DNA UPTAKE SPECIFICITY OF *HAEMOPHILUS INFLUENZAE*

by

GRANT ALEXANDER POJE

B.Sc., Simon Fraser University, 1996

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Department of Zoology)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

June, 2000

© Grant Alexander Poje, 2000
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of **Zoology**

The University of British Columbia
Vancouver, Canada

Date **June 20, 2007**
Abstract

DNA binding and uptake by the naturally transformable Gram-negative bacterium *Haemophilus influenzae* Rd has been studied for over twenty years. It is well characterized that *H. influenzae* cells preferentially bind and take up DNA from their own species. This preferential uptake is dependent on the ability of cells to recognize a 9-base pair uptake signal sequence (USS), 5’-AAGTGCAGG, in the DNA molecule. Genomic analysis has shown that there are 1465 copies of the 9-base pair uptake sites. Further analysis revealed an extended consensus region of 29 base pairs which includes the core region and two down stream 6-base pair A/T-rich regions, each spaced about one helix turn apart.

To determine properties of the DNA molecule that, in addition to the presence of the USS, are necessary for uptake by *H. influenzae* I designed oligonucleotides with variations in the regions flanking the core USS. Oligonucleotides were varied in both the length and base composition of 3' and 5' flanking sequences. I showed that nucleotides 5' to the USS are required for high levels of binding and uptake of DNA. Also, I showed that both length and base composition of the 3' flanking region greatly affect binding and uptake. If sequence 3' to the USS is G/C rich, uptake proceeds at a very low level. However, if DNA lacks a 3' sequence, both binding and uptake are abolished. Based on these findings I propose a speculative model of how cells bind and take up DNA in a sequence specific manner. I attempt with this model to supplement other proposed models which do not address the initial steps of binding and uptake by *H. influenzae*.
A second goal of my research was to isolate the receptor that allows competent cells to preferentially bind and take up sequence specific DNA. The method I used was UV laser crosslinking. Conditions used in crosslinking experiments were varied including the time of incubation of DNA and cells, the presence and absence of competing DNAs and attachment of bulky groups to the DNA to prevent uptake. Following many attempts to isolate the receptor I concluded that under the conditions used, it was impossible to isolate the receptor using laser crosslinking.
TABLE OF CONTENTS

ABSTRACT.................................................................................................................. ii
TABLE OF CONTENTS.................................................................................................. iv
LIST OF TABLES.......................................................................................................... viii
LIST OF FIGURES......................................................................................................... ix
LIST OF ABBREVIATIONS........................................................................................... xi
ACKNOWLEDGMENTS................................................................................................... xii

CHAPTER ONE............................................................................................................. 1
Introduction
I Competence, Transformation and DNA uptake.......................................................... 2
  1.1 Natural competence............................................................................................... 2
  1.2 DNA uptake and natural transformation............................................................... 2
  1.3 Uptake and transformation by gram-positive bacteria........................................... 3
  1.4 Uptake and transformation by gram-negative bacteria.......................................... 3
  1.5 Uptake and transformation by H. influenzae......................................................... 4
  1.6 Processing and translocation of DNA................................................................. 5
  1.7 Uptake requirements............................................................................................. 7
  1.8 Proteins involved in translocation and recombination......................................... 8
II UV Laser Crosslinking............................................................................................... 13
  2.1 Background........................................................................................................... 15
  2.2 Mechanism of crosslinking.................................................................................... 16
  2.3 Stability................................................................................................................ 16
  2.4 Efficiency............................................................................................................. 16
  2.5 Specificity............................................................................................................. 17

CHAPTER TWO............................................................................................................. 18
Materials and Methods
  2.1 Strains, Plasmids and Oligonucleotides............................................................... 18
  2.2 Culture conditions............................................................................................... 18
  2.3 Media.................................................................................................................. 21
2.4 Transformation of *H. influenzae* ................................................................. 23
  2.4.1 Competence induction ............................................................................. 23
  2.4.2 Transformation using linear DNA .......................................................... 23
  2.4.3 Transformation using circular plasmid DNA .......................................... 23
2.5 *E. coli* plasmid transformation ................................................................. 24
2.6 Nucleic Acid Techniques ............................................................................. 24
  2.6.1 Isolation of plasmid DNA ...................................................................... 24
  2.6.2 Isolation of *H. influenzae* chromosomal DNA ...................................... 24
  2.6.3 DNA labeling ......................................................................................... 25
  2.6.4 DNA sequencing .................................................................................... 26
  2.6.5 Oligonucleotides .................................................................................. 26
  2.6.6 Biotinylated oligonucleotides ................................................................. 26
  2.6.7 Cloning of Oligonucleotides ................................................................ 27
    2.6.7.1 Cloning USS-1 and USS-R ............................................................... 27
    2.6.7.2 Cloning USS-50 ............................................................................ 29
2.7 Electrophoresis ............................................................................................. 29
  2.7.1 Agarose ................................................................................................. 30
  2.7.2 Polyacrylamide Gel Electrophoresis (PAGE) ........................................... 30
    2.7.2.1 DNA ............................................................................................. 30
    2.7.2.2 Crosslinked DNA-protein complexes .............................................. 30
2.8 Uptake experiments ...................................................................................... 31
2.9 Binding experiments .................................................................................... 31
2.10 UV laser crosslinking ................................................................................ 32

CHAPTER THREE .................................................................................................. 35
Characterization of uptake requirements of *H. influenzae*
  3.1 Is the 29 bp USS sufficient for uptake? ...................................................... 35
    3.1.1 Uptake of a 'perfect' USS ................................................................. 35
  3.2 Effect of USS position on binding and uptake ............................................ 38
    3.2.1 Oligonucleotides .............................................................................. 38
      3.2.1.1 Binding and uptake of 30 bp DNA fragments ............................... 38
      3.2.1.2 Binding and uptake of 50 bp fragments ....................................... 39
    3.2.2 Plasmid DNA .................................................................................. 42
3.3 Length dependence of uptake. Is USS-1 too short to be taken up?........ 44
3.4 Effect of 3' consensus flanking sequence on DNA binding and uptake ............................................. 46
  3.4.1 Binding and uptake of DNA containing the lowest frequency base composition in 3' flanking consensus region .......... 46
  3.4.2 Binding and uptake of DNA lacking sequence 3' to the 9 bp core........................................................................... 48
3.5 Competition of various DNAs for uptake of chromosomal DNA .................................................................................. 50
  3.5.1 Competition by chromosomal DNAs and 29 bp fragments .................................................................................. 52
  3.5.2 Competition by chromosomal DNAs and 50 bp fragments .................................................................................. 52
3.6 Discussion of uptake results ....................................................................................................................................... 55
3.7 Future experiments .................................................................................................................................................. 56

CHAPTER FOUR .............................................................................................................................................. 57
UV Laser Crosslinking
Crosslinking labeled DNA to the receptor ........................................................................................................... 57
  4.1 Time interval for uptake of USS-50 ....................................................................................................................... 59
  4.2 Does crosslinking increase the amount of DNA associated with the outside of cells? .................................................... 59
  4.3 Crosslinking using labeled USS-50 ....................................................................................................................... 61
  4.4 Crosslinking following varied time of incubation .............................................................................................. 64
  4.5 Crosslinking in a rec-2 mutant background ........................................................................................................ 66
    4.5.1 Crosslinking in wild-type and rec-2 backgrounds ........................................................................................ 66
  4.6 Biotinylated oligonucleotides .................................................................................................................................... 68
    4.6.1 Uptake experiments ............................................................................................................................................. 68
    4.6.2 Crosslinking experiments .................................................................................................................................... 72
  4.7 Calculations .......................................................................................................................................................... 74
    4.7.1 Scenario 1; estimate 100 receptors/cell and a 200 kDa receptor ........................................................................ 74
    4.7.2 Scenario 2; estimate 10 receptors/cell and a 50 kDa receptor ........................................................................... 75
  4.8 Discussion of crosslinking results ......................................................................................................................... 75
  4.9 Future experiments .................................................................................................................................................. 76
CHAPTER FIVE.................................................................................................................. 78

Hypothetical model for uptake by *H. influenzae*

5.1 Discussion of uptake characterization results......................................................... 78

5.1.1 Model for uptake of *H. influenzae*..................................................................... 78

5.1.2 Interaction of oligonucleotides with the receptor complex......................... 82

BIBLIOGRAPHY............................................................................................................. 87
LIST OF TABLES

Table 2.1 Bacterial strains used in this study................................................................. 19
Table 2.2 Plasmids used in this study.............................................................................. 19
Table 2.3 Oligonucleotides used in this study................................................................. 20
Table 2.4 Components of non-commercial media......................................................... 21
LIST OF FIGURES

Figure 1.1 Uptake and incorporation of homologous DNA by *H. influenzae*...... 6

Figure 1.2 Alignment of the core 9-bp USS and flanking regions.................. 10

Figure 1.3. Proteins thought to be responsible for uptake of DNA by *H. influenzae*.................................................................................................................. 12

Figure 1.4 General strategy for identification of the DNA binding receptor of *H. influenzae* using UV laser crosslinking.................................................. 14

Figure 2.1 Cloning of USS-1, USS-R and USS-50............................................. 28

Figure 2.2 Schematic drawing of apparatus used for UV laser crosslinking.... 33

Figure 3.1 Uptake of 'perfect' USS containing DNA is not greater than a DNA that does not contain a USS................................................................. 37

Figure 3.2 Binding and uptake of DNAs by competent cells.......................... 40

Figure 3.3 Binding and uptake by competent cells........................................ 41

Figure 3.4 Uptake of cloned USS-1................................................................. 43

Figure 3.5 Uptake of different length oligos.................................................. 45

Figure 3.6 Binding of 50 bp DNA fragments by competent cells................... 47

Figure 3.7 Binding of 50 bp DNA fragments by competent cells................... 49

Figure 3.8 Schematic representation of a competition assay.......................... 51

Figure 3.9 Chromosomal DNA and 29 bp fragments.................................... 53

Figure 3.10 Chromosomal DNA and 50 bp fragments.................................... 54

Figure 4.1 Uptake of labeled USS-50 as a function of time............................. 58

Figure 4.2 Crosslinking increases the amount of DNA associated with the outside of cells............................................................................................................. 60
Figure 4.3 Crosslinking using short and long DNA fragments........................................ 62

Figure 4.4 Crosslinking USS-50 to wild-type *H. influenzae* cells........................................ 63

Figure 4.5 Crosslinking experiments using USS-50, USS-1 and MAP7 DNAs as bait........................................ 65

Figure 4.6 Crosslinking using MAP7 chromosomal DNA........................................ 67

Figure 4.7 Crosslinking in rec-2 and wild-type backgrounds........................................ 69

Figure 4.8 Biotinylated oligonucleotide........................................ 71

Figure 4.9 Illustration of interaction of competent *H. influenzae* cells with biotinylated DNA attached to agarose beads........................................ 73

Figure 5.1 Hypothetical model for binding and uptake by *H. influenzae*........................................ 79

Figure 5.2 Model for uptake of oligos lacking 5' sequence........................................ 81

Figure 5.3 Model for binding of an oligo having a G/C rich 3' flanking region........................................ 83

Figure 5.4 Binding of USS-50-Ri........................................ 84
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>BHI</td>
<td>brain heart infusion (rich culture medium)</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>cfu</td>
<td>colony forming unit</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNasel</td>
<td>deoxyribonuclease I</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>μg</td>
<td>microgram</td>
</tr>
<tr>
<td>MIV</td>
<td>&quot;M-four&quot;; a nutrient-limited <em>H. influenzae</em> competence induction medium</td>
</tr>
<tr>
<td>ml</td>
<td>milliliter</td>
</tr>
<tr>
<td>NAD</td>
<td>nicotinamide adenine dinucleotide</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram</td>
</tr>
<tr>
<td>nov</td>
<td><em>H. influenzae</em> novobiocin resistance allele</td>
</tr>
<tr>
<td>novF</td>
<td>novobiocin resistant</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>oligo</td>
<td>oligonucleotide</td>
</tr>
<tr>
<td>r</td>
<td>IUB code for purine</td>
</tr>
<tr>
<td>sBHI</td>
<td>brain heart infusion supplemented with haemin and NAD</td>
</tr>
<tr>
<td>w</td>
<td>IUB code for adenosine or thymidine</td>
</tr>
<tr>
<td>y</td>
<td>IUB code for cytidine or thymidine</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I would like to thank Dr. R. Redfield for providing excellent guidance and instruction on the work performed in this thesis. I would also like to thank the members of my supervisory committee for their insightful comments: Dr. Beatty, Dr. Brock and Dr. Roberge. Also, I would like to acknowledge all the members of the Redfield lab who made coming to lab everyday an enjoyable experience. I would like to acknowledge Lisette Ungethum for her assistance on experiment 3.2.2.

Finally, I would like to thank my supporting cast. They remain in the background, but support me in everything I do and deserve much of the credit for who I have become. Thank you Sonia, Mom and Tracy.
CHAPTER ONE

Introduction

Competence is the ability of cells to bind and take up DNA from their environment [1, 2]. The ability of cells to develop competence has received much attention in the past and continues to attract investigation. In this thesis I have tried to clarify some of the requirements for DNA uptake in *Haemophilus influenzae*. In addition to this I have attempted to isolate the DNA binding receptor that allows *H. influenzae* to recognize and preferentially take up homologous DNA from its environment. The technique I chose to use to isolate the receptor is UV laser crosslinking.

This project is important for a number of reasons. *H. influenzae* is a naturally transformable gram-negative facultative anaerobe of the family Pasteurellaceae. It is commensal in the upper respiratory tract of humans and is responsible for causing such diseases as meningitis and otitis media [3]. Currently there are vaccines for some strains of *H. influenzae*. However, isolating the receptor may provide a new target for drugs in an attempt to eliminate strains of *H. influenzae* for which there are currently no vaccines available. Also, isolating the receptor may assist others trying to identify proteins responsible for sequence-specific uptake in other gram-negative bacteria.

Where possible I have cited original work. However, there are also many useful review articles that I have used. Two comprehensive, though slightly dated, review articles on transformation are “Genetic Transformation” by Smith and Danner [2], and “Transformation in *Haemophilus*: a problem in membrane biology” by Kahn and Smith [4]. Two useful review articles on DNA uptake by bacteria are “DNA uptake
in *Haemophilus* transformation" by Goodgal [1] and a very recent article by Dubnau called “DNA uptake in bacteria” [5].

I Competence, Transformation and DNA uptake

1.1 Natural competence.
Competence development requires proper expression of proteins within the cytoplasm and delivery of specific proteins to the cell surface. Functions performed by these competence proteins include DNA binding and transport across cell walls and membranes. Natural competence differs from artificial competence in that artificial measures such as treatment with lysozyme [6] or calcium ions [7] are not required.

1.2 DNA uptake and natural transformation
Transformation arises when naturally competent bacteria take up DNA and integrate new alleles into their genome. Natural transformation can be divided into three steps: 1) competence development, 2) DNA uptake 3) DNA integration. Natural transformation is widespread, occurring in both gram-positive and gram-negative bacteria. *Bacillus subtilis* and *Streptococcus pneumoniae* are examples of naturally competent gram-positive bacteria; *H. influenzae* and *Neisseria gonorrhoeae* are gram-negative examples.

DNA uptake is an integral part of transformation. Interestingly, gram-positive and gram-negative bacteria have evolved significantly different systems of binding and uptake, possibly due to differences in their cell wall composition and the presence of a thick peptidoglycan layer surrounding gram-positive bacteria. There are two main differences in DNA uptake between the two types of bacteria. The first is the ability of some gram-negative cells to take up specific DNA molecules, which is
dependent on the presence of short repeated sequences in the DNA. Gram-positive bacteria take up DNA without sequence preference. The second difference is the ability of gram-negative bacteria to take up both strands of DNA into a nuclease resistant form. Gram-positive bacteria degrade one strand on their cell surface prior to uptake [2]. These points will be covered in more detail below.

1.3 Uptake and transformation by gram-positive bacteria.
Transformation in both *B. subtilis* and *S. pneumoniae* has been studied extensively. Upon development of competence in these bacteria, double stranded DNA is bound to the outside of the cell [8]. In *S. pneumoniae*, following binding the DNA is fragmented on the cell surface, producing single strand nicks which are converted to double strand breaks [9-11]. Cleavage at the cell surface may serve to generate new ends near proteins involved in uptake, allowing entry of the DNA into the cell by its newly formed end. If uptake were to proceed only from pre-existing ends, then as the length of the DNA fragment increased, search time for an end would also increase, which could place limitations on the ability of cells to hold onto DNA long enough for uptake to occur. As DNA moves across the cytoplasmic membrane one strand is degraded and the remaining strand is coated by Ssb, a single stranded binding protein. Following this, DNA is incorporated into the host chromosome by RecA, a protein required for homologous recombination [12].

1.4 Uptake and transformation by gram-negative bacteria.
Gram-negative bacteria have a three-layered cell envelope composed of an inner cytoplasmic membrane, a peptidoglycan layer, and an outer membrane. The mechanism of DNA movement across these layers is unclear.
Uptake systems of gram-negative bacteria can be divided into two general categories: homologous and heterologous. *Acinetobacter calcoaceticus* is a naturally transformable gram-negative bacterium with heterologous uptake. It takes up DNA without regard to the source [13, 14]. Conversely, in both *Haemophilus* and *Neisseria*, DNA uptake is homospecific; that is, competent cells recognize and preferentially take up DNA from their own genus [15]. These gram-negative bacteria take up both strands, converting it into a nuclease resistant form. It is not known if DNA taken up by *A. calcoaceticus* is taken up as a double stranded molecule.

Both *H. influenzae* and *N. gonorrhoeae* have been used as models to study the transformation of gram-negative bacteria. The sequence specific binding in both bacteria results from recognition of short uptake sequences which are abundant in their genomes. DNA that lacks these uptake sequences transforms poorly and competes inefficiently against uptake sequence-containing DNA [15-17].

### 1.5 Uptake and transformation of *H. influenzae*.

*H. influenzae* strains are classified into six serotypes (a-f) based on the antigenic properties of their capsule. The capsule is composed of a negatively charged porous matrix consisting of a phosphodiester linked ribose-ribitol copolymer [18]. Isolates that lack this polysaccharide capsule are classified as ‘nontypable’. These are also termed rough because absence of the capsule causes colonies to lose the sheen of smooth capsulated colonies. The capsule does not prevent translocation of DNA into the cell, as it has been shown that encapsulated cells transform as efficiently as non-encapsulated strains [4, 19]. The lab strain we use (Rd) is a non-encapsulated, non-pathogenic derivative of a serotype d strain.
The proteins that allow sequence specific binding in *H. influenzae* have not yet been characterized, but many features of uptake and transformation have been studied (transformation of *H. influenzae* is outlined in Figure 1.1). DNA containing uptake sequences interacts with a postulated receptor protein complex and crosses the outer membrane as intact double stranded DNA (Figure 1.1, Steps 1 and 2). Work by Deich and Smith (1980) has clarified many aspects of the uptake process. They found that cells were only able to take up 3 to 8 DNA molecules and suggested that cell surface receptors acted only once. From this, they estimated an uptake rate of 500-1000 nucleotides per second. After DNA is taken up by cells it becomes resistant to external nucleases and cannot be eluted from cells by high salt washes.

### 1.6 Processing and translocation of DNA

Translocation across the inner membrane of *H. influenzae* requires a free end (Figure 1.1 step 3) [20, 21]. Circular DNA or hairpin structures are able to move across the outer membrane, but are not transported into the cytoplasm [22, 23]. Transport across the inner membrane results in complete degradation of the 5' incoming strand and partial degradation of the 3' strand [4, 22, 24]. After crossing the inner membrane, the remaining 3' leading strand is available for homologous recombination. Transport of DNA into *N. gonorrhoeae* occurs in the same way as it does in *H. influenzae* [22].

Single stranded DNA has not been isolated from the cytoplasm of cells. This may be because only a short length of single-stranded DNA is present at any given time as the DNA crosses the inner membrane [5]. Homologous recombination occurs when incoming DNA replaces regions of the recipient chromosome containing sequence
Figure 1.1. Uptake and incorporation of homologous DNA by H. influenzae.

Step 1; Receptor binds DNA containing a USS.

Step 2; Movement of DNA across outer membrane into periplasmic space/transformasome. DNA becomes resistant to external and internal nucleases.

Step 3; Translocation of DNA across inner membrane into the cytoplasm. The 5' leading strand is completely degraded whereas the 3' leading strand is degraded slowly.

Step 4; Integration of the 3' donor strand (A) into the homologous region of the recipient chromosome (a).
similarity (Figure 1.1, Step 4) [21, 25]. An average of 1.5 kb of the 3' strand of the donor DNA is degraded during the search for homology [22]. Once recombination is initiated, it is rapidly completed and proceeds to the 5' end of the incoming DNA [22].

1.7 Uptake requirements

As mentioned above, DNA uptake by *H. influenzae* is sequence specific. Cells have the ability to recognize and preferentially take up homologous DNA. The uptake specificity of *H. influenzae* depends on the ability of the short 9-base-pair (bp) sequence (5' AAGTGCGGT 3'), called an uptake signal sequence (USS), to interact with the cell surface receptor [16, 26-29]. The USS flanking sequence has also been shown to be important for efficient uptake in *H. influenzae*. The more A/T rich the sequence flanking the USS, the greater the amount of DNA taken up by cells [30]. *N. gonorrhoeae* also binds homospecific DNA by recognizing an unrelated 10 bp USS [16, 28].

*H. influenzae* was the first "living" organism to have its complete genome sequenced [31] following the sequencing of several viral and organellar genomes [32-34]. After sequencing, the frequency and distribution of USSs within the genome was analyzed. Previous uptake and competition experiments had estimated the number of USSs to be close to 600 [26], although the 62% A/T base composition of *H. influenzae* predicted that about 8 USSs were expected to occur by chance. In fact, upon examining the sequence, it was found that 1465 USSs were present in the genome, occurring in both orientations with equal frequency [35]. The over-representation of USSs is also reflected in the frequency of 9 bp sequences differing from the consensus at a single position. There are 764 copies of these singly-mismatched 9 bp consensus sequences, where 254 would be expected by chance [2].
Introduction

One hypothesis to explain the over-representation of USSs in the genome is that their occurrence results from selection for uptake of homologous DNA, which would require selection for both a biased receptor and over-representation of USSs within the *H. influenzae* genome. Others have postulated that USSs might function intracellularly as transcription termination sequences or chi sites [35]. Genomic analysis has shown that these specific roles of USSs are unlikely, however a structural role for the USS has not been ruled out [2]. A third hypothesis is that a biased receptor directly causes USSs to accumulate in the genome [2, 36]. The receptor preferentially binds DNA containing the USS, allowing cells take up that DNA and integrate it into their chromosome.

Within the *H. influenzae* genome, flanking the USSs, there are also regions of consensus (Figure 1.2). When all 1465 copies of the USS were aligned in the plus direction a 29 bp consensus USS was identified, that had the sequence 5' aAAGTGCGGT . rwwwww rwwwww 3', where uppercase letters represent the bases that define the USS, lowercase letters are bases that occur in >50% of the USSs, a dot is any base, r is purine, and w is A or T [35]. According to Smith et al., if the lengths of all the 29-bp USSs and the singly mutated sites are added, there are a total of 2229 sites occupying approximately 3.5% of the genome [35]. In actuality, the USSs do not fully constrain 3.5% of the genome. This is because, as illustrated in Figure 1.2, each position, outside the 9 bp core, has some flexibility in the nucleotide that can occupy it.

1.8 Proteins involved in translocation and recombination.

The DNA-binding receptor has not been isolated from either *H. influenzae* or *N. gonorrhoeae*. However, a number of other proteins required for processing and translocation of DNA in these and other bacteria have been identified. Although the DNA uptake mechanisms of gram-positive and gram-negative bacteria are thought
to be different, many of the proteins known to be involved in uptake and transformation are similar in both bacteria. I will briefly cover the function of proteins involved in competence in gram-positive bacteria and illustrate how they may also function in gram-negative bacteria.

The *B. subtilis* comE operon contains genes involved in different aspects of competence and transformation. *comEA* is the first open reading frame of the *comE* operon [37, 38]. In *B. subtilis* ComEA is required for both DNA binding and transport into the cell [38]. ComEA has homologs in both *N. gonorrhoeae* and *H. influenzae*, however a function of these proteins has not yet been assigned in these organisms. In these bacteria ComEA could potentially act as part of the receptor protein complex as it does in gram-positive bacteria. However, it is probably not the USS recognition protein, since gram-positive bacteria do not preferentially recognize or take up homologous DNA.

In *B. subtilis* ComEC, encoded by the third open reading frame of the *comE* operon, has been shown to be required for transport of DNA, but dispensable for binding. ComEC may form part of an aqueous channel since it contains 6 potential membrane spanning segments [5]. The ComEC homologs, Rec-2 in *H. influenzae* and ComA in *N. gonorrhoeae*, have been shown to play essential roles in transformation [39-41]. Mutations in these proteins do not affect binding or uptake of DNA but greatly reduce transformation frequencies by preventing transport of DNA across the inner membrane. Like Rec-2 mutants, *H. influenzae* cells with mutations in DprA take up DNA into a DNase resistant form but are unable to transform [42, 43]. Thus, in conjunction with Rec-2, DprA may function to transport DNA across the inner membrane of *H. influenzae.*
**Figure 1.2. Alignment of the core 9-bp USS and flanking regions.**
This data arises from analysis of 1465 USSs that occur in the *H. influenzae* genome.
Numbers below each column represent percent base frequencies at each position.
(modified from Smith et al., 1995)
Another protein with an essential role in transformation in *H. influenzae* is PorA, a disulfide oxidoreductase that localizes to the periplasm and is required for competence-associated changes in the protein composition of the membrane [44]. In its absence, DNA binding is abolished. It is possible that PorA is needed for the correct folding of one or more membrane proteins during competence development.

Another gene, comFC [45] has been characterized in *B. subtilis*. The product of this gene resembles the ComF protein of *H. influenzae* [46, 47]. Mutations of comFC decrease transformation 5-10 fold in *B. subtilis*, decreasing transport slightly but not binding. Deletion of comF in *H. influenzae* does not impair the ability of cells to bind DNA [46]. A specific role for the gene product has not yet been elucidated in either organism.

Rec-1, a homolog of *E. coli* RecA, is the only known recombination protein with an identified role in *H. influenzae* transformation. Mutations in Rec-1 lead to a transformation deficient phenotype, not because cells are defective in binding, uptake or transport, but because they are defective in integration of donor DNA into the chromosome [41, 48, 49].

A model of DNA uptake in *H. influenzae* has been proposed (Figure 1.3) [5]. In this model a large protein complex is assembled below the receptor in the periplasmic space. This complex may contain as yet unidentified proteins as well as those listed above (PorA and ComF). The receptor then binds DNA on the outside of the cell and passes the DNA across the outer membrane. DprA and Rec-2 are located on the inner membrane and act to feed DNA through it. These proteins may associate with a yet unidentified nuclease that cleaves the 5' strand of the incoming DNA. Upon entry, Rec-1 aids in the incorporation of DNA into the chromosome.
Figure 1.3 Proteins thought to be responsible for uptake of DNA by *H. influenzae*. 
II UV Laser Crosslinking of DNA to Proteins

Researchers have tried to isolate the receptor that allows *H. influenzae* to take up homologous DNA by employing different approaches with no success. Mutagenesis screens using transposons have been performed, isolating genes that when mutated lead to defects in transformation [50, 51]. This technique has been useful in finding proteins involved in the regulation of competence, nutritional state sensing, DNA binding, uptake and translocation [44, 51-54]

Biochemical approaches have included a comparison of the polypeptides in the membrane fraction of competent and non-competent cells [55, 56] and the isolation of DNA binding proteins from membranes of competent cells [57]. Sutrina and Scocca [58] reported the isolation of a periplasmic protein fraction possessing DNA-binding activity from non-competent cells. They suggested that these proteins may become associated with the cell membrane during competence development. Kahn et al. [59] reported the presence of DNA-binding activity in the cell culture supernatants of certain competence mutants after they were subjected to competence-inducing procedures. Subsequent to this, Concino and Goodgal developed a procedure to label specific cell surface proteins implicated in DNA uptake [60]. These attempts to isolate the receptor were promising in that they isolated proteins with the ability to bind DNA, however a lack of reproducibility has raised doubt to the validity of these results.

The general strategy I employed to isolate the receptor was UV laser crosslinking (Figure 1.4). The basis for crosslinking is that if DNA and proteins are in contact, a laser pulse of UV-light can cause covalent crosslinks to form between them. Further to this, if the DNA is radioactively labeled then the protein will also become labeled when crosslinked to the DNA. Such labeled proteins could be detected when samples are separated by polyacrylamide gel electrophoresis and exposed to film.
Introduction

Figure 1.4 General strategy for identification of the DNA binding receptor of *H.influenzae* using UV laser cross-linking.
Using crosslinking to isolate the receptor has three main benefits over the previously used techniques. The first and primary benefit is its simplicity. It does not require purification of membrane extracts or other complicated procedures. Second, it uses in vivo conditions. This is important because it eliminates the loss of protein-protein contacts which can occur in membrane preparations. Thirdly, it is a relatively rapid procedure that can produce results in a matter of days.

2.1 Background

Ultraviolet light is a 'zero-length' crosslinking agent which creates bonds between proteins and DNA at contact sites. Unlike chemical crosslinking there is no need for exogenous crosslinking agents that may disrupt the protein-nucleic acid complex [61, 62]. First used in the 1960s, UV irradiation was shown to cause the formation of protein-DNA crosslinks in bacteria [63, 64].

The original crosslinking experiments used broad band germicidal lamps to irradiate samples. This was a weak source of UV light, and so these experiments required irradiation times ranging from minutes to several hours [65]. Such prolonged times of irradiation created conditions for the redistribution of proteins. This problem has been addressed by using lasers as the source of UV light.

UV lasers confer a number of benefits. The laser allows the number of photons needed for crosslinking to be delivered in nano- or picoseconds. Since this reaction is several orders of magnitude faster than macromolecular rearrangements between protein and nucleic acid molecules (100 μs or greater), UV-induced crosslinking essentially “freezes” interactions between two molecules, allowing researchers to examine instantaneous DNA-protein interactions [66]. Another benefit to using a laser is that samples are irradiated with a beam of monochromatic light. Through the use of frequency modulators a wavelength can be used that maximizes the
number of crosslinks per pulse and minimizes the amount of protein degradation (see section 2.4).

2.2 Mechanism of crosslinking

Crosslinking occurs in two steps. The first step is the absorption of photons by a nucleic acid base, causing the base to change from its ground state to an excited, highly reactive state. The second step involves the conversion of the energy of the excited base into the protein-nucleic acid crosslink.

Initially, photons produced by the laser excite nucleotide bases into singlet (S_1) and triplet (T_1) states. This increases the possibility of absorption of a second photon, and transition to even higher excited states (T_n and S_n) [62]. These T_n and S_n states have energies of 8-9 eV which exceeds the ionization potential of the bases in solution and leads to the generation of purine and pyrimidine cation radicals [67]. Details of the crosslink formation step are unclear but are thought to involve the pyrimidine and purine cationic radicals which have the potential to crosslink to amino acids [62, 67].

2.3 Stability

UV induced crosslinks between proteins and nucleic acids are covalent [68]. Crosslinks generated by low intensity UV lamps are known to be resistant to both heat and alkali [69], but are completely broken down by treatment with either 1M acetic acid or 6M HCl for 15 min. at 25°C [70]. UV laser generated crosslinks are expected to have similar properties.

2.4 Efficiency

The efficiency of crosslink formation has been studied in vitro by many groups. Efficiency is highest between 245 and 280 nm, wavelengths where the UV light is
primarily absorbed by the nucleic acids (specifically the thymidine residues) [71]. Crosslinking can also be obtained using other wavelengths (200-240 nm), but in this range a high amount of protein degradation also occurs [61]. Depending on the conditions outlined below, irradiation by a UV-laser can cause 1 to 20% of a protein sample to become crosslinked to DNA [62, 71].

This range illustrates that many factors must be considered when using crosslinking to study protein-DNA interactions. For example, crosslinking efficiency is a function of the number of favorable contacts that occur between protein and nucleic acid [71]. Specifically, when the binding site of the protein is completely filled with nucleic acids, binding (and crosslinking) should be maximal. In addition to the strength of the DNA-protein interaction, the efficiency of crosslinking also depends on the wavelength of the exciting radiation, the nucleotide composition of the DNA, and the total number of photons applied to the sample [61].

2.5 Specificity

Crosslinking between DNA and proteins occurs through single nucleotide residues [61]. Crosslinking shows nucleotide preference, with thymidine being the most reactive. The nucleotide residues can be ranked in order of decreasing photoreactivity: dT >> dC > rU > rC, dA, dG [61]. It has been demonstrated that uracil can be crosslinked to 12 different amino acids and thymidine to five [72-74]. Cytosine is also able to be crosslinked to amino acids, however, single purine bases appear to be un-reactive [75]. Therefore, if crosslinking is to be successful in isolating the receptor, DNA that is rich in thymidine should be used as the 'bait'.
CHAPTER TWO
Materials and Methods

2.1 Strains, Plasmids and Oligonucleotides

Strains, plasmids and oligonucleotides (oligos) used in this study are listed in Tables 2.1, 2.2 and 2.3, respectively. All *H. influenzae* strains are descendants of Alexander and Leidy's original Rd strain [76]. Plasmid pGEM7- was obtained from Promega. All oligonucleotides used in this study were purchased from Alpha DNA.

2.2 Culture conditions

*H. influenzae* strains were cultured at 37°C in brain heart infusion (BHI; Difco) supplemented with hemin (10 µg/ml) and nicotinamide adenine dinucleotide (NAD; 2 µg/ml). Cultures were innoculated from either a single colony or a frozen 1 ml aliquot of an early exponential phase culture. *E. coli* strains were cultured in Luria-Bertani (LB) broth (Difco) or Terrific broth (Table 2.4) at 37°C [77].

When high levels of aeration were required for either bacterial species, cultures were grown in Erlenmeyer flasks (of at least 5X the culture volume) shaken at 200 rpm in a shaking water bath (Innova 3000, New Brunswick Scientific). If only gentle aeration was needed, cultures were grown in loosely-capped test tubes (18mm X 150 mm) and rolled (60 rpm) using a tissue culture roller (Lab-line) placed in a 37°C incubator.
Table 2.1 Bacterial strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>H. influenzae</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KW20</td>
<td>Wild-type</td>
<td>Alexander and Leidy [78]</td>
</tr>
<tr>
<td>MAP7</td>
<td><em>kan</em> <em>nal</em> <em>nov</em> <em>str</em> <em>spp</em> <em>rif</em> <em>vio</em></td>
<td>J. Setlow [79]</td>
</tr>
<tr>
<td>RR622</td>
<td><em>rec-2::MiniTn10 kan</em></td>
<td>Mini Tn10kan plasmid from D. McCarthy [80] integrated into KW20 chromosome by P. Williams.</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5a</td>
<td><em>supE44 recA1</em></td>
<td>D. Hanahan [81]</td>
</tr>
</tbody>
</table>

Table 2.2 Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Genotype</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGEM7-Zf-</td>
<td>pBR322 derivative (<em>amp</em>')</td>
<td>Promega</td>
</tr>
<tr>
<td>pGP1</td>
<td>pGEM7-Zf-::USS-1</td>
<td>This study</td>
</tr>
<tr>
<td>pGPR</td>
<td>pGEM7-Zf-::USS-R</td>
<td>This study</td>
</tr>
<tr>
<td>pGP50</td>
<td>pGEM7-Zf-::USS-50</td>
<td>This study</td>
</tr>
</tbody>
</table>
### Table 2.3 Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>USS-1-W</td>
<td>5'TAAAGTGCGGTAAATTITTAAGTATTTTT 3'</td>
</tr>
<tr>
<td>USS-1-C</td>
<td>3'TTTCAACGCCAATTAAAATTTCATAAAAA 5'</td>
</tr>
<tr>
<td>USS-R-W</td>
<td>5'TCTTGTTAGAATCTGAGTGTTATTTTT 3'</td>
</tr>
<tr>
<td>USS-R-C</td>
<td>3'AGAACAATCTTNGACTCACAATAAATTTA 5'</td>
</tr>
<tr>
<td>USS-30-W</td>
<td>5'GGTACCATATAAAGTGCGGTTAATTIT 3'</td>
</tr>
<tr>
<td>USS-30-C</td>
<td>3'CCATGGTATATTTCAACGCCAATTAAAAAATG 5'</td>
</tr>
<tr>
<td>USS-40-W</td>
<td>5'TGGTACCATAAAGTGCGGTTAATTITTTAC 3'</td>
</tr>
<tr>
<td>USS-40-C</td>
<td>3'ACCATGGTATATTTCAACGCCAATTAAAAAATG 5'</td>
</tr>
<tr>
<td>USS-50-W</td>
<td>5'TAAATGGTACCATATAAAGTGCGGTTAATTITTTAAGTATTTTTTAATTCCATAATTTG 3'</td>
</tr>
<tr>
<td>USS-50-C</td>
<td>3'ATTATCATGATATTACATGGTATATTTCAACGCCAATTAAAAAATG 5'</td>
</tr>
<tr>
<td>USS-50 LE-W</td>
<td>5'AAATGGTACCATATAAAGTGCGGTTAATTITTTAAGTATTTTTTAATTCCATAATTTG 3'</td>
</tr>
<tr>
<td>USS-50 LE-C</td>
<td>3'TTTCAACGCCAATTAAAAAATTTCCATAAAAACTTAAAGGATTACCATGGTATATTTCACGCCA 5'</td>
</tr>
<tr>
<td>USS-50 RI-W</td>
<td>5'TAAATTTTTAAGTATTTTTTTATATTTCCATAATTTTAAAGTATTTTTTAATTCCATAATTTG 3'</td>
</tr>
<tr>
<td>USS-50 RI-C</td>
<td>3'ATTAAAAATTTCCATAAAAACTTAAAGGATTACCATGGTATATTTCACGCCA 5'</td>
</tr>
<tr>
<td>USS-50 RC-W</td>
<td>5'TAAATGGTACCATATAAAGTGCGGTTAATTITTTAAGTATTTTTTAATTCCATAATTTG 3'</td>
</tr>
<tr>
<td>USS-50 RC-C</td>
<td>3'ATTATCATGATATTACATGGTATATTTCAACGCCAATTAAAAAATG 5'</td>
</tr>
<tr>
<td>USS-50 R-W</td>
<td>5'ATTATCATGATATTACATGGTATATTTCAACGCCAATTAAAAAATG 5'</td>
</tr>
<tr>
<td>USS-50 R-C</td>
<td>3'TAAATTTTTAAGTATTTTTTTATATTTCCATAATTTTAAAGTATTTTTTAATTCCATAATTTG 3'</td>
</tr>
<tr>
<td>D-USS-50-W-Biotin</td>
<td>5'TAAATGGTAAAGTGCGGTTAATTITTTAAGTATTTTTTAATTCCATAATTTG 3'</td>
</tr>
<tr>
<td>D-USS-50-C</td>
<td>3'ATTATCATTACATGATATTACATGGTATATTTCAACGCCAATTAAAAAATG 5'</td>
</tr>
</tbody>
</table>

All oligonucleotides listed were synthesized as single stranded molecules. Oligos were designed to allow annealing with complementary strands to form double stranded DNA molecules. The 9 bp core USS is indicated in bold. Restriction enzyme sites are shown above their recognition sequence.
Agar plates were prepared by the addition of 12 g/L Bacto-Agar (Difco) to liquid media prior to autoclaving. Additional hemin was applied to BHI plates older than 24 hours. For plating of *H. influenzae*, cells were serially diluted in ‘dilution solution’ containing 1X phosphate-buffered saline and 10% BHI, then plated on sBHI plates (BHI plates supplemented with hemin and NAD). When screening for transformants, cells were plated on sBHI containing the following concentrations of antibiotics: novobiocin, 2.5 μg/ml and chloramphenicol, 1 μg/ml. For LB plates the concentrations were: ampicillin, 100 μg/ml and chloramphenicol, 25 μg/ml.

### 2.3 Media

All media were sterilized by autoclaving. The ingredients of non-commercial media are described in Table 2.4.

<table>
<thead>
<tr>
<th>Table 2.4 Components of non-commercial media</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a) Terrific broth [77]</strong></td>
</tr>
<tr>
<td>Bacto-tryptone</td>
</tr>
<tr>
<td>Bacto-yeast extract</td>
</tr>
<tr>
<td>Glycerol</td>
</tr>
<tr>
<td>KH₂PO₄</td>
</tr>
<tr>
<td>K₂HPO₄</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Table 2.4 continued

b) MIV medium for competence induction [82].

<table>
<thead>
<tr>
<th>Solution</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution 21</td>
<td></td>
</tr>
<tr>
<td>Distilled water</td>
<td>850 ml</td>
</tr>
<tr>
<td>L-Aspartic acid</td>
<td>4.0 g</td>
</tr>
<tr>
<td>L-Glutamic acid</td>
<td>0.2 g</td>
</tr>
<tr>
<td>Furmaric Acid</td>
<td>1.0 g</td>
</tr>
<tr>
<td>NaCl</td>
<td>4.7 g</td>
</tr>
<tr>
<td>Tween 80</td>
<td>0.2 ml</td>
</tr>
<tr>
<td>K$_2$HPO$_4$</td>
<td>0.87 g</td>
</tr>
<tr>
<td>KH$_2$PO$_4$</td>
<td>0.67 g</td>
</tr>
<tr>
<td>Solution 22</td>
<td></td>
</tr>
<tr>
<td>L-Cystine</td>
<td>0.04 g</td>
</tr>
<tr>
<td>L-Tyrosine</td>
<td>0.1 g</td>
</tr>
<tr>
<td>L-Citruline</td>
<td>0.06 g</td>
</tr>
<tr>
<td>L-Phenylalanine</td>
<td>0.2 g</td>
</tr>
<tr>
<td>L-Serine</td>
<td>0.3 g</td>
</tr>
<tr>
<td>L-Alanine</td>
<td>0.2 g</td>
</tr>
<tr>
<td>Solution 23</td>
<td></td>
</tr>
<tr>
<td>CaCl$_2$</td>
<td>0.1 M solution</td>
</tr>
<tr>
<td>Solution 24</td>
<td></td>
</tr>
<tr>
<td>MgSO$_4$</td>
<td>0.1 M solution</td>
</tr>
<tr>
<td>Solution 40</td>
<td>5% (w/v) solution of vitamin-free casamino acids (Difco) in distilled water.</td>
</tr>
</tbody>
</table>

MIV is made by adding 1 ml of each of solutions 22, 23, 24 and 40 to 100 ml of solution 21.
2.4 Transformation of *H. influenzae*

2.4.1 Competence induction

Competence was induced in *H. influenzae* by transfer of cells to MIV starvation medium [82] as outlined in Barcak et al. [79]. Cells were grown overnight in sBHI to stationary phase. The following day cultures were diluted 1000 fold into sBHI and incubated at 37°C to permit several generations of growth during exponential phase. Cells were grown to a density of approximately $10^9$ cfu/ml ($OD_{600}$ of 0.2-0.25) and collected by filtration using a 100 ml Nalgene Analytical Test Filter Funnel (0.2 μm pore size). Cells were rinsed once with MIV [83] and transferred to a flask containing one volume of MIV (equal to the amount of culture that was originally filtered). Cells were shaken at 100 rpm for 100 minutes, by which time they achieved a maximal level of competence.

2.4.2 Transformation using linear DNA

DNA (200 ng) containing antibiotic resistance markers was incubated with 200 μl of MIV competent cells. Cells were rolled in a tissue culture roller at 37°C for 15 minutes to allow sufficient time for uptake of the DNA [29]. DNase I was added and the mixture was rolled for an additional 10 minutes. The transformation mixture was serially diluted in dilution solution and each dilution plated on medium containing the appropriate antibiotic.

2.4.3 Transformation using circular plasmid DNA.

Circular plasmid DNA transforms *H. influenzae* poorly. Plasmid transformations were carried out using a previously described method of enhancing plasmid transformation by treating MIV-competent cells with 32% glycerol [84].
2.5 *E. coli* plasmid transformation

*E. coli* cells were made competent by treatment with cold 100 mM CaCl₂, and otherwise transformed by standard procedures [85]. All clonings were carried out in *E. coli* strain DH5α.

2.6 Nucleic Acid Techniques

2.6.1 Isolation of plasmid DNA

*H. influenzae* cells were grown in sBHI and *E. coli* cells were grown in Terrific Broth overnight to stationary phase. Plasmid DNA was extracted using the alkaline lysis procedure [77]. If plasmids were to be sequenced, DNA was further purified by precipitation with LiCl and polyethylene glycol (PEG 8000; Sigma) [85].

2.6.2 Isolation of *H. influenzae* chromosomal DNA

Chromosomal DNA was isolated as previously described [79]. Bacterial cultures (35 ml) were grown overnight in sBHI, pelleted and resuspended in 0.15M NaCl, 0.1M ethylenediamine-tetraacetic acid (EDTA), pH 8.0. Cells were lysed by adding 1% sodium dodecyl sulfate (SDS) for 10 minutes at 52°C. Lysed cells were treated with proteinase K (50 µg/ml) for 1 hour at 37°C, followed by extraction with one volume of phenol-chloroform (1:1). DNA was precipitated by adding 2 volumes of 95% ethanol and collected by spooling on a glass rod. DNA was dried for 1 hour at room temperature and dissolved in 500 µl TE pH 8.0 [10 mM Tris-HCl, pH8; 1 mM EDTA]. Dissolved DNA was treated with RNase A (0.2 mg/ml, Sigma) at 37°C for 30 minutes. DNA was additionally purified by extraction with equal volumes of phenol and phenol-chloroform. Extracted DNA was precipitated with 2 volumes of 95% ethanol and 0.15M NaCl, dried, then resuspended in 10 ml of TE pH 8.0.
2.6.3 DNA labeling
Oligos and linearized plasmids were end labeled using γ\(^{33}\)P-ATP. DNA (4-40 μg) was incubated in polynucleotide kinase buffer (70 mM Tris-HCl, 10 mM MgCl\(_2\), 5 mM DTT; pH 7.6) with 5-25 μl γ\(^{33}\)P-ATP (10 mCi/ml) and 1 unit of T4 polynucleotide kinase in a 50 μl volume. The reaction proceeded for 30 minutes and was halted by transfer to a 65°C heat block for 20 minutes.

Nick translation was used to label chromosomal DNA [85]. Chromosomal DNA (10-20 μg) was incubated in *E. coli* DNA polymerase buffer (10 mM Tris-HCl; pH 7.5, 5 mM MgCl\(_2\), 7.5 mM dithiothreitol) with 5-20 μl α\(^{33}\)P-dATP (10 mCi/ml), 1 unit of *E. coli* DNA polymerase, 1-5 μl of DNase I (0.01 μg/ml) and 20 mM each of dTTP, dGTP and dCTP, in a volume of 200-250 μl. The reaction was incubated at 12-15°C for 30 minutes. The reaction was stopped with the addition of EDTA to 0.5 M and 100 μl TE pH 8.0. After incubation, all labeling reactions were extracted once with an equal volume of phenol-chloroform. Labeled DNAs were purified from unincorporated nucleotides by using either 10 ml chromatography columns (Bio-Rad) packed with Sephadex (G-15; Pharmacia) or ‘MicroSpin’ G-50 columns (Amersham).

The incorporation of label into the DNA was determined by placing labeled sample (0.5-10 μl) into a scintillation vial, adding 1 ml of scintillation fluid (ACS Scintillation cocktail; Amersham) and counting in a Beckman scintillation counter. The specific activity was calculated by dividing the number of counts per minute by the volume of the sample analyzed, then dividing by the concentration of DNA. The concentration of DNA was estimated by dividing the amount initially used in the labeling reaction by the volume of the labeling reaction. The specific activity of the DNA is expressed as the number of counts per minute per microgram of DNA.
2.6.4 DNA sequencing

Automated DNA sequencing of plasmids was carried out by the Nucleic Acid-Protein Service (NAPS) unit at UBC using AmpliTaq Dye Terminator Cycle Sequencing chemistry.

2.6.5 Oligonucleotides

All oligonucleotides were purchased as single strand DNA molecules. Concentrations of oligonucleotides were determined by their absorbance at 260 nm in a Beckman Du-65 spectrophotometer, and application of the formula \[ \text{DNA concentration} = A_{260} \times \text{dilution} \times 37 \mu g/ml \]. Oligos were made double stranded by adding equal amounts of each complementary oligo (40-100 µg, 4 µg/µl in dH₂O) in a microfuge tube and placing in a beaker containing boiling water for 10 minutes. Immediately after this incubation period tubes were spun for 5 seconds in a microfuge, to remove condensation from walls and caps of tubes, then replaced in the water bath.

To facilitate annealing of the single strands the water bath was removed from the hot plate and allowed to cool at room temperature to 50°C, at which time the oligos were expected to be double stranded. To confirm this a 1 µg aliquot of the boiled sample, as well as equal amounts of the single strand precursors, were run on a polyacrylamide gel and stained with ethidium bromide. Ethidium bromide intercalates between the stacked bases of nucleic acids [86]. Single stranded oligos were visible in lanes where they were added individually, but not in the boiled sample lanes. From this, the DNA was considered to be primarily double stranded.

2.6.6 Biotinylated oligonucleotides

D-USS-50-W-Biotin is a single stranded oligonucleotide with a biotin molecule attached to the 3’ terminus by a 15 atom spacer arm (CH₂-CH₂-NH-CO-CH₂-CH₂-CH₂-CH₂-CH₂-NH-CO-CH₂-CH₂-CH₂-CH₂-CH₂-). It is complementary to D-USS-50-C (see
Materials and Methods

Table 2.3). These oligos were made double stranded by the same procedure as in section 2.6.5 and were end labeled at the 5' ends as in section 2.6.3.

To join the biotinylated oligo to streptavidin agarose beads (Gibco BRL) a mixture of the two components was incubated in 1X phosphate-buffered saline and rolled in a tissue culture roller for 30 minutes. Beads were pelleted by centrifugation and resuspended in MIV (1 ml) after the supernatant liquid was removed. This was repeated four times in total to remove all unbound oligo from the beads. The number of molecules of DNA bound to the streptavidin-agarose beads was determined by scintillation counting, using the specific activity of radiolabelling and molecular weight of the oligo.

2.6.7 Cloning of Oligonucleotides (See Figure 2.1)
DNA ligations were carried out according to Sambrook et al. [77] using T4 DNA ligase (Boehringer Mannheim). PEG 8000 (15 % w/v) was added to blunt end ligations.

2.6.7.1 Cloning USS-1 and USS-R
Plasmid pGEM7- (1-3 μg) was digested with Smal, a blunt cutter which cuts once within the multiple cloning site in the lacZ gene. To minimize re-annealing, plasmids were dephosphorylated for 30 minutes at 37°C with 0.1 units of alkaline phosphatase in 50 μl of one-Phor-All PLUS buffer (10 mM Tris-acetate, 10 mM magnesium acetate and 50 mM potassium acetate). This reaction was stopped by heat inactivation of the enzyme at 85°C for 15 minutes. Water (150 μl) was added to the reaction and DNA was extracted with 200 μl of phenol-chloroform (1:1).
Materials and Methods

Figure 2.1 Cloning of USS-1, USS-R and USS-50.
A) To clone USS-1 and USS-R pGEM7- was digested with Smal then dephosphorylated. The oligos were phosphorylated then ligated into the Smal site of pGEM7- and scored for white phenotype.
B) To clone USS-50, both pGEM7- and USS-50 were digested with EcoRI and KpnI then ligated together.
Plasmids that contained inserts were identified by scoring for a white phenotype in the presence of X-GAL.
Oligos (0.5-1 μg) were phosphorylated in a 50 μl reaction volume of T4 polynucleotide kinase buffer using T4 polynucleotide kinase (10 units) in the presence of 1 mM ATP.

Phosphorylated oligos were combined with digested plasmids in T4 ligase buffer (66 mM Tris-HCl, pH 7.6, 6.6 mM MgCl₂, 10 mM DTT, 66 μM ATP) containing 1 unit of T4 DNA ligase (Boehringer Mannheim). Reactions were performed at room temperature for 16-18 hours. Ligations were transformed into *E. coli* and cells were plated on LB containing ampicillin. Colonies were scored for a white phenotype when grown in the presence of X-GAL. Plasmid DNA was isolated from white colonies, digested and run on 0.8 - 2 % agarose gels. After restriction mapping, inserts were sequenced using the T7 and SP6 primers [87].

2.6.7.2 Cloning of USS-50

Plasmid pGEM7- (1-3 μg) and USS-50 (5-15 μg) were each digested with *EcoRI* and *KpnI*. These enzymes cut once in each DNA molecule. After digestion, plasmid DNA and oligos were combined in T4 ligase buffer with 1 unit of T4 DNA ligase. Ligations were performed at 14°C for 16-24 hours. As above, ligations were transformed into *E. coli* and grown in the presence of X-GAL. As above, plasmid DNA was isolated from white colonies and extensive restriction digestions were performed to confirm the presence of an insert. After restriction mapping, inserts were sequenced using the same primers as above.

2.7 Electrophoresis

DNA fragments were separated on either agarose or polyacrylamide gels. Gels were stained with ethidium bromide (0.25 μg/ml) and separated DNA fragments were visualized under UV light. Either 100 bp or 1 kb ladders (New England Biolabs) were used as size standards.
2.7.1 Agarose gel electrophoresis
DNA fragments were separated on either 0.8 or 2 % agarose (Gibco BRL) Tris-Acetate EDTA (TAE, pH 8) gels [77]. DNA fragments were purified from agarose gels using a ‘Gene-Clean’ kit (Bio 101 Inç.).

2.7.2 Polyacrylamide Gel Electrophoresis (PAGE)
2.7.2.1 DNA
Small fragments of DNA were separated by PAGE using either 15 or 20 % polyacrylamide gels in 1X Tris Borate EDTA (TBE) and run at 100-150 V [77].

2.7.2.2 Crosslinked DNA-protein complexes
Samples were analyzed by SDS-PAGE using either 8 or 12% polyacrylamide with 5% stacking gels in a vertical minigel electrophoresis system (Owl Scientific) [77]. Following irradiation of DNA-protein mixtures, samples were mixed with 1/2 volume of 3X sample buffer (187.5 mM Tris-HCl (pH 6.8), 6% (w/v) SDS, 30% glycerol and 0.03% bromphenol blue). Samples were boiled for 10 minutes prior to loading. For unknown reasons, many crosslinked protein-DNA complexes do not enter polyacrylamide gels when solubilized in sample buffer that does not contain at least 1M urea [88]; therefore after boiling 1/5 volume of 5 M urea was added to samples prior to loading. Gels were electrophoresed (200-210 V) then dried using a gel dryer (BioRad) on Whatman DE 81 filter paper. Dried gels were exposed to a phosphoimager screen (Molecular Dynamics) for 24-96 hours.
2.8 Uptake experiments
Radioactively labeled DNA (1 μg) was incubated with 1 ml of freshly made MIV-competent cells. Cells and DNA were rolled at 37°C for 10 minutes. DNase I (50 μg) was added and the mixture was placed on ice. After 5 minutes 100 μl of NaCl (5M) was added and cells were pelleted by centrifugation at 13,000 rpm for 1 minute at 4°C (Canlab Biofuge A). The supernatant fluid was removed and cells were resuspended in cold MIV containing 1 M NaCl. Cells were again pelleted and the supernatant liquid removed. The pellet was resuspended in 200 μl of MIV at room temperature and transferred to a scintillation vial. Scintillation fluid (1 ml) was added and the radioactivity of the sample was counted using the scintillation counter.

2.9 Binding experiments
For some experiments it was necessary to determine the amount of DNA bound by cells. This was achieved using a binding assay. This procedure differs from uptake experiments in that a binding assay determines the amount of DNA that is able to be removed from the outside of cells, after non-specifically bound DNA is removed. Briefly, for binding experiments, cells were incubated with labeled DNA for 10 minutes, then washed with MIV to remove any DNA that was non-specifically associated with the outside of the cell. After the non-specific DNA was thought to be removed, cells were treated with DNase I to remove specifically bound DNA. Following DNase I treatment, cells were washed with a high salt solution to facilitate removal of DNA that remained bound to cells. Following these treatments the amount of label released into the supernatant fluid was counted.
2.10 UV laser crosslinking

Crosslinking experiments used a Quanta-Ray Model GCR14S pulsed Nd:YAG (neodymiumyttrium-aluminum-garnet) laser (Spectra Physics) (illustrated in Figure 2.2). Access to this laser was generously provided by Dr. M. Roberge (Dept. of Biochemistry and Molecular Biology, UBC). This laser emits photons with a wavelength of 1064 nm. It is equipped with an HG-2 harmonic generator (Spectra Physics) containing KD*P (potassium dideuterium phosphate) crystals, which reduces the wavelength of the emitted light from 1064 to 266 nm. Dichroic mirrors (DHS-2 Quanta-Ray dichroic harmonic separator) are used to eliminate residual 532 and 1064 nm light to a beam dump and to reflect monochromatic 266 nm light. This configuration allows the laser to emit 5-6 ns pulses with an energy of up to 60 mJ and a beam diameter of 6.4 mm. The energy was measured with an Astral AA30 power and energy meter equipped with an AC25 UV sensor (Scientech) [88].

Freshly made MIV competent *H. influenzae* cells were incubated with labeled DNA for 10 seconds, 1 minute or 30 minutes prior to crosslinking. The rationale for varying the time of incubation was to maximize the probability that the receptor was in contact with the labeled DNA when the pulses from the laser were applied. When crosslinking experiments were performed using biotinylated oligos, the labeled oligos were incubated with cells for 10 minutes prior to crosslinking.

Cells (50 μl, OD 0.2) were added to DNA (0.1 μg - 1 μg) in a microfuge tube and mixed with a pipette tip. Open 1.5 ml microfuge tubes were placed horizontally in the path of the laser. Tubes were stabilized by placing them in a 10-mm hole drilled
Figure 2.2 Schematic drawing of apparatus used for UV laser cross-linking. (Modified from Ho et al., (1994). This is a top view, not to scale.

The laser delivers pulses of 1064 nm. The harmonic generator reduces the wavelength of input light to 266 nm. The dichroic mirrors allow passage of only 266 nm monochromatic light to the sample. Other wavelengths (1064 nm and 532 nm) are directed to the beam dump. The sample is held horizontally by a plexiglass sheet attached to a micromanipulator.
in a small plexiglas sheet held in place by a Brinkmann micromanipulator. Crosslinking experiments were performed with three to six 50-mJ pulses.

For crosslinking experiments using streptavidin-agarose beads, samples were irradiated with 6 pulses of UV light from the laser. Following crosslinking, 1% SDS was added to samples then incubated at 37°C for 10 minutes. The supernatant liquid was removed and streptavidin-agarose beads were washed once with MIV. Samples were then treated with DNase (see section 2.8) for 5 minutes on ice. Sample buffer was added and the entire sample was loaded into the well. After electrophoresis gels were stained using a silver staining kit (BioRad silver stain plus).
CHAPTER THREE

Characterization of DNA sequences required for uptake by *H. influenzae*

Competent *H. influenzae* cells preferentially take up DNA which contains a specific uptake signal sequence (USS). In addition to the 9 bp core, other factors may also affect uptake. In this chapter I describe experiments that address how flanking sequences affect DNA uptake by competent *H. influenzae* cells.

3.1 Is the 29 bp USS sufficient for uptake?

USSs in the genome are flanked by regions of conserved sequence (see Figure 1.2). The 9 bp core and regions flanking the USS total 29 base pairs. Previous research treated DNA uptake in a qualitative manner [30]. However, this did not address the wide range (2 ng - 150 ng) of uptake which occurs depending on DNA length, the sequence of the 9 bp core and variations in the A/T richness flanking the USS. Since uptake occurs over such a range, it is more accurate to present it as a quantitative value rather than a qualitative one.

3.1.1 Uptake of a 'perfect' USS

Uptake can be quantified in two ways: µg of DNA/ml of competent cells or number of DNA molecules/cell. It is unclear whether the total amount of DNA taken up by cells, or the number of molecules taken up per cell, is the most biologically relevant way to present the data. Therefore, in this chapter data will be presented in both ways.

For clarity, in the following sections I will include small diagrams of the oligos used. Black boxes represent the 9 bp core, hatched regions represent the consensus base and white regions represent non-conserved sequences.
Is the 29 bp 'perfect' USS sufficient for uptake? An oligo (USS-1) was designed based on the information in Figure 1.2. It contains a 9 bp core USS and in the 3' flanking region it contains the consensus base (the base that occurs most frequently at that position following a USS in the genome) \([\text{USS-1}]\) (See Table 2.3). For consistency, I will refer to the orientation of the USS as presented in Figure 1.2. Sequences to the left of the 9 bp core will be termed 5' and sequences to the right will be referred to as 3'. In addition to USS-1 a negative control oligo (USS-R) was also designed. USS-R is the same length and base composition as USS-1 but does not contain a 9 bp core \([\text{USS-R}]\) (See Table 2.3). Since USS-1 contains the consensus base at each position in the core and flanking regions, I predicted that cells would bind and take it up preferentially over USS-R.

Uptake of USS-1 and USS-R was tested using nick translated \(H.\ influenzae\) and \(B.\ subtilis\) chromosomal DNA as positive and negative controls, respectively. The results of this experiment are shown in Figure 3.1.

Figure 3.1 (a) shows that USS-1 is not taken up better than negative control DNAs. Approximately 100 fold more labeled \(H.\ influenzae\) chromosomal DNA was taken up than all other DNAs. This correlates well with published results [15].

Figure 3.1 (b) represents uptake expressed as numbers of molecules per cell. It is clear that \(H.\ influenzae\) chromosomal DNA is taken up better than \(B.\ subtilis\) chromosomal DNA. From these data it appears that six fold more molecules of USS-1 and USS-R were taken up per cell than MAP7 DNA. This is because of the length of the chromosomal DNA, which was approximately 20 kb (estimated by electrophoresis). Since these chromosomal DNA fragments were much longer than the oligo, fewer molecules per cell were taken up. Like figure 3.1 (a) this figure shows that USS-1 is not taken up better than its negative control USS-R.
Figure 3.1. Uptake of ‘perfect’ USS oligo is not greater than an oligo that does not contain a USS.

1 ml of freshly made MIV competent cells were incubated with 1 μg of DNA in each case. Uptake procedures outlined in section 2.8 were followed. The mean value of duplicate trials are shown and error bars represent the standard deviation between the samples. Where error bars are not shown the error was too small to graph.
These results imply that uptake by *H. influenzae* requires more than the presence of a USS on the DNA molecule. Thus, the previous assumption that the 9 bp core USS is necessary and sufficient for uptake [30] appears to be incorrect.

Two possible explanations exist for the lack of recognition of USS-1. First, sequence 5’ to the core USS may be necessary for high levels of uptake and second, the 29 bp molecule may be too short to be taken up by competent cells. If the first explanation is correct, placing a USS in the middle of a 30 bp DNA fragment should increase uptake of the fragment over negative control levels. However, if the second explanation is correct, uptake should be equal for all 30 bp fragments regardless of the position of the USS. Experiments testing these two possibilities are presented in the following sections.

### 3.2 Effect of USS position on binding and uptake

Does the position of the USS affect recognition and uptake of DNA? In this section I will address this question and test the hypothesis that high levels of uptake require sequences 5’ to the 9 bp core USS.

In most of the following experiments in this chapter, both binding and uptake results will be presented, because in determining how cells interact with a DNA molecule, both the amount of DNA bound and the portion of the bound DNA that is taken up are important.

#### 3.2.1 Oligonucleotides

**3.2.1.1 Binding and uptake of 30 bp DNA fragments**

In order to address the possibility that sequences 5’ to the core USS are necessary for uptake, a 30 bp oligo was designed. USS-30 is 30 bp in length, contains a 9 bp core and in part of the 3’ flanking region it contains the USS consensus base
Uptake Results and Discussion

[USS-30] (See Table 2.3). Unlike USS-1, USS-30 has 11 bp 5' to the 9 bp core and contains only 10 bp of the 3' flanking consensus (See Table 2.3).

Figure 3.2 shows the results of binding and uptake experiments performed using USS-30, USS-1 and USS-R. These data show that, compared to USS-1 or USS-R, binding and uptake are approximately three fold and two fold greater for USS-30. This implies that nucleotides 5' to the USS are necessary for uptake.

3.2.1.2 Binding and uptake of 50 bp fragments

The result obtained with USS-30 was further examined using longer oligonucleotides. USS-50 is a 50 bp fragment with 15 bp 5' to the 9 bp core and 26 bp 3'. [USS-50] (See Table 2.3). Like USS-1, USS-50 contains the consensus base in the 3' flanking region. For ease of cloning it also contains Kpnl and EcoRI restriction sites at its ends.

I also designed a 50 bp oligo that consists of the 9 bp core starting one bp from the 5' end [USS-50 L] and the consensus base at each position in the 3' flanking region (See Table 2.3). A negative control was also designed, USS-50-R [USS-50-R] (See Table 2.3). Like USS-R, this oligo does not contain a USS or flanking consensus but has the same base composition as USS-50.

As in the preceding section, both binding and uptake were tested. Figure 3.3 illustrates the results of these experiments. Figure 3.3 (a) shows that binding decreases 1.5 fold when the USS is located at the 5' terminus and 12 fold when the 9 bp core is absent. Unlike the 30 bp DNA fragments, positioning of the USS at the 5' terminus did not decrease binding to negative control levels. Cells bound 8.3 fold
**Figure 3.2.** Binding and uptake of 29 and 30 bp oligonucleotides by competent cells.
The mean value of duplicate trials are shown and error bars represent the standard deviation between the samples. Where error bars are not shown the error was too small to graph.
Figure 3.3. Binding and uptake of 50 bp oligonucleotides by competent *H. influenzae* cells.
The mean value of duplicate trials are shown and error bars represent the standard deviation between the samples. Where error bars are not shown the error was too small to graph.
more of USS-50-Le than of the negative control USS-50-R. Uptake experiments in part (b) show that uptake of USS-50 and USS-50-Le were 3.2 and 2.2 fold greater than the negative control levels, respectively.

These results, like the results using the 30 bp DNA fragments, show that the position of the USS within a fragment affects uptake. In order for maximal uptake to be achieved, DNA must have sequence 5' to the USS (the minimum length of 5' sequence needed for optimal uptake was not tested). However, unlike the 30 bp oligos, if the 5' sequence is absent in 50 bp molecules, uptake and binding are decreased but not abolished.

3.2.2 Uptake of plasmid DNA modified by insertion of USS sequences
The results presented above are based on changes of the position the USS within short DNA fragments. To determine if these results are applicable to longer fragments, alteration of USS placement within a 3 kb fragment was tested. USS-1 was cloned into the multiple cloning site of pGEM7 and the plasmid was linearized separately with three restriction enzymes *SacI*, *KpnI* and *ClaI*, such that the USS and flanking consensus were localized to the middle, 5' and 3' ends respectively. Cleavage of the plasmid with *KpnI* and *SacI* placed the USS within four bases of the 5' end and 30 bp of the 3' end, respectively, whereas all linearized plasmids contained the entire 3' flanking consensus region (Figure 3.4 a). The results of these uptake experiments are shown in Figure 3.4 (b).

Placement of the USS near the 3' end or the middle of the linear plasmid after cleavage with *ClaI* or *SacI* did not significantly affect uptake. Uptake of the linear plasmid containing the USS on the 5' end was less than other fragments. This is most likely due to the proximity of the USS to the 5' end. Demonstrated with 30 and
Figure 3.4. Uptake of cloned USS-1.
a) Schematic diagram of cloned USS-1. 9 bp core USS is shown in bold. Extended consensus sequence is underlined. Cut sites are shown with an arrow. Scal cut site is approximated.
b) USS-1 was cloned into pGEM7- digested with the indicated enzymes and end-labeled. One μg of 3 kb linear DNA was incubated with cells as outlined in section 2.8. The mean value of duplicate trials are shown and error bars represent the standard deviation between the samples. Where error bars are not shown the error was too small to graph.
50 bp fragments, nucleotides 5' to the USS are necessary for efficient uptake. From these results I concluded that the necessity of 5' sequence for maximal uptake of USS containing DNAs applies to long DNA fragments as well as short oligos.

3.3 Length dependence of uptake: Is USS-1 too short to be taken up?

I tested the possibility that the low uptake of USS-1 was because a 29 bp fragment was too short to be recognized and taken up by cells. If this explanation were correct a 29-30 bp fragment should not be taken up by cells regardless of the position of the USS. As demonstrated in section 3.2.1.1 this is not the case. Thus, if 29-30 bp is long enough to be taken up by cells, then uptake should be similar for 30, 40 and 50 bp fragments containing USSs in the middle. This prediction was tested by use of a 40 bp oligo, designed to contain 11 bases 5' to the USS core and the consensus bases in the 3' flanking region [USS-40].

Figure 3.5 shows the results of uptake experiments. Figure 3.5 (a) shows that the amount of DNA taken up per ml of cells increases as the length of the DNA fragments increase. This resembles the finding that more chromosomal DNA is taken up than short oligos on a mass basis. However, when these data are presented as numbers of molecules per cell it appears that uptake is equal for each oligo. This indicates that USS-30 is long enough to be taken up by cells. Thus the likely reason for low uptake of USS-1 is due to the absence of sequences 5' to the USS.

From Figure 3.5 (b) it appears that 30 molecules of USS-R are taken up per cell. This is likely a case of non-specific interaction of USS-R with the cells. A lipopolysaccharide layer covers gram-negative bacteria [89, 90], and so I speculate that this apparent uptake may be a consequence of DNA that has become
Figure 3.5. Uptake of different length oligos.
The mean value of duplicate trials are shown and error bars represent the standard deviation between the samples. Where error bars are not shown the error was too small to graph.
Uptake Results and Discussion

intertwined with the lipopolysaccharide layer making it inaccessible to DNase I. If this is the case, then DNA would appear to be taken up by cells when in fact it is simply bound to the lipopolysaccharide region of the outer membrane. The explanation for this apparent uptake was not directly tested although, section 3.5.1 addresses this indirectly.

3.4 Effect of 3’ consensus flanking sequence on DNA binding and uptake

This section outlines experiments that tested how changing both the length and base composition of the 3’ flanking region affects uptake.

3.4.1 Binding and uptake of 50 bp oligos containing the non-consensus bases in the 3’ flanking region.

Uptake of a DNA fragment increases when the USS is flanked by an A/T rich sequence [30, 91]. These experiments, performed before the *H. influenzae* genome was sequenced, tested naturally occurring and artificially designed sequences. To test the effect of base composition on uptake, a 50 bp oligo was synthesized (USS-50-RC). This oligo contains a 9 bp core USS. However, this oligo instead of containing the consensus bases in the flanking region has the least common base at each position (shaded region) (See Table 2.3). For direct comparison of the effect of the 3’ sequence, the sequence 5’ to the USS is identical to USS-50. The results of uptake and binding experiments are illustrated in Figure 3.6.

Figure 3.6 (a) shows that binding of USS-50 was 2 fold greater than for USS-50-RC and 2.8 fold greater than USS-50-R. Binding of USS-50-RC was 1.4 fold greater than USS-50-R. Uptake of USS-50 was 2.3 fold greater than uptake of USS-50-RC, whereas USS-50-RC was taken up 2.3 fold better than USS-50-R. From these results
Figure 3.6. Binding and uptake of 50 bp oligos.
USS-50 contains the USS in the middle of the DNA fragment. USS-50-RC contains the USS in the middle of the fragment but contains the least common nucleotide at each position in the 3' flanking consensus region. USS-50-R does not contain a USS. The mean value of duplicate trials are shown and error bars represent the standard deviation between the samples. Where error bars are not shown the error was too small to graph.
I conclude that cells are able to take up DNAs containing the least common base in the 3' flanking consensus region, although at a much lower level than DNA containing an A/T rich flanking region. These data agree well with previous results of analysis of the effect of G/C flanking richness on uptake [30, 91].

One reason put forward to explain the greater uptake of DNA containing A/T rich flanking sequences was that the receptor makes contact with the minor groove of the incoming DNA [30]. In the minor groove, A/T base pairs differ from G/C pairs by the absence of the centrally located 2-amino group of guanine [92]. It is thought that after the receptor contacts the DNA it may induce partial melting in the region of the 9 bp core, a process facilitated by the lower melting point of A+T-rich DNA [30].

3.4.2 Binding and uptake of DNA lacking sequences 3' to the 9 bp core

The previous section illustrated that base composition of the 3' flanking region affects uptake. To further explore this, I examined how the absence of 3' sequence affects uptake. A 50 bp oligo (USS-50-Ri) was designed that contained a 9 bp core at the 3' terminus. To maintain fragment length and the same base composition as USS-50, sequences 3' of the core in USS-50 were placed 5' to the core of USS-50-Ri [USS-50-Ri]. It was predicted that removal of the 3' flanking consensus sequence would decrease the cell's ability to bind and take up the oligo. The results of binding and uptake experiments are presented in Figure 3.7.

These results show that DNA lacking 3' sequence is not bound or taken up better than a negative control. Binding of USS-50-Ri was only slightly better (1.4 fold) than USS-50-R. Uptake of USS-50 was 6.8 fold greater than USS-50-Ri. USS-50-R was
Figure 3.7. Binding and uptake of 50 bp oligos.
USS-50 contains the USS in the middle of the DNA fragment. USS-50-Ri contains the USS at the 3' terminus. USS-50-R does not contain a USS.
The mean value of duplicate trials are shown and error bars represent the standard deviation between the samples. Where error bars are not shown the error was too small to graph.
taken up 1.3 fold better than USS-50-Ri. These results indicate that the absence of 3' flanking sequence decreases binding to nearly the negative control levels and completely abolishes uptake. I conclude that sequences 3' to the core USS are absolutely required for uptake.

3.5 Competition of various DNAs for uptake of chromosomal DNA

As a different kind of test of the ability of various DNAs to be bound and taken up by competent *H. influenzae* cells, competition tests were performed. This assay was used to test whether or not the binding and uptake results described in the preceding sections were representative of actual binding and uptake by cells.

The general basis of a competition assay is that a labeled DNA fragment, specifically taken up by cells, is incubated with cells in the presence or absence of unlabeled competing DNAs (See Figure 3.8). We can consider two cases. First, if unlabeled DNAs are taken up by cells and there are a limited number of receptors per cell, the unlabeled DNA will compete with the labeled DNA for access to receptors and decrease the amount of labeled DNA taken up (Figure 3.8 b). In the second case, if the unlabeled DNA is not taken up by cells, competition for access to the receptor will not occur and labeled DNA will be taken up at a high level (Figure 3.8 a and c). Using this model, I expect that DNAs taken up by cells in the previous sections should compete for uptake of labeled MAP7 and those that were not taken up should not.
Figure 3.8 Schematic representation of a competition assay.

a) Cells are incubated with labeled DNA containing a USS. Cells take-up 4 molecules of labeled DNA.

b) Cells are incubated with labeled DNA containing a USS and unlabeled DNA that also contains a USS. As a result of competition for access to the receptor, cells only take up 2 molecules of labeled DNA.

c) Cells are incubated with labeled DNA containing a USS and unlabeled DNA that does not contain a USS. Since the competing DNA does not contain a USS cells are able to take-up 4 molecules of labeled DNA.
3.5.1 Competition between chromosomal DNA and 29 bp fragments
Labeled MAP7 chromosomal DNA (1μg, approximately 20 kb) and unlabeled competing DNAs (MAP7, *E. coli* DH5α, USS-1 and USS-R) were mixed in tubes prior to the addition of competent cells. Cells were incubated with the DNAs for 10 minutes and the general uptake procedure was followed.

Figure 3.9 shows the results of this experiment. Unlabeled MAP7 DNA competes for uptake of labeled MAP7 DNA. This is expected because MAP7 DNA is specifically bound and taken up by competent *H. influenzae* cells. DH5α chromosomal DNA and USS-R do not interfere with uptake of MAP7 DNA. This is also expected since these DNAs are not taken up by cells.

USS-1 competes weakly with uptake of MAP7 chromosomal DNA better than USS-R. This is surprising since figure 3.2 shows that USS-1 is not bound or taken up better than USS-R. This competition might arise if USS-1 is bound weakly by the receptor. It was theorized by Deich and Smith (1980) that DNA uptake occurs in three stages [29]. The first is weak/reversible binding of DNA by receptors. The second step is strong/irreversible binding in which the DNA is committed to uptake. The third step is the conversion of DNA into a DNase resistant, non-elutable form. Thus, if USS-1 is able to be bound weakly by the receptor it would interfere with uptake of labeled MAP7 DNA.

3.5.2 Competition between chromosomal DNA fragments and 50 bp fragments
Using the same assay as above, the ability of 50 bp fragments to compete for uptake of 20 kb chromosomal DNA was studied. Figure 3.10 shows that USS-50 and USS-50-Le compete for uptake of labeled chromosomal DNA, however, USS-50-RC, USS-50-Ri and USS-50-R do not.
Figure 3.9. Competition for uptake of labeled chromosomal DNA.

Competing DNAs were mixed with 1 µg of labeled MAP7 DNA prior to the addition of cells. Uptake of the labeled DNA was measured by the standard assay. The mean value of duplicate trials are shown and error bars represent the standard deviation between the samples. Where error bars are not shown the error was too small to graph.
Figure 3.10. Competition for uptake of labeled chromosomal DNA.

Competing DNAs were mixed with 1 μg of labeled MAP7 DNA prior to the addition of cells. Uptake of the labeled DNA was measured by the standard assay. The mean value of duplicate trials are shown and error bars represent the standard deviation between the samples. Where error bars are not shown the error was too small to graph.
This information correlates well with the information in Figure 3.9 which shows that sequence 5' to the USS is not needed for competing DNAs to interfere with the uptake of chromosomal DNA. However, sequence 3' to the 9 bp core is needed, and must be A/T rich if a DNA molecule is to compete. Variations in the counts due to incomplete removal of labeled MAP7 DNA from cells could account for the apparent increase in uptake of labeled MAP7 chromosomal DNA in the presence of USS-50-R and USS-50-RC.

3.6 Discussion of uptake results.

Experimental investigations of DNA binding and uptake, prior to completion of the \textit{H. influenzae} genome sequence, indicated that binding is saturable, reversible and specific. This suggested that a receptor protein, or complex of proteins, are responsible for the observed sequence specific uptake [41]. In this thesis I have attempted to expand on previous results, guided by information gained from the sequencing of the \textit{H. influenzae} genome.

Throughout this thesis I have addressed a number of questions in an attempt to assemble a model of preferential uptake of DNA in \textit{H. influenzae}. The general focus of these research questions address the importance of flanking sequence on uptake, in an attempt to determine how proteins that make up the receptor interact with the DNA molecule. USSs within the genome are flanked by regions of conserved sequence. With this information I designed oligos with variations in these flanking regions. Oligos were varied in both length and composition of 3' and 5' flanking sequence. I have been able to show that nucleotides 5' to the USS are required for high levels of binding and uptake of DNA however, the length and base composition of the 5' region was not tested. Also, I have shown that both length and base composition of the 3' flanking region greatly affect binding and uptake. If
sequence 3’ to the USS is G/C rich, uptake proceeds at a very low level. However, if DNA lacks a 3’ sequence, both binding and uptake are abolished.

3.7 Further research
Many questions concerning the effect of flanking DNA on binding and uptake remain unanswered. I will briefly cover two points that may be addressed in the future. First, oligos containing a single bp 5’ to the USS, and plasmids that contain 4 bp 5’ to the USS, are not taken up at wildtype levels by cells. One question to address could be: What is the minimum length of the 5’ sequence needed for high levels of uptake to occur? Second, one can ask: How does a G/C rich 5’ flanking sequence affect uptake, if the 3’ flanking consensus is A/T rich? If the DNA is denatured 5’ to the USS then increasing the G/C richness will decrease uptake, as it does in the 3’ flanking region. However, if denaturation does not occur in the 5’ region uptake should not be affected.
CHAPTER FOUR

UV Laser Crosslinking

Crosslinking labeled DNA to the receptor.

UV laser crosslinking is a powerful technique to study DNA-protein interactions. The design of my crosslinking experiments was simple. Labeled DNA was incubated with competent cells. Samples were irradiated with three to six pulses from the laser, then boiled in sample buffer. Proteins were separated by polyacrylamide gel electrophoresis, stained with either coomassie blue or silver and exposed to a phosphoimager screen. Variables adjusted were the time of incubation with DNA and also the presence or absence of unlabeled competing DNAs. This chapter outlines the results attained in attempting to crosslink DNA to the *H. influenzae* receptor protein.

4.1 Time interval for uptake of USS-50

For UV light to crosslink the receptor to DNA, the pulses must be delivered when the two molecules are in contact. To approximate when the DNA contacted the receptor, I tested when uptake begins and for how long uptake proceeds. Previous experiments have demonstrated that uptake of chromosomal DNA is complete within 5 minutes [15, 29]. However, this had not been tested with very short DNA molecules.

Figure 4.1 shows that uptake of USS-50 begins soon after incubation with cells. The DNase pre-treatment point corresponds to DNA being treated with DNase for 10 minutes prior to incubation with cells. Little DNA is taken up in the first minute and
Figure 4.1. Uptake of labeled USS-50 as a function of time. Cells were incubated with 1μg of labeled USS-50 for various time intervals, preceded or followed by treatment with DNase I as outlined in section 2.8.
uptake is complete after 10 minutes. From this data I concluded that irradiation of samples should take place before 10 minutes, prior to completion of uptake.

4.2 Does crosslinking increase the amount of DNA associated with cells?
To isolate the receptor, laser crosslinking must create crosslinks between labeled DNA and outer membrane proteins. To test if this was possible, DNA was incubated with cells with and without irradiation from the laser, and the amount of DNA bound to the outside of cells was studied. Cells were first incubated with DNA. At specified times (10 sec, 1 min and 30 min) laser pulses were delivered to individual samples. After incubation, cells were washed and the amount of DNA able to be removed by DNase treatment was tested. I predicted that if UV light caused the formation of crosslinks, then the amount of DNA bound to the outside of cells should be increased by irradiation. The results are shown in Figure 4.2.

The amount of DNA bound to the surface of cells (i.e. accessible to DNase I) was increased by irradiation after 10 seconds and 1 minute of incubation. However, irradiation after 30 minutes of incubation did not increase the amount of DNA bound, compared to the non-irradiated control. These results indicate that UV irradiation increases the amount of DNA bound to the outside of cells at short but not long incubation times.

The increase in the amount of DNA bound to cells at short incubation times is most likely a result of DNA being in contact with the receptor at those time points. As is shown in Figure 4.1, after 30 minutes cells have completed uptake of DNA, and so when these cells were irradiated there was little if any DNA bound to the outside of cells.
Figure 4.2. Crosslinking increases the amount of DNA associated with the outside of cells. Cells (50 µl) were incubated with 0.05 µg of labeled USS-50, then irradiated with a laser pulse at the indicated times. DNA was removed from the outside of cells by treatment with DNase I and high salt washes.
4.3 Crosslinking using labeled USS-50

Initial crosslinking experiments used end labeled USS-50 as the 'bait'. A short oligo was used, rather than labeled chromosomal DNA, in an attempt to minimize the amount of non-specific crosslinking (see Figure 4.3).

Competing DNA was used in the following experiments to differentiate between sequence-specific and nonspecific crosslinking. If cells are incubated with equal amounts of labeled and unlabeled DNA that are specifically recognized, the unlabeled DNA should compete with the labeled DNA for access to the receptor. This competition would lead to a decrease in the amount of label associated with proteins that bind DNA in a sequence specific manner. However, if unlabeled competing DNA is not recognized there will be no change in the amount of label associated with the receptor. Further to this, when a competing DNA is used that is not recognized by cells, less label should become associated with proteins that bind DNA independent of its sequence. Figure 4.4 shows the results of crosslinking experiments performed using USS-50. Two competing DNAs were used in these experiments, one that would compete with labeled USS-50 for access to the receptor (unlabeled USS-50) and another that would not (unlabeled USS-R). To minimize the amount of background, a serial dilution of labeled USS-50 was used in crosslinking experiments.

The migration of DNA in a polyacrylamide gel changes when it is bound to a protein [93, 94], so I expected that these short DNA fragments should migrate more slowly when crosslinked to a protein. However, the non-crosslinked DNA in lane 9 migrates no faster than the crosslinked bands in lanes 1-8. From this I concluded that crosslinking had not occurred and the bands seen in lanes 1-8 were most likely un-reacted labeled DNA. No other bands were visible in the image produced by the phosphoimager screen.
Figure 4.3. Representation of hypothetical crosslinking using short and long DNA fragments.

1. Long nick translated chromosomal DNA used in crosslinking experiments has a high probability of coming into contact with and becoming crosslinked to non-receptor membrane proteins.
2. Short oligonucleotides in crosslinking experiments are less likely to come into contact with non-receptor membrane proteins.
Figure 4.4. Crosslinking USS-50 to wild-type *H. influenzae* cells.

Phosphorimager detection of $^{33}$P radioactivity associated with bands in an SDS-PAGE gel. Numbers represent the amount of DNA incubated with cells prior to crosslinking. Dashes indicate no DNA added. Labeled USS-50 (lanes 1-4) was crosslinked to competent KW20 cells (50 µl) in the presence of competing unlabeled DNA (USS-50, lanes 5 and 7; USS-50-R, lanes 6 and 8). Crosslinking was performed after 1 minute of incubation of cells with DNA. Lane 9 contains labeled USS-50 loaded directly onto the gel.
4.4 Crosslinking following varied times of incubation

Initial attempts did not produce enough crosslinking to allow visualization of labeled proteins, perhaps because the receptor and DNA were not in contact at the time of irradiation. This could be because the DNA had not yet contacted the receptor or because it had already been internalized when the laser pulses were delivered. To control for this, the time of incubation was varied. DNAs were incubated with cells for either 10 seconds, 1 minute or 30 minutes prior to crosslinking. Different DNAs were also used in case some might provide a better substrate for crosslinking than others. USS-50, USS-1 and MAP7 chromosomal DNA were used in the following crosslinking experiments. The results are shown in Figure 4.5.

Crosslinking using *H. influenzae* chromosomal DNA (MAP7) showed a high molecular weight streak in the 10 second, 1 minute and 30 minute lanes (lanes 2, 5 and 8). Since this band was also visible in the non-crosslinked lane I concluded it was likely due to either incorporation into the chromosome or simply re-isolation of input DNA. It has been demonstrated that integration of DNA into the chromosome can take place within 10 minutes of incubation with cells [41]. Though irradiated samples were placed on ice after crosslinking, cells were stored for up to 30 minutes prior to boiling in SDS. This may have allowed time for the DNA to integrate into the chromosome. Variation in the intensity of bands could have arisen due to differences in uptake by cells.

Crosslinking using USS-1 (lanes 3, 6, and 9) did not produce any bands. This is most likely due to low uptake of the USS-1 oligo, outlined in section 3.1.1.

Use of USS-50 as the 'bait' in crosslinking experiments gave a high molecular weight band present in lanes 1 and 4 that is not present in the non-crosslinked lane (lane 10).
Figure 4.5. Crosslinking experiments using USS-50, USS-1 and MAP7 DNAs as bait.

Phosphorimager detection of $^{33}$P radioactivity associated with bands in an SDS-PAGE gel. The upper and lower panels represent high and low molecular weight DNA from the same polyacrylamide gel, respectively. 0.1 μg of labeled DNAs were incubated with 50 μl of competent cells then crosslinked after the times listed above. Lanes 1 - 3 represent crosslinking experiments performed after 10 seconds of incubation using. Lanes 4 - 6 are crosslinking experiments performed after 1 minute using the same DNAs as lanes 1-3. Lanes 7 - 9 are crosslinking experiments performed after 30 minutes of incubation using the same DNAs as lanes 1-3. Lanes 10 - 12 are control lanes, which are DNAs incubated with cells but not crosslinked.
The band is most prominent at 10 seconds, decreasing in intensity at 1 minute and is not visible after 30 minutes (lane 7). Possible explanations for these bands include, crosslinking to proteins, incorporation into the chromosome or re-isolation of input DNA. To test which explanation was correct, crosslinking experiments were performed in a rec-2 background.

4.5 Crosslinking in a rec-2 mutant background

As outlined in the Introduction, the Rec-2 protein is required for movement of DNA across the inner membrane of H. influenzae. Mutations in Rec-2 lead to DNA being localized in the periplasmic space, unable to move into the cytoplasm [39, 41]. Crosslinking experiments were performed in a rec-2 mutant strain to prevent incorporation of labeled DNA into the chromosome. It was expected that high molecular weight bands seen in Figure 4.5 should disappear if they were due to incorporation into the chromosome, but remain visible if they were a result of crosslinking.

4.5.1 Crosslinking in wild-type and rec-2 backgrounds.

MAP7 chromosomal DNA, USS-50 and USS-1 were incubated with freshly made MIV competent wild-type or rec-2 cells, and irradiated with 3 pulses from the laser. Results are shown in Figure 4.6 and 4.7 for chromosomal DNA and oligos respectively.

In Figure 4.6, lanes 1 - 4 (wildtype) and 5 - 8 (rec-2) show similar banding patterns. There are bands in non-crosslinked lanes (4 and 8), which indicates that these bands are not due to crosslinking. Since the bands are visible in the rec-2 strain it is likely that these bands are not due to incorporation into the chromosome but arise as a result of isolation of input DNA. The pattern in lanes 1 - 8 resemble that in lane 10,
Figure 4.6. Crosslinking using MAP7 chromosomal DNA.  
Phosphorimager detection of $^{33}$P radioactivity associated with bands in an SDS-PAGE gel.  
0.1 μg of labeled 20 kb chromosomal DNA was incubated with 50 μl of competent cells then crosslinked after the times listed below.  
Lanes 1-4 are the result of crosslinking in a wild-type background.  
Incubation times were: lane 1, 10 seconds; lane 2, 1 minute; lane 3, 30 minutes; lane 4 non-crosslinked.  
Lanes 5-8 are the result of crosslinking in a rec-2 background.  
Incubation times were as above.  
The DNA in lane 9 was treated with DNase I prior to crosslinking.  
Lane 10 has MAP7 DNA loaded directly.
which contained MAP7 DNA loaded directly onto the gel. The decreased intensity likely reflects the small fraction of DNA taken up by cells under saturating conditions. The DNA in Lane 9 was digested with DNase for 10 minutes prior to crosslinking. This served as a control since digested DNA should not be taken up by cells. I concluded that the high molecular weight bands seen in figures 4.5 and 4.6 are the result of re-isolation of the DNA taken up by cells from the periplasm, rather than crosslinking or chromosomal incorporation.

This experiment was repeated using oligos instead of chromosomal DNA (Figure 4.7). Crosslinking experiments in a rec-2 mutant background are shown in lanes 1 - 3. Lane 4 shows the non-crosslinked control. High molecular weight bands were not visible, probably because USS-50 cannot cross the inner membrane and integrate into the chromosome. Low intensity bands, which have the same mobility as that in lane 6, indicate the presence of USS-50 in these lanes. These low molecular weight bands are likely a result of re-isolation of USS-50. Using USS-1 failed to yield any useful information. The faintness of the USS-1 band when incubated with cells (lane 8) likely reflects low uptake and subsequent re-isolation from the periplasm.

4.6 Biotinylated oligonucleotides

4.6.1 Uptake experiments

Lack of crosslinking indicated that the previous approach would not label enough protein to identify the receptor. I hypothesized that the failure to crosslink sufficient receptor was because once bound, DNA is rapidly taken up and no longer available for crosslinking. Cells start and complete uptake at slightly different times. Thus, only a small percentage of receptors will be in contact with DNA at any one time. In the previously described experiments some cells might not yet have contacted a DNA molecule and others would have completed uptake at the time of irradiation.
Crosslinking Results and Discussion

Figure 4.7. Crosslinking in rec-2 and wild-type backgrounds. Phosphorimager detection of $^{33}\text{P}$ radioactivity associated with bands in an SDS-PAGE gel. Incubation times for lanes 1-4 are 10 sec, 1 minute, 30 minutes and non-crosslinked respectively. Lanes 6 and 10 contain 0.1 μg of USS-50 and USS-1 respectively, loaded directly onto the gel.
Crosslinking Results and Discussion

If this were true, only a small percentage of receptors would have been available for crosslinking when the laser delivered its pulses. It is not possible to simply irradiate samples for extended periods with many pulses from the laser due to the significant amount of protein degradation that occurs when greater than six pulses are delivered to a sample (M. Roberge personal communication). Therefore, I designed an experiment that should allow cells to partially take up DNA then stop midway through uptake. This would essentially ‘freeze’ uptake at a point where DNA was in contact with the receptor, increasing the amount of DNA crosslinked by the laser pulse. An oligo with a biotin molecule attached to its 3’ terminus was designed to prevent complete uptake of the DNA molecule (Figure 4.8 and Table 2.3).

The DNA is spaced from the biotin by a 15 atom spacer arm (CH₂-CH₂-NH-CO-CH₂-CH₂-CH₂-CH₂-NH-CO-CH₂-CH₂-CH₂-CH₂-). This spacer was attached between the DNA and the biotin to decrease interference by the DNA in biotin-streptavidin interactions.

Uptake experiments were performed using biotinylated DNA. It was found that biotinylation did not prevent uptake of the DNA molecule. However, uptake experiments simply determine the amount of label associated with cells. Therefore if the biotinylated DNA is taken up until the spacer arm is reached it might appear as if the molecule is completely taken up by cells. Thus I tested whether biotin was accessible to streptavidin-agarose beads after cells were incubated with DNA for 10 minutes. Streptavidin-agarose consists of agarose beads covered with streptavidin molecules, which bind tightly to biotin. The results from this experiment showed that biotin was not accessible to streptavidin after incubation with cells (21 cpm associated with streptavidin-agarose after addition). From this, I concluded that the
Figure 4.8. Biotinylated oligonucleotide.
A. Magnified view showing biotin attached to a nucleotide
B. Size relationship of 50 bp DNA molecule to biotin.
biotinylated DNA had been completely taken up by cells. However, it is possible that the lipopolysaccharide layer may have interfered with streptavidin-biotin binding, making it appear as if the biotin molecule was taken up by cells, though it still remained on the outside of cells. Re-isolation of the biotinylated DNA was not performed.

4.6.2 Crosslinking experiments

Because biotinylation of the double stranded oligo did not prevent complete uptake of the DNA, the bulkier streptavidin-agarose molecule was attached prior to incubation with cells. This association essentially attaches an agarose bead to the end of the DNA molecule. Complete uptake of the complex is impossible since the agarose beads are approximately 100 times larger than an *H. influenzae* cell (100 μm vs 1 μm).

The maximum volume of a crosslinking experiment is 50 μl. Therefore, I attempted to determine the number of cells that are able to bind to 50 μl of streptavidin-agarose beads. First, biotinylated DNA was incubated with streptavidin agarose beads for 30 minutes. Unbound DNA was removed by extensive washing (Figure 4.9 a). The complex was then incubated with freshly made MIV-competent cells for 10 minutes (Figure 4.9 b). After this incubation, the streptavidin-agarose beads were washed extensively to remove unbound cells. Following the removal of unbound cells, plating showed that there were 1.72 X 10^6 cfu remaining associated with 50 μl of streptavidin-agarose beads (approximately 1 X 10^3 cfu were associated with beads in the absence of DNA). This is the maximum number of cells available for crosslinking in a 50 μl volume.
Crosslinking Results and Discussion

Figure 4.9. Illustration of the presumed interaction of competent *H. influenzae* cells with biotinylated DNA attached to agarose beads. Relationship not to scale.
Crosslinking experiments were performed using the cell-bead complex and SDS-PAGE gels were stained using a silver stain kit. Bands were not visible in crosslinked or non-crosslinked lanes (gels not shown).

4.7 Calculations

Calculations were performed to evaluate why proteins were not visible after crosslinking. I predicted the reason for this was that the total mass of all the receptors in 1.72 \times 10^6 cells was not enough to visualize the band on silver stained polyacrylamide gel. Sensitivity tests of the silver staining indicated a detection limit of approximately 5 ng of protein per band.

To perform these calculations, some estimates were made of the size and number of receptors. Calculations were performed using two different estimates for these variables. Previous results estimated between three and eight receptors per cell. In the calculations below I used an estimate of 10 and 100 receptors per cell. Also, I used two estimates for the size of the receptor protein, 50 and 200 kDa.

4.7.1 Scenario 1; estimate 100 receptors/cell and a 200kDa receptor.

\[
1.72 \times 10^6 \text{ cells} \times 100 \text{ receptors/cell} = 1.72 \times 10^8 \text{ receptors}
\]

\[
1.72 \times 10^8 \text{ receptors} \times 1 \text{ mol/6.02} \times 10^{23} \text{ receptors} = 2.86 \times 10^{-16} \text{ moles}
\]

\[
2.86 \times 10^{-16} \text{ moles} \times 200000 \text{ g/mol} = 5.71 \times 10^{-11} \text{ g}
\]

\[
5.71 \times 10^{-11} \text{ g} = 57 \times 10^{-12} \text{ g} = 57 \text{ pg}
\]

The minimum amount of protein per band needed to be detected by silver staining: 5000 pg/band

\[
\frac{5000 \text{ pg}}{57 \text{ pg}} = 88^* \]
4.7.2 Scenario 2; estimate 10 receptors/cell and a 50kDa receptor.

Same calculations as above

\[
\frac{5000 \text{ pg}}{1.43 \text{ pg}} = 3500^* 
\]

*This is the number of 50 \( \mu \)l aliquots needed, to have 5 ng (5000 pg) of receptor in a single band on a polyacrylamide gel.

*This assumes 100% of receptors become crosslinked to a DNA fragment.

From these calculations I concluded that in order to visualize the receptor, between 100 and 3000 samples must be loaded into a single well of a polyacrylamide gel. As outlined in the Introduction the efficiency of crosslinking is 1-20%. Therefore, the number of samples needed to isolate 5 ng of receptor from 1.72 \( \times \) 10\(^6\) cells is likely between 1000 - 30000. With this information I concluded that it would be impossible to isolate the receptor by UV laser crosslinking with this strategy. Further attempts to isolate the receptor were not performed.

4.8 Discussion of crosslinking results.

The use of laser crosslinking to isolate the receptor did not yield enough protein for further study. A potential explanation for this is that the interaction of the DNA with the receptor is transient. Transient binding could arise if the interaction of the DNA and receptor is reversible or if the bound DNA molecule is translocated rapidly across the outer membrane. In either case the DNA would be in contact with the receptor for only a short time, decreasing the fraction of USSs in contact with the receptor at the time of UV irradiation.
Complete uptake of the DNA molecule was prevented by using streptavidin-agarose beads. As in the previous experiments this did not isolate enough protein to allow further study because only 1.7% of the cells originally incubated with the streptavidin-agarose beads remained associated after removal of non-specifically bound cells. Calculations performed in chapter four illustrate that with this number of cells it would be nearly impossible to isolate enough protein to visualize a band on a polyacrylamide gel stained with silver. In addition to this, it is unlikely that each receptor of every cell was associated with a DNA molecule. This arises because if we imagine a cell bound to a streptavidin-agarose bead by receptors on one side, it is likely that receptors on the other side of the cell will not be in contact with the bead, thus limiting the number of receptors in contact with DNA when the sample was irradiated. Cells are unlikely to be in contact with more than one bead, due to the beads' large size.

4.9 Future experiments

If crosslinking is to be used again to try to isolate the receptor, a procedure will need to be devised that prevents complete uptake of the DNA molecule, yet allows a high proportion of the receptors to contact DNA. It may be possible to achieve this by linking gold particles to DNA molecule. Previous experimentation has shown that gold labeled DNA remains on the outside of cells after incubation (R. Redfield personal communication). This procedure might prevent complete uptake of the DNA molecule, yet allow a high proportion of receptors to contact DNA, increasing the probability of isolating the receptor.

Also, genomic analysis could be used to identify the receptor. If a computer program were designed that could search the *H. influenzae* genome for proteins that contain transmembrane domains as well as DNA binding domains it could provide a subset of genes, one of which might be the receptor. Deletions could be made of
each open reading frame followed by binding and uptake assays to determine the phenotype of the mutant. The *N. gonorrhoeae* genome sequence may also serve to narrow the number of potential genes that encode the receptor. Since both bacteria are able to bind and take up sequence specific DNA, searching for genes in *N. gonorrhoeae* that have homology to the DNA-binding-membrane proteins found in the original search of the *H. influenzae* genome might assist in isolating the receptor. However, since the USS of *N. gonorrhoeae* is unrelated to the USS of *H. influenzae*, this may not aid in decreasing the number of proteins which could potentially be the receptor.
CHAPTER FIVE

Hypothetical model for uptake by *H. influenzae*

Using the information gained from my research and previous findings, I have formulated a speculative model for how *H. influenzae* might recognize, bind and take up DNA in a sequence specific manner. This model differs from the model proposed by Dubnau (1999) (Figure 1.3 p.12) in a number of ways. Dubnau’s model does an excellent job of addressing the protein composition of the receptor complex inside the outer membrane, therefore I will not discuss the proteins involved in transport of DNA after it crosses the outer membrane. A shortcoming of Dubnau’s model however, is that it does not address how the USS is recognized at the outside of cells. His model hypothesizes that DNA enters the periplasm through an outer membrane pore consisting of the secretin protein PilQ [5]. A more complete model for uptake by *H. influenzae* should address how cells are able to preferentially bind and take up sequence specific DNA and also address how flanking sequence affects uptake. Therefore, my model will focus on the initial steps of binding and uptake in an attempt to supplement the model proposed by Dubnau.

5.1.1 Model for uptake by *H. influenzae*

Figure 5.1 illustrates the proposed model. DNA is bound by four hypothetical proteins on the cell surface (Figure 5.1 (a)). For simplicity I will refer to each protein separately (A - D), although fewer proteins could be involved, each having more than one DNA binding motif. Previous research has shown that, as competence develops, the outer membrane of *H. influenzae* changes in protein composition [55, 56]. Therefore, it is likely that as competence develops, genes are transcribed and the mRNAs encoding these binding proteins are translated then inserted into the
Figure 5.1. Hypothetical model for binding and uptake by *H. influenzae*. 
A) DNA is bound by proteins A-D 
B) Proteins C and D cause melting of the DNA in the A/T rich 3' flanking region eventually leading to the formation of the single stranded core region. 
C) Protein B takes the DNA in to the periplasmic space by forming a bend in the DNA.
membrane. Unpublished experiments have found that when membrane proteins of H. influenzae are solubilized, sequence specific DNA binding is lost (R. Redfield personal communication). Thus, I propose that each protein must contact others of the complex before uptake can commence and if interaction between these proteins is prevented, binding and uptake will not occur.

Uptake begins when the USS recognition protein B binds to the core USS on the DNA. This is followed by proteins A, C and D binding to the DNA (Figure 5.1 a). I hypothesize that proteins C and D contact the minor groove of the DNA since, as illustrated in Figure 1.2 (p.9), the distance from the center of one A/T rich (rwwwww) region to the next is 12 bp, roughly corresponding to one turn of the DNA helix (10.3 bp). Proteins C and D may recognize the A/T rich region by the absence of 2-amino groups in the minor groove of A/T base pairs [91]. I predict that another protein A binds to the 5' region of the DNA, supported by the observation that DNA lacking 5' sequence is not taken up well by competent cells.

In the A/T rich (rwwwww) region, proteins C and D begin to unwind the DNA eventually making the 9 bp core single stranded (Figure 5.1 b). Binding proteins A, C and D hold the single stranded core in close proximity to the USS recognition protein B. This protein then begins to translocate the DNA through the outer membrane (Figure 5.1 c). The mechanism of this movement is unknown but could involve bending of the single stranded DNA to allow passage through the membrane. As DNA is moved into the periplasmic space it re-anneals the single stranded regions. Following movement into the periplasm DNA is translocated across the inner membrane.
5.1.2 Interaction of oligos with the receptor complex

If the proposed model is correct, it should be possible to explain the uptake and binding characteristics of the different oligos used throughout this study. If a DNA molecule lacks sequence 5' to the USS, binding by protein A should not occur. This would prevent the USS from being held in close proximity to protein B (See Figure 5.2 a). Proteins C and D would denature the double stranded DNA. However, protein B, since it is not in contact with the core USS, would not internalize the DNA molecule. This model correlates well with the behavior of USS-1 and USS-50-Le which lack sequence 5' to the USS. USS-1 is not taken up better than its negative control DNA USS-R. USS-50-Le is bound and taken up at a level much lower than USS-50, which has sequence 5' to the USS. It is also possible that if the DNA is denatured in the A/T rich regions (Figure 5.2 b) the 9 bp core may also become denatured. This would lead to the denaturation of the entire oligo which might cause cells to release the single stranded oligo (Figure 5.2 c), preventing uptake.

To explain the low uptake of USS-50-RC, which had an A/T rich region replaced with a G/C rich 3' flanking consensus, it is possible that binding occurs at proteins A - D (see Figure 5.3). The energy required to denature the G/C rich region is higher than an A/T rich 3' flanking consensus. If proteins C and D are unable to denature the DNA, the single stranded USS will not be accessible to protein B and uptake will not occur. If uptake does not begin, cells may release the DNA molecule, leading to the low levels of binding, uptake and interference observed.
Figure 5.2. Model for uptake of oligos lacking 5' sequence
a) DNA is bound by proteins C and D. Protein A is unable to bind to the DNA because the oligo lacks sequence 5' to the USS. Since the DNA is not bound by protein a the 9bp core is not in close contact with protein b therefore uptake does not occur.
b) DNA is bound by proteins B-D. Denaturation of A/T rich regions begins.
c) Denaturation continues to the 9 bp core USS. DNA becomes single stranded and is released by binding proteins.
Model of Uptake

Figure 5.3. Model for binding of an oligo having a G/C rich 3' flanking region.
DNA is bound by proteins A-D. Since this DNA has a G/C rich 3' flanking region, melting in that region does not occur. This prevents the 9bp core from becoming single stranded, preventing uptake of the DNA.
Figure 5.4 shows binding of DNA that lacks 3' flanking sequence \[E^3\]. Since the 3' flanking sequence is missing, if protein B binds the USS the DNA will not be bound by proteins C and D. As above, if this happens then denaturation and uptake will not occur.

If my model is a true description of uptake in *H. influenzae* one might expect that cells should be able to take up single-stranded DNA. However, under 'natural' conditions this does not occur. Previous research has shown that single-stranded DNA can be taken up, although in a non-sequence specific fashion, if cells are first incubated with DNA at low pH, followed by a second period of incubation at neutral pH [95]. An explanation for the lack of uptake under 'natural' conditions could be that proteins C and D bind only double stranded DNA. If so, the core USS would not be localized near protein B and uptake would not occur.
Figure 5.4. Binding of USS-50-Ri.
DNA is bound at 9bp core by protein B. The 3' flanking region is not in contact with proteins C and D therefore denaturation and uptake do not occur.
Bibliography


