Williams College Library

Your unpublished thesis, submitted for a degree at Williams College and deposited in the Williams College Library, will be made available for research use. You may, through this form, determine copyright and reproduction of your thesis. Please check the statements that reflect your wishes.

Publication and Quotation: Literary Property Rights

___ I do not choose to retain literary property rights to my thesis and wish to assign them immediately to Williams College.

___ I wish to retain literary property rights to my thesis for a period of___ years, or until my death, whichever is the later, at which time the literary property rights shall be assigned to Williams College.

___ For so long as I retain literary property rights to my thesis, readers may make brief quotations from it without my permission. Extensive quotations may not be made without my permission.

Copying and Dissemination

___ Copies of my thesis may be provided to any researcher.

___ Copies of my thesis may be deposited only at other institutions (such as public, college or university libraries).

___ Copying of my thesis, except as needed to maintain an adequate number of research copies available in the Williams College Library, is expressly prohibited, and copies may not be disseminated outside the Library.

___ Limitations (if any) to copying and dissemination shall be in effect until_____, at which time current Archives practice shall prevail.

Signed Signature removed. See Archives for hard copy.

Thesis title Purification of Bacillus Subtilis Comk Protein

Date 5-20-02

Accepted Signature removed. See Archives for hard copy.

Date 5-20-02

Purification of *Bacillus subtilis* ComK Protein

by

Carol Lynn Higgins

A thesis submitted in partial fulfillment
of the requirements for the
Degree of Bachelor of Arts with Honors
in Chemistry

Williams College
Williamstown, Massachusetts

May 2002

Abstract

Competence is a state that develops in *Bacillus subtilis* as a result of high cell density and limited nutrients. When cells achieve competence, they produce a complex DNA binding and uptake system that allows them to take up exogenous DNA, which can then be recombined into their chromosomes. A host of proteins regulate competence, and the regulation of competence is dominated by the competence transcription factor ComK. Although ComK has been identified for approximately 10 years, researchers have had difficulty in purifying the protein due to a variety of factors. Methods including His-Tags and maltose binding proteins have been used with limited success and the lack of a good method of purification has hindered the study of ComK. In addition, due to poor solubility, no structural analysis of ComK has been done. The objective of this research was to clone and produce truncated versions of the ComK DNA binding domain and to determine a procedure for the purification of wild-type protein and the truncated version. ComK was overexpressed in a strain containing plasmid pET21a-*comK*, in which the *comK* gene is under the control of the T7 promoter. Cells were induced with IPTG and harvested. Protein purification of the wild-type protein was achieved using a High Q ion exchange column, a step-wise ammonium sulfate precipitation to a concentration of 50% ammonium sulfate and a gel filtration column.

Acknowledgements

Professor Lovett for introducing me to the world of biochemistry and protein purification, for patiently answering all of my questions and for passing on his love of teaching chemistry.

Tom O'Gara for all of his help in lab, good music and lots of jokes.

Tracey and Eli for being the best labmates anyone could ask for.

The Chemistry Department for constant support over four wonderful years.

Spencer House Girls for all of the laughter and support throughout the years.

The Women's Crew Team for being constant sources of inspiration and pride for four incredible years.

And, of course, my family, for their overwhelming love and support always.

Table of Contents

Abstract	i
Acknowledgements	ii
Introduction	1
Competence Development	1
Regulation of Competence	2
DNA Uptake System	4
Expression of Genes Encoding the DNA Uptake System	5
ComK and DNA binding	6
RNA Polymerase	7
ComK and RNA Polymerase Interactions	12
Proposed Research	14
Experimental Procedures	17
Materials	17
PCR Reactions	17
Preparation of the Vector and PCR Products	18
Ligation and Transformation	19
Large Scale Induction and Fractionation	20
Ammonium Sulfate Precipitation	21
Chromatography	22
DNA Cellulose Column	24
Electroelution	25
Gel Filtration Chromatography	26
Results and Discussion	27
PCR Reactions and Cloning	27
Protein Purification Techniques	30
Conclusions	37
References	38
Appendices	43
Standard Procedures	43
Recipes	47

Introduction

Bacillus subtilis is a Gram-positive soil bacterium that can naturally differentiate into a state of competence. Genetic competence is characterized by the cell's ability to synthesize a complex DNA binding and uptake system and by activating recombination genes (Dubnau, 1993). Competence development can be initiated by a starvation-induced initiation process which is optimized at high cell densities with glucose as the main carbon source. When nutrients are limited, *B. subtilis* cells synthesize a motility and chemotaxis system allowing them to search for a more hospitable environment. Under the conditions mentioned above, the cells can develop genetic competence as they enter the post-exponential growth phase. Competence in *B. subtilis* is regulated by several factors including the synthesis of specific proteins and the need for specific environmental stimuli.

Competence Development

B. subtilis must attain several conditions in order to develop competence. The optimal time for competence development is two hours after exponential growth ceases, and a minimal media with glucose as the main carbon source (Dubnau, 1993). There must also be an environment with high cell density. When these prerequisites are achieved, competence can develop, but even in optimal conditions there is still only about a 10% rate of competence development. It is possible to distinguish between competent cells and non-competent cells with a density gradient and centrifugation, however this is not required for competence development. Competent cells have been found to be more buoyant than non-competent cells. In addition to these prerequisites, and in order to

reach competence, cells must accomplish two crucial things. In order to acquire new genetic material they must first synthesize a complex DNA binding and uptake apparatus.

Regulation of Competence

The key protein in competence development is the protein ComK. The regulation of ComK and competence overall is highly regulated as a two-fold mechanism in *B. subtilis*, the first part activated by quorum sensing and nutritional limitation, producing the protein ComS (Dubnau and Lovett, 2002). ComS goes on to activate the second part of the mechanism. The overall mechanism is begun by the release of the C-terminal 10 residues of the protein ComX, into the surrounding medium, where it acts as a pheromone signal (Lazazzera, et al., 1999). It is almost certain that ComX then binds directly to the histidine kinase, ComP, a membrane protein (Tortosa, et al., 2001). This pathway then converges with a second quorum-sensing pathway, at the point of ComA phosphorylation (Solomon, et al., 1995). ComA~P binds to the multi-gene operon *srfA* at two upstream palindromic ComA boxes, activating the *comS* gene and producing the protein ComS (Figure 1) (Weinrauch, et al., 1990).

ComK binds upstream from the *comK* promoter and initiates *comK* transcription and a relatively small amount of ComK can initiate a positive feedback loop (Hamoen, et al., 1998; Van Sinderen, et al., 1994). Under normal conditions, in order to prevent inappropriate induction of competence, a stable ternary structure in which the N-terminal domain of MecA binds ComK and the C-terminal domain of MecA binds ClpC (Persuh, et al., 1999). In addition, ComS binds the N-terminal domain of MecA which initiates the release of ComK from the ternary complex. ComK is then available to bind to its promoter and initiate the expression of the transformation genes (Figure 2).

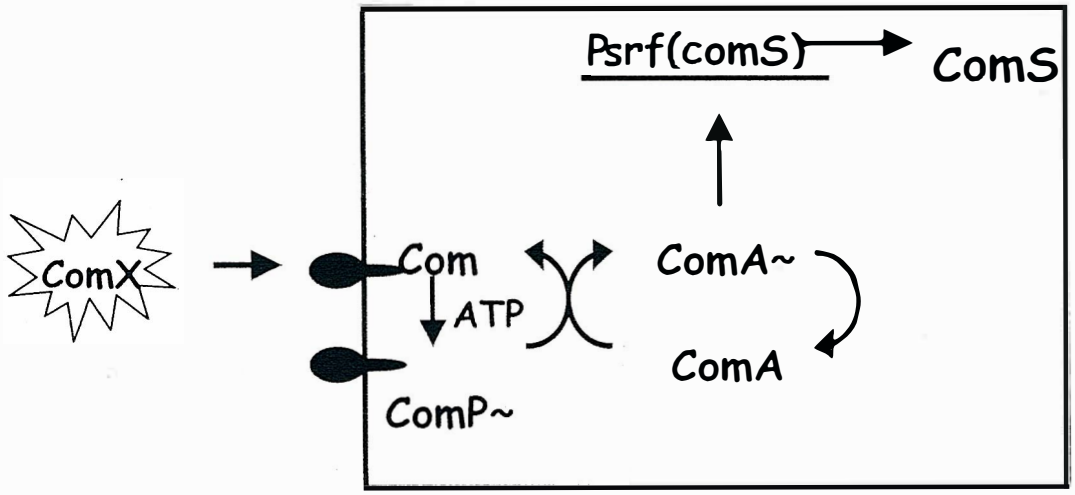

FIGURE 1: The quorum-sensing pathways of competence regulation. The box represents a cell. Adapted from Dubnau and Lovett (2002).

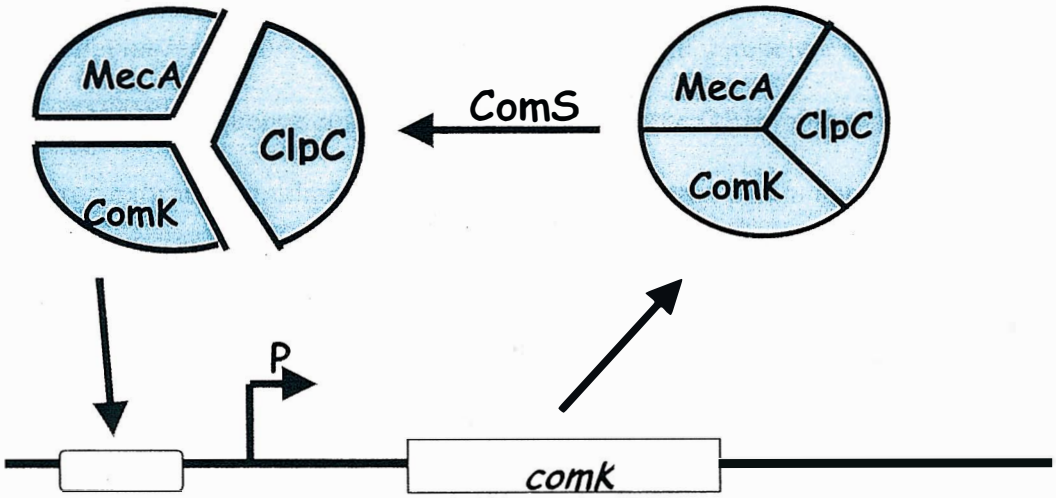

FIGURE 2: Regulation of ComK via MecA. Adapted from Dubnau and Lovett (2002).

DNA Uptake System

The overall mechanics of DNA uptake in Gram-positive bacteria begins with the binding of DNA to the cell surface proteins. *B. subtilis* has approximately 50 of these proteins (Dubnau and Cirigliano, 1972; Singh, 1972). There is subsequent double-strand cleavage of the DNA, usually beginning at the binding point of the DNA to the cell surface (Arwert and Venema, 1973; Dubnau, 1976; Dubnau and Cirigliano, 1972). One strand is taken up while the other is degraded, and the degradation products are released into the extracellular medium. After the single strand is internalized, it is presumed that it interacts with RecA, stimulating recombination. In *B. subtilis* it has been reported that uptake occurs in both the 5'-3' direction and the 3'-5' direction, unlike other similar model systems (Vagner, et al., 1990).

Research has shown that four loci code the DNA binding and uptake apparatus: *comC, comE, comF* and *comG*, which are all labeled as late-competence genes (Dubnau and Lovett, 2002). Sequence comparison of ComC and the ORFs of the heptacistronic *comG* operon exhibited similarities with type-IV pilins from *Pseudomonas* species and other pilin-like complexes. Type-IV pilins are proteins needed for synthesis and assembly of the DNA uptake system. Combined with experimental data, these observations led to a model in which DNA binding is accomplished by a pilin-like structure, composed of several *comG*-encoded proteins that anchors in the cytoplasmic membrane and possibly into the cell wall (Dubnau, 1999). ComEC is an integral membrane protein with at least six membrane-spanning segments, a hydrophilic C-terminal domain and a hydrophobic N-terminal domain. This protein appears to form a water-filled passageway that the DNA moves through to enter the cell. ComFA is

another integral membrane protein, however it is only partly in the bilayer and is on the inner face of the membrane. This protein may provide the energy for the transport of the single-stranded DNA across the membrane (Dubnau and Lovett, 2002). Two additional proteins coded by competence operons are NucA and Nin. NucA is a nuclease thought to be responsible for introducing double-stranded breaks in DNA at the cell surface and Nin forms a complex with NucA that limits the activity of NucA (Smith, et al., 1985; Vosman, et al., 1988). There are more proteins coded for by competence operons, but their function is not fully understood as of yet.

Expression of Genes Encoding the DNA Uptake System

When multiple copies of the *comC* promoter were introduced into *B. subtilis*, Mohan and Dubnau (1990) observed reduced competence and reduced expression of the late competence genes *comC, comG* and *comE*, which suggests the titration of a specific transcription factor. Based on these results Mohan and Dubnau proposed the existence of a competence transcription factor (CTF) that is required for the expression of late competence genes, as discussed above. Van Sinderen, et al. (1994) identified the gene responsible, *comK*, and it was found to be responsible for the expression of late-competence genes, for the expression of recombination genes such as *recA* or *addAB* and for its own expression (Haijema, et al., 1995; Haijema, et al., 1996; Van Sinderen, et al., 1995; Van Sinderen and Venema, 1994). ComK-footprinting analysis with the promoters of *comC, comG, comE, comF, comK, addAB* and *recA* found a conserved AT-rich palindromic sequence that is the recognition site for ComK (Hamoen, et a., 1998). It was also found that transcription stimulating activity of ComK can be demonstrated *in vitro*.

Based on this research it is apparent that the protein ComK is responsible for the expression of all late genes necessary for competence.

ComK and DNA Binding

ComK is a 192-residue protein that binds to an $AAAAN_5TTTT$ sequence (ComK box) upstream of its target genes (Hamoen, et al., 1998). The ComK box is always found in duplicate, on the same face of the DNA helix, separated by 2 to 4 helical turns. The variation in distance between the two ComK boxes distinguishes three classes of ComK-regulated promoters. The first class of promoters includes *recA, addAB, din A* and *nucA* and has a distance between ComK boxes of 21 nucleotides, two helical turns. The second class of promoters is the late competence genes (*comC, comG, comE,* and *comF*) and have the AT-boxes separated by 31 nucleotides, which is three helical turns. The third class is the *comK* promoter, which has the AT-boxes repeating every 4 helical turns, or 44 nucleotides. Despite the variations in length and number of turns, in all classes the AT-boxes appear on the same face of the DNA. The presence of the ComK dimers on the same face of the DNA allows the dimers to form tetramers (Hamoen, et al., 1998), a process shown to be essential for maximum efficiency in ComK binding. In class 2 and class 3 promoters, the formation of tetramers is accomplished only when two ComK dimers stretch across a large distance, which suggests some sort of DNA contortion. ComK has been shown to induce DNA bending by an estimated 75 degrees in the promoter's *comG* and *comF* (Hamoen, et al., 1998). Hydroxyl radical analysis shows that protected bases in the upper and lower strands of DNA closest to the center of the AT-box dyad symmetries are offset in the 3' direction relative to the centers of the AT-

boxes. This indicates that ComK monomers contact each other across the major groove of DNA, which in turn implies that ComK recognizes the AT-box consensus sequence in the minor groove of DNA (Hamoen, et al., 2001). DNA bending is observed with transcriptional activators and is thought to facilitate the wrapping of DNA around RNA polymerase.

RNA Polymerase

In order to understand the effect of ComK binding and DNA bending on the binding of RNA polymerase, it is necessary to be familiar with RNA polymerase's structure and function. Most of the research done with RNA polymerase has been done in *E. coli*. The structure of RNA polymerase in *E. coli* is conserved in *B. subtilis* except for the presence of the σ subunit. RNA polymerase is a multisubunit enzyme consisting of five subunits, $\alpha_2\beta\beta'\sigma$ (Murakami, et al., 1997). The entire structure is called a holoenzyme and the core enzyme consists of all of the subunits except σ. The σ subunit recognizes the promoter region on DNA and is weakly associated with RNA polymerase. The remaining subunits bind to the promoter region after recognition, then transcription is initiated.

Initiation begins when subunits bind up to three promoter elements: the –10 and –35 bp, recognized by the σ subunit and may involve the UP element, recognized by the α subunit (DeHaseth, et al., 1998). The UP element is located upstream of the –35 region between –40 and –60 bp. After RNA polymerase binds to the DNA, the RNA polymerase-DNA structure changes conformation and forms an open complex. The open complex allows the double-stranded DNA to begin to partially separate and RNA

polymerase gains access to the bases. After approximately 10 bp's are transcribed, the σ subunit dissociates from the promoter, allowing the RNA polymerase to bind more tightly and elongation can continue.

The UP element has been shown to stimulate promoter activity ~30-fold in certain genes (Ross, et al., 2001). The consensus sequence of the UP element has been identified as containing alternating A and T tracts in proximal and distal subunits from which the α subunits of RNA polymerase bind to the DNA. Each α subunit is 329 amino acid residues and is involved in RNA polymerase assembly in addition to DNA binding (Murakami, et al., 1997). The α subunits contain two independently folded domains with specific functions. The amino-terminal domain (α-NTD, residues 8-233) is responsible for the dimerization and interactions between β and β', and the carboxy-terminal domain (α-CTD, residues 245-329) contains DNA binding determinants (Hayward, et al., 1991; Blatter, et al., 1994). A protease-sensitive flexible linker connects the two α subunits, which presumably allows transcription activators to bind to a variety of places on DNA and still have access to the α subunits (Blatter, et al., 1994; Negishi, et al., 1995).

The α-CTD also plays an important role in communication with class I transcription factors (i.e. those that require the α-CTD subunit for activation) such as cAMP receptor protein (CRP), MarA, OxyR, OmpR, Rob, SoxS, TyrR and GalR (Murakami, et al., 1997). The targeting of the α subunits transcription factor action led to the discovery of α-DNA interactions in the early 1990s (Russo and Silhavy, 1992). The importance of the α subunits in transcription initiation became evident after the successful production and purification of RNA polymerases with mutant α in which the

CTD section of the α subunit was missing (Estrem, et al., 1999). These mutants were defective in rRNA promoter function which led to the conclusion that the α-DNA interactions play an important role in transcription (Gourse, et al., 2000).

There has been a lot of work done trying to elucidate the interaction between the α subunits and the UP element through extensive footprinting and crosslinking studies. One model system used in studying these interactions contains the gene *rrnB*. Footprints of RNA polymerase on *rrnB*P1 and other promoters extend about 60 bp upstream of the transcription starting point and protection upstream of ~-40 is due to interactions with α (Kolb, et al., 1993, Ross, et al., 1993). The DNA backbone regions centered at ~-42 and ~-52 of the *rrnB* P1 UP element are protected against hydroxyl radical cleavage by α or RNA polymerase. The specifics of α binding have been complicated by specific features of interactions with different promoter upstream regions. The α subunits can bind to DNA in either a sequence-specific or a sequence-nonspecific manner, and some footprinting analysis has shown multiple regions of protection within one upstream region (Ross, et al., 2001). In addition, comprehensive cross-linking studies suggest that the two α-CTDs may oscillate among different binding sites or alternatively may bind to different sites in different molecules in the population (Naryshkin, et al., 2000).

The variability of binding sites is due to the long, flexible connector between the α-NTD and α-CTD in addition to intrinsic DNA distortions and protein-induced DNA bending in initiation complexes (Gourse, et al., 2000; Naryshkin, et al., 2000). The linker allows the CTD to move at least 30 bp along the promoter region of the DNA. This is supported by reports that a single CRP site can be shifted from –61.5 to ~-90 which retaining transcriptional activity and the critical determinants of the CRP. NMR studies

also indicate the presence of a linker that is at least 13 residues long (Murakami, et al., 1997).

Research has also shown that the two α-CTD subunits from β and β' bind separately to the UP element. It was first determined that α-CTD is a minor groove binding protein through hydroxyl radical footprinting experiments (Murakami, et a., 1997). Protected regions were separated by 3 bp, which suggests binding across the minor groove (Newlands, et al., 1991; Ross, et al., 1993; Estrem, et al., 1998). In studies with distamycin, an antibiotic that binds tightly in the minor groove, the α subunit was unable to bind to the UP element. Distamycin also inhibits UP-dependent transcription when UP-independent transcription was unaffected, indicating that the minor groove plays a key role with the UP element (Gourse, et al., 2000). The possible role of major groove contacts in α-UP element interactions was tested by interference footprinting using RNA polymerase and promoter fragments prepared by limited incorporation of dUTP or 7-deaza-7-nitro-dATP into the DNA. These experiments showed that the major groove contacts contribute little if any to α-UP element interactions (Ross, et al., 2001).

After determining that the minor groove was the main site of binding for the α-CTD on the UP element, research was done to determine where the individual units bind. Although α-CTD purifies as a dimer in solution, there is no evidence that the molecules interact when bound to DNA (Gourse, et al., 2000). UP element subsites of *rrnB* were tested for the interaction with one copy of α-CTD using α RNA polymerases lacking one α-CTD using DNA affinity cleavage by a reagent attached to a specific site on the proteins (Murakami, et al., 1997). The two α-CTDs were differentiated based on the fact that one α binds to β and the other to β'. There are three possible configurations of

binding interactions, the first being when the promoter has a full UP element with two consensus subsites. In this case both α-CTDs must bind, one at the distal subsite and one at the proximal. Both consensus proximal sites and consensus distal sites require only one α-CTD (Murakami, et al., 1997). The α-CTDs from the β and β' subunit can function interchangeably (Estrem et al., 1999). Murakami, et al. (1997) has proposed a model for the mechanism of independent positioning of two α-CTD subunits by two CRP dimers. They found that in the case of promoters carrying tandem DNA sites for CRP, each of the α-CTD subunits interacts with each of two CRP dimers. The α-CTD from the β subunit has a constant site of binding, while the α-CTD subunit from β' shifts binding locations depending on the location of the upstream CRP dimer.

Before the α-UP element interactions were discovered, it was found that some naturally occuring promoters have phased A-tracts upstream of the –35 hexamer and that these A-tracts increase promoter activity when fused to core promoters in synthetic promoter hybrids (McAllister and Achberger, 1988). Due to the fact that A-tracts phased with a helical repeat results in DNA bending of approximately 18 degrees per A-tract, it was proposed that DNA curvature *per se* facilitated transcription (Gourse, et al., 2000). The presence of A-tracts in UP elements offers an alternative explanation: A-tracts directly affect transcription because of DNA-α interactions. Evidence for this model includes the failure of phased A-tracts to stimulate expression when the promoters were transcribed by mutant RNA polymerases and footprinting analysis showing that protection of the A-tracts was α-CTD dependent (Aiyar, et al., 1998). These results demonstrated that in most cases A-tracts increase transcription through DNA-protein interactions, such as the binding of α-CTD to DNA, by causing bending of the DNA.

Therefore phased A-tracts should be considered a subset of UP element sequences (Gourse et al. 2000). In addition, the presence of DNAse I hypersensitive sites in footprints of the UP element suggest that after RNA polymerase binding there is DNA distortion within the α binding site (Gourse, 1988; Estrem, et al., 1999). Therefore, it seems that some portion of the wrapping of DNA around RNA polymerase in the transcription initiation complex is due to α-DNA interactions.

ComK and RNA Polymerase Interactions

In *B. subtilis*, the SOS system is an inducible DNA repair system that is inhibited by the protein LexA. Under normal conditions, LexA binds to the *recA* promoter and acts as a repressor (Raymond-Denise, et al., 1991). When there is DNA damage, LexA is activated by RecA and undergoes autocleavage which stops LexA's repression of recA and initiates the SOS system (Miller, et al., 1996). Hamoen, et al., (2001) have found that ComK can override the transcription inhibition by LexA without displacing LexA. Mobility shift assays showed that despite overlapping recognition sites both ComK and LexA in the presence of each other are able to bind to the *recA* promoter with equal binding affinity to when the proteins are alone. In addition, *in vitro* transcription assays showed that the presence of ComK was sufficient for competence induced *recA* expression.

By looking specifically at the interaction at the *recA* promoter, Hamoen, et al., also found that the SOS box in this promoter is not found in the typical locations of –10 and –35, rather it is located at –52, which corresponds to the distal site of the UP element. Just as in *E. coli*, the UP elements of *B. subtilis* are needed for high level expression of

certain genes and have been shown to interact with the *B. subtilis* α-CTD (Hamoen, et al., 1998). In addition, the sequence of the *recA* promoter between –46 and –59 is AT rich, resembling the *E. coli* UP element consensus sequence. This study also proposes the idea that the bending of DNA by ComK is an essential feature to stopping LexA inhibition, either by allowing the RNA polymerase to come in and displace the LexA or by providing an alternative binding interaction allowing RNA polymerase to carry out transcription (Figure 3) (Hamoen et al., 2001). Specifically in the *recA* promoter, ComK must bind to all four AT boxes in order provide the bending necessary for the wrapping of DNA around RNA polymerase. Further support comes from evidence that deleting the most upstream AT box causes a 70% reduction in competence-related *recA* expression (Cheo et al., 1993).

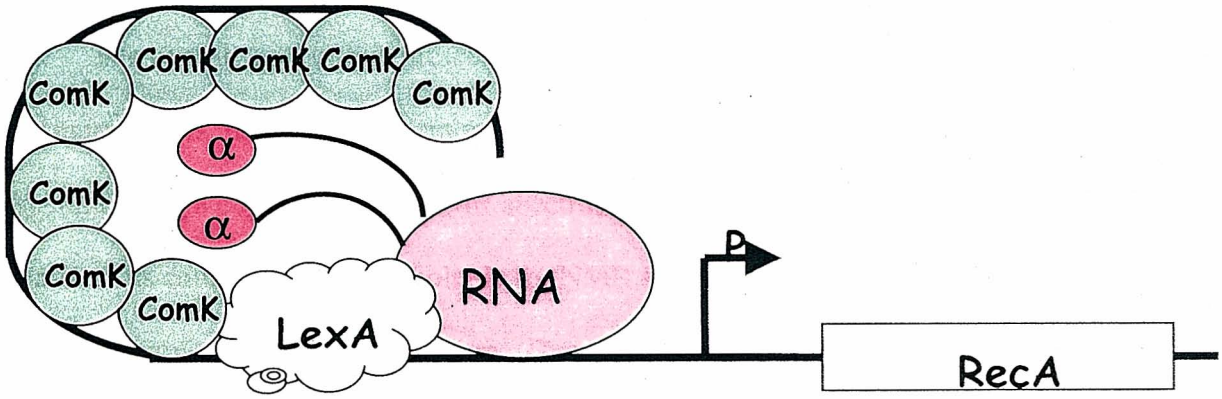

FIGURE 3: Model of possible ComK-LexA-RNAP interactions on the *recA* promoter.

In addition to DNA wrapping, the binding of RNA polymerase to the *recA* promoter could result from direct interactions between ComK and α-CTD (Hamoen, et al 2001). Based on evidence that *E. coli* α-CTDs interact with the proximal UP element subsite at position –42 and the distal UP element subsite at position –52, it appears significant that the downstream AT boxes of the *comC*, *comE*, *comF* and *comG* promoters are centered around position –48. The location of these binding sites relative to the α-CTD binding sites suggests a mechanism in which ComK affects the binding of α-CTD. This could be a direct interaction in which ComK provides an additional binding site for the α-CTD when LexA inhibits the UP element. Alternatively it could be an indirect interaction in which ComK induces a DNA conformation change that allows α-CTD binding.

Proposed Research

I. Cloning *B. subtilis comK* and *rpoA* α genes

The overall goal of this project is to characterize the binding of ComK and RNA polymerase in *B. subtilis*, to determine any interactions, and to then determine their structure using a combination of NMR spectroscopy and molecular modeling. However, ComK in its complete form is insoluble for NMR spectroscopy because it is too big and has a tendency to aggregate in solution. Therefore we propose to attempt to clone and purify the DNA binding domain of ComK in order to determine protein-protein interactions and protein-DNA interactions. The DNA binding domain of ComK has been identified as the last 97 amino acid residues of the protein. In conjunction with Tracey Jackson, we propose to clone the ComK DNA binding domain and the RNA polymerase

α-CTD subunit into a pET21a+ vector downstream of a T7 promoter. The recombinant vectors will then be transformed into BL21-Gold competent cells. These cells encode for T7 RNA polymerase under the control of a *lac* promoter. After transformation and growth in liquid media this *E. coli* expression system can be induced using IPTG.

II. Purification of Wild-Type ComK

The purification of wild-type ComK has been attempted in the past using methods such as His-Tag purification and maltose binding proteins. These methods have proved to be unsuccessful with low yields of purified protein. Purification has also proved difficult due to the solubility of ComK in various solvents. As a result, a major component of this thesis is the purification of ComK using more traditional methods of protein purification, including affinity columns, gel filtration columns and ammonium sulfate precipitation. Amish Shah (2000) successfully cloned *comK* into a pet21a+ vector, ready for induction using IPTG. Following successful induction of the protein and extraction of a crude extract, the various purification techniques will be tested on the crude extract in order to determine the best means of purification at the highest possible yield.

III. Targeted Footprinting using FeBABE

Once the DNA binding sites of the two proteins have been cloned and purified, the next step of the project is to begin to analyze the protein-protein interactions of ComK and the α-CTD of RNA polymerase and the protein-DNA interactions between the two proteins and the DNA. A current method is known as targeted footprinting and

in our case we will use the cutting reagent iron (S)-1-(*p*-bromoacetamidobenzyl)-ethylenediaminetetraacetate (FeBABE) (Traviglia et al., 1999). This method involves attaching the cutting reagents randomly to lysine residues on a proteins surface and then utilizing this lysine-labeled protein to cleave the polypeptide backbone of the other protein at exposed residues adjacent to is binding site (Traviglia et al, 1999).

This method is a straightforward means of labeling proteins because lysine residues are common on the surface of proteins and the cleavage of nucleic acids by lysine-labeled proteins has been documented (Hanai and Wang, 1994). An important feature of this procedure is that a macromolecule bound to any site on the lysine-labeled protein will be cleaved. As a result, when preparing the lysine-labeled proteins, it is necessary to place cutters in sites over the entire surface so that the entire protein is reached by the cutting reagents. Crystal structures of proteins have shown that lysines are distributed across the protein surface, with the average lysine ϵ-amino group about 12 A away from the corresponding group on the nearest lysine. This technology allows for efficient screening of protein-protein interactions and makes it easy to assess where multiple factors bind to a common target (Traviglia et al., 1999).

Experimental Procedures

Materials

Oligonucleotide primers in the forward direction were special ordered from Sigma Genosys. These primers included ComK1, 5'CACAAGCCGCCGCATATGGTGG ACCC-3'; and ComK2, 5-CCAATGGGAAACCATATGGAGCTGCCG-3'. The oligonucleotide primer in the reverse direction was oligo 104, 5'-GGAATTCCGACGG CAAAGCTTTATG-3' from Shah (2000). PCR reaction materials came from Stratagene and materials for the site-directed mutagenesis was from Stratagene. The pET 21a(+) vector and the DH5α cell and the BL21-Gold cells were from Novagen. The Wizard Plus SV MiniPreps DNA Purification Systems was from Promega. FPLC columns (1mL) were from Bio-Rad Laboratories, dsDNA cellulose was from USB Laboratories and Sephadose was from Sigma.

PCR Reactions

Coding portions for the last 111 amino acids (ComK1) and the last 62 amino acids (ComK2) of *comK* were amplified from chromosomal *B. subtilis* DNA using polymerase chain reaction (PCR) as described by Shah (2000). The oligos in the forward direction were mutated to include an *NdeI* site preceding the start codon for ease in cloning and the reverse oligo included a *HindIII* site. The mutations are outside of the open reading frame and have no affect on the makeup of the desired protein. A control reaction was prepared containing 5 μL of 10x reaction buffer, 2 μL (10 ng) of pWhitescript 4.5-kb control plasmid (5 ng/μL), 1.25 μL (125 ng) of oligo control primer #1 (34-mer, 100 μg/μL), 1.25 μL (125 ng) of oligo control primer #2 (34-mer, 100 μg/μL), 1 μL of dNTP

mix and distilled water (ddH_2O) to a final volume of 50 μL. Five sample reactions of varying chromosomal DNA concentration were set up for oligonucleotide primer ComK1 and five for ComK2. Each reaction mixture included 5 μL of 10x reaction buffer, x μL (5-50 ng) of chromosomal DNA, x μL (125 ng) of oligonucleotide primer #1 (forward direction, ComK1 or ComK2), x μL (125 ng) of oligonucleotide primer 104, 1 μl of dNTP mix and ddH_2O to a final volume of 50 μL. The volumes of chromosomal DNA varied from 2-20 μL and the final volume of ddH_2O was adjusted accordingly. After setting up both the control and sample reactions, 1 μL of Pfu DNA polymerase (2.5 U/μL) was added to each reaction mixture, followed by an overlay of 30 μL of mineral oil. The reactions were cycled 30 times through 96^0C (1 min), 55^0C (1 min) and 72^0C (2 min) from Shah (2000). Following temperature cycling, the reactions were cooled to less that 37^0C and stored at -20^0C. The reaction products were run on a 1% agarose gel with 1:5000 ethidium bromide (10 mg/mL). Samples were made of 10 μL aliquots mixed with 2 μL of bromophenol blue loading buffer and were electrophoresed at 100V with TBE. Results were seen under UV light.

Preparation of the Vector and PCR Products

Restriction enzyme digests were run on both the PCR products and the pET 21a(+) vectors. Preparation of the vector was done by setting up a reaction mixture of 5.6 μL pET 21a(+) vector, 3 μL multicore buffer, 1 μL NdeI enzyme, 1 μL HindIII enzyme and 19.4 μL ddH_2O for a total volume of 30 μL. The reaction mixture for digesting the PCR product was 10 μL PCR product, 3 μL multicore buffer, 1 μL NdeI enzyme, 1 μL HindIII enzyme and 15 μL ddH_2O for a total volume of 30 μL. The reaction mixtures

were incubated at 37^0C for 4 hours and then gel purification and gel extraction were done on the products.

Ligation and Transformation

The ligation reaction mixture was made up of 5 μL 10x Ligase Buffer, 2 μL pET vector, 1 μL T4 DNA ligase diluted to 0.2-0.4 Weiss units/μL, 2 μL target gene insert (purified PCR product) and 10 μL ddH_2O to a final reaction volume of 20 μL. The ligase was added last and the mixture was mixed gently with a pipet tip. The reaction was incubated at 16^0C overnight. In addition a control reaction was set up with the gene insert omitted and replaced with a larger volume of ddH_2O.

To begin the transformation process, the DH5α competent cells were gently thawed on ice. For each reaction, aliquots of 50 μL of the cells were put into prechilled 15 mL Falcon polypropylene tubes. A sample of 1 μL of the ligation product or control DNA was added to separate aliquots of the competent cells. The reactions were mixed gently and incubated on ice for 30 minutes followed by a heat pulse for 45 seconds at 42^0C. The reaction mixtures were put on ice for an additional 2 minutes. Preheated NZY^+ broth at 42^0C was added to each reaction in a volume of 0.5 mL and the reactions were incubated at 37^0C for 1 hour with shaking at 225-250 rpm. The transformation reactions were immediately plated using 10 μL of the reaction mixture on LB agar plates with 50 μg/μL of carbenicillen. The transformation plates were incubated at 37^0C overnight.

After the successful growth of colonies on the LB-agar plates, 5 mL overnight cultures of LB broth and 50μg/μL of carbenicillen were set up and each was innoculated

with one colony. 10 μL aliquots from each overnight culture were plated on new LB agar plates with 50 μg/μl of carbenicillen and incubated overnight at 37^0C. Wizard Plus SV MiniPreps were done on individual colonies to isolate the plasmids with the desired gene. The products from the MiniPreps were run on an agarose gel to confirm that the plasmid has been separated and purified. A subsequent transformation was done with the purified DNA plasmids into BL21-Gold cells, using the protocol outlined above. The BL21-Gold competent cells contain the T7 RNA polymerase, which upon induction result in the overproduction of the desired protein. The transformation products were plated on LB agar plates with 50 μg/μL of carbenicillen and incubated overnight at 37^0C.

Large Scale Induction and Fractionation

The colonies from the plated BL21-Gold competent cells were then available for large-scale growth and induction of the plasmid for overproduction of the desired proteins. In addition to the vectors coding for the shortened segments of ComK, the vectors with the entire sequence for ComK cloned by Amish Shah (2000) were induced via the following procedure. Individual colonies were inoculated into 5 mL starter culture of LB broth and 50 μg/μL of carbenicillin and incubated in 37^0C with shaking at 225-250 rpm to an $OD_{600} \leq 0.5$. LB broth with 50 μg/μL of carbenicillin were prepared in 1 L amounts and inoculated with the full amount of the 5 mL starter cultures. These volumes were incubated with at 37^0C with shaking at 225-250 rpm to an OD_{600} of 0.5-1.0. To induce the expression of the desired gene IPTG was added to the mixture to a final concentration of 1 mM and incubation was continued for two hours. In order to quickly test for induction, 1 mL aliquots of culture were taken before induction and at the

end of the incubation. The samples were spun in a microcentrifuge, the pellets were resuspended in SDS loading buffer and were run on an SDS-PAGE gel.

The induced cells were harvested using centrifugation at 6,500 x g for 15 minutes at 4^0C. The supernatant was discarded and the cells were thoroughly resuspended in ice-cold resuspension buffer using 2 mL of buffer per gram of cells (wet weight). The suspension was cooled on ice to 4^0C to prevent heating during cell breakage. Lysis of the cells was done using lysozyme plus sonication. Lysis was started by adding lysozyme to a final concentration of 100 µg/mL from a freshly prepared stock of 10 mg/mL in water. The mixtures were incubated at 30^0C for 15 minutes. They were mixed by swirling and sonicated on ice until cells were lysed and the solution was no longer viscous. It is important to keep the cells cool during sonication to prevent heat denaturation of the proteins. The homongenate was centrifuged at 12,000 x g at 4^0C for 30 minutes. The soluble cell extract was removed and stored for purification at 4^0C. The pellet was also saved in case all of the cells were not lysed or the protein was insoluble in the Resuspension buffer. Samples of the crude extract were run on SDS-PAGE gels to test for the presence of the induced proteins.

Ammonium Sulfate Precipitation

Ammonium sulfate precipitation was done in a step-wise manner, beginning with 1 mL of crude extract protein solution using a protocol from Doonan (1996). The precipitation was done in the cold room at 0^0C. The initial precipitation was done with samples at 30%, 50% and 75% saturation at 0^0C. Subsequent precipitations included a 20% and 40% saturation at 0^0C. Appropriate amounts of ammonium sulfate were

weighed out (Table 1) and then ground into a fine powder before being added slowly to the crude extract. For the initial 30% precipitation, 0.164 g of ammonium sulfate was added. The crude extract was put in a beaker with a stir bar at a low speed. The ammonium sulfate was added slowly in small batches over a period of several minutes. After the ammonium sulfate was added, the solution was equilibrated for 10 minutes and then spun down in a microcentrifuge in the cold room for 10 minutes. The supernatant was separated from the pellet and an aliquot of the supernatant was collected for running on a 15% SDS-PAGE gel. The remainder of the supernatant was used in subsequent precipitations. In subsequent precipitations, additional ammonium sulfate should be added to the solution up to the desired amount (i.e. when going from 30% to 50% only add 0.127 g, rather than an addition 0.291 g). After the precipitations were completed, a 15% SDS-PAGE gel was run to show the results. It is important that this precipitation be done in a stepwise manner and that it be done in the cold room.

Table 1
Amounts of Solid Ammonium Sulfate Required to Change the Concentration of a Solution from a Given Starting Value to a Desired Target Value at 0^0C

	Target percentage saturation at 0^0C (g of solid ammonium sulfate/mL)						
Initial % saturation at 0^0C	20%	30%	40%	50%	60%	70%	75%
0%	0.106 g	0.164 g	0.226 g	0.291 g	0.361 g	0.436 g	0.476 g

Chromatography

Fast Pressure Liquid Chromatography (FPLC) was done using a variety of BioRad 1 mL prepacked columns. The first column used was a t-butyl HIC column.

Buffers (high salt buffer A: 100 mM sodium phosphate, 2.4 M ammonium sulfate, pH 6.8; low salt buffer B: 100 mM sodium phosphate, pH 6.8) were made in 1 L volumes and filtered through a 0.45 μm filter and stored at 4^0C. The t-butyl HIC columns are unique amount the columns used in this research in that they begin with a high salt buffer and elute with a low salt buffer. The FPLC was equilibrated using the automated wash program and the column was prepared by washing the cartridge with the low salt buffer B for 2 minutes at 2 mL/min, then washing the cartridge with high salt buffer A for 10 minutes at 6 mL/min. The cartridge was inverted to allow any excess air to be expelled from the cartridge and the cartridge was washed again with low salt buffer B for 10 minutes at 6 mL/min. The cartridge was then equilibrated with high salt buffer A for 2 minutes at 6 mL/min. The cartridge was returned to the proper starting point, prepared for use.

Filtered 50% ammonium sulfate sample was loaded into a 200 μL sample loop and run on program 1 (35 minutes at 0.7 mL/min on a linear gradient from 100% buffer A to 100% buffer B). Samples were collected in 0.3 mL amounts and every other sample was run on a 15% SDS-PAGE gel.

In addition to the t-butyl HIC column, three ion exchange columns were run, eluting with a low salt to high salt gradient. The first was a High Q column, an anion exchange column. Buffers (low salt buffer A: 20 mM Tris-HCl pH 8.0, 0.1 mM EDTA, 50 mM NaCl; high salt buffer: 20 mM Tris-HCl pH 8.0, 0.1 mM EDTA, 1.0 M NaCl) were mixed in 1 L volumes and filtered through a 0.45 μm filter and then stored at 4^0C. The FPLC was equilibrated using the automated wash program and then the cartridge was prepared first by washing the cartridge with low salt buffer A for 2 minutes at 2

mL/min, followed by washing with the high salt buffer B for 10 minutes at 6 mL/min. The cartridge was then inverted to allow excess air to escape and then it was run with low salt buffer A for 10 minutes at 6 mL/min. The cartridge was then returned to its correct position. A filtered sample of crude extract was loaded into a 100 µL sample loop and run on program 1 (35 minutes at 0.7 mL/min on a linear gradient from 100% buffer A to 100% buffer B). Samples were collected in 0.3 mL volumes and aliquots of samples were run on an SDS-PAGE gel, based on the locations of peaks on the graph.

The flow-through fractions from the High Q column were combined and a 100 µL sample was subsequently run on a High S cation exchange column. The High S column was run with the same buffers from the High Q column and the cartridge was prepared via the same method. Program 1 was run and samples were collected in 0.3 mL volumes and run on a SDS-PAGE gel. Another cation exchange column, a CM cartridge, was run with a 100 µL sample of filtered crude extract. The cartridge was prepared with the same method as the High Q and High S columns and program 1 was used. Aliquots of 0.3 mL samples were run on a SDS-PAGE gel.

DNA Cellulose Column

In addition to FPLC, a DNA cellulose column was made using dsDNA cellulose. The column was made using 0.3 grams of the dsDNA cellulose and buffers (low salt buffer: 20 mM Tris-HCl pH 8.0, 0.1 mM EDTA, 0.5 mM NaCl; high salt buffer: 20 mM Tris-HCl pH 8.0, 0.1 mM EDTA, 1 M NaCl). The DNA cellulose was added to the empty cartridge followed by 2 mL of low salt buffer that was allowed to settle. After

settling, the column was washed using three column volumes of the low salt buffer. A 100 μL sample of crude extract was added to the column and 100 μL fractions were collected as an additional three column volumes of low salt buffer was added. After the low salt buffer, three column volumes of high salt buffer were added and run to elute proteins and 100 μL fractions were also collected.

Electroelution

Electroelution is an alternative protein purification method to columns. In this case a crude extract sample was run on a 15% SDS-PAGE gel. The band containing ComK at ~22 kD was cut out of the gel using a spatula and placed in an Elutrap made by Schleicher & Schuell. The Elutrap was prepared by inserting two BT1 filters and one BT2 filter as laid out in manual and then the Elutrap was placed in an electrophoresis chamber with 1x TG buffer. The end of the Elutrap with the gel sample in it was placed in the direction of the anode (positive charge) and the system was run at 100 V for 4-6 hours. The solvent in the small collection chamber between one BT1 filter and the BT2 filter was extracted using a Pasteur pipette and then guanidine hydrochloride was added to a concentration of 6 M in order to replace all of the SDS that is bound to the protein, denaturing it. The solution was then dialyzed overnight in resuspension Buffer. The resuspension Buffer was changed twice to ensure total dialysis of the guanidium hydrochloride out of the solution and out of the protein. The dialyzed solution was then run on a 15% SDS-PAGE gel to see if the protein had been separated successfully.

Gel Filtration Chromatography

A gel filtration column was also run using Sephadose beads from Sigma and the resuspension buffer as the column buffer. An empty column was set up in a refrigerator and the bottom tube was clamped shut. 3 mL of resuspension buffer was added to the column followed by slow addition of the Sephadose beads in 1 mL amount until the beads were approximately 2 cm from the top of the clear column. The slurry was allowed to settle and then a minimum of three column volumes of the resuspension buffer was run over the column to thoroughly wash the beads. 100 μL of the High Q flow-through fraction was added to the top of the column head along with 2 μL each of vitamin B_{12} and blue dextran as molecular weight markers. The column was then run with resuspension buffer and fractions were collected in one drop amounts (approximately 50 μL).

Results and Discussion

PCR Reactions and Cloning

PCR reactions were run with oligos 186 or 187 in the forward direction and oligo 104 in the reverse direction in order to amplify the genes for two truncated versions of the protein ComK. Oligos 186 and 187 contained the Nde I restriction site, and oligo 104 contained the HindIII restriction site (Table 2). Truncated versions of the sequence were produced because it is known that the DNA binding site of ComK is the last 97 amino acids of the protein. Oligo 186 and oligo 104 were used to create a truncated gene of the last 111 amino acids by having oligo 186 cut at base pair 244, while oligo 187 and oligo 104 were used for the gene coding for the last 62 amino acids of the sequence by cutting at base pair 390.

Table 2
Nucleotide sequence of primers used for PCR amplification

Primer Name	Sequence
186 (ComK1)	5'-CACAAGCCGCCGCATATGGTGGACCC-3'
187 (ComK2)	5'-CCAATGGGAAACCATATGGAGCTGCCG-3'
104	5'-GGAATTCCGACGGCAAAGCTTTATG-3'

The highlighted nucleotide sequences of primers ComK1 and ComK2 are *NdeI* restriction sites and the highlighted sequence of primer 104 is the *HindIII* restriction site.

The products of these PCR reactions were run on a 1% agarose gel to see if the PCR reactions worked and they were found to be successful. Both of these gene inserts and the pET21a+ vector were digested with the restriction enzymes NdeI and HindIII. In addition, the inserts had λ DNA added to the reaction mixture before digestion as a control in order to make sure that the enzymes were working properly (Figure 4).

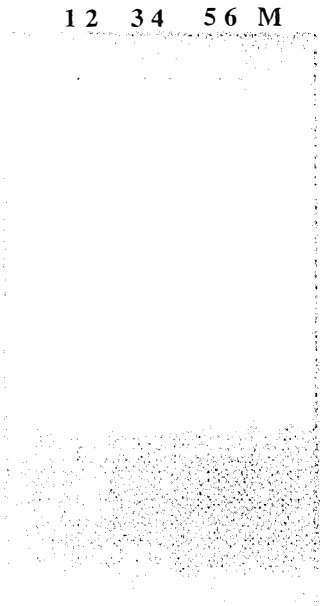

FIGURE 4: 1% agarose gel. Lanes 1-2: ComK1 digest products. Lanes 3-4: ComK1 digest products with λ-DNA as a control. Lanes 5-6: ComK2 digest products. Lane M: Kb ladder with DNA fragments.

After the insert and vector were digested, they were gel-purified using a 1% agarose gel. It is possible to see the size difference between the two inserts in this gel (Figure 5). The insert and the vector were then cut out of the gel and eluted from the gel overnight. Standard ethanol precipitation was done on the products and the ligation mixture was set up as described in Experimental Procedures. The truncated genes of the DNA binding domain of ComK were inserted into the pET21a+ vector behind the T7 promoter. The ligation was deemed successful based on a small sample run on a 1% agarose gel.

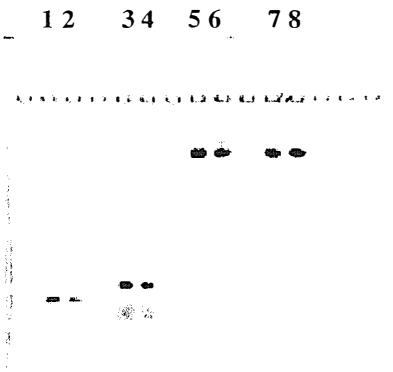

FIGURE 5: 1% agarose gel. Lanes 1-2: ComK2 inserts. Lanes 3-4: ComK1 inserts. It is possible to see that ComK2 is the smaller insert. Lanes 5-6: pET21a+ digest #1. Lanes 7-8: pET21a+ digest #2.

Following ligation, the products of the reaction were transformed into DH5-α cells. Transformation was done four times and was successful in one of the attempts. DNA extraction was done using a Wizard MiniPrep and the DNA was digested with NdeI and HindIII. However, no inserts were found in the products and the research moved on to the purification of the wild-type ComK protein.

Amish Shah '00 successfully cloned the entire *comK* gene into a pET21a+ vector behind the T7 promoter for his honors thesis in 2000, however, he experienced difficulty in purifying the protein. He created two strains, 447 and 448, each containing the pET21a+-*comK* plasmid in BL21-Gold cells. Both strains were grown as described in Experimental Procedures and induced with IPTG. The IPTG serves to derepress the T7 RNA polymerase gene on the chromosome and the *comK* gene, both of which are under

the control of the *lac* repressor. Aliquots taken before and after induction were run on a 15% SDS-PAGE gel and successful induction of ComK was found in both of these strains (Figure 6).

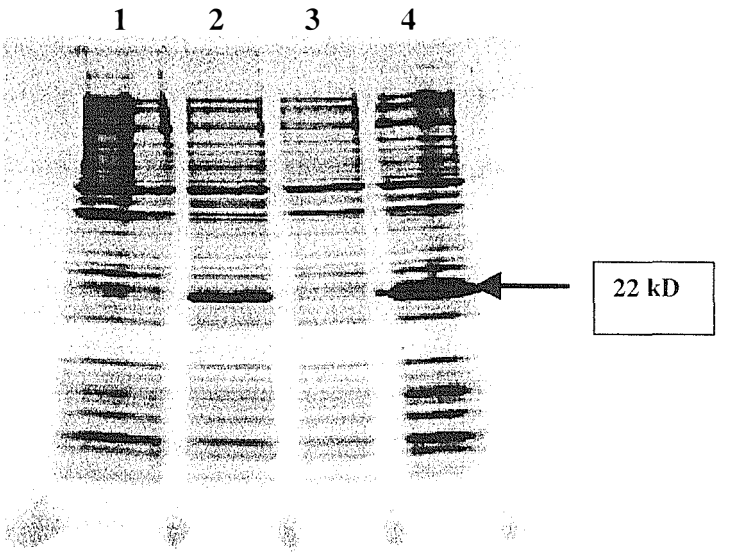

FIGURE 6: Induction of wild-type ComK strains 447 and 448. 5 mL cultures were grown to an O.D. of 0.5 and induced with 100 mM IPTG for 2 hours. Lanes 1 and 2: Uninduced and induced (respectively) of strain 447. Lanes 3 and 4: Uninduced and induced (respectively) of strain 448. ComK is approximately 22 kD in size.

Protein Purification Techniques

Following the successful induction of full-length ComK, research shifted to the purification of ComK. Since its identification as a transcription factor approximately seven years ago, many labs have attempted to purify ComK using relatively new techniques such as His-Tags and maltose binding proteins. However, neither of these methods has produced good results, so this project was focused on more traditional methods of protein purification. The first method for purification that was used was an

ammonium sulfate precipitation. Ammonium sulfate is added in a stepwise manner and proteins are precipitated out of solution based on their solubility in various concentrations of the salt. Ammonium sulfate was initially added to 30% (w/v), 50%, and 75% and the precipitated protein was separated from soluble proteins by centrifugation. Samples from each step were run on a 15% SDS-PAGE gel. The precipitate from the 50% cut contained most of the ComK. To achieve a better purification, another ammonium sulfate precipitation was done to find ammonium sulfate concentrations of 20%, 40% and 50%. These results showed that the ComK was still partially in solution at 40% ammonium sulfate. The 50% pellet containing ComK was resuspended in resuspension buffer as described in Experimental Procedures and a sample of the resuspended pellet, the supernatants at 20% and 40%, and a crude extract sample were run on a 15% SDS-PAGE gel (Figure 7). There was definite enhancement of the protein found in the resuspended 50% pellet and this was used for further purification. At this point the protein is in a concentration of ammonium sulfate that will preclude any binding that involves electrostatic interactions. Therefore the salt must be removed before the application of purification techniques that are based on such binding.

One method in more traditional protein purification is the use of columns that separate protein based on factors such as charge, affinity, specific binding and size. In this project, samples were run on three ion exchange columns and a DNA cellulose column to see if ComK would bind to the columns or come out in the flow-through. The ammonium sulfate product was run on a t-butyl hydrophobic column because any potential hydrophobic interactions between the protein and the column are enhanced by the presence of ammonium sulfate. In addition, a gel filtration column was also run in

order to try to separate the protein out based on size. For each column except for the hydrophobic column, a sample of crude extract was loaded on the column and run as described in Experimental Procedures. Each of the columns has a specific function and it is possible to try to predict the behavior of a protein on a column based on certain information. In the case of ComK, it is likely that the protein binds to DNA in a cationic exchange interaction between the negative charges on DNA and the positive charges that are found in the ComK DNA binding domain. Using this information we began testing a variety of columns.

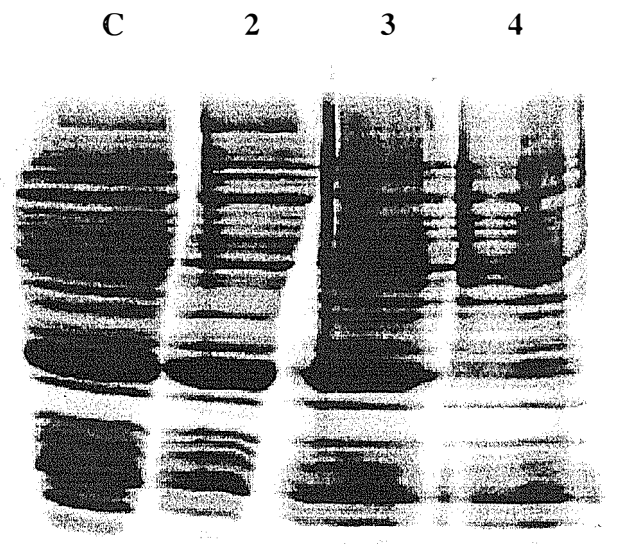

FIGURE 7: Ammonium sulfate precipitation of crude extract of wild-type ComK. Lane C: Crude extract with ComK present at ~22 kD. Lane 2: Resuspended pellet from 50% concentration. Lane 3: Supernatant from 20% saturation. Lane 4: Supernatant from 40% saturation.

Three ion exchange columns were run, a High Q column, a High S column and a CM column. The High Q column contains a strongly basic anionic exchanger. It was not known if ComK would bind to this column, but it is known that High Q columns have a tendency to bind a large number of cellular proteins. As a result, it was predicted that even if the column did not bind ComK, hopefully it would bind many of the other proteins in solution. A 100 μL sample of filtered crude extract was run on a low salt to high salt gradient and selected fractions were run on a 15% SDS-PAGE gel (Figure 8).

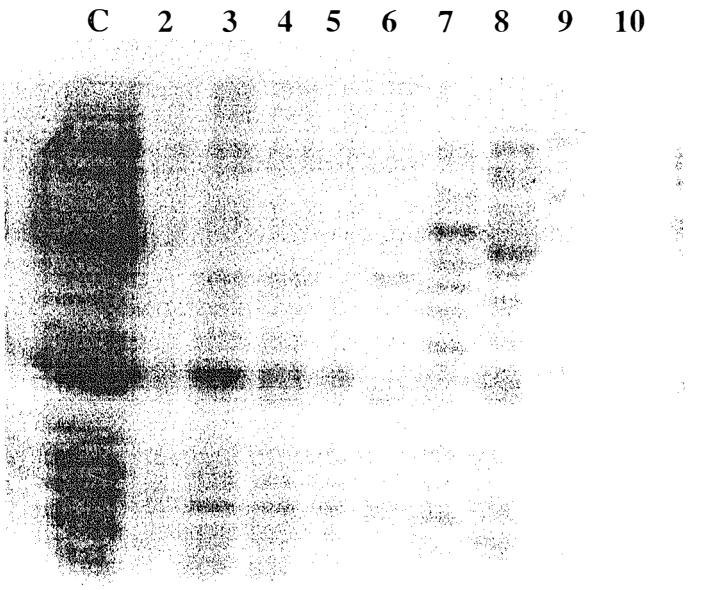

FIGURE 8: High Q fractions. Lane 1: Crude extract. Lanes 2-5: Flow-through from the High Q column with ComK at 22 kD. Lanes 6-10: Fractions from the salt gradient.

ComK eluted in the flow-through fractions and its purification was significantly enhanced. This result shows that there were not strong interactions between the positively charged resin and any exposed negative charges on ComK and that many of the other proteins in solution were bound to the column, as predicted.

The flow-through fractions were combined and subsequently run on a High S column, which is a cationic exchanger, in hopes that the protein would bind to the column in a similar fashion to the way it binds to DNA and then be eluted with the NaCl gradient. There were no significant amounts of protein in any of the fractions from the High S column. This may have been due to the dilution of the protein in the sample volume or that there are not enough cationic exchange interactions on ComK to bind to the column. In addition, the protein may have been bound so tightly to the column that a 1M concentration of NaCl might not have been enough to elute ComK off the column. However, when a 3M concentration of NaCl was added to the column no protein eluted. When crude extract was run on the CM column was run ComK came out in the flow-through without any enhancement of protein purification.

Another column that was tried was a t-butyl hydrophobic FPLC column. This column produces hydrophobic interactions with proteins that bind to the column at high salt concentrations and then be eluted out as the salt gradient decreases. The column was prepared and run as described in Experimental Procedures and fractions were collected and based on absorbance at 280 nm, fractions containing protein peaks were run on a 15% SDS-PAGE gel. None of the fractions had any significant amounts of ComK in them.

Because ComK is a DNA binding protein, a DNA cellulose column was tried. A sample of crude extract was run on the DNA cellulose column however it appeared that the majority of the proteins came out in the flow-through and there was no significant amount of purification of ComK.

Another protein purification technique is electroelution. This technique involves cutting a band of the desired protein out of a 15% SDS-PAGE gel and eluting it out of the gel using an electrophoresis chamber and a special device called an Elutrap. The drawback to this method is that even if you successfully elute the protein out of the gel, it is in a denatured form because of all of the SDS that is bound to the protein. To remove the SDS, guanidinium chloride was added to solution to a concentration of 6M (Doonan, 1996). It is thought that is this situation all of the guanidinium chloride will replace the SDS in the protein. This aids the purification process because guanidinium chloride is much easier to dialyze out of the protein than SDS. The hope is then that as the guanidinium chloride is slowly dialyzed out of the protein and the solution, the protein will renature and return to an active form. In this case, only a small amount of protein was eluted out of the gel and despite slow dialysis in resuspension buffer, it did not appear that a significant amount of ComK was recovered.

After all of these methods were tested and the results showed that only the ammonium precipitation and the High Q column had shown any significant purification of the protein, an additional round of purification was done on a crude extract. A sample of filtered crude extract in a 1 mL volume was run on the High Q column and the flow-through fractions containing a high concentration of partially purified ComK were obtained in 300 μL amounts. From the most concentrated fraction, 100 μL was removed

and an ammonium sulfate precipitation was done. The results from this precipitation were that only a very few additional proteins were separated out from the solution with ComK (Figure 9). Following this precipitation, a 100 μL sample from the High Q flow-through was run on a size exclusion column made of Sephadose beads. It was expected that ComK would move through the column and come out toward the end

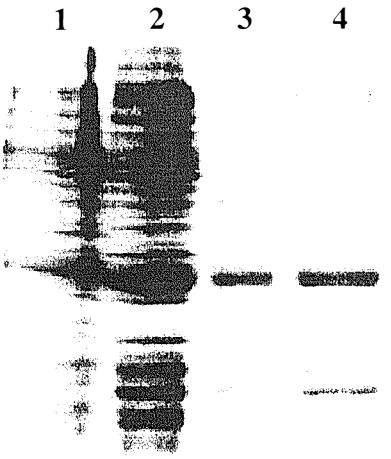

FIGURE 9: Lane 1: Uninduced crude extract. Lane 2: Induced crude extract. Lane 3: High Q flow-through. Lane 4: Ammonium sulfate precipitation of High Q flow-through.

of the run due to its rather small size. However, fractions run on a 15% SDS-PAGE gel showed that ComK came out very early in the column run (Figure 10). Because ComK eluted before proteins with higher molecular weights, ComK must exist primarily as a multimer.

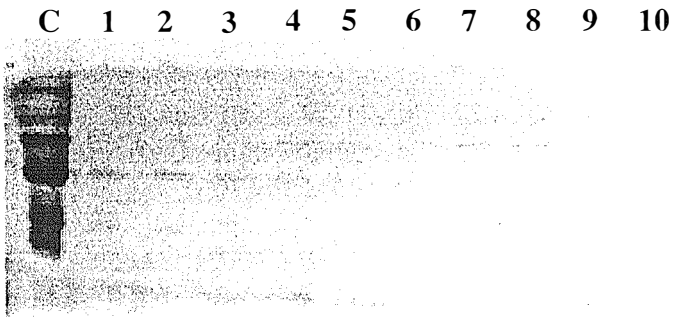

FIGURE 10: Gel filtration column fractions. Lane C: Flow-through from the High Q column. Lanes 1-4: Fractions containing ComK in a well-purified form. Lanes 5-10: Later fractions with no ComK found in them.

Conclusions

ComK was successfully induced with IPTG and partial protein purification was done using a High Q ion exchange column, gel filtration column and ammonium sulfate precipitation. Purification of ComK is almost complete and the next step in this project to begin testing the binding capabilities of purified ComK using FeBABE and targeted footprinting in order to study the interactions between ComK, the α subunit of RNA polymerase and LexA when all three are interacting with DNA at the *recA* promoter site. Further work must also be done with the cloning of the truncated forms of ComK using oligos 186 or 187 and oligo 104. The truncated forms of ComK are needed in order to perform structural analysis on ComK as a whole and specifically on the DNA binding domain.

References

Aiyar, S. E., Gourse, R. L., and Ross, W. (1998) Upstream A-tracts increase prokaryotic promoter activity through interactions with the RNA polymerase alpha subunit. *Proc. Natl. Acad. Sci. USA* **95:** 14652-14657.

Arwert, F., and G. Venema. (1973) Transformation in *Bacillus subtilis.* Fate of newly introduced transforming DNA. *Mol. Gen. Genet.* **123:** 185-198

Blatter, E. E., Ross, W., Tang, H., Gourse, R. L. and Ebright, R. H. (1994) Domain organization of RNA polymerase alpha subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell.* **78:** 889-896.

Cheo, D. L., Bayles, K. W., and Yasbin, R. E. (1993) *J. Bacteriol.* **175:** 5907-5915.

DeHaseth, P. L., Zupanic, M. L., and Record, M. T. (1998) RNA polymerase-promoter interactions: The comings and goings of RNA polymerase. *J. Bacteriol.* **180:** 3019-3025.

Doonan, S. (1996) in *Protein Purification Protocols* (Doonan, S., ed) pp. 135-144, Humana Press, Totowa, NJ.

Dubnau, D. and Lovett, C. M. (2002) Transformation and Recombination, p. 453-471. *In* A. L. Sonenshein (ed.), *Bacillus subtilis* and Its Closest Relatives: From Genes to Cells. ASM Press, Washington, D.C.

Dubnau, D. (1999) DNA uptake in bacteria. *Annu. Rev. Microbiol.* **53:** 214-244.

Dubnau, D. (1976) Genetic transformation of *Bacillus subtilis*: a review with emphasis on the recombination mechanism, p. 14-27. *In* D. Schlessinger (ed.), *Microbiology—1976.* American Society for Microbiology, Washington, D.C.

Dubnau, D., and Cirigliano, C. (1972) Fate of transforming DNA following uptake by competent *Bacillus subtilis.* III. Formation and properties of products isolated from transformed cells which are derived entirely from donor DNA. *J. Mol. Biol.* **64:** 9-29.

Dubnau, D., and Cirigliano, C. (1972) Fate of transforming DNA following uptake by competent *Bacillus subtilis.* IV. The endwise attachment and uptake of transforming DNA. *J. Mol. Biol.* **64:** 31-46.

Estrem, S. T., Gaal, T., Ross, W., and Gourse, R. L. (1998) Identification of an UP Element Consensus Sequence for Bacterial Promoters. *Proc. Natl. Acad. Sci. USA* **95:** 9761-9766.

Estrem, S. T., Ross, W., Gaal, T., Chen, Z. W. S., Niu, W., Ebright, R. H., *et al.* (1999) Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase α subunit. *Genes Dev.* **13:** 2134-2147.

Gourse, R. L., Ross, W., Gaal, T. (2000) UP s and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Mol. Microbiol.* **37:** 687-695.

Gourse, R. L. (1988) Visualization and quantitative analysis of complex formation between *E. coli* RNA polymerase and an rRNA promoter *in vitro*. *Nucleic Acids Res.* **16:** 9789-9809.

Haijema, B. J., Hamoen, L.W., Kooistra, J., Venema, G., and Van Sinderen, D. (1995) Expression of the ATP-dependent deoxyribonuclease of *Bacillus subtilis* is under competence-mediated control. *Mol. Microbiol.* **15:** 203-211.

Haijema, B. J., Van Sinderen, D., Winterling, K., Kooistra, J., Venema, G., and Hamoen, L.W. (1996) Regulated expression of the *dinR* and *recA* genes during competence development and SOS induction in *Bacillus subtilis*. *Mol. Microbiol.* **22:** 75-85.

Hamoen, L. W., Van Werkhoven, A. F., Bijilsma, J. J., Dubnau, D., and Veneman, G. (1998) The competence transcription factor of *Bacillus subtilis* recognizes short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix. *Genes Dev.* **12:** 1539-1550.

Hamoen, L. W., Haijema, B., Bijlsma, J. J., Venema, G. and Lovett, C. M. (2001) The *Bacillus subtilis* Competence Transcription Factor, ComK, Overrides LexA-imposed Transcriptional Inhibition without Physically Displacing LexA. *J. Biol. Chem.* **46:** 42901-42907.

Hanai, R., and Wang, J.C. (1994) *Proc. Natl. Acad. Sci. USA* **91**: 11904-11908.

Hayward, R. S., Igarashi, K., and Ishihama, A. (1991) Functional specialization within the α-subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **221:** 23-29.

Kolb, A., Igarashi, K., Ishihama, A. Lavigne, M., Buckle, M. and Buc, H. (1993) *E. coli* RNA polymerase, deleted in the C-terminal part of its α-subunit, interacts differently with the c.AMP-CRP complex at the *lac*P1 and at the *gal*P1 promoter. *Nucleic Acids Res.* **21:** 319-326.

Lazazzera, B. A., Palmer, T., Quisel, J., and Grossman, A. D. (1999) in *Cell-Cell Signaling in Bacteria* (Dunny, G. M. and Winans, S. C., ed) pp. 27-46, American Society for Microbiology, Washington, D. C.

McAllister, C. F., and Achberger, E. C. (1988) Effects of polyadenine-containing curved DNA on promoters utilization in *Bacillus subtilis*. *J. Biol. Chem.* **263:** 11743-11749.

Miller, M. C., Resnick, J. B., Smith, B. T., and Lovett, C. M. J. (1996) *J. Biol. Chem.* **271**: 33502-33508.

Mohan, S. and Dubnau, D. (1990) Transcriptional regulation of *comC*: evidence for a competence-specific factor in *Bacillus subtilis*. *J. Bacteriol.* **172**: 4064-4071.

Murakami, K., Kimura, M., Owens, J. T., Meares, C. F., and Ishihama, A. (1997) The two α subunits of *Escherichia coli* RNA polymerase are asymmetrically arranged and contact different halves of the DNA upstream element. *Proc. Natl. Acad. Sci.* **94:** 1709-1714.

Murakami, K., Owens, J. T., Belyaeva, T. A., Meares, C. F., Busby, S. J. W. and Ishihama, A. (1997) Positioning of two alpha subunit carboxy-terminal domains of RNA polymerase at promoters by two transcription factors. *Proc. Natl. Acad. Sci.* **94:** 11274-11278.

Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V., and Ebright, R. H. (2000) Structural organization of the RNA polymerase-promoter open complex. *Cell* **101:** 601-611.

Negishi, T., Fujita, N. and Ishihama, A. (1995) *J. Mol. Biol.* 723-728.

Newlands, J. T., Gaal, T., Mecsas, J., and Gourse, R. T. (1993) Transcription of the *Escherichia coli rrnB* promoter by the heat shock RNA polymerase $E\sigma^{32}$ *in vitro*. *J. Bacteriol.* **175:** 661-668.

Persuh, M., Turgay, K., Mandic-Mulec, I., and Dubnau, D. (1999) The N- and C-terminal domains of MecA recognize different partners in the competence molecular switch. *Mol. Microbiol.* **33:** 886-894.

Raymond-Denise, A., and Guillen, N. (1991) *J. Bacteriol.* **173**: 7084-7091.

Ross, W., Ernst, A., and Gourse, R. L. (2001) Fine structure of *E. coli* RNA polyermase-promoter interactions: α subunit binding to the UP element minor groove. *Genes Dev.* **15:** 491-506.

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

Russo, F. D. and Silhavy, T. J. (1992) Alpha: the Cinderella subunit of RNA polymerase. *J. Biol. Chem.* **267:** 14515-14518.

Singh, R. M. (1972) Number of deoxyribonucleic acid uptake sites in competent cells of *Bacillus subtilis*. *J. Bacteriol.* **110**: 266-272.

Smith, H., Wiersma, K., Venema, G., and Bron, S. (1985) Transformation in *Bacillus subtilis*: further characterization of a 75,000-dalton protein complex involved in binding and entry of donor DNA. *J. Bacteriol.* **164**: 201-206.

Solomon, J., Magnuson, R., Srivastava, A., and Grossman, A. D. (1995) Convergent sensing pathways mediate response to two extracellular competence factors in *Bacillus Subtilis*. *Genes Dev.* **9**: 547-558.

Tortosa, P., Logsdon, L., Kraigher, B., Itoh, Y., Mandic-Mulec, I., and Dubnau, D. (2001) Specificity and genetic polymorphism of the Bacillus competence quorum-sensing system. *J. Bacteriol.* **183**: 451-460.

Traviglia, S. L., Datwyler, S. A., Yan, D., Ishihama, A., and Meares, C. F. (1999) Targeted Protein Footprinting: Where Different Transcription Factors Bind to RNA Polymerase. *Biochem.* **38:** 15774-15778.

Traviglia, S. L., Datwyler, S. A., and Meares, C. F. (1999) Mapping Protein-Protein Interactions with a Library of Tethered Cutting Reagents: The Binding Site of σ^{70} on *Escherichia coli* RNA polymerase. *Biochem.* **38**: 4259-4265.

Vagner, V., Claverys, J.-P., Ehrlich, S. D., and Mejean, V. (1990) Direction of DNA entry in competent cells of *Bacillus subtilis*. *Mol. Microbiol.* **4:** 1785-1788.

Van Sinderen, D., Ten Berge, A., Hayema, B. J., Hamoen, L., and Venema, G. (1994). Molecular cloning and sequence of *comK*, a gene required for genetic competence in *Bacillus subtilis*. *Mol. Microbiol.* **11:** 695-703.

Van Sinderen, D., and Venema, G. (1994) ComK acts as an autoregulatory control switch in the signal transduction route to competence in *Bacillus subtilis*. *J. Bacteriol.* **176:** 5762-5770.

Van Sinderen, D., Luttinger, A., Kong, L., Dubnau, D., Venema, G., and Hamoen, L. (1995) *ComK* encodes the competence transcription factor, the key regulatory protein for competence development in *Bacillus subtilis*. *Mol. Microbiol.* **15:** 455-462.

Vosman, B., Kuiken, G., and Venema, G. (1988) Transformation in *Bacillus subtilis*: involvement of the 17-kilo-dalton DNA-entry nuclease and the competence-specific 18 kilo-dalton protein. *J. Bacteriol.* **170:** 3703-3710.

Weinrauch, Y., Penchev, R., Dubnau, E., Smith, I., and Dubnau, D. (1990) A *Bacillus subtilis* regulatory gene product for genetic competence and sporulation resembles sensor protein members of the bacterial two-component signal-transduction systems. *Genes Dev.* **4**: 860-872.

APPENDICES

Standard Procedures

Agarose Gels

1. Measure 0.4 g Agarose into 50 mL of TBE buffer.
2. Melt Agarose in microwave. Add 5 µL of 10 mg/mL Ethidium Bromide, then allow to cool to touch.
3. Set up casting platform, securely sealing the ends with tape. Insert appropriate comb into the casting platform, then pour the gel 0.5-1.0 cm thick.
4. After gel has set, remove comb and tape. Place gel in gel box with buffer. Buffer should cover top of gel.
5. Add loading dye to samples and apply samples to the gel.
6. Attach leads to the gel box. DNA is negatively charged and will migrate toward the anode (positive lead). Set voltage to between 1-10 V per cm of gel length (about 80 for 8-well mini gels, 100 for 14-well mid-size gels, 120-150 for 22-well large gels). Running times is 45 minutes to 1 hour, or until the dye travels 2/3 the length of the gel. DNA is visualized under UV light.

WIZARD MINI-PREP

I. *Preparation of E. coli*

1. Choose at least 15 colonies from transformation plates to inoculate 1-10 mL of LB + carbenicillin (50 µL/mL) for pET21a+.
2. Use a sterile toothpick or inoculating loop to transfer each colony onto the master plate and then to the test tube.

3. Incubate with shaking at 37^0C for approximately 24 hours.

II. *Production of Clear Lysate*

1. Pellet 1-5 mL of bacterial culture by centrifugation for 10 minutes. Pour off supernatant.
2. Add 250 μL of Wizard Cell Resuspension Solution. Completely resuspend pellet by vortexing.
3. Add 350 μL of Wizard Cell Lysis Solution and mix by inverting 4 times. Caution not to vortex. Incubate 1-5 minutes until suspension clears (do not exceed 5 minutes).
4. Add 10 μL of Alkaline Protease solution and mix by inverting 4 times. Again, remember not to vortex. Incubate for maximum 5 minutes.
5. Add 350 μL of Wizard Neutralization solution and mix by inverting 4 times. Remember not to vortex.
6. Centrifuge bacterial lysate for 10 minutes at room temperature.

III. *Spin Down and Purify by Centrifugation*

1. Transfer cleared lysate to prepared and labeled Spin Column. Avoid transferring white precipitate.
2. Centrifuge supernatant in Spin Column for 1 minutes at room temperature. Remove collection tube and discard flow-through. Place column back in tube.
3. Add 750 μL of Wizard Column Wash Solution. Centrifuge for 2 minutes at room temperature. Remove collection tube and discard flow-through. Place column back in tube.
4. Centrifuge columns for 1 minute. Remove collection tubes and place columns in sterile 1.5 mL microcentrifuge tubes.

5. Elute DNA by adding 100 µL of nuclease-free water to column. Centrifuge for 1 minute at room temperature. Store DNA at -20^{0}C.

SDS-PAGE (15%):

Clean small glass plates with ethanol. Use Vaseline at each corner of the plates and thinly coat spacers. Place spacers on plates and form mold. Clamp with binding clips; stand upright and fill with water to test seal. If the seal leaks, repeat the process.

Running Gel:

	1x	4x
dH_2O	1.3 mL	5.2 mL
1.5 M Tris Cl pH 8.8	2.8 mL	11.2 mL
10% SDS	83 µL	332 µL
30% (37:1) acrylamide stock	4.2 mL	16.8 mL
TEMED	2.5 µL	10.0 µL
10% ammonium persulfate	50 µL	200 µL

Pour into mold, stopping 1-2 cm from the top of the mold. Lay a small amount of t-butanol on top. Allow to polymerize and then rinse butanol with ddH_2O.

Stacking Gel:

	1x	4x
dH_2O	2.6 mL	10.4 mL
2 M Tris Cl pH 6.8	115 µL	460 µL
10% SDS	38 µL	152 µL
30% (37:1) acrylamide stock	0.940 mL	3.76 mL
TEMED	2.5 µL	10.0 µL
10% ammonium persulfate	50 µL	200 µL

Pour on top of Running gel, insert comb and allow to polymerize.

Remove bottom spacer and comb from gel plates and mount gel in small gel box. Add SDS Running Buffer to both reservoirs. Samples are prepared in microfuge tubes by mixing with 1/6 volume SDS Loading Buffer and then boiling for 2 minutes, followed by a brief vortex. Samples are loaded and gel run at 100-200 V until dye reaches the bottom of the gel. After electrophoresis, stain gel with Coomasie Brilliant Blue for 20-40 minutes with gentle shaking. Cover with Destain and shake gently overnight to destain.

RECIPES

Coomasie Stain:

Mix: 2.5 g Coomasie Brilliant Blue
450 mL Methanol
100 mL Acetic Acid
450 mL ddH_2O

Destain:

Mix: 300 mL Methanol
400 mL Acetic Acid
Bring to 4 L with ddH_2O

LB Broth and Plates:

Per 1liter mix:

10 g Tryptone Peptone
5 g Yeast Extract
10 g NaCl
15 g Agar (for plates only)
Bring to 1L with ddH_2O, autoclave.
Cool to 60^0C, add antibiotic and pour.

SDS Running Buffer:

Mix: 12.1 g Tris
57.6 g Glycine
4 g SDS
Bring to 4 L with ddH_2O

5x TBE

Mix: 54 g Tris (base)
27.5 g Boric Acid
20 mL 0.5 M EDTA, pH 8.0
Bring to 1 L with ddH_2O

5x TG

Mix: 30.28 g Tris
142.7 g Glycine
3.92 g EDTA
Bring to 1 L with ddH_2O

Resuspension Buffer

Mix: 20 mL Tris Cl, pH 8.0

0.2 mL EDTA, pH 8.0

2.922 g NaCl

Bring to 1 L with ddH_2O