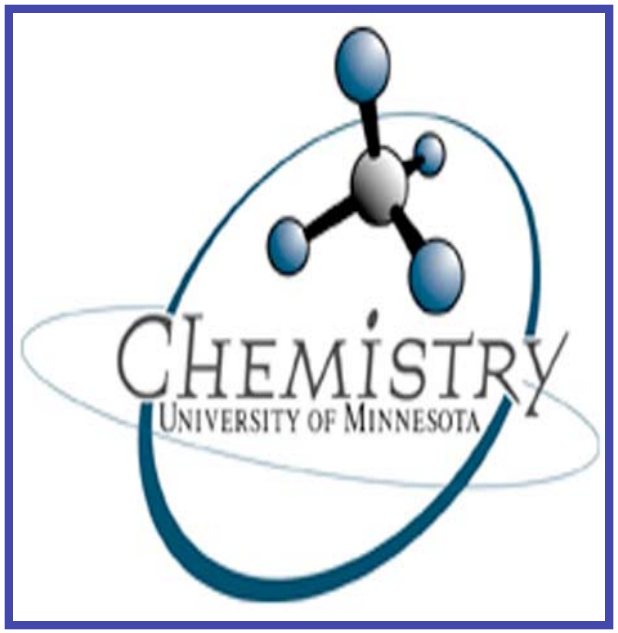
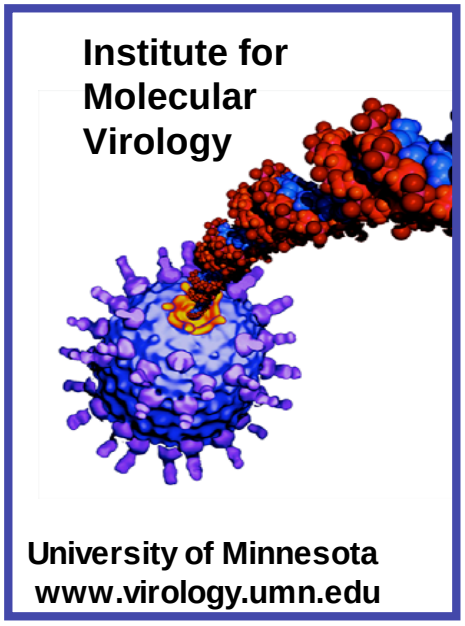
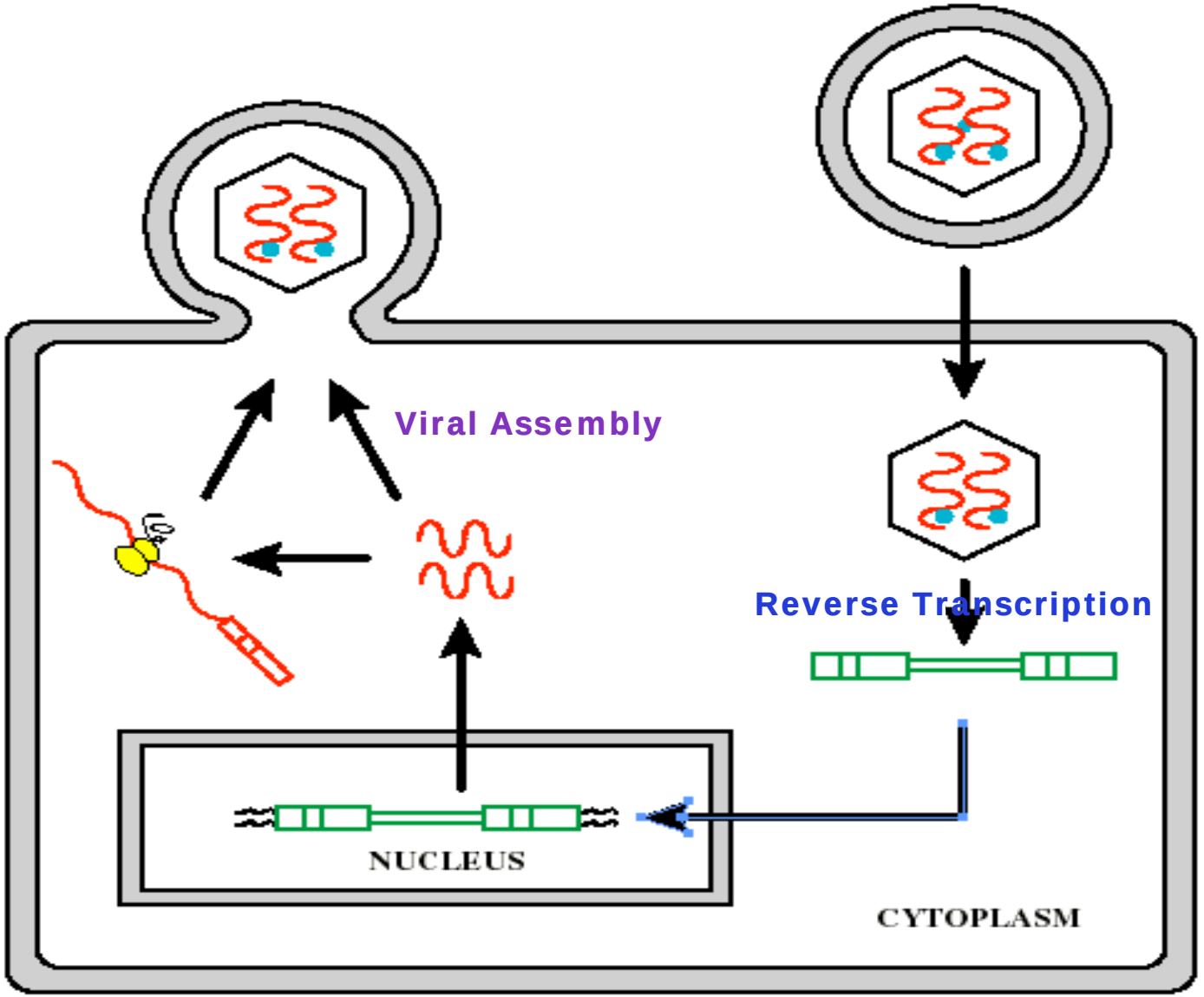
Mechanism of NC's nucleic acid chaperone activity in reverse transcription

Ioulia Rouzina

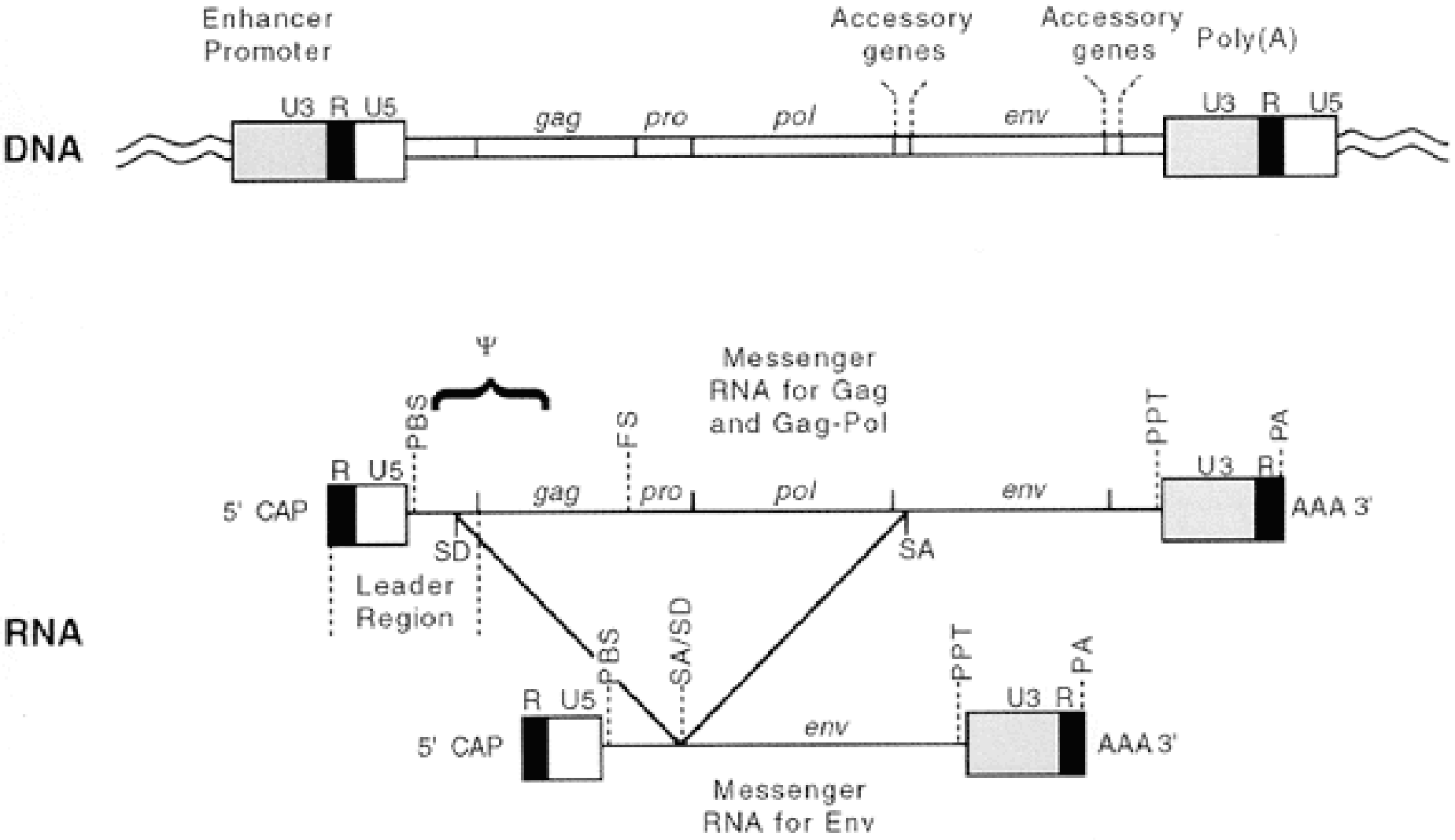




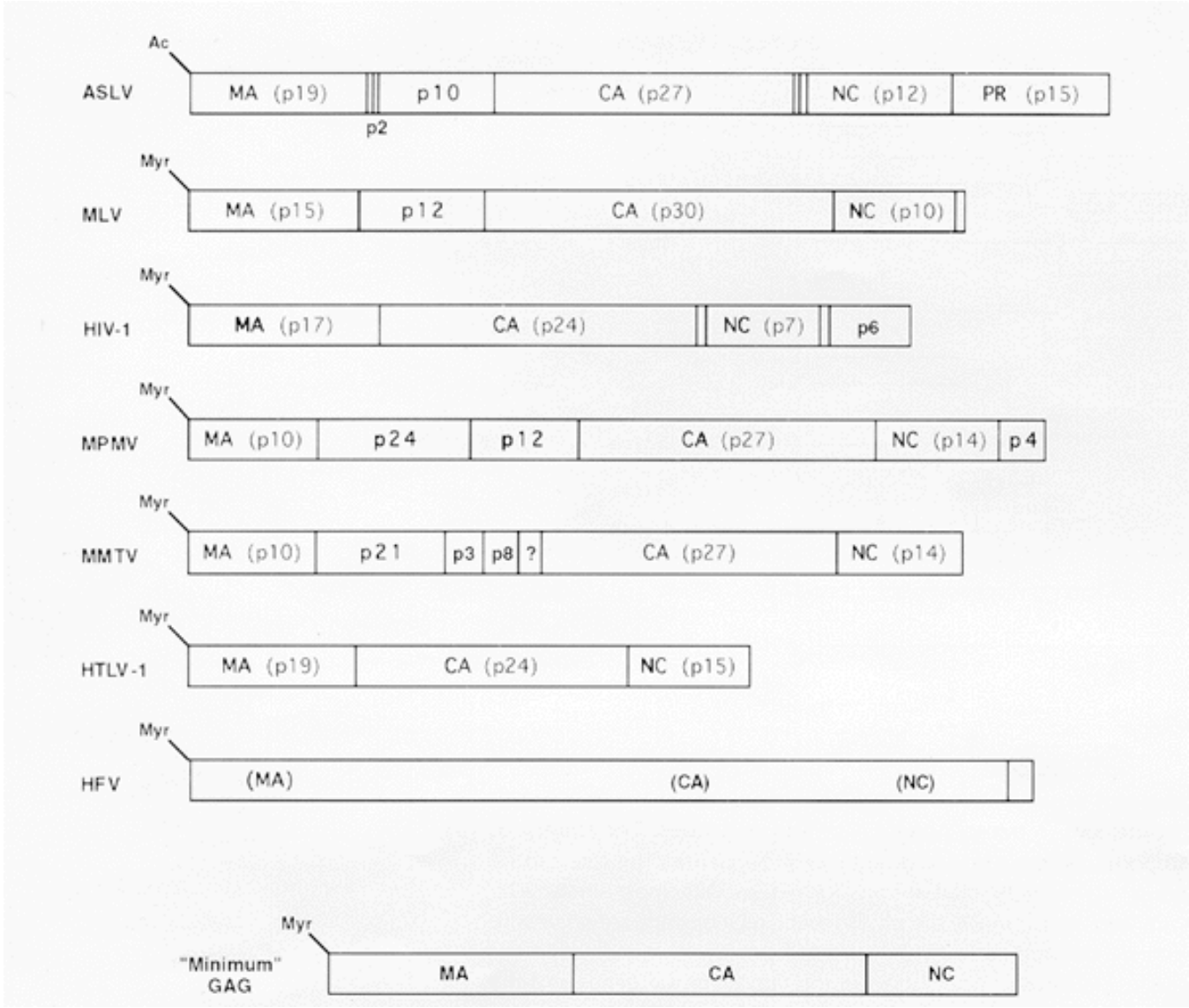
Retrovirus life cycle



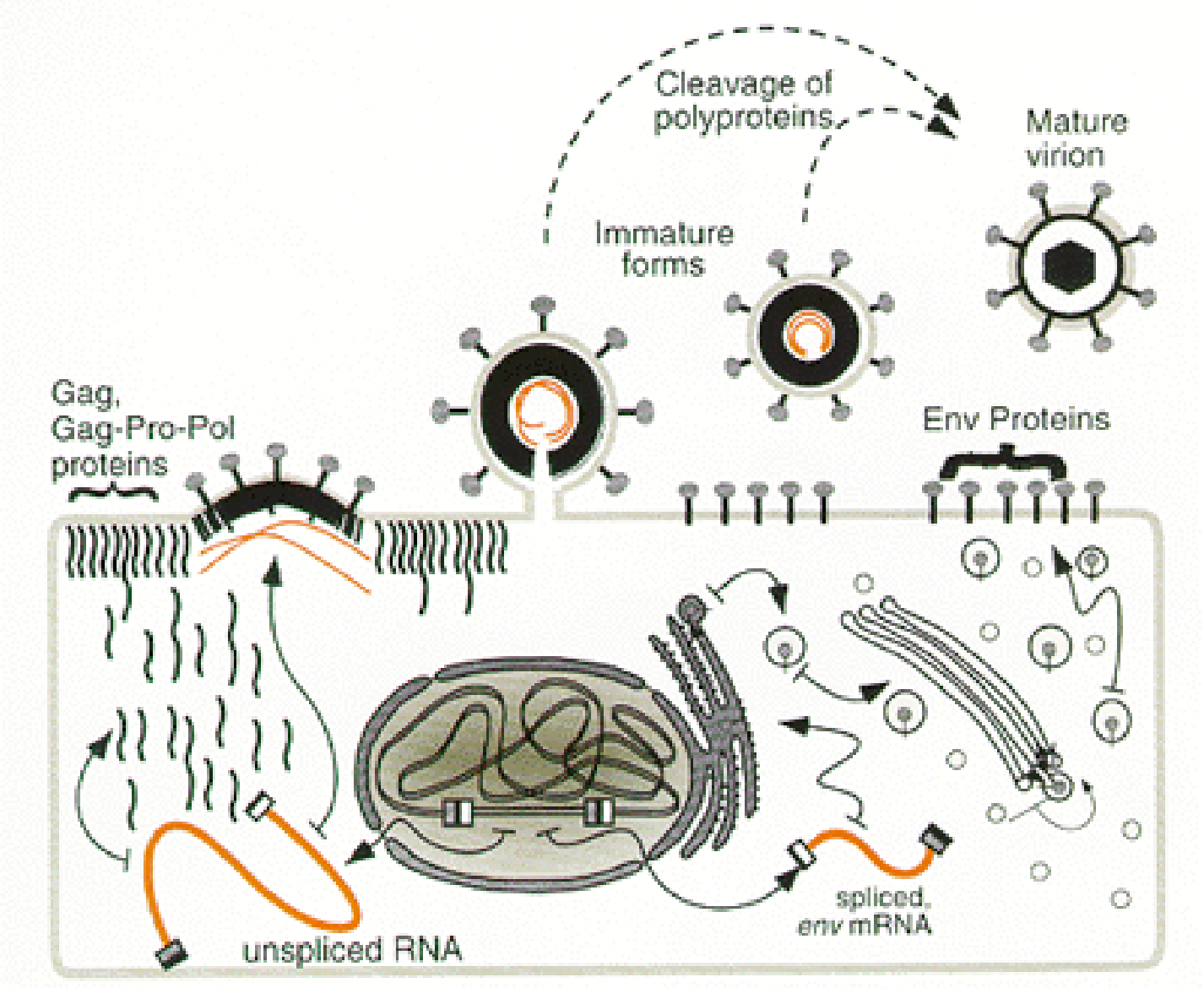
Genetic organization of generalized provirus.



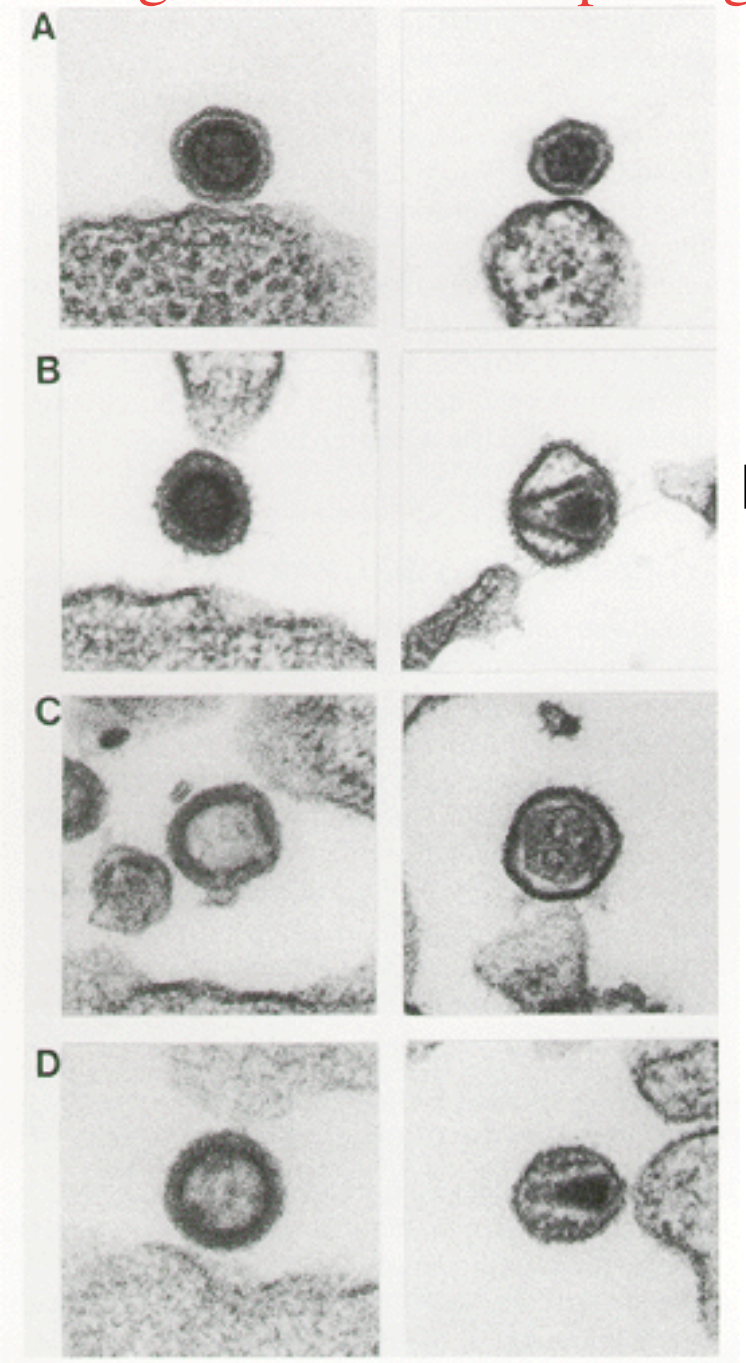
Organization of Gag proteins in retroviruses.



Overview of retroviral assembly showing the universal features.



Changes in virion morphology with maturation are virus specific.



MLV

M-PMV

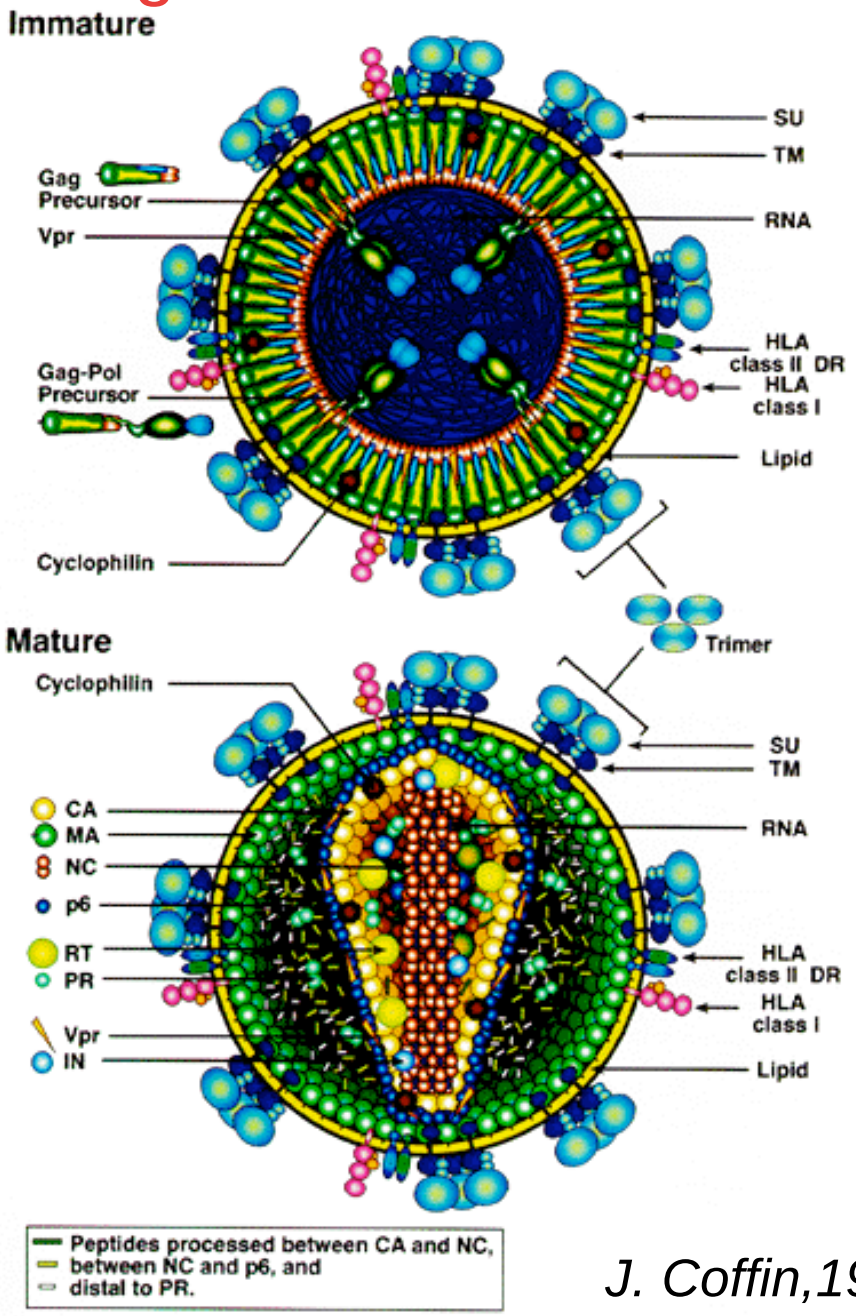
HTLV-1

HIV-1

CA protein structures of different viruses are very similar, but mature viral morphology is different

W.Sundquist, J.Briggs

Organization of immature and mature HIV-1 virion

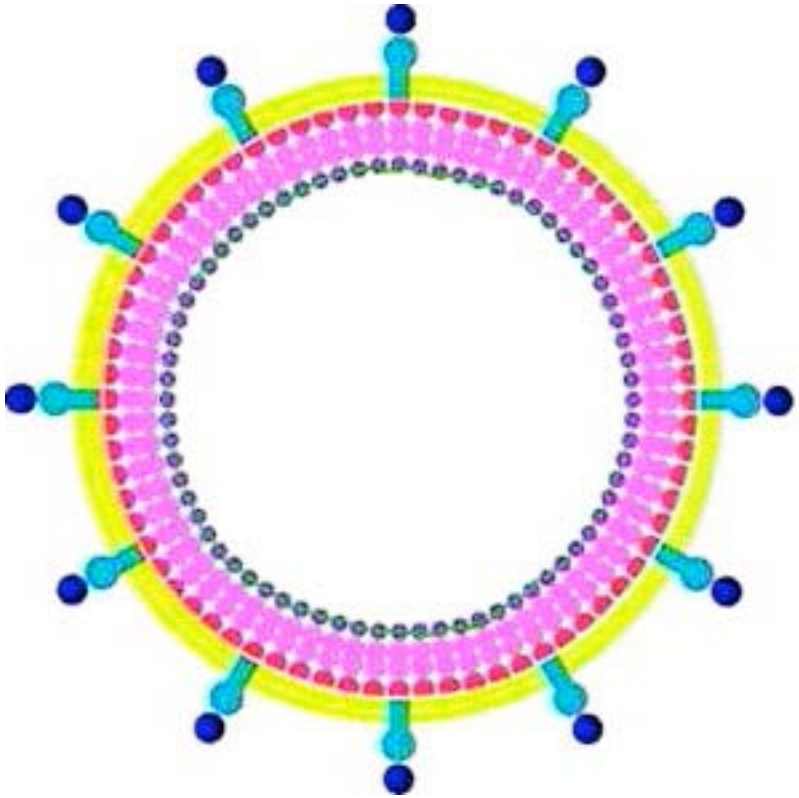


J. Coffin,1997

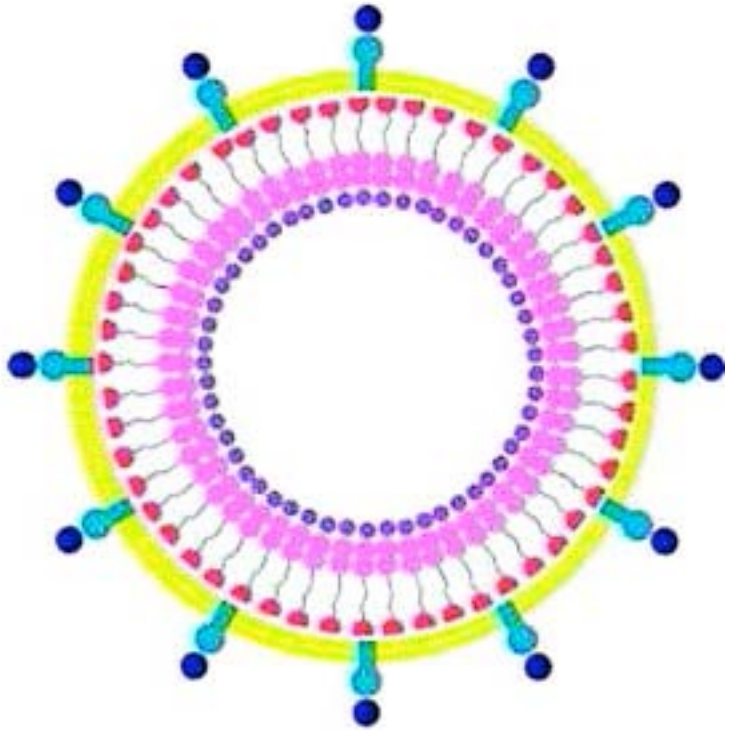
- 4000-5000 CA in spherical capsid ~120-200nm. Size induced by CA-CA curvature(?)
- Assembly is RNA-dependent (also assembles on DNA, ds,ss, etc)
- Stiochiometric amounts of nt packaged (1 CA/6 nt)
- High salt supports assembly w/o RNA
- Assmembles at membrane in vivo In vitro can be assembled without membraneis in the presence of Inositol(6-)
- Kinetics of assembly/CA-dependence never studied in vitro

- 1000-1500 CA in mature capsid; often Immature core has few mature cores;
- RNA/NC aggregated within the core
- Weak core stability
- Core none transparent to RT, IN, but transparent to NTP and ATP, to NC? ~250 RT, IN only

Immature HIV-1 and RSV virions



HIV-1

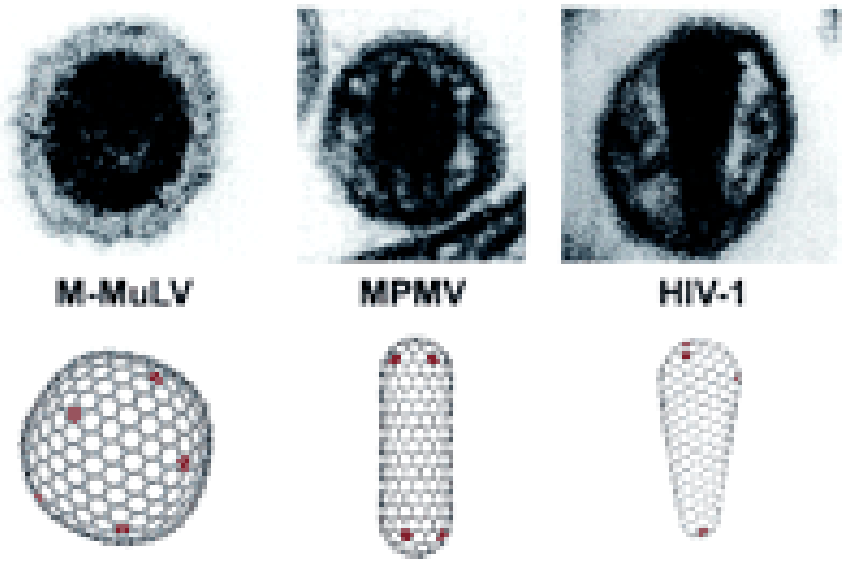


RSV

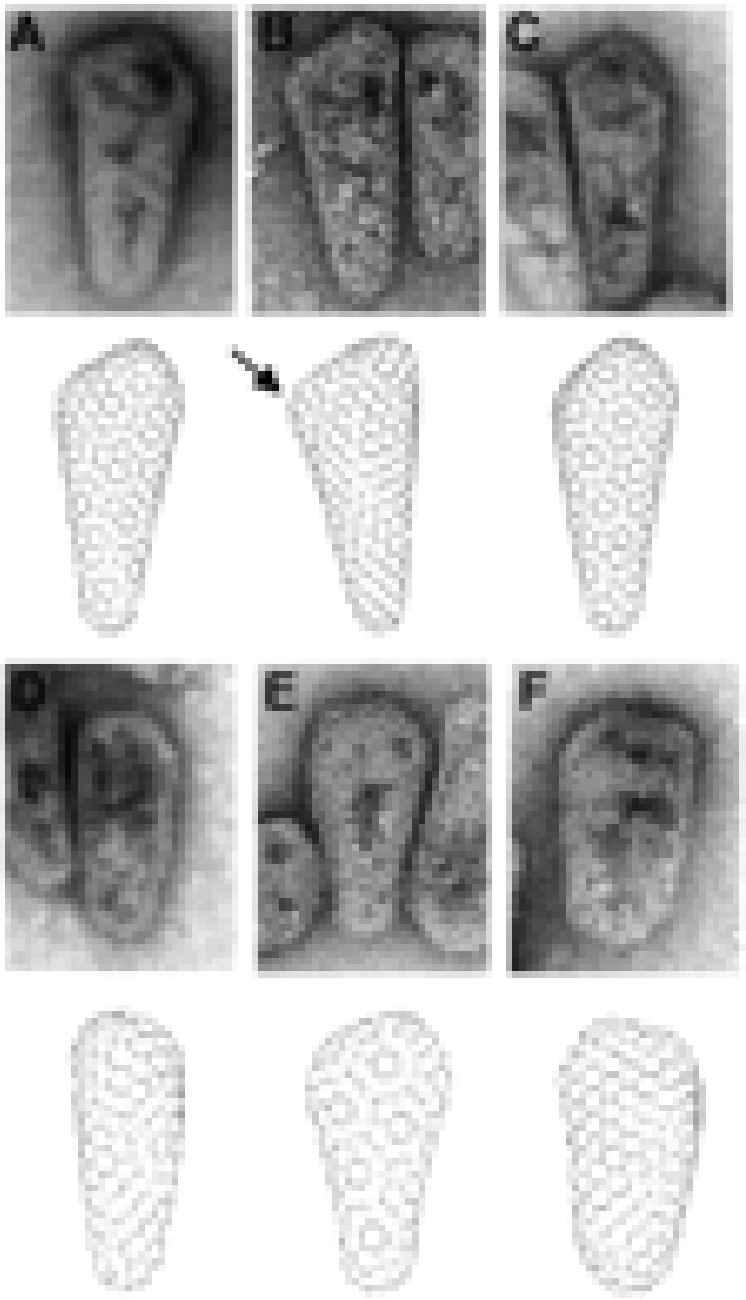
CA-CA interactions determine the curvature/size of virions

J. Briggs and V.Vogt 2005 JMB

Comparisons of HIV-1 capsids with idealized models.



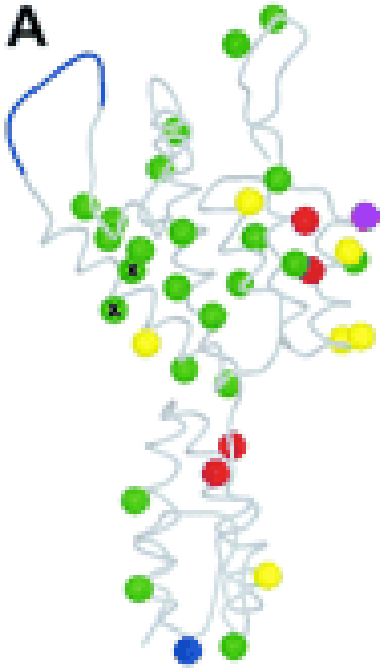
Sundquist, 2004



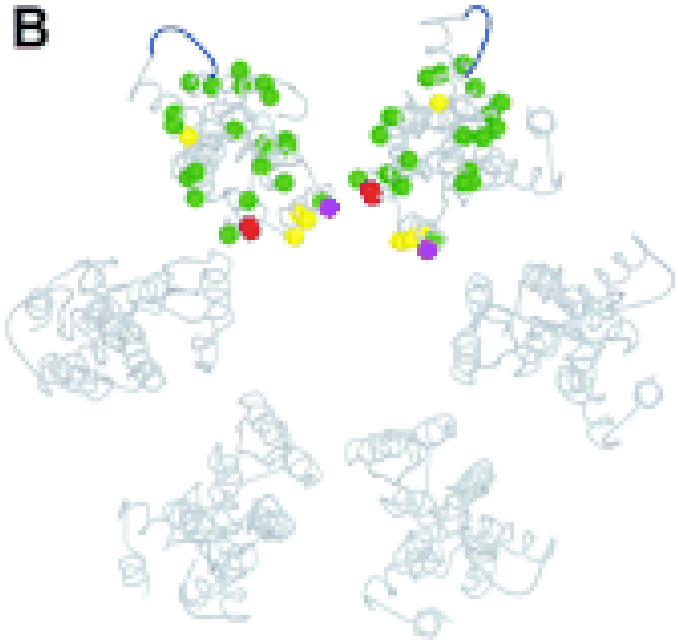
Upon gag cleavage capsid disintegrates And re-assembles; nucleation governed by RNA starts from narrow end and grows till it reaches outside membrane. Capsid formed CA of hexamers with 5/7 pentamers at the ends. ~20^0angle cone

Capsid -capsid interaction

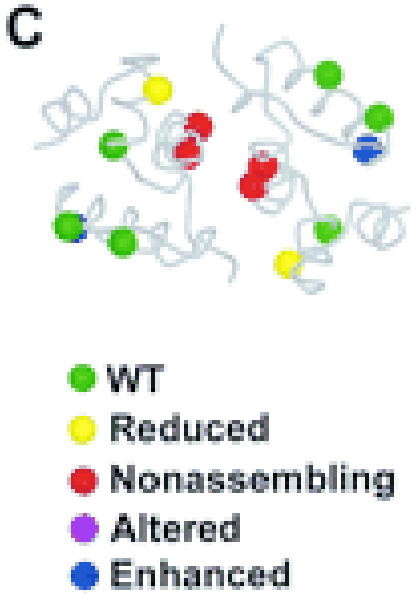
HIV-1 CA protein



model of the hexameric rings formed by the CA NTD



Crystallographic model of the dimer formed by the CA CTD



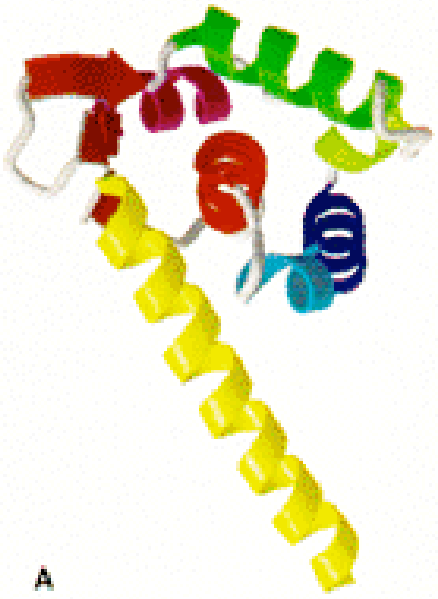
NTD mutations mostly affect mature virus structure

CTD mutations mostly affect immature virus structure

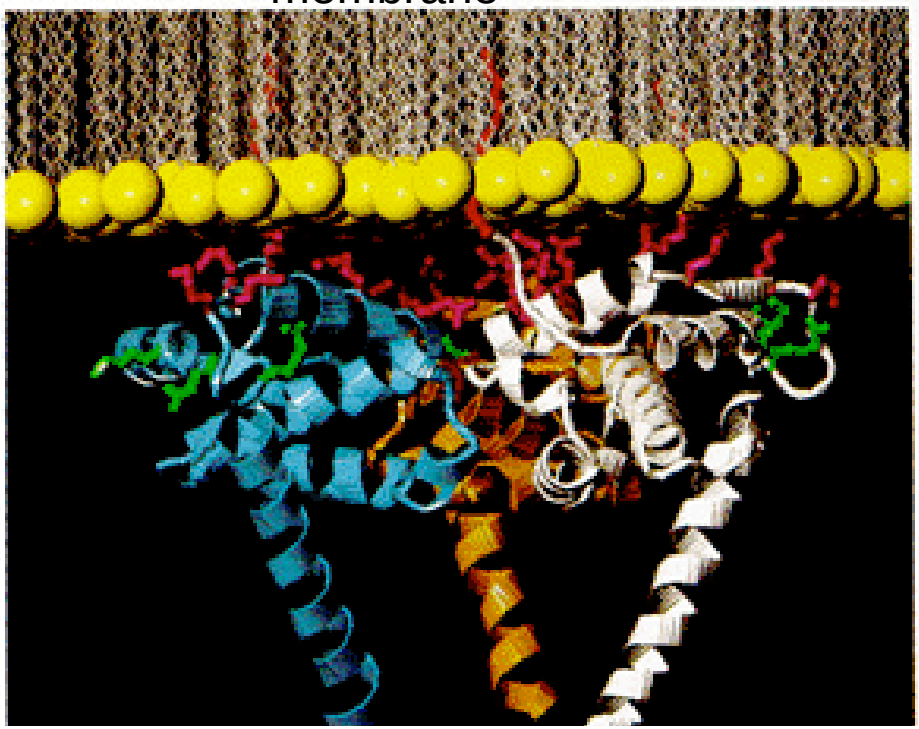
In high salt CA self assembles into cylinders made of CA hexamers
Alternative head-to-tail models of CA assembly still exist (20 Å resolution)

Three-dimensional structures of MA.

MA
monomer



MA trimer
associated with cell
membrane



Cationic matrix domains are myristilated,
Hydrocarbon tails intercalate membrane.

Open questions of viral self-assembly

Spontaneous or driven?

Kinetics of self assembly?

Gag dependence of self-assembly?

CA-CA interactions?

Solution ionic strength effects on assembly?

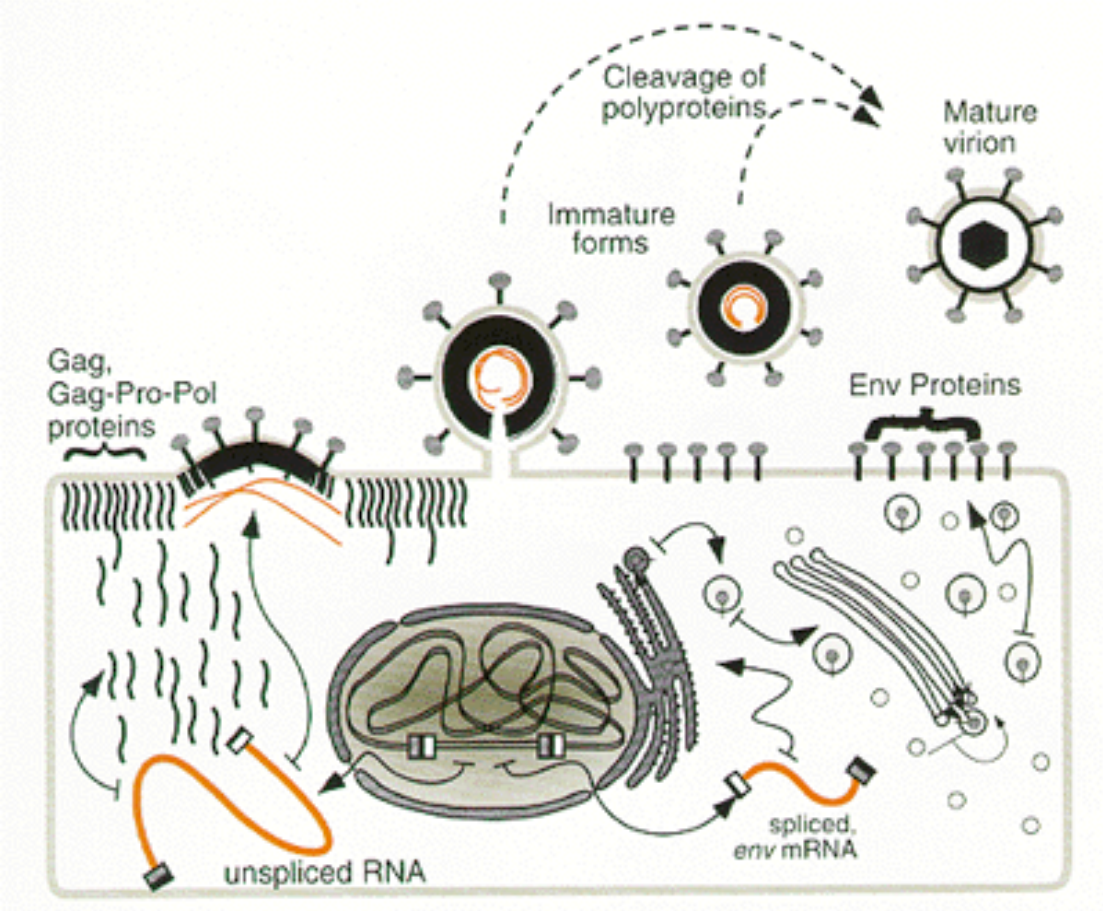
Exact NC/nt binding stiochiometry in packaging?

Problem of preferred viral RNA packaging is still unresolved

Kinetic hypothesis:
Newly synthesized gag packages its own viral RNA
Counter arguments:
•VLP assembled on random RNA will slowly exchange it for viral RNA
•Direct experiment, HU WS, 2006

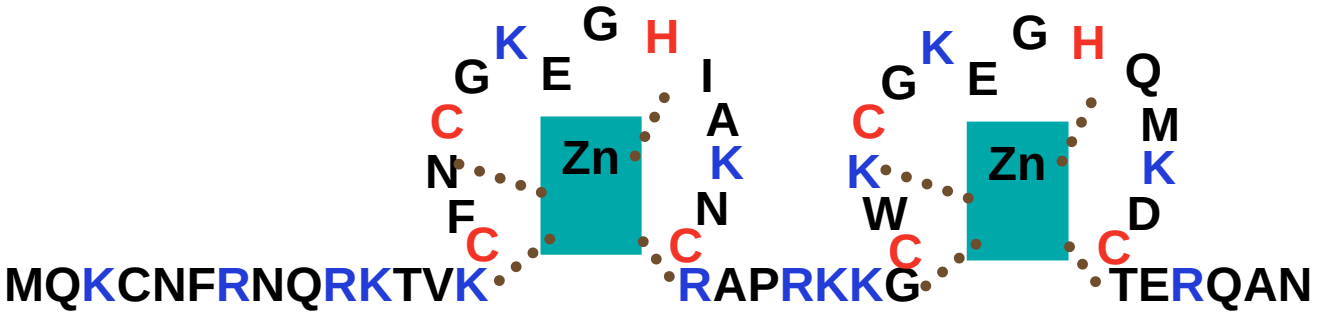
Probably equilibrium specificity

Packaging signal ψ is typically long 100-1000 nt
No strong binding sequence-specificity observed

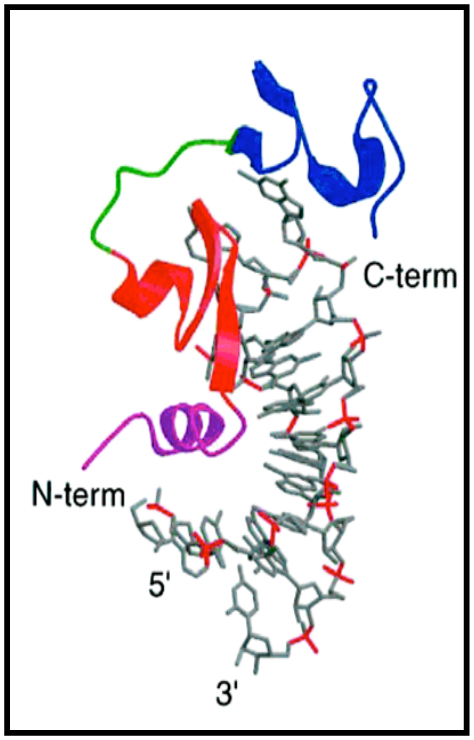


HIV-1 Nucleocapsid Protein (NC) is responsible For viral RNA packaging

15 basic residues (pI = 9.93)
2 nonequivalent CCHC Zn^{2+} "knuckles"



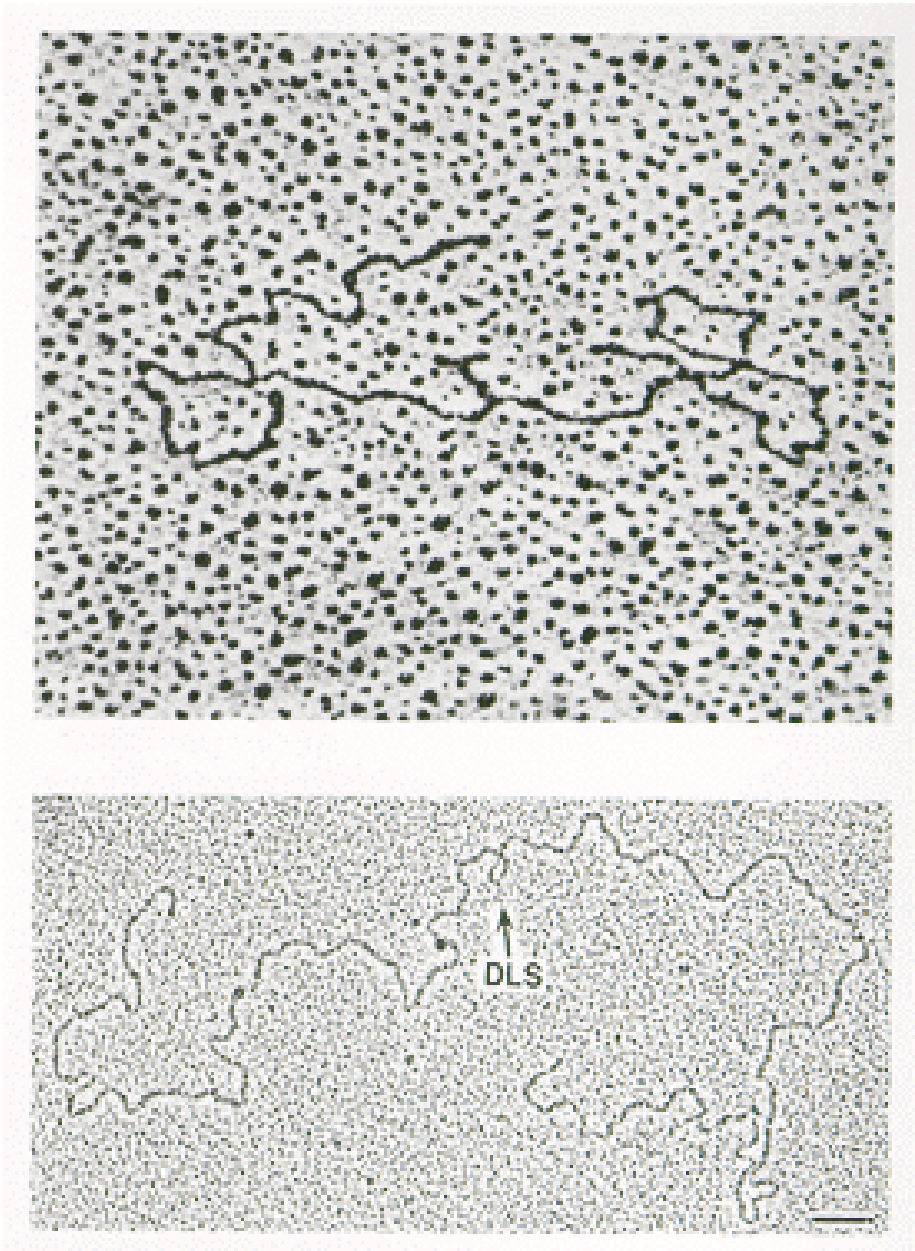
- Cases of packaging viral RNA by nonhomolohous gag with homologous NC are known
- NC binds nonspecifically all RNA/DNA;
- Has some preference for Gs;
- Not enough sequence-specificity to explain packaging;
- Viral RNA packaging usually coupled with dimerization



NC-SL3 complex

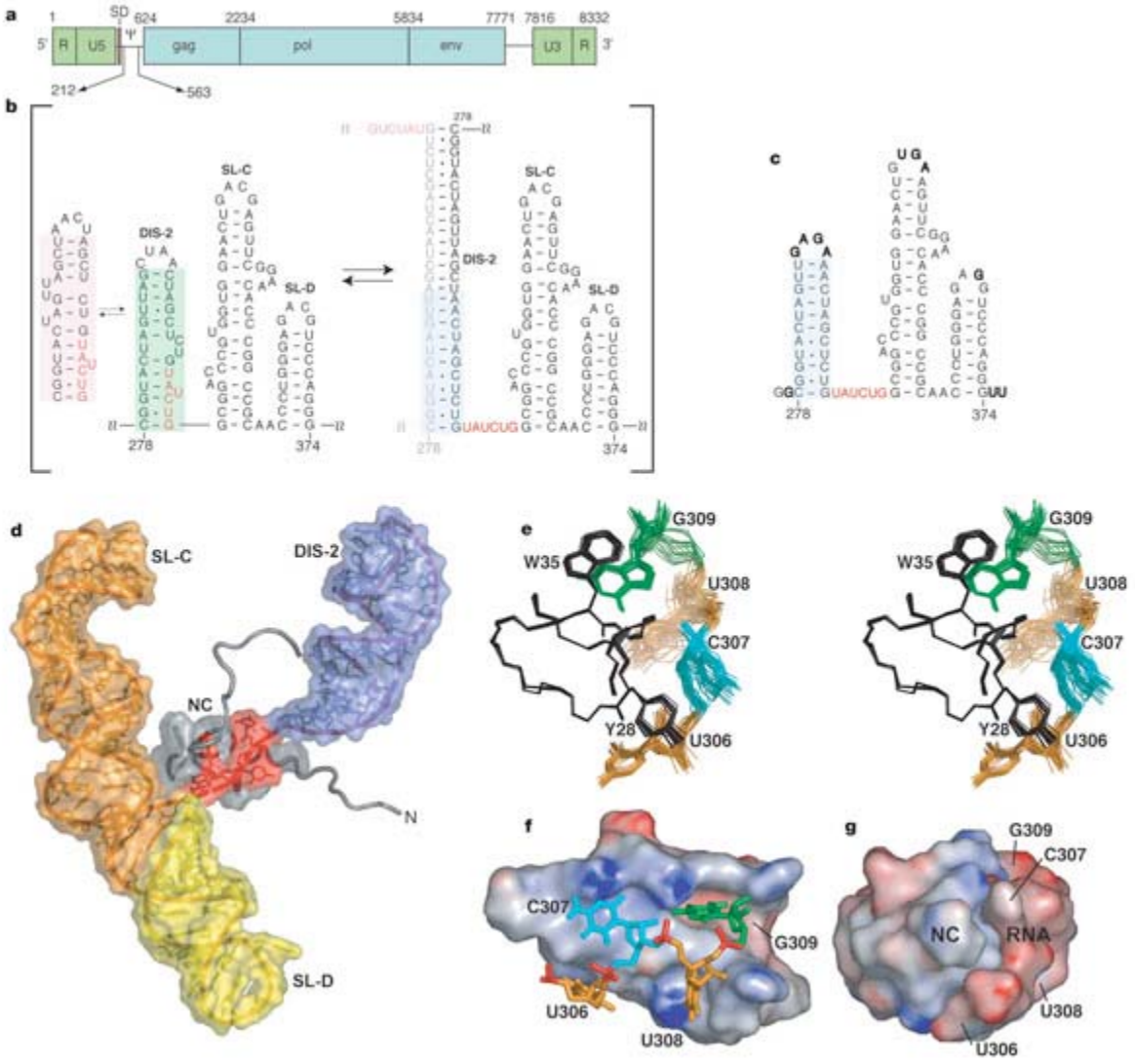
(M. Summers & co-workers, *Science,* 1998)

Dimer of retroviral RNA



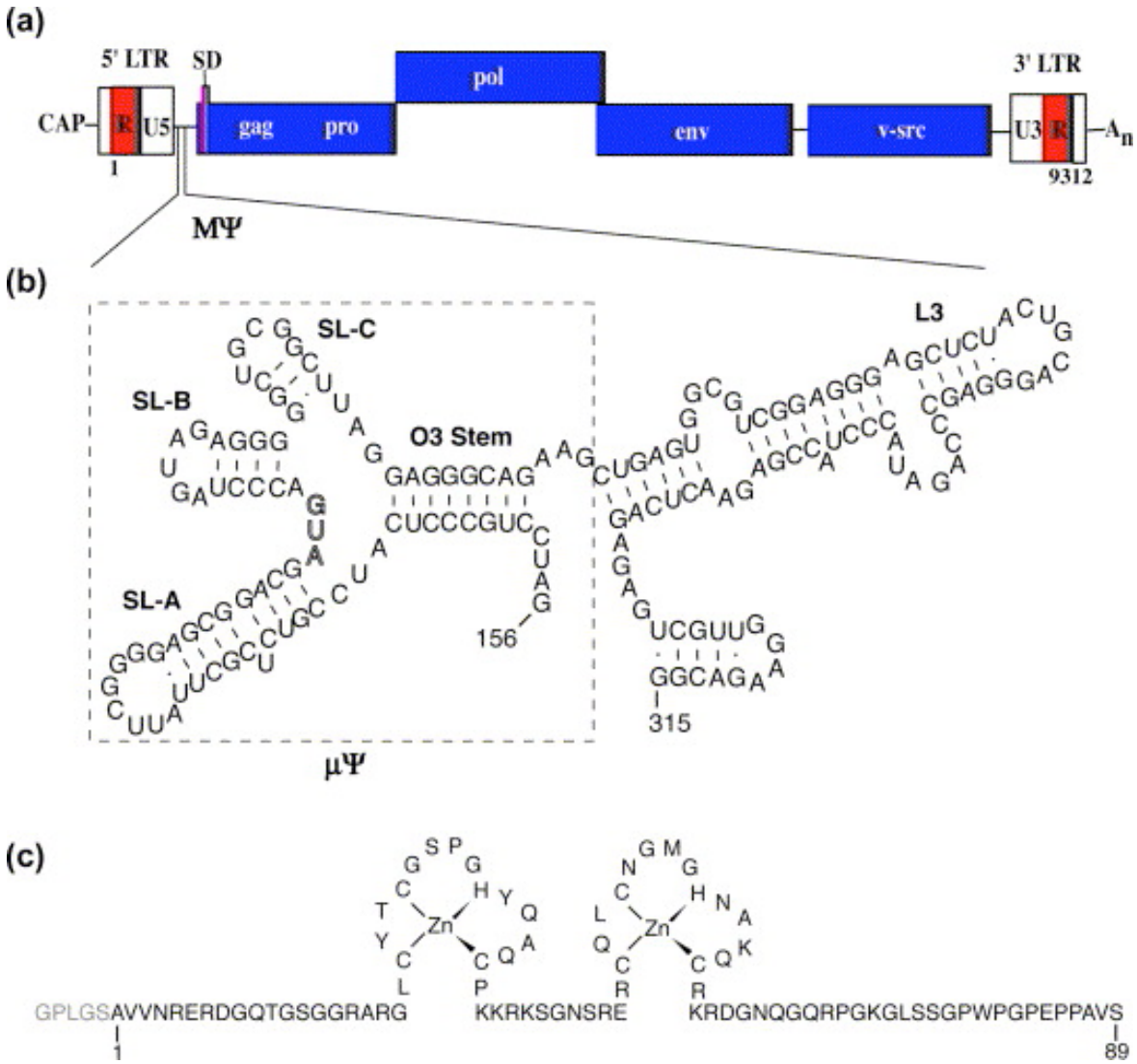
(Top panel) An electron micrograph of virion RNA from the feline retrovirus RD114. The dimer linkage structure appears as a rod-like structure near the center of the figure with short single-stranded tails representing the 5′ end. Each RNA subunit has a looped feature of secondary structure. (Courtesy of Hsing-Jien Kung.) (Bottom panel) An electron micrograph of virion RNA from HIV-1. The position of the dimer linkage structure (DLS) is shown, which for HIV-1 forms a small loop. Bar, 100 nm.

MLV RNA is suggested to be selected by UCUG exposure due to genome dimerization



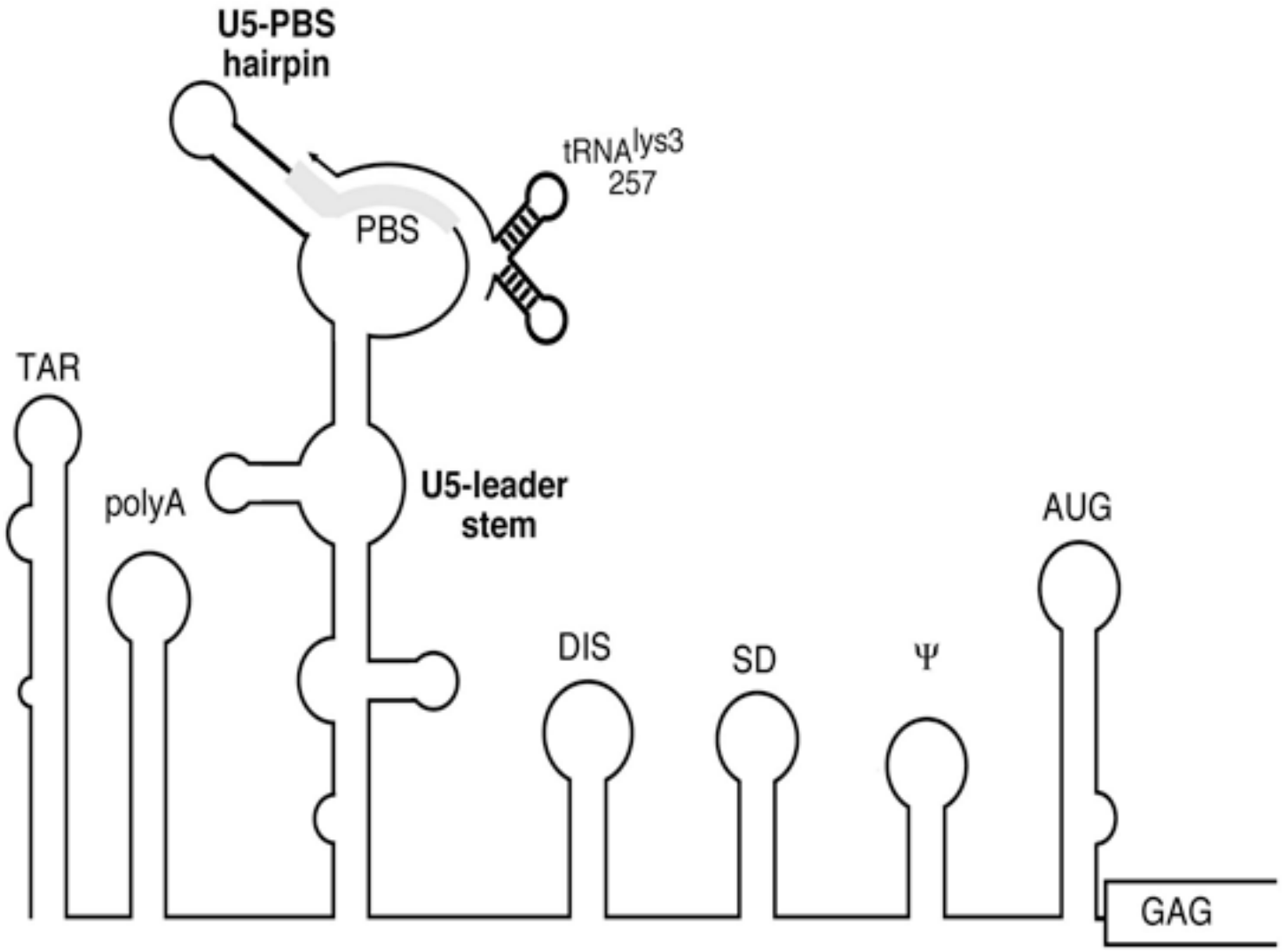
Summers et.al. 2005 Nature

Ψ region of RNV is short (80 nt) and highly structured.
RSV genome does not have to dimerize to be packaged



Summers et.al. 2006

RNA secondary structure model of HIV-1 5’ untranslated leader region



In HIV-1 Ψ region contains dimerization signal in the region of high local stem density

My hypothesis for viral RNA recognition:

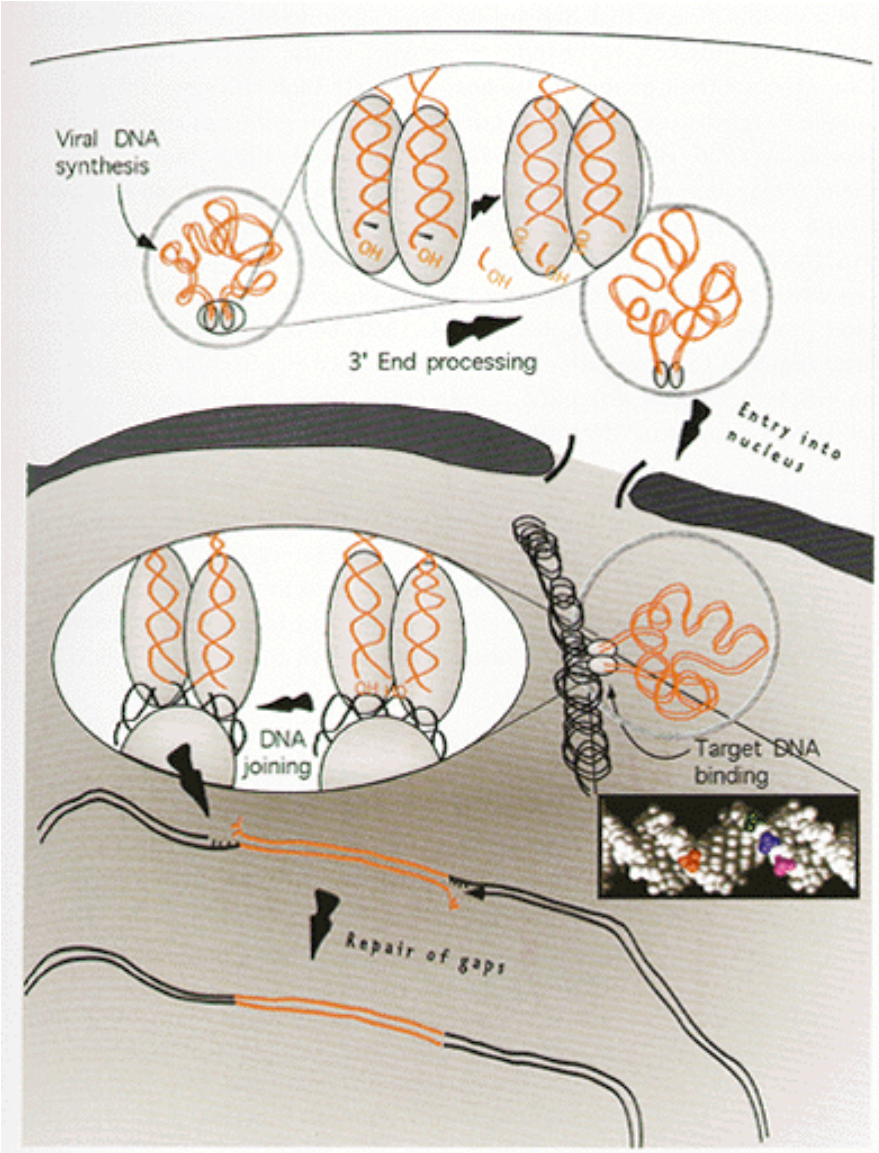
NC binds preferentially to RNA regions of high stem-loop
Concentration + some sequence specificity in ss regions=

Strong electrostatics + specific zinc-finger binding

Just electrostatics is insufficient to explain selectivity:
Virus-specific RNA/NC interactions

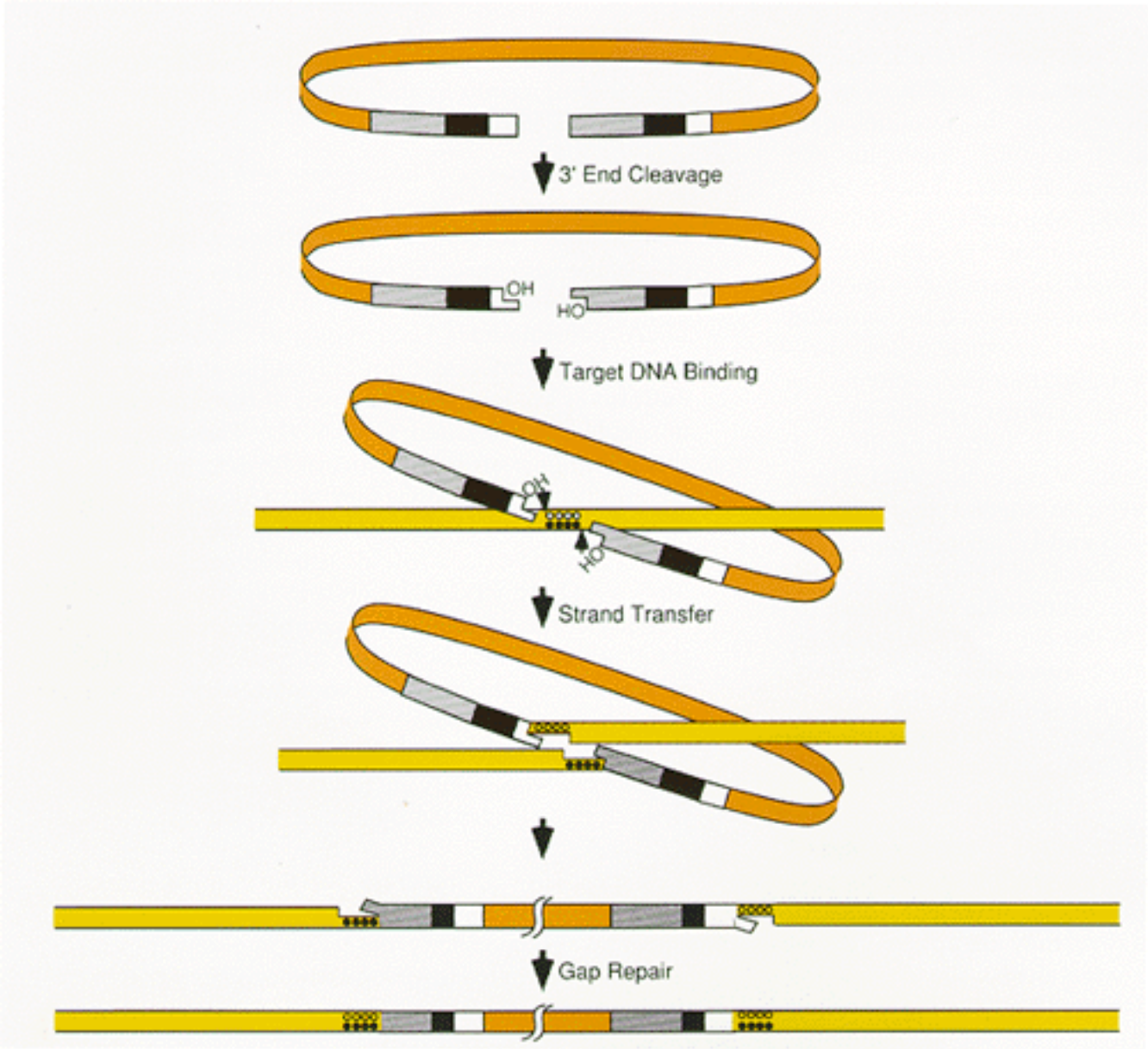
Observed slightly preferential packaging of some highly structured mRNA and tRNA

The integration process.

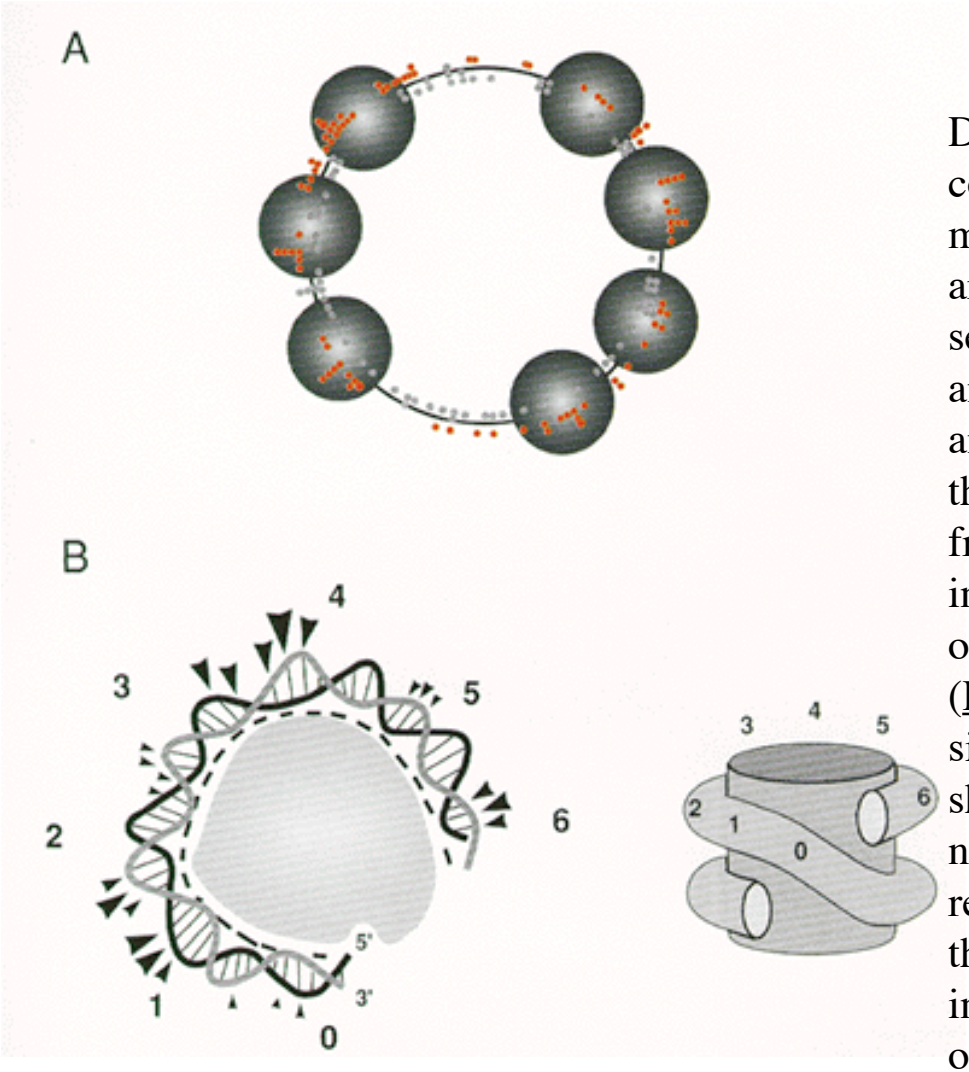


The protein components of the preintegration complex are represented by the gray circle surrounding the viral DNA molecule (color). Integrase is represented by the ovoids at the ends of the viral DNA. The host DNA is shown assembled into nucleosomes, reflecting the predilection for integration to occur at sites exposed on the surface of nucleosomes. The steps illustrated are described in the accompanying text. (Inset) Representation of B-form DNA, with one phosphate colored in red, representing one of the phosphates attacked by a viral DNA 3′ end during integration. The phosphates on the opposite strand, separated by a 4-, 5-, or 6-base 5′ stagger, are colored in green, blue, and magenta, respectively. These correspond to the spacing of target phosphates attacked during integration of MLV, HIV, and RSV, respectively, accounting for the 4-, 5-, or 6-base-pair target sequence repeats that flank the corresponding proviruses. Note that the two target phosphates in each case are on the same face of the B-form double helix, flanking the major groove.

Schematic outline of the principal steps in retroviral DNA integration.

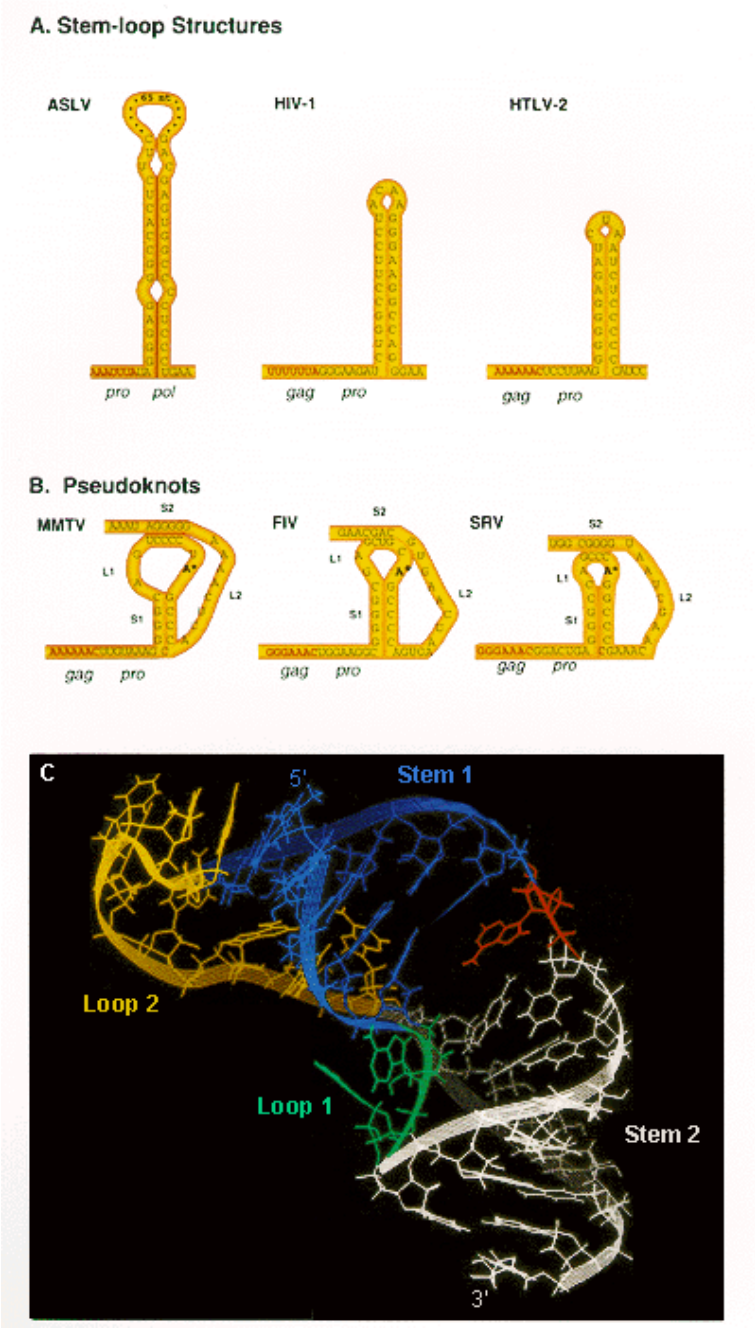


Nucleosome structure influences local target site specificity of integration.



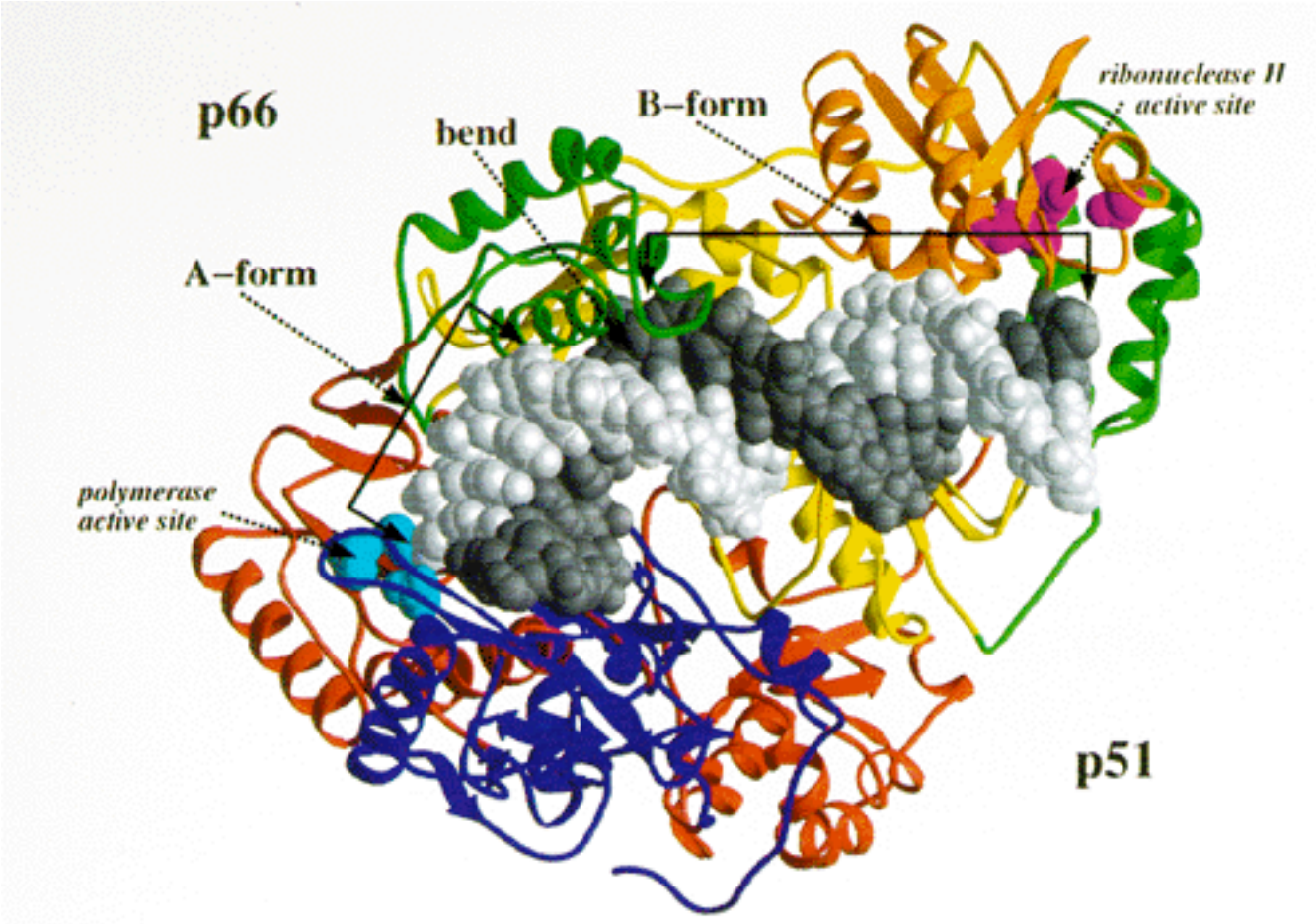
Distribution of targets used by MLV integration complexes in an in vitro assay using a 1453-base-pair minichromosome target. The positions of nucleosomes are indicated by the large gray circles. The positions of sequenced sites of integration into minichromosomes and the naked plasmid DNA are indicated by the red and gray dots, respectively. Note that integration into the minichromosomal DNA occurs significantly more frequently into nucleosomal than into internucleosomal regions, whereas that bias is not observed for integration into the naked plasmid (Pryciak et al. 1992b). (B) Distribution of integration sites in nucleosomal DNA (Pruss et al. 1994a). The shaded area represents the histone octamer. The numbers in the inset diagram of a nucleosome represent helical turns, counting from the dyad axis of the nucleosome. Arrowheads indicate sites of integration of HIV-1 DNA ends in vitro, with the size of the arrowheads representing the relative frequency of integration at each site. Note that positions flanking the most severely distorted, widened major grooves are highly preferred as integration targets.

Regulation of frame shift by viral RNA structure



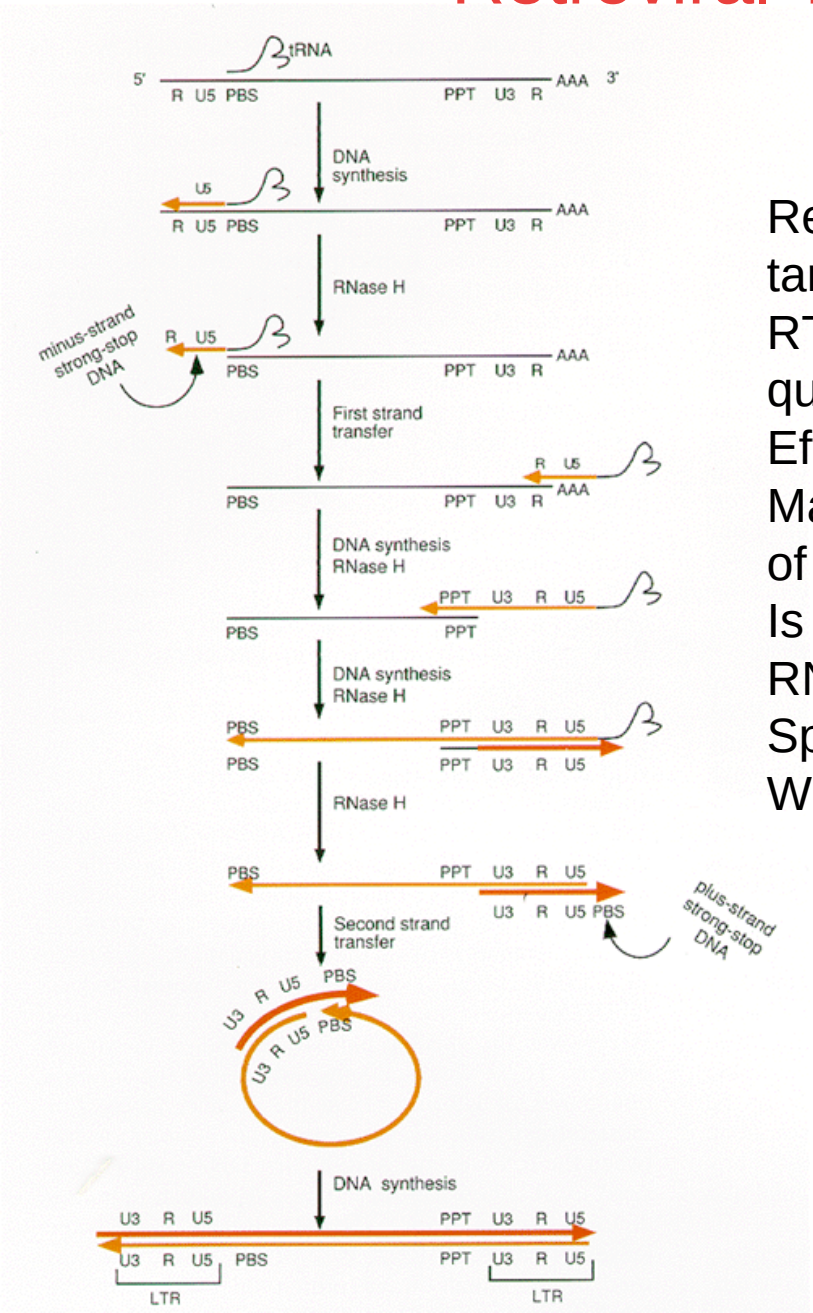
Secondary structure downstream from frameshifting sites. (A) Hairpins downstream from the ASLV, HIV-1, and HTLV-2 frameshift sites. (Modified from Jacks 1990.) (B) Pseudoknots downstream from the MMTV, FIV, and SRV-1 frameshift sites. Elements of secondary structure are as described in Chen et al. (1995) and the legend to Fig. 6. The asterisks indicate the position of an unpaired adenosine which is in position to disrupt the coaxial stacking of the two stems. (C) Structure of the pseudoknot at the gag-pro boundary of MMTV as determined by NMR (Reprinted, with permission, from Shen and Tinoco 1995).

Reverse Transcriptase in complex with RNA/DNA substrate



RT has two functions: polymerase and RNase

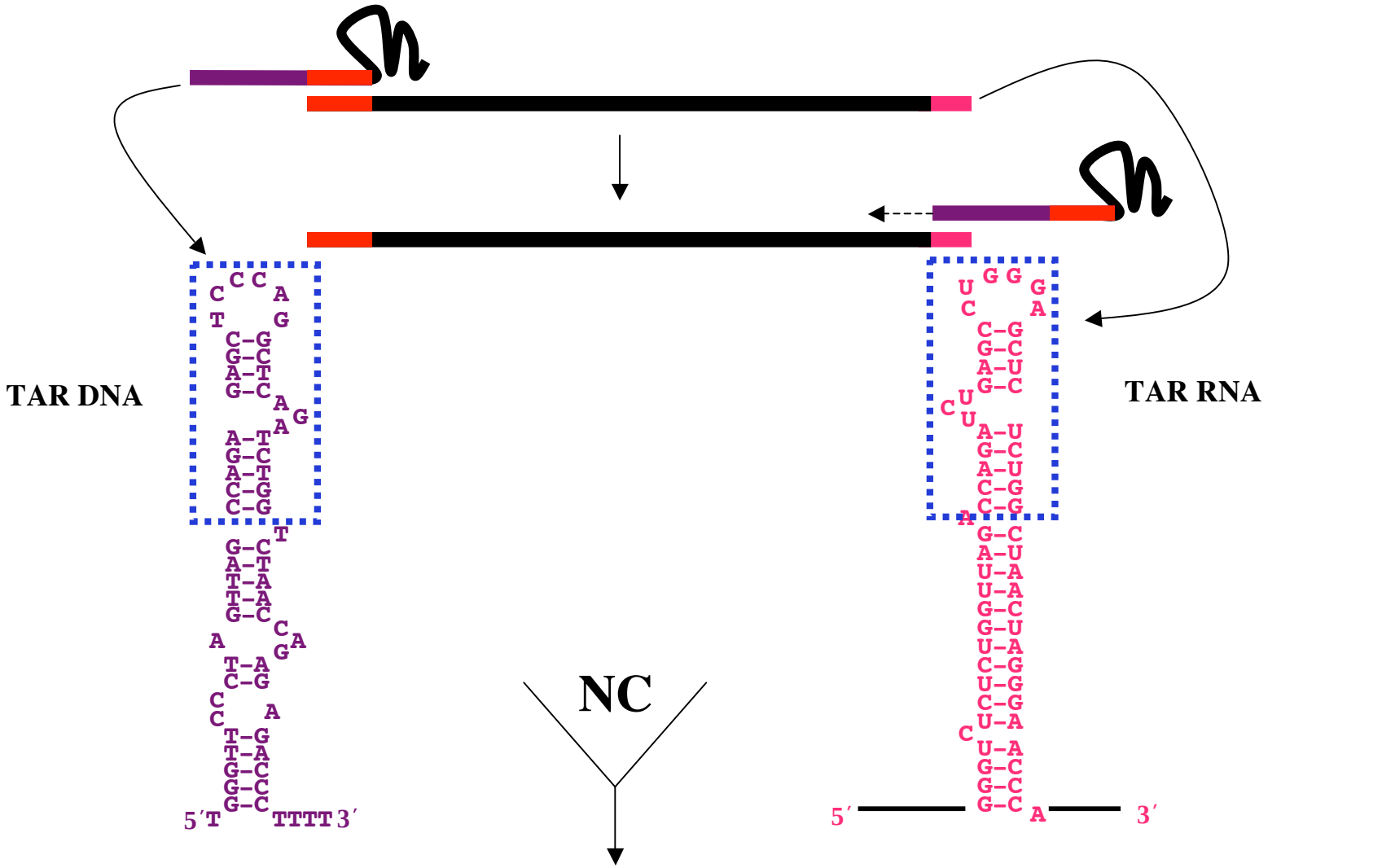
Retroviral Transcription



Reverse transcription is a main antiviral target: nt substitutes
RT activity is well studies, but still questions remain: initiation, fidelity, rate
Effect of NC;
Major question: efficiency and patterns of viral strand recombination:
Is it affected by RNA sequence?
RNA structure?
Spatial proximity due to RNA dimerization?
What creates inter viral recombination block?

(Black line) RNA;
(light color) minus-strand DNAs;
(dark color) plus-strand DNA.

The annealing step in minus-strand transfer depends on NC

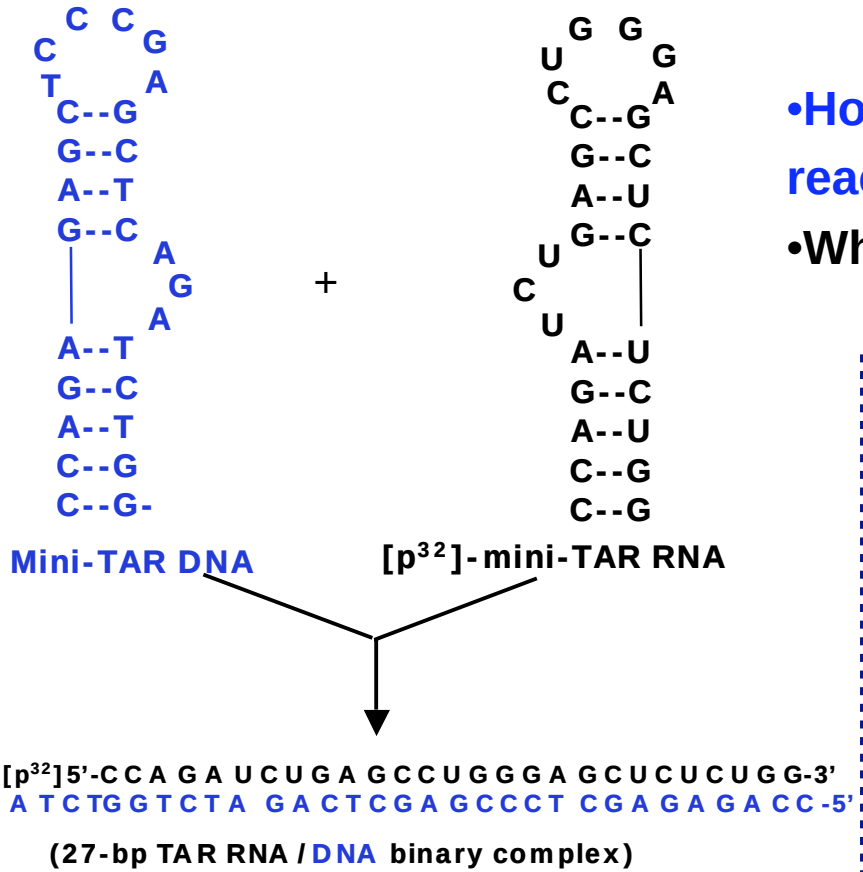


3′——CCCAGAGAGACCAATCGGTCTAGACTCGGACCCTCGAGAGACCGATTGATCCCTTTGGGT 5′

5′——GGGUCUCUCUGGUUAGACCAGAUCUGAGCCUGGGAGCUCUCUGGCUAACUAGGGAACCCA——3′

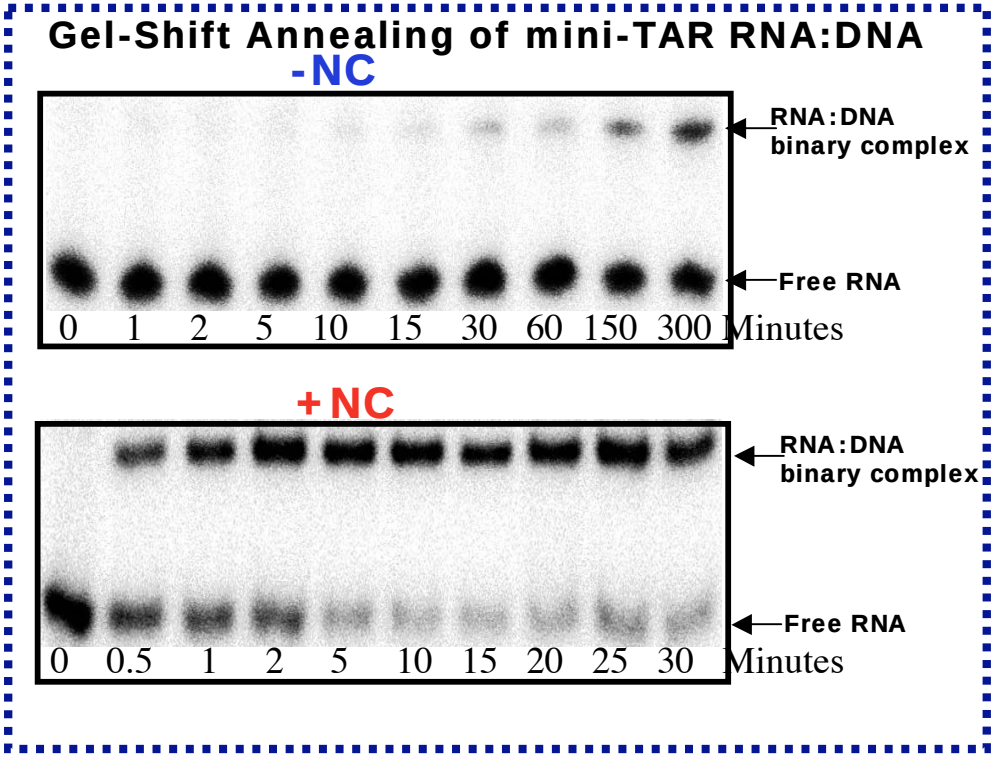
TAR RNA-DNA binary complex

Mini-TAR RNA/DNA:
A minimal system to dissect the annealing pathway



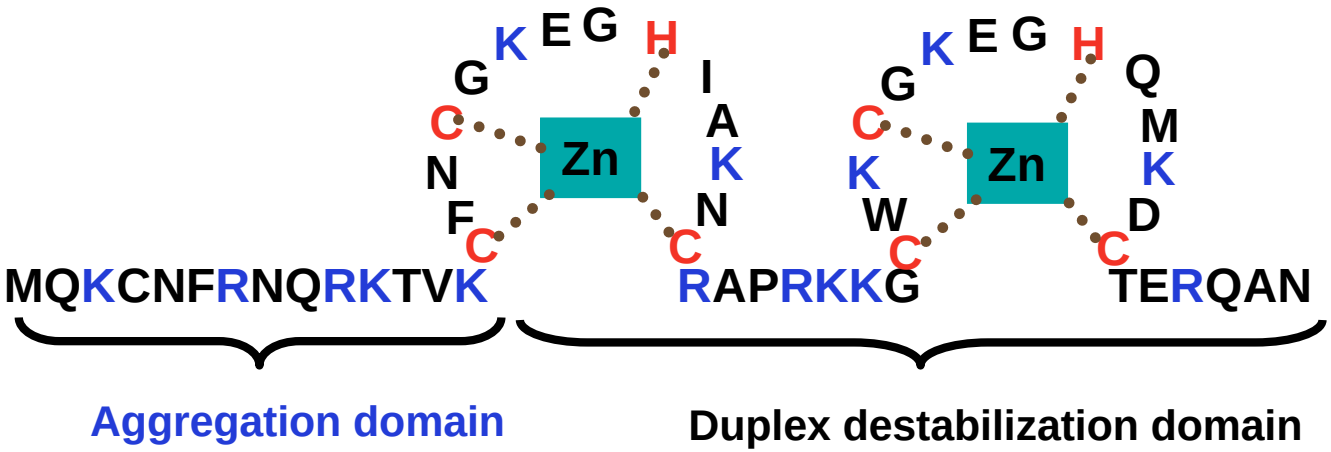
Questions

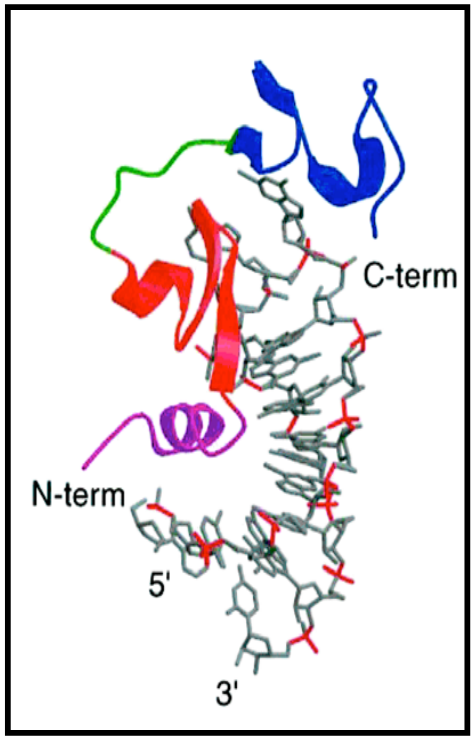
- How much does NC enhance the rate of this reaction?
- What is the annealing mechanism?



HIV-1 Nucleocapsid Protein (NC)
A Nucleic Acid Chaperone

15 basic residues (pI = 9.93)
2 nonequivalent CCHC Zn^{2+} "knuckles"



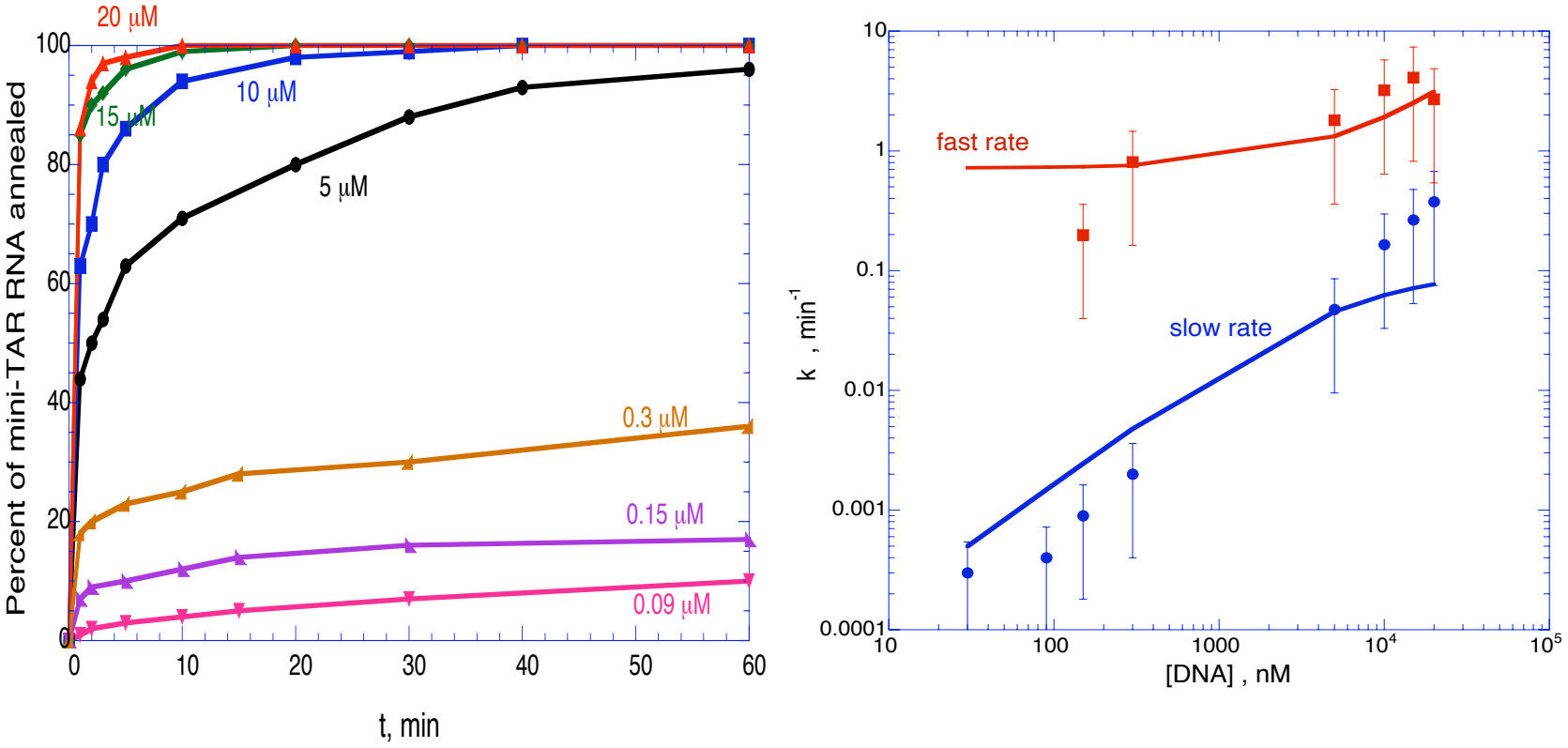


NC-SL3 complex

(M. Summers & co-workers, *Science,* 1998)

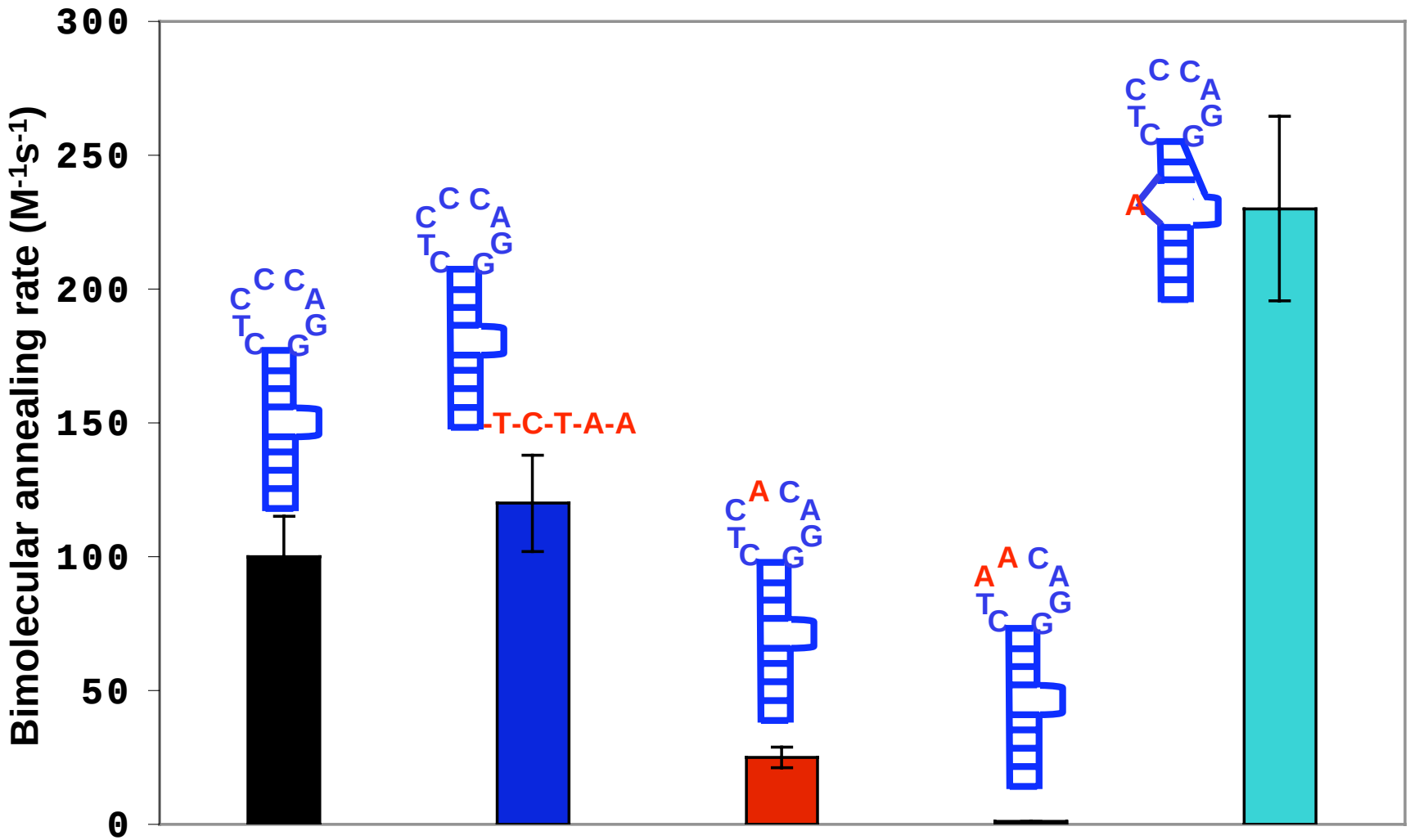
What is mechanism of NC's *ATP-independent* chaperone activity?

Dependence of mini-TAR RNA/DNA annealing on excess [DNA] in the absence of NC

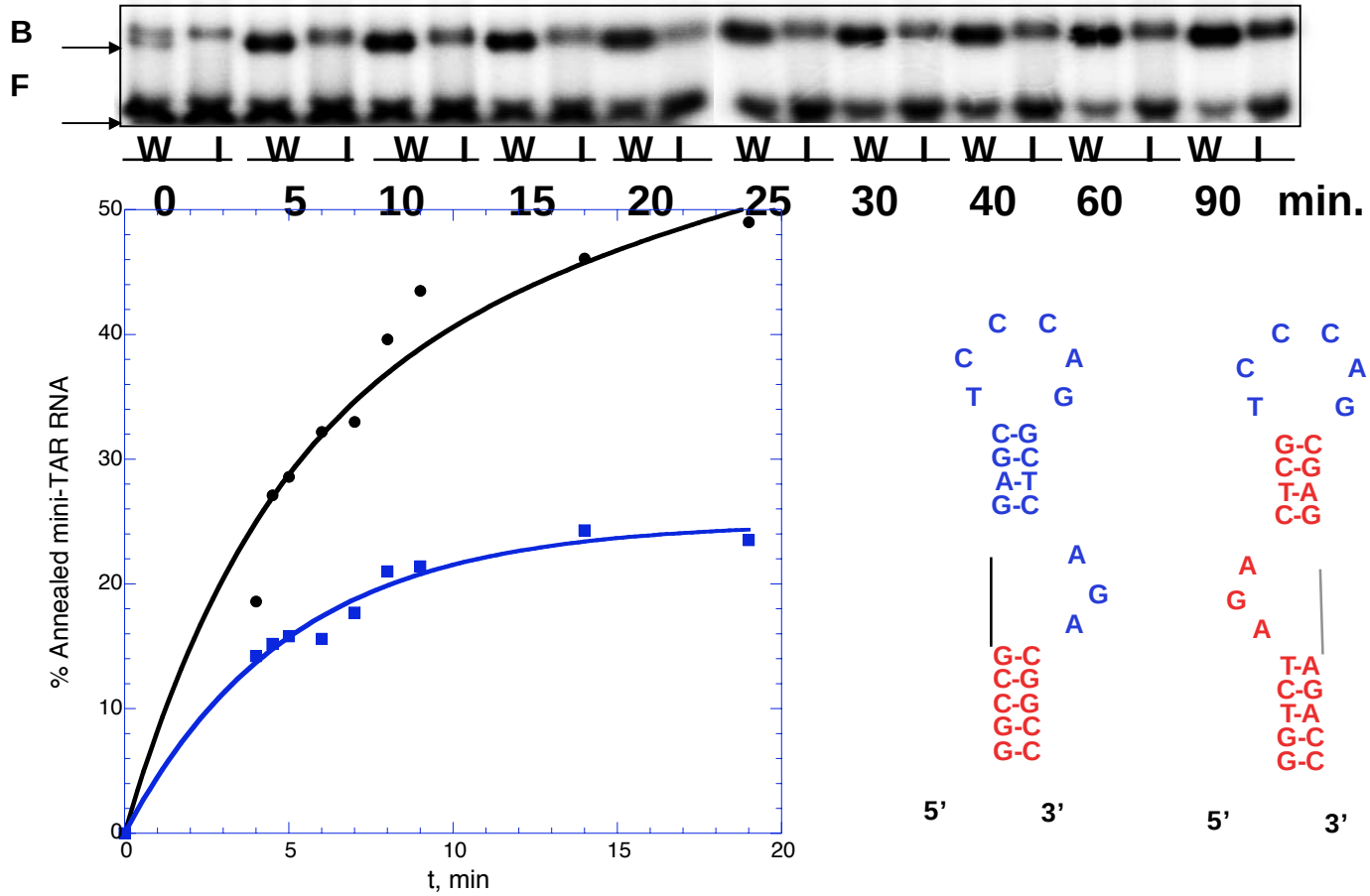


Annealing is bi-exponential with fraction of fast component growing with stability of annealing intermediate

Mutations in the loop dramatically decrease mini-TAR annealing in the absence of NC



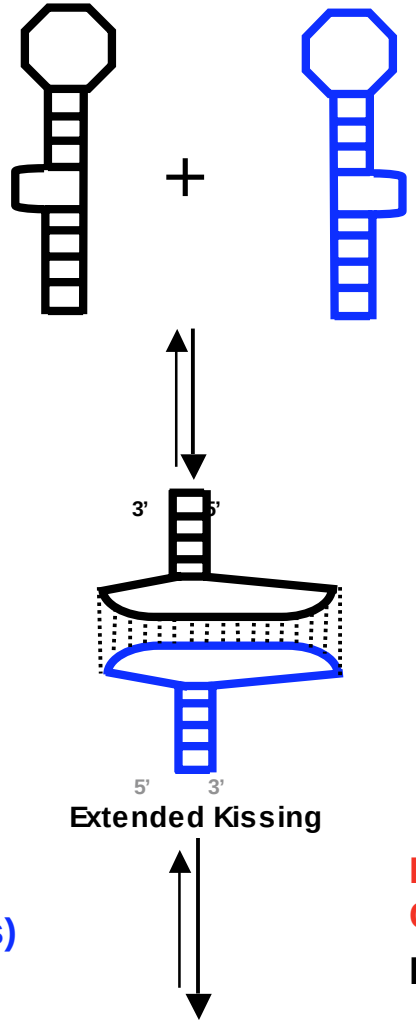
Two annealing rates correspond to fast formation of extended kissing complex and its slower conversion to duplex



- No annealing of mini-TAR RNA to mini-TAR DNA with only loop complementarity is observed
- Annealing kinetics of mini-TAR DNA with terminal inverted stem is similar to fast component of wt annealing.
- Annealing intermediate is extended kissing that co-migrates on the gel with wt

HIV-1 TAR RNA/TAR DNA annealing proceeds via a two-step process

$$R + D \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} RD^* \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} RD$$



bimolecular association

Nucleic acid aggregating ability (Cationic amino acids)

unstable extended kissing in pre-equilibrium 100-fold facilitation

k_1~10^3 $M^{-1}s^{-1}$, k_{-1}~1 min^{-1} (- NC)

k_1~10^5 $M^{-1}s^{-1}$, k_{-1}~1 min^{-1} (+ NC)

Extended Kissing

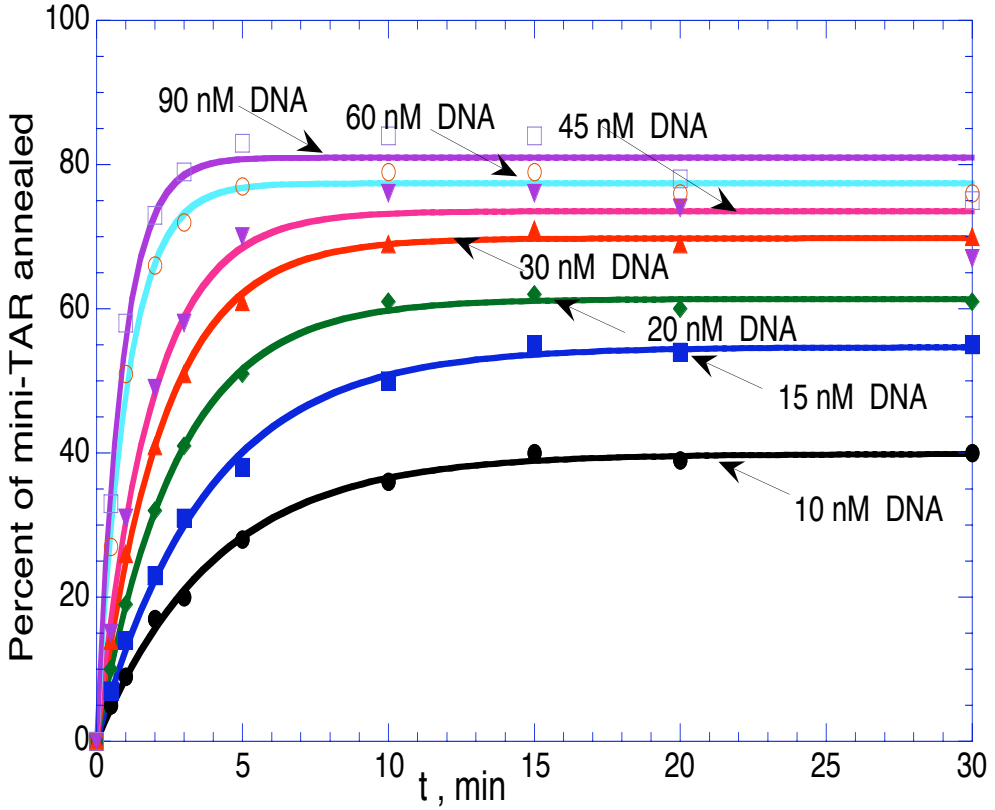
Duplex destabilizing activity (Zinc fingers)

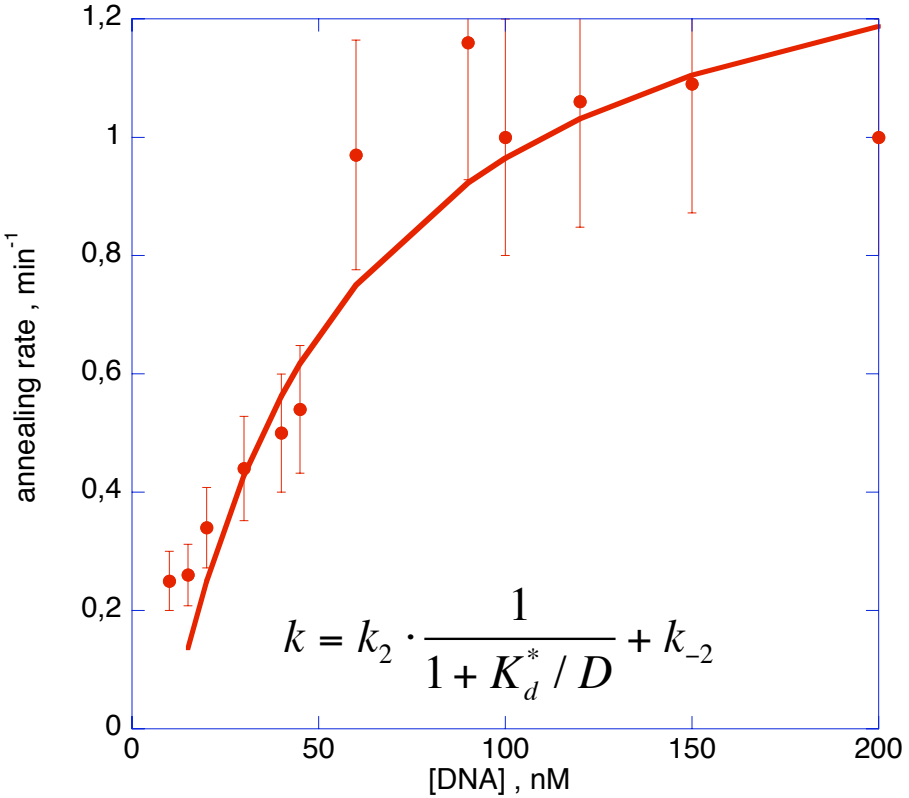
Rate-limiting Conversion 10-fold facilitation

k_2~0.1 min^{-1} (- NC); k_2~1 min^{-1} (+ NC)

RNA:DNA Binary Complex

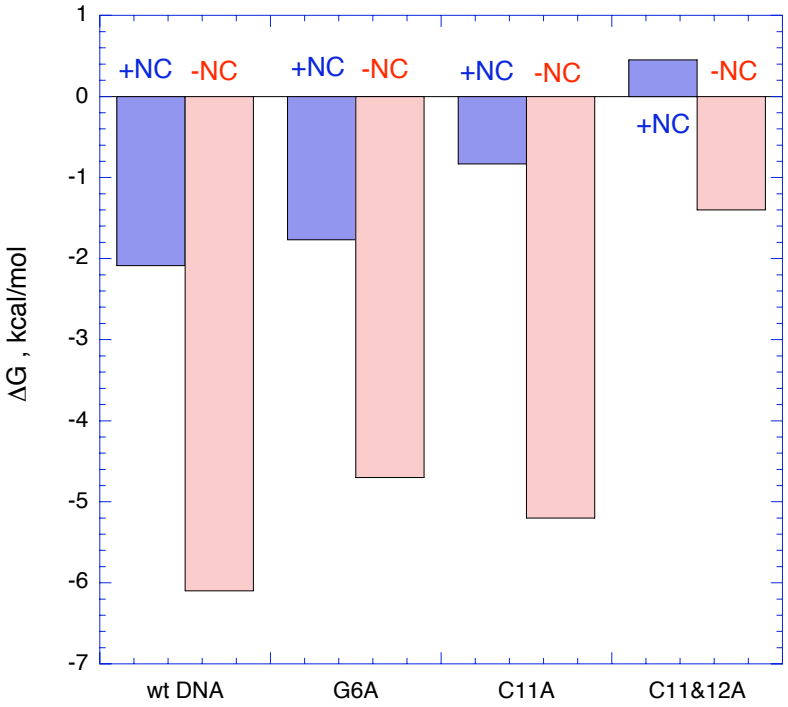
Dependence of mini-TAR RNA/DNA annealing on [DNA] in the presence of saturating [NC].

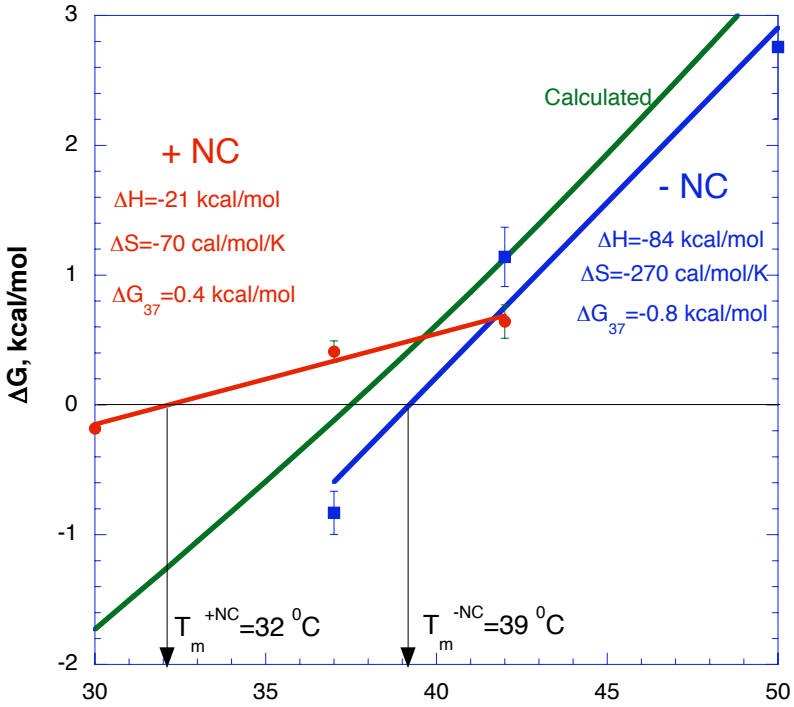




- 1000-fold facilitated single exponential annealing.
- Annealing becomes reversible and is incomplete in equilibrium.
- The duplex dissociation rate ~0.1 min^{-1}
- Annealing rate saturates at high [DNA] at conversion rate~1min^{-1}

Effect of NC on mini-TAR RNA/DNA annealing equilibrium



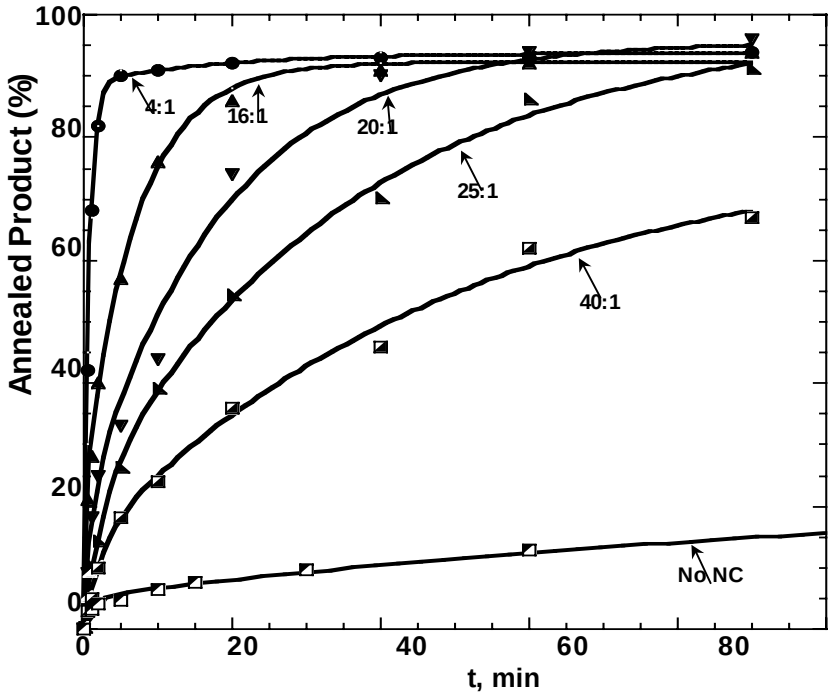


$$\Delta G = -RT\ln(K) \qquad K = \frac{P_\infty}{100 - P_\infty} = \frac{D}{K_d} \qquad \Delta G \overset{T,C}{=} \Delta H - T\Delta S$$

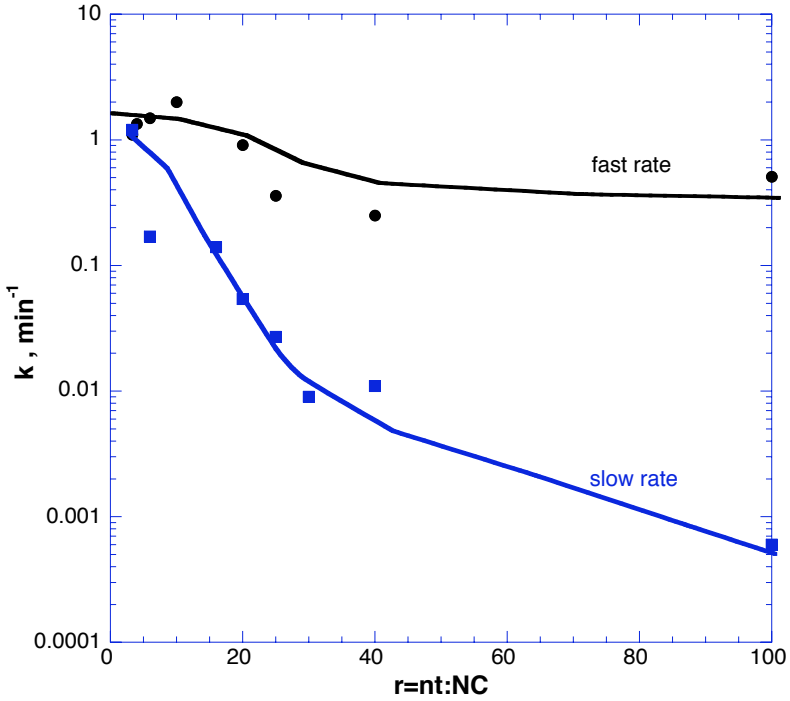
- Incomplete equilibrium annealing level of mini-TAR RNA/DNA can be used to estimate annealing free energy ΔG; ΔG can be calculated in the absence of NC;
- NC destabilizes duplexes by ~1 kcal/mol/bp & decreases enthalpy and entropy of duplex melting

Annealing depends strongly on nt:NC ratio

a

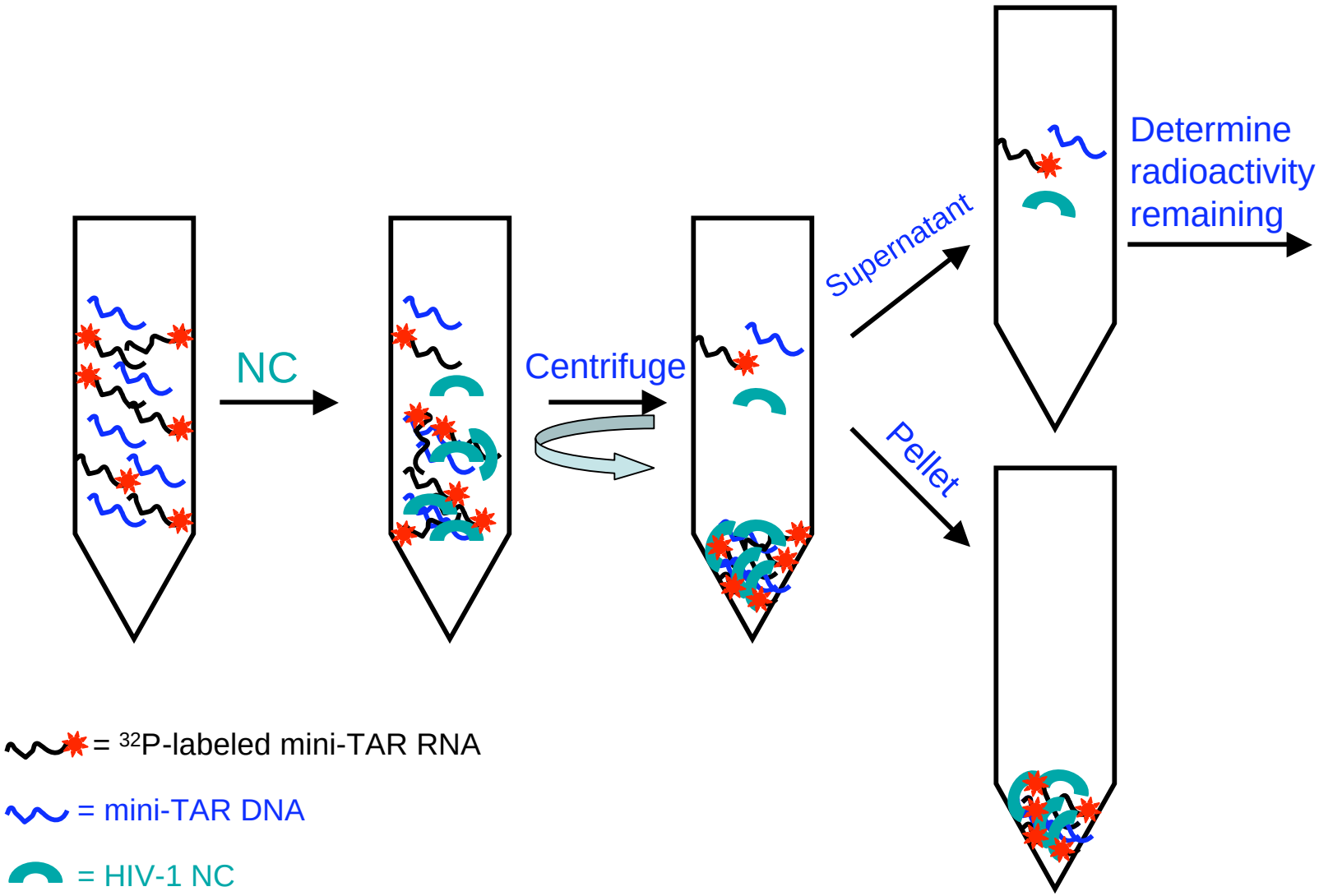


100

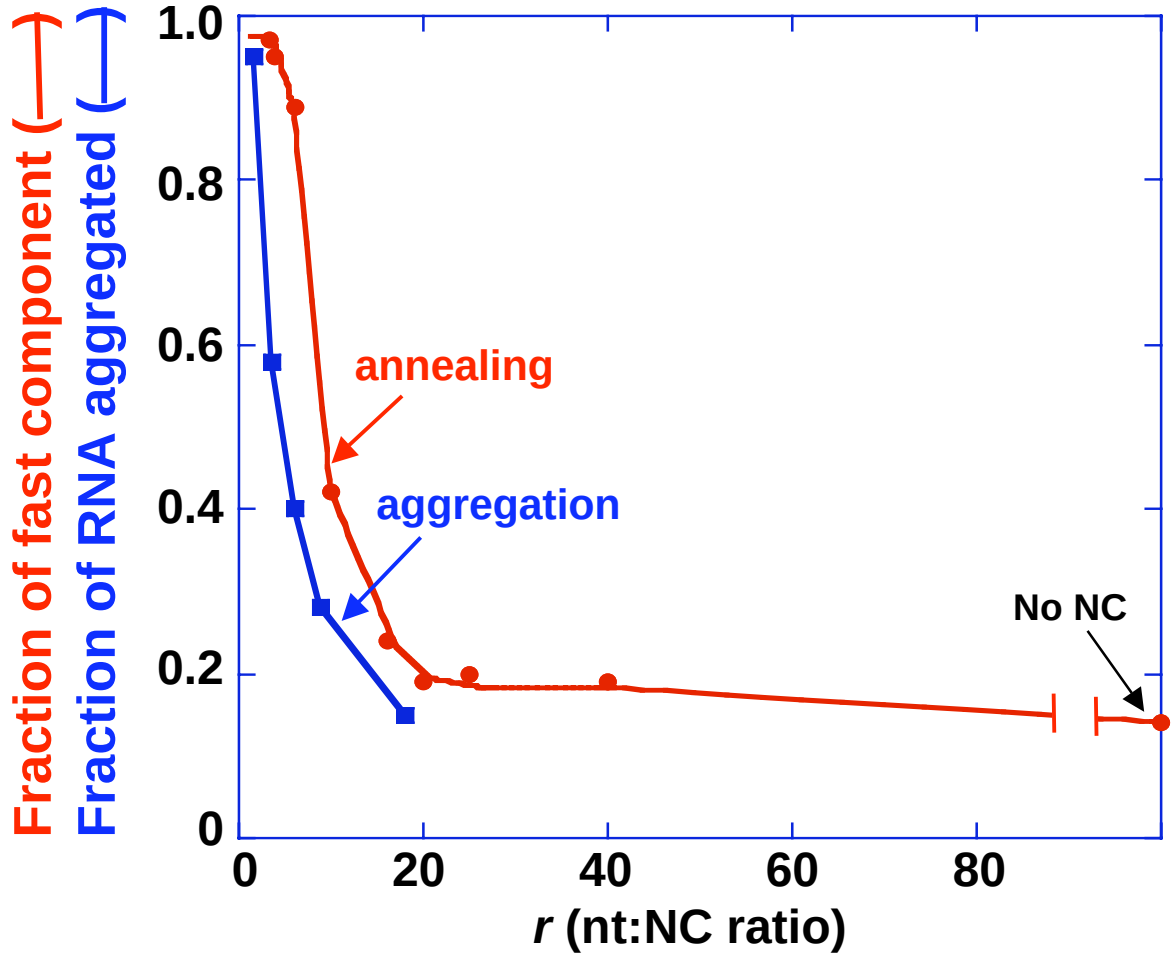


- Subsaturated [NC] slows down all elementary reaction rates, with the largest effect on conversion;
- Facilitation of bi-molecular step requires saturation with NC, i.e.r=nt:NC<6;
- Significant duplex destabilization is observed at much higher r;
- Fast fraction and slow rate grow dramatically within narrow range 4<r<15;
- Slow and fast rates converge at saturated NC;
- annealing intermediate is stabilized at saturated NC.

Aggregation Assay

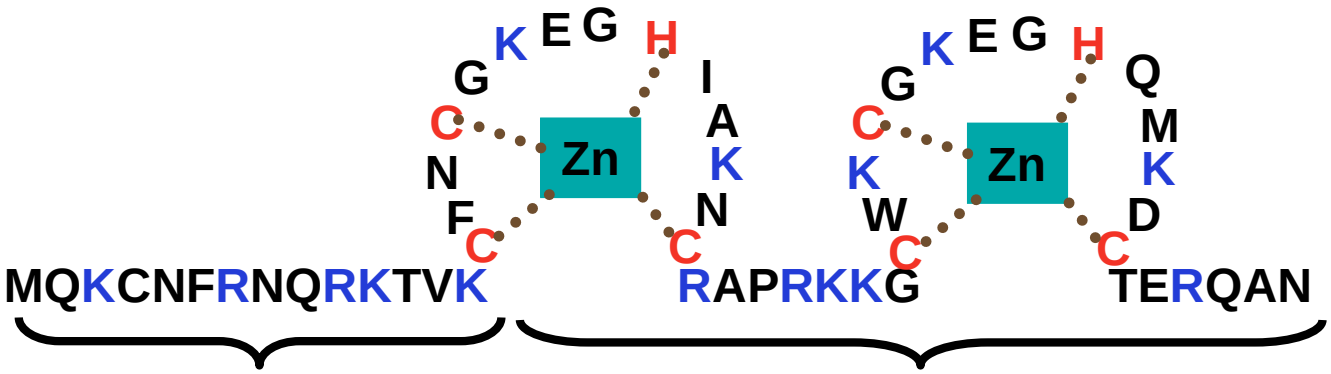


Direct correlation between the intermediate stabilization and RNA aggregation



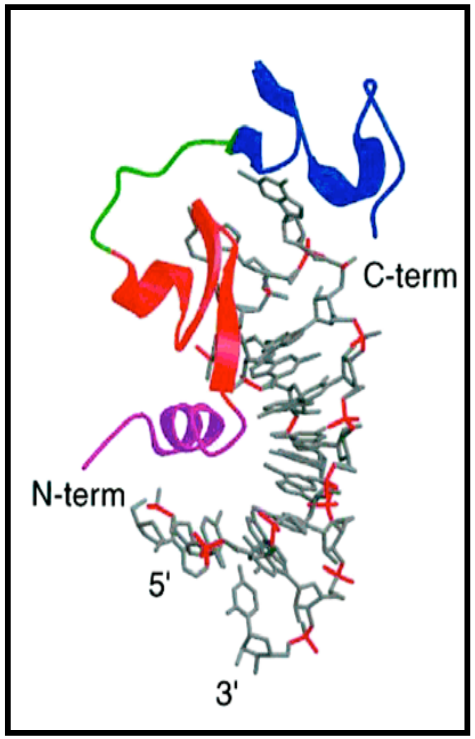
HIV-1 Nucleocapsid Protein (NC)
A Nucleic Acid Chaperone

15 basic residues (pI = 9.93)
2 nonequivalent CCHC Zn^{2+} "knuckles"



Aggregation domain
~100-fold facilitation

Duplex destabilization domain
~10-fold facilitation



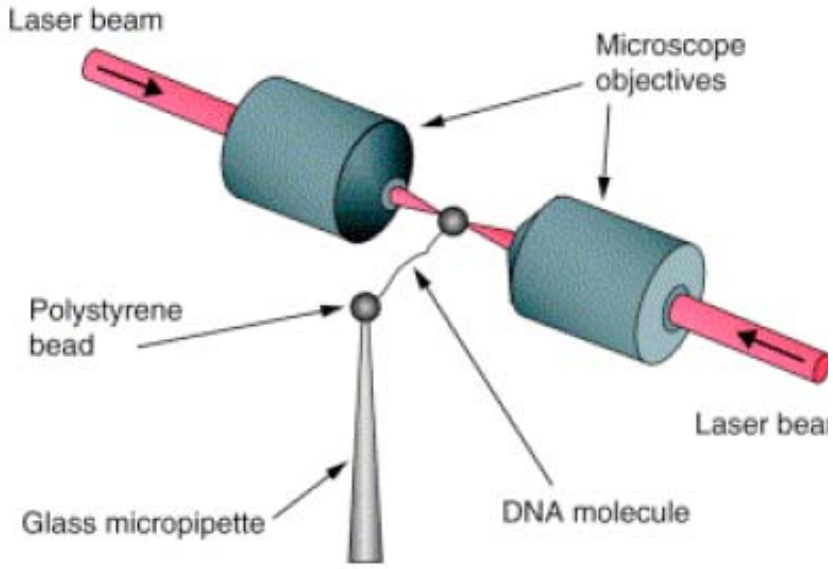
NC-SL3 complex

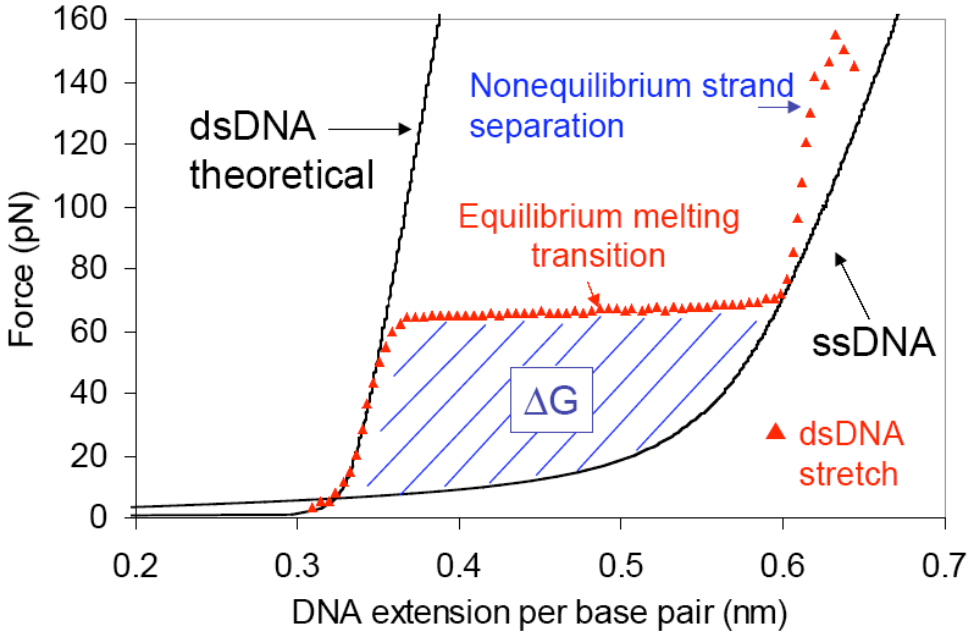
(M. Summers & co-workers, *Science,* 1998)

What is mechanism of NC's *ATP-independent* chaperone activity?

NC's effect on DNA duplex stability and strand annealing kinetics via single-molecule DNA stretching

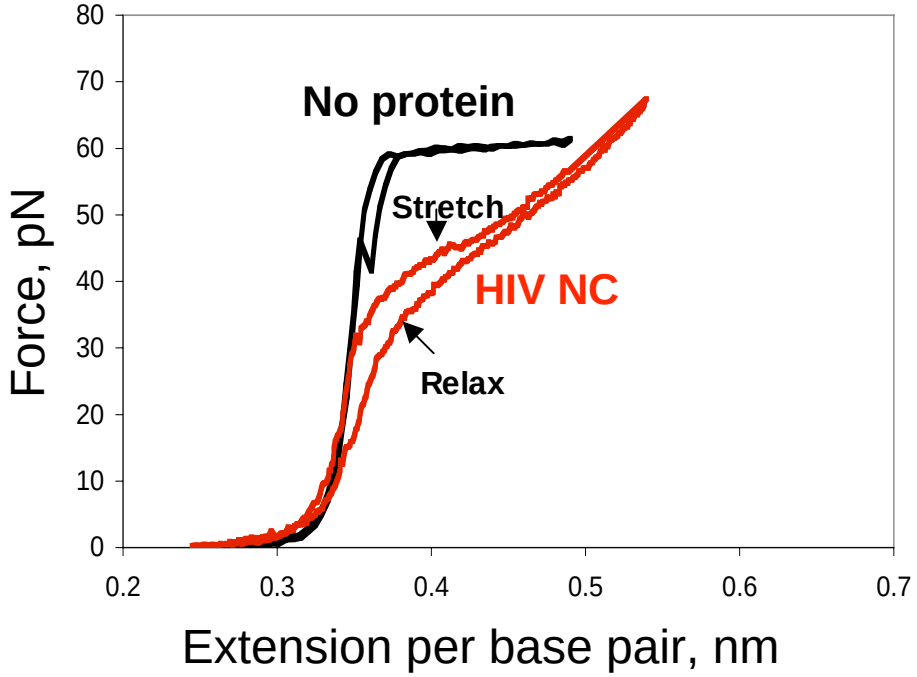
Collaborator: Prof. Mark Williams (Northeastern University)





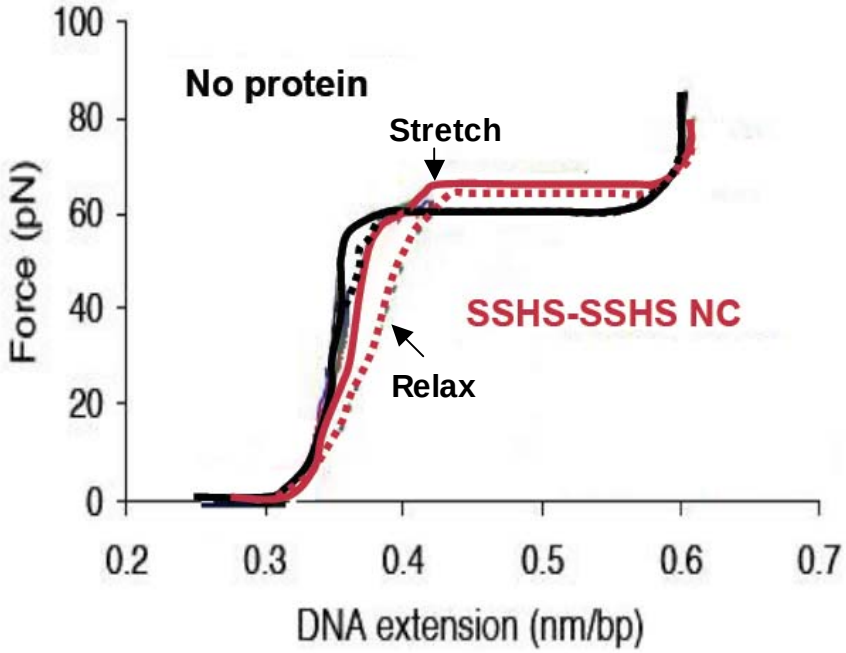
- Shape of stretching transition provides information on *destabilization* activity.
- Hysteresis (i.e., lack of an exact match between stretch and relax curves) provides information on strand *annealing* activity

Wild-type HIV-1 NC: effective duplex destabilization and strand re-annealing

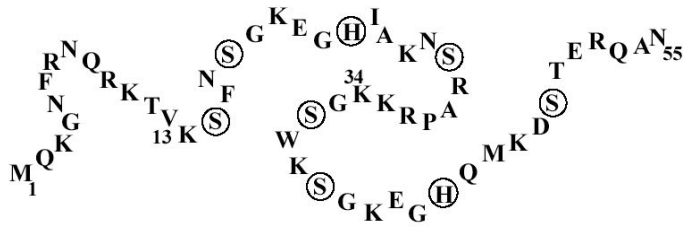


- Reduced melting force
- Small hysteresis upon relaxation

Zinc-less NC: no duplex destabilization but effective strand annealing agent

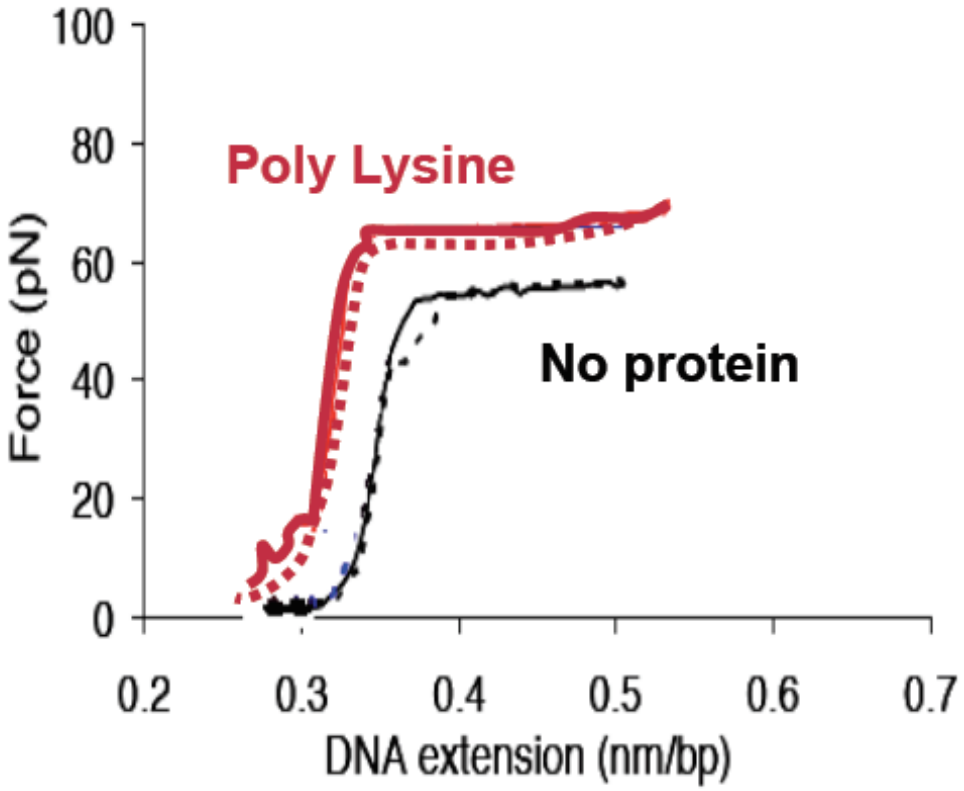


SSHS NC



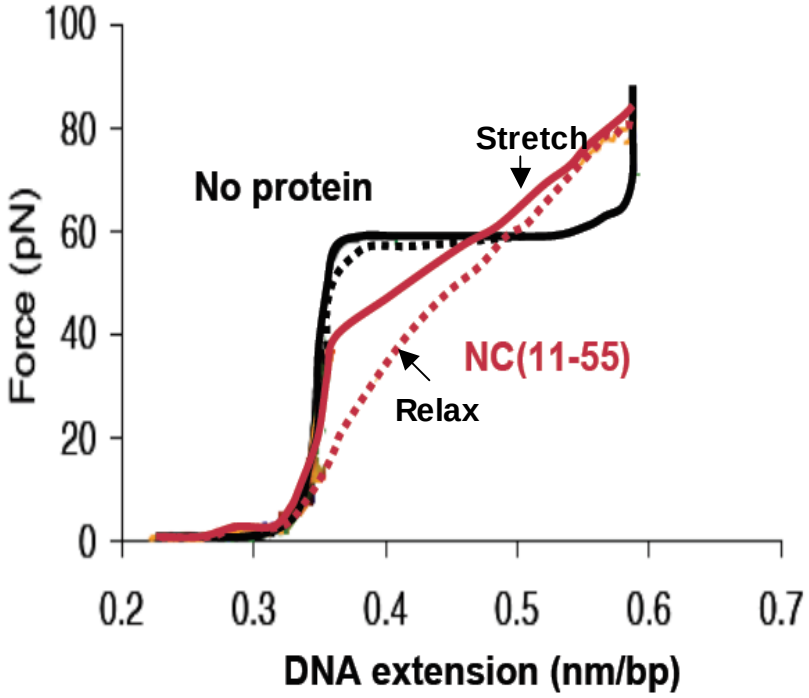
- Increased melting force
- Small hysteresis upon relaxation

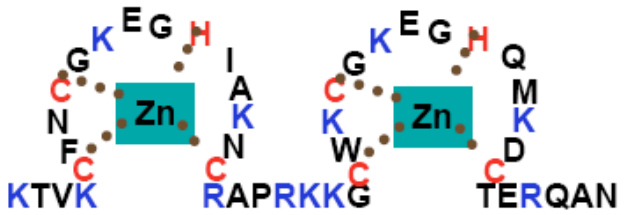
Poly Lysine: no duplex destabilization but more effective DNA strand re-annealing than NC or SSHS NC



- Increased melting force
- No hysteresis upon relaxation

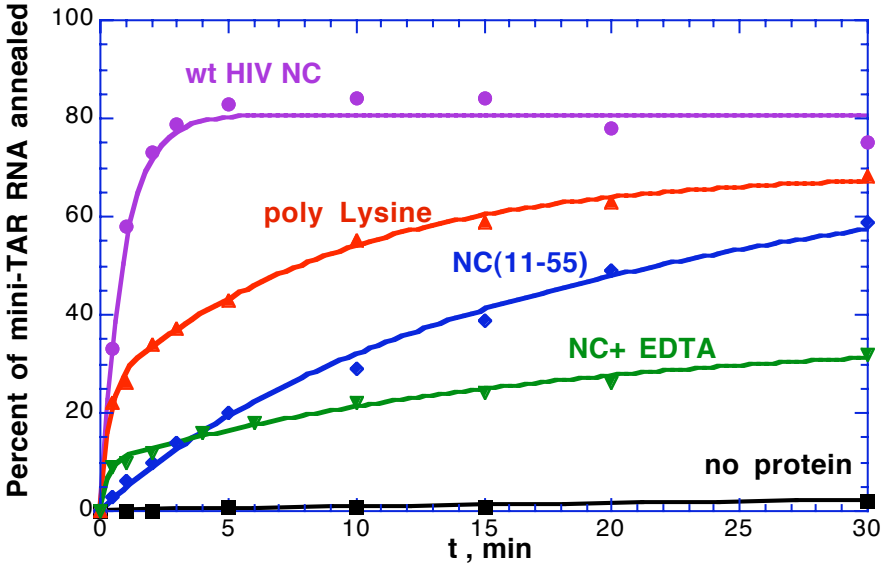
(11-55)NC: effective duplex destabilization but inefficient strand re-annealing

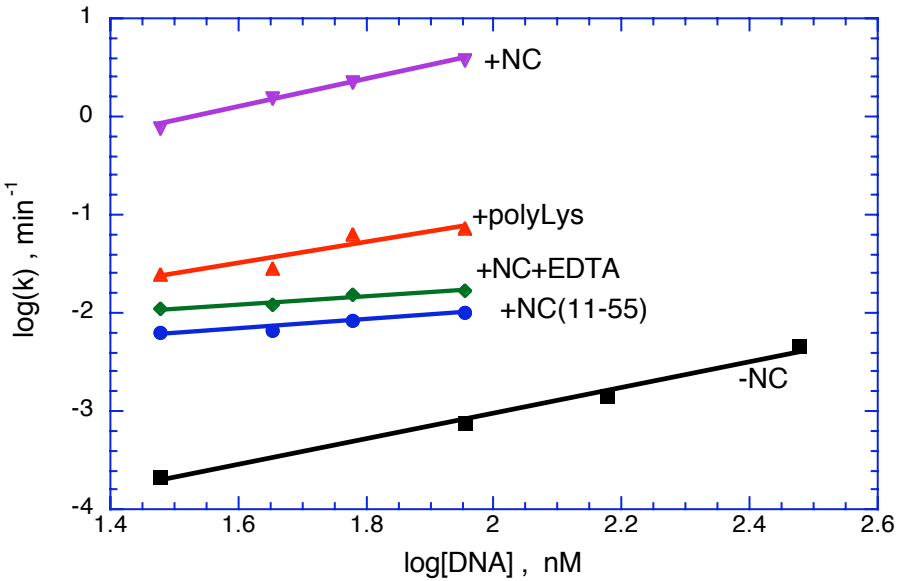




- Reduced melting force
- Greater hysteresis than WT NC

HIV-1 NC's zinc fingers and N-terminal domain contribute to facilitation of mini-TAR RNA/DNA annealing

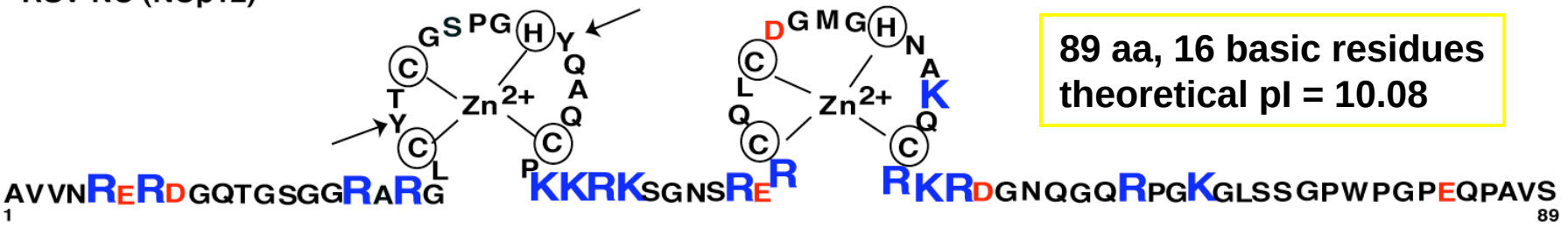




	rate enhancement
NC	~1000-fold
poly Lys	~100-fold
NC+EDTA	~70-fold
NC(11-55)	~50-fold

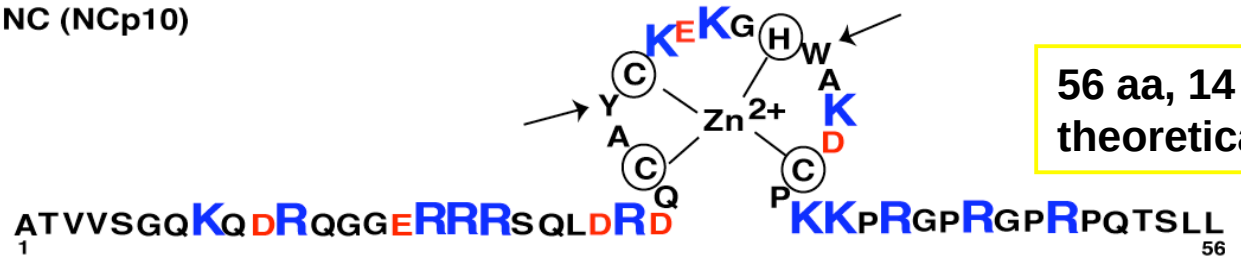
Comparative studies with RSV, MLV and HTLV NC Proteins

RSV NC (NCp12)



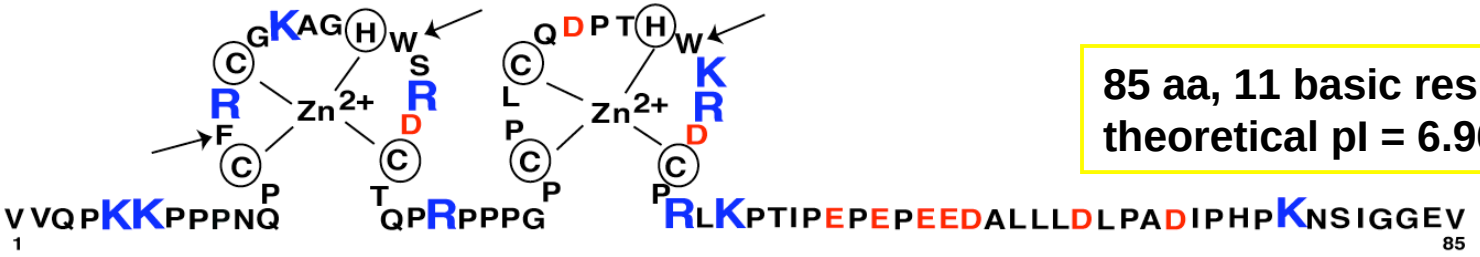
89 aa, 16 basic residues
theoretical pI = 10.08

Mo-MLV NC (NCp10)



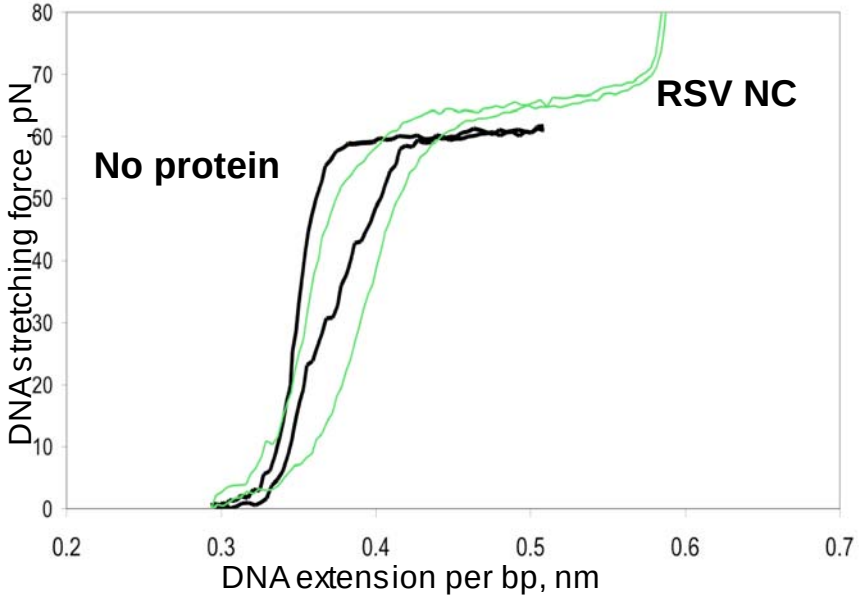
56 aa, 14 basic residues
theoretical pI = 10.15

HTLV-1 NC (NCp15)



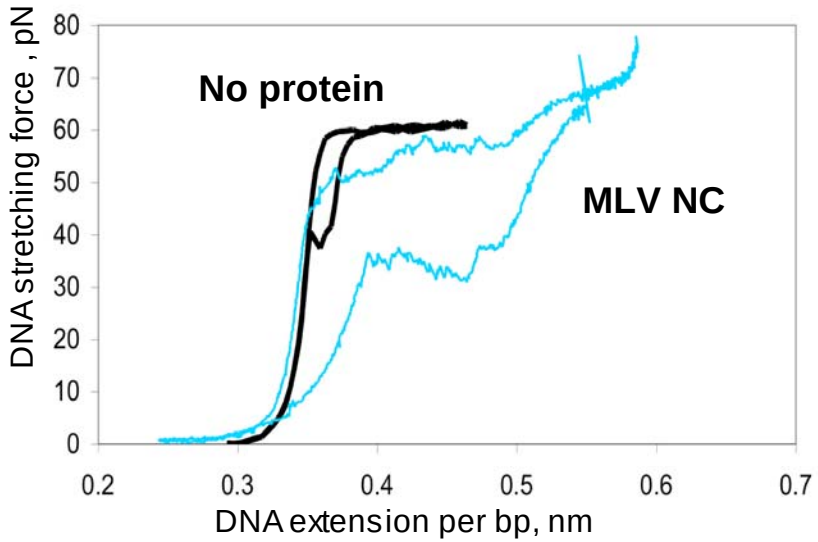
85 aa, 11 basic residues
theoretical pI = 6.96

RSV NC promotes nucleic acid aggregation but not duplex destabilization



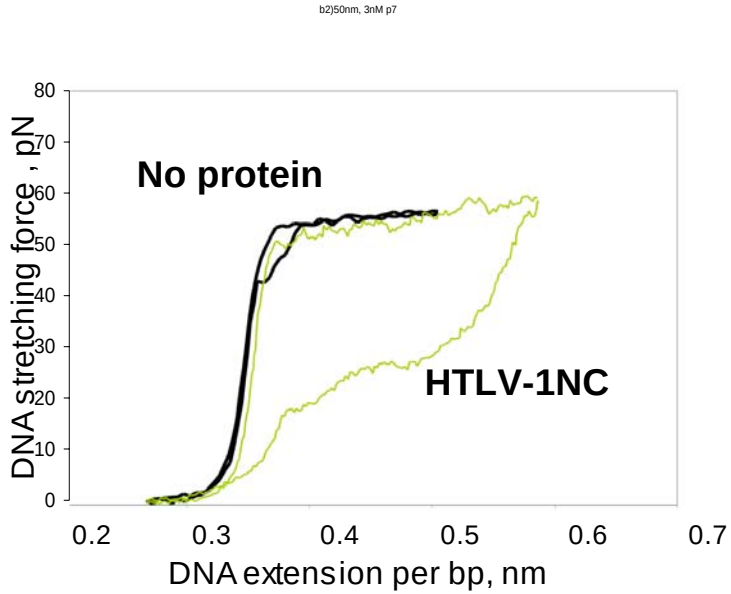
- **Increased melting force**
- **Small hysteresis**

MLV NC is a good duplex destabilizer but an inefficient aggregating agent



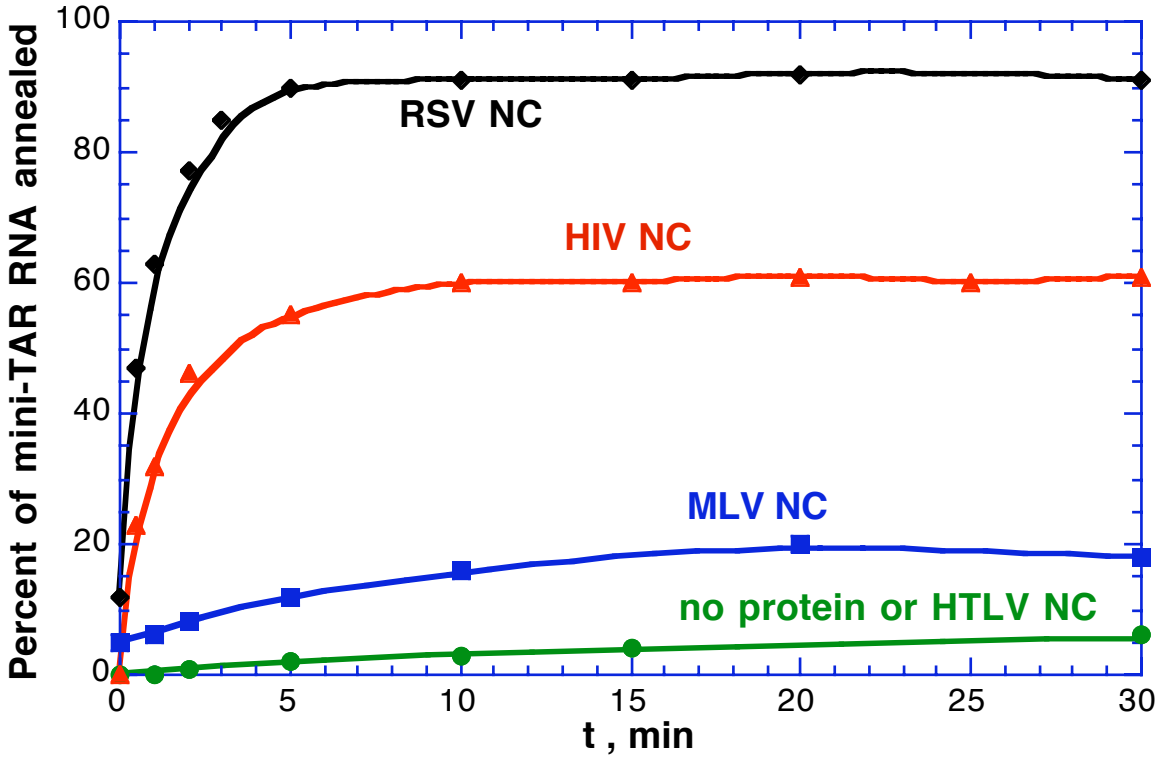
- **decreased melting force**
- **large rugged hysteresis**

HTLV-1 NC behaves like a single-stranded binding protein and is a poor annealing agent



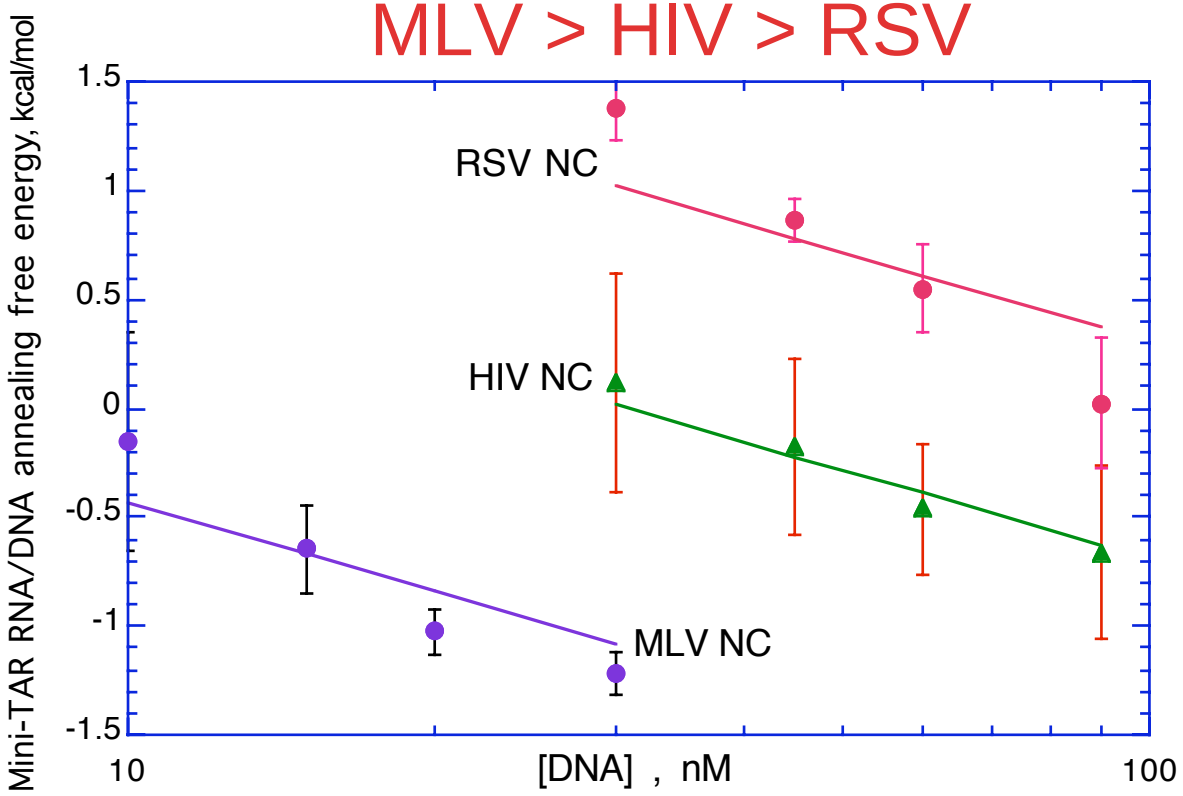
- large hysteresis indicates that HTLV-1 NC dissociates from single-stranded DNA very slowly
- Lack of effect on melting force is due to slow protein binding

Mini-TAR RNA/DNA annealing assays also reveal significant differences in chaperone activities of retroviral NC proteins



Conditions: 5 nM mini-TAR RNA+30 nM mini-TAR DNA
at 37^0C in 20 mM NaCl and 1 mM Mg
1 μM RSV NC; 1 μM HIV NC; 3 μM MLV NC; 5 μM HTLV NC

Strength of duplex destabilization activity of NC proteins: MLV > HIV > RSV

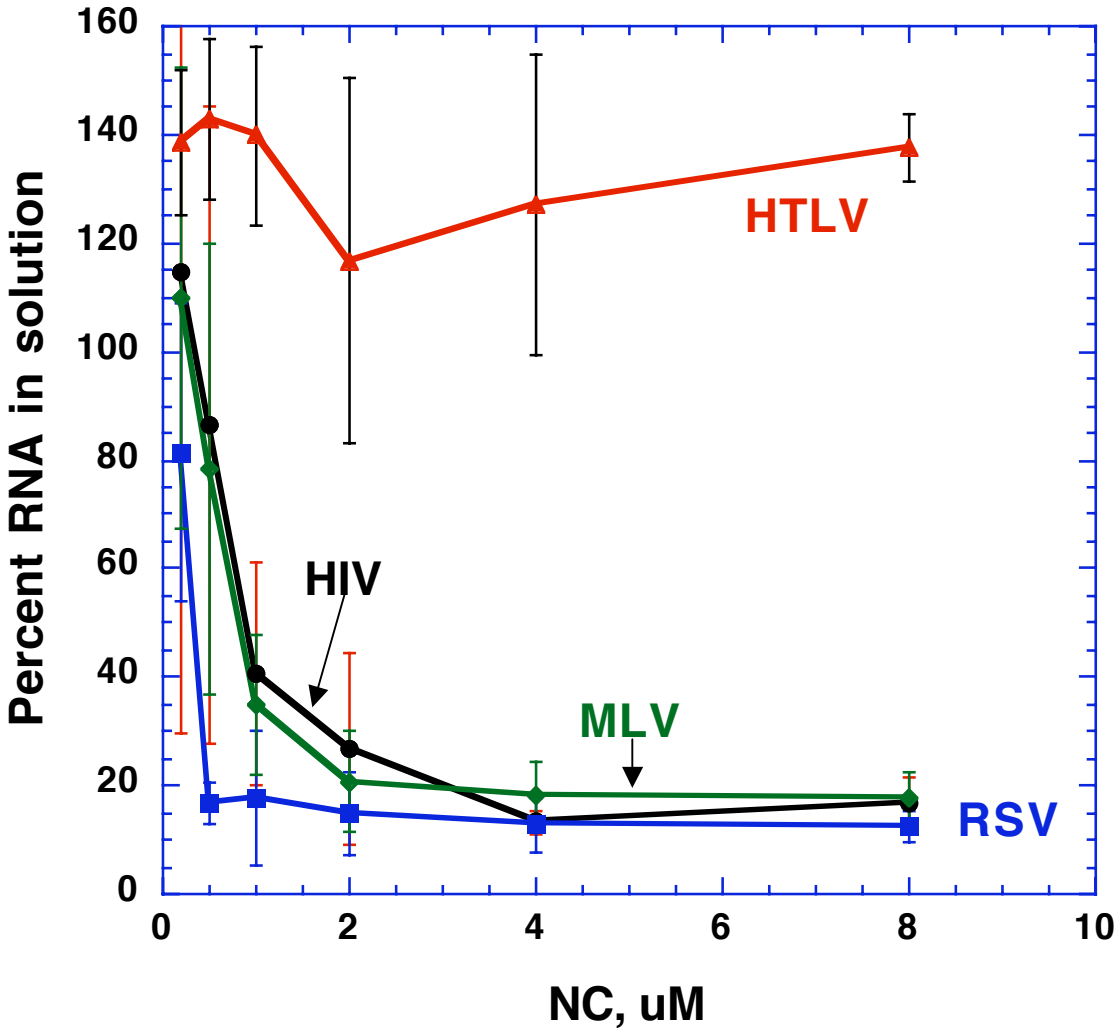


$$\Delta G[DNA] = -RT \ln\left(\frac{\text{annealed fraction}}{\text{not annealed fraction}}\right) = -RT \ln\left(\frac{[DNA]}{[DNA]_o}\right) + \Delta G_o$$

Annealing free energy at high strand concentration, ΔG_0:

With MLV NC	- 9.2 kcal/mol
With HIV NC	-10.5 kcal/mol
With RSV NC	-11.3 kcal/mol

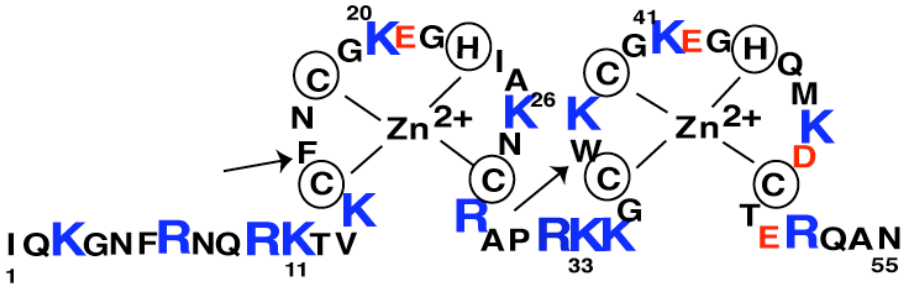
WT HTLV NC is a very poor aggregating agent



RSV > HIV ~ MLV >> HTLV

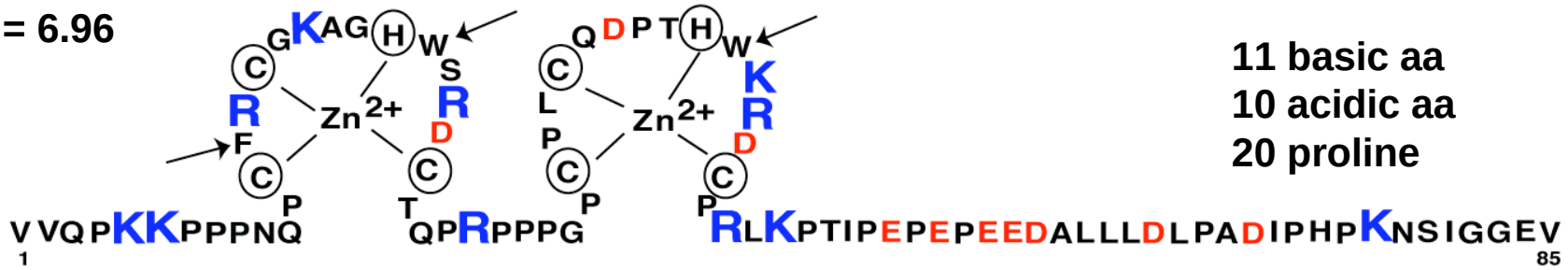
HTLV-1 NC truncation mutant

HIV-1 NC
pI = 9.93



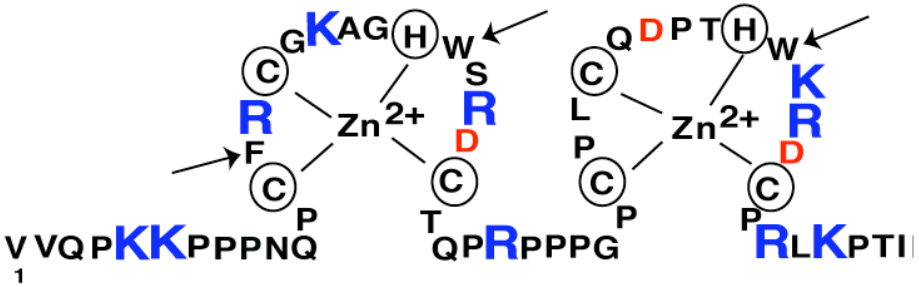
14 basic aa
4 acidic aa

HTLV-1 NC
pI = 6.96



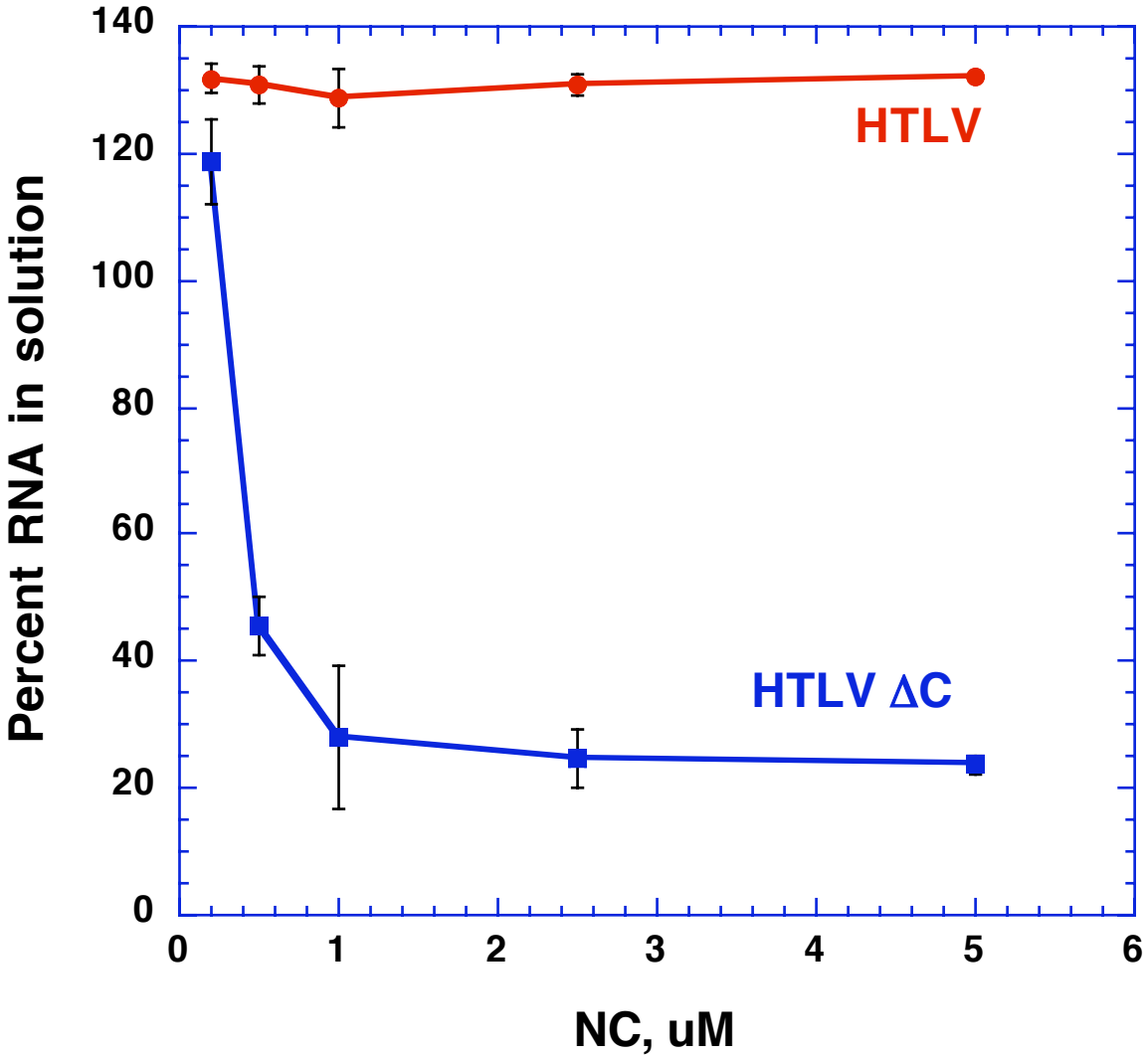
11 basic aa
10 acidic aa
20 proline

HTLV-1 NC-ΔC-term
pI = 9.55

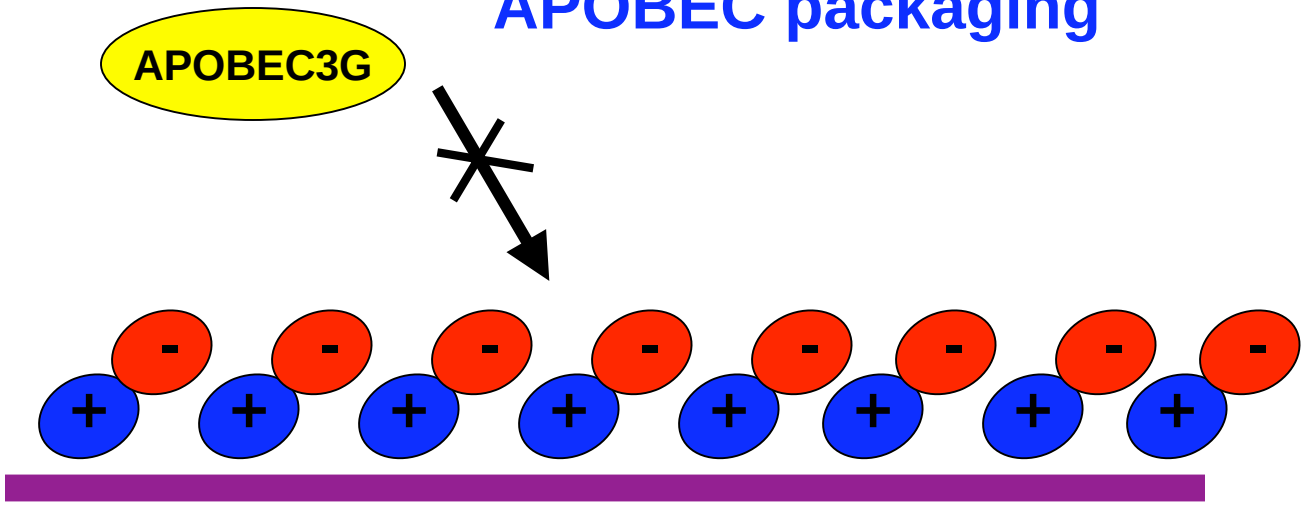


10 basic aa
3 acidic aa
14 proline

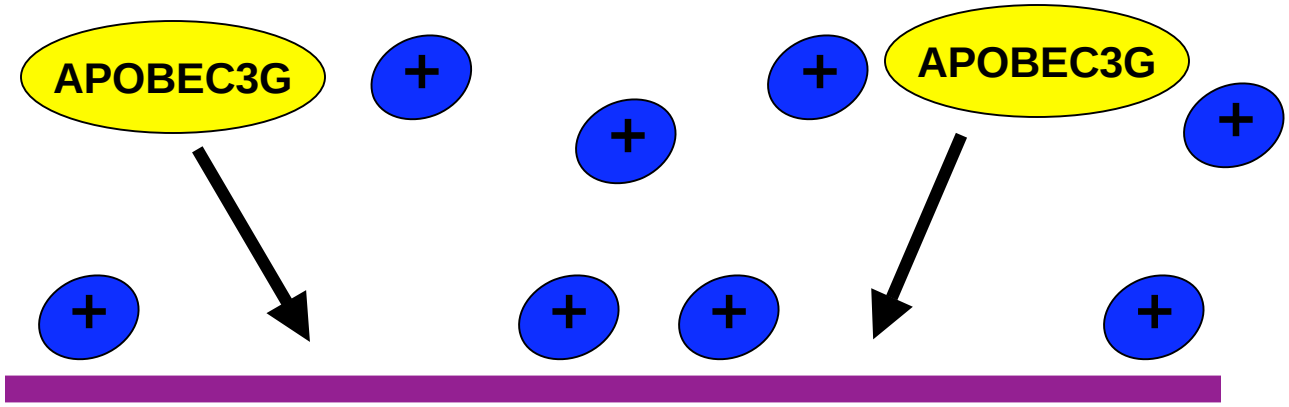
Aggregation activity is restored upon truncation of acidic C-terminal domain



Results may be consistent with RNA-bridge model for APOBEC packaging



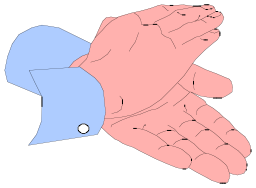
HTLV NC protects RNA from APOBEC



ΔC-NC can dissociate and APOBEC can bind RNA and be packaged via Gag

Conclusions

- **Three main properties are essential for nucleic acid chaperone activity of the protein:**
  - **Ability to weakly destabilize all nucleic acid secondary structures**
  - **Ability to aggregate nucleic acids via none specific electrostatic interactions**
  - **Fast kinetics of protein/nucleic acid association and dissociation**
- **Chaperone activities of retroviral NC proteins vary widely:**
  - **Relative strength of duplex destabilization activity of NC proteins: MLV > HIV > RSV > HTLV**
  - **Kinetics of protein/nucleic acid interaction becomes slower as: HIV>RSV>MLV>HTLV**
  - **HTLV NC is a poor chaperone protein with properties similar to a slow single-stranded binding protein. C terminal deletion makes this protein a much faster ss DNA binder.**
- **All NC proteins except for HTLV NC aggregate nucleic acids, but slower protein also immobilize strands, thus interfering with their annealing.**
- **HIV-1 NC is an most effective nucleic acid chaperone, due to its ability to facilitate both nucleic acid aggregation and duplex destabilization, and by being the "fastest" nucleic acid binder.**



Acknowledgements

Present NC Group (Minnesota):

Prof. Karin Musier-Forsyth
Ms. My-Nuong Vo
Dr. Besik Kankia
Dr. Kristen Stewart
Mr. Mithun Mitra

Single Molecule Stretching:

Prof. Mark Williams (Northeastern)
Dr. Margareta Cruceanu

NC Protein:

Dr. Rob Gorelick (NCI, Frederick)
Prof. George Barany (Minnesota)
Dr. Dan Mullen

Funding:

NIH