Chapter 26 RNA metabolism

26.0 Intro

RNA appears to be very similar to DNA
  has the 2’ OH on the ribose
  has U instead of T

Most RNA functions as single stand
  so folds back on itself
  Gives more structural diversity

RNA functions in information storage, information transmission and as enzyme.
  Suspect it may have been ‘The’ molecule in primitive cell

Will see RNA interaction with protein for most functions
  RNA will have both structural and catalytic role in these structure

All RNA except RNA of RNA viruses derived from DNA.
  DNA sequence copied onto RNA through process of transcription

3 major kinds of RNA
  Messenger mRNA carries genetic information from nucleus to ribosome for translation
  Transfer tRNA associates specific AA with specific triplets on mRNA
  Ribosomal rRNA integral structural part of ribosome

Also lots of RNA that does not fit above categories
  Regulatory, catalytic functions or precursors
  In mammalian cells there are more of these than m,t,&r!

Transcription more selective than replication
  will only transcribe one gene or one group of genes, not entire DNA
  regulation of this process described in chapter 28

Sum of all RNA synthesized in cell is called transcriptome

From last chapter know that in mammalian cells the amount of DNA that actually codes for protein is small, so might expect transcriptome to be small. But you would be wrong. It is actually quite large. A wide range of special function RNA’s are made, and we really don’t know what all of them do!
Here will concentrate on synthesis of RNA from DNA and processing and turnover of RNA molecules

Will also look at RNA template/ DNA copy system

26.1 DNA Dependent RNA synthesis
Will compare to DNA that we just did similar
chemical mechanism
polarity
use of template
initiation, elongation, termination stages
different
does NOT require a primer
uses a limited segment of DNA
only 1 strand of DNA is the template

A. RNA synthesized by RNA polymerases
DNA dependent RNA polymerase
Isolated by 4 different groups in 1960's
Needs DNA template
4 NTP's
Mg$^{2+}$
Protein also binds Zn$^{2+}$!

Diagram figure 26-1a
Add units to the 3' end (same mech as DNA, the OH acts as a nucleophile to attach the $\alpha$ PO$_4$ so releases PP$_i$
So 5'→3' polarity

\[(\text{NMP})_n + \text{NTP} \rightarrow (\text{NMP})_{n+1} + \text{PPi}\]

Requires DNA for activity
Most active on double stranded DNA

Most active when bound to double stand DNA, yet only 1 of 2 strands copied in 3'→5' direction using template

No primer required
RNA polymerase binds at specific sequences called promoters
5'PPP of first residue not cleaved but remains

Only about 8 bp of RNA are bound to DNA before peels off and DNA duplex reforms
E coli RNA polymerase keep about 17 bp of DNA unwound for transcription

Proceeds at a rate of 50-90 nucleotides/sec
(DNA Pol III 1000/sec DNA Pol I 16-20/sec)

DNA can’t unwind this fast so observe + supercoiled ahead and - supercoiled behind (26.1B,C)
So need topoisomerases to relieve strain

2 strand of DNA different functions
  Template strand - the one RNA codes from
  Non-template strand or coding strand - identical in sequence to RNA except for T/U switch
Either stand can be template strand
  Regulatory sequences, by convention are the sequence in the nontemplate (coding) strand

DNA dependent RNA polymerase in E coli
  5 core subunits $\alpha_2\beta\beta'\omega$ 390,000 MW
  A 6th unit $\sigma$
    Part of a family of similar units
    Bind transiently to core, to direct binding to different promoters
    Most common $\sigma$ is $\sigma^{70}$ (70,000 MW)
  No proofreading ability. 1 mistake in $10^4$ to $10^5$
  Not as critical
    Many RNA’s made from 1 DNA
    RNA’s are used for a while then degraded, so not permanent
  Have observed RNA polymerase pausing after incorporating
    Mismatch, and reaction can be reversed,
    But don’t know if this is real or artifact

**B. RNA synthesis initiated at promoters**
RNA is initiated at specific sites called promoters
many differ promoters, so are of intense research

In E coli promoter region runs from about 70 bp upstream to 30 bp downstream from actual initiation site

Promoter for $\sigma^{70}$ has features in figure 26-5
  2 short similar sequenced at -10 and -35
    (10 and 35 before the transcribed sequence)
Not always the same but get a consensus
-10 is TATAAT
-35 is TTGACA
Highly expressed genes may have additional AT-rich upstream
recognition element called UP in -40 to -60
Bound by α subunit of polymerase
Efficiency that bind and initiate transcription determined by how close
sequence is to consensus, if they are spaced correctly, and how far
they are from actual start of transcription

A change of 1 base in -10 and -35 can drop efficiency 10-100 times

Overall initiation broken into 2 phases Figure 26-6
Binding - binding of RNA polymerase to promoter on DNA
  Directed by σ factor
  First a closed complex (DNA is intact)
  Then DNA is unwound in -10 region to form ‘open’ complex
Initiation
  Movement to the actual start place of transcription
  Start of transcription
  σ drops off randomly after elongation begins
  Is replaced with NusA protein
  NusA only comes off when transcription is complete

Other σ factors Table 26-1
  other promoter than one bound by σ^{70}
in previous chapter talked about heat shock proteins
these are made when σ^{32} is expressed
Use of different σ used to coordinate expression of different sets of
genes

C. Transcription is Regulated at several levels
Transcription regulated so only those genes needed by cell are expressed
But most control is targeted at regulation of RNA polymerase binding,
Additional regulation occur in all steps of transcription including:
  initiation, elongation and termination
Control mechanisms
  1. Different σ bind to different promoters
  2. binding of regulatory proteins
    Activators - bind to a DNA to increase binding of polymerase
cAMP receptor protein (CRP) expressed when cell
doesn’t have glucose as a source of E, Turn on
genes needed to utilize other sugars
Repressors - bind to DNA to block synthesis of RNA
Lac repressor. if lactose not available, block
expression of these genes (will see in chapter 28)

D. Specific Sequences Signal Termination of RNA Synthesis
RNA synthesis is highly processive - Protein can't fall off till job is done
Certain DNA sequence pause in RNA synthesis
If pause is long enough, will terminate
not well understood in eukaryotes so will look at E coli
2 mechanisms
- \( \rho \)-independent (Figure 26-7a)
  RNA contain self complementary sequences
  Hairpin form 15-20 based before end of strand
  End of message has 3 conserved A's that are transcribed to U's
  Polymerase hits poly A it pauses
  When pauses hairpin forms
  Hairpin disrupts RNA/polymerase binding and complex falls off

- \( \rho \)- dependent Figure 26-7b
  No poly A
  Usually a CA rich sequence called 'rut'
  Rho utilization element
  \( \rho \) protein binds to RNA at a specific site
  Move 5'→3' on RNA until it reach polymerase that is paused
  In releases the RNA polymerase with use of ATP
  Mech not known

E. Eukaryotic cells have three kinds of Nuclear RNA polymerases
More complicated system
3 different complexes - share several components
  RNA polymerase I (pol I) - only for synthesis of preribosomal RNA
    1 long piece that contains 3 smaller pieces.
    The the pieces are 3 different ribosomal RNA’s
      Called 18S 5.8S and 28S
      more later in chapter
    Great variety in sequence of promoters between species

  RNA polymerase II (Pol II) primarily for synthesis of mRNA’s
    And some specialized RNA’s
    Recognizes thousands of promoters
    Common feature is TATA box at -30 and Inr sequence at +1
    See figure 26-8
RNA polymerase III (Pol III) makes tRNA’s, 5S rRNA and some specialized RNA’s.
  5S is a fourth rRNA and ~5.8s
Promoters well characterized
Some control points are inside the gene rather than at the starting end!

**F. RNA Polymerase II requires many other proteins**
Polymerase itself 12 subunits (prokaryote was only 5)
Largest subunit, RBP1, high degree of homology with bacterial β’
  Carboxy terminal tail with many repeats of sequence YSPTSPS
  27 repeats, 18 exact in yeast
  52 repeats, 21 exact in mouse and human
Give domain special name
  CTD (carboxyl-terminal domain)
  Separated from main body with unstructured region
  *Will see several special roles for this region*
RBP2 similar to bacterial β
RBP 3 and RBP 11 - some similarity to bacterial α

Whole host of other proteins called *transcription factors (TF...)*
To form active transcription complex
  General Transcription factors - required for all pol II promoters
  TFIIA -transcription factor A for Pol II etc
  Highly conserved for all eukaryotes
Overall transcription split into many phases
  Assembly
  Initiation
  Elongation
  Termination
*Keep an eye on figure 26-9 and table 26-2 as discuss details*
In process many proteins may be present in preassembled complexes

**I. Assembly**
  TBP (TATA binding protein) binds to TATA box
  (If no TATA box TBP comes in with TFIIID, but this process not well understood and not in figure 26-9)
  Bind TFIIB to TBP - this enlarges extend of DNA binding
  Bind TFIIA not essential but tightens binding at weaker sites
  Bind TFIIF and RNA polymerase II (This factor helps make more specific to site)
  Finally TFIIE and TFIIH bind to make complete closed complex
  H is a helicase - starts unwinding DNA near initiation site
  Requires ATP
Now get open complex
Didn’t show you all the details actually >30 peptides involved!

II. Initiation
TFIIH also acts as a kinase
   Phosphorylates several sites on Pol II in CTD
   This activates pol II to start transcription
   CTD can also be phosphorylated by other enzymes like CDK9
   Phosphorylation may have other effects as well in elongation
As first 60-70 nucleotides synthesized TFIIE and TFIIH are released factors marks entry into elongation

III. Elongation
TFIIF stays attached
Other factors - elongation factors - bind to enhance activity
(Don’t confuse with protein synthesis elongation factors !)
Some of these factors bind to phosphates on CTD tail
   Keep elongation from pausing
Some factors involved in post transcriptional processing

IV. Termination
When transcription is complete and complex terminated, PolII is dephosphorylated so can begin another transcript Not understood

V. Regulation
Elaborate, involved many proteins with initiation complex
Will study details in chapter 28
Some proteins interact with transcription factors
   Some with PolII itself

VI. TFIIH - Many functions
   Have seen several activities
   Helicase to open DNA and kinase to phosphorylate Pol II
   Is a complex and some subunits of complex are also used in excision repair of DNA
   If transcription hits a damaged DNA TFIIH initiates repair mech
   Lack of TFIIH associated with some human diseases
      Xeroderma pigmentosum and Cockaynes syndrome
**G. DNA-dependent RNA polymerase can be selectively inhibited**

Inhibited by actinomycin D in both prokaryotes and eukaryotes

**Figure 26-10**

inserts between 2 GC pairs in DNA

prevents movement of polymerase

this is only thing it does so can be used as diagnostic for things linked to RNA synthesis

Acridine is the same

Rifampicin - binds to β subunit of bacterial RNA polymerase

Prevents release of promoter

Sometimes used as an antibiotic

Mushroom Amanita phalloides makes α-amanitin

Stops mRNA synthesis in animal by blocking RNA pol II can also block pol III in higher conc, but does not block RNA pol I or bacterial RNA polymerase, or the mushrooms own pol II

**26.2 RNA processing**

newly synthesized RNA call **Primary Transcript**

Many RNA molecules in bacteria and almost all RNA in eukaryotes need to be processed after synthesis

- fair amount of process in both pro and eukaryotic tRNA - cleavage of ends and lots of base and sugar modification

- eukaryotic mRNA - cutting out regions and adding caps at both ends

Also the final processing - want to destroy mRNA’s after they have been used!

Some of the enzymes that do this processing are RNA’s not proteins call these ribozymes

Newly synthesized RNA molecule called **primary transcript**

Most processing done in eukaryotic mRNA, pro and eukary t-RNA special function RNA’s also processed

Let’s look at overview **Figure 26-11**

Start with Eukaryotic mRNA

5’ end

Splice

3’ end

Elaborate complexes perform each processing step, sometimes all happening at once, Most all interacting with that phosphorylated CTD tail on polymerase. Even tied with proteins that bind to RNA in nucleus and transport into cytosol and delivered to ribosome

then go to pro and eu t-RNA
A. Eukaryotic mRNA’s capped at 5’ end
   5’ cap - 7-methyl guanine linked to 5’ terminal residue
   (See figure 26-12)
   Protects from ribonucleases
   Used for specific binding interactions in ribosome (next chapter)
   G added first using GTP
   Then methylated
   Additional methylations on 1st and 2nd sugar 2’ OH’s
   Processing done very early
   After first 20 or 30 residues added
   All enzymes needed bound to CTD domain
   Once cap is added
      Release from enzymes but bound to CTD by cap-binding complex (CBC)

B. Both Introns and Exons are transcribed from DNA to RNA
   in bacteria RNA is collinear with proteins sequence and protein can be
   synthesized directly from RNA (in fact Proteins synthesis may be initiated
   before the mRNA is off the DNA) (no separate nucleus so this is easy)

   Genes in eukaryotes contain introns - noncoding regions in RNA
   in vertebrates almost all proteins EXCEPT Histones have introns
   bacteria have a few introns

   In transcription entire DNA sequence introns and exons are transcribed by
   RNA polymerase

   Introns then cut out and exons spliced together

   Weren’t discovered until late 1970’s when tried to hybridize RNA back to
   its DNA

   Most eukaryotic exons <1000 bases, many in 100-200 range
   (30-60 AA’s)
   introns vary from 50 to 20,000 nucleotides
   In higher eukaryotes may have more bases in introns than exons

C. RNA Catalyzed Splicing
   4 types of introns
   I & II share key characteristics but some differences in details of splicing
   I some nuclear, mitochondrial, and chloroplast r, t, and mRNA
   II usually mitochondrial or chloroplast mRNA in fungi algae and plants
   Some I and II found in rare bacterial splicing events
   Neither required ATP for splicing
Both use similar mech - figure 26-13 - a transesterification reaction
OH of one ribose attacks phosphate of another, no E is lost so
that’s why don’t need ATP

Type I mech Figure 26-14
need GTP, but not for energy (can use GMP, or GDP!)
Uses 3'OH of ribose as nucleophile
Forms a 3'-5' bond with the 5' end of the intron
The 3'OH on the exon then attaches the splice point
And get the exon spliced out

Type II mech (figure 26-15)
Uses a 2' OH from an A inside the exon to start it off
End up with a lariat

NO ENZYMES REQUIRED - self splicing!

Third type of splicing - requires protein - Most common mechanism
Called Spliceosomal intron
Because splicing is done in a protein complex called
spliceosome
Figure 26-16 a&b
Largest class
Nuclear mRNA’s
Ends up with lariats (so resembles type II)
But requires RNA-protein complexes called
Small nuclear ribonucleoproteins snRNP’s or ‘snurps’

Each snRPP contains has its own RNA
Called snRNA for small nuclear
snRNA of 100-200 residue
5 snRNA’s (U1,U2,U4,U5,U6) are abundant in nucleus

And several proteins
Both protein and RNA parts are highly conserved

Spliceosomal introns usually
GU at 5' end
AG at 3' end
U1snRNA has sequence complimentary to 5' splice site of many
introns
U1snRNP binds to this region

Then add U2, U4,U5,and U6 snRNP’s to make
Spliceosome
Contains the 5 RNA’s and about 50 proteins
Need ATP so assemble spliceosome, but not for splicing reaction.

A less common spliceosome contains U11 and U12 instead of U1 and U2 snRNP’s

5’ AU….AC(3’) sequence

Some splicing components tethered to CTD?

Figure 26-16c
First 5’ junction and spliceosome tethered to CTD
Then 3’ end comes through and splice removed

After splicing intron remains in nucleus and is degraded

Fourth type of splicing
Used in certain tRNA’s
Requires a splicing endonuclease
Requires ATP
Endonuclease cleaves at both end of intron
2 ends are joined like in a ligase with ATP for E

Type III splicing with spliceosome limited to eukaryotes
Type I & II splicing seen in some bacteria and viruses
Splicing more common in archaea than in bacteria

D. Eukaryotic mRNA also has a distinctive 3’ end
80 to 250 residues of A
Binding site for specific proteins
Some protein help protect from degradation
Some used in interaction with ribosome
Some prokariots also get a poly A tail
But in prokariots this stimulates degradation!

Mech

Figure 26-17
RNA extended past site where put on tail
Cleaved at addition site by an endonuclease
Endonuclease part of large enzyme complex associated with CTD
Site marked by sequence
AAUAAA 10 to 30 residues on 5’ side (upstream)
G-U rich region 20-40 residues downstream
Cleavage leave a 3’ OH open
Polyadenylate polymerase then used ATP to add 80-250 A tail
Does not require a template
Does require a primer
Overall processing summarized in figure 26-18

E. Multiple Products are derived from 1 gene via differential splicing
Splicing seem to be a pretty wasteful process
don’t really know why at this point
some cells have used to their advantage

Most eukaryotic mRNAs produce a single mature RNA
Some can be processed differently to make different mature RNA and different proteins

**COMPLEX TRANSCRIPTS** Figure 26-19
Processing to obtain different mRNA
1 way differentiates for attachment of poly A tail 26-19a
Used in variable domain of immunoglobins
Another way alternate splicing patterns 26-19b
This is observed in 3 different myosins found in fruit flies
Both are used in production of two different hormones from rat thyroid and brain (figure 26-20) (calcitonin & calcitonin-gene-related peptide CGRP)

Many, perhaps most genes in mammalian genomes use alternative splicing
Not used as much in lower eukaryotes

F. **rRNA and tRNA also undergo processing**
rRNA of both prokaryote and eukaryotes made from large precursors that must be cleaved
Also many methyl groups added and bases modified Figure 26-22
Long pieces called *pre-ribosomal RNA*’s

**Bacteria**
5S, 16S and 23S and a few tRNA’s all come from one 30S

*Does anybody know what S means?*
RNA of about 6,500 bases
RNA removed at both ends and in between
16S RNA
11 base modifications
10 bases methylated

23S RNA
11 bases modified
12 methylated

Ecoli has 7 such preribosomal RNA sequences scattered throughout the genome
See figure 26-23
rRNA sequences the same
In between sequences differ
Cleaved by RNAses, not by splicing

Eukaryotes

RNA synthesized using RNA Polymerase I
45S preribosomal RNA processed in nucleolus
Figure 26-24
To make 18,28 and 5.8 rRNA’s
Ribosome also need a 5S rRNA
Entirely separate gene and uses RNA polymerase III

Ribosome assembled in complex in Nucleolus
Includes all nucleases required for cleavages
All enzymes required for base modifications
Up to 200 base modifications
Each relies on a different snoRNA-Protein complex (snoRNP)
Typically 1 snoRNA & 4-5 proteins

Yeast complex 170 nonribosomal proteins 70 snoRNA’s, 78 ribosomal proteins
Humans have more base modifications so even more snoRNA’s

SnoRNA (figure 26-25)
Small nucleolar RNA’s snoRNAs
Guide modification, cleavage and some protein interaction
60-300 nucleotides
Typically an intron from another gene
Contains 10-21 sequence that is complementary to site on rRNA where interaction occurs

t-RNA’s Figure 26-26
Most cells have 40-50 distinct tRNA’s (20 AA’s, 64 anticodons)
Eukaryotic have multiple copies of many t-RNA genes

Usually have to take off RNA from both ends and eukaryotes may have introns to splice out

Sometimes may have 2 or more tRNAs on a single RNA that need to be separated
RNAse P removes 5' end
   Found in all organisms
   Has both protein and RNA
   Can function in bacteria without protein!
3' end removed by other endonucleases including RNaseD

Further processing
   Amino acid will be attached to a CCA sequence on th 3' end
      many bacterial and ALL eukaryotic don’t have this sequence
   Must be added by tRNA nucleotidyl transferase
      Has separate sites for all 3 bases
      Binds of RNA and adds
      No template needed

Base modification
   Methylation, deamindation, reduction of some bases
   Examples figure 26-22
   Occur at characteristic positions in structure
      Sometimes need for structure
   Inosine will be important in next chapter as the ‘wobble’ base

G. Special-function RNA processing
   The variety and number of special function RNA’s is rapidly expanding
      (Was not even in 4th edition of text)
   Many of these RNA’s require processing
   snRNA’s
      SnRNA’s synthesized as large precursor by RNA polymerase II
         Ribonucleases remove excess from ends
            11 base modifications
   snoRNA
      Some snoRNA are introns of other genes
      Again precursor is larger than actual product
      As splicing occurs, sno protein binds and ribonuclese cuts the excess RNA off at both ends
   Micro RNA’s (miRNA) (new 5th edition)
      ~22 nucleotides long
      Complementary to sequence in particular regions of mRNA
      Regulate expression by cleaving or suppressing translation
      Involved in gene regulation (Chapter 28)
      Present in all multicellular organisms
      May be 1% of human genome
         And targets on 1/3 of the mRNA’s
   Figure 26-27
Synthesized as large precursor
Will not worry about details of processing

**H. Some Event in RNA metabolism are catalyzed by RNA enzymes**

3 best studied
- Self splicing of type I introns
- RNase P
- Hammerhead ribozyme

Do transesterification or phosphodiester bond cleavage

Self splicing group I intron 400 bases
RNase P 337 bases
Hammerhead ribozyme total of 41 bases in 2 strands
(Critical portion of larger enzyme but can function alone)
activity depend on structure
Heat or solvent denature and destroy activity

**I. Group I introns**

Had Mechanism earlier [Figure 26-14]
Secondary structure shown 26-28

**Enzymatic properties**
- Acts like enzyme
- Binding of GTP can be saturated, has a Km of 30μM
- Can be competitively inhibited by inhibitor (3’ deoxyG)
- Very precise in activity
- Uses it’s own RNA to make alignments
  - **internal sequence guide**
- Originally thought that self splicing technically not an enzyme because can’t be reused
  - But intron from Teterahymena can guide other splicing events

**II. Characteristic of other Ribozymes**

RNase P
- 337 nucleotides 17,500 MW protein (13%) (~ 160AA’s)
- 111,000 MW RNA (330X 337) (86%)
- 156 AA’s

Under some conditions RNA can work alone
RNA can recognize shape of pre-tRNA and CCA sequence
  (Implies reaction is second in processing?)
Protein stabilizes RNA and facilitates function
Can cleave CCA from diverse sequences
Hammerhead ribozyme
From a plant virusoid
Small RNA associated with a plant virus
Promotes self splicing reaction

I. Cellular mRNA degradation
want RNA synthesis to balance RNA degradation
part of control on protein levels

Rate of RNA degradation vary greatly for different genes
Protein needed briefly, RNA may last a minute or a few seconds
Protein needed continuously may be stable over several cell lifetimes!

In eukaryotes average lifetime about 3 hours
Pool usually exchanged about 10x / generation

Half life of bacterial mRNA about 1.5 min in bacteria

In E coli usually start with one or more cuts by endoribonuclease
Followed by 3'--5' exoribonuclease

In lower eukaryotes
Shorten poly A tail
Decap 5' end
Degrade 5'--3'

All eukaryotes have complex of up to 10 3'--5' exoribonucleases called exosome
Usually involved in processing 3' ends of r- and t- and special function RNA's like sn and sno
May also be major degradation pathway in higher Eukaryotes

Some bacterial mRNA have a hairpin loop in ρ-independent terminator
adds stability

Similar hairpins also seen in more stable primary transcripts

I. Polynucleotide Phosphorylase
A curiosity
discovered in 1955
could synthesize RNA randomly from 5' diphosphates (NOT TRI)
Needed no template
can’t use NTP’s

Reaction is reversible
so think is really and endonuclease

Can be used to synthesize your own random RNA to your specifications

26.3 RNA-Dependent Synthesis of RNA and DNA

So far DNA has been the template
there are some cases where RNA is the template
In particular RNA viruses
So not just DNA to RNA to Protein
but RNA to RNA
and RNA to DNA

A. Reverse Transcriptase produced DNA from viral RNA

Figure 26-32 Retrovirus life cycle
Certain RNA viruses that infect animal cells carry a RNA dependent DNA polymerase or reverse transcriptase in the virus particle
Genome is a single stranded RNA, usually about 10,000 nucleotides long

1st step in infection is to copy viral RNA into DNA
Then degrade RNA and replace with DNA
Then often incorporate DNA into host DNA
When host transcribes this gene then start making viral particles

These are the retroviruses
Predicted to exist in 1960's
Detected in 1970's

Typically have 3 genes Figure 26-33
Gag, pol, and env
Single transcript carries gag and pol, LTR’s & $\psi$
(LTR Long terminal repeat)
Translated as a single protein
Then cleaved into 6 proteins with different functions

Protein from gag portion make up core of viral particle
Pol gene contains
protease to cleave the long peptide
Integrase for integrating DNA into host
Reverse transcriptase

Env gene is viral coat protein
At end of RNA are **long terminal repeats (LTR’s)** of a few hundred bases
Used to help integrate DNA into host genome
Contain promoters for viral genes

Reverse Transcriptase
Many reverse transcriptases have 2 subunits, α&β
β is 90,000
α is 65,000 fragment of beta
Like many other RNA and DNA polymerases, reverse transcriptase contain Zn^{2+}
Works best with its own RNA but can be used with others

Does 3 different reactions
1. Synthesis of DNA on RNA template
2. Degradation of RNA template
3. DNA synthesis on DNA template

DNA/RNA synthase and RNase different parts of enzyme

For DNA synthesis needs a cellular tRNA primer
Is a cellular tRNA carried in virus, that came from previously infected cell
tRNA 3’ end complimentary with a sequence in viral RNA
DNA synthesized in 5’-3’ direction (like everybody else)

Reverse transcriptase does not have a 3’-5’ proofreading
Error rate 1/20,000 bases
Very high error rate seems typical for RNA viruses
(Error rate for HIV virus is even worse!)
That is why RNA viruses have faster viral evolution

Reverse transcriptases useful for biotech industry
If you have a protein you want to express
You isolate the RNA on the ribosome as the protein is being synthesized
Use the reverse transcriptase to make back into DNA for your work

**B. Retroviruses cause cancer and AIDS**
most retroviruses do not kill host, but integrate into host genome and get expressed when host divides

Some viruses - RNA tumor viruses - contain an oncogene
Cause cell to grow abnormally
Human immunodeficiency virus (HIV) causes acquired immune deficiency syndrome (AIDS) - a retrovirus
A fairly standard retrovirus with a few extra genes
Not standard because kill host cells (T lymphocytes) instead of making tumors
By killing T lymphocytes kills immune system
Not standard because reverse transcriptase 10x less accurate than most
1 or more errors in very copy!
Any 2 viral copies are different
Makes hard to make a vaccine because coat protein is always changing

C. Transposons, Retroviruses and Introns - a Common origin?
Some Transposons (See DNA chapter) have structure very similar to retroviruses
Sometimes transposons closely resemble retroviruses called retrotransposons
Code for an enzyme homologous to reverse transcriptase
Flanked by LTR’s
Transpose from one location to another using RNA intermediate (RNA→RNA/DNA→DNA/DNA)
Probably using their reverse transcriptase
Most eukaryotic transposons use this mech
Prokaryotic use a different mech (DNA intermediate)

Retrotransposons lack an env gene so no viral particle
A defective virus caught in a cell!? (poor thing)

Sequence comparisons suggest that reverse transcriptase is ancient enzyme
May predate multicellular organisms

Many Group I and Group II introns are also genetically mobile!
Spliced out intron codes for a DNA endonuclease

Allows these intron to be introduced into homologous genes that don’t have introns

Process called **homing** (figure 26-37)
Group I homing is DNA based
Group II homing is RNA based

Suggest introns are artifact of ancient retrovirus and transposons???
D. Telomerase is a specialized reverse transcriptase

Telomere - a specialized structure at the ends of linear eukaryotic chromosomes

generally many copies of short oligonucleotide sequence
\[ T_xG_y \] on one 3’ strand \[ C_yA_x \] on 5’ x and y between 1 and 4

Range in length form a few dozen base pairs (ciliated protozoa) to 10 of thousands in mammals
TG side longer than other side so has region of single stranded DNA
Can be up to a few hundred bases on 3’ end of strand

Ends of DNA not easily replicated by cellular polymerase
DNA replication needs template and primer

Look back at figure 25-12 synthesis of okazaki fragments

How do you replicate the 3’ strand and an end?
Need to get an RNA primer
RNA primer then cleaved off
This would shorten the end every time!!
(The 5’ end is OK because Okazaki fragment formed in interior then DNA out to end)

This problem solved by telomerase that adds telomeres to end of DNA

Uses an unusual mech
Is a RNA/protein complex
RNA part 150 nucleotides
Contains about 1.5 copies of a \[ C_yA_x \] repeat

Figure 26-38

Its RNA base pairs with 3’ end of DNA
Uses its own RNA as template for making DNA to fill in missing end on 3’ DNA
Thus is a reverse transcriptase
But unlike others can only use its own RNA as a short template
After extension of \[ T_xG_y \] strand, \[ C_yA_x \] synthesized by polymerases
Single stranded region protected by specific binding protein in lower Eukariots (especially if < a few 100 bp)
In higher eukaryots >1000 bp
Single strand end sequestered in T-loop Figure 26-38b
Loop bound by protein TRF1 & TRF2 to protect from nucleases and repair enzymes
Telomeres and aging
Protozoans can lose telomerase activity
If this happens genes get shorter and shorter
Eventually cell line dies

Also happens in cell lines cultured from humans
Germ-lines - cell lines that reproduce forever
Contain telomerase activity
Try to culture somatic cells
They lack telomerase activity
They die quickly
Fibroblast
See a link between age of individual cell is take from
and Length of telomere
If introduce telomerase activity into the cell line
Cells live much longer

Definitely an area for future research

E. Some Viral RNA’s are replicated by RNA dependent RNA polymerase
some E. coli bacteriophages, and some eukariotic viruses contain RNA
genomes
RNA serves as mRNA for viral proteins
RNA synthesized by an RNA dependent RNA polymerase
Never a DNA intermediate, that is why not called a retrovirus

Protein is 210,000 MW
Has 4 subunits
One subunit 65,000 MW
Active site of replication
Product of viral rep gene
3 other units normally part of e. coli protein synthesis!
Tu and Ts factors used in carrying tRNA to ribosome
S1 normal part of 30S ribosome
? help rep protein find and binds viral RNA??

Makes RNA complementary to viral RNA
Looks like a normal polymerase except requirement or RNA
template
No proofreading ability so high error rate
Will not replicate DNA
Will not replicate other RNA’s, specific for it’s own
F. RNA Synthesis offers clues to evolution
interesting speculations but few hard facts for teaching
but it is an interesting read