

FitA and FitB, a *Neisseria gonorrhoeae* Protein Complex Involved in the Regulation of
Transcellular Migration

by

John Scott Wilbur

A DISSERTATION

Presented to the Department of Molecular Microbiology and Immunology
and the Oregon Health and Sciences University

School of Medicine

in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

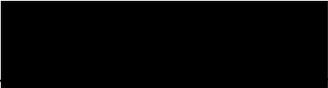
March 2006

School of Medicine
Oregon Health and Sciences University

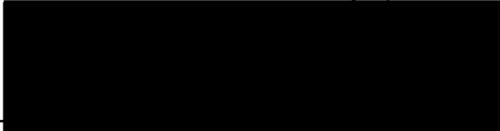
CERTIFICATE OF APPROVAL

This is to certify that the Ph.D. Thesis of
John Scott Wilbur
has been approved


Dr. Magdalene So, Thesis Advisor


Dr. Fred Heffron, Member


Dr. Jay Mellies, Member


Dr. Jorge Crosa, Member

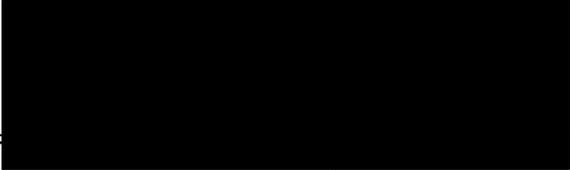


TABLE OF CONTENTS

Table of contents	<i>i</i>
Acknowledgments	<i>iv</i>
Abstract	<i>v</i>
Chapter 1: Introduction	1
1.0 Neisseria	2
1.1 Clinical manifestation	2
1.2 The GC infection process	3
1.2.1 Loose adherence and microcolony formation	3
1.2.2 Tight adherence and invasion	4
1.2.3 Intracellular replication and transcytosis	5
1.3 The role of transcytosis in infection	6
1.4 Antigenic variation of GC virulence factors	7
1.4.1 <i>PilE</i> recombination with <i>pilS</i> locus	7
1.4.2 Opa phase variation	9
1.5 Tissue culture models of infection	10
1.6 Identification of the fit locus	10
1.7 FitA and ribbon-helix-helix DNA binding proteins	11
1.8 FitB and PIN domains	12

1.9 Present work	14
References	17
Chapter 1 Figure legends	29
Chapter 1 Figures	33
Chapter 2: Manuscript I: <i>Neisseria gonorrhoeae</i> FitA interacts with FitB to bind DNA through its ribbon-helix-helix motif	36
Abstract	39
Introduction	40
Experimental procedures	42
Results	47
Discussion	54
References	55
Tables 1-3	59
Figure legend	62
Figures 1-6	64
Chapter 3: Manuscript II: Structure of FitAB from <i>Neisseria gonorrhoeae</i> bound to DNA reveals a heterodimer of PIN and ribbon-helix-helix domains	69
Abstract	71
Introduction	72
Results and Discussion	75
Materials and Methods	83

References	88
Table and Figure Legends	94
Figure 1-5	98
Table 1	103
Chapter 4: Additional Results: examination of the relationship between the <i>fit</i> locus and <i>pilE</i> recombination	104
Results and discussion	105
References	109
Figure legend	113
Figure 1-2	114
Table 1	116
Chapter 5: Conclusions	117
1.1 Opa expression drives transcytosis	118
1.2 <i>PilE</i> recombination slows transcytosis rate	119
1.3 FitA functions to abrogate FitB activity	121
1.4 FitB is responsible for the fast trafficking phenotype	122
1.5 Final Remarks	124
References	124
Figure legends	128
Figure 1-2	130
Table 1	132

Acknowledgments

It has been my honor and pleasure to have spent the last ten years working for Dr. Maggie So. She hired me as an undergraduate dishwasher and has encouraged and mentored me as research assistant and then as a graduate student. Throughout my years working with her, Maggie has given me tremendous freedom to explore whatever crazy idea or blind ally I wished to explore.

I would also like to acknowledge my many mentors, collaborators, teachers and friends who have taught me so much about how this world we live in works. Specifically I need to thank Sunghee Chai for teaching me most of what I know about bench work. Brian Ahmer and Alex Merz for teaching me the joys of the hidden world a bacteria. And my three favorite, and worst role models Dr. Fred Heffron, Dr. Jorge Crosa, and Dr. Richard Brennan also need to be mentioned, as they were always willing to tell me to keep at it in my moments of greatest disappointment and despair.

I would also like to thank all of my friends and coworkers over the many years I have spent on the sixth floor of CROET. So many have come and gone that I cannot name them all. I thank you for your patents with me and your companionship throughout. I must mention Shaun Lee, George Neimann, Duncan Parson, Brian Ahmer, Jason Larson, Dustin Higashi, Patty Ayalla and Rob Pearson.

Finally I must thank my parents, David and Linda Wilbur. I would like to dedicate this work to Hettie-Ray Wilbur who never had a doubt. Thank you for listening.

Abstract

Neisseria gonorrhoeae (GC) is the causative agent of the sexually transmitted disease gonorrhea. This gram-negative diplococcus is responsible for over 700,000 infections in the US in 2006. In this work I characterize the GC *fit* locus. Mutants in this locus increased the speed of bacterial transcytosis across polarized monolayers of human cells. The *fit* locus encodes two proteins, FitA and FitB. FitA is a DNA binding protein that contains a ribbon-helix-helix DNA binding motif (RHH). FitA forms a heterooctamer with FitB and binds the DNA sequence TGCTATCA (FitID) found in the *fit* promoter region. The FitA/B heterooctamer binds the FitID sequence with 38-fold higher affinity than FitA alone. This is the first reported interaction between a RHH protein and another protein that caused an increase in DNA binding. The 3D crystal structure of the FitA/B heterooctamer bound to DNA confirmed that FitA was a RHH protein and revealed that FitB had a putative PIN domain. This structure also reveals that arginine 68 of FitA inserts into the putative PIN domain active site of FitB. In addition to increasing the rate of transcytosis, I also found the *fit* mutant was also deficient in *pilE/S* recombination during infection. Based upon the above findings, I propose that the function of the *fit* locus is to stimulate this recombination event during transcytosis as a means of host immune evasion and this activity is responsible for the observed trafficking and replication phenotypes.

Chapter 1: Introduction

1.0 *Neisseria gonorrhoeae*

Neisseria gonorrhoeae (gonococcus, GC), a gram-negative bacterium, is the etiological agent of the sexually transmitted disease (STD) gonorrhea. In 2004, there were 680,000 reported GC infections in the US, making it the second most common sexually transmitted disease in the nation (CDC, www.cdc.gov/std). The high prevalence of gonorrhea is surprising in light of the fact that the infection is easily cleared with a short regimen of antibiotics. Given the ready availability of healthcare in the US, why is this bacterium able to persist and thrive in human society? The success of this organism is due in no small part to its ability to create and maintain asymptomatic infections that are referred to as the carrier state.

1.1 Clinical manifestation

Gonococcal infections can manifest themselves in a number of different ways depending on the sex of the individual and the tissue infected. In males, infections are usually localized to the urethral epithelial cells, although infections of the rectal and pharyngeal epithelial cells also occur (13, 29, 30). Urethritis in men usually presents as a purulent discharge five to seven days after initial exposure, and is caused by localized inflammation and the recruitment of neutrophils.

In females, GC infections are more complicated due to the number of different tissues that can be infected. As observed in males, GC infects the female urethral, pharyngeal and rectal epithelia (13), though females can also present with cervicitis and pelvic inflammatory disease (PID), depending on how much of the female genital tract is infected. GC infections are the leading cause of sterility among women in the US (CDC).

Disseminated gonococcal infections (DGI), where the bacteria infect tissues beyond the urogenital tract, are rare but are seen in both sexes, and can lead to arthritis, septicemia and meningitis (9, 43).

The asymptomatic carrier also plays a role in person-to-person transmission of the disease. A study of people 18-35 years of age in Baltimore, MD found that 5.3% of the population carried *Neisseria gonorrhoeae* without exhibiting overt symptoms of disease (83). Of those surveyed, 1.5% reported having had symptomatic gonorrheal disease in the past 12 months, but had sought treatment and were cleared of the infection. This study found that there are 3.5 times as many carriers of GC as there are people who show overt symptoms of gonorrhea. This makes sense from the point of view of the pathogen, as it will be more successful if it can be transmitted from individual to individual without causing inflammation. A host response, such as inflammation of the genital tract, is likely to inhibit further passage of any STD.

1.2 The GC infection process

The process of GC infection of human cells can be broken down into six major steps (Figure 1; (54): loose adherence, microcolony formation, tight adherence, invasion, intracellular replication and transcytosis or egress. These stages of GC infection are defined both by timing after initial contact and the bacterial and host factors involved.

1.2.1 Loose adherence and microcolony formation

The first contact between bacterial and host cells is termed loose adherence and is promoted by interactions between the bacterial type IV pilus (Tfp) and host cell proteins

(47, 56, 82). One example of this interaction is between the minor pilin subunit PilC and CD46 (65). This interaction has been implicated in activation of host cell signaling cascades (41), the trigger of a host cytosolic calcium flux and endosome exocytosis (2). By three to four hours after initial attachment of the bacteria to the host cell, the bacteria organize into microcolonies on the cell surface (51, 53). This process involves the type IV pilus and is dependent on pilus retraction. The microcolony is thought to consist of between 10-100 bacteria evenly spaced into a raspberry shaped complex, though larger microcollonies are observed at later time points (18, 54), Merz, Lee and So unpublished observations). Though the exact function of the microcolony is unknown, it has been observed that the region of the host cell beneath the microcolony, called a cortical plaque, is highly enriched in a number of host cell proteins (51). These processes have been well characterized and reviewed (13, 54).

1.2.2 Tight adherence and invasion

Between 6 and 8 hours after initial attachment the microcolonies disperse, the pili are thought to be shed and the bacteria form a monolayer on the host cell surface (54, 58). At this point the bacteria and host cell membranes are found in close proximity. It has been proposed that this interaction is governed by interactions between the bacterial Opa adhesin and host cell proteins CEACAM (CEA-related cell adhesion molecule or CD66) and/or the Heparan sulphate proteoglycan receptor (HSPG; (24, 55, 84). The bacteria then enter the host cells in a process similar to macropinocytosis (23, 88).

A number of different mechanisms have been proposed to explain how GC invade epithelial cells. The first of these involves interactions between one of the Opa proteins

and its receptor, either CEACAM or HSPG (11, 12, 24). The second involves interactions between bacterial lipooligosaccharide (LOS) and the host asialoglycoprotein receptor (31). The third method of entry is thought to be pilus (Tfp) dependent, and presumably involving the pilus binding to CD46 (46, 65). In none of these cases is it known how receptor activation by the bacterial ligand leads to invasion. Also unclear is the relative importance of these interactions to entry, but the bacteria used in this thesis were piliated, non-Opa expressing with an undefined LOS profile, so the interactions we present here may be the result of Tfp/CD46 mediated entry.

1.2.3 Intracellular replication and transcytosis

Once the bacteria enter the host cell, they traffic from a gentamicin-sensitive compartment to a gentamicin-resistant compartment. Because gentamicin is an antibiotic that does not readily cross eukaryotic membranes (66), we surmise that the compartment containing gentamicin-resistant GC is isolated from the endocytic pathway. Beyond that, the intracellular environment of GC is not defined. Once inside the gentamicin-resistant compartment GC is able to replicate (42, 52). Intracellular replication has been observed in polarized and non-polarized cells as well as cells in fallopian tube explants (5, 22, 42). Approximately 18-36 hours after apical infection of polarized cells GC can be recovered in the basal media (34, 35, 37, 52, 87). This process of crossing monolayers is known as transcytosis and does not disrupt the tight junctions or cell viability. The bacteria do not pass between the cells of the monolayer (52). Intracellular GC has also been speculated to exit the apical membrane, but there is little evidence to support this activity.

It has been reported that expression of any one of the 11 Opa proteins promotes transcytosis of both GC and *E. coli* (87). Scanning EM studies also found that intracellular GC expressed Opa (22). As mentioned above, Opa have been shown to interact with different receptors on the apical surface and this interaction promotes entry into the cell. What role, if any, these interactions play in directing GC transcytosis is still unclear.

1.3 The role of transcytosis in infection

The primary symptoms of a gonorrhea infection are inflammation and the recruitment of neutrophils. Both of these symptoms are appear 3 to 7 days after initial exposure to the bacteria and are concurrent with localization of the bacteria to the basolaminar space (29, 45). This observation suggests transcytosis of host epithelial cells plays an important part in disease progression.

The process of transcytosis may also play a role in GC antigenic and phase variation. Ilver et al. found that 100% of GC that had crossed a polarized monolayer of Hec-1B cells had undergone *pilE* antigenic variation and 50% had undergone *pilC* phase variation. In a human challenge study, GC that were recovered after the onset of urethritis had also undergone phase and antigenic variation of the outer membrane proteins PilC and PilE, respectively (28). These two studies suggest a connection between the regulation of transcytosis and phase and antigenic variation.

1.4 Antigenic variation of GC virulence factors

GC encodes many virulence factors whose expression undergoes phase or antigenic variation. Varying the antigenic structure of surface proteins is hypothesized to help bacteria avoid the immune response. Two of the most important genes involved in GC-host cell interactions are *pilE* and *opa* are an example of this (75, 80). The type IV pilus is a filamentous structure that is involved in adhesion to host cells, natural competence and twitching motility (54, 63, 64). The major structural subunit of the pilus is encoded by *pilE*. This gene undergoes high frequency antigenic variation (AV) (17, 36, 75).

1.4.1 *PilE* recombination with *pilS* loci

When GC are grown in GCB (Gibco) media, *pilE* undergoes high-frequency AV at a rate between 10^{-4} and 10^{-3} event/cell/generation (76). *PilE* AV is a RecA-dependent gene conversion event between *pilE* and one of the non-transcribed *pilS* (silent pilin locus) pseudo-genes (Figure 2A) (36). The gonococcal strain FA1090 has one actively transcribed *pilE* locus. There are also 19 different copies of *pilS* distributed among 5 loci in the genome. The *pilS* alleles are truncated at their 5' end, do not encode a functional pilin, are not transcribed and do not directly contribute to the formation of the type IV pilus. The *pilS* loci function as a reservoir of sequences with which the actively transcribed *pilE* locus recombines. Recombination between these loci can occur at multiple sites, leading to an almost limitless number of possible *pilE* sequences.

The *pilE* gene has three domains, including a 5' constant region, a central semi-variable region (SV), and a 3' hypervariable region (HV_L and HV_T, Figure 2) (36, 39, 74).

The exchange of DNA or recombination between *pilE* and *pilS* occurs in regions that share sequence homology, e.g. *cys1*, *cys2* or downstream of the ORF. This process is a gene conversion event because the *pilS* sequence is inserted into the *pilE* locus without a reciprocal exchange of *pilE* sequence into the *pilS* locus.

PilE recombination is dependent on many different components of the DNA repair and recombination pathways. The GC RecA homologue is required for all homologous recombination (40), including *pilE/pilS* recombination. RecX is also required for efficient recombination (81), as are members of the RecF-like recombination pathway, including *recJ*, *recO*, and *recQ* (33, 50, 77).

Efficient *pilE* antigenic variation also requires a number of *cis* elements. Within the *pilE* and *pilS* sequences are a number of invariant sequences thought to be involved in homologous pairing during recombination. There are also four *cis*-acting sequences outside the *pilE* open reading frame, including the two *Sma/Cla* repeats and the RS1 and RS2 upstream repeat sequences (Figure 2e) (73, 86). These putative control elements do not effect the transcription of *pilE* but appear to be required for efficient AV. They are not, however, involved in generalized recombination and DNA repair (73).

There is some controversy about the process by which the *pilS* sequences are donated for recombination with *pilE*. Usually recombination occurs between two sister chromosomes during replication. Thus, it is likely that this is when *pilE* AV occurs. However, GC has an additional reservoir of DNA that can participate in recombination. GC is naturally competent for DNA uptake, that is, the bacteria can import neisserial DNA from its surroundings and recombines it with homologous sequences of its chromosome with high frequency (17, 32, 67). GC is prone to autolysis and some strains

have a type IV secretion system that can export chromosomal DNA into the surroundings (26, 27). These two pools of “free” DNA can potentially function as the donor of *pilS* sequences for *pilE* AV (17, 68). It is not clear at this time to what extent either of these two DNA reservoirs contributes to AV.

1.4.2 *Opa* phase variation

Another group of adhesins on the bacterial surface are the Opa proteins. There are up to 11 different *opa* genes in the GC genome and they have been shown to play a role in complement resistance and adhesion to host cells (6, 24, 55). Each Opa binds to either one of the CEACAM receptors or the HSPG receptor (25, 84, 85). These interactions are primarily thought to be involved in tight adhesion to host cells but they have also been shown to play a role in invasion, host cell signaling and transcytosis. Expression of Opa is phase variable due to a mechanism known as slip-strand misrepair (79).

Slip-strand misrepair occurs within the coding region of each *opa* gene at a series of repeated five base pairs (pentameric repeats). During replication of this region, mistakes can be made such that the new and old strands of DNA do not align properly. When this misalignment is repaired each daughter cell will have a different number of pentameric repeats within the given *opa*, changing its reading frame. Only one of the possible reading frames encodes a functional protein, the other two reading frames encode a protein that is truncated and/or non-functional. Phase variation occurs at a rate of 10^{-3} per generation (3, 79).

1.5 Tissue culture models of infection

Humans are the only known host of *Neisseria gonorrhoeae*, and this is presumed to be due to its specificity for human receptors and nutritional requirements (49, 54). This fact complicates the study of the infection process, as there is no good animal model that recapitulates the process of colonization and disease progression (38). It is therefore necessary to rely on the infection of tissue explants and immortalized cell lines to understand the myriad interactions between GC and its host. In the work described in this thesis two cell lines were used to study GC-host cell interactions: polarized T84 epithelial cells and A431 epithelial cells. A431 cells are a human cell line derived from an epidermoid carcinoma of the vulva of an 85-year old female (ATCC). These cells are useful for studies of GC adhesion, invasion and intracellular behaviors as they produce little mucus and are thus ideal for differentiating adhered bacteria from intracellular populations (35, 42). T84 epithelial cells are derived from of colonic carcinoma origin and polarize when seeded on transwell filters (35, 44, 52) (Millipore, Costar). Polarized cells form tight junctions with each other, and have proteins partitioned to apical and basolateral sites similar to those in native epithelial cells in the mucosa. They have therefore been useful in studying bacterial transcytosis or traversal across epithelial cells (44, 52).

1.6 Identification of the Fit locus

In infection of polarized T84 cells and human cervical explanted tissues, GC are seen to traffic across monolayers in 24-36 hours (47, 48, 52). The closely related pathogen *Neisseria meningitidis* (96% sequence identity) is able to cross polarized

monolayers in six to eight hours (52). Given this temporal difference it was hypothesized that GC would have genes that actively slowed the rate of transcytosis. Hopper et al. 2000 undertook a screen of transposon mutants to identify GC mutations that enabled the bacteria to cross polarized monolayers of host cells more quickly than their wildtype parents. This screen identified four transposon insertions that increased the speed of transcytosis. Two of these are in a locus called *fit* (fast intracellular trafficking). Interestingly, the *fit* mutants not only crossed the monolayer more quickly than wild type GC, they also divided more quickly within cells (35). It is unclear whether these two seemingly disparate phenotypes are functionally related.

The *fit* locus is predicted to encode two proteins, FitA and FitB (Figure 3). In one *fit* mutant the transposon has inserted into the putative *fitB* ORF and in the other the transposon has inserted into the predicted *fitA* promoter region (Figure 3a). Homologues of *fitA* and *fitB* are found in a wide variety of bacteria ranging from human pathogens to plant symbionts to archaea. These homologues are almost always observed with the same bicistronic arrangement as *fit*: the last nucleotide of the *fitA* stop codon is the first nucleotide of the *fitB* start codon. Many *fitA/B* homologues are encoded on plasmids, but GC *fitA* and *fitB* are chromosomal.

1.7 FitA and ribbon-helix-helix DNA binding proteins

FitA is predicted to be a 78 amino acid protein with a molecular weight of 8.6 kDa. Secondary structural predictions suggest a helical conformation for the first 50 amino acids similar to those found in ribbon-helix-helix (RHH) DNA binding proteins (Manuscript I, Figure 1). Of the many FitA homologues, the only protein with an

assigned function is MvpA (32% similarity to FitA), which is encoded on the *Shigella flexneri* plasmid pRS4000 and is just upstream of the FitB homologue MvpT (60, 70). MvpA has been reported to function as the “antidote” in a toxin-antidote plasmid stability locus. When MvpA (*fitA* homologue) and MvpT (*fitB* homologue) are both expressed, the bacterium is viable. If MvpT is expressed in the absence of MvpA, the cells do not divide, though they do recover and divide if expression of MvpT is stopped. Expression of MvpA alone had no effect on bacterial viability. In GC we see the same transcytosis phenotype for the *fitAB* mutant (A9, Figure 2) and *fitB* mutant (A1, Figure 2), supporting the idea that the phenotype is dependent on the presence of *fitB* (35). The *fit* locus mutations do not affect the stability of the GC cryptic plasmid (35).

The RHH motif is a well characterized DNA binding motif, exemplified by the bacterial repressors Arc, Mnt, CopG, MetJ, NikR and ParG (7, 19-21, 57, 62, 69, 78). These proteins form dimers and bind DNA in a sequence specific manner. Arc, Mnt and CopG bind DNA alone, while MetJ and NikR show enhanced binding to DNA in the presence of small molecular cofactor (s-adenosylmethionine and nickel, respectively) (8, 57). ParG, in contrast, binds DNA but its binding affinity is decreased in the presence of ParF (19).

RHH proteins are a family of small DNA binding proteins that have a structure consisting of a short β -strand followed by two α -helices (57, 69). The α -helices are primarily involved in dimerization. RHH proteins form dimers in solution and the formation of the dimer is required for proper folding and DNA binding activity (69). The antiparallel β -strands of the dimer contact the major groove of DNA in a sequence-specific manner. The RHH motif can be differentiated from other similar DNA binding

motifs by substitution mutation of polar residues in the β -strand. This type of mutations will abrogate sequence specific DNA binding without affecting dimer formation if the protein has an RHH motif (61, 62, 71).

RHH proteins bind to DNA in a sequence specific manner at inverted repeats. The inverted repeats consist of two 6-8 base pair (bp) recognition sequences or halfsites that are spaced by an intervening sequence ranging from 1 to 16 bp (72). Many RHH proteins are autoregulatory. Therefore, RHH binding sites are often discovered within their promoter region (72).

1.8 FitB and PIN domains

FitB has a predicted length of 189 amino acids and a molecular weight of 14 kDa. It is a highly hydrophobic protein with a predicted PIN domain motif (see below, Pfam - PF01850, www.ncbi.nlm.nih.gov/Structure). FitB homologues are found in a wide variety of bacteria, including pathogens, plant symbionts and extremophiles. Each of these homologues have a PIN domain motif and a *fitA* homologue. The best characterized FitB homologue is *Shigella flexneri* MvpT (34% similarity, 20% identity) (70). This protein is found on the plasmid pRS4000 and is reported to be the toxin in a toxin-antitoxin plasmid stability locus; see above. Two other PIN domain motif-containing proteins also have this toxin-antitoxin relationship with a RHH motif containing protein. *Salmonella dublin* VagD, the archetypal PIN domain motif protein (59), and *Leptospira interrogans* VapC (89) were shown to inhibit growth when expressed in the absence of their “antitoxins” (VagC and VapB respectively).

The function of the PIN domain is not well understood, but a number of PIN domain containing proteins have been described. 1,486 proteins in both prokaryotes and eukaryotes are predicted to have PIN domain motifs (www.ncbi.nlm.nih.gov/Structure). 3D crystal structural studies of PAE2754, from the thermophile *Pyobaculum aerophilum* revealed the PIN domain to contain a five β -strand pocket with four highly conserved acidic residues surrounding a Mg^{++} ion (1). This structure is similar to the nuclease domains of *Taq* polymerase, phage T4 RNase, and 5'-3' flap endonuclease (1). PAE2754 has weak nuclease activity on double stranded DNA at 37° C. Another PIN domain-containing protein, the *Saccharomyces cerevisiae* Nob1p, is involved in 20s pre-rRNA cleavage and it acts as a D-site endonuclease (14). These data, taken together, suggest that proteins with the PIN domain motif will have RNase or DNase activity.

1.9 Present work

When I began my thesis work, *fit* was identified as a GC locus that controlled the seemingly disparate activities of transcytosis and intracellular growth. Sequence analysis suggested that this locus encode two proteins, FitA and FitB. My thesis project was to prove that the *fit* locus encoded these two proteins and to examine their biochemical properties. The aim of this work was to gain a better understanding of how this locus affects transcytosis and intracellular growth and thus gain a better understanding of the intracellular life of *Neisseria gonorrhoeae*.

In Manuscript I, I reported that FitA is a DNA binding protein. I found that FitA binds an inverted repeat FitIS (FitA/B interacting sequence) just upstream of the *fitA* open reading frame (ORF). Sequence predictions suggested that FitA has a ribbon-helix-

helix DNA binding motif and I confirmed the presence of this motif by mutational analysis. I also found that FitA and FitB formed a heterooctamer complex with a stoichiometry of 4 FitA and 4 FitB molecules. The complex binds to the same inverted repeat with a 38-fold higher affinity than FitA dimers. FitB is the first protein shown to increase the DNA binding affinity of a RHH containing protein (FitA).

In Manuscript II, Mattison et al. solved the 3D crystal structure of the FitA/B heterooctamer bound to a 36 bp fragment of DNA. This crystal structure confirmed that FitA made specific contact with FitIS via a RHH motif. This structure also confirmed the prediction that FitB contained a PIN domain. The majority of this work was done by Kirsten Mattison. She crystallized FitA and FitB and she solved their 3D crystal structure when bound to the 36 base pair DNA fragment. My contribution to this work was the initial cloning and I found conditions to express and purify the proteins. I also identified the 38 bp DNA fragment bound by the complex and determined the conditions and thermal dynamics of binding. We collaborated intellectually throughout this work, interpreting the results of each experiment, and generated our model of FitA/B function.

We were unable to demonstrate any nuclease activity from the FitA/B heterooctamer. Interestingly, the crystal structure revealed that an arginine residue (arg68) from FitA appears in the putative active site of the FitB PIN domain. If FitB has nuclease activity, its activity would be inhibited by the positioning of the FitA arginine. These data recall the interaction between the *S. flexneri* MvpA and MvpT, where MvpA (the FitA homologue) inhibits the activity of MvpT, a homologue of FitB and putative PIN domain-containing protein (70). This led us to our working model that FitA and FitB interact with each other at its own promoter to repress *fit* transcription as well as to

inhibit FitB nuclease activity. We also propose that, by some unknown mechanism, FitA and FitB separate when the bacteria is within host cells, freeing up FitB nuclease activity and subsequently resulting in the phenotypes of slow intracellular growth and slow trafficking across host epithelial monolayers.

In our working model we have attempted to connect the intracellular growth and fast trafficking phenotypes of the *fit* mutants. We propose that they are the consequence of FitB nuclease activity. I predict that FitB will have nuclease activity when the bacteria are within host cells. When DNA damage occurs in *E. coli*, replication is halted by the activated SOS pathway (15, 16). Perhaps FitB slows intracellular growth by causing single or double stranded breaks. Ilver et al reported that GC undergoes *pilE* antigenic variation during transcytosis. This recombination would cause breaks in the DNA, halting replication much like the SOS response. GC does not have a classic SOS cascade though it is reasonable to predict it contains a similar response to DNA damage in order to maintain sequence fidelity during replication (4, 10). Thus, if the *fit* locus is involved in the stimulation of intracellular *pilE* AV, it would make sense that the mutant would have a increased replication phenotype within cells.

In chapter four of this thesis I describe experiments that examined the relationship between the *fit* mutant and *pilE* AV during transcytosis. I found that 100% of wildtype GC underwent *pilE* AV during transcytosis of T84 monolayers. In contrast the *fit* mutant did not undergo *pilE* AV at all during transcytosis. This indicates a strong relationship between the activity of the *fit* locus and *pilE* AV. These results were written in the form of a note in the hope that with a few more expected results the study can be submitted as a short manuscript.

In the final chapter, I discuss what these results suggest about the function(s) of the *fit* locus in GC pathogenesis. I hypothesize that FitA and FitB play an important role in the establishment of the carrier state by slowing transcytosis and stimulating antigenic variation.

1. **Arcus, V. L., K. Backbro, A. Roos, E. L. Daniel, and E. N. Baker.** 2004. Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J Biol Chem* **279**:16471-8.
2. **Ayala, B. P., B. Vasquez, S. Clary, J. A. Tainer, K. Rodland, and M. So.** 2001. The pilus-induced Ca²⁺ flux triggers lysosome exocytosis and increases the amount of Lamp1 accessible to Neisseria IgA1 protease. *Cell Microbiol* **3**:265-75.
3. **Belland, R. J., S. G. Morrison, J. H. Carlson, and D. M. Hogan.** 1997. Promoter strength influences phase variation of neisserial opa genes. *Mol Microbiol* **23**:123-35.
4. **Black, C. G., J. A. Fyfe, and J. K. Davies.** 1998. Absence of an SOS-like system in *Neisseria gonorrhoeae*. *Gene* **208**:61-6.
5. **Casey, S., D. Veale, and H. Smith.** 1979. Demonstration of Intracellular Growth of Gonococci in Human Phagocytes using Spectinomycin to Kill Extracellular Organisms. *J. Gen. Microbiol.* **113**:395-398.
6. **Chen, T., J. Swanson, J. Wilson, and R. J. Belland.** 1995. Heparin protects Opa⁺ *Neisseria gonorrhoeae* from the bactericidal action of normal human serum. *Infect Immun* **63**:1790-5.

7. **Chivers, P. T., and R. T. Sauer.** 1999. NikR is a ribbon-helix-helix DNA-binding protein. *Protein Sci* **8**:2494-500.
8. **Chivers, P. T., and R. T. Sauer.** 2000. Regulation of high affinity nickel uptake in bacteria. Ni²⁺-Dependent interaction of NikR with wild-type and mutant operator sites. *J Biol Chem* **275**:19735-41.
9. **Cucurull, E., and L. R. Espinoza.** 1998. Gonococcal arthritis. *Rheum Dis Clin North Am* **24**:305-22.
10. **Davidson, T., and T. Tonjum.** 2006. Meningococcal genome dynamics. *Nat Rev Microbiol* **4**:11-22.
11. **Dehio, C., E. Freissler, C. Lanz, O. G. Gomez-Duarte, G. David, and T. F. Meyer.** 1998. Ligation of cell surface heparan sulfate proteoglycans by antibody-coated beads stimulates phagocytic uptake into epithelial cells: a model for cellular invasion by *Neisseria gonorrhoeae*. *Exp Cell Res* **242**:528-39.
12. **Dehio, C., S. D. Gray-Owen, and T. F. Meyer.** 1998. The role of neisserial Opa proteins in interactions with host cells. *Trends Microbiol.* **6**:489-495.
13. **Edwards, J. L., and M. A. Apicella.** 2004. The molecular mechanisms used by *Neisseria gonorrhoeae* to initiate infection differ between men and women. *Clin Microbiol Rev* **17**:965-81, table of contents.
14. **Fatica, A., D. Tollervey, and M. Dlakic.** 2004. PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA. *Rna* **10**:1698-701.
15. **Foster, P. L.** 2005. Stress responses and genetic variation in bacteria. *Mutat Res* **569**:3-11.

16. **Fry, R. C., T. J. Begley, and L. D. Samson.** 2005. Genome-wide responses to DNA-damaging agents. *Annu Rev Microbiol* **59**:357-77.
17. **Gibbs, C., B. Reimann, E. Schultz, A. Kaufmann, R. Haas, and T. Meyer.** 1989. Reassortment of pilin genes in *Neisseria gonorrhoeae* occurs by two distinct mechanisms. *Nature* **338**:651-652.
18. **Gill, D. B., D. Spitzer, M. Koomey, J. E. Heuser, and J. P. Atkinson.** 2005. Release of host-derived membrane vesicles following pilus-mediated adhesion of *Neisseria gonorrhoeae*. *Cell Microbiol* **7**:1672-83.
19. **Golovanov, A. P., D. Barilla, M. Golovanova, F. Hayes, and L. Y. Lian.** 2003. ParG, a protein required for active partition of bacterial plasmids, has a dimeric ribbon-helix-helix structure. *Mol Microbiol* **50**:1141-53.
20. **Gomis-Ruth, F. X., M. Sola, P. Acebo, A. Parraga, A. Guasch, R. Eritja, A. Gonzalez, M. Espinosa, G. del Solar, and M. Coll.** 1998. The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *Embo J* **17**:7404-15.
21. **Gomis-Ruth, F. X., M. Sola, R. Perez-Luque, P. Acebo, M. T. Alda, A. Gonzalez, M. Espinosa, G. del Solar, and M. Coll.** 1998. Overexpression, purification, crystallization and preliminary X-ray diffraction analysis of the pMV158-encoded plasmid transcriptional repressor protein CopG. *FEBS Lett* **425**:161-5.
22. **Gorby, G. L., and G. B. Schaefer.** 1992. Effect of attachment factors (pili plus Opa) on *Neisseria gonorrhoeae* invasion of human fallopian tube tissue in vitro: quantitation by computerized image analysis. *Microb Pathog* **13**:93-108.

23. **Grassme, H. U., R. M. Ireland, and J. P. van Putten.** 1996. Gonococcal opacity protein promotes bacterial entry-associated rearrangements of the epithelial cell actin cytoskeleton. *Infect Immun* **64**:1621-30.
24. **Gray-Owen, S. D.** 2003. Neisserial Opa proteins: impact on colonization, dissemination and immunity. *Scand J Infect Dis* **35**:614-8.
25. **Gray-Owen, S. D., C. Dehio, A. Haude, F. Grunert, and T. F. Meyer.** 1997. CD66 carcinoembryonic antigens mediate interactions between Opa- expressing *Neisseria gonorrhoeae* and human polymorphonuclear phagocytes. *Embo J* **16**:3435-45.
26. **Hamilton, H. L., and J. P. Dillard.** 2006. Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. *Mol Microbiol* **59**:376-85.
27. **Hamilton, H. L., N. M. Dominguez, K. J. Schwartz, K. T. Hackett, and J. P. Dillard.** 2005. *Neisseria gonorrhoeae* secretes chromosomal DNA via a novel type IV secretion system. *Mol Microbiol* **55**:1704-21.
28. **Hamrick, T. S., J. A. Dempsey, M. S. Cohen, and J. G. Cannon.** 2001. Antigenic variation of gonococcal pilin expression in vivo: analysis of the strain FA1090 pilin repertoire and identification of the pilS gene copies recombining with pilE during experimental human infection. *Microbiology* **147**:839-49.
29. **Handsfield, H., and P. Sparling.** 1995. Principles and Practice of Infectious Diseases, 4 ed, vol. 2. Churchill Livingstone, New York.
30. **Harkness, A. H.** 1948. The pathology of gonorrhoea. *Br J Vener Dis* **24**:137-146.

31. **Harvey, H. A., M. P. Jennings, C. A. Campbell, R. Williams, and M. A. Apicella.** 2001. Receptor-mediated endocytosis of *Neisseria gonorrhoeae* into primary human urethral epithelial cells: the role of the asialoglycoprotein receptor. *Mol Microbiol* **42**:659-72.
32. **Hebeler, B., and F. Young.** 1975. Autolysis of *Neisseria gonorrhoeae*. *J. Bacteriol.* **122**:385-392.
33. **Hill, S. A.** 2000. Opa expression correlates with elevated transformation rates in *Neisseria gonorrhoeae*. *J Bacteriol* **182**:171-8.
34. **Hopper, S., B. Vasquez, A. Merz, S. Clary, J. S. Wilbur, and M. So.** 2000. Effects of the immunoglobulin A1 protease on *Neisseria gonorrhoeae* trafficking across polarized T84 epithelial monolayers. *Infect Immun* **68**:906-11.
35. **Hopper, S., J. S. Wilbur, B. L. Vasquez, J. Larson, S. Clary, I. J. Mehr, H. S. Seifert, and M. So.** 2000. Isolation of *Neisseria gonorrhoeae* mutants that show enhanced trafficking across polarized T84 epithelial monolayers. *Infect Immun* **68**:896-905.
36. **Howell-Adams, B., and H. S. Seifert.** 2000. Molecular models accounting for the gene conversion reactions mediating gonococcal pilin antigenic variation. *Mol Microbiol* **37**:1146-58.
37. **Ilver, D., H. Kallstrom, S. Normark, and A. B. Jonsson.** 1998. Transcellular passage of *Neisseria gonorrhoeae* involves pilus phase variation. *Infect Immun* **66**:469-73.

38. **Johansson, L., A. Rytönen, P. Bergman, B. Albiger, H. Kallstrom, T. Hokfelt, B. Agerberth, R. Cattaneo, and A. B. Jonsson.** 2003. CD46 in meningococcal disease. *Science* **301**:373-5.
39. **Kline, K. A., E. V. Sechman, E. P. Skaar, and H. S. Seifert.** 2003. Recombination, repair and replication in the pathogenic *Neisseriae*: the 3 R's of molecular genetics of two human-specific bacterial pathogens. *Mol Microbiol* **50**:3-13.
40. **Koomey, J. M., and S. Falkow.** 1987. Cloning of the *recA* gene of *Neisseria gonorrhoeae* and construction of gonococcal *recA* mutants. *J Bacteriol* **169**:790-5.
41. **Lee, S. W., R. A. Bonnah, D. L. Higashi, J. P. Atkinson, S. L. Milgram, and M. So.** 2002. CD46 is phosphorylated at tyrosine 354 upon infection of epithelial cells by *Neisseria gonorrhoeae*. *J Cell Biol* **156**:951-7.
42. **Lin, L., P. Ayala, J. Larson, M. Mulks, M. Fukuda, S. R. Carlsson, C. Enns, and M. So.** 1997. The *Neisseria* type 2 IgA1 protease cleaves LAMP1 and promotes survival of bacteria within epithelial cells. *Mol Microbiol* **24**:1083-94.
43. **Masi, A. T., and B. I. Eisenstein.** 1981. Disseminated gonococcal infection (DGI) and gonococcal arthritis (GCA): II. Clinical manifestations, diagnosis, complications, treatment, and prevention. *Semin Arthritis Rheum* **10**:173-97.
44. **McCormick, B. A., S. P. Colgan, C. Delp-Archer, S. I. Miller, and J. L. Madara.** 1993. *Salmonella Typhimurium* attachment to human intestinal epithelial monolayers: transcellular signalling to subepithelial neutrophils. *J Cell Biol* **123**:895-907.

45. **McGee, Z., and D. Stephens.** 1984. Common Pathways of Invasion of Mucosal Barriers by *Neisseria gonorrhoeae* and *Neisseria meningitidis*. *Surv. Synth. Path. Res.* **3**:1-10.
46. **McGee, Z., D. Stephens, L. Hoffman, W. Schlech III, and R. Horn.** 1983. Mechanisms of mucosal invasion by pathogenic *Neisseria*. *Rev. Infect. Dis.* **5**:S708-S714.
47. **McGee, Z. A., A. P. Johnson, and D. Taylor-Robinson.** 1981. Pathogenic mechanisms of *Neisseria gonorrhoeae*: observations on damage to human Fallopian tubes in organ culture by gonococci of colony type 1 or type 4. **143**:413-422.
48. **McGee, Z. A., and D. S. Stephens.** 1984. Common pathways of invasion of mucosal barriers by *Neisseria gonorrhoeae* and *Neisseria meningitidis*. *Surv Synth Pathol Res* **3**:1-10.
49. **McKenna, W. R., P. A. Mickelsen, P. F. Sparling, and D. W. Dyer.** 1988. Iron uptake from lactoferrin and transferrin by *Neisseria gonorrhoeae*. *Infect Immun* **56**:785-91.
50. **Mehr, I. J., and H. S. Seifert.** 1998. Differential roles of homologous recombination pathways in *Neisseria gonorrhoeae* pilin antigenic variation, DNA transformation and DNA repair. *Mol Microbiol* **30**:697-710.
51. **Merz, A. J., C. A. Enns, and M. So.** 1999. Type IV pili of pathogenic neisseriae elicit cortical plaque formation in epithelial cells. *Mol Microbiol* **32**:1316-32.

52. **Merz, A. J., D. B. Rifken, C. G. Arvidson, and M. So.** 1996. Traversal of a polarized epithelium by pathogenic *Neisseriae*: facilitation by type IV pili and maintenance of epithelial barrier function. *Mol Med* **2**:745-54.
53. **Merz, A. J., and M. So.** 1997. Attachment of piliated, Opa- and Opc- gonococci and meningococci to epithelial cells elicits cortical actin rearrangements and clustering of tyrosine-phosphorylated proteins. *Infect Immun* **65**:4341-9.
54. **Merz, A. J., and M. So.** 2000. Interactions of pathogenic *neisseriae* with epithelial cell membranes. *Annu Rev Cell Dev Biol* **16**:423-57.
55. **Meyer, T. F.** 1999. Pathogenic *neisseriae*: complexity of pathogen-host cell interplay. *Clin Infect Dis* **28**:433-41.
56. **Mosleh, I. M., H. J. Boxberger, M. J. Sessler, and T. F. Meyer.** 1997. Experimental infection of native human ureteral tissue with *Neisseria gonorrhoeae*: adhesion, invasion, intracellular fate, exocytosis, and passage through a stratified epithelium. *Infect Immun* **65**:3391-8.
57. **Phillips, S. E.** 1994. The beta-ribbon DNA recognition motif. *Annu Rev Biophys Biomol Struct* **23**:671-701.
58. **Pujol, C., E. Eugene, M. Marceau, and X. Nassif.** 1999. The meningococcal PilT protein is required for induction of intimate attachment to epithelial cells following pilus-mediated adhesion. *Proc Natl Acad Sci U S A* **96**:4017-22.
59. **Pullinger, G. D., and A. J. Lax.** 1992. A *Salmonella dublin* virulence plasmid locus that affects bacterial growth under nutrient-limited conditions. *Mol Microbiol* **6**:1631-43.

60. **Radnedge, L., M. A. Davis, B. Youngren, and S. J. Austin.** 1997. Plasmid maintenance functions of the large virulence plasmid of *Shigella flexneri*. *J Bacteriol* **179**:3670-5.
61. **Raumann, B. E., K. L. Knight, and R. T. Sauer.** 1995. Dramatic changes in DNA-binding specificity caused by single residue substitutions in an Arc/Mnt hybrid repressor. *Nat Struct Biol* **2**:1115-22.
62. **Raumann, B. E., M. A. Rould, C. O. Pabo, and R. T. Sauer.** 1994. DNA recognition by beta-sheets in the Arc repressor-operator crystal structure. *Nature* **367**:754-7.
63. **Rudel, T., D. Facius, R. Barten, I. Scheuerpflug, E. Nonnenmacher, and T. F. Meyer.** 1995. Role of pili and the phase-variable PilC protein in natural competence for transformation of *Neisseria gonorrhoeae*. *Proc Natl Acad Sci U S A* **92**:7986-90.
64. **Rudel, T., I. Scheuerpflug, and T. F. Meyer.** 1995. *Neisseria* PilC protein identified as type-4 pilus tip-located adhesin. *Nature* **373**:357-9.
65. **Rytkonen, A., L. Johansson, V. Asp, B. Albiger, and A. B. Jonsson.** 2001. Soluble pilin of *Neisseria gonorrhoeae* interacts with human target cells and tissue. *Infect Immun* **69**:6419-26.
66. **Sandoval, R. M., and B. A. Molitoris.** 2004. Gentamicin traffics retrograde through the secretory pathway and is released in the cytosol via the endoplasmic reticulum. *Am J Physiol Renal Physiol* **286**:F617-24.
67. **Sarubbi, F., and F. Sparling.** 1974. Transfer of Antibiotic Resistance in Mixed Cultures of *Neisseria gonorrhoeae*. *J. Infect. Dis.* **130**:660-663.

68. **Sarubbi, F. A., Jr., and P. F. Sparling.** 1974. Transfer of antibiotic resistance in mixed cultures of *Neisseria gonorrhoeae*. *J Infect Dis* **130**:660-3.
69. **Sauer, R. T., M. E. Milla, C. D. Waldburger, B. M. Brown, and J. F. Schildbach.** 1996. Sequence determinants of folding and stability for the P22 Arc repressor dimer. *Faseb J* **10**:42-8.
70. **Sayeed, S., L. Reaves, L. Radnedge, and S. Austin.** 2000. The stability region of the large virulence plasmid of *Shigella flexneri* encodes an efficient postsegregational killing system. *J Bacteriol* **182**:2416-21.
71. **Schildbach, J. F., M. E. Milla, P. D. Jeffrey, B. E. Raumann, and R. T. Sauer.** 1995. Crystal structure, folding, and operator binding of the hyperstable Arc repressor mutant PL8. *Biochemistry* **34**:1405-12.
72. **Schreiter, E. R., M. D. Sintchak, Y. Guo, P. T. Chivers, R. T. Sauer, and C. L. Drennan.** 2003. Crystal structure of the nickel-responsive transcription factor NikR. *Nat Struct Biol* **10**:794-9.
73. **Sechman, E. V., M. S. Rohrer, and H. S. Seifert.** 2005. A genetic screen identifies genes and sites involved in pilin antigenic variation in *Neisseria gonorrhoeae*. *Mol Microbiol* **57**:468-83.
74. **Seifert, H. S.** 1996. Questions about gonococcal pilus phase- and antigenic variation. *Mol Microbiol* **21**:433-40.
75. **Seifert, H. S., R. S. Ajioka, C. Marchal, P. F. Sparling, and M. So.** 1988. DNA transformation leads to pilin antigenic variation in *Neisseria gonorrhoeae*. *Nature* **336**:392-395.

76. **Serkin, C. D., and H. S. Seifert.** 1998. Frequency of pilin antigenic variation in *Neisseria gonorrhoeae*. *J Bacteriol* **180**:1955-8.
77. **Skaar, E. P., M. P. Lazio, and H. S. Seifert.** 2002. Roles of the *recJ* and *recN* genes in homologous recombination and DNA repair pathways of *Neisseria gonorrhoeae*. *J Bacteriol* **184**:919-27.
78. **Somers, W. S., and S. E. Phillips.** 1992. Crystal structure of the *met* repressor-operator complex at 2.8 Å resolution reveals DNA recognition by beta-strands. *Nature* **359**:387-93.
79. **Stern, A., M. Brown, P. Nickel, and T. F. Meyer.** 1986. Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. *Cell* **47**:61-71.
80. **Stern, A., P. Nickel, T. F. Meyer, and M. So.** 1984. Opacity determinants of *Neisseria gonorrhoeae*: gene expression and chromosomal linkage to the gonococcal pilus gene. *Cell* **37**:447-56.
81. **Stohl, E. A., and H. S. Seifert.** 2001. The *recX* gene potentiates homologous recombination in *Neisseria gonorrhoeae*. *Mol Microbiol* **40**:1301-10.
82. **Swanson, J.** 1973. Studies on gonococcus infection. IV. Pili: their role in attachment of gonococci to tissue culture cells. *J Exp Med* **137**:571-89.
83. **Turner, C. F., S. M. Rogers, H. G. Miller, W. C. Miller, J. N. Gribble, J. R. Chromy, P. A. Leone, P. C. Cooley, T. C. Quinn, and J. M. Zenilman.** 2002. Untreated gonococcal and chlamydial infection in a probability sample of adults. *Jama* **287**:726-33.

84. **van Putten, J. P., T. D. Duensing, and R. L. Cole.** 1998. Entry of OpaA+ gonococci into HEp-2 cells requires concerted action of glycosaminoglycans, fibronectin and integrin receptors. *Mol Microbiol* **29**:369-79.
85. **Virji, M., S. M. Watt, S. Barker, K. Makepeace, and R. Doyonnas.** 1996. The N-domain of the human CD66a adhesion molecule is a target for Opa proteins of *Neisseria meningitidis* and *Neisseria gonorrhoeae*. *Mol Microbiol* **22**:929-39.
86. **Wainwright, L. A., K. H. Pritchard, and H. S. Seifert.** 1994. A conserved DNA sequence is required for efficient gonococcal pilin antigenic variation. *Mol Microbiol* **13**:75-87.
87. **Wang, J., S. D. Gray-Owen, A. Knorre, T. F. Meyer, and C. Dehio.** 1998. Opa binding to cellular CD66 receptors mediates the transcellular traversal of *Neisseria gonorrhoeae* across polarized T84 epithelial cell monolayers. *Mol Microbiol* **30**:657-71.
88. **Zenni, M. K., P. C. Giardina, H. A. Harvey, J. Shao, M. R. Ketterer, D. M. Lubaroff, R. D. Williams, and M. A. Apicella.** 2000. Macropinocytosis as a mechanism of entry into primary human urethral epithelial cells by *Neisseria gonorrhoeae*. *Infect Immun* **68**:1696-9.
89. **Zhang, Y. X., J. Li, X. K. Guo, C. Wu, B. Bi, S. X. Ren, C. F. Wu, and G. P. Zhao.** 2004. Characterization of a novel toxin-antitoxin module, VapBC, encoded by *Leptospira interrogans* chromosome. *Cell Res* **14**:208-16.

Figure legends

Figure 1 Model of the *Neisseria gonorrhoeae* infection process.

1. Loose adherence, 0-3 hours post infection – Initial contact with the host cell, mediated by the type IV pilus and involves interactions with the host-cell receptor CD46 (43, 65). This interaction is marked by a calcium flux, activation of host-cell kinases and endosome and lysosome exocytosis.

2. Microcolony formation, 1-4 hours post infection – GC form a microcolony of the host-cell surface following initial attachment. This process is dependent on a functioning, retractile type IV pilus. The host cell membrane is enriched for many host-cell proteins under the microcolony (merz,). This enriched region is called the cortical plaque.

3. Tight adherence, 4-6 hours post infection – The microcolony spreads out on the cell surfaces, pili are shed and the bacterial and host-cell membranes are found in close proximity. This interaction has been proposed to involve an interaction between one of the Opa and host-cell receptors or the bacterial LPS and host-cell Toll-like receptors.

4. Invasion, 6-8 hours post infection – GC invade in a actin dependent process similar to macropinocytosis.

5. Intracellular replication, 6-24 hours post infection – GC traffic to a compartment within the host-cell that is not accessible to the antibiotic gentamicin and undergo replication. The *fit* mutant has been shown to have increased intracellular replication.

6. Transcytosis, 24+ hours post infection – GC exocytose from the basolateral membrane of host cells. The bacteria have undergone *pilE* antigenic variation (AV) and are expressing Opa proteins. The process of crossing polarized epithelial cells (transcytosis) is thought to be dependent on the expression of one or more of the Opa proteins.

Fig. 2. Pilin antigenic variation hybrid intermediate model 1.

A and B. Recombination between *pilS1* and *pilE* to form a *pilE/pilS* hybrid intermediate.

A. Recombination at a short region of shared identity (filled diamond) between *pilS1* (open boxes) copy 4 and *pilE* (grey box) to form a hybrid intermediate. Each box in *pilS1* represents the six silent copies as indicated above.

B. More detailed view of interactions between the donor *pilS* copy and *pilE*. The structure of the pilin gene family: constant (const.), semivariable (SV), conserved cysteine region 1 (cys1), hypervariable loop (HVL), conserved cysteine region 2 (cys2) and hypervariable tail (HVT) are shown on the top. The cross (X) shows the crossover site at a short region of identity shared between the *pilS* copy and *pilE*. Constant or short regions of identical

sequences are represented by black boxes, *pilS*-specific sequences are white and *pilE*-specific sequences are grey. The filled double-headed arrows show the proposed interactions between the conserved *cys1* and *cys2* sequences (cross-hatched boxes) (Howell-Adams and Seifert, 1999). The white double-headed arrows show putative interactions with other conserved sequences. The four point star represents the *Sma/Cla* repeat, which in *pilS1* is unlinked to five of the six pilin copies.

C and D. Recombination of a *pilE/pilS* hybrid intermediate with a recipient *pilE*.

C. Recombination between a hybrid intermediate and the recipient *pilE*. Recombination between a *pilS* portion of the hybrid locus and a shared region of identity in the recipient *pilE* (white diamond). The crossover (thick X) upstream of the *pilE* sequences shows the homologous recombination event required to insert *pilS* sequences into *pilE*.

D. More detailed view of recombination between a hybrid intermediate and a recipient *pilE*. The thin X shows a crossover between two short regions of identity. The white double-headed arrows show putative interactions with other conserved sequences.

E. Sequence map of *pilE* region showing location of cis acting elements. Large black arrow is *pilE* ORF. Hooked arrow is the promoter. Light Grey diamond is cis acting *Sma/Cla* repeat. Dark grey hexagon is RS1 (regulatory sequence 1). Black oval is RS2 (regulatory sequence 2). The FitIS and FitPP, sequences bound by FitA/B are not present in this region.

This figure was adapted from Howell-Adams B and Seifert HS, Mol Microbiol 2000

Figure 3 Sequence of *fit* locus.

A. Schematic map of predicted *fitA* and *fitB* ORFs. A1 and A9 mark locations of transposon insertion. The A9 mutant was characterized in Hopper et al 2000 and Chapter 4.

B. Sequence of *fit* locus. In the predicted promoter region of *fitA* the predicted RNA polymerase binding site (-10) and ribosome binding site (RBS) are underlined. Arrows show locations of A1 and A9 transposon insertion. Predicted amino acid translation of *fitA* and *fitB* are below the DNA sequence. Box shows the overlapping stop codon of *fitA* and the start codon of *fitB*.

Figure 1

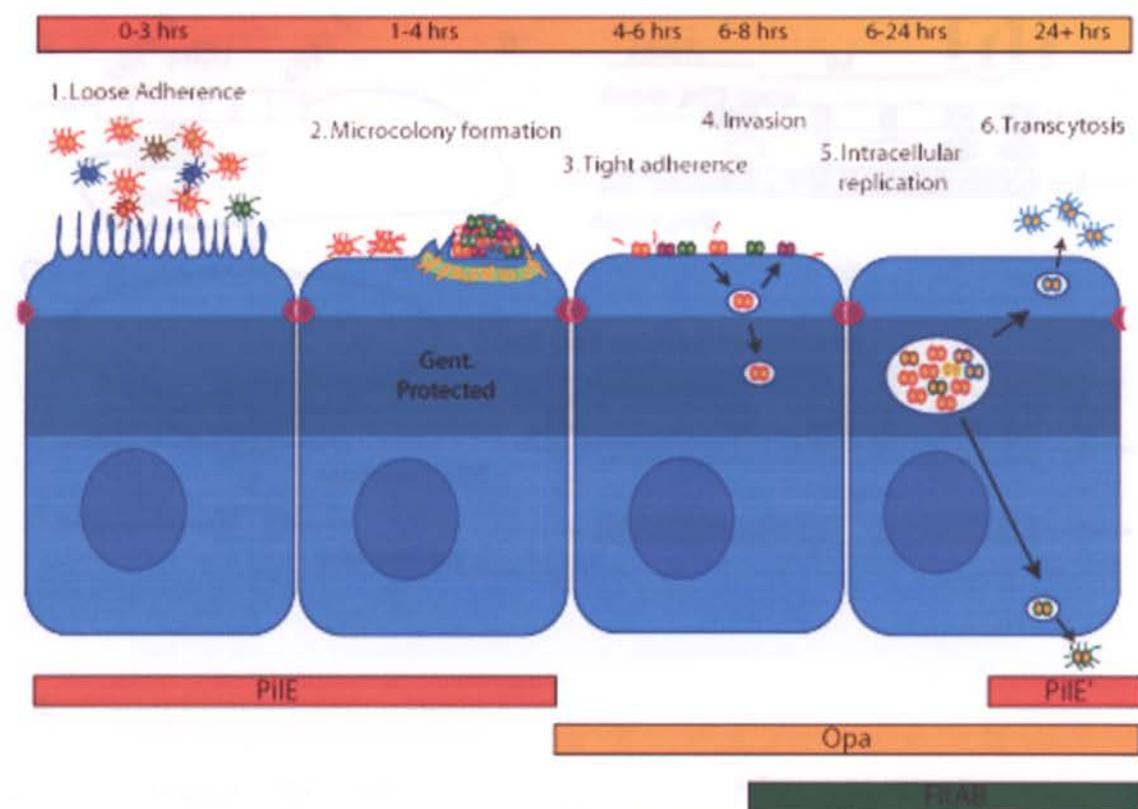


Figure 2

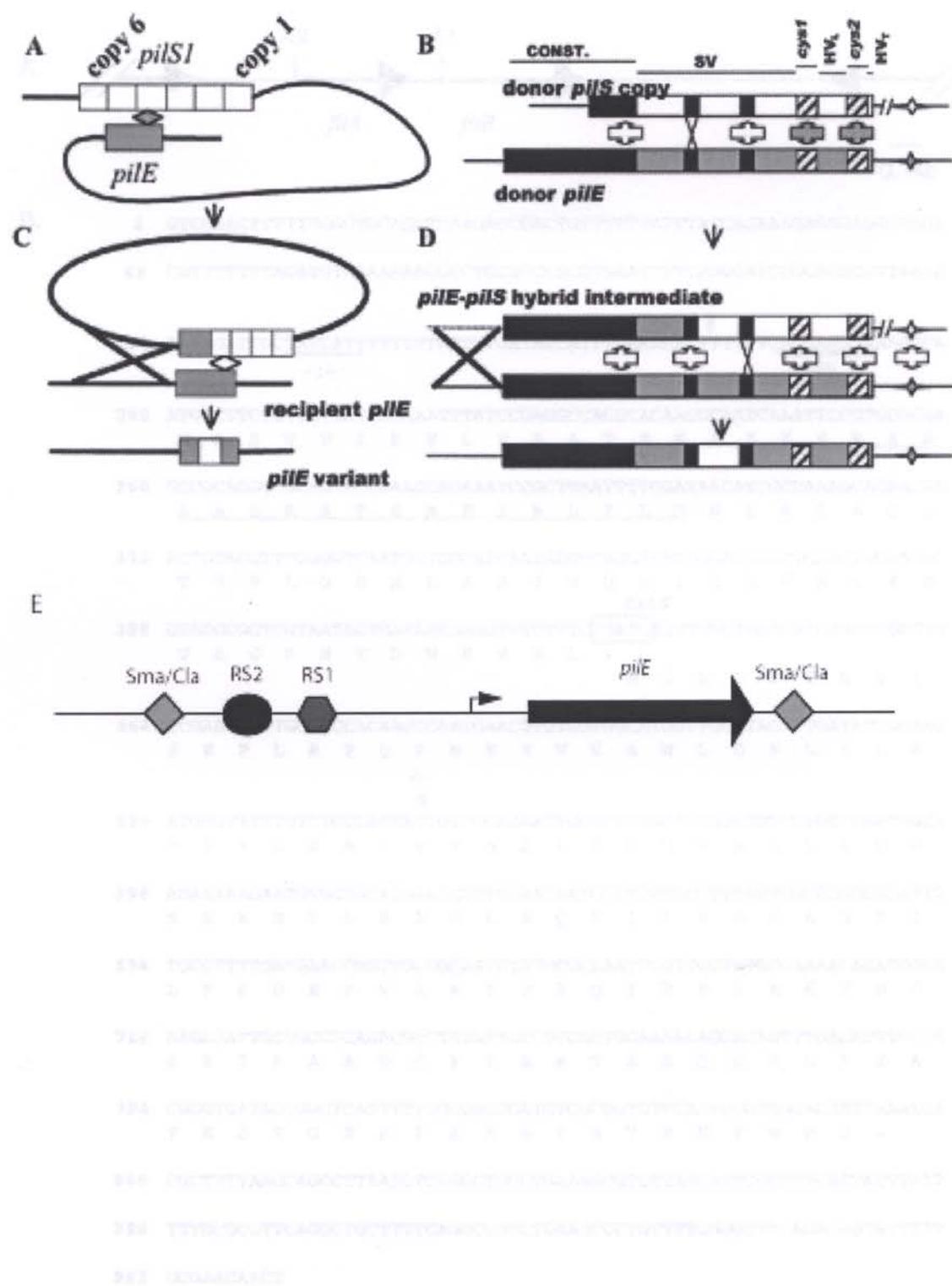
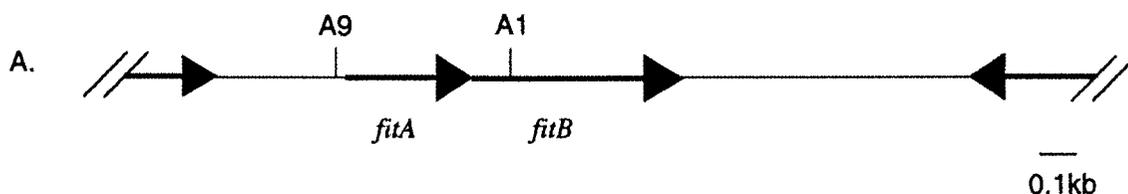


Figure 3



B.

1 GTCCGACTTTTTGGGGTGCAGTTCAAGAGCGACTGTTTTTTATTTATCAGAAAAGAGGGAGGGGAGA

68 CGTTTTTTTAGATGTAAAAAACGCCTGCGGCCACCTGAATTTTCAGGCATCCGCAGGCGTTAAAA

134 AATAGATTGCTATCATTTTTTTTATTTTGATAGCATTGAAAAACTTTTTTCAGGAAGGAAAGGCA

-10

fitA

RBS

200 ATGGCTTCTGTTGTGATTAGAAATTTATCCGAGGCCACGCACAACGCAATCAAATCCGTGCGCGA

M A S V V I R N L S E A T H N A I K F R A R

266 GCCGCAGGGCGCAGTACCGAAGCAGAAATCCGCTTAATTTGGATAACATCGCCAAAGCACAACAA

A A G R S T E A E I R L I L D N I A K A Q Q

332 ACTGTACGTTTGGGGTCAATGTTGGCATCAATAGGGCAGGAAATCGGAGGTGTTGAGCTGGAAGAC

T V R L G S M L A S I G Q E I G G V E L E D

fitB

398 GTACGCGGTCGTAATACTGATAACGAGGTTTCTTTGATGATTTTTGCTGGACACGAATGTGATTT

V R G R N T D N E V S L .

M I L L D T N V I

464 CCGAACCTTTGCGCCCACAACCCAATGAACGTGTGGTGGCATGGTTGGATAGTTTGATATTGGAAG

S E P L R P Q P N E R V V A W L D S L I L E

A1

530 ATGTGTATTTGCTGCCATTACTGTTGCAGAAATGCGTTTGGGTGTGGCGTTGTTGCTCAATGGCA

D V Y L S A I T V A E L R L G V A L L L N G

596 AGAAAAAGAATGTGCTGCACGAACGTTTGGAAACAATCCATTTGCTTTATTTGCGGGGCGGATTC

K K K N V L H E R L E Q S I L P L F A G R I

638 TGCCTTTTGATGAACCGGTTGCCGAATCTATGCGCAAATTCGTTCTATGCCAAAACACATGGCA

L P F D E P V A A I Y A Q I R S Y A K T H G

728 AAGAGATTGCTGCCGACACGGCTATATTGCCCCACTGCAAAACAGCACAGTTTGACAGTTGCTA

K E I A A A D G Y I A A T A K Q H S L T V A

794 CGCGTGATACCGGCTCATTTTTTGCGGCCGATGTCGCGGTGTTCAATCCGTGGCAGCATTAAAAGA

T R D T G S F F A A D V A V F N P W H D .

860 CGCTTTTAAGCAGCCTTAATGTCAGGCTGCCTGAAAGCGTCTTAGCATTCCGTTTCATTATTTATT

926 TTTGCGCGTTCAGGCTGCTTTTCAGGCCGTCTGAAGCCCTGTTTTCGGGTTTCAGACGGTATTTTT

992 GCGAACAACT

Chapter 2: Manuscript I
***Neisseria gonorrhoeae* FitA interacts with FitB to Bind DNA through
a ribbon-helix-helix motif**

***Neisseria gonorrhoeae* FitA interacts with FitB to bind DNA through its ribbon-
helix-helix motif[†]**

J. Scott Wilbur^{†}, Peter T. Chivers[§], Kirsten Mattison^{||}, Laura Potter[‡],
Richard G. Brennan³ and Magdalene So[‡]*

[‡] Department of Molecular Microbiology and Immunology, Oregon Health and Sciences University, Portland, OR 97239. wilburj@ohsu.edu

[§]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110

^{||} Department of Biochemistry, School of Medicine, Oregon Health and Sciences University, Portland, OR 97239

[†]This work was funded by the NIH grant 5RO1 AI47260-05

Running Title: FitA/B DNA binding complex

Abbreviations and Textual Footnotes

¹ Abbreviation

GC – *Neisseria gonorrhoeae*, gonococcus

RHH – Ribbon-helix-helix motif

FitIS – Fit interaction sequence

FitPP – Fit perfect palindrome

Abstract: The *fit* locus, encoding two proteins, FitA and FitB, was identified in a genetic screen for *Neisseria gonorrhoeae* determinants that affect trafficking across polarized epithelial cells. To better understand how the locus may control these activities, we have undertaken a biochemical analysis of FitA and FitB. FitA is a DNA binding protein with a ribbon-helix-helix (RHH) motif. Purified FitA forms a homodimer that binds a 150 bp *fit* promoter sequence containing the translational start site. A putative β -strand mutant, FitA(R7A), is unable to bind this DNA, supporting further that FitA is a RHH protein. FitB interacts with FitA to form a 98 kDa complex. FitA/B binds DNA with a 38-fold higher affinity than the FitA homodimer. To our knowledge, FitB is the only protein that increases the DNA-binding affinity of an RHH protein. In DNase I footprint assays, FitA/B protects a 62-bp region within the *fit* promoter containing the predicted -10 sequence and an 8-bp inverted repeat, TGCTATCA-N₁₂-TGATAGCA. FitA/B_{His} is able to bind to either half-site alone with high affinity.

Neisseria gonorrhoeae (GC, gonococcus) is a sexually-transmitted pathogen that infects only man. Infection is usually initiated at the mucosal epithelium of the urogenital tract (1). Because of its specificity for a number of human proteins (2-7) interactions between GC and host tissues have been studied mainly using human cell culture systems (8-12). These studies have shed light on the molecular and cell biology of GC attachment, cell entry and intracellular survival (13-17). In contrast, the mechanisms used by GC to cross the epithelial barrier are not well understood. Studies of transcellular trafficking using monolayers of polarized T84 human colorectal carcinoma cells have shown that piliated GC not expressing Opa (a protein that also influences attachment and invasion, (18) cross the epithelial barrier in 24-36 hours after inoculation onto the apical membrane (9). In contrast, *N. meningitidis*, a closely related pathogen with very similar virulence factors, crosses this monolayer within 6 hours.

We identified a chromosomal locus in GC that plays a role in transcellular trafficking and intracellular growth called *fit* (fast intracellular trafficking), (9, 14, 19). Mutants with transposon insertions in *fit* cross polarized T84 monolayers 5-6 times more quickly than their wt parent strain. Moreover, *fit* mutants are altered in their intracellular behavior: they replicate within epithelial cells more quickly than the wt parent, though their replication in liquid medium is unaffected.

The *fit* locus encodes two open reading frames (ORF), *fitA* and *fitB*. The ORFs are tandemly arranged with a +1 base overlap, i.e., the last base of the *fitA* stop codon serves as the first base of the *fitB* start codon. FitA has a predicted molecular weight of 8.4 kDa and a pI of 6.6. FitB is predicted to be a 15.3 kDa protein with a pI of 5.8.

Neither FitA nor FitB has a Sec signal sequence. Preliminary analysis suggested that the products of the *fit* locus are cytosolic and that FitA may interact with DNA (19).

Detailed biochemical knowledge of the *fit* proteins is crucial for understanding the role of the locus in gonococcal pathogenesis. We had hypothesized that FitA binds DNA and, through this interaction, regulate bacterial division within the host cell. Here we report that FitA binds to its own promoter through a ribbon-helix-helix motif (RHH). RHH motifs have been identified and studied in numerous proteins with DNA binding activity (20-24). This motif consists of a β -strand followed by two α -helices (25). The β -strand inserts into the major groove of the DNA and is involved in sequence recognition (21, 26). The two α -helices are primarily involved in dimer formation. Many members of this family of proteins are autoregulatory. Arc, NikR, CopG and Mnt bind as homotetramers to sequences within their own promoters (21, 26-29). The binding sites of these proteins consist of two 6-8 bp halfsites, separated by 1-16 bp of intervening sequence.

Our studies revealed that FitB forms a complex with FitA and this complex increase the DNA-binding affinity of FitA. All other RHH proteins bind DNA without the help of a second protein, although the MetJ and NikR repressors require small molecule activators for DNA-binding (S-adenosylmethionine and nickel, respectively; 30, 31), and ParG is negatively affected in its DNA-binding activity upon complexing with ParF (32). Thus, FitA appears to be a novel member of the RHH family of proteins whose DNA-binding activity is positively affected by FitB.

Experimental procedures

Preparation of DNA—PCR assays were performed using Pfu polymerase (Stratagene). The template for PCR amplifications was chromosomal DNA from the *Neisseria gonorrhoeae* strain FA1090 and purified as described (19). PCR was performed for 30 cycles: 30 s at 96 °C, 30 s at 50 °C, 60 s at 72 °C, unless otherwise indicated. DNA for gel retardation assays was prepared by PCR amplification, end-labeled with [$\alpha^{32}\text{P}$] ATP using T4 kinase (Invitrogen), and purified on ProbQuant G-50 micro columns (Amersham Biosciences) as per manufacturer's instructions.

Oligonucleotide pairs for fragments were as follows: *fit_{pro}* Stbpre402 and StbL108, Fragment 1 StbL108 and JSWanti-64, Fragment 2 StbL108 and JSWanti-65, Fragment 3 JSW 64 and Stbpre402, Fragment 4 JSW 65 and Stbpre402, Fragment 5 JSW 64 and JSWanti-65. PCR annealing temperature for fragments 1-5 was 45 °C.

JSW₂₂₀ was generated by PCR amplification using the oligonucleotide primers JSW65 and StbR205. The amplified fragment was then cloned into the vector pCR-Blunt II (Invitrogen). DNA fragments were excised from this vector by restriction digest and the single stranded 3' overhang of one of the sites was filled with Klenow fragment (Roche) and [$\gamma^{32}\text{P}$]d-ATP (Invitrogen) in restriction digest buffer. "Plus" strand labeled DNA was generated by cutting with HindIII (filled) then EcoRV in buffer B (Roche). "Minus" strand labeled DNA was generated by cutting with XbaI (filled) then SpeI in buffer H (Roche). Labeled DNA fragments were purified on a 6% TBE (90 mM Tris/borate, 2 mM EDTA) polyacrylamide gel and resuspended in H₂O.

Overexpression and purification of histidine fusion proteins—The *fitA* ORF was amplified by PCR using the oligonucleotides PETFit5'.NcoI and FitA.XhoI-3'. PCR was performed as described above. These primers amplify a DNA fragment that begins at the *fitA* start codon and ends at its last codon, just prior to the stop codon; it also includes the restriction sites NcoI and XhoI at each end of the ORF to facilitate cloning. The *fitA* insert, excised with NcoI and XhoI, was ligated into the appropriately digested vector pET28b (Invitrogen) for overexpression. This vector added a T7 promoter and sequences encoding six histidine residues to the 5' terminus of the *fitA* ORF. The resulting plasmid (pET28b::*fitA*) was transformed into the *E. coli* strain BL21 (pLysE, Invitrogen) for overexpression. This protocol was also used to construct vectors for overexpression of FitA/B_{His} (pET28b::*fitA/B_{His}*) and FitA(R7A)_{His} (pET28b::*fitA(R7A)_{His}*). Oligonucleotides used for construction of *fitA/B_{His}* were PETFit5'.NcoI and PETFit3'.XhoI. Oligonucleotides for construction of *fitA(R7A)_{His}* were FitA.R7A.NcoI and FitA.XhoI-3'. This last pair of oligonucleotides contains an alanine codon at the 7th position instead of arginine. FitA(R7A)/B_{His} (pET28b::*fitR7A/B_{His}*) was constructed using the oligonucleotides FitA.R7A.NcoI and PETFit3'.XhoI. All constructs were confirmed by sequencing analysis.

Protein overexpression and purification were performed as follows. 14 hour Luria broth (LB) cultures of each strain were diluted 1:100 in the same medium supplemented with Kanamycin (60 µg/ml) and incubated for two hours at 37 °C. Protein expression was induced by the addition of IPTG (0.5 mM) and further incubation at 37 °C for two hours. Bacteria were then collected by centrifugation and the pellets were resuspended and lysed in Bugbuster Protein Extraction Reagent with Benzonase (Novagen) per

manufacturer's instruction. Protein was then purified using a His Bind Kit (Novagen) per manufacturer's instructions. Briefly, bacterial lysates were loaded onto a 10 ml column from the His Bind Kit and the column was washed with 10 column volumes of Binding Buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9) and then 6 column volumes of Wash Buffer (60 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9). The his-tagged protein was eluted with 5 ml of Elution Buffer (1 M imidazole, 0.5 M NaCl, 20mM Tris-HCl pH 7.9), then concentrated on a 10 kDa Centricon Centrifugal Filter Device (Amicon). FitA/B_{His} was eluted from the nickel column in the Elution Buffer described above with the addition of 10 mM MgCl₂. Protein concentration was determined with a Bio-Rad protein assay (Bio-rad), using BSA as a standard.

Gel retardation assays—Binding reactions were performed according to published protocol (33). Briefly, binding was determined by incubating purified His-tagged protein with radiolabeled DNA in binding buffer (0.1mM poly-dIdC, 25mM Tris, 10mM MgCl₂, 50 mM NaCl, pH 7.5) in a total volume of 20 µl at 37°C for 30 minutes. Samples were then incubated at 4 °C for 10 minutes, and the products were separated in a native 4% polyacrylamide gel (80:1 acrylamide: bis, 12 mM tris/AC pH 7.5, 1 mM EDTA), 25-30 volts at 4 °C for 8-10 hours. Gels were then anchored to 3MM Whatman paper, covered in plastic wrap and exposed to film (X-Omat Blue XB1, Kodak) at -80 °C.

DNase I footprinting assays—20-µl binding reactions were prepared as above. Samples were incubated at 37 °C for 30 minutes, then with 1 µl DNase I (0.01unit/µl; Roche) at 25 °C for 2 minutes (33). Reaction mixtures were terminated by adding 1 µl of 0.5 M EDTA and incubated at 4 °C for 10 minutes. DNA fragments were then precipitated with ethanol (75%), sodium acetate (90 mM) and glycogen (1%) and frozen

for >14 hours at -80 °C. Pellets were collected by centrifugation, then resuspended in Stop Buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF), then incubated at 80 °C for 5 minutes. The reaction products were then separated by electrophoresis on a sequencing gel (6% 30:1 acrylamide, 6 M urea, TBE buffer) and the radioactive bands were detected by autoradiography (X-Omat AR, Kodak) at -80 °C.

Secondary structure and thermal stability—Circular dichroism experiments were performed on a Jasco J715 spectropolarimeter in 20 mM phosphate buffer (pH 7.5) using 1 cm pathlength. Far UV CD spectra were collected in triplicate from 260-200 nm (0.5 nm step size, 4 s averaging time). Thermal denaturation curves were measured at 222 nm (1 °C step size, 16 s averaging time) using a NESLAB RTE111 water bath.

Oligomeric state—Sedimentation equilibrium experiments were performed on a Beckman XL-A analytical ultracentrifuge. FitA protein samples (10-100 μM) were in 20 mM Tris (pH 8.0), 100 mM NaCl. FitA/B_{His} samples (50-100 μM) were in 100 mM sodium phosphate (pH 8.0), 100 mM sodium chloride, and 100 mM imidazole, with or without 10 mM MgCl₂. Equilibrium experiments were carried out at 13000, 19000, and 24000 rpm for FitA and 11000, 18000, and 25000 rpm for FitA/B_{His}. Equilibrium was reached by 12 h at each speed for all samples. Data were fit to a single oligomeric species model using SCIENTIST. The calculated molecular weights of FitA (18,997 Da) and FitA/B_{His} (24,787 Da) were determined using the amino acid sequence plus the sequence LEHHHHHH, which was introduced in the construction of FitA_{His} and FitA/B_{His}. The partial molar volume of FitA_{His} (0.7315 mg/ml), FitA/B_{His} (0.7444 mg/ml)

and buffer densities (1.007 g/ml, FitA_{His}; 1.018 g/ml, FitA/B_{His}) were calculated using the program SEDNTERP.

The molecular weight of the FitA/B_{His} complex was also determined by dynamic light scattering on DynaPro molecular size instrument (Protein Solutions). Protein was concentrated to 7 mg/ml (~70 mM), filtered and analyzed in a 10 mm cuvette. The data was then analyzed with a filter for di-dispersion with 95% 98 kDa and 5% larger aggregates.

Fluorescence anisotropy/polarization—Fluorescence polarization (FP) was performed on a Beacon 2000 (Panvera, Invitrogen). All experiments were done with 5' fluorescein labeled DNA fragments and millipolarization (mP) and millianisotropy (mA) values were monitored in response to increases in protein concentration. Oligonucleotide pairs are listed in Table 1. Each pair of complementary fragments was suspended in H₂O at a 5 mM concentration, and annealed following incubation at 100 °C for 20 minutes, and slowly cooled to ambient temperature (~23 °C) over a period of 60 minutes. All experiments were done in Fit binding buffer (25 mM Tris-pH 7.5, 10 mM MgCl₂ 50 mM NaCl).

FP experiments were done as previously described (34). Baseline was determined for 5 nM fluorescein labeled DNA followed by addition of up to 3000 nM protein. Dissociation constants (K_d) were calculated using the equation $P = \frac{(P_{\text{bound}} - P_{\text{free}})[\text{Protein}]}{K_d + [\text{Protein}]} + P_{\text{free}}$ where P is the polarization measured at a given total protein concentration, P_{free} is the initial polarization of free fluorescein labeled DNA, and P_{bound} is the maximum polarization of specifically bound DNA. The resulting binding curve was plotted using KaleidaGraph™ 3.5.

Results

FitA binds DNA as a homodimer through a ribbon-helix-helix (RHH)

motif—Sequence profiling (28) of published bacterial protein sequences uncovered a potential RHH motif in FitA located between residues 1 and 45 (Figure 1b). RHH motifs consist of a β -strand that is involved in DNA sequence recognition, followed by two α -helices that function in dimer formation. The putative FitA RHH-domain contains these motifs: the Chou-Fasman algorithm predicted a β -strand between residues 1 and 11, and an α -helical region between residues 15 and 39, centered by a turn (Figure 1b).

The primary function of an RHH-domain is to confer sequence specific DNA binding. Moreover, many RHH proteins are autoregulatory, binding to inverted repeats within their own promoters to control transcription. We therefore examined the ability of FitA to bind its own promoter. The C-terminus of FitA was tagged with 6 histidine residues by cloning the *fitA* ORF downstream of the T7 promoter in pET28b (Invitrogen). FitA_{His} was overexpressed in *E. coli* BL21 by IPTG induction and purified (Figure 2a) as described in Experimental procedures. Sedimentation equilibrium experiments carried out on FitA_{His} revealed that the protein has an apparent molecular weight of 18,573 Da (Table 2) compared to the predicted value of 18,997 Da for a dimer. Additionally, the thermal stability of FitA_{His} was concentration dependent (Table 2), as would be expected with a transition from folded dimer to unfolded monomers. These data suggest that FitA_{His} forms a homodimer in the absence of FitB, consistent with the obligate dimeric nature of the RHH fold.

FitA_{His} was next examined for its DNA-binding activity using a gel retardation assay. Varying amounts of FitA_{His} were incubated, in the presence of poly dIdC, with a fixed amount of ³²P-labelled *fit_{pro}*, a 320-bp DNA fragment encompassing the region immediately upstream of the *fit* ORF (Figure 2f). FitA_{His} bound to *fit_{pro}* at ≥ 2.2 μ M concentration (Figure 2b). This is a weak affinity for specific binding.

Deletion analysis of the *fit_{pro}* fragment was next performed to determine if FitA_{His} recognized a specific region of *fit_{pro}*. FitA_{His}, at a 2.2 μ M concentration, was assessed for its ability to bind five overlapping ³²P-labelled fragments covering the *fit_{pro}* region (Figure 2c and f). FitA_{His} shifted the mobility of fragments 3 and 4 but not the other fragments. This localizes the binding site to a 150 bp region immediately upstream of the *fitA* ORF. Gel retardation assays were also done with the fragment *JSW₂₂₀*. This fragment contained the 150-bp of fragment 4 and extends an additional 68-bp into the *fitA* ORF (Figure 2f). FitA_{His} bound *JSW₂₂₀* with a K_d of approximately 1.5 μ M (Figure 2d).

We next determined whether the putative RHH motif influences the DNA binding activity of FitA. Alanine substitutions in the β -strand within the RHH of Arc and NikR abrogate DNA binding activity without affecting the DNA binding domain fold or the ability of these proteins to dimerize (28, 35). The Chou-Fasman algorithm predicts that FitA residues 1-11 form part of a β -strand (Figure 1b), and a comparison of the FitA homologues reveals that the arginine in residue 7 is absolutely conserved in this region (Figure 1a). This arginine was replaced with alanine by site-directed mutagenesis and the C-terminus was His-tagged as described in Experimental procedures. FitA(R7A)_{His} was overexpressed in *E. coli* BL21 and purified as described for wt-FitA_{His} (Figure 2a). The secondary structure, thermal stability, and oligomeric state of FitA(R7A)_{His} were virtually

identical to FitA_{His} (Table 2, Figure 3), indicating that the arginine to alanine mutation did not destabilize the RHH domain.

The effect of the R7A mutation on DNA binding was investigated using a gel retardation assay. FitA(R7A)_{His} did not bind *JSW*₂₂₀ at a concentration of 12.0 μM (Figure 2D), indicating that this mutation decreased the DNA binding affinity of the protein at least 8-fold. Additionally, FitA(R7A)_{His} did not bind *fit*_{pro} under any conditions tested (data not shown). Taken together, these results strongly suggest that FitA is a DNA-binding protein that makes specific contact with DNA through the RHH motif.

FitB forms a complex with FitA—FitB was His-tagged by cloning the *fitB* ORF downstream of the T7 promoter in pET28b, as described in Experimental procedures. However, all attempts to overexpress soluble FitB_{His} failed; the protein was partitioned into inclusion bodies (data not shown).

We next co-expressed *fitA* and *fitB* by cloning the *fitA/B* ORFs downstream of the T7 promoter in pET28b. This construction also placed a His-tag at the C-terminus of FitB. IPTG induction of this construct in *E. coli* BL21 yielded large amounts of soluble FitB_{His} that could be recovered by nickel affinity column chromatography (Figure 2a). Furthermore, fractions containing FitB_{His} also contained a ~10 kDa protein. N-terminal protein sequencing (Edman degradation) revealed the two proteins to be FitA and FitB in a 1:1 ratio. FitA therefore appears to increase the solubility of FitB. Fractions containing FitA and FitB_{His} were examined further.

Dynamic light scattering (DLS) and analytical ultra centrifugation (AUC) were performed on the complex to determine its molecular weight. DLS was done on freshly purified protein and the protein was found to be 95% monodisperse in a complex with a

molecular weight of 98 kDa. AUC studies revealed the complex to sediment with an apparent molecular weight of 95 kDa. If the two proteins exist in a 1:1 ratio, the stoichiometry for the complex is likely to be A₄B₄, with a calculated molecular weight of 98,247 Da. We therefore used a molecular weight of 98 kDa when calculating FitA/B_{His} molarity.

FitB increases FitA DNA-binding affinity—FitA/B_{His} was next examined for its DNA-binding activity by gel retardation analysis. Increasing concentrations of the FitA/B_{His} complex were incubated with ³²P-labelled *fit_{pro}* in the presence of poly dIdC, and the products separated on a native gel. The FitA/B_{His} complex bound *fit_{pro}* at concentrations $\geq 0.190 \mu\text{M}$ (Figure 4a). FitA/B_{His} formed two nucleoprotein complexes with *fit_{pro}*. The faster-migrating complex was observed at a FitA/B_{His} concentration of $\sim 0.190 \mu\text{M}$, and the slower-migrating complex at $\geq 0.765 \mu\text{M}$ (Figure 4a). The nature of these two complexes is unclear.

The FitA/B_{His} binding site in the *fit* promoter region was also examined. FitA/B_{His} (0.19 μM) was added to each of the five overlapping, PCR-generated, ³²P-labelled fragments covering the *fit_{pro}* region (see Figure 2f) and binding was assessed using the gel retardation assay. Like FitA_{His} alone, FitA/B_{His} bound only fragments 3 and 4 (Figure 4b).

We noted that FitA/B_{His} has a higher affinity for binding to the *fit_{pro}* fragment than FitA alone (2.2 μM FitA_{His} vs. 0.19 μM FitA/B_{His}). This was confirmed in additional binding assays using the *JSW₂₂₀* fragment as probe. FitA_{His} bound *JSW₂₂₀* at a K_d of 1.50 μM . In contrast FitA/B_{His} bound *JSW₂₂₀* with a K_d of 0.048 μM , indicating that the FitA/B_{His} complex has a 20-30 fold higher affinity for this DNA than FitA_{His} alone (Figure 2d and e).

When FitA(R7A) was co-expressed with FitB_{His}, both proteins could be recovered from the column (Figure 2a). The R7A mutation, therefore, does not affect the ability of FitA to interact with FitB. In addition, the FitA(R7A)/FitB_{His} complex did not bind *JSW*₂₂₀ at a concentration of 1.56 μM (Figure 2e) or *fit*_{pro} (data not shown). Compared to FitA/B_{His}, the affinity of the FitA(R7A)/FitB_{His} mutant complex to DNA is decreased >30-fold. These experiments provide additional evidence that the DNA-binding activity of FitA is dependent on its RHH motif and that the FitA/B complex binds DNA primarily through this motif.

FitA/B_{His} binds DNA in a sequence-specific manner—DNase I footprinting analysis was employed to determine the sequences in the *fit* promoter region bound by FitA/B_{His}. FitA/B_{His} bound both the “plus” and “minus” strands of fragment *JSW*₂₂₀ (Figure 5a and b). The complex protected the region between – 60 and –17 bp in the minus strand, and the region between – 79 and – 45 bp in the plus strand (Figure 5c). The protected region covers the putative –10 promoter sequence (TATCATT).

These FitA/B_{His} binding studies also revealed DNase I-hypersensitive sites in *JSW*₂₂₀, three on the minus strand at –24 (T), –41 (A) and –57 (G), and three on the plus strand at –47 (T) –55 (A) and –79 (G). Although these results suggest that the DNA bound by FitA/B it undergoes a structural modification, further analysis will be required to determine how FitA/B causes this hypersensitivity.

Repeated attempts to obtain a footprint of FitA_{His} bound to *JSW*₂₂₀ were unsuccessful. FitA_{His} alone marginally protected *JSW*₂₂₀ from DNase I digestion but we were unable to resolve a clear footprint. Failure of FitA_{His} to footprint is likely due to the

lower DNA-binding affinity of FitA in the absence of FitB, and this complex is not sufficiently stable to protect the DNA from cleavage by DNase I.

FitA_{His} and FitA/B_{His} bind the sequence TGCTATCA—The DNase I footprinting experiments localized binding of the FitA/B_{His} complex to a 62-bp region of DNA. This fragment contains the inverted repeat TGCTATCA-N₁₂-TGATAGCA. A double stranded DNA fragment corresponding to this inverted repeat with 4 bp of adjacent sequence was synthesized, termed IR36, fluorescein tagged on the 5' end and analyzed by fluorescence anisotropy/polarization (FP) for its ability to bind FitA_{His} and FitA/B_{His} (Table 3). FitA/B_{His} bound the IR36 fragment with a very high affinity ($K_d = 4.5 \pm 0.7$ nM); in comparison, FitA_{His} bound this fragment with a $K_d = 178 \pm 34$ nM, a decrease of 38-fold. This enhancement of FitA DNA binding by FitB is consistent with our gel shift findings (Figure 6).

We hypothesized that FitA_{His} and FitA/B_{His} were binding to the inverted repeat sequence. To demonstrate this interaction we analyzed the binding properties of oligodeoxynucleotides derived from IR36 with alterations in the inverted repeat sequence. FitA_{His} and FitA/B_{His} bound to fragments with either half site scrambled (IR36-1 and IR36-2, Table 3), but not to the fragment with both half sites scrambled (IR36-3). FitA_{His} bound with equal affinity to the half site mutants and to the wild type fragment. FitA/B_{His} bound these fragments with a 2.5-fold lower affinity, suggesting that unlike FitA_{His}, the FitA/B_{His} complex may be interacting with both half sites.

We also examined the contribution of the 12-bp A/T rich spacer region between the two half-sites to FitA/B_{His} binding. This was done by either replacing the A/T₁₂ intervening sequence with G/C₁₂ (IR36-4) or changing the spacing and phase of the two

half sites with respect to each other. A 2-bp deletion brings the half sites moderately out of phase (IR36-5). A 6-bp deletion orients the half sites on opposite sides of the helix (IR36-6). Finally, a 10-bp deletion removes a whole turn, bringing the half sites back into phase (IR36-7). FitA/B_{His} bound to each of these four fragments with similar affinity (Table 3). These data suggest that FitA and FitA/B bind to the half sites independently.

In order to test this, DNA binding assays were done on smaller oligodeoxynucleotides containing the 8-bp sequence TGCTATCA, which we termed Fit interaction sequence, FitIS, (Table 3). FitA/B_{His} did not bind to the 8-bp fragment alone, so longer fragments were analyzed. We found that FitA/B_{His} bound a single half site if the oligodeoxynucleotide length was 12-bp or greater (data not shown). FitA/B_{His} bound a 16-bp fragment with an affinity of 70.7 ± 11 nM (FitIS16, Table 3). These data indicate that FitA/B_{His} is able to bind the half site independently.

The FitIS sequence is a pseudo-palindrome, with 6 of its 8-bp form an inverted repeat, TGCTATCA. Given the anti-parallel arrangement of the β -strands involved in sequence recognition of RHH proteins (21, 23), we thought FitA/B would bind a perfect palindrome better than the FitIS sequence. There are two possible perfect palindromes that can be made from FitIS with a single base pair change (FitPP16-A/T and FitPP16-G/C). FitA/B_{His} bound the FitPP16-A/T fragment with an affinity of 45.7 ± 2.3 nM (1.5 fold better than FitIS16) but did not bind the FitPP16-G/C fragment.

Taken together these data indicate that FitA and FitB form a complex with each other that binds the sequence TGCTATCA, FitIS. The GC genome has 14 copies of FitID and four copies of TGATATCA, termed Fit perfect palindrome, FitPP (Strain FA1090, <http://www.stdgen.lanl.gov/>). Two copies of FitIS form an inverted repeat in

the *fit* promoter region. All other copies of the FitIS and FitPP occur as single copies. We have yet to determine if FitA/B binds to the other copies to the FitIS or FitPP and what consequence this binding may have *in vivo*.

Discussion

We have provided evidence that FitA binds its own promoter DNA, and that the DNA-recognition activity involves a ribbon-helix-helix motif. In the absence of FitB, FitA forms a homodimer in aqueous solution that binds the DNA sequence TGCTATCA (FitIS) upstream of the *fitA* ORF.

FitB interacts with FitA and increases its DNA binding affinity. In turn, FitA greatly enhances the solubility of FitB. The FitA/B complex has a molecular weight of 98 kDa likely with a stoichiometry of A₄B₄. The FitA/B complex protects a 62-bp region of its promoter from DNase I digestion and the complex binds the FitIS within this region with a 38-fold higher affinity than the FitA homodimer. Moreover, the interaction of FitA/B with DNA was dependent on the RHH domain of FitA. FitA/B binds with a 2.5 fold higher affinity to DNA fragments containing two copies of FitIS than to fragments with a single copy. This difference in affinity was not observed with the FitA homodimer, indicating that the FitA/B complex may be interacting with both of the FitIS halfsites. Additionally, we found that FitA/B was able to interact with palindromic sequence FitPP with higher affinity.

The sequence TGCTATCA (FitIS) occurs 14 times in the GC genome, and the palindromic sequence TGATATCA (FitPP) four times. Of these 18 potential binding

sites, only two are located within promoter regions – those upstream of *fitA*. Of the rest, 12 are located within ORFs and six are found in predicted intergenic regions. These observations raise the question of whether FitA/B interacts with these sequences, and if so, whether such an interaction reflects a transcriptional regulatory function.

These studies demonstrate a novel interaction between a ribbon-helix-helix DNA binding protein and a protein that positively affects its binding affinity. Crystallography studies are ongoing to determine how the interaction between FitA and FitB increases the affinity of FitA for DNA.

Acknowledgment

We would especially like to thank Sunghee Chai, Ken Haak and Jay Mellies for their advice and technical assistance in the preparation of this manuscript. We would also like to thank Shaun Lee, Susan Clary, Nathan Weyand and Eric Barklis for their critical reading and ideas.

References

1. **Handsfield, H., and Sparling, P. (1995) *Principles and Practice of Infectious Diseases*, Vol. 2, 4 ed., Churchill Livingstone, New York.**
2. **Cornelissen, C. N., Biswas, G. D., Tsai, J., Paruchuri, D. K., Thompson, S. A., and Sparling, P. F. (1992) *J Bacteriol* 174, 5788-97.**
3. **Popp, A., Dehio, C., Grunert, F., Meyer, T. F., and Gray-Owen, S. D. (1999) *Cell Microbiol* 1, 169-81.**

4. Chen, T., Grunert, F., Medina-Marino, A., and Gotschlich, E. C. (1997) *J Exp Med* 185, 1557-64.
5. Kallstrom, H., Liszewski, M. K., Atkinson, J. P., and Jonsson, A. B. (1997) *Mol Microbiol* 25, 639-47.
6. Kallstrom, H., Blackmer Gill, D., Albiger, B., Liszewski, M. K., Atkinson, J. P., and Jonsson, A. B. (2001) *Cell Microbiol* 3, 133-43.
7. Bonna, R. A., Yu, R., and Schryvers, A. B. (1995) *Microb Pathog* 19, 285-97.
8. McGee, Z. A., Johnson, A. P., and Taylor-Robinson, D. (1981) *J Clin Invest* 87, 413-422.
9. Merz, A. J., Rifkin, D. B., Arvidson, C. G., and So, M. (1996) *Mol Med* 2, 745-54.
10. Merz, A. J., and So, M. (1997) *Infect Immun* 65, 4341-9.
11. Harvey, H. A., Jennings, M. P., Campbell, C. A., Williams, R., and Apicella, M. A. (2001) *Mol Microbiol* 42, 659-72.
12. Harvey, H. A., Ketterer, M. R., Preston, A., Lubaroff, D., Williams, R., and Apicella, M. A. (1997) *Infect Immun* 65, 2420-7.
13. Meyer, T. F. (1999) *Clin Infect Dis* 28, 433-41.
14. Merz, A. J., and So, M. (2000) *Annu Rev Cell Dev Biol* 16, 423-57.
15. Gray-Owen, S. D. (2003) *Scand J Infect Dis* 35, 614-8.
16. Lin, L., Ayala, P., Larson, J., Mulks, M., Fukuda, M., Carlsson, S. R., Enns, C., and So, M. (1997) *Mol Microbiol* 24, 1083-94.
17. Ayala, P., Lin, L., Hopper, S., Fukuda, M., and So, M. (1998) *Infect Immun* 66, 5001-7.

18. Wang, J., Gray-Owen, S. D., Knorre, A., Meyer, T. F., and Dehio, C. (1998) *Mol Microbiol* 30, 657-71.
19. Hopper, S., Wilbur, J. S., Vasquez, B. L., Larson, J., Clary, S., Mehr, I. J., Seifert, H. S., and So, M. (2000) *Infect Immun* 68, 896-905.
20. Raumann, B. E., Brown, B.M. and Sauer R.T. (1994) *Curr Opin Struct Biol* 4, 36-43.
21. Raumann, B. E., Rould, M. A., Pabo, C. O., and Sauer, R. T. (1994) *Nature* 367, 754-7.
22. Sauer, R. T., Milla, M. E., Waldburger, C. D., Brown, B. M., and Schildbach, J. F. (1996) *Faseb J* 10, 42-8.
23. Phillips, S. E. (1994) *Annu Rev Biophys Biomol Struct* 23, 671-701.
24. Chivers, P. T., and Sauer, R. T. (2000) *J Biol Chem* 275, 19735-41.
25. Somers, W. S., and Phillips, S. E. (1992) *Nature* 359, 387-93.
26. Gomis-Ruth, F. X., Sola, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R., Gonzalez, A., Espinosa, M., del Solar, G., and Coll, M. (1998) *Embo J* 17, 7404-15.
27. Knight, K. L., and Sauer, R. T. (1992) *Embo J* 11, 215-23.
28. Chivers, P. T., and Sauer, R. T. (1999) *Protein Sci* 8, 2494-500.
29. Schreiter, E. R., Sintchak, M. D., Guo, Y., Chivers, P. T., Sauer, R. T., and Drennan, C. L. (2003) *Nat Struct Biol* 10, 794-9.
30. Phillips, K., and Phillips, S. E. (1994) *Structure* 2, 309-16.
31. Chivers, P. T., and Sauer, R. T. (2002) *Chem Biol* 9, 1141-8.

32. Golovanov, A. P., Barilla, D., Golovanova, M., Hayes, F., and Lian, L. Y. (2003) *Mol Microbiol* 50, 1141-53.
33. Chai, S., and Alsonso, J. C. (1996) *Nucleic Acids Res* 24, 282-8.
34. Newberry, K. J., and Brennan, R. G. (2004) *J Biol Chem* 279, 20356-62.
35. Brown, B. M., Milla, M. E., Smith, T. L., and Sauer, R. T. (1994) *Nat Struct Biol* 1, 164-8.

TABLE 1. Oligonucleotide primers used for construction of probes for gel retardation and footprint assays and Fluorescence anisotropy.^a

PETFit.NcoI	CCATGGCTTCTGTTGTGATTAG
FitA.XhoI-3'	CTCGAGCAAAGAAACCTCGTTATCAG
PETFit3'.XhoI	CTCGAGGTGCCACGGATTGAACACCG
FitA.R7A.NcoI	CCATGGCTTCTGTTGTGATCGCGAATTTATCC
Stbpre402	TGCCTTTCCTTCCTGAAAAAAGTTTTTC
StbL108	TACCGAATCGGAAGCGGCAC
JSW65	ATTTATCAGAAAGCGGGAGGGGAGACG
JSWanti-65	CGTCTCCCCTCCCTCTTTCTG
JSW64	ACAGAATTCAGTCCCTGAAAGCCGC
JSWanti-64	GCGGCTTTCAGGGACTGAATTCTG
StbR205	GCACGGAATTTGATTGCGTTG
IR 36 fwd	F-AGATTGCTATCATTTTTTTTTATTTTGATAGCATTG
IR 36 rev	CAAATGCTATCAAATAAAAAAAAAATGATAGCAATCT
IR36-1 fwd	F-AGATATACTGTCTTTTTTTTTATTTTGATAGCATTG
IR36-1 rev	CAAATGCTATCAAATAAAAAAAAAAACTATCGTATCT
IR36-2 fwd	F-AGATTGCTATCATTTTTTTTTATTTGACAGTATTTG
IR36-2 rev	CAAAACTATCGTAAATAAAAAAAAAATGATAGCAATCT
IR36-3 fwd	F-AGATATACTGTCTTTTTTTTTATTTGACAGTATTTG
IR36-3 rev	CAAATACTGTCAAATAAAAAAAAAAAGACAGTATATCT
IR36-4 fwd	F-AGATTGCTATCAGGGGGGGGGGGGTGATAGCATTG
IR36-4 rev	CAAATGCTATCACCCCCCCCCCCTGATAGCAATCT
IR36-5 fwd	F-TAGATTGCTATCATTTTTTTATTTTGATAGCATTG
IR36-5 rev	ACAAATGCTATCAAATAAAAAAATGATAGCAATCTA
IR36-6 fwd	F-TTTTAGATTGCTATCATTATTTTGATAGCATTGTTT
IR36-6 rev	AAACAAATGCTATCAAATAATGATAGCAATCTAAA
IR36-7 fwd	F-TTTTTAGATTGCTATCATTTGATAGCATTGTTTAT
IR36-7 rev	AATAACAAATGCTATCAAATGATAGCAATCTTTTTT
FitIS16 fwd	F-AGATTGCTATCATTTT
FitIS16 rev	AAAATGATAGCAATCT
FitPP16 A/T fwd	F-AGATTGATATCATTTT
FitPP16 A/T rev	AAAATGATATCAATCT
FitPP16 G/C fwd	F-AGATTGCTAGCATTTT
FitPP16 G/C rev	AAAATGCTAGCAATCT

^aAll oligonucleotides are listed 5' to 3'. F- = 5' fluorescein tag.

TABLE 2. Biophysical data for FitA_{his} and FitA(R7A)_{his}

Protein	T_m (2 μ M)	T_m (5 μ M)	T_m (10 μ M)	Mw aa	Mw AUC
FitA _{his}	53.9	57.5	60.4	18997	18573
FitA(R7A) _{his}	53.6	56.7 (4 μ M)	nd	18827	18255

TABLE 3. Fluorescence anisotropy with IR-prox and FitA_{his} or FitA/B_{his}

Fragment name	Sequence of top strand oligonucleotide ^a	K _d FitA _{his} (nM) ^{bc}	K _d FitA/B _{his} (nM) ^c
IR36	5' -AGATTGCTATCATTTTTTTTATTTT <u>GATAGC</u> ATTTG	176 ±34	4.5 ±0.7
IR36-1	5' -AGATa <u>tactgtc</u> TTTTTTTTATTTT <u>GATAGC</u> ATTTG	155 ±33	14.1 ±2.9
IR36-2	5' -AGATTGCTATCATTTTTTTTATTTgacagtatTTTG	132 ±25	12.4 ±1.5
IR36-3	5' -AGATa <u>tactgtc</u> TTTTTTTTATTTgacagtatTTTG	>3000	>3000
IR36-4	5' -AGATTGCTATCAgggggggggggTGATAGC <u>ATTTG</u>	n.d.	3.5 ±0.5
IR36-5	5' -tAGATTGCTATCATTTTTtATTTT <u>GATAGC</u> ATTTGt	n.d.	4.2 ±0.8
IR36-6	5' -tttAGATTGCTATCAttATTTT <u>GATAGC</u> ATTTGttt	n.d.	7.6 ±1.2
IR36-7	5' -tttttAGATTGCTATCAttTGATAGC <u>ATTTG</u> tttat	n.d.	3.0 ±0.6
FitIS16	5' -AGATTGCTATCATTTT	n.d.	70.7 ±11
FitPP16-A/T	5' -AGATTGaTATCATTTT	n.d.	45.7 ±2.3
FitPP16-G/C	5' -AGATTGCTAgCATTTT	n.d.	>500

^a Top strand of double stranded oligodeoxynucleotide DNA target listed (UNDERLINE - inverted repeat, lower case – sequence changed from wildtype). 5' end of top strand was fluorescein tagged.

^b n.d. – not determined

^c K_d calculated using a hyperbolic curve fit algorithm (KaleidaGraph™), error is mP value deviation from the curve.

FIG 1. Bioinformatics analysis of FitA. (A) Clustal W alignment of the first 65 amino acids of FitA with homologues deduced from the genome sequences of other bacteria. Note that these homologues are hypothetical proteins. Homology is markedly reduced in the regions not pictured here. White text on Black boxes: residues with 100% sequence similarity; Black text on Gray boxes: denote highly conserved residues; Black text on white boxes: non-conserved residues. Conserved arginine at position 7 is denoted by arrow. (B) Chou-Fasman secondary structure prediction of the first 45 amino acids of FitA (the α -helix region is denoted by the black box above; the β -sheet region is denoted by the dark gray box; and the turn region is denoted by the light grey box).

FIG 2. Analysis of the DNA-binding activity of FitA_{His}. (A) Purified his tagged Fit proteins separated by SDS-PAGE and stained with Coomassie blue. Lane 1: molecular weight markers; Lane 2: purified FitA_{His}; Lane 3: FitA(R7A)_{His}; Lane 4: FitA/B_{His}; Lane 5: FitA(R7A)/B_{His}. (B) Gel retardation assay showing binding of varying amounts of FitA_{His} to the ³²P-labeled *fit_{pro}* probe in the presence of poly dIdC (see diagram below). (C) Deletion analysis of the region of *fit_{pro}* bound by 2.2 μ M of FitA_{His}. (D) Gel retardation analysis of the binding affinity of FitA_{His} and FitA(R7A)_{His} to ³²P-labeled fragment *JSW₂₂₀* in the presence of poly dIdC. (E) Gel retardation analysis of the binding affinity of the FitA/B_{His} and FitA(R7A)/B_{His} complexes to ³²P-labeled fragment *JSW₂₂₀* in the presence of poly dIdC. (F) Schematic of the *fit* locus and its upstream sequences (bottom diagram), and DNA fragments encoding segments of this region (Fragments 1-5; *JSW₂₂₀*) used in gel-retardation and footprinting assays.

FIG 3. Circular dichroism analysis of FitA_{His} (Black) and FitA(R7A)_{His} (Grey).

FIG 4. Gel retardation analysis of the DNA binding activity of varying amounts of the FitA/B_{his} complex with ³²P-labeled *fit_{pro}* (A), and FitA/B_{His} with ³²P-labeled fragments 1-5 at 1.53 μM (B).

FIG 5. DNaseI footprinting analysis of the FitA/B_{His} complex to the “plus” (A) and “minus” (B) strand of *JSW₂₂₀*. Protein concentrations used were 0.675 μM, 1.35 μM and 2.73 μM in (A) and 0.43 μM, 0.675 μM, 1.35 μM and 2.73 μM in (B). The region protected from DNaseI digestion is indicated by the black line and DNaseI hypersensitive cleavage sites created by FitA/B_{His} binding are shown with asterisks. Black rectangle marks the extent of protection from DNase I cleavage. (C) Diagram of *fit* promoter sequences protected by FitA/B_{His} in the “plus” strand (Gray box) and “minus” strand (open box). “IR_{prox}” and divergent arrows indicate the inverted repeat at the putative -10 sequence.

FIG 6. Binding isotherms of FitA_{His} and FitA/B_{His} to the IR36 DNA fragment. FitA_{His} (●) bound this fragment with a K_d of 176 nM and FitA/B_{His} (□) bound with a K_d of 4.5 nM, a 38-fold higher affinity. These binding data were generated using the fluorescence anisotropy/polarization based DNA-binding assay, as described in Experimental Procedures. Note that FitA binding data above 150 nM is shown in inserted graph.F

Figure 1

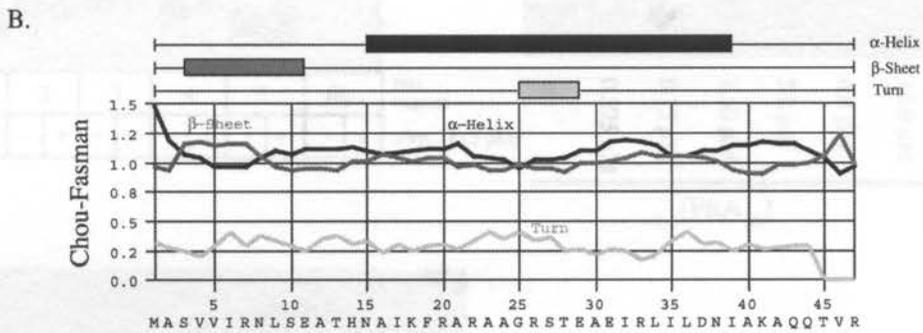
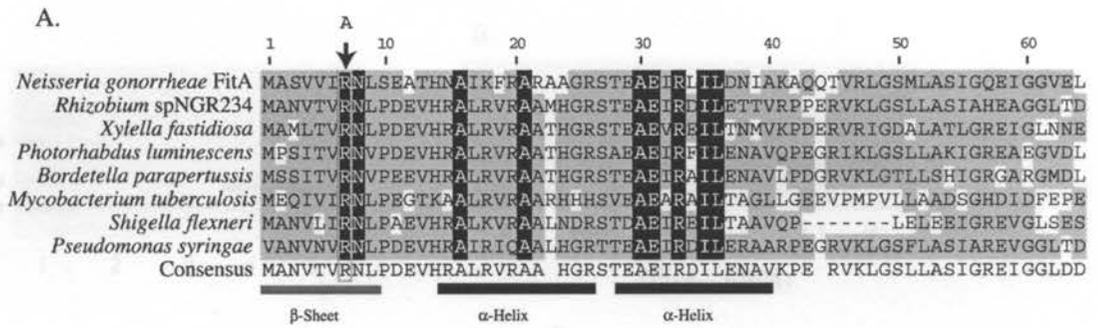


Figure 2

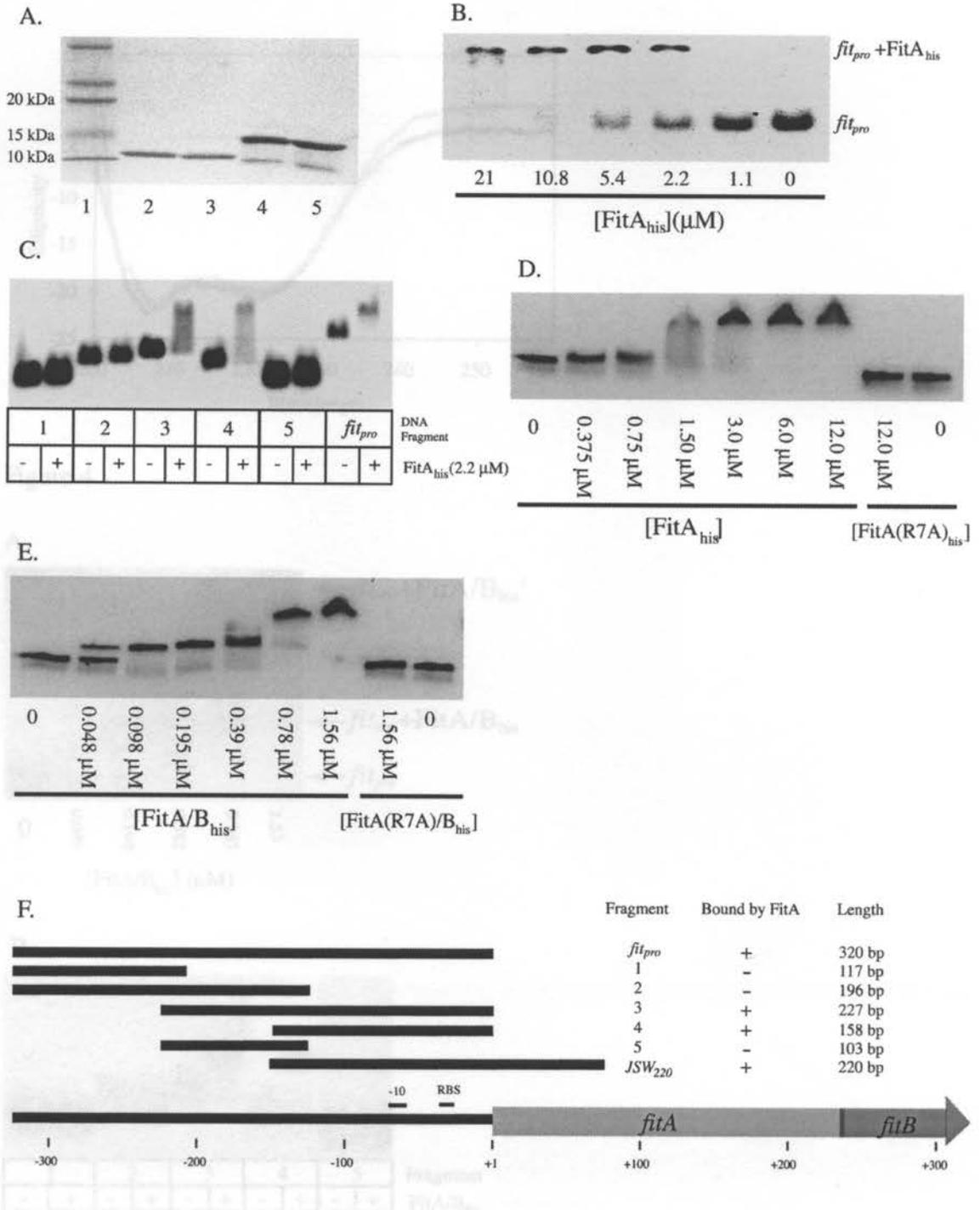


Figure 3

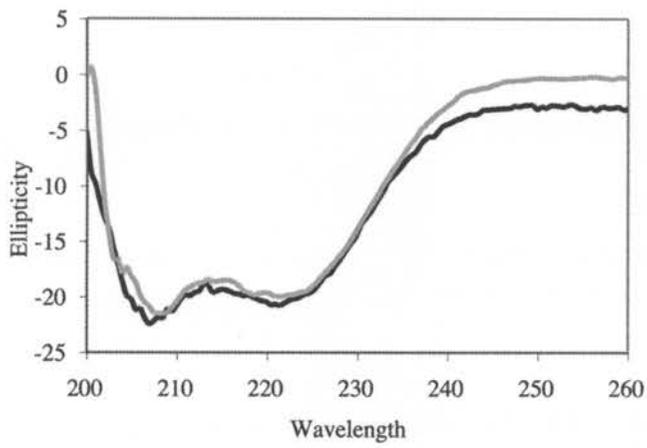
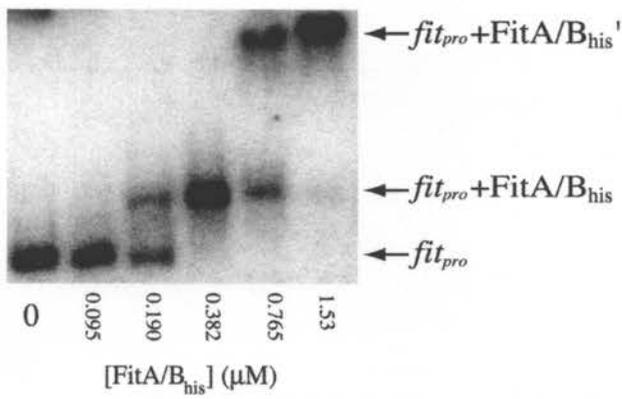


Figure 4

A.



B.

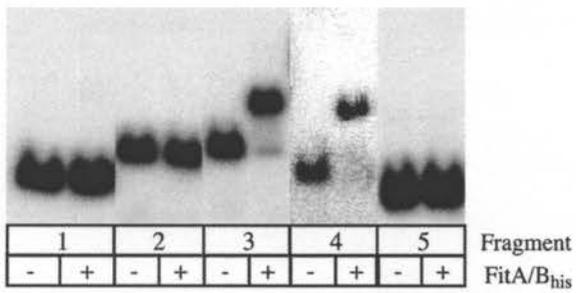


Figure 5

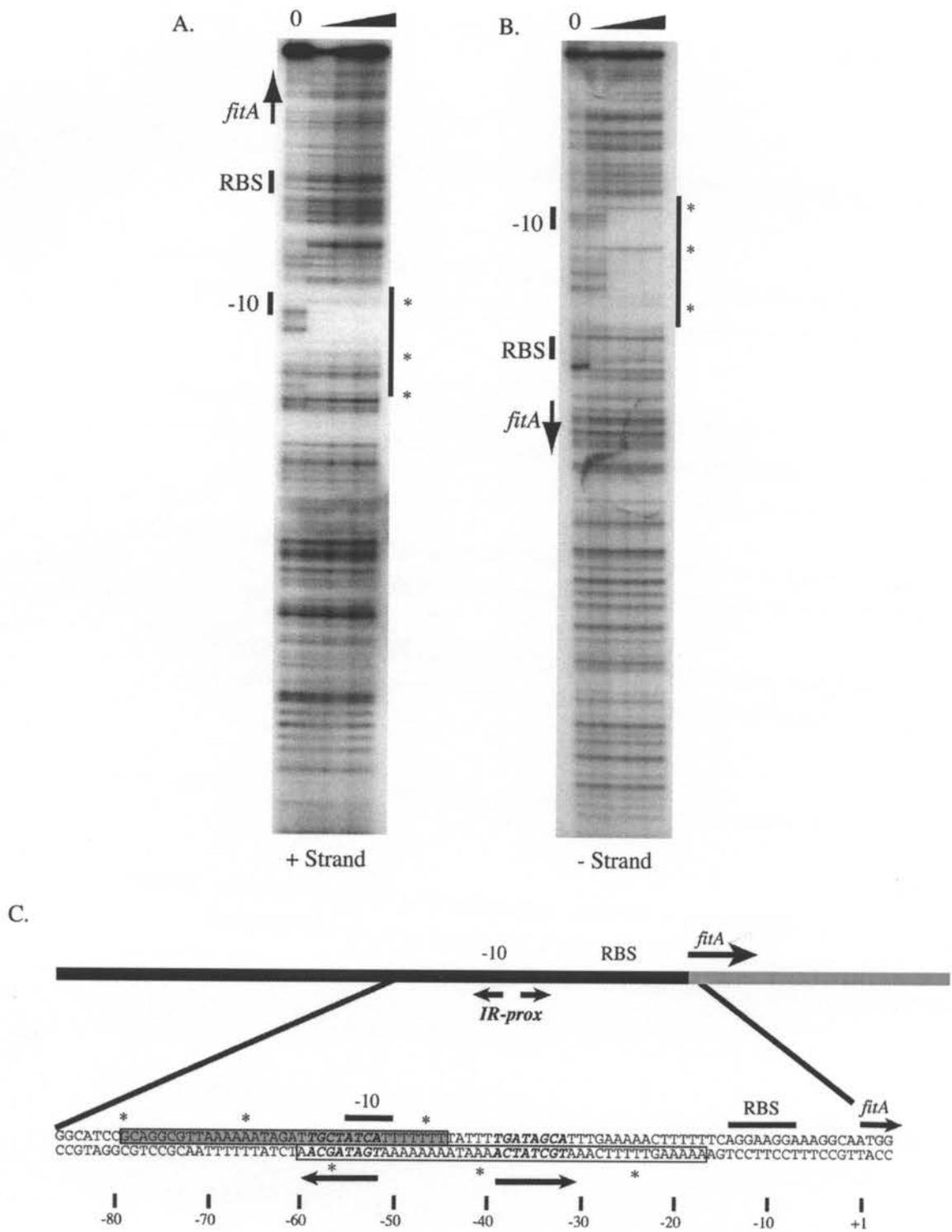
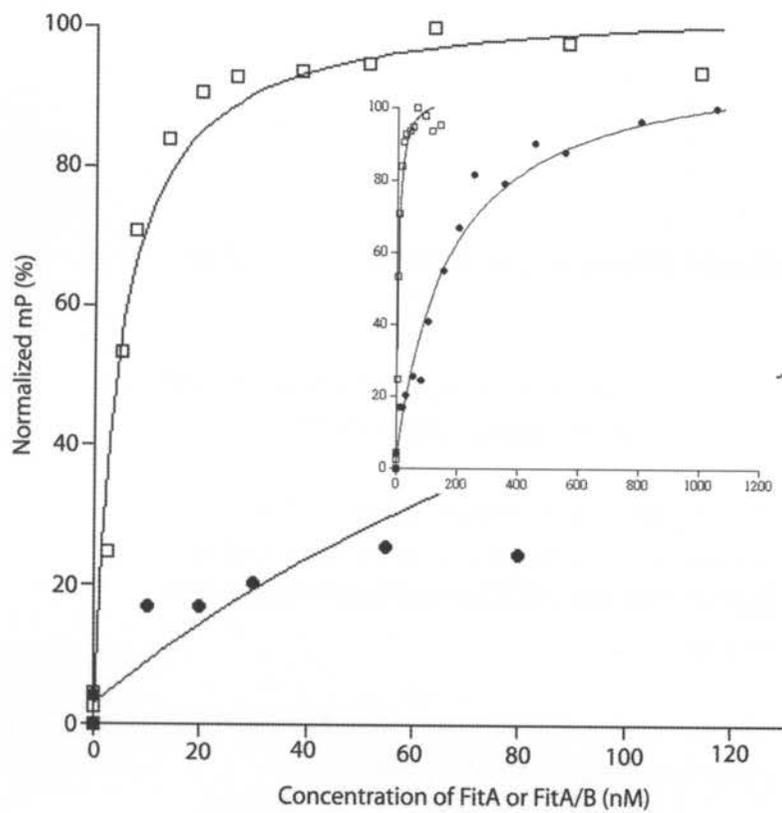


Figure 6



Chapter 3: Manuscript II

**Structure of FitAB from *Neisseria gonorrhoeae* bound to DNA
reveals a heterodimer of PIN and ribbon-helix-helix domains**

Structure of FitAB from *Neisseria gonorrhoeae* bound to DNA reveals a heterodimer of PIN and ribbon-helix-helix domains

Kirsten Mattison^{1,3}, J. Scott Wilbur², Magdalene So² and Richard G. Brennan¹

Departments of ¹Biochemistry and Molecular Biology and ²Molecular Microbiology and Immunology, Oregon Health & Science University, Portland, OR, USA, 97239. ³Present address: Bureau of Microbial Hazards, Health Products and Food Branch, Health Canada, Ottawa, ON, Canada, K1A 0L2.

Character count of manuscript: 43,780 (with spaces), 36, 897(without spaces).

Running title: Structure of FitAB from *N. gonorrhoeae*

Address correspondence to: Richard G. Brennan, Department of Biochemistry and Molecular Biology, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Mail Code L224, Portland, OR, 97239-3098

Tel. 503 494-4427

Fax. 503 494-8393

E-Mail: brennanr@ohsu.ed

Neisseria gonorrhoeae is a sexually transmitted pathogen that initiates infections in humans by adhering to the mucosal epithelium of the urogenital tract. The bacterium then enters the apical region of the cell and traffics across the cell to exit into the subepithelial matrix. Mutations in the fast intracellular trafficking (*fitAB*) locus cause the bacteria to transit a polarized epithelial monolayer more quickly than the wild type parent and to replicate within cells at an accelerated rate. Here, we present the crystal structure of the FitAB complex bound to a DNA fragment from its promoter region. FitA binds DNA through its ribbon-helix-helix motif and is tethered to FitB by the insertion of a four turn α helix into an extensive FitB hydrophobic pocket. FitB is composed of a PIN (PiIT N-terminus) domain, with a central, twisted, 5 stranded parallel β sheet that is open on one side and flanked by five α helices. A tetramer of FitAB heterodimers binds to the 36 bp DNA fragment, explaining how FitB enhances the DNA binding affinity of FitA. (175 words maximum)

Subject categories: Structural Biology; Microbiology and Pathogens

Keywords: DNA binding, PIN domain, RHH domain

Introduction

Neisseria gonorrhoeae (GC) is the agent of the sexually transmitted disease, gonorrhea. The mechanisms used by GC to initiate infection have been very well characterized. Gonococci adhere via a multi-step cascade and subsequently enter cells forming the epithelial barrier of the urogenital tract, traffic across these cells and exit into the subepithelial matrix (Apicella et al., 1996; Merz and So, 2000). Although studies have identified many of the molecular mechanisms used by GC to adhere to and enter cells, our knowledge of the mechanisms that operate in the later stages of infection is limited.

GC are able to survive and grow within epithelial cells (Hopper et al., 2000b); they also traverse the epithelial monolayer to infect the stromal tissue of the subepithelium (Merz and So, 2000). The immune response to bacteria in the subepithelium produces the inflammation and purulent discharge characteristic of gonorrhea (Edwards and Apicella, 2004; Holmes et al., 1971). On occasion, GC establish a carrier state in which an asymptomatic individual harbors culturable and transmissible bacteria. These carriers are key to the spread of gonococcal disease, as humans are the only known reservoir for these organisms (Turner et al., 2002). The mechanisms by which GC maintains this persistent state is unknown. One hypothesis is that the organism resides within the epithelial cells instead of crossing into the subepithelium, thus evading the host immune response. The gene product(s) that affect GC intracellular growth and transcytosis are therefore important for the maintenance of gonococci in the human population.

The *fitAB* operon was identified in a screen for GC mutants with a fast intracellular trafficking phenotype across polarized epithelial monolayers (Hopper et al., 2000b). A GC mutant that lacks *fitAB* grows normally extracellularly, but has an accelerated rate of intracellular replication with a concomitant increase in the rate at which this mutant traverses a monolayer of polarized epithelial cells. Thus, either FitA or FitB, or their complex, is hypothesized to slow intracellular replication and intracellular trafficking of GC.

FitA is an 8.4 kDa protein with a predicted N-terminal ribbon-helix-helix (RHH) DNA binding motif (Raumann et al., 1994; Wilbur et al., 2005). FitB is a 15.3 kDa protein with a predicted PIN (PilT-N terminus) domain according to the BLAST search tool at <http://www.ncbi.nlm.nih.gov/BLAST> (Altschul et al., 1990). The function of the PIN domain is unknown, however many proteins that contain a PIN domain are thought to perform roles in nucleic acid metabolism including synthesis and remodelling. In genome studies on archaea and thermophilic bacteria, sequences predicted to encode PIN-domain containing proteins are found in regions predicted to encode DNA polymerases, helicases and nucleases (Makarova et al., 2002). In addition, the Dis3p exonucleases from *Saccharomyces cerevisiae* and nonsense-mediated mRNA decay (NMD) proteins in *Caenorhabditis elegans* are predicted to have PIN domains (Clissold and Ponting, 2000; Makarova et al., 1999). Bi-cistronic operons with both a FitA and a FitB homologue are limited to plasmid-encoded gene clusters that control plasmid stability (Freiberg et al., 1997; Hanekamp et al., 1997). In *Salmonella dublin*, when the FitB-like protein VagD was expressed in the absence of the FitA-like protein VagC, the bacteria failed to grow on minimal medium and a high proportion of cells were elongated

(Pullinger and Lax, 1992). Such defects suggest that the *S. dublin* FitB homologue VagD inhibits cell division *in vivo*, while the GC FitAB slows replication only when the bacteria are within epithelial cells (Hopper et al., 2000a).

FitA and FitB form a heterodimer, and copurify after overexpression in *Escherichia coli*. The FitAB complex binds DNA from the *fitAB* upstream region with high affinity (Wilbur et al., 2005). In our current model the FitAB complex binds to the *fitAB* promoter when GC are in an extracellular environment. This results in both sequestration of FitAB and repression of *fitAB* transcription. Upon invasion, we propose that FitAB is released from the DNA to slow both GC replication and transcytosis by an undefined mechanism.

In order to understand the structural basis of FitAB function, we determined the x-ray crystallographic structure of a FitAB complex bound to a 36 base pair DNA fragment from the *fitAB* upstream region. Four FitA and four FitB proteins form a unique tetramer of heterodimers structure that explains the ability of FitAB to bind DNA with higher affinity than FitA alone (Wilbur et al., 2005). As expected, FitA was found to contain a ribbon-helix-helix domain that forms specific contacts with the DNA. The structure of FitB reveals a tightly packed α - β - α fold with structural homology to PIN domains, as predicted by our primary sequence analysis. However, unlike other PIN containing proteins, including its closest structural homologues, FitB does not display nuclease activity. An arginine residue from FitA blocks the putative active site of FitB in the FitAB heterodimer. From these structural studies, we hypothesize that FitB might slow intracellular GC replication when this FitA-mediated inhibition of its nuclease activity is released by complex dissociation.

Results and Discussion

Overall structure of FitAB

The structure of the FitB protein complexed with a C-terminal fragment of FitA (FitAcB) was determined to 1.8 Å resolution by multiple wavelength anomalous diffraction using selenomethionine substituted proteins (Materials and Methods, Table I). The asymmetric unit of the native crystals contains 158 residues: the C-terminus of FitA (FitcA, residues 46-64) and one molecule of FitB (residues 1-139). The R_{work} and R_{free} are 19.1% and 22.4%, respectively. This structure was used as a model to solve the structure of the full length FitAB complex bound to a 36-bp DNA molecule to 3.0 Å resolution by molecular replacement (Materials and Methods, Table I). The asymmetric unit of these crystals contains 828 protein residues: four molecules of FitA (residues residues 2-69, 2-65, 2-68, 2-64) and four molecules of FitB (residues 1-143, 1-140, 1-143, 1-140). The R_{work} and R_{free} are 21.2% and 26.9%, respectively.

The FitA monomer has an extended structure, with the topology $\beta 1-\alpha 1-\alpha 2-\alpha 3$ (Figure 1A). Electron density is visible for the intact N-terminus of FitA, beginning at Ala2. We have confirmed by amino terminal sequencing that Met1 is not present in our preparation, presumably it is N-formyl-methylated and cleaved *in vivo* during overexpression of the proteins in *Escherichia coli* (data not shown). At the C-terminal end, variable electron density is seen for the four molecules of the asymmetric unit, the final 9-14 residues are not visible (depending on the monomer) and are thought to be flexible in the crystals. The first 45 residues of this protein ($\beta 1-\alpha 1-\alpha 2$) are highly

homologous to the ribbon-helix-helix (RHH) class of DNA binding proteins, which includes the bacteriophage P22 proteins Mnt and Arc (Figure 1B) (Burgering et al., 1994; Raumann et al., 1994; Wilbur et al., 2005). An arginine found in $\beta 1$ of the RHH proteins is conserved in FitA (Figure 1B, highlighted in blue).

FitB forms a compact domain with an $\alpha/\beta/\alpha$ fold (Figure 1A). This protein consists of a central 5-stranded parallel β sheet with four α helices packed on one side of the sheet and three α helices on its other side, with topology of $\beta 1-\alpha 1-\alpha 2-\beta 2-\alpha 3-\alpha 4-\beta 3-\alpha 5-\alpha 6-\beta 4-\alpha 7-\beta 5$ (Figure 1A). Electron density for the entire FitB protein is visible, with only engineered histidine residues at the C-terminus of FitB not observed in the structure (missing 3/6 His in 2 monomers and 6/6 His in 2 monomers). Searches, using both the DALI server (Holm and Sander, 1995) and the protein structure comparison service SSM (Krissinel and Henrick, 2004) at the European Bioinformatics Institute (www.ebi.ac.uk), found structural homologues of FitB in PIN-domain containing proteins. The canonical PIN domain is found in the PiIT N-terminus, and the closest FitB structural homolog is PAE2754 from *Pyrobaculum aerophilum* (Figure 1B)(Arcus et al., 2004; Kaiser and Wall, 1999). The functional significance of this domain is unknown. However, the PIN domain is found in a wide variety of systems, from bacterial FitB-like genes thought to be involved in plasmid maintenance, to the yeast Dis3p exonucleases (Clissold and Ponting, 2000; Freiberg et al., 1997; Hanekamp et al., 1997). Despite a lack of primary sequence similarity, the four acidic residues absolutely conserved among PIN domains are present in FitB (Figure 1B, highlighted in red).

In addition to the RHH and PIN domain containing proteins, there are a group of prokaryotic proteins with a high level of sequence homology to FitAB (Figure 1B). These

typically consist of both a FitA and a FitB homologue in a conserved operon organization and little is understood about their function, other than that they play a general role in plasmid stability and/or partition (Brendler et al., 2004; Freiberg et al., 1997; Hanekamp et al., 1997; Lan et al., 2003; Pullinger and Lax, 1992). The structure of the FitAB complex is likely to predict the structures of these proteins. We propose that the biochemical function of this group of proteins within prokaryotic cells is the same as that performed by FitAB. Sequence alignments of FitA and FitB with two examples of such systems are shown in Figure 1B.

The FitA-FitB interface

The FitA and FitB proteins form a tightly associated dimer and the FitAB structure reveals the heterodimerization interface. The C-terminal portion of FitA (residues 46-78) does not have obvious sequence homologues or predicted secondary structure (Wilbur et al., 2005). However, residues 46 through 60 of this C-terminal region form a four turn α helix (Figure 1A), and the interface between the C-terminus of FitA and FitB buries 1900 \AA^2 of accessible surface area. The extensive buried surface area of the FitAB complex suggests that the C-terminal FitA helix is tightly bound to FitB. The FitA helix fills a hydrophobic pocket in FitB resulting in a globular heterodimeric domain (Figure 1A and C). The functional significance of the increased stability of FitA in the FitAB complex is underscored by the observation that the FitAB complex binds DNA with 40-fold higher affinity than FitA alone (Wilbur et al., 2005).

The FitA-FitB heterodimer interface is formed predominantly by contacts between $\alpha 3$ and the C-terminal extended coil region of FitA and $\alpha 1$, $\alpha 2$, and $\alpha 4$ of FitB.

Much of the interaction can be attributed to shape complementarity; a number of individual hydrophobic residues can be identified that contribute to the interface. For example, the nonpolar side chain of residue Leu52 of the FitA helix sits between Ile9 and Pro12 of FitB α 1 (Figure 1C). At the C-terminus of FitA, outside of the helical region, the side chains of residue Ile59 contacts the side chains of residues Val58 and Leu59 from α 4 of FitB (Figure 1C). Other residues of the α 4 and α 4- β 3 turn of FitB that are important components of the dimer interface are Ile67, Leu70 and Phe71, which contact Leu48, Met51 and Ile55 of FitA (Figure 1C).

The FitA helix is also stabilized by four ionic interactions, which likely serve to orient the two molecules properly, thereby facilitating a tight association between the complementary hydrophobic surfaces (Figure 1C). Interestingly, two of these electrostatic interactions are analogous to the charge clamp described in structures of the human nuclear receptors bound to co-activator helices (Nolte et al., 1998; Watkins et al., 2003). In the charge clamp, charged residues engage in favourable electrostatic interactions with the oppositely charged dipole of the helical termini. The guanidinium group of FitA residue Arg47 interacts with the backbone carbonyl of residue Asp26 at the C-terminal end of the FitB α 2 helix. The N ζ of FitB residue Lys55 interacts with a backbone carbonyl of residues Glu60 and Glu61 at the C-terminal end of the FitA helix. Other stabilizing electrostatic interactions occur between the side chains of FitA Glu58 and FitB Arg62 and FitA Glu63 and FitB Arg14.

Overall structure of the FitAB-IR36 complex

Four FitAB heterodimers form a unique tetramer structure that binds to one 36 base pair DNA fragment (Figure 2). FitA proteins from complexes I and IV form a tightly associated domain that binds to one of the inverted repeat half sites, as do the FitA portions of complexes II and III. The FitB-FitB dimer interfaces, by contrast, are formed between proteins from complexes I and II or complexes III and IV (Figure 2A).

The FitA homodimer is predicted to be extremely stable, with a buried accessible surface area of 2850 \AA^2 . This compact dimer is characteristic of the RHH DNA binding proteins (Burgering et al., 1994) and involves almost every amino acid residue in $\beta 1$, $\alpha 1$ and $\alpha 2$ of FitA (Figure 3A and B). The $\beta 1$ strands from each FitA monomer combine to form a two stranded anti-parallel β sheet, in which backbone carbonyls form hydrogen bonds with the backbone amides from the neighboring strand (Figure 1A). The two α helices of the RHH domain also contribute to the extensive interface. Thr13, Ala16 and Ile17 of $\alpha 1$ from one monomer form hydrophobic contacts to Leu36, Ile39 and Ala40 of the other monomer (Figure 3B). This interface occurs twice in each RHH domain. In addition to the van der Waals contacts, a salt bridge is formed between the $N\epsilon$ of Arg20 from $\alpha 1$ and the $O\epsilon$ of Gln43 from $\alpha 2$ (Figure 3B).

The FitB homodimer also buried a large accessible surface area, 1870 \AA^2 . At this interface, $\alpha 3$ from one FitB monomer contacts $\alpha 5$ from the other FitB monomer. As above for FitA, there are two such interfaces per homodimer. Phe78 in the $\alpha 4$ - $\alpha 5$ loop is a key residue for FitB-FitB dimerization. The backbone amide and carbonyl atoms of Phe78 form hydrogen bonds to their counterparts in the adjacent monomer and in addition, the Phe78 aromatic ring contributes to the hydrophobic interface by

approaching the C β of Ala37 from α 3 (Figure 3C). Other residues that form key hydrophobic contacts are Ala41, Leu45 and Ala48 from α 3 and Tyr86, Ala87 and Ser91 from α 5 (Figure 3C). Arg44 and Glu80 form a salt bridge between the two helices that also serves to stabilize this extensive interface.

Together, the interfaces depicted in Figures 1C, 3A, 3B and 3C create a stable tetramer of FitAB heterodimers that has been observed in solution studies (Wilbur et al., 2005). This structure accounts for the differences in DNA binding affinity observed for FitA versus FitAB complexes. FitA dimers bind the IR36 site with an affinity of approximately 180 nM, while the FitAB tetramer of heterodimers binds with an affinity of 4.5 nM (Wilbur et al., 2005). The increase in stability provided by FitB to the tetrameric complex easily explains this 40-fold increase in DNA binding affinity, even though FitB does not interact directly with the DNA molecule (Figure 2B).

The IR36 DNA fragment is shown in Figure 2C. This sequence is found upstream of the *fitAB* operon in *N. gonorrhoeae*. The inverted repeat half sites were defined in biochemical studies as the specific bases required for FitAB binding to this region (Wilbur et al., 2005). In our structure, the two FitA β sheets bind on the same face of the IR36 DNA molecule (Figure 2B) however, this positioning not necessary for high affinity DNA binding. Inverted repeats with various spacer lengths between the half sites all bind FitAB with high affinity (Wilbur et al., 2005). It is possible that the flexible loop joining FitA α 2 and α 3 would allow rotation of the tetramer complex to accommodate the observed binding with other half site orientations.

The IR36 fragment is unusual because the two half sites are separated by a long (14 base pair) spacer region of AT-rich DNA. Sequences like this, containing four or

more consecutive A·T base pairs are known as A-tracts (Han et al., 1997). These sequences adopt a slightly different structure than typical B-form DNA in that they are essentially straight and rigid (DiGabrieli et al., 1989; Leroy et al., 1988). In addition, these sequences deviate from B-form DNA by having a compressed minor groove and a shorter helical repeat of only 10 bp (Alexeev et al., 1987; Han et al., 1997). The rigidity of A-tracts is predicted to allow for sharp bends at their edges (Han et al., 1997), however we do not see any pronounced local bending in the structure of IR36 (Figure 2A and 2B). In addition, the central 14 base pairs do not appear to be more straight than the rest of the molecule; there is a gradual 1° bend per base pair that is evenly spread throughout the sequence. The central sequence shares only one of the characteristics said to be found in A-tract DNA, these A·T base pairs have a compressed minor groove when compared to the flanking sequences. Minor groove widths in the A-tract range from 2.9 to 4.6 Å, while in the flanking sequences the width of the minor groove ranges from 5.9 to 8.9 Å.

FitA DNA contacts

FitA makes very few specific contacts to the DNA, which provides additional clarification of the biochemical result that FitA dimers bind IR36 with 40-fold lower affinity than the tetramer of FitAB complexes (Wilbur et al., 2005). All contacts to base pairs are mediated by FitA residues from the conserved β sheet and as predicted, the highly conserved Arg7 is crucial for DNA recognition (Figure 4A). Arg7 from each FitA subunit contacts the 5' most guanine base of the inverted repeat sequence as well as the thymine base on the 5' side of this guanine (Figure 4A). In subunits I and IV, this thymine base forms a water-mediated hydrogen bond to Asn8 as well. The other specific

contact is a van der Waals interaction between Val5 from FitA and the thymine from the T/A sequence central to each inverted repeat (Figure 4A). In addition to these base contacts, a number of residues from the amino terminal end of FitA $\alpha 2$ contact the phosphate-sugar backbone of the DNA molecule. The side chains of Arg33 and Ser27 from all FitA subunits are involved in such interactions, as are the amide nitrogens of Thr28 and Glu29 (Figure 4A). Figure 4B highlights the interactions between $\beta 1$ from FitA IV and bases 30' through 32', and provides visual confirmation of the important role of Arg7 in making direct base contacts.

FitB forms a PIN domain

FitB has a high degree of structural homology to the PIN domain containing protein PAE2754. 93 C α atoms of these two proteins overlay with an r.m.s.d. of 1.9 Å (Figure 5A). PIN domains contain four highly conserved acidic residues that cluster at the C-terminal end of the β sheet and form a negatively charged pocket in the center of the molecule (Arcus et al., 2004; Hwang et al., 1998). This acidic pocket of the flap endonuclease-1 is the active site for its exonuclease activity (Hwang et al., 1998) and that of PAE2754 has been proposed to carry out an exonuclease function as well (Arcus et al., 2004). In FitB these conserved residues are Asp5, Glu42, Asp104 and Asp122 and they cluster to form an acidic pocket just as those from PAE2754 (Figure 5B).

We performed *in vitro* assays to characterize any nuclease activity that might be associated with FitAB, however we were unable to detect any nucleic acid degradation in the presence of the FitAB complex. Unlike homologous exonucleases, FitB has a dimerization partner, FitA. An arginine residue at the C-terminus of FitA is located in the FitB acidic pocket, potentially blocking this region from performing its enzymatic

function (Figure 5C). We propose that the FitB nuclease is activated in the cell when the FitA-FitB complex dissociates in response to some unknown signal. Preliminary experiments with fragments of FitAB or FitB alone have also failed to reveal this activity, future studies will focus on elucidation of the appropriate target nucleic acid for FitB.

In conclusion, FitA and FitB form a heterodimer in which FitA is the DNA binding domain and FitB may have nuclease activity that is blocked by the presence of FitA. Four such heterodimers associate into a tetrameric structure that binds to the IR36 sequence from the *fitAB* promoter region with high affinity. Many PIN domain-containing proteins are involved in some aspect of nucleic acid metabolism and/or remodelling, with the prokaryotic FitAB and its homologues responsible for controlling rates of DNA replication and/or plasmid maintenance (Freiberg et al., 1997; Hanekamp et al., 1997; Hopper et al., 2000b). Future studies on the activity of FitB will help us understand both generally how PIN domains control such diverse processes as replication and nonsense mediated mRNA decay and specifically the role of FitAB in GC virulence (Clissold and Ponting, 2000; Hopper et al., 2000b).

Materials and Methods

Protein preparation, crystallization and intensity data collection

FitA and FitB were overexpressed in *Escherichia coli* using a pET28b vector (Invitrogen) that encodes the intact FitAB operon (Wilbur et al., 2005). This vector incorporates a T7 promoter at the 5' end of FitA and sequences encoding six histidine residues at the 3' end of FitB. The C-terminal amino acid sequence of FitB was also

slightly altered by the addition of an XhoI restriction site, changing from ...NPWHD to ...NPWHLEHHHHHH. The overexpressed FitAB complex was purified using nickel affinity chromatography (Qiagen). Purified FitAB complex was concentrated to 5 mg/mL in 25 mM Tris, pH 7.5, 500 mM NaCl, 200 mM imidazole. Using 0.1 mg/mL trypsin (Sigma) the complex was digested for 30 minutes at 22 °C before crystallization trials. Trypsin inhibitor cross-linked to agarose beads (Sigma) was used to remove trypsin from the FitAB solution after digestion. Polyacrylamide gel electrophoresis and mass spectrometry analysis revealed that this treatment removed the ribbon-helix-helix motif of FitA and the resulting complex does not bind DNA (data not shown). Crystallization was carried out using hanging drop-vapour diffusion where FitAcB was mixed 1:1 (v:v) with a reservoir solution of 0.26 M sodium phosphate/citrate pH 4.7 and 2.0 M ammonium sulphate. Crystals appeared overnight and grew to final dimensions of 0.1 mm x 0.1 mm x 0.02 mm in approximately 3 days.

In order to generate selenomethionine (SeMet) substituted FitAcB complex, the expression vector described above was used as a template for standard PCR mutagenesis (Stratagene) of FitB to yield a construct encoding FitAB where residues Leu 43, Leu 63 and Leu 116 of FitB were substituted with methionines (FitAB3(LxM)). The FitAB3(LxM) protein was purified as described above. DNA binding assays confirmed that this FitAB complex has the same affinity as wild type FitAB for DNA (data not shown). For overexpression of SeMet substituted FitAB3(LxM), *Escherichia coli* harbouring the expression vector were grown in minimal medium lacking methionine with added SeMet as described (Doublet, 1997). Using nickel affinity column chromatography, SeMet-containing FitA and FitB3(LxM) co-purify as do the wild type

proteins. The SeMet containing heterodimer was concentrated and trypsinized as described for wild type FitAB. Crystallization of the SeMet-FitAcB3(LxM) complex employed 0.26 M sodium citrate pH 5.6 and 2.0 M ammonium sulphate. Crystals with dimensions 0.2 mm x 0.2 mm x 0.2 mm were obtained after four weeks.

For FitAB + DNA crystals, 5 mg/ml of purified native FitAB3(LxM) was mixed in a 4:1 molar ratio with IR36 DNA (Wilbur et al., 2005), where two of the thymine bases were replaced with 5-iodouracil (I) (top strand, 5'-

AGATTGCTATCATTTTTTTTTATTTTGATAGCATITG; bottom strand, 5' -

CAAATGCTATCAAAAIAAAAAAATGATAGCAATCT). The protein/DNA

complex was then mixed 1:1 (v:v) with a reservoir solution of 0.1 M sodium acetate, pH 4.0, 0.27 M sodium acetate, pH 7.0, 7.2 % PEG 20,000, 7.2 % PEG monomethyl ether 550. Crystals with dimensions 0.02 x 0.1 x 0.5 mm were obtained after one week.

Cryoprotection conditions for crystals of both native and SeMet substituted FitcAB were established by soaking crystals in 20% glycerol, 0.26 M sodium phosphate/citrate pH 4.7 and 2.2 M ammonium sulphate for approximately 30 seconds. Cryoprotection was achieved for FitAB + IR36 crystals by soaking the crystals in 20 % 2-methyl-2,4-pentanediol, 0.1 M sodium acetate, pH 4.0, 0.27 M sodium acetate, pH 7.0, 7.2 % PEG 20,000, 7.2 % PEG monomethyl ether 550 for 30 seconds. All crystals were flash frozen in a nitrogen stream at 100 K. X-ray intensity data were collected at the Advanced Light Source beamline 8.2.1 (Berkeley, CA) and processed using MOSFLM (Powell, 1999) as implemented in the CCP4 suite (Potterton et al., 2003). The native FitAcB crystals are monoclinic, space group C2 with $a = 70.0 \text{ \AA}$, $b = 50.7 \text{ \AA}$, $c = 48.3 \text{ \AA}$ and $\beta = 118.5^\circ$. The SeMet substituted FitAcB crystals take space group P2₁2₁2₁ with cell

dimensions $a = 49.1 \text{ \AA}$, $b = 68.5 \text{ \AA}$, $c = 104.5 \text{ \AA}$. FitAB + IR36 crystals are monoclinic, space group $P2_1$ with cell edges $a = 75.0 \text{ \AA}$, $b = 82.4 \text{ \AA}$, $c = 135.5 \text{ \AA}$ and $\beta = 94.2^\circ$.

Structure determination and refinement

The structure of SeMet-FitAcB3(LxM) was solved by multiple wavelength anomalous diffraction (MAD) methods using the SeMet-FitAcB3(LxM) data collected at three wavelengths (Table 1). Initial phases were calculated with SOLVE (Terwilliger and Berendzen, 1999) using data from 20.0 to 3.0 \AA resolution and improved by solvent flipping (with 45% solvent content) as implemented in the crystallography and NMR system (CNS) (Brunger et al., 1998). The handedness was determined by inspection of electron density maps where the initial phases were derived either from the Se atom sites found by SOLVE or sites with the inverted coordinates. Electron density maps for the entire resolution range (58.0-2.0 \AA) were then calculated using CNS. FitB and amino acid residues 46-65 of FitA were built into the map using O (Jones et al., 1991), and the location of the selenomethionine residues as reference points. Some water molecules were added to the model and simulated annealing (SA), positional and thermal parameter refinement using CNS were performed, followed by rebuilding of the model in O. When the R_{work} and R_{free} were 28.2% and 29.7%, respectively, the coordinates were used to solve the structure of the wild type, native FitcAB molecule by molecular replacement using MOLREP (Vagin and Teplyakov, 1997). Multiple rounds of refinement in CNS followed by model rebuilding in O were done using the native data to the 1.8 \AA limiting resolution until the R_{free} converged to 22.4%. The final model contains residues 1-139 of FitB, 46-64 of FitA, 92 water molecules, 1 acetate ion, 2 sulphate ions and 3 magnesium ions. The

final model was verified by inspection of $2F_o-F_c$ simulated annealing-composite omit maps. The quality of the stereochemistry, assessed by PROCHECK (Laskowski et al., 1993) is excellent, with 94.2 % of all ϕ/ψ angles in the most favored region of a Ramachandran plot and none in generously allowed or disallowed regions of the plot.

The high resolution FitcAB structure was used as a model to solve the structure of the FitAB + IR36 complex by molecular replacement using MOLREP (Vagin and Teplyakov, 1997). The remainder of the FitA sequence and the DNA molecules were built into the map using O (Jones et al., 1991). After extensive simulated annealing (SA), positional and thermal parameter refinement using CNS followed by model rebuilding in O, R_{work} and R_{free} converged to 21.2% and 26.9%, respectively. The model was verified by inspection of the $2F_o-F_c$ simulated annealing-composite omit maps. The stereochemistry is good for this 3.0 Å model, with 88.5 % of all ϕ/ψ angles in the most favored region of a Ramachandran plot and only 0.3 % (2 residues out of 740 non glycine, non proline residues) in generously allowed regions as determined by PROCHECK (Laskowski et al., 1993). The final model contains four molecules of FitA (residues 2-69, 2-65, 2-68, 2-64), four molecules of FitB (residues 1-143, 1-140, 1-143, 1-140), the 36 base pair double-stranded IR36 DNA fragment and 54 water molecules.

Figures were made using Swiss-PDB Viewer (Guex and Peitsch, 1997) and POV-Ray (www.povray.org).

Acknowledgements

The authors thank Drs. Hans Peter Bächinger, Kerry Maddox and Cory Bystrom for N-terminal sequencing and mass spectroscopy analysis of proteins. The Advanced Light

Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory. This work was supported by NIH grant AI47260 to MS. RGB is the Richard T. Jones Professor of Structural Biology at OHSU. KM is a postdoctoral fellow of the American Heart Association - Pacific Mountain Affiliate.

Data deposition

The coordinates and structure factors for the FitcAB fragment and the FitAB + IR36 complex have been deposited at the Protein Data Bank under the accession codes 1YH4 and 2BSQ, respectively.

References

- Alexeev, D.G., Lipanov, A.A. and Skuratovskii, I. (1987) Poly(dA)·poly(dT) is a B-type double helix with a distinctively narrow minor groove. *Nature*, **325**, 821-823.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol*, **215**, 403-410.
- Apicella, M.A., Ketterer, M., Lee, F.K., Zhou, D., Rice, P.A. and Blake, M.S. (1996) The pathogenesis of gonococcal urethritis in men: confocal and immunoelectron microscopic analysis of urethral exudates from men infected with *Neisseria gonorrhoeae*. *J Infect Dis*, **173**, 636-646.

- Arcus, V.L., Backbro, K., Roos, A., Daniel, E.L. and Baker, E.N. (2004) Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J Biol Chem*, **279**, 16471-16478.
- Brendler, T., Reaves, L. and Austin, S. (2004) Interplay between plasmid partition and postsegregational killing systems. *J Bacteriol*, **186**, 2504-2507.
- Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*, **54 (Pt 5)**, 905-921.
- Burgering, M.J.M., Boelens, R., Gilbert, D.E., Breg, J.N., Knight, K.L., Sauer, R.T. and Kaptein, R. (1994) Solution structure of dimeric Mnt repressor (1-76). *Biochemistry*, **33**, 15036-15045.
- Clissold, P.M. and Ponting, C.P. (2000) PIN domains in nonsense-mediated mRNA decay and RNAi. *Curr Biol*, **10**, R888-890.
- DiGabrieli, A.D., Sanderson, M.R. and Steitz, T.A. (1989) Crystal lattice packing is important in determining the bend of a DNA dodecamer containing an adenine tract. *Proc Natl Acad Sci U S A*, **86**, 1816-1820.
- Doublet, S. (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol*, **276**, 523-530.
- Edwards, J.L. and Apicella, M.A. (2004) The molecular mechanisms used by *Neisseria gonorrhoeae* to initiate infection differ between men and women. *Clin Microbiol Rev*, **17**, 965-981, table of contents.

- Freiberg, C., Fellay, R., Bairoch, A., Broughton, W.J., Rosenthal, A. and Perret, X. (1997) Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature*, **387**, 394-401.
- Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714-2723.
- Han, G.W., Kopka, M.L., Cascio, D., Grzeskowiak, K. and Dickerson, R.E. (1997) Structure of a DNA analog of the primer for HIV-1 RT second strand synthesis. *J Mol Biol*, **269**, 811-826.
- Hanekamp, T., Kobayashi, D., Hayes, S. and Stayton, M.M. (1997) Avirulence gene D of *Pseudomonas syringae* pv. tomato may have undergone horizontal gene transfer. *FEBS Lett*, **415**, 40-44.
- Holm, L. and Sander, C. (1995) Dali: a network tool for protein structure comparison. *Trends Biochem Sci*, **20**, 478-480.
- Holmes, K.K., Counts, G.W. and Beaty, H.N. (1971) Disseminated gonococcal infection. *Ann Intern Med*, **74**, 979-993.
- Hopper, S., Vasquez, B., Merz, A., Clary, S., Wilbur, J.S. and So, M. (2000a) Effects of the immunoglobulin A1 protease on *Neisseria gonorrhoeae* trafficking across polarized T84 epithelial monolayers. *Infect Immun*, **68**, 906-911.
- Hopper, S., Wilbur, J.S., Vasquez, B.L., Larson, J., Clary, S., Mehr, I.J., Seifert, H.S. and So, M. (2000b) Isolation of *Neisseria gonorrhoeae* mutants that show enhanced trafficking across polarized T84 epithelial monolayers. *Infect Immun*, **68**, 896-905.

- Hwang, K.Y., Baek, K., Kim, H.Y. and Cho, Y. (1998) The crystal structure of flap endonuclease-1 from *Methanococcus jannaschii*. *Nat Struct Biol*, **5**, 707-713.
- Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A*, **47 (Pt 2)**, 110-119.
- Kaiser, D. and Wall, D. (1999) Type IV pili and cell motility. *Mol Microbiol*, **32**, 1-10.
- Krissinel, E. and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr*, **60**, 2256-2268.
- Lan, R., Stevenson, G. and Reeves, P.R. (2003) Comparison of two major forms of the *Shigella* virulence plasmid pINV: positive selection is a major force driving the divergence. *Infect Immun*, **71**, 6298-6306.
- Laskowski, R.A., Moss, D.S. and Thornton, J.M. (1993) Main-chain bond lengths and bond angles in protein structures. *J Mol Biol*, **231**, 1049-1067.
- Leroy, J.L., Charretier, E., Kochoyan, M. and Gueron, M. (1988) Evidence from base-pair kinetics for two types of adenine tract structures in solution: their relation to DNA curvature. *Biochemistry*, **27**, 8894-8898.
- Makarova, K.S., Aravind, L., Galperin, M.Y., Grishin, N.V., Tatusov, R.L., Wolf, Y.I. and Koonin, E.V. (1999) Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. *Genome Res*, **9**, 608-628.

- Makarova, K.S., Aravind, L., Grishin, N.V., Rogozin, I.B. and Koonin, E.V. (2002) A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res*, **30**, 482-496.
- Merz, A.J. and So, M. (2000) Interactions of pathogenic neisseriae with epithelial cell membranes. *Annu Rev Cell Dev Biol*, **16**, 423-457.
- Nolte, R.T., Wisely, G.B., Westin, S., Cobb, J.E., Lambert, M.H., Kurokawa, R., Rosenfeld, M.G., Willson, T.M., Glass, C.K. and Milburn, M.V. (1998) Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature*, **395**, 137-143.
- Potterton, E., Briggs, P., Turkenburg, M. and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr*, **59**, 1131-1137.
- Powell, H.R. (1999) The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallogr D Biol Crystallogr*, **55 (Pt 10)**, 1690-1695.
- Pullinger, G.D. and Lax, A.J. (1992) A Salmonella dublin virulence plasmid locus that affects bacterial growth under nutrient-limited conditions. *Mol Microbiol*, **6**, 1631-1643.
- Raumann, B.E., Rould, M.A., Pabo, C.O. and Sauer, R.T. (1994) DNA recognition by beta-sheets in the Arc repressor-operator crystal structure. *Nature*, **367**, 754-757.
- Terwilliger, T.C. and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr*, **55 (Pt 4)**, 849-861.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence

weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, **22**, 4673-4680.

Turner, C.F., Rogers, S.M., Miller, H.G., Miller, W.C., Gribble, J.N., Chromy, J.R., Leone, P.A., Cooley, P.C., Quinn, T.C. and Zenilman, J.M. (2002) Untreated gonococcal and chlamydial infection in a probability sample of adults. *Jama*, **287**, 726-733.

Vagin, A. and Teplyakov, A. (1997) MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Cryst.*, **30**, 1022-1025.

Watkins, R.E., Maglich, J.M., Moore, L.B., Wisely, G.B., Noble, S.M., Davis-Searles, P.R., Lambert, M.H., Klierer, S.A. and Redinbo, M.R. (2003) 2.1 Å crystal structure of human PXR in complex with the St. John's wort compound hyperforin. *Biochemistry*, **42**, 1430-1438.

Wilbur, J.S., Chivers, P.T., Mattison, K., Potter, L., Brennan, R.G. and So, M. (2005) *Neisseria gonorrhoeae* FitA interacts with FitB to bind DNA through its ribbon-helix-helix motif. *Biochemistry*.

Table and Figure Legends

Table I

^a $R_{\text{sym}} = \frac{\sum \sum |I_{\text{hkl}} - I_{\text{hkl}}(j)|}{\sum I_{\text{hkl}}}$, where $I_{\text{hkl}}(j)$ is observed intensity and I_{hkl} is the final average value of intensity.

^bvalues in parentheses are for the highest resolution shell

^cFigure of Merit = $\langle |\sum P(\alpha)e^{i\alpha} / \sum P(\alpha)| \rangle$, where α is the phase and $P(\alpha)$ is the phase probability distribution.

^d $R_{\text{work}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$ and $R_{\text{free}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$; where all reflections belong to a test set of 10% randomly selected data.

Figure 1 The structure of the FitAB heterodimer. In all panels, FitA is shown in magenta and FitB is in cyan. (A) Ribbon diagram of the FitAB heterodimer. The α helices and β sheets for both proteins are labeled as are the N- and C-termini. (B) Sequence alignment of FitA and FitB proteins with homologues. The secondary structural elements of the proteins are shown above the alignment. Y4jJK from *Rhizobium sp.* and StbCB from *Pseudomonas syringae* are operons highly homologous to FitAB by sequence analysis. Arc and Mnt from Enterobacteria phage P22 are ribbon-helix-helix proteins with structural homology to FitA. PiIT from *Azotobacter vinelandii* and PAE2754 from *Pyrobaculum aerophilum* are PIN domain containing proteins with structural homology to FitB. Alignments were generated with ClustalW (Thompson et al., 1994). Residues involved in FitA-FitB heterodimerization are highlighted in yellow, those forming the FitA-FitA dimer interface are highlighted in green, and those making up the FitB-FitB

dimer interface are highlighted in orange. Conserved residues involved in specific protein-DNA contacts are highlighted in blue and the conserved residues forming the putative FitB active site are highlighted in red. (C) Stereo view of the FitA-FitB heterodimer interface. Residues involved in ionic interactions contributing to the stability of the interface are colored according to grey = carbon, red = oxygen and blue = nitrogen. This extensive interface buries up to 1900 Å² of accessible surface area.

Figure 2 Ribbon diagram of the FitAB-IR36 complex. (A) The FitA and FitB proteins are colored magenta and cyan, respectively. DNA is colored according to grey = carbon, red = oxygen, blue = nitrogen, yellow = phosphorus. The four FitAB heterodimers are numbered from I to IV. (B) View of Figure 2A rotated to demonstrate that the two FitA β sheets bind on the same face of the DNA helix. (C) The 36 bp IR36 site used for crystallization. The 8 bp inverted repeat half sites are highlighted in yellow.

Figure 3 Stereoviews of dimerization interfaces. FitA is in magenta and FitB is in cyan. Some residues are colored according to grey = carbon, red = oxygen and blue = nitrogen. (A) FitA dimerization forms the DNA binding β sheet, with β1 from each of the FitA subunits (I and IV or II and III). (B) The two FitA monomers form a tightly packed globular domain, with interactions between α1 of one subunit and α2 of the dimer partner (I with IV and II with III). Two identical interfaces are formed in this manner per FitA dimer, which together with the β sheet stabilize the domain and bury 2850 Å² of accessible surface area. (C) Two FitB monomers have an extensive dimerization interface (I and II or III and IV), where α3 from one subunit contacts α5 from the other subunit.

Half of this interface is shown here, as there is 2-fold symmetry around F78 in the complete interface. FitB homodimerization buries 1870 Å² of accessible surface area.

Figure 4 FitA-DNA contacts. (A) Schematic diagram of the FitA-DNA contacts. The deoxyriboses of each nucleotide are numbered and shown as pentagons. Side chain-DNA hydrogen bonds are indicated by blue arrows, backbone amide-DNA hydrogen bonds are green arrows and van der Waals contacts are shown as yellow arrows. In brackets beside each FitA residue number is the roman numeral describing the subunit to which it belongs, as defined in Figure 2. (B) Stereo view of one FitA-DNA interface with the 3.0 Å resolution composite omit map (green mesh contoured at 1.0 σ). The FitA protein is shown as magenta balls and sticks and the DNA is shown as balls and sticks where grey = carbon, red = oxygen, blue = nitrogen, yellow = phosphorus. Wat1 is a water molecule forming hydrogen bonds with both Asn8 and Thy32'.

Figure 5 FitB is homologous to a protein with nuclease activity. (A) Superimposition of the structure of FitB (cyan) and PAE2754 from *Pyrobaculum aerophilum* (blue)(PDB accession #1V8P). The alignment is based on optimized superimposition of 93 C α atoms (FitB residues 1-11, 29-51, 54-71, 74-80, 99-126, 131-136), resulting in a final r.m.s.d. of 1.9 Å between the two structures. (B) The strictly conserved acidic residues that define the PIN domain cluster in a surface pocket at the C-terminal end of the central, 5 stranded β sheet. Residues and ribbons from FitB are in cyan and from PAE2754 are in blue. (C) In FitB, this putative active site for nuclease activity is blocked by the presence of Arg68

from FitA. The FitA ribbon is in magenta and the FitB ribbons are in cyan. All residues are colored according to where grey = carbon, red = oxygen, blue = nitrogen.

Figure 1

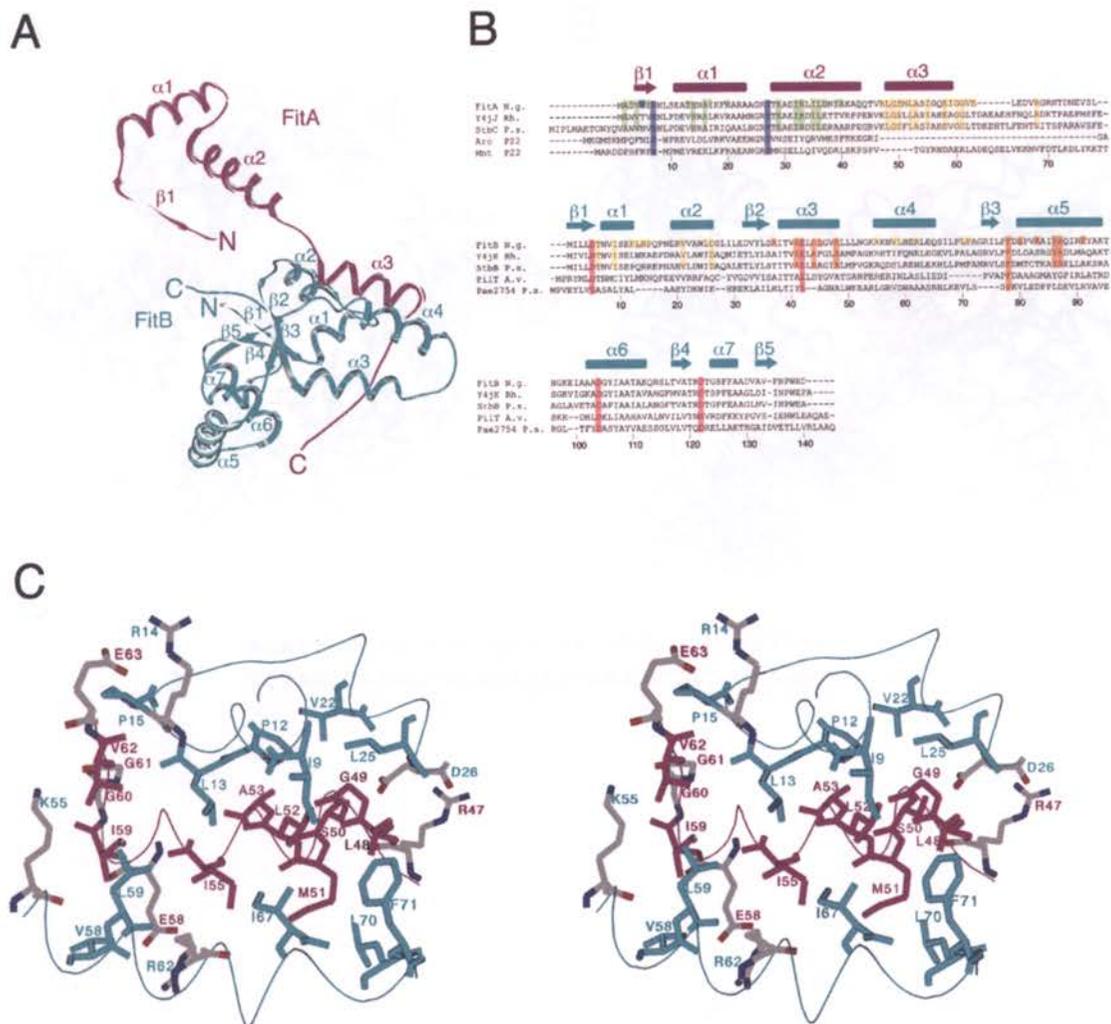


Figure 2

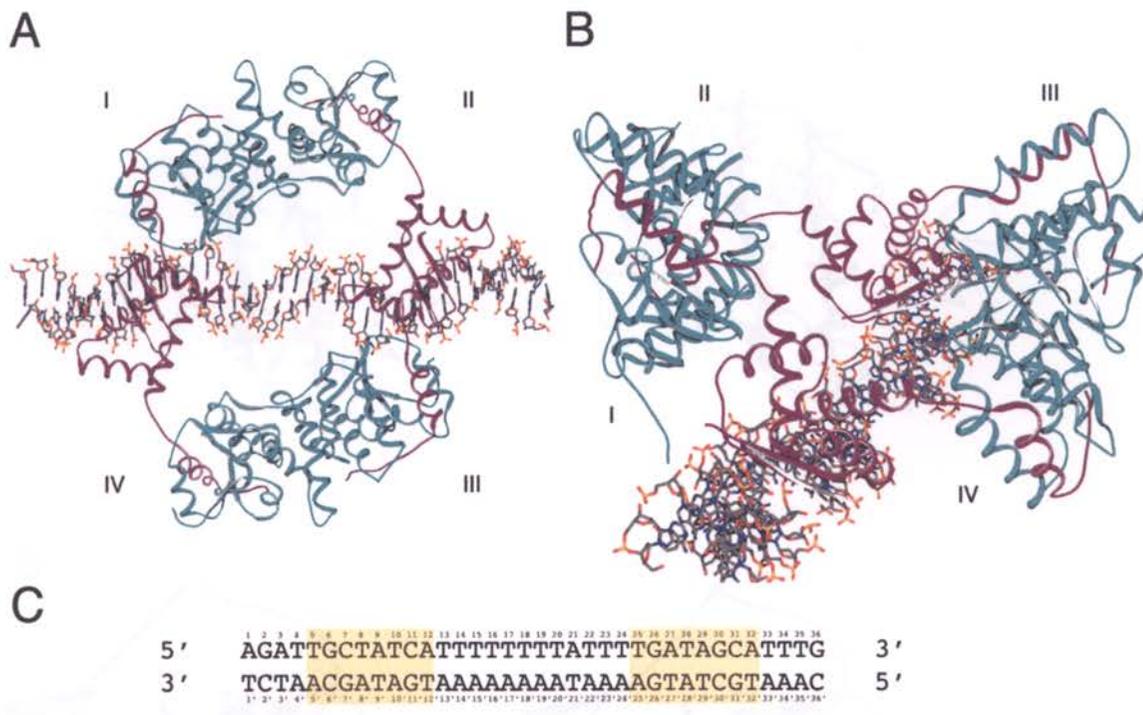
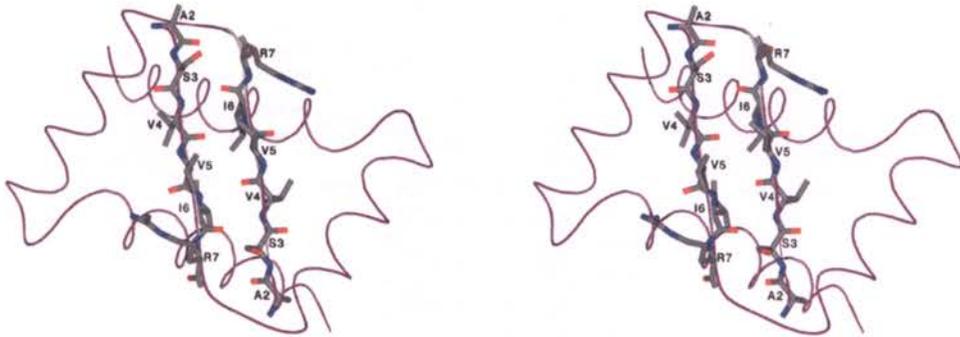
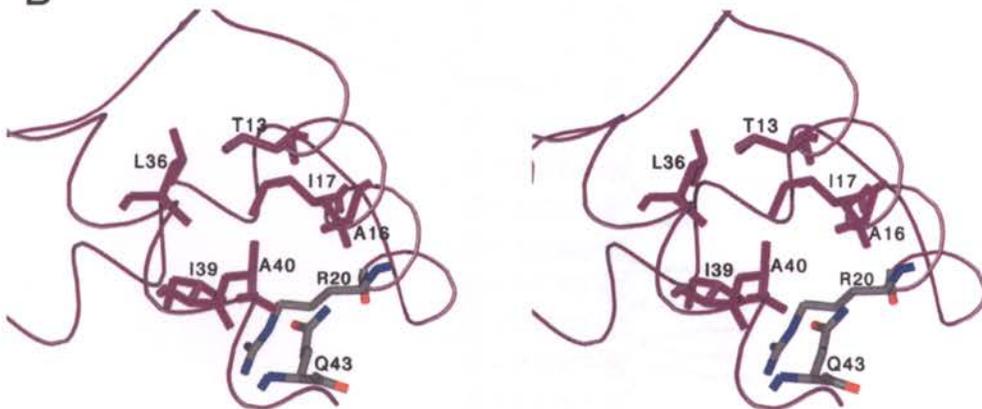


Figure 3

A



B



C

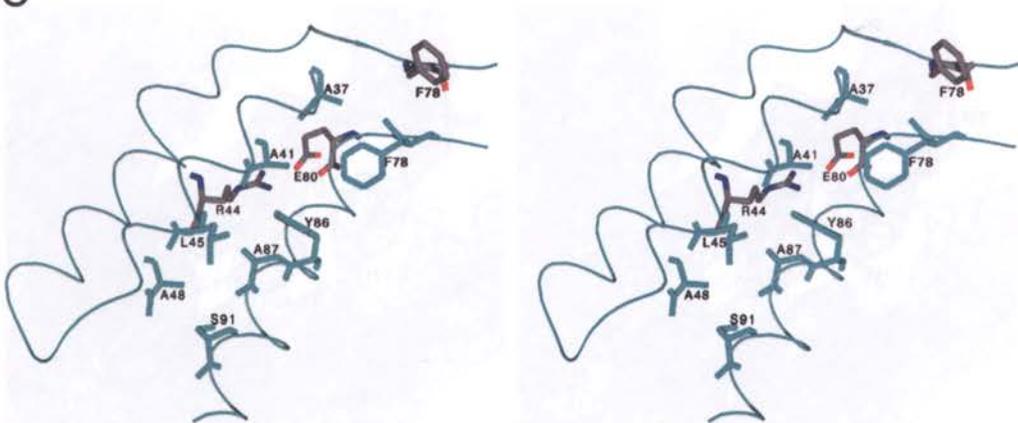


Figure 4

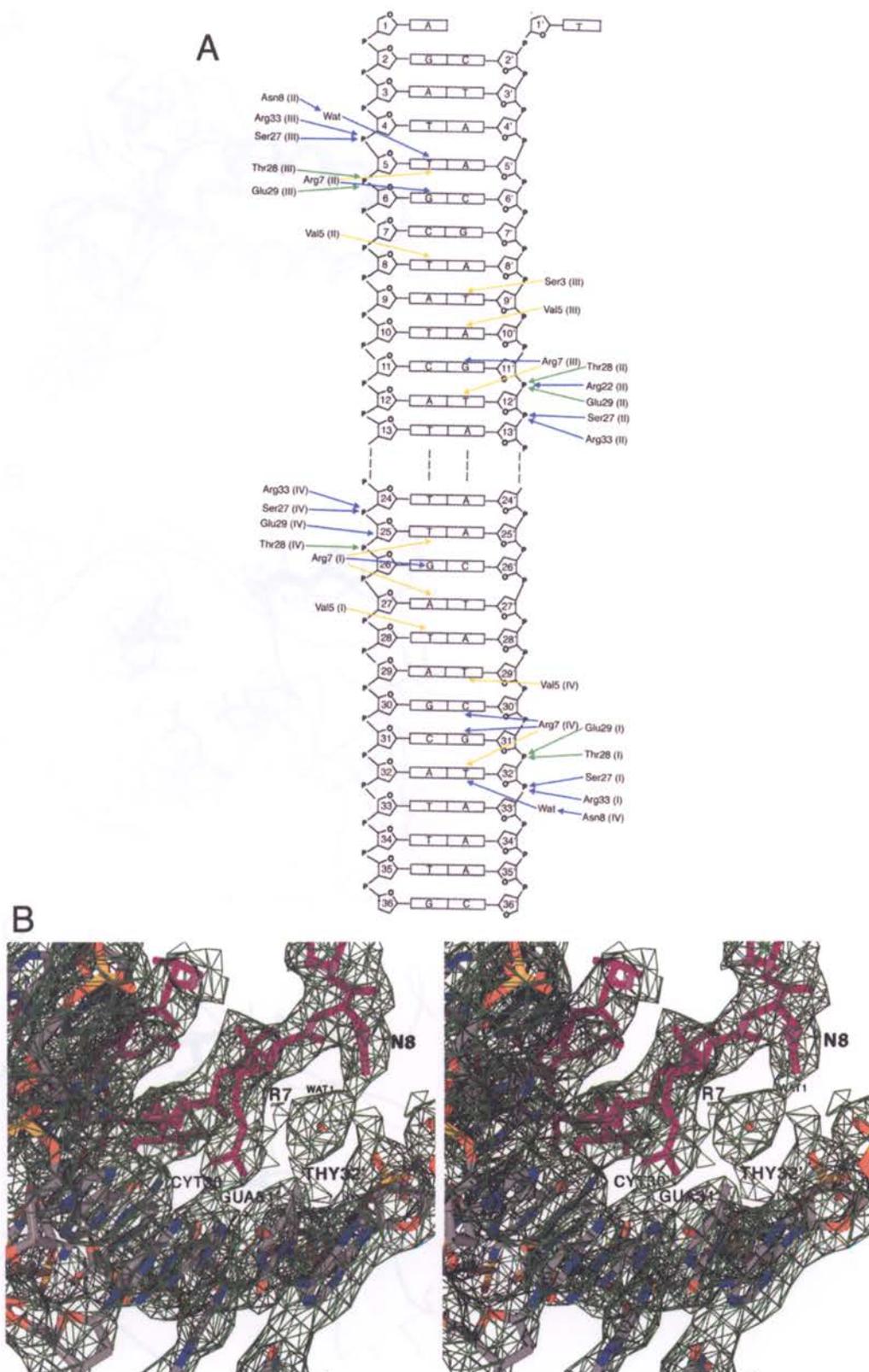
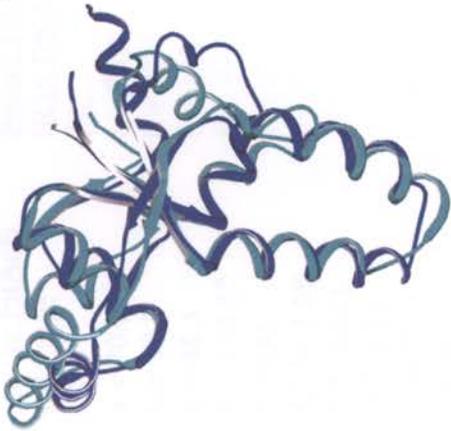
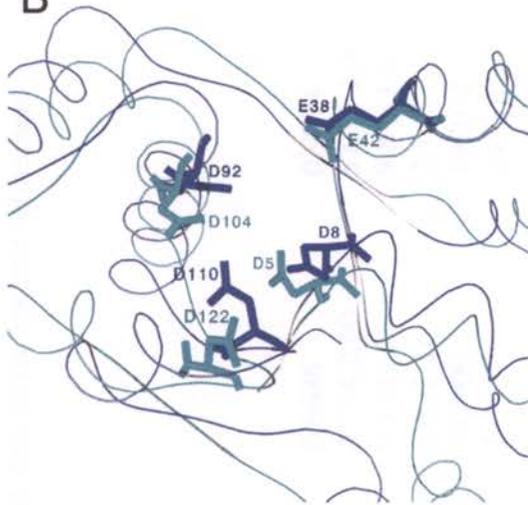


Figure 5

A



B



C

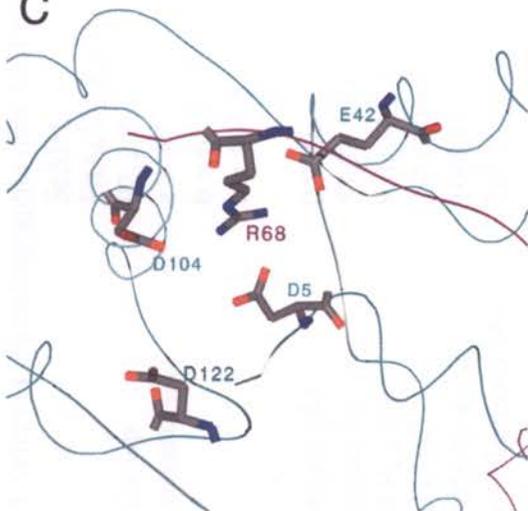


Table I Selected crystallographic data for FitAB data collection and refinement

	<u>SeMet- FitAcB3(LxM)</u>			<u>Native FitAcB wild type</u>	<u>Native FitAB + IR36 DNA</u>
Space Group	P2 ₁ 2 ₁ 2 ₁			C2	P2 ₁
Cell constants (Å)	a = 49.1 b = 68.5 c = 104.5			a = 70.0 b = 50.7 c = 48.3 β = 118.5°	a = 75.0 b = 82.4 c = 135.5 β = 94.2°
Wavelength (Å)	0.9796	0.9686	0.9794	1.0000	1.0332
Resolution (Å)	57.74-2.00	57.74-2.00	57.74-2.00	42.26-1.80	81.65-3.00
Overall R _{sym} (%) ^a	9.0 (32.5) ^b	9.5 (33.2)	9.8 (33.5)	5.5 (22.2)	17.8 (44.0)
Overall I/σ(I)	3.7 (2.3)	3.4 (2.2)	3.1 (2.2)	9.7 (3.1)	3.1 (1.5)
Total reflections (#)	196843	197909	197072	31002	33243
Unique reflections (#)	24539	24542	24537	13340	16708
Completeness (%)				96.5 (95.3)	99.9 (99.9)
Phasing Resolution (Å)	20.00-3.00				
Sites (#)	7				
Overall Figure of Merit ^c	0.720				
Refinement					
R _{work} /R _{free} (%) ^d				19.1/22.4	21.2/26.9
rmsd					
Bond Angles (°)				1.28	1.27
Bond Lengths (Å)				0.007	.008
B values (Å ²)				1.90	1.60
Average B-values (Å ²)				25	41
Ramachandran Analysis					
Most Favoured (%/#)				94.2/129	88.5/655
Additional Allowed (%/#)				5.8/8	11.2/83
Generously Allowed (%/#)				0/0	0.3/2
Disallowed (%/#)				0/0	0/0

Chapter 4: Additional Results

Examining the relationship between the *fit* locus and *pilE* recombination

In this section I will discuss unpublished work I have done examining the relationship between *pilE* AV and the *fit* locus. During transcytosis GC undergoes a specific recombination event at the *pilE* locus (10, 11, 16). *PilE* encodes the major structural subunit of the type IV pilus (4, 6). This event is termed high frequency antigenic variation (AV), and involves the directed recombination between hyper-variable regions of *pilE* and homologous regions of the chromosome collectively called *pilS* (or silent pilin loci) (1, 10, 17). These recombination events are well characterized and function primarily in a RecA dependent process (12).

GC undergoes *pilE* AV during urethral infection of human males as well as during tissue culture models of polarized Hec1B cells (human endocervical carcinoma cell line, (ATCC HTB113; (5, 11)). In the human infection studies, GC were inoculated into the urethra and urine from the subjects was collected periodically after challenge and cultured. GC were not recovered from the urine samples until the subject had developed urethritis, 4 or more days after initial exposure. The *pilE* sequences from the collected bacteria were different from the sequence in the input strain, indicating that *pilE* AV had occurred during infection. Moreover, in infection studies of polarized Hec1B monolayers 100% of transcytosed bacteria had undergone AV. It was also reported that a *recA* mutant failed to cross the monolayer (11). These results strongly suggest a connection between transcytosis and *pilE* AV.

When GC are grown in GCB media (Gibco), *pilE* undergoes AV at a rate between 10^{-4} and 10^{-3} event/cell/generation (18). This is a gene conversion event between *pilE* and

one of the 19 non-transcribed *pilS* pseudo-genes, because the sequence of *pilE* is changed without any changes to the *pilS* sequences (5, 10). The *pilS* alleles are truncated at their 5' end, do not encode a functional pilin, and do not directly contribute to the formation of the type IV pilus. The consequence of these recombination events is a change in the Pile amino acid sequence. The variable regions are predicted to be surface exposed along the pilus (2, 3). In this way a bacterium is thought to change the Pile antigens exposed to the host immune system after each passage through a monolayer of epithelial cells.

Previously, we demonstrated that the *fit* locus of GC increased the speed of bacterial traversal across polarized monolayers of T84 cells (ATCC CCL248) as well as decreasing bacterial growth inside host cells (8). We wanted to determine if there was a connection between *fit* and *pilE* AV during transcytosis. We infected T84 polarized monolayers with FA1090 (wildtype) and A9 (FA1090-*fitAB* mutant) bacteria (piliated). Bacteria that had traversed the monolayers in the first 36 hours were collected and the *pilE* sequences from these strains were determined. In each experiment, four wells of polarized T84 cells were infected with FA1090 and four wells with A9. Monolayers were prepared and polarity was determined as previously described (7, 8, 13, 15). The infected monolayers were transferred to fresh DMEM-F12 media every 8 hours for 36 hours. The basal media from each time point was plated on 15 cm GCB agar plates and incubated at 37° for 18-24 hours. Bacterial colonies were picked with sterile 3MM filter paper (Watson) and stored at -80°C in 200µL of GCB+20% glycerol. A total of 54 wildtype FA1090 colonies and 53 A9 post-transcytosis colonies were isolated and

analyzed. These colonies were collected from three different experiments and from multiple wells in each experiment.

The *pilE* sequence from the above strains was amplified by PCR (Primer: Pile-start and Sp3A) (9, 20) and the products were sequenced. We found that *pilE* locus of all 54 FA1090 (wt) colonies had sequence changes (Table 1). The pilin sequences from five variant *pilE* sequences are shown in Figure 1. The sequence of the FA1090 Pile at the beginning of the infection is also displayed. The residues in red are those that have changed from the original input residues. These changes occurred in the region of *pilE* that have previously been shown to vary during *pilE* AV. In contrast, 52 of the 53 A9 *pilE* sequences post transcytosis were the same as the input strain's sequence. The one different A9 *pilE* sequence had a single point mutation that did not change the amino acid sequence. This sequence variation is therefore not likely due to *pilE* AV. Thus, *fit* locus appears to stimulate *pilE* AV during transcytosis.

We demonstrated whether A9 was capable of generalized recombination by transforming the strain with the plasmid pGCC5 (14). This plasmid recombines with a homologous sequence in the GC chromosome and inserts a chloramphenicol resistance marker between the *aspC* and *lctP* genes. Insertion into this locus has not had measurable effect on replication (Mehr and Seifert, unpublished result). This plasmid had similar transformation efficiency into both FA1090 and A9 strains (data not shown). These data indicate that the A9 strain is able to undergo *recA* dependent recombination.

These data also show that the A9 mutation does not affect pilus assembly or function, as GC transformation requires a functional, retractile pilus.

We next determined if A9 was capable of *pilE* AV outside of host cells. In order to test this, the A9 strain was passaged on GCB agar plates for 10 days; at the 10th passage *pilE* from five of colonies was sequenced. We found that the A9 strain had undergone *pilE* AV in one of these colonies (data not shown). Thus the mutation in A9 did not affect *recA* dependent recombination, transformation, or *pilE* AV, except during transcytosis.

Ilver et al 1998 found that a *recA* mutant would not traffic across polarized Hec1B cells, suggesting the dependence of transcytosis on the general recombination pathway. We attempted to repeat this in our polarized T84 cell infection model (8, 15). Polarized T84 cell monolayers were infected with either MS11A (wildtype) or MS11VD302 (*recA*). Basal media was collected at 12, 24, 36 and 48 hours, as described in Hopper et al. Not only was the mutant able to cross the monolayer, it crossed more quickly than wildtype (Figure 2). We cannot explain the differences between Ilver et al.'s results and ours. Perhaps they are due to the different host cell types infected.

We also attempted to examine the relationship between the *fit* locus and the phase variation rate of the Opa (opacity proteins). Opa expression has been shown to stimulate transcytosis of GC across monolayers of T84 cells (19). Opa phase variation is controlled by a process known as slip-strand misrepair, a replication dependent process

that does not involve *recA*. Unfortunately, due to some technical difficulties we have been unable to verify a connection between Opa expression and the *fit* locus.

Our data suggest that the *fit* locus stimulates *pilE* AV during transcytosis. The mechanism of this activity is unclear though the fact the *recA* mutant had a fast trafficking phenotype, similar to the *fit* mutants, suggests that FitA/B may regulate transcytosis and intracellular replication through stimulation of RecA activity.

1. **Aho, E. L., and J. G. Cannon.** 1988. Characterization of a silent pilin gene locus from *Neisseria meningitidis* strain FAM18. *Microb Pathog* **5**:391-8.
2. **Forest, K. T., S. L. Bernstein, E. D. Getzoff, M. So, G. Tribbick, H. M. Geysen, C. D. Deal, and J. A. Tainer.** 1996. Assembly and antigenicity of the *Neisseria gonorrhoeae* pilus mapped with antibodies. *Infect Immun* **64**:644-52.
3. **Forest, K. T., S. A. Dunham, M. Koomey, and J. A. Tainer.** 1999. Crystallographic structure reveals phosphorylated pilin from *Neisseria*: phosphoserine sites modify type IV pilus surface chemistry and fibre morphology. *Mol Microbiol* **31**:743-52.
4. **Forest, K. T., and J. A. Tainer.** 1997. Type-4 pilus-structure: outside to inside and top to bottom--a minireview. *Gene* **192**:165-9.
5. **Hamrick, T. S., J. A. Dempsey, M. S. Cohen, and J. G. Cannon.** 2001. Antigenic variation of gonococcal pilin expression in vivo: analysis of the strain FA1090 pilin repertoire and identification of the *pilS* gene copies recombining with *pilE* during experimental human infection. *Microbiology* **147**:839-49.

6. **Hermodson, M., K. Chen, and T. Buchanan.** 1978. *Neisseria pili* Proteins: Amino-Terminal Amino Acid Sequences and Identification of an Unusual Amino Acid. *Biochem.* **17**:442-445.
7. **Hopper, S., B. Vasquez, A. Merz, S. Clary, J. S. Wilbur, and M. So.** 2000. Effects of the immunoglobulin A1 protease on *Neisseria gonorrhoeae* trafficking across polarized T84 epithelial monolayers. *Infect Immun* **68**:906-11.
8. **Hopper, S., J. S. Wilbur, B. L. Vasquez, J. Larson, S. Clary, I. J. Mehr, H. S. Seifert, and M. So.** 2000. Isolation of *Neisseria gonorrhoeae* mutants that show enhanced trafficking across polarized T84 epithelial monolayers. *Infect Immun* **68**:896-905.
9. **Howell-Adams, B., and H. S. Seifert.** 1999. Insertion mutations in *pilE* differentially alter gonococcal pilin antigenic variation. *J Bacteriol* **181**:6133-41.
10. **Howell-Adams, B., and H. S. Seifert.** 2000. Molecular models accounting for the gene conversion reactions mediating gonococcal pilin antigenic variation. *Mol Microbiol* **37**:1146-58.
11. **Ilver, D., H. Kallstrom, S. Normark, and A. B. Jonsson.** 1998. Transcellular passage of *Neisseria gonorrhoeae* involves pilus phase variation. *Infect Immun* **66**:469-73.
12. **Koomey, J. M., and S. Falkow.** 1987. Cloning of the *recA* gene of *Neisseria gonorrhoeae* and construction of gonococcal *recA* mutants. *J Bacteriol* **169**:790-5.
13. **McCormick, B., P. Hofman, J. Kim, D. Carnes, S. Miller, and J. Madara.** 1995. Surface attachment of *Salmonella typhimurium* to intestinal epithelia

imprints the subepithelial matrix with gradients chemotactic for neutrophils. *J. Cell Biol.*:1599-1608.

14. **Mehr, I. J., C. D. Long, C. D. Serkin, and H. S. Seifert.** 2000. A homologue of the recombination-dependent growth gene, *rdgC*, is involved in gonococcal pilin antigenic variation. *Genetics* **154**:523-32.
15. **Merz, A. J., D. B. Rifken, C. G. Arvidson, and M. So.** 1996. Traversal of a polarized epithelium by pathogenic *Neisseriae*: facilitation by type IV pili and maintenance of epithelial barrier function. *Mol Med* **2**:745-54.
16. **Plant, L., and A. B. Jonsson.** 2003. Contacting the host: insights and implications of pathogenic *Neisseria* cell interactions. *Scand J Infect Dis* **35**:608-13.
17. **Seifert, H. S., R. S. Ajioka, C. Marchal, P. F. Sparling, and M. So.** 1988. DNA transformation leads to pilin antigenic variation in *Neisseria gonorrhoeae*. *Nature* **336**:392-395.
18. **Serkin, C. D., and H. S. Seifert.** 1998. Frequency of pilin antigenic variation in *Neisseria gonorrhoeae*. *J Bacteriol* **180**:1955-8.
19. **Wang, J., S. D. Gray-Owen, A. Knorre, T. F. Meyer, and C. Dehio.** 1998. Opa binding to cellular CD66 receptors mediates the transcellular traversal of *Neisseria gonorrhoeae* across polarized T84 epithelial cell monolayers. *Mol Microbiol* **30**:657-71.
20. **Wright, C. J., A. E. Jerse, M. S. Cohen, J. G. Cannon, and H. S. Seifert.** 1994. Nonrepresentative PCR amplification of variable gene sequences in clinical

specimens containing dilute, complex mixtures of microorganisms. *J Clin Microbiol* **32**:464-8.

Figure legends

Figure 1 Alignment of predicted PilE amino acid sequences post transcytosis. The sequence of five colonies are displayed, aligned to the input sequence (FA1090). Amino acids in red are sequences different from input sequence. All A9 isolates had the input sequence when recovered from the basal media.

Figure 2 Transcytosis experiment with *recA* mutant strain. Red striped bars are *recA* strain and solid bars are wildtype strain. Forty T84 monolayers were infected and transferred to fresh media every 12 hours. Basal media was then further incubated at 37° C for 24 hours and monitored for color change due to acidification, indicating the presence of GC in the media. Displayed here is the percent of wells that had bacteria in the basal media at each time point.

Figure 2

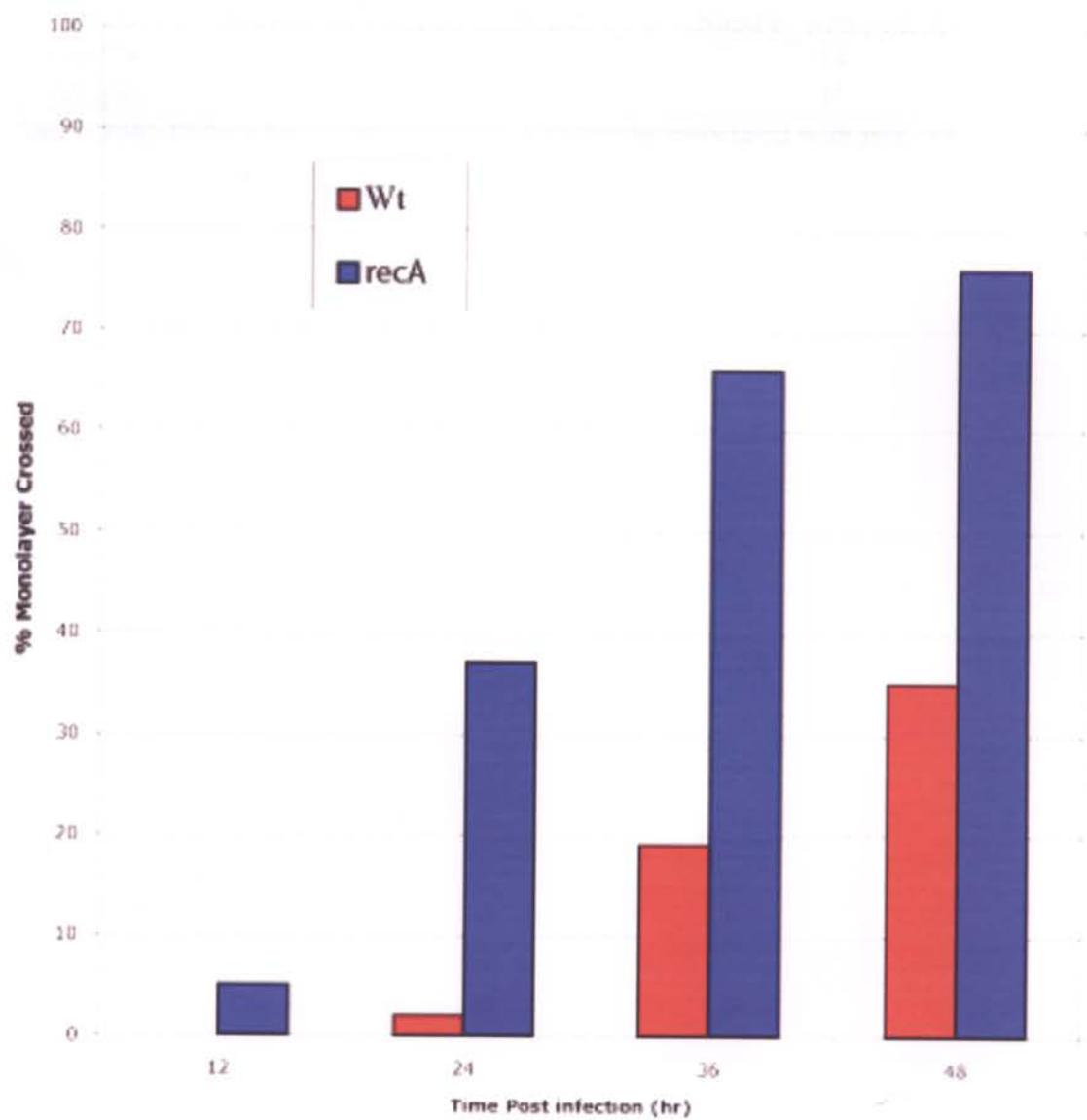


Table 1. Results of *pilE* AV post-transcytosis analysis

Strain	Number of colonies analyzed	Number with <i>pilE</i> AV
FA1090	54	54
A9 (<i>fit</i>)	53	1 ^a

^aone colony had a single point mutation, not usually associated with *pilE* AV

Chapter 5: Conclusions

Conclusions

In this thesis I examined the biochemical properties of the fast trafficking proteins FitA and FitB in order to gain a better understanding of the intracellular life cycle of *Neisseria gonorrhoeae* and the establishment of a carrier state infection. Manuscript I shows that FitA and FitB form a heterooctamer complex that binds to DNA in the promoter region of *fitA*. In manuscript II the crystal structure of the FitA/B heterooctamer bound to DNA was solved. This structure suggests that FitB is a PIN domain containing protein, yet the PIN domains putative active site is blocked by Arg68 of FitA when the proteins are in complex and bound to DNA. Chapter four shows that the *fit* mutant does not undergo *pilE* AV during transcytosis. In this chapter I will discuss what these findings suggest and how they relate to the original goal of understanding the intracellular life of GC.

1.1 Opa expression drives transcytosis

The *fit* locus was originally identified in a mutant screen for GC determinants that affect bacterial transcellular trafficking across polarized human epithelial cell monolayers (8). The *fit* mutants transcytosed 5-6 times faster than wildtype. These mutants also exhibited a second phenotype: fast intracellular replication compared to wildtype, but not in liquid media. Is the process of transcytosis functionally related to intracellular replication? These two processes could be related if transcytosis is dependent on a gene whose expression is tied to replication.

The Opa (opacity) proteins are a class of outer membrane proteins that have been implicated in transcytosis. Interestingly, the expression of these proteins are phase

variable by process called slip-strand misrepair (16, 17, 20). Slip-strand misrepair is a process by which mistakes in the pairing of repeated sequences within the open reading frame of the *opa* change whether a functional protein is expressed. These changes occur during DNA replication, therefore Opa phase variation is dependent on replication. Opa proteins interact with host cell receptors and are the only GC proteins that have been shown to stimulate transcytosis of polarized cells (5, 18-20).

The transcytosis assays in this thesis were started with pilated bacteria not expressing Opa. If transcytosis is dependent on Opa expression the rate of transcytosis would be dependent upon on the rate of bacterial replication in our transcytosis assays. In Wang et. al. (1998) transcytosis experiments were done with Opa expressing GC and they found that these bacteria crossed the monolayer in as little as six hours, similar to the rate of both MC and our fast trafficking mutants (7, 8, 12). We analyzed GC colonies after transcytosis in our experiments (chapter 4) and all of these colonies expressed at least one Opa. These data taken together strongly support the model that Opa expression is a primary requirement for transcytosis across T84 cells. Since Opa phase variation is coupled to replication, a strain of GC that replicates more quickly inside cells would also have a fast trafficking phenotype, as is observed with the *fit* mutants.

1.2 Pile recombination slows transcytosis rate

Previous work has shown that GC undergoes *pile* recombination during transcytosis (10). One step of recombination involves the cleavage of the DNA phosphate backbone to allow strand exchange between the donor and template strands (Chapter 1 - Figure 1). When the DNA phosphate backbone is broken, in either a single

or a double stranded break, the replication complex cannot advance without a loss of sequence integrity (4). It is therefore crucial that the bacteria stop replication whenever there is a break in the DNA.

In *E. coli*, the SOS pathway stops replication in the presence of a DNA break (3, 4). Whenever there is a single or double stranded DNA break, the multifunctional protein RecA becomes activated. Activated RecA interacts with LexA, the global repressor of SOS genes, causing LexA to undergo autocleavage and degradation. Once LexA is degraded the SOS proteins are transcribed and DNA damage is repaired. One of the proteins regulated by LexA is SfiA. SfiA sequesters FtsZ, which prevents septal ring formation and thus cell division (9). This allows recombination between the two copies of the chromosome to occur and lesions to be repaired. In this way sequence fidelity is maintained between the two daughter cells.

GC does not have a homologue of *lexA* or *sfiA*, nor does it have an SOS response to DNA damage, though it does have a functioning excision repair system (2). This does not mean that GC does not have a regulated response to DNA breakage, just that the system may not be regulated in the same way as *E. coli*. GC does have *recA* and a functioning excision repair system that responds to DNA damage from a variety of environmental insults, including UV radiation and mitogenic chemicals (11). It is therefore a reasonable hypothesis that GC would have another system to ensure fidelity during DNA repair. Perhaps a different regulatory system is required because GC is naturally competent and is capable of taking up free DNA and recombining it with its genome (6, 13, 14). The bacteria therefore would need a regulatory system that responds to a signal other than free 3' and 5' DNA ends.

As mentioned above GC undergoes specific recombination at the *pilE* locus during transcytosis. This process causes single and double strand breaks in the chromosome, which would antagonize replication. If *pilE* recombination is stimulated when the bacteria are within host cells this should slow down replication, *opa* phase variation, and transcytosis. This idea is supported by the unpublished finding that the fast trafficking *fit* mutant did not undergo *pilE* recombination during transcytosis and the *recA* mutant strain had a fast trafficking phenotype (Chapter 4). These data strongly support a model where *fit* stimulation of *pilE* recombination causes slower GC trafficking (Figure 2).

1.3 FitA functions to abrogate FitB activity

Part of this thesis examines the structure and function of FitA and FitB *in vitro*. The FitA/B heterooctomer complex can bind to its own promoter through the RHH domain of FitA. This binding is likely to have an autoregulatory function, though we have not been able to demonstrate it. Additionally, the crystal structure revealed that the N-terminus of FitA interacts with FitB. The N-terminus of FitA consists of an α -helix interacting with a hydrophobic pocket in FitB and an unstructured strand that folds over the FitB PIN domain. Most importantly, Arg68 of FitA was found to insert into the proposed PIN domain and interact with the four highly conserved acidic residues thought to make up the active site. These two lines of data suggest that the primary function of FitA is to block the activity of FitB.

1.4 *FitB is responsible for the fast trafficking phenotype*

FitB is likely responsible for the fast trafficking phenotypes we have observed with the *fit* mutant. A number of lines of evidence support this hypothesis. The first comes from the types of mutations in the *fit* locus identified in the initial screen for fast trafficking GC (8). Two mutations were isolated, one in the promoter of *fitAB* and one within the coding region of *fitB*. These mutants had the same fast trafficking phenotype suggesting that FitB is sufficient for normal regulation of transcellular trafficking speed. A second line of evidence comes from the FitAB homologues in *Shigella flexneri*. The *S. flexneri* FitB homologue MvpT has toxin activity, while its FitA homologue MvpA antagonizes MvpT toxin activity (15). In this system it is the FitB homologue that has activity and the FitA homologue functions only to modulate that activity. Finally FitB contains a proposed PIN domain. The PIN domains in other proteins are thought to have nuclease activity (1). If FitB is a nuclease, its nicking of DNA would inhibit replication, perhaps even stimulating recombination, and give rise to the phenotypes observed with the *fit* mutants. The only drawback to this model is that we have not been able to detect any nuclease activity for FitB. Our best explanation for this lack of observable activity is that we have only been able to solubilize FitB in the presence of FitA, whose function seems to be to inhibit FitB activity.

I have developed two hypotheses to explain how *in vivo* FitB could act like a nuclease. First, FitB may be soluble in the absence of FitA, perhaps by interacting with another protein. Through this interaction the PIN domain would be exposed and it would have nuclease activity. We do not have any evidence at this time to suggest that FitB

interacts with other proteins. Second, FitA blocks FitB activity only when bound to the FitID inverted repeat. There are 12 other sequences within in GC genome that FitA could potentially bind to (Table 1). However, all of these are only an eight base pair halfsite (FitID or FitPP), not the inverted repeat found in the *fit* promoter region. We would predict that these halfsites would only interact with half of the FitA/B heterooctomer. Perhaps FitA/B can bind to two of these halfsites, acting as a scaffold between distant portions of the chromosome (figure 2). This new conformation could put strain on the heterooctomer, changing how FitA and FitB interact so that the PIN domain can become active. Surprisingly, we found FitID in close proximity to two other loci that also affect GC trafficking speed: *mod* (NG0641) and *pglC* (NG0084, Table 1). In both of these, the FitID is found within the coding region of the ORF, suggesting that the FitID, and predicted FitAB binding at this site is not involved in transcriptional regulation of *mod* or *pglC*. This circumstantially supports our model of FitA/B scaffold activity at these sites.

One additional possible activity of the FitA/B complex that I have not yet discussed is transcriptional regulation outside of the *fit* operon. FitA is a DNA binding protein. All of the phenotypes associated with this locus (regulation of trafficking, intracellular growth and *pilE* AV) could be attributed to transcriptional regulation. However, FitID and FitPP sequences not associated with the *fit* locus are not in predicted promoter regions (Table 1). If the FitA/B complex is a transcriptional regulator it would likely interact with a sequence(s) I have not yet identified.

1.5 Final remarks

In this thesis I set out to better understand the intracellular life of GC through the biochemical analysis of the fast trafficking proteins FitA and FitB. I found that these proteins stimulate *pilE* AV in addition increasing trafficking and intracellular growth rates. Our crystal structure analysis predicts that FitB is a nuclease. It is through this activity that I propose FitB stimulates *pilE* recombination, subsequently causing a decrease in intracellular growth and trafficking speed. In Figure 2, I summarize how I think FitA and FitB regulates the various processes involved in transcytosis.

This model of *fit* activity show the connection between all of the phenotypes we have observed. Yet there are still questions to be answered. How does FitB stimulate the two disconnected processes of recombination and slip-strand misrepair? What is the stimulus within the cell that activates FitB? What is the target of FitB nuclease activity and how does this stimulate *pilE* recombination? Further experimentation will elucidate the answer to these questions and thus gain a better understanding of how GC interacts with the host cell and how it establishes and maintains the carrier state.

1. **Arcus, V. L., K. Backbro, A. Roos, E. L. Daniel, and E. N. Baker.** 2004. Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J Biol Chem* **279**:16471-8.
2. **Black, C. G., J. A. Fyfe, and J. K. Davies.** 1998. Absence of an SOS-like system in *Neisseria gonorrhoeae*. *Gene* **208**:61-6.
3. **Foster, P. L.** 2005. Stress responses and genetic variation in bacteria. *Mutat Res* **569**:3-11.

4. **Fry, R. C., T. J. Begley, and L. D. Samson.** 2005. Genome-wide responses to DNA-damaging agents. *Annu Rev Microbiol* **59**:357-77.
5. **Gray-Owen, S. D., C. Dehio, A. Haude, F. Grunert, and T. F. Meyer.** 1997. CD66 carcinoembryonic antigens mediate interactions between Opa- expressing *Neisseria gonorrhoeae* and human polymorphonuclear phagocytes. *Embo J* **16**:3435-45.
6. **Hamilton, H. L., and J. P. Dillard.** 2006. Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. *Mol Microbiol* **59**:376-85.
7. **Hopper, S., B. Vasquez, A. Merz, S. Clary, J. S. Wilbur, and M. So.** 2000. Effects of the immunoglobulin A1 protease on *Neisseria gonorrhoeae* trafficking across polarized T84 epithelial monolayers. *Infect Immun* **68**:906-11.
8. **Hopper, S., J. S. Wilbur, B. L. Vasquez, J. Larson, S. Clary, I. J. Mehr, H. S. Seifert, and M. So.** 2000. Isolation of *Neisseria gonorrhoeae* mutants that show enhanced trafficking across polarized T84 epithelial monolayers. *Infect Immun* **68**:896-905.
9. **Huisman, O., R. D'Ari, and S. Gottesman.** 1984. Cell-division control in *Escherichia coli*: specific induction of the SOS function SfiA protein is sufficient to block septation. *Proc Natl Acad Sci U S A* **81**:4490-4.
10. **Ilver, D., H. Kallstrom, S. Normark, and A. B. Jonsson.** 1998. Transcellular passage of *Neisseria gonorrhoeae* involves pilus phase variation. *Infect Immun* **66**:469-73.

11. **Kline, K. A., E. V. Sechman, E. P. Skaar, and H. S. Seifert.** 2003. Recombination, repair and replication in the pathogenic *Neisseriae*: the 3 R's of molecular genetics of two human-specific bacterial pathogens. *Mol Microbiol* **50**:3-13.
12. **Merz, A. J., D. B. Rifkenberg, C. G. Arvidson, and M. So.** 1996. Traversal of a polarized epithelium by pathogenic *Neisseriae*: facilitation by type IV pili and maintenance of epithelial barrier function. *Mol Med* **2**:745-54.
13. **Sarubbi, F. A., Jr., and P. F. Sparling.** 1974. Transfer of antibiotic resistance in mixed cultures of *Neisseria gonorrhoeae*. *J Infect Dis* **130**:660-3.
14. **Saunders, J. R.** 1979. Specificity of DNA uptake in bacterial transformation. *Nature* **278**:601-2.
15. **Sayeed, S., L. Reaves, L. Radnedge, and S. Austin.** 2000. The stability region of the large virulence plasmid of *Shigella flexneri* encodes an efficient postsegregational killing system. *J Bacteriol* **182**:2416-21.
16. **Stern, A., M. Brown, P. Nickel, and T. F. Meyer.** 1986. Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. *Cell* **47**:61-71.
17. **Stern, A., and T. F. Meyer.** 1987. Common mechanism controlling phase and antigenic variation in pathogenic *neisseriae*. *Mol Microbiol* **1**:5-12.
18. **van Putten, J. P., T. D. Duensing, and R. L. Cole.** 1998. Entry of OpaA+ gonococci into HEP-2 cells requires concerted action of glycosaminoglycans, fibronectin and integrin receptors. *Mol Microbiol* **29**:369-79.

19. **Virji, M.** 1996. Studies on the molecular mechanisms of meningococcal interactions with human cells. Towards anti-adhesion measures for the control of meningococcal disease. *Adv Exp Med Biol* **408**:113-22.
20. **Wang, J., S. D. Gray-Owen, A. Knorre, T. F. Meyer, and C. Dehio.** 1998. Opa binding to cellular CD66 receptors mediates the transcellular traversal of *Neisseria gonorrhoeae* across polarized T84 epithelial cell monolayers. *Mol Microbiol* **30**:657-71.

Figure legends

Figure 1 FitA/B scaffold model

A. Schematic representation of the interaction of the FitA/B heterooctomer interacting with the GC chromosome at the *fitA-B* promoter.

B. Schematic of the Heterooomer interacting with two FitID sequences on distant regions of the chromosome.

In A the PIN domain of FitB is interacting with the tails of FitA, representing “off” conformation with Arg68 of FitA interacting with the four Acidic residues of the active site. In B the PIN domain is open, representing the “on” conformation when the heterooctomer confirmation is changed due to the stress of binding different pieces of DNA. In this predicted conformation the change of binding would alter the interaction between FitA and FitB such that Arg68 is no longer in the PIN domain active site and thus the “cryptic” nuclease activity is possible.

Figure 2 Web of FitA/B activity

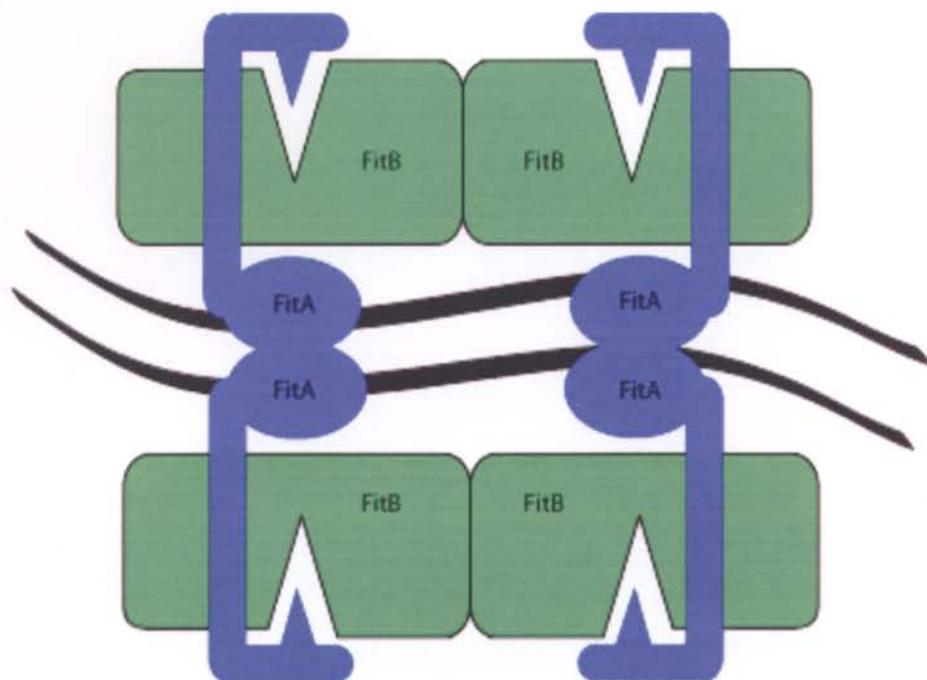
This is a summery of the finding of this thesis with respect to how FitA/B controls transcytosis rate. Arrows indicate stimulation and black squares indicate inhibition. The outside ring shows the two fold connection between replication and transcytosis.

1. In the right arm of the ring replication stimulates slip-strand misrepair
2. Slip-strand misrepair controls Opa phase variation and expression.
3. Opa expression is sufficient to drive transcytosis.
4. Transcytosis stimulates *pilE* antigenic variation

5. In the left arm of the ring replication and RecA activity have a reciprocal relationship where recombination is dependent on replication, yet activated RecA slows replication.
6. *PilE* antigenic variation is dependent on RecA.
 - 6.1 RecA mutant has a fast trafficking phenotype, therefore RecA slows transcytosis in this model.
7. FitA inhibits FitB activity, by repressing *fitA/B* transcription and blocking the PIN active site.
8. Summary of FitB activity: FitB slows intracellular growth, Increases *pilE* AV, Slows transcytosis and has some unknown effect on Opa expression. The relationship with RecA and or Slip-strand missrepair is still unclear.

Figure 1

A.



B.

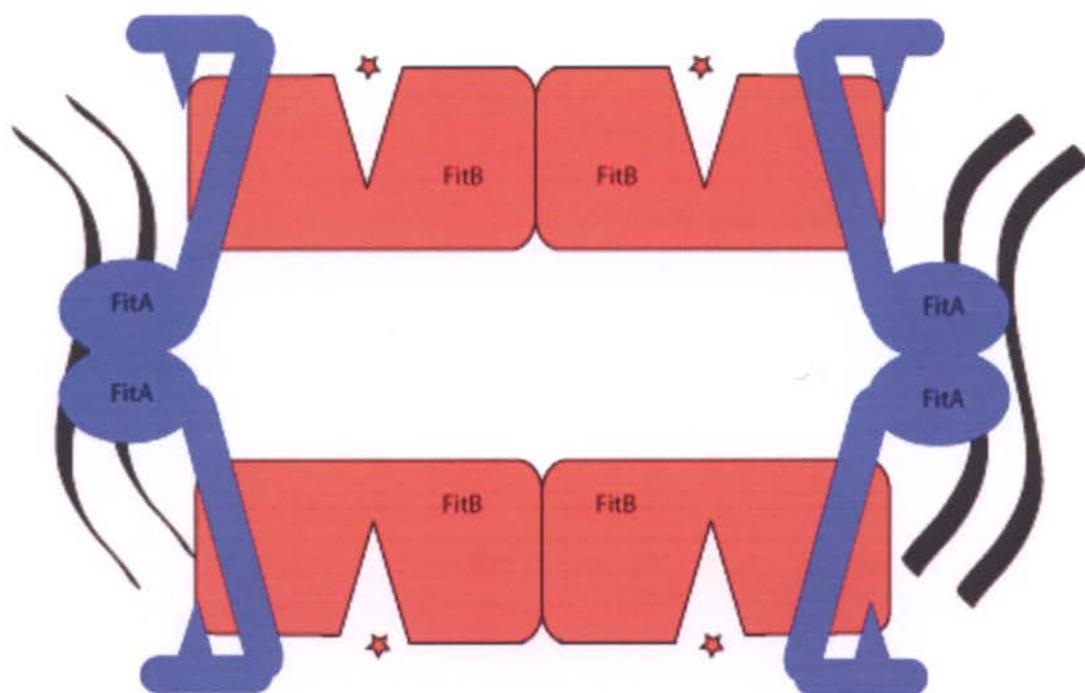


Figure 2

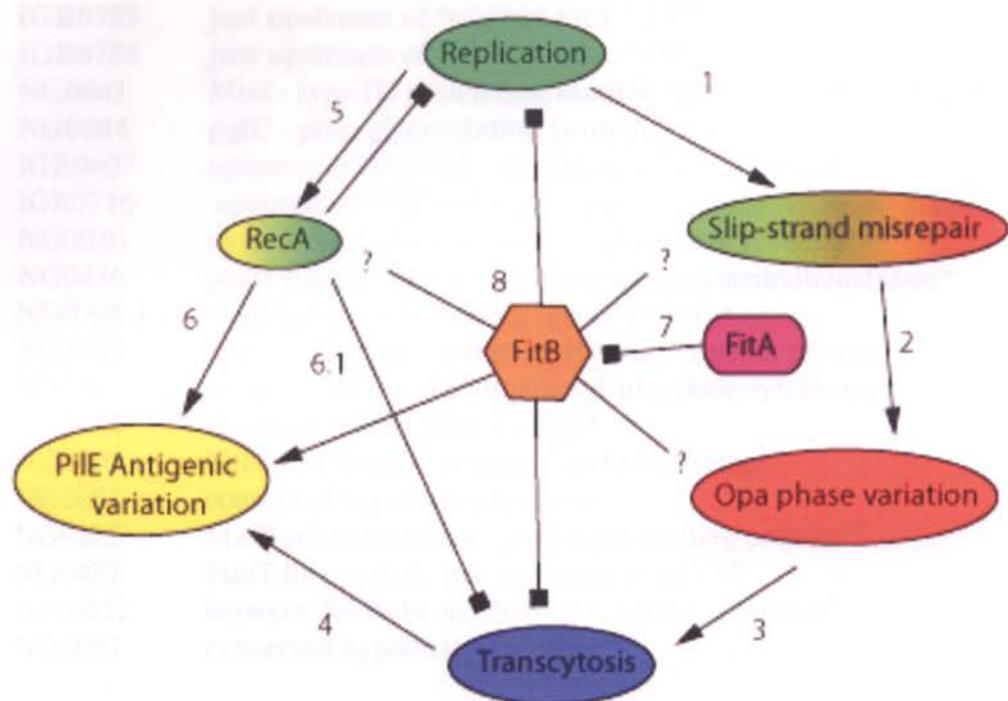


Table 1 Location of FitIS and FitPP on the FA1090 genome

IGR0755	just upstream of NG0908 FitA “A9”^a
IGR0755	just upstream of NG0908 FitA “A9”^a
NG0641	Mod - type III restriction/modification methylase, “A11”^{*a}
NG0084	pglC - pilin glycosylation protein, “A12”^{*a}
IGR0407	upstream of NG0492 - hypothetical protein (Akt/PKB)
IGR0716	upstream of NG0864.1 ggt - gamma-glutamyltranspeptidase
NG0036	dxps, dxs a 1-deoxyxylulose-5-phosphate synthase*
NG0436	panB 3-methyl-2-oxobutanoate hydroxymethyltransferase*
NG0706.1	hemagglutinin/hemolysin related protein*
NG1433	ABC transporter - nitrate or taurine permease protein*
NG1541	murE - UDP-N- acetylmuramyl-tripeptide synthetase*
NG1648	invertase related gene 7 - Irg7*
NG1805	Neisseria-specific protein - uncharacterized*
NG2051	conserved hypothetical protein*
NG0225	MafB related adhesin - glycolipid binding protein* [@]
NG0452	FimT like protein, just upstream of pilV* [@]
IGR0532	between NG0634 and NG0635, neither promoter [@]
NG2051	conserved hypothetical protein* [@]

* within a predicted coding region of a protein

@ FitPP binding site

^a A9, A11 and A12 are three transposon mutations isolated causing fast trafficking phenotype in Hopper et. al. 2000