Western Graduate&PostdoctoralStudies

Western University
Scholarship@Western

Electronic Thesis and Dissertation Repository

2-23-2012 12:00 AM

Regulation of the bacterial transposon Tn5

Crystal R. McLellan, *The University of Western Ontario*

Supervisor: Dr. David Haniford, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Biochemistry

© Crystal R. McLellan 2012

Follow this and additional works at: https://ir.lib.uwo.ca/etd

 Part of the Molecular Biology Commons

Recommended Citation

McLellan, Crystal R., "Regulation of the bacterial transposon Tn5" (2012). *Electronic Thesis and Dissertation Repository*. 391.
https://ir.lib.uwo.ca/etd/391

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlswadmin@uwo.ca.

REGULATION OF THE BACTERIAL TRANSPOSON TN5

(Spine Title: Regulation of the bacterial transposon Tn5)

(Thesis format: Integrated Article)

by

Crystal R. McLellan

Graduate Program in Biochemistry

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Crystal R. McLellan 2012

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisor

Dr. David Haniford

Supervisory Committee

Dr. Gregory Gloor

Dr. David Edgell

Examiners

Dr. David Edgell

Dr. Christopher Brandl

Dr. Stephen Barr

Dr. Kerri Kobryn

The thesis by

Crystal R. McLellan

entitled:

REGULATION OF THE BACTERIAL TRANSPOSON TN5

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Date

Chair of the Thesis Examination Board

ABSTRACT

Transposons play an integral role in bacterial adaptation, therefore it is important to understand the chemical steps and regulatory factors that govern their mobility. The properties that drive insertion of transposons into subsequent DNA sites have been studied for few elements. The transposon Tn5 in particular has few defined insertion sites and little is known about how its targets are selected. An objective of this thesis was to determine the molecular details of target insertion in Tn5 transposition *in vitro*. To this end, I assessed whether Tn5 could preferentially insert into short oligonucleotide substrates *in vitro*. I detected a Tn5 transpososome complex bound to a short target molecule *in vitro* and observed specific insertion into target sites. I also determined that distortion of the target backbone at precise locations is a determinant for insertion efficiency. This flexibility may facilitate a unique mechanism for specific target selection in Tn5. In addition, the regulation of transposons by bacterial proteins remains elusive for many transposons. A secondary objective of this thesis was to investigate the impact of the host proteins H-NS and Hfq on Tn5 transposition in *Escherichia coli*. H-NS and Hfq were determined to regulate Tn5 transposition *in vivo* in opposing directions. An Δ*hns* strain displayed lower transposition than an *hns*+ strain and an Δ*hfq* strain supported higher levels than an *hfq*+ strain. H-NS was found to specifically bind the transpososome *in vitro* and stimulate transpososome assembly under conditions that would normally disfavour assembly. I determined potential H-NS-DNA binding sites in the transpososome, and found that these sequences were critical for the interaction, and for H-NS up-regulation of Tn5 *in vivo*. Modeling of the H-NS protein onto the structure of the Tn5 transpososome revealed potential interactions. Hfq was found to strongly inhibit

transposase expression under certain growth conditions, specifically from a transposase-*lacZ* translational fusion. Hfq also bound to the 5' region of transposase RNA *in vitro*, providing evidence that Hfq acts directly to repress transposase expression and down-regulate Tn5 transposition. As these host proteins are key mediators of stress responses and virulence, they may act to link cellular physiology to transposition rates.

KEYWORDS: bacteria, *Escherichia coli*, resistance, DNA transposition, Tn5 transposon, transpososome, transposase, strand transfer, regulation, H-NS, assembly, Hfq, translation

CO-AUTHORSHIP STATEMENT

This thesis contains material from published (Chapters 2 and 3) and submitted (Chapter 4) manuscripts. On all manuscripts Crystal R. Whitfield (married name McLellan) was the first author and David B. Haniford was the corresponding author. C. Whitfield carried out all experimental work and prepared all tables and figures presented within this thesis except for the experimental work presented in Figure 5.7, which was performed by Joseph A. Ross, and Figure 5.1 is an adaptation of work prepared by J. Ross.

Chapter 2

A version of this chapter has been published. Reproduced with permission from Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2006) Formation, characterization and partial purification of a Tn5 strand transfer complex. *Journal of molecular biology*, **364**, 290-301. S. Wardle contributed to this work by purifying proteins and assisting with experiments. D. Haniford and C. Whitfield contributed to writing of the first draft and subsequent revisions.

Chapter 3

A version of this chapter has been published. Reproduced with permission from Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2009) The global bacterial regulator H-NS promotes transpososome formation and transposition in the Tn5 system. *Nucleic acids research*, **37**, 309-21. S. Wardle contributed to this work by purifying proteins. C. Whitfield wrote the first draft and subsequent revisions were made by C. Whitfield and D. Haniford.

Chapter 4

A version of this chapter has been submitted for publication.

Whitfield,C.R., Shilton,B.S., and Haniford,D.B. (2012) Identification of basepairs within Tn5 termini that are critical for H-NS binding to the transpososome and regulation of Tn5 transposition. *Mobile DNA*. B. Shilton was co-author of the manuscript and contributed by conducting protein modeling and preparing Figures 4.8 and 4.10. C. Whitfield wrote the initial draft of the manuscript and subsequent revisions were made by C. Whitfield and D. Haniford.

ACKNOWLEDGMENTS

My journey through graduate school has been supported by many people. I would like to thank my supervisor, lab associates, friends, and family for their encouragement.

To my graduate supervisor Dave, thank you for your help and patience. As a supervisor you were always dependable and provided a critical eye when necessary. You nurtured my desire to work independently, and for that I also thank you.

I would also like to thank my graduate advisory committee members, Dr. Dave Edgell and Dr. Greg Gloor, for important discussions and advice during my PhD candidacy. My thanks also go out to other members of the biochemistry department who offered assistance to my work specifically, Dr. Chris Brandl, Dr. Megan Davey, Dr. Brian Shilton and all the wonderful ladies in the biochemistry head office.

Although I thank many members past and present of the Haniford and Brandl labs for making this journey less stressful, I would like to address the following people: Simon, Michelle, Julie, Janine, Joe, Ryan, and Brian. Thank you for your assistance when it came to research projects, and as importantly, for your antics when I needed a break.

To my wonderful parents, Barbara and Paul, thank you for being there for me during these past few years. I also want to thank my parents-in-law Colleen and John for their continuous support and prayers, and to all my extended family whom have eagerly been waiting for this day with me.

Finally, I need to thank my husband Brodie for standing by me through this journey. Despite the challenges we have faced in the last six years, your unending encouragement and love has helped me through the difficult times.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF APPENDICES

Page

LIST OF ABBREVIATIONS

aa	amino acid
bp	base pair
bs	bottom strand
BSA	bovine serum albumin
dsDNA	double-stranded DNA
DNA	2’ deoxyribonucleic acid
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EK	glutamate to lysine substitution
EMSA	electrophoretic mobility shift assay
FD	flanking donor
HA	hyperactive
het	heterogeneous
IE	inside end
IN	integrase
IS	insertion sequence
IS50R	insertion sequence 50 right
IS50L	insertion sequence 50 left
kb	kilobase
kDa	kilodalton
LGT	lateral gene transfer
LP	leucine to proline substitution
MA	methionine to alanine substitution
MALDI-TOF	matrix-assisted laser desorption/ionization-time of flight
MC	monomer complex or ‘single-end’ complex
ME	mosaic end
MP	missing phosphate
mRNA	messenger RNA
NHS	N-hydroxysulfosuccinimide
nt	nucleotide
NTS	non-transferred strand
OE	outside end
PAGE	polyacrylamide gel electrophoresis
PAPI	polyadenylate polymerase I
pBpa	p-benzoyl-L-phenylalanine
pBpaRS	pBpa tRNA synthetase
PEG	polyethylene glycol
PC	pre-cleaved
RB	reaction buffer

RNA	ribonucleic acid
RT-PCR	reverse transcription-polymerase chain reaction
STC	strand transfer complex
STP	strand transfer products
sRNA	small RNA
T3-37	target 3, 37 nt
T2-41	target 2, 41 nt
T3-17	target 3, 17 nt
T'ase	transposase
T'some	transpososome
TE	Tris-EDTA buffer
ts	top strand
TS	transferred strand
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
UM	unmodified
UTR	untranslated
WT	wild-type

Ap	ampicillin
Ble	bleomycin
Cm	chloramphenicol
Kan	kanamycin
Sm	streptomycin
Tet	tetracycline

Arg	arginine
Gln	glutamine
His	histidine
Ile	isoleucine
Phe	phenylalanine
Pro	proline
Ser	serine
Tyr	tyrosine

CHAPTER 1

1 GENERAL INTRODUCTION

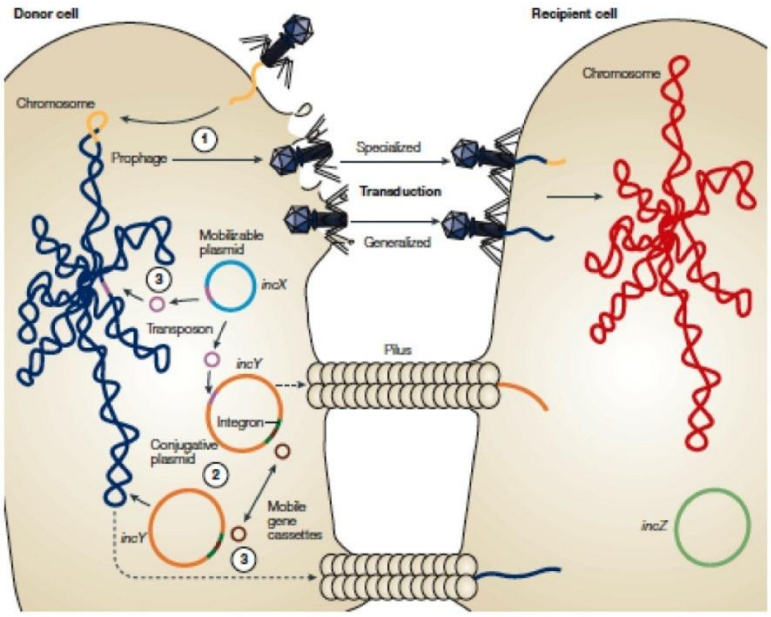
1.1 Lateral gene transfer and transposons in bacteria

Lateral gene transfer (LGT) has played a significant role in the evolution, diversification, and speciation of bacteria. It is presumed to be responsible for the high adaptability of bacteria to new environments as opposed to the accumulation of point mutations (1). In fact, none of the major distinguishing phenotypes of *Escherichia coli* and *Salmonella enterica* arose through point mutations of common ancestral genes (2). Furthermore this type of genetic transmission facilitates the spread of antibiotic resistance determinants through bacterial populations (and among diverse taxonomical and ecological species) and has been implicated in cancer in humans (3, 4).

LGT refers to the movement of genetic material that does not involve transmission of DNA through generations. There are several mechanisms responsible for this type of genetic transfer in bacteria (Figure 1.1) (5). Transduction involves incorporation of foreign phage DNA into a new bacterial genome typically through site-specific recombination as for λ phage. Conjugation results in the transfer of unique 'F' plasmids from one bacterial cell to another through direct contact. Certain bacteria can also take up DNA from the surrounding environment through natural transformation. Importantly, processes such as transduction and conjugation can mediate shuffling of genetic information between cells as removal of phage DNA can accidentally package bacterial DNA in the capsid, and F plasmids can accumulate chromosomal segments (6). Other processes such as transposition of transposable elements or integrons also contribute to the horizontal movement of genetic information. Transposons are discrete mobile

elements that encode proteins called transposases that catalyze the movement of the element from one location to another (i.e. a 'transposition' event). The impact of transposon dissemination on LGT is compounded by the fact that transposons can insert into plasmids, phage genomes, or other mobile elements, which can then be transferred among bacterial cells. For example, recombination of prophages with resident transposons contributes to the well-documented mosaic structure of phages (7). Transposition differs from the classical mechanisms of gene transfer described above because: 1) transposition events can occur within the resident genome, providing a source of intrinsic mutability; 2) they often transpose autonomously and do not require homologous recombination for integration; and 3) they must be contained within a plasmid or genome to be maintained.

Transposons specifically are talented at disseminating antibiotic resistance (8, 9). This is partly because selection has accumulated resistance factors within numerous transposons (10). It has been proposed that approximately 50% of antibiotic resistance cases identified recently are caused by transposons (3) and the rise in antibiotic resistance may be viewed as transposon evolution as opposed to purely bacterial evolution (11). For example, the IncM plasmids encoding resistance to gentamicin first isolated in 1972 were similarly isolated years later however, contained many additional resistance determinants and transposons such as Tn6, Tn7 and Tn9 (12). This is not surprising as transposons are abundant in bacteria, they can integrate multiple resistant determinants in tandem where expression is driven by a single promoter, and they can easily accumulate new resistance factors.



Nature Publishing Group

Figure 1.1. Mechanisms of DNA transfer in bacteria.

1. Transduction: The DNA genome (yellow) of a temperate phage inserts into the chromosome (dark blue); it later replicates, occasionally packaging host DNA alone (generalized transduction) or with its own DNA (specialized transduction), lyses the cell, and infects a recipient cell in which the novel DNA recombines into the recipient host cell chromosome (red). 2. Conjugation: Large, conjugative plasmids (orange) use a protein structure known as a pilus to establish a connection with the recipient cell and transfer themselves to the recipient. Alternatively, a copy of a small, multicopy plasmid, defective genomic island or a copy of the entire bacterial chromosome can be transferred to a recipient cell where the mobile genetic elements either insert into the recipient chromosome or replicate independently if compatible with the resident plasmids (light green). 3. Transposition: Transposons (pink) integrate into new sites on the chromosome or plasmids by non-homologous recombination and are subsequently transferred to recipient cells by the mechanisms described above. Integrons (dark green) use similar mechanisms to exchange gene cassettes (brown). Reprinted by permission from Macmillan Publishers Ltd: [Nature Reviews Microbiology] (6), copyright (2005).

1.2 A history of mobile DNA elements

The first evidence of mobile DNA elements surfaced early in the 20th century. The maize geneticist Emerson presumed that ‘unstable mutations’ that did not appear to obey Mendelian rules of inheritance were due to temporary locus restrictions caused by some type of irregular inhibition (13). Approximately 30 years later McClintock noticed a similar pattern. As a well-established geneticist she was interested in determining a pattern of inheritance in maize; she attributed the phenotypes to particular gene fragments that ‘moved’ that she termed ‘controlling elements’ (i.e. Ac/Ds). The Ds locus, in addition to the unlinked Ac locus, had a propensity to cause breakage of chromosome 9; this coincided with the ‘transposition’ or movement of Ds from its original position to a new location (14, 15).These elements were not termed mobile DNA until well after their discovery and the concept of a dynamic genome, consisting of distinct DNA regions that could move, was not easily accepted by the scientific community.

Fortunately, numerous mobile elements in almost every organism have been isolated and studied since her discovery. In the early 1960’s researchers discovered a particular bacteriophage that caused a ‘mutator’ phenotype in bacteria due to insertion of the phage DNA (16). This phage was subsequently named Mu and provided the first example of a DNA-based transposon that transposed via a replicative mechanism. Mu remains one of the best-studied transposition sBT1 0 0 1 223.54 26 Tm.srs dDWHba [illegible]

comprised of two near-identical ISs that encompass a relatively short region of DNA were discovered in the early 1970's (18). In eukaryotes, the first Ty retrotransposons were discovered in yeast early in the 1980's (19). These RNA-mediated elements termed long terminal repeat (LTR)-containing retrotransposons are similar to animal retroviruses, but do not encode *env* proteins or have an infection cycle. Many more examples since then have been documented, such as the P-element in *Drosophila*, mariner elements found in several animal species and Alu sequences in humans (20-23).

Advancements in genome sequencing have revealed an association with what is termed 'junk' DNA and the presence of mobile DNA elements in many organisms. Sequencing uncovered that almost half of the 'junk' in the human genome contains long-inactivated transposable elements, RNA viruses, retrotransposons, and other mobile genetic elements (24). It also led to the discovery of hidden regulatory systems, such as non-coding RNAs in bacteria and microRNAs in mammals, and allowed computational biologists to establish the lineage of elements in prokaryotic and eukaryotic organisms. For example, the Alu retrotransposons were used to trace the human lineage to Africa independently of mitochondrial DNA analysis by analyzing the pattern of polymorphic sequences at sites where Alu elements previously transposed (25). These discoveries have placed further importance on transposable elements and how they contribute to the biology and evolution of organisms.

1.3 Impact of transposons on the medical community

Initially, the medical impact of transposons was focused on the direct consequences of moving discrete DNA fragments throughout a genome. Transposition events can

negatively affect the stability of a genome by creating double-strand breaks, insertions or deletions, and other chromosomal rearrangements that can occur from a number of complex joining reactions. There are many documented cases of disease states in humans that result from transposon/retrotransposon-mediated rearrangements: L1 element insertions have been known to cause many diseases such as hemophilia, muscular dystrophy and colon cancer, over 20 different Alu element insertions have been noted to cause disease states such as Apert's syndrome and breast cancer (26).

Mobile elements can also affect gene expression by shuffling regulatory elements or by inserting any of the promoters found within the elements directly adjacent unregulated genes or open reading frames. In human cases, an L1 insertion was found responsible for enhancing transcription from the apolipoprotein(a) gene, and insertion of a region from a LINE 3' UTR segment was determined to cause transcriptional activation of a growth hormone gene cluster (27, 28). In bacteria, a transposon insertion in the species *Burkholderia* places a promoter from the transposon in correct orientation upstream of a coding sequence that metabolizes a particular toxin (29), and activation of the cryptic *bgl* operon in *E. coli* involves insertion of IS1, IS5 and Mu (30-32)

Interestingly, DNA transposons have the potential to be used as vehicles for gene delivery to treat a variety of diseases. A clear advantage to using transposons compared to viruses as carriers is that they do not incur potentially fatal immunological reactions and they insert more randomly, whereas viruses pose risks for insertional mutation of endogenous genes resulting in serious consequences. Recently, The Sleeping Beauty Transposon System (SBTS) was created as a non-viral carrier of genes to be incorporated into human chromosomes (33, 34). Interestingly the SB transposon was generated by

reassembly of inactive mariner-like elements in salmon to form a complete and functional element in humans and other eukaryotic cells; it has been used to successfully treat Fanconi anemia in humans (35) and haemophilia A in mice (36, 37).

As previously mentioned, transposons are also intimately involved in antimicrobial resistance, which is a mounting health concern. In the past decade the number of resistant organisms, and the geographic locations affected by these organisms has increased dramatically in part because bacteria become resistant to antimicrobial agents within a short time of exposure. Penicillin-resistant *Staphylococcus aureus* appeared in the 1940's only years after the discovery of penicillin (38). This pattern persisted over time with new bacteria and led health workers to believe that drug resistance is mobile; this hypothesis was later confirmed by scientific research and medical observations (39, 3). Many transposons were discovered directly from patient samples containing resistant bacteria, and more are being recovered with antibiotic resistance each year (e.g. Tn5382 (40), Tn502 and Tn512 (41), and CTn6002 (42))

1.4 Host interactions with transposition

Transposons and other mobile genetic elements present a dilemma to the host cell. Although they are contributors of genetic variability and many carry potentially beneficial genes, high levels of transposition can cause an overwhelming number of destructive chromosomal rearrangements. Therefore a balance must be maintained between activity and inactivity that is crucial for the survival of the transposon. Transposon sequences are determined by the same selective pressure as other genes in bacteria, and for many elements random mutations over time have created strong

regulatory forces against activity at the transcriptional and translational level (43, 44). For example many bacterial transposons have weak promoters that drive expression of transposase proteins and the presence of Dam methylation sites or early transcription termination signals further down-regulate transposase levels. Weak translational initiation, frame-shifting and protein instability can also influence transposition. Even assembling the nucleo-protein complexes that carry out transposition chemistry is difficult due to sub-optimal binding sites for transposases and sub-optimal transposase sequences among other factors (45). Regulatory elements that have accumulated within the transposon sequence such as those described above are termed intrinsic regulators, however regulation is also exerted by bacterial host proteins.

Evolutionary pressure has recruited different types of host proteins with established cellular functions to influence transposition. Obvious examples include the use of replication proteins and other proteins that are required to repair gaps and double-strand breaks due to transposition. Other less obvious host factors can either promote or down-regulate transposition either directly or indirectly. In the context of this introduction, indirect regulation represents host factors that affect the transcription or translation of transposition proteins (or other regulators) and any regulation that does not occur at the level of transposition (e.g. supercoiling). Direct regulation represents host factors that must contact the transposon sequence, transposase, or transpososome complexes in order to influence transposition.

Several proteins responsible for the maintenance of nucleoid architecture in bacteria, including IHF, HU, Fis and H-NS, influence the activity of transposons. IHF, HU, and H-NS have all been shown to bind transposon sequences directly to stimulate

transpososome assembly or transposition either *in vitro* or *in vivo* (e.g. Mu, Tn10, Tn1000) (46-55). Conversely, H-NS indirectly affects IS1 transposition by protecting the protein stability of the IS1 transposase *insAB'* (56, 57). Indirect regulation of Mu is also mediated by H-NS and Fis by stabilizing the Mu repressor-DNA complex that controls activity of transposase (58) or affecting phage gene expression that in turn influences transposition, respectively (59, 60).

Other host factors that affect transposition are the supercoiling agents DNA Gyrase and Topoisomerase I, the replicative factor DnaA, the chaperone ClpX, acyl carrier protein (ACP) and ribosomal protein L29. The supercoiled state of either the transposon or the target location influences activity, and Gyrase binding sites in Mu prophage genome are required for transpososome assembly of the ends (61). DnaA (62) and ClpX (63) are presumed to assist with disassembly of transpososome complexes after integration of Tn5 and Mu, respectively, and L29 and ACP simulate recognition of the specific Tn7 target site *attTn7* by the transposition protein TnsD (64). Even an RNA-associated protein, Hfq, has been shown to directly influence transposition (e.g. Tn10/IS10) (65). Although past studies have fully explored the intrinsic control of many transposons, a full regulatory network of host proteins or pathways that extrinsically govern any transposon has yet to be revealed.

Another variable that must be addressed when discussing transposon regulation by host factors is the growth environment of the bacterial cell. Interestingly, many of the host proteins discussed above such as the nucleic acid binding proteins H-NS and Hfq are intimately involved in stress response pathways that respond to fluctuations in temperature, osmolarity, oxidation or phosphate (66-71). Transposon regulation in

eukaryotes has identified many examples of how transposons can cause rearrangements when stress occurs (72-74). It is presumed that in bacteria under certain environmental conditions, such as nutrient-starvation or entry into stationary phase, the balance between activity and inactivity of may shift (an idea first proposed by Barbara McClintock). Despite much effort little direct evidence exists to link environmental factors with induction of transposition (75). Further evidence is required to substantiate the claim that environmental changes and stress can induce transposition.

1.5.0 The Tn5 transposon

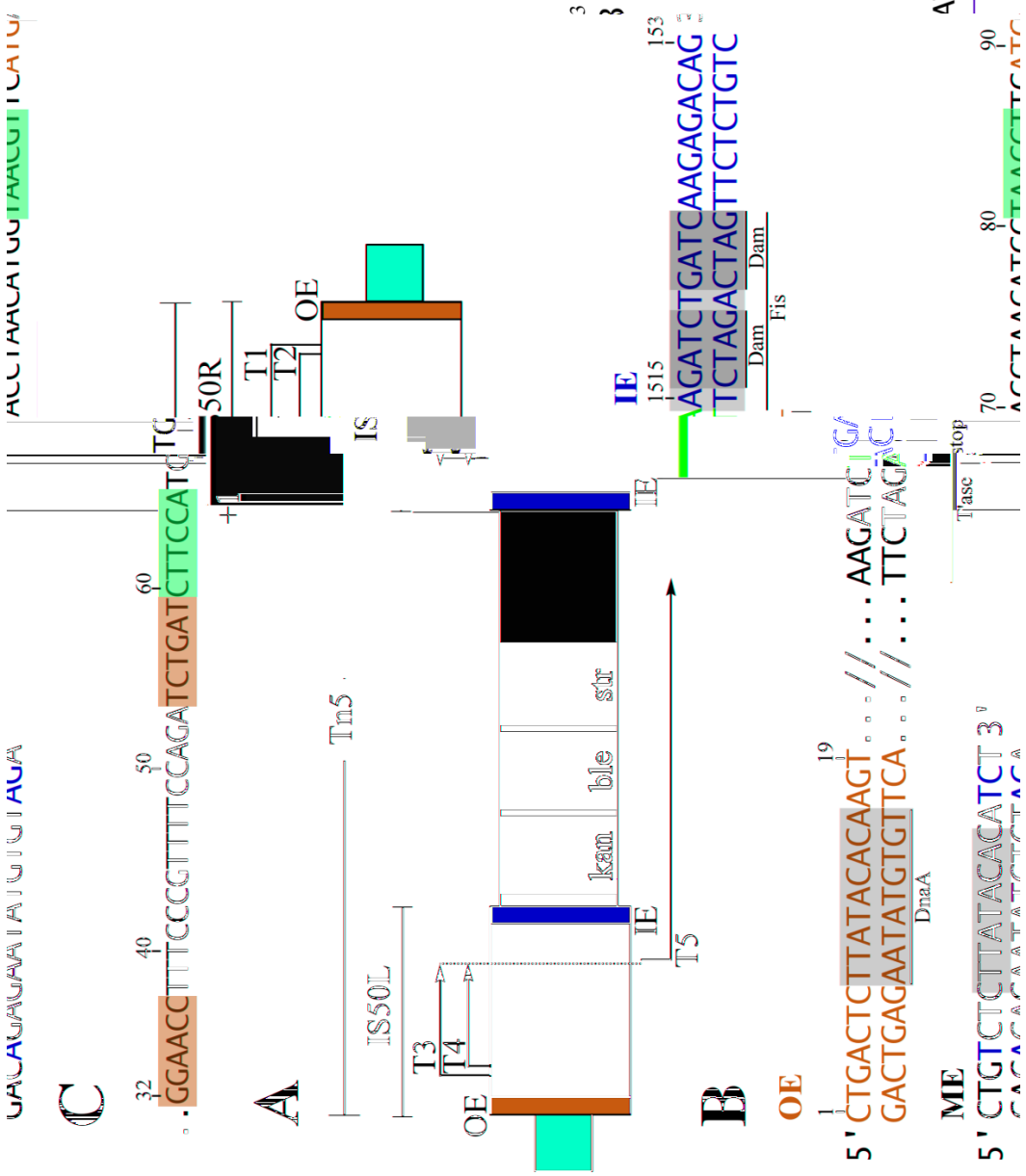
Tn5 was one of the first transposons discovered. Berg *et al* noticed that insertions of the R factor kanamycin resistance determinant into the bacteriophage λ genome differed structurally from the insertion of *E. coli* chromosomal genes (76). Sequencing of the transferred DNA determined that it was composed of two inverted repeats that encompassed three genes encoding antibiotic resistance. Subsequent work has verified Tn5 as a composite bacterial transposon and a member of the IS4 family of transposable elements. It has served as a model system to study transposition *in vivo* and *in vitro* for several decades (77).

1.5.1 Tn5: a model system

This thesis focuses on two main areas: the characterization of determinants for target selection by a bacterial transposon and the host regulation of bacterial transposition by global regulators. The Tn5 transposon has many attractive features that make it an appropriate model system to study these phenomena. Tn5 is a relatively simple and

Figure 1.2. Structure and regulatory sequences of the Tn5 transposon.

(A) Schematic representation of the 5818 bp Tn5 transposon. Inverted insertion sequences (IS) 50R (right) and 50L (left), represented by large white rectangles, are 1533 bp in size. These flank the transposon and consist of specific recognition sequences (19 bp) that define the inside end (IE; green rectangle) and the outside end (OE; purple rectangle), in addition to promoters and sequences that express transposition proteins. Flanking DNA is represented by gold rectangles. Only the proteins produced from IS50R (transposase and inhibitor) are involved in the transposition process. A single nucleotide change in IS50L (represented by the vertical dashed line) creates an ochre stop codon that renders the proteins P3 and P4 inactive. Fortuitously, this nucleotide change creates a better transcriptional signal that produces a transcript corresponding to the remainder of the IS and into the intervening sequence. The intervening sequence contains three open reading frames that allow resistance to the antibiotics kanamycin, bleomycin, and streptomycin (black rectangles) that are transcribed by the fortuitous promoter. Transcriptional start sites are indicated by black arrows. (B) Sequence composition of the OE and IE in the context of IS50R. OE sequence and IE sequence are represented by purple and green lettering respectively. The stop codon for transposase is indicated in red. DnaA and Fis binding sites are indicated in addition to Dam methylation sites (grey boxes and underlined). The mosaic end (ME) sequence consisting of sequences from both the outside and inside ends is also depicted. Sequence in black is shared by the OE and in green shared by the IE. (C) Regulatory sequences governing transposase and inhibitor synthesis of IS50R. '-35' and '-10' transcriptional sequences are indicated by purple and orange translucent rectangles and horizontal arrows for T1 and T2 transcripts (indicated in (A)) respectively. Purple and orange 'ATG' represent start codons for transposase and inhibitor proteins. Numbering above the sequence begins at the 5' end of the OE.


C
A
B
OE
ME
IS50R
T1
T2
OE
IS
Tn5
IS50L
T3
T4
OE
km
ble
str
IE
T5
IE
IE
Dam
Dam
Fis
DnaA
Tase
stop
+1
5'
3'

efficient ‘cut-and-paste’ transposon that has a convenient *in vitro* set-up to study the mechanistic details of transposition. The reaction requires only short double-strand DNA that represents the ends of the transposon and purified ‘hyperactive’ transposase protein; unfortunately the wild-type transposase has little activity *in vitro* (78, 79). Recent biochemical and structural studies using X-ray crystallography have provided detailed molecular structures for Tn5 transposase and the Tn5 transpososome structure (80, 81). This makes the Tn5 system useful for inferring structural changes from biochemical evidence.

Additionally, we chose the Tn5 system to study the influence of the bacterial global regulator H-NS on transposition because several other similar elements such as IS903 and Tn10/IS10, which both transpose similarly to Tn5, had been determined to be influenced by this host protein. Similarly, we chose Tn5 as a model system to study the impact of Hfq on transposition because Hfq was found to drastically impact Tn10/IS10 transposition. Hfq was inadvertently found to regulate Tn10/IS10 transposition while determining the effects of H-NS on transposition *in vivo*. An Δ*hfq* strain supported 80-fold higher transposition levels and this was later partially attributed to Hfq affecting transposase expression by promoting the pairing of an antisense RNA to the Tn10 transposase transcript. Another unknown mechanism that down-regulates transposase expression was presumed to exist because transposition was still increased (8-fold) in an Δ*hfq* strain when a mutation was placed in the antisense RNA that does not confer any RNA pairing. Tn5 was an excellent choice to examine the possibilities of this unknown regulatory mechanism as Tn5 is similar in genetic structure and mechanism to the Tn10

transposon, and no evidence of antisense or *trans*-encoded RNA regulation have been reported for Tn5.

1.5.2 Tn5 and IS50 structure

Tn5 is 5.8 kb in size and consists of two near-identical bacterial insertion sequences (IS50Right and IS50Left) of 1.5 kb in inverted orientation surrounding open reading frames for three antibiotic resistance determinants: kanamycin, bleomycin and streptomycin (Figure 1.2) (82). The promoter for these genes is located within IS50L and results from a single base pair difference compared to IS50R that creates a sequence closer to the canonical σ70 promoter sequence (83).

The genetic requirements for Tn5/IS50 transposition are located within the IS50 elements. Each IS element consists of a 19 bp outside end (OE) and inside end (IE) sequence that define the boundary of the element and the OE sequences together define the boundaries of the Tn5 transposon. These act as binding determinants for the transposase protein in order to catalyze transposition. The two ends differ in only 3 positions however both function efficiently in the presence of wild-type transposase (84). Interestingly these sequences contain recognition sites for host proteins such as DnaA (OE), Dam methylase (IE) and Fis (IE).

There are two protein products encoded in each element. The single base pair difference in IS50L discussed above also encodes a premature stop codon for the encoded proteins yielding them inactive. Both proteins expressed from IS50R, transposase and inhibitor, are involved in transposition, but only transposase is absolutely required for catalysis (85). The inhibitor protein is an N-terminal truncation of transposase that is

produced with different transcriptional and translational sequences compared to transposase, but in the same reading frame. It acts a *trans*-active inhibitor of transposition by non-productive multimerization with wild-type transposase (86).

1.5.3 Transpososome assembly

The term ‘transpososome’ was formed to describe the complex nucleo-protein structure that mediates all chemical stages of transposition reactions. Extensive protein-DNA contacts as well as protein-protein contacts provide stability to the structure. Due to the fact that the products of each chemical stage are not released, the transpososome is essential for orchestrating the reaction chemistry and offers protection for the cleaved DNA strands from nuclease activity (45, 87, 88). Because no covalently linked DNA-transposase intermediates exist for bacterial transposons nor do they require an external energy source, energy is presumably supplied by transposase binding and the stereoselectivity of the phosphate at each subsequent step determines the order of the reaction steps and drives the reaction forward. These characteristics create a transpososome that becomes more robust as the reaction occurs (45).

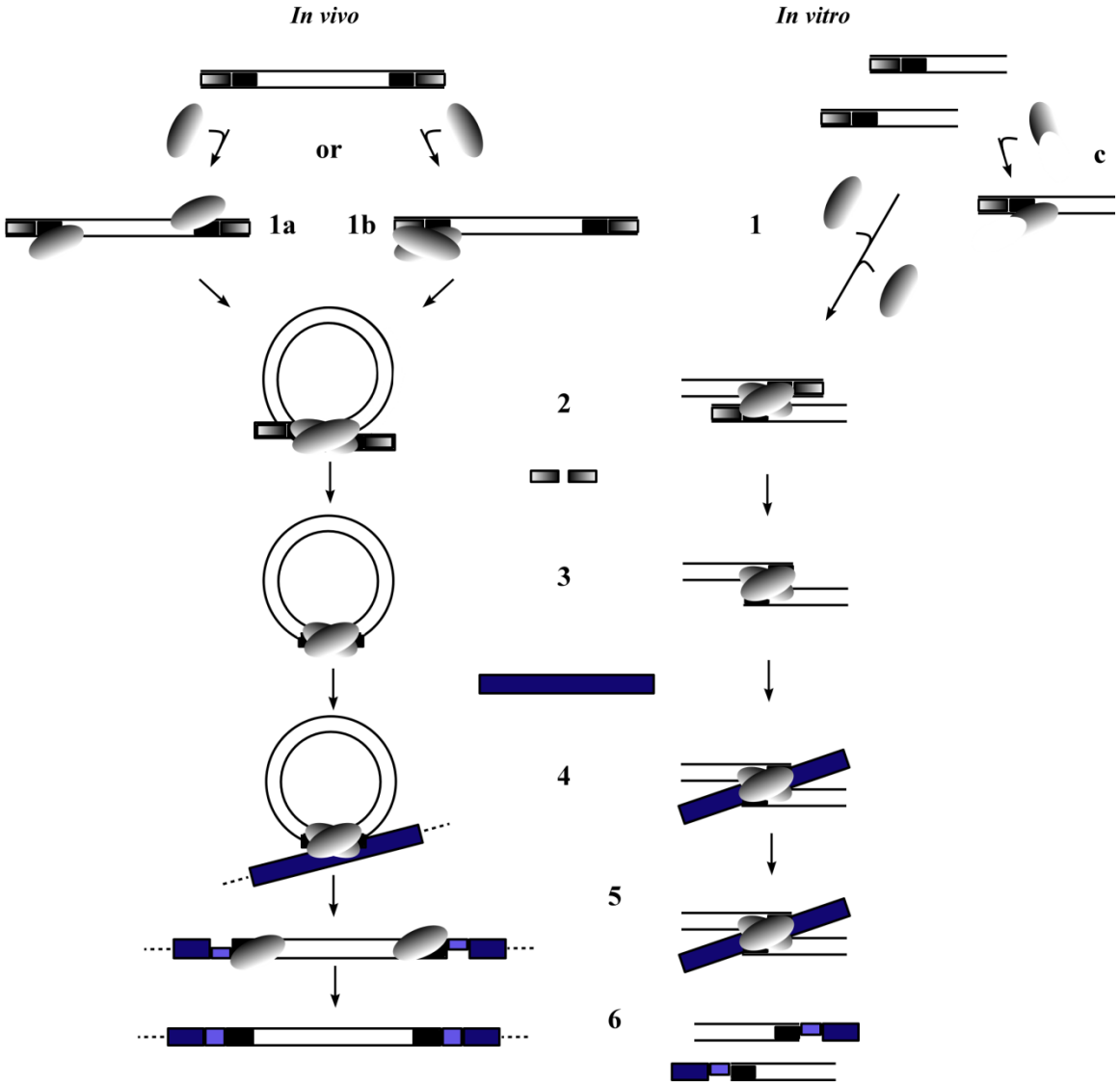
The only factors absolutely resquired for this assembly are the Tn5 transposase protein and the outside/inside end sequences to be transposed. The exact steps of transpososome assembly for Tn5 are largely unknown, and presumed to be complicated especially in a biological context, but ultimately produce a dimer of transposase bound to two end sequences. Fortunately the co-crystal structure of a Tn5 pre-cleaved transpososome and results from genetic and biochemical experiments using mutant transposase proteins have provided insights into this process (84, 80, 89-91, 81).

In order for transpososome assembly to initiate the transposase protein must overcome an inhibitory interaction between residues in its N-terminal DNA-binding domain and C-terminal dimerization domain (92). This inhibitory function is presumed to explain the '*cis*' preference for Tn5 transposition. During translation of the protein, the N-terminal can associate with DNA before the C-terminal is translated. In the event that the DNA-binding domain is in the active, released state, the initial step of transpososome assembly is proposed to be non-specific DNA binding by transposase. This binding is presumed to be important in localization of the end sequences through a direct transfer mechanism, particularly if they are within approximately 450 bp of the non-specific binding site (this may also provide evidence for the *cis* bias in Tn5 transposition) (93, 94). It is unknown whether the two transposase proteins dimerize before or after this initial binding, however biochemical and structural studies indicate that oligomerization occurs between the two C-terminal α-helices of the monomers (90, 80, 95, 96). If we assume that transposase initially binds DNA as a monomer, the two protein molecules dimerize or 'pair' after each end sequence is bound. The structural conformations involved in this step are unknown.

Biochemical and structural data suggest that at this stage one transposase interacts with the end it initially bound as a monomer (*cis*) and the same monomer interacts with the opposite end bound by the other transposase monomer (*trans*) (Figure 1.3). The initial *cis*-contacts occur between DNA positions 6 to 19 of the end sequence (where the first position is 1) and several important residues of transposase (84, 97, 91, 98). The X-ray crystal structure of a pre-cleaved Tn5 transpososome displays these *cis*-contacts clearly (80, 81). *Trans*-contacts between transposase dimer and DNA end sequences occur at

Figure 1.3. Proposed modes of Tn5 transpososome assembly *in vivo* and *in vitro*.

The initial stage (1) requires association of the transposase protein (grey gradient oval) with the end sequences (black rectangles), either outside ends (OE) for Tn5 transposition or an OE and an inside end (IE) for IS50 transposition. It is unknown whether a monomer or dimer of transposase first contacts an end, therefore each possibility is illustrated in 1a. and 1b. *In vitro*, interaction between Tn5 transposase and a single end is not detectable unless a C-terminally truncated transposase is used as depicted in (c). When each end is bound by a single molecule of transposase, the monomer complexes (MC) are brought together by a complex series of dynamic protein-DNA and protein-protein interactions to form the initial transpososome, also termed the synaptic complex (2.). A similar process is presumed to occur even if a dimer of transposase initially contacts one end. Within the initial transpososome complex, excision removes the transposon from the flanking donor DNA sequences (grey gradient rectangles) to form a different category of complex called the cleaved transpososome (3.). In the presence of a double-stranded DNA target molecule (blue rectangle) the complex captures the target, presumbly by binding of the target DNA in a positively charged cleft (4.). Integration of the transposon into the target DNA occurs within the strand transfer complex and represents a complex between (4.) and (5.). After insertion, repair proteins fix the remaining gaps. *In vitro* the stages from 2 to 5 are similar to the proposed *in vivo* events, however after strand joining the resulting *in vitro* products are linear and remain unrepaired (6.).


In vivo
In vitro
or
1a
1b
1
c
2
3
4
5
6

positions 1-5 of the partner end and are inserted into the active site, which consists of the catalytic 'DDE' residues consistent with other IS4 family transposases. This structure represents the initial transpososome and must be assembled correctly in order for transposition to occur.

A consequence of proper transpososome assembly in Tn5 is that catalysis also occurs in *trans* where one end bound to one transposase is catalyzed by the other transposase monomer. This has been shown to occur in Mu, Tn7 and Tn5 and allows for the coupling and correct placement of cleavage and joining reactions at each end (99-101, 80). Although an orchestrated set of structural changes must occur before catalysis, important conformational changes also occur during catalysis. For example the β-loop of transposase (residues 242-333) is not directly involved in catalysis nor required for transpososome assembly, but it is necessary for each subsequent stage presumably by supporting conformational changes (102).

1.5.4 Mechanism of transposition

The 'cut-and-paste' or conservative transposition mechanism presumed to be used by Tn5 is shared by many other elements such as Tn10 (103, 104), Tn7 (105, 106), P-element (107, 108) and Tc1/mariner elements (109, 110). The mechanism was investigated thoroughly in Tn5 after establishment of an *in vitro* transposition system using purified transposase and short oligonucleotides (78, 111). The process involves excision and insertion of the element by three metal-mediated (Mg^{2+} or Mn^{2+}) phosphoryl transfer reactions: 3' strand nicking, hairpin formation, and hairpin resolution (Figure 1.4) (112-114). The transposase active site 'DDE' motif likely co-ordinates two divalent

metals that play an important role in cleavage of the transposon from the donor by activating the oxygen atoms in water molecules and by physically contacting the P-O bond to be broken (115). Oxygen atoms from water molecules are used to nick the backbone phosphate group on the transferred strands at the flanking donor/transposon junctions. Note that the transferred strand of each end refers to the strand that is chemically joined to the target site. This first step is shared by all DNA transposases, retroviral integrases and the LTR retrotransposases (116). The second step involves nucleophilic attack of the free 3'-OH on the transferred strand in-line to the phosphate backbone on the non-transferred strand at the donor/end junction generating a hairpin.

Hairpin formation is used by many bacterial transposons and is the hallmark of IS4 family elements such as the Tn10 transposon and RAG-1-mediated DNA cleavage in V(D)J recombination (117, 118). For Tn5, this releases the donor DNA from the transposon and protects the transposon end sequences within the covalent DNA linkage. Hairpin formation requires extensive distortion of the backbone of the non-transferred strand near the nicked site to form the inter-strand phosphodiester bond. This involves stacking of a 'flipped out' thymine base with a tryptophan residue of transposase and the 5' phosphate of the thymine is held in position with a tyrosine and arginine residue (part of the YREK signature found in IS4 transposases) (119). Studies have shown this phenomenon occurs in Tn10 and V(D)J transposition (120-122). The hairpins are resolved by nucleophilic attack of two metal-activated water molecules releasing the transposon with double-strand breaks at each end. The resulting two double-strand breaks in the flanking DNA are repaired by the host's double-strand break repair mechanism of homologous recombination.

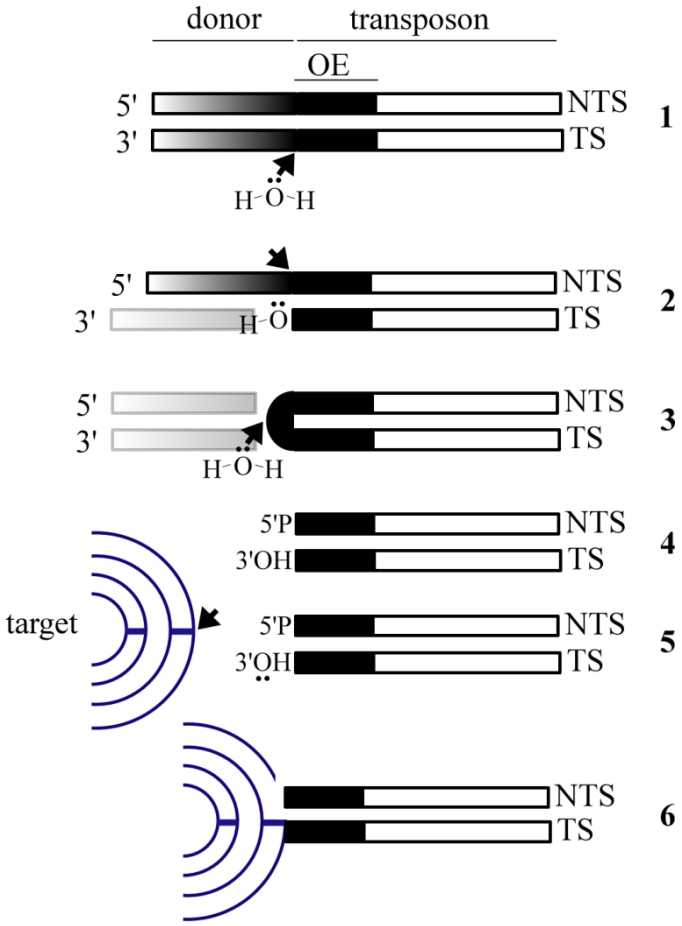


Figure 1.4. Mechanism of Tn5 transposition.

In the context of the initial transpososome (see Figure 1.3) an activated water molecule performs nucleophilic attack (at the black arrow) and hydrolyzes one strand (the transferred strand, TS) of the DNA (1.), which exposes a 3'OH group at the end of the transposon adjacent to the flanking donor DNA (2.). Individual strands are represented by rectangles. The 3'OH is then activated to perform a nucleophilic attack on the opposite strand of DNA (at the black arrow), forming a hairpin structure and excising the transposon from donor DNA (3.). Accurate hydrolysis of the hairpin results in blunt-ended DNA at the transposon end, free from flanking DNA, yet still held tightly in the transpososome structure (4.). After binding of the transpososome to a target molecule, the activated 3'OH groups at the ends of the transposon (5.) perform nucleophilic attacks on the 9 bp target site, to successfully insert the transposon into the new location (6.). NTS, non-transferred strand.

Integration of the cleaved transposon bound to the transposase dimer occurs into a 9 bp target site for Tn5, however there is little known about target binding or capture (123). It is hypothesized that DNA phosphate coordination by residues R210, Y319 and R322 of transposase results in movement of the 5' non-transferred strand DNA phosphate away from the 3' end allowing for efficient target DNA binding (81). After target binding, the two free 3'OH of the transposon attack the phosphate backbone of the target site 9 bp apart, resulting in insertion of the element and duplication of the target sequence on either side of the element after repair by the host replication machinery.

1.5.5 Target selection

Proper target selection in bacterial transposition systems ensures that the element is propagated successfully and the element-host relationship is maintained (124). Elements show insertion bias into certain sequences, genomic regions or DNA with particular structural features directly or through the use of accessory proteins. For example, retrotransposons such as Ty1 and Ty3 in yeast prefer to transcribe away from coding regions by inserting upstream of promoters (124). Tn7 uses the transposon-encoded protein TnsE to directly interact with the processivity factor β-clamp to insert into DNA that is being replicated (125). The Tn10 transpososome directly contacts 9 bp target sites with high specificity; a consensus sequence for insertion has been determined for Tn10 (126, 127). Similarly to Mu, there are few characterized 'hot-spot' target sites for Tn5 transposition, making Tn5 the element of choice for transposition mutagenesis.

In vivo studies have shown that Tn5 inserts into a sequence in the promoter of the *tet* gene in pBR322 (128, 129). G/C base pairs at the ends of the 9 bp site as well as the

extent of local negative supercoiling are important for specificity in this site. Further *in vivo* and *in vitro* work confirmed that a G nucleotide at position 1 and a C nucleotide at position 9 are critical for specific insertion (130). In addition, by analyzing a compilation of different insertions by Tn5 into short 1 kb regions *in vivo* and *in vitro*, researchers were able to define a weak 9 bp consensus sequence for insertion as 5'-GNTYWRANC-3'. Interestingly, overlapping clusters of target sites are also an important determinant for target selection in Tn5 as poor consensus sequences that overlap are ideal for Tn5 insertion. Target binding is presumed to be facilitated by the co-operativity generated from transposase-transposase interactions in a transpososome 'microfilament' located along several overlapping sites. The molecular basis for this model is unknown, however it is fundamentally different from the way that transposase recognizes the end sequences in Tn5 transposon. Although an *in vitro* system has been set up to characterize the Tn5 transpososome using short DNA substrates *in vitro*, a complex consisting of target capture or strand transfer had not been studied to validate these results before the work in this thesis (84).

1.5.6 Regulation of Tn5 transposition by host factors in *Escherichia coli*

The Tn5 transposon is regulated by several host proteins. Many play subtle roles in transposition however none have been determined to be essential. Some have already been referred to in an earlier section (1.4) and the mechanisms used by the following proteins to influence Tn5 will be discussed here: LexA, Topoisomerase I, DnaA, Fis (Factor for Inversion Stimulus), and IHF (Integration Host Factor). The LexA protein, a regulator of the SOS response in *E. coli*, directly represses transposase synthesis to

regulate transposition (131). *E. coli* Topoisomerase I positively regulates transposition *in vivo* and *in vitro* and is hypothesized to directly interact with transposase to impact transposition (as opposed to affecting supercoiling) (132). Some evidence suggests that DnaA also interacts directly with the Tn5 transpososome by competing with transposase for a DnaA binding site found in the OE sequence (62). This is presumed to assist in removal of DNA from the complex after transposition. Similarly to DnaA, Fis and IHF binding sites are located within the Tn5 transposons. Fis influences both Tn5 and IS50 transposition differently, promoting Tn5 transposition and inhibiting IS50 transposition by binding to unmethylated Fis binding site within the IE (133). The molecular details of Fis regulation are unclear. Although IHF was found to bind a site within the IS50 sequence using electrophoretic mobility shift analysis (EMSA), the physiological role of this interaction has not been elucidated. IHF is proposed to promote transposition of Tn5 and IS50 in strains that lack Dam methylase however the basis for this regulation is unclear (134).

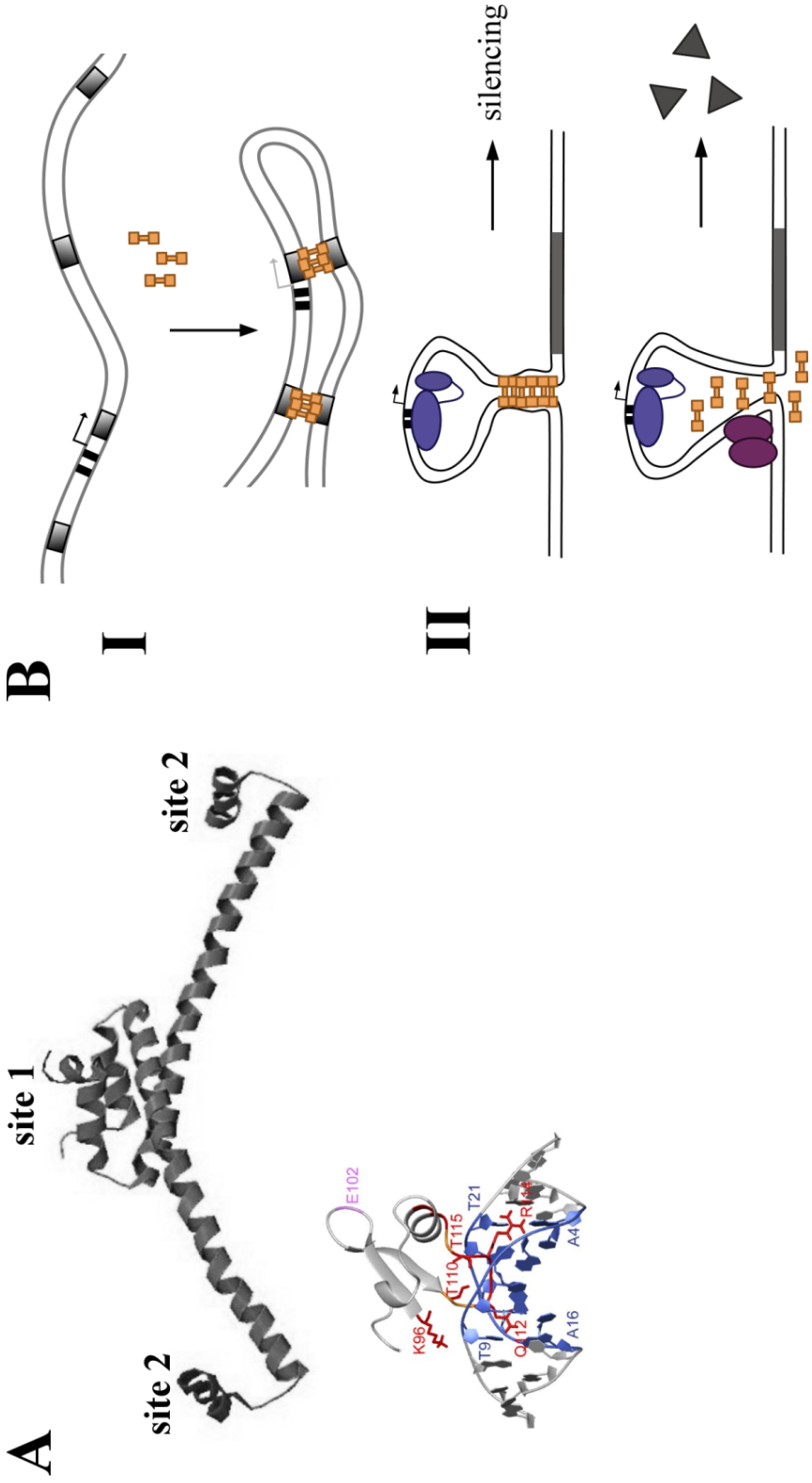
1.6.0 Introduction to host factors investigated in this thesis

1.6.1 The bacterial transcriptional regulator H-NS

The histone-like nucleoid structuring protein H-NS has been studied in bacteria for decades. As a nucleoid-associated factor H-NS assists in the maintenance of chromatin structure, and as a transcriptional repressor H-NS regulates >100 genes in bacteria particularly those involved in responding to new environments (68, 135). H-NS has also been identified as a protein that aids in the acquisition of foreign genes through

Figure 1.5. Structure of H-NS and function in lateral gene transfer.

(A) Structural representation of the individual N-terminal oligomerization domain (NTD) (PDB: ID 3NR7) (136), and C-terminal DNA-binding domain (CTD) docked to a DNA substrate (adapted from PDB: ID 2L93) (137). The solution structure of the CTD was determined using *Salmonella* H-NS_{CTD} (residues 91-137) and the NTD was determined with X-ray crystallography using *Salmonella* H-NS (residues 1-83). In (A) higher-order oligomerization would occur between H-NS monomers at sites 1 and sites 2. Below this, the modeled structure of the CTD bound to DNA shows the protein backbone and side chains that are involved in DNA-binding in red. Residues that may be involved in DNA-binding are labeled in orange. The binding interface of DNA is shown in blue. These domains are connected through a flexible linker region, which is largely unaccounted for in either structure. The N- and C-terminals are indicated. (B) Model for H-NS facilitation of lateral gene transfer. (I) AT-rich sequences characteristic of foreign DNA (black gradient boxes) are bound by H-NS (orange) once integrated into the host chromosome and H-NS polymerization and DNA bridging silences foreign promoters (black rectangles). (II) Subsequent expression of high-affinity DNA-binding proteins (purple ovals) that can outcompete H-NS binding allow gene expression (grey triangles) under appropriate conditions. In this way the bacteria is protected from foreign gene products and over time the foreign system can be re-established and built into host regulatory networks. dsDNA is represented in grey. Black rectangles, transcriptional signals; blue ovals, RNA Polymerase; grey rectangle, target gene. Transcriptional signals and start site are represented by black rectangles and a black or grey arrow. Adapted by permission from John Wiley and Sons (138) copyright (2008).


A
site 1
site 2
site 2
E102
K96
T110
T115
T21
R24
A4
T9
Q12
A16
B
I
II
silencing

lateral gene transfer, as well as a protein that influences recombination and transposition systems.

H-NS is a relatively small (15.6 kDa) basic protein that is highly abundant in bacterial cells (~20 000 molecules per cell) (139). Although a structure of full-length H-NS has yet to surface, many studies provide evidence for a C-terminal DNA-binding domain of unique structure, and an N-terminal oligomerization domain consisting of a coiled-coil α-helical structure (Figure 1.5A). These domains are connected by a presumably flexible linker region whose presence is intimately involved in function. There is conflicting evidence about the active form of the protein however at physiological temperature many results indicate that H-NS functions as a dimer (140-144). Unlike IHF and Fis, H-NS typically binds DNA non-specifically by associating with A/T-rich and odd DNA structural distortions or curved regions (145-149). However *in vitro* binding studies have been able to determine a consensus sequence for H-NS binding (150).

Within the last year, NMR and X-ray crystallography have produced the best resolution structures of both the N- and C-terminal domains providing key information on how H-NS binds DNA and how oligomerization is mediated (136, 137). It appears that the most important binding determinant for H-NS is the width of the minor groove, which is determined by sequence and not curvature (137). H-NS is proposed to contact DNA through a 'hook-like' motif (Figure 1.5.A, bottom). Oligomerization is proposed to occur in a 'head-to-head' and 'tail-to-tail' fashion where protein-protein contacts mediate the interaction at the terminal regions of the oligomerization domain. These recent developments have provided new insights into the function of H-NS as a 'scaffolding'

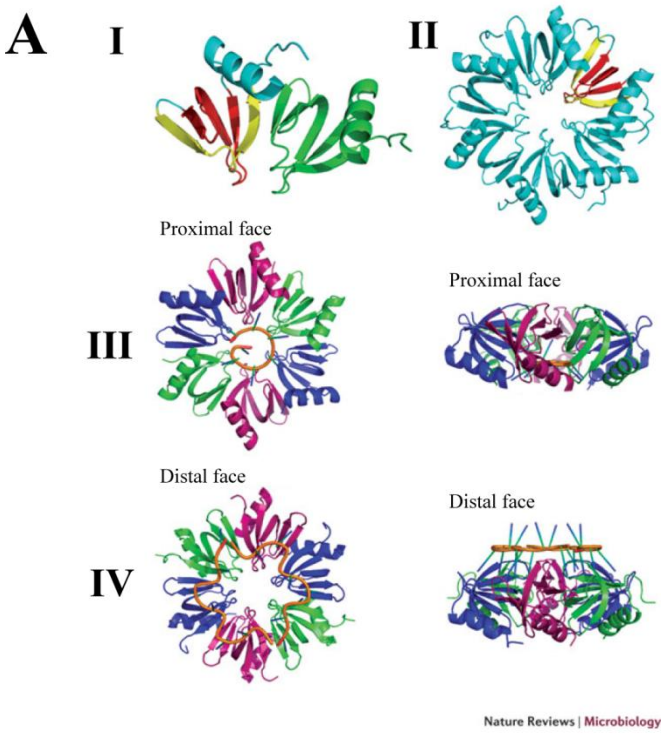
protein to condense DNA and are likely critical for other functions of the protein such as transcriptional repression.

H-NS is a key factor in the adaptation of bacteria to the environment. In addition to its role in maintaining proper architecture of the nucleoid, H-NS is responsible for regulating approximately 5% of the genes in *E. coli*, including pathogenesis determinants and stress response genes to osmotic and acid stress (151). H-NS functions largely as a transcriptional silencer and can interact with promoter regions in order to trap RNA polymerase (152), and can bind downstream regulatory elements in coding regions to effectively block transcription elongation independent of the promoter (153, 154).

H-NS facilitates lateral gene transfer by utilizing a DNA 'bridging' mechanism to loop segments of A/T-rich foreign DNA in order to silence promoters (Figure 1.5B) (155, 138, 156). The bridging capability of H-NS is likely mediated by the flexible linker region and allows a dimer bound at one segment through the DNA-binding domain to be united to a second dimer at a distantly located region. Recombination systems are also regulated by H-NS such as the FimB-mediated inversion of *fimS* responsible for expression of type 1 fimbriae in *E. coli* (157) as well as transposition systems (158-160, 51, 57, 58, 56). H-NS-dependent regulation of transposition varies in mechanism however for many elements H-NS interacts directly with the transposition complex through DNA and/or protein contacts. For the Tn10 transposon both DNA and protein interactions are presumed to be necessary for stimulation of intermolecular events. This represents one of few proteins that directly associate with H-NS; the only other protein shown to directly associate with H-NS is the bacterial motility protein FliG (161). As a

Figure 1.6. Structure of Hfq and proposed mechanisms of riboregulation.

(A) An individual monomer of the Hfq hexamer consists of an α-β(1-5) structure where the β strands form antiparallel sheets (I). The second β-strand is twisted such that it contributes to both sheets and forms a squat barrel (II). (III) and (IV) Each face on opposite sides of the Hfq hexamer can interact with RNAs (orange). The proximal face represents the surface on which the amino-terminal α-helix is exposed and includes residues in the Sm2 sequence motif whereas the distal face is on the opposite side. Disordered protomer 'tails' are likely to emanate from the equator of the Hfq ring and may form electrostatic interactions with nucleic acids. Biological and biochemical evidence suggests that each face is proposed to bind particular RNAs. (B) Proposed mechanisms of Hfq activity. (a) Hfq in association with a small RNA (sRNA) may sequester the ribosome-binding site (RBS) of a target mRNA, thus blocking binding of the 30S and 50S ribosomal subunits and repressing translation. (b) In some mRNAs, a secondary structure in the 5′ untranslated region (UTR) can mask the RBS and inhibit translation. A complex formed by Hfq and a specific RNA may activate the translation of one of these mRNAs by exposing the translation initiation region for 30S binding. (c) Hfq may protect some sRNAs from riboendonuclease cleavage, which is carried out by ribonuclease E (RNase E) in many cases. Hfq may induce the cleavage (often by RNase E) of some sRNAs and their target mRNAs. Hfq may stimulate the polyadenylation of an mRNA by poly(A) polymerase (PAP), which in turn triggers 3′-to 5′ degradation by an exoribonuclease (Exo). Reprinted by permission from Macmillan Publishers Ltd: [Nature Reviews Microbiology] (162), copyright (2011).



Nature Publishing Group

major host regulator involved in adaptation of the bacterial cell to new environments, H-NS activity may serve to stimulate transposon activity in response to cellular physiology.

1.6.2 The bacterial riboregulator Hfq

Hfq was discovered as a bacterial host factor required for replication of the RNA phage Qβ (163). Disruptions of the *hfq* gene in many bacterial species create pleiotropic effects signifying the impact of Hfq on several cellular processes such as virulence and stress response (164). As an RNA-binding protein Hfq is a key riboregulator in bacteria that is intimately involved in post-transcriptional regulation of many genes (162, 165).

Hfq is a small (11 kDa) abundant (30 000-60 000 molecules per cell) nucleic-acid binding protein. Phylogenetic analysis has found Hfq orthologues in approximately half of all sequenced Gram-negative and Gram-positive bacteria, and there is at least one Hfq-like protein in an archaeal species *Methanococcus jannaschii* (166, 167). Hfq architecture is reminiscent of eukaryotic Sm-like proteins involved in RNA splicing (Figure 1.6A). High-resolution structures of Hfq suggest that it acts as a doughnut-shaped hexamer (at least in many species) wherein the monomer folds autonomously into a domain of a bent five-stranded antiparallel β-sheet and an N-terminal α-helix (168-171). Despite variability in sequence the monomeric fold is conserved from bacteria to archaea. RNA binds the doughnut-shape in *E. coli* not through the pore, but to either or both of the two faces termed the proximal and distal faces (Figure 1.6A, III and IV). The structure of Hfq from *Staphylococcus aureus* indicates that the proximal face consists of six near identical binding pockets for U or A nucleotides however not all pockets are predicted to be bound by the same RNA and other residues are implicated in RNA binding in *E. coli* Hfq (168).

In addition, the RNA is bound in a circular unwound manner around the pore. The distal face is known to bind poly(A) sequences or ARN/ARNN tracts (R is either purine residue) using a network of six tripartite binding motifs to recognize specific RNA bases.

Hfq uses these binding faces to interact with RNA and proteins (e.g. RnaseE and polyadenylate polymerase I) to mediate many post-transcriptional effects (summarized in Figure 1.6B). Examples often involve Hfq enhancing the non-perfect pairing of an sRNA to its mRNA target(s) through three possible avenues: 1) by increasing the on rate for an sRNA annealing to a target (this may be accomplished by bringing the two RNA molecules into proximity on one Hfq molecule); 2) by altering the structure of the RNA molecule such that new single-stranded regions become exposed for pairing; or 3) by increasing the stability of the sRNA-mRNA pair even after dissociation of Hfq (162). For example, Hfq binds the 5' leader region of *rpoS* mRNA and the sRNA DsrA in order to accelerate the pairing of a downstream *rpoS* transcript region with DsrA to stimulate translation of the RpoS stress response factor (σS) in *E. coli* by opening a stem-loop that represses translation initiation (172). This type of pairing between an sRNA and mRNA is well-documented for Hfq, however the consequences vary. For example, the interaction between the sRNA MicC and the mRNA *ompC* (outer membrane protein C) requires Hfq to repress translation initiation; the sRNA in conjunction with Hfq inhibits ribosome binding *in vitro* (173). Hfq can also affect stability of RNA species. The sRNA RyhB is induced when iron is limiting and Hfq is required for pairing to its target mRNAs (proteins involved in iron metabolism and storage) where the transcripts and the sRNA are consequently degraded through an RnaseE-dependent mechanism (174).

1.7 Topics addressed in this thesis

Despite the wealth of data concerning the chemical steps of excision and joining in bacterial transposition and other related systems, little is known about target site

determinants and how these regulate the final step of transposition. Chapter 2 describes the experiments performed to isolate a Tn5 transpososome bound to target DNA. I determined that target DNA bending is an important characteristic of Tn5 insertion. Since H-NS, a protein that prefers to bind bent or distorted DNA, was being used simultaneously in other experiments in our lab, I used it to probe for these structures in the strand transfer complex --- H-NS binding would support the hypothesis that DNA is distorted in the strand transfer complex. After discovering that H-NS had an affinity for the complex, I questioned if H-NS binding to the transpososome impacts Tn5 transposition similarly to the Tn10 system, that is, by positively affecting the strand transfer reaction. The work presented in Chapters 3 and 4 provide evidence for a previously unknown regulatory mechanism of Tn5 by the protein H-NS, which is mediated by direct contact between H-NS and the Tn5 transpososome. In Chapter 5 another previously unknown mechanism of host regulation in Tn5 transposition is presented. I provide evidence that the RNA chaperone Hfq represses transposase expression and down-regulates transposition. Chapter 6, as a general discussion, unites the previous chapters and addresses the broader impact of this work, and future research goals.

1.8 References

1. Ochman,H., Lawrence,J.G. and Groisman,E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature*, **405**, 299-304.
2. Lawrence,J.G. and Ochman,H. (1998) Molecular archaeology of the Escherichia coli genome. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 9413-7.
3. Levy,S.B. and Marshall,B. (2004) Antibacterial resistance worldwide: causes, challenges and responses. *Nature medicine*, **10**, S122-9.
4. Khan,A.A. and Shrivastava,A. (2010) Bacterial infections associated with cancer: possible implication in etiology with special reference to lateral gene transfer. *Cancer metastasis reviews*, **29**, 331-7.
5. Thomas,C.M. and Nielsen,K.M. (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature reviews. Microbiology*, **3**, 711-21.
6. Frost,L.S., Leplae,R., Summers,A.O. and Toussaint,A. (2005) Mobile genetic elements: the agents of open source evolution. *Nature reviews. Microbiology*, **3**, 722-32.
7. Hendrix,R.W. (2003) Bacteriophage genomics. *Current opinion in microbiology*, **6**, 506-11.
8. Gentry,L.O. (1991) Bacterial resistance. *The Orthopedic clinics of North America*, **22**, 379-88.
9. Lupski,J.R. (1987) Molecular mechanisms for transposition of drug-resistance genes and other movable genetic elements. *Reviews of infectious diseases*, **9**, 357-68.
10. Kleckner,N. (1981) Transposable elements in prokaryotes. *Annual review of genetics*, **15**, 341-404.
11. Reanney,D. (1976) Extrachromosomal elements as possible agents of adaptation and development. *Bacteriological reviews*, **40**, 552-90.
12. Datta,N., Dacey,S., Hughes,V., Knight,S., Richards,H., Williams,G., Casewell,M. and Shannon,K.P. (1980) Distribution of genes for trimethoprim and gentamicin resistance in bacteria and their plasmids in a general hospital. *Journal of general microbiology*, **118**, 495-508.
13. Emerson,R.A. (1914) Multiple factors vs. “Golden Mean” in size inheritance. *Science*, **40**, 57-8.
14. McClintock,B. (1950) The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences of the United States of America*, **36**, 344-55.
15. McClintock,B. (1956) Controlling elements and the gene. *Cold Spring Harbor symposia on quantitative biology*, **21**, 197-216.
16. Taylor,A.L. (1963) Bacteriophage-induced mutation in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, **50**, 1043-51.

17. Galas,D. and Chandler,M. (1989) Bacterial Insertion Sequences. In Berg,D., How,M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, D.C., pp. 109-162.
18. Scott,J.R. and Churchward,G.G. (1995) Conjugative transposition. *Annual review of microbiology*, **49**, 367-97.
19. Eickbush,T. and Malik,H. (2002) Origins and Evolution of Retrotransposons. In Craig,N.L., Craigie,R., Gellert,M., Lambowitz,A.M. (eds), *Mobile DNA II*. ASM Press, pp. 1111-1139.
20. Engels,W.R. (1983) The P family of transposable elements in Drosophila. *Annual review of genetics*, **17**, 315-44.
21. Bryan,G.J., Jacobson,J.W. and Hartl,D.L. (1987) Heritable somatic excision of a Drosophila transposon. *Science*, **235**, 1636-8.
22. Kramerov,D.A. and Vassetzky,N.S. (2005) Short retroposons in eukaryotic genomes. *International review of cytology*, **247**, 165-221.
23. Hormozdiari,F., Alkan,C., Ventura,M., Hajirasouliha,I., Malig,M., Hach,F., Yorukoglu,D., Dao,P., Bakhshi,M., Sahinalp,S.C., et al. (2011) Alu repeat discovery and characterization within human genomes. *Genome research*, **21**, 840-9.
24. Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W., et al. (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.
25. Terreros,M.C., Alfonso-Sánchez,M.A., Novick,G.E., Luis,J.R., Lacau,H., Lowery,R.K., Regueiro,M. and Herrera,R.J. (2009) Insights on human evolution: an analysis of Alu insertion polymorphisms. *Journal of human genetics*, **54**, 603-11.
26. Kazazian,H.H. (1998) Mobile elements and disease. *Current opinion in genetics & development*, **8**, 343-50.
27. Yang,Z., Boffelli,D., Boonmark,N., Schwartz,K. and Lawn,R. (1998) Apolipoprotein(a) gene enhancer resides within a LINE element. *The Journal of biological chemistry*, **273**, 891-7.
28. Shewchuk,B.M., Cooke,N.E. and Liebhaber,S.A. (2001) The human growth hormone locus control region mediates long-distance transcriptional activation independent of nuclear matrix attachment regions. *Nucleic acids research*, **29**, 3356-61.
29. Hübner, a and Hendrickson,W. (1997) A fusion promoter created by a new insertion sequence, IS1490, activates transcription of 2,4,5-trichlorophenoxyacetic acid catabolic genes in Burkholderia cepacia AC1100. *Journal of bacteriology*, **179**, 2717-23.
30. Reynolds,A.E., Felton,J. and Wright,A. (1981) Insertion of DNA activates the cryptic bgl operon in E. coli K12. *Nature*, **293**, 625-9.
31. Schnetz,K. and Rak,B. (1992) IS5: a mobile enhancer of transcription in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 1244-8.

32. Manna,D., Wang,X. and Higgins,N.P. (2001) Mu and IS1 transpositions exhibit strong orientation bias at the Escherichia coli bgl locus. *Journal of bacteriology*, **183**, 3328-35.
33. Hackett,P.B., Largaespada,D.A. and Cooper,L.J.N. (2010) A transposon and transposase system for human application. *Molecular therapy : the journal of the American Society of Gene Therapy*, **18**, 674-83.
34. Ivics,Z. and Izsvák,Z. (2011) Non-viral Gene Delivery with the Sleeping Beauty Transposon System. *Human gene therapy*, **22**, 1043-51.
35. Hyland,K.A., Olson,E.R., Clark,K.J., Aronovich,E.L., Hackett,P.B., Blazar,B.R., Tolar,J. and Scott McIvor,R. (2011) Sleeping Beauty-mediated correction of Fanconi anemia type C. *The journal of gene medicine*, **13**, 462-9.
36. Aronovich,E.L., McIvor,R.S. and Hackett,P.B. (2011) The Sleeping Beauty transposon system: a non-viral vector for gene therapy. *Human molecular genetics*, **20**, R14-20.
37. Liu,L., Mah,C. and Fletcher,B.S. (2006) Sustained FVIII expression and phenotypic correction of hemophilia A in neonatal mice using an endothelial-targeted sleeping beauty transposon. *Molecular therapy : the journal of the American Society of Gene Therapy*, **13**, 1006-15.
38. Medeiros,A.A. (1997) Evolution and dissemination of beta-lactamases accelerated by generations of beta-lactam antibiotics. *Clin Infect Dis*, **24 (S1)**, S19-45.
39. Watanabe,T. (1963) Infectious heredity of multiple drug resistance in bacteria. *Bacteriological reviews*, **27**, 87.
40. Carias,L.L., Rudin,S.D., Donskey,C.J. and Rice,L.B. (1998) Genetic linkage and cotransfer of a novel, vanB-containing transposon (Tn5382) and a low-affinity penicillin-binding protein 5 gene in a clinical vancomycin-resistant Enterococcus faecium isolate. *Journal of bacteriology*, **180**, 4426-34.
41. Petrovski,S., Blackmore,D.W., Jackson,K.L. and Stanisich,V.A. (2011) Mercury(II)-resistance transposons Tn502 and Tn512, from Pseudomonas clinical strains, are structurally different members of the Tn5053 family. *Plasmid*, **65**, 58-64.
42. Warburton,P.J., Palmer,R.M., Munson,M.A. and Wade,W.G. (2007) Demonstration of in vivo transfer of doxycycline resistance mediated by a novel transposon. *The Journal of antimicrobial chemotherapy*, **60**, 973-80.
43. Nagy,Z. and Chandler,M. (2004) Regulation of transposition in bacteria. *Research in microbiology*, **155**, 387-98.
44. Kleckner,N. (1990) Regulation of transposition in bacteria. *Annual review of cell biology*, **6**, 297-327.
45. Gueguen,E., Rousseau,P., Duval-Valentin,G. and Chandler,M. (2005) The transpososome: control of transposition at the level of catalysis. *Trends in microbiology*, **13**, 543-9.

46. Surette,M.G., Lavoie,B.D. and Chaconas,G. (1989) Action at a distance in Mu DNA transposition: an enhancer-like element is the site of action of supercoiling relief activity by integration host factor (IHF). *The EMBO journal*, **8**, 3483-9.
47. Signon,L. and Kleckner,N. (1995) Negative and positive regulation of Tn10/IS10-promoted recombination by IHF: two distinguishable processes inhibit transposition off of multicopy plasmid replicons and activate chromosomal events that favor evolution of new transposons. *Genes & development*, **9**, 1123-36.
48. Chalmers,R., Guhathakurta,A., Benjamin,H. and Kleckner,N. (1998) IHF modulation of Tn10 transposition: sensory transduction of supercoiling status via a proposed protein/DNA molecular spring. *Cell*, **93**, 897-908.
49. Crellin,P., Sewitz,S. and Chalmers,R. (2004) DNA looping and catalysis; the IHF-folded arm of Tn10 promotes conformational changes and hairpin resolution. *Molecular cell*, **13**, 537-47.
50. Wardle,S.J., O'Carroll,M., Derbyshire,K.M. and Haniford,D.B. (2005) The global regulator H-NS acts directly on the transpososome to promote Tn10 transposition. *Genes & development*, **19**, 2224-35.
51. Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2009) The global bacterial regulator H-NS promotes transpososome formation and transposition in the Tn5 system. *Nucleic acids research*, **37**, 309-21.
52. Lavoie,B.D. and Chaconas,G. (1993) Site-specific HU binding in the Mu transpososome: conversion of a sequence-independent DNA-binding protein into a chemical nuclease. *Genes & development*, **7**, 2510-9.
53. Lavoie,B.D. and Chaconas,G. (1994) A second high affinity HU binding site in the phage Mu transpososome. *The Journal of biological chemistry*, **269**, 15571-6.
54. Craigie,R. and Mizuuchi,K. (1985) Mechanism of transposition of bacteriophage Mu: structure of a transposition intermediate. *Cell*, **41**, 867-76.
55. Morisato,D. and Kleckner,N. (1987) Tn10 transposition and circle formation in vitro. *Cell*, **51**, 101-11.
56. Shiga,Y., Sekine,Y., Kano,Y. and Ohtsubo,E. (2001) Involvement of H-NS in transpositional recombination mediated by IS1. *Journal of bacteriology*, **183**, 2476-84.
57. Rouquette,C., Serre,M.-C. and Lane,D. (2004) Protective role for H-NS protein in IS1 transposition. *Journal of bacteriology*, **186**, 2091-8.
58. Falconi,M., McGovern,V., Gualerzi,C., Hillyard,D. and Higgins,N.P. (1991) Mutations altering chromosomal protein H-NS induce mini-Mu transposition. *The New biologist*, **3**, 615-25.
59. Bétermier,M., Poquet,I., Alazard,R. and Chandler,M. (1993) Involvement of Escherichia coli FIS protein in maintenance of bacteriophage mu lysogeny by the repressor: control of early transcription and inhibition of transposition. *Journal of bacteriology*, **175**, 3798-811.

60. Drunen,C.M. van, Zuylen,C. van, Mientjes,E.J., Goosen,N. and Putte,P. van de (1993) Inhibition of bacteriophage Mu transposition by Mu repressor and Fis. *Molecular microbiology*, **10**, 293-8.
61. Pato,M.L. and Banerjee,M. (1996) The Mu strong gyrase-binding site promotes efficient synapsis of the prophage termini. *Molecular microbiology*, **22**, 283-92.
62. Yin,J.C. and Reznikoff,W.S. (1987) dnaA, an essential host gene, and Tn5 transposition. *Journal of bacteriology*, **169**, 4637-45.
63. Levchenko,I., Luo,L. and Baker,T.A. (1995) Disassembly of the Mu transposase tetramer by the ClpX chaperone. *Genes & development*, **9**, 2399-408.
64. Sharpe,P.L. and Craig,N.L. (1998) Host proteins can stimulate Tn7 transposition: a novel role for the ribosomal protein L29 and the acyl carrier protein. *The EMBO journal*, **17**, 5822-31.
65. Ross,J.A., Wardle,S.J. and Haniford,D.B. (2010) Tn10/IS10 transposition is downregulated at the level of transposase expression by the RNA-binding protein Hfq. *Mol Micro*, **78**, 607-21.
66. Amit,R., Oppenheim,A.B. and Stavans,J. (2003) Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophysical journal*, **84**, 2467-73.
67. Stella,S., Falconi,M., Lammi,M., Gualerzi,C.O. and Pon,C.L. (2006) Environmental control of the in vivo oligomerization of nucleoid protein H-NS. *Journal of molecular biology*, **355**, 169-74.
68. Dorman,C.J. (2004) H-NS: a universal regulator for a dynamic genome. *Nature reviews. Microbiology*, **2**, 391-400.
69. Vanderpool,C.K. (2007) Physiological consequences of small RNA-mediated regulation of glucose-phosphate stress. *Current opinion in microbiology*, **10**, 146-51.
70. Massé,E. and Gottesman,S. (2002) A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 4620-5.
71. Vytvytska,O., Jakobsen,J.S., Balcunaite,G., Andersen,J.S., Baccarini,M. and Gabain, a von (1998) Host factor I, Hfq, binds to Escherichia coli ompA mRNA in a growth rate-dependent fashion and regulates its stability. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 14118-23.
72. Capy,P., Gasperi,G., Biémont,C. and Bazin,C. (2000) Stress and transposable elements: co-evolution or useful parasites? *Heredity*, **85 (Pt 2)**, 101-6.
73. Labrador,M. and Corces,V.G. (1997) Transposable element-host interactions: regulation of insertion and excision. *Annual review of genetics*, **31**, 381-404.
74. Federoff,N. (2002) Control of mobile DNA. In Nancy L. Craig, Robert Craigie, Martin Gellert,A.M.L. (ed), *Mobile DNA II*. ASM Press, Washington, D.C., pp. 997-1001.

75. Twiss,E., Coros,A.M., Tavakoli,N.P. and Derbyshire,K.M. (2005) Transposition is modulated by a diverse set of host factors in Escherichia coli and is stimulated by nutritional stress. *Molecular microbiology*, **57**, 1593-607.
76. Berg,D.E., Davies,J., Allet,B. and Rochaix,J.D. (1975) Transposition of R factor genes to bacteriophage lambda. *Proceedings of the National Academy of Sciences of the United States of America*, **72**, 3628-32.
77. Reznikoff,W.S. (2003) Tn5 as a model for understanding DNA transposition. *Molecular microbiology*, **47**, 1199–1206.
78. Goryshin,I.Y. and Reznikoff,W.S. (1998) Tn5 in vitro transposition. *The Journal of biological chemistry*, **273**, 7367-74.
79. Steiniger-White,M., Bhasin,A., Lovell,S., Rayment,I. and Reznikoff,W.S. (2002) Evidence for "Unseen" Transposase–DNA Contacts. *Journal of Molecular Biology*, **322**, 971-982.
80. Davies,D.R., Goryshin,I.Y., Reznikoff,W.S. and Rayment,I. (2000) Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science*, **289**, 77-85.
81. Klenchin,V.A., Czyz,A., Goryshin,I.Y., Gradman,R., Lovell,S., Rayment,I. and Reznikoff,W.S. (2008) Phosphate coordination and movement of DNA in the Tn5 synaptic complex: role of the (R)YREK motif. *Nucleic acids research*, **36**, 5855-62.
82. Rothstein,S.J., Jorgensen,R.A., Yin,J.C., Yong-di,Z., Johnson,R.C. and Reznikoff,W.S. (1981) Genetic organization of Tn5. *Cold Spring Harbor symposia on quantitative biology*, **45 (Pt 1)**, 99-105.
83. Rothstein,S.J. and Reznikoff,W.S. (1981) The functional differences in the inverted repeats of Tn5 are caused by a single base pair nonhomology. *Cell*, **23**, 191-9.
84. Bhasin,A., Goryshin,I.Y., Steiniger-White,M., York,D. and Reznikoff,W.S. (2000) Characterization of a Tn5 pre-cleavage synaptic complex. *Journal of molecular biology*, **302**, 49-63.
85. Isberg,R.R., Lazaar,A.L. and Syvanen,M. (1982) Regulation of Tn5 by the right-repeat proteins: control at the level of the transposition reaction? *Cell*, **30**, 883-92.
86. Cruz,N.B. de la, Weinreich,M.D., Wiegand,T.W., Krebs,M.P. and Reznikoff,W.S. (1993) Characterization of the Tn5 transposase and inhibitor proteins: a model for the inhibition of transposition. *Journal of bacteriology*, **175**, 6932-8.
87. Mizuuchi,K. (1992) Transpositional recombination: mechanistic insights from studies of mu and other elements. *Annual review of biochemistry*, **61**, 1011-51.
88. Mizuuchi,K. (1997) Polynucleotidyl transfer reactions in site-specific DNA recombination. *Genes to cells : devoted to molecular & cellular mechanisms*, **2**, 1-12.
89. Peterson,G. and Reznikoff,W.S. (2003) Tn5 transposase active site mutations suggest position of donor backbone DNA in synaptic complex. *The Journal of biological chemistry*, **278**, 1904-9.

90. Steiniger-White,M. and Reznikoff,W.S. (2000) The C-terminal alpha helix of Tn5 transposase is required for synaptic complex formation. *The Journal of biological chemistry*, **275**, 23127-33.
91. Vaezeslami,S., Sterling,R. and Reznikoff,W.S. (2007) Site-directed mutagenesis studies of tn5 transposase residues involved in synaptic complex formation. *Journal of bacteriology*, **189**, 7436-41.
92. Weinreich,M.D., Gasch,A. and Reznikoff,W.S. (1994) Evidence that the cis preference of the Tn5 transposase is caused by nonproductive multimerization. *Genes & development*, **8**, 2363-74.
93. Steiniger,M., Adams,C.D., Marko,J.F. and Reznikoff,W.S. (2006) Defining characteristics of Tn5 Transposase non-specific DNA binding. *Nucleic acids research*, **34**, 2820-32.
94. Adams,C.D., Schnurr,B., Skoko,D., Marko,J.F. and Reznikoff,W.S. (2006) Tn5 transposase loops DNA in the absence of Tn5 transposon end sequences. *Molecular microbiology*, **62**, 1558-68.
95. Weinreich,M.D., Mahnke-Braam,L. and Reznikoff,W.S. (1994) A functional analysis of the Tn5 transposase. Identification of domains required for DNA binding and multimerization. *Journal of molecular biology*, **241**, 166-77.
96. York,D. and Reznikoff,W.S. (1996) Purification and biochemical analyses of a monomeric form of Tn5 transposase. *Nucleic acids research*, **24**, 3790-6.
97. Phadnis,S.H. and Berg,D.E. (1987) Identification of base pairs in the outside end of insertion sequence IS50 that are needed for IS50 and Tn5 transposition. *Proceedings of the National Academy of Sciences of the United States of America*, **84**, 9118-22.
98. Makris,J.C., Nordmann,P.L. and Reznikoff,W.S. (1988) Mutational analysis of insertion sequence 50 (IS50) and transposon 5 (Tn5) ends. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 2224-8.
99. Aldaz,H., Schuster,E. and Baker,T.A. (1996) The interwoven architecture of the Mu transposase couples DNA synapsis to catalysis. *Cell*, **85**, 257-69.
100. Savilahti,H., Rice,P.A. and Mizuuchi,K. (1995) The phage Mu transpososome core: DNA requirements for assembly and function. *The EMBO journal*, **14**, 4893-903.
101. Naumann,T.A. and Reznikoff,W.S. (2000) Trans catalysis in Tn5 transposition. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 8944-9.
102. Steiniger,M., Metzler,J. and Reznikoff,W.S. (2006) Mutation of Tn5 transposase beta-loop residues affects all steps of Tn5 transposition: the role of conformational changes in Tn5 transposition. *Biochemistry*, **45**, 15552-62.
103. Benjamin,H.W. and Kleckner,N. (1989) Intramolecular transposition by Tn10. *Cell*, **59**, 373-83.
104. Bender,J. and Kleckner,N. (1986) Genetic evidence that Tn10 transposes by a nonreplicative mechanism. *Cell*, **45**, 801-15.

105. Bainton,R., Gamas,P. and Craig,N.L. (1991) Tn7 transposition in vitro proceeds through an excised transposon intermediate generated by staggered breaks in DNA. *Cell*, **65**, 805-16.
106. Hagemann,A.T. and Craig,N.L. (1993) Tn7 transposition creates a hotspot for homologous recombination at the transposon donor site. *Genetics*, **133**, 9-16.
107. Engels,W.R., Johnson-Schlitz,D.M., Eggleston,W.B. and Sved,J. (1990) High-frequency P element loss in Drosophila is homolog dependent. *Cell*, **62**, 515-25.
108. Kaufman,P.D. and Rio,D.C. (1992) P element transposition in vitro proceeds by a cut-and-paste mechanism and uses GTP as a cofactor. *Cell*, **69**, 27-39.
109. Vos,J.C., Baere,I. De and Plasterk,R.H. (1996) Transposase is the only nematode protein required for in vitro transposition of Tc1. *Genes & development*, **10**, 755-61.
110. Plasterk,R.H. and Groenen,J.T. (1992) Targeted alterations of the Caenorhabditis elegans genome by transgene instructed DNA double strand break repair following Tc1 excision. *The EMBO journal*, **11**, 287-90.
111. Reznikoff,W.S., Goryshin,I.Y. and Jendrisak,J.J. (2004) Tn5 as a molecular genetics tool: In vitro transposition and the coupling of in vitro technologies with in vivo transposition. *Methods Mol Biol*, **260**, 83-96.
112. Reznikoff,W.S., Bhasin,A., Davies,D.R., Goryshin,I.Y., Mahnke,L. a, Naumann,T., Rayment,I., Steiniger-White,M. and Twining,S.S. (1999) Tn5: A molecular window on transposition. *Biochemical and biophysical research communications*, **266**, 729-34.
113. Kennedy,A.K., Haniford,D.B. and Mizuuchi,K. (2000) Single active site catalysis of the successive phosphoryl transfer steps by DNA transposases: insights from phosphorothioate stereoselectivity. *Cell*, **101**, 295-305.
114. Bhasin,A., Goryshin,I.Y. and Reznikoff,W.S. (1999) Hairpin formation in Tn5 transposition. *The Journal of biological chemistry*, **274**, 37021-9.
115. Lovell,S., Goryshin,I.Y., Reznikoff,W.R. and Rayment,I. (2002) Two-metal active site binding of a Tn5 transposase synaptic complex. *Nature structural biology*, **9**, 278-81.
116. Craig,N.L. (1995) Unity in transposition reactions. *Science*, **270**, 253-4.
117. Kennedy,A.K., Guhathakurta,A., Kleckner,N. and Haniford,D.B. (1998) Tn10 transposition via a DNA hairpin intermediate. *Cell*, **95**, 125-34.
118. McBlane,J.F., Gent,D.C. van, Ramsden,D.A., Romeo,C., Cuomo,C.A., Gellert,M. and Oettinger,M.A. (1995) Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell*, **83**, 387-95.
119. Ason,B. and Reznikoff,W.S. (2002) Mutational analysis of the base flipping event found in Tn5 transposition. *The Journal of biological chemistry*, **277**, 11284-91.
120. Bischerour,J. and Chalmers,R. (2009) Base flipping in tn10 transposition: an active flip and capture mechanism. *PloS one*, **4**, e6201.

121. Bischerour,J., Lu,C., Roth,D.B. and Chalmers,R. (2009) Base flipping in V(D)J recombination: insights into the mechanism of hairpin formation, the 12/23 rule, and the coordination of double-strand breaks. *Molecular and cellular biology*, **29**, 5889-99.
122. Bischerour,J. and Chalmers,R. (2007) Base-flipping dynamics in a DNA hairpin processing reaction. *Nucleic acids research*, **35**, 2584-95.
123. Lodge,J.K., Weston-Hafer,K. and Berg,D.E. (1988) Transposon Tn5 target specificity: preference for insertion at G/C pairs. *Genetics*, **120**, 645-50.
124. Craig,N.L. (1997) Target site selection in transposition. *Annual review of biochemistry*, **66**, 437-74.
125. Parks,A.R., Li,Z., Shi,Q., Owens,R.M., Jin,M.M. and Peters,J.E. (2009) Transposition into replicating DNA occurs through interaction with the processivity factor. *Cell*, **138**, 685-95.
126. Kleckner,N., Steele,D.A., Reichardt,K. and Botstein,D. (1979) Specificity of insertion by the translocatable tetracycline-resistance element Tn10. *Genetics*, **92**, 1023-40.
127. Halling,S.M. and Kleckner,N. (1982) A symmetrical six-base-pair target site sequence determines Tn10 insertion specificity. *Cell*, **28**, 155-63.
128. Berg,D.E., Schmandt,M.A. and Lowe,J.B. (1983) Specificity of transposon Tn5 insertion. *Genetics*, **105**, 813-28.
129. Lodge,J.K. and Berg,D.E. (1990) Mutations that affect Tn5 insertion into pBR322: importance of local DNA supercoiling. *Journal of bacteriology*, **172**, 5956-60.
130. Goryshin,I.Y., Miller,J. a, Kil,Y.V., Lanzov,V. a and Reznikoff,W.S. (1998) Tn5/IS50 target recognition. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 10716-21.
131. Kuan,C.T. and Tessman,I. (1991) LexA protein of Escherichia coli represses expression of the Tn5 transposase gene. *Journal of bacteriology*, **173**, 6406-10.
132. Yigit,H. and Reznikoff,W.S. (1998) Escherichia coli DNA topoisomerase I and suppression of killing by Tn5 transposase overproduction: topoisomerase I modulates Tn5 transposition. *Journal of bacteriology*, **180**, 5866-74.
133. Weinreich,M.D. and Reznikoff,W.S. (1992) Fis plays a role in Tn5 and IS50 transposition. *Journal of bacteriology*, **174**, 4530-7.
134. Makris,J.C., Nordmann,P.L. and Reznikoff,W.S. (1990) Integration host factor plays a role in IS50 and Tn5 transposition. *Journal of bacteriology*, **172**, 1368-73.
135. Dorman,C.J. (2007) H-NS, the genome sentinel. *Nature reviews. Microbiology*, **5**, 157-61.
136. Arold,S.T., Leonard,P.G., Parkinson,G.N. and Ladbury,J.E. (2010) H-NS forms a superhelical protein scaffold for DNA condensation. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 15728-32.

137. Gordon,B.R.G., Li,Y., Cote,A., Weirauch,M.T., Ding,P., Hughes,T.R., Navarre,W.W., Xia,B. and Liu,J. (2011) Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 10690-5.
138. Dorman,C.J. and Kane,K.A. (2009) DNA bridging and antibridging: a role for bacterial nucleoid-associated proteins in regulating the expression of laterally acquired genes. *FEMS microbiology reviews*, **33**, 587-92.
139. Spassky,A., Rimsky,S., Garreau,H. and Buc,H. (1984) H1a, an E. coli DNA-binding protein which accumulates in stationary phase, strongly compacts DNA in vitro. *Nucleic acids research*, **12**, 5321-40.
140. Stella,S., Spurio,R., Falconi,M., Pon,C.L. and Gualerzi,C.O. (2005) Nature and mechanism of the in vivo oligomerization of nucleoid protein H-NS. *The EMBO journal*, **24**, 2896-905.
141. Atlung,T. and Ingmer,H. (1997) H-NS: a modulator of environmentally regulated gene expression. *Molecular microbiology*, **24**, 7-17.
142. Falconi,M., Gualtieri,M.T., Teana,A. La, Losso,M.A. and Pon,C.L. (1988) Proteins from the prokaryotic nucleoid: primary and quaternary structure of the 15-kD Escherichia coli DNA binding protein H-NS. *Molecular microbiology*, **2**, 323-9.
143. Ceschini,S., Lupidi,G., Coletta,M., Pon,C.L., Fioretti,E. and Angeletti,M. (2000) Multimeric self-assembly equilibria involving the histone-like protein H-NS. A thermodynamic study. *The Journal of biological chemistry*, **275**, 729-34.
144. Smyth,C.P., Lundbäck,T., Renzoni,D., Siligardi,G., Beavil,R., Layton,M., Sidebotham,J.M., Hinton,J.C., Driscoll,P.C., Higgins,C.F., et al. (2000) Oligomerization of the chromatin-structuring protein H-NS. *Molecular microbiology*, **36**, 962-72.
145. Spurio,R., Falconi,M., Brandi,A., Pon,C.L. and Gualerzi,C.O. (1997) The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *The EMBO journal*, **16**, 1795-805.
146. Yamada,H., Yoshida,T., Tanaka,K., Sasakawa,C. and Mizuno,T. (1991) Molecular analysis of the Escherichia coli hns gene encoding a DNA-binding protein, which preferentially recognizes curved DNA sequences. *Mol Gen Genet*, **230**, 332-6.
147. Owen-Hughes,T.A., Pavitt,G.D., Santos,D.S., Sidebotham,J.M., Hulton,C.S., Hinton,J.C. and Higgins,C.F. (1992) The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell*, **71**, 255-65.
148. Lucht,J.M., Dersch,P., Kempf,B. and Bremer,E. (1994) Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled proU operon of Escherichia coli. *The Journal of biological chemistry*, **269**, 6578-8.

149. Tupper,A.E., Owen-Hughes,T.A., Ussery,D.W., Santos,D.S., Ferguson,D.J., Sidebotham,J.M., Hinton,J.C. and Higgins,C.F. (1994) The chromatin-associated protein H-NS alters DNA topology in vitro. *The EMBO journal*, **13**, 258-68.
150. Bouffartigues,E., Buckle,M., Badaut,C., Travers,A. and Rimsky,S. (2007) H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nature structural & molecular biology*, **14**, 441-8.
151. Hommais,F., Krin,E., Laurent-Winter,C., Soutourina,O., Malpertuy,A., Caer,J.P. Le, Danchin,A. and Bertin,P. (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Molecular microbiology*, **40**, 20-36.
152. Rimsky,S., Zuber,F., Buckle,M. and Buc,H. (2001) A molecular mechanism for the repression of transcription by the H-NS protein. *Molecular microbiology*, **42**, 1311-23.
153. Dole,S., Nagarajavel,V. and Schnetz,K. (2004) The histone-like nucleoid structuring protein H-NS represses the Escherichia coli bgl operon downstream of the promoter. *Molecular microbiology*, **52**, 589-600.
154. Nagarajavel,V., Madhusudan,S., Dole,S., Rahmouni,A.R. and Schnetz,K. (2007) Repression by binding of H-NS within the transcription unit. *The Journal of biological chemistry*, **282**, 23622-30.
155. Navarre,W.W., McClelland,M., Libby,S.J. and Fang,F.C. (2007) Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes & development*, **21**, 1456-71.
156. Lucchini,S., Rowley,G., Goldberg,M.D., Hurd,D., Harrison,M. and Hinton,J.C.D. (2006) H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS pathogens*, **2**, e81.
157. Corcoran,C.P. and Dorman,C.J. (2009) DNA relaxation-dependent phase biasing of the fim genetic switch in Escherichia coli depends on the interplay of H-NS, IHF and LRP. *Molecular microbiology*, **74**, 1071-82.
158. Liu,D., Haniford,D.B. and Chalmers,R.M. (2011) H-NS mediates the dissociation of a refractory protein-DNA complex during Tn10/IS10 transposition. *Nucleic acids research*, **39**, 6660-8.
159. Swingle,B., O'Carroll,M., Haniford,D. and Derbyshire,K.M. (2004) The effect of host-encoded nucleoid proteins on transposition: H-NS influences targeting of both IS903 and Tn10. *Molecular microbiology*, **52**, 1055-67.
160. Wardle,S.J., Chan,A. and Haniford,D.B. (2009) H-NS binds with high affinity to the Tn10 transpososome and promotes transpososome stabilization. *Nucleic acids research*, **37**, 6148-60.
161. Donato,G.M. and Kawula,T.H. (1998) Enhanced binding of altered H-NS protein to flagellar rotor protein FliG causes increased flagellar rotational speed and

hypermotility in Escherichia coli. *The Journal of biological chemistry*, **273**, 24030-6.

162. Vogel,J. and Luisi,B.F. (2011) Hfq and its constellation of RNA. *Nature reviews. Microbiology*, **9**, 578-89.
163. Franze de Fernandez,M.T., Hayward,W.S. and August,J.T. (1972) Bacterial proteins required for replication of phage Q ribonucleic acid. Pruification and properties of host factor I, a ribonucleic acid-binding protein. *The Journal of biological chemistry*, **247**, 824-31.
164. Tsui,H.C., Leung,H.C. and Winkler,M.E. (1994) Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. *Molecular microbiology*, **13**, 35-49.
165. Chao,Y. and Vogel,J. (2010) The role of Hfq in bacterial pathogens. *Current opinion in microbiology*, **13**, 24-33.
166. Brennan,R.G. and Link,T.M. (2007) Hfq structure, function and ligand binding. *Current opinion in microbiology*, **10**, 125-33.
167. Sun,X., Zhulin,I. and Wartell,R.M. (2002) Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic acids research*, **30**, 3662-71.
168. Schumacher,M.A., Pearson,R.F., Møller,T., Valentin-Hansen,P. and Brennan,R.G. (2002) Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *The EMBO journal*, **21**, 3546-56.
169. Sauter,C., Basquin,J. and Suck,D. (2003) Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from Escherichia coli. *Nucleic acids research*, **31**, 4091-8.
170. Nikulin,A., Stolboushkina,E., Perederina,A., Vassilieva,I., Blaesi,U., Moll,I., Kachalova,G., Yokoyama,S., Vassylyev,D., Garber,M., et al. (2005) Structure of Pseudomonas aeruginosa Hfq protein. *Acta crystallographica. Section D, Biological crystallography*, **61**, 141-6.
171. Link,T.M., Valentin-Hansen,P. and Brennan,R.G. (2009) Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19292-7.
172. Soper,T.J. and Woodson,S.A. (2008) The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA*, **14**, 1907-17.
173. Chen,S., Zhang,A., Blyn,L.B. and Storz,G. (2004) MicC, a second small-RNA regulator of Omp protein expression in Escherichia coli. *Journal of bacteriology*, **186**, 6689-97.
174. Massé,E., Escorcia,F.E. and Gottesman,S. (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. *Genes & development*, **17**, 2374-83.

CHAPTER 2[1]

2 FORMATION, CHARACTERIZATION AND PARTIAL PURIFICATION OF A TN5 STRAND TRANSFER COMPLEX

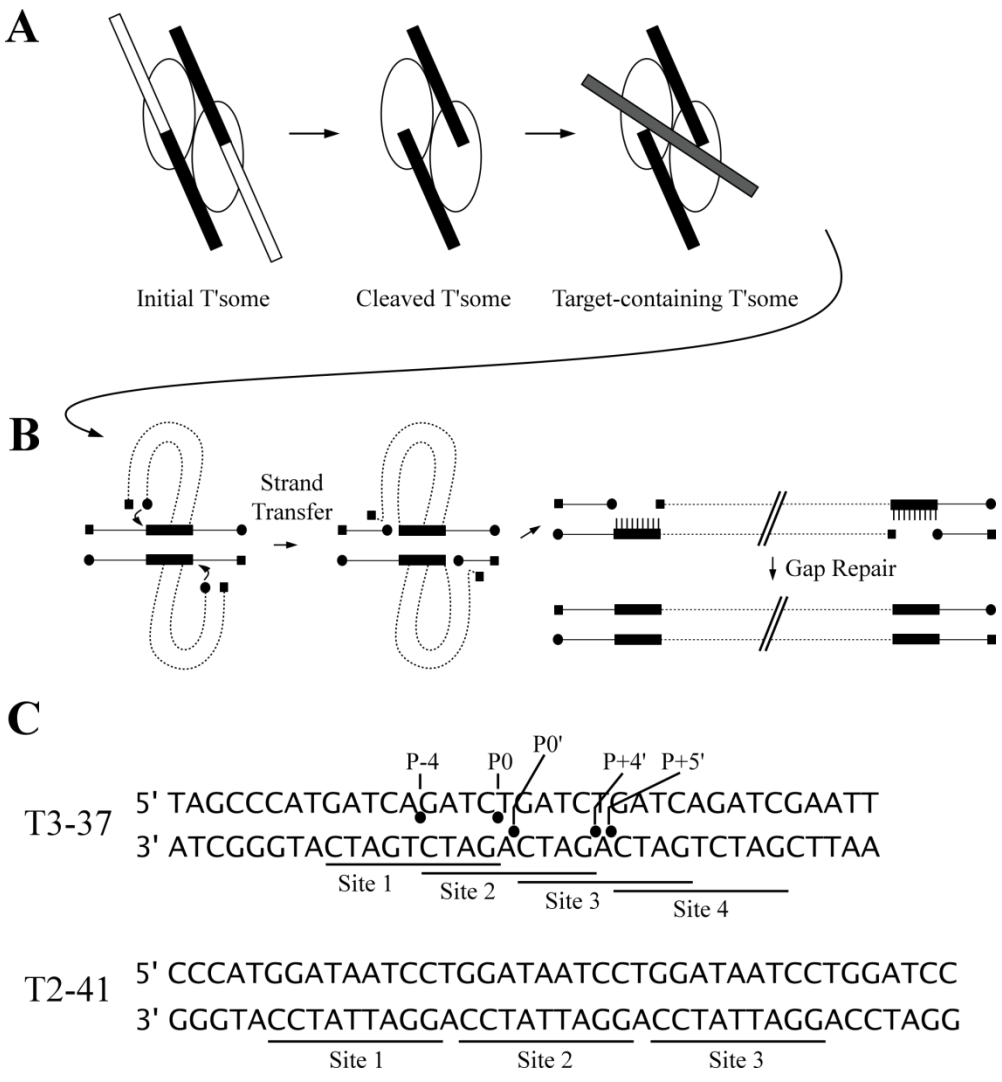
2.1 Introduction

DNA transposition reactions take place in the context of higher order protein–DNA complexes called transpososomes (1). A number of transposons form at least three distinct types of transpososomes, including a pre-cleavage complex, a cleaved complex in which the transposon ends are separated completely or partially from the flanking donor DNA, and a strand transfer complex in which the transposon ends are joined to a single, short target site (Figure 2.1A) (2). Within these transpososomes, transposon-encoded proteins, and sometimes host proteins, bind to the terminal DNA sequences (or transposon ends) of the transposon (3, 4). Through a complex network of protein-DNA and protein-protein interactions, these proteins hold the transposon ends together and direct the transpososome through the chemical steps in transposition. The transpososome is a dynamic entity, as it must coordinate the entry and exit of several different DNA strands into the active sites of its protein components and ensure that these events occur in the appropriate order. While much is known about the individual chemical steps in transposition reactions, little is known about how these steps are coordinated (5). To gain further insight into the changes in transpososome structure that drive transposition reactions forward, it is desirable to obtain high-resolution structures of different transpososomes from a given transposition pathway. Tn5 is one of the only systems that has yielded a high-resolution X-ray structure of a transpososome (6, 7).

[1]A version of this chapter has been published: Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2006) Formation, characterization and partial purification of a Tn5 strand transfer complex. *Journal of molecular biology*, **364**, 290-301.

Figure 2.1. Tn5 insertion and the sequence of two classes of target sites.

(A) During the course of Tn5 transposition the initial Tn5 transpososome is first converted to a fully cleaved transpososome and then to a transpososome that has incorporated a target DNA. Transposase, ovals; transposon end DNA, black rectangles; flanking donor DNA, open rectangles. The steps in strand transfer and gap repair are also shown, but not in the context of the transpososome. The curved arrows show 3′OH termini (black circle) of the transposon (dotted lines) being joined to a specific target site (black rectangles). The gaps produced by strand transfer are repaired by host proteins, leading to a duplication of the original target site. Squares represent terminal $5'PO_4$ groups. (B) The DNA sequence of two different classes of target DNAs used in the current work is shown. The scheme for numbering phosphate groups (black dots) is indicated for the T3-37 target DNA. The phosphate groups on either side of the centre of symmetry of site 2 are designated 0 and 0’ (the ‘ symbol indicates the bottom strand of the target DNA). Moving in the right-hand direction away from the 0 and 0′ positions, phosphate groups are given a + symbol, and moving in the left-hand direction a - symbol. The repeats of DNA sequences (5’-GATCNGATC-3’ and 5’-GGATNATCC-3’) previously shown to act as hot-spots for Tn5 insertion in the context of plasmid DNAs are underlined (8).


A
Initial T'some
Cleaved T'some
Target-containing T'some
B
Strand
Transfer
Gap Repair
C
P-4
P0
P0'
P+4'
P+5'
T3-37
5' TAGCCCATGATCAGATCTGATCTGATCAGATCGAATT
3' ATCGGGTACTAGTCTAGACTAGACTAGTCTAGCTTAA
Site 1
Site 2
Site 3
Site 4
T2-41
5' CCCATGGATAATCCTGGATAATCCTGGATAATCCTGGATCC
3' GGGTACCTATTAGGACCTATTAGGACCTATTAGGACCTAGG
Site 1
Site 2
Site 3

The transpososome structure that was solved likely represents an excision intermediate, as the transposon ends used to build the transpososome did not include flanking donor DNA. A major limitation to using Tn5 as a model for studying transpososome dynamics is that there has been no stable target DNA-containing transpososome generated. The objective of the current study was to generate such a transpososome in the Tn5 system.

Tn5 is a composite bacterial transposon composed of two copies of IS50; IS50Right encodes the transposase protein. The terminal DNA sequences of IS50, which are called the outside and inside ends (OE and IE, respectively), include the specificity determinants for transposase binding. Only the terminal 19 bp of the OEs and IEs are required for full end function. The two ends differ in sequence at seven of 19 positions, but both support transposition with wild-type (WT) transposase (9). A hybrid (or mosaic) end (ME) composed of OE and IE sequences provides the most active end sequence, and is largely used *in vitro* to increase the overall efficiency of the transposition reaction. Similarly, particular mutant forms of transposase, referred to as hyperactive transposases, have been shown to be more active than WT transposase and are thus also widely used *in vitro* (10, 11).

Tn5 transposes by a cut-and-paste mechanism (10). The transposon is first completely excised from flanking donor DNA by a pair of double-strand breaks made at each end. The excised transposon then binds and inserts into a target DNA. Host repair proteins then process the strand transfer junctions to give rise to short direct repeats of 9 bp (Figure 2.1A). Target site selection in the Tn5 system is not highly sequence-specific or DNA structure specific (8, 12, 13) and this has been a major obstacle in preparing a

homogeneous population of target DNA-containing transpososomes in this system that are suitable for biochemical and/or structural characterization.

Several studies have been carried out to establish the degree of randomness for Tn5 insertion. Preferred sites of insertion have been identified in the tetracycline resistance gene of pBR322 (14), the chloramphenicol resistance gene of pACYC184, and the kanamycin resistance gene derived from Tn903 (8). Moreover, it has been shown that certain preferred sites become even 'stronger' sites for Tn5 insertion when these sites are embedded within a cluster of overlapping sequences that resemble the initial insertion hot-spot (8). The most extensive analysis to date of Tn5 target choice has come from a study in which a Tn5 derivative was used to generate disruptions of all of the annotated open reading frames in *Escherichia coli*. A total of 1976 Tn5 insertions were characterized at the DNA sequence level and from this a consensus sequence for Tn5 insertion was developed. This analysis revealed that Tn5 preferentially recognizes a highly degenerate, asymmetric sequence located in the 5' half of the target site (13).

Here, we show that a stable Tn5 strand transfer complex (STC) can be generated *in vitro* by mixing a transpososome, generated with a pre-cleaved ME and a hyperactive transposase protein, with a short, double-stranded oligonucleotide containing tandem, overlapping repeats of a preferred Tn5 insertion sequence. The insertion specificity is surprisingly high, with approximately 70% of events going into a single site. Modifications to the target DNA that increase the above specificity, as well as the overall efficiency of target DNA utilization for strand transfer, have been identified. Finally, we show that a substantial enrichment/purification of the STC can be achieved in a single

step using polyethylene glycol precipitation, an observation that may greatly facilitate structural studies of this complex.

2.2.0 Materials and methods

2.2.1 Target DNA and transposon end substrates

The oligonucleotides listed in Table 2.1 were used for the assembly of target DNA substrates and the Tn5 MEs. Each oligonucleotide was gel-purified, and the amount of oligonucleotide recovered was quantified spectroscopically by measuring the absorbance at 260 nm. Equimolar amounts of oligonucleotide pairs were annealed under standard conditions and concentrations of duplex DNAs were estimated based on the amount of input oligonucleotides. The yield of duplex DNA was typically greater than 95% as judged by native gel electrophoresis. Pre-cleaved Tn10 OEs used in the assembly of the Tn10 transpososome were generated by digestion of pWY1005 (details are available from the authors upon request) with PvuII. Subsequently the 70 bp OE-containing fragment was isolated by HPLC. DNAs were labeled at their 5′ termini by treatment with bacteriophage T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]-ATP (Amersham Biosciences).

2.2.2 Protein purification

Tn5 transposase (EK54/LP372 and EK54/LP372/DA97) was purified as described (15) to a purity of greater than 98% based on SDS-PAGE analysis (see Figure 3.4, lane 2 in Chapter 3). Tn10 transposase was purified as described (16). IHF was purified as

described (17). Protein concentrations were determined by Bradford assay for Tn5 transposase and by BCA assay (Pierce) for Tn10 transposase and IHF, with bovine serum albumin (BSA) as a standard.

2.2.3 Transpososome assembly and strand transfer assays

For the assembly of the pre-cleaved Tn5 transpososome, 50 nM ME DNA was mixed with 511 nM Tn5 transposase protein in reaction buffer (RB) containing 25 mM Hepes (pH 7.5), 100 mM KCl, 1 mM β-mercaptoethanol, 50 μg ml^{-1} of BSA, 75 μg ml^{-1} of tRNA, as described (15). Typically, this yielded approximately 6 nM pre-cleaved transpososome. After incubation at 37°C for 1 hour, target DNA and 10 mM $MgCl_2$ or 10 mM $CaCl_2$ were added and incubation was continued for 16 hours at 25°C. Unless indicated otherwise, target DNA was present at 4 μM, which represents a 666-fold excess relative to the pre-cleaved transpososome. The Tn10 STC was formed as described (18). For native gel analysis of the above reactions, samples were mixed at a ratio of 4:1 (v/v) with non-denaturing load mix and applied to a 5% (w/v) polyacrylamide gel as described (19). For denaturing gel analysis, the above reactions were extracted with phenol, precipitated with ethanol, resuspended in water, mixed at a ratio of 1:1 (v/v) with denaturing load dye and then applied to an 8% polyacrylamide sequencing gel using 1xTBE buffer. In experiments where the STC was excised from a native gel, the DNA was eluted for 16 hours at 42°C in 5 ml of chelexed 10 mM Tris-HCl (pH 8), 1 mM EDTA. The eluate was applied to a Polyprep column (Bio-Rad) and then subjected to ultrafiltration using a 5 kDa cut-off filter (Amicon/Millipore) where the volume was reduced to approximately 200 μl. Subsequently, the samples were extracted with phenol,

precipitated with ethanol, resuspended in a mixture of water and denaturing load dye, and subjected to denaturing gel electrophoresis as described above. For ^{32}P-labeled substrates, gels were analyzed by phosporimaging (Storm 860, Molecular Dynamics) and quantified using ImageQuant software. For unlabeled substrates, DNA gels were stained with ethidium bromide and imaged and quantified using the AlphaImager 3400 (AlphaInnotech). All concentrations given are final concentrations.

2.2.4 Precipitation of STC with PEG

Unlabeled STC was formed as described above in a total volume of 500 µl. PEG-8K (200 mg; 40%, w/v) (Sigma-Aldrich) was added to an STC reaction and the sample was mixed with gentle agitation until the PEG was fully dissolved. The mixture was incubated for 16 hours at 25°C without further agitation, and then subjected to centrifugation at 14,000 rpm for 20 minutes at 25°C. After gently removing the supernatant, centrifugation was repeated to facilitate complete removal of the supernatant. The pellet was then resuspended in one-fifth of the initial reaction volume using RB (see above) plus 1 mM Tris-HCl (pH 8), 0.1 mM EDTA, 0.05 mM NaCl. As described above, samples were applied to a native 5% polyacrylamide gel, stained with ethidium bromide and quantified using the AlphaImager.

2.3.0 Results

2.3.1 Short DNA fragments are suitable targets for Tn5 insertion *in vitro*

Our initial objective was to see if a stable target containing transpososome could be produced in the Tn5 system using a short linear target DNA containing a preferred insertion site. We generated a Tn5 transpososome with a 30 bp ME (lacking flanking donor DNA) and mixed this transpososome (referred to as the pre-cleaved transpososome or PC t'some) with an excess of a 37 bp DNA fragment (T3-37; see Figure 2.1B) containing a previously defined Tn5 insertion hot-spot (8). After incubation in the presence or in the absence of a divalent metal ion (either $CaCl_2$ or $MgCl_2$), samples were applied to a native polyacrylamide gel to separate transpososomes from protein-free ME substrate. We show that there is efficient conversion (>90%) of the PC t'some to a faster migrating species, and that this conversion is dependent on both the addition of a target DNA fragment and $MgCl_2$. Note that $MgCl_2$, but not $CaCl_2$, supports catalysis in the Tn5 system (compare lanes 1–4 and lane 6 with lane 5 in Figure 2.2A) (10). There is some radioactive signal in the vicinity of the STC in samples where $MgCl_2$ but not target DNA was added (lane 2). This signal is more diffuse than the STC in lane 5, and may be due to either a low level of nuclease activity in our transposase preparation or the occurrence of a low level of strand transfer into the unreacted ME DNA.

While the dependence of the formation of the above alternative form of the transpososome is consistent with this species being an STC, it was somewhat unexpected that an STC would migrate through the gel at a faster rate than the PC t'some. Conversion of the Tn5 PC t'some to an STC is expected to increase the mass of the transpososome from 146 kDa to 170 kDa, and this could reduce the mobility of the complex. For example, when the analogous experiment was performed in the Tn10 system, the STC was found to migrate at a slower rate relative to the PC t'some (compare

lanes 1 and 2, and lanes 3–5 in Figure 2.2B) (19). In this case, the mass of the Tn10 transpososome increased from 184 kDa to 210.8 kDa. However, in the bacteriophage Mu system, the STC migrates further in a native gel than the cleaved donor complex (20). Since the mobility of a protein-nucleic acid complex in a native gel is determined by the mass to charge ratio and the DNA is highly charged, it may be that the amount of charge neutralization afforded by protein binding to target DNA is significantly different in the Tn10 versus the Mu and Tn5 systems, and that this accounts for the different pattern of transpososome migration in the former (see 'Discussion').

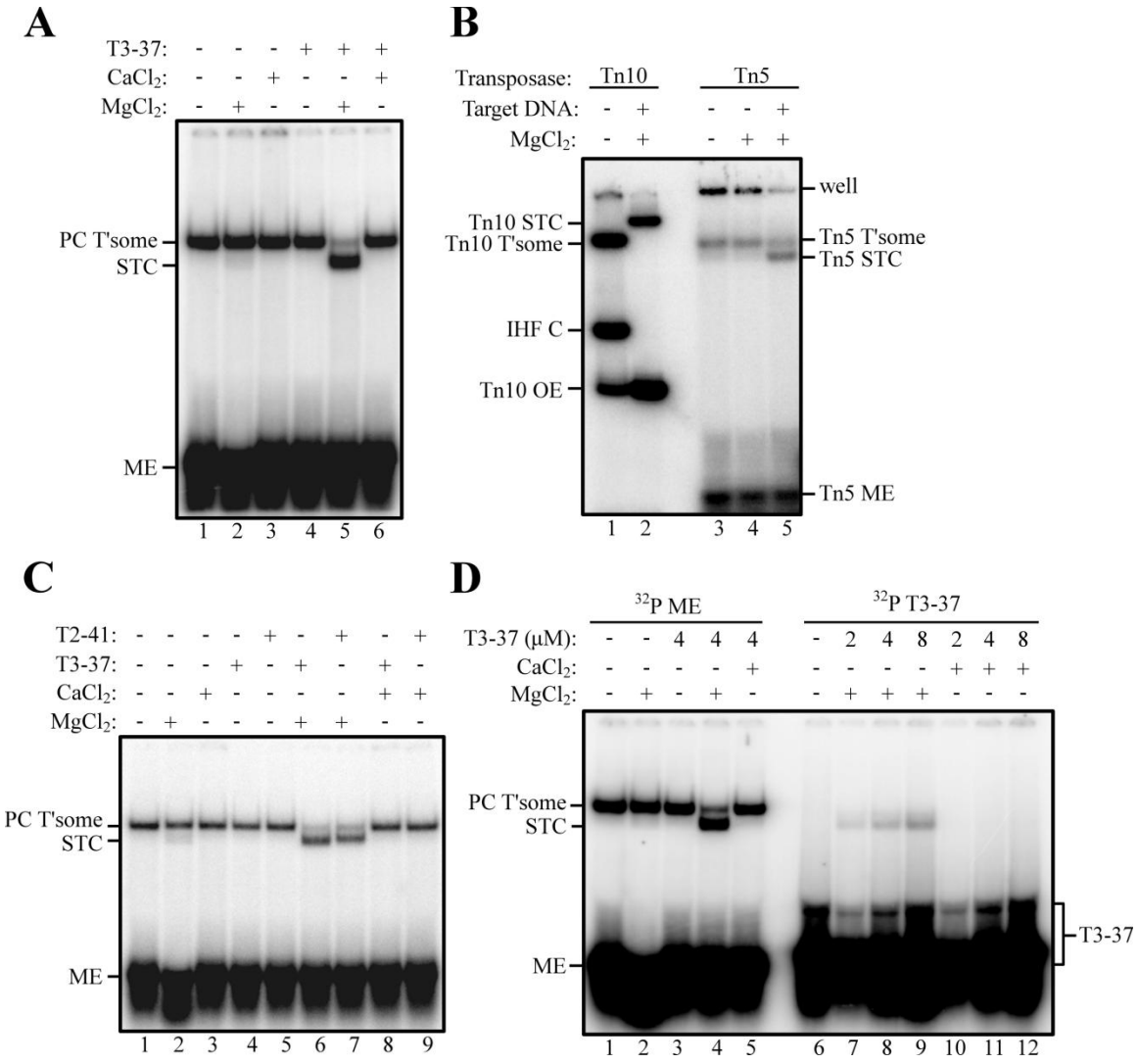
We have also varied the size and sequence of the target fragment in order to determine if either of these factors has an impact on the mobility of the putative STC in the Tn5 system. We show that addition of a different target DNA fragment (T2-41) that is 41 bp in length (Figure 2.1B) and contains tandem, but not overlapping, repeats of a hot-spot for Tn5 insertion (5'-GGATNATCC-3'), generates a transpososome that migrates faster than the PC t'some (compare lane 7 with lanes 1, 5 and 9 in Figure 2.2C).

We also switched the DNA labeling set-up so that the target, instead of the ME DNA, contained ^{32}P at its 5' termini. We show that a species with the same mobility as the putative STC is observed when the target DNA is 5' ^{32}P-labeled, and that detection of this species is dependent on the presence of $MgCl_2$ (compare lanes 7–9 with lane 4 in Figure 2.2D).

Notably, the putative STC is not observed when the appropriate assembly reaction is treated with phenol (data not shown). In addition, when the putative STC band was cut out of the native gel and the eluted material was subjected to SDS-PAGE, only full-

Figure 2.2. Formation of a Tn5 strand transfer complex.

(A) A pre-cleaved transpososome (PC T'some) containing (5′-^{32}P)-labeled mosaic end (ME) DNA (30 bp) was mixed where indicated with T3-37 target DNA, $CaCl_2$ or $MgCl_2$. After incubation for 16 hours at 25°C reactions were applied to a native 5% polyacrylamide gel. STC is the putative strand transfer complex. Evidence that this species is a STC is provided in Figure 2.3. (B) The relative mobility of STCs formed in the Tn5 and Tn10 systems are compared. The outside ends (OE) and target DNA used in the Tn10 system are 70 bp and 40 bp in length, respectively, and the OE is (5′ ^{32}P)-labeled. Integration host factor (IHF) is required for assembly of the Tn10 transpososome and accordingly some residual IHF-OE complex (IHF C) is detected (lane 1). Target DNA binding requires prior removal of IHF from the Tn10 transpososome and therefore IHF is not expected to be present in the Tn10 STC. (C) Products of reactions containing two different Tn5 target DNAs, T3-37 and T2-41 (see Figure 2.1B), are shown. (D) Products of reactions generated with the ME or T3-37 labeled at their 5′ termini with ^{32}P are shown. Reactions in (B)-(D) were analyzed as in (A). Species were visualized by phosphorimaging.


A
T3-37: - - - + + +
CaCl2: - - + - - +
MgCl2: - + - - + -
PC T'some
STC
ME
1 2 3 4 5 6
B
Transposase: Tn10 Tn5
Target DNA: - + - - +
MgCl2: - + - + +
well
Tn10 STC
Tn10 T'some
Tn5 T'some
Tn5 STC
IHF C
Tn10 OE
Tn5 ME
1 2 3 4 5
C
T2-41: - - - - + - + - +
T3-37: - - - + - + - + -
CaCl2: - - + - - - - + +
MgCl2: - + - - - + + - -
PC T'some
STC
ME
1 2 3 4 5 6 7 8 9
D
32P ME
32P T3-37
T3-37 (μM): - - 4 4 4 - 2 4 8 2 4 8
CaCl2: - - - - + - - - - + + +
MgCl2: - + - + - - + + + - - -
PC T'some
STC
ME
T3-37
1 2 3 4 5 6 7 8 9 10 11 12

length Tn5 transposase was detected. This latter experiment rules out the possibility that faster migration of the putative STC is due to proteolysis of the transposase component of the transpososome (data not shown). We have shown that the putative STC species is not formed when reactions are carried out with the catalytically inactive Tn5 transposase mutant D97A(21), and that the relative mobility of the PC t'some and the putative STC does not change when the percentage of polyacrylamide in the native gel is varied from 4% to 8% (w/v) (data not shown).

2.3.2 Insertion into T3-37 is highly specific

To further characterize the putative STC, we excised this species from a native gel, purified the DNA from protein and subjected the DNA to electrophoresis on a denaturing polyacrylamide gel (Figure 2.3A). As would be expected for an STC, the denaturing gel analysis reveals the presence of DNA species that are larger than the input ME DNA; note that 50% of the labeled DNA is 30 nt in length because only one of the ME strands (the transferred strand) is joined to a target DNA in the strand transfer reaction, but both strands DNA are labeled with ^{32}P at their 5′ termini.

The denaturing gel analysis reveals that ME joining to the target DNA is highly sequence specific, as only six bands are observed when both target DNA and $MgCl_2$ are present (lane 2). Bands labeled 2ts and 2bs are of a size consistent with insertion into the region of the target DNA designated site 2 (ts and bs indicate top and bottom strand insertion products, respectively) (see the diagram in Figure 2.3B). Whereas bands labeled 1ts/1bs and 3ts/3bs are of a size consistent with insertions into sites 1 and 3, respectively.

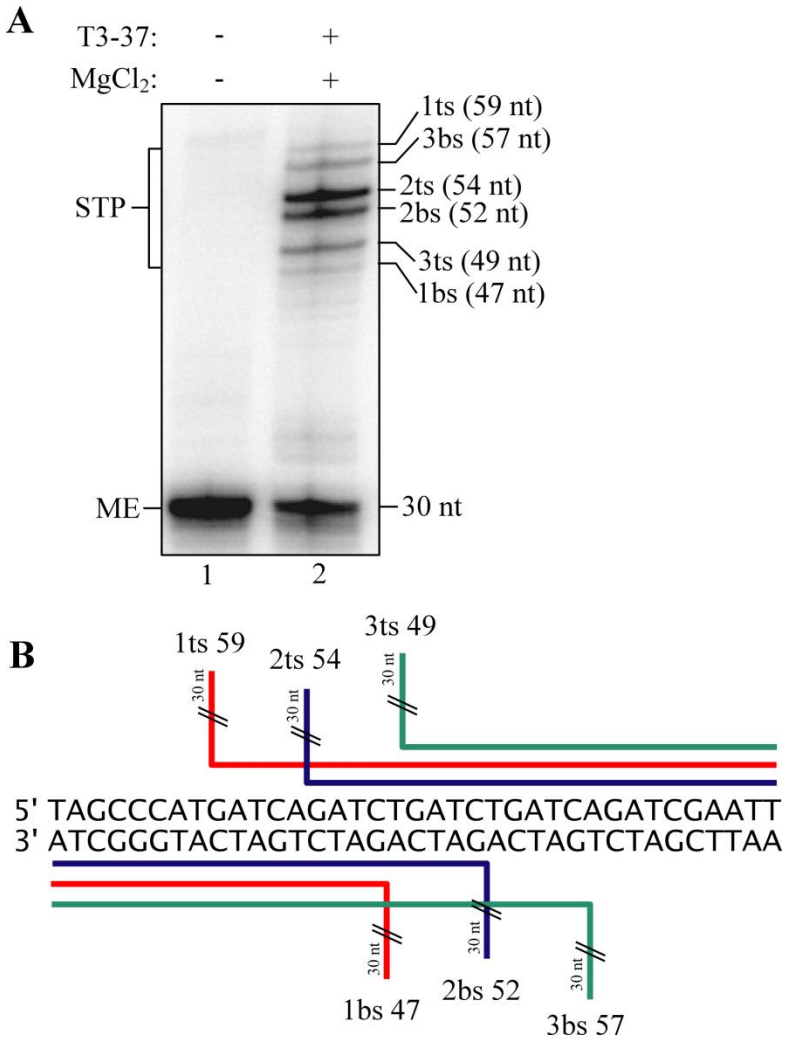


Figure 2.3. Sites of insertion within T3-37.

(A) Species labeled STC and ME in Tn5 reactions in Figure 2.2A were extracted from the native gel. The purified DNA was analyzed on a high-resolution, denaturing 8% polyacrylamide gel. A background ladder of bands was visible upon over-exposing the image shown and this served as a counting ladder for determining the size of each of the major species labeled. These assignments were confirmed by repeating the experiment with a variant of T3-37 where the terminal 3 bp on the left-hand side of T3-37 was absent (see the text). For each STP, the number refers to the usage of site 1, 2 or 3 and ts and bs indicate a top and a bottom strand STP respectively. (B) The illustration shows the predicted origin of each of the strand transfer products (STP) in (A). The vertical line represents the transferred strand of the ME (30 nt) and the horizontal line represents the target DNA component (the length of the latter is indicated in each case). Site 1, 2 and 3 STPs are coloured red, blue and green, respectively.

To confirm that these assignments are correct, we altered the size of the target DNA and observed the expected change in the product sizes; for example, when we removed 3 bp from the left end of T3-37, a 49 nt instead of 52 nt product was detected, as well as a 54 nt product. The former matches the expected product size for a bottom strand joining event at site 2 and the latter matches the expected product size for a top strand joining event at site 2 (data not shown).

Quantification of strand transfer product (STP) band intensities from the above experiment indicates that approximately 70% of the insertions went into site 2, 23% went into site 3, and 7% went into site 1. Given that Tn5 insertion specificity is relatively low, it was quite surprising that a single 5'-GATCNGATC-3' site within an overlapping array of this sequence would support such a high level of insertion. There is also a slight strand bias towards top strand events for site 2. Notably, it has been shown that top and bottom strand joining events are not tightly coupled in the Tn5 system (21). The results presented thus far indicate that a short linear DNA fragment with a hot-spot for Tn5 insertion (previously defined in a plasmid system) is a very efficient substrate for Tn5 insertion. The resulting STC is stable and migrates as a homogeneous species on native polyacrylamide gel electrophoresis, where its mobility is distinct from that of the PC t'some. Furthermore, there is a high degree of target site selectivity in this system.

2.3.3 Defining a minimal target sequence for STC formation

We were interested in determining the minimal size of a target fragment for STC formation, as reducing the target DNA size could be an important factor in optimizing crystallization conditions. In addition, we wanted to know if reducing the target DNA

size would further increase the specificity of insertion so that essentially all of the insertion events would go into one site. We tested STC formation and insertion specificity with two shorter variants of T3-37, named T3-17 and T3-21, the sequences of which are shown in Figure 2.4A. T3-17 has two instead of four overlapping repeats of the target sequence 5'-GATCNGATC-3', whereas T3-21 has three overlapping repeats of this sequence. From the band shift experiment in Figure 2.4B, it is apparent that T3-21 supports STC formation at efficiency comparable to that observed for T3-37 (compare lanes 4 and 8; also see Figure 2.4C, lane 8). In contrast, there does not appear to be efficient conversion of the PC t'some to STC in the T3-17 reaction (lane 6); the appearance of a diffuse band just below the PC t'some species could be a STC, although, as shown below, there is no evidence for strong target site selectivity in this reaction.

To evaluate insertion specificity, portions of the above reactions were analyzed on a denaturing gel. The denaturing gel analysis shows that there is strong target site selectivity in the T3-21 reaction, but not in the T3-17 reaction (compare lanes 7 and 8, and lanes 5 and 6 in Figure 2.4C). The most abundant STP in the T3-21 reaction is of a size consistent with insertion into the bottom strand of the site 2 equivalent in this target DNA (see the diagram in Figure 2.4C). The species designated 2bs accounts for approximately 16% of the total counts in lane 8. Surprisingly, the corresponding top strand STP (2ts) is much less abundant, constituting only 2% of the total counts. No other species stand out above background in lane 8. In comparison, 19% of the total counts in the T3-37 reaction are due to STPs, and these products result from insertion into sites 1, 2 and 3, with site 2 being the predominant site used. Thus, relative to the T3-37 reaction, the T3-21 reaction exhibits a higher degree of target site selectivity. We have not further

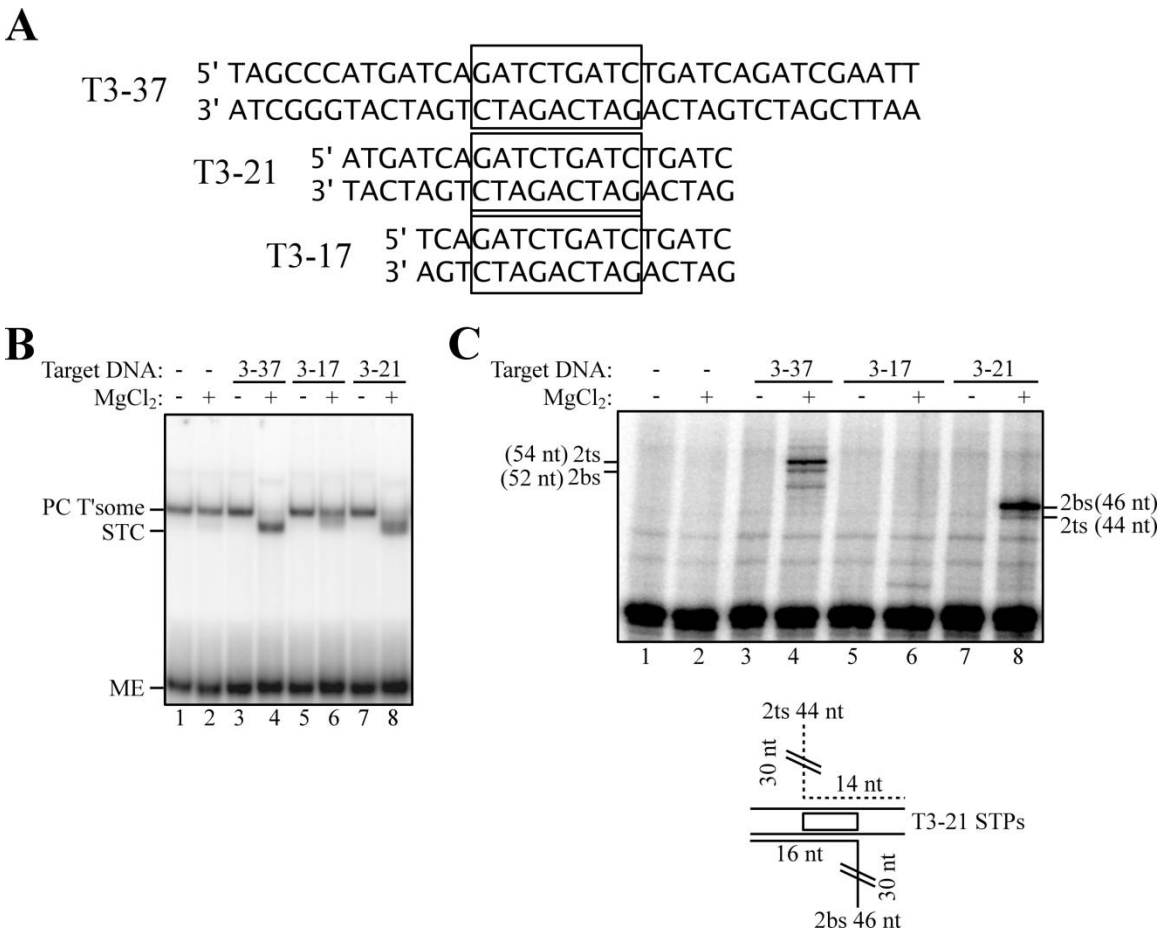


Figure 2.4. Effect of target DNA length on efficiency of strand transfer.

(A) The DNA sequences of T3-37 variants used to examine the impact of reducing the length of T3-37 on STC formation are shown. The box indicates the position of site 2 in T3-37 and the site 2 equivalent in T3-21 and T3-17. (B) and (C) Native and denaturing gel analyses, respectively, of strand transfer reactions performed with the target DNAs in (A). The size of the predominant STP in the T3-21 reaction (2bs) is consistent with efficient strand transfer into only the bottom strand of the site 2 equivalent in this target (see the accompanying diagram - note that the broken line indicates the top strand product of 44 nt is formed much less efficiently than the bottom strand product of 46 nt). The box indicates the position of the site 2 equivalent in T3-21, as in (A).

investigated the basis for the strand bias in site 2 usage in the T3-21 reaction. With regard to defining determinants in the target DNA that are important for strand transfer, the results of the above experiment indicate that at least three repeats of 5'-GATCNGATC-3' are required to support efficient usage of a short DNA fragment for Tn5 insertion.

2.3.4 Utilization of target DNAs with strategically positioned missing phosphates greatly reduces the amount of input target DNA required for efficient STC formation.

In earlier studies directed at defining important determinants in target DNAs for Tn10 insertion, we demonstrated that target DNAs missing a phosphate group at one of the sites of strand joining exhibited a greatly increased ability to: (1) form a stable complex with the transpososome; and (2) to support strand joining to the intact scissile phosphate group of the modified target site (22). We considered the possibility that such a modification might also increase target site selectivity in the Tn5 system. Notably, in all of the experiments shown thus far target DNA was present in large excess (~700-fold) relative to the transpososome. This could prove problematic for structural studies and, consequently, it would be advantageous to use a 'modified' target fragment, if the modification significantly reduces the amount of input target DNA required for efficient STC formation.

Initially, we generated a variant of T3-37, which is missing a phosphate group at the scissile position on the bottom strand of insertion site 2 (designated MP +4′; see Figure 2.1B). When we compared the ability of this variant to participate in strand transfer, relative to T3-37, we found that T3-37 MP+4′ supported efficient STP formation

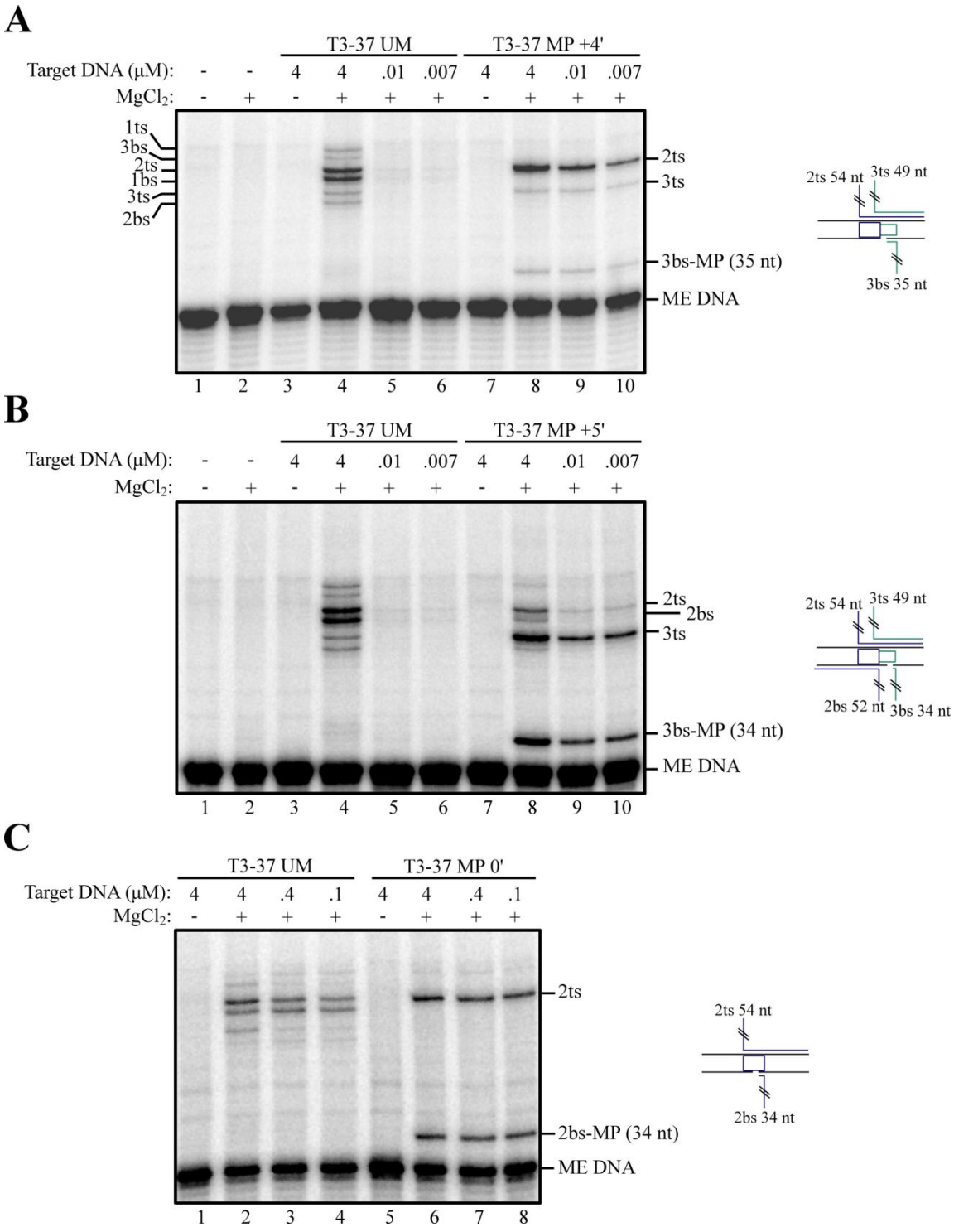
at much lower concentrations of target relative to the unmodified target DNA (T3-37 UM). For example, when the concentration of target DNA (either T3-37 MP+4′ or T3-37 UM) was reduced 400-fold (from 4.0 μM to 0.01 μM), there was a decrease of at least tenfold in the site 2 STPs in the T3-37UM reaction and only a 1.7-fold decrease in the single site 2 STP in the T3-37MP+4′ reaction (compare lanes 5 and 9 in Figure 2.5A). In addition, the specificity of insertion is increased into T3-37 MP+4′ as site 1 STPs are not detected. A similar effect was observed when then there was a missing phosphate group at the other position of strand transfer chemistry in site 2, the −4 phosphate (see Figure 2.1B) (data not shown). Note that the absence of the position +4′ phosphate group prevents the formation of the bs 52 STP, because a strand-joining event requires an intact phosphodiester bond at the site of joining (23).

Interestingly, if the position of the MP is moved over one nucleotide to the +5′ site, there is a significant decrease in site 2 usage and a corresponding increase in site 3 usage (Figure 2.5B). In this situation, both scissile phosphates of site 2 are intact and there is still a gap in the central region of site 3 (see Figure 2.1B). Thus, it would appear that, with respect to enhancing target site utilization, the position where a gap is introduced into a target site is critical.

We have looked at the effect of introducing a gap into the central region of site 2. The T3-37 MP0′ target supports efficient STC formation at low concentrations (Figure 2.5C). For example, the decrease in the 2ts STP for the T3-37 MP0′ reaction was 1.2-fold over a 40-fold decrease in target concentration, whereas the decrease in the T3-37 UM reaction was two-fold over the same target concentration range. Moreover, the specificity of insertion has increased relative to the reaction with T3-37 UM, as only site 2 and site 1

Figure 2.5. Effect of 'missing phosphates' on target site usage.

(A-C) Forms of T3-37 were generated in which a single phosphate group at a defined position of one strand was absent. These 'missing phosphate' (MP) target DNAs were used in strand transfer reactions and the reaction mixes were subjected to denaturing gel electrophoresis. Each MP target was analyzed in parallel with unmodified (UM) T3-37. The amount of target DNA used per reaction is shown. Note that in the bottom panel there was a 40-fold difference in the range of target concentrations used with the lowest concentration being 100 nM. In the upper two panels there was a 570-fold difference in the range of target concentrations used with the lowest concentration being 7 nM. The illustrations show each of the major strand transfer products (STP) detected. Colour-coding is as in Figure 2.3 and the location of each of the MPs is indicated in Figure 2.1C.


A
T3-37 UM
T3-37 MP +4'
Target DNA (μM): - - 4 4 .01 .007 4 4 .01 .007
MgCl2: - + - + + + - + + +
1ts
3bs
2ts
1bs
3ts
2bs
2ts
3ts
3bs-MP (35 nt)
ME DNA
1 2 3 4 5 6 7 8 9 10
2ts 54 nt
3ts 49 nt
3bs 35 nt
B
T3-37 UM
T3-37 MP +5'
Target DNA (μM): - - 4 4 .01 .007 4 4 .01 .007
MgCl2: - + - + + + - + + +
2ts
2bs
3ts
3bs-MP (34 nt)
ME DNA
1 2 3 4 5 6 7 8 9 10
2ts 54 nt
3ts 49 nt
2bs 52 nt
3bs 34 nt
C
T3-37 UM
T3-37 MP 0'
Target DNA (μM): 4 4 .4 .1 4 4 .4 .1
MgCl2: - + + + - + + +
2ts
2bs-MP (34 nt)
ME DNA
1 2 3 4 5 6 7 8
2ts 54 nt
2bs 34 nt

STPs are detected. Notably, the gap present in T3-37 MP0′ is not in site 1 and is directly opposite a scissile phosphate group in site 3 (see Figure 2.1B). We found that introduction of a gap opposite either of the site 2 scissile phosphate groups did not increase site 2 utilization at low target DNA concentrations (data not shown). In addition to providing a means of limiting the amount of input target required to promote strand transfer, the above missing phosphate experiments present a means of increasing target site selectivity. Furthermore, these experiments provide fundamental information regarding determinants that are important in the Tn5 strand transfer reaction. We presume that strategic removal of phosphate groups enhances the utilization of a particular target site by increasing the conformational flexibility of the DNA in that site in a position-specific manner.

2.3.5 Concentration and partial purification of the T3-37 STC by precipitation with PEG

To facilitate structural studies on the Tn5 STC generated in this work, we have attempted to purify the STC formed with T3-37. We found that the most effective method for doing this was to selectively precipitate the STC out of solution with polyethylene glycol (PEG). In an initial experiment, we added different amounts of solid PEG to STC assembly reactions containing a fixed amount of unlabeled ME DNA, target DNA and Tn5 transposase. These reactions were carried out on a scale where it was possible to detect both the input DNAs and transpososomes by ethidium bromide staining of a native gel (data not shown). We found that under our standard conditions for assembly, addition of PEG to 40% (w/v) was optimal for selectively precipitating the STC. Comparison of

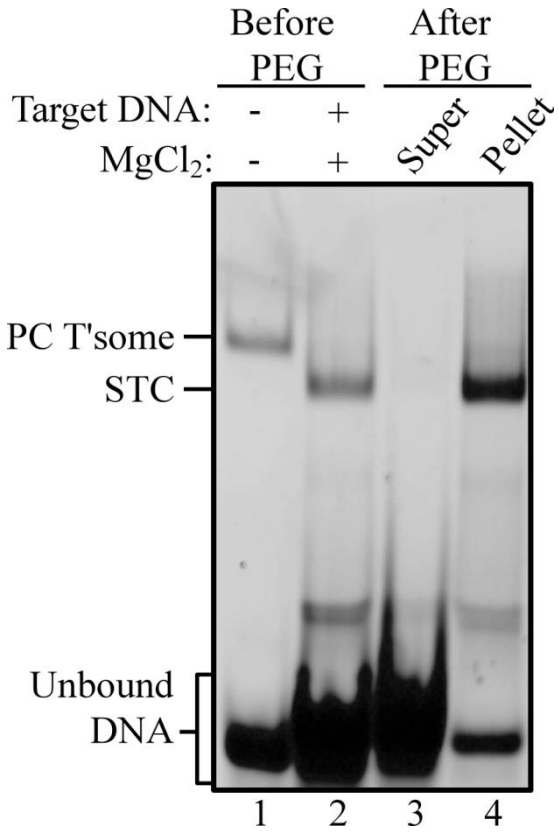


Figure 2.6. Concentration and partial purification of the STC.

The STC was formed with T3-37 under standard conditions in a volume of 500 µl. Aliquots (20 µl) were removed from the reaction to monitor the efficiency of transpososome formation in the starting material using native polyacrylamide gel electrophoresis (lanes 1 and 2). Polyethylene glycol was added to the remainder of the assembly reaction. After incubation for 16 hours the sample was centrifuged, the supernatant was decanted and the pellet was resuspended in 100 µl of reaction buffer: 20 µl of the supernatant (Super) and 10 µl of the pellet fraction were analyzed as described above (lanes 3 and 4). The material in lanes 2 and 4 represent 2% and 10% of the initial reaction volume respectively.

lanes 1 and 2 in Figure 2.6 shows that addition of T3-37 and $MgCl_2$ to the PC t'some results in essentially 100% conversion of the former to STC. Lane 3 shows what is left in the supernatant after subjecting the mixture in lane 2 to treatment with PEG. Lane 4 shows a portion of the pellet fraction recovered. The results show that the supernatant fraction contains large amounts of unbound DNA and very little STC, whereas the pellet fraction is greatly enriched in STC. Compared to the untreated sample, where only 2-3% of the total DNA detected by ethidium bromide staining is STC, approximately 55% of the total DNA detected in the PEG pellet is STC. This represents ~20-fold enrichment in STC. It is worth noting that precipitation of the STC with PEG could be applied to larger-scale reactions than those performed here to generate, in one simple step, quantities of the STC that are suitable for crystallization trials.

2.4.0 Discussion

The determination of the structure of the Tn5 pre-cleaved transpososome provided detailed information on the protein-DNA and protein-protein interactions that govern the assembly of the transpososome (6, 7). However, molecular details regarding the protein-DNA and protein-protein interactions governing target DNA-binding and the joining of transposon ends to a target site have remained elusive. In addition to providing insight into the nature of both the target capture and strand transfer steps, determining the structure of a Tn5 transpososome with target DNA bound may aid in the design of HIV integrase inhibitors, as the structures of the Tn5 transposase and HIV integrase are very similar and both proteins catalyze a strand transfer reaction. We have been able to generate a strand transfer complex in the Tn5 system using a pre-cleaved transpososome

and a target DNA made from short oligonucleotides. Insertion into the target DNA chosen is surprisingly specific, thus allowing the preparation of what appears to be a highly homogenous population of transpososomes. We have been able to generate a highly enriched fraction of the Tn5 STC by precipitation with PEG and anticipate that this fraction will provide a good starting point for crystallization trials.

2.4.1 Evidence for STC formation

We assessed the ability of the Tn5 transpososome to form a stable complex with a short target DNA fragment by mixing the transpososome with the target DNA and asking if this addition changed the mobility of the transpososome in an electrophoretic mobility shift assay. A clear change in transpososome mobility was detected, specifically under conditions where strand transfer would be expected. We had anticipated that the increase in mass associated with end joining to the target DNA would reduce the rate of transpososome migration, as we had observed in the analogous experiment in the Tn10 system (19). However, in this respect, the Tn5 system behaved more like the Mu system, as the STC migrated through a native polyacrylamide gel at a faster rate than the cleaved donor complex (20).

The efficiency of conversion of the initial transpososome to the faster migrating form was close to 100% when a large excess of target DNA was used and $MgCl_2$, but not $CaCl_2$, was present. The dependence of the formation of the novel transpososome on both target DNA and $MgCl_2$ suggested that this transpososome is an STC as opposed to a target capture complex, because only the former would show an absolute dependence on the presence of both target DNA and a catalytic divalent metal cation. We verified this by

purifying the faster migrating species from a native gel and analyzing the DNA on a denaturing gel. This analysis revealed an increase in the length of the input ME fragment, as would be expected if the ME DNA were joined to target DNA.

Importantly, only six different species longer than the ME DNA were observed in the above assay, indicating that joining events did not occur randomly throughout the target sequence. In fact, the sizes of these species were indicative of the ME being joined to three different target sites, designated sites 1, 2 and 3. This is in stark contrast to what is observed in retroviral integration systems, where an extensive ladder of strand transfer products is observed in a similar experiment, which is indicative of target site selectivity being very low (24-26).

We presume that we failed to detect a target capture complex, because this species is significantly less stable than the STC. This is consistent with our observation that a large excess of target DNA is required for the complete conversion of the PC t'some to STPs. If the transpososome were to bind weakly to the target DNA, then a high dissociation rate or low association rate would be partially offset by the increased opportunity for binding events afforded by a high concentration of target DNA. Strand transfer would then stabilize target-containing transpososomes by covalently joining transposon ends to the target DNA, effectively driving the reaction to completion.

It is not clear at this point why theTn5 STC migrates faster than the PC t'some, whereas the Tn10 STC migrates slower than the corresponding PC t'some. There was no evidence for proteolysis of the Tn5 transposase protein over the course of the incubation of PC t'some, target DNA and $MgCl_2$. In principle, there could be additional protomers of transposase bound to the PC t'some that are subsequently titrated off the

transpososome in the presence of excess target DNA. In which case, the gain of mass resulting from incorporation of target DNA into the transpososome could be offset by the loss of mass due to dissociation of transposase protomers. Precedent for this comes from studies in the bacteriophage Mu system, where 'extra' molecules of MuA were found associated with the initial transpososome. In this case, treatment of these transpososomes with heparin, a polyanionic polymer that can compete with DNA for protein binding, was found to be sufficient to dissociate these extra MuA molecules from the transpososome (27). However, we found that treatment of the Tn5 STC and the PC t'some with heparin did not alter the relative mobilities of these species (data not shown), so this possibility seems unlikely. Alternatively, it may be that in the Tn10 system the transposase contacts a greater number of phosphate residues in the target DNA backbone in the STC, relative to what occurs in theTn5 and Mu systems and, accordingly, there is significantly more charge neutralization in the target DNA, an occurrence that should reduce the mobility of the complex. In fact, it has been shown that Tn10 transposase contacts approximately 24 base pairs of target DNA in the Tn10 target capture complex, which is over half the length of the target DNA used in the current work (28).

2.4.2 Basis for target site selectivity

We initially chose the T3-37 target fragment for study because a closely related sequence, when present in a plasmid DNA, was identified as a hot-spot for Tn5 insertion *in vivo* (8). Both T3-37, and the sequence it was designed from, contain four overlapping repeats of the DNA sequence 5'-GATCNGATC-3'. For three of the four overlapping repeats there is a perfect match to the preferred base pairs at each of the five positions that are most

highly conserved in Tn5 target sites (see Figure 5C of (13)). This match would be expected to aid in transpososome binding to these sites. Interestingly, single 5'-GATCNGATC-3' sequences found in the inside end of IS50 (29) the Tn5 transposase promoter and a kanamycin resistance gene, have not been found to act as hot-spots for Tn5 insertion (8). Consistent with this, we found in the current work that two overlapping 5'-GATCNGATC-3' repeats were insufficient to impose high target site selectivity (T3-17 in Figure 2.4). Thus, the presence of base pairs matching the consensus sequence is not on its own sufficient to confer hot-spot status to a site. Importantly, hot-spot status was restored by adding a 5'-GATC-3' sequence to each side of the 9 base pair core site (T3-21 in Figure 2.4), indicating that the presence of at least three overlapping repeats of 5'-GATCNGATC-3' is sufficient to generate an insertion hot-spot.

What might be the basis for the above effect? It has been proposed that base-specific interactions between transposase and the consensus base pairs of target sites, present within an array of repeated sequences, facilitates the binding of multiple transpososomes to the region (8). Transposase molecules within adjacent transpososomes would then be able to interact, thereby strengthening the binding of each individual transpososome. Cooperative binding of transpososomes would then increase the probability of a productive strand-joining event within the array, presumably by increasing the length of time each transpososome spends on a particular target site. If this were the case, the most central site within the array of repeats would be expected to support the largest number of insertion events because the transpososome bound to the central site would be tethered to this site by transpososomes bound on either side. In contrast, the outermost sites would bind a transpososome that is tethered on only one side

and therefore would be used less frequently. Consistent with this model, we have shown that site 2, the most central repeat in the T3-37 target, was used seven times more frequently than the outer repeats comprising site 1 and site 3. In addition, there was no evidence of usage of a fourth repeat at the extreme right end of T3-37.

If the above model is correct, then there is the concern that further information obtained from the experimental system employed here may apply to only a minor subset of target sites that are capable of binding multiple transpososomes. Furthermore, the presence of multiple transpososomes on a target DNA would likely complicate high-resolution structural studies. However, based on the mobility of the STC, it is unlikely that the STC formed with T3-37 and T3-21 contains more than one transpososome. Thus, while adjacent transpososomes could increase the stability of the most central transpososome on the target DNA, we anticipate that only the most central transpososome participates in strand transfer and therefore remains stably linked to the target DNA.

2.4.3 Increased conformational flexibility at specific locations within a Tn5 target site aids in STC formation

We have shown in the Tn10 system that target DNAs missing phosphate groups at specific locations are better substrates for target capture and strand transfer than the equivalent target molecules with intact DNA strands (22). There is also evidence in the Tn10 system that target capture is facilitated by bending of the target DNA (18). Furthermore, the bending loci coincide with positions in the backbone where removal of a phosphate group enhances target site utilization (22). Taken together, these results

support the idea that missing phosphate groups increase target capture and strand transfer by making it easier for the transpososome to bend the target DNA. The increased propensity of target DNA to bend is also expected to enhance target interactions in the bacteriophage Mu system (30) and in DNA transposition directed by the RAG proteins (31). Interestingly, the sites of target bending and missing phosphate enhancement are also the sites of strand transfer reaction chemistry in the Tn10 system (i.e. the scissile phosphate groups). We have observed the same phenomenon in the Tn5 system, as we have shown that target DNAs with a missing phosphate group at either of the sites of chemistry in site 2 of the T3-37 target DNA (i.e. the +4′ and the −4 positions), permit strand transfer into the intact scissile phosphate group at a greatly reduced concentration of target DNA. In addition, as in the Tn10 system, generating a missing phosphate group at one of the adjacent residues (the +5′ position in Tn5) failed to enhance target site 2 usage, as site 3 became the most frequently used target site when this alteration was introduced. This indicates that increasing the conformational flexibility of the DNA, at only very specific positions, has a positive effect on STC formation in the two systems.

It is probably not a coincidence that increased conformational flexibility at the sites of strand joining is beneficial to STC formation in both Tn10 and Tn5 systems. It may be that, in both systems, the strand transfer chemistry is dependent on the transpososome adopting a conformation wherein the chemical bonds of the scissile phosphate groups are under strain. This same conformation might represent the most stable form of the target capture complex, as target bending could increase the number of contacts made between the target DNA and transposase (18, 29).

The site 2 +5′ missing phosphate group is located in the centre of site 3 and, because this alteration enhanced site 3 usage, it seemed likely that increased conformational flexibility in the central portion of a Tn5 target site would also be beneficial for target site selection. This hypothesis was confirmed by testing the ability of a target DNA with a missing phosphate group at the central position of site 2 (the 0′ position) to act in strand transfer. This alteration significantly increased the ability of this site to participate in strand transfer at reduced target concentrations, and it increased the specificity of the system as events into alternative sites were reduced. This observation could prove useful for structural studies, where it will be advantageous to generate as homogenous a population of STCs as possible. Notably, the presence of a discontinuity in the DNA backbone of a protein-DNA complex is not expected to cause problems with regard to crystallization. In fact, obtaining a crystal structure of IHF in complex with an IHF binding site was dependent on the presence of a nick in one of the DNA strands (32).

2.4.4 Prospects for high-resolution structural studies

Acquiring a high-resolution structure for a Tn5 transpososome bound to target DNA presents some novel challenges. Compared to other transposition systems, such as Tn7 (33) or Tn10 (34), Tn5 is much less selective with regard to where it inserts (13). Thus, at the outset of the current work it was not clear if it would be possible to generate a population of target containing molecules where the Tn5 transpososome was bound or integrated into a single site. In addition, while there has been some evidence for highly specific, albeit inefficient, Tn5 insertion into a short (40 bp) cleaved Tn5 end sequence (35) it had not been shown that Tn5 could bind and insert with high efficiency into a

short target DNA of a sequence unrelated to Tn5 ends. We have shown that short target DNA fragments can serve as very efficient substrates for strand transfer in the Tn5 system. Moreover, by using a short target fragment with overlapping repeats of a preferred Tn5 insertion site, we were able to produce a fairly homogeneous population of STCs in which at least 70% of the strand transfer events occurred at a single site. Furthermore, we have shown that the degree of target site selectivity could be increased by reducing the size of the target DNA or by removing phosphate groups at specific locations in the target DNA backbone.

A possible limitation of our short target fragment system is that, in order to drive the conversion of PC t'some to STC to completion, it was necessary to use a large excess of target DNA. Since only a small fraction of the total input target DNA is actually converted to STC, we anticipate that it will be necessary to purify the STC away from the unreacted target DNA for crystallographic studies. We have found that PEG can selectively precipitate the STC from unbound DNA (see 'Materials and methods'), and we have used this approach to produce a fraction that is highly enriched in STC and thus overcome the above problem. We also anticipate that selective precipitation of the STC by treatment with PEG will be useful in concentrating the STC from 'scaled-up' reactions to generate sufficient quantities of this complex for crystallization trials.

In addition to obtaining fundamental information on how transposition reactions occur, defining a high-resolution structure of a Tn5 target DNA containing transpososome could be useful in anti-AIDS drug development. Tn5 represents a surrogate model system for developing HIV integrase (IN) inhibitors, because the structure of the catalytic core domains of HIV-IN and Tn5 transposase are remarkably

similar (6, 36). In proof of principle of the Tn5 system serving as a surrogate system for HIV-IN drug development, it has been shown recently that a subclass of compounds that inhibit initial transpososome formation in the Tn5 system are also effective inhibitors of HIV-IN (37).

Table 2.1. Oligonucleotides used for assembly of the ME substrate and target DNAs.

Name	Sequence (5'-3')
Unmodified oligos	
ME NTS 30	CTGTCTCTTATACACATCTTGAGTGAGTGA
ME TS 30	TCACTCACTCAAGATGTGTATAAGAGACAG
T3-37 T	TAGCCCATGATCAGATCTGATCTGATCAGATCGAATT
T3-37 B	AATTCGATCTGATCAGATCAGATCTGATCATGGGCTA
T2-41 T	CCCATGGATAATCCTGGATAATCCTGGATAATCCTGGATCC
T2-41 B	GGATCCAGGATTATCCAGGATTATCCAGGATTATCCATGGG
T3-21 T	ATGATCAGATCTGATCTGATC
T3-21 B	GATCAGATCAGATCTGATCAT
T3-17 T	TCAGATCTGATCTGATC
T3-17 B	GATCAGATCAGATCTGA
HisG1 40 T	AAAATTAATTTACACACTCAGCGCCTGATTGCGATGGCGG
HisG1 40 B	TCCGCCATCGCAATCAGGCGCTGAGTGTGTAAAATTAATTT
[a]MP oligos	
T3-37 +4' L 22	GATCAGATCTGATCATGGGCTA
T3-37 +4' R 15	AATTCGATCTGATCA
T3-37 +5' L 23	AGATCAGATCTGATCATGGGCTA
T3-37 +5' R 14	AATTCGATCTGATC
T3-37 +3' L 21	ATCAGATCTGATCATGGGCTA
T3-37 +3' R 16	AATTCGATCTGATCAG

Each pair shown (e.g. T3-37 +4' L 22 and T3-37 +4' R 15) was mixed with the corresponding complementary strand (e.g. T3-37 T) of unmodified oligonucleotide to generate the missing phosphate targets.

a. MP indicates that oligonucleotides in this group were used to make missing phosphate (MP) targets.

2.5 References

1. Haniford,D.B. and Chaconas,G. (1992) Mechanistic aspects of DNA transposition. *Current opinion in genetics & development*, **2**, 698-704.
2. Mizuuchi,K. (1997) Polynucleotidyl transfer reactions in site-specific DNA recombination. *Genes to cells : devoted to molecular & cellular mechanisms*, **2**, 1-12.
3. Haren,L., Ton-Hoang,B. and Chandler,M. (1999) Integrating DNA: transposases and retroviral integrases. *Annual review of microbiology*, **53**, 245-81.
4. Nagy,Z. and Chandler,M. (2004) Regulation of transposition in bacteria. *Research in microbiology*, **155**, 387-98.
5. Curcio,M.J. and Derbyshire,K.M. (2003) The outs and ins of transposition: from mu to kangaroo. *Nature reviews. Molecular cell biology*, **4**, 865-77.
6. Davies,D.R., Goryshin,I.Y., Reznikoff,W.S. and Rayment,I. (2000) Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science*, **289**, 77-85.
7. Lovell,S., Goryshin,I.Y., Reznikoff,W.R. and Rayment,I. (2002) Two-metal active site binding of a Tn5 transposase synaptic complex. *Nature structural biology*, **9**, 278-81.
8. Goryshin,I.Y., Miller,J. a, Kil,Y.V., Lanzov,V. a and Reznikoff,W.S. (1998) Tn5/IS50 target recognition. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 10716-21.
9. Reznikoff,W.S., Bhasin,A., Davies,D.R., Goryshin,I.Y., Mahnke,L. a, Naumann,T., Rayment,I., Steiniger-White,M. and Twining,S.S. (1999) Tn5: A molecular window on transposition. *Biochemical and biophysical research communications*, **266**, 729-34.
10. Goryshin,I.Y. and Reznikoff,W.S. (1998) Tn5 in vitro transposition. *The Journal of biological chemistry*, **273**, 7367-74.
11. Zhou,M., Bhasin,A. and Reznikoff,W.S. (1998) Molecular genetic analysis of transposase-end DNA sequence recognition: cooperativity of three adjacent base-pairs in specific interaction with a mutant Tn5 transposase. *Journal of molecular biology*, **276**, 913-25.
12. Shevchenko,Y., Bouffard,G.G., Butterfield,Y.S.N., Blakesley,R.W., Hartley,J.L., Young,A.C., Marra,M.A., Jones,S.J.M., Touchman,J.W. and Green,E.D. (2002) Systematic sequencing of cDNA clones using the transposon Tn5. *Nucleic acids research*, **30**, 2469-77.
13. Kang,Y., Durfee,T., Glasner,J.D., Qiu,Y., Frisch,D., Winterberg,K.M. and Blattner,F.R. (2004) Systematic mutagenesis of the Escherichia coli genome. *Journal of bacteriology*, **186**, 4921-30.
14. Berg,D.E., Schmandt,M.A. and Lowe,J.B. (1983) Specificity of transposon Tn5 insertion. *Genetics*, **105**, 813-28.

15. Peterson,G. and Reznikoff,W.S. (2003) Tn5 transposase active site mutations suggest position of donor backbone DNA in synaptic complex. *The Journal of biological chemistry*, **278**, 1904-9.
16. Chalmers,R.M. and Kleckner,N. (1994) Tn10/IS10 transposase purification, activation, and in vitro reaction. *The Journal of biological chemistry*, **269**, 8029-35.
17. Surette,M.G. and Chaconas,G. (1989) A protein factor which reduces the negative supercoiling requirement in the Mu DNA strand transfer reaction is Escherichia coli integration host factor. *The Journal of biological chemistry*, **264**, 3028-34.
18. Pribil,P.A. and Haniford,D.B. (2003) Target DNA Bending is an Important Specificity Determinant in Target Site Selection in Tn10 Transposition. *Journal of Molecular Biology*, **330**, 247-259.
19. Junop,M.S. and Haniford,D.B. (1997) Factors responsible for target site selection in Tn10 transposition: a role for the DDE motif in target DNA capture. *The EMBO journal*, **16**, 2646-55.
20. Savilahti,H., Rice,P.A. and Mizuuchi,K. (1995) The phage Mu transpososome core: DNA requirements for assembly and function. *The EMBO journal*, **14**, 4893-903.
21. Naumann,T.A. and Reznikoff,W.S. (2002) Tn5 transposase active site mutants. *The Journal of biological chemistry*, **277**, 17623-9.
22. Pribil,P.A., Wardle,S.J. and Haniford,D.B. (2004) Enhancement and rescue of target capture in Tn10 transposition by site-specific modifications in target DNA. *Mol Micro*, **52**, 1173-86.
23. Knowles,J.R. (1980) Enzyme-catalyzed phosphoryl transfer reactions. *Annual review of biochemistry*, **49**, 877-919.
24. Craigie,R., Fujiwara,T. and Bushman,F. (1990) The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. *Cell*, **62**, 829-37.
25. Katz,R.A., Merkel,G., Kulkosky,J., Leis,J. and Skalka,A.M. (1990) The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. *Cell*, **63**, 87-95.
26. Bushman,F.D. and Craigie,R. (1991) Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 1339-43.
27. Wu,Z. and Chaconas,G. (1997) The Mu transposase tetramer is inactive in unassisted strand transfer: an auto-allosteric effect of Mu A promotes the reaction in the absence of Mu B. *Journal of molecular biology*, **267**, 132-41.
28. Pribil,P.A. and Haniford,D.B. (2000) Substrate recognition and induced DNA deformation by transposase at the target-capture stage of Tn10 transposition. *Journal of molecular biology*, **303**, 145-59.

29. Lodge,J.K., Weston-Hafer,K. and Berg,D.E. (1991) Tn5 insertion specificity is not influenced by IS50 end sequences in target DNA. *Mol Gen Genet*, **228**, 312-5.
30. Yanagihara,K. and Mizuuchi,K. (2002) Mismatch-targeted transposition of Mu: a new strategy to map genetic polymorphism. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 11317-21.
31. Lee,G.S., Neiditch,M.B., Sinden,R.R. and Roth,D.B. (2002) Targeted Transposition by the V(D)J Recombinase. *Mol Cell Biol*, **22**, 2068-2077.
32. Rice,P.A., Yang,S., Mizuuchi,K. and Nash,H.A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell*, **87**, 1295-306.
33. Craig,N.L. (1991) Tn7: a target site-specific transposon. *Molecular microbiology*, **5**, 2569-73.
34. Halling,S.M. and Kleckner,N. (1982) A symmetrical six-base-pair target site sequence determines Tn10 insertion specificity. *Cell*, **28**, 155-63.
35. Ason,B. and Reznikoff,W.S. (2004) DNA sequence bias during Tn5 transposition. *Journal of molecular biology*, **335**, 1213-25.
36. Dyda,F., Hickman,A.B., Jenkins,T.M., Engelman,A., Craigie,R. and Davies,D.R. (1994) Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science*, **266**, 1981-6.
37. Ason,B., Knauss,D.J., Balke,A.M., Merkel,G., Skalka,A.M. and Reznikoff,W.S. (2005) Targeting Tn5 transposase identifies human immunodeficiency virus type 1 inhibitors. *Antimicrobial agents and chemotherapy*, **49**, 2035-43.

CHAPTER 3[1]

3 THE GLOBAL BACTERIAL REGULATOR H-NS PROMOTES TRANSPOSOSOME FORMATION AND TRANSPOSITION IN THE TN5 SYSTEM

3.1 Introduction

The H-NS protein is a global transcriptional repressor that plays an important role in lateral gene transfer (LGT) in gram-negative bacteria by temporarily silencing newly acquired genes (1-4). In addition, there is emerging evidence that H-NS can directly contribute to LGT by acting as a positive regulator of several different DNA transposition systems (5-9). In the case of Tn10, H-NS appears to up-regulate transposition by binding directly to the transposition complex (or transpososome) and altering the conformation of this complex in a manner that promotes intermolecular transposition events while at the same time inhibiting self-destructive intramolecular events (9-11). This represents a new role for H-NS and at the present time it is unclear how many other transposition systems are regulated by H-NS via a direct interaction with the transposition machinery.

H-NS is a highly abundant DNA-binding protein in gram-negative bacteria. It readily forms dimers in solution and can form tetramers and higher order oligomers. H-NS binds regulatory regions of genes and through higher order oligomerization is thought to bring distant segments of DNA together to block RNA polymerase progression (12, 13). H-NS preferentially binds bent and deformed DNA sequences (14). The ability of H-NS to recognize structural features of DNA, as opposed to recognition of specific DNA sequences, may make H-NS particularly well suited to regulate transposition reactions

[1]A version of this chapter has been published: Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2009) The global bacterial regulator H-NS promotes transpososome formation and transposition in the Tn5 system. *Nucleic acids research*, **37**, 309-21.

because the available evidence is consistent with transpososomes containing distorted DNA structures (15-19). In the current work we have asked if H-NS binds to and influences the activity of the Tn5 transposition machinery.

The overall frequency of Tn5 transposition is just under one event per 10^5 cells per generation. This relatively low transposition frequency reflects the occurrence of negative regulation primarily at three levels: (1) the expression of the transposase gene; (2) the expression of an inhibitor protein; and (3) the formation of the transpososome (20-24). With regard to the latter, assembly of the transpososome is inefficient. Tn5 transpososome assembly may involve the interaction of two monomer complexes (MCs), each of which consists of a single molecule of transposase bound to a Tn5 end sequence (Figure 3.1), although there is also evidence for another similar but more complicated assembly pathway (see 'Discussion'). MC formation is directed by an interaction between the N-terminal DNA-binding domain of transposase and a portion of the transposon end spanning roughly base pair 6 to 19 (19, 25, 26). These contacts are referred to as *cis* contacts. In contrast, pairing of two MCs is directed by *trans* contacts involving the C-terminal portion of a monomer bound to one end with base pair 1–5 of the partner end. In addition, an α-helix in the C-terminal portion of transposase mediates protein-protein interactions between transposase monomers bound to separate ends (27). Prior to DNA binding, the N- and C-terminal domains of transposase have a tendency to interact with each other and this interaction is inhibitory to both MC formation and transpososome formation (24). In addition to Tn5 transposase being suboptimal for transposition, the OE and IE sequences of Tn5 are also suboptimal for transposition. A mosaic end (ME) sequence made of a combination of OE and IE sequences has been

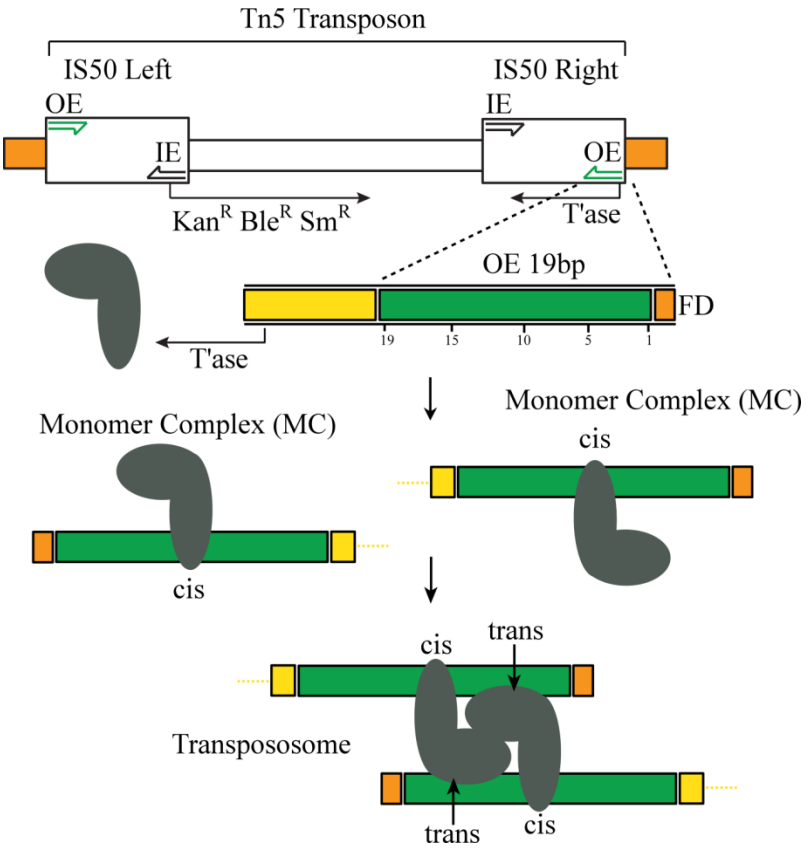


Figure 3.1. Structure of Tn5 and transposase-DNA interactions in transpososome assembly.

Tn5 is a 5.8 kb composite transposon encoding resistance to kanamycin (Kan), bleomycin (Ble) and streptomycin (Sm). These antibiotic resistance genes are flanked by nearly identical copies of the ~1.5 kb IS50 insertion sequences. IS50R encodes functional transposase (T'ase). Outside (OE) and inside (IE) sequences contain determinants for transposase binding and are shown as green and black half arrows, respectively. Orange rectangles represent flanking donor DNA (FD). A possible pathway for transpososome assembly is shown below the illustration of Tn5. An OE sequence (numbered green rectangle) is depicted binding a single molecule of transposase (grey shape) forming a monomer complex (MC). Yellow rectangles represent transposon DNA beyond OE. At this stage the transposase-OE contacts are defined as *cis* contacts. Subsequently, two MCs interact to form a transpososome. This interaction is driven by *trans* contacts between protein and DNA, and protein-protein contacts. In an alternative model (not shown) transposase is transferred to a transposon end sequence in dimer form after initially pairing two non-transposon end DNA segments. The dimer-bound transposon end then captures a second transposon end (28, 29).

shown to increase the transposition frequency substantially (12- to 45-fold) by increasing transpososome formation (30).

With the above considerations in mind, a highly active *in vitro* transposition system was developed by employing an optimized end sequence (i.e. the ME) and a hyperactive (HA) transposase (31). This HA transposase contains three mutations: one at position 372 that prevents the inhibitory N- and C-terminal domain interaction (LP372); one at position 56 that prevents synthesis of the inhibitor protein (MA56); and a mutation at position 54 that enhances the transposase-end interaction (EK54) (24, 32, 33). This system was employed to generate X-ray crystal structures of the Tn5 transpososome, which have provided a wealth of information regarding the protein-DNA and protein-protein interactions that govern transpososome formation (34, 35). These structures, as well as footprinting studies (19), reveal that portions of the DNA within the Tn5 transpososome have a distorted DNA structure making this complex a potential substrate for H-NS binding.

We show in the current work that H-NS binds to the Tn5 transpososome *in vitro* and that under certain conditions H-NS can stimulate transpososome formation. We also show that H-NS is able to bind to the MC, an observation that is consistent with H-NS acting at an early stage to facilitate transpososome formation. Finally, we show that *in vivo*, the presence of H-NS stimulates the level of Tn5 transposition.

3.2.0 Materials and methods

Chemicals and oligonucleotides were from Sigma-Aldrich. BCA reagent was from Pierce. Growth media was from Becton Dickinson. Enzymes were from New England

Biolabs. Radio-nucleotides were from Amersham Biosciences.

3.2.1 Strains and plasmids

Escherichia coli strains DH10B (36) and NK5830 (37) were transduced to Δ*hns* by generalized transduction using phage P1 (38) and M182Δ*hns*::Kan^R (39) as described in (7). MDW320 (21) contains a Tn5-derived transposon with a promoterless *lacZ* gene inserted into the F' plasmid pOX38-Gen and this strain was mated to DH10B and DH10BΔ*hns*::Kan^R to generate strains CRW31 and CRW51, respectively.

Plasmid pDH551 encodes MA56/LP372 transposase and was constructed by ligating an 8078 bp NheI-MfeI backbone fragment from pGRTYB35 (40) to the 815 bp NheI-MfeI fragment from pRZ9905 (41). pDH389 encodes a mini-Tn5-Tet^R element and was constructed by cloning a BamHI fragment containing Tn10 Tet^R from pNK861 (42) into BamHI digested pMOD-2 (Epicenter). pDH390 encodes an arabinose-inducible Tn5 transposase gene and was constructed by cloning a ClaI-SalI fragment containing AraC-AraOp-Tn5 HA (hyperactive) transposase from pGRAra2 into ClaI-SalI digested pACYC184. pGRAra2 was derived from pBAD18 (43). pDH508 encodes a mini-Tn5-Cm^R element and was constructed by cloning a BamHI fragment from pNK1210 (44) containing the Cm^R gene of IS903 into BamHI digested pMOD-2. pDH533 encodes Tn5 transposase (MA56) under the control of its native promoter and a mini-Tn5-Cm^R element. This plasmid is a derivative of pRZ9905 in which the PshAI fragment encoding the mini-Tn5-Cm^R gene from pDH508 was ligated into the filled-in BglII site of pRZ9905. pDH548 is a derivative of pRZ9905 in which HA transposase was substituted for the MA56 transposase. This was accomplished by cutting pRZ9905 with MfeI and

BsaAI and ligating the 3331 bp backbone to a 1052 bp BsaAI-MfeI fragment encoding HA transposase obtained from pGRTYB35. pDH549 is a derivative of pDH548 in which the mini-Tn5-Cm^R PshAI fragment of pDH508 was ligated into the filled-in BglII site of pDH548. pDH550 is a derivative of pDH549 encoding MA56/LP372 transposase and was constructed by ligating the large NheI-MfeI backbone fragment from pDH549 to the 815 bp NheI-MfeI fragment from pRZ9905.

3.2.2 Protein purification

Tn5 HA transposase and derivatives were over-expressed as C-terminal chitin binding domain fusions and purified using the IMPACT system (New England Biolabs) as described in (45) and determined to be over 98% pure based on SDS-PAGE analysis (see Figure 3.4, lane 2). H-NS (WT and P116S) was over-expressed from pET3a- H-NS in BL21(DE3)*hns*::Kan^R and purified to 98% purity as described in (39) (see Figure 3.4, lane 3). StpA was over-expressed from pET3a-StpA in BL21(DE3)*hns*::Kan^R and purified as described in (39). Protein concentrations were determined by Bradford assay for Tn5 transposase and by BCA assay for H-NS and StpA.

3.2.3 Formation of transposition complexes

DNA substrates for the assembly of transposition complexes were generated by annealing the following 5’ ^{32}P-labeled, gel-purified oligonucleotides: 20 bp ME-5’-CTGTCTCTTATACACATCTT-3’ (NTS)/5’-AAGATGTGTATAAGAGACAG-3’ (TS); 53 bp ME (WT)-5’-CCCTGCAGGTCGACTGTCTCTTATACACATCTTGAGTGAGTGAGCATGCATGT

- 3’ (NTS)/5’-
ACATGCATGCTCACTCACTCAAGATGTGTATAAGAGACAGTCGACCTGCAGG
G-3’ (TS); 53 bp ME (3/6)-5’-
CCCTGCAGGTCGACTCTCACTTATACACATCTTGAGTGAGTGAGCATGCATGT
- 3’ (NTS)/5’ –
ACATGCATGCTCACTCACTCAAGATGTGTATAAGTGAGAGTCGACCTGCAGG
G-3’ (TS). Transposition complexes were assembled by mixing oligonucleotide substrates (0.05 µM) with transposase (0.4 µM) and where indicated, variable amounts of H-NS or StpA in Hepes-KCl buffer for 1 hour at 37°C as previously described (45). Where H-NS was added subsequent to transposase, ME DNA and transposase were first incubated for 1 hour at 37°C and then H-NS was added and incubation was carried out for an additional 30 minutes at 25°C. When heparin sulphate was present in assembly reactions it was added before transposase, H-NS or StpA to a final concentration of 2 µM. Samples were mixed with non-denaturing load dye and subjected to electrophoresis on a 5% TAE gel as previously described (46). Gel images were obtained by phosphorimaging using the Storm 680 PhosphorImager (Molecular Dynamics) and species levels were quantified using ImageQuant software (Molecular Dynamics).

3.2.4 Stoichiometric analysis

Large-scale transpososome assembly reactions (275 µl) were set up with H-NS at 0.36 and 1.4 µM, as described above. Reactions were concentrated using Vivaspin 3 kDa cut-off filters (Vivascience) to approximately 65 µl. Samples were loaded in duplicate onto a 5% native polyacrylamide gel. After electrophoresis a portion of the gel with one set of

samples was stained with ethidium bromide and this stained gel was used as a guide to identify the position of H-NS-transpososomes in the unstained portion of the gel; we avoided staining the analytical portion of the gel because we found that ethidium bromide can cause proteins to dissociate from DNA. Proteins were eluted out of gel slices containing H-NS-transpososomes by incubating gel slices in 1ml of elution buffer (0.1% SDS, 250mM NH4Ac) for 16 hours at 42°C. Eluted proteins were then concentrated to approximately 15 ml using Vivaspin 3 kDa filters and the entire sample was subjected to SDS-PAGE using a 12% gel. Proteins were visualized by Coomassie blue staining and levels quantified using the AlphaImager (AlphaInnotech).

3.2.5 Hydroxyl radical footprinting

Transpososomes were assembled as previously described except that the 53 bp ME contained a 5' ^{32}P-label on either the transferred (TS) or non-transferred (NTS) strand and H-NS was added at a concentration of 0.6 or 1.8 μM (approximately equivalent to reactions in lanes 8 and 10 of Figure 3.2B). Hydroxyl radical treatment of assembly reactions was performed essentially as described (19), except that the chemical treatment was carried out at 25°C for 10 minutes and then samples were directly applied to a 5% native polyacrylamide gel. Transpososomes, H-NS-transpososomes and unbound ME DNA were isolated from the wet gel after exposure of the gel to film. DNA was eluted from gel slices in elution buffer (0.1% SDS, 250 mM NH_4Ac) for 16 hours at 42°C. Eluted DNA was then treated with phenol, and ethanol precipitated. The purified DNA was resuspended in water and denaturing loading dye and the volume adjusted so that 1 μl of each sample contained approximately equal numbers of radioactive counts. Samples

were then applied to a 10% high-resolution denaturing gel using 1xTBE buffer, along with a Maxim and Gilbert 'G' ladder prepared from the appropriate ME DNA, and a gel image was obtained using phosphorimaging as described previously. Protected bands and bands of hypersensitivity in the DNA upon addition of H-NS to the transpososome were determined by using densitometry to graph the band intensities in both the unbound transpososome and the H-NS-bound transpososome resulting DNA cleavages, aligning the graphs with the unreacted ME substrate length and comparing peak/band intensities. A protected phosphate has a weaker band intensity compared to the control and a hypersensitive phosphate has a higher band intensity compared to the control.

3.2.6 *In vitro* excision assay

Transposition time courses following donor cleavage were carried out by initially forming transpososomes with or without HNS (2 µM) and/or heparin (2 µM) as described in the 'Materials and methods'. $MgCl_2$ (10 mM) was added to each transpososome sample and the reactions were incubated at 37°C for up to 16 hours. At the indicated time points (0, 1, 3, 16 hours) equal volume aliquots were removed, treated with phenol and the DNA was recovered by ethanol precipitation. DNA was resuspended in denaturing loading dye and analyzed on an 8% high-resolution denaturing gel as above. The amount of transpososome formed per reaction condition was quantified using ImageQuant software (Molecular Dynamics) based on the percentage of transpososome signal compared to the total amount of signal in the entire lane. Similarly, ImageQuant was used to quantify the amount of cleavage products generated in each reaction at each time point as a percentage of the total signal in the entire lane.

3.2.7 *In vivo* transposition assays

The papillation assay was carried out by transforming CRW31 (*hns*+) and CRW51 (Δ*hns*) with pGRPET2 (Ap^R) and selecting for transformants on MacConkey lactose plates containing 100 µg ml^{-1} ampicillin. pGRPET2 (31) is a pET-21d derivative encoding HA transposase under the control of a T7 promoter and was used to achieve a relatively low level of transposase expression *in vivo* in strains lacking the T7 RNA polymerase gene. Plates were incubated for up to 4 days at 37°C. The mating out assay where transposase was provided in c*is* was performed as described in (47) except that donor transformants (pDH533 or pDH549 in NK5830 and NK5830Δ*hns*::Kan^R) were selected on M9 plates supplemented with glucose, arginine, ampicillin (50 µg ml^{-1}) and chloramphenicol (20 µg ml^{-1}). In addition, HB101 was used as the recipient and mating mixtures were pelleted and resuspended in 0.85% saline before plating on M9 plates supplemented with glucose, leucine and streptomycin (150 µg ml^{-1}) for total exconjugants, or glucose, leucine, streptomycin (150 µg ml^{-1}) and chloramphenicol (20 µg ml^{-1}) for transposition events. For the mating out assay where transposase was provided in *trans* pDH389 and pDH390 were co-transformed into NK5830 and NK5830Δ*hns*::Kan^R and transformants were selected as above. Transformants were grown overnight in SOC media to inhibit transposase synthesis. Cells were then pelleted, washed with saline and subcultured (1:20 dilution) into 1 ml of M9 media supplemented with arginine and 0.2% arabinose; growth in arabinose induces transposase synthesis. Donors were then grown for 4 hours on fast roll followed by 4 hours on slow roll. After mixing donor cultures (1 ml) with 2.5 ml of HB101 recipient cells (OD_{600}~0.6), cultures

were grown for an additional hour on slow roll. Mating mixtures were then processed as above and plated on M9 plates supplemented with glucose, leucine and streptomycin (150 µg ml^{-1}) for total exconjugants, or glucose, leucine, streptomycin (150 µg ml^{-1}) and tetracycline (10 µg ml^{-1}) for transposition events.

3.3.0 Results

3.3.1 H-NS binds with high specificity to the Tn5 transpososome

To determine if H-NS binds the Tn5 transpososome, we performed electrophoretic mobility shift assays (EMSA) with purified H-NS and mixtures containing short linear ^{32}P-labeled ME DNA and the HA (hyperactive) transposase. In one arrangement we incubated H-NS, HA transposase and ME DNA (ME+T'ase+H-NS) and in another we pre-mixed HA transposase and ME DNA before adding H-NS (ME+T'ase/+H-NS). In the former arrangement there is the potential for H-NS to bind either the transpososome or a pre-transpososome MC (monomer complex). We also performed incubations with H-NS and only the ME (ME+H-NS). Also, the size of the ME DNA was varied in different experiments in order to determine the minimal end sequence requirements for H-NS binding.

The results in Figure 3.2A show that with a 20 bp ME, H-NS binds to the Tn5 transpososome irrespective of whether the transpososome was formed before or after H-NS addition (compare lane 6 with lanes 7–10 and lane 11 with lanes 12–15). Based on our titration data (Figure 3.2 and data not shown) we estimate that a molar excess of H-NS to transpososome of 20-fold is optimal for the formation of the H-NS-transpososome.

As the concentration of H-NS present in the reaction increased the mobility of the H-NS-transpososome further decreased, implying that at the lower H-NS concentrations H-NS binding sites within the transpososome were not saturated; note that additional binding sites might be located within the DNA or the protein components of the transpososome – the latter might include H-NS oligomerization. Operationally, we refer to the reduced mobility forms of the H-NS-transpososome as 'supershifted' forms. Notably, at the same range of H-NS concentrations used to detect H-NS binding to the transpososome, we did not detect H-NS binding to free ME DNA (lanes 1–5). These results indicate that H-NS binds the transpososome with considerably higher affinity than it binds unbound ME DNA. These results also show that a transpososome formed with the least amount of ME sequence required for transposase binding (20 bp) contains sufficient determinants for H-NS binding. ME DNA was also substituted in these reactions with an authentic OE DNA substrate and similar H-NS binding was observed (data not shown).

Similar results were obtained with a 53 bp substrate containing 40 bp of ME DNA and 13 bp of flanking donor DNA (Figure 3.2B); although for this substrate there is some indication that the order of H-NS addition did have an impact on the amount of H-NS-transpososome formed. When ME DNA, transposase and H-NS were added at the same time there was as much as a 2-fold increase in the amount of H-NS-transpososome formed versus when H-NS was added after ME and transposase (compare lanes 7–8 and 12–13 in Figure 3.2B). This raises the possibility that H-NS is capable of binding the MC and facilitating transpososome assembly (see also Figures 3.5 and 3.7).

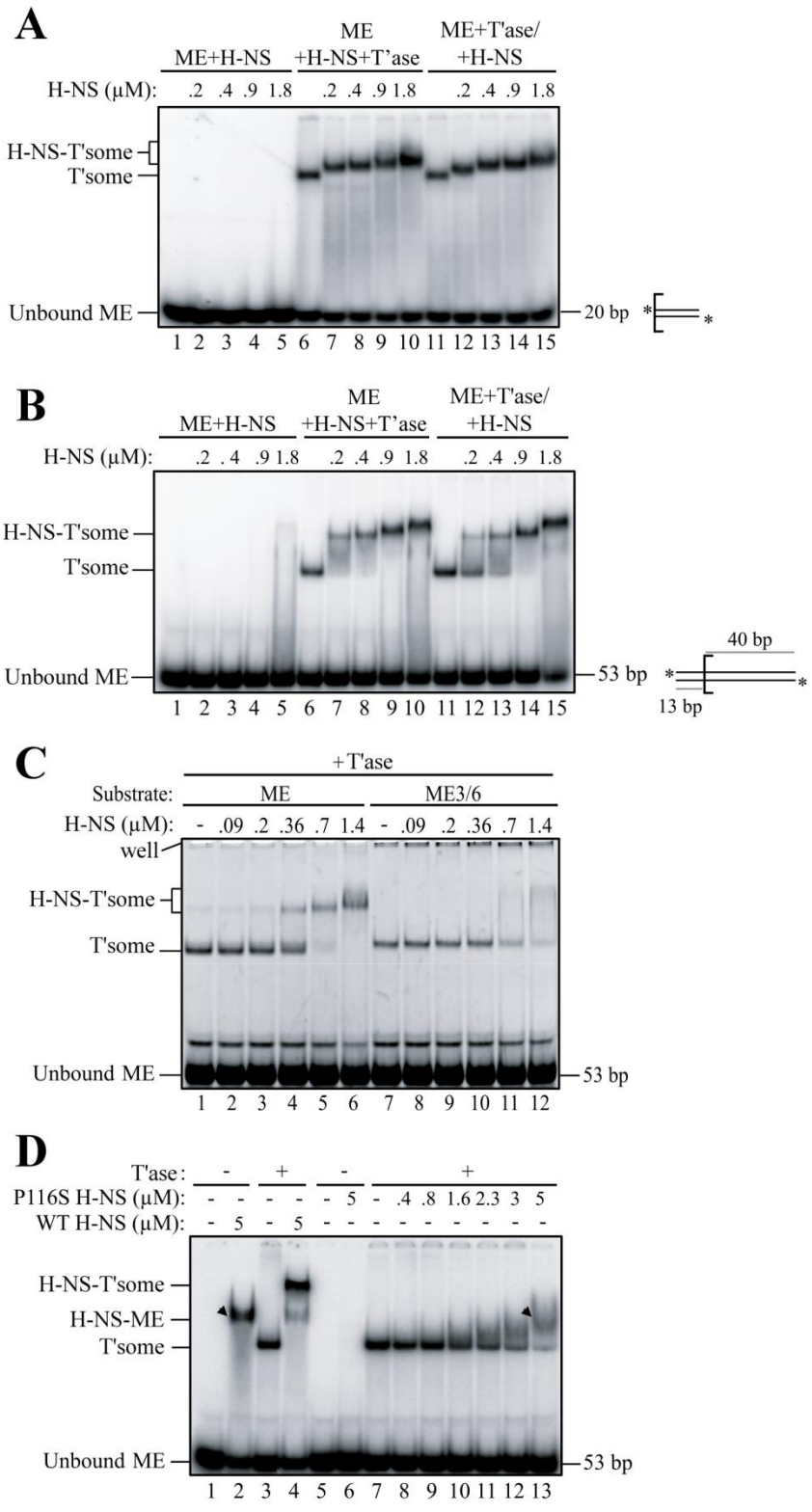
Strikingly, there was a significant reduction in the ability of H-NS to bind a transpososome formed with the 53 bp ME DNA containing mutations at positions 3 and 6

(compare lanes 4–6 to 10–12 in Figure 3.2C). Positions 3 and 6 have been shown to contribute to *trans* contacts with transposase (34) and accordingly we found that these mutations reduce the amount of transpososome formed by about 50%. That these mutations reduce the binding of H-NS to the transpososome is suggestive of the region immediately adjacent to the transposon terminus (i.e. where *trans* contacts occur with the transposase) providing important determinants for H-NS binding.

A mutant form of H-NS (P116S H-NS) has been identified that has lost the capacity to bind DNA in a structure-specific manner but retains non-specific DNA binding (48, 49). We used this mutant to ask if H-NS binding to the Tn5 transpososome involves a DNA structure-specific interaction between H-NS and the transpososome. We show in Figure 3.2D that P116S H-NS has a greatly reduced binding affinity for the Tn5 transpososome, as even at the highest concentration of P116S H-NS used (500-fold molar excess relative to the transpososome), only a small mobility shift was observed (compare lanes 2 and 13); this species has a greater gel mobility relative to the H-NS-transpososome generated with wild-type H-NS (lane 4) and is only observed at high H-NS concentrations (>5 μM) in our standard reaction conditions. This result supports the idea that H-NS recognizes specific structural determinants in the Tn5 transpososome that are not present in the unbound ME DNA. An analogous experiment in the Tn10 system led to the same conclusion (10).

Figure 3.2. Electrophoretic mobility shift assays with Tn5 transpososomes and H-NS.

(A) and (B) H-NS mobility shifts of transpososomes formed with 20 and 53 bp ME substrates, respectively. Where indicated, H-NS was added to transpososome assembly reactions either at the same time as transposase (ME+T'ase+H-NS) or after transposase (i.e. subsequent to transpososome formation) (ME+T'ase/+H-NS). An illustration of the respective substrates is shown beside the gel image; the bracket defines the beginning of the ME sequence and asterisks indicate the position of ^{32}P-labels. H-NS-T'some, H-NS- bound transpososome; T'some, transpososome. (C) H-NS mobility shift of a transpososome formed with a ME substrate containing a G to C and a T to A mutation at positions 3 and 6, respectively (ME3/6). (D) H-NS mobility shift with P116S H-NS. The arrowheads show the mobility of a non-specific H-NS-ME complex (lane 2) and a P116S H-NS-transpososome complex (lane 13). H-NS binds the ME in the absence of transposase only at relatively high (>5 µM) H-NS concentrations. Similarly, P116S H-NS binds to the transpososome only at relatively high concentrations (>3 µM). Note that in (A)-(D) HA transposase was used.

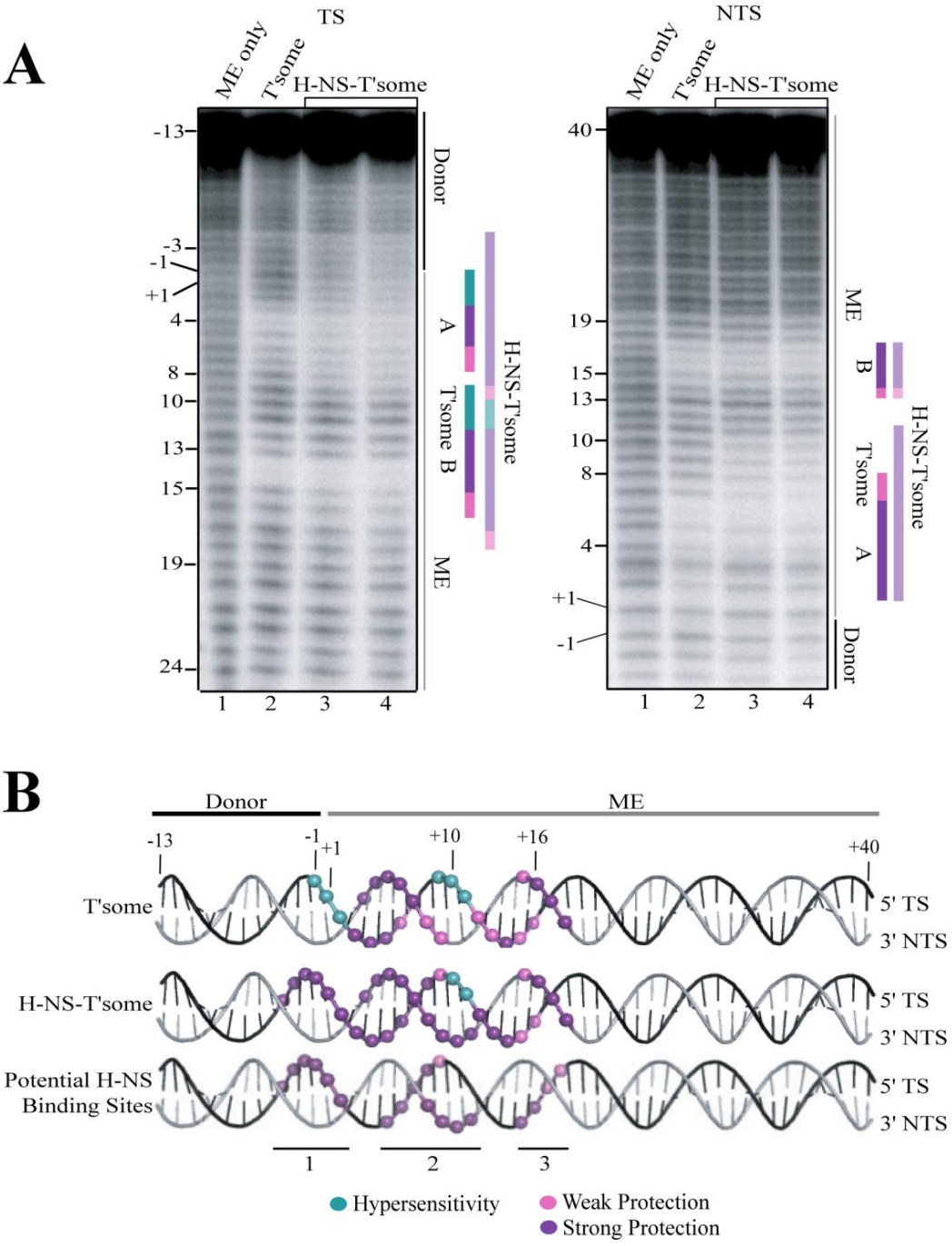

A
ME+H-NS
ME +H-NS+T'ase
ME+T'ase/ +H-NS
H-NS (μM): .2 .4 .9 1.8 .2 .4 .9 1.8 .2 .4 .9 1.8
H-NS-T'some
T'some
Unbound ME
20 bp
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
B
ME+H-NS
ME +H-NS+T'ase
ME+T'ase/ +H-NS
H-NS (μM): .2 .4 .9 1.8 .2 .4 .9 1.8 .2 .4 .9 1.8
H-NS-T'some
T'some
Unbound ME
53 bp
40 bp
13 bp
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
C
+T'ase
Substrate: ME ME3/6
H-NS (μM): - .09 .2 .36 .7 1.4 - .09 .2 .36 .7 1.4
well
H-NS-T'some
T'some
Unbound ME
53 bp
1 2 3 4 5 6 7 8 9 10 11 12
D
T'ase: - + - +
P116S H-NS (μM): - - - - - 5 - .4 .8 1.6 2.3 3 5
WT H-NS (μM): - 5 - 5 - - - - - - - - -
H-NS-T'some
H-NS-ME
T'some
Unbound ME
53 bp
1 2 3 4 5 6 7 8 9 10 11 12 13

3.3.2 Hydroxyl radical footprinting of H-NS-transpososomes

We characterized the protein-DNA contacts within the H-NS-transpososome by performing hydroxyl radical DNA footprinting as described in the 'Materials and methods'. Incorporation of H-NS into the transpososome significantly altered the hydroxyl radical footprint as regions of hydroxyl radical protection typically observed in the transpososome were expanded, and regions of hypersensitivity were protected from cleavage (Figure 3.3A). In our transpososome footprint there are two zones of protection attributable to transposase: zone A includes nucleotides 3–7 on the TS and nucleotides 2–7 on the NTS; zone B includes nucleotides 12–16 on the TS and nucleotides 14–17 on the NTS. As well, residues 2-(-1) and 9–11 on the TS show hypersensitivity to hydroxyl radicals. These results match well with the results of previously reported hydroxyl radical footprinting experiments performed on the Tn5 transpososome and contact data inferred from the Tn5 transpososome crystal structure (19, 34). In the H-NS-transpososome footprints, protection in zone A was extended in both directions to include residues 3-(-4) and 8–9 on the TS and 8–11 on the NTS. This includes loss of hydroxyl radical hypersensitivity at residues 2-(-1) and residue 9 of the TS. In addition, protection in zone B was extended from residues 17–18 on the TS. Overall, the footprinting data suggests that H-NS may have as many as three distinct binding sites within the transpososome (Figure 3.3B). Notably, estimates from stoichiometric analysis of the H-NS-transpososome formed at 1.4 μM H-NS (similar to the H-NS-transpososome formed for each footprint shown in lanes 4 of Figure 3.3A) suggest that four dimers of H-NS may be present per transpososome (Figure 3.4).

Figure 3.3. Hydroxyl radical footprinting of H-NS-transpososomes.

(A) Transferred (TS) and non-transferred (NTS) strand footprints are shown. Transpososome assembly reactions were set up as described in the 'Materials and methods' with the 53 bp ME substrate (Figure 3.2B) labeled with ^{32}P at the 5' terminus of either the TS or the NTS. Samples in lanes 3 and 4 were from H-NS-transpososome assembly reactions performed with 0.6 and 1.8 µM H-NS, respectively (these reflect the H-NS mobility shifts of lanes 8 and 10, respectively in Figure 3.2B). The +1 position is the first base pair of the ME and the –1 position is the first base pair of the flanking donor DNA. Regions of strong and weak hydroxyl radical protection are indicated by purple and pink vertical bars, respectively. Regions of hydroxyl radical hypersensitivity are represented by vertical green bars. Binding sites of transposase, referred to in the text as zones A and B, are labeled. (B) Positions of hydroxyl radical protection and hypersensitivity are summarized on helical representations of DNA for the transpososome and H-NS-transpososome. The three potential binding sites for H-NS (sites 1–3) are shown on the bottom helical representation. These sites are defined by differences in hydroxyl radical protection and hypersensitivity in the transpososome versus the H-NS-transpososome.


A
TS
ME only
T'some
H-NS-T'some
-13
-3
-1
+1
4
8
10
13
15
19
24
Donor
A
T'some B
ME
H-NS-T'some
1 2 3 4
NTS
ME only
T'some
H-NS-T'some
40
19
15
13
10
8
4
+1
-1
ME
B
T'some
A
H-NS-T'some
Donor
1 2 3 4
B
Donor
ME
-13
-1
+1
+10
+16
+40
T'some
H-NS-T'some
Potential H-NS Binding Sites
5' TS
3' NTS
1
2
3
Hypersensitivity
Weak Protection
Strong Protection

3.3.3 H-NS facilitates transpososome assembly

Over the course of studying the interactions between H-NS and the Tn5 transpososome we frequently observed an increase in the yield of transpososome when H-NS was included in transpososome assembly reactions as opposed to being added to a mixture of pre-formed transpososome (compare lanes 6 and 10 in Figure 3.2B). This led us to speculate that H-NS might increase the efficiency of transpososome formation. We also reasoned that it might be difficult to observe a strong effect of H-NS on transpososome assembly under our standard reaction conditions because the *in vitro* Tn5 system has been optimized for transpososome assembly (see 'Introduction'). We therefore asked if H-NS might have a more substantial impact on transpososome assembly under conditions where it is more difficult for the transpososome assembly to take place.

One way we reduced the efficiency of transpososome formation was by performing assembly reactions in the presence of heparin, a low molecular weight polyanion. Heparin is expected to compete with the ME DNA for binding of transposase and therefore its addition should reduce transpososome formation. Consistent with this expectation, we show in Figure 3.5A that addition of heparin reduced transpososome formation by 22-fold (compare lanes 3 and 10). When the identical reaction was carried out in the presence of H-NS, the efficiency of transpososome formation was greatly increased. At the highest concentration of H-NS used, transpososome formation was increased 11-fold relative to the heparin-treated sample that did not receive H-NS (compare lanes 10 and 14). In contrast, in the absence of heparin the increase in

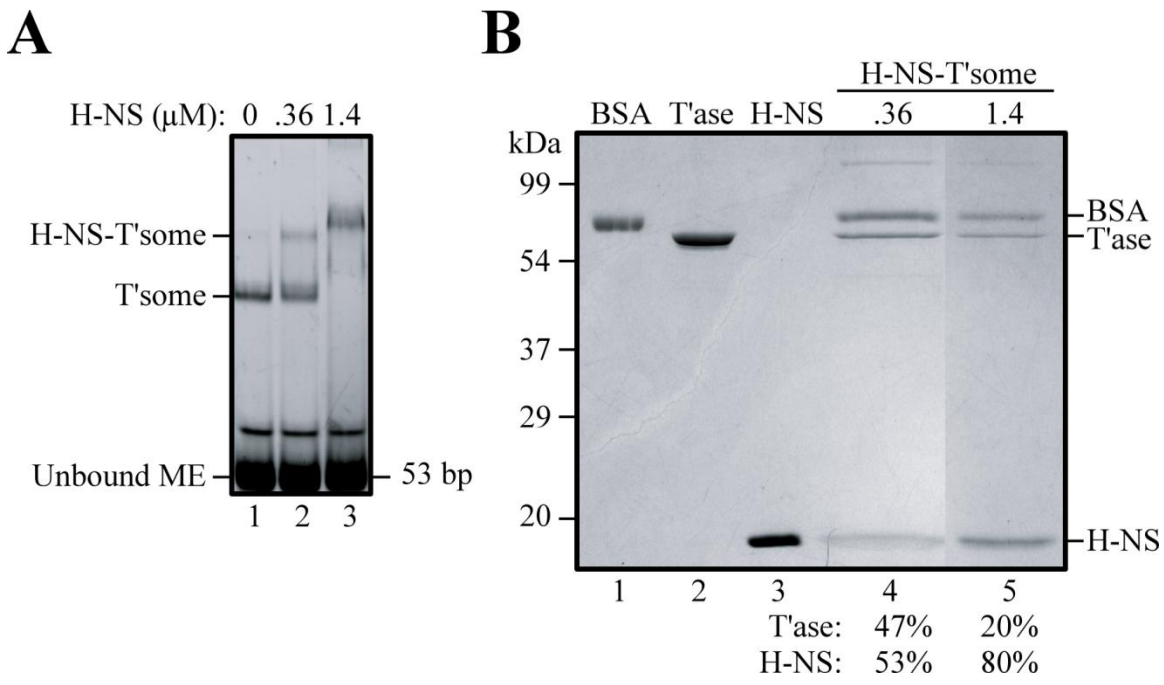


Figure 3.4. Stoichiometric analysis of H-NS-transpososomes.

(A) H-NS-transpososomes were formed as in Figure 3.2B, except that the ME DNA was not labeled with ^{32}P. (B) Gel slices from the gel shown in (A) containing H-NS-transpososome formed with two different amounts of H-NS were obtained and proteins were eluted from the slices as described in the 'Materials and methods'. Eluates were analyzed by SDS-PAGE and proteins detected by Coomassie Blue staining. Note that BSA is present in the transpososome assembly reaction and some BSA co-purifies with the H-NS-transpososomes. Samples in the first three lanes include purified BSA, Tn5 transposase and H-NS, respectively. The amounts of H-NS used to form the H-NS-T'some purified for SDS-PAGE gel analysis are indicated above the gel and are in μM amounts. The relative protein composition of the H-NS-transpososomes, calculated from the quantified band intensities of H-NS and transposase species in the 'H-NS-T'some' lanes (lanes 4 and 5) of the gel image in (B), are indicated as percentages of total protein within the respective H-NS-transpososomes.

transpososome formation was less than 2-fold (compare lanes 3 and 7). Moreover, the vast majority of the transpososome detected in reactions with heparin were in the H-NS-bound form. This is suggestive of H-NS increasing transpososome assembly by binding directly to the transpososome (as opposed to titrating out heparin and thus acting in an indirect manner).

Further evidence that H-NS is acting directly to increase transpososome formation comes from two additional experiments. In one experiment, we used StpA instead of H-NS to look for stimulation of transpososome assembly in the presence of heparin. StpA is a paralogue of H-NS, sharing 58% amino acid identity in *E. coli* (39). However, StpA is more basic than H-NS with a predicted pI of 9.08 compared to 5.25 for H-NS, and thus is expected to bind significantly more heparin relative to H-NS (39, 50). If an increase in transpososome formation was simply the result of H-NS titrating heparin out of the reaction, then StpA should be more effective than H-NS at promoting transpososome formation. However, when StpA was added instead of H-NS, only a moderate increase in transpososome formation was observed (3-fold as opposed to 17-fold when H-NS was used in this experiment), suggesting that the H-NS effect is direct (Figure 3.5B). It is likely that the relatively small StpA enhancement in transpososome formation observed in this experiment is indirect because only a small percentage (approximately 5%) of the transpososome formed in the StpA reaction was in StpA-bound form.

In a second experiment we asked if H-NS could rescue transpososome formation in a reaction where ME3/6 (described in Figure 3.2C) was used as substrate. Recall that the ME3/6 transpososome was largely defective for binding H-NS. We failed to see a

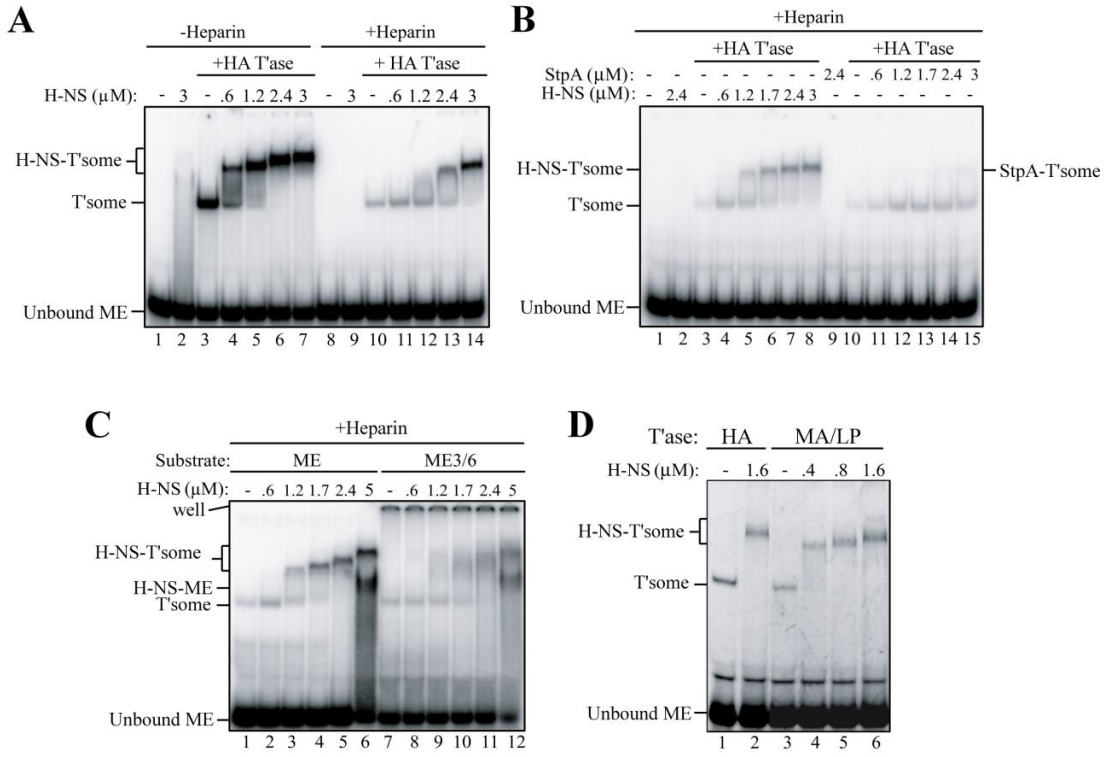


Figure 3.5. Effect of H-NS on transpososome formation under heparin challenge or in the presence of MA/LP transposase.

In (A) and (B) transpososome assembly reactions were carried out with ^{32}P-labeled 53 bp ME substrate following the 'ME+T'ase+H-NS' regimen (see Figure 3.2) in the presence and absence of heparin. Note that in (B) either H-NS or StpA were added to the reactions. The same reaction set-up, including HA transposase, was used in (C) except that in addition, the ME3/6 substrate described in Figure 3.2C was used. (D) H-NS was added to transpososome assembly reactions carried out with HA transposase and MA/LP transposase. Note that the only difference between these proteins is that HA transposase contains the EK mutation and MA/LP transposase does not. Species are labeled as in Figure 3.2. StpA-T'some, StpA-bound transpososome.

significant increase in transpososome formation (<2-fold) in this reaction (Figure 3.5C), suggesting that H-NS binding to the transpososome is required for the stimulation in transpososome formation seen in the presence of H-NS under heparin competition conditions.

An alternative approach to reducing the efficiency of transpososome formation was to utilize a form of transposase that contains the wild-type residue at position 54, but still contains the MA56 and LP372 mutations, hereafter referred to as MA/LP transposase. The glutamate at position 54 hinders transposase-end interactions and impedes transpososome formation. The efficiency of transpososome assembly was reduced 3-fold when this form of transposase was used (compare lanes 1 and 3 in Figure 3.5D). In the presence of H-NS the efficiency of transpososome formation in the MA/LP reaction was increased by up to 2.5-fold (compare lanes 3 and 6), which coincided with the formation of H-NS-transpososome.

Taken together, the results in this section show that H-NS can stimulate transpososome formation and that this stimulation correlates with H-NS binding to the transpososome. It is also apparent that the region where transposase makes contacts in *trans* with the ME is not only critical for H-NS binding but also important in H-NS-directed stimulation of transpososome formation.

The above stimulation of transpososome formation by H-NS is only functionally significant if the H-NS bound form of the transpososome is capable of undergoing the chemical steps in transposition. We tested this by adding $MgCl_2$ to a reaction equivalent to that shown in lane 14 of Figure 3.5A (i.e. a transpososome assembly reaction carried out in the presence of heparin and H-NS). Upon incubating this reaction for varying

amounts of time and then analyzing the DNA on a denaturing gel, we detected the expected distribution of transposon excision products (Figure 3.6). Furthermore, the efficiency of excision product formation, per amount of transpososome formed, was comparable with that observed in a control reaction performed without heparin and H-NS. Thus, we conclude that H-NS does not have an inhibitory effect on transposon excision in the Tn5 system.

3.3.4 H-NS binds to a Tn5 MC

We can think of two general ways in which H-NS could promote transpososome formation in the Tn5 system. H-NS might increase the stability of the transpososome or it might play a role in transpososome assembly. The latter possibility is much more likely because the Tn5 transpososome is very stable and over the time course of our experiments is not expected to dissociate to any significant degree (28, 51). We have tested the idea that H-NS facilitates transpososome assembly by asking if H-NS binds to the MC, a species that may either be a bona fide assembly intermediate or closely resemble such an intermediate (Figure 3.1 and see 'Discussion'). Under standard *in vitro* reaction conditions the MC is unstable and thus difficult to work with. In fact, analysis of this species requires the utilization of a mutant form of transposase (Δ369 EK/MA) that lacks 108 amino acids from the C-terminus, which includes the dimerization subdomain, and cannot go on to form a transpososome (23, 18). Importantly, formation of the MC with Δ369 EK/MA displays many of the DNA sequence requirements for transpososome formation and thus provides a reasonable model for the study of a transpososome precursor (19).

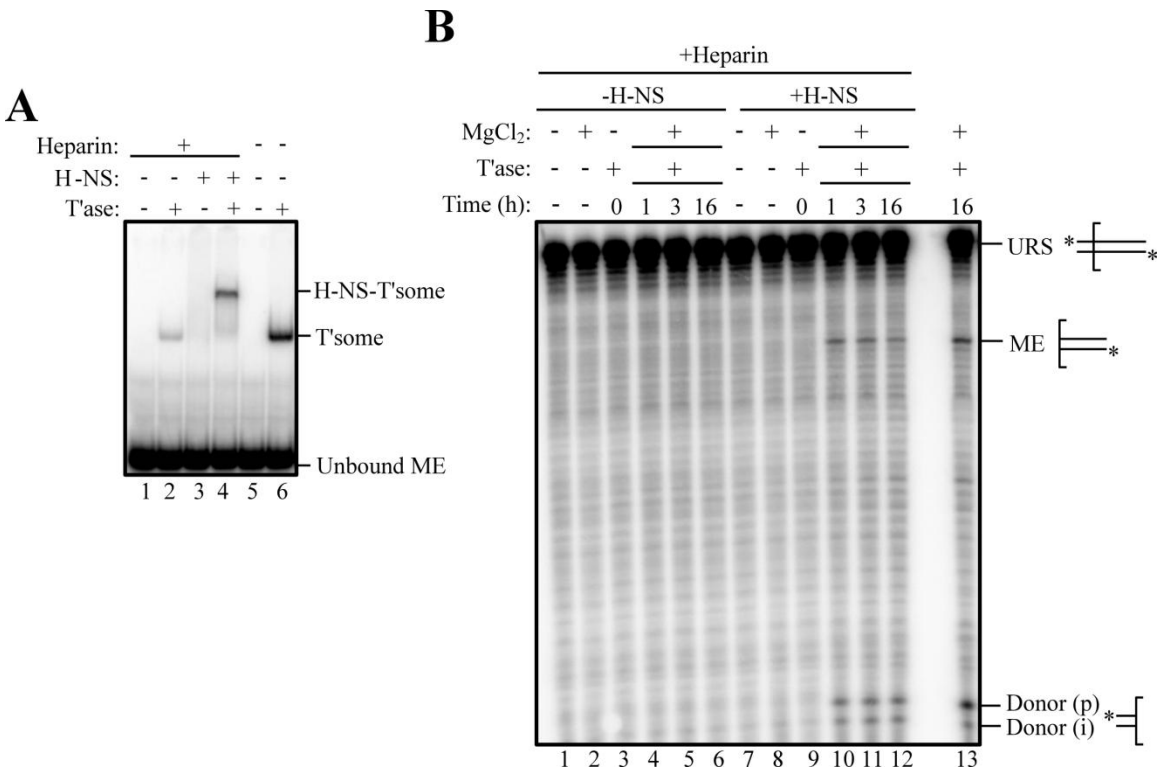


Figure 3.6. Donor cleavage assay for the H-NS-transpososome formed in the presence of heparin.

(A) Transpososome assembly reactions were carried out with and without H-NS in the presence of heparin as in Figure 3.5A. An additional reaction was also carried out in the absence of heparin and H-NS. A mobility-shift assay is shown. (B) An aliquot of each of the reactions shown in lanes 2, 4 and 6 of part (A) was mixed with $MgCl_2$ (final concentration 10 mM) for the indicated times to initiate donor cleavage. Reactions were terminated and the DNA was on a high resolution denaturing gel. The radioactive counts loaded per lane were approximately equivalent for the full set of samples. URS, unreacted substrate; ME, cleaved mosaic end; Donor (p), cleaved donor DNA from precise excision; Donor (i), cleaved donor DNA from imprecise excision (illustrations beside the gel image show each of these species in black lines). The transposon end hairpin species was resolved at the first time point and is therefore not shown on the gel.

In assembly reactions with Δ369 EK/MA a MC is readily detected (lane 3, Figure 3.7A). Addition of H-NS to the Δ369 EK/MA MC resulted in a mobility shift, clearly showing that H-NS is able to bind this species. At the lower H-NS concentrations the H-NS bound form of the complex constitutes a heterogeneous mixture of species as indicated by the presence of a smear above the MC (lanes 5–8). At the highest H-NS concentration (lane 9) there is a dramatic shift in the distribution of H-NS-bound MC towards a more homogeneous species. Given the extent of the mobility shift, which is almost as large as that observed for the transpososome versus free ME DNA (compare lanes 9 and 2), and the concentration dependence of this shift, it is likely that H-NS is forming higher order oligomers in this complex. Notably, H-NS binding to the MC occurs at concentrations of H-NS that appear to be insufficient to form a complex with the unbound ME. This implies that the MC has binding determinants for H-NS that are not present in unbound ME DNA. Consistent with this, we show in Figure 3.7B that P116S H-NS failed to bind to the MC.

3.3.5 Disruption of the *hns* gene reduces the frequency of Tn5 transposition *in vivo*

The observations presented thus far raise the possibility that H-NS may act as a positive regulator of Tn5 transposition *in vivo*. We have tested this idea by measuring the relative transposition frequency of Tn5 in isogenic *hns*+/Δ*hns* strains of *E. coli*. This was done using two well-established transposition assays, a papillation assay and a mating out assay. In the papillation assay used here, the experimental read-out for transposition is a '*lacZ* turn-on' event that results from mobilization of a Tn5 derivative encoding a truncated *lacZ* gene into an expressed gene. When plated on MacConkey lactose

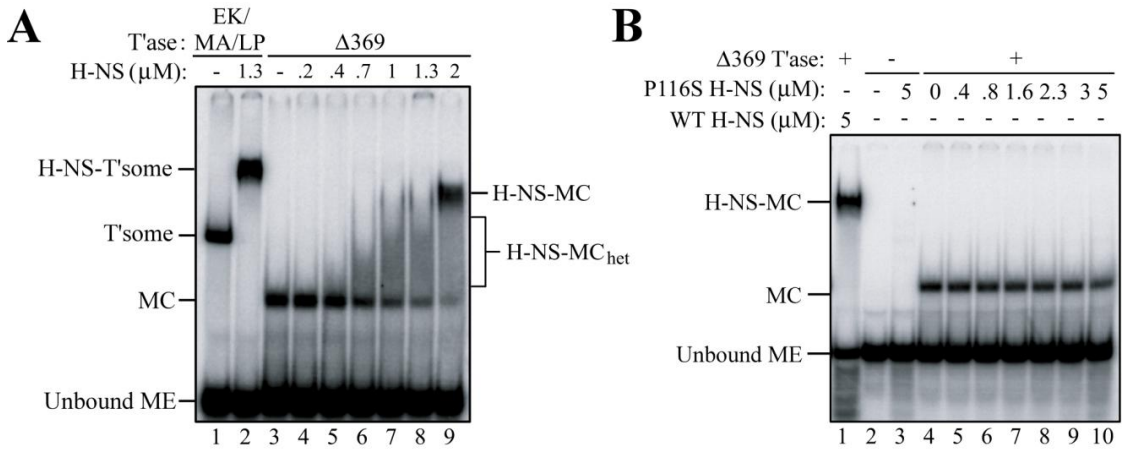


Figure 3.7. Electrophoretic mobility shift assays of transpososomes formed with Δ369 EK/MA transposase and H-NS.

(A) H-NS mobility shifts with transpososomes formed with either HA transposase or Δ369 EK/MA transposase. (B) Mobility shift with Δ369 EK/MA and either wild-type (WT) or P116S H-NS. H-NS-transpososomes in (A) and (B) were assembled under the 'ME+T'ase+H-NS' regimen, and the substrate is the same as in Figure 3.2B. Species are labeled as in Figure 3.2. MC, monomer complex; H-NS-MC, H-NS-bound monomer complex; H-NS-MC het, a heterogeneous mixture of H-NS-monomer complexes.

indicator plates, cells that have had a Tn5 transposition event form red papillae (or out-growths) on a background of *lacZ*- cells (white). The number of LacZ+ papillae formed per colony is roughly proportional to the frequency of Tn5 transposition within the colony (52). When we compared the average number of papillae formed in the *hns*+ and Δ*hns* strains expressing HA transposase encoded by a multicopy plasmid, we observed roughly 5-fold fewer LacZ+ papillae per colony in the Δ*hns* strain. Representative colonies are shown in Figure 3.8A. Note that we have compared the number of papillae formed in the two strains when colonies were of a similar size; the Δ*hns* strain grew at approximately half the rate of the *hns*+ strain. We also compared LacZ+ papillae formed in an Δ*hns* strain to those formed in the same strain but complemented with a plasmid-encoded source of *hns* under control of the *lac* promoter (Figure 3.8B). The representative colonies illustrate that transposition is enhanced 4- to 5-fold in the complemented Δ*hns* strain. Finally, when we transformed the *trans*-encoded *hns* source into an *hns*+ strain we saw a strong positive effect on transposition (Figure 3.8C).

In the mating out assay the frequency of transposon insertions into an F' plasmid is measured by mating a transposon-containing donor strain (F+) with a suitable recipient strain (F-) and then selecting for recipient cells that have acquired the transposon. A relative transposition frequency is then obtained by dividing the number of transposon-containing exconjugants by the number of total exconjugants. In one experiment, we transformed into isogenic *hns*+ and Δ*hns* strains a plasmid encoding a mini-Tn5-Tet^R element and a compatible plasmid encoding HA transposase under the control of an

Table 3.1. *In vivo* transposition of a mini-Tn5-Tet^{R} element catalyzed in *trans* by hyperactive transposase in *hns*+ and Δ*hns* strains.

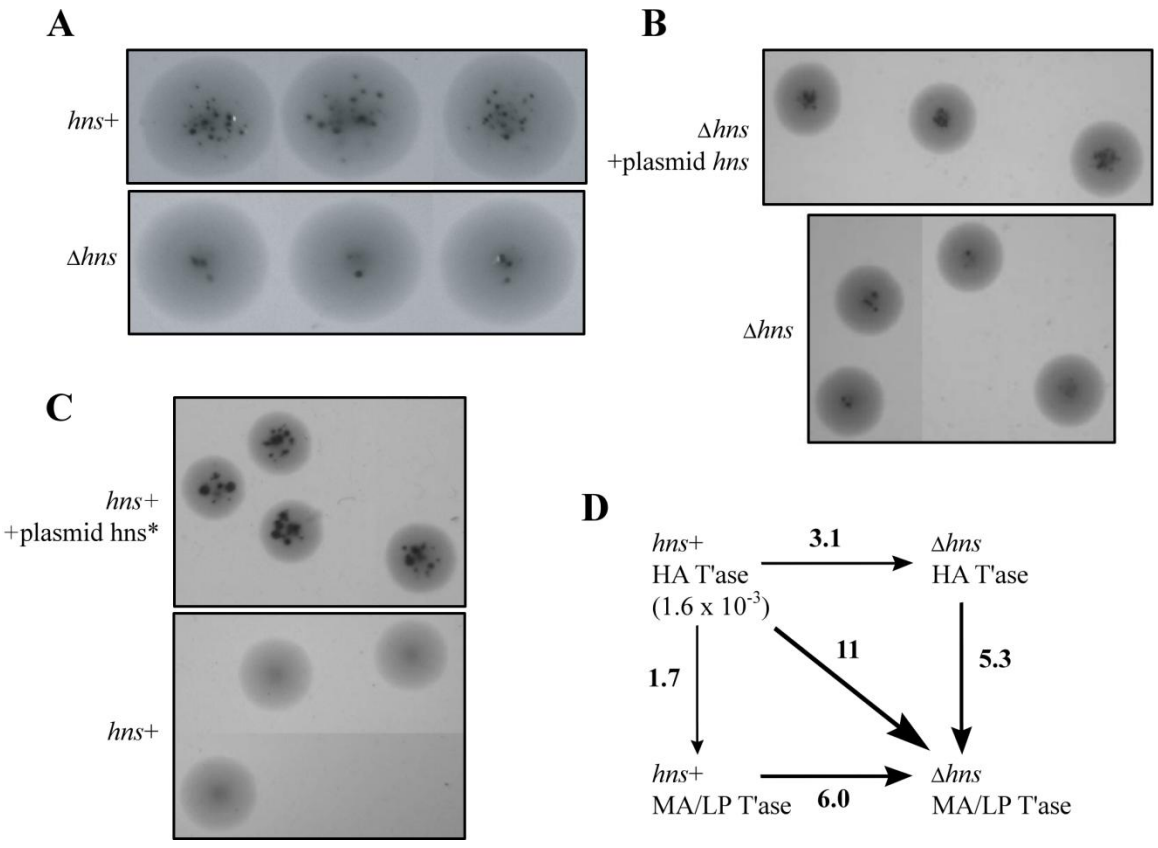
Transposase	Strain	Transposition Frequency	Normalized Frequency[a]
pDH390	NK5830F'	$8.1\pm2.0\text{x}10^{-2}$	1.0
pDH390	NK5830*hns*::Kan^{R}	$2.6\pm0.97\text{x}10^{-2}$	0.3

a. Relative transposition frequencies were calculated by dividing the number of SmRTetR colonies (transposition events) by the number of SmR colonies (total exconjugants) obtained per 0.1 ml of mating mix. Transposition frequencies represent an average value obtained from three independent experiments wherein matings with at least eight different donor transformants were carried out in each experiment.

b. Transposition frequencies were normalized to the level of transposition catalyzed by the HA transposase in the hns+ strain.

Figure 3.8. *In vivo* transposition assays in isogenic *hns*+ and Δ*hns E. coli* strains.

(A)-(C) Relative transposition frequency of mini-Tn5-*lacZ* in isogenic *hns*+ and Δ*hns* strains as measured using a papillation assay. (A) Three colonies of isogenic *hns*+ and Δ*hns* strains transformed with a plasmid encoding HA transposase are shown. The dark spots are LacZ+ papillae; the number of papillae formed per colony provides a measure of the relative frequency of transposition of a mini-Tn5-*lacZ* transposon within the colony. Note that the photographs were taken when colonies were roughly the same size, which corresponded to 50 and 90 hours of growth, respectively, for the *hns*+ versus the Δ*hns* strain. (B) Representative colonies from an *hns* 'complemented' papillation assay conducted with Δ*hns* strains transformed with HA transposase and either an empty vector (bottom) or a plasmid encoding *hns* under control of the *lac* promoter (top). Photographs were also taken when colonies were roughly the same size at approximately 80 hours growth. (C) Representative colonies from an *hns* over-expression papillation assay conducted with *hns*+ strains transformed with HA transposase and either an empty vector (bottom) or a plasmid encoding *hns* under control of the *lac* promoter. Photographs were also taken when colonies were roughly the same size, at 40 and 50 hours respectively. (D) Box diagram showing the fold-changes in transposition frequency in *cis* as measured by mating out for *hns*+ and Δ*hns* strains expressing HA transposase or MA/LP transposase. Transposition frequencies were calculated by dividing the number of Sm^RCm^R colonies (transposition events) by the number of Sm^R colonies (total exconjugants) obtained per 0.1 ml of mating mix. Transposition frequencies represent an average from at least two independent experiments with four donor transformants. The transposition frequency for the *hns*+/HA transposase is shown. Arrows point in the direction of the fold-decrease (e.g. *hns*+/HA transposase supported transposition at a 3.1-fold higher level relative to Δ*hns*/HA transposase).


A
hns+
Δhns
B
Δhns
+plasmid hns
Δhns
C
hns+
+plasmid hns*
hns+
D
hns+
HA T'ase
(1.6 x 10-3)
3.1
Δhns
HA T'ase
1.7
11
5.3
hns+
MA/LP T'ase
6.0
Δhns
MA/LP T'ase

arabinose inducible promoter. In this set-up transposase synthesis is induced by growth in arabinose. This system measures Tn5 transposition in *trans* since the mini-Tn5 element and the transposase source are on separate plasmids. The results presented in Table 3.1 show an increase in transposition of about 3-fold in the *hns*+ versus the Δ*hns* strain under conditions of arabinose induction.

We also tested the idea that utilization of HA transposase might make Tn5 transposition less sensitive to the H-NS status of the cell. We constructed plasmids containing either HA transposase (pDH549) or MA/LP transposase (pDH550), and a mini-Tn5-Cm^R element, and used these plasmids in a mating out experiment. In this case, transposase expression was under the control of the native Tn5 transposase promoter and transposition occurs in *cis* because the transposase and the mini-Tn5 element are on the same plasmid. The EK54 mutation was chosen for reversion to wild-type because we previously showed that H-NS can increase the level of transpososome formed *in vitro* with the MA/LP transposase (Figure 3.5D). The results show that the defect in transposition is twice as large (6-fold versus 3-fold) in the Δ*hns* versus the *hns*+ strain in the presence of MA/LP compared to the HA transposase (Figure 3.8D). This result is consistent with the possibility that H-NS partially compensates for the suboptimal transposase.

3.4.0 Discussion

We have shown that H-NS binds with high specificity to the Tn5 transpososome. Characterization of the H-NS interaction with the transpososome by DNA footprinting analysis provided evidence that H-NS binds in very close proximity to transposase and

stoichiometric analysis indicated that up to four H-NS dimers can bind the transpososome. H-NS binding to the Tn5 transpososome was shown to affect the *in vitro* transposition reaction in a positive way. Under conditions where transpososome formation was rendered suboptimal, the presence of H-NS increased the yield of transpososome formed and this transpososome was shown to be catalytically competent. H-NS was also shown to bind a MC, raising the possibility that this host factor can act early in the transpososome assembly process. Finally, we have shown genetically that H-NS has the potential to act as a positive regulator of Tn5 transposition *in vivo* as the frequency of Tn5 transposition was reduced in an *hns* disruption strain.

3.4.1 H-NS interactions with the transpososome

We used EMSA to show that H-NS binds preferentially to the Tn5 transpososome. At H-NS concentrations that did not significantly affect the mobility of ME DNA, we saw a full shift in the transpososome to a form with a lower electrophoretic mobility. We examined the determinants for H-NS binding to the transpososome by increasing the ME DNA size and adding flanking donor DNA. We observed that H-NS did not require the presence of flanking donor DNA in the transpososome in order to bind, but did require the presence of at least 20 bp of ME DNA (i.e. the minimum amount of transposon required for transpososome formation). Given that transposase makes contacts over most of the ME sequence (19, 34) this observation suggested that H-NS must be in close association with transposase in the context of the transpososome, an inference that was confirmed by DNA footprinting studies. We observed an expansion of the ‘transposase footprint’ when H-NS was incorporated into the transpososome, as summarized in Figure

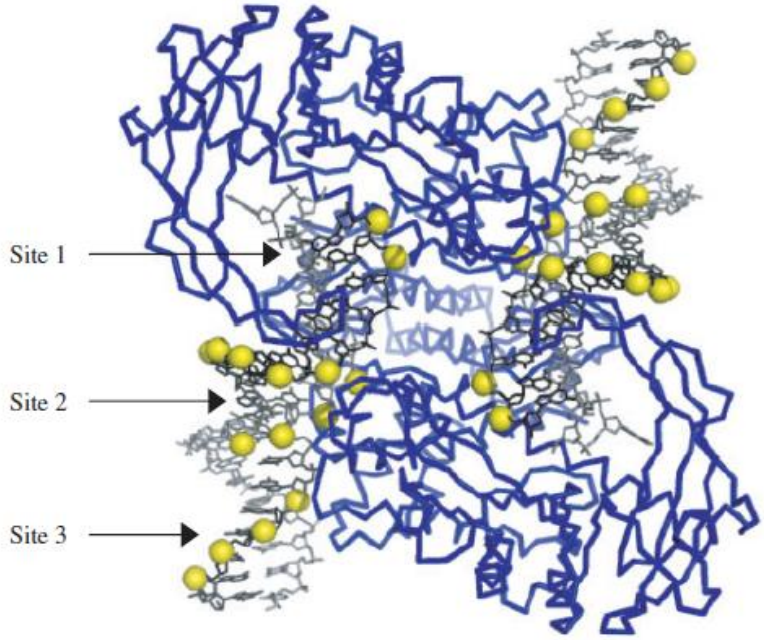


Figure 3.9. Three-dimensional representation of the Tn5 transpososome with potential H-NS binding sites indicated.

A model of the Tn5 transpososome formed with the HA transposase and 20 bp OEs is shown along with potential H-NS binding sites 1 to 3. Note that all three sites include portions of the DNA that are not encompassed by transposase on at least one surface. Transposase is shown as blue. The OE DNA strands are depicted in black and grey for the transferred strand and non-transferred strands, respectively. Yellow spheres indicate the position of the phosphate backbone where H-NS binding may occur. Adapted from PDB: ID 1MUH (32).

3.3B. We infer from this data that there may be as many as three distinct H-NS binding sites within the transpososome (listed as sites 1–3 in Figure 3.3B). Site 1 includes residues (-4) to 2 of the TS (i.e. the proximal region of zone A), site 2 includes residues 5–9 of the TS and 8–11 of the NTS (i.e. the distal region of zone A), and site 3 includes residues 15 to 18 of the TS (i.e. the distal region of zone B). Interestingly, there is evidence that all three of these potential H-NS binding sites include DNA with a distorted structure (Figure 3.9). Site 1 includes a DNA kink between residues 1 and 2 and work from solution studies revealed hydroxyl radical hypersensitivity extending into the flanking donor DNA (19, 34, 53). Site 2 is immediately adjacent to a 418 bend in the helical axis of the DNA between residues 11–12 (34). Site 3 is coincident with a second region of hydroxyl radical hypersensitivity that extends from residues 16–20 (19). This correspondence of potential H-NS binding sites inferred from footprinting studies with regions of deformed DNA structure fits well with the known preference of H-NS binding to deformed DNA structures.

Based on the structure of the Tn5 transpososome we can infer that the major groove of the DNA is available for H-NS binding in sites 1 and 3 and the minor groove is available in site 2. Interestingly, the hydroxyl radical protection pattern (Figures 3.3 and 3.9) is consistent with H-NS binding in both the major and minor groove in different segments of transpososome. Notably, it remains controversial as to whether H-NS binds to the major or the minor groove (54, 55), although we have previously detected H-NS binding to the Tn10 transpososome using a minor groove-specific DNA footprinting reagent (9, 10). A possible complicating factor in the current analysis is that by binding immediately adjacent to transposase, or possibly interacting with transposase, as has been

shown in the Tn10 system (9), H-NS may change the contacts transposase makes with the Tn5 end DNA. Thus, at this point we do not know if specific hydroxyl radical protections shown in Figure 3.9 are a consequence of only H-NS contacts. Additional studies will be required to sort out the mode of H-NS binding in the Tn5 transpososome.

H-NS also bound with high specificity to a Tn5 MC formed with ME DNA and the Δ369 EK/MA transposase. The fact that H-NS selectively bound the MC suggests that H-NS is not restricted to binding the fully assembled transpososome in the Tn5 system. We infer that transposase binding to the ME is likely sufficient to generate determinants for H-NS binding. However, at this point it is unclear if the potential H-NS binding determinants described above for the transpososome are also present in the MC. Nevertheless, our finding that P116S H-NS failed to bind the MC and the transpososome with high specificity reinforces our conclusion that deformed DNA structures constitute important H-NS binding determinants in both complexes.

We found that one of the potential H-NS binding sites (Site 1) was critical for H-NS binding to the transpososome. When positions 3 and 6 of the ME DNA were mutated, we noted a small decrease in transpososome levels formed, however these mutations significantly reduced H-NS binding to the transpososome. While we have not yet determined if both mutations are necessary to prevent H-NS binding, it is intriguing that at least one mutation (base pair 3) is at a position expected to be part of a distorted DNA structure. It is also interesting that mutations within only one of the three potential H-NS binding sites would significantly abrogate H-NS binding to the transpososome. One possible explanation for this is that H-NS binding to the proximal portion of zone A affects the conformation of the transpososome in a manner that permits additional H-NS

binding events. This would be analogous to the situation in the Tn10 system where it appears that H-NS binding to a distorted DNA structure in the flanking donor DNA portion of the transpososome is critical for additional H-NS binding events within the terminal inverted repeat (10).

3.4.2 Functional consequences of H-NS binding to the Tn5 transpososome

Our binding assays showed that addition of H-NS had a small positive effect on the *in vitro* Tn5 transpososome assembly when the reaction was carried out under conditions that have been optimized for transpososome formation (i.e. with the HA transposase and the ME). In principle, this optimization might render the transpososome assembly reaction insensitive to host factors. With this in mind we set out to identify a means of making transpososome formation less efficient. We found that the addition of the polyanion heparin greatly reduced transpososome formation. Heparin is expected to compete with the ME DNA for transposase binding, thereby making it more difficult for the transpososome to form. At a heparin concentration sufficient to reduce transpososome formation by roughly 22-fold, we found that addition of H-NS resulted in an 11-fold increase in transpososome formation. Thus, under these reaction conditions, transpososome formation became strongly dependent on the presence of H-NS. Control reactions were included that demonstrated that H-NS binding to the transpososome was a prerequisite for stimulation of transpososome formation in the presence of heparin.

We also reduced the efficiency of transpososome formation by using a form of transposase that did not include one of the mutations (EK54) that leads to hyperactivity. In this situation there was a 2.5-fold increase in transpososome formation when H-NS

was added to the reaction. Substitution of lysine for glutamate at position 54 creates a favourable base-specific *cis* contact with thymine 10 and likely removes an unfavourable contact between E54 and the phosphate backbone at positions 10 to 12 (30). It is intriguing that putative H-NS binding sites 2 and 3 are located within a few residues on either side of positions 10-12. H-NS binding to sites 2 and 3 could alter the DNA structure in a manner that prevents the unfavourable contact between E54 and positions 10-12 from forming. Alternatively, H-NS binding to these sites may cause a conformational change in transposase that partially shields the phosphate backbone from E54. In either case, H-NS would make the transposase-ME interaction less sensitive to the nature of the amino acid at position 54.

It is intriguing that H-NS can both stimulate transpososome formation and bind the MC. These observations lead us to propose that the association between H-NS and the MC might assist in transpososome assembly. It is well established that H-NS is capable of forming bridges through self-oligomerization between separate or distantly spaced DNA molecules (12, 13). H-NS-H-NS interactions could therefore help drive the association between two MCs. While it is clear that this type of interaction has a minimal effect on the Tn5 system with the HA transposase, it may be important in assembly reactions carried out with wild-type transposase (see 'Discussion'). Notably, there is evidence from studies on longer DNA molecules (relative to that used in the current study) that transpososome assembly in the Tn5 system may be more complicated than outlined above and in Figure 3.1. It has been shown that transposase is capable of synapsing nontransposon end segments and that a transposase dimer formed during this interaction can be transferred to a transposon end sequence. Subsequently, the

transposase dimer would capture the second transposon end, thereby forming a stable transpososome (28, 29). As described earlier, H-NS could facilitate the initial pairing of separate, non-transposon end DNA segments. Alternatively, or in addition, H-NS might promote the capture of the second transposon end by helping to stabilize a bend in the intervening DNA between the two transposon ends.

3.4.3 H-NS acts as a positive regulator of Tn5 transposition *in vivo*

We used two different transposition assays to determine if the absence of H-NS has an impact on the Tn5 transposition frequency *in vivo*. We found that in the absence of H-NS Tn5 transposition was reduced approximately 5-fold in the papillation assay and up to 11-fold in the mating out assay. We also provide evidence that the effect is not caused by downstream 'polar' effects of the insertion used to disrupt *hns* as complementation of the Δ*hns* strain with an external source of *hns* restored transposition levels. These results are consistent with H-NS being a positive regulator of Tn5 transposition *in vivo*.

Furthermore, we have provided evidence that is consistent with this positive role working at the level of transpososome assembly. We found that addition of H-NS to an *in vitro* assembly reaction with MA/LP transposase partially compensated for the absence of the EK54 mutation, a mutation that is known to increase transposase binding to Tn5 end sequences. In addition, we found that the fold-decrease in transposition frequency *in vivo* in an Δ*hns* strain encoding a transposase lacking the EK54 mutation was larger than in the same strain encoding a transposase with the EK54 mutation. If H-NS acts indirectly to promote Tn5 transposition, it is not obvious why the absence of the EK54 mutation would produce this larger effect.

3.4.4 A broader role for H-NS in lateral gene transfer

The work presented here provides the second documented case of H-NS promoting a transposition reaction through a direct interaction with a transposition complex. In the Tn10 system the available evidence is consistent with H-NS promoting transposition by binding to the initial transpososome and ultimately promoting a conformational change in this complex that inhibits self-destructive intramolecular transposition events and promotes intermolecular transposition events. While H-NS appears to stimulate Tn5 transposition by a different mechanism, the common thread in the two systems is that H-NS binds selectively to transpososomes and this raises the possibility that the activity of many other transposition systems may be modulated by H-NS. If this turns out to be the case, then H-NS would play an even broader role in lateral gene transfer than has previously been predicted through the xenogeneic silencing model because transposons are major drivers of lateral gene transfer.

3.5 References

1. Lucchini,S., Rowley,G., Goldberg,M.D., Hurd,D., Harrison,M. and Hinton,J.C.D. (2006) H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS pathogens*, **2**, e81.
2. Navarre,W.W., Porwollik,S., Wang,Y., McClelland,M., Rosen,H., Libby,S.J. and Fang,F.C. (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in Salmonella. *Science*, **313**, 236-8.
3. Navarre,W.W., McClelland,M., Libby,S.J. and Fang,F.C. (2007) Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes & development*, **21**, 1456-71.
4. Doyle,M., Fookes,M., Ivens,A., Mangan,M.W., Wain,J. and Dorman,C.J. (2007) An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. *Science*, **315**, 251-2.
5. Shiga,Y., Sekine,Y., Kano,Y. and Ohtsubo,E. (2001) Involvement of H-NS in transpositional recombination mediated by IS1. *Journal of bacteriology*, **183**, 2476-84.
6. Rouquette,C., Serre,M.-C. and Lane,D. (2004) Protective role for H-NS protein in IS1 transposition. *Journal of bacteriology*, **186**, 2091-8.
7. Swingle,B., O'Carroll,M., Haniford,D. and Derbyshire,K.M. (2004) The effect of host-encoded nucleoid proteins on transposition: H-NS influences targeting of both IS903 and Tn10. *Molecular microbiology*, **52**, 1055-67.
8. Twiss,E., Coros,A.M., Tavakoli,N.P. and Derbyshire,K.M. (2005) Transposition is modulated by a diverse set of host factors in Escherichia coli and is stimulated by nutritional stress. *Molecular microbiology*, **57**, 1593-607.
9. Wardle,S.J., O'Carroll,M., Derbyshire,K.M. and Haniford,D.B. (2005) The global regulator H-NS acts directly on the transpososome to promote Tn10 transposition. *Genes & development*, **19**, 2224-35.
10. Ward,C.M., Wardle,S.J., Singh,R.K. and Haniford,D.B. (2007) The global regulator H-NS binds to two distinct classes of sites within the Tn10 transpososome to promote transposition. *Molecular microbiology*, **64**, 1000-13.
11. Singh,R.K., Liburd,J., Wardle,S.J. and Haniford,D.B. (2008) The nucleoid binding protein H-NS acts as an anti-channeling factor to favor intermolecular Tn10 transposition and dissemination. *Journal of molecular biology*, **376**, 950-62, 10.1016/j.jmb.2007.12.035.
12. Fang,F.C. and Rimsky,S. (2008) New insights into transcriptional regulation by H-NS. *Current opinion in microbiology*, **11**, 113-20.
13. Rimsky,S., Zuber,F., Buckle,M. and Buc,H. (2001) A molecular mechanism for the repression of transcription by the H-NS protein. *Molecular microbiology*, **42**, 1311-23.

14. Dame,R.T., Wyman,C. and Goosen,N. (2001) Structural basis for preferential binding of H-NS to curved DNA. *Biochimie*, **83**, 231-4.
15. Lavoie,B.D., Chan,B.S., Allison,R.G. and Chaconas,G. (1991) Structural aspects of a higher order nucleoprotein complex: induction of an altered DNA structure at the Mu-host junction of the Mu type 1 transpososome. *The EMBO journal*, **10**, 3051-9.
16. Savilahti,H., Rice,P.A. and Mizuuchi,K. (1995) The phage Mu transpososome core: DNA requirements for assembly and function. *The EMBO journal*, **14**, 4893-903.
17. Wang,Z., Namgoong,S.Y., Zhang,X. and Harshey,R.M. (1996) Kinetic and structural probing of the precleavage synaptic complex (type 0) formed during phage Mu transposition. Action of metal ions and reagents specific to single-stranded DNA. *The Journal of biological chemistry*, **271**, 9619-26.
18. York,D. and Reznikoff,W.S. (1997) DNA binding and phasing analyses of Tn5 transposase and a monomeric variant. *Nucleic acids research*, **25**, 2153-60.
19. Bhasin,A., Goryshin,I.Y., Steiniger-White,M., York,D. and Reznikoff,W.S. (2000) Characterization of a Tn5 pre-cleavage synaptic complex. *Journal of molecular biology*, **302**, 49-63.
20. Krebs,M.P. and Reznikoff,W.S. (1986) Transcriptional and translational initiation sites of IS50. Control of transposase and inhibitor expression. *Journal of molecular biology*, **192**, 781-91.
21. Weinreich,M.D., Mahnke-Braam,L. and Reznikoff,W.S. (1994) A functional analysis of the Tn5 transposase. Identification of domains required for DNA binding and multimerization. *Journal of molecular biology*, **241**, 166-77.
22. Mahnke Braam,L.A., Goryshin,I.Y. and Reznikoff,W.S. (1999) A mechanism for Tn5 inhibition. carboxyl-terminal dimerization. *The Journal of biological chemistry*, **274**, 86-92.
23. Reznikoff,W.S. (1982) Tn5 transposition and its regulation. *Cell*, **31**, 307-8.
24. Weinreich,M.D., Gasch,A. and Reznikoff,W.S. (1994) Evidence that the cis preference of the Tn5 transposase is caused by nonproductive multimerization. *Genes & development*, **8**, 2363-74.
25. York,D. and Reznikoff,W.S. (1996) Purification and biochemical analyses of a monomeric form of Tn5 transposase. *Nucleic acids research*, **24**, 3790-6.
26. Jilk,R. a, Makris,J.C., Borchardt,L. and Reznikoff,W.S. (1993) Implications of Tn5-associated adjacent deletions. *Journal of bacteriology*, **175**, 1264-71.
27. Steiniger-White,M. and Reznikoff,W.S. (2000) The C-terminal alpha helix of Tn5 transposase is required for synaptic complex formation. *The Journal of biological chemistry*, **275**, 23127-33.
28. Steiniger,M., Adams,C.D., Marko,J.F. and Reznikoff,W.S. (2006) Defining characteristics of Tn5 Transposase non-specific DNA binding. *Nucleic acids research*, **34**, 2820-32.

29. Adams,C.D., Schnurr,B., Skoko,D., Marko,J.F. and Reznikoff,W.S. (2006) Tn5 transposase loops DNA in the absence of Tn5 transposon end sequences. *Molecular microbiology*, **62**, 1558-68.
30. Zhou,M., Bhasin,A. and Reznikoff,W.S. (1998) Molecular genetic analysis of transposase-end DNA sequence recognition: cooperativity of three adjacent base-pairs in specific interaction with a mutant Tn5 transposase. *Journal of molecular biology*, **276**, 913-25.
31. Goryshin,I.Y. and Reznikoff,W.S. (1998) Tn5 in vitro transposition. *The Journal of biological chemistry*, **273**, 7367-74.
32. Wiegand,T.W. and Reznikoff,W.S. (1992) Characterization of two hypertransposing Tn5 mutants. *Journal of bacteriology*, **174**, 1229-39.
33. Zhou,M. and Reznikoff,W.S. (1997) Tn5 transposase mutants that alter DNA binding specificity. *Journal of molecular biology*, **271**, 362-73.
34. Davies,D.R., Goryshin,I.Y., Reznikoff,W.S. and Rayment,I. (2000) Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science*, **289**, 77-85.
35. Lovell,S., Goryshin,I.Y., Reznikoff,W.R. and Rayment,I. (2002) Two-metal active site binding of a Tn5 transposase synaptic complex. *Nature structural biology*, **9**, 278-81.
36. Hanahan,D. (1983) Studies on transformation of Escherichia coli with plasmids. *Journal of molecular biology*, **166**, 557-80.
37. Foster,T.J., Davis,M.A., Roberts,D.E., Takeshita,K. and Kleckner,N. (1981) Genetic organization of transposon Tn10. *Cell*, **23**, 201-13.
38. Rosner,J.L. (1972) Formation, induction, and curing of bacteriophage P1 lysogens. *Virology*, **48**, 679-89.
39. Zhang,A., Rimsky,S., Reaban,M.E., Buc,H. and Belfort,M. (1996) Escherichia coli protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *The EMBO journal*, **15**, 1340-9.
40. Bhasin,A., Goryshin,I.Y. and Reznikoff,W.S. (1999) Hairpin formation in Tn5 transposition. *The Journal of biological chemistry*, **274**, 37021-9.
41. Naumann,T.A. and Reznikoff,W.S. (2000) Trans catalysis in Tn5 transposition. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 8944-9.
42. Way,J.C., Davis,M.A., Morisato,D., Roberts,D.E. and Kleckner,N. (1984) New Tn10 derivatives for transposon mutagenesis and for construction of lacZ operon fusions by transposition. *Gene*, **32**, 369-79.
43. Guzman,L.M., Belin,D., Carson,M.J. and Beckwith,J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of bacteriology*, **177**, 4121-30.

44. Kleckner,N., Bender,J. and Gottesman,S. (1991) Uses of transposons with emphasis on Tn10. *Methods in enzymology*, **204**, 139-80.
45. Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2006) Formation, characterization and partial purification of a Tn5 strand transfer complex. *Journal of molecular biology*, **364**, 290-301.
46. Cruz,N.B. de la, Weinreich,M.D., Wiegand,T.W., Krebs,M.P. and Reznikoff,W.S. (1993) Characterization of the Tn5 transposase and inhibitor proteins: a model for the inhibition of transposition. *Journal of bacteriology*, **175**, 6932-8.
47. Goryshin IYu, Kil,Y.V. and Reznikoff,W.S. (1994) DNA length, bending, and twisting constraints on IS50 transposition. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 10834-8.
48. Spurio,R., Falconi,M., Brandi,A., Pon,C.L. and Gualerzi,C.O. (1997) The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *The EMBO journal*, **16**, 1795-805.
49. Stella,S., Spurio,R., Falconi,M., Pon,C.L. and Gualerzi,C.O. (2005) Nature and mechanism of the in vivo oligomerization of nucleoid protein H-NS. *The EMBO journal*, **24**, 2896-905.
50. Sonnenfield,J.M., Burns,C.M., Higgins,C.F. and Hinton,J.C. (2001) The nucleoid-associated protein StpA binds curved DNA, has a greater DNA-binding affinity than H-NS and is present in significant levels in hns mutants. *Biochimie*, **83**, 243-9.
51. Adams,C.D., Schnurr,B., Marko,J.F. and Reznikoff,W.S. (2007) Pulling apart catalytically active Tn5 synaptic complexes using magnetic tweezers. *Journal of molecular biology*, **367**, 319-27.
52. Krebs,M.P. and Reznikoff,W.S. (1988) Use of a Tn5 derivative that creates lacZ translational fusions to obtain a transposition mutant. *Gene*, **63**, 277-85.
53. Bischerour,J. and Chalmers,R. (2007) Base-flipping dynamics in a DNA hairpin processing reaction. *Nucleic acids research*, **35**, 2584-95.
54. Tippner,D., Afflerbach,H., Bradaczek,C. and Wagner,R. (1994) Evidence for a regulatory function of the histone-like Escherichia coli protein H-NS in ribosomal RNA synthesis. *Molecular microbiology*, **11**, 589-604.
55. Tippner,D. and Wagner,R. (1995) Fluorescence analysis of the Escherichia coli transcription regulator H-NS reveals two distinguishable complexes dependent on binding to specific or nonspecific DNA sites. *The Journal of biological chemistry*, **270**, 22243-7.

CHAPTER 4[1]

4 REGULATION OF TN5 BY H-NS INVOLVES A DIRECT INTERACTION BETWEEN H-NS AND THE TRANSPOSOSOME

4.1 Introduction

Most bacteria harbor a variety of different types of transposons (1). While transposons can compromise genome stability through the various types of DNA rearrangements they promote, they can also confer a selective advantage to their hosts. This can come about through transposons acquiring genes that encode resistance to antibiotics and other environmental toxins or through transposon insertion events that alter the expression of key host genes. In order for transposons and their hosts to coexist transposition levels must be tightly regulated (2). There are several examples where host proteins have been co-opted to down-regulate transposition. For instance, Dam methylase of *E. coli* methylates GATC sequences found in both the promoters controlling the expression of some transposase genes and in the transposon ends of several transposons. The former inhibits transposase expression and the latter inhibits transposase binding to transposon 'end' sequences (3, 4). Other examples of host proteins that limit transposition include proteins that are global regulators of gene expression in bacteria including IHF (5), RNaseE (6) and Hfq (7). It is also apparent in some cases that there has been strong selective pressure for transposons to contain regulatory sequences for transposase genes that are suboptimal for transposase expression. For example, both Tn10 and Tn5 have weak promoter sequences and suboptimal translation initiation regions for their

[1]A version of this chapter has been submitted for publication: Whitfield,C.R., Shilton,B.S., and Haniford,D.B. Regulation of Tn5 transposition by H-NS involves a direct interaction between H-NS and the transpososome. *Mobile DNA.*

transposase genes. On the other hand there are some examples where transposons appear to have co-opted host proteins to promote their transposition. Examples of such proteins include IHF, HU, H-NS, Fis, Topoisomerase I, DNA Gyrase and DnaA (2, 8).

It is often unclear as to whether host proteins directly or indirectly regulate transposition reactions. The development of *in vitro* transposition reactions for systems such as Mu, Tn10 and Tn5 has allowed host factors implicated genetically as regulators of transposition reactions to be tested for their potential to directly interact with the transposition machinery. IHF, H-NS and HU are all DNA-binding proteins that have been shown to directly interact with transposition complexes *in vitro* (2, 8, 9). The distinction between a direct versus an indirect regulatory pathway could be important with regard to how efficiently and quickly a transposon can respond to changing physiological conditions in the cell.

In the current work we have focused on understanding the role of H-NS in Tn5 transposition. H-NS is a highly expressed DNA-binding protein that is present in many proteobacteria (10). H-NS binding to high affinity sites embedded within A-T rich sequences is thought to nucleate polymerization of H-NS on DNA (11). In solution, H-NS exists as a dimer at physiological concentrations but upon binding DNA H-NS forms higher order oligomers through head-to-head and tail-to-tail interactions between adjacent dimers (12). Oligomerization of H-NS on promoter regions of genes results in gene silencing probably through exclusion of RNA polymerase (13). H-NS influences the expression of a large number of genes in *E. coli* and is therefore considered a global regulator of gene expression (14, 15). H-NS also plays an important role in lateral gene transfer in some proteobacteria as it has a propensity to silence newly acquired genes,

which tend to be A-T rich, permitting bacteria to gradually integrate the new DNA into existing regulatory circuits (13).

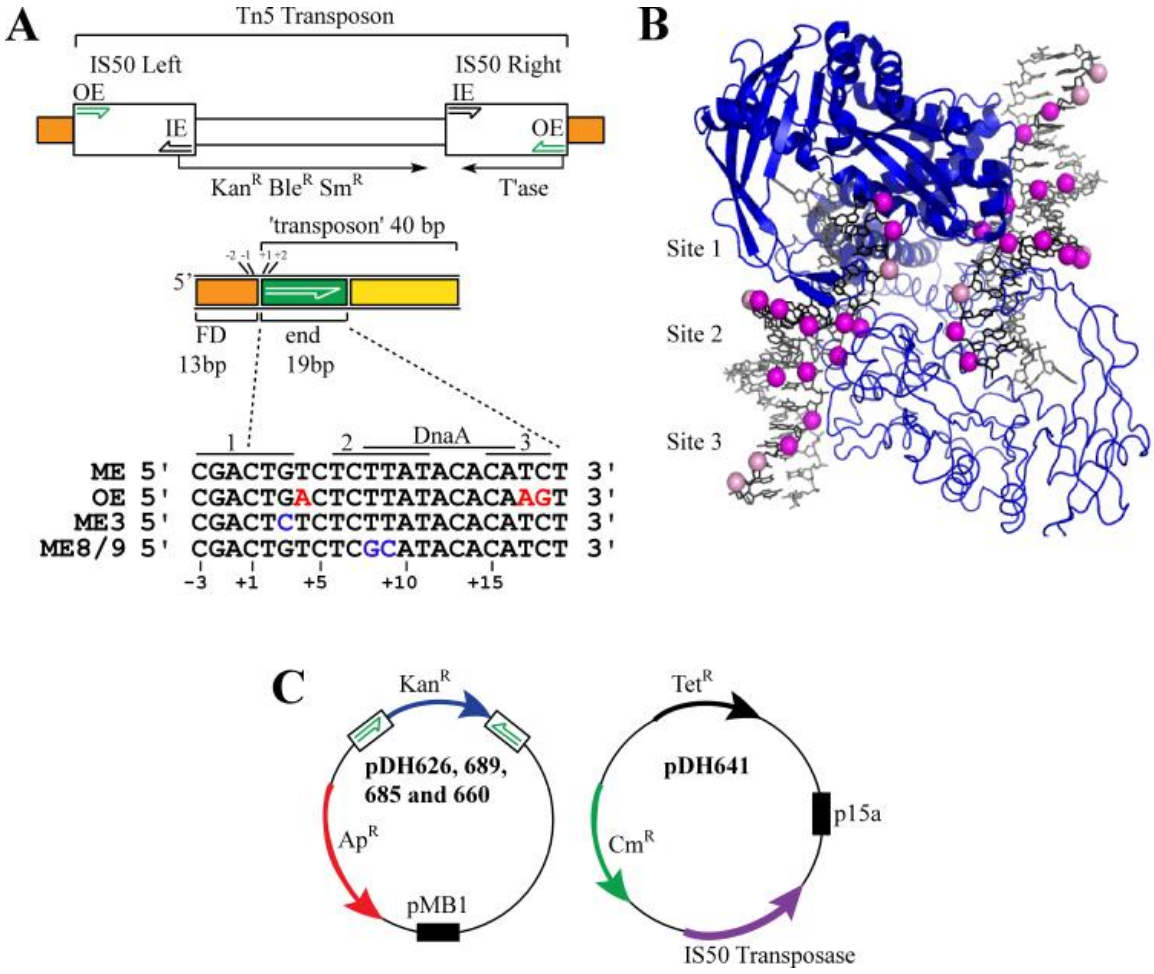
Tn5 is a composite transposon made up of three antibiotic resistance genes encompassed by insertion sequences IS50Right and IS50Left (Figure 4.1A). Tn5/IS50 is widely distributed in proteobacteria. Tn5 and IS50 transposition is tightly regulated with events occurring at a frequency of roughly one event per element per generation in 10^7 and 10^3 cells, respectively. A high-resolution structure of the Tn5 transpososome has provided a wealth of information regarding protein-protein and protein-DNA interactions within the transpososome (16), but details still remain to be elucidated with regard to how the transpososome is initially assembled.

In previous work we have shown that inclusion of H-NS in Tn5 transpososome assembly reactions resulted in incorporation of H-NS into the transpososome (9). Moreover, when such assembly reactions were performed under conditions where transpososome assembly was suboptimal (i.e. in the presence of a DNA competitor) it was found that inclusion of H-NS greatly facilitated transpososome formation. H-NS did not directly impact the efficiency of transposon excision when it was added to reactions containing pre-assembled transpososomes under normal conditions (9). The positive effect of H-NS on transpososome formation could result from H-NS promoting: (1) the formation of a pre-transpososome complex (for example, a complex where a monomer of transposase binds a single transposon end); (2) the assembly of pre-transpososome complexes into transpososome; and/or (3) the stabilization of the transpososome.

Tn5 transposition is reduced ~3-5-fold *in vivo* in *E. coli* containing a disruption ofthe *hns* gene (Δ*hns*) (9). It remains to be seen if this reduction in transposition is the

Figure 4.1. Tn5 genetic structure, transpososome structure and transposon substrates.

(A) The organization of Tn5 (5.8 kb) is shown. The outside (OE) and inside ends (IE) of Tn5 (half arrows) contain determinants for transposase binding. IS50Right encodes transposase and other ORFs are shown (thin arrows show the corresponding transcription unit). Flanking donor DNA is represented as orange rectangles. Below the schematic of Tn5 an illustration of the basic transposon end fragment (53 bp) used for *in vitro* binding studies is shown. In addition, the sequence of the terminal 19 nucleotides, as well as 3 flanking nucleotides, of the non-transferred strand of each of the four transposon end substrates used in this work is shown (ME – mosaic end). Base pair changes between the WT ME and mutant MEs (ME 3 and ME 8/9) are in blue. Base pair differences between OE and ME sequences are in red. The DnaA binding site that overlaps the Tn5 transposon sequence is indicated above the sequences with a black vertical line. Potential H-NS binding sites inferred from hydroxyl radical footprinting are also represented in this way and identified as either sites 1, 2 or 3. (B) Sites 1, 2 and 3 represent potential H-NS binding sites of the H-NS-Tn5 transpososome as described in (A); the light and dark pink spheres placed in the structural model of the Tn5 transpososome represent positions of weak and strong protection against hydroxyl radical attack, respectively. (C) Schematic of plasmids used for mating out assays. The plasmid on the left contains the mini-Tn5-Kan^R element with arrows in boxes depicting the end sequence (WT ME – pDH626, OE – pDH689, ME 3 – pDH685, or ME 8/9 – pDH660) and the thick blue arrow indicates the kanamycin resistance gene. The plasmid on the right encodes transposase (MA56) (purple arrow) under control of its native promoter. Black boxes represent the origins of replication, pMB1 and p15A; Ap^R, Cm^R and Tet^R encode resistance genes for ampicillin, chloramphenicol and tetracycline, respectively.


A
Tn5 Transposon
IS50 Left
OE
IE
IS50 Right
IE
OE
Kan^R Ble^R Sm^R
T'ase
'transposon' 40 bp
-2 -1 +1 +2
5'
FD
13bp
end
19bp
1
2
DnaA
3
ME 5' CGACTGTCTCTTATACACATCT 3'
OE 5' CGACTGACTCTTATACACAAGT 3'
ME3 5' CGACTCTCTCTTATACACATCT 3'
ME8/9 5' CGACTGTCTCGCATACACATCT 3'
-3 +1 +5 +10 +15
B
Site 1
Site 2
Site 3
C
Kan^R
pDH626, 689, 685 and 660
Ap^R
pMB1
Tet^R
pDH641
p15a
Cm^R
IS50 Transposase

result of a transpososome assembly defect or an indirect effect of H-NS on gene expression.

Our finding that H-NS bound to the Tn5 transpososome *in vitro* led us to perform DNA footprinting experiments in an attempt to localize the site(s) of H-NS binding. The reactivity of the transpososome DNA to hydroxyl radical cleavage was altered at three sites, which we have designated sites 1, 2 and 3 (9) (Figure 4.1). Based on the transpososome crystal structure, sites 2 and 3 are the most obviously accessible sites for H-NS binding. In the current study we have further investigated the interaction between H-NS and the Tn5 transpososome by measuring the affinity of H-NS for the Tn5 transpososome and using site-directed mutagenesis and protein-protein cross-linking studies to define determinants for H-NS binding to the transpososome. We have also used the information gained from site directed mutagenesis to test the idea that H-NS regulates Tn5 transposition *in vivo* by acting directly on the transpososome.

4.2.0 Materials and methods

Chemicals and oligonucleotides were from Sigma-Aldrich. Growth media was from Becton Dickinson. Enzymes were from New England Biolabs. Radio-nucleotides were from Amersham Biosciences.

4.2.1 Plasmids and transposon DNAs

Plasmid-based mini-Tn5 elements were constructed using a three-fragment cloning strategy. Primers containing a Pst1 site (CW1 and CW2) were designed to the 5’ and 3’ ends of the kanamycin resistance gene from pNK1182 (17) and a PCR reaction was

performed to amplify the Kan^R fragment. After digestion with PstI the Kan^R fragment was ligated to the linear fragment of KpnI-digested pTZ18U (18) and to a 20 bp ME fragment (CW3/4 for WT ME, CW5/6 for OE, CW7/8 for ME 3 and CW9/10 for ME 8/9) containing PstI and KpnI overhangs. The resulting plasmids, pDH626, pDH689, pDH685 and pDH660, contain identical mini-Tn5 transposons except that the transposon ends were either wild-type ME, OE, ME 3 and ME 8/9 sequences, respectively. To provide a source of Tn5 transposase on a compatible plasmid to the above transposon substrates, Tn5 transposase DNA from pRZ9905 (19) was cloned into pACYC184 as follows: restriction sites for either HindIII or XbaI were incorporated into primers (CW11 and 12) complementary to the 5' and 3' ends of the Tn5 transposase gene and following amplification and digestion of the transposase fragment with HindIII and XbaI the transposase fragment was cloned into HindIII/XbaI-digested pACYC184. The transposase gene used contains a mutation that eliminates synthesis of the inhibitor protein, but otherwise is wild-type in sequence.

Transposon end substrates used for binding assays contain 13 bp of donor DNA and 40 bp of transposon DNA. These substrates were generated by annealing complementary, gel-purified oligonucleotides (Table 4.1) and were subsequently 5' end-labeled with T4 polynucleotide kinase (New England Biolabs) and γ^{32}P-ATP (Amersham Biosciences) using standard procedures.

4.2.2 Protein purification

Tn5 transposase and H-NS were purified to 98% purity as shown in Figure 4.7A (lanes 3 and 4 respectively) as described previously (20, 21). Tn5 transposase concentration was

determined using the Bradford assay (Pierce) and H-NS concentration was determined using the BCA assay (Pierce).

4.2.3 H-NS binding assays

Transpososome assembly reactions were performed by mixing ^{32}P-labeled transposon end fragments (2 nM) and purified transposase (200 nM) as described previously (9), except that transposase was added at 400 nM in the case of reactions with the ME 8/9 substrate in order to obtain roughly equivalent amounts of transpososome in all *in vitro* reactions. H-NS was added to the reactions at the same time as transposase in varying concentrations (normally 23 nM to 849 nM) and after incubation for 30 minutes at 37˚C reactions were mixed with load dye and applied to a 5% native polyacrylamide gel. Gels were run and analyzed as previously described (9). The typical transpososome yield was about 10% of input DNA giving a final concentration of 0.2 nM transpososome per assembly reaction. ImageQuant v5.1 software was used to analyze H-NS-bound fractions of transpososome based on the proportion of labeled DNA present in the mobility shifts compared to the overall total labeled DNA in each lane. The equilibrium dissociation constant (K_d) was generated using the equation:

$$\theta^{-1} = 1 + (\frac{K_d}{[P_t]})$$

where θ is the fraction of H-NS-bound transpososome and P_t is the total H-NS concentration. At least three independent binding experiments were performed for each transposon end substrate. Data for WT ME and OE substrates was fit to a quadratic equation using non-linear regression and binding curves were generated using GraphPad Prism v5.0.

4.2.4 Mating out assay

Mating out assays were performed in isogenic *hns*+ (DBH33) and Δ*hns* (DBH1) donor strains (Table 4.1) as previously described (9). These strains were transformed with a 'transposon substrate' plasmid containing a Kan^R gene between the transposon ends (either pDH626, pDH689, pDH685 and pDH660) and a compatible 'transposase' plasmid (pDH641). Donor cells were mixed with recipient cells (HB101) and after growth in LB mating mixes were pelleted and resuspended in 0.85% saline whereupon cells were plated on M9-glucose plates supplemented with leucine, thiamine and streptomycin sulphate (150 μg ml^{-1}) for measuring total exconjugants and the above plus kanamycin (50 μg ml^{-1}) for measuring transposon hops. Transposition frequencies were calculated by dividing the number of Kan^RSm^R colonies by the number of Sm^R colonies.

4.2.5 EDC/NHS chemical cross-linking

100x volume (1 ml) transpososome assembly reactions were prepared with unlabeled ME DNA (500 nM), purified transposase (565 nM), and WT H-NS (1.2 μM) or P116S H-NS (7.4 μM) as previously described (9). Reactions were concentrated by microfiltration from the initial volume to 0.045 ml (Millipore Vivaspin 30000 kDa cut-off). Samples were then treated with 9.5 μl of the chemical cross-linker 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide or EDC (50 mM) plus N-hydroxysulfosuccinimide or NHS (12.5 mM) (both prepared in water) for 2.5 minutes at room temperature. Non-denaturing load dye was added to the cross-linking reactions and samples were applied to a 5% native polyacrylamide gel. After staining the gel with ethidium bromide the

transpososome and H-NS-bound transpososome were isolated based on mobility differences. Proteins were eluted out of the gel slices at 42°C with 1 ml of elution buffer (0.5% SDS and 1M Sodium Acetate), concentrated as described above and subjected to immunoblot analysis as previously described (22). Purified Tn5 transposase and H-NS were also subjected to EDC/NHS cross-linking as above except that 0.5 µg (1.2 µM of transposase and 3.6 µM of H-NS) of each protein was used either separately or together with or without 2.5 µl of the chemical cross-linker EDC (50 mM) plus NHS (12.5 mM) and incubated at room temperature for 3 minutes. SDS load dye was added and samples were immediately loaded on the SDS protein gel for immunoblot analysis. For complexes that were treated with micrococcal nuclease, the cross-linking protocol was similar except for the following changes. After cross-linker treatment the reactions were stopped with addition of Tris-HCl (pH 8) to a final concentration of 500 nM. Micrococcal nuclease (100 units) and 5 mM $CaCl_2$ were added to the reactions and incubated at 37°C for 15 minutes. At the same time, additional samples were mock-treated with only 5 mM $CaCl_2$ and no nuclease. A small amount of each sample (4 µl) was removed and analyzed on a 1% agarose gel to ensure that no DNA remained in the samples treated with the nuclease and that complexes were not disrupted in the mock-treated samples. The remainder of the samples were concentrated as described above to 20 µl and loaded on an SDS protein gel.

4.2.6 MALDI-TOF MS analysis of cross-linked products

400x volume (4 ml) H-NS-transpososome assembly reaction was concentrated to 225 µl, divided into 4 equivalent aliquots and each aliquot was treated similarly to above. The cross-linking reactions were quenched by addition of Tris-HCl (pH 8) to a final

concentration of 500 nM and then applied to a 5% native polyacrylamide gel. H-NS-transpososome was eluted from the native gel after staining with ethidium bromide and fractions were pooled, concentrated and applied to a single lane of an SDS protein gel, which was stained with Coomassie Blue. Transposase monomer, H-NS monomer and cross-linked product 'a' were gel-isolated from the same lane using an Ettan Spot-picker (GE Healthcare). In-gel digestion was performed using a MassPREP automated digester station (PerkinElmer). Gel pieces were Coomassie destained using 50mM ammonium bicarbonate and 50% acetonitrile, which was followed by protein reduction using 10 mM dithiothreitol (DTT), alkylation using 55 mM iodoacetamide (IAA), and tryptic digestion. Peptides were extracted using a solution of 1% formic acid and 2% acetonitrile and lyophilized. Prior to mass spectrometry analysis, dried peptide samples were re-dissolved in 50% acetonitrile and 0.1 % trifluoroacetic acid (TFA). A saturated solution of the MALDI matrix, α-cyano-4-hydroxycinnamic acid (CHCA), was prepared in 67% acetonitrile and 0.05% TFA, diluted to 70% saturation, mixed with the samples at 1:1 ratio (v/v) and 1 µl samples were spotted on the MALDI target. Mass spectrometry data were obtained using a 4700 Proteomics Analyzer, MALDI-TOF (Applied Biosystems, Foster City, CA, USA). Data acquisition and data processing were done using MassLynx 3.5 Mass Spectrometry Software (Waters), respectively. The instrument is equipped with a 355 nm Nd:YAG laser; the laser rate is 200 Hz. Reflectron and linear positive ion modes were used. Reflectron mode was calibrated at 50 ppm mass tolerance. Each mass spectrum was collected as a sum of 1000 shots. Theoretical masses of the peptides produced from a tryptic digestion of Tn5 transposase and H-NS were calculated using

PROWL's ProteinInfo peptide mass prediction tool (Rockefeller University, Laboratory of Mass Spectrometry and Gaseous Ion Chemistry).

4.2.7 Quantitative RT-PCR and analysis

1-2 ml of cells were removed from donor strain cultures immediately before the remaining 1 ml were mixed with recipient strain for the mating out. The cells were gently centrifuged (4000 x g for 4 minutes), resuspended in 200 µl of RNALater (Ambion) and incubated at 4°C overnight. The following day, RNA was extracted from each sample using the RNeasy Mini-kit (Qiagen). The quality and quantity of the final RNA samples were assessed using agarose gel electrophoresis and a NanoDrop spectrophotometer (IMPLEN) to measure A_{260}, 260/230 and 260/280 ratios. A portion of the RNA was treated with the RNase-free TURBO DNA-free Kit (Applied Biosystems) as per instructions for typical amounts of contaminating genomic DNA. The resulting RNA was quantified again as above to ensure 260/230 and 260/280 ratios were in the range of 1.5-2.0. RT-PCR was performed with a portion of this RNA (0.5-1 µg) using the protocol and reagents in the High Capacity RT-PCR Kit (Applied Biosystems). The final cDNA concentrations of 25-50 ng μl^{-1} (depending on the starting amount of RNA) were diluted to 25 ng μl^{-1} (if necessary) for use in real-time PCR reactions. TaqMan primers and probes (see Table 4.1) were designed using the Applied Biosystem Primer Express 2.0 software to the very 5' end (nucleotides 5-87) of the transposase transcript and the endogenous control 16S rRNA. 16S has been used as 'normalizing' control for Δ*hns* strains in other work (23). Reactions for real-time PCR were done in 20 µL volumes in 384 well clear plates using the protocol in the TaqMan Gene Expression MasterMix

guide (Applied Biosystems). Three biological replicates were tested per strain and three technical replicates were used for real-time PCR for each biological replicate. Appropriate control reactions were conducted where RNA was omitted, or reverse transcriptase was omitted for each RNA sample. Control real-time PCR reactions containing no cDNA (1X TE replacement) and no TaqMan 'enzyme mix' were also conducted. Reactions were run using standard cycle parameters on an Applied Biosystems 7900-HT Real-Time System. The Pfaffl mathematical model of relative quantification was used to determine the relative amounts of transposase mRNA (24).

$$Relative\ amount = \frac{(E_{transposase})^{\Delta Ct_{transposase}}}{(E_{16S})^{\Delta Ct_{16S}}}$$

where *E* is the PCR amplification efficiencies of the transposase transcript and 16S rRNA transcript that were determined by creating several standard curves with known diluted target concentrations ($E = 10^{(-1/slope)}$). The amplification efficiencies for the transposase target and 16S gene using these primers and probes were determined to be 1.99 and 1.92 respectively. The Ct represents the cycle threshold; the cycle at which the fluorescence (ΔR_n) detected is statistically significant as computed automatically by the real-time system. ΔCt of transposase or 16S represents the average sample Ct for each condition subtracted from the average Ct of the reference condition to which all samples will be quantified relative to (i.e. NK5830F' (*hns*+ strains) transformed with pDH533-4 (WT ME DNA)).

4.2.8 *In vitro* excision and strand transfer assays

For assembly of transpososomes ^{32}P-labeled mosaic end (ME) DNA (50 nM) and transposase (EK54/MA56/LP372) (500 nM) with or without purified H-NS (400 nM)

were mixed simultaneously in reaction buffer containing 25 mM Hepes (pH 7.5), 100 mM KCl, 1 mM β-mercaptoethanol, 50 μg mL^{-1} of BSA, 75 μg mL^{-1} of tRNA, as described in (9). After incubation at 37°C for 1 hour, target DNA (4 μM) and $MgCl_2$ (10 mM) were added to reactions as indicated in the figure and incubation was continued for up to 16 hours at 25°C. The target DNA 'T3-37' was used for insertion *in vitro* and was present at ~666-fold excess relative to the transpososome as previously described (25). At the indicated times aliquots were removed, treated with phenol, and the DNA was recovered by ethanol precipitation. DNA pellets were resuspended in denaturing loading dye and applied to an 8% sequencing gel. Gels were analyzed by phosphorimaging. Before the addition of target and metal transpososome complexes were analyzed to ensure quality and relative amounts. Transpososomes formed with and without H-NS were mixed at a ratio of 4:1 (v/v) with non-denaturing load mix and applied to a 5% polyacrylamide gel as previously described (26).

4.2.9 Modeling the H-NS-transpososome complex

The H-NS DNA-binding domain was manually positioned into the minor groove of the transposon end DNA, using the H-NS DNA-binding domain in complex with duplex DNA as a guide (adapted from PDB: ID 2L93) (27). The symmetry of the Tn5 transpososome was used to place a second H-NS DNA-binding domain in an equivalent position on the second transposon end. For the 'front face' complex, the structure of the H-NS N-terminal domain (NTD) dimer was manually positioned to place the C-terminal ends close to the DNA-binding domains (NTD PDB: ID 3NR7) (12). Residues 83-90 were added to the C-terminal end of the H-NS NTD and the connection between the two

domains was made using the loop-building utility in SwissPDBViewer (28). A similar process was used to position the H-NS dimer on the opposite side ('back face') of the transpososome, but in this case H4 was manually positioned into the major groove of the transposon end DNA to facilitate the connection between the N- and C-terminal domains. Minor adjustments and corrections to the stereochemistry were made using Coot (29).

4.3.0 Results

4.3.1 Base pair mutations within two putative H-NS binding sites reduce the affinity of H-NS for the Tn5 transpososome

We previously used hydroxyl radical footprinting to characterize the H-NS interaction with a Tn5 transpososome assembled with mosaic end (ME) sequences (9); the ME is a chimeric end composed of nucleotides from both the outside (OE) and inside ends (IE) of Tn5 and is optimized for use *in vitro* (30). Three potential H-NS binding sites were defined previously (see model in Figure 4.1B) (9). In this work we have tested the importance of each of the three sites for H-NS binding by introducing one or more base pair mutations into these sites. As H-NS binding to ME sequences is dependent on the presence of transposase (9), we targeted residues expected not to be critical for transposase-end interactions. This greatly limited the number of mutations we tested as transposase makes extensive contacts with the transposon end DNA (16, 31). The three different transposon end substrates we tested for H-NS binding are shown in Figure 4.1A.

We used an electrophoretic mobility shift assay (EMSA) to measure the impact of the above basepair changes on incorporation of H-NS into the Tn5 transpososome. In the

EMSAs shown in Figure 4.2A we simultaneously mixed H-NS and transposase with ^{32}P-labeled DNA in buffer lacking a divalent metal ion. These conditions favor the formation of an H-NS-bound transpososome, but no chemical steps in transposition take place. Under these conditions H-NS binding reached saturation for the WT ME and OE transpososomes at concentrations of 112 nM (Figure 4.2A, lane 6) and 340 nM (lane 14), respectively. This is manifested as an up-shift in the gel mobility of the respective transpososomes. Notably there was no appreciable binding of H-NS to OE substrate DNA that had not associated with transposase (Figure 4.2A, lane 8). This is indicative of H-NS binding to the transpososome with high specificity. In contrast, under the same reaction conditions we did not detect H-NS binding to either transpososomes formed with ME 3 (Figure 4.2A, lanes 16-20) or ME 8/9 (lanes 22-26) substrates. This clearly shows that the selected site 1 (ME 3) and site 2 (ME 8/9) mutations drastically impair H-NS binding to the transpososome without interfering with transpososome assembly.

We used binding data from EMSAs shown in Figure 4.2A and other similar experiments to generate binding curves from which we could calculate K_d values for the respective H-NS-transpososome interactions (Figure 4.2B). H-NS addition to transpososomes at concentrations below 100 nM caused a heterogeneous mixture of complexes (visualized by a 'smear') when analyzed by native gel, which made the two different transpososomes, and despite the difficulties in measuring H-NS-bound transpososome, we provide evidence for apparent K_d values as 51 ± 6.6 nM for H-NS to WT ME transpososome and 232 ± 67.1 for H-NS to the OE transpososome.

It should be noted that in the above experiments we don't know if H-NS is binding directly to the transpososome or to a pre-transpososome complex that is unstable

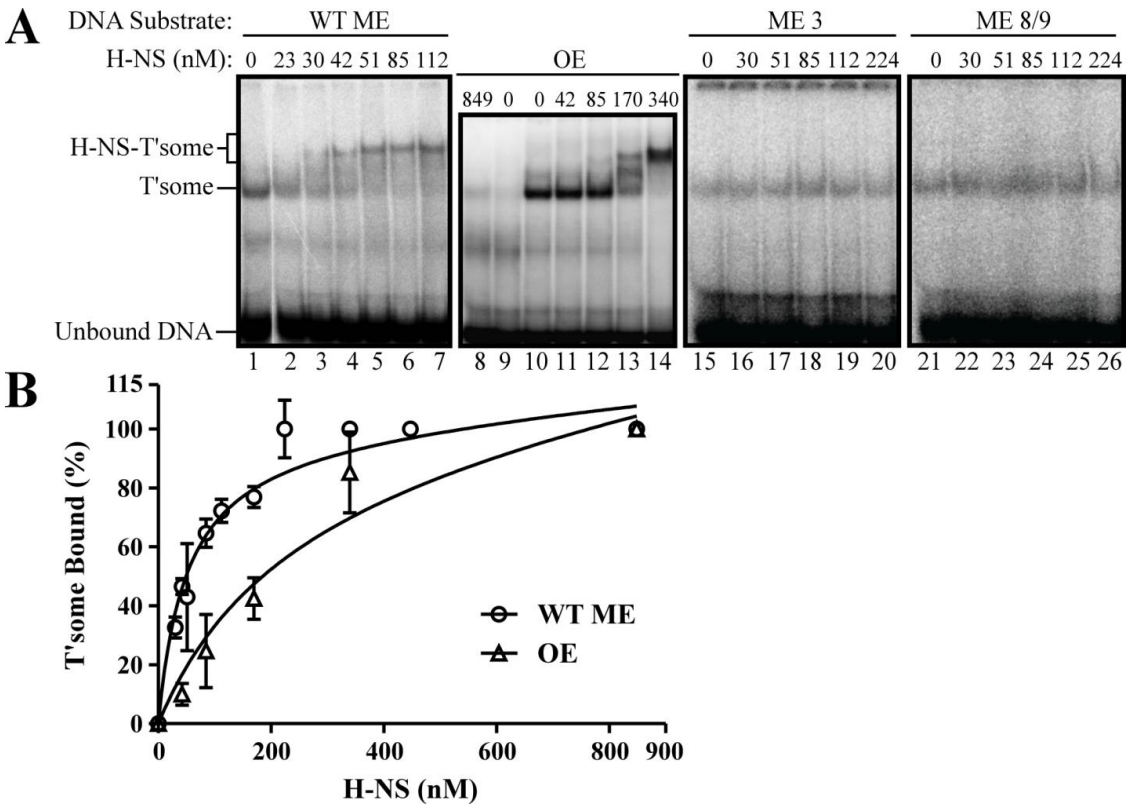


Figure 4.2. H-NS binding assays.

(A) H-NS and transposase were added simultaneously to the indicated ^{32}P-labeled transposon end substrate DNAs where indicated. Binding reactions were analyzed by gel electrophoresis on a native 5% polyacrylamide gel and subject to phosphorimager analysis. In each experiment transposase and substrate DNA concentrations were kept constant, while the concentration of H-NS was varied as indicated. Positions of transpososome (T'some), H-NS-shifted transpososome (H-NS-T'some) and unbound substrate DNA are indicated. In (A) lane 1 of the 'WT ME' gel was originally loaded in the last lane of the gel and was moved without alteration to the first lane of the gel. (B) Binding isotherms of fractional saturation as a function of H-NS concentration for H-NS binding to the wild-type ME and the OE transpososomes. Each binding isotherm was derived from four independent binding experiments similar to those depicted in (A). The percent transpososome complex shifted by H-NS was determined and these data were fit to a quadratic equation as described in Methods. The fits were used to provide estimates for the observed dissociation constant, obs K_d.

in the mobility shift assay and/or converts rapidly and irreversibly to a transpososome. For convenience we will, throughout the rest of the paper, refer to the mobility shift results as H-NS binding to the transpososome.

Taken together, the results in this section are consistent with sites 1 and 2 being particularly important for H-NS binding to the transpososome. It remains to be seen if site 3 plays a critical role.

4.3.2 Mutations within H-NS binding sites 1 and 2 reduce the ability of H-NS to regulate Tn5 transposition

Identifying transposon end mutations that reduce H-NS binding to the transpososome provided us with an opportunity to test the idea that the role H-NS plays in Tn5 transposition *in vivo* is a direct result of H-NS binding to the transpososome, as opposed to being an indirect effect. The latter is most definitely a concern given the large number of genes H-NS regulates and the pleiotropic nature of *hns* mutations (10, 13). If H-NS regulates Tn5 transposition directly by acting on the transpososome, then mutations that reduce H-NS binding to the transpososome should reduce the sensitivity of transposition *in vivo* to the *hns* status of the cell. In other words, the roughly 3- to 5-fold increase in transposition of Tn5 observed in *hns*+ versus Δ*hns* should be reduced when Tn5 contains a mutation in the transposon ends that reduces H-NS binding. To test our hypothesis we generated a series of plasmids containing mini-Tn5 transposons with either WT MEs, mutant MEs or OEs. We then measured the transposition frequency of these transposons *in vivo* in isogenic *hns*+/Δ*hns* strains of *E. coli* using a mating out assay. In this experimental set-up transposase (M56A) was provided on a second plasmid under the

control of its native promoter (the M56A mutation in transposase ensures that the transposase inhibitor protein is not synthesized making it easier to detect transposition events) (32).

The results for this experiment are presented in graphical form in Figure 4.3. Each graph shows the range of transposition frequencies obtained for at least 3 and up to 5 independent donor strain clones in multiple pair-wise (*hns*+ versus Δ*hns*) comparisons for a given transposon (each pair contains results from a single experiment). For each pair the 'fold difference' in transposition observed is indicated on the scatter plot. The results show a trend that roughly matches the *in vitro* H-NS-transpososome binding data for the different substrates. H-NS had the highest affinity for the WT ME and there was a consistent trend of reduced transposition in Δ*hns* versus *hns*+ (on average 5.6-fold) for this substrate (Figure 4.3A). Relative to the WT ME, the OE transpososome had the next highest affinity for H-NS and the OE transposon was slightly less sensitive relative to the WT ME transposon to *hns* status as transposition was reduced to a lesser degree (on average 3.6-fold) in Δ*hns* versus *hns*+ (Figure 4.3B). Finally, ME 8/9 and ME 3 transpososomes had the lowest affinity for H-NS and the *hns* status of the donor strain had little impact on the transposition frequency of the corresponding transposon substrates (no effect for ME 3 and a 2-fold effect for ME 8/9) (Figures 4.3C and D). It should be noted that each of the mutant ME transposons did transpose at a lower frequency than the WT ME transposon (10-fold for ME 3 and 50-fold for ME 8/9), indicating that the mutations do have a negative impact on transposition independent of their effects on H-NS binding. As an additional control we measured relative copy levels of the two transposition plasmids to ensure that the mating out results were not due to

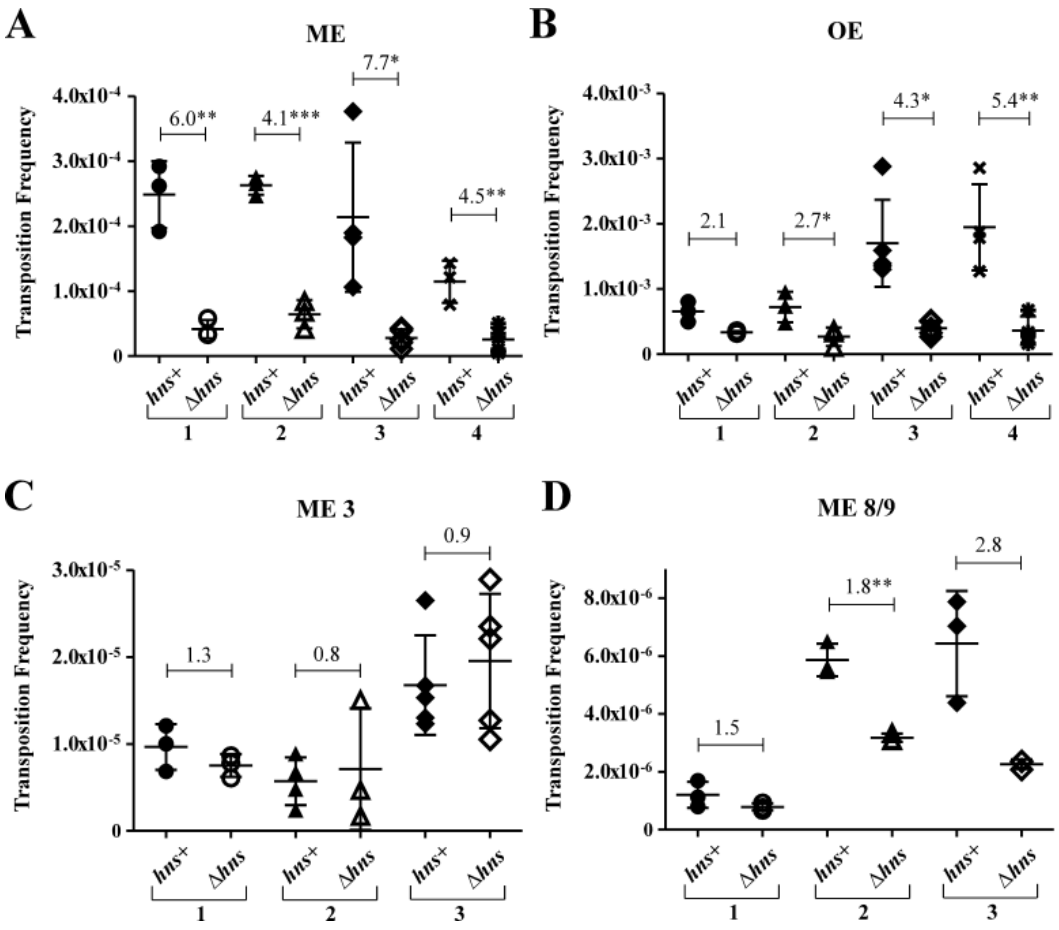


Figure 4.3. Transposition frequencies of mini-Tn5 derivatives in isogenic *hns*+ and Δ*hns* strains.

Mating out experiments were carried out on *hns*+ and Δ*hns* strains of *E. coli* containing a plasmid encoding one of the mini-Tn5 derivatives indicated and a compatible plasmid encoding transposase (M56A) expressed from its native promoter (see 'Materials and methods' for details). Each graph provides transposition frequencies of individual clones (3-5 per strain) for at least 3 and up to 4 independent experiments. Each experiment compared transposition levels for a pair of isogenic strains containing the same mini-Tn5 substrate. Horizontal and vertical lines are the mean and the standard deviation, respectively, for the transposition frequencies. The fold difference based on the mean transposition frequency between *hns*+ and Δ*hns* strains for each experiment is indicated in the upper portion of each graph. The t-test to measure variance was used to determine if differences between frequencies within each mating out experiment were/were not due to chance. *$p<0.05$, **$p<0.01$, ***$p<0.001$.

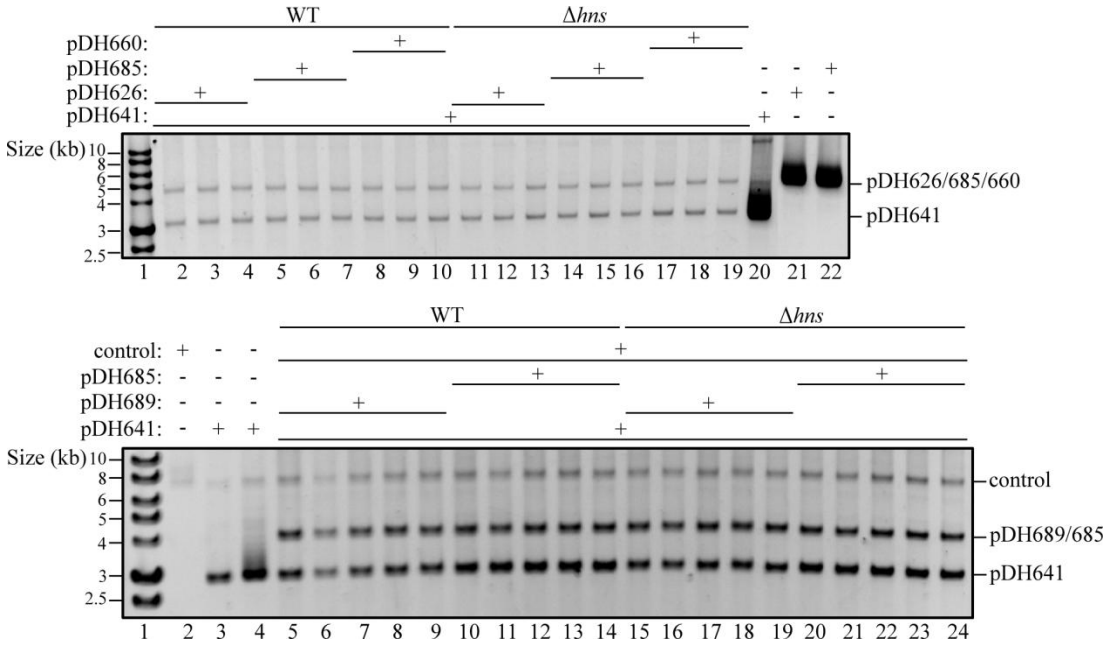


Figure 4.4. Relative copy number of transposition plasmids transformed into *hns*+ and Δ*hns* strains used for mating out assays.

Relative copy number levels of the plasmids transformed into mating out donor strains are shown for 3-4 biological replicates for each of the substrate plasmids (pDH626 - ME, pDH689 - OE, pDH685 – ME3, and pDH660 – ME8/9). Copy number was determined immediately prior to mating to ensure that Δ*hns* status did not influence the copy number or maintenance of any plasmid. Note that pDH641 encoding transposase is present in every sample. Although only representative gels are shown, plasmid levels were determined for each donor transformant each time a mating out assay was conducted. Briefly, the optical density was determined for each sample to provide an estimate of the amount of cells. Equivalent cell density amounts from each sample were used for plasmid isolation using Mini-preps (Sigma-Aldrich). In the bottom gel, an additional control was carried out to account for differences in plasmid recovery. An equivalent density of cells from DH5α transformed with a control plasmid of different size than the transposition plasmids was added to each donor sample before plasmid isolation was conducted. DNA from each sample (approximately 5-10 µl), along with plasmid control samples from previous plasmid isolations to indicate proper sizes, were run on a 1% agarose gel, stained with ethidium bromide, and viewed using an AlphaImager (AlphaInnotech).

differences in plasmid copy amounts. Figure 4.4 indicates that copy levels are not significantly different between *hns*+ and Δ*hns* donor strains used in the mating out assays and therefore the resulting transposition frequencies are not affected by different amounts of transposition plasmids.

4.3.3 H-NS status does not influence Tn5 transposase steady-state mRNA levels

If H-NS stimulates Tn5 transposition by acting directly on the transpososome, we would not expect the difference in transposition frequency in *hns*+ versus Δ*hns* to be linked to differential expression of the transposase gene. As H-NS is a global regulator of gene expression we thought it important to rule this possibility out. To test this we measured steady-state transposase transcript levels by quantitative RT-PCR using RNA isolated from *hns*+ and Δ*hns* cells used in mating out experiments. The results of this analysis show that transposase expression levels are marginally lower in *hns*+ versus Δ*hns* cells (Figure 4.5). The data shown has been normalized to 16S rRNA levels measured in the two strains; 16S rRNA levels are known not to be influenced by H-NS (23). Thus, the higher transposition frequencies observed in *hns*+ versus Δ*hns* strains cannot be accounted for by reduced transposase gene expression in Δ*hns*.

4.3.4 Chemical cross-linking indicates that transposase and H-NS interact in the context of the transpososome

The results from our binding studies suggest that H-NS is in close proximity to transposase within the transpososome. Determining if there are direct interactions between transposase and H-NS has mechanistic implications for how H-NS promotes

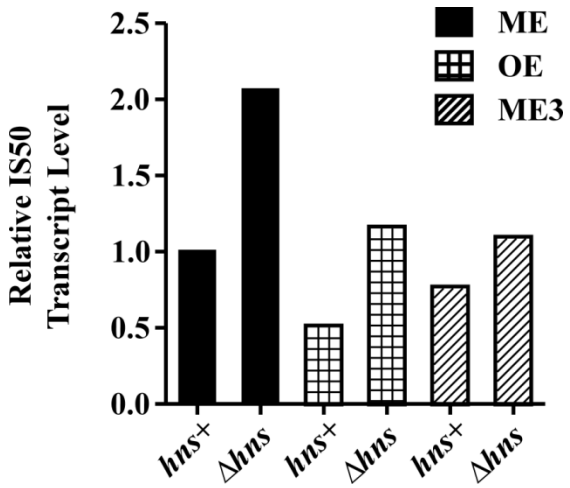


Figure 4.5. Quantitative RT-PCR analysis of transposase transcript levels in *hns*+ and Δ*hns* strains.

Total RNA from donor strains used in mating out experiments was isolated at mid-log phase of growth. Roughly equivalent amounts of RNA were reverse transcribed using random primers and transposase levels were quantified using real-time PCR (see 'Materials and methods'). We simultaneously measured 16S rRNA levels from *hns*+ and Δ*hns* cells. The relative levels of IS50 transposase transcript presented have been normalized to the corresponding levels of 16S rRNA transcript. Each normalized value represents an average from 4 independent transformants used in a single mating out experiment and all are compared to the reference strain, *hns*+/ME strain.

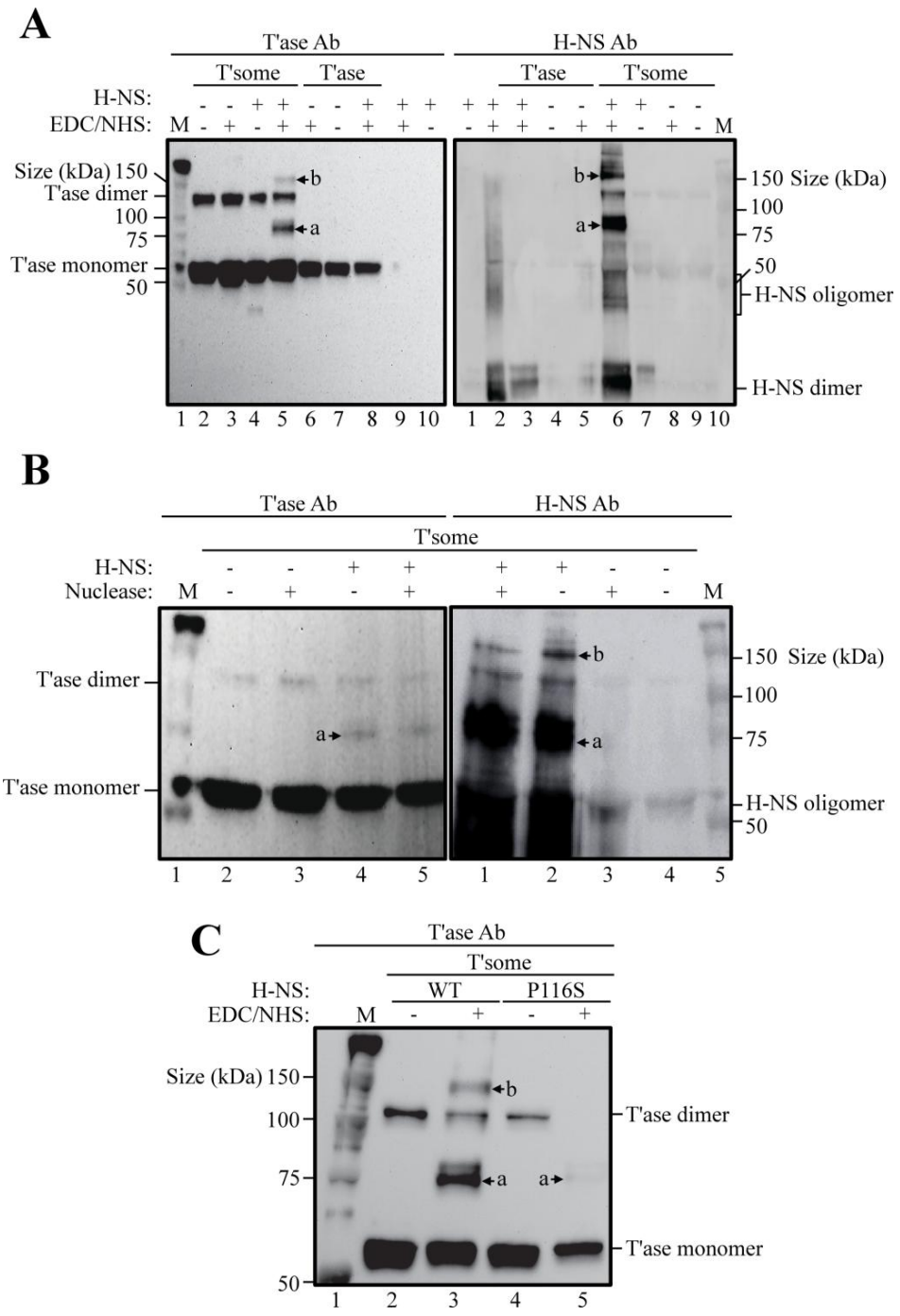
Tn5 transposition. Only a few proteins have been shown to directly interact with H-NS (33, 34, 22). Intriguingly, one of these proteins is the Tn10 transposase. We used the same approach previously reported for the Tn10 system for assessing if Tn5 transposase directly interacts with H-NS (22). Briefly, we treated transpososome assembly reactions (+/- H-NS) or purified Tn5 transposase (+/- H-NS) with the protein cross-linking reagent EDC/NHS; this reagent is a zero-length cross-linker that covalently links carboxyl and amino groups. Subsequently, EDC/NHS or mock treated transpososomes were gel-purified and proteins eluted from these gel slices were analyzed by Western blotting and mass spectrometry.

In the Western blot analysis we were looking for a product(s) that in the presence of EDC/NHS had a reduced mobility on an SDS gel relative to transposase (the larger of the two proteins – 60 kDa versus 16 kDa) and was detected by both antibodies to transposase and H-NS. Note that we used a version of Tn5 transposase protein containing the T7 *gene 10* peptide as an N-terminal epitope tag, thus allowing us to use a commercially available monoclonal antibody for transposase detection. For H-NS detection we used a polyclonal H-NS antibody. After probing a blot with one antibody, the blot was stripped and re-probed with the other antibody.

We show in Figure 4.6A that EDC/NHS treatment of a transpososome assembly reaction containing H-NS yielded two prominent novel products ‘a’ and ‘b’ that were detected by both antibodies and have an apparent molecular mass greater than the mass of monomeric transposase (lane 5, left panel; lane 6, right panel). Products ‘a’ and ‘b’ were not detected in the mock cross-linking reaction where EDC/NHS was omitted (lane 4, left panel; lane 7, right panel) and when H-NS was not included in the assembly reaction

Figure 4.6. H-NS-transposase cross-linking.

(A) and (B) Transpososomes assembled in the presence of H-NS were subjected to cross-linker (EDC/NHS) treatment as described in 'Materials and methods'. The resulting H-NS-transpososome was purified from a native polyacrylamide gel and proteins eluted from this gel slice were subjected to Western blot analysis using a monoclonal antibody to the T7 *gene 10* epitope present on the N-terminus of transposase. Subsequently, the blot was stripped and re-probed with a polyclonal antibody to H-NS. Signal from the marker lane was used to align each pair of blots. In (A) EDC/NHS cross-linking was also carried out on reactions containing purified transposase and H-NS in the absence of DNA. In (B) aliquots of cross-linked H-NS-transpososome were treated with micrococcal nuclease prior to Western blot analysis. In (C) either WT H-NS or a mutant form of H-NS (P116S) was included in the transpososome assembly reaction and cross-linking and Western blot analysis was carried out as above except that the blot was probed only with the antibody to the epitope tag on transposase. H-NS monomer is not detected in any of the blots because it was run off the gel. M, molecular weight markers; T'some, gel-purified H-NS-transpososome; T'ase, purified transposase protein; 'a' and 'b', species inferred to contain transposase cross-linked to H-NS.


A
T'ase Ab
H-NS Ab
T'some
T'ase
T'ase
T'some
H-NS:
EDC/NHS:
M
Size (kDa) 150
T'ase dimer
100
75
T'ase monomer
50
b
a
150 Size (kDa)
100
75
50
H-NS oligomer
H-NS dimer
1 2 3 4 5 6 7 8 9 10
B
T'ase Ab
H-NS Ab
T'some
H-NS:
Nuclease:
T'ase dimer
T'ase monomer
150 Size (kDa)
100
75
H-NS oligomer
50
1 2 3 4 5
C
T'ase Ab
T'some
H-NS:
WT
P116S
EDC/NHS:
Size (kDa) 150
100
75
50
T'ase dimer
T'ase monomer

(lane 3, left panel; lane 8, right panel). Products ‘a’ and ‘b’ were also not detected when transposase was mixed with H-NS in the absence of ME DNA (lane 8, left panel; lane 3, right panel), indicating that the appearance of these products is dependent on transpososome formation. However, products ‘a’ and ‘b’ were detected after micrococcal nuclease treatment of gel-purified EDC/NHS-treated H-NS-transpososomes indicating that these products do not include a DNA component (Figure 4.6B). Finally, gel slices containing products ‘a’ and ‘b’ were digested with trypsin and analyzed using mass spectrometry (MALDI-TOF) (Figure 4.7). The spectra obtained were compared to a theoretical prediction of mass-to-charge ratios of the resulting peptides from a tryptic digest of transposase and H-NS. Several peaks/peptides corresponding to both proteins were detected in each of ‘a’ and ‘b’ samples providing definitive proof that H-NS and transposase are present in products whose formation is dependent on EDC/NHS treatment. The mass spectrometry did not provide information regarding residues involved in the cross-linking as there were no peaks/peptides that were specific to the cross-linked products.

At this point we can only speculate on the precise composition of products ‘a’ and ‘b’ based on their relative abundances and apparent molecular weights. Product ‘a’ is the more abundant product and accordingly is most likely the simplest in terms of composition. With an apparent molecular weight greater than 60 and less than 120 kDa we expect this product includes a monomer of transposase cross-linked to either a monomer of H-NS (76 kDa) or a cross-linked dimer of H-NS (92 kDa). Product ‘b’ could be a cross-linked transposase dimer that is itself cross-linked to an H-NS monomer or a cross-linked H-NS dimer. As H-NS readily forms dimers in solution that are efficiently

cross-linked by EDC/NHS (35) it is tempting to speculate that each of products ‘a’ and ‘b’ includes a cross-linked H-NS dimer.

We also tested the ability of a mutant version of H-NS (H-NS P116S) to form cross-links with transposase. H-NS P116S is defective in both structure-specific DNA-binding and higher order oligomerization (35, 36) and we have previously shown that it binds weakly to the Tn5 transpososome, probably through non-specific interactions (9). Upon treating a transpososome assembly reaction that included P116S H-NS with EDC/NHS we saw only a trace amount of product ‘a’ (Figure 4.6C, compare lanes 3 and 5) suggesting that H-NS must be precisely positioned within the transpososome in order for it to stably interact with transposase.

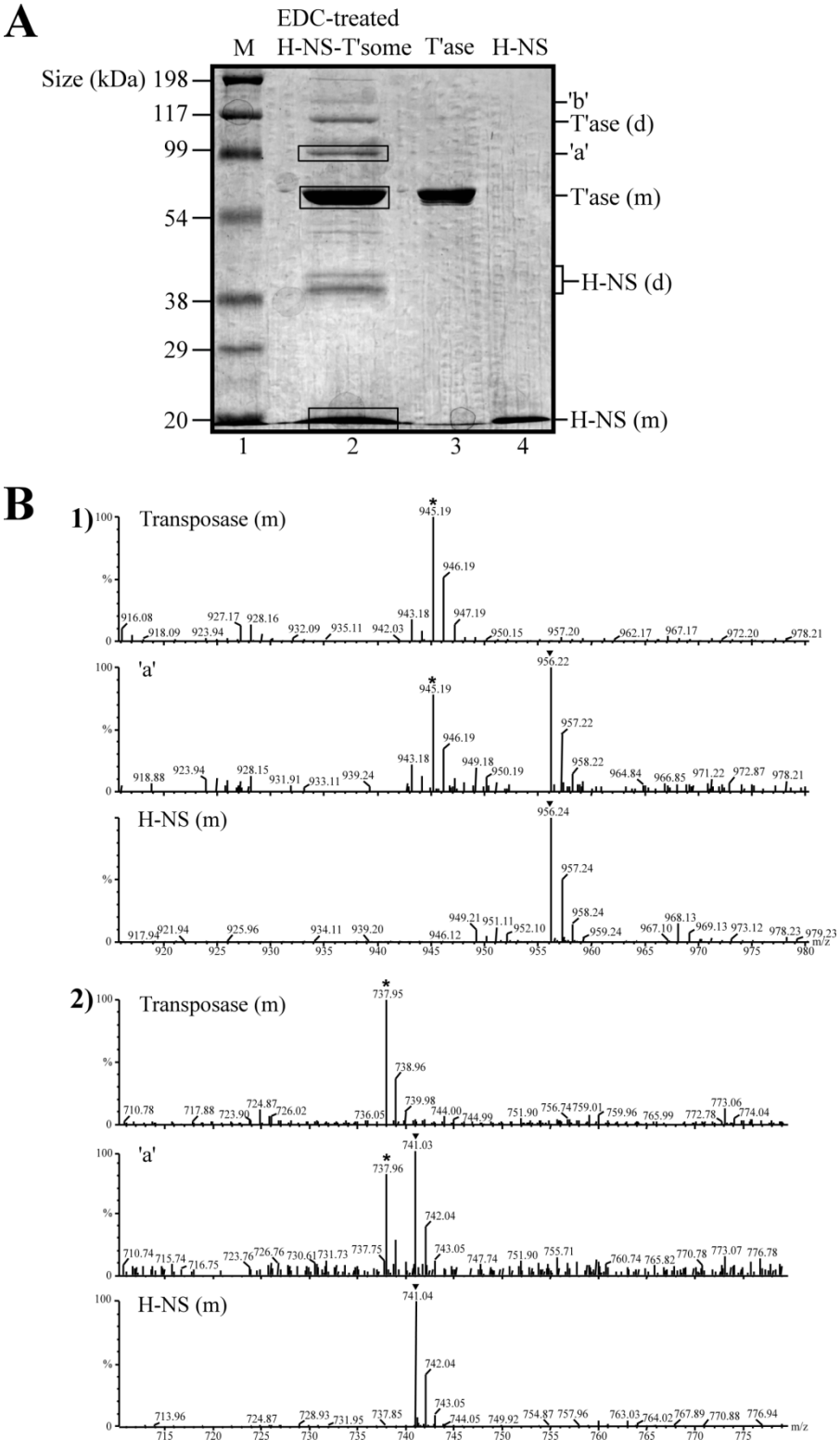
Overall the results in this section indicate that within the Tn5 transpososome H-NS is in close enough proximity to directly interact with transposase. Given that H-NS is a DNA-binding protein this raises the possibility that H-NS may help tether transposase to transposon end sequences by interacting both with transposase and DNA. Finally, based on ‘a’ being the most abundant cross-linked product it seems likely that a single dimer of H-NS is most frequently associated with the Tn5 transpososome under the reaction conditions used here.

4.3.5 Modeling H-NS into the Tn5 transpososome structure

Our biochemical data is most consistent with a dimer of H-NS binding the Tn5 transpososome through interactions with segment 2 of the transposon end DNA. In an attempt to integrate this data with the available structural data, we asked if an H-NS dimer could be docked into the existing Tn5 transpososome structure. An NMR structure

Figure 4.7. Mass spectrometry of cross-linked product ‘a’.

(A) Coomassie blue-stained SDS gel containing gel-purified EDC/NHS cross-linked H-NS-transpososome (lane 2), purified transposase (lane 3) and purified H-NS (lane 4). Plugs of acrylamide from the boxed areas including transposase, H-NS and product ‘a’ were excised using a ‘spot-picker’, digested with trypsin, and analyzed by MALDI-TOF. (B) Two sets of spectra derived from transposase monomer (m), H-NS monomer and ‘a’ bands are shown. In both sets the spectra from ‘a’ have peptides that are also found in the spectra derived from transposase and H-NS. Peaks are separated on the x-axis based on mass/charge ratio. *, peptides corresponding to transposase; ▼, peptides corresponding to H-NS.


A
EDC-treated
M
H-NS-T'some
T'ase
H-NS
Size (kDa)
198
117
99
54
38
29
20
'b'
T'ase (d)
'a'
T'ase (m)
H-NS (d)
H-NS (m)
1
2
3
4
B
1)
Transposase (m)
'a'
H-NS (m)
2)
Transposase (m)
'a'
H-NS (m)
m/z

is available for the H-NS DNA-binding domain, residues 91 to 137 (27). On the basis of chemical shift experiments, Gordon and co-workers modeled the H-NS binding domain into a DNA duplex containing a 5'-ATATAT-3' sequence. We used this model of the DNA/H-NS complex to help position the H-NS DNA-binding domain in the minor groove of the transposon end DNA in the Tn5 transpososome (16), close to positions 8 and 9; recall that mutations at these positions abrogated H-NS binding to the transpososome. A second H-NS DNA-binding domain was placed in the equivalent position on the second transposon end. Positioned in this manner, the distance between the two DNA-binding domains is approximately 80 Å (Figure 4.8). A crystal structure is available for residues 2 to 82 of the N-terminal oligomerization domain of H-NS (12): this domain consists of 4 α-helices (H1, H2, H3 and H4) that form an extended dimer. In the crystal, H4 mediates oligomerization between symmetry-related H-NS dimers; this interaction is thought to be biologically important because of its potential role in DNA compaction. However, our available data suggests that the Tn5 transpososome interacts with a single H-NS dimer. The distance between the C-terminal ends of H3 in the H-NS dimer is approximately 100 Å, which is close to the distance of 80 Å between the H-NS DNA-binding domains that were positioned in the Tn5 transpososome (Figure 4.8). On this basis, we were able to connect the ends of the H-NS N-terminal domain dimer to the C-terminal DNA-binding domains using the flexible linker, residues 83 to 91 (Figure 4.8A). In this configuration, the H-NS dimer runs across the region of the transpososome implicated in target capture (37). Given the flexibility in the position of H4 and residues 83 to 91, the H-NS dimer could also cross the opposite side of the transpososome (Figure 4.8B). In this case, H4 was positioned to pass through the major groove of the transposon

end DNA to facilitate connection with the DNA-binding domain. In this configuration H-NS could remain bound to the transpososome without interfering with target capture.

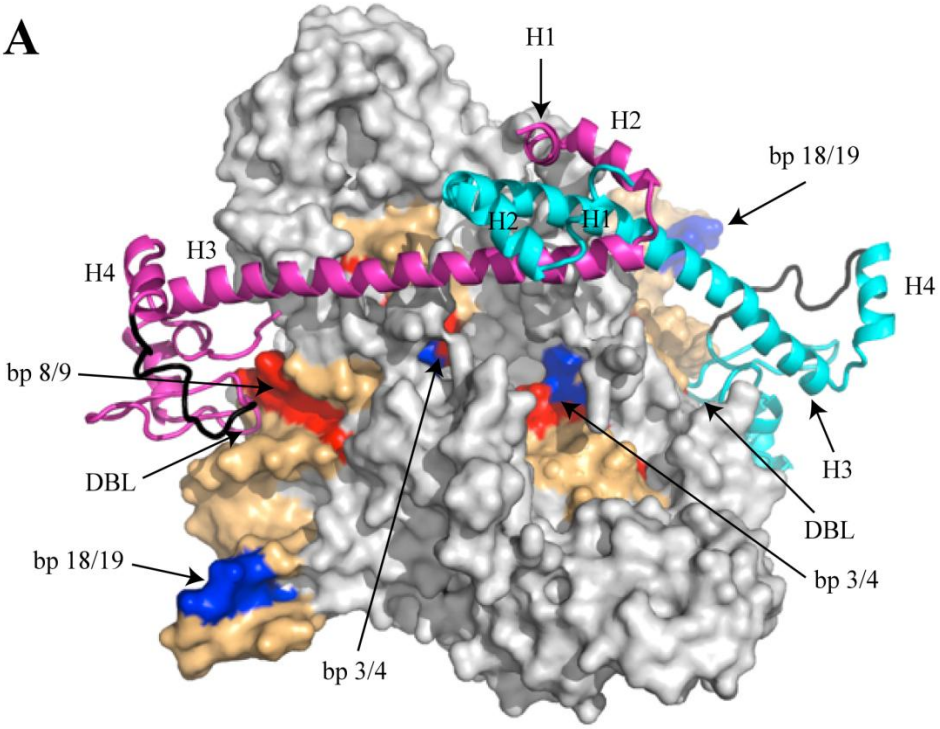
There are many possible ways the dimerization domain can be oriented relative to the DNA-binding domain and this orientation will dictate where along the transpososome face H1-H3 is positioned. At present we have chosen an orientation between the domains that limits the number of steric clashes between H1-H3 and either the 'front' or 'back' face of the transposase dimer. While the models are obviously preliminary, we think they are useful because they show that the dimensions of the head-to-head H-NS dimer and the transpososome are compatible for binding in a manner where residues established to be important in transpososome-H-NS interactions (positions 8 and 9 of the transposon ends) are the main anchor points of the structure.

4.4.0 Discussion

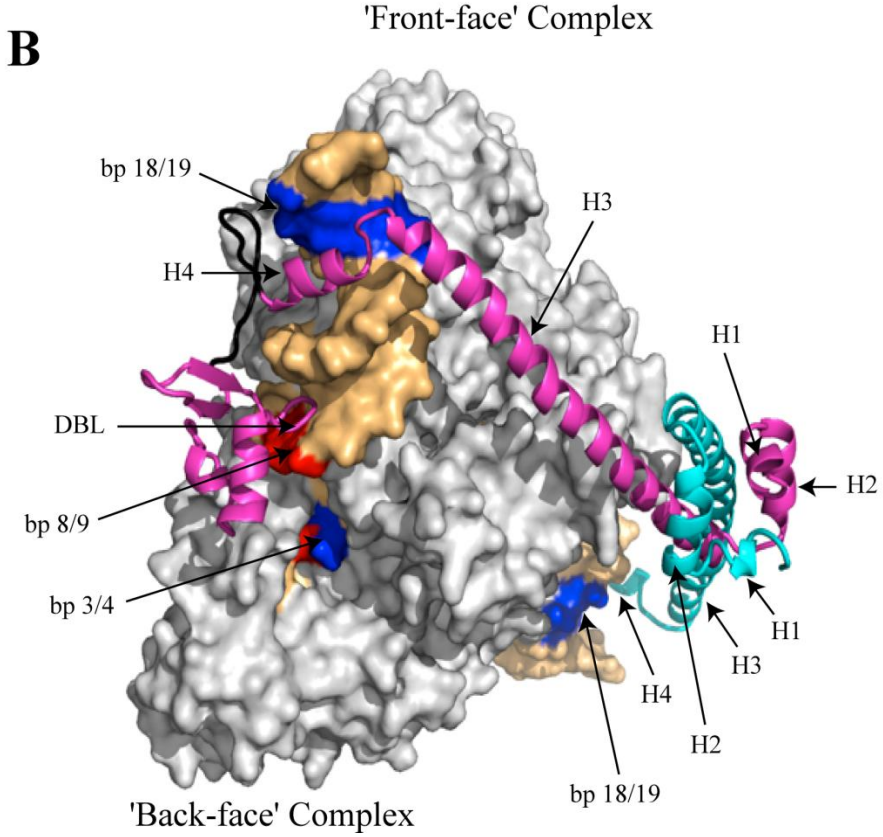
H-NS promotes Tn5 transpososome formation through an as yet undefined mechanism that involves incorporation of H-NS into the transpososome. In this work we measured the binding affinity of H-NS for the Tn5 transpososome (or possibly a pre-transpososome transposition complex) and went on to identify through mutational analysis base pairs within Tn5 end sequences that play an important role in this interaction. We have also shown through protein-protein cross-linking analysis that H-NS directly interacts with transposase specifically in the context of the transpososome and presume that this interaction contributes significantly to the relatively high affinity with which H-NS binds the transpososome. Defining mutations within the Tn5 ends that strongly decreased H-NS binding to the transpososome afforded us the opportunity to ask if H-NS promotes Tn5

Figure 4.8. Models for H-NS binding to the Tn5 transpososome.

(A) and (B) Models wherein a head-to-head dimer of H-NS is docked to either the 'front' or 'back' side of the transpososome. Helices 1-4 of H-NS are labeled, as are specific base pairs in the transposon ends that were mutated in this study (red). Individual monomers of H-NS are in different colors (cyan and magenta) and the residues linking H4 and the DNA-binding domain of H-NS are in black. Tn5 transposase residues are in grey and end DNA sequence is in gold, blue or red. DBL; DNA-binding loop.



'Front-face' Complex



'Back-face' Complex

transposition *in vivo* by directly binding the transpososome. We found mutations that inhibited H-NS-transpososome interactions *in vitro* reduced the sensitivity of transposition reactions to the *hns* status of the cell, a finding consistent with H-NS acting directly on the transpososome to promote transposition.

4.4.1 Transposon end sequences and transposase provide determinants for H-NS binding to the Tn5 transpososome

H-NS typically binds AT-rich DNA within promoter sequences independent of interactions with other proteins (13). Within this context a wide range of binding affinities have been reported ranging from micromolar to low nanomolar. One of the tightest interactions reported to date for *E. coli* H-NS involves the promoter sequence of the *proU* operon where a 10 bp segment within this regulatory sequence was found to bind H-NS with a K_d of 15 nM and to serve as a nucleating sequence for H-NS binding to nearby sites possessing intrinsically lower H-NS binding affinities (11). The highest affinity interaction for H-NS reported to date is with a Tn10 transpososome where the reported K_d was 0.3 nM. In this case H-NS binding determinants included both transposon end sequences and the Tn10 transposase protein (22). In the current study we measured the binding strength of H-NS for Tn5 transpososomes containing different transposon end sequences and reported a K_d value of approximately 51 nM for the WT ME end sequence. We have also shown that H-NS interacts with Tn5 transposase and accordingly infer that, as in the Tn10 system, the transposase protein provides determinants for H-NS binding.

H-NS could promote transpososome assembly by directly binding to the fully-assembled transpososome and stabilizing this structure. Alternatively, H-NS could promote transposase binding to a transposon end thereby facilitating formation of a single-end complex and/or promoting the pairing of single-end complexes. As transpososomes tend to be inherently stable structures (38, 39), we favor the idea that H-NS acts prior to transpososome formation. Consistent with this we have previously shown that H-NS can bind a single-end Tn5 transpososome complex (9). Unfortunately, to this point, a Tn5 single-end complex has only been detected using a mutant form of transposase that is unable to transition into a transpososome (40), so it has not yet been possible to further dissect the role of H-NS in transpososome formation.

Two host proteins, DnaA and Fis, interact with Tn5 end sequences and could potentially compete with H-NS for binding. However, Fis binds only the inside end sequence of Tn5 with high affinity and therefore is not directly relevant to the current work where we have focused on mosaic and outside ends (41). In contrast, both the ME and OE contain a single DnaA binding site (42) and this site overlaps putative H-NS binding sites 2 and 3 (Figure 4.1A). Interestingly it has been reported that DnaA can out-compete transposase for OE binding even after a transpososome has been formed (43). This raises the possibility that *in vivo* transposase loading onto an OE or ME may require an additional host factor such as H-NS. The K_d of DnaA for a DNA sequence with a single DnaA site is 30-50 nM (44) and keeping in mind our determination of a K_d of 51 nM for the H-NS interaction with Tn5 ME-transpososome, it is possible that H-NS would be able to effectively compete with DnaA for ME binding *in vivo* and thereby significantly contribute to transpososome assembly.

4.4.2 Point mutations in putative H-NS binding sites within ME sequences greatly reduce H-NS binding to the Tn5 transpososome

We used information from hydroxyl radical footprinting on Tn5 transpososomes formed with ME DNA (plus and minus H-NS) to guide us in making point mutations in ME DNA that potentially would be useful in: (1) further defining an H-NS binding site(s) in the transpososome and (2) assessing if *in vivo* H-NS acts directly on the transpososome to regulate Tn5 transposition. A single base pair change in putative binding site 1 and a double base pair change in putative binding site 2 both strongly reduced transpososome formation; we were not able to reliably measure K_d values for the H-NS interaction with these transposon end substrates because band shifts for these transpososomes were only observed at H-NS concentrations where non-specific binding of H-NS to unbound end sequences occurred (data not shown). In contrast, the 3 base pair differences in the OE versus the ME sequence did not affect the affinity of H-NS for the transpososome as strongly as the other mutations. Thus, to this point the data are consistent with sites 1 and 2 but not site 3 harboring the most critical determinants for H-NS binding to the transpososome.

NMR analysis of the C-terminal DNA-binding domain of *E. coli* H-NS in complex with a 15 bp DNA fragment has provided important new insight into how H-NS contacts DNA. Based on the mapped binding interfaces and the structure of the DNA-binding domain, a docking model for the H-NS C-terminal domain bound to DNA was generated. The model is consistent with H-NS interacting with DNA through the minor groove over 5 consecutive residues. Moreover, results from a protein-binding microarray

study revealed that the major determinant for optimal H-NS binding is the shape of the minor groove, which is dictated by local DNA sequence (27). Mixed AT-rich sequences or A-tracts within a GC-rich sequence appear to have the optimal minor groove geometry for H-NS binding. Additionally, the presence of GC base pairs in the centre of an AT-tract is unfavorable for H-NS binding because of the less optimal electrostatic potential for binding arginine residues, which are present in the DNA-binding motif (an AT-hook-like loop) of H-NS (45). Also, the presence of a 2-NH_2 group on G that protrudes into the minor groove may provide a steric block to H-NS binding. Within the Tn5 transpososome transposase makes extensive contacts in the major groove, including base-specific contacts spanning residues 7-13. In contrast, there are fewer minor groove contacts made in the transpososome and the only A-T rich segment, which spans residues 8-12, faces away from the transposase dimer on the minor groove side (16). As we have mutated residues 8 and 9 from T:A to G:C and C:G base pairs respectively and seen a drastic reduction in H-NS binding, it seems likely, given the preference of H-NS for AT-rich sequences embedded in GC-rich sequences, that these mutations would directly disrupt an H-NS binding site. It is less obvious why the mutation in putative H-NS binding site 1 would strongly inhibit H-NS binding. The mutation does not alter the AT-content and it is therefore likely that the geometry of the minor groove is not significantly altered by this mutation. Furthermore, the minor groove spanning positions 1-3 faces into the space occupied by the transposase dimer and although there are no direct contacts with the minor groove in this segment it is not obvious that there is room to accommodate a molecule of H-NS at this position. As both bases at position 3 normally make transient

contacts with transposase it is also possible that the position 3 mutation interferes with H-NS binding indirectly by altering the conformation of transposase within the complex.

While we think it is more likely that H-NS binds directly to site 2 versus site 1, it is remarkable that a mutation in either site can abrogate H-NS binding to the Tn5 transpososome. This result implies that H-NS contacts with DNA residues in site 2 are not on their own sufficient to provide stable H-NS binding. If stable binding requires both H-NS contacts with DNA and transposase, then the above conundrum can be solved by arguing that the site 1 mutation interferes with the ability of transposase to contact H-NS, while the site 2 mutations interfere with the ability of H-NS to contact DNA (within site 2), and that both types of contacts are necessary for tight H-NS binding. Clearly, it will be important in the future to define the relative contribution H-NS-transposase interactions make towards the energetics of H-NS binding in this system.

4.4.3 Genetic evidence that H-NS regulates Tn5 transposition by binding directly to transposition complexes

Our identification of transposon end mutations that strongly interfere with H-NS binding to the Tn5 transpososome afforded us the opportunity to test the idea that H-NS up-regulates Tn5 transposition *in vivo* by acting directly on transposition complexes. Our expectation was that the transposition frequency of a transposon harboring such a mutation(s) would not be influenced by the *hns* status of the cell. In other words, the transposition frequency of such a transposon would not increase in an *hns*+ relative to an isogenic Δ*hns* strain. Consistent with the idea that H-NS acts directly on the transpososome to promote Tn5 transposition, we found that the transposition frequency

of Tn5 elements harboring end mutations that strongly interfered with H-NS binding to the transpososome *in vitro* was largely insensitive to *hns* status. Interestingly, the OE transposon, which exhibited an intermediate level of H-NS binding *in vitro* (relative to the ME and ME 3 and ME 8/9) also showed an intermediate response to *hns* status. That is, transposition increased in *hns*+ but the increase was about half the value for the WT ME element and close to double the value for the ME 8/9 element.

Another possible reason for why ME 3 and ME 8/9 transposition levels were insensitive to *hns* status might be that the mutations essentially crippled the transposition machinery. As an extreme example one would not expect to see a significant impact of *hns* status on transposition in the case of a catalytically dead transposase mutant. However, the site 1 mutation (present in ME 3) only reduced the transposition frequency about 10-fold and *in vitro* had little effect on transpososome assembly. As there is not a severe block imposed by this mutation, one might anticipate that a factor that increases the efficiency of transpososome formation would ultimately increase the transposition frequency. However, this would not occur if that factor failed to bind the mutant transpososome.

The site 2 mutation (in ME 8/9) had a more severe effect on transposition (~50-fold reduction) and did show a reduced ability to form transpososome *in vitro*, as we had to increase the transposase concentration approximately 2-fold to get levels of transpososome equivalent to those formed in a WT ME reaction. Base pairs 8 and 9 combined make three base-specific *cis*-contacts with transposase (16) and therefore it is likely that these mutations primarily act at the level of initial binding of transposase to a transposon end. In principle, H-NS could suppress this type of defect but such

suppression would not occur in the instance where H-NS cannot bind the transposon ends (as is the case *in vitro* for the ME 8/9 substrate).

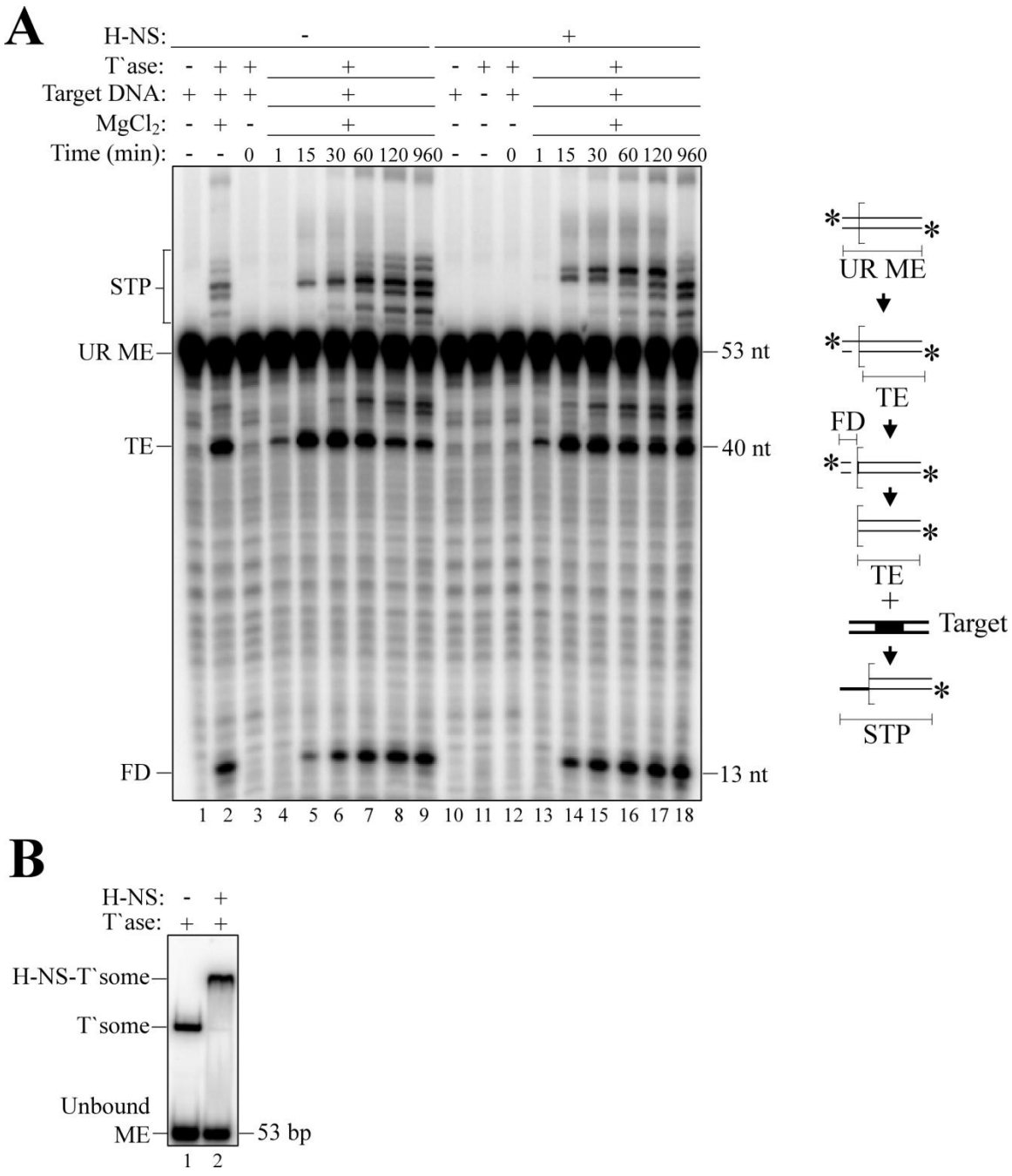
While our genetic data is consistent with H-NS acting directly on the transpososome to up-regulate Tn5 transposition, we also thought it was important to assess if the H-NS status of our strains influenced transposase mRNA levels as transposase expression is expected to be the major level of control for Tn5 transposition (43). For example, the 3- to 5-fold increase in Tn5 transposition in the *hns*+ versus Δ*hns* strain could be due to H-NS reducing the expression of a negative regulator of Tn5 transposase transcription. However, our quantitative RT-PCR analysis revealed this not to be the case as Δ*hns* strains actually had marginally higher (close to 2-fold) steady-state levels of transposase transcript compared to *hns*+ strains.

4.4.4 Model for H-NS function in Tn5 transposition

We favor the idea that H-NS functions at an early stage in Tn5 transposition because we previously showed that addition of H-NS to a transpososome assembly reaction under conditions suboptimal for transpososome formation resulted in a strong stimulation in transpososome formation. Additionally, when we added H-NS to reactions with pre-formed transpososome, we did not see an increase in either excision or strand transfer product levels (Figure 4.9). In fact, we do not see an increase in either excision or strand transfer *in vitro* no matter what the order of H-NS addition is unless the addition occurs under reaction conditions where transpososome assembly is suboptimal (9). As it has not been possible to trap a single-end complex (presumably one transposase monomer bound to one transposon end) in the Tn5 system, we presume that such a complex is inherently

Figure 4.9. Kinetic analysis of *in vitro* Tn5 transposition with H-NS.

(*A*) Illustrations of the substrate and products are shown to the *right* of the gel. The square bracket defines the position of the transposon-donor junction. The strand transfer products (STP) result from joining of the transferred strand of the transposon end (TE) to either the top or bottom strand of the T3-37 target site (see Chapter 2, Figure 2.1C). Different sized STPs are generated because T3-37 contains three overlapping 9 bp target sites that are asymmetrically positioned within the target DNA fragment. *, 5′ ^{32}P labels; UR ME, unreacted mosaic end substrate; FD, flanking donor DNA; nt, nucleotides; T'some, transpososome; T'ase, transposase. (B) Native gel analysis of transpososomes assembled for the transposition analysis carried out in (A).


A
H-NS:
T`ase:
Target DNA:
MgCl2:
Time (min):
STP
UR ME
TE
FD
53 nt
40 nt
13 nt
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
UR ME
TE
FD
TE
+
Target
STP
B
H-NS:
T`ase:
H-NS-T`some
T`some
Unbound
ME
53 bp
1 2

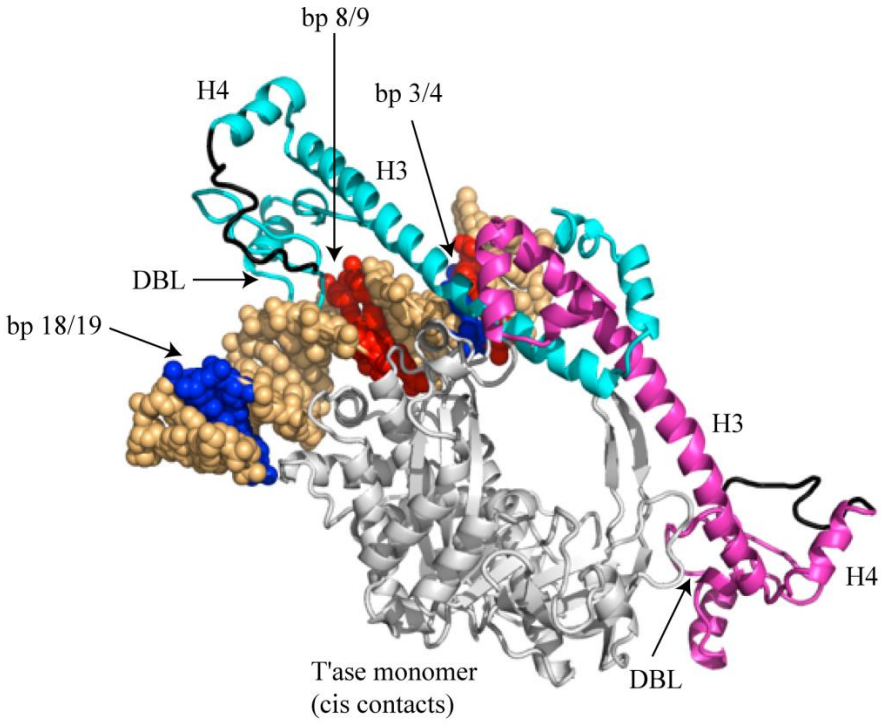


Figure 4.10. Model for H-NS binding to a 'single-end' complex.

A head-to-head dimer of H-NS is shown bound to a transposon end-transposase monomer complex. The transposase monomer is making *cis* contacts with the transposon end. This diagram was generated by simply removing one transposase monomer and one transposon end from the model in Figure 4.8A. Labeling is as in Figure 4.8.

unstable. We suggest that binding of H-NS to a single-end complex could increase the stability of the complex by helping to tether a transposase molecule to a single end through making interactions both with the same DNA molecule and transposase. This is modeled in Figure 4.10 where we have simply removed one transposase monomer and one transposon end from the complex in Figure 4.8A. We presume the order of binding on a single end would be transposase first followed by a dimer of H-NS because we don't see a stable interaction of H-NS with transposon end DNA under our standard transpososome assembly conditions in the absence of transposase. Once transposase is stably bound to a single end (through the help of H-NS), we suggest the stably bound end could capture a second transposon end to which transposase is only weakly bound. This would involve both the formation of *trans* contacts between transposase and end sequences as well as engagement of a free H-NS DNA-binding domain with the second end. One caveat to this model is that addition of H-NS to a transpososome assembly reaction has not yielded a single end-transposase-H-NS complex in our gel shift assays. However, this could be explained if the rate of capturing a second end is fast relative to the time scale of electrophoresis.

4.4.5 Biological significance of the H-NS-transpososome interaction

Transpososome assembly is a critical early step in most transposition reactions. Thus, modulation of this step by host proteins represents an attractive means of fine-tuning transposition frequencies to fit both the needs of the transposon and the host. While H-NS protein levels remain relatively constant throughout the different stages of bacterial growth the oligomeric state of H-NS is sensitive to various environmental conditions

Table 4.1. Oligonucleotides used in this study[a].

Name	Sequence (5' to 3')
CW1	CGCGTTTAATCTGCAGCACAGTCGTGATGGC
CW2	CCCTGCGCAGCGCGCAGCTGCAGCCTGAATACGCG
CW3	CTCGA**C**TGTCTCTTATACACATCTAGCGTCCTGAACGGAACCTTCTGCA
CW4	GAAGGTTCCGTTCAGGACGCTAGATGTGTATAAGAGACA**G**TCGAGGTAC
CW5	CTCGA**C**TGACTCTTATACACAAGTAGCGTCCTGAACGGAACCTTCTGCA
CW6	GAAGGTTCCGTTCAGGACGCTACTTGTGTATAAGAGTCA**G**TCGAGGTAC
CW7	CTCGA**C**TCTCTCTTATACACATCTAGCGTCCTGAACGGAACCTTCTGCA
CW8	GAAGGTTCCGTTCAGGACGCTAGATGTGTATAAGAGAGA**G**TCGAGGTAC
CW9	CTCGA**C**TGTCTCGCATACACATCTAGCGTCCTGAACGGAACCTTCTGCA
CW10	GAAGGTTCCGTTCAGGACGCTAGATGTGTATGCGAGACA**G**TCGAGGTAC
CW11	NNNAAGCTTGGGTAACGCCAGGGTTTTCCCACTC
CW12	NNNTCTAGACGCCAAGCTTGCATGCCTGCAGGTC
WT ME 53NTS	CCCTGCAGGTCGA**C**TGTCTCTTATACACATCTTGAGTGAGTGAGCATGCA
WT ME 53TS	ACATGCATGCTCACTCACTCAAGATGTGTATAAGAGACA**G**TCGACCTGCAGGG
OE 53NTS	CCCTGCAGGTCGA**C**TGACTCTTATACACAAGTTGAGTGAGTGAGCATGCA
OE 53TS	ACATGCATGCTCACTCACTCAACTTGTGTATAAGAGTCA**G**TCGACCTGCAGGG
ME 3 53NTS	CCCTGCAGGTCGA**C**TCTCTCTTATACACATCTTGAGTGAGTGAGCATGCA
ME 3 53TS	ACATGCATGCTCACTCACTCAAGATGTGTATAAGAGACA**G**TCGACCTGCAGGG
ME 8/9 53NTS	CCCTGCAGGTCGA**C**TGTCTCGCATACACATCTTGAGTGAGTGAGCATGCA
ME 8/9 53TS	ACATGCATGCTCACTCACTCAAGATGTGTATGCGAGACA**G**TCGACCTGCAGGG
T1TaseF	GACCTCTTAAGATGGTAACGTTCATG
T1TaseR	GCCGAAGAGAACACAGATTTAGC
T1TaseProbe[b]	6FAM-TAACTTCTGCTCTTCATCGTG-MGBNFQ
16SF	ACCAGGGCTACACACGTGCTA
16SR	TCTCGCGAGGTCGCTTCT
16SProbe[b]	6FAM- AATGGCGCATACAAA MGBNFQ

a. The first nucleotide of the transposon end sequence (1 or +1 position) of each strand is in bold. The mutated nucleotides are underlined.

b. 6FAM represents 6-carboxyfluorescein. MGBNFQ represents the quencher dye.

(15, 46). Based on our model for H-NS binding to the Tn5 transpososome where we suggest only a dimer of H-NS is accommodated into the transpososome, we predict that conditions that disfavor higher order oligomerization of H-NS would increase the potential of H-NS to regulate transpososome assembly in Tn5 transposition. The availability of dimeric H-NS might also be important in limiting the binding of DnaA protein to Tn5 end sequences at early stages of transposition when DnaA binding would be disruptive to transposition.

4.5 References

1. Curcio,M.J. and Derbyshire,K.M. (2003) The outs and ins of transposition: from mu to kangaroo. *Nature reviews. Molecular cell biology*, **4**, 865-77.
2. Nagy,Z. and Chandler,M. (2004) Regulation of transposition in bacteria. *Research in microbiology*, **155**, 387-98.
3. Roberts,D., Hoopes,B.C., McClure,W.R. and Kleckner,N. (1985) IS10 transposition is regulated by DNA adenine methylation. *Cell*, **43**, 117-30.
4. Yin,J.C., Krebs,M.P. and Reznikoff,W.S. (1988) Effect of dam methylation on Tn5 transposition. *Journal of molecular biology*, **199**, 35-45.
5. Signon,L. and Kleckner,N. (1995) Negative and positive regulation of Tn10/IS10-promoted recombination by IHF: two distinguishable processes inhibit transposition off of multicopy plasmid replicons and activate chromosomal events that favor evolution of new transposons. *Genes & development*, **9**, 1123-36.
6. Jain,C. and Kleckner,N. (1993) IS10 mRNA stability and steady state levels in Escherichia coli: indirect effects of translation and role of rne function. *Molecular microbiology*, **9**, 233-47.
7. Ross,J.A., Wardle,S.J. and Haniford,D.B. (2010) Tn10/IS10 transposition is downregulated at the level of transposase expression by the RNA-binding protein Hfq. *Mol Micro*, **78**, 607-21.
8. Wardle,S.J., O'Carroll,M., Derbyshire,K.M. and Haniford,D.B. (2005) The global regulator H-NS acts directly on the transpososome to promote Tn10 transposition. *Genes & development*, **19**, 2224-35.
9. Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2009) The global bacterial regulator H-NS promotes transpososome formation and transposition in the Tn5 system. *Nucleic acids research*, **37**, 309-21.
10. Dorman,C.J. (2004) H-NS: a universal regulator for a dynamic genome. *Nature reviews. Microbiology*, **2**, 391-400.
11. Bouffartigues,E., Buckle,M., Badaut,C., Travers,A. and Rimsky,S. (2007) H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nature structural & molecular biology*, **14**, 441-8.
12. Arold,S.T., Leonard,P.G., Parkinson,G.N. and Ladbury,J.E. (2010) H-NS forms a superhelical protein scaffold for DNA condensation. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 15728-32.
13. Navarre,W.W., McClelland,M., Libby,S.J. and Fang,F.C. (2007) Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes & development*, **21**, 1456-71.
14. Hommais,F., Krin,E., Laurent-Winter,C., Soutourina,O., Malpertuy,A., Caer,J.P. Le, Danchin,A. and Bertin,P. (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Molecular microbiology*, **40**, 20-36.

15. Ono,S., Goldberg,M.D., Olsson,T., Esposito,D., Hinton,J.C.D. and Ladbury,J.E. (2005) H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. *The Biochemical journal*, **391**, 203-13.
16. Davies,D.R., Goryshin,I.Y., Reznikoff,W.S. and Rayment,I. (2000) Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science*, **289**, 77-85.
17. Benjamin,H.W. and Kleckner,N. (1992) Excision of Tn10 from the donor site during transposition occurs by flush double-strand cleavages at the transposon termini. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 4648-52.
18. Mead,D.A., Szczesna-Skorupa,E. and Kemper,B. Single-stranded DNA "blue" T7 promoter plasmids: a versatile tandem promoter system for cloning and protein engineering. *Protein engineering*, **1**, 67-74.
19. Naumann,T.A. and Reznikoff,W.S. (2000) Trans catalysis in Tn5 transposition. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 8944-9.
20. Peterson,G. and Reznikoff,W.S. (2003) Tn5 transposase active site mutations suggest position of donor backbone DNA in synaptic complex. *The Journal of biological chemistry*, **278**, 1904-9.
21. Zhang,A., Rimsky,S., Reaban,M.E., Buc,H. and Belfort,M. (1996) Escherichia coli protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *The EMBO journal*, **15**, 1340-9.
22. Wardle,S.J., Chan,A. and Haniford,D.B. (2009) H-NS binds with high affinity to the Tn10 transpososome and promotes transpososome stabilization. *Nucleic acids research*, **37**, 6148-60.
23. Moreira,R.N., Dressaire,C., Domingues,S. and Arraiano,C.M. (2011) A new target for an old regulator: H-NS represses transcription of bolA morphogene by direct binding to both promoters. *Biochemical and biophysical research communications*, **411**, 50-5.
24. Pfaffl,M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research*, **29**, e45.
25. Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2006) Formation, characterization and partial purification of a Tn5 strand transfer complex. *Journal of molecular biology*, **364**, 290-301.
26. Junop,M.S. and Haniford,D.B. (1997) Factors responsible for target site selection in Tn10 transposition: a role for the DDE motif in target DNA capture. *The EMBO journal*, **16**, 2646-55.
27. Gordon,B.R.G., Li,Y., Cote,A., Weirauch,M.T., Ding,P., Hughes,T.R., Navarre,W.W., Xia,B. and Liu,J. (2011) Structural basis for recognition of AT-rich

DNA by unrelated xenogeneic silencing proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 10690-5.
28. Guex,N. and Peitsch,M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714-23, 10.1002/elps.1150181505.
29. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography*, **60**, 2126-32.
30. Zhou,M., Bhasin,A. and Reznikoff,W.S. (1998) Molecular genetic analysis of transposase-end DNA sequence recognition: cooperativity of three adjacent base-pairs in specific interaction with a mutant Tn5 transposase. *Journal of molecular biology*, **276**, 913-25.
31. Phadnis,S.H. and Berg,D.E. (1987) Identification of base pairs in the outside end of insertion sequence IS50 that are needed for IS50 and Tn5 transposition. *Proceedings of the National Academy of Sciences of the United States of America*, **84**, 9118-22.
32. Wiegand,T.W. and Reznikoff,W.S. (1992) Characterization of two hypertransposing Tn5 mutants. *Journal of bacteriology*, **174**, 1229-39.
33. Ali,S.S., Beckett,E., Bae,S.J. and Navarre,W.W. (2011) The 5.5 Protein of Phage T7 Inhibits H-NS through Interactions with the Central Oligomerization Domain. *Journal of bacteriology*, **193**, 4881-92.
34. Donato,G.M. and Kawula,T.H. (1998) Enhanced binding of altered H-NS protein to flagellar rotor protein FliG causes increased flagellar rotational speed and hypermotility in Escherichia coli. *The Journal of biological chemistry*, **273**, 24030-6.
35. Badaut,C., Williams,R., Arluison,V., Bouffartigues,E., Robert,B., Buc,H. and Rimsky,S. (2002) The degree of oligomerization of the H-NS nucleoid structuring protein is related to specific binding to DNA. *The Journal of biological chemistry*, **277**, 41657-66.
36. Spurio,R., Falconi,M., Brandi,A., Pon,C.L. and Gualerzi,C.O. (1997) The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *The EMBO journal*, **16**, 1795-805.
37. Gradman,R.J., Ptacin,J.L., Bhasin,A., Reznikoff,W.S. and Goryshin,I.Y. (2008) A bifunctional DNA binding region in Tn5 transposase. *Molecular microbiology*, **67**, 528-40.
38. Haniford,D.B., Benjamin,H.W. and Kleckner,N. (1991) Kinetic and structural analysis of a cleaved donor intermediate and a strand transfer intermediate in Tn10 transposition. *Cell*, **64**, 171-9.
39. Surette,M.G., Buch,S.J. and Chaconas,G. (1987) Transpososomes: stable protein-DNA complexes involved in the in vitro transposition of bacteriophage Mu DNA. *Cell*, **49**, 253-62.

40. Steiniger-White,M. and Reznikoff,W.S. (2000) The C-terminal alpha helix of Tn5 transposase is required for synaptic complex formation. *The Journal of biological chemistry*, **275**, 23127-33.
41. Weinreich,M.D. and Reznikoff,W.S. (1992) Fis plays a role in Tn5 and IS50 transposition. *Journal of bacteriology*, **174**, 4530-7.
42. Yin,J.C. and Reznikoff,W.S. (1987) dnaA, an essential host gene, and Tn5 transposition. *Journal of bacteriology*, **169**, 4637-45.
43. Reznikoff,W. (2002) Mobile DNA II Nancy L. Craig, Robert Craigie, Martin Gellert,A.M.L. (ed) ASM Press, Washington, D.C.
44. Schaper,S. and Messer,W. (1995) Interaction of the initiator protein DnaA of Escherichia coli with its DNA target. *The Journal of biological chemistry*, **270**, 17622-6.
45. Rohs,R., West,S.M., Sosinsky,A., Liu,P., Mann,R.S. and Honig,B. (2009) The role of DNA shape in protein-DNA recognition. *Nature*, **461**, 1248-53.
46. Stella,S., Falconi,M., Lammi,M., Gualerzi,C.O. and Pon,C.L. (2006) Environmental control of the in vivo oligomerization of nucleoid protein H-NS. *Journal of molecular biology*, **355**, 169-74.

CHAPTER 5

5 THE RNA-BINDING PROTEIN HFQ NEGATIVELY REGULATES TN5 TRANSPOSITION BY DOWN-REGULATING TRANSPOSASE EXPRESSION

5.1 Introduction

The spread of antibiotic resistance genes in bacteria is mediated largely by transposons (1, 2). Transposons are mobile genetic elements and their movement can be regulated by host proteins (3). It is therefore important to understand what host factors regulate particular transposons and how this regulation is achieved. Many host factors that modulate transposition also play important cellular functions, such as in chromosomal maintenance, supercoiling and DNA replication. However, in many transposition systems we know very little about the host factors that are key regulators. Interestingly, several host factors involved in stress response pathways (e.g. H-NS and IHF) have been proposed to link environmental stimuli to changes in transposition rates (4-8). The host protein Hfq, which is widely implicated in responding to bacterial stress, has recently been identified as a potent negative regulator of Tn10/IS10 transposition (9, 10). This is the first transposition system in which Hfq has been determined to be a regulator and this observation raises the question as to whether other bacterial transposition systems might likewise be regulated by Hfq.

After its initial discovery as a host protein that aids in phage Qβ replication, evidence suggested that Hfq was involved in critical biological processes in many bacterial and archaeal species, for instance Hfq was shown to regulate most virulence genes in *Salmonella enterica* (11, 12). It is now considered a major bacterial

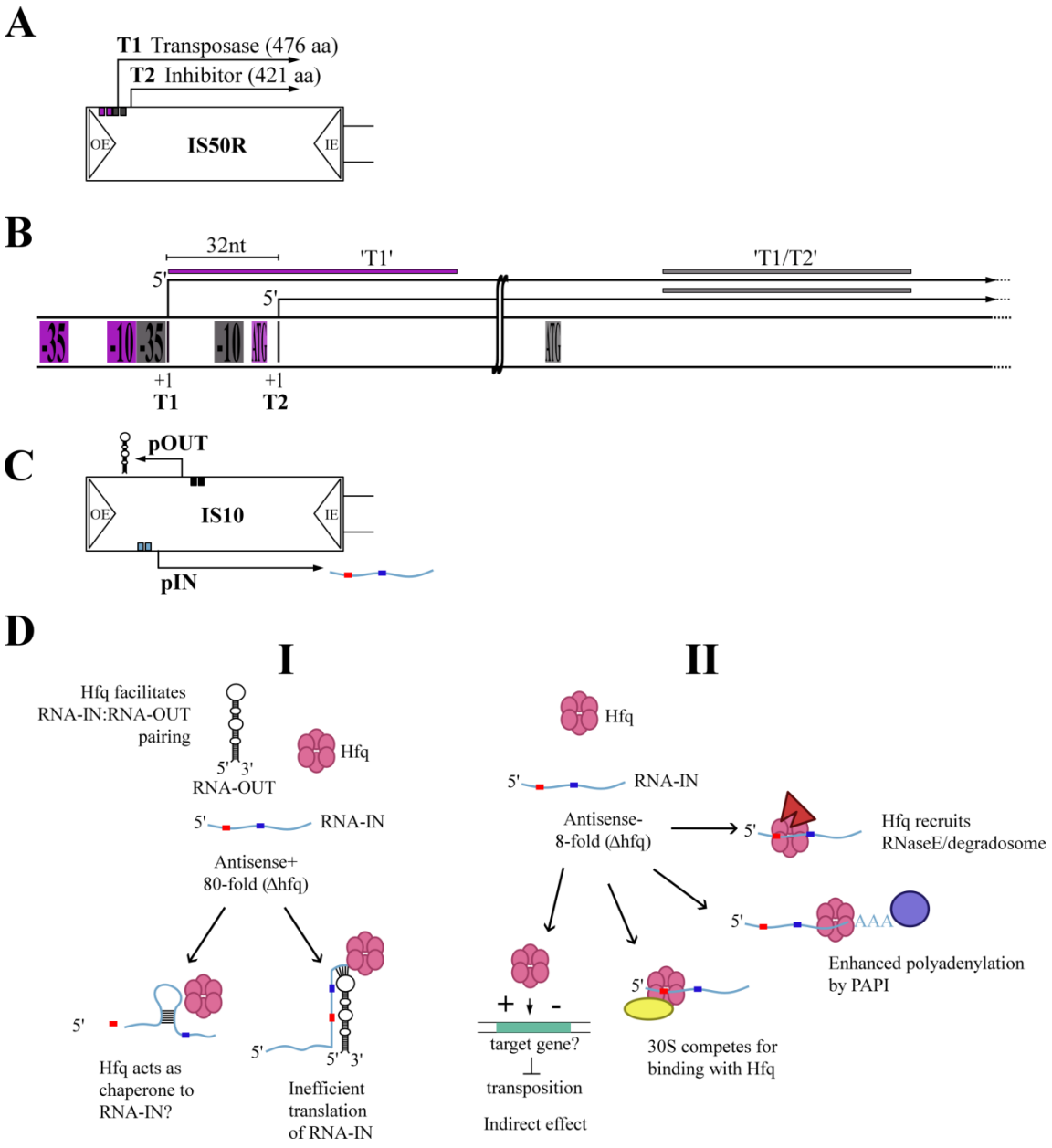
riboregulator that largely regulates gene expression at the post-transcriptional level by mediating the interaction between sRNAs and mRNA targets (e.g. DsrA-*rpoS* (13), OxyS-*rpoS* (14), RprA-*rpoS* (15), and RyhB-*sodB* (16) (17, 18). Hfq also affects gene expression by additional mechanisms that involve its binding to RNA species and/or proteins such as the degradosome and polyadenylate polymerase I (PAPI) (19-22). These activities typically result in differential gene regulation in response to particular stresses.

With respect to transposition, the presence of Hfq was determined to aid the pairing of a cis-endcoded antisense RNA (RNA-OUT) with partial complementarity to the transcript of the Tn10 transposase (RNA-IN) *in vitro* (10) (Figure 5.1C). This pairing is known to decrease transposase expression by inhibiting the translation efficiency of RNA-IN. The mechanism by which Hfq does this has not been established however, a decrease in the available transposase pool causes a decrease in the overall level of transposition. Interestingly even in the absence of the antisense RNA, an hfq disruption increases transposase expression and transposition, although the magnitude of this increase was smaller.

In the current work we ask if another transposon is regulated by Hfq in order to begin to assess if Hfq is a general regulator of transposition reactions. We chose to investigate Tn5 because this transposon has been well-characterized in terms of transposase expression and its mechanism of transposition. Moreover, there is no existing evidence that an antisense RNA plays a role in regulating Tn5 transposition, therefore determining the impact of Hfq on Tn5 may provide evidence for a mechanism of Hfq-dependent regulation of Tn10 that does not rely on an RNA regulator.

Figure 5.1. Transcriptional units of the Tn5 transposon and proposed regulation of Tn10 transposition by Hfq.

(A) Structure of IS50Right. Transcriptional units are indicated with black arrows, and rectangles in purple and grey represent the transcriptional signals for the promoters T1 and T2 respectively. The outside end (OE) and inside end (IE) sequences are indicated with inverted triangles. (B) Linear representation and relative location of Tn5 regulatory elements. Note that there are two promoters located in IS50R. Transcriptional sequences for T1 and T2 are indicated in purple and grey boxes respectively, and the start sites for transcription are indicated with a black vertical lines and '+1'. T1 contains an additional 32 nt at the 5' end, but is otherwise identical to T2. Transposase is translated from T1; an inhibitor protein is translated from T2. Start codons of transposase and inhibitor are represented by light purple and light grey boxes respectively. Long thin rectangles represent positions of the TaqMan primers and probes used for real-time RT-PCR detection of the T1 transcript and the T1/T2 transcripts, respectively. (C) Schematic of the promoters and structure of IS10 labeled as (A). Black and cyan boxes represent promoters for pOUT (antisense RNA) and (pIN) (the transposase transcript) respectively. (D) Proposed regulation of IS10 transposition by Hfq. Several possible mechanisms of Hfq action either in the presence (I) or absence (II) of a *cis*-encoded antisense RNA are depicted. The Hfq hexamer is represented by mauve ovals, and the transcripts of IS10, RNA-OUT (antisense RNA) and RNA-IN (transposase transcript), are depicted in black and cyan respectively. The Shine-Dalgarno site and translational start codon of RNA-IN are represented by small red and blue boxes. RnaseE, polyadenylate polymerase I and 30S ribosome are represented by a dark red shape, a blue circle and a yellow oval respectively.


A
T1 Transposase (476 aa)
T2 Inhibitor (421 aa)
OE
IS50R
IE
B
32nt
'T1'
'T1/T2'
5'
5'
-35
-10
-35
-10
ATG
ATG
+1
T1
+1
T2
C
pOUT
OE
IS10
IE
pIN
D
I
Hfq facilitates
RNA-IN:RNA-OUT
pairing
5'
3'
RNA-OUT
Hfq
5'
RNA-IN
Antisense+
80-fold (Δhfq)
5'
Hfq acts as
chaperone to
RNA-IN?
5'
5'
3'
Inefficient
translation
of RNA-IN
II
Hfq
5'
RNA-IN
Antisense-
8-fold (Δhfq)
5'
Hfq recruits
RNaseE/degradosome
5'
AAA
Enhanced polyadenylation
by PAPI
+
-
target gene?
transposition
Indirect effect
5'
30S competes for
binding with Hfq

5.2.0 Materials and methods

Chemicals and oligonucleotides were from Sigma-Aldrich. Growth media was from Becton Dickinson. Enzymes were from New England Biolabs. Radio-nucleotides were from Amersham Biosciences.

5.2.1 Plasmids

Plasmids used in this work were constructed using standard cloning techniques. The construct used for mating out assays pDH533-4 (Ap^R; Cm^R) was constructed from pRZ9905 (23) and a fragment from an altered pMOD^{TM} plasmid (Epicentre), pDH508, which contains the gene for chloramphenicol resistance from Tn9 (24) inserted into a BamHI restriction site between the mosaic end sequences for Tn5 transposition. The transposon fragment was cut from pDH508 using PshA1 and ligated to a filled-in BglII site in pRZ9905. This produced a final construct consisting of the Tn5 transposase, with an alanine substitution at position 56 that removes the start codon for the inhibitor protein, under control of its native promoter directly adjacent to a mini-Tn5 transposon consisting of mosaic ends marked with a chloramphenicol resistance gene. The MA56 mutation does not alter transposase or transposition rates and is present to remove the start codon for the inhibitor protein (25). Therefore the term wild-type refers to transposase with the MA56 substitution. Sequencing was used to verify the plasmid; orientation of the mini-transposon was determined to position transcription of the chloramphenicol gene in the same direction as the transposase gene.

The *lacZ*-translational fusion constructs are pDH658, pDH659 and pDH679. The *lacZ* fragment used for each construct was PCR-amplified using primers CW1 and CW2

with SacII and BglII restriction sites engineered at the 5' ends respectively, and the template pDH611 (Tn10 transposase-*lacZ* transcriptional fusion) (10). The PCR product was digested with SacII and BglII, gel-purified, and ligated to the large fragment of SacII/BglII-digested pRZ9905 to create pDH658, or pDH623 to create pDH659. pDH623 is essentially pRZ9905 except that it contains *rrnB* termination signals immediately upstream (~50 bp) of the IS50 promoter. The fragment containing the *rrnB* termination signals was amplified using the primers CW3 and CW4 (engineered to have EcoRI restriction sites on the 5' ends) and the template pBAD24 (26). The PCR product was digested with EcoRI and inserted into the EcoRI site of pRZ9905. The orientation of the insertion was verified by sequencing. The *lacZ* fragment was then inserted into pDH623 as described above to create plasmid pDH659. The resulting translational fusions consist of the first 366 bp of the Tn5 transposase coding region followed in-frame by base pairs 28 to 3075 of the *lacZ* coding sequence. pDH679 is essentially the same as pDH658 except that 297 bp of the transposase coding region is deleted (positions 51-347, where position 1 is the transcriptional start) using SapI digestion and self-ligation of the large fragment.

For the complementation mating out assays, strains were co-transformed with pDH708 (Kan^R; Cm^R) and either the empty vector pWKS30 (27) or pDH700 (Ap^R) (Hfq under control of its native promoter). pDH708 is a derivative of pDH533-4 where the ampicillin resistance gene was essentially replaced with the kanamycin resistance gene. pDH533-4 was digested with XmnI and ligated to a blunt-ended fragment containing the kanamycin gene that was produced using PCR and primers from pNK1182 (28). pDH700 was cloned using a PCR fragment containing full-length Hfq and its 5' regulatory

sequences and pWKS30. The *hfq* gene was amplified from genomic DNA of NK5830 (29) with forward (CW5) and reverse (CW6) primers containing 5' XbaI and HindIII restriction sites, respectively. The PCR product was digested with HindIII and XbaI and ligated to the large fragment of HindIII/XbaI-digested pWKS30.

5.2.2 Protein purification

Hfq protein was purified exactly as previously described to 99% purity (10). Note that contaminating RNA was removed with an RNase A treatment.

5.2.3 Mating out assay

Mating out assays used NK5830F' (DBH33) and derivatives (DBH16) (10) as donors and HB101 as recipient. pDH533-4 was used as the source of mini-Tn5 transposon and transposase. However for mating out assays complemented with Hfq from a plasmid, pDH708 was used as the source of transposon and transposase, and either pDH700 or pWKS30 were used as the *trans* source of Hfq or the empty vector respectively. Plasmids were transformed into donor strains and transformants were selected on either M9-glucose plates supplemented with arginine (40 μg ml^{-1}) and thiamine (1 μg ml^{-1}) with ampicillin (50 μg ml^{-1}) and chloramphenicol (20 μg ml^{-1}), or for the complementation assays, kanamycin (50 μg ml^{-1}) and ampicillin (50 μg ml^{-1}). Donors were grown overnight to saturation in 1 ml of M9-glucose media supplemented with arginine and thiamine with appropriate antibiotics. Donors were subcultured 1:10 the following day in a total of 4 ml in M9 supplemented with arginine and thiamine without antibiotic. Subculturing involved an initial 2 hours of growth for *hfq*+ and 4 hours of growth for

Δ*hfq* on a fast roller, followed by an additional 2 hours of growth for *hfq*+ and 4 hours of growth for Δ*hfq* at on a slow roller. The extra time allotted for growth of Δ*hfq* cells ensured that donor strains were at roughly the same cell density prior to mixing with the recipient. The recipient strain was grown overnight to saturation in 5 ml of LB with streptomycin (150 μg ml^{-1}) subcultured 1:10 with LB, and grown to an optical density of 0.6. Recipient was then carefully mixed with donors and cultures were grown on a slow roller for 1 hour. Before addition of recipient, 1-2 ml of each donor sample transformed with pDH533-4 in each strain was removed for RNA isolation. After the mating was complete, cultures were mixed vigorously, 1 ml of mating mix was removed and cells were pelleted and then resuspended in 1 ml of 0.85% saline. Cells were then plated at appropriate dilutions to grow single colonies on M9-glucose plates supplemented with leucine and streptomycin (150 μg ml^{-1}) for determining the number of exconjugants, and leucine, streptomycin and chloramphenicol (20 μg ml^{-1}) for measuring transposition events. The transposition frequency was obtained by dividing the number of Cm^R Sm^R colonies ('hops') by the number of Sm^R colonies (total exconjugants). We routinely monitored the copy number of the transposition plasmids to determine if the *hfq* status of the donor strain influenced plasmid copy number. One ml of each donor sample containing pDH533-4 was removed based on the cell density as determined by optical density; equivalent cell amounts for each sample were used for plasmid DNA isolation (Sigma-Aldrich). Frequently equal volumes of a separate strain (DH5α) transformed with a plasmid of different size compared to pDH533-4 was added in equal volumes to each of the donor samples as a control for efficiency/recovery differences among the columns

used. The plasmids recovered were analyzed on a 1% agarose gel with ethidium bromide staining.

5.2.4 β-galactosidase assays

The same isogenic *hfq*+ and Δ*hfq* strains of *E. coli* used for the mating out assays were transformed with pDH658, pDH659 or pDH679 (described above). Cells were grown as described for the mating out assay entirely in M9 minimal supplemented with glucose, arginine, and thiamine. After reaching mid-log-phase cells were placed on ice for 20 minutes, pelleted and resuspended in liquid M9 medium supplemented with glucose but lacking amino acids and antibiotics. β-galactosidase activity in cell extracts was determined as previously described (30). At least three different transformants of each strain were tested in at least three independent assays and mean values plus standard deviation on the mean were calculated.

5.2.5 Quantitative RT-PCR and analysis

To measure steady-state Tn5 transcript levels in *hfq*+ and Δ*hfq* strains, approximately 1-2 ml of cells was removed from NK5830F' and DBH16 donor samples transformed with pDH533-4 immediately before the remaining 1 ml was used for mating out. The cells were gently centrifuged (4000 x g for 4 minutes), resuspended in 200 µl of RNALater (Ambion) and incubated at 4°C overnight. The following day, RNA was extracted from each sample using the RNeasy Mini-kit (Qiagen). The quality and quantity of the final RNA samples were assessed using 1% agarose gel electrophoresis and a Nanophotometer (IMPLEN) to measure A_{260}, 260/230 and 260/280 ratios. A portion of the RNA was

treated with the RNase-free TURBO DNase Kit (Applied Biosystems) as per instructions for typical amounts of contaminating genomic DNA. The resulting RNA was quantified again as above to ensure 260/230 and 260/280 ratios were in the range of 1.5-2.0. Reverse transcription was performed with a portion of this RNA (1-2 µg) using random primers and reagents in the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Appropriate control reactions were conducted where RNA was omitted, or reverse transcriptase was omitted for each RNA sample. The final cDNA concentrations of 50-100 ng μl^{-1} (depending on the starting amount of RNA) were diluted to 10 ng μl^{-1} for use in real-time PCR reactions. TaqMan primers and probes are listed in Table 5.1 and were designed using the Applied Biosystem Primer Express 2.0 software to the transposase transcript and to the endogenous control 16S rRNA. 16S has been used as an endogenous control for Δ*hfq* strains previously (31, 32). Reactions for real-time PCR were typically done in triplicate for each of the four biological samples per experiment in 20 µl volumes in 384-well clear plates using the protocol in the TaqMan Gene Expression MasterMix Protocol for custom primers and probe (Applied Biosystems). Control real-time PCR reactions containing no cDNA (1x TE replacement) and no TaqMan 'enzyme mix' were also conducted. Reactions were run using standard cycle parameters on an Applied Biosystems 7900-HT Real-Time System (50°C, 2 minutes; 95°C, 10 minutes; 95°C, 15 seconds; 60°C, 60 seconds). The Pfaffl mathematical model of relative quantification was used to determine the relative amounts of transposase mRNA (33):

$$\frac{(E_{transposase})^{\Delta Ct_{transposase}}}{(E_{16S})^{\Delta Ct_{16s}}}$$

E is the PCR amplification efficiencies of the transposase transcript and 16S rRNA transcript that were determined by creating several standard curves with known target concentrations as described by Applied Biosystems. Ct is the cycle threshold or the cycle at which the fluorescence (ΔR_n) detected is statistically significant as computed automatically by the real-time system. The amplification efficiency is determined as:

$$E = 10^{-(\frac{1}{slope})}$$

ΔCt of transposase or 16S represents the sample Ct subtracted from the Ct of the reference condition to which all samples will be quantified relative to (i.e. NK5830F' transformed with pDH533-4 or pDH658).

To measure the half-life of Tn5 transposase transcript, RNA was isolated at different time points in the following rifampicin sensitive strains: DBH116 (wild-type) and DBH117 (Δ*hfq*; Cm^R), which were gifts from H. Aiba (34). Strains were transformed with pDH533-4 and transformants were grown on M9-glucose with ampicillin (50 μg ml^{-1}) and for the case of DBH117 also chloramphenicol (20 μg ml^{-1}). Half-life analysis was conducted with two different transformants from each strain. Single transformants were grown as overnights in 5 ml to saturation in M9-glucose with appropriate antibiotics. Samples were subcultured in 40 ml as in the mating out protocol with M9-glucose and appropriate antibiotics, and grown to OD_{600} of 0.4-0.6. Five ml of cells was removed before addition of rifampicin as a control and 5 ml aliquots were removed at varying time points after addition of rifampicin to 200 μg ml^{-1} and immediately frozen in an ethanol/dry ice mixture and kept on ice to thaw for approximately 2 hours. Rifampicin acts within 2-4 seconds to inhibit transcription initiation (35). RNA isolation was performed from this point on as described previously except that after the final ethanol

precipitation, RNA was resuspended in 50 µl (30). RNA was then assessed for quality and quantity using a Nanophotometer (IMPLEN). A portion of this RNA was then treated with DNase as per normal contaminating DNA using the TURBO DNA-free Kit (Applied Biosystems). The DNase-treated RNA was then assessed for quality and quantity by analyzing a small portion on a 1% agarose gel and with a Nanophotometer. After this point 1-2 µg of RNA was reverse transcribed into cDNA using the same protocol as above for the steady-state analysis of transposase mRNA. TaqMan primers and probes used for half-life analysis are shown in Table 5.1 and represented in Figure 5.1B. Quantitative PCR and analysis was done the same as above except that the '-Rifampicin' condition (or time=0') served as the reference sample for both *hfq*+ and Δ*hfq* strains. Transcript levels were normalized to corresponding 16S rRNA levels. Data were further normalized by assuming that the level of RNA obtained by extrapolation of the predicted curve to t=0 was 100% and all other levels are plotted relative to this; resulting amounts were plotted as a function of time of RNA extraction after rifampicin addition, and non-linear regression was used to fit a one phase decay curve to the data.

5.2.6 *In vitro* transcription and RNA labeling

Linear DNA templates for run-off transcription of Tn5 transposase RNA (nucleotides 1-193) was synthesized by PCR using primers CW7 and CW8 respectively (note that the 'forward' primer includes the T7 consensus promoter and the T7 polymerase adds 2 additional nucleotides to the 5' region of the transposase RNA). The amplicon was gel-purified and used as template for *in vitro* transcription using T7 RNA polymerase (New England Biolabs). For *in vitro* transcription approximately 1 µg of template DNA was

mixed with 0.5 mM of each rNTP, 50 mM DTT, 0.5 ml of RNasin (Promega) and 50 units of T7 RNA polymerase in supplied reaction buffer (New England Biolabs) in a total volume of 20 µl. RNA was ^{32}P-labeled by including [α-^{32}P]-CTP by in the transcription reaction instead of CTP. The mixture was incubated at 37°C for 2 hours before adding 1 unit of RQ1 RNase free DNase (Promega) and continuing incubation for 0.5 hours. The mixture was then ethanol precipitated and resuspended in 20 µl of Hfq binding buffer; 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 250 mM NH4Cl, 10% glycerol. RNAs were subject to denaturing PAGE and purified from gel slices. Eluates containing RNA were run through a Biospin 6 column (Bio-Rad) and then concentrated by ethanol precipitation. RNA was resuspended in Hfq binding buffer and the concentration was determined using a Nanophotometer (IMPLEN).

5.2.7 Hfq binding assays

^{32}P-labeled transposase RNA was incubated at 95°C for 2 minutes and then placed on ice for 1 minute ('pre-treated') before being mixed with Hfq protein in Hfq binding buffer as previously described (10). The final concentration of RNA in a 10 µl binding reaction was 0.4 nM. Binding reactions were incubated at 25°C for 15 minutes before being mixed with load dye (20 mM Tris-HCl (pH 7.5), 10 mM DTT, 100 mM KCl, 30% glycerol, 0.05% w/v bromophenol blue) and loaded onto a 6% native polyacrylamide gel. Electrophoresis took place in 1x TAE buffer and was carried out at 200 V for 1.5 hours. Following electrophoresis the gel was analyzed as described previously (36). The competitive binding assays were conducted similarly, except that Hfq was incubated initially with varying amounts of unlabeled DsrA or poly(A)$_{18}$ ('pre-treated' with heat

followed by ice, similarly to above) in 10 µl volumes for 15 minutes at 25°C, followed by addition of equal volumes of 'pre-treated' ^{32}P-labeled transposase RNA. Reactions were done in the same Hfq binding buffer and analyzed by gel electrophoresis as above.

5.3.0 Results

5.3.1 Hfq negatively regulates Tn5 transposition

We asked whether another transposon that is not expected to be regulated by an antisense RNA is regulated by Hfq. To this end we measured the frequency of Tn5 transposition in isogenic *hfq*+ (DBH33) and Δ*hfq* (DBH16) *Escherichia coli* strains using a mating out assay (see 'Materials and methods'). Briefly, we transformed a plasmid (pDH533-4) encoding a mini-Tn5 element and a source of transposase into the *hfq*+/Δ*hfq* strains. Transformants were selected on minimal media plates (M9) supplemented with glucose and ampicillin for plasmid selection. Subsequently, transformants were grown in liquid media (LB) and mixed with a recipient strain (HB101). The results from this experiment are shown in Figure 5.2 and reveal that Tn5 transposition increases 5-fold in the Δ*hfq* strain relative to *hfq*+. The magnitude of this effect is comparable to that previously observed in the Tn10 system under conditions where the antisense RNA is not expressed. When we repeated the above experiment but conducted each growth step of the donor strains in minimal media we observed an even stronger transposition-up phenotype in the Δ*hfq* strain (~50-fold).

We considered the possibility that altering the growth conditions for the donor strains might influence plasmid copy number in the isogenic strains. For example, an

increased plasmid copy number in Δ*hfq* versus *hfq*+ could increase the transposition frequency due to presumably more transposase produced and more copies of the Tn5 derivative. To compare the relative copy numbers of the plasmid in the different strains under the different growth conditions we removed an equal density of cells from donor cultures used for mating out experiments, and mixed these cells with an equal amount of cells from a control strain (DH5α) transformed with a control plasmid of different size. Subsequently, we extracted total plasmid DNA from these mixtures and analyzed this material on an agarose gel (see 'Materials and methods'). The results indicate a small increase in the copy number (~1.5-fold) of the Tn5 plasmid in the Δ*hfq* strain under both growth conditions (Figure 5.3). It is unlikely that these small increases can account for the 5- to 50-fold increase in Tn5 transposition reported above if we assume that a transposase increase is responsible for the increased transposition.

The Δ*hfq* strain used contains a disruption of the *hfq* gene. In order to rule out the possibility that changes in transposition rates reported above were due to downstream polar effects of the cassette insertion, we performed a 'complementation' experiment wherein we introduced a low copy plasmid encoding the *hfq* gene (pDH700) into the Δ*hfq* strain and measured the Tn5 transposition frequency. To maintain both the '*hfq*' and the Tn5 plasmids in the donor strains with antibiotic selection it was necessary to create a modified version of pDH533-4 that is ampicillin sensitive (pDH708). We co-transformed pDH708 and pDH700 (or the vector only control pDH337) into the isogenic *hfq*+ and Δ*hfq* strains and performed all of the growth for the respective donor strains in minimal media supplemented with glucose and the appropriate antibiotics. The results from this mating out experiment (Figure 5.2) show that transposition was significantly reduced in

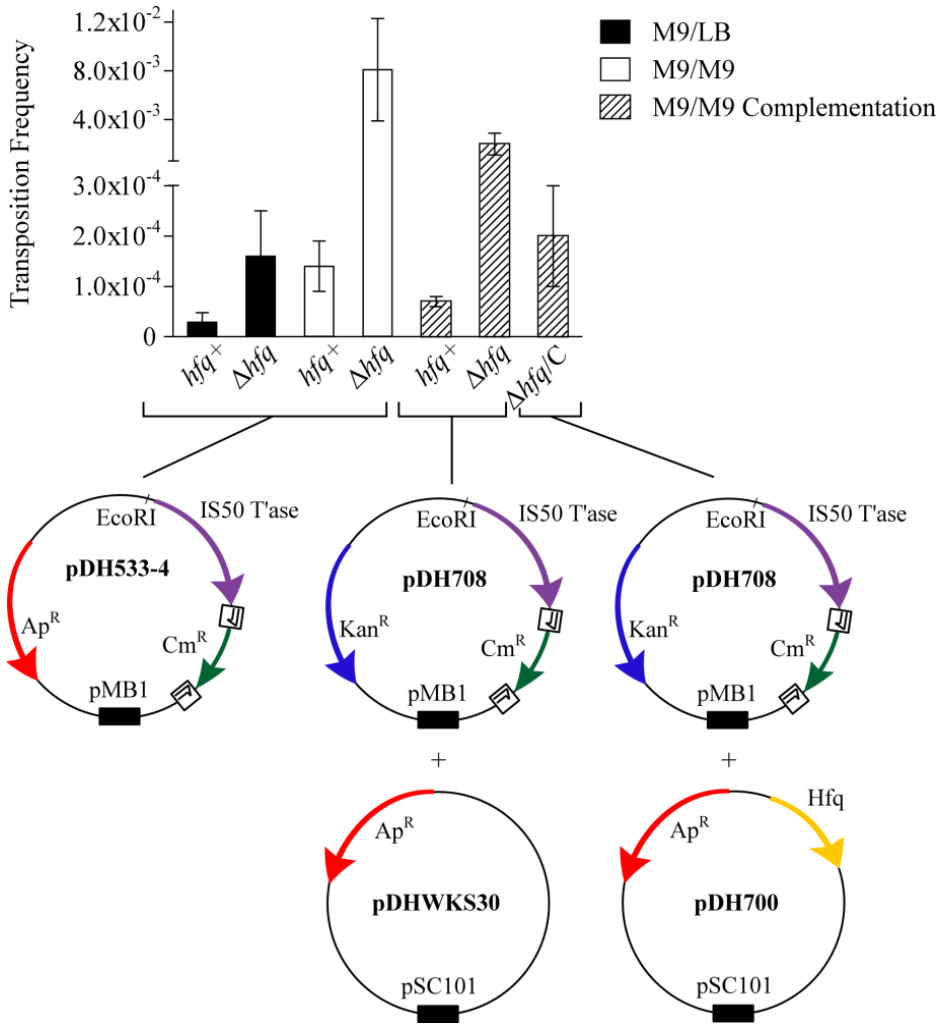


Figure 5.2. Transposition frequencies in isogenic *hfq*+ and Δ*hfq E. coli* strains.

Strains (indicated on the x-axis) were transformed with the plasmids depicted below the bar graph as indicated and mating out assays were used to measure transposition as described in the 'Materials and methods'. Frequencies represent the mean of at least three independent assays where four individual transformants were used. 'Δ*hfq*/C' represents the Δ*hfq* strain transformed with transposition plasmid pDH708 and a complementation vector containing the *hfq* gene under control of its native promoter (pDH700). The legend next to the bar graph indicates the type of agar growth followed by the type of liquid growth during the mating out assay. Note that the M9 minimal media used contained 0.1% glucose. Coloured arrows indicate transcriptional units and are labeled appropriately. Black boxes represent the origin of replication. Half-arrows in white boxes represent the Tn5 mosaic end sequences.

the presence of Hfq (provided by the low copy plasmid) compared to when Hfq was absent (vector only control).

Taken together the results in this section show that Tn5 transposition is negatively regulated by Hfq; this led us to probe the mechanism by which Hfq exerts this regulation.

5.3.2 Tn5 transposase expression from a translational fusion is increased in an Δ*hfq* strain

Hfq can influence bacterial gene expression by regulating translation and/or the stability of mRNAs. Accordingly, we constructed a series of Tn5 transposase-*lacZ* translational fusions to measure transposase gene expression in isogenic *hfq*+ and Δ*hfq* strains. In the standard fusion the transposase gene up to the 123rd codon was fused to the 11th codon of the *lacZ* gene and this fusion was cloned into a plasmid (pDH658). After transforming pDH658 into *hfq*+ (DBH33) and Δ*hfq* (DBH16) and growing the transformants on M9 media supplemented with glucose, β-galactosidase assays were conducted. As depicted in Figure 5.4 β-galactosidase levels were approximately 17-fold higher in the Δ*hfq* compared to *hfq*+, indicating that Hfq strongly down-regulates Tn5 transposase expression.

To be sure that transposase expression in the above experiment was not directed by transcriptional read-through we introduced a strong transcriptional terminator in front of the Tn5 promoter in our translational fusion construct and performed β-galactosidase assays. Prior to performing these assays we confirmed by semi-quantitative RT-PCR that the transcriptional terminator prevented read-through transcripts (data not shown). The results of the β-galactosidase assay show that we observed similar increases in

transposase expression with this translational fusion as with the translational fusion lacking the upstream terminator indicating that the transposase promoter is controlling expression of the transposase gene (Figure 5.4A).

Repression of Tn5 transposase expression by Hfq could be direct or indirect. In the case of the former possibility we would expect Hfq to bind the Tn5 transposase mRNA. Accordingly, we looked for likely Hfq binding sites within the region of the transposase transcript included in the translational fusion (the first 393 nt of the transcript). Typically Hfq binds A/U/AU-rich nucleotides following or preceding a short hairpin. We identified only one particular region of the 5' end that matched this description located at nucleotides 144 to 180 by using an RNA folding prediction program (RNAfold, University of Vienna) and by manually examining the sequence. In addition, we searched the transcript for the Hfq binding consensus sequence 5'-AAYAAYAA-3'. We found a good match to this consensus sequence in the 5' region of the transposase gene starting at nucleotide 155; 5'-AATCAATAA-3'. There were few restriction enzyme sites flanking the potential Hfq binding site in pDH533-4. Instead we constructed a large deletion that removed this sequence and maintained the reading frame of the message. Although the mRNA produced from this deletion construct lacks a sizable portion of the transposase transcript (nucleotides 51 to 347) the 5' UTR and the start codon are intact. The results presented in Figure 5.4A show that deletion of this segment of the transposase transcript did not result in dampened expression by Hfq as would be expected if a critical regulatory binding site for Hfq had been removed.

In the above experiments we routinely measured plasmid copy number in *hfq*+ and Δ*hfq* strains (Figure 5.3). As previously reported plasmid copy number increased

Figure 5.3. Plasmid levels in isogenic *hfq*+ and Δ*hfq* strains used for mating out and β-galactosidase assays.

(A) The relative means and standard deviations of plasmid levels determined in *hfq*+ and Δ*hfq* strains are shown for the mating out (pDH533-4) and translational fusion (pDH658/659/679) plasmids. The mean plasmid level determined for the Δ*hfq* strain was plotted relative to the *hfq*+ strain. Equivalent cell densities from samples used in the mating out and β-galactosidase assays conducted in minimal media were used to isolate plasmids (see 'Material and methods'). Equivalent amounts of a different strain transformed with a control plasmid of different size were added to each sample before plasmid isolation to account for differences in plasmid recovery. Resulting plasmids were compared using agarose gel electrophoresis. At least three transformants were analyzed for *hfq*+ and Δ*hfq* strains and at least three independent experiments for each assay were conducted. (B) Representative agarose gels stained with ethidium bromide indicating the relative plasmid levels of pDH533-4 (top panel) and pDH658 (bottom panel) in *hfq*+ and Δ*hfq* strains are also shown. The t-test was used to determine if the difference in means of *hfq*+ compared to Δ*hfq* for either condition were statistically significant. *$p<0.05$

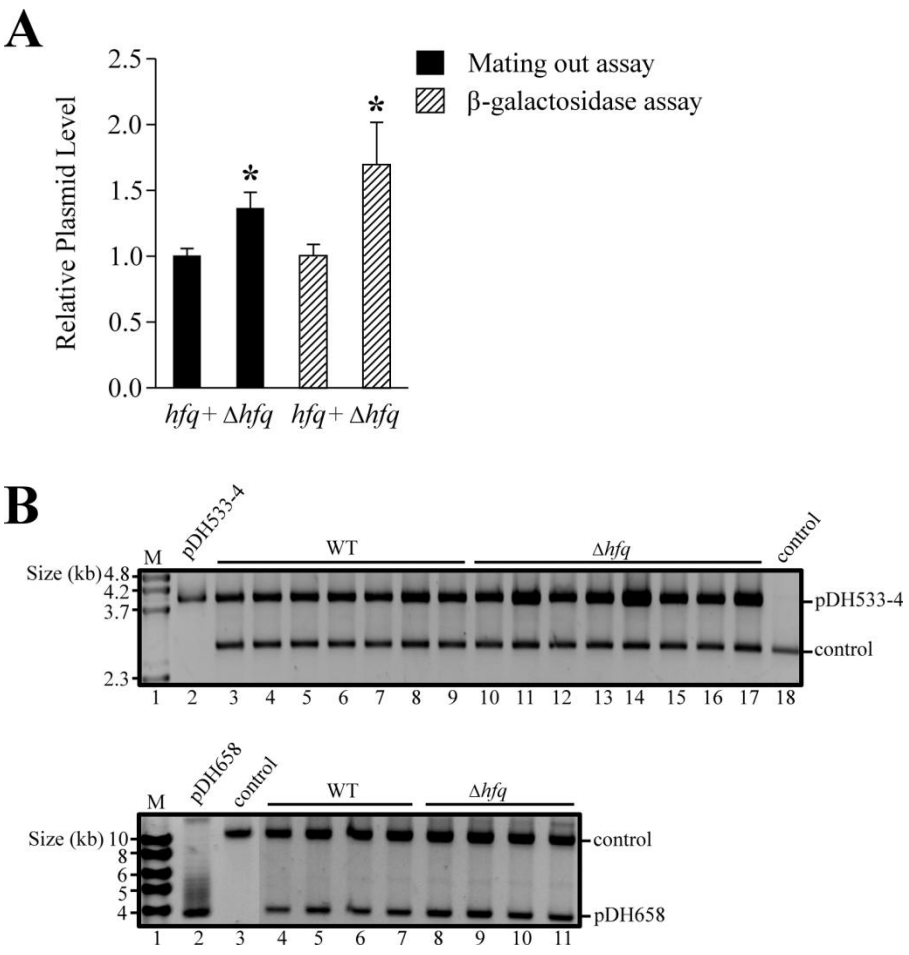
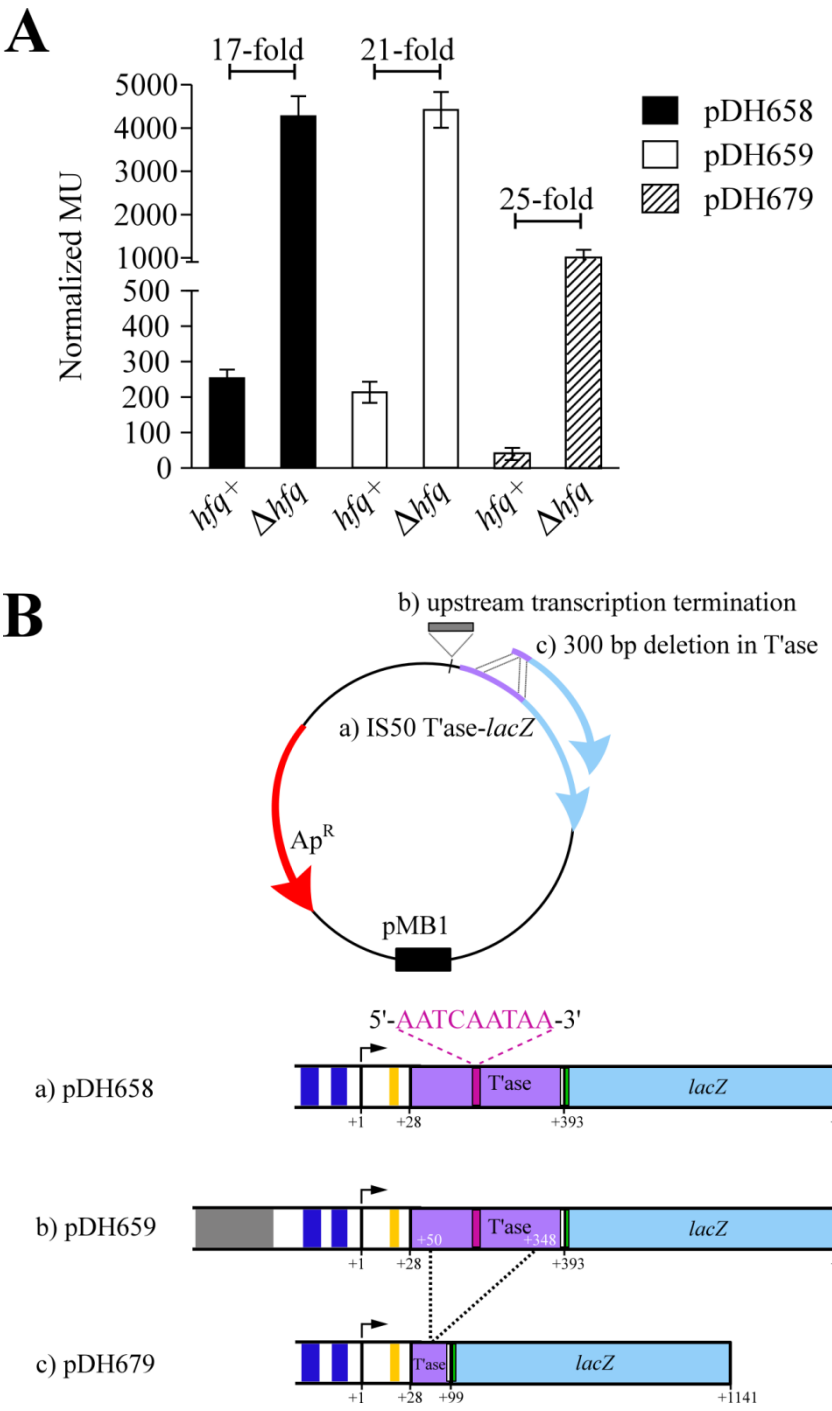

A
Relative Plasmid Level
2.5
2.0
1.5
1.0
0.5
0.0
Mating out assay
β-galactosidase assay
*
*
hfq+ Δhfq
hfq+ Δhfq
B
M
pDH533-4
WT
Δhfq
control
Size (kb) 4.8
4.2
3.7
2.3
pDH533-4
control
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
M
pDH658
control
WT
Δhfq
Size (kb) 10
8
6
5
4
control
pDH658
1 2 3 4 5 6 7 8 9 10 11

Figure 5.4. β-galactosidase expression of transposase-*lacZ* translational fusions in isogenic *hfq*+ and Δ*hfq* strains.

(A) Each translational fusion was transformed into *hfq*+ and Δ*hfq* strains and grown in minimal media with glucose. β-galactosidase assays were performed as in the 'Materials and methods'. Background β-galactosidase activity of each strain (with the empty vector pRZ9905) was subtracted from the raw Miller Units (MU) to give the 'Normalized MU' reported. The fold-increase in Δ*hfq* strain compared to *hfq*+ is given above each set of data. The assays were conducted independently at least twice with 4-5 different transformants. (B) Schematic of the plasmids containing modified Tn5 elements used for β-galactosidase assays. The transposase coding sequence is translationally fused to the coding sequence of the *lacZ* gene. Three translational fusions were used in this work and are depicted both on and below the plasmid substrate. Construct a) produces a transcript consisting of the 123rd codon of transposase (thin white rectangle) followed by the 11th codon of β-galactosidase (thin green rectangle). Construct b) is essentially the same as a) except that it contains a 325 base pair insertion containing a transcription termination signal (grey rectangle) upstream of the transposase promoter (light blue rectangles represent -35 and -10 transcriptional signals). Construct c) is essentially the same as a) except that 297 bp of the transposase coding sequence was removed from positions 51-347 of the transcriptional unit (indicated by white numbering in construct b), which contained a putative binding site for Hfq of 5'-AAYAAYAA-3' (pink rectangle). Numbering below the constructs begins at the transcriptional start site for transposase (black arrow and +1) and continues in sequence to the end of the *lacZ* gene. Purple and cyan rectangles represent the transposase and *lacZ* coding sequences. Light purple and gold rectangles represent the promoter and Shine-Dalgarno sequences respectively. The black box represents the origin of replication (pMB1). Ap^R, ampicillin; Cm^R, chloramphenicol; T'ase, transposase.


A
17-fold
21-fold
25-fold
Normalized MU
5000
4000
3000
2000
1000
500
400
300
200
100
0
pDH658
pDH659
pDH679
hfq+
Δhfq
hfq+
Δhfq
hfq+
Δhfq
B
b) upstream transcription termination
c) 300 bp deletion in T'ase
a) IS50 T'ase-lacZ
ApR
pMB1
5'-AATCAATAA-3'
a) pDH658
T'ase
lacZ
+1
+28
+393
+1441
b) pDH659
+50
T'ase
+348
lacZ
+1
+28
+393
+1441
c) pDH679
T'ase
lacZ
+1
+28
+99
+1141

slightly (~1.5-fold) in the Δ*hfq* strain. This cannot account for the large increases in β-galactosidase activity reported for Δ*hfq* versus *hfq*+ strains in Figure 5.4.

5.3.3 Hfq does not significantly impact the steady state levels or stability of the Tn5 transposase transcript

The derepression of transposase expression observed in the translational fusion experiments (in Δ*hfq*) could be due to Hfq reducing transcript levels and/or translation efficiency in the wild-type strain. We looked at the former possibility by comparing steady state levels of the Tn5 transposase transcript as well as the stability of this transcript in *hfq*+ and Δ*hfq* strains. IS50 contains two promoters that produce overlapping transcripts, one that produces an inhibitor protein (T2) and another that produces transposase (T1) (Figure 5.1). To look at steady-state levels of transposase transcript we performed quantitative RT-PCR on RNA isolated from *hfq*+ and Δ*hfq* strains using a specific primer pair that only amplifies the transcript that produces transposase. Notably, we isolated RNA from the same cultures used for both mating out and translational fusion expression experiments and normalized the transposase transcript levels to the amount of 16S rRNA produced in the strains. The results of this analysis (Figure 5.5) show that the steady-state transposase transcript levels were slightly increased (~1.5-fold) in the Δ*hfq* cells compared to *hfq*+ cells used in the mating out experiment, and that there was essentially no difference in transcript levels in the two cell types used in the translational fusion experiments (Figure 5.5).

We then examined transposase transcript stability in another set of isogenic *hfq*+ (DBH115) and Δ*hfq* (DBH117) strains transformed with the mini-Tn5 plasmid. In order

to measure the half-life of transposase mRNA in *hfq*+ and Δ*hfq* strains it was necessary to use a different set of strains from those used in the mating out assays that were verified to be rifampicin-sensitive. RNA was isolated from the above strains transformed with pDH533-4 at various time points after rifampicin treatment and transposase transcript levels were determined by quantitative RT-PCR. We used two sets of primers for detection of the transcript because we were concerned that in amplifying only one portion of the transposase RNA, our read-out might not be identifying the primary RNA processing event (see Figure 5.1B).

Briefly, RNA half-lives were determined by normalizing the quantity of transposase transcript at each time point to the quantity of 16S rRNA at each time point and fitting the resulting plot using non-linear regression with a one-phase decay curve. In experiments with the T1 probe set half-lives of the transposase transcript in cells grown in M9 media supplemented with glucose were 2.8±0.24 and 2.1±0.22 minutes in *hfq*+ and Δ*hfq* strains, respectively. In experiments with the central probe set these values were 1.9±0.27 and 1.8±0.0028 minutes, respectively (Figure 5.6). These half-lives are not drastically different from those determined previously for Tn5 transcripts (37).

Overall, the results in this section suggest that the *hfq* status of the cell has only a minor effect on transposase mRNA synthesis and metabolism and therefore point to the possibility that Hfq represses transposase expression primarily at the level of translation.

5.3.4 Hfq binds the 5’ region of the Tn5 transposase transcript

To further probe the mechanism by which Hfq represses Tn5 transposase expression we asked if Hfq binds to the 5’ end of the transposase transcript *in vitro*. A relatively strong

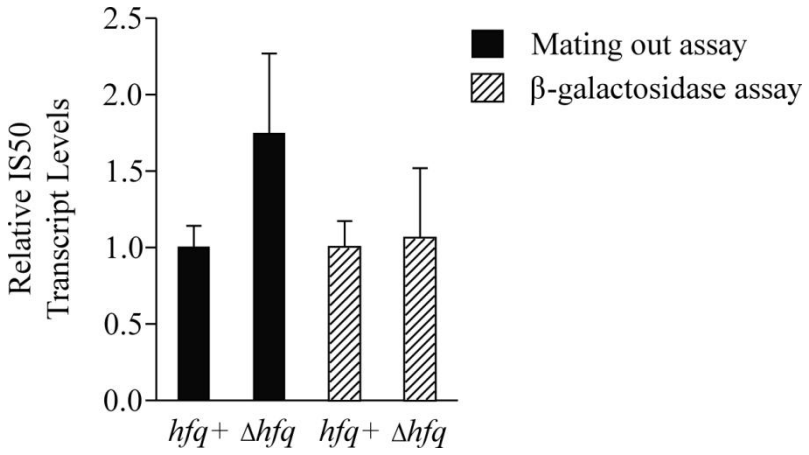


Figure 5.5. Relative transposase transcript levels in isogenic *hfq*+ and Δ*hfq* strains used for mating out and β-galactosidase assays.

Total RNA from donor strains used in mating out and β-galactosidase experiments was isolated at mid-log phase of growth. Equivalent amounts of RNA were reverse transcribed using random primers and transcript levels were quantified using TaqMan primers and probe 'T1' illustrated in Figure 5.1B and listed in Table 5.1 as described in the 'Materials and methods'. The relative levels of transcripts presented in the graphs were normalized to the corresponding levels of 16S rRNA transcript using the Pfaffl method, and the transcript levels are plotted relative to the *hfq*+ strain for each assay. Each normalized value represents an average from at least three technical replicates from three independent clones used in at least two mating out and β-galactosidase experiments. The t-test was used to determine that the difference in means of *hfq*+ and Δ*hfq* in both data sets were not statistically significant.

and specific interaction would provide support for the possibility that Hfq acts directly on the transposase transcript to negatively regulate its expression. To this end we produced via *in vitro* transcription ^{32}P-labeled transposase RNA that includes the first 193 nt of the transcript. Hfq protein was added in varying concentrations to this RNA under conditions where Hfq was in significant excess relative to the RNA. We then performed an electrophoretic mobility shift assay (EMSA) to test for Hfq binding. As shown in Figure 5.7A we observed almost a complete shift in the RNA mobility when Hfq was present at a concentration of 65 nM with two distinct Hfq-bound species being formed (complex 1 and 2). If we presume that complex 1 consists of a hexamer of Hfq bound to a single RNA molecule, the reduced mobility in complex 2 may be due to more than one Hfq hexamer bound to the RNA species. An additional possibility involves the fact that two different conformations of the single transposase RNA species are present when analyzed by native gel electrophoresis analysis (Figure 5.7A). Therefore it is possible that the different complexes result from a single Hfq hexamer binding two different conformations of the same RNA species.

We also performed binding assays in the presence of competitor RNA. In one experiment we used the sRNA DsrA as a competitor. DsrA is known to bind the proximal RNA binding site in Hfq and the affinity of this RNA for Hfq has previously been determined (K_d=21 nM) (38). In this experiment Hfq was added at a concentration that gave exclusively complex 2 in the absence of competitor (lane 4). Inclusion of DsrA at a concentration of 2.5 nM (6.3-fold excess to transposase RNA) reduced the amount of complex 2 formed by ~50%. Under these conditions complex 1 was formed. As the amount of DsrA increased complex 2 was lost and complex 1 remained. At the highest

Figure 5.6. Transposase transcript half-lives in isogenic *hfq*+ and Δ*hfq* strains.

Rifampicin-sensitive isogenic *hfq*+ and Δ*hfq* strains, DBH116 and DBH117, were transformed with pDH533-4 and used for transcript turnover experiments. Strains were grown identically to the mating out and β-galactosidase assays. At mid-log phase after subculturing an aliquot was removed before rifampicin addition (time=0'). After addition of rifampicin, equivalent aliquots were removed at varying time points (indicated in the plots) and treated as in the 'Materials and methods'. (A) 'T1' and (B) 'T1/T2' primers and probes sets depicted in Figure 5.1B and Table 5.1 were used for amplification of the transcript. Transcript levels were normalized to corresponding 16S rRNA levels. Data were further normalized by assuming that the level of RNA obtained by extrapolation of the predicted curve to t=0 was 100% and all other amounts are plotted relative to this; resulting amounts were plotted as a function of time of RNA extraction after rifampicin addition, and non-linear regression was used to fit a one phase decay curve to the data. Half-life analysis was conducted twice with independent transformants (shown as 1 and 2 on the line graphs). Error bars on each data point represent the 3-4 technical replicates that were used for the quantitative PCR analysis at each time point.

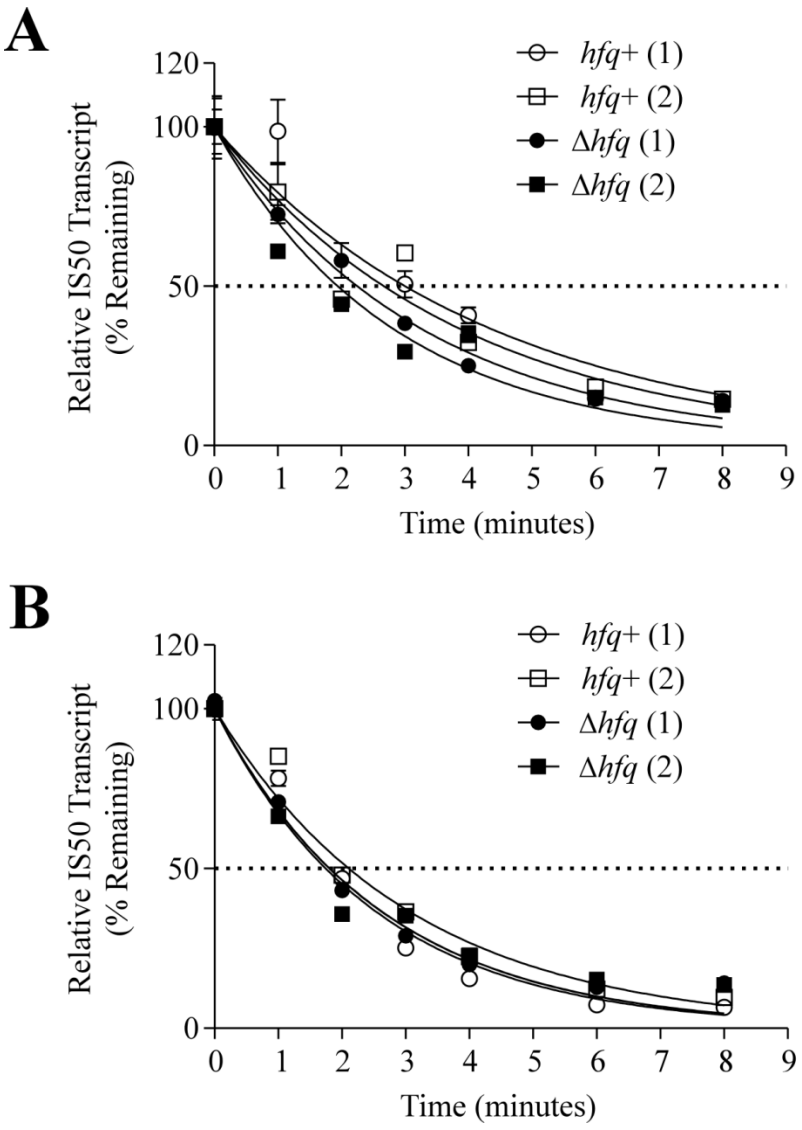
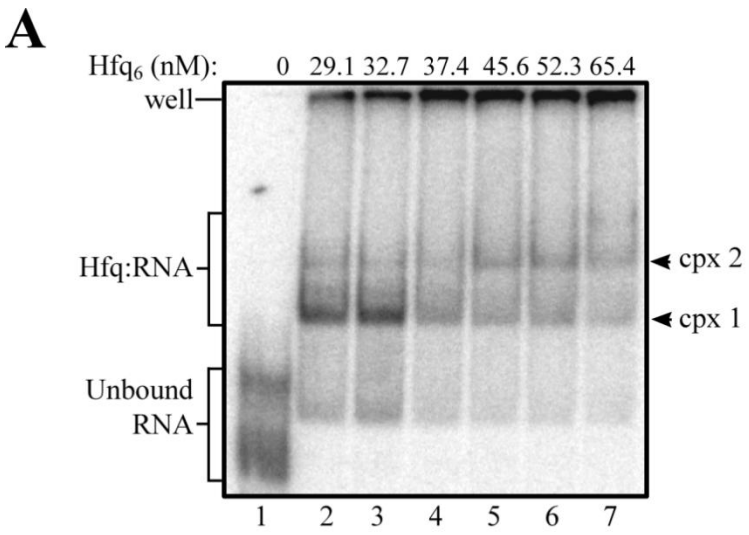
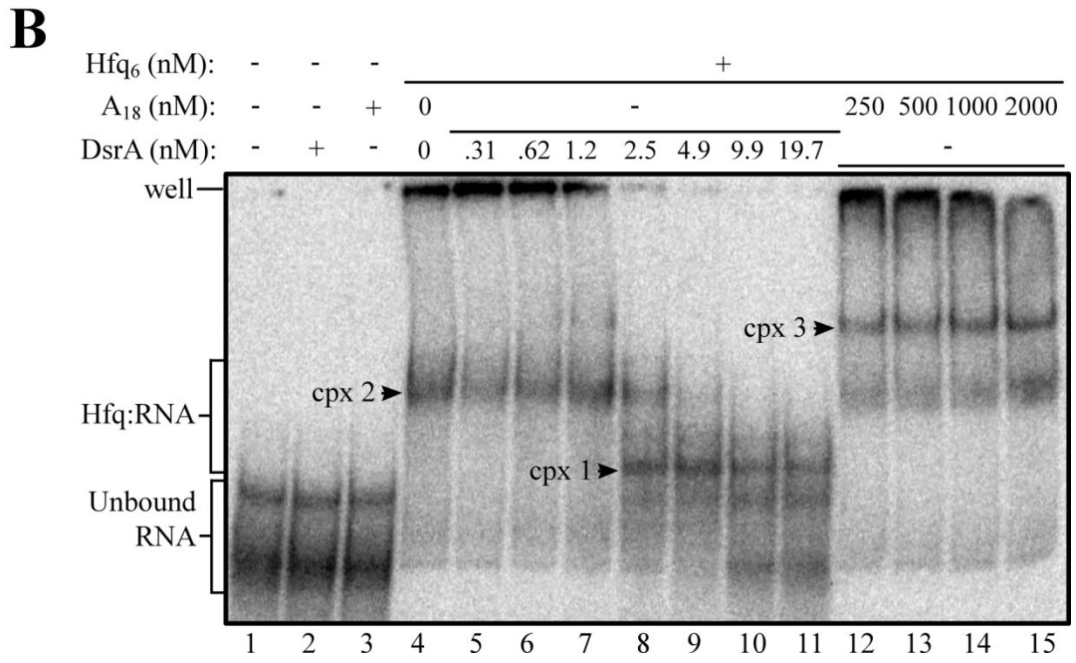

A
Relative IS50 Transcript
(% Remaining)
120
100
50
0
0 1 2 3 4 5 6 7 8 9
Time (minutes)
hfq+ (1)
hfq+ (2)
Δhfq (1)
Δhfq (2)
B
Relative IS50 Transcript
(% Remaining)
120
100
50
0
0 1 2 3 4 5 6 7 8 9
Time (minutes)
hfq+ (1)
hfq+ (2)
Δhfq (1)
Δhfq (2)

Figure 5.7. Hfq binds the transposase transcript *in vitro*.

(A) ^{32}P-labeled RNA (195 nt) from the 5' end of the transposase transcript was mixed with varying amounts of purified Hfq protein as indicated and subjected to EMSA as described in the 'Materials and methods'. Transposase RNA was present at 0.4 nM in each binding reaction. Note that Hfq binding produces two distinct shifted complexes labeled 1 and 2. (B) Hfq binding assays with transposase RNA in the presence of competitor species DsrA or poly(A)$_{18}$. A constant concentration of Hfq (29 nM) was added to unlabeled competitor species at concentrations indicated and incubated in Hfq binding buffer. A constant amount of ^{32}P-labeled transposase RNA (0.4 nM) was then mixed with the reactions, incubated, and subjected to EMSA as in (A). The competitor sRNA DsrA is in excess to transposase RNA over the range of 0.8- to 50-fold, whereas poly(A)$_{18}$ is in excess to transposase RNA over a range of 625- to 5000-fold. Concentrations of Hfq are given per hexamer.


A
Hfq6 (nM): 0 29.1 32.7 37.4 45.6 52.3 65.4
well
Hfq:RNA
Unbound RNA
cpx 2
cpx 1
1 2 3 4 5 6 7


B
Hfq6 (nM): - - - +
A18 (nM): - - + 0 - 250 500 1000 2000
DsrA (nM): - + - 0 .31 .62 1.2 2.5 4.9 9.9 19.7 -
well
cpx 3
cpx 2
Hfq:RNA
cpx 1
Unbound RNA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

concentration of DsrA used (20 nM or 50-fold excess to transposase RNA) complex 1 formation was inhibited by approximately 50%. We think the most straightforward interpretation of these results is as follows: the 195 nt transposase RNA has two distinct binding sites for Hfq and Hfq binds to these sites with different affinities. At the lower concentrations of DsrA there is sufficient DsrA to prevent Hfq binding to the low but not the high affinity site on the transposase RNA. At the higher DsrA concentrations there is competition for Hfq binding to both sites on the transposase RNA. As the input amount of transposase RNA was 0.4 nM, it appears that the affinity of the 'low' affinity site for Hfq on the transposase RNA is marginally better than the affinity of DsrA for Hfq. It follows that the affinity of the 'high' affinity site on the transposase RNA for Hfq is substantially better than the affinity of DsrA for Hfq. It is also possible that DsrA can compete more readily with complex 2 because it represents a weaker association between Hfq and a particular conformation of the RNA species, whereas Hfq binding to the conformation of RNA present in complex 1 can resist DsrA competition more strongly. In a second experiment we used poly$(A)_{18}$ RNA as a competitor. Poly(A) RNA containing stretches of 16 to 27 adenines are known to bind with high affinity (1.4-1.6 nM) to the distal RNA binding face on Hfq (39, 40). Interestingly we did not see competition in this experiment. Instead, at relatively high concentrations of poly$(A)_{18}$ RNA we observed both complex 2 and what appears to be the formation of a ternary complex (complex 3) in which Hfq presumably binds both the transposase RNA and poly$(A)_{18}$.

5.4.0 Discussion

We show in the current work that Tn5 transposition is significantly increased in an Δ*hfq* strain and that the extent of increase is dependent on the bacterial growth conditions. These results indicate that Hfq status plays an important role in the regulation of Tn5 transposition despite the fact that Tn5 is not known to be regulated by an antisense or *trans*-encoded RNA. This work suggests that Hfq status has a substantial effect on the translational expression of transposase and that the expression of transposase correlates with *in vivo* transposition levels. To our knowledge this is only the second example of a bacterial transposon that is modulated at the level of transposase production by Hfq. These experiments reveal Hfq may act as a regulator of transposition in general.

5.4.1 Hfq negatively regulates Tn5 transposition *in vivo*

We observed that disruption of the *hfq* gene resulted in either a moderate or strong increase in Tn5 transposition (either 5- or 58-fold). This identifies Hfq as a potent negative regulator of Tn5 transposition, which likely limits potentially destructive chromosomal rearrangements in the host caused by unregulated transposition events. Furthermore, we present the first example of Hfq negatively regulating a transposon to different degrees depending on the growth conditions used. This may not be surprising as Hfq regulation is often triggered by environmental changes (41-43).

Hfq-dependent regulation occurs by three main avenues under stress conditions: 1) indirect regulation that results from maintained translation efficiency of the *rpoS* mRNA (encoding the stress sigma factor σ^S) (44, 45), the down-regulation of σ^E-mediated envelope stress response (46-48) and the maintenance of σ^H-mediated cytoplasmic stress response (47); 2) regulation of genes that require only Hfq under certain conditions (for

example Hfq is involved in up-regulating the membrane transporter genes such as AcrAB, part of the AcrAB-TolC efflux system, in response to antibiotics) (49); and 3) regulation of genes that depend on the differential expression of regulatory sRNAs whose biological function requires Hfq. For example, Hfq is involved in the phosphosugar stress response when bacterial cells accumulate high levels of glucose-6-phosphate that cannot be metabolized. In this situation, Hfq mediates pairing between the sRNA SgrS and the mRNA *ptsG* (which produces a protein in the glucose-specific PTS transporter system) and causes rapid loss of *ptsG* mRNA through RNaseE-dependent degradation (50, 51). As a regulator of transposons and an integral factor in responding to changes in growth conditions, Hfq may provide a link between environmental stimuli and transposition rates.

In contrast there are few examples demonstrating the direct impacts of environmental stimuli on transposition. In response to high levels of certain metals, transcription is increased for several IS element transposases resulting in metal tolerance (52) and the bacteriophage Mu transposon is proposed to be involved in transposon-based adaptive mutations by spontaneously forming *araB-lacZ* fusions in nutrient-starved cultures (53, 54). Furthermore there are even fewer examples of host factors that induce conditional regulation of transposition. Nonetheless, it has been proposed that host proteins involved in cellular pathways such as metabolism, transport and redox form a complex group of proteins that detect the cell status and somehow relay this information to transposition systems (55).

5.4.2 Hfq down-regulates IS50 transposase expression

We observed dramatic increases in β-galactosidase activity in the Δ*hfq* strain compared to the *hfq*+ strain with transposase-*lacZ* translational fusions, indicating that Hfq down-regulates Tn5 transposase expression. These observations were not dependent on read-through transcription of the Tn5 transcript however nucleotides 1-46 of the transposase transcript are presumed to contain determinants for Hfq-repression of transposase. We determined that the expression of transposase from these assays in *hfq*+ and Δ*hfq* strains was not due to differences in transcription initiation or half-life of the transposase transcript. These results implicate Hfq in translational regulation of transposase.

Although our results are consistent with the fact that increased transposase expression in the Δ*hfq* strain leads to increased transposition, previous results have suggested that Tn5 transposition is not affected by increasing transposase levels. However these experiments were not conducted in an Δ*hfq* strain and it is possible that increasing transposase levels under the conditions used in this work does result in increased transposition (56). Also, conflicting evidence was presented in the previous experiments as the results from one assay did indicate that a strong promoter of transposase led to increased transposition. It was proposed that increasing transposase concentration beyond a certain threshold causes non-productive self-multimerization, but that this regulatory mechanism exists to mainly limit transposition in *trans*. In the current work however, transposition was measured in *cis*. Nonetheless, in the wild-type strain Hfq mediates strong negative regulation of transposase expression. We presume that this effect is biologically significant, as there is evidence that other host factors known to negatively regulate transposase expression (Dam methylase and LexA transcriptional regulator) also cause decreases in transposition rates (57, 58).

Since there is no evidence to suggest that Tn5 transposition is regulated by an RNA species, we favour a mechanism of translational regulation that does not involve sRNA-mRNA pairing. It is therefore important to know that Hfq can mediate this type of post-transcriptional regulation. In fact, there are several examples of proteins modulated by Hfq at the translational level that do not involve a *cis*- or *trans*-encoded RNA species. For example, Hfq autoregulates its own translational synthesis by interacting with two binding sites in the transcript near the ribosome binding site to hinder translation initiation (59). In addition, Hfq binding to the 5' untranslated region of the outer membrane protein (*ompA*) transcript alters the RNA structure such that formation of the translation initiation complex is obstructed (60, 61). At this time we have little direct evidence for this type of mechanism. However, the sequence determinants for translational regulation by Hfq may lie close to the ribosome binding site and/or the initiator codon of transposase and we provide *in vitro* evidence that Hfq binds the 5' region of the transposase transcript. Clearly further analysis is needed to define the most minimal region of transposase transcript necessary for Hfq binding and translational repression (for example, by introducing point mutations into the transposase RNA for *in vitro* binding assays or translational fusion analysis) to begin to assess the mechanism responsible for down-regulating transposase translation efficiency. Several other mechanistic possibilities for Hfq-dependent regulation that is not mediated by a regulatory RNA are depicted in Figure 1D, II.

5.4.3 Hfq interacts with the transposase transcript through the proximal site

In vitro binding assays indicated that Hfq has a relatively strong affinity for the 5' region of the Tn5 transposase transcript (K_d <60 nM) providing evidence that Hfq directly affects transposase expression. The binding of Hfq to the transposase transcript occurs in the range of binding affinities reported for other sRNA and mRNA species that bind Hfq. For example, the observed binding affinities for Hfq to *rpoS* mRNA, RydC sRNA and DsrA sRNA are approximately 280 nM (62), 50 nM and 21 nM (38) respectively, per Hfq hexamer. Although accurate binding affinities have not been determined for Hfq to the Tn10 RNA species RNA-IN (transposase mRNA) and RNA-OUT (sRNA), estimated K_d values for these RNAs are approximately <15 nM and 120 nM respectively (J. Ross, personal communication). Although there is a putative Hfq binding site within the 5' region of the transposase transcript, deletion analysis revealed that this site is dispensable for Hfq-dependent translational regulation of transposase, indicating that at least one other binding site exists. Future experiments will need to determine the exact binding sites for Hfq in the transposase transcript in order to establish a mechanism through which translational control is mediated.

Of the two potential RNA-binding faces on Hfq, results from the competitive binding assays suggest that the transposase transcript binds the proximal face, which predominately binds sRNAs as opposed to mRNAs based on previous experiments. Thus it is not surprising that the putative binding site for Hfq in the transposase transcript (5'-AAUCAAUAA-3') was dispensable for regulation of translation, as this sequence has been found to bind the distal face specifically (63, 39). Furthermore, these results suggest that a *trans*-encoded sRNA or antisense RNA is not required to regulate transposase expression as this type of RNA would presumably contact the proximal site, which would

be occupied with the transposase mRNA. However, we do not want to discount the possibility that several RNA molecules may be bound to the same site as the transposase transcript or bound to the distal site in order to regulate translation via Hfq.

5.4.4 Future aims and biological significance of Hfq regulation

We provide evidence for a model of Tn5 regulation by Hfq that involves down-regulation of Tn5 transposase expression and decreased transposition. We also provide evidence that this occurs at the stage of translation through a direct contact between the transcript and Hfq. Future experiments must determine whether Hfq is directly involved in this regulation or if these observations are due to an indirect factor regulated by Hfq. It is also important to determine if Hfq directly affects translation of transposase, for example by using an *in vitro* translation system, and to determine how Hfq affects translation of transposase by probing for changes in mRNA secondary structure upon Hfq association. As mentioned previously it would be beneficial to know the specific Hfq binding site(s) in the transposase transcript, and determine the importance of the site(s) for *in vivo* regulation. Future experiments may also focus on determining why and how bacterial growth differences are responsible for the Hfq-dependent differential regulation in Tn5 transposition, and the role of Hfq in other transposon systems.

The biological significance of Hfq regulation of transposition may be two-fold: as a potent negative regulator Hfq limits transposition to protect the cell, and as a protein intimately involved in stress response, Hfq links external stimuli to transposition rates. Because stress induction typically enhances the transcription of several sRNA species that bind Hfq, the amount of unbound Hfq protein may change during different growth

conditions. Although Hfq is a highly expressed protein, there are many binding sites in the nucleoid, RNA pool, and ribosomes, therefore it is possible that the amount of unbound Hfq changes significantly upon induction of RNA species (64). Under conditions or stimuli that require Hfq to be bound to many different sRNA species at once, transposition rates may change. This could serve to link the external stimuli directly to transposition of those transposons regulated by Hfq.

Table 5.1 Oligonucleotides used in this study.

Name	Sequence 5'-3'
CW1	NNNCCGCGGAGTCGTTTTACAACGTCGTGACTGGGAAAACCC
CW2	NNNAGATCTTTATTTTTGACACCAGACCAACTGGTAATGGTAGCG
CW3	NNNNNGAATTCTGCCTGGCGGCAGTAGCGCGGTGG
CW4	NNNNNGAATTCAAAAGGCCATCCGTCAGGATGGCCTTC
CW5	NNTCTAGANN CAGGTTGTTGGTGCTATC
CW6	NNAAGCTTNNTTATTCGGTTTCTTCGCT
CW7	TAATACGACTCACTATAGGCAGGTGACCTCTTAAGATGGTAACGTTCATGATAAC
CW8	GGCGGCTTCACTACCCTCTGATGAGATGG
T1TaseF	GACCTCTTAAGATGGTAACGTTCATG
T1TaseR	GCCGAAGAGAACACAGATTTAGC
T1TaseProbe[a]	6FAM-TAACTTCTGCTCTTCATCGTG-MGBNFQ
16SF	ACCAGGGCTACACACGTGCTA
16SR	TCTCGCGAGGTCGCTTCT
16SProbe[a]	6FAM- AATGGCGCATACAAA MGBNFQ

a. 6FAM represents 6-carboxyfluorescein. MGBNFQ represents the quencher dye.

5.5 References

1. Gentry,L.O. (1991) Bacterial resistance. *The Orthopedic clinics of North America*, **22**, 379-88.
2. Olsen,J.E. (1999) Antibiotic resistance: genetic mechanisms and mobility. *Acta Vet Scand Suppl*, **92**, 15-22.
3. Nagy,Z. and Chandler,M. (2004) Regulation of transposition in bacteria. *Research in microbiology*, **155**, 387-98.
4. Hengge-Aronis,R. (1999) Interplay of global regulators and cell physiology in the general stress response of Escherichia coli. *Current opinion in microbiology*, **2**, 148-52.
5. Suzuki,T., Ueguchi,C. and Mizuno,T. (1996) H-NS regulates OmpF expression through micF antisense RNA in Escherichia coli. *Journal of bacteriology*, **178**, 3650-3.
6. Dorman,C.J. (2004) H-NS: a universal regulator for a dynamic genome. *Nature reviews. Microbiology*, **2**, 391-400.
7. Hommais,F., Krin,E., Laurent-Winter,C., Soutourina,O., Malpertuy,A., Caer,J.P. Le, Danchin,A. and Bertin,P. (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Molecular microbiology*, **40**, 20-36.
8. Amit,R., Oppenheim,A.B. and Stavans,J. (2003) Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophysical journal*, **84**, 2467-73.
9. Vogel,J. and Luisi,B.F. (2011) Hfq and its constellation of RNA. *Nature reviews. Microbiology*, **9**, 578-89.
10. Ross,J.A., Wardle,S.J. and Haniford,D.B. (2010) Tn10/IS10 transposition is downregulated at the level of transposase expression by the RNA-binding protein Hfq. *Mol Micro*, **78**, 607-21.
11. Ansong,C., Yoon,H., Porwollik,S., Mottaz-Brewer,H., Petritis,B.O., Jaitly,N., Adkins,J.N., McClelland,M., Heffron,F. and Smith,R.D. (2009) Global systems-level analysis of Hfq and SmpB deletion mutants in Salmonella: implications for virulence and global protein translation. *PloS one*, **4**, e4809.
12. Sittka,A., Pfeiffer,V., Tedin,K. and Vogel,J. (2007) The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. *Molecular microbiology*, **63**, 193-217.
13. Sledjeski,D.D., Whitman,C. and Zhang,A. (2001) Hfq is necessary for regulation by the untranslated RNA DsrA. *Journal of bacteriology*, **183**, 1997-2005.
14. Zhang,A., Altuvia,S., Tiwari,A., Argaman,L., Hengge-Aronis,R. and Storz,G. (1998) The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *The EMBO journal*, **17**, 6061-8.

15. Majdalani,N., Chen,S., Murrow,J., St John,K. and Gottesman,S. (2001) Regulation of RpoS by a novel small RNA: the characterization of RprA. *Molecular microbiology*, **39**, 1382-94.
16. Massé,E. and Gottesman,S. (2002) A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 4620-5.
17. Majdalani,N., Vanderpool,C.K. and Gottesman,S. Bacterial small RNA regulators. *Critical reviews in biochemistry and molecular biology*, **40**, 93-113.
18. Waters,L.S. and Storz,G. (2009) Regulatory RNAs in bacteria. *Cell*, **136**, 615-28.
19. Butland,G., Peregrín-Alvarez,J.M., Li,J., Yang,W., Yang,X., Canadien,V., Starostine,A., Richards,D., Beattie,B., Krogan,N., et al. (2005) Interaction network containing conserved and essential protein complexes in Escherichia coli. *Nature*, **433**, 531-7.
20. Morita,T. and Aiba,H. (2011) RNase E action at a distance: degradation of target mRNAs mediated by an Hfq-binding small RNA in bacteria. *Genes & development*, **25**, 294-8.
21. Mohanty,B.K., Maples,V.F. and Kushner,S.R. (2004) The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. *Molecular microbiology*, **54**, 905-20.
22. Hajnsdorf,E. and Régnier,P. (2000) Host factor Hfq of Escherichia coli stimulates elongation of poly(A) tails by poly(A) polymerase I. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 1501-5.
23. Naumann,T.A. and Reznikoff,W.S. (2000) Trans catalysis in Tn5 transposition. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 8944-9.
24. Alton,N.K. and Vapnek,D. (1979) Nucleotide sequence analysis of the chloramphenicol resistance transposon Tn9. *Nature*, **282**, 864-9.
25. Wiegand,T.W. and Reznikoff,W.S. (1992) Characterization of two hypertransposing Tn5 mutants. *Journal of bacteriology*, **174**, 1229-39.
26. Guzman,L.M., Belin,D., Carson,M.J. and Beckwith,J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of bacteriology*, **177**, 4121-30.
27. Wang,R.F. and Kushner,S.R. (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in Escherichia coli. *Gene*, **100**, 195-9.
28. Morisato,D. and Kleckner,N. (1987) Tn10 transposition and circle formation in vitro. *Cell*, **51**, 101-11.
29. Foster,T.J., Davis,M.A., Roberts,D.E., Takeshita,K. and Kleckner,N. (1981) Genetic organization of transposon Tn10. *Cell*, **23**, 201-13.

30. Jain,C. and Kleckner,N. (1993) IS10 mRNA stability and steady state levels in Escherichia coli: indirect effects of translation and role of rne function. *Molecular microbiology*, **9**, 233-47.
31. Baker,C.S., Eöry,L. a, Yakhnin,H., Mercante,J., Romeo,T. and Babitzke,P. (2007) CsrA inhibits translation initiation of Escherichia coli hfq by binding to a single site overlapping the Shine-Dalgarno sequence. *Journal of bacteriology*, **189**, 5472-81.
32. Liu,Y., Wu,N., Dong,J., Gao,Y., Zhang,X., Mu,C., Shao,N. and Yang,G. (2010) Hfq is a global regulator that controls the pathogenicity of Staphylococcus aureus. *PloS one*, **5**.
33. Pfaffl,M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research*, **29**, e45.
34. Morita,T., Kawamoto,H., Mizota,T., Inada,T. and Aiba,H. (2004) Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in Escherichia coli. *Molecular microbiology*, **54**, 1063-75.
35. Pedersen,S. and Reeh,S. (1978) Functional mRNA half lives in E. coli. *Mol Gen Genet*, **166**, 329-36.
36. Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2009) The global bacterial regulator H-NS promotes transpososome formation and transposition in the Tn5 system. *Nucleic acids research*, **37**, 309-21.
37. McCommas,S.A. and Syvanen,M. (1988) Temporal control of transposition in Tn5. *Journal of bacteriology*, **170**, 889-94.
38. Mikulecky,P.J., Kaw,M.K., Brescia,C.C., Takach,J.C., Sledjeski,D.D. and Feig,A.L. (2004) Escherichia coli Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nature structural & molecular biology*, **11**, 1206-14.
39. Link,T.M., Valentin-Hansen,P. and Brennan,R.G. (2009) Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19292-7.
40. Sun,X. and Wartell,R.M. (2006) Escherichia coli Hfq binds A18 and DsrA domain II with similar 2:1 Hfq6/RNA stoichiometry using different surface sites. *Biochemistry*, **45**, 4875-87.
41. Aiba,H. (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Current opinion in microbiology*, **10**, 134-9.
42. Storz,G. and Gottesman,S. (2006) Versatile Roles of Small RNA Regulators in Bacteria. In Gesteland,R., Cech,T., Atkins,J. (eds), *The RNA World*. Cold Spring Harbor Lab Press, New York, pp. 567-594.
43. Gottesman,S. (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends in genetics*, **21**, 399-404.
44. Muffler,A., Traulsen,D.D., Fischer,D., Lange,R. and Hengge-Aronis,R. (1997) The RNA-binding protein HF-I plays a global regulatory role which is largely, but not

exclusively, due to its role in expression of the sigmaS subunit of RNA polymerase in Escherichia coli. *Journal of bacteriology*, **179**, 297-300.

45. Brown,L. and Elliott,T. (1996) Efficient translation of the RpoS sigma factor in Salmonella typhimurium requires host factor I, an RNA-binding protein encoded by the hfq gene. *Journal of bacteriology*, **178**, 3763-70.
46. Figueroa-Bossi,N., Lemire,S., Maloriol,D., Balbontín,R., Casadesús,J. and Bossi,L. (2006) Loss of Hfq activates the sigmaE-dependent envelope stress response in Salmonella enterica. *Molecular microbiology*, **62**, 838-52.
47. Guisbert,E., Rhodius,V.A., Ahuja,N., Witkin,E. and Gross,C.A. (2007) Hfq modulates the sigmaE-mediated envelope stress response and the sigma32-mediated cytoplasmic stress response in Escherichia coli. *Journal of bacteriology*, **189**, 1963-73.
48. Thompson,K.M., Rhodius,V.A. and Gottesman,S. (2007) SigmaE regulates and is regulated by a small RNA in Escherichia coli. *Journal of bacteriology*, **189**, 4243-56.
49. Yamada,J., Yamasaki,S., Hirakawa,H., Hayashi-Nishino,M., Yamaguchi,A. and Nishino,K. (2010) Impact of the RNA chaperone Hfq on multidrug resistance in Escherichia coli. *The Journal of antimicrobial chemotherapy*, **65**, 853-8.
50. Vanderpool,C.K. and Gottesman,S. (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Molecular microbiology*, **54**, 1076-89.
51. Maki,K., Uno,K., Morita,T. and Aiba,H. (2008) RNA, but not protein partners, is directly responsible for translational silencing by a bacterial Hfq-binding small RNA. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 10332-7.
52. Brocklehurst,K.R. and Morby,A.P. (2000) Metal-ion tolerance in Escherichia coli: analysis of transcriptional profiles by gene-array technology. *Microbiology*, **146 (Pt 9)**, 2277-82.
53. Casadaban,M.J. (1976) Transposition and fusion of the lac genes to selected promoters in Escherichia coli using bacteriophage lambda and Mu. *Journal of molecular biology*, **104**, 541-55.
54. Shapiro,J.A. (1997) Genome organization, natural genetic engineering and adaptive mutation. *Trends in genetics*, **13**, 98-104.
55. Twiss,E., Coros,A.M., Tavakoli,N.P. and Derbyshire,K.M. (2005) Transposition is modulated by a diverse set of host factors in Escherichia coli and is stimulated by nutritional stress. *Molecular microbiology*, **57**, 1593-607.
56. Weinreich,M.D., Gasch,A. and Reznikoff,W.S. (1994) Evidence that the cis preference of the Tn5 transposase is caused by nonproductive multimerization. *Genes & development*, **8**, 2363-74.

57. Yin,J.C., Krebs,M.P. and Reznikoff,W.S. (1988) Effect of dam methylation on Tn5 transposition. *Journal of molecular biology*, **199**, 35-45.
58. Kuan,C.T. and Tessman,I. (1991) LexA protein of Escherichia coli represses expression of the Tn5 transposase gene. *Journal of bacteriology*, **173**, 6406-10.
59. Vecerek,B., Moll,I. and Bläsi,U. (2005) Translational autocontrol of the Escherichia coli hfq RNA chaperone gene. *RNA*, **11**, 976-84.
60. Vytvytska,O., Moll,I., Kaberdin,V.R., Gabain,A. von and Bläsi,U. (2000) Hfq (HF1) stimulates ompA mRNA decay by interfering with ribosome binding. *Genes & development*, **14**, 1109-18.
61. Moll,I., Afonyushkin,T., Vytvytska,O., Kaberdin,V.R. and Bläsi,U. (2003) Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA*, **9**, 1308-14.
62. Soper,T.J. and Woodson,S.A. (2008) The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA*, **14**, 1907-17.
63. Brennan,R.G. and Link,T.M. (2007) Hfq structure, function and ligand binding. *Current opinion in microbiology*, **10**, 125-33.
64. Ali Azam,T., Iwata,A., Nishimura,A., Ueda,S. and Ishihama,A. (1999) Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *Journal of bacteriology*, **181**, 6361-70.

CHAPTER 6

6 GENERAL DISCUSSION

During the past several decades the Tn5 transposon has served as an excellent system for studying bacterial transposition. The main objectives of this research are to provide evidence of target selection and strand transfer *in vitro* and discover new host regulatory mechanisms of Tn5. The results presented in Chapters 2-5 further explain how the movement of bacterial transposons is controlled, which has important implications for understanding transposon biology and the host-transposon relationship. The following sections emphasize the main conclusions of this thesis, relate these to other transposition systems, attempt to explain the regulatory mechanisms, and address the importance of this work.

6.1 Formation and characterization of a Tn5 strand transfer complex

Despite the wide use of transposons as mutagenic agents, they generally display a preference for selecting target sites. For transposons such as Tn5 and Mu, insertion occurs at many sequence-unrelated targets and both transposons display low-specificity consensus targets (1). This suggests that structural characteristics of the target DNA or transpososome-induced target deformations, rather than sequence of the target, may assist in insertion. Biochemical evidence indicates that many transposase proteins are required to ‘read-out’ the target shape before insertion. However, the phenomenon of selecting deformed targets appears to be only somewhat conserved among transposons. Analysis of *in vivo* target site duplications, based on computer algorithms that detect curvature and ‘bendability’, determined that duplications were often more flexible than an average

sequence. Transposons such as bacteriophage Mu, the bacterial transposons Tn7 and Tn10, and the eukaryotic elements Sleeping Beauty and P-element were determined to preferentially insert into these types of sequences (2-6). Despite the fact that Mu target sites have low specificity, nucleotides mismatches can create a strong preference for insertion, thus it follows that target distortion may dictate preference for Mu (7). Interestingly V(D)J recombination and HIV-1 integration are also targeted into distorted sites (8). A distinguishing feature of the Tn10 transposon is that target DNA bending is induced upon target capture/insertion by protein-DNA interactions in the transpososome complex. The molecular basis of target DNA bending was investigated by Pribil *et al*. and it was discovered that removal of the scissile phosphates could dramatically enhance insertion efficiency into a site (9).

Because of the relaxed insertion specificity for the Tn5 transposon the molecular details of target insertion were never addressed. Although a few 'hot-spots' were isolated these have never been used to study Tn5 strand transfer *in vitro*. The work in this thesis presents the first experiments investigating Tn5 target insertion *in vitro* and validates the use of an *in vivo* target for *in vitro* studies of transposon integration. Similarly to Tn10, the observations described in Chapter 2 indicate that target DNA bending is critical for insertion efficiency. The results provide evidence that Tn5 target insertion occurs readily into specific sequences and the requirements of this step are similar to the structurally and functionally related transposon Tn10.

The flexibility required at the site of DNA breakage and joining is presumed to lower the energy barrier to the chemically reactive groups and/or be preferred for transpososome binding during insertion. This feature has not been documented for many

transposons, possibly due to the lack of information about target sequences, however a similar phenomenon was found for the Tn7 transposon. The Tn7-encoded TnsD protein binds to a specific site adjacent to the Tn7 insertion site, resulting in a distortion of the DNA structure and subsequent binding by the TnsC protein that recruits the co-transposase complex for insertion (10).

For the Tn5 transposon, flexibility at the centre of a target site was also determined to be important for strand transfer efficiency. Since the centrally located phosphate is not a reactive group, distortion at this region of a target site may facilitate the structural requirements of multiple transpososomes bound to the target site as presumed by the microfilament model (see Chapter 2 Discussion) (11). This mechanism likely depends on the presence of transposase, the target site sequence, and the presence of overlapping target sites, in addition to flexibility at the scissile and central positions. It is possible that a single transpososome interacts with the target site, but in this case, it is less obvious why distortion in the centre of the site would benefit insertion, especially without a structural model to consult.

Obtaining a high-resoluation structure of a transpososome bound to a target molecule would be useful for many transposition systems; the implications for understanding HIV-integration and designing inhibitor molecules to this process would also be advantageous. To this end, I attempted to create a relatively pure mixture of Tn5 strand transfer complex. I used an overlapping target site where insertion occurred primarily into one site and utilized a purification technique to eliminate unbound, contaminating DNA. However, there were difficulties obtaining a truly homogeneous

sample that would have facilitated crystallography trials, and concurrent work to understand the effect of H-NS on Tn5 transposition became a priority.

6.2.0 The importance of H-NS and Hfq regulation of transposition systems

Host proteins represent a critical level of transposon regulation; however their impact on transposition in general is not well-understood. The results presented in Chapters 3 to 5 provide important insights into two new regulatory mechanisms of the Tn5 transposon by the proteins H-NS and Hfq. The results indicate that the transcriptional repressor H-NS promotes Tn5 activity, whereas the RNA chaperone Hfq inhibits Tn5 activity. I provide evidence that H-NS promotes Tn5 transposition by directly assisting with transpososome assembly whereas Hfq inhibits Tn5 transposition by altering expression of the transposase protein at the translational level. Although these mechanisms of transposon regulation are well-documented, the importance of these specific host factors to Tn5 was previously unknown. Furthermore, no proteins had been discovered that specifically impact Tn5 transpososome assembly prior to this work. Although H-NS also regulates other transposition (Tn10, IS1, IS903) and recombination (FimS) systems, Hfq has been found to regulate only one other transposition system (Tn10).

6.2.1 The transpososome represents a critical stage for regulation by host factors

Host proteins can affect any aspect of transposon biology. The majority of these proteins are nucleic-acid binding proteins that impact assembly of transpososomes required for activity. Bacteriophage Mu is assisted by specific DNA and protein co-factors during transpososome assembly. HU, in addition to the transposase protein MuA, are integral in

forming the higher-order transpososome complex (12). MuA contacts both ends of the Mu genome at specific sequences (*att* sites). The N-terminal domain of MuA interacts with an internal activating sequence (IAS) located approximately 1 kb from the left end of the Mu genome while another region interacts directly with the *att* sequence. Footprinting experiments located a binding site for the HU protein between two MuA binding sites in the *att* region and also detected a strong cleavage pattern in the area (suggestive of a strong bend). HU, which happens to share many features with H-NS, brings the two *att* sequences together by wrapping around the spacer DNA. This aids in the formation of the first transpososome complex of the pathway at the left end of the genome where the MuA protein must interact with both the *att* sequences and the IAS. Similarly to that proposed for H-NS and Tn5 transposition, HU is required at an early stage of Mu transpososome assembly and is not required for the chemical step of strand transfer. In contrast, H-NS association with the Tn10 transpososome directly impacts downstream events such as stabilization of the complex (likely through protein-protein contacts) and promoting productive strand transfer events.

IHF, a protein similar in activity to H-NS and HU, also plays a role in transpososome assembly of both Mu and Tn10 transposition systems. Under certain conditions IHF binds the IAS element in Mu and induces a sharp bend to facilitate transpososome formation (13). For Tn10, IHF is necessary to by-pass the requirement for negative supercoiling *in vitro* likely through IHF-mediated bending that assembles higher-order nucleo-protein complexes (14, 15). IHF is required for Tn10 transpososome assembly *in vitro* by promoting DNA bends at IHF-binding sites near the transposase binding sites (14). The ‘folded’ transposon ends are proposed to create a molecular spring

that when released provides energy for a conformational change in the transpososome that facilitates target capture (16).

Host factor regulation is also prevalent in eukaryotic transposition systems. For example, transposition of the P-element found in *Drosophila* species requires the host enzyme, IRBP (Inverted Repeat Binding Protein), which binds the inverted terminal repeats and is responsible for the element's initial excision from the chromosome (17). HMGB1 is a non-histone nuclear protein associated with chromatin that has the ability to bend DNA and is implicated in the Sleeping Beauty transposon (18, 19). For the reconstituted 'cut-and-paste' transposon, Sleeping Beauty (SB), mouse cells lacking HMGB1 have diminished SB activity. *In vivo* HMGB1 was found to interact directly with SB transposase, suggesting that protein-protein interactions recruit HMGB1 to the transpososome complex (20). Once bound by transposase, HMGB1 stimulates specific binding of the transposase to a set of 'inner' direct repeats near the end of the transposon. HMGB1 is proposed to induce a bend at both ends in the spacer region between the inner and outer direct repeats, which allows for the correct binding of transposase to the outer direct repeats. This specific sequence of events is essential for a catalytically competent transpososome and requires the HMGB1 protein.

The phenomenon of host proteins aiding in transpososome formation is common in DNA transposition systems. Host factors are often involved with transpososome assembly because it allows the transposon to take advantage of already present proteins and thus decreases the need for specialized protein expression from the transposon itself. In addition, host proteins are often critical to bring together the inert substrates (at least transposase and transposon DNA) often in a sequential order to form a catalytically ready

complex. The complex must be assembled correctly with the transposon ends before transposition can occur for several reasons: 1) it orients the ends properly relative to one another and increases the likelihood of accurate cleavage at the transposon/donor junction and ensures that both ends can be joined to a relatively short target DNA site; 2) it ensures that cleavage and joining is spatially and temporally linked; and, 3) it prevents joining of one end and another independent end or non-specific sequence to a target site as well as single-end insertion, which could both be destructive. These demands can be easily met with the use of nucleic-acid binding proteins or chromatin remodelling-like proteins in bacteria that can interact directly with DNA or protein-DNA complexes.

6.2.2 Expression of transposition proteins is a critical stage of regulation

As 'selfish DNA', transposons aim to propagate themselves as often as possible without harming the host cell via destructive chromosomal rearrangements. This evolutionary pressure has resulted in tight negative control of transposons. The most widely used method to effectively down-regulate transposition is limiting the expression of transposon-encoded proteins, such as transposases (21). Many endogenous promoters of bacterial transposase genes are transcribed inefficiently, such as IS10, IS50, IS21 and IS911 (22). Host factors such as transcriptional or translational regulators can be co-opted by the transposon system to facilitate repressive mechanisms. The work presented in Chapter 5 indicates that the host protein Hfq plays a substantial role in down-regulation of transposase synthesis.

There are at least two other host proteins that also limit Tn5 transposase. IS50 contains sequences for Dam methylation nearby or overlapping the transposase

transcriptional sequences that when fully methylated restrict promoter usage (23). A Δ*dam* strain enhances transposition of Tn5 by up to 10-fold. A similar phenomenon exists for IS10 and IS903. LexA, a transcriptional repressor of the SOS regulon, has also been implicated in Tn5 transposition. A *lexA*-deficient strain supports 3-fold higher transposition levels, and LexA has been proposed to repress transposase transcription initiation through a *recA*-independent mechanism that likely involves binding of LexA directly to the IS50 regulatory sequences (24).

As described in Chapter 5, the impact of Hfq on transposase expression appears to be greater than that of either Dam methylase or LexA. Hfq-mediated repression of transposase levels represents the strongest effect of a host factor on the expression of Tn5 transposase to date. Hfq exerts its effect at translation efficiency of transposase; post-transcriptional regulation of transposase synthesis is common to other transposable elements. Tn10 transposase is regulated by an antisense RNA, which is proposed to affect translation of the transposase (25, 26). Elements such as IS10 and IS50 protect themselves from impinging 'read-through' transcription that would increase transposase synthesis by encoding sequences that form an RNA secondary structure upon read-through that sequester the ribosome binding site (27, 28). In addition, programmed translational frame-shifting is common to several bacterial transposons such as IS1 and members of the IS3 family, where slippage of the ribosome one base upstream, typically due to stem-loop or pseudoknot RNA secondary structures, and resumption of translation, produces a different yet related protein (29). This mechanism combines information contained in two open reading frames and allows expression of different proteins with different functions.

6.3 Utilizing H-NS and Hfq as modulators of transposition systems

H-NS and Hfq both play important regulatory and maintenance roles in bacteria. Chapters 3-5 provide evidence that these proteins are also representative of a relatively small group of host factors that regulate bacterial transposition. What features of these proteins make them useful to regulate transposition? First, many share common properties, such as the ability to bind DNA. The nucleoid-associated proteins H-NS, Hfq, HU, IHF and Fis are small basic proteins that bind nucleic acids and play functional roles (in additional to their chromosome structuring roles) in the global regulation of essential DNA processes such as replication, recombination, and transcription. Similar characteristics exist for host proteins that affect DNA transposition in eukaryotic organisms, such as the HGMB1 protein. Second, the protein levels of these factors vary throughout the growth cycle of bacteria but are typically expressed during exponential growth. H-NS and Hfq are highly expressed in the bacterial cell and their levels are mostly stable throughout the cell cycle, particularly in exponential phase. Third, these proteins are conserved among many bacterial species. Finally, H-NS and Hfq (in addition to Fis and IHF) are pleiotropic regulators that are involved in the expression of many (>50) genes, including many whose expression is triggered by stress.

Why are these characteristics shared by many host proteins that regulate bacterial transposition? As previously stated, a large number of these proteins are nucleic-acid binding proteins that have the ability to interact directly with the transposition machinery. These proteins could therefore interact directly with transposition complexes to fine-tune transposon activity instead of the transposon maintaining its own *cis*-encoded modulating

factors. Transposition systems can ‘rely’ on these proteins since they have established functions in the cell and are generally well-expressed. Some of these host factors play roles in facilitating lateral gene transfer, which can in turn facilitate the spread of transposons to plasmids or other mobile cassettes (e.g. H-NS). Since transposition can be considered a type of lateral gene transfer, it follows that proteins involved in one aspect may also control another. Additionally, because these proteins are conserved across bacterial species, it reasons that regulation could continue after a successful lateral gene transfer event, which is critical if the transposon is to survive in a new strain or species. This has obvious implications for the spread of transposons and resistance. In eukaryotic organisms, lateral movement of transposable elements has been proposed to be limited by host factors. For example, the spread of P-elements to species beyond the drosophilids is unusual, whereas Tc1 and mariner elements have been found in many animal phyla. Since P-element requires a specific host protein (IRBP) to increase in copy number once acquired by a new organism, this can only occur in organisms that have that (or an equivalent) protein. Tc1 and mariner elements however only require transposase for transposition and thus their spread is not limited (30).

An important possibility is that many of these proteins may offer communication between the cell and the transposon, either through their cellular concentrations or by being directly involved in stress responses. Transposon regulation by host proteins such as Fis or IHF may change throughout the growth cycles as protein levels fluctuate. That is, their availability may ‘inform’ the transposon of the cell status (31). For other factors involved in stress response pathways like H-NS or Hfq, transposition rates may vary depending on the presence or absence of external conditions that can change the protein

function or availability. Communication between the bacterial state and transposon activity could assist in the survival of the transposon. For example, the survival of a mobile element may depend on increased activity in situations where the host cell is about to die. Enhanced mobility may increase the chance of the element successfully undergoing a lateral gene transfer event and thus escaping the cell.

H-NS and Hfq are now identified as members of a global network of proteins that affect individual transposon activity by different mechanisms. It is unlikely that all the proteins in the network have been discovered. Assembling these networks will have implications for understanding the activity of transposons, and the host cellular pathways involved in transposon regulation.

6.4 Evidence of differential regulation of transposition

The results presented in Chapter 5 provide further evidence that host factors can differentially regulate transposition depending on the nutritional conditions presented to the bacteria. Hfq was found to down-regulate transposition to the greatest extent in minimal media with glucose where growth of the bacterial cell is slower due to lack of abundant nutrients.

The host protein Fis also regulates Tn5 differentially based on the physiology of the host, although it does not do so in the presence or absence of a particular component (32). Fis exerts its optimal regulation only during the exponential phase of growth when its levels are significantly high. During this time, Tn5 and IS50 transposition frequencies are presumed to be stimulated by the presence of Fis; however when the inside end (IE) is unmethylated, which would normally allow for an even higher level of IS50

transposition, the presence of Fis actually inhibited this activity, presumably by directly interacting with the transposon IE DNA. This mechanism is proposed to limit transposition in conditions when the IS element would be 'over-active', while maintaining moderate positive regulation of the full transposon. Another example demonstrating this phenomenon is that mutations in the *aspA* gene were determined to induce transposition earlier in the growth of a colony, indicating that transposition of some transposons is restricted early in growth and a connection exists between nutrient availability and transposition (33). Aspartase is important in fumarate respiration during anaerobic growth and the mutations are proposed to disperse a signal that triggers transposition. Unfortunately the pathway or factors that link these processes are unknown.

I did not examine the possibility that H-NS differentially regulates transposition. It is possible that factors such as salt or temperature could lead to different transposition rates in the Δ*hns* strain compared to *hns*+, since these environmental factors signal regulation of target genes by H-NS (34, 35). In fact, insertion sequences such as IS1, IS30 and IS911 in addition to Tn3 display varied transposition frequencies dependent on temperature growth of the bacterium (36-38). Based on the observations in Chapters 5 and previous work, experiments should be conducted in varying media types when assessing the impact of host factors on transposition so that conditional regulation does not go undetected.

6.5 Possible explanations of Hfq-dependent negative regulation of Tn5 transposition in minimal media with glucose

If we presume that Hfq-dependent regulation of Tn5 in minimal media with glucose is caused by an environmental stimuli it is worth discussing the possible factor responsible. Although there are several differences in composition between the rich and minimal media used in these studies and any number of them may be responsible for the Hfq-dependent differential regulation, here we have focussed the discussion on the most obvious difference: glucose. The rich media used for bacterial growth detailed in Chapter 5 is based on the Luria Broth described by Miller and does not contain any sugar compounds, whereas the minimal media used contained glucose (39).

Interestingly, the global protein expression profiles of an Δ*hfq* strain of *Salmonella* grown in rich versus minimal media suggest that Hfq plays a significant role in central metabolism (40). At the protein level, Hfq differentially regulates numerous proteins involved in glycolysis and gluconeogenesis, the TCA cycle, pyruvate decarboxylation and pentose phosphate shunt in addition to arginine catabolism (AstABCD) and propanoite catabolism (PrpBCDE). This suggests that Hfq levels may be regulated by sugar molecules such as glucose in order to differentially regulate the above pathways. In fact, it was recently determined that Hfq levels are inhibited when growth occurs in the absence of glucose by the CRP-cAMP complex (41). cAMP (cyclic AMP) and CRP (cyclic AMP receptor protein) form a complex that acts as a pleiotropic transcriptional regulator for glucose metabolism. The CRP-cAMP complex binds a particular region of the *hfq* promoter and represses Hfq expression at the transcriptional level. This means that in the presence of glucose repression of the *hfq* promoter is released and Hfq levels are increased. Increased levels of Hfq could allow for further inhibition of transposition by a direct mechanism where Hfq interacts with the

transposase mRNA to repress transposase expression post-transcription. In a strain deficient in Hfq grown with glucose in the media, transposition would appear quite high (as I observed in Chapter 5). Additionally, microarray analysis revealed that *hfq* transcripts are increased by the presence of RpoS when experimental strains are grown in minimal media with glucose (42). Determining Hfq protein levels in different media used in our experiments and assessing transposition rates in CRP-cAMP-altered Δ*hfq* strains could provide evidence for or against this model.

A consequence of CRP-cAMP-mediated repression of *hfq* in the absence of glucose is stabilization of *ompA* mRNA, as decreased amounts of Hfq occupy less molecules of *ompA* mRNA. In the presence of glucose and increased Hfq concentration, Hfq binds to an RNaseE cleavage site in the 5'UTR of *ompA* mRNA and sequesters the ribosome binding site resulting in destabilization and stimulated decay (41). This suggests that the presence of glucose alters the stability of other mRNA species, which may indirectly affect transposition in an Hfq-dependent manner (either directly by affecting transposase expression or indirectly by affecting an unknown aspect of transposition). Additionally, the expression profile of sRNA species is known to change drastically in different media conditions (such as glucose) suggesting that an sRNA may contribute to the change in transposition rates in an Hfq-dependent manner. One possibility is that an sRNA interacts with the transposase transcript to further inhibit translation of transposase *in vivo* in addition to the action of Hfq. This would be an interesting possibility as *trans*-encoded RNA regulation of bacterial DNA transposons has not been documented.

6.6 Concluding remarks and future directions

The work presented in this thesis proposes that host networks that govern transposition are complex and under-appreciated with respect to bacterial transposon regulation. I provide evidence to support the claim that host proteins and physiology are critical factors that impact the activity of transposons. Without a clear understanding of how, when and why transposons are mobilized, we cannot presume to explain their rapid spread through populations or their contribution to bacterial evolution and disease. Although it is unlikely that a unified host network regulates each bacterial transposon, many host proteins have been implicated in several transposable systems. The possibility exists that proteins involved in these regulatory networks are closely connected to the process of lateral gene transfer in bacteria, which also contributes to the rapid evolution and adaptability of bacteria.

Identifying the components and pathways that govern transposons presents a challenge. Systematically testing candidate host factors for an impact on transposition is a tedious task because strains must be created and validated, and measuring transposition rates of many transposons can be difficult. In addition it is difficult to hypothesize which candidate host genes may be implicated in transposition. Conducting large-scale mutational analyses and determining the effect on multiple transposition systems at one time appears attractive. However the high-throughput nature of these assays can be compromised due to the difficulty in understanding the results because of interplay between the systems one is studying and the consequences or stress on the strain in which transposition is being measured. Nonetheless, this type of experiment has been previously documented in bacteria (33). Twiss *et al*. detected previously unknown host factors, such

as H-NS, as regulators of several model transposable elements such as Tn10 and IS903. Experiments similar to these will provide more information on the depth of host factor regulation of transposable elements.

There are several other implications for future research in this area. An in-depth analysis of the mechanism of Hfq- and H-NS-mediated regulation of transposition is necessary. This could involve the use of an *in vitro* translation system to test the hypothesis that Hfq directly affects transposase expression and the altering of growth conditions to determine the factor responsible for Hfq regulation of Tn5 transposition. Structural probing using X-ray crystallography may be able to determine the Tn5 transpososome bound to the H-NS protein. This information would provide critical protein-protein and protein-DNA interactions that may facilitate transpososome assembly by H-NS and could provide the first structure of full-length H-NS. It is possible that once H-NS is bound to the transpososome the flexible linker region is 'held' in place through interactions with the transposase protein. Structural studies could also produce the first Tn5 transpososome structure bound to target DNA, which has implications for all related transposons and HIV-integration. Due to the fact that several Tn5 transpososome structures have been solved using this method it follows that modified complexes may also be determined. Finally, further experiments must address whether transposon activity can be modulated by simple changes in media and the mechanisms that drive this regulation. It would be interesting to determine how many transposable elements are affected by these measures and how many different factors can affect their activity. Although we presume that this occurs for many systems, no concrete evidence suggests it is a general phenomenon.

Although the Tn5 transposon has proven to be a good model system for studying the regulation of host factors in transposition, it will be worthwhile to discover the depth of host factor regulation of other bacterial elements. Identification of these proteins and their cellular pathways will shed light on the intimate host-transposon relationship that is critical for the mobility of these elements.

6.7 References

1. Mizuuchi,M. and Mizuuchi,K. (1993) Target site selection in transposition of phage Mu. *Cold Spring Harbor symposia on quantitative biology*, **58**, 515-23.
2. Craig,N.L. (1991) Tn7: a target site-specific transposon. *Molecular microbiology*, **5**, 2569-73.
3. Pribil,P.A. and Haniford,D.B. (2003) Target DNA Bending is an Important Specificity Determinant in Target Site Selection in Tn10 Transposition. *Journal of Molecular Biology*, **330**, 247-259.
4. Haapa-Paananen,S., Rita,H. and Savilahti,H. (2002) DNA transposition of bacteriophage Mu. A quantitative analysis of target site selection in vitro. *The Journal of biological chemistry*, **277**, 2843-51.
5. Vigdal,T.J., Kaufman,C.D., Izsvák,Z., Voytas,D.F. and Ivics,Z. (2002) Common physical properties of DNA affecting target site selection of sleeping beauty and other Tc1/mariner transposable elements. *Journal of molecular biology*, **323**, 441-52.
6. Liao,G.C., Rehm,E.J. and Rubin,G.M. (2000) Insertion site preferences of the P transposable element in Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 3347-51.
7. Yanagihara,K. and Mizuuchi,K. (2002) Mismatch-targeted transposition of Mu: a new strategy to map genetic polymorphism. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 11317-21.
8. Lee,G.S., Neiditch,M.B., Sinden,R.R. and Roth,D.B. (2002) Targeted Transposition by the V(D)J Recombinase. *Mol Cell Biol*, **22**, 2068-2077.
9. Pribil,P.A., Wardle,S.J. and Haniford,D.B. (2004) Enhancement and rescue of target capture in Tn10 transposition by site-specific modifications in target DNA. *Mol Micro*, **52**, 1173-86.
10. Kuduvalli,P.N., Rao,J.E. and Craig,N.L. (2001) Target DNA structure plays a critical role in Tn7 transposition. *The EMBO journal*, **20**, 924-32.
11. Goryshin,I.Y., Miller,J. a, Kil,Y.V., Lanzov,V. a and Reznikoff,W.S. (1998) Tn5/IS50 target recognition. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 10716-21.
12. Chaconas,G. and Harshey,R.M. (2002) Transposition of phage Mu DNA. In Craig,N.L., Craigie,R., Gellert,M., Lambowitz,A.M. (eds), *Mobile DNA II*. ASM Press, pp. 384-402.
13. Surette,M.G., Lavoie,B.D. and Chaconas,G. (1989) Action at a distance in Mu DNA transposition: an enhancer-like element is the site of action of supercoiling relief activity by integration host factor (IHF). *The EMBO journal*, **8**, 3483-9.
14. Sakai,J., Chalmers,R.M. and Kleckner,N. (1995) Identification and characterization of a pre-cleavage synaptic complex that is an early intermediate in Tn10 transposition. *The EMBO journal*, **14**, 4374-83.

15. Crellin,P. and Chalmers,R. (2001) Protein-DNA contacts and conformational changes in the Tn10 transpososome during assembly and activation for cleavage. *The EMBO journal*, **20**, 3882-91.
16. Chalmers,R., Guhathakurta,A., Benjamin,H. and Kleckner,N. (1998) IHF modulation of Tn10 transposition: sensory transduction of supercoiling status via a proposed protein/DNA molecular spring. *Cell*, **93**, 897-908.
17. Beall,E.L., Admon,A. and Rio,D.C. (1994) A Drosophila protein homologous to the human p70 Ku autoimmune antigen interacts with the P transposable element inverted repeats. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 12681-5.
18. Bustin,M. (1999) Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Molecular and cellular biology*, **19**, 5237-46.
19. Zayed,H., Izsvák,Z., Khare,D., Heinemann,U. and Ivics,Z. (2003) The DNA-bending protein HMGB1 is a cellular cofactor of Sleeping Beauty transposition. *Nucleic acids research*, **31**, 2313-22.
20. Ivics,Z., Kaufman,C.D., Zayed,H., Miskey,C., Walisko,O. and Izsvák,Z. (2004) The Sleeping Beauty transposable element: evolution, regulation and genetic applications. *Current issues in molecular biology*, **6**, 43-55.
21. Kleckner,N. (1990) Regulation of transposition in bacteria. *Annual review of cell biology*, **6**, 297-327.
22. Nagy,Z. and Chandler,M. (2004) Regulation of transposition in bacteria. *Research in microbiology*, **155**, 387-98.
23. Yin,J.C., Krebs,M.P. and Reznikoff,W.S. (1988) Effect of dam methylation on Tn5 transposition. *Journal of molecular biology*, **199**, 35-45.
24. Kuan,C.T. and Tessman,I. (1991) LexA protein of Escherichia coli represses expression of the Tn5 transposase gene. *Journal of bacteriology*, **173**, 6406-10.
25. Simons,R.W. and Kleckner,N. (1983) Translational control of IS10 transposition. *Cell*, **34**, 683-91.
26. Ma,C. and Simons,R.W. (1990) The IS10 antisense RNA blocks ribosome binding at the transposase translation initiation site. *The EMBO journal*, **9**, 1267-74.
27. Davis,M.A., Simons,R.W. and Kleckner,N. (1985) Tn10 protects itself at two levels from fortuitous activation by external promoters. *Cell*, **43**, 379-87.
28. Schulz,V.P. and Reznikoff,W.S. (1991) Translation initiation of IS50R read-through transcripts. *Journal of molecular biology*, **221**, 65-80.
29. Chandler,M. and Fayet,O. (1993) Translational frameshifting in the control of transposition in bacteria. *Molecular microbiology*, **7**, 497-503.
30. Silva,J.C., Loreto,E.L. and Clark,J.B. (2004) Factors that affect the horizontal transfer of transposable elements. *Current issues in molecular biology*, **6**, 57-71.

31. Ali Azam,T., Iwata,A., Nishimura,A., Ueda,S. and Ishihama,A. (1999) Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *Journal of bacteriology*, **181**, 6361-70.
32. Weinreich,M.D. and Reznikoff,W.S. (1992) Fis plays a role in Tn5 and IS50 transposition. *Journal of bacteriology*, **174**, 4530-7.
33. Twiss,E., Coros,A.M., Tavakoli,N.P. and Derbyshire,K.M. (2005) Transposition is modulated by a diverse set of host factors in Escherichia coli and is stimulated by nutritional stress. *Molecular microbiology*, **57**, 1593-607.
34. Ono,S., Goldberg,M.D., Olsson,T., Esposito,D., Hinton,J.C.D. and Ladbury,J.E. (2005) H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. *The Biochemical journal*, **391**, 203-13.
35. Lucht,J.M., Dersch,P., Kempf,B. and Bremer,E. (1994) Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled proU operon of Escherichia coli. *The Journal of biological chemistry*, **269**, 6578-8.
36. Reif,H.J. and Saedler,H. (1975) IS1 is involved in deletion formation in the gal region of E. coli K12. *Mol Gen Genet*, **137**, 17-28.
37. Kretschmer,P.J. and Cohen,S.N. (1979) Effect of temperature on translocation frequency of the Tn3 element. *Journal of bacteriology*, **139**, 515-9.
38. Haren,L., Bétermier,M., Polard,P. and Chandler,M. (1997) IS911-mediated intramolecular transposition is naturally temperature sensitive. *Molecular microbiology*, **25**, 531-40.
39. Miller,J.H. (1972) Experiments in Molecular Genetics Cold Spring Harbor Laboratory, New York.
40. Ansong,C., Yoon,H., Porwollik,S., Mottaz-Brewer,H., Petritis,B.O., Jaitly,N., Adkins,J.N., McClelland,M., Heffron,F. and Smith,R.D. (2009) Global systems-level analysis of Hfq and SmpB deletion mutants in Salmonella: implications for virulence and global protein translation. *PloS one*, **4**, e4809.
41. Lin,H.-H., Hsu,C.-C., Yang,C.-D., Ju,Y.-W., Chen,Y.-P. and Tseng,C.-P. (2011) Negative effect of glucose on ompA mRNA stability: A potential role of cAMP on the repression of hfq in Escherichia coli. *Journal of bacteriology*.
42. Dong,T. and Schellhorn,H.E. (2009) Control of RpoS in global gene expression of Escherichia coli in minimal media. *Mol Gen Genet*, **281**, 19-33.

APPENDICES

Appendix A: Rights and Permissions

Permission is granted from Oxford University Press and Elsevier to use the following published articles as part of my written thesis for academic purposes without written permission as rights retained by the article authors (see Elsevier and Oxford University Press):

Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2006) Formation, characterization and partial purification of a Tn5 strand transfer complex. *Journal of molecular biology*, **364**, 290-301.

Whitfield,C.R., Wardle,S.J. and Haniford,D.B. (2009) The global bacterial regulator H-NS promotes transpososome formation and transposition in the Tn5 system. *Nucleic acids research*, **37**, 309-21.

[Rights and Permissions for Figure 1.1]

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

Nov 29, 2011

This is a License Agreement between Crystal McLellan ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 2798560388699

License date Nov 29, 2011

Licensed content publisher Nature Publishing Group

Licensed content publication Nature Reviews Microbiology

Licensed content title Mobile genetic elements: the agents of open source evolution

Licensed content author Laura S. Frost, Raphael Leplae, Anne O. Summers, Ariane Toussaint

Licensed content date Sep 1, 2005

Type of Use reuse in a thesis/dissertation

Requestor type academic/educational

Format print and electronic

Portion figures/tables/illustrations

Number of figures/tables/illustrations 1

High-res required no

Figures Figure 1. Transfer of DNA between bacterial cells.

Author of this NPG article no

Your reference number

Title of your thesis / dissertation REGULATION OF THE BACTERIAL TRANSPOSON TN5: TARGET DEFORMATION DRIVES INTEGRATION AND THE *ESCHERICHIA COLI* HOST PROTEINS H-NS AND HFQ MODULATE TRANSPOSITION

Expected completion date Feb 2012

Estimated size (number of pages) 250

Total 0.00 USD

[Rights and Permissions for Figure 1.6B]

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Dec 06, 2011

This is a License Agreement between Crystal McLellan ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 2803361306383

License date Dec 06, 2011

Licensed content publisher John Wiley and Sons

Licensed content publication FEMS Microbiology Reviews

Licensed content title DNA bridging and antibridging: a role for bacterial nucleoid-associated proteins in regulating the expression of laterally acquired genes

Licensed content author Charles J. Dorman,Kelly A. Kane

Licensed content date May 1, 2009

Start page 587

End page 592

Type of use Dissertation/Thesis

Requestor type University/Academic

Format Print and electronic

Portion Figure/table

Number of figures/tables 1

Number of extracts Original Wiley figure/table number(s)

Figure 1 The figure will be adapted in my thesis.

Will you be translating? No

Order reference number

Total 0.00 USD

[Rights and Permissions for Figure 1.7]

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

Nov 29, 2011

This is a License Agreement between Crystal McLellan ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 2798491219053

License date Nov 29, 2011

Licensed content publisher Nature Publishing Group

Licensed content publication Nature Reviews Microbiology

Licensed content title Hfq and its constellation of RNA

Licensed content author Jörg Vogel, Ben F. Luisi

Licensed content date Aug 15, 2011

Type of Use reuse in a thesis/dissertation

Requestor type academic/educational

Format print and electronic

Portion figures/tables/illustrations

Number of figures/tables/illustrations 4

High-res required no

Figures Figures 1 and 2 will be required and possibly figures 3 and 4.

Author of this NPG article no

Your reference number

Title of your thesis / dissertation REGULATION OF THE BACTERIAL TRANSPOSON TN5: TARGET DEFORMATION DRIVES INTEGRATION AND THE *ESCHERICHIA COLI* HOST PROTEINS H-NS AND HFQ MODULATE TRANSPOSITION

Expected completion date Feb 2012

Estimated size (number of pages) 250

Total 0.00 USD

[Rights and Permissions for Appendix E, Figure E1A]

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

Dec 13, 2011

This is a License Agreement between Crystal McLella n ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 2807350663391

License date Dec 13, 2011

Licensed content publisher Nature Publishing Group

Licensed content publication Nature Methods

Licensed content title Photo-cross-linking interacting proteins with a genetically encoded benzophenone

Licensed content author Ian S Farrell, Rebecca Toroney, Jennifer L Hazen, Ryan A Mehl, Jason W Chin

Licensed content date May 1, 2005

Type of Use reuse in a thesis/dissertation

Requestor type academic/educational

Format print and electronic

Portion figures/tables/illustrations

Number of figures/tables/illustrations 1

High-res required no

Figures Figure 1

Author of this NPG article no

Your reference number

Title of your thesis / dissertation REGULATION OF THE BACTERIAL TRANSPOSON TN5: TARGET DEFORMATION DRIVES INTEGRATION AND THE *ESCHERICHIA COLI* HOST PROTEINS H-NS AND HFQ MODULATE TRANSPOSITION

Expected completion date Feb 2012

Estimated size (number of pages) 250

Total 0.00 USD

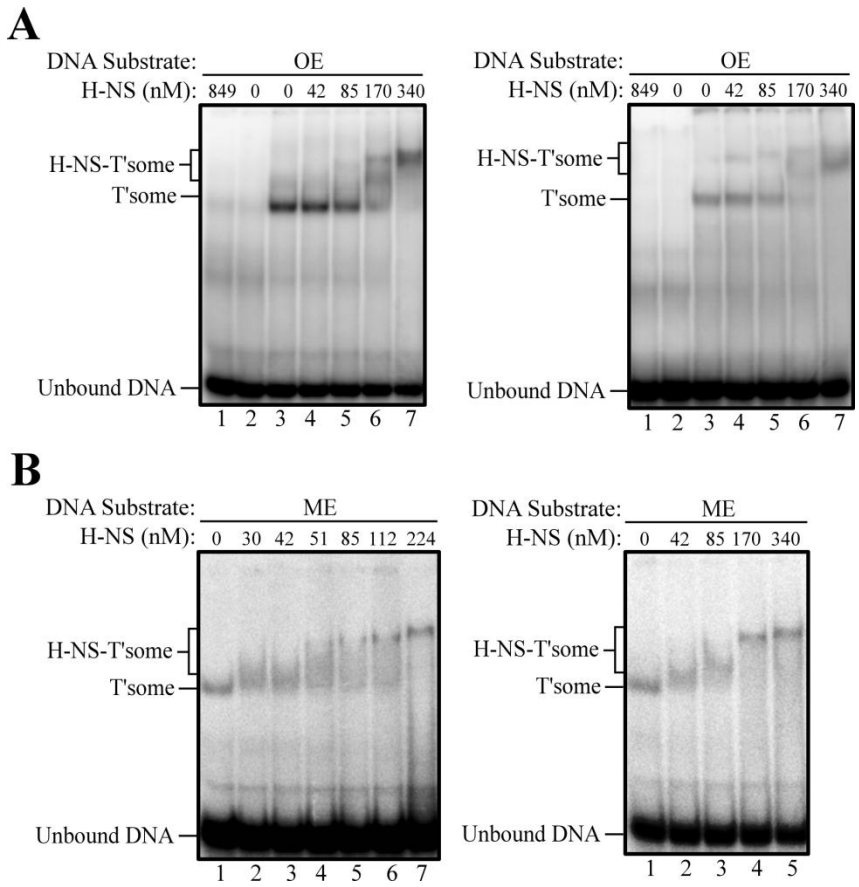
Appendix B:

Table S4.1. *In vivo* transposition of wild-type and mutant mini-Tn5-Kan^R elements in isogenic *hns*+ and Δ*hns* strains with transposase supplied in *trans* (pDH641).

Transposon Source	Strain	Transposition frequency[a]	Normalized frequency[b]	Normalized frequency/pair[c]
pDH626(WT ME)	NK5830F'wild-type	2.1 (±0.9) x10^{-4}	1.0	1.0
pDH626	NK5830F'*hns*::Kan^R	3.7 (±2.1) x10^{-5}	0.18	0.18***
pDH689(OE)	NK5830F'wt	1.2 (±0.7) x10^{-3}	5.7	1.0
pDH689	NK5830F'Δ*hns*	3.2 (±1.2) x10^{-4}	1.5	0.26**
pDH660(ME 8/9)	NK5830F'wt	4.5 (±2.7) x10^{-6}	0.021	1.0
pDH660	NK5830F'Δ*hns*	2.1 (±1.0) x10^{-6}	0.010	0.48*
pDH685(ME 3)	NK5830F'wt	1.2 (±0.6) x10^{-5}	0.057	1.0
pDH685	NK5830F'Δ*hns*	1.3 (±0.8) x10^{-5}	0.062	1.1

a. Relative transposition frequencies were calculated by dividing the number of Sm^{R}Kan^R colonies (transposition events) by the number of Sm^R colonies (total exconjugants) obtained per 0.1 mL of mating mix. Transposition frequencies represent an average value obtained from three independent experiments wherein matings with 3-5 different donor transformants were carried out.
b. Transposition frequencies were normalized to the level of transposition in the WT strain transformed with pDH626.
c. Each pair of transposition frequencies for each transposon was normalized to the WT strain for that specific pair. Asterisks indicate statistical significance for normalized transposition frequencies for each transposon pair as determined using the t-test. *p<0.05, **p<0.005, ***p<0.0001

Appendix C:



Supplementary H-NS binding assays. H-NS and transposase were added simultaneously to (A) OE and (B) ME ^{32}P-labelled transposon end substrate DNA. In each experiment transposase and substrate DNA concentrations were kept constant, while the concentration of H-NS was varied as indicated. For a detailed explanation of the reaction assembly see the 'Materials and methods' section 4.2.3 in Chapter 4. Binding reactions were analyzed by gel electrophoresis on a native 5% polyacrylamide gel and subject to phosphorimager analysis. Percent transpososome complex shifted by H-NS was determined from these gels, along with those presented in Figure 4.2, and data were fit to a quadratic equation as described in 'Materials and methods' section 4.2.3. The fits were used to provide estimates for the observed dissociation constant, obs K_d. Positions of transpososome (T'some), H-NS-shifted transpososome (H-NS-T'some) and unbound substrate DNA are indicated.

Appendix D:

Table S5.1. *In vivo* transposition of a mini-Tn5 element catalyzed in *cis* with IS50 transposase in *hfq*+ and Δ*hfq* strains grown in different conditions.

Substrate	Media[a]	Strain	Transposition Frequency[b]	Normalized Frequency[c]
pDH533-4	M9/LB	DBH33 *hfq*+	$2.9(\pm1.9)\times10^{-5}$	1
		DBH16 *hfq-1*::Ωcat	$1.6(\pm0.9)\times10^{-4}$	5.5*
	LB/LB	*hfq*+	$3.4(\pm0.7)\times10^{-5}$	1
		Δ*hfq*	$4.0(\pm1.0)\times10^{-6}$	0.12*
	M9/M9	*hfq*+	$1.4(\pm0.5)\times10^{-4}$	1
		Δ*hfq*	$8.1(\pm4.2)\times10^{-3}$	58*

a. Media growth (x/y) represents initial transformant growth on agar plates (x) followed by liquid growth (y) during the mating out assay.
b. Transposition frequencies were determined by dividing the number of Sm^RCm^R ('hops') colonies by the number of Sm^R colonies (exconjugants) on plates after mating. Frequencies represent the mean of at least three independent mating out experiments where four individual donor transformants were used.
c. Transposition frequencies were compared to the *hfq*+ strain for each media condition used. Asterisks indicate statistical significance for normalized transposition frequencies for each transposon pair as determined using the t-test. *$p<0.001$

Appendix E: Determining the interacting protein residues between H-NS and transposase by utilizing a photo-cross-linkable residue incorporated into the peptide sequence of Tn5 transposase

As well as understanding the protein-DNA interactions between H-NS and the transpososome required for positive regulation of Tn5 transposition, we also attempted to define the H-NS-transposase interactions. The chemical cross-linking studies (Figure 4.5) and peptide analysis of cross-linked species (Figure 4.6) described in Chapter 4 indicate that H-NS and transposase make protein-protein contacts in the H-NS-bound transpososome. H-NS was recently found to interact with the Tn10 transposase in the context of the transpososome however the details and significance of these interactions are unknown for either system with respect to H-NS function in transposition (1).

Mass spectrometry did not provide a location of the cross-links within the major cross-linked species produced between H-NS and the Tn5 transposase upon addition of the cross-linker EDC. That is, there were no peaks/peptides that were specific to the cross-linked species, and absent from the peptide analysis of H-NS and transposase alone, that indicated a stable link between peptides from both proteins. This may have been due to several reasons. For example, the cross-linked species may have contained too many inter-cross-links between H-NS and transposase, and the MS analysis was not sensitive to detect the low levels of several independent linkages. In addition, too few (less than 50-60%) of the peptides from tryptic digests of the cross-linked species, transposase, and H-NS were detected by the analysis, thus the cross-linked region may also not have been detected. If we had determined the residues involved in the interaction, we could have specifically tested if they were required for H-NS binding to the transpososome *in vitro*

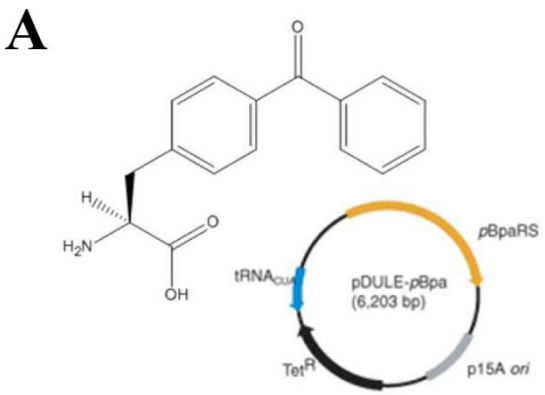
and H-NS regulation of the Tn5 system *in vivo*. This would have been an important step in understanding how H-NS functions as a positive regulator in Tn5 and possibly Tn10 transposition.

We implemented an *in vitro* site-directed photo-cross-linking approach to determine the residues of Tn5 transposase involved in contacting H-NS through protein-protein interactions. The approach we used involved purification of a protein that has the photo-cross-linkable residue, p-benzoyl-L-phenylalanine (pBpa), incorporated into the protein as described previously (2). Briefly, the method relies on the expression of an orthogonal aminoacyl-tRNA synthetase-$tRNA_{CUA}$ pair that incorporates the pBpa residue at an amber codon (UAG) in any gene product expressed in *E. coli*. The aminoacyl-tRNA synthetase and tRNA are expressed from a p15A-based plasmid that must be transformed into the *E. coli* strain along with the plasmid expressing the amber mutant protein of interest. After purification of the protein, mixing with the putative ligand partner and exposure to UV light (pBpa cross-links to C-H groups when exposed to UV light), subsequent SDS-PAGE or Western blotting analysis can detect the cross-linked product. This approach has been successfully used by previous work (3, 4). To be sure this approach could work in our laboratory, we validated that the control protein, histidinol dehydrogenase (HDH), was able to be expressed and purified upon incorporation of pBpa, and form cross-links with other monomers in the presence of UV (Figure E1).

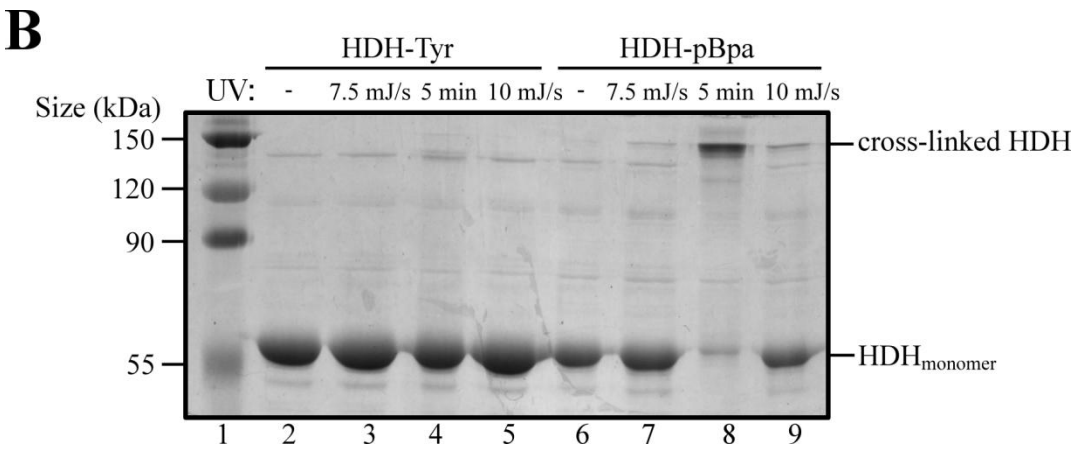
We decided to use transposase for incorporation of pBpa in the directed cross-linking for several reasons. We presumed that we would need to test several positions in order to find an interacting residue with H-NS therefore we needed an easy method of

Figure E1. Establishing a site-directed photo-cross-linkable approach.

(A) Structure of the pBpa residue and schematic of the plasmid that encodes pBpa-specific aminoacyl-tRNA synthetase and $tRNA_{CUA}$. C-H groups within 3Å of the carbonyl oxygen are targets for UV-induced cross-linking. The tRNA synthetase is located between the *lpp* promoter and *rrnB* terminator and is represented by a yellow arrow. The tRNA is located between an *lpp* promoter and *rrnC* terminator and is represented by a blue arrow. pBpaRS indicates the aminoacyl-tRNA synthetase that is necessary for addition of the pBpa residue into the growing polypeptide chain. Reprinted by permission from Macmillan Publishers Ltd: [Nature Methods] (2), copyright (2005). (B) Photo-cross-linking of histidinol dehydrogenase-pBpa. Plasmids that encode a histidine-tagged version of histidinol dehydrogenase (HDH) with an amber codon at position 225 (pTrc-HDH225-TAG) and the pBpa-specific aminoacyl-tRNA synthetase-$tRNA_{CUA}$ pair (pDULE-pBpa) or tyrosine-specific aminoacyl-tRNA synthetase-$tRNA_{CUA}$ pair (pDULE-Tyr) were transformed into *E. coli* (ER2566) and grown according to Farrell *et al.* (2005). An additional strain was transformed with the plasmid pTrc-HDH, which encodes the wild-type histidine-tagged HDH. Strains were grown and HDH was induced as described in Farrell *et al.*, and purified as described previously for a histidine-tagged protein (5). Equivalent amounts of purified proteins were exposed to UV light in clear Eppendorf tubes surrounded by a polystyrene Petri dish to filter out protein-damaging light below 280 nm for the indicated amounts or time as set by the equipment (Stratalinker UV Crosslinker, Stratagene). SDS denaturing load mix was added to the samples in a 1:2 (v/v) ratio and samples were loaded on a 12% SDS protein gel. The gel was subsequently stained with Coomassie Blue to visualize proteins using an AlphaImager (AlphaInnotech). All plasmid listed above were gifts from P. Schultz.



[Nature Publishing Group]



purifying the modified protein and we already had a simple method of purifying transposase using the IMPACT protocol (New England Biolabs). In addition, transposase is a moderately large protein (476 residues) and has several hydrophobic/bulky residues in areas lacking secondary structure we assumed that insertion of the photo-cross-linkable residue would not considerably alter global or local structure. H-NS is one-third the size of transposase and is composed of important α-helices in the dimerization domain and in the DNA-binding motif. We were wary that insertion of the bulky pBpa residue at any position in H-NS would alter the activity of the H-NS protein.

One caveat of this approach is that information must be known about the location of the interaction sites before designing the experiment. In order to determine the positions of transposase that may interact with H-NS we used molecular modeling based on footprinting data (Chapter 3, Figure 3.3) and the co-crystal structure of a Tn5 transposition intermediate, as well as the solution solved structures of the DNA-binding and dimerization domains of H-NS (6). This was done similarly to the modeling presented in Chapter 4 however for this work, previous versions of structures from the H-NS dimerization (7) and DNA-binding domains were used (8). The model used for this work was structured differently compared to those presented in Chapter 4 since it was unknown how the H-NS DNA-binding domain interacts with DNA and a shorter construct was used for determination of the oligomerization domain (Figure E2). Notably this model contains two dimers of H-NS interacting on one transpososome complex.

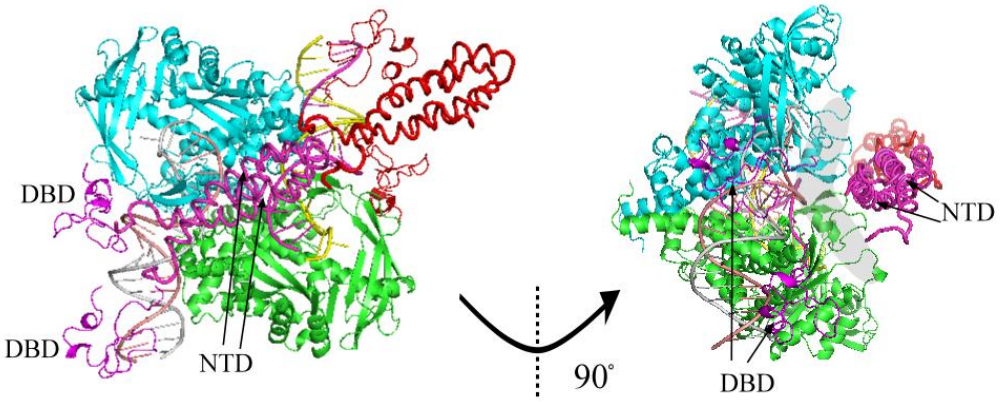


Figure E2. Model of the H-NS-transpososome complex used for determining potential residues of contact between H-NS and transposase.

The model was assembled similarly to that described in Chapter 4 (Materials and methods, section 4.2.8) however different structures were used to represent the two H-NS domains. Residues 1-57 were used to solve the structure of the H-NS N-terminal domain of *E. coli* (PDB: ID 1LR1) (7). Residues 91-137 were used to solve the structure of the H-NS C-terminal domain of *E. coli* (PDB: ID 1HNR) (8). H-NS domains were assembled onto the co-crystal structure of the pre-cleaved Tn5 transpososome (PDB: ID 1MUH) (6). The N-terminal domains of each dimer were manually positioned to place the C-terminal DNA-binding domains close to the DNA regions where footprinting analysis detected additional protected nucleotides when the transpososome was formed in the presence of H-NS. Transposase monomers are represented by aqua and green ribbon molecules. H-NS dimers are represented by red and pink molecules; the N- and C-terminals are indicated. The grey shaded region indicates where residues were chosen to be modified to pBpa for cross-linking. DBD, DNA-binding domain; NTD, N-terminal domain.

Briefly, the following procedure was used for purifying transposase with incorporated pBpa (2). Initially we selected residues that we wished to substitute for pBpa in transposase. The protocol indicated that large, preferably hydrophobic residues should be chosen if they are presumed to be near the interaction. Based on these criteria we selected the following residues: His105, Gln106, Ser116, Ile117, Gln118, Pro154, Arg250, and Phe463 (Figure E2). There were not many hydrophobic side chains in the regions we presumed were interacting with H-NS. Next we utilized site-directed mutagenesis to individually change the corresponding codons to amber stop codons (QuikChange, Stratagene) within the vector used for purification of IPTG-inducible transposase (pGRTYB35) (9). We individually transformed these plasmids, in addition to the plasmid that encodes the synthetase and $tRNA_{CUA}$ for recognition of pBpa and the amber codon (pDULE-pBpa), into the purification strain ER2566 (New England Biolabs). In a control strain, the plasmid that encodes the synthetase-tRNA pair for incorporation of pBpa was substituted with a plasmid encoding the aminoacyl-tRNA synthetase and $tRNA_{CUA}$ pair (pDULE-Tyr) for incorporation of a tyrosine residue at the amber codon. This set-up acts as a control for induction, activity and potential cross-linking of the protein of interest. In two other control strains, the purification vector expressing transposase with or without the amber codon was transformed to test protein expression in wild-type conditions (no protein expression should be detected with the amber mutant in the absence of pDULE-Tyr or pDULE-pBpa). Transformants were selected in rich media (LB) with appropriate antibiotics (ampicillin at 100 mg l^{-1} and tetracycline at 25 mg l^{-1}) and subsequently grown as described in Farrell *et al.* from Step 5 to Step 8 (2005). In the strain where pBpa was to be incorporated into transposase, the

non-natural amino acid was supplemented in the media. We purified transposase protein using a previously described protocol (10). We observed little induction of modified transposase with all the individual transposase constructs containing amber codons in the presence of pBpa by analyzing induced cell extracts compared to non-induced extracts using SDS-PAGE and Coomassie Blue staining (Figure E3).

We altered many aspects of the protocol to remedy this problem. We suspect that toxicity was not an issue because although the strains containing pBpa substrate and the plasmid expressing aminoacyl-tRNA synthetase-$tRNA_{CUA}$ grew slower than the wild-type strain, growth continued past an optical density of 0.6 in the minimal media and when rich media conditions were substituted. Previous work suggested that the nucleotide immediately 3' to the amber codon may prevent expression of the mutant protein by inhibiting suppression of the stop codon (11). Therefore we repeated site-directed mutagenesis in several of our constructs to ensure that a purine was directly downstream of the amber codon, as this had been shown to increase expression of the modified protein. This did not improve the expression of pBpa-transposase. We also increased the volume of media used and altered the level of pBpa residue in the media to enhance protein levels, and determined induction levels over a time-course to determine if the protein was being produced and subsequently degraded over the course of induction. Neither of these avenues improved induction or resulted in better yields of protein. Although over a year was spent attempting to induce protein, we were not able to purify enough modified protein to conduct *in vitro* binding experiments satisfactorily.

We suspect that incorporation of pBpa into transposase caused structural instability. For several amber mutants, induction of transposase was better when the

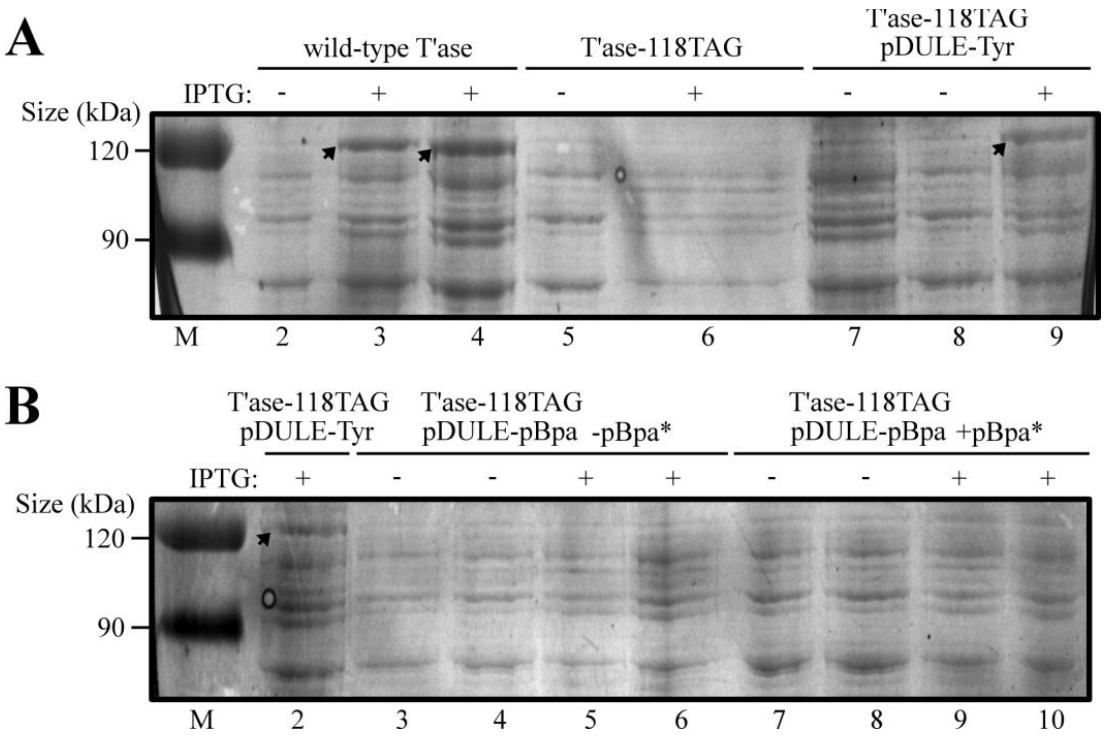
strain was transformed with pDULE-Tyr, which incorporates tyrosine at the amber codon, compared to DULE-pBpa (Figure E3). Induction of modified proteins was typically lower when pDULE-Tyr was used compared to the wild-type conditions or in the presence of pDULE-pBpa, unless pBpa was affecting the tertiary structure of the protein.

While tyrosine is an amino acid with an uncharged polar side chain, the side chain of pBpa is approximately twice as large as that of tyrosine containing two aromatic groups that likely destabilize the overall integrity of the transposase protein. It is possible that the large size and aromatic character of pBpa make unfavourable charge-charge interactions or provide steric clash between nearby side chains, and disrupt the structure or proper folding of the protein, which ultimately becomes degraded.

Although our data suggests that pBpa cannot be structurally accommodated in transposase, it is odd that a variety of amino acid residues (His, Ile, Ser, Gln, Arg, Phe, Pro) found in different locations of transposase would similarly cause structural issues upon substitution of pBpa into the polypeptide. Overall these results likely indicate that proper folding of the individual transposase protein is critical for expression, and the large bulky side chain of pBpa cannot be sterically accommodated in the protein. Possible remedies for this problem include substituting pBpa for a different photo-cross-linkable residue such as p-azido-L-phenylalanine or attempting to incorporate the pBpa residue into H-NS.

Figure E3. Expression of transposase with pBpa or tyrosine incorporated at codon 118.

For a detailed account of growth and induction of proteins please see the text. The only exception is that strains were grown in rich (LB) media and induced for 5 hours once growth reached an optical density of 0.6. Before induction, a small amount of cells (500 µl) were removed from each strain as a control. Cells were gently centrifuged (4000 x g for 4 minutes) and the supernatant was removed before storing pellets on ice. After induction with IPTG (1 mM) and growth, 300 µl of cells were removed from each strain, centrifuged as above and kept on ice. Before SDS-PAGE analysis, pellets were gently thawed, mixed with approximately 20 µl of filtered water and 10 µl of SDS denaturing load mix. Roughly equivalent amounts (approximately 15 µl) were loaded on a 12% SDS protein gel and subsequently stained with Coomassie Blue to determine induction levels. Where two lanes represent IPTG-induced cells, the second lane was loaded with slightly more sample (approximately 20 µl). The induced protein has a considerably slower mobility than the predicted size of transposase (56 kDa) because the purification vector expresses transposase with a chitin-binding domain fused to the C-terminal for ease of purification using the IMPACT system (New England Biolabs). The induced protein is indicated by black arrows. Wild-type T'ase represents the purification vector pGRTYB35. T'ase-118TAG represents the amber mutant purification vector with TAG at codon 118 (the resulting transcript contains UAG). pDULE-Tyr and pDULE-pBpa are described in the text. Note that in (B) media used for growth of cells corresponding to lanes 3 to 6 did not contain the non-natural amino acid pBpa, therefore induction was not expected; in lanes 7-10 pBpa was included in the media used for growth therefore induction was expected.


A
wild-type T'ase
T'ase-118TAG
T'ase-118TAG
pDULE-Tyr
IPTG: - + + - + - - +
Size (kDa)
120
90
M 2 3 4 5 6 7 8 9
B
T'ase-118TAG
pDULE-Tyr
T'ase-118TAG
pDULE-pBpa -pBpa*
T'ase-118TAG
pDULE-pBpa +pBpa*
IPTG: + - - + + - - + +
Size (kDa)
120
90
M 2 3 4 5 6 7 8 9 10

Appendix E: References

1. Wardle,S.J., Chan,A. and Haniford,D.B. (2009) H-NS binds with high affinity to the Tn10 transpososome and promotes transpososome stabilization. *Nucleic acids research*, **37**, 6148-60.
2. Farrell,I.S., Toroney,R., Hazen,J.L., Mehl,R.A. and Chin,J.W. (2005) Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nature methods*, **2**, 377-84.
3. Lee,H.S., Dimla,R.D. and Schultz,P.G. (2009) Protein-DNA photo-crosslinking with a genetically encoded benzophenone-containing amino acid. *Bioorganic & medicinal chemistry letters*, **19**, 5222-4.
4. Mori,H. and Ito,K. (2006) Different modes of SecY-SecA interactions revealed by site-directed in vivo photo-cross-linking. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 16159-64.
5. Ward,C.M., Wardle,S.J., Singh,R.K. and Haniford,D.B. (2007) The global regulator H-NS binds to two distinct classes of sites within the Tn10 transpososome to promote transposition. *Molecular microbiology*, **64**, 1000-13.
6. Davies,D.R., Goryshin,I.Y., Reznikoff,W.S. and Rayment,I. (2000) Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science*, **289**, 77-85.
7. Esposito,D., Petrovic,A., Harris,R., Ono,S., Eccleston,J.F., Mbabaali,A., Haq,I., Higgins,C.F., Hinton,J.C.D., Driscoll,P.C., et al. (2002) H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. *Journal of molecular biology*, **324**, 841-50.
8. Shindo,H., Iwaki,T., Ieda,R., Kurumizaka,H., Ueguchi,C., Mizuno,T., Morikawa,S., Nakamura,H. and Kuboniwa,H. (1995) Solution structure of the DNA binding domain of a nucleoid-associated protein, H-NS, from Escherichia coli. *FEBS letters*, **360**, 125-31.
9. Bhasin,A., Goryshin,I.Y. and Reznikoff,W.S. (1999) Hairpin formation in Tn5 transposition. *The Journal of biological chemistry*, **274**, 37021-9.
10. Peterson,G. and Reznikoff,W.S. (2003) Tn5 transposase active site mutations suggest position of donor backbone DNA in synaptic complex. *The Journal of biological chemistry*, **278**, 1904-9.
11. Miller,J.H. and Albertini,A.M. (1983) Effects of surrounding sequence on the suppression of nonsense codons. *Journal of molecular biology*, **164**, 59-71.

CURRICULUM VITAE

NAME: Crystal R. McLellan

EDUCATION

Doctor of Philosophy, Biochemistry (to be awarded February 2012)
Schulich School of Medicine and Dentistry, University of Western Ontario, Canada
Advisor: Dr. David Haniford

Bachelor of Medical Sciences, Honours Biochemistry, 2004
Schulich School of Medicine and Dentistry, University of Western Ontario, Canada

FELLOWSHIPS, AWARDS, and HONOURS

Natural Sciences & Engineering Research Council of Canada Postgraduate Scholarship - $63,000	2008-2011
Natural Sciences & Engineering Research Council of Canada Alexander Graham Bell Canadian Graduate Scholarship - $17,500	2007-2008
Ontario Graduate Scholarship - $10,000	2007
Schulich Graduate Research Scholarship - $8,400	2007-2010
Western Graduate Research Scholarship - $30,000	2006-2010
Research Assistantship - $33,600	2006-2010
Rossiter Award (for best Honours thesis)	2004
University of Western Ontario Dean's Honour List	2000-2004

TEACHING EXPERIENCE

Teaching Assistant Biological Macromolecules/Biochemistry 3381A	2007-2010
Teaching Assistant Biochemistry and Molecular Biology/Biochemistry 2280A	2006

RESEARCH EXPERIENCE

Research Assistant Department of Biochemistry, UWO, Dr. David Haniford Lab	2004-2005
Research Assistant Department of Biochemistry, UWO, Dr. Christopher Brandl Lab	2004

PUBLICATIONS

Whitfield, C.R., Wardle, S.J., and Haniford, D.B. (2006) Formation, characterization, and partial purification of a Tn5 strand transfer complex. *J Mol Biol.* **364**: 290-301.

Whitfield, C.R., Wardle, S.J., and Haniford, D.B. (2009) The global bacterial regulator H-NS promotes transpososome formation and transposition in the Tn5 system. *Nuc Acids Res.* **37**: 309-21.

Whitfield, C.R., Shilton, B.S., and Haniford, D.B. Regulation of Tn5 transposition by H-NS involves a direct interaction between H-NS and the transpososome. *Mobile DNA* (submitted for publication)