Structural Virology

Lecture 8

Pavel Plevka

Retroviridae

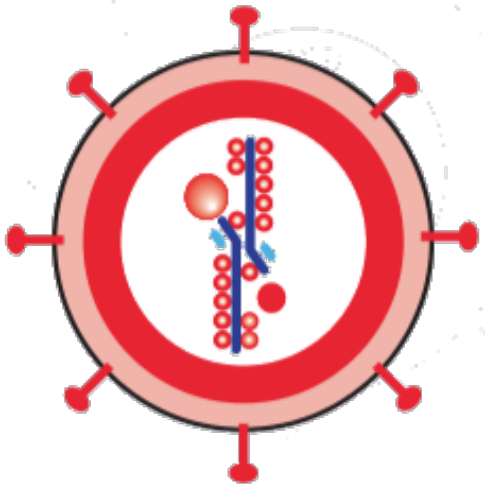
***retro* (Latin) = backwards**

Hosts: mammals
birds
other vertebrate animals

Diseases: immunodeficiency diseases
leukaemias
solid tumours

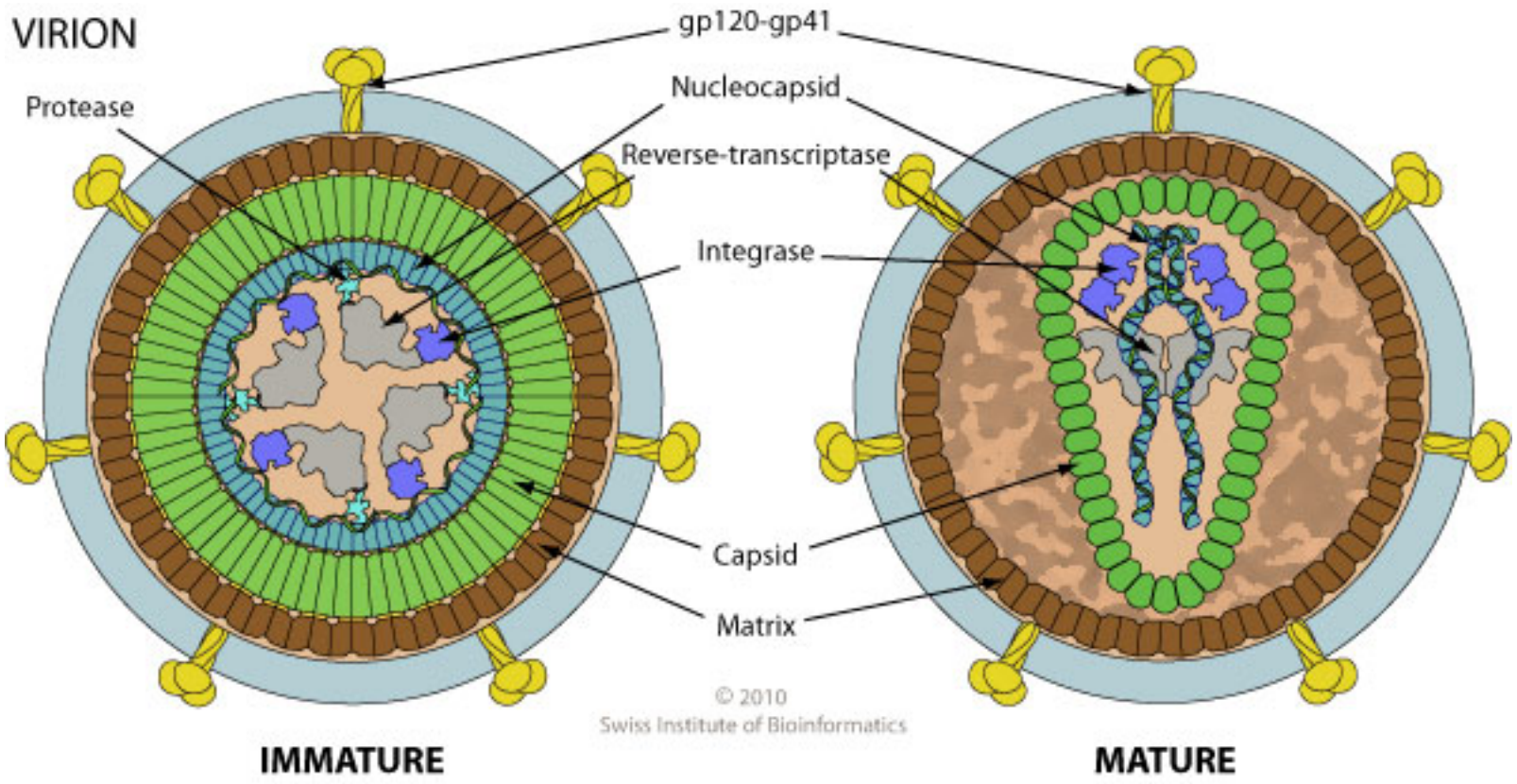
Virion

- Enveloped
- 80–110 nm diameter
- Genome: single-stranded RNA
  plus polarity
  9–10 kb
- Contains reverse transcriptase

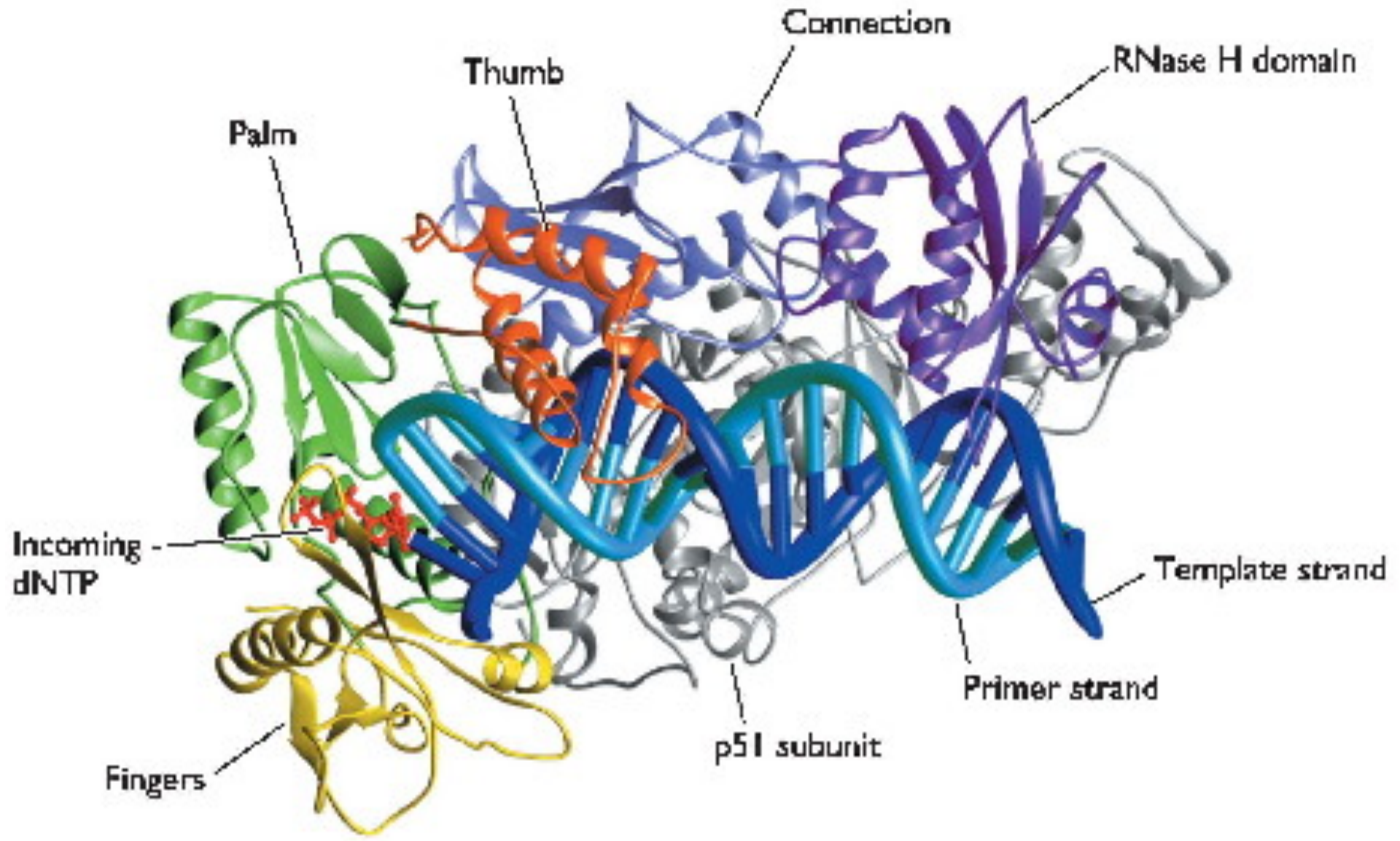


Reverse transcriptases

Murine leukemia virus (MLV)



Reverse transcriptase



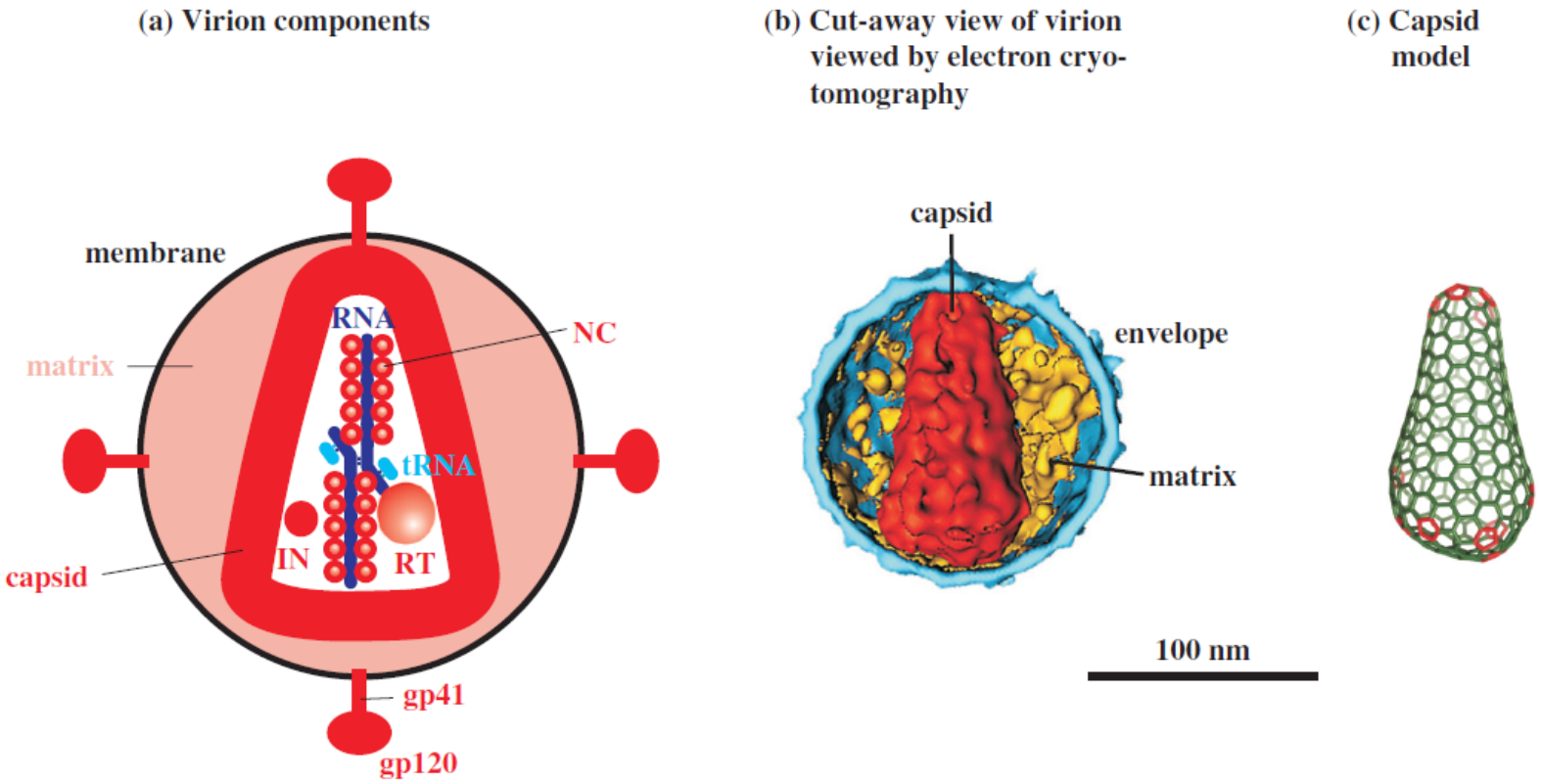
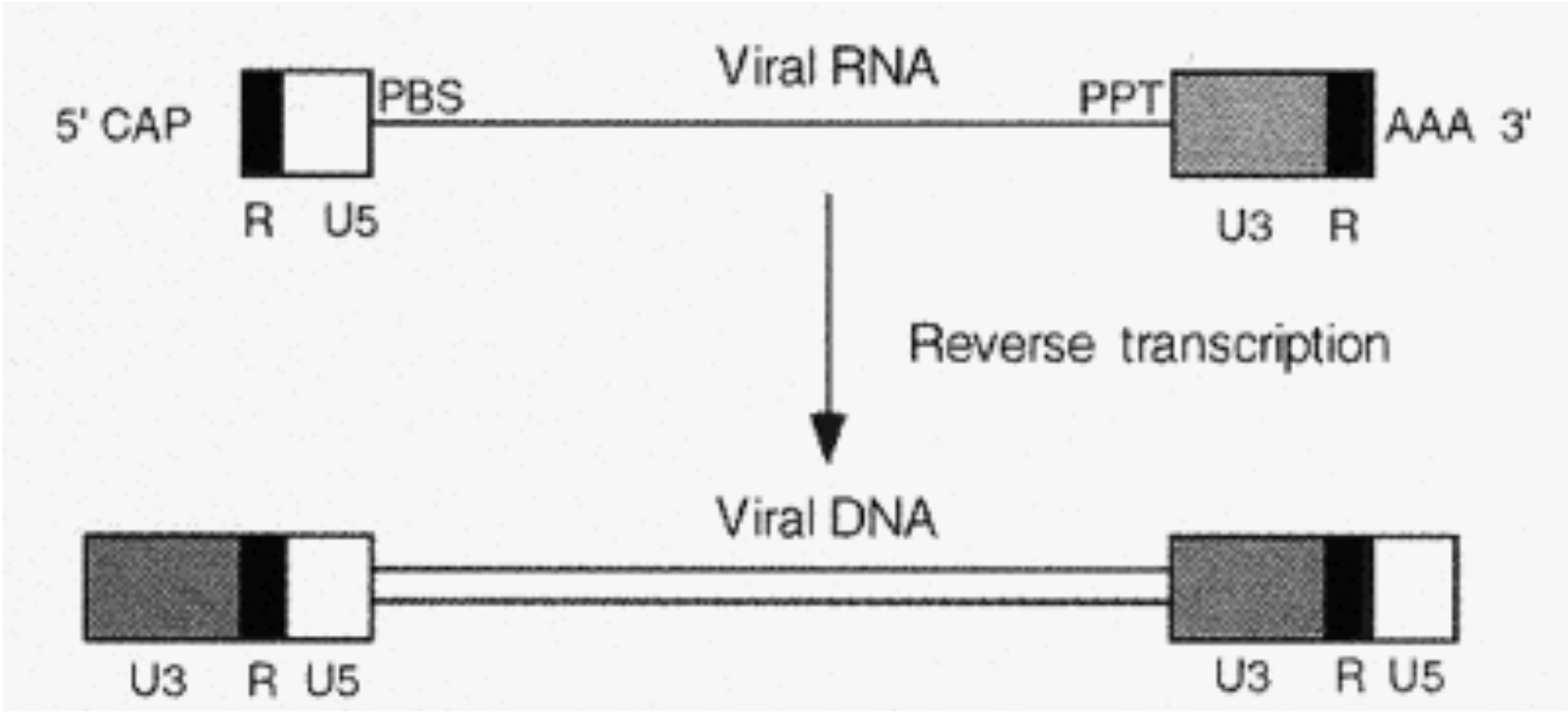
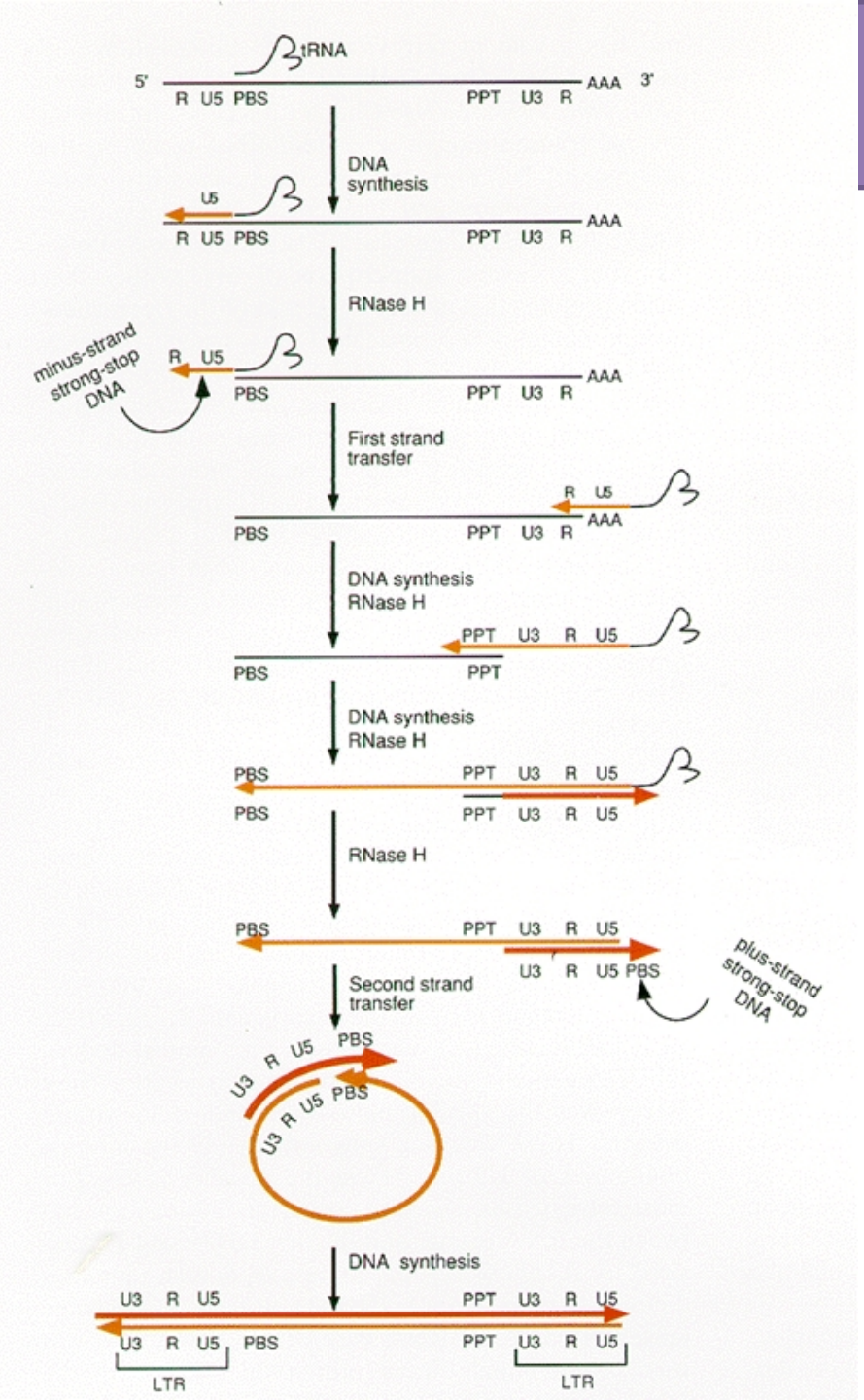


Figure 18.1 **HIV virion.** (a) Virion components. IN: integrase. NC: nucleocapsid protein. RT: reverse transcriptase. The TM and SU glycoproteins indicated are those of HIV-1 (gp41 and gp120). (c) Capsid model, showing protein hexamers in green and pentamers in red.

Sources: (b) Grünewald and Cyrklaff (2006) *Current Opinion in Microbiology*, 9, 437. (c) Ganser-Pornillos, Yeager, and Sundquist (2008) *Current Opinion in Structural Biology*, 18, 203. (b) and (c) reproduced by permission of Elsevier Limited and the authors

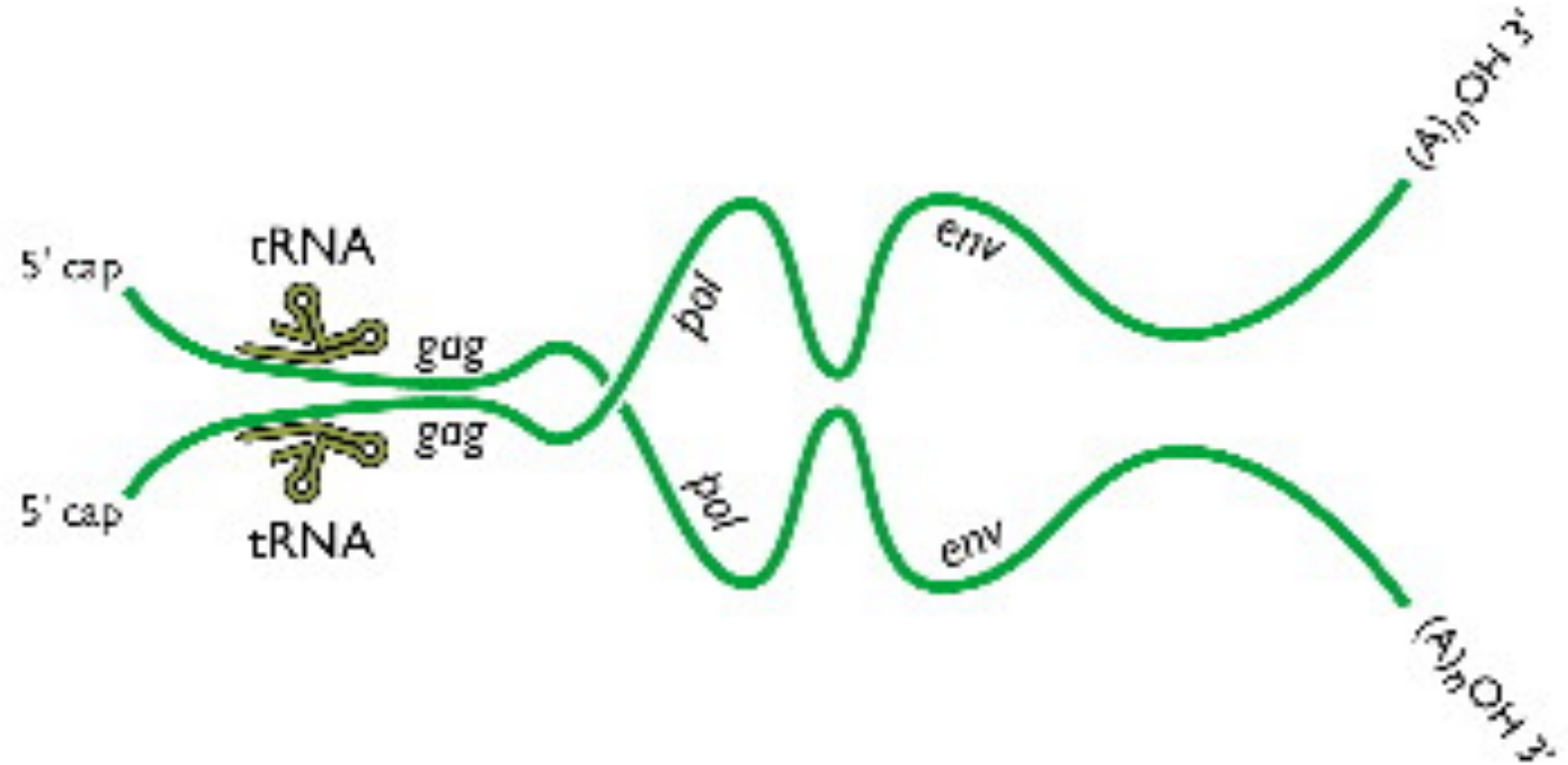
Retrovirus RNA and DNA forms of genome

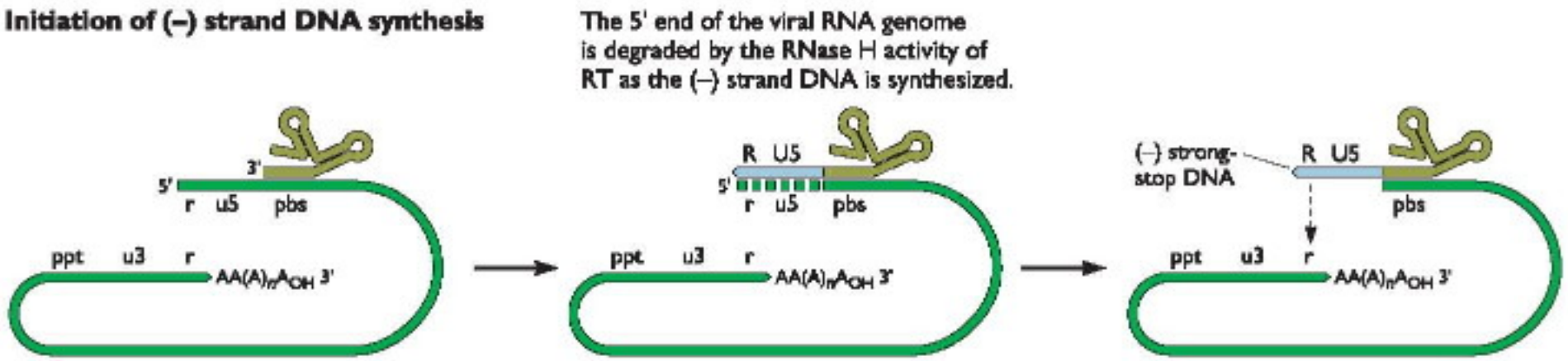


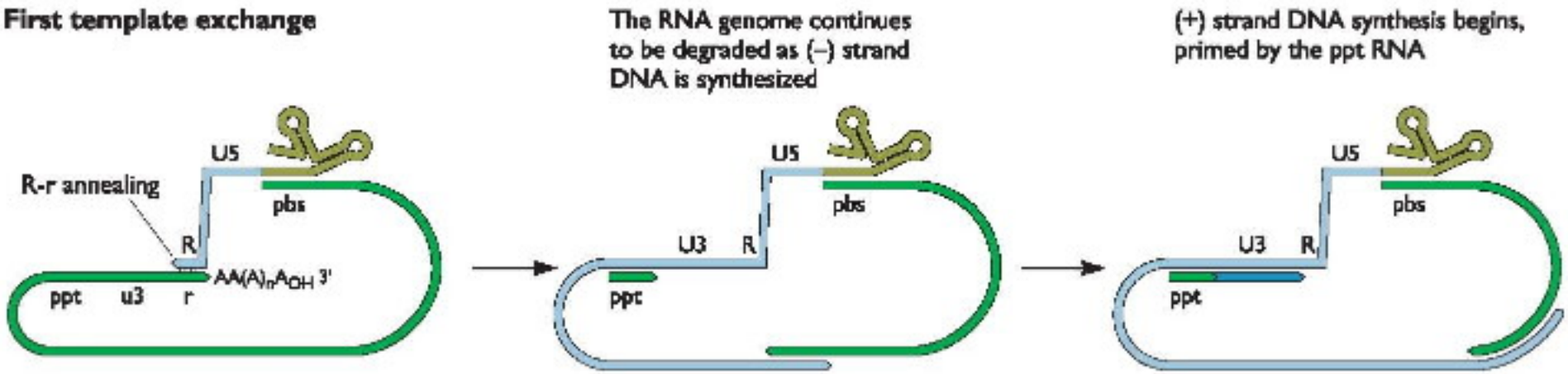


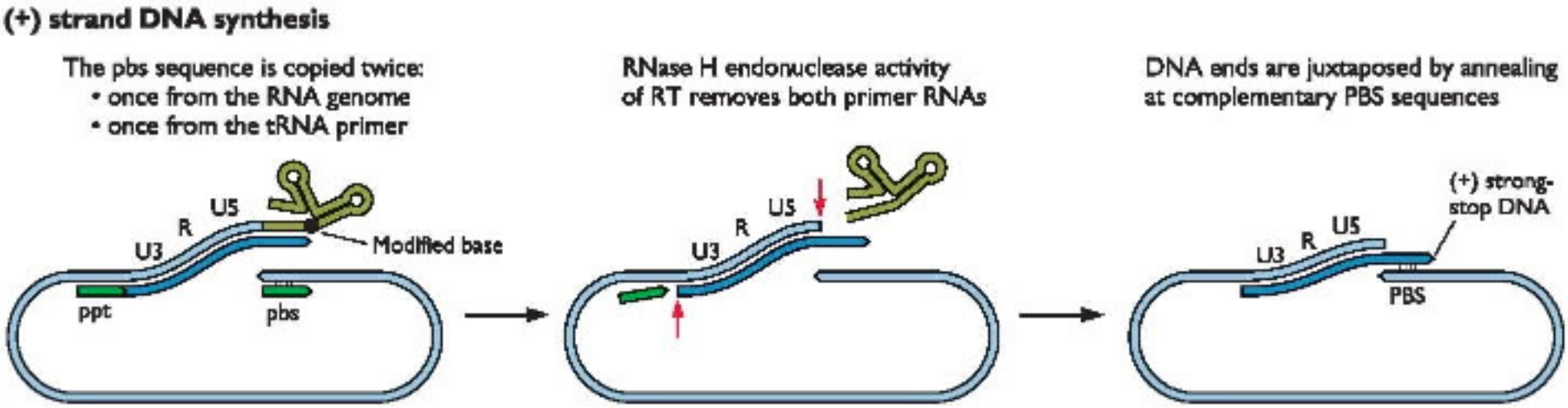
Reverse transcription of retrovirus genome

Genomes inside virion

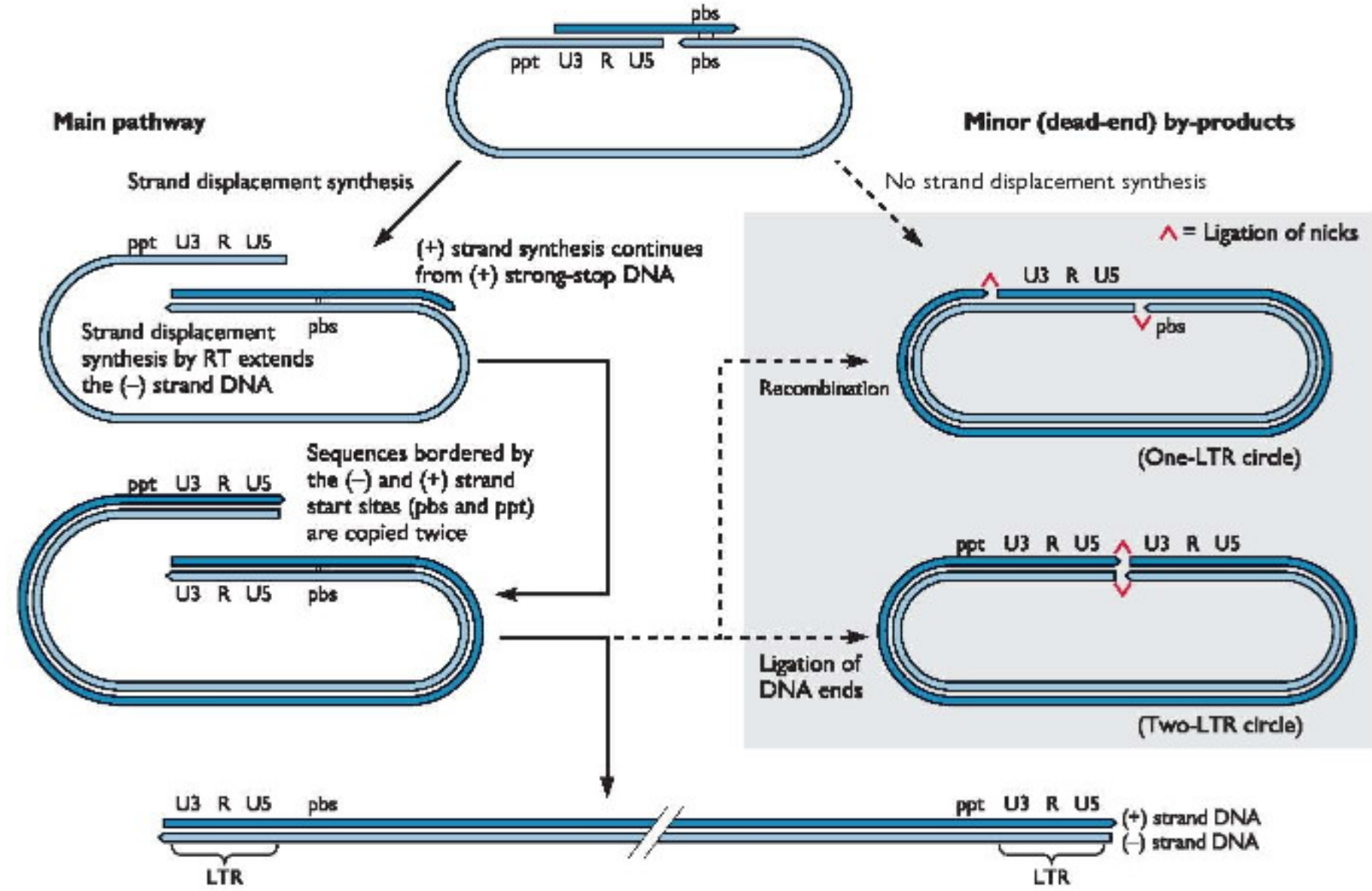






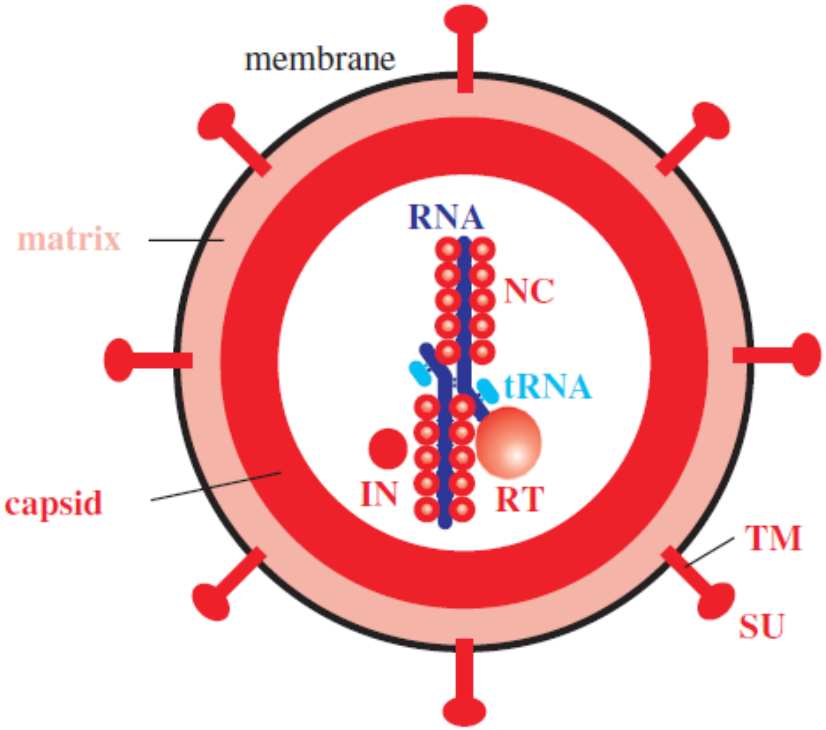


Second template exchange is facilitated by annealing of PBS sequences

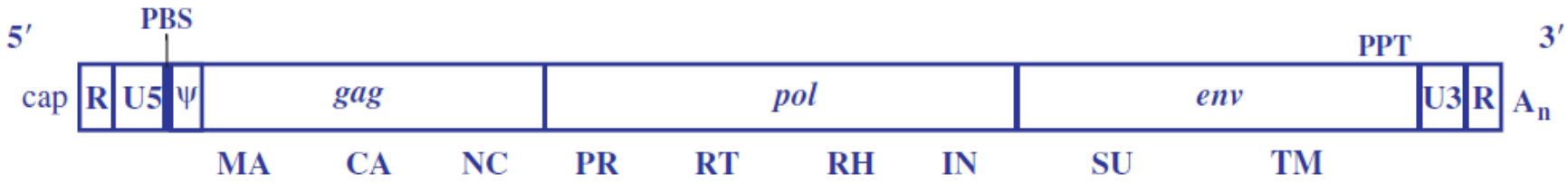


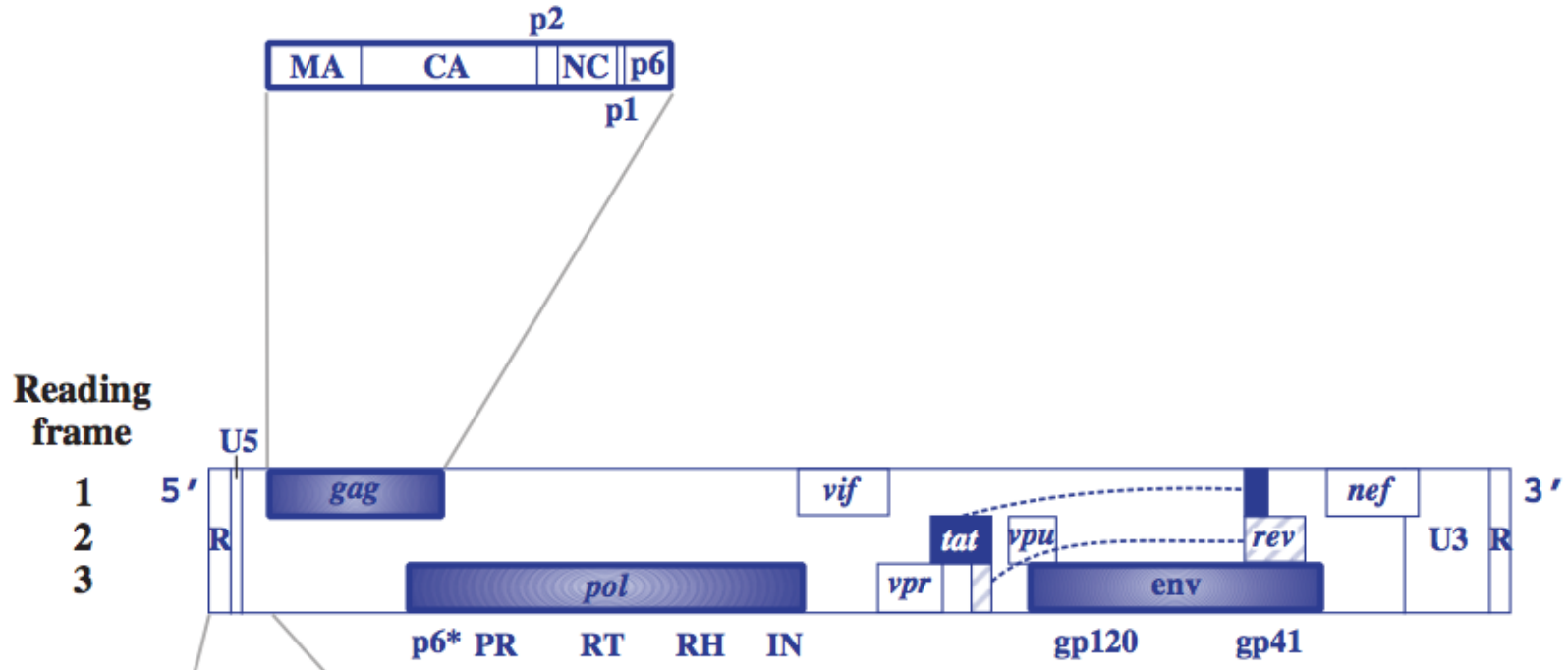
Virion and genome organization

Virion components

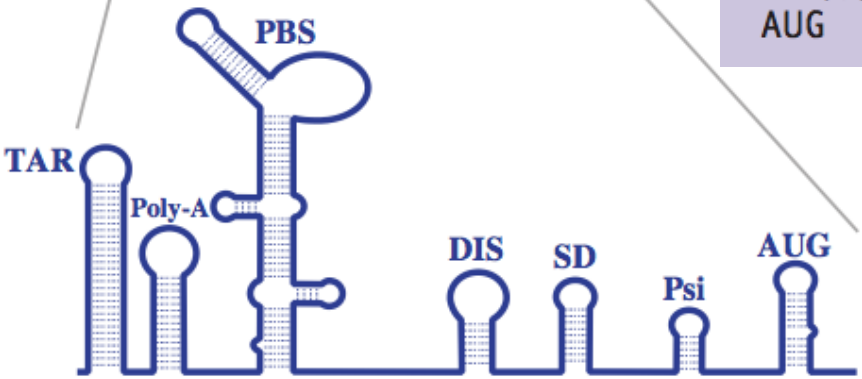


Genome organization and gene products

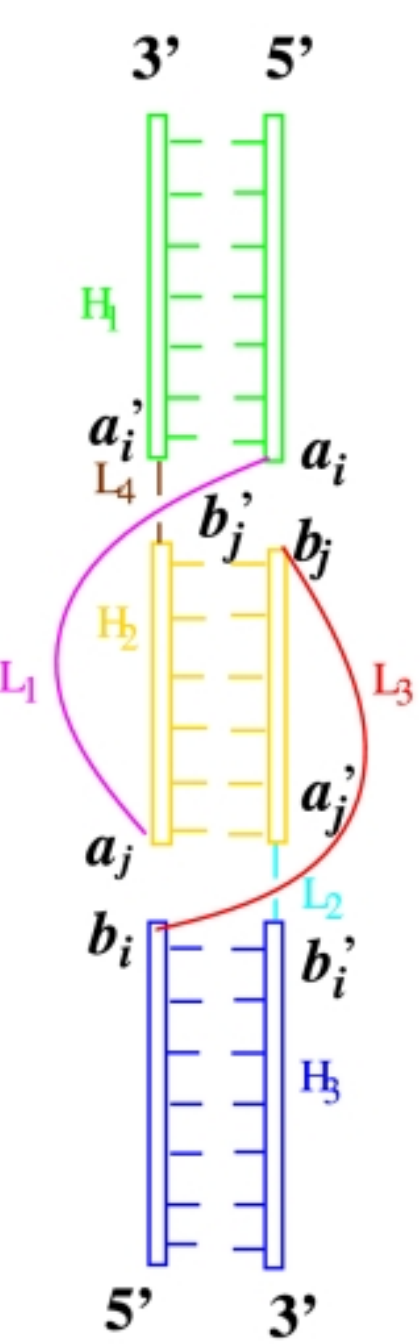




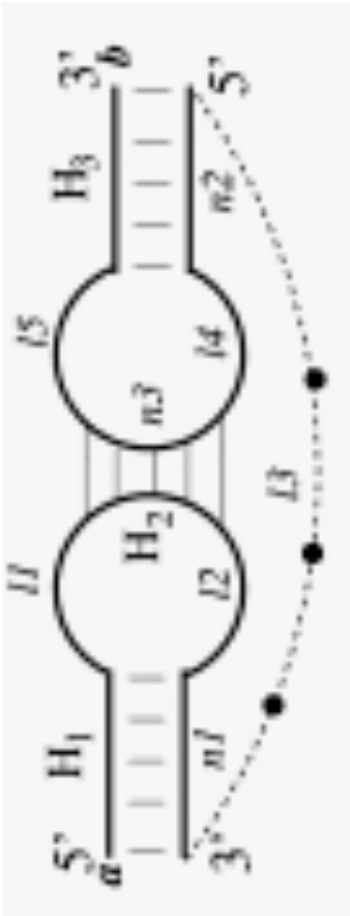
TAR	trans-acting response element
Poly-A	polyadenylation signal
PBS	primer-binding site
DIS	dimerization initiation site (involved in formation of kissing loop complex)
SD	splice donor site
Psi (ψ)	main part of the packaging signal
AUG	start codon of the *gag* gene



Main genes	*gag*	*g*roup-specific *a*nti*g*en (encodes matrix, capsid, p2, nucleocapsid, p1 and p6)
	pol	*pol*ymerase (encodes p6*, protease, reverse transcriptase, RNase H, integrase)
	env	*env*elope
Auxiliary genes	*nef*	*ne*gative regulatory *f*actor
	rev	*r*egulator of *e*xpression of *v*irion proteins
	tat	*t*rans*a*ctivator of *t*ranscription
	vif	*v*irion *i*nfectivity *f*actor
	vpr	*v*iral *p*rotein *R*
	vpu	*v*iral *p*rotein *U*
Non-coding sequences	R	repeat sequence
	U3	unique sequence at 3′ end of genome
	U5	unique sequence at 5′ end of genome
Domains at the 5′ end of the genome	TAR	trans-acting response element
	Poly-A	polyadenylation signal
	PBS	primer-binding site
	DIS	dimerization initiation site (involved in formation of kissing loop complex)
	SD	splice donor site
	Psi (ψ)	main part of the packaging signal
	AUG	start codon of the *gag* gene



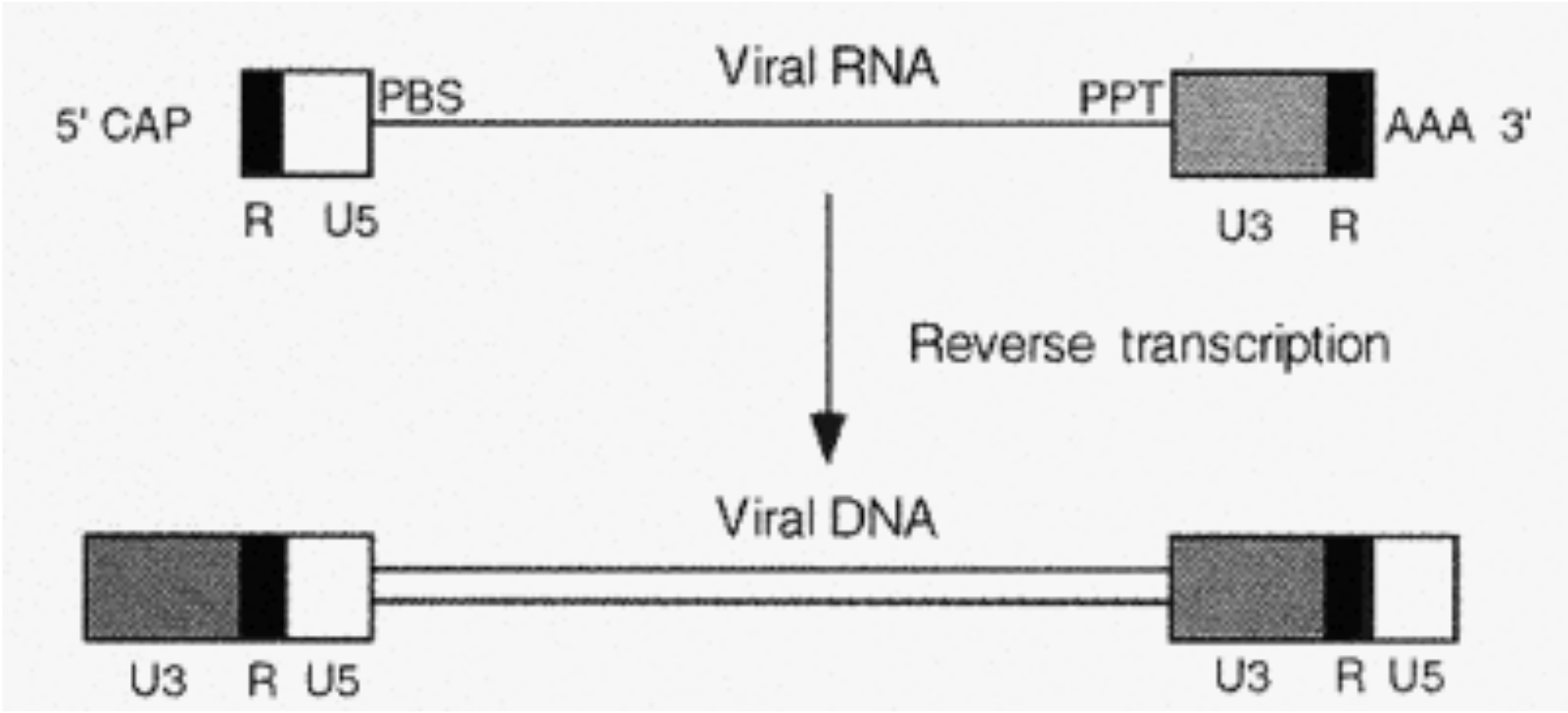
Kissing RNA loops



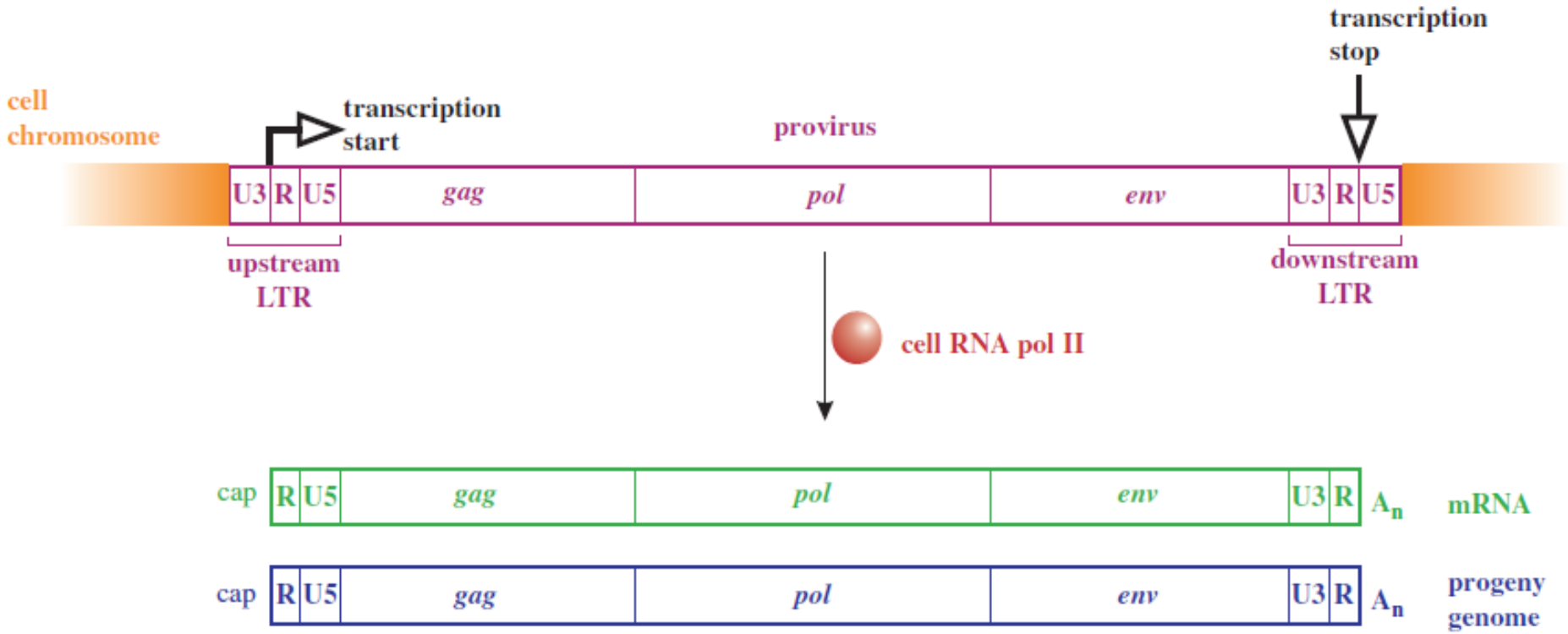
The HIV trans-activation response (TAR) element is an RNA element which is known to be required for the trans-activation of the viral promoter and for virus replication. The TAR hairpin is a dynamic structure[1] that acts as a binding site for the Tat protein, and this interaction stimulates the activity of the long terminal repeat promoter.[2]

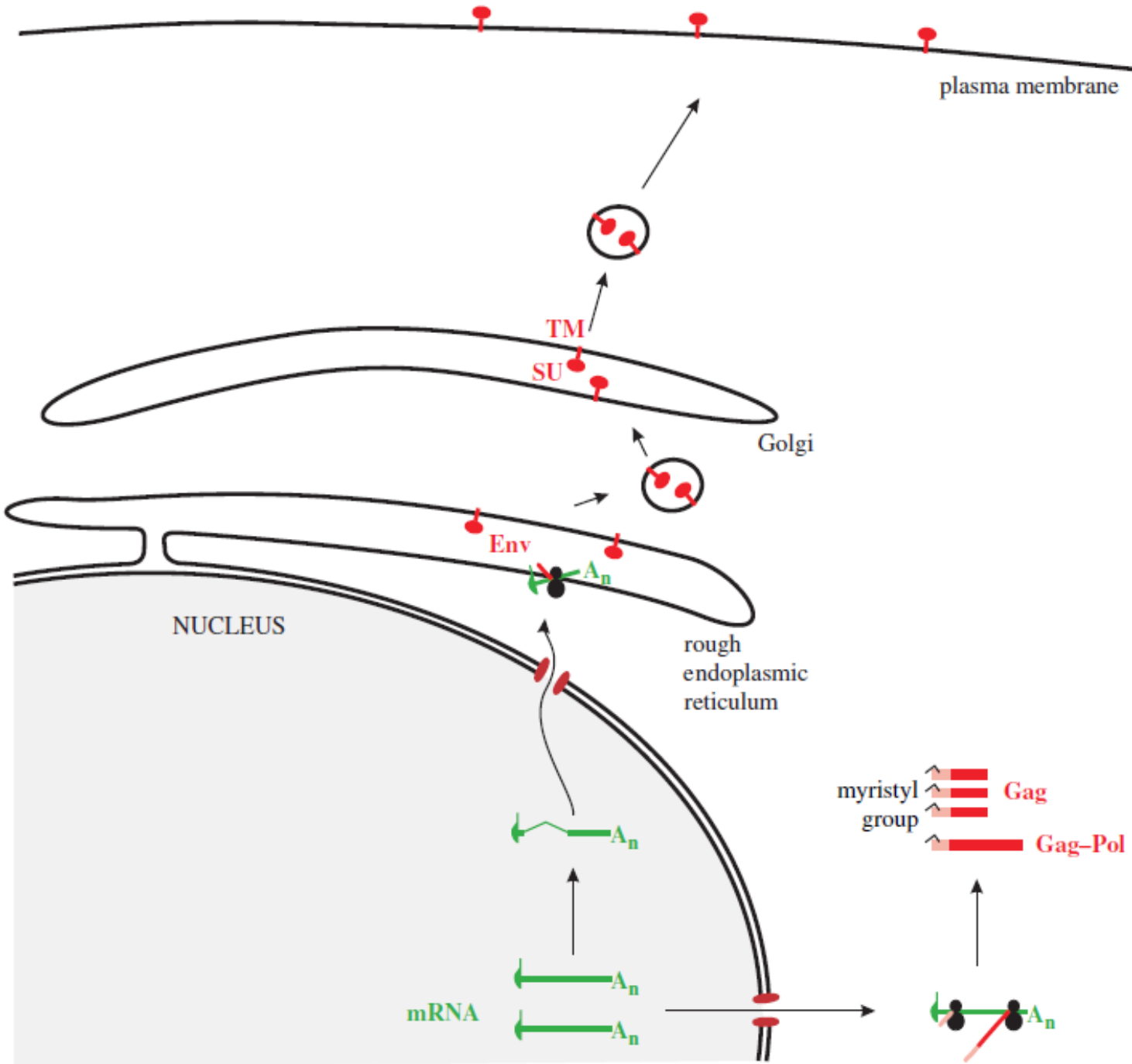
Further analysis has shown that TAR is a pre-microRNA that produces mature microRNAs from both strands of the TAR stem-loop.[3] These miRNAs are thought to prevent infected cells from undergoing apoptosis by downregulating the genes ERCC1 and IER3.[4]

Retrovirus transcription start and terminator?



Retrovirus transcription




plasma membrane
TM
SU
Golgi
Env
An
NUCLEUS
rough
endoplasmic
reticulum
myristyl
group
Gag
Gag–Pol
An
An
mRNA
An
An

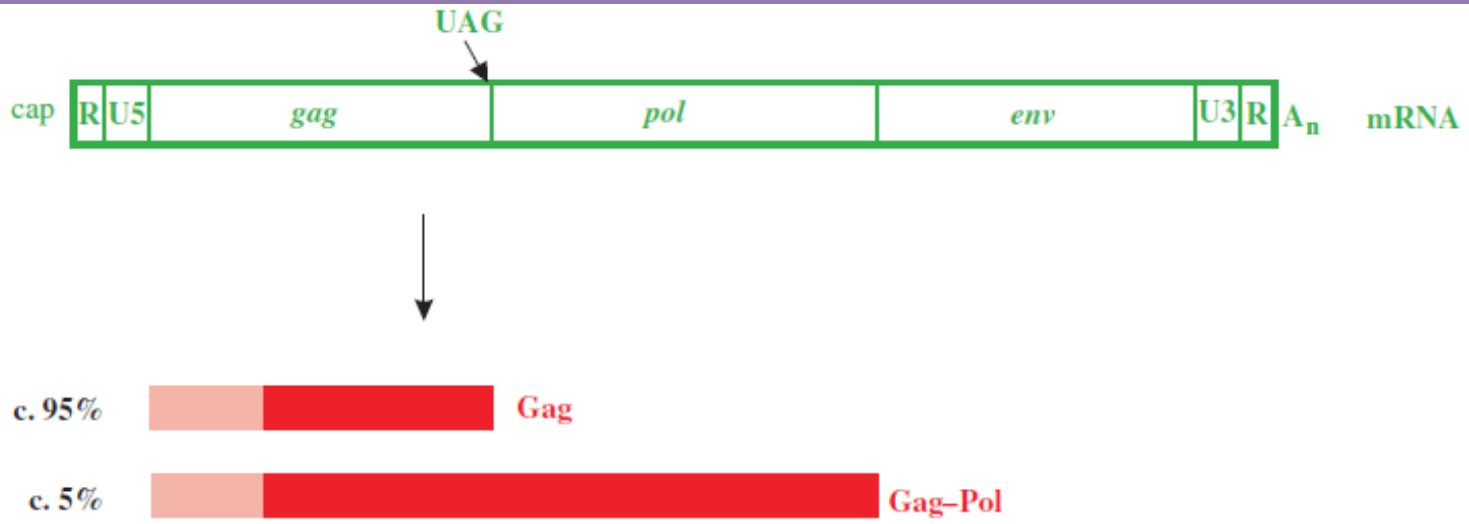


Figure 17.7 Translation of Gag–Pol by reading through a stop codon (UAG). Approximately 5% of ribosomes incorporate an amino acid at the *gag* stop codon and translate *pol* too.

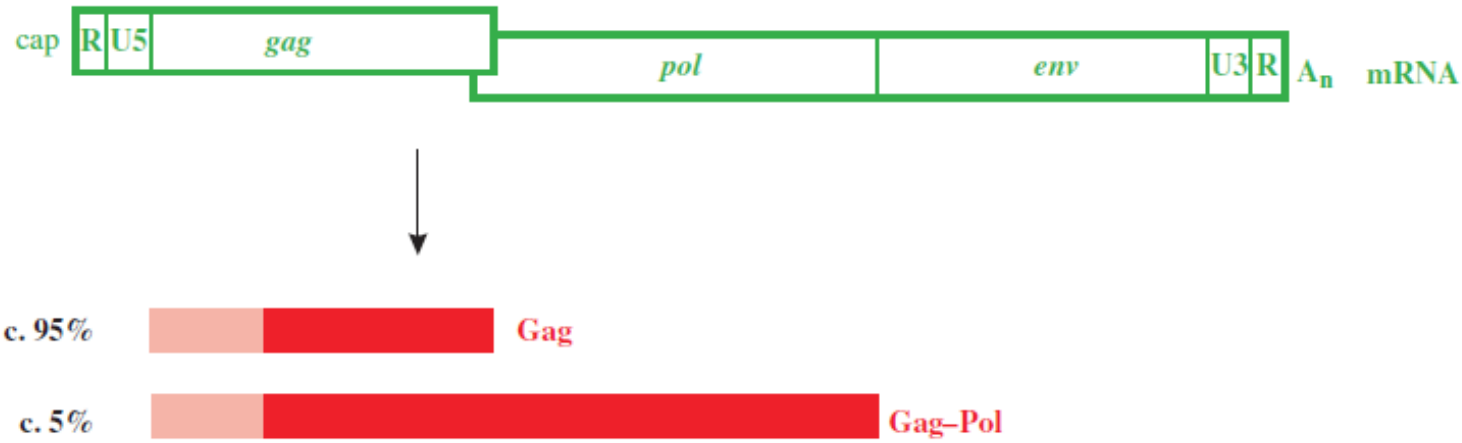


Figure 17.8 Translation of Gag–Pol by ribosomal frameshifting. Approximately 5% of ribosomes shift into a different reading frame before the *gag* stop codon and translate *pol* too.

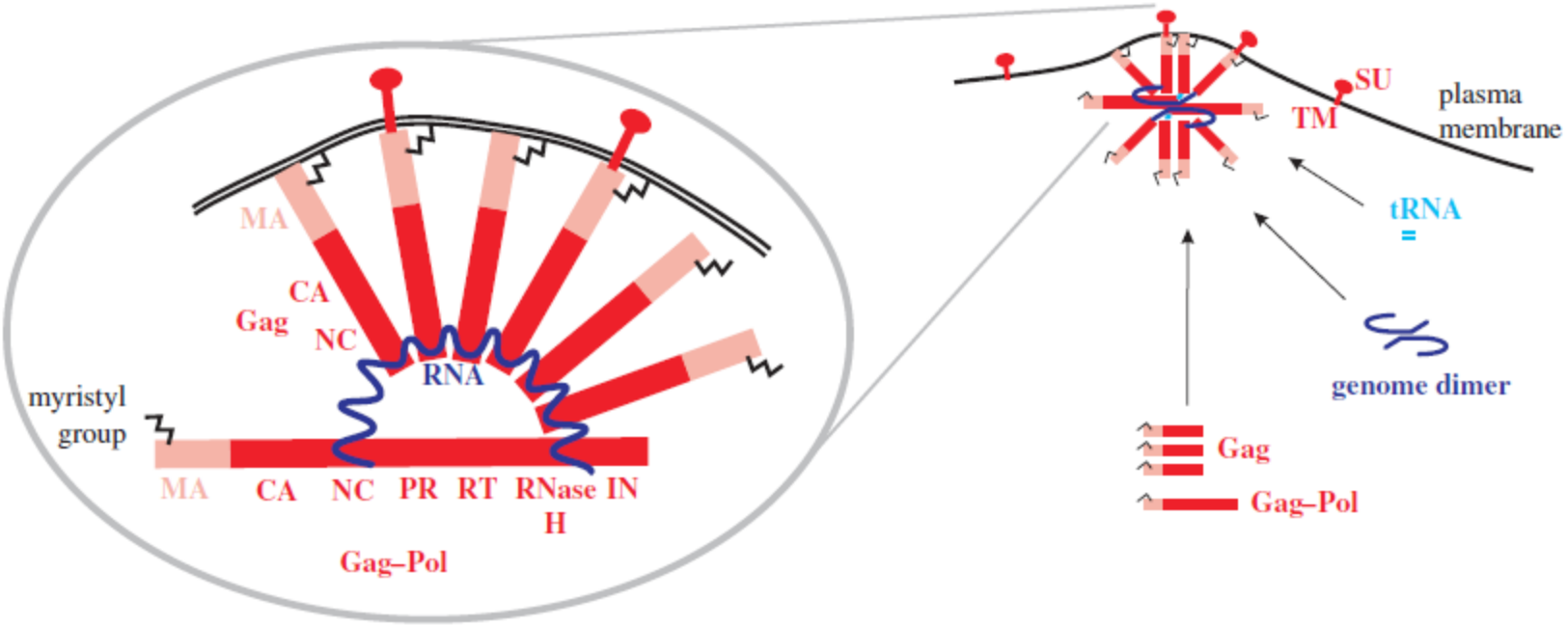


Figure 17.9 **Retrovirus assembly – early stages.** A genome dimer associates with cell tRNAs and with Gag and Gag–Pol proteins. The domains of Gag and Gag–Pol are indicated in the inset. The order of the Gag domains MA–CA–NC is the same as the exterior-to-interior order of the proteins in the virion (Figure 17.1).

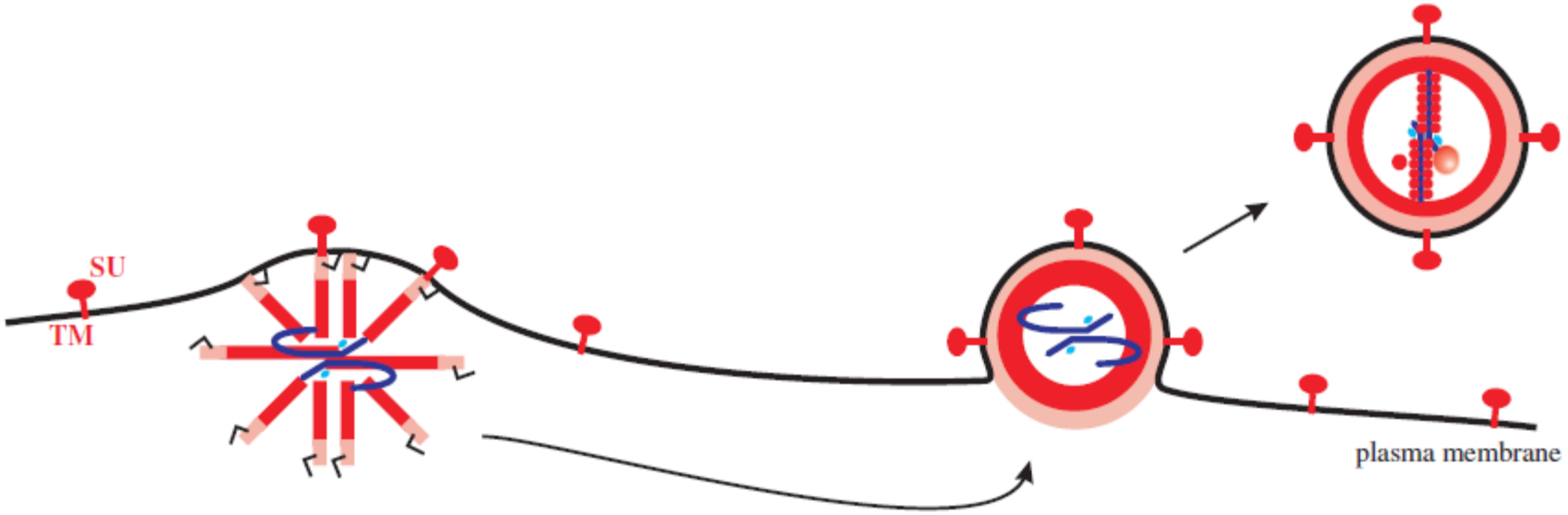
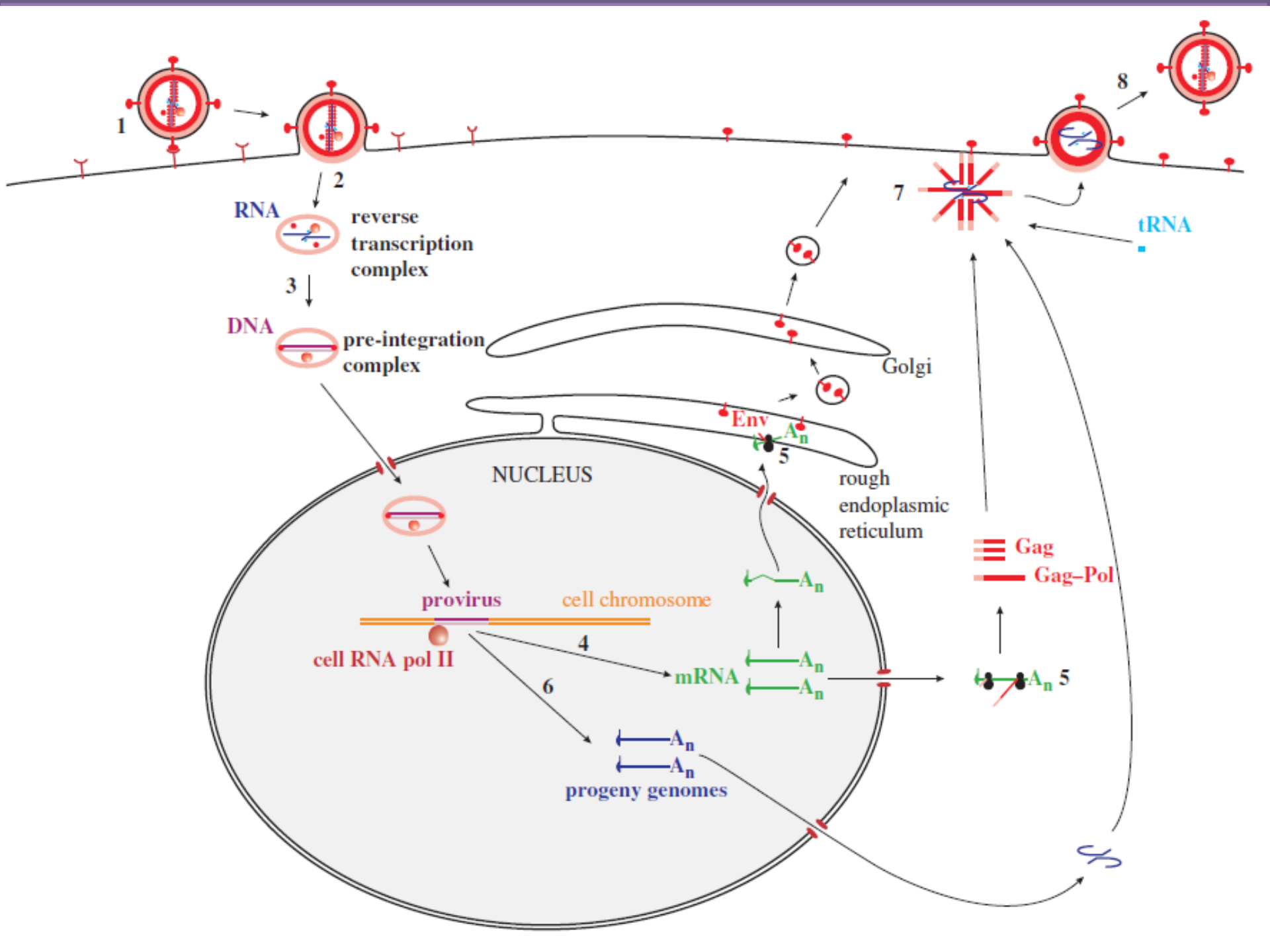
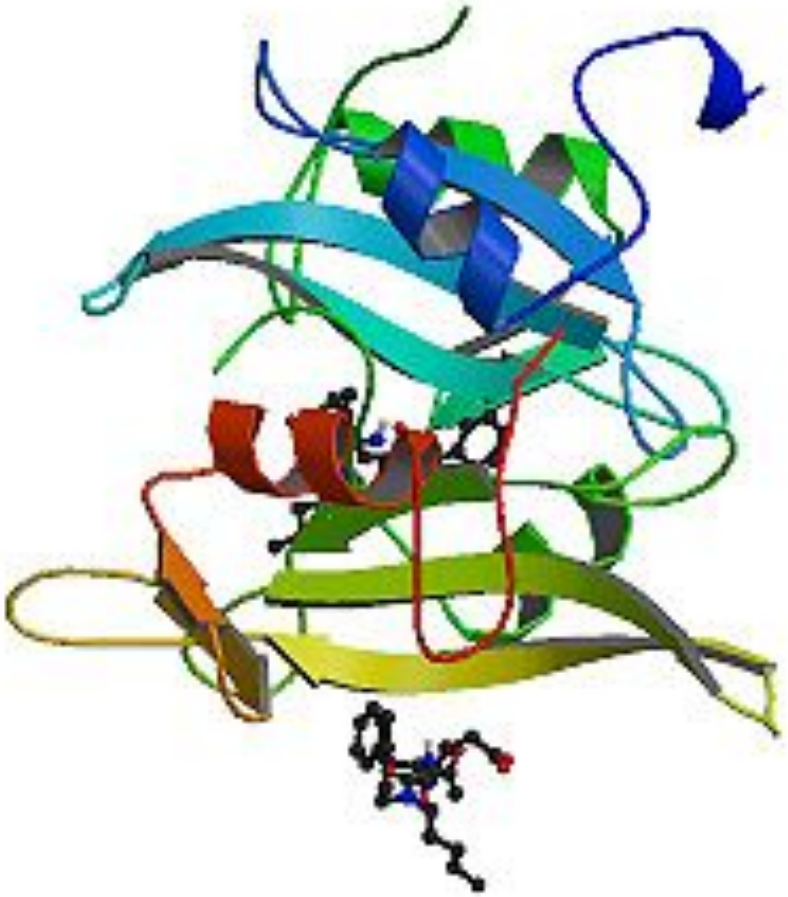


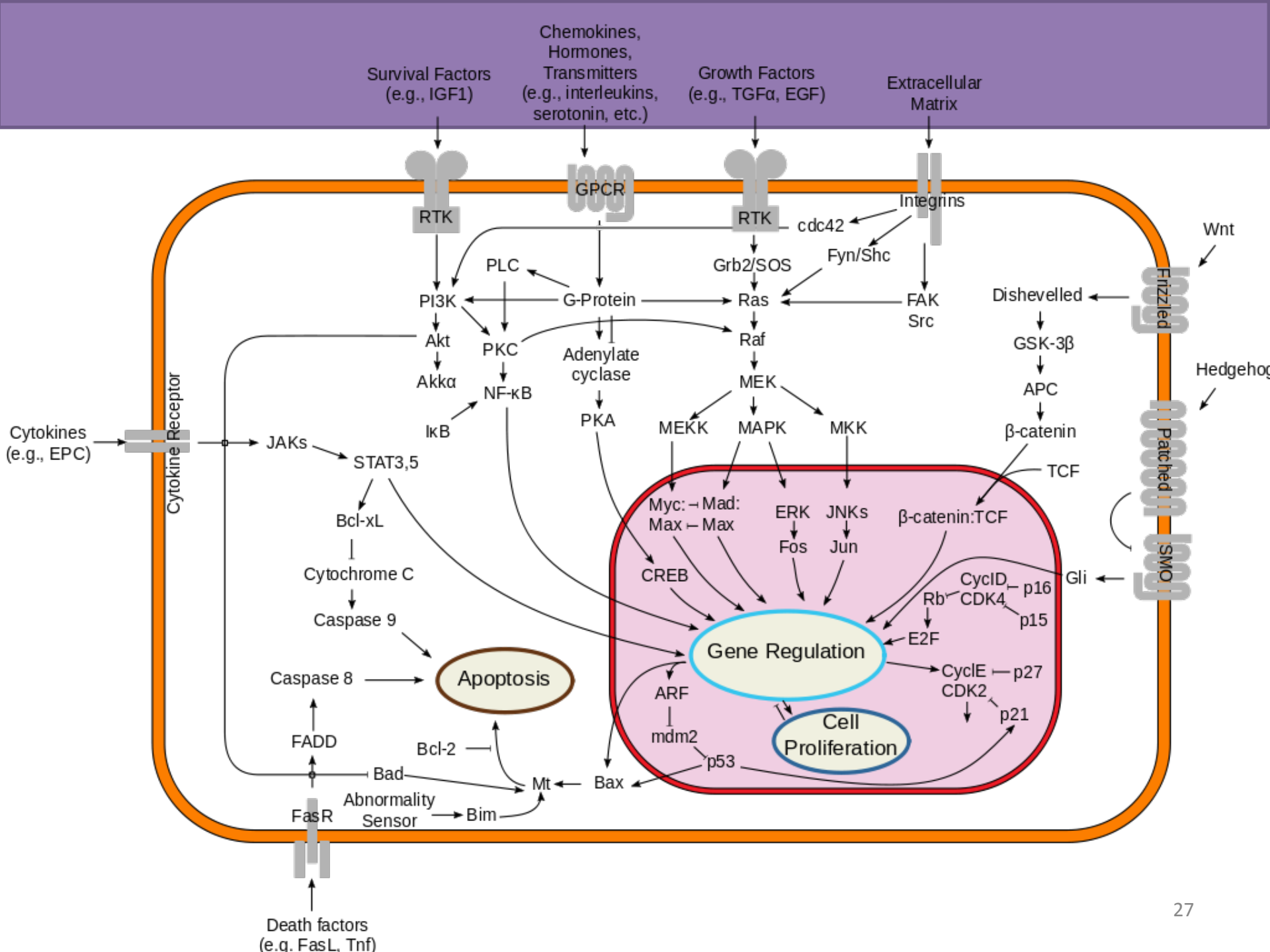
Figure 17.10 **Retrovirus assembly – late stages.** The envelope is acquired by budding from the plasma membrane. During and after budding Gag and Gag–Pol are cleaved to form the virion proteins.

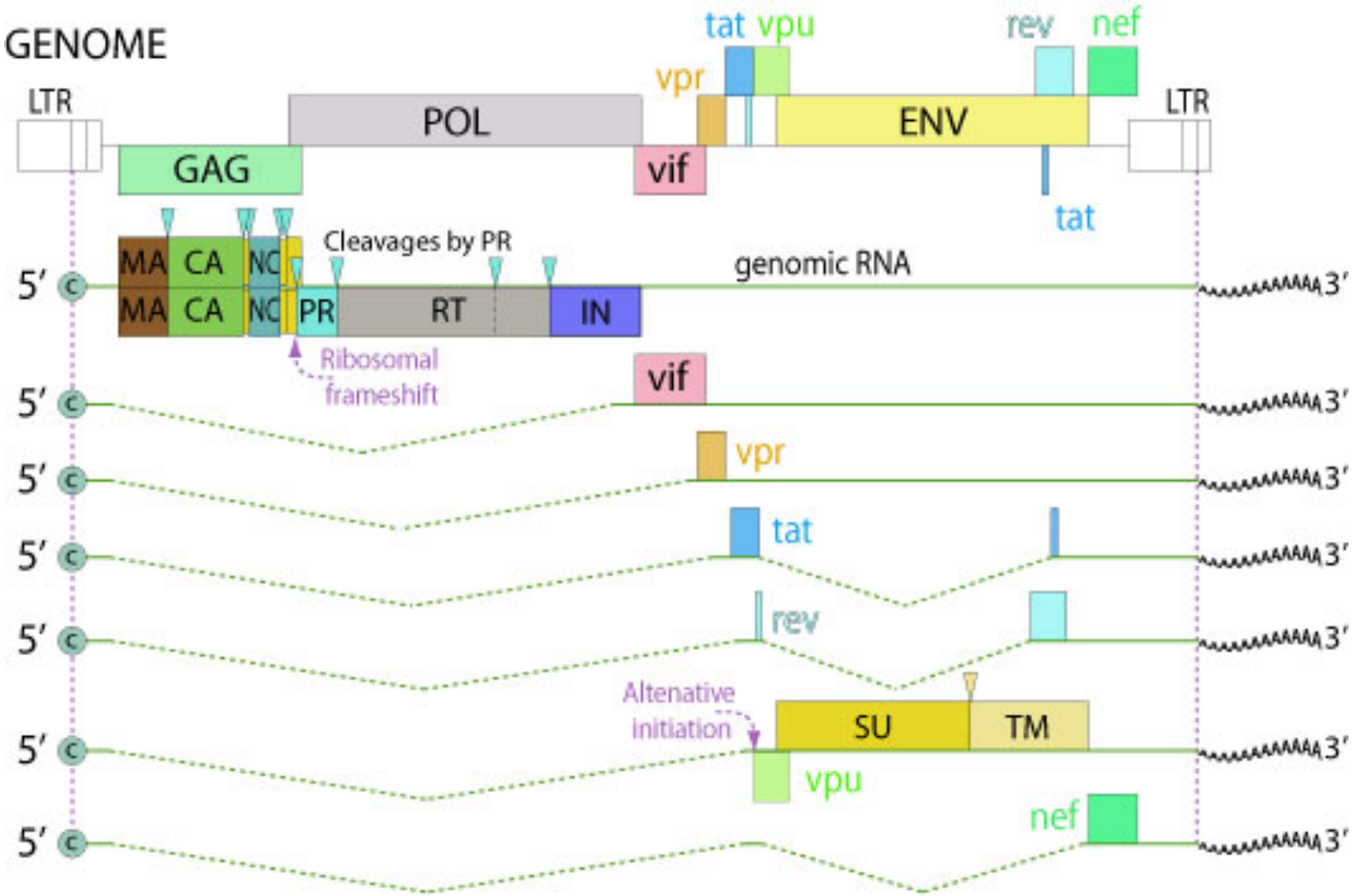

1
2
RNA
reverse
transcription
complex
3
DNA
pre-integration
complex
NUCLEUS
provirus
cell chromosome
cell RNA pol II
4
mRNA
6
progeny genomes
Golgi
Env
5
rough
endoplasmic
reticulum
7
tRNA
8
Gag
Gag–Pol

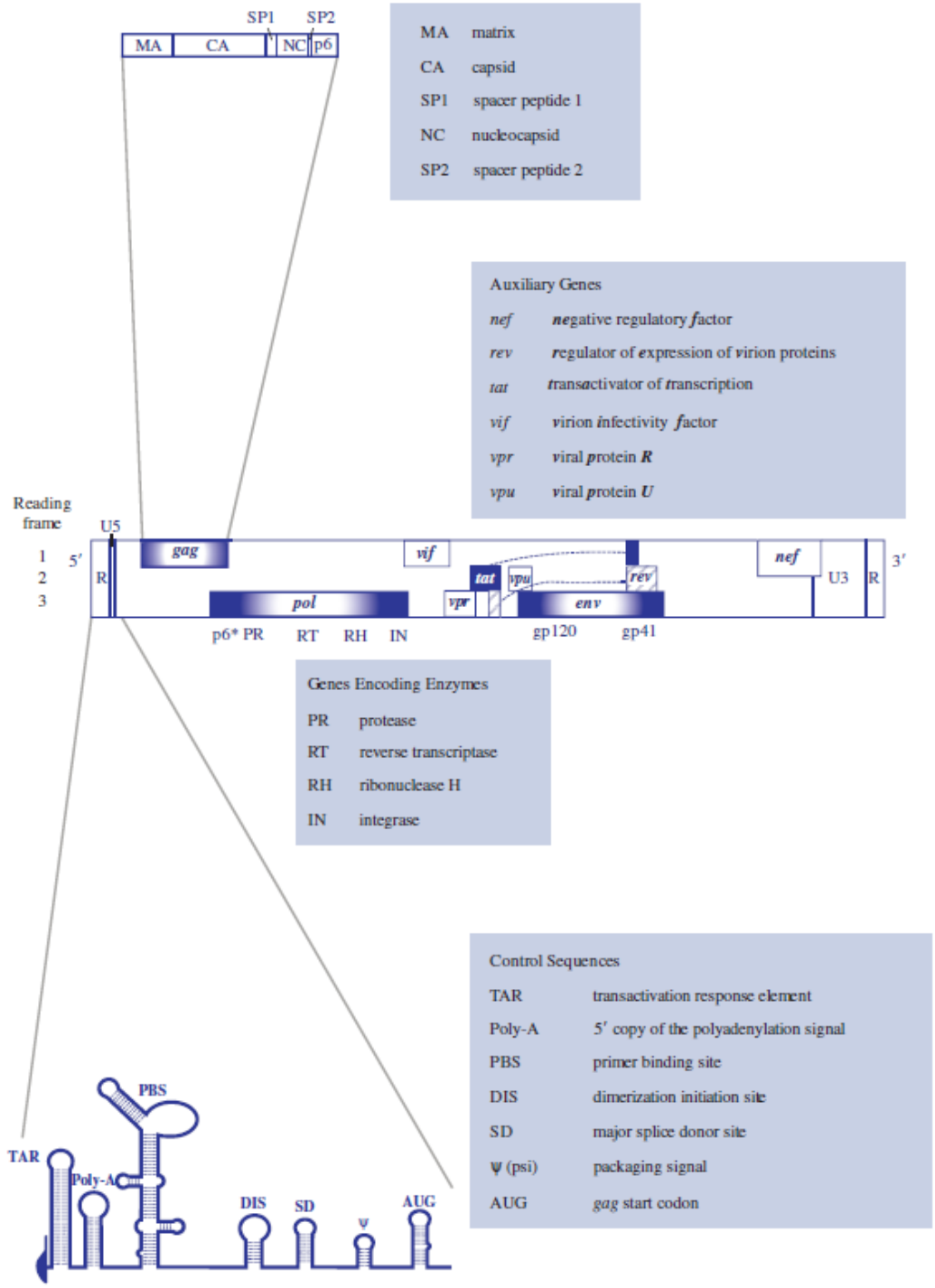
SRC

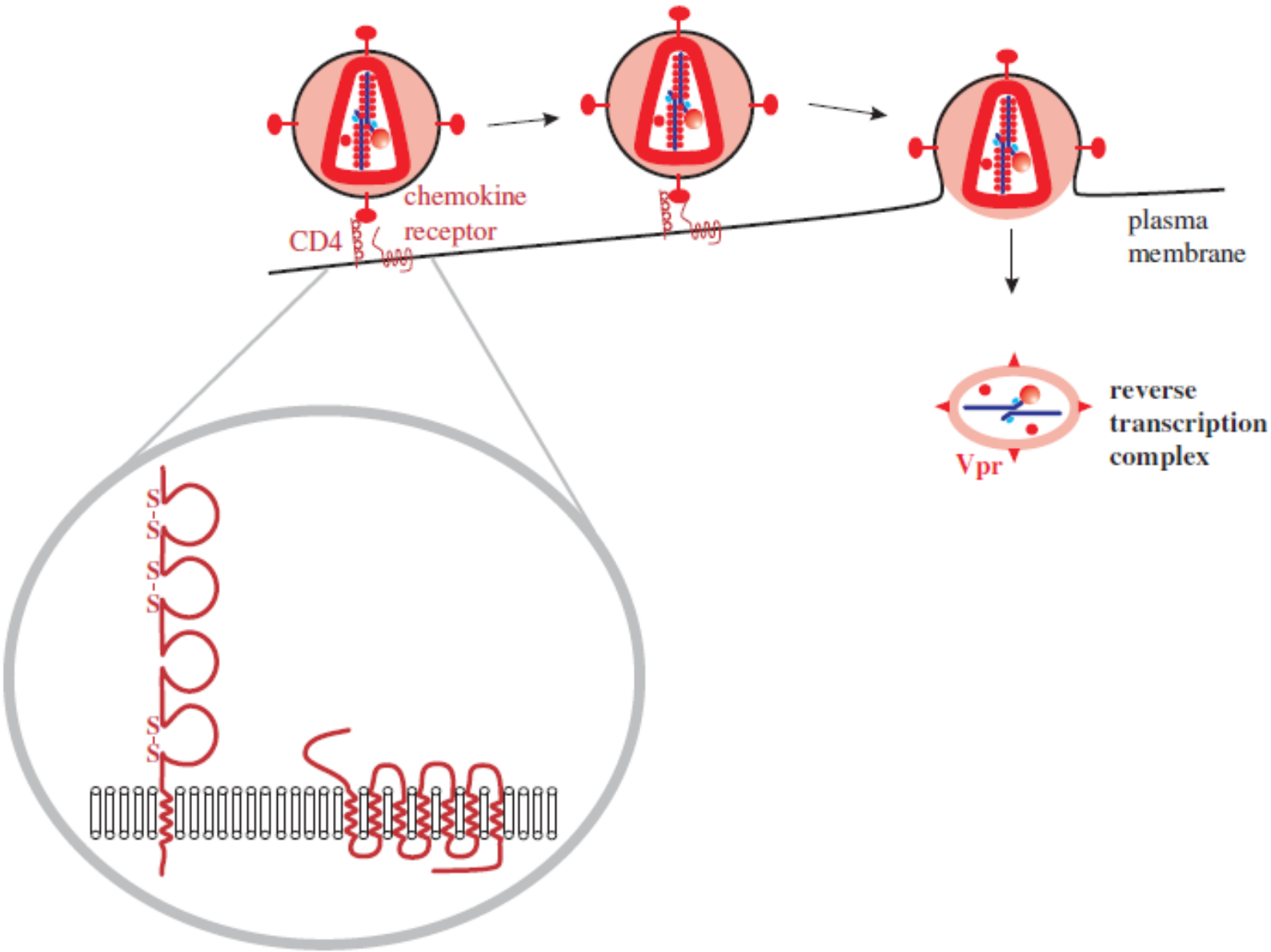
Figure 17.12 **Rous sarcoma virus genome.** There is an oncogene (*src*) in addition to the three standard retrovirus genes.

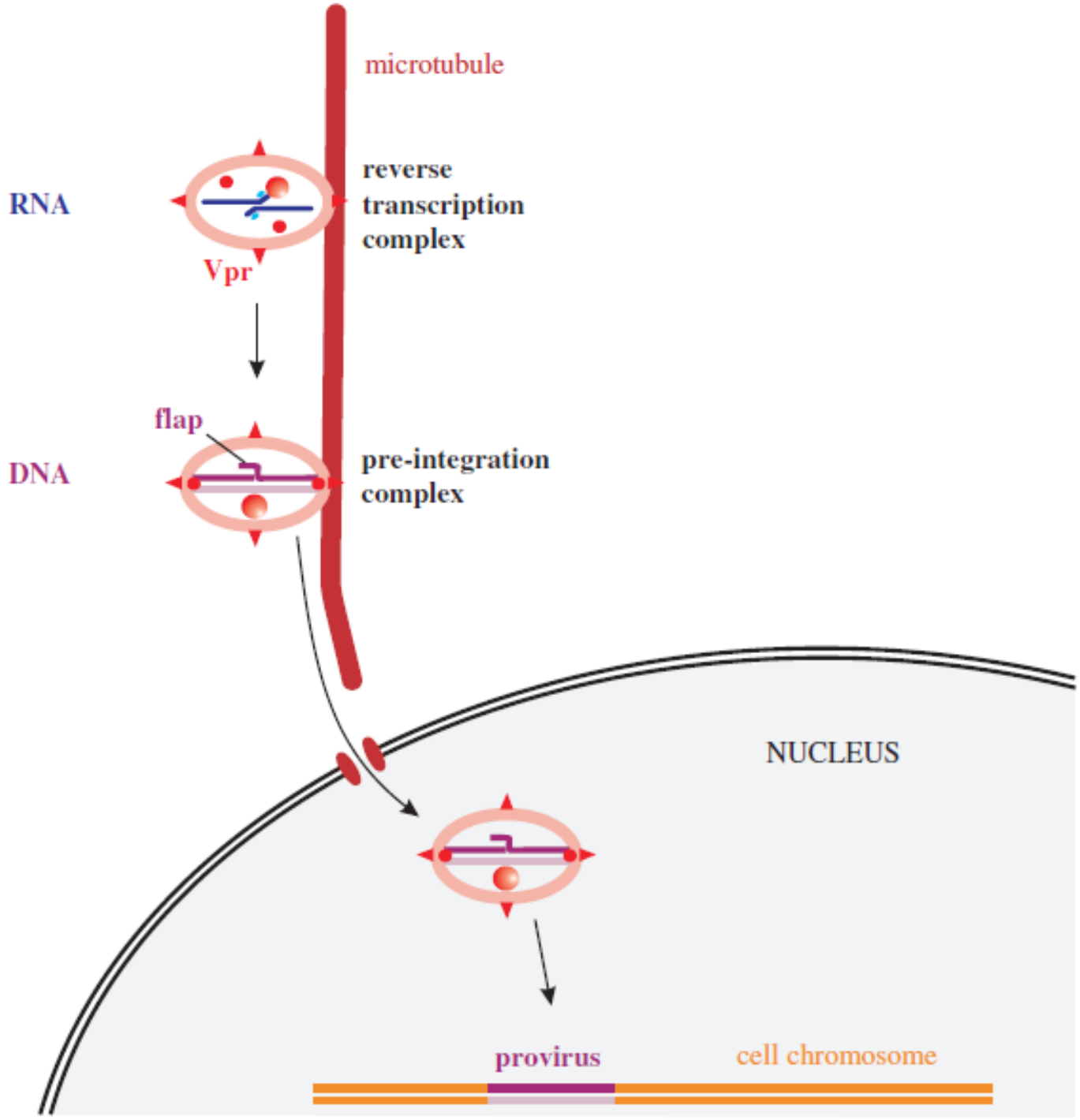



Survival Factors
(e.g., IGF1)
Chemokines,
Hormones,
Transmitters
(e.g., interleukins,
serotonin, etc.)
Growth Factors
(e.g., TGFα, EGF)
Extracellular
Matrix
RTK
GPCR
RTK
Integrins
cdc42
Fyn/Shc
Grb2/SOS
PLC
PI3K
G-Protein
Ras
FAK
Src
Wnt
Frizzled
Dishevelled
GSK-3β
APC
β-catenin
Hedgehog
Patched
SMO
Gli
TCF
β-catenin:TCF
Akt
PKC
Adenylate
cyclase
Raf
Akkα
NF-κB
IκB
PKA
MEK
MEKK
MAPK
MKK
Cytokines
(e.g., EPC)
Cytokine Receptor
JAKs
STAT3,5
Bcl-xL
Cytochrome C
Caspase 9
Caspase 8
Apoptosis
FADD
Bcl-2
Bad
Mt
Bax
FasR
Abnormality
Sensor
Bim
Death factors
(e.g. FasL, Tnf)
Myc:
Max
Mad:
Max
ERK
JNKs
Fos
Jun
CREB
Gene Regulation
CycID
CDK4
p16
p15
Rb
E2F
CyclE
CDK2
p27
p21
ARF
mdm2
p53
Cell
Proliferation

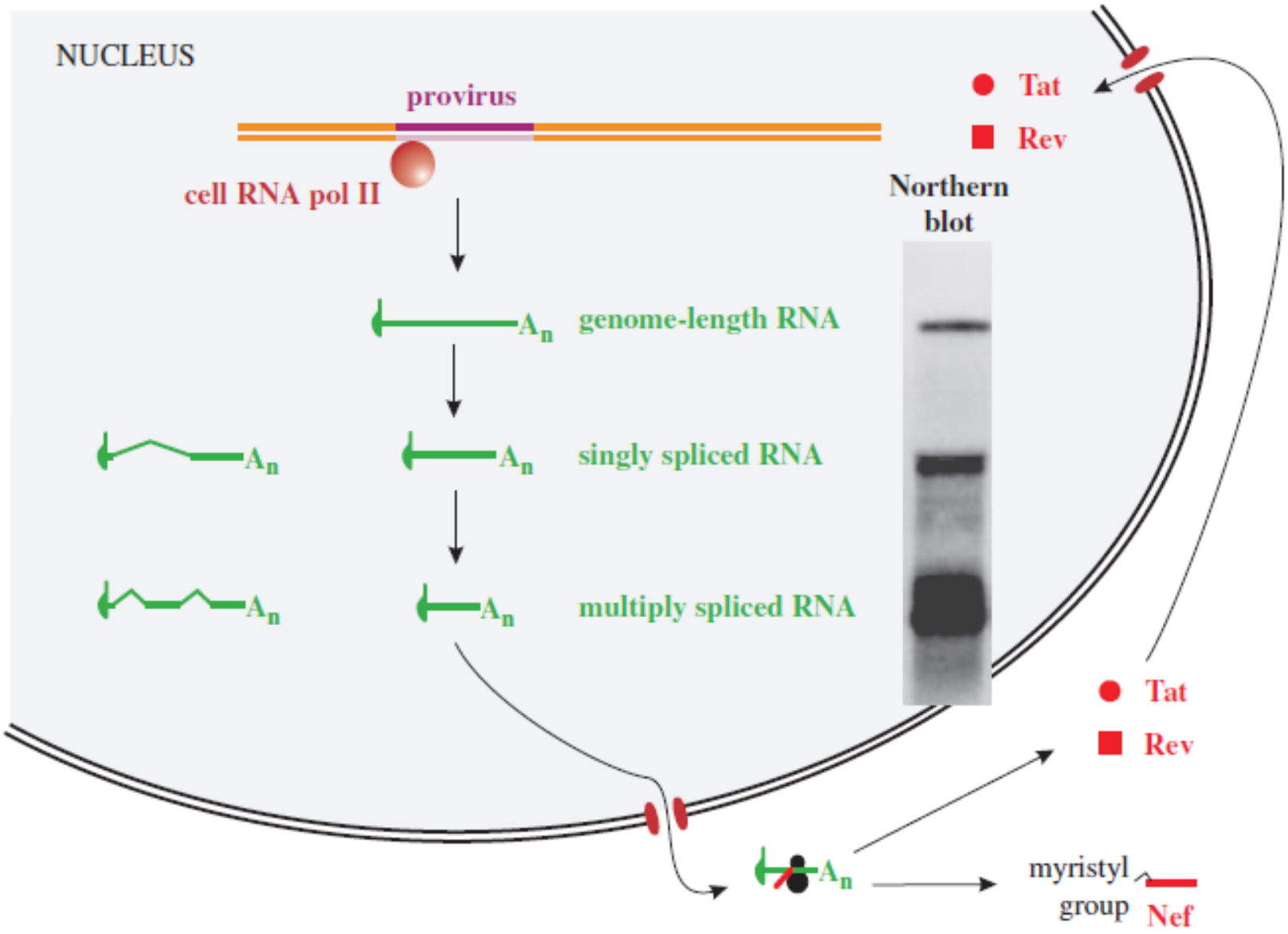

GENOME
LTR
GAG
POL
vif
vpr
tat vpu
rev
nef
ENV
LTR
tat
MA CA NC PR RT IN
Cleavages by PR
genomic RNA
Ribosomal frameshift
5′
3′
vif
vpr
tat
rev
Alternative initiation
SU
TM
vpu
nef


SP1
SP2
MA
CA
NC
p6
MA matrix
CA capsid
SP1 spacer peptide 1
NC nucleocapsid
SP2 spacer peptide 2
Auxiliary Genes
nef negative regulatory factor
rev regulator of expression of virion proteins
tat transactivator of transcription
vif virion infectivity factor
vpr viral protein R
vpu viral protein U
Reading frame
1
2
3
5′
U5
R
gag
pol
vif
tat
vpu
rev
vpr
env
nef
U3
R
3′
p6* PR
RT
RH
IN
gp120
gp41
Genes Encoding Enzymes
PR protease
RT reverse transcriptase
RH ribonuclease H
IN integrase
Control Sequences
TAR transactivation response element
Poly-A 5′ copy of the polyadenylation signal
PBS primer binding site
DIS dimerization initiation site
SD major splice donor site
Ψ (psi) packaging signal
AUG gag start codon
TAR
Poly-A
PBS
DIS
SD
Ψ
AUG

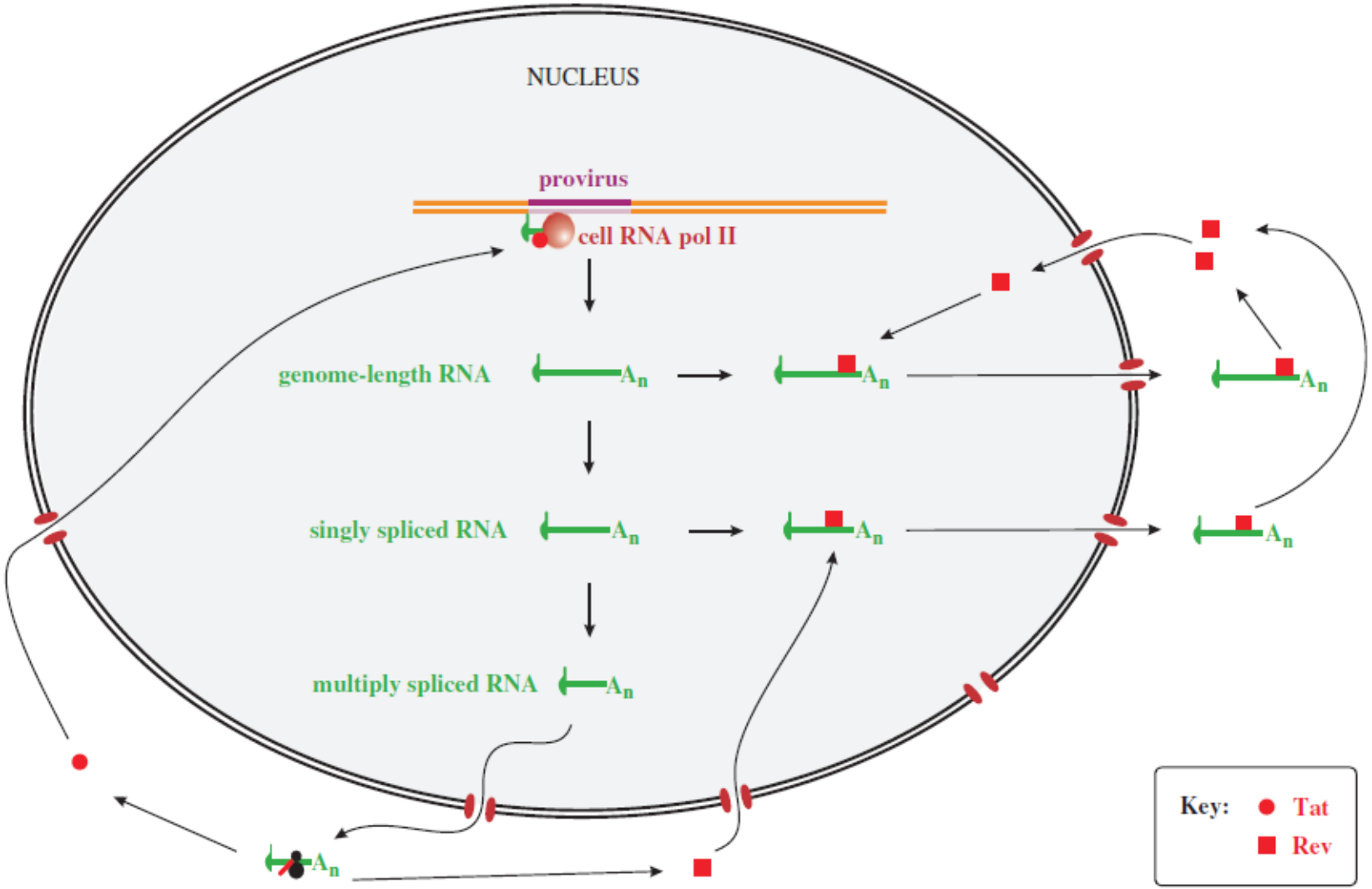

chemokine
receptor
CD4
plasma
membrane
reverse
transcription
complex
Vpr
S-S
S-S
S-S


microtubule
RNA
reverse
transcription
complex
Vpr
flap
DNA
pre-integration
complex
NUCLEUS
provirus
cell chromosome

Early transcription



Functions of Tat and Rev



Tat – translation of whole genome
Rev – transport of mRNAs to cytoplasm

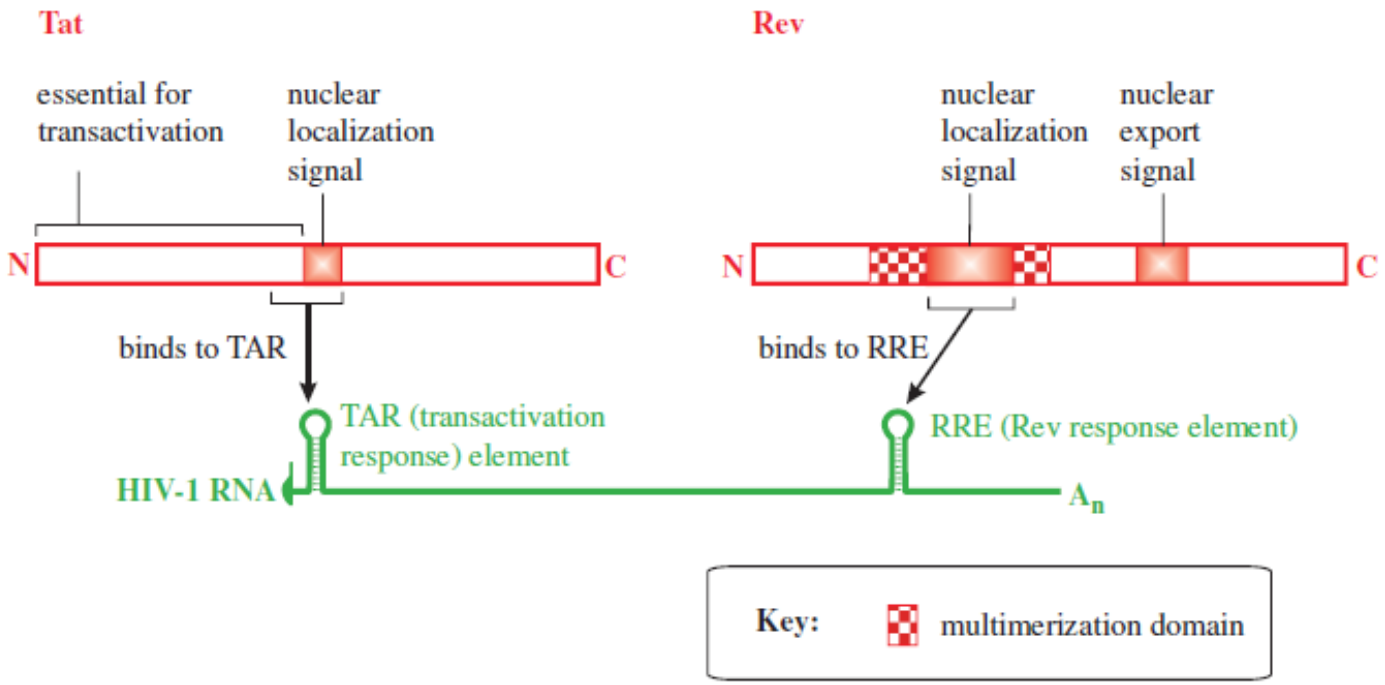


Figure 18.8 **HIV-1 Tat and Rev proteins and their binding sites in the virus RNA.** The TAR and RRE regions of the RNA have complex secondary structures. The RRE is present in genome-length RNA and the singly spliced RNAs, but it is absent from the multiply spliced RNAs.

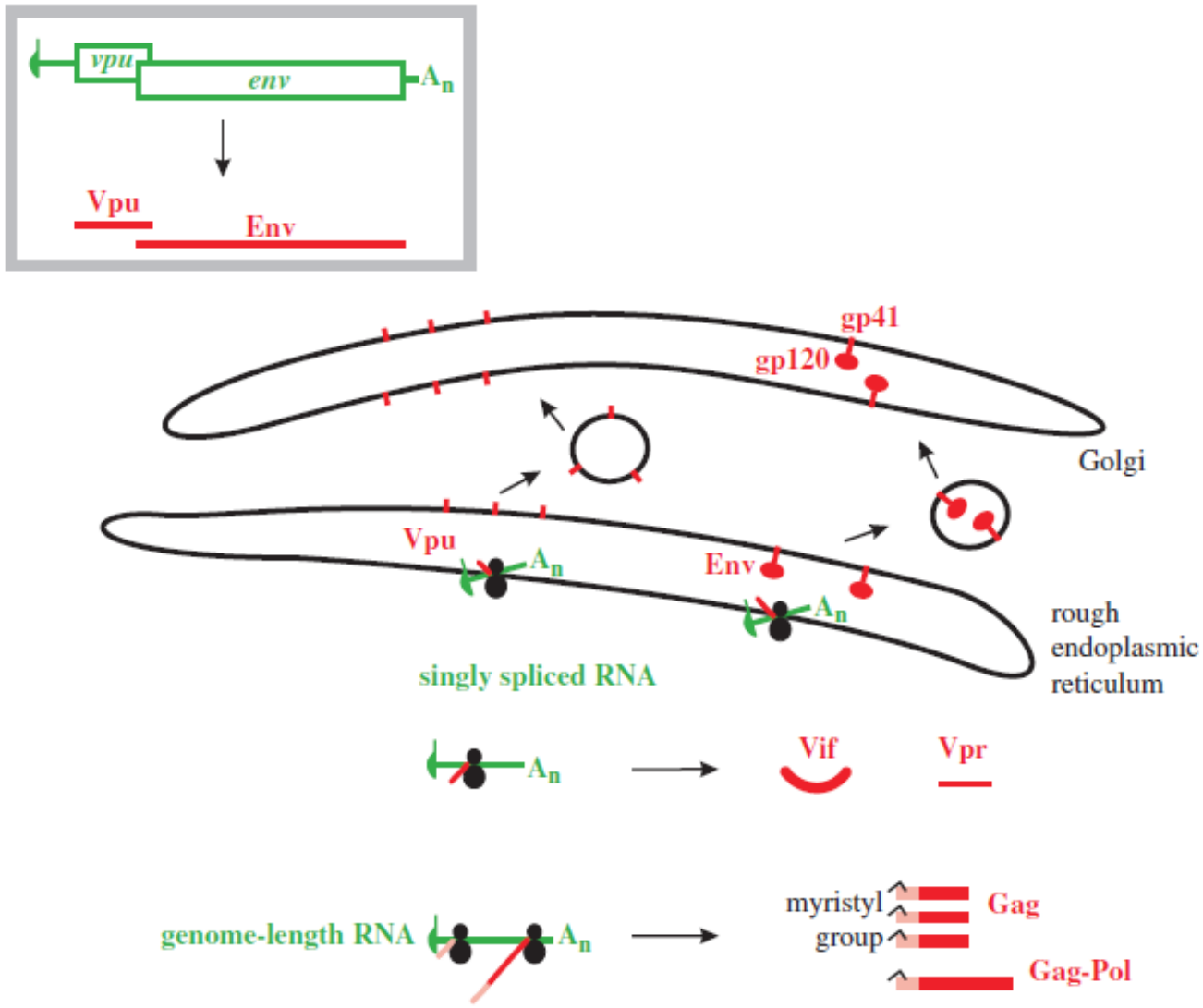


Figure 18.9 **HIV-1 late gene expression.** Vpu and Env are translated from singly spliced RNAs in the rough endoplasmic reticulum. The inset shows translation of Vpu and Env from a bicistronic mRNA. Env is synthesized when the *vpu* start codon is bypassed during leaky scanning. The remaining proteins are translated on free ribosomes: Vif and Vpr from singly spliced RNAs, and Gag and Gag–Pol from genome-length RNAs.

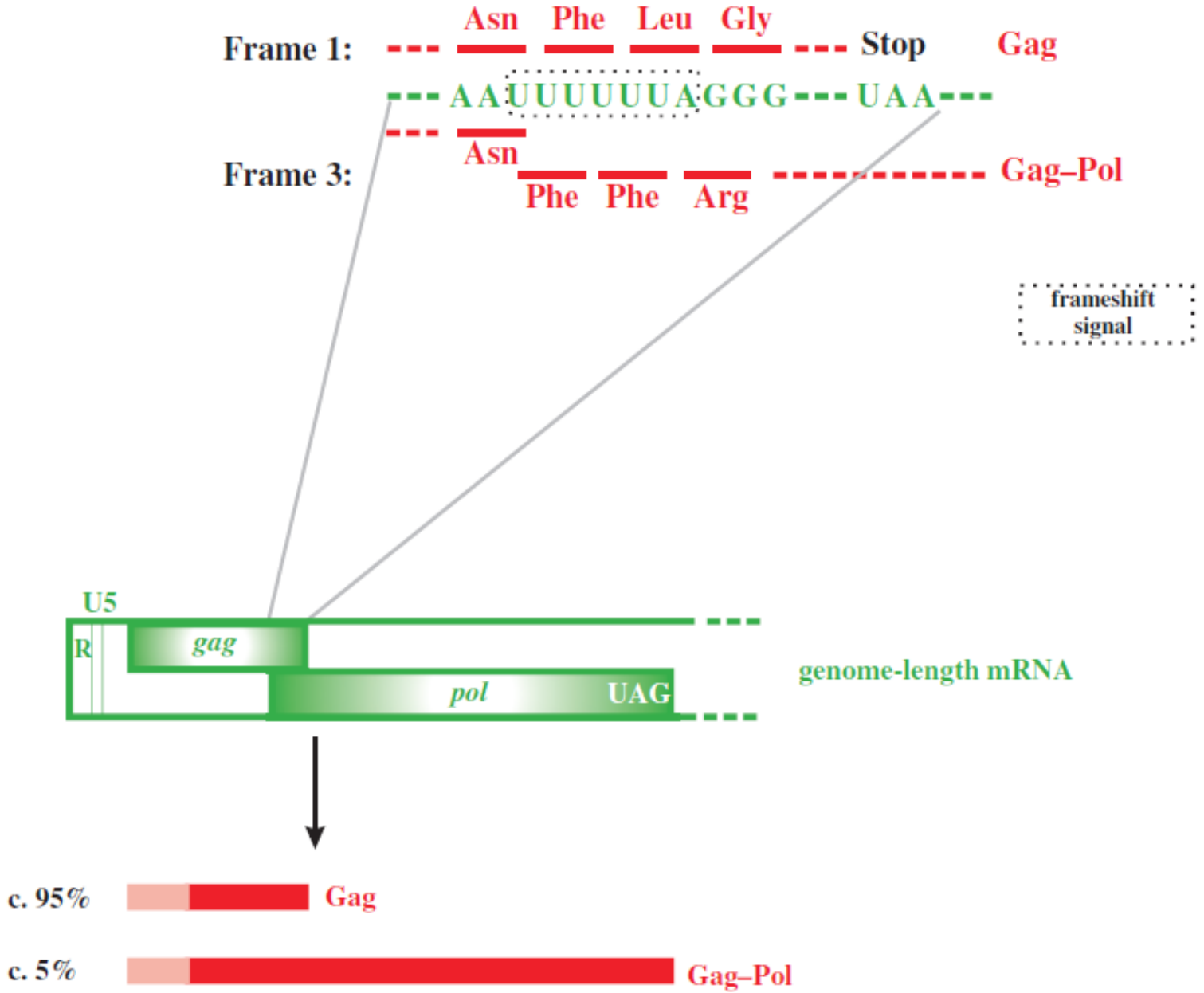
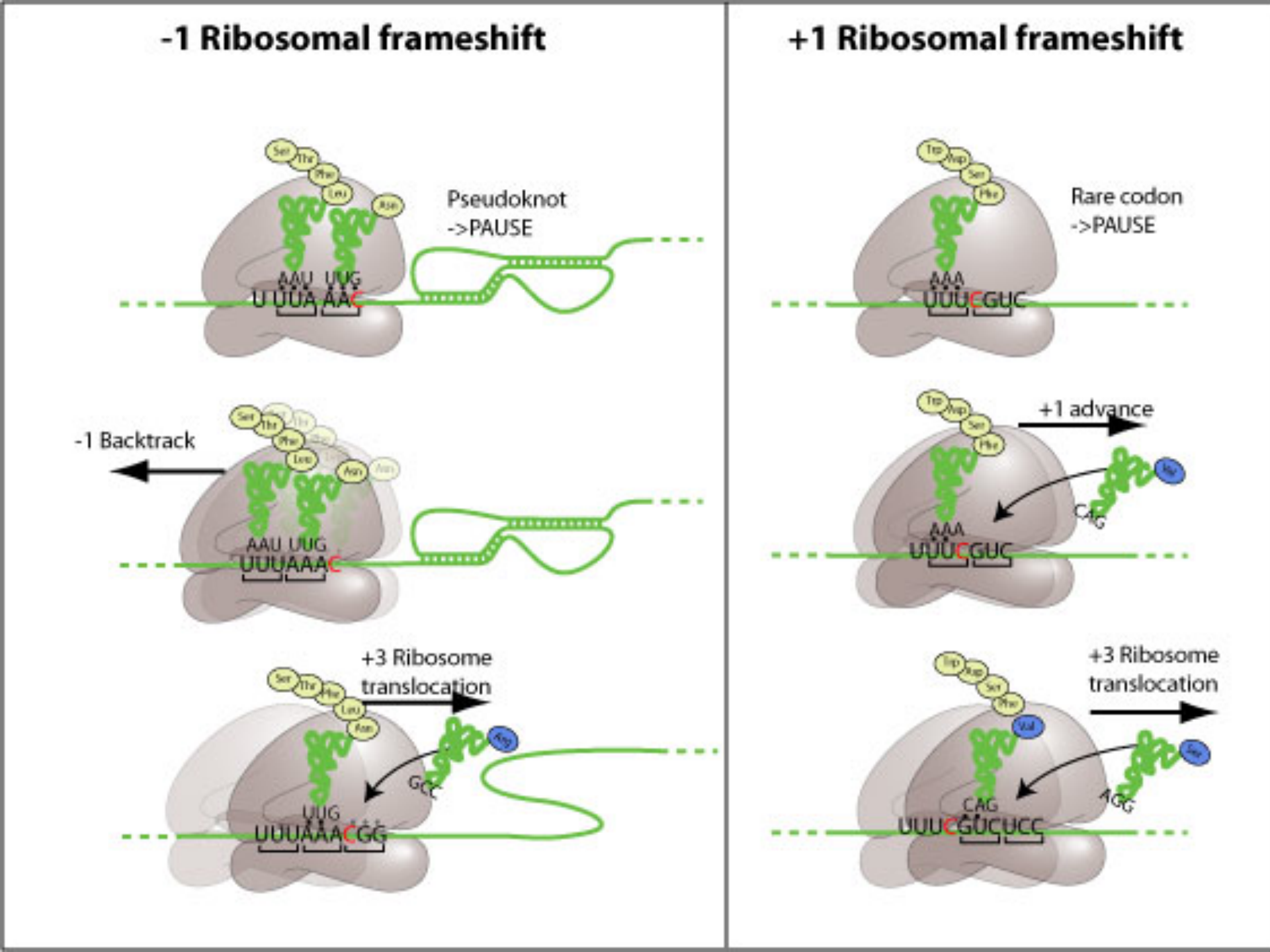
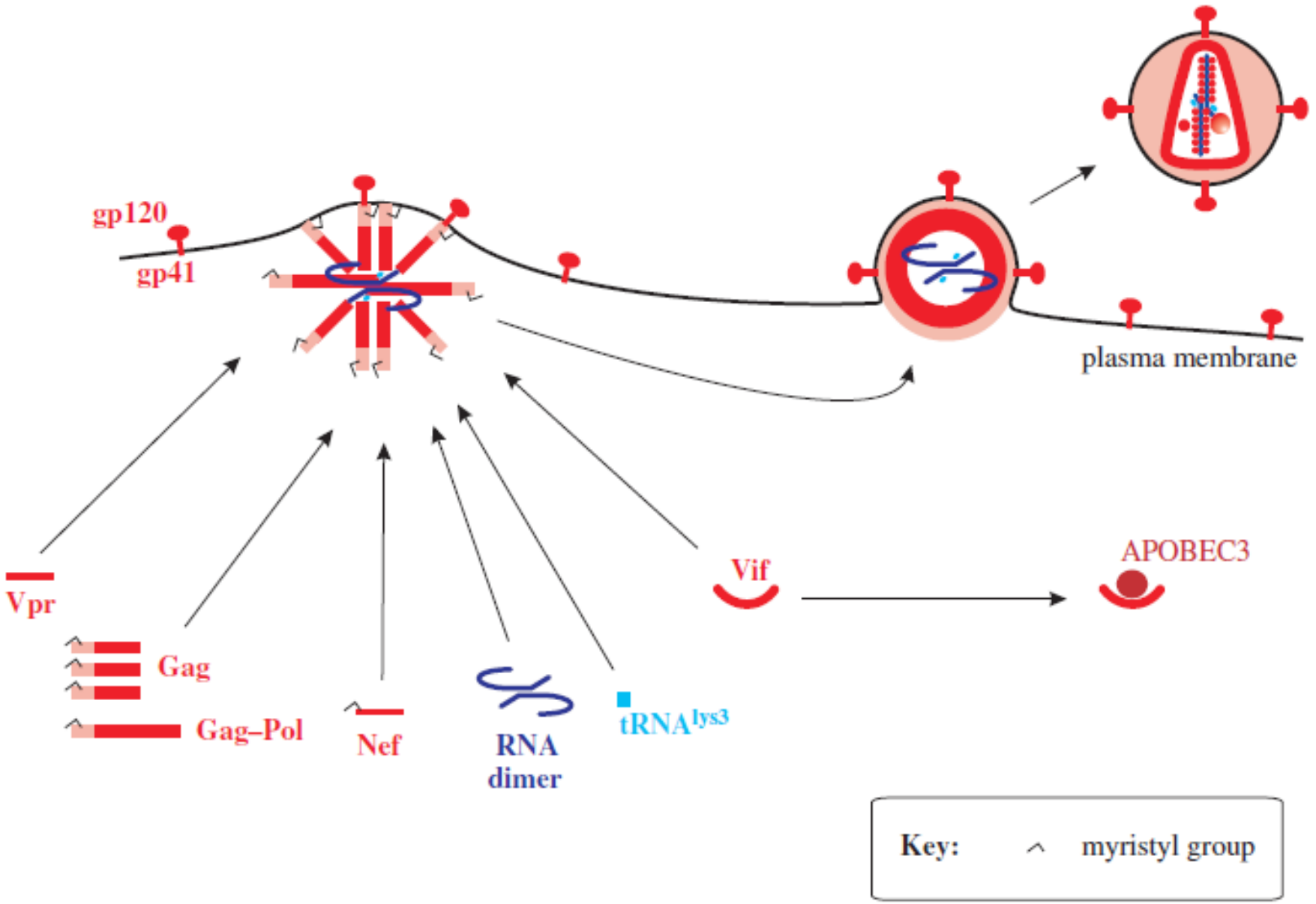
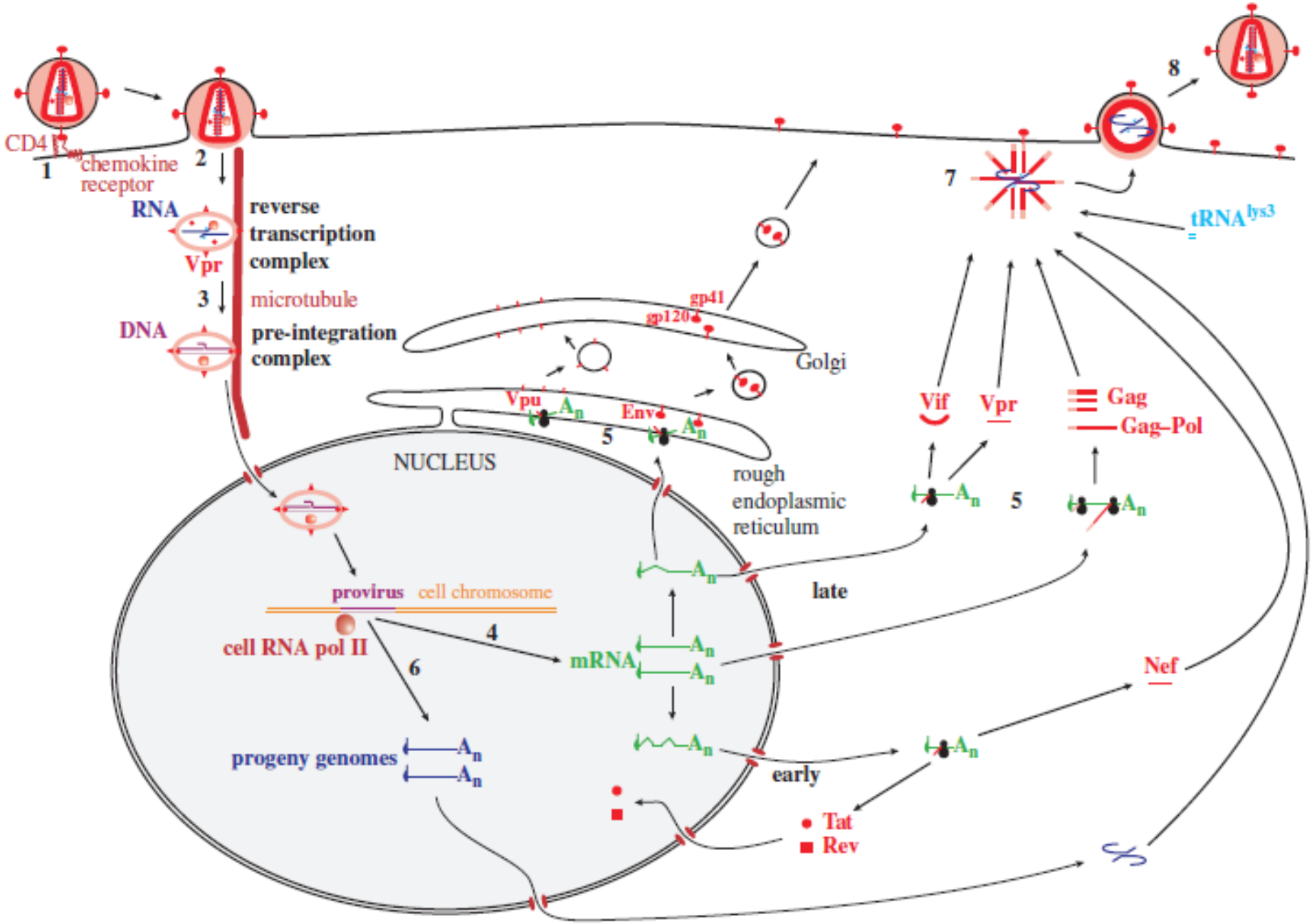


Figure 18.10 Expression of HIV-1 Gag–Pol by ribosomal frameshifting. A ribosome reading in frame 1 shifts at the slippery sequence UUUUUUA to reading in frame 3.


-1 Ribosomal frameshift
+1 Ribosomal frameshift
Pseudoknot
->PAUSE
Rare codon
->PAUSE
-1 Backtrack
+1 advance
+3 Ribosome
translocation
+3 Ribosome
translocation

Virion assembly




CD4
chemokine receptor
1
2
RNA
Vpr
reverse transcription complex
3
microtubule
DNA
pre-integration complex
gp41
gp120
Golgi
Vpu
Env
5
NUCLEUS
rough endoplasmic reticulum
provirus
cell chromosome
cell RNA pol II
4
6
mRNA
progeny genomes
late
early
Tat
Rev
Nef
Vif
Vpr
Gag
Gag–Pol
7
8
tRNA lys3

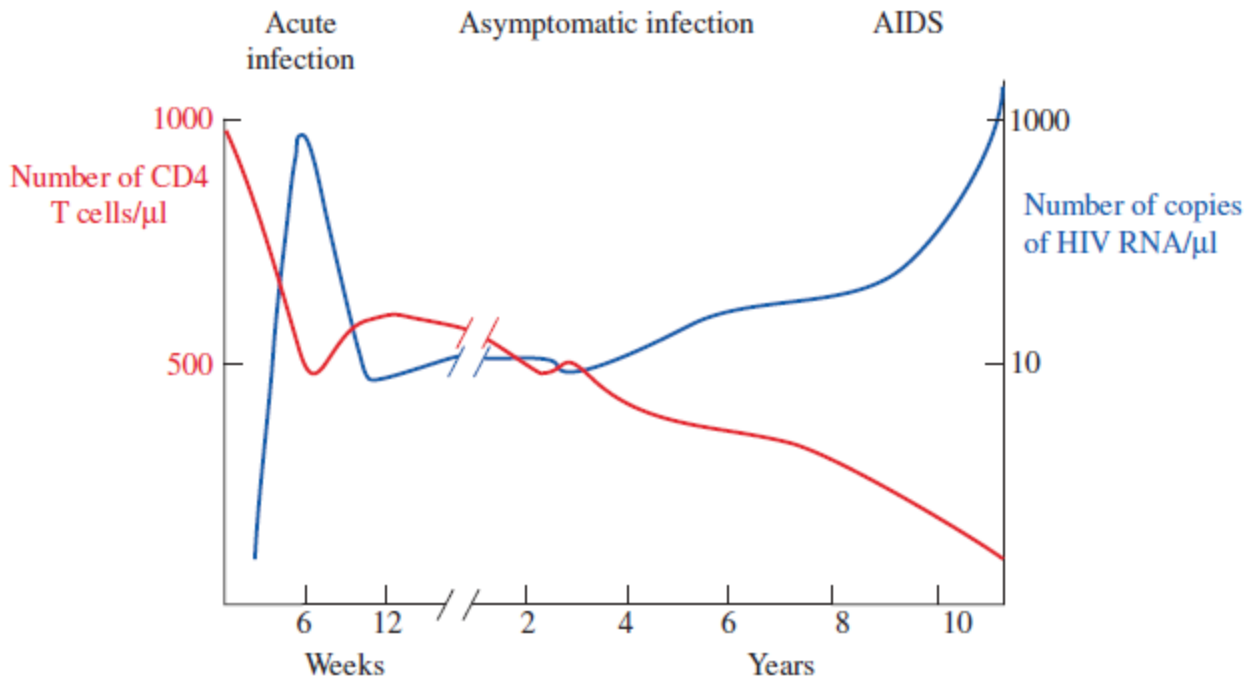
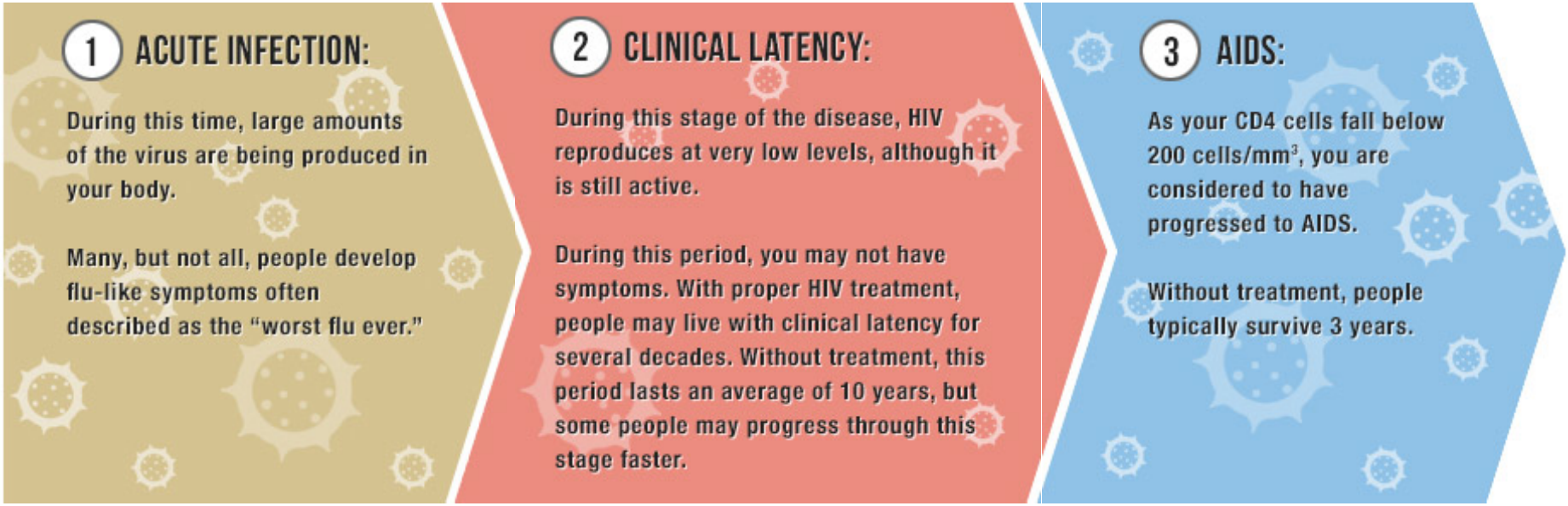


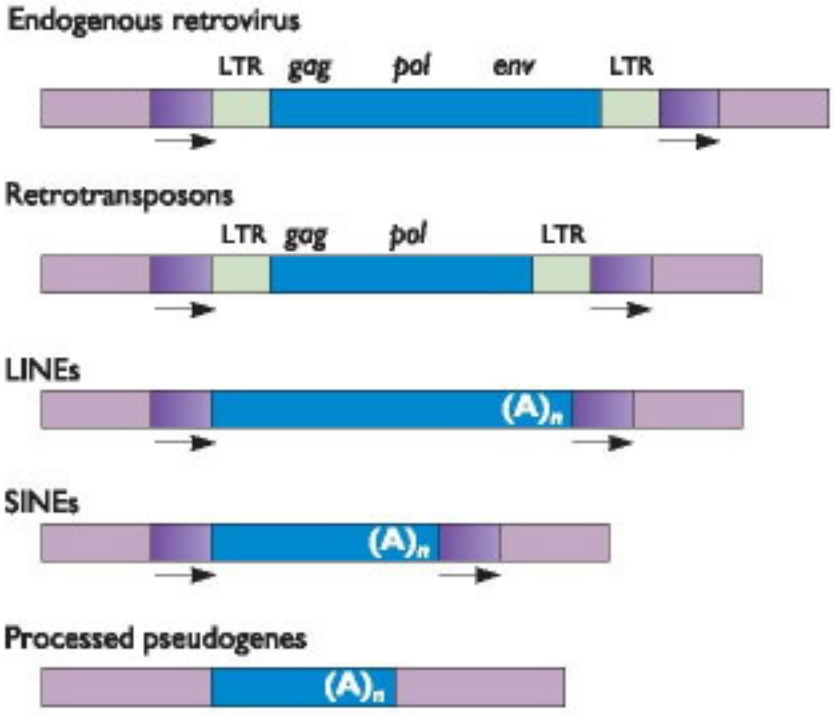
Figure 18.13 Levels of CD4 T cells and HIV RNA in blood during HIV infection. The concentration of HIV RNA is a measure of viremia. Shortly after infection viremia rises, then it falls off and relatively low levels are detectable throughout the asymptomatic period. A rise in viremia heralds the onset of AIDS.

HIV infection - AIDS



Simple retroviruses		Complex retroviruses	
Genus	Virus examples	Genus	Virus examples
Alpharetrovirus	Rous sarcoma virus	*Deltaretrovirus*	Human T-lymphotropic viruses 1 & 2
Betaretrovirus	Mouse mammary tumour virus	*Epsilonretrovirus*	Walleye dermal sarcoma virus
		Lentivirus	Human immunodeficiency virus 1
Gammaretrovirus	Murine leukaemia virus Feline leukaemia virus	*Spumavirus*	Chimpanzee foamy virus

Characteristics of retroelements resident in eukaryotic genomes



Designation	Characteristic	Example	Copy no.
Endogenous retroviruses	RT, LTR (internal Pol II promoter), and *env*	HERVs (human)	$1–10^2$
Retrotransposons	RT, LTR (internal Pol II promoter)	Ty3 (yeast)	$10^2–10^4$
Retroposons (LINEs)	RT, internal Pol III promoter, A-rich sequence at end	LINE 1 (human)	$10^4–10^5$
Retrosequences (SINEs)	A-rich sequence at end, internal Pol III promoter, but no RT	*Alu* (human)	$10^5–10^6$
Processed pseudogenes	A-rich sequence at end, no internal promoter, no RT	β-Tubulin (human)	$1–10^2$

Learning outcomes

- describe the retrovirus virion
- describe the main features of the retrovirus genome
- explain the main features of the retrovirus replication cycle
- give examples of retroviruses and explain their importance
- discuss endogenous retroviruses

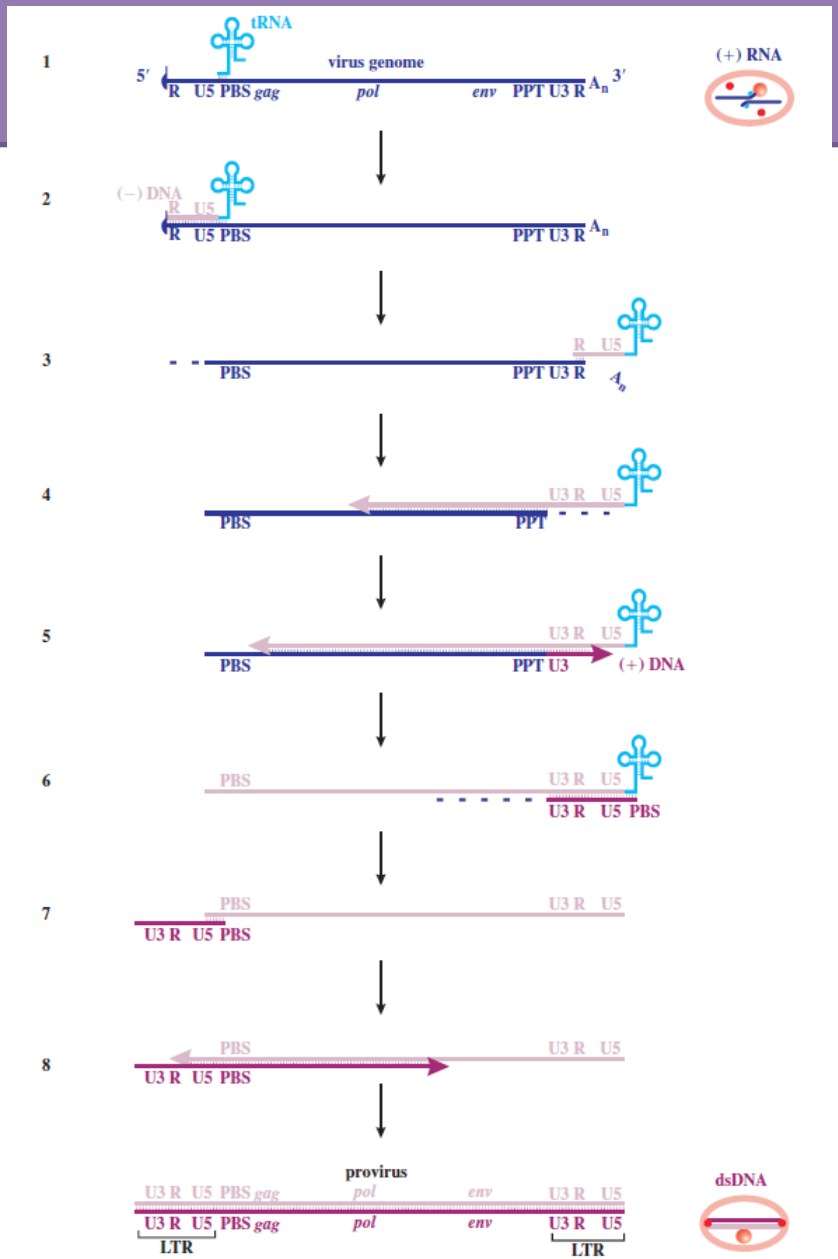


Figure 17.3

Figure 17.3 Retroviral reverse transcription. LTR: long terminal repeat. PBS: primer binding site. PPT: polypurine tract (a sequence made up entirely, or almost entirely, of purine residues). R: repeat sequence. U3: unique sequence at 3′ end of genome. U5: unique sequence at 5′ end of genome.

1. A copy of the virus genome with a tRNA bound at the PBS.
2. The reverse transcriptase begins (–) DNA synthesis at the 3′ end of the tRNA.
3. The RNase H digests the RNA from the RNA–DNA duplex. The (–) DNA attaches at the 3′ end of either the same RNA strand or the second copy of the genome.
4. Elongation of the (–) DNA continues, while the RNase H degrades the template RNA from the 3′ end as far as the PPT.
5. Synthesis of (+) DNA begins.
6. The remaining RNA is degraded.
7. The (+) DNA detaches from the 5′ end of the (–) DNA template and attaches at the 3′ end.
8. Synthesis of both DNA strands is completed.

Learning outcomes

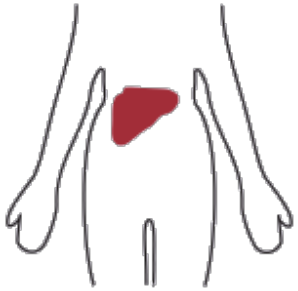
- explain the importance of HIV
- describe, with the aid of a labelled diagram, the HIV-1 virion
- describe the HIV-1 genome
- write an illustrated account of the replication cycle of HIV-1
- discuss the variability of HIV-1
- discuss the effects of HIV infection on the host
- evaluate approaches to the prevention of HIV transmission

Hepadnaviruses and other reverse transcribing DNA viruses

Hepatitis-causing **DNA** viruses

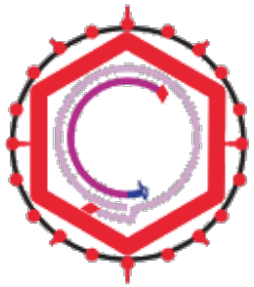
Hosts: humans and other primates
birds

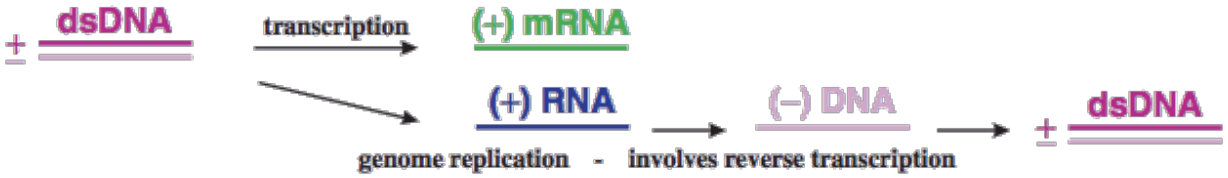
Cause diseases of the liver.



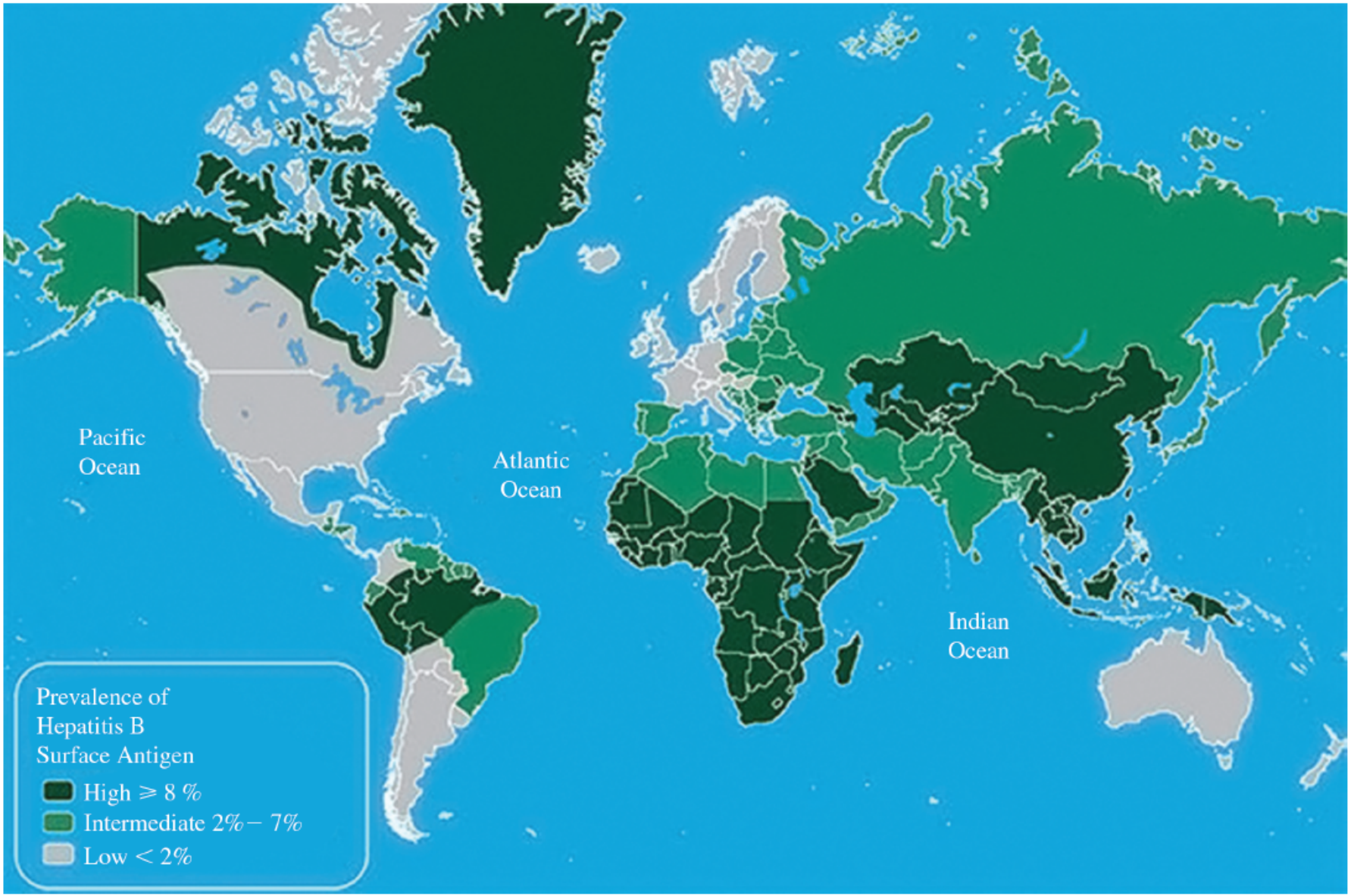
Virion

- Enveloped
- 40–48 nm diameter
- Icosahedral capsid
- Genome: DNA (partly single-stranded)
  3 kb(p)

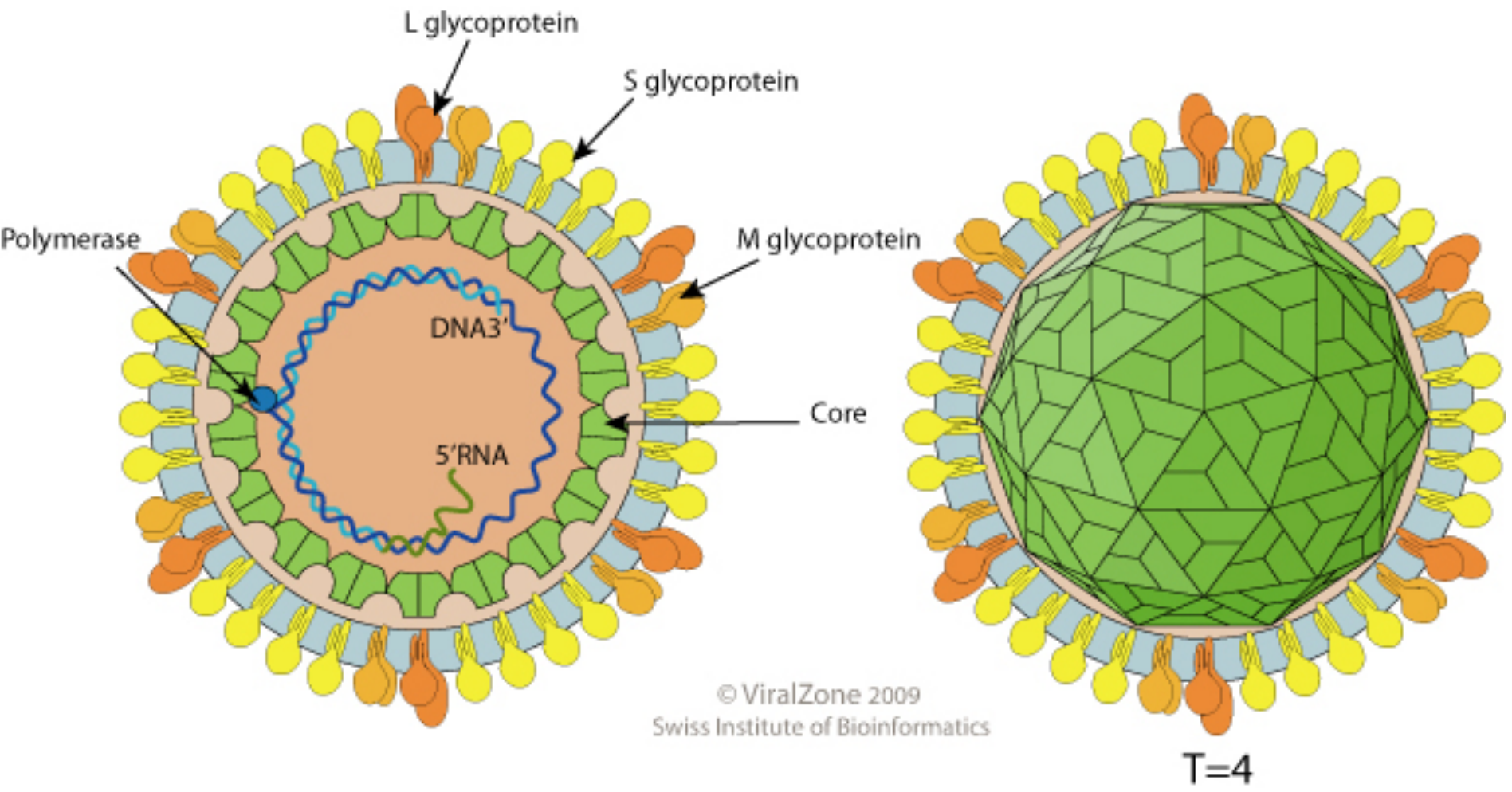


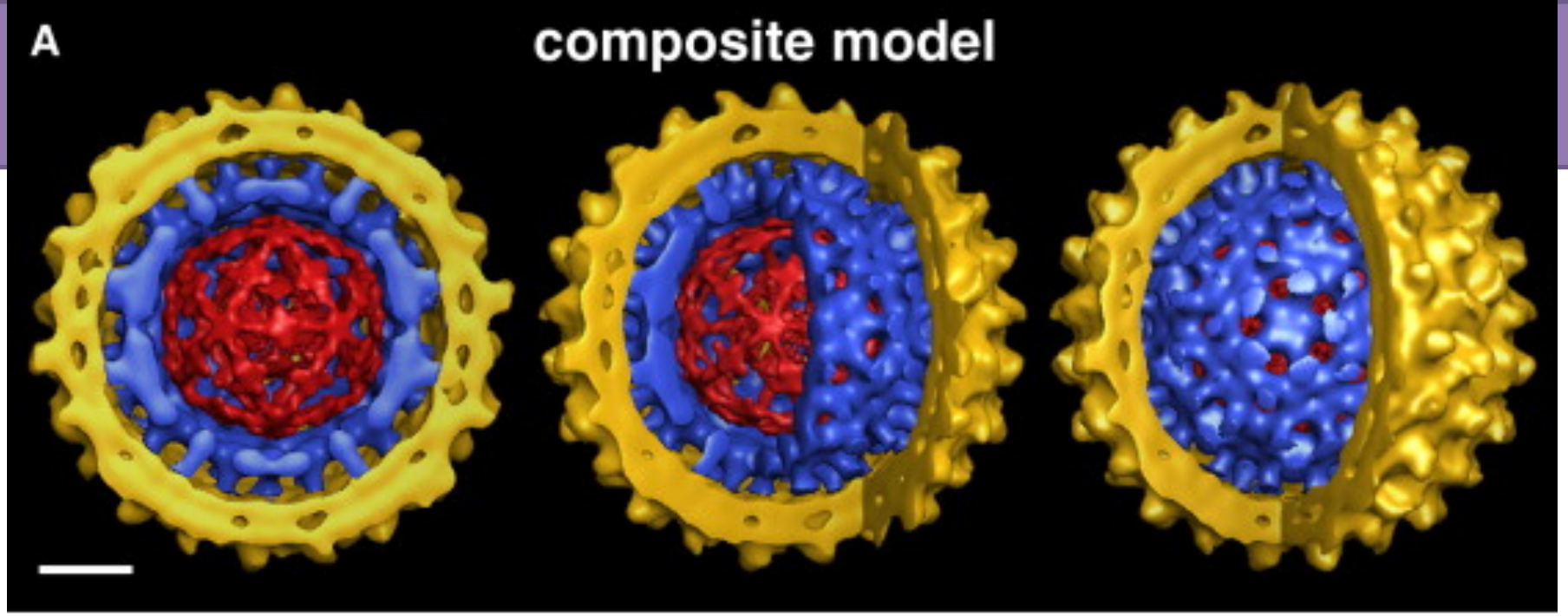


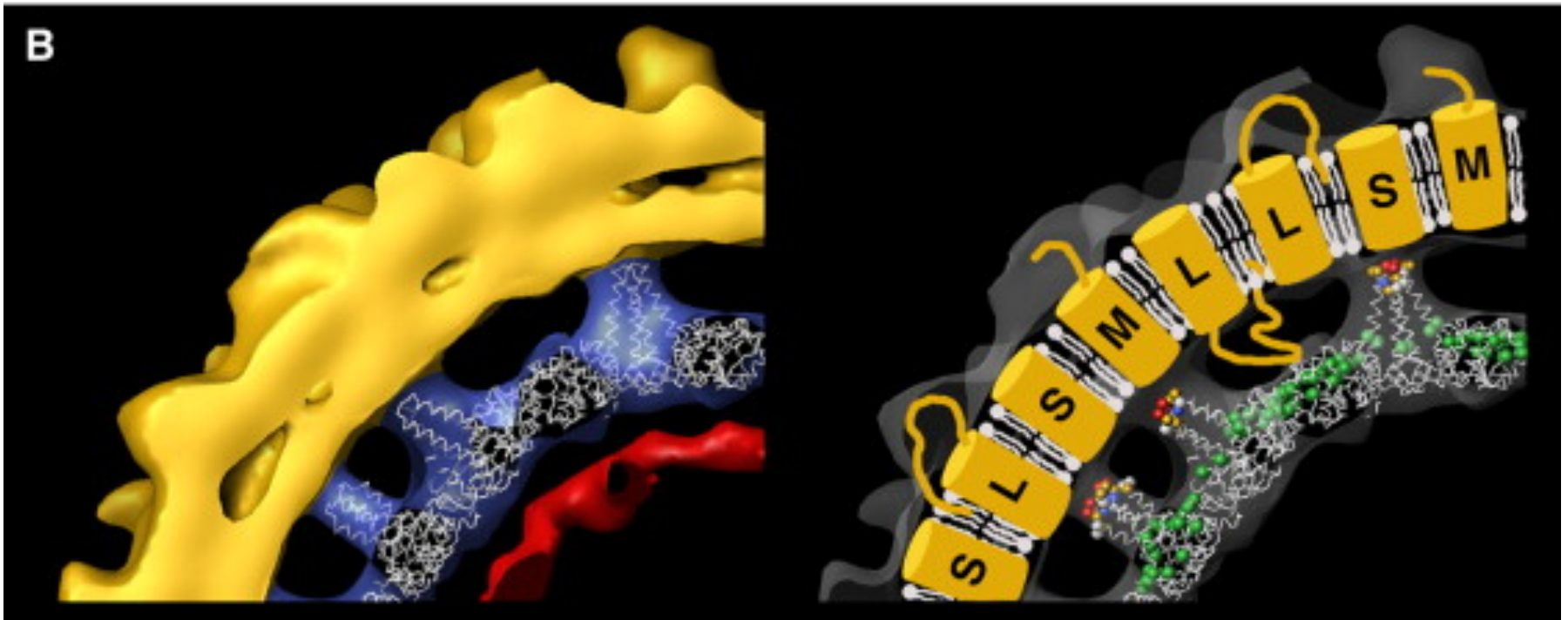
World distribution of Hepatitis B

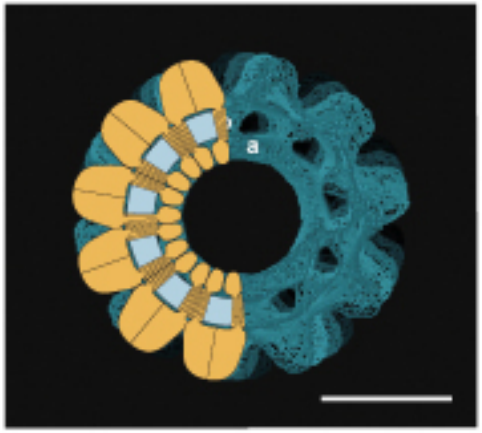


Hepatitis B virion structure

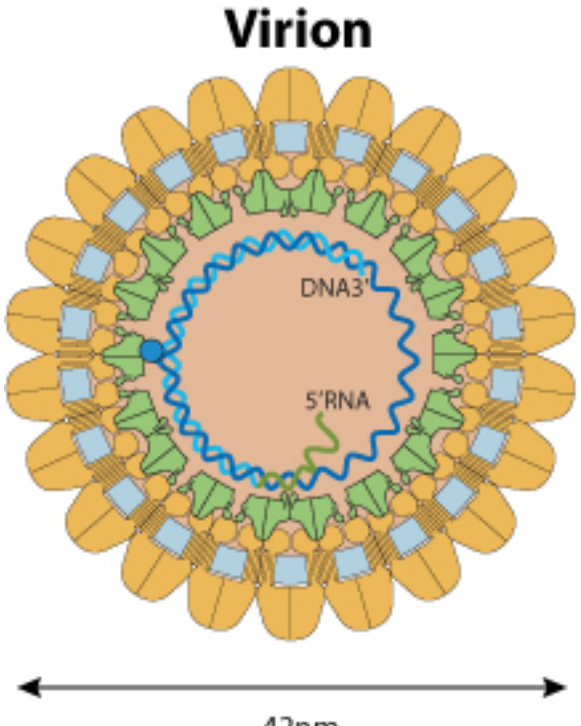


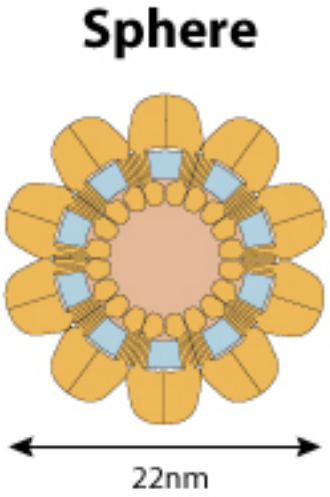

A
composite model

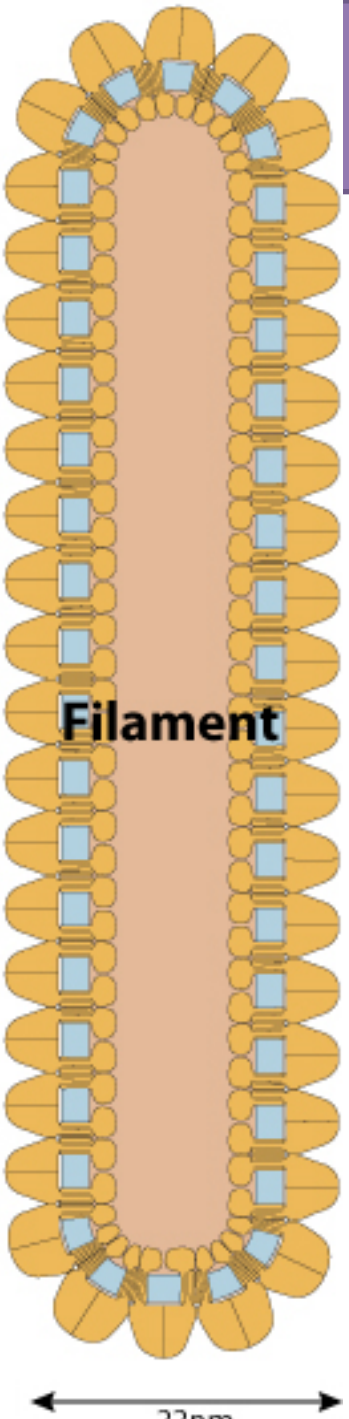

B
S
L
S
M
L
L
S
M



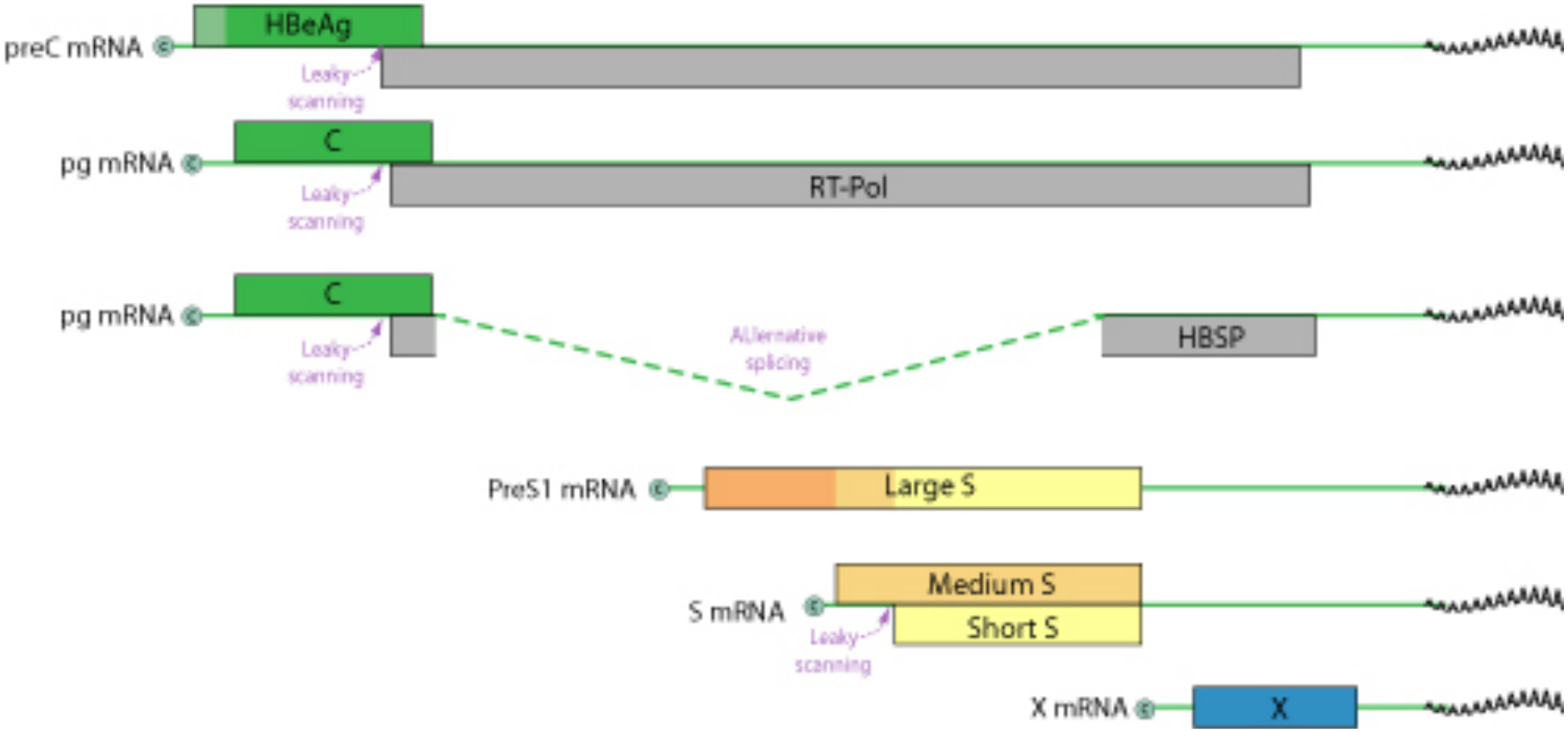
picture from Short et al. 2009



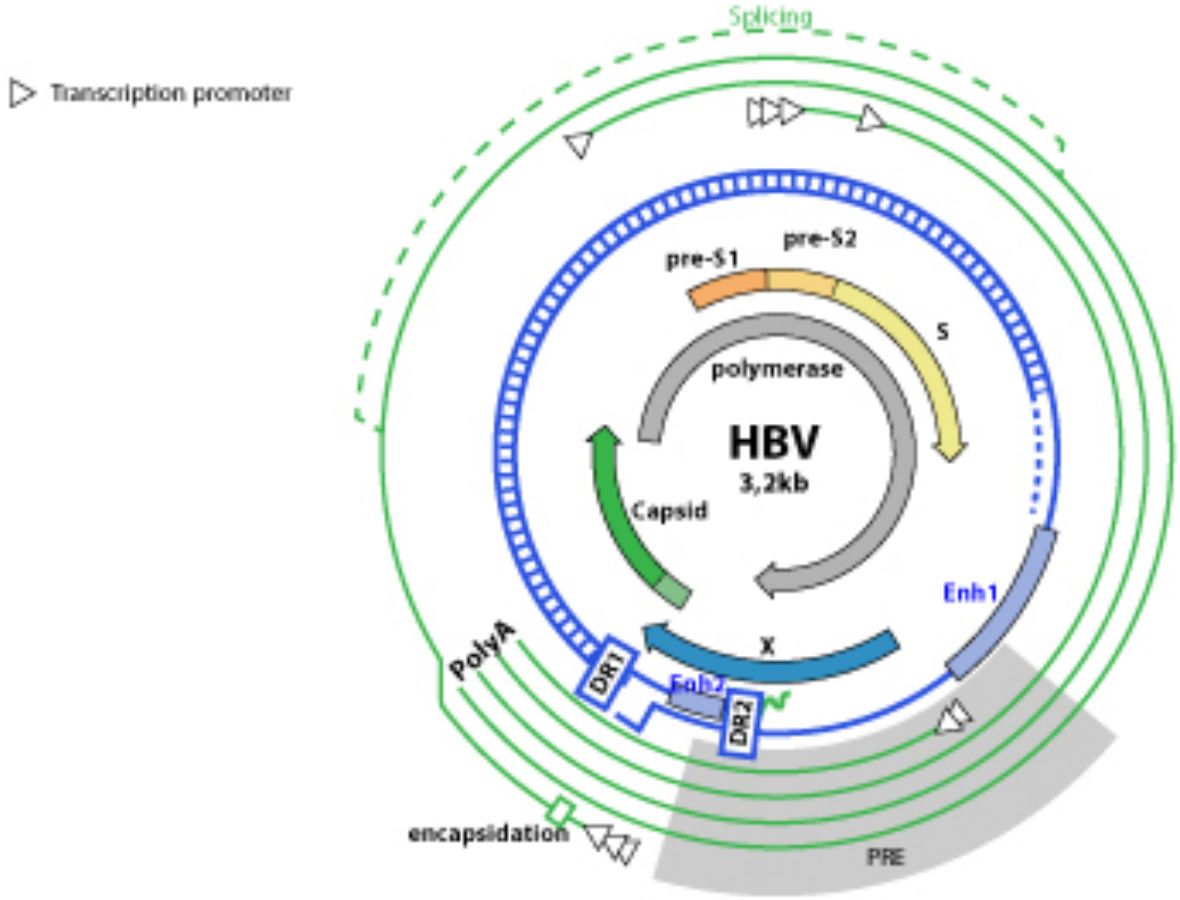




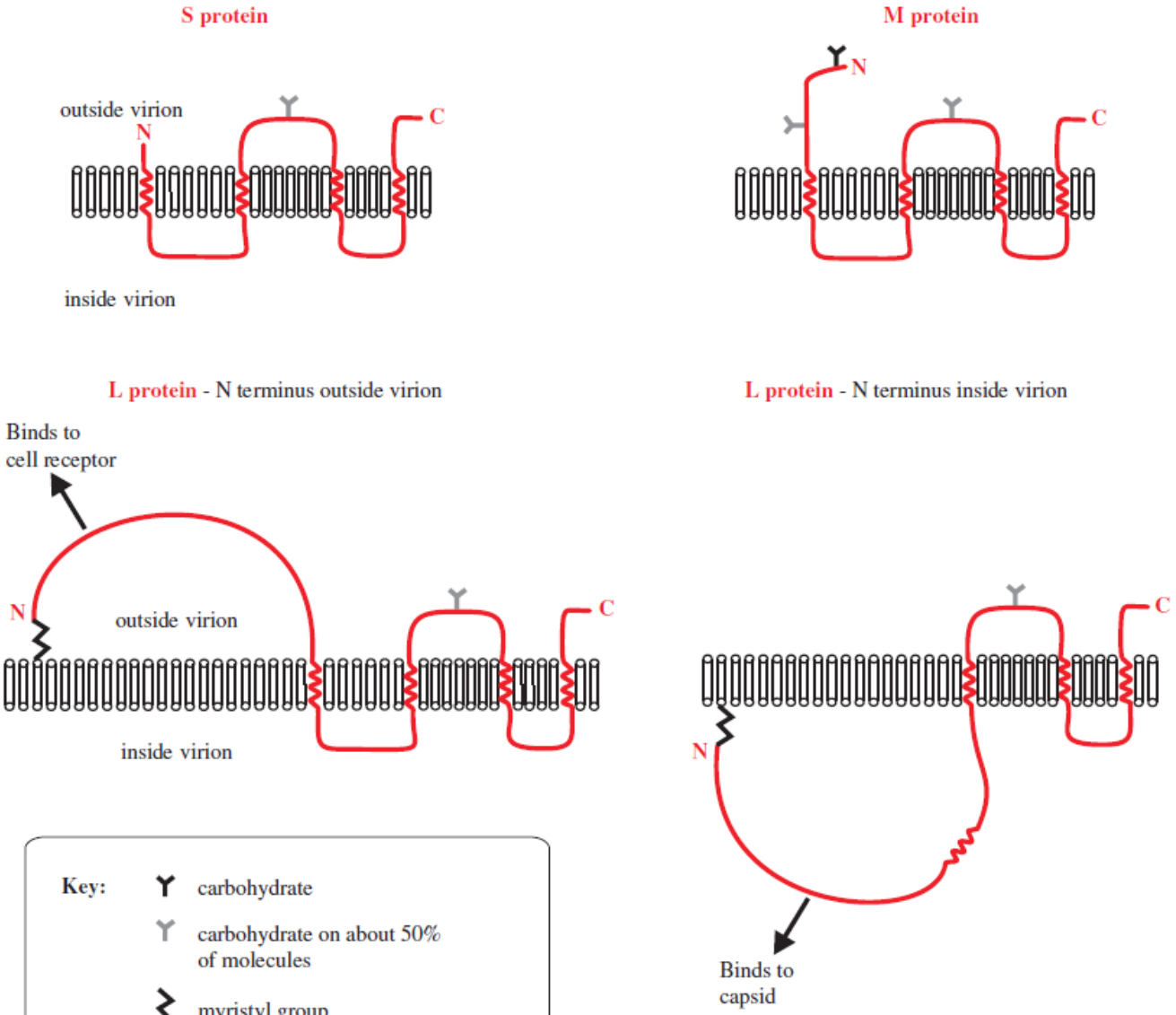
HBV transcripts

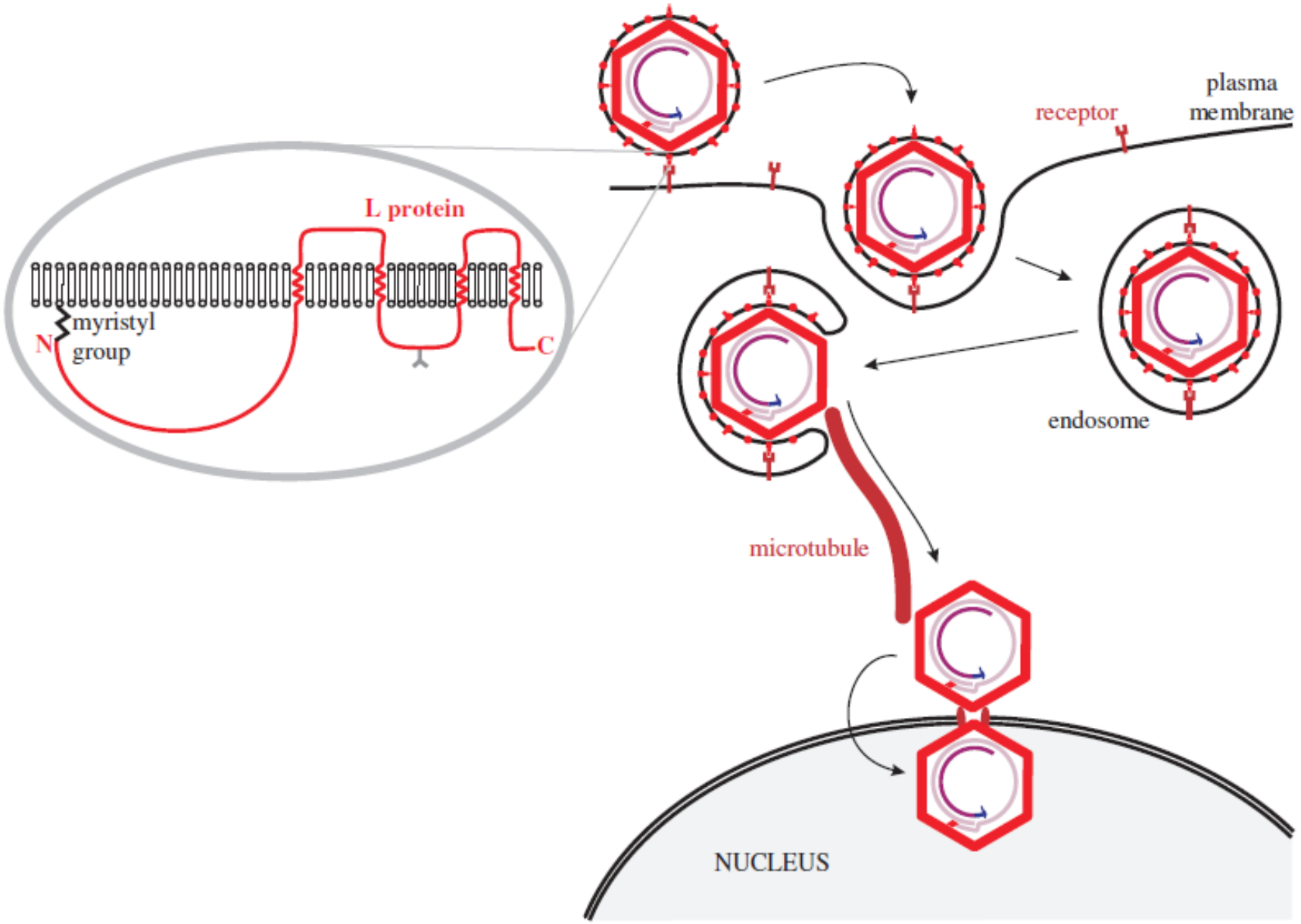


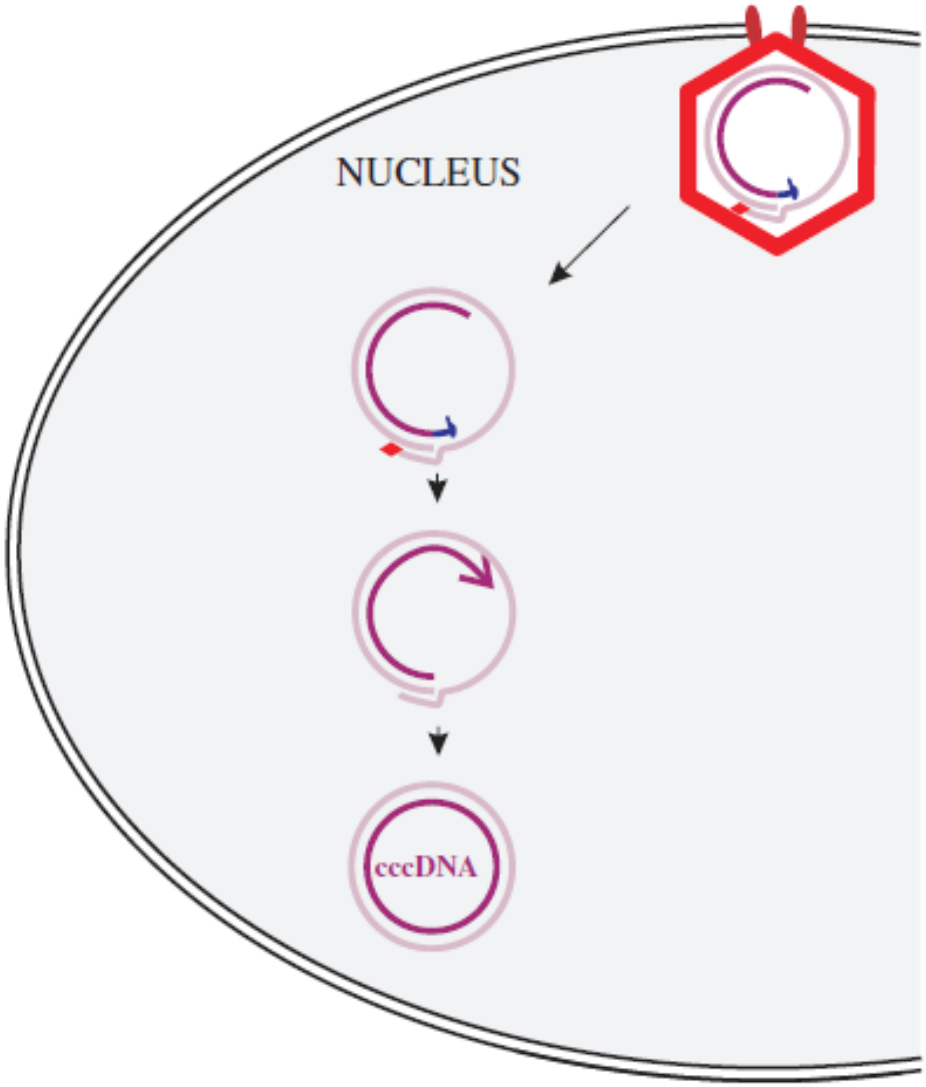
Hepatitis B genome organization

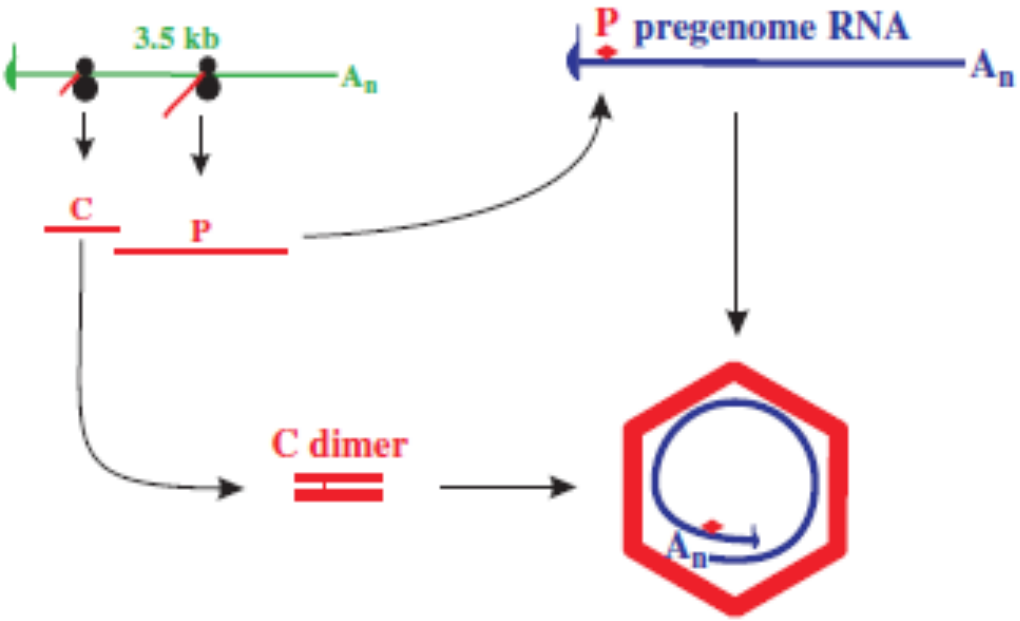


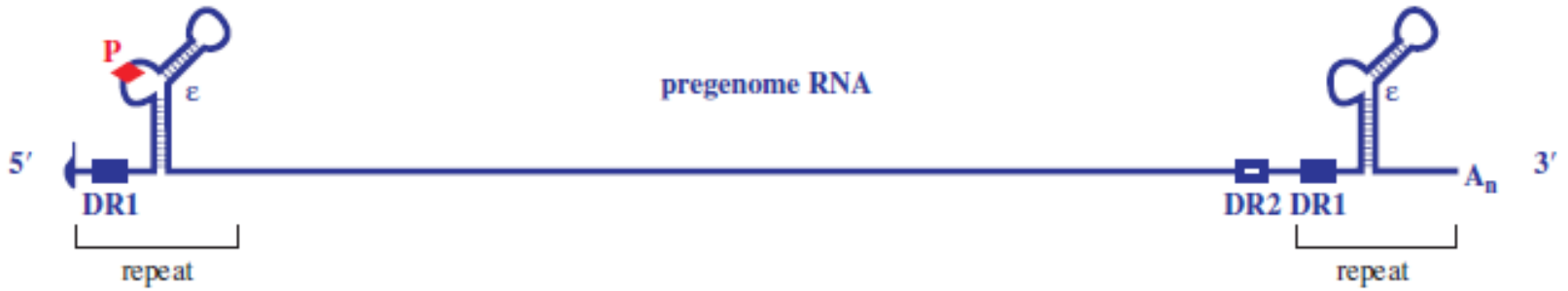
HBV S, M, L proteins

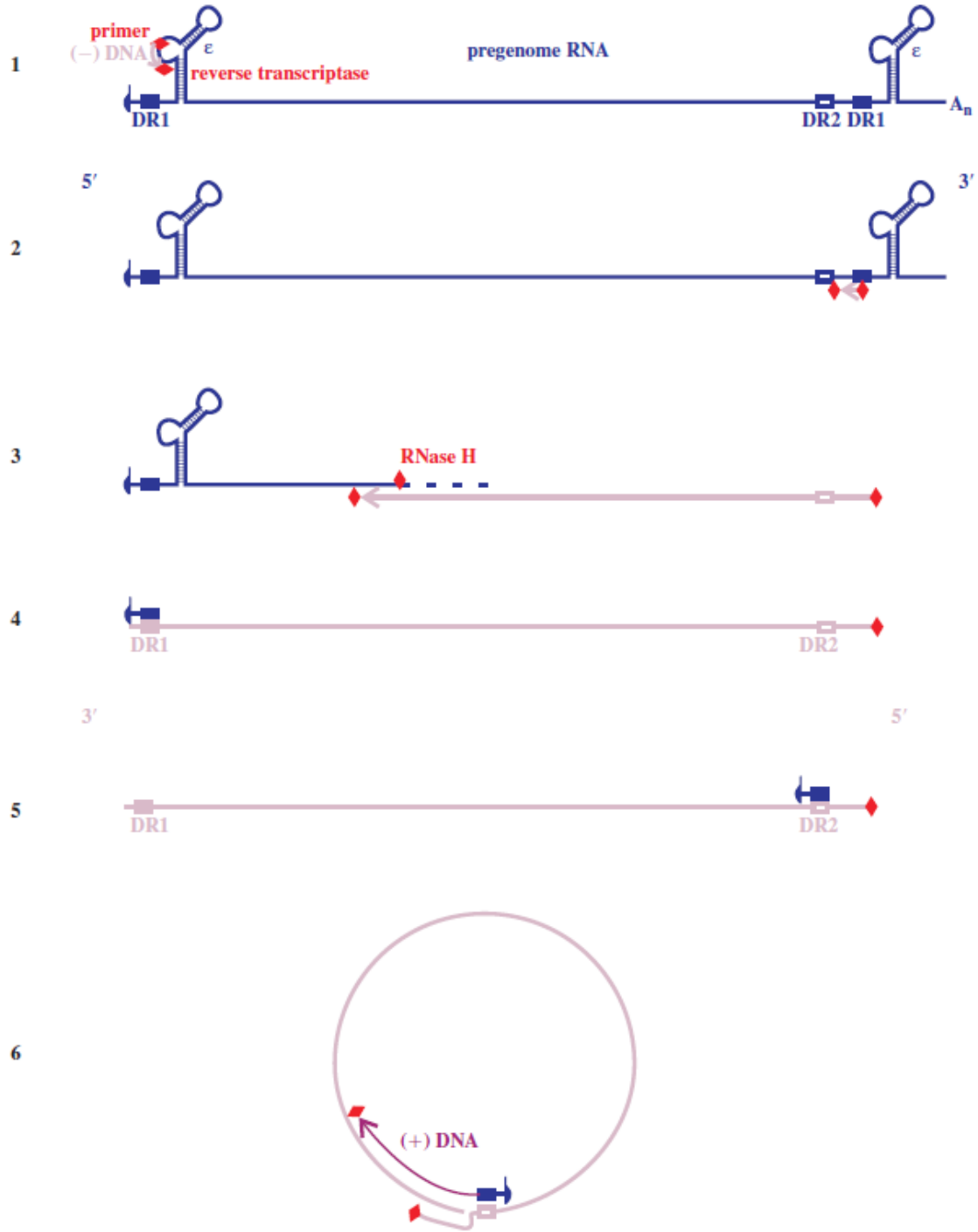


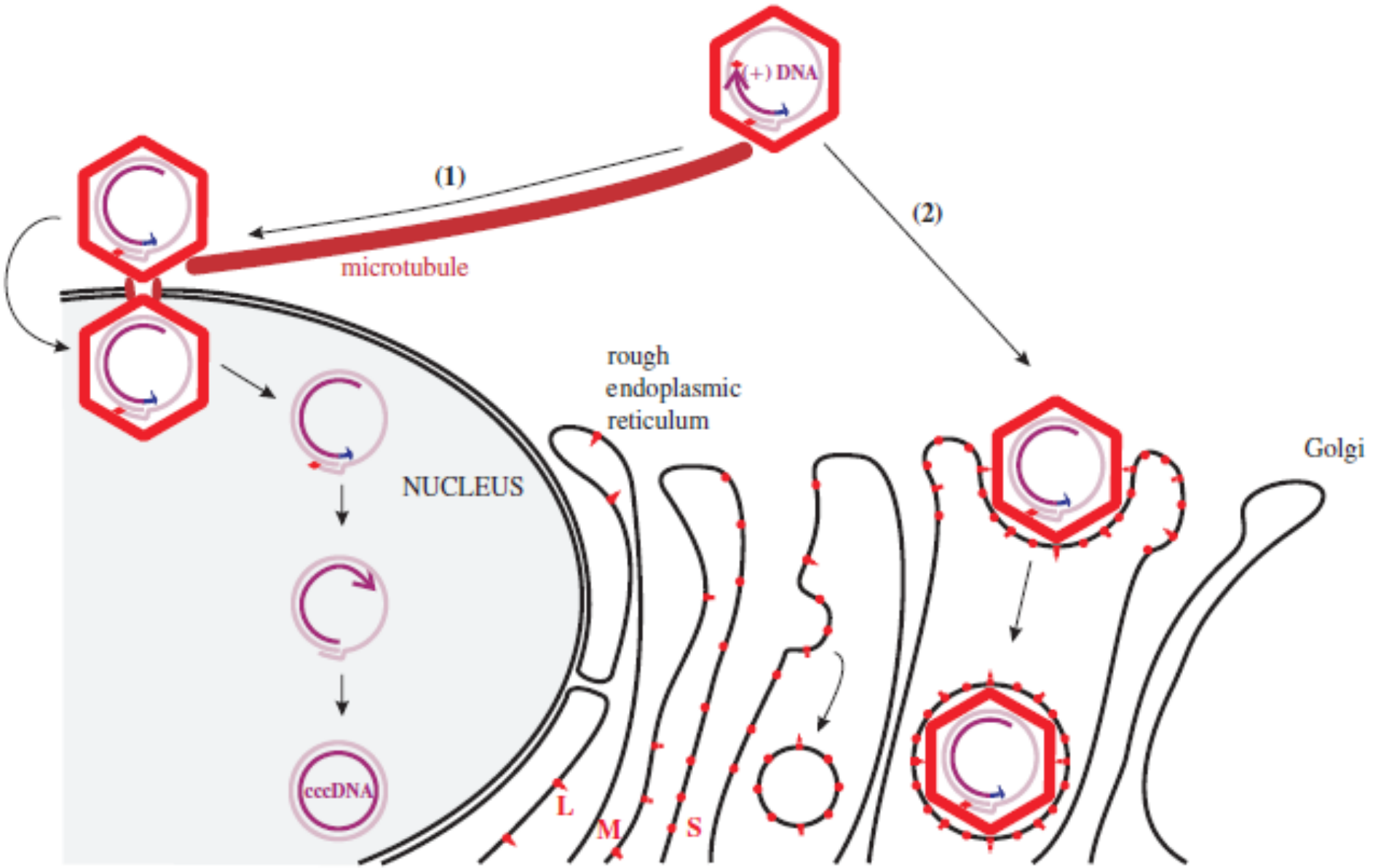

L protein
myristyl
group
N
C
plasma
membrane
receptor
endosome
microtubule
NUCLEUS

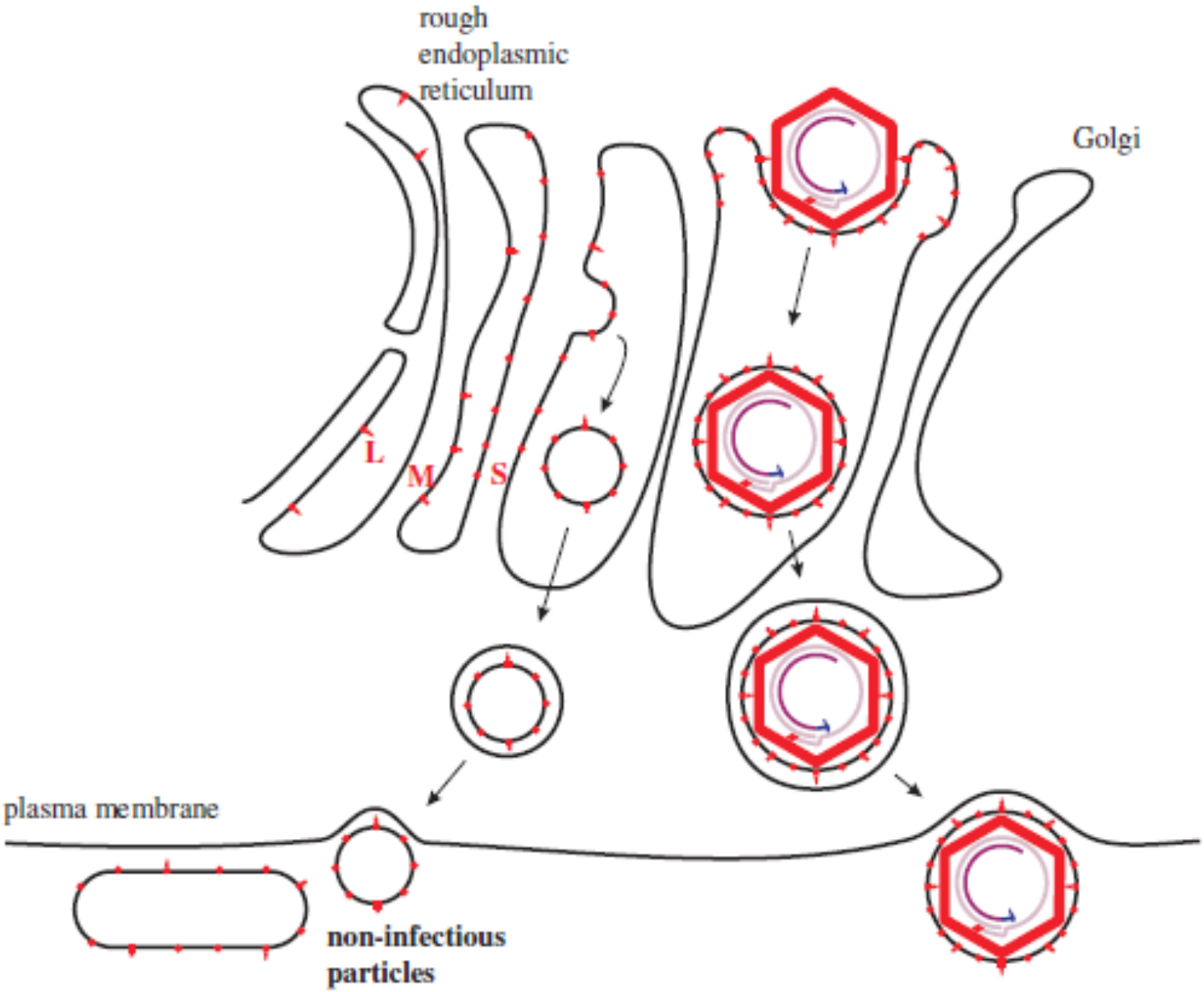

NUCLEUS
cccDNA

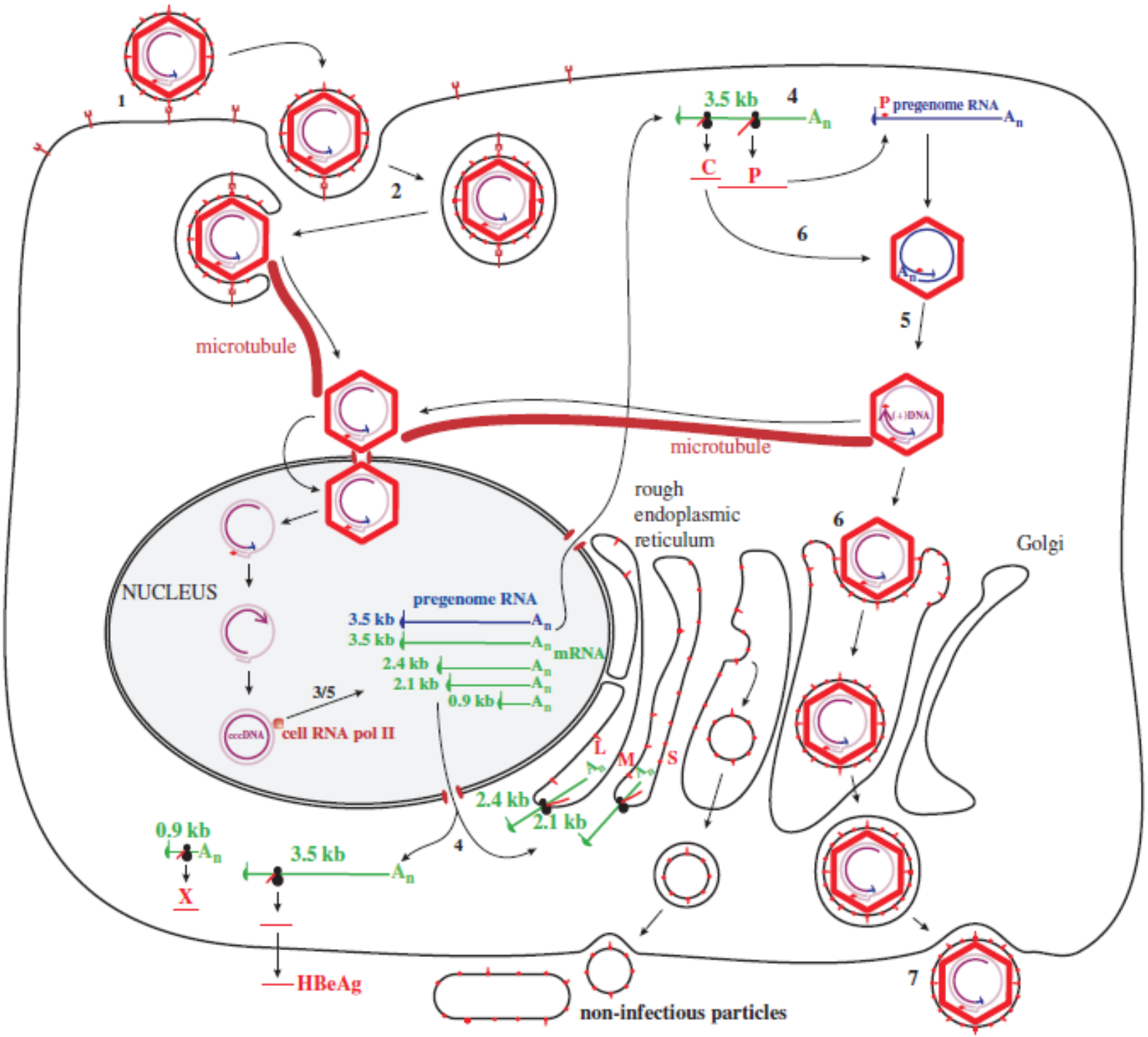

3.5 kb
An
P pregenome RNA
An
C
P
C dimer
An


P
ε
pregenome RNA
ε
5′
An
3′
DR1
DR2 DR1
repeat
repeat


1
primer
(–) DNA
ε
reverse transcriptase
pregenome RNA
ε
DR1
DR2 DR1
An
5′
3′
2
3
RNase H
4
DR1
DR2
3′
5′
5
DR1
DR2
6
(+) DNA


(+) DNA
(1)
(2)
microtubule
rough
endoplasmic
reticulum
Golgi
NUCLEUS
cccDNA
L
M
S


rough
endoplasmic
reticulum
Golgi
L
M
S
plasma membrane
non-infectious
particles


1
2
3.5 kb
4
An
P pregenome RNA
An
C
P
6
5
microtubule
microtubule
rough
endoplasmic
reticulum
Golgi
6
NUCLEUS
pregenome RNA
3.5 kb
An
3.5 kb
An
mRNA
2.4 kb
An
2.1 kb
An
0.9 kb
An
3/5
cccDNA
cell RNA pol II
L
M
S
2.4 kb
2.1 kb
4
0.9 kb
An
X
3.5 kb
An
HBeAg
non-infectious particles
7

Learning outcomes

- explain the importance of HBV
- describe the HBV virion and non-infectious particles
- outline the main features of the HBV genome
- describe the HBV replication cycle
- evaluate means of preventing and treating HBV infection