PRODUCTION OF RECOMBINANT *PSEUDOMONAS AERUGINOSA* RNA POLYMERASE HOLOENZYME COMPLEX AND ITS UTILIZATION FOR *IN VITRO* TRANSCRIPTIONAL REGULATION STUDIES

by

DERRICK AFFUL

(Under the Direction of Cory Momany, PhD)

ABSTRACT

Pseudomonas aeruginosa is an important public health pathogen that is well-known for its high ability to adapt to a diverse range of environmental conditions, including those encountered in human hosts. The remarkably high ecological success, versatility, and pathogenicity of *P. aeruginosa* has been attributed to its harboring of a large repertoire of regulatory proteins coupled with an intricate genetic network. Yet, many aspects of the *P. aeruginosa* regulatory network remain largely unexplored for antibiotic drug discovery. Therefore, this dissertation describes studies that contribute to addressing some of the current unmet needs in *P. aeruginosa* regulatory network research. First, the methods utilized to produce high yields of highly pure and active *P. aeruginosa* RNA polymerase (RNAP) will be described, and previously unknown activity and kinetic information for the enzyme will be reported. The successful production of the central enzyme mediating transcription in *P. aeruginosa* sets the foundation to structural characterization of transcriptional regulation in the bacterium, which could provide novel insights into species-specific therapeutic drug development. For further probing transcriptional regulation in *P. aeruginosa*, this dissertation also focuses on two

LysR-type Transcriptional Regulators (LTTRs) in the pathogen: FinR and CysB. Both proteins were found to be central to the *P. aeruginosa* virulence network through their regulation of sulfur metabolism in the pathogen. We identified a previously unreported role for FinR and found CysB to be a likely global regulator that mediated several virulent pathways. A central component of the experiments that led to these findings was our utilization of the *P. aeruginosa* RNAP to probe transcription in the bacterium. Overall, the findings from this dissertation provide the foundation for new studies targeting the *P. aeruginosa* regulatory network for antibiotic drug discovery and also provide a gateway for the elucidation of LTTR-mediated transcriptional regulation in *P. aeruginosa*.

INDEX WORDS: *Pseudomonas aeruginosa*, transcriptional regulation, RNA polymerase, *P. aeruginosa* RNA polymerase, RNAP, bacteria transcription, LysR-type transcriptional regulator, FinR, CysB, alternate sigma factor

PRODUCTION OF RECOMBINANT PSEUDOMONAS AERUGINOSA RNA POLYMERASE HOLOENZYME COMPLEX AND ITS UTILIZATION FOR *IN VITRO* TRANSCRIPTIONAL REGULATION STUDIES

by

DERRICK AFFUL

B.S., Kwame Nkrumah University of Science and Technology, Ghana, 2009

M.S., Kwame Nkrumah University of Science and Technology, Ghana, 2013

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

PRODUCTION OF RECOMBINANT *PSEUDOMONAS AERUGINOSA* RNA POLYMERASE HOLOENZYME COMPLEX AND ITS UTILIZATION FOR *IN VITRO* TRANSCRIPTIONAL REGULATION STUDIES

by

DERRICK AFFUL

Major Professor: Cory Momany

Committee: Scott Pegan
George Zheng
Robert Phillips

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2019

DEDICATION

This dissertation is dedicated to my late mother, Esther Nana Serwah Afful and the rest of my loving family.

ACKNOWLEDGMENTS

This dissertation and the research contained in it would not have been possible without the favor and grace of the Lord, and I wish to first and foremost thank Him for the strength through this past 5 years. The document holds far more than the culmination of five years of study, as it also reflects within its pages, the relationships with the many generous and inspiring people who in one way or another contributed and extended their valuable assistance in the preparation and completion of the studies described here.

I am truly indebted to my major professor, Dr. Cory Momany for all his support, guidance, encouragement, and advice through these years. His belief in me is what kept me going, and I will be forever grateful for the wonderful opportunity to have him as a major professor. I would also like to acknowledge the support and guidance received from the members of my graduate committee. To Dr. Ellen Neidle, Dr. Scott Pegan, Dr. George Zheng, and Dr. Robert Phillips, I say, thank you!

It is also my pleasure to acknowledge a long list of special people whose prayers, love, generosity and support have helped me survive these past five years. I am particularly thankful to my wife, Lawrenda Appiah, for the support and love that saw me through the years. Additionally, I am thankful to special friends like Nadia Nkansah, Emmanuel Kenneth Cudjoe, Kevin Osei Affram, Aboagyewaa Oppong-Damoah, Felix Acheampong, Eric Mintah, Victor Lincha, Eugene Fletcher, Nicole Laniohan, Jaeah Kim, Sukneeraj Kaur, Emmanuel Timmy Donkoh, Courage Kweku Dzandu, Daniel Sarfo Marfo, David Baidoe-Ansah, and a plethora of others whose support in numerous ways

would be forever acknowledged by me. To my lab mates, Nicole Laniohan, Ajay Arya, and Juliana Brondani, as well as all the undergraduate students and high school students who have worked on numerous projects with me to make this dissertation possible, I say, thank you for being part of my story these five years. I am also grateful to the rest of my classmates, staff, and professors in the Pharmaceutical and Biomedical Sciences program for their support in various ways throughout these years.

Finally, I acknowledge the unflagging love of my family. To my wife, dad, late mother who this thesis is dedicated to, siblings, uncles, aunties, mothers-in-laws, siblings-in-laws, I say, thank you for the love and unrelenting support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION AND LITERATURE REVIEW

1.1 BACTERIAL REGULATORY CIRCUITS

Historical Perspectives

Regulatory circuits provide the remarkable and essential ability to modulate cellular processes in all organisms in response to physicochemical and nutritional changes in the environment [1, 2]. Transcriptional regulation was initially demonstrated in bacteria, which have an extraordinary repertoire of proteins to control gene expression [3, 4]. Nobel laureates François Jacob and Jacques Monod first focused on lactose metabolism to demonstrate that bacteria respond to environmental stimuli by expressing or repressing specific genes [5]. Subsequently, it was postulated that genes can be organized into functional units, called operons, to enable the co-expression and co-regulation of physiologically related gene products [6]. Studies of the *lac* operon set the foundation for our current knowledge of transcriptional control in bacteria and other living organisms.

In 1965, following the work of Jacob and Monod, Ptashne showed that a genetic switch dictated whether *E. coli* cells infected by λ-bacteriophage would undergo lysogeny (genomic incorporation) or lysis [7, 8]. The recognition that genetic networks are integral to cellular function increased, and further studies with theoretical models, such as Boolean dynamic systems [9], lent

support to the idea that functional states in the cell can be represented by a giant genetic network. Notably, bacterial RNA polymerase (RNAP) was found to be essential for transcription and was established as the central element coordinating gene expression within this network [10]. Other factors, including sigma factors and anti-sigma factors, were discovered to help coordinate bacterial response to environmental changes [11] [10, 12]. In 1975, Pribnow discovered a conserved regulatory element to which RNAP binds to facilitate transcription initiation [13]. Different elements were elucidated that jointly control transcription and contribute to the notion of modularity and varied modes/mechanisms of regulation. Structural studies of some of these elements were critical to an improved understanding of transcription. Initial structures were that of a transcriptional factor-effector complex that was bound to its DNA template whereas a latter structure depicted the ternary complex between catabolite activator protein (CAP), the alpha C-terminal domain (α-CTD) of RNAP, and DNA [14]. These studies provided a physical representation of the induction of a regulator by its respective inducer leading to affinity and specificity changes for the target gene.

Next, the need for motif searching resulted in the development of several computational sequence-searching programs. Prominent among them was MEME, developed by Bailey and Elkan in 1994 [15], which is currently one of the most widely used bioinformatic tools. Around the same time, the development of DNA microarrays by [16] opened the door for genome-wide probing of transcription, which also made high-throughput experimentation of transcriptional networks possible. Microarrays and ChIP-chip data enabled the confirmation of the longstanding theory that genes are highly interrelated [17-21]. Subsequently, the first databases of transcriptional regulation were developed for *E. coli*, prominent among them being RegulonDB [22, 23]. Other important advances in the bacterial regulatory network timeline include the structural characterization of the *E. coli* and *T. thermophilus* RNA polymerases [24-29], as well as the discovery of regulatory RNAs

[30-34] and alternate sigma factors [35-37], all of which further affirmed the complexities of gene regulatory networks.

1.2 BACTERIAL TRANSCRIPTION FACTORS

Transcription factors (TFs) represent one of the largest protein classes in most organisms [38]. TFs bind to specific recognition sequences in the vicinity of their target genes to activate or repress the expression of the genes [39]. Typically, these proteins comprise two domains: the sensing domain that receives internal and external signals, and the DNA-interacting domain that recognizes and binds to DNA [40]. The coordination of roles between these two domains allows TFs to act in their capacity as regulatory switches. Additionally, most transcription factors harbor these domains in one single protein, except for two component systems where a sensor protein receives a triggering signal and then transduces the signal to its receiving partner that enforces the required transcriptional regulatory activity [41, 42]. In bacteria, the helix-turn-helix (HTH) motif is the most common element in the DNA-binding domain (DBD) as up to 84% of DBDs in one-component TFs have been found to contain this signature [41, 43]. Specifically, a subtype of HTH motifs, the winged-HTH motif (wHTH) has been proposed to be the most abundant DBD in prokaryotes, accounting for close to 45% of total TFs [44]. The wHTH motif contains two small beta-sheets acting as the “wings” or loops (W1, W2) in addition to three alpha helices (H1, H2, H3) and three beta-sheets (S1, S2, S3) arranged in the order H1-S1-H2-H3-S2-W1-S3-W2 [45]. One of the three alpha-helices act as a recognition element and forms specific contacts with DNA by fitting into the major groove [46]. Prokaryotic transcription factors are grouped into various families based on the similarities between the different DBDs. In their study, [47] identified about 75 families of bacterial and archaeal TFs

and these have also been classified under various superfamilies. Out of the 75 TF families, 17 were specific to bacteria, with 10 out of the 17 belonging to proteobacteria. The largest family among the prokaryotic TFs is the LysR transcriptional regulators, followed by the AraC and TetR families [48].

1.3 THE LYSR-TYPE TRANSCRIPTIONAL REGULATOR FAMILY OF BACTERIAL TRANSCRIPTION FACTORS

The LysR-type Transcriptional Regulator (LTTR) family of transcriptional factors are the largest class of prokaryotic DNA-binding proteins [49]. Henikoff and colleagues [50] were the first to document this class of transcriptional regulators following their identification of nine homologous proteins within three bacterial species that exhibited a high degree of DNA-binding domain conservation. One of the proteins, LysR, the regulator of *lysA* in *E. coli*, was the best characterized of the group at the time [51] and its name was adopted to represent the whole class. Today, more than 40,000 potential LTTRs have been identified across bacteria, archaea, and even some eukaryotes [52, 53]. Perhaps the most intriguing characteristic about LTTRs is the level of diversity exhibited in the processes they regulate despite the high sequence similarities among them.

Structurally, LTTRs contain between 280 to 350 amino acids and comprise two domains: an N-terminal DNA-binding domain (DBD) containing a winged-helix-turn-helix (wHTH) motif, a bridging linker helix (LH), and a C-terminal effector-binding domain (EBD), often referred to as the regulatory domain [54-56] (**Figure 1-1A and 1-1B**). The DBD, involved in major groove DNA interactions and promoter recognition specificity, is highly conserved at a sequence level among all LTTRs. This domain also appears to be critically involved in transcriptional activation through its direct interactions with RNA polymerase [57]. The DBD is linked to the regulatory domain by a

long linker helix with variable length, bending properties, and sequence (**Figure 1-1B**). The more sequence diverse regulatory or effector-binding domain typically comprises two distinct subdomains that fold to reveal a cleft where effector molecules typically bind. However, not all LTTRs are regulated by effectors (*e.g.* NAC). Effector recognition and binding triggers conformational or oligomerization changes in the protein that affect transcription [58-61] Characteristically, the genes for LTTRs are organized divergently from the gene or operon they regulate. As a result, many of them are involved in autoregulation where they repress their own expression [56].

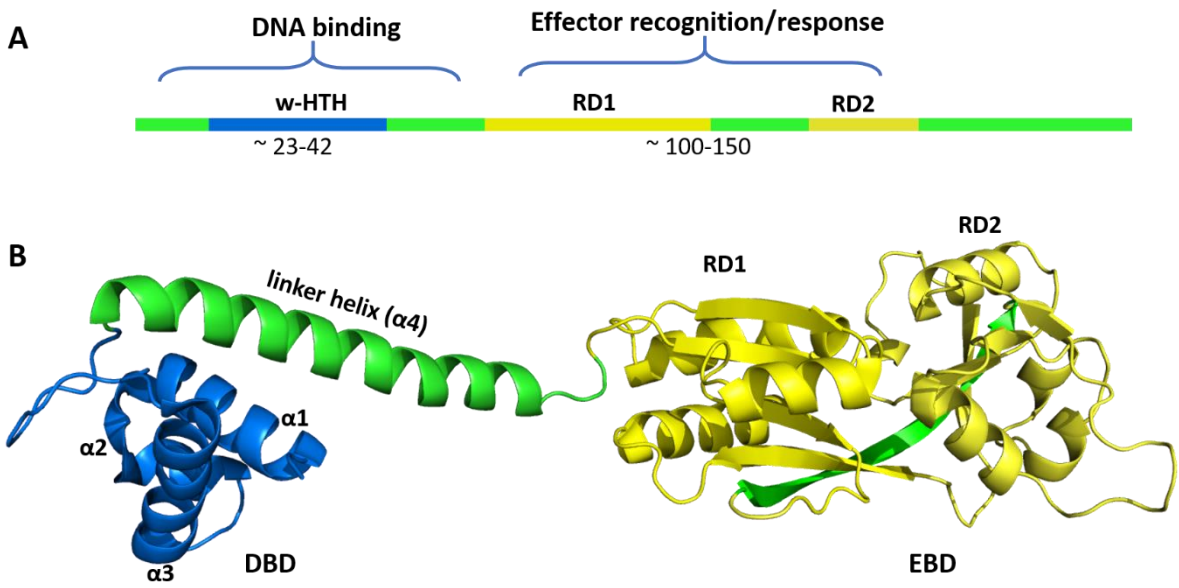


Figure 1-1. Structure of LysR-type transcriptional regulators. **(A)** LTTRs comprise a highly conserved DNA binding domain containing a winged helix-turn-helix domain at the N-terminus, and a less conserved effector binding domain at the C-terminus. The EBD or regulatory domain (RD) contains two subdomains, RD1 and RD2. **(B)** X-ray structure of a monomer unit of OxyR, a *P. aeruginosa* LTTR involved in oxidative stress regulation [62].

LysR-Type Transcriptional Regulators in *P. aeruginosa*

The genomes of Pseudomonads have been found to contain a large repertoire of regulatory proteins. These metabolically versatile organisms therefore occupy and survive in a diverse range of environments and niches [63]. The pathogenic pseudomonad, *Pseudomonas aeruginosa*, the most studied pathogenic bacterial model due to its metabolic and pathogenic abilities [63], is an important opportunistic public health pathogen that causes both acute and chronic human infections and is the leading cause of nosocomial infections worldwide [64]. Behind *E. coli* and *B. subtilis*, *P. aeruginosa* contains the third largest known regulatory network in any bacteria [65]. As a result, at least 489 genes have been identified in the *P. aeruginosa* genome that encode putative transcriptional regulators [66]. These transcriptional regulators, coupled with their respective sigma factors as well as various regulatory RNA and RNA-binding proteins, work in concert to produce an extensive network of virulence factors. Unsurprisingly, as many as 125 of the *P. aeruginosa* transcriptional regulators (over 25%) have been identified as putative LTTRs, reflecting the reliance of this organism on these LTTRs particularly for disease pathogenesis and survival.

LysR-type transcriptional regulators have attracted substantial research attention over the years [49, 51, 52, 56, 58, 59, 61, 62, 67-84], and they make attractive drug targets for two major reasons. First, the regulation of essential cellular processes by some of these LTTRs suggests that targeting them could be fatal to the cells. Second, and more importantly, many of these LTTRs regulate non-essential processes like virulence and antibiotic drug resistance. This suggests that that these processes could be targeted without compromising bacterial growth, a strategy that has been considered promising for overcoming drug resistance (Mandal et al., 2016). The first *P. aeruginosa* LTTR to be characterized was TrpI, which was found to be involved in tyrosine synthesis [67]. Since

then, a total of twenty-four other *P. aeruginosa* LTTRs have been characterized, most of them within the last decade (**Table 1-1**).

Functional classification of *P. aeruginosa* LTTRs

A search in the Pseudomonas Database (www.pseudomonas.com) for putative *P. aeruginosa* LTTRs identified 125 of them in the genome of the bacterium, of which 24 had been characterized. A classification of the characterized *P. aeruginosa* LTTRs into four main functional categories is shown in the Venn Diagram of **Figure 1-2**. Many of the identified LTTRs have roles that overlap two or more of the functional categories. Additionally, a few LTTRs in the pathogen have also been regarded as global regulators that control a host of genes/processes. **Table 1-1** contains a summary of all known information regarding the currently characterized LTTRs in *P. aeruginosa*.

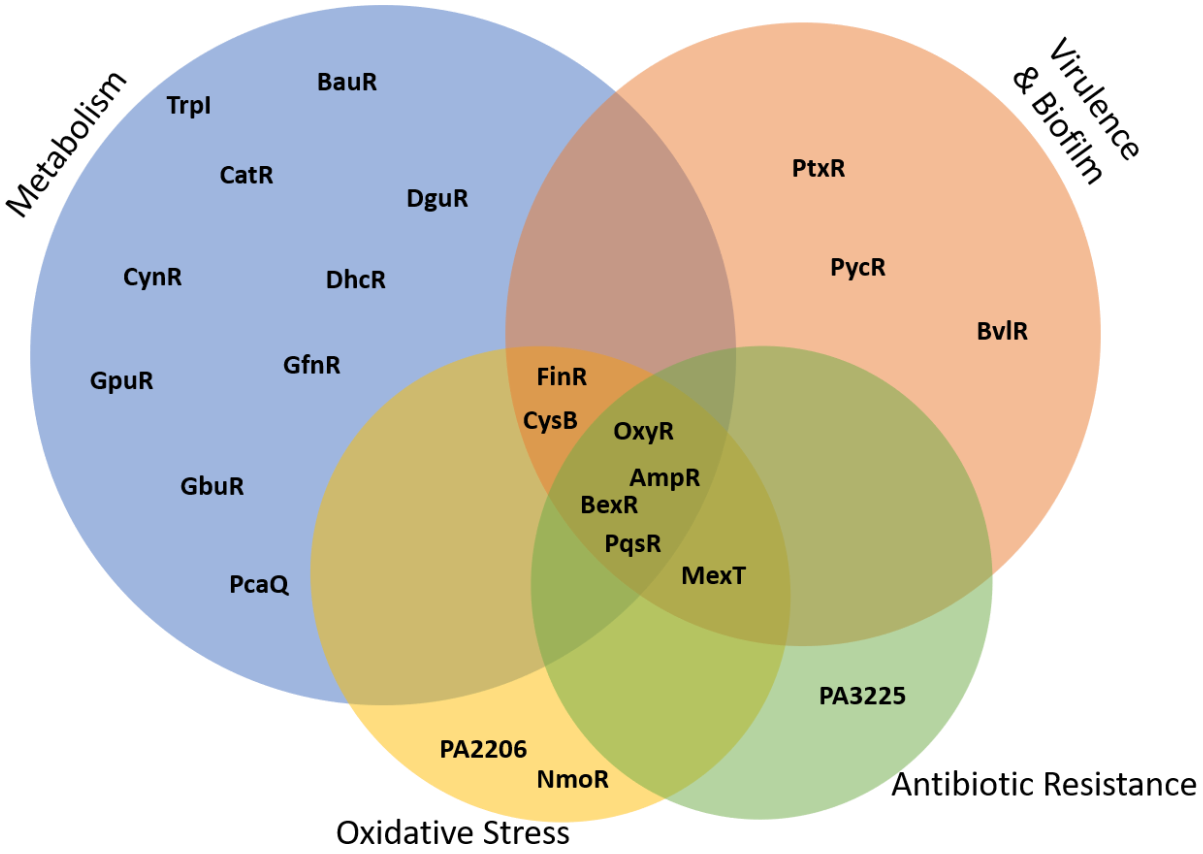


Figure 1-2. Classification of the characterized *P. aeruginosa* LTTRs under four functional categories. Most LTTRs in *P. aeruginosa* are involved in various metabolic pathways (n = 17) as well as the regulation of virulence (n = 10), whereas a few others are involved in oxidative stress regulation (n = 8) and antibiotic resistance (n = 6). The four global LTTRs, OxyR, AmpR, PqsR, and BexR, have been associated with roles overlapping all four categories, while others like FinR, CysB, and MexT have roles overlapping at least 2 categories.

Table 1-1. *P. aeruginosa* LysR-Type Transcriptional Regulators that are currently characterized

LTTR	Locus tag	Gene/Operon	Known/Predicted function(s)	Known/Predicted ligand(s)	Expression	Regulation	Known structure(s) (PDB code)
AmpR	PA4109	*ampRCGDE*, Global	Global, β-lactamase expression	Peptidoglycan degradation products*	Divergent	Autoregulation	EBD** (5MMH)
BauR	PA0133	*bauRABCD*	Polyamine utilization (C and N sources)	β-alanine*	Divergent	Autoregulation	No
BexR	PA2432	Global	Phenotypic variation through bistable expression	None	Divergent	Autoregulation	No
BvlR	PA2877	Global	Repression of virulence factor production	None	Divergent	Autoregulation	No
CatR	PA2510	*catBC*	Benzoate utilization (characterized in *P. putida*)	None	Divergent	Unknown	No
CynR	PA2054	*cynTS*	Cynate catabolism (characterizaed in *E. coli*)	None	Divergent	Unknown	No
CysB	PA1754	*msuEDC*, global	Aliginate synthesis, Cysteine biosynthesis	Cysteine*, sulfur esters*, organosulfates*, alginate*	Divergent	Autoregulation	No
DguR	PA5085	*dguRABC*	D-Glutamate/D-Glutamine utilization	D-Glu*, D-Gln*	Divergent	Autoregulation	No
DhcR	PA1998	*dhcRAB*	Carnitine catabolism	3-hydrocarnitine	Divergent	Autoregulation	No
FinR	PA3398	*fprA*	Oxidative stress	None	Divergent	Autoregulation	No
GbuR	PA1422	*gbuRA*	Arginine utilization via 4-guanidinobutyrate	Arginine	Divergent	Autoregulation	No
GfnR	PA3630	*souA*	Sarcosine metabolism, Formaldehyde detoxification	Sarcosine*	Divergent	Autoregulation	No
GpuR	PA0289	*gpuPAR*	Arginine utilization via 3-guanidinobutyrate	Arginine	Divergent	Autoregulation	No

iciA	PA4363	*oriC*	Chromosomal regulation (characterized in *E. coli*)	None	Not divergent	Unknown	No
MexT	PA2492	*mexEF-oprN*	Drug efflux, Global	None	Divergent	Autoregulation	No
MvfR	PA1003	*pqsABCDE*	Quorum sensing	Alkyl quinolones Homoserine lactones	Divergent	*lasR* (+regulation) *rhiR* (-regulation)	EBD** (4JVC, 4JVD, 4JVI)
NmoR	PA4203	*nmorRA*	Proprionate-3-nitronate detoxification	None	Divergent	Autoregulation	No
OxyR	PA5344	katABC, aphABCF, global	Oxidative stress, global	None	Divergent	Autoregulation	FL*** (4X6G, 4XWS, 4Y0M)
PA2206	PA2206	Global	Oxidative stress, Virulence	None	Not divergent	Unknown	No
PA3225	PA3225	*tssABC1*	Antibiotic resistance repressor	None	Divergent	Unknown	No
PcaQ	PA0152	*pcaDHG*	Catabolism of phenolic compounds (characterized in *A. tumefaciens*)	None	Not divergent	Unknown	No
PtxR	PA2258	*toxA*	Endotoxin A expression/virulence	None	Divergent	Autoregulation	No
PycR	PA5437	*pycAB*	Pyruvate carboxylase expression Virulence/maintenance of infections	None	Divergent	Autoregulation	No
TrpI	PA0037	*trpAB*	Tryptophan biosynthesis	Indoleglycerol phosphate	Divergent	Autoregulation	No

*Based on prediction from orthologs (no experimental data available). **EBD: Effector binding domain; ***FL: Full-length

P. aeruginosa LTTRs involved in metabolic pathways

Almost half of the currently characterized *P. aeruginosa* LTTRs play various roles in metabolic pathways. The high environmental adaptability of the bacterium is possible due to its ability to metabolize various compounds as nutrients and incorporate others into diverse metabolic pathways. Most of the LTTRs in this category are involved in utilization or biosynthesis of specific amino acids. TrpI (PA0037), the tryptophan biosynthesis regulator, was the first LTTR to be characterized from *P. aeruginosa*. The LTTR was found be involved in tryptophan biosynthesis through its regulation of *trpBA*, the tryptophan synthase operon [67, 68]. TrpI was one of the early LTTRs from which the classical DNA bending LTTR regulatory model was deduced. In the model, one of two binding positions on the target operon is favored depending on the presence or absence of a co-inducer. The presence of an inducer cause DNA bending that allows transcription of the downstream operon [52]. In the case of TrpI, the co-inducer was identified to be indoleglycerol [69-72].

The other major LTTR in the category is CysB (PA1754), the cysteine biosynthesis regulator. The LTTR controls the cysteine regulon that oversees L-cysteine biosynthesis in many bacteria [73]. The regulon comprises the genes for L-cystine, glutathione, sulfate, and thiosulfate uptake; those for sulfate activation and reduction to sulfide; genes for both O-acetylserine (thiol)-lyase isozymes, and the genes involved in alkanesulfonate utilization. All these genes have been theorized to be under positive CysB regulation with N-acetyl-L-serine as its inducer [49, 50, 74]. The *P. aeruginosa* CysB was first shown to be involved to be involved in expression of AlgD, the enzyme that mediates alginate synthesis during colonization of the cystic fibrosis lung [75]. The LTTR was also implicated in the regulation of sulfate-starvation-induced (SSI) proteins, which was consistent with the observations in other species like *E. coli* whose *cysB* regulon has been widely studied [76].

Additionally, CysB was found to be involved in sulfate ester desulfurization and transport [77] through its regulation of the sulfur-regulated arylsulfate gene cluster (*ats* genes). Levels of CysB increase in response to low intracellular L-cysteine levels [78], but the mechanism of the regulation is unknown. So CysB acts as a signal molecule that can in turn tune acquisition of sulfur and dampen non-essential pathways that deplete L-cysteine. The LTTR has also been shown to mediate the production of various *P. aeruginosa* virulence factors. One of these is the iron starvation sigma factor gene, *pvdS,* which controls the expression of several host virulence factors [79, 80]. This new CysB role linked the iron and sulfur metabolic regulons in *P. aeruginosa* to a single LTTR. Another virulent process CysB has been shown to be involved in is quorum sensing in *P. aeruginosa*. The LTTR was found to repress the expression of PqsR (MvfR), the multiple virulence factor regulator [81, 82]. In this way, CysB represses the production of the quinolone signal required for quorum sensing. In a very recent study, CysB was found to mediate virulence through its regulation of the *P. aeruginosa* type III secretion system [83]. These studies have established CysB as a master regulator in *P. aeruginosa* due to its broad regulon. From its regulon, CysB also seems to play an important role in disease progression and lifestyle choice of *P. aeruginosa,* making it an ideal target for drug discovery.

Three other LTTRs in this category that are involved in amino acid biosynthetic/utilization processes include DguR (PA5085), GbuR (PA1422), GpuR (PA0289). DguR regulates the *dguRABC* operon, which is involved in D-glutamine and D-glutamate utilization in *P. aeruginosa* [84]. GbuR the guanidinobutyrase regulator, and GpuR, the guanidinoprioprionase regulator, are both involved in the catabolism of arginine or other guanidino-containing compounds for their use as carbon and nitrogen sources in *P. aeruginosa* [85, 86]. The final set of LTTRs in this category regulate specific catabolic processes in *P. aeruginosa*. BauR (PA0133) was identified as part of the

bauRABCD operon and was shown to be broadly involved in regulation of polyamine utilization in *P. aeruginosa* [87]. The LTTR is divergently expressed upstream of *bauA* and it was found to be responsive to β-alanine. Another *P. aeruginosa* LTTR, DhcR (PA1998**)**, the dehydrocarnitine regulator, was found to be involved in the catabolism of carnitine, which is used as a carbon and nitrogen source in *P. aeruginosa* [88]. The LTTR is divergently expressed from the *dhcAB* operon and was shown to be regulated by 3-hydrocarnitine. The glutathione-dependent formaldehyde neutralization regulator, GfnR (PA3630), is another *P. aeruginosa* LTTR in this category. The LTTR was first reported by [89] and it was shown to be involved in sarcosine catabolism.

P. aeruginosa LTTRs involved in virulence and biofilm formation

The second largest functional class of LTTRs in *P. aeruginosa* comprises ones that mediate the production of various virulence factors, a notorious hallmark of this pathogen. The host of virulence factors the bacterium relies on to cause infections include lipopolysaccharides, phospholipases, exoproteases, siderophores, phenazines, outer membrane vesicles, exotoxins, type III secreted effectors, flagella, hemolysins, and pili [90, 91]. LTTRs play a critical role in the *P. aeruginosa* virulence factor network and it is no surprise that these proteins are now being regarded as better drug discovery targets than proteins whose targeting have fatal consequences for the bacteria [92, 93].

The most widely studied *P. aeruginosa* LTTR is the regulator of quorum sensing signal, PqsR (PA1003). Also called the multiple virulence factor regulator (MvfR), this transcription factor plays a central role in virulence and biofilm formation in the bacteria through its regulation of quorum sensing, the phenomenon where bacteria cells coordinate gene expression in a population

density-dependent manner [94]. Originally characterized by [95], the LTTR was shown to positively regulate a host of virulence factors including elastase, phospholipase, homoserine lactone and quinolone inducers, as well as the *phnAB* operon which is involved in phenazine biosynthesis. It is now known that the primary role of PqsR is to initiate the transcription of the *pqsABCDE* operon following its activation by the pseudomonas quinolone signal (PQS), which comprises various quinolone compounds [96]. Somehow uncharacteristically for LTTRs, PqsR has been shown to not undergo autoregulation but rather, positive regulation by another gene, *lasR* as well as negative regulation by the *rhiR* gene [97]. Over the years, the LTTR has been established as a global transcriptional regulator in *P. aeruginosa* [98] that works in concert with other notable LTTRs in the bacteria like AmpR [99], CysB [81, 82], and OxyR [100]. The LTTR also plays a central role in biofilm formation by *P. aeruginosa*, a popular characteristic of the bacteria that is responsible for its persistent infections and resistance to antibiotics [101-103]. PqsR is perhaps the only *P. aeruginosa* LTTR that has been thoroughly researched in terms of potential ligands. Many studies have focused on assessing the potential of various quinolones as activators of the transcription factor while a host of others have explored various compounds as potential antagonists for inhibiting the protein [104-107]. Structurally, PqsR is one of only three *P. aeruginosa* LTTRs whose structures have been determined and are available in the Protein Data Bank (**Table 1-1**), albeit only as an effector binding domain without ligand [108] and with 2-nonylquinolin-4(1H)-one [109].

Another LTTR that mediates virulence in *P. aeruginosa* is PtxR (PA2258), the regulator of exotoxin A, which is the most toxic of all the virulence factors in the bacteria [110]. Exotoxin is an ADP-ribosyl transferase enzyme that plays a major role in the disruption of protein synthesis, ultimately resulting in cell death [111]. The LTTR was first reported by [112] who showed it to regulate the exotoxin A gene, *toxA*. In follow-up studies, the same group discovered an adjacent

gene, *ptxS*, that interfered with the function of PtxR on exotoxin A production [113, 114]. Another major role PtxR has been implicated with is the regulation of the *pvcABCD* operon, which is responsible for the synthesis of the chromophore moiety of the *P. aeruginosa* siderophore pyoverdine [115]. In a more recent study, PtxR was found to regulate some of the genes involved in carbon metabolism in *P. aeruginosa* [116]. PtxR has also been linked with the production of quorum sensing-controlled virulence factors like rhamnolipid and pyocyanin [117], suggesting a potential cross-talk with PqsR. Although a study by [118] found that natriuretic peptides, a family of eukaryotic hormones that mediate cytotoxicity, can modulate exotoxin A production, later studies suggest that the ligand for PtxS is 2-ketogluconate [116, 119].

Among the currently characterized *P. aeruginosa* LTTRs, the only one to be implicated in epigenetic control in the bacteria is BexR (PA2432), the bistability expression regulator. Phenotypic variation has been thought to help bacteria survive in harsh environmental conditions by allowing a small subset of the population to adapt and thrive [120]. Bacteria utilize bistable systems, among others, to undergo phenotypic variation. These are systems that exist simultaneously in two states and can reversibly switched anytime. In such systems, half the genes are in an 'on' state and half are in an 'off' state. BexR was identified as a bistable switch that is itself bistably expressed and is positively autoregulated [121]. The mRNA expression profile of a *bexR*-knockout compared to wild-type cells revealed a diverse set of 71 genes that were potentially under BexR control. Notable among these genes were the *mexEF* operon (reviewed below), quorum sensing genes, and the *prXEFA* genes, which encode components of the alkaline protease production and secretion system that have been implicated in *P. aeruginosa* virulence. It therefore seems probable that BexR is a global regulator that cross-talks with many of the other transcriptional regulators in the bacteria (**Figure 1-2**). No potential effector ligands have been reported for the protein.

Another *P. aeruginosa* LTTR that participates in virulence is BvlR (PA2877), the regulator of pathogenicity. The protein was found to be involved in the regulation of several virulence factors as well as tight microcolony formation during the biofilm formation process [122]. Interestingly, BvlR was found to downregulate the type III secretion system (T3SS), which is an important component of the *P. aeruginosa* virulence factor arsenal. The finding that BvlR was involved in the repression of virulence factor production but was required for biofilm formation in *P. aeruginosa* was an unusual one for an LTTR; however, the results were consistent with the observations in bacteria that many virulence factor production systems are shut down once cells transition into the biofilm state [123]. The final LTTR in this category is PycR (PA5437), the pyruvate carboxylase regulator. The protein found to be heavily involved in maintenance of *P. aeruginosa* infections in the cystic fibrosis lung such that its inactivation led to a 100,000-fold decrease in virulence in a rat disease model [124]. PycR is divergently expressed from its two genes, *pycA* and *pycB*, both of which were predicted to encode two subunits of pyruvate carboxylase.

P. aeruginosa LTTRs involved in oxidative stress regulation

Reactive oxygen species (ROS) in the environment of bacteria cells can cause damage to various biomolecules. As a result, bacterial genomes encode various genes that can be used to mount an antioxidant defense mechanism [125]. Bacteria rely on several antioxidant enzymes that neutralize specific oxidative stressors, and these include catalases, superoxide dismutases, peroxiredoxins, alkyl hydroperoxide reductases, and thiol peroxidases [126]. The metabolic versatility and high environmental adaptability of *P. aeruginosa* ensures that the bacterium survives nutrient-limited conditions and grows in stressful environments. During infections, the bacterium is exposed to toxic

levels of various ROS generated by macrophages and other immune cells [127]. As expected, *P. aeruginosa* is armed with an impressive antioxidant defense system that relies mainly on LTTRs and it is no surprise that the master regulator of oxidative stress in the bacteria, OxyR (PA5344), is an LTTR.

The *P. aeruginosa* antioxidant defense system comprises two superoxide dismutases, three catalases (encoded by *katA*, *katB*, *katC*), as well as four alkyl hydroperoxide reductases (AhpA, AhpB, AphCF) [128]. The expression of all these genes is under OxyR regulation, making the protein one of the most widely studied *P. aeruginosa* LTTRs. Studies and characterization of the OxyR regulon began primarily in *E. coli* [129]. Through the works of [128] and several others [130-133], OxyR was established as a central regulator of the *P. aeruginosa* oxidative stress response. Over the years, several studies have further contributed to expanding the OxyR regulon and its regulation, which has established the transcriptional regulator as a global regulator with roles overlapping the other LTTR functional classifications. Disease model studies found OxyR to be required for full virulence in the bacteria [134], while other studies have also reported the involvement of the protein in virulence [135, 136] and virulence factor production [137]. OxyR has also been associated with roles encompassing metabolic processes like the production of pyocyanin and rhamnolipids [100] or iron homeostasis [138]. In their study [125] found a total of 56 genes under OxyR regulation that were previously unknown. These included genes involved in iron homeostasis, quorum sensing, protein synthesis, and even oxidative phosphorylation. OxyR has also been implicated in *P. aeruginosa* antibiotic resistance through its involvement in aminoglycoside resistance in the pathogen [139, 140]. One of the earliest details observed about OxyR was how its oxidation led to activation. The formation of a disulfide bond between two cysteine residues, C199

and C208 was found to be crucial to activation [131]. The protein is also the only *P. aeruginosa* LTTR whose full-length protein structure has been determined [62, 141].

Another major LTTR in this category is FinR (PA3398), the ferredoxin reductase regulator. Although the *P. aeruginosa* FinR is among the most recently characterized LTTRs in the bacterium [142], the protein has been well-researched in other proteobacteria, particularly in other pseudomonads like *P. putida*. The FinR proteins regulate the expression of the ferredoxin-$NADP^+$ reductases encoded by the *fpr* genes. Ferredoxin reductases (Fprs) are ubiquitous, monomeric, and reversible flavin-dependent enzymes that catalyze the reversible redox reaction between NADPH or NADP+ and one-electron carriers such as ferredoxin or flavodoxin [143]. Fpr enzymes therefore play a crucial role in maintaining the NADP+/NADPH ratio and redox state associated with important cellular processes like siderophore synthesis/regulation, iron acquisition, sulfur assimilation and cysteine biosynthesis, and oxidative stress [126, 144-146]. These roles for FinR have been confirmed in other pseudomonads like *P. putida* [143, 147-149]. In their study, [142], who reported the characterization of the *P. aeruginosa* FinR, found the LTTR to positively regulate *fprA* expression while undergoing autoregulation. The study also showed that *finR* mutants were associated with an increased sensitivity to paraquat, which was reversed upon overexpression of *fprA*.

Two other LTTRs that have been implicated in oxidative stress regulation in *P. aeruginosa* include the unnamed but characterized PA2206and NmoR (PA4203). PA2206 was found to be required during the oxidative stress response as well as for full virulence in a zebrafish disease model [150]. NmoR on the other hand was involved in indirect oxidative stress regulation through its role in oxidative detoxification [151].

P. aeruginosa LTTRs involved in antibiotic resistance

Over the years, *P. aeruginosa* has earned its designation as a 'superbug', joining the ranks of bacteria that have proven resistant to essentially all antibiotics on the market [152]. Consequently, the bacterium displays a high resistance to a diverse range of antibiotics, including aminoglycosides, quinolones and β-lactams [153]. Antibiotic resistance mechanisms in *P. aeruginosa* have been classified into intrinsic, acquired, and adaptive antibiotic resistance mechanisms [154]. Intrinsic antibiotic resistance mechanisms comprise those that are engraved in the bacterium's genetic makeup and are mediated mainly by two transcriptional regulators, MexT and AmpR, both LTTRs. Expectedly, both LTTRs are among the most widely studied *P. aeruginosa* transcriptional regulators. AmpR (PA4109), the ampicillin resistance regulator, controls the *amp* genes that partly mediate the resistance *P. aeruginosa* to the β-lactam class of antibiotics. The role of AmpR and its regulon in conferring antibiotic resistance on *P. aeruginosa* is well-documented [155-161]. In their study, [162] found over 500 different genes that were dysregulated following transcriptome analysis of *ampR*-knockout mutants. In a subsequent study, LTQ-XL mass spectrometry was used for transcriptome analysis, which showed that as much as 53% of total *P. aeruginosa* proteins were under AmpR regulation [163]. Processes AmpR was found to regulate include biofilm formation, quorum sensing, virulence factor production, and expression of the MexEF efflux pump. In another study, the LTTR was implicated in additional processes including oxidative stress, heat shock, and iron uptake [99]. Recently, the X-ray crystal structure of the *P. aeruginosa* AmpR effector binding domain was determined by Dik and colleagues [164], making AmpR the second LTTR to be characterized structurally as an effector binding domain. While it had been proposed that various peptidoglycan degradation products could act as ligands [165], the structural studies with the EBD revealed that the

likely activator ligand is the muropeptide, UDP-N-acetyl-beta-d-muramyl-l-Ala-gamma-d-Glu-meso-DAP-d-Ala-d-Ala.

The multidrug efflux transporter regulator, MexT (PA2492), is another widely studied *P. aeruginosa* LTTR and its involvement in antibiotic resistance is well-documented [166-174]. The LTTR was first characterized by [175] as part of a complex system of three operons that were each multidrug efflux pumps. The three operons have been shown to have a similar organization: the first of each contains a periplasmic fusion protein (MexA, MexC, or MexE), the second gene expresses a cytoplasmic membrane protein (MexB, MexD, or MexF), while the final gene in each operon encodes an outer membrane protein (MexB, MexD, or MexF). The MexB, MexD, and MexF proteins are thought to be the actual efflux pumps [175] and together, they mediate the transport of antibiotics such the β-lactams, quinolones, tetracyclines, chloramphenicol, and macrolides [176]. Like the other major LTTRs in *P. aeruginosa*, MexT has been established as a global regulator in the bacteria, with roles encompassing virulence [177-181], quorum sensing [182, 183], and even oxidative/metabolic stress [176, 184, 185]. Effectors of MexT have yet to be reported in the literature.

The last LTTR in this category is the unnamed **PA3225**, a transcriptional regulator that was found to repress antibiotic resistance mechanisms in *P. aeruginosa* [186]. This makes PA3225 a particularly interesting and somewhat unusual protein in comparison with the others in this category. The LTTR's ability to suppress antibiotic resistance in *P. aeruginosa* was found to be attributed to its repression of the MexAB-OprM efflux pump and other efflux transporters. This makes PA3225 an LTTR worthy of further research as its antibiotic resistance suppression role can be explored for drug discovery.

1.4 BACTERIAL RNA POLYMERASES

The transcription of DNA to RNA is the first step in the central dogma of biology. This makes RNA polymerase (RNAP) the central protein in gene expression, one of the most important proteins in all kingdoms of life [187]. RNA polymerases are complex DNA-dependent molecular machines that synthesize RNA chains complementary to DNA template strands by utilizing ribonucleoside triphosphate (NTP) substrates. As a machine with the active center at its core, RNAP is comprised of a number of distinct immobile and moving parts that receive various regulatory signals [188]. Furthermore, as a molecular vehicle, RNAP generates force that allows it to incorporate ribonucleotides unidirectionally into RNA at rates between 15 and 80 nt/s [189, 190]. RNA polymerases are some of the most complex enzymes in living organisms, and this is attributed to their multi-subunit and large nature. Bacterial RNAPs contain five units that are largely conserved across different species in terms of primary sequence, ternary structure, and function, and each RNAP has a molecular weight around 500 kDa [191]. Bacteria contain a single RNAP dedicated to making all RNA species, while eukaryotes have multiple RNAPs dedicated to specific types of transcripts. RNA polymerases were discovered in the 1960s by Hurwitz and colleagues whose studies were focused on the incorporation of ribonucleotides into RNA [11, 192, 193].

Architecture of Bacterial RNA Polymerases

Early studies of the bacterial RNAP revealed it to be a multisubunit enzyme that existed in two complexes: the core enzyme ($\alpha_2\beta\beta'$) that was catalytically competent but incapable of initiation, and the holoenzyme ($\alpha_2\beta\beta'\sigma$) that was specialized for initiation [194]. Subsequent research has identified

a fifth subunit, ω, whose role is not completely understood but is thought to be involved in regulation and assembly [195]. The organization of the different RNAP subunits to form the complete holoenzyme is shown in **Figure 1-3**.

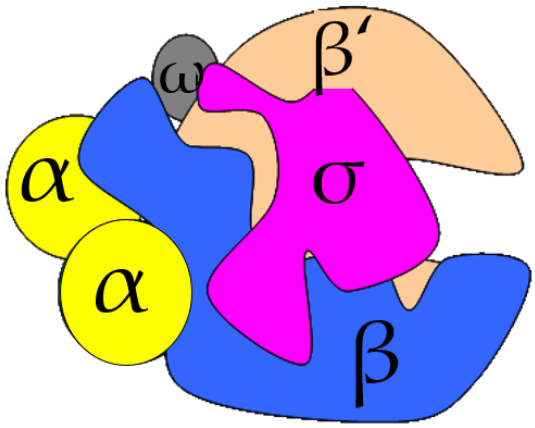


Figure 1-3. The prokaryotic RNAP holoenzyme organization showing the five subunits.

As shown in **Figure 1-3**, the overall structure of the ~400 kDa bacterial RNAP core enzyme resembles a 'crab claw' [24] and this shape is conserved in both the archaeal [196] and eukaryotic RNAPs [197, 198]. However, unlike the eukaryotic RNAPs, there is only one form of core RNAP enzyme [199]. Analysis and discovery of the various RNAP subunits began in the late 1970s, championed by Burgess and his group [200-203]. The β subunit was the first subunit to be identified and this was through the finding that rifamycins and streptovaricins, potent antibiotics, inhibited transcription initiation by binding to the RNAP enzyme instead of DNA [204]. Further studies confirmed that the two genes encoding the β and β' subunits, *rpoB* and *rpoC*, respectively, were indeed related and that the two were on the same operon under a common promoter. The β and β'

subunits are the two largest subunits of RNAP (~150 – 160 kD each), and they together make up the catalytic core of the enzyme [194]. Both subunits provide sites for binding of the DNA template and RNA transcript during transcription, while they're also involved in chain elongation and termination.

The next RNAP subunit to be identified was the ~40 kDa α subunit (encoded by *rpoA*). Unlike the well-known notion that functionally related genes in bacteria are closely located in proximity to each other on the chromosome, *rpoA* was found to be nowhere near the *rpoBC* operon [205]. Functionally, the α subunit has been shown to be involved in three distinct processes: RNAP assembly, DNA binding at some promoters, and transcriptional activation [206-210]. The finding that the α subunit was key to RNAP assembly was one of the early roles associated with the protein and the role was found to be exclusive to the N-terminus [206]. During RNAP assembly, two moles of α are required for every molecule of RNAP formed. The α-C-terminal domain (α-CTD) has been established as an essential part of the alpha subunit that mediates activation at some promoters. This role of the subunit is also linked to its role regarding DNA binding, which also lies within the α-CTD. The protein binds to sites called UP elements located upstream of the – 10 and – 35 promoter regions [210]. Interestingly, the α-CTD is not conserved between prokaryotes and eukaryotes.

The discovery of the σ subunit (σ^{70}, encoded by *rpoD*) of RNAP was made serendipitously through studies that aimed to understand RNAP subunit structure [211]. It is now known that prokaryotes contain multiple σ factors, with σ^{70} being the most abundant that transcribes most of the genome and handles most housekeeping functions [211]. The sigma factors have been found to share four conserved regions that contribute to their various functions [212, 213]. The protein is unique in prokaryotes, although it has been suggested that its roles are similar to that of the TATA binding protein in eukaryotes [194]. Although sigma factors bind DNA, many of them cannot do this in the absence of core RNA polymerase [214]. In addition to DNA binding leading to transcription

initiation, sigma factors have also been associated with transcription activation through activator binding [215]. Indeed, the architecture of σ^{70} reveals several activator binding sites in the C-terminal domain [194].

1.5 BACTERIAL TRANSCRIPTION

The transcription of DNA to RNA is one of the most fundamental process of life in all organisms. The multistep nature of the process renders it susceptible to high regulation at multiple points, a necessity considering how fundamental gene expression is to any organism. Therefore, transcription is controlled by an array of factors and substances like promoter DNA sequence, transcriptional activators, coactivators, repressors, accessory ligands, nucleotide concentration, temperature, salt and solute concentrations, and other environmental variables [216].

The first and key step in transcription is the recognition of a promoter sequence by RNAP. Promoters are cis-acting elements located upstream of the transcription start site (TSS) [217]. Promoter recognition by RNAP critically relies on the sigma factors. Therefore, the specific type of sigma factor that associates with the core enzyme to form the holoenzyme defines transcriptional specificity. Regardless of the sigma factor, most bacterial promoters can be classified into two functional sites known as the −35 and −10 regions upstream of the TSS, both regions separated by 15-31 bp [38]. Each sigma factor recognizes specific consensus sequences in these two regions and similarity to the sequence has been found to correlate positively with promoter strength. Conversely, mismatches in the promoter consensus sequence can affect the level of gene expression, although the gene products themselves will remain unaltered [218]. In addition to the 10 and -35 elements,

other components of the promoter that must be recognized by the RNAP holoenzyme prior to transcription initiation include, the UP element, as well as the $^{-15}$TGnT^{-12} motif [219].

The first step in the transcription process is initiation. This is a complex process that occurs in several steps: promoter recognition by RNAP, formation of initiation complex, synthesis of the initial phosphodiester bonds, and movement of RNAP from the promoter to begin elongation [220]. Initial recognition of the promoter elements (P) by RNAP (R) leads to the formation of the closed initiation complex (RPc) to reflect the fully double-stranded nature of the template DNA at this point [221]. Formation of the RPc is the beginning of a series of structural transitions that culminates in the formation of the processive transcription elongation complex (TEC). The promoter elements that make the formation of the RPc possible include the UP elements, the –35 and –10 elements, and the TGnT motif region. The formation of RPc is a readily reversible process and this renders the process susceptible to competitors like heparin that possess the high ability to reverse the process [219]. Particularly, addition of heparin at a specified time after transcription initiation ensures that only one round of transcription is possible, which can be a valuable approach for studying transcription kinetics. RNAP typically contacts close to 80 bp of template DNA during initiation which extends from about 20 bp downstream of the transcription start site up to ~60 bp upstream of it [183].

Except for the interactions between the –35 element as well as the –10 element with sigma, all other DNA-RNAP interactions in the closed initiation complex are mostly non-specific and weak, making the RPc largely unstable. Consequently, aromatic residues within sigma function to induce destabilizations and local distortions within the DNA duplex at the conserved A or T residue occupying the –11 position. The resulting thermal fluctuation and free energy that accumulates in within the complex leads to DNA melting, strand separation, and the insertion of the template strand into the active site [187, 222]. The formation of the transcription bubble results from these processes

and the various kinetic changes that precede its formation is referred to as isomerization. Isomerization leads to the formation of an open promoter-RNAP initiation complex (RPo) where the RNAP jaw is fully closed to stabilize the template DNA in the active site. RPo is the stable species at 37°C and it is normally not susceptible to competition by polyanionic agents like heparin [223]. These extensive insights into the steps underlying the formation of the RPo and the initiation complex has been possible due to the available of several high-resolution structures of the *T. thermophilus*, *T. aquaticus*, and *E. coli* DNA-RNAP complexes [25, 26, 224, 225].

Conversion of RPo to RPinit, the initiation complex requires RNAP to bind to NTP molecules that are complementary to the nucleotides at the +1 and +2 positions on the template strand. Purines are typically preferred for the +1 position, the reason transcription start codons begin with an A or G nucleotide, and the site they occupy (the i site) has a relatively high Km for these NTPs [226] (Helmann, 2009). Following the occupation of the i +1 site by the NTP complementary to the +2 nucleotide, the formation of the first phosphodiester bond between the two substrate NTPs begins. The phosphodiester bond catalytic process leads to the release of a pyrophosphate (PPi) by-product and the subsequent translocation of the terminal NMP from the i +1 site to the i site to make way for the catalysis of the +3 NTP. The cyclic process of NTP binding, catalysis, and translocation continues and is repeated after every NTP incorporation [227]. Initial stages of elongation typically do not result in the formation of productive transcripts. This is because an initial transcript made must displace the sigma σ3.2 domain loop encountered in its path as shown in **Figure 1-4A**; however, many short transcripts (~ 2-10 nucleotides long) are incapable of doing this and are therefore released as short abortive transcripts. Eventually, the RNA transcript elongates to about 12 nucleotides, which is a sufficient length to fill completely the RNA-DNA hybrid as well as upstream RNA exit channel under the β flap, leading to a displacement of the σ3.2 loop and bringing abortive

initiation to an end [28, 226] (**Figure 1-4B**). The growth of the transcription bubble and the RNA-DNA hybrid requires both the template and non-template DNA strands to be looped out of the main channel, resulting in a 'DNA scrunching' process during the early stages of elongation. Consequently, this, coupled with the displacement of σ3.2 domain loop leads to destabilization of the contacts between the sigma σ4 domain and the β flap as shown in **Figure 1-4C**. The loss of contact with the β flap in turn leads to unstable interactions between σ4 and the –35 element, which subsequently causes RNAP to release the promoter, a process referred to as promoter clearance [28, 187]. Following promoter clearance, the RNAP core enzyme translocates further downstream to complete the formation of the transcription elongation complex (TEC) and continue RNA elongation (**Figure 1-4D**). These insights into the bacterial transcription elongation process have been available by courtesy of a number of high-resolution structures of the *T. thermophilus* TEC comprising synthetic DNA/RNA scaffolds and NTPs [228-231].

The final step in transcription is termination, an essential step for accurate gene expression as well as the removal of RNAP from transcript products. Programmed transcription termination occurs by one of three potential pathways: 1) intrinsic termination requiring only the RNAP, RNA, and DNA, 2) Rho-dependent termination that requires the presence of the Rho protein, and 3) Mdf-dependent termination where damaged elongation complexes are targeted [232].

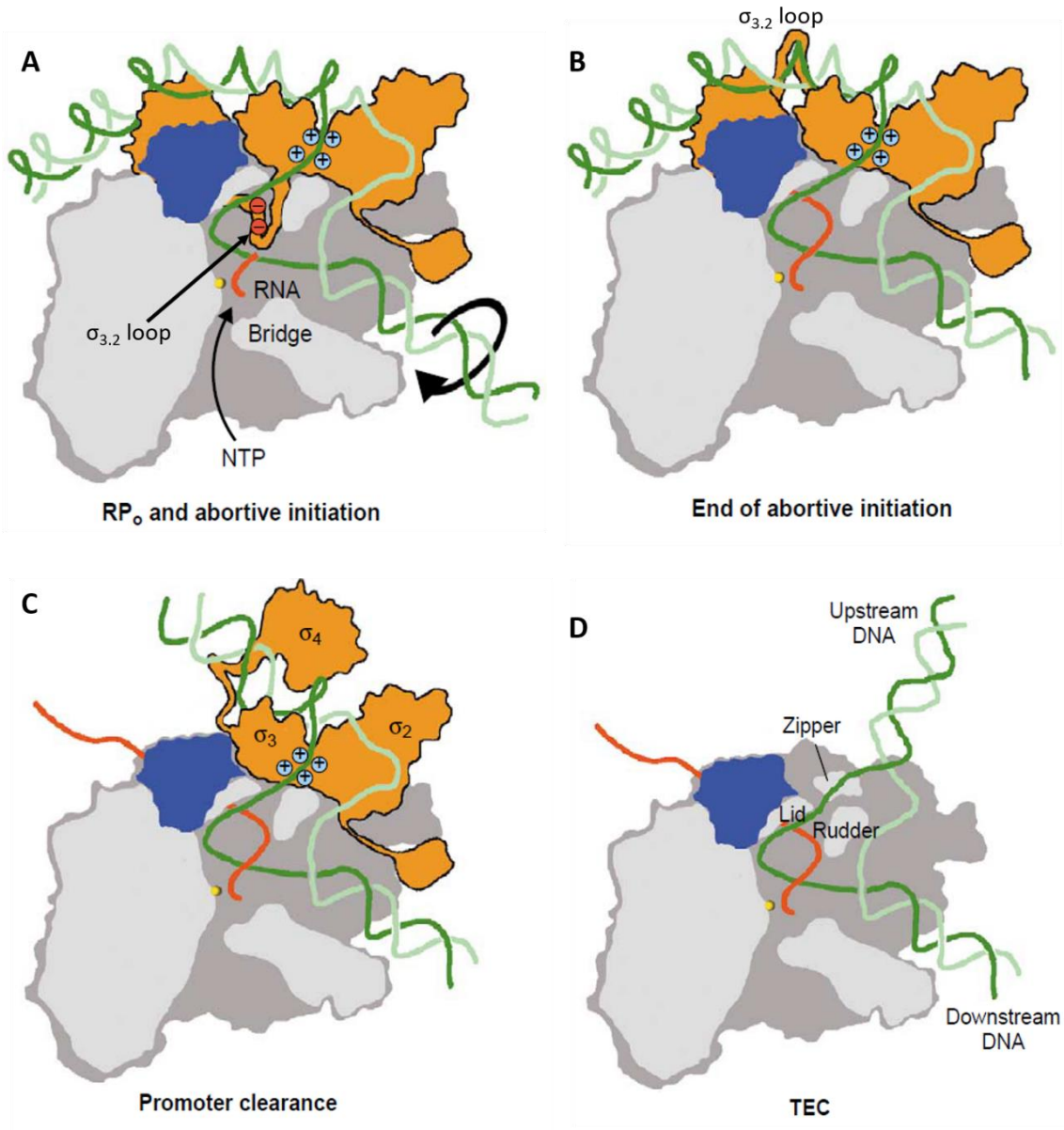


Figure 1-4. Stages involved in the later parts of transcription initiation and the early stages of elongation. A cross-sectional view of the RNAP holoenzyme is shown, containing the β flap (blue), σ (orange), catalytic Mg^{2+} (yellow) and rest of RNAP (gray). The template strand is shown in dark green, the NT strand in light green, and the RNA transcript is shown in red (Figure reused with permission from [28]).

1.6 INTRODUCTION OF STUDY

Research described in this thesis focuses on *P. aeruginosa*. The Gram-negative pseudomonad is an aerobic bacillus that is present ubiquitously in diverse environments [233]. *P. aeruginosa* is particularly a stubborn bacterium because its infections are difficult to treat, owing to its exhaustive intrinsic and acquired antibiotic resistance mechanisms. Unsurprisingly, the bacterium is among the ESKAPE pathogens, a list of the most pathogenic nosocomial bacteria for which only few antibiotics are effective against [234]. *P. aeruginosa* has therefore been the focus of several studies focusing on antibiotic drug discovery targeted at the pathogen. Nevertheless, one relatively unexplored target for *P. aeruginosa* drug discovery is its regulatory network, the heart of the bacterium's high metabolic and pathogenic abilities. In this regard, the various research described in the thesis is aimed at making important contributions to existing knowledge regarding the *P. aeruginosa* transcriptional regulatory network. Because RNA polymerase is essential for any transcriptional regulation studies targeted at the pathogen, the first study described (**Chapter 2**), will focus on the biochemical characterization of recombinatnly produced *P. aeruginsa* RNA polymerase and its utilization for transcription regulation studies. This will be the first description of the successful production of RNAP from this bacterium and it opens a multitude of opportunities for studying transcription in the *P. aeruginosa*. Next, **Chapters 3** and **4** will focus on transcriptional regulation of two *P. aeruginosa* LTTRs, FinR and CysB, and probe the regulatory roles of the two proteins in the pathogen. Overall, the research described provides an important biochemical reagent for probing transcriptional regulation in *P. aeruginosa* (RNAP), while presenting insights into the regulatory role of two prominent LTTRs in the pathogen that are viable targets for drug discovery due to the processes we have found them to regulate.

1.7 REFERENCES

[1] R. Silva-Rocha, V. de Lorenzo, Noise and robustness in prokaryotic regulatory networks. Annu Rev Microbiol 64 (2010) 257-275.

[2] I. Cases, V. de Lorenzo, Promoters in the environment: transcriptional regulation in its natural context. Nat Rev Microbiol 3 (2005) 105-118.

[3] E. Perez-Rueda, M.A. Martinez-Nuñez, The repertoire of DNA-binding transcription factors in prokaryotes: functional and evolutionary lessons. Science progress 95 (2012) 315-329.

[4] F. Jacob, J. Monod, Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 3 (1961) 318-356.

[5] F. Jacob, J. Monod, [Genes of structure and genes of regulation in the biosynthesis of proteins]. C R Hebd Seances Acad Sci 249 (1959) 1282-1284.

[6] J.A. Freyre-Gonzalez, L. Trevino-Quintanilla, Analyzing regulatory networks in bacteria. Nat Educ 3 (2010) 24.

[7] M. Ptashne, Replication and Host Modification of DNA Transferred during Bacterial Mating. J Mol Biol 11 (1965) 829-838.

[8] M. Ptashne, The Detachment and Maturation of Conserved Lambda Prophage DNA. J Mol Biol 11 (1965) 90-96.

[9] S.A. Kauffman, Metabolic stability and epigenesis in randomly constructed genetic nets. Journal of theoretical biology 22 (1969) 437-467.

[10] R.R. Burgess, A.A. Travers, J.J. Dunn, E. Bautz, Factor stimulating transcription by RNA polymerase. Nature 221 (1969) 43.

[11] J. Hurwitz, The enzymatic incorporation of ribonucleotides into polydeoxynucleotide material. Journal of Biological Chemistry 234 (1959) 2351-2358.

[12] A. Stevens, J. Rhoton, Characterization of an inhibitor causing potassium chloride sensitivity of an RNA polymerase from T4 phage-infected Escherichia coli. Biochemistry 14 (1975) 5074-5079.

[13] D. Pribnow, Bacteriophage T7 early promoters: nucleotide sequences of two RNA polymerase binding sites. Journal of molecular biology 99 (1975) 419-443.

[14] B. Benoff, H. Yang, C.L. Lawson, G. Parkinson, J. Liu, E. Blatter, Y.W. Ebright, H.M. Berman, R.H. Ebright, Structural basis of transcription activation: the CAP-αCTD-DNA complex. Science 297 (2002) 1562-1566.

[15] T.L. Bailey, C. Elkan, Fitting a mixture model by expectation maximization to discover motifs in bipolymers. (1994).

[16] M. Schema, D. Shalon, R. Davis, P. Brown, Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270 (1995) 467-470.

[17] A.L. Barabasi, R. Albert, Emergence of scaling in random networks. Science 286 (1999) 509-512.

[18] N. Guelzim, S. Bottani, P. Bourgine, F. Kepes, Topological and causal structure of the yeast transcriptional regulatory network. Nat Genet 31 (2002) 60-63.

[19] P.L. DeRisi J, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM, Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nature Genetics 14 (1996) 457-460.

[20] D.J. Lockhart, H. Dong, M.C. Byrne, M.T. Follettie, M.V. Gallo, M.S. Chee, M. Mittmann, C. Wang, M. Kobayashi, H. Horton, E.L. Brown, Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat Biotechnol 14 (1996) 1675-1680.

[21] B. Ren, F. Robert, J.J. Wyrick, O. Aparicio, E.G. Jennings, I. Simon, J. Zeitlinger, J. Schreiber, N. Hannett, E. Kanin, T.L. Volkert, C.J. Wilson, S.P. Bell, R.A. Young, Genome-wide location and function of DNA binding proteins. Science 290 (2000) 2306-2309.

[22] A.M. Huerta, H. Salgado, D. Thieffry, J. Collado-Vides, RegulonDB: a database on transcriptional regulation in Escherichia coli. Nucleic Acids Res 26 (1998) 55-59.

[23] H. Salgado, A. Santos, U. Garza-Ramos, J. van Helden, E. Diaz, J. Collado-Vides, RegulonDB (version 2.0): a database on transcriptional regulation in Escherichia coli. Nucleic Acids Res 27 (1999) 59-60.

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

[25] Y. Zhang, Y. Feng, S. Chatterjee, S. Tuske, M.X. Ho, E. Arnold, R.H. Ebright, Structural basis of transcription initiation. Science 338 (2012) 1076-1080.

[26] R.S. Basu, B.A. Warner, V. Molodtsov, D. Pupov, D. Esyunina, C. Fernandez-Tornero, A. Kulbachinskiy, K.S. Murakami, Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme. J Biol Chem 289 (2014) 24549-24559.

[27] A.H. Finney, R.J. Blick, K. Murakami, A. Ishihama, A.M. Stevens, Role of the C-terminal domain of the alpha subunit of RNA polymerase in LuxR-dependent transcriptional activation of the lux operon during quorum sensing. J Bacteriol 184 (2002) 4520-4528.

[28] K.S. Murakami, S.A. Darst, Bacterial RNA polymerases: the wholo story. Curr Opin Struct Biol 13 (2003) 31-39.

[29] K.S. Murakami, S. Masuda, S.A. Darst, Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. Science 296 (2002) 1280-1284.

[30] S. Gottesman, The small RNA regulators of Escherichia coli: roles and mechanisms*. Annu Rev Microbiol 58 (2004) 303-328.

[31] K.M. Wassarman, F. Repoila, C. Rosenow, G. Storz, S. Gottesman, Identification of novel small RNAs using comparative genomics and microarrays. 15 (2001) 1637-1651.

[32] L. Argaman, R. Hershberg, J. Vogel, G. Bejerano, E.G. Wagner, H. Margalit, S. Altuvia, Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. Curr Biol 11 (2001) 941-950.

[33] S. Chen, E.A. Lesnik, T.A. Hall, R. Sampath, R.H. Griffey, D.J. Ecker, L.B. Blyn, A bioinformatics based approach to discover small RNA genes in the Escherichia coli genome. Biosystems 65 (2002) 157-177.

[34] E. Rivas, R.J. Klein, T.A. Jones, S.R. Eddy, Computational identification of noncoding RNAs in E. coli by comparative genomics. Curr Biol 11 (2001) 1369-1373.

[35] D.W. Cowing, J. Mecsas, M.T. Record, Jr., C.A. Gross, Intermediates in the formation of the open complex by RNA polymerase holoenzyme containing the sigma factor sigma 32 at the groE promoter. J Mol Biol 210 (1989) 521-530.

[36] D.W. Cowing, J. Bardwell, E.A. Craig, C. Woolford, R.W. Hendrix, C.A. Gross, Consensus sequence for Escherichia coli heat shock gene promoters. Proceedings of the National Academy of Sciences 82 (1985) 2679-2683.

[37] D.W. Cowing, C.A. Gross, Interaction of Escherichia coli RNA polymerase holoenzyme containing sigma 32 with heat shock promoters. DNase I footprinting and methylation protection. J Mol Biol 210 (1989) 513-520.

[38] S.C. Janga, J. Collado-Vides, Structure and evolution of gene regulatory networks in microbial genomes. Res Microbiol 158 (2007) 787-794.

[39] G. Lloyd, P. Landini, S. Busby, Activation and repression of transcription initiation in bacteria. Essays Biochem 37 (2001) 17-31.

[40] A. Ishihama, Prokaryotic genome regulation: multifactor promoters, multitarget regulators and hierarchic networks. FEMS microbiology reviews 34 (2010) 628-645.

[41] L.E. Ulrich, E.V. Koonin, I.B. Zhulin, One-component systems dominate signal transduction in prokaryotes. Trends in microbiology 13 (2005) 52-56.

[42] T. Mascher, J.D. Helmann, G. Unden, Stimulus perception in bacterial signal-transducing histidine kinases. Microbiology molecular biology reviews 70 (2006) 910-938.

[43] A.S. Seshasayee, P. Bertone, G.M. Fraser, N.M. Luscombe, Transcriptional regulatory networks in bacteria: from input signals to output responses. Current opinion in microbiology 9 (2006) 511-519.

[44] V. Charoensawan, D. Wilson, S. Teichmann, Genomic repertoires of DNA-binding transcription factors across the tree of life. Nucleic acids research 38 (2010) 7364-7377.

[45] K.S. Gajiwala, S.K. Burley, Winged helix proteins. Current opinion in structural biology 10 (2000) 110-116.

[46] D.A. Rodionov, Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. Chemical reviews 107 (2007) 3467-3497.

[47] E. Pérez-Rueda, J. Collado-Vides, L. Segovia, Phylogenetic distribution of DNA-binding transcription factors in bacteria and archaea. Computational biology chemistry 28 (2004) 341-350.

[48] S.K. Kummerfeld, S.A. Teichmann, DBD: a transcription factor prediction database. Nucleic acids research 34 (2006) D74-D81.

[49] M.A. Schell, Molecular biology of the LysR family of transcriptional regulators. Annual Reviews in Microbiology 47 (1993) 597-626.

[50] S. Henikoff, G.W. Haughn, J.M. Calvo, J.C. Wallace, A large family of bacterial activator proteins. Proceedings of the National Academy of Sciences 85 (1988) 6602-6606.

[51] P. Stragier, F. Richaud, F. Borne, J.C. Patte, Regulation of diaminopimelate decarboxylase synthesis in Escherichia coli. I. Identification of a lysR gene encoding an activator of the lysA gene. J Mol Biol 168 (1983) 307-320.

[52] S.E. Maddocks, P.C. Oyston, Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology 154 (2008) 3609-3623.

[53] F.J. Reen, M. Barret, E. Fargier, M. O'Muinneachain, F. O'Gara, Molecular evolution of LysR-type transcriptional regulation in Pseudomonas aeruginosa. Mol Phylogenet Evol 66 (2013) 1041-1049.

[54] S. Muraoka, R. Okumura, Y. Uragami, T. Nonaka, N. Ogawa, K. Miyashita, T. Senda, Purification and crystallization of a LysR-type transcriptional regulator CBNR from Ralstonia eutropha NH9. Protein Pept Lett 10 (2003) 325-329.

[55] S. Muraoka, R. Okumura, N. Ogawa, T. Nonaka, K. Miyashita, T. Senda, Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. J Mol Biol 328 (2003) 555-566.

[56] J. Zaim, A.M. Kierzek, The structure of full-length LysR-type transcriptional regulators. Modeling of the full-length OxyR transcription factor dimer. Nucleic acids research 31 (2003) 1444-1454.

[57] A. Lochowska, R. Iwanicka-Nowicka, J. Zaim, M. Witkowska-Zimny, K. Bolewska, M.M. Hryniewicz, Identification of activating region (AR) of Escherichia coli LysR-type transcription factor CysB and CysB contact site on RNA polymerase alpha subunit at the cysP promoter. Mol Microbiol 53 (2004) 791-806.

[58] O.C. Ezezika, S. Haddad, E.L. Neidle, C. Momany, Oligomerization of BenM, a LysR-type transcriptional regulator: structural basis for the aggregation of proteins in this family. Acta Crystallogr Sect F Struct Biol Cryst Commun 63 (2007) 361-368.

[59] O.C. Ezezika, S. Haddad, T.J. Clark, E.L. Neidle, C. Momany, Distinct effector-binding sites enable synergistic transcriptional activation by BenM, a LysR-type regulator. J Mol Biol 367 (2007) 616-629.

[60] S. Muraoka, R. Okumura, N. Ogawa, T. Nonaka, K. Miyashita, T. Senda, Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. Journal of molecular biology 328 (2003) 555-566.

[61] A. Ruangprasert, S.H. Craven, E.L. Neidle, C. Momany, Full-length structures of BenM and two variants reveal different oligomerization schemes for LysR-type transcriptional regulators. J Mol Biol 404 (2010) 568-586.

[62] I. Jo, I.Y. Chung, H.W. Bae, J.S. Kim, S. Song, Y.H. Cho, N.C. Ha, Structural details of the OxyR peroxide-sensing mechanism. Proc Natl Acad Sci U S A 112 (2015) 6443-6448.

[63] D. Balasubramanian, S.K. Murugapiran, E. Silva-Herzog, L. SCHNEGER, X. Yang, G. Tatke, G. Narasimhan, K. Mathee, Transcriptional regulatory network in Pseudomonas aeruginosa, Caiser Academic Press, United Kingdom, 2013.

[64] A.L. Goodman, S. Lory, Analysis of regulatory networks in Pseudomonas aeruginosa by genomewide transcriptional profiling. Current Opinion in Microbiology 7 (2004) 39-44.

[65] E. Galán-Vásquez, B. Luna, A. Martínez-Antonio, The regulatory network of Pseudomonas aeruginosa. Microbial informatics and experimentation 1 (2011) 3.

[66] G.L. Winsor, E.J. Griffiths, R. Lo, B.K. Dhillon, J.A. Shay, F.S. Brinkman, Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. Nucleic acids research 44 (2016) D646-D653.

[67] M. Schell, P.H. Brown, S. Raju, Use of saturation mutagenesis to localize probable functional domains in the NahR protein, a LysR-type transcription activator. Journal of biological chemistry 265 (1990) 3844-3850.

[68] Z.X. Tian, E. Fargier, M. Mac Aogain, C. Adams, Y.P. Wang, F. O'Gara, Transcriptome profiling defines a novel regulon modulated by the LysR-type transcriptional regulator MexT in Pseudomonas aeruginosa. Nucleic Acids Res 37 (2009) 7546-7559.

[69] F. Imperi, F. Tiburzi, G.M. Fimia, P. Visca, Transcriptional control of the pvdS iron starvation sigma factor gene by the master regulator of sulfur metabolism CysB in Pseudomonas aeruginosa. Environ Microbiol 12 (2010) 1630-1642.

[70] R. Tyrrell, K.H. Verschueren, E.J. Dodson, G.N. Murshudov, C. Addy, A.J. Wilkinson, The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. Structure 5 (1997) 1017-1032.

[71] M.P. Koentjoro, N. Ogawa, STRUCTURAL STUDIES OF TRANSCRIPTIONAL REGULATION BY LysR-TYPE TRANSCRIPTIONAL REGULATORS IN BACTERIA. Reviews in Agricultural Science 6 (2018) 105-118.

[72] G.S. Knapp, J.C. Hu, Specificity of the E. coli LysR-Type Transcriptional Regulators. PLOS ONE 5 (2010) e15189.

[73] T.A. Lewis, A. Glassing, J. Harper, M.J. Franklin, Role for ferredoxin:NAD(P)H oxidoreductase (FprA) in sulfate assimilation and siderophore biosynthesis in Pseudomonads. J Bacteriol 195 (2013) 3876-3887.

[74] R.S. Mandal, A. Ta, R. Sinha, N. Theeya, A. Ghosh, M. Tasneem, A. Bhunia, H. Koley, S. Das, Ribavirin suppresses bacterial virulence by targeting LysR-type transcriptional regulators. Sci Rep 6 (2016) 39454.

[75] L.S. Collier, G.L. Gaines, 3rd, E.L. Neidle, Regulation of benzoate degradation in Acinetobacter sp. strain ADP1 by BenM, a LysR-type transcriptional activator. J Bacteriol 180 (1998) 2493-2501.

[76] H. Cao, G. Krishnan, B. Goumnerov, J. Tsongalis, R. Tompkins, L.G. Rahme, A quorum sensing-associated virulence gene of Pseudomonas aeruginosa encodes a LysR-like transcription regulator with a unique self-regulatory mechanism. Proceedings of the National Academy of Sciences 98 (2001) 14613-14618.

[77] A. Stintzi, Z. Johnson, M. Stonehouse, U. Ochsner, J.M. Meyer, M.L. Vasil, K. Poole, The pvc gene cluster of Pseudomonas aeruginosa: role in synthesis of the pyoverdine chromophore and regulation by PtxR and PvdS. J Bacteriol 181 (1999) 4118-4124.

[78] K. Vercammen, Q. Wei, D. Charlier, A. Dotsch, S. Haussler, S. Schulz, F. Salvi, G. Gadda, J. Spain, M.L. Rybtke, T. Tolker-Nielsen, J. Dingemans, L. Ye, P. Cornelis, Pseudomonas aeruginosa LysR PA4203 regulator NmoR acts as a repressor of the PA4202 nmoA gene, encoding a nitronate monooxygenase. J Bacteriol 197 (2015) 1026-1039.

[79] Y. Song, C. Yang, G. Chen, Y. Zhang, Z. Seng, Z. Cai, C. Zhang, L. Yang, J. Gan, L. Haihua, Molecular insights into the master regulator CysB-mediated bacterial virulence in Pseudomonas aeruginosa. Molecular microbiology 0 (2019).

[80] S.H. Craven, O.C. Ezezika, C. Momany, E.L. Neidle, LysR homologs in Acinetobacter: insights into a diverse and prevalent family of transcriptional regulators. Molecular Biology Edited by U Gerischer Norfolk, UK: Caister Scientific Press (2008).

[81] S.H. Craven, O.C. Ezezika, S. Haddad, R.A. Hall, C. Momany, E.L. Neidle, Inducer responses of BenM, a LysR-type transcriptional regulator from Acinetobacter baylyi ADP1. Mol Microbiol 72 (2009) 881-894.

[82] S. Boonma, A. Romsang, J. Duang-Nkern, S. Atichartpongkul, W. Trinachartvanit, P. Vattanaviboon, S. Mongkolsuk, The FinR-regulated essential gene fprA, encoding ferredoxin NADP+ reductase: Roles in superoxide-mediated stress protection and virulence of Pseudomonas aeruginosa. PLoS One 12 (2017) e0172071.

[83] A.M. Alanazi, E.L. Neidle, C. Momany, The DNA-binding domain of BenM reveals the structural basis for the recognition of a T-N11-A sequence motif by LysR-type transcriptional regulators. Acta Crystallogr D Biol Crystallogr 69 (2013) 1995-2007.

[84] O.C. Ezezika, L.S. Collier-Hyams, H.A. Dale, A.C. Burk, E.L. Neidle, CatM regulation of the benABCDE operon: functional divergence of two LysR-type paralogs in Acinetobacter baylyi ADP1. Appl Environ Microbiol 72 (2006) 1749-1758.

[85] M. Chang, A. Hadero, I. Crawford, Sequence of the Pseudomonas aeruginosa trpI activator gene and relatedness of trpI to other procaryotic regulatory genes. Journal of bacteriology 171 (1989) 172-183.

[86] J. Gao, G.N. Gussin, Mutations in TrpI binding site II that differentially affect activation of the trpBA promoter of Pseudomonas aeruginosa. EMBO J 10 (1991) 4137-4144.

[87] M. Chang, I.P. Crawford, The roles of indoleglycerol phosphate and the TrpI protein in the expression of trpBA from Pseudomonas aeruginosa. Nucleic Acids Res 18 (1990) 979-988.

[88] S. Pineiro, I. Olekhnovich, G.N. Gussin, DNA bending by the TrpI protein of Pseudomonas aeruginosa. J Bacteriol 179 (1997) 5407-5413.

[89] M. Chang, I.P. Crawford, In vitro determination of the effect of indoleglycerol phosphate on the interaction of purified TrpI protein with its DNA-binding sites. J Bacteriol 173 (1991) 1590-1597.

[90] I. Olekhnovich, G.N. Gussin, Recognition of binding sites I and II by the TrpI activator protein of pseudomonas aeruginosa: efficient binding to both sites requires InGP even when site II is replaced by site I. Gene 223 (1998) 247-255.

[91] N. Kredich, Biosynthesis of Cysteine. (2008).

[92] J. Ostrowski, N. Kredich, Negative autoregulation of cysB in Salmonella typhimurium: in vitro interactions of CysB protein with the cysB promoter. Journal of bacteriology 173 (1991) 2212-2218.

[93] I. Delic-Attree, B. Toussaint, J. Garin, P.M. Vignais, Cloning, sequence and mutagenesis of the structural gene of Pseudomonas aeruginosa CysB, which can activate algD transcription. Mol Microbiol 24 (1997) 1275-1284.

[94] M.A. Kertesz, K. Schmidt-Larbig, T. Wuest, A novel reduced flavin mononucleotide-dependent methanesulfonate sulfonatase encoded by the sulfur-regulated msu operon of Pseudomonas aeruginosa. J Bacteriol 181 (1999) 1464-1473.

[95] J. Hummerjohann, S. Laudenbach, J. Retey, T. Leisinger, M.A. Kertesz, The sulfur-regulated arylsulfatase gene cluster of Pseudomonas aeruginosa, a new member of the cys regulon. J Bacteriol 182 (2000) 2055-2058.

[96] N.M. Kredich, Biosynthesis of Cysteine. EcoSal Plus 3 (2008).

[97] L. Leoni, N. Orsi, V. de Lorenzo, P. Visca, Functional analysis of PvdS, an iron starvation sigma factor of Pseudomonas aeruginosa. Journal of bacteriology 182 (2000) 1481-1491.

[98] J.M. Farrow, 3rd, L.L. Hudson, G. Wells, J.P. Coleman, E.C. Pesci, CysB Negatively Affects the Transcription of pqsR and Pseudomonas Quinolone Signal Production in Pseudomonas aeruginosa. J Bacteriol 197 (2015) 1988-2002.

[99] J.M. Farrow, 3rd, E.C. Pesci, Distal and proximal promoters co-regulate pqsR expression in Pseudomonas aeruginosa. Mol Microbiol 104 (2017) 78-91.

[100] Y. Song, C. Yang, G. Chen, Y. Zhang, Z. Seng, Z. Cai, C. Zhang, L. Yang, J. Gan, L. Haihua, Molecular insights into the master regulator CysB-mediated bacterial virulence in Pseudomonas aeruginosa. Molecular microbiology Accepted Author Manuscript. doi:10.1111/mmi.14200 (2019).

[101] W. He, G. Li, C.-K. Yang, C.-D. Lu, Functional characterization of the dguRABC locus for d-Glu and d-Gln utilization in Pseudomonas aeruginosa PAO1. Microbiology 160 (2014) 2331-2340.

[102] Y. Nakada, Y. Itoh, Characterization and regulation of the gbuA gene, encoding guanidinobutyrase in the arginine dehydrogenase pathway of Pseudomonas aeruginosa PAO1. Journal of bacteriology 184 (2002) 3377-3384.

[103] Y. Nakada, Y. Itoh, Pseudomonas aeruginosa PAO1 genes for 3-guanidinopropionate and 4-guanidinobutyrate utilization may be derived from a common ancestor. Microbiology 151 (2005) 4055-4062.

[104] X. Yao, W. He, C.-D. Lu, Functional characterization of seven γ-glutamylpolyamine synthetase genes and the bauRABCD locus for polyamine and β-alanine utilization in Pseudomonas aeruginosa PAO1. Journal of bacteriology 193 (2011) 3923-3930.

[105] M.J. Wargo, D.A. Hogan, Identification of genes required for Pseudomonas aeruginosa carnitine catabolism. Microbiology 155 (2009) 2411-2419.

[106] G.G. Willsey, M.J. Wargo, Sarcosine catabolism in Pseudomonas aeruginosa is transcriptionally regulated by SouR. Journal of bacteriology 198 (2016) 301-310.

[107] A.E. Ballok, G.A. O'Toole, Pouring salt on a wound: Pseudomonas aeruginosa virulence factors alter Na+ and Cl− flux in the lung. Journal of bacteriology 195 (2013) 4013-4019.

[108] T. Strateva, I. Mitov, Contribution of an arsenal of virulence factors to pathogenesis of Pseudomonas aeruginosa infections. Annals of microbiology 61 (2011) 717-732.

[109] A.R. Hauser, Pseudomonas aeruginosa: so many virulence factors, so little time. Critical care medicine 39 (2011) 2193.

[110] J.A. Bartell, A.S. Blazier, P. Yen, J.C. Thøgersen, L. Jelsbak, J.B. Goldberg, J.A. Papin, Reconstruction of the metabolic network of Pseudomonas aeruginosa to interrogate virulence factor synthesis. Nature communications 8 (2017) 14631.

[111] G. Xiao, J. He, L.G. Rahme, Mutation analysis of the Pseudomonas aeruginosa mvfR and pqsABCDE gene promoters demonstrates complex quorum-sensing circuitry. Microbiology 152 (2006) 1679-1686.

[112] E. Deziel, S. Gopalan, A.P. Tampakaki, F. Lepine, K.E. Padfield, M. Saucier, G. Xiao, L.G. Rahme, The contribution of MvfR to Pseudomonas aeruginosa pathogenesis and quorum sensing circuitry regulation: multiple quorum sensing-regulated genes are modulated without affecting lasRI, rhlRI or the production of N-acyl-L-homoserine lactones. Mol Microbiol 55 (2005) 998-1014.

[113] G. Xiao, E. Deziel, J. He, F. Lepine, B. Lesic, M.H. Castonguay, S. Milot, A.P. Tampakaki, S.E. Stachel, L.G. Rahme, MvfR, a key Pseudomonas aeruginosa pathogenicity LTTR-class regulatory protein, has dual ligands. Mol Microbiol 62 (2006) 1689-1699.

[114] D. Maura, R. Hazan, T. Kitao, A.E. Ballok, L.G. Rahme, Evidence for Direct Control of Virulence and Defense Gene Circuits by the Pseudomonas aeruginosa Quorum Sensing Regulator, MvfR. Sci Rep 6 (2016) 34083.

[115] D. Balasubramanian, H. Kumari, M. Jaric, M. Fernandez, K.H. Turner, S.L. Dove, G. Narasimhan, S. Lory, K. Mathee, Deep sequencing analyses expands the Pseudomonas aeruginosa AmpR regulon to include small RNA-mediated regulation of iron acquisition, heat shock and oxidative stress response. Nucleic Acids Res 42 (2014) 979-998.

[116] T. Vinckx, Q. Wei, S. Matthijs, P. Cornelis, The Pseudomonas aeruginosa oxidative stress regulator OxyR influences production of pyocyanin and rhamnolipids: protective role of pyocyanin. Microbiology 156 (2010) 678-686.

[117] Q. Guo, W. Kong, S. Jin, L. Chen, Y. Xu, K. Duan, PqsR-dependent and PqsR-independent regulation of motility and biofilm formation by PQS in Pseudomonas aeruginosa PAO1. J Basic Microbiol 54 (2014) 633-643.

[118] Z. Cai, Y. Liu, Y. Chen, J.K. Yam, S.C. Chew, S.L. Chua, K. Wang, M. Givskov, L. Yang, RpoN Regulates Virulence Factors of Pseudomonas aeruginosa via Modulating the PqsR Quorum Sensing Regulator. Int J Mol Sci 16 (2015) 28311-28319.

[119] S. Boopathi, R. Vashisth, P. Manoharan, R. Kandasamy, N. Sivakumar, Stigmatellin Y - An anti-biofilm compound from Bacillus subtilis BR4 possibly interferes in PQS-PqsR mediated quorum sensing system in Pseudomonas aeruginosa. Bioorg Med Chem Lett 27 (2017) 2113-2118.

[120] J.M. Farrow, 3rd, Z.M. Sund, M.L. Ellison, D.S. Wade, J.P. Coleman, E.C. Pesci, PqsE functions independently of PqsR-Pseudomonas quinolone signal and enhances the rhl quorum-sensing system. J Bacteriol 190 (2008) 7043-7051.

[121] T. Klein, C. Henn, J.C. de Jong, C. Zimmer, B. Kirsch, C.K. Maurer, D. Pistorius, R. Muller, A. Steinbach, R.W. Hartmann, Identification of small-molecule antagonists of the Pseudomonas aeruginosa transcriptional regulator PqsR: biophysically guided hit discovery and optimization. ACS Chem Biol 7 (2012) 1496-1501.

[122] C. Lu, B. Kirsch, C. Zimmer, J.C. de Jong, C. Henn, C.K. Maurer, M. Musken, S. Haussler, A. Steinbach, R.W. Hartmann, Discovery of antagonists of PqsR, a key player in 2-alkyl-4-quinolone-dependent quorum sensing in Pseudomonas aeruginosa. Chem Biol 19 (2012) 381-390.

[123] C. Lu, B. Kirsch, C.K. Maurer, J.C. de Jong, A. Braunshausen, A. Steinbach, R.W. Hartmann, Optimization of anti-virulence PqsR antagonists regarding aqueous solubility and biological properties resulting in new insights in structure-activity relationships. Eur J Med Chem 79 (2014) 173-183.

[124] K. Kefala, D. Kotsifaki, M. Providaki, E.G. Kapetaniou, L. Rahme, M. Kokkinidis, Purification, crystallization and preliminary X-ray diffraction analysis of the C-terminal fragment of the MvfR protein from Pseudomonas aeruginosa. Acta Crystallogr Sect F Struct Biol Cryst Commun 68 (2012) 695-697.

[125] A. Ilangovan, M. Fletcher, G. Rampioni, C. Pustelny, K. Rumbaugh, S. Heeb, M. Camara, A. Truman, S.R. Chhabra, J. Emsley, P. Williams, Structural basis for native agonist and synthetic inhibitor recognition by the Pseudomonas aeruginosa quorum sensing regulator PqsR (MvfR). PLoS Pathog 9 (2013) e1003508.

[126] J.M. Gaines, N.L. Carty, J.A. Colmer-Hamood, A.N. Hamood, Effect of static growth and different levels of environmental oxygen on toxA and ptxR expression in the Pseudomonas aeruginosa strain PAO1. Microbiology 151 (2005) 2263-2275.

[127] J. Morlon-Guyot, J. Méré, A. Bonhoure, B. Beaumelle, Processing of Pseudomonas aeruginosa exotoxin A is dispensable for cell intoxication. Infection and immunity 77 (2009) 3090-3099.

[128] A.N. Hamood, J.A. Colmer, U.A. Ochsner, M.L. Vasil, Isolation and characterization of a Pseudomonas aeruginosa gene, ptxR, which positively regulates exotoxin A production. Mol Microbiol 21 (1996) 97-110.

[129] J.A. Colmer, A.N. Hamood, Characterization of ptxS, a Pseudomonas aeruginosa gene which interferes with the effect of the exotoxin A positive regulatory gene, ptxR. Mol Gen Genet 258 (1998) 250-259.

[130] J.A. Colmer, A.N. Hamood, Molecular analysis of the Pseudomonas aeruginosa regulatory genes ptxR and ptxS. Can J Microbiol 47 (2001) 820-828.

[131] A. Daddaoua, S. Fillet, M. Fernandez, Z. Udaondo, T. Krell, J.L. Ramos, Genes for carbon metabolism and the ToxA virulence factor in Pseudomonas aeruginosa are regulated through molecular interactions of PtxR and PtxS. PLoS One 7 (2012) e39390.

[132] N.L. Carty, N. Layland, J.A. Colmer-Hamood, M.W. Calfee, E.C. Pesci, A.N. Hamood, PtxR modulates the expression of QS-controlled virulence factors in the Pseudomonas aeruginosa strain PAO1. Mol Microbiol 61 (2006) 782-794.

[133] A.S. Blier, W. Veron, A. Bazire, E. Gerault, L. Taupin, J. Vieillard, K. Rehel, A. Dufour, F. Le Derf, N. Orange, C. Hulen, M.G. Feuilloley, O. Lesouhaitier, C-type natriuretic peptide modulates quorum sensing molecule and toxin production in Pseudomonas aeruginosa. Microbiology 157 (2011) 1929-1944.

[134] A. Daddaoua, T. Krell, J.L. Ramos, Transcriptional control by two interacting regulatory proteins: identification of the PtxS binding site at PtxR. Nucleic Acids Res 41 (2013) 10150-10156.

[135] I. Vallet-Gely, J.S. Sharp, S.L. Dove, Local and global regulators linking anaerobiosis to cupA fimbrial gene expression in Pseudomonas aeruginosa. Journal of bacteriology 189 (2007) 8667-8676.

[136] K.H. Turner, I. Vallet-Gely, S.L. Dove, Epigenetic control of virulence gene expression in Pseudomonas aeruginosa by a LysR-type transcription regulator. PLoS Genet 5 (2009) e1000779.

[137] R.R. McCarthy, M.J. Mooij, F.J. Reen, O. Lesouhaitier, F. O'Gara, A new regulator of pathogenicity (bvlR) is required for full virulence and tight microcolony formation in Pseudomonas aeruginosa. Microbiology 160 (2014) 1488-1500.

[138] H. Sondermann, N.J. Shikuma, F.H. Yildiz, You've come a long way: c-di-GMP signaling. Current opinion in microbiology 15 (2012) 140-146.

[139] I. Kukavica-Ibrulj, F. Sanschagrin, A. Peterson, M. Whiteley, B. Boyle, J. Mackay, R.C. Levesque, Functional genomics of PycR, a LysR family transcriptional regulator essential for maintenance of Pseudomonas aeruginosa in the rat lung. Microbiology 154 (2008) 2106-2118.

[140] Q. Wei, P.N. Minh, A. Dotsch, F. Hildebrand, W. Panmanee, A. Elfarash, S. Schulz, S. Plaisance, D. Charlier, D. Hassett, S. Haussler, P. Cornelis, Global regulation of gene expression by OxyR in an important human opportunistic pathogen. Nucleic Acids Res 40 (2012) 4320-4333.

[141] A. Romsang, S. Atichartpongkul, W. Trinachartvanit, P. Vattanaviboon, S. Mongkolsuk, Gene expression and physiological role of Pseudomonas aeruginosa methionine sulfoxide reductases during oxidative stress. Journal of bacteriology (2013) JB. 00167-00113.

[142] L. Lan, T.S. Murray, B.I. Kazmierczak, C. He, Pseudomonas aeruginosa OspR is an oxidative stress sensing regulator that affects pigment production, antibiotic resistance and dissemination during infection. Molecular microbiology 75 (2010) 76-91.

[143] U.A. Ochsner, M.L. Vasil, E. Alsabbagh, K. Parvatiyar, D.J. Hassett, Role of the Pseudomonas aeruginosa oxyR-recG operon in oxidative stress defense and DNA repair: OxyR-dependent regulation of katB-ankB, ahpB, and ahpC-ahpF. J Bacteriol 182 (2000) 4533-4544.

[144] G. Storz, L.A. Tartaglia, B.N. Ames, The oxyR regulon. Antonie Van Leeuwenhoek 58 (1990) 157-161.

[145] W. Panmanee, F. Gomez, D. Witte, V. Pancholi, B.E. Britigan, D.J. Hassett, The peptidoglycan-associated lipoprotein OprL helps protect a Pseudomonas aeruginosa mutant devoid of the transactivator OxyR from hydrogen peroxide-mediated killing during planktonic and biofilm culture. J Bacteriol 190 (2008) 3658-3669.

[146] Y.J. Heo, I.Y. Chung, W.J. Cho, B.Y. Lee, J.H. Kim, K.H. Choi, J.W. Lee, D.J. Hassett, Y.H. Cho, The major catalase gene (katA) of Pseudomonas aeruginosa PA14 is under both positive and negative control of the global transactivator OxyR in response to hydrogen peroxide. J Bacteriol 192 (2010) 381-390.

[147] Z.C. Yuan, R. Zaheer, T.M. Finan, Phosphate limitation induces catalase expression in Sinorhizobium meliloti, Pseudomonas aeruginosa and Agrobacterium tumefaciens. Mol Microbiol 58 (2005) 877-894.

[148] Y.S. Choi, D.H. Shin, I.Y. Chung, S.H. Kim, Y.J. Heo, Y.H. Cho, Identification of Pseudomonas aeruginosa genes crucial for hydrogen peroxide resistance. J Microbiol Biotechnol 17 (2007) 1344-1352.

[149] G.W. Lau, B.E. Britigan, D.J. Hassett, Pseudomonas aeruginosa OxyR is required for full virulence in rodent and insect models of infection and for resistance to human neutrophils. Infect Immun 73 (2005) 2550-2553.

[150] H.W. Bae, Y.H. Cho, Mutational analysis of Pseudomonas aeruginosa OxyR to define the regions required for peroxide resistance and acute virulence. Res Microbiol 163 (2012) 55-63.

[151] W. Panmanee, N. Charoenlap, S. Atichartpongkul, A. Mahavihakanont, M.D. Whiteside, G. Winsor, F.S.L. Brinkman, S. Mongkolsuk, D.J. Hassett, The OxyR-regulated phnW gene encoding 2-aminoethylphosphonate:pyruvate aminotransferase helps protect Pseudomonas aeruginosa from tert-butyl hydroperoxide. PLoS One 12 (2017) e0189066.

[152] K.A. Melstrom, Jr., R. Kozlowski, D.J. Hassett, H. Suzuki, D.M. Bates, R.L. Gamelli, R. Shankar, Cytotoxicity of Pseudomonas secreted exotoxins requires OxyR expression. J Surg Res 143 (2007) 50-57.

[153] T. Vinckx, S. Matthijs, P. Cornelis, Loss of the oxidative stress regulator OxyR in Pseudomonas aeruginosa PAO1 impairs growth under iron-limited conditions. FEMS Microbiol Lett 288 (2008) 258-265.

[154] X. Wang, M. Li, L. Liu, R. Mou, X. Zhang, Y. Bai, H. Xu, M. Qiao, DsbM, a novel disulfide oxidoreductase affects aminoglycoside resistance in Pseudomonas aeruginosa by OxyR-regulated response. J Microbiol 50 (2012) 932-938.

[155] M. Li, X. Guan, X. Wang, H. Xu, Y. Bai, X. Zhang, M. Qiao, DsbM affects aminoglycoside resistance in Pseudomonas aeruginosa by the reduction of OxyR. FEMS Microbiol Lett 352 (2014) 184-189.

[156] I. Jo, N. Park, I.Y. Chung, Y.H. Cho, N.C. Ha, Crystal structures of the disulfide reductase DsbM from Pseudomonas aeruginosa. Acta Crystallogr D Struct Biol 72 (2016) 1100-1109.

[157] J. Yeom, C.O. Jeon, E.L. Madsen, W. Park, In vitro and in vivo interactions of ferredoxin-NADP+ reductases in Pseudomonas putida. Journal of biochemistry 145 (2009) 481-491.

[158] A. Romsang, J. Duang-nkern, W. Wirathorn, P. Vattanaviboon, S. Mongkolsuk, Pseudomonas aeruginosa IscR-Regulated Ferredoxin NADP (+) Reductase Gene (fprB) Functions in Iron-Sulfur Cluster Biogenesis and Multiple Stress Response. PloS one 10 (2015) e0134374.

[159] R. Yan, S. Adinolfi, A. Pastore, Ferredoxin, in conjunction with NADPH and ferredoxin-NADP reductase, transfers electrons to the IscS/IscU complex to promote iron-sulfur cluster assembly. Biochim Biophys Acta 1854 (2015) 1113-1117.

[160] J. Kim, W. Park, Oxidative stress response in Pseudomonas putida. Applied microbiology biotechnology 98 (2014) 6933-6946.

[161] S. Yeom, J. Yeom, W. Park, Molecular characterization of FinR, a novel redox-sensing transcriptional regulator in Pseudomonas putida KT2440. Microbiology 156 (2010) 1487-1496.

[162] Y. Lee, S. Peña-Llopis, Y.-S. Kang, H.-D. Shin, B. Demple, E.L. Madsen, C.O. Jeon, W. Park, Expression analysis of the fpr (ferredoxin-NADP+ reductase) gene in Pseudomonas putida KT2440. Biochemical biophysical research communications 339 (2006) 1246-1254.

[163] F.J. Reen, J.M. Haynes, M.J. Mooij, F. O'Gara, A non-classical LysR-type transcriptional regulator PA2206 is required for an effective oxidative stress response in Pseudomonas aeruginosa. PLoS One 8 (2013) e54479.

[164] E.B. Breidenstein, C. de la Fuente-Nunez, R.E. Hancock, Pseudomonas aeruginosa: all roads lead to resistance. Trends Microbiol 19 (2011) 419-426.

[165] P. Lambert, Mechanisms of antibiotic resistance in Pseudomonas aeruginosa. Journal of the royal society of medicine 95 (2002) 22.

[166] Z. Pang, R. Raudonis, B.R. Glick, T.-J. Lin, Z. Cheng, Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies. Biotechnology advances (2018).

[167] J.I. Campbell, O. Ciofu, N. Hoiby, Pseudomonas aeruginosa isolates from patients with cystic fibrosis have different beta-lactamase expression phenotypes but are homogeneous in the ampC-ampR genetic region. Antimicrob Agents Chemother 41 (1997) 1380-1384.

[168] C. Juan, M.D. Macia, O. Gutierrez, C. Vidal, J.L. Perez, A. Oliver, Molecular mechanisms of beta-lactam resistance mediated by AmpC hyperproduction in Pseudomonas aeruginosa clinical strains. Antimicrob Agents Chemother 49 (2005) 4733-4738.

[169] J. Quale, S. Bratu, J. Gupta, D. Landman, Interplay of efflux system, ampC, and oprD expression in carbapenem resistance of Pseudomonas aeruginosa clinical isolates. Antimicrob Agents Chemother 50 (2006) 1633-1641.

[170] V.H. Tam, A.N. Schilling, M.T. LaRocco, L.O. Gentry, K. Lolans, J.P. Quinn, K.W. Garey, Prevalence of AmpC over-expression in bloodstream isolates of Pseudomonas aeruginosa. Clin Microbiol Infect 13 (2007) 413-418.

[171] D.J. Wolter, A.J. Schmidtke, N.D. Hanson, P.D. Lister, Increased expression of ampC in Pseudomonas aeruginosa mutants selected with ciprofloxacin. Antimicrob Agents Chemother 51 (2007) 2997-3000.

[172] H. Kumari, D. Balasubramanian, D. Zincke, K. Mathee, Role of Pseudomonas aeruginosa AmpR on beta-lactam and non-beta-lactam transient cross-resistance upon pre-exposure to subinhibitory concentrations of antibiotics. J Med Microbiol 63 (2014) 544-555.

[173] G. Cabot, C. Lopez-Causape, A.A. Ocampo-Sosa, L.M. Sommer, M.A. Dominguez, L. Zamorano, C. Juan, F. Tubau, C. Rodriguez, B. Moya, C. Pena, L. Martinez-Martinez, P. Plesiat, A. Oliver, Deciphering the Resistome of the Widespread Pseudomonas aeruginosa Sequence Type 175 International High-Risk Clone through Whole-Genome Sequencing. Antimicrob Agents Chemother 60 (2016) 7415-7423.

[174] D. Balasubramanian, L. Schneper, M. Merighi, R. Smith, G. Narasimhan, S. Lory, K. Mathee, The regulatory repertoire of Pseudomonas aeruginosa AmpC ss-lactamase regulator AmpR includes virulence genes. PLoS One 7 (2012) e34067.

[175] H. Kumari, S.K. Murugapiran, D. Balasubramanian, L. Schneper, M. Merighi, D. Sarracino, S. Lory, K. Mathee, LTQ-XL mass spectrometry proteome analysis expands the Pseudomonas aeruginosa AmpR regulon to include cyclic di-GMP phosphodiesterases and phosphoproteins, and identifies novel open reading frames. J Proteomics 96 (2014) 328-342.

[176] D.A. Dik, T. Dominguez-Gil, M. Lee, D. Hesek, B. Byun, J. Fishovitz, B. Boggess, L.M. Hellman, J.F. Fisher, J.A. Hermoso, S. Mobashery, Muropeptide Binding and the X-ray Structure of the Effector Domain of the Transcriptional Regulator AmpR of Pseudomonas aeruginosa. J Am Chem Soc 139 (2017) 1448-1451.

[177] P.D. Lister, D.J. Wolter, N.D. Hanson, Antibacterial-resistant Pseudomonas aeruginosa: clinical impact and complex regulation of chromosomally encoded resistance mechanisms. Clinical microbiology reviews 22 (2009) 582-610.

[178] H. Maseda, K. Saito, A. Nakajima, T. Nakae, Variation of the mexT gene, a regulator of the MexEF-oprN efflux pump expression in wild-type strains of Pseudomonas aeruginosa. FEMS Microbiol Lett 192 (2000) 107-112.

[179] H. Maseda, M. Uwate, T. Nakae, Transcriptional regulation of the mexEF-oprN multidrug efflux pump operon by MexT and an unidentified repressor in nfxC-type mutant of Pseudomonas aeruginosa. FEMS Microbiol Lett 311 (2010) 36-43.

[180] M. Uwate, Y.K. Ichise, A. Shirai, T. Omasa, T. Nakae, H. Maseda, Two routes of MexS-MexT-mediated regulation of MexEF-OprN and MexAB-OprM efflux pump expression in Pseudomonas aeruginosa. Microbiol Immunol 57 (2013) 263-272.

[181] M.L. Sobel, S. Neshat, K. Poole, Mutations in PA2491 (mexS) promote MexT-dependent mexEF-oprN expression and multidrug resistance in a clinical strain of Pseudomonas aeruginosa. J Bacteriol 187 (2005) 1246-1253.

[182] P.M. Luong, B.D. Shogan, A. Zaborin, N. Belogortseva, J.D. Shrout, O. Zaborina, J.C. Alverdy, Emergence of the P2 phenotype in Pseudomonas aeruginosa PAO1 strains involves various mutations in mexT or mexF. J Bacteriol 196 (2014) 504-513.

[183] P. Juarez, I. Broutin, C. Bordi, P. Plesiat, C. Llanes, Constitutive Activation of MexT by Amino Acid Substitutions Results in MexEF-OprN Overproduction in Clinical Isolates of Pseudomonas aeruginosa. Antimicrob Agents Chemother 62 (2018).

[184] J. Weile, R.D. Schmid, T.T. Bachmann, M. Susa, C. Knabbe, DNA microarray for genotyping multidrug-resistant Pseudomonas aeruginosa clinical isolates. Diagn Microbiol Infect Dis 59 (2007) 325-338.

[185] L. Pumbwe, L.J. Piddock, Two efflux systems expressed simultaneously in multidrug-resistant Pseudomonas aeruginosa. Antimicrob Agents Chemother 44 (2000) 2861-2864.

[186] V.C. Thai, H.V. Pham, D.N.M. Nguyen, P. Lambert, T.T.H. Nguyen, A deletion mutation in nfxB of in vitro-induced moxifloxacin-resistant Pseudomonas aeruginosa confers multidrug resistance. Acta Microbiol Immunol Hung 64 (2017) 245-253.

[187] T. Kohler, S.F. Epp, L.K. Curty, J.C. Pechere, Characterization of MexT, the regulator of the MexE-MexF-OprN multidrug efflux system of Pseudomonas aeruginosa. J Bacteriol 181 (1999) 6300-6305.

[188] H. Fetar, C. Gilmour, R. Klinoski, D.M. Daigle, C.R. Dean, K. Poole, mexEF-oprN multidrug efflux operon of Pseudomonas aeruginosa: regulation by the MexT activator in response to nitrosative stress and chloramphenicol. Antimicrob Agents Chemother 55 (2011) 508-514.

[189] Y. Jin, H. Yang, M. Qiao, S. Jin, MexT regulates the type III secretion system through MexS and PtrC in Pseudomonas aeruginosa. J Bacteriol 193 (2011) 399-410.

[190] Z.X. Tian, M. Mac Aogain, H.F. O'Connor, E. Fargier, M.J. Mooij, C. Adams, Y.P. Wang, F. O'Gara, MexT modulates virulence determinants in Pseudomonas aeruginosa independent of the MexEF-OprN efflux pump. Microb Pathog 47 (2009) 237-241.

[191] A.D. Olivas, B.D. Shogan, V. Valuckaite, A. Zaborin, N. Belogortseva, M. Musch, F. Meyer, W.L. Trimble, G. An, J. Gilbert, O. Zaborina, J.C. Alverdy, Intestinal tissues induce an SNP mutation in Pseudomonas aeruginosa that enhances its virulence: possible role in anastomotic leak. PLoS One 7 (2012) e44326.

[192] M. Hogardt, J. Heesemann, Adaptation of Pseudomonas aeruginosa during persistence in the cystic fibrosis lung. Int J Med Microbiol 300 (2010) 557-562.

[193] R.D. Oshri, K.S. Zrihen, I. Shner, S. Omer Bendori, A. Eldar, Selection for increased quorum-sensing cooperation in Pseudomonas aeruginosa through the shut-down of a drug resistance pump. ISME J 12 (2018) 2458-2469.

[194] C. Zaoui, J. Overhage, D. Lons, A. Zimmermann, M. Musken, P. Bielecki, C. Pustelny, T. Becker, M. Nimtz, S. Haussler, An orphan sensor kinase controls quinolone signal production via MexT in Pseudomonas aeruginosa. Mol Microbiol 83 (2012) 536-547.

[195] E. Fargier, M. Mac Aogain, M.J. Mooij, D.F. Woods, J.P. Morrissey, A.D. Dobson, C. Adams, F. O'Gara, MexT functions as a redox-responsive regulator modulating disulfide stress resistance in Pseudomonas aeruginosa. J Bacteriol 194 (2012) 3502-3511.

[196] A. Kumar, H.P. Schweizer, Evidence of MexT-independent overexpression of MexEF-OprN multidrug efflux pump of Pseudomonas aeruginosa in presence of metabolic stress. PLoS One 6 (2011) e26520.

[197] C.W. Hall, L. Zhang, T.F. Mah, PA3225 Is a Transcriptional Repressor of Antibiotic Resistance Mechanisms in Pseudomonas aeruginosa. Antimicrob Agents Chemother 61 (2017).
[198] S. Borukhov, E. Nudler, RNA polymerase: the vehicle of transcription. Trends in microbiology 16 (2008) 126-134.

[199] E. Nudler, RNA Polymerase Active Center: The Molecular Engine of Transcription. 78 (2009) 335-361.

[200] E.A. Galburt, S.W. Grill, A. Wiedmann, L. Lubkowska, J. Choy, E. Nogales, M. Kashlev, C. Bustamante, Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. Nature 446 (2007) 820.

[201] M.D. Wang, M.J. Schnitzer, H. Yin, R. Landick, J. Gelles, S.M. Block, Force and velocity measured for single molecules of RNA polymerase. Science 282 (1998) 902-907.

[202] S.A. Darst, Bacterial RNA polymerase. Current opinion in structural biology 11 (2001) 155-162.

[203] J.J. Furth, J. Hurwitz, M. Anders, The role of deoxyribonucleic acid in ribonucleic acid synthesis. I. The purification and properties of ribonucleic acid polymerase. J Biol Chem 237 (1962) 2611-2619.

[204] J. Hurwitz, A. Bresler, R. Diringer, The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA. Biochemical biophysical research communications 3 (1960) 15-19.

[205] C.A. Gross, C.L. Chan, A structure/function analysis of Escherichia coli RNA polymerase. Phil Trans R Soc Lond B 351 (1996) 475-482.

[206] K.S. Murakami, Structural biology of bacterial RNA polymerase. Biomolecules 5 (2015) 848-864.

[207] A. Hirata, B.J. Klein, K.S. Murakami, The X-ray crystal structure of RNA polymerase from Archaea. Nature 451 (2008) 851.

[208] P. Cramer, D.A. Bushnell, J. Fu, A.L. Gnatt, B. Maier-Davis, N.E. Thompson, R.R. Burgess, A.M. Edwards, P.R. David, R.D. Kornberg, Architecture of RNA polymerase II and implications for the transcription mechanism. Science 288 (2000) 640-649.

[209] P. Cramer, D.A. Bushnell, R.D. Kornberg, Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution. Science 292 (2001) 1863-1876.

[210] T. Ghosh, D. Bose, X. Zhang, Mechanisms for activating bacterial RNA polymerase. FEMS microbiology reviews 34 (2010) 611-627.

[211] R.R. Burgess, RNA polymerase. Annu Rev Biochem 40 (1971) 711-740.

[212] R.R. Burgess, C. Gross, F. Engbaek, Cysteine and methionine content of the Escherichia coli ribonucleic acid polymerase subunits. J Bacteriol 128 (1976) 390-393.

[213] F. Engbaek, C. Gross, R.R. Burgess, Biosynthesis of Escherichia coli RNA polymerase subunits upon release of rifampicin inhibition. Mol Gen Genet 143 (1976) 297-299.

[214] C. Gross, F. Engbaek, T. Flammang, R. Burgess, Rapid micromethod for the purification of Escherichia coli ribonucleic acid polymerase and the preparation of bacterial extracts active in ribonucleic acid synthesis. J Bacteriol 128 (1976) 382-389.

[215] (!!! INVALID CITATION !!! {}).

[216] S.R. Jaskunas, R.R. Burgess, M. Nomura, Identification of a gene for the alpha-subunit of RNA polymerase at the str-spc region of the Escherichia coli chromosome. Proceedings of the National Academy of Sciences 72 (1975) 5036-5040.

[217] K. Igarashi, A. Ishihama, Bipartite functional map of the E. coli RNA polymerase α subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. Cell 65 (1991) 1015-1022.

[218] S. Busby, R. Ebright, Promoter structure, promoter recognition, and transcription activation in prokaryotes. Cell 79 (1994) 743-746.

[219] Y. Chen, Y.W. Ebright, R.H. Ebright, Identification of the target of a transcription activator protein by protein-protein photocrosslinking. Science 265 (1994) 90-92.

[220] M. Kimura, A. Ishihama, Functional Map of the Alpha Subunit ofEscherichia coliRNA Polymerase: Amino Acid Substitution within the Amino-terminal Assembly Domain. Journal of molecular biology 254 (1995) 342-349.

[221] W. Ross, K.K. Gosink, J. Salomon, K. Igarashi, C. Zou, A. Ishihama, K. Severinov, R.L. Gourse, A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. Science 262 (1993) 1407-1413.

[222] A. Feklistov, B.D. Sharon, S.A. Darst, C.A. Gross, Bacterial sigma factors: a historical, structural, and genomic perspective. Annual review of microbiology 68 (2014) 357-376.

[223] M. Lonetto, M. Gribskov, C.A. Gross, The sigma 70 family: sequence conservation and evolutionary relationships. Journal of bacteriology 174 (1992) 3843.

[224] J.D. Helmann, M.J. Chamberlin, DNA sequence analysis suggests that expression of flagellar and chemotaxis genes in Escherichia coli and Salmonella typhimurium is controlled by an alternative sigma factor. Proceedings of the National Academy of Sciences 84 (1987) 6422-6424.

[225] M. Paget, Bacterial sigma factors and anti-sigma factors: structure, function and distribution. Biomolecules 5 (2015) 1245-1265.

[226] M. Li, H. Moyle, M. Susskind, Target of the transcriptional activation function of phage lambda cI protein. Science 263 (1994) 75-77.

[227] E. Ruff, M. Record, I. Artsimovitch, Initial events in bacterial transcription initiation. Biomolecules 5 (2015) 1035-1062.

[228] S.d.A. e Silva, S. Echeverrigaray, Bacterial promoter features description and their application on E. coli in silico prediction and recognition approaches, Bioinformatics, InTech, 2012.

[229] B. Lewin, Genes, IX, Jones and Bartlett Publishers, 2008.

[230] I.G. Hook-Barnard, D.M. Hinton, s. biology, Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. Gene Regulation and Systems Biology 1 (2007) 117762500700100020.

[231] P.L. deHaseth, M.L. Zupancic, M.T. Record, RNA Polymerase-Promoter Interactions: the Comings and Goings of RNA Polymerase. 180 (1998) 3019-3025.

[232] K.B. Decker, D.M. Hinton, Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. Annual review of microbiology 67 (2013) 113-139.

[233] J. Lee, S. Borukhov, Bacterial RNA Polymerase-DNA Interaction—The Driving Force of Gene Expression and the Target for Drug Action. Frontiers in molecular biosciences 3 (2016) 73.

[234] S.P. Haugen, M.B. Berkmen, W. Ross, T. Gaal, C. Ward, R.L. Gourse, rRNA promoter regulation by nonoptimal binding of σ region 1.2: an additional recognition element for RNA polymerase. Cell 125 (2006) 1069-1082.

[235] Y. Zhang, D. Degen, M.X. Ho, E. Sineva, K.Y. Ebright, Y.W. Ebright, V. Mekler, H. Vahedian-Movahed, Y. Feng, R. Yin, S. Tuske, H. Irschik, R. Jansen, S. Maffioli, S. Donadio, E. Arnold, R.H. Ebright, GE23077 binds to the RNA polymerase 'i' and 'i+1' sites and prevents the binding of initiating nucleotides. Elife 3 (2014) e02450.

[236] Y. Zuo, T.A. Steitz, Crystal structures of the E. coli transcription initiation complexes with a complete bubble. Molecular cell 58 (2015) 534-540.

[237] J.D. Helmann, RNA polymerase: a nexus of gene regulation, Elsevier, 2009.

[238] D.G. Vassylyev, I. Artsimovitch, Tracking RNA polymerase, one step at a time. Cell 123 (2005) 977-979.

[239] S.-i. Sekine, Y. Murayama, V. Svetlov, E. Nudler, S. Yokoyama, The ratcheted and ratchetable structural states of RNA polymerase underlie multiple transcriptional functions. Molecular cell 57 (2015) 408-421.

[240] A. Weixlbaumer, K. Leon, R. Landick, S.A. Darst, Structural basis of transcriptional pausing in bacteria. Cell 152 (2013) 431-441.

[241] D.G. Vassylyev, M.N. Vassylyeva, A. Perederina, T.H. Tahirov, I. Artsimovitch, Structural basis for transcription elongation by bacterial RNA polymerase. Nature 448 (2007) 157.

[242] D.G. Vassylyev, M.N. Vassylyeva, J. Zhang, M. Palangat, I. Artsimovitch, R. Landick, Structural basis for substrate loading in bacterial RNA polymerase. Nature 448 (2007) 163.

[243] J.W. Roberts, S. Shankar, J.J.J.A.R.M. Filter, RNA polymerase elongation factors. Annual Review of Biochemistry 62 (2008) 211-233.

[244] L. Nguyen, J. Garcia, K. Gruenberg, C. MacDougall, Multidrug-resistant pseudomonas infections: hard to treat, but hope on the horizon? Current infectious disease reports 20 (2018) 23.

[245] S. Santajit, N. Indrawattana, Mechanisms of antimicrobial resistance in ESKAPE pathogens. BioMed research international 2016 (2016).

Chapter 2

OVERPRODUCTION AND PURIFICATION OF HIGHLY ACTIVE RECOMBINANT PSEUDOMONAS AERUGINOSA STR. PAO1 RNA POLYMERASE HOLOENZYME COMPLEX

Derrick Afful, Liming Cai, and Cory Momany

Submitted to Protein Expression and Purification

2.1 ABSTRACT

The bacterial RNA polymerase (RNAP) is a large, complex molecular machine that is the engine of gene expression. Despite global conservation in their structures and function, RNAPs from different bacteria can have unique features in promoter and transcription factor recognition. Therefore, availability of purified RNAP from different bacteria is key to understanding these species-specific aspects and will be valuable for antibiotic drug discovery. *Pseudomonas aeruginosa* is one of the leading causes of hospital and community acquired infections worldwide - making the organism an important public health pathogen. We developed a method for producing high quantities of highly pure and active recombinant *P. aeruginosa* str. PAO1 RNAP core and holoenzyme complexes that employed two-vector systems for expressing the core enzyme (α, β, β', and ω subunits) and for expressing the holoenzyme complex (core + σ^{70}). Unlike other RNAP expression approaches, we used a low temperature autoinduction system in *E. coli* with T7 promoters that produced high cell yields and stable protein expression. The purification strategy comprised of four chromatographic separation steps (metal chelate, heparin, and ion-exchange) with yields of up to 10 mg per 500 mL culture. Purified holoenzyme and reconstituted holoenzyme from core and σ^{70} were highly active at transcribing both small and large-sized DNA templates, with a determined elongation rate of ~18 nt/s for the holoenzyme. The successful purification of the *P. aeruginosa* RNAP provides a gateway for studies focusing on *in vitro* transcriptional regulation in this pathogen.

Keywords:

RNA Polymerase; *Pseudomonas aeruginosa*; transcription

Topics:

Escherichia coli; Nucleic Acid Metabolism; Co-Expression of Several Protein; Nucleic Acid Binding; Transcription; Anion Exchange; Multisubunit-Complex

Highlights

This represents the first report of recombinant *Pseudomonas aeruginosa* RNA polymerase.

Plasmids for expression of both the core and holoenzymes were designed and built.

A low temperature autoinduction expression approach was used that resulted in high protein yields.

A four-step chromatographic purification scheme produced highly pure and stable RNA polymerase complexes.

Pseudomonas aeruginosa RNA polymerase was comparable in activity to commercial *E. coli* RNA polymerase.

2.2 INTRODUCTION

Bacterial RNA polymerases (RNAPs) play a cardinal role in gene expression and control as the source of newly synthesized RNA. Unsurprisingly, they have been the focus of numerous scientific studies since their discovery in the 1960s by Hurwitz [1]. The protein complex, considered one of the most structurally complicated enzymes in cells, is a large, multisubunit, molecular machine whose roles in regulatory networks are controlled by several cellular as well as environmental factors [2]. The RNAP basic architecture is conserved throughout bacteria. The two large subunits (β and β', encoded by *rpoB* and *rpoC*) make up the bulk of the enzyme and are together responsible for catalysis of RNA polymerase from ribonucleotide triphosphates, whereas the α subunit (encoded by *rpoA*) forms a dimer at the periphery of the complex and is responsible for assembly of the complex and regulatory interactions. An accessory subunit (ω), whose role has not been fully elucidated but is thought to be involved in regulation and assembly [3] , completes the formation of the approximately 400 kDa core enzyme complex ($\alpha_2\beta\beta'\omega$). The core enzyme is catalytically competent, but promoter-specific initiation requires the recruitment of a σ subunit, completing the formation of the holoenzyme complex [3-8]. Despite the high level of conservation in architecture and catalytic mechanism, RNAPs in different bacteria have subtly unique differences and exhibit species-specific properties with regards to promoter recognition and their interactions in the transcription activation complex [1, 27]. Particularly for pathogenic bacteria, structural and mechanistic elucidation of their specific RNAPs may provide a foundation for targeting of the enzymes for species-specific antibiotic drug discovery.

Pseudomonas aeruginosa str. PAO1, a Gram-negative bacterium, is an important public health pathogen and the leading cause of both community-acquired and nosocomial infections

worldwide [9, 10] and is the major cause of morbidity and mortality in cystic fibrosis patients [11]. Additionally, the organism is listed among the CDC's "ESKAPE" pathogens, the list of highly pathogenic bacteria for which only few medications are effective against [12]. These characteristics, together with its high antibiotic resistance properties [13-15], have made *P. aeruginosa* the target of many studies focusing on gene expression regulation as well as drug discovery [16-20]. However, a major impediment for such research groups has been the current unavailability *P. aeruginosa* RNAP. The only other pathogenic bacterium whose recombinant RNAP production and purification has been described is *Mycobacterium tuberculosis* [21-23]. The other three bacterial RNAPs that have been purified and characterized are the RNAPs from *Escherichia coli* [16, 24, 25], *Thermus aquaticus*/*Thermus thermophiles* [26, 27], and *Bacillus subtilis* [6, 28, 29].

In this study, we describe the overproduction and purification of the recombinant *Pseudomonas aeruginosa* str. PAO1 RNA polymerase core and holoenzyme complexes. The approach utilizes expression in *E. coli* from two plasmids encoding the genes for the different subunits of the enzyme behind T7 bacteriophage promoters and rare tRNAs. The use of autoinduction medium coupled with a multi-step purification approach produced high yields of purified enzyme. Various experiments validated the activity of the complexes.

2.3 MATERIALS AND METHODS

Cloning and Construction of Plasmids

All plasmids and strains used in the study, along with a description of each one, are listed in **Table 2-1**. Oligonucleotides, synthesized by Integrated DNA Technologies, are listed in **Table 2-2**. The *rpoA*, *rpoB*, *rpoC*, *rpoD* and rpoZ genes were PCR amplified from *P. aeruginosa* PAO1 genomic DNA using Phusion Hot Start High-Fidelity DNA Polymerase (Thermo Scientific). PCR amplicons were purified using a DNA Clean & Concentrator kit (Zymo Research) before cloning steps. All plasmid constructs were transformed into XL1-blue *E. coli* cells by electroporation to allow verification of the respective constructs by restriction digestion and DNA sequencing. Initially, a two-vector system was created for expression of RNAP comprised of plasmids pDAP5 and pDAP7. Plasmid pDAP5 is based on the pRARE2 plasmid that was isolated from Rosetta™ 2 (DE3) cells into which the *rpoB* and *rpoC* genes were cloned behind a single T7 promoter. Plasmid pDAP7 is based on a modified pET28b backbone into which *rpoA*, *rpoZ*, and *rpoD* genes were cloned behind a single T7 promoter. To create pDAP5 (**Figure 2-1A**), an intermediate pRARE2 vector containing the T7 promoter/T7 terminator region of pET28b was created. The PCR amplicon generated by primers T7QI-into-pACYC-F and T7QI-into-pACYC-R using plasmid pET28b-SapKO-WT as the PCR template (to extract the T7 promoter/terminator) was cut and blunt ligated into the BsaBI site (GATNN|NNATC) of pRARE2 following procedures outlined in [30] for BspQI/BsaI cloning (50°C precut, room temperature cut/ligation).

Table 2-1. Bacterial strains and plasmids used in this study

Name	Description	Source
***E. coli* strain**		
XL1-Blue	F^+ recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIq ZΔM15 Tn10 (Tet^R)]	Agilent
BL21(DE3)	F^- *fhuA2 lon ompT gal (λ DE3) [dcm]* Δ*hsdSλ* *DE3 = λ sBamHIo* Δ*EcoRI-B int::(lacI::PlacUV5::T7 gene1) i21* Δ*nin5*	New England Biolabs
Rosetta™ 2 (DE3)	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal (λ DE3) dcm* / pRARE2 (Cam^R)	Novagen
Plasmids		
pET28b(+)	T7*lac* promoter, Kan^R, pBR322 ori, 6x-His	Novagen
pRARE2	Contains rare codons inserted into pACYC184 backbone, Cam^R, p15A ori, isolated from Rosetta™ 2 *E. coli* cells	Novagen
pCDFDuet-1	T7*lac* promoter, Str/Spc^R, CloDF13 ori, 6x-His, S-Tag	Novagen
pET28b-SapKO-WT	pET28b(+) containing dual BspQI cloning sites inserted between the start codon and two stop codons.	[25]
pET28b-SapKO-CH	pET28b(+) containing dual BspQI cloning sites inserted between the start codon and a C-terminal 5X-His with two stop codons.	[25]
pET28b-SapKO-BsaI	pET28b(+) containing dual BsaI cloning sites inserted between the start codon and a C-terminal 5X-His with two stop codons.	[25]
pDAP5	*rpoB* and rpoC cloned into pRARE2 between T7 promoter/terminator	This work
pDAP7	*rpoA, rpoZ, rpoD* with N-terminal 5x-His-TEV in pET28b-SapKO-CH	This work
pDAP10*	rpoAZ in pET28b-SapKO-BsaI	This work
pDAP15*	N-terminal 10xHis-TEV, *rpoD*, in pCDFDuet-1	This work
pDAP18*	pDAP5 with T7 promoter and N-terminal10xHis-TEV sequence inserted upstream of *rpoC*	This work
pDAP22*	*rpoA* and *rpoZ* amplified from pDAP10 and ligated to PCR-amplified pDAP15. Contains *rpoA*, *rpoZ*, and N-terminal 11xHis-TEV-*rpoD*	This work
T7 DNA	T7 bacteriophage genomic DNA	Boca Scientific

*Final plasmids used for RNAP core and holoenzyme production.

Table 2-2. Oligonucleotides used in building plasmids

Name	Sequence*
T7QI-into-pACYC-F	CAGGTCTCCGATCCCGCGAAATTAATACG
T7QI-into-pACYC-R	GGGGTCTCAGATCCGGATATAGTTCCTCCTTTCAGC
rpoZ-Gblock1 [+]	GAACGCTCTTCAATGAGAGACCAAGGTCTCGTAACCTCTTGAAGGA GATATACCATGGCCCGCGTCACCGTTGAAGACTGCCTGGACAACGT CGATAACCGTTTCGAGCTGGTCATGCTCGCCACCAAGCGCGCCCGT CAGCTGGCTACCGGCGGCAAGGAGCCGAAAGTGGCCTGGGAAAAC GACAAGCCGACCGTCGTCGCCCTGCGCGAGATCGCTTCCGGCCTGG TCGATGAGAACGTCGTCCAGCAGGAAGACATCGTCGAGGACGAAC CGCTGTTCGCAGCGTTCGACGACGAGGCCAACACCGAGGCCCTGT AACAAGAAGGAGATATACCATGCACCATCATCATCACGAGAATCT CTACTTCCA**AGGCCT**CACAGAAGAGCGGG
Gblock2 [+]	GGGGATTTAAATCGTATTGTACACGGCCGCATAATCGAAATTAATA CGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCATCTT AGTATATTAGTTAAGTATAAGAAGGAGATATACATATGCACCATCA TCATCACCACCATCATCATCACGAGAATCTCTACTTCCA**AGGCCT**g Cgg
PAO1_rpoA-F	GGGGTCTCACATGCAGAGTTCGGTAAATGAG
PAO1_rpoA-R	GGGGTCTCGGTTATGCAGTGGCCTTGTCGTC
PAO1_rpoB_partial-F	GGGCTCTTCAATGGCTTACTCATACACTGAGAAAAAACG
PAO1_rpoB_partial-R	GCTGCTCTTCGAGCATGCGG
PAO1_rpoBC-F	GGGCTCTTCTGCTCGAGGAACAGCGCAAGGTCG
PAO1_rpoBC-R	GGGCTCTTCAGTGTTAGTTACCGCTCGAGTTCAG
PAO1_rpoD-F	CATGTCCGGAAAAGCGCAACAGC
PAO1_rpoD-R	TTATTACTCGTCGAGGAAGGAGCGAAG
T7Tev-pDAP5-F	C**ATG**AAAGACTTGCTTAATCTGTTGAAAAACC
T7Tev-pDAP5-F	A**TTA**TTCGGTTTCCAGTTCGATGTCG
T7TEV_from_Gblock2-F	GGGGATTTAAATCGTATTGTACAC
T7TEV_from_Gblock2-R	CCGCAGGCCTTGGAAGTAG
rpoAZ_from_pDAP7-F	TAGGTCTCACATGCAGAGTTCGGTAAATGAGTTC
rpoAZ_from_pDAP7-R	TAGGTCTCTGGTGTTATTACAGGGCCTCGGTGTTGG
pCDFDuet_PmlI-F	ATGCTCTTCTGTGGCCAGGATCCGAATTCGAG
pCDFDuet_PmlI-R	ATGCTCTTCACACGTGGTGATGATGGTGATGGCT
rpoD_from_pDAP7-F	CACCATCATCATCACGAGAATCTC
rpoD_from_pDAP7-R	TTATTACTCGTCGAGGAAGGAGC
rpoAZ-HiFi-F	TGCTTCCGGTAGTCAATAAAGGTGATGTCGGCGATATAGGCG
rpoAZ-HiFi-R	GTCTATTGCTGGTTTACCGGCCCATTCGCCAGACTGGACATC
rpoD-HiFi-F	GCCGGTAAACCAGCAATAGACATAAGC
rpoD-HiFi-R	CTTTATTGACTACCGGAAGCAGTGTGACC
Sequencing primers	**Appendix Table A2-1**

*BspQI restriction endonuclease sites (GCTCTTC) are single underlined; BsaI restriction endonuclease sites (GGTCTC) are double underlined; StuI restriction endonuclease site (AGGCCT) is denoted underlined bold; initiation and termination codons within oligonucleotides are bold

[+] Only the portion of the DNA sequence relevant to this study is shown

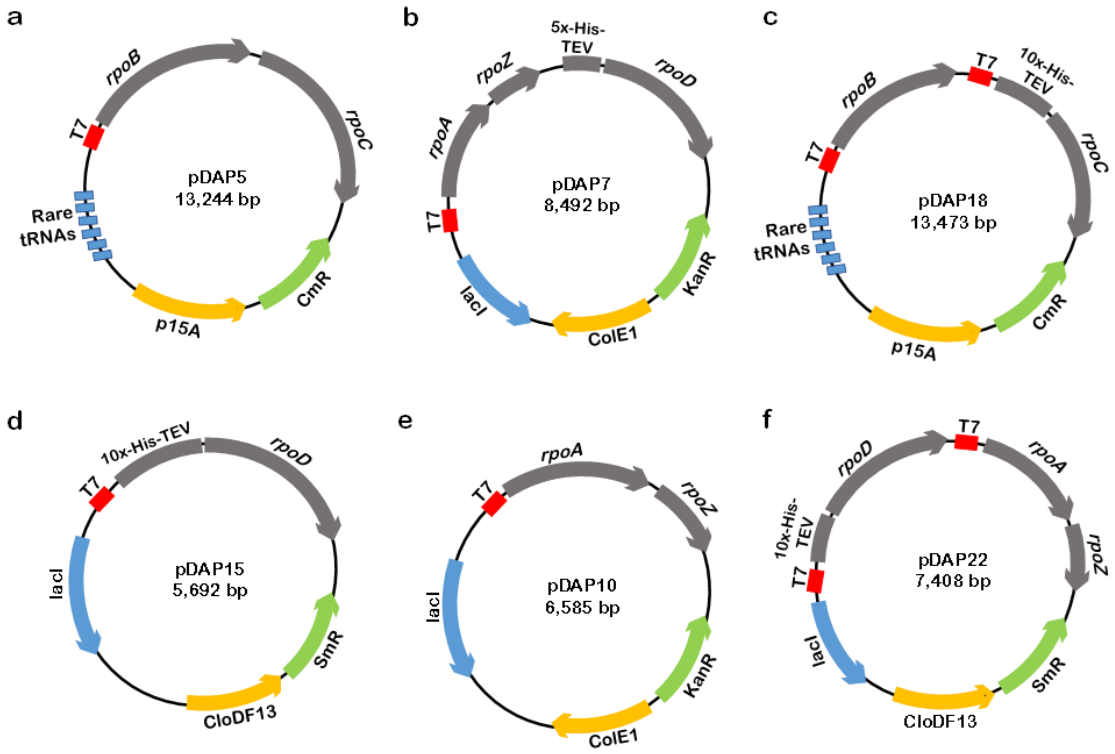


Figure 2-1. Vectors used for RNAP core and holoenzyme expression. **(a)** *rpoB* and *rpoC* were cloned into a modified pRARE2 vector using engineered BsaI sites. **(b)** *rpoA* and *rpoZ* were cloned into a modified pET28b(+) vector using the type IIS enzyme BsaI. *rpoD* was cloned behind *rpoZ* using by doing a cut/ligate with StuI. **(c)** A gblock fragment containing a T7 promoter sequence and a 10x-His tag with a TEV protease cleavage site was introduced upstream of *rpoC* **(d)** *rpoD* containing a N-terminal 10x-His tag and TEV protease cleavage site was cloned into a PmlI-modified pCDFDuet-1 vector. **(e)** Moving *rpoD* from pDAP7 led to the creation of pDAP10 which contained *rpoA* and *rpoZ* **(f)** *rpoA* and *rpoZ* were moved from pDAP10 and cloned into pDAP15 using NEB's HiFi assembly cloning kit. All genes were under expression of the T7 RNA polymerase. pDAP10 and pDAP18 were co-expressed to produce RNAP core enzyme whereas pDAP22 and pDAP18 were co-expressed to produce RNAP holoenzyme. RpoD (σ^{70}) produced from pDAP15 was used together with the core enzyme for *in vitro* RNAP holoenzyme assembly.

RpoB and *rpoC* are polycistronic in *P. aeruginosa* PA01, however, there is a single BspQI site in *rpoB*. To insert the *rpoB/rpoC* genes, a three-piece BspQI cut/ligation was performed with the partial *rpoB* PCR amplicon (PAO1_rpoB_partial-F/–R primers), the remaining *rpoB* PCR amplicon, the full *rpoC* (PAO1_rpoBC-F/-R primers) PCR amplicon, and the T7-pRARE2 intermediate vector to create the final pDAP5. The cloning process replaced the single BspQI site of *rpoB* with a silent mutation. To create plasmid pDAP7 (**Figure 2-1A**), rpoZ-gBlock1 (**Table 2-1**) was synthesized to encode two divergent BsaI restriction sites, a ribosome binding site (RBS), *rpoZ*, a 2nd RBS, an initiating codon followed by a 5x-His tag sequence, and a TEV protease cleavage site with a StuI restriction site (AGG|CCT), all flanked by BspQI sites for insertion into the plasmid pET28b-SapKO-CH as per [30]. The *rpoA* PCR amplicon (generated from PCR primers PAO1_rpoA-F and PAO1_rpo-R) was cut/ligated upstream of *rpoZ* using BsaI I. This positioned *rpoA* directly behind the T7 promoter and RBS of the plasmid with double stop codons (TAATAA) at the 3' end of *rpoA*. Finally, the *rpoD* PCR amplicon (primers PAO1_rpoD-F/-R) was phosphorylated (End-It DNA End-Repair Kit™, Epicentre) and blunt cloned downstream of *rpoZ* into the StuI restriction site to make the final plasmid pDAP7 (**Figure 2-1B**). *RpoA*, *rpoZ* and *rpoD* are in a polycistronic organization in pDAP7. Plasmid pDAP18 was created to improve expression of *rpoC* from plasmid pDAP5 by introducing a T7 promoter and an additional 10 x polyhistidine tag (10x-his tag)/TEV protease site on the 5' position of *rpoC*. A PCR amplicon containing a 10x-his tag with a TEV protease site was generated by PCR amplification of Gblock2 with primers T7TEV_from_Gblock2-F/-R. A second PCR amplicon with pDAP5 as template DNA was generated using primers T7TEV-pDAP5-F/-R, which divergently flank the intergenic region between *rpoB* and *rpoC*. Plasmid pDAP18 was then generated by blunt ligation of the two amplicons after 5' phosphorylation of the insert (**Table 2-1**, **Figure 2-1C**). To allow the separate expression of the core enzyme, *rpoD* was

moved from pDAP7 and introduced into the pCDFDuet-1 vector to create plasmid pDAP15 (**Table 2-1**, **Figure 2-1D**). To make plasmid pDAP15, an intermediate plasmid, pDAP11, was created. To create plasmid pDAP11, a PmlI restriction site was introduced into pCDFDuet-1 by PCR amplification of the vector with primers, pCDFDuet_PmlI-F/-R, followed by digestion at 37 °C. The PCR amplicon of the 5x-His-TEV-*rpoD* region of pDAP7 (primers rpoD_from_pDAP7-F/-R) was blunt ligated into the *Pml*I restriction site introduced downstream of the pDAP11 6x-His tag. The resulting plasmid, pDAP15, contained *rpoD* with an N-terminal 11x-His tag and TEV protease cleavage site (**Table 2-1**, **Figure 2-1D**). Finally, the region containing *rpoA* and *rpoZ* was PCR amplified from pDAP7 (primers rpoAZ_from_pDAP7-F-/-R) and cloned into BsaI sites of plasmid pET28b-SapKO-BsaI to generate pDAP10 (**Table 2-1, Figure 2-1E**).The three plasmids (pDAP10, pDAP15, and pDAP18) allowed either the core RNAP enzyme (pDAP10 and pDAP18) or holoenzyme (all three) to be expressed and purified after co-transformation by electroporation into BL21(DE3) cells (New England Biolabs) and selection on all relevant antibiotics. A significant amount of *E. coli* α subunit contamination was apparent in the holoenzyme preparations (confirmed by mass spectrometry; data shown in **Appendix**). To address this contamination issue, *rpoA* and *rpoZ* were transferred into the pCDFDuet-1-based pDAP15 plasmid that contained *rpoD*. The new plasmid, pDAP22 (**Table 2-1, Figure 2-1F**), was created using NEBuilder HiFi DNA Assembly (New England Biolabs) combining the PCR amplicon of a region containing the T7 promoter, *rpoA*, *rpoZ*, and the T7 terminator using pDAP10 as the template DNA (primers rpoAZ-HiFi-F/-R) with an amplicon of pDAP15 (primers pDAP15-rpoD-HiFi-F/-R).

All plasmids were sequenced fully in both directions (GeneWiz and ACGT Inc) using standard T7 promoter and terminator primers and the sequencing primers (for *rpoB* and *rpoC*) listed **in Appendix Table A2-1**. Sequences were identical to the genomic sequences reported for *Pseudomonas*

aeruginosa PA01 (GenBank accession AE004091.2) with the exception that two silent mutations were introduced into two adjacent codons in the BspQI site of *rpoB* (GAAGAGCAG mutated to GAGGAACAG).

Protein Expression and Purification

Production of TEV protease, RNAP holoenzyme and core, and σ^{70}. TEV protease, used to remove the N-terminal polyhistidine purification tags on RpoC and RpoD, was prepared in house as described in [31]. For production of the recombinant RNAP core and holoenzyme complexes, either pDAP5 and pDAP7, pDAP22 and pDAP18 (to express the holoenzyme), or pDAP18 and pDAP10 (to express the core enzyme) were co-transformed into electrocompetent *E. coli* BL21(DE3) cells (New England Biolabs) with selection on appropriate antibiotic Luria-Bertani agar plates (pDAP22 on 50 µg/mL streptomycin/spectinomycin; pDAP10 on 50 µg/mL kanamycin; pDAP18 on 33 µg/mL chloramphenicol). Plasmid pDAP15 (selected on streptomycin-spectinomycin plates) was utilized for the expression of σ^{70} (RpoD) needed for *in vitro* RNAP holoenzyme assembly with the core enzyme. Approximately 10 colonies were picked from each plate and inoculated into 10 mL LB broth containing the relevant antibiotics. The starter culture was incubated with shaking at 37 °C for about 5-7 hours (OD_{600} = 0.5 – 0.8). For overproduction of the RNAP core and holoenzyme complexes, the autoinduction method as described by Studier [32] was used with modifications. The 10 mL starter culture was inoculated into 500 mL PA-5052 autoinduction media supplemented with 1% tryptone and 2% yeast extract and containing appropriate antibiotics (100 µg/mL for kanamycin). The culture was incubated at 25 °C with shaking (300 RPM) until saturation (approximately 30 hours; OD_{600} ~10). The cells were then harvested by centrifugation at 6000 × *g* for 10 min, and the pellet

was resuspended in 40 mL of lysis buffer (50 mM tris, 500 mM NaCl, 5% v/v glycerol, 10 mM 2-mercaptoethanol, pH 8.0). A protease inhibitor cocktail (EDTA-free) mix from Bimake (1:100 v/v) and PMSF in ethanol (1 mM final) were added to the extract, after which the cells were lysed by an ice-chilled French press (Thermo Electron) at 16,000 psi. The crude cell-free extract was prepared by high speed centrifugation (60,000 × *g*) of the lysate in a Beckman Avanti^{TM} J-25I centrifuge for 30 min at 4 °C.

Purification of RNAP from the cell-free supernatant followed the purification scheme outlined in **Figure 2-2,** which also details the buffers used. All HPLC purification was done using an ÄKTA Purifier system (GE Biosciences) at room temperature. First, an Ni-NTA purification was performed using a 5 mL HisTrap HP column (GE Healthcare Life Sciences) pre-equilibrated with 5 column volumes (CV) of buffer A. After loading the extract on the column, unbound proteins were washed with 10 CV of the binding buffer. A linear gradient of 0% to 100% buffer B was applied over 40 CV at 3 mL/min and peak fractions were analyzed by a 10% SDS-PAGE gel electrophoresis (EZ-Run Protein Gel Solution, Fisher Scientific) with Coomassie brilliant blue G250 staining [33]. Fractions containing RNAP subunits were pooled together and dialyzed into 4L of buffer C for 4 hours. Following dialysis, heparin affinity chromatography was done using a 5 mL HiTrap Heparin HP column (GE Healthcare Life Sciences). The column was pre-equilibrated with 5 CV of buffer D (**Figure 2-2**). After loading of the dialyzed protein on to the column, a gradient of 0 – 100% of buffer E (**Figure 2-2**) was applied over 40 CV at 3 mL/min. Protein fractions were analyzed by SDS-PAGE as before. The relevant fractions were then pooled together for TEV cleavage and dialyzed into buffer F.

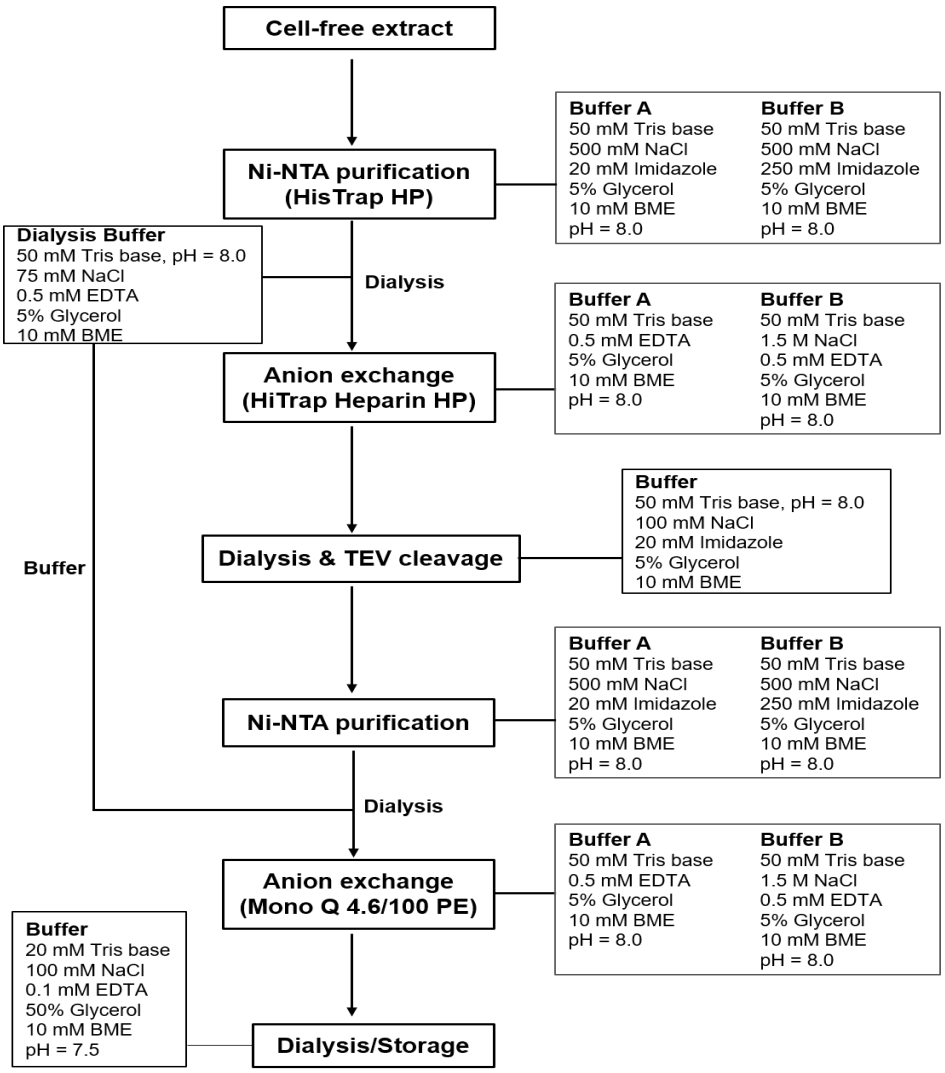


Figure 2-2. Scheme used for *P. aeruginosa PA01* RNAP purification from *E. coli*. After obtaining the cell-free extract, four chromatographic purification steps were performed: Ni-NTA, heparin affinity, TEV cleavage-coupled Ni-NTA, and anion exchange chromatography. The specific column type used for each of these is shown in parenthesis. Using this multistep purification approach, highly pure RNAP free of the 10x-His tags on RpoD and RpoC subunits was obtained. The HPLC and dialysis buffers used for the different purification steps are shown on the left and right of the scheme (BME represents 2-mercaptoethanol). The final RNAP sample was dialyzed in a pH 7.5 buffer containing 50% glycerol and then stored at -20°C until use.

To remove the N-terminal polyhistidine tags via TEV protease, an amount of TEV protease (made in-house) was added to achieve a 1:10 TEV: protein mass ratio. Protein concentration was calculated prior to this using the Bradford Assay (Bio-Rad). The TEV cleavage was left to proceed overnight at 4 °C. To separate the RNAP from the His tag fragment, another Ni-NTA chromatography step was done. After column pre-equilibration as described above, the sample was loaded on the column and the flow-through fractions which contained the protein were saved. Confirmation of the presence of the protein in the flow-through fractions was done by SDS-PAGE gel analysis, and the relevant fractions were pooled and dialyzed in to buffer D for the final purification step. This step involved the use of a 4.6/100 PE MonoQ column (GE Healthcare Life Sciences). A 0% to 20% linear gradient (buffer D to buffer E) was applied over 5 CV, this was followed by second gradient segment of 20% to 30% applied over 20 CV, and then the final gradient step was from 30% to 100% applied over 5 CV. The peak fractions were collected and analyzed by SDS-PAGE gels as before. The relevant fractions were pooled and dialyzed into the final storage buffer, buffer G, which contained 50% glycerol. The protein was stored at -20 °C until use. Recombinant *P. aeruginosa* sigma 70 was prepared as above for the RNAP except that the heparin chromatography step was omitted. Protein purity was estimated for the different purification stages by densitometry analysis of SDS-PAGE gels using NIH's ImageJ software [34].

Mass Spectrometry Analysis

20 µg of purified RNAP was electrophoresed on a 10% SDS-PAGE gel (EZ-Run Protein Gel Solution, Fisher Scientific), and the gel was stained with Coomassie brilliant blue G250. Bands corresponding to the subunits (α, β, β', σ, ω) were excised from the gel and then completely de-

stained twice with a solution containing 50 mM ammonium bicarbonate in 50% methanol for 20 min at room temperature with rocking. The gel bands were then dehydrated in a solution containing 75% acetonitrile for 20 min before transferring to fresh tubes where they were dried at 40 °C. The dried gels were incubated in a solution containing 20 μg/mL trypsin and 20 mM ammonium bicarbonate at 37 °C for 2 hours. Two extractions were then done in a solution containing 50% acetonitrile/0.1% TFA for 20 min and the combined extract solutions were dried. The proteins were identified by peptide mass fingerprinting analysis at the UGA Proteomics and Mass Spectrometry (PAMS) facility using a MALDI-MS (Bruker Autoflex TOF mass spectrometer). The Mascot database search program in conjunction with Proteome Discoverer™ Software (ThermoFisher Scientific) was used to identify the protein subunits.

In vitro Transcription Assays

Each *in vitro* transcription assay had a 20 μL reaction volume containing the transcription reaction buffer, which was adapted from New England Biolabs' *E. coli* RNAP transcription reaction buffer (final condition, 40 mM tris, 150 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 0.01% Triton X-100, pH 7.5), as well as 10 U of RNase Inhibitor Murine (New England Biolabs), 10 μg/mL BSA, 1.25 mM of each ribonucleotide trisphosphate (New England Biolabs), RNAP, and T7 Phage DNA (Boca Scientific) as template DNA.

To measure the activity of RNAP, a modification of the procedure described by [35] was used. First, to determine the lag time for RNAP, a 200 μL reaction mixture was prepared on ice that contained 25 nM (1.25 μg) of T7 DNA, 10 nM RNAP, and the rest of the transcription reaction components as listed above, with the ribonucleotide triphosphate mix added last. Immediately following the addition

of the rNTPs, a 20 μL aliquot was taken and immediately transferred into a PCR tube incubated at 75°C on a Thermal Cycler (Eppendorf Master Cycler Personal) to heat kill the RNAP and prevent any transcriptional activity. The incubation time for heat killing the RNAP was 10 min, and this sample represented the time zero time point. The remaining reaction mixture was incubated at 37°C and time points were taken at 10 sec, 30 sec, 1 min, 1.5 min, 2 min, 2.5 min, 3 min, 4 min, 5 min, 6 min, 7 min, 8 min, and 9 min. At each time point, a 20 μL aliquot was removed, and the RNAP was immediately heat killed for 10 min as described above. The samples were then cooled to room temperature and 10 U of RNase-free DNase I (New England Biolabs) was added together with the DNase I reaction buffer (to 1 x final) and the reaction volume was adjusted to 50 μL with nuclease-free water. DNase I digestion was done for 45 min at 37°C. To concentrate the RNA transcripts made and eliminate degraded DNA fragments which could interfere with RNA quantification, Microcon-100 concentration columns (Millipore Sigma) were used. For each 50 μL reaction mixture, 250 μL of TE buffer (20 mM tris, 1 mM EDTA, pH 7.5) was added before transferring the sample to the columns. The columns were then spun at 500 ×g in a microcentrifuge (Eppendorf Centrifuge 5415 C) for 12 minutes at room temperature. To facilitate sample recovery, 20 μL of TE buffer was added into each column before the columns were inverted into new RNase-free vials and sample recovered by 5 minutes of centrifugation at 500 ×g. The 20 μL filtered and concentrated RNA samples were then added to 80 μL of TE buffer in an opaque 96-well microtiter plate (Costar), and then 100 μL of Quant-iT™ RiboGreen® RNA Reagent (ThermoFisher Scientific) previously diluted 200-fold in TE buffer (final RiboGreen concentration, 150 nM) was added. The samples were incubated at room temperature for 5 minutes protected from light, and then RiboGreen fluorescence was determined with SPECTRAmax M2e microplate reader (Molecular Devices) using machine default settings for RiboGreen. The fluorescence of the time zero sample represented the background fluorescence that

was subtracted from each fluorescence measurement to correct for background fluorescence. Quantification of total RNA synthesized at each time was done using standard curves generated by RNA standards contained in the Quant-iT™ RiboGreen® RNA Reagent and Kit (ThermoFisher Scientific).

To determine other RNAP kinetic parameters (elongation rate, fraction of active enzyme, termination efficiency), the procedure was followed as above for the determination of the lag time, but with some modifications. First, the time zero sample was taken as described and the remaining reaction mixture was also incubated at 37°C as described; however, after 2 min of incubation, heparin (0.1 mg/mL) was added to prevent RNAP re-initiation and ensure that only one round of transcription occurred. Second, samples were taken over a broader time range (2.5, 3, 4, 5, 10, 15, 20, and 25 min), and each was heat killed and treated as described earlier. The experiment was repeated with commercial *E. coli* RNAP (New England Biolabs) for comparison. The assay was repeated with *in vitro*-assembled RNAP. To reconstitute RNAP, RNAP core enzyme was added to purified σ70 protein in a 50 μL reaction mixture containing transcription reaction buffer to obtain 1 μM RNAP core and 5 μM σ70. The mixture was then diluted 5-fold with transcription reaction buffer to reduce the glycerol levels before incubation at room temperature for 60 min. A volume of the assembled mixture that corresponds to a final RNAP concentration of 10 nM in the transcription assay was used and calculations for the determination of the various kinetic parameters also followed the approach proposed by [35].

2.4 RESULTS AND DISCUSSION

Construction of plasmids for RNAP core and holoenzyme production and expression analysis

We created several vectors with the *Pseudomonas aeruginosa* PA01 *rpo* genes arranged, in some cases as polycistronic transcripts, behind T7 promoters for high-level expression of the different RNAP subunits in *E. coli*. N-terminal polyhistidine tags were encoded in the *rpoC* and *rpoD* genes in the plasmids with intervening TEV protease sites to allow removal of the purification tags (leaving N-terminal glycines). Initially, our strategy involved using two compatible plasmids. The *rpoA*, *rpoZ*, and *rpoD* genes with a removable 5xHis-tag/TEV protease site on RpoD were cloned into a pET28-based plasmid (pET28b-SapKO-CH) and the *rpoB* and *rpoC* genes were cloned into a pRARE2 vector modified with a T7 promoter (**Table 2-1**). The plasmid pET28b-SapKO-CH was chosen because it works well for T7-based expression in the laboratory and the dual BspQI sites make cloning straightforward [30]. Recombinant protein yields from this plasmid are routinely high using autoinduction medium; however, amounts of plasmid isolated from cells are routinely low, consistent with reports that the pET28b-derived vector has a relatively low copy number [36-38]. The plasmid pRARE2 was chosen for use as it has an origin of replication and antibiotic selection compatible with the pET28b vector, and pRARE2 encodes rare *E. coli* tRNAs (**Appendix Table A2-2** contains a list of rare tRNAs and their abundance in the *P. aeruginosa* PA01 *rpo* genes). Unfortunately, the protein yields were low when the proteins were co-expressed from these two plasmids (**Appendix Table A2-3**). To address the low expression issue and simplify purification further, we decoupled *rpoD* from *rpoA* and *rpoZ* expression by moving *rpoD* into a different plasmid, pCDFDuet-1 with a longer polyhistidine tag (10x-His) and introduced a TEV protease site

into *rpoC* (**Figure 2-1A-C**). The expression levels of the different RNAP subunits are known to follow the order *rpoZ* > *rpoA* > *rpoB* > *rpoC* [39]; therefore, we thought that tagging *rpoC* will ensure only the fully assembled core/holoenzyme is purified in high yields. The yield improved substantially following the tagging of *rpoC* and shift of the *rpoD*- supporting the decision to rebuild the vectors. Since all three plasmids (pDAP18, pDAP10, pDAP15) were compatible in *E. coli*, we did not encounter any expression problems with this strategy. However, SDS PAGE gels of purified RNAP samples identified a co-purified species with a lower molecular weight than RpoA (**Appendix Figure A2-1**). Mass spectrometry analysis revealed that the lower MW band was *E. coli* RpoA. *E. coli* RNAP subunit contamination of heterologously overexpressed RNAPs is a common concern. We hypothesized that this *E. coli* contamination could be partially attributed to the different copy numbers of the three plasmids used for expressing the five subunits. Of the three plasmids, the pET28(b+)-based pDAP10 carrying *rpoA* and *rpoZ* had the lowest copy number (inferred based on plasmid yields) and hence in retrospect was not the ideal plasmid for expression of the RpoA subunit, which was responsible for assembly of the other subunits. More significantly, 2 moles of RpoA are required for every 1 mole of RNAP holoenzyme enzyme assembled and therefore the inadequacy in the amount of RpoA produced compared to the other subunits might have led to sequestration of compensatory *E. coli* RpoA into the PA01 RNAP. In their study, [40] compared the purity and yield levels of RNAP subunits expressed on different polycistronic plasmid combinations and found that the best yield was obtained by cloning *rpoA* and *rpoZ* into the higher copy number pACYCDuet vector (which is fundamentally the pRARE2 without rare codons), and cloning *rpoB* and *rpoC* into the pETDuet vector. We therefore tested our hypothesis that RpoA expression could be improved by moving both *rpoA* and *rpoZ* into the pCDFDUET-based pDAP15 vector containing *rpoD*, creating pDAP22 (**Figure 2-1D**). Co-expression of this higher copy number plasmid with the

pRARE2-based pDAP18 containing *rpoB* and *rpoC* eliminated any measurable contaminating *E. coli* RpoA and led to the highest RNAP yields. It is important to point out however that the MonoQ anion exchange column used in the final purification step was able to resolve RNAP with contaminating *E. coli* subunits from RNAP containing homogenous *P. aeruginosa* subunits. As a result, our triple plasmid expression system can still be utilized if final RNAP yield is not of concern. Hence for production of the RNAP core enzyme, the pET28-based pDAP10 containing *rpoA* and *rpoZ*, and the pRARE2-based pDAP18 containing *rpoB* and *rpoC* were used. σ^{70} (RpoD) expressed from pDAP15 was used together with the core enzyme for *in vitro* RNAP holoenzyme assembly.

Expression and purification of RNAP core and holoenzymes

Due to the high complexity and multi-subunit nature of RNA polymerases, the expression and purification approach used for obtaining highly pure and active RNAP enzyme preparations is critical. Such an approach must take into consideration factors such as plasmid compatibility, protein stability following expression, correct assembly to yield active protein, and purification steps. To address these factors, we implemented a room temperature autoinduction protein production approach coupled with a multistep purification scheme. To our knowledge, this is the first report to describe the use of an autoinduction method for RNAP expression. Proteins produced in this manner have been shown to be better behaved and more stable compared to those expressed by IPTG induction [41, 42]. This is especially essential for an enzyme as large and complex as RNAP. Supplementation of the standard PA-5052 autoinduction media with 1% tryptone and 2% yeast extract ensured both a high cell mass (20 – 25 g/L) and corresponding high protein (20 – 30 mg/L) yield (**Table 2-3**).

Table 2-3. Purification profile of RNAP at the different chromatographic stages

Sample	Total protein (mg)*	Purity (%)**
Cell free extract	3412.32	13.3
Ni-NTA	28.96	80.2
Heparin affinity	25.5	83.3
TEV cleavage Ni-NTA	12.45	86.8
MonoQ	10.8	96.5

*The data shown here represent the statistics from a 500 mL culture.
** Purity was determined by densitometry from SDS-PAGE gels using NIH's ImageJ program [30].

The induction process was done at room temperature and typically required growing the cultures for about 30 hours at 25 °C. We found that the induction time could be significantly reduced to 15 – 20 hours by initially incubating the culture with shaking at 37 °C until OD_{600} = 0.6 – 1.0, before transferring to 25 °C.

Following protein expression, our purification approach involved four different chromatographic steps (**Figure 2-2**), with preliminary experiments done to identify the ideal buffer and separating conditions for each step. Since RNAP is a DNA binding protein, many approaches described in the literature for purification of the *E. coli*, *T. aquatics*, and *B. subtilis* RNAPs employ polyethyleneimine (Polymin-P) precipitation to remove potential DNA contamination. We instead relied on the final MonoQ anion exchange chromatography for this purpose, but also employed an intermediate heparin affinity chromatography column and the combination of metal-chelate chromatography, TEV protease treatment to remove the polyhistidine tags that bind the proteins to

metal chelate media, and a 2nd metal chelate chromatography step with the complexes flow directly through without retention. The four-step chromatographic purification process ensured the production of highly pure RNAP core and holoenzyme preparations (**Figure 2-3**).

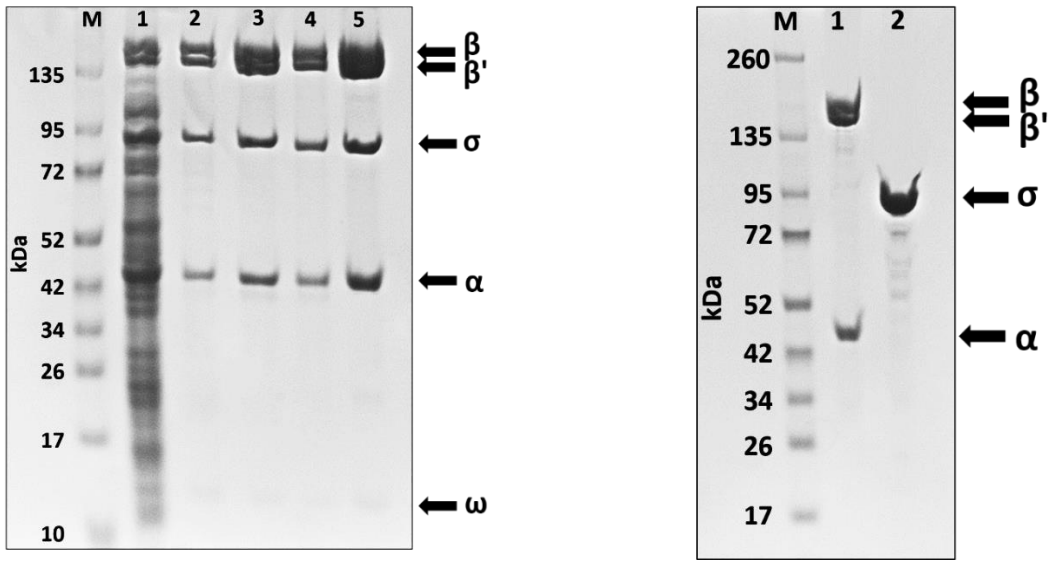


Figure 2-3. Purification of *P. aeruginosa PA01* RNAP core and holoenzymes. **(a)** RNAP holoenzyme purification from BL21(DE3) *E. coli* cells. Lane 1: cell-free extract after cell lysis and high-speed centrifugation. Lane 2: Ni-NTA fractions following elution with 125 mM imidazole. Lane 3: heparin affinity chromatography. Lane 4: TEV cleavage-coupled Ni-NTA purification. Lane 5: final RNAP sample after anion-exchange purification on the MonoQ column. (**b**) Purification of the core enzyme and σ^{70} followed the same chromatographic steps as the holoenzyme. Only the final core enzyme (Lane 1) and σ^{70} (Lane 2) samples are shown. The MW marker is shown as Lane M in

Mass spectrometry analysis confirmed that our RNAP preparation contained homogenous *P. aeruginosa* subunits with no detectable *E. coli* RNAP subunit contamination (**Appendix A**). We also found that removal of the 10xHis tag on the RpoC and RpoD subunits was critical to activity of the enzyme (data not shown). Lastly, *in vitro* transcription assays that lacked DNA templates showed no RNA synthesis (data not shown). This is evidence that there is little contaminating DNA in the final RNAP product.

Transcription assays and activity of RNAP

We used a standard RNAP transcription assay using T7 Phage DNA as template to assess the activity of the *P. aeruginosa* RNAP and compared it with that of commercial *E. coli* RNAP holoenzyme. The assay procedure was a modification of the method described by [35] and all calculations and determinations were done in line with the stipulations in the paper. Synthesis of RNA chains from T7 phage DNA *in vitro* has been shown to occur exclusively at three promoter sites, Al, A2, and A3, located near the 5' end of the linear DNA sequence [43, 44]. Additionally, located at 7720 bp of the DNA is a strong termination signal at which ~80% of RNAP molecules terminate, leading to the generation of a 7133 nt RNA transcript from transcription initiation at the A promoters [45]. As a result, the predicted kinetics of T7 RNA synthesis by RNAP follows a four-phase scheme as outlined in **Figure 2-4A.** Phase I involves the initial stages of initiation where promoter location, melting and the synthesis of short abortive products occur, the length of which represents the lag time of the enzyme. Phase II represents a linear phase where each RNAP molecule is assumed to be actively involved in constant RNA chain elongation.

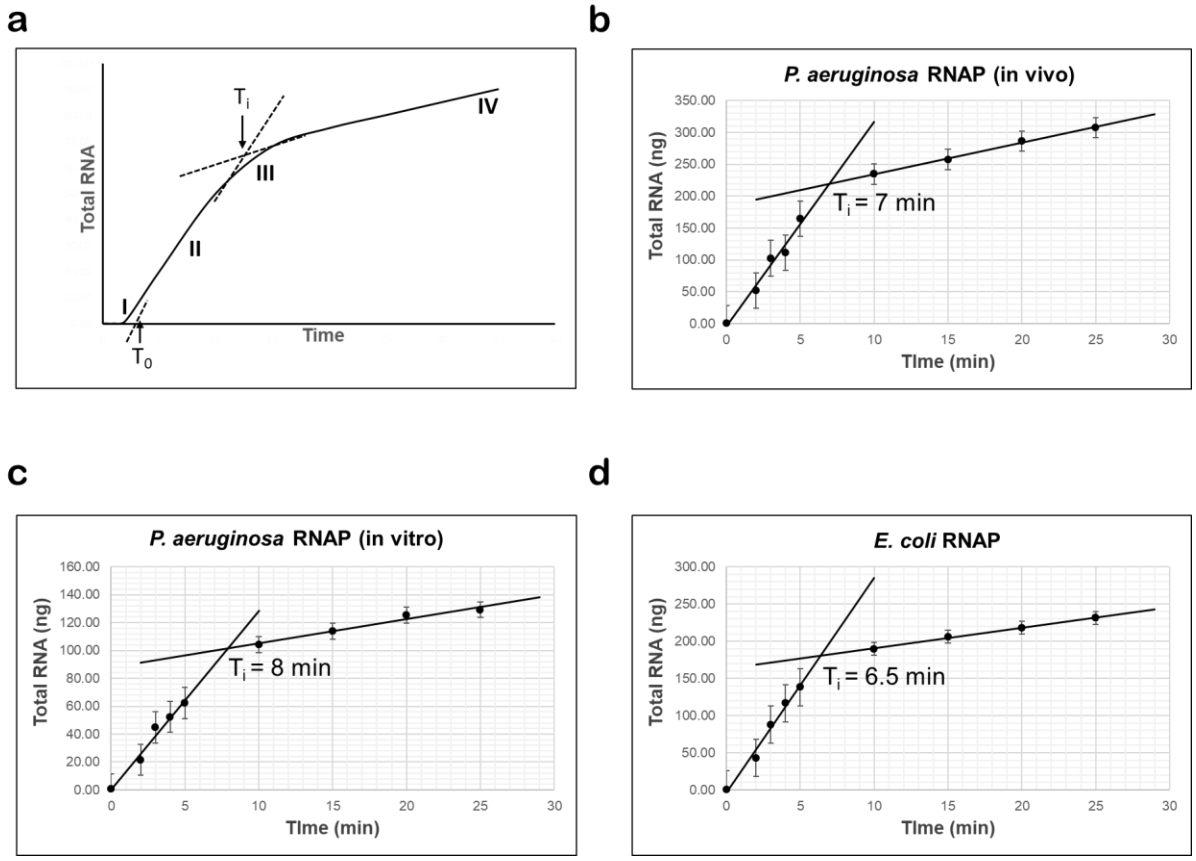


Figure 2-4. Kinetics of RNAP activity. **(a)** Predicted kinetics of RNA synthesis from T7 Phage DNA as outlined in [35]. Phase I represents the initial stages of transcription initiation where promoter recognition, melting, and synthesis of short abortive transcripts occur. In Phase II, constant rNTP incorporation occurs and RNA is synthesized in a linear fashion until RNAP encounters the strong terminator at 7720 bp. Addition of heparin during the early phases of stage II precludes RNAP molecules that terminate at the terminator from re-initiation and Phase III therefore represents the small inflection period during which majority of the RNAP molecules terminate, resulting in a significant reduction in RNA synthesis. During Phase IV, the remaining RNAP molecules that failed to terminate continue another linear phase of RNA synthesis that could span the entire length of the T7 genome. The lag time is indicated by T_0 whereas the termination time (T_i) is determined as the intersection of the Phase II and IV curves. **(b)** Kinetics

of the in *P. aeruginosa* RNAP (*in vivo*-assembled) followed the expected kinetics described above and the T_i was determined as 7 min. **(c)** The *in vitro*-assembled *P. aeruginosa* RNAP was associated with a slightly higher T_i (8 min) which translated into a slightly lower elongation rate compared to the *in vivo*-assembled one. **(d)** The kinetics of the *P. aeruginosa* RNAP was closely comparable to that of commercial *E. coli* RNAP (T_i = 6.5). Each graph represents data from duplicate experiments using the same batch of purified (or purchased) RNAP.

The addition of heparin during the early stage of Phase II ensures the inhibition of chain re-initiation by RNAP molecules that synthesize abortive transcripts. Particularly, the heparin also prevents re-initiation of the ~80% of RNAP molecules that encounter the terminator at 7720 bp, leading to a significant reduction in the rate of rNTP incorporation (Phase III). The remaining ~20% of RNAP molecules are involved in a second linear phase of chain elongation (Phase IV) that could extend to the length of the entire T7 DNA. Two time points on the graph are important as shown in **Figure 2-4A**: the lag time (T_0) and the time to reach the 7720 bp terminator (T_i), determined as the intersection of the Phase II and IV curves. As shown in **Figure 2-4B**, RNA synthesis from T7 DNA by the *P. aeruginosa* RNAP (*in vivo*-assembled) followed this predicted kinetics. The graph is valuable for calculating important kinetic parameters associated with RNAP activity, particularly the chain elongation rate. Because RNAP molecules that terminate at the 7720 bp terminator would have synthesized an RNA product of average length 7133, this can be divided by $T_i - T_0$ to yield the chain elongation rate. Using the T_i determined for the *P. aeruginosa* RNAP (7 min) and a lag time of 30 seconds, the chain elongation rate was determined to be ~18 nucleotides per second. To compare,

the assay was repeated with commercial *E. coli* RNAP (**Figure 2-4D**). The results indicated that the two holoenzymes had comparable activities, with the *E. coli* RNAP being associated with an elongation rate of 19 nucleotides per second. In their determination, [35] found a chain growth of 17-18 nt/s for the *E. coli* RNAP as well as a rate within a range of 15-19 nt/s for five other RNAPs tested. As shown in **Table 2-4**, the amount of RNA product produced per second per mg of RNAP was also closely comparable for the two enzymes. Two other kinetic parameters that can be calculated are the fraction of active polymerase in a preparation as well as the termination efficiency. The termination efficiency is calculated as the 1 – (slope of IV/slope of II), and the *E. coli* RNAP was associated with a slightly better termination at the 7720-terminator compared to the *P. aeruginosa* RNAP (**Table 2-4**).

Table 2-4. Activity profile of the *P. aeruginosa* RNAP in comparison with the *E. coli* RNAP *

RNAP sample	Elongation rate (nt/s)	Activity (ng RNA/min)	Termination Efficiency (%)
E. coli	19 ± 0.24	26.7 ± 2.97	90.4 ± 0.03
P. aeruginosa (in vivo)	18 ± 0.20	31.2 ± 1.15	84.1 ± 0.58
P. aeruginosa (*in vitro*)	16 ± 0.15	14.1 ± 1.77	85.6 ± 1.37

*Calculations are based on duplicate experiments using the same batch of purified (or purchased) RNAP.

Comparison of *in vivo*- vs. *in vitro*-assembled RNAP holoenzyme

Previous reports on the production of recombinant *E. coli*, *T. aquaticus*, *B. subtilis*, or *M. tuberculosis* RNAPs included studies where the core enzyme and sigma factor were purified separately, then assembled *in vitro* to make the holoenzyme [21, 24, 27, 29]. This approach is especially desirable if different sigma factors are to be used with their corresponding promoters for transcriptional analysis. Additionally, the approach might also be better at producing a high yield of the assembled holoenzyme compared to the yield if all five subunits were expressed and assembled *in vivo*. Finally, it is relatively less challenging to devise a strategy to clone and express the core enzyme and sigma factor separately than it is to devise one for all five subunits to be expressed and assembled *in vivo*. However, the efficiency of the *in vitro* assembly could pose a major challenge. For instance, it is conceivable that certain nuances during *in vivo* assembly would not be accounted for during *in vitro* assembly, leading to a less active *in vitro*-assembled enzyme. This was our observation when we compared the RNAP activity of the enzyme assembled *in vivo* vs. *in vitro* (**Table 2-4, Figure 2-4C, Figure 2-5**). The determined elongation rate for the *in vitro*-assembled holoenzyme was 16 nt/s, a number slightly lower than that for the *in vivo* assembled one (18 nt/s). Of more significance, the total RNA made per minute from the *in vivo*-assembled enzyme was 2 times higher than that of the reconstituted holoenzyme. This suggests that either the assembly conditions we used were not efficient (*i.e.* a large proportion of core remains unassociated with sigma) or possibly, the purified core or sigma factor is not stable in the storage conditions (in 50% glycerol at -20 °C,) and only a fraction reassembles as a functional holoenzyme.

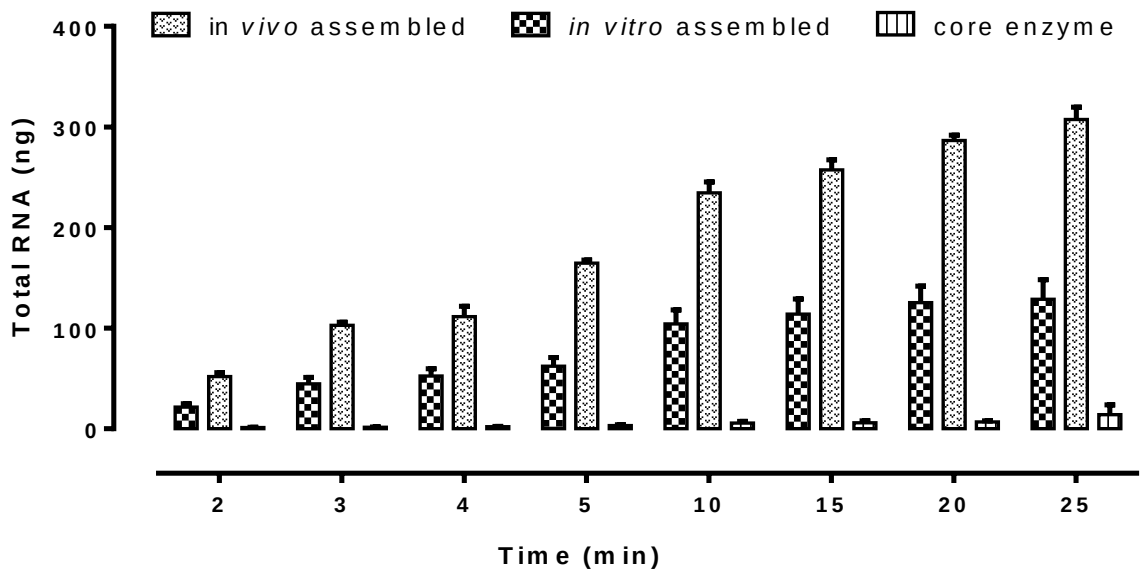


Figure 2-5. Activity of the *in vitro* assembled RNAP holoenzyme (5-fold excess of σ^{70}) in comparison with the *in vivo* assembled RNAP. The *in vivo*-assembled enzyme had an activity that was about 2 times higher than that of the *in vivo* assembled enzyme. As expected, the core enzyme lacking the sigma factor, which is largely incapable of promoter-specific initiation, had very little activity. Figure is based on data from duplicate experiments using the same batch of RNAP.

2.5 CONCLUSIONS

The large size, high complexity, and multisubunit nature of RNA polymerases present major challenges in producing the enzymes- either recombinantly or by another means. Unsurprisingly, only four bacterial RNAP enzymes have been over-produced and purified. Structural and functional studies of transcription in any specific bacteria require a reproducable approach for producing

adequate amounts of the specific bacterial RNAP. The expression plasmids described here make it relatively straightforward to produce either the core or holoenzyme in abundant quantities and at the high level of purity necessary for structural and biochemical studies. Furthermore, low temperature autoinduction for expression of the proteins [32] means no IPTG is necessary. Our successful production of recombinant *P. aeruginosa* RNA polymerase presents an important breakthrough for studies focused on transcriptional regulation in the pathogenic bacteria.

Acknowledgements

The *P. aeruginosa str.* PAO1 genomic DNA used to isolate the genes was a kind gift from Dr. Balazs Rada (radab@uga.edu). We are also grateful to Nicklaus Galloway for helping with the purification of the TEV protease, as well as David Seth Adams and the Young Scholars Program at the University of Georgia. Finally we are grateful to the UGA Proteomics and Mass Spectrometery (PAMS) Core Facility, particularly, Dennis Phillips (drp@uga.edu) and Chau-Wen Chou (cwchou@uga.edu) for help with the Mass Spectrometry analysis. 5% of the funding for this work was provided by the College of Pharmacy Wilson Pharmacy Fund, University of Georgia. 95% of the research reported in this publication was supported by the United States, National Institutes of Health, award number: AWD00007016. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

2.6 REFERENCES

1] J. Hurwitz, A. Bresler, R.J.B. Diringer, b.r. communications, The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA. 3 (1960) 15-19.

[2] K. Kuznedelov, K. Severinov, Recombinant bacterial RNA polymerase: preparation and applications. Methods 47 (2009) 44-52.

[3] K.S. Murakami, Structural biology of bacterial RNA polymerase. Biomolecules 5 (2015) 848-864.

[4] S.A. Darst, Bacterial RNA polymerase. Current opinion in structural biology 11 (2001) 155-162.

[5] K. Geszvain, R. Landick, The structure of bacterial RNA polymerase, The bacterial chromosome, American Society of Microbiology, 2005, pp. 283-296.

[6] J.D. Helmann, Purification of Bacillus subtilis RNA polymerase and associated factors, Methods in enzymology, Elsevier, 2003, pp. 10-24.

[7] S.A. Kumar, The structure and mechanism of action of bacterial DNA-dependent RNA polymerase. Progress in biophysics and molecular biology 38 (1981) 165-210.

[8] K.S. Murakami, S.A. Darst, Bacterial RNA polymerases: the wholo story. Current opinion in structural biology 13 (2003) 31-39.

[9] S.L. Gellatly, R.E. Hancock, Pseudomonas aeruginosa: new insights into pathogenesis and host defenses. Pathogens and disease 67 (2013) 159-173.

[10] C.K. Stover, X.Q. Pham, A. Erwin, S. Mizoguchi, P. Warrener, M. Hickey, F. Brinkman, W. Hufnagle, D. Kowalik, M. Lagrou, Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. Nature 406 (2000) 959.

[11] G.P. Bodey, R. Bolivar, V. Fainstein, L. Jadeja, Infections caused by Pseudomonas aeruginosa. Reviews of infectious diseases 5 (1983) 279-313.

[12] S. Santajit, N. Indrawattana, Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens. Biomed Res Int 2016 (2016) 2475067.

[13] P.D. Lister, D.J. Wolter, N.D. Hanson, Antibacterial-resistant Pseudomonas aeruginosa: clinical impact and complex regulation of chromosomally encoded resistance mechanisms. Clin Microbiol Rev 22 (2009) 582-610.

[14] K. Poole, Pseudomonas aeruginosa: resistance to the max. Frontiers in microbiology 2 (2011) 65.

[15] J. Yayan, B. Ghebremedhin, K. Rasche, Antibiotic resistance of Pseudomonas aeruginosa in pneumonia at a single university hospital center in Germany over a 10-year period. Plos one 10 (2015) e0139836.

[16] K.S. Murakami, S. Masuda, S.A. Darst, Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. Science 296 (2002) 1280-1284.

[17] P.K. Taylor, A.T. Yeung, R.E. Hancock, Antibiotic resistance in Pseudomonas aeruginosa biofilms: towards the development of novel anti-biofilm therapies. Journal of biotechnology 191 (2014) 121-130.

[18] R.S. Mandal, A. Ta, R. Sinha, N. Theeya, A. Ghosh, M. Tasneem, A. Bhunia, H. Koley, S. Das, Ribavirin suppresses bacterial virulence by targeting LysR-type transcriptional regulators. Scientific reports 6 (2016) 39454.

[19] M. Chatterjee, C. Anju, L. Biswas, V.A. Kumar, C.G. Mohan, R. Biswas, Antibiotic resistance in Pseudomonas aeruginosa and alternative therapeutic options. International Journal of Medical Microbiology 306 (2016) 48-58.

[20] M.M. Saleh, H.A. Abbas, M.M. Askoura, Repositioning secnidazole as a novel virulence factors attenuating agent in Pseudomonas aeruginosa. Microbial pathogenesis 127 (2019) 31-38.

[21] R. Banerjee, P. Rudra, R.K. Prajapati, S. Sengupta, J. Mukhopadhyay, Optimization of recombinant Mycobacterium tuberculosis RNA polymerase expression and purification. Tuberculosis 94 (2014) 397-404.

[22] R. Harshey, T. Ramakrishnan, Purification and properties of DNA-dependent RNA polymerase from Mycobacterium tuberculosis H37RV. Biochimica et Biophysica Acta (BBA)-Nucleic Acids and Protein Synthesis 432 (1976) 49-59.

[23] J.-F. Jacques, S. Rodrigue, R. Brzezinski, L. Gaudreau, A recombinant Mycobacterium tuberculosis in vitro transcription system. FEMS microbiology letters 255 (2006) 140-147.

[24] S. Borukhov, A. Goldfarb, Recombinant Escherichia coli RNA polymerase: purification of individually overexpressed subunits and in vitro assembly. Protein expression and purification 4 (1993) 503-511.

[25] K.S. Murakami, X-ray crystal structure of Escherichia coli RNA polymerase σ70 holoenzyme. Journal of Biological Chemistry 288 (2013) 9126-9134.

[26] K. Kuznedelov, V. Lamour, G. Patikoglou, M. Chlenov, S.A. Darst, K. Severinov, Recombinant Thermus aquaticus RNA polymerase for structural studies. Journal of molecular biology 359 (2006) 110-121.

[27] L. Minakhin, S. Nechaev, E.A. Campbell, K. Severinov, Recombinant Thermus aquaticus RNA polymerase, a new tool for structure-based analysis of transcription. Journal of bacteriology 183 (2001) 71-76.

[28] L.C. Anthony, I. Artsimovitch, V. Svetlov, R. Landick, R.R. Burgess, Rapid Purification of His6-Tagged Bacillus subtilis Core RNA Polymerase. Protein expression and purification 19 (2000) 350-354.

[29] X. Yang, P.J. Lewis, Overproduction and purification of recombinant Bacillus subtilis RNA polymerase. Protein expression and purification 59 (2008) 86-93.

[30] N.R. Galloway, H. Toutkoushian, M. Nune, N. Bose, C. Momany, Rapid cloning for protein crystallography using type IIS restriction enzymes. Crystal Growth & Design 13 (2013) 2833-2839.

[31] P.G. Blommel, K.J. Becker, P. Duvnjak, B.G. Fox, Enhanced bacterial protein expression during auto-induction obtained by alteration of lac repressor dosage and medium composition. Biotechnology progress 23 (2007) 585-598.

[32] F.W. Studier, Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41 (2005) 207-234.

[33] I. Syrový, Z. Hodný, Staining and quantification of proteins separated by polyacrylamide gel electrophoresis. Journal of Chromatography B: Biomedical Sciences and Applications 569 (1991) 175-196.

[34] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. Nature methods 9 (2012) 671.

[35] M.J. Chamberlin, W.C. Nierman, J. Wiggs, N. Neff, A quantitative assay for bacterial RNA polymerases. Journal of Biological Chemistry 254 (1979) 10061-10069.

[36] P. Dersch, H. Fsihi, E. Bremer, Low-copy-number T7 vectors for selective gene expression and efficient protein overproduction in Escherichia coli. FEMS microbiology letters 123 (1994) 19-26.

[37] A.V. Mardanov, T.S. Strakhova, V.A. Smagin, N.V. Ravin, Tightly regulated, high-level expression from controlled copy number vectors based on the replicon of temperate phage N15. Gene 395 (2007) 15-21.

[38] W.N. Wood, K.D. Smith, J.A. Ream, L.K. Lewis, Enhancing yields of low and single copy number plasmid DNAs from Escherichia coli cells. Journal of microbiological methods 133 (2017) 46-51.

[39] V. Svetlov, I. Artsimovitch, Purification of bacterial RNA polymerase: tools and protocols. Methods Mol Biol 1276 (2015) 13-29.

[40] R. Banerjee, P. Rudra, R.K. Prajapati, S. Sengupta, J. Mukhopadhyay, Optimization of recombinant Mycobacterium tuberculosis RNA polymerase expression and purification. Tuberculosis (Edinb) 94 (2014) 397-404.

[41] L. Briand, G. Marcion, A. Kriznik, J.-M. Heydel, Y. Artur, C. Garrido, R. Seigneuric, F. Neiers, A self-inducible heterologous protein expression system in Escherichia coli. Scientific reports 6 (2016) 33037.

[42] F.W. Studier, Stable expression clones and auto-induction for protein production in E. coli, Structural Genomics, Springer, 2014, pp. 17-32.

[43] M.J. Chamberlin, The selectivity of transcription. Annual review of biochemistry 43 (1974) 721-775.

[44] F.W. Studier, J.J. Dunn, Organization and expression of bacteriophage T7 DNA, Cold Spring Harbor symposia on quantitative biology, Cold Spring Harbor Laboratory Press, 1983, pp. 999-1007.

[45] R.C. Williams, M.J. Chamberlin, Electron microscope studies of transient complexes formed between Escherichia coli RNA polymerase holoenzyme and T7 DNA. Proceedings of the National Academy of Sciences 74 (1977) 3740-3744.

Chapter 3

SULFUR ASSIMILATION IN *PSEUDOMONAS AERUGINOSA*: THE ROLE OF FINR, A LYSR-TYPE TRANSCRIPTIONAL REGULATOR

3.1 SYNOPSIS

The LysR-type Transcriptional Regulators (LTTR) are the largest class of transcriptional factors in bacteria, and they mediate diverse cellular processes. LTTRs are abundant in *Pseudomonas aeruginosa*, an important pathogen that is especially problematic because it can survive in diverse environmental conditions. The adaptability of this bacterium, which has been attributed to its regulatory agility, contributes to its high pathogenicity. Therefore, studies of regulatory circuits and LTTRs in *P. aeruginosa* may provide new treatment targets to control this infectious pathogen. The focus of this study was to probe the regulatory role of the FinR, a *P. aeruginosa* LTTR that was recently shown to be involved in oxidative stress regulation through its control of *fprA*, the ferredoxin-NADP+ reductase enzyme. Although FinR has so far only been implicated in oxidative stress regulation in *P. aeruginosa*, this is likely a secondary role. Instead, we suspect that its primary role is to contribute to sulfur assimilation in the pathogen. In the sulfur assimilation pathway where sulfur is incorporated into L-cysteine, FprA can act in the capacity of CysJ, a sulfite reductase whose ortholog is missing from the *P. aeruginosa* genome. This places FinR, which regulates FprA expression, at a crucial regulatory point in the pathway. To assess this previously unreported role of

FinR in *P. aeruginosa*, we probed binding of FinR to different regions of the *finR–fprA* genetic region and found that the LTTR binds tightly to the intergenic region. Bioinformatic analysis identified several conserved LTTR boxes in the intergenic region that were putative FinR binding sites and binding studies revealed that the LTTR likely regulates *fprA* expression through a 'sliding dimer' regulatory mechanism. Using transcription assays, we also showed that sulfite is an inducer for FinR, further validating the involvement of the LTTR in sulfur assimilation in *P. aeruginosa*. These findings establish FinR as an important LTTR in *P. aeruginosa* and identify the protein as a potential target for antibiotic drug discovery.

3.2 INTRODUCTION

The ferredoxin reductase regulator, FinR, is one of the best characterized LTTRs in pseudomonads [1-5]. The FinRs regulate the expression of the ferredoxin-NADP+ reductases encoded by the *fpr* genes [3, 6]. The Fpr proteins are ubiquitous, monomeric, and reversible flavin enzymes that catalyze the reversible redox reaction between NADPH or NADP+ and one-electron carriers such as ferredoxin or flavodoxin [7] . These enzymes play a crucial role in maintaining the NADP+/NADPH redox ratio, hence placing them at critical positions in important cellular processes like siderophore synthesis/regulation, iron acquisition, sulfur assimilation and cysteine biosynthesis, and oxidative stress [1, 2, 8, 9]. The involvement of FprA in some of these roles has been confirmed in *P. putida* [3, 5, 7, 10]; however, the physiological function of the *P. aeruginosa* FprA was previously largely unknown because, unlike the *P. putida fprA*, every attempt by different groups to successfully construct the *fprA* knockout mutant failed [11]. Indeed, [4], who reported the characterization of the *P. aeruginosa* FinR (PA3398) found *fprA* to be an essential gene for the pathogen such that it could

not be successfully knocked-out without fatal consequences for the cells. Although the *P. aeruginosa* FinR was found to be non-essential in the study, its ortholog from a more distantly-related proteobacteria, *Acinetobacter baylyi* ADP1, has been found to be an essential protein [12]. In *P. aeruginosa*, *finR* is expressed divergently from the *fprA* gene, with a 272 bp intergenic DNA between them. This genetic organization is commonly observed for LTTRs that regulate the divergent nearby gene. FinR was found to positively regulate *fprA* expression [4] while undergoing negative autoregulation. Additionally, *finR* mutants were associated with an increased sensitivity to paraquat and this effect was reversed upon increased expression of *fprA*, implicating FinR in the *P. aeruginosa* oxidative stress response. This role for FinR also seemed to be essential for virulence as *in vivo* studies in a *P. aeruginosa* Drosophila host model showed *finR* mutants to be associated with attenuated virulence. Importantly, the study further determined the +1 *fprA* transcription start site and proposed a location for the FinR binding site within the intergenic region.

Even though the role of FinR in *P. aeruginosa* has so far been restricted to the oxidative stress response, it is clear from studies focusing on FinR in other pseudomonads that the transcriptional regulator might also play a significant role in cysteine biosynthesis through the sulfur assimilation pathway. In the pathway (**Figure 3-1**), sulfate and its derivatives are converted to sulfite, which is reduced by CysI/CysJ, both sulfite reductases, to sulfide before subsequent conversion to L-cysteine by the cysteine synthases (CysM/CysK) [1]. The absence of an obvious CysJ ortholog in the *P. aeruginosa* genome [13, 14] means that a different redox protein is necessary for this critical role. We propose that FprA acts in the capacity of CysJ in a ferredoxin-dependent manner. It is therefore conceivable that the oxidative stress regulatory function associated with FinR is only a secondary role triggered by processes dependent on the sulfur synthesis pathway. Oxidative stress and sulfur metabolism are linked through iron-sulfur cluster and L-cysteine/glutathione turnover

within the cells following intracellular oxidative damage [1, 10]. Coupling all these information with the finding that FprA is an essential enzyme in *P. aeruginosa* suggest clearly that the cysteine biosynthesis pathway is a critical process in the pathogen, further establishing FinR, the regulator of FprA expression, as an important transcriptional regulator in the organism.

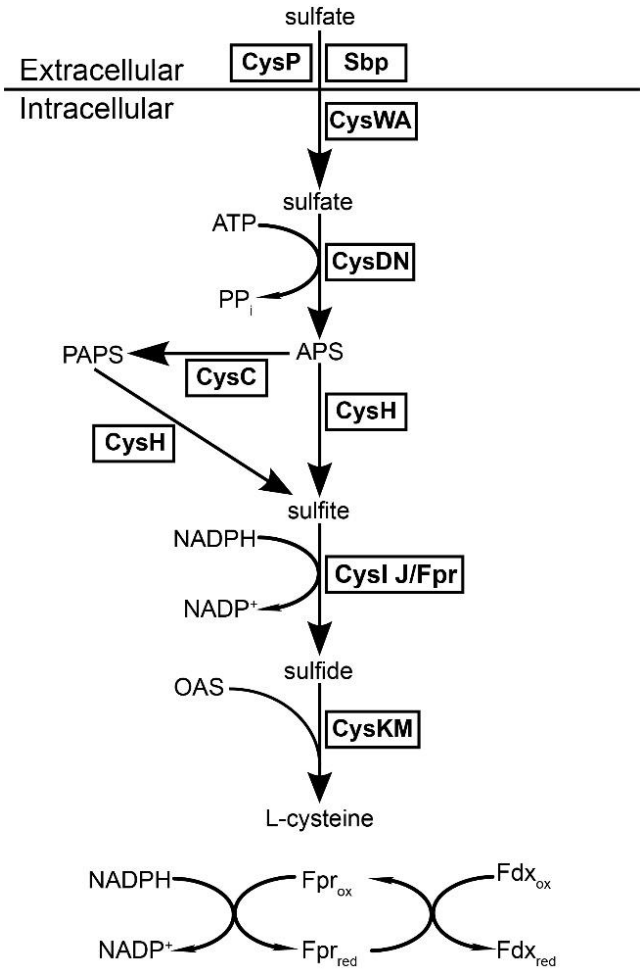


Figure 3-1. The sulfur assimilation pathway in bacteria. Sulfur is incorporated as sulfate and enters the cells through various transporters in the membrane. Ultimately, the sulfate is converted to sulfite, the substrate for the two sulfite reductase enzymes: CysI and the ferredoxin-dependent CysJ. Reduction of sulfite converts it to sulfide, which is subsequently converted to L-cysteine by

cysteine synthase (CysK). Pseudomonads do not contain a gene identified as CysJ [1]. FrpA likely acts in a CysJ capacity.

In this study, we probed the potential involvement of FinR in the *P. aeruginosa* sulfur assimilation pathway and show that sulfite is a potential effector for the protein. We identified various FinR binding sites within the *finR–fprA* intergenic region that supported a 'sliding dimer' regulatory mechanism for FinR. Using electrophoretic mobility shift assays to probe various regions of the intergenic region, we showed that FinR indeed likely relies on this mechanism in its regulation of FprA. Transcription assays that utilized the *P. aeruginosa* RNA polymerase were then used to further validate these findings. Overall the study expands the regulatory role of the *P. aeruginosa* FinR and identifies a potential effector for the transcriptional regulator.

3.3 MATERIALS AND METHODS

Construction of the *Pseudomonas aeruginosa finR* Expression Plasmid

All bacterial strains, plasmids, primers, and DNA used in the study are listed in **Table 3-1.** The creation of the FinR and CysB expression plasmids entailed PCR amplifying the *finR* or *cysB* genes by polymerase chain reaction and insertion into a pET28b-based vector with a C-terminal polyhistidine purification tag. *P. aeruginosa* str. PAO1 genomic DNA was used as template for a

PCR amplification (primers, FPM1 and FPM2 for FinR; PAO1_cysB_BspQI-F and PAO1_cysB_BspQI-R for CysB) using Phusion® Hot Start II High-Fidelity DNA polymerase (Themo Scientific™). The cycling conditions were as follows: denaturation at 95°C for 45 sec, annealing at 50 – 60°C for 30 sec, extension at 72°C for 30 sec, and a final extension at 72°C for 10 min. Following 30 PCR cycles, the reactions were assessed by agarose gel electrophoresis (1% agarose gel with 1X TBE as running buffer). The resulting PCR amplicons were purified using a PCR Cleanup Kit (Zymo Research) and the concentration of the purified DNA was checked on a NanoDrop™ Lite Spectrophotometer (ThermoFisher Scientific). The *finR* PCR products were inserted into pET28b(+).SapKO−CH.BspQI plasmid using a cut/ligation protocol with BspQI (New England Biolabs) as described by [15].

Table 3-1. Bacterial strains, plasmids, and oligonucleotides used in the study

Name	Description	Source
***E. coli* Strains**		
XL1-Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F′ proAB lacIq ZΔM15 Tn10 (Tetr)]	Agilent
BL21-CodonPlus(DE3)-RIPL	*E. coli* B F– ompT hsdS(rB – mB –) dcm+ Tetr gal λ(DE3) endA Hte [argU proL Camr] [argU ileY leuW Strep/Specr]	Agilent
Plasmids		
pET28b(+)	T7lac promoter, Kan^R , CloDF13 ori, 6x-His	Novagen
pET28b(+).SapKO−CH.BspQI	pET28b(+) modified with BspQI cloning site	[15]
pDAP16	pDAP16.pET28b.PAO1.finR.CHx5	This work
pDAP20	pDAP20.pUC18.PAO1.finR-fpR	This work

Primers		
FPM1	ACGCTCTTCCATGAATTCACCCTCCGCC	IDT DNA
FPM2	TGCTCTTCTGTGGGGAATCAGCGGACGCACC	IDT DNA
FPM3	CAGATGAAGTAGAACTGCCG	IDT DNA
FPM4	CATGCAACGTCATACCCTC	IDT DNA
FPM5	CCAGGCTTCCTCGTCTAGAGC	IDT DNA
FPM6	CAAAGACTCCTAGGAAAACGC	IDT DNA
FPM7	CCATATCCATATTCTGGATAAGCATTATC CAGACAATTCATTTTG	IDT DNA
FPM8	CAAAATGAATTGTCTGGATAATGCTTAT CCAGAATATGGATATGG	IDT DNA
FPM9	CCATATCCATATTCTGGATAAGCATTATCC AGACAATTCATTTTGCGGATATTT	IDT DNA
FPM10	AAATATCCGCAAAATGAATTGTCTGGAT AATGCTTATCCAGAATATGGATATGG	IDT DNA
FPM11	GGGACCTTGATGCTGAAGAACTCCAG	IDT DNA
PAO1_cysB_BspQI-F	GGGCTCTTCCATGAAGCTTCAGCAATTGCG	IDT DNA
PAO1_cysB_BspQI-R	GGGCTCTTCAGTGGTAGACCGGCAGTTCGATG	IDT DNA

Following cloning, the reaction was transformed by electroporation into XL1 Blue *E. coli* cells (Agilent) on LB agar plates containing 50 μg/mL kanamycin. Clones were miniprepped and restriction digested with XbaI and XhoI to identify appropriate clones. Clones were sequenced using the T7 promoter and terminator primers (GENEWIZ). The generated plasmid, pDAP16, contained the gene encoding FinR with a 5x C-terminal His tag behind a T7 promoter (**Table 3-1**).

FinR Protein Expression and Purification

Protein expression was done using *E. coli* BL21-CodonPlus(DE3)-RIPL cells (Agilent) transformed with pDAP16 . The following antibiotic concentrations were used in LB agar plates for all transformations and protein expression cultures: 50 μg/mL of kanamycin or 100 μg/mL for the autoinduction culture, 75 μg/mL of streptomycin/spectinomycin, and 33 μg/mL of chloramphenicol. Following a fresh transformation, a 5 mL starter culture containing about 10 colonies and all three antibiotics was incubated with shaking at 37°C for 5 hr, after which the culture was inoculated into a 500 mL PA-5052 autoinduction culture media [16] containing the relevant antibiotics. The autoinduction culture was incubated at 25°C with shaking (300 RPM) until saturation (~20 hr). The cells were then harvested by centrifugation at 6000 x g for 10 min and the pellet was resuspended in 40 mL of Ni-NTA binding buffer A (20 mM tris, 25 mM imidazole, 500 mM NaCl, 10% glycerol, 10 mM 2-mercaptoethanol, pH 8.0). Then, PMSF in ethanol (1 mM final) was added to the cell extract, and the cells were lysed by French press (Thermo Electron). The cell-free crude extract was prepared by high speed centrifugation (60,000g) of the lysate in a Beckman AvantiTM J-25I centrifuge for 30 min. Protein purification was performed an ÄKTApurifier system (GE Healthcare Life Sciences) at room temperature. The crude extract was eluted through a 5 mL HisTrap HP column (GE Healthcare Life Sciences) equilibrated with the Ni-NTA binding buffer A at a flow rate of 3 ml/min. 10 column volumes (CV) of buffer A were eluted to wash off unbound contaminants before a gradient of 0% to 100% of elution buffer B (20 mM tris, 500 mM imidazole, 500 mM NaCl, 10% glycerol, 10 mM 2-mercaptoethanol, pH 8.0) was applied over 40 CV at a flow rate of 3 mL/min. Peak fractions containing the protein (~100 mM imidazole) were collected and analyzed by a 10% SDS-PAGE gel (EZ-Run Protein Gel Solution, Fisher Scientific) stained with Coomasie

blue G-250 [17]. Fractions containing FinR protein were then dialyzed overnight in buffer C (20 mM tris, 50 mM NaCl, 100 mM imidazole, 0.5 mM EDTA, 10% glycerol, and 10 mM 2-mercarptoethanol, pH 8.0). The dialysate was then loaded onto a 5 mL HiTrap Q column from GE Healthcare Life Sciences equilibrated in buffer C and eluted with a linear gradient (0-100%) with buffer C containing 1 M NaCl instead of 50mM NaCl over 40 CV at 3 mL/min. FinR protein fractions were collected and analyzed as before. The purified FinR samples were then pooled and extensively dialyzed in four successive 5 L of buffer C (20 L total) to ensure complete removal of any co-purified ligands. The protein sample was then quantified with a Bradford assay before storing it at –20°C in 20 mM tris, 50 mM NaCl, 100 mM imidazole, 50 % glycerol, 10 mM 2-mercarptoethanol, and 0.1 mM EDTA, pH 8.0. This was done by a 1:1 dilution of the protein sample with 100% glycerol.

Alignment and Motif Analysis

The *P. aeruginosa* FinR protein sequence was aligned with that of orthologs from other pseudomonads using Clustal Omega to assess the homology among them [18]. Genetic organization of the *P. aeruginosa finR–fprA* genetic region in comparison with that of other pseudomonads and select *Acinetobacter sp.* was assessed using Ortholuge DB [19] and Absynte [20]. Identification of conserved regulatory binding site motifs in the *finR–fprA* intergenic region of pseudomonads was performed using the motif discovery component of the MEME suite [21] using selected organisms (*P. aeruginosa, P. fluorescens, P. putida, P. stutzeri, Azoarcus* sp. *BH72, Azotobacter vinelandii*) sharing the same divergent orientation of *finR-fpr* identified using the SEED database [22].

Electrophoretic Mobility Shift Assays (EMSAs)

A DNA fragment containing the entire *finR* and *fprA* region of the *P. aeruginosa* PA01 genome (including their intergenic region) was PCR-amplified from *P. aeruginosa* str. PAO1 genomic DNA using the primers, FPM3 and FPM4 (**Table 3-1**). To generate plasmid pDAP20, the DNA fragment was then cloned into a pUC18 plasmid by blunt ligation following linearization of the plasmid by ZraI (New England Biolabs) using Fast-Link DNA Ligase (Epicentre). After confirming the sequence of the PCR product within the plasmid, pDAP20 was used as template for PCR to generate different DNA fragments of the intergenic regions with the following primer sets: FPM5 and FPM6 (PCR fragment: MX-1), FPM5 and FPM8 (PCR fragment: MX-2), FPM5 and FPM10 (PCR fragment: MX-3), FPM9 and FPM6 (PCR fragment: MX-4) (**Table 3-1**, **Figure 3-5A**). Additionally, two short synthetic DNA oligonucleotides, MP1 and MP2, were generated by providing conditions for the annealing of FPM7 and FPM8 (generating MP1) as well as FPM9 and FPM10 (generating MP2). MP1 (45 bp) and MP2 (54 bp) contained just enough nucleotides encompassing the putative FinR binding sites but differed in the presence or absence of the – 35 region and the binding site downstream of it. The oligonucleotides were annealed according to Integrated DNA Technologies' protocol (www.idtdna.com).

The EMSA buffer contained 10 mM tris acetate (pH 8.0), 1 mM potassium acetate, 5 mM ammonium acetate, 5 mM magnesium acetate, 1 mM calcium chloride, and 10 mM 2-mercaptoethanol. In addition to the buffer, each 10 µL volume reaction contained specific concentrations of DNA (2 nM or 10 nM) incubated with varying FinR concentrations (1 nM – 1000 nM) at 25°C for 60 min. After incubation, the reactions were loaded onto 6% PAGE gels (National Diagnostics) and electrophoresis was performed in a 0.5 x TAE buffer (pH 8.0) at a voltage of 200

V and a current of 22 mA. Gels were then stained with SYBR™ Green I Nucleic Acid Gel Stain (Invitrogen) before imaging on a Storm 825 Imaging System (GE Healthcare Life Sciences). In cases where the effect of a ligand (sulfite) was assessed, each 10 μL EMSA reaction contained 1 mM sodium sulfite final concentration, while the running buffer also contained 1 mM of the ligand. To determine the binding constants (Kd values) the intensities of the gel bands were determine with ImageJ [23], which allowed the fraction of protein bound to be determined. By following the procedure described by [24], and using GraphPad Prism 6, saturation binding curves were generated that were fitted to ligand depletion model (Y= (-b + sqrt(b*b-4*a*c))/(2*a), where a = -1 (assuming negligible non-specific binding); b = Kd + X + Bmax; c = Bmax(X)). Each Kd value was calculated based on data from duplicate experiments.

In vitro Transcription Assays

All reactions were performed with RNAse-free buffers and with conditions to minimize RNAse contamination. Template DNA for *in vitro* transcription assays was generated by PCR amplification using pDAP20 as template with the primers, FPM5 and FPM11 (**Table 3-1**). The resulting linear DNA template contained the *finR–fprA* intergenic region and the first 215 nucleotides of *fprA*. Initially, the template DNA (20 nM), purified FinR (100 nM), NTPs (1.25 mM), and sulfite where necessary (1 mM) were incubated for 30 min at 25°C in a 20 μL volume in transcription buffer T (40 mM tris base, 150 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 0.01% Triton X-100, pH = 7.5), 10 U of RNase Inhibitor Murine (NEB), and BSA (10 μg/mL, New England Biolabs). The transcription reaction was initiated by addition of *P. aeruginosa* RNA polymerase (300 nM) purified as described in **Chapter 2** and the reactions were incubated at 37°C for 1 hour. After incubation at 75°C for 10

min to denature the RNA polymerase, 1x DNase I reaction buffer (New England Biolabs) and 5U of DNase I (New England Biolabs) were added, and the reaction incubated at 37°C for 15 min. To visualize the RNA transcripts on a gel, aliquots of each sample (10 µL) were added to 10 µL of 2x RNA loading dye (95% formamide, 0.02% SDS, 0.02% bromophenol blue, 0.001% xylene, cyanol, 1 mM EDTA). The samples were then heated at 90°C for 2 min, chilled on ice, and run on a 6% Urea PAGE gel (SequaGel - UreaGel System, National Diagnostics). The gels were then stained with SYBR™ Gold Nucleic Acid Gel Stain (ThermoFisher Scientific) and imaged on a Storm 825 Imaging System (GE Healthcare Life Sciences). For quantification of the synthesized RNA products, the Quant-iT™ RiboGreen™ RNA Assay Kit (ThermoFisher Scientific) was used and fluorescence readings were taken with a SPECTRAmax M2e microplate reader (Molecular Devices).

3.4 RESULTS

Bioinformatic analysis

Bacteria commonly express ferredoxin-NADP+ reductases. Unlike many bacteria however, the pseudomonads were found to have a very similar genetic organization in the region expressing the *fprA*-encoded orthologs. **Figure 3-2** shows the synteny in this gene region among different *Pseudomonas sp.* as well as in two *Acinetobacter sp.* and it reveals that the *fprA* orthologs in pseudomonads are divergently expressed from another gene that is slightly larger, *finR*. These genes are annotated as encoding putative transcriptional regulators of the LysR-type family. Multiple sequence alignment of the various *finR* orthologs in different *Pseudomonas sp.* demonstrated their high relatedness (**Figure 3-3**).

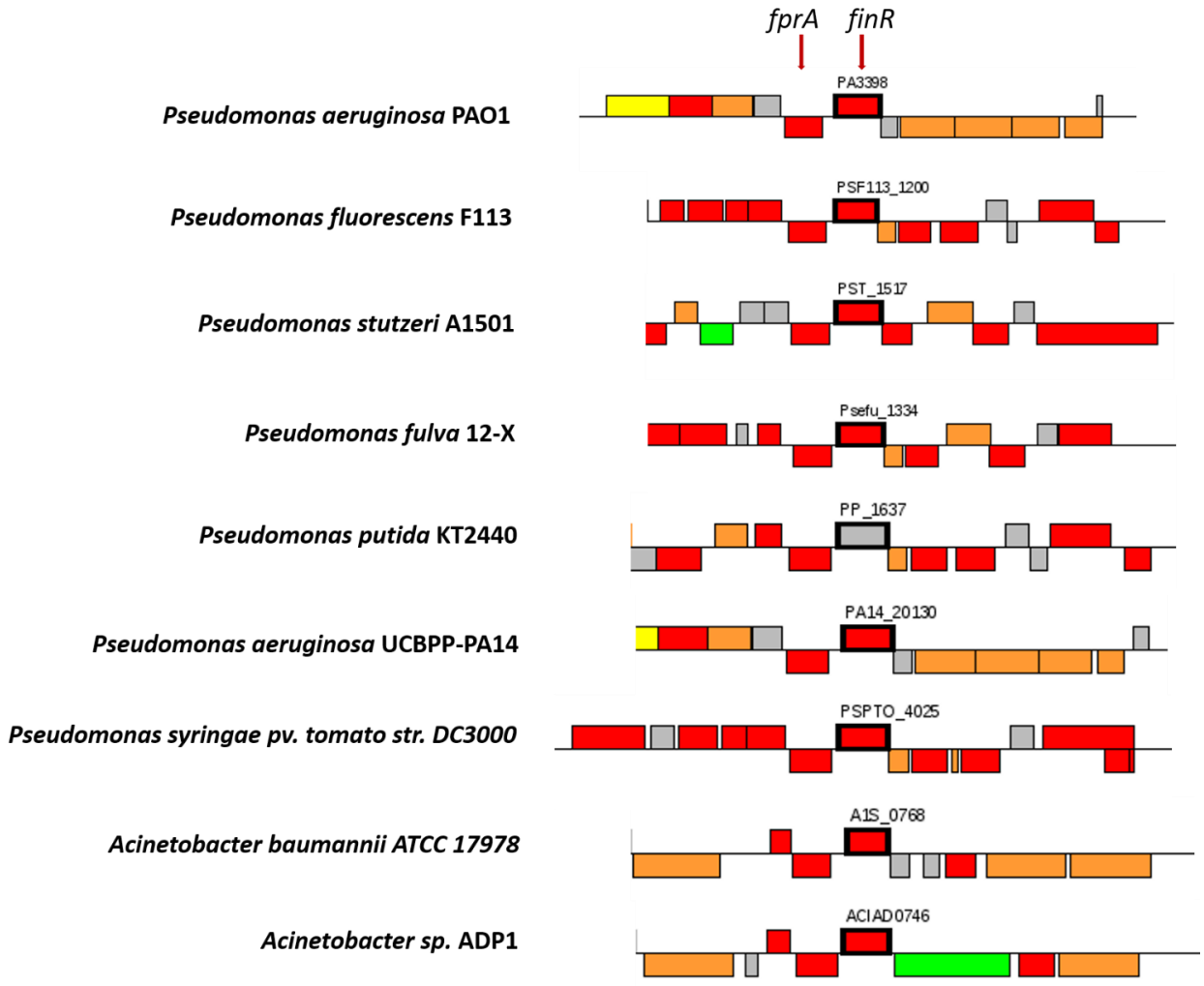


Figure 3-2. Genetic organization of the *fprA-finR* region showing synteny in multiple *Pseudomonas sp.* as well as in *Acinetobacter baylyi* ADPI and *Acinetobacter baumannii*. Rectangle colors indicate cellular location of the gene products: Red – cytoplasmic; Yellow – periplasmic membrane; Orange – cytoplasmic membrane; Green – outer membrane; Gray – unknown.

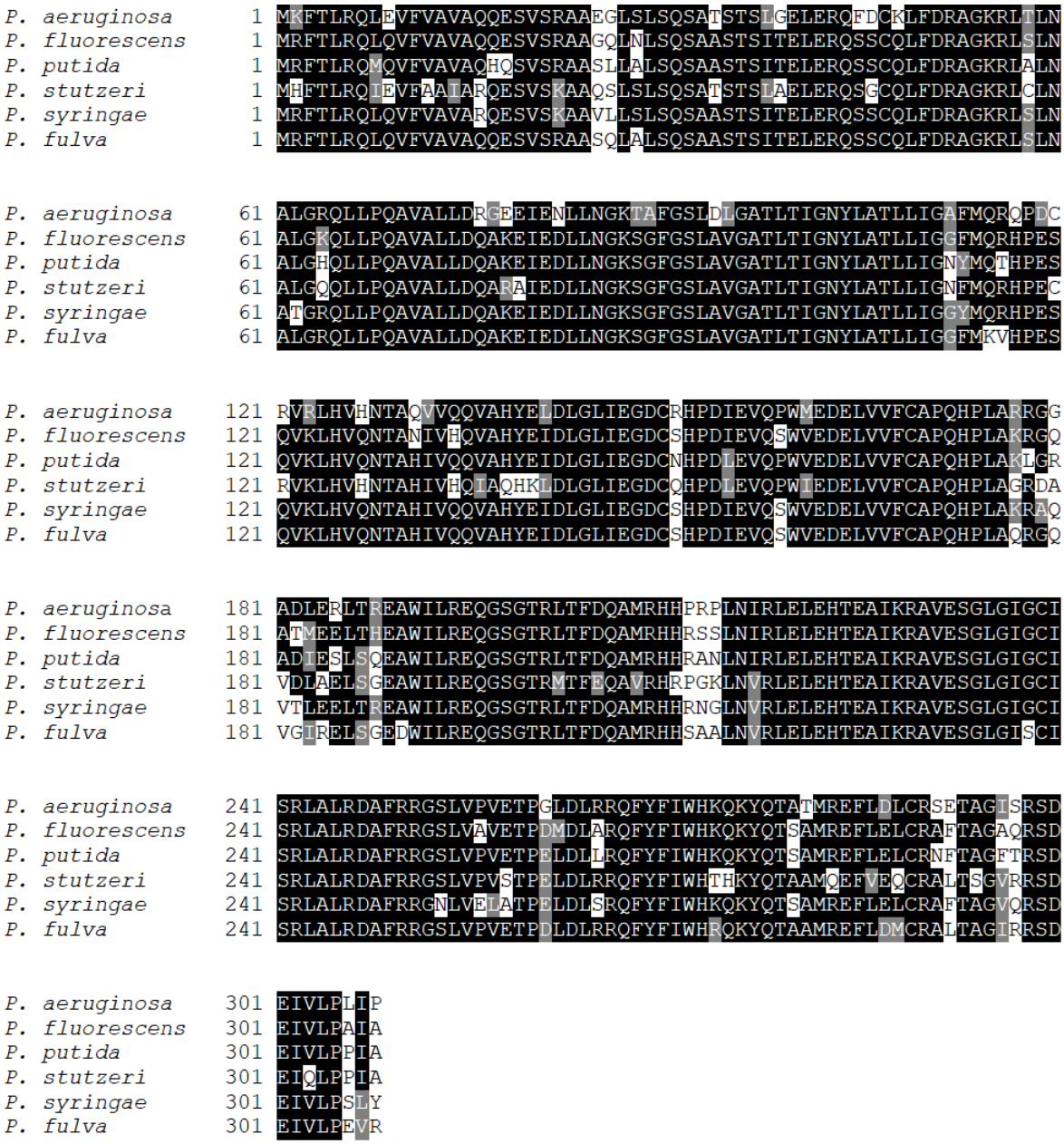


Figure 3-3. Multiple sequence alignment of FinR orthologs in six *Pseudomonas sp.* reveals high sequence similarities among the proteins that were predicted to be putative LTTRs.

To investigate whether the similar genetic organization in the *fprA–finR* region of the different *Pseudomonas sp.* translated into a potentially similar LTTR regulatory mechanism for the FprA enzyme, a MEME motif search was performed. LTTRs have been observed to typically bind regions containing the T-N_{11}-A motif, also called the LTTR box [25]. For the analysis, the *fprA–finR* intergenic regions in a set of pseudomonads were searched for the occurrence of conserved sequence motifs. As shown in **Figure 3-4A**, binding sites that comprised the characteristic inverted repeat sequence associated with LTTR boxes were found. The DNA sequence TATCC and its inverted repeat, GGATA, were prominent in the region. Some regions contained a strong palindrome (*e.g.* the sequence ATATCCATATTCTGGATAA, between the -10 and -35 regions of the *finR* promoter), while others contained significant variations from this unit with one side or the other having the site in a manner like the LTTR BenM from *Acinetobacter baylyi* [26]. If FinR regulates *fprA* expression and functions as a classical tetrameric LTTR in its mechanism of action, it is expected that a FinR DBD dimeric unit will bind to the perfect palindromic repeat sequence with the other DBD dimer binding to the imperfect sequences. This genetic arrangement allows the LTTR to switch from one sequence to another depending on whether the LTTR is activated or not by an inducer. These multiple T-N_{11}-A motifs are consistent with a classic LTTR regulatory mechanism for FinR that involves either the DNA bending or the sliding dimer model. In contrast to the pseudomonads, the *fprA–finR* intergenic region of *Acinetobacter* species did not contain any recognizable consensus motifs that could be potential FinR binding sites (**Figure 3-4B**). This was particularly an interesting observation because as shown in **Figure 3-2**, *A. baylyi* ADP1. and *A. baumannii* (and others not shown), have genetic organizations of their *finR* and *fprA* orthologs that are very similar to that of pseudomonads.

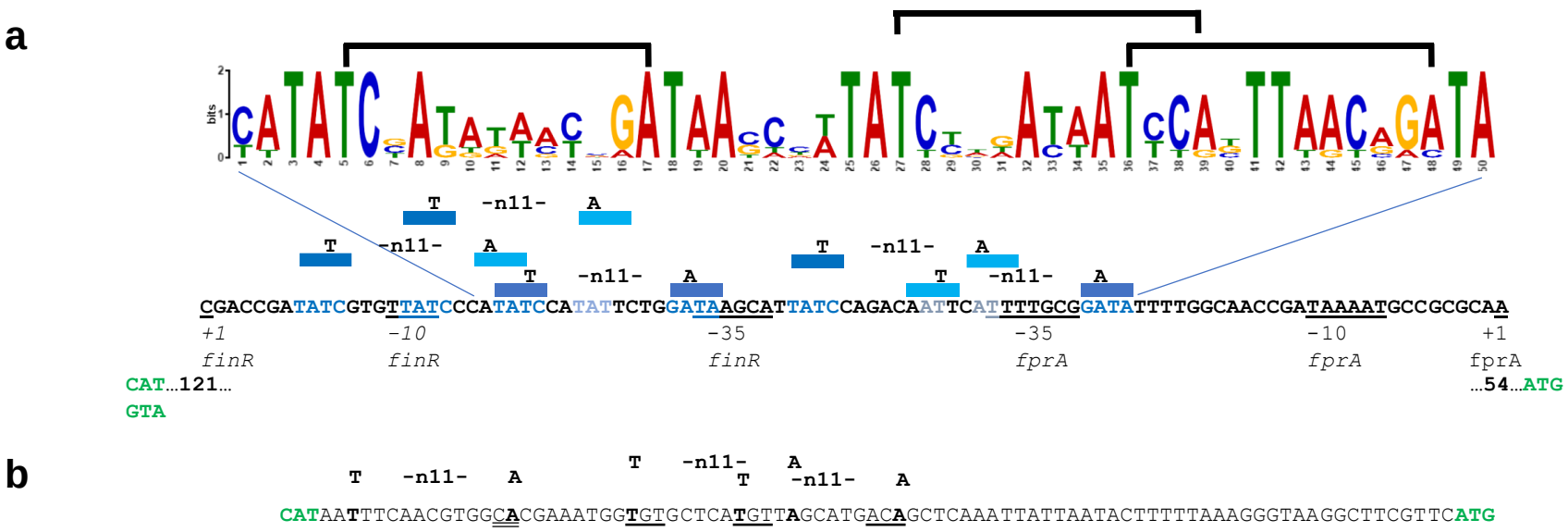


Figure 3-4. Results of MEME motif search for potential FinR binding sites in the *finR–fprA* intergenic region of various bacteria. **(A)** Pseudomonads contain a classic T-N_{11}-A motif in the region with a TATC consensus sequence. Each pair of dark blue rectangles indicate perfect TATC repeat sequences whereas the light blue rectangles indicate near perfect or imperfect TATC sequences. **(B)** Unlike pseudomonads, no FinR binding boxes were found in the *finR–fprA* intergenic region of *Acinetobacter baylyi* ADP1

Probing of FinR binding sites in the *finR–fprA* intergenetic region

Our finding of conserved LTTR boxes in the *finR–fprA* intergenic region hinted at the potential for FinR to regulate the expression of both genes. The intergenic region was therefore probed for FinR binding by incubating various concentrations of purified FinR (**Appendix Figure A3-1**) with DNA fragments (2 nM or 10 nM) that matched various parts of the intergenic region. The DNA fragments covered the intergenic region (MX-1), a region spanning the 5' end of the intergenic region through to the nucleotides just before the *fprA* –35 promoter element (MX-2), a region comprising MX-2 together with the *fprA* –35 element and the putative FinR binding site immediately downstream of it (MX-3), and a region containing about 40 upstream nucleotides of the *fprA* –35 element through to the 3' end of the intergenic region (MX-4) (**Figure 3-5a**).

Gel shift assays revealed that FinR binds strongly to the intergenic region with a Kd of about 3.6 nM (**Figure 3-5b, Table 3-2**), supporting the existence of FinR boxes in the intergenic region. Interestingly, shortening the intergenic DNA fragment to exclude the –35 promoter element of *fprA* and the FinR box immediately downstream of it (MX-2) reduced the affinity of FinR for the DNA substantially as shown in both **Figure 3-5c** and **Table 3-2**. In contrast, a shortened intergenic DNA that still contained the *fprA* –35 promoter element and the downstream FinR box (MX-3) had a better affinity for FinR compared to the affinity of MX-2 (Kd = 6.3 nM) (**Figure 3-5d**). This suggested that the region immediately upstream of the *fprA* –35 element was essential for FinR binding. A further confirmation of this theory was made by binding studies with DNA fragment MX-4 which excluded the first 140 nucleotides of the intergenic region. The first 54 nucleotides of MX-4 were therefore the last 54 nucleotides of MX-3 (**Figure 3-5a**). The MX-4 DNA was associated with strong binding to FinR (Kd = 2.7 nM), with binding affinity close to that of the full intergenic region (**Figure**

3-5e, Table 3-2). The binding studies with these four DNA fragments revealed that the binding sites of FinR in the *finR–fprA* intergenic region was restricted to a region of about 50 nucleotides spanning from the – 20 *fprA* position to the – 10 *finR* position.

Table 3-2. FinR binding constants determined for the different DNA fragments*

DNA	FinR Kd (nM)
MX-1	3.6 ± 0.7
MX-2	119.5 ± 29.9
MX-3	6.3 ± 1.2
MX-4	2.7 ± 0.3

*Based on data from duplicate experiments.

*Binding curves used for the Kd determination are shown in **Appendix Figure A3-3**.

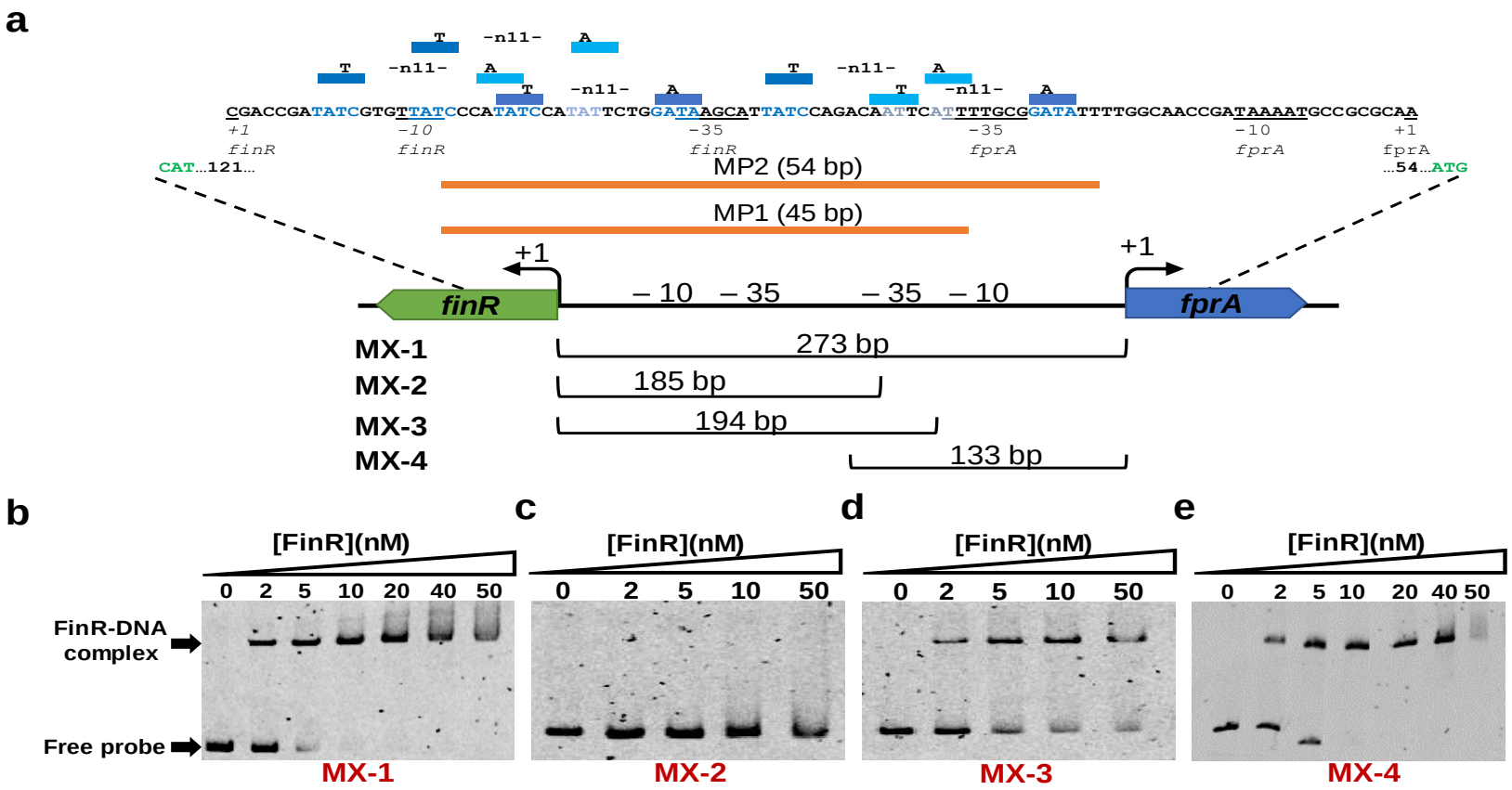


Figure 3-5. Probing of the *P. aeruginosa finR–fprA* intergenic region for FinR binding sites. **(a)** Various DNA fragments matching various regions of the intergenic region (MX1 – MX4) were amplified by PCR using the respective primers listed in **Table 3.1**. **(b) - (e)** 2 nM amount of each fragment was incubated with the indicated amounts of FinR. The binding constant (Kd) was determined from each gel (duplicate experiments) by fitting the band intensities to a ligand depletion model.

Sulfite is a potential FinR inducer

Because FinR seemed to act like a classic LTTR, we postulated that it will follow a classic LTTR regulatory mechanism called the 'sliding dimer' model in its regulation of FprA expression. In that model, an LTTR, in the absence of an inducer, will bind to the DNA in a way that precludes RNAP access to the –35 region (or other sigma factor binding sites, like an UP region) and hence prevent transcription of the downstream gene. Binding of an inducer causes conformational changes that result in the 'sliding' of the LTTR dimer to a secondary binding position that exposes the –35 element for RNAP-mediated transcription [25, 27, 28]. The identification of potential primary and secondary FinR boxes in the *finR–fprA* intergenic region (**Figure 3-4a**) are consistent with a sliding dimer mechanism of FinR activation. Additionally, the binding studies with MX-2 and MX-3 (**Figure 3-5c-d**) further hinted at a sliding dimer regulatory model for FinR since the two DNA fragments contained FinR binding sites favorable for binding depending on whether the LTTR was in the induced or the uninduced state. However, the important question that remained unanswered was the inducer that mediated FinR activation. We hypothesized that sulfite, the substrate of FprA could be this inducer.

Therefore, to assess the potential of sulfite as a FinR inducer that activates the LTTR through a 'sliding dimer' regulatory mechanism, more binding studies were done with oligonucleotides MP1 and MP2 (**Figure 3-5a**). Both MP1 and MP2 were designed to contain just enough nucleotides that encompassed all identified primary and secondary FinR boxes. Just like MX-2 and MX-3, MP1 and MP2 also contained regions that allowed FinR binding depending on whether the LTTR was induced or not. Using BenM again as a model, one would expect the binding sites of the induced protein to lie within the binding sites of the uninduced protein [26].

Therefore, because the 45 bp MP1 fragment lacked the putative FinR-binding site immediately downstream of the –35 element, binding of this fragment to FinR was expected to be less favorable in the uninduced state. In contrast, binding to the 54 bp long MP2 fragment was expected to be high in the uninduced state as this DNA contained the critical binding site downstream of the –35 element that was required for binding in this state. As shown in Figure **3-6a-b**, FinR bound very poorly to MP1 in the absence of inducer whereas strong binding was observed for MP2. These observations were similar to the results observed for the FinR-MX-2 and FinR-MX-3 binding studies (**Figure 3-5c** and **3-5d**) and it was not surprising since MX-2 and MX-3 are simply longer DNA fragments of MP1 and MP2 (respectively) that extend to the 5' end of the intergenic region. The results suggested that uninduced FinR had a higher affinity for DNA containing the – 35 region and the binding site immediately downstream of it compared to DNA lacking these regions. However, upon introduction of a FinR inducer, the LTTR is expected to have an increased affinity for both MP1 and MX-2 since both DNA fragments contain all the binding sites required for FinR binding in the induced state (the binding site downstream of the –35 region is not required for binding in this state). This is demonstrated in **Figure 3-6c** where MP1, which previously showed poor binding to FinR in the absence of sulfite, was found to have a substantially increased affinity for the LTTR in the presence of 1 mM sulfite. The binding affinity of FinR for MP2 was not expected to change significantly in the presence of sulfite since the DNA still contained the binding sites required for FinR binding in the induced state, which is what was observed (**Figure 3-6d**).

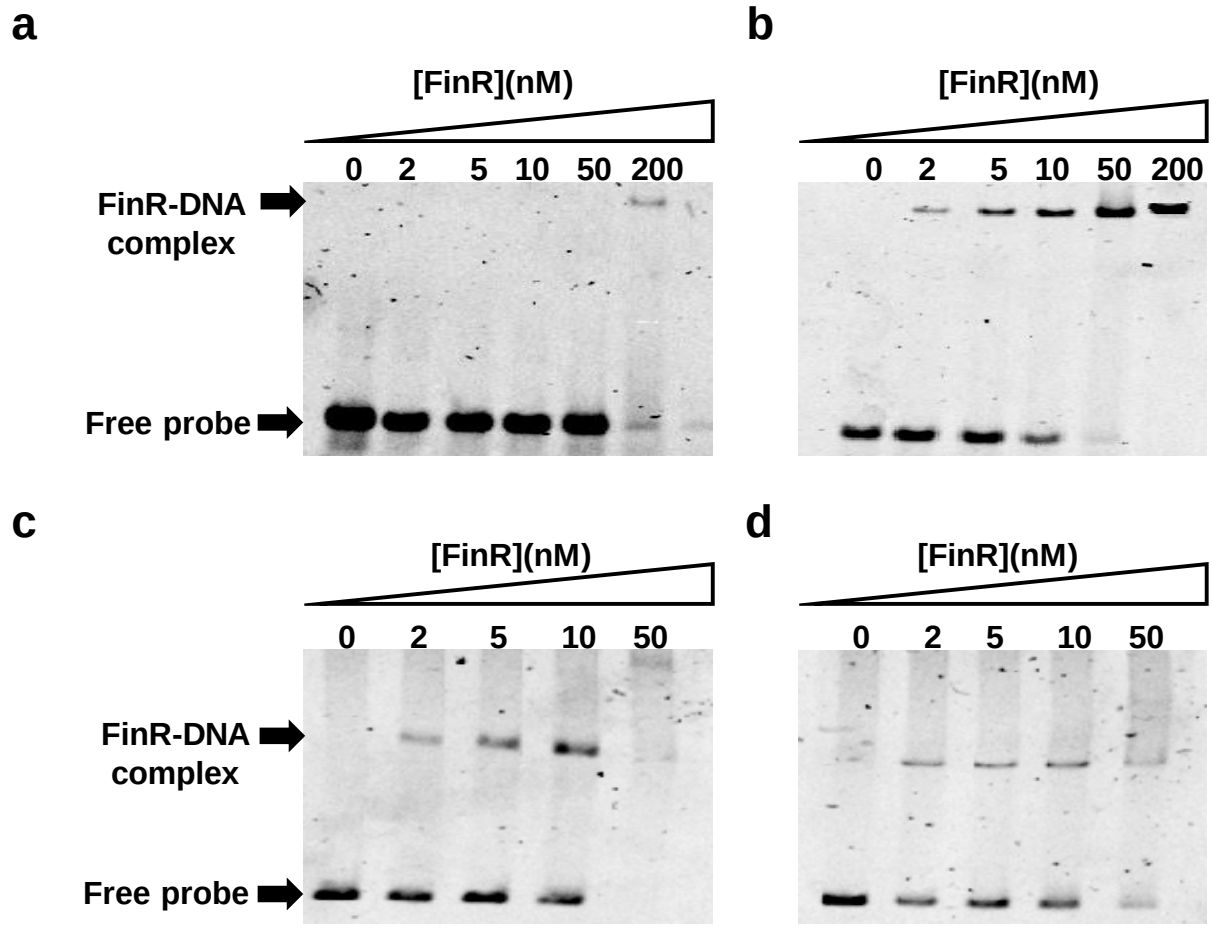


Figure 3-6. Assessment of sulfite as a potential inducer of FinR in its regulation of *fprA* expression. **(a)** The shorter MP1 fragment (45 bp), which lacked a –35 region and the FinR box immediately downstream of it that was required for binding in the uninduced state bound poorly to FinR. **(b)** In contrast, MP2 (54 bp), which contained all the binding sites required for FinR binding in the uninduced state, was associated with strong FinR binding. **(c)** In the presence of sulfite, the affinity of FinR for MP1 increases compared to its affinity for the DNA in the uninduced state because MP1 DNA contains all the binding sites required for FinR binding in the induced state. **(d)** The affinity of FinR for MP2 in the presence of sulfite expectedly did not change much.

The *in vitro* gel shift data suggested that FinR controlled *fprA* expression through a sliding dimer regulatory mechanism that was sulfite dependent. To further assess the potential of sulfite as a FinR inducer, an *in vitro* transcription assay was performed in the presence and absence of sulfite and the expression of downstream *fprA* vs. upstream *finR* genes were evaluated. As shown in **Figure 3-7a**, the DNA template (**TRX-5**) used was such that a 122 nt RNA runoff product (TX-1) would be synthesized from the *finR* promoter, while a 270 nt product (TX-2) would be made from the *fprA* promoter (*e.g.* the sulfite-induced state). In the absence of any FinR (**Figure 3-7b lane F3**), *finR* is transcribed but not *fprA*, mimicking the physiological state where basal levels of FinR are expressed to block *fprA* expression in the uninduced state. When the transcription reaction contains FinR, the LTTR autorepresses its own expression, while also either repressing *fprA* expression or not activating the expression due to the absence of an inducer. This is shown in **Figure 3-7b, lane F4** where neither *finR* nor *fprA* is transcribed. When sulfite is introduced into the reaction, transcriptional activation of *fprA* occurs due to the conformational changes that allow RNAP access to the full promoter by uncovering the -35 region of *fprA* (**Figure 3-7b, lane F5**). In this state, *finR* expression is also repressed as demonstrated by the lack of a TX-1 transcript product. As a control experiment, *in vitro* transcription assays were performed in parallel with the well-studied LTTR, CysB, a regulator of sulfate metabolism in bacteria. The sulfite-dependent changes in *fprA* transcription were not observed when the transcription reaction was repeated with CysB (**Figure 3-7b, Lanes C1-C4**). In fact, CysB did not transcribe *fprA* at all. CysB has been proposed to be involved in several cellular processes, but its most prominent roles have been demonstrated in sulfur and iron regulation [29-31].

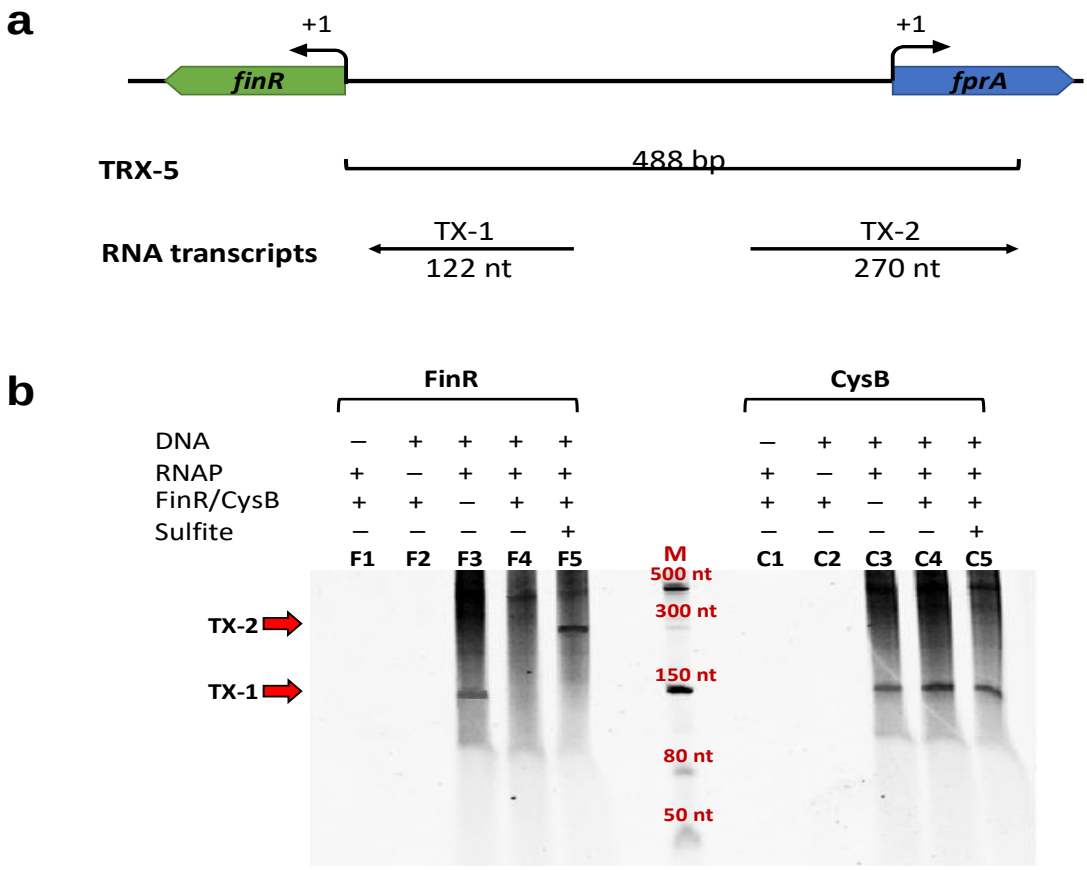


Figure 3-7. Transcription assay demonstrating the regulation of *fprA* expression by FinR with sulfite as the inducer. **(a)** A 488 bp DNA that comprised the *finR–fprA* intergenic region as well as the first 215 nucleotides of *fprA* was used as template for transcription assays. One of two transcript products, TX-1 or TX-2, is synthesized depending on whether transcription is initiated at the *finR* promoter or the *fprA* promoter (induced state), respectively. **(b)** Template DNA (20 nM) was incubated with 50 mM FinR or CysB with and without sulfite (1 mM) for 30 min before RNA synthesis was initiated by addition of 300 nM RNAP. The samples were run on a 6% urea PAGE gel which shows the two expected transcript products, TX-1 and TX-2, synthesized in the uninduced and induced FinR states, respectively. Lanes F1 and C1 are the negative control lanes lacking template DNA, whereas Lanes F2 and C2 are for the reactions without any RNAP.

The fact that neither FinR or CysB transcribed *fprA* suggests that this promoter, in the absence of an induced FinR, is not efficiently transcribed by a σ^{70}-RNAP. Interestingly, we found CysB to bind very tightly to the *finR–fprA* intergenic region. But it showed no discrimination in its binding affinity for MX2 and MX3, unlike FinR (data not shown). We also assessed FinR as a potential regulator of CysI, the major sulfite reductase enzyme in the sulfur pathway in many bacteria. Our results showed that FinR does not bind to the *cysI* promoter region and the presence of sulfite did not change its low affinity for the DNA (see **Appendix Figure A3-2**).

3.5 DISCUSSION

Ferredoxin-NADP+ reductases catalyze the reversible redox reaction between NADPH and one electron carriers like ferredoxin and they are encoded by the *fpr* genes, which are widespread in bacteria, archaea, eukaryotes, and plants [2]. In many bacteria, *fpr* is divergently expressed from a transcriptional regulator of the LysR-type [6]. The regulator, commonly called the ferredoxin reductase regulator (FinR), is one of the best characterized LysR-type Transcriptional Regulators (LTTRs) in pseudomonads, especially *P. putida* [1, 3, 4]. In *P. aeruginosa*, FinR was found to regulate one of the *fpr* genes, *fprA*, in an oxidative stress dependent manner [4]. Because Fpr functions in several redox reactions by mediating the donation of electrons to ferredoxin-dependent enzymes like nitrite reductases, sulfite reductases, glutamate synthases, and ferredoxin-thioredoxin reductases [32], the finding that its regulator–FinR–mediates the oxidative stress response in *P. aeruginosa* was not surprising. However, *P. aeruginosa*, like most bacteria, express another LTTR (OxyR) that is regarded as the master regulator of oxidative stress in bacteria [33-35]. Therefore, we suspected that the oxidative stress regulatory role of FinR is only a corollary of its involvement in

another major cellular process: sulfur assimilation. Indeed, *fprA* has been implicated in cellular processes like siderophore synthesis/regulation and iron acquisition, [1, 2, 8, 9, 36], both of which are connected to sulfur assimilation/cysteine biosynthesis.

We therefore investigated the potential involvement of the *P. aeruginosa* FinR in the sulfur assimilation pathway by assessing it ability to regulate the expression of *fprA*, which we suspect acts in the capacity of CysJ, a sulfite reductase whose ortholog is missing in the *P. aeruginosa* genome. Our finding of a conserved genetic organization in the *finR–fprA* region among different *Pseudomonas sp.* suggested a potentially similar regulatory mechanism. Unsurprisingly, bioinformatic analysis identified several consensus sequences (FinR boxes) that were conserved in the intergenic regions of the different pseudomonads. Expectedly, purified FinR was found to bind to both the *finR* and *fprA* promoters located within the intergenic region. Specifically, the putative FinR binding sites were located between positions –20 and –53 of the *finR* promoter, as well as –43 and –85 of the *fprA* promoter using the +1 *finR* and *fprA* positions as predicted by [4]. Binding of FinR to the *fprA* promoter in *P. aeruginosa* is consistent with findings regarding the *P. putida* FinR [3]. Although *P. aeruginosa* and *P. putida* are currently the only pseudomonads whose FinR regulatory role on FprA have been assessed, the conserved genetic organization of the *finR–fprA* region coupled with the presence of conserved FinR boxes in the intergenic region suggest that other pseudomonads likely undergo a similar FinR regulatory mechanism. Interestingly, although *Acinetobacter sp.* were found to contain a similar genetic arrangement in the *finR–fprA* region like that observed for *Pseudomonas sp.*, no FinR boxes were identified in the intergenic region by bioinformatic analysis. This observation is consistent with findings from another study in our lab that found that the *A. baylyi* ADP1 FinR binds to neither the *fprA* nor the *finR* promoter, although it binds to many other genes in the sulfur pathway (unpublished data). It is therefore possible that

FprAs in pseudomonads play a more critical role in sulfur assimilation compared to their orthologs in *E. coli* and other enteric bacteria. This reasoning is further supported by the fact that *fprA* is as essential gene that cannot be knocked-out in *P. aeruginosa* without fatal consequences for the cells. The essentially of FprA as a major player of the *P. aeruginosa* sulfur assimilation pathway could be explained by the absence from its genome of CysJ, a major sulfite reductase enzyme in *E. coli* and other enteric bacteria that complement the role of another sulfite reductase, CysI. Unlike the orthologs from other bacteria, the CysI from pseudomonads seems to be a distinct sulfite reductase that partners with either reduced ferredoxin or FprA directly rather than be CysJ dependent. FprA in *P. aeruginosa* therefore likely transfers electrons directly to CysI, which might explain how the enzyme is a critical one in the bacterium. In their study, [37] found that the *P. aeruginosa* FprA efficiently transferred electrons from NADPH to heme oxygenase.

In [4] where the *P. aeruginosa* FinR was characterized, although FinR knockout mutants were found to be susceptible to paraquat-induced oxidative stress, the study did not establish paraquat as a FinR inducer. Because we suspected FprA to be involved in the reduction of sulfite to sulfide in the *P. aeruginosa* sulfur assimilation pathway, we explored sulfite as the FinR inducer. We found sulfite to mediate FinR activation, subsequently leading to *fprA* expression, which was consistent with findings from other studies in our lab regarding the binding of sulfite to FinR orthologs from *A. baylyi* and *E. coli* (unpublished data). As the substrate of FprA in pseudomonads, sulfite was the most ideal candidate among the list of potential FinR effectors. Although it is a normal metabolite during reductive sulfate assimilation in bacteria that is generated through reduction of adenosine phosphosulfate (APS) by an adenylylsulfate reductase (CysH) [38], sulfite is also potentially toxic to cells beyond certain concentrations [39]. The antimicrobial and antioxidative properties of sulfite is therefore reason for its use as a preservative [40, 41]. These properties of

sulfite explain why its reduction to sulfide is essential in bacteria, and further establishes FprA and FinR as important proteins in *P. aeruginosa*. Sulfite is also the one metabolite in the sulfur assimilation pathway that could function as an oxidative stressor and could be the link between the oxidative stress regulatory role for FinR as demonstrated by [1, 3, 4] and the LTTR's role in sulfur assimilation as demonstrated in this study. Furthermore, the turnover of iron-sulfur clusters could lead to the generation of oxidized sulfur species like sulfite that could trigger the expression of FprA through FinR activation. Interestingly, *finR*-dependent, paraquat-induced *fprA* expression in *P. putida* was suppressed by reduced sulfur sources [1]. From the findings of this study, *finR*-dependent, sulfite-induced *fprA* expression will also likely be suppressed by reduced sulfur sources like sulfide. In this study, FinR was found to regulate *fprA* expression by following a classic LTTR 'sliding dimer' regulatory model of transcriptional activation. In the model, an LTTR oligomerizes to form a tetramer such that one dimer precludes RNAP-mediated transcription by preventing access to the –35 promoter element. However, the presence of an effector or some other stimuli induces conformational changes that results in sliding of the dimer to a new position to allow RNAP-mediated transcription [25]. A schematic representation of this mechanism as proposed for FinR-dependent, sulfite-induced *fprA* expression shown in **Figure 3-8**. FinR is proposed to exist as a tetramer of two homodimers such that one dimer binds at the perfect palindromic repeat sequence that makes up the recognition binding site (RBS). The location of the RBS between the –10 and –35 *finR* promoter elements also ensures the repression of *finR* expression upon binding of the FinR dimer to the site (**Figure 3-8a**). Downstream of the RBS are two activation binding sites (ABS) that contain imperfect palindromic repeats to allow easy movement of the second FinR dimer from one ABS site to the other.

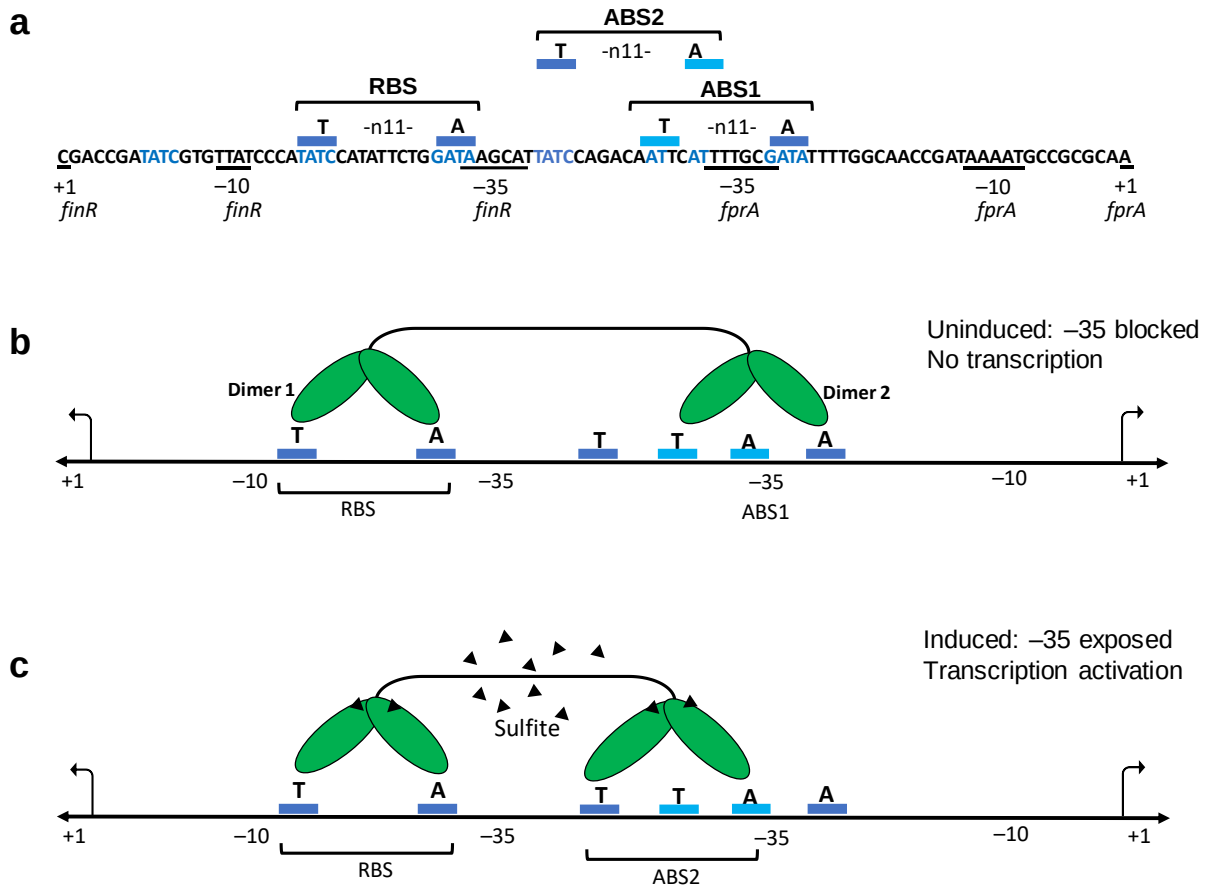


Figure 3-8. Regulatory mechanism of sulfite-mediated FinR regulation of *fprA* expression in *P. aeruginosa*. **(a)** The *finR–fprA* intergenic region showing the three predicted FinR binding sites. The recognition binding site (RBS) contains a perfect palindromic repeat as serves as a binding site for one of the FinR dimers whereas two activation binding sites (ABS) that contain imperfect palindromic repeats allow sliding of the second FinR dimer from one site to the next depending on the presence or absence of an inducer. **(b)** In the uninduced state, one dimer binds at the RBS and the other binds at the ABS site that stretches across the –35 element (ABS1). In this way, RNA polymerase is prevented from transcribing *fprA*. **(c)** However, in the presence of sulfite, conformational changes in the tetramer causes a sliding of the ABS1 dimer to a new position, ABS2 that exposes the –35 and hence leading to transcriptional activation.

In the uninduced state, the second FinR dimer binds to the ABS (ABS1) that spans the –35 *fprA* promoter element. Oligomerization of the two dimers leads to the formation of a closed tetrameric complex that blocks RNAP access to the –35 element (**Figure 3-8b**). However, binding of sulfite triggers a quaternary structural change of the LTTR tetramer that results in the formation of the open tetrameric LTTR form. Consequently, a rearrangement of the DNA-binding domains occurs that results in a sliding movement of the ABS1 dimer to the secondary binding position, ABS2, which exposes the –35 for RNAP-mediated transcription (**Figure 3-8c**). The location of ABS2 just upstream of the –35 element means RNAP can begin transcribing *fprA* uninhibited. This LTTR regulatory mechanism has been demonstrated with classic LTTRs like CbnR [42, 43] as well as BenM [44, 45] and CatM [46]. Although this FinR regulatory mechanism has not been demonstrated in any other pseudomonad, the conserved arrangement of the *finR–fprA* genetic region suggests that all pseudomonads likely follow a similar phenomenon.

As shown in **Figure 3-7b**, no *fprA* transcript product was synthesized in the absence of FinR, although the –35 promoter was available for RNAP-mediated transcriptional initiation. This suggests that the *fprA* –35 promoter is a weak one that requires the presence of a transcriptional regulator to activate transcription. In this way, the *P. aeruginosa* FinR acts as a transcriptional activator of FprA expression, rather than as a de-repressor.

3.6 CONCLUSIONS

The *P. aeruginosa* FinR is involved in the sulfur assimilation pathway through its regulation of the expression of FprA, an enzyme that acts in the capacity of a sulfite reductase in pseudomonads and therefore catalyzes the reduction of sulfite to sulfide in the pathway. Sulfite, the substrate of FprA,

was found to be the inducer of FinR in its regulation of *fprA* expression. The regulatory mechanism underlying the process conformed to the classic LTTR sliding dimer model where transcription of *fprA* was blocked in the absence of sulfite but activated following binding of sulfite to FinR. These findings establish a new, previously unreported role for the *P. aeruginosa* FinR that implicates the LTTR together with *fprA* as part of the cys regulon in pseudomonads. The findings also establish FinR as an important LTTR in *P. aeruginosa* and reveal the protein as a potential target for antibiotic drug discovery.

3.7 REFERENCES

[1] T.A. Lewis, A. Glassing, J. Harper, M.J. Franklin, Role for ferredoxin:NAD(P)H oxidoreductase (FprA) in sulfate assimilation and siderophore biosynthesis in Pseudomonads. J Bacteriol 195 (2013) 3876-3887.

[2] A. Romsang, J. Duang-nkern, W. Wirathorn, P. Vattanaviboon, S. Mongkolsuk, Pseudomonas aeruginosa IscR-Regulated Ferredoxin NADP (+) Reductase Gene (fprB) Functions in Iron-Sulfur Cluster Biogenesis and Multiple Stress Response. PloS one 10 (2015) e0134374.

[3] S. Yeom, J. Yeom, W. Park, Molecular characterization of FinR, a novel redox-sensing transcriptional regulator in Pseudomonas putida KT2440. Microbiology 156 (2010) 1487-1496.

[4] S. Boonma, A. Romsang, J. Duang-Nkern, S. Atichartpongkul, W. Trinachartvanit, P. Vattanaviboon, S. Mongkolsuk, The FinR-regulated essential gene fprA, encoding ferredoxin NADP+ reductase: Roles in superoxide-mediated stress protection and virulence of Pseudomonas aeruginosa. PLoS One 12 (2017) e0172071.

[5] Y. Lee, S. Peña-Llopis, Y.-S. Kang, H.-D. Shin, B. Demple, E.L. Madsen, C.O. Jeon, W. Park, Expression analysis of the fpr (ferredoxin-NADP+ reductase) gene in Pseudomonas putida KT2440. Biochemical biophysical research communications 339 (2006) 1246-1254.

[6] Y. Lee, J. Yeom, Y.S. Kang, J. Kim, J.S. Sung, C.O. Jeon, W. Park, Molecular characterization of fprB (ferredoxin-NADP+ reductase) in Pseudomonas putida KT2440. Journal of microbiology and biotechnology 17 (2007) 1504-1512.

[7] J. Yeom, C.O. Jeon, E.L. Madsen, W. Park, In vitro and in vivo interactions of ferredoxin-NADP+ reductases in Pseudomonas putida. Journal of biochemistry 145 (2009) 481-491.

[8] A. Romsang, S. Atichartpongkul, W. Trinachartvanit, P. Vattanaviboon, S.J.J.o.b. Mongkolsuk, Gene expression and physiological role of Pseudomonas aeruginosa methionine sulfoxide reductases during oxidative stress. (2013) JB. 00167-00113.

[9] R. Yan, S. Adinolfi, A. Pastore, Ferredoxin, in conjunction with NADPH and ferredoxin-NADP reductase, transfers electrons to the IscS/IscU complex to promote iron-sulfur cluster assembly. Biochim Biophys Acta 1854 (2015) 1113-1117.

[10] J. Kim, W. Park, Oxidative stress response in Pseudomonas putida. Applied microbiology and biotechnology 98 (2014) 6933-6946.

[11] S.A. Lee, L.A. Gallagher, M. Thongdee, B.J. Staudinger, S. Lippman, P.K. Singh, C.J.P.o.t.N.A.o.S. Manoil, General and condition-specific essential functions of Pseudomonas aeruginosa. (2015) 201422186.

[12] V. de Berardinis, D. Vallenet, V. Castelli, M. Besnard, A. Pinet, C. Cruaud, S. Samair, C. Lechaplais, G. Gyapay, C. Richez, M. Durot, A. Kreimeyer, F. Le Fevre, V. Schachter, V. Pezo, V. Doring, C. Scarpelli, C. Medigue, G.N. Cohen, P. Marliere, M. Salanoubat, J. Weissenbach, A complete collection of single-gene deletion mutants of Acinetobacter baylyi ADP1. Mol Syst Biol 4 (2008) 174.

[13] G.L. Winsor, E.J. Griffiths, R. Lo, B.K. Dhillon, J.A. Shay, F.S. Brinkman, Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. Nucleic acids research 44 (2016) D646-D653.

[14] G.L. Winsor, D.K. Lam, L. Fleming, R. Lo, M.D. Whiteside, N.Y. Yu, R.E. Hancock, F.S.J.N.a.r. Brinkman, Pseudomonas Genome Database: improved comparative analysis and population genomics capability for Pseudomonas genomes. 39 (2010) D596-D600.

[15] N.R. Galloway, H. Toutkoushian, M. Nune, N. Bose, C. Momany, Rapid cloning for protein crystallography using type IIS restriction enzymes. Crystal Growth Design 13 (2013) 2833-2839.

[16] F.W. Studier, Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41 (2005) 207-234.

[17] I. Syrový, Z. Hodný, Staining and quantification of proteins separated by polyacrylamide gel electrophoresis. Journal of Chromatography B: Biomedical Sciences and Applications 569 (1991) 175-196.

[18] S. Chojnacki, A. Cowley, J. Lee, A. Foix, R. Lopez, Programmatic access to bioinformatics tools from EMBL-EBI update: 2017. Nucleic acids research 45 (2017) W550-W553.

[19] M.D. Whiteside, G.L. Winsor, M.R. Laird, F.S. Brinkman, OrthologeDB: a bacterial and archaeal orthology resource for improved comparative genomic analysis. Nucleic acids research 41 (2012) D366-D376.

[20] A. Despalins, S. Marsit, J. Oberto, Absynte: a web tool to analyze the evolution of orthologous archaeal and bacterial gene clusters. Bioinformatics 27 (2011) 2905-2906.

[21] T.L. Bailey, C. Elkan, Fitting a mixture model by expectation maximization to discover motifs in bipolymers. (1994).

[22] R. Overbeek, T. Begley, R.M. Butler, J.V. Choudhuri, H.-Y. Chuang, M. Cohoon, V. de Crécy-Lagard, N. Diaz, T. Disz, R. Edwards, The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic acids research 33 (2005) 5691-5702.

[23] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. Nature methods 9 (2012) 671.

[24] M.A. Heffler, R.D. Walters, J.F. Kugel, Using electrophoretic mobility shift assays to measure equilibrium dissociation constants: GAL4-p53 binding DNA as a model system. Biochemistry and Molecular Biology Education 40 (2012) 383-387.

[25] S.H. Craven, O.C. Ezezika, C. Momany, E.L. Neidle, LysR homologs in Acinetobacter: insights into a diverse and prevalent family of transcriptional regulators. Molecular Biology Edited by U Gerischer Norfolk, UK: Caister Scientific Press (2008).

[26] A.M. Alanazi, E.L. Neidle, C. Momany, The DNA-binding domain of BenM reveals the structural basis for the recognition of a T-N11-A sequence motif by LysR-type transcriptional regulators. Acta Crystallogr D Biol Crystallogr 69 (2013) 1995-2007.

[27] S.E. Maddocks, P.C. Oyston, Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology 154 (2008) 3609-3623.

[28] M.P. Koentjoro, N. Ogawa, STRUCTURAL STUDIES OF TRANSCRIPTIONAL REGULATION BY LysR-TYPE TRANSCRIPTIONAL REGULATORS IN BACTERIA. Reviews in Agricultural Science 6 (2018) 105-118.

[29] F. Imperi, F. Tiburzi, G.M. Fimia, P. Visca, Transcriptional control of the pvdS iron starvation sigma factor gene by the master regulator of sulfur metabolism CysB in Pseudomonas aeruginosa. Environ Microbiol 12 (2010) 1630-1642.

[30] J. Hummerjohann, S. Laudenbach, J. Retey, T. Leisinger, M.A. Kertesz, The sulfur-regulated arylsulfatase gene cluster of Pseudomonas aeruginosa, a new member of the cys regulon. J Bacteriol 182 (2000) 2055-2058.

[31] M.A. Kertesz, K. Schmidt-Larbig, T. Wuest, A novel reduced flavin mononucleotide-dependent methanesulfonate sulfonatase encoded by the sulfur-regulated msu operon of Pseudomonas aeruginosa. J Bacteriol 181 (1999) 1464-1473.

[32] N. Carrillo, E.A. Ceccarelli, Open questions in ferredoxin-NADP+ reductase catalytic mechanism. European Journal of Biochemistry 270 (2003) 1900-1915.

[33] U.A. Ochsner, M.L. Vasil, E. Alsabbagh, K. Parvatiyar, D.J. Hassett, Role of the Pseudomonas aeruginosa oxyR-recG operon in oxidative stress defense and DNA repair: OxyR-dependent regulation of katB-ankB, ahpB, and ahpC-ahpF. J Bacteriol 182 (2000) 4533-4544.

[34] Y.S. Choi, D.H. Shin, I.Y. Chung, S.H. Kim, Y.J. Heo, Y.H. Cho, Identification of Pseudomonas aeruginosa genes crucial for hydrogen peroxide resistance. J Microbiol Biotechnol 17 (2007) 1344-1352.

[35] Y.J. Heo, I.Y. Chung, W.J. Cho, B.Y. Lee, J.H. Kim, K.H. Choi, J.W. Lee, D.J. Hassett, Y.H. Cho, The major catalase gene (katA) of Pseudomonas aeruginosa PA14 is under both positive and negative control of the global transactivator OxyR in response to hydrogen peroxide. J Bacteriol 192 (2010) 381-390.

[36] J. Yeom, C.O. Jeon, E.L. Madsen, W. Park, Ferredoxin-NADP+ reductase from Pseudomonas putida functions as a ferric reductase. Journal of bacteriology 191 (2009) 1472-1479.

[37] A. Wang, Y. Zeng, H. Han, S. Weeratunga, B.N. Morgan, P. Moënne-Loccoz, E. Schönbrunn, M. Rivera, Biochemical and structural characterization of Pseudomonas aeruginosa Bfd and FPR: ferredoxin NADP+ reductase and not ferredoxin is the redox partner of heme oxygenase under iron-starvation conditions. Biochemistry 46 (2007) 12198-12211.

[38] N. Kredich, Biosynthesis of Cysteine. (2008).

[39] F. Hennicke, M. Grumbt, U. Lermann, N. Ueberschaar, K. Palige, B. Böttcher, I.D. Jacobsen, C. Staib, J. Morschhäuser, M. Monod, Factors supporting cysteine tolerance and sulfite production in Candida albicans. Eukaryotic cell 12 (2013) 604-613.

[40] I.S. Chang, B.H. Kim, P.K. Shin, Use of sulfite and hydrogen peroxide to control bacterial contamination in ethanol fermentation. Appl Environ Microbiol 63 (1997) 1-6.

[41] S.V. Irwin, P. Fisher, E. Graham, A. Malek, A.J.P.o. Robidoux, Sulfites inhibit the growth of four species of beneficial gut bacteria at concentrations regarded as safe for food. 12 (2017) e0186629.

[42] S. Muraoka, R. Okumura, Y. Uragami, T. Nonaka, N. Ogawa, K. Miyashita, T. Senda, Purification and crystallization of a LysR-type transcriptional regulator CBNR from Ralstonia eutropha NH9. Protein peptide letters 10 (2003) 325-329.

[43] S. Muraoka, R. Okumura, N. Ogawa, T. Nonaka, K. Miyashita, T. Senda, Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. Journal of molecular biology 328 (2003) 555-566.

[44] S.H. Craven, O.C. Ezezika, S. Haddad, R.A. Hall, C. Momany, E.L. Neidle, Inducer responses of BenM, a LysR-type transcriptional regulator from Acinetobacter baylyi ADP1. Mol Microbiol 72 (2009) 881-894.

[45] A. Ruangprasert, S.H. Craven, E.L. Neidle, C. Momany, Full-length structures of BenM and two variants reve
al different oligomerization schemes for LysR-type transcriptional regulators. J Mol Biol 404 (2010) 568-586.

[46] O.C. Ezezika, L.S. Collier-Hyams, H.A. Dale, A.C. Burk, E.L. Neidle, CatM regulation of the benABCDE operon: functional divergence of two LysR-type paralogs in Acinetobacter baylyi ADP1. Appl Environ Microbiol 72 (2006) 1749-1758.

Chapter 4

THE POTENTIAL ROLE OF CYSB, A LYSR-TYPE TRANSCRIPTIONAL REGULATOR, AS A GLOBAL REGULATOR IN *PSEUDOMONAS AERUGINOSA*

4.1 SYNOPSIS

The cysteine biosynthesis regulator, CysB, is a LysR-type Transcriptional Regulator (LTTR) that controls the expression of genes associated with the biosynthesis of cysteine in bacteria. In *P. aeruginosa*, LTTRs make up close to 30% of the transcriptional regulator genes and they mediate the pathogen's high virulence, pathogenicity, as well as environmental and metabolic prowess. Therefore, characterization of the different LTTRs in *P. aeruginosa* may shed more light on LTTR-mediated transcriptional regulation in the bacterium while identifying new therapeutic targets for species-specific antibiotic drug discovery. In this study, we sought to functionally characterize the *P. aeruginosa* CysB by assessing its potential as a global regulator in the bacterium. Using a predicted CysB binding motif, we identified 25 putative CysB-co-regulated genes, many of which were involved in virulent processes. Gel shift assays revealed CysB to be an unusually promiscuous LTTR that formed multiple complexes with DNA fragments containing the promoters for the putative co-regulated genes. In σ^{70}-mediated transcription assays, CysB was found to regulate some of the genes including *cysI* and *pvdS*, whereas other genes like *mvfR* and *algD* were not found to be under CysB regulation. Overall, these findings support CysB as a global regulator in *P. aeruginosa*

that mediate various processes, primary virulent ones. Therefore, as a global regulator, CysB could be an important antibiotic drug discovery target whose targeting could lead to the dysregulation of several cellular processes.

4.2 INTRODUCTION

CysB, the cysteine biosynthesis regulator, is one of well-researched bacterial LysR-Type Transcriptional Regulators (LTTRs). The LTTR controls the cysteine regulon that oversees L-cysteine biosynthesis in many bacteria [1]. The regulon comprises the genes for L-cystine, glutathione, sulfate, and thiosulfate uptake; those for sulfate activation and reduction to sulfide; genes for both O-acetylserine (thiol)-lyase isozymes, and the genes involved in alkanesulfonate utilization. All these genes have been theorized to be under positive CysB regulation with N-acetyl-L-serine as its inducer [2-4].

Unlike the *E. coli* CysB, the *P. aeruginosa* CysB (PA1754) has not been extensively studied, even though it was one of the earliest LTTRs to be characterized in the pathogen. The transcriptional regulator was first shown to be involved in regulating the expression of AlgD, the enzyme that mediates alginate biosynthesis during colonization of the cystic fibrosis lung [5]. DNA encoding the *algD* gene regulatory region was used in pull down experiments with *P. aeruginosa* cell extracts, and CysB was identified as one of two proteins that was pulled down. Regarding its role in cysteine biosynthesis in *P. aeruginosa*, one of the earliest studies found CysB to be involved in the regulation of sulfate-starvation-induced (SSI) proteins, a role consistent with findings from *E. coli* studies [6]. These SSIs are expressed to enable cells to grow on organosulfates and are repressed by sulfate, cysteine, and thiocyanate. Through mutation studies, another operon, *msuEDC*, was discovered and

found to be under CysB regulation. The *msuD* gene was found to express an enzyme that catalyzed the desulfonation of alkanesulfonates and required O_2 and $FMNH_2$. The *msuEDC* operon has therefore been classified as being part of the CysB regulon. In another study, the *P. aeruginosa* CysB regulon was expanded to include the sulfur-regulated arylsulfate gene cluster (*ats* genes) which are required for sulfate ester desulfurization and transport [7]. The CysB regulon was further expanded to include the iron starvation sigma factor gene, *pvdS* [8]. PvdS is expressed in response to iron starvation and been shown to be involved in virulence [9] by mediating the expression of a host virulence factors. This new CysB role was pivotal as it linked the iron and sulfur regulons in *P. aeruginosa* to a single LTTR. CysB has also been implicated in *P. aeruginosa* quorum sensing through its interaction with the *pqsR* (also referred to as *mvfR*) promoter [10, 11]. The LTTR was shown to bind to the *pqsR* promoter to repress transcription of PqsR and hence repress production of the quinolone signal required for quorum sensing. Finally, CysB was found to mediate virulence through its regulation of the *P. aeruginosa* type III secretion system [12].

In this study, we investigated the role of CysB as a global regulator in *P. aeruginosa*. Using various bioinformatic tools, we identified putative CysB-coregulated genes in the *P. aeruginosa* genome and assessed their regulation by CysB using gel shift assays and transcription assays. The findings complement structural data on the full-length structure of the *Acinetobacter baylyi* ADP1 CysB (unpublished data) that point to an unusual regulatory mechanism for the LTTR.

4.3 MATERIALS AND METHODS

Cloning, Expression, and Purification of CysB

Table 4-1 lists all the bacterial strains, PCR primers, and plasmids used/generated in the study. The *cysB* gene was amplified from *P. aeruginosa* str. PAO1 genomic DNA to generate a DNA fragment with flanking BspQI sites to aid in restriction cloning. The amplification conditions and cloning approach was followed as described in **Chapter 3** for FinR. The resulting CysB expression plasmid, pDAP21, was transformed into BL21-CodonPlus(DE3)-RIPL cells and selected on agar plates containing 50 μg/mL kanamycin, 33 μg/mL chloramphenicol, and 75 μg/mL streptomycin/spectinomycin. The autoinduction method as described for FinR expression in **Chapter 3** was modified to ensure the absence of any sulfur-containing compounds that could result in the production of inactive CysB protein. This approach stemmed from the hypothesis that sulfate is an effector of CysB that could lock the protein in the inactive conformation during or after protein synthesis in the over-expression system. Therefore, our standard PA-5052 autoinduction media was modified such that all components contributing sulfur sources were replaced with sulfur-free alternatives. Because commercial kanamycin is only available as a sulfate salt, the sulfate was exchanged with chloride using Dowex® 1X8 100-200 (Cl-form) chromatographic resin (ThermoFisher Scientific). The resulting kanamycin chloride was quantified by a modification of the method described by [13]. The components of the modified M63 media are listed in the **Appendix** section (**Table A4-1**). All other protein expression and purification procedures were followed as described in **Chapter 3** for FinR. The purified low-sulfate CysB protein was stored in a

buffer containing 20 mM tris, 0.1 mM EDTA, 10 mM 2-mercaptoethanol, and 50% glycerol, pH 8.0, at -20 °C by doing a 1:1 dilution with 100% glycerol.

Table 4-1. Bacterial strains, plasmids, and primers used in the study.

Name	Description	Source
***E. coli* Strains**		
XL1-Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F´ proAB lacIq ZΔM15 Tn10 (Tetr)]	Agilent
BL21-CodonPlus(DE3)-RIPL	*E. coli* B F– ompT hsdS(rB – mB –) dcm+ Tetr gal λ(DE3) endA Hte [argU proL Camr] [argU ileY leuW Strep/Specr]	Agilent
Plasmids		
pET28b(+)	T7lac promoter, Kan^R , CloDF13 ori, 6x-His	Novagen
pET28b(+).SapKO−CH.BspQI	pET28b(+) modified with BspQI cloning site	Lab plasmid
pDAP21	pDAP21.pET28b.PAO1.cysB.CHx5	This work
pDAP33	pDAP33.pUC18.PAO1.algD_prom	This work
pDAP34	pDAP34.pUC18.PAO1.pvdS_prom	This work
pDAP35	pDAP35.pUC18.PAO1.mvfR_prom	This work
pDAP36	pDAP36.pUC18.PAO1.cysB_prom	This work
pDAP37	pDAP37.pUC18.PAO1.cysI_prom	This work
pDAP38	pDAP38.pUC18.PAO1.alkS_prom	This work
Primers		
PAO1_cysB_BspQI-F	GGGCTCTTCCATGAAGCTTCAGCAATTGCG	IDT DNA
PAO1_cysB_BspQI-R	GGGCTCTTCAGTGGTAGACCGGCAGTTCGATG	IDT DNA

PAO1_algDprom-F	CAATGCCCACGGCTATTACTTC	IDT DNA
PAO1_algDprom-R	TCGTAGTCCTTGATCGCGGTG	IDT DNA
PAO1_pvdSprom-F	CCGAATAAGGCAGGCAGAAC	IDT DNA
PAO1_pvdSprom-R	GCTGAGATGGGTGACGTTGTC	IDT DNA
PAO1_mvfRprom-F	CCGTGTCCCCTTGAGCAAAC	IDT DNA
PAO1_mvfRprom-R	CGAGCACGCACTGGTTGAAG	IDT DNA
PAO1_cysBprom-F	CCTCGATGAGGCGGAAAAGC	IDT DNA
PAO1_cysBprom-R	CGATGACGCAGCGATTCCAG	IDT DNA
PAO1_cysIprom-F	ATCCCCTACAACAGCATGGAC	IDT DNA
PAO1_cysIprom-R	GGATGATCTCGCACCAGGGG	IDT DNA
PAO1_alkSprom-F	GAGATAGTGGTTGACCCCTCG	IDT DNA
PAO1_alkSprom-R	CTGGAAGCCAGCAACTCGAC	IDT DNA
PAO1_retSprom-F	CACCGCGCTGAAGGATGGCCAGGTGG	IDT DNA
PAO1_retSprom-R	CCAGTTGGCCGTCCTGCACCAGGTAGTAGTCC	IDT DNA
PAO1_tauAprom-F	CTGTAGGGACGGGGGCATGGAGCAAGG	IDT DNA
PAO1_tauAprom-R	CGAACGGGGTAGCGATGGTCTTGCCGAC	IDT DNA

Search and Identification of CysB Co-Regulated Genes

Bioinformatic searches were done in the Pseudomonas Database (www.pseudomonas.com) to look for similarities among CysB orthologs among pseudomonads, while OrthoDB [14], together with Absynte [15] and the SEED database [16] were used for further bioinformatic assessments. Putative CysB-co regulated genes were predicted based on available RNA-seq analyses data, while the MEME suite [17] was used to search, analyze, and compare motifs among the different predicted

genes. Further CysB co-regulated genes were identified from a literature search for studies that have associated the *P. aeruginosa* CysB with specific regulatory processes. To predict σ70 binding sites in identified CysB gene targets, the BPROM Bacterial Promoters prediction program was used [18].

Electrophoretic Mobility Shift Assay Analysis

After identifying putative CysB co-regulated genes, PCR primers (**Table 4-1**) were used to amplify regions of the genome containing the target gene and its upstream intergenic region. Each PCR amplicon was ligated into a pUC18 plasmid by cut/ligation with the blunt end restriction endonuclease, ZraI, with incubation temperature modifications (37 °C instead of 50 °C) of [19]. The resulting plasmids are shown in **Table 4-1.** The generated plasmids were used as the templates for PCR reactions to generate the DNA fragments used for gel shifts and transcription assays. The buffer used for the gel shift assays contained 10 mM tris acetate, 1 mM potassium acetate, 5 mM ammonium acetate, 5 mM magnesium acetate, 1 mM calcium chloride, and 10 mM 2-mercaptoethanol (pH 8.0). In addition to the buffer, each reaction contained 2 nM of DNA as well as variable concentrations of CysB (1 – 500 nM) in a 10 μL total reaction mix. After 60 minutes incubation at room temperature, the reactions were loaded onto 4% native PAGE gels and electrophoresis was done in a 0.5x TAE buffer (pH 8.0) at a voltage of 200 V and a current of 22 mA. Gels were stained with SYBR™ Green I Nucleic Acid Gel Stain (Invitrogen) before imaging on a Storm 825 Imaging System (GE Healthcare Life Sciences). To test the effect of ligands on binding of CysB to the respective DNA fragments, 1 mM of the specific ligand was added to the 10 μL EMSA reaction mix as well as to the electrophoresis running buffer.

Transcription Assays

The PCR fragments generated for the gel shift analysis were used as template for transcription assays. The transcription assays were done as described in **Chapter 3** for FinR using 20 nM of each DNA template, 100 nM of purified CysB, and 200 nM of RNAP. To visualize the RNA transcripts on a gel, aliquots of each sample (10 µL) was added to 10 µL of 2x RNA loading dye (contains 95% formamide, 0.02% SDS, 0.02% bromophenol blue, 0.001% xylene, cyanol, 1 mM EDTA), heated at 90°C for 2 min, chilled on ice, then run on a 6% Urea PAGE gel (SequaGel - UreaGel System, National Diagnostics). The gels were then stained with SYBR™ Gold Nucleic Acid Gel Stain (ThermoFisher Scientific) and imaged on a Storm 825 Imaging System (GE Healthcare Life Sciences).

4.4 RESULTS

Identification of CysB co-regulated genes in *P. aeruginosa*

CysB is one of few LTTRs that have been implicated in several regulatory functions, suggesting its potential as global regulator in *P. aeruginosa*. To probe this further, bioinformatic searches were done to identify putative CysB-co-regulated genes in *P. aeruginosa*. In their study, Imperi and colleagues [8] reported the role of CysB in the regulation of PvdS, a Group 4 alternate sigma factor in *P. aeruginosa*. Using the proposed CysB binding sites on the *pvdS* promoter from the study, a search was done in RNAseq databases to identify putative genes harboring these binding sites, and the results are indicated in **Figure 4-1**. The putative co-regulated genes were found to contain the

predicted binding sites in their upstream intergenic regions. Furthermore, some of them were found to contain dyad repeat sequences that suggested the location of LysR boxes in the upstream intergenic regions. A list of the predicted genes together with their respective roles if known is outlined in **Table 4-2.** MEME analysis also identified the ATGC motif as being widespread in the upstream intergenic region of the different genes (**Figure 4-1**). Interestingly however, using TOMTOM (MEME Suite) to search for potential matches of the predicted CysB motifs in prokaryotic databases did not yield any convincing results, as it did not identify the motif as one belonging to a transcription factor.

>PA0185/4/3 fig|208964.1.peg.185 Alkanesulfonates transport system permease protein
catggggcggctccttggaacgggaagagagggattcccgccgggagcgggtgtgcggctagctatgcacagcggcgcctcgaaaataaataataaaaattcatttttt
tataaccacaagcagtatgacaggccgcacgggcctctccgcgtcatatgcagccagagcatggaaaatattctccaaatgcatttgccgctgcctggggccttttctttt
aaatggcccttcttagcgaaaaaatgcatttattcggcgatttatattcccATG

>PA0192/PA0191 fig|208964.1.peg.192 Putative TonB-dependent receptor
gccgcgcggacgaaactgcggtgttcggcgagggcgatgaagtaacgcagttggcgcagatccatgatgctttaccggcattggaaatattggacaaaggcatttgccg
ttcgggatgtgctggttttaaatgcaagctcttattccgtcaacaaagaacaaattagaaaaatatgcaatttagtaataagcatagatagaccccaggagctgaacat
ggtccaccgtgccttgccgccccacctccgccgttccgaatgcccttccgcgtcgccggcggctgtcgcccctgacagtcgcccggggctgagccgtccgccggcccc
cttcgaacctttcgccagatcgctgccttgcgcggcatggccgcctgctgcctgcgtccggctgcgccggaaccgaggacattcacgATG

>PA3930 (CioA) DOOR transcript cioA/cioB
ccaacacgcgaggggcggcttgaaggggagagtagcgactctttttattccataaaaatattatttgtttatttttatattctattaatgaatatgagagagctgaggc
aatttgtctcgcgacaaccgaccacagcaggacggtctaggctaacgggtaagattttcgtaccgaattgctctgcatcaaggcgggcggtccgggaatctgcttgcga
cccttcccgcattcctgccatccagggctgcggtcgctttccaatgacaagaacgatagctgaaggatgtgttccaaaaccttgcgtagcgttgcgcgtgaggaagcac
ggccgaatccgcacgtaagccggaagggccgagaacctgaagaggagttgccatg

>PA2335/PA2336/ PA2334 transcript DOOR fig|208964.1.peg.2335 TonB-dependent receptor
tatgcattttctacatggaaaatatcgggcgaatgcaattgccggggacggcgtgggcttttttataaaggcaatctcttattccgtaaaagcaatataaaacataaata
aataaccatacggaattagcagtcacccccttcatccaggacccacacATG

>PA2482/PA2481 transcript DOOR fig|208964.1.peg.2482 FIG135464: Cytochrome c4
gtctttcggtgataggtttttttggttctaaacttagaactaaaaaacaataagccagagaggaaccctgccATG

>PA2331 (transcript PA2331/30/29/28 fig|208964.1.peg.2331 Macrophage infectivity potentiator-related protein
tttagaatattgaatggatttcttattaccaagagacttcttcgacacctcaaccaccggagtcgaaagcATG

Figure 4-1. Identification of CysB-co-regulated genes in *P. aeruginosa*. The search was done in RNAseq databases for the CysB binding site (colored in yellow) as predicted by [8]. The blue-colored sequences indicate identified palindromic repeats that could confirm binding of CysB to these regions to regulate the downstream genes. Start codons of the target genes are colored in green. Predicted σ^{70} -35 and -10 regions are bold/underlined.

Table 4-2. CysB-co regulated genes in *P. aeruginosa* as identified from literature or predicted from RNAseq databases using the putative CysB binding site as predicted by [8]

Locus Tag	Gene name	Product name/predicted function
PA0185	*atsB*	probable permease of ABC transporter
PA0186	*atsR*	probable binding protein component of ABC transporter
PA0191	-	probable transcriptional regulator
PA0192	-	probable TonB-dependent receptor
PA0197	*tonB2*	transporter protein TonB2
PA1003	*pqsR*	LysR-type transcriptional regulator MvfR [11]
PA1431	*rsaL*	regulatory protein RsaL
PA1432	*lasI*	autoinducer synthesis protein LasI
PA1756	*cysH*	3'-phosphoadenosine-5'-phosphosulfate reductase
PA1838	*cysI*	sulfite reductase CysI
PA2009	*hmgA*	homogentisate 1,2-dioxygenase
PA2310	*tauD*	taurine dioxygenase TauD
PA2331	-	probable transcriptional regulator
PA2335	-	probable TonB-dependent receptor
PA2432	*bexR*	bistable expression regulator, BexR
PA2482	-	probable cytochrome c
PA2601	-	probable transcriptional regulator
PA3397	*fprA*	ferredoxin NADPH reductase FprA
PA3540	*algD*	GDP-mannose 6-dehydrogenase AlgD
PA3709	-	probable major facilitator superfamily (MFS) transporter [5]
PA3930	*cioA*	cyanide insensitive terminal oxidase
PA3936	*TauC*	probable permease of ABC taurine transporter
PA3938	*tauA*	probable periplasmic taurine-binding protein precursor
PA5483	*algB*	two-component response regulator AlgB
PA4856	*retS*	Regulator of Exopolysaccharide and Type III Secretion

Binding of CysB to different DNA fragments

To assess the potential regulation of the predicted genes by CysB, a subset of these genes were obtained by PCR amplification using primers such that the generated product contained the upstream intergenic region harboring the putative CysB binding sites as well as sequences in both directions of the intergenic region to enable an assessment of the direction of transcription. The selected genes included, *cysI*, *pvdS*, *mvfR*, *algD*, *atsB*, *cysB*, and *fprA*. Of these genes, *pvdS*, *mvfR*, *algD*, and *atsB* were chosen because, apart from being predicted among the list of CysB-co-regulated genes by bioinformatic analysis, at least one reported study has also demonstrated the regulation of these genes by CysB. The *cysB* gene was chosen to assess whether CysB was involved in autoregulation, while *cysI* and *fprA* were chosen to assess the involvement of CysB in the regulation of the *P. aeruginosa* sulfur assimilation pathway, the pathway the LTTR has been proposed to regulate in multiple bacteria. Results from gel shift analysis of these genes using purified low sulfate CysB (**Appendix Figure A4-1**) are shown in **Figure 4-2**. CysB was found to bind tightly to each of the DNA fragments in a similar manner that was particularly uncharacteristic of LTTRs. The protein formed multiple complexes with each DNA as depicted by the multi-band nature of the gels.

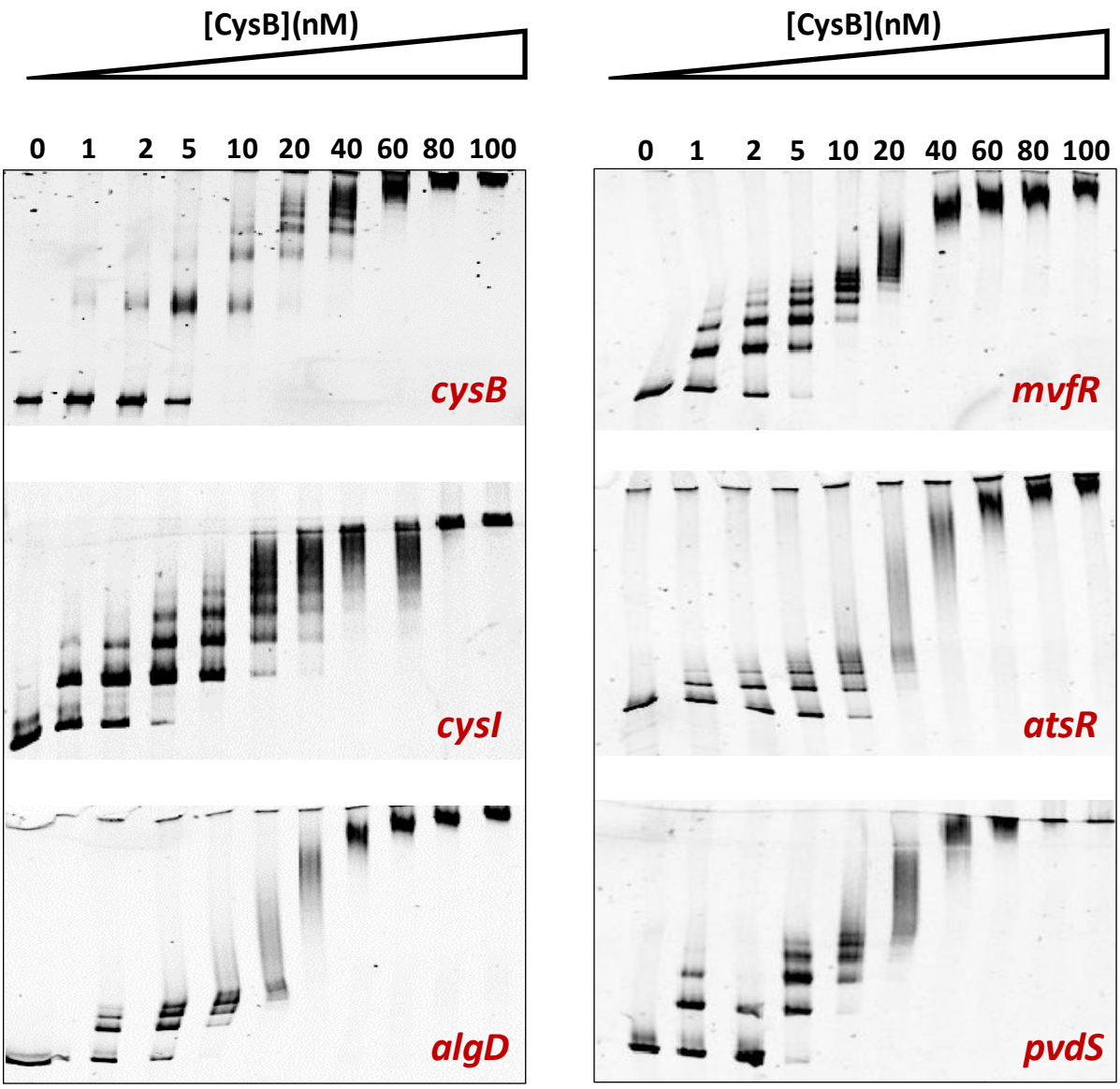


Figure 4-2. Binding of CysB to different potential gene targets. 2 nM of each DNA fragment was incubated with 0 – 500 nM of purified CysB (low sulfate) and the reactions were assessed by native PAGE.

CysB transcription studies

The binding studies with CysB seem to suggest that the LTTR regulates all these genes, lending further support to the theory that the protein is a global regulator. However, it is also very possible that these observations are attributable to the promiscuity of the LTTR in its DNA binding properties. To distinguish true regulation from mere binding *in vitro*, one approach is an *in vitro* transcription assay. Having produced the *P. aeruginosa* RNAP, such experiments are therefore possible. To illustrate the concept that binding of a transcription factor to DNA does not necessarily equate to transcriptional regulation, the *fprA* gene with its promoter was used. This DNA fragment was described in **Chapter 3** for use in assessing the regulatory effect of FinR on *fprA* expression. It was found that CysB bound very tightly to the *fprA* promoter (**Figure 4-3**). As evident from **Figure 4-3A** the multi-band gel shift phenomenon was again observed for CysB on this gene, although such observations were not made with FinR (**Chapter 3**). Interestingly, although CysB was associated with tighter binding to the *fprA* gene region than FinR, the LTTR was not associated with transcriptional activation of *fprA* expression, unlike FinR (**Figure 4-3B**). Instead, as observed in **Lanes C3-C5** of **Figure 4-3B**, *finR* expression was always activated regardless of the presence or absence of CysB or sulfite. These results therefore do not support CysB as a regulator of *fprA* expression.

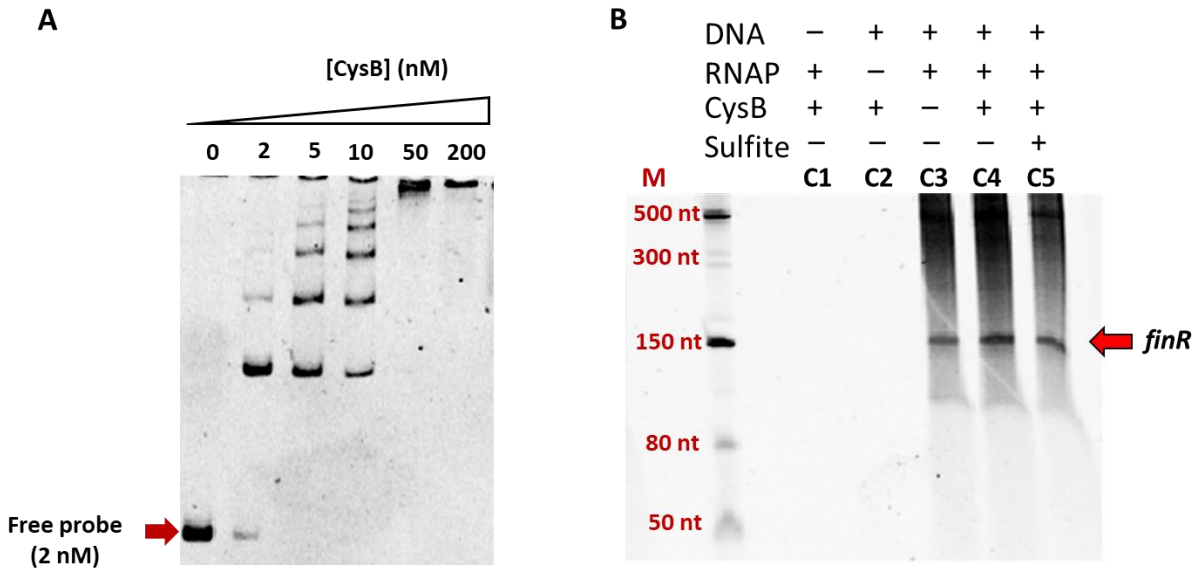


Figure 4-3. Binding of *P. aeruginosa* CysB to the *fprA* promoter **(A)**. Gel shift assays showed tight CysB binding to the DNA fragment. **(B)** Transcription assays with the DNA did not result in the transcriptional activation of *fprA* expression as observed for FinR in **Chapter 3**.

To further assess the potential regulation of these gene targets by CysB, transcription assays were performed. The –35 and –10 consensus σ^{70} binding sites on each of the gene targets were predicted to enable an analysis of the expected RNA products from the transcription. For the transcription assays, the potential of various ligands to serve as inducers or repressors of CysB was assessed. The different ligands- sulfite, sulfate, thiosulfate, adenosine 5′-phosphosulfate (APS), O-acetyl serine (OAS), and N-acetyl serine (NAS) - are intermediates in the sulfur assimilation pathway in bacteria. As shown in **Figure 4-4**, various transcription outcomes were observed for the different gene targets. For some of them like *cysI* and *pvdS*, differential transcription depending on the presence or absence

of CysB or some of the ligands were observed, whereas others like *algD* and *mvfR* were not associated with any reasonable differential transcription. From the assays, sulfite, sulfate, and thiosulfate were associated with an inhibitory action on transcription in the presence of CysB, suggesting that these ligands were potential negative CysB regulators. In all the transcription assays, end-to-end transcript products, (which are typical in *in vitro* transcription assays) were observed. These products form due to failure of some RNAP molecules to fall off at the end of transcription, leading to a looping of the template DNA such that the RNAP is permitted to transcribe in the reverse direction, forming a long transcript product typically the length of the template DNA or by initiation of transcription at the ends of the linear DNA [20]. Notably, transcription of these end-to-end RNA products (indicated by the blue arrows in **Figure 4-4**) were not affected by the inhibitory effects of sulfite, sulfate, and thiosulfate, whereas the opposite was observed regarding the transcription of the expected products (indicated by the red arrows). This suggested that the inhibitory action observed for these ligands on transcription were not due to an inhibition of RNAP activity. For some other gene targets like *algD* and *mvfR*, no reasonable gene transcripts were observed. Results of further gel shift experiments to probe the effect of sulfate and sulfite on CysB function are shown in **Figure 4-5**. Although sulfate (and sulfite) was not found to completely repress CysB binding to these DNA fragments, the affinity of CysB for each DNA fragment was slightly reduced in the presence of the ligands compared to its affinity for the DNA fragments in the absence of the ligands (**Figure 4-5**). Particularly, the addition of sulfite or sulfate both seemed to eliminate some of the complexes observed in the absence of the ligands, as the multi-band phenomenon was less pronounced.

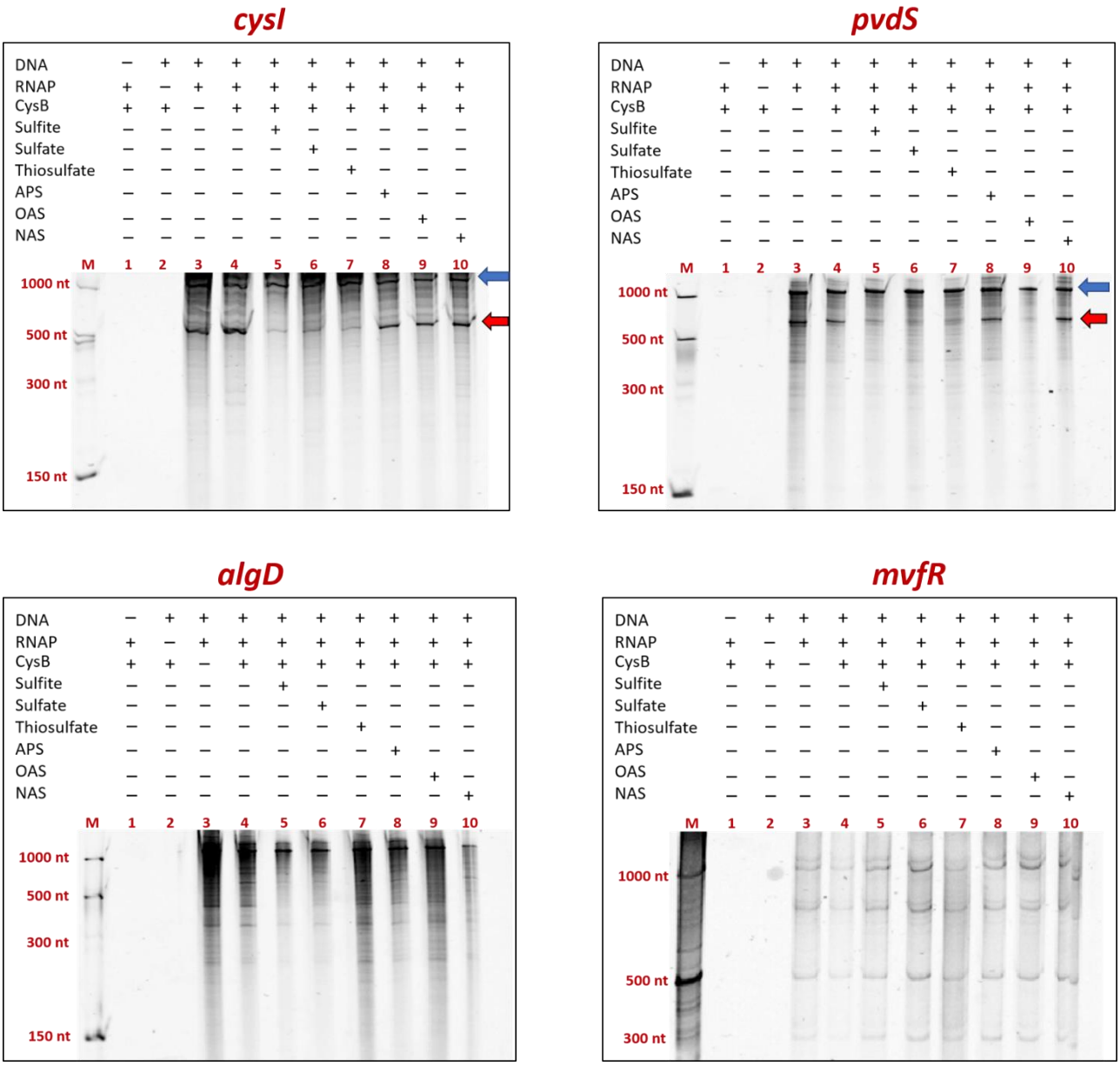


Figure 4-4. CysB-mediated transcription of various CysB-co-regulated genes in *P. aeruginosa*. CysB was first incubated with the DNA fragment and the respective ligand for 30 min before RNAP was added for transcription to continue for another 60 min. The RNA products were then extracted from the reaction and assessed by urea PAGE. The red arrows indicate the expected transcript products and the blue arrows indicate end-to-end transcript products.

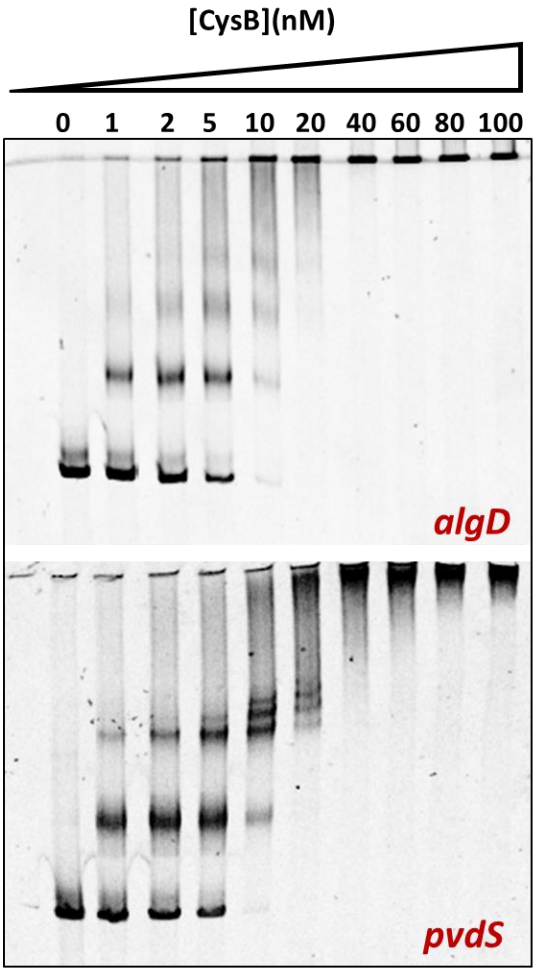


Figure 4-5. Binding of CysB to the algD and pvdS promoters in the presence of 1 mM sulfate.

4.5 DISCUSSION

The *P. aeruginosa* CysB was first described by [5] who found the protein to regulate the synthesis of alginate, a virulence factor produced by the pathogen during cystic fibrosis development. Prior to this discovery, CysB was regarded as the master regulator of cysteine biosynthesis in

enterobacteria bacteria because it had been thought to regulate many genes and steps of the sulfur assimilation/cysteine biosynthesis pathway [1]. Although CysB has been implicated in other cellular processes like siderophore biosynthesis [8], quorum sensing, [10, 11], and the *P. aeruginosa* type II secretion system [12], the LTTR is still relatively less studied in *P. aeruginosa* compared to studies in *E. coli* and *S. typhimurium*. In this study, we sought to assess the potential of CysB as a global regulator in *P. aeruginosa*. This was done by assessing the affinity of CysB for putative co-regulated genes that were identified, as well as by assessing the transcription of these genes in a CysB dependent manner.

Unlike most LTTRs, CysB is not divergently expressed from any gene in its neighborhood. Additionally, genetic analysis of the CysB genetic region found that the region is not conserved in pseudomonads. LTTRs are typically divergently expressed from their target genes and are conserved in their genetic arrangement among multiple organisms [21, 22]. Therefore, the findings regarding CysB suggested that the protein was a non-classic LTTR in its properties and potentially in its behavior. In their study, [8] assessed regulation of the expression of PvdS, an alternative sigma factor in *P. aeruginosa*, and found CysB to be involved in the regulatory process. The CysB binding site on the *pvdS* promoter was found in the study to encompass the sequence ATTTCTAACTAGCTGATTCCTAAAGATTATTAAAAAAATTT. Using this sequence, a search was done in RNAseq databases to identify putative CysB-co-regulated genes that contained this sequence or slight variations of it. As shown in **Table 4-2**, several genes were identified based on this motif. Many of the genes have been implicated in virulent processes in *P. aeruginosa* and they include BexR [23], PqsR [10, 11, 24], LasI [25, 26], AlgD [5], RetS [12] and PvdS [8, 9, 27]. CysB therefore likely plays an important role in disease progression and lifestyle choice of *P. aeruginosa*, making the LTTR one that could be an ideal target for drug discovery. The identified genes were

also found to contain a somewhat conserved repeat sequence (ATGC) within their upstream intergenic regions, further supporting them as CysB-co-regulated genes. However, unlike typical LTTRs, these palindromic sequences were asymmetric. Furthermore, the motif was not identified as a transcription factor motif using a transcription factor prediction program (TOMTOM). Overall, the findings from the bioinformatic analysis revealed CysB to be an LTTR with irregular properties that potentially regulated several diverse genes, many of which are involved in the production of virulent factors.

To further probe the potential of the identified genes as co-regulated genes of CysB, gel shift assays were done. As shown in **Figure 4-2**, CyB was associated with an unusual gel shift behavior where multiple DNA bands were observed on the gels. This alluded to the formation of multiple CysB-DNA complexes and this was consistent with other CysB studies that have demonstrated CysB binding through gel shift assays [8, 11, 12]. Interestingly, these findings regarding the binding properties of CysB are seemingly consistent with the X-ray structure we have obtained for the *A. baylyi* ADP1 CysB full-length protein (unpublished data) that reveals an uncharacteristic (for LTTRs) "donut"-shaped structure. The structure supports the possibility that the protein could potentially slide onto the DNA and load up multiple DNA fragments with the DNA passing through the donut hole, explaining the multiple complexes observed. It is possible then that such a structure is necessary to support CysB's role as a global LTTR, although no known similar LTTR structures have ever been reported. In *P. aeruginosa*, the only LTTR whose full-length structure has been reported is OxyR, the master regulator of oxidative stress in bacteria [28]. Although OxyR has similarly been considered a global regulator like CysB due to its regulation of multiple genes involved in diverse processes, its structure was found to conform to classic LTTR structure. Additionally, X-ray structures of two bacterial CysBs have so far been reported but both also

conform to classic LTTR structure. The first structure is that of the regulatory domain the *Klebsiella aerogenes* CysB [29] whereas the second is that of the DNA-binding domain of the *P. aeruginosa* CysB bound to DNA [12]. None of these structures bear any similarities to that obtained by us for the *A. baylyi* CysB, probably because none of these structures are that of the full-length protein. Unfortunately, every attempt by us to crystalize the full-length *P. aeruginosa* CysB proved futile, although such structural data will be the key to understanding the unusual behavior of the LTTR as observed in the study.

The gel shift experiments with CysB (**Figure 4-2**) also suggests that the LTTR is a promiscuous protein that can bind to any DNA it contacts with. Such an unusual LTTR behavior as demonstrated by CysB can be attributed to its role as a global regulator. However, no other LTTRs have been identified in the literature that demonstrate a binding behavior similar to CysB. In this study, we utilized transcription assays to differentiate true transcriptional regulation from mere DNA binding. Such a distinction is particularly necessary for promiscuous transcriptional regulators like CysB. To demonstrate that DNA binding of CysB does not necessarily equate to transcriptional regulation, the *fprA* promoter was utilized in a transcription assay with CysB. In **Chapter 3**, FinR was found to regulate *fprA* expression in a sulfite dependent manner. As shown in **Figure 4-3A**, CysB was associated with very tight binding to the DNA in a similar manner as observed for other DNA fragments. However, no CysB-dependent *fprA* expression was observed. One of three explanations are true in this case: 1) FprA is not under CysB regulation, 2) CysB indirectly regulates FprA expression by regulating FinR expression, and 3) CysB directly regulates FprA expression but other regulatory elements must be present. Indirect CysB regulation of FprA expression through FinR will also support CysB's role as the master regulator of sulfur assimilation and hence cysteine biosynthesis in *P. aeruginosa*. In such a manner, CysB could be acting as a positive regulator of

FinR, and hence *fprA* expression. Another possibility is that CysB is required for an enhanced FinR activity on the *fprA* promoter. This will further support CysB as a global regulator as it means the LTTR exhibits one more characteristic of global LTTRs: the need for co-regulating transcription factors.

From the transcription assay experiments, the CysB co-regulated genes identified in this study can be grouped under two categories. The first category comprises genes that are likely under CysB regulation as deduced by the formation of the expected transcript products in a manner that was ligand-dependent (**Figure 4-4**). Such genes included CysI, the other major sulfite reductase in *P. aeruginosa*, and PvdS, the alternative sigma factor. The second category comprises genes like AlgD and PqsR (MvfR) that were not associated with any reasonable expected transcript products in CysB transcription assays. Interestingly, these two genes are ones where at least one study has reported CysB to be involved in their regulation. In the case of PqsR, the quorum sensing regulator, CysB was found to be a negative regulator [11], which might explain why no transcript products were observed in the presence of CysB. However, the absence of CysB also did not result in any reasonable transcript product(s). It is therefore possible that these genes are not under direct CysB control; however, it is also possible that other regulatory proteins like alternate sigma factors are involved in the transcriptional regulation of these genes. Indeed, for *algD*, the extracytoplasmic function (ECF) protein, AlgU, has been implicated in its regulation [30]. In the transcription assay experiments, the activity of CysB was found to be consistently reduced in the presence of sulfite, sulfate, and thiosulfate as effectors. This observation is consistent with studies of CysB in *E. coli* and *S. typhimurium* [2, 31-34]. Sulfite, sulfate, and thiosulfate have all been considered anti-inducers that repress binding of CysB to the *cys* promoters [32, 34, 35]. Generally, sulfur starvation in bacteria is a signal to globally repress all systems and processes like siderophore biosynthesis that utilize

cysteine, whereas it is a signal to activate expression of genes that increase sulfur assimilation and cysteine biosynthesis. As revealed by the gel shift assays, the affinity of CysB for the DNA fragments was modified slightly in the presence of sulfate (**Figure 4-5)** as well as sulfite (data not shown). As evident from **Figure 4-5**, the presence of any of these ligands led a reduction in the number of CysB-DNA complexes observed in any lane compared to similar experiments in the absence of the ligands (**Figure 4-2**). This could be related to conformational changes in the 'donut' CysB structure that also contribute to the observed reduction in the CysB transcription activity. These observations need to be investigated further. Additionally, one way to further validate the identified genes as true CysB-co-regulated genes would be create the *P. aeruginosa cysB* knockout mutant and assess the expression and regulation of these genes.

4.6 CONCLUSIONS

The *P. aeruginosa* CysB is a non-classic LTTR that likely acts in a global capacity. We identified several potential CysB co-regulated genes in the *P. aeruginosa* genome that contained predicted CysB binding sites, notably the ATGC repeat sequence. CysB also seems to be promiscuous in its DNA binding ability and the protein was found to bind tightly to a wide array of gene targets. Unlike classically-behaved LTTRs, multiple CysB-DNA complexes were observed in binding studies, supporting an uncharacteristic LTTR structure that we have observed in the X-ray structure of another CysB ortholog. Transcription assays revealed that CysB activity was repressed by sulfite and sulfate, consistent with expectations for these ligands in regulating CysB-mediated transcriptional regulation of the cysteine biosynthesis genes. Overall, these studies expand current

knowledge on characteristics and function of the *P. aeruginosa* CysB and present complementary data that support existing research data on CysB proteins in the lab.

4.7 REFERENCES

[1] N. Kredich, Biosynthesis of Cysteine. (2008).

[2] J. Ostrowski, N. Kredich, Negative autoregulation of cysB in Salmonella typhimurium: in vitro interactions of CysB protein with the cysB promoter. Journal of bacteriology 173 (1991) 2212-2218.

[3] S. Henikoff, G.W. Haughn, J.M. Calvo, J.C. Wallace, A large family of bacterial activator proteins. Proceedings of the National Academy of Sciences 85 (1988) 6602-6606.

[4] M.A. Schell, Molecular biology of the LysR family of transcriptional regulators. Annual Reviews in Microbiology 47 (1993) 597-626.

[5] I. Delic-Attree, B. Toussaint, J. Garin, P.M. Vignais, Cloning, sequence and mutagenesis of the structural gene of Pseudomonas aeruginosa CysB, which can activate algD transcription. Mol Microbiol 24 (1997) 1275-1284.

[6] M.A. Kertesz, K. Schmidt-Larbig, T. Wuest, A novel reduced flavin mononucleotide-dependent methanesulfonate sulfonatase encoded by the sulfur-regulated msu operon of Pseudomonas aeruginosa. J Bacteriol 181 (1999) 1464-1473.

[7] J. Hummerjohann, S. Laudenbach, J. Retey, T. Leisinger, M.A. Kertesz, The sulfur-regulated arylsulfatase gene cluster of Pseudomonas aeruginosa, a new member of the cys regulon. J Bacteriol 182 (2000) 2055-2058.

[8] F. Imperi, F. Tiburzi, G.M. Fimia, P. Visca, Transcriptional control of the pvdS iron starvation sigma factor gene by the master regulator of sulfur metabolism CysB in Pseudomonas aeruginosa. Environ Microbiol 12 (2010) 1630-1642.

[9] L. Leoni, N. Orsi, V. de Lorenzo, P. Visca, Functional analysis of PvdS, an iron starvation sigma factor of Pseudomonas aeruginosa. Journal of bacteriology 182 (2000) 1481-1491.

[10] J.M. Farrow, 3rd, E.C. Pesci, Distal and proximal promoters co-regulate pqsR expression in Pseudomonas aeruginosa. Mol Microbiol 104 (2017) 78-91.

[11] J.M. Farrow, 3rd, L.L. Hudson, G. Wells, J.P. Coleman, E.C. Pesci, CysB Negatively Affects the Transcription of pqsR and Pseudomonas Quinolone Signal Production in Pseudomonas aeruginosa. J Bacteriol 197 (2015) 1988-2002.

[12] Y. Song, C. Yang, G. Chen, Y. Zhang, Z. Seng, Z. Cai, C. Zhang, L. Yang, J. Gan, L. Haihua, Molecular insights into the master regulator CysB-mediated bacterial virulence in Pseudomonas aeruginosa. Molecular microbiology 0 (2019).

[13] N. Shehzadi, K. Hussain, M.T. Khan, M. Salman, M. Islam, H.M. Khan, Development of a Validated Colorimetric Assay Method for Estimation of Amikacin Sulphate in Bulk and Injectable Dosage Form. J Chem Soc Pak 38 (2016) 63.

[14] M.D. Whiteside, G.L. Winsor, M.R. Laird, F.S. Brinkman, OrtholugeDB: a bacterial and archaeal orthology resource for improved comparative genomic analysis. Nucleic acids research 41 (2012) D366-D376.

[15] A. Despalins, S. Marsit, J. Oberto, Absynte: a web tool to analyze the evolution of orthologous archaeal and bacterial gene clusters. Bioinformatics 27 (2011) 2905-2906.

[16] R. Overbeek, T. Begley, R.M. Butler, J.V. Choudhuri, H.-Y. Chuang, M. Cohoon, V. de Crécy-Lagard, N. Diaz, T. Disz, R. Edwards, The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic acids research 33 (2005) 5691-5702.

[17] T.L. Bailey, M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, J. Ren, W.W. Li, W.S. Noble, MEME SUITE: tools for motif discovery and searching. Nucleic acids research 37 (2009) W202-W208.

[18] V.S.A. Salamov, A. Solovyevand, Automatic annotation of microbial genomes and metagenomic sequences. Metagenomics and its applications in agriculture Nova Science Publishers, Hauppauge, NY, USA (2011) 61-78.

[19] N.R. Galloway, H. Toutkoushian, M. Nune, N. Bose, C. Momany, Rapid cloning for protein crystallography using type IIS restriction enzymes. Crystal Growth Design 13 (2013) 2833-2839.

[20] Y. Gholamalipour, A. Karunanayake Mudiyanselage, C.T. Martin, 3′ end additions by T7 RNA polymerase are RNA self-templated, distributive and diverse in character—RNA-Seq analyses. Nucleic acids research 46 (2018) 9253-9263.

[21] S.H. Craven, O.C. Ezezika, C. Momany, E.L. Neidle, LysR homologs in Acinetobacter: insights into a diverse and prevalent family of transcriptional regulators. Molecular Biology Edited by U Gerischer Norfolk, UK: Caister Scientific Press (2008).

[22] S.E. Maddocks, P.C. Oyston, Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology 154 (2008) 3609-3623.

[23] K.H. Turner, I. Vallet-Gely, S.L. Dove, Epigenetic control of virulence gene expression in Pseudomonas aeruginosa by a LysR-type transcription regulator. PLoS genetics 5 (2009) e1000779.

[24] G. Xiao, E. Déziel, J. He, F. Lépine, B. Lesic, M.H. Castonguay, S. Milot, A.P. Tampakaki, S.E. Stachel, L.G. Rahme, MvfR, a key Pseudomonas aeruginosa pathogenicity LTTR-class regulatory protein, has dual ligands. Molecular microbiology 62 (2006) 1689-1699.

[25] J.P. Pearson, K.M. Gray, L. Passador, K.D. Tucker, A. Eberhard, B.H. Iglewski, E. Greenberg, Structure of the autoinducer required for expression of Pseudomonas aeruginosa virulence genes. Proceedings of the National Academy of Sciences 91 (1994) 197-201.

[26] L. Passador, J.M. Cook, M.J. Gambello, L. Rust, B.H. Iglewski, Expression of Pseudomonas aeruginosa virulence genes requires cell-to-cell communication. Science 260 (1993) 1127-1130.

[27] P.A. Beare, R.J. For, L.W. Martin, I.L. Lamont, Siderophore-mediated cell signalling in Pseudomonas aeruginosa: divergent pathways regulate virulence factor production and siderophore receptor synthesis. Mol Microbiol 47 (2003) 195-207.

[28] I. Jo, I.Y. Chung, H.W. Bae, J.S. Kim, S. Song, Y.H. Cho, N.C. Ha, Structural details of the OxyR peroxide-sensing mechanism. Proc Natl Acad Sci U S A 112 (2015) 6443-6448.

[29] R. Tyrrell, K.H. Verschueren, E.J. Dodson, G.N. Murshudov, C. Addy, A.J. Wilkinson, The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. Structure 5 (1997) 1017-1032.

[30] S. Chevalier, E. Bouffartigues, A. Bazire, A. Tahrioui, R. Duchesne, D. Tortuel, O. Maillot, T. Clamens, N. Orange, M.G.J.B.e.B.A.-G.R.M. Feuilloley, Extracytoplasmic function sigma factors in Pseudomonas aeruginosa. (2018).

[31] A. Lochowska, R. Iwanicka-Nowicka, D. Plochocka, M.M. Hryniewicz, Functional dissection of the LysR-type CysB transcriptional regulator Regions important for DNA binding, inducer response, oligomerization, and positive control. Journal of Biological Chemistry 276 (2001) 2098-2107.

[32] J. Ostrowski, N. Kredich, In vitro interactions of CysB protein with the cysJIH promoter of Salmonella typhimurium: inhibitory effects of sulfide. Journal of bacteriology 172 (1990) 779-785.

[33] A. Lochowska, R. Iwanicka-Nowicka, J. Zaim, M. Witkowska-Zimny, K. Bolewska, M. Hryniewicz, Identification of activating region (AR) of Escherichia coli LysR-type transcription factor CysB and CysB contact site on RNA polymerase alpha subunit at the cysP promoter. Molecular microbiology 53 (2004) 791-806.

[34] M.M. Hryniewicz, N.M. Kredich, The cysP promoter of Salmonella typhimurium: characterization of two binding sites for CysB protein, studies of in vivo transcription initiation, and demonstration of the anti-inducer effects of thiosulfate. Journal of bacteriology 173 (1991) 5876-5886.

[35] A. Kouzuma, T. Endoh, T. Omori, H. Nojiri, H. Yamane, H. Habe, Transcription factors CysB and SfnR constitute the hierarchical regulatory system for the sulfate starvation response in Pseudomonas putida. Journal of bacteriology 190 (2008) 4521-4531.

Chapter 5

DISCUSSION AND DISSERTATION SUMMARY

The bacterial regulatory network comprises various proteins, nucleic acids, and compounds that coordinate to endow on these organisms the ability to contend with various metabolic and environmental demands. Particularly, pathogenic bacteria rely on these network for disease pathogenesis and virulence. The gram-negative pathogenic pseudomonad, *P. aeruginosa*, is a stubborn pathogen whose infections are difficult to treat, owing to its harboring of exhaustive intrinsic and acquired antibiotic resistance mechanisms. The appearance of a high proportion of multidrug resistant (MDR) *P. aeruginosa* strains has in recent times earned the pathogen an addition to the list of 'superbugs', bacteria that are resistant to practically all antibacterials currently on the market [1, 2]. The *P. aeruginosa* antibiotic resistance arsenal comprises a well-coordinated network of systems incorporating various intrinsic, acquired, and adaptive resistance mechanisms that confer on the bacterium its notorious formidability [2] [3]. These abilities have been attributed to the extensiveness and complexity of the pathogen's regulatory and gene networks. The *P. aeruginosa* genome has been found to contain a large repertoire of regulatory proteins that contribute to its high metabolic and pathogenic versatility [4] [5] and it is no surprise that the pathogen contains the third largest known regulatory network in any bacteria [6]. Interestingly, despite being the focus of several genetic and biochemical studies, approximately one-third of the *P. aeruginosa* functional gene remains unknown [7]. To this end, the research described in this dissertation aims to make

contributions towards addressing some of the current unmet research needs regarding the *P. aeruginosa* transcriptional regulatory network.

One of these needs, the biochemical characterization of the *P. aeruginosa* RNA polymerase (RNAP), is the focus of **Chapter 2** of this dissertation. Being the primary protein that mediates the central dogma of biology, studies focusing on RNAP are important for understanding the bacterial transcription process. Particularly, RNAP-focused studies are desirable because the enzyme is a proven target for broad-spectrum antibacterial therapy [8-10]. The suitability of the bacterial RNAP as a target of broad-spectrum drug discovery stems from three main facts: the essentiality of the enzyme, the high conservation of RNAP subunits, and the lack of subunit conservation between prokaryotic and eukaryotic RNAPs. Despite these desirable attributes, only two antibiotic drugs are currently available that target bacterial transcription: rifamycin and fidaxomicin. Rifamycin was developed as an inhibitor for the *M. tuberculosis* RNAP and was also found to be effective against other bacterial RNAPs. However, subtle differences in the catalytic subunits of different RNAPs coupled with bacterial-mediated mutations in the catalytic region have resulted in reduced efficacy of the drug in recent times [8, 11-14]. Although various agents have been described that target RNAP activity [15-19], these studies have utilized the *E. coli* RNAP, one of only two bacterial RNAPs whose full-length structure is known (*T. thermophilus* being the other one). With the potential for the development of resistance, species-specific antibiotic drug discovery is now being considered the most ideal approach [20-24]. For targeting transcription via RNAP, this is especially important as RNAPs different bacteria are thought to be uniquely different and exhibit species-specific properties with regards to promoter recognition and their interactions in the transcription activation complex [25]. For instance, despite the similarities in the structures of the *E. coli,* and *T. thermophiles* RNAPs, differences have been identified in their initiation and elongation complexes as well as their

rate of activities [26]. Additionally, RNAP regions distant from the active site are quite divergent, and many species-specific insertions are located in the β and β′ subunits [27], complicating sequence alignments. Furthermore, regulatory strategies and auxiliary factors, which in many cases bind to divergent sites, appear to be quite distinct among *E. coli* vs. *T. thermophiles* species even though the global structures of the RNAPs from these two bacteria are similar [25]. Differences have also been observed with regards to inhibitors of different RNAPs. For instance, *Thermus* RNAPs are resistant to several inhibitors of the *E. coli* enzyme, including fidaxomycin, which was approved in 2011 for treatment of *Clostridium difficile* infections. Additionally, ligands that bind to both RNAPs could be such that the binding sites are far apart due to subtle structural differences [28, 29]. These facts further back the need for species-specific transcription studies in bacteria. Of the various tools and methods available for doing such species-specific transcription studies, structural analysis is indispensable in providing the atomic level of detail required for understanding of catalysis and design of antibiotics targeting the transcription apparatus. However, one major impediment for a successful structural analysis is the availability of a method for the production of the specific bacterial RNAP of interest.

For *P. aeruginosa*, we have contributed significantly in this regard by designing an approach for the recombinant production of the pathogen's RNAP as well as by providing important biochemical data on the enzyme. Although recombinant production is the most ideal method for high yield protein production, utilization of such an approach for RNAP is typically a difficult feat owning to the enzyme's high complexity and multi-subunit nature. This is the reason for which only four bacterial RNAP enzymes have been recombinantly produced and purified till date. Our described expression and purification approach ensures the production of highly pure, stable and active *P. aeruginosa* RNAP fractions. Importantly, our method allows the *in vivo* assembly of the complete

RNAP holoenzyme that precludes the need for further *in vitro* assembly and subsequent purification as described in other studies [30-33]. Moreover, our method also utilizes a room temperature autoinduction approach which ensures straightforward protein production that results in stable and well-behaved protein. We have also determined various activity and kinetic parameters for the *P. aeruginosa* RNAP that address a current knowledge gap regarding RNAP-mediated transcription in the bacterium. The elongation rate for the *P. aeruginosa* enzyme (~18 nt/s) was found to be slightly less compared to that of commercial *E. coli* RNAP (~19 nt/s), although both are within the broad range elongation rates as determined by [34] for six different bacterial RNAPs. The elongation rate for the *E. coli* RNAP as determined in this study is about 2 nt/s higher than what was reported by [34]. This could be attributed to differences in activity as a result of purity differences of the two preparations. Such a theory is supported by the fact that the study by [34] was done in the 1970s where current advances in recombinant protein expression and purification were unavailable. Our study also presents previously unavailable information regarding the activity differences in RNAP preparations assembled *in vitro* vs. the *in vivo* assembled enzyme. We found that the *in vitro*-assembled enzyme was twice less active and had an elongation rate that was ~2 nt/s less than the *in vivo*-assembled enzyme. This is a relevant finding as most available approaches for producing recombinant bacteria RNAP have utilized an *in vitro* reconstitution approach because of its relative simplicity and high success rate. We suspect that the differences observed can be attributed to the differences in the assembly conditions between the two approaches, a theory that would need further investigation to validate.

The successful production of the *P. aeruginosa* RNAP presents a multitude of opportunities for further research. Firstly, it enables structural characterization of RNAP-transcription factor interactions, which is the first step toward transcription-targeted drug discovery in the pathogen.

Because LysR-type Transcriptional Regulators (LTTRs) play important roles in *P. aeruginosa* disease pathogenesis, virulence, and antibiotic resistance, structural characterization of various RNAP-LTTR-DNA complex will present novel insights into common regulatory mechanisms that could be targeted for drug discovery. Particularly, it seems reasonable that groups of LTTRs will utilize common interfaces with RNAP to activate transcription. For instance, we have identified a conserved 'HPLA' amino acid motif in the effector binding domain of several LTTRs following sequence alignment that we suspect could play an essential role in RNAP contact and hence transcription activation. Interestingly, many of the LTTRs that were associated with this motif were found to be involved in fundamental metabolic pathways. Therefore, this cannot be explained simply as a structural feature. The availability of the *P. aeruginosa* RNAP also affords the opportunity to do *in vitro* transcription regulation studies. Such studies will involve the use of transcription assays featuring RNAP, promoters of target genes, respective transcription factors, and potential ligands/inhibitors. In so doing, the assay can be applied to high-throughput screens for evaluation of compounds as inhibitors of RNAP activation on the genes of interest. Similar high-throughput screening approaches have been described [35-39] but these did not incorporate transcription factors (e.g. LTTRs) and therefore run into the challenge of missing inhibitors that block transcription factor-mediated transcription activation. In this study, we have also designed plasmids for the production of just the RNAP core enzyme. Production of the separate core enzyme in this manner allows the *in vitro* reconstitution of the holoenzyme fractions of any bacterial sigma factor. The holoenzyme enzyme preparations of specific *P. aeruginosa* sigma factors allow an investigation of transcriptional specificity on promoters recognized by the sigma factors. In one such study, we assembled the core enzyme with sigma 70 as well as separately with two *P. aeruginosa* alternative sigma factors (RpoS and AlgU). Then, a DNA fragment that contained RpoS, AlgU, and RpoD recognition sites such that

the different promoters had different strengths were used. The generation of the DNA fragment by PCR was based off a study by [40] that found that AlgU contributed to posttranscriptional activity by increasing the expression of *rsmA*, a posttranscriptional regulator that controls virulence factor production and biofilm formation in *P. aeruginosa*. A search in the intergenic region upstream of *rsmA* found the consensus sequences for AlgU (GAACTT-16/17N-TCTGA), RpoS (CTATATC), and RpoD (TTGACA-17N-TATAAT). A 471 bp DNA fragment containing these putative binding sites was then generated and used as template for transcription assays. The RNAP holoenzymes generated for each respective sigma factor were utilized in the assays and the synthesized RNA was extracted and visualized on gels (**Appendix Figure E-2**). The results showed the selective transcription of the same *rsmA* gene target by the different sigma factors in a manner corresponding to the different promoter strengths. This was a proof-of-concept study that demonstrated the utilization of RNAP for transcription specificity studies. The experiments therefore set the foundation for future studies that aim to further probe transcriptional regulation in *P. aeruginosa*. The successful *in vitro* assembly of RNAP holoenzyme of specific *P. aeruginosa* sigma factors also set the foundation for future studies that will aim at structural characterization of core RNAP interactions with the different sigma factor types. Such studies are particularly essential because interaction with sigma factors is one aspect of transcriptional regulation which we suspect is strictly species-specific. For instance, *E. coli* contains only 6 sigma factors in comparison to the 24 in *P. aeruginosa* [41]. It is therefore conceivable that specific structural contacts between RNAP core enzyme and specific *P. aeruginosa* sigma factors are absence in *E. coli*. Furthermore, structural characterization of specific sigma factor interaction with RNAP core enzyme will provide insights to potential inhibitors that could block RNAP holoenzyme assembly. Such compounds have the potential to be developed into potent therapeutic agents that target global transcription in the

pathogen. For instance, the SB series of compounds was shown to inhibit RNAP holoenzyme formation by [39]; however, the assay was based on the *E. coli* RNAP while only sigma 70 inhibition was assessed. If our theory that RNAP core enzyme-sigma factor interactions are very species-specific is true, then it is possible that these compounds will be ineffective against *P. aeruginosa* strains. In preliminary studies, we have produced recombinantly, the seven currently characterized *P. aeruginosa* sigma factors (**Appendix Figure E-1**) and showed successful *in vitro* assembly with the core enzyme. This paves the way for structural characterization of the different RNAP core enzyme–sigma factor complexes. Finally, we have also utilized heterologous *P. aeruginosa* RNAP core enzyme–sigma factor preparations for the *in vitro* production of small regulatory bacteria RNAs for structural characterization (Nicole Laniohan Dissertation, 2019). This novel approach utilizes plasmid constructs containing consensus promoter sequences for the specific sigma factors of interest for *in vitro* transcription assays. This way, we have circumvented the use of the widely described T7-RNAP system [42-46] that is fraught with various problems [47].

Chapter 3 and **Chapter 4** of this dissertation focused on probing the regulatory role of two *P. aeruginosa* LTTRs. Out of the 125 putative LTTRs we identified in the *P. aeruginosa* genome, 25 have currently been characterized and they were found to be involved broadly in metabolic processes, virulence and biofilm formation, oxidative stress, or antibiotic resistance. *P. aeruginosa* LTTRs therefore make ideal therapeutic targets because these are processes that are relied on by the pathogen for disease pathogenesis and antibiotic resistance. On one hand, targeting the essential LTTRs (mostly the ones involved in metabolic processes) will come with fatal consequences for the cells. However, on the other hand, targeting the LTTRs that are involved in non-essential processes like virulence, antibiotic resistance, and oxidative stress could simply take away the ability of the bacterium to cause infections without necessarily killing the cells. This approach of "disarming"

bacteria is being proposed as the most ideal antibiotic drug discovery approach because it does not put a selective pressure on the bacteria to develop resistance [1, 48-50]. Currently, only one *P. aeruginosa* LTTR, PqsR, has been thoroughly explored for potential inhibitors in this manner [50-55], underscoring how largely unexplored the *P. aeruginosa* LTTR-focused drug discovery research field is. FinR and CysB, the two LTTRs of focus in **Chapters 3 & 4** of this dissertation, were both found to have properties that make them ideal targets for drug discovery. FinR, implicated originally in the *P. aeruginosa* oxidative stress response [56], is a classically-behaved LTTR that is divergently expressed from its target gene, *fprA.* We identified putative FinR binding sites in the *finR–fprA* intergenic region that suggested binding of the LTTR to the region and that also alluded to a 'sliding dimer' regulatory mechanism. Using gel shift assays and transcription assays, we showed that FinR does indeed regulate *fprA* expression, with sulfite as its inducer. These studies with FinR have identified a previously unreported role for the LTTR: its involvement in the *P. aeruginosa* sulfur assimilation pathway, hence expanding current knowledge regarding the regulatory role of the protein. No other studies have reported the direct involvement of FinR in the sulfur assimilation pathway nor elucidated the potential regulatory mechanisms involved in FinR-dependent, sulfite-induced FprA expression in pseudomonads, although some studies have indirectly implicated FprA in sulfur assimilation in *P. putida* [57-60]. We speculate that all pseudomonads will undergo a similar regulatory mechanism as observed here for *P. aeruginosa*. Two facts back this speculation: our finding of a very highly conserved *finR–fprA* genetic region among pseudomonads, and our finding of highly conserved FinR boxes in the intergenic region of all pseudomonads. The regulatory mechanism also seems to be restricted to only pseudomonads as studies in our lab (unpublished) have found no binding of the *A. baylyi* and *E. coli* FinRs to the *finR–fprA* gene region. Interestingly however, our solved X-ray structure of the *A. baylyi* full-length FinR protein (unpublished) reveals

an oligomerization scheme that supports the regulatory mechanism proposed here for the *P. aeruginosa* FinR. The structure also supports our finding of sulfite as the inducer for FinR because a binding site for the metabolite was identified in the regulatory domain. Sulfite as the inducer for FinR is logical not just because it is the substrate of FprA, the gene regulated by FinR, but also because it is the one metabolite in the sulfur pathway that could function as an oxidative stressor. Sulfite is toxic to bacteria beyond certain concentrations and it could be the link between the oxidative stress regulatory role for FinR as demonstrated by [56, 58, 61] and the LTTR's role in sulfur assimilation as demonstrated in this study. The need to immediately eliminate any accumulated sulfite in cells might explain why FprA is an essential enzyme in *P. aeruginosa* as found by [56]. Although FinR was found to be nonessential, its absence was found to increase susceptibility of the cells to oxidative damage, which led to reduced virulence in a Drosophila disease model [56]. FinR is therefore a viable drug discovery target whose targeting could make *P. aeruginosa* cells more susceptible to reactive oxygen species (ROS)-mediated immune attack during human infection. Potential drugs targeting FinR could therefore be used as adjuvants or in combination therapies to reduce *P. aeruginosa* virulence and drug resistance. Targeting of FinR, which will result in a repression of FprA expression, will also likely deprive the cells of the ability to synthesize L-cysteine, an essential amino acid. However, pseudomonads express another sulfite reductase, CysI that could provide a compensatory pathway for sulfite reduction. That said, we also suspect that FprA acts to donate electrons to CysI; therefore, repression of FprA expression through FinR targeting could deprive CysI of the electrons needed for sulfite reduction. To investigate these further, a *P. aeruginosa finR* knockout mutant will need to be created and assessed for *cysI* expression. Additionally, growing the *finR* mutant on different oxidized and reduced sulfur sources will present important regulatory insights. In related experiments in the lab, FinR is being used in

pull-down experiments to extract potential inhibitors from plant extracts, while another study is assessing the potential of various isolated compounds to inhibit FinR-dependent *fprA* expression. To complement the biochemical data generated for the *P. aeruginosa* FinR as described here, structural information data will be needed to provide a big picture understanding of FinR-dependent, sulfite-induced *fprA* expression, albeit multiple attempts by us to crystallize the *P. aeruginosa* FinR full-length protein have proved futile. However, the study has laid the foundation for cyro electron microscopy (cryo-EM) structural characterization of the FinR–*fprA*–RNAP complex, the ultimate goal of the project. Since no current structure exists of a ternary RNAP-DNA-LTTR complex, such a structure is bound to present highly significant previously unknown insights to LTTR-RNAP interactions. Genetic studies have shown that LTTRs directly contact either the C-terminal domain of the α-subunit of RNAP (αCTD) as shown in MetR [62] and CysB [63] or the σ70 subunit as demonstrated in OxyR [64]. Notwithstanding this, the type of RNAP interactions with an individual LTTR may differ for distinct promoter regions. An RNAP-DNA-LTTR structure would therefore help to fill some of these knowledge gaps and expand current understanding of TF-mediated transcriptional regulation in bacteria. Preceding any cryo-EM experiments would be an exploration of the formation of a stable RNAP-DNA-LTTR ternary complex. We have demonstrated the formation of such a stable complex by gel shift assays using a DNA fragment containing all putative FinR binding sites as identified from this study (**Appendix Table A5-1**). With stable complex formation, negative EM data can be obtained that will provide the foundation for cryo-EM experiments.

Unlike FinR, the *P. aeruginosa* CysB was found to be a non-classical LTTR that was not divergently expressed from any gene in its neighborhood. Furthermore, the genetic arrangement in the *cysB* region was not conserved among pseudomonads nor in other bacteria. Using a putative

CysB binding site as proposed by [65], we identify as many as 25 putative CysB-co-regulated genes. These findings suggested that CysB could be a global regulator with a broad regulon in *P. aeruginosa*. Indeed, this theory is supported by the fact that at least 5 reported studies have implicated the LTTR in various processes in the bacteria [65-70]. Additionally, CysB is regarded as the master regulator of sulfur metabolism in bacteria and it has been predicted to regulate all the genes in the cysteine biosynthesis pathway [71]. According to [72], a transcription factor can be classified as global if, 1) it demonstrates the ability to regulate genes belonging to target groups of different sigma factors, 2) it demonstrates the ability to regulate genes with different functional classifications, 3) it has the potential to respond to a diverse range of both environmental and exogenous stimuli, and 4) it has multiple co-regulating transcription factors. Findings from this study coupled with findings from other CysB-focused studies [65-70, 73] suggest that the transcription factor is associated with at least two of these characteristics. Many of the putative CysB-co-regulated genes we found are predicted to be involved in various virulent processes, suggesting that the LTTR likely plays an important role in disease progression and lifestyle choice of *P. aeruginosa*. This also backs CysB as a suitable drug discovery target. Targeting a global regulator like CysB could potentially lead to the dysregulation of a plethora of genes, particularly those involved in virulence. In gel shift assays, CysB was found to form multiple complexes with each DNA fragment in a manner that suggested promiscuity in DNA binding, an unusual characteristic for LTTRs. This finding is consistent with an X-ray structure we have obtained for the *A. baylyi* CysB (unpublished) that reveals an unusual 'donut' shaped LTTR structure. Interestingly, neither the effector binding domain X-ray structure of the *Klebsiella aerogenes* CysB [74] nor the DNA binding domain X-ray structure of the *P. aeruginosa* CysB [67] conforms to our *A. baylyi* structure. Nonetheless, the unusual gel shift data together with the structural data backs the theory that CysB is a global regulator that must have

structural features to accommodate different, functionally unrelated genes. In such a case, the LTTR also likely relies on a host of other transcription factors, sigma factors, regulatory RNAs, other regulatory proteins, and ligands to ensure transcriptional control on different promoters. Of the bacterial LTTRs that have been structurally characterized, none has been predicted to be a global regulator like CysB and so no structures exist that support the biochemical findings observed here. It is worth mentioning that we attempted multiple times to crystallize the *P. aeruginosa* CysB (produced under both sulfur limiting and non-limiting conditions) but were unsuccessful in this regard. Although the *P. aeruginosa* OxyR, an LTTR regarded as the master regulator of oxidative stress in bacteria, has been considered as a somewhat global regulator, the full-length OxyR structure as solved by [75] was found to conform to classic LTTR structural features.

In this study, we utilized transcription assays to further probe the regulatory role of CysB on the putative co-regulated genes we identified. Such transcription regulation studies were only possible because of our successful production of the *P. aeruginosa* RNAP. Classically, studies have utilized commercial *E. coli* RNAP in such assays; however, unless the genes being studies are of *E. coli* origin, such an assay will likely not give trustworthy data. This is especially highlighted by the facts described earlier in this Chapter that support the existence of species-specific differences in transcriptional regulation in different organisms. In the case of *P. aeruginosa*, only two studies were identified in the literature that utilized RNAP from the bacterium for transcription experiments [76, 77]. In both studies, RNAP from cell extracts were utilized. However, such an approach is not ideal for selective transcription as the cell extract might contain many confounding elements. Additionally, not every laboratory will have the capacity to grow the pathogenic *P. aeruginosa* in the lab. Therefore, our successful description of a method for producing highly pure and active fractions of RNAP enzyme addresses these concerns. Using our transcription assays, we assessed

the regulatory role of CysB on the different putative co-regulated genes and tested the potential of several ligands as CysB inducers. The ligands, sodium sulfite, sodium sulfate, sodium thiosulfate, N-acetyl serine, O-acetyl serine, and adenosine 5′-phosphosulfate are various metabolites in the sulfur assimilation pathway. From our assays, some of the genes were found to possibly be under CysB regulation, while others were not. Because CysB is suspected to be a global regulator, it is conceivable that certain regulatory elements are required on some of the promoters that were found in transcription assays to not be under CysB regulation. For instance, some alternative sigma factors like AlgU and PvdS might be involved respectively in CysB-mediated regulation of AlgD and MvfR, two genes that were found to not be under CysB regulation in this study but for which at least one study each has implied otherwise [65, 70]. In our CysB studies, we also identified three potential inducers for the protein: sulfite, sulfate, and thiosulfate. This further supports CysB as a global regulator as the ability to be modulated by a different ligands and other stimuli is an important characteristic of a global transcriptional regulator. To further validate the global regulatory role of CysB, a *cysB* knockout mutant must be created and assessed for the expression and regulation of the various putative co-regulated genes identified here. Growing the mutant on various sulfur source will also allow an assessment of potential CysB inducers. However, because *cysB* is likely an essential gene, a knock out mutant might be challenging to create, the likely reason for which no such studies have been reported till date. Furthermore, additional transcriptional regulation experiments can be done with specific alternative sigma factors and other putative regulatory proteins to rule out genes that are definitively not under CysB regulation. Ultimately, structural characterization of the CysB-RNAP-DNA complex by cryo-EM as described for FinR above will present in-depth insights into the LTTR's regulatory role and offer suggestions for targeting the protein for drug discovery.

Overall, the research described in this dissertation addresses various unmet needs in *P. aeruginosa* regulatory network research and provides an important blueprint and foundation for probing transcriptional regulation in the pathogen. Ultimately, the insights presented here, coupled with data from future experiments, will provide novel ideas into species-specific targeting of the *P. aeruginosa* regulatory network, with the goal to develop potent therapeutic agents to control this infectious and problematic pathogen.

REFERENCES

[1] Z. Pang, R. Raudonis, B.R. Glick, T.-J. Lin, Z. Cheng, Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies. Biotechnology advances (2018).

[2] E.B. Breidenstein, C. de la Fuente-Núñez, R.E. Hancock, Pseudomonas aeruginosa: all roads lead to resistance. Trends in microbiology 19 (2011) 419-426.

[3] P.A. Lambert, Mechanisms of antibiotic resistance in Pseudomonas aeruginosa. J R Soc Med 95 Suppl 41 (2002) 22-26.

[4] D. Balasubramanian, S.K. Murugapiran, E. Silva-Herzog, L. SCHNEGER, X. Yang, G. Tatke, G. Narasimhan, K. Mathee, Transcriptional regulatory network in Pseudomonas aeruginosa, Caiser Academic Press, United Kingdom, 2013.

[5] G.L. Winsor, E.J. Griffiths, R. Lo, B.K. Dhillon, J.A. Shay, F.S. Brinkman, Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. Nucleic acids research 44 (2016) D646-D653.

[6] E. Galán-Vásquez, B. Luna, A. Martínez-Antonio, The regulatory network of Pseudomonas aeruginosa. Microbial informatics and experimentation 1 (2011) 3.

[7] S. Hwang, C.Y. Kim, S.-G. Ji, J. Go, H. Kim, S. Yang, H.J. Kim, A. Cho, S.S. Yoon, I. Lee, Network-assisted investigation of virulence and antibiotic-resistance systems in Pseudomonas aeruginosa. Scientific reports 6 (2016) 26223.

[8] I. Chopra, Bacterial RNA polymerase: a promising target for the discovery of new antimicrobial agents. Current opinion in investigational drugs (London, England: 2000) 8 (2007) 600-607.

[9] P. Villain-Guillot, L. Bastide, M. Gualtieri, J.-P. Leonetti, Progress in targeting bacterial transcription. Drug discovery today 12 (2007) 200-208.

[10] R. Mariani, S. Maffioli, Bacterial RNA polymerase inhibitors: an organized overview of their structure, derivatives, biological activity and current clinical development status. Current medicinal chemistry 16 (2009) 430-454.

[11] H.G. Floss, T.-W. Yu, Rifamycin mode of action, resistance, and biosynthesis. Chemical reviews 105 (2005) 621-632.

[12] E.A. Campbell, N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb, S.A. Darst, Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. Cell 104 (2001) 901-912.

[13] A. Bryskier, Anti-MRSA agents: under investigation, in the exploratory phase and clinically available. Expert review of anti-infective therapy 3 (2005) 505-553.

[14] S.A. Darst, New inhibitors targeting bacterial RNA polymerase. Trends in biochemical sciences 29 (2004) 159-162.

[15] C. Ma, X. Yang, P.J. Lewis, Bacterial transcription as a target for antibacterial drug development. Microbiol Mol Biol Rev 80 (2016) 139-160.

[16] A. Srivastava, M. Talaue, S. Liu, D. Degen, R.Y. Ebright, E. Sineva, A. Chakraborty, S.Y. Druzhinin, S. Chatterjee, J. Mukhopadhyay, New target for inhibition of bacterial RNA polymerase:'switch region'. Current opinion in microbiology 14 (2011) 532-543.

[17] C. Ma, X. Yang, P.J. Lewis, Bacterial transcription inhibitor of RNA polymerase holoenzyme formation by structure-based drug design: from in silico screening to validation. ACS infectious diseases 2 (2015) 39-46.

[18] R.H. Ebright, Novel RNA polymerase inhibitor found in soil extracts provides hope for future antibacterial drugs. Future medicinal chemistry 9 (2017) 1857-1861.

[19] W. Zhu, J. Haupenthal, M. Groh, M. Fountain, R.W. Hartmann, New insights into the bacterial RNA polymerase inhibitor CBR703 as a starting point for optimization as an anti-infective agent. Antimicrobial agents and chemotherapy 58 (2014) 4242-4245.

[20] C.-J. Lee, X. Liang, X. Chen, D. Zeng, S.H. Joo, H.S. Chung, A.W. Barb, S.M. Swanson, R.A. Nicholas, Y. Li, Species-specific and inhibitor-dependent conformations of LpxC: implications for antibiotic design. Chemistry & biology 18 (2011) 38-47.

[21] R.L. Then, H.-G. Sahl, Anti-infective strategies of the future: is there room for species-specific antibacterial agents? Current pharmaceutical design 16 (2010) 555-566.

[22] T. Maxson, D.A. Mitchell, Targeted treatment for bacterial infections: prospects for pathogen-specific antibiotics coupled with rapid diagnostics. Tetrahedron 72 (2016) 3609.

[23] R.J. Melander, D.V. Zurawski, C. Melander, Narrow-spectrum antibacterial agents. MedChemComm 9 (2018) 12-21.

[24] J.A. Shapiro, A.R. Kaplan, W.M. Wuest, From General to Specific: Can Pseudomonas Primary Metabolism Be Exploited for Narrow-Spectrum Antibiotics? ChemBioChem 20 (2019) 34-39.

[25] V. Svetlov, I. Artsimovitch, Purification of bacterial RNA polymerase: tools and protocols, Bacterial Transcriptional Control, Springer, 2015, pp. 13-29.

[26] K.S. Murakami, Structural biology of bacterial RNA polymerase. Biomolecules 5 (2015) 848-864.

[27] W.J. Lane, S.A. Darst, Molecular evolution of multisubunit RNA polymerases: structural analysis. J Mol Biol 395 (2010) 686-704.

[28] Y. Zuo, Y. Wang, T.A. Steitz, The mechanism of E. coli RNA polymerase regulation by ppGpp is suggested by the structure of their complex. Mol Cell 50 (2013) 430-436.

[29] W. Ross, C.E. Vrentas, P. Sanchez-Vazquez, T. Gaal, R.L. Gourse, The magic spot: a ppGpp binding site on E. coli RNA polymerase responsible for regulation of transcription initiation. Mol Cell 50 (2013) 420-429.

[30] S. Borukhov, A. Goldfarb, Recombinant Escherichia coli RNA polymerase: purification of individually overexpressed subunits and in vitro assembly. Protein Expr Purif 4 (1993) 503-511.

[31] A. China, V. Nagaraja, Purification of RNA polymerase from mycobacteria for optimized promoter-polymerase interactions. Protein Expr Purif 69 (2010) 235-242.

[32] X. Yang, P.J. Lewis, Overproduction and purification of recombinant Bacillus subtilis RNA polymerase. Protein Expr Purif 59 (2008) 86-93.

[33] R. Banerjee, P. Rudra, R.K. Prajapati, S. Sengupta, J. Mukhopadhyay, Optimization of recombinant Mycobacterium tuberculosis RNA polymerase expression and purification. Tuberculosis (Edinb) 94 (2014) 397-404.

[34] M.J. Chamberlin, W.C. Nierman, J. Wiggs, N. Neff, A quantitative assay for bacterial RNA polymerases. J Biol Chem 254 (1979) 10061-10069.

[35] N.T. Scharf, V. Molodtsov, A. Kontos, K.S. Murakami, G.A. Garcia, Novel Chemical Scaffolds for Inhibition of Rifamycin-Resistant RNA Polymerase Discovered from High-Throughput Screening. SLAS Discov 22 (2017) 287-297.

[36] S.A. El-Mowafi, E. Sineva, J.N. Alumasa, H. Nicoloff, J.W. Tomsho, S.E. Ades, K.C. Keiler, Identification of inhibitors of a bacterial sigma factor using a new high-throughput screening assay. Antimicrob Agents Chemother 59 (2015) 193-205.

[37] V. Bergendahl, T. Heyduk, R.R. Burgess, Luminescence resonance energy transfer-based high-throughput screening assay for inhibitors of essential protein-protein interactions in bacterial RNA polymerase. Appl Environ Microbiol 69 (2003) 1492-1498.

[38] G. Campagnola, P. Gong, O.B. Peersen, High-throughput screening identification of poliovirus RNA-dependent RNA polymerase inhibitors. Antiviral Res 91 (2011) 241-251.

[39] E. Andre, L. Bastide, S. Michaux-Charachon, A. Gouby, P. Villain-Guillot, J. Latouche, A. Bouchet, M. Gualtieri, J.P. Leonetti, Novel synthetic molecules targeting the bacterial RNA polymerase assembly. J Antimicrob Chemother 57 (2006) 245-251.

[40] S.D. Stacey, C.L. Pritchett, Pseudomonas aeruginosa AlgU contributes to posttranscriptional activity by increasing rsmA expression in a mucA22 strain. Journal of bacteriology 198 (2016) 1812-1826.

[41] S. Schulz, D. Eckweiler, A. Bielecka, T. Nicolai, R. Franke, A. Dötsch, K. Hornischer, S. Bruchmann, J. Düvel, S. Häussler, Elucidation of sigma factor-associated networks in Pseudomonas aeruginosa reveals a modular architecture with limited and function-specific crosstalk. PLoS pathogens 11 (2015) e1004744.

[42] V. Guerniou, D. Gasparutto, T. Douki, J. Cadet, S. Sauvaigo, Enhancement of the in vitro transcription by T7 RNA polymerase of short DNA templates containing oxidative thymine lesions. C R Biol 328 (2005) 794-801.

[43] A.A. Elov, E.M. Volkov, T.G. Reintamm, T.S. Oretskaia, Z.A. Shabarova, [RNA synthesis using T7 phage RNA polymerase: transcription of synthetic DNA templates in solution and on polymer support]. Bioorg Khim 15 (1989) 159-165.

[44] S.T. Nath, L.J. Romano, Transcription by T7 RNA polymerase using benzo[a]pyrene-modified templates. Carcinogenesis 12 (1991) 973-976.

[45] J.A. Pleiss, M.L. Derrick, O.C. Uhlenbeck, T7 RNA polymerase produces 5' end heterogeneity during in vitro transcription from certain templates. RNA 4 (1998) 1313-1317.

[46] F.J. Triana-Alonso, M. Dabrowski, J. Wadzack, K.H. Nierhaus, Self-coded 3'-extension of run-off transcripts produces aberrant products during in vitro transcription with T7 RNA polymerase. J Biol Chem 270 (1995) 6298-6307.

[47] R.T. Batey, Advances in methods for native expression and purification of RNA for structural studies. Curr Opin Struct Biol 26 (2014) 1-8.

[48] R.C. Allen, R. Popat, S.P. Diggle, S.P. Brown, Targeting virulence: can we make evolution-proof drugs? Nature Reviews Microbiology 12 (2014) 300.

[49] M. Totsika, Benefits and challenges of antivirulence antimicrobials at the dawn of the post-antibiotic era. Drug Delivery Letters 6 (2016) 30-37.

[50] S.W. Dickey, G.Y. Cheung, M. Otto, Different drugs for bad bugs: antivirulence strategies in the age of antibiotic resistance. Nature reviews Drug discovery 16 (2017) 457.

[51] T. Klein, C. Henn, J.C. de Jong, C. Zimmer, B. Kirsch, C.K. Maurer, D. Pistorius, R. Muller, A. Steinbach, R.W. Hartmann, Identification of small-molecule antagonists of the Pseudomonas aeruginosa transcriptional regulator PqsR: biophysically guided hit discovery and optimization. ACS Chem Biol 7 (2012) 1496-1501.

[52] C. Lu, B. Kirsch, C. Zimmer, J.C. de Jong, C. Henn, C.K. Maurer, M. Musken, S. Haussler, A. Steinbach, R.W. Hartmann, Discovery of antagonists of PqsR, a key player in 2-alkyl-4-quinolone-dependent quorum sensing in Pseudomonas aeruginosa. Chem Biol 19 (2012) 381-390.

[53] F. Soukarieh, P. Williams, M.J. Stocks, M. Camara, Pseudomonas aeruginosa quorum sensing systems as drug discovery targets: current position and future perspectives. Journal of medicinal chemistry 61 (2018) 10385-10402.

[54] M. Zender, T. Klein, C. Henn, B. Kirsch, C.K. Maurer, D. Kail, C. Ritter, O. Dolezal, A. Steinbach, R.W. Hartmann, Discovery and biophysical characterization of 2-amino-oxadiazoles

as novel antagonists of PqsR, an important regulator of Pseudomonas aeruginosa virulence. J Med Chem 56 (2013) 6761-6774.

[55] C. Lu, C.K. Maurer, B. Kirsch, A. Steinbach, R.W. Hartmann, Overcoming the unexpected functional inversion of a PqsR antagonist in Pseudomonas aeruginosa: an in vivo potent antivirulence agent targeting pqs quorum sensing. Angew Chem Int Ed Engl 53 (2014) 1109-1112.

[56] S. Boonma, A. Romsang, J. Duang-Nkern, S. Atichartpongkul, W. Trinachartvanit, P. Vattanaviboon, S. Mongkolsuk, The FinR-regulated essential gene fprA, encoding ferredoxin NADP+ reductase: Roles in superoxide-mediated stress protection and virulence of Pseudomonas aeruginosa. PLoS One 12 (2017) e0172071.

[57] J. Kim, W. Park, Oxidative stress response in Pseudomonas putida. Applied microbiology and biotechnology 98 (2014) 6933-6946.

[58] S. Yeom, J. Yeom, W. Park, Molecular characterization of FinR, a novel redox-sensing transcriptional regulator in Pseudomonas putida KT2440. Microbiology 156 (2010) 1487-1496.

[59] J. Yeom, C.O. Jeon, E.L. Madsen, W. Park, In vitro and in vivo interactions of ferredoxin-NADP+ reductases in Pseudomonas putida. Journal of biochemistry 145 (2009) 481-491.

[60] Y. Lee, S. Peña-Llopis, Y.-S. Kang, H.-D. Shin, B. Demple, E.L. Madsen, C.O. Jeon, W. Park, Expression analysis of the fpr (ferredoxin-NADP+ reductase) gene in Pseudomonas putida KT2440. Biochemical biophysical research communications 339 (2006) 1246-1254.

[61] T.A. Lewis, A. Glassing, J. Harper, M.J. Franklin, Role for ferredoxin:NAD(P)H oxidoreductase (FprA) in sulfate assimilation and siderophore biosynthesis in Pseudomonads. J Bacteriol 195 (2013) 3876-3887.

[62] P.S. Fritsch, M.L. Urbanowski, G.V. Stauffer, Role of the RNA polymerase α subunits in MetR-dependent activation of metE and metH: important residues in the C-terminal domain and orientation requirements within RNA polymerase. Journal of bacteriology 182 (2000) 5539-5550.

[63] A. Lochowska, R. Iwanicka-Nowicka, J. Zaim, M. Witkowska-Zimny, K. Bolewska, M. Hryniewicz, Identification of activating region (AR) of Escherichia coli LysR-type transcription

factor CysB and CysB contact site on RNA polymerase alpha subunit at the cysP promoter. Molecular microbiology 53 (2004) 791-806.

[64] X. Wang, P. Mukhopadhyay, M.J. Wood, F.W. Outten, J.A. Opdyke, G. Storz, Mutational analysis to define an activating region on the redox-sensitive transcriptional regulator OxyR. Journal of bacteriology 188 (2006) 8335-8342.

[65] F. Imperi, F. Tiburzi, G.M. Fimia, P. Visca, Transcriptional control of the pvdS iron starvation sigma factor gene by the master regulator of sulfur metabolism CysB in Pseudomonas aeruginosa. Environ Microbiol 12 (2010) 1630-1642.

[66] J. Hummerjohann, S. Laudenbach, J. Retey, T. Leisinger, M.A. Kertesz, The sulfur-regulated arylsulfatase gene cluster of Pseudomonas aeruginosa, a new member of the cys regulon. J Bacteriol 182 (2000) 2055-2058.

[67] Y. Song, C. Yang, G. Chen, Y. Zhang, Z. Seng, Z. Cai, C. Zhang, L. Yang, J. Gan, L. Haihua, Molecular insights into the master regulator CysB-mediated bacterial virulence in Pseudomonas aeruginosa. Molecular microbiology 0 (2019).

[68] J.M. Farrow, 3rd, E.C. Pesci, Distal and proximal promoters co-regulate pqsR expression in Pseudomonas aeruginosa. Mol Microbiol 104 (2017) 78-91.

[69] J.M. Farrow, 3rd, L.L. Hudson, G. Wells, J.P. Coleman, E.C. Pesci, CysB Negatively Affects the Transcription of pqsR and Pseudomonas Quinolone Signal Production in Pseudomonas aeruginosa. J Bacteriol 197 (2015) 1988-2002.

[70] I. Delic-Attree, B. Toussaint, J. Garin, P.M. Vignais, Cloning, sequence and mutagenesis of the structural gene of Pseudomonas aeruginosa CysB, which can activate algD transcription. Mol Microbiol 24 (1997) 1275-1284.

[71] N. Kredich, Biosynthesis of Cysteine. (2008).

[72] A. Martınez-Antonio, J.J.C.o.i.m. Collado-Vides, Identifying global regulators in transcriptional regulatory networks in bacteria. 6 (2003) 482-489.

[73] A. Kouzuma, T. Endoh, T. Omori, H. Nojiri, H. Yamane, H. Habe, Transcription factors CysB and SfnR constitute the hierarchical regulatory system for the sulfate starvation response in Pseudomonas putida. Journal of bacteriology 190 (2008) 4521-4531.

[74] R. Tyrrell, K.H. Verschueren, E.J. Dodson, G.N. Murshudov, C. Addy, A.J. Wilkinson, The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. Structure 5 (1997) 1017-1032.

[75] I. Jo, I.Y. Chung, H.W. Bae, J.S. Kim, S. Song, Y.H. Cho, N.C. Ha, Structural details of the OxyR peroxide-sensing mechanism. Proc Natl Acad Sci U S A 112 (2015) 6443-6448.

[76] J.G. Gao, G.N. Gussin, RNA polymerases from Pseudomonas aeruginosa and Pseudomonas syringae respond to Escherichia coli activator proteins. J Bacteriol 173 (1991) 394-397.

[77] C.A. Vakulskas, K.M. Brady, T.L. Yahr, Mechanism of transcriptional activation by Pseudomonas aeruginosa ExsA. J Bacteriol 191 (2009) 6654-6664.

APPENDIX

Appendix A (Chapter 2)

Table A2-1. Sequencing primers used to confirm constructs*

	Primer name	Sequence
RpoBC	PAO1_rpoBC_seg2421-F	ACTTTAAGAAGGAGATATACCATGGC
	PAO1_rpoBC_seg4036-R	TTGTGGGTGATCTCGGAAAGC
	PAO1_rpoBC_seg3985-F	AGCTGTCGCAGTTCATGGACC
	PAO1_rpoBC_seg5603-R	CTGCAGCTTGCGCTTCTTGTC
	PAO1_rpoBC_seg5517-F	GAACAGCTGGAGAAGGCCC
	PAO1_rpoBC_seg7149-R	TTGAGCAGTTCGTGAACAGC
	PAO1_rpoBC_seg7096-F	CGGTGACGATTTCGATGCTC
	PAO1_rpoBC_seg8815-R	GCCACGCATACCTGCCAGC
	PAO1_rpoBC_seg8748-F	ACATGATGGCTGACTCGGGTG
	PAO1_rpoBC_seg10810-R	CGCAAGCTTGTCGACGGAG
	RpoBC.pRARE2.seq360-F	CCATCAACTACCGTACCTTCAAGCCG
	RpoBC.pRARE2.seq2160-F	CTACTCGACCATTTCCGGCGTGTCC
	RpoBC.pRARE2.seq3840-F	GCTGATTCCGAAGTGGCGTCACCTGAACGTG
	RpoBC.pRARE2.seq10560-F	CGTTCATCTCCGACACCCTGAAGATCG
	RpoBC.pRARE2.seq12240-F	GTCATCGACGTTCAGGTCTTCACCCGC
RpoAZD	PAO1_rpoAZD_seg5943-F	CCGAACCTGGGCAAGAAGTCC
	PAO1_rpoAZD_seg6712-F	GGTGGAAAGCGACATCGGTC
	PAO1_rpoAZD_seg7482-F	GCAGAAGCTGGCGGCCCTG
	DuetUP2	TTGTACACGGCCGCATAATC
	DuetDOWN1	GATTATGCGGCCGTGTACAA

ACYCDuetUP1	GGATCTCGACGCTCTCCCT
T7term	GCTAGTTATTGCTCAGCGG

Table A2-2. List of rare codons in the *P. aeruginosa rpo* genes and their abundance*

Codon	Codes as	*E. coli* Usage frequency (%)	Count				
			RpoA	RpoB	RpoC	RpoD	RpoZ
CGG	Arg	4.1	0	4	5	3	0
CGA	Arg	4.3	0	0	0	2	0
TCC	Ser	5.5	6	25	30	13	1
TCT	Ser	5.7	0	0	1	1	0
CCC	Pro	6.4	3	3	3	4	0
ACA	Thr	6.4	0	0	0	1	0
AGT	Ser	7.2	3	0	1	0	0
CCA	Pro	6.6	0	1	0	0	0
TCA	Ser	7.8	0	1	0	0	0
TGC	Cys	8	2	10	12	1	0
ACT	Thr	8	3	5	1	2	0
TCG	Ser	8	5	17	23	6	1
CCT	Pro	8.4	3	2	2	1	0
GGG	Gly	8.6	0	3	0	1	0
GGA	Gly	9.2	0	0	0	2	0

*Calculated with ATGme [1]

[1] E. Daniel, G.U. Onwukwe, R.K. Wierenga, S.E. Quaggin, S.J. Vainio, M. Krause, ATGme: Open-source web application for rare codon identification and custom DNA sequence optimization. BMC bioinformatics 16 (2015) 303.

Table A2-3. RNAP protein yield from co-expression of plasmids pDAP5 and pDAP7

Sample	Total protein (mg)
Cell free extract	3234
Ni-NTA	10
Heparin affinity	6
TEV cleavage Ni-NTA	3

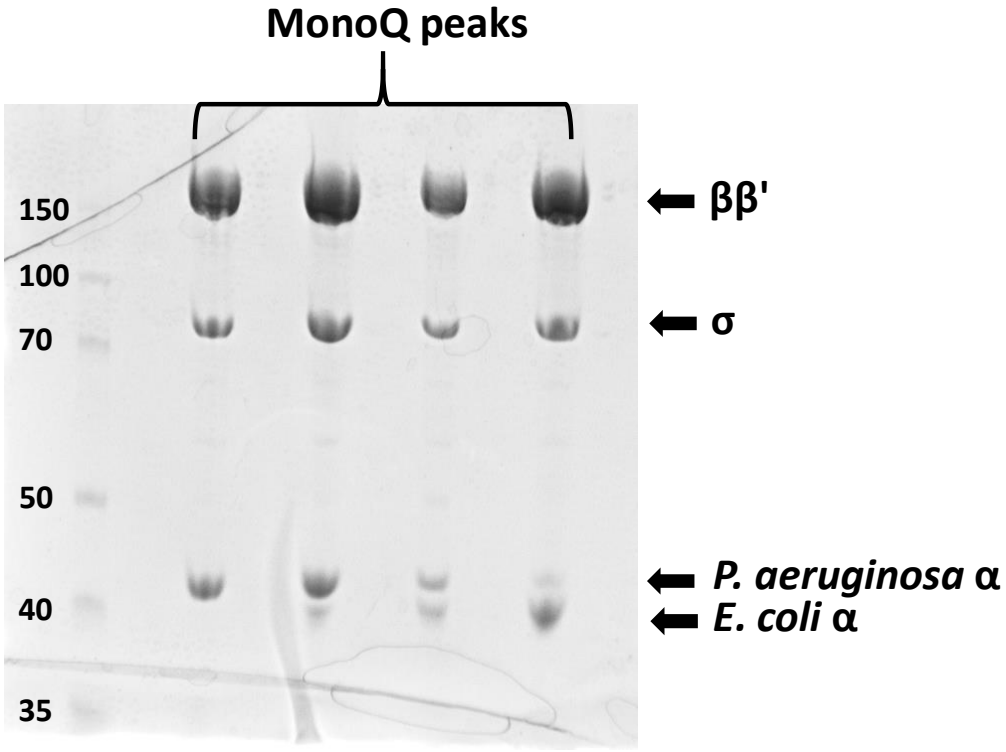
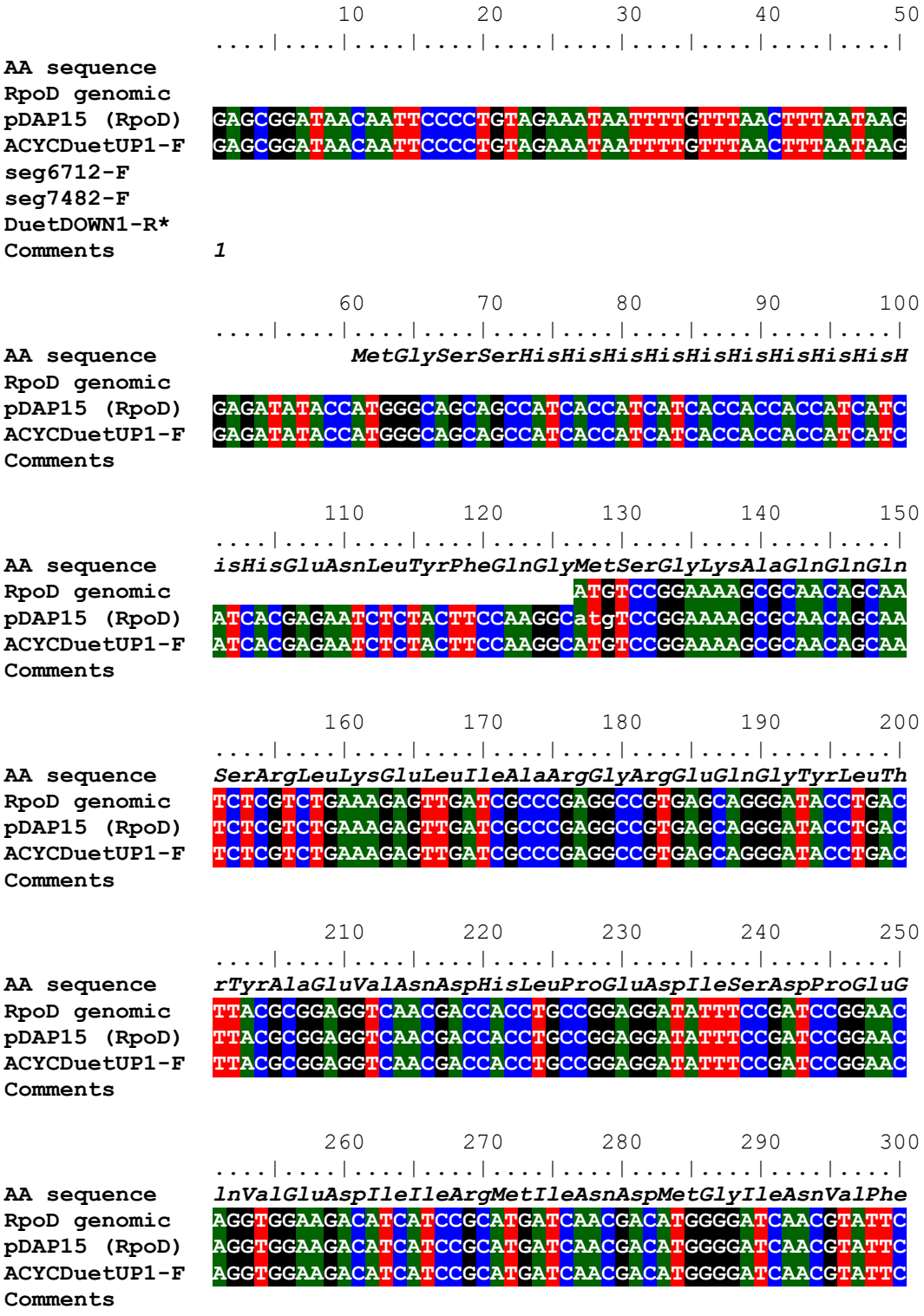
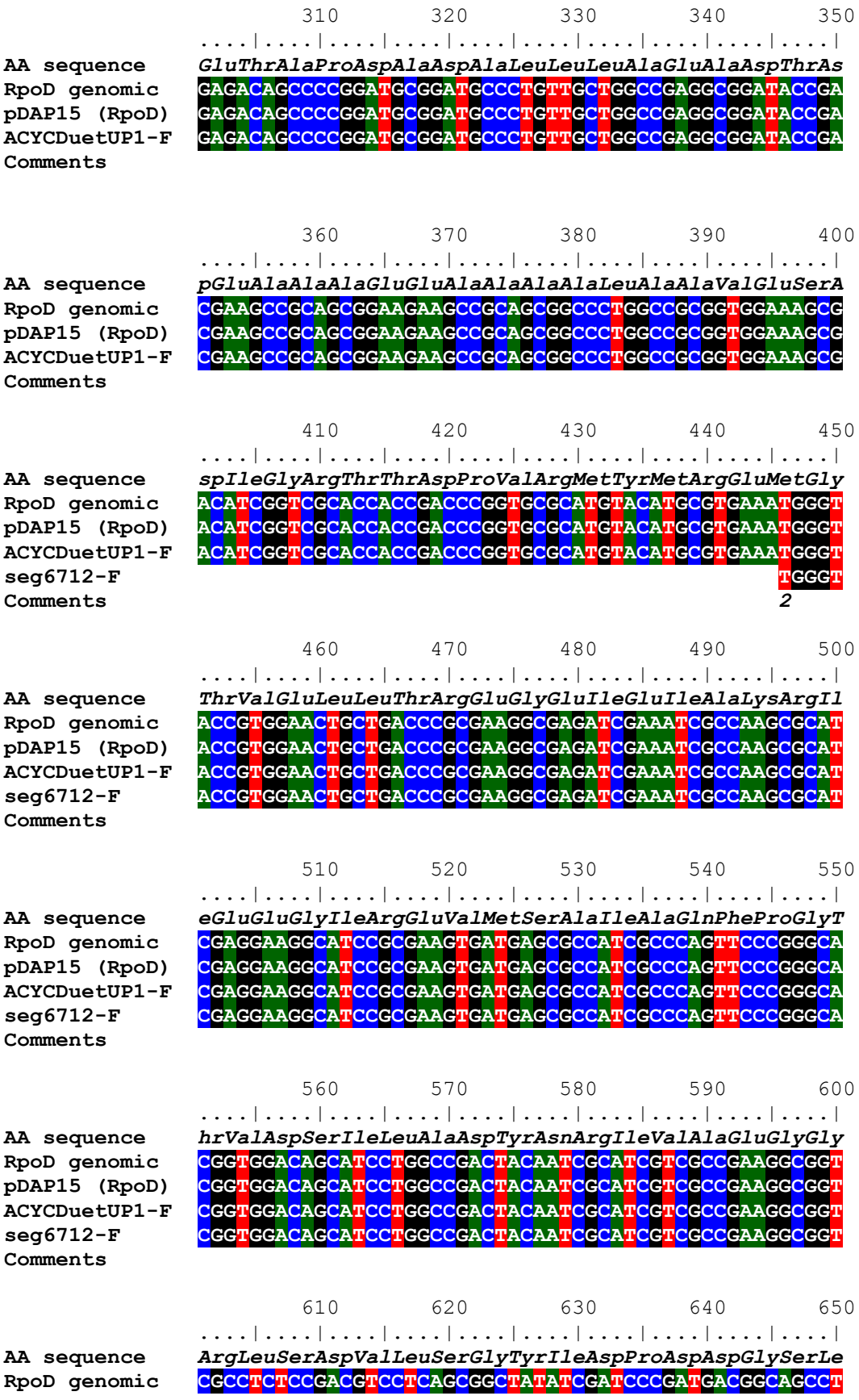


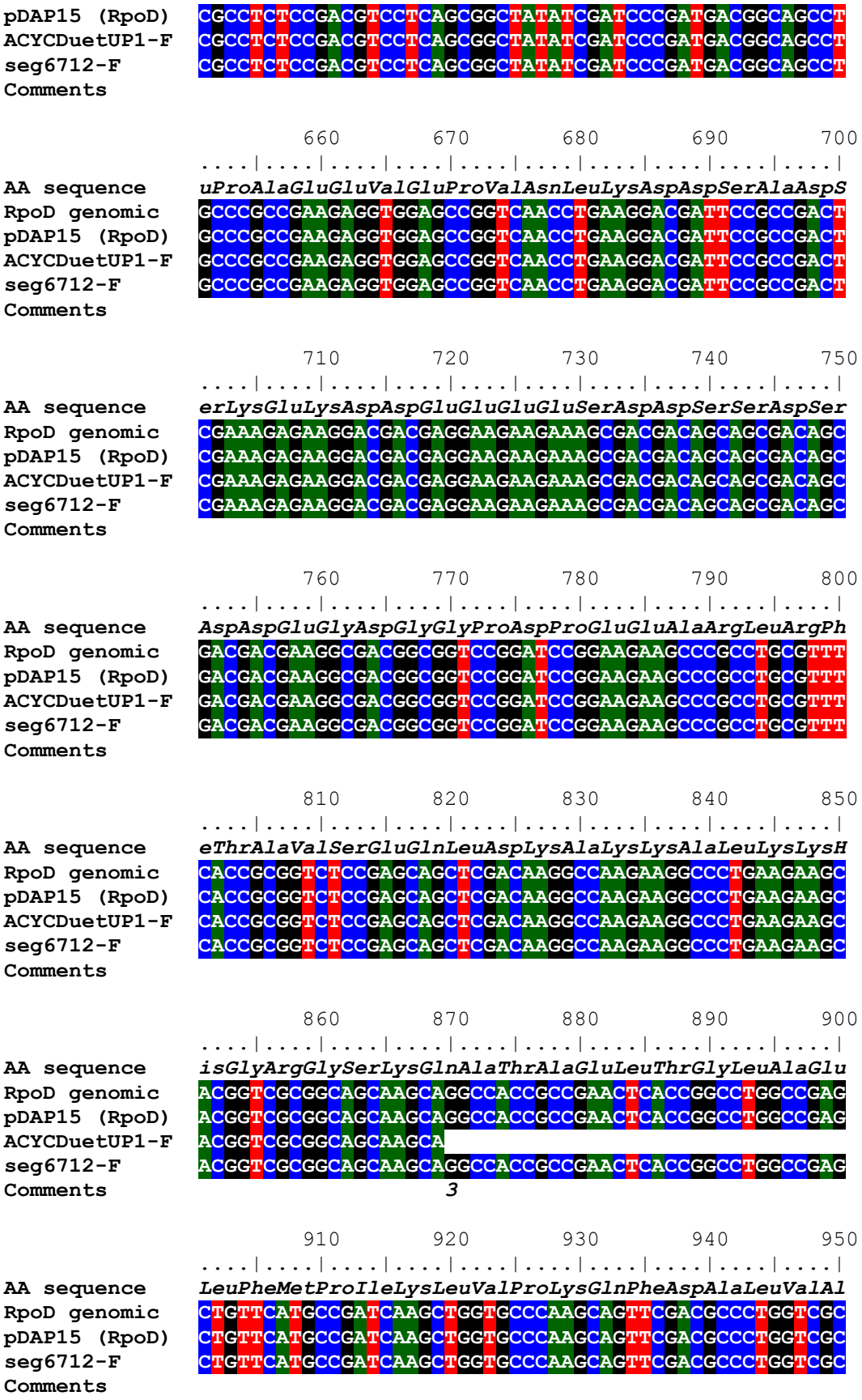
Figure A2-1. *E. coli* α contamination of *P. aeruginosa* RNAP fractions using the triple plasmid set (pDAP18, pDAP10, pDAP15) for heterologous expression. The different lanes represent the peaks obtained from the MonoQ purification step, which was able to resolve the sample into homogenous *P. aeruginosa* fractions and various heterogenous fractions. The identity of the *E. coli* α subunit contamination was revealed by MALDI mass spectrometry of bands isolated from similar SDS-PAGE gels.

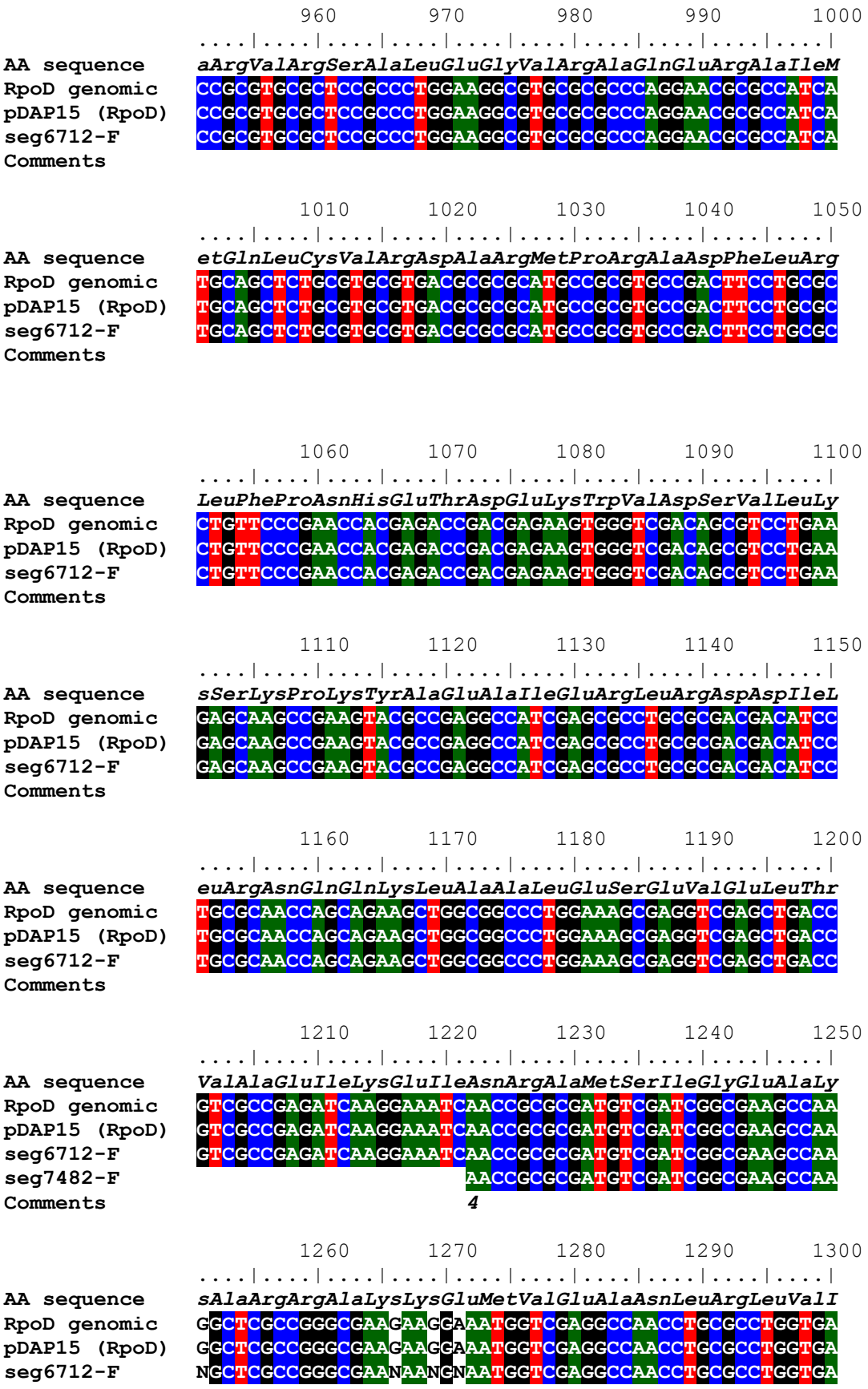
RNAP Plasmids Sequence Alignments

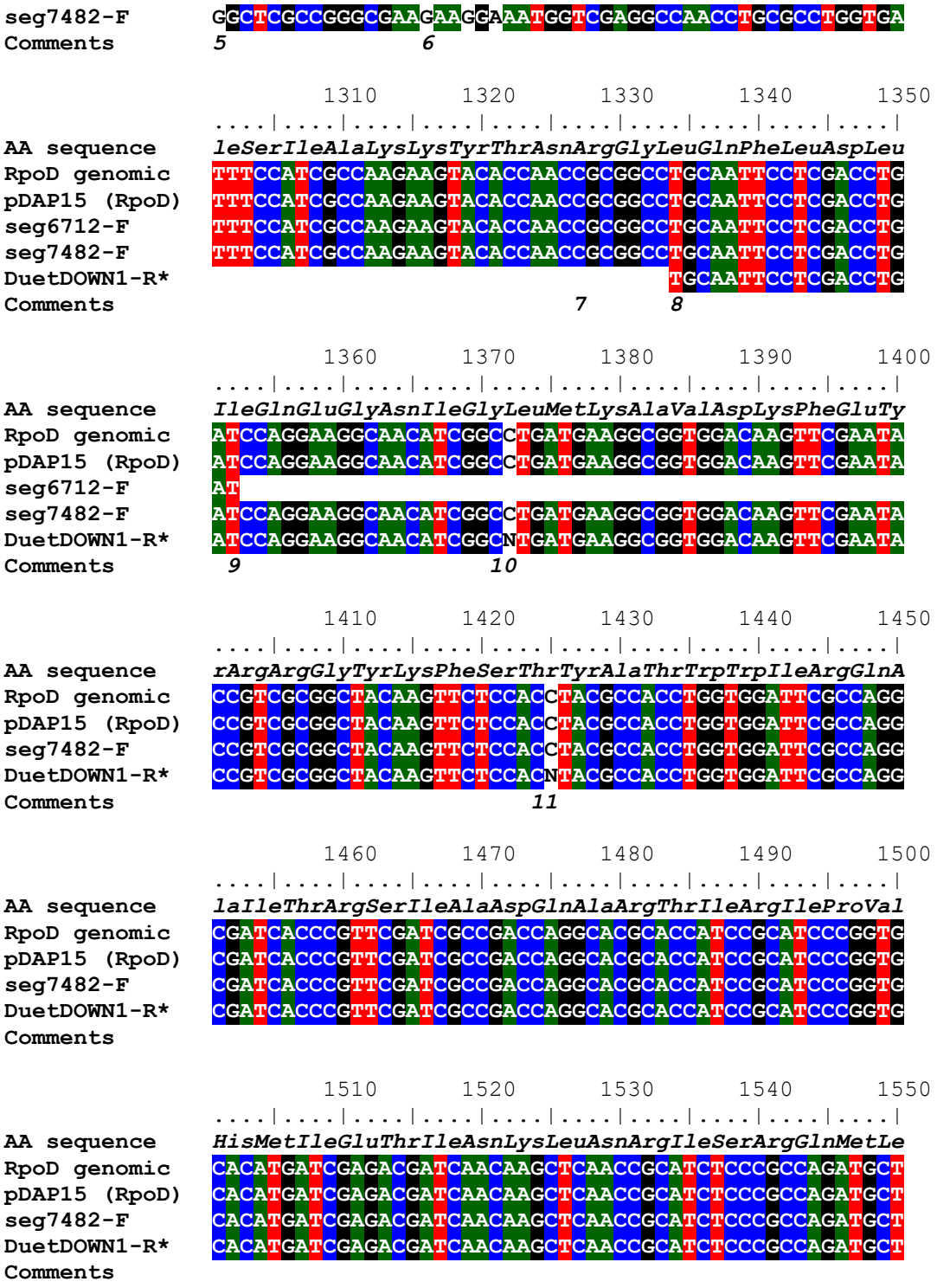
Sequencing analysis for plasmid pDAP10 (RpoAZ pET28b)

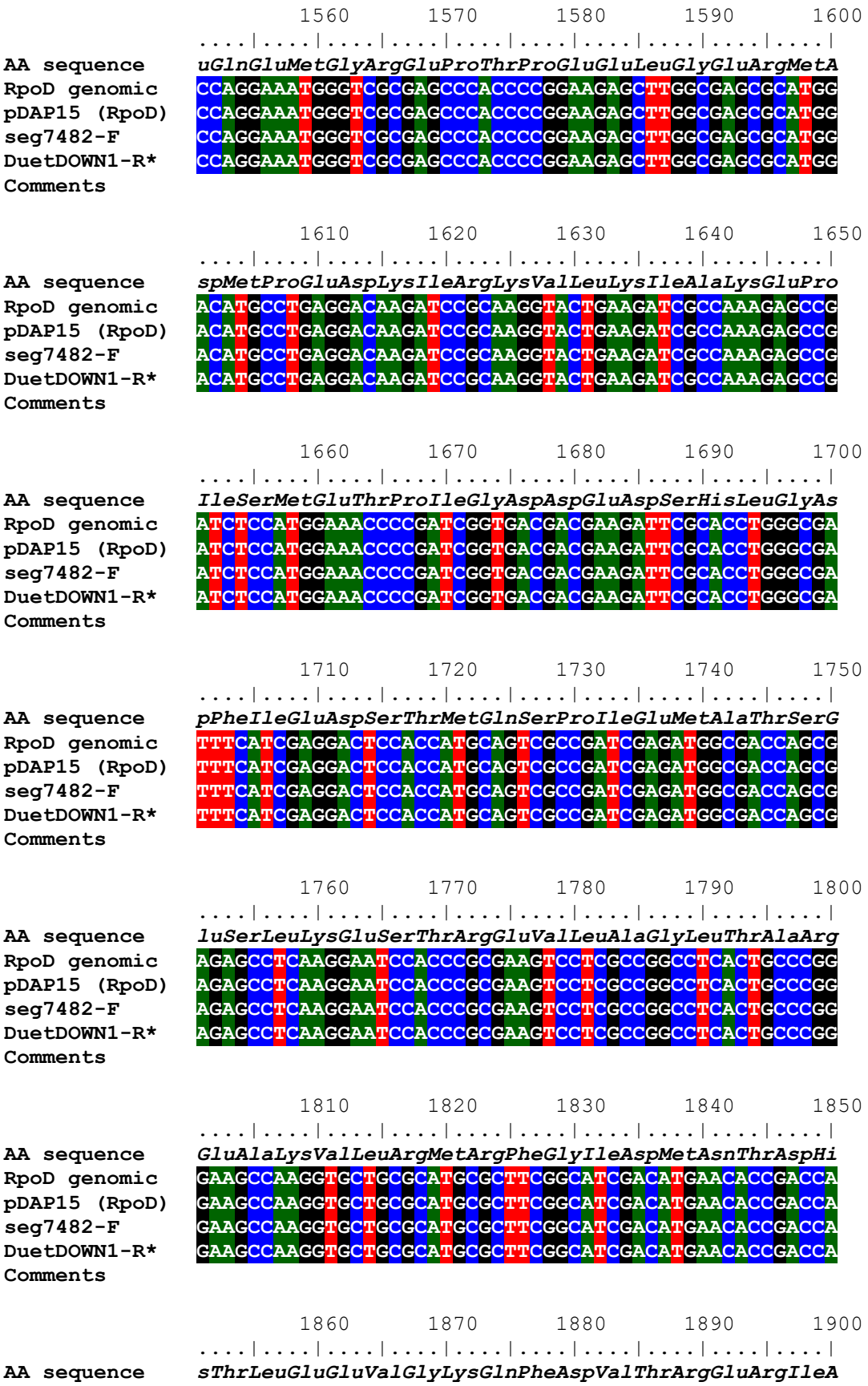


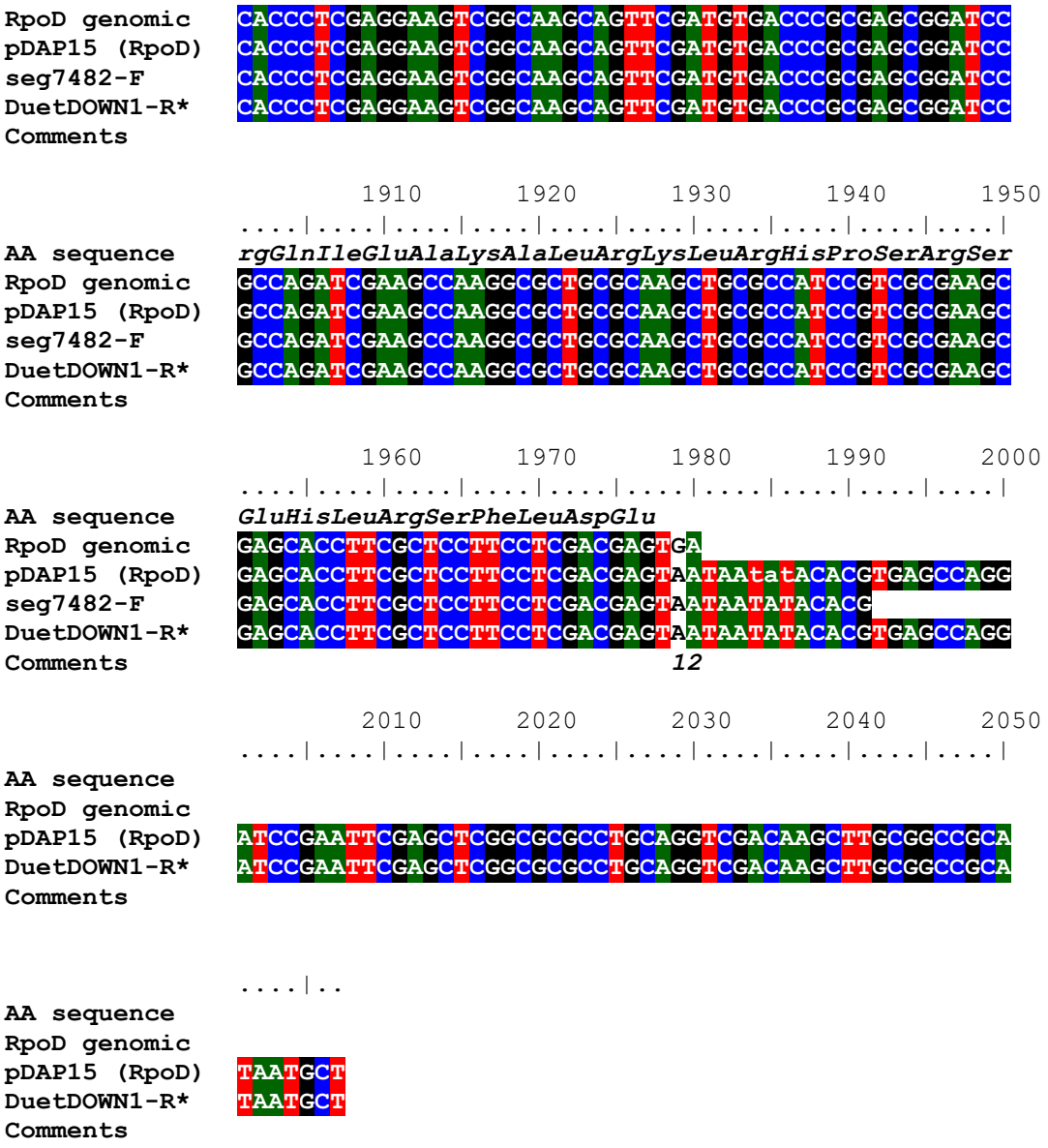

310 320 330 340 350
AA sequence GluThrAlaProAspAlaAspAlaLeuLeuLeuAlaGluAlaAspThrAs
RpoD genomic GAGACAGCCCCGGATGCGGATGCCCTGTTGCTGGCCGAGGCGGATACCGA
pDAP15 (RpoD) GAGACAGCCCCGGATGCGGATGCCCTGTTGCTGGCCGAGGCGGATACCGA
ACYCDuetUP1-F GAGACAGCCCCGGATGCGGATGCCCTGTTGCTGGCCGAGGCGGATACCGA
Comments
360 370 380 390 400
AA sequence pGluAlaAlaAlaGluGluAlaAlaAlaAlaLeuAlaAlaValGluSerA
RpoD genomic CGAAGCCGCAGCGGAAGAAGCCGCAGCGGCCCTGGCCGCGGTGGAAAGCG
pDAP15 (RpoD) CGAAGCCGCAGCGGAAGAAGCCGCAGCGGCCCTGGCCGCGGTGGAAAGCG
ACYCDuetUP1-F CGAAGCCGCAGCGGAAGAAGCCGCAGCGGCCCTGGCCGCGGTGGAAAGCG
Comments
410 420 430 440 450
AA sequence spIleGlyArgThrThrAspProValArgMetTyrMetArgGluMetGly
RpoD genomic ACATCGGTCGCACCACCGACCCGGTGCGCATGTACATGCGTGAAATGGGT
pDAP15 (RpoD) ACATCGGTCGCACCACCGACCCGGTGCGCATGTACATGCGTGAAATGGGT
ACYCDuetUP1-F ACATCGGTCGCACCACCGACCCGGTGCGCATGTACATGCGTGAAATGGGT
seg6712-F TGGGT
Comments 2
460 470 480 490 500
AA sequence ThrValGluLeuLeuThrArgGluGlyGluIleGluIleAlaLysArgIl
RpoD genomic ACCGTGGAACTGCTGACCCGCGAAGGCGAGATCGAAATCGCCAAGCGCAT
pDAP15 (RpoD) ACCGTGGAACTGCTGACCCGCGAAGGCGAGATCGAAATCGCCAAGCGCAT
ACYCDuetUP1-F ACCGTGGAACTGCTGACCCGCGAAGGCGAGATCGAAATCGCCAAGCGCAT
seg6712-F ACCGTGGAACTGCTGACCCGCGAAGGCGAGATCGAAATCGCCAAGCGCAT
Comments
510 520 530 540 550
AA sequence eGluGluGlyIleArgGluValMetSerAlaIleAlaGlnPheProGlyT
RpoD genomic CGAGGAAGGCATCCGCGAAGTGATGAGCGCCATCGCCCAGTTCCCGGGCA
pDAP15 (RpoD) CGAGGAAGGCATCCGCGAAGTGATGAGCGCCATCGCCCAGTTCCCGGGCA
ACYCDuetUP1-F CGAGGAAGGCATCCGCGAAGTGATGAGCGCCATCGCCCAGTTCCCGGGCA
seg6712-F CGAGGAAGGCATCCGCGAAGTGATGAGCGCCATCGCCCAGTTCCCGGGCA
Comments
560 570 580 590 600
AA sequence hrValAspSerIleLeuAlaAspTyrAsnArgIleValAlaGluGlyGly
RpoD genomic CGGTGGACAGCATCCTGGCCGACTACAATCGCATCGTCGCCGAAGGCGGT
pDAP15 (RpoD) CGGTGGACAGCATCCTGGCCGACTACAATCGCATCGTCGCCGAAGGCGGT
ACYCDuetUP1-F CGGTGGACAGCATCCTGGCCGACTACAATCGCATCGTCGCCGAAGGCGGT
seg6712-F CGGTGGACAGCATCCTGGCCGACTACAATCGCATCGTCGCCGAAGGCGGT
Comments
610 620 630 640 650
AA sequence ArgLeuSerAspValLeuSerGlyTyrIleAspProAspAspGlySerLe
RpoD genomic CGCCTCTCCGACGTCCTCAGCGGCTATATCGATCCCGATGACGGCAGCCT


pDAP15 (RpoD)
ACYCDuetUP1-F
seg6712-F
Comments
660 670 680 690 700
AA sequence
uProAlaGluGluValGluProValAsnLeuLysAspAspSerAlaAspS
RpoD genomic
pDAP15 (RpoD)
ACYCDuetUP1-F
seg6712-F
Comments
710 720 730 740 750
AA sequence
erLysGluLysAspAspGluGluGluGluSerAspAspSerSerAspSer
RpoD genomic
pDAP15 (RpoD)
ACYCDuetUP1-F
seg6712-F
Comments
760 770 780 790 800
AA sequence
AspAspGluGlyAspGlyGlyProAspProGluGluAlaArgLeuArgPh
RpoD genomic
pDAP15 (RpoD)
ACYCDuetUP1-F
seg6712-F
Comments
810 820 830 840 850
AA sequence
eThrAlaValSerGluGlnLeuAspLysAlaLysLysAlaLeuLysLysH
RpoD genomic
pDAP15 (RpoD)
ACYCDuetUP1-F
seg6712-F
Comments
860 870 880 890 900
AA sequence
isGlyArgGlySerLysGlnAlaThrAlaGluLeuThrGlyLeuAlaGlu
RpoD genomic
pDAP15 (RpoD)
ACYCDuetUP1-F
seg6712-F
Comments
3
910 920 930 940 950
AA sequence
LeuPheMetProIleLysLeuValProLysGlnPheAspAlaLeuValAl
RpoD genomic
pDAP15 (RpoD)
seg6712-F
Comments


960 970 980 990 1000
AA sequence aArgValArgSerAlaLeuGluGlyValArgAlaGlnGluArgAlaIleM
RpoD genomic
pDAP15 (RpoD)
seg6712-F
Comments
1010 1020 1030 1040 1050
AA sequence etGlnLeuCysValArgAspAlaArgMetProArgAlaAspPheLeuArg
RpoD genomic
pDAP15 (RpoD)
seg6712-F
Comments
1060 1070 1080 1090 1100
AA sequence LeuPheProAsnHisGluThrAspGluLysTrpValAspSerValLeuLy
RpoD genomic
pDAP15 (RpoD)
seg6712-F
Comments
1110 1120 1130 1140 1150
AA sequence sSerLysProLysTyrAlaGluAlaIleGluArgLeuArgAspAspIleL
RpoD genomic
pDAP15 (RpoD)
seg6712-F
Comments
1160 1170 1180 1190 1200
AA sequence euArgAsnGlnGlnLysLeuAlaAlaLeuGluSerGluValGluLeuThr
RpoD genomic
pDAP15 (RpoD)
seg6712-F
Comments
1210 1220 1230 1240 1250
AA sequence ValAlaGluIleLysGluIleAsnArgAlaMetSerIleGlyGluAlaLy
RpoD genomic
pDAP15 (RpoD)
seg6712-F
seg7482-F
Comments
4
1260 1270 1280 1290 1300
AA sequence sAlaArgArgAlaLysLysGluMetValGluAlaAsnLeuArgLeuValI
RpoD genomic
pDAP15 (RpoD)
seg6712-F


seg7482-F
Comments
1310 1320 1330 1340 1350
AA sequence
RpoD genomic
pDAP15 (RpoD)
seg6712-F
seg7482-F
DuetDOWN1-R*
Comments
1360 1370 1380 1390 1400
AA sequence
RpoD genomic
pDAP15 (RpoD)
seg6712-F
seg7482-F
DuetDOWN1-R*
Comments
1410 1420 1430 1440 1450
AA sequence
RpoD genomic
pDAP15 (RpoD)
seg7482-F
DuetDOWN1-R*
Comments
1460 1470 1480 1490 1500
AA sequence
RpoD genomic
pDAP15 (RpoD)
seg7482-F
DuetDOWN1-R*
Comments
1510 1520 1530 1540 1550
AA sequence
RpoD genomic
pDAP15 (RpoD)
seg7482-F
DuetDOWN1-R*
Comments

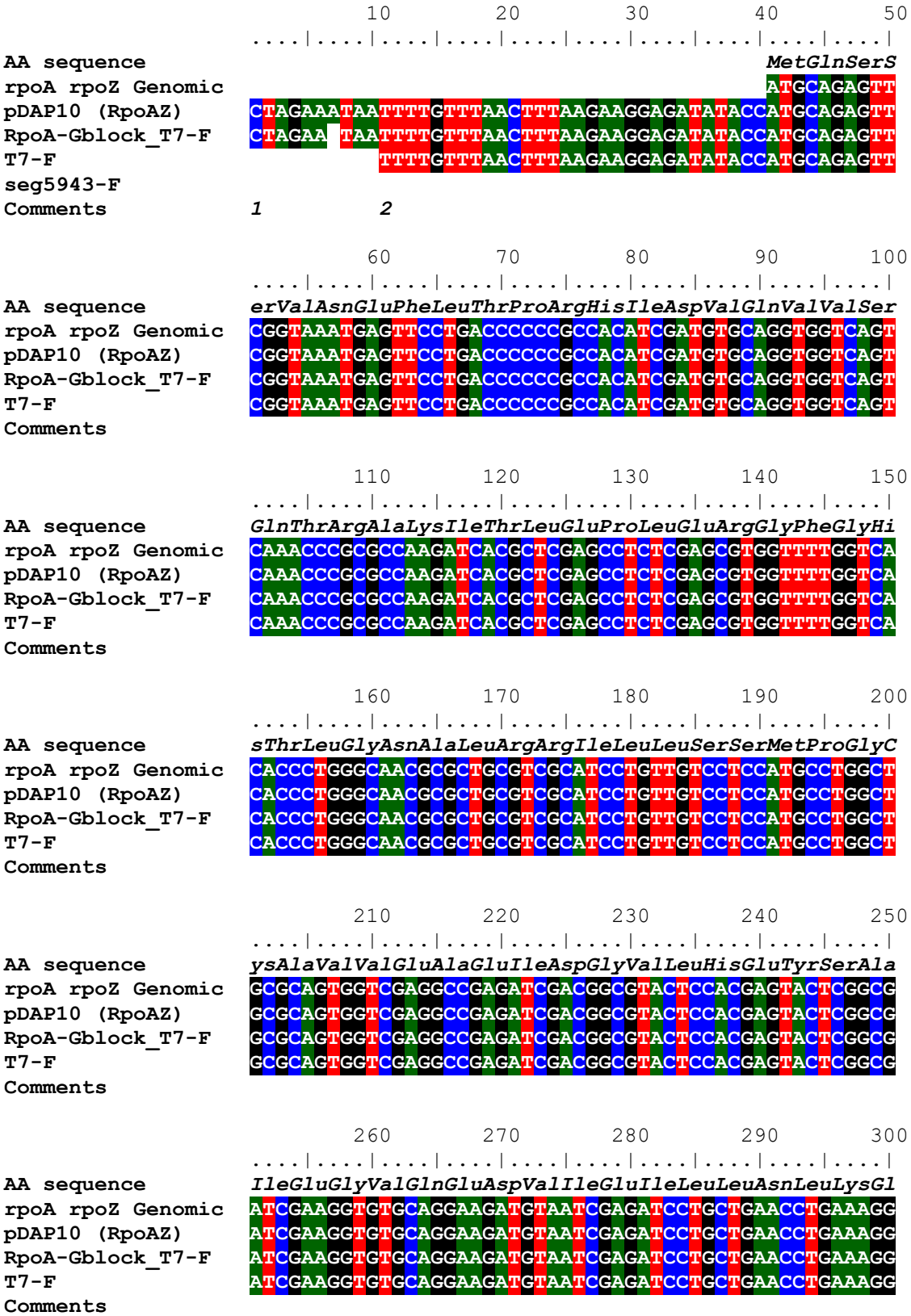

1560 1570 1580 1590 1600
AA sequence uGlnGluMetGlyArgGluProThrProGluGluLeuGlyGluArgMetA
RpoD genomic CCAGGAAATGGGTCGCGAGCCCACCCCGGAAGAGCTTGGCGAGCGCATGG
pDAP15 (RpoD) CCAGGAAATGGGTCGCGAGCCCACCCCGGAAGAGCTTGGCGAGCGCATGG
seg7482-F CCAGGAAATGGGTCGCGAGCCCACCCCGGAAGAGCTTGGCGAGCGCATGG
DuetDOWN1-R* CCAGGAAATGGGTCGCGAGCCCACCCCGGAAGAGCTTGGCGAGCGCATGG
Comments
1610 1620 1630 1640 1650
AA sequence spMetProGluAspLysIleArgLysValLeuLysIleAlaLysGluPro
RpoD genomic ACATGCCTGAGGACAAGATCCGCAAGGTACTGAAGATCGCCAAAGAGCCG
pDAP15 (RpoD) ACATGCCTGAGGACAAGATCCGCAAGGTACTGAAGATCGCCAAAGAGCCG
seg7482-F ACATGCCTGAGGACAAGATCCGCAAGGTACTGAAGATCGCCAAAGAGCCG
DuetDOWN1-R* ACATGCCTGAGGACAAGATCCGCAAGGTACTGAAGATCGCCAAAGAGCCG
Comments
1660 1670 1680 1690 1700
AA sequence IleSerMetGluThrProIleGlyAspAspGluAspSerHisLeuGlyAs
RpoD genomic ATCTCCATGGAAACCCCGATCGGTGACGACGAAGATTCGCACCTGGGCGA
pDAP15 (RpoD) ATCTCCATGGAAACCCCGATCGGTGACGACGAAGATTCGCACCTGGGCGA
seg7482-F ATCTCCATGGAAACCCCGATCGGTGACGACGAAGATTCGCACCTGGGCGA
DuetDOWN1-R* ATCTCCATGGAAACCCCGATCGGTGACGACGAAGATTCGCACCTGGGCGA
Comments
1710 1720 1730 1740 1750
AA sequence pPheIleGluAspSerThrMetGlnSerProIleGluMetAlaThrSerG
RpoD genomic TTTCATCGAGGACTCCACCATGCAGTCGCCGATCGAGATGGCGACCAGCG
pDAP15 (RpoD) TTTCATCGAGGACTCCACCATGCAGTCGCCGATCGAGATGGCGACCAGCG
seg7482-F TTTCATCGAGGACTCCACCATGCAGTCGCCGATCGAGATGGCGACCAGCG
DuetDOWN1-R* TTTCATCGAGGACTCCACCATGCAGTCGCCGATCGAGATGGCGACCAGCG
Comments
1760 1770 1780 1790 1800
AA sequence luSerLeuLysGluSerThrArgGluValLeuAlaGlyLeuThrAlaArg
RpoD genomic AGAGCCTCAAGGAATCCACCCGCGAAGTCCTCGCCGGCCTCACTGCCCGG
pDAP15 (RpoD) AGAGCCTCAAGGAATCCACCCGCGAAGTCCTCGCCGGCCTCACTGCCCGG
seg7482-F AGAGCCTCAAGGAATCCACCCGCGAAGTCCTCGCCGGCCTCACTGCCCGG
DuetDOWN1-R* AGAGCCTCAAGGAATCCACCCGCGAAGTCCTCGCCGGCCTCACTGCCCGG
Comments
1810 1820 1830 1840 1850
AA sequence GluAlaLysValLeuArgMetArgPheGlyIleAspMetAsnThrAspHi
RpoD genomic GAAGCCAAGGTGCTGCGCATGCGCTTCGGCATCGACATGAACACCGACCA
pDAP15 (RpoD) GAAGCCAAGGTGCTGCGCATGCGCTTCGGCATCGACATGAACACCGACCA
seg7482-F GAAGCCAAGGTGCTGCGCATGCGCTTCGGCATCGACATGAACACCGACCA
DuetDOWN1-R* GAAGCCAAGGTGCTGCGCATGCGCTTCGGCATCGACATGAACACCGACCA
Comments
1860 1870 1880 1890 1900
AA sequence sThrLeuGluGluValGlyLysGlnPheAspValThrArgGluArgIleA

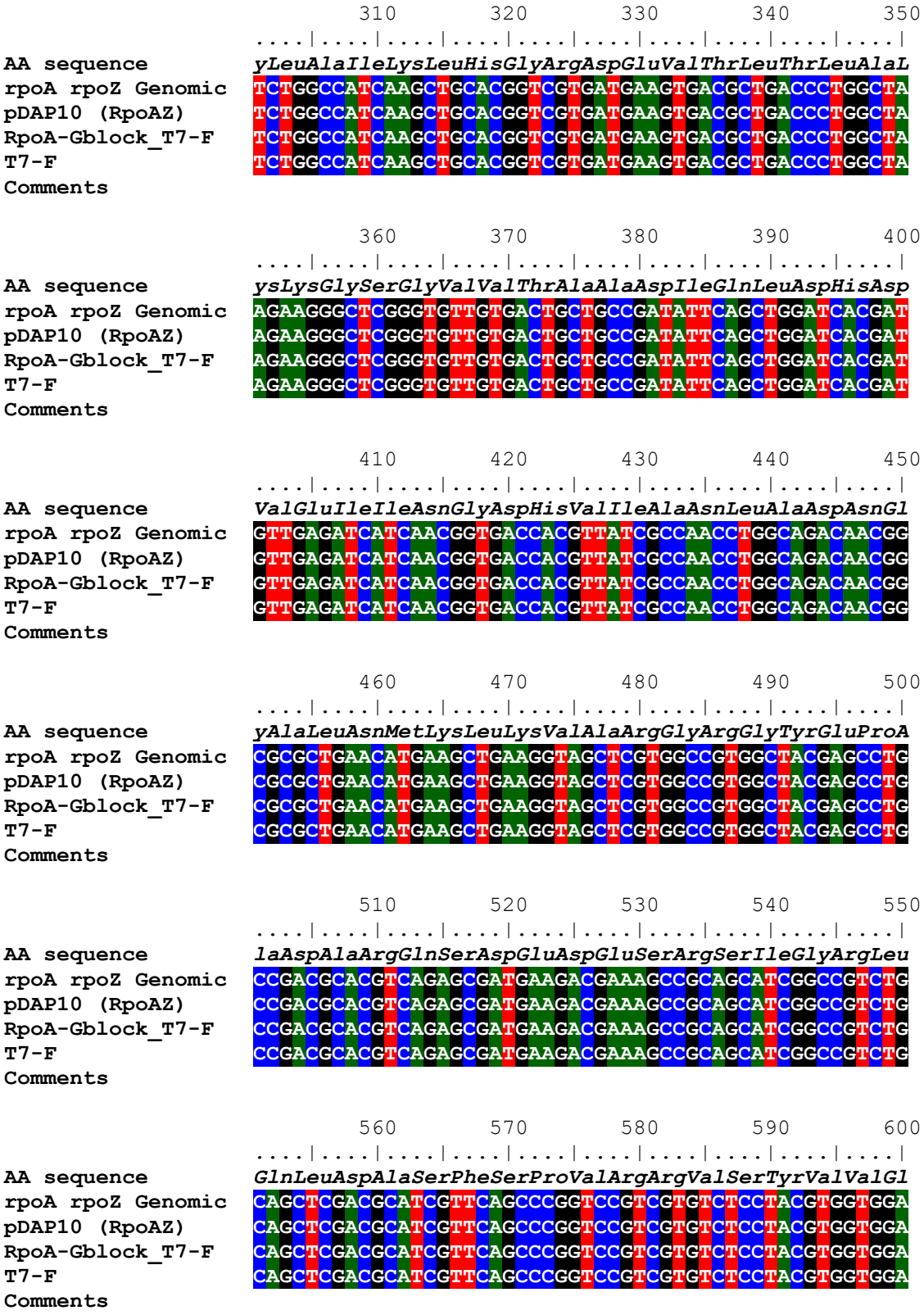

RpoD genomic
pDAP15 (RpoD)
seg7482-F
DuetDOWN1-R*
Comments
CACCCTCGAGGAAGTCGGCAAGCAGTTCGATGTGACCCGCGAGCGGATCC
1910
1920
1930
1940
1950
AA sequence
rgGlnIleGluAlaLysAlaLeuArgLysLeuArgHisProSerArgSer
GCCAGATCGAAGCCAAGGCGCTGCGCAAGCTGCGCCATCCGTCGCGAAGC
1960
1970
1980
1990
2000
GluHisLeuArgSerPheLeuAspGlu
GAGCACCTTCGCTCCTTCCTCGACGAGTGA
GAGCACCTTCGCTCCTTCCTCGACGAGTAATAAtatACACGTGAGCCAGG
GAGCACCTTCGCTCCTTCCTCGACGAGTAATAATATACACG
GAGCACCTTCGCTCCTTCCTCGACGAGTAATAATATACACGTGAGCCAGG
12
2010
2020
2030
2040
2050
ATCCGAATTCGAGCTCGGCGCGCCTGCAGGTCGACAAGCTTGCGGCCGCA
....|..
TAATGCT

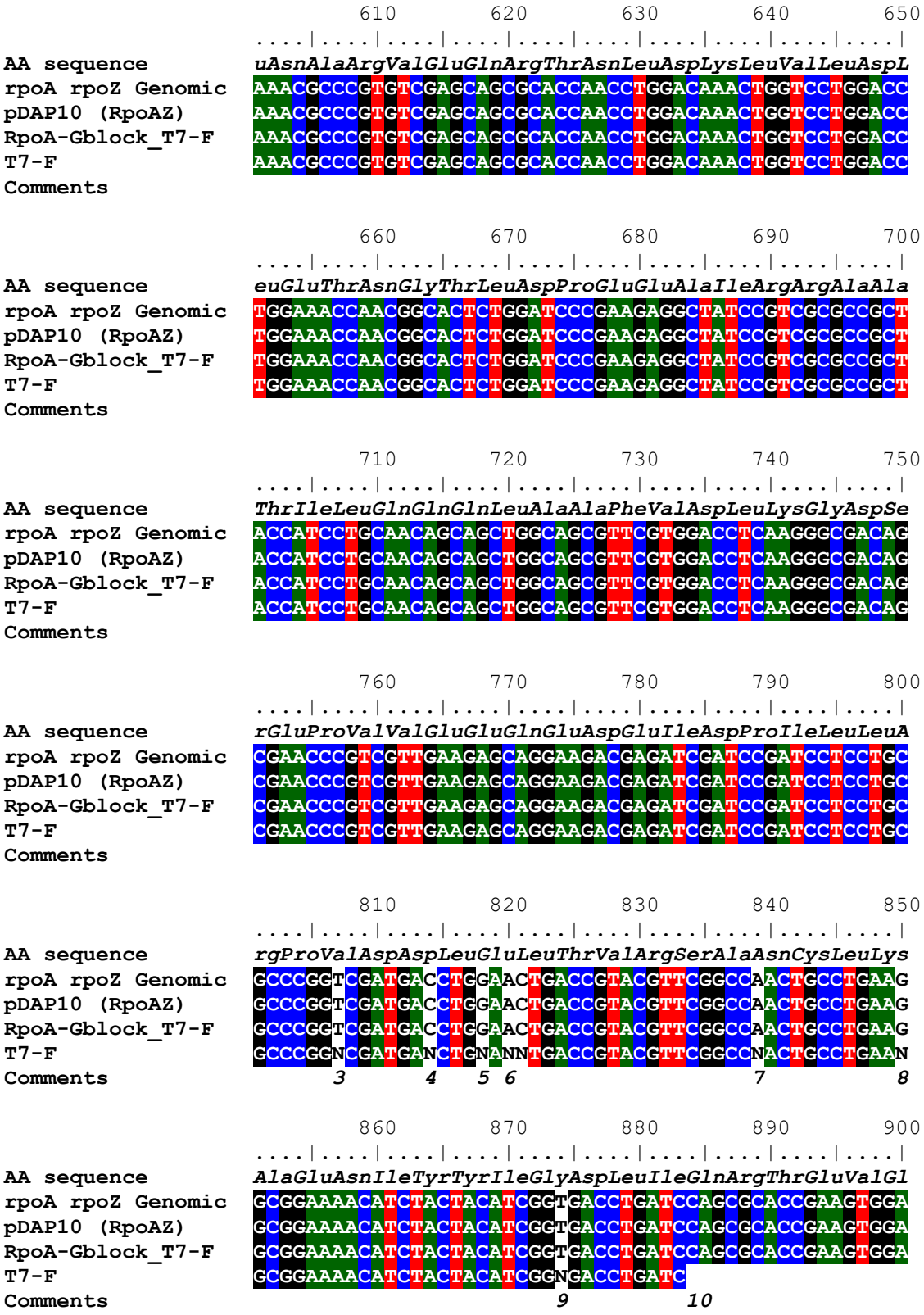
pDAP15 Sequencing Corrections plasmid RpoD DUET

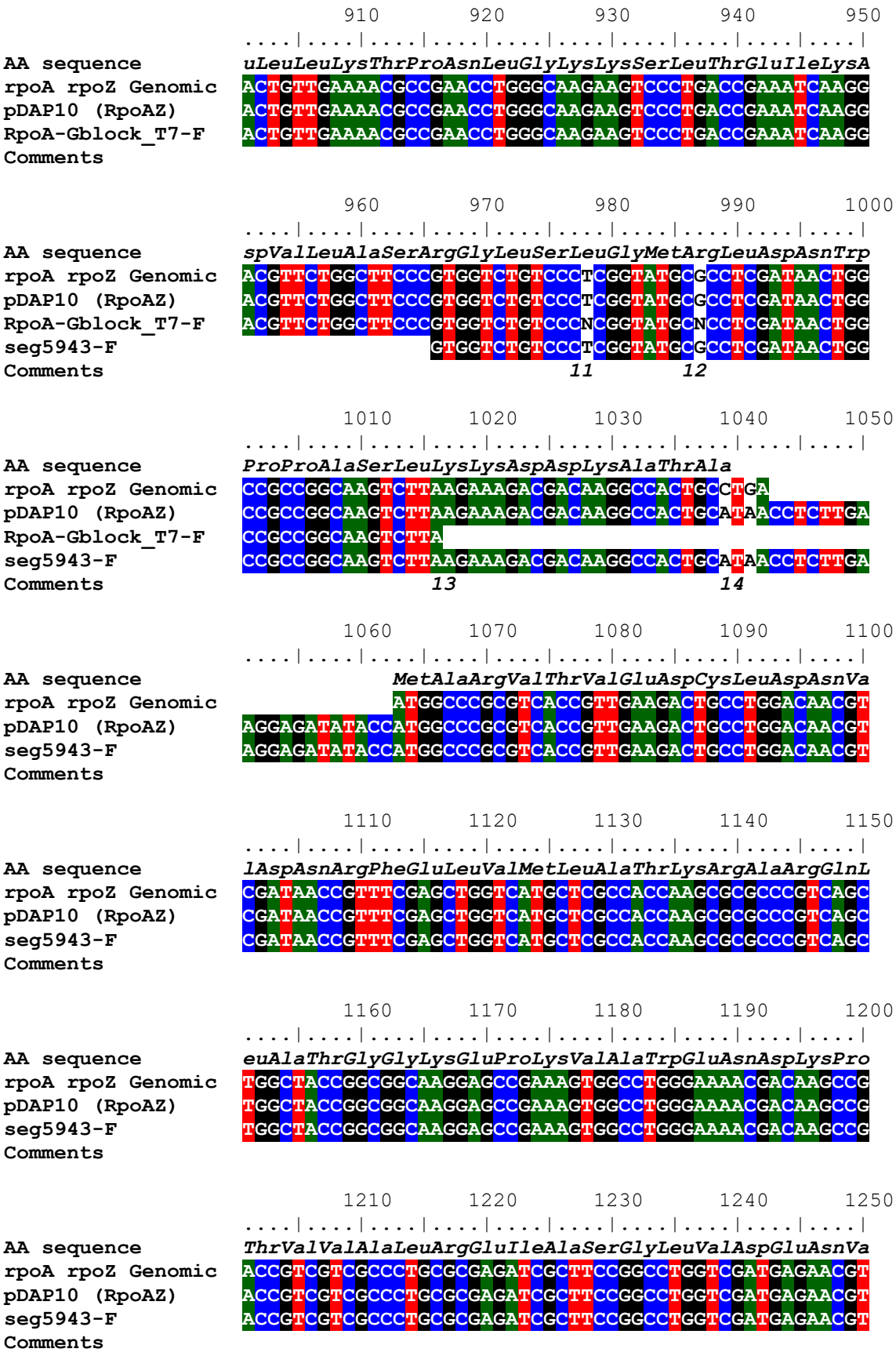
Comment #	**Sequencing primer**	Comments
1	ACYCDuetUP1-F	1-32 removed
2	seg6712-F	1-34 removed
3	ACYCDuetUP1-F	901-end removed
4	seg7482	1-33 removed
5	seg6712	G double is evident. N is a G.
6	seg6712	trace is consistent with correct sequence (G peaks and A) through the area
7	seg6712	trace is consistent with N removed (CCGCGG seq)
8	DuetDOWN1-R*	1-32 removed
9	seg6712	trace is poor after 908; end nucleotides removed
10	DuetDOWN1-R*	G (C*) has shoulder peak consistent with correct
11	DuetDOWN1-R*	G (C*) has shoulder peak consistent with correct
12		TGA stop codon replaced with TAATAA

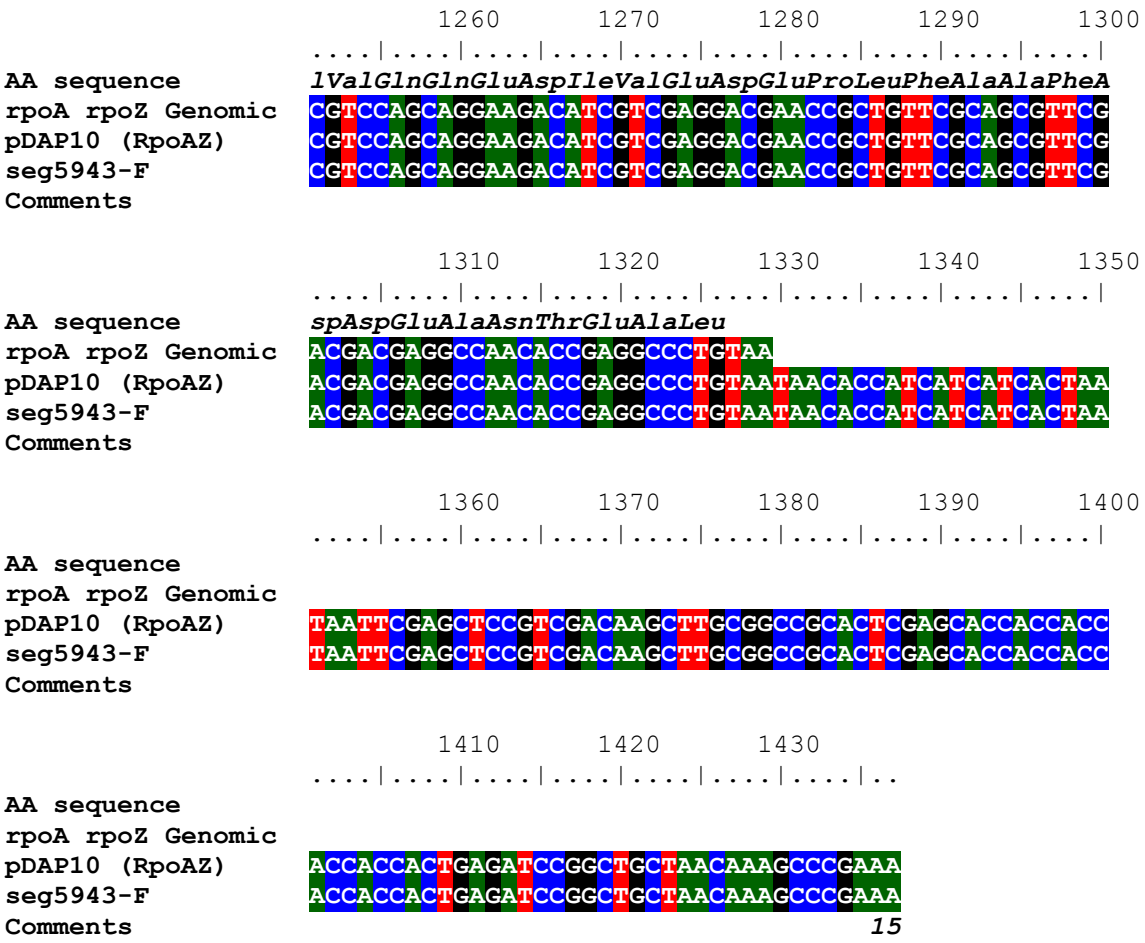
Sequencing analysis for plasmid pDAP10 (RpoAZ pET28b)




310 320 330 340 350
AA sequence
yLeuAlaIleLysLeuHisGlyArgAspGluValThrLeuThrLeuAlaL
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
RpoA-Gblock_T7-F
T7-F
Comments
360 370 380 390 400
AA sequence
ysLysGlySerGlyValValThrAlaAlaAspIleGlnLeuAspHisAsp
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
RpoA-Gblock_T7-F
T7-F
Comments
410 420 430 440 450
AA sequence
ValGluIleIleAsnGlyAspHisValIleAlaAsnLeuAlaAspAsnGl
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
RpoA-Gblock_T7-F
T7-F
Comments
460 470 480 490 500
AA sequence
yAlaLeuAsnMetLysLeuLysValAlaArgGlyArgGlyTyrGluProA
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
RpoA-Gblock_T7-F
T7-F
Comments
510 520 530 540 550
AA sequence
laAspAlaArgGlnSerAspGluAspGluSerArgSerIleGlyArgLeu
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
RpoA-Gblock_T7-F
T7-F
Comments
560 570 580 590 600
AA sequence
GlnLeuAspAlaSerPheSerProValArgArgValSerTyrValValGl
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
RpoA-Gblock_T7-F
T7-F
Comments


610 620 630 640 650
AA sequence uAsnAlaArgValGluGlnArgThrAsnLeuAspLysLeuValLeuAspL
rpoA rpoZ Genomic AAACGCCCGTGTCGAGCAGCGCACCAACCTGGACAAACTGGTCCTGGACC
pDAP10 (RpoAZ) AAACGCCCGTGTCGAGCAGCGCACCAACCTGGACAAACTGGTCCTGGACC
RpoA-Gblock_T7-F AAACGCCCGTGTCGAGCAGCGCACCAACCTGGACAAACTGGTCCTGGACC
T7-F AAACGCCCGTGTCGAGCAGCGCACCAACCTGGACAAACTGGTCCTGGACC
Comments
660 670 680 690 700
AA sequence euGluThrAsnGlyThrLeuAspProGluGluAlaIleArgArgAlaAla
rpoA rpoZ Genomic TGGAAACCAACGGCACTCTGGATCCCGAAGAGGCTATCCGTCGCGCCGCT
pDAP10 (RpoAZ) TGGAAACCAACGGCACTCTGGATCCCGAAGAGGCTATCCGTCGCGCCGCT
RpoA-Gblock_T7-F TGGAAACCAACGGCACTCTGGATCCCGAAGAGGCTATCCGTCGCGCCGCT
T7-F TGGAAACCAACGGCACTCTGGATCCCGAAGAGGCTATCCGTCGCGCCGCT
Comments
710 720 730 740 750
AA sequence ThrIleLeuGlnGlnGlnLeuAlaAlaPheValAspLeuLysGlyAspSe
rpoA rpoZ Genomic ACCATCCTGCAACAGCAGCTGGCAGCGTTCGTGGACCTCAAGGGCGACAG
pDAP10 (RpoAZ) ACCATCCTGCAACAGCAGCTGGCAGCGTTCGTGGACCTCAAGGGCGACAG
RpoA-Gblock_T7-F ACCATCCTGCAACAGCAGCTGGCAGCGTTCGTGGACCTCAAGGGCGACAG
T7-F ACCATCCTGCAACAGCAGCTGGCAGCGTTCGTGGACCTCAAGGGCGACAG
Comments
760 770 780 790 800
AA sequence rGluProValValGluGluGlnGluAspGluIleAspProIleLeuLeuA
rpoA rpoZ Genomic CGAACCCGTCGTTGAAGAGCAGGAAGACGAGATCGATCCGATCCTCCTGC
pDAP10 (RpoAZ) CGAACCCGTCGTTGAAGAGCAGGAAGACGAGATCGATCCGATCCTCCTGC
RpoA-Gblock_T7-F CGAACCCGTCGTTGAAGAGCAGGAAGACGAGATCGATCCGATCCTCCTGC
T7-F CGAACCCGTCGTTGAAGAGCAGGAAGACGAGATCGATCCGATCCTCCTGC
Comments
810 820 830 840 850
AA sequence rgProValAspAspLeuGluLeuThrValArgSerAlaAsnCysLeuLys
rpoA rpoZ Genomic GCCCGGTCGATGACCTGGAACTGACCGTACGTTCGGCCAACTGCCTGAAG
pDAP10 (RpoAZ) GCCCGGTCGATGACCTGGAACTGACCGTACGTTCGGCCAACTGCCTGAAG
RpoA-Gblock_T7-F GCCCGGTCGATGACCTGGAACTGACCGTACGTTCGGCCAACTGCCTGAAG
T7-F GCCCGGNCGATGANCTGNANNTGACCGTACGTTCGGCCNACTGCCTGAAN
Comments 3 4 5 6 7 8
860 870 880 890 900
AA sequence AlaGluAsnIleTyrTyrIleGlyAspLeuIleGlnArgThrGluValGl
rpoA rpoZ Genomic GCGGAAAACATCTACTACATCGGTGACCTGATCCAGCGCACCGAAGTGGA
pDAP10 (RpoAZ) GCGGAAAACATCTACTACATCGGTGACCTGATCCAGCGCACCGAAGTGGA
RpoA-Gblock_T7-F GCGGAAAACATCTACTACATCGGTGACCTGATCCAGCGCACCGAAGTGGA
T7-F GCGGAAAACATCTACTACATCGGNGACCTGATC
Comments 9 10

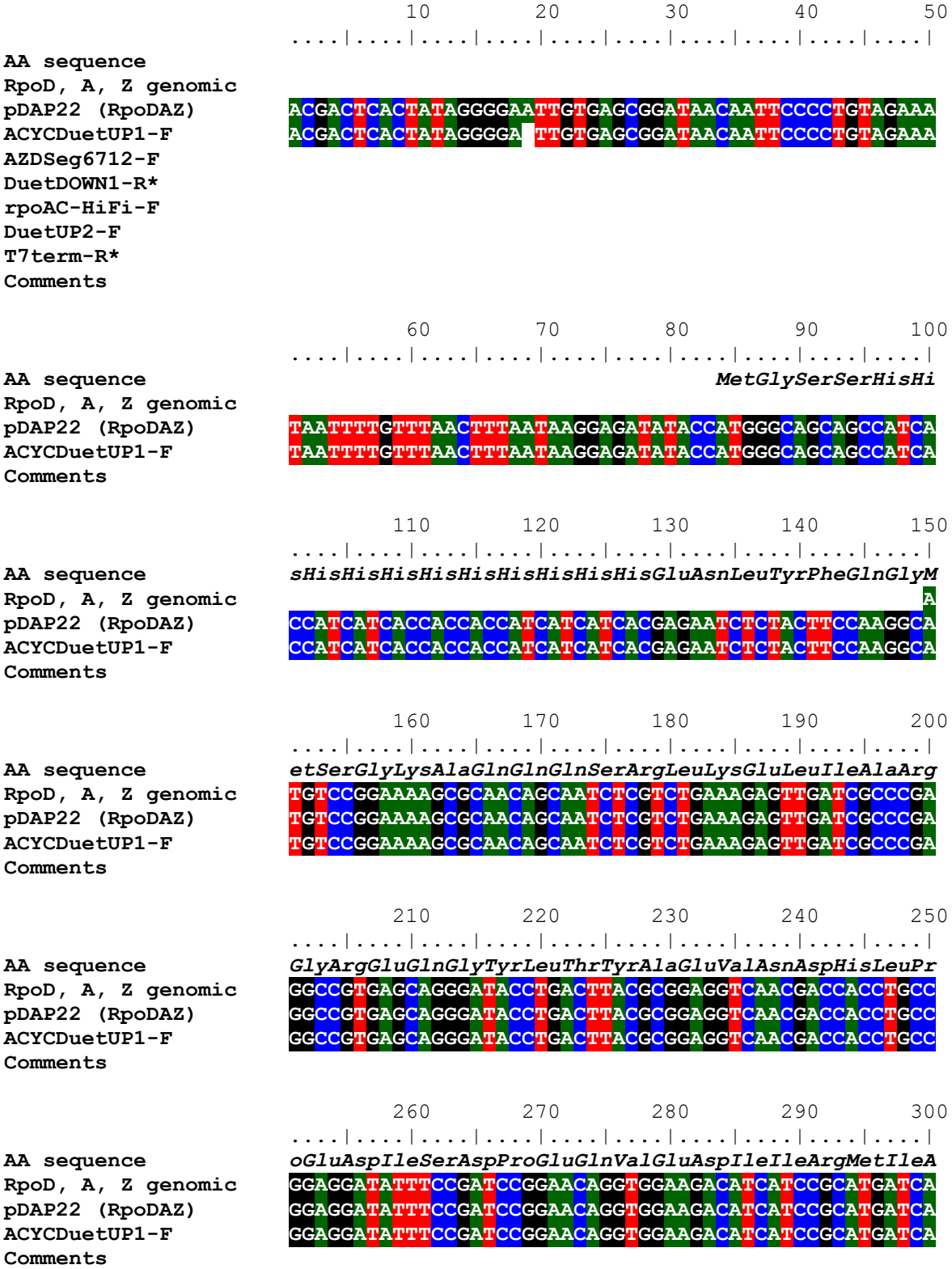

910 920 930 940 950
AA sequence uLeuLeuLysThrProAsnLeuGlyLysLysSerLeuThrGluIleLysA
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
RpoA-Gblock_T7-F
Comments
960 970 980 990 1000
AA sequence spValLeuAlaSerArgGlyLeuSerLeuGlyMetArgLeuAspAsnTrp
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
RpoA-Gblock_T7-F
seg5943-F
Comments
11 12
1010 1020 1030 1040 1050
AA sequence ProProAlaSerLeuLysLysAspAspLysAlaThrAla
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
RpoA-Gblock_T7-F
seg5943-F
Comments
13 14
1060 1070 1080 1090 1100
AA sequence MetAlaArgValThrValGluAspCysLeuAspAsnVa
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
seg5943-F
Comments
1110 1120 1130 1140 1150
AA sequence lAspAsnArgPheGluLeuValMetLeuAlaThrLysArgAlaArgGlnL
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
seg5943-F
Comments
1160 1170 1180 1190 1200
AA sequence euAlaThrGlyGlyLysGluProLysValAlaTrpGluAsnAspLysPro
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
seg5943-F
Comments
1210 1220 1230 1240 1250
AA sequence ThrValValAlaLeuArgGluIleAlaSerGlyLeuValAspGluAsnVa
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
seg5943-F
Comments

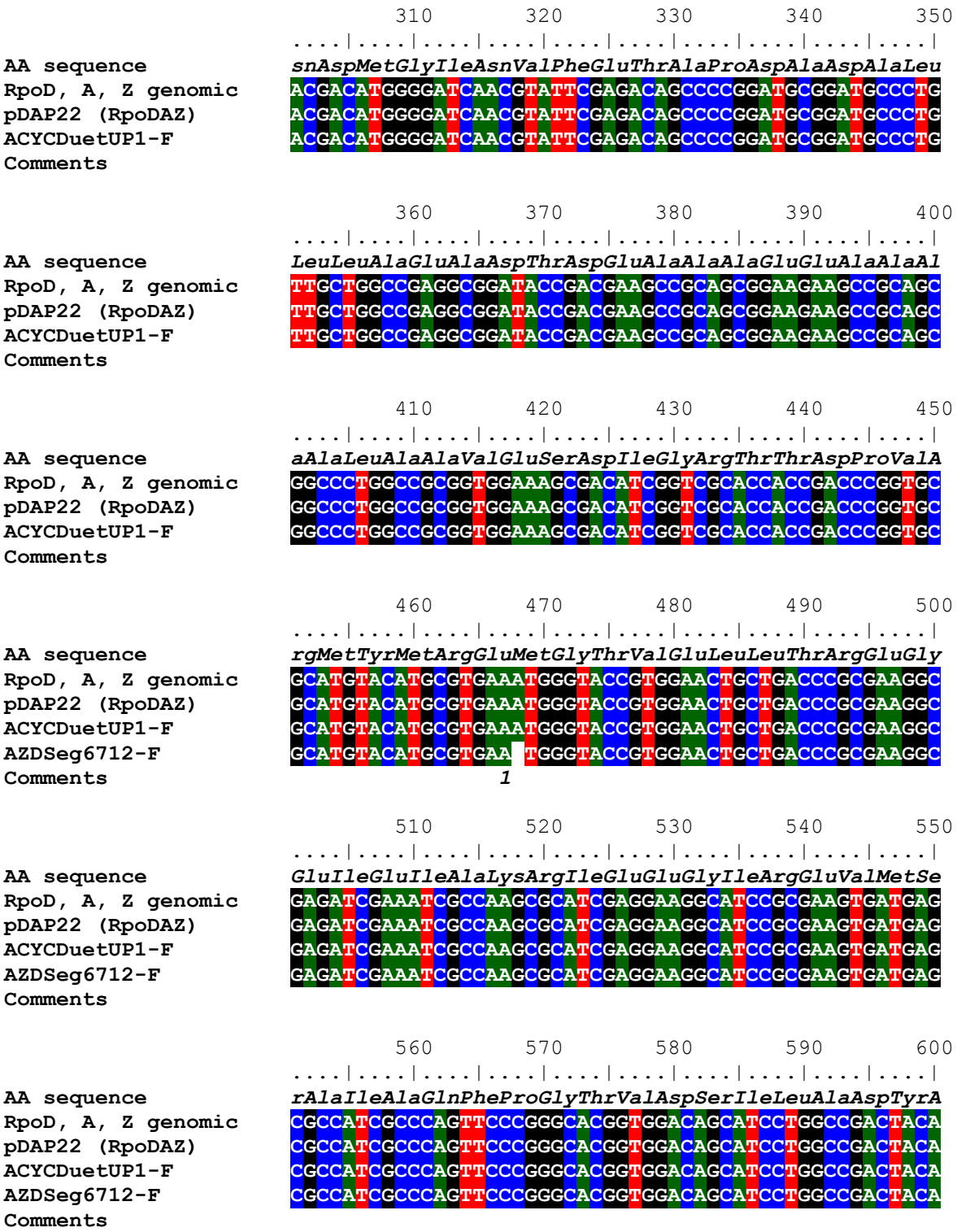

1260 1270 1280 1290 1300
AA sequence
lValGlnGlnGluAspIleValGluAspGluProLeuPheAlaAlaPheA
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
seg5943-F
Comments
1310 1320 1330 1340 1350
AA sequence
spAspGluAlaAsnThrGluAlaLeu
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
seg5943-F
Comments
1360 1370 1380 1390 1400
AA sequence
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
seg5943-F
Comments
1410 1420 1430
AA sequence
rpoA rpoZ Genomic
pDAP10 (RpoAZ)
seg5943-F
Comments
15

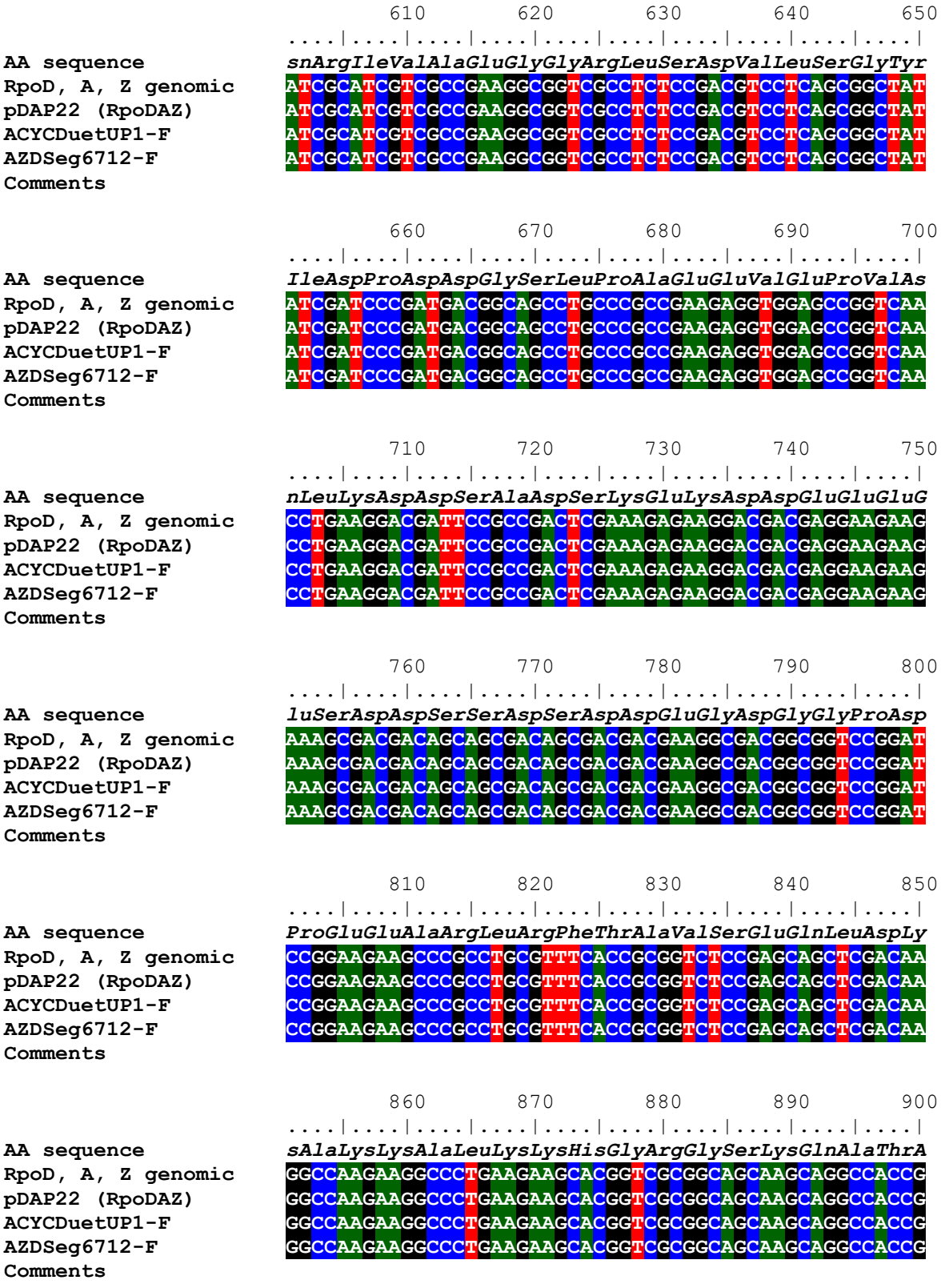
pDAP10 Sequencing Corrections plasmid RpoAZ pET28b

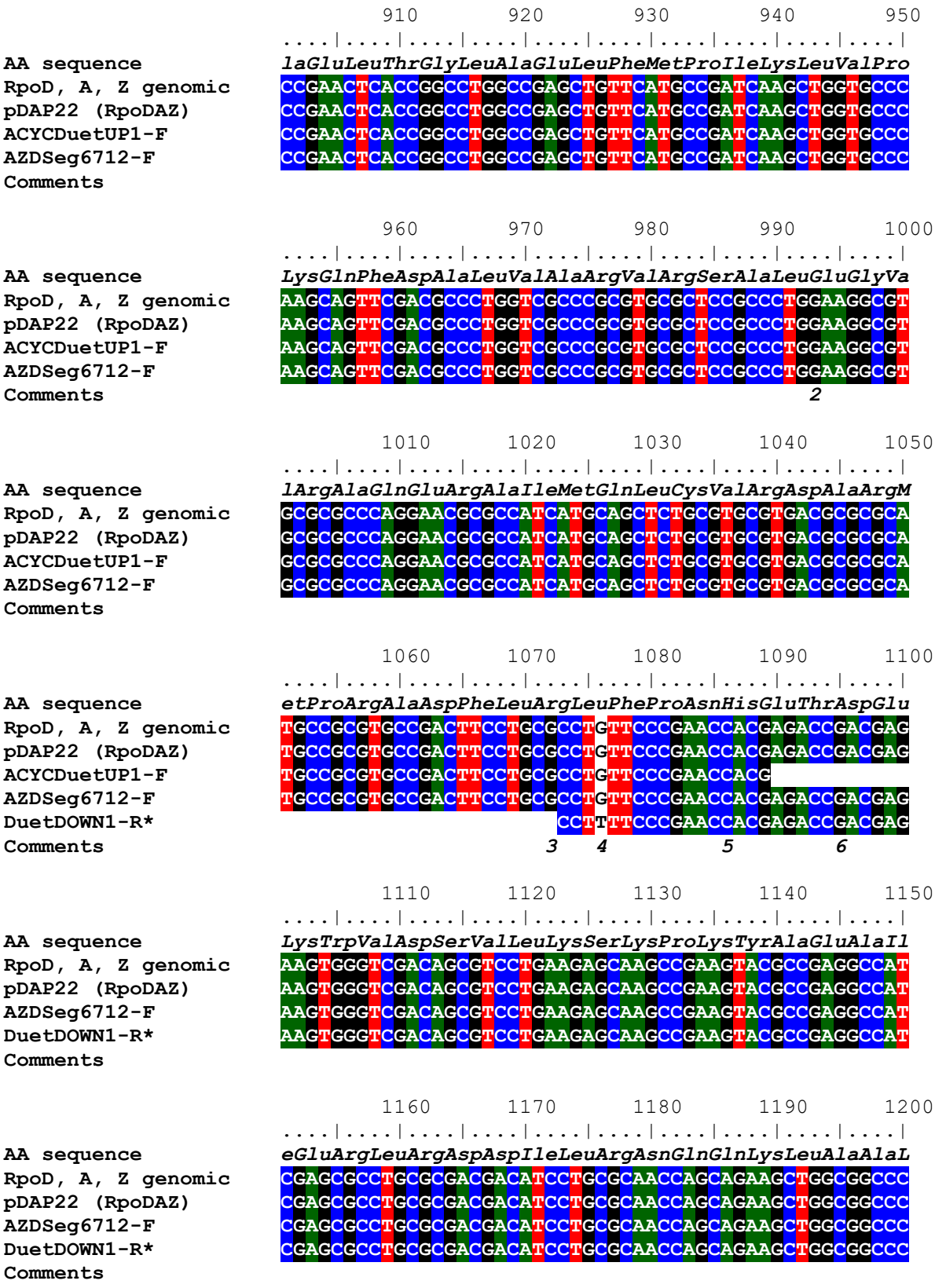
Comment #	**Sequencing primer**	Comments
1	RpoA-Gblock_T7-F	3' 20 nucleotides deleted
2	T7-F	3' 35 nucleotides deleted
3	T7-F	N is clearly a T.
4	T7-F	Wide CC peak makes N a C
5	T7-F	G shoulder makes N consistent with G, but not clear
6	T7-F	Wide A peak and clear C peak make NN consistent with AC
7	T7-F	A doublet supports N as A
8.	T7-F	G has clear shoulder making N a G
	RpoA-Gblock_T7-F	Wide A peak is consistent with two A's instead of triplet. Extra A deleted.
9	T7-F	T peak is present. N is T.
10	T7-F	Sequence is poor. Remaining sequence deleted from 909.
11.	RpoA-Gblock_T7-F	T peak is clearly present
12.	RpoA-Gblock_T7-F	G peak is clearly present
13.	RpoA-Gblock_T7-F	Sequence is poor; peaks spread. Deleted from 1036
14.		Silent mutation introduced in cloning. And weak TGA stop codon replaced with stronger TAA.
15.	seg5943-F	Sequence continues beyond and is consistent with plasmid. Remaining sequence deleted.

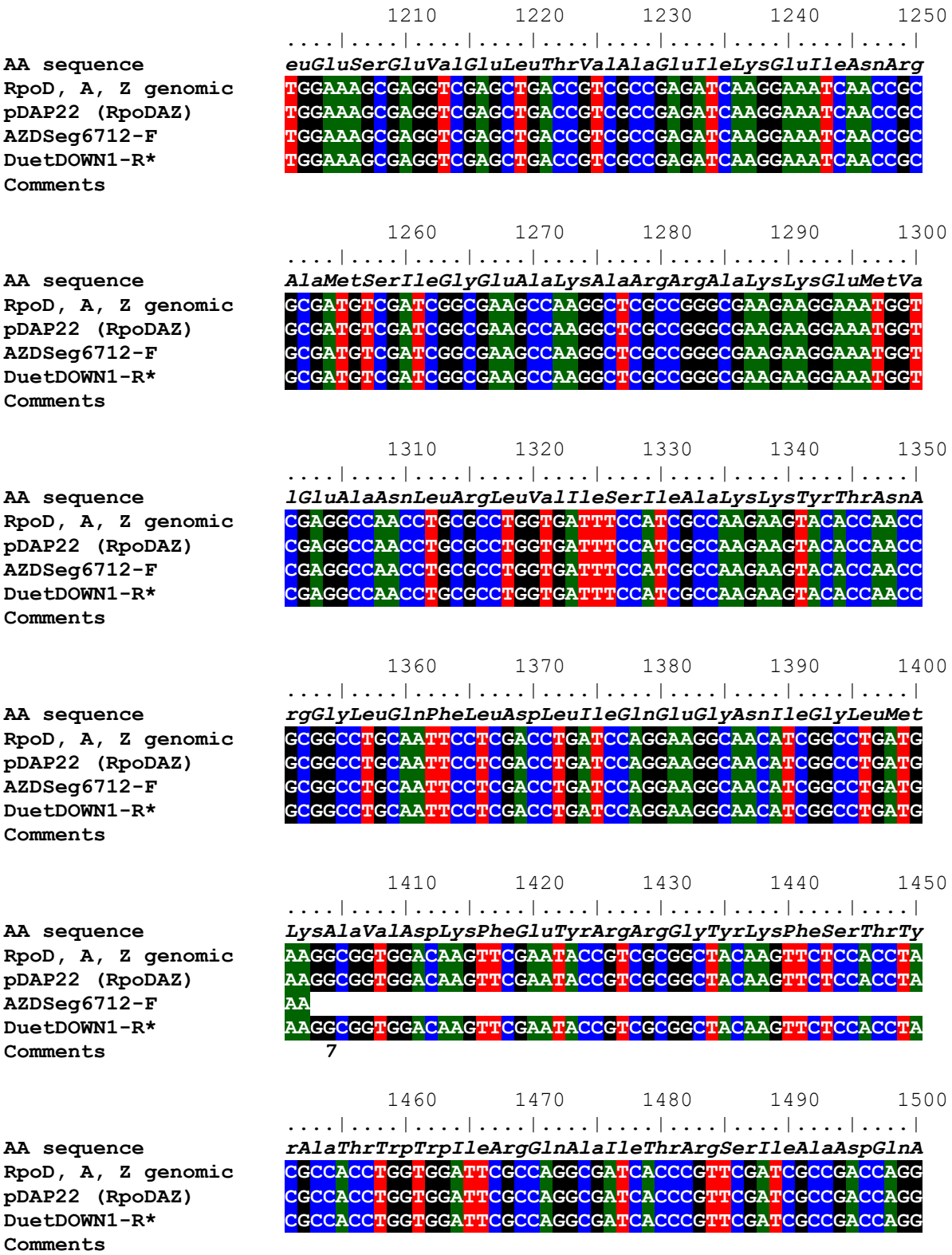
Sequencing analysis of plasmid pDAP22 (RpoAZD Duet)

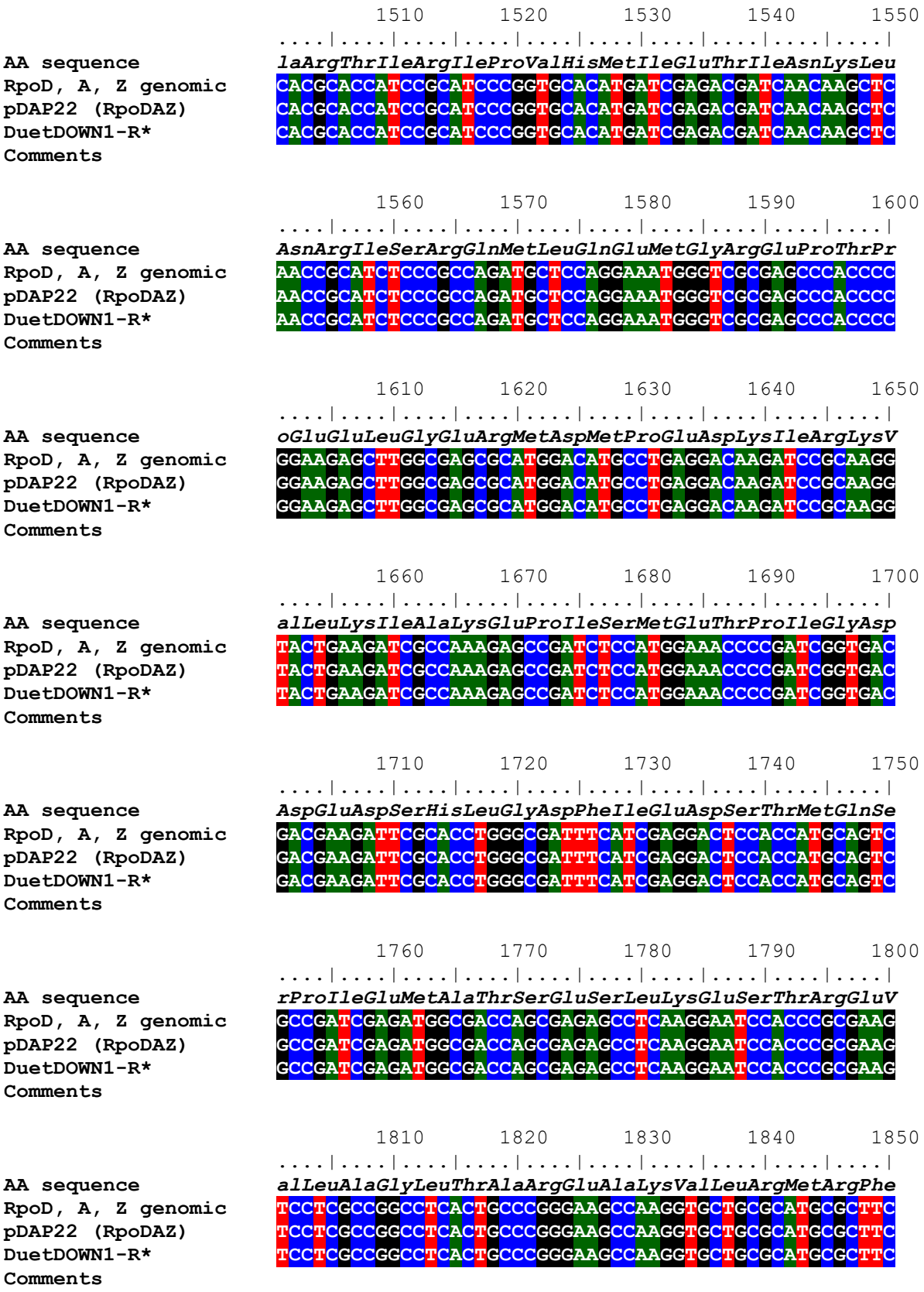


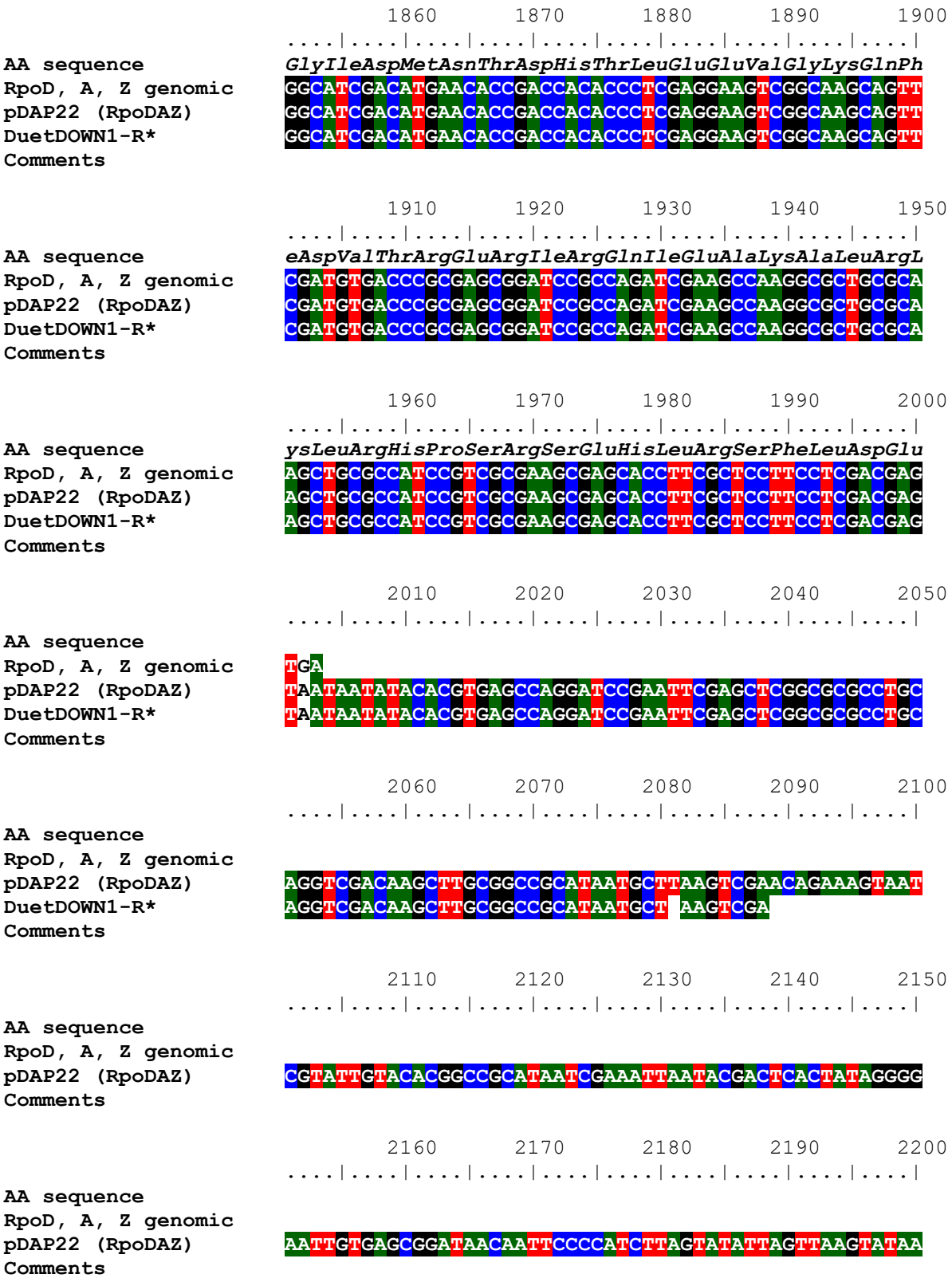

310 320 330 340 350
AA sequence
snAspMetGlyIleAsnValPheGluThrAlaProAspAlaAspAlaLeu
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
Comments
360 370 380 390 400
AA sequence
LeuLeuAlaGluAlaAspThrAspGluAlaAlaAlaGluGluAlaAlaAl
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
Comments
410 420 430 440 450
AA sequence
aAlaLeuAlaAlaValGluSerAspIleGlyArgThrThrAspProValA
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
Comments
460 470 480 490 500
AA sequence
rgMetTyrMetArgGluMetGlyThrValGluLeuLeuThrArgGluGly
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments
1
510 520 530 540 550
AA sequence
GluIleGluIleAlaLysArgIleGluGluGlyIleArgGluValMetSe
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments
560 570 580 590 600
AA sequence
rAlaIleAlaGlnPheProGlyThrValAspSerIleLeuAlaAspTyrA
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments

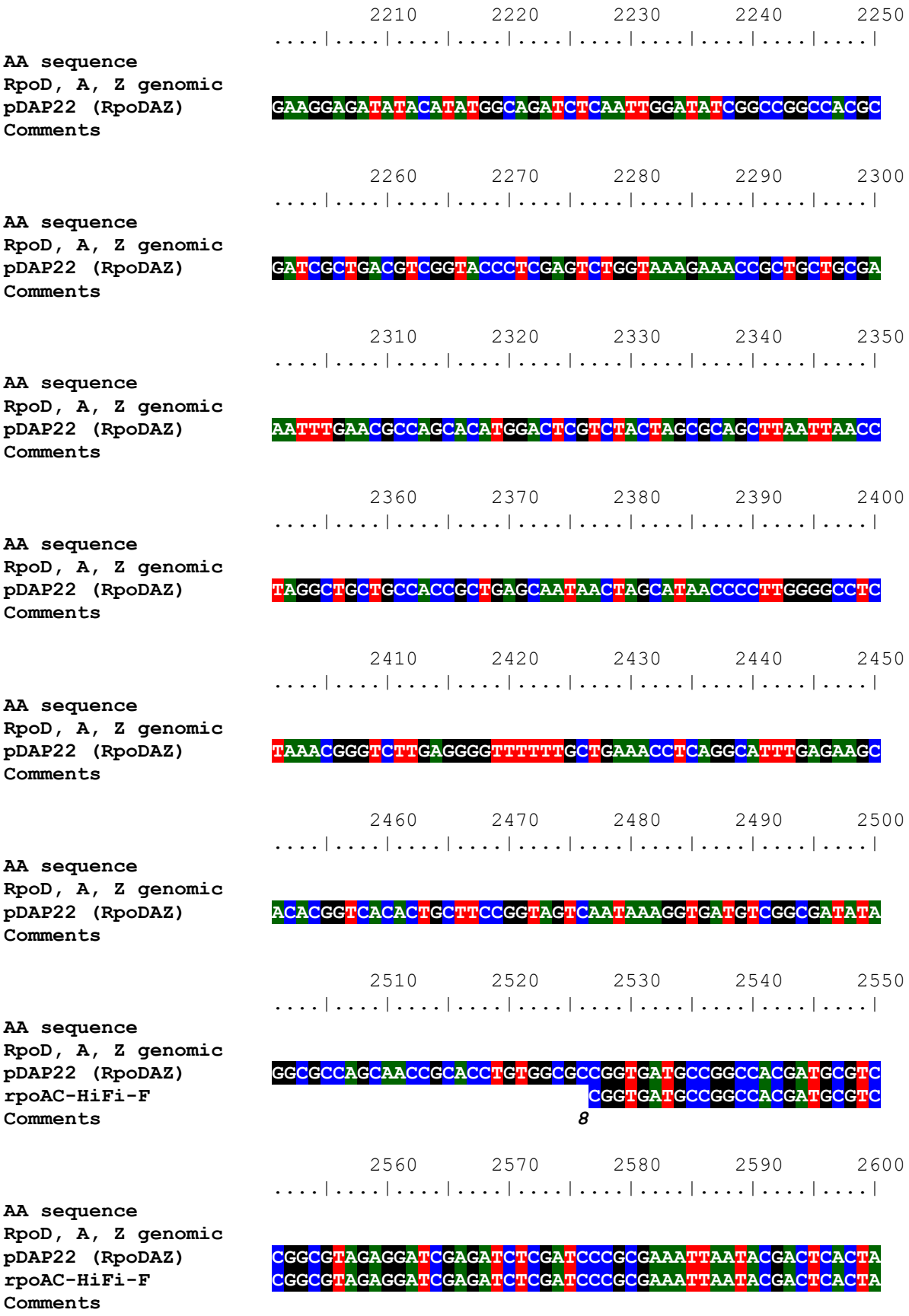

610 620 630 640 650
AA sequence
snArgIleValAlaGluGlyGlyArgLeuSerAspValLeuSerGlyTyr
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments
660 670 680 690 700
AA sequence
IleAspProAspAspGlySerLeuProAlaGluGluValGluProValAs
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments
710 720 730 740 750
AA sequence
nLeuLysAspAspSerAlaAspSerLysGluLysAspAspGluGluGluG
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments
760 770 780 790 800
AA sequence
luSerAspAspSerSerAspSerAspAspGluGlyAspGlyGlyProAsp
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments
810 820 830 840 850
AA sequence
ProGluGluAlaArgLeuArgPheThrAlaValSerGluGlnLeuAspLy
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments
860 870 880 890 900
AA sequence
sAlaLysLysAlaLeuLysLysHisGlyArgGlySerLysGlnAlaThrA
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments

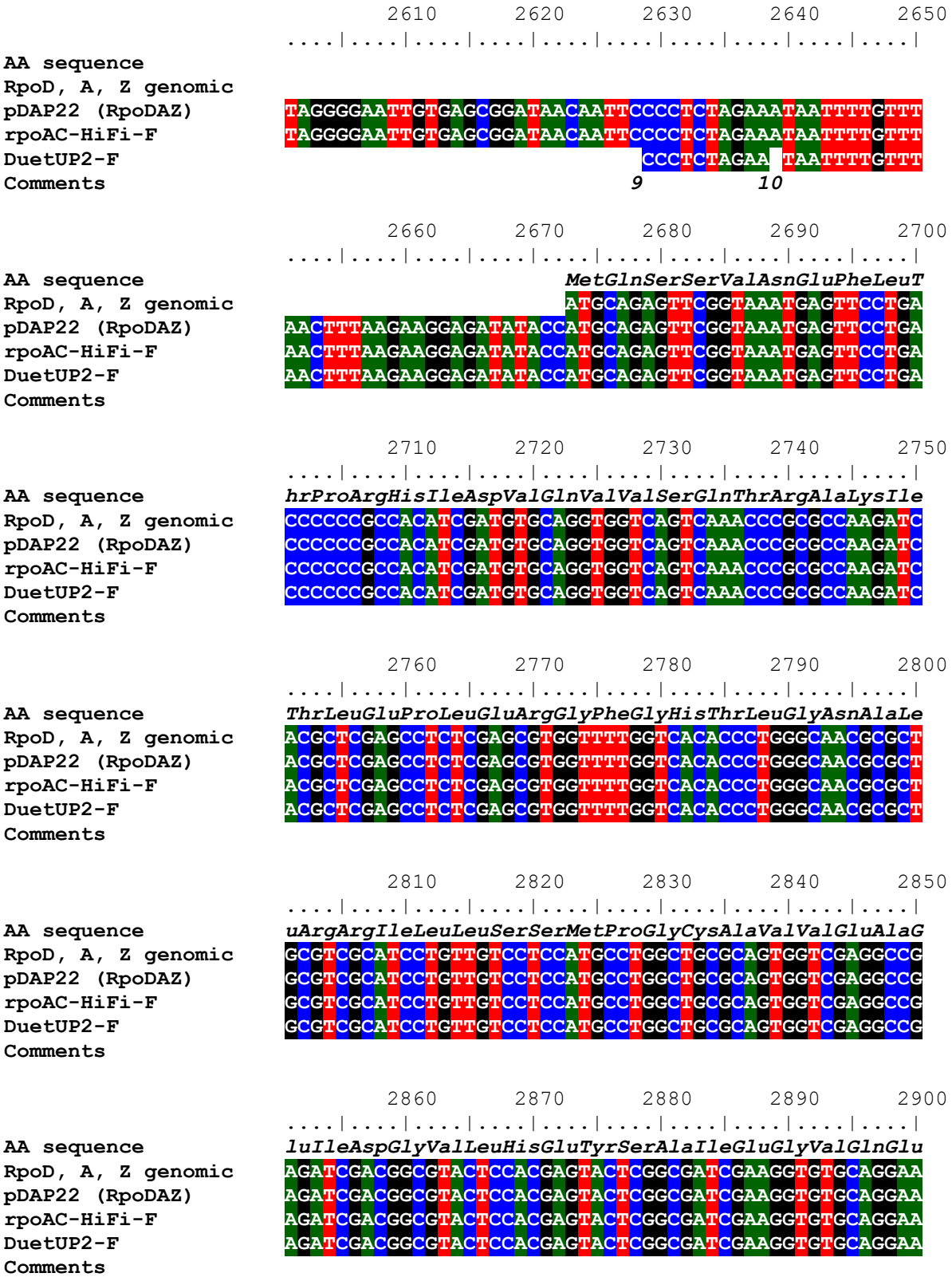

910 920 930 940 950
AA sequence laGluLeuThrGlyLeuAlaGluLeuPheMetProIleLysLeuValPro
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments
960 970 980 990 1000
AA sequence LysGlnPheAspAlaLeuValAlaArgValArgSerAlaLeuGluGlyVa
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments
2
1010 1020 1030 1040 1050
AA sequence lArgAlaGlnGluArgAlaIleMetGlnLeuCysValArgAspAlaArgM
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
Comments
1060 1070 1080 1090 1100
AA sequence etProArgAlaAspPheLeuArgLeuPheProAsnHisGluThrAspGlu
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
ACYCDuetUP1-F
AZDSeg6712-F
DuetDOWN1-R*
Comments
3 4 5 6
1110 1120 1130 1140 1150
AA sequence LysTrpValAspSerValLeuLysSerLysProLysTyrAlaGluAlaIl
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
AZDSeg6712-F
DuetDOWN1-R*
Comments
1160 1170 1180 1190 1200
AA sequence eGluArgLeuArgAspAspIleLeuArgAsnGlnGlnLysLeuAlaAlaL
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
AZDSeg6712-F
DuetDOWN1-R*
Comments

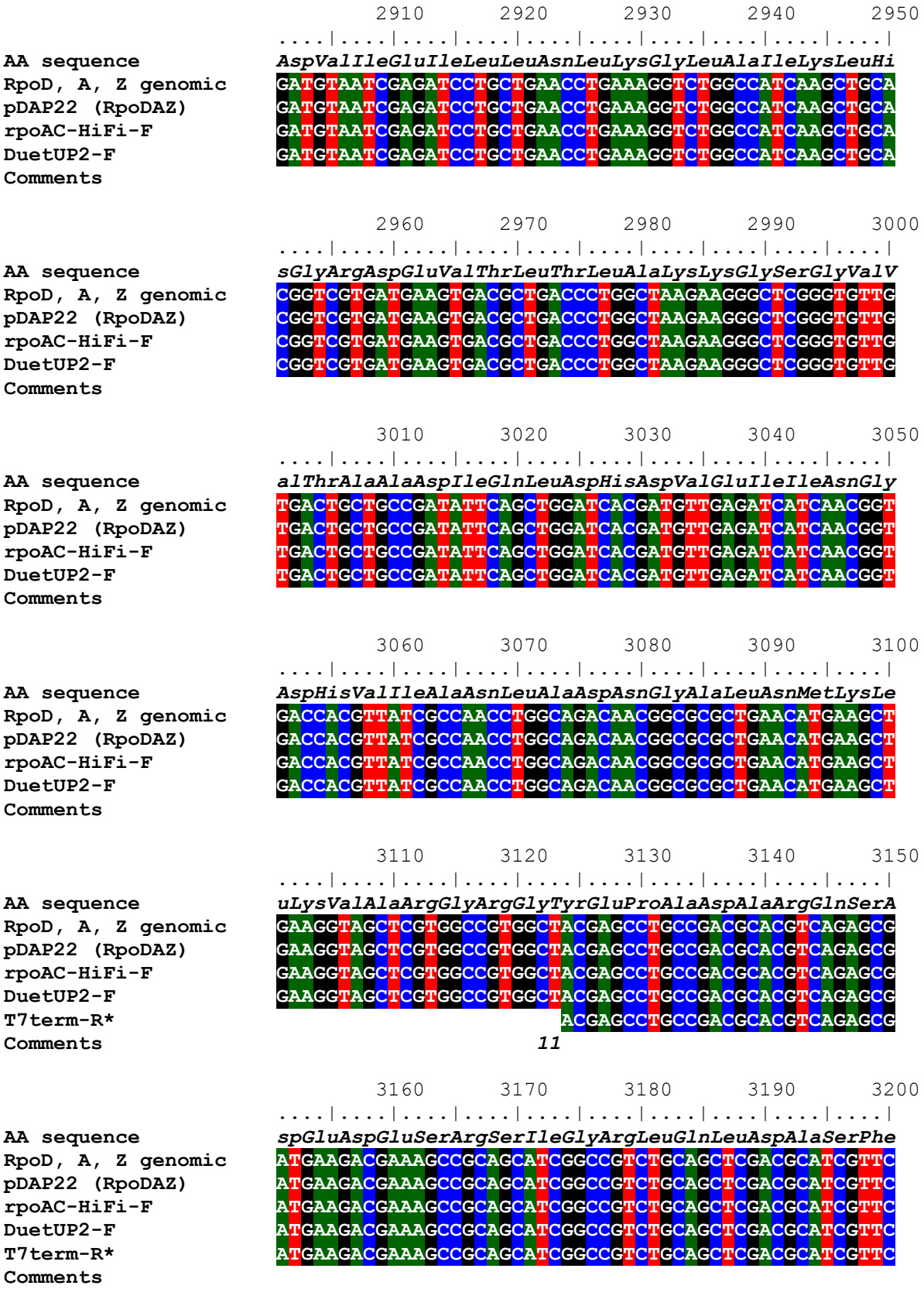

1210 1220 1230 1240 1250
AA sequence
euGluSerGluValGluLeuThrValAlaGluIleLysGluIleAsnArg
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
AZDSeg6712-F
DuetDOWN1-R*
Comments
1260 1270 1280 1290 1300
AA sequence
AlaMetSerIleGlyGluAlaLysAlaArgArgAlaLysLysGluMetVa
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
AZDSeg6712-F
DuetDOWN1-R*
Comments
1310 1320 1330 1340 1350
AA sequence
lGluAlaAsnLeuArgLeuValIleSerIleAlaLysLysTyrThrAsnA
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
AZDSeg6712-F
DuetDOWN1-R*
Comments
1360 1370 1380 1390 1400
AA sequence
rgGlyLeuGlnPheLeuAspLeuIleGlnGluGlyAsnIleGlyLeuMet
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
AZDSeg6712-F
DuetDOWN1-R*
Comments
1410 1420 1430 1440 1450
AA sequence
LysAlaValAspLysPheGluTyrArgArgGlyTyrLysPheSerThrTy
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
AZDSeg6712-F
DuetDOWN1-R*
Comments
7
1460 1470 1480 1490 1500
AA sequence
rAlaThrTrpTrpIleArgGlnAlaIleThrArgSerIleAlaAspGlnA
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments

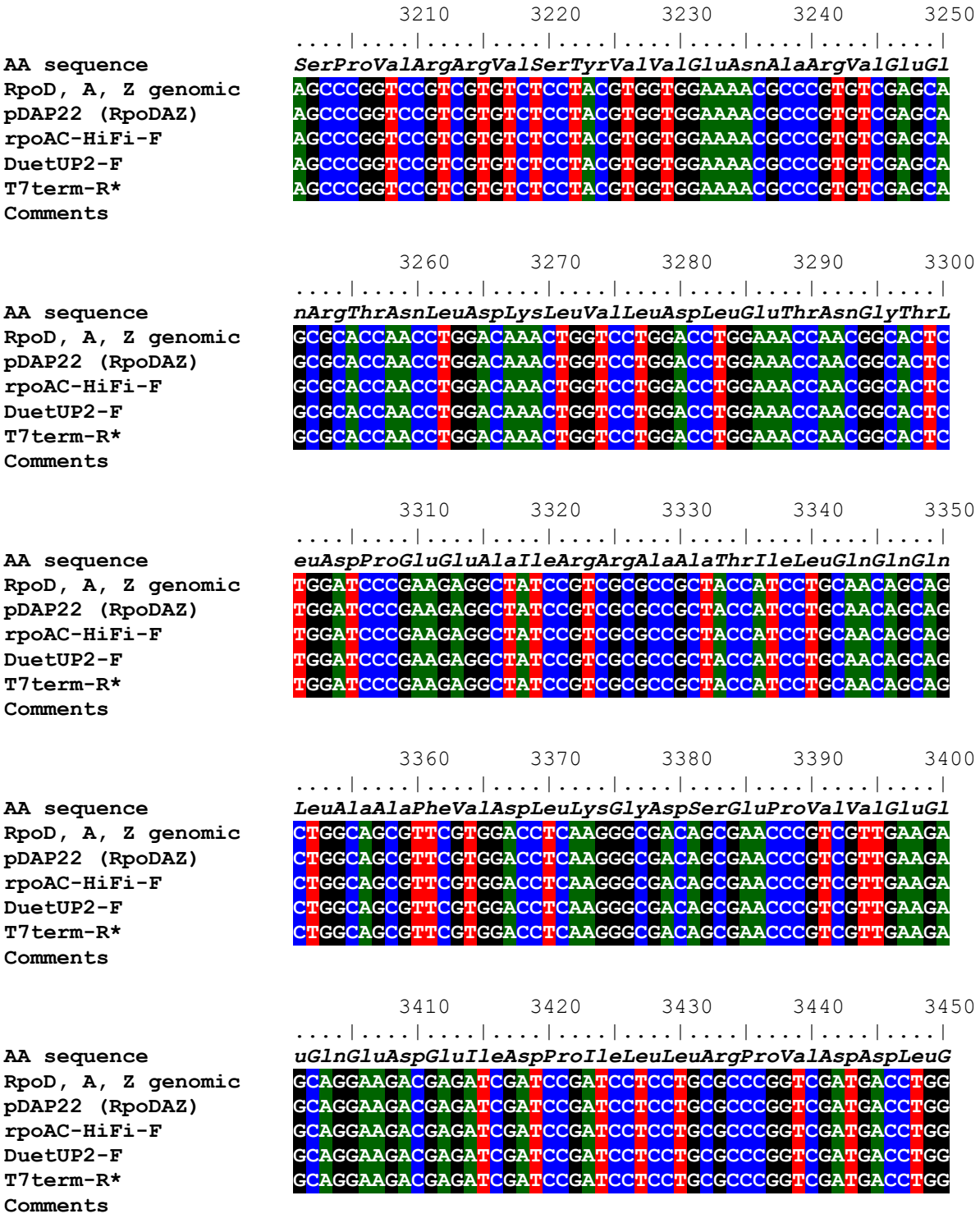

1510 1520 1530 1540 1550
AA sequence
laArgThrIleArgIleProValHisMetIleGluThrIleAsnLysLeu
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments
1560 1570 1580 1590 1600
AA sequence
AsnArgIleSerArgGlnMetLeuGlnGluMetGlyArgGluProThrPr
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments
1610 1620 1630 1640 1650
AA sequence
oGluGluLeuGlyGluArgMetAspMetProGluAspLysIleArgLysV
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments
1660 1670 1680 1690 1700
AA sequence
alLeuLysIleAlaLysGluProIleSerMetGluThrProIleGlyAsp
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments
1710 1720 1730 1740 1750
AA sequence
AspGluAspSerHisLeuGlyAspPheIleGluAspSerThrMetGlnSe
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments
1760 1770 1780 1790 1800
AA sequence
rProIleGluMetAlaThrSerGluSerLeuLysGluSerThrArgGluV
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments
1810 1820 1830 1840 1850
AA sequence
alLeuAlaGlyLeuThrAlaArgGluAlaLysValLeuArgMetArgPhe
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments

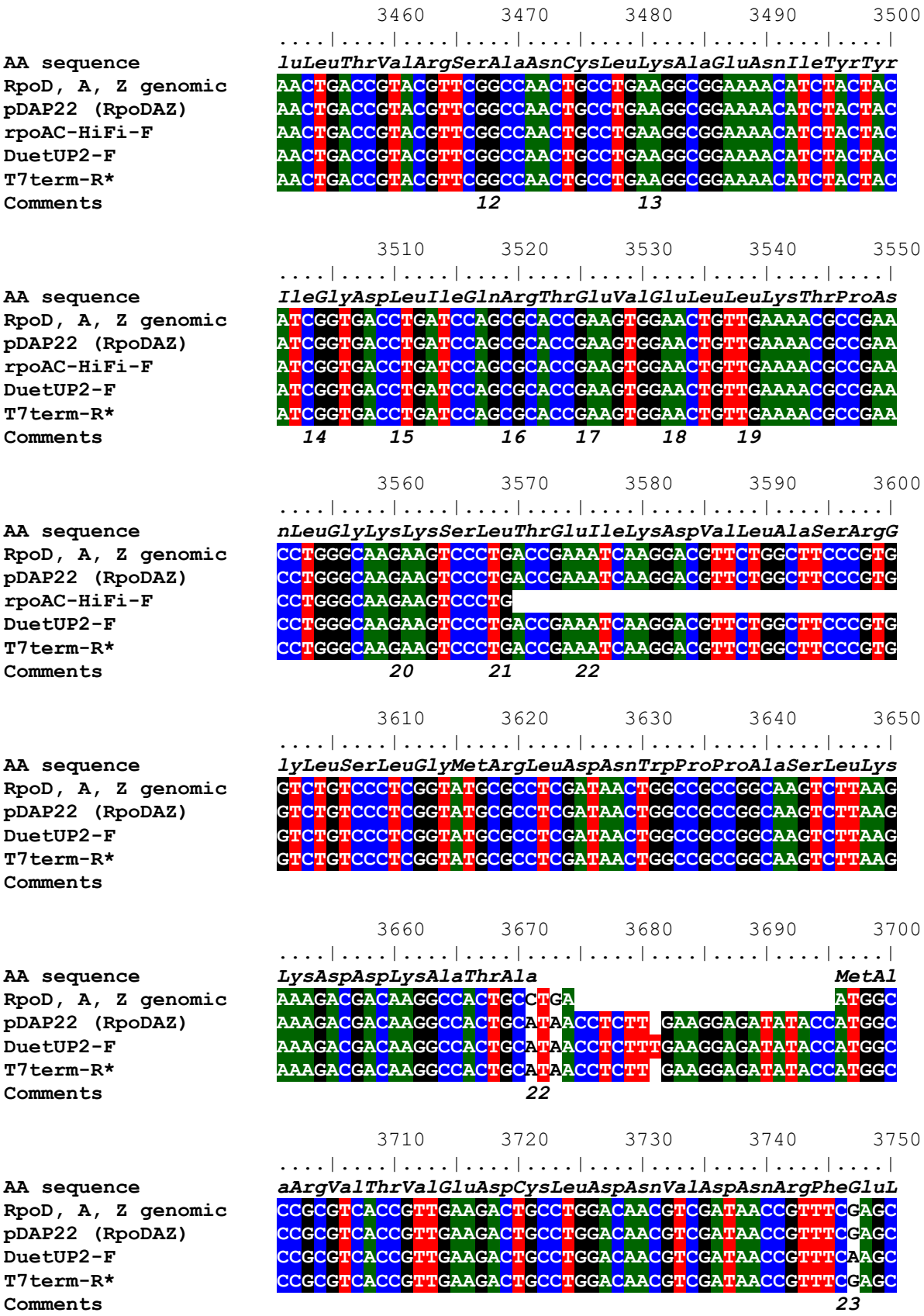

1860 1870 1880 1890 1900
AA sequence
GlyIleAspMetAsnThrAspHisThrLeuGluGluValGlyLysGlnPh
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments
1910 1920 1930 1940 1950
AA sequence
eAspValThrArgGluArgIleArgGlnIleGluAlaLysAlaLeuArgL
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments
1960 1970 1980 1990 2000
AA sequence
ysLeuArgHisProSerArgSerGluHisLeuArgSerPheLeuAspGlu
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetDOWN1-R*
Comments
2010 2020 2030 2040 2050
AA sequence
RpoD, A, Z genomic
TGA
pDAP22 (RpoDAZ)
TAATAATATACACGTGAGCCAGGATCCGAATTCGAGCTCGGCGCGCCTGC
DuetDOWN1-R*
TAATAATATACACGTGAGCCAGGATCCGAATTCGAGCTCGGCGCGCCTGC
Comments
2060 2070 2080 2090 2100
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
AGGTCGACAAGCTTGCGGCCGCATAATGCTTAAGTCGAACAGAAAGTAAT
DuetDOWN1-R*
AGGTCGACAAGCTTGCGGCCGCATAATGCT AAGTCGA
Comments
2110 2120 2130 2140 2150
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
CGTATTGTACACGGCCGCATAATCGAAATTAATACGACTCACTATAGGGG
Comments
2160 2170 2180 2190 2200
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
AATTGTGAGCGGATAACAATTCCCCATCTTAGTATATTAGTTAAGTATAA
Comments

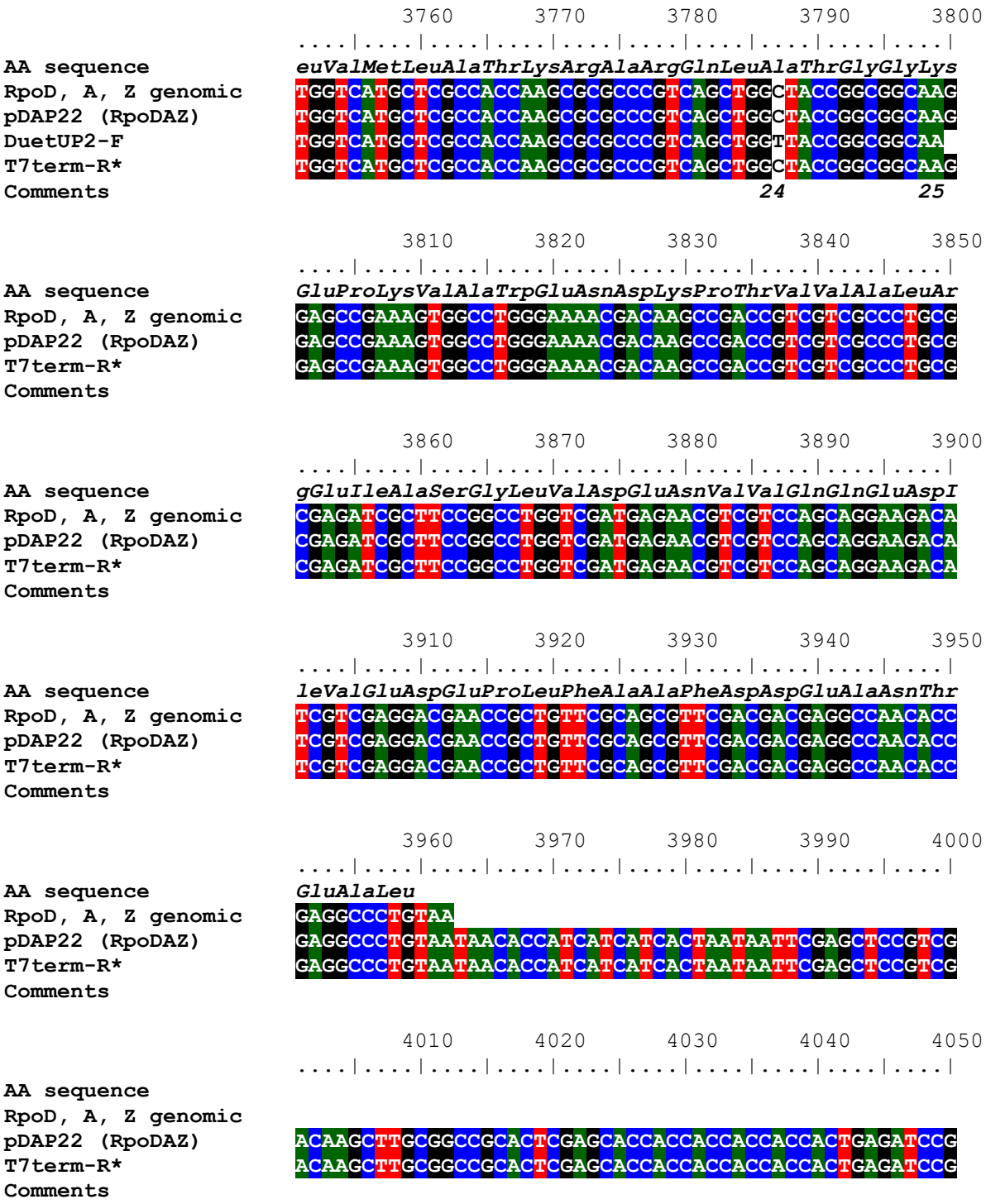

2210 2220 2230 2240 2250
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
Comments
GAAGGAGATATACATATGGCAGATCTCAATTGGATATCGGCCGGCCACGC
2260 2270 2280 2290 2300
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
Comments
GATCGCTGACGTCGGTACCCTCGAGTCTGGTAAAGAAACCGCTGCTGCGA
2310 2320 2330 2340 2350
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
Comments
AATTTGAACGCCAGCACATGGACTCGTCTACTAGCGCAGCTTAATTAACC
2360 2370 2380 2390 2400
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
Comments
TAGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTC
2410 2420 2430 2440 2450
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
Comments
TAAACGGGTCTTGAGGGGTTTTTTGCTGAAACCTCAGGCATTTGAGAAGC
2460 2470 2480 2490 2500
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
Comments
ACACGGTCACACTGCTTCCGGTAGTCAATAAAGGTGATGTCGGCGATATA
2510 2520 2530 2540 2550
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
Comments
GGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTC
CGGTGATGCCGGCCACGATGCGTC
8
2560 2570 2580 2590 2600
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
Comments
CGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
CGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA

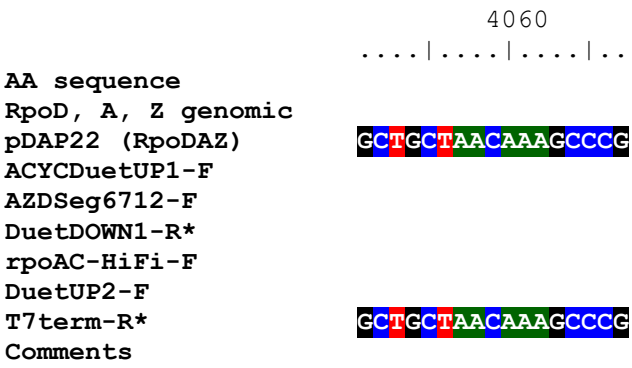

2610 2620 2630 2640 2650
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
Comments
9 10
2660 2670 2680 2690 2700
AA sequence
MetGlnSerSerValAsnGluPheLeuT
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
Comments
2710 2720 2730 2740 2750
AA sequence
hrProArgHisIleAspValGlnValValSerGlnThrArgAlaLysIle
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
Comments
2760 2770 2780 2790 2800
AA sequence
ThrLeuGluProLeuGluArgGlyPheGlyHisThrLeuGlyAsnAlaLe
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
Comments
2810 2820 2830 2840 2850
AA sequence
uArgArgIleLeuLeuSerSerMetProGlyCysAlaValValGluAlaG
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
Comments
2860 2870 2880 2890 2900
AA sequence
luIleAspGlyValLeuHisGluTyrSerAlaIleGluGlyValGlnGlu
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
Comments


2910 2920 2930 2940 2950
AA sequence
AspValIleGluIleLeuLeuAsnLeuLysGlyLeuAlaIleLysLeuHi
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
Comments
GATGTAATCGAGATCCTGCTGAACCTGAAAGGTCTGGCCATCAAGCTGCA
2960 2970 2980 2990 3000
AA sequence
sGlyArgAspGluValThrLeuThrLeuAlaLysLysGlySerGlyValV
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
Comments
CGGTCGTGATGAAGTGACGCTGACCCTGGCTAAGAAGGGCTCGGGTGTTG
3010 3020 3030 3040 3050
AA sequence
alThrAlaAlaAspIleGlnLeuAspHisAspValGluIleIleAsnGly
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
Comments
TGACTGCTGCCGATATTCAGCTGGATCACGATGTTGAGATCATCAACGGT
3060 3070 3080 3090 3100
AA sequence
AspHisValIleAlaAsnLeuAlaAspAsnGlyAlaLeuAsnMetLysLe
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
Comments
GACCACGTTATCGCCAACCTGGCAGACAACGGCGCGCTGAACATGAAGCT
3110 3120 3130 3140 3150
AA sequence
uLysValAlaArgGlyArgGlyTyrGluProAlaAspAlaArgGlnSerA
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
T7term-R*
Comments
GAAGGTAGCTCGTGGCCGTGGCTACGAGCCTGCCGACGCACGTCAGAGCG
ACGAGCCTGCCGACGCACGTCAGAGCG
11
3160 3170 3180 3190 3200
AA sequence
spGluAspGluSerArgSerIleGlyArgLeuGlnLeuAspAlaSerPhe
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
T7term-R*
Comments
ATGAAGACGAAAGCCGCAGCATCGGCCGTCTGCAGCTCGACGCATCGTTC


3210 3220 3230 3240 3250
AA sequence
SerProValArgArgValSerTyrValValGluAsnAlaArgValGluGl
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
T7term-R*
Comments
3260 3270 3280 3290 3300
AA sequence
nArgThrAsnLeuAspLysLeuValLeuAspLeuGluThrAsnGlyThrL
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
T7term-R*
Comments
3310 3320 3330 3340 3350
AA sequence
euAspProGluGluAlaIleArgArgAlaAlaThrIleLeuGlnGlnGln
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
T7term-R*
Comments
3360 3370 3380 3390 3400
AA sequence
LeuAlaAlaPheValAspLeuLysGlyAspSerGluProValValGluGl
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
T7term-R*
Comments
3410 3420 3430 3440 3450
AA sequence
uGlnGluAspGluIleAspProIleLeuLeuArgProValAspAspLeuG
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
T7term-R*
Comments


3460 3470 3480 3490 3500
AA sequence
luLeuThrValArgSerAlaAsnCysLeuLysAlaGluAsnIleTyrTyr
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
T7term-R*
Comments
12
13
3510 3520 3530 3540 3550
AA sequence
IleGlyAspLeuIleGlnArgThrGluValGluLeuLeuLysThrProAs
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
T7term-R*
Comments
14
15
16
17
18
19
3560 3570 3580 3590 3600
AA sequence
nLeuGlyLysLysSerLeuThrGluIleLysAspValLeuAlaSerArgG
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
rpoAC-HiFi-F
DuetUP2-F
T7term-R*
Comments
20
21
22
3610 3620 3630 3640 3650
AA sequence
lyLeuSerLeuGlyMetArgLeuAspAsnTrpProProAlaSerLeuLys
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetUP2-F
T7term-R*
Comments
3660 3670 3680 3690 3700
AA sequence
LysAspAspLysAlaThrAla
MetAl
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetUP2-F
T7term-R*
Comments
22
3710 3720 3730 3740 3750
AA sequence
aArgValThrValGluAspCysLeuAspAsnValAspAsnArgPheGluL
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetUP2-F
T7term-R*
Comments
23


3760 3770 3780 3790 3800
AA sequence
euValMetLeuAlaThrLysArgAlaArgGlnLeuAlaThrGlyGlyLys
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
DuetUP2-F
T7term-R*
Comments
24
25
3810 3820 3830 3840 3850
AA sequence
GluProLysValAlaTrpGluAsnAspLysProThrValValAlaLeuAr
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
T7term-R*
Comments
3860 3870 3880 3890 3900
AA sequence
gGluIleAlaSerGlyLeuValAspGluAsnValValGlnGlnGluAspI
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
T7term-R*
Comments
3910 3920 3930 3940 3950
AA sequence
leValGluAspGluProLeuPheAlaAlaPheAspAspGluAlaAsnThr
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
T7term-R*
Comments
3960 3970 3980 3990 4000
AA sequence
GluAlaLeu
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
T7term-R*
Comments
4010 4020 4030 4040 4050
AA sequence
RpoD, A, Z genomic
pDAP22 (RpoDAZ)
T7term-R*
Comments

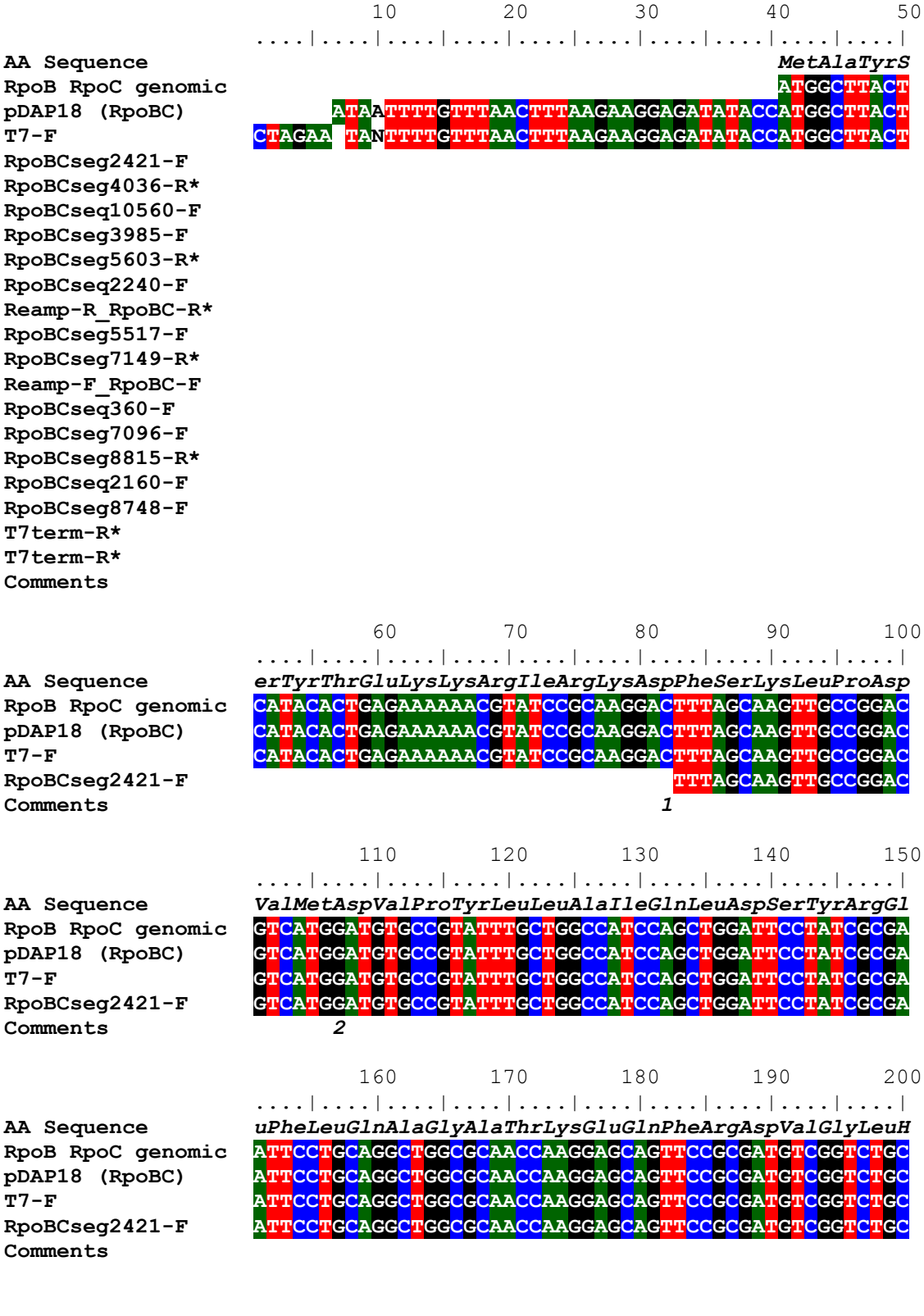


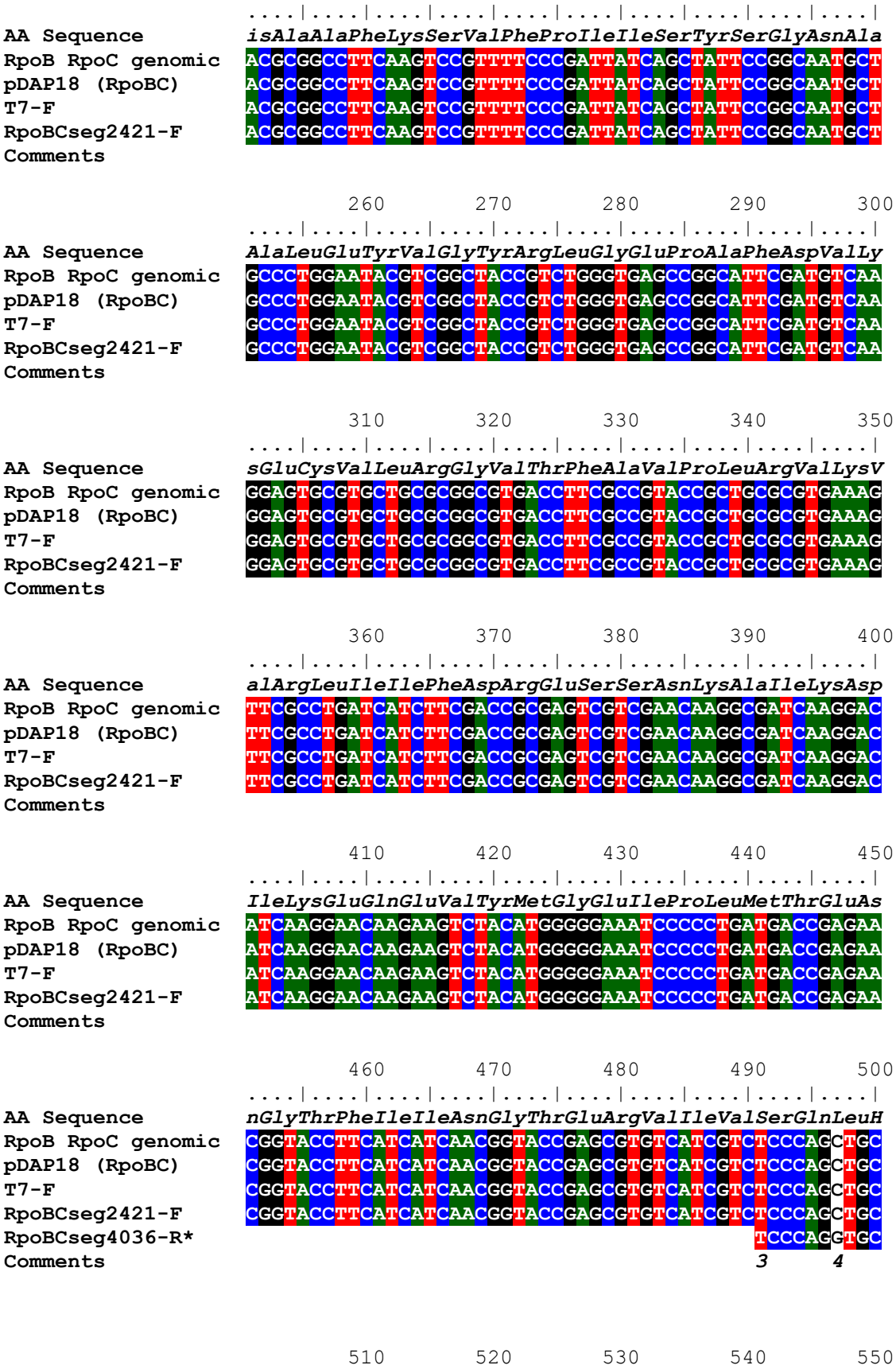
pDAP22 Sequencing Corrections plasmid RpoAZD

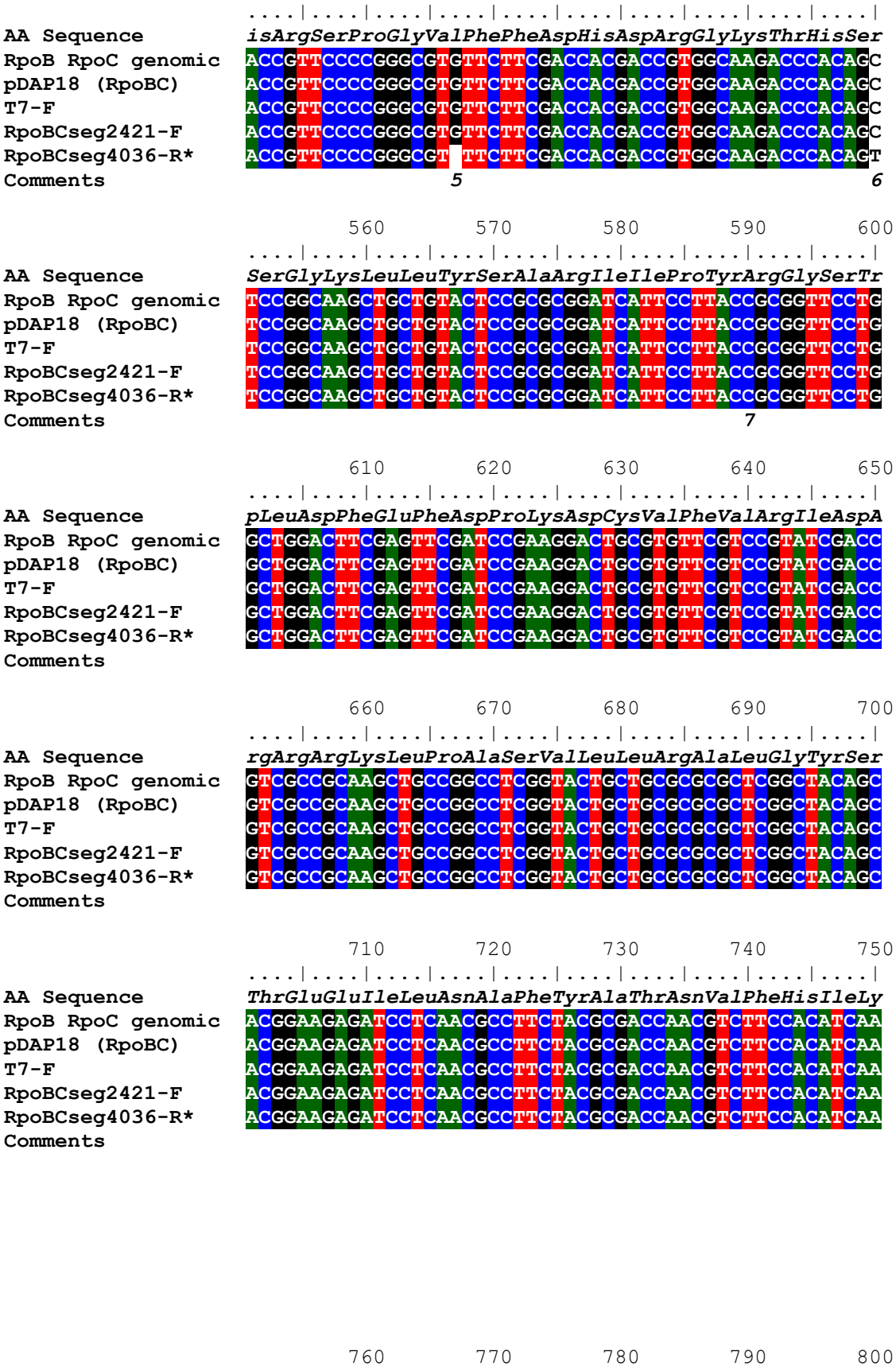
Comment #	Sequencing primer	Comments
1	AZDSeg6712-F	AA doublet not consistent with sequence, AAA is correct
2	ACYCDuetUp1-F	(995) triplet A peak is evident, but-First A is a shoulder. A deleted
3	DuetDOWN1-R*	peaks are smeared past 1011, removed last 90 nucleotides (beginning of sequence in reverse complement)
4	DuetDOWN1-R*	C (G*) is clearly present
5	DuetDOWN1-R*	doublet G (not triplet). C* deleted
6	DuetDOWN1-R*	triplet G appears correct, but is not send in AZDSeq6712
7	AZDSeg6712-F	peaks are now speary and many misreads (952 in sequence), remaining sequence deleted (most were correctly called)
8.	rpoAC-HiFi-F	first 13 nucleotides not read well, deleted
9	DuetUP2-F	first 38 nucleotides not read well, deleted
10	rpoAC-HiFi-F	extra G removed
10.	DuetUP2-F	A peak is smeared, consistent with a missing A
11.	T7term-R*	nucleotides removed after 960
12.	rpoAC-HiFi-F	run of Gs peaks smeared at 981, two G's probable. G removed
13.	rpoAC-HiFi-F	run of three A's last A peak is shoulder. A deleted
14.	rpoAC-HiFi-F	C doublet peak is barely split and appears compressed. Is consistent with single C. Extra C deleted.
15.	rpoAC-HiFi-F	G single peak is spread. Sequence is compressed. Consistent with single G in correct sequence.
16	rpoAC-HiFi-F	G single peak is spread. Sequence is compressed. Consistent with single G in correct sequence.
17	rpoAC-HiFi-F	A triplet has a double peak. Sequence is compressed. Consistent with two As in correct sequence.
18	rpoAC-HiFi-F	A triplet has a double peak. Sequence is compressed. Consistent with two As in correct sequence.

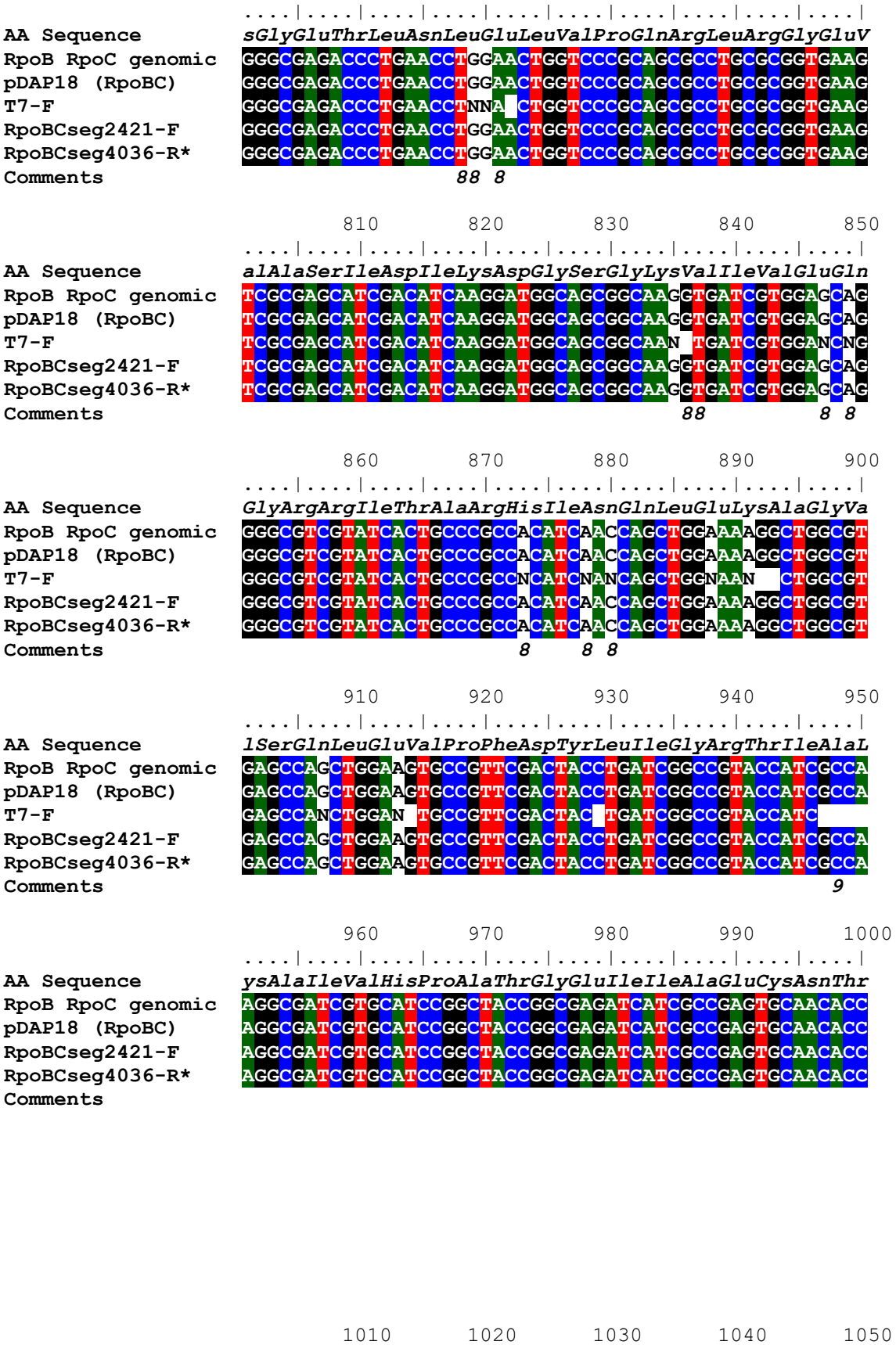
19	rpoAC-HiFi-F	G doublet has a single peak, but extended. Sequence is consistent with one G. G delected
20	rpoAC-HiFi-F	A quartet has a double peak, but extended. Sequence is compressed-For-Four A's. Consistent with three A's in correct sequence. A delected
21	rpoAC-HiFi-F	many sequence errors after 1092. Nucleotides after 1092 deleted.
22		TAA stop codon replaced TGA stop codon. Silent mutation encoding ALA, GCC>GCA, introduced during cloning.
23	DuetUP2-F	Double A peak has G peak under, so correct GAG sequence is compatible, but unclear.
24	DuetUP2-F	TT is single peak with small C peak under.
25	DuetUP2-F	trace has many errors. Deleted after 1181

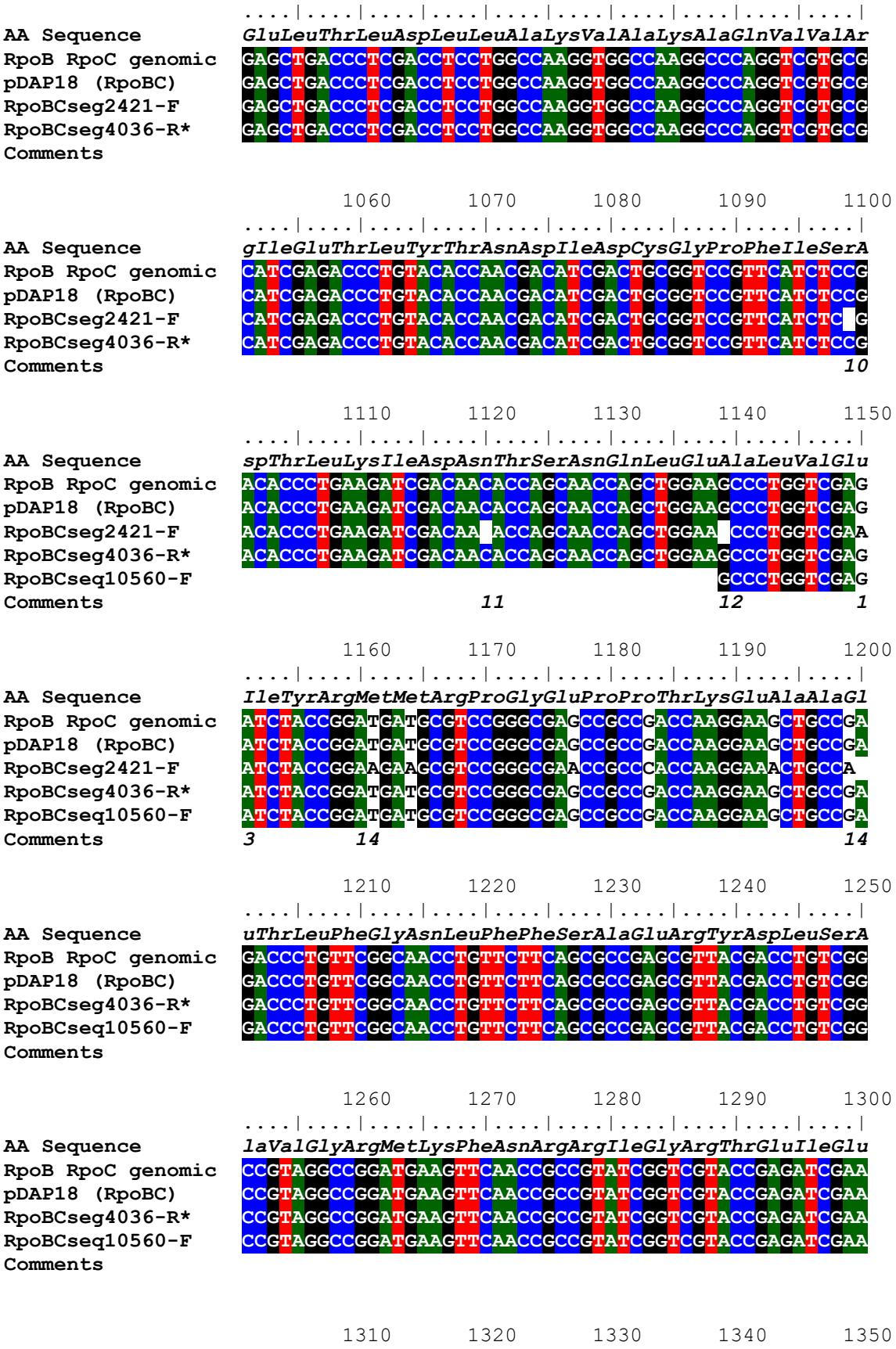
Sequencing analysis of plasmid pDAP18 (pRARE RpoBC)

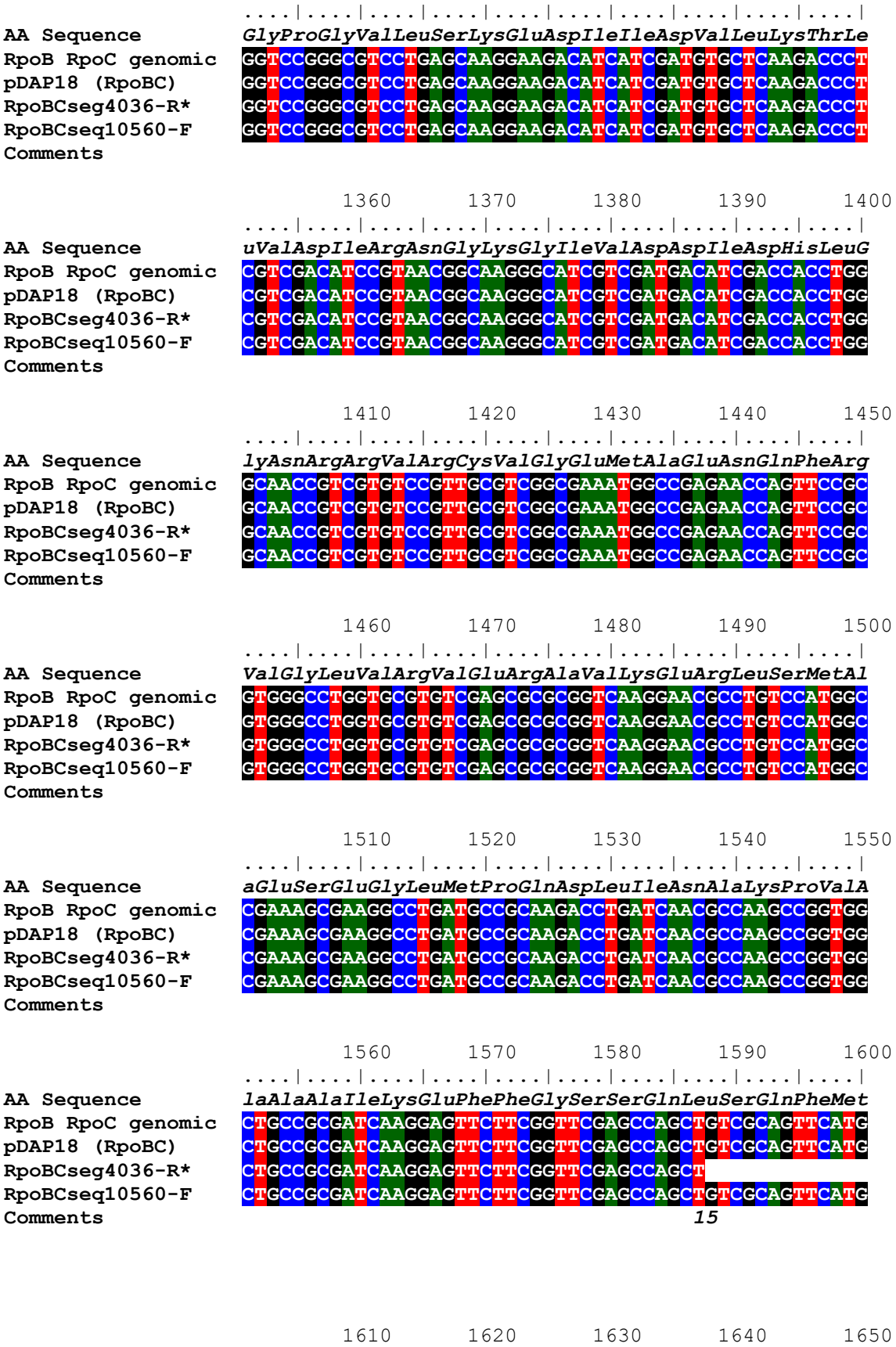


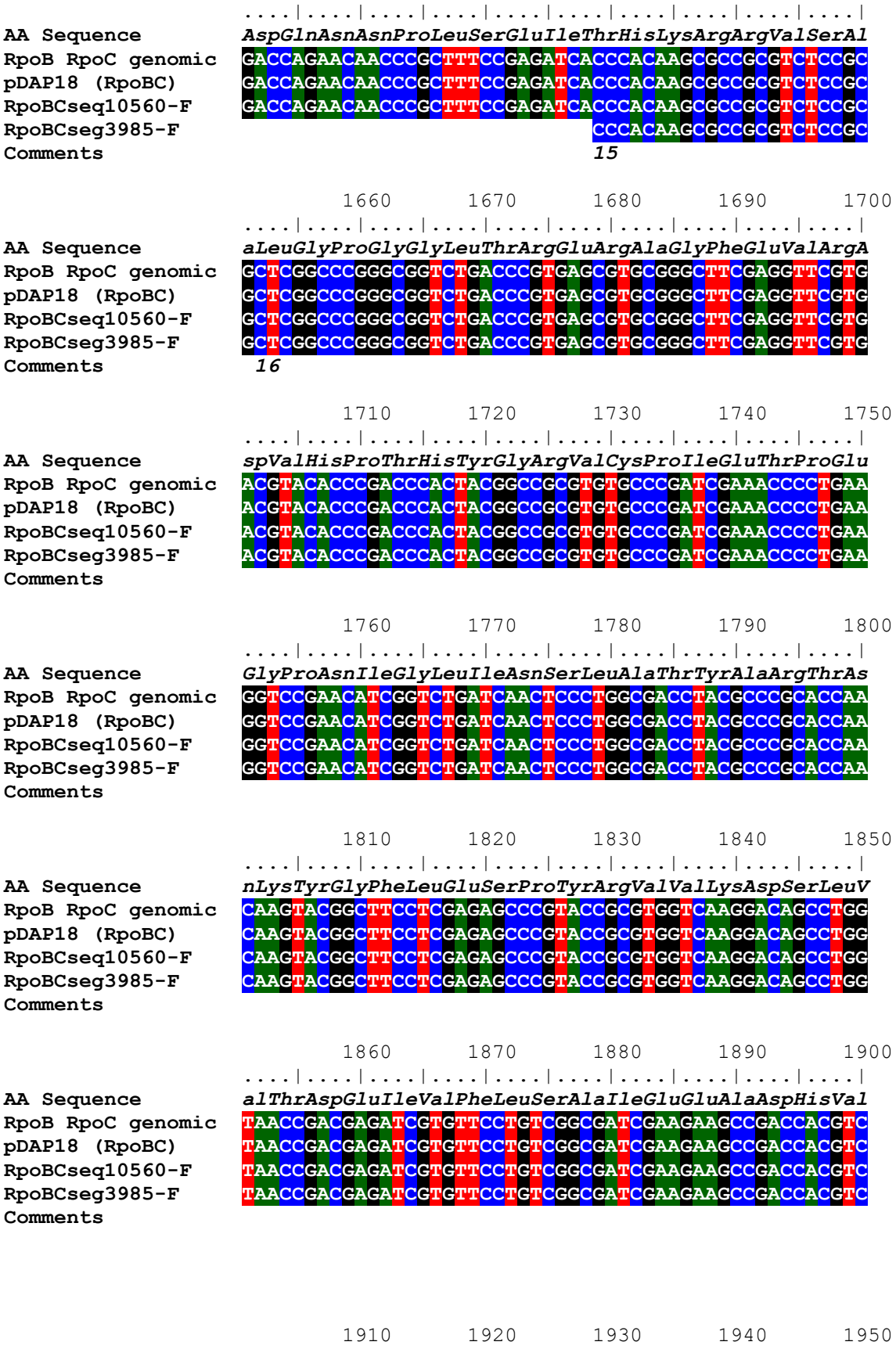

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
T7-F
RpoBCseg2421-F
RpoBCseg4036-R*
Comments
isAlaAlaPheLysSerValPheProIleIleSerTyrSerGlyAsnAla
AlaLeuGluTyrValGlyTyrArgLeuGlyGluProAlaPheAspValLy
sGluCysValLeuArgGlyValThrPheAlaValProLeuArgValLysV
alArgLeuIleIlePheAspArgGluSerSerAsnLysAlaIleLysAsp
IleLysGluGlnGluValTyrMetGlyGluIleProLeuMetThrGluAs
nGlyThrPheIleIleAsnGlyThrGluArgValIleValSerGlnLeuH
3
4
260
270
280
290
300
310
320
330
340
350
360
370
380
390
400
410
420
430
440
450
460
470
480
490
500
510
520
530
540
550

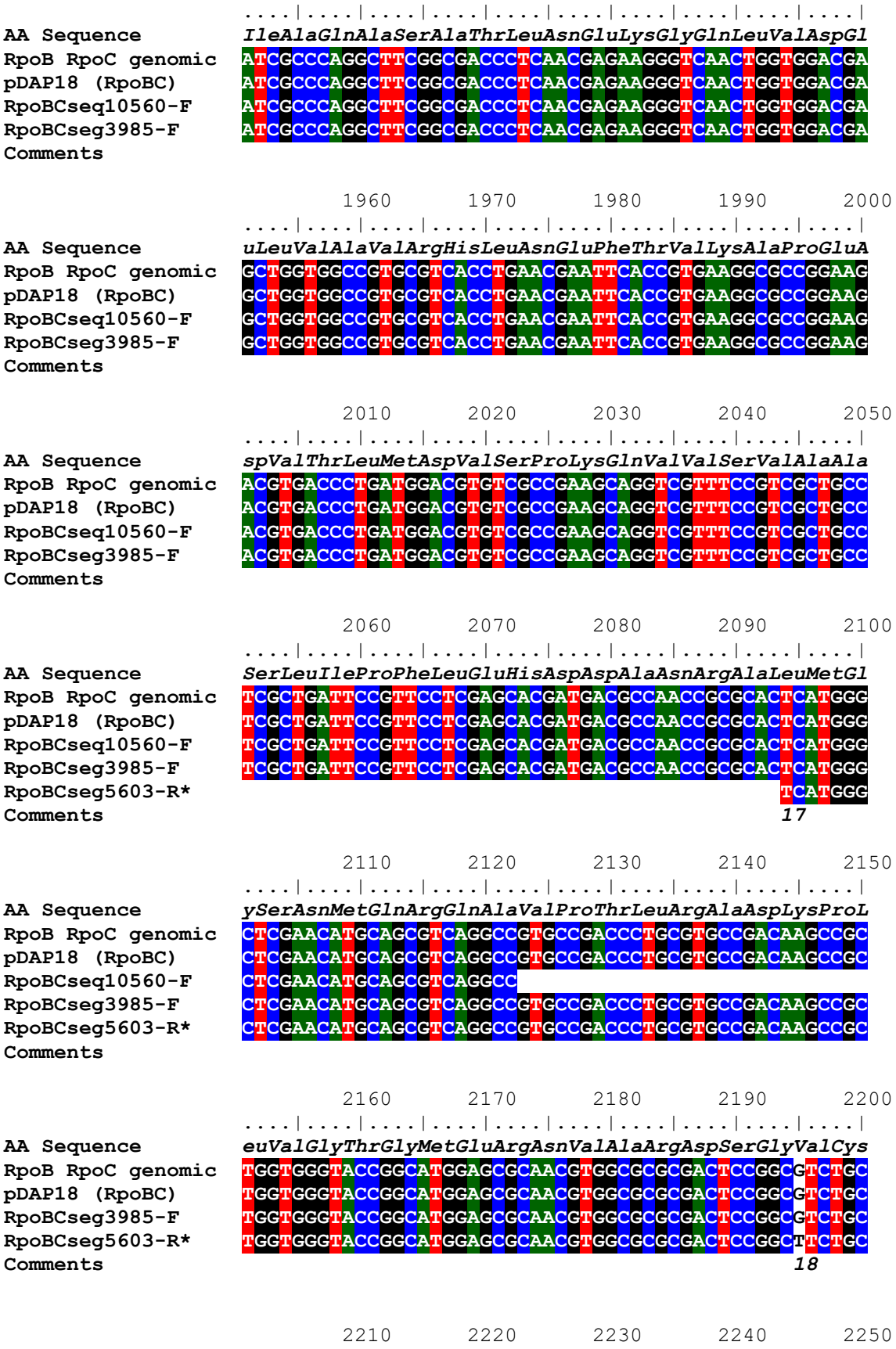

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
T7-F
RpoBCseg2421-F
RpoBCseg4036-R*
Comments
isArgSerProGlyValPhePheAspHisAspArgGlyLysThrHisSer
5
6
560
570
580
590
600
SerGlyLysLeuLeuTyrSerAlaArgIleIleProTyrArgGlySerTr
7
610
620
630
640
650
pLeuAspPheGluPheAspProLysAspCysValPheValArgIleAspA
660
670
680
690
700
rgArgArgLysLeuProAlaSerValLeuLeuArgAlaLeuGlyTyrSer
710
720
730
740
750
ThrGluGluIleLeuAsnAlaPheTyrAlaThrAsnValPheHisIleLy
760
770
780
790
800

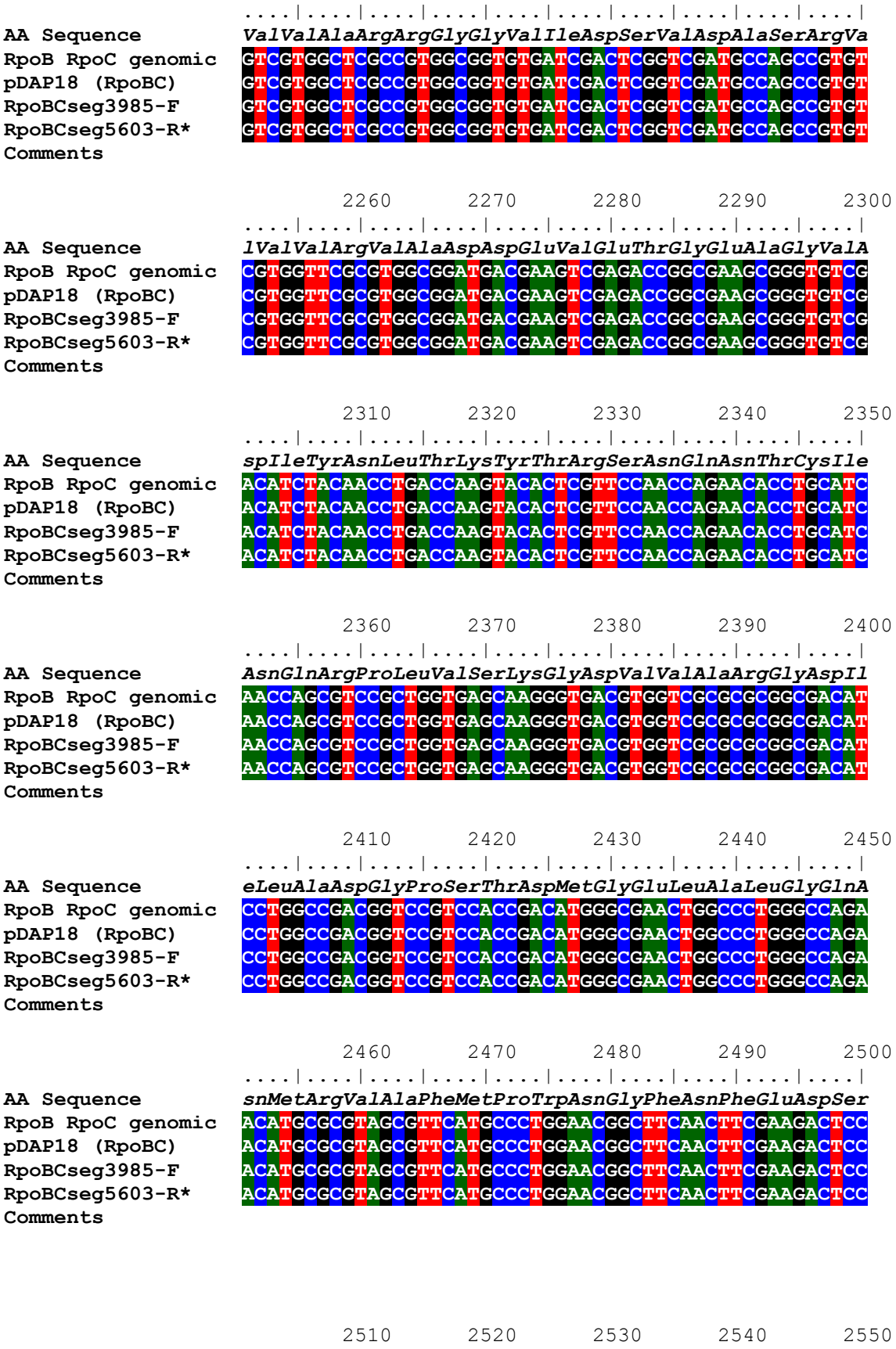

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
T7-F
RpoBCseg2421-F
RpoBCseg4036-R*
Comments
sGlyGluThrLeuAsnLeuGluLeuValProGlnArgLeuArgGlyGluV
alAlaSerIleAspIleLysAspGlySerGlyLysValIleValGluGln
GlyArgArgIleThrAlaArgHisIleAsnGlnLeuGluLysAlaGlyVa
lSerGlnLeuGluValProPheAspTyrLeuIleGlyArgThrIleAlaL
ysAlaIleValHisProAlaThrGlyGluIleIleAlaGluCysAsnThr
810
820
830
840
850
860
870
880
890
900
910
920
930
940
950
960
970
980
990
1000
1010
1020
1030
1040
1050

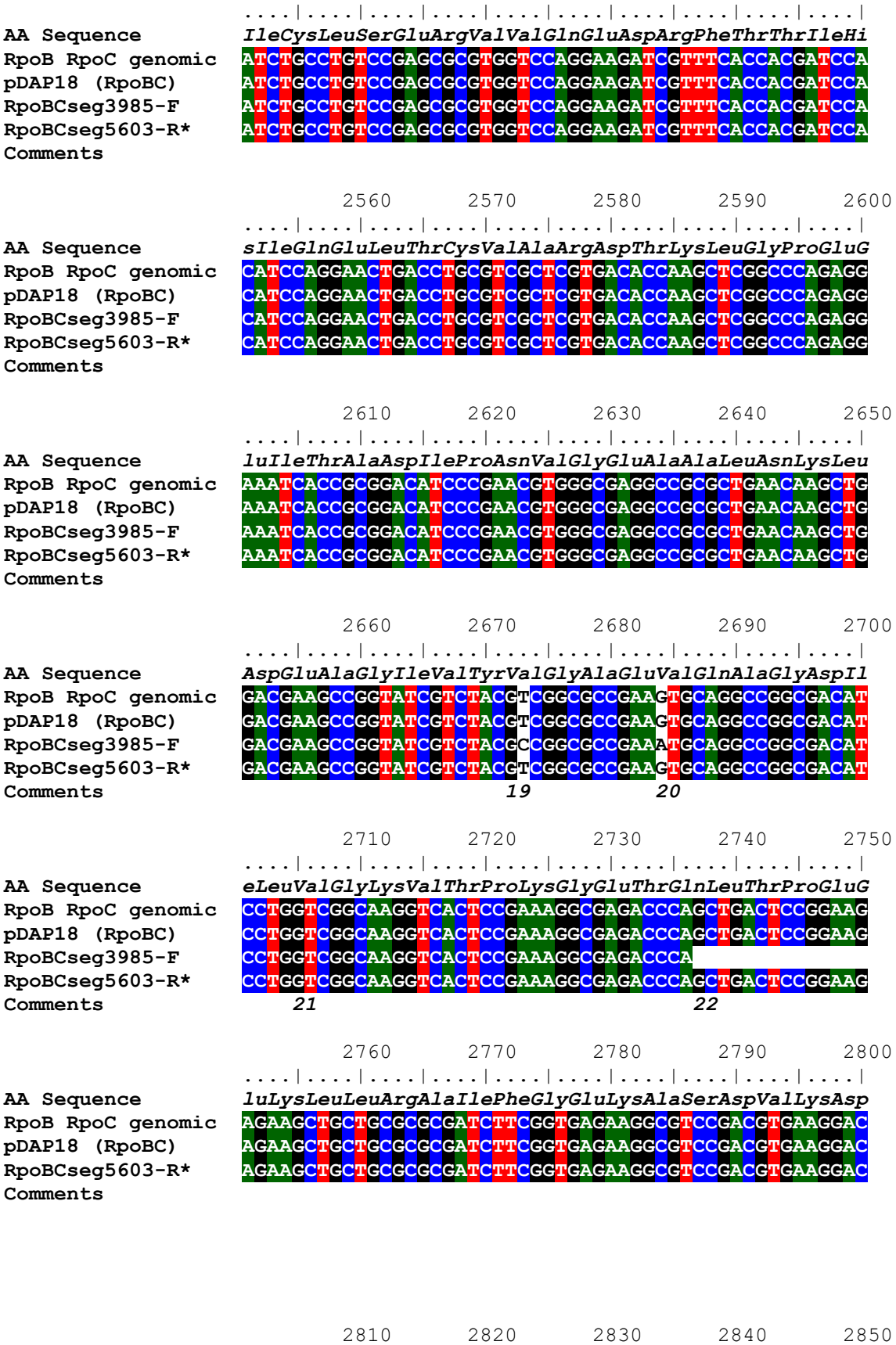

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg2421-F
RpoBCseg4036-R*
RpoBCseq10560-F
Comments
GluLeuThrLeuAspLeuLeuAlaLysValAlaLysAlaGlnValValAr
gIleGluThrLeuTyrThrAsnAspIleAspCysGlyProPheIleSerA
spThrLeuLysIleAspAsnThrSerAsnGlnLeuGluAlaLeuValGlu
IleTyrArgMetMetArgProGlyGluProProThrLysGluAlaAlaGl
uThrLeuPheGlyAsnLeuPhePheSerAlaGluArgTyrAspLeuSerA
laValGlyArgMetLysPheAsnArgArgIleGlyArgThrGluIleGlu
1060
1070
1080
1090
1100
1110
1120
1130
1140
1150
1160
1170
1180
1190
1200
1210
1220
1230
1240
1250
1260
1270
1280
1290
1300
1310
1320
1330
1340
1350

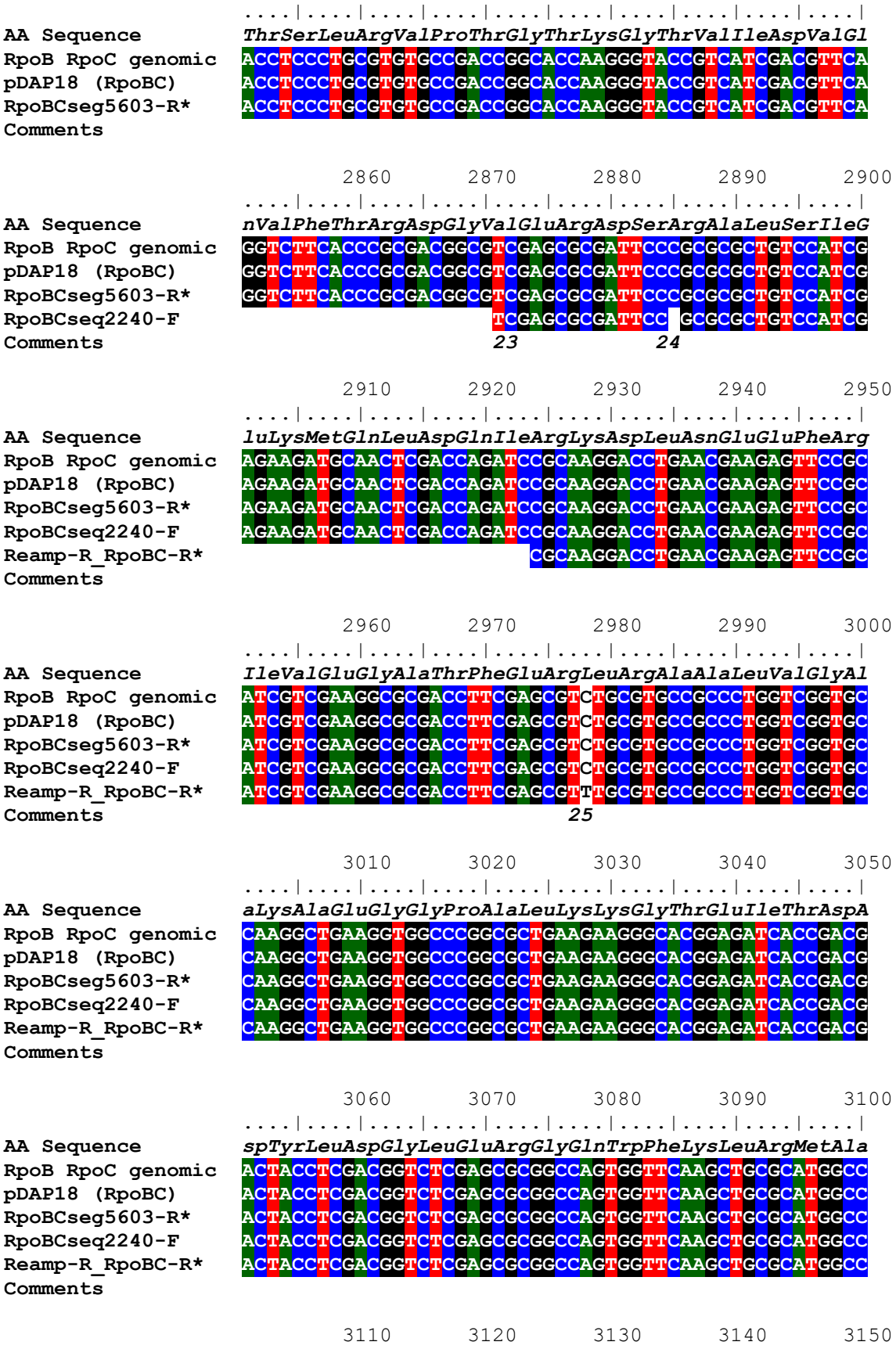

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg4036-R*
RpoBCseq10560-F
Comments
GlyProGlyValLeuSerLysGluAspIleIleAspValLeuLysThrLe
uValAspIleArgAsnGlyLysGlyIleValAspAspIleAspHisLeuG
lyAsnArgArgValArgCysValGlyGluMetAlaGluAsnGlnPheArg
ValGlyLeuValArgValGluArgAlaValLysGluArgLeuSerMetAl
aGluSerGluGlyLeuMetProGlnAspLeuIleAsnAlaLysProValA
laAlaIleLysGluPhePheGlySerSerGlnLeuSerGlnPheMet
15
1360
1370
1380
1390
1400
1410
1420
1430
1440
1450
1460
1470
1480
1490
1500
1510
1520
1530
1540
1550
1560
1570
1580
1590
1600
1610
1620
1630
1640
1650

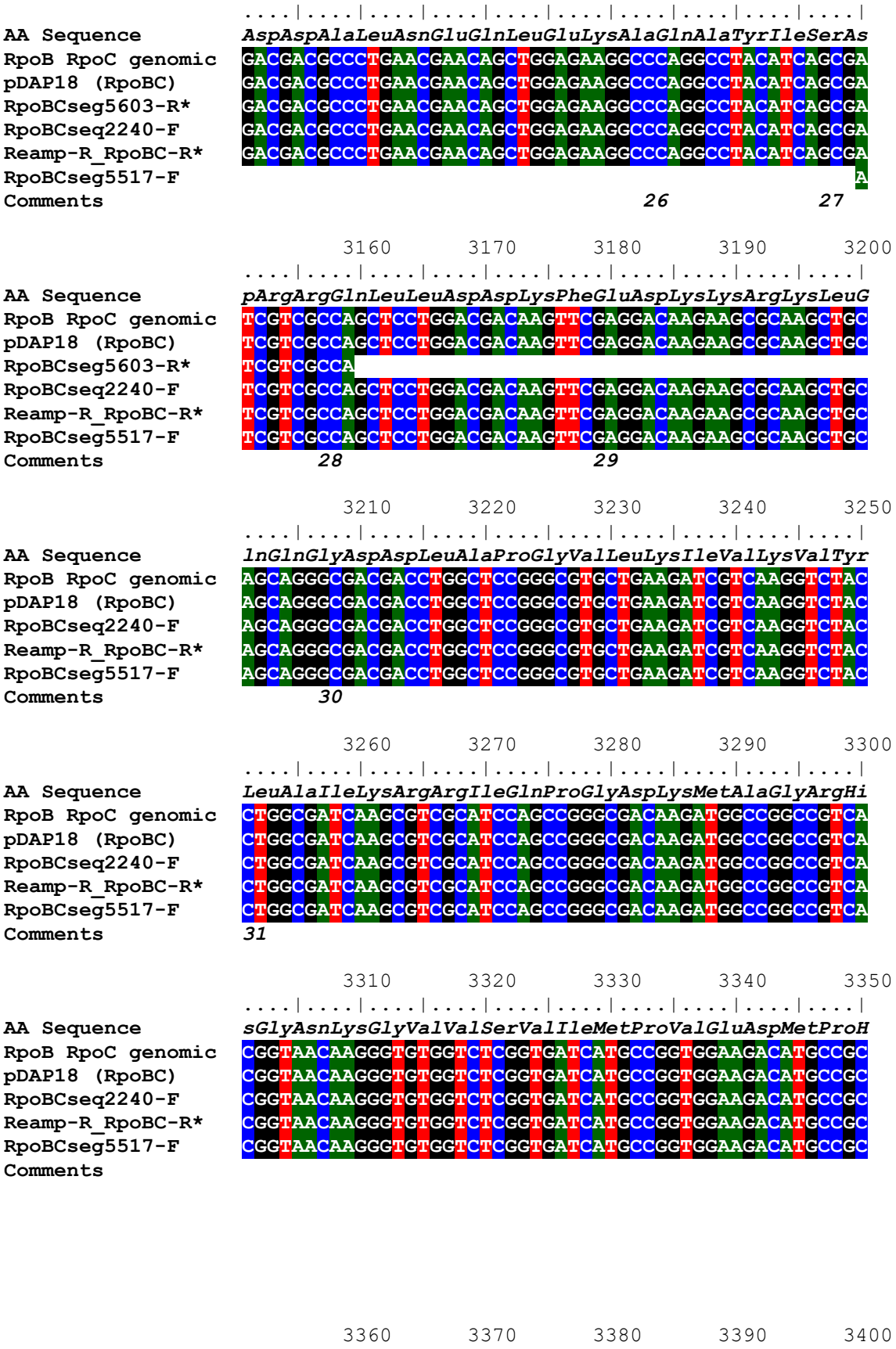

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseq10560-F
RpoBCseg3985-F
Comments
AspGlnAsnAsnProLeuSerGluIleThrHisLysArgArgValSerAl
15
1660
1670
1680
1690
1700
aLeuGlyProGlyGlyLeuThrArgGluArgAlaGlyPheGluValArgA
16
1710
1720
1730
1740
1750
spValHisProThrHisTyrGlyArgValCysProIleGluThrProGlu
1760
1770
1780
1790
1800
GlyProAsnIleGlyLeuIleAsnSerLeuAlaThrTyrAlaArgThrAs
1810
1820
1830
1840
1850
nLysTyrGlyPheLeuGluSerProTyrArgValValLysAspSerLeuV
1860
1870
1880
1890
1900
alThrAspGluIleValPheLeuSerAlaIleGluGluAlaAspHisVal
1910
1920
1930
1940
1950

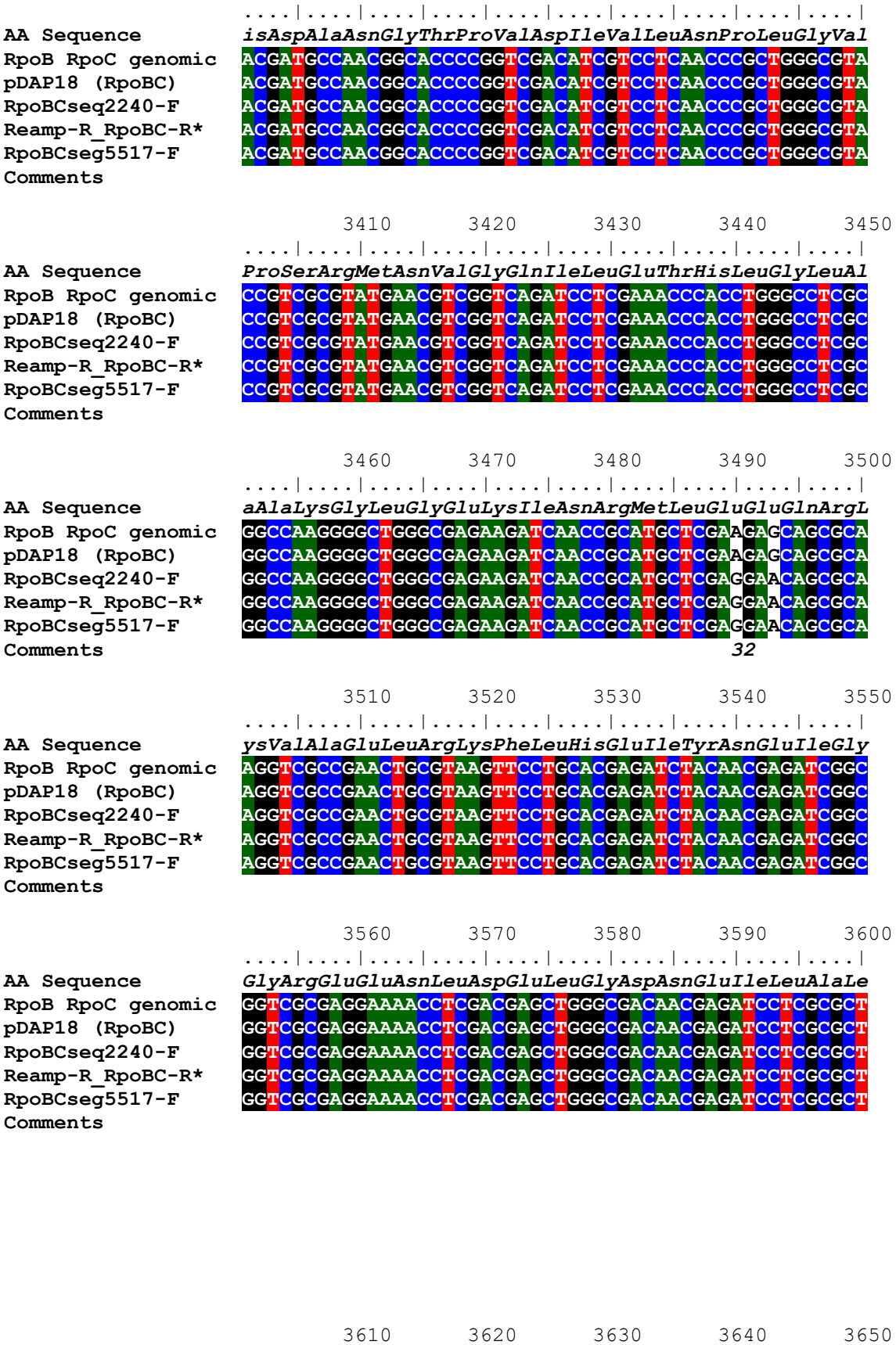

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseq10560-F
RpoBCseg3985-F
RpoBCseg5603-R*
Comments
IleAlaGlnAlaSerAlaThrLeuAsnGluLysGlyGlnLeuValAspGl
uLeuValAlaValArgHisLeuAsnGluPheThrValLysAlaProGluA
spValThrLeuMetAspValSerProLysGlnValValSerValAlaAla
SerLeuIleProPheLeuGluHisAspAspAlaAsnArgAlaLeuMetGl
17
ySerAsnMetGlnArgGlnAlaValProThrLeuArgAlaAspLysProL
euValGlyThrGlyMetGluArgAsnValAlaArgAspSerGlyValCys
18
1960
1970
1980
1990
2000
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100
2110
2120
2130
2140
2150
2160
2170
2180
2190
2200
2210
2220
2230
2240
2250

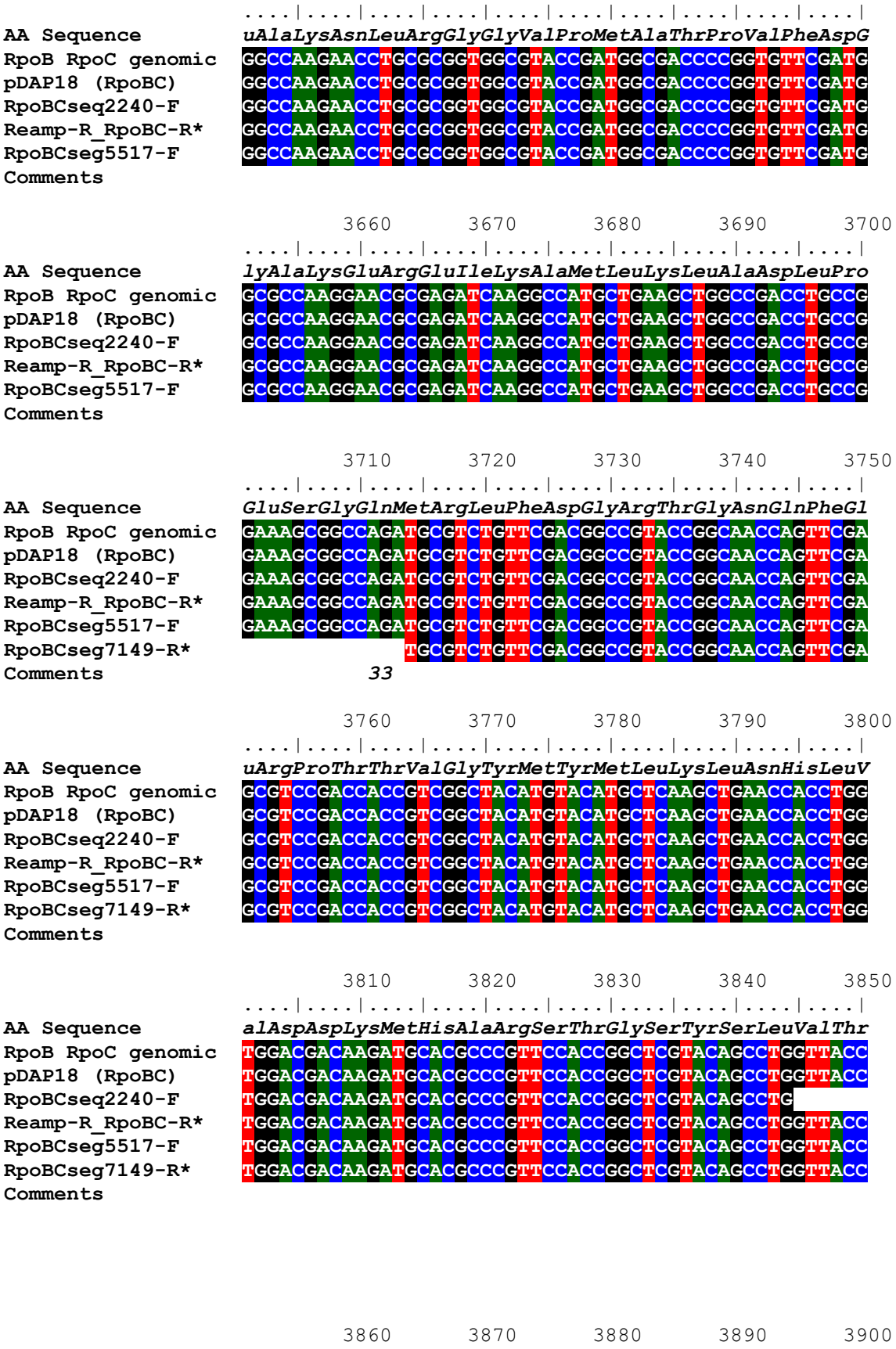

....|....|....|....|....|....|....|....|....|....|
AA Sequence ValValAlaArgArgGlyGlyValIleAspSerValAspAlaSerArgVa
RpoB RpoC genomic GTCGTGGCTCGCCGTGGCGGTGTGATCGACTCGGTCGATGCCAGCCGTGT
pDAP18 (RpoBC) GTCGTGGCTCGCCGTGGCGGTGTGATCGACTCGGTCGATGCCAGCCGTGT
RpoBCseg3985-F GTCGTGGCTCGCCGTGGCGGTGTGATCGACTCGGTCGATGCCAGCCGTGT
RpoBCseg5603-R* GTCGTGGCTCGCCGTGGCGGTGTGATCGACTCGGTCGATGCCAGCCGTGT
Comments
2260 2270 2280 2290 2300
....|....|....|....|....|....|....|....|....|....|
AA Sequence lValValArgValAlaAspAspGluValGluThrGlyGluAlaGlyValA
RpoB RpoC genomic CGTGGTTCGCGTGGCGGATGACGAAGTCGAGACCGGCGAAGCGGGTGTCG
pDAP18 (RpoBC) CGTGGTTCGCGTGGCGGATGACGAAGTCGAGACCGGCGAAGCGGGTGTCG
RpoBCseg3985-F CGTGGTTCGCGTGGCGGATGACGAAGTCGAGACCGGCGAAGCGGGTGTCG
RpoBCseg5603-R* CGTGGTTCGCGTGGCGGATGACGAAGTCGAGACCGGCGAAGCGGGTGTCG
Comments
2310 2320 2330 2340 2350
....|....|....|....|....|....|....|....|....|....|
AA Sequence spIleTyrAsnLeuThrLysTyrThrArgSerAsnGlnAsnThrCysIle
RpoB RpoC genomic ACATCTACAACCTGACCAAGTACACTCGTTCCAACCAGAACACCTGCATC
pDAP18 (RpoBC) ACATCTACAACCTGACCAAGTACACTCGTTCCAACCAGAACACCTGCATC
RpoBCseg3985-F ACATCTACAACCTGACCAAGTACACTCGTTCCAACCAGAACACCTGCATC
RpoBCseg5603-R* ACATCTACAACCTGACCAAGTACACTCGTTCCAACCAGAACACCTGCATC
Comments
2360 2370 2380 2390 2400
....|....|....|....|....|....|....|....|....|....|
AA Sequence AsnGlnArgProLeuValSerLysGlyAspValValAlaArgGlyAspIl
RpoB RpoC genomic AACCAGCGTCCGCTGGTGAGCAAGGGTGACGTGGTCGCGCGCGGCGACAT
pDAP18 (RpoBC) AACCAGCGTCCGCTGGTGAGCAAGGGTGACGTGGTCGCGCGCGGCGACAT
RpoBCseg3985-F AACCAGCGTCCGCTGGTGAGCAAGGGTGACGTGGTCGCGCGCGGCGACAT
RpoBCseg5603-R* AACCAGCGTCCGCTGGTGAGCAAGGGTGACGTGGTCGCGCGCGGCGACAT
Comments
2410 2420 2430 2440 2450
....|....|....|....|....|....|....|....|....|....|
AA Sequence eLeuAlaAspGlyProSerThrAspMetGlyGluLeuAlaLeuGlyGlnA
RpoB RpoC genomic CCTGGCCGACGGTCCGTCCACCGACATGGGCGAACTGGCCCTGGGCCAGA
pDAP18 (RpoBC) CCTGGCCGACGGTCCGTCCACCGACATGGGCGAACTGGCCCTGGGCCAGA
RpoBCseg3985-F CCTGGCCGACGGTCCGTCCACCGACATGGGCGAACTGGCCCTGGGCCAGA
RpoBCseg5603-R* CCTGGCCGACGGTCCGTCCACCGACATGGGCGAACTGGCCCTGGGCCAGA
Comments
2460 2470 2480 2490 2500
....|....|....|....|....|....|....|....|....|....|
AA Sequence snMetArgValAlaPheMetProTrpAsnGlyPheAsnPheGluAspSer
RpoB RpoC genomic ACATGCGCGTAGCGTTCATGCCCCTGGAACGGCTTCAACTTCGAAGACTCC
pDAP18 (RpoBC) ACATGCGCGTAGCGTTCATGCCCCTGGAACGGCTTCAACTTCGAAGACTCC
RpoBCseg3985-F ACATGCGCGTAGCGTTCATGCCCCTGGAACGGCTTCAACTTCGAAGACTCC
RpoBCseg5603-R* ACATGCGCGTAGCGTTCATGCCCCTGGAACGGCTTCAACTTCGAAGACTCC
Comments
2510 2520 2530 2540 2550

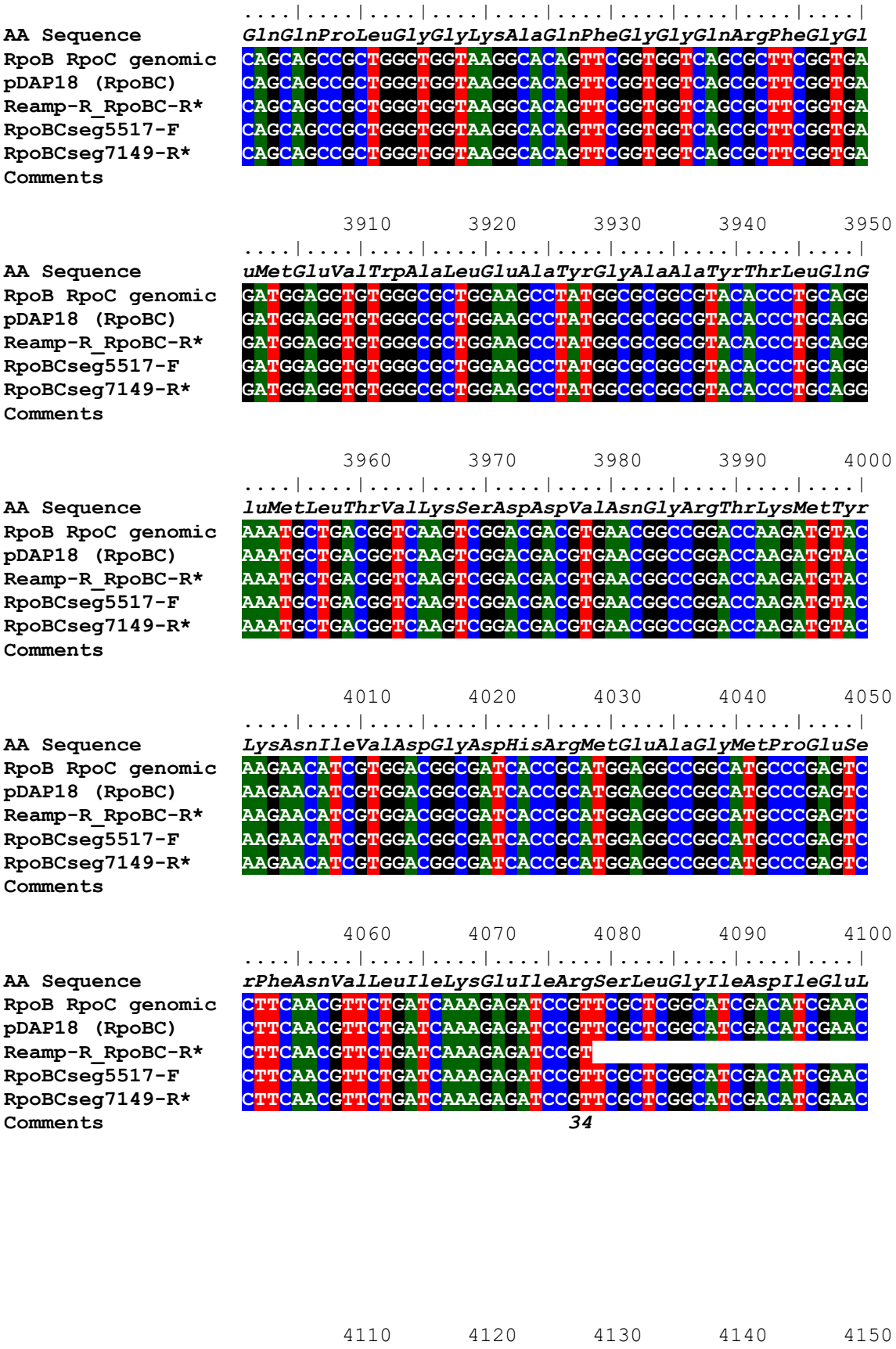

....|....|....|....|....|....|....|....|....|....|
AA Sequence IleCysLeuSerGluArgValValGlnGluAspArgPheThrThrIleHi
RpoB RpoC genomic ATCTGCCTGTCCGAGCGCGTGGTCCAGGAAGATCGTTTCACCACGATCCA
pDAP18 (RpoBC) ATCTGCCTGTCCGAGCGCGTGGTCCAGGAAGATCGTTTCACCACGATCCA
RpoBCseg3985-F ATCTGCCTGTCCGAGCGCGTGGTCCAGGAAGATCGTTTCACCACGATCCA
RpoBCseg5603-R* ATCTGCCTGTCCGAGCGCGTGGTCCAGGAAGATCGTTTCACCACGATCCA
Comments
2560 2570 2580 2590 2600
....|....|....|....|....|....|....|....|....|....|
AA Sequence sIleGlnGluLeuThrCysValAlaArgAspThrLysLeuGlyProGluG
RpoB RpoC genomic CATCCAGGAACTGACCTGCGTCGCTCGTGACACCAAGCTCGGCCCAGAGG
pDAP18 (RpoBC) CATCCAGGAACTGACCTGCGTCGCTCGTGACACCAAGCTCGGCCCAGAGG
RpoBCseg3985-F CATCCAGGAACTGACCTGCGTCGCTCGTGACACCAAGCTCGGCCCAGAGG
RpoBCseg5603-R* CATCCAGGAACTGACCTGCGTCGCTCGTGACACCAAGCTCGGCCCAGAGG
Comments
2610 2620 2630 2640 2650
....|....|....|....|....|....|....|....|....|....|
AA Sequence luIleThrAlaAspIleProAsnValGlyGluAlaAlaLeuAsnLysLeu
RpoB RpoC genomic AAATCACCGCGGACATCCCGAACGTGGGCGAGGCCGCGCTGAACAAGCTG
pDAP18 (RpoBC) AAATCACCGCGGACATCCCGAACGTGGGCGAGGCCGCGCTGAACAAGCTG
RpoBCseg3985-F AAATCACCGCGGACATCCCGAACGTGGGCGAGGCCGCGCTGAACAAGCTG
RpoBCseg5603-R* AAATCACCGCGGACATCCCGAACGTGGGCGAGGCCGCGCTGAACAAGCTG
Comments
2660 2670 2680 2690 2700
....|....|....|....|....|....|....|....|....|....|
AA Sequence AspGluAlaGlyIleValTyrValGlyAlaGluValGlnAlaGlyAspIl
RpoB RpoC genomic GACGAAGCCGGTATCGTCTACGTCGGCGCCGAAGTGCAGGCCGGCGACAT
pDAP18 (RpoBC) GACGAAGCCGGTATCGTCTACGTCGGCGCCGAAGTGCAGGCCGGCGACAT
RpoBCseg3985-F GACGAAGCCGGTATCGTCTACGCCGGCGCCGAAATGCAGGCCGGCGACAT
RpoBCseg5603-R* GACGAAGCCGGTATCGTCTACGTCGGCGCCGAAGTGCAGGCCGGCGACAT
Comments 19 20
2710 2720 2730 2740 2750
....|....|....|....|....|....|....|....|....|....|
AA Sequence eLeuValGlyLysValThrProLysGlyGluThrGlnLeuThrProGluG
RpoB RpoC genomic CCTGGTCGGCAAGGTCACTCCGAAAGGCGAGACCCAGCTGACTCCGGAAG
pDAP18 (RpoBC) CCTGGTCGGCAAGGTCACTCCGAAAGGCGAGACCCAGCTGACTCCGGAAG
RpoBCseg3985-F CCTGGTCGGCAAGGTCACTCCGAAAGGCGAGACCCA
RpoBCseg5603-R* CCTGGTCGGCAAGGTCACTCCGAAAGGCGAGACCCAGCTGACTCCGGAAG
Comments 21 22
2760 2770 2780 2790 2800
....|....|....|....|....|....|....|....|....|....|
AA Sequence luLysLeuLeuArgAlaIlePheGlyGluLysAlaSerAspValLysAsp
RpoB RpoC genomic AGAAGCTGCTGCGCGCGATCTTCGGTGAGAAGGCGTCCGACGTGAAGGAC
pDAP18 (RpoBC) AGAAGCTGCTGCGCGCGATCTTCGGTGAGAAGGCGTCCGACGTGAAGGAC
RpoBCseg5603-R* AGAAGCTGCTGCGCGCGATCTTCGGTGAGAAGGCGTCCGACGTGAAGGAC
Comments
2810 2820 2830 2840 2850

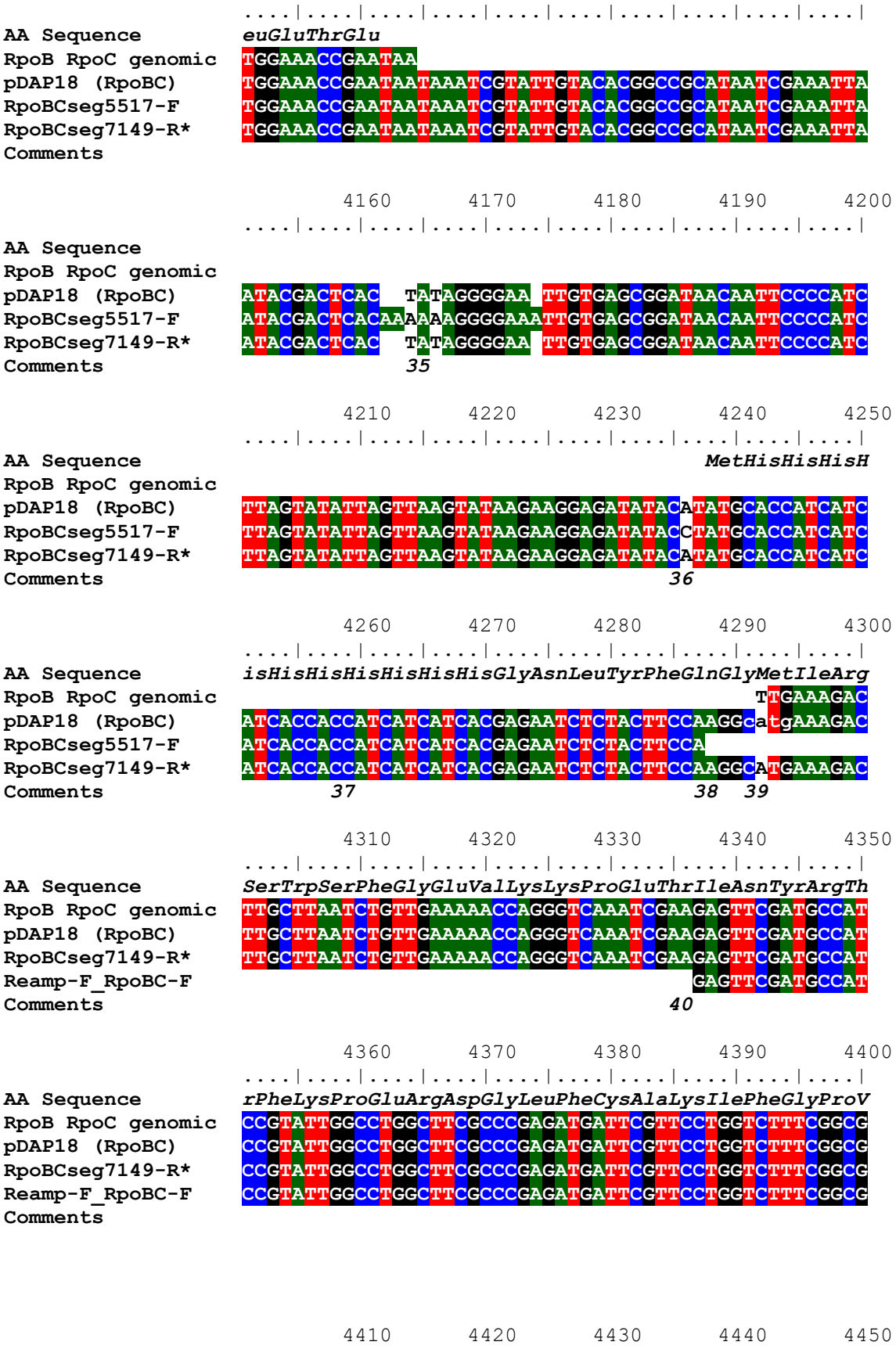

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg5603-R*
RpoBCseq2240-F
Reamp-R_RpoBC-R*
Comments
ThrSerLeuArgValProThrGlyThrLysGlyThrValIleAspValGl
nValPheThrArgAspGlyValGluArgAspSerArgAlaLeuSerIleG
luLysMetGlnLeuAspGlnIleArgLysAspLeuAsnGluGluPheArg
IleValGluGlyAlaThrPheGluArgLeuArgAlaAlaLeuValGlyAl
aLysAlaGluGlyGlyProAlaLeuLysLysGlyThrGluIleThrAspA
spTyrLeuAspGlyLeuGluArgGlyGlnTrpPheLysLeuArgMetAla
2860
2870
2880
2890
2900
2910
2920
2930
2940
2950
2960
2970
2980
2990
3000
3010
3020
3030
3040
3050
3060
3070
3080
3090
3100
3110
3120
3130
3140
3150
23
24
25

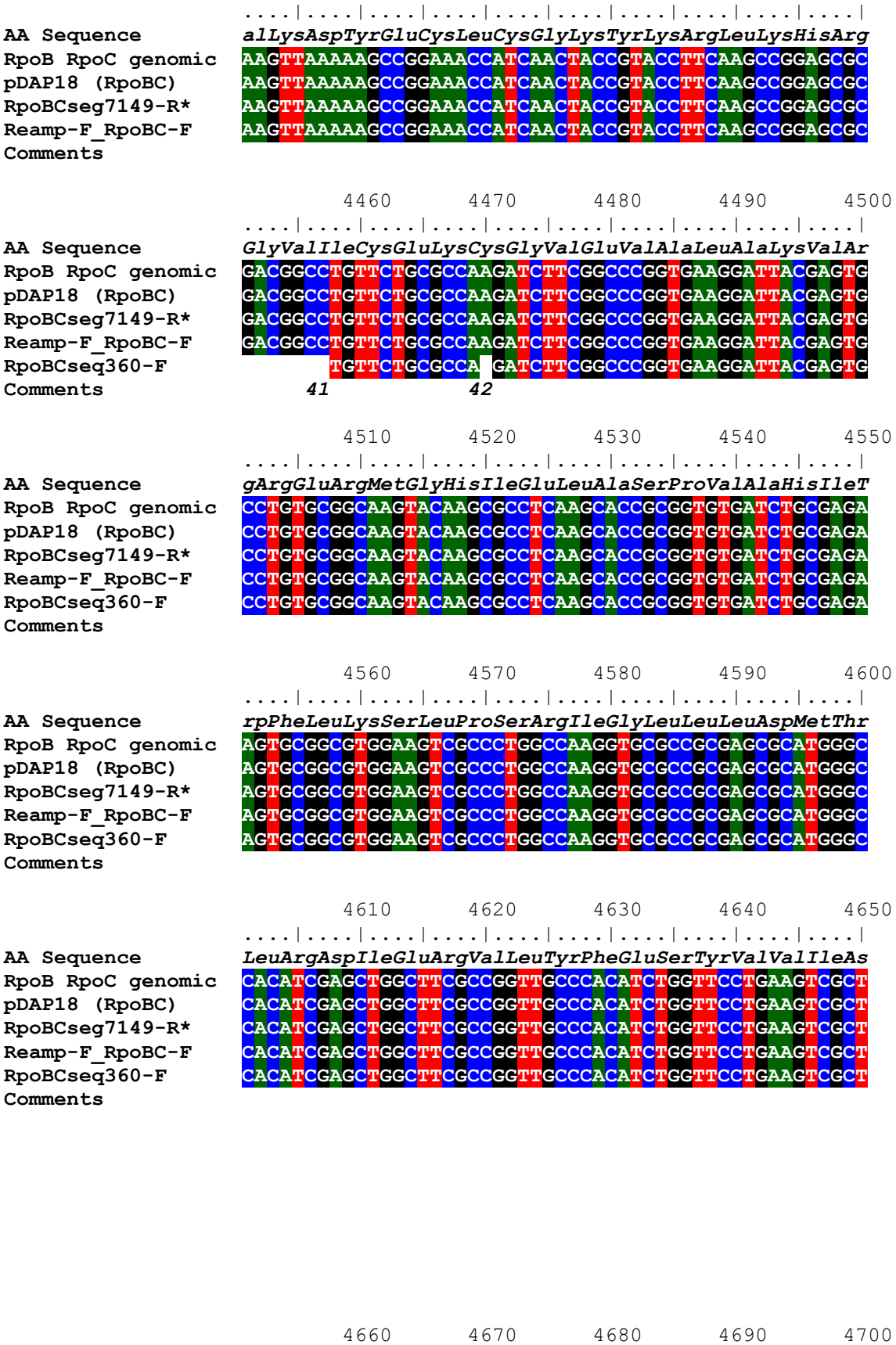

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg5603-R*
RpoBCseq2240-F
Reamp-R_RpoBC-R*
RpoBCseg5517-F
Comments
AspAspAlaLeuAsnGluGlnLeuGluLysAlaGlnAlaTyrIleSerAs
26
27
3160
3170
3180
3190
3200
pArgArgGlnLeuLeuAspAspLysPheGluAspLysLysArgLysLeuG
28
29
3210
3220
3230
3240
3250
lnGlnGlyAspAspLeuAlaProGlyValLeuLysIleValLysValTyr
30
3260
3270
3280
3290
3300
LeuAlaIleLysArgArgIleGlnProGlyAspLysMetAlaGlyArgHi
31
3310
3320
3330
3340
3350
sGlyAsnLysGlyValValSerValIleMetProValGluAspMetProH
3360
3370
3380
3390
3400

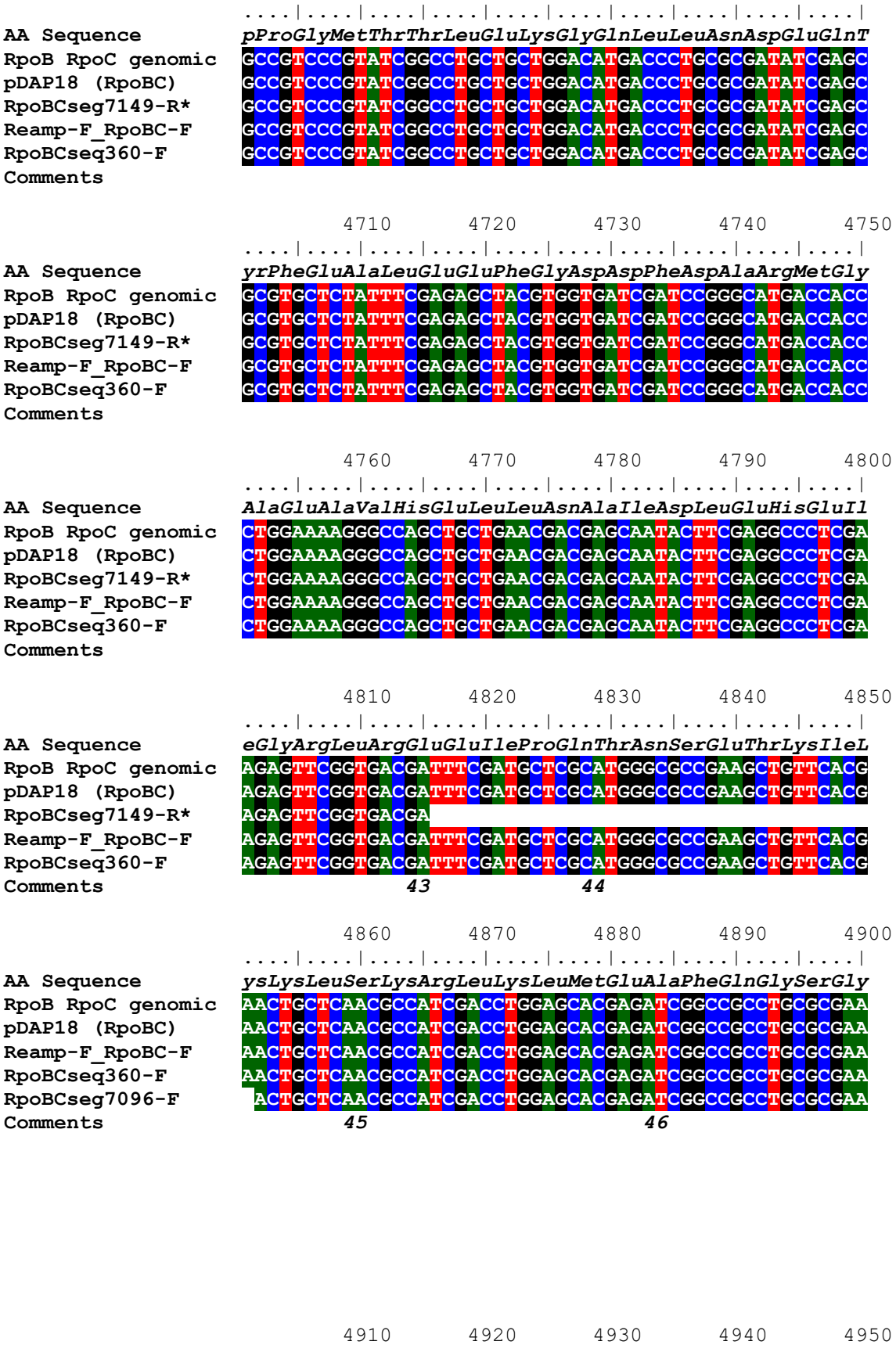

....|....|....|....|....|....|....|....|....|....|
AA Sequence isAspAlaAsnGlyThrProValAspIleValLeuAsnProLeuGlyVal
RpoB RpoC genomic ACGATGCCAACGGCACCCCGGTCGACATCGTCCTCAACCCGCTGGGCGTA
pDAP18 (RpoBC) ACGATGCCAACGGCACCCCGGTCGACATCGTCCTCAACCCGCTGGGCGTA
RpoBCseq2240-F ACGATGCCAACGGCACCCCGGTCGACATCGTCCTCAACCCGCTGGGCGTA
Reamp-R_RpoBC-R* ACGATGCCAACGGCACCCCGGTCGACATCGTCCTCAACCCGCTGGGCGTA
RpoBCseg5517-F ACGATGCCAACGGCACCCCGGTCGACATCGTCCTCAACCCGCTGGGCGTA
Comments
3410 3420 3430 3440 3450
....|....|....|....|....|....|....|....|....|....|
AA Sequence ProSerArgMetAsnValGlyGlnIleLeuGluThrHisLeuGlyLeuAl
RpoB RpoC genomic CCGTCGCGTATGAACGTCGGTCAGATCCTCGAAACCCACCTGGGCCTCGC
pDAP18 (RpoBC) CCGTCGCGTATGAACGTCGGTCAGATCCTCGAAACCCACCTGGGCCTCGC
RpoBCseq2240-F CCGTCGCGTATGAACGTCGGTCAGATCCTCGAAACCCACCTGGGCCTCGC
Reamp-R_RpoBC-R* CCGTCGCGTATGAACGTCGGTCAGATCCTCGAAACCCACCTGGGCCTCGC
RpoBCseg5517-F CCGTCGCGTATGAACGTCGGTCAGATCCTCGAAACCCACCTGGGCCTCGC
Comments
3460 3470 3480 3490 3500
....|....|....|....|....|....|....|....|....|....|
AA Sequence aAlaLysGlyLeuGlyGluLysIleAsnArgMetLeuGluGluGlnArgL
RpoB RpoC genomic GGCCAAGGGGCTGGGCGAGAAGATCAACCGCATGCTCGAAGAGCAGCGCA
pDAP18 (RpoBC) GGCCAAGGGGCTGGGCGAGAAGATCAACCGCATGCTCGAAGAGCAGCGCA
RpoBCseq2240-F GGCCAAGGGGCTGGGCGAGAAGATCAACCGCATGCTCGAGGAACAGCGCA
Reamp-R_RpoBC-R* GGCCAAGGGGCTGGGCGAGAAGATCAACCGCATGCTCGAGGAACAGCGCA
RpoBCseg5517-F GGCCAAGGGGCTGGGCGAGAAGATCAACCGCATGCTCGAGGAACAGCGCA
Comments 32
3510 3520 3530 3540 3550
....|....|....|....|....|....|....|....|....|....|
AA Sequence ysValAlaGluLeuArgLysPheLeuHisGluIleTyrAsnGluIleGly
RpoB RpoC genomic AGGTCGCCGAACTGCGTAAGTTCCTGCACGAGATCTACAACGAGATCGGC
pDAP18 (RpoBC) AGGTCGCCGAACTGCGTAAGTTCCTGCACGAGATCTACAACGAGATCGGC
RpoBCseq2240-F AGGTCGCCGAACTGCGTAAGTTCCTGCACGAGATCTACAACGAGATCGGC
Reamp-R_RpoBC-R* AGGTCGCCGAACTGCGTAAGTTCCTGCACGAGATCTACAACGAGATCGGC
RpoBCseg5517-F AGGTCGCCGAACTGCGTAAGTTCCTGCACGAGATCTACAACGAGATCGGC
Comments
3560 3570 3580 3590 3600
....|....|....|....|....|....|....|....|....|....|
AA Sequence GlyArgGluGluAsnLeuAspGluLeuGlyAspAsnGluIleLeuAlaLe
RpoB RpoC genomic GGTCGCGAGGAAAACCTCGACGAGCTGGGCGACAACGAGATCCTCGCGCT
pDAP18 (RpoBC) GGTCGCGAGGAAAACCTCGACGAGCTGGGCGACAACGAGATCCTCGCGCT
RpoBCseq2240-F GGTCGCGAGGAAAACCTCGACGAGCTGGGCGACAACGAGATCCTCGCGCT
Reamp-R_RpoBC-R* GGTCGCGAGGAAAACCTCGACGAGCTGGGCGACAACGAGATCCTCGCGCT
RpoBCseg5517-F GGTCGCGAGGAAAACCTCGACGAGCTGGGCGACAACGAGATCCTCGCGCT
Comments
3610 3620 3630 3640 3650

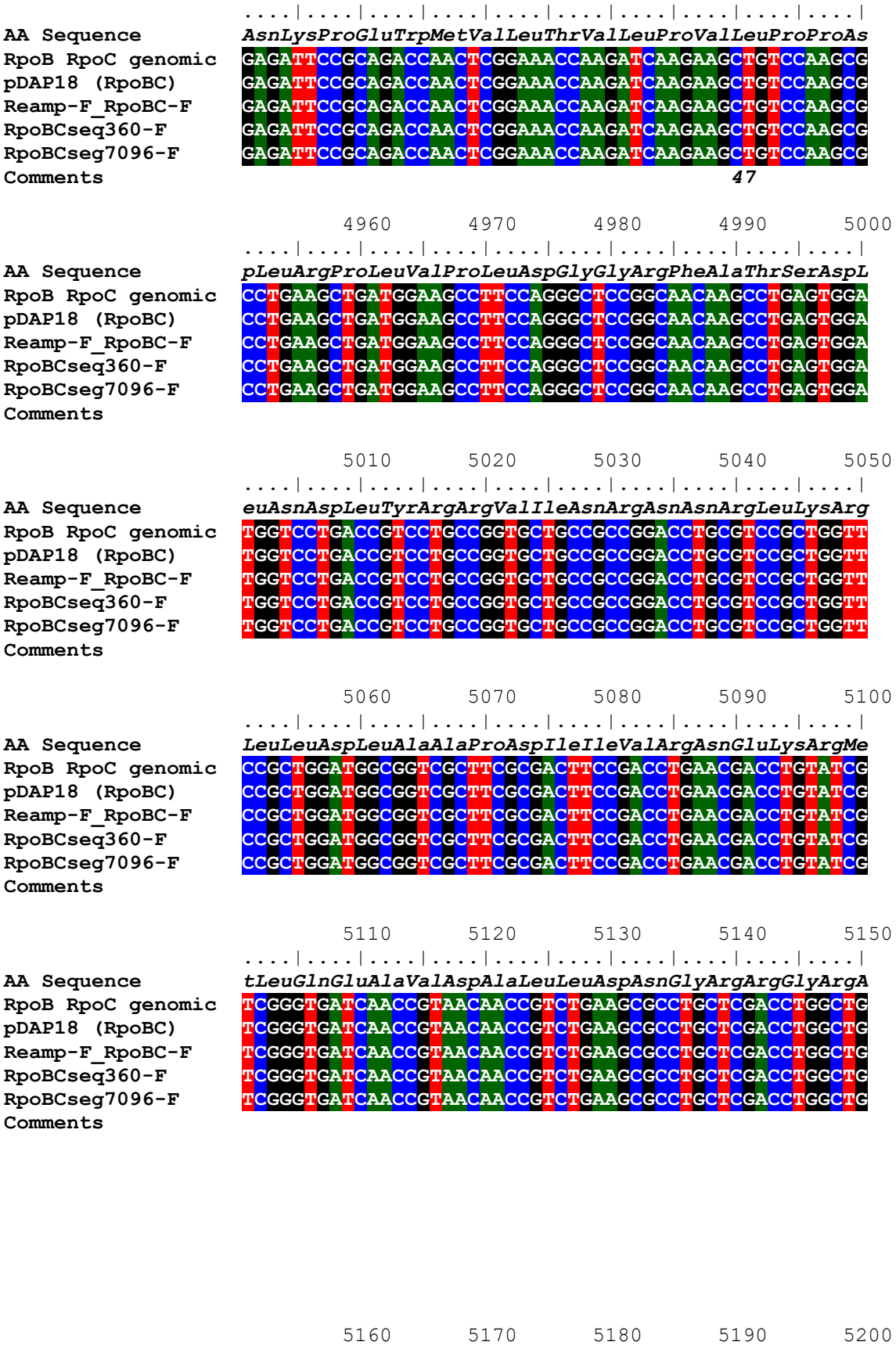

....|....|....|....|....|....|....|....|....|....|
AA Sequence uAlaLysAsnLeuArgGlyGlyValProMetAlaThrProValPheAspG
RpoB RpoC genomic GGCCAAGAACCTGCGCGGTGGCGTACCGATGGCGACCCCGGTGTTCGATG
pDAP18 (RpoBC) GGCCAAGAACCTGCGCGGTGGCGTACCGATGGCGACCCCGGTGTTCGATG
RpoBCseq2240-F GGCCAAGAACCTGCGCGGTGGCGTACCGATGGCGACCCCGGTGTTCGATG
Reamp-R_RpoBC-R* GGCCAAGAACCTGCGCGGTGGCGTACCGATGGCGACCCCGGTGTTCGATG
RpoBCseg5517-F GGCCAAGAACCTGCGCGGTGGCGTACCGATGGCGACCCCGGTGTTCGATG
Comments
3660 3670 3680 3690 3700
....|....|....|....|....|....|....|....|....|....|
AA Sequence lyAlaLysGluArgGluIleLysAlaMetLeuLysLeuAlaAspLeuPro
RpoB RpoC genomic GCGCCAAGGAACGCGAGATCAAGGCCATGCTGAAGCTGGCCGACCTGCCG
pDAP18 (RpoBC) GCGCCAAGGAACGCGAGATCAAGGCCATGCTGAAGCTGGCCGACCTGCCG
RpoBCseq2240-F GCGCCAAGGAACGCGAGATCAAGGCCATGCTGAAGCTGGCCGACCTGCCG
Reamp-R_RpoBC-R* GCGCCAAGGAACGCGAGATCAAGGCCATGCTGAAGCTGGCCGACCTGCCG
RpoBCseg5517-F GCGCCAAGGAACGCGAGATCAAGGCCATGCTGAAGCTGGCCGACCTGCCG
Comments
3710 3720 3730 3740 3750
....|....|....|....|....|....|....|....|....|....|
AA Sequence GluSerGlyGlnMetArgLeuPheAspGlyArgThrGlyAsnGlnPheGl
RpoB RpoC genomic GAAAGCGGCCAGATGCGTCTGTTCGACGGCCGTACCGGCAACCAGTTCGA
pDAP18 (RpoBC) GAAAGCGGCCAGATGCGTCTGTTCGACGGCCGTACCGGCAACCAGTTCGA
RpoBCseq2240-F GAAAGCGGCCAGATGCGTCTGTTCGACGGCCGTACCGGCAACCAGTTCGA
Reamp-R_RpoBC-R* GAAAGCGGCCAGATGCGTCTGTTCGACGGCCGTACCGGCAACCAGTTCGA
RpoBCseg5517-F GAAAGCGGCCAGATGCGTCTGTTCGACGGCCGTACCGGCAACCAGTTCGA
RpoBCseg7149-R* TGCGTCTGTTCGACGGCCGTACCGGCAACCAGTTCGA
Comments 33
3760 3770 3780 3790 3800
....|....|....|....|....|....|....|....|....|....|
AA Sequence uArgProThrThrValGlyTyrMetTyrMetLeuLysLeuAsnHisLeuV
RpoB RpoC genomic GCGTCCGACCACCGTCGGCTACATGTACATGCTCAAGCTGAACCACCTGG
pDAP18 (RpoBC) GCGTCCGACCACCGTCGGCTACATGTACATGCTCAAGCTGAACCACCTGG
RpoBCseq2240-F GCGTCCGACCACCGTCGGCTACATGTACATGCTCAAGCTGAACCACCTGG
Reamp-R_RpoBC-R* GCGTCCGACCACCGTCGGCTACATGTACATGCTCAAGCTGAACCACCTGG
RpoBCseg5517-F GCGTCCGACCACCGTCGGCTACATGTACATGCTCAAGCTGAACCACCTGG
RpoBCseg7149-R* GCGTCCGACCACCGTCGGCTACATGTACATGCTCAAGCTGAACCACCTGG
Comments
3810 3820 3830 3840 3850
....|....|....|....|....|....|....|....|....|....|
AA Sequence alAspAspLysMetHisAlaArgSerThrGlySerTyrSerLeuValThr
RpoB RpoC genomic TGGACGACAAGATGCACGCCCGTTCCACCGGCTCGTACAGCCTGGTTACC
pDAP18 (RpoBC) TGGACGACAAGATGCACGCCCGTTCCACCGGCTCGTACAGCCTGGTTACC
RpoBCseq2240-F TGGACGACAAGATGCACGCCCGTTCCACCGGCTCGTACAGCCTG
Reamp-R_RpoBC-R* TGGACGACAAGATGCACGCCCGTTCCACCGGCTCGTACAGCCTGGTTACC
RpoBCseg5517-F TGGACGACAAGATGCACGCCCGTTCCACCGGCTCGTACAGCCTGGTTACC
RpoBCseg7149-R* TGGACGACAAGATGCACGCCCGTTCCACCGGCTCGTACAGCCTGGTTACC
Comments
3860 3870 3880 3890 3900

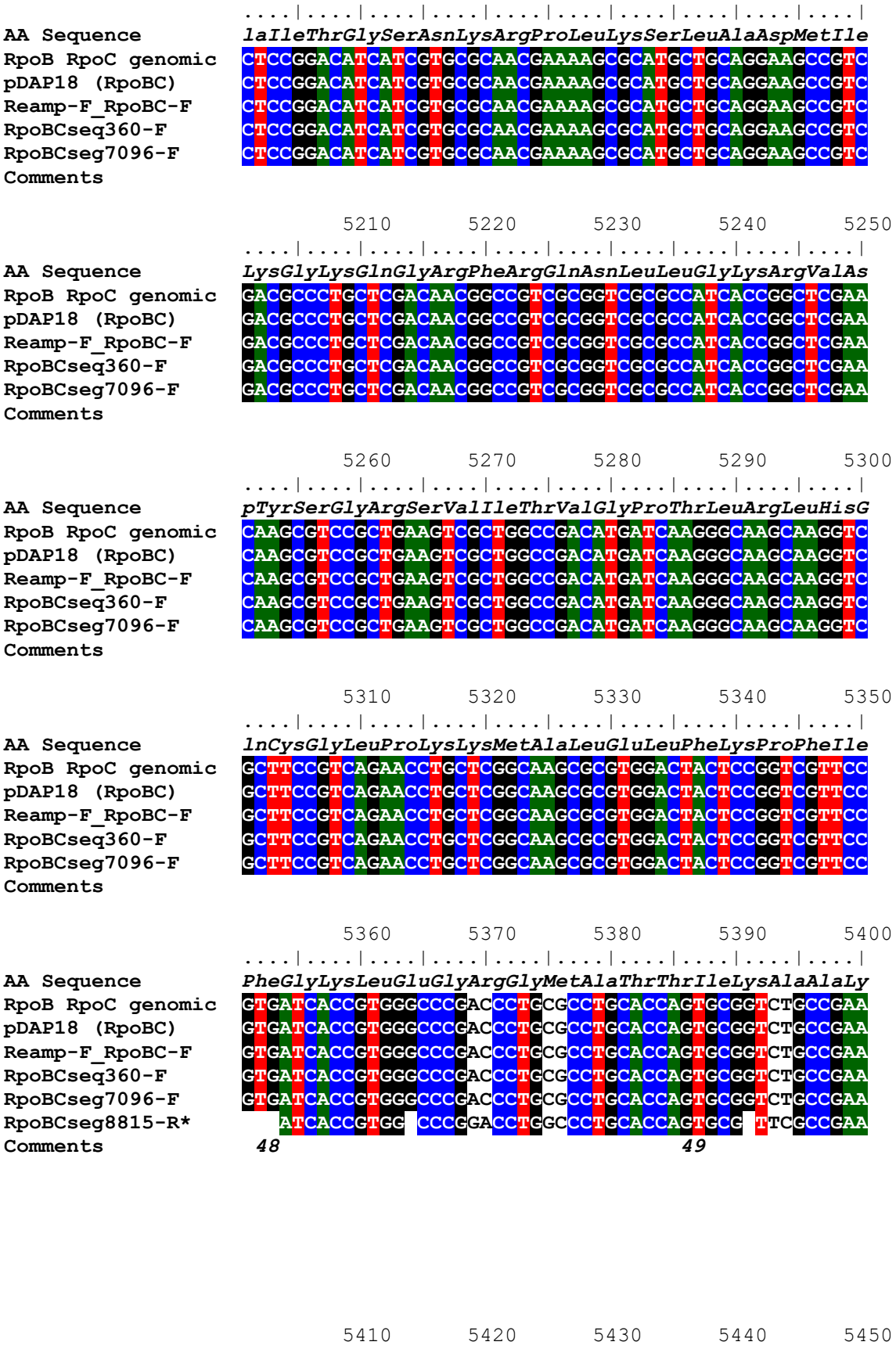

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
Reamp-R_RpoBC-R*
RpoBCseg5517-F
RpoBCseg7149-R*
Comments
GlnGlnProLeuGlyGlyLysAlaGlnPheGlyGlyGlnArgPheGlyGl
CAGCAGCCGCTGGGTGGTAAGGCACAGTTCGGTGGTCAGCGCTTCGGTGA
3910
3920
3930
3940
3950
uMetGluValTrpAlaLeuGluAlaTyrGlyAlaAlaTyrThrLeuGlnG
GATGGAGGTGTGGGCGCTGGAAGCCTATGGCGCGGCGTACACCCTGCAGG
3960
3970
3980
3990
4000
luMetLeuThrValLysSerAspAspValAsnGlyArgThrLysMetTyr
AAATGCTGACGGTCAAGTCGGACGACGTGAACGGCCGGACCAAGATGTAC
4010
4020
4030
4040
4050
LysAsnIleValAspGlyAspHisArgMetGluAlaGlyMetProGluSe
AAGAACATCGTGGACGGCGATCACCGCATGGAGGCCGGCATGCCCGAGTC
4060
4070
4080
4090
4100
rPheAsnValLeuIleLysGluIleArgSerLeuGlyIleAspIleGluL
CTTCAACGTTCTGATCAAAGAGATCCGTTCGCTCGGCATCGACATCGAAC
CTTCAACGTTCTGATCAAAGAGATCCGT
34
4110
4120
4130
4140
4150

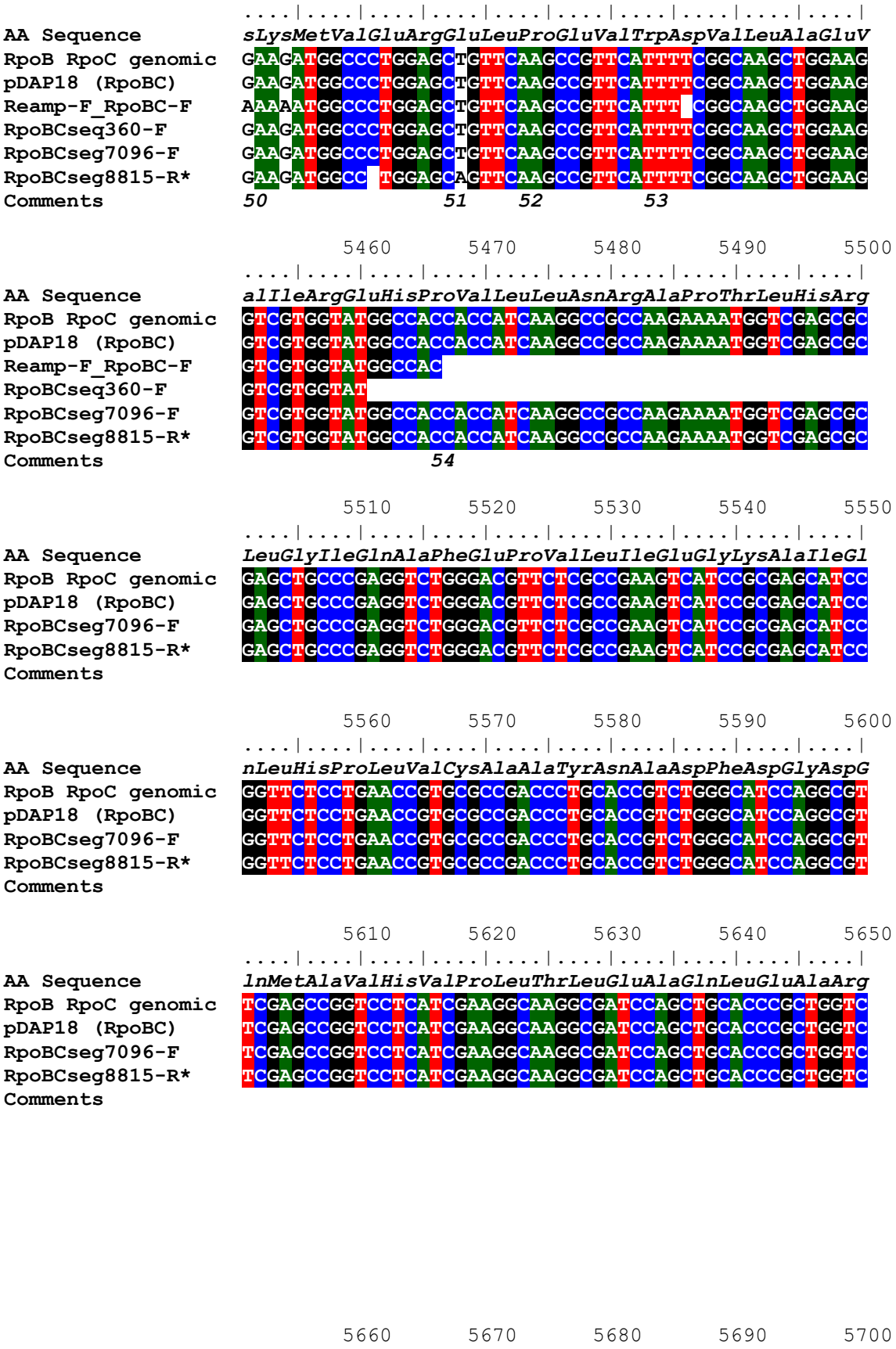

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg5517-F
RpoBCseg7149-R*
Comments
Reamp-F_RpoBC-F
euGluThrGlu
TGGAAACCGAATAA
TGGAAACCGAATAATAAATCGTATTGTACACGGCCGCATAATCGAAATTA
TGGAAACCGAATAATAAATCGTATTGTACACGGCCGCATAATCGAAATTA
TGGAAACCGAATAATAAATCGTATTGTACACGGCCGCATAATCGAAATTA
4160
4170
4180
4190
4200
ATACGACTCAC TATAGGGGAA TTGTGAGCGGATAACAATTCCCCATC
ATACGACTCACAAAAAGGGGAAATTGTGAGCGGATAACAATTCCCCATC
ATACGACTCAC TATAGGGGAA TTGTGAGCGGATAACAATTCCCCATC
35
4210
4220
4230
4240
4250
MetHisHisHisH
TTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGCACCATCATC
TTAGTATATTAGTTAAGTATAAGAAGGAGATATACCTATGCACCATCATC
TTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGCACCATCATC
36
4260
4270
4280
4290
4300
isHisHisHisHisHisHisGlyAsnLeuTyrPheGlnGlyMetIleArg
TTGAAAGAC
ATCACCACCATCATCATCACGAGAATCTCTACTTCCAAGGCatgAAAGAC
ATCACCACCATCATCATCACGAGAATCTCTACTTCCA
ATCACCACCATCATCATCACGAGAATCTCTACTTCCAAGGCATGAAAGAC
37
38
39
4310
4320
4330
4340
4350
SerTrpSerPheGlyGluValLysLysProGluThrIleAsnTyrArgTh
TTGCTTAATCTGTTGAAAAACCAGGGTCAAATCGAAGAGTTCGATGCCAT
TTGCTTAATCTGTTGAAAAACCAGGGTCAAATCGAAGAGTTCGATGCCAT
TTGCTTAATCTGTTGAAAAACCAGGGTCAAATCGAAGAGTTCGATGCCAT
GAGTTCGATGCCAT
40
4360
4370
4380
4390
4400
rPheLysProGluArgAspGlyLeuPheCysAlaLysIlePheGlyProV
CCGTATTGGCCTGGCTTCGCCCGAGATGATTCGTTCCTGGTCTTTCGGCG
CCGTATTGGCCTGGCTTCGCCCGAGATGATTCGTTCCTGGTCTTTCGGCG
CCGTATTGGCCTGGCTTCGCCCGAGATGATTCGTTCCTGGTCTTTCGGCG
CCGTATTGGCCTGGCTTCGCCCGAGATGATTCGTTCCTGGTCTTTCGGCG
4410
4420
4430
4440
4450

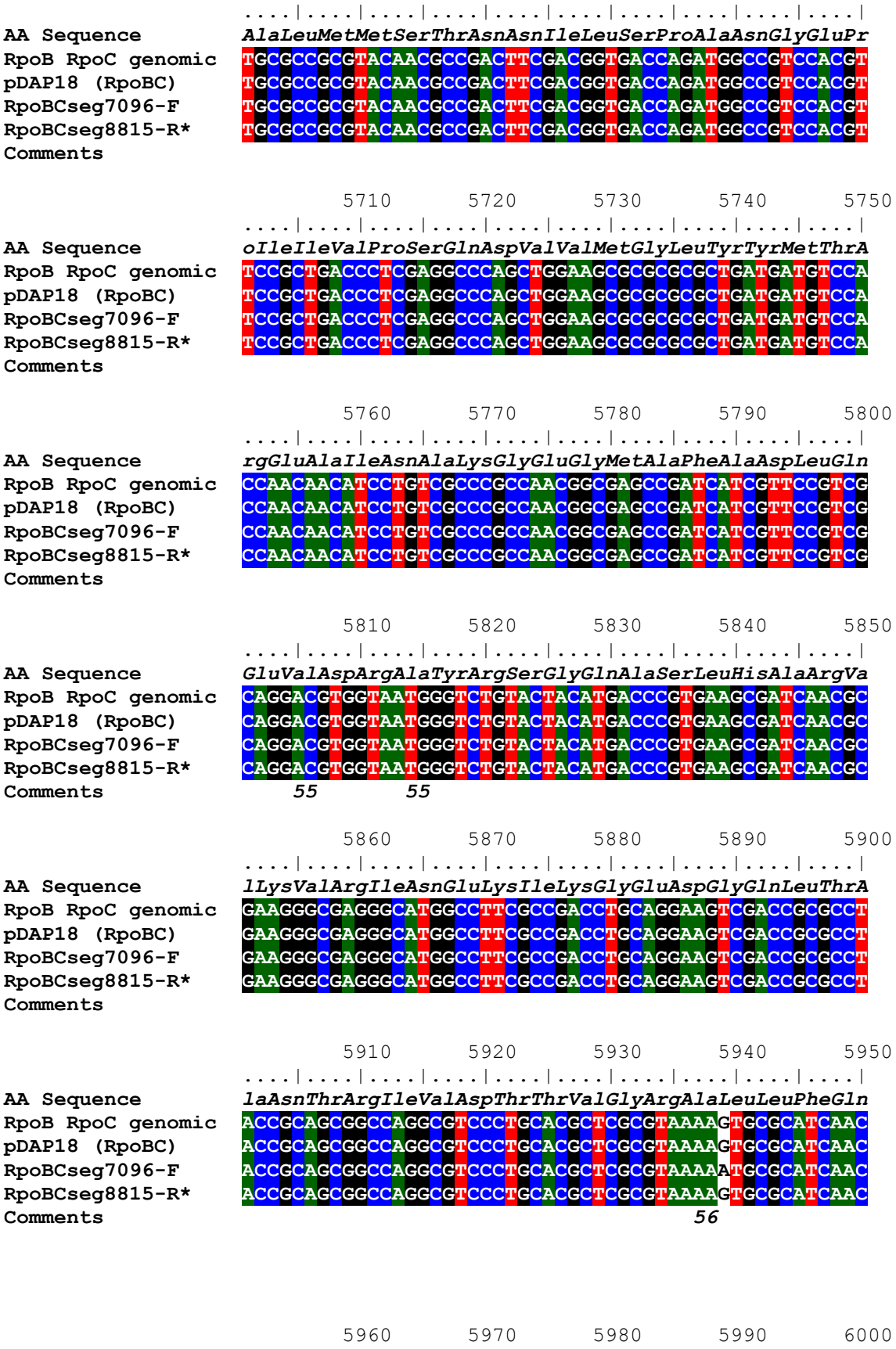

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg7149-R*
Reamp-F_RpoBC-F
RpoBCseq360-F
Comments
4460
4470
4480
4490
4500
4510
4520
4530
4540
4550
4560
4570
4580
4590
4600
4610
4620
4630
4640
4650
4660
4670
4680
4690
4700
41
42

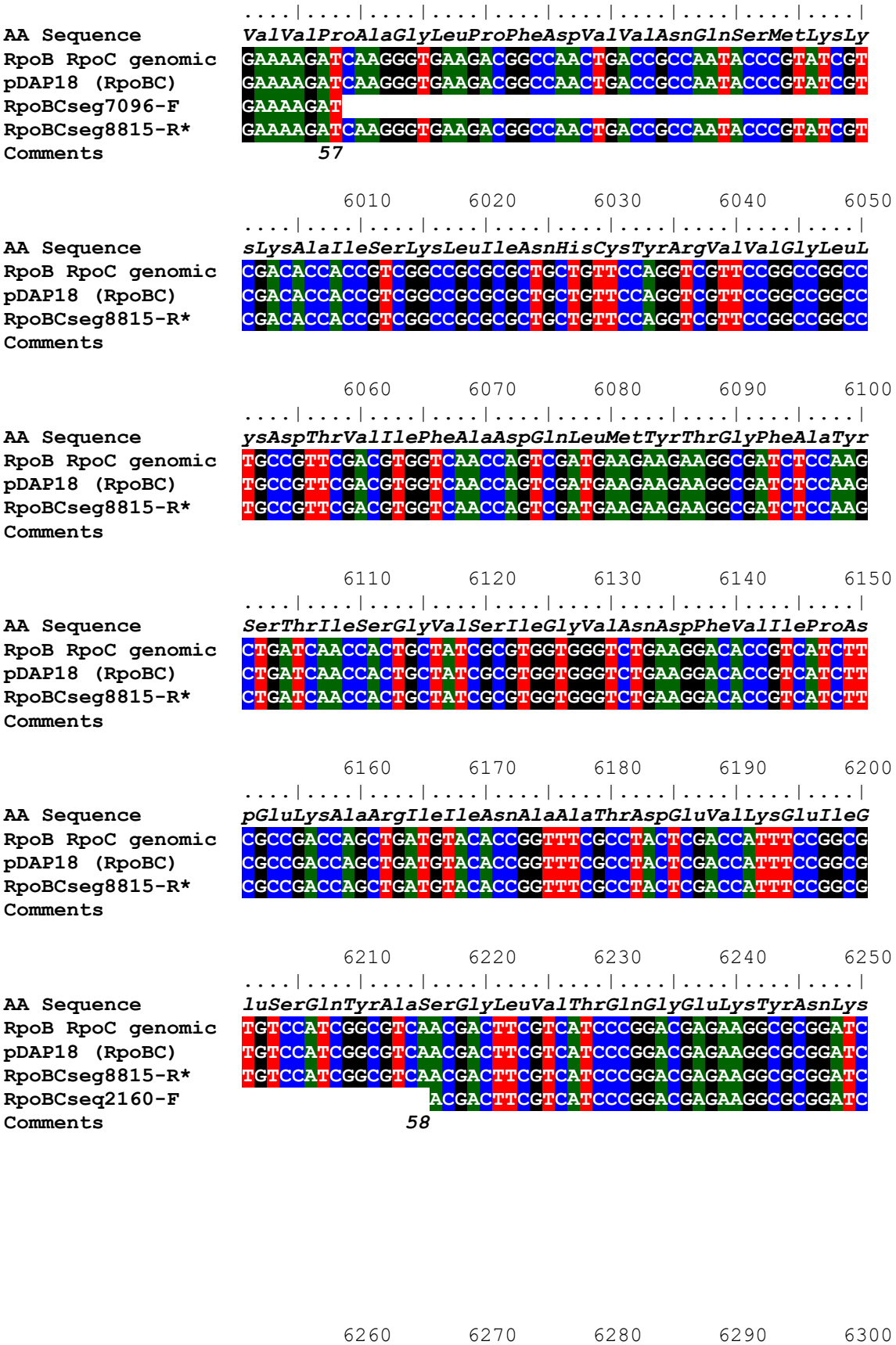

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg7149-R*
Reamp-F_RpoBC-F
RpoBCseq360-F
RpoBCseq7096-F
Comments
4710
4720
4730
4740
4750
4760
4770
4780
4790
4800
4810
4820
4830
4840
4850
4860
4870
4880
4890
4900
4910
4920
4930
4940
4950
43
44
45
46

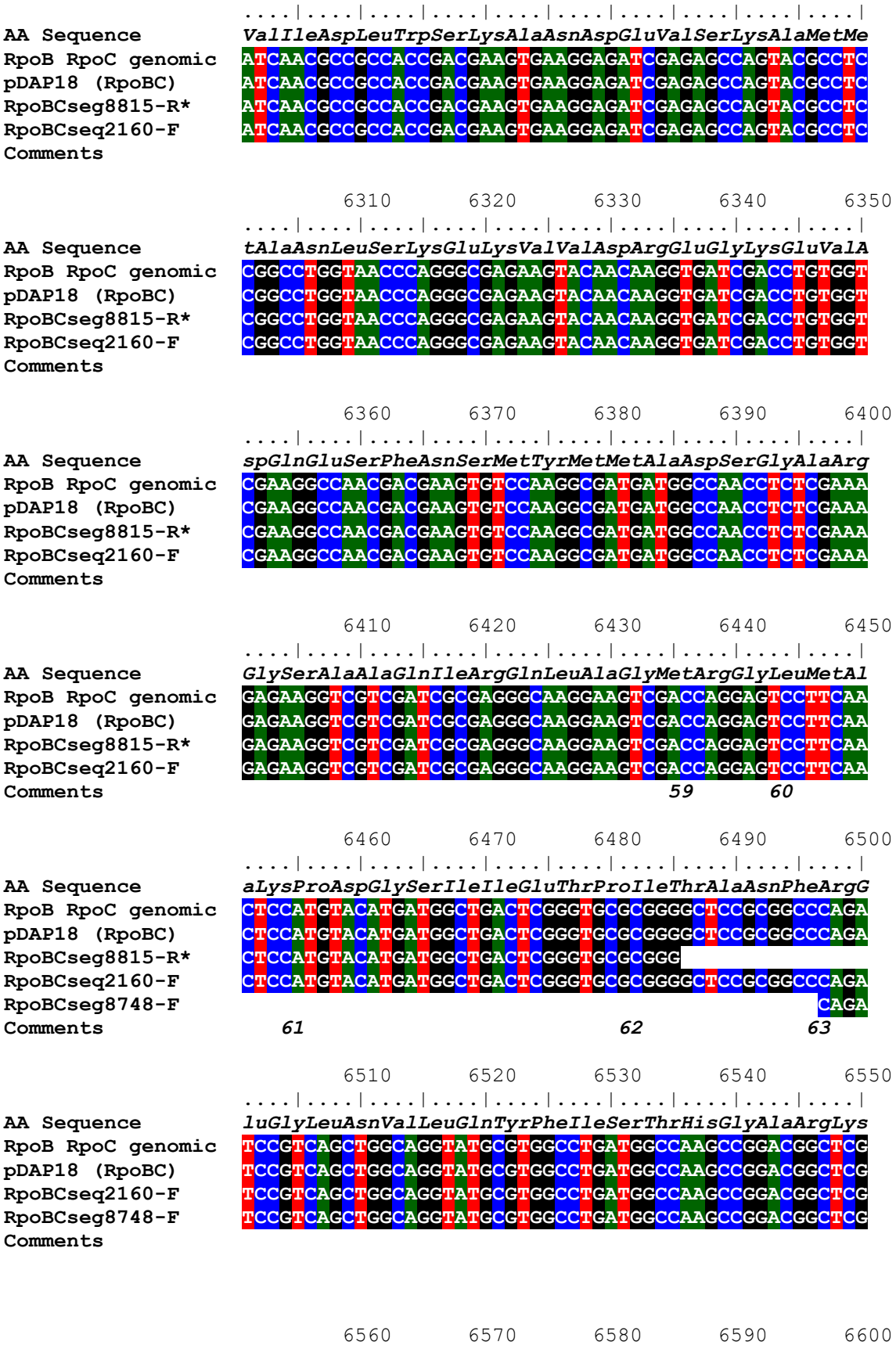

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
Reamp-F_RpoBC-F
RpoBCseq360-F
RpoBCseg7096-F
Comments
AsnLysProGluTrpMetValLeuThrValLeuProValLeuProProAs
47
4960
4970
4980
4990
5000
pLeuArgProLeuValProLeuAspGlyGlyArgPheAlaThrSerAspL
5010
5020
5030
5040
5050
euAsnAspLeuTyrArgArgValIleAsnArgAsnAsnArgLeuLysArg
5060
5070
5080
5090
5100
LeuLeuAspLeuAlaAlaProAspIleIleValArgAsnGluLysArgMe
5110
5120
5130
5140
5150
tLeuGlnGluAlaValAspAlaLeuLeuAspAsnGlyArgArgGlyArgA
5160
5170
5180
5190
5200

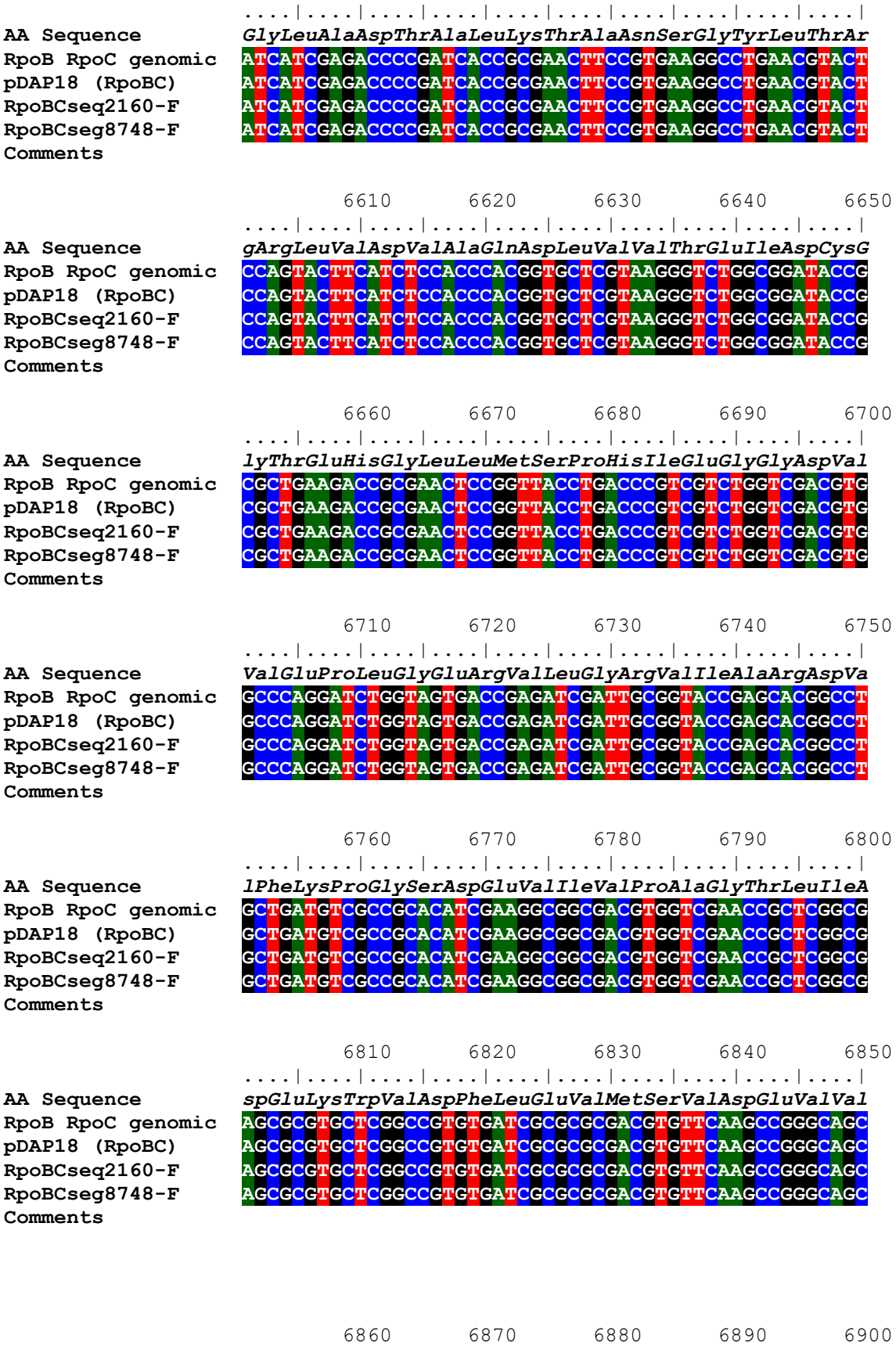

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
Reamp-F_RpoBC-F
RpoBCseq360-F
RpoBCseg7096-F
RpoBCseg8815-R*
Comments
5210
5220
5230
5240
5250
5260
5270
5280
5290
5300
5310
5320
5330
5340
5350
5360
5370
5380
5390
5400
5410
5420
5430
5440
5450
48
49

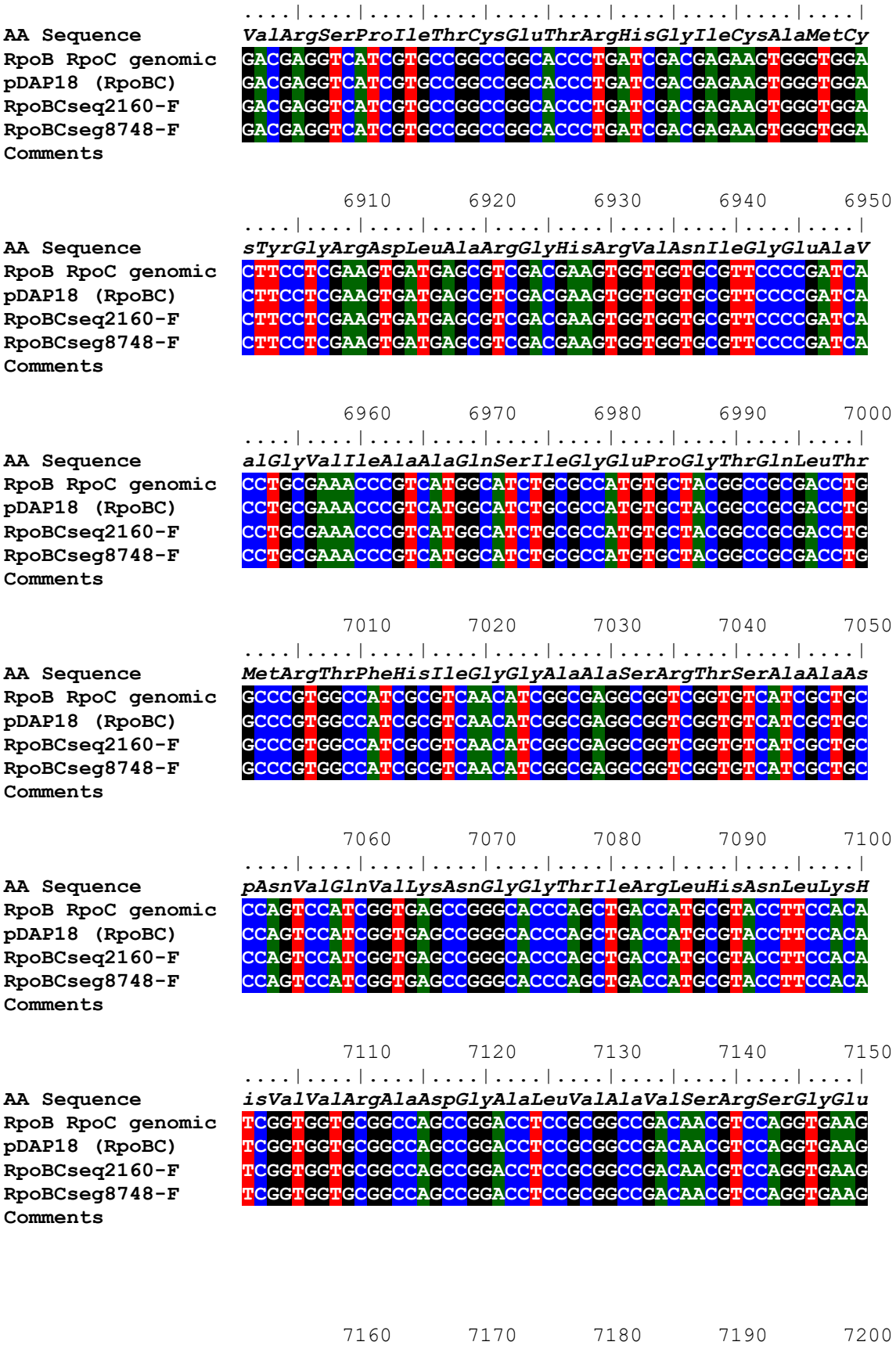

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
Reamp-F_RpoBC-F
RpoBCseq360-F
RpoBCseg7096-F
RpoBCseg8815-R*
Comments
sLysMetValGluArgGluLeuProGluValTrpAspValLeuAlaGluV
50
51
52
53
5460
5470
5480
5490
5500
alIleArgGluHisProValLeuLeuAsnArgAlaProThrLeuHisArg
54
5510
5520
5530
5540
5550
LeuGlyIleGlnAlaPheGluProValLeuIleGluGlyLysAlaIleGl
5560
5570
5580
5590
5600
nLeuHisProLeuValCysAlaAlaTyrAsnAlaAspPheAspGlyAspG
5610
5620
5630
5640
5650
lnMetAlaValHisValProLeuThrLeuGluAlaGlnLeuGluAlaArg
5660
5670
5680
5690
5700

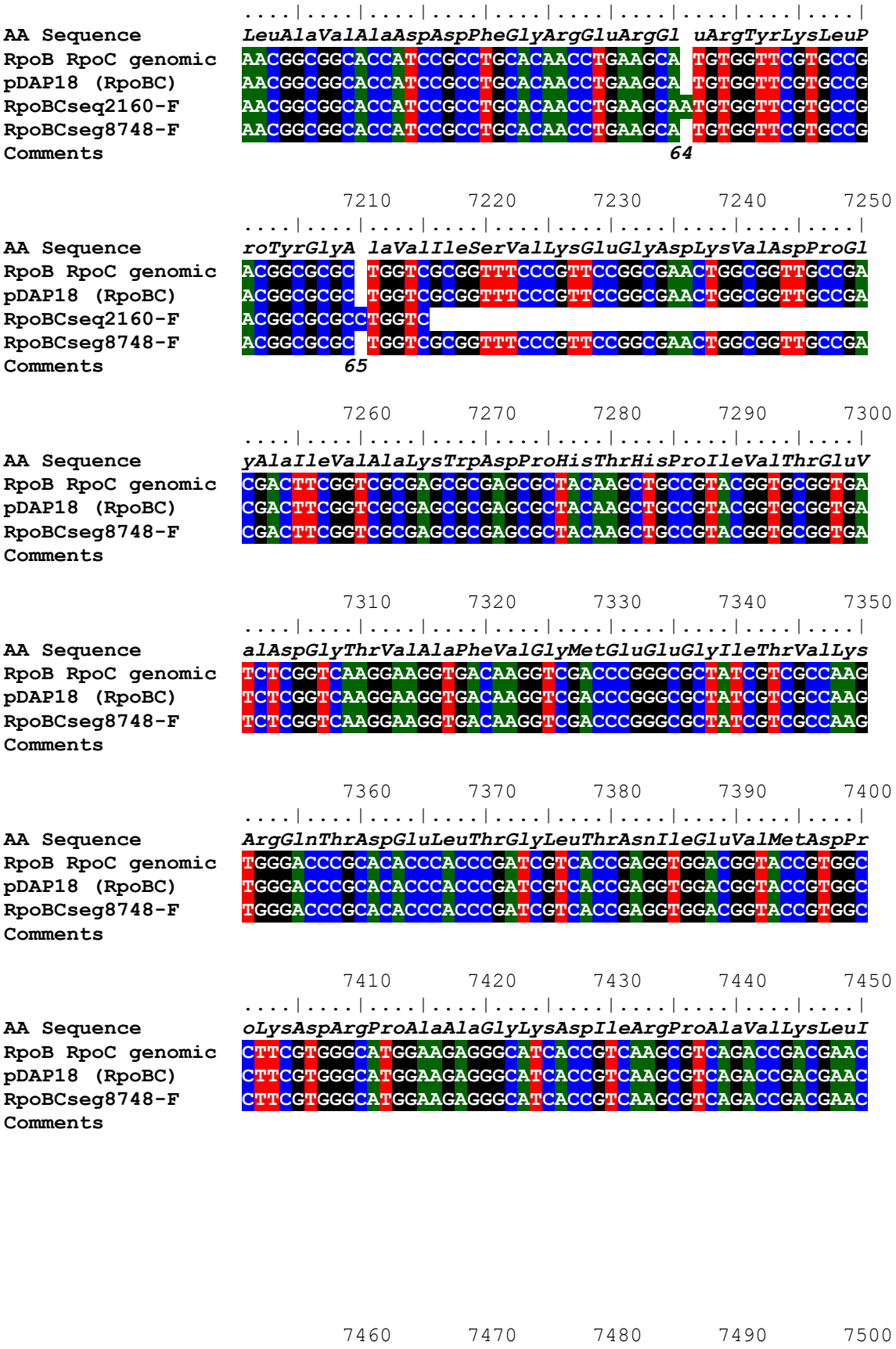

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg7096-F
RpoBCseg8815-R*
Comments
AlaLeuMetMetSerThrAsnAsnIleLeuSerProAlaAsnGlyGluPr
oIleIleValProSerGlnAspValValMetGlyLeuTyrTyrMetThrA
rgGluAlaIleAsnAlaLysGlyGluGlyMetAlaPheAlaAspLeuGln
GluValAspArgAlaTyrArgSerGlyGlnAlaSerLeuHisAlaArgVa
55
55
lLysValArgIleAsnGluLysIleLysGlyGluAspGlyGlnLeuThrA
laAsnThrArgIleValAspThrThrValGlyArgAlaLeuLeuPheGln
56
5710
5720
5730
5740
5750
5760
5770
5780
5790
5800
5810
5820
5830
5840
5850
5860
5870
5880
5890
5900
5910
5920
5930
5940
5950
5960
5970
5980
5990
6000

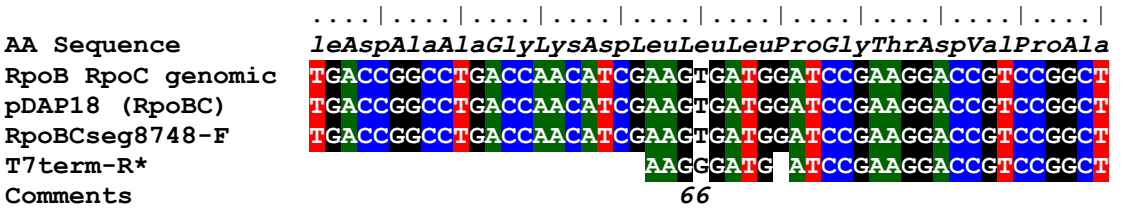

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg7096-F
RpoBCseg8815-R*
RpoBCseq2160-F
Comments
ValValProAlaGlyLeuProPheAspValValAsnGlnSerMetLysLy
sLysAlaIleSerLysLeuIleAsnHisCysTyrArgValValGlyLeuL
ysAspThrValIlePheAlaAspGlnLeuMetTyrThrGlyPheAlaTyr
SerThrIleSerGlyValSerIleGlyValAsnAspPheValIleProAs
pGluLysAlaArgIleIleAsnAlaAlaThrAspGluValLysGluIleG
luSerGlnTyrAlaSerGlyLeuValThrGlnGlyGluLysTyrAsnLys
57
58
6010
6020
6030
6040
6050
6060
6070
6080
6090
6100
6110
6120
6130
6140
6150
6160
6170
6180
6190
6200
6210
6220
6230
6240
6250
6260
6270
6280
6290
6300

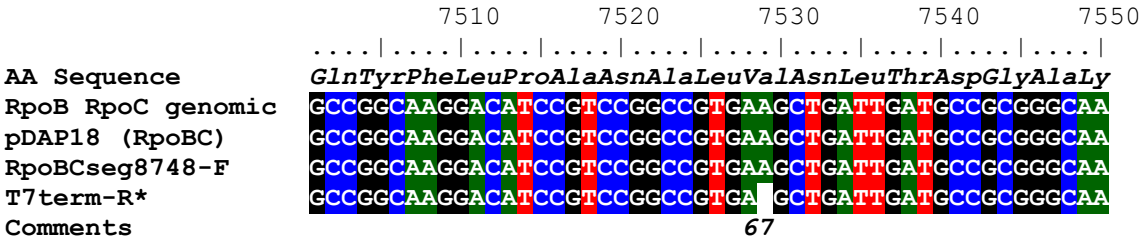

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg8815-R*
RpoBCseq2160-F
RpoBCseg8748-F
Comments
ValIleAspLeuTrpSerLysAlaAsnAspGluValSerLysAlaMetMe
tAlaAsnLeuSerLysGluLysValValAspArgGluGlyLysGluValA
spGlnGluSerPheAsnSerMetTyrMetMetAlaAspSerGlyAlaArg
GlySerAlaAlaGlnIleArgGlnLeuAlaGlyMetArgGlyLeuMetAl
aLysProAspGlySerIleIleGluThrProIleThrAlaAsnPheArgG
luGlyLeuAsnValLeuGlnTyrPheIleSerThrHisGlyAlaArgLys
59
60
61
62
63
6310
6320
6330
6340
6350
6360
6370
6380
6390
6400
6410
6420
6430
6440
6450
6460
6470
6480
6490
6500
6510
6520
6530
6540
6550
6560
6570
6580
6590
6600

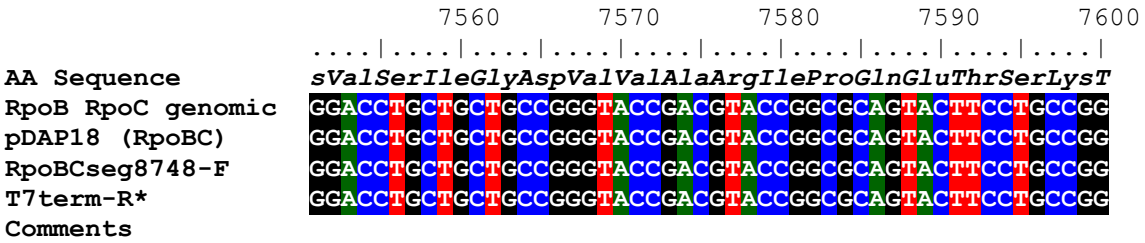

....|....|....|....|....|....|....|....|....|....|
AA Sequence GlyLeuAlaAspThrAlaLeuLysThrAlaAsnSerGlyTyrLeuThrAr
RpoB RpoC genomic ATCATCGAGACCCCGATCACCGCGAACTTCCGTGAAGGCCTGAACGTACT
pDAP18 (RpoBC) ATCATCGAGACCCCGATCACCGCGAACTTCCGTGAAGGCCTGAACGTACT
RpoBCseq2160-F ATCATCGAGACCCCGATCACCGCGAACTTCCGTGAAGGCCTGAACGTACT
RpoBCseg8748-F ATCATCGAGACCCCGATCACCGCGAACTTCCGTGAAGGCCTGAACGTACT
Comments
6610 6620 6630 6640 6650
....|....|....|....|....|....|....|....|....|....|
AA Sequence gArgLeuValAspValAlaGlnAspLeuValValThrGluIleAspCysG
RpoB RpoC genomic CCAGTACTTCATCTCCACCCACGGTGCTCGTAAGGGTCTGGCGGATACCG
pDAP18 (RpoBC) CCAGTACTTCATCTCCACCCACGGTGCTCGTAAGGGTCTGGCGGATACCG
RpoBCseq2160-F CCAGTACTTCATCTCCACCCACGGTGCTCGTAAGGGTCTGGCGGATACCG
RpoBCseg8748-F CCAGTACTTCATCTCCACCCACGGTGCTCGTAAGGGTCTGGCGGATACCG
Comments
6660 6670 6680 6690 6700
....|....|....|....|....|....|....|....|....|....|
AA Sequence lyThrGluHisGlyLeuLeuMetSerProHisIleGluGlyGlyAspVal
RpoB RpoC genomic CGCTGAAGACCGCGAACTCCGGTTACCTGACCCGTCGTCTGGTCGACGTG
pDAP18 (RpoBC) CGCTGAAGACCGCGAACTCCGGTTACCTGACCCGTCGTCTGGTCGACGTG
RpoBCseq2160-F CGCTGAAGACCGCGAACTCCGGTTACCTGACCCGTCGTCTGGTCGACGTG
RpoBCseg8748-F CGCTGAAGACCGCGAACTCCGGTTACCTGACCCGTCGTCTGGTCGACGTG
Comments
6710 6720 6730 6740 6750
....|....|....|....|....|....|....|....|....|....|
AA Sequence ValGluProLeuGlyGluArgValLeuGlyArgValIleAlaArgAspVa
RpoB RpoC genomic GCCCAGGATCTGGTAGTGACCGAGATCGATTGCGGTACCGAGCACGGCCT
pDAP18 (RpoBC) GCCCAGGATCTGGTAGTGACCGAGATCGATTGCGGTACCGAGCACGGCCT
RpoBCseq2160-F GCCCAGGATCTGGTAGTGACCGAGATCGATTGCGGTACCGAGCACGGCCT
RpoBCseg8748-F GCCCAGGATCTGGTAGTGACCGAGATCGATTGCGGTACCGAGCACGGCCT
Comments
6760 6770 6780 6790 6800
....|....|....|....|....|....|....|....|....|....|
AA Sequence lPheLysProGlySerAspGluValIleValProAlaGlyThrLeuIleA
RpoB RpoC genomic GCTGATGTCGCCGCACATCGAAGGCGGCGACGTGGTCGAACCGCTCGGCG
pDAP18 (RpoBC) GCTGATGTCGCCGCACATCGAAGGCGGCGACGTGGTCGAACCGCTCGGCG
RpoBCseq2160-F GCTGATGTCGCCGCACATCGAAGGCGGCGACGTGGTCGAACCGCTCGGCG
RpoBCseg8748-F GCTGATGTCGCCGCACATCGAAGGCGGCGACGTGGTCGAACCGCTCGGCG
Comments
6810 6820 6830 6840 6850
....|....|....|....|....|....|....|....|....|....|
AA Sequence spGluLysTrpValAspPheLeuGluValMetSerValAspGluValVal
RpoB RpoC genomic AGCGCGTGCTCGGCCGTGTGATCGCGCGCGACGTGTTCAAGCCGGGCAGC
pDAP18 (RpoBC) AGCGCGTGCTCGGCCGTGTGATCGCGCGCGACGTGTTCAAGCCGGGCAGC
RpoBCseq2160-F AGCGCGTGCTCGGCCGTGTGATCGCGCGCGACGTGTTCAAGCCGGGCAGC
RpoBCseg8748-F AGCGCGTGCTCGGCCGTGTGATCGCGCGCGACGTGTTCAAGCCGGGCAGC
Comments
6860 6870 6880 6890 6900

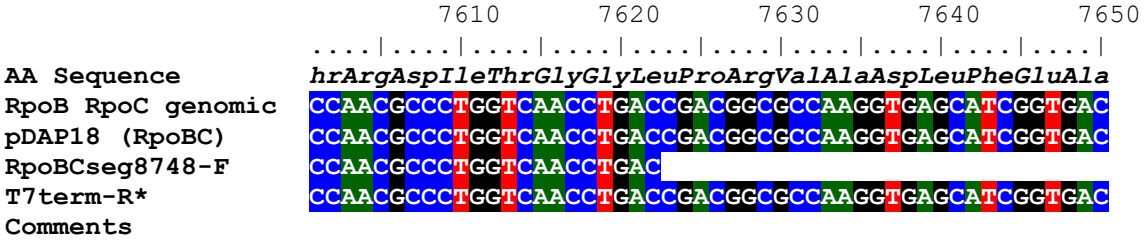

....|....|....|....|....|....|....|....|....|....|
AA Sequence ValArgSerProIleThrCysGluThrArgHisGlyIleCysAlaMetCy
RpoB RpoC genomic GACGAGGTCATCGTGCCGGCCGGCACCCTGATCGACGAGAAGTGGGTGGA
pDAP18 (RpoBC) GACGAGGTCATCGTGCCGGCCGGCACCCTGATCGACGAGAAGTGGGTGGA
RpoBCseq2160-F GACGAGGTCATCGTGCCGGCCGGCACCCTGATCGACGAGAAGTGGGTGGA
RpoBCseg8748-F GACGAGGTCATCGTGCCGGCCGGCACCCTGATCGACGAGAAGTGGGTGGA
Comments
6910 6920 6930 6940 6950
....|....|....|....|....|....|....|....|....|....|
AA Sequence sTyrGlyArgAspLeuAlaArgGlyHisArgValAsnIleGlyGluAlaV
RpoB RpoC genomic CTTCCTCGAAGTGATGAGCGTCGACGAAGTGGTGGTGCGTTCCCCGATCA
pDAP18 (RpoBC) CTTCCTCGAAGTGATGAGCGTCGACGAAGTGGTGGTGCGTTCCCCGATCA
RpoBCseq2160-F CTTCCTCGAAGTGATGAGCGTCGACGAAGTGGTGGTGCGTTCCCCGATCA
RpoBCseg8748-F CTTCCTCGAAGTGATGAGCGTCGACGAAGTGGTGGTGCGTTCCCCGATCA
Comments
6960 6970 6980 6990 7000
....|....|....|....|....|....|....|....|....|....|
AA Sequence alGlyValIleAlaAlaGlnSerIleGlyGluProGlyThrGlnLeuThr
RpoB RpoC genomic CCTGCGAAACCCGTCATGGCATCTGCGCCATGTGCTACGGCCGCGACCTG
pDAP18 (RpoBC) CCTGCGAAACCCGTCATGGCATCTGCGCCATGTGCTACGGCCGCGACCTG
RpoBCseq2160-F CCTGCGAAACCCGTCATGGCATCTGCGCCATGTGCTACGGCCGCGACCTG
RpoBCseg8748-F CCTGCGAAACCCGTCATGGCATCTGCGCCATGTGCTACGGCCGCGACCTG
Comments
7010 7020 7030 7040 7050
....|....|....|....|....|....|....|....|....|....|
AA Sequence MetArgThrPheHisIleGlyGlyAlaAlaSerArgThrSerAlaAlaAs
RpoB RpoC genomic GCCCGTGGCCATCGCGTCAACATCGGCGAGGCGGTCGGTGTCATCGCTGC
pDAP18 (RpoBC) GCCCGTGGCCATCGCGTCAACATCGGCGAGGCGGTCGGTGTCATCGCTGC
RpoBCseq2160-F GCCCGTGGCCATCGCGTCAACATCGGCGAGGCGGTCGGTGTCATCGCTGC
RpoBCseg8748-F GCCCGTGGCCATCGCGTCAACATCGGCGAGGCGGTCGGTGTCATCGCTGC
Comments
7060 7070 7080 7090 7100
....|....|....|....|....|....|....|....|....|....|
AA Sequence pAsnValGlnValLysAsnGlyGlyThrIleArgLeuHisAsnLeuLysH
RpoB RpoC genomic CCAGTCCATCGGTGAGCCGGGCACCCAGCTGACCATGCGTACCTTCCACA
pDAP18 (RpoBC) CCAGTCCATCGGTGAGCCGGGCACCCAGCTGACCATGCGTACCTTCCACA
RpoBCseq2160-F CCAGTCCATCGGTGAGCCGGGCACCCAGCTGACCATGCGTACCTTCCACA
RpoBCseg8748-F CCAGTCCATCGGTGAGCCGGGCACCCAGCTGACCATGCGTACCTTCCACA
Comments
7110 7120 7130 7140 7150
....|....|....|....|....|....|....|....|....|....|
AA Sequence isValValArgAlaAspGlyAlaLeuValAlaValSerArgSerGlyGlu
RpoB RpoC genomic TCGGTGGTGCGGCCAGCCGGACCTCCGCGGCCGACAACGTCCAGGTGAAG
pDAP18 (RpoBC) TCGGTGGTGCGGCCAGCCGGACCTCCGCGGCCGACAACGTCCAGGTGAAG
RpoBCseq2160-F TCGGTGGTGCGGCCAGCCGGACCTCCGCGGCCGACAACGTCCAGGTGAAG
RpoBCseg8748-F TCGGTGGTGCGGCCAGCCGGACCTCCGCGGCCGACAACGTCCAGGTGAAG
Comments
7160 7170 7180 7190 7200

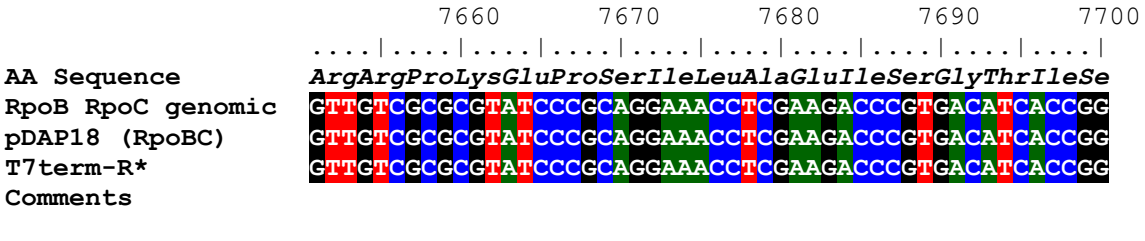

AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseq2160-F
RpoBCseg8748-F
Comments
64
65
7210
7220
7230
7240
7250
7260
7270
7280
7290
7300
7310
7320
7330
7340
7350
7360
7370
7380
7390
7400
7410
7420
7430
7440
7450
7460
7470
7480
7490
7500

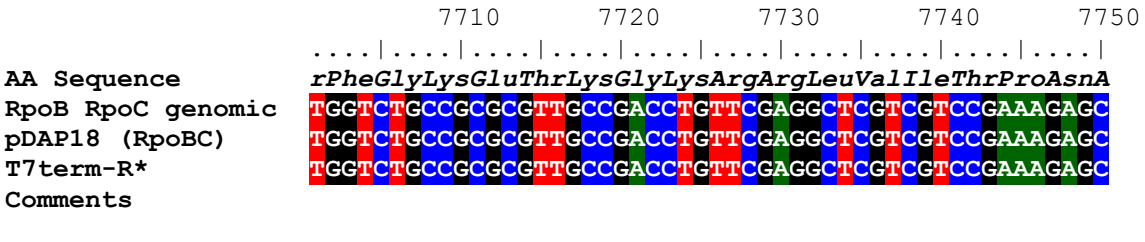

AA Sequence
leAspAlaAlaGlyLysAspLeuLeuLeuProGlyThrAspValProAla
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg8748-F
T7term-R*
Comments
66


7510
7520
7530
7540
7550
AA Sequence
GlnTyrPheLeuProAlaAsnAlaLeuValAsnLeuThrAspGlyAlaLy
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg8748-F
T7term-R*
Comments
67

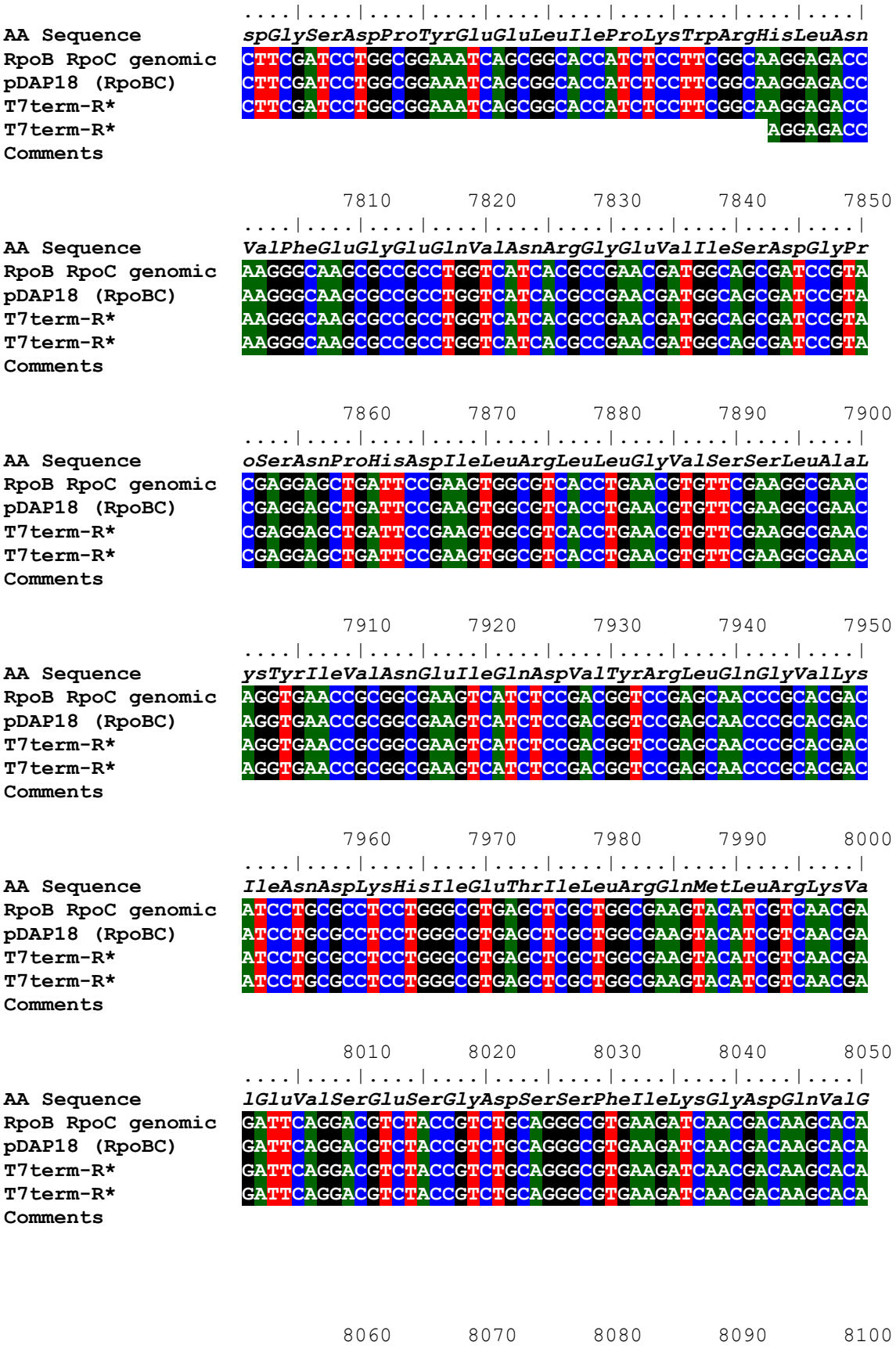

7560
7570
7580
7590
7600
AA Sequence
sValSerIleGlyAspValValAlaArgIleProGlnGluThrSerLysT
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg8748-F
T7term-R*
Comments

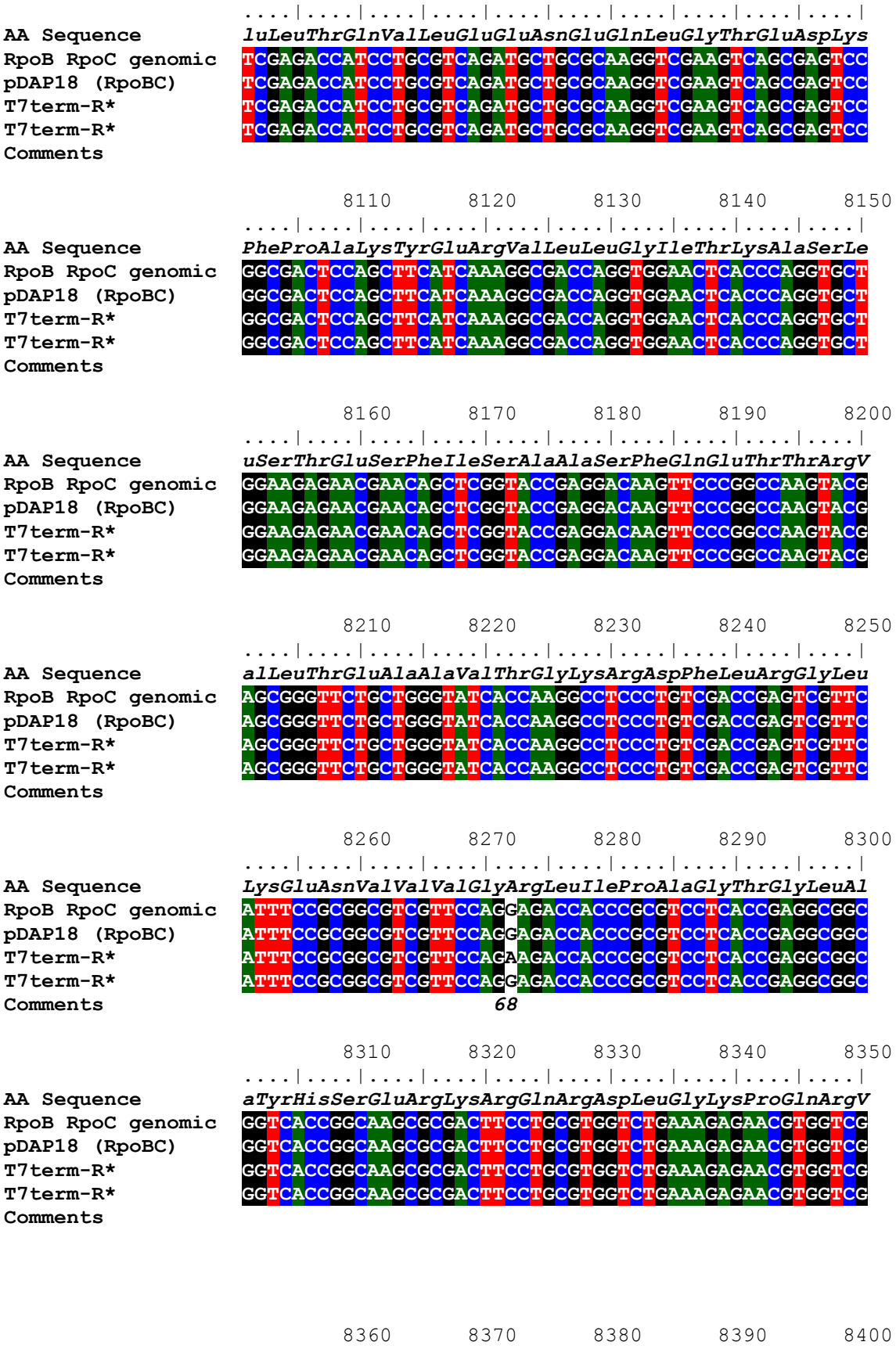

7610
7620
7630
7640
7650
AA Sequence
hrArgAspIleThrGlyGlyLeuProArgValAlaAspLeuPheGluAla
RpoB RpoC genomic
pDAP18 (RpoBC)
RpoBCseg8748-F
T7term-R*
Comments

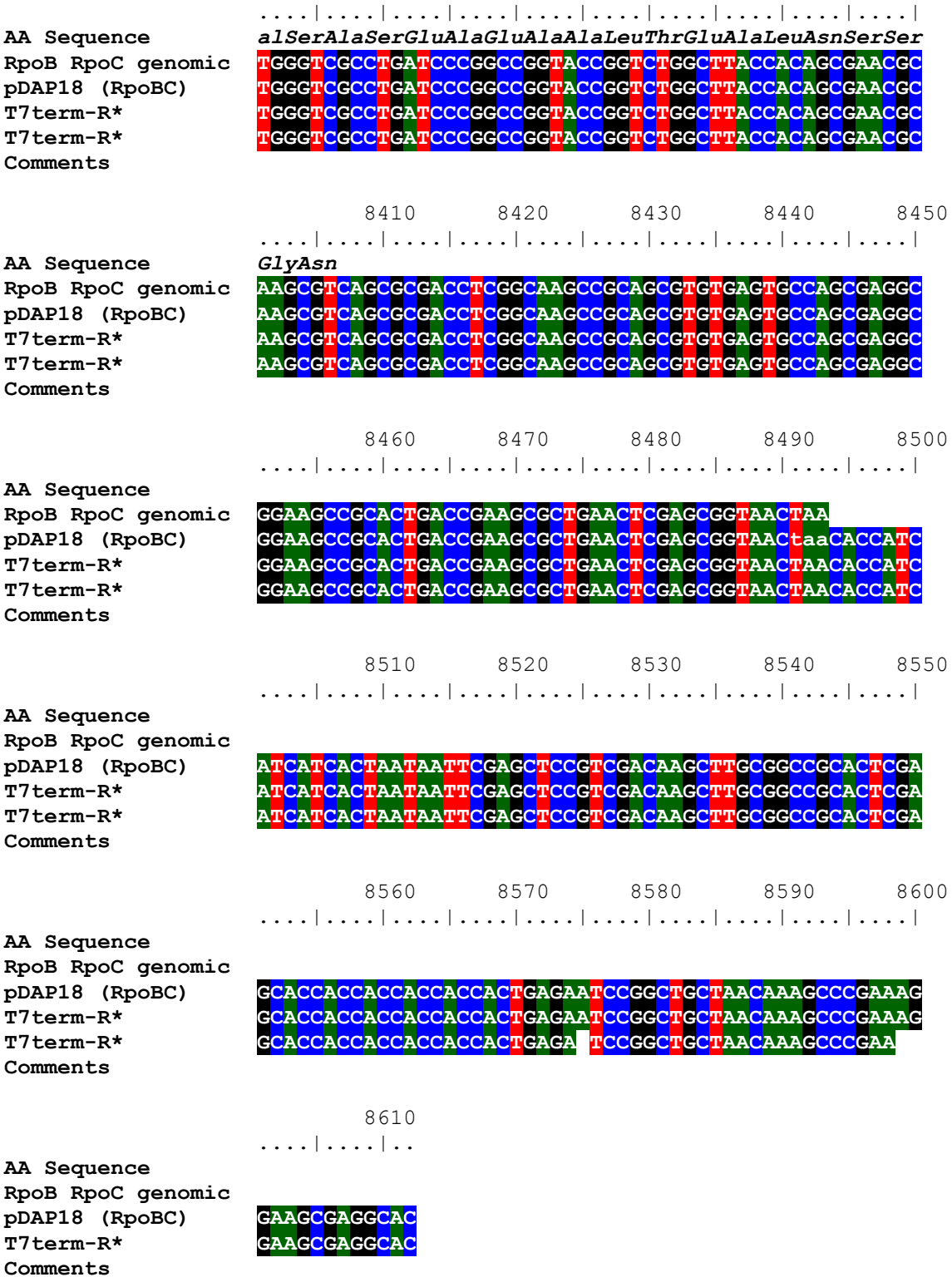

7660
7670
7680
7690
7700
AA Sequence
ArgArgProLysGluProSerIleLeuAlaGluIleSerGlyThrIleSe
RpoB RpoC genomic
pDAP18 (RpoBC)
T7term-R*
Comments


7710
7720
7730
7740
7750
AA Sequence
rPheGlyLysGluThrLysGlyLysArgArgLeuValIleThrProAsnA
RpoB RpoC genomic
pDAP18 (RpoBC)
T7term-R*
Comments


7760
7770
7780
7790
7800


AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
T7term-R*
T7term-R*
Comments
spGlySerAspProTyrGluGluLeuIleProLysTrpArgHisLeuAsn
CTTCGATCCTGGCGGAAATCAGCGGCACCATCTCCTTCGGCAAGGAGACC
7810 7820 7830 7840 7850
ValPheGluGlyGluGlnValAsnArgGlyGluValIleSerAspGlyPr
AAGGGCAAGCGCCGCCTGGTCATCACGCCGAACGATGGCAGCGATCCGTA
7860 7870 7880 7890 7900
oSerAsnProHisAspIleLeuArgLeuLeuGlyValSerSerLeuAlaL
CGAGGAGCTGATTCCGAAGTGGCGTCACCTGAACGTGTTCGAAGGCGAAC
7910 7920 7930 7940 7950
ysTyrIleValAsnGluIleGlnAspValTyrArgLeuGlnGlyValLys
AGGTGAACCGCGGCGAAGTCATCTCCGACGGTCCGAGCAACCCGCACGAC
7960 7970 7980 7990 8000
IleAsnAspLysHisIleGluThrIleLeuArgGlnMetLeuArgLysVa
ATCCTGCGCCTCCTGGGCGTGAGCTCGCTGGCGAAGTACATCGTCAACGA
8010 8020 8030 8040 8050
lGluValSerGluSerGlyAspSerSerPheIleLysGlyAspGlnValG
GATTCAGGACGTCTACCGTCTGCAGGGCGTGAAGATCAACGACAAGCACA
8060 8070 8080 8090 8100


....|....|....|....|....|....|....|....|....|....|
AA Sequence luLeuThrGlnValLeuGluGluAsnGluGlnLeuGlyThrGluAspLys
RpoB RpoC genomic TCGAGACCATCCTGCGTCAGATGCTGCGCAAGGTCGAAGTCAGCGAGTCC
pDAP18 (RpoBC) TCGAGACCATCCTGCGTCAGATGCTGCGCAAGGTCGAAGTCAGCGAGTCC
T7term-R* TCGAGACCATCCTGCGTCAGATGCTGCGCAAGGTCGAAGTCAGCGAGTCC
T7term-R* TCGAGACCATCCTGCGTCAGATGCTGCGCAAGGTCGAAGTCAGCGAGTCC
Comments
8110 8120 8130 8140 8150
....|....|....|....|....|....|....|....|....|....|
AA Sequence PheProAlaLysTyrGluArgValLeuLeuGlyIleThrLysAlaSerLe
RpoB RpoC genomic GGCGACTCCAGCTTCATCAAAGGCGACCAGGTGGAACTCACCCAGGTGCT
pDAP18 (RpoBC) GGCGACTCCAGCTTCATCAAAGGCGACCAGGTGGAACTCACCCAGGTGCT
T7term-R* GGCGACTCCAGCTTCATCAAAGGCGACCAGGTGGAACTCACCCAGGTGCT
T7term-R* GGCGACTCCAGCTTCATCAAAGGCGACCAGGTGGAACTCACCCAGGTGCT
Comments
8160 8170 8180 8190 8200
....|....|....|....|....|....|....|....|....|....|
AA Sequence uSerThrGluSerPheIleSerAlaAlaSerPheGlnGluThrThrArgV
RpoB RpoC genomic GGAAGAGAACGAACAGCTCGGTACCGAGGACAAGTTCCCGGCCAAGTACG
pDAP18 (RpoBC) GGAAGAGAACGAACAGCTCGGTACCGAGGACAAGTTCCCGGCCAAGTACG
T7term-R* GGAAGAGAACGAACAGCTCGGTACCGAGGACAAGTTCCCGGCCAAGTACG
T7term-R* GGAAGAGAACGAACAGCTCGGTACCGAGGACAAGTTCCCGGCCAAGTACG
Comments
8210 8220 8230 8240 8250
....|....|....|....|....|....|....|....|....|....|
AA Sequence alLeuThrGluAlaAlaValThrGlyLysArgAspPheLeuArgGlyLeu
RpoB RpoC genomic AGCGGGTTCTGCTGGGTATCACCAAGGCCTCCCTGTCGACCGAGTCGTTC
pDAP18 (RpoBC) AGCGGGTTCTGCTGGGTATCACCAAGGCCTCCCTGTCGACCGAGTCGTTC
T7term-R* AGCGGGTTCTGCTGGGTATCACCAAGGCCTCCCTGTCGACCGAGTCGTTC
T7term-R* AGCGGGTTCTGCTGGGTATCACCAAGGCCTCCCTGTCGACCGAGTCGTTC
Comments
8260 8270 8280 8290 8300
....|....|....|....|....|....|....|....|....|....|
AA Sequence LysGluAsnValValValGlyArgLeuIleProAlaGlyThrGlyLeuAl
RpoB RpoC genomic ATTTCCGCGGCGTCGTTCCAGGAGACCACCCGCGTCCTCACCGAGGCGGC
pDAP18 (RpoBC) ATTTCCGCGGCGTCGTTCCAGGAGACCACCCGCGTCCTCACCGAGGCGGC
T7term-R* ATTTCCGCGGCGTCGTTCCAGAAGACCACCCGCGTCCTCACCGAGGCGGC
T7term-R* ATTTCCGCGGCGTCGTTCCAGGAGACCACCCGCGTCCTCACCGAGGCGGC
Comments 68
8310 8320 8330 8340 8350
....|....|....|....|....|....|....|....|....|....|
AA Sequence aTyrHisSerGluArgLysArgGlnArgAspLeuGlyLysProGlnArgV
RpoB RpoC genomic GGTCACCGGCAAGCGCGACTTCCTGCGTGGTCTGAAAGAGAACGTGGTCG
pDAP18 (RpoBC) GGTCACCGGCAAGCGCGACTTCCTGCGTGGTCTGAAAGAGAACGTGGTCG
T7term-R* GGTCACCGGCAAGCGCGACTTCCTGCGTGGTCTGAAAGAGAACGTGGTCG
T7term-R* GGTCACCGGCAAGCGCGACTTCCTGCGTGGTCTGAAAGAGAACGTGGTCG
Comments
8360 8370 8380 8390 8400


AA Sequence
RpoB RpoC genomic
pDAP18 (RpoBC)
T7term-R*
T7term-R*
Comments
alSerAlaSerGluAlaGluAlaAlaLeuThrGluAlaLeuAsnSerSer
TGGGTCGCCTGATCCCGGCCGGTACCGGTCTGGCTTACCACAGCGAACGC
TGGGTCGCCTGATCCCGGCCGGTACCGGTCTGGCTTACCACAGCGAACGC
TGGGTCGCCTGATCCCGGCCGGTACCGGTCTGGCTTACCACAGCGAACGC
TGGGTCGCCTGATCCCGGCCGGTACCGGTCTGGCTTACCACAGCGAACGC
8410
8420
8430
8440
8450
GlyAsn
AAGCGTCAGCGCGACCTCGGCAAGCCGCAGCGTGTGAGTGCCAGCGAGGC
AAGCGTCAGCGCGACCTCGGCAAGCCGCAGCGTGTGAGTGCCAGCGAGGC
AAGCGTCAGCGCGACCTCGGCAAGCCGCAGCGTGTGAGTGCCAGCGAGGC
AAGCGTCAGCGCGACCTCGGCAAGCCGCAGCGTGTGAGTGCCAGCGAGGC
8460
8470
8480
8490
8500
GGAAGCCGCACTGACCGAAGCGCTGAACTCGAGCGGTAACTAA
GGAAGCCGCACTGACCGAAGCGCTGAACTCGAGCGGTAACtaaCACCATC
GGAAGCCGCACTGACCGAAGCGCTGAACTCGAGCGGTAACTAACACCATC
GGAAGCCGCACTGACCGAAGCGCTGAACTCGAGCGGTAACTAACACCATC
8510
8520
8530
8540
8550
ATCATCACTAATAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGA
ATCATCACTAATAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGA
ATCATCACTAATAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGA
8560
8570
8580
8590
8600
GCACCACCACCACCACCACTGAGAATCCGGCTGCTAACAAAGCCCGAAAG
GCACCACCACCACCACCACTGAGAATCCGGCTGCTAACAAAGCCCGAAAG
GCACCACCACCACCACCACTGAGA TCCGGCTGCTAACAAAGCCCGAA
8610
GAAGCGAGGCAC
GAAGCGAGGCAC

Sequencing corrections and observations for plasmid pDAP18 (pRARE RpoBC)

Comment #	**Sequencing Reaction**	Changes/Observations
1	RpoBCseg2421-F	removed first 37. Sequence is consistent, but many errors in first part of sequence read.
2	RpoBCseg2421-F	GGG read reduced to GG since a doublet G is evident in sequencing read.
3	RpoBCseg4036-R*	many sequence errors, first 50 deleted. But trace is consistent with correct sequence.
4	RpoBCseg4036-R*	error G (C*) appears correctly sequenced, but does not agree with two other sequencing reactions
5	RpoBCseg4036-R*	G (C*) peak is present
6	RpoBCseg4036-R*	G (C*) peak is present under "A" (T*)
7	RpoBCseg4036-R*	G (C*) triplet is compressed, extra G removed.
8	T7-F	sequence from here on is poor.
9	T7-F	deleted remainder of sequence data.
10	RpoBCseg2421-F	CC peak is wide, compatible with correct seq
11	RpoBCseg2421-F	peak for missing C is present
12	Rp0BCpRA2seq7_seq10560-F	First 20 removed from 5'
	RpoBCseg2421-F	G peak is present under A's
13	RpoBCseg2421-F	G peak is present under A's
14	RpoBCseg2421-F	multiple sequence read errors at end of sequence (ignoring)
15	RpoBCseg4036-R* & RpoBCseg3985-F,	ends of reads deleted
16	RpoBCseg3885-F	only one T peak in sequencing read, T deleted from
17	RpoBCseg5603-R*	first 43 nucleotides are poor, deleted
18	RpoBCseg5603-R*	Single A peak (T*) instead of doublet and a C (G*) peak is present, but not counted. Genomic sequence is correct.
19	RpoBCseg3985-F	T peak was ignored, but is present. C peak is smeared
20	RpoBCseg3985-F	A triple has a A doublet and small G peak (not read)
21	RpoBCseg3985-F	three G's are a doublet, G deleted.
22	RpoBCseg3985-F	last nucleotides are poor, deleted to end
23	RpoBCseq12240-F	First 6 nucleotides removed from 5'
24	RpoBCseq12240-F	Missing C. Peak is wide, consistent with CCC.
25	Reamp-R_RpoBC-R*	tiny G (C*) present between A's. Genomic correct.
26	Reamp-R_RpoBC-R*	quartet of G's (C*) is really three peaks with compressed G (removed extra C)
27	RpoBCseg5517-F	first 6 nucleotides removed at beginning
28	Reamp-R_RpoBC-R*	double of T's (A*) is one peak (removed extra A)
	RpoBCseg5603-R*	first 17 nucleotides removed- poor read (end of complement)
29	RpoBCseg5517-F	double of A's is one peak (removed extra A)
30	RpoBCseg5517-F	G quartet is three peaks (extra G removed)
31	RpoBCseg5517-F	no clear A peak, A removed
32		BspQI RE site knocked out in cloning procedure- silent mutations introduced.

33	RpoBCseg7149-R*	C peak (G*) is evident in run of A's, added G; small G peak is evident and 4A's are not supported. Deleted last reads as smearing.
34	Reamp-R_RpoBC-R*	beginning of read is poor, deleted first 11 nucleotides (end of complement)
35	RpoBCseg5517-F	sequence around 1016-1028 is misread in trace due to low peaks. Sequence corrected. Sequence does not matter in this region.
36	RpoBCseg5517-F	small A peak is present under small C peak. Sequence difference does not matter.
37	RpoBCseg5517-F	small A peak is present under dip in C peaks. Sequence changed to A.
38	RpoBCseg5517-F	sequence read to end is poor (but consistent with correct sequence), nucleotides at 3' deleted.
39		Start codon in rpoB is TTG. This was mutated to ATG for expression and a 10xHis-TEV site encoded into the DNA.
40	Reamp-F_RpoBC-F	sequence read at beginning is poor. Removed first 20 nucleotides at 5' (removed end of complement). Sequence was compatible with correct.
41	RpoBCSeq360-F	10 nucleotides deleted
42	RpoBCSeq360-F	A peak is wide, consistent with correct seq
43	RpoBCseg7149-R*	sequence read at beginning is poor. Removed first 20 nucleotides at 5' (removed end of complement). Sequence was compatible with correct.
44	RpoBCseg7096-F	first 18 removed
45	RpoBCseg7096-F	A peak is wide, consistent with AA and not single A
46	RpoBCseg7096-F	double T in sequence is not consistent with single T peak.
47	RpoBCseg7096-F	double T in sequence is not consistent with single T peak. Removed 2nd T.
48	RpoBCseg8815-R*	End of read deleted
49	RpoBCseg8815-R*	sequence between 43 and 45 is poor with peaks smeared (at end of read, beginning of complement)
50	Reamp-F_RpoBC-F	Run of A's have G's and gaps in trace (small peaks) consistent with sequence
51	RpoBCseg8815-R*	small peck for A under T peak. Sequence is consistent with correct.
52	Reamp-F_RpoBC-F	single A is wide peak consistent with double A
53	Reamp-F_RpoBC-F	very wide A peaks are consistent with missing A
54	Reamp-F_RpoBC-F	sequence is getting poor, though consistent with correct. Deleted last 14 nucleotides.
55	RpoBCseg7096-F	peaks are smeary and overlapping in this region with doublets for single nucleotides. Extra A's removed from sequence.
56	RpoBCseg7096-F	G peak is present in trace but covered by large A peak. A replaced with G.
57	RpoBCseg7096-F	3' reads are now getting poor, but are consistent with correct sequence. Deleted from this point.
58	RpoBCseq2160-F	Beginning of read 11 nucleotides deleted

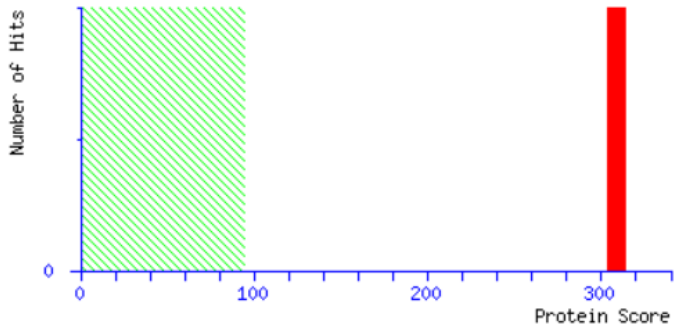
59	RpoBCseg8815-R*	CCC (GGG*) is two peaks. Replaced with two C's. A peak is sitting under T peak. Removed T* from sequence.
60	RpoBCseg8815-R*	Extra A added sitting under G peak too close, T* removed from sequence.
61	RpoBCseg8815-R*	Clear T peak where two A's reported. A replaced T
62	RpoBCseg8815-R*	read is poor before nucleotide 14. 14 deleted at 5' .
63	RpoBCseg8748-F	read is poor before 15. First 15 deleted.
64	RpoBCseq2160-F	Double A has clear double peak. Near end of run. Ambiguous Seq8748-F is very clear single peak.
65	RpoBCseq2160-F	Double C has clear double peak. Near end of run. Ambiguous. Seq8748-F is very clear single peak.
66	pDAP18 T7term-R*	A signal is evident under C peak. Region is not clear.
	pDAP18 T7term-R*	C doublet (G*G*) is not evident. Other sequencing (RpoBCseg8748-F) is correct.
67	pDAP18 T7term-R*	Missing T (A*) in sequence. T peak is wide, consistent with missing T.
68	pDAP18 T7term-R*	T doublet is clear in pDAP18 T7term-R*. C double is clear in pDAP18 (2) T7term-R*. Ambiguous.

Mass spectrometry analysis of RNA polymerase (snapshot data)

Mass fingerprinting analysis of *P. aeruginosa* RpoA subunit

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 94 are significant (p<0.05).



Concise Protein Summary Report

Format As	Concise Protein Summary	Help
	Significance threshold p< 0.05	Max. number of hits AUTO
	Preferred taxonomy All entries	

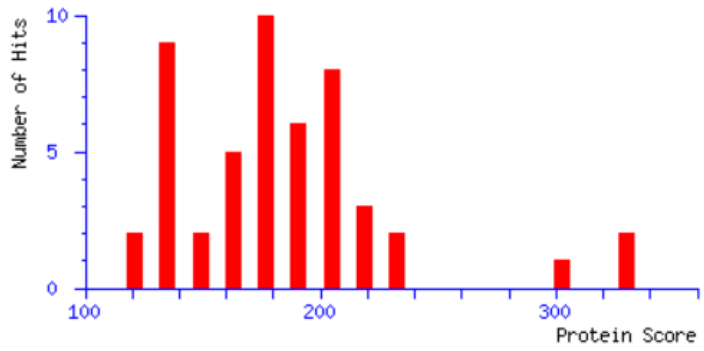
Re-Search All Search Unmatched

1. WP_023114709.1 **Mass:** 36657 **Score:** 309 **Expect:** 1.6e-23 **Matches:** 26
DNA-directed RNA polymerase subunit alpha [Pseudomonas aeruginosa]
WP_034046828.1 **Mass:** 36569 **Score:** 309 **Expect:** 1.6e-23 **Matches:** 26
DNA-directed RNA polymerase subunit alpha [Pseudomonas aeruginosa]
WP_034053091.1 **Mass:** 36626 **Score:** 309 **Expect:** 1.6e-23 **Matches:** 26
DNA-directed RNA polymerase subunit alpha [Pseudomonas aeruginosa]
WP_058011971.1 **Mass:** 36613 **Score:** 309 **Expect:** 1.6e-23 **Matches:** 26
DNA-directed RNA polymerase subunit alpha [Pseudomonas aeruginosa]
WP_058157974.1 **Mass:** 36555 **Score:** 309 **Expect:** 1.6e-23 **Matches:** 26
DNA-directed RNA polymerase subunit alpha [Pseudomonas aeruginosa]
WP_058163594.1 **Mass:** 36613 **Score:** 309 **Expect:** 1.6e-23 **Matches:** 26
MULTISPECIES: DNA-directed RNA polymerase subunit alpha [Pseudomonas]

Mass fingerprinting analysis of *P. aeruginosa* RpoD subunit

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 94 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

Preferred taxonomy: All entries

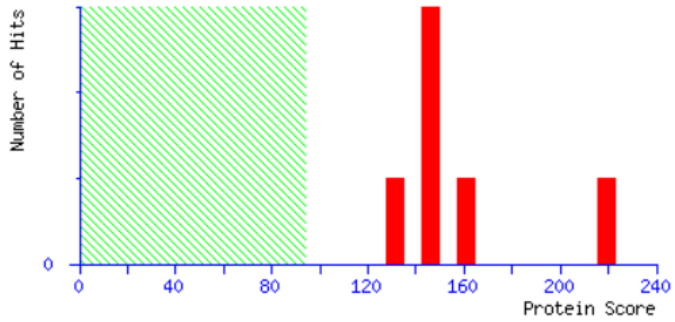
Re-Search All | Search Unmatched

1. WP_023130542.1 **Mass:** 69581 **Score:** 330 **Expect:** 1.3e-25 **Matches:** 48
RNA polymerase sigma factor RpoD [Pseudomonas aeruginosa]
WP_049269744.1 **Mass:** 69557 **Score:** 330 **Expect:** 1.3e-25 **Matches:** 48
RNA polymerase sigma factor RpoD [Pseudomonas aeruginosa]
WP_023112708.1 **Mass:** 69627 **Score:** 329 **Expect:** 1.6e-25 **Matches:** 48
RNA polymerase sigma factor RpoD [Pseudomonas aeruginosa]
WP_023083946.1 **Mass:** 69600 **Score:** 328 **Expect:** 2e-25 **Matches:** 48
RNA polymerase sigma factor RpoD [Pseudomonas aeruginosa]
WP_003113215.1 **Mass:** 69612 **Score:** 327 **Expect:** 2.5e-25 **Matches:** 48
RNA polymerase sigma factor RpoD [Pseudomonas aeruginosa]

Mass fingerprinting analysis of *P. aeruginosa* RpoBC subunit

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 94 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

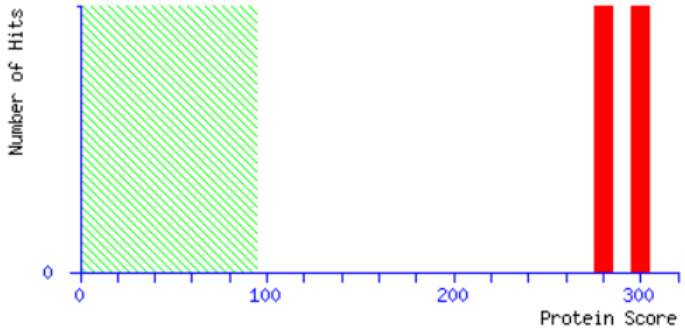
Re-Search All | Search Unmatched

1. **Mixture 1** **Total score:** 219 **Expect:** 1.6e-14 **Matches:** 117
Components (only one family member shown for each component):
CDH75131.1 **Mass:** 148202 **Score:** 157 **Expect:** 2.5e-08 **Matches:** 69
DNA-directed RNA polymerase subunit beta [Pseudomonas aeruginosa MH27]
WP_058148956.1 **Mass:** 154288 **Score:** 128 **Expect:** 2e-05 **Matches:** 64
DNA-directed RNA polymerase subunit beta' [Pseudomonas aeruginosa]

Mass fingerprinting analysis of RNAP contaminated with *E. coli* RNAP

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 94 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary — Help

Significance threshold p< 0.05 — Max. number of hits AUTO

Preferred taxonomy: All entries

Re-Search All | Search Unmatched

1. WP_001162098.1 **Mass:** 36508 **Score:** 300 **Expect:** 1.3e-22 **Matches:** 25
DNA-directed RNA polymerase subunit alpha [Salmonella enterica]
WP_001162094.1 **Mass:** 36489 **Score:** 281 **Expect:** 1e-20 **Matches:** 24
MULTISPECIES: DNA-directed RNA polymerase subunit alpha [Proteobacteria]
CAA37838.1 **Mass:** 36436 **Score:** 281 **Expect:** 1e-20 **Matches:** 24
unnamed protein product, partial [Escherichia coli]
WP_001404504.1 **Mass:** 36485 **Score:** 281 **Expect:** 1e-20 **Matches:** 24
DNA-directed RNA polymerase subunit alpha [Escherichia coli]
WP_001416928.1 **Mass:** 36449 **Score:** 281 **Expect:** 1e-20 **Matches:** 24
DNA-directed RNA polymerase subunit alpha [Escherichia coli]
WP_001420238.1 **Mass:** 36547 **Score:** 281 **Expect:** 1e-20 **Matches:** 24

Appendix B (Chapter 3)

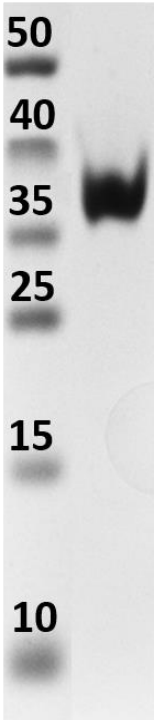


Figure A3-1. SDS-PAGE analysis of purified *P. aeruginosa* FinR (~ 37 kDa). Gel shows the final purified protein fraction from the Q column.

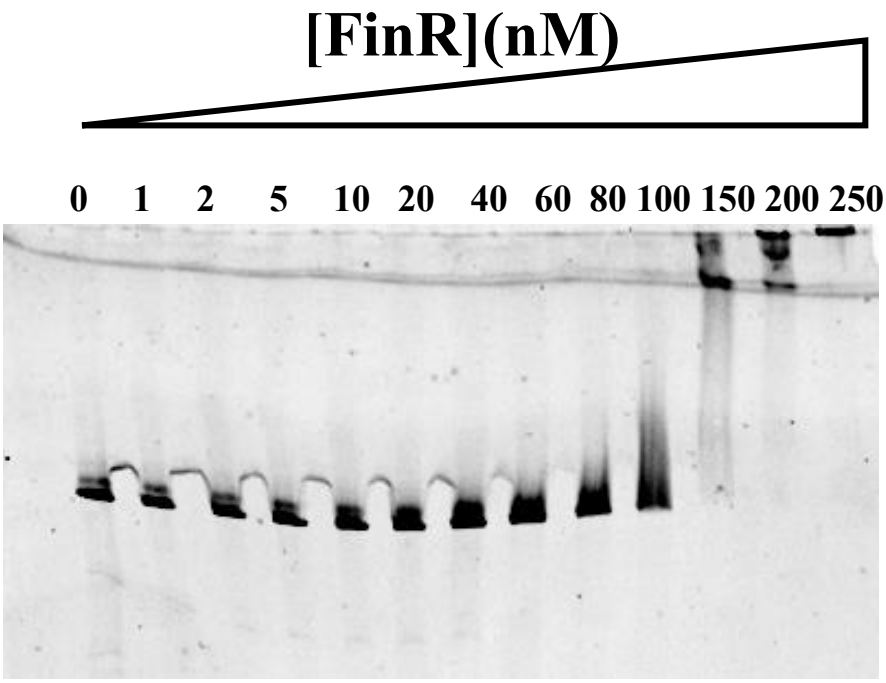


Figure A3-2. Binding of purified *P. aeruginosa* FinR to 2 nM of a DNA fragment containing the *P. aeruginosa cysI* promoter.

EMSA binding curves used for Kd determination.

(FinR concentrations: 0, 1, 2, 5, 10, 25, 50, 100, 150, 200, 250, 500, 750, 1000 nM)

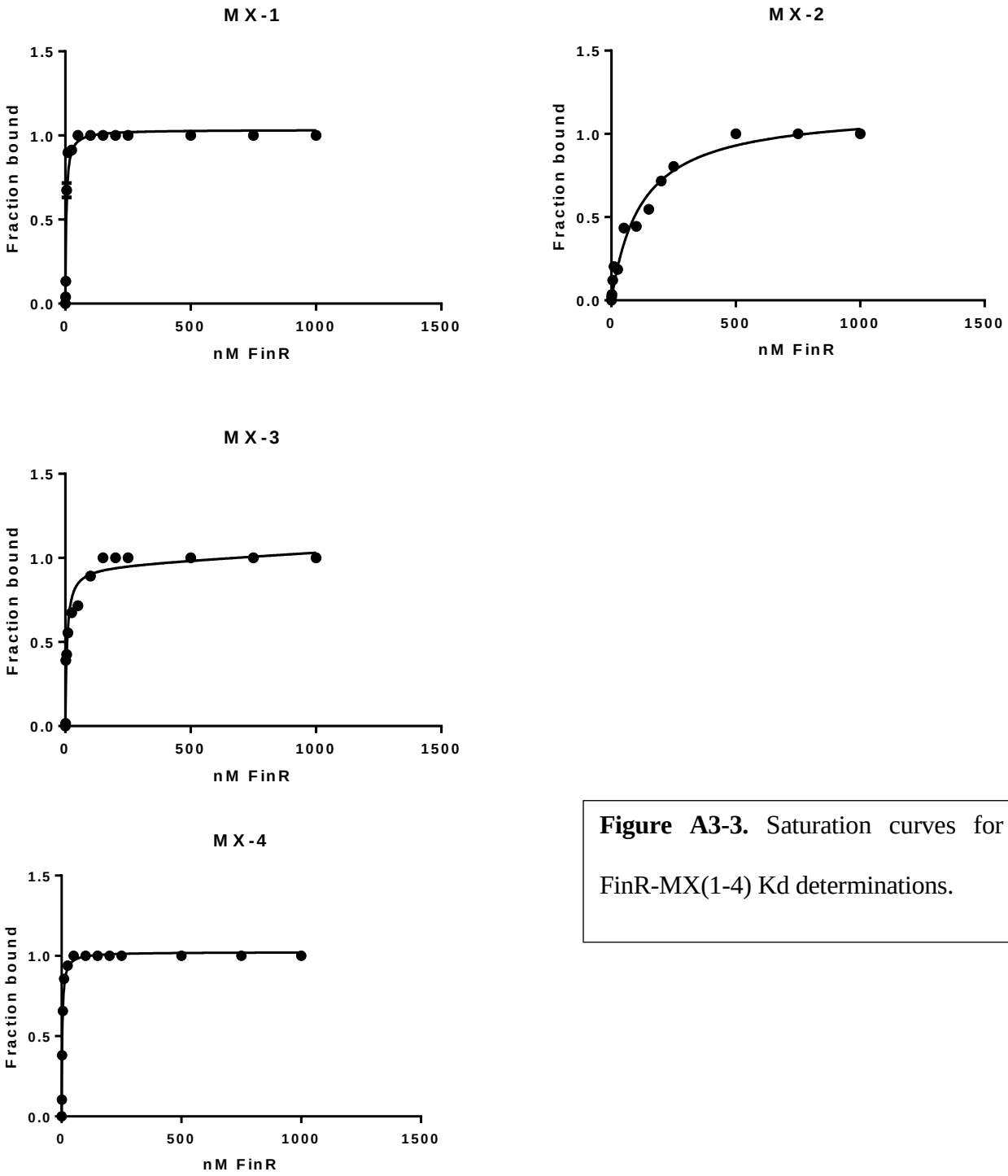
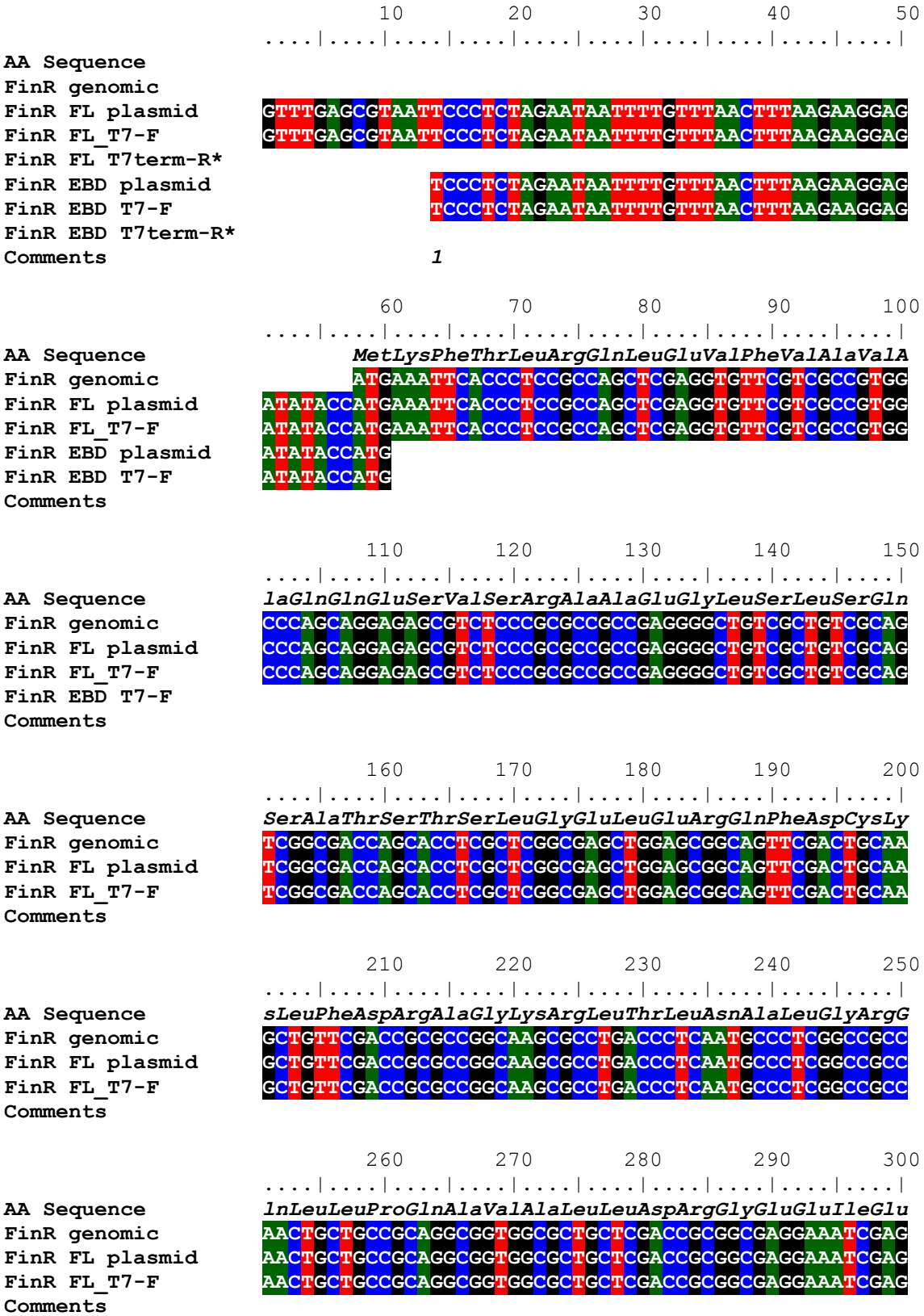
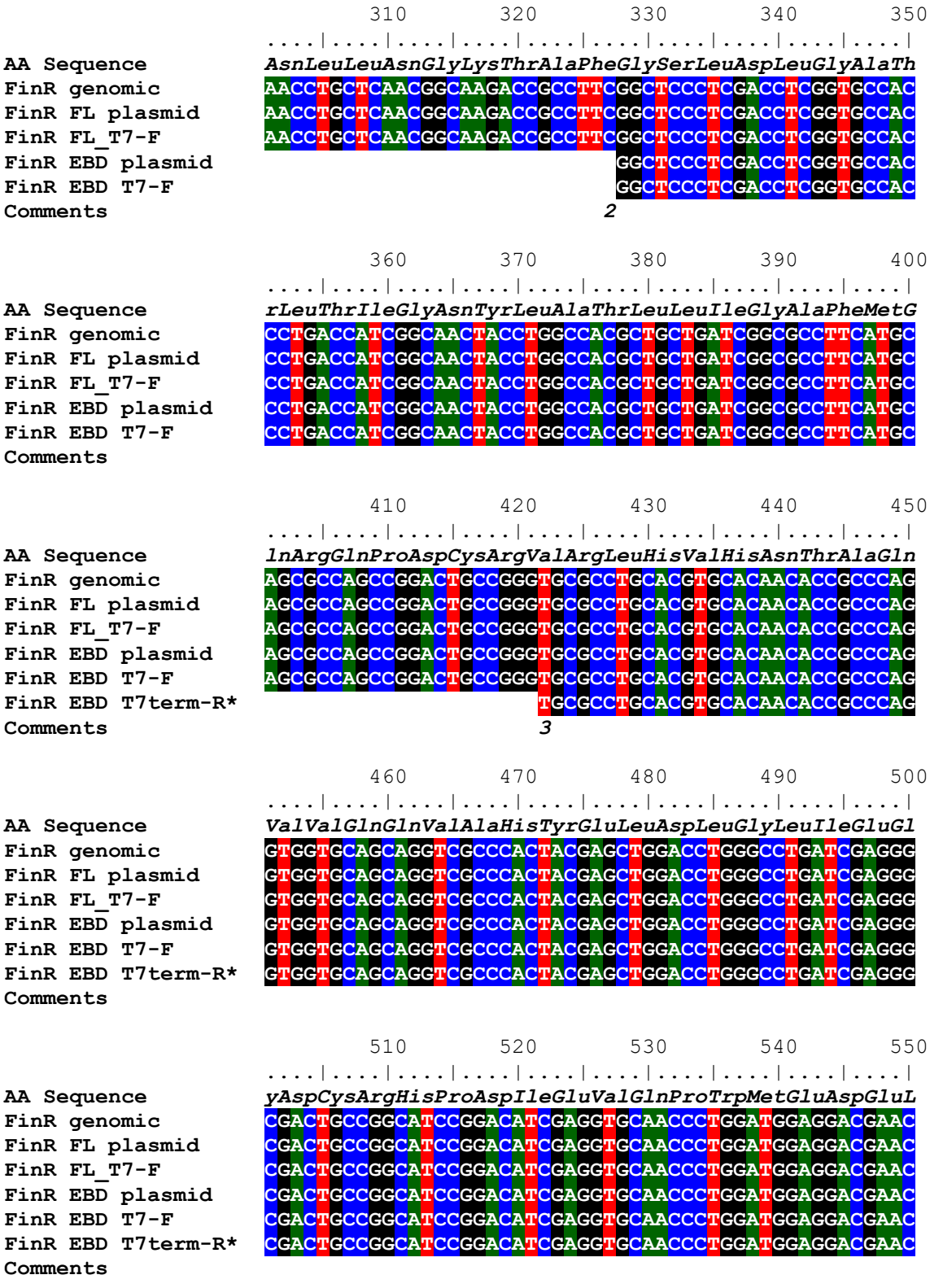
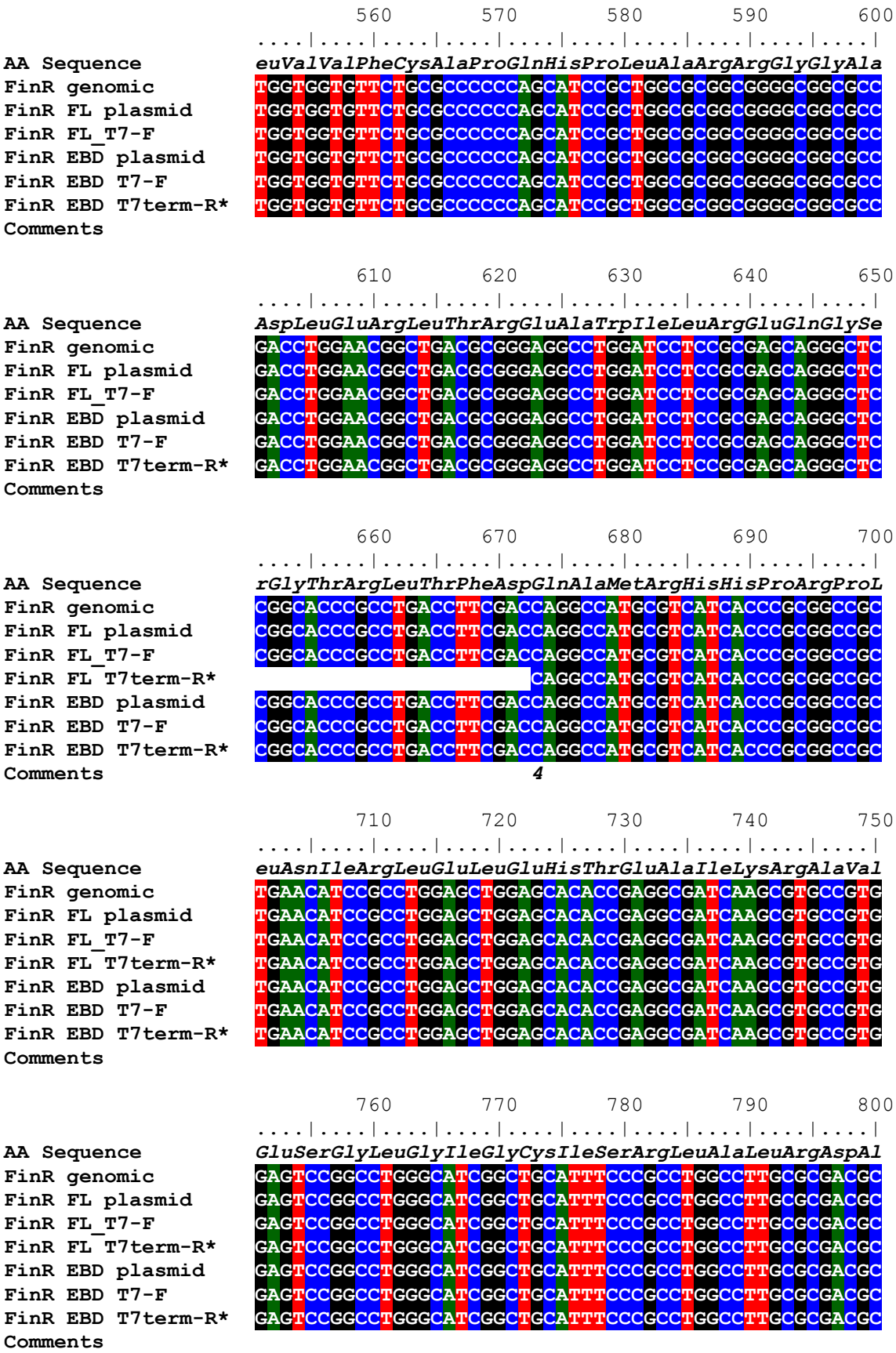


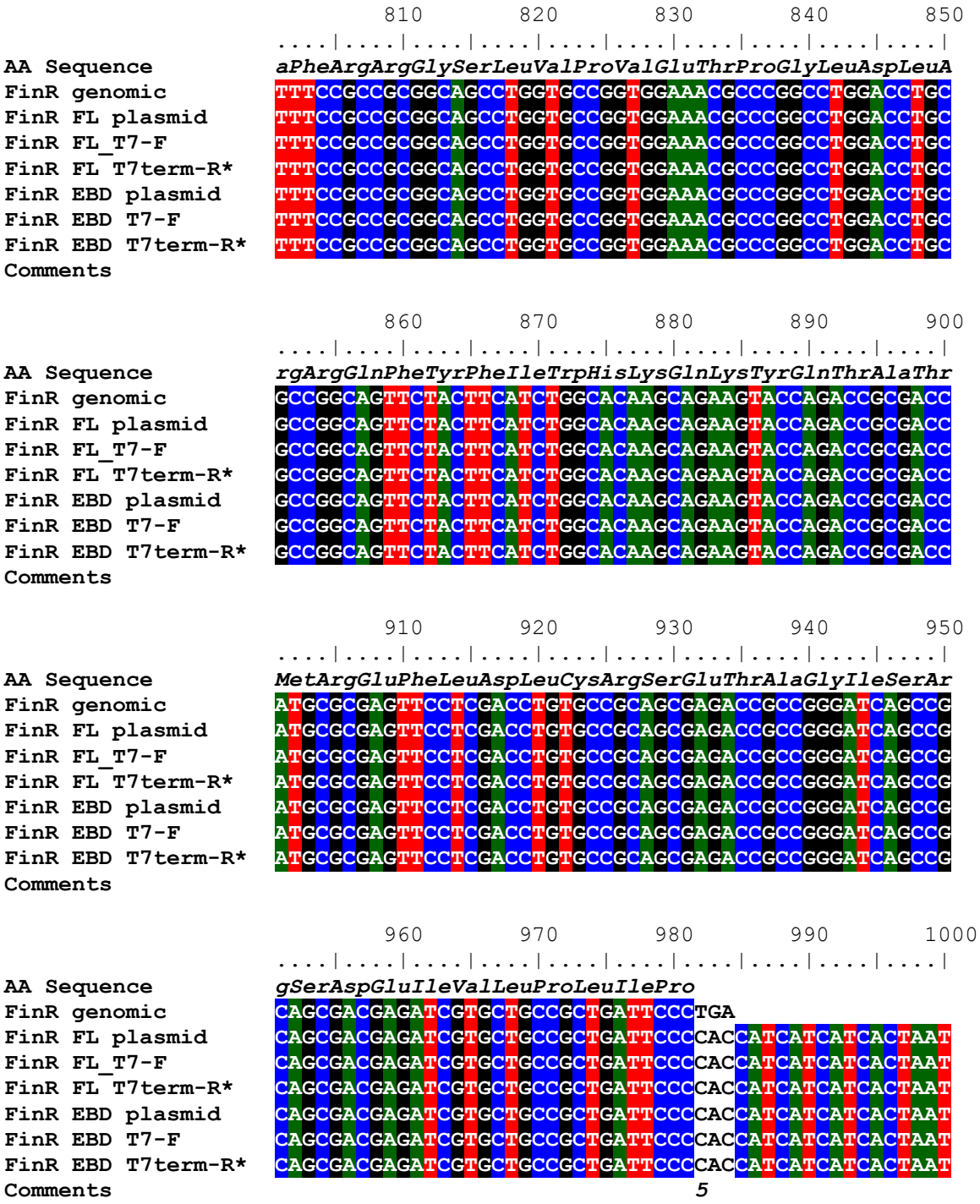
Figure A3-3. Saturation curves for FinR-MX(1-4) Kd determinations.

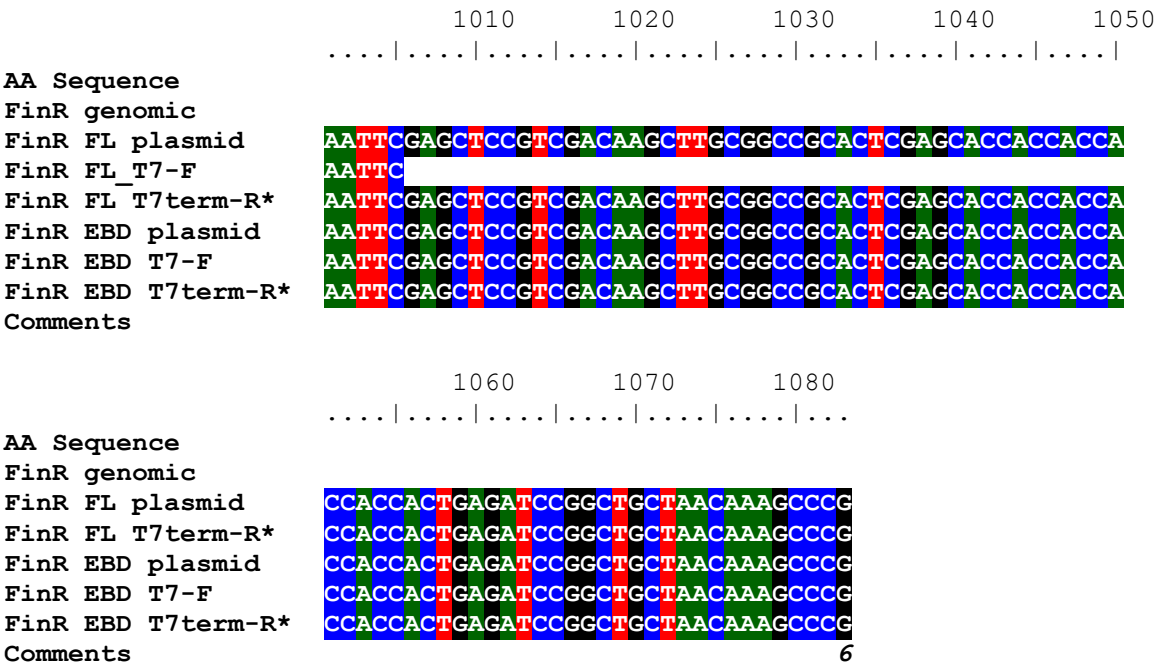
Sequencing analysis for FinR expression plasmids




310 320 330 340 350
AA Sequence AsnLeuLeuAsnGlyLysThrAlaPheGlySerLeuAspLeuGlyAlaTh
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR EBD plasmid
FinR EBD T7-F
Comments 2
360 370 380 390 400
AA Sequence rLeuThrIleGlyAsnTyrLeuAlaThrLeuLeuIleGlyAlaPheMetG
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR EBD plasmid
FinR EBD T7-F
Comments
410 420 430 440 450
AA Sequence lnArgGlnProAspCysArgValArgLeuHisValHisAsnThrAlaGln
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments 3
460 470 480 490 500
AA Sequence ValValGlnGlnValAlaHisTyrGluLeuAspLeuGlyLeuIleGluGl
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments
510 520 530 540 550
AA Sequence yAspCysArgHisProAspIleGluValGlnProTrpMetGluAspGluL
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments


560 570 580 590 600
AA Sequence euValValPheCysAlaProGlnHisProLeuAlaArgArgGlyGlyAla
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments
610 620 630 640 650
AA Sequence AspLeuGluArgLeuThrArgGluAlaTrpIleLeuArgGluGlnGlySe
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments
660 670 680 690 700
AA Sequence rGlyThrArgLeuThrPheAspGlnAlaMetArgHisHisProArgProL
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR FL T7term-R*
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments
4
710 720 730 740 750
AA Sequence euAsnIleArgLeuGluLeuGluHisThrGluAlaIleLysArgAlaVal
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR FL T7term-R*
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments
760 770 780 790 800
AA Sequence GluSerGlyLeuGlyIleGlyCysIleSerArgLeuAlaLeuArgAspAl
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR FL T7term-R*
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments


810 820 830 840 850
AA Sequence aPheArgArgGlySerLeuValProValGluThrProGlyLeuAspLeuA
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR FL T7term-R*
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments
860 870 880 890 900
AA Sequence rgArgGlnPheTyrPheIleTrpHisLysGlnLysTyrGlnThrAlaThr
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR FL T7term-R*
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments
910 920 930 940 950
AA Sequence MetArgGluPheLeuAspLeuCysArgSerGluThrAlaGlyIleSerAr
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR FL T7term-R*
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments
960 970 980 990 1000
AA Sequence gSerAspGluIleValLeuProLeuIlePro
FinR genomic
FinR FL plasmid
FinR FL_T7-F
FinR FL T7term-R*
FinR EBD plasmid
FinR EBD T7-F
FinR EBD T7term-R*
Comments
5



Sequencing Corrections for FinR expression plasmids

Comment #	Sequencing primer	Comments
1	FinR-EBD T7-F	First 24 nucleotides deleted
2		Effector binding domain lacks region encoding amino acids 2-88
3	FinR-EBD T7term-R*	Sequence poor after read 680 (5' end). Deleted
4	FinR-FL T7term- R*	Sequence poor after read 428 (5' end). Deleted
5		TGA stop codon mutated to encode 5xHis tag with TAATAA double stop (C-terminal his tag)
5	FinR-EBD T7term R*	First 24 nucleotides deleted
6	FinR-EBD T7-F, FinR-EBD T7term R*, FinR-FL T7term- R*	Sequences trimmed to this position. Sequence is consistent with plasmid beyond this.

Appendix C (Chapter 4)

Table A4-1. Low sulfate defined medium for auto-induction

	1000 ml total	500 ml total	Final concentration
Sterile water	583 ml	450 ml	-
1 M $MgCl_2$	1.0 ml	0.5 ml	1 mM
1000x metals mix (low sulfate)	1.0 ml	0.5 ml	1x
50x 5052	20 ml	10 ml	1x
20x KPMN	50 ml	25 ml	1x
L-cysteine (250 mM, 30.29mg/ml)	2 ml	1 mM	0.5 mM
L-leucine (125 mM, 10X)	100 ml	50 ml	12.5 mM
L-Isoleucine (125 mM, 10X)	100 ml	50 ml	12.5 mM
L-Valine (250 mM, 10X)	100 ml	50 ml	25 mM
Methionine (25mg/ml)	8 ml	4 ml	200 µg/ml
17aa (CYM; each 10mg/ml)	20 ml	10 ml	200 µg/ml each
Antibiotics:			
Kanamycin Cl (50 mg/ml)	2 ml	1 ml	100 µg/ml
Chloramphenicol (100mg/ml)	0.34 ml	0.5 ml	34 µg/ml
Ampicillin (50mg/ml)	1.0 ml	0.5 ml	50 µg/ml
Strep/Spec (100mg/ml)	0.75 ml		75 µg/ml
Thiamine HCl (300 mM)	1 ml	0.5 ml	3 mM
Vitamins mix	2 ml	1 ml	-

For all media, add 1 M $MgCl_2$ and 1000x metals mix before adding 20x NPS to avoid precipitate

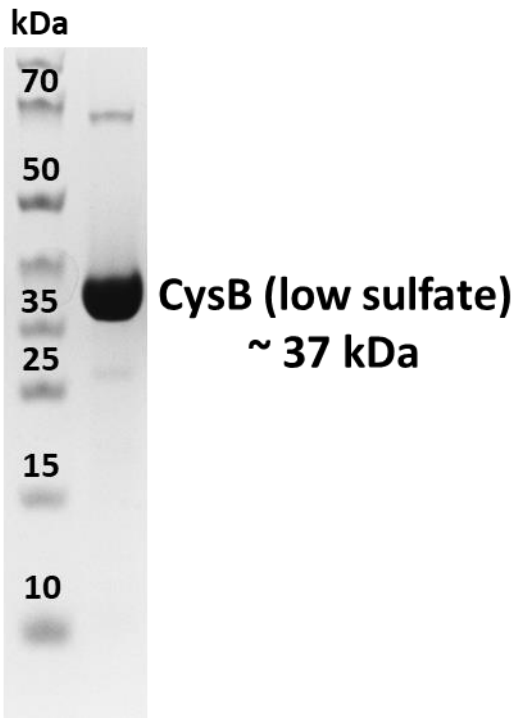
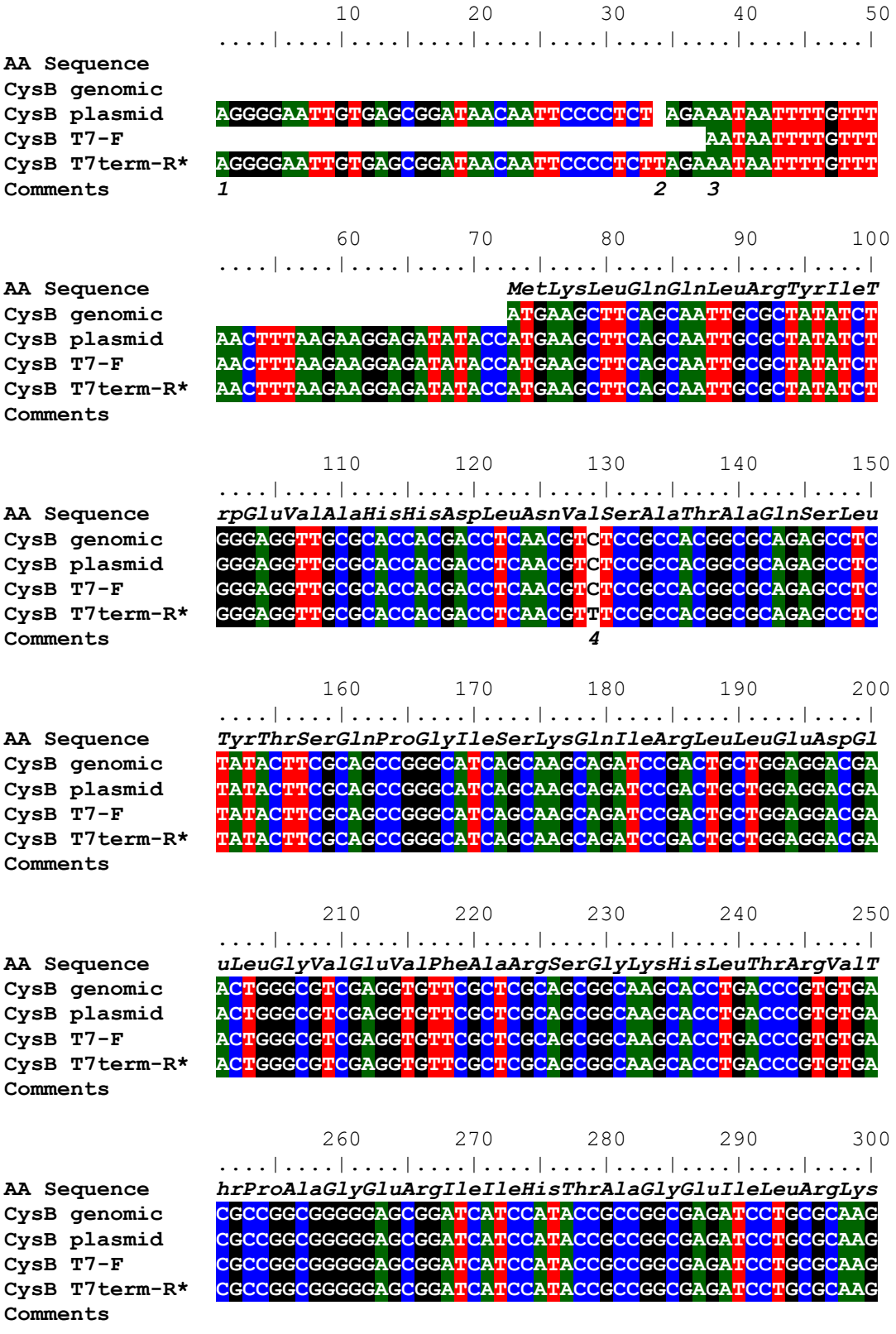
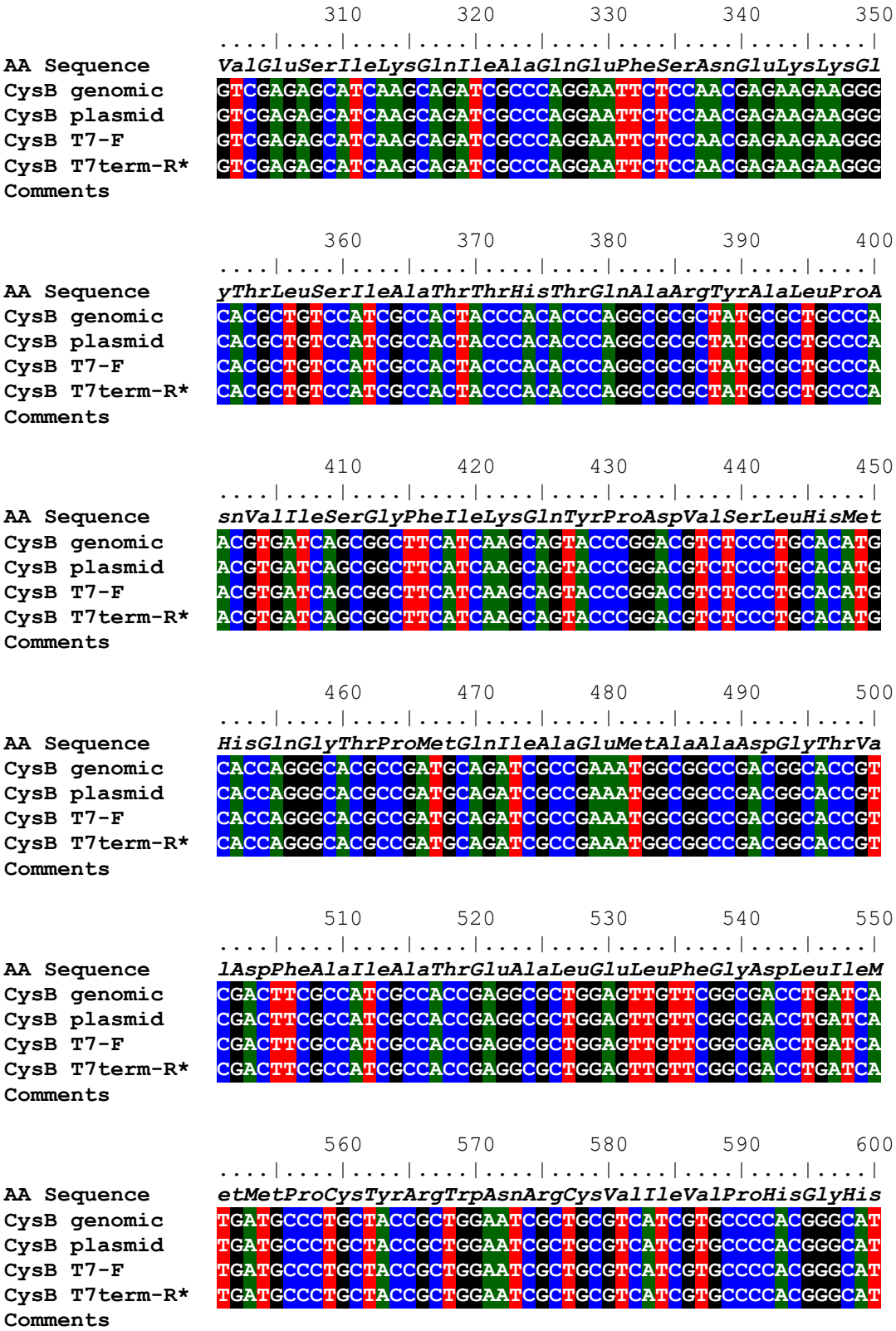
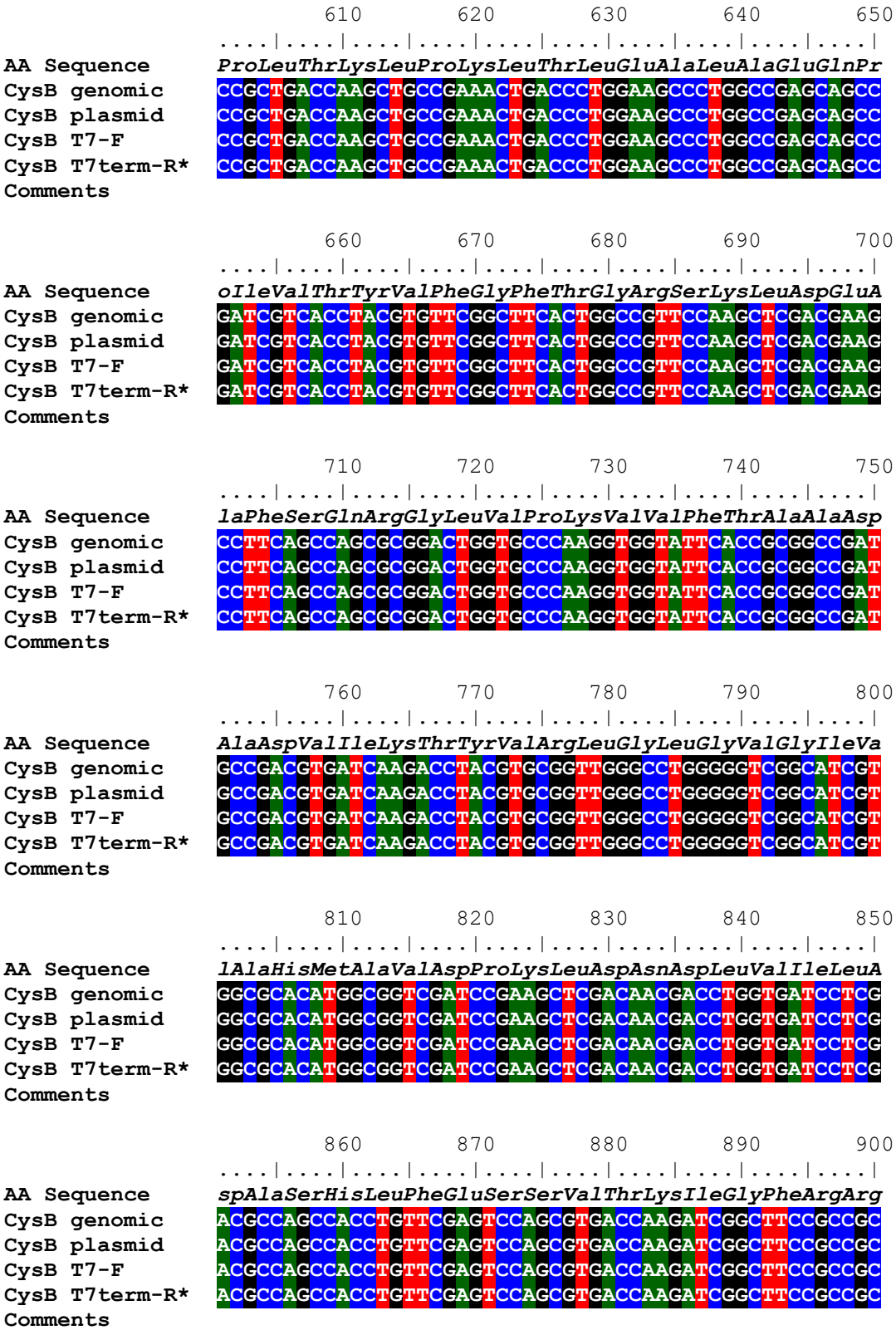


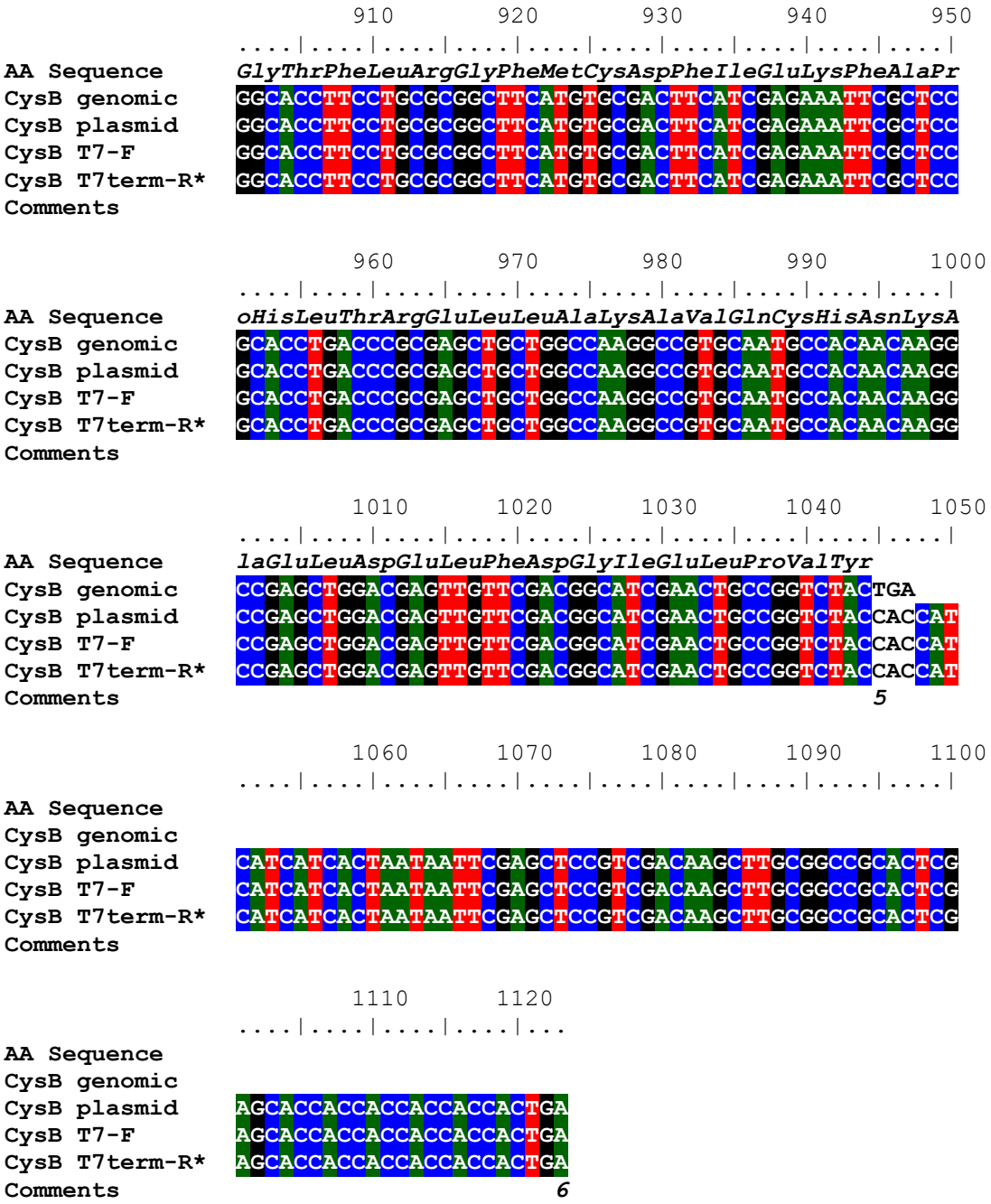
Figure A4-1. SDS Page analysis of purified *P. aeruginosa* CysB (low sulfate).

Sequencing analysis for CysB expression plasmid




310 320 330 340 350
AA Sequence ValGluSerIleLysGlnIleAlaGlnGluPheSerAsnGluLysLysGl
CysB genomic GTCGAGAGCATCAAGCAGATCGCCCAGGAATTCTCCAACGAGAAGAAGGG
CysB plasmid GTCGAGAGCATCAAGCAGATCGCCCAGGAATTCTCCAACGAGAAGAAGGG
CysB T7-F GTCGAGAGCATCAAGCAGATCGCCCAGGAATTCTCCAACGAGAAGAAGGG
CysB T7term-R* GTCGAGAGCATCAAGCAGATCGCCCAGGAATTCTCCAACGAGAAGAAGGG
Comments
360 370 380 390 400
AA Sequence yThrLeuSerIleAlaThrThrHisThrGlnAlaArgTyrAlaLeuProA
CysB genomic CACGCTGTCCATCGCCACTACCCACACCCAGGCGCGCTATGCGCTGCCCA
CysB plasmid CACGCTGTCCATCGCCACTACCCACACCCAGGCGCGCTATGCGCTGCCCA
CysB T7-F CACGCTGTCCATCGCCACTACCCACACCCAGGCGCGCTATGCGCTGCCCA
CysB T7term-R* CACGCTGTCCATCGCCACTACCCACACCCAGGCGCGCTATGCGCTGCCCA
Comments
410 420 430 440 450
AA Sequence snValIleSerGlyPheIleLysGlnTyrProAspValSerLeuHisMet
CysB genomic ACGTGATCAGCGGCTTCATCAAGCAGTACCCGGACGTCTCCCTGCACATG
CysB plasmid ACGTGATCAGCGGCTTCATCAAGCAGTACCCGGACGTCTCCCTGCACATG
CysB T7-F ACGTGATCAGCGGCTTCATCAAGCAGTACCCGGACGTCTCCCTGCACATG
CysB T7term-R* ACGTGATCAGCGGCTTCATCAAGCAGTACCCGGACGTCTCCCTGCACATG
Comments
460 470 480 490 500
AA Sequence HisGlnGlyThrProMetGlnIleAlaGluMetAlaAlaAspGlyThrVa
CysB genomic CACCAGGGCACGCCGATGCAGATCGCCGAAATGGCGGCCGACGGCACCGT
CysB plasmid CACCAGGGCACGCCGATGCAGATCGCCGAAATGGCGGCCGACGGCACCGT
CysB T7-F CACCAGGGCACGCCGATGCAGATCGCCGAAATGGCGGCCGACGGCACCGT
CysB T7term-R* CACCAGGGCACGCCGATGCAGATCGCCGAAATGGCGGCCGACGGCACCGT
Comments
510 520 530 540 550
AA Sequence lAspPheAlaIleAlaThrGluAlaLeuGluLeuPheGlyAspLeuIleM
CysB genomic CGACTTCGCCATCGCCACCGAGGCGCTGGAGTTGTTCGGCGACCTGATCA
CysB plasmid CGACTTCGCCATCGCCACCGAGGCGCTGGAGTTGTTCGGCGACCTGATCA
CysB T7-F CGACTTCGCCATCGCCACCGAGGCGCTGGAGTTGTTCGGCGACCTGATCA
CysB T7term-R* CGACTTCGCCATCGCCACCGAGGCGCTGGAGTTGTTCGGCGACCTGATCA
Comments
560 570 580 590 600
AA Sequence etMetProCysTyrArgTrpAsnArgCysValIleValProHisGlyHis
CysB genomic TGATGCCCTGCTACCGCTGGAATCGCTGCGTCATCGTGCCCCACGGGCAT
CysB plasmid TGATGCCCTGCTACCGCTGGAATCGCTGCGTCATCGTGCCCCACGGGCAT
CysB T7-F TGATGCCCTGCTACCGCTGGAATCGCTGCGTCATCGTGCCCCACGGGCAT
CysB T7term-R* TGATGCCCTGCTACCGCTGGAATCGCTGCGTCATCGTGCCCCACGGGCAT
Comments


610 620 630 640 650
AA Sequence ProLeuThrLysLeuProLysLeuThrLeuGluAlaLeuAlaGluGlnPr
CysB genomic CCGCTGACCAAGCTGCCGAAACTGACCCTGGAAGCCCTGGCCGAGCAGCC
CysB plasmid CCGCTGACCAAGCTGCCGAAACTGACCCTGGAAGCCCTGGCCGAGCAGCC
CysB T7-F CCGCTGACCAAGCTGCCGAAACTGACCCTGGAAGCCCTGGCCGAGCAGCC
CysB T7term-R* CCGCTGACCAAGCTGCCGAAACTGACCCTGGAAGCCCTGGCCGAGCAGCC
Comments
660 670 680 690 700
AA Sequence oIleValThrTyrValPheGlyPheThrGlyArgSerLysLeuAspGluA
CysB genomic GATCGTCACCTACGTGTTCGGCTTCACTGGCCGTTCCAAGCTCGACGAAG
CysB plasmid GATCGTCACCTACGTGTTCGGCTTCACTGGCCGTTCCAAGCTCGACGAAG
CysB T7-F GATCGTCACCTACGTGTTCGGCTTCACTGGCCGTTCCAAGCTCGACGAAG
CysB T7term-R* GATCGTCACCTACGTGTTCGGCTTCACTGGCCGTTCCAAGCTCGACGAAG
Comments
710 720 730 740 750
AA Sequence laPheSerGlnArgGlyLeuValProLysValValPheThrAlaAlaAsp
CysB genomic CCTTCAGCCAGCGCGGACTGGTGCCCAAGGTGGTATTCACCGCGGCCGAT
CysB plasmid CCTTCAGCCAGCGCGGACTGGTGCCCAAGGTGGTATTCACCGCGGCCGAT
CysB T7-F CCTTCAGCCAGCGCGGACTGGTGCCCAAGGTGGTATTCACCGCGGCCGAT
CysB T7term-R* CCTTCAGCCAGCGCGGACTGGTGCCCAAGGTGGTATTCACCGCGGCCGAT
Comments
760 770 780 790 800
AA Sequence AlaAspValIleLysThrTyrValArgLeuGlyLeuGlyValGlyIleVa
CysB genomic GCCGACGTGATCAAGACCTACGTGCGGTTGGGCCTGGGGGTCGGCATCGT
CysB plasmid GCCGACGTGATCAAGACCTACGTGCGGTTGGGCCTGGGGGTCGGCATCGT
CysB T7-F GCCGACGTGATCAAGACCTACGTGCGGTTGGGCCTGGGGGTCGGCATCGT
CysB T7term-R* GCCGACGTGATCAAGACCTACGTGCGGTTGGGCCTGGGGGTCGGCATCGT
Comments
810 820 830 840 850
AA Sequence lAlaHisMetAlaValAspProLysLeuAspAsnAspLeuValIleLeuA
CysB genomic GGCGCACATGGCGGTCGATCCGAAGCTCGACAACGACCTGGTGATCCTCG
CysB plasmid GGCGCACATGGCGGTCGATCCGAAGCTCGACAACGACCTGGTGATCCTCG
CysB T7-F GGCGCACATGGCGGTCGATCCGAAGCTCGACAACGACCTGGTGATCCTCG
CysB T7term-R* GGCGCACATGGCGGTCGATCCGAAGCTCGACAACGACCTGGTGATCCTCG
Comments
860 870 880 890 900
AA Sequence spAlaSerHisLeuPheGluSerSerValThrLysIleGlyPheArgArg
CysB genomic ACGCCAGCCACCTGTTCGAGTCCAGCGTGACCAAGATCGGCTTCCGCCGC
CysB plasmid ACGCCAGCCACCTGTTCGAGTCCAGCGTGACCAAGATCGGCTTCCGCCGC
CysB T7-F ACGCCAGCCACCTGTTCGAGTCCAGCGTGACCAAGATCGGCTTCCGCCGC
CysB T7term-R* ACGCCAGCCACCTGTTCGAGTCCAGCGTGACCAAGATCGGCTTCCGCCGC
Comments


910 920 930 940 950
AA Sequence GlyThrPheLeuArgGlyPheMetCysAspPheIleGluLysPheAlaPr
CysB genomic GGCACCTTCCTGCGCGGCTTCATGTGCGACTTCATCGAGAAATTCGCTCC
CysB plasmid GGCACCTTCCTGCGCGGCTTCATGTGCGACTTCATCGAGAAATTCGCTCC
CysB T7-F GGCACCTTCCTGCGCGGCTTCATGTGCGACTTCATCGAGAAATTCGCTCC
CysB T7term-R* GGCACCTTCCTGCGCGGCTTCATGTGCGACTTCATCGAGAAATTCGCTCC
Comments
960 970 980 990 1000
AA Sequence oHisLeuThrArgGluLeuLeuAlaLysAlaValGlnCysHisAsnLysA
CysB genomic GCACCTGACCCGCGAGCTGCTGGCCAAGGCCGTGCAATGCCACAACAAGG
CysB plasmid GCACCTGACCCGCGAGCTGCTGGCCAAGGCCGTGCAATGCCACAACAAGG
CysB T7-F GCACCTGACCCGCGAGCTGCTGGCCAAGGCCGTGCAATGCCACAACAAGG
CysB T7term-R* GCACCTGACCCGCGAGCTGCTGGCCAAGGCCGTGCAATGCCACAACAAGG
Comments
1010 1020 1030 1040 1050
AA Sequence laGluLeuAspGluLeuPheAspGlyIleGluLeuProValTyr
CysB genomic CCGAGCTGGACGAGTTGTTCGACGGCATCGAACTGCCGGTCTACTGA
CysB plasmid CCGAGCTGGACGAGTTGTTCGACGGCATCGAACTGCCGGTCTACCACCAT
CysB T7-F CCGAGCTGGACGAGTTGTTCGACGGCATCGAACTGCCGGTCTACCACCAT
CysB T7term-R* CCGAGCTGGACGAGTTGTTCGACGGCATCGAACTGCCGGTCTACCACCAT
Comments 5
1060 1070 1080 1090 1100
AA Sequence
CysB genomic
CysB plasmid CATCATCACTAATAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCG
CysB T7-F CATCATCACTAATAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCG
CysB T7term-R* CATCATCACTAATAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCG
Comments
1110 1120
AA Sequence
CysB genomic
CysB plasmid AGCACCACCACCACCACCACTGA
CysB T7-F AGCACCACCACCACCACCACTGA
CysB T7term-R* AGCACCACCACCACCACCACTGA
Comments 6

Sequencing Corrections for CysB plasmid

Comment #	Sequencing primer	Comments
1	CysB T7term-R*	Deleted 3' after nucleotide 1162 (sequence covers lac operator)
2	CysB T7term-R*	AA (TT*) doublet peak is not clear
3	CysB T7-F	Deleted first 22
3	CysB-EBD T7term-R*	Sequence poor after read 680 (5' end). Deleted
4	CysB-FL T7term- R*; CysB T7-F	A (T*) triplet peaks in -R* is clear; TCT peaks in -F are clear. GTC/GTT codons are both Val, so not a problem.
5		TGA stop codon mutated to encode 5xHis tag with TAATAA double stop (C-terminal his tag)
6	CysB-EBD T7-F, CysB T7term R*	Sequences trimmed to this position. Sequence is consistent with plasmid beyond this.

Appendix D (Chapter 5)

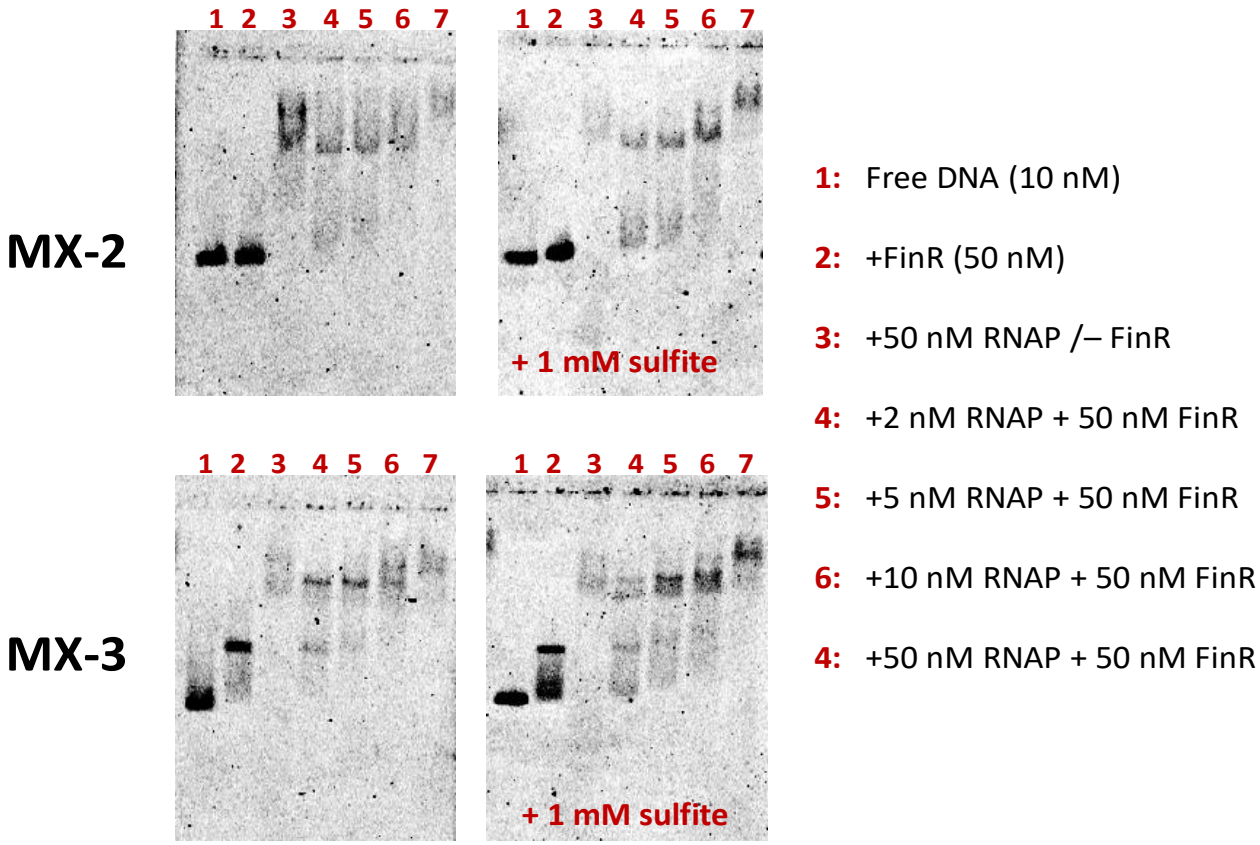


Figure A5-1. Complex formation between FinR-RNAP and MX-2/MX-3

Appendix E (Preliminary studies on *P. aeruginosa* alternative sigma factors)

Cloning, expression, purification of *P. aeruginosa* alternative sigma factors

Table E-1. Bacterial strains, plasmids, and primers used for cloning and expression of alternative sigma factors

Name	Description	Source
***E. coli* Strains**		
XL1-Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F′ proAB lacIq ZΔM15 Tn10 (Tetr)]	Agilent
BL21-CodonPlus(DE3)-RIPL	*E. coli* B F– ompT hsdS(rB – mB –) dcm+ Tetr gal λ(DE3) endA Hte [argU proL Camr] [argU ileY leuW Strep/Specr]	Agilent
Plasmids		
pET28b(+)	T7lac promoter, Kan^R , CloDF13 ori, 6x-His	Novagen
pET28b(+).SapKO−CH.BspQI	pET28b(+) modified with BspQI cloning site	Lab plasmid
pDAP24	pDAP24.pET28b.PAO1.rpoS.CHx5	This work
pDAP25	pDAP25.pET28b.PAO1.algU.CHx5	This work
pDAP26	pDAP26.pET28b.PAO1.pvdS.CHx5	This work
pDAP27	pDAP26.pET28b.PAO1.fliA.CHx5	This work
pDAP28	pDAP26.pET28b.PAO1.sigX.CHx5	This work
pDAP29	pDAP26.pET28b.PAO1.rpoN.CHx5	This work
pDAP30	pDAP26.pET28b.PAO1.fvpI.CHx5	This work
Primers		
PAO1 RpoS-F	GGGCTCTTCAATGGCACTCAAAAA AGAAGGG	IDT DNA
PAO1 RpoS-R	GGGCTCTTCAAGTGACCTTGTCGCGCAG	IDT DNA

PAO1 AlgU-F	GGGCTCTTCAGGGATGCTAACCCAG GAACAGGATCAACAAC	IDT DNA
PAO1 AlgU-R	GGGCTCTTCATTAGGCTTCTCGCAA CAAAGGCTGCAGAGCTTCG	IDT DNA
PAO1 PvdS-F	GGGCTC TTCCGGGATGTCGGAAC AACTGTCTACCCGCAGATGC	IDT DNA
PAO1 PvdS-R	GGGCTCTTCATTATCGGCGGGCG CTGAGATG	IDT DNA
PAO1 FliA-F	GGGCTCTTCAGGGATGACAGCGGC CTCTGGAGTGCGTATGTATAGC	IDT DNA
PAO1 FliA-R	GGGCTCTTCATTAGGCGGACCGCCA ATCGGCCAGGC	IDT DNA
PAO1 SigX-F	GGGCTCTTCCGGGATGACGCGTGCC TATGAAGAATTGATGCGGCG	IDT DNA
PAO1 SigX-R	GGGCTCTTCATTATGTCTCGGTGGCAT CTGAAAACTTTTCGCGC	IDT DNA
PAO1 RpoN-F	GGGCTCTTCAGGGATGAAACCGT CGCTAGTCCTCAAGATGGGCCAGC	IDT DNA
PAO1 RpoN-R	GGGCTCTTCATTACACCAGTCG CTTGCGCTCGCTCG	IDT DNA
PAO1 fvpI-F	GGGCTCTTCAGGGATGGAAAACCAT TACCGGGAGCTGCTGCGATTCC	IDT DNA
PAO1 fvpI-R	GGGCTCTTCATTAGTCGGCTTCCC ATTCGCGCATGC	IDT DNA
RsmAprom-F	GCCGGCGCGACAGGGTGAGTG ACGCTGAC	IDT DNA
RsmAprom-R	cGCATGATACCCATCTTTAC CCCGTTTGCAAAGGG	IDT DNA

PCR primers containing terminal BspQI restriction sites for seven (7) of the currently characterized *P. aeruginosa* alternate sigma factors (RpoS, AlgU, PvdS, FvpI, FliA, RpoN, SigX) were used to generate the DNA fragments used for restriction cloning. The RpoD expression construct (pDAP15), generated in Chapter 2, was also used in parts of this study. The restriction cloning method followed the approach described by [19] and utilized the BspQI restriction sites in the pET28b(+).SapKO−CH.BspQI plasmid. The generated expression plasmids (**Table E-1**) each contained a 5x His tag and a TEV protease cleavage site. Each plasmid was sequenced (ACGT Inc) to confirm a successful cloning of the correct gene fragments. The approach for protein expression and purification for each protein followed the multistep approached described for expressing and purifying RNA polymerase in Chapter 2. SDS PAGE analysis was done to assess the quality of the purification process for each protein, and the protein samples were stored in in the –20°C freezer in a buffer containing 20 mM tris (pH 7.5), 100 mM NaCl, 0.1 mM EDTA, 10 mM 2-mercaptoethanol, and 50% glycerol. RpoD (σ^{70}) was cloned, expressed, and purified as described in Chapter 2.

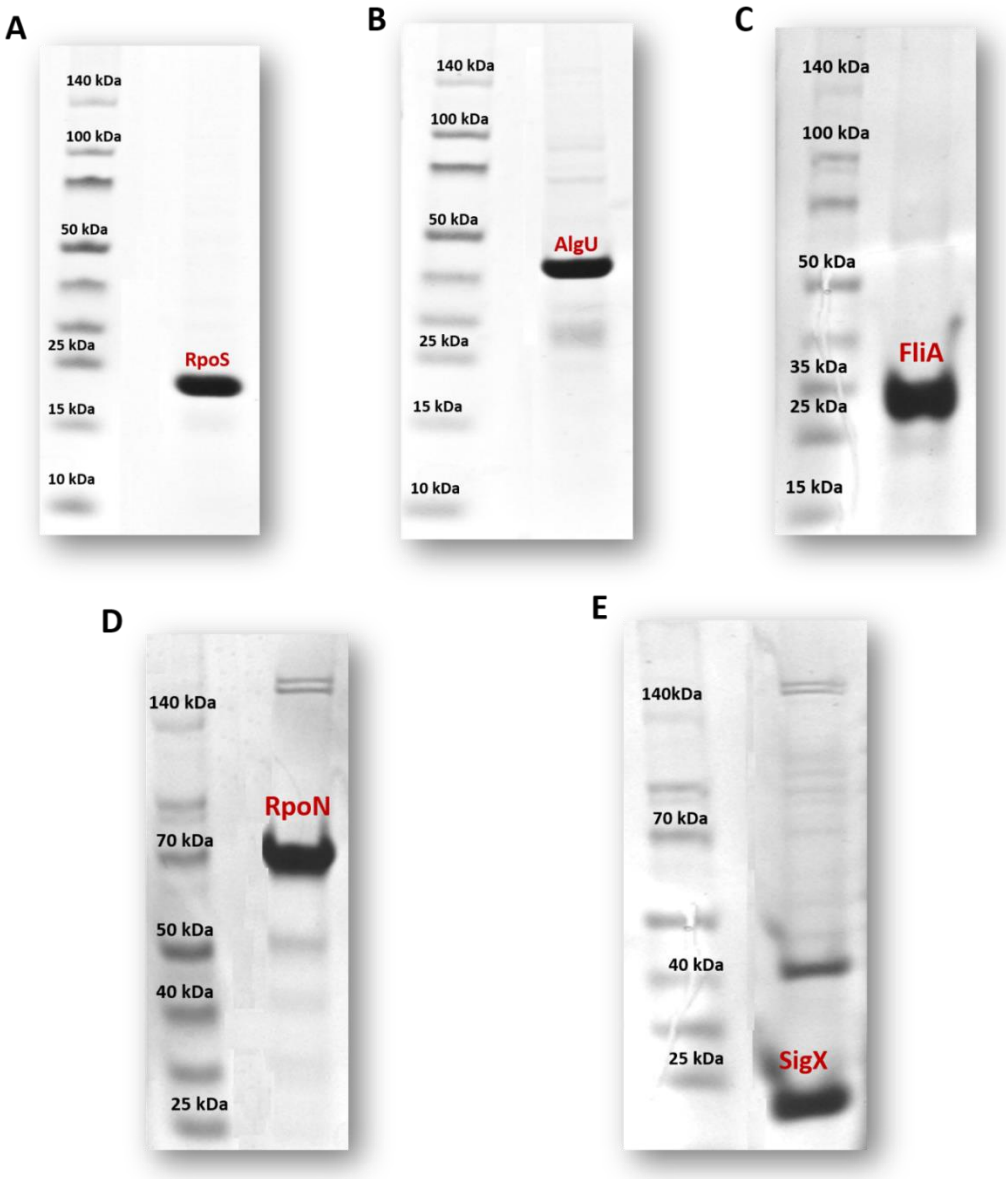


Figure E-1. Purification of alternate sigma factors in *P. aeruginosa*. **(A-E)** SDS PAGE gel analysis of the five proteins, RpoS, AlgU, FliA, RpoN, and SigX, that were soluble and well behaved. Some of the sigma factors like SigX, RpoN, and AlgU were associated with co-elution with some *E. coli* RNAP subunits while others like RpoS and FliA were not able to co-elute with any *E. coli* RNAP subunits.

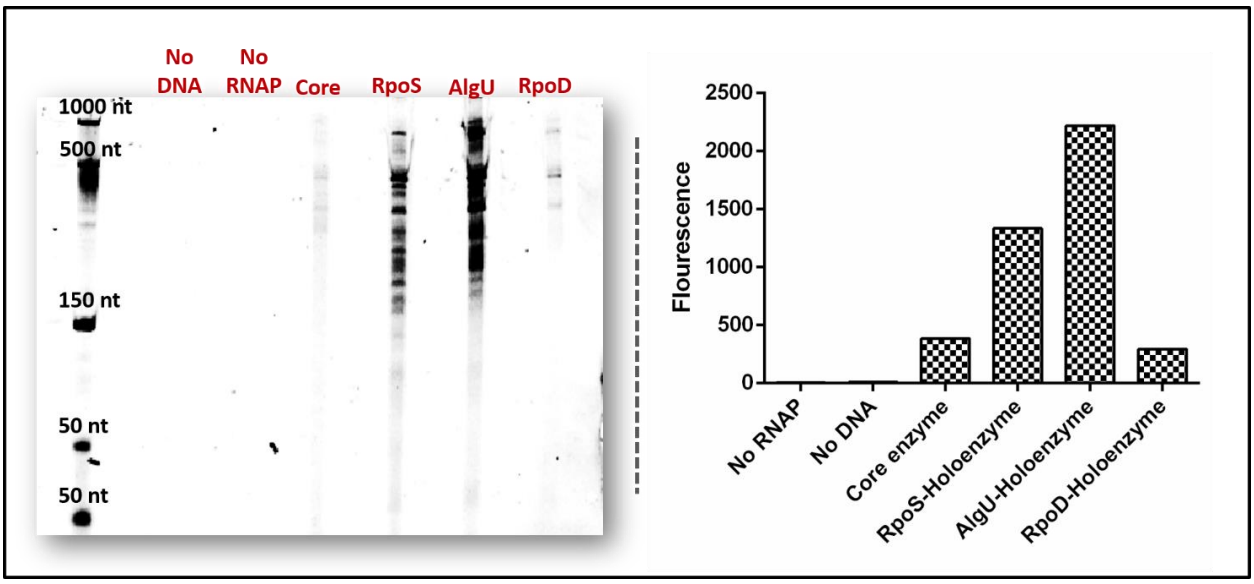


Figure E-2. Selective transcription of RNAP based on promoter strength for different sigma factors. A DNA fragment containing the consensus sequences for RpoS, AlgU, and RpoD such that each promoter had a different strength for expression of the target gene, *rsmA*, was used. **Left**: PAGE gel analysis of synthesized RNA transcripts. **Right**: Quantification of RNA transcripts by measuring Ribogreen fluorescence. Data is based on a single experiment.

Appendix F

Plasmid constructs created by Derrick Afful in the course of this dissertation research

pDA #	Base Plasmid	Organism	Gene Insert(s)	CH/NH	# of His	TEV	Full Name
pDAP1*	pET28b	PAO1	*oxyR*	CH	5x	No	pDAP1.pET28b.PAO1.oxyR.CHx5
pDAP2	pET28b	PAO1	*mvfR*	CH	5x	No	pDAP2.pET28b.PAO1.mvfR.CHx5
pDAP3	pET28b	PAO1	*gblock, rpoZ*	-	-	-	pDAP3.pET28b.PAO1.rpoZ-gblock
pDAP4	pET28b.rpoZ.gblock	PAO1	*rpoA*	-	-	-	pDAP4.pET28b.PAO1.rpoA-rpoZ
pDAP5	pRARE2 CH1	PAO1	*rpoB,rpoC*	No	-	No	pDAP5.pRARE2.PAO1.rpoB-rpoC.CH1
pDAP6	pRARE2 CH2	PAO1	*rpoB, rpoC*	No	-	No	pDAP6.pRARE2.PAO1.rpoB-rpoC.CH2
pDAP7	pET28b.rpoAZ	PAO1	*rpoD*	NH	5x	Yes	pDAP7.pET28b.PAO1.rpoA-rpoZ-rpoD.NHx5.TEV
pDAP8	pET28b	A.bau	*bfmR*	CH	5x	No	pDAP8.pET28b.ABAU.bfmR.CHx5
pDAP9	pET28b	PAO1	*bexR*	CH	5x	No	pDAP9.pET28b.PAO1.bexR.CHx5
pDAP10	pET28b	PAO1	*rpoZ, rpoA*	No	-	No	pDAP10.pET28b.PAO1.rpoA-rpoZ.ver2
pDAP11	pCDFDuet	-	-	-	-	-	pDAP11.pCDFDuet.PmlI.NHx6
pDAP12	pRARE2	PAO1	*rpoB, rpoC*	NH	10x	Yes	pDAP12.pRARE2.PAO1.RpoB-RpoC.ver2.NHx10.TEV
pDAP13	pET28b	PAO1	*bvlR*	CH	5x	No	pDAP13.pET28b.PAO1.bvlR.CHx5
pDAP14	pET28b	PAO1	*mexT-EBD*	CH	5x	No	pDAP14.pET28b.PAO1.mexT-EBD.CHx5
pDAP15	pCDFDuet	PAO1	*rpoD*	NH	10x	Yes	pDAP15.pCDFDuet.PAO1.rpoD.NHx10.TEV
pDAP16	pET28b	PAO1	*finR*	CH	5x	No	pDAP16.pET28b.PAO1.finR.CHx5
pDAP17*	pET28b	PAO1	*finR-EBD*	CH	5x	No	pDAP17.pET28b.PAO1.finR-EBD.CHx5

pDAP18	pRARE2	ADP1	*rpoB, rpoC*	NH	10x	Yes	pDAP18.pRARE2.ADP1.rpoB-rpoC.NHx10.TEV
pDAP19	pET28b	ADP1	*rpoA, rpoZ, rpoD*	NH	5x	Yes	pDAP19.pET28b.ADP1.rpoA-rpoZ-rpoD.NHx5.TEV
pDAP20	pUC18	PAO1	*finR-fpR*	-	-	-	pDAP20.pUC18.PAO1.finR-fpR
pDAP21	pET28b	PAO1	*cysB*	CH	5x	No	pDAP21.pET28b.PAO1.cysB.CHx5
pDAP22	pCDFDuet	PAO1	*rpoA, rpoZ, rpoD*	NH	10x	Yes	pDAP22.pCDFDuet.PAO1.rpoA-rpoZ-rpoD.NHx10.TEV
pDAP23	pUC18	PAO1	*rpoA, rpoZ*	-	-	-	pDAP23.pUC18.PAO1.rpoA-rpoZ
pDAP24	pET28b	PAO1	*rpoS (sigma fac.)*	NH	5x	Yes	pDAP24.pET28b.PAO1.rpoS.CHx5
pDAP25	pET28b	PAO1	*algU (sigma fac.)*	NH	5x	Yes	pDAP25.pET28b.PAO1.algU.CHx5
pDAP26	pET28b	PAO1	*pvdS (sigma fac.)*	NH	5x	Yes	pDAP26.pET28b.PAO1.pvdS.CHx5
pDAP27	pET28b	PAO1	*fliA*	NH	5x	Yes	pDAP27.pET28b.PAO1.fliA.CHx5
pDAP28	pET28b	PAO1	*sigX*	NH	5x	Yes	pDAP28.pET28b.PAO1.sigX.CHx5
pDAP29	pET28b	PAO1	*rpoN*	NH	5x	Yes	pDAP29.pET28b.PAO1.rpoN.CHx5
pDAP30	pET28b	PAO1	*fvpI*	NH	5x	Yes	pDAP30.pET28b.PAO1.fvpI.CHx5
pDAP31	pCDFDuet	PAO1	*hrpL*	NH	10x	Yes	pDAP31.pCDFDuet.PAO1.rpoAZ.DC3000.hrpL.NHx10.TEV
pDAP32	pET28b	PAO1	*hrpL*	NH	5x	Yes	pDAP32.pCDFDuet.DC3000.hrpL.NHx5.TEV
pDAP33	pUC18	PAO1	*algD_prom*	-	-	-	pDAP33.pUC18.PAO1.algDprom
pDAP34	pUC18	PAO1	*pvdS_prom*	-	-	-	pDAP34.pUC18.PAO1.pvdSprom
pDAP35	pUC18	PAO1	*mvfR_prom*	-	-	-	pDAP35.pUC18.PAO1.mvfRprom
pDAP36	pUC18	PAO1	*cysB_prom*	-	-	-	pDAP36.pUC18.PAO1.cysBprom
pDAP37	pUC18	PAO1	*cysI_prom*	-	-	-	pDAP37.pUC18.PAO1.cysIprom
pDAP38	pUC18	PAO1	*atsR_prom*	-	-	-	pDAP38.pUC18.PAO1.atsRprom
pDAP39	pCDFDuet	ADP1	*rpoA, rpoZ, rpoD*	NH	10x	Yes	pDAP39.pCDFDuet.ADP1.rpoA-rpoZ-rpoD.NHx10.TEV

Legend: Yellow highlight − soluble/purifiable protein; Yellow highlight with * − crystallizable protein

Abbreviations used: PAO1 − *P. aeruginosa*; ADP1 − *A. baylyi*; A.bau − *A. baumannii*