation. Because full inhibition of Eσ⁷⁰ RPo formation by Gp2 depends on R1.1 of σ⁷⁰, we considered whether interactions between Eσ⁷⁰ and the consensus promoter DNA sequences play any role in the mechanism by which inhibition of RPo formation by Gp2 at σ⁷⁰-dependent promoters occurs. To address this issue experimentally, we determined whether Gp2 could inhibit the catalytic activity of

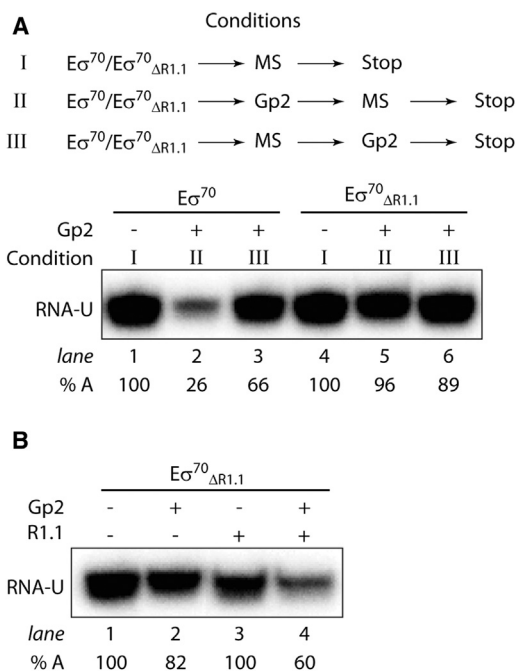


Figure 4. Gp2 Requires R1.1 of σ^{70} , but Not the Consensus Promoter DNA Sequences, to Fully Inhibit RPo Formation by $E\sigma^{70}$

(A) Autoradiograph of a 20% (v/v) denaturing urea gel showing the synthesis of RNA-U from the MS probe by $E\sigma^{70}$ and $E\sigma^{70}_{\Delta R1.1}$ in the absence (condition I) and presence (conditions II and III) of Gp2. The percentage of RNA-U synthesized (% A) in the reactions with Gp2 with respect to reactions with no Gp2 is given at the bottom of the gel for each reaction.

(B) As in (A), except that the reaction was conducted with $E\sigma^{70}_{\Delta R1.1}$ in the absence and presence (at ~8-fold molar excess over $\sigma^{70}_{\Delta R1.1}$) of isolated R1.1 domain added in *trans*. For (A) and (B), all data obtained in at least three independent experiments fell within 5% of the % A value shown. See also Figure S3.

$E\sigma^{70}$ on a promoterless minimal nucleic acid scaffold template (hereafter called the minimal scaffold [MS] probe). The MS probe consists of an 18-nucleotide-long DNA duplex and an 8-nucleotide-long RNA-DNA heteroduplex separated by two unpaired DNA bases (Kulbachinskiy et al., 2004) (Figure S3A). Thus, the MS probe lacks the consensus promoter DNA sequences recognized by σ^{70} . The addition of $\alpha^{32}\text{P}$ -UTP to the $E\sigma^{70}$ -MS probe complex results in the synthesis of a nine-nucleotide-long $\alpha^{32}\text{P}$ -UTP-labeled RNA product, hereafter called RNA-U (Kulbachinskiy et al., 2004). As shown in Figure 4A, lanes 1 and 2, Gp2 inhibits the synthesis of RNA-U from the $E\sigma^{70}$ -MS probe complex by preventing $E\sigma^{70}$ from binding to the MS probe (see also Figure S3B, lanes 1–4). In contrast, and as expected, the addition of Gp2 to the preformed $E\sigma^{70}$ -MS probe complex had relatively little effect on the amount of RNA-U synthesized (Figure 4A, lane 3). Under identical conditions, when $E\sigma^{70}_{\Delta R1.1}$ was used, no inhibition of RNA-U synthesis by Gp2 was observed (Figure 4A, lanes 4–6). The *in trans* addition of the isolated R1.1 domain to $E\sigma^{70}_{\Delta R1.1}$ conferred a significant degree of Gp2 sensitivity upon the $E\sigma^{70}_{\Delta R1.1}$ -dependent transcription from the MS probe (Figure 4B, lane 4). Further, in the absence of σ^{70} , the catalytic activity of core RNAP on the MS probe was unaffected by Gp2 (Fig-

ure S3C). In summary, the results obtained with the MS probe corroborate the view that full inhibition of $E\sigma^{70}$ RPo formation by Gp2 depends on σ^{70} (specifically the R1.1 domain) but occurs independently of the consensus promoter DNA elements.

Inhibition of RPo Formation by Gp2 Involves a Long-Range, R1.1-Dependent, Antagonistic Effect on $E\sigma^{70}$ -Promoter Interactions

In the RPC, the promoter DNA does not interact with the dwDBC, and consistent with previous results, RPo formation is not inhibited by Gp2 (Cámara et al., 2010; Mekler et al., 2011b). We next conducted experiments to investigate whether the binding of Gp2 to the β' jaw domain influences RNAP-promoter interaction outside of the dwDBC during RPo formation, and determine what role (if any) R1.1 plays in this process. We conducted electrophoretic gel mobility shift assays (EMSAs) to determine whether $E\sigma^{70}$ binding to shortened versions of the σ^{70} -dependent *lacUV5*, λP_{R3} , and T7A1 promoter probes truncated at position -7 of both strands (the -7/-7 probes) is inhibited by Gp2. Note that the -7/-7 probes contain the -35 and -10 consensus promoter DNA elements recognized by σ^{70} regions 4.2 and 2.4, respectively, and lack the dwDNA segment. The results reveal that the binding of $E\sigma^{70}$ to the -7/-7 probes is not inhibited by Gp2 (Figure 5A, lanes 3 and 4, and Figure S4A), and are thus consistent with the view that interactions between the RNAP and the promoter in the RPC are not affected by Gp2. In contrast, under identical conditions, the binding of $E\sigma^{70}$ to the corresponding +20/+20 probes is abolished by Gp2, as expected (Figure 5A, lanes 1 and 2, and Figure S4A).

We next determined the minimum length of the promoter template at which inhibition by Gp2 starts to occur. To that end, we extended the *lacUV5* -7/-7 probe in one-basepair increments and monitored the Gp2 sensitivity of complex formation with these probes by EMSA. The results show that binding of $E\sigma^{70}$ to the -2/-2 probe is reduced by 50% in the presence of Gp2, and binding of $E\sigma^{70}$ to the -1/-1 and +1/+1 probes is inhibited by 85% and 100%, respectively (Figure 5B and Figure S4B). Because the site of Gp2 binding, the β' jaw, is located downstream of the +1 position, the inhibitory effect of Gp2 evidently extends beyond the inhibition of dwDNA interactions with the β' jaw during RPo formation (see above). In support of the above view, whereas inhibition of $E\sigma^{70}$ binding to the +20/+20 DNA depends on the order of addition, Gp2 inhibited the binding of $E\sigma^{70}$ to the +1/+1 probe independently of the order of addition (Figure 5C; compare lanes 1–3 and 4–6). The order of addition-independent inhibitory effect of Gp2 on the binding of $E\sigma^{70}$ to the +1/+1 probe is specific, because the binding of $E\sigma^{70}$ to the +1/+1 probe is not inhibited by an RNAP-binding mutant of Gp2 (R56E) (Cámara et al., 2010) (Figure S4C). The binding of Gp2 to the β' jaw per se was not the cause of inhibition of $E\sigma^{70}$ binding to the +1/+1 probe, because $E\sigma^{70}$ bound the +1/+1 probe in the presence of a Gp2 mutant (Mut7) that binds RNAP normally but is functionally attenuated (Sheppard et al., 2011) (Figure S4C). Furthermore, the inhibitory effect of Gp2 was markedly reduced when RNAP containing σ^{70} lacking R1.1 was used to bind the +1/+1 probe (Figure 5D, lanes 3 and 4) and, as expected, was partially restored when the isolated R1.1 domain was added in *trans* to

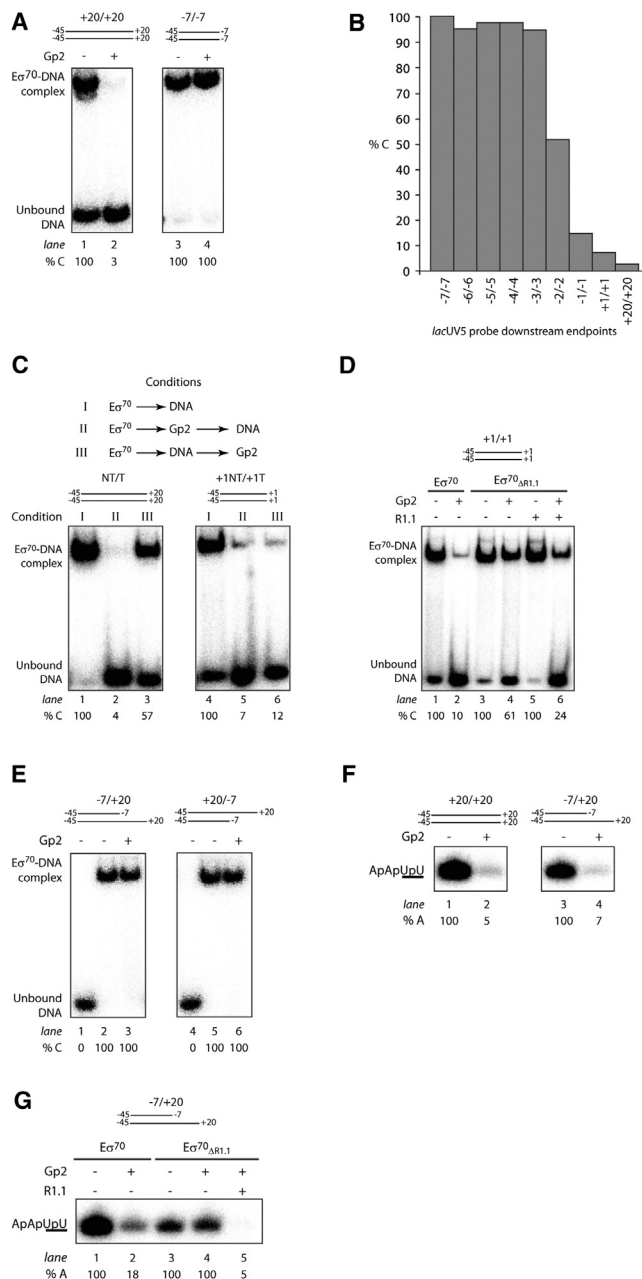


Figure 5. Inhibition of RPo Formation by Gp2 Involves Long-Range Antagonistic Effects on $E\sigma^{70}$ -Promoter Interactions

(A) Autoradiograph of a 4% (v/v) native polyacrylamide gel comparing binding of $E\sigma^{70}$ to the +20/+20 (lanes 1 and 2) and -7/-7 (lanes 3 and 4) probes in the presence (lanes 2 and 4) and absence (lanes 1 and 3) of Gp2.

(B) Graph showing the percentage of DNA bound by $E\sigma^{70}$ in the presence of Gp2 compared with reactions with no Gp2. The dsDNA probes with different downstream end points are indicated in the x axis of the graph.

(C) Autoradiograph of 4% (v/v) native polyacrylamide gels comparing binding of $E\sigma^{70}$ with the +20/+20 (lanes 1–3) and +1/+1 (lanes 4–6) probes in the absence (condition I) and presence (conditions II and III) of Gp2.

(D) Autoradiograph of a 4% (v/v) native polyacrylamide gel comparing binding of $E\sigma^{70}$ (lanes 1 and 2) and $E\sigma^{70}\Delta R1.1$ (lanes 3–6) in the presence (lanes 2, 4, and 6) and absence (lanes 1, 3, and 5) of Gp2 to the +1/+1 probe. In lanes 5 and 6, the isolated R1.1 domain is present in *trans*.

this reaction (Figure 5D, lanes 5 and 6). Overall, we conclude that inhibition of the RPo formation by Gp2 also involves a long-range, R1.1-dependent antagonistic effect on $E\sigma^{70}$ interactions with DNA around the RNAP active center.

Gp2 Interferes with the Promoter DNA Template Strand Accessing the RNAP Active Center

We next wanted to determine the effect of Gp2 on $E\sigma^{70}$ binding to variants of the -7/-7 *lacUV5* probe with either template or nontemplate single-strand downstream extensions to position +20 (-7/+20 or +20/-7 promoter probes, respectively; recall that RPo formation on the +20/+20 probe is efficiently inhibited by Gp2). The EMSA results reveal that Gp2 had no detectable effect on the binding of $E\sigma^{70}$ to either of these probes (Figure 5E). To determine whether the RNAP active site can access the +1 position on the template strand of the -7/+20 probe in the presence of Gp2, we performed transcription-initiation assays. The results, shown in Figure 5F, indicate that even though Gp2 does not inhibit the binding of $E\sigma^{70}$ to this probe (Figure 5E, lanes 1–3), the synthesis of the ApApUpU transcript is effectively abolished in the presence of Gp2. The inhibitory effect of Gp2 on transcription from the -7/+20 probe is R1.1 dependent: ApApUpU transcript synthesis by $E\sigma^{70}\Delta R1.1$ is not inhibited by Gp2 (Figure 5G, lanes 3 and 4), whereas the *in trans* addition of the isolated R1.1 domain to $E\sigma^{70}\Delta R1.1$ -containing reactions abolishes ApApUpU synthesis (Figure 5G, lane 5). Therefore, we conclude that the RNAP active center cannot productively access the template strand of the -7/+20 promoter probe when Gp2 is bound to the β' jaw domain. This conclusion is consistent with the view that the binding of Gp2 to the β' jaw has a long-range, R1.1-mediated antagonistic effect on $E\sigma^{70}$ -promoter interactions near the RNAP active center. In other words, it seems that the binding of Gp2 to the β' jaw in $E\sigma^{70}$ restricts single-stranded DNA from accessing the RNAP active center in an R1.1-dependent manner.

Gp2 Appropriates R1.1 to Efficiently Inhibit RPo Formation by $E\sigma^{70}$

Previously, Mekler et al. (2002) showed that FRET can be used to monitor the displacement of R1.1 from near the dwDBC in free

(E) Autoradiograph of 4% (v/v) native polyacrylamide gels comparing binding of $E\sigma^{70}$ to promoter probes with either the nontemplate or template strand ending at the -7 position (-7/+20 and +20/-7) in the absence (lanes 2 and 5) and presence (lanes 3 and 6) of Gp2. In (A–E), the percentage of DNA bound by $E\sigma^{70}$ or $E\sigma^{70}\Delta R1.1$ (% C) in the reactions with Gp2 with respect to reactions with no Gp2 is given at the bottom of the gels. The data obtained in at least two independent experiments fell within 3% of the % C value shown.

(F) Autoradiographs of 20% (v/v) denaturing urea gels showing the synthesis of the ApApUpU transcript (underlined nucleotides are $\alpha^{32}P$ labeled) from the *lacUV5* promoter +20/+20 and -7/+20 probes by $E\sigma^{70}$ in the absence (lanes 1 and 3) and presence (lanes 2 and 4) of Gp2.

(G) As in F, but using the -7/+20 probe comparing the activity of $E\sigma^{70}$ and $E\sigma^{70}\Delta R1.1$ in the absence (lanes 1 and 3) and presence (lanes 2, 4, and 5) of Gp2. In lane 5, the isolated R1.1 domain is present in *trans* (at ~8-fold molar excess over $\sigma^{70}\Delta R1.1$). In (F) and (G), the percentage of ApApUpU transcript synthesized (% A) in the reactions with Gp2 with respect to reactions with no Gp2 is given at the bottom of the gel for each reaction. The data obtained in at least two independent experiments fell within 5% of the % A value shown. See also Figure S4.

$E\sigma^{70}$ to the tip of the β pincer in the RPo (see Introduction). Because Gp2 efficiently inhibits transcription initiation in an R1.1-dependent manner, we wanted to determine whether Gp2 antagonizes the obligatory displacement of R1.1 during RPo formation. Initially, we calculated the distance between R1.1 and the RNAP active center by measuring FRET between a fluorescein probe incorporated at amino acid position 36 in σ^{70} R1.1 (hereafter called σ^{70*}) and rifampicin (Rif), an antibiotic that binds RNAP between the upstream and downstream lobes of the β subunit and effectively quenches the fluorescence in $E\sigma^{70*}$ (Figure S5A and Figure 6A; compare lines labeled $E\sigma^{70*}$ and $[E\sigma^{70*}\text{-Rif}]$). As expected, the quenching efficiency is much lower in the RPo formed on the N25cons promoter than with $E\sigma^{70*}$ (Figure 6A; compare lines labeled $E\sigma^{70*}$ and $[E\sigma^{70*}\text{-Rif}]$ with lines labeled $E\sigma^{70*}\text{-N25cons}$ and $[E\sigma^{70*}\text{-Rif}]\text{-N25cons}$, respectively), thus indicating displacement of R1.1 from near the dwDBC upon RPo formation. In control experiments in which Rif was replaced with the colorless RNAP inhibitor sorangicin-A (Sor), which binds RNAP in the Rif-binding site (Campbell et al., 2005), no influence on the fluorescence of the fluorescein probe attached to position 36 in R1.1 was detected (Figure S5B; compare lines $E\sigma^{70*}$ and $[E\sigma^{70*}\text{-Sor}]$). Further, the addition of Rif to the preformed $E\sigma^{70*}\text{-Sor}$ complex also resulted in a negligible decrease in fluorescence intensity (Figure S5B; compare lines $[E\sigma^{70*}\text{-Sor}]$ and $[E\sigma^{70*}\text{-Sor}]\text{-Rif}$). Thus, the Rif-mediated decrease in fluorescence intensity of the fluorescein probe attached to amino acid position 36 in R1.1 is specific to the binding of Rif and is a consequence of quenching via the FRET mechanism. The addition of Gp2 caused an $\sim 13\%$ decrease in the fluorescence intensity of $E\sigma^{70*}$ (Figure 6B; compare lines labeled $E\sigma^{70*}$ and $E\sigma^{70*}\text{-Gp2}$), suggesting that the binding of Gp2 to the β' jaw domain changes the microenvironment of R1.1 and/or its positions with respect to the RNAP active center. These effects result from specific Gp2 binding to the β' jaw domain, because no change in fluorescence spectra was observed in the presence of the R56E Gp2 mutant (Sheppard et al., 2011), which does not bind to the RNAP (compare Figure S5C and Figure 6B). The calculated distances between the fluorescein probe attached to amino acid position 36 in R1.1 and Rif are 41 Å and 65 Å in $E\sigma^{70*}$ and RPo, respectively (Figure 6C). However, in the $[E\sigma^{70*}\text{-Rif}]\text{-Gp2}$ complex, this distance is longer than in the $[E\sigma^{70*}\text{-Rif}]$ complex by 9 Å (Figure 6C). When the N25cons probe is added to the $[E\sigma^{70*}\text{-Rif}]\text{-Gp2}$ complex, the calculated distance between fluorescein at amino acid position 36 in R1.1 and Rif is 53 Å (Figure 6B, compare lines labeled $[E\sigma^{70*}\text{-Rif}]\text{-Gp2}$ and $[E\sigma^{70*}\text{-Rif}]\text{-Gp2}\text{-N25cons}$, and Figure 6C). In contrast, as mentioned above, when the N25cons probe is added to the $[E\sigma^{70*}\text{-Rif}]$ complex (i.e., in the absence of Gp2), the corresponding calculated distance is 65 Å (Figure 6A, compare lines labeled $[E\sigma^{70*}\text{-Rif}]$ and $[E\sigma^{70*}\text{-Rif}]\text{-N25cons}$, and Figure 6C). Thus, Gp2 reduces the distance of R1.1 displacement during RPo formation by 12 Å. This effect is specific, because no change in fluorescence spectra was observed in control experiments with R56E Gp2 mutant (compare Figure S5C and Figure 6B). Because $E\sigma^{70}$ forms a complex with the N25cons promoter, which contains Gp2 and bears some hallmarks of the RPo formed in the absence of Gp2 (see Introduction), our results indicate that ternary

($E\sigma^{70}$ -Gp2-N25cons promoter) complex formation is not accompanied by the characteristic long-distance displacement of R1.1 that normally occurs during RPo formation (Figures 6A and 6C). To avoid a possible error related to uncertainty of the Rif-fluorescein distance determination (Knight et al., 2005), we performed control experiments conducted with fluorescein probe attached to a different position in R1.1 (amino acid 59), which further corroborated our conclusion (Figure 6C and Figure S5E). When the N25cons probe is added to the $[E\sigma^{70}\text{-Rif}]\text{-Gp2}$ complex, the calculated distance between the fluorescein at amino acid position 59 in R1.1 and Rif is 40 Å (Figure S5F, compare lines labeled $[E\sigma^{70}\text{-Rif}]\text{-Gp2}$ and $[E\sigma^{70}\text{-Rif}]\text{-Gp2}\text{-N25cons}$, and Figure 6C). In contrast, when the N25cons probe is added to the $[E\sigma^{70}\text{-Rif}]$ complex (i.e., in the absence of Gp2), the corresponding calculated distance is 61 Å (Figure S5E, compare lines labeled $[E\sigma^{70}\text{-Rif}]$ and $[E\sigma^{70}\text{-Rif}]\text{-N25cons}$, and Figure 6C). Thus, Gp2 appropriates R1.1 to efficiently inhibit RPo formation by $E\sigma^{70}$.

DISCUSSION

The interaction between dwDNA and dwDBC in the RNAP is important for establishing the RPo for transcription initiation at bacterial promoters. The *Ec* RNAP β' jaw domain, a pivotal feature of the dwDBC, is a multifunctional domain whose role in transcription extends beyond RPo formation (Ederth et al., 2006). In previous studies, Ederth et al., (2006) and our group (Wigneshweraraj et al., 2006) showed that the β' jaw domain is involved in regulatory interplay with other parts of the RNAP that extend to the distally located nascent RNA-binding site. In vitro, deletion of the β' jaw domain in the *Ec* RNAP destabilizes RPo, suppresses transcriptional pausing, increases the overall elongation rate, and decreases intrinsic termination (Ederth et al., 2002, 2006). Here, we have demonstrated that the β' jaw domain makes sequence-nonspecific contacts with dsDNA. *Ec* RNAP mutants with the G1161R mutation or the deletion of residues 1149–1190 in the β' jaw domain form RPo with significantly reduced half-lives compared with RPo formed by the wild-type RNAP. Our data provide an explanation for the observed phenotypes of the mutant RNAP in a structural context: (1) the β' jaw domain residues that undergo significant chemical shift changes upon interaction with dsDNA (T1169, R1174, and M1180) are all located within the deleted region (i.e., 1149–1190), and (2) the invariant G1161 residue (Figure S1A) is located in the $\beta 1$ sheet of the β' jaw domain structure and is facing away from the DNA-binding surface toward the β' insertion 6 domain; thus, an arginine side chain at this position could compromise the overall structural integrity of the dwDBC. During T7 phage infection of *Ec*, the sequence-nonspecific interaction between the β' jaw domain and dsDNA is subject to regulation by the small, non-DNA-binding T7 transcription factor Gp2, which inhibits RPo formation by the host RNAP. One strategy used by Gp2 to inhibit RPo formation by the host RNAP is to sterically occlude dsDNA from binding to the β' jaw domain by competing with dsDNA for overlapping interaction surfaces on the β' jaw domain.

The narrow width of the dwDBC observed in structures of the bacterial RNAP is thought to constrain the entry of DNA into the catalytic cleft of RNAP for RPo formation. An obligatory step

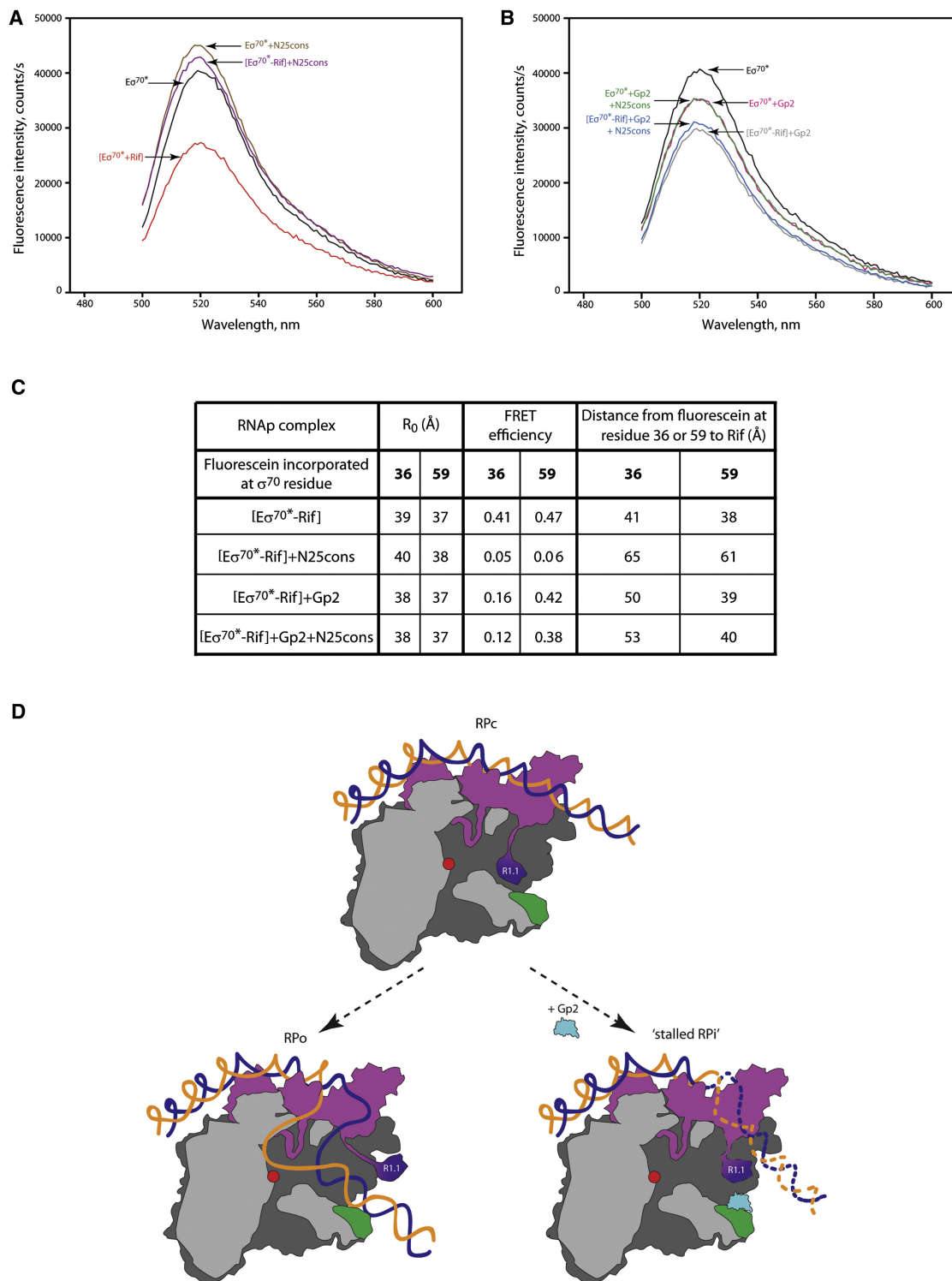


Figure 6. Gp2 Appropriates R1.1 to Efficiently Inhibit RPo Formation by E σ^{70}

(A and B) Measurement of FRET between fluorescein incorporated into σ^{70} at position 36 (σ^{70*}) and Rif during RPo formation in the absence (A) and presence (B) of Gp2. The fluorescence emission spectra are recorded with 482 nm excitation.

(C) The FRET efficiency values and distance calculations are tabulated (see also Figure S5) and the values presented are averages obtained from two to three individual experiments; the estimated error in R_0 is $\sim 10\%$.

(D) Cartoon (as in Figure 1A) depicting the mechanism by which RPo formation is inhibited by Gp2 at σ^{70} -dependent promoters.

during RPo formation at σ^{70} -dependent promoters is the displacement of the R1.1 domain of σ^{70} from near the dwDBC to the tip of the β pincer, which occurs concomitantly with the loading of DNA in the DBC/dwDBC. The negatively charged R1.1 domain is believed to act as a molecular placeholder for dsDNA in $E\sigma^{70}$; therefore, the displacement of R1.1 would facilitate the loading of dwDNA into the dwDBC and subsequently stabilize the formation of RPo (Mekler et al., 2002). Thus, it was proposed that R1.1 can facilitate DNA entry into the dwDBC by holding the β and β' pincers open so that the promoter DNA can enter the dwDBC and the template promoter strand can access the RNAP active center (Saecker et al., 2011). Our results demonstrate that R1.1 of σ^{70} has an important functional role in the mechanism by which Gp2 inhibits RPo formation at σ^{70} -dependent promoters. Consistent with this view, Gp2 inhibits RPo formation by RNAP containing alternative σ factors such as σ^{38} or σ^{54} less efficiently or not detectably, respectively (see above) (Wigneshweraraj et al., 2004). We envisage a model in which Gp2 repositions R1.1 (see below) and/or stabilizes R1.1 so that it can no longer be displaced to the tip of the β pincer. This would allow RPo to form and thus stall the $E\sigma^{70}$ -promoter complex in an intermediate state (intermediate promoter complex [RPI] in Figure 6D) at σ^{70} -dependent promoters. Furthermore, the extended negatively charged patch in the dwDBC that results upon binding of Gp2 to the β' jaw domain could thus help electrostatically reposition R1.1 and/or mimic the presence of R1.1 in the dwDBC. In support of the former view, the removal of the negatively charged side chains of the amino acids in Gp2 that contribute to the NCS does not affect the affinity of Gp2 to the RNAP or its overall structural integrity, but it does severely decrease its ability to inhibit RPo formation (Sheppard et al., 2011). Intriguingly, the R1.1-dependent inhibitory effect of Gp2 on the binding of dwDNA to the β' jaw domain extends beyond the dwDBC and has long-range antagonistic effects on $E\sigma^{70}$ -promoter interactions that extend up to the RNAP active center. As a consequence, $E\sigma^{70}$ -promoter DNA interactions become stalled en route to the RPo (Figure 6D). Because not even single-stranded DNA can access the active center of the RNAP in the presence of Gp2, it seems that Gp2 when bound to the β' jaw domain restricts the conformational flexibility and changes in the RNAP that normally accompany RPo formation. Because the interaction between Gp2 and the isolated β' jaw domain fragment is several orders of magnitude weaker than that between Gp2 and the whole enzyme, we cannot exclude the possibility that Gp2 undergoes positional rearrangements and makes additional contacts with the other domains of the RNAP that surround the dwDBC, notably the β' insertion 6 and/or the β downstream lobe domains. It is therefore likely that the solution structure of the Gp2- β' jaw fragment likely reports on an early encounter complex between Gp2 and RNAP.

The biological role of Gp2 is to inhibit the transcription of early T7 genes from strong σ^{70} -dependent promoters present on the T7 genome by the host RNAP (Savalia et al., 2010). The absence of Gp2 results in unsuccessful infection because the antiterminated host RNAP moves into regions of the T7 genome that are normally transcribed by the T7 RNAP. The interference of the fast-moving T7 RNAP, which is responsible for the transcrip-

tion of middle and late T7 genes, by the slow-moving host RNAP results in aberrant packaging of concatemeric viral DNA into virion heads, and thus unsuccessful infection. The two-pronged strategy used by Gp2 to inhibit the host RNAP, through occlusion of dwDNA from binding to the dwDBC and appropriation of a σ^{70} -specific domain, leads to highly efficient inhibition of $E\sigma^{70}$ -dependent transcription from very strong early T7 promoters to ensure successful infection. The inhibition of host RNAP by Gp2 thus defines an as yet uncharacterized mechanism by which bacterial transcription is regulated by a nonbacterial factor.

EXPERIMENTAL PROCEDURES

Proteins and Promoter Templates

Details of the proteins and promoter templates used in this study are provided in the Supplemental Experimental Procedures.

NMR Spectroscopy and Structure Calculation

Details about the NMR solution structure calculation of the Gp2- β' jaw fragment complex are provided in the Supplemental Experimental Procedures.

In Vitro Transcription Assays

In vitro transcription assays were conducted essentially as previously described (Cámara et al., 2010). Reactions (10 μ l) were conducted using final concentrations of 100 nM $E\sigma^{70}$, 20 nM unlabeled promoter DNA probes, 0.5 mM dinucleotide primer ApA, 100 μ g/ml heparin, and 3 μ Ci of [α - 32 P]-UTP (for *lacUV5*) or [α - 32 P]-ATP and 0.5 μ M CTP (for *osmE*). Unless otherwise indicated, Gp2 and $E\sigma^{70}$ (at a 2:1 molar ratio) were always preincubated before the promoter DNA was added to the reaction. When present, R1.1 was always preincubated with $E\sigma^{70}_{\Delta R1.1}$ (at the indicated amounts) before addition of Gp2 and/or promoter DNA to the reaction. The reactions were resolved on a 20% (w/v) urea-denaturing polyacrylamide gel. The dried gel was visualized and quantified with the use of an FLA-5000 PhosphorImager.

Native Gel Mobility Assays

All native mobility shift assays were conducted essentially as described previously (Cámara et al., 2010). Binding reactions (10 μ l) were set up as described above.

FRET Assays

Fluorescence emission spectra of $E\sigma^{70}$ reconstituted with σ^{70} subunit labeled at position 36 or 59 with fluorescein (σ^{70}_F) were recorded with 482 excitation as previously described (Knight et al., 2005). When present, Rif and Sor (at 1 μ M) were incubated with $E\sigma^{70}_F$ for 10 min at 37°C. The $E\sigma^{70}_F$ -Gp2 complex was obtained by incubation of 1 nM $E\sigma^{70}_F$ and 200 nM Gp2 for 10 min at 37°C. Promoter complexes were obtained by incubation of 1 nM $E\sigma^{70}$ or $E\sigma^{70}$ -Gp2 with 5 nM N25cons for 15 min at 37°C. The FRET efficiency, critical FRET radius (R_0), and distance between fluorescein and Rif were determined as previously described (Knight et al., 2005). A previous analysis of the accuracy of similar FRET-based measurements of distances between $E\sigma^{70}$ -bound Rif and fluorescein incorporated at different positions in σ^{70} indicated that the uncertainty of such a distance determination is 11%–25%, with a mean of 15% (Knight et al., 2005).

ACCESSION NUMBERS

The coordinates and structure factors for the Gp2- β' jaw fragment complex have been deposited in the Protein Data Bank (accession code 2LMC).

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and Supplemental Experimental Procedures and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2012.06.013>.

ACKNOWLEDGMENTS

This work was supported by grants from the Biotechnology and Biological Sciences Research Council and the Wellcome Trust to S.W. and S.M., and from the National Institutes of Health (GM59295 and GM64530) and the Presidium Molecular and Cellular Biology Program of the Russian Academy of Sciences to K.S. S.W. received a David Phillips Fellowship from the Biotechnology and Biological Sciences Research Council (BB/E023703). We thank Dr. Debbie Hinton for providing the plasmids encoding the isolated domain of R1.1 and $\sigma^{70}_{\Delta R1.1}$.

Received: November 22, 2011

Revised: January 4, 2012

Accepted: June 12, 2012

Published online: July 19, 2012

REFERENCES

- Bordes, P., Repoila, F., Kolb, A., and Gutierrez, C. (2000). Involvement of differential efficiency of transcription by σ^{70} and σ^{32} RNA polymerase holoenzymes in growth phase regulation of the *Escherichia coli* *osmE* promoter. *Mol. Microbiol.* 35, 845–853.
- Cámara, B., Liu, M., Reynolds, J., Shadrin, A., Liu, B., Kwok, K., Simpson, P., Weinzierl, R., Severinov, K., Cota, E., et al. (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. USA* 107, 2247–2252.
- Campbell, E.A., Pavlova, O., Zenkin, N., Leon, F., Irschik, H., Jansen, R., Severinov, K., and Darst, S.A. (2005). Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *EMBO J.* 24, 674–682.
- Ederth, J., Artsimovitch, I., Isaksson, L.A., and 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., and 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.
- Gruber, T.M., and Gross, C.A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* 57, 441–466.
- Haugen, S.P., Ross, W., and 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.
- Knight, J.L., Mekler, V., Mukhopadhyay, J., Ebright, R.H., and Levy, R.M. (2005). Distance-restrained docking of rifampicin and rifamycin SV to RNA polymerase using systematic FRET measurements: developing benchmarks of model quality and reliability. *Biophys. J.* 88, 925–938.
- Kulbachinskiy, A., Feklistov, A., Krashennikov, I., Goldfarb, A., and Nikiforov, V. (2004). Aptamers to *Escherichia coli* core RNA polymerase that sense its interaction with rifampicin, sigma-subunit and GreB. *Eur. J. Biochem.* 271, 4921–4931.
- Mekler, V., Kortkhonja, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R., and 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., and Severinov, K. (2011a). A critical role of downstream RNA polymerase-promoter interactions in the formation of initiation complex. *J. Biol. Chem.* 286, 22600–22608.
- Mekler, V., Minakhin, L., Sheppard, C., Wigneshweraraj, S., and Severinov, K. (2011b). Molecular mechanism of transcription inhibition by phage T7 gp2 protein. *J. Mol. Biol.* 413, 1016–1027.
- Murakami, K.S., and Darst, S.A. (2003). Bacterial RNA polymerases: the whole story. *Curr. Opin. Struct. Biol.* 13, 31–39.
- Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002a). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* 296, 1285–1290.
- Murakami, K.S., Masuda, S., and Darst, S.A. (2002b). Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* 296, 1280–1284.
- Nechaev, S., and Severinov, K. (1999). Inhibition of *Escherichia coli* RNA polymerase by bacteriophage T7 gene 2 protein. *J. Mol. Biol.* 289, 815–826.
- Nechaev, S., and Severinov, K. (2003). Bacteriophage-induced modifications of host RNA polymerase. *Annu. Rev. Microbiol.* 57, 301–322.
- 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, 8.
- Saecker, R.M., Record, M.T., Jr., and 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., and Severinov, K. (2010). The role of the T7 Gp2 inhibitor of host RNA polymerase in phage development. *J. Mol. Biol.* 402, 118–126.
- Schwartz, E.C., Shekhtman, A., Dutta, K., Pratt, M.R., Cowburn, D., Darst, S., and Muir, T.W. (2008). A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. *Chem. Biol.* 15, 1091–1103.
- Sheppard, C., Cámara, B., Shadrin, A., Akulenko, N., Liu, M., Baldwin, G., Severinov, K., Cota, E., Matthews, S., and Wigneshweraraj, S.R. (2011). Inhibition of *Escherichia coli* RNAP by T7 Gp2 protein: role of negatively charged strip of amino acid residues in Gp2. *J. Mol. Biol.* 407, 623–632.
- Vassilyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassilyeva, M.N., Borukhov, S., and Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417, 712–719.
- Wigneshweraraj, S.R., Burrows, P.C., Nechaev, S., Zenkin, N., Severinov, K., and Buck, M. (2004). Regulated communication between the upstream face of RNA polymerase and the beta' subunit jaw domain. *EMBO J.* 23, 4264–4274.
- Wigneshweraraj, S.R., Savalia, D., Severinov, K., and Buck, M. (2006). Interplay between the beta' clamp and the beta' jaw domains during DNA opening by the bacterial RNA polymerase at sigma54-dependent promoters. *J. Mol. Biol.* 359, 1182–1195.

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

Correspondence

Sivaramesh Wigneshweraraj
s.r.wig@imperial.ac.uk

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' jaw domain, which is part of the DNA binding channel. The binding of Gp2 to the β' 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' jaw, amino acid residues in the $\beta 3$ strand of Gp2 contribute to the primary interaction interface with the β' jaw. The 7009 *E. coli* strain is resistant to T7 because it carries a charge reversal point mutation in the β' jaw that prevents Gp2 binding. However, a T7 phage encoding a mutant form of Gp2, called Gp2 ^{β} , 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 ^{β} , 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' jaw is compromised. The proposed additional interaction interface between RNAP and Gp2 may contribute to the multipronged mechanism of transcription inhibition by Gp2.

Received 26 July 2012

Revised 7 September 2012

Accepted 11 September 2012

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 (σ) factor subunit that determines the specificity of the RNAP-promoter interaction. In *Escherichia coli*, the σ^{70} -containing RNAP ($E\sigma^{70}$) is responsible for the transcription of most genes during exponential growth (Haugen *et al.*,

2008). Transcription by $E\sigma^{70}$ begins with the formation of the initial $E\sigma^{70}$ -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, ~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' 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.

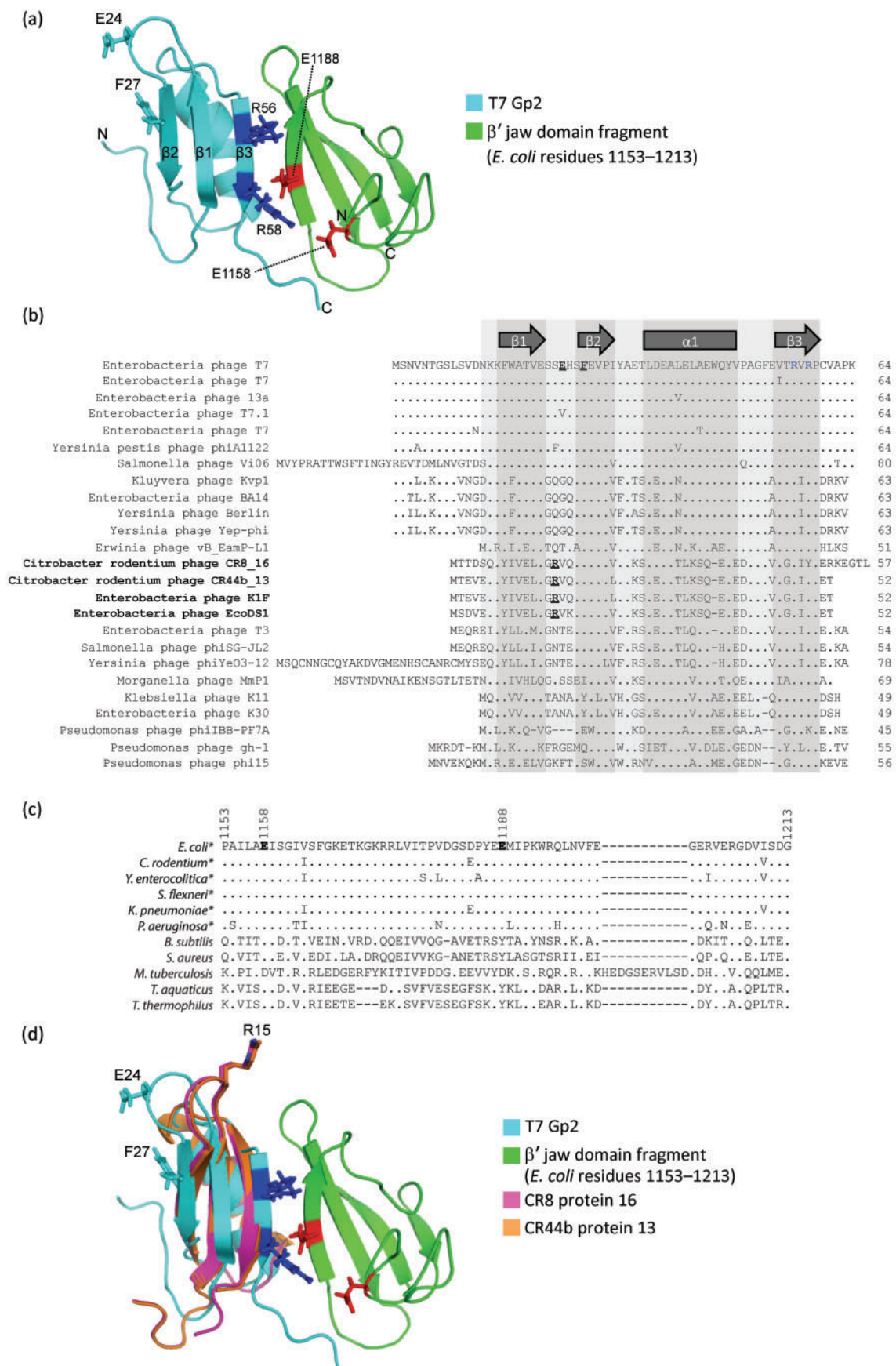


Fig. 1. Sequence alignment analysis of T7 Gp2-like proteins. (a) Ribbon representation of the T7 Gp2–*E. coli* RNAP β' jaw domain fragment (residues 1153–1213) complex. Gp2 is shown in cyan and the β' jaw domain fragment is shown in green. The positions of the amino acids E24, F27 and R56 in T7 Gp2 are indicated along with the E1158 and E1188 residues in the β' jaw domain. (b) Alignment of amino acid sequences from known Gp2-like proteins. The sequences are displayed in single-letter amino acid code, with the length of the sequence indicated on the far right. The localization of the β -strands and α -helix based on the structure of T7 Gp2 is indicated. The highly conserved R56 and R58 residues are highlighted in blue. The positions of amino acids residues E24 and F27 in T7 Gp2 are underlined and in bold type. The arginine residue, corresponding to T7 Gp2 amino acid residue S23 in 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; Cámara *et al.*, 2010).

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 transcription regulatory factors, which repress, stimulate or modulate its ability to form the RPo (Browning & Busby, 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, 2003). T7 phage encodes a 7 kDa protein, called Gp2, which binds to a domain in the β' subunit, called the 'jaw' domain, and potentially inhibits RPo formation by the *E. coli* RNAP (Nechaev & Severinov, 1999). The β' jaw domain is an integral component of the dwDBC and plays a major role during RPo formation and stable maintenance of the RPo (Ederth *et al.*, 2002, 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 the EBI database at the time of writing) located in the β_3 strand of Gp2 are in close proximity to E1158 and E1188 in the β' jaw, and thereby contribute to significant ionic interactions across the interface (Fig. 1a, b) (James *et al.*, 2012). Alanine or charge reversal substitutions at either R56 or R58 prevent Gp2 from binding to the RNAP (Cámara *et al.*, 2010). Conversely, RNAP containing a lysine substitution at either E1158 (BR3 *E. coli*) or E1188 (7009 *E. coli*) is resistant to inhibition by Gp2 (Nechaev & Severinov, 1999). Thus, the primary interaction interface between Gp2 and the RNAP involves R56 and R58 in the β_3 strand of Gp2 and E1158 and E1188 in the β' jaw domain of the RNAP.

Wild-type T7 is unable to productively infect *E. coli* 7009 (Chamberlin, 1974; Nechaev & Severinov, 1999; Schmitt *et al.*, 1987). A mutant phage, called T7 β , that successfully infects the 7009 *E. coli* strain (which has a mutant RNAP containing the E1188K substitution in the β' jaw domain), encodes a triple

mutant form of Gp2 (harbouring substitutions E24K, F27Y and R56C; hereafter called Gp2 $^\beta$) (Chamberlin, 1974; Schmitt *et al.*, 1987). 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 β_1 and β_2 strands in Gp2, i.e. at the opposite side to the β_3 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 $^\beta$ inhibits wild-type, 7009 and BR3 *E. coli* RNAP. The results strongly suggest that the E24K and F27Y mutations facilitate the interaction between Gp2 $^\beta$ and RNAP when the primary interaction interface between Gp2 and RNAP is compromised. Experiments addressing the loop interconnecting β_1 and β_2 strands in Gp2 homologues from two T7-like phages that infect *Citrobacter rodentium* further corroborate our results.

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 *Xba*I and *Hind*III, 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 *Nde*I and *Bam*HI 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

exactly as T7 Gp2 (Cámara *et al.*, 2010). Wild-type and mutant forms of the *E. coli* core RNAP were overexpressed from pVS10 and purified by affinity chromatography exactly as described previously (Cámara *et al.*, 2010). The E1188K and E1158K mutant *E. coli* RNAPs were constructed by using the QuikChange Site-Directed PCR Mutagenesis kit using pVS10 as the template (Belogurov *et al.*, 2007). Purified proteins were then dialysed into storage buffer [10 mM Tris/HCl (pH 8.0), 50 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 50 % (v/v) glycerol] and aliquots were stored at -20°C (short-term) and -80°C (long-term). The sequences of primers used for the cloning and mutagenesis in this study are available from the authors upon request.

Bacterial growth attenuation assays. Seed cultures were grown at 37°C , shaking at 700 r.p.m. for 6–7 h in a THERMOstar (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 multiwell 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 OD_{600} of ~ 0.2 – 0.25 by adding 0.04 % (w/v) arabinose for MG1655, JE1134 and BR3 cultures, and 0.4 % (w/v) arabinose for 7009 cultures. The doubling time (Dt) was calculated from the gradient of the growth curve typically between the OD_{600} corresponding to induction time and OD_{600} 0.7–0.8. The R^2 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 2.5×10^7 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 [$\text{mass (g)}/M_r$ (10 000 Da)], multiplying by Avogadro's constant ($6.022 \times 10^{23} \text{ mol}^{-1}$) and then by dividing the total number of Gp2 molecules by the number of cells in the sample (i.e. 2.5×10^7 cells).

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

Efficiency of plaque formation assays. These were performed exactly as previously described (Cámara *et al.*, 2010; Rokyta *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 β_3 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 Gp2 $^{\beta}$ (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 β_1 and β_2 strands of the proteins (Fig. 1d). Thus, it is possible that Gp2 homologues from CR8, CR44b, K1F and EcoDS1 phages, like T7 Gp2 $^{\beta}$, 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 T7 $^{\beta}$.

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 Gp2 $^{\beta}$ 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

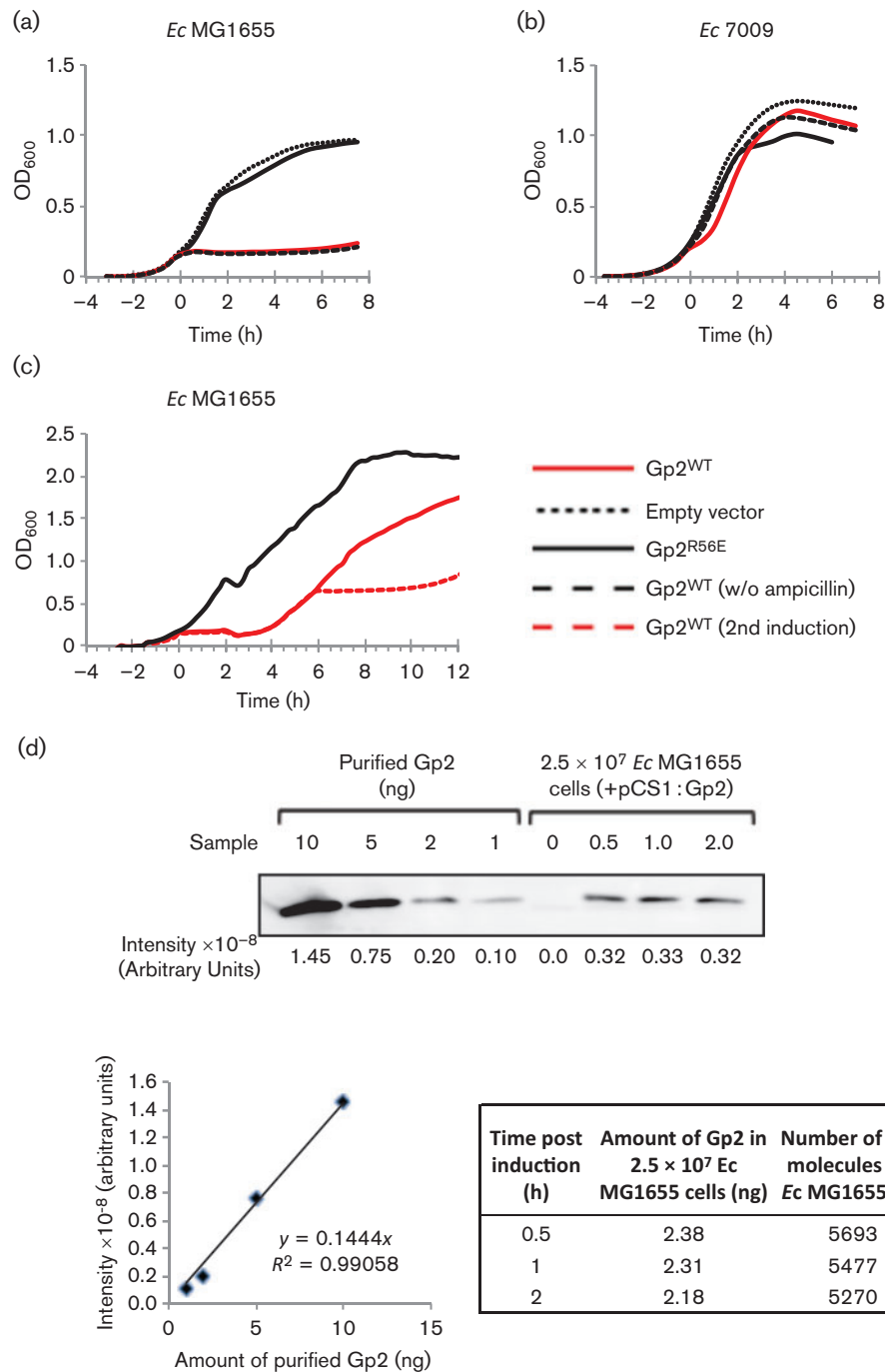


Fig. 2. Establishing a simple *in vivo* assay to measure the activity of Gp2 in the absence of T7 phage infection. Growth curves of (a) *E. coli* MG1655 and (b) 7009 cells expressing Gp2 upon induction with arabinose at $t=0$. The solid red, dotted and solid black lines represent cells harbouring pCS1 : Gp2^{WT}, pBAD18 (empty vector) and pCS1 : Gp2^{R56E}, respectively; the dashed line represents cells harbouring pCS1 : Gp2^{WT} 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 Gp2^β

ND, Not determined.

Gp2	<i>Ec</i> MG1655 Dt* (min) ± SD†	FI‡	<i>Ec</i> 7009 Dt (min) ± SD	FI§	<i>Ec</i> BR3 Dt (min) ± SD	FII
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
Gp2 ^β +R58E	52 ± 1	1.0	59 ± 8	1.1	73 ± 5	0.6

*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.§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:Gp2^{R56E} plasmid, which encodes a functionally defective Gp2 mutant (Cámara *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 Gp2^β to the 7009 RNAP. Therefore, we constructed single, double and triple mutant versions of Gp2 based on Gp2^β 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 (Cámara *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 Gp2^{R56C} 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,

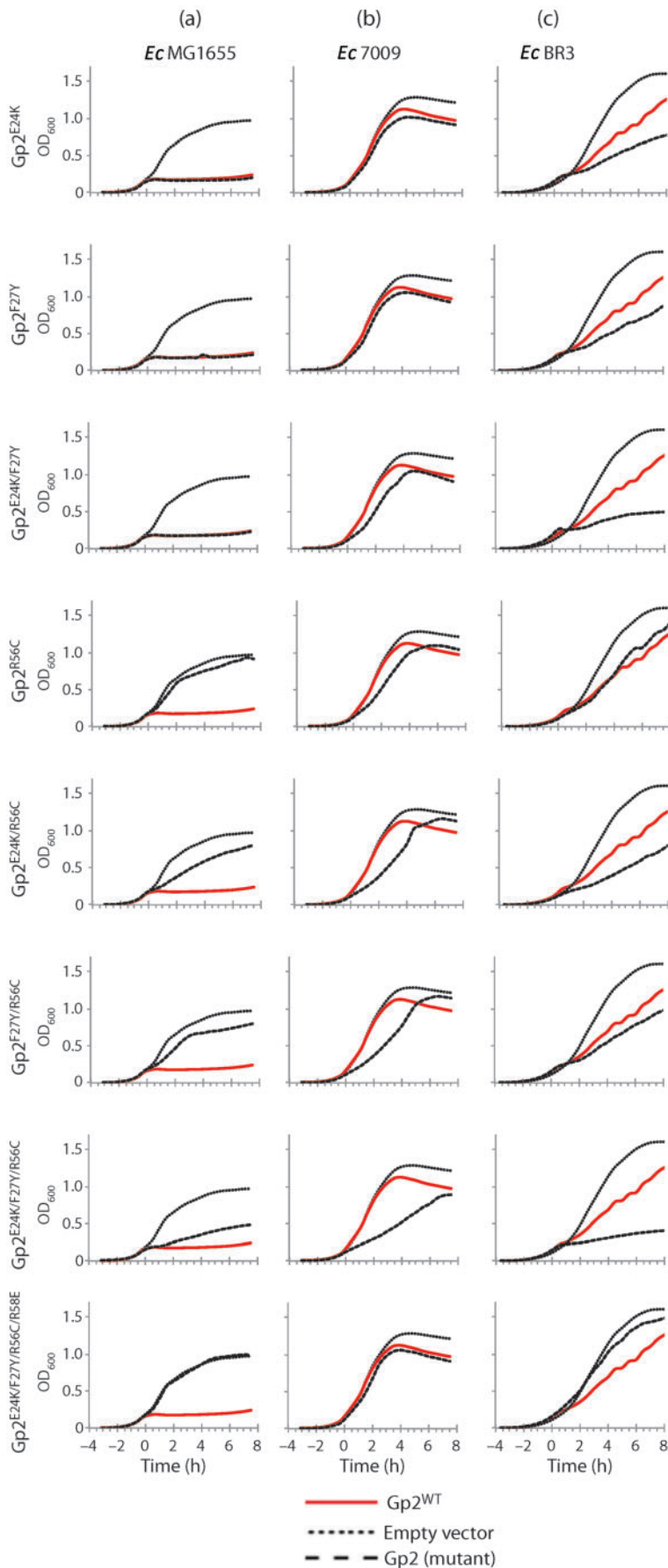


Fig. 3. Ability of single, double and triple mutant versions of Gp2 based on Gp2^β to attenuate bacterial growth upon induction. Growth curves of (a) *E. coli* MG1655, (b) 7009 and (c) BR3 expressing wild-type Gp2 and Gp2 mutants containing various combinations of mutations based on Gp2^β. Gp2 expression was induced with arabinose at $t=0$. In each graph, the solid red, dotted and dashed black lines represent cells harbouring pCS1:Gp2^{WT}, 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^{WT} and pBAD18 (empty vector) were used.

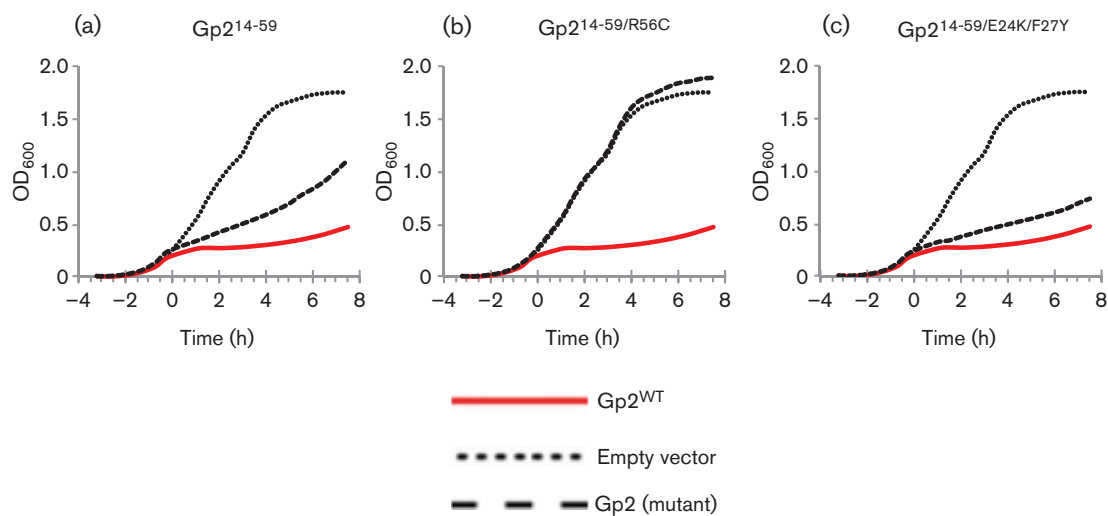


Fig. 4. Amino acids 14–59 of Gp2 are sufficient to inhibit the *E. coli* RNAP. Growth curves of *E. coli* MG1655 cells in which the expression of (a) Gp2^{14–59}, (b) Gp2^{14–59/R56C} and (c) Gp2^{14–59/E24K/F27Y} was induced at *t*=0 (dashed lines). In each graph, the solid red and dotted lines represent cells harbouring pCS1 :Gp2^{WT} and pBAD18 (empty vector), respectively.

Table 1), the effect was highly reproducible. When present together, the E24K and F27Y mutations further improved the ability of Gp2^{R56C} (i.e. Gp2^β) to attenuate wild-type *E. coli* growth (Fig. 3a), and the Dt of cells expressing Gp2^β 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 β1 and β2 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 β' jaw domain. Based on Dt values calculated for MG1655 *E. coli* cells after induction of Gp2^{E24K/R56C}, Gp2^{F27Y/R56C} and Gp2^β (Table 1), it seems that the charge reversal substitution at E24 (i.e. E24K) in the loop interconnecting the β1 and β2 strands in Gp2 rather than the F27Y mutation at the beginning of the β2 strand is more important in enabling the interaction between Gp2 and the RNAP when the primary interaction interface with the β' jaw is compromised by the R56C substitution. However, since a glutamate substitution at position R58 in the context of Gp2^β effectively abolishes the ability of Gp2^β to attenuate the growth of wild-type *E. coli*, it would seem that an interaction interface between the β3 strand and β' jaw domain still persists in the Gp2^β–RNAP complex (Fig. 3a). Experiments with the 7009 and BR3 *E. coli* strains, which encode Gp2-resistant forms of RNAP (see Introduction), further corroborated this view: even though the Dt of 7009 and BR3 *E. coli* cells, when compared with the MG1655 *E. coli* cells, is relatively unaffected by overproduction of wild-type Gp2, Gp2^β increases the Dt of 7009 and BR3 *E. coli* cells by ~2.6- and 2.2-fold, respectively, compared with corresponding cells expressing wild-type Gp2 (Fig. 3b, c, Table 1).

Interestingly, in the context of the BR3 *E. coli* cells, expression of Gp2^{E24K/F27Y} attenuates growth as efficiently as that of Gp2^β (Fig. 3c, Table 1). In both cases, the R58E substitution abolishes the ability of Gp2^β to attenuate cell growth (Fig. 3b, c). Overall, the results strongly suggest that the E24K and F27Y substitutions facilitate an interaction between Gp2 and RNAP when the primary interaction interface in the β3 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}

ND, Not determined.

Gp2	<i>Ec</i> 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 evolutionarily 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 (CR44b₁₃ and CR8₁₆), 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 CR44b₁₃ and CR8₁₆ can complement a T7 mutant phage (T7 Δ am64) harbouring a mutation in *gene 2* (Burck & Miller, 1978). T7 Δ am64 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 T7 Δ am64, 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: $\text{EOP} = (\text{number of T7}\Delta\text{am64 plaques on strain} + \text{test plasmid}) / (\text{number of T7}\Delta\text{am64 plaques on strain} + \text{pSW33:Gp2}^{\text{WT}})$. 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 CR44b₁₃ (pAS33:CR44b₁₃) were transformed into *E. coli* BL21, T7 Δ am64 formed plaques, albeit at 40 % reduced efficiency compared with *E. coli* BL21 cells harbouring pSW33:Gp2^{WT} (EOP of 0.61 ± 0.02). However, T7 Δ am64 was able to form plaques equally well on *E. coli* BL21 cells harbouring a plasmid encoding CR8₁₆ (pAS33:CR8₁₆) (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 CR44b₁₃ and CR8₁₆, like T7 Gp2, bind to the β' jaw domain of the *E. coli* RNAP. As shown in Fig. 5(a), induction of CR44b₁₃ and CR8₁₆ expression with 0.04 % (w/v) arabinose from pAS1:CR44b₁₃ and pAS1:CR8₁₆, 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 CR44b₁₃ and CR8₁₆ 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 CR44b₁₃ or CR8₁₆, 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 CR44b₁₃ and CR8₁₆ proteins to inhibit the synthesis of the tetranucleotide RNA product (ApApUpU) from the *lacUV5* 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 CR44b₁₃ 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 CR44b₁₃ 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 CR44b₁₃ 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 $\beta 3$ strand of Gp2 contribute to the primary interaction interface with the β' jaw domain

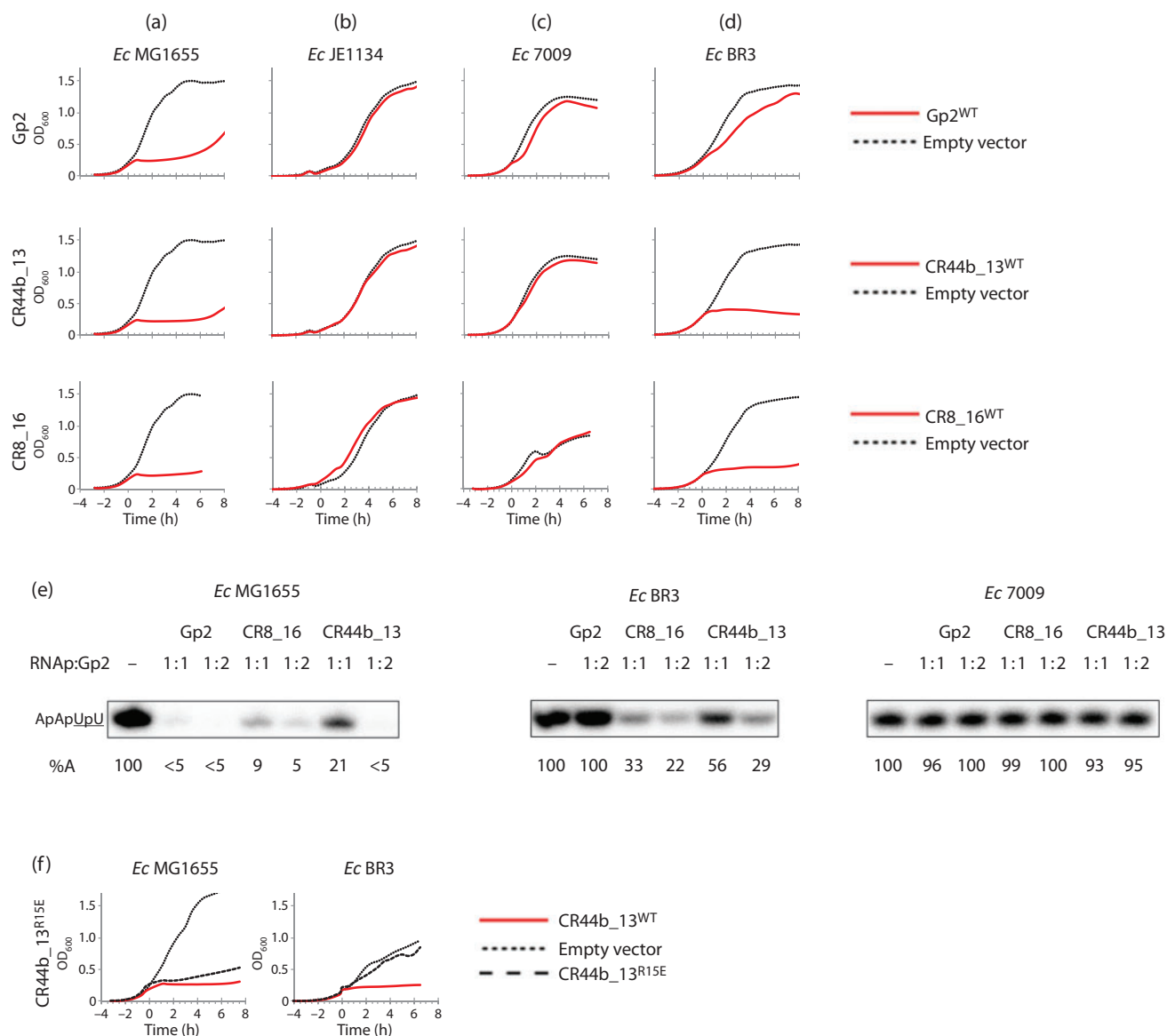


Fig. 5. Experiments with Gp2-like proteins from *C. rodentium* phages CR44b and CR8. Growth curves of (a) *E. coli* MG1655, (b) JE1134, (c) 7009 and (d) BR3 cells in which the expression of T7 Gp2 and Gp2-like proteins CR44b_13 and CR8_16 was induced at $t=0$ (solid red lines). In each graph, the dotted line represents cells harbouring pBAD18 (empty vector). (e) Autoradiograph of a 20% (w/v) denaturing gel showing synthesis of the transcript ApApUpU (the ³²P-labelled nucleotides underlined) from the *lacUV5* promoter by the σ^{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 $t=0$.

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 β_1 and β_2 strands, and, therefore, located on the opposite side to the β_3 strand, can compensate for amino acid substitutions in Gp2 and/or the β' jaw domain that

compromise the primary interaction interface between Gp2 and the RNAP. This seems to constitute the molecular basis by which T7 β (encoding Gp2 ^{β} 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.

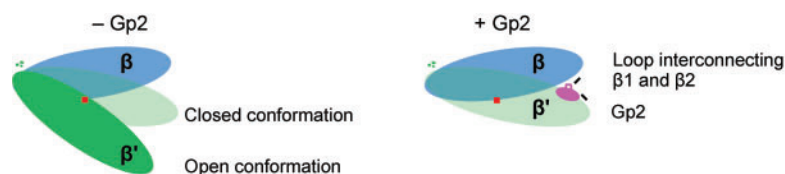


Fig. 6. Gp2 induces RNAP to adopt the 'closed state' conformation. Cartoon showing how the region including and surrounding the β 1 and β 2 strands in Gp2 could contribute to inducing the closed state conformation in the RNAP.

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, Gp2^{14–59}, 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 β 1 and β 2 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. Furthermore, the binding of dwDNA to the dwDBC and the concomitant displacement of the σ^{70} domain, known as region 1.1, from the catalytic cleft are the minimal pre-requisites for efficient and stable RPo formation at promoters utilized by the RNAP containing the σ^{70} factor (Mekler *et al.*, 2002). Gp2 seems to employ a three-pronged strategy to inhibit RPo formation by (i) sterically antagonizing the interaction between dwDNA and the β' jaw domain, (ii) preventing obligatory displacement of σ^{70} region 1.1, and (iii) inducing RNAP to adopt the closed state, thereby restricting DNA access to the DBC and dwDBC (Cámara *et al.*, 2010; Chakraborty *et al.*, 2012; James *et al.*, 2012; Mekler *et al.*, 2011). Thus, the interface between Gp2 and the RNAP could consist of at least two interaction interfaces: a primary one with the β' jaw and one or more auxiliary interaction interfaces. We hypothesize that these interfaces collectively contribute to the high affinity between Gp2 and the RNAP and orchestrate the multipronged mechanism of transcription inhibition by Gp2. Specifically, we envisage a model in which, when Gp2 is bound to the jaw domain on the β' pincer, the region surrounding and including the loop interconnecting the β 1 and β 2 of Gp2 interacts with the β pincer (which is located directly across from the β' jaw on

other side of the DBC) and thereby 'locks' the RNAP in the closed state, leading to very efficient inhibition of RPo formation (Fig. 6). This mechanism of inhibition of RPo formation by Gp2 is analogous to how some antibiotics (such as myxopyronin, coralopyronin and ripostatin) and the non-DNA-binding transcription factor DksA inhibit RPo formation by allosterically inducing RNAP to adopt 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* phages CR8 and CR44b, and providing CR8 and CR44b phage lysates. 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., Vassilyeva, M. N., Svetlov, V., Klyuyev, S., Grishin, N. V., Vassilyev, 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., Cámara, 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., Revyakin, 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 Biosyst* **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., Beck, 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

Key Roles of the Downstream Mobile Jaw of *Escherichia coli* RNA Polymerase in Transcription Initiation

Amanda Drennan,[†] Mark Kraemer,[†] Michael Capp,[†] Theodore Gries,^{†,‡} Emily Ruff,[‡] Carol Sheppard,[§] Sivaramesh Wigneshweraraj,[§] Irina Artsimovitch,^{||} and M. Thomas Record, Jr.^{*,†,‡}

[†]Department of Biochemistry, The University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

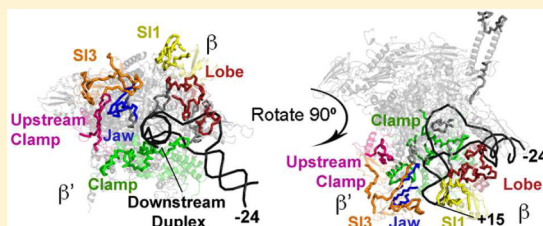
[‡]Department of Chemistry, The University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

[§]Department of Microbiology and Centre for Molecular Microbiology and Infection, Imperial College London, SW7 2AZ

^{||}Department of Microbiology and Center for RNA Biology, The Ohio State University, Columbus, Ohio 43210, United States

S Supporting Information

ABSTRACT: Differences in kinetics of transcription initiation by RNA polymerase (RNAP) at different promoters tailor the pattern of gene expression to cellular needs. After initial binding, large conformational changes occur in promoter DNA and RNAP to form initiation-capable complexes. To understand the mechanism and regulation of transcription initiation, the nature and sequence of these conformational changes must be determined. *Escherichia coli* RNAP uses binding free energy to unwind and separate 13 base pairs of λP_R promoter DNA to form the unstable open intermediate I_2 , which rapidly converts to much more stable open complexes (I_3 , RP_o). Conversion of I_2 to RP_o involves folding/assembly of several mobile RNAP domains on downstream duplex DNA. Here, we investigate effects of a 42-residue deletion in the mobile β' jaw (ΔJAW) and truncation of promoter DNA beyond +12 (DT+12) on the steps of initiation. We find that in stable ΔJAW open complexes the downstream boundary of hydroxyl radical protection shortens by 5–10 base pairs, as compared to wild-type (WT) complexes. Dissociation kinetics of open complexes formed with ΔJAW RNAP and/or DT+12 DNA resemble those deduced for the structurally uncharacterized intermediate I_3 . Overall rate constants (k_a) for promoter binding and DNA opening by ΔJAW RNAP are much smaller than for WT RNAP. Values of k_a for WT RNAP with DT+12 and full-length λP_R are similar, though contributions of binding and isomerization steps differ. Hence, the jaw plays major roles both early and late in RP_o formation, while downstream DNA functions primarily as the assembly platform after DNA opening.



Escherichia coli RNA polymerase (RNAP) is a molecular catalytic machine that synthesizes RNA from a DNA template. Kinetic-mechanistic studies with the WT RNAP and the λP_R promoter as a function of RNAP concentration, upstream DNA length, and solution variables (temperature, salt and solute concentrations), together with footprinting of key intermediates, provide strong evidence that RNAP is also a molecular “isomerization” machine, opening and stabilizing the initiation bubble of promoter DNA in the active site cleft to form a series of open complexes of increasing stability (I_2 , I_3 , and RP_o ; see Figure 1A).^{1–4} Early in the mechanism, RNAP bends the upstream ends of both the –35 and –10 regions of the promoter to wrap upstream DNA on the back side of RNAP and to put the downstream duplex in contact with the cleft in the first kinetically significant intermediate I_1 .^{5,6} Subsequently RNAP actively opens the initiation bubble (–11 to +2) in the cleft to form the initial open complex I_2 , in which the start site region of the template strand may be positioned as in RP_o , but the discriminator region of the nontemplate strand is not.¹

The initial, unstable open complex I_2 is greatly stabilized in the steps converting it to I_3 and RP_o (Figure 1B).^{2,3} Analysis of large effects of urea and other solutes on the composite

dissociation rate constant k_d reveals that ~120 amino acid residues fold or are buried in interfaces in the large scale conformational changes that convert I_2 to RP_o , including ~80 residues in converting I_2 to the subsequent (I_2 – I_3)[‡] transition state and ~40 residues in converting this transition state to RP_o via I_3 .^{3,7} We proposed that conversion of I_2 to RP_o involves rearrangement of the nontemplate strand in the cleft¹ and assembly of the β' jaw (*E. coli* β' 1150–1208) and other mobile domains of RNAP (called downstream mobile elements, or DMEs; defined in Figure 2A) on downstream duplex DNA.^{2,3} Although structural studies of RNAP suggest that the DMEs are rigid bodies, crystallization conditions may select for certain ordered conformations of the DMEs,⁷ and PONDR⁸ calculations suggest that extensive regions of *E. coli* DMEs are not stably folded in apo-holoenzyme (see Figure 2B and Discussion).⁷

In the present study, we focus on the role of the jaw in the steps of open complex formation at the λP_R promoter. Deletion

Received: September 14, 2012

Revised: October 31, 2012

Published: November 1, 2012



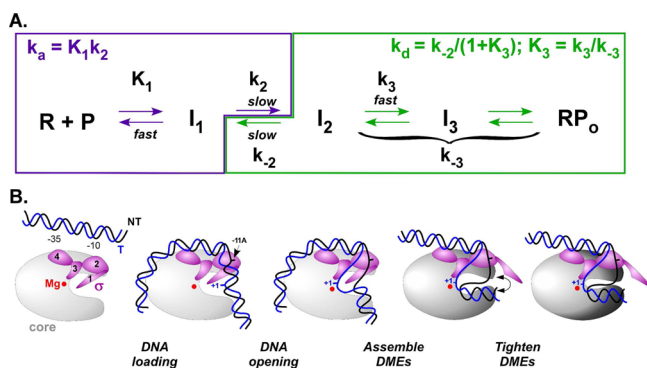


Figure 1. Proposed mechanism of forming and stabilizing the open complex between *E. coli* RNAP and λP_R promoter DNA. (A) Investigation of the kinetics of association of free RNAP (R) with λP_R promoter DNA (P) to form open complexes provides information about the first half of this mechanism, including the equilibrium constant K_1 for binding, bending, and wrapping the DNA duplex to form the advanced closed complex I_1 , and the rate constant k_2 for the rate-determining isomerization step (converting I_1 to I_2) in which the DNA duplex is opened in the cleft by RNAP.^{1,4,6} Investigation of the dissociation of open complexes provides information about the late steps of the DNA opening mechanism: k_3 is the rate constant for conversion of I_2 to I_3 ; and k_{-3} is the composite rate constant for conversion of $R P_o$ (or an initial $I_3/R P_o$ mixed population) to I_2 . The ratio k_3/k_{-3} is defined as K_3 , the equilibrium constant for conversion of I_2 to the equilibrium mixture of I_3 and $R P_o$.³ The rate-determining step in dissociation is DNA closing (rate constant k_{-2}). (B) A cartoon model describing the molecular transitions of this mechanism. The active site Mg^{2+} ion is shown as a red ball. The β' mobile jaw is a downstream mobile element (DME) located in the vicinity of the downstream duplex. Here, we propose that the jaw and other DMEs assemble and tighten around the downstream duplex in the late steps of the mechanism to stabilize the late open complexes.

of the jaw ($\Delta JAW = \beta' \Delta 1149-1190$) is known to reduce the lifetimes of open complexes at several promoters,⁹ but its effects on the steps that form and stabilize open complexes have not been previously characterized. The far downstream region of the promoter (+10 to +20), engulfed¹⁰ by interactions with DMEs (Figure 2) in stable open complexes ($R P_o$), corresponds to the region that has been proposed to interact with the jaw in elongation complexes and heteroduplex promoter complexes.^{9,11,12} Importantly, although deletion of the jaw has pleiotropic effects on transcription initiation, elongation, and termination, it does not compromise *E. coli* viability,⁹ indicating the lack of gross structural changes in ΔJAW transcription complexes.

Recent studies by James et al. are consistent with a direct interaction between the jaw and the promoter DNA.¹³ What is the role of these interactions? Do they participate in promoter opening,¹² or are they responsible for stabilization of the initial unstable open complex (I_2)?^{2,3,7} On the basis of comparison of properties of open complexes formed by ΔJAW RNAP and/or a λP_R promoter variant truncated at +12 (DT+12), we conclude that interactions of the jaw, and/or of neighboring DMEs that become repositioned when the jaw is removed, with the far downstream duplex DNA occur primarily in the late conversion of I_3 to $R P_o$, while interactions of the jaw with incleft elements of RNAP, including other DMEs and/or region 1.1 of σ^{70} ,¹⁴ are responsible for its effects on the early steps of open complex formation.

EXPERIMENTAL PROCEDURES

Buffers. RNAP storage buffer contains 50% glycerol (v/v), 0.01 M Tris (pH 7.5 at 4 °C), 0.1 M NaCl, 0.1 mM DTT, and 0.1 mM Na_2EDTA . Wash buffer for nitrocellulose filter binding experiments contains 0.01 M Tris (pH 8.0 at room temperature), 0.1 M NaCl, and 0.1 mM Na_2EDTA . Phosphate buffer (PB) for dissociation kinetic experiments in the absence of Mg^{2+} is 0.01 M $Na_2HPO_4^-$, 0.16 M NaCl (0.14–0.30 M NaCl in salt dependence experiments), 1 mM dithiothreitol (DTT), and 100 $\mu g/mL$ BSA. Tris buffer (TB), a common transcription buffer used here for kinetics experiments in the presence of Mg^{2+} , is 0.04 M Tris (adjusted to pH 8.0 at the experimental temperature with HCl), 0.12 M KCl, 0.01 M $MgCl_2$, 1 mM DTT, and 100 $\mu g/mL$ BSA. Low salt Tris buffer (LSTB) is the same as TB, but with 0.06 M KCl and 0.005 M $MgCl_2$. RNAP storage buffer was added to keep the glycerol concentration constant in a series of association experiments (where the RNAP concentration was varied) and in dissociation experiments (6% total RNAP storage buffer in TB and LSTB; 13% in PB). For MnO_4^- footprinting experiments, DTT was omitted; for $HO\bullet$ footprinting experiments, the Tris concentration was reduced to 0.01 M.

RNAP Preparation. *E. coli* ΔJAW RNAP core was prepared as described previously.¹⁵ *E. coli* WT RNAP core, σ^{70} , and holoenzyme were endogenously expressed and purified from *E. coli* MG1655¹⁶ as described previously.¹ ΔJAW holoenzyme was reconstituted by incubating ΔJAW RNAP core (500 nM) and WT σ^{70} (2 μM) for 90 min at 37 °C.

λP_R Promoter DNA Preparation. Full-length (FL) λP_R DNA duplex fragments (~180 bp long; approximately -120 to +60) were PCR-amplified from plasmid pPR59 (λP_R -59 to +34, a gift from Dr. Wilma Ross)¹⁷ using downstream primer 5'-CAGGACCCGGGCGCGCTTAATTAACACTCTTA-TACATTATTCC-3' and upstream primer 5'-GTAC-GAATTCGATATCCAGCTATGACCATGATTACGC-CAGC...3'. Fragments were isolated using the Qiagen QIAquick PCR Purification Kit (Valencia, CA). To end label the template strand downstream, duplex fragments were cleaved with BssHII, and the 5' phosphates were removed with Antarctic Phosphatase and replaced with [γ -³²P] ATP using T4 polynucleotide kinase. The 3' ends were filled in using Sequenase with 1 mM dGTP + dCTP, and the upstream end was cut with EcoRV to remove the upstream label. For nontemplate strand downstream fill-in label, duplex fragments were cleaved with EcoRV and BssHII, and filled in with [α -³²P] dCTP (along with 1 mM cold dGTP + dCTP) using Sequenase.

Downstream-truncated (at +12; DT+12) λP_R DNA duplex fragments (~140 bp long; approximately -120 to +12) were obtained by PCR as above, using the same upstream primer as for full-length λP_R (see above) and downstream primer 5'-GAGGACCCGGGTTACTACATGCAACCATTAT-CACCGCCAGAG-3'. To fill-in label the upstream end of the template strand, the PCR product was cleaved with EcoRI and SmaI, and labeled with [α -³²P] dATP (added with 1 mM cold dATP + dTTP) using Sequenase.

Enzymes used in λP_R promoter DNA preparation and labeling were obtained from New England Biolabs (Ipswich, MA), except for Sequenase, which was obtained from Affymetrix (Santa Clara, CA). All labeled fragments were purified on a 5% acrylamide gel and isolated using an Elutip-d

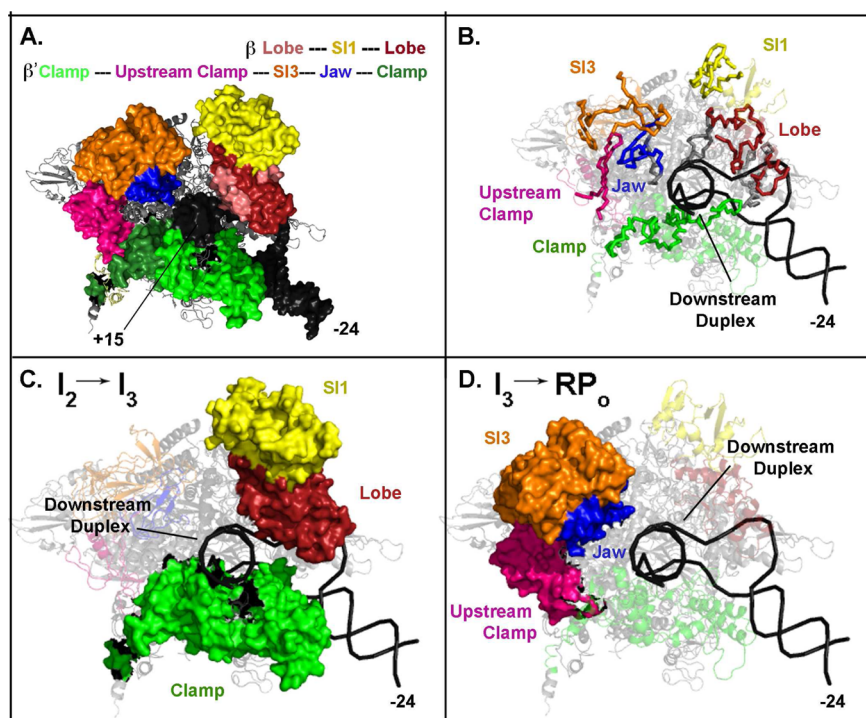


Figure 2. DMEs and their proposed stepwise assembly on downstream DNA to stabilize the initial open complex (I_2) in transcription initiation. (A) DMEs of the β and β' subunits (surface representation) surrounding the downstream DNA duplex (black, to +15) are defined on a model of the *E. coli* RNAP transcription elongation complex.²² DMEs include the β lobe (red, β 184–224, 344–438); the β sequence insertion (SI)1 (yellow, β 225–343); the β' clamp (green, β' 131–347, 1260–1368); the β' upstream clamp (pink, β' 808–912); β' SI3 (orange, β' 943–1130); and the β' jaw (blue, β' 1149–1208). All residue numbers refer to *E. coli* RNAP. (B) Regions of β and β' in *E. coli* RNAP²² that are predicted to be unstructured (primarily DMEs) by the PONDR (Predictor of Natural Disordered Regions) VL-XT program are represented by brightly shaded ribbons (see Supporting Information). (C) In the conversion of I_2 to I_3 (or the preceding transition state), a step in which approximately 80 RNAP residues are buried by folding/assembly/binding,³ we propose that the β' clamp and/or β lobe assemble on the proximal DNA duplex (+3 to +10). (D) In the conversion of I_3 (or the previous transition state; see text) to RP_0 , a step in which approximately 40 RNAP residues are buried by folding/assembly/binding,³ we deduce that the β' jaw and neighboring DMEs assemble on the distal downstream DNA (+10 to +20).

Purification Minicolumn (Whatman (GE Healthcare, Pittsburgh, PA)).

Dissociation Assays. For dissociation kinetic assays, an initial population of $90 \pm 10\%$ open complexes was obtained by incubation of excess WT or Δ JAW RNAP holoenzyme (5–20 nM active) with 32 P-labeled λ P_R promoter (~ 150 pM) at 25 or 37 °C. Irreversible dissociation was initiated by addition of 50–100 μ g/mL heparin, a nondisplacing competitor for both WT¹⁸ and Δ JAW RNAP (Figure S2, Supporting Information). For studies of effects of urea on dissociation, urea was added at the same time as heparin. As a function of time, 100 μ L aliquots (~ 3000 cpm) were removed and applied to nitrocellulose filters on a vacuum manifold.⁴ Filters were dried under heat lamps and counted with a Hewlett-Packard Tri-Carb 2100 TR scintillation counter. In TB, about 20% of open complexes formed by Δ JAW RNAP are found to be nondissociating. Addition of urea reduces the initial population of Δ JAW open complexes, making it difficult to quantify the dissociation rate constant k_d accurately as a function of urea in TB. Experiments with WT and Δ JAW RNAP to determine effects of urea on k_d were therefore performed at a convenient NaCl concentration (0.16 M) in PB, the buffer used to determine the effects of univalent salt concentration on k_d .

Analysis of Dissociation Data. Filter-retained counts were corrected by subtracting background counts (determined by filtering solutions in the absence of RNAP) to determine cpm_f, the filter-retained counts due to open complexes. For each data

set, the dissociation rate constant k_d was determined by fitting cpm_f vs time to the first order rate equation

$$\text{cpm}_t = A + B \exp(-k_d t) \quad (1)$$

where $A = \text{cpm}_{t=\infty}$ (which can differ from background because of a subpopulation of nondissociating complexes) and $B = \text{cpm}_{t=0} - \text{cpm}_{t=\infty}$. The fraction of promoter DNA in the form of open complexes at time t (θ_t) is calculated as $\theta_t = (1/E)(\text{cpm}_{t=0}/\text{cpm}_{\text{tot}})$ where cpm_{tot} is the total cpm in the aliquot filtered and E is the efficiency of retention of open complexes on the nitrocellulose filter membrane, typically in the range 0.6–0.9 (see below) depending on the details of how the filters are prepared and how the filter bound complexes are treated with wash buffer.³ In Figures 5 and 6, the quantity plotted is $\Delta\theta = \theta_t - \theta_{t=\infty}$, where $\theta_{t=\infty} = (1/E)(\text{cpm}_{t=\infty}/\text{cpm}_{\text{tot}})$. All data fitting, analysis and normalization were performed using Igor Pro Version 5.03.

Determination of the Equilibrium Constant K_3 for Stabilization of the Initial Open Complex by Assembly of Downstream Mobile Elements. Values of K_3 (defined in Figure 1) were determined from experimental values of dissociation rate constants k_d using the relationship³

$$K_3 = k_{-2}/k_d - 1 \cong k_{-2}/k_d \text{ for } k_d < k_{-2} \quad (2)$$

The value of the DNA closing rate constant k_{-2} for WT RNAP and FL λ P_R promoter DNA at 25 °C was interpolated from results at 10 and 37 °C,³ using the experimentally determined

Table 1. Dissociation Rates of ΔJAW and WT RNAP Open Complexes and Their Interpretation: Effects of [Urea]^a

RNAP	λPR	relative dissociation rate ^b	10 ⁻³ K ₃ (M ⁻¹) ^c	ΔG ₃ ^{0c} (kcal/mol)	urea dependence of k _d (d ln k _d /d[urea]) (M ⁻¹) ^d
WT	FL	1	~63	-6.5 ± 0.3	3.1 ± 0.1
WT	DT+12	23 ± 3	~2.7	-4.7 ± 0.3	1.9 ± 0.2
ΔJAW	FL	67 ± 7	~0.9	-4.1 ± 0.3	2.1 ± 0.1
ΔJAW	DT+12	52 ± 11	~1.2	-4.2 ± 0.4	2.1 ± 0.1

^aConditions: PB (0.16 M NaCl), 25 °C. ^bCalculated for WT/FL dissociation rate constant k_d = (2.7 ± 0.2) × 10⁻⁵ s⁻¹ for these conditions (Figure 5). ^cK₃ and ΔG₃⁰ defined as in Figure 1. K₃ is the composite equilibrium constant and ΔG₃⁰ is the corresponding standard free energy change for the process of converting the initial open complex I₂ to the equilibrium population of stable open complexes (RP₀, I₃) for λP_R at 25 °C. K₃ calculated from eq 2 using a DNA closing rate constant k₋₂ = 1.7 ± 0.4 s⁻¹ at 25 °C, obtained by interpolation of WT/FL results.³ Uncertainty in K₃ ~ 30–45%. ^dSlopes of plots (Figure 6D) of ln k_d vs [urea].

activation energy (9.9 kcal/mol). This DNA closing rate for WT RNAP and FL λP_R is independent of urea concentration; the entire effect of urea on k_d is on K₃.³ A preliminary dissection of the dissociation kinetics for ΔJAW RNAP–FL λP_R complexes (E. Ruff, unpublished) indicates that the difference in k_d between ΔJAW and WT RNAP is largely in K₃, and that the effect of urea in both cases is primarily on K₃. We assume that this situation also applies for DT+12 λP_R. For the dissociation of WT RNAP from FL λP_R the uncertainty in k₋₂ is approximately 20%,³ and that in K₃ is approximately 30%.

Association Assays. To initiate association, excess WT or ΔJAW RNAP (1–50 nM active) was added to ³²P-labeled λP_R promoter (~150 pM) in TB at 37 °C. At selected times after mixing, 100 μL aliquots (~3000 counts) were withdrawn. Heparin (50 μg/mL final) was added for 10 s and each aliquot was filtered and counted as described above.

Analysis of Association Data. Association experiments for ΔJAW RNAP at low concentrations (e.g., in Figure 8 in Results) do not go to completion under the conditions investigated, and kinetic data were therefore analyzed as reversible associations⁴ according to eq 3:

$$\text{cpm}_t = \text{cpm}_{t=0} + (\text{cpm}_{\text{eq}} - \text{cpm}_{t=0})(1 - \exp(-\beta t)) \quad (3)$$

where β is the decay to equilibrium rate constant, cpm_t is the background-corrected counts retained on the filter at time t, and cpm_{eq} and cpm_{t=0} are the values of cpm_t at equilibrium and extrapolated to t = 0, respectively. The decay rate constant β is related to the rate constants k_d and α of dissociation and association, respectively, by

$$\beta = k_d + \alpha = \alpha/\theta_{\text{eq}} \quad (4)$$

where θ_{eq} is the fraction of promoter DNA in the form of competitor-resistant (open) complexes at equilibrium, and

$$\alpha = K_1 k_2 [\text{RNAP}] / (1 + K_1 [\text{RNAP}]) \quad (5)$$

where K₁ is the equilibrium constant for forming the I₁ intermediate closed complex and k₂ is the isomerization (DNA opening) rate constant.⁴ The composite association rate constant k_a (Table 1) is given by⁴

$$k_a = K_1 k_2 \quad (6)$$

If association is irreversible, then θ_{eq} = 1, k_d = 0, and β = α. More generally, determination of α from β requires knowledge of either θ_{eq} or k_d. Determination of θ_{eq} from the experimental quantities cpm_{eq} and cpm_{tot} (the background-corrected total cpm in each filtered aliquot) requires knowledge of the filter efficiency E, the fraction of competitor resistant (open) complexes in solution that are retained on the filter:⁴

$$\text{cpm}_{\text{eq}} / \text{cpm}_{\text{tot}} = \theta_{\text{eq}}^{\text{obs}} = E \theta_{\text{eq}} \quad (7)$$

Therefore, from eqs 4 and 7,

$$\beta \theta_{\text{eq}}^{\text{obs}} = E(\beta - k_d) \quad (8)$$

Analysis of the variation in experimental values of βθ_{eq}^{obs} vs β for a series of experiments at different [RNAP] yields best-fit values of E and k_d, allowing calculation of θ_{eq} and α from each determination of β using eq 4. For the sets of experiments analyzed here, efficiencies E of 0.62, 0.77, and 0.81 were obtained for WT-DT+12 complexes, ΔJAW-FL complexes, and ΔJAW-DT+12 complexes, respectively. Values of E and θ_{eq} were used to calculate θ_i for Figure 8. For WT RNAP and DT+12 promoter DNA, values of α obtained from eq 4 were analyzed as a hyperbolic function of [RNAP] using eq 5 to obtain k_a, K₁, and k₂. For ΔJAW RNAP, where α is observed to be proportional to [RNAP] over the accessible concentration range, only k_a is obtained from the analysis. Igor Pro Version 5.03 was used for data fitting and analysis.

KMnO₄ Footprinting. Open complexes (5 nM active RNAP + ~1 nM ³²P-λP_R promoter DNA) were formed in TB (-DTT) or LSTB (-DTT) by incubation for 30 min at 37 °C. After incubation, heparin (100 μg/mL) was added to some samples for 10 s before addition of KMnO₄ to eliminate short-lived complexes. KMnO₄ (1.25 mM final concentration) was added for 10 s (determined to be a single-hit dose) and reactions were quenched with 1 M β-mercaptoethanol. Samples were then cleaved with 1 M piperidine, purified, and run on an 8% acrylamide gel together with control (-RNAP) lanes.^{19,20}

HO• Footprinting. Open complexes were formed in LSTB at 37 °C as described for KMnO₄ footprinting, except that buffers contained DTT and were prepared with a reduced Tris concentration (10 mM) because Tris scavenges HO•.²¹ After a brief (10 s) challenge with 100 μg/mL heparin, [Fe-EDTA]⁻² (1.8 mM Fe²⁺, 4.2 mM EDTA) and sodium ascorbate (1.2 mM) were added, followed by addition of H₂O₂ (0.1%) to generate HO•. After 10 s, a short enough exposure time to be in the single-hit regime, reactions were quenched (11.7 mM thiourea + 2.0 μM EDTA) and worked up by phenol extraction and ethanol precipitation. Samples were run on an 8% acrylamide gel.^{6,10}

Gel Imaging and Analysis. Footprinting and transcription gels were imaged using either a Typhoon 9410 Variable Mode Imager or a Typhoon FLA-9000 (both from GE Healthcare) and analyzed using Image Quant TL. HO• line scans were normalized to a short region of the footprint upstream of -60. This upstream normalization provides a better match of downstream as well as upstream baselines for the different line scans on each gel than is achieved by normalization to any group of downstream bands or to total cpm. While small drifts

in the downstream baselines are observed with this normalization (visible in the difference line scans in Figure 4E,F), they are not large enough to affect the semiquantitative conclusions of our analysis. To reduce noise, a moving average of 10 positions (approximately 0.7% of the length of the gel) was applied to each line scan (Excel). Difference line scans for each DNA strand were determined by measuring the distance between the free DNA backbone and RNAP-complexed DNA backbone for each DNA base position +1 to +22. For each strand, distances were normalized to the maximum distance, averaged with the independent (duplicate) experiment, and plotted in a bar graph to more explicitly portray the relative protection of WT and Δ JAW RNAP open complexes.

Prediction of RNAP Disordered Regions. The *E. coli* RNAP β and β' subunits²² were analyzed using VL-XT, the most stringent program in PONDR⁸ (Predictions of Naturally Disordered Regions). Only sequences in which at least 15 sequential residues are predicted to be disordered were considered, and predicted disordered sequences known to be in subunit–subunit interfaces were excluded.

RESULTS AND ANALYSIS

KMnO₄ Footprints of λ P_R Promoter DNA in Δ JAW Complexes. How does the highly conserved bacterial jaw domain affect the distribution and structure of RNAP–promoter complexes? Ederth et al. observed that Δ JAW RNAP forms unstable complexes at the λ P_R promoter at 30 °C at moderate salt concentration (0.01 M MgCl₂, 0.15 M KCl).⁹ To eliminate the possibility that a mixed population of complexes and free promoter DNA exists under these conditions, we sought conditions to increase the stability and lifetime of Δ JAW RNAP complexes at the λ P_R promoter (Figure 3A) for footprinting

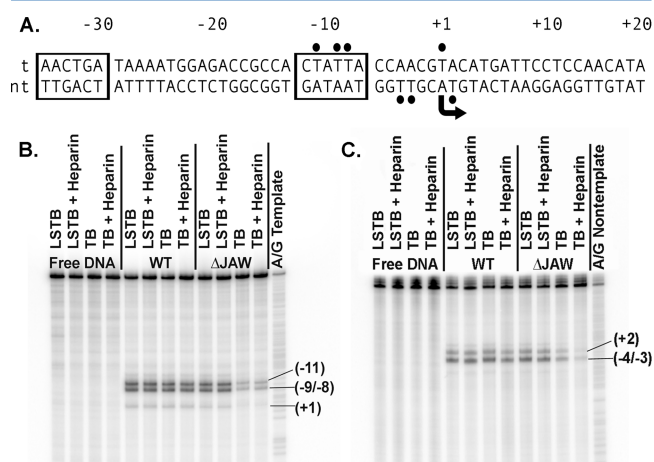


Figure 3. MnO₄[–] cleavage of unstaked (single-stranded) thymine residues of λ P_R promoter DNA–RNAP open complexes. (A) Sequence of λ P_R promoter DNA positions –35 to +20. The –35 and –10 elements of the promoter, which interact with σ regions 4.2 and 2.3/2.4, respectively,⁵ are boxed. MnO₄-reactive thymines in WT RNAP open complexes are denoted by bullets above and below the template (t) and nontemplate (nt) strands, respectively. (B, C) Comparison of KMnO₄ footprints of λ P_R promoter DNA in Δ JAW and WT RNAP open complexes in TB (0.01 M Mg²⁺, 0.12 M K⁺) and in LSTB (0.005 M Mg²⁺, 0.06 M K⁺) at 37 °C: (B) template strand; (C) nontemplate strand. Positions of reactive thymine residues are indicated; nontemplate strand thymines at positions –7 and –10 are not reactive to MnO₄[–] because they are protected by interactions with σ .³⁷

experiments. At 37 °C in lower-[salt] buffer (LSTB: 0.005 M MgCl₂, 0.06 M KCl), Δ JAW RNAP– λ P_R open complexes are long-lived (Figure 3B,C; LSTB lanes), suggesting that all bound DNA is in the form of an open complex.

To compare extents of DNA opening, we exposed Δ JAW and WT RNAP complexes to KMnO₄ in either LSTB or TB (0.01 M MgCl₂, 0.12 M KCl). KMnO₄ preferentially oxidizes unstacked thymines, providing a method of detecting and quantifying DNA opening in RNAP–promoter complexes.^{1,20,23} We find that Δ JAW– λ P_R complexes are open: thymines at –11, –9, –8 and +1 on the template strand (Figure 3B) and –4, –3 and +2 (but not –7 and –10) on the nontemplate strand (Figure 3C) exhibit the same KMnO₄ reactivity, and therefore are exposed to the same extents in the open regions of WT– λ P_R and Δ JAW– λ P_R open complexes. In TB, detected thymines on both strands in Δ JAW open complexes are significantly less KMnO₄-reactive than in WT open complexes (Figure 3B,C; TB lanes). Though Δ JAW open complexes are less stable and shorter-lived in TB than in LSTB, this cannot explain the large reductions in KMnO₄ reactivity (see below). For both WT and Δ JAW RNAP, λ P_R promoter DNA opening and closing rate constants are only weakly (if at all) dependent on salt concentration³ (E. Ruff, unpublished), so this is not a salt effect on DNA opening. Nor can it be a salt effect on KMnO₄ reactivity, since the KMnO₄ reactivities of these thymines in WT open complexes are the same in TB as in LSTB.

HO• Footprints of λ P_R Promoter DNA in Δ JAW Open Complexes. The jaw domain is located near the downstream DNA (Figure 2A), and can interact with DNA nonspecifically.^{13,24,25} Furthermore, a large β' region that includes the jaw cross-links to the DNA backbone.²⁶ However, direct evidence demonstrating jaw contacts to DNA in *any* RNAP–DNA complex (e.g., closed, open, initiation, or elongation states) is lacking. Here, we ask whether the jaw deletion alters protection of the DNA backbone in open promoter complexes. Figure 4 compares normalized line scans from quantitative single-hit HO• footprints of both template (panels A and C) and nontemplate (panels B and D) strands of FL λ P_R in stable open complexes formed by either WT or Δ JAW RNAP in LSTB at 37 °C. (Sequencing gels for the experiments in panels A and B are shown in Figure S1, Supporting Information.) Bar graphs of average differences in HO• protection between open complexes and free promoter DNA at each downstream position from +1 to +22 are shown in Figure 4E,F.

Two very significant differences are observed in HO• footprints of Δ JAW and WT open complexes. First, the downstream boundaries of HO• protection of both strands in the Δ JAW RNAP open complex are shorter than for WT RNAP. The boundary on the template strand (Figure 4E) is approximately 1 helical turn shorter for the Δ JAW open complex than for that formed by WT RNAP, ending near +10 instead of +20. For the nontemplate strand (Figure 4F), the downstream boundary of the Δ JAW RNAP open complex is near +15, about 0.5 helical turn shorter than that observed for WT RNAP (+20). (The difficulty in quantitatively normalizing the entire length of these >100 nucleotide line scans is revealed in the drift in the region downstream of the boundary of the WT footprint at +20, observed in the difference plots in Figure 4E,F. This drift is not large enough to affect the semiquantitative results presented above, though it does increase the uncertainty in assignment of the downstream boundary of the Δ JAW footprints (sequence position +10 $\pm$ 2 bases on the template strand and position +15 $\pm$ 3 bases on the nontemplate

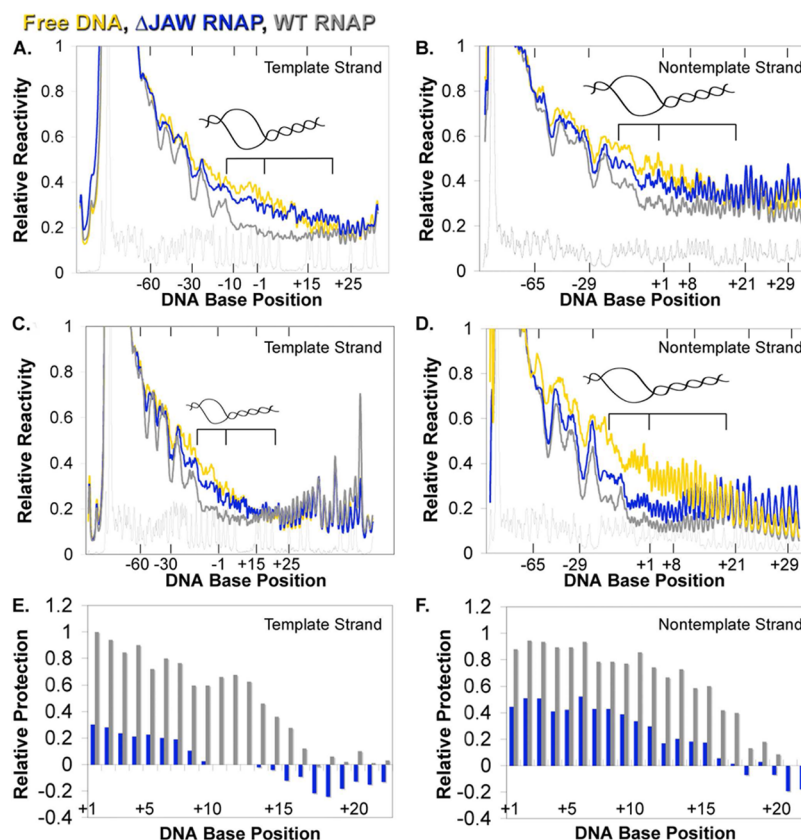


Figure 4. Line scans of lanes in single-hit HO• footprinting experiments (Figure S1) performed in duplicate for the template (A,C) and nontemplate (B, D) strands of λP_R promoter DNA in open complexes with ΔJAW RNAP (blue) or WT RNAP (gray) and free λP_R DNA (gold) in LSTB at 37 °C. The regions of the DNA initiation bubble (−11 to +2) and the downstream DNA duplex (+3 to +20) are indicated by brackets below a model of these DNA regions. Bar graphs in panels E and F display the relative protection of individual downstream positions of template (E) and nontemplate (F) strands in ΔJAW and WT open complexes, calculated as the average of difference line scans for each strand (see Experimental Procedures). Each panel is normalized so the maximum difference in protection in this interval is set equal to 1.0.

strand.) The reduction in HO• protection of the ΔJAW RNAP open complex provides direct evidence that the jaw, and/or neighboring regions of RNAP perturbed by deleting the jaw, likely contacts one side of the duplex DNA from +10 to +20 in WT RP_o . Direct contact between the jaw and the distal downstream duplex (+10 to +20) was proposed previously.^{9,24,26} However, structural studies of bacterial RNAP suggested that the interaction between the jaw and the DNA downstream duplex is proximal to the transcription start site, forming an interface only to about +12.^{22,25} In this case, the presence of the jaw might act indirectly by influencing the concurrent folding/assembly of one or more additional DMEs onto the distal downstream duplex. At this time, we are not able to distinguish between these two possibilities.

A second very striking difference between HO• footprints of ΔJAW and WT open complexes, shown for the downstream region of the promoter DNA analyzed in Figure 4E,F (but also clearly observable in the upstream promoter regions in Figure 4A–D), is that all regions of both DNA strands in the ΔJAW open complex are significantly less protected from HO• than is the case for WT RP_o . This effect extends at least from −45 to the downstream boundary of the footprint and is especially pronounced in the open region of the promoter (initiation bubble; −11 to +2) in the active site cleft. This global reduction in protection cannot be attributed simply to the shorter lifetime of the ΔJAW open complex, which is more than 1 h in LSTB and therefore much longer than the time scale of the

footprinting experiments (10 s). In TB, the ΔJAW open complex exhibits no significant HO• protection of either template or nontemplate strand (data not shown), even though the lifetime of the ΔJAW complex in TB (30 min) is much greater than the time required to footprint (10 s; see below), and $KMnO_4$ footprints (Figure 2) reveal that this complex is an open complex.

A plausible explanation of the reduced protection observed for the ΔJAW RNAP open complex in both buffers is that more water remains in the ΔJAW RNAP–promoter interface than in the WT RNAP–promoter interface, and that HO• gains access to the DNA in the interface by facilitated diffusion on a network of residual water of hydration. Alternatively, or in addition, the ΔJAW RNAP–promoter interface may be more dynamic than that of the WT enzyme, exhibiting normal modes of vibration that result in local separations of RNAP and promoter DNA that allow HO• access to all portions of the interface in a 10 s footprinting experiment without any global dissociation.

Large Effects of Deletion of the Jaw or of Downstream Promoter DNA on the Rate of Dissociation of Open Complexes. To investigate when, where, and how strongly the jaw and neighboring DMEs assemble on the downstream duplex in the process of forming the stable open complex at λP_R , we compared the kinetics of dissociation of open complexes formed by WT and ΔJAW RNAP on FL and truncated (at +12) λP_R promoter DNA (DT+12). Dissociation

time courses in PB (0.16 M NaCl) at 25 °C are shown in Figure 5A. Dissociation rate constants k_d calculated from these

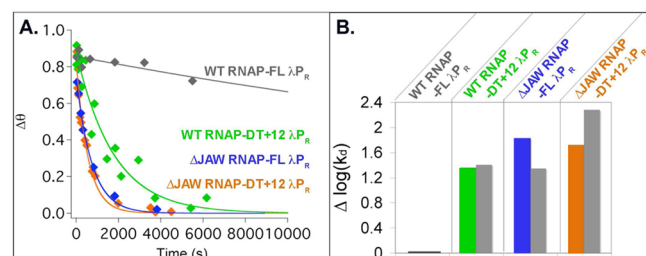


Figure 5. Kinetics of dissociation of WT and variant open complexes in PB (0.16 M Na⁺) at 25 °C (panel A), as compared to results in TB (0.01 M Mg²⁺, 0.12 M K⁺) at 37 °C (panel B). (A) Fraction of dissociable open complexes remaining at time t ($\Delta\theta$; see Experimental Procedures), determined using the nitrocellulose filter assay in PB with excess heparin as competitor (Table 1). Complexes investigated include WT RNAP at FL (gray) or DT+12 (green) λP_R promoter DNA and Δ JAW RNAP at FL (blue) and DT+12 (orange) λP_R promoter DNA. (B) Relative dissociation rate constants of variant complexes expressed relative to WT-FL complexes. Results in PB (green, blue, and orange bars, as determined by studies shown in panel A) are compared with those in TB (gray bars, as determined by the nitrocellulose dissociation assay (WT complexes) or decay to equilibrium data (Δ JAW complexes; see Table 2)).

data (see Table 1) are compared with one another, and with the corresponding quantities determined at 37 °C in TB (0.12 M KCl, 0.01 M Mg²⁺), in Figure 5B. Values of k_d in TB, obtained by analysis of reversible association kinetic experiments (shown in Figure 8 below) are listed in Table 2. In both PB (25 °C) and TB (at 37 °C), WT-FL open complexes are long-lived. Dissociation rate constants k_d for WT-FL open complexes are small and of comparable magnitude ($k_d = 2.7 \times 10^{-5} \text{ s}^{-1}$ in PB at 25 °C; $k_d = 2.2 \times 10^{-5} \text{ s}^{-1}$ in TB at 37 °C¹⁸), permitting convenient study of the destabilizing effects of urea on k_d of the less stable variant complexes.

For both sets of solution conditions, effects of deletion of the jaw or of the downstream DNA on k_d are large (~20–60-fold increase in k_d from that observed for dissociation of WT RNAP from FL λP_R). Details differ somewhat in the two conditions investigated. In TB, each single variant increases k_d approximately 25-fold; in PB, deleting the jaw increases k_d even more dramatically (67-fold), while the effect of deleting the downstream duplex is similar to that observed in TB. The

additional effect of the second deletion (comparing k_d for Δ JAW-DT+12 with k_d for Δ JAW-FL or WT-DT+12) is small in PB at 25 °C, but larger in TB at 37 °C. In PB at 25 °C, deletion of the downstream DNA in Δ JAW open complexes does not affect the dissociation rate, and deletion of the jaw in open complexes formed with DT+12 DNA increases the dissociation rate by only a small amount (~2-fold) compared to WT RNAP. In TB at 37 °C, however, deletion of the downstream DNA in Δ JAW open complexes increases the dissociation rate by approximately 8-fold. These differences between the variants in PB and TB probably result from specific effects of Mg²⁺; previous dissociation studies with WT RNAP and FL λP_R as a function of Mg²⁺ concentration provided evidence for binding of 2–3 Mg²⁺ ions in the conversion of I_2 to RP_o .²⁷

Because $k_d = k_{-2}/(1 + K_3)$, where k_{-2} is the DNA closing rate constant (the bottleneck step in dissociation) and K_3 is the composite equilibrium constant for forming the stable open complex (RP_o , I_3 , or an equilibrium mixture of the two) from I_2 (Figure 1), these large effects on k_d of deleting either the jaw or the downstream DNA could result from a combination of increases in k_{-2} and/or reductions in K_3 . For WT RNAP, the large effect of urea on k_d arises from K_3 , while k_{-2} is independent of urea concentration.³ Because deletion of either the jaw or the downstream duplex significantly reduces the urea dependence of k_d (as shown in Figure 6D below), we infer that the primary effect of these deletions must be on K_3 . Consistent with this deduction, preliminary measurements of k_{-2} for dissociation of Δ JAW RNAP from FL λP_R promoter DNA (E. Ruff, unpublished) reveal that it is similar to (indeed somewhat smaller than) k_{-2} for WT RNAP.

Values of the equilibrium constant K_3 and the standard free energy change ΔG_3^0 for conversion of I_2 to the stable open complex obtained from these data using eq 2 are listed in Tables 1 and 2 for the two conditions investigated. For WT RNAP and FL λP_R , the full stabilization of the open complex is ~−7 kcal/mol (K_3 of order 10^5), as determined previously.³ A similar conclusion was obtained from dissociation kinetic studies with DT heteroduplexes.¹² Tables 1 and 2 show that deletion of either the far downstream duplex (at +12) or of the jaw eliminates approximately one-third of the stabilization provided by the DME assembly. In PB, deletion of *both* the downstream duplex (at +12) and the jaw has no additional effect, as expected if the main effect of deleting the jaw is to eliminate interactions of DMEs with region +12 to +20 of the

Table 2. Rate and Equilibrium Constants for the Steps of Transcription Initiation at the λP_R Promoter at 37 °C^a

RNAP	λP_R	relative dissociation rate ^b	$10^{-3} K_3 \text{ (M}^{-1}\text{)}^c$	$\Delta G_3^0 \text{ (kcal/mol)}$	$10^{-6} k_a^d \text{ (M}^{-1} \text{ s}^{-1}\text{)}$	$10^{-6} K_1^d \text{ (M}^{-1}\text{)}$	$k_2 \text{ (s}^{-1}\text{)}^d$
WT	FL	1	~150	−6.7 ± 0.3	3.8 ± 0.7 ^e	6.6 ± 2.7 ^e	0.66 ± 0.27 ^e
	DT+12	25 ± 3	~5.9	−5.4 ± 0.3	5 ± 2	40 ± 20	0.14 ± 0.09
Δ JAW	FL	23 ± 3	~6.6	−5.5 ± 0.3	0.33 ± 0.05 ^f	nd ^g	nd ^g
	DT+12	180 ± 80	~0.8	−4.2 ± 0.8	0.33 ± 0.05 ^f	nd ^g	nd ^g
WT	UT-47 ^h	0.9 ± 0.3	~170	−7.4 ± 0.5	0.5 ± 0.1	26 ± 2	0.02 ± 0.003

^aConditions: TB (0.01 M MgCl₂, 0.12 M KCl), 37 °C. ^bCalculated for WT/FL dissociation rate constant $k_d = 2.2 (\pm 0.3) \times 10^{-5} \text{ s}^{-1}$ for these conditions.¹⁸ Values of k_d for variants were determined from reversible association kinetic data of Figure 8 using eq 8 (see Experimental Procedures). ^c K_3 and ΔG_3^0 are defined in Figure 1 and Table 1 footnote c; at 37 °C for λP_R , these quantities presumably refer to the process of converting I_2 to RP_o . K_3 is calculated from eq 2 using a DNA closing rate constant $k_{-2} = 3.3 (\pm 0.7) \text{ s}^{-1}$ at 37 °C.³ Uncertainty in K_3 is ~30% except for Δ JAW-DT+12 and WT-UT-47 where it is ~75% and ~60%, respectively. ^dValues of k_a , K_1 , and k_2 obtained from eqs 5 and 6; k_a is the composite second-order association rate constant for open complex formation, K_1 is the composite equilibrium constant for forming the I_1 closed complex, and k_2 is the isomerization rate constant for converting closed intermediate I_1 to open intermediate I_2 . ^eRef 4. ^fAssociation results for Δ JAW RNAP at FL and DT+12 λP_R promoters were fit together. ^gNot able to be determined from these data. ^hDissociation and association rate constants for WT RNAP at upstream-truncated (at −47; UT-47) λP_R .¹⁹

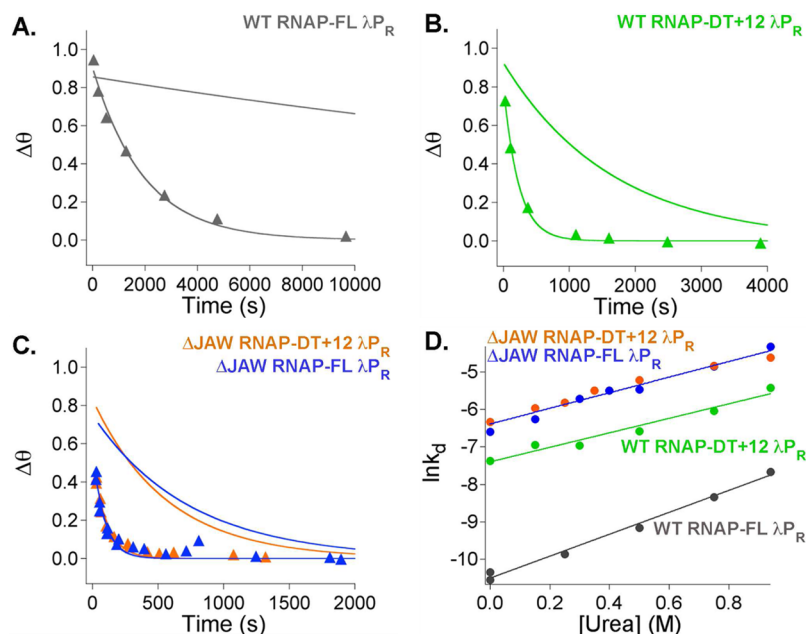


Figure 6. Effects of urea on the dissociation rate constant k_d , used as a probe of coupled folding and assembly of DMEs (Figure 2), including the jaw, on downstream DNA. Experiments were performed using the nitrocellulose filter assay in PB at 25 °C. Panels A–C compare dissociation time courses in 0.94 M urea (triangles), plotted as in Figure 5A, with those in the absence of urea (fits reproduced from Figure 5A). (A) WT-FL λP_R (gray). (B) WT-DT+12 λP_R (green). (C) Δ JAW-FL (blue) or -DT+12 λP_R (orange). (D) Plot of $\ln k_d$ vs urea molarity, comparing dissociation of WT and Δ JAW RNAP from FL and DT+12 λP_R promoter DNA. Colors are coded to panels A–C. The fitted line for the combined Δ JAW-FL and Δ JAW-DT+12 data set is blue. Slopes ($d \ln k_d / d[\text{urea}]$) are listed in Table 1.

downstream duplex. Effects observed in both buffers indicate that removal of the jaw disrupts a salt- and Mg^{2+} -dependent network of contacts of DMEs with each other and with the distal downstream duplex, greatly reducing K_3 and the lifetime of the open complex and greatly increasing its $\text{HO}\bullet$ reactivity.

Effects of Urea on Dissociation Rate Constants k_d for WT vs Variant Open Complexes. The effect of urea on dissociation rate constants for WT RNAP-FL λP_R open complexes in PB is as large as that observed for unfolding a small globular protein (100–150 residues),²⁸ which we interpreted in terms of the folding and assembly of DMEs on the downstream duplex DNA (+3 to +20) in the process of converting I_2 to RP_o .^{3,7} If deletion of the jaw and/or the downstream duplex reduces the folding and assembly of DME on the downstream DNA in these late steps, then dissociation of these variant open complexes should exhibit a smaller dependence on urea concentration than for the WT-FL RP_o . Panels A–C of Figure 6 compare the kinetics of dissociation of WT-FL and variant open complexes in 0.94 M urea with that observed in the absence of urea (from Figure 5A). These panels illustrate the general observation that the effect of urea on the rate of dissociation of WT-FL RP_o is much larger than the effect of urea on the rates of dissociation of the Δ JAW RNAP and DT+12 downstream deletion variants. Dissociation rate constants k_d obtained from these data and analogous experiments at intermediate urea concentrations are plotted on a log scale ($\ln k_d$) vs urea concentration for these four situations in Figure 6D. Linear fits represent the data well in all cases, with intercepts (compared with WT RNAP-FL λP_R in Table 1) that agree within uncertainty with the experimentally determined dissociation rate constants in the absence of urea (Figure 5A).

Slopes obtained from linear fits ($d \ln k_d / d[\text{urea}]$) are listed in Table 1. For dissociation of WT-FL RP_o in PB, $d \ln k_d / d[\text{urea}] = 3.1 \text{ M}^{-1}$, the same within uncertainty as that reported

previously in TB at 17 °C (3.0 M^{-1}).⁷ For the three single and double variant open complexes formed with Δ JAW RNAP and/or DT+12 λP_R , $d \ln k_d / d[\text{urea}]$ is only 2/3 as large ($d \ln k_d / d[\text{urea}] = 1.9\text{--}2.1 \text{ M}^{-1}$). From eq 2, these urea derivatives of k_d are equal in magnitude and opposite in sign to the corresponding urea derivatives of K_3 , because the DNA closing rate constants k_{-2} are independent of urea concentration for WT³ and Δ JAW (E, Ruff, unpublished) RNAP. Deletion of the jaw, of the downstream duplex at +12, or of both, eliminates one-third of this urea effect.

Dissection of urea effects on the dissociation rate constant k_d for WT RNAP-FL λP_R open complexes reveals that two-thirds of it is in conversion of I_2 to the $(I_2 - I_3)^\ddagger$ transition state, with the other one-third in the conversion of this transition state to I_3 and then to RP_o .³ Both the urea results and the reductions in open complex lifetime observed for the jaw deletion and/or downstream DNA truncation indicate that by these criteria the Δ JAW-FL λP_R final open complex is more similar to $(I_2 - I_3)^\ddagger$ or I_3 than to RP_o . Comparison of the effects of urea on dissociation of open complexes formed by WT and Δ JAW RNAP at FL and DT+12 λP_R with these kinetic-mechanistic results argues strongly that the jaw assembles with other DMEs on the downstream duplex promoter DNA late in the mechanism of RP_o formation, most likely in the conversion of I_3 to RP_o , and that the final open complex formed by Δ JAW RNAP is a plausible model of the intermediate I_3 formed by WT RNAP at FL λP_R . Relative dissociation rate constants (Tables 1 and 2) are consistent with this conclusion, as are the different effects of salt concentration on dissociation rate constants of WT and Δ JAW open complexes, presented next.

Effects of Salt Concentration on Dissociation Kinetics of the Δ JAW RNAP–FL λP_R Open Complex are Consistent with an I_3 -like Open Complex. Figure 7 compares the effects of NaCl concentration on dissociation

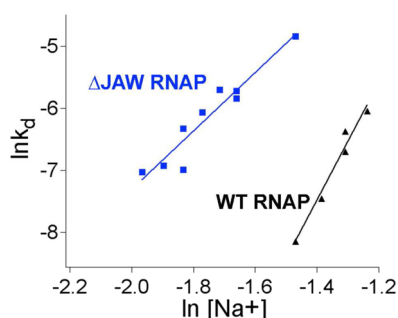


Figure 7. Comparison of effects of $[Na^+]$ on dissociation rate constants k_d of open complexes formed by Δ JAW (blue squares) and WT RNAP (black triangles)² at FL λP_R in PB, 25 °C. Slopes of these plots ($d \ln k_d / d \ln [Na^+]$) are listed in Table 3.

rate constants for WT-FL open complexes at 25 °C² and Δ JAW-FL open complexes at 25 °C in PB. Paralleling the reduced effect of urea on k_d for Δ JAW open complex dissociation (33% reduction compared to WT RNAP), dissociation of the Δ JAW open complex in PB is much less salt concentration dependent (50% reduction in $d \ln k_d / d \ln [Na^+]$; see Table 3) than dissociation of RP_o formed by WT

Table 3. Salt Effects on Dissociation Kinetics of Δ JAW and WT RNAP Open Complexes^a

RNAP	λP_R	relative dissociation rate ^b	$10^{-3} K_3$ (M ⁻¹) ^{b,c}	ΔG_3^0 (kcal/mol) ^d	$d \ln k_d / d \ln [Na^+]$ ^e
WT	FL	1	6.0 ± 1.0	-5.2 ± 0.3	9.5 ± 1.0
Δ JAW	FL	28 ± 3	0.21 ± 0.05	-3.2 ± 0.2	4.7 ± 0.7

^aConditions: PB, 25 °C. ^bAt 0.23 M Na^+ where, for WT RNAP, $k_d = (2.8 \pm 0.1) \times 10^{-4} s^{-1}$, interpolated from Figure 7. ^cCalculated using $k_{-2} = 1.7 \pm 0.4 s^{-1}$ at 25 °C, interpolated from results at 10 and 37 °C.³ ^dCalculated from K_3 . ^eDetermined from Figure 7.

RNAP at FL λP_R . While the previous kinetic data for WT RNAP did not allow the salt effect on k_d to be uniquely dissected into contributions from the individual steps of dissociation, as was done for the urea effect,³ the greatly reduced salt effect on dissociation of the Δ JAW open complex is completely consistent with the proposal that this complex is I_3 -like, with a salt dependence which is halfway between that of RP_o ($d \ln k_d / d \ln [Na^+] = 9$) and that of the initial open complex I_2 ($d \ln k_d / d \ln [Na^+] = 0$).

Early Steps of Δ JAW- λP_R Open Complex Formation.

We previously proposed that DMEs have large effects on early as well as late steps of open complex formation.^{5,6} Here, we use association kinetic studies to test our hypothesis that the jaw participates in the steps leading up to and/or including opening ($R + P \rightleftharpoons I_1 \rightarrow I_2$; Figure 1). For a series of concentrations of Δ JAW RNAP, all in excess over promoter DNA, time courses of open complex formation at FL and DT+12 λP_R at 37 °C in TB are plotted in Figure 8A,C. Analogous results for association of WT RNAP at DT+12 λP_R are plotted in Figure 8B. In all cases the kinetics are single-exponential, demonstrating that the initial binding step(s) ($R + P \rightleftharpoons I_1$) are in rapid equilibrium on the time scale of the subsequent slow step ($I_1 \rightarrow I_2$) involving DNA opening.²⁹

At all but the highest RNAP concentrations investigated, the kinetics of association of Δ JAW RNAP with both FL and DT+12 (Figure 8A,C) are reversible, forming equilibrium mixtures of open complexes and unbound promoter DNA. By contrast,

time courses of association of WT RNAP with FL λP_R ⁴ and with DT+12 λP_R (Figure 8B) are irreversible. Values of RNAP-concentration-dependent composite association rate constants α (eq 5) for Δ JAW RNAP at FL and DT+12 λP_R obtained as described in Experimental Procedures from the decay-to-equilibrium data in Figure 8A,C, are plotted vs concentration of Δ JAW RNAP in Figure 8D. Also plotted in Figure 8D are values of α for the association of WT RNAP with DT+12 λP_R (Figure 8B), and fitted curves summarizing the RNAP-concentration-dependences of α previously reported for association of WT RNAP with FL λP_R ⁴ and with upstream-truncated (UT-47) λP_R .¹⁹

From Figure 8D, rates of open complex formation from free reactants at all but the lowest RNAP concentrations are slower for all variants than for the WT-FL case. In particular, values of the rate constant α for the Δ JAW variant and either FL or DT+12 λP_R are as small as those observed previously for association of WT RNAP with UT-47 λP_R .¹⁹ Removing either the jaw or the upstream DNA reduces the rate of open complex formation to a similarly large extent. This finding supports our previous proposal that the upstream clamp, jaw, and other DMEs function as a relay system, in which the upstream clamp interacts with the far upstream DNA in the wrapped I_1 intermediate, and this interaction shifts the positions of the jaw and other DMEs to open a gate in the downstream cleft and allow entry of the downstream duplex prior to opening.^{5,6} Removing only the downstream DNA (i.e., open complex formation by WT RNAP with DT+12 λP_R) reduces α to a much smaller though still significant extent at higher RNAP concentrations.

Initial slopes of plots of α vs RNAP concentration (Figure 8D) are overall second-order rate constant for open complex formation k_a (eq 6). Values of k_a for WT RNAP and either FL or DT+12 λP_R are similar, and are approximately 10-fold larger than those for Δ JAW RNAP and either FL or DT+12 λP_R , or for WT RNAP and UT-47 λP_R . This large effect on k_a of deleting the jaw cannot be explained entirely by an interaction of RNAP with the downstream duplex in I_1 or the subsequent transition state, since the same behavior is observed with both FL and DT+12 λP_R . This is consistent with the deduction from the k_d data that the jaw deletion has large effects on the interaction of DMEs with the downstream duplex late in RP_o formation, in the conversion of I_3 to RP_o .

From eq 6, k_a is the product of the equilibrium constant K_1 for formation of the I_1 closed complex and the rate constant k_2 for the isomerization step, including DNA opening, that converts I_1 to I_2 . For situations where K_1 and the range of RNAP concentrations investigated are large enough, individual values of K_1 and k_2 are obtained from analysis of the dependence of α on $[RNAP]$ using eq 5. For open complex formation by WT RNAP, truncation of the downstream duplex (DT+12), like upstream truncation (UT-47), increases K_1 and reduces k_2 . For DT+12, these effects compensate, so the overall k_a is unaffected by downstream truncation; for UT-47, Davis et al. found that the reduction in k_2 was the dominant effect, greatly reducing k_a even though K_1 increased.¹⁹ For open complex formation by Δ JAW RNAP with either FL or DT+12 λP_R , the data do not permit dissection of k_a into K_1 and k_2 .

The finding that association rate constants k_a are the same for open complex formation by Δ JAW RNAP and FL or DT+12 λP_R promoter (Table 2) appears to contradict what is observed in Figure 8, where decay to equilibrium rates at a specified $[RNAP]$ are faster for Δ JAW-DT+12 complexes (Figure 8C)

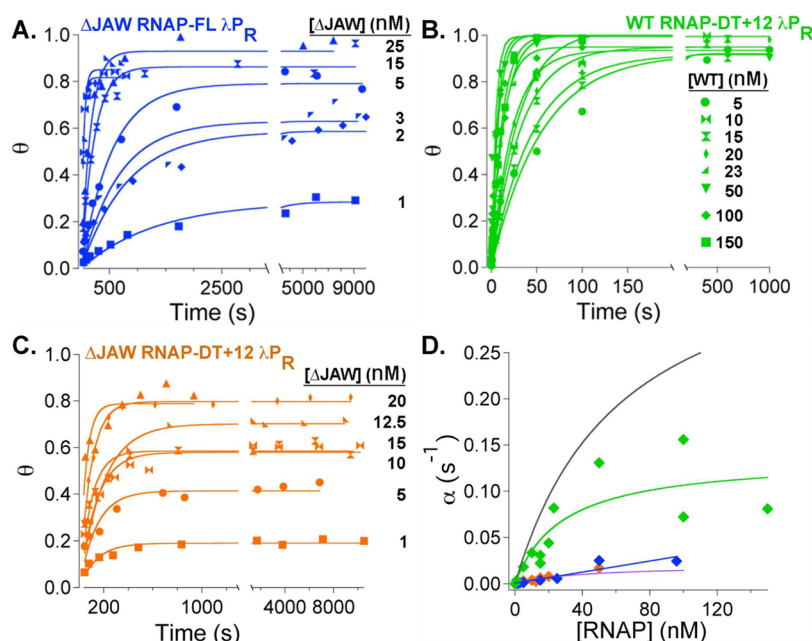


Figure 8. Comparison of association kinetics of Δ JAW and WT RNAP with FL and DT+12 λP_R promoter DNA in TB at 37 °C. Panels A–C display time courses of the fraction of λP_R promoter DNA in the form of heparin-resistant open complexes (θ) for different RNAP concentrations for (A) Δ JAW RNAP and FL λP_R promoter (blue), (B) WT RNAP and DT+12 λP_R promoter (green), and (C) Δ JAW RNAP and DT+12 λP_R promoter (orange). Curves are single exponential fits to eq 3, yielding decay to equilibrium rate constants β , from which the association rate constant α is obtained (eq 4). (D) Plots of α vs $[\text{RNAP}]$ for the data of panels A–C, compared with published fits for association of WT RNAP with FL (gray)⁴ and UT-47 λP_R (purple).¹⁹ The axes of the inset figure are the same as in the main panel. A hyperbolic fit (eq 5) to the WT-DT+12 data set (green curve) and a linear fit to the combined Δ JAW-FL and Δ JAW-DT+12 data set (blue line) are shown. Values obtained from these fits are reported in Table 2.

than for Δ JAW-FL complexes (Figure 8A). This is not a contradiction; the faster decay to equilibrium is entirely because k_d for dissociation of Δ JAW-DT+12 open complexes is larger than for Δ JAW-FL open complexes in TB (Table 2).

DISCUSSION

In fundamental and significant ways, the mechanism of forming and stabilizing the open RNAP- λP_R promoter complex resembles the mechanism of an enzyme-catalyzed reaction.^{1,5} First, in a series of steps that are rapidly reversible, like those that form a precatalytic complex between enzyme and substrates, RNAP binds the start-site region of duplex promoter DNA, bends it, and inserts it into the active site cleft.^{3,6,19} Then, in a slow (rate-determining) step occurring in the active site cleft, like the catalytic step of an enzyme-catalyzed reaction, RNAP opens the initiation bubble (−11 to +2) and appears to place the +1 template strand base in the active site.¹ Subsequent to opening, the downstream half of the melted nontemplate strand appears to be repositioned, and RNAP DMEs assemble and fold on the distal downstream duplex.^{1–3,7} For the conditions used to investigate dissociation of open complexes here, the steps of DME disassembly are rapidly reversible on the time scale of the rate determining DNA closing step. For both enzyme catalyzed reactions and transcription initiation, the early and late reversible steps are the targets of most regulatory factors and ligands.

In this study, we probed the role of RNAP interactions with the downstream DNA by removing either the DNA downstream beyond +12 or the jaw, which is thought to interact with the downstream DNA. Footprinting reveals that Δ JAW open complexes differ structurally from those formed by the WT RNAP: although the DNA is melted to the same extent, both

DNA strands are globally less protected from HO• attack, and the downstream boundary of HO• protection shortens by 5–10 bp. Dissociation kinetics of Δ JAW open complexes are faster and less dependent on urea concentration, whereas promoter binding-DNA opening is much slower than for WT RNAP. These results implicate the jaw role during both early and late steps in RP_o formation. By contrast, the downstream DNA appears to function primarily after DNA opening: while DT+12 truncation has the same effect on dissociation kinetics as the removal of the jaw, it does not affect association rate constant k_a for WT RNAP.

Consistent with the reported effects of the jaw deletion on the lifetime of the open complex,^{9,14} we deduce that the jaw is required for RNAP contacts with the distal downstream region (+10 to +20) of the promoter established in the last step of stabilization of the open complex (conversion of I_3 to RP_o). Strikingly, however, we also find that deletion of the jaw greatly reduces the rate of the early steps of initiation. We deduce that this effect is not simply due to loss of interactions with the downstream DNA; a loss of interactions of the jaw with in-cleft elements such as other DMEs and/or with region 1.1 of σ^{70} must also be involved.¹⁴ This interpretation is consistent with observations of Mekler et al. that gp2, a phage inhibitor that binds to the jaw, destabilizes RNAP/scaffold complexes even in the absence of downstream DNA.²⁴ In addition, results of HO• footprinting suggest that the jaw is required to seal the downstream end of the cleft in RP_o , eliminating water and forming a tight interface of RNAP-DNA backbone interactions that is highly protected from HO• attack. By contrast, the open DNA strands and surrounding duplex in the I_3 -like Δ JAW open complex appear to be more hydrated, as judged by their increased HO• reactivity. Expulsion of water during RP_o

formation presumably contributes significantly to the entropic driving force for the endothermic steps of opening and conversion of the initial open complex (I_2) to RP_o , as previously proposed.²

Our current and previous results indicate that assembly of DMEs on the downstream duplex stabilizes the initial open complex I_2 in converting it to RP_o .^{2,3} Analysis of RNAP binding to heteroduplex downstream junction templates with progressively shortened downstream DNA duplex revealed that many interactions contribute to this stabilizing effect.¹² The largest incremental effect was observed when the duplex was extended from +8 to +9. Structural studies suggest that DNA near +9 contacts the jaw directly.²⁵ This interaction likely plays a role in formation and stabilization of RP_o and is subject to control by gp2¹³ and likely other regulators.

However, structural analysis of RNAP shows that the jaw does not extend beyond +10, raising a question of how the far downstream protection is established. Collectively, the available data suggest that contacts between the jaw and promoter DNA set the trajectory for DNA entry into a trough formed by the jaw and clamp domains. Single-molecule FRET experiments demonstrate that the clamp is open in free RNAP but closes upon the formation of RP_o .³⁰ While the open state of the clamp permits DNA movement in and out of the cleft, interactions of DNA with RNAP trigger an inward 16° rotation of the clamp, effectively sealing the cleft in RP_o . The observed protection from HO• is likely due to a combination of the cleft closure and direct contacts between DMEs and DNA.

Cross-linking data support the existence of direct interactions between DNA backbone from +10 to +20 and a large fragment of the β' subunit (residues 821–1407), which includes the clamp, the jaw, and SI3.²⁶ Given the range of motions demonstrated³⁰ and proposed²² for these DMEs, we cannot pinpoint specific interactions of RNAP regions with each position in the promoter DNA. Furthermore, DMEs are proposed to function together in several relay systems (Figure 2),^{5,13} and changes in different elements may give rise to similar effects on transcription, e.g., in the case of the jaw and SI3, which share an extensive interface.³¹ Thus, it is difficult to distinguish between direct (mediated via DNA contacts) and indirect effects of deletions in individual DMEs. A more refined mapping of the DME contacts with the promoter DNA awaits high-resolution structural and cross-linking analysis of RP_o .

We propose that in Δ JAW RNAP-promoter complexes, the downstream DNA may not be properly positioned in the trough, and clamp closure is subsequently inhibited. The same effect is observed upon addition of gp2, which blocks DNA interactions with the jaw³⁰ and locks the clamp in the open state.¹³

The Role of the Jaw in Early (Pre-Opening) Steps. The profound effects of deleting the jaw on the early steps of open complex formation, observed for both FL and DT+12 promoter DNA, indicate that the jaw and/or neighboring regions affected by its deletion are a key part of the proposed relay system^{5,6} by which the upstream promoter sequence, transcription factors, ligands and solution variables work together to regulate the rate of forming the initial open complex (I_2), and thereby regulate the rate of transcription initiation. The limited evidence available to date suggests that much of this regulation occurs in the steps that put the downstream DNA duplex in the cleft. The subsequent opening step is relatively independent of solute concentration^{2,7} suggesting that, like catalytic steps of enzymes, it occurs within

a somewhat sealed cleft that is not readily accessible to solutes and regulatory ligands.

We have proposed that these early acting upstream regulatory factors have a common basis of action: together they set or modify the trajectory of far-upstream DNA in early, relatively unstable closed intermediate complexes (designated RP_c) in which the promoter DNA lies outside the cleft.⁵ The trajectory of far-upstream DNA determines its ability to interact with the upstream clamp, which in turn affects the positioning of the jaw and other DMEs.⁶ More specifically, we proposed that when −35 and UP elements are recognized by σ^{70} region 4 and the α CTDs (respectively), far-upstream DNA is bent and wrapped around the back of RNAP, in close proximity to the downstream end of the active site cleft and the upstream clamp. These interactions trigger release of DMEs from their blocking positions, opening the downstream gate to allow entry of the downstream duplex.

Interestingly, yeast RNAP II may also use a molecular relay to mediate loading of the duplex DNA into the active site cleft, which is obstructed in the free enzyme in all multisubunit RNAPs. The Pol II subunit Rpb4/7 (absent in bacterial RNAP) binds to the RNAP core, facilitating clamp closure.³² Recent studies postulate that subsequent binding of the multisubunit mediator complex attracts Rpb4/7, resulting in formation of the open clamp state and thus allowing for the entry of duplex DNA into the RNAP active site cleft.³³

The Role of the Jaw in Late (Post-Opening) Steps. The observed dissociation rate constant k_d of stable open complexes (mainly RP_o) to free promoter and RNAP provides information about the steps that convert RP_o back to I_2 and the rate of DNA closing ($I_2 \rightarrow I_1$) (Figure 1). The DNA closing rate (k_{-2}) is not a function of salt or urea concentration.³ Therefore, solute and salt effects on k_d arise from their effect on the equilibrium constant for conversion of I_2 to RP_o (K_3). For the conditions of our experiments, steps by which RP_o converts to I_2 rapidly equilibrate on the time scale of DNA closing ($I_2 \rightarrow I_1$).²⁹ We proposed that the conversion of I_2 to RP_o involves assembly of DMEs on the downstream duplex, and movement of the downstream portion of the nontemplate strand in the cleft.^{1–3} Our current studies test these proposals by deleting the jaw and/or the downstream duplex (DT+12).

Kontur et al. concluded from the urea effect on k_d that approximately 100–150 residues of RNAP are buried by folding/assembly in the steps converting the initial open complex I_2 to the final open complex RP_o (Figure 2).³ Roughly two-thirds of this urea effect occurs in the conversion of I_2 to the subsequent (I_2 – I_3)[‡] transition state, and the remaining 40 residues are buried in the conversion of this transition state to RP_o . In the current study, we find that, in contrast to WT enzyme, the Δ JAW variant does not protect the distal downstream DNA from HO• attack. The effect of urea on dissociation of Δ JAW RNAP from either FL or DT+12 λP_R and of WT from DT+12 λP_R indicates that ~40 fewer residues are buried in the open complex formed by these variants. PONDR predicts that 33 residues near the jaw are unstructured in RNAP (Figure 2B). The simplest explanation of these observations is that the jaw folds and assembles with other DMEs on the distal promoter DNA in the final step of open complex formation (Figure 2D), and that the open complex formed by Δ JAW may therefore be a stable mimic of the relatively unstable I_3 intermediate open complex formed by WT RNAP on the path to RP_o at λP_R .

What events (involving burial of approximately 80 residues by folding/assembly) occur in the conversion of I_2 to I_3 ? We propose that folding and assembly of regions of the β' clamp and/or the β lobe, DMEs proposed to interact with the proximal downstream DNA duplex (+3 to +9),⁹ occur in this step (Figure 2C). PONDR predicts that 99 residues of the β' clamp and 57 residues of the β lobe are unstructured in RNAP. Formation of the proximal downstream contacts would logically precede assembly of the jaw and/or other DMEs to the distal downstream duplex, as stability increases in the conversion of I_2 to RP_o . Experiments are in progress to test these predictions for interactions of the DMEs with downstream DNA in I_2 and I_3 , and to determine the effect of promoter sequence and interactions in the cleft on the extent of assembly and tightening of the jaw and other DMEs.

Our studies reveal the existence of three types of “open” promoter complexes that differ in interactions between the downstream duplex DNA with DMEs. We hypothesize that the distribution among these complexes differs at different cellular promoters, underlying their regulation by transcription factors such as DksA, which allosterically alters RNAP contacts with DNA downstream from the active site.³⁴ We propose that this mode of regulation may be utilized by eukaryotic RNAPs, which have structural elements that are homologous (β' clamp, β lobe) or functionally analogous (β' jaw) to bacterial DMEs.³⁵ Furthermore, similar to our footprinting results obtained at λP_R , recent single-molecule studies of yeast RNAP II open complexes revealed very different positions of the downstream DNA,³⁶ supporting our idea that promoter DNA/RNAP interactions in this region are highly dynamic and suggesting that they may contribute to regulation of eukaryotic gene expression.

■ ASSOCIATED CONTENT

■ Supporting Information

PONDR predictions of naturally disordered regions of the DMEs. Figure S1: Representative HO• footprinting gels; Figure S2: Dissociation of Δ JAW-FL and WT-DT+12 open complexes are not heparin-dependent. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: (608) 262-5332; fax: (608) 262-3453; e-mail: mtrecord@wisc.edu.

Present Address

#Department of Chemistry, Beloit College, 700 College St., Beloit, WI 53511.

Funding

This work was supported by NIH Grants GM023467 and GM047022 to M.T.R., NSF Grant MCB-0949569 to I.A., grants from the BBSRC and Wellcome Trust to S.W., and by UW-Madison. S.W. is a recipient of a BBSRC David Phillips Fellowship (BB/E023703).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge numerous contributions of Dr. Ruth Saecker to the design of this research and comments on the manuscript. We also thank Dr. Wilma Ross for extensive and very helpful comments on an earlier draft of this manuscript.

Finally, we thank the reviewers and editor for their careful reading of the manuscript, comments, and constructive criticism.

■ ABBREVIATIONS USED

RNAP, RNA polymerase; WT, wild-type RNAP; DME, downstream mobile element; Δ JAW, deletion of the RNAP jaw ($\Delta\beta'$ 1149–1190); SI, sequence insertion; FL, full-length λP_R promoter DNA; DT+12, truncation of λP_R promoter DNA downstream at +12; UT-47, truncation of λP_R promoter DNA upstream at -47; TB, Tris buffer; LSTB, low salt Tris buffer; PB, phosphate buffer

■ REFERENCES

- (1) Gries, T. J., Kontur, W. S., Capp, M. W., Saecker, R. M., and Record, M. T., Jr. (2010) One-step DNA melting in the RNA polymerase cleft opens the initiation bubble to form an unstable open complex. *Proc. Natl. Acad. Sci. U. S. A.* 107, 10418–10423.
- (2) Kontur, W. S., Capp, M. W., Gries, T. J., Saecker, R. M., and Record, M. T., Jr. (2010) Probing DNA binding, DNA opening, and assembly of a downstream clamp/jaw in *Escherichia coli* RNA polymerase- λP_R promoter complexes using salt and the physiological anion glutamate. *Biochemistry* 49, 4361–4373.
- (3) Kontur, W. S., Saecker, R. M., Capp, M. W., and Record, M. T., Jr. (2008) Late steps in the formation of *E. coli* RNA polymerase- λP_R promoter open complexes: Characterization of conformational changes by rapid [perturbant] upshift experiments. *J. Mol. Biol.* 376, 1034–1047.
- (4) Saecker, R. M., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr, Capp, M. W., and Record, M. T., Jr. (2002) Kinetic studies and structural models of the association of *E. coli* σ^{70} RNA polymerase with the λP_R promoter: Large scale conformational changes in forming the kinetically significant intermediates. *J. Mol. Biol.* 319, 649–671.
- (5) Saecker, R. M., Record, M. T., Jr, and 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.
- (6) Davis, C. A., Bingman, C. A., Landick, R., Record, M. T., Jr, and Saecker, R. M. (2007) Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* 104, 7833–7838.
- (7) Kontur, W. S., Saecker, R. M., Davis, C. A., Capp, M. W., and Record, M. T., Jr. (2006) Solute probes of conformational changes in open complex (RP_o) formation by *Escherichia coli* RNA polymerase at the λP_R promoter: Evidence for unmasking of the active site in the isomerization step and for large-scale coupled folding in the subsequent conversion to RP_o . *Biochemistry* 45, 2161–2177.
- (8) Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J., and Dunker, A. K. (2001) Sequence complexity of disordered protein. *Proteins* 42, 38–48.
- (9) Ederth, J., Artsimovitch, I., Isaksson, L. A., and Landick, R. (2002) The downstream DNA jaw of bacterial RNA polymerase facilitates both transcriptional initiation and pausing. *J. Biol. Chem.* 277, 37456–37463.
- (10) Craig, M. L., Suh, W. C., and Record, M. T., Jr. (1995) HO• and DNase I probing of *E. coli* RNA polymerase- λP_R promoter open complexes: Mg^{2+} binding and its structural consequences at the transcription start site. *Biochemistry* 34, 15624–15632.
- (11) Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S. A. (2000) A structural model of transcription elongation. *Science* 289, 619–625.
- (12) Mekler, V., Minakhin, L., and Severinov, K. (2011) A critical role of downstream RNA polymerase-promoter interactions in the formation of initiation complex. *J. Biol. Chem.* 286, 22600–22608.
- (13) James, E., Liu, M., Sheppard, C., Mekler, V., Camara, B., Liu, B., Simpson, P., Cota, E., Severinov, K., Matthews, S., and

Wigneshweraraj, S. (2012) Structural and mechanistic basis for the inhibition of *Escherichia coli* RNA polymerase by T7 Gp2. *Mol. Cell* 47, 755–766.

(14) Wigneshweraraj, S. R., Burrows, P. C., Severinov, K., and Buck, M. (2005) Stable DNA opening within open promoter complexes is mediated by the RNA polymerase β' -jaw domain. *J. Biol. Chem.* 280, 36176–36184.

(15) Camara, B., Liu, M., Reynolds, J., Shadrin, A., Liu, B., Kwok, K., Simpson, P., Weinzierl, R., Severinov, K., Cota, E., Matthews, S., and Wigneshweraraj, S. R. (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.

(16) Burgess, R. R., and Jendrisak, J. J. (1975) A procedure for the rapid, large-scale purification of *Escherichia coli* DNA-dependent RNA polymerase involving polymin P precipitation and DNA-cellulose chromatography. *Biochemistry* 14, 4634–4638.

(17) Mangiarotti, L., Cellai, S., Ross, W., Bustamante, C., and Rivetti, C. (2009) Sequence-dependent upstream DNA-RNA polymerase interactions in the open complex with λP_R and λP_{RM} promoters and implications for the mechanism of promoter interference. *J. Mol. Biol.* 385, 748–760.

(18) Roe, J. H., Burgess, R. R., and Record, M. T., Jr. (1984) Kinetics and mechanism of the interaction of *Escherichia coli* RNA polymerase with the λP_R promoter. *J. Mol. Biol.* 176, 495–522.

(19) Davis, C. A., Capp, M. W., Record, M. T., Jr., and Saecker, R. M. (2005) The effects of upstream DNA on open complex formation by *Escherichia coli* RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* 102, 285–290.

(20) Craig, M. L., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr., Capp, M. W., Saecker, R. M., and Record, M. T., Jr. (1998) DNA footprints of the two kinetically significant intermediates in formation of an RNA polymerase-promoter open complex: Evidence that interactions with start site and downstream DNA induce sequential conformational changes in polymerase and DNA. *J. Mol. Biol.* 283, 741–756.

(21) Tullius, T. D., Dombroski, B. A., Churchill, M. E., and Kam, L. (1987) Hydroxyl radical footprinting: A high-resolution method for mapping protein-DNA contacts. *Methods Enzymol.* 155, 537–558.

(22) 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.

(23) Sasse-Dwight, S., and Gralla, J. D. (1989) $KMnO_4$ as a probe for lac promoter DNA melting and mechanism in vivo. *J. Biol. Chem.* 264, 8074–8081.

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

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

(26) Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V., and Ebright, R. H. (2000) Structural organization of the RNA polymerase-promoter open complex. *Cell* 101, 601–611.

(27) Suh, W. C., Leirimo, S., and Record, M. T., Jr. (1992) Roles of Mg^{2+} in the mechanism of formation and dissociation of open complexes between *Escherichia coli* RNA polymerase and the λP_R promoter: Kinetic evidence for a second open complex requiring Mg^{2+} . *Biochemistry* 31, 7815–7825.

(28) Hong, J., Capp, M. W., Saecker, R. M., and Record, M. T., Jr. (2005) Use of urea and glycine betaine to quantify coupled folding and probe the burial of DNA phosphates in lac repressor-lac operator binding. *Biochemistry* 44, 16896–16911.

(29) Tsodikov, O. V., and Record, M. T., Jr. (1999) General method of analysis of kinetic equations for multistep reversible mechanisms in the single-exponential regime: Application to kinetics of open complex formation between $E\sigma^{70}$ RNA polymerase and λP_R promoter DNA. *Biophys. J.* 76, 1320–1329.

(30) Chakraborty, A., Wang, D., Ebright, Y. W., Korlann, Y., Kortkhonja, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., Nixon, B. T., Knight, J., Weiss, S., and Ebright, R. H. (2012) Opening and closing of the bacterial RNA polymerase clamp. *Science* 337, 591–595.

(31) Conrad, T. M., Frazier, M., Joyce, A. R., Cho, B. K., Knight, E. M., Lewis, N. E., Landick, R., and Palsson, B. O. (2010) RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. *Proc. Natl. Acad. Sci. U. S. A.* 107, 20500–20505.

(32) Armache, K. J., Kettenberger, H., and Cramer, P. (2003) Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A.* 100, 6964–6968.

(33) Cai, G., Imasaki, T., Yamada, K., Cardelli, F., Takagi, Y., and Asturias, F. J. (2010) Mediator head module structure and functional interactions. *Nat. Struct. Mol. Biol.* 17, 273–279.

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

(35) Werner, F., and Grohmann, D. (2011) Evolution of multi-subunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* 9, 85–98.

(36) Treutlein, B., Muschielok, A., Andrecka, J., Jawhari, A., Buchen, C., Kostrewa, D., Hog, F., Cramer, P., and Michaelis, J. (2012) Dynamic architecture of a minimal RNA polymerase II open promoter complex. *Mol. Cell* 46, 136–146.

(37) Feklistov, A., and Darst, S. A. (2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* 147, 1257–1269.

Overexpression of *Escherichia coli* *udk* mimics the absence of T7 Gp2 function and thereby abrogates successful infection by T7 phage

Andrey Shadrin,¹ Carol Sheppard,¹ Dhruti Savalia,²
Konstantin Severinov^{2,3,4} and Sivaramesh Wigneshweraraj¹

Correspondence

Sivaramesh Wigneshweraraj
s.r.wig@imperial.ac.uk

¹Section of Microbiology, Faculty of Medicine and MRC Centre for Molecular Bacteriology and Infection, Imperial College 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

³Gene Biology of the Russian Academy of Sciences, Moscow, Russia

⁴Molecular Genetics of the Russian Academy of Sciences, Moscow, Russia

Successful infection of *Escherichia coli* by bacteriophage T7 relies upon the transcription of the T7 genome by two different RNA polymerases (RNAPs). The bacterial RNAP transcribes early T7 promoters, whereas middle and late T7 genes are transcribed by the T7 RNAP. Gp2, a T7-encoded transcription factor, is a 7 kDa product of an essential middle T7 gene, 2, and is a potent inhibitor of the host RNAP. The essential biological role of Gp2 is to inhibit transcription of early T7 genes that fail to efficiently terminate in order to facilitate the coordinated usage of the T7 genome by both host and phage RNAPs. Overexpression of the *E. coli* *udk* gene, which encodes a uridine/cytidine kinase, interferes with T7 infection. We demonstrate that overexpression of *udk* antagonizes Gp2 function in *E. coli* in the absence of T7 infection and thus independently of T7-encoded factors. It seems that overexpression of *udk* reduces Gp2 stability and functionality during T7 infection which consequently results in inadequate inhibition of host RNAP and in the accumulation of early T7 transcripts. In other words, overexpression of *udk* mimics the absence of Gp2 during T7 infection. Our study suggests that the transcriptional regulation of the T7 genome is surprisingly complex and might potentially be affected at many levels by phage- and host-encoded factors.

Received 31 October 2012

Revised 5 December 2012

Accepted 6 December 2012

INTRODUCTION

The *Escherichia coli* bacteriophage T7 relies upon the multisubunit RNA polymerase (RNAP) of its host and a single subunit RNAP encoded by its own genome for transcription of its genes. During the infection process, the host RNAP transcribes early T7 phage genes from the genetic left end of the viral genome, whereas the T7 RNAP transcribes the middle and late T7 phage genes. Whilst activities of both the host and T7 RNAP on the T7 phage genome are likely to be regulated and co-ordinated at multiple levels involving host and phage factors, it is generally accepted that the products of T7 phage early gene 0.7 and middle gene 2 are responsible for inhibition of host RNAP at later stages of infection such that host resources are shifted towards the production of viral transcripts synthesized by the T7 RNAP. T7 Gp0.7 is a serine/

threonine kinase that phosphorylates a specific threonine residue (T1068) in the evolutionarily variable region (β' GNCD/I6 domain) of the host RNAP, which results in increased ρ -independent termination during transcription of host genes (Severinova & Severinov, 2006). T7 Gp2 is a 7 kDa protein that binds to the structurally conserved β' jaw domain of the host RNAP that is adjacent to the β' GNCD/I6 domain. Gp2 is a potent inhibitor of the *E. coli* RNAP and antagonizes conformational changes in the RNAP and interactions between the RNAP and the promoter DNA that are necessary for the engagement of the RNAP with the promoter DNA for transcription initiation (James *et al.*, 2012; Mekler *et al.*, 2011).

Unlike Gp0.7, Gp2 is essential for T7 phage growth in *E. coli*. Recently, it was demonstrated that the essential function of Gp2 during T7 phage infection is to abolish transcription from the strong host RNAP-dependent early T7 phage promoters A0–A3 (Savalia *et al.*, 2010). Down-mutations in the T7 A3 promoter compensate for the

Abbreviation: RNAP, RNA polymerase.

Supplementary methods and two supplementary figures are available with the online version of this paper.

absence of Gp2 (Savalia *et al.*, 2010). Likewise, mutations in *boxA* anti-termination element located downstream of early T7 phage promoters also make Gp2 dispensable (Savalia *et al.*, 2010). It therefore follows that, in the absence of Gp2, T7 A0-A3 promoter-initiated and *boxA* anti-terminated transcription by host RNAP interferes with T7 phage infection. The exact mechanism of this interference remains unknown.

Qimron *et al.* (2006) showed that mutant T7Δ0.7 phage, and to a relatively lesser extent wild-type T7 phage, is unable to grow on wild-type *E. coli* cells overexpressing *udk*, a uridine/cytidine kinase, which converts uridine and cytidine into the ribonucleoside monophosphates UMP and CMP, respectively. More recently, Qimron *et al.* (2008) identified missense mutations in gene 2 and gene 3.5 (which encodes T7 lysozyme – an inhibitor of the T7 RNAP) that enable T7Δ0.7 mutant and wild-type T7 phage growth on *E. coli* overexpressing *udk*. Overproduction of Gp2 was sufficient to suppress the growth defect of T7Δ0.7 mutant and wild-type T7 phage in *udk* overexpressing *E. coli* cells (Qimron *et al.*, 2008). Thus, it appears that during T7 phage infection of *udk* overexpressing *E. coli* cells, Gp2 is not able to adequately inhibit the host RNAP (Qimron *et al.*, 2008). However, whether *udk* overexpression directly affects Gp2 function or is mediated via an unidentified T7 phage and/or host factor(s) during T7 phage infection remains unknown. In addition, the consequence of *udk* overexpression on the biological role of Gp2 during T7 phage infection remains uncharacterized. Here, we present results from experiments in which we have addressed these issues.

METHODS

Plasmids. Plasmid pAS2:Gp2 and pAS2:CR44b_13 was constructed by digesting pSW33:Gp2 (Cámara *et al.*, 2010) and pAS33:CR44b_13 (Shadrin *et al.*, 2012) with *Xba*I and *Hind*III restriction endonucleases and the resulting 0.3 kb fragment containing T7 gene 2 and CR44b_13 was cloned into the same sites in pBAD33 (ATCC). pAS3 was constructed from pQE30 (Qiagen) by digesting with *Bsm*I and further circularization of the largest fragment. The resulting pAS3 plasmid is distinct from pQE30 only by 407 bp (*Bsm*I–*Bsm*I deletion) in the coding region of the chloramphenicol acetyltransferase gene. To make pAS3:*udk*, *E. coli udk* gene was amplified by PCR using primers 5'-TTTGGATCCATGACTGATCAGTCTCACCAGTG-3' and 5'-ATCAAGCITTATTCAAAGAACTGACTTATTTTCGCT-3' (italic underlined type indicates restriction enzyme sites), digested with *Bam*HI and *Hind*III and cloned into the same sites of pAS3 vector.

Bacterial growth attenuation assays. These assays were conducted in 250 ml flasks in a water bath at 37 °C with agitation at 180 r.p.m. Each flask contained 25 ml Luria broth (LB) growth medium supplemented with 30 mg chloramphenicol l⁻¹ (to select for plasmids pAS2:Gp2 and pAS2:CR44b_13) and 100 mg ampicillin l⁻¹ (to select for plasmids pAS3 and pAS3:*udk*), which was inoculated with 0.25 ml overnight culture grown in the same medium additionally supplemented with 0.5 % glucose to prevent leaky expression of proteins. The overexpression of *udk* was induced by adding IPTG to 1 mM final concentration when the culture reached OD₆₀₀~0.15. Thirty-five minutes after induction with IPTG the cells were pelleted and

resuspended in the same volume of the medium containing antibiotics. At this point a sample corresponding to *t*=0 was taken and Gp2 synthesis was induced by adding L⁺-arabinose solution to a final concentration of 0.1 % (w/v). For the calculation of Gp2 stability, a second washing step (as above) was performed at *t*=1.5 h to remove arabinose. At least three biological replicates were performed for each growth curve shown in Figs 1, 2 and S1, available with the online version of this paper. 2

Western blotting. Western blots were conducted exactly as described by Shadrin *et al.* (2012) but using polyclonal antibodies against Gp2 (Eurogentec) and using anti-rabbit-horseradish peroxidase-conjugated antibodies (Dakocytomation) as secondary antibodies. Western blots were conducted for two samples obtained from two of the three biological replicates from the growth curve experiments (see above) and the best representative blot is shown. The Gp2 half-life was calculated by measuring the relative intensities of the Gp2 signal from the Western blots assuming that the Gp2 degradation begins immediately after removing arabinose.

Primer extension analysis of early T7 transcripts. The reactions were conducted essentially as described by Savalia *et al.* (2010) but in *E. coli* strain BL21 containing the *udk* overexpression ASKA plasmid pCA24N:*udk* (Kitagawa *et al.*, 2005) in the presence of 0.5 mM IPTG.

RESULTS

Overexpression of *udk* antagonizes Gp2 activity in *E. coli* in the absence of T7 phage infection

We recently described a simple bacterial growth attenuation assay to study the activity of wild-type and mutant Gp2 forms *in vivo* in the absence of a T7 phage infection (Shadrin *et al.*, 2012). In this assay, plasmid-borne Gp2 is expressed from the arabinose-inducible *araBAD* promoter in *E. coli* strain MG1655. The induction of Gp2 expression with arabinose results in efficient attenuation of bacterial growth. We adapted this assay to find out if overexpression of *udk* can antagonize Gp2 activity in the absence of a T7 phage infection. In our experimental setup, *E. coli* strain MG1655 contained two compatible plasmids: pAS2:Gp2 in which gene 2 is placed under the control of the arabinose-inducible *araBAD* promoter and either pAS3:empty or pAS3:*udk* in which *udk* is placed under the control of the IPTG-inducible T5 promoter. The *E. coli* cells were grown for 1 h in LB at 37 °C to OD₆₀₀~0.15 and *udk* overexpression was induced with 1 mM IPTG for 45 min. The cells were then washed with fresh LB media and resuspended in LB containing 0.1 % w/v L⁺-arabinose to induce Gp2 expression and allowed to grow at 37 °C. As shown in Fig. 1(a), the expected attenuation of bacterial growth upon induction of Gp2 in the absence of *udk* overexpression was detected under our experimental conditions (compare solid and dotted red lines). Overexpression of *udk*, in the absence of Gp2 expression did not lead to attenuation of bacterial growth (Fig. 1a, dotted blue line). However, we failed to detect the expected attenuation of bacterial growth upon Gp2 induction in cells overexpressing *udk* (Fig. 1a, solid blue line). Similar results were obtained when experiments were repeated in

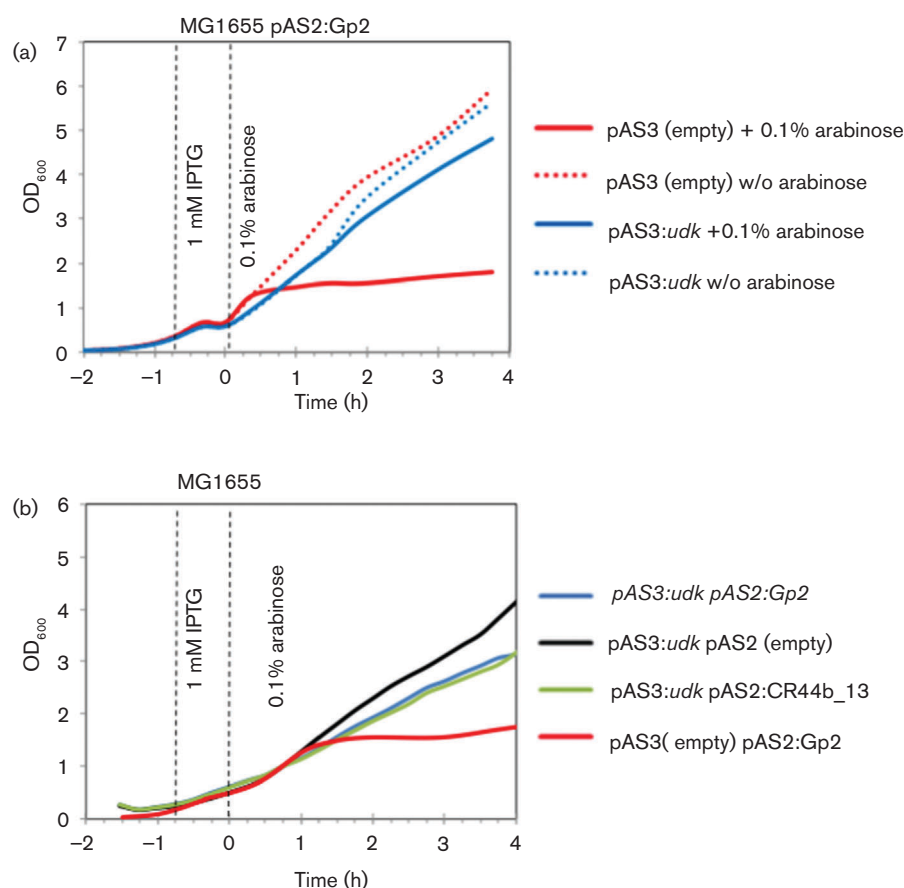


Fig. 1. Overexpression of *udk* antagonizes Gp2 activity in *E. coli* in the absence of T7 phage infection. (a) Growth curves of *E. coli* MG1655 containing different plasmid combinations and growth conditions (as indicated). (b) Growth curves of *E. coli* cells containing pAS2:CR44b_13 encoding Gp2-like protein CR44b_13 from CR44b phage, and with different plasmid combinations.

the context of the Gp2-like protein from *Citrobacter rodentium* (CR44b_13) which shares 47 % identity with Gp2 that also binds to the β' jaw domain of the *E. coli* RNAP and inhibits transcription initiation (Fig. 1b, c) (Shadrin *et al.*, 2012). Overall, we conclude that the overexpression of *udk* antagonizes Gp2 function in *E. coli* in the absence of any other T7 phage encoded factor(s). In other words, it seems that the antagonistic effect of *udk* overexpression on Gp2 activity in *E. coli* during T7 phage infection occurs independently of any other T7 phage-encoded factors.

Overexpression of *udk* in *E. coli* reduces Gp2 stability and functionality in the absence of T7 phage infection

We extended the Gp2-mediated growth attenuation assay to monitor accumulation and stability of Gp2 in the absence and presence of overexpressed Udk. The experiment shown in Fig. 2 was setup exactly as in Fig. 1, but the *E. coli* cells were subjected to additional washing step in

fresh LB to remove arabinose that is used to induce Gp2 expression. Samples were taken at $t=0$, 0.5, 1 and 1.5 h to monitor Gp2 accumulation post-induction and at $t=1.8$, 2.3 and 2.8 h to monitor Gp2 stability after removal of arabinose (note that all time points are relative to $t=0$ when Gp2 expression was induced). Whole-cell extract from $\sim 10^8$ *E. coli* cells from each sample was prepared and analysed by Western blotting using polyclonal antibodies against Gp2. As shown in Fig. 2, increased (by $\sim$ four- to fivefold) accumulation of Gp2 was observed in whole-cell extracts prepared from *udk* overexpressing *E. coli* cells when compared with whole-cell extracts prepared from *udk* non-overexpressing *E. coli* cells. This is as expected because induction of Gp2 expression will eventually lead to inhibition of transcription of gene 2 from the pAS2:Gp2 by the *E. coli* RNAP, which, in the presence of *udk* overexpression will not occur (see above). However, the relative stability of Gp2 in whole-cell extracts prepared from *udk* overexpressing *E. coli* cells appears to be significantly reduced upon removal of the inducer (arabinose) when compared with the stability of Gp2 in whole-cell extracts

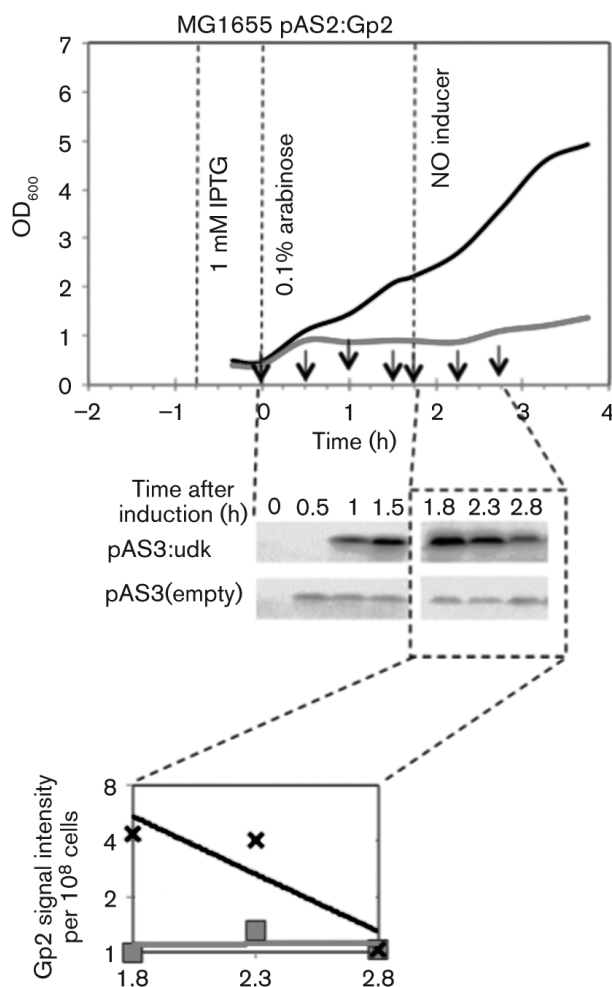


Fig. 2. Overexpression of *udk* in *E. coli* reduces Gp2 stability and functionality in the absence of T7 phage infection. As in Fig. 1, growth curves (top panel) of *E. coli* MG1655 containing pAS2:Gp2 and pAS3:empty (grey line) and pAS2:Gp2 and pAS3:*udk* (black line). Samples for preparing whole-cell extracts were taken at the indicated time points and analysed by Western blotting (middle panel) and signal intensity (relative to Gp2 signal at $t=1.8$ in whole-cell extracts from cells containing pAS2:Gp2 and pAS3:empty) on the Western blot was quantified and plotted on a graph (bottom panel).

prepared from *udk* non-overexpressing *E. coli* cells (Fig. 2). Under our experimental conditions, the estimated half-life of Gp2 in *udk* overexpressing *E. coli* cells is ~49 min whereas it exceeds 3 h in *udk* non-expressing *E. coli* cells. Intriguingly, the relative amount of total Gp2 molecules in whole-cell extracts from *udk* overexpressing and non-expressing cells at $t=2.8$ (i.e. 1 h after removal of the inducer) is comparable, even though Gp2 is unable to attenuate *E. coli* cell growth under *udk* overexpressing conditions. Since the binding of Gp2 and the promoter DNA downstream of the transcription start site to the β' jaw domain (which is necessary for transcription initiation

to occur) are mutually exclusive events (Mekler *et al.*, 2011), this observation suggests that Gp2 is somehow prevented from interacting with the RNAP under *udk* overexpressing conditions. Overall, it seems that *udk* overexpression is likely to result in a condition that mimics the absence of Gp2 function in *E. coli*.

Overexpression of *udk* mimics the absence of Gp2 function during T7 phage infection

The essential biological role for Gp2 during T7 infection of *E. coli* is the prevention of anti-terminated transcription from the early T7 promoters A0-A3 (Savalia *et al.*, 2010) into regions of the T7 phage genome normally transcribed by the T7 RNAP. Inspired by the finding that overexpression of *udk* could create a condition that mimics the absence of Gp2 in *E. coli*, we set out to determine the effect of *udk* overexpression on Gp2 functionality during T7 phage infection. As shown in Fig. 3 (inset) and consistent with the observation made by Savalia *et al.* (2010), in the course of wild-type T7 phage infection of *E. coli*, transcription from early T7 promoters A0-A3 by the host RNAP is eliminated shortly after initiation of transcription by the T7 RNAP (Savalia *et al.*, 2010). In contrast, when *E. coli* is infected by T7 carrying an amber mutation in gene 2, host RNAP transcription from early phage promoters continues late in infection (Savalia *et al.*, 2010). A similar effect is observed when wild-type T7 infects a host bacterium whose RNAP lacks the β' jaw domain (Savalia *et al.*, 2010). In agreement with the results shown in Figs 1 and 2, transcripts from early T7 promoters accumulated continuously throughout T7 phage infection in cells overexpressing *udk*. This result thus clearly indicates that overexpression of *udk* effectively prevents Gp2 function *in vivo* during T7 phage infection. In other words, overexpression of *udk* mimics the absence of Gp2 during T7 phage infection and leads to accumulation of anti-terminated transcription from early T7 promoters, which as a consequence abrogates successful T7 phage gene expression and infection.

DISCUSSION

In this study we demonstrate that overexpression of *E. coli udk* allows transcription from early T7 promoters to continue late in infection, thus mimicking infections that occur in the absence of Gp2 function i.e. when *E. coli* cells are infected by gene 2 mutant T7 phage or cells whose RNAP carries lesions in the Gp2-binding β' jaw domain is infected by the wild-type T7 phage (Savalia *et al.*, 2010). It therefore follows that *udk*, when overexpressed, obliterates the essential *in vivo* function of Gp2, i.e. inhibition of transcription from early T7 phage promoters (Savalia *et al.*, 2010). The mechanism by which overexpression of *udk* obliterates Gp2 function remains largely elusive; yet, our results indicate that it is unlikely to involve any T7 phage-encoded factors. However, we cannot exclude the possibil-

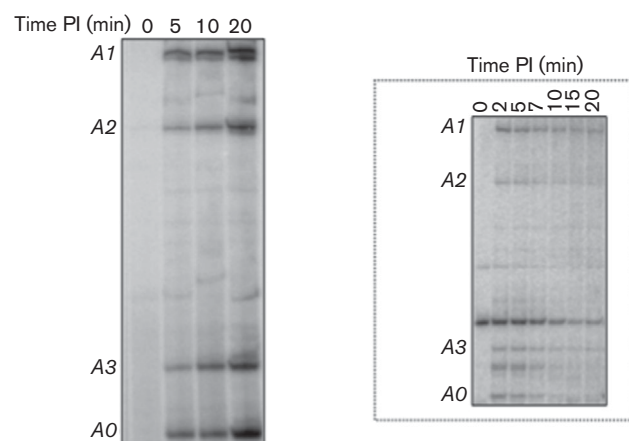


Fig. 3. Overexpression of *E. coli* *udk* mimics the absence of Gp2 during T7 phage infection. Autoradiograph (8%, w/v) showing transcripts originating from the T7 A0, A1, A2 and A3 promoters observed by primer extension analysis of RNA isolated from T7 phage infected *E. coli* cells under *udk* overexpressing conditions at the indicated times post-infection (PI). The inset shows the same experiment conducted with RNA isolated from T7 infected *E. coli* cells in the absence of *udk* overexpression.

ity of the involvement of host factors that could affect Gp2 functionality during T7 phage infection under *udk* overexpressing conditions. This view is consistent with the observation that purified recombinant Udk has no detectable effect on *E. coli* RNAP transcription from the early T7 phage promoter T7 A1 on its own and does not influence the ability of Gp2 to either bind to or inhibit the *E. coli* RNAP *in vitro* (Fig. S2a, b). Further, since Udk converts cytidine and uridine into their respective ribonucleoside monophosphates, CMP and UMP, we considered whether the significantly increased intracellular concentration of CMP and GMP under *udk* overexpressing conditions could potentially affect the ability of Gp2 to bind to and inhibit the *E. coli* RNAP. Thus, we investigated the ability of Gp2 to inhibit transcription initiation by the *E. coli* RNAP in the presence of CMP and GMP. As shown in Fig. S2(c), Gp2 is able to inhibit transcription initiation from the *lacUV5* promoter equally well in the absence and presence of CMP and GMP. However, we note that overexpression of *udk* appears to significantly reduce the half-life of Gp2 in *E. coli* when compared with the half-life of Gp2 under *udk* non-overexpressing conditions. This could indicate that Gp2 is somehow prevented from interacting with the RNAP under *udk* overexpressing conditions. We propose that RNAP-unbound Gp2 is more susceptible to proteolysis under *udk* overexpressing conditions than Gp2 that is bound to the RNAP and thus relatively better protected from proteolysis. In support of the latter, Gp2 is relatively less stable (estimated half-life of 16 min) in mutant *E. coli* cells JE1134 (Ederth *et al.*, 2002) containing mutant RNAP without the β' jaw domain (this is the binding site for Gp2) (Fig. S1) than in wild-type *E.*

coli cells (estimated half-life of >3 h; Fig. 2). Furthermore, the view that Gp2 is somehow prevented from interacting with the RNAP under *udk* overexpressing conditions is consistent with the observation made by Qimron *et al.* (2008) demonstrating that the overexpressed *udk*-mediated antagonist effect on Gp2 function during T7 phage infection can be overcome by supplying excess plasmid-borne Gp2.

The fact that we detect continuous accumulation of transcription from early T7 phage promoters under *udk* overexpressing conditions during T7 infection is thus consistent with the view that overexpression of *udk* creates a condition that mimics the absence of Gp2 function. This provides experimental evidence in support of the model proposed by Qimron *et al.* (2008), which predicts that under *udk* overexpressing conditions, transcription by the 'slow-elongating' host RNAP from the early T7 promoters into regions of the T7 genome normally transcribed by the 'fast-elongating' T7 RNAP, causes the latter to pause, which in turn results in undue recruitment of T7 genome packaging factors Gp18 and Gp19 and thus the production of less than unit length phage genomes, which as a consequence abrogates successful infection of *E. coli* by T7 phage. We speculate that *udk* could have been part of a past mechanism in *E. coli* that existed to thwart infection by T7-like phages, but is no longer available as a mechanism for suppression of T7 infection. We propose that the T7 phage has evolved a strategy involving the serine/threonine kinase Gp0.7 to counteract the effect of Udk. Two lines of evidence support this view. Firstly, Udk is phosphorylated at a threonine residue in stationary phase *E. coli* cells (Macek *et al.*, 2008), suggesting that Udk activity could be somehow antagonized by phosphorylation. Secondly, mutant T7 Δ 0.7 phage is severely impaired for growth in *E. coli* cells overexpressing *udk*, whereas wild-type T7 is relatively less impaired compared with T7 Δ 0.7 phage. In other words, we propose that Gp0.7-mediated phosphorylation of Udk could prevent the Udk-mediated antagonism of Gp2 function during T7 phage infection of *E. coli*. Global transcriptome and phosphoproteome analysis in *E. coli* under *udk* overexpressing conditions and during wild-type T7 and T7 Δ 0.7 infections might provide further insights into overexpressed Udk-mediated antagonism of Gp2 function in *E. coli* during T7 phage infection. Such experiments are currently under way in our laboratory.

ACKNOWLEDGEMENTS

This work was supported by grants from the BBSRC and Wellcome Trust to S. W.

REFERENCES

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.

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.

James, E., Liu, M., Sheppard, C., Mekler, V., Cámara, 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.

Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H. & Mori, H. (2005). Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res* **12**, 291–299.

Macek, B., Gnad, F., Soufi, B., Kumar, C., Olsen, J. V., Mijakovic, I. & Mann, M. (2008). Phosphoproteome analysis of *E. coli* reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Mol Cell Proteomics* **7**, 299–307.

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.

Qimron, U., Marintcheva, B., Tabor, S. & Richardson, C. C. (2006). Genomewide screens for *Escherichia coli* genes affecting growth of T7 bacteriophage. *Proc Natl Acad Sci U S A* **103**, 19039–19044.

Qimron, U., Kulczyk, A. W., Hamdan, S. M., Tabor, S. & Richardson, C. C. (2008). Inadequate inhibition of host RNA polymerase restricts T7 bacteriophage growth on hosts overexpressing *udk*. *Mol Microbiol* **67**, 448–457.

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.

Severinova, E. & Severinov, K. (2006). Localization of the *Escherichia coli* RNA polymerase β' subunit residue phosphorylated by bacteriophage T7 kinase Gp0.7. *J Bacteriol* **188**, 3470–3476.

Shadrin, A., Sheppard, C., Severinov, K., Matthews, S. & Wigneshweraraj, S. (2012). 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. *Microbiology* **158**, 2753–2764.

Edited by: D. Grainger