

Prelim review, BIOM 555

Miler T. Lee

24 March 2005

1 Control of DNA replication (Calvi)

1.1 Eukaryotic elongation

Replication is a semi-conservative, semi-discontinuous process

1. Helicase unwinds DNA at the fork, which creates supercoiling ahead of the fork
2. TopI/TopII remove the supercoils, decatenate
3. RPA (Replication protein A) binds to single stranded DNA to prevent breakage
4. Leading strand initiated by RNA primer made by Primase, which makes a short RNA fragment then a DNA fragment
5. RFC recognizes the DNA fragment, loads PCNA sliding clamp
6. PCNA clamps around the DNA making contact with DNA Pol δ and ϵ , which perform 5' to 3' synthesis
7. Lagging strand involves piecewise replication: RNaseH, FEN1 remove the RNA primer, ligase ligates the Okazaki fragments together

1.2 Initiation

DNA replication is initiated at origins, a term that refers to both the physical replication site and the sequence motif required for initiation. This is a major point for control of replication:

1. Origin recognition and binding by Initiator
2. Origin unwinding
3. Recruitment of replication fork proteins, which establishes two forks that go outward

1.3 E. coli origins

E. coli have a single discrete origin at OriC site – site of sequential loading of proteins.

1. DnaA recognizes and binds the A boxes at the origin, initially melts the DNA
2. DnaC loads DnaB (helicase), which wraps around the DNA and further melts
3. terminates at Ter site

1.4 *S. cerevisiae* origins

S. cerevisiae have multiple origins, ~420. The pre-replicative complex (pre-RC) binds an origin in G1:

1. ARS consensus sequence (part of A element) is bound by Origin recognition complex (ORC 1-6), which recognizes the A and B1 elements and serves as a scaffold for other proteins to load
2. CDT1, CDC6 load, which in turn load MCM (mini chromosome maintenance) helicase complex
3. ABF1 transcription factor binds to B3 stimulates the origin

1.5 Control of multiple origins

- G1 nucleus + S phase cell induced to replicate DNA: S phase has diffusible factor (cdks)
- G2 nucleus + S phase cell not induced to replicate DNA: rereplication block mechanisms in G2 nucleus

Pre-RC regulation is mediated by CDK concentrations during the cell cycle. Re-replication is inhibited during S, G2, and early M phases in a semi-redundant fashion.

1. M-G1 boundary, CDKs are low, pre-RC assembles onto origins: "Licensing"
2. late G1, CDKs high, CDC7 kinase stimulate initiation
3. S-phase, pre-RC is destroyed, so license is lost
4. G2, CDKs inhibit reassembly of pre-RC and geminin binds CDT1 to prevent reassembly semi-redundantly

1.6 Metazoan origins

There is no known consensus sequence for origins in animals. ORC binds DNA non-specifically in vitro, though replication does begin at specific preferred sites on chromosomes.

- origins are large and modular with partial functional redundancy (e.g., deletions 50kb away can reduce initiation)
- origins have different strengths, so some are passively replicated by incoming forks
- origins have different timings for when they fire (late origins tend to be in heterchromatin)
- location of origins can change during development
- chromatin structure influences both origin firing and elongation; often maintained on daughter chromosomes

2 Prokaryotic extrachromosomal elements (Yuk)

2.1 Plasmids

- autonomous replicating DNA
- accessory functions: antibiotic resistance, virulence factors, metabolism of xenobiotics
- Host range can be narrow or broad, determined by ori region and genes present on plasmid (broad range bacteria have more genes)
- Plasmids can be maintained by addiction: e.g., F plasmid encoding Ccd proteins, one very stable toxin, one labile antitoxin
- Plasmid copy number is regulated; high copy (relaxed) or low copy (stringent)

2.2 Regulation by antisense RNA

- RNA II is an RNA primer that initiates normal replication
- Under normal conditions, rop protein is translated
- RNA I promoter lies on the opposite strand from RNA II and encodes an RNA that overlaps the coding region of RNA II
- at high concentrations of plasmid, there is accumulation of RNA I and Rop; RNA I will duplex with RNA II, assisted by Rop, thus inhibiting initiation of DNA synthesis by RNA II
- deleting Rop can increase the copy number of the plasmid

2.3 Regulation by antisense RNA and protein

- primary player is protein RepA, binds to ori to initiate DNA synthesis
- repA transcript can be transcribed from either repA promoter or an upstream copB promoter
- copB promoter makes dicistronic transcript, also codes for CopB protein
- CopB inhibits repA promoter, allowing antisense product to be transcribed from copA promoter that overlaps the repA transcript
- double stranded rna is formed and degraded

2.4 Regulation by protein, iterons

- different RepA required for DNA synthesis
- iterons: short repeat sequences near ori
- repA has own feedback loop, represses its own promoter at high concentrations
- additional exogenous RepA does not increase plasmid copy number
- exogenous iterons inhibit replication of the plasmid
- high concentrations of RepA also “handcuff” plasmids by binding to iterons and multiplexing the plasmids, inhibiting synthesis
- mutation screen identified repA mutants that could increase copy number, as they were mutated only in the handcuffing domains

2.5 Incompatibility groups

Plasmids in the same incompatibility group share the same method of replication control, so cannot coexist in the same bacterium – over time, one or the other plasmid will be lost

3 Prokaryotic transcription (Decatur)

3.1 RNA polymerase

RNAP core enzyme has five subunits – β' , β , α , α , ω . Core binds to a sigma factor to form holoenzyme.

3.2 Sigma factors

- dissociable subunit, swapped depending on situation
- each sigma factor has different promoter specificity
- binds to core and recognizes specific DNA sequences in promoter
- melts DNA (open complex)
- ejected during elongation
- major vegetative sigma is σ_{70}

3.3 bacterial promoter

- consists of -35, -10 recognition elements with 17-base spacer, whose sequence is unimportant
- basal level transcription strength depends on the particular sequence of -35 and -10 sites

3.4 σ only binds DNA when associated with core

- sigma-70 has a long N-term region 1.1
- blocks region 4.2 from binding to -35 sequence
- does not block region 2.4 from interacting with -10
- sigma without 1.1 region will bind specifically to DNA
- nontethered trans 1.1 regions can inhibit binding of sigma
- once bound to core, the 1.1 region folds away so binding interaction can occur
- non-binding without core prevents sigma from being associated with DNA unproductively

3.5 UP elements

The c-term of alpha can bind specifically to UP elements, which are AT-rich sequences upstream of -35 in certain promoters (e.g., some rRNA promoters). Binding increases transcription levels in the absence of activators, since the additional DNA interaction strengthens RNAP binding.

3.6 alternative sigma factors

- way to reprogram gene expression
- produced in response to changing environmental conditions, developmental stages
- different sigma factors have different intrinsic affinities for core
- interactions between sigma factors and core can be modulated by other proteins or small RNAs

3.7 anti-sigma factors

- binding of anti sigma factor can sequester sigma from core RNAP, e.g. SpoIIAB and SigF
- anti-sigma factor can also pull away sigma from core, e.g. FlgM and Sig28

3.8 CAP

CAP is a transcriptional activator. It is a homodimer that binds to a perfect inverted repeat upstream of the core binding site via a helix-turn-helix motif.

- Class I CAP dependent promoter: CAP binding site at -61
- Class II CAP dependent promoter: CAP overlaps -35 site upstream

3.9 Class I CAP-dependent promoters

AR1 is the activating region on CAP. AR1 interacts with RNAP α -C terminal domain. AR1 must be function on the proximal subunit of CAP, irrelevant in the distal subunit.

alphaCTD will bind nonspecifically to the DNA between CAP and the rest of the core enzyme. The tether between α -CTD and α -NTD allows α -CTD to stretch and bind specifically to an UP element, with a correspondingly upstream CAP binding event.

3.10 Class II CAP-dependent promoters

In the absence of CAP, α -CTD can make nonspecific DNA interactions with the CAP binding site. CAP binds between α -CTD and the rest of core. AR1 on the distal subunit of CAP binds α -CTD. AR2 on the proximal subunit binds to the α -NTD. In addition, AR3 interacts with sigma.

4 Prokaryotic and eukaryotic DNA repair (Brown)

4.1 Types of DNA damage

4.1.1 Base hydrolysis

- Caused by reaction with H_2O , at biological pH (7.4)
- Leads to mismatches (deamination) and stalled DNA replication and transcription (depurination)
- Prevalence
 - Depurination: 5000 bases/genome/day
 - Depyrimidation: 300 bases/genome/day
 - Cytosine deamination: 100 bases/genome/day
- Repair: Base excision repair (predominantly), Mismatch repair

4.1.2 Oxidative DNA damage

- Reaction: $RH_2 + \bullet OH \rightarrow \bullet RH + H_2O$
- Various base damage (thymine glycol, 8-oxoguanine) and strand breaks (electron attacks ribose)
- Repair: Base excision repair, SSBR

4.1.3 Mismatches

- Causes
 - Errors during replication (E. coli DNA polymerase with exonuclease error rate = 10^{-7} , mammalian 600 mismatched bases/cycle)
 - Base deamination (5-methyl-cytosine deamination to thymine creates a T/G mismatch)
 - Translesion synthesis: error-prone replication to bypass damaged bases during DNA replication, puts in a patch; counts on base repair to come by later to correct the base
- Repair: Mismatch repair

4.1.4 Ultraviolet light

- Photoelectric Diels-Alder reaction, converts neighboring Pyrimidines to a cyclobutane dimer; equilibrium, though favored to the right
- prevalence: 10,000 dimers/cell/sunburn
- Repair: Nucleotide excision repair

4.1.5 Strand breaks

- Causes: oxygen radicals, ionizing radiation, β -elimination
- Single-strand or double-strand; double-strand breaks easy during DNA replication, when one hit on the template strand can cause a DSB
- breaks can cause translocations, loss of heterozygosity

4.2 DNA repair mechanisms

4.2.1 Base excision repair

- Mechanism
 1. Detection of Uracil where deaminated Cytosine was; DNA glycosylase flips out the base and cleaves it off, leaving an intact ribose bond
 2. Excise the ribose using AP endonuclease
 3. DNA polymerase β (for single base) fills in the base
 4. Ligation via XRCC1/LIG1
- Genes
 - Glycosylases, very well conserved: 8-20 specific for different common types of damage (UNG, Uracil N-glycosylase; OGG1, 8-oxoguanine glycosylase; NEIL1, removes thymine glycol); some overlap of functionality
 - AP endonuclease: Ape1, Ape2 – general mechanism for other pathways too
 - Polymerase: β (for short patch), δ/ϵ (for long patch)

4.2.2 Mismatch repair

Prokaryotic

1. **MutS** (ADP-bound homodimer) recognizes a mismatch based on the bulge that it creates in the double strand, binds to the bulge; ATP binds; in vitro, you can block a mismatched-strand of DNA with a block on both ends, and the proteins keep binding onto the DNA – dynamic process
2. **MutL** homodimer binds and stabilizes MutS
3. **MutH** is recruited, recognizes a methyl group on the older strand somewhere downstream; MutS/MutL activate MutH to nick the DNA on the newer strand; entire complex dissociates (ATP → ADP)
4. **UvrD** helicase to unwind the DNA
5. **Exonuclease I or VII** comes in to chew back the cleaved strand
6. Resynthesis of the DNA (DNA polymerase) and ligation (ligase)

Eukaryotic probably following right behind DNA polymerase

1. MutS homologs heterodimerize to bind to different sorts of lesions
 - **MSH2-MSH6**: base-base mispair, 1-base pair loop
 - **MSH2-MSH3**: 1 base pair loop, 2-4 base pair loop
2. **MLH1-PMS2** is recruited, complex runs into DNA polymerase (as opposed to a MutH homolog)
3. helicase, exonuclease, DNA polymerase

4.2.3 Nucleotide excision repair

Global genome repair fixes a dimer in either strand, so not transcriptionally dependent

1. XPC-HR23B heterodimer recognizes the dimer and opens the double helix
2. XPB, XPD and TFIIH bind and open the DNA further, 20-30 nucleotides around the lesion
3. XPA and XPG bind, as well as RPA, which stabilizes the single-stranded DNA and prevents hairpins from forming
4. XPF loads; XPF and XPG nick the DNA on the appropriate strand
5. Damaged fragment is excised
6. DNA Pol δ/ϵ fills the gap, ligase I seals

Transcription coupled repair fixes only in the sense (transcribed) strand during transcription

1. XPC is not required; rather, lesions detected during transcription
2. CSA, CSB thought to assist the removal of RNA polymerase and placement of TFIIH
3. XPB, XPD, XPA, XPG load and proceeds as with GGR

4.2.4 Strand break repair

SSBs are easily corrected via XRCC1, LIG1. DSBs occur using three strategies

Homologous recombination (HDR) requires perfect copy of the region to serve as template, so more closely associated with G2 cells (sister chromatid cohesion keeps the homologs together)

1. exonuclease cleaves back a 3' overhang
2. Rad51 binds to the single stranded DNA and “coats” the DNA so it can invade into the sister chromatid; assisted by Rad52, BRCA2
3. DNA polymerase δ/ϵ copies the DNA
4. Holiday junctions formed, Blm/TOP3a migrates them to the appropriate place
5. Holiday junction resolvase (Rad51C involved?) to cleave the junctions
6. either a gene conversion or crossover event can result

Single-strand annealing (SSA) mutagenic pathway, since it causes a deletion, but on average it'll hit an intergenic region anyway

1. exonuclease cleaves back a 3' overhang until it reveals a homology region downstream
2. Rad52 allows the homology regions to basepair, causing massive 3' overhangs
3. Rad1 and Rad10 cause removal of the overhang, resulting in a deletion

Non-homologous end joining (NHEJ) mutagenic pathway, causes deletion of a few base pairs

1. Artemis, MRE complex, Wru may be required to polish the ends (removal of some bases)
2. Ku70/Ku80 binds to the blunt ends, allowing LIG4/XRCC4 to bind
3. DSB is ligated

| | |
|----------------------|----------------------------|
| Base hydrolysis | Base excision repair |
| Oxidative damage | Base excision repair |
| Mismatches | Mismatch repair |
| UV | Nucleotide excision repair |
| Single strand breaks | SSBR |
| Double strand breaks | Homologous recomb and NHEJ |

Table 1: Summary of repair mechanisms

4.3 Effects of cell cycle and checkpoints on genome maintenance

- DNA repair takes time
- Passing into next phase of cell cycle before completion of repair could be mutagenic
- Cell-cycle checkpoints triggered by DS break in G1, G2; pyrimidine dimer in S phase

5 Eukaryotic transcription (Kadesh)

5.1 Eukaryotic class II gene transcription

Basal apparatus consists of TFIID, RNA Pol II, and a host of other proteins. Pol II has a CTD with repeated amino acids that is a site for protein interaction. TFIID binds the TATA box via a TBP (TATA binding protein) in a rate-limiting step. Enhancer proteins (transcription factors), which bind to heterogeneous upstream sequences, help speed up this process (in the case of activators/coactivators) or repress it (repressors/corepressors). TFs generally have a way to bind DNA and a separate domain subject to regulation, or multimerize with other elements to put together such functionality.

5.2 Examples of transcription factors

- MyoD is a master regulatory protein at the top of a regulatory cascade that leads to cell identity specification. Binds DNA by dimerizing with E47
- Max forms heterodimers with each of Myc, Mad, Mga, and Mnt; the complex binds specific enhancer regions and recruits coactivators/corepressors
- VP16 is a strong viral transcriptional activator that can only bind DNA upon dimerization with host factors.
- CREB must be phosphorylated at S133 to be active (by CaM kinases, PKA)
- STAT contains SH2 domains, which interact with phosphotyrosines. It is only active as a dimer, and it must be phosphorylated by Jak to dimerize, in response to extracellular signals (IFN, IL-6)
- TGF- β is a ligand that usually inhibits cell cycle. Binding to receptor complex activates SMAD 1, 2, 3 (receptor SMADs) by phosphorylation, which are now competent to form heterodimers with SMAD-4 (Co-SMAD). SMAD-4 localizes to the nucleus and usually piggy-backs on existing DNA binding proteins such as AP1, TFE3, FAST-1
- NF- κ Bs are sequestered in the cytoplasm by I κ Bs. Proteolytic cleavage of I κ B causes release of NF- κ B, which can then translocate to the nucleus, form heterodimers, and activate transcription
- FOXO is an apoptosis activator and is regulated by phosphorylation through several pathways, including AKT (PKB). Phosphorylation leads to recognition by 14-3-3 and sequestering in the cytoplasm.
- SREBPs activate sterol metabolism pathways. SREBP is a long polypeptide that is ER-membrane bound. In high cholesterol conditions, SREBP is stabilized by SCAP. In low cholesterol conditions, S1P and S2P cleave the lumen portion of SREBP, releasing a cleaved and active transcriptional activator. SREBP activates lipid synthesis/uptake genes such as HMG CoA reductase.
- Notch is a membrane receptor that upon contact with ligand bound by a neighboring cell (Delta/Serrate, Jagged) is cleaved to release NICD, a TF that piggy backs on DNA binding protein CSL/Su(H) to control cell-fate genes.

5.3 Coactivators, corepressors

5.3.1 Mediator

Mediator is a coactivator protein complex (subunits include TRAP, DRIP/ARC, CRSP) that promotes assembly of Pol II pre-initiation complexes and recruitment of activators and chromatin-repressing factors to the site of transcription. Normally, TBP is not bound to TATA box until it binds an activator. Binding recruits mediator, which in turn stimulates assembly of the basal machinery. The component Sur2 in Mediator is required, for example, to bind to E1A and Elk-1 for them to activate transcription.

5.3.2 CBP

CBP is a coactivator for the TF CREB and is only bound when CREB is phosphorylated at S133. p300 is similar to CBP. Both are examples of histone-acetyl transferases, which have an effect on chromatin structure.

5.3.3 SMRT

SMRT, N-CoR interact with mSin3A, which is a corepressor for Mad-Max. Additionally, SMRT associates with a histone deacetylase

5.3.4 MEF2

MEF2s are involved in skeletal myogenesis and can function as either activators or repressors. MEF2 interacts with HDAC5 to inhibit transcription. Cam Kinases phosphorylate HDAC5, which leads to binding of 14-3-3 and the export of HDAC5 out of the nucleus. Unbound, MEF2 activates transcription.