

# NST PART 1B

## Cell and Developmental Biology

Michaelmas Term 2024

### LECTURES 1 - 6

#### Molecular Biology of the Cell Nucleus

**Dr Torsten Krude**  
Department of Zoology

**Lectures:** Biffen Lecture Theatre, Department of Genetics  
Tuesday, Thursday, Saturday at 10 am

#### LITERATURE REFERENCES

(i) General references. There are some excellent textbook introductions to the cell biological topics of this course. Relevant page numbers are given for each lecture in this handout.

1. '**Molecular Biology of the Cell**', 6<sup>th</sup> edition (MBC VI).  
B Alberts et al., 2015, Garland Science, New York & Abingdon.
2. '**Molecular Cell Biology**', 8<sup>th</sup> edition (MCB VIII).  
H Lodish et al., 2016, Freeman & Company, New York.
3. '**Lewin's Genes XI**'.  
JE Krebs et al., 2014. Jones & Bartlett, Sudbury MA.

For the first part of lecture 1, an additional good and relevant introduction into the structures of DNA and RNA can be found in chapter 4 of:

4. '**Biochemistry**', 6<sup>th</sup> or 7<sup>th</sup> edition.  
JM Berg, JL Tymoczko & L Stryer, 2007 or 2012.  
Freeman & Company, New York.

(ii) References to reviews and individual papers should be available online. If you have difficulties finding a reference, please tell the lecturer.

# LECTURE 1: ARCHITECTURE OF THE NUCLEUS I

## Introduction, DNA, Nucleosomes and Chromatin

**References:** Lodish et al., MCB VIII, pages 167-176

Alberts et al., MBC VI, pages 175-193

Krebs et al., Lewin's Genes XI, pages 223-242

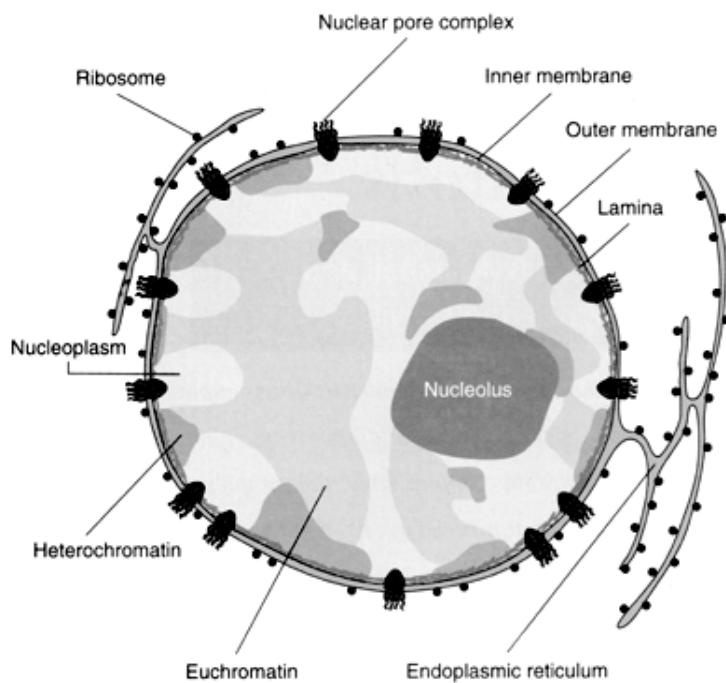
Further reading: Rhodes (1997) Nature 389, 231-233

### Introduction

At first approach the main function of the eukaryotic nucleus is the containment of the genomic DNA. In a human cell's nucleus, about two metres of DNA are contained within a sphere of approximately 10µm in diameter. This amazing degree of compaction is brought about by association of the genomic DNA with the structural components of chromatin, most prominently the histone proteins, to form ordered subunit structures called nucleosomes. Arrays of nucleosomes are further compacted into higher order structures to form chromosomes. Chromatin is surrounded by a double lipid bilayer membrane system forming the nuclear membrane, supported by a protein meshwork known as the nuclear envelope. Regulated transport of macromolecules into and out of the nucleus is achieved through nuclear pores.

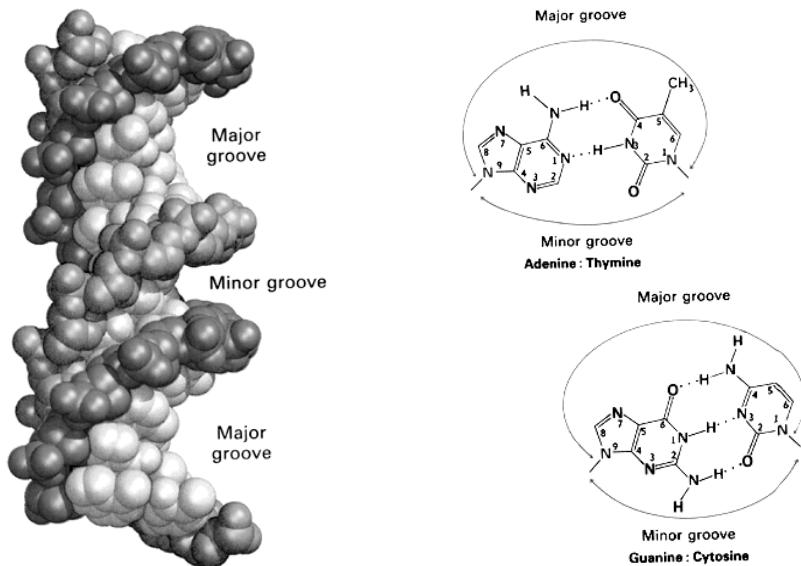
The nucleus is the information centre of the eukaryotic cell, synthesising the mRNA for thousands of proteins to be exported into the cytoplasm for directing protein synthesis. In addition, the nucleus must replicate its entire structure accurately during each cell division cycle.

In the following 6 lectures, we will first consider structural elements of the nucleus, then discuss their replication during the cell cycle and we will finish by considering transport across the nuclear membrane. The subsequent 3 lectures will address the regulation of gene expression (i.e. transcription).



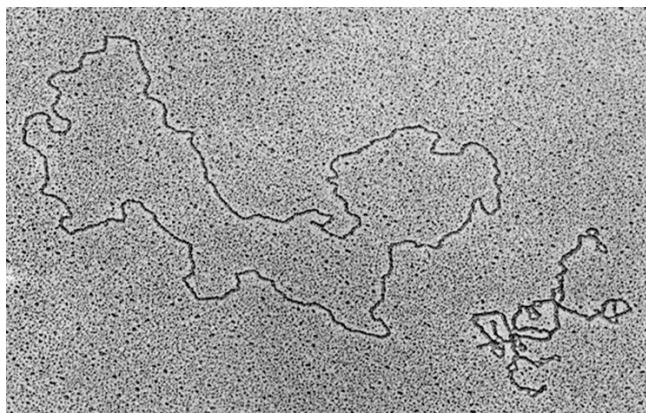
## DNA and its structure

Under physiological pH and salt concentrations as found in the nucleus DNA assumes the 'B-Form', a right-handed helix. The outer sugar-phosphate backbone chains are connected via hydrogen-bonded base-pairs forming a major and a minor groove. Access to genetic information (e.g. as required for sequence-specific binding proteins) is via the base-pair specific surface of the DNA helix in the grooves. Interaction of DNA with sequence-independent binding proteins (e.g. histones) can also be achieved through interaction with the backbone.



(Note: additional cartoons of DNA/RNA structures can be found in textbooks etc)

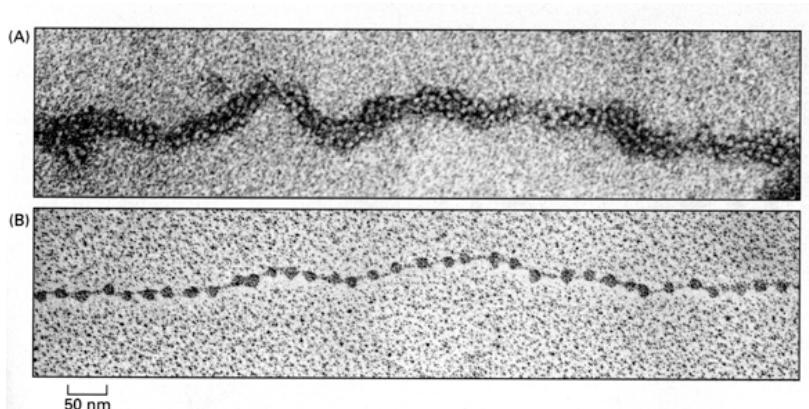
DNA can be visualised in the electron microscope. The small circular genomes of eukaryotic DNA viruses (such as SV40, polyoma etc.) can be seen in two configurations: relaxed and supercoiled.



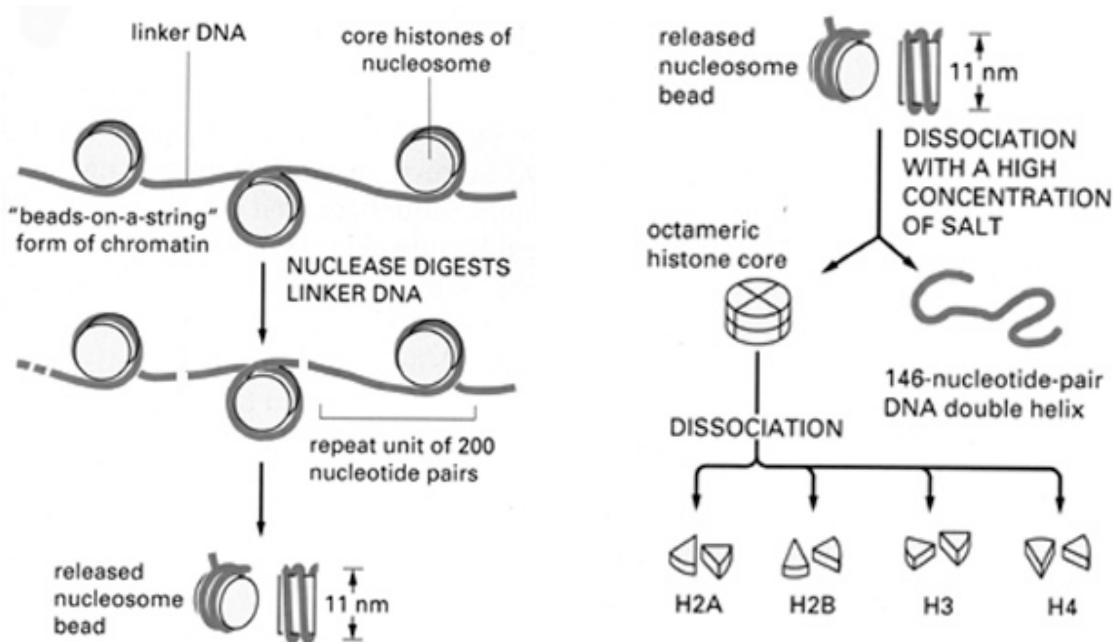
In a supercoiled circle, the DNA helix is again coiled around itself, forming a more compact structure when compared to the open relaxed circle. (Bacteria use this supercoiling to fit their DNA into the small bacterial cell, and in contrast, supercoiling as found in eukaryotic circular DNA is a consequence of its association with histone proteins to form nucleosomes.) DNA Topoisomerases are enzymes that can relax supercoiled DNA involving a transient breaking and resealing of the DNA backbone (as discussed in detail later in lecture 4).

## Nucleosomes and Chromatin

Electron microscopy has provided evidence for regular arrays of nucleosomes along the genomic DNA, constituting a chromatin fibre.

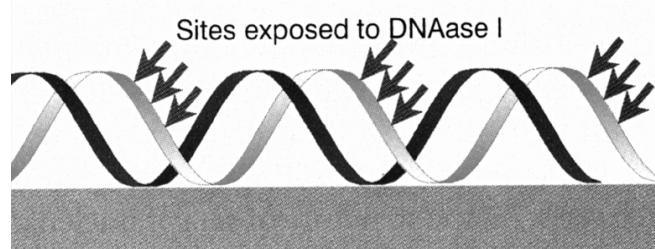


Treatment of cell nuclei with certain DNA endonucleases (e.g. Micrococcal Nuclease) releases nucleosomes as discrete substructures, containing a fragment of genomic DNA and structural histone proteins. The DNA and the histones can be separated from each other further by high salt treatment or denaturation.

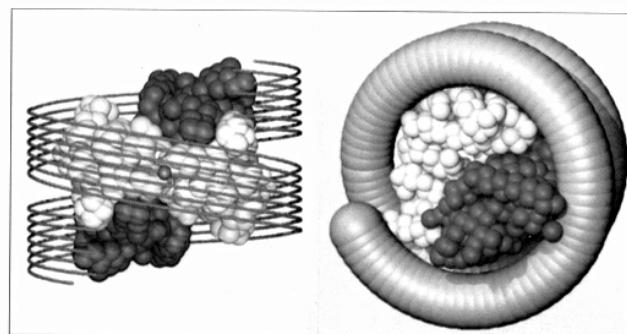


Nucleosomes core particles consist of 146bp of DNA and an octameric histone core, containing two copies of each core histone (H2A, H2B, H3, H4) in all eukaryotic organisms from yeast to humans. Histone proteins form stable substructures: a heterotetramer of two molecules each of H3 and H4 and heterodimers of one molecule each of H2A and H2B. One tetramer and two dimers form a histone octamer, which provides the structural core of the nucleosome.

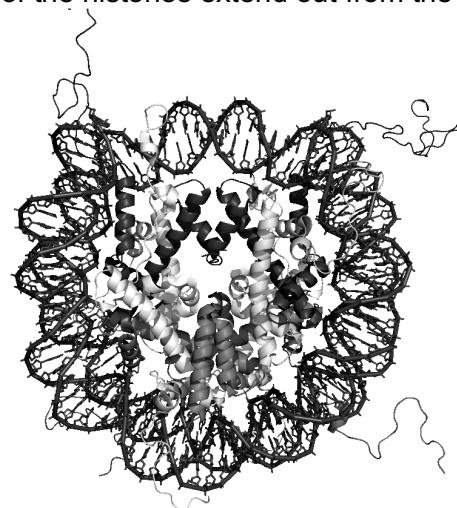
Further digestion of core particles with DNase I results in a regular periodic pattern of about 10-12bp fragments (i.e. the number of base pairs per full helical turn), indicating that the DNA is wrapped around the surface of the histone core octamer.



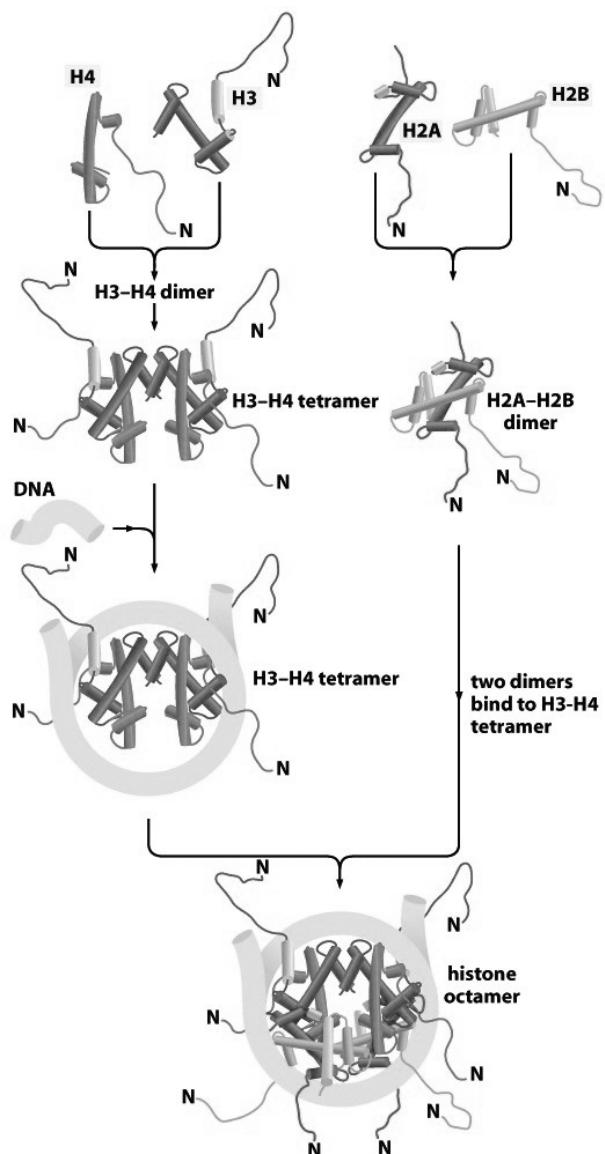
This low-resolution structure shown in a simplified diagram below. The nucleosomal wrapping of DNA around the histone core leads to the formation of one constrained supercoil per nucleosome (which is released when the histones are removed).



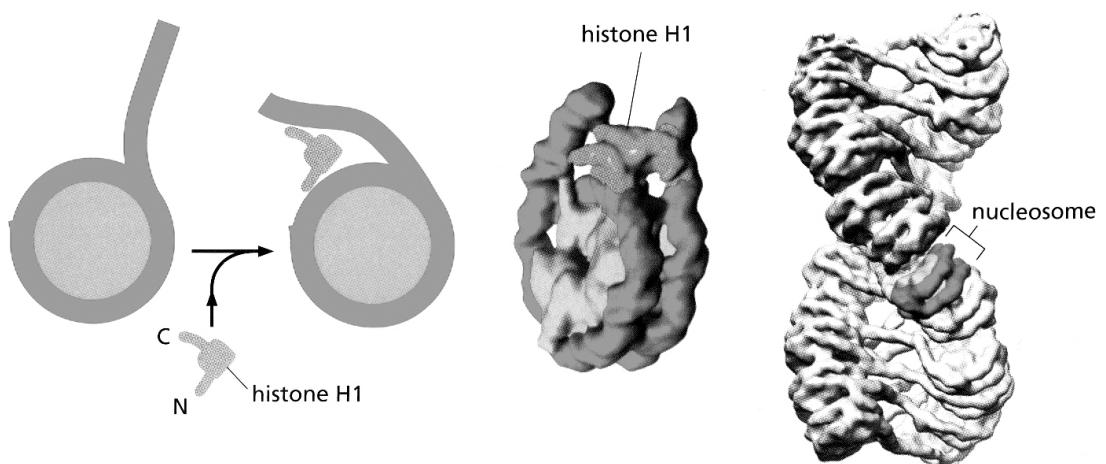
The structure of the nucleosome core has been resolved at atomic resolution by X-ray diffraction of pure nucleosome crystals. This high-resolution structure is shown below. The central globular domains of the histone proteins are located on the inside, whereas the N and C terminal tails of the histones extend out from the core structure.



Nucleosomes are assembled from histone subcomplexes. A tetramer of  $(H3-H4)_2$  contacts DNA, and this is subsequently joined by two H2A-H2B dimers. This reaction can occur with purified components *in vitro* if the salt conditions are carefully controlled. However, *in vivo* it is mediated by histone chaperones and chromatin assembly factors (as discussed in detail later in lecture 5).

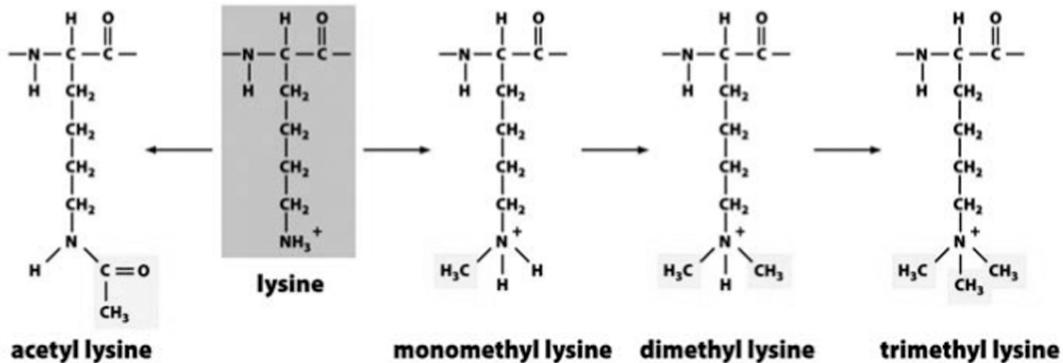


Adjacent nucleosomes are connected by linker DNA (its size varies between different cell types, tissues and species). Associations of linker histone H1 with the linker DNA and the nucleosome core are involved in forming higher order structures.

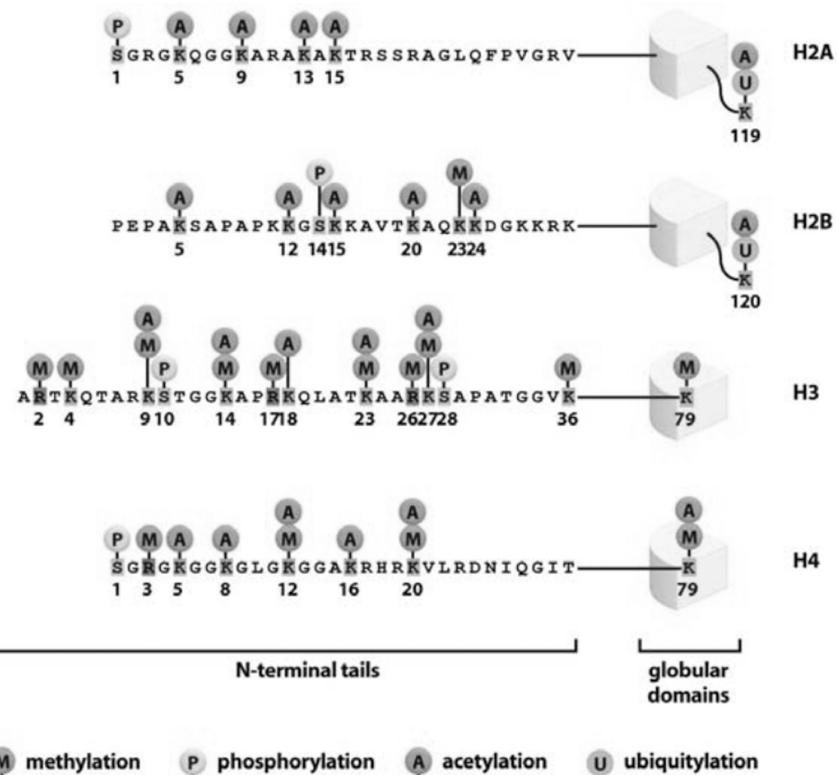


## Modifications of histones: the ‘histone code’

The tails of the four core histones extend out from the nucleosome core and are available to be acted upon by protein modifying enzymes. Specific amino acids in these tails are therefore often post-translationally modified by a wide variety of modifications, including acetylation, methylation, phosphorylation and ubiquitination.



These modifications comprise epigenetic information (heritable information that is not encoded in the DNA sequence) that can influence, for instance, whether a particular genetic locus is active or inactive under a given set of cellular circumstances. This concept is known as the histone code.



The histone code is “read” by proteins that can bind specifically to the modified histone tails. For example lysine 9 of histone H3 specifies an inactive state when trimethylated, but an active state when acetylated, by attracting different binding proteins.

These states are reversible as the modifications can be removed and new modifications can be added. Thus the cell’s chromatin is a dynamic entity, and regulated changes in its structure occur as cells respond to their environment, grow and differentiate.

## LECTURE 2: ARCHITECTURE OF THE NUCLEUS II

### Chromosomes and organisation of the nucleus

**References:** Alberts et al., MBC VI, pages 194-216

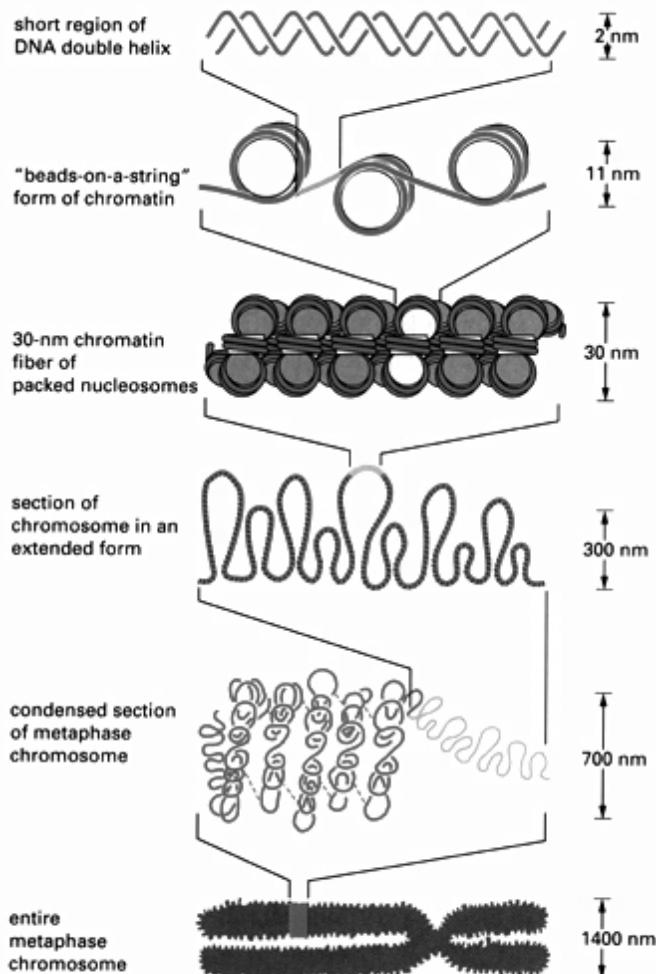
Lodish et al., MCB VIII, pages 327-341

Krebs et al., Lewin's Genes XI, pages 200-210

#### High-order chromatin structure

The nucleosome subunits of chromatin allow for hierarchies of folding chromatin fibres. Dynamic accessibility of certain regions of DNA in the chromatin fibres has to be provided for gene expression or replication.

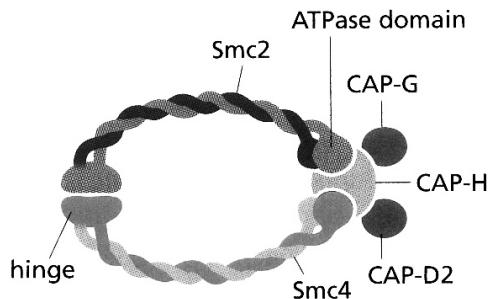
The most condensed state of chromatin is found during metaphase of mitosis in the form of chromosomes (mitotic events involving reversible chromosome condensation will be discussed later in lecture 5). Key principles of higher-order packing include loop formation and attachment of chromatin fibres to an underlying 'scaffold' or 'matrix' structure.



Packing DNA into mitotic chromosomes

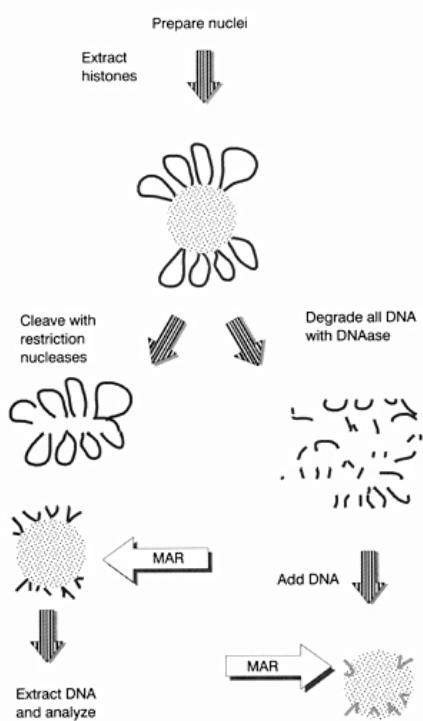
### Condensation of the chromatin fibre

Chromatin condensation is mediated by condensins, large proteins belonging to the class of 'structural maintenance of chromosomes' proteins (SMCs). Complexes of the condensins SMC2 and SMC4 with kleisin proteins (CAPs) can clamp chromatin fibres, thus mediating their condensation.



### Nuclear Matrix or Nuclear Scaffold

Depleting entire interphase nuclei or mitotic chromosomes of histones (e.g. by detergents and high salt) liberates the chromosomal DNA from an underlying residual protein structure. This structure is referred to as nuclear matrix, or chromosome scaffold. The DNA is attached in loops of ~60kbp to the residual matrix or scaffold. The DNA that is attached to the matrix or scaffold is referred to as either MAR or SAR (matrix- or scaffold-associated regions, respectively) and can be isolated after digesting the DNA loops with nuclease and sequenced.

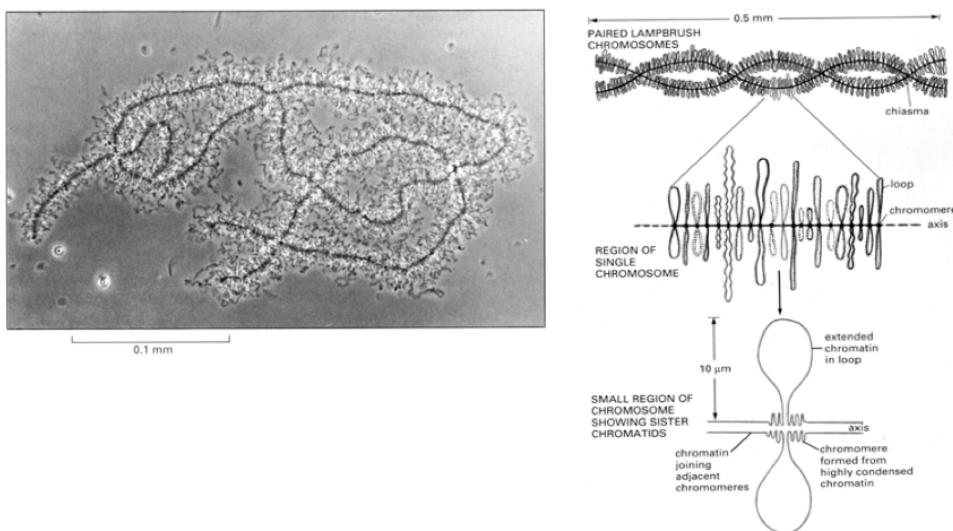


MARs are very A/T rich and contain weak 'consensus' sites for DNA topoisomerase II, suggesting that DNA topoisomerase II might be involved in controlling the coiling of (topologically closed) DNA loops on the matrix. Generally speaking, the constituents of the nuclear matrix or scaffold are ill-defined. However, a few specific factors are being characterised, such as scaffold-attachment factor A (SAF-A), which can directly bind to SAR elements.

### Principles of chromosome architecture during interphase

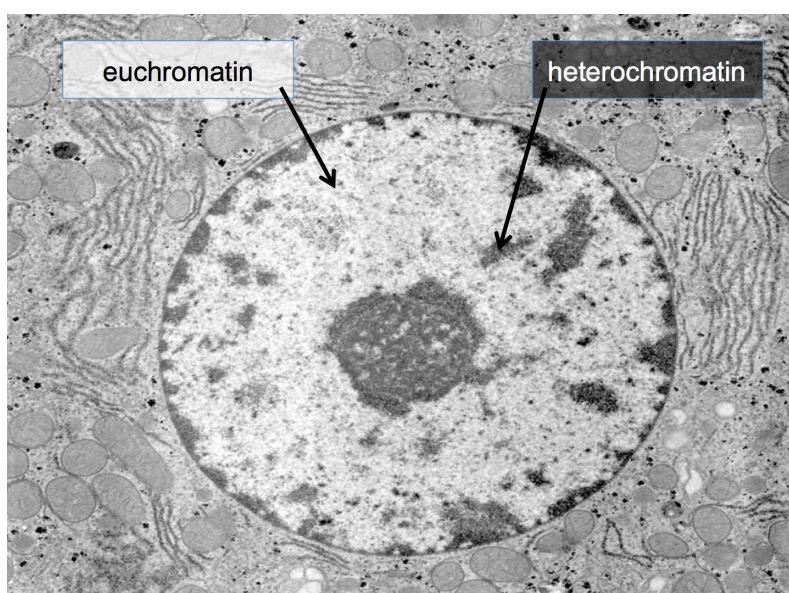
In special cases, individual chromosomes can be visualised during interphase and be analysed by light microscopy. Lampbrush chromosomes of newt oocytes can highlight essential features of interphase chromatin organisation.

These chromosomes exist in some amphibian oocytes during early meiotic prophase in a partially condensed state (which can persist for several months!). They are exceptionally active in transcription so that densely packed nascent RNP particles coat the loops allowing them to be seen in the light microscope. Hybridisation of DNA probes to chromosome preparations confirms that the organisation is strictly sequence specific and individual decondensed loops can correspond to particular active genes.



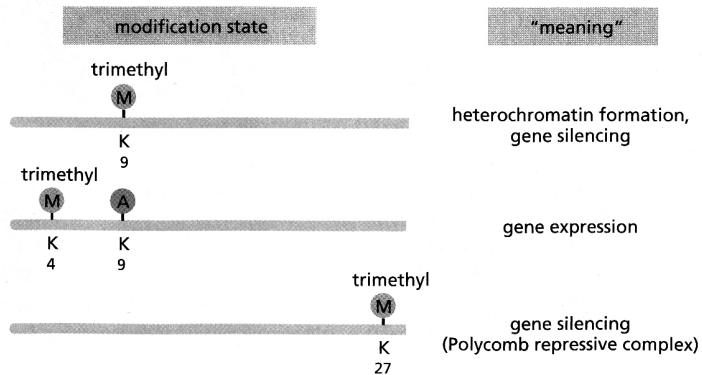
### The interphase nucleus

Transmission electron micrographs of cell nuclei in interphase cannot provide a differentiated image of individual chromosomes. However, two distinct morphological states of chromatin are seen: densely packed heterochromatin and relatively decondensed euchromatin. Active genes are present in euchromatin, but not in heterochromatin.



Formation of euchromatin and heterochromatin involves the post-translational modifications of histone tails. These modifications produce altered binding surfaces for

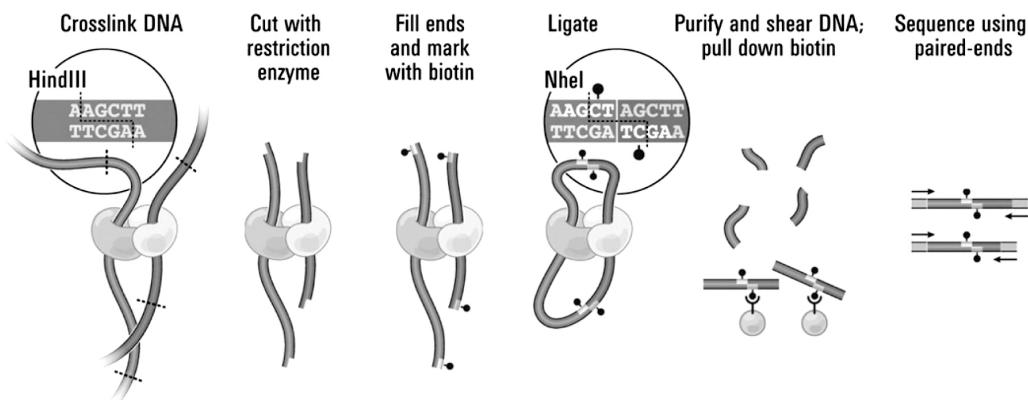
effector proteins that influence chromatin structure. For instance, histone H3 that is methylated at lysine residue 9 (H3K9me) is bound by the HP1 heterochromatin protein, which brings with it other proteins that act to compact and silence the DNA. On the other hand, histone H3K9 acetylation is an activating modification, which brings in chromatin remodelling enzymes that open the chromatin structure and enable access of the transcription machinery to the DNA.



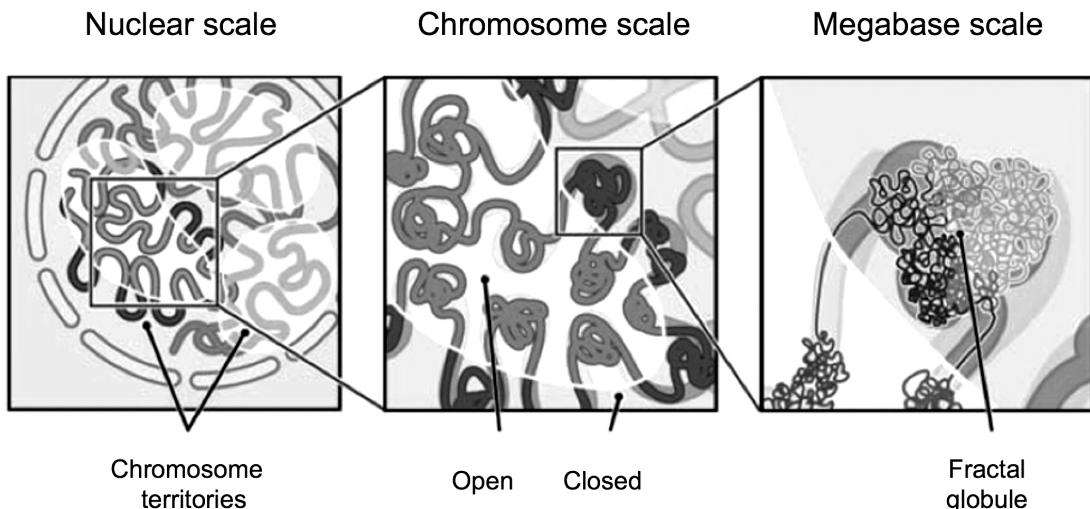
### Functional compartmentalisation of the interphase cell nucleus

Hybridisation experiments of whole nucleus preparations with probes specific for individual chromosomes ('chromosome painting') have provided evidence for the concept that individual chromosomes occupy discrete 'territories' within the interphase nucleus. Locations of individual chromosomes and individual genetic loci within the nucleus, however, are dynamic.

The recent developments of high throughput DNA sequencing techniques have enabled the mapping of three-dimensional DNA-DNA interactions within the nucleus. This technique is known as 'Hi-C' (High-resolution Chromosome conformation capture). It is based on the ligation of proximal DNA segments and the sequencing of these junctions.



The evidence obtained by Hi-C has been used for modelling interactions within chromosome territories at chromosomal resolution, and for a dynamic spatial segregation of chromatin fibres into open and closed domains at megabase resolution.



### Structural patterns of nuclear activity

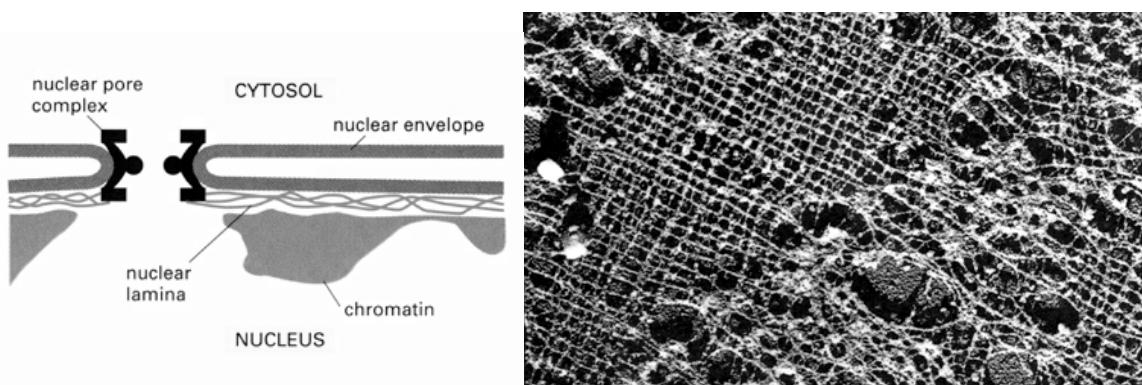
Evidence is growing that many activities of the interphase nucleus occur at discrete subnuclear sites, or foci, rather than dispersed or in solution.

Examples of such nuclear compartmentalization are clusters of DNA replication forks and clusters of the RNA synthesis and processing machinery. (Mechanistic details of DNA replication and transcription will be discussed in subsequent lectures.) The intracellular locations and patterns of these 'replication foci' and 'transcription foci' are highly dynamic. These foci seem to be associated with the nuclear matrix.

The nucleolus is a large subnuclear compartment that is separated from the rest of the nucleus by a region of condensed heterochromatin. Within the nucleolus the rDNA is transcribed by RNA polymerase I to generate pre rRNA. These RNAs are then processed to generate mature rRNAs and assembled with imported ribosomal proteins to generate ribosome subunits. These are then exported out of the nucleus and complete ribosomes are assembled in the cytoplasm.

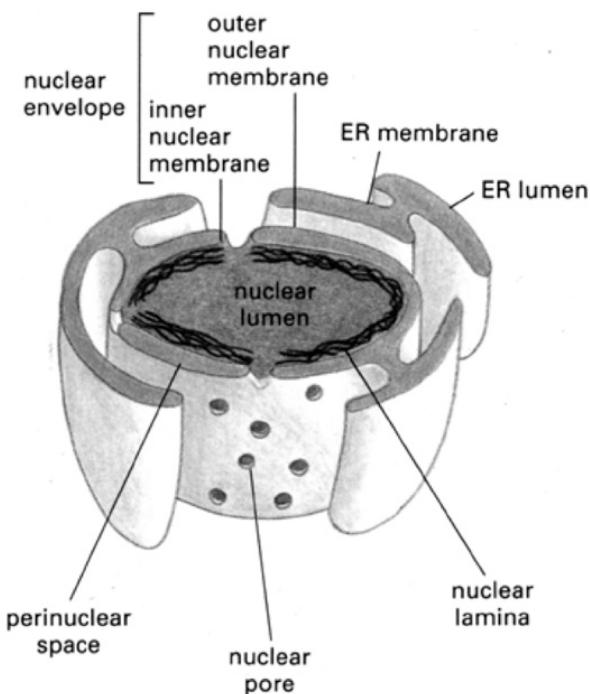
### The nuclear frontier: nuclear membrane and lamina

In the interphase nucleus, chromatin is surrounded by the nuclear envelope. It consists of two membranes and the underlying lamina, a two-dimensional meshwork made of a special kind of intermediate filament proteins: lamins A, B and C. It makes direct contact to chromatin as well as to the inner nuclear membrane. Some evidence suggests that lamin fibres extend into the lumen of the nuclei, possibly contributing to the architecture of the nuclear matrix.



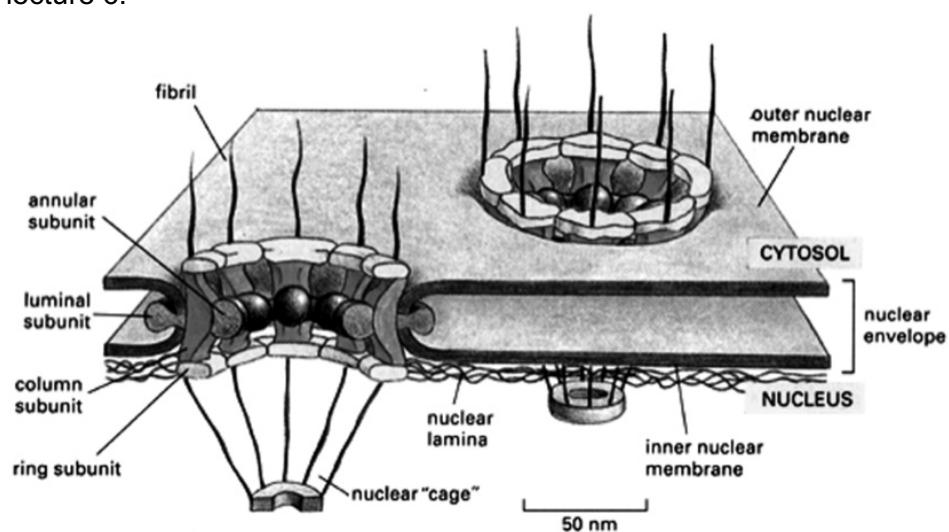
The nuclear envelope: cartoon and electron micrograph of the lamina

The inner nuclear membrane contains receptors that bind the nuclear lamina. The outer membrane is continuous with the rough endoplasmatic reticulum and contains ribosomes. Both membranes encapsulate the perinuclear space, which is continuous with the lumen of the ER.



The membrane system of the nuclear envelope

Both nuclear membranes are perforated by nuclear pore complexes. These are huge macromolecular complexes of a molecular weight of ~150MDa and are made of 50-100 different proteins, the nucleoporins. Nuclear pore complexes form aqueous channels across the nuclear envelope allowing diffusion of small molecules into and out of the nucleus. Their main function, however, is the regulated transport of macromolecules into and out of the nucleus, which we will discuss in detail later in lecture 6.



## LECTURE 3: EUKARYOTIC CHROMOSOME REPLICATION I

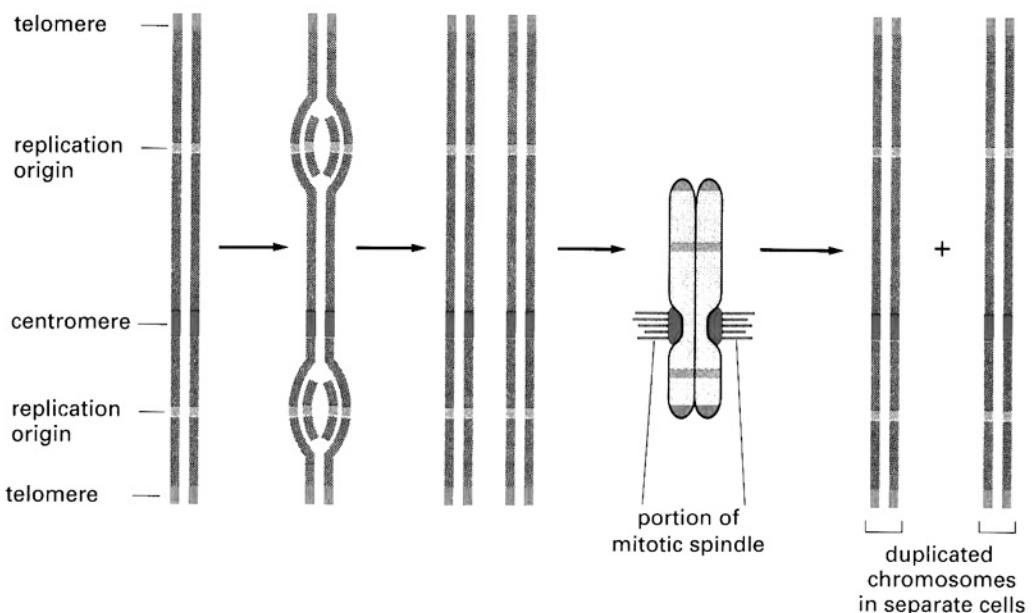
### Replication Origins, Centromeres, Telomeres

**References:** Alberts et al., MBC VI, pages 254-266  
Lodish et al. MCB VIII, pages 345-349  
Krebs et al., Lewin's Genes XI, pages 295-302, 210-220

*Further reading:* Hu & Stillman (2023) Molecular Cell 83, 352-372.

Chromosome replication in eukaryotes differs from prokaryotes in several ways, e.g. 1) replication initiates at many sites per chromosome; 2) one round of replication and cell division is completed before the next starts; 3) parental & progeny DNA is assembled into nucleosomes and higher order chromatin structures.

We will discuss first the elements required for replication of chromosomes during S phase of the cell division cycle, and then those required for segregation of two replicated daughter chromosomes at mitosis.

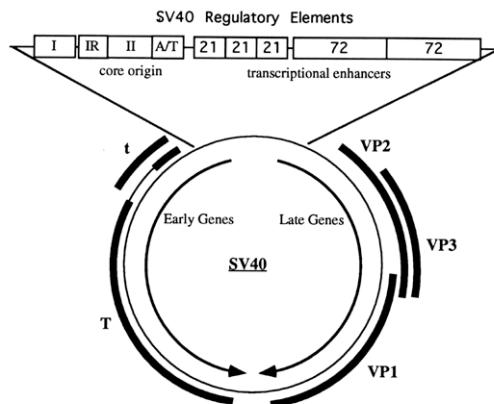


The chromosomal sites where DNA replication initiates are operationally defined as replication origins. What specifies these origins?

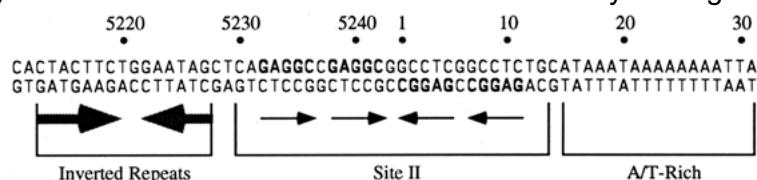
Origins of replication are better understood in prokaryotes, animal viruses and lower eukaryotes (e.g. yeast) than in higher eukaryotes. Knowledge of higher eukaryotic origins comes from biochemical and genetic analyses of a few example cases, and from high throughput DNA sequencing analyses at genome-wide resolution.

#### Replication origins of animal viruses

SV40 and polyoma virus are organised as minichromosomes and only provide the control elements for their own replication: origin sequences and initiator protein. The other factors required for their replication are recruited from the infected host cell (monkey or human), hence they were the first simple and useful model systems for eukaryotic chromosome replication.



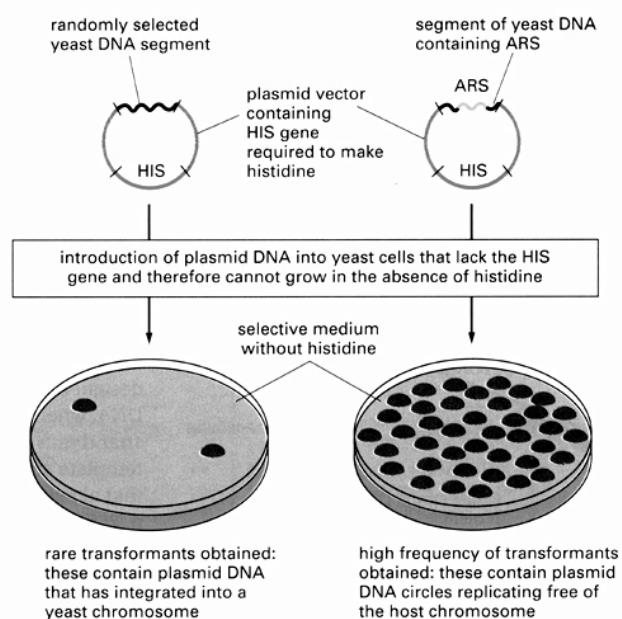
The initiator protein 'large T antigen' must bind to a unique site on the viral genome to initiate replication; a 65bp control region containing a 27bp inverted repeat (potential hairpin) and a conserved A/T-rich element is necessary for origin function.



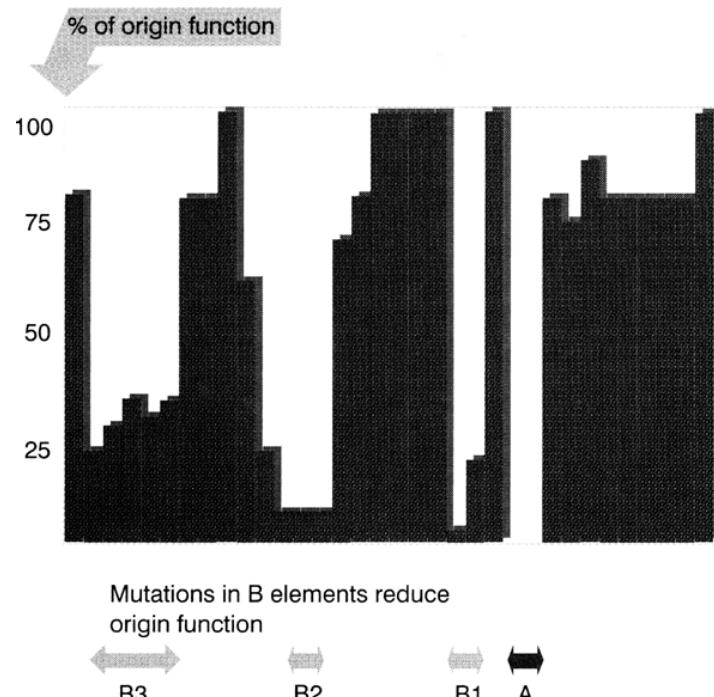
These viral origins provide no valid models for the cell's chromosomal origins because the virus must replicate many times in one cell cycle for propagation, thereby defying cellular controls (see Lecture 5).

### Autonomously replicating sequences in yeast: origins of replication

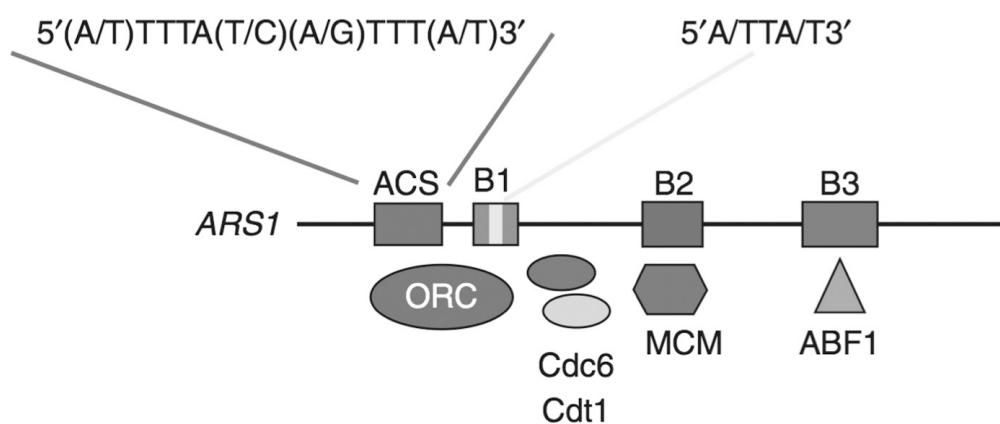
Yeast plasmids (e.g. 2 $\mu$  plasmid) are organised as episomal minichromosomes. Their replication requires a specific DNA sequence element called an autonomously replicating sequence, or ARS element. These act as replication origins allowing initiation. ARS elements were isolated from random genomic DNA fragments by their ability to confer the ability to replicate as an autonomous plasmid in yeast cells after ligation to an origin-less circular plasmid.



Deletion and point mutants in a typical ARS element define an essential consensus (A box) for ARS activity. Flanking sequences (B elements) also affect the efficiency of ARS function. More than 200 ARS elements are found in the chromosomal DNA of budding yeast and they also function as origins in this context. However, replication does not always initiate at all possible yeast origins.

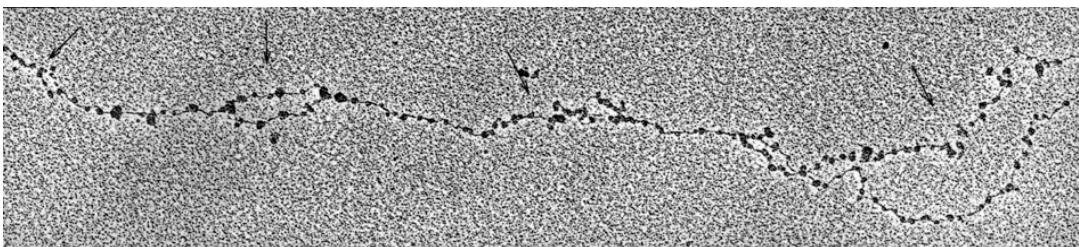


Nuclease digestion experiments *in vivo* have shown that the DNA sequences of the A and B elements are protected from nuclease degradation by associated proteins. Yeast origins are recognised in a sequence-dependent manner by an initiator protein complex called the origin recognition complex (ORC), which binds the A box in the presence of ATP. Additional factors such as Cdc6, Cdt1 and the MCM complex associate with ORC in a stepwise manner and occupy additional positions such as the B elements (more about the function of this so-called pre-replicative complex later in lectures 4 and 5).



## Replication Origins in Higher Eukaryotes

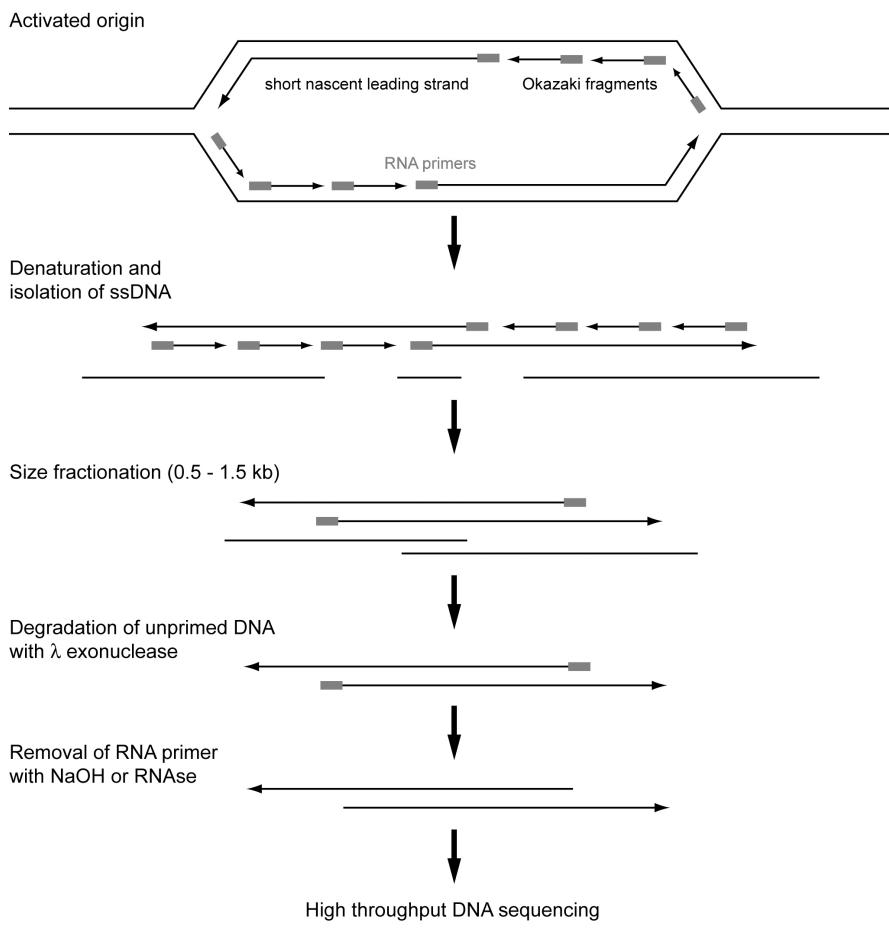
An average human chromosome would take about a month to replicate from a single initiation site with two divergent forks moving at 3kb/min. However, S phase usually lasts for less than 10 hours. Therefore, several thousand replication origins must be present per chromosome, and tens or even hundreds of thousand origins per genome. Indeed, EM analysis of replicating chromosomes provided evidence for replication bubbles emanating from activated origins at intervals of about 30-300kb.



A site-specific replication origin can be defined as, and mapped at high precision by determining the transition point between leading and lagging strand synthesis (more about DNA replication forks later in lecture 4).

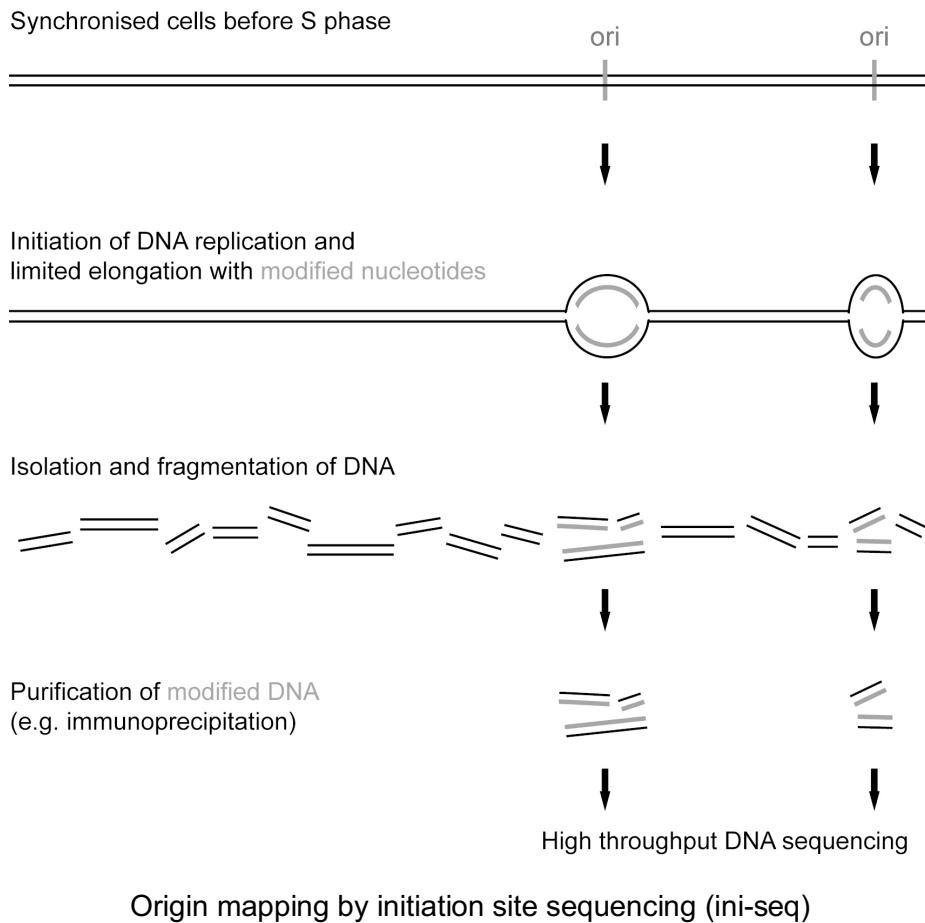
The task of mapping DNA replication origins in vertebrate cells has been revolutionised recently by technological developments in high-throughput DNA sequencing and computational analyses. These techniques (described below) have identified some tens of thousand potential origin sites in the human genome.

In one experimental technique, small nascent DNA leading strands are isolated from replicating cells and then sequenced to localise replication origins (SNS-seq).



Origin mapping by short nascent strand sequencing (SNS-seq)

In a second technique, DNA replication is initiated in nuclei of cells synchronised in late G1 phase of the cell cycle (just before DNA replication begins). DNA is allowed to replicate only for a short time following initiation and labelled by a modified nucleotide. This initiation-site associated replicated DNA is isolated and sequenced (ini-seq).



Both of the above techniques, SNS-seq and ini-seq, have yielded consistent results, detailing tens of thousands of defined replication origins in the human genome. They often correlate with genomic sites of high GC content, including CpG islands. These sites may form G quadruplex structures, and they often contain modified bases, such as methylated and hydroxymethylated cytosines. Additional experimental approaches such as the genome-wide profiling of Okazaki fragments, or the isolation of small replication bubbles have provided complementary datasets, that suggest that several individual origins tend to aggregate into larger initiation zones.

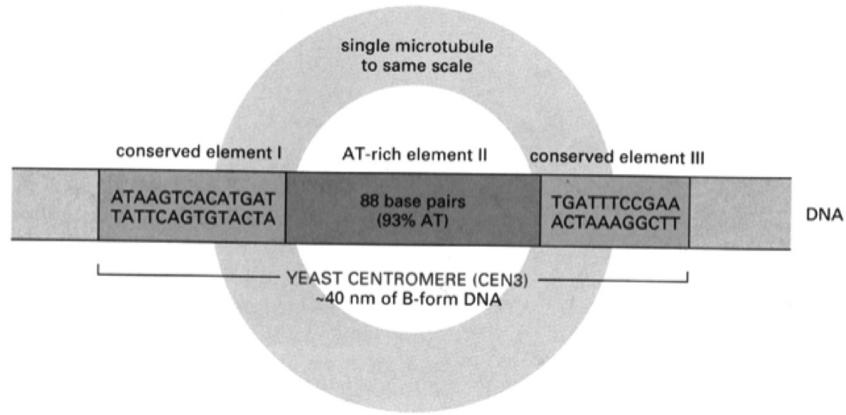
Prominent origin sites (and initiation zones) are often present in actively transcribed euchromatic regions of the genome. They are frequently located at or upstream of transcription start sites. Therefore, when DNA replication initiates at or near active gene promoters, replication and transcription elongation complexes move in the same direction and any detrimental head-on collisions between are avoided over that transcribed gene body.

The genetic and epigenetic factors that determine, or influence, the specification of an origin site on a higher eukaryotic genome are still unclear and currently under investigation. No strict DNA consensus motifs like the yeast ars elements are known in higher eukaryotes to date. However, likely candidates of origin specification

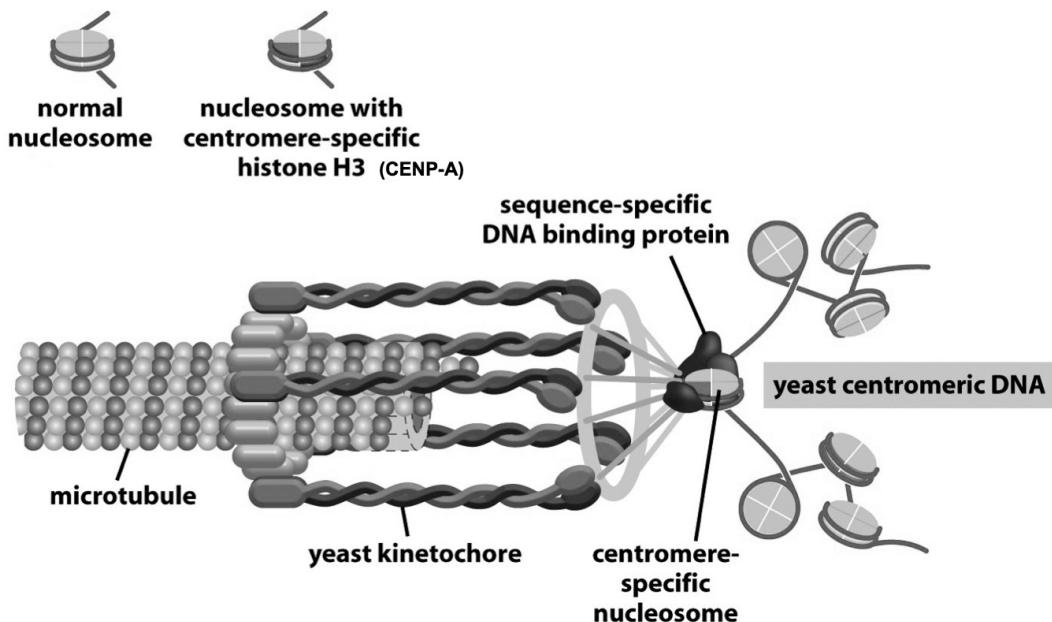
elements include short GC-rich DNA sequence motifs that may lead to unusual DNA structures (i.e. G quadruplexes), and include specific DNA and histone modifications.

### Centromeres attach chromosomes to mitotic or meiotic spindles

Back to yeast: Plasmids containing ARS elements replicate, but they are gradually lost without strong selective pressure. They can be stabilised by a centromere sequence (CEN). A CEN consist of three conserved DNA sequence elements I, II and III.

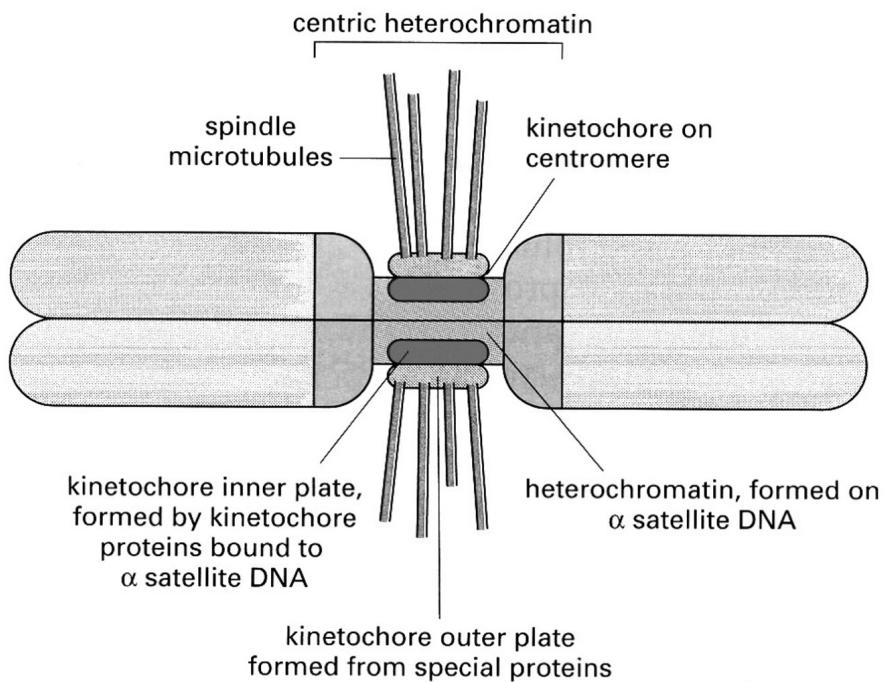


Centromeres serve as the attachment sites for centromeric proteins and spindle microtubules to form part of the kinetochore complex, where in mitosis the chromatids of condensed chromosomes are attached to the mitotic spindle.



Centromeres in higher eukaryotes are much larger and more complex. Centromeric DNA contains alpha-satellite DNA elements and is furthermore assembled into specific centromeric heterochromatin that spreads even further into the chromosomes.

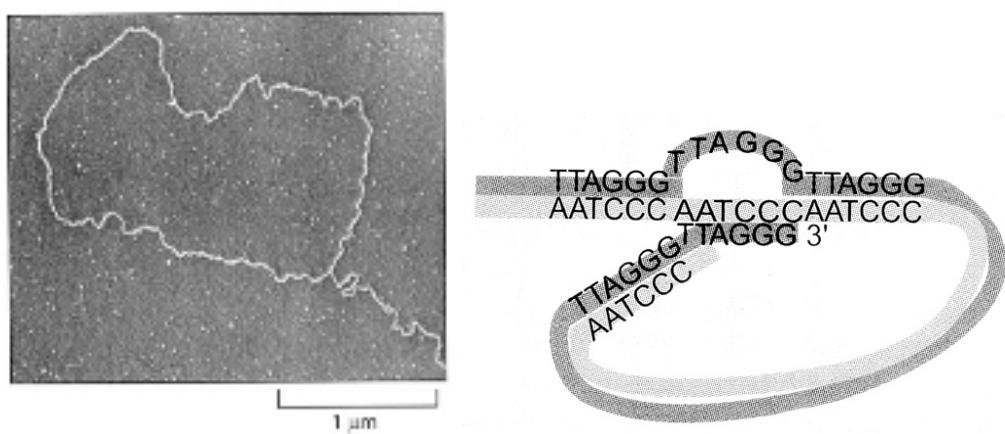
Centric heterochromatin contains sections with centromere-specific histone H3 variant, and sections with normal histone H3 that is di-methylated at lysine 4.



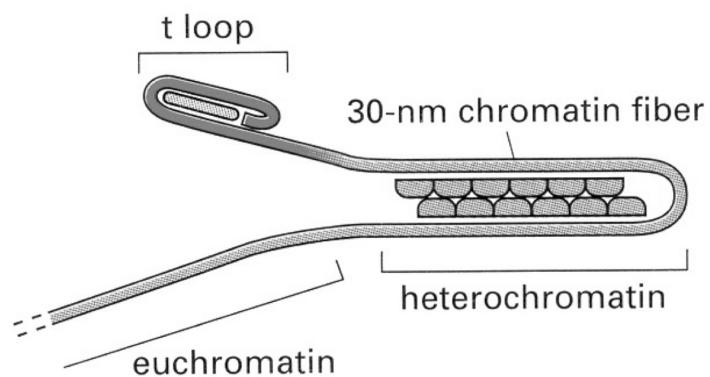
### Telomeres provide stable chromosome ends

Linear plasmids in yeast need CEN, ARS and a functional chromosome end, a telomere (TEL), to successfully replicate and segregate to daughter cells. Telomeres contain simple repeating sequences that form 3'-single stranded DNA overhang ends: yeast repeat TG(1-3): human repeat TTAGGG: Arabidopsis repeat TTTAGGG. (The complex replication of telomeres will be discussed in lecture 4).

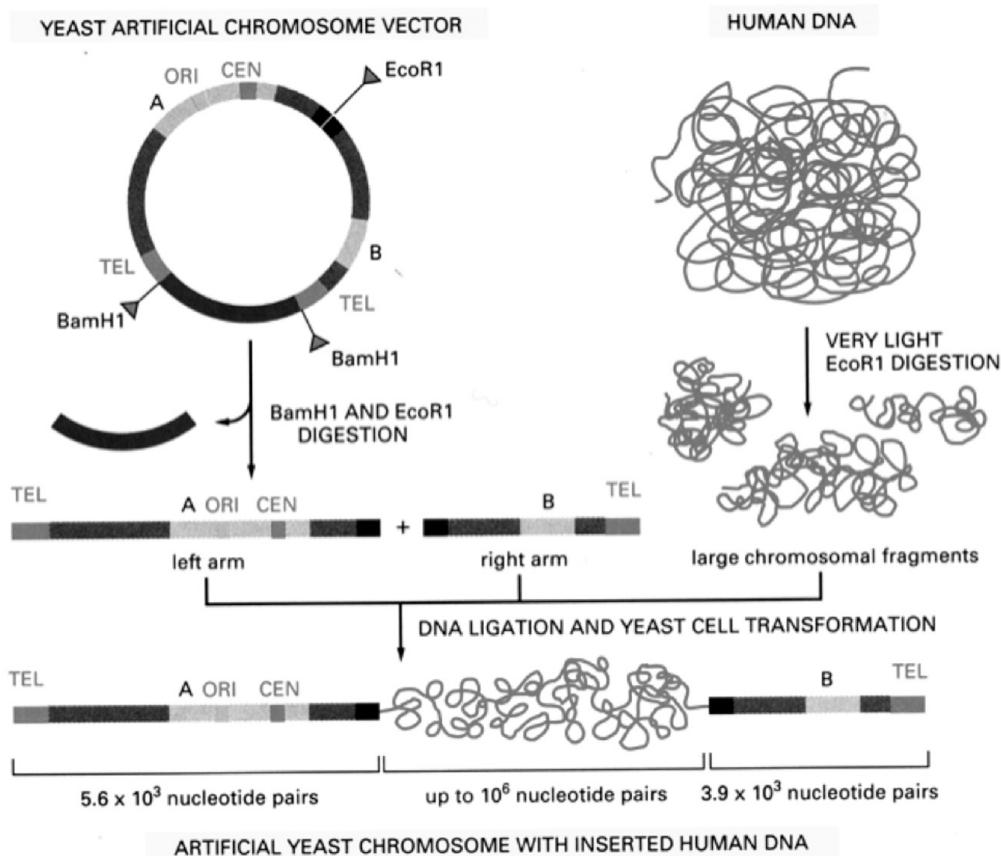
Hairpin and loop models have been put forward to explain the unusual structure of the DNA ends, as no free ends are detectable. At human telomeres a terminal loop, the 'T-loop', has been observed in the EM.



Telomeric DNA is also assembled into a specialised telomeric heterochromatin and attracts a protective protein complex called 'shelterin'. Telomeric heterochromatin spreads from the telomeric DNA further into the chromosomal DNA.



It is thus possible to generate in yeast artificial chromosomes (YAC) using all of these three elements. These YACs have been used successfully in cloning huge DNA fragments, for instance during the original human genome sequencing project.



## LECTURE 4:

## EUKARYOTIC CHROMOSOME REPLICATION II

### Replication forks and telomere maintenance

#### References:

Alberts et al., MBC VI, pages 237-254

Lodish et al., MCB VIII, pages 197-203

#### For enthusiasts:

Bell & Dutta (2002) Annu. Rev. Biochem. 71, 333-374

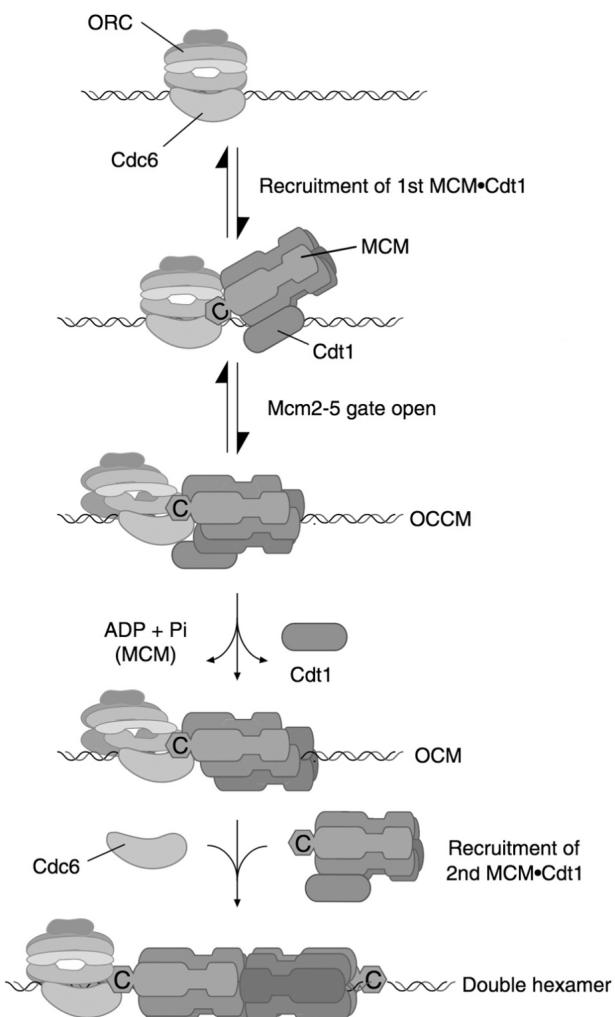
MacNeill (2001) Curr. Biol. 11, R842-R844

Deegan & Diffley (2016) Curr Opinion Struc Biol. 37, 145-151

The first enzymatic step of DNA replication is a localised separation, or unwinding of the two DNA strands at the replication origin. This event is catalysed by DNA helicases. Unwound DNA is stabilised by single-strand binding proteins. From here, DNA polymerases and additional proteins are recruited that build up active DNA replication forks.

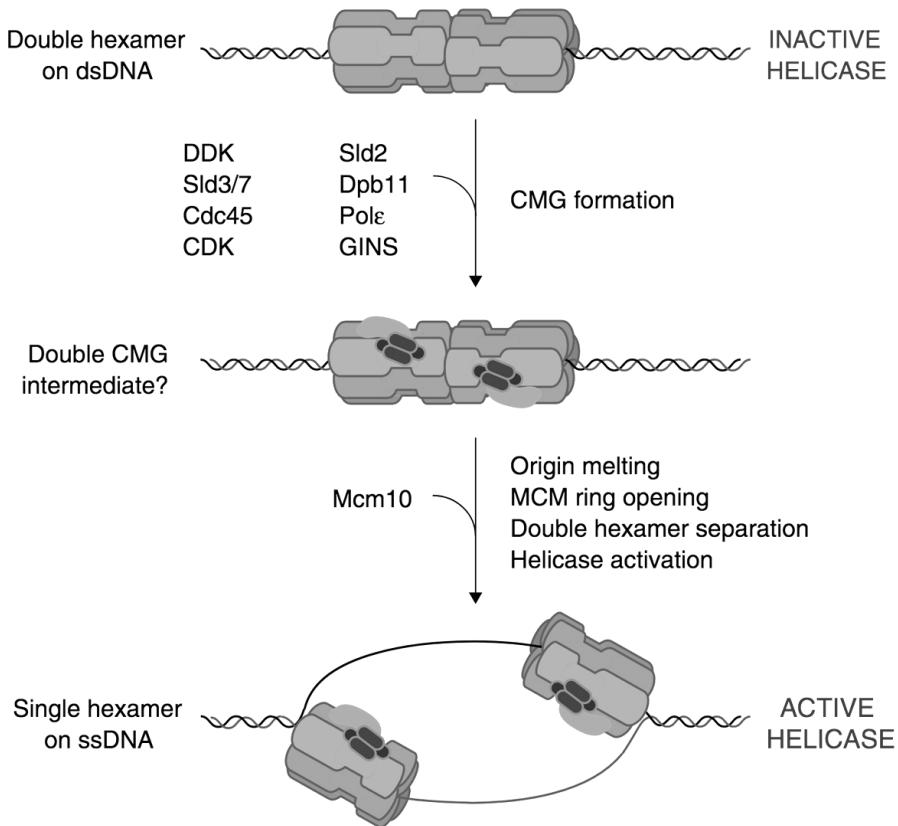
#### The replicative DNA helicase

The eukaryotic helicase is a core complex of six MCM proteins with many other associated proteins, including Cdc45 and the GINS complex. A large MCM double hexamer complex is loaded at replication origins in an ATP-dependent manner by the origin recognition complex ORC, involving Cdc6 and Cdt1 proteins.



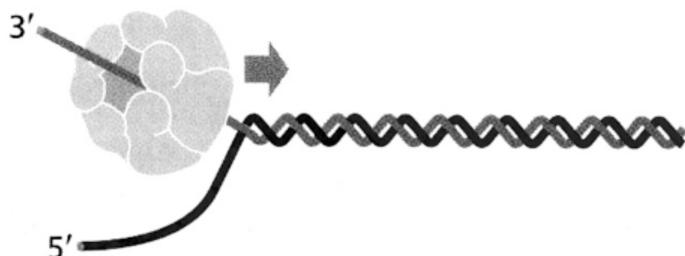
### Activation of the MCM helicase in S phase

The loaded, but enzymatically inactive MCM double hexamer complex is converted into an active helicase involving the activities of protein kinases CDK and DDK, and the association with several additional proteins, including Cdc45 and GINS. This complex is known as the CMG complex (from Cdc45, MCM, GINS). DNA helicase activation leads to local DNA strand unwinding and the separation of the double hexamer complexes. Subsequently each one of the two CMG helicase complexes travels with one of the two emerging replication forks away from the initiation site.



### Model of DNA helicase activity

After DNA unwinding, the active CMG DNA helicase translocates on the DNA leading strand in 3' to 5' direction dependent on ATP hydrolysis, thereby displacing the complementary DNA strand.



Around activated helicases, functional DNA replication fork complexes are assembled, involving the recruitment of DNA polymerases and replication factors.

## DNA polymerases

There are six major DNA polymerases (pol) in eukaryotes:

pol $\alpha$	contains primase; priming and primer extension of 'leading' and 'lagging' strand
pol $\gamma$	mitochondrial DNA synthesis
*pol $\delta$	replication of Okazaki fragments on lagging strand
*pol $\epsilon$	replication of continuous leading strand
	*DNA polymerases $\delta$ and $\epsilon$ contain proof-reading exonucleases
pol $\beta$	DNA base excision repair, gap-filling synthesis
pol $\eta$	DNA translesion synthesis, error-prone

## Hi-Fi Replication

DNA polymerases make mistakes. They are corrected by proofreading 3'-5' exonuclease activity and by mismatch repair systems:

Error rates:	pol $\alpha$	1 in $10^4$
	pol $\delta$ or $\epsilon$ (proofreading)	1 in $10^7$
	pol $\delta$ or $\epsilon$ + mismatch repair	1 in $10^9$

## Replication factors

Other crucial proteins of the eukaryotic DNA replication fork include:

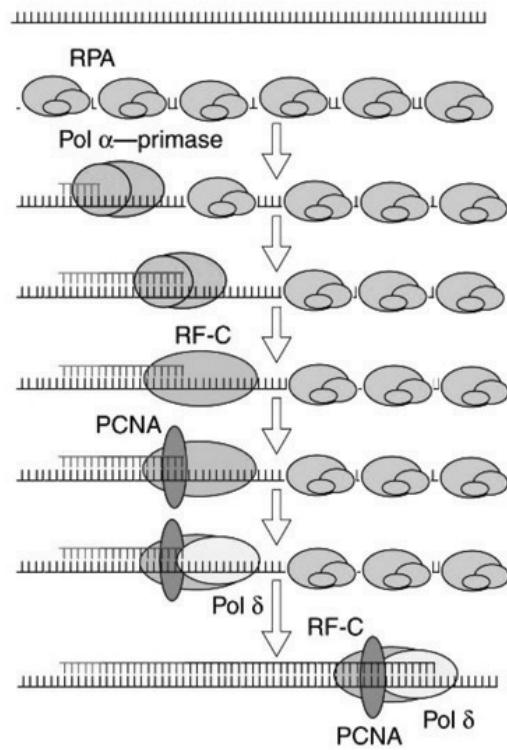
DNA helicases	unwind the two DNA strands, generate single-stranded DNA templates
RPA	single strand binding protein, stabilises the unwound strands, recruits pol $\alpha$ /primase
PCNA	sliding clamp, binds to pol $\delta$ and $\epsilon$ , Fen-1 and others
RFC	loads and unloads PCNA
Fen-1	flap endonuclease, removes short primers
Dna2	endonuclease, removes long primer flaps
DNA ligase I	joins Okazaki fragments
DNA topoisomerases	release superhelical stress

Recent proteomic analysis of isolated DNA replication fork complexes have identified large amounts of additional proteins that play roles in maintaining replication fork stability, facilitating replication of damaged DNA and replication of chromatin templates (more about the replication of chromatin templates in lecture 5).

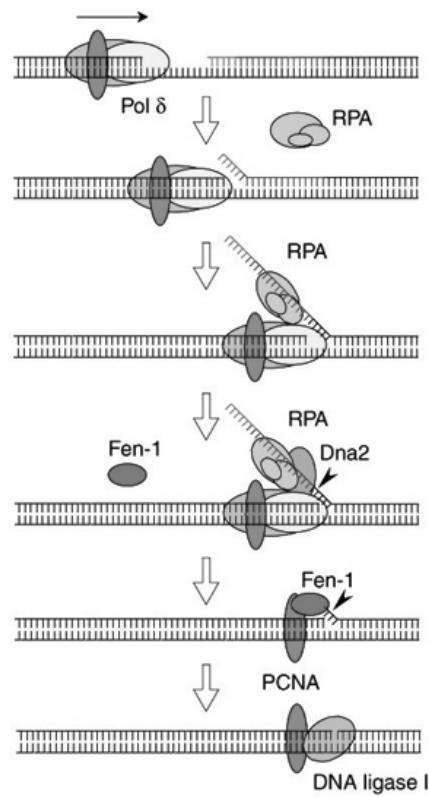
## DNA synthesis at DNA replication forks

The following model shows the concerted action of these core replication proteins during DNA strand synthesis in eukaryotes. Please note the mechanistic similarities between the equivalent steps in prokaryotes (which you have discussed last year in NST1A Biology of Cells) and eukaryotes (below):

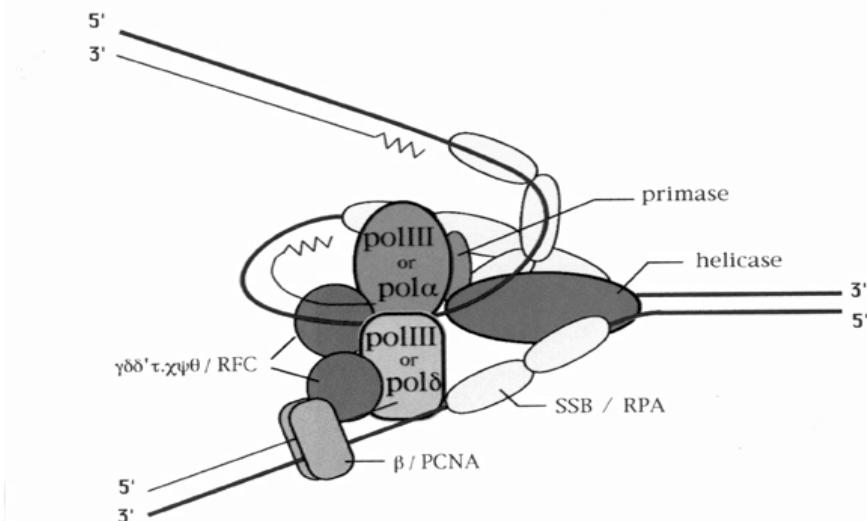
1) Initiation and elongation of DNA strand synthesis (applies to both leading and lagging strand)



2) Maturation of Okazaki fragments (lagging strand)



To establish a DNA replication fork, both leading and lagging strand synthesis are coupled and the lagging strand is looped back to obtain co-linearity ('trombone model'). The proteins of the replication fork thus form a complex 'molecular machine'.

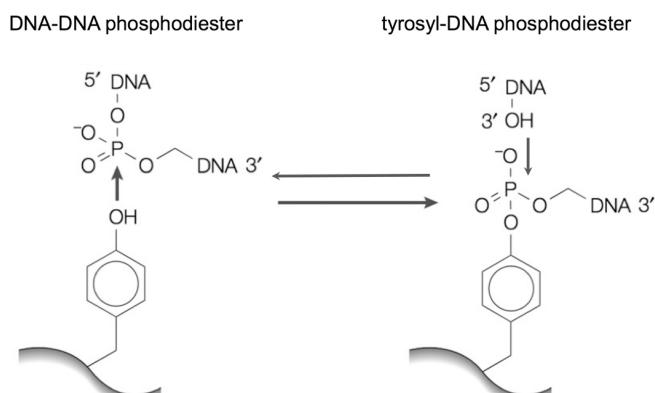


A model of a DNA replication fork (integrating data from SV40 and *E. coli*)  
(for dynamic animations please follow: <http://www.dnai.org/a/index.html>)

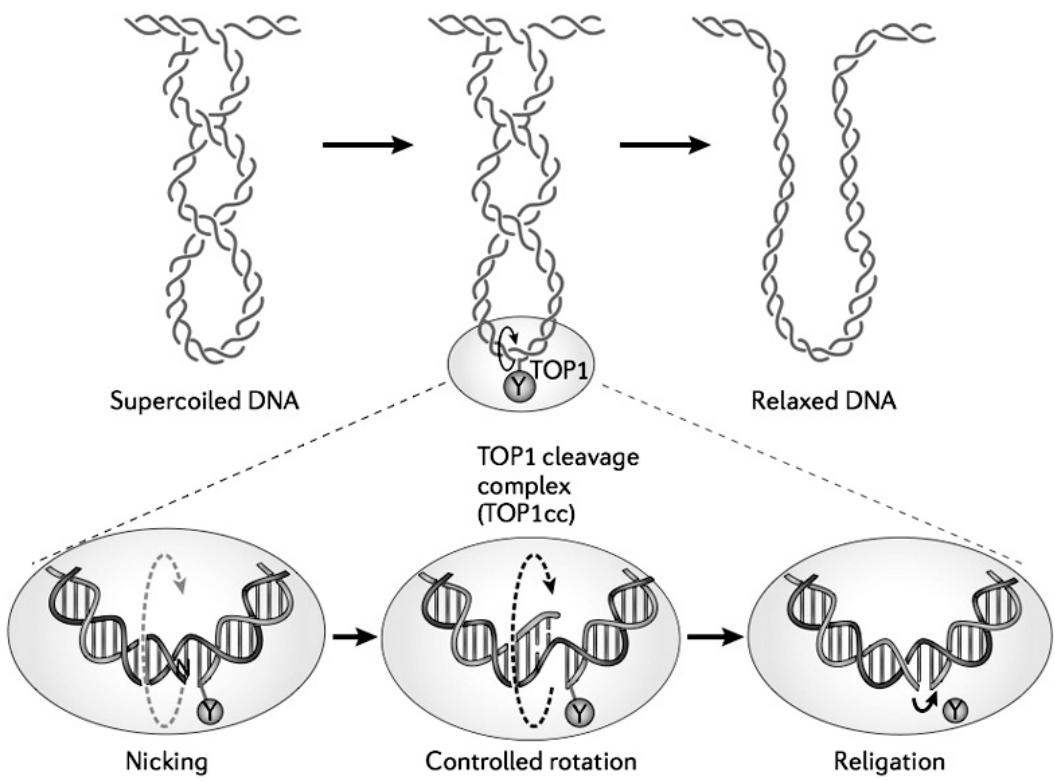
### DNA topoisomerases

The immense length of DNA in the nucleus generates topological problems: In the average human chromosome (~150 Mbp), the DNA strands wind around each other  $\sim 1.4 \times 10^7$  times. These turns must be removed during replication. In addition, separation of DNA strands during transcription and DNA replication generates (positive) supercoils ahead of the moving polymerase complexes, which would eventually prevent further elongation. DNA topoisomerases resolve these by altering the number of times DNA strands wind around each other.

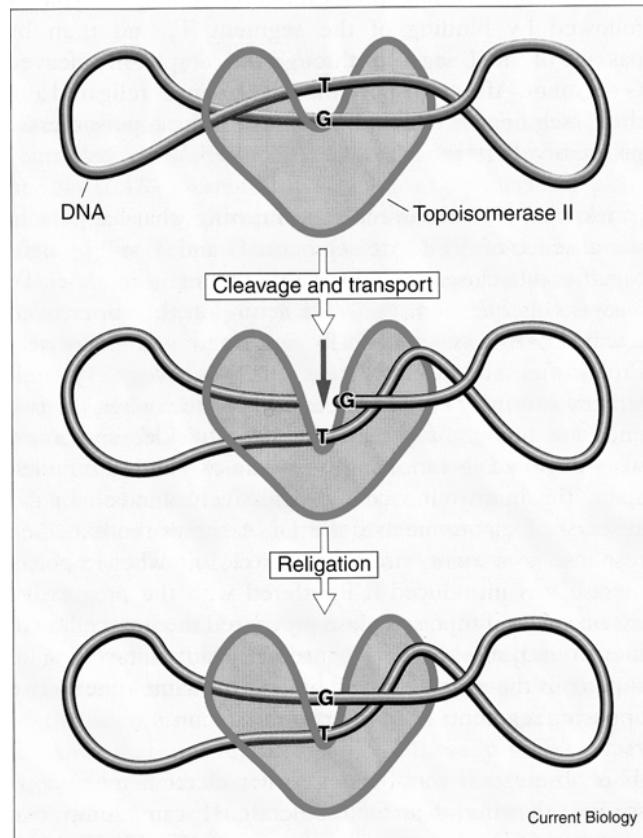
DNA topoisomerase I nicks one strand of a DNA duplex and attaches a DNA phosphate group to a tyrosine residue in its active centre covalently forming a new ester bond. This allows rotation of the free end of the cut strand around the uncut single-strand. It then seals the nick by breaking the ester bond of the DNA with the tyrosine and re-ligating the DNA, without requiring ATP. These chemical reactions are trans-esterifications.



This process can therefore remove strain imposed on a molecule by local helix unwinding as found in front of active DNA or RNA polymerases.

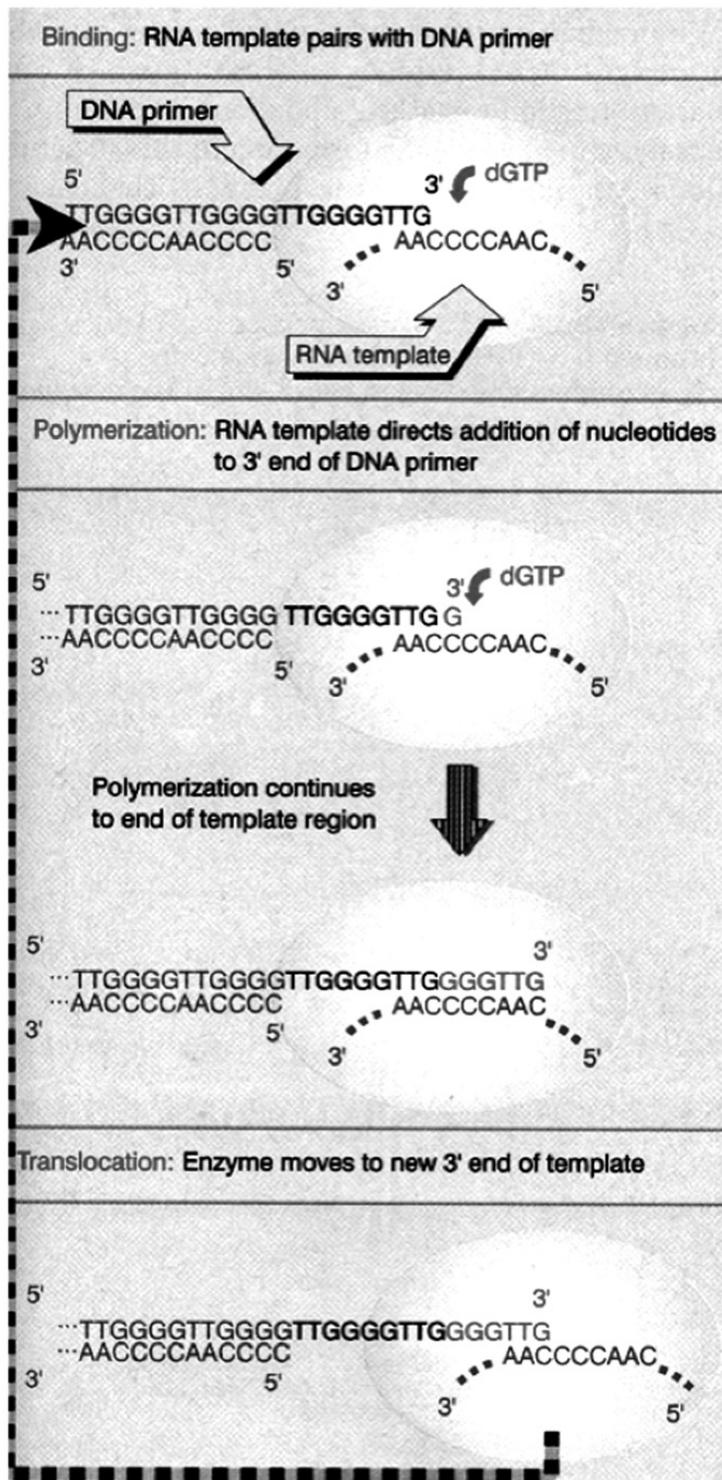


DNA topoisomerase II cuts both strands and it bridges the gap allowing other regions of DNA duplex to pass through before resealing, removing supercoils from the DNA. This enzyme can also separate interlocked DNA rings (concatemers or catenanes), a property which is essential in the final stages of DNA replication and during mitosis.



### Replication of telomeres

The mechanism of co-ordinated leading and lagging strand synthesis leads to a loss of DNA at the linear ends of the chromosomes, the telomeres. During every replication cycle the ends would be shortened. To counteract this loss, the enzyme telomerase can elongate the ends by synthesising and adding new telomere repeats onto the ends, by using its own RNA template.



## LECTURE 5

## EUKARYOTIC CHROMOSOME REPLICATION III

### Cell cycle control and chromatin dynamics

#### References:

Alberts et al., MBC VI, pages 974-982  
Lodish et al., MCB VIII, pages 874-904

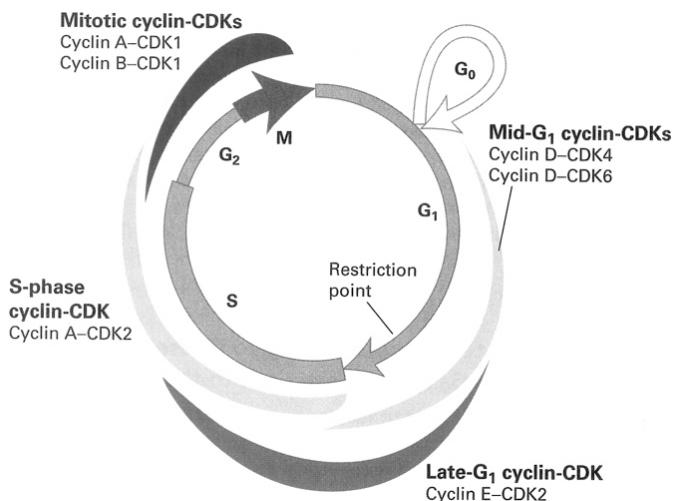
#### For enthusiasts:

Sclafani and Holzen (2007) Ann. Rev. Genet. 41, pp237-280  
Corpet and Almouzni (2009) Trends Cell Biol. 19, 29-41  
Webster et al. (2009) J. Cell Sci. 122, 1477-1486

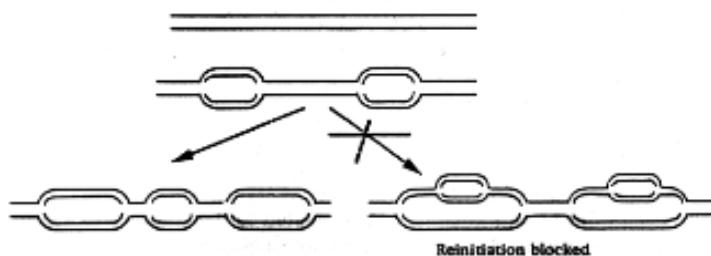
#### Cell cycle control of DNA replication

DNA replication is tightly controlled during the cell division cycle: the entire genome is replicated precisely once in S phase and separation of the replicated chromatids occurs in mitosis. Both events are strictly separated!

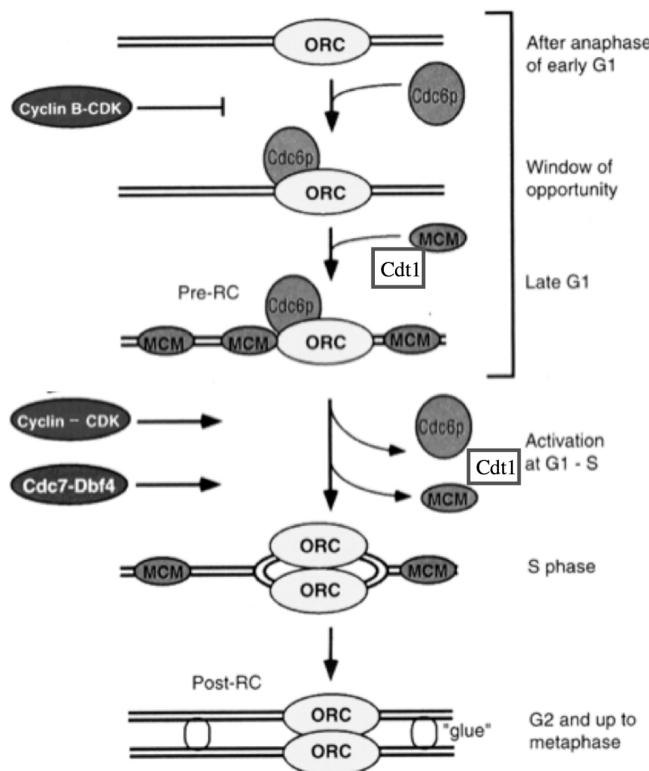
What controls the start of DNA synthesis in eukaryotic cells? Growing evidence points to cyclin-dependent protein kinase (CDK) complexes, related to the complex of cyclin B and CDK1 that controls mitosis. In vertebrates, prime candidates are cyclin A-CDK2 and cyclin E-CDK2 complexes. The Dbf4-Cdc7 protein kinase (DDK) is also crucial for origin activation and the initiation of DNA replication.



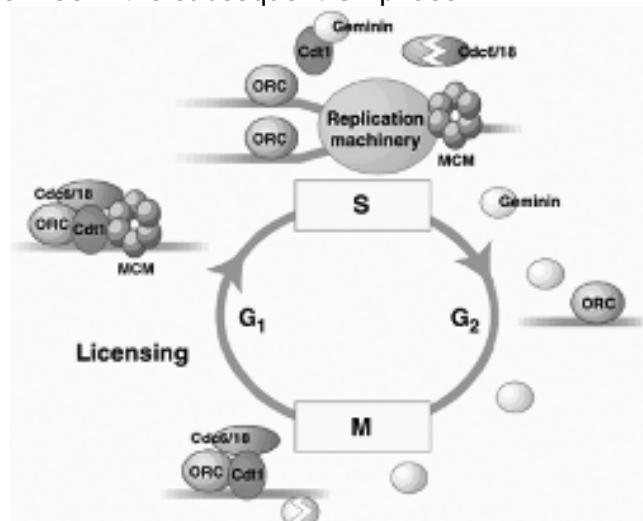
Once replication is initiated at an origin, re-initiation is prevented:



How is the re-initiation of DNA replication prevented? At each replication origin the essential pre-replication complex (pre-RC) or 'replication licence' is assembled following exit from mitosis. As discussed in lecture 4, it consists of ORC, Cdc6, Cdt1 and MCM (minichromosome maintenance) proteins and is required for initiation. After DNA replication is initiated, the pre-RC is dismantled (Cdc6 and Cdt1 are degraded by proteolysis and MCM complexes are displaced from replicated DNA). Reformation of new pre-RCs and re-reinitiation of DNA replication are thus prevented until exit from mitosis. A first level of control is exerted by CDKs. High CDK activity is essential for origin firing in S phase and for preventing pre-RC re-assembly in S and G2 (CDK activities are low in G1). This basic mechanism is conserved from yeast to humans.

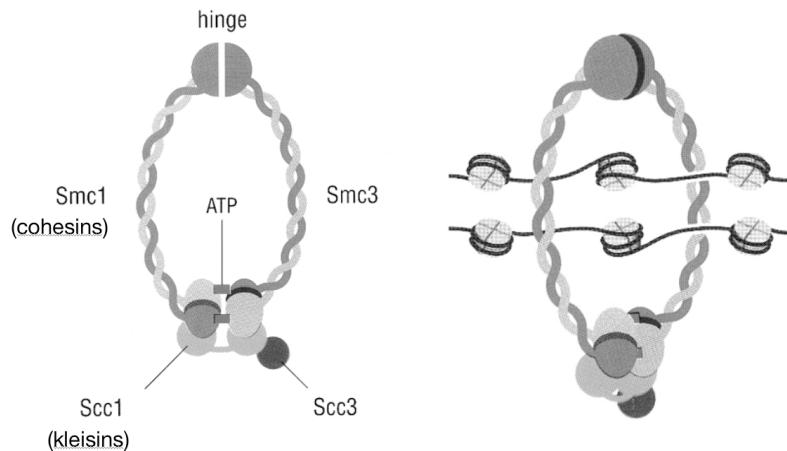


A second, independent level of control has recently been found in multicellular organisms, involving Cdt1 and the protein Geminin. Geminin binds to and inactivates remaining Cdt1 in S and G2 phase, thus preventing a re-assembly of new pre-RCs after initiation of DNA replication. In mitosis, Geminin is degraded allowing Cdt1 to assemble new pre-RCs in the subsequent G1 phase.



### Sister chromatid cohesion

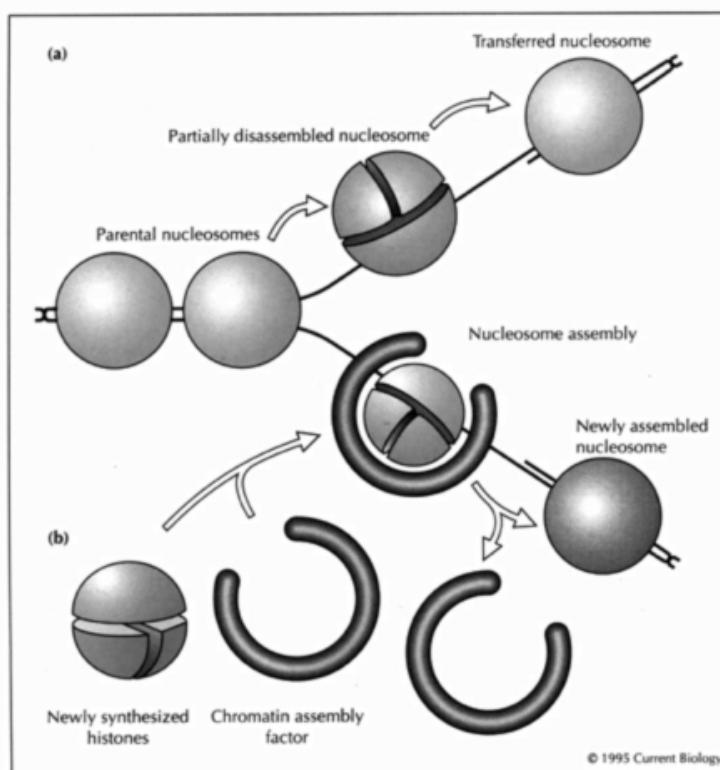
Newly replicated sister chromatid fibres are physically held together until the metaphase to anaphase transition in mitosis. This sister chromatid cohesion is mediated by cohesins, proteins also belonging to the class of 'structural maintenance of chromosomes' proteins (SMCs).



### Chromatin assembly

During chromosome replication in S phase not only must the entire genomic DNA be replicated, but also the chromatin structure it is assembled into. In front of an advancing replication fork, chromatin partially disassembles, and parental nucleosomes are transferred past the replication fork machinery and their histones are recycled.

New histones are synthesised during S phase of the cell cycle and assembled into nucleosomes on replicated DNA by assembly factors.

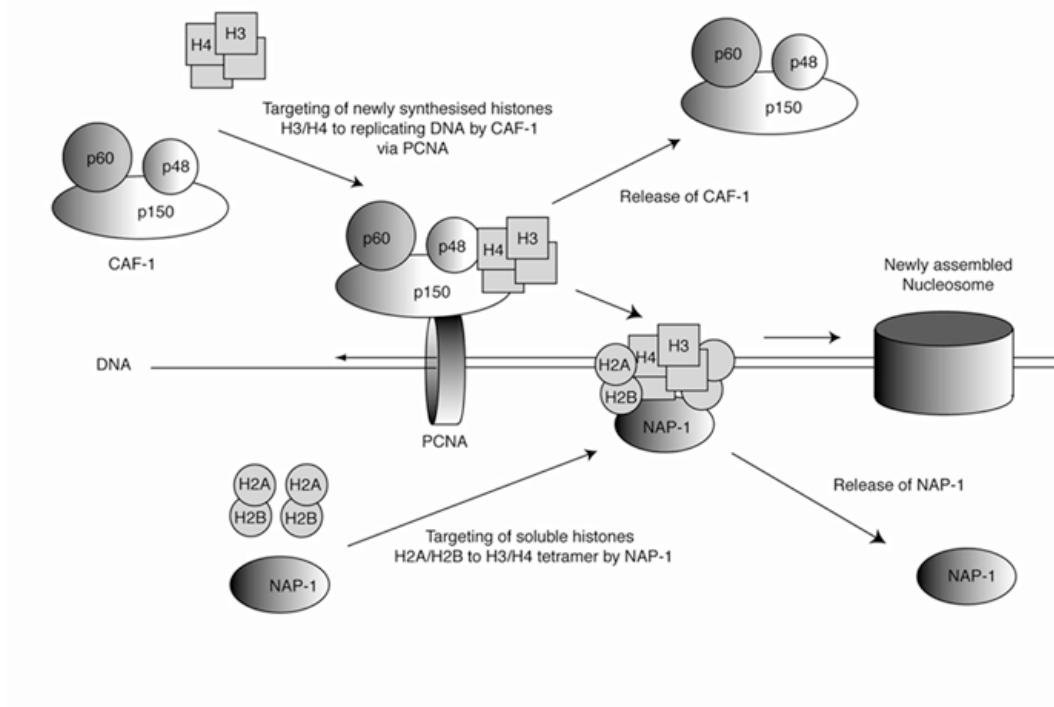


Chromatin assembly at the replication fork

Histones and DNA can, in principle, self assemble to form nucleosome cores but this process is mediated by other proteins in the cell. In Xenopus embryos proteins called N1 and nucleoplasmin are associated with histones and will assemble nucleosome cores at physiological ionic strength in vitro.

The chromatin assembly factor CAF-1 has been isolated from human and other cells, which facilitates replication-dependent nucleosome assembly in the SV40 replication system. It interacts with the replication fork protein PCNA and thus targets newly synthesised histones H3 and H4 to the replication fork. Other assembly proteins (such as Asf1 and NAP-1/2) act synergistically with CAF-1 to assemble entire new nucleosomes.

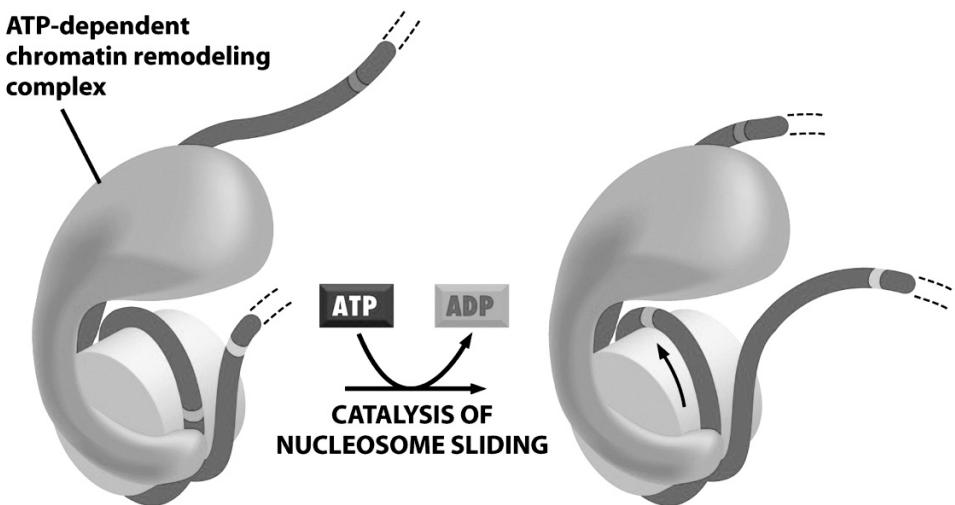
Old tetramers of histones H3 and H4 stay together during replication, and after transfer to a replicated DNA daughter strand they can associate with either new or old dimers of histones H2A and H2B. Linker histones (H1) associate later and higher order structures are formed.



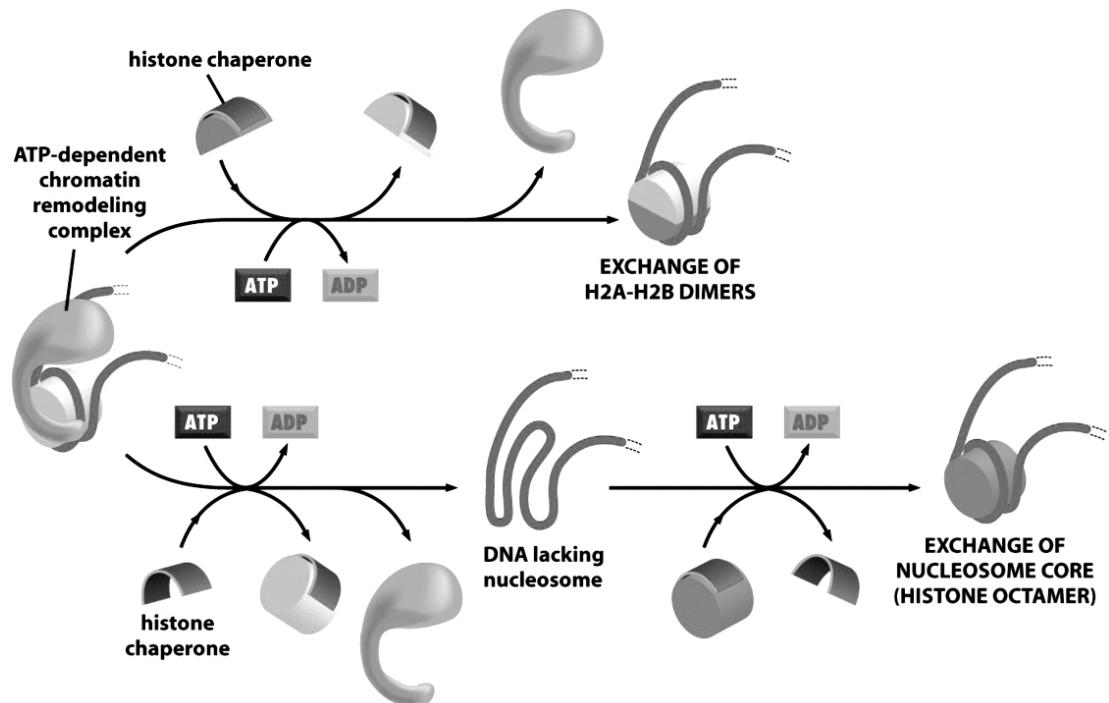
Chromatin assembly factors in action

### Chromatin remodelling

Once assembled, chromatin fibres are not static. Factors have been discovered that remodel chromatin in an ATP-dependent manner by sliding nucleosomes along the DNA fibre. These remodelling factors are usually large multi-subunit complexes.



Several remodelling factors are capable of exchanging (and evicting) histones within assembled nucleosomes, utilising histone chaperones as co-factors. This allows for an exchange of histone types or reprogramming of epigenetic marks such as histone modifications.



In conclusion, chromatin remodelling allows the chromatin fibre to be dynamic and therefore able to react to metabolic requirements arising from DNA replication, repair and transcription. A key mechanistic feature is that DNA binding factors will thus be able to gain access to sites on DNA, which might be otherwise occluded by nucleosomes.

## LECTURE 6: NUCLEAR TRANSPORT

### Import and Export

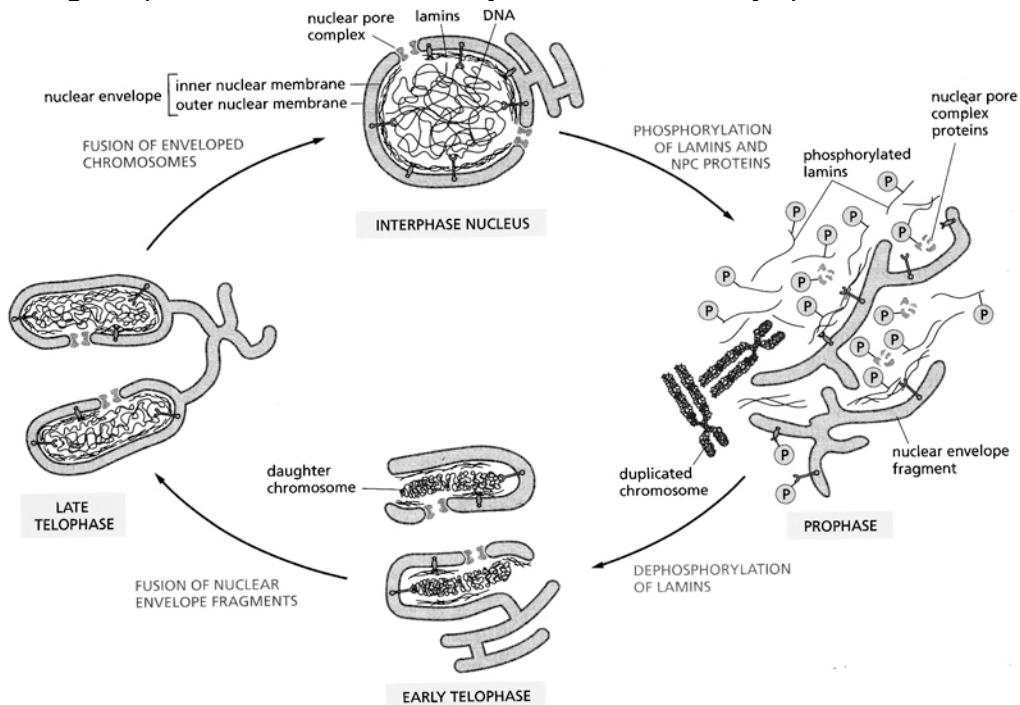
**References:** Alberts et al., MBC VI, pages 649-658  
Lodish et al., MCB VIII, pages 622-628

**Further reading:** Görlich & Kutay (1999), Annu. Rev. Cell Dev Biol 15, 607-660  
Allen et al. (2000), J Cell Sci 113, 1651-1659

**For enthusiasts:** Weis (2007), Cell 130, 405-407  
Schmidt & Görlich (2016), Trends Biochem Sci 41, 46-61

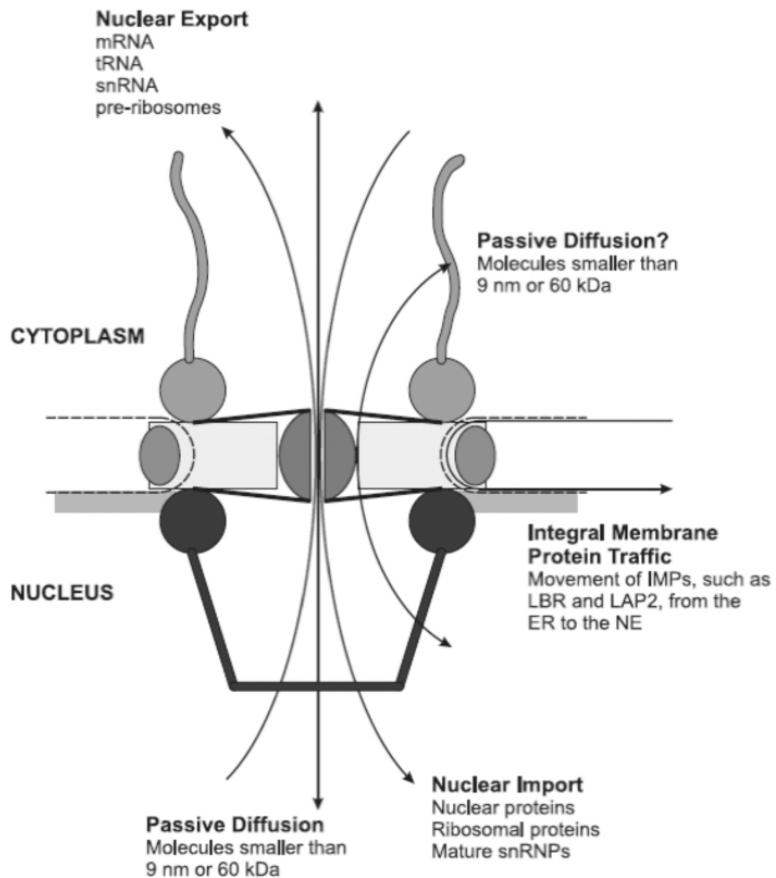
### Nuclear disassembly and re-assembly during mitosis

In higher eukaryotic cells, the nucleus disassembles in the prophase of mitosis, following complete chromosome replication in the previous S phase. Increased protein phosphorylation (mostly by cyclin B-CDK1) result in nuclear envelope breakdown. The lamina depolymerises into soluble lamin A/C and membrane-associated lamin B. Nuclear pore complexes (NPCs) disassemble into soluble nucleoporin subcomplexes. The nuclear membranes fragment into vesicles or tubes. Chromatin condenses into separated chromosomes until metaphase and associates with the spindle apparatus. During this period, chromatin is directly accessible to the cytoplasm.



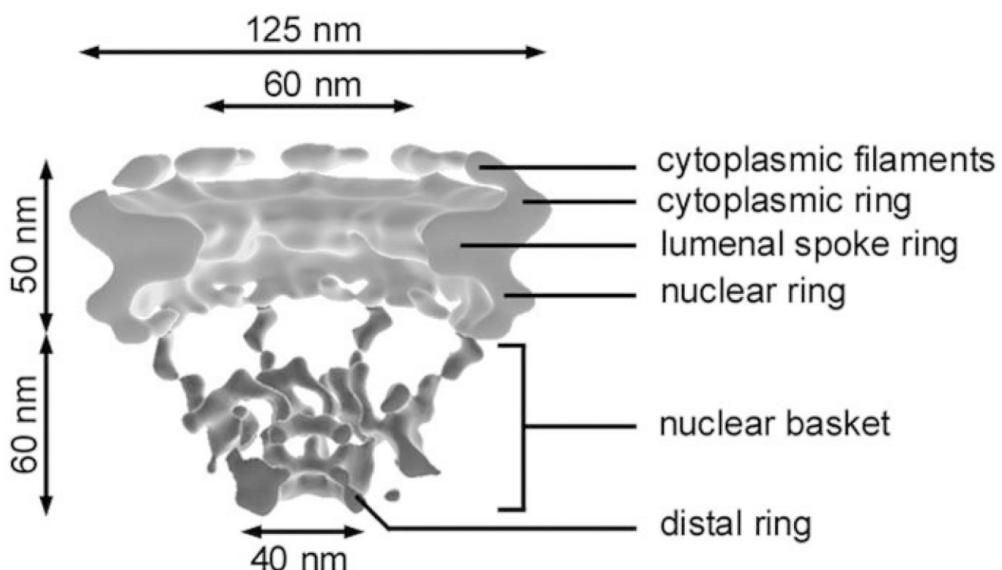
Following inactivation of CDK-complexes and a sharp fall of kinase activity in anaphase, lamins and NPC proteins become dephosphorylated. Nuclei re-assemble in telophase after successful sister chromatid separation, in a phenomenological inversion of the disassembly process. It involves chromosome decondensation, membrane assembly from vesicles, lamina polymerisation and NPC assembly from soluble nucleoporin subcomplexes.

Following nuclear envelope assembly, several million proteins and RNA molecules must travel across the nuclear envelope per minute during interphase. All of this nuclear transport is mediated through the nuclear pore complexes. How?



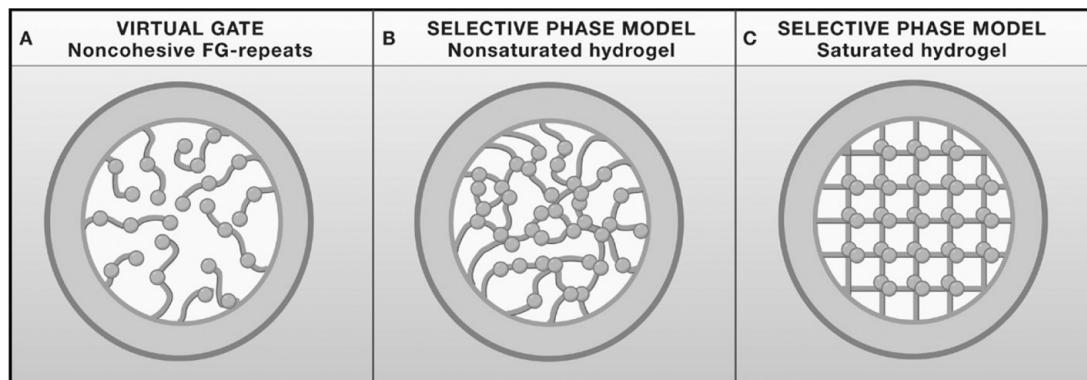
### Nuclear pore complexes

Image reconstructions from electron micrographs of isolated NPCs show a ring of eight subunits surrounding a central channel through which proteins and RNA pass. High-resolution studies have provided a good understanding of the internal structure of the nuclear pore complex and its physical dimensions. Fibrils project from both surfaces of the nuclear pore complex. Those on the inside are organised as baskets.



### **What occupies the central channel (gate) through the NPC?**

Structural studies have provided evidence for the presence of a protein structure lining and occupying the central gate of the NPC. It consists of so-called nucleoporin proteins, which are very rich in repeats of the two hydrophobic amino acids phenylalanine and glycine (FG). FG repeats can interact with each other and form a dynamic hydrogel. Several models about this structure have been postulated:



The selective phase model is currently favoured: the hydrogel of cohesive FG-repeat proteins would form a barrier against diffusion of general macromolecules whilst providing a solvent for translocating molecules.

### **Size limit to rapid entry**

The nuclear envelope prevents nuclear entry of larger molecules, e.g. labelled dextrans of various sizes, but allows passive diffusion of small molecules. This allows calculation of an effective pore diameter of 9nm. Therefore the effective pore diameter, which is controlled by the physical properties of the hydrogel, is much smaller than the 60nm channel of the pore complex. This is too small for proteins larger than 40-60kDa. Therefore, how do large proteins and large RNA molecules interact with the hydrogel and cross the NPC?

### **Import of nuclear proteins**

Nuclear proteins of all sizes can accumulate in the nucleus by >100-fold. In microinjection experiments into the cytoplasm, nuclear proteins re-accumulate in the nucleus. This is achieved by selective entry through the nuclear pore complexes. The underlying molecular mechanism has been elucidated by studying the import of the nuclear proteins nucleoplasmin and SV40 large T antigen:

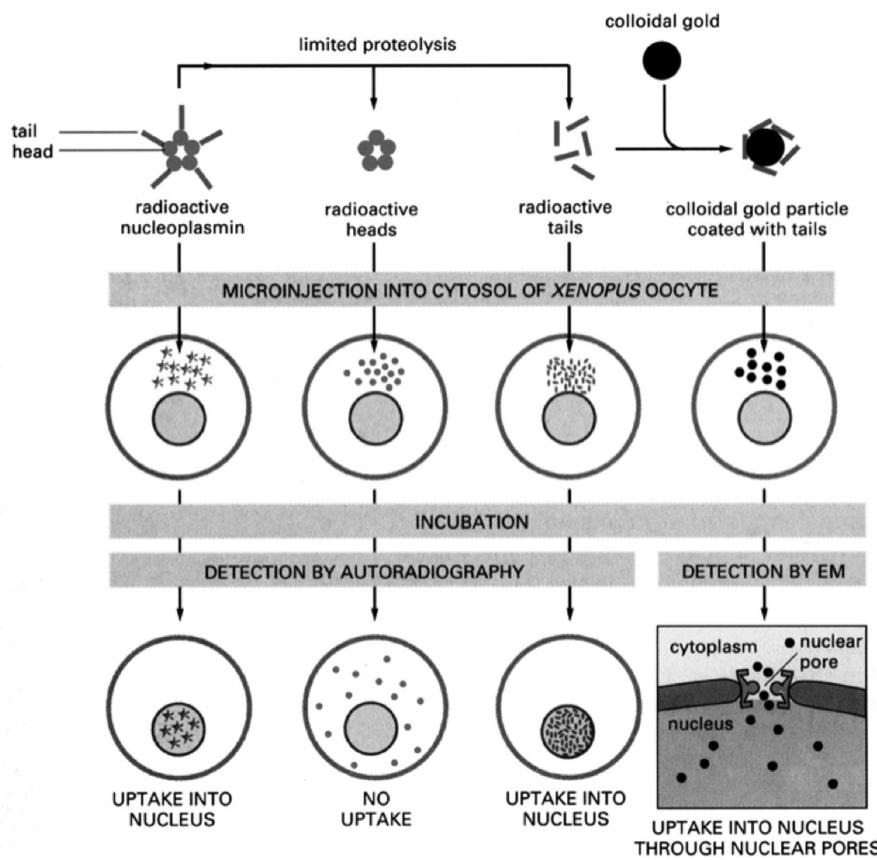
### **Nuclear localisation signals (NLS)**

Proteins localising to the nucleus (e.g. nucleoplasmin and SV40 large T antigen) contain small peptide motifs that cause their import. When fused to other non-nuclear proteins (e.g. BSA) they cause nuclear accumulation of the fusion protein. Mutations of certain lysines of nucleoplasmin or SV40 T antigen to threonine or asparagine abolish transport. The bold amino acids of the respective NLS were found to be essential:

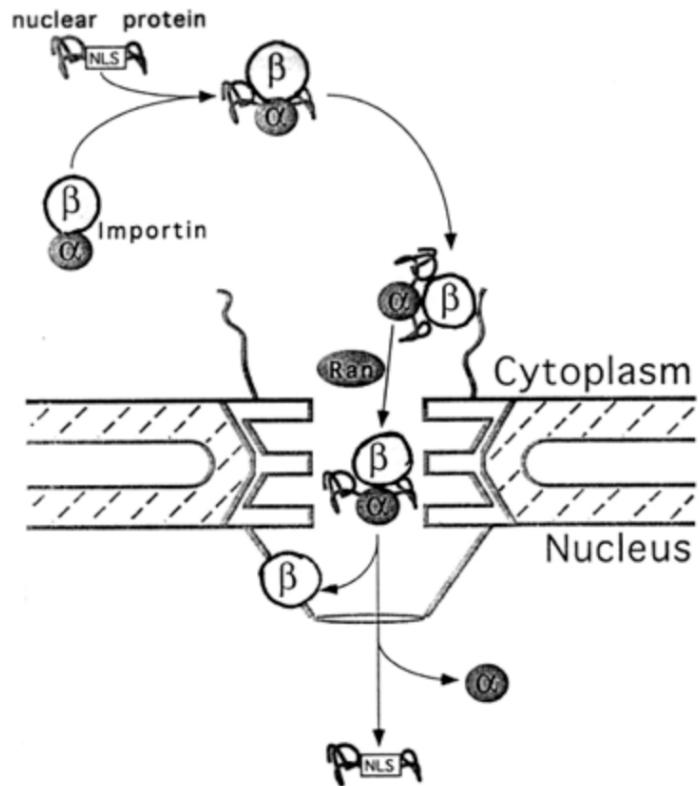
Nucleoplasmin:           **lys arg** ---- 10 ---- **lys lys lys lys**

SV40 T antigen:          pro **lys lys lys arg lys val**

Injection of colloidal gold coated with nucleoplasmin into the cytoplasm reveals passage through the central channel of the NPC.

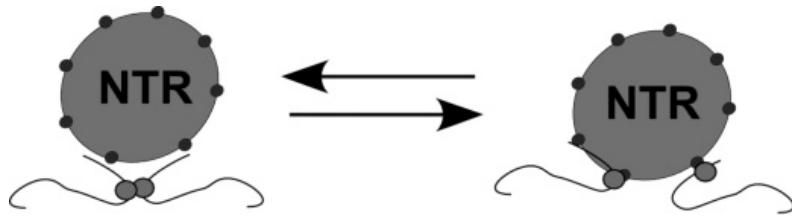


Transport occurs in two steps: rapid binding of the cargo to the cytoplasmic side of the nuclear pores, followed by a slower energy-dependent translocation through pores. Two soluble key proteins are required for these two steps, importin and Ran, a small GTP'ase.

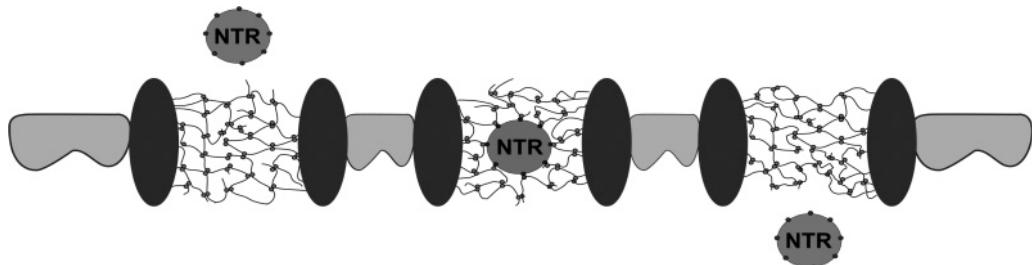


The  $\alpha$  subunit of importin binds the NLS of the cargo and the  $\beta$  subunit docks at the nuclear pore complex. Following transport through the NPC, nuclear ran-GTP causes dissociation of importin from the NLS-cargo. Importin  $\alpha$  and  $\beta$  are separately exported back into the cytoplasm in a ran-GTP dependent manner.

Importin  $\beta$  reversibly interacts with FG repeats of the NPC.



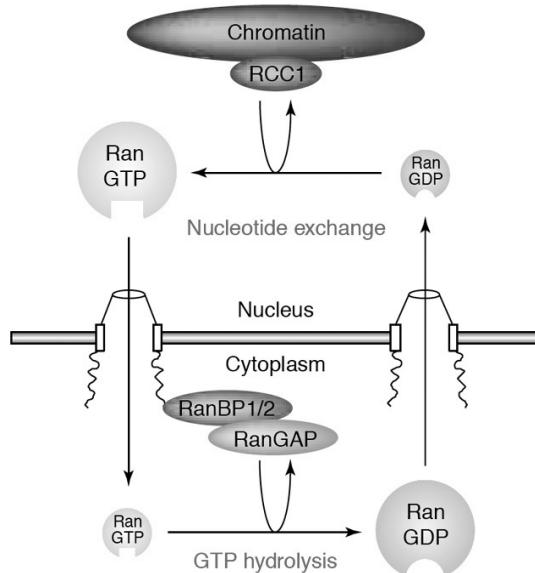
It can therefore cross the hydrogel and act as the import mediator through the 'gate'.



### The ran cycle

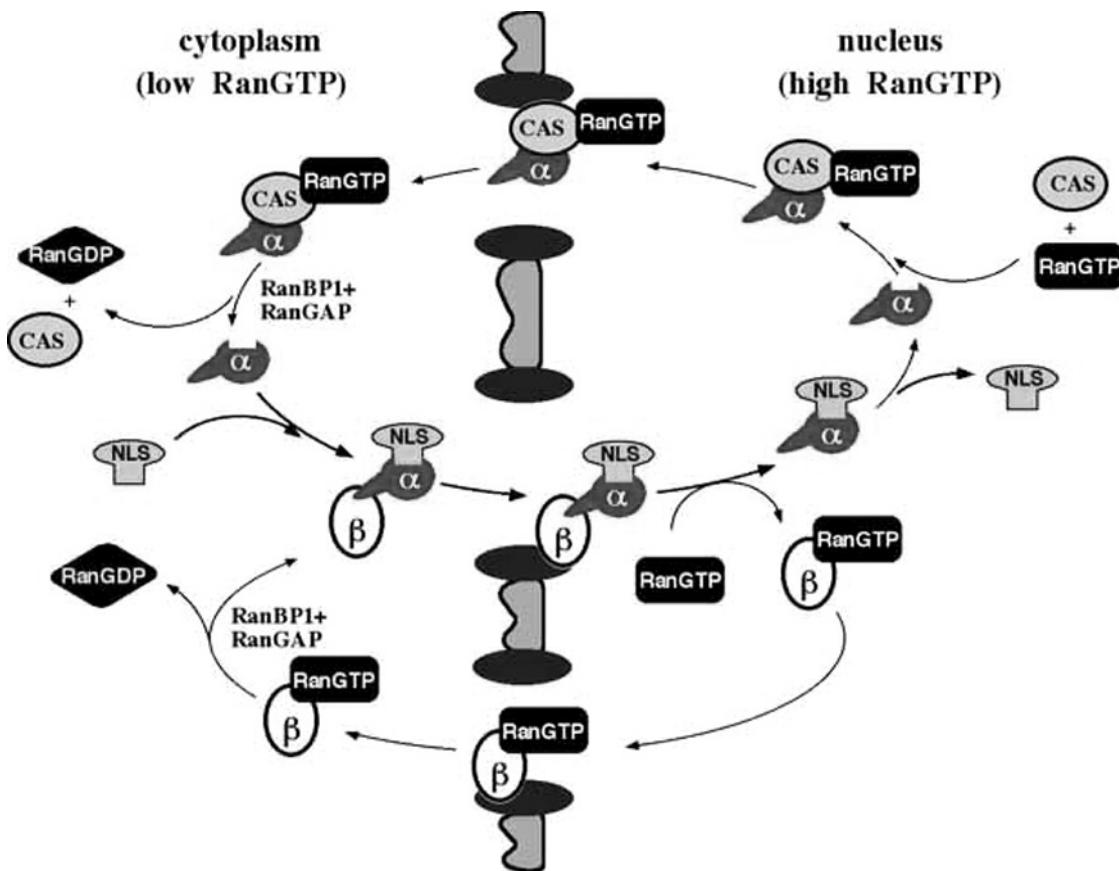
As seen above, the small GTP'ase ran is involved in nuclear import (and export). A ran-cycle model has been proposed:

Ran can bind GDP or GTP. Ran GDP is predominantly cytosolic and ran GTP is predominantly nuclear. In the cytosol, a ran-specific GAP (GTPase activating protein) stimulates the endogenous GTPase of ran and converts it to the GDP form. In the nucleus the chromatin-bound nucleotide exchange factor RCC1 promotes exchange of the bound GDP to GTP, thus building a ran-GDP/ran-GTP gradient across the nuclear envelope. Ran GDP is required for cargo binding to importins in the cytosol and ran GTP promotes cargo dissociation from importins in the nucleus. Conversely, ran-GTP is involved in cargo binding to exportins in the nucleus (see below).



## Regulated import

To regulate nuclear import of NLS-containing cargo, the ran cycle is coupled to the importin cycle, as summarised in this diagram:



## Export of Proteins and RNA

RNA is exported as protein complexes. It is signal-dependent and carrier-mediated. It occurs through the nuclear pore complex. Giant mRNA transcripts from insect salivary glands (35–40kb) unfold to pass through the pore, always 5' end first. Before mRNA is exported, it must be processed correctly, including splicing and poly A addition.

## Nuclear Export Signals (NES)

Export signals have recently been identified in several exported proteins, including some that bind RNA.

HIV-1 Rev:

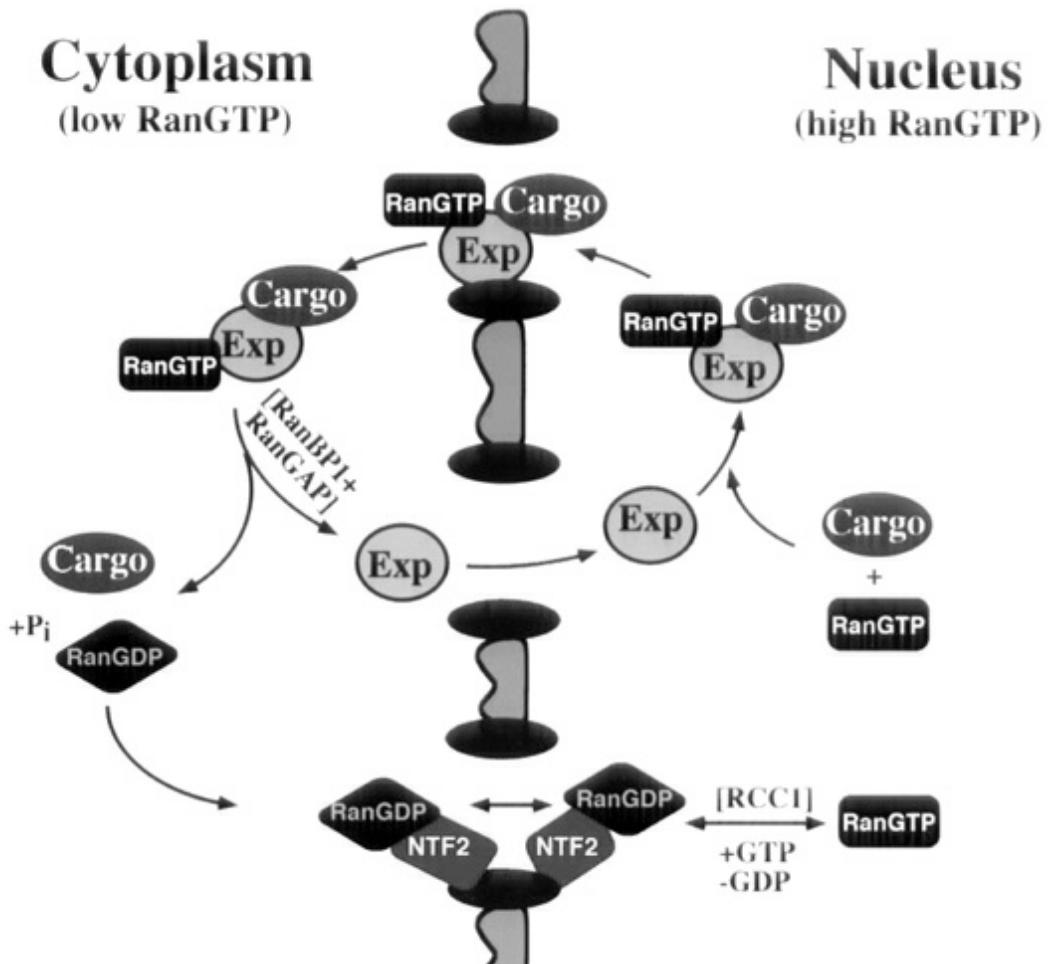
Leu Pro Pro Leu Glu Arg Leu Thr Leu

Protein kinase inhibitor:

Leu Ala Leu Lys Leu Ala Gly Leu Asp Ile

## Regulated export

Export signals are recognised by export receptors (exportins) that are related to importin  $\beta$ . The regulation of cargo export again depends on the ranGTP/ranGDP gradient across the nuclear envelope. Ran protein levels are equilibrated across the envelope by the nuclear transport factor NTF2.



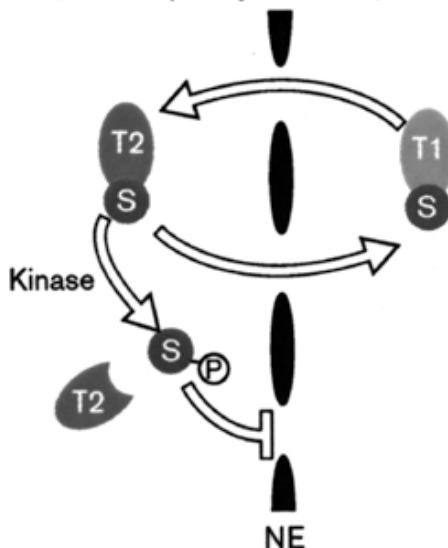
HIV unspliced transcripts bind to the Rev protein, which mediates export via interaction with an exportin. Small RNAs bind their own exportins directly: tRNAs interact with exportin-t, and other small non-coding RNAs such as microRNA precursors interact with exportin 5.

### Higher levels of regulated transport

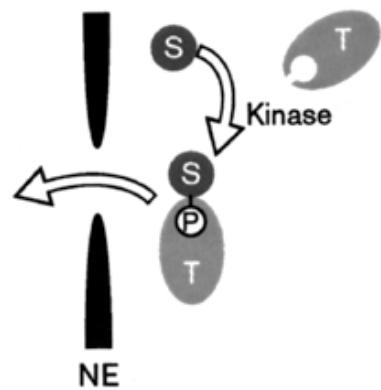
Some proteins can only enter the nucleus when released from cytoplasmic anchors, e.g. several transcription factors, providing a possible level of gene regulation. Phosphorylation of NLS or NES elements is currently discussed as another means to regulate nuclear transport. For example, cyclin B1 constantly shuttles between nucleus and cytoplasm in S and G2 phase, but phosphorylation of a NES in early mitosis blocks export and results in nuclear accumulation.

**(a) Unidirectional control**

(i) Shuttling substrate  
(for example, cyclin B, Pap1)

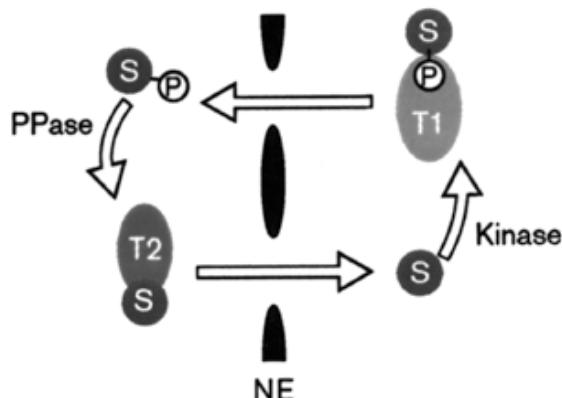


(ii) Non-shuttling substrate



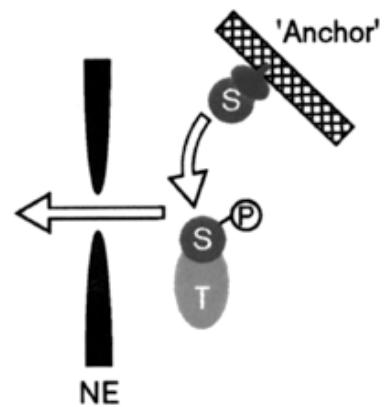
**(b) Binary switch**

(for example, Hog1p, Pho4p)



**(c) Anchor/release**

(for example, Spc1[Sty1]?)



End of Lectures 1-6