Abstract

Title of dissertation: HUMAN IMMUNODEFICIENCY VIRUS NUCLEOCAPSID

PROTEIN: ANALYSIS OF THE MECHANISM OF STRAND

EXCHANGE AND THE ROLE OF THE ZINC FINGERS IN

NUCLEIC ACID CHAPERONE ACTIVITY.

Megan Joy Heath, Doctor of Philosophy, 2004

Thesis directed by: Professor Jeffrey J. DeStefano
Department of Cell Biology and Molecular Genetics

The human immunodeficiency virus genome is coated by the nucleocapsid protein (NC). NC is a 55 amino acid highly basic protein. It has two zinc fingers that differ by five amino acids. NC contains nucleic acid chaperone activity that aids in the formation of the highly stable nucleic acid structures by destabilizing and preventing the formation of weaker structures. This activity is important for genome dimerization and maturation, tRNA:primer binding site annealing, and many steps in reverse transcription. Annealing experiments were performed with four different RNA structures and complementary DNAs. NC enhanced annealing of all structures showing that NC enhances both unwinding of nucleic acid structure and hybridization of unstructured sequences. NC mutant proteins were used in annealing assays. 1.1 NC had two copies of the first zinc finger, 2.2 NC had two copies of the second zinc finger, and 2.1 NC had both zinc fingers with their positions switched. Experiments showed that all mutants could enhance the

annealing of weakly structured nucleic acids but only 1.1 NC and 2.1 NC enhanced

annealing of strongly structured nucleic acids. Results suggest that finger one is

important for nucleic acid unwinding while finger two plays an accessory role in

annealing. The mechanism of strand exchange, another important aspect of NC

chaperone activity, was also investigated. Experiments were performed using RNA:DNA

hybrids with either the DNA or RNA radioactively labeled. Hybrids were incubated with

different types of RNA acceptor molecules to which the DNA could transfer. The

transfer of the DNA or the displacement of the original donor RNA was monitored.

Experiments showed that optimal enhancement of strand exchange by NC occurred with

acceptors that had more than 22 nucleotides that could anneal to the single stranded

region of the DNA. Also, experiments with acceptors that had point mutations showed

that the region of the acceptor that binds to the single stranded region of the DNA should

be completely complementary for optimal NC stimulation. These results indicate the

annealing of the acceptor and DNA outside the donor:DNA hybrid region can be

important initiation step for NC enhanced strand exchange.

HUMAN IMMUNODEFICIENCY VIRUS NUCLEOCAPSID PROTEIN: ANALYSIS

OF THE MECHANISM OF STRAND EXCHANGE AND THE ROLE OF THE ZINC

FINGERS IN NUCLEIC ACID

CHAPERONE ACTIVITY

by

Megan Joy Heath

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2004

Advisory Committee:

Dr. Jeffrey J. DeStefano, Chair
Dr. James Culver
Dr. Siba K. Samal
Dr. Daniel C. Stein
Dr. Barbara Thorne

Dedication

This dissertation is dedicated to my Grandfather, John Patrick Fitzgerald, for instilling in

his family the importance of education and for being a constant source of encouragement

to me throughout my educational pursuits.

Acknowledgements

I would like to start by thanking the members of my committee. Over the past few years and committee meetings I have seen each member contribute specifically to my education at Maryland. I would like to thank Dr. Siba Samal for teaching me the importance of having a broad knowledge within the field of virology and not narrowing my focus to retroviruses. I thank Dr. Barbara Thorne for challenging my writing capabilities and encouraging me to continue to develop my writing skills. I know my dissertation is better because of this. I appreciate Dr. Dan Stein's ability to put my projects in a different perspective and think of alternative ways to try experiments such that the best possible research is performed. I thank Dr. James Culver, for pushing me at every meeting to explain why my research is useful to the field of retrovirology and how it is going to impact this field. It was vital that I ask that question of myself and my research continually to pursue knowledge that could enhance our chances of finding a better drug therapy for those suffering with AIDs. I also would like to thank Dr. Rick Stewart, who is not on my committee, but spent many hours teaching me how to use the Fluoromax 2 and examining my data.

I need to express unending gratitude to the last member of my committee, my advisor, Dr. Jeffrey DeStefano. Many times I have talked with incoming students and told them that I ended up with the best advisor I could imagine. I have learned so much from my advisor, and he has helped to shape my professional outlook. The one thing that stands out the most in my mind is that every single time I was ready to give up on part of my project he would challenge me with trying one last thing. Inevitably it would always

work and would bear much fruit. I only hope I can remember this attitude of perseverance as long as I work in the research field.

Many thanks to Nancy and Lori who have helped with so many questions and tasks. None of us would ever graduate without you. Also, to Jeanie, Claudine, Karen, Loraine, Elizabeth, whose help was also vital to my success. Thank you to Dr. Todd Cooke, Dr. Elizabeth Gantt, and Dr. Albert Ides for guiding me through the necessary steps to obtain my degree.

I am so thankful to the members of my lab for keeping me motivated when things would not work properly for months at a time. To Aarti, Jason and Bill for taking me under their wing and teaching most of the techniques I would use in my research. To Jason for remaining my friend even though we see eye to eye on almost nothing. To Bill for being a true friend, and a great coworker, ordering anything at a moments notice and always being willing to lend a hand. You are one of the most altruistic people I have ever met. Niru thanks for keeping us all laughing. Reeshma, thank you for remembering every important date possible. Aminata, and Bridget thank you for keeping our dishes so clean. And also to Aminata and Titi for keeping things light-hearted in the lab when the graduate students were too stressed about their work. I have the utmost gratitude for Suchi. We met my first day at Maryland and we shared in this journey together. I don't think I could've done this without you. You made bad days good and good days great. I admire your work as a scientist and look forward to seeing all that you will accomplish. Also, the many other graduate students who I got a chance to commiserate and laugh with, Nate, Mike, Jamie, Liliana, Beth, Nandini, Sabrina, Rob, Jason, and many others. Thank you for being such a great team!

Halfway through graduate school I was blessed to marry into a wonderful family. I am so thankful for my in-laws, Rich and Judy for encouraging me to remember my spiritual growth and not get bogged down by discouraging seasons in my life. I am so thankful to Ben and Rachel for countless meals and Friday night games. I also appreciate Sarah accepting me as a sister right from the start. To Jeff's many cousins and aunts and uncles, thank you for your encouragement and interest in my pursuits. Let's go camping! Thank you Grandma Nettie, Grandpa Charlie and Grandma Betty for your genuine interest in my education. I also have many friends from church and my specifically my home group who I need to thank for their constant prayers and encouragement.

Lastly I need to express gratitude to the person I am convinced I would not be graduating without. To Jeffrey, my loving husband. You are one of a kind. You never doubted my work, and that your wife would soon have her PhD. I am certain that I would've been too afraid and would have bailed out long ago, if I didn't have your support. I am eternally grateful to you. Even more than your support I am just so thankful that I have you to share this life with. I know that I will accomplish many things because you believe in me and you won't let me back down. I hope you know I believe in you in the same way and I can't wait to see the many things you achieve. I thank God for you every day, you are my greatest gift.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIDS: Acquired immunodeficiency syndrome

ARV: AIDS associated retrovirus

BSA: Bovine Serum Albumin

CIP: Calf intestinal alkaline phosphatase

CNBPs: Cellular nucleic acid binding proteins

CRF: Circulating recombinant form

DABCYL: 4-[[(4-dimethylamino)phenyl]-azo] benzenesulfonicamino

DIS: Dimerization initiation site

DNA: Deoxyribonucleic acid

DTT: Dithiothreitol

EIAV: Equine infectious anemia virus

Env: Envelope

ER: Endoplasmic reticulum

FAM: Fluorescein-6-carboxamidohexyl

FRET: Fluorescence resonance energy transfer

Gag: Group antigen

HIV: Human immunodeficiency virus

HIV-1: Human immunodeficiency virus Type 1

HIV-2: Human immunodeficiency virus Type 2

HSV: Harvey sarcoma virus

IN: Integrase

L-domain: Late domain

MA: Matrix protein

mRNA: messenger RNA

MPMV: Mason-Phizer monkey virus

MuLV: Moloney murine leukemia virus

NC: Nucleocapsid protein

Nef: negative factor

PBS: Primer binding site

PIC: Preintegration complex

PK: Proteinase K

PNK: Polynucleotide kinase

Pol: Polymerase

PPT: Polypurine tract

PR: Protease

Ψ-site: Psi site

R: Repeat region

RNA: Ribonucleic acid

RRE: Rev response element

RT: Reverse transcriptase

SELEX: Systemic evolution of ligands by exponential enrichment

SIV: Simian immunodeficiency virus

SL1: Stem loop 1

SL2: Stem loop 2

SL3: Stem loop 3

SL4: Stem loop 4

SNV: Spleen necrosis virus

-ssDNA: Minus strand strong stop DNA

+ssDNA: Plus strand strong stop DNA

SU: Surface protein

TAR: *Trans*-activation region

Tat: Transactivator of transcription

TM: Transmembrane protein

tRNA: Transfer RNA

U3: Unique 3' region

U5: Unique 5' region

USA: United States of America

Vif: Viral infectivity factor

Vpr: Viral Protein R

Vpu: Viral Protein U

Chapter 1

General Introduction

Acquired Immune Deficiency Syndrome (AIDS) is caused by infection with the human immunodeficiency virus (HIV). Since its emergence in 1981 nearly 18 million people have died from opportunistic infections because their immune systems were compromised from AIDS. This epidemic is particularly threatening in developing countries like those in sub-Saharan Africa where AIDS education and prevention programs are lacking. In 2003 alone Sub-Saharan Africa lost between 2.5 and 3.5 million people to AIDS (135). In the same year there were between 25 and 28 million living with HIV/AIDS in this region. This is in stark contrast to the prevalence of AIDS in high income countries like the United States of America (USA) where in 2003 there were 12,000-18,000 deaths from AIDS and 36,000-54,000 newly infected individuals (135). Developing countries contain 12 % of the world's population, and 95 % of the new cases of HIV occur in these countries (125). These numbers show that AIDS education, prevention, and quality antiretroviral treatment can minimize the number of people lost each year to this horrific disease. In the last 21 years HIV has spread to nearly every country and has infected all age groups. In fact 11 % of people living with HIV in the USA are over the age of 50 (125). It is clear that AIDS must be an area of intense research such that this global epidemic can be halted and lives can be saved.

HIV is a member of the family *Retroviridae* (22). There are two types of this virus, HIV-1 and HIV-2. HIV-1 is the strain responsible for the majority of AIDS cases worldwide, though in recent years the prevalence of HIV-2 has grown (125). HIV-1 is

further broken down into three groups, M for the main HIV genotypes, O for the types

with genomes that vary substantially from the M group, and N for a rare virus that is

neither M or O. The M group contains 11 subtypes (A-K) which vary in their group-

specific antigen (Gag) and envelope (Env) protein sequences. The prevalence of the

various subtypes is different throughout the world. In the USA the predominant subtype

is B whereas in Africa subtypes A, C, and D are predominant (125). The many different

subtypes of HIV make production of successful vaccines and drug therapies difficult. For

this reason it is important to know as much as possible about this virus and its life cycle

such that drug therapies and vaccines can be created to target numerous subtypes.

The HIV Life Cycle

The life cycle of HIV is shown in Figure 1. HIV infects CD4+ cells. A number

of CD4+ cell lines have been infected with HIV *in vitro*, but *in vivo* the cells primarily

infected with HIV are CD4+ T lymphocytes and macrophage-lineage cells (22).

Infection begins when the viral protein gp120 (SU) associates with CD4 on the cell

membrane. However, while this association occurs with high affinity, it is not sufficient

for HIV infection. A co-receptor also associates with gp120 and CD4. There are two

primary co-receptors, CXCR4 in T cells, and CCR5 in macrophages (22). The co-

receptors are members of a family of proteins that contain nine membrane spanning

domains and normally function as chemokine receptors (22). After a stable complex is

formed between gp120, CD4 and the co-receptor, the viral and cellular phospholipid

membranes fuse. This fusion allows the entry of the virion core into the cell. After entry

of the core the process of reverse transcription begins. Inside the core the two copies of
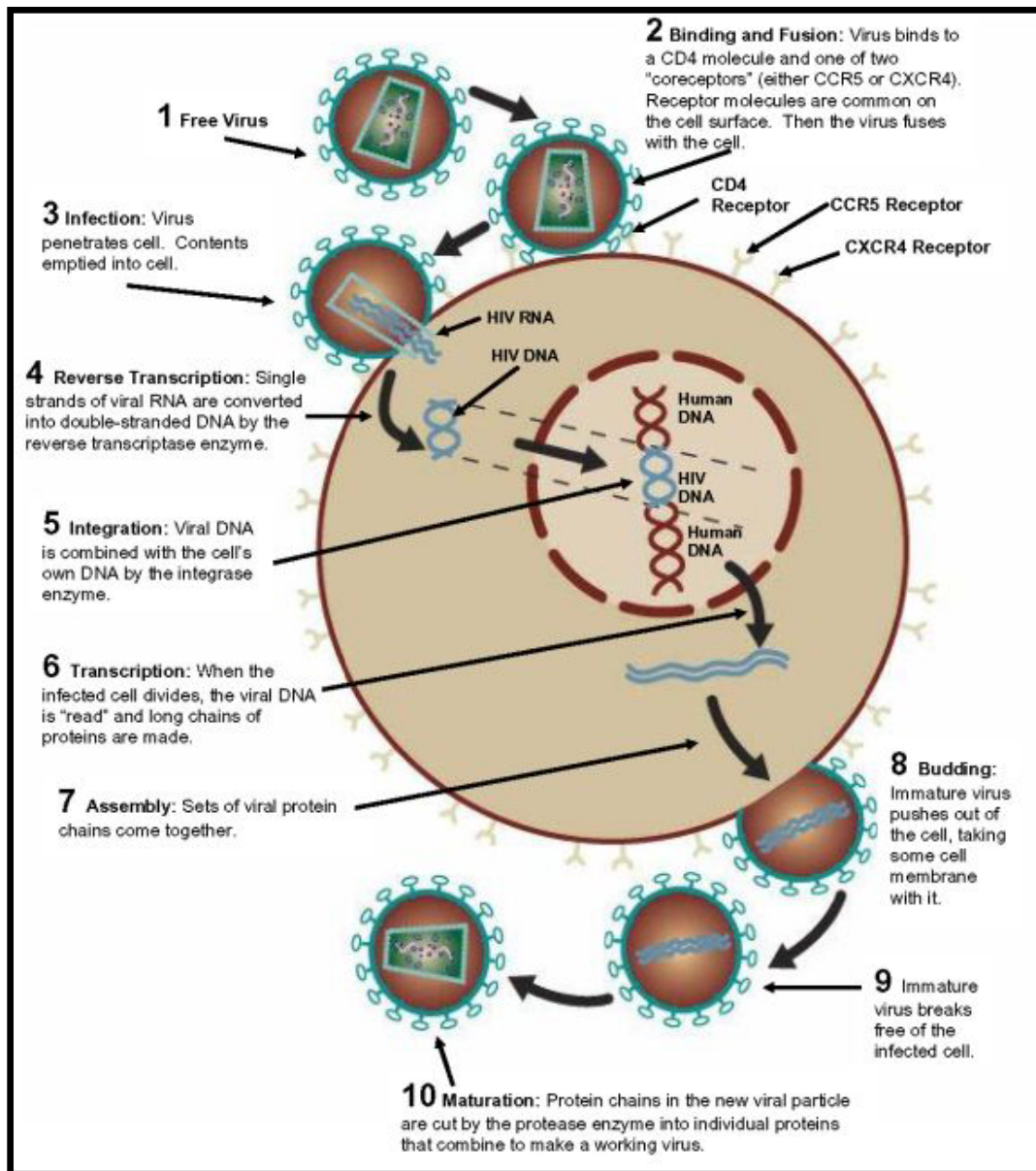
Figure 1: HIV life cycle. The life cycle is depicted beginning with the attachment of free virus to the cell membrane. Infection continues with the release of the core into the cell and reverse transcription. The double stranded DNA genome then integrates into the host genome and cellular transcription machinery synthesizes mRNAs used as the new RNA genomes and to synthesize viral proteins. Viral components assemble at the membrane where budding occurs. Virus maturation takes place after budding. Figure obtained from New Mexico Aids Training Center, The University of New Mexico Health Sciences Center.
http://www.aidsinfonet.org/articles.php?articleID=106

the RNA genome form a dimer through interactions at the 5' ends. Reverse transcription is thoroughly described later, and results in the replication of the dimeric RNA genome into a double stranded DNA. This DNA is then transported to the nucleus as part of the preintegration complex (PIC). The PIC is a large complex containing the DNA, integrase (IN), and potentially a number of other viral and cellular proteins. The PIC migrates into the nucleus through the nuclear pore. In the nucleus IN inserts the DNA into the host genome.

The DNA provirus is then transcribed by host transcriptional machinery. The U3 region of the provirus contains a number of *cis*-acting sequences important for the initiation of transcription. The cellular RNA polymerase transcribes the provirus into mRNAs which can serve two purposes: either as a template for translation of the new viral proteins, or as the new RNA genome to be packaged into progeny virions. HIV-1 is a complex retrovirus which has a number of viral proteins that are transcribed from singly or multiply spliced mRNAs. The functions of each of these proteins are discussed further in the next section. However, it is important to note that the Tat, Rev, and Nef proteins are synthesized from multiply spliced mRNAs while Vif, Vpr, Vpu and Env are translated from singly spliced mRNAs. Gag and Pol gene products are produced from unspliced mRNAs that also serve as new genomes. The Tat protein is transcribed early because it is an important activator for transcription of the longer, unspliced viral mRNAs. The Rev protein is also transcribed early because this protein binds to the Rev-response element (RRE) found in singly and unspliced viral mRNAs. This association is necessary for the export of these longer mRNAs from the nucleus to the cytoplasm. Each

mRNA is transported to the cytoplasm where it either undergoes translation or, in the case of unspliced mRNA, is packaged into the virion (22, 49).

After the viral structural proteins have been translated in the cytoplasm and the viral genomes have been synthesized these components associate at the cell membrane. The genomic RNA is packaged by association with the nucleocapsid protein (NC) portion of the Gag precursor protein. The HIV genome packaging signal, referred to as the Psi-site (Ψ–site) consist of four hairpins which are found near the 5' end of the genome. Also, the dimerization initiation sequence (DIS), found in stem loop one (SL1) of the Ψ-site contains a palindrome sequence that enables the dimerization of the two RNA genomes. This is important for efficient packaging of the genomes.

The envelope proteins are translated in the endoplasimic reticulum (ER) and then transported through the ER and golgi apparatus where they are glycosolated and incorporated into the phospholipid membrane. The Gag molecules associate with the cell membrane, where the Env molecules are present and budding begins. Sequences in the matrix (MA) region of Gag have been shown to be important for association with Env proteins and the cell membrane (22). The virions bud from the membrane of the cell and the virion core is encased in the phospholipid membrane. After budding the viral protease (PR) cleaves the Gag and Gag-Pol proteins and virion core undergoes maturation. During maturation the viral genomes condense and the core takes on a cone-like structure. The progeny virions are then able to infect new CD4+ cells. The mature virus particle is shown in Figure 2.
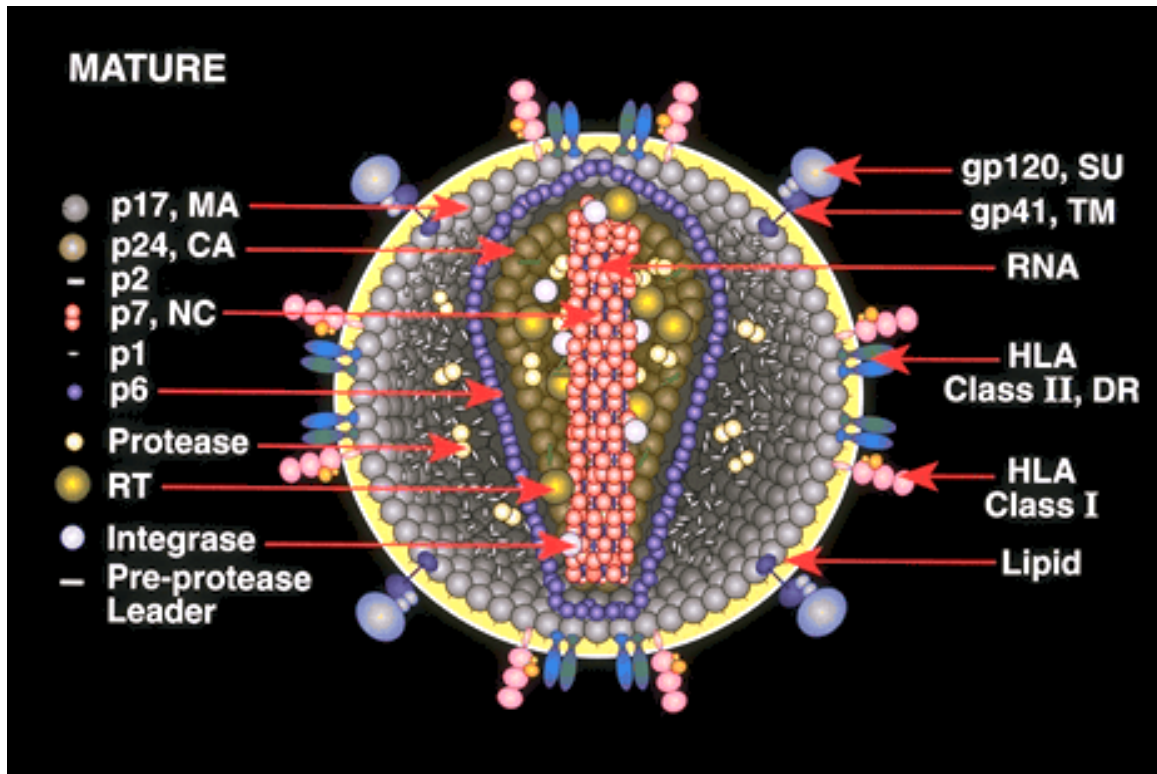
Figure 2: Structure of HIV. The structure of a mature HIV particle is shown. The viral envelope is composed of a phospholipid bilayer associated with transmembrane and peripheral envelope proteins. The core of HIV packs into a cone shape due to the interactions between capsid molecules and the condensing of genomic RNA. The genome is a dimeric RNA which is coated with nucleocapsid protein molecules. Within the core are the three enzymes, protease, reverse transcriptase, and integrase. Image obtained from AIDS Research Information Center, ARIC's AIDS Medical Information Library. http://www.critpath.org/aric/library/img002.htm

HIV genome structure and proteins

The HIV genome codes for structural, enzymatic and accessory proteins. The genome organization in the proviral DNA is shown in Figure 3. This figure also illustrates some of the key nucleic acid segments in the genome.

The Gag precursor protein contains the HIV structural proteins. These include the MA, capsid and NC proteins (22, 49). These proteins each provide the virus with the proper structure for infection and replication. However, the structural proteins can have other functions as well. The MA is structural protein which associates with the lipid bilayer. Also, there is evidence for interactions between MA and the envelope proteins (22). MA is myristilated near the N-terminal end and the hydrophobic myristate is thought to insert into the lipid bilayer stabilizing Gag binding to the membrane (22). Adjacent to the MA sequences in Gag are the amino acid sequences that make up the capsid protein. Capsid-capsid interactions cause a tight virion core to be formed in the shape of a cone. This occurs after the protease cleavage and the maturation of the viral genome. Maturation of the genome is enhanced by NC. This protein is a RNA chaperone which enhances the correct folding of the genome such that it can pack tightly into the viral core (110). The last protein in the Gag precursor that has gained some recent recognition is the p6 peptide. This peptide is found at the C-terminal end of Gag and has been shown to be important for budding from the cellular membrane. The p6 protein contains the HIV late (L) domain. Some L domain deletion mutants do not bud properly and remain tethered to the membrane (22).

HIV-1 Genome Organization

Figure 3: HIV genome organization. The genome organization is shown as found in the proviral DNA. Rev, Tat, and Nef proteins are synthesized from multiply spliced mRNAs. The accessory proteins, Vif, Vpu and Vpr, and envelope proteins are synthesized from singly spliced mRNAs. Gag and Gag-Pol proteins are synthesized from unspliced RNAs. Also, the location of important RNA structures, such as TAR and RRE are shown. The p9 protein in this figure represents NC. NC is further cleaved to just under 7 kilodaltons. The p7 peptide in this image represents a protein that is cleaved again to p6. Image obtained from http://www.blc.arizona.edu/Marty/429/Lectures/Figures/HIV1_genome.GIF

Due to a frameshift during translation of unspliced HIV mRNA about ten percent of the Gag molecules are synthesized as Gag-Pol precursors (121). The polymerase (pol) region encodes the PR, reverse transcriptase (RT) and IN enzymes. PR cleaves the Gag and Gag-Pol molecules into their mature proteins. PR is thought to have protease activity in both the precursor and mature forms because it must cleave itself (or neighboring protease molecules) when it is in the precursor form. RT is responsible for the replication of the genomic RNA into the double stranded DNA. RT functions as a dimer, with a p66 and p51 subunit (22). This enzyme has RNA and DNA dependant polymerase activity. It also has a RNase H region which will cleave RNA that is part of a RNA:DNA hybrid. This activity is important for the successful replication of the dimeric RNA genomes into the proviral DNA. After the double stranded DNA genome is synthesized it is integrated into the host genome by IN.

There are two auxiliary proteins, Tat and Rev, which serve regulatory functions in the HIV life cycle. That Tat protein is found mainly in the nucleus and is important for transcription of the provirus DNA. This protein associates with the trans-activation region (TAR) located near the 5' end of the viral RNA. Tat also associates with the cellular cyclin protein. These associations enhance elongation 100 fold during transcription of viral sequences (49). The Rev protein associates with the RRE, which is a sequence found in the *env* portion of the genome. An arginine rich region of Rev binds to a high-affinity binding sight in the RRE. Rev also contains a leucine-rich nuclear export sequence on its C-terminal. This is recognized by cellular export proteins which then export Rev and the transcript it is bound to from the nucleus (49). Overall Rev serves to enhance the export of singly and unspliced mRNA. This leads to an increase in

9

the production of viral structural and enzymatic proteins and a decrease in Tat, Rev, and Nef which are translated from multiply spliced mRNAs.

As was previously mentioned, HIV is a complex retrovirus containing a number of accessory proteins. The Nef (for negative factor) protein helps the virus to evade recognition by the immune system. It decreases the surface expression of the major histocompatibility class I molecules and down-regulates CD4 expression (22, 49). Nef is also important for the pathogenesis of HIV. Viral infectivity factor (Vif) has sparked much interest in recent years. Vif enhances viral DNA synthesis during rounds of infection subsequent to Vif's inclusion in the virion core. Vif can cause viral growth in previously non-permissive cell lines. Vif also protects the viral geneome from the deamination activity of a cellular protein, APOBEC3G, by inducing ubiquitination of this protein, marking it for degradation (23, 76, 88, 91, 148). In the absence of Vif nucleotide deamination leads to high mutation levels that severely hamper viral replication. This is currently an area of extensive study. The viral protein U (Vpu) is found only in HIV and one SIV strain. Vpu is a hydrophobic integral protein that forms oligomeric complexes. When Vpu is deleted from the virus multiple cores are found in progeny virions and these virions are found in the intracellular compartments rather than exterior to the cell (49). Also, when phosphorylated, Vpu enhances the degradation of CD4 in the ER. Viral protein R (Vpr) has been shown to be a part of the PIC and also has been shown to cause the infected cell to arrest in G_2 of the cell cycle (49). Vpr is packaged into the virion at a concentration similar to the Gag molecules. Vpr has also been shown to have some transcriptional activation activity (22).

The exterior of the virus is made up of a phospholipid bilayer membrane. The two envelope proteins are found associated with this membrane. The transmembrane protein (TM) is the C-terminal portion of the envelope protein. The surface protein (SU) is the smaller protein, derived from the N-terminal portion of the envelope protein. TM is an integral protein while SU forms non-covalent interactions and disulfide bonds with TM to associate with the peripheral of the viral membrane (49). These proteins are glycosolated and interact with CD4 and the coreceptor on the cell membrane to initiate infection.

Reverse Transcription

The HIV replication process, reverse transcription, is shown in Figure 4. During reverse transcription the single stranded RNA genome is copied into a double stranded DNA provirus (21, 22, 49). The HIV genome consists of dimeric plus stranded RNA. Each RNA genome is approximately 9.1 kb. At the onset of reverse transcription host cell tRNA^{Lys3} binds to the primer binding site (PBS). The tRNA is used as a primer for minus strand DNA synthesis. RT is a RNA-dependant, and DNA-dependant DNA polymerase. Using the tRNA as a primer it begins to synthesize DNA on the RNA template. As the DNA is being synthesized, RT's RNase H region cleaves the previously copied regions of the RNA template from the newly synthesized DNA. DNA synthesis continues until RT reaches the 5' end of the RNA genome. At this point the first of two obligatory strand transfers must occur for synthesis to continue.
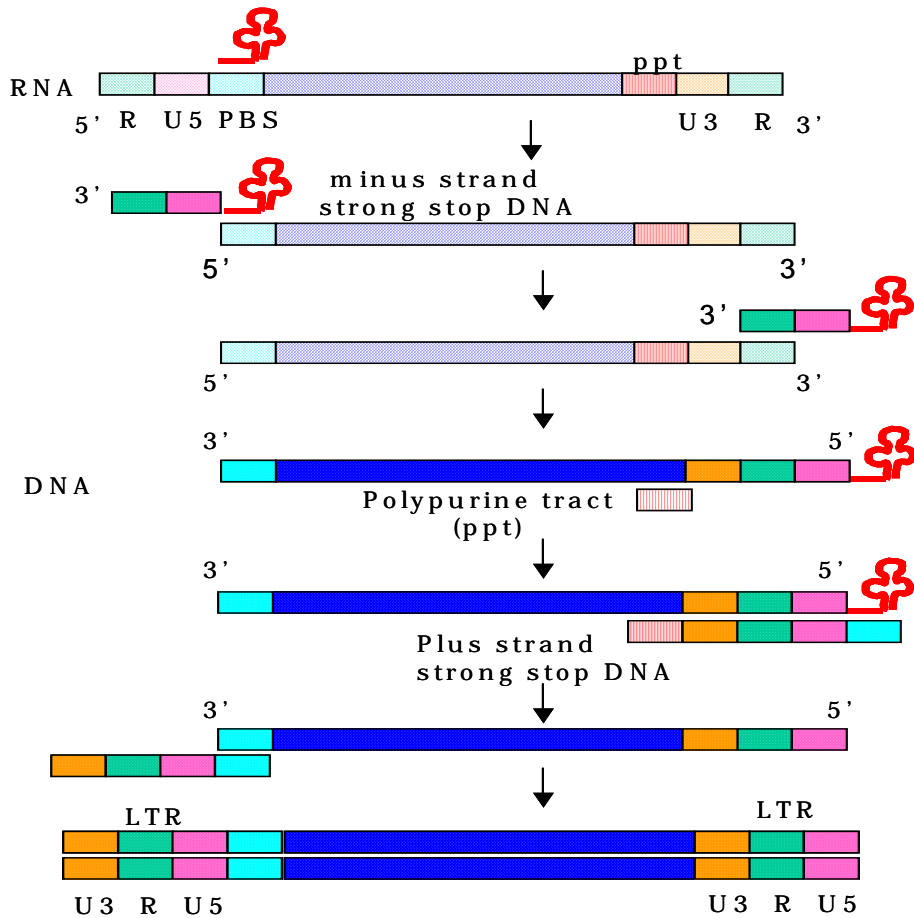
Figure 4: HIV reverse transcription. The replication process of HIV is shown. This begins with the binding of the tRNA to the primer binding site on the RNA genome. RT begins DNA synthesis and continues to the 5' end where minus strand DNA transfer occurs to the 3' end of the RNA template. RT RNase H cleaves the RNA after it is copied but does not cleave the polypurine tract. The polypurine tract serves as the primer for plus strand DNA synthesis. Plus strand synthesis continues to the 5' end of the minus strand DNA where it transfers to the primer binding site sequences on the minus strand DNA. Circulization of the genome occurs so that synthesis can continue with each DNA using the other as a template. The double stranded DNA that is synthesized has two copies of the U3, R and U5 sequences. Image obtained from

http://www.tulane.edu/~dmsander/WWW/335/335Replication.html

On each end of the RNA genome there are repeat (R) sequences. These sequences provide the necessary homology for strand transfer to occur. When the DNA has been synthesized to the 5' end of the genome it is then transferred to either the 3' end of the same genome or the 3' end of the other RNA genome in the virion core. This provides a template for continued synthesis. As DNA synthesis continues the RNase H region of RT continues to cleave the RNA template. However the polypurine tract (PPT) is initially resistant to RNase H activity. Therefore this sequence of RNA remains bound to the minus strand DNA and can then serve as the primer for plus strand DNA synthesis.

Plus strand synthesis is primed from the PPT and continues towards the 5' end of the minus strand DNA. After the PBS is copied from the tRNA on the 5' end of the minus strand DNA the tRNA is removed by RNase H. This provides a region of homology for the second obligatory strand transfer. Plus strand transfer occurs when this homologous region of the plus strand binds to the PBS causing circulization of the templates and synthesis continues near the 3' end of the minus strand DNA. Synthesis then continues with each DNA using the ends of the opposing DNA as a template. It is important to note that throughout both minus and plus strand transfer recombination can occur. Reverse transcription is complete when both the minus and plus strands have been synthesized and contain the 3' unique (U3), R , and the 5' unique (U5) regions on each end. This double stranded DNA is then integrated into the host genome and cellular machinery will be used to synthesize new genomes.

Recombination

There are currently 11 HIV-1 subtypes in group M. The variety between the subtypes makes the design of successful HIV-1 drug therapy or a successful vaccine difficult. However, this problem is heightened when recombination between and within subtypes is considered. Because one retroviral RNA is thought to be enough to complete synthesis of viral DNA, recombination may not be vital for replication of retroviruses (75). However, recombination is a major cause of the genetic diversity that occurs between quasispecies within an individual infected with HIV (131). Recombination occurs when a cell is infected with two viruses that differ genetically and each make a provirus that integrates into the cell genome. These proviruses can make new RNA genomes. It was initially thought that recombination could occur when the two different viruses infect the cell, leading to the creation of one recombinant provirus. However, evidence has been generated that shows that there are two different proviruses made which lead to the packaging of two different RNA genomes in one virion, and recombination occurs when the provirus is made from these genomes (70). When two different genomes (one from each provirus) are packaged there is a potential for recombination creating a provirus with sequences from each of the different genomes. Recombination occurs by the process of strand transfer, also called template switching. Strand transfer occurs when the DNA being synthesized transfers from one template to another and continues synthesis, thereby incorporating genetic sequences from two separate templates.

Recombination is an important process for the persistence of HIV and poses a problem in drug development. Mutations have been found that give resistance to

compounds considered for drug treatment of HIV (20). Kellam and Larder showed that when cells infected with HIV were treated with zidovudine, a reverse transcriptase inhibitor, recombinant forms of the virus arose that were resistant to zidovudine (77). It is thought that in the current virus population there is a strain that exists that is resistant to each potential drug treatment (20). However, these strains are not initially prevalent because the mutations are slightly detrimental to replication. Only when there is an advantage for the replication of these viruses do these recombinant forms carry on (20).

The other important effect that recombination can have is to hamper vaccine development. The variation in the HIV genome creates many difficulties for vaccine development and recombination only adds to this difficulty because it leads to increased variation (85). Data has shown recombination between different subtypes in infected individuals. Analysis of genomic sequences from eight subtypes showed that they had all been involved in recombination events (112). Also phylogenic analysis has shown that recombination has occurred between subtypes B and F in Brazil and between subtypes A and E in Thailand (52, 117). In fact the recombinant form of A and E is the virus strain which is predominantly infecting individuals in Thailand (52). Recombination has lead to the persistence of circulating recombinant forms (CRF) of HIV which provide another hurdle in the development of a globally effective HIV vaccine (107). Virus pools analyzed from a premature infant that was infected by two different subtype B strains (due to blood transfusions with infected blood) showed recombination had occurred between the strains. These results show that dual infection can occur and that recombination *in vivo* leads to the generation of quasispecies within an infected individual (40). Quasispecies are species of viruses that are very similar but have genetic

differences due to mutations that occurs within an individual. Also simian immunodeficiency virus (SIV) recombination has been shown to occur in a rhesus monkey infected with two different SIV strains. Recombination was shown by 2 weeks post infection indicating that recombination is a quick process that will occur with ease when it is favorable to virus survival (141).

Recombination has been thought to occur by two different mechanisms: strand displacement assimilation, and the copy choice mechanism (21, 70). Strand displacement assimilation states that recombination occurs when the plus strand DNA is being synthesized. As one plus strand is being synthesized it displaces another plus strand which is also being synthesized on the minus strand DNA. It must do this by switching from one region of the minus strand to another region before continuing synthesis. This type of recombination has been shown to occur in a single round of replication using spleen necrosis virus (SNV), a C-type retrovirus. Hu and Temin designed viral constructs that contained all the *cis*-acting sequences needed for replication but no coding sequences for viral proteins. They used two different constructs that contained resistance for hygromycin phosphotransferase B (*hygro*) and neomycin (*neo*). However, in one construct *hygro* was disabled by a frameshift mutation and in the other *neo* was disabled. These constructs were used to infect cells and a single round of replication was able to proceed. Because no viral proteins were made replication was limited to one round. The presence of recombinants was found by selection using both neomycin and hygromycin. Restriction enzymes were used to evaluate the sequences of the recombinant viruses and differentiate between minus and plus strand transfer. It was shown that both plus strand

and minus strand transfer occurred and the strand displacement assimilation model was implicated for plus strand transfer (70).

These experiments and others have also shown that recombination can occur during the synthesis of the minus strand by a copy-choice mechanism. This mechanism was introduced as forced copy choice stating that when reverse transcriptase reaches a break in the template RNA while synthesizing minus strand DNA the DNA will transfer to either another region on that template or the other RNA template and continue synthesis (21). This model has been modified from forced copy choice to copy choice because these transfers do not always occur at breaks in the template. Many *in vitro* experiments have demonstrated copy-choice recombination as well (7, 13, 32-34, 94, 98, 99, 113). Overall, experiments indicate the copy-choice type recombinations occurring during minus strand synthesis are the predominate mode of recombination.

The recombination frequency has been determined for a number of retroviruses. It has been shown that HIV recombines more frequently than SNV and Moloney murine leukemia virus (MuLV) (70, 72, 100). HIV recombination has been estimated to occur about two or three times per replication cycle which is 10-20 times greater than SNV and MuLV (72). The results show that HIV recombines about one time for every 3000 bases in the genome. This is a very high recombination rate that can lead to much variation among quasispecies.

Non-homologous transfer occurs approximate 0.1 – 1 % as much as homologous transfers (149). *In vivo* experiments have been done using the R regions of different retroviruses that have very little homology. Non-homologous transfer was observed, though the frequency was about 1000 to 100 times less than homologous transfer. When

the transfer zone was examined, small regions of sequence similarity (less than eight nucleotides) were found to be the site of transfer (144, 149). Phylogenic analysis has indicated that non-homologous transfer between two SIV strains may be responsible for the acquisition of the vpx gene (an additional gene found in SIV, not found in HIV-1) in the SIV sooty mangabey strain (120). These results indicate that both non-homologous and homologous transfer can occur and give rise to alterations in progeny virus' genetic code. However, homologous recombination is far more frequent than non-homologous.

HIV-1 Nucleocapsid Protein

The research presented in this dissertation examines the chaperone activity of HIV-1 nucleocapsid protein. HIV-1 NC is shown in Figure 5. This protein contains 55 amino acids, of which 15 are basic arginine or lysine residues. HIV-1 NC contains two zinc finger motifs which have the amino acid sequence $CX_2CX_4HX_4C$, where X is any amino acid. The three cysteine residues and the histidine residue bind zinc with high affinities (24, 48, 123, 126, 128). The zinc fingers are connected by a seven amino acid linker region, which contains four basic amino acids. The linker region contains a proline residue which has been shown to kink the linker region such that the two zinc fingers are spatially close, particularly the aromatic residues, Phe16 and Trp37, of finger 1 and 2, respectively (93). Also, the folding of the first zinc finger around the zinc ion has shown that there is a hydrophobic patch on the surface of the zinc finger (126, 127). NMR studies of NC in solution have shown that the N-terminal and C-terminal regions are flexible and have no predicted structure (93). However, circular dichroism studies have
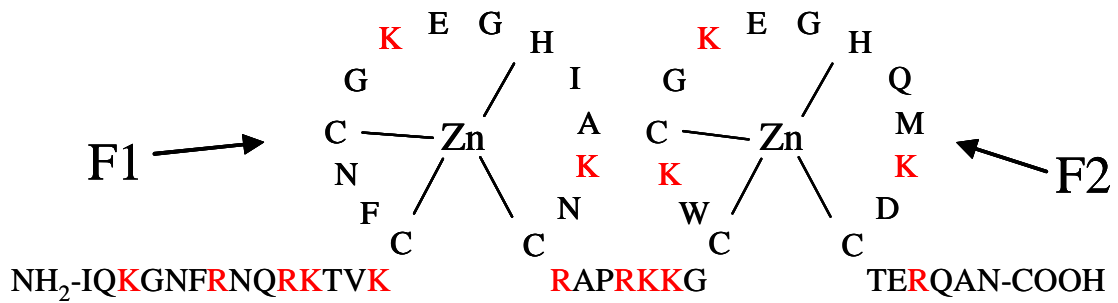
Figure 5: Schematic of HIV-1 nucleocapsid protein from pNL4-3. NC contains 15 basic amino acids, which are shown in red. It also has two CCHC motifs that each coordinate a zinc ion. These zinc fingers are labeled F1 and F2 for finger 1 and finger 2, respectively. The N terminal and C-terminal regions are flexible whereas the zinc fingers form a globular structure in the center of the molecule.

shown that these regions can form helices, and that the predicted N-terminal helix would cause the five basic residues to cluster on one side of the helix (128).

NC structure has also been studied when bound to different portions of the HIV packaging signal. NMR studies performed with the N-terminal zinc finger and a five nucleotide sequence that is a portion of the Ψ-site, dACGCC, showed that the nucleic acid was bound within the four amino acid hydrophobic cleft in the finger. This positioning enabled the formation of hydrogen bonds and electrostatic interactions between the positively charged arginine residue and the backbone of the nucleic acid (124). The binding of NC to stem loop 3 (SL3) of the Ψ-site showed that the zinc fingers form interactions with two guanine residues in the loop region while the N-terminal region of the protein formed a helix that inserted into the major groove of the stem (28). Similarly, NMR studies with NC and stem loop 2 (SL2) of the Psi-RNA showed that the zinc fingers interact with the guanine residues in the loop region while the N-terminal helix interacts with the stem (4). However, the orientation of the zinc fingers on the loop region differed slightly between the two stem loops.

HIV-1 NC mutagenesis studies have shown that the basic residues are vital for NC binding to nucleic acids (119, 128, 136). Filter binding assays and northern blot analysis was done with 5' leader RNA and mutant NC proteins. The basic residues in these proteins were substituted with neutral amino acids. These experiments showed that R-7, R-32 and K-33 were of particular importance for NC nucleic acid binding (119). R-7 is in the N-terminal region while R-32 and K-33 are in the basic linker region. R-7 was also shown to be involved in electrostatic contacts with RNA when tested with poly($1,N^6$-ethenoadenylic acid) (poly(εA)) fluorescence. Poly(εA) is a homopolymeric fluorescent

RNA and when NC bound to the RNA fluorescence intensity increased. The effect of electrostatic interactions in NC:RNA binding was measured by adding salt back into the reaction after NC binding. These results showed that the R-7 residue was important for binding. However, they tested a number of other mutants as well and showed that two basic residues in the N-terminal zinc finger also engage in electrostatic interactions with the NC (136). Experiments with zinc bound and zinc free forms of NC show that NC which is bound to zinc has a higher ability to bind and unwind DNA (106).

HIV-1 NC is a nucleic acid chaperone (24, 110). This means that NC will unfold weak secondary RNA and DNA structures in order to enhance the formation of stronger structures (67). This activity has a number of implications in the viral life cycle, and this is discussed at length in Chapter 2. The chaperone activity has been studied extensively and NC has been shown to unwind a number of different structures to enhance formation of more favorable structures (6, 66, 133, 137). NC has been shown to unwind DNA using optical tweezers to stretch single DNA molecules (140). In this experiment, optical tweezers were used to hold a stretavidin bead. Another streptavidin-coated bead was trapped by laser beams. A single DNA molecule that was labeled on each end with biotin was captured between the two streptavidin-coated beads. Force extension measurements were taken by moving the pipette, stretching the DNA molecule. These experiments were done in the presence and absence of NC and NC lowered the cooperativity of the helix to coil transition. NC lowered the force necessary for the helix to coil transition indicating that NC aids in unwinding the DNA (140). HIV-1 NC and Mason-Pfizer monkey virus (MPMV) NC have both been shown to enhance the renaturation of a 149 base pair DNA indicating that chaperone activity is not individual to HIV-1 NC (41).

Golinelli *et al.* performed annealing experiments of small oligonucleotides (< 30 nucleotides) and observed that NC enhanced annealing of structured oligonucleotides but had no effect on the annealing of unstructured sequences (55). However, results presented in this dissertation show a clear stimulation of RNA:DNA annealing of unstructured sequences in the presence of NC. The zinc fingers of HIV-1 NC have been shown to be important to the chaperone activity (62, 64, 66, 139). Other experiments have suggested it is the basic residues that are important for NC chaperone activity (29, 83). The basic residues are vital for NC binding to nucleic acids. Therefore, these residues would certainly be necessary for the chaperone activity. However, recent experiments indicate that the basic residues alone are not sufficient for HIV-1 NC chaperone activity.

HIV-1 NC RNA binding is vital for the viral life cycle. NC coats the genome of HIV within the virion core. The positively charged residues can easily associate with the negative charge of the phosphate backbone of the genome and the hydrophobic residues can associate with the hydrophobic bases. HIV-1 NC is incorporated into the virion as part of the precursor Gag protein. The initial PR cleavage of Gag releases p15 from the C-terminal end of Gag. NC is the N-terminal peptide in p15 and two more PR cleavages release NC (p7) from p1 and p6 (78). After NC has initially been cleaved from the Gag precursor protein it has been shown to have an occluded binding site of fifteen nucleotides. When it is further cleaved to the mature 55 amino acid form the occluded binding site is between six and eight nucleotides (78, 79, 136, 145).

NC binding to RNA is slightly cooperative (145). Because of the highly basic nature of the protein, NC-NC association is not likely without the negative charges of the

phosphate backbone present. These negative charges neutralize the repulsion between basic NC molecules. Within the core there are approximately 2000 molecules of NC (24). One major function of NC is to protect the genome from any RNase cleavages. NC bound to RNA stem loops was shown to be bound to bulges on the stem loops because RNase T1 cleavage was inhibited in these locations with NC present (3). Also, NC was shown to protect linear DNA from nuclease digestion (82).

HIV-1 NC binds to nucleic acids nonspecifically. Nitrocellulose binding assays have shown that NC binds with higher affinity to single stranded RNA than single stranded DNA. Also binding affinity was additionally decreased when double stranded DNA was tested (128). However, fluorescence studies have shown that NC has similar binding affinities for single and double stranded DNA (137). Though NC will bind to all different types of nucleic acids some experiments have shown preference for different nucleic acid sequences or structures. Fluorescence studies completed to test NC binding to SL3 of the packaging signal showed that NC does bind preferentially to TG sequences and that the stem-loop structure of SL3 stabilized NC binding. NC mutagenesis in these experiments showed again that Trp37 interaction with a G residue in the loop was important for binding (138). Additionally, Fisher *et al*. showed that NC binds preferentially to TG repeats in experiments using surface plasmon resonance. This was tested using RNA as well and NC bound to UG repeats preferentially (47). Systematic evolution of ligands by exponential enrichment (SELEX) experiments have shown that NC binds preferentially to sequences that form stem loop structures (3, 86). Also, RNase digestion experiments showed that NC binds specifically to bulges within the stem loops (3). Other SELEX experiments show that NC binds preferentially to stretches of

guanosines and uridines and that this is enhanced when these sequences are found at the end of a long and stable stem loop structure (10). Lastly, experiments examining the binding of NC to stem loop four (SL4) of the Ψ-site have conflicting results. Nitrocellulose binding assays have shown that NC binds to SL4 with high affinity, but NMR and gel shift studies dispute this finding (5, 18).

It is therefore evident that HIV-1 NC binds to all types of nucleic acids, but does have a preference for some of the stem loops in the HIV-1 packaging signal. This is important because the NC region of the Gag precursor is necessary for efficient packaging of the viral genome into progeny virions. Early experiments used the baculovirus expression system to determine the role of the NC region of Gag in virus formation. When Gag protein was expressed using baculovirus in insect cells virus like particles bud from the cell surface. However, when a modified Gag protein was expressed that had p15 sequences deleted from the C-terminal, the Gag molecules would associate with the cell membrane but no particle assembly occurred (54). This shows that the NC region of Gag is vital for virion particle assembly. Also, experiments have been done to examine what region of the packaging signal is most important for NC packaging. A RNA stem loop which was shown to be a high-affinity ligand for HIV-1 NC was substituted for SL3 or SL1 and packaging was evaluated. It was shown that the ligand could be substituted for SL3 and efficient packaging would occur. However, this did not hold true for SL1, which contains the DIS. The ligand abolished dimerization, and therefore packaging was not efficient using this ligand substitution for SL1 (19).

HIV-1 NC enhances the dimerization of the genomes as well. The dimerization palindrome, found in SL1, is the initiation site for dimerization. NC enhances the

binding of these sequences between SL1 of adjacent genomes. However, NC has also been shown to enhance the transition of the loop-loop complex to an extended dimer (97). This includes the unwinding of the SL1, such that the stem sequences can bind to the complementary sequences of the other genome. Further experiments showed that it is the basic residues in the N-terminal and linker region that are important for NC enhancement of the change from loop-loop dimer to extended-duplex dimer (129). Also, Gag polyprotein has been shown to enhance dimerization (45). This is important because NC enhancement of dimer formation can occur before cleavage of NC from the polyprotein.

There has also been a great deal of study regarding what residues of NC are important for RNA packaging. Alanine scanning mutagenesis has been performed to determine the involvement of the basic residues in packaging. These experiments showed that the basic residues in the N-terminal tail, basic linker and the zinc fingers were all important for efficient packaging (104). Experiments have also shown the importance of the zinc finger structure on RNA packaging. Morellet *et al.* mutated the linker region of NC by substituting a leucine for proline in position 31. The proline residue provides a kink in the linker that causes the two zinc fingers to cluster together. This substitution led to the formation of immature particles that were not infectious (92). Also, the first zinc finger was disrupted in experiments where cysteine was substituted for His23. This CCCC motif still binds zinc, but the spatial proximity of the two fingers is interrupted. This mutation did not disrupt virion formation but the virions formed did not package viral RNA efficiently and were not infectious (31). Also, experiments have been completed with NC proteins that have the position of the zinc fingers switched or altered.

The two zinc fingers have five amino acids that differ between them. When mutants were used that had two copies of the first zinc finger (1.1 NC) 70 % of wild-type levels of RNA was packaged. However, when a finger switch mutant (2.1 NC) or a mutant with two copies of the second zinc finger (2.2 NC) was used less than 15 % of wild-type levels of RNA was packaged (58). This indicates that the N-terminal finger is important for packaging, and this finger should be in its native position for efficient packaging. However, other experiments show that certain mutations in the first zinc finger can still lead to viruses that package high levels of viral RNA. When residues in the first zinc finger were substituted with residues from seven different cellular nucleic acid binding proteins (CNBPs) that contained zinc fingers, RNA was efficiently packaged with six of the seven mutants (90). Therefore the first zinc finger can tolerate certain modifications and continue to package viral RNA. The importance of the retroviral NC protein in genomic RNA packaging has also been shown with SIV NC. Mutational analysis has indicated that the first zinc finger in SIV NC is important for RNA encapsidation (1). Interestingly, when the single MuLV zinc finger was mutated from the CCHC motif to CCCC or CCHH packaging was not affected (57). This supports the notion that packaging of the viral RNA may be reduced by mutations that alter the spatial proximity of the two zinc fingers because this interaction does not exist in the MuLV NC protein.

After the virus buds from the cell membrane the maturation process ensues. This process involves the tight condensing of the viral RNA and the formation of the core structure. HIV-1 NC has been shown to be important for both the core structure formation and maturation of viral RNA. *In vivo* experiments have shown that certain HIV-1 NC mutations can alter the viral core structure. When the N-terminal CCHC

motif was either deleted, or changed to CCCC, the diameter of progeny virion was larger, and the core morphology was abnormal (130). Also, mutational analysis showed some basic residues in the N-terminal region of NC were important for core morphology. The mutant R7R10K11S, which had the arginines in positions 7 and 10 and the lysine in position 11 changed to serine, showed stick-shaped cores. On the other hand, K14D had either aberrant, noncircular particles or normal appearance where the dense nucleocapsid structure occupied only part of the conical core (12). The viruses in these experiments were examined by electron microscopy. The linker region was also mutated to determine the importance of this region on core structure. When the proline at position 31 was mutated to a leucine residue the viruses produced had immature core morphology (101). Also, *in vitro* experiments have shown that HIV-1 NC enhances the maturation of dimeric RNA from Harvey Sarcoma Virus (HSV). These experiments showed this RNA was converted from a less stable to more stable form with the addition of NC (46). Also the processing between p2 and NC in Gag has been shown to be critical for the maturation of the dimeric genome (110, 122). This cleavage releases NC from the Gag protein, and therefore the mature NC protein is implicated in RNA maturation.

After the virus buds from the cell membrane, and maturation has occurred, the new virus can then infect a new cell. The synthesis of the proviral DNA occurs when the core is released into the newly infected cell. HIV-1 NC has been shown to enhance a number of steps in reverse transcription. Reverse transcription begins with the binding of the tRNA primer to the PBS on the RNA genome. HIV-1 and MuLV NC both enhance the binding of the tRNA^{Lys3} to the PBS (8, 29, 30, 83, 111). CD spectrum of NC:tRNA complexes indicate that NC enhances the unwinding of tRNA (79). Also terbium

cleavage assays have shown that NC binding alters the cleavage profile showing that the tRNA unwinds due to NC binding (65). This would make the 3' end of the tRNA available for annealing to the PBS. However, other experiments have indicated that NC does not unwind the tRNA (60). It is somewhat unclear whether NC enhancement of tRNA binding is due solely to enhancement of the annealing of the two RNAs or if NC also enhances the unwinding of tRNA. There is also discrepancy about the residues of NC that are important for tRNA annealing to the PBS. When zinc fingers were deleted from NC protein no tRNA:PBS annealing was observed by gel-shift analysis (111). Also reverse transcription experiments examining the initiation step showed that initiation was reduced when a zinc finger deletion NC mutant was used (115). Contrary to these results De Rocquigny *et al.* showed that the basic residues exterior to the zinc fingers are the important residues for tRNA:PBS annealing (30). This was also shown for MuLV NC protein (29).

Throughout the process of reverse transcription NC enhances different types of strand transfer. NC has been shown to enhance the obligatory minus strand transfer (2, 43, 61, 80, 146). One important role NC may have in enhancing minus strand transfer is that it inhibits self-priming, a process that prevents the DNA from being synthesized correctly (43, 61). Self-priming occurs at the 5' end of the template RNA because the minus strand DNA forms a hairpin complementary to the TAR RNA. When the RT has synthesized the TAR DNA it sometimes continues synthesis using the minus strand DNA as a template. This occurs because the hairpin conformation brings the 5' base of the hairpin in close proximity to the RT:DNA complex. Therefore synthesis continues using the 5' end of the minus strand DNA as a template. Investigators have shown that when

oligonucleotides are present that are complementary to the TAR DNA self-priming is

inhibited by NC. Plus strand transfer is also stimulated in the presence of NC (62, 142).

Also, internal strand transfers, which may not be obligatory, but are important for the

generation of genetic diversity have been shown to be enhanced by NC (7, 33, 69, 98, 99,

108, 113, 133). NC enhances the RNase H activity of RT which stimulates the transfer of

the nascent DNA (15, 42, 62). Also the zinc fingers of NC are important for the

enhancement of strand transfer (62, 64). Often in reverse transcription RT will reach a

site on the RNA that has a strong secondary structure. This region can cause the RT to

pause, lowering the efficiency of reverse transcription. NC has been shown to minimize

pausing in DNA extension experiments (32, 81). This indicates that strand transfer may

be important for overcoming pausing during reverse transcription. The effect of NC on

strand transfer is discussed in greater detail in Chapter 3. NC has been shown to have a

variable affect on RT processivity, enhancing processivity in some regions and having no

affect on other regions (33, 73). Processivity is the number of nucleotides synthesized

without enzyme dissociation. This is a measure of how quickly synthesis occurs. Taken

together, these experiments indicate that NC is vital to the replication of genomic RNA

into proviral DNA.

Another role for NC in the viral life cycle could be its involvement in proviral

integration. NC was first shown to enhance the IN cleavage reaction *in vitro*. IN cleaves

two nucleotides from the ends of the proviral DNA to generate recessed 3' ends for

integration into the host DNA. *In vitro* experiments tested the cleavage of two

nucleotides of a DNA oligonucleotide pair. When NC was included in the reaction

cleavage was increased nearly two fold (82). Other experiments have shown that NC

also enhances the integration of DNAs mimicking the long terminal repeats into target DNA (16, 53). In one case coupled integration was increased 1000 fold in the presence of HIV-1 NC (16). Also, NC has been found as part of the preintegration complex (14, 51). These experiments all suggest that HIV-1 NC may have an important role in the integration of proviral DNA into the host genome.

Chapter 2

HIV-1 NC zinc fingers have differing roles in nucleic acid annealing.

Introduction

As discussed in the General Introduction, HIV-1 NC has been shown to possess nucleic acid chaperone activity (41, 82, 83, 110, 133, 137, 140). Nucleic acid chaperones aid in the unfolding of nucleic acid structures to enhance the annealing of more thermodynamically favorable structures (containing more base pairs) (67). This chaperone activity aids in tRNA:PBS annealing, annealing of RNA and DNA in strand transfer, genome dimerization and maturation.

Genome maturation requires the condensing of the genome by NC. NC has been shown to enhance the maturation of HSV RNA from a less stable dimer to a more stable dimer (46). NC has been shown to enhance dimerization of HIV RNA and enhance the conversion to an extended duplex formation rather than a loop-loop dimer conformation (45, 97). This enables the RNA interactions to take on a more stable form condensing the genome during maturation. NC has also been shown to enhance the annealing of the tRNA^{Lys3} to the PBS (8, 17, 83). There is discrepancy regarding whether NC is able to unwind the tRNA or just enhance the hybridization to PBS (60, 65, 79). However, there is a consensus that the overall tRNA:PBS annealing is enhanced by NC. This leads to a stimulation of the initiation of reverse transcription. Both terminal strand transfers (minus and plus strong-stop DNA transfers) have been shown to be stimulated when NC is included in the reaction *in vitro* (2, 7, 25, 61, 74). Also, internal strand transfer has been shown to be enhanced by NC *in vitro* (32, 33, 108). Furthermore, different types of

nucleic acid annealing reactions have been done to examine NC's chaperone activity. These experiments have been done with nucleic acids that are not derived from the genome but are useful in understanding NC chaperone activity. NC has been shown to enhance annealing in almost every circumstance (55, 133, 137). The only report where NC did not enhance nucleic acid annealing involved very short (less than 30 nucleotides) unstructured oligonucleotides (55). Though NC did not enhance annealing in these experiments it could have been because NC binds with less affinity to shorter oligonucleotides (109). Therefore, NC chaperone activity has been clearly demonstrated in many different types of experiments using many different nucleic acids.

Though NC has been shown in a variety of experiments to be a nucleic acid chaperone, it is not clear if there are particular regions of NC that are responsible for this activity or if various components (helix-destabilization (unwinding) or hybridization of complements) of its' chaperone activity are catalyzed by different regions of NC. A number of investigators have completed annealing assays with mutant NC proteins to determine what residues are necessary for NC chaperone capability. Studies examining the effect of MuLV NC mutants on the annealing of tRNA[Pro] to MuLV RNA, as well as the dimerization of the genomic RNA segments showed that the basic regions of MuLV NC were necessary for enhancement (29, 105). These studies indicated that the single zinc finger of MuLV NC could be removed and chaperone activity was retained. Similar experiments were completed with genomic RNA sequences and NC derived from HIV. RNA dimerization and tRNA[Lys3] binding were observed in the presence of HIV-1 NC mutants that contain only amino acid sequences exterior to the zinc fingers (28, 30, 69, 83, 129). Mutants with zinc fingers replaced by glycine-glycine linkers were able to

32

enhance annealing, though the concentration of these proteins was often much higher than used with wild type NC. *In vitro* experiments have also been completed with ribozymes that will cleave RNA sequences upon annealing to them (96, 134). RNA cleavage was observed in the presence of wild type NC as well as mutants containing the conserved basic residues with deleted zinc fingers (96). Though these reports indicate that only the basic residues of NC are important for chaperone activity it is also known that these residues are vital for the RNA and DNA binding activity of NC (29, 30, 96, 136). Therefore the reduction in annealing due to deletion of basic residues could merely be a result of lowered NC binding to the nucleic acid. If the protein does not bind properly to the nucleic acids it may not enhance annealing. Therefore, the deletion of basic residues does not rule out the importance of zinc fingers in annealing because the mutant proteins may not bind to the nucleic acids appropriately.

In contrast to the above results, similar annealing experiments have also been completed with NC proteins that retained the zinc fingers but contained mutated residues within these regions. Guo *et al.* (64) completed annealing assays using sequences from the TAR region of the HIV-1 genome. They used three types of proteins with mutations that alter either the structure of the zinc finger or the residues within the zinc finger that are not involved in zinc binding. Optimal annealing activity was only observed when the N-terminal zinc finger was not mutated. The effect of mutations to the C-terminal zinc finger were not as pronounced, though these results varied with the type of mutant constructed. The results are supported in experiments that tested the effect of altering the zinc finger structure on tRNA^{Lys3} annealing to the PBS (111). In this case the native zinc fingers were shown to be necessary for the annealing reaction. Also, annealing assays

33

completed with the nucleic acid sequences from the regions of minus and plus strand transfer, in the presence of NC mutants unable to bind zinc, showed that the requirement for the zinc finger structure differed due to the nucleic acid sequence being used (62). Lastly, single DNA molecule stretching experiments have been completed with NC to examine its ability to destabilize the helix and aid in transition to a coil structure. HIV-1 NC mutants were used in these experiments and it was observed that the N-terminal finger of HIV-1 NC must be in the native position for optimal chaperone activity (139, 140).

Though the chaperone activity of HIV-1 NC has been extensively studied we wanted to evaluate the possible reasons for the current discrepancies regarding the role of the zinc fingers. The work presented here investigates how NC enhanced nucleic acid annealing differs for unstructured and structured nucleic acids. Work presented in this thesis showed that NC increases the annealing of both types of nucleic acids, indicating that it stimulates both the unfolding of structured nucleic acids as well as the direct hybridization of complements. NC zinc finger mutants were also used in annealing assays that provided additional insight into the potential role of zinc fingers in annealing. The N-terminal zinc finger was shown to be necessary for the annealing of sequences with a high degree of secondary structure, while the C-terminal zinc finger was shown to enhance the annealing of the unstructured sequence. Overall the results suggest that the N-terminal finger is more important in unwinding secondary structures while the C-terminal finger has an accessory role in annealing.

Materials and Methods

Preparation of wild type and mutant NC proteins- HIV-1 NC from the HIV-1

AIDS associated retrovirus (ARV) strain was prepared as described (145). The only

adjustment made was that it was done in $1/35^{th}$ the proportions used in the protocol. Wild

type and mutant NC from the HIV-1 NL4-3 strain was prepared as explained previously

(16). There are four differing amino acids between the NC proteins derived from the two

HIV strains. NC from the ARV strain contains two arginine residues in the C-terminal

zinc finger that are substituted with lysine residues in the NC from the NL4-3 strain.

Aliquots of HIV-1 NC were prepared and stored in 50 mM Tris-HCl (pH 7.5), 10%

glycerol and 5 mM 2-mercaptoethanol at -80 ^{o}C. Fresh aliquots were used for each

experiment.

Preparation of oligonucleotides- Any chemicals or enzymes that do not

specifically mention where they were obtained from were from Sigma or Fisher

Scientific. RNA oligonucleotide 0.0rna was purchased from Dharmicon Research, Inc.,

and DNA oligonucleotide 0.0dna was from Invitrogen Life Technologies. The most

highly structured pair of oligonucleotides, 21.7rna and 21.7dna was purchased from

Integrated DNA technologies. The RNA contained a 5' fluorescein-6-carboxamidohexyl

(FAM) end label while the DNA contained a 3' 4-[[(4-dimethylamino)phenyl]-azo]

benzenesulfonicamino (DABCYL) group. Oligonucleotides 7.5rna, 16.3rna, and the

21.7rna used in nuclease mapping experiments, were transcribed from DNA

oligonucleotide pairs. One DNA strand of each pair was 61 nucleotides long and

contained the sequence for the SP6 promoter at the 5' end followed by the DNA

sequence corresponding to the desired RNA. Twenty pmoles of the 61-mer was

combined with 40 pmoles of a complementary 42-mer DNA in a hybridization buffer (50 mM Tris-HCl (ph 8.0), 1 mM DTT, and 80 mM KCl). The hybrid reaction was heated to 65 $^{\circ}$C for five minutes and then cooled slowly to room temperature. The 5' overhang was then filled in using Klenow (New England Biolabs) at 37 $^{\circ}$C. The hybrid was incubated for 1 hour with 10 units of Klenow and 200 μM dNTPs (Roche Applied Science). The double stranded DNA product was extracted by phenol:chloroform:isoamyl alcohol (25:24:1) and precipitated with ethanol. Forty units of SP6 polymerase (New England Biolabs) was then used to transcribe the 42 nucleotide RNA product. The transcription reaction conditions were 2 units/μL RNase inhibitor, 40 mM Tris-HCl PH 8.0, 6 mM MgCl$_2$, 10 mM DTT, 2 mM spermidine, 1 mM dNTPs and the reaction was run for 2 hours at 37 $^{\circ}$C. DNA was digested with 15 units of DNase 1 (Boehringer Manheim) for 10 minutes at 37 $^{\circ}$C. The reaction was stopped by adding an equal volume of 2 X formamide dye (95 % formamide, 18 mM EDTA, 0.025 % SDS, 0.025%, xylene cyanol, bromophenol blue). The RNA was heated to 90 $^{\circ}$C for 3 minutes and then subjected to gel electrophoresis on an RNase free 10% denaturing polyacrylamide gel. RNA was excised and eluted in formamide elution buffer (80% formamide, 400 mM NaCl, 1 mM EDTA, 40 mM Tris-HCl (pH 7.0)) then precipitated with ethanol. The 5' end of the RNA was dephosphorylated using calf intestinal alkaline phosphatase (CIP). Seventy-five pmoles of RNA was incubated with CIP for one hour at 37 $^{\circ}$C. The dephosphorylated RNA was extracted by phenol:chloroform:isoamyl alcohol (25:24:1) and precipitated with ethanol. All of the aforementioned oligonucleotides were resuspended in 30 μL of water and quantified spectrophotometrically using the optical density.

RNA end labeling- RNAs 0.0rna, 7.5rna and 16.3rna were labeled at the 5' end

with $(\gamma-^{32}P)$ ATP. 21.7rna used in mapping experiments was also labeled at the 5' end.

Fifty pmoles of dephosphorylated RNA was end labeled using T4 polynucleotide kinase.

The 5' end of the most highly structured sequence, 21.7rna, was connected to a FAM

molecule. Therefore, this RNA was 3' end labeled with 5'-^{32}P cytidine-3',5' bis

phosphate as described. 10 mM Cytidine-3'phosphate was labeled with 15 µL g-P^{32} ATP

(10 µCi/µL) and 20 units of T4 PNK in reaction buffer (50 mM Tris-HCl pH 8.0, 10 mM

$MgCl_2$, 5 mM DTT, 1.5 mM spermidine-HCl. The reaction was incubated at 37 ^{o}C

overnight and then heated to 90 ^{o}C for 3 minutes. The reaction was centrifuged and 17.5

µL of the supernatant was removed and added to 50 pmol of 21.7rna, and 25 units of

RNA ligase. The ligase buffer contained 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10

mM DTT, and 30 mg/L BSA. The ligation was incubated at 4 ^{o}C overnight. The

reaction was stopped by adding an equal volume of 2 X formamide dye. All labeled

RNAs were purified on a 10% polyacrylamide gel, eluted and precipitated as described

above. RNAs were resuspended in 70 µL of TE buffer (10 mM Tris-HCl pH=8, 1 mM

EDTA pH=8) and quantified spectrophometrically.

RNase mapping experiments- 5' end P^{32}-labeled RNA was mapped to using T1

nuclease and RNase A. To begin reactions, reaction buffer was made (final concentration

50 mM Tris-HCl pH 7, 6 mM $MgCl_2$ and 80 mM KCl) and 7 µL was aliquoted into each

reaction tube. To each tube 2 µL of RNA was added and 1 µL of either RNase A or T1

RNase was added. Reaction continued at room temperature for two minutes. Reactions

were stopped by adding 5 µL of 2 X formamide dye. The concentrations of RNase A

used were 0, 0.25, 0.5, 1, 2.5 and 5 units. T1 RNase was used at 0, 0.1, 0.5, 1, 5, 10 and

50 units. Two ladders were used, a G-ladder and a base hydrolysis ladder. The T1

RNase digestion ladder was created by incubating 2 µL of RNA with 1 µL of enzyme in

T1 buffer (300 mM NaCl, 10 mM Tris pH 7.6, 5 mM EDTA) at 37 ^{o}C for 10 minutes.

Then 15 µL of 1 X 80 (50 mM Tris-HCl pH 8, 0.1 mM EDTA, 1 mM DTT, 80 mM KCl)

and 20 µL of 2 X formamide dye was added to stop the reaction. The base hydrolysis

ladder was made by incubating 1 µL of RNA in 3 µL of water. The reaction was heated

to 65 ^{o}C for five seconds and 1 µL of NaOH was added. Reaction was heated to 65 ^{o}C

for fifteen seconds and 1 µL of HCl was added. Reaction was stopped by adding 15 µL

of 1 X 80 and 20 mL of 2 X formamide dye. Eight microliters from each reaction was

run on a 12 % denaturing polyacrylamide gel. Gels were dried and autoradiograms were

obtained.

 NC titration- An NC titration was performed with 0.0rna to determine the

appropriate amount of NC to use in annealing assays. Separate RNA and DNA reactions

were made. Reactions were preincubated with the specified concentrations of NC.

Reactions were started by adding 4.5 µL of the DNA:NC reaction to 10.5 µL of the

RNA:NC reaction. Final concentrations were 5 nM RNA, 10 nM DNA, 50 mM Tris-HCl

(pH 8.0), 0.1 mM EDTA, 1 mM DTT, 6 mM $MgCl_2$, 80 mM KCl, and 100 µM $ZnCl_2$.

NC concentrations were 0, 0.5, 1.0, 2.0, 4.0, 8.0 or 15.0 µM. Reactions were stopped

after 15 seconds by adding 7.5 µL of stop solution (0.25% bromophenol blue, 20%

glycerol, 20 mM EDTA (pH=8), 0.2% SDS, 0.4 mg of yeast tRNA per mL) (133).

Reactions were run on 15 % native polyacrylamide gels and autoradiograms were

obtained. Gels were quantified using a BioRad GS-525 phosphorimager.

RNA:DNA annealing assay- End ^{32}P-labeled RNA and the complementary DNA were diluted in TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA) and separately heated to 90 ^{o}C for 3 minutes and then transferred to ice for five minutes. The RNA was then preincubated at 37 ^{o}C in reaction buffer in the presence or absence of mutant or wild type HIV-1 NC for two minutes. Complementary DNA was separately preincubated at 37 ^{o}C for two minutes in the presence or absence of NC. To start the reactions 17 μL of DNA solution was added to 90 μL of reaction mixture containing the RNA. The final concentration of RNA was 5 nM oligonucleotide, which was 0.21 μM nucleotide and the final concentration of DNA was 10 nM oligonucleotide, which was 0.42 μM nucleotide. NC final concentration was 2 μM and the buffers were 50 mM Tris-HCl (pH 8.0), 0.1 mM EDTA, 1 mM DTT, 6 mM MgCl$_2$, 80 mM KCl, and 100 μM ZnCl$_2$. Aliquots of 15 μL were removed at specific time points (as indicated) and added to 7.5 μL of the stop solution (0.25% bromophenol blue, 20% glycerol, 20 mM EDTA (pH=8), 0.2% SDS, 0.4 mg of yeast tRNA per mL) (133). All reactions were incubated in the stop solution at 37 ^{o}C for 1 minute before being transferred to ice. Reactions were then subject to electrophoresis on 12% or 15% native polyacrylamide gels. Gels were dried and subjected to autoradiography (118) or phosphorimager analysis using a BioRad GS-525 phosphorimager. Percent annealing was determined by dividing the amount of annealed product (A) in each lane by the total RNA (annealed and single stranded (S)) in each lane and multiplying by 100 (Percent Annealed = A/(A+S) * 100).

Annealing detected by fluorescence resonance energy transfer (FRET)- As was previously noted, 21.7rna was purchased with a FAM molecule on its 5' end. The complementary DNA was purchased with a 3' DABCYL. Annealing assays were

completed at 25 $^{\circ}$C using a Fluoromax-2 spectrofluorometer (Jobin Yvon Instruments, S.A. Inc.). The RNA and DNA were separately incubated in the presence or absence of wild type NC or mutants 1.1 and 2.2. The reactions were started by mixing 10.5 µL of the DNA-NC solution and 59.5 µL of the RNA-NC solution. The final concentration of the RNA and DNA was 5 nM and 10 nM, respectively (total nucleotide concentration 0.63 µM). The NC was used at a final concentration of 2 µM and the final concentration of the reagents in the buffer was as follows: 50 mM Tris-HCl (pH 8.0), 0.1 mM EDTA, 1 mM DTT, 6 mM $MgCl_2$, 80 mM KCl, and 100 µM $ZnCl_2$. Excitation wavelength was 494 nm with a bandwidth of 1 nm. The emission bandwidth was 5 nm and the spectrum was observed from 508 to 570 nm. The emission spectrum was taken at times 0, 1, 2, 4, 8 and 16 minutes. The intensity of the emission peaked at 517 nm. An intensity ratio (I_r) was determined by dividing the peak intensity at a given time (I_t) by the peak intensity at time zero (I_0) ($I_r = I_t/I_0$). Three experiments were taken for each type of NC and an average intensity ratio was determined and plotted versus time.

Results

Structure of substrate RNAs. The RNAs used in these experiments each contained 42 nucleotides. The structures are shown in Figure 6. These structures were determined using both RNAdraw and mfold using mfold default conditions (89, 150). The Gibbs free energy for unfolding at these conditions is shown next to the structure. Nucleic acids were designed such that they would contain the same number of nucleotides but have an increasing strength of secondary structures. Structures were chosen that did not have a

40
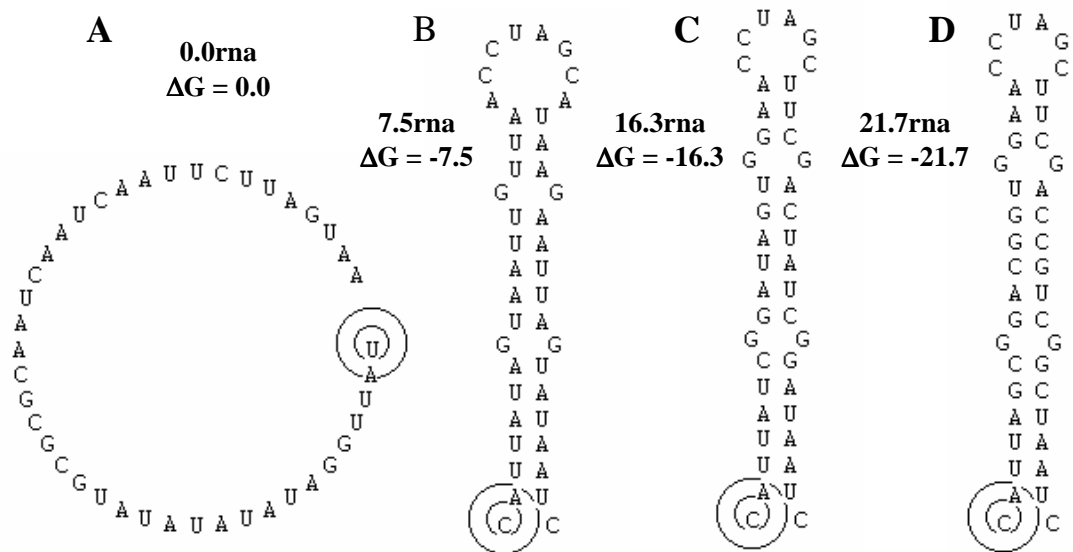
Figure 6: Structure of RNAs used in annealing assays. RNAs were designed to contain the same number of nucleic acids (42) though the sequence composition differed. Free energies for unfolding are given in kcal/mol using mfold default conditions, 37 °C and 1 M NaCl. Actual free energies for unfolding would be altered due to conditions used in the reactions. A) 0.0rna does not have any predicted base pairs. B) 7.5rna contained 14 A-U base pairs and no G-C pairs. C) G-C content was increased in 16.3rna which had 4 G-C base pairs and 11 A-U base pairs. E) The strongest structure, 21.7rna, contained 15 base pairs with 7 G-C pairs and 8 A-U pairs. Concentric circles denote the 5' end.

high degree of GU (or GT in the complementary DNA) repeats due to reports indicating that NC may preferentially bind to these sequences (47, 138).

Each structure was subject to RNase mapping using T1 RNase and RNase A. RNase A cleaves 3' of uracil and cytosine residues in a single stranded region of the RNA whereas T1 RNase cleaves 3' of single stranded guanine residues. The reactions were run on 10 % denaturing gels. Gels were dried and autoradiograms were obtained. The mapping profiles are shown in Figure 7 for the weakly structured substrates and Figure 8 for the stronger substrates. The profiles showed that the structures were consistent with the structures predicted by RNAdraw and mFold. For each of the stem loops there were cytosine residues in the loop region that were not cleaved by RNase A. However these residues were adjacent to uracil residues that were predominantly cleaved. This could potentially be due to a preference for RNase A to cleave uracil residues rather than cytosine residues, or the orientation of the binding of the enzyme to the loop region. 7.5rna did show cleavages in regions of the stem that were predicted to be double stranded. However, this is a weak stem loop so it is expected that the RNA would fluctuate between the open and closed state. The stronger structures had faint cleavages in regions that should be double stranded, but these cleavages always appeared after the predominant cleavage at G23 or U21, depending on the enzyme used. This indicates that once the RNA was cleaved unwinding occurred making other residues accessible for nuclease degradation.
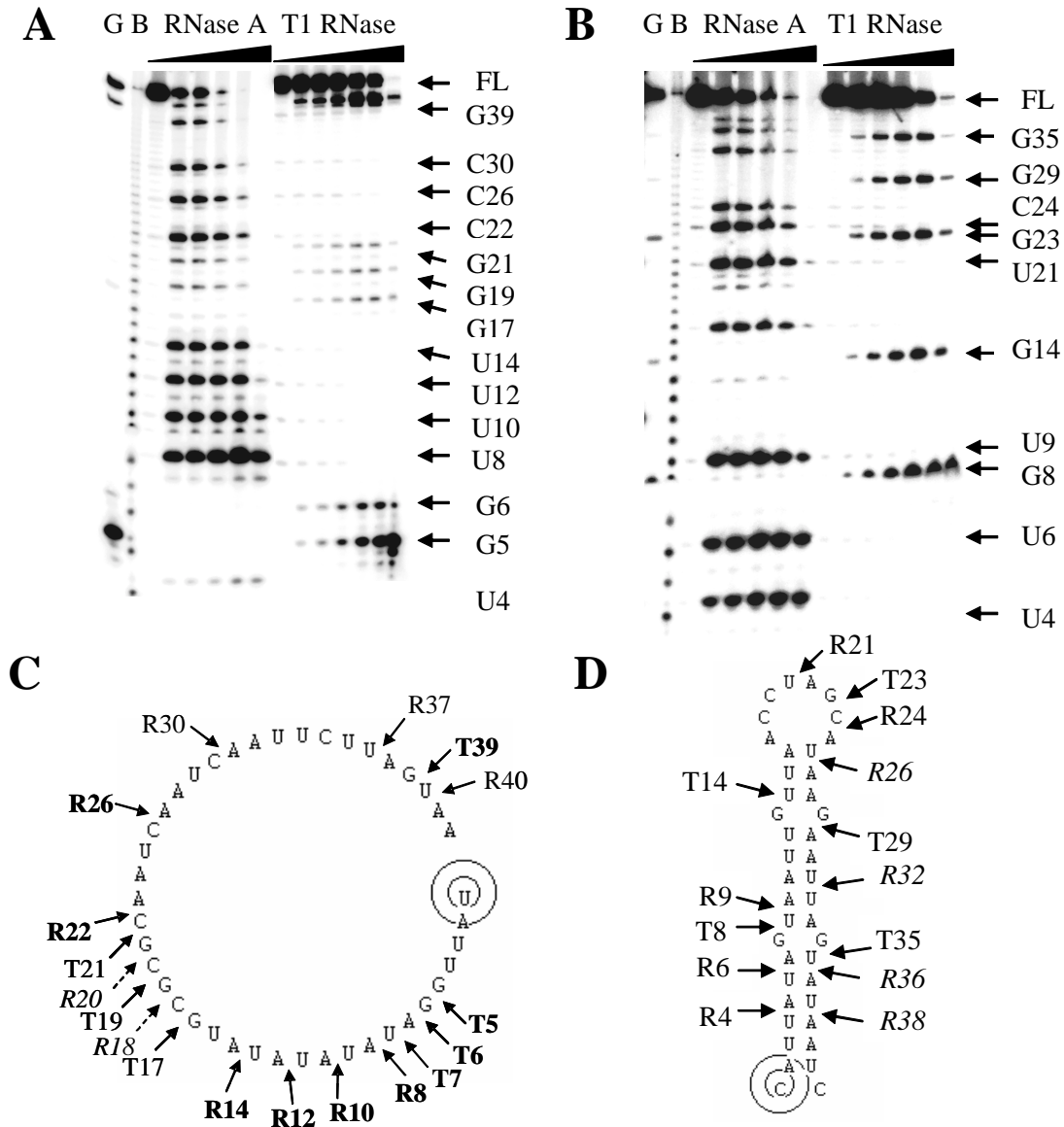
Figure 7: RNase A and T1 RNase mapping of weak structures. Panels A and B show the RNase digestion maps with 0.0rna and 7.5rna, respectively. The T1 digestion ladder is shown in the lane marked G and the base hydrolysis ladder in the lane marked B. The amount of RNase A used was 0, 0.25, 0.5, 1, 2.5 and 5 units. T1 RNase was used at 0, 0.1, 0.5, 1, 5, 10, 50 units. The major cleavages are marked at the right of the gels. FL denotes the full length RNA. The structures shown in C (0.0rna) and D (7.5rna) show the position of the cleavage points. Cleavage points from RNase A digestion are marked with an R and those from T1 RNase digestion are marked with a T. Bold markings indicate predominant cleavage points, normal font indicates mild cleavage points and italics indicate very faint cleavages.
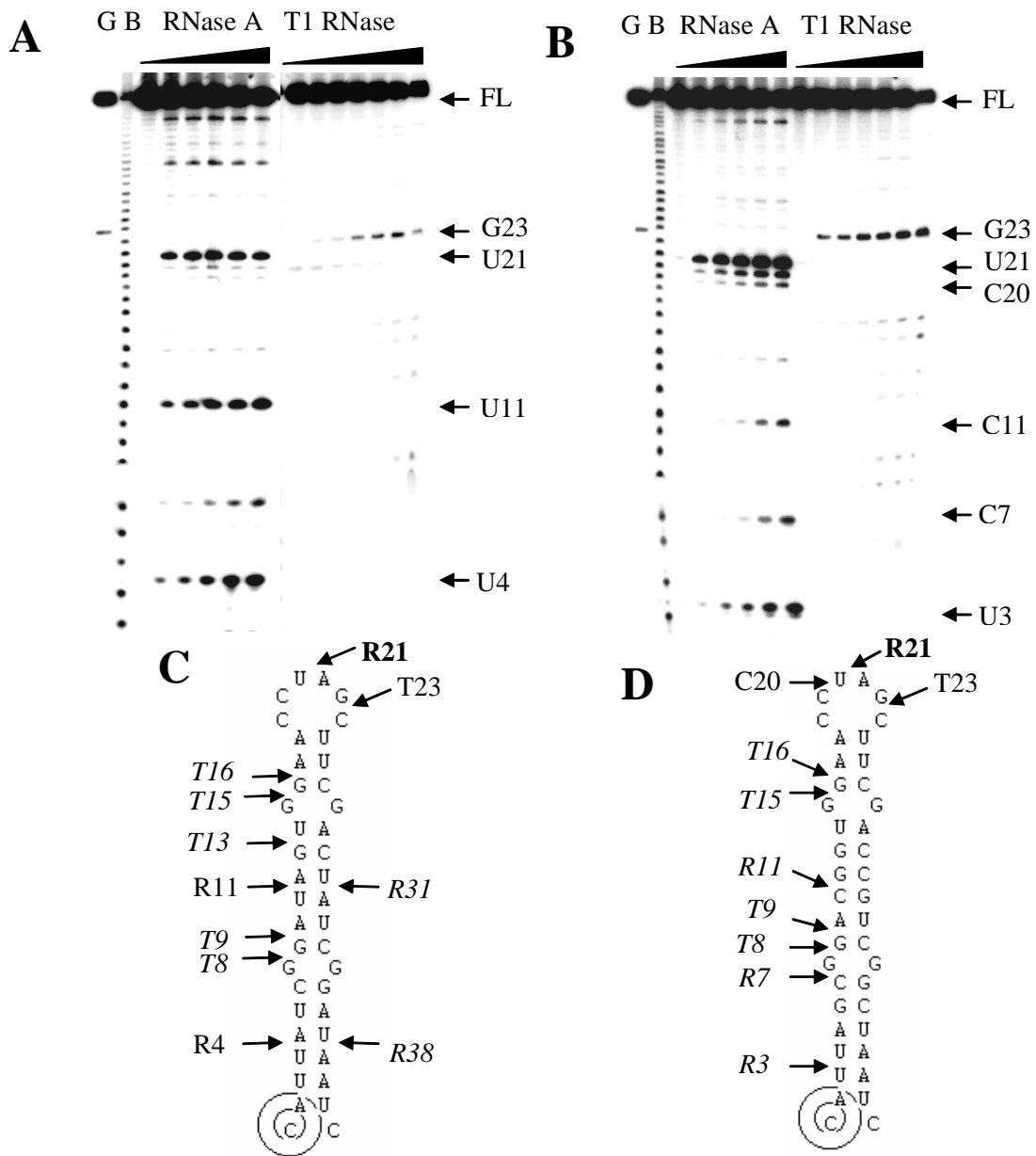
Figure 8: RNase A and T1 RNase mapping of strong structures. Panels A and B show the RNase digestion maps with 16.3rna and 21.7rna, respectively. The T1 digestion ladder is shown in the lane marked G and the base hydrolysis ladder in the lane marked B. The amount of RNase A used was 0, 0.25, 0.5, 1, 2.5 and 5 units. T1 RNase was used at 0, 0.1, 0.5, 1, 5, 10, 50 units. The major cleavages are marked at the right of the gels. FL shows the position of the full length RNA. The structures shown in C (16.3rna) and D (21.7rna) show the position of the cleavage points. Cleavage points from RNase A digestion are marked with an R and those from T1 RNase digestion are marked with a T. Bold markings indicate predominant cleavage points, normal font indicates mild cleavage points and italics indicate very faint

HIV-1 NC titration to determine NC concentration for annealing experiments. A titration was performed to determine the appropriate amount of NC to use in annealing experiments. The titration was performed with 0.0rna. The titration was completed by making separate RNA and DNA reactions and aliquoting them into seven separate tubes each. To each of the fourteen tubes NC was added in the designated concentration and the RNA and DNA were separately preincubated with NC for two minutes. The DNA aliquot was added to the RNA to begin the reaction. The reaction progressed for 15 seconds when stop solution was added. Reactions were run on 15 % non-denaturing gels and autoradiograms were obtained. The NC titration is shown in Figure 9. From the titration 2 µM NC was the concentration chosen for annealing reactions. This concentration was chosen because the annealing was highest with 1 and 2 µM NC. Lower concentrations did not show as much annealing and higher concentrations also showed a slight decrease in annealing. Also, higher concentrations tend to aggregate and therefore the product sticks in the wells. Two µM NC was the highest concentration where there was no aggregate present.

HIV-1 NC enhanced annealing of structured and unstructured RNAs. Annealing assays were completed to examine how HIV-1 NC effected the annealing of unstructured and structured nucleic acids. The end labeled RNA and its complementary DNA were separately preincubated at 37 ^{o}C in the presence or absence of NC for two minutes as indicated in the Methods. The DNA was then added to the RNA reaction mixture and time courses were completed. The final concentration of DNA was twice the final concentration of RNA. The salt conditions in these experiments were chosen to be as close to *in vivo* conditions as possible without interfering with NC:nucleic acid binding

45

A   C   0.25   0.5   1   2   4   8   15

← R:D

← ssRNA

B
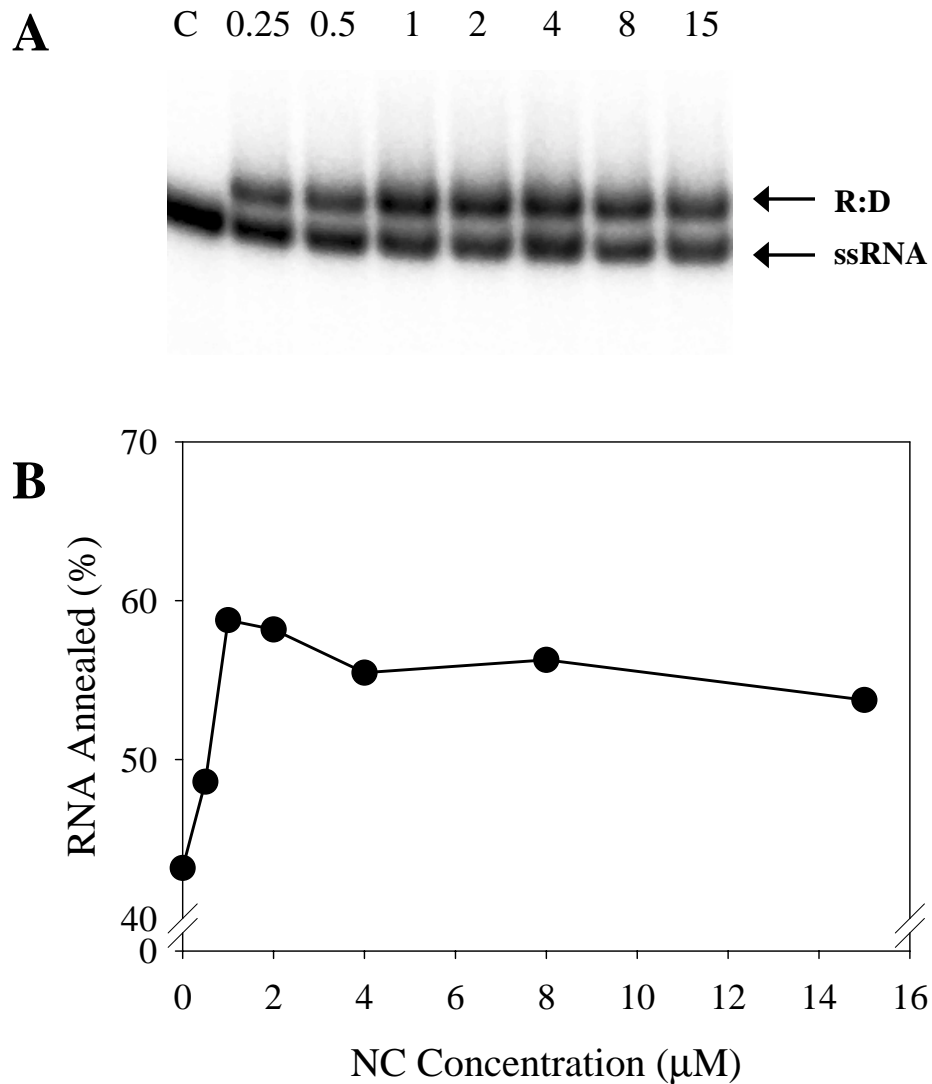
Figure 9: HIV-1 NC titration with 0.0rna. Annealing experiments were done with 0.0rna and 0.0dna in the presence of increasing concentration of NC. The autoradiogram is shown in panel A. Reactions were stopped at 15 seconds. The position of the annealed product (R:D) and the single stranded RNA is shown. The concentration is of NC used in each reaction is shown above each lane in micromolar concentrations. The control lane (C) contained no NC and no DNA. The percent RNA annealed to DNA was obtained and is shown *versus* time in panel B.

in vitro. The KCl concentration in our experiments was 80 mM versus 150 mM normally found in mammalian cells. However to much KCl in our reactions could have prevented efficient NC:nucleic acid binding. We used 100 µM $ZnCl_2$ to provide adequate zinc to maintain zinc finger structure. The $MgCl_2$ concentration (6 mM) was within the normal range for mammalian cells (approximately 10 mM). Tris-HCl was used as a buffer to maintain pH level. Time courses for structures 0.0rna and 7.5rna were run for four minutes. The time was extended to 16 and 32 minutes for 16.3rna and 21.7rna, respectively. The annealing reactions were subject to gel electrophoresis which allowed for the determination of percent of the RNA that annealed to the complementary DNA. Annealed RNA shifted up in the autoradiogram because the double stranded nucleic acid migrated slower through the native gel. One reaction containing just the RNA, DNA and hybridization buffer was heated and slowly cooled and used as a hybrid control which showed where the annealed RNA and DNA ran in the gel (data not shown).

Figure 10 shows autoradiograms of experiments with each of the RNA substrates. NC increased the annealing of RNA to DNA relative to reactions without NC for all the substrates. Note also that for the more strongly structured substrates (16.3rna and 21.7rna), a band was observed above the RNA:DNA hybrid species (marked as "Dimer" on the figure). This species migrating at this position was observed most prominently in control reactions that had RNA alone and likely represents dimerized RNA. Dimers, though less thermodynamically stable than hybrids (ΔG=-34.7 vs. -57.8 for 21.7rna dimer and hybrid, respectively), are still quite stable and would be expected to form. The results from experiments with all the substrates were quantified using a phosphorimager, and the data was displayed graphically.
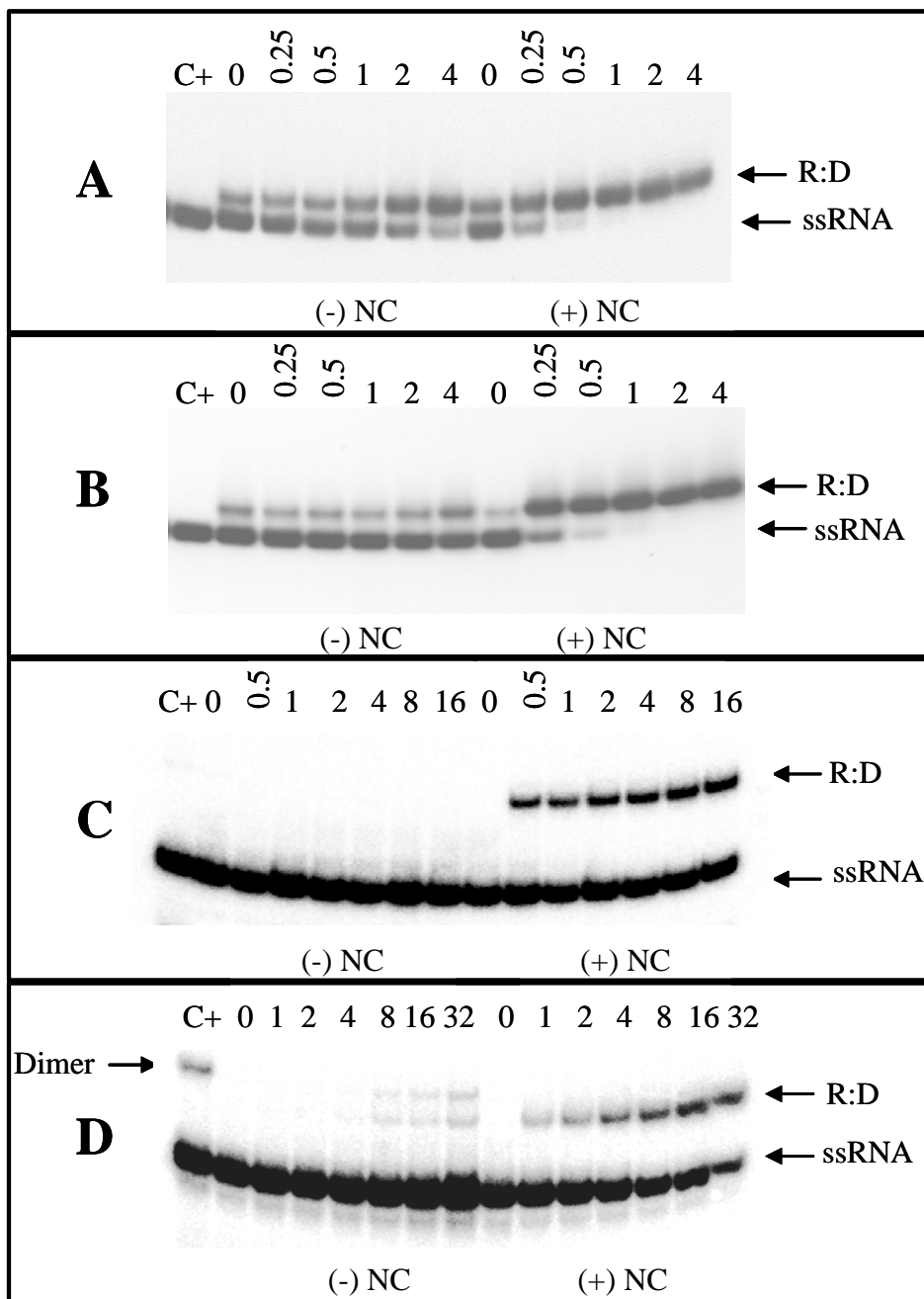
Figure 10: Autoradiograms showing enhanced nucleic acid annealing in the presence of HIV-1 NC. RNA was end-labeled with P^{32} and annealing assays were completed in the presence or absence of NC. Time courses were completed as described in Methods, with time varying for each structure. Time is denoted on each panel. Autoradiograms are as follows: A) 0.0rna. B) 7.5rna. C) 16.3rna. D) 21.7rna. The positions of the single stranded RNA (ssRNA), RNA:DNA hybrid (R:D) and RNA:RNA dimers (dimer) are shown. C+ is the control sample containing RNA and NC but no complementary DNA. Time is in minutes.

The graphs shown in Figure 11 display the percent of RNA annealed in the presence or absence of HIV-1 NC for the weaker structures. Annealing at time zero was only detected for the 0.0rna and 7.5rna. This was probably due to the ease with which these substrates are able to anneal either in the sample buffer or at the beginning of electrophoresis before the RNA and DNA completely migrate into the gel. The annealing reaction for the 0.0rna structure is shown in Figure 11A. There was a clear increase in the RNA:DNA annealing even without NC. Over four minutes annealing increases to 65%. However, when NC was added the rate of annealing was accelerated considerably. At four minutes the annealing of the RNA and DNA increased substantially in reactions with NC compared to those without NC. When NC was included, maximum annealing was nearly reached by one minute. These results demonstrate a role for NC in annealing that is separate from a helix destabilizing (unwinding) activity since this substrate presumably has no helices that require unwinding.

Experiments completed with 7.5rna are shown in Figure 11B. In this case stimulation by NC was more evident than with 0.0rna. Annealing increased to 98% with NC at 4 min compared to 35% in the absence of NC. Comparing Figure 11B with Figure 12A and B, the graphs for the stronger structures, it was evident that annealing in the absence of NC was much greater for 7.5rna than for the stronger structures. In reactions with 7.5rna but without NC there was a considerable amount of annealing that occurred even over the shorter times used, indicating that this weak stem loop does not need substantial unwinding activity to hybridize to the complement.

Figure 12 shows the graphical representation of the strand annealing for the strongest structures. For each of these structures a substantial amount of annealing did

49
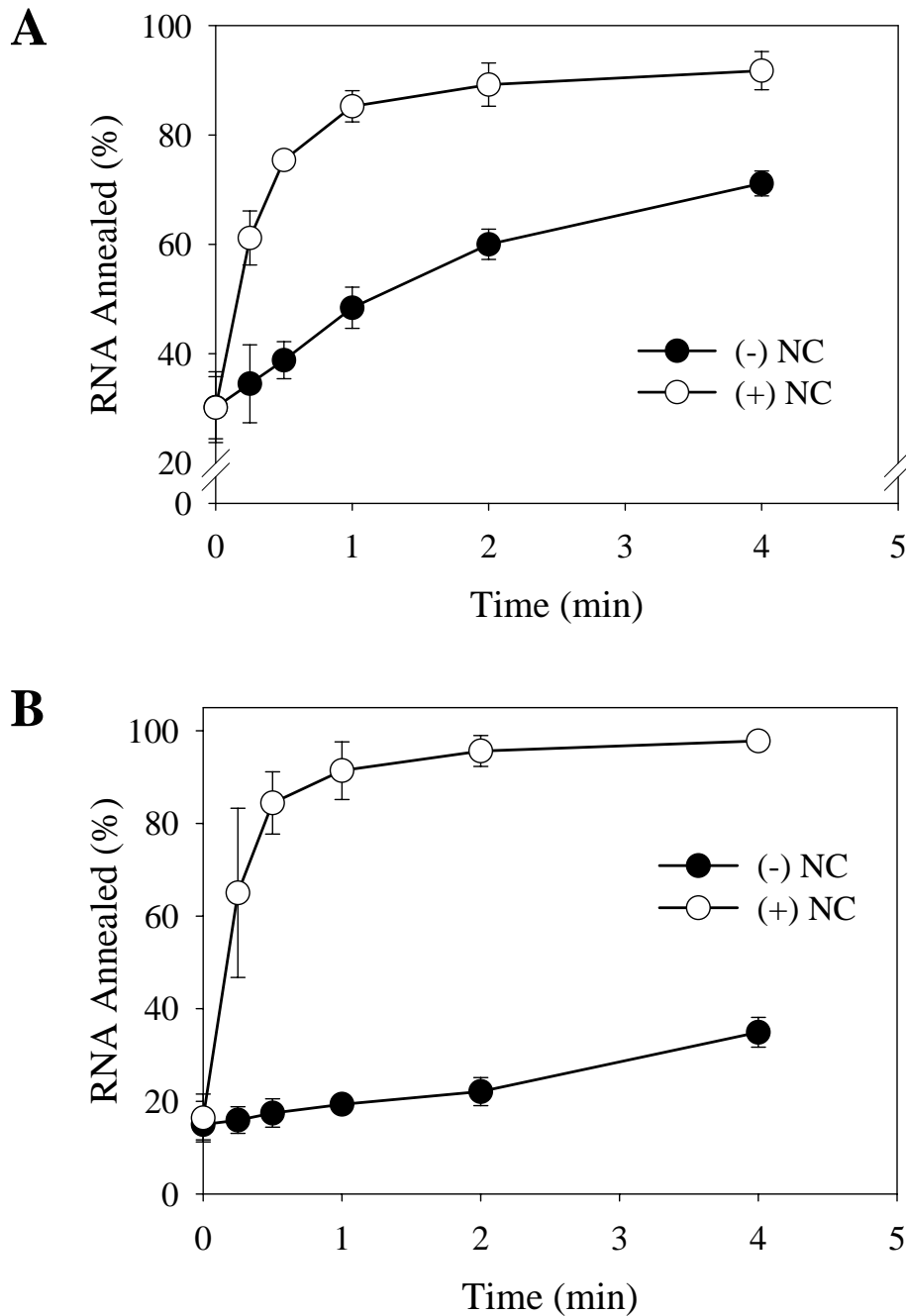
Figure 11: NC enhances the rate of nucleic acid annealing in unstructured and weakly structured sequences. Autoradiograms shown in Figure 10 were quantified and data was displayed graphically. Graphs represent the average of three experiments and the standard deviation is shown with error bars. Annealing assays are shown with (open circles) and without (filled circles) NC for A) 0.0rna, B) 7.5rna.
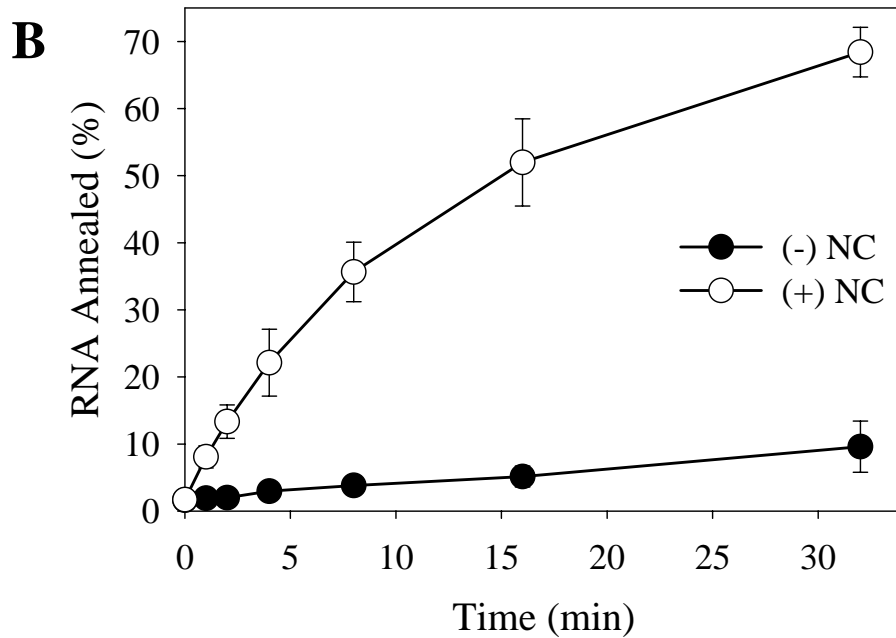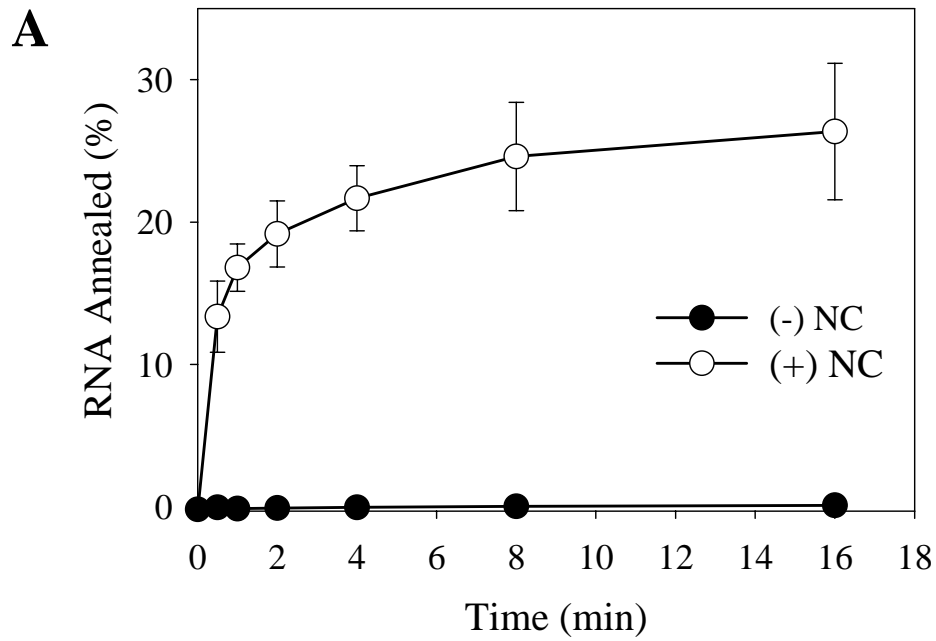
Figure 12: NC enhances the rate of nucleic acid annealing in assays with strong structures. Autoradiograms shown in Figure 10 were quantified and data was displayed graphically. Graphs represent the average of three experiments and the standard deviation is shown with error bars. Annealing assays are shown with (open circles) and without (closed circles) NC for A) 16.3rna, B) 21.7rna.

not occur in the absence of HIV-1 NC. Annealing assays for 16.3rna and 21.7rna were

run over longer periods of time due to slower rates of annealing. In each case the

structure prevented high levels of annealing in the absence of NC. However, the addition

of NC allowed the annealing of complementary RNA and DNA, presumably because NC

aided in the unfolding of secondary structure. The level of annealing was less than was

observed with the weaker structures suggesting that the rate limiting step for these

substrates was destabilization of intra- or intermolecular interactions. Though NC was

clearly able to stimulate this step, it was still the slow step in the reaction for all but the

two weakest structures used.

Mutant NC proteins display differing effects on annealing, depending on the type

of nucleic acid structure. We next wanted to investigate what roles the zinc fingers of

NC have in nucleic acid annealing. Strand annealing assays were completed with three

mutant NC proteins and wild type NC derived from the NL4-3 strain of HIV. Wild type

NC from the NL4-3 strain had similar annealing activity to the wild type NC from the

ARV strain (compare Figures 11, and 12 with 14 and 16). As was previously noted the

N-terminal and C-terminal zinc fingers of NC are not biologically equivalent (58). It was

therefore possible that the fingers may have differential activities with respect to the

hybridization (stimulation of hybrid formation in the absence of structure) *vs.* unwinding

(helix destabilization) activities of NC. Three NC mutants, 1.1 NC, 2.2 NC and 2.1 NC,

were analyzed. The first mutant, 1.1 NC, contains two N-terminal zinc fingers, the first

in its native position, and the second in the position of the native C-terminal zinc finger.

Two copies of the C-terminal zinc finger are present in the mutant 2.2 NC, and 2.1 NC is

a finger switch mutant where the positions of the zinc fingers are switched. Annealing

assays were completed for each structure in the presence of the mutant NC proteins. Autoradiograms of mutant assays with 0.0rna and 7.5rna are shown in shown in Figure 13. It was evident that the mutants had differential effects on annealing of the different types of structures. Graphical representations of mutant annealing assays with unstructured and weakly structured sequences are shown in Figure 14.

Results of annealing assays completed with 0.0rna showed all mutants enhanced annealing of this substrate. Both of the mutants containing the original C-terminal zinc finger clearly display annealing comparable to wild type. Our quantification also indicated that 1.1 NC significantly enhances annealing of 0.0rna and 0.0dna. However Fig 13A shows that the annealed products did not form a discrete band, instead shifting up to form two smeared bands. These products were both quantified as annealed product. These results indicate that all mutants contain hybridization activity which enhances the annealing of unstructured sequences.

Though 7.5rna is weakly structured the results of the mutant assays with this substrate differed from the results with 0.0rna. The annealing assays completed in the presence and absence of HIV-1 NC with 7.5rna showed that annealing will occur without NC, but at a much-reduced rate compared to reactions with wild type NC. Reactions completed in the presence of each mutant showed an increase in the rate of annealing compared with reactions in the absence of NC. However, 2.2 NC had a much lower ability to enhance annealing with 7.5rna as compared with 0.0rna. On this substrate both 1.1 and 2.1 NC were considerably better than 2.2 NC while they were only slightly below wild type. The results suggest that the first zinc finger may be important in unwinding nucleic acid strands, even for relatively weak stem loop structures.

53

Figure 13: Autoradiograms of annealing assays with weak structures in the presence of HIV-1 NC mutants. Annealing assays were performed with zinc finger mutant NC proteins. Autoradiograms showed the presence of single stranded RNA (ssRNA) and RNA:DNA hybrids (R:D). A) 0.0rna annealing assays with wt and 2.1 NC. B) 0.0rna annealing assays with 1.1 and 2.2 NC. C) 7.5rna annealing assays with wt and 2.1 NC. D) 7.5rna annealing assays with 1.1 and 2.2 NC. The time is labeled in minutes and control labels are as follows: C_a: Control with wild type NC, C_b: Control with 2.1 NC, C_c: Control without NC, C_d: control with 1.1 NC and C_e: Control with 2.2 NC. Control reactions were completed for the full length of the time course without complementary DNA.

Figure 14: HIV-1 NC mutant enhanced nucleic acid annealing assays with weakly structured substrates. Strand annealing assays completed with HIV-1 NC mutants were quantified and displayed graphically. Graphs are the average of three experiments. Error bars indicate the standard deviation. The legend in A indicates which mutant was used in each experiment. Graphs are as follows: A) 0.0rna, B) 7.5rna.

To test the importance of finger one in nucleic acid unwinding annealing assays were completed with the stronger structures and mutant NC proteins. Results of the annealing assays with mutant NC proteins and strongly structured nucleic acids are shown in Figures 15 and 16. Consistent with the trend observed with 7.5rna, the ability of 2.2 NC to enhance hybrid formation decreased with an increase in the strength of the nucleic acid structure. This mutant stimulated hybridization to the least extent with each of the structured RNAs. In fact, with the strongest structure there was no significant difference between annealing reactions without NC and with 2.2 NC. In contrast 1.1 and 2.1 retained some activity on all the structures with 1.1 showing an increased rate of activity on the more structured substrates. The results further support the hypothesis that finger one is more important for unwinding. Also, the results with the switch mutant, 2.1 NC, indicate that the location of the fingers is also important. If this were not the case then 2.1 should have had wild type levels of activity since it possesses both fingers. This mutant did have wild type levels of annealing activity as judged from assays with 0.0rna, but unwinding was reduced. This may suggest that the context of finger one is particularly important to its function while finger two is less context-dependent.

RNA:DNA annealing detected by FRET supports gel based assays. In order to confirm the above results by a second technique, fluorescence quench experiments were also completed using 21.7rna and 21.7dna. The RNA was synthesized with a FAM molecule on its 5' end. This molecule fluoresces when excited by light at 494 nm. However, if this molecule comes within close proximity to a DABCYL molecule, which is a dark quencher, the fluorescence will be quenched. The 21.7dna in these assays contained a 3' DABCYL molecule. Annealing of the RNA and DNA would therefore
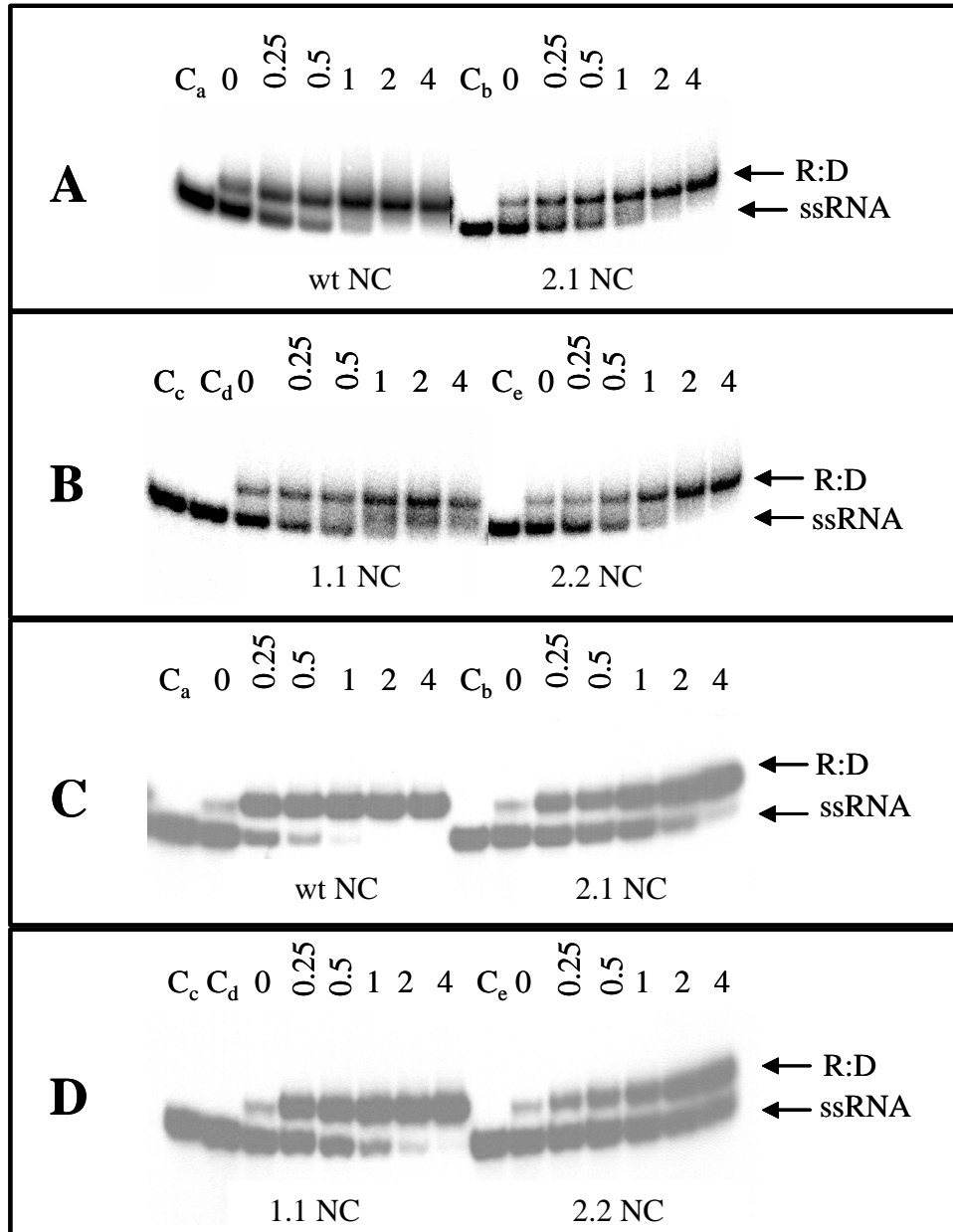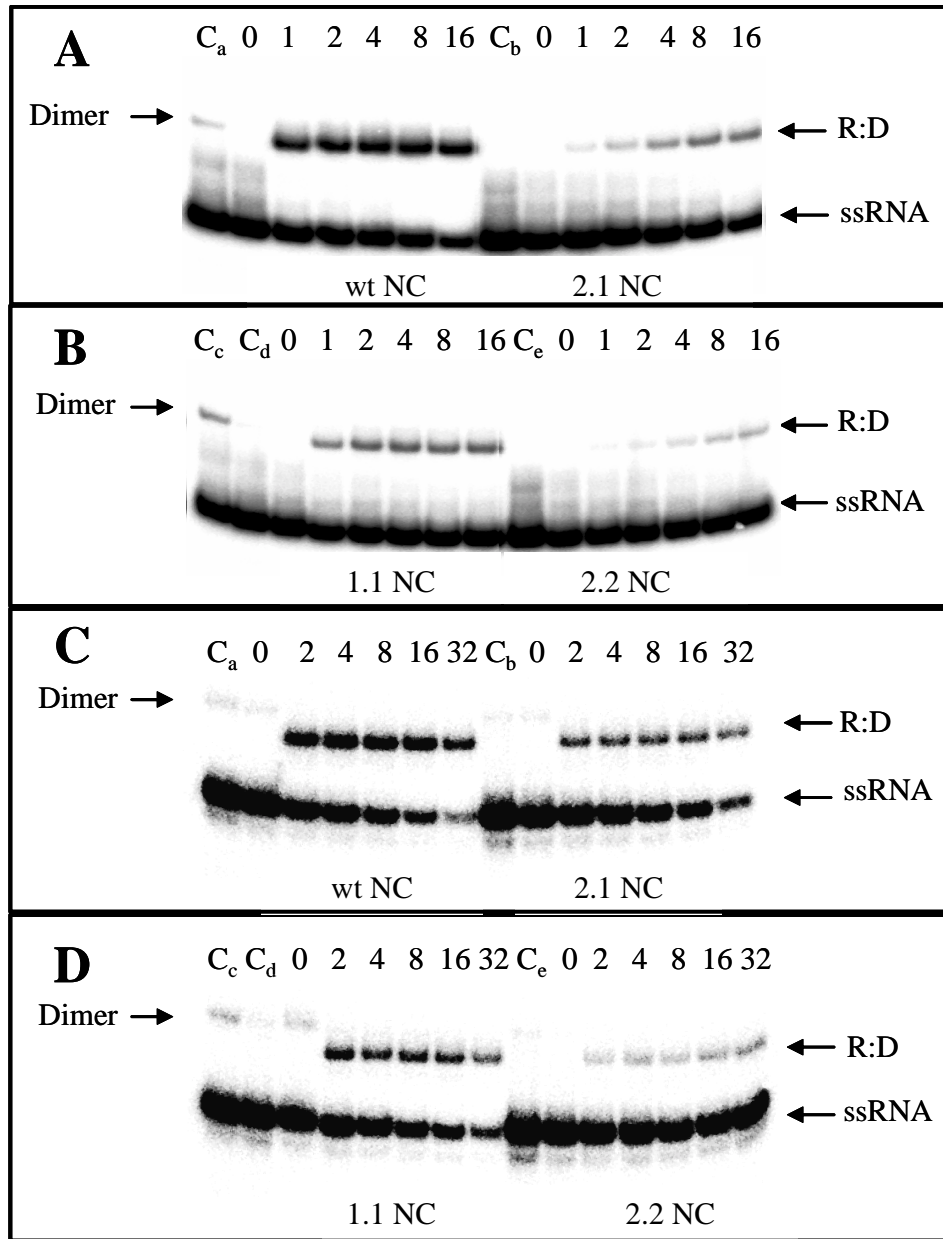
Figure 15: Autoradiograms of annealing assays with strong structures in the presence of HIV-1 NC mutants. Strand annealing assays were performed with zinc finger mutant NC proteins. Autoradiograms showed the presence of single stranded RNA (ssRNA), RNA:DNA hybrids (R:D) and RNA:RNA dimers (Dimer). A) 16.3rna annealing assays with wt and 2.1 NC. B) 16.3rna annealing assays with 1.1 and 2.2 NC. C) 21.7rna annealing assays with wt and 2.1 NC. D) 21.7rna annealing assays with 1.1 and 2.2 NC. The time is labeled in minutes and control labels are as follows: C_a: Control with wild type NC, C_b: Control with 2.1 NC, C_c: Control with no NC, C_d: control with 1.1 NC and C_e: Control with 2.2 NC. Control reactions were completed for the full length of the time course without complementary DNA.

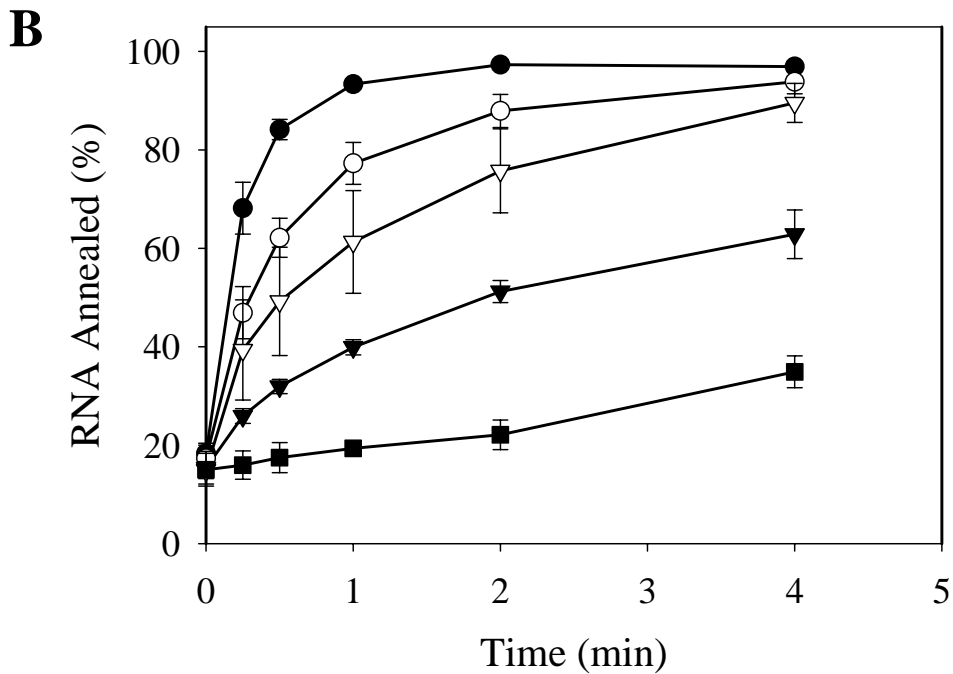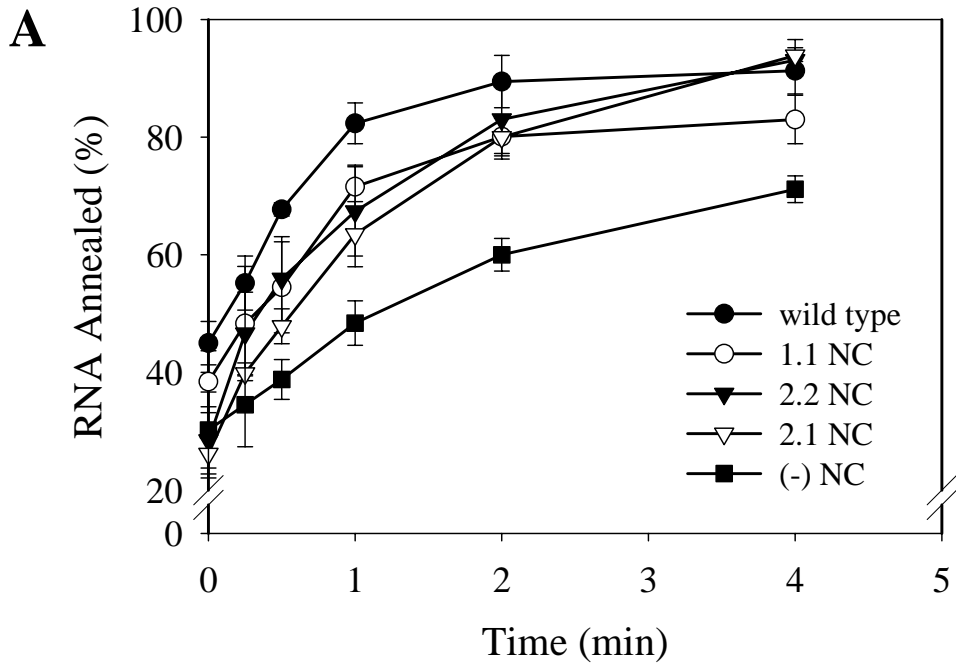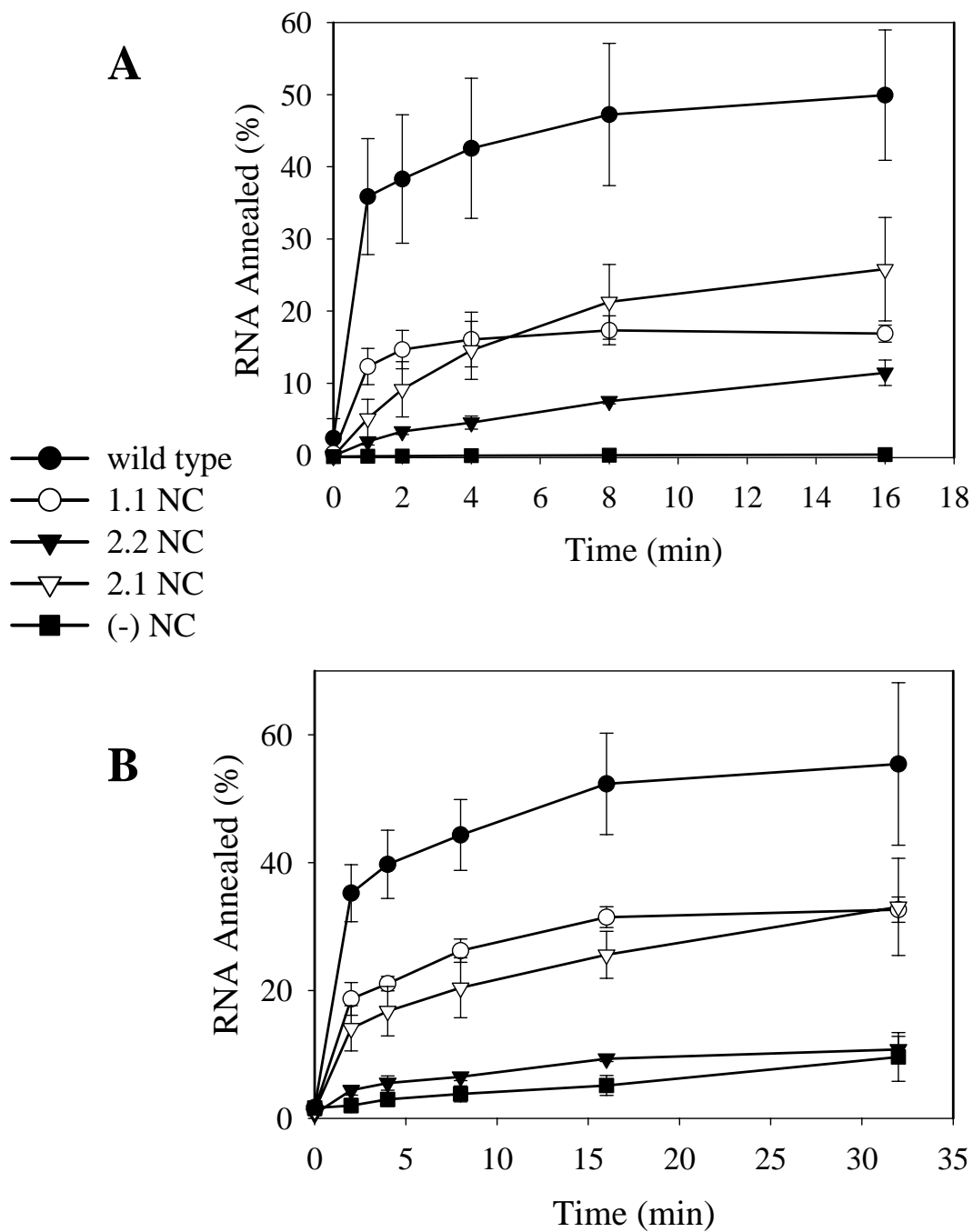**Figure 16: HIV-1 NC mutant enhanced nucleic acid annealing assays with strongly structured substrates.** Strand annealing assays completed with HIV-1 NC mutants were quantified and displayed graphically. Graphs are the average of three experiments. Error bars indicate the standard deviation. The legend indicates which mutant was used in each experiment. Graphs are as follows: A) 16.3rna, B) 21.7rna.

lead to a quenching of the FAM fluorescence. Fluorescence decay experiments were completed in the absence of NC and in the presence of wild type, 1.1 and 2.2 NC. A graph showing fluorescence decay is shown in Figure 17. Wild type NC showed a large and relatively rapid decrease in fluorescence, whereas no major decrease in intensity was shown without NC. This shows that the addition of NC increases the amount of FAM molecules that are quenched by coming into close proximity to the DABCYL molecule, more specifically by the annealing of the RNA and DNA. The mutants, 1.1 and 2.2 NC, showed very different intensity ratio profiles. The intensity ratio with 2.2 was similar to that observed without NC while 1.1 decreased less than wild type but considerably greater than 2.2. These assays showed that 1.1 NC was able to enhance a decrease in fluorescence, indicating an increase in annealing while 2.2 NC showed no stimulation of annealing. This data clearly supports the findings from experiments performed in the gel-based annealing assay.
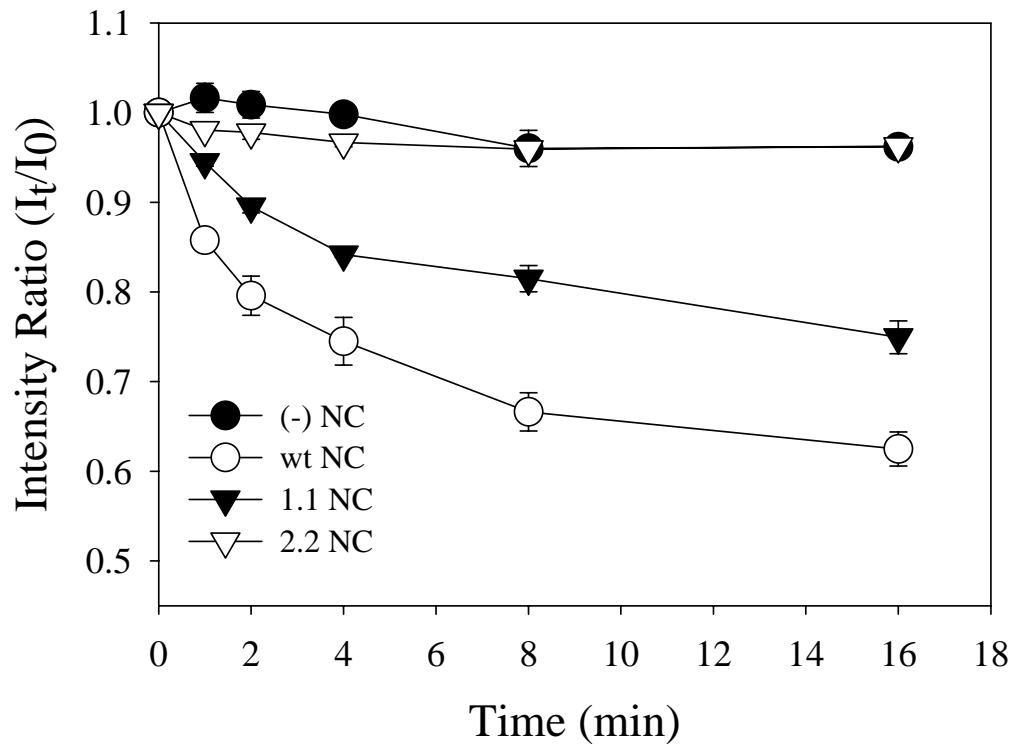
Figure 17: Annealing of 21.7rna and 21.7dna detected by fluorescence resonance energy transfer analysis. 5' FAM-labeled 21.7rna was incubated with 3' DABCYL-labeled 21.7dna in the presence and absence of wild-type and NC mutants. The intensity ratio was determined as described in the Methods section. Intensity ratio is plotted versus time (in minutes). The legend indicates which reaction is plotted in each line. Each line is the average of three experiments and the standard deviations are shown by error bars.

Discussion

HIV-1 NC has been shown to contain chaperone activity that can aid in the many nucleic acid annealing steps that occur during the viral life cycle. This protein contains two zinc fingers, which have been studied extensively. The zinc finger motifs cannot be substituted for one another in strand transfer assays that mimic retrovirus recombination and mutants with alterations to the zinc fingers have been shown to reduce or eliminate viral infectivity (58, 64). However, the role of the individual zinc fingers in nucleic acid annealing is unclear. In this report, we show that the hybridization and unwinding components of NC annealing activity are proportioned unequally, with the latter being more prevalent in finger one. Finger 2, although apparently possessing little unwinding activity, clearly enhances the overall activity of NC since both fingers in the appropriate context are required for full annealing of strongly structured RNAs.

Annealing assays presented here with wild type HIV-1 NC showed that NC enhanced the formation of hybrids between complementary RNAs and DNAs that contained varying degrees of secondary structure. Structured sequences require unwinding to be available for annealing to their complement. In the presence of HIV-1 NC the rate of annealing was dramatically increased for all the substrates that were tested. NC even enhanced annealing with the 0.0rna substrate that had no predicted structure (Figs. 10 and 11). This result strongly supports a role for NC in enhancing hybridization of nucleic acids irrespective of secondary structure. Golinelli *et al*. showed that NC did not stimulate annealing of unstructured sequences (55). However, their sequences were smaller, all less than 31 nucleotides. NC binds with lower affinity to

61

shorter oligonucleotides and this could be a reason for the discrepancy between these results (109). In addition, when we did experiments with oligonucleotides less than 30 nucleotides in length NC inhibited annealing, probably because the annealing of complements was not favorable (data not shown).

NC also stimulated annealing using highly structured substrates (Figs. 10 and 12). On structures like 16.3rna and 21.7rna destabilization of the strong secondary structures limit the rate of annealing and NC's unwinding activity would be required for faster annealing. This suggests that NC possesses two related activities, hybridization and helix-destabilizing activity. It is possible that there is only one activity that stimulates hybridization and then unwinds secondary structure through strand invasion (114). However, others have clearly shown that NC destabilizes secondary structure in tRNAs and other nucleic acids in the absence of a complementary strand (65, 79). This provides clear evidence for unwinding activity while enhanced hybridization in the absence of structure shown here support a separate annealing activity. In addition, other investigators have completed fluorescence quenching experiments to examine strand-exchange of oligonucleotides. These experiments showed that HIV-1 NC highly enhanced the rate of annealing of complementary strands while moderately enhancing unwinding of the helix DNA (137). Again, this argues for two related but separate activities. Our results show that these activities can actually be distinguished using specific NC mutants.

As discussed in the introduction there are many reports where NC mutants were used to determine what amino acid residues are important for nucleic acid annealing. Experiments showed that the binding of tRNA^{Lys,3} to the primer binding site, and

dimerization of sequences containing the Ψ-site, required only peptides outside of the zinc fingers which included the basic backbone residues (30, 105). It is known that the basic residues of NC are vital for RNA binding *in vitro* (119). Therefore the decreased amount of annealing in the presence of mutants lacking basic backbone residues could be primarily due to a lack of NC binding to the RNA. Lapadat-Tapolsky *et al*. (83) completed annealing experiments with oligonucleotides containing sequences from the R region of the HIV-1 genome. This report also indicated that it is the backbone residues that are important for annealing rather than the zinc fingers. In agreement with this, binding assays completed with NC mutants containing only the zinc finger residues showed a highly reduced binding affinity when compared with wild type (137). Therefore, there is a clear role for the basic amino acids in the backbone of NC in annealing, however, these reports do not preclude involvement of the zinc fingers in annealing or other NC activities such as helix destabilization. Studies conducted with NC mutants that contain the zinc fingers but have alterations that affect the structure or sequence of the fingers show clear effects on annealing activity of the protein. Single molecule stretching experiments completed with 2.1 NC and 1.1 NC showed that it was important to have the N-terminal zinc finger in its native position to enhance the helix-coil transition (139). These findings are in agreement with the work presented here. We have shown that the N-terminal zinc finger must be present for annealing of the strongest nucleic acid structures, and the highest activity is observed when this finger is in its native position (mutant 1.1 or wild type, see Figs. 15, 16, and 17). This is also in accordance with annealing assays completed with the TAR region. This region of the genome is highly structured and annealing TAR RNA to complementary DNA was only

substantial in mutants that contained the N-terminal zinc finger in the position observed in wild type NC (64). It is important to note that of the viruses containing the NC mutations used in this report, only mutant 1.1 was infectious, though there was a substantial reduction in infectivity when compared with wild type (58, 59, 64). In addition, viruses containing 1.1 NC reverted to a wild type phenotype over time clearly indicating that this configuration is sub-optimal. Taken together these results indicate that the N-terminal zinc finger of HIV-1 NC is important for the unfolding of strong secondary nucleic acid structure and double stranded DNA unwinding while the basic backbone and finger 2 probably play a role in enhancing duplex formation by stimulating hybridization and perhaps augmenting the activity of finger 1. It is clear from the infection experiments that the role of finger 2 is either non-essential or can be partly compensated by finger 1 and the protein backbone.

Finger 2 of NC has not previously been shown to be important for annealing, although annealing assays have not been conducted with unstructured RNAs and mutant NCs. The results we obtained indicate that finger 2 may enhance the hybridization of unfolded RNAs, though it cannot enhance the unfolding of structured RNAs. Therefore, it is possible that this zinc finger is involved in enhancing the collisions and binding of complementary nucleic acids. It should be noted that it is also possible that this effect is due solely, or at least mostly to the basic backbone residues exterior to the zinc fingers in 2.2 NC. However, annealing assays completed with structured nucleic acids and 1.1 NC showed that this protein did not enhance annealing as much as wild type NC. This result supports a possible role for finger 2 in the annealing step. However, other NC mutants will have to be examined to substantiate this hypothesis.

While HIV-1 NC contains two zinc fingers, MuLV NC protein has only one. For -ssDNA transfer to occur in MuLV replication, the only structure that must unfold is much weaker than the HIV-1 TAR hairpin required for –ssDNA transfer in HIV (132). Also, Williams *et al.* (139) point out that retroviruses which contain long repeat regions with numerous hairpin structures have NC proteins with two zinc fingers whereas those with short repeat regions and minimal structure have NC proteins with only one zinc finger (11, 56). Consistent with this is the possibility that a second finger (finger 1) in HIV is required in order to enhance the unwinding activity of NC.

It is not clear what causes the apparent differences between the activities of the two zinc fingers. There are five amino acid residues that differ between the N- and C-terminal zinc fingers of HIV (NL4-3 strain). These include from finger 1 to 2: phenylalanine to tryptophan, asparagine to lysine, isoleucine to glutamine, alanine to methionine, and asparagine to aspartate. Nucleic acid annealing would be favored by a minimization of the electrostatic repulsion between the acidic phosphate groups in complementary strands. HIV-1 NC is highly basic molecule with a net charge of +13 that would be capable of minimizing that repulsion. The apparent unwinding activity in the N-terminal zinc finger could be related to the presence of more hydrophobic residues in this finger. These residues may play a role in disrupting hydrophobic base stacking interactions by associating with the hydrophobic rings of the bases. We are currently designing mutants in which the N-terminal zinc finger amino acids will be incrementally replaced with those found in the C-terminal zinc finger. These mutants will be used in annealing assays with structured RNA sequences to determine which residues are important for the nucleic acid unfolding capability of the N-terminal zinc finger.

The results presented here indicate that the C-terminal zinc finger of NC has an accessory role in annealing, while the N-terminal zinc finger has a role in unfolding of secondary structure. Understanding the methods by which NC displays chaperone activity, and what roles the zinc fingers have in this activity could be very useful to the design of therapeutic agents that could inhibit the function of NC throughout the viral life cycle.

Chapter 3

Analysis of the mechanism of NC enhanced strand exchange.

Introduction

HIV is known to undergo the process of recombination during reverse transcription. This process can lead to genetic diversity when recombination occurs between two heterozygous genomes. It has been shown that HIV can undergo recombination as many as 2-3 times per replication cycle, or at a rate of 4×10^{-4} switches/base pair (147). Most recombination events in retroviruses occur by a mechanism referred to as strand transfer or template switching. This occurs when the nascent DNA switches from the original template RNA to another RNA or another region of the same template RNA and DNA synthesis continues in the new region.

Recombination has been shown to be a method HIV uses for building up drug resistance. Mutations in HIV-1 RT can render HIV resistant to zidovudine treatment. When two separate viruses were used to infect cells that were later treated with zidovudine progeny virion were resistant to the treatment. Sequencing results showed that recombination had taken place to create viruses that had the mutations necessary in RT for zidovudine resistance (77). Also in regions where two or more subtypes of HIV exist it has been shown that intersubtype recombination can occur in individuals who have been infected with both types (52, 117). When these recombinants persist they can become CRFs which are stable viruses with segments from two or more subtypes (107). This clearly presents a problem for the development of a successful HIV vaccine or drug treatments (85). The more variation there is among circulating forms of HIV the harder it

will be to produce successful therapies. It is therefore imperative that this process be thoroughly studied to understand what factors enhance strand transfer and how these factors can be prevented.

HIV-1 NC has been shown to enhance strand transfers. This protein is a nucleic acid chaperone which can enhance the rearrangement of nucleic acids to encourage formation of the most stable hybrid. There are two obligatory strand transfers that occur during the replication of HIV. The first, occurs when the -ssDNA has been synthesized from the PBS to the 5' end of the RNA template. The TAR hairpin is a strong hairpin sequence at this end of the genome and after or during the copying of this RNA sequence –ssDNA transfer occurs. HIV-1 NC has been shown to enhance –ssDNA transfer and minus strand transfer in general (52, 61, 62, 68, 80, 117, 146). In addition to directly stimulating the –ssDNA transfer event, NC also enhances transfer by inhibiting self priming which occurs when the TAR DNA (DNA sequence copied from the TAR RNA) has been synthesized and RT continues to synthesize the DNA using the DNA itself as a template (61). Self-priming (also referred to as "fold-back synthesis") can occur because the TAR DNA is in a hairpin structure after it has been copied. Therefore when RT reaches the base of the hairpin it uses the DNA sequences as a template. This results in "dead end" products that cannot undergo transfer. However, HIV-1 NC can inhibit this process when a complementary nucleic acid is present in the reaction that the DNA can switch to and on which synthesis can continue (43). In fact, even if the complementary oligonucleotide only binds to the end of the TAR hairpin, and does not have sequences extending past the 3' end of the DNA self-priming is still inhibited.

A FRET based assay has shown that NC enhances the unfolding the TAR DNA in order to enhance minus strand transfer and this unfolding occurs more readily with NC in the presence of a complementary nucleic acid (68). The TAR hairpin contains terminal bulges which have been shown to be important for NC unfolding of this hairpin (9), indicating that the nucleic acid properties are also important to NC enhancement of – ssDNA transfer. Lastly it has also been shown that the zinc fingers, particularly finger one are important for –ssDNA transfer. When the zinc finger structure is disturbed by using mutant NC proteins which do not bind zinc, NC's ability to enhance transfer is inhibited (62).

The second obligatory strand transfer that occurs during HIV replication is +ssDNA transfer. This occurs when the +ssDNA which is being copied on the minus strand DNA template reaches the 5' end of that template and switches to the PBS site of minus strand DNA to continue synthesis (see Fig. 4). HIV-1 NC has also been shown *in vitro* to enhance this process. It does so by first enhancing HIV-1 RT RNase H activity (102), which cleaves the $tRNA^{Lys3}$ from the 5' end of the template DNA. NC enhances the dissociation of the $tRNA^{Lys3}$ from the plus strand DNA and enables this region of the DNA to anneal to the PBS of the minus strand DNA. NC has been shown to enhance the annealing of the PBS and +ssDNA as well (142). Again, when NC mutants were used that were not able to bind zinc, NC was not able to enhance the removal of the tRNA from the minus strand DNA (62).

The last type of strand transfer has not been shown to be obligatory to successful HIV replication. However, internal strand transfer (transfers occurring during minus strand DNA synthesis within the internal regions of the genome) is crucial for generating

diversity among these viruses. Internal transfer has been an area of extensive research because of the importance it plays in allowing the virus to evade effective treatment. After –ssDNA transfer synthesis of the minus strand continues on the RNA genome along with RNase H degradation of the RNA. At some point DNA can dissociate from the original template and bind to a homologous region on the second genome in the virion to complete synthesis. If the two genomes are not identical this can result in a "chimeric" minus strand containing information from both genomes but identical to neither. This would result in new genetically diverse progeny viruses. The genome on which DNA synthesis initiates is referred to as the "donor" while the genome to which the DNA transfers is the "acceptor". These events can be mimicked *in vitro* using short segments of the genome as donor and acceptor.

Although internal transfers can occur essentially anywhere within the genome, certain hot spots have been identified, including the R and *env* regions (94, 107). The DIS sequence has also been shown to enhance strand transfer because the dimerization of the genomes brings the acceptor into close proximity with the donor:DNA complex (7). Also, RT pausing (temporary stalling of DNA synthesis that generally occurs at secondary structures with the genome) can enhance strand transfer. This presumably occurs because the paused RT can catalyze extensive degradation of the RNA template near the pause site which can lead to dissociation of the donor and DNA. Dissociation may also be facilitated by the acceptor binding to the DNA upstream of the pause site. The DNA can then anneal to the acceptor for completion of synthesis (35, 37, 143). It has also been shown that DNAs with terminal 3' mismatches bound to a donor RNA can enhance transfer of the DNA to the acceptor (38, 39). This may result from the mismatch

producing an "artificial" pause due to the relatively slow rate of mismatch extension by RT. The pause by RT can then lead to transfer of the DNA.

The acceptor has been shown to be important for transfer, and in some systems it will invade the synthesis of the DNA on the donor RNA by binding to the DNA after RNase H cleavage of the template RNA (7, 35, 95). This initiates transfer of the DNA from the donor to the acceptor. Also, specific structures that can form in the acceptor have been implicated in initiating invasion of the donor-DNA complex (80, 95). Finally, the overall folding of the acceptor can influence the extent of strand transfer. In general, less structured acceptors can more easily associate with the DNA and this can increase the efficiency of strand transfer in these regions (32). Taken together these results indicate that there are a number of nucleic acid properties that affect the efficiency of strand transfer.

In addition to the nucleic acid properties that can effect the efficiency of strand transfer, HIV-1 NC has also been shown to enhance internal strand transfer (33). Experiments have demonstrated that the structure of the acceptor is modified by NC binding and this in turn enhances strand exchange (98). NC can promote interactions between the DNA and the acceptor that stimulate transfer by accelerating binding of the DNA and acceptor or promoting invasion of the DNA-donor complex by the acceptor (80, 95, 99, 146). Stimulation of RNase H activity by NC can also promote internal transfers similar to what was shown for +ssDNA.

NC zinc fingers are important for a number of different steps in internal strand transfer. Some experiments indicate that the N- and C-terminal regions of the protein, exterior to the zinc fingers, are important for NC enhancement of extension of strand

transfer DNA, but that the zinc finger regions were important for the annealing of DNA and acceptor RNA (69). Guo *et. al.* showed that the first zinc finger is important for RNase H activity and annealing, both components leading to a more efficient transfer process (64). Also many experiments examining the chaperone activity of HIV-1 NC have shown the importance of finger 1 in nucleic acid annealing (64, 66, 111, 139, 140), and our previous results have shown finger 2 plays an accessory role in annealing (66).

There are two current models proposed for copy choice recombination (35). The donor dissociation model, states that the donor RNA first dissociates from the nascent DNA before transfer occurs. A schematic depicting this model is shown in Figure 18. In this model the donor RNA and nascent DNA within the donor-DNA complex dissociate. After dissociation the DNA anneals to the acceptor RNA. RT then associates with the complex and synthesis continues using the acceptor RNA as the template. In this model the acceptor plays a passive role in strand transfer. Early results indicated that strand transfer mediated by pause sites may occur by the donor dissociation mechanism, though more recent work has shown in experiments with strong hairpins, the acceptor does play a role in stimulating transfer (35, 37, 143). Presumably dissociation would be enhanced near pause sites since the stalled RT can carry out extensive RNase H cleavage in the vicinity of the pause site. It may be that some pause sites induce donor dissociation while others simply stall the polymerase and allow time for binding of the acceptor to the donor-DNA complex as described in the model below.

The acceptor facilitated model of strand transfer indicates that the acceptor has a more active role in initiating strand transfer. This model is shown in Figure 19. In this model the acceptor binds to the single stranded region of the DNA, from which the
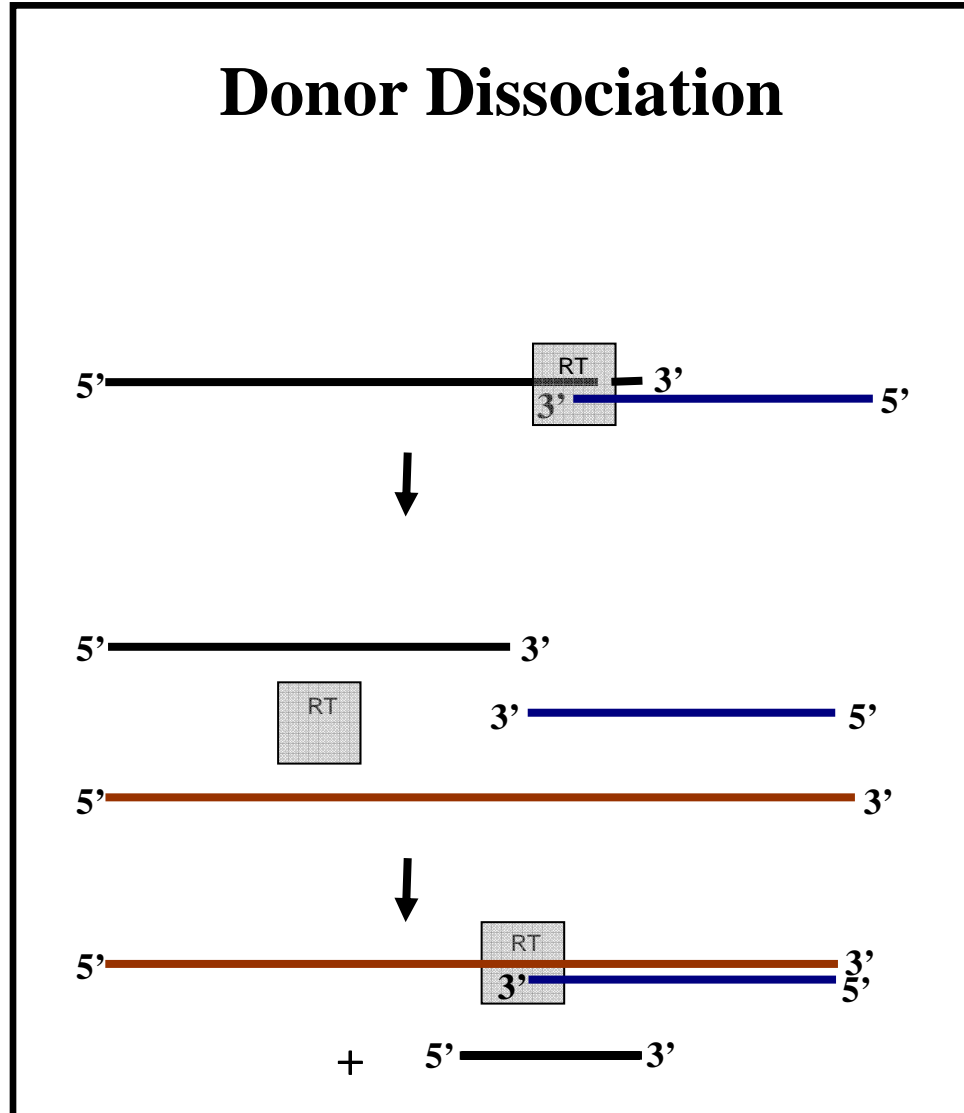
Figure 18: Donor dissociation model of strand transfer. One possible mechanism of copy choice internal transfer is shown. Initially the DNA (blue) is being synthesized on the donor RNA template (black). RT RNase H cleavages of the donor RNA occur. At some point the DNA dissociates from the donor RNA. The DNA is then free to anneal to the acceptor RNA (red) template. RT associates with the RNA:DNA complex and synthesis continues on the new template.

Acceptor Facilitated
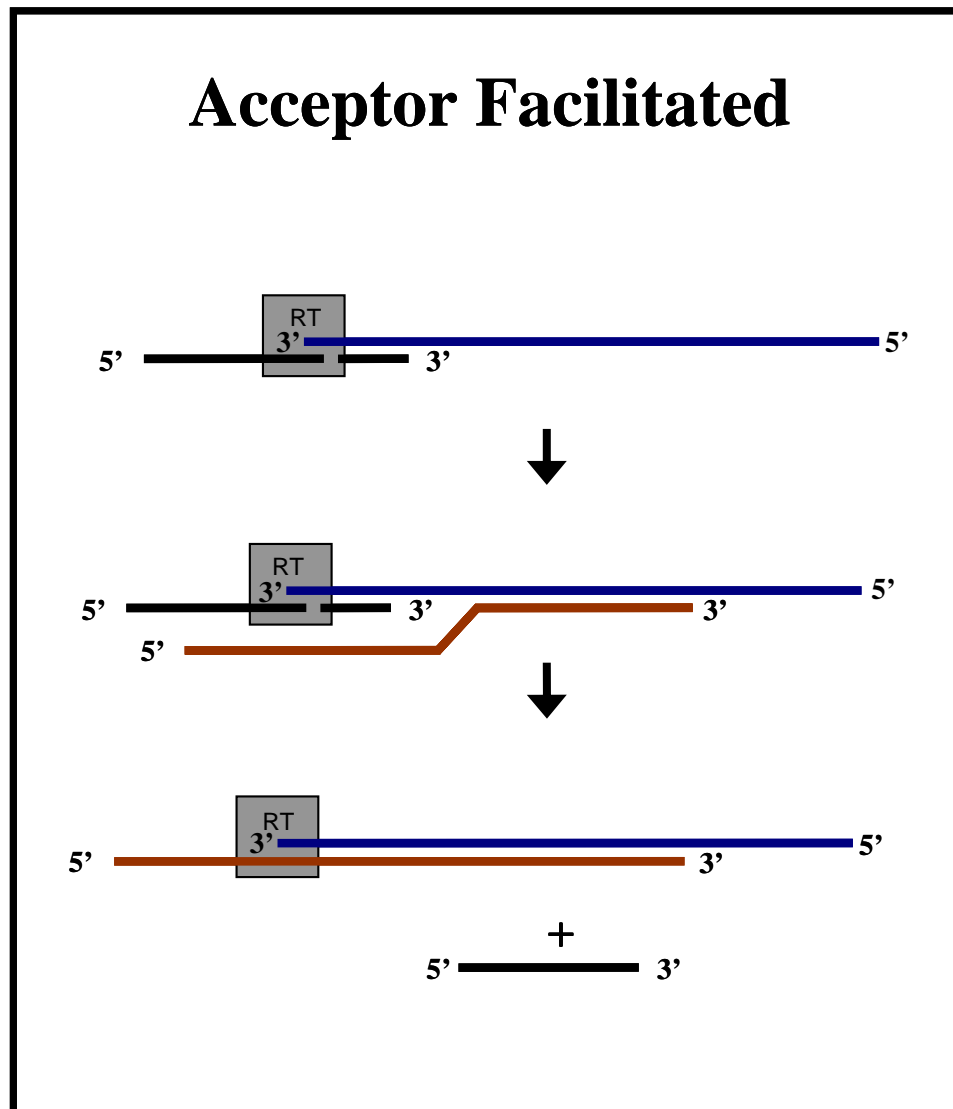
Figure 19: Acceptor facilitated model for strand transfer. In this model the acceptor RNA (red) binds to the single stranded region of the DNA (blue) that was released from the donor template (black) by RT RNase H activity. The annealing of the acceptor RNA and nascent DNA displaces the donor from the DNA. After the acceptor is completely annealed to the 3' end of the DNA, RT continues DNA synthesis on the acceptor template.

original RNA template was cleaved by RNase H. After binding it continues to anneal and displaces the donor from the nascent DNA. The acceptor and DNA then complete annealing. RT associates with this complex and synthesis continues on the acceptor template. Many recent experiments using the TAR hairpin, the DIS, and sequences from EIAV PBS have implicated the acceptor facilitated model for strand transfer (7, 95, 103, 113). The proposed dock and lock model for transfer suggest that pausing enables RT RNase H to cleave the donor to shorter fragments making the DNA available for the acceptor invasion (113).

These results indicate that internal strand transfer can occur by a number of mechanisms and that there are a number of factors that can enhance transfer. In this thesis strand exchange and donor displacement experiments in the presence of acceptors with different length 3' ends and varying degrees of complementarity to the DNA are presented. These experiments were designed to test the importance of the 3' end of the acceptor, which would bind to the nascent DNA, after RNase H cleavage of the donor template. The results indicated that optimal enhancement of strand transfer by HIV-1 NC occurs when the acceptor extends past the donor:DNA hybrid region for at least 22 nucleotides. They also indicated that it is vital for the single stranded region of the DNA, and the region of the acceptor that binds to this portion of the DNA, to have a high degree of complementarity for enhanced transfer with NC. This indicates that the homology between the acceptor and the donor upstream of the transfer point is very important for efficient transfer. Overall the results support a mechanism where NC stimulates binding of the acceptor to the upstream single stranded region of DNA then facilitates migration of the acceptor through the donor-DNA duplex region resulting in donor displacement

and strand exchange. No evidence for direct invasion of the duplex in the presence or absence of NC was observed.

Materials and Methods

Preparation of RNAs- Any chemicals or enzymes that do not specifically mention where they came from were obtained from Fisher Scientific or Sigma. RNAs were prepared by run off transcription completed on double stranded DNA products synthesized by PCR (see below). In strand exchange experiments PCR was completed using the pBSM13+, and pBSM13Δ for the acceptors and donor, respectively. The pBSM13Δ plasmid was created from pBSM13+ by removing 50 base pairs when cut with *Eco*RI and *Hind*III restriction enzymes. These extra base pairs gave the acceptors 50 additional nucleotides which allowed for a shift in gels when the DNA was hybridized to the acceptors as opposed to the donor. In donor displacement experiments pBSM13Δ was used to synthesize both the donor and acceptors. A list of the primers used in the PCR reactions is shown in Table 1. The various types of acceptors used are described and illustrated more specifically in the results section.

PCR reactions were run to amplify the desired DNA fragments. The PCR reaction included 0.1 μg template DNA, 50 pmoles of each primer, 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$ and 200 mM dNTPs and 5 units of Taq DNA polymerase (New England Biolabs). The heating temperature used was 94 $^{\circ}$C, with an annealing temperature of 48 $^{\circ}$C and an extension temperature of 72 $^{\circ}$C. Each temperature was held for one minute and 35 cycles were conducted. PCR reactions were run on 8 % non-denaturing polyacrylamide gels and the DNA bands were excised and eluted

Table 1:   PCR Primers used to synthesize DNAs used as templates for donor and acceptor RNA synthesis.

Acceptor	Primer 1[a] 5' to 3'	Primer 2[b] 5' to 3'
25acc	TTGTAATACGACTCACTATA	TAACCCTCACTAAAGGGAAC
35acc	TTGTAATACGACTCACTATA	AGCTCGAAATTAACCCTCACT
47acc	TTGTAATACGACTCACTATA	ATGATTACGCCAAGCTCGAA
63acc	TTGTAATACGACTCACTATA	GGAAACAGCTATGACCATGA
80acc	TTGTAATACGACTCACTATA	GTGTGGAATTGTGAGCGGAT

[a] This primer was used to create the end of the PCR DNA near the T7 start site. The same primer was used for each acceptor because all acceptors have the same 5' end which starts at the T7 site.

[b] This primer was used to create the end of the PCR DNA to which RNA transcription would run. This end was increased in length to increase the 3' end of the RNA, increasing the region of the acceptor that would anneal to the single stranded region of the DNA.

overnight in TE buffer (10 mM Tris-HCl pH 8, 1 mM EDTA). DNA was filtered and

precipitated as described in Chapter 2. After resuspending the DNA in water it was used

as a template for run-off transcription (see below).

For mutated acceptors, DNA primers were purchased from Integrated DNA

Technologies. These primers contained the desired mutations. Comp3mjf, comp4mut,

and comp6mut were 58-mer DNAs that were complementary to the 3' end of Mjf3hyb,

mjf4mut and mjf6mut, respectively. Mjf3hyb, mjf4mut, and mjf6mut were each 77-mer

DNAs that contained the SP6 promoter and the DNA sequence to yield the desired RNA.

Each set of primers was hybridized in a hybridization buffer (50 mM Tris-HCl pH 8, 80

mM KCl) by heating to 65 ^{o}C and slowly cooling to below 37 ^{o}C. The 5' end was then

filled in using 10 units of Klenow polymerase and 200 µM dNTPs. The hybrids were

phenol-chloroform extracted and precipitated. The pellets were resuspended in water and

the DNA was used as template for transcription reactions.

Transcription reactions for acceptors derived from the plasmids were prepared

using 1 µg of template DNA and 40 units of T7 RNA polymerase (Boehringer Manheim)

and run for two hours at 37 ^{o}C. The final concentration of components in the reaction

was as follows: 2 units/µL RNase inhibitor, 40 mM Tris-HCl PH 8.0, 6 mM MgCl$_2$, 100

mM DTT, 2 mM spermidine, 1 mM dNTPs. Fifteen Units of DNase 1 was added for ten

minutes to digest template DNA. Transcription reactions for mutated acceptors were run

in the same way except SP6 polymerase was used. Transcription reactions were phenol-

chloroform extracted and precipitated. RNA was run on 8 or 10 % denaturing

polyacrylamide gels. RNA was excised and the crushed gel was suspended in RNA

elution buffer (80% formamide, 400 mM NaCl, 1 mM EDTA, 40 mM Tris-HCl (pH

7.0)). After overnight elution RNA was precipitated as described previously. RNA was quantified by optical absorbance using a GeneQuant Spectrophotometer.

Preparation of 80-mer DNA for strand transfer experiments- DNA was purchased and purified on 8 % denaturing polyacrylamide gels. The prominent band was excised and eluted in TE buffer overnight. The DNA suspended in TE was filtered and precipitated. Fifty pmols of DNA was 5'- P^{32} end labeled using T4 PNK. The labeling reaction was done at 37 ^{o}C for thirty minutes. The T4 PNK was heat inactivated by incubating the reaction at 65 ^{o}C for fifteen minutes. The DNA was centrifuged on a G-25 sephadex column for four minutes at 3 X G.

Preparation of hybrid- 5' end labeled DNA was resuspended with 75 to 100 pmols of donor RNA in hybridization buffer (50 mM Tris-HCl pH 8, 80 mM KCl). The solution was heated to 65 ^{o}C for five minutes and then slowly cooled to below 37 ^{o}C. The hybrid was run on 10 % native polyacrylamide gel along with a sample of 5' end labeled single stranded DNA. This allowed for accurate retrieval of the hybrid. The gel was exposed to film with markers to enable the film to be lined up accurately on the gel. The hybrid was retrieved. An example film is shown in Figure 20. The hybrid band was removed and eluted for five to six hours in the hybridization buffer. The eluted hybrid was filtered and quantified. The amount of hybrid retrieved was determined using a scintillation counter to measure a known quantity of the labeled DNA and the unknown quantity of the hybrid. The counts per minute/pmol was determined using the known quantity of DNA. This could then be used to determine how many pmols of hybrid were recovered given the counts per minute of hybrid from the scintillation counter.

Figure 20: Shift of P^{32} labeled DNA due to hybridization to donor RNA. Gel shows the shift up of the DNA due to annealing to RNA. The donor:DNA band was excised and eluted for use in strand exchange experiments. The position of the single stranded DNA is also shown (ssDNA).

NC titration- An NC titration was performed at 37 $^{\circ}$C. Sixty microliter acceptor and hybrid master reactions were made. Six microliter aliquots from the hybrid reaction were removed placed in seven separate tubes. The same was done with the acceptor reactions. Each tube was then preincubated with 1.5 microliters of the given NC concentration for 2 minutes. Final NC concentrations were 0, 0.125, 0.25, 0.5, 1, 2, 4 µM. The reactions were started by adding 7.5 µL from the acceptor tube to the corresponding hybrid tube. The reactions ran for 8 minutes and then were stopped with 15 µL of Proteinase K (PK) (Kodak) stop solution (2 mg/mL PK, 10 mM Tris pH 8, 1.25 % SDS, 15 mM EDTA). Six microliters of 6 X native dye was added and reactions were placed on ice. Reactions were run on a 10 % native polyacrylamide gel and dried. Autoradiograms were obtained and quantified.

Strand transfer reactions- Strand transfer reactions were performed at 37 $^{\circ}$C. Four solutions of equal volume were made. Two solutions contained the hybrid while the others contained the acceptor. Six and a half µL of NC or control buffer (50 mM Tris-HCl pH 8, 2 mM DTT, 0.1 mM EDTA) was added to 58.5 µL of the hybrid and acceptor reactions. The reactions were preincubated with NC or buffer for two minutes. After two minutes 7.5 µL of acceptor reaction with NC was added to 7.5 µL of the hybrid reaction with NC in PK stop solution for time zero. The same was done for the reactions without NC. To start the reactions 57 µL of the acceptor solution with NC was added to 57 µL of the hybrid solution containing NC. The same was done for the reactions without NC. The final concentrations were 2 nM hybrid, 4 nM acceptor RNA, 1 µM NC (or 0 µM NC for reactions without NC), 50 mM Tris-HCl pH 8, 0.1 mM EDTA, 2 mM DTT, 0.1 µg/µL bovine serum albumin (BSA), 100 µM $ZnCl_2$, 6 mM $MgCl_2$, and 80 mM

KCl. At specific time points (1, 2, 4, 8, 16 and 32 min) 15 µL of each reaction was

removed and added to 15 µL of PK solution. Each reaction was incubated in PK for 4

minutes to digest NC protein and then 6 µL of 6X native dye was added. Reactions were

placed on ice immediately. Control reactions were performed by adding 6.75 µL of

hybrid reaction to 6.75 µL of buffer without acceptor and 1.5 µL of NC or control buffer.

Controls were incubated at 37 ^{o}C for 32 minutes and then PK solution was added. After

4 minutes 6 µL of 6X native dye was added and control reactions were placed on ice.

Reactions were then run on 10 % native polyacrylamide gels. Gels were dried and

autoradiographs were obtained. Gels were quantified using a phosphorimager. The

percent transfer was obtained by dividing the amount of DNA hybridized to the acceptor

(T) by the amount of DNA hybridized to the donor (H) and that transferred to the

acceptor(T) and multiplying by 100 (% Transferred = T/ (H+T) * 100).

   DNA:RNA annealing experiments- The DNA was 5' end labeled as described for

the preparation of the hybrid. The DNA was not hybridized to the donor RNA and was

used in the single stranded form. Four separate master reactions were made, two with

DNA and two with RNA (either 25don or 25acc). Six and a half microliters of NC or

control buffer (50 mM Tris-HCl pH 8, 2 mM DTT, 0.1 mM EDTA) was added to 58.5

µL of the DNA and RNA reactions. The reactions were preincubated with NC or buffer

for two minutes at 37 ^{o}C. After two minutes 7.5 µL of the RNA reaction with NC was

added to 7.5 µL of the DNA reaction with NC in PK stop solution for time zero. The

same was done for the reactions without NC. The reactions were started by adding 57 µL

of the RNA solution with NC to 57 µL of the DNA solution with NC. The same was

done for the reactions without NC. Final concentrations in the reaction were 2 nM DNA,

4 nM RNA, 1 µM NC (or 0 µM NC for reactions without NC), 50 mM Tris-HCl pH 8.0,

0.1 mM EDTA, 2 mM DTT, 0.1 mg/mL BSA, 100 mM $ZnCl_2$, 6 mM $MgCl_2$ and 80 mM

KCl. At each time point (1, 2, 4, 8, 16, 32 min) 15 µL of each reaction was removed and

added to 15 µL of PK solution. Each reaction was incubated in PK solution for 4 minutes

to digest NC protein and then 6 µL of 6 X native dye was added. Reactions were

immediately placed on ice. Control reactions were performed by adding 6.75 µL of the

DNA reaction to 6.75 µL of buffer without RNA and 1.5 µL of NC or control buffer.

Controls were incubated at 37 ^{o}C for 32 minutes and then PK solution was added. After

4 minutes 6 µL of 6 X native dye was added and control reactions were placed on ice.

Reactions were then run on 10 % native polyacrylamide gels. Gels were dried and

autoradiograms were obtained. Gels were quantified using a phosphorimager. The

percent DNA annealed was obtained by dividing the annealed DNA (A) by the total DNA

(single stranded DNA (S) + the annealed DNA) and multiplying by 100 (Percent

Annealed = (A/(A+S)*100)).

   Preparation of RNA for donor displacement experiments- For hybrids 25hyb and

35hyb the RNA was internally labeled for donor displacement experiments. This was

done by synthesizing the RNA in run-off transcription reactions which included α-P^{32}

UTP. Reactions included 1 µg template DNA, 40 mM Tris-HCl pH 8.0, 6 mM $MgCl_2$,

10 mM DTT, 2 mM sperimidine, 10 nmol cold ATP, CTP and GTP, 2 nmol cold UTP,

100 pmol labeled UTP, 20 units of T7 RNA polymerase, and 20 units of RNase inhibitor.

Transcription reactions were run for 1 hour at 37 ^{o}C. Fifteen units of DNase 1 were

added and the reaction was incubated for ten additional minutes. Twice the volume of 2

X formamide dye was added and the RNA was purified on a 10 % polyacrylamide

83

denaturing gel. The RNA was excised and eluted in RNA elution buffer as previously described. The RNA eluate was removed from the gel and filtered. Two volumes of ethanol were added and the RNA was precipitated as described in Chapter 2. RNA was quantified using a scintillation counter and the amount of RNA was determined by the specific activity of the radiolabel. The equation used was $A = A_o e(-\lambda t)$ where A_o is the activity on the reference date, λ is the decay factor, 0.04854, and t is the number of days of decay.

Preparation of hybrid for donor displacement experiments- Fifty pmoles of labeled donor RNA was combined with 75 pmoles of 80-mer DNA in hybridization buffer. The solution was heated to 65 ^{o}C for five minutes and then slowly cooled below 37 ^{o}C. The hybrid was run on a 10 % native polyacrylamide gel adjacent to single stranded labeled donor RNA. The gel was exposed to film and the hybrid was excised and eluted in hybridization buffer for five to six hours. Experiments done with 25hyb and 47acc3mut were completed with a hybrid that was eluted for a half hour in 1 X TBE (89 mM Tris-HCl pH 8.0, 89 mM Boric Acid, 2 mM EDTA). In this case the gel was excised and placed in dialysis tubing in 1 mL 1 X TBE. The tubing was suspended in a liter of 1 X TBE and current was applied to draw the hybrid out of the gel into the buffer in the tubing. This was done because during elution there was repeatedly more than 15 % of the hybrid falling apart. Therefore, to obtain a stable hybrid elution conditions were altered. The amount of hybrid retrieved in all cases was quantified as described for the strand transfer hybrids using known quantities of single stranded donor for comparison.

Donor displacement experiments- Donor displacement experiments were completed similarly to the strand transfer experiments. Hybrid solutions and separate

acceptor solutions were made. Six and a half µL of NC or control buffer (50 mM Tris-HCl, 0.1 mM EDTA, 2 mM DTT) was added to 58.5 µL of the hybrid and acceptor reactions. The reactions were preincubated with NC (at 37 ^{o}C) or buffer for two minutes. After two minutes 7.5 µL of acceptor reaction with NC was added to 7.5 µL of the hybrid reaction with NC in PK stop solution. This was the time zero reaction. The same was done for the reactions without NC. To start the reactions 57 µL of the acceptor solution with NC was added to 57 µL of the hybrid solution containing NC. The same was done for the reactions without NC. The final concentrations were 2 nM hybrid, 4 nM acceptor RNA, 1 µM NC (or 0 µM NC for reactions without NC), 50 mM Tris-HCl pH 8, 0.1 µg/µL BSA, 0.1 EDTA, 2 mM DTT, 100 µM $ZnCl_2$, 6 mM $MgCl_2$, and 80 mM KCl. As mentioned previously, the experiments done with 25hyb and 47acc3mut were performed with slightly altered buffer conditions due to a change in hybrid elution conditions. Final reaction concentrations in this experiment was 37 mM Tris-HCl pH 8.0, 37 mM Boric Acid, 0.1 µg/µL BSA, 0.8 mM EDTA, 7.6 mM $MgCl_2$, 100 mM $ZNCl_2$, 80 and mM KCl. At specific time points (1, 2, 4, 8, 16, 32 minutes) 15 µL of each reaction was removed and added to 15 µL of PK solution. Each reaction was incubated in PK for 4 minutes to digest NC protein and then 6 µL of 6 X native dye was added and reactions were placed on ice. Control reactions were performed by adding 6.75 µL of hybrid reaction to 6.75 µL of buffer without acceptor and 1.5 µL of NC or control buffer. Controls were incubated at 37 ^{o}C for 32 minutes and then PK solution was added. After 4 minutes 6 µL of 6 X native dye was added and control reactions were placed on ice. Reactions were then run on 10 % native polyacrylamide gels. Gels were dried and autoradiographs were obtained. Gels were quantified using a phosphorimager. The percent donor displacement

was obtained by dividing the amount of single stranded RNA (S) by the amount of RNA hybridized to the 80-mer (H) added to the amount of single stranded RNA and multiplying by 100 (% Donor Displaced = S/ (H+S) * 100). When 80acc and 47acc3mut were used there was some donor that was displaced from the hybrid which bound to the acceptor. For these reactions the displaced donor was calculated by adding the single stranded RNA (S) to the RNA that transferred to the acceptor (T) and dividing by the total RNA (S + H + T) and multiplying by 100. With every acceptor used the percent RNA displaced at time zero was subtracted from each time point. This was done because during hybrid purification a small amount of the hybrid (less than 15 %) would fall apart. The donor displaced after hybrid purification was subtracted so that it would not be considered displaced due to exchange of the DNA to the acceptor in the reactions.

Results

The experiments completed to test the mechanism for NC catalyzed internal strand exchange pointed to a complex mechanism. It appeared that this enhancement is not singly driven by thermodynamics or by structural features of the acceptor. Rather, both components play a role and optimal transfer will only be observed when the acceptor has a long 3' end that extends at least 22 base pairs beyond the 3' end of the donor.

Strand-exchange to acceptors forming hybrid regions with at least 22 additional base pairs is enhanced by NC. The first experiments completed were strand exchange experiments with a labeled DNA. The 80-mer DNA was 5' end labeled with P^{32} and hybridized to 36-mer RNA by being heated for five minutes to 65 ^{o}C and then slowly

86

cooled to room temperature. This hybrid, 25hyb, contained 25 base pairs. A schematic

of the hybrid is shown in Figure 21 along with the different acceptors used in these

experiments. After purification the hybrid was used to examine transfer of the labeled

DNA to acceptors with different length hybrid regions.

The first experiment performed was an NC titration to determine how much NC

was necessary for optimal stimulation of transfer. The hybrid was incubated for eight

minutes with a long acceptor, 63acc that binds all but 17 nucleotides of the DNA. This

acceptor binds to 38 nucleotides of the single stranded region of the DNA. This was

completed for seven concentrations of NC varying from 0 to 4 µM. An autoradiogram

and graphical representation of the NC titration is shown in Figure 22. From the NC

titration 1 µM NC concentration was chosen for the experiments. This was chosen

because optimal stimulation occurred with 0.5 and 1 µM NC. The higher concentrations

showed a decrease in strand transfer which is probably due to the fact that NC can

aggregate nucleic acids when present at a high concentration. In the autoradiogram the

highest concentration (4 µM) of NC smears the DNA on the gel, probably because there

is too much NC to effectively degrade. Though transfer was also high with 0.5 µM NC,

which is the concentration closest to a 1:7 NC:nucleotide ratio (ratio corresponding to the

binding site size of NC protein), we did not choose this concentration because there was a

slight increase with 1 µM NC. Therefore, 1 µM NC would yield the best information

when used with each different acceptor.
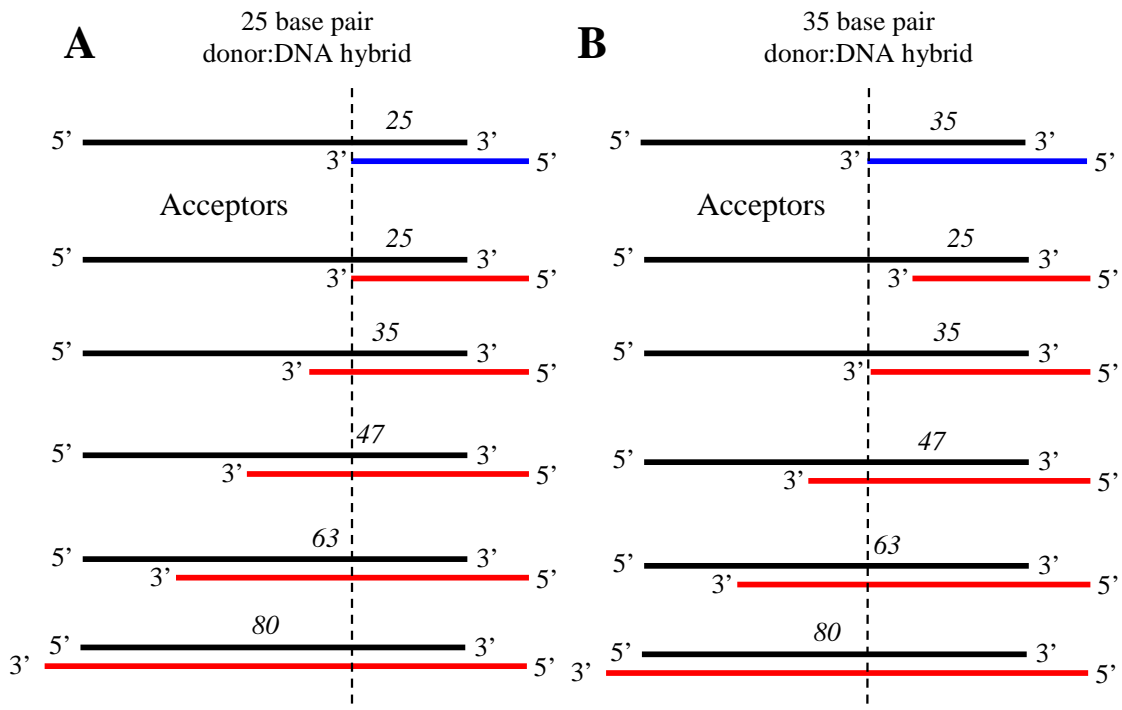
Figure 21: Schematic of hybrid regions between donor and DNA and acceptors and DNA. The hybrid regions between the donor RNA and DNA are shown at the tops of panels A and B. The DNA is black and the donor is blue. Panel A shows the hybrid with 25don which base pairs to 25 nucleotides of the DNA. Panel B shows the hybrid with 35don which base pairs to 35 nucleotides of the DNA. The hybrid regions with each acceptor are shown below the donor:DNA hybrids. From top to bottom the DNA is shown bound to: 25acc, 35acc, 47acc, 63acc, and 80acc. The DNA is black and the acceptors are red. The vertical dashed line in each panel illustrates the point on the DNA where the 3' end of the donor is bound. This is used to highlight the 3' end of the acceptor relative to the 3' end of the donor. The number of base pairs each donor or acceptor forms with the DNA is marked above the DNA. All acceptors are shown with the short 5' end that was used in the donor displacement experiments. Strand exchange acceptors all had a 50 nucleotide addition on the 5' end.
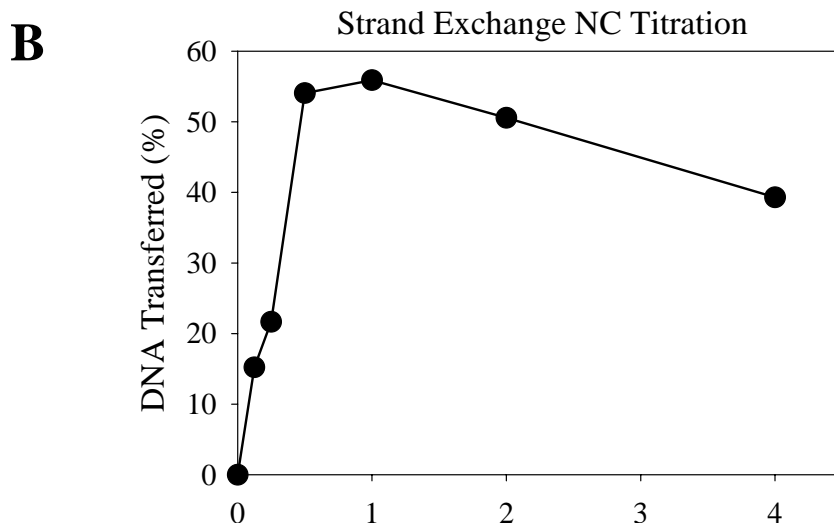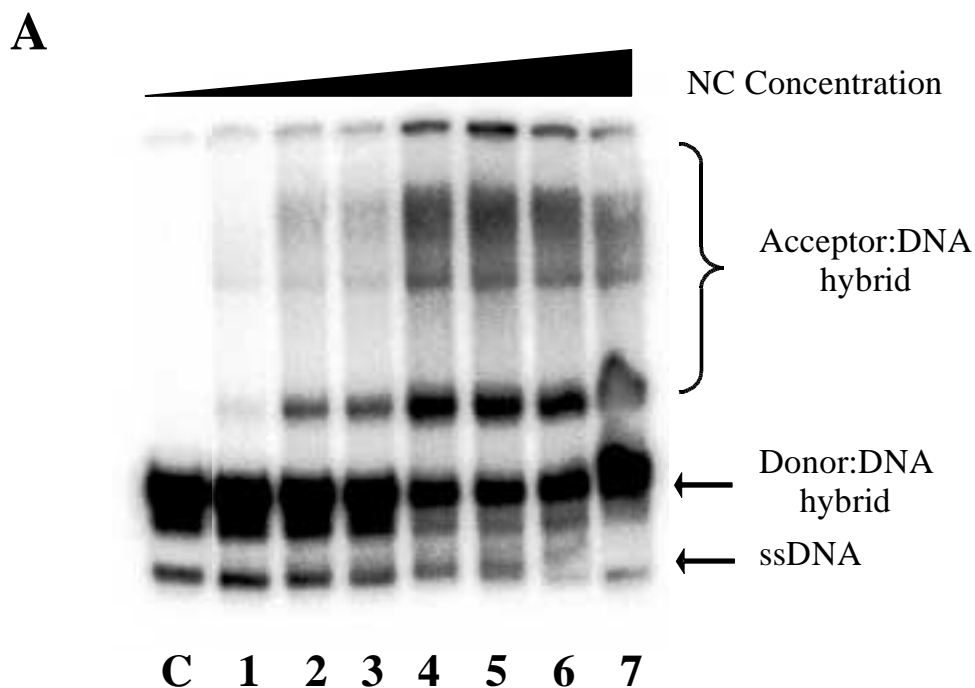
Figure 22: NC strand transfer titration using 63acc. The autoradiogram of the NC titration is shown in panel A. Lane C, control without acceptor or NC. Lanes 1 through 7 have the following NC concentrations: 0, 0.125, 0.250, 0.5, 1, 2, and 4 μM. Positions where the acceptor:DNA hybrid, donor:DNA hybrid and single stranded DNA run are marked. Panel B is the graph showing the percent of DNA transferred. The graph shows one representative experiment, though the titration was repeated and similar results were obtained.

To test DNA strand transfer to different acceptors the hybrid was incubated with each acceptor in the presence or absence of NC over 32 minutes. Reaction conditions were chosen for the same reasons discussed in Chapter 2. The experiments were run on native polyacrylamide gels and autoradiograms were obtained. Figures 23 and 24 show the transfer from the 36-mer RNA to five different acceptors. These acceptors increase in length on the 3' end as illustrated in Figure 21. This increases the number of nucleotides that can bind to the DNA outside the donor:DNA hybrid region. In panel B of Figure 23 it is clear that NC enhances transfer to 63acc, which can base pair with 38 additional nucleotides of the DNA. The gel shows that as the DNA transfers from the donor RNA to the acceptor it shifts up in the native gel. This allowed for quantification of the transferred DNA. Also, in Figure 24, the experiment completed using 25acc shows no transfer. This acceptor has the same hybrid region as the donor. In this case the DNA does not transfer from the donor to the acceptor.

These experiments were completed with five different acceptors. The results are displayed graphically in Figure 25. This figure shows that substantial transfer is not observed until the acceptor can form an additional 22 base pairs with the DNA outside the hybrid region. When the acceptor forms no additional base pairs with the DNA transfer is not observed in the presence or absence of NC. When ten additional nucleotides are added to the acceptor no transfer is seen without NC and with NC there is just a very slight enhancement in transfer. However when the acceptors with 22 additional base pairs is used NC enhances transfer substantially. There is just a slight transfer without NC, yet when NC is added the rate of transfer is greatly increased. This
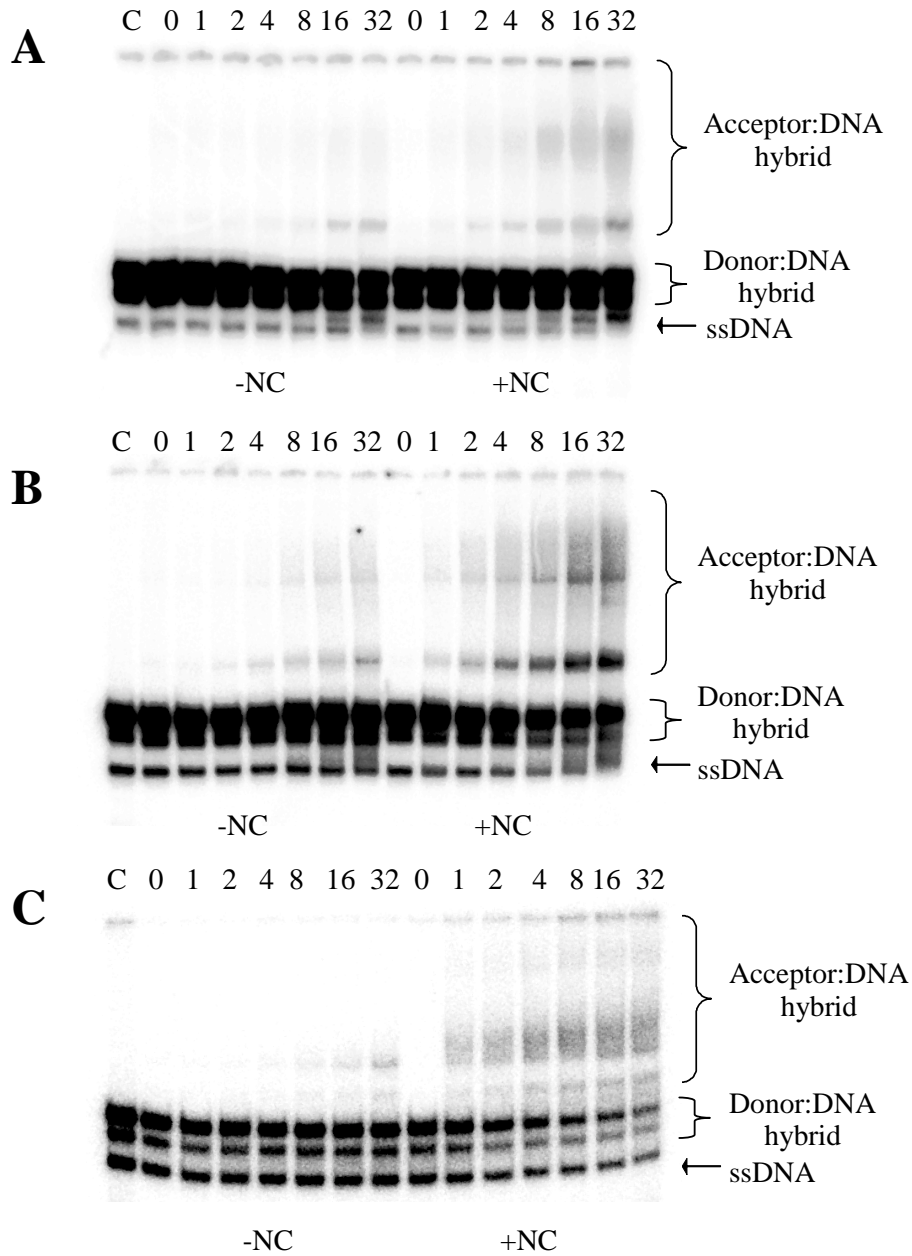
Figure 23: Autoradiograms showing DNA exchange from donor RNA to long acceptors. Autoradiograms are shown with 47acc (A), 63acc (B) and 80acc (C). Positions where the acceptor:DNA hybrid, donor:DNA hybrid and single stranded DNA (ssDNA) run are marked. The time is marked above the lane. The first seven time points on each gel were completed in the absence of NC while the last seven were done in the presence of NC. The control lane (C) had NC, but no acceptor. Control reactions were completed over 32 minutes.
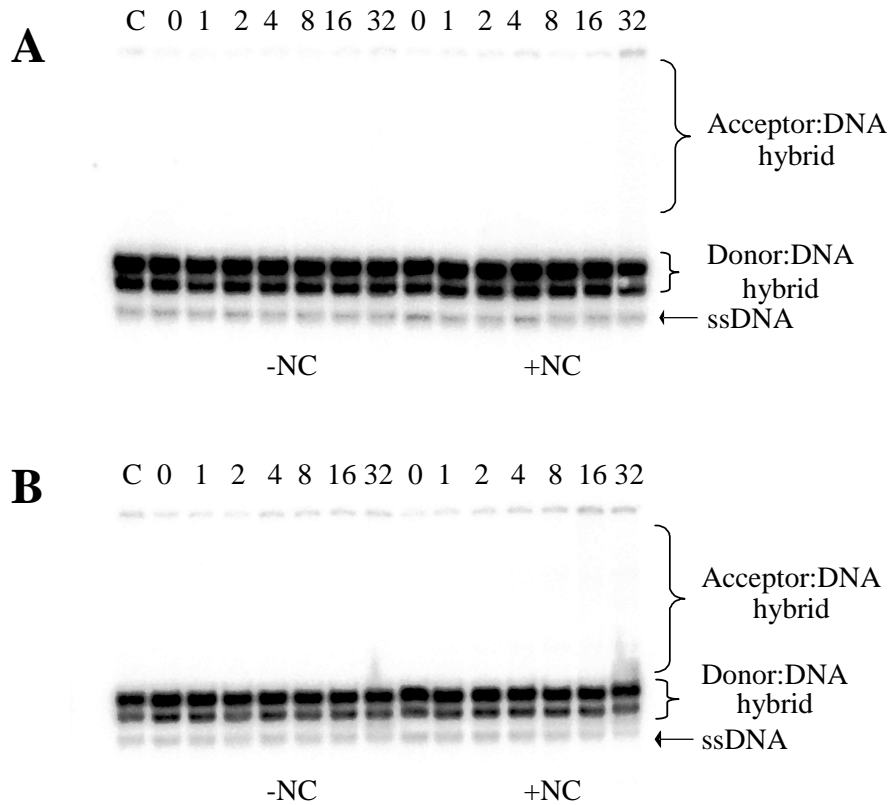
Figure 24: Autoradiograms showing DNA strand exchange to short acceptors.
Autoradiograms are shown with 25acc (A), and 35acc (B). Positions where the
acceptor:DNA hybrid, donor:DNA hybrid and single stranded DNA (ssDNA) run are
marked. The time is marked above the lane. The first seven time points on each gel
were completed in the absence of NC while the last seven were completed in the
presence of NC. The control lane (C) had NC, but no acceptor. Control reactions
were carried out for 32 minutes.
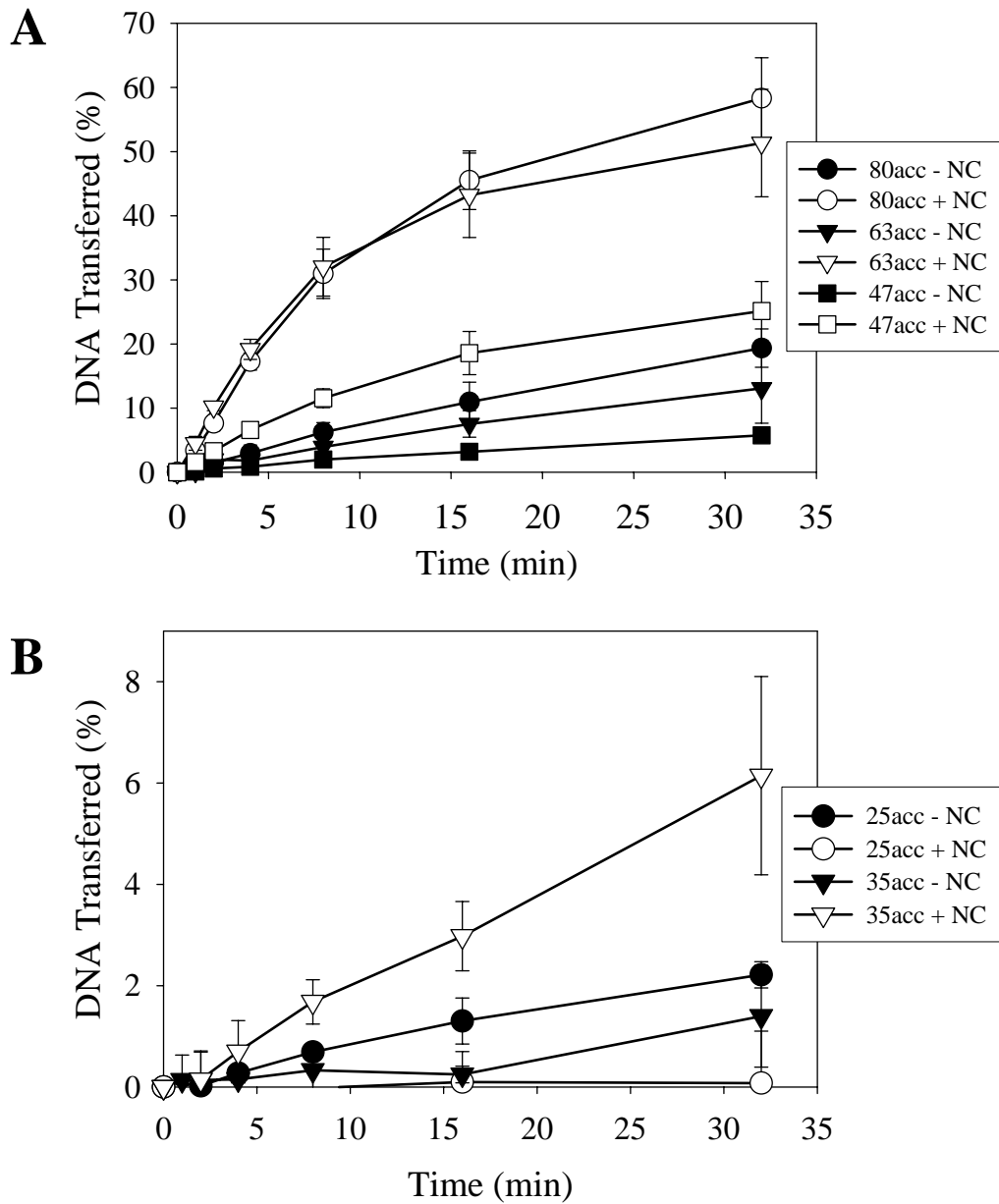
Figure 25: Graphs showing DNA transfer from donor to acceptor RNAs. Graphs show the percent of DNA transferred from the donor RNA to acceptor RNA. The acceptor used in each experiment is given in the legends. Open symbols are with NC and closed symbols without NC. Acceptors with hybrid regions greater than 47 base pairs are in panel A and acceptors with smaller regions in panel B. Note the difference in the Y-axis values in each graph. The average of three experiments is shown and the error bars represent the standard deviation.

trend is also observed with 63acc and 80acc. In both cases there is transfer without NC, but the addition of NC enhances the rate of transfer considerably.

These experiments clearly suggest that the acceptor must have additional base pairs for NC to enhance transfer. However, it was important to determine if NC was not enhancing transfer to the shorter acceptors because NC would not even enhance annealing between the 80-mer DNA and these RNAs. Therefore, annealing experiments were completed using 5' end labeled 80-mer DNA and 25don or 25acc. These results are shown in Figure 26. The graph shows that annealing will occur between the 80-mer DNA and both RNAs even without NC. NC further enhances annealing to both RNAs. Therefore, hybrid formation between these nucleic acids is enhanced by NC suggesting that the absence of strand exchange is due to the lack of additional nucleotides on the acceptor.

Donor displacement experiments with NC show optimal enhancement of RNA exchange with annealing regions greater than 22 nucleotides. Strand exchange experiments showed that the length of the complementary region between the DNA and acceptor is important in NC enhanced strand exchange. This indicates that the acceptor plays a more active role in strand transfer, supporting the acceptor-facilitated model of strand transfer. If acceptor binding to the single stranded region of the DNA is a precursor to strand transfer then we should not see donor being displaced from the original hybrid in the presence of acceptors where strand exchange was not observed. We therefore employed a different method to test strand exchange. In the donor displacement experiments the donor RNA (25don) was labeled with P^{32} and the 80-mer DNA was not labeled. The hybrid was formed in the same way as it was with the labeled
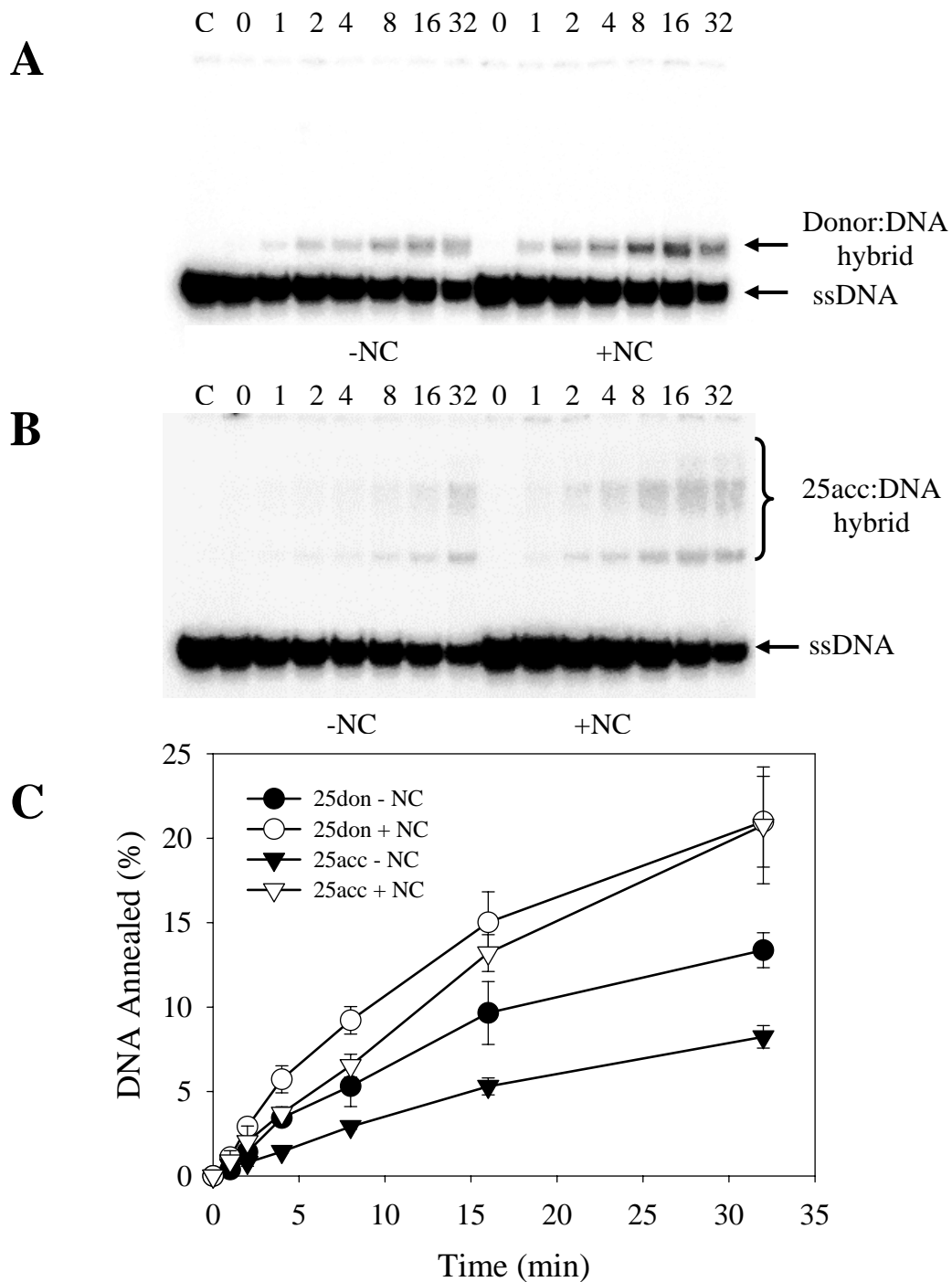
Figure 26: Annealing of 80mer DNA to RNAs with 25 nucleotide complementary regions. Panel A shows the autoradiogram of annealing to the 25don. This RNA has 36 nucleotides. Panel B shows the autoradiogram of annealing to 25acc which has 86 nucleotides. In both panels the position of the single stranded (ssDNA) and RNA-bound DNA are marked. The time is shown above the gel and the NC lanes are marked. Panel C is the graph displaying the percent of DNA annealed to the RNAs in the presence (open symbols) and absence (closed symbols) of NC. The average of three experiments is plotted and the error bars show the standard deviation.

DNA and cold donor RNA. The hybrid was then incubated with different acceptors in the presence and absence of NC over a 32 minute time course. The experiments were also run on non-denaturing polyacrylamide gels and autoradiograms were obtained. In these experiments we looked at the amount of the donor that was displaced from the hybrid.

Using the donor displacement method was beneficial for a number of reasons. First we could see what types of acceptors enhanced the removal of the donor from the DNA. This allowed us to differentiate between the two models proposed for transfer and to determine how NC enhances transfer. Second, this technique allowed us to remove a variable that could potentially have an affect on the strand exchange experiments. For the strand exchange experiments it was necessary to have an additional 50 nucleotides on the 5' end of the acceptors so that the DNA would adequately shift up in the gel once bound to the acceptor. However, in the donor displacement experiments the 50 nucleotides could be removed from the 5' end of the acceptor because the observed change was the donor being removed from the hybrid. This is beneficial because the extra 50 nucleotides could change the secondary structure of the acceptor RNA, and in most cases allow for a much stronger structure. An example of this is shown in Figure 27 with 63acc. Because RNA structure could play a role in strand transfer, and our experiments are not intended to test differences in structure it was desirable for us to remove these 50 nucleotides. Due to these advantages we used the donor displacement method to further examine the effects of NC on strand transfer.

Donor displacement experiments were first completed with the 80-mer DNA and 36-mer RNA hybrid. This hybrid, 25hyb, contains 25 base pairs. The donor
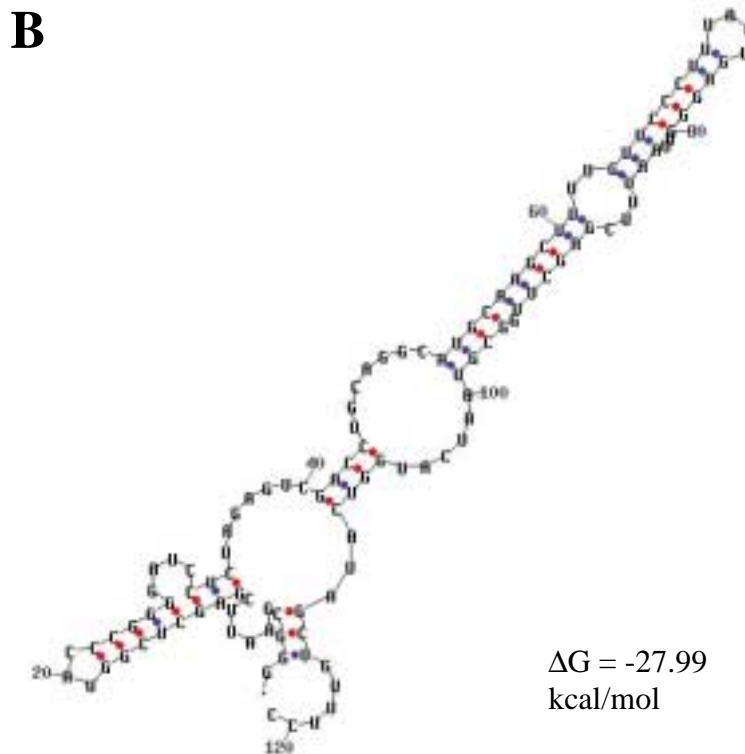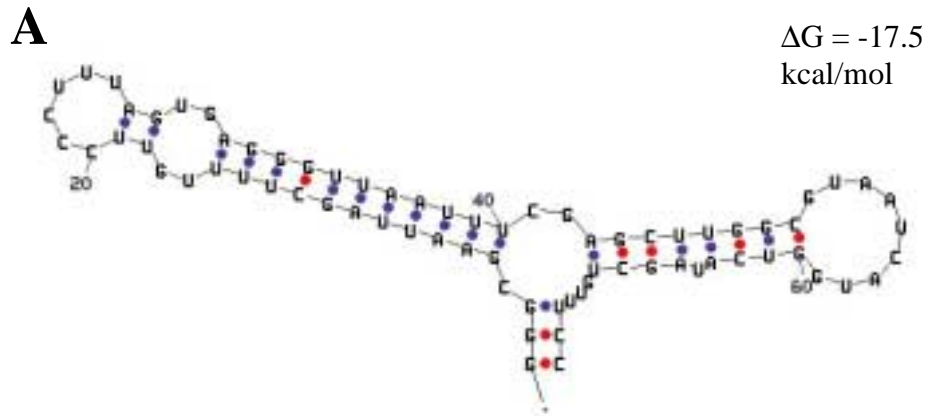
Figure 27: Predicted folding of 63acc in the short and long form. A) The structure of 63acc derived from the pbsM13Δ plasmid is shown. B) The structure of 63acc derived from the pbsM13+ plasmid is shown. The structure in B contains 50 more nucleotides than that in A. The predicted free energy for each structure is also shown. Structures were predicted using mFold.

displacement autoradiograms for 25hyb are shown in Figure 28. These results are displayed graphically in Figure 29. The graphs support the strand exchange experiments discussed previously. Again there is not a substantial amount of displacement until 47acc is used. This acceptor can hybridize to an additional 22 nucleotides of the DNA beyond the 3' end of the donor RNA. The donor is not displaced when the shorter version of 25acc is used.

When 35acc was used, which can hybridize to an additional ten nucleotides when compared with the donor, there was a very slight displacement of the donor RNA and minimal stimulation with NC. This indicates that the complementary region outside the donor:DNA hybrid region is essential for stimulation of transfer, and that an additional ten nucleotides is not enough for optimal NC stimulation of transfer. However, when an additional 22 nucleotides was available in the acceptor the amount of donor displacement increased dramatically. Also a clear stimulation with NC was observed. This indicates that NC will increase strand exchange when there is an additional 22 base pairs that can be formed. The increase in stability between the donor:DNA and 47acc:DNA hybrids was 26.7 kcal/mol as predicted by MELTING (84).

These results were supported when the length of the 3' end of the acceptor was extended (see Fig. 28). When experiments were completed with 63acc and 80acc displacement was observed without NC for both acceptors. However, when NC was added the reactions proceeded much quicker. Interestingly there was not a statistical difference between the stimulation observed in the displacement with either 63acc or 80acc. Therefore, though 80acc can hybridize to the full length of the DNA, 17 nucleotides more than 63acc, NC does not enhance exchange to this acceptor more than

Figure 28: Autoradiograms of donor displacement experiments with 25hyb.
Autoradiograms show the displacement of the donor RNA in the presence of 25acc
(A), 35acc (B), 47acc (C), 63acc (D) and 80acc (E). The position of the donor:DNA
hybrid is shown as well as the displaced RNA (ssRNA). NC was included in the
reactions with a plus (+) symbol. Time is marked above each gel in minutes.

Figure 29: Graphs showing donor displacement in the presence and absence of NC. Graphs show the RNA displaced in the presence (open symbols) and absence (closed symbols) of NC. A) Acceptors with hybrid regions less than ten base pairs. B) Acceptors with hybrid regions longer than ten base pairs. The acceptor used in each experiment is marked in the legend. The average of three experiments is shown and the error bars represent the standard deviation.

63acc. In fact when these acceptors are compared to 47acc, NC stimulation of donor displacement is highest with 47acc. Displacement in the absence of NC is also highest with 47acc.

Though it is clear that the length of the complementary region is important for displacement to occur, it does not seem that the longest complementary region will enhance displacement, and exchange of the RNAs, the most in the presence of NC. This is probably because the shorter acceptor, 47acc, folds with less thermodynamic stability than the longer acceptors. Therefore the nucleotides in this acceptor would be more available for annealing to the DNA and initiating transfer. Also, because the acceptors were shorter in the displacement experiments than in the strand exchange experiments this effect would be more pronounced, whereas the acceptors with the longer 5' ends all seem to form strong structures that can interfere with transfer. The structures for each acceptor, in both the short and long form, were predicted using mfold and the stabilities are listed in Table 2.

These results point to another interesting fact regarding NC stimulation of strand exchange. NC enhancement is only observed in reactions where displacement can also occur in the absence of NC. For example displacement is not observed without NC when 25acc is used. When NC is added there is still no displacement which occurs. However, when 35acc is used there is a slight amount of displacement without NC. NC enhances this displacement. As the length of the acceptor increases to a length where the displacement is considerable without NC the effect of NC is greater. Therefore, if the donor:DNA hybrid would not normally fall apart in the presence of 25acc, as we observe, NC will not enhance the displacement of the donor from the DNA with this acceptor.

Table 2: Free energy predictions for folding of acceptor RNAs.

Acceptor	Free Energy of Folding of Short Acceptor[a]	Free Energy of Folding of Long Acceptor[a]
25acc	-8.3	- 21.0
35acc	-11.2	-25.0
47acc	-14.0	-30.7
63acc	-17.5	-27.9
80acc	-25.7	-39.7

[a]Free Energies given are in kcal/mol. They are predicted using mFold using default conditions. Some sequences could fold into more than one structure. The free energy given for these structures was from the strongest predicted structure.

However, if the hybrid will fall apart in the presence of a longer acceptor NC will enhance this effect. This suggests that NC is primarily enhancing the annealing of the acceptor outside the donor:DNA hybrid region and therefore stimulating displacement and exchange of the RNAs. This indicates how important the annealing of the single stranded region of the DNA and acceptor is in strand exchange.

Donor displacement from longer hybrid optimally enhanced by NC with annealing regions longer than 28 nucleotides. We next wanted to determine whether the number of nucleotides in the acceptor necessary for transfer would be the same if the original donor:DNA hybrid was longer. Therefore we performed donor displacement experiments using a 35 base pair hybrid (35hyb). In this case 35acc (the short version) was internally labeled with P^{32} and hybridized to the 80-mer DNA. A schematic of this hybrid is also shown in Figure 21. Displacement experiments were completed with and without NC with the same acceptors used in the previous reactions. Figure 30 shows the autoradiograms of displacement assays with 35hyb and the different acceptors. The graphs are shown in Figure 31. Displacement was not observed with either 25acc or 35acc. This again supports the notion that the acceptor must be able to anneal outside of the donor:DNA hybrid for NC enhanced transfer. Otherwise, we would expect to see displacement of the 35 nucleotide donor with either acceptor because transfer of the DNA would be initiated by the donor being displaced not the annealing of the acceptor outside the original hybrid region. However, displacement is not seen with either 35acc, which has the same hybrid region as the donor:DNA hybrid in these experiments, or with 25acc which would have a shorter hybrid region. Also, it is important to note that in the previous displacement experiments completed with the 25 base pair hybrid and 35acc,
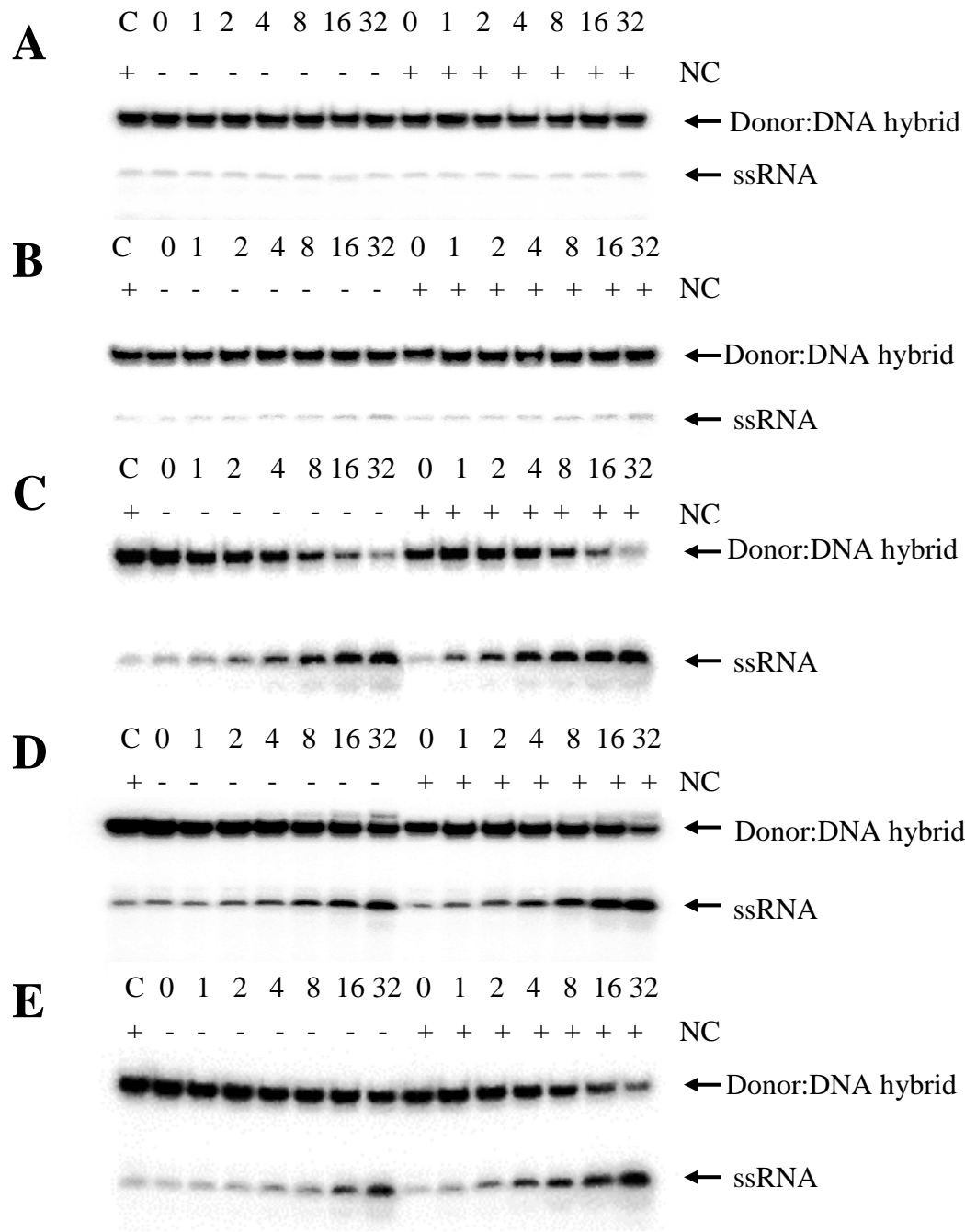
103
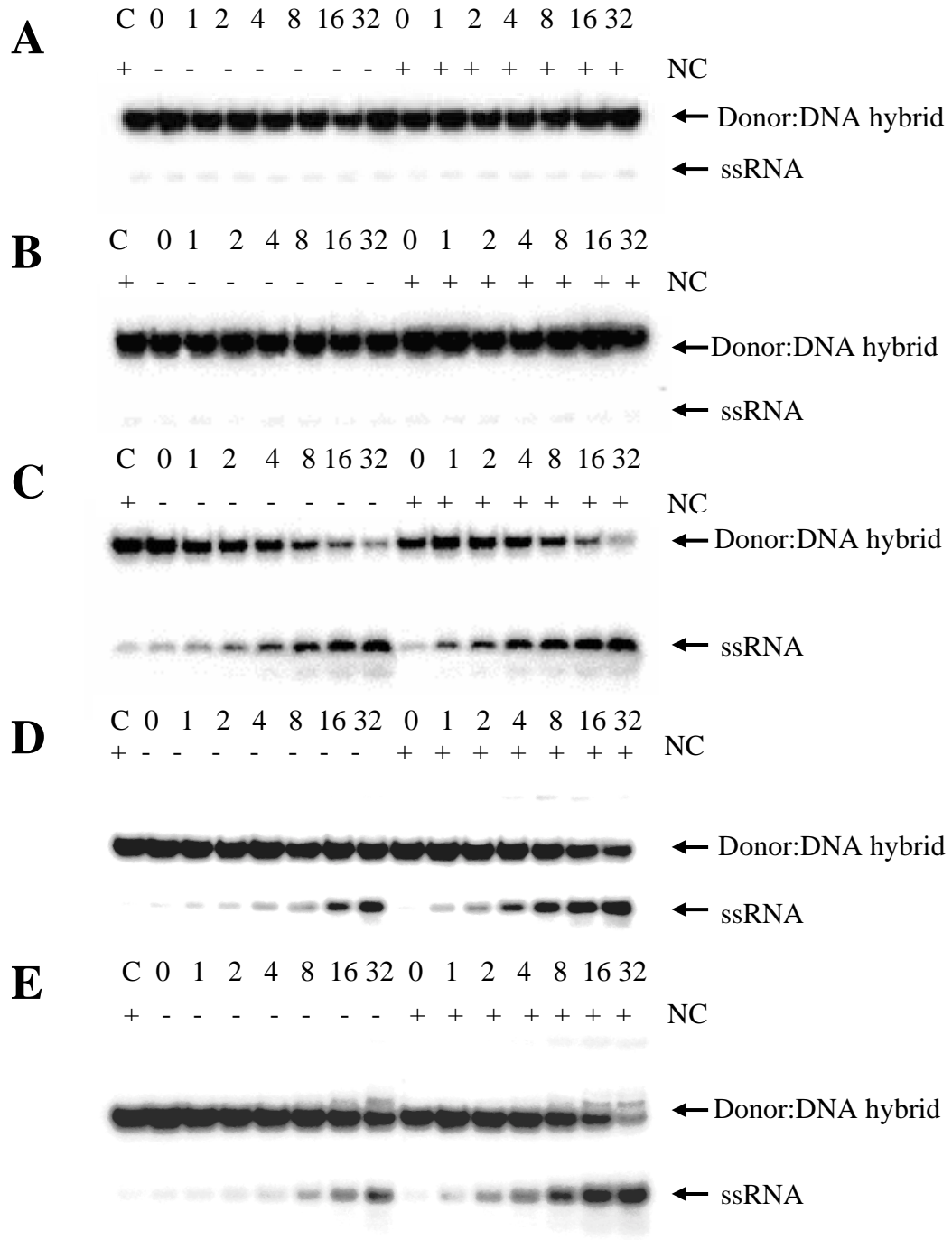
Figure 30: Autoradiograms showing donor displacement experiments with 35hyb. Autoradiograms show the displacement of the 35don RNA in the presence of 25acc (A), 35acc (B), 47acc (C), 63acc (D) and 80acc (E). The position of the donor:DNA hybrid is shown as well as the displaced RNA (ssRNA). NC was included in the reactions with a plus (+) symbol. Time is marked above each gel in minutes.
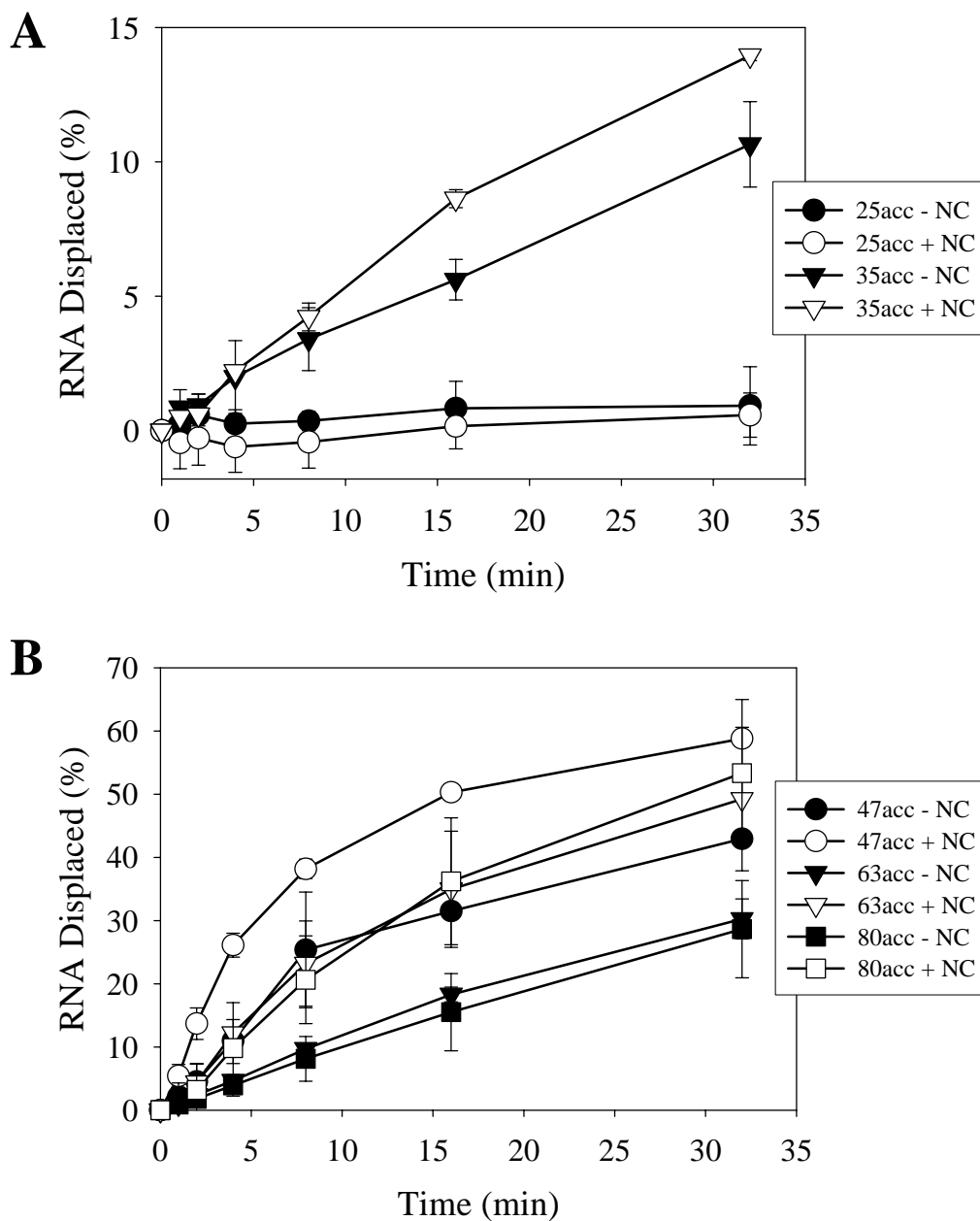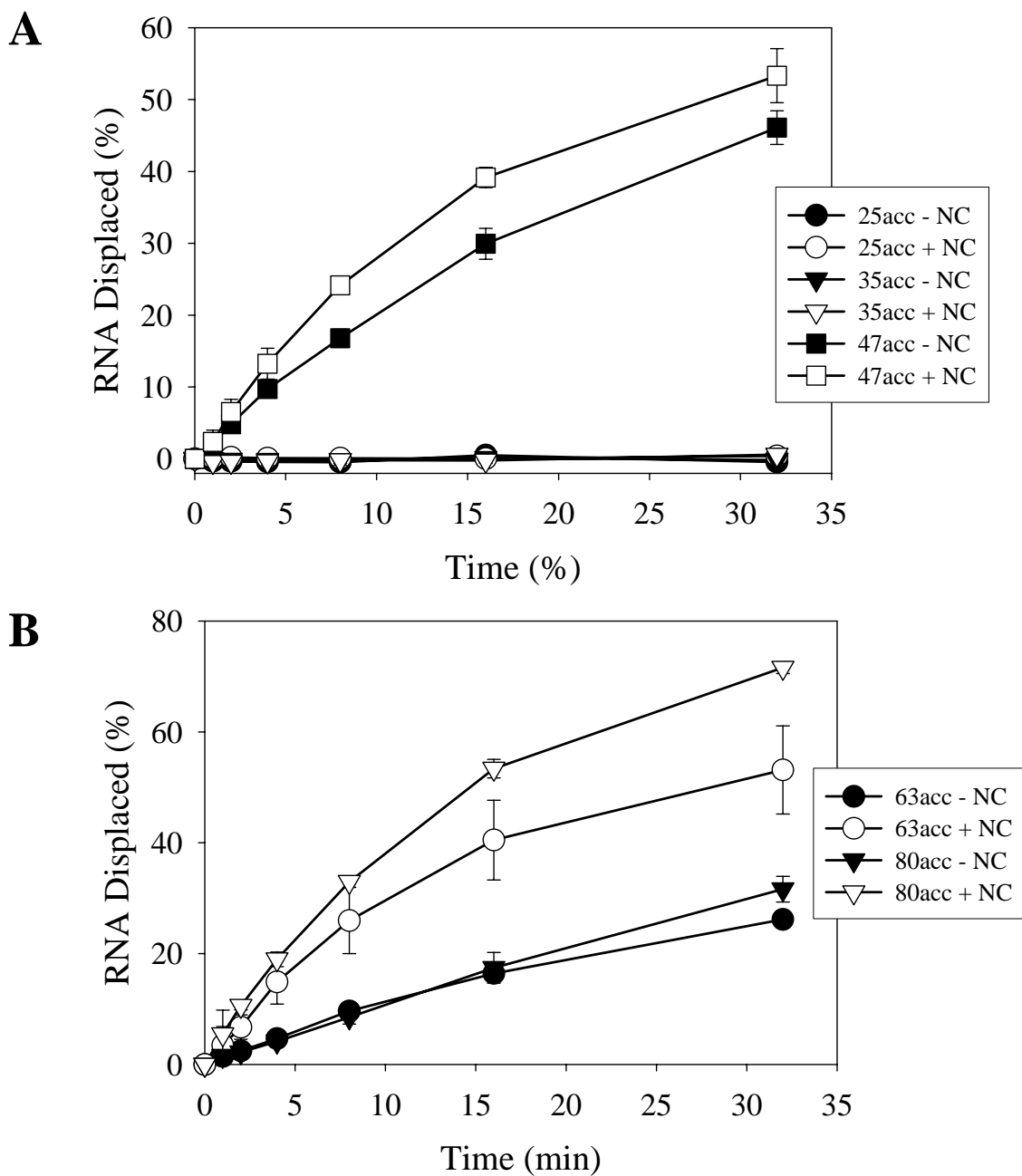
Figure 31: Donor displacement graphs of experiments completed with 35hyb. Graphs show the displacement of 35don in the presence (open symbols) and absence (closed symbols) of NC. A) Acceptors that form hybrids with less than 12 additional base pairs than the donor:DNA hybrid. B) Acceptors with hybrids longer than 28 additional base pairs. The legends show which acceptor was used in each experiment. Each line represents the average of three experiments and the error bars show the standard deviation.

there was a small amount of displacement observed. This acceptor could hybridize to the DNA for 10 base pairs outside the donor:DNA hybrid region. If the displacement observed in those reactions was simply due to some property of the acceptor other than the ability to anneal to those 10 nucleotides, we would expect to see a similar level of displacement in these experiments with 35hyb. However, no displacement is observed either with or without NC. This reiterates the fact that in this system the acceptor needs to hybridize outside the donor:DNA hybrid region to initiate strand exchange.

Displacement experiments were also done with the 35 base pair hybrid in the presence of the longer acceptors. These results were similar to the results observed with the 25 base pair hybrid. In Figures 30 and 31 the 35 nucleotide donor was displaced in the presence of 47acc, 63acc and 80acc. NC enhanced the displacement of the 35don in each of these reactions. In this case 47acc only had an additional twelve nucleotides which could anneal to the DNA outside the region of the donor:DNA hybrid. Similar to experiments with 25hyb and the 35acc which had an additional ten nucleotides, stimulation of this reaction with NC was evident but minimal. Together these results indicate that when there is only a short extension of nucleotides, less than 13, which can anneal to the DNA outside the hybrid region, NC will enhance transfer to this acceptor but not substantially. However, when the acceptor can anneal to regions longer than 22 nucleotides outside the hybrid region, optimal NC stimulation occurs but does not necessarily increase considerably with the length of the potential acceptor:DNA hybrid region.

NC promotes strand exchange to acceptors with complementary sequences in the region that anneals to the single stranded portion of the DNA. The results obtained from

displacement experiments indicated that the annealing of the region outside of the

donor:DNA hybrid is vital to the enhancement of transfer with NC. For considerable

stimulation of transfer an additional 22 nucleotides in the acceptor and DNA hybrid is

necessary. Therefore, we wanted to determine the importance of complete

complementarity in this region. We designed three mutant acceptors that were derived

from 47acc. These acceptors all had different point mutations in 47acc. The sequences

of the hybrid region between the mutant acceptors and the DNA are shown in Figure 32.

The mutations disturbed the complementarity between the acceptor and the DNA. The

first acceptor, 47acc6mut, had six mutations which were in the region that would anneal

to the single stranded DNA as well as the hybrid region between the donor and the DNA

(in the 25 base pair hybrid). The second acceptor, 47acc4mut, had four mutations. Three

mutations were in the region that would anneal to the single stranded region of the DNA

and one in the hybrid region of 25hyb. The last mutant acceptor, 47acc3mut, had three

mutations in the hybrid region of 25hyb. This acceptor was completely complementary

in the 22 nucleotide region that binds to the single stranded region of the DNA. The

change in thermodynamic stability between the donor:DNA and acceptor:DNA using

these acceptors with each of the hybrids is shown in Table 3. This table also summarizes

the results for strand exchange and donor displacement experiments. It is important to

note that though 47acc3mut and 47acc4mut have mutations in different regions the

change in the stability between the original hybrid used, either 25hyb or 35hyb, and the

hybrid with these acceptors, is essentially the same. This change is -11.2 and + 0.3

kcal/mol, with 25hyb and 35hyb, respectively. Therefore experiments with these

47acc6mut:

5' CUU UUG UUC C**G**U UUA GUG **U**GG GUU AAU U**A**C GAG **G**UU GG**G** GUA A**A**C AU 3'
*3' GAA AAC AAG GGA AAT CAC TCC CAA TTA AAG CTC GA*A CCG CAT TAG TA 5'

47acc4mut:

5' CUU UUG UUC CCU UUA GUG **U**GG GUU AAU U**A**C GAG **G**UU GGC GUA A**A**C AU 3'
*3' GAA AAC AAG GGA AAT CAC TCC CAA TT*A AAG CTC GA*A CCG CAT TAG TA 5'

47acc3mut:

5' CUU UU**C** UUC C**G**U UUA GUG AG**C** GUU AAU UUC GAG CUU GGC GUA AUC AU 3'
*3' GAA AAC AAG GGA AAT CAC TCC CAA TT*A AAG CTC   GA*A CCG CAT TAG TA 5'

Figure 32: Sequence of the hybrid regions of mutant acceptors and DNA. The sequence of the mutant acceptor is shown just above the DNA sequence the acceptor will bind to. From top to bottom the sequence of 47acc6mut, 47acc4mut and 47acc3mut are presented. The red nucleotides are the point mutations. The DNA that is bold italics is the portion that is originally bound to 25don in 25hyb, and the plain italics is the additional ten nucleotides that are initially bound to 35don in 35hyb.

Table 3: Increase in stability when DNA transfers from donor RNA to each acceptor, and summary of results of strand exchange and donor displacement experiments.

Hybrid	Acceptor	RNA Displaced - NC	RNA Displaced + NC	Strand Exchange - NC	Strand Exchange + NC	Change in Stability[a]
25 base pair	25acc	-	-	-	-	**0**
	35acc	+	+	-	+	**+ 11.5**
	47acc	++	+++	+	++	**+ 26.7**
	47acc3mut	+	++	X	X	**+11.1**
	47acc4mut	++	+	X	X	**+ 11.2**
	47acc6mut	-	-	X	X	**- 1.6**
	63acc	++	+++	+	+++	**> + 44.8[b]**
	80acc	++	+++	++	+++	**> + 44.8[b]**
35 base pair	25acc	-	-	X	X	**0**
	35acc	-	-	X	X	**0**
	47acc	++	+++	X	X	**+ 15.2**
	47acc3mut	+	+	X	X	**-0.3**
	47acc4mut	-	-	X	X	**- 0.3**
	47acc6mut	-	-	X	X	**-13.1**
	63acc	++	+++	X	X	**> + 33.3[b]**
	80acc	++	+++	X	X	**> + 33.3[b]**

Symbols are as follows: - none observed, + max observed < 20%, ++ max observed greater than 20%, less than 50 %, +++ max observed > 50 %, X = not tested

[a]Thermodynamic stabilities predicted by MELTING Program. An increase (+) in the change of stability indicates an increase in the negative value, therefore becoming more stable. A decrease (-) indicates a less stable hybrid.

[b]MELTING can only compute the stability of a hybrid with a maximum of 59 base pairs. Therefore the increase in the stability of these hybrids is greater than the given increase which was calculated for 59 base pairs of these hybrids.

acceptors test transfer to acceptors with no difference in thermodynamic stability but different complementarity.

Experiments completed with 35hyb and the mutant acceptors are shown in Figure 33. Also the results are displayed graphically in Figure 34A. Displacement in the presence of 47acc (Figure 31) showed that NC will enhance transfer to this acceptor, but only minimally, because there are only 12 additional nucleotides. However, when the mutated acceptors were used NC did not show any stimulation of transfer. For 47acc4mut and 47acc6mut no displacement was observed either with or without NC. With 47acc3mut a small amount of displacement was seen, but this displacement was not increased with NC. These results show that with acceptors where there is only a small number of additional nucleotides, less than 13, a completely complementary acceptor is vital for enhanced transfer with NC. It also indicates that the single stranded region of the DNA that anneals to the acceptor outside of the donor:DNA hybrid has a greater need for complete complementarity than the hybrid region. This is indicated because exchange is only observed in the mutant where the mutations are in the donor:DNA hybrid region, though this exchange is not enhanced by NC.

Displacement experiments were also completed with the 25hyb and the mutant acceptors. The results are shown graphically in Figure 34B and the autoradiograms are shown in Figure 35. With 47acc there is a clear stimulation of displacement, indicating exchange of the RNAs, in the presence of NC. When 47acc6mut was used with 25hyb no transfer was observed with or without NC. This result indicates that disturbing the complementarity between the acceptor and the DNA is detrimental to transfer. However, we wanted to pinpoint what region of complementarity is important for enhanced transfer

Figure 33: Autoradiograms of donor displacement with 35hyb and mutant acceptors. Autoradiograms show the displacement of 35don RNA in the presence of 47acc3mut (A), 47acc4mut (B), and 47acc6mut (C). The position of the donor:DNA hybrid is shown as well as the displaced RNA (ssRNA). NC was included in the reactions with a plus (+) symbol. Time is marked above each gel in minutes.
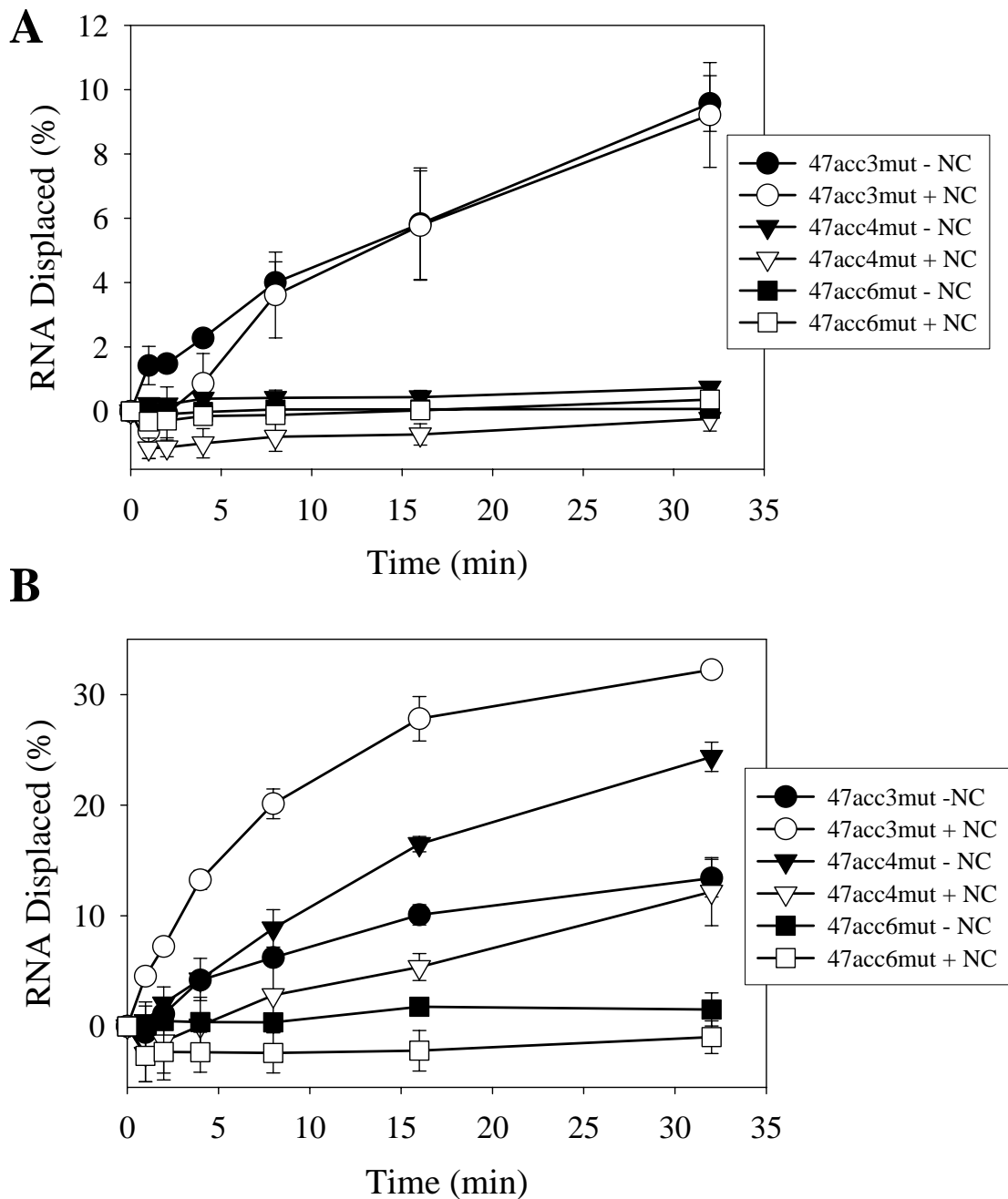
Figure 34: Graphs of donor displacement assays with mutant acceptors. Graphs show the percent RNA displaced over 32 minutes. A) Displacement experiments using 35hyb. B) Displacement experiments using 25hyb. Experiments done in the presence of NC are shown with open symbols and those without NC have filled circles. Legends show which acceptors are represented by each symbol. Experiments were repeated three times and the average is shown. Error bars represent the standard deviation.

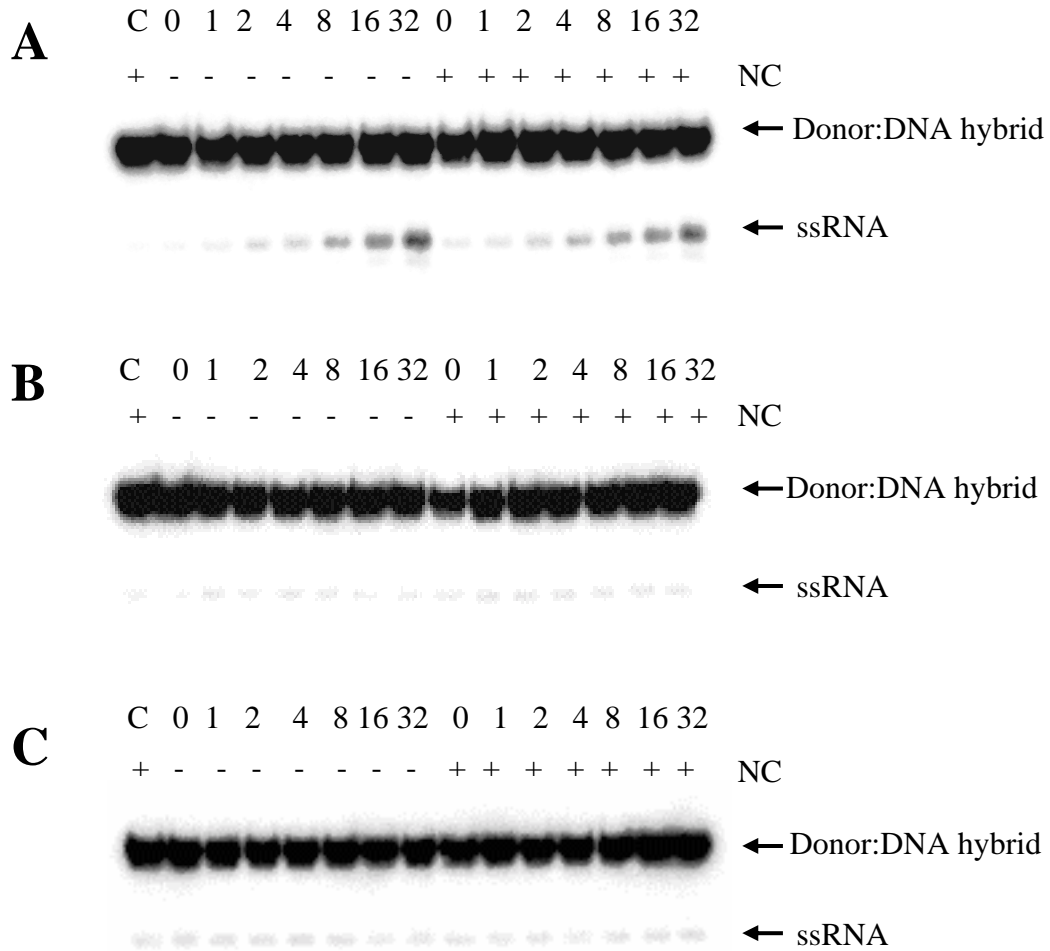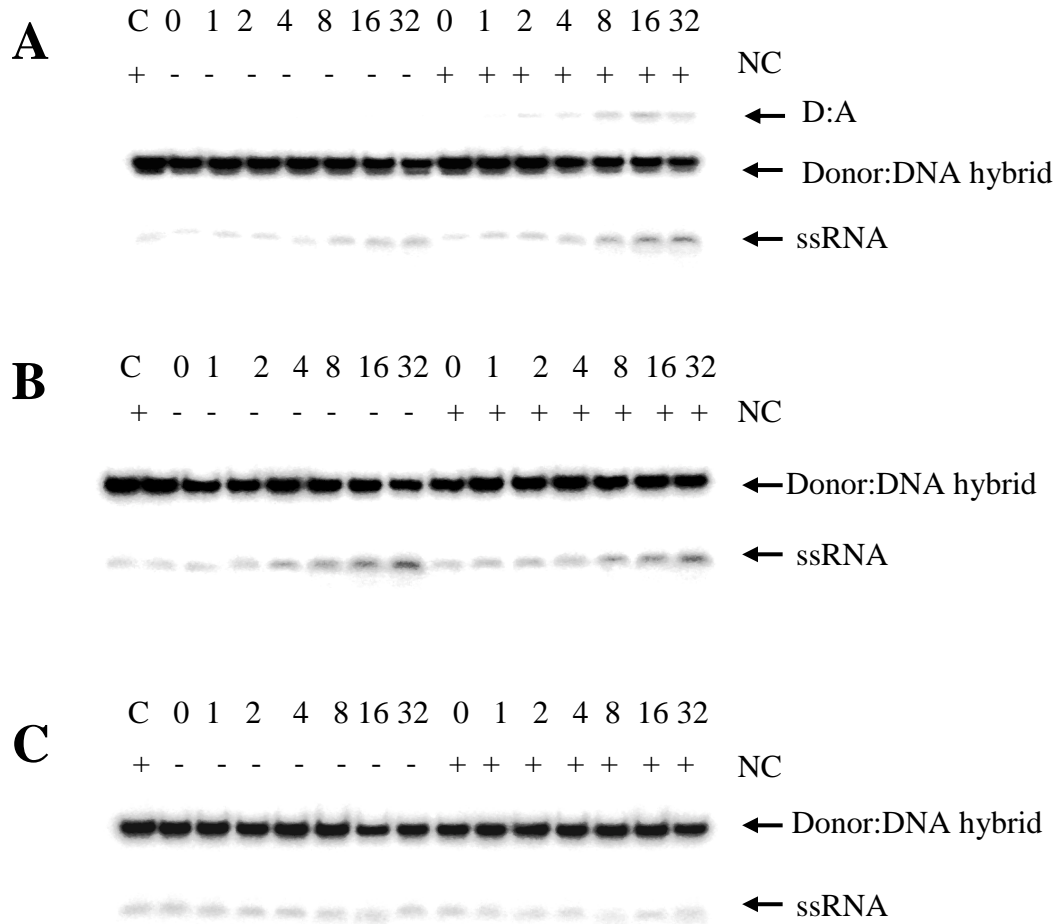**Figure 35: Autoradiograms of donor displacement assays with 25hyb and mutant acceptors.** Autoradiograms show the displacement of 25don RNA in the presence of 47acc3mut (A), 47acc4mut (B), and 47acc6mut (C). The position of the donor:DNA hybrid, donor bound to acceptor (D:A), and single stranded RNA (ssRNA) are highlighted. NC was included in the reactions with a plus (+) symbol. Time is marked above each gel in minutes.

with NC. Therefore experiments were completed with 47acc3mut and 47acc4mut. As was stated earlier, 47acc3mut has three mutations in the region of the donor:DNA hybrid, while 47acc4mut has one mutation in that region and three mutations in the single stranded region of the DNA. The results with 47acc3mut show a similar, though somewhat decreased, displacement profile to 47acc. With 47acc3mut there is a substantial amount of displacement without NC but the rate of displacement is considerably increased with NC. Therefore, when the mutations are in the donor:DNA hybrid region there is not a great effect on NC enhancement of exchange. NC will still greatly enhance exchange of RNAs with an acceptor that has three mutations that are not in the 22 nucleotide region that hybridizes to the single stranded portion of the DNA. The results with 47acc4mut are drastically different. With this acceptor without NC there is again a substantial amount of displacement. However, when the reaction is completed in the presence of NC, donor displacement is considerably inhibited. This result clearly shows the importance of the complementarity between the single stranded region of DNA and the acceptor. When this region is completely complementary NC will enhance transfer. However, when this region is not complementary transfer is actually inhibited with NC.

It is also important to note that the increase in stability of the hybrid from 25hyb to these acceptors is -11.2 kcal/mol. This is similar to the increase in the stability between 25hyb and the DNA bound to 35acc, which is -11.5 kcal/mol (see Table 3). However, NC did not have a large effect on the displacement with 35acc. These results clearly indicate that NC is not solely increasing exchange of RNAs based on the increase in the stability of the hybrid but rather on the potential to form a longer hybrid. Transfer

with NC was minimal to 35acc, and was inhibited to 47acc4mut, when the complementarity was disrupted. However, when the complementarity to the 22 nucleotide single stranded region of the DNA was not disrupted, as in 47acc3mut, NC clearly enhanced transfer, in a manner similar to 47acc, which has no mutations. These results show how vital the single stranded region of the DNA, which can anneal to a complementary region in the acceptor, is for NC enhancement of transfer. This also supports the acceptor facilitated model of transfer showing that in this system the acceptor must be complementary outside the donor:DNA hybrid for efficient stimulation of transfer by NC.

Discussion

HIV-1 NC is a nucleic acid chaperone that is able to enhance the process of strand transfer (recombination) during the viral replication cycle. Recombination leads to genetic diversity in HIV. This genetic diversity is important because it can help the virus evade effective drug therapies or vaccines.

Though NC has been shown to enhance strand transfer *in vitro*, there are still questions pertaining to the conditions which make strand transfer optimal. For example, does the acceptor play a role in strand exchange? If so what types of acceptors would be more likely to engage in transfer? As was stated in the Introduction there are two current models for internal strand transfer. The first model is the donor dissociation model (see Figure 18) where the donor is initially displaced from the nascent DNA before the DNA binds to the acceptor where synthesis continues. This model suggests that the acceptor does not play an active role in the strand transfer process, rather is simply available for

the DNA to anneal to after the RT and donor have fallen off of the DNA. The acceptor facilitated model (see Figure 19) suggests that the acceptor does in fact play an active role in the strand transfer process. This model implies that the acceptor initiates transfer when it anneals to the single stranded region of the nascent DNA. After the acceptor anneals to the DNA the donor is displaced by the closing of the DNA:acceptor hybrid and synthesis continues on the acceptor.

In both the acceptor facilitated and donor dissociation models for strand transfer it is easy to speculate where the chaperone activity of NC would be able to enhance transfer. In the donor displacement model NC could enhance the destabilization of the donor:DNA hybrid allowing this hybrid to fall apart quicker. This would enhance the number of DNA molecules available for transfer to the acceptor RNA. If this were the function of NC in strand transfer reactions, transfer of the DNA to acceptor RNAs would not be as dependant on the type of acceptor used in the experiment. NC would enhance the destabilization of the hybrid in the same manner using all types of acceptors. Experiments have shown that NC functions to weaken the interaction between the donor and nascent DNA using model substrates *in vitro* (35). In addition, NC can enhance the annealing of single stranded DNA to acceptor (see Fig 26). Therefore a role for NC in the donor dissociation model seems possible. In the acceptor facilitated model of strand exchange NC could play a role in both the destabilization of the donor:DNA hybrid as well as stimulating the annealing of the acceptor to the single stranded region of the DNA. As shown in the previous chapter NC clearly plays a role in enhancing the direct annealing of single stranded RNA and DNA, whether structure must first be destabilized by NC or not. This indicates that NC may enhance the annealing of the acceptor RNA

116

and DNA outside the region of the donor:DNA hybrid. If NC enhances this annealing, then the properties of the acceptor (i.e. length, complementarity) would be vital to the reaction. It would be expected that transfer to a longer acceptor would occur more readily in the presence of NC than to an acceptor with the same potential hybrid region as the donor.

The results presented in this chapter support the acceptor facilitated model for strand exchange (although they do not necessarily argue against the donor dissociation model (see below)). Figures 25, 29, and 31 show that in both the strand exchange and donor displacement experiments, acceptors that had the same hybrid region as the donor would not show strand transfer in the presence or absence of NC. Also, acceptors that had only ten additional nucleotides to hybridize to the single stranded region of DNA had just a slight enhancement in strand transfer with NC. With longer acceptors, that had at least 22 additional nucleotides, enhancement by NC was much greater. These results indicate first that the acceptor must have additional nucleotides that can anneal to the nascent DNA outside the region of the donor:DNA hybrid for strand exchange to occur and be enhanced by NC. They also show that there seems to be a threshold in the number of nucleotides that must be available outside the donor:DNA hybrid such that NC enhancement of strand exchange will be considerable. Once this number of nucleotides is reached there is not necessarily an increase in exchange with NC if additional nucleotides are added. These results support an active role for the acceptor in strand exchange. They also suggest that a major role of NC may be to enhance the annealing of the acceptor to the single stranded region of the DNA.

Since it is apparent that the annealing of the acceptor to the single stranded region of the DNA is vital for enhanced transfer of the DNA we wanted to determine whether the number of potential base pairs in this region held true for a stronger donor:DNA hybrid. Therefore we increased the hybrid from 25 base pairs to 35 base pairs. When displacement was tested, with an acceptor that had only 12 additional nucleotides that could anneal to the DNA, NC stimulation was evident, but not optimal. However when greater than 22 nucleotides was used (28, and 45) NC stimulation of strand exchange was high. These results show that the acceptor role in strand exchange is important even when a stronger donor:DNA hybrid is used. Also, because the DNA did not transfer to 35acc in these experiments, when it did slightly with the 25 base paired hybrid, the acceptor role is highlighted. This shows the significance of the single stranded region because 35acc was able to displace one donor (in 25hyb) but not the other (in 35hyb). The only difference in these experiments was that with 25hyb, 35acc could bind to ten nucleotides outside the donor:DNA hybrid, whereas there were no available nucleotides in the DNA in experiments with 35hyb.

The experiments discussed here clearly show that the acceptor is vital to the enhancement of strand exchange by NC. However, we also investigated how important the complementarity of the single stranded region of the DNA to the acceptor was by mutating the acceptor in different regions. These results indicated that the complementarity of the single stranded region is important for exchange to be stimulated by NC. When the 35 base pair hybrid was used no displacement was observed with either of the acceptors that had mutations in the region that would bind to the single stranded portion of the DNA. However, with acceptor that had mutations in the

118

donor:DNA hybrid region, but was free of mutations outside this region there was some displacement observed, though it was not enhanced by NC. This result indicates that when there are only 12 additional nucleotides in the acceptor:DNA versus donor:DNA hybrid complete complementarity is important for NC stimulation. This stimulation was slight with the completely complementary acceptor so when any mutations were introduced exchange was drastically decreased.

Displacement experiments with the 25 base pair hybrid and mutated acceptors revealed even more about the importance of complementarity between the acceptor and the DNA. When the acceptor was used that had mutations in both the regions inside and outside the donor:DNA hybrid no displacement was observed either with or without NC. When the acceptor was used which had mutations in the donor:DNA hybrid region but not the region exterior to that, NC clearly enhanced strand exchange, shown by a large stimulation of donor displacement. However, when the acceptor was used that had mutations within the region that binds to the single stranded portion of the DNA, there was a clear inhibition of donor displacement with NC. This indicates that NC inhibited transfer of the DNA to the acceptor with mutations outside the donor:DNA hybrid region whereas it enhanced transfer to the acceptor with the mutations inside this region. Both of these acceptors had the same increase in thermodynamic stability of the acceptor:DNA hybrid when compared to the donor:DNA hybrid (-11.2 kcal/mol). These results suggest that NC enhances transfer by enhancing the annealing of the complementary regions of the acceptor and single stranded DNA. When the complementary sequences are disturbed, NC actually can inhibit transfer. NC has previously been shown to inhibit some annealing reactions (55). These *in vitro* experiments showed that NC can inhibit

the rate of annealing between oligonucleotides with regions of complementarity less than 12 nucleotides, or longer oligonucleotides that do not form stable secondary structures. The results of these annealing experiments, and our assays, show that NC can prevent the formation of weak hybrids. As suggested in Chapter 2, NC may have both unwinding and hybridization activity. Specifically, the basic nature of NC may act to neutralize the positive charge of complementary nucleic acid strands. It may be that this hybridization activity can be prevented if the potential hybrid formed will not be strong enough. The unwinding activity of NC could, in this case, prevent the formation of weak hybrids. This is illustrated in Figure 36. NC chaperone activity lowers the activation energy for unwinding nucleic acid hybrids. There could be a threshold in the stability of a hybrid above which NC will allow hybrid formation and below which NC promotes destabilization. In the strand exchange experiments presented here, this threshold could have been reached for experiments with the completely complementary 22 nucleotide annealing region. However, when the complementarity was disturbed the threshold was not reached, and therefore NC unwinding activity prevented annealing of the acceptor and DNA in this region, and strand exchange was inhibited.

The results presented in this chapter reveal an interesting aspect of the acceptor properties involved in strand exchange. First it is clear that the binding of the acceptor to the single stranded region outside the donor:DNA hybrid region is essential for NC stimulated strand exchange. Also it is preferable if this region is longer than 20 nucleotides for maximal NC stimulation. Lastly these results show that the complementarity of this region is even more important than an increase in the thermodynamic stability from the donor:DNA to acceptor:DNA hybrids. For maximal
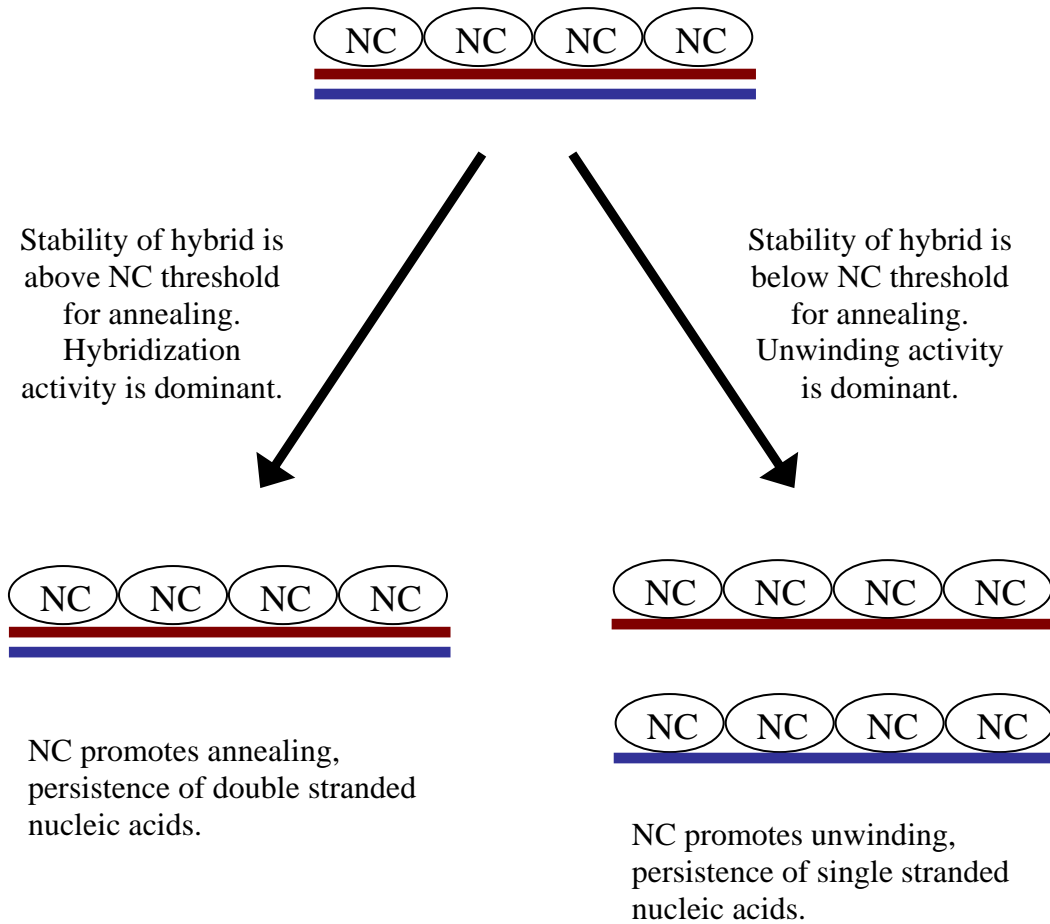
Figure 36: Model for NC chaperone activity on hybrids with different stabilities. The model presented here suggests that NC chaperone activity can vary due to potential hybrid being formed. NC lowers the activation energy for unfolding for all hybrids. This model indicates that there may be a threshold in this energy above which NC favors hybridization (left panel) and below which NC favors destabilization of the hybrid (right panel). NC has two activities, hybridization and unwinding. This model indicates that NC will use one of those activities depending on the type of hybrid in question, and the stability of that hybrid. The stability can differ between hybrids due to length, sequence and complementarity.

stimulation of strand transfer with NC the region of the DNA outside the donor:DNA should be completely complementary to the portion of the acceptor that can bind this region. Interestingly NC may even inhibit transfer of the DNA from the donor to a substantially longer acceptor with disturbed complementarity in the region of the single stranded DNA.

These results are interesting when considering the known facts about recombination. Nonhomologous recombination has been shown to occur approximately 1/100 – 1/1000 as frequently as homologous recombination (149). These results suggest that NC may even be inhibiting strand transfers in nonhomologous sequences whereas it enhances recombination in homologous sequences. When we used the acceptor with mutations inside the donor:DNA hybrid region that was completely complementary outside that region NC substantially stimulated transfer. This is important when considering genetic diversity because while NC can enhance transfer between homologous sequences it is not vital that the regions downstream the acceptor:DNA annealing site be completely complementary for enhanced transfer. The DNA being synthesized on the donor RNA can transfer to a homologous region on another RNA (or a homologous region on the same template) where synthesis continues. It can then obtain new sequences downstream from the homologous region where transfer was initiated. This can clearly occur close to the region of homology if NC will enhance transfer to a sequence that is complementary for just 22 nucleotides before having mismatched sequences in a 25 base pair hybrid region. This is important considering experiments that showed recombination occurs in nonhomologous regions at points of less than eight

nucleotide sequence homology (144, 149). The annealing of the DNA and acceptor at these sequences can be enhanced by NC due to the complementary sequences.

It should be noted that although the results support the acceptor facilitated model they do not argue against strand transfer by donor dissociation. Previous *in vitro* results have demonstrated both types of transfer depending on the nucleotide sequences that were examined (35). In addition, recent results from our lab suggest a donor dissociation type mechanism occurs at a strong pause site in the *gag-pol* frameshift region while acceptor facilitation mediated transfer in a weakly structured region of the *env* gene was observed (manuscript in preparation). It is likely that both mechanisms can occur and the type would depend on the structure and sequence of the acceptor and donor. Also, in one report donor dissociation mediated transfer was shown to occur at a strong pause site after RT RNase H activity had shortened the region of homology between the donor and nascent DNA to about 13 bases (35). Since the hybrids used in this thesis were considerably longer and RT was not added to the reactions it is not surprising that dissociation of the donor in the absence of acceptor with longer hybrid regions was not observed.

These results highlight the importance of even short sequences of complementarity in a stretch of non-complementary sequences. It is these sequences in which NC can enhance annealing to stimulate strand transfer. Strand exchange is stimulated by NC when there is a stretch of complementary sequences between the acceptor and the DNA that can anneal outside the donor:DNA hybrid region. When the complementary sequence is longer than 22 nucleotides NC stimulation of strand transfer can be quite high. Internal strand transfer can lead to the generation of progeny viruses

with increased genetic diversity. Therefore, it is important to understand the process of internal strand transfer and what can be done to manipulate this process such that we can minimize the ability of the virus to evade drug treatments and effective vaccines.

Chapter 4

General Discussion

The results presented here broaden our understanding of HIV-1 NC chaperone activity. This chaperone activity is vital for the survival of HIV and therefore, understanding it in its entirety may be important to overcoming the virus. The ultimate goal for researchers studying HIV is to develop better drug therapies and effective vaccines such that this virus can be eradicated. In order to obtain this goal every step of the viral life cycle must be studied extensively such that it is clear which steps can be exploited to terminate the illness caused by this virus.

HIV-1 NC chaperone activity is a process utilized in a number of steps throughout the life cycle. This activity is important for enhancing optimal interactions between RNA and RNA, as well as DNA and RNA, and DNA and DNA. Nucleic acid chaperones act to prevent and unfold nucleic acid misfolds (67, 87). Misfolds are folds that are not thermodynamically favored but may occur due to interactions of the nucleic acids that occur previous to the interactions that would lead to the formation of more thermodynamically favorable structures.

There are a number of known RNA chaperones that are not derived from a virus. The *Escherichia coli* S12 ribosomal protein binds nucleic acids nonspecifically, similar to HIV-1 NC (67). *In vitro* assays have shown that this protein can aid in the folding of group I introns for efficient splicing. Also, the heterogeneous nuclear ribonucleoprotein A1, which binds mRNAs *in vivo*, has been shown to have chaperone activity. This protein has been shown to facilitate annealing of complementary nucleic acids (87).

Also, recent work has shown that the RNA-dependant ATPase CYT-19 has chaperone activity that aids in the splicing of the mitochondrial large subunit rRNA intron (87). In evaluating cellular processes, from mRNA synthesis to translation it is clear that RNA chaperones must be important to proper protein synthesis. Also, protein chaperones, such as the heat shock proteins, aid in the proper folding of other proteins during synthesis (116). These chaperones are important for the production of functional proteins. Therefore, it is not surprising that viruses have also obtained a need for chaperones.

HIV-1 has a diploid RNA genome that forms a dimer within the virion. Each RNA contains approximately 9.1 kb. The virus must undergo a number of steps to effectively replicate the genome. Therefore, the need for a nucleic acid chaperone is quite apparent. This dimeric genome packs into the virion core. The virus has a diameter of about 100 nanometers. Consequently, this packing is very compact, but also must be functional such that the virus will be able to undergo reverse transcription upon entry into the newly infected cell. The correct dimer must form with the interactions beginning at the DIS. Also tRNA annealing must occur exactly at the PBS such that the correct proviral DNA is synthesized. Interestingly, though there are several steps where nucleic acid chaperone activity is important, this activity is contained in a single viral protein. Taking this into account it is not surprising many mutations in NC sequences leads to formation of non-infectious particles or prevents virion formation completely, depending on the severity of the mutation (1, 12, 54, 58, 59, 101, 130).

HIV-1 NC zinc fingers display different types of chaperone activity.

Since chaperone activity affects a number of processes that involve different combinations of nucleic acids we investigated both nucleic acid annealing and strand transfer (strand exchange). Though strand transfer is specific to reverse transcription, the annealing assays provide information that pertains generally to the chaperone activity. The results from the annealing assays give insight into the importance of zinc fingers in steps such as genome dimerization and maturation and tRNA:PBS annealing. In summary, these results showed that the first zinc finger is important for the unwinding of strong nucleic acid structures. Though the second zinc finger alone can not enhance the unfolding of strong structures, it was shown to have an accessory role in annealing. Wild type levels of annealing with strong structures only occurred when both fingers were present in their native position (*ie.* with wild type protein).

Evaluating the different steps of the life cycle that have been discussed we can see a role for each of the fingers. In order for dimerization to occur the DIS sequences in each RNA must come together. The stem loops then unfold to form the extended duplex dimer. Annealing assays have been performed with the DIS and NC peptides containing only the N-terminal region and the linker peptide. These experiments indicated that the zinc fingers were not necessary for dimerization (129). However, ten times as much of the mutant peptide compared to wild type NC was necessary for the formation of the extended duplex dimer. Therefore, it would be interesting to see if the extended duplex dimer would form with 1.1, 2.1 or 2.2 NC. Our results indicate that 1.1 and 2.1 NC would probably enhance the formation of this dimer because the first zinc finger would be able to unfold the stem loops and enhance complete annealing of the dimer. Also, the

first zinc finger may be important for maturation of the genome. NC has been shown to enhance genome maturation, which is the process by which the genome packs into a more condensed and stable structure that can fit inside the virion core. Because both virions are packaged into the virus when budding from the membrane, before the compact core forms, it is likely that there are a number of misfolds that form between the two RNAs or within one RNA. Because finger one was shown to be important for unfolding, this finger is most likely critical during the maturation process when a number of misfolds must be corrected. Also, there has been some discrepancy in the role of NC in unfolding the tRNA to prepare it for annealing to the PBS. Our results indicate that if in fact NC is important for the unfolding of tRNA, finger one must be present for efficient unfolding to occur. Taking these steps into account it is understandable that substitution of finger two for finger one in *in vivo* infection assays leads to creation of a virus that is replication deficient (58).

However, our results also indicated an accessory role for the second zinc finger of HIV. Our assays showed that 2.2 NC enhanced annealing of unfolded nucleic acids, and weakly structured nucleic acids. This indicates that the second zinc finger may also be important for a number of annealing steps. In each annealing step that requires the unfolding of nucleic acids there are two steps, the unfolding and the hybridization. It is not clear whether these steps occur concurrently or whether unfolding is complete before hybridization occurs. Either way the hybridization of unfolded portions can be enhanced by the second zinc finger. There is no discrepancy regarding the importance of the basic residues of NC in nucleic acid annealing. Assays have been done with NC mutants that do not contain the zinc fingers. Some of these assays show that the residues outside the

128

zinc fingers can enhance annealing. However, when basic regions are deleted annealing is diminished (29, 30, 83). This is not at all surprising considering that the basic residues are also completely vital for NC binding to nucleic acids. Therefore, it could be that the basic residues are important for binding while the zinc fingers enhance annealing. More likely, the basic residues do have some annealing activity as well. Either way our results indicate that mutant proteins containing two copies of the C-terminal zinc finger do have some annealing activity with weakly structured nucleic acids. This mutant contained all the native basic residues so this could be an effect of these residues or the zinc finger. This is going to be tested further. However, if the second zinc finger were not at all important in annealing we would have expected 1.1 NC to have wild type level of activity with all of the structures, and interestingly this protein was only comparable to wild type with the unstructured nucleic acid, thereby implicating the second zinc finger with an accessory role for annealing. This accessory role in annealing may not be important for all parts of the life cycle that require chaperone activity, though some steps may proceed at lowered efficiency if the second zinc finger were not present. While the unfolding activity of the first zinc finger is essential for annealing when strong structures are present in the nucleic acids, the activity of finger 2 may not be essential to annealing. *In vivo* assays with viruses containing 1.1, 2.2 and 2.1 NC have shown that only viruses with 1.1 NC were infectious (58). These viruses were initially defective, replicating at much lower levels than wild type viruses. However, over 35 days the viruses reverted to wild type phenotype and were able to replicate efficiently.

The cause of the unwinding activity found in the first zinc finger is unknown. There are five amino acids that differ between finger 1 and finger 2 in the protein derived

from strain NL4-3: Phenylalanine to tryptophan, asparagine to lysine, isoleucine to glutamine, alanine to methionine and asparagine to aspartate. Overall finger 1 is slightly more hydrophobic. Hydrophobic interactions with the bases could aid in hybrid destabilization by interfering with base stacking. We are currently designing NC mutants with single and double point mutations that will help to investigate which of these residues play a role in unwinding nucleic acids. These proteins will be tested in annealing assays to determine their ability to enhance annealing between RNAs and DNAs with a high degree of secondary structure. It will also be important to test these mutant NC proteins *in vivo*. Mutant viruses will be created with the different NC proteins to determine the affect the mutations have on virus formation, RNA packaging, reverse transcription and virus infectivity. If the *in vitro* annealing assays indicate that certain amino acids have a role in unwinding we would expect that the mutations would be severely detrimental to virus survival *in vivo*. All NC mutants should be tested both *in vitro* and *in vivo* to be certain that the results of annealing assays correlate with NC activity in replicating viruses.

Taken together the results presented here suggest that both zinc fingers are important for annealing, though the role of the first zinc finger is vital while the second is merely accessory. This agrees with *in vivo* results that indicate that the native sequence of the first zinc finger is necessary for the persistence of a replication competent virus (58, 59). Therefore, considering drug therapies it would be more important to focus on the first zinc finger. If the activity of the N-terminal zinc finger can be disrupted, perhaps the virus will not persist. Some NC inhibitors have been tested *in vitro* and *in vivo* (71). Four compounds, 3-nitrosobenzamid (NOBA), disulfide benzamide (DIBA), dithiane,

130

and azodicarbonamide (ADA) were investigated. They tested the effect of each treatment on nucleic acid binding, and viral replication. These compounds initiate the release of zinc from the NC zinc fingers, thereby disturbing their structure and potentially their chaperone activity. When compared with three zinc finger proteins found in the cell, three of these compounds were shown to be unable to initiate release of zinc from the cellular proteins (71). The investigators also performed molecular modeling analysis to investigate the potential interactions of the compounds with each of the zinc fingers. They found that the second zinc finger is probably more reactive than the first zinc finger, mostly due the serine in C49 being the most nucleophilic site of the zinc fingers (71). Because the first zinc finger is more likely the finger that would need to be disturbed for efficient inactivation of the protein, more inhibitors should be evaluated. It would be favorable to find an inhibitor that can inactivate the unfolding capability of finger one. It is important to determine which residues in finger 1 are important for the unwinding activity of this finger. These residues could potentially be specifically targeted in order to increase the specificity of the treatment on interfering with NC unwinding activity but not interfering with cellular zinc finger protein function. Even with the accessory role that finger two can play, if unfolding activity is removed, most, if not all, of the steps on which the chaperone activity is displayed would be disturbed.

Due to the importance of NC to several steps throughout the viral life cycle it is a desirable target for vaccine development. However, this is a difficult task for a number of reasons. HIV-1 NC would most likely not be a good protein peptide vaccine. NC is bound to the RNA within the virion core so it is not presented to the immune system on the viral surface. Though there are antibodies against NC, in viral infections it is difficult

131

to imagine that these antibodies are highly effective in fighting infection because NC is enclosed in the virion. Considering live attenuated vaccines it would be more important to inactivate finger one rather than finger two. This is because there is evidence that infection of virus with a mutated finger two but wild type finger one can revert back to a virus with wild type phenotype (58). However, due to the high mutation rate of HIV and its ability to revert to wild type after many different types of mutations it is unlikely that a live attenuated vaccine will ever prove safe for protection against this virus. Though it might be difficult to discern how NC may be targeted in vaccine development, it should not be ruled out as a target because of the importance of this protein throughout the life cycle.

HIV-1 NC enhances annealing of acceptor and DNA to stimulate strand exchange.

Strand transfer is also a process in the life cycle where HIV NC chaperone activity plays a key role. There are three kinds of strand transfer which NC has been shown to enhance, minus and plus strand strong stop transfers, and internal transfer. The last type of transfer is in theory, not obligatory to successful replication but leads to recombination between the two genomes in the core. HIV undergoes an average of approximately three internal transfers occur during each replication cycle (72). Therefore, although the proposed replication mechanism does not include a specific internal transfer step, such transfer may well be necessary for efficient replication. In any regard, internal strand transfer is important for the generation of viral genetic diversity. The implications of this type of transfer are obvious for vaccine and drug development. Recombination increases genetic diversity and the potential evolution of resistant viruses.

Therefore, it is important to understand by what mechanism internal strand transfer occurs and what can possibly be done to prevent the occurrence of this process.

The results presented here are somewhat unexpected when taking into account NC chaperone activity. It might be expected that NC would always enhance the formation of the most thermodynamically stable hybrid. However, these results show that thermodynamics are not always the ultimate determinant in NC enhanced strand exchange. There were some instances where NC did not substantially increase transfer to an acceptor that would form a significantly more thermodynamically stable hybrid. When mutant acceptors were used results showed that NC will enhance transfer to one mutant acceptor, and inhibit transfer to another, even when both have the same increase in thermodynamic stability. These experiments indicated that the parameter most important for strand exchange enhanced by NC was a complementary region on the acceptor that could bind to the single stranded region of the DNA.

The results from experiments with mutated acceptors illustrate an important point regarding NC chaperone activity. It was evident that NC did not strictly enhance formation of the strongest hybrid. Instead, NC promoted formation of hybrids that had complementary sequences in the single stranded region where annealing occurred between the DNA and acceptor (that was greater than or equal to 22 nucleotides). This is most likely because NC chaperone activity may have a threshold of stability above which it favors hybrid formation and below which it favors hybrid destabilization (see Fig. 36). Therefore, in experiments where there was complete complementarity in the region where the DNA and acceptor could anneal, exchange was promoted by NC due to a stimulation of annealing. This was not true when this region was not completely

complementary. In experiments with disturbed complementarity in this region NC inhibited hybrid formation.

It would be advantageous to test the hypothesis that NC has a stability threshold above which NC will favor hybrid formation and below which it favors destabilization. One way this could be tested would be to use the unstructured RNA sequence from Chapter 2. When this sequence is shortened incrementally by five nucleotides to 37, 32, 27, 22, 17, 12 and 7 nucleotides it remains unstructured as predicted by mFold and RNAdraw. These oligonucleotides could be synthesized and annealing assays could be performed with each of the RNAs and their complementary DNAs. If NC has a stability threshold like we propose, we would expect there would be a point where NC no longer stimulates annealing of the RNA and DNA sequences, but rather inhibits hybrid formation.

Strand exchange experiments indicated that NC enhanced exchange was initiated by the acceptor binding to the single stranded region of the DNA. This supports the acceptor facilitated model for strand transfer which suggests that the acceptor binds to the single stranded region of the DNA exterior to the donor:DNA hybrid (see Figure 19). After binding to the single stranded region the annealing of the acceptor migrates downstream through the donor:DNA hybrid region and displaces the donor. Then synthesis can continue with the acceptor as the new template. In the acceptor facilitated model of transfer NC most likely enhances the binding of the acceptor and the single stranded region of the DNA. The closing of the acceptor:DNA hybrid displaces the donor rather than NC initiated donor displacement. However NC chaperone activity

could also enhance this displacement after the closing of the acceptor:DNA hybrid begins.

The donor dissociation model states that the donor and the DNA first dissociate before the annealing of the acceptor and the DNA. There are two steps in the donor displacement model that could potentially be enhanced by NC chaperone activity. The first is the destabilization of the donor:DNA hybrid and secondly the annealing of the DNA and the acceptor. Experiments presented here showed that NC did not enhance any dissociation of the donor:DNA hybrid when short acceptors were used. The hybrid used in these experiments had 25 base pairs. Others have shown that when RT pauses at a hairpin additional RNase H cleavages can occur that will cleave the hybrid to ten to fifteen base pairs (36, 42, 142). NC would most likely enhance the dissociation of a weak hybrid with less than fifteen base pairs. Donor dissociation has been shown to be a mechanism of transfer occurring at strong pause sites (13, 35, 37, 113). Our experiments did not test this mechanism directly due to the stronger hybrid that was used and the lack of RT in these reactions.

HIV-1 NC enhanced exchange to mutant acceptors that had a completely complementary acceptor:DNA binding region exterior to the donor:DNA hybrid. This gives insight to why homologous transfer occurs 100 to 1000 times as much as nonhomologous transfer (149). Homologous transfer could be favorable because the annealing of the acceptor and DNA is an important initiator for transfer. When the complementarity was disturbed in this region NC even inhibited transfer in some cases.

This result, in conjunction with the previous results regarding zinc fingers in annealing, indicate that both zinc fingers may be important for strand transfer. Because

135

both fingers played a role in enhancing annealing, though only the first zinc finger was important for unfolding, both fingers are probably important for enhancing the annealing of the acceptor and the DNA. This step is vital to enhanced transfer by NC in the experiments shown here, and most likely many cases *in vivo* as well. Therefore it is easy to speculate that the zinc fingers are important for strand transfer because they enhance the annealing of the acceptor and the DNA. Others have shown the zinc fingers are necessary for optimal minus and plus strand transfer (62). The results shown here may show one of the reasons they are important for transfer, being that they aid in the annealing of the acceptor and DNA to initiate transfer. Also, the nucleic acids involved in transfer could potentially be folded in conformations that inhibit transfer and therefore the first zinc finger is probably important as well for unfolding the acceptor and DNA structures making them available for annealing.

One important question for those investigating recombination is how it can be prevented. Others have shown that Actinomycin D can inhibit strand transfer by inhibiting the annealing of the RNA acceptor and the DNA (26, 63). Even when NC was included in reactions Actinomycin D was able to inhibit transfer. Also 4-chlorophenylhydrazone of mesoxalic acid inhibits the RNase H activity of RT, thereby inhibiting transfer (27, 50). From our results it would be suggested that one way to minimize transfer is to package two highly non-homologous genomes in one virus. However, this is not something that can be controlled, and all HIV strains within each group have some level of sequence homology. Because it is not easy to manipulate the genome and packaging in this way, the NC protein may be a more realistic target when trying to prevent recombination.

Clearly NC enhances recombination. In the results presented here NC enhances transfer by enhancing the annealing of the acceptor and single stranded region of the DNA. This is optimal when the acceptor and DNA are completely complementary and when there are long regions, in excess of 20 nucleotides that can anneal. As nucleic acids increase in length they increase in their ability to form secondary structures and potentially strong structures. Therefore, most often the acceptor RNA *in vivo* will need to be unfolded to be available for annealing to the DNA. These issues point to an important target for drug therapy to fight HIV. The first zinc finger in HIV-1 NC clearly plays an important role in the chaperone activity of NC. Other experiments have also exhibited the unwinding activity of the first zinc finger. Single molecule DNA stretching experiments showed that the first zinc finger must be present for unwinding of DNA (139).

HIV-1 NC zinc finger 1 is an important target for drug therapy.

If the unfolding activity of NC can be disturbed the HIV life cycle will quickly be halted. This activity is important for so many steps and the replication process in particular will not proceed without NC unwinding activity. Therefore finger one of NC is one of the best targets to inactivate NC's chaperone activity. Though disruption of the basic residues can cause disruption of NC:nucleic acid binding, there are many cellular proteins that have a high content of basic residues that could interfere with targeting NC basic residues with drug therapy. The same can be said for the zinc fingers. In fact in a recent report the zinc fingers found in CNBPs were substituted for the NC zinc fingers and most of the mutants were found to produce infectious viruses (90). Therefore, it is

137

quite difficult to target NC zinc fingers without affecting CNBPs and potentially a number of other cellular zinc finger proteins. However, other compounds have been investigated that do target NC zinc fingers while having no effect on specific cellular zinc fingers (71). Additionally, others have used another peptide, RB 2121, that has all of the structural features of NC, to compete with NC in replication assays (44). This peptide inhibited the amount of DNA synthesized, probably by inhibiting the formation of the RT:NC:nucleic acid complex necessary for reverse transcription. This indicates that a peptide could potentially be constructed that would be similar to NC but may not have the unfolding ability of the first zinc finger. Therefore, it is possible that a compound may possibly be devised that would interfere with the activity of finger one but have minimal side effects on cellular zinc fingers.

The work presented here further defines and clarifies the mechanism for the chaperone activity of HIV-1 NC. The zinc fingers found in this protein are important for its chaperone activity and finger one is particularly important for nucleic acid unfolding. Internal strand transfer, which leads to recombination is initiated by the acceptor:DNA annealing. The complementary region outside the donor:DNA hybrid is critical for transfer enhanced by NC. Annealing is most likely facilitated by the zinc fingers of NC and the unfolding of the acceptor and DNA facilitated by finger one. This makes finger one a prime target and finger two a potential target for drug therapy. If NC chaperone activity can be disturbed then the persistence of HIV infections within and individual can be eliminated.

References

1. **Akahata, W., E. Ido, and M. Hayami.** 2003. Mutational analysis of two zinc-finger motifs in the nucleocapsid protein of simian immunodeficiency virus mac239. J. Gen. Virol. **84:**1641-8.

2. **Allain, B., M. Lapadat-Tapolsky, C. Berlioz, and J. L. Darlix.** 1994. Transactivation of the minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the retroviral genome. EMBO J. **13:**973-81.

3. **Allen, P., B. Collins, D. Brown, Z. Hostomsky, and L. Gold.** 1996. A specific RNA structural motif mediates high affinity binding by the HIV-1 nucleocapsid protein (NCp7). Virology **225:**306-15.

4. **Amarasinghe, G. K., R. N. De Guzman, R. B. Turner, K. J. Chancellor, Z. R. Wu, and M. F. Summers.** 2000. NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. J. Mol. Biol. **301:**491-511.

5. **Amarasinghe, G. K., J. Zhou, M. Miskimon, K. J. Chancellor, J. A. McDonald, A. G. Matthews, R. R. Miller, M. D. Rouse, and M. F. Summers.** 2001. Stem-loop SL4 of the HIV-1 psi RNA packaging signal exhibits weak affinity for the nucleocapsid protein. structural studies and implications for genome recognition. J. Mol. Biol. **314:**961-70.

6. **Azoulay, J., J. P. Clamme, J. L. Darlix, B. P. Roques, and Y. Mely.** 2003. Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations. J. Mol. Biol. **326:**691-700.

7. **Balakrishnan, M., B. P. Roques, P. J. Fay, and R. A. Bambara.** 2003. Template dimerization promotes an acceptor invasion-induced transfer mechanism during human immunodeficiency virus type 1 minus-strand synthesis. J. Virol. **77:**4710-21.

8. **Barat, C., V. Lullien, O. Schatz, G. Keith, M. T. Nugeyre, F. Gruninger-Leitch, F. Barre-Sinoussi, S. F. LeGrice, and J. L. Darlix.** 1989. HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. EMBO J. **8:**3279-85.

9. **Beltz, H., J. Azoulay, S. Bernacchi, J. P. Clamme, D. Ficheux, B. Roques, J. L. Darlix, and Y. Mely.** 2003. Impact of the terminal bulges of HIV-1 cTAR DNA on its stability and the destabilizing activity of the nucleocapsid protein NCp7. J. Mol. Biol. **328:**95-108.

10.   **Berglund, J. A., B. Charpentier, and M. Rosbash.** 1997. A high affinity binding site for the HIV-1 nucleocapsid protein. Nucleic Acids Res. **25:**1042-9.

11.   **Berkowitz, R., J. Fisher, and S. P. Goff.** 1996. RNA packaging. Curr. Top. Microbiol. Immunol. **214:**177-218.

12.   **Berthoux, L., C. Pechoux, M. Ottmann, G. Morel, and J. L. Darlix.** 1997. Mutations in the N-terminal domain of human immunodeficiency virus type 1 nucleocapsid protein affect virion core structure and proviral DNA synthesis. J. Virol. **71:**6973-81.

13.   **Buiser, R. G., R. A. Bambara, and P. J. Fay.** 1993. Pausing by retroviral DNA polymerases promotes strand transfer from internal regions of RNA donor templates to homopolymeric acceptor templates. Biochim. Biophys. Acta. **1216:**20-30.

14.   **Bukrinsky, M. I., N. Sharova, T. L. McDonald, T. Pushkarskaya, W. G. Tarpley, and M. Stevenson.** 1993. Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. Proc. Natl. Acad. Sci. USA **90:**6125-9.

15.   **Cameron, C. E., M. Ghosh, S. F. Le Grice, and S. J. Benkovic.** 1997. Mutations in HIV reverse transcriptase which alter RNase H activity and decrease strand transfer efficiency are suppressed by HIV nucleocapsid protein. Proc. Natl. Acad. Sci. USA **94:**6700-5.

16.   **Carteau, S., R. J. Gorelick, and F. D. Bushman.** 1999. Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase in vitro: stimulation by the viral nucleocapsid protein. J. Virol. **73:**6670-9.

17.   **Cen, S., Y. Huang, A. Khorchid, J. L. Darlix, M. A. Wainberg, and L. Kleiman.** 1999. The role of Pr55(gag) in the annealing of tRNA3Lys to human immunodeficiency virus type 1 genomic RNA. J. Virol. **73:**4485-8.

18.   **Clever, J., C. Sassetti, and T. G. Parslow.** 1995. RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. J. Virol. **69:**2101-9.

19.   **Clever, J. L., R. A. Taplitz, M. A. Lochrie, B. Polisky, and T. G. Parslow.** 2000. A heterologous, high-affinity RNA ligand for human immunodeficiency virus Gag protein has RNA packaging activity. J. Virol. **74:**541-6.

20.   **Coffin, J. M.** 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science **267:**483-9.

21. **Coffin, J. M.** 1979. Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. J. Gen. Virol. **42:**1-26.

22. **Coffin, J. M., S. H. Hughes, and H. E. Varmus.** 1997. Retroviruses. Cold Spring Harbor Laboratory Press, United States of America.

23. **Conticello, S. G., R. S. Harris, and M. S. Neuberger.** 2003. The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. Curr. Biol. **13:**2009-13.

24. **Darlix, J. L., M. Lapadat-Tapolsky, H. de Rocquigny, and B. P. Roques.** 1995. First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses. J. Mol. Biol. **254:**523-37.

25. **Darlix, J. L., A. Vincent, C. Gabus, H. de Rocquigny, and B. Roques.** 1993. Trans-activation of the 5' to 3' viral DNA strand transfer by nucleocapsid protein during reverse transcription of HIV1 RNA. C. R. Acad. Sci. III **316:**763-71.

26. **Davis, W. R., S. Gabbara, D. Hupe, and J. A. Peliska.** 1998. Actinomycin D inhibition of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase and nucleocapsid protein. Biochemistry **37:**14213-21.

27. **Davis, W. R., J. Tomsho, S. Nikam, E. M. Cook, D. Somand, and J. A. Peliska.** 2000. Inhibition of HIV-1 reverse transcriptase-catalyzed DNA strand transfer reactions by 4-chlorophenylhydrazone of mesoxalic acid. Biochemistry **39:**14279-91.

28. **De Guzman, R. N., Z. R. Wu, C. C. Stalling, L. Pappalardo, P. N. Borer, and M. F. Summers.** 1998. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. Science **279:**384-8.

29. **De Rocquigny, H., D. Ficheux, C. Gabus, B. Allain, M. C. Fournie-Zaluski, J. L. Darlix, and B. P. Roques.** 1993. Two short basic sequences surrounding the zinc finger of nucleocapsid protein NCp10 of Moloney murine leukemia virus are critical for RNA annealing activity. Nucleic Acids Res. **21:**823-9.

30. **De Rocquigny, H., C. Gabus, A. Vincent, M. C. Fournie-Zaluski, B. Roques, and J. L. Darlix.** 1992. Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers. Proc. Natl. Acad. Sci. USA **89:**6472-6.

31. **Demene, H., C. Z. Dong, M. Ottmann, M. C. Rouyez, N. Jullian, N. Morellet, Y. Mely, J. L. Darlix, M. C. Fournie-Zaluski, S. Saragosti, and et al.** 1994. 1H NMR structure and biological studies of the His23-->Cys mutant nucleocapsid protein of HIV-1 indicate that the conformation of the first zinc finger is critical for virus infectivity. Biochemistry **33:**11707-16.

32. **Derebail, S. S., M. J. Heath, and J. J. DeStefano.** 2003. Evidence for the differential effects of nucleocapsid protein on strand transfer in various regions of the HIV genome. J. Biol. Chem. **278:**15702-12.

33. **DeStefano, J. J.** 1995. Human immunodeficiency virus nucleocapsid protein stimulates strand transfer from internal regions of heteropolymeric RNA templates. Arch. Virol. **140:**1775-89.

34. **DeStefano, J. J.** 1994. Kinetic analysis of the catalysis of strand transfer from internal regions of heteropolymeric RNA templates by human immunodeficiency virus reverse transcriptase. J. Mol. Biol. **243:**558-67.

35. **DeStefano, J. J., R. A. Bambara, and P. J. Fay.** 1994. The mechanism of human immunodeficiency virus reverse transcriptase-catalyzed strand transfer from internal regions of heteropolymeric RNA templates. J. Biol. Chem. **269:**161-8.

36. **DeStefano, J. J., L. M. Mallaber, P. J. Fay, and R. A. Bambara.** 1993. Determinants of the RNase H cleavage specificity of human immunodeficiency virus reverse transcriptase. Nucleic Acids Res. **21:**4330-8.

37. **DeStefano, J. J., L. M. Mallaber, L. Rodriguez-Rodriguez, P. J. Fay, and R. A. Bambara.** 1992. Requirements for strand transfer between internal regions of heteropolymer templates by human immunodeficiency virus reverse transcriptase. J. Virol. **66:**6370-8.

38. **Diaz, L., J. V. Cristofaro, and J. J. DeStefano.** 2000. Human immunodeficiency virus reverse transcriptase base misincorporations can promote strand transfer. Arch. Virol. **145:**1117-31.

39. **Diaz, L., and J. J. DeStefano.** 1996. Strand transfer is enhanced by mismatched nucleotides at the 3' primer terminus: a possible link between HIV reverse transcriptase fidelity and recombination. Nucleic Acids Res. **24:**3086-92.

40. **Diaz, R. S., E. C. Sabino, A. Mayer, J. W. Mosley, and M. P. Busch.** 1995. Dual human immunodeficiency virus type 1 infection and recombination in a dually exposed transfusion recipient. The Transfusion Safety Study Group. J. Virol. **69:**3273-81.

41. **Dib-Hajj, F., R. Khan, and D. P. Giedroc.** 1993. Retroviral nucleocapsid proteins possess potent nucleic acid strand renaturation activity. Protein Sci. **2:**231-43.

42. **Driscoll, M. D., M. P. Golinelli, and S. H. Hughes.** 2001. In vitro analysis of human immunodeficiency virus type 1 minus-strand strong-stop DNA synthesis and genomic RNA processing. J. Virol. **75:**672-86.

43.   **Driscoll, M. D., and S. H. Hughes.** 2000. Human immunodeficiency virus type 1 nucleocapsid protein can prevent self-priming of minus-strand strong stop DNA by promoting the annealing of short oligonucleotides to hairpin sequences. J. Virol. **74:**8785-92.

44.   **Druillennec, S., C. Z. Dong, S. Escaich, N. Gresh, A. Bousseau, B. P. Roques, and M. C. Fournie-Zaluski.** 1999. A mimic of HIV-1 nucleocapsid protein impairs reverse transcription and displays antiviral activity. Proc. Natl. Acad. Sci. USA **96:**4886-91.

45.   **Feng, Y. X., S. Campbell, D. Harvin, B. Ehresmann, C. Ehresmann, and A. Rein.** 1999. The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. J. Virol. **73:**4251-6.

46.   **Feng, Y. X., T. D. Copeland, L. E. Henderson, R. J. Gorelick, W. J. Bosche, J. G. Levin, and A. Rein.** 1996. HIV-1 nucleocapsid protein induces "maturation" of dimeric retroviral RNA in vitro. Proc. Natl. Acad. Sci. USA **93:**7577-81.

47.   **Fisher, R. J., A. Rein, M. Fivash, M. A. Urbaneja, J. R. Casas-Finet, M. Medaglia, and L. E. Henderson.** 1998. Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides. J. Virol. **72:**1902-9.

48.   **Fitzgerald, D. W., and J. E. Coleman.** 1991. Physicochemical properties of cloned nucleocapsid protein from HIV. Interactions with metal ions. Biochemistry **30:**5195-201.

49.   **Flint, S. J., L. W. Enquist, R. M. Krug, V. R. RAcaniello, and A. M. Skalka.** 2000. Principles of Virology. American Society for Microbiology, Washington, DC.

50.   **Gabbara, S., W. R. Davis, L. Hupe, D. Hupe, and J. A. Peliska.** 1999. Inhibitors of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. Biochemistry **38:**13070-6.

51.   **Gallay, P., S. Swingler, J. Song, F. Bushman, and D. Trono.** 1995. HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. Cell **83:**569-76.

52.   **Gao, F., D. L. Robertson, S. G. Morrison, H. Hui, S. Craig, J. Decker, P. N. Fultz, M. Girard, G. M. Shaw, B. H. Hahn, and P. M. Sharp.** 1996. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. J. Virol. **70:**7013-29.

53. **Gao, K., R. J. Gorelick, D. G. Johnson, and F. Bushman.** 2003. Cofactors for human immunodeficiency virus type 1 cDNA integration in vitro. J. Virol. **77:**1598-603.

54. **Gheysen, D., E. Jacobs, F. de Foresta, C. Thiriart, M. Francotte, D. Thines, and M. De Wilde.** 1989. Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. Cell **59:**103-12.

55. **Golinelli, M. P., and S. H. Hughes.** 2003. Secondary structure in the nucleic acid affects the rate of HIV-1 nucleocapsid-mediated strand annealing. Biochemistry **42:**8153-62.

56. **Golinelli, M. P., and S. H. Hughes.** 2001. Self-priming of retroviral minus-strand strong-stop DNAs. Virology **285:**278-90.

57. **Gorelick, R. J., D. J. Chabot, D. E. Ott, T. D. Gagliardi, A. Rein, L. E. Henderson, and L. O. Arthur.** 1996. Genetic analysis of the zinc finger in the Moloney murine leukemia virus nucleocapsid domain: replacement of zinc-coordinating residues with other zinc-coordinating residues yields noninfectious particles containing genomic RNA. J. Virol. **70:**2593-7.

58. **Gorelick, R. J., D. J. Chabot, A. Rein, L. E. Henderson, and L. O. Arthur.** 1993. The two zinc fingers in the human immunodeficiency virus type 1 nucleocapsid protein are not functionally equivalent. J. Virol. **67:**4027-36.

59. **Gorelick, R. J., T. D. Gagliardi, W. J. Bosche, T. A. Wiltrout, L. V. Coren, D. J. Chabot, J. D. Lifson, L. E. Henderson, and L. O. Arthur.** 1999. Strict conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes: characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences. Virology **256:**92-104.

60. **Gregoire, C. J., D. Gautheret, and E. P. Loret.** 1997. No tRNA3Lys unwinding in a complex with HIV NCp7. J. Biol. Chem. **272:**25143-8.

61. **Guo, J., L. E. Henderson, J. Bess, B. Kane, and J. G. Levin.** 1997. Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA. J. Virol. **71:**5178-88.

62. **Guo, J., T. Wu, J. Anderson, B. F. Kane, D. G. Johnson, R. J. Gorelick, L. E. Henderson, and J. G. Levin.** 2000. Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. J. Virol. **74:**8980-8.

63.   **Guo, J., T. Wu, J. Bess, L. E. Henderson, and J. G. Levin.** 1998. Actinomycin D inhibits human immunodeficiency virus type 1 minus-strand transfer in in vitro and endogenous reverse transcriptase assays. J. Virol. **72:**6716-24.

64.   **Guo, J., T. Wu, B. F. Kane, D. G. Johnson, L. E. Henderson, R. J. Gorelick, and J. G. Levin.** 2002. Subtle alterations of the native zinc finger structures have dramatic effects on the nucleic acid chaperone activity of human immunodeficiency virus type 1 nucleocapsid protein. J. Virol. **76:**4370-8.

65.   **Hargittai, M. R., A. T. Mangla, R. J. Gorelick, and K. Musier-Forsyth.** 2001. HIV-1 nucleocapsid protein zinc finger structures induce tRNA(Lys,3) structural changes but are not critical for primer/template annealing. J. Mol. Biol. **312:**985-97.

66.   **Heath, M. J., S. S. Derebail, R. J. Gorelick, and J. J. DeStefano.** 2003. Differing roles of the N- and C-terminal zinc fingers in human immunodeficiency virus nucleocapsid protein-enhanced nucleic acid annealing. J. Biol. Chem. **278:**30755-63.

67.   **Herschlag, D.** 1995. RNA chaperones and the RNA folding problem. J. Biol. Chem **270:**20871-4.

68.   **Hong, M. K., E. J. Harbron, D. B. O'Connor, J. Guo, P. F. Barbara, J. G. Levin, and K. Musier-Forsyth.** 2003. Nucleic acid conformational changes essential for HIV-1 nucleocapsid protein-mediated inhibition of self-priming in minus-strand transfer. J. Mol. Biol. **325:**1-10.

69.   **Hsu, M., L. Rong, H. de Rocquigny, B. P. Roques, and M. A. Wainberg.** 2000. The effect of mutations in the HIV-1 nucleocapsid protein on strand transfer in cell-free reverse transcription reactions. Nucleic Acids Res. **28:**1724-9.

70.   **Hu, W. S., and H. M. Temin.** 1990. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. Proc. Natl. Acad. Sci. USA **87:**1556-60.

71.   **Huang, M., A. Maynard, J. A. Turpin, L. Graham, G. M. Janini, D. G. Covell, and W. G. Rice.** 1998. Anti-HIV agents that selectively target retroviral nucleocapsid protein zinc fingers without affecting cellular zinc finger proteins. J. Med. Chem. **41:**1371-81.

72.   **Jetzt, A. E., H. Yu, G. J. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty.** 2000. High rate of recombination throughout the human immunodeficiency virus type 1 genome. J. Virol. **74:**1234-40.

73.	**Ji, X., G. J. Klarmann, and B. D. Preston.** 1996. Effect of human immunodeficiency virus type 1 (HIV-1) nucleocapsid protein on HIV-1 reverse transcriptase activity in vitro. Biochemistry **35:**132-43.

74.	**Johnson, P. E., R. B. Turner, Z. R. Wu, L. Hairston, J. Guo, J. G. Levin, and M. F. Summers.** 2000. A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription. Biochemistry **39:**9084-91.

75.	**Jones, J. S., R. W. Allan, and H. M. Temin.** 1994. One retroviral RNA is sufficient for synthesis of viral DNA. J. Virol. **68:**207-16.

76.	**Kao, S., M. A. Khan, E. Miyagi, R. Plishka, A. Buckler-White, and K. Strebel.** 2003. The human immunodeficiency virus type 1 Vif protein reduces intracellular expression and inhibits packaging of APOBEC3G (CEM15), a cellular inhibitor of virus infectivity. J. Virol. **77:**11398-407.

77.	**Kellam, P., and B. A. Larder.** 1995. Retroviral recombination can lead to linkage of reverse transcriptase mutations that confer increased zidovudine resistance. J. Virol. **69:**669-74.

78.	**Khan, R., and D. P. Giedroc.** 1994. Nucleic acid binding properties of recombinant Zn2 HIV-1 nucleocapsid protein are modulated by COOH-terminal processing. J. Biol. Chem. **269:**22538-46.

79.	**Khan, R., and D. P. Giedroc.** 1992. Recombinant human immunodeficiency virus type 1 nucleocapsid (NCp7) protein unwinds tRNA. J. Biol. Chem. **267:**6689-95.

80.	**Kim, J. K., C. Palaniappan, W. Wu, P. J. Fay, and R. A. Bambara.** 1997. Evidence for a unique mechanism of strand transfer from the transactivation response region of HIV-1. J. Biol. Chem. **272:**16769-77.

81.	**Klasens, B. I., H. T. Huthoff, A. T. Das, R. E. Jeeninga, and B. Berkhout.** 1999. The effect of template RNA structure on elongation by HIV-1 reverse transcriptase. Biochim. Biophys. Acta. **1444:**355-70.

82.	**Lapadat-Tapolsky, M., H. De Rocquigny, D. Van Gent, B. Roques, R. Plasterk, and J. L. Darlix.** 1993. Interactions between HIV-1 nucleocapsid protein and viral DNA may have important functions in the viral life cycle. Nucleic Acids Res. **21:**831-9.

83.	**Lapadat-Tapolsky, M., C. Pernelle, C. Borie, and J. L. Darlix.** 1995. Analysis of the nucleic acid annealing activities of nucleocapsid protein from HIV-1. Nucleic Acids Res. **23:**2434-41.

146

84.   **Le Novere, N.** 2001. MELTING, computing the melting temperature of nucleic acid duplex. Bioinformatics **17:**1226-7.

85.   **Levy, J. P.** 1998. AIDS vaccine development. Science **280:**806-7.

86.   **Lochrie, M. A., S. Waugh, D. G. Pratt, Jr., J. Clever, T. G. Parslow, and B. Polisky.** 1997. In vitro selection of RNAs that bind to the human immunodeficiency virus type-1 gag polyprotein. Nucleic Acids Res. **25:**2902-10.

87.   **Lorsch, J. R.** 2002. RNA chaperones exist and DEAD box proteins get a life. Cell **109:**797-800.

88.   **Marin, M., K. M. Rose, S. L. Kozak, and D. Kabat.** 2003. HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. Nat. Med. **9:**1398-403.

89.   **Matzura, O., and A. Wennborg.** 1996. RNAdraw: an integrated program for RNA secondary structure calculation and analysis under 32-bit Microsoft Windows. Comput. Appl. Biosci. **12:**247-9.

90.   **McGrath, C. F., J. S. Buckman, T. D. Gagliardi, W. J. Bosche, L. V. Coren, and R. J. Gorelick.** 2003. Human cellular nucleic acid-binding protein Zn2+ fingers support replication of human immunodeficiency virus type 1 when they are substituted in the nucleocapsid protein. J. Virol. **77:**8524-31.

91.   **Mehle, A., B. Strack, P. Ancuta, C. Zhang, M. McPike, and D. Gabuzda.** 2004. Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway. J. Biol. Chem. **279:**7792-8.

92.   **Morellet, N., H. de Rocquigny, Y. Mely, N. Jullian, H. Demene, M. Ottmann, D. Gerard, J. L. Darlix, M. C. Fournie-Zaluski, and B. P. Roques.** 1994. Conformational behaviour of the active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by 1H NMR. J. Mol. Biol. **235:**287-301.

93.   **Morellet, N., N. Jullian, H. De Rocquigny, B. Maigret, J. L. Darlix, and B. P. Roques.** 1992. Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by 1H NMR. EMBO J. **11:**3059-65.

94.   **Moumen, A., L. Polomack, B. Roques, H. Buc, and M. Negroni.** 2001. The HIV-1 repeated sequence R as a robust hot-spot for copy-choice recombination. Nucleic Acids Res. **29:**3814-21.

95.   **Moumen, A., L. Polomack, T. Unge, M. Veron, H. Buc, and M. Negroni.** 2003. Evidence for a mechanism of recombination during reverse transcription dependent on the structure of the acceptor RNA. J. Biol. Chem. **278:**15973-82.

96. **Muller, G., B. Strack, J. Dannull, B. S. Sproat, A. Surovoy, G. Jung, and K. Moelling.** 1994. Amino acid requirements of the nucleocapsid protein of HIV-1 for increasing catalytic activity of a Ki-ras ribozyme in vitro. J. Mol. Biol. **242:**422-9.

97. **Muriaux, D., H. De Rocquigny, B. P. Roques, and J. Paoletti.** 1996. NCp7 activates HIV-1Lai RNA dimerization by converting a transient loop-loop complex into a stable dimer. J. Biol. Chem. **271:**33686-92.

98. **Negroni, M., and H. Buc.** 2000. Copy-choice recombination by reverse transcriptases: reshuffling of genetic markers mediated by RNA chaperones. Proc. Natl. Acad. Sci. USA **97:**6385-90.

99. **Negroni, M., and H. Buc.** 1999. Recombination during reverse transcription: an evaluation of the role of the nucleocapsid protein. J. Mol. Biol. **286:**15-31.

100. **Onafuwa, A., W. An, N. D. Robson, and A. Telesnitsky.** 2003. Human immunodeficiency virus type 1 genetic recombination is more frequent than that of Moloney murine leukemia virus despite similar template switching rates. J. Virol. **77:**4577-87.

101. **Ottmann, M., C. Gabus, and J. L. Darlix.** 1995. The central globular domain of the nucleocapsid protein of human immunodeficiency virus type 1 is critical for virion structure and infectivity. J. Virol. **69:**1778-84.

102. **Peliska, J. A., S. Balasubramanian, D. P. Giedroc, and S. J. Benkovic.** 1994. Recombinant HIV-1 nucleocapsid protein accelerates HIV-1 reverse transcriptase catalyzed DNA strand transfer reactions and modulates RNase H activity. Biochemistry **33:**13817-23.

103. **Peliska, J. A., and S. J. Benkovic.** 1992. Mechanism of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. Science **258:**1112-8.

104. **Poon, D. T., J. Wu, and A. Aldovini.** 1996. Charged amino acid residues of human immunodeficiency virus type 1 nucleocapsid p7 protein involved in RNA packaging and infectivity. J. Virol. **70:**6607-16.

105. **Prats, A. C., V. Housset, G. de Billy, F. Cornille, H. Prats, B. Roques, and J. L. Darlix.** 1991. Viral RNA annealing activities of the nucleocapsid protein of Moloney murine leukemia virus are zinc independent. Nucleic Acids Res. **19:**3533-41.

106. **Priel, E., E. Aflalo, I. Seri, L. E. Henderson, L. O. Arthur, M. Aboud, S. Segal, and D. G. Blair.** 1995. DNA binding properties of the zinc-bound and zinc-free HIV nucleocapsid protein: supercoiled DNA unwinding and DNA-protein cleavable complex formation. FEBS Lett. **362:**59-64.

107. **Quinones-Mateu, M. E., Y. Gao, S. C. Ball, A. J. Marozsan, A. Abraha, and E. J. Arts.** 2002. In vitro intersubtype recombinants of human immunodeficiency virus type 1: comparison to recent and circulating in vivo recombinant forms. J. Virol. **76:**9600-13.

108. **Raja, A., and J. J. DeStefano.** 1999. Kinetic analysis of the effect of HIV nucleocapsid protein (NCp) on internal strand transfer reactions. Biochemistry **38:**5178-84.

109. **Rein, A.** 2003. Personal Communication.

110. **Rein, A., L. E. Henderson, and J. G. Levin.** 1998. Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. Trends Biochem. Sci. **23:**297-301.

111. **Remy, E., H. de Rocquigny, P. Petitjean, D. Muriaux, V. Theilleux, J. Paoletti, and B. P. Roques.** 1998. The annealing of tRNA3Lys to human immunodeficiency virus type 1 primer binding site is critically dependent on the NCp7 zinc fingers structure. J. Biol. Chem. **273:**4819-22.

112. **Robertson, D. L., P. M. Sharp, F. E. McCutchan, and B. H. Hahn.** 1995. Recombination in HIV-1. Nature **374:**124-6.

113. **Roda, R. H., M. Balakrishnan, J. K. Kim, B. P. Roques, P. J. Fay, and R. A. Bambara.** 2002. Strand transfer occurs in retroviruses by a pause-initiated two-step mechanism. J. Biol. Chem. **277:**46900-11.

114. **Roda, R. H., M. Balakrishnan, J. K. Kim, B. P. Roques, P. J. Fay, and R. A. Bambara.** 2002. Strand Transfer Occurs in Retroviruses by a Pause-initiated Two-step Mechanism. J. Biol. Chem. **277:**46900-46911.

115. **Rong, L., C. Liang, M. Hsu, X. Guo, B. P. Roques, and M. A. Wainberg.** 2001. HIV-1 nucleocapsid protein and the secondary structure of the binary complex formed between tRNA(Lys.3) and viral RNA template play different roles during initiation of (-) strand DNA reverse transcription. J. Biol. Chem. **276:**47725-32.

116. **Rutherford, S. L.** 2003. Between genotype and phenotype: protein chaperones and evolvability. Nat. Rev. Genet. **4:**263-74.

117. **Sabino, E. C., E. G. Shpaer, M. G. Morgado, B. T. Korber, R. S. Diaz, V. Bongertz, S. Cavalcante, B. Galvao-Castro, J. I. Mullins, and A. Mayer.** 1994. Identification of human immunodeficiency virus type 1 envelope genes recombinant between subtypes B and F in two epidemiologically linked individuals from Brazil. J. Virol. **68:**6340-6.

149

118. **Sambrook, J., D. W. Russell, N. Irwin, and K. A. Janssen.** 2001. Molecular Cloning A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

119. **Schmalzbauer, E., B. Strack, J. Dannull, S. Guehmann, and K. Moelling.** 1996. Mutations of basic amino acids of NCp7 of human immunodeficiency virus type 1 affect RNA binding in vitro. J. Virol. **70:**771-7.

120. **Sharp, P. M., E. Bailes, M. Stevenson, M. Emerman, and B. H. Hahn.** 1996. Gene acquisition in HIV and SIV. Nature **383:**586-7.

121. **Shehu-Xhilaga, M., S. M. Crowe, and J. Mak.** 2001. Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. J. Virol. **75:**1834-41.

122. **Shehu-Xhilaga, M., H. G. Kraeusslich, S. Pettit, R. Swanstrom, J. Y. Lee, J. A. Marshall, S. M. Crowe, and J. Mak.** 2001. Proteolytic processing of the p2/nucleocapsid cleavage site is critical for human immunodeficiency virus type 1 RNA dimer maturation. J. Virol. **75:**9156-64.

123. **South, T. L., P. R. Blake, R. C. Sowder, 3rd, L. O. Arthur, L. E. Henderson, and M. F. Summers.** 1990. The nucleocapsid protein isolated from HIV-1 particles binds zinc and forms retroviral-type zinc fingers. Biochemistry **29:**7786-9.

124. **South, T. L., and M. F. Summers.** 1993. Zinc- and sequence-dependent binding to nucleic acids by the N-terminal zinc finger of the HIV-1 nucleocapsid protein: NMR structure of the complex with the Psi-site analog, dACGCC. Protein Sci. **2:**3-19.

125. **Stine, G. J.** 2001. AIDS Update 2001. Prentice Hall, Upper Saddle River, New Jersey.

126. **Summers, M. F., L. E. Henderson, M. R. Chance, J. W. Bess, Jr., T. L. South, P. R. Blake, I. Sagi, G. Perez-Alvarado, R. C. Sowder, 3rd, D. R. Hare, and et al.** 1992. Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution-state structure of the nucleocapsid protein from HIV-1. Protein Sci. **1:**563-74.

127. **Summers, M. F., T. L. South, B. Kim, and D. R. Hare.** 1990. High-resolution structure of an HIV zinc fingerlike domain via a new NMR-based distance geometry approach. Biochemistry **29:**329-40.

128. **Surovoy, A., J. Dannull, K. Moelling, and G. Jung.** 1993. Conformational and nucleic acid binding studies on the synthetic nucleocapsid protein of HIV-1. J. Mol. Biol. **229:**94-104.

129. **Takahashi, K., S. Baba, Y. Koyanagi, N. Yamamoto, H. Takaku, and G. Kawai.** 2001. Two basic regions of NCp7 are sufficient for conformational conversion of HIV-1 dimerization initiation site from kissing-loop dimer to extended-duplex dimer. J. Biol. Chem. **276:**31274-8.

130. **Tanchou, V., D. Decimo, C. Pechoux, D. Lener, V. Rogemond, L. Berthoux, M. Ottmann, and J. L. Darlix.** 1998. Role of the N-terminal zinc finger of human immunodeficiency virus type 1 nucleocapsid protein in virus structure and replication. J. Virol. **72:**4442-7.

131. **Temin, H. M.** 1993. Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. Proc. Natl. Acad. Sci. USA **90:**6900-3.

132. **Trubetskoy, A. M., S. A. Okenquist, and J. Lenz.** 1999. R region sequences in the long terminal repeat of a murine retrovirus specifically increase expression of unspliced RNAs. J. Virol. **73:**3477-83.

133. **Tsuchihashi, Z., and P. O. Brown.** 1994. DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. J. Virol. **68:**5863-70.

134. **Tsuchihashi, Z., M. Khosla, and D. Herschlag.** 1993. Protein enhancement of hammerhead ribozyme catalysis. Science **262:**99-102.

135. **UNAIDS.** 2003. AIDS epidemic update. Joint United Nations Programme on HIV/AIDS, World Health Organization.

136. **Urbaneja, M. A., B. P. Kane, D. G. Johnson, R. J. Gorelick, L. E. Henderson, and J. R. Casas-Finet.** 1999. Binding properties of the human immunodeficiency virus type 1 nucleocapsid protein p7 to a model RNA: elucidation of the structural determinants for function. J. Mol. Biol. **287:**59-75.

137. **Urbaneja, M. A., M. Wu, J. R. Casas-Finet, and R. L. Karpel.** 2002. HIV-1 nucleocapsid protein as a nucleic acid chaperone: spectroscopic study of its helix-destabilizing properties, structural binding specificity, and annealing activity. J. Mol. Biol. **318:**749-64.

138. **Vuilleumier, C., E. Bombarda, N. Morellet, D. Gerard, B. P. Roques, and Y. Mely.** 1999. Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence study. Biochemistry **38:**16816-25.

139. **Williams, M. C., R. J. Gorelick, and K. Musier-Forsyth.** 2002. Specific zinc-finger architecture required for HIV-1 nucleocapsid protein's nucleic acid chaperone function. Proc. Natl. Acad. Sci. USA **99:**8614-9.

140. **Williams, M. C., I. Rouzina, J. R. Wenner, R. J. Gorelick, K. Musier-Forsyth, and V. A. Bloomfield.** 2001. Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching. Proc. Natl. Acad. Sci. USA **98:**6121-6.

141. **Wooley, D. P., R. A. Smith, S. Czajak, and R. C. Desrosiers.** 1997. Direct demonstration of retroviral recombination in a rhesus monkey. J. Virol. **71:**9650-3.

142. **Wu, T., J. Guo, J. Bess, L. E. Henderson, and J. G. Levin.** 1999. Molecular requirements for human immunodeficiency virus type 1 plus-strand transfer: analysis in reconstituted and endogenous reverse transcription systems. J. Virol. **73:**4794-805.

143. **Wu, W., B. M. Blumberg, P. J. Fay, and R. A. Bambara.** 1995. Strand transfer mediated by human immunodeficiency virus reverse transcriptase in vitro is promoted by pausing and results in misincorporation. J. Biol. Chem. **270:**325-32.

144. **Yin, P. D., V. K. Pathak, A. E. Rowan, R. J. Teufel, 2nd, and W. S. Hu.** 1997. Utilization of nonhomologous minus-strand DNA transfer to generate recombinant retroviruses. J. Virol. **71:**2487-94.

145. **You, J. C., and C. S. McHenry.** 1993. HIV nucleocapsid protein. Expression in Escherichia coli, purification, and characterization. J. Biol. Chem. **268:**16519-27.

146. **You, J. C., and C. S. McHenry.** 1994. Human immunodeficiency virus nucleocapsid protein accelerates strand transfer of the terminally redundant sequences involved in reverse transcription. J. Biol. Chem. **269:**31491-5.

147. **Yu, H., A. E. Jetzt, Y. Ron, B. D. Preston, and J. P. Dougherty.** 1998. The nature of human immunodeficiency virus type 1 strand transfers. J. Biol. Chem. **273:**28384-91.

148. **Yu, X., Y. Yu, B. Liu, K. Luo, W. Kong, P. Mao, and X. F. Yu.** 2003. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. Science **302:**1056-60.

149. **Zhang, J., and H. M. Temin.** 1993. Rate and mechanism of nonhomologous recombination during a single cycle of retroviral replication. Science **259:**234-8.

150. **Zuker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res **9:**133-48.