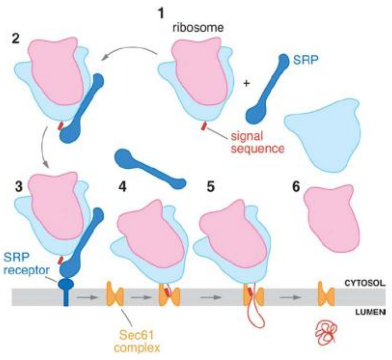
Targeting and Transport of Proteins to the ER

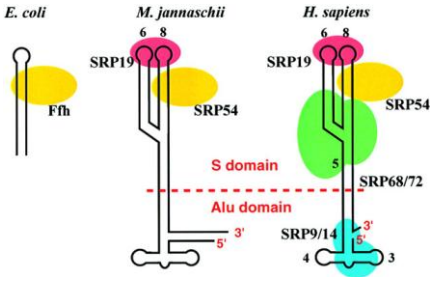
Translocation of Proteins across the ER Membrane



Proteins are synthesized by ribosomes in the cytoplasm, further modified and finally brought to their right location. The signal hypothesis describes how targeting signals direct the transport of proteins. Example: SRP (signal recognition particle) binds to a signal sequence in a growing polypeptide chain as well as to the ribosome, then the entire complex is targeted to the membrane by an interaction of the SRP with its receptor. Via a translocon complex, the ribosome synthesizes the protein directly into the ER (cotranslational translocation).

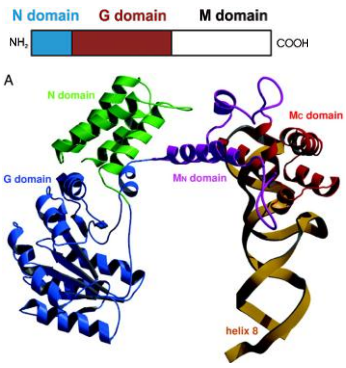
ER Signal Sequence: Hydrophobic core (7-15 amino acid residues), hydrophilic and basic N-terminal and a polar C-terminal (2-9 polar, small amino acid residues, consensus site for cleavage by signal peptidase). The ER signal sequence targets proteins for secretion and membrane insertion and is located at the N-terminus of a pre-protein. Its total length is 15-25 amino acid residues.

Signal Recognition Particle (SRP): Recognizes signal sequence, conserved, found in the cytoplasm. The S-domain recognizes the signal peptide and interacts with the SRP receptor (SR) and the Alu-domain mediates translational arrest (elongation is inhibited). The conserved component is SRP54 (Ffh=fifty four homologue).



Signal Recognition Particle Cycle:

- Binding of SRP to the ribosome and the signal peptide via SRP54
- Elongation arrest via the Alu domain
- Docking on the SR via SRP54-SRα interactions
- Docking of ribosome on translocon and release of SRP from SR by GTP hydrolysis



SRP54: Three different domains, N, G and M. The methionine-rich, hydrophobic M domain binds RNA and recognizes the hydrophobic core of signal peptide (SP). The G domain has a GTPase function (unusually low

affinity for GTP). GTP-binding is required for the interaction with the receptors SRα on the ER membrane. Between M and G there is a flexible linker region which likely communicates SP-binding to the G domain. After binding of the signal peptide, GTPase is activated and the SRP undergoes a conformational change. The Alu domain then competitively inhibits elongation factor binding (eEF2 promotes the translocation step of aminoacyl-tRNA from the A to the P site during protein synthesis) by covering the same site on the ribosome and therefore blocks elongation. Now, the SRP can target the ribosome-protein complex to the SRP receptor. This co-translational targeting couples translation and translocation and therefore helps prevent misfolding of the nascent protein in the cytosol.

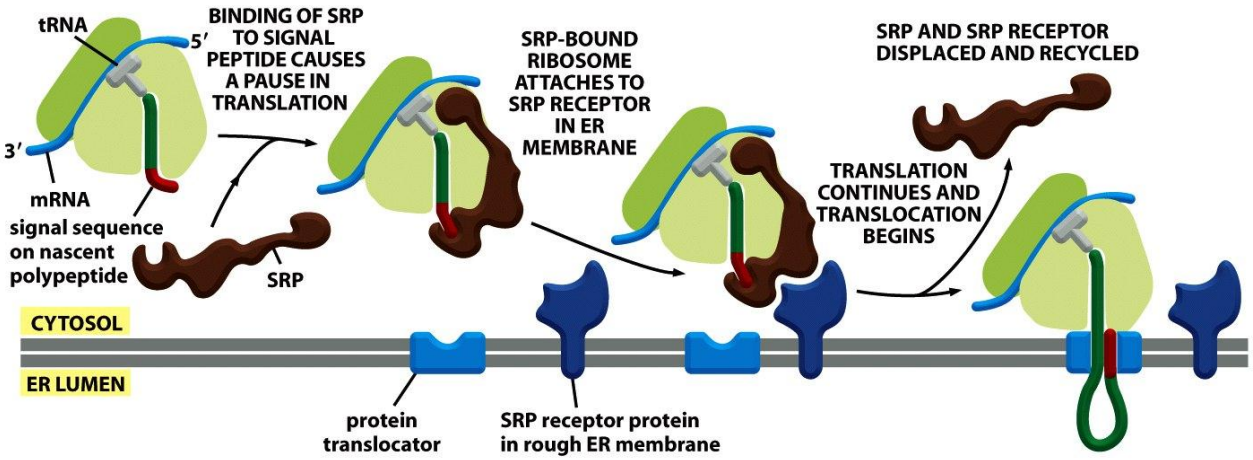


Figure 12-40 Molecular Biology of the Cell 5/e (© Garland Science 2008)

The SRP Receptor: Consisting of two subunits, SRα (FtsY) and SRβ, and SRα resembles SRP54 (both have the N and the G domain). The complex of Ffh (SRP54) and FtsY (SRα) NG leads to a quasi-two-fold symmetrical heterodimer. Primarily the G domains have an extensive interaction surface and there are major conformational changes in the dimer form compared to the free proteins. SRP and SR reciprocally stimulate each other's GTPase activity and are dissociated after GTP hydrolysis. The two GTPase sites form a composite active site with the nucleotides packed in a head-to-tail manner. Hydrogen bonds form between the 3'OH ribose of one nucleotide and the γ-phosphate of the other. GTP hydrolysis loosens the interactions of the γ-phosphate which leads to the dissociation of the complex. Summarized: Complex formation, GTPase activation -> GTP hydrolysis -> complex dissociation (all highly coupled).

Shortly:

- Conformational changes in SR and SRP -> ready for interaction
- SR and SRP bind at translocon, ribosome nascent chain complex (RNC) possibly delays GTP hydrolysis
- SRP delivers the RNC to the translocon before dissociation from the SR, GTPase activation

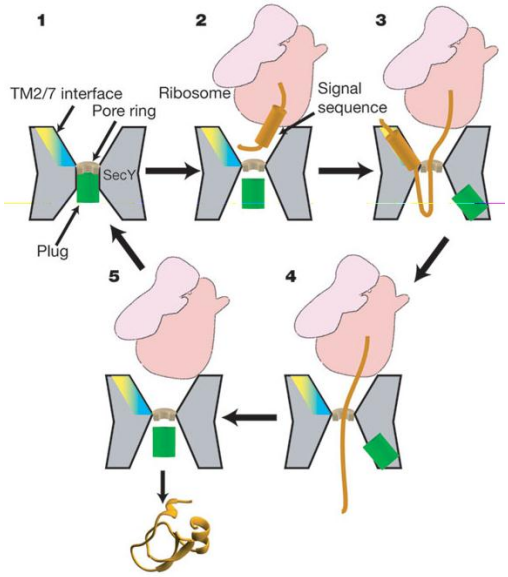
- Dissociation upon GTP hydrolysis

SRP searches for signal peptides on nascent chains, upon signal sequence binding, conformational changes are transmitted to the GTPase core, then binding of SR. Association between the NG domains of SRP54 and SR releases SRP54 N domain from 60S, translocon-binding site on ribosome exposed now, then docking of RNC to membrane.

Translocon:

- SecY, SecE and SecG in bacteria
- Sec61α, Sec61β and Sec61γ in eukaryotes

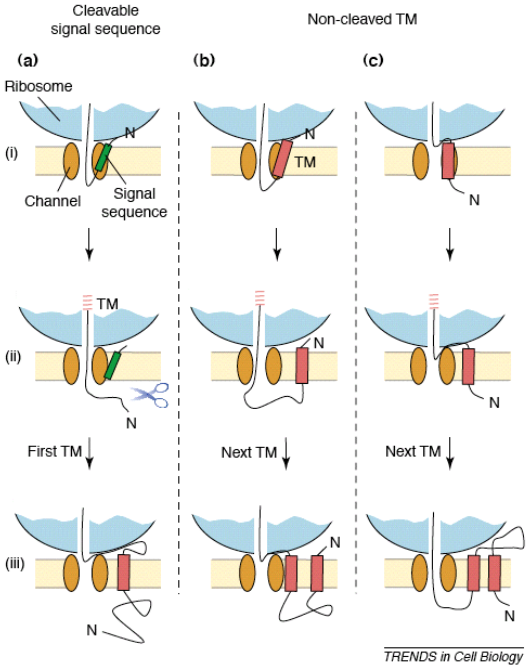
a-subunit made up of 10 transmembrane helices (TMH) constituting two halves (TMH1-5 and TMH 6-10). Cytosolic loops constitute binding site for the ribosome. In the closed state, helix 2a serves as a plug. In the open state, six hydrophobic residues work as a seal. These two features likely maintain a membrane barrier during membrane protein synthesis. The pore size is 5-8 Å which does not allow folded domains to get through.



Model:

1. Channel closed, plug blocks the pore
2. Binding of RNC
3. SP inserts between TM2b and TM7, plug moves away, pore opens during translocation
4. Pore forms a seal around chain, SP cleaved by SP peptidase
5. Plug returns to closed position

Membrane Protein Integration: All models on the right require lateral (sideways) exit of transmembrane (TM) regions to the lipid bilayer. Apolar sites two halves of the translocon complex open like a clam shell through hinge movements and allow TMs to get out.



Topogenesis of Membrane Proteins at the ER:

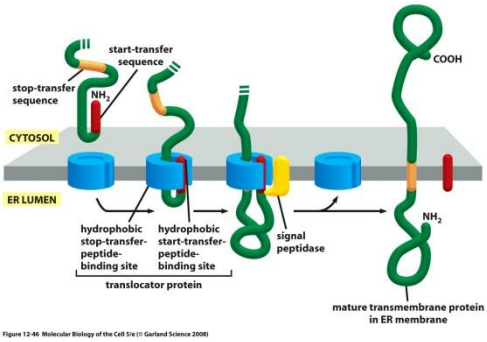


Figure 12-46 Molecular Biology of the Cell 5/e (© Garland Science 2008)

- Type I: C-terminal on the cytosolic side, N-terminal in the ER lumen, signal sequence at the N-terminus
- Type II: N-terminal on the cytosolic side, C-terminal in the ER lumen
- Type III (type Ia): as type I, but with signal in the middle of the peptide

What determines the orientation of TMHs?

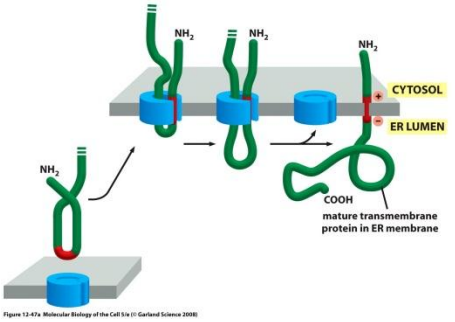


Figure 12-47a Molecular Biology of the Cell 5/e (© Garland Science 2008)

- Positive-inside rule: the more positively charged segment stays in the cytosol (charged residues flanking the hydrophobic core of the signal are important)
- Hydrophobicity of the signal:
  - N-terminal signals initially insert in the N=exo and C=cyt orientation and then invert based on their charge distribution
  - The more hydrophobic the signal, the harder to invert due to higher affinity for the translocon
- Protein folding (internal signals): Folding of hydrophilic sequences N-terminal to a signal sterically hinder N-terminal translocation

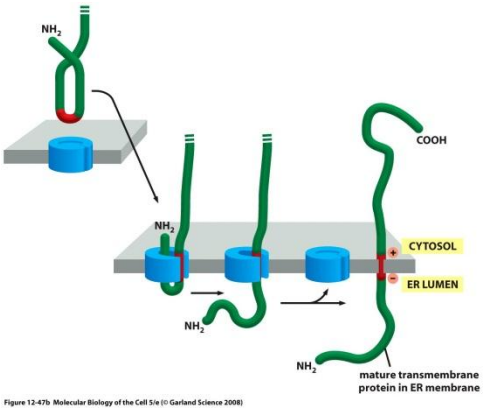


Figure 12-47b Molecular Biology of the Cell 5/e (© Garland Science 2008)

Multipass transmembrane proteins contain start and stop transfer sequences to get through the membrane several times.

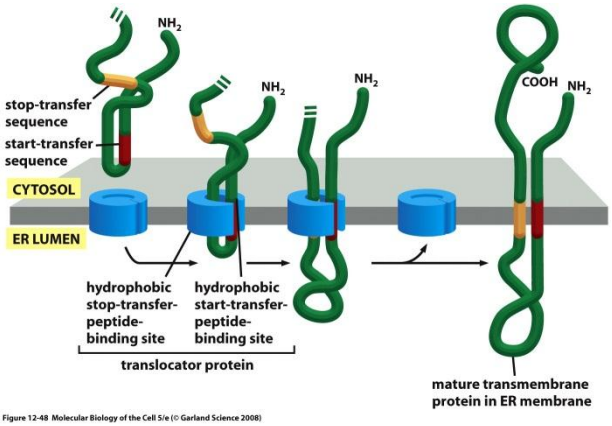


Figure 12-48 Molecular Biology of the Cell 5/e (© Garland Science 2008)

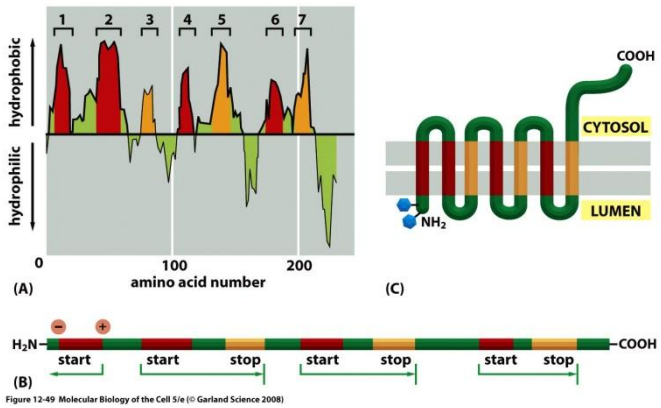
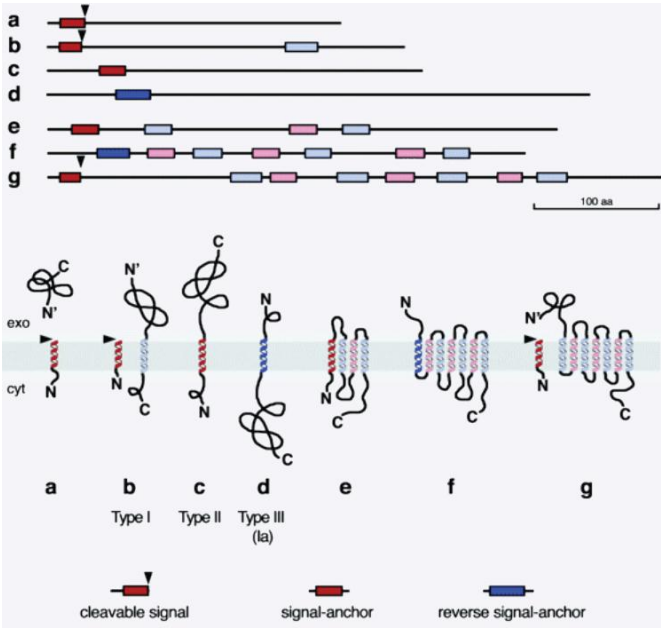
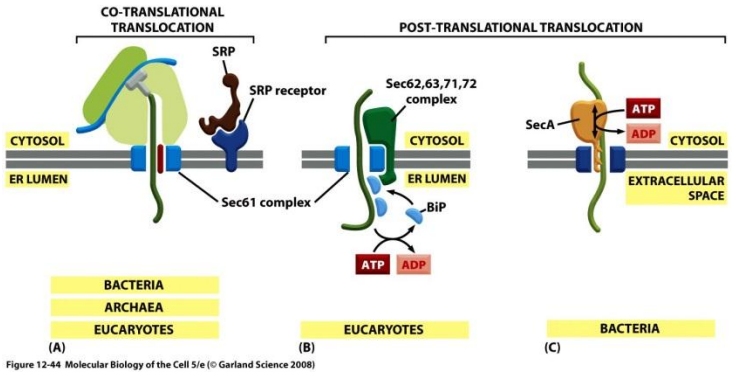


Figure 12-49 Molecular Biology of the Cell 5/e (© Garland Science 2008)



Co- and Posttranslational Targeting:

A. No additional energy source than input into translation needed
B. ATP hydrolysis by BiP (HSP70) in the ER lumen
C. ATPase activity of SecA pushes the peptide chain through the pore of translocator



Targeting and Transport of Proteins to the Mitochondria, Chloroplasts and Peroxisomes

Elements for protein transport systems:

- Signal (address)
- Chaperones and targeting factors
- Translocation machinery (receptors and channels)
- Translocation motor (hydrolysis of nucleoside triphosphates)
- Protein folding machinery

Export: Transport of proteins from the cytosol to an extracytosolic compartment (ER, bacterial plasma membrane, thylakoids). Export signals (signal sequences) mediate transport across the bacterial plasma membrane, the ER membrane and into thylakoids of chloroplasts. They are hydrophobic and usually cleaved from the pre-protein. If they are not cleaved, they can function as a signal anchor.

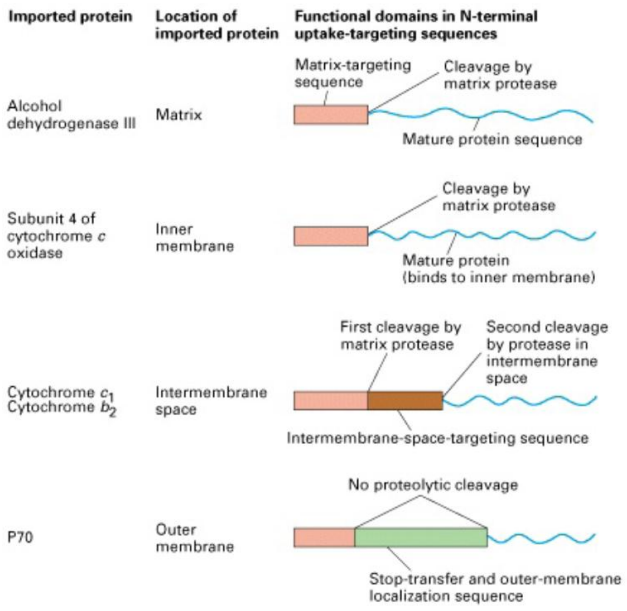
Import: Transport of proteins to compartments functionally equivalent to the cytosol (mitochondria, chloroplasts). Import signals mediate the transport into mitochondria, chloroplasts and peroxisomes and are less hydrophobic.

Target Organelle	Location of Signal Sequence	Signal Removal	Nature of Signal
ER	N-terminal	yes	Hydrophobic core
Mitochondria	N-terminal	yes	Amphipathic helix
Chloroplasts	N-terminal	yes	rich in Ser and Thr
Peroxisome	C-terminal, N-terminal	no	-SKL; N-ter nonapeptide
Nucleus	Mostly internal	no	Diverse, e.g. classical basic NLS

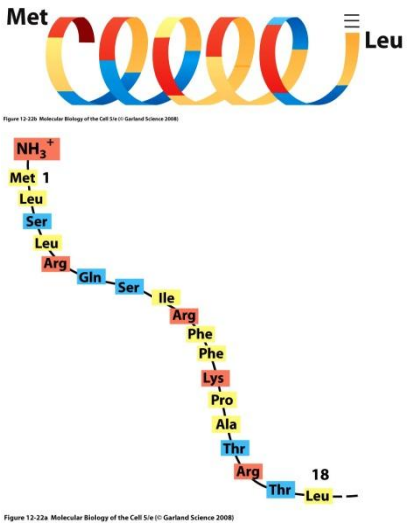
Protein Transport into Mitochondria

Some proteins that are synthesized in the cytoplasm and have to get into mitochondria are citrate synthase and other citric acid enzymes (matrix), ADP/ATP antiporter and cytochrome c oxidase subunits (inner membrane), cytochrome c and cytochrome x peroxidase (intermembrane space) and mitochondrial porin (P70, outer membrane).

To get the proteins to their different location within the mitochondria, different targeting signals are used.



The matrix targeting signal (MTS) is N-terminal and has an amphiphilic α-helix, with hydrophobic residues on one side and hydrophilic ones on the other side. It is 15-30 amino acids long (blue amino acids are uncharged, red ones are positively charged).

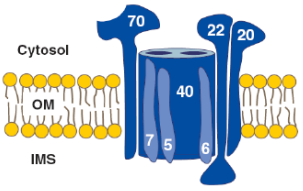


Mitochondrial Translocation Channel:

TOM: Protein translocase of the Outer Membrane

TIM: Protein Translocase of the Inner Membrane

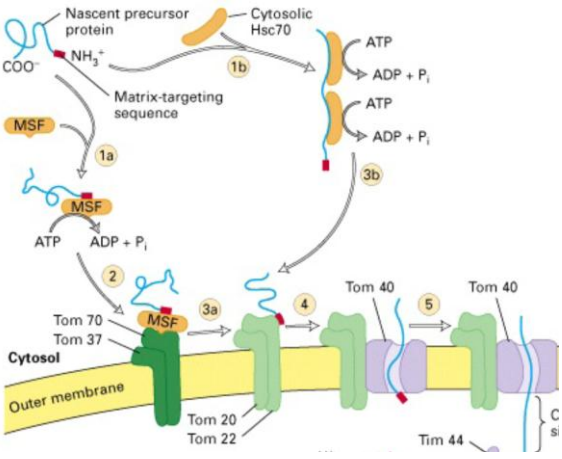
Targeting to the mitochondrial surface: Needs the soluble factors Hsc70 and MSF and the membrane receptors Tom70/Tom37 and Tom20/Tom22. The preprotein receptors Tom20/22 and Tom70 are associated with the general insertion pore complex (GIP). Tom70 mainly recognizes internal signal sequences within preproteins and Tom20 preferentially binds preproteins with N-terminal presequences. Both Tom20 and Tom70 transfer their substrates to Tom22 which functions as a central receptor within the GIP and delivers preproteins to the actual protein-conducting element of the TOM complex. There are two major ways of transporting proteins:



TOM complex

Tom70	Surface receptor
Tom40	Translocation pore
Tom22	Surface receptor
Tom20	Surface receptor
Tom7	Translocation pore
Tom6	Translocation pore
Tom5	Translocation pore

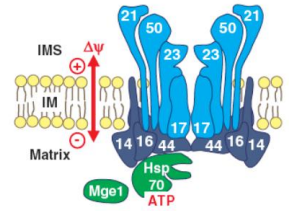
1. Chaperone binding and unfolding of protein, therefore exposure of signal sequence, then transport through Tom20/22 to Tom40 (mostly used)
2. Usage of MSF, hydrolysis of ATP, Tom70/37, then Tom20/22 and Tom40



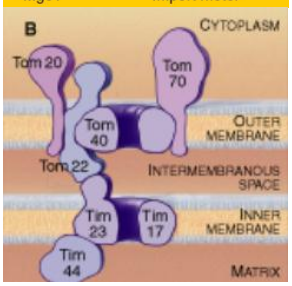
Tom20 recognizes MTS through a binding groove, which is made up mainly of hydrophobic amino acids (2 acidic conserved residues (Glu78 and Glu79), recognizing the helix), while the hydrophilic residues are found in the periphery of the groove. The MTS is bound in a helical structure with three strongly hydrophobic amino acids (leucines) oriented towards the hydrophobic patch of the binding groove.

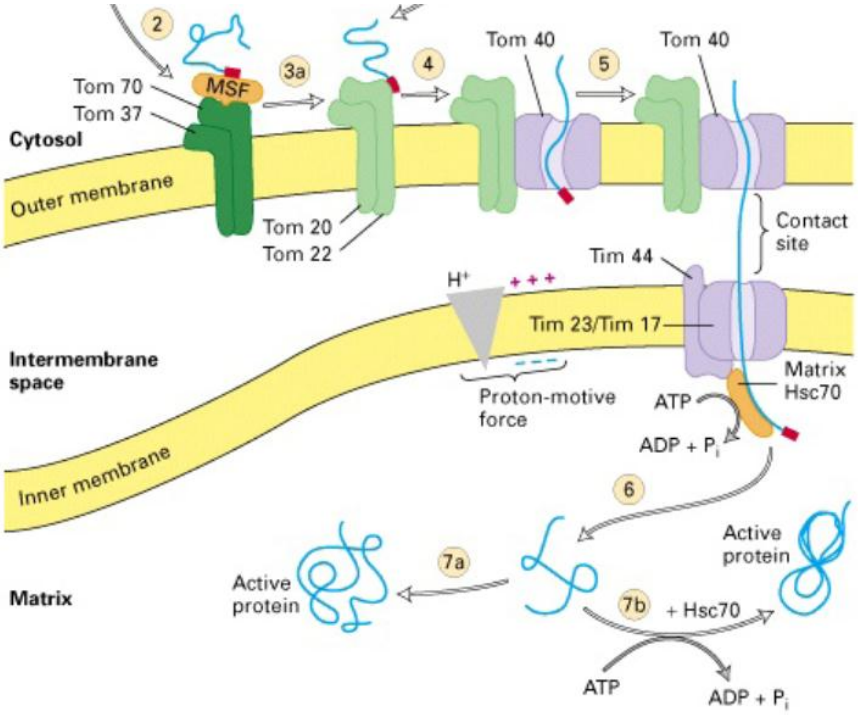
Translocation through TIM: Linked to TOM, Tom40 cooperates with the Tim17/23 complex, there is no stop of the peptide in the intermembrane space but a direct translocation into the matrix. The

membrane potential (positive on the inner side, negative in the matrix) is used to drive the translocation, together with hydrolysis of ATP.



TIM23 complex	
Tim50	Membrane sector
Tim44	Import motor
Tim23	Membrane sector
Tim21	Membrane sector
Tim17	Membrane sector
Tim16/Pam16	Import motor
Tim14/Pam18	Import motor
mtHsp70	Import motor
Mge1	Import motor





There are two models on how the ATP-driven import motor could work:

1. Thermal ratchet: Close binding of mtHSP70 to the incoming peptide prevents backsliding (more evidence)
2. Cross-bridge ratchet: Conformational change of mtHSP70 brings in the peptide with power strokes

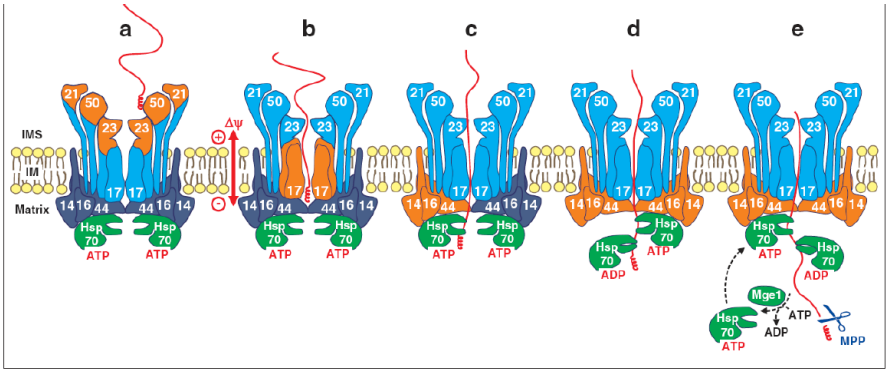


Figure 4

Stages of TIM23-mediated protein translocation. (*a*) Preproteins are directed by the intermembrane space (IMS)-exposed domains of Tim21, Tim50, and Tim23 (21, 50, and 23, respectively) to the protein-conducting channel of the TIM23 translocase (step 1). (*b*) Upon gating of the TIM23 channel, the presequences reach the matrix in an electrical membrane potential (Δψ)-dependent step (step 2), (*c*) where they interact with the ATP-bound state of mtHsp70 (step 3). (*d*) Upon ATP hydrolysis, mtHsp70 tightly binds the incoming chain and is released from the TIM23 complex (step 4). (*e*) Repeated mtHsp70-binding cycles drive complete translocation of the precursor protein. mtHsp70 is finally released upon exchange of the bound ADP for ATP, a reaction stimulated by Mge1 (step 5). In most cases, the presequence is proteolytically removed by the mitochondrial-processing peptidase (MPP). Abbreviations: 14, Tim14; 16, Tim16; and 44, Tim44.

Variety of Complexes:

- TIM 23: Transport to the matrix, membrane protein insertion into inner membrane, intermembrane space (in conjunction with TOM)
- TIM 22: Insertion of multipass membrane proteins into the inner membrane (in conjunction with TOM)
- OXA: Transport of proteins from the matrix, membrane protein insertion into inner membrane, translocation to intermembrane space

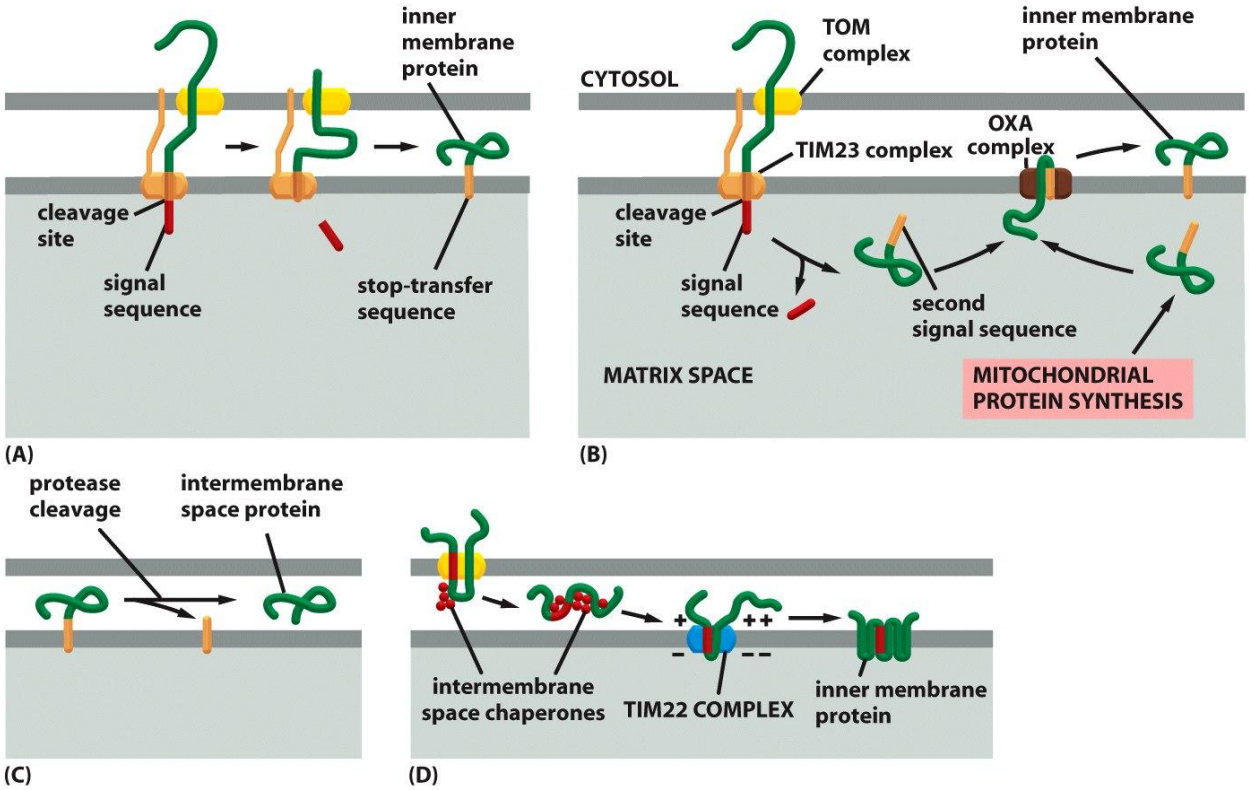


Figure 12-28 Molecular Biology of the Cell 5/e (© Garland Science 2008)

A. Two signals, firs one MTS, second one hydrophobic signal, just after MTS
B. Initial translocation, then unmasking of second signal sequence after first one gets cut away
C. Soluble intermembrane proteins are cut from their signal sequence
D. Multi-spanning membrane proteins are first translocated by TOM, chaperones then lead the peptide to TIM 22 complex

Summary:

- Protein import into mitochondria occurs post-translationally across two membranes

- Transport of nuclear encoded preproteins into the mitochondria is mediated by TOM and TIM translocases in the outer and inner mitochondrial membranes
- Preproteins that are destined for the matrix carry a cleavable amino-terminal matrix targeting signal (MTS), and they are transported across TOM and TIM23 in a coordinated fashion and in an unfolded state
- An electrical membrane potential is necessary for the translocation of the MTS; import of the rest of the preproteins is mediated by an ATP-driven import motor
- The import motor consists of the peripheral inner membrane protein Tim44, the mitochondrial chaperone mtHsp70 and its co-factor Mge1
- MtHsp70 binds a precursor that emerges from the import channel, hydrolyses ATP and is released from Tim44; The preprotein chain, in complex with mtHsp70, moves further inward. Retrograde movement is blocked by bound mtHsp70. Another mtHsp70 is binding and the precursor is fully imported by several cycles of mtHsp70 binding. Exchange of ADP for ATP leads to polypeptide chain release, a reaction stimulated by Mge1.

Protein Transport into Chloroplasts

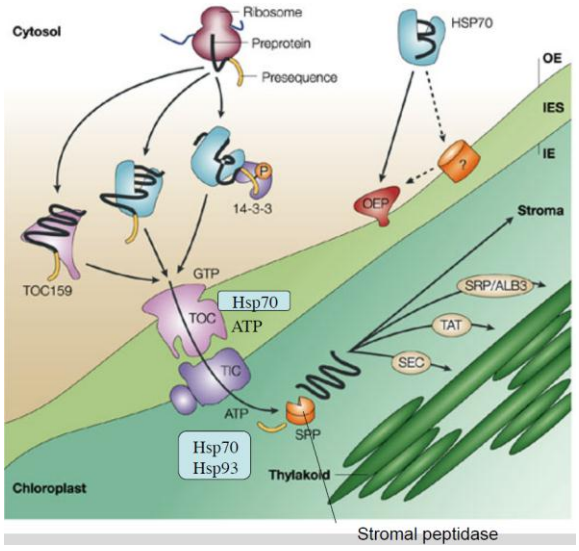
Occurs post-translational with unfolded proteins, is energy-dependent (ATP and GTP) and requires a N-terminal transit peptide. There is no electrochemical gradient involved. Proteins involved are:

- **TOC:** Translocon at the Outer Membrane of Chloroplasts
- **TIC:** Translocon at the Inner Membrane of Chloroplasts
- Hsp molecules

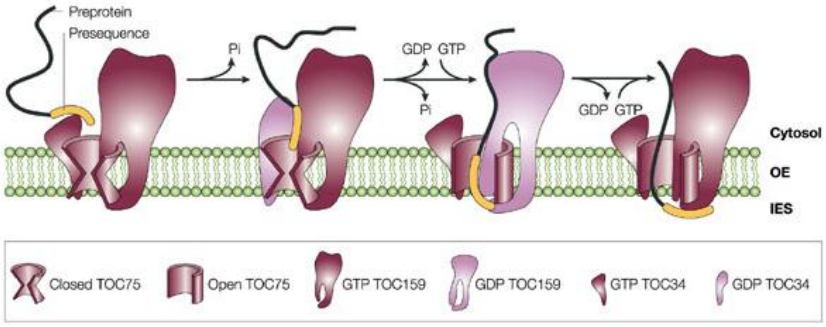
Translocation:

a) HSP70 keeps preproteins with an amino-terminal chloroplast targeting signal (or transit sequence) unfolded in the cytoplasm; preproteins bind to TOC receptor subunits in a GTP-dependent manner

b) Some preproteins can form a cytosolic guidance complex (composed of HSP70 and a 14-3-3 dimer); 14-3-3 recognize phosphorylated proteins and receptors then recognize the 14-3-3 bound to phosphorylated proteins



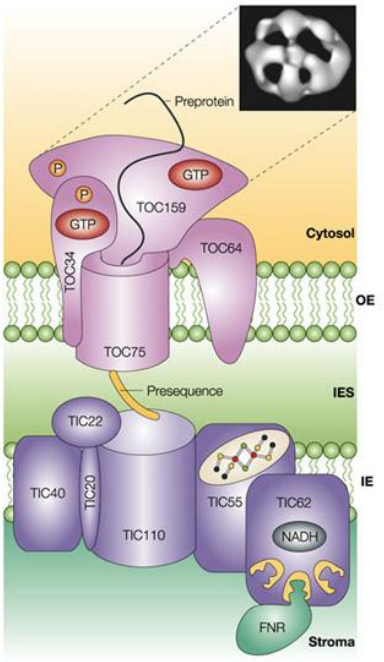
c) Not really important, a soluble TOC159 receptor might shuttle between the cytosol and the TOC complex bringing preproteins to the organelle surface



Motor Hypothesis for the Action of TOC159: First, a preproteins binds to TOC34-GTP and is transferred to TOC159-GTP upon GTP hydrolysis. TOC34 releases GDP and rebinds GTP. Upon hydrolysis, TOC159 undergoes a conformational change and pushes the preproteins through the TOC75 import channel. TOC159 finally releases GDP and rebinds GTP.

TOC Core Complex:

- GTP-dependent receptor TOC34
- Import channel TOC75
- GTP-driven motor protein TOC159
- TOC64 might function as a docking site for HSP70



TIC Core Complex:

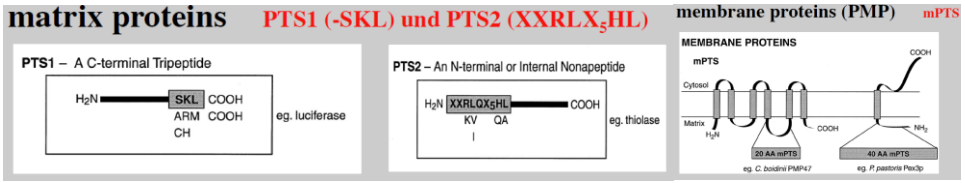
- TIC110 might form the translocation channel
- TIC20 might also be part of this channel
- TIC40 might function as a chaperone-recruiting site
- TIC22 is supposed to coordinate TIC and TOC function

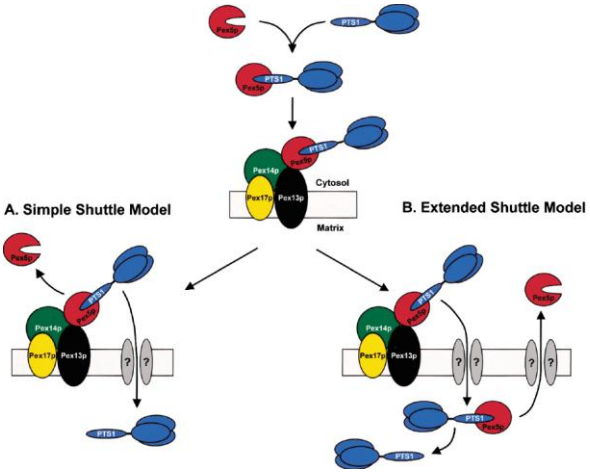
Summary:

- Protein import into chloroplasts occurs post-translationally across two membranes; the molecular machines are called TOC and TIC, and are located in the outer and inner membranes
- TOC consists of three functionally distinct subunits; a GTP-dependent receptor (TOC34), a barrel-type translocation channel (TOC75) and a GTP-dependent molecular motor (TOC159)
- Translocation across the TIC complex is driven by ATP hydrolysis, which is probably required for the action of molecular chaperones

Protein Transport into Peroxisomes (NO NOTES!)

Occurs post-translationally with already folded proteins and oligomers. The import is mediated by peroxins and proteins harbor a peroxisomal targeting signal (PTS).

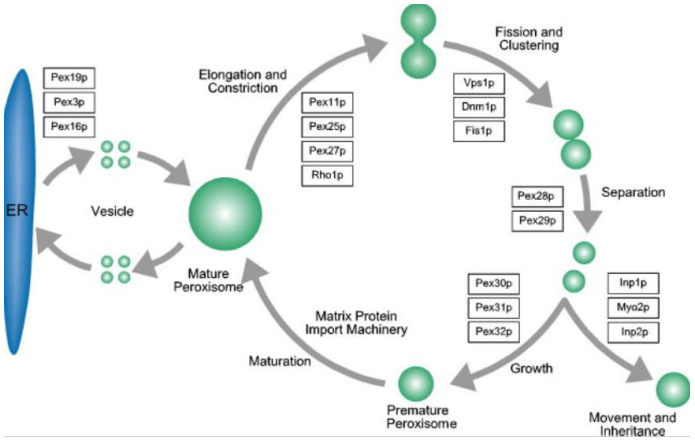




Pex5p (PTS1R) is responsible for the transport of PTS1 proteins to peroxisomes (Pex7 for PTS2). In the simple shuttle model, Pex5p only helps PTS1 to get through the membrane. In the extended shuttle model, Pex5p is translocated across the membrane as well.

Peroxisomal Biogenesis: Some lethal human diseases are connected to failures in peroxisomal synthesis, e.g. Zellweger syndrome or neonatal adrenoleukodystrophy.

Precursor peroxisomal vesicles bud from the ER (Pex19p, Pex3p and Pex16p are required for peroxisomal membrane formation, facilitate insertion of peroxisomal membrane proteins). New peroxisomes are believed to arise primarily b duplication of pre-existing peroxisomes.

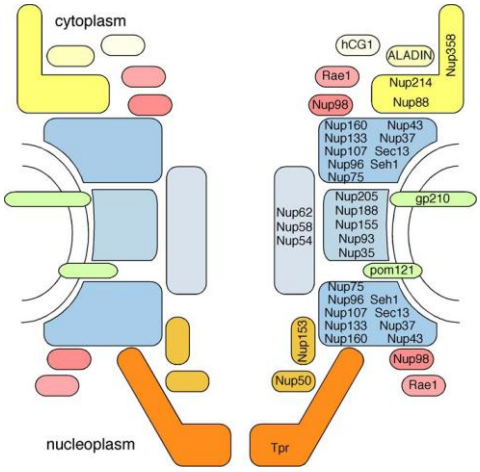


Summary:

- Post-translationally, peroxins facilitate import of folded or already oligomerized proteins
- PTS1 and PTS2, recognized by either Pex5p or Pex7p
- Hypothesis of shuttling receptors: PTS receptors cycle between cytosol and peroxisomal membrane. Extended shuttle hypothesis suggests that receptors do not release cargo after docking but reach peroxisomal lumen together with cargo, then dissociation takes place, the receptors shuttled back to cytoplasm
- Protein import into peroxisome and targeting and insertion of membrane proteins are performed by distinct machineries

Transport into and out of the Nucleus

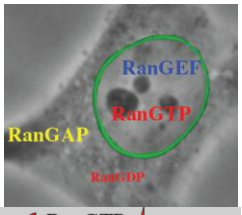
1. Introduction

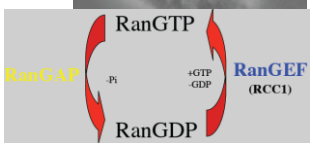


Nucleo-cytoplasmic communication is crucial for many different processes (apoptosis, viral replication, tumor suppression etc.). Nuclear pore complexes (NPCs) are the gateways. They have an octagonal symmetry, 60-100 MDa, huge complexes, gigantic macromolecule, FG repeats, passive diffusion can take place till 40 kDa, larger proteins have to be transported actively.

Cargo for nuclear transport includes: Ribosomal proteins, RNA binding proteins, snRNPs, tRNAs, mRNAs etc.

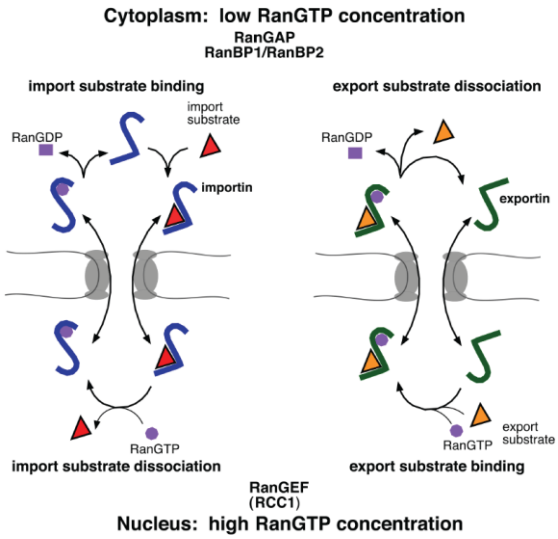
Shuttling is achieved by nuclear transport receptors, so called importins (recognize import substrates) and exportins (bind export substrates in the nucleus). Both interact with the FG repeats. An important player in the uptake and release of cargo is the small ras-like GTPase Ran. Ran exists in a GDP and in a GTP bound form (very slow transition). RanGAP is is important for the activation of RanGTP (hydrolysis) whereas RanGEF exchanges GDP with GTP. RanGAP is exclusively found in the cytoplasm whereas RanGEF is only found in the nucleus, therefore, the RanGTP form is available in much higher concentrations in the nucleus.





Importins: Have a high affinity for RanGTP and RanGTP binding leads to dissociation of the import substrate. **Exportins:** Have a low affinity for RanGTP, substrate and RanGTP bind cooperatively. Both shuttle continuously between the nucleus and the cytoplasm. They interact with the NPC and drive their own import and export. RanGTP binding controls transport of substrate association.

Example: Importin α and β bind NLS, are transported into the nucleus, then importin β dissociates which lowers α importins affinity for NLS -> release into nucleus.



There are many different im- and exportins and many corresponding nuclear localization and nuclear export signals (NLS and NES); different import signals are recognized by different transport receptors, which can be looked at in in *vitro* nuclear import assays. Here, cells are semi-permeabilized and incubated with fluorescent import substrate and treated with different factors. One round of transport does not need energy, for recycling, GTP hydrolysis is needed. Transport through NPCs is very fast, about 100 molecules/NPC/second (up to 1000); as there are about 3000-4000 NPC/nucleus, this gives us about 300'000-400'000 molecules/nucleus/second.

Transport receptor	function/substrate
Importin β	Import of ribosomal proteins, HIV-Rev, Tat
Importin β/α	Import of NLS proteins
Importin β/Snurportin	Import of U snRNPs with m_3G-cap
Importin β/Importin 7	Import of linker histone H1
Transportin 1	Import of hnRNP proteins, Import of ribosomal proteins, Import of core histones
Transportin SR	Import of SR proteins (splicing factor)
Importin 5	Import of ribosomal proteins, core histones
Importin 7	Import of ribosomal proteins, core histones
Importin 9	Import of core histones H2A and H2B
CRM1	Export of NES proteins (L-X_{2-3}-(L,I,M,F,M)-X_{2-3}-L-X-(L,I,V)
CAS	Export of Importin α
Exportin-t	Export of tRNAs
Exportin 4	Export of eIF-5A
Exportin 5	Export of small structured RNAs

Models of Translocation:

- Virtual Gate: Unstructured noncohesive FG-repeat-containing filaments fill the NPC which entropically excludes cargo not bound by receptors.
- Selective Phase Model: Saturated hydrogel within NPC, all FG domains engage in a maximum number of interactions, highly ordered mesh, receptors are thought to dissolve the FG mesh and catalyze the entry of cargo.

2. The Small GTPase Ran

Functions: Nucleo-cytoplasmic transport, mitotic spindle assembly, nuclear envelope reassembly after mitosis.

Family	Example	Function
Small GTPases	Rab	Vesicle targeting and fusion
	Ran	Nuclear transport
	Ras	Signal transduction
	Rho	Reg. of actin cytoskeleton
	Sar	Vesicle formation
Trimeric G proteins	$G_i\alpha$	Signal transduction
Translation factors	EF-G	Protein synthesis
	EF-Tu	
Translocation GTPases	SRP54	Protein transport into the ER
	SRP receptor	
	Dynamin	Endocytosis

2.1 Molecular Insights into GTPases

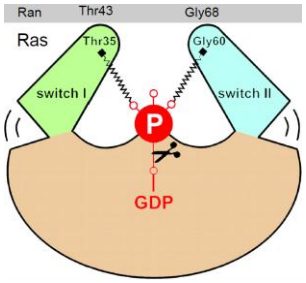
The molecular switch is done by GAP and GEF (as mentioned above). The G domains are conserved and consist of six-stranded parallel β-sheets sandwiched between six α-helices. GTP binds to a groove formed largely by loops and hydrolysis and phosphate dissociation cause major changes in the conformation of the switch loops. There are five conserved motifs in this GTP-binding groove:

- G1: α- and β-phosphate binding, also called P-loop

- G2: γ-phosphate binding (the one that is taken off from GTP), Mg^{2+}, switch I, effector loop
- G3: γ-phosphate binding, catalytically active Q residue, switch II
- G4 and G5: Guanine recognition

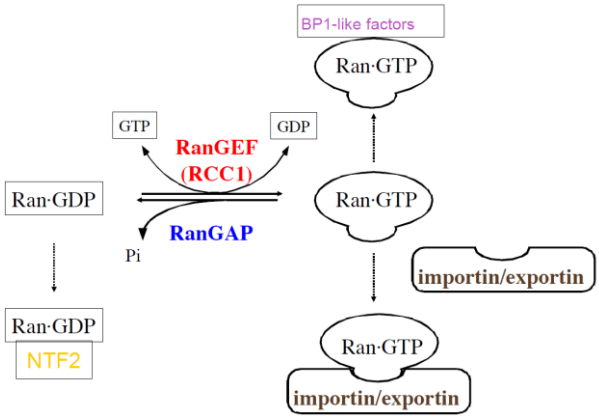
Protein	Seqence	G1 (P-loop) GXXXXGK(S/T)	G2 (switch I) XTX	G3 (switch II) DXXG	G4 (N/T)(K/Q)D	G5 (T/G/C)(C/S)A
H-Ras	P01112	GAGGVGKS	YDPTIED	ILDTAGQE	VGNKCD	YIETSAK
RhoA	P06749	GDGACGKT	YVPTVFE	LWDTAGQE	VGNKKD	YMECSAK
Cdc42	P25763	GDGAVGKT	YVPTVFD	LFDTAGQE	VGTQID	YVECSAL
Rac1	P15154	GDGAVGKS	YIPTVFD	LWDTAGQE	VGTKLD	YLECSAL
$G_{i\alpha}$	P04898	GAGESGKS	RVKTTGI	LFDVGGQR	FLNKKD	THFTCAT
EF-Tu	P20001	GHVDHGKT	RGITINT	HVDGPGHA	FLNKCD	IVRGSAL
Arf-1A	P32889	GLGAAGKS	TIPTIGF	VWDVGGQD	FANKQD	IQATCAT

GTP hydrolysis and phosphate dissociation cause major changes in the conformations of the switch loops (switch I and switch II). The γ-phosphate determines the structure of the flexible switch regions.



Loaded Spring Mechanism: When GTP-bound, switch I and II are bound to the γ-phosphate via the main chain NH groups of the invariant Thr and Gly residues. Release of γ-phosphate after GTP hydrolysis allows the switch regions to relax into a different conformation.

2.2 How Ran Works: From Structure to Function



Factor	Interaction RanGTP/ GDP	Localization	Function
Ran	n.a.	Nuc/Cyt	Directionality of nuclear transport
Transport receptors superfamily	RanGTP	Nuc/ NPC/ Cyt	Translocation of macromolecules through the NPC
RanGAP	RanGTP	Cyt/ NPC	Stimulation of GTP hydrolysis on Ran
RanGEF (RCC1)	Nucleotide free form of Ran	Nuc	Nucleotide exchange
RanBP- family: RanBP1 RanBP2 RanBP3	RanGTP	 Cyt NPC (cyt. side) Nuc	 Export complex dissociation Export complex dissociation Formation of spec. export complexes
NTF2	RanGDP	Cyt/ NPC	Nuclear import of Ran

Ran belongs to the superfamily of Ras-related proteins. It is the most abundant family member (0.4% of cellular protein). Ran stands for Ras-related nuclear protein, has no membrane anchor, is soluble and 25

kDa. It binds guanine nucleotides with high affinity and specificity (pico- to nanomolar range), has a low GTP dissociation rate and the intrinsic nucleotide hydrolysis rate is very slow ($1.8*10^{-5}s^{-1}$). The 3D structure of Ran is very similar to the small GTPase Ras. One can find a G domain with 6 stranded β-sheets surrounded by 5 α-helices (α,β protein).

With GDP bound, the switch I and II regions move away from the nucleotide, adopting an open conformation, whereas the carboxy-terminal tail (switch III) packs as an α-helix against the G domain. When GTP is bound, the switch regions I and II move closer to the nucleotide to interact with its γ-phosphate, whereas the carboxy-terminal tail is flipped away from the G domain. The special feature of Ran is the C-terminal helix (switch III), extending from the globular part (this helix is switched in the GDP-bound form).

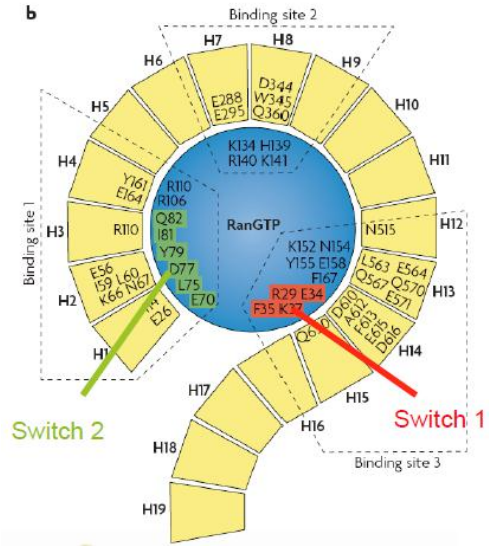
-> So, Ran GDP has the two switch regions in a more open conformation than observed in the RanGTP structures. In RanGDP, the carboxy-terminal tail is packed against the core G domain of Ran. This tail is displaced from the G-domain in the GTP-bound form.

Summary:

- Ran is a Ras family member (with a conserved G domain)
- Contains a C-terminal helix that responds to the nucleotide bound state
- Binds GDP and GTP with high affinity (nM to pM)
- Low intrinsic rate of GTP hydrolysis ($t_{1/2}$: hours)

2.3 Effectors and Regulators

Effectors of small GTPases interact more tightly with the GTP-bound form. For Ran, the effectors are transport receptors (Imp and Exp) and RanBP1 family members).



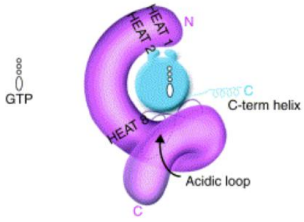
Importin-β:

19 HEAT repeats (one repeat is about 40 amino acids, made up of two anti-parallel α-helices), forming a superhelical structure in which Ran is embedded. For binding importin-β, only the switch I and II regions are used, whereas the switch III region (carboxy-

terminal tail of Ran) is not involved in these interactions and stays flexible and solvent exposed.

Binding of RanGTP to transport receptors results in the inhibition of the enzymatic activity: intrinsic and RanGAP induced GTP hydrolysis and RanGEF stimulated nucleotide exchange inhibited. This is because Ran is very tightly bound and not accessible to GAP/GEF.

There is also some overlap between the RanGTP and the IBB-domain (N-terminus of importin α) binding sites on importin β, so RanGTP and importin α cannot bind at the same time (sterical hindrance). RanGTP binding induces a conformational change in importin-β, which increases the helicoidal pitch of importin β, thereby releasing the IBB domain by an allosteric mechanism.



Example transportin (Ran/importin): Best known NLS recognized by transportin is the M9 sequence of the mRNA-binding protein hnRNPA1. This NLS is 38 amino acids long and has an overall positive charge. It binds as a linear extended peptide to the C-terminal of transportin. Recognition of the NLS occurs at the inner concave of transportin. The overall shape of the superhelix is different from importin β. The two arches have a distinct kink, so that instead of the tight snaillike conformation observed in the importin β:IBB complex, transportin has a rather open Z-like structure. The acidic loop at HEAT8 interacts with bound RanGTP and interferes with cargo binding in the RanGTP-bound conformation as it protrudes into the cargo binding site (acidic loop appears not to contribute to substrate binding!).

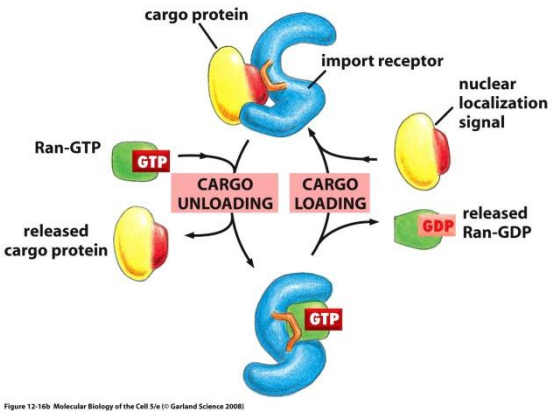
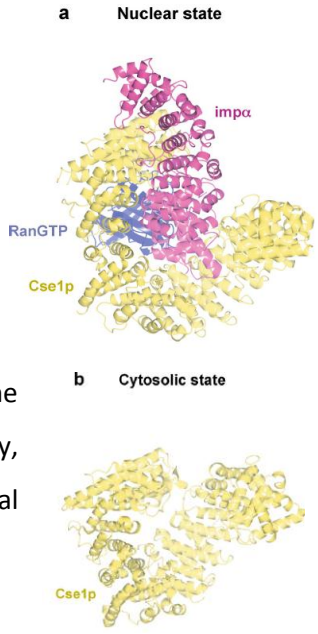


Figure 12-16b Molecular Biology of the Cell 5/e (© Garland Science 2008)

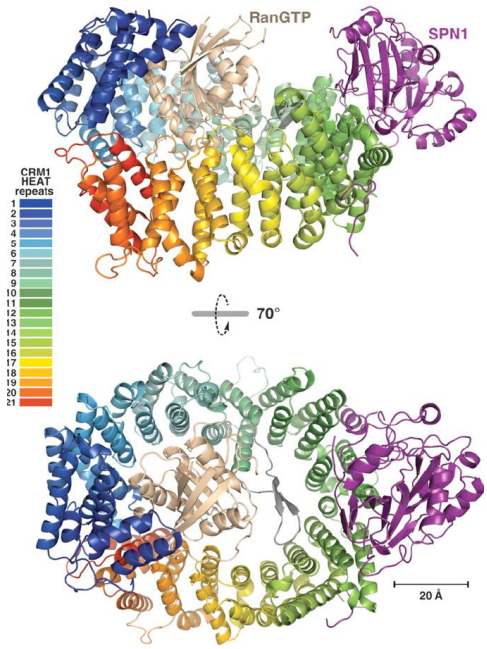


Example CAS (Cse1p, Ran/exportin): Cse1 is composed of 19 consecutive HEAT repeats. RanGTP binds at the N-terminal arch of the superhelix. Importin α contacts both the N-terminal and C-terminal arches of Cse1. Importin α also forms extensive interactions with RanGTP which partly explain the mutual need of Ran and cargo in the formation of the export complex.

The unbound form of Cse1 has a more closed conformation as compared with the bound state. The N-terminal and C-terminal arches of Cse1 interact intramolecularly, with HEAT repeats 1-3 contacting repeats 14-16. The displacement of the N-terminal

arch away from the C-terminal one is required for binding of the export cargo and this conformation is likely induced and/or stabilized by RanGTP.

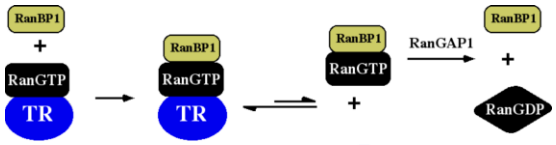
Example CRM1 (Xpo1p, Ran/exportin): Bent into a toroid structure, HEAT21 touches 2B and 5 and Ran is enclosed into the toroid and stabilizes ring closure by extensive contacts. NES cargo is not enveloped by CRM1 but binds outside. CRM1 contacts both switch regions of Ran (I and II).



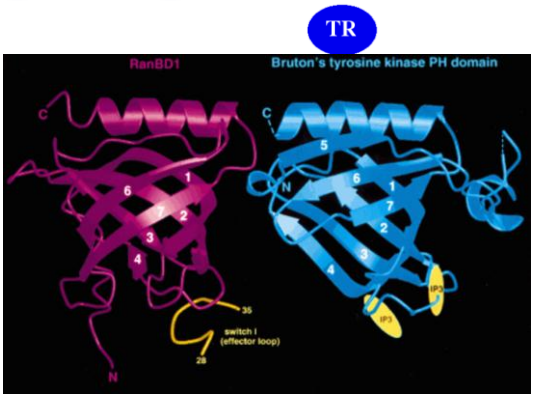
Summary:

- N-terminal helical repeat motifs of import receptors interact with switch I and II regions of Ran
- Interaction of RanGTP sterically interferes with binding of import cargo molecules (release of import cargo upon RanGTP binding)
- Extensive interactions between exportin, RanGTP and cargo ensure cooperative complex formation
- Binding of RanGTP to transport receptors results in the inhibition of all known enzymatic activities, such as the intrinsic and RanGAP induced GTP hydrolysis and RanGEF stimulated nubleatide exchange (this results from overlap of binding sites)

RanBP1 Family:



Binding of importins and exportins to RanGTP inhibits RanGAP-induced GTP hydrolysis by Ran. Cytoplasmic RanBP1-family members help to dissociate RanGTP from transport receptors (TR), they stimulate GTPase activity by about an order of magnitude.

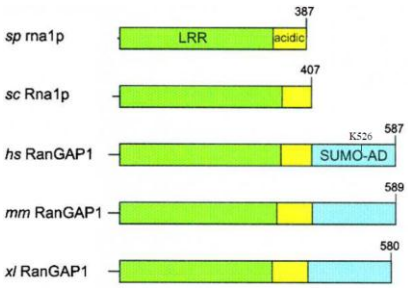


RanBP1: Cytoplasmic, 23 kDa, containing a RBD (Ran binding domain).

RanBP2: Part of the nucleoporin complex, 4 RBDs, 358 kDa .

The RanBD has the β-barrel topology of a pleckstrin-homology (PH) domain. The acidic C-terminus of Ran is bound to the basic patch on the RanBD. RanBP1 removes RanGTP from importin β-like receptors.

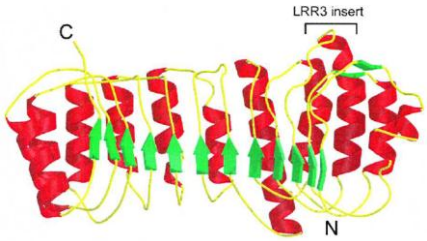
Summary:

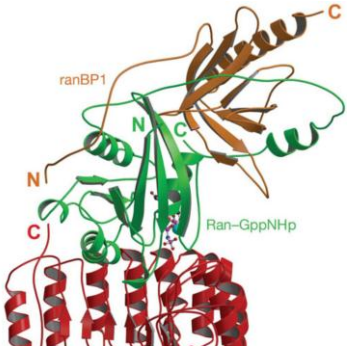


- RanBP1 and RanBP2 contain RBDs, PH domain fold
- Tight molecular embracement between RanGTP and RBD
- RBD stimulates dissociation of RanGTP from transport receptors
- Stimulation of GTP hydrolysis on Ran by one order of magnitude

GAP:

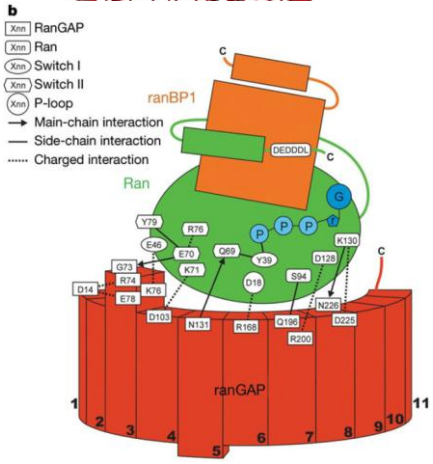
Ran hydrolyzes GTP only slowly, the half time of RanGTP is several hours. GAPs increase the rate of GTP hydrolysis by five orders of magnitude (from $1.8*10^{-5}s^{-1}$ to $2.1*10s^{-1}$ at 25°). Hydrolyzation is achieved by transfer of the γ-phosphate to water and involves a conserved Gln residue in small GTPases (Q61 Ras, Q69 Ran). Q-residue is required for stabilization of the transition state. GAPs are required for GTPase to reach transition state conformation (either by changing conformation of GTPase or by directly providing residues required for the transition state (Arg in RasGAP)). In yeast, the RanGAP Rna1p is a cytoplasmic protein whereas the RanGAP1 in metazoans is attached to the NPC by SUMO (bound to RanBP2). RanGAPs contain N-terminal 11 LRRs (leucine-rich repeats) which gives them a crescent-like structure (β-α hairpin repeat unit). GAPs have possible SUMO-attachment sites (important for localization).





The RanGAP/RanGTP/RanBP1 complex:

- RanGTP interacts with 7 of 11 LRRs
- LRR3 loop interacts with and stabilizes switch II
- RanGAP helps to orient the catalytic Q residue
- Basic machinery for fast GTP hydrolysis is provided exclusively by Ran, correct positioning of the catalytic glutamine (Q69) is essential for catalysis
- Gln69 is locked in a non-reactive conformation in Ran, only in the GAP complex it is in the active position
- Tyr39 of Ran forms hydrogen bonds with both the γ-phosphate oxygen and the Gln69 side chain of Ran
- RanGAP stimulates GTP hydrolysis without providing an Arg finger
- Ran Q69L mutants bind GTP but are GTPase deficient

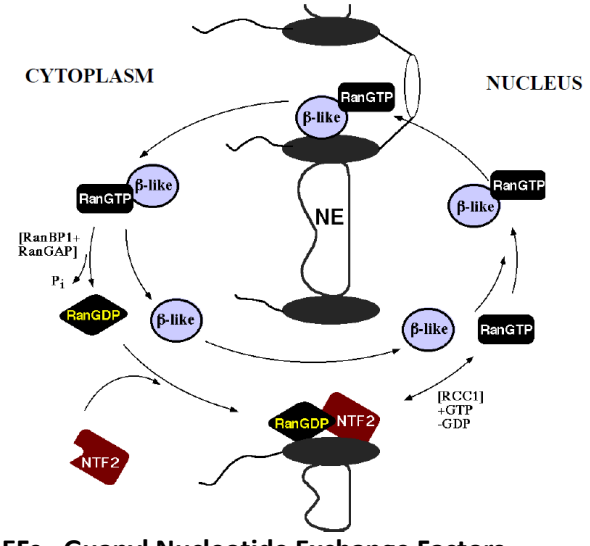


Summary:

- Ran is resistant to RanGAP-mediated GTP hydrolysis when bound to importins/exportins, due to inaccessibility of the switch region in such a complex
- The importin-RanGTP complex is recycled to the cytoplasm where GTP hydrolysis on bound Ran is stimulated by the concerted action of RanGAP and RanBP1 or RanBP2
- The RanBDs can access RanGTP even when it is in a complex with importins, as they recognize the carboxy-terminal helix of Ran, which is exposed to solvent in RanGTP-importin complexes
- Ran becomes susceptible to RanGAP-mediated GTP hydrolysis when bound to the RanBD of RanBP1 or RanBP2
- RanGAP stimulates GTP hydrolysis on Ran by 5 orders of magnitude

NTF2

Nuclear import of RanGDP is mediated by NTF2 (back and forth shuttling). NTF2 is a 15 kDa homodimer protein and binds both two RanGDP and FxFG repeat-containing nucleoporins. NTF2 binds switch II region of RanGDP.



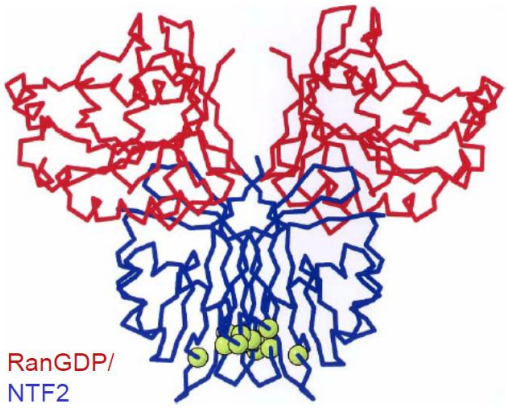


Fig. 4. Crystal structure of NTF2-RanGDP complex in which NTF2 is blue and Ran is red. The hydrophobic residues near Trp7 on NTF2 implicated in its interaction with FxFG repeats are shown as yellow spheres.

GEFs - Guanyl Nucleotide Exchange Factors

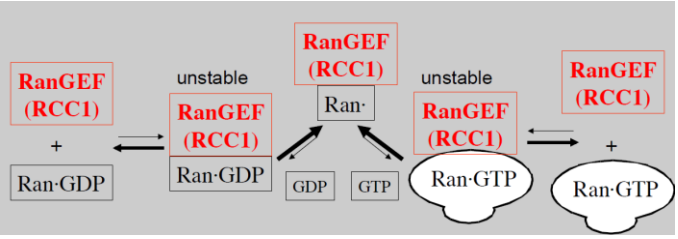
- Small GTPases bind their nucleotides stably, low dissociation rate leads to an almost irreversible binding of the guanine nucleotide
    - E.g. Ran: Ran binds GTP with high affinity (> 10^9 M^{-1})
    - Half-life of RanGDP and RanGTP complexes being in the range of several hours

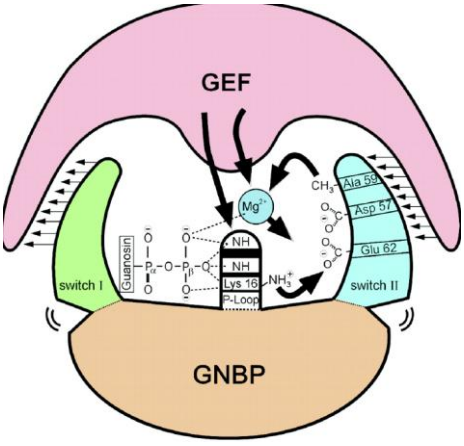
- Ran has a higher dissociation rate of GTP ($1.1*10^{-4}s^{-1}$) as compared to that of GDP ($1.5*10^{-5}s^{-1}$ at 25°) which indicates a 7-fold higher affinity of Ran for GDP

- GEFs accelerate GDP/GTP exchange by several orders
- Two different principles:
  - Stabilization of the nucleotide-free form (GEF merely acts as a catalyst to increase the rate at which equilibrium between the GDP- and GTP-bound forms of the proteins is reached
  - Disruption of the GDP binding site

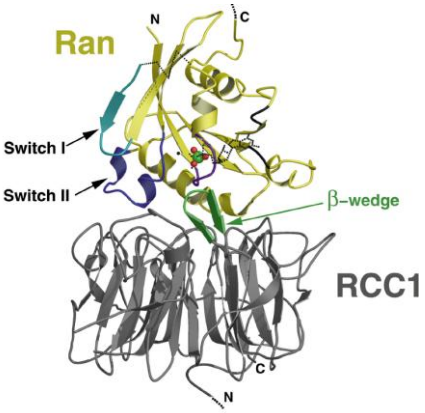
GEFs decrease the affinity for nucleotide by stabilizing the nucleotide free form:

1. A ternary G protein-GEF-nucleotide complex of low nucleotide affinity is formed
2. The mutual competition of GEF and nucleotide leads to the formation of a stable binary G protein-GEF complex
3. This complex reverts back to the binary G protein-nucleotide complex in the presence of nucleotide





The most important contribution to high-affinity binding of the guanine nucleotide to Ras-like GTPases is due to interaction of the phosphates with the P-loop and the Mg-ion. Upon GEF binding, Mg^{2+} is pushed out of its position by elements of the GTPase itself, i.e. the Ala59 (in Ras and Rac) and from residues of GEF. Residues of the P-loop are disturbed, and its lysine is reoriented toward invariant carboxylates from the switch II region, the invariant Asp (Asp57 in ras) or a highly conserved Glu (Glu62). In a push-and-pull mechanism, switch I is pushed out of its normal position, whereas switch II is pulled toward the nucleotide-binding site.



RCC1 (RanGEF) was initially identified as a cell cycle regulator. It is a nuclear protein of 45 kDa, chromatin-bound and increases guanine

nucleotide dissociation by more than five orders of magnitude. A surface exposed β-wedge on RCC1 plays the decisive role in catalysis by pushing aside the P loop needed for nucleotide binding.

Summary:

- RanGEF, RCC1 (metazoans) stimulates dissociation of Ran-bound nucleotide (GDP or GTP) and the subsequent exchange reaction by stabilizing the nucleotide-free form of Ran
- *In vivo*, Ran binds GTP which is present in high concentration in the cell
- RanGEF is restricted to the nucleus and bound to chromatin
- RanGEF is composed of β-strands that form a propeller of seven blades; all blades of the RCC1 β-propeller participate in the Ran interaction
- Key element of the interface is a protruding β-wedge, which interacts with P loop and plays the decisive in catalysis by pushing aside the P-loop

Regulation of Nucleo-Cytoplasmic Transport

Introduction

Many biological processes are controlled by transport of cargo out and into the nucleus, such as cell cycle progression and proliferation, signal transduction (MAPK, NFκB, GC receptor…), response to environmental stimuli in yeast (Pho4, Yap1…), circadian clock, organogenesis, muscle differentiation, bone morphogenesis, development of nervous, muscular, immune and vascular systems (NF-ATs).

Steady-state localization of a protein is often determined by its relative rates of nuclear import and export. Retention in the one or the other compartment plays an equally important role.

There are different ways for regulation:

1) Regulation of cargo-transport receptor complex formation (e.g. phosphorylation of cargo)
2) Regulation of cytoplasmic or nuclear anchoring/retention (also often by phosphorylation)
3) Regulation of the soluble transport machinery
4) Regulation of the NPC

Phosphorylation regulates transport by either activating or inactivating NLS (Dorsal, SV50 T-ag activation of NLS, NF-AT, Pho4p inactivation) and NES (Pho4p activated, cycling B inactivated). Also important for complex formation (dimerization of MAPK, period and timeless), degradation (NFκB, IκB, period and timeless) and binding or release from anchoring partners (e.g. from 14-3-3 proteins).

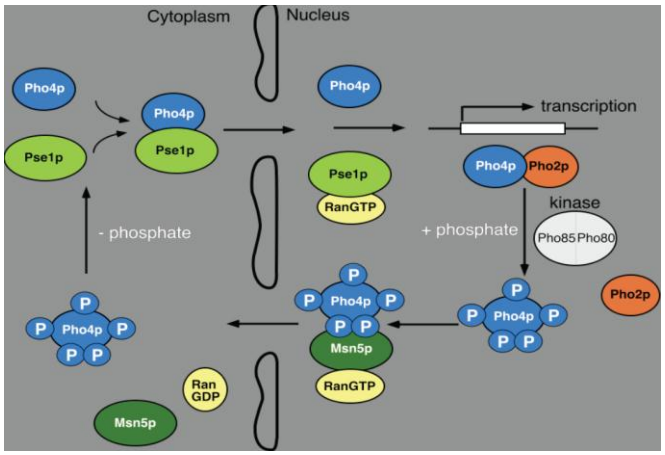
1. Regulation of Cargo-Receptor Complex Formation

1.1. Phosphorylation of Cargo

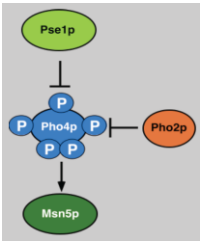
1.1.1. Pho4 (Phosphate Starvation in Yeast)

Pho4 is a transcriptional activator that is needed for the expression of genes under phosphate starvation in yeast. In low phosphate environment, Pho4p is imported into the nucleus by the importin Pse1p. Binding of Pho4p to the transcription factor Pho2p leads to activation of transcription.

If high amounts of phosphate are available, Pho4p is phosphorylated by a kinase complex (5x on five Ser residues) in the nucleus which triggers binding to the exportin Msn5p. Phosphorylated Pho4p is exported to the cytoplasm.



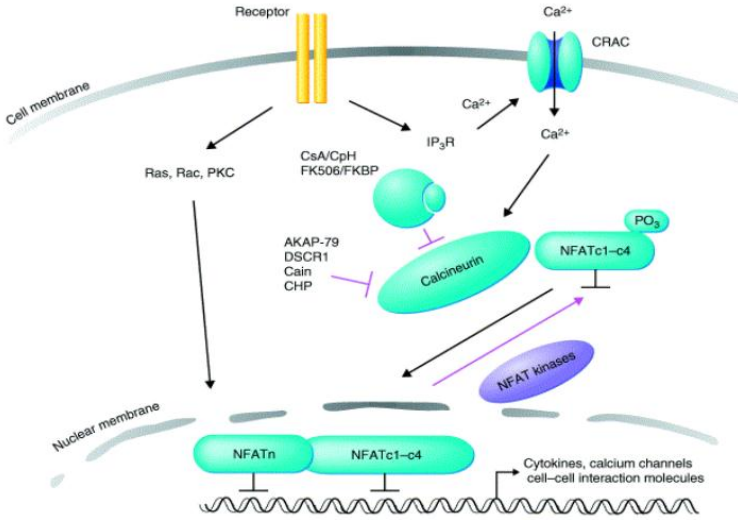
Phosphorylation inactivates Pho4 by preventing its reimport (Pse1p cannot bind phosphorylated Pho4p), dissociating from Pho2 and promoting its export by Msn5.

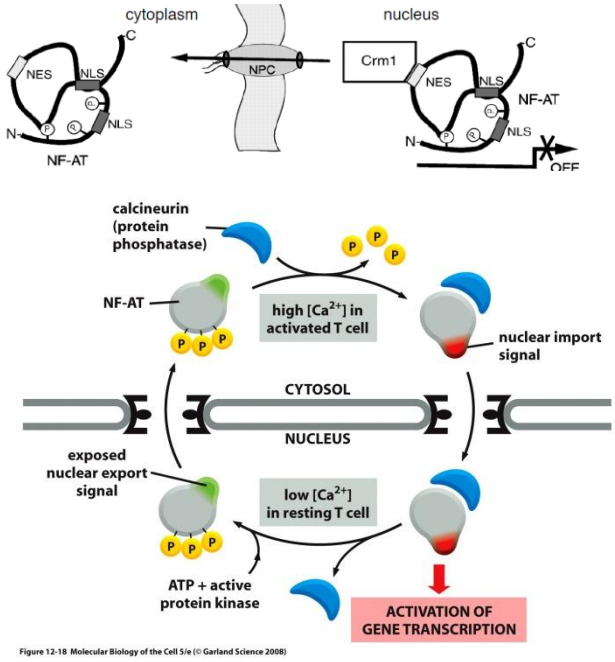


1.1.2. NF-AT (Nuclear Factor of Activated T Cells

First discovered in lymphocytes, expressed in most immune system cells. Pivotal role in transcription of cytokine genes and other genes critical for the immune response. NF-AT also regulates cell differentiation and development. The activity of NF-AT proteins is tightly regulated by the calcium/calmodulin-dependent phosphatase calcineurin (Ca^{2+} dependent phosphatase) on the level of nucleo-cytoplasmic transport.

First, NF-AT is phosphorylated and exported to the cytoplasm by virtue of a leucine-rich NES recognized by the exportin CRM1 and therefore kept out of the nucleus. Then, dephosphorylation triggers its import. Elevation of the cytoplasmic Ca^{2+} concentration causes activation of calcineurin which then dephosphorylates NF-AT and masks its NES. Nuclear import and activation of transcription.





Immunosuppressive drugs such as cyclosporine A or FK506 prevent nuclear accumulation of NF-AT by inhibiting calcineurin.

1.2. Complex Formation

Proteins can form complexes with themselves, with another protein or with an RNA.

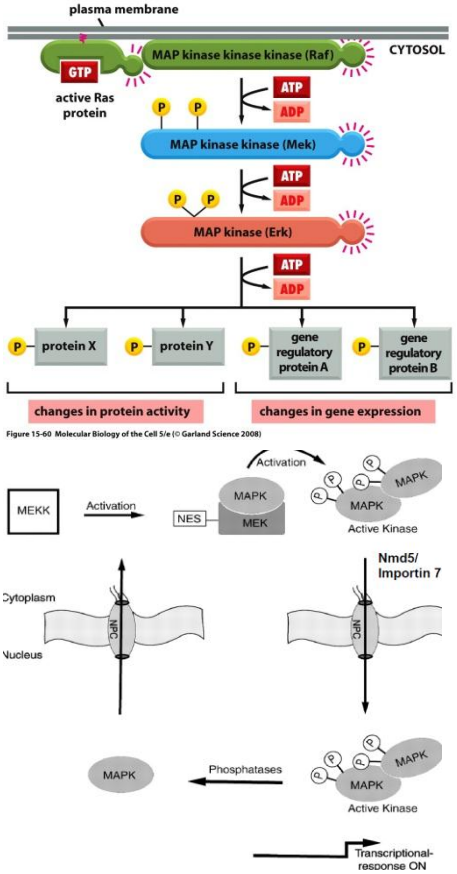
1.2.1. MAPK

The family of MAPKs is divided into two large subfamilies:

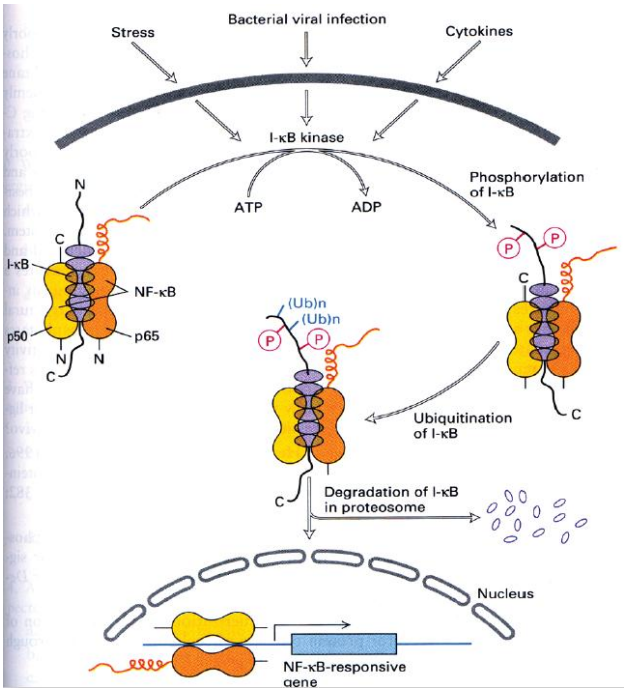
- The ERKs (extracellular signal-regulated kinase) involved in cell proliferation and differentiation
- p38 and the Jun N-terminal kinase (JNK) activated in response to cellular stress (UV, heat, high osmolarity)

MAPKs are involved in many different signaling pathways (RTK, GPCR…). MAP is activated in the cytosol by phosphorylation on conserved Thr and Tyr residues by MEK. Phosphorylated MAPK forms dimers and is then translocated into the nucleus (e.g. Hog1p and ERK2). Phosphorylation of MAPK accomplishes three tasks:

1) It turns in into an active kinase
2) It promotes dissociation from the cytoplasmic MEK
3) It promotes dimerization that is required for its nuclear localization



1.2.2. NF-κB



NF-κB is a heterodimeric transcription factor composed of the RelA (p65) and the p50 subunits. It induces transcription of a large number of genes involved in the immune, inflammatory and apoptotic responses. Activation occurs upon stimulation (stress, injuries, infections) by cytokines (interleukin 1b), TNFα, lipopolysaccharides, dsRNA (via PKR) and pro-apoptotic and necrotic stimuli such as oxygen radicals, UV or γ-radiation.

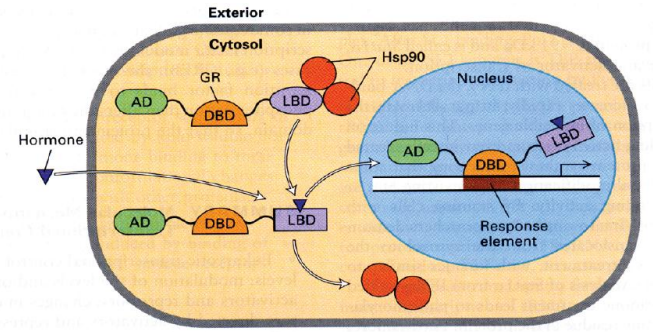
In the cytoplasm, NF-κB is kept inactive by association with IκB. Different environmental stimuli activate a kinase (IKK) that phosphorylates IκB on two serine residues and targets it for degradation. Degradation of IκB allows for the rapid translocation of NF-κB into the nucleus and the onset of a transcriptional response. Before, the NLS of NF-κB is not accessible (because in complex with

IκB). IκB inactivates NF-κB by restricting its access to the nucleus (NLS shielding to block import, NES in IκB to promote NF-κB export) and by inhibiting its DNA-binding activity.

1.2.3. Glucocoricoid Receptor (GR)

The GR belongs to the nuclear receptor superfamily and translocates from the cytoplasm to the nucleus in response to ligand binding. In unstimulated cells, GR is inactive in the cytoplasm and is complexed with heat shock proteins such as a Hsp90 or Hsp56. Binding to these proteins accomplishes three tasks:

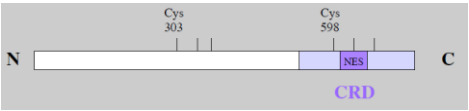
- Restricting GR localization to the cytoplasm
- Inhibiting its DNA-binding activity
- Maintaining GR in a state of high affinity for its ligand



Ligand binding to the LBD (ligand binding domain) causes dissociation of Hsp90, this allows GR to enter the nucleus and bind DNA with its DBD and therefore leading to transcriptional regulation.

1.3. Conformational Change in Cargo

1.3.1. YAP1 (Response to Oxidative Stress)

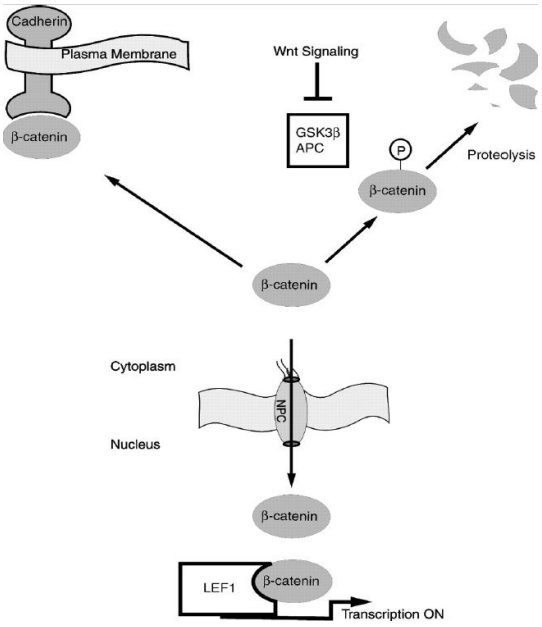


Yeast transcription factor YAP1 translocates to the nucleus in response to oxidative stress and a variety of toxic compounds. YAP1 induces transcription of antioxidative genes (e.g. thioredoxin and catalase) and genes that belong to the multidrug-resistancee family, to provide protection against toxic compounds. Cytoplasmic localization of YAP1 in unstressed cells requires a small region known as the cysteine-rich domain (CRD). The CRD acts as a redox sensor. Oxidation (disulfide bond between Cys303 and 598) leads to a masking of the NES. This inhibits the interaction between YAP1 and Crm1, leading to a rapid nuclear accumulation of YAP1 and activation of the antitoxic transcriptional response.

The activity of YAP1 is controlled primarily at the level of nuclear localization. In response to hydrogen peroxide, a disulfide bond is formed involving the oxidant receptor protein Orp1, leading to nuclear accumulation and activation of Yap1 target genes.

2. Cytoplasmic Anchoring: Example β-Catenin

- Anchoring of adaptor linking cadherins, which are surface membrane proteins involved in cell adhesion, to the actin cytoskeleton
- Role in the Wnt pathway as a transcriptional coactivator to inducea transcriptional program required for cell fate determination and dorsoventral axis formation



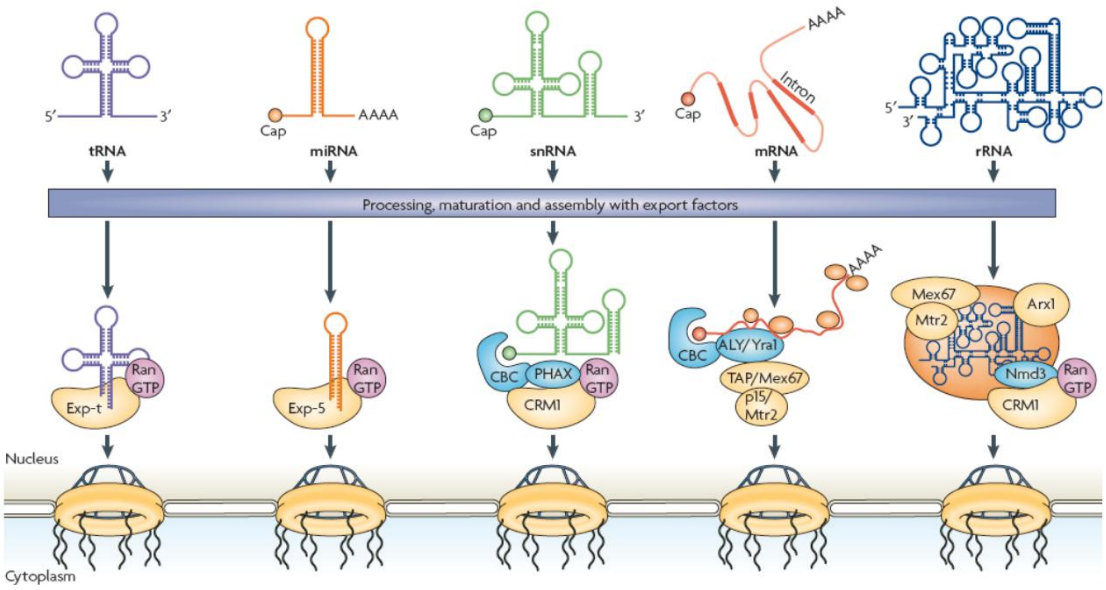
Without the Wnt signal, β-catenin is bound to cadherins and free β-catenin is phosphorylated and degraded. With the Wnt signal, β-catenin is no longer phosphorylated and degraded; any β-catenin released from retention is imported into the nucleus.

3./4. Regulation of the Soluble Transport Machinery and of the Nuclear Pore Complex

- VSV-M (matrix protein of vesicular stomatitis virus): VSV is an RNA virus that replicates in the cytoplasm. VSV-M inhibits nuclear export of almost all forms of cellular RNA and interacts with the nucleoporin Nup98 suggesting that VSV-M inhibits host cell gene expression by targeting a nucleoporin.
- Quiescent cells show a slower rate of nuclear import of large gold particles coated with nucleoplasmin, compared with proliferating cells.
- Yeast Nup53 has been shown to be phosphorylated during mitosis, the importin Pse1 can then no longer associate with the NPC.

Export of Cellular and Viral mRNAs

Introduction



Pre-mRNA particles are dynamic structures that are remodeled on their way to the cytoplasm. mRNAs are always covered with dozens of particles.

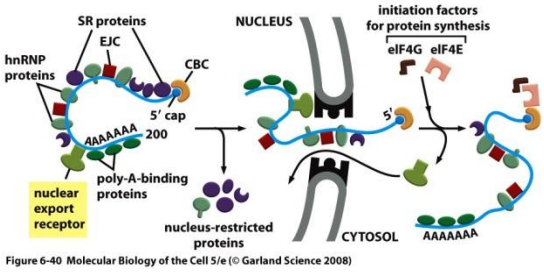
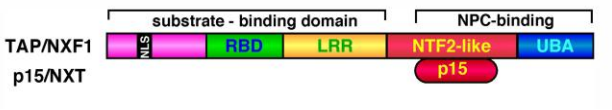


Figure 6-40 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Export Signals

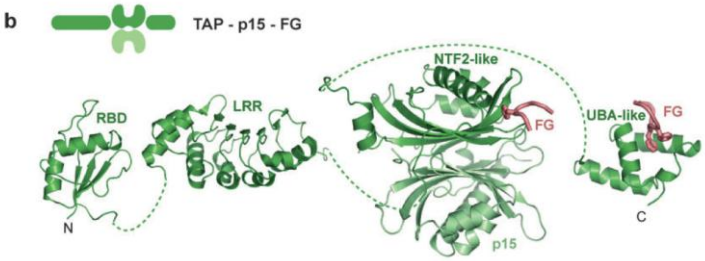
Both the 5'm^7G cap and the 3' poly(A) tail stimulate mRNA export, but neither structure appears to be essential for transport (even if the cap or the tail is removed, mRNA is still exported).

Export is mediated by a distinct type on nuclear transport receptor. Poly(A) mRNA export is mediated by the transport receptor heterodimers TAP/p15 (metazoans) and Mex67p/Mtr2p (yeast). These transport receptors constitute the NXF family (nuclear export factor). They bind directly to nucleoporins and only bind weakly do RNA.



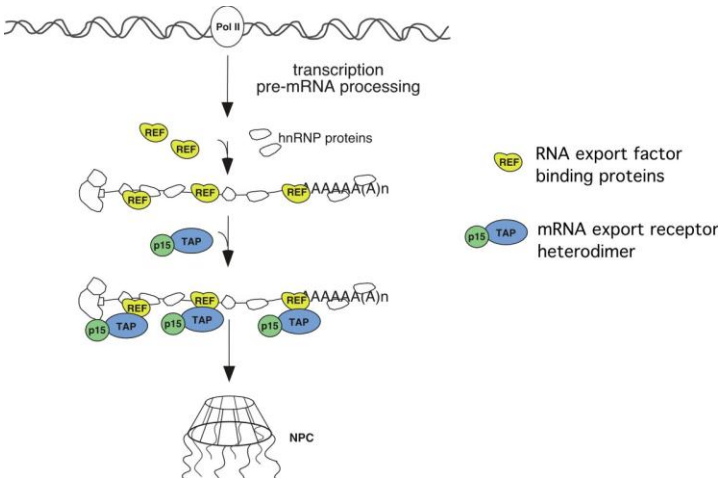
Experiments (fluorescence *in situ* hybridization) showed that at non-permissive temperature in a tsMEX67 mutant, mRNA accumulates in the nucleus. The same thing happens, if TAP is knocked-down by RNAi. In *Xenopus* oocytes it has been shown that injected radioactively labeled mRNA is expored when coinjected with TAP. Mex67p (yeast homologue of TAP) and TAP are required for poly(A) mRNA export.

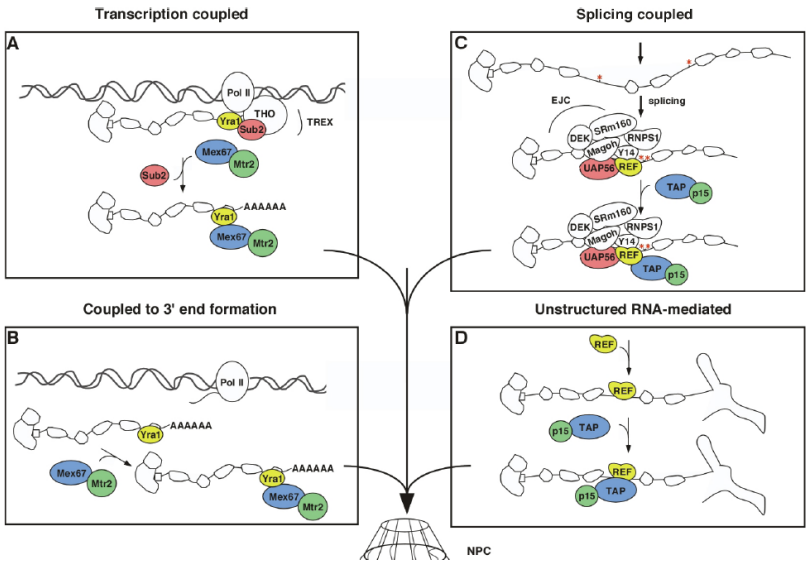
TAP and p15 resemble the NTF2 dimer, NTF2-like domains allow p15/TAP to bind to NPC but they do not interact with RanGDP. TAP harbours two FG binding domains, one in the NTF2-like domain and one in the UBA-like domain (ubiquitin associated fold).



How is the mRNA bound?

- TAP and Mex67p can be crosslinked to polyadenylated RNAs
- The affinity for cellular poly(A) mRNA is low -> need for RNA binding proteins to serve as adaptor
- Best characterized adaptor: ALY/REF alias Yra1p; ALY/REF belong to the so-called REF family (RNA export factor binding proteins) and individual REFs are not essential (redundancy!)





TAP/p15 is a Ran-independent export receptor for poly(A) mRNA. The export signal must be recognized by ALY/REF. mRNA is unstructured, therefore they recruit REFs.

There are 4 models to describe export factor recruitment to poly(A) mRNAs in the nucleus:

a) Transcription coupled
b) Coupled to 3’ end formation
c) Splicing coupled (most common principle)
d) Unstructured RNA-mediated

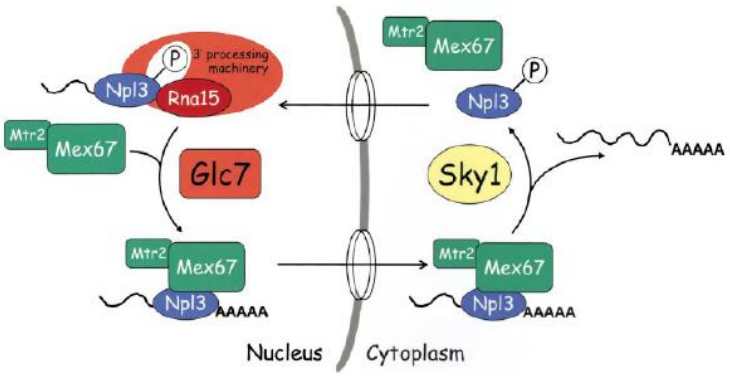
Another stimulator of mRNA export is splicing (experiments showed, that if equal amounts of pre-mRNA and corresponding i-mRNAs were mixed with U6 snRNA (control) and injected into oocyte nuclei, spliced mRNA is faster exported).

What is the reason for that? EJC (exon junction complex)! The EJC is a multiprotein complex, deposited by the spliceosome 20-24 nucleotides upstream of the mRNA exon-exon junctions in a sequence-independent manner. The EJC plays a role in post-splicing mRNA metabolism by influencing the transport, translation and stability of spliced mRNAs. The EJC contains at least five proteins: SRm160, RNPS1, REF/ALY, DEK and Y14-Mago heterodimers.

Summary:

- Only mature, fully processed mRNA is exported to the cytoplasm
- 5' and poly(A) tail are not required for export but may facilitate it (e.g. via CBC), both stabilize
- Removal of introns (splicing) is required for export of cellular mRNA (viruses have developed special tricks)
- Splicing has two consequences:
    - Release of mRNA from nuclear retention
    - Stimulates recruitment of export factors (EJC)
- mRNA export receptors (TAP/p15 human; Mex67p/Mtr2p yeast) do not belong to the exportin family and do not bind RanGTP (their NPC-binding domain resembles NTF2)
- TAP need REFs (ALY/Yra1p) to associate with cellular poly(A) mRNA

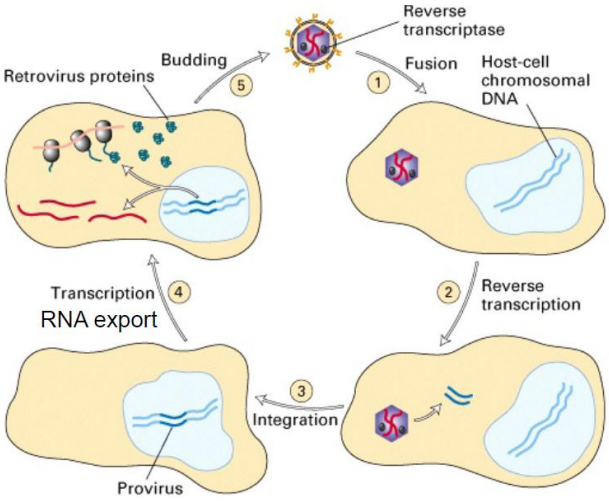
Interference with RanGTP leads to a disturbed snRNA and mRNA export. This is due to an indirect pathway. The export itself does not need RanGTP, but recycling of the export factors does.



What makes mRNA transport directional? Cytoplasmic removal of mRNA export factors! Pioneering round of translation, ATP-dependent enzymes (DEAD box proteins) at the NPC -> Dbp5, phosphorylation/dephosphorylation of RNA export factors (cytoplasm by Sky1p kinase, nucleus by phosphatase Glc7). Phosphorylated Npl3p associates with nascent transcripts in

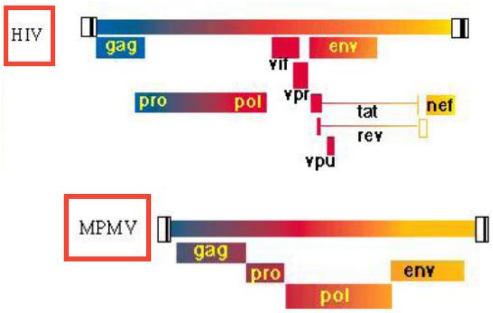
the nucleus, subsequently, 3' end processing of the transcript stimulates Glc7p to dephosphorylate Npi3p, resulting in release of the mRNP from the 3' processing machinery and association with Mex67p, thereby making the mRNP competent for export. In the cytoplasm, phosphorylation of Npi3p by Sky1p promotes the disassembly of exported mRNPs.

Retroviral RNA export



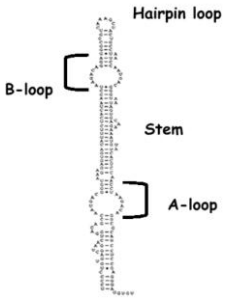
Retroviruses possess an RNA genome. The steps of the retroviral life cycle are entry, uncoating, reverse transcription, nuclear import, transcription and RNA processing, translation, assembly and budding and maturation.

Complex and simple retroviruses can be distinguished from each other by their genome organization. Complex viruses do have additional accessory proteins (such as HIV) whereas simple viruses only express gag, pro, pol and env (e.g. MPMV).

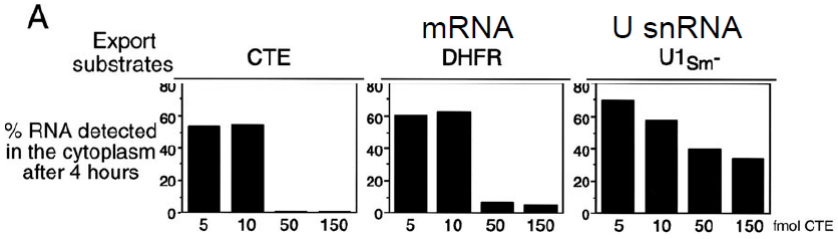


Retroviruses need to export unspliced, genomic RNA. Replication of retroviruses requires export of unspliced viral RNAs, both as templates for the synthesis of structural proteins and as genomic RNA to be packaged in progeny virions. All retroviruses therefore evolved some mechanism to bypass the requirement of splicing prior to export.

Example Mason-Pfizer Monkey Virus (MPMV): Contains mRNA export element CTE (constitutive transport element). It is required for export of unspliced, genomic MPMV RNA, but not for export of fully spliced RNAs. A and B loop, important for activity of CTE. CTE hijacks the cellular mRNA export.

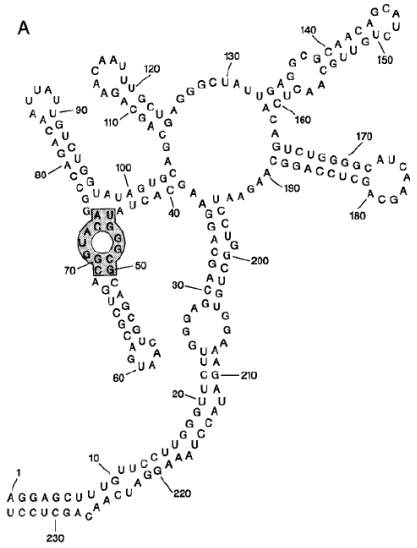


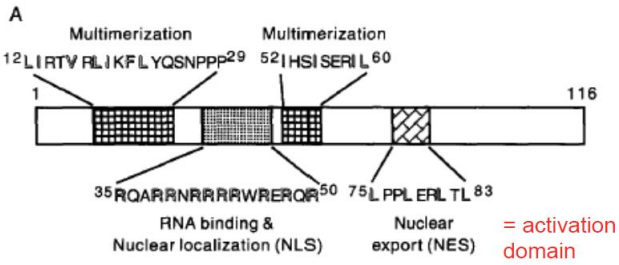
Experiment: 32P-labeled CTE, DHFR and U1Sm-RNAs were coinjected into the nuclei of *Xenopus* oocytes with increasing amounts of unlabeled CTE RNA. Phospho-imager analysis was performed to quantitate the extents of RNA export at 4 h in the presence of increasing amounts of CTE RNA competitor. This analysis is expressed as the percentage of RNA detected in the cytoplasm.



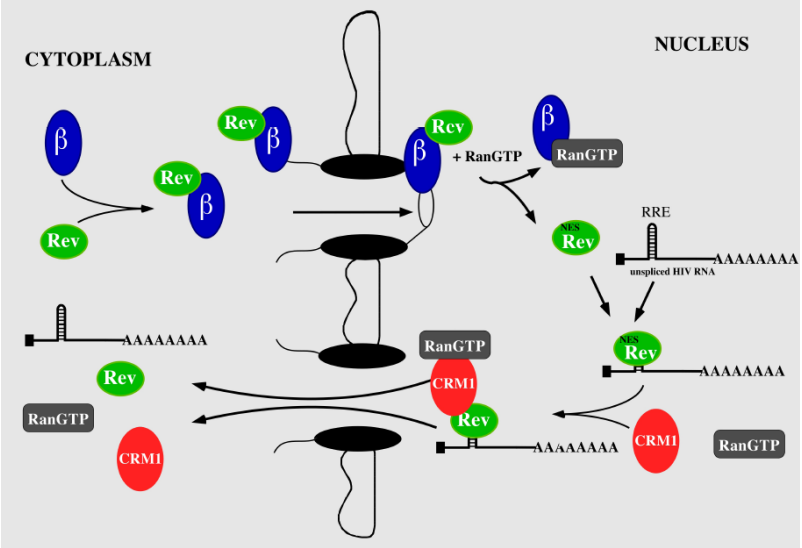
It seems that CTE and mRNA share a common component required for export. It has been shown, that TAP serves as the CTE export receptor. Recombinant TAP has been shown to stimulate CTE RNA export.

Example HIV (Human Immunodeficiency Virus): Rev is an accessory protein of HIV. The rev response element, RRE, is the strategy HIV uses to hijack mRNA export. Proposed structure of the HIV-1 RRE showing the location of the primary HIV-Rev RNA binding site. The Rev activation domain is a leucine-rich NES. It is a domain, which is required for Rev function *in vivo*, the activation domain in Rev drives nuclear export of unspliced genomic HIV RNA and it has been shown later to harbor nuclear export activity. As the activation domain is the classical leucine-rich NES it binds the exportin CRM1.



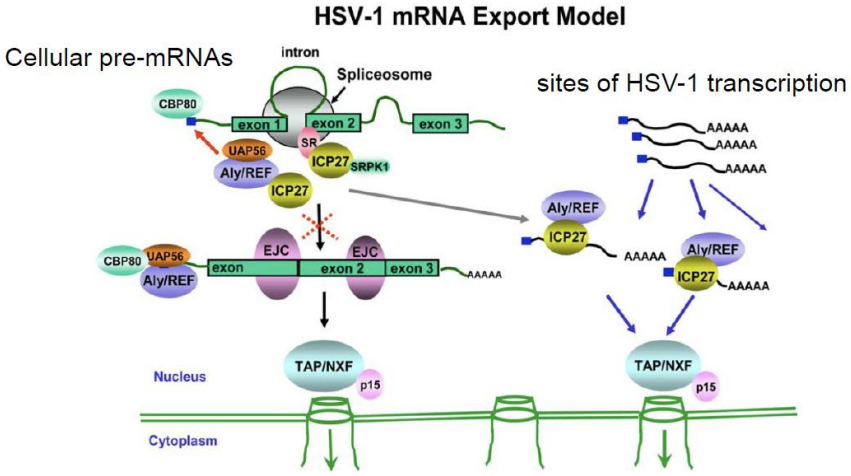


HIV-1 replication requires the cytoplasmic expression of unspliced, singly spliced and fully spliced viral mRNAs. Early in the viral life cycle, in the absence of Rev function, only fully spliced viral mRNAs, encoding the regulatory proteins Tat, Nef and Rev itself, are exported from the nucleus and expressed. Incompletely spliced viral mRNAs, encoding primarily viral structural proteins, are retained in the nucleus. Later, in the presence of Rev, these incompletely spliced viral

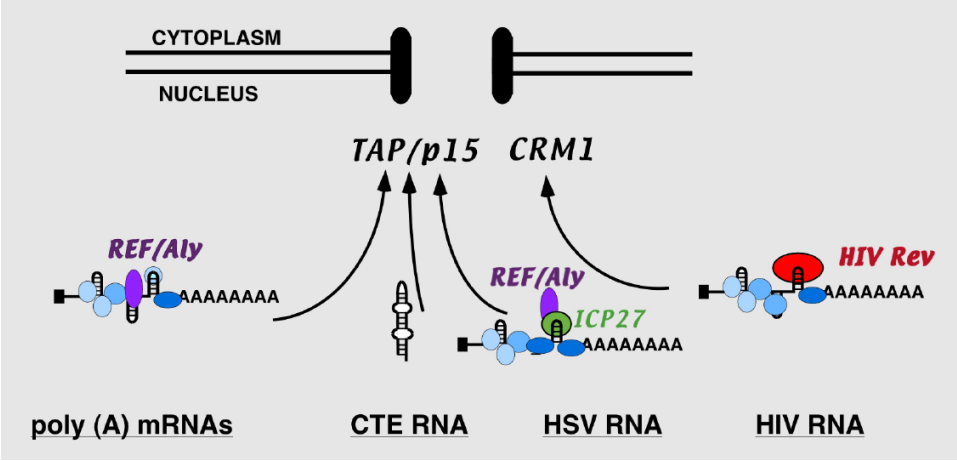


mRNAs are exported and expressed due to the recruitment of Rev and its associated cellular cofactors to the cis-acting Rev response element (RRE) RNA target.

Example Herpes Simplex Virus 1 (HSV-1): Enveloped DNA virus, capsid is transported through the cytosol to the NPCs, DNA is then released from the NPC bound capsid into the nucleus. Most transcripts encoded by HSV-1 are intronless, and therefore require a means to access cellular export pathways efficiently. HSV-1 encodes an essential protein termed ICP27, which mediates the export of viral intronless mRNAs (mutants of ICP27 show reduced levels of viral mRNAs and are defective in viral DNA replication). ICP27 shuttles between the nucleus and the cytoplasm. It interacts with Aly/REF and recruits it to sites of HSV-1 transcription. ICP27 binds viral mRNAs and the ICP27-Aly/REF-RNA complex is directed to TAP/NXF1. ICP27 also inhibits splicing of cellular mRNAs.



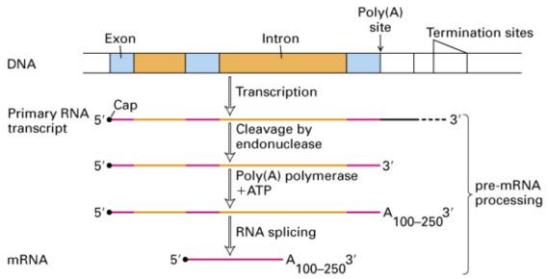
Summary:



Post-Transcriptional Control of Gene Expression

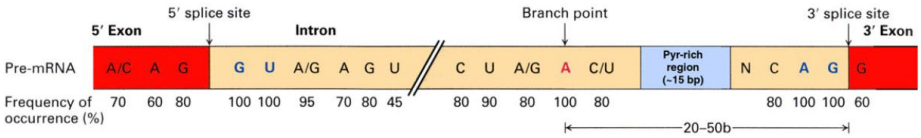
1. mRNA Processing

1.1. pre-mRNA Splicing



Consensus sequences define the splice sites:

- 5’ splice site (donor site) is a conserved GU dinucleotide
- 3’ splice site (acceptor site) is a conserved AG dinucleotide
- Polypyrimidine tract upstream of the 3’ splice site
- Adenine residue is used as a branch point in the splicing reaction



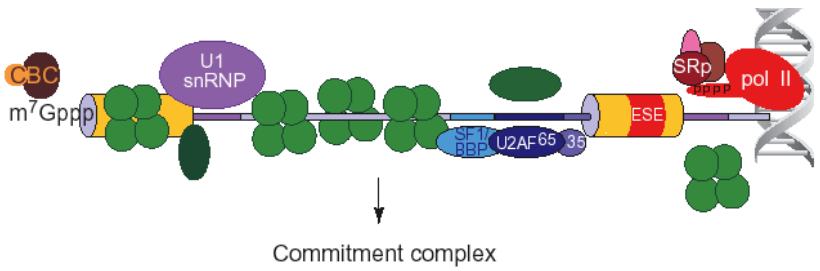
There is some autocatalysation in splicing, but mostly, a splicing machinery, the so called **spliceosome**, is needed, which is made up of:

- snRNPs (snurps): Main compoments of spliceosome, snRNA (U1, U2, U4, U5 or U6), Sm proteins, snRNP-specific proteins (e.g. U1: 70k, A, C)
- SR proteins: Contain Ser/Arg dipeptides, bind RNA (mostly), function as auxillary factors, recognize splice sites, important for assembly of the spliceosome
- hnRNP: Heterogenous nuclear RNA binding proteins, package nascent mRNA, more than 20 different ones, different functions.

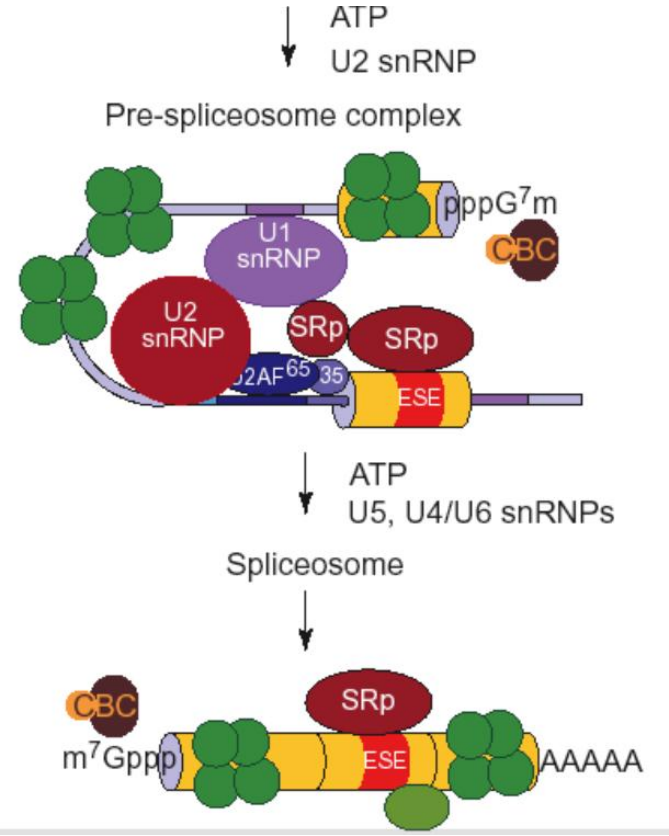
The spliceosome is a huge complex that catalyses the excision of introns and the ligation of the exons.

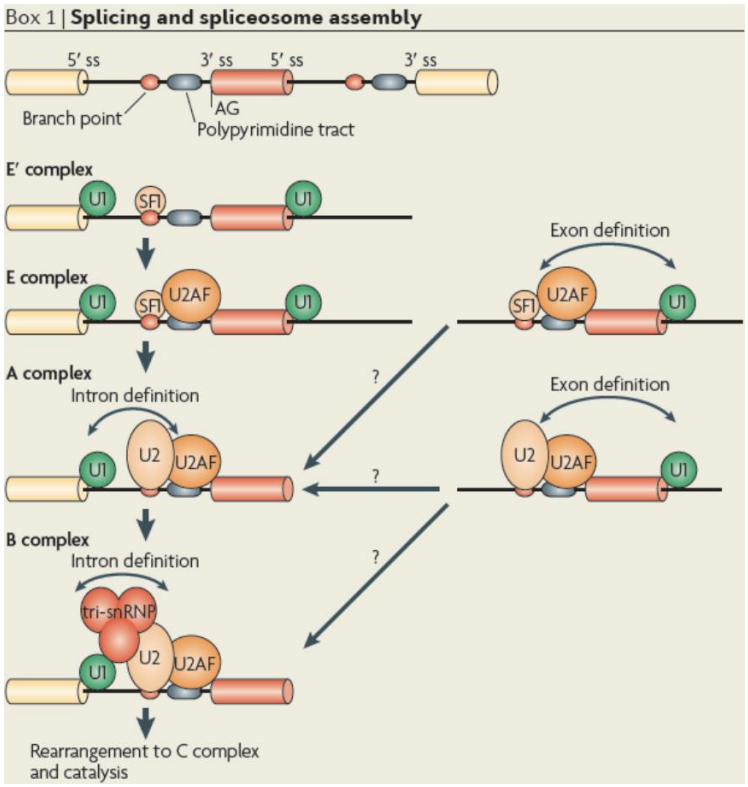
Spliceosome Assembly: As the exons are normally quite short (about 500 bp) and the introns can be up to several kb long, the exon has to be defined, which happens here through U1 and U2AF35:

1. U1 snRNP binds to the 5’ splice site
2. SR protein SF1 binds to branch point A
3. SR protein U2AF65 binds the polypyrimidine tract and SR protein U2AF35 the 3’ splice site
4. U2 snRNP associates with the branch point



5. U2 binds to the branch point, ATP hydrolysis to form the pre-spliceosomal complex
6. SR proteins play a role in promoting assembly by assisting binding of U snRNPs and branch point recognition (guiding U2AF)
7. Subsequent binding of U5 and U4/6 snRNPs allows the assembly of a mature spliceosome
8. Spliced mRNA then released from spliceosome, changes in hnRNP composition can occur (facilitate nucleocytoplasmic export and cytoplasmic events)





Alternative Splicing

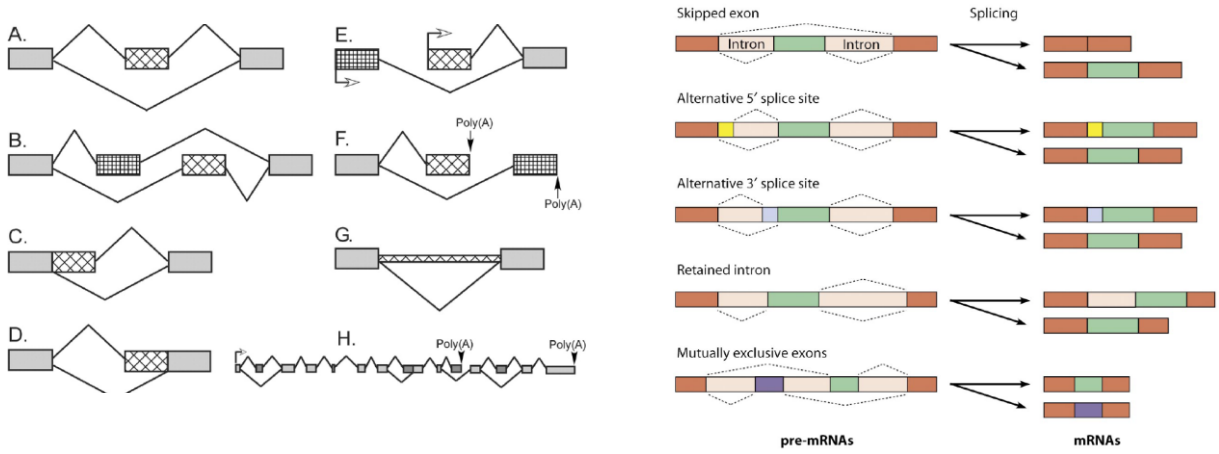
100 different isoforms of 1 gene -> protein diversity. Alternative mRNA splicing is the regulated process of differential inclusion or exclusion of regions of the pre-mRNA. It is an important source for protein diversity in eukaryotes and is often regulated in a temporal or tissue-specific fashion giving rise to different protein isoforms in different tissues or developmental states. Normally, splicing proceeds in a sequential order and all 5' splice junctions are functionally equivalent as are all 3' junctions. Variations in the use of splice junctions can produce different polypeptides from a single gene.

Example inner ear: Alternative splicing of the pre-mRNA for the K^+ channel leads to at least 576 different forms. These are expressed in a gradient along the sensory-receptor cells in the inner ear, which enables the perception of a wide range of sound frequencies.

There are different patterns of alternative splicing:

A. An exon can be either included or excluded
B. Mutually exclusive exons
C. Alternative 5’ and 3’ splice sites allow the lengthening or shortening of a particular exon
D. See C.
E. Alternative promoters can alternative poly(A) sites switch the 5’- or 3’-most exons
F. See E.
G. A retained intron can be excised or can be retained in the translated mRNA
H. A single pre-mRNA can exhibit multiple sites of alternative splicing

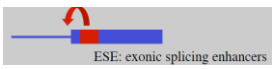
Together, combinations of these patterns allow to produce many different mRNAs. Boxes represent exons (grey ones depict those that are in all mRNAs) and lines the way they are joined together.



Mechanisms of alternative splicing: Trans-acting proteins (repressors and activators), which bind to cis-acting RNA elements regulatory sites (silencers and enhancers) on the RNA. Such elements may also influence how splicing will occur independently of proteins by their secondary structures. Together, these elements form a “splicing code” that governs how splicing will occur under different cellular conditions.

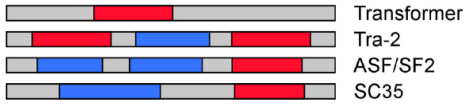
Positive regulation by Cis-enhancer elements:

- Purine-rich splicing enhancers: Mostly in exons, often activate splicing of the upstream intron, function as binding sites for trans-acting factors, such as SR proteins, ESE -> exonic splicing enhancers



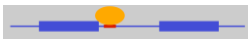
- Pyrimidine-rich enhancers: In introns close to the 5’ splice site, help to recruit U1 snRNP to it, ISE -> intronic splicing enhancers

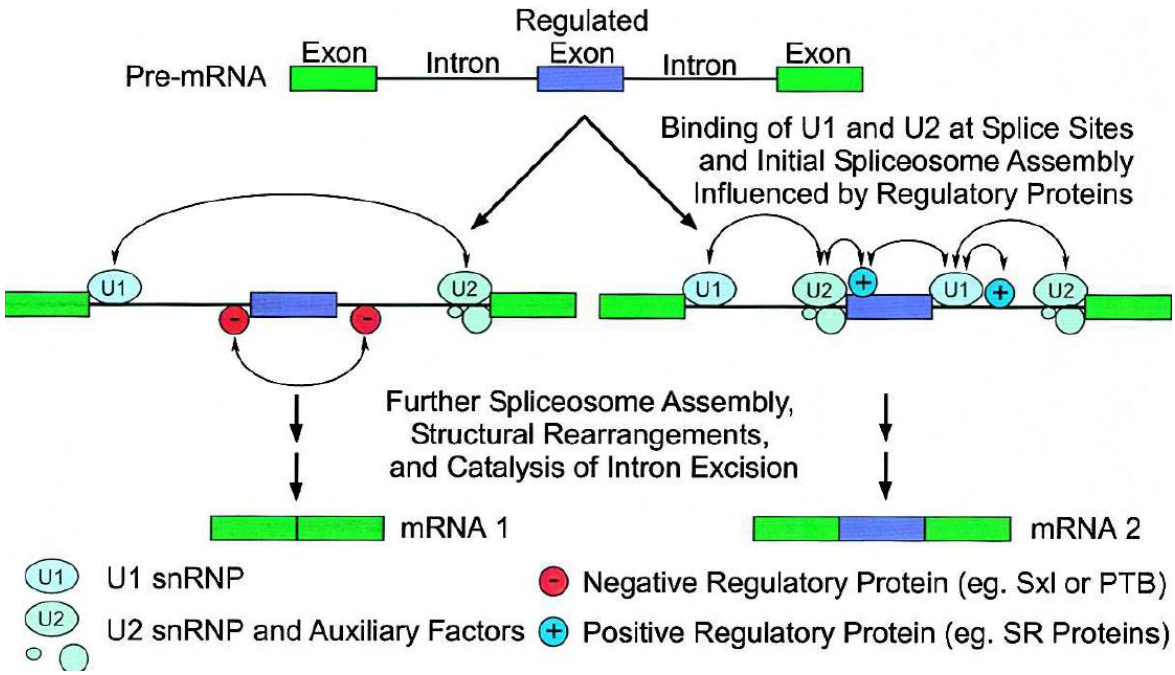
SR Proteins: Family of highly conserved proteins in metazoans, required for constitutive splicing and influence alternative splicing. SR proteins contain one or more copies of an RNA recognition mofit (RRM, in blue) and an RS domain (red) rich in alternating Arg and Ser residues. The RRM determines RNA binding specificity and the RS mediates protein-protein interactions to recruit splicing machinery. SR proteins often recognize exonic splicing enhancers (ESE) leading to the activation of suboptimal splice sites. Phosphorylation can modulate SR protein activity.



In the traditional enhancer model, the SR protein binds to the splicing enhancer and recruits U2AF via protein interactions between the RS domains of the SR protein and U2AF35. U2 snRNP is subsequently recruited by virtue of interactions between U2AF65 and components of U2 snRNP and by the interaction between the U2AS65 RS domain and the branch point.

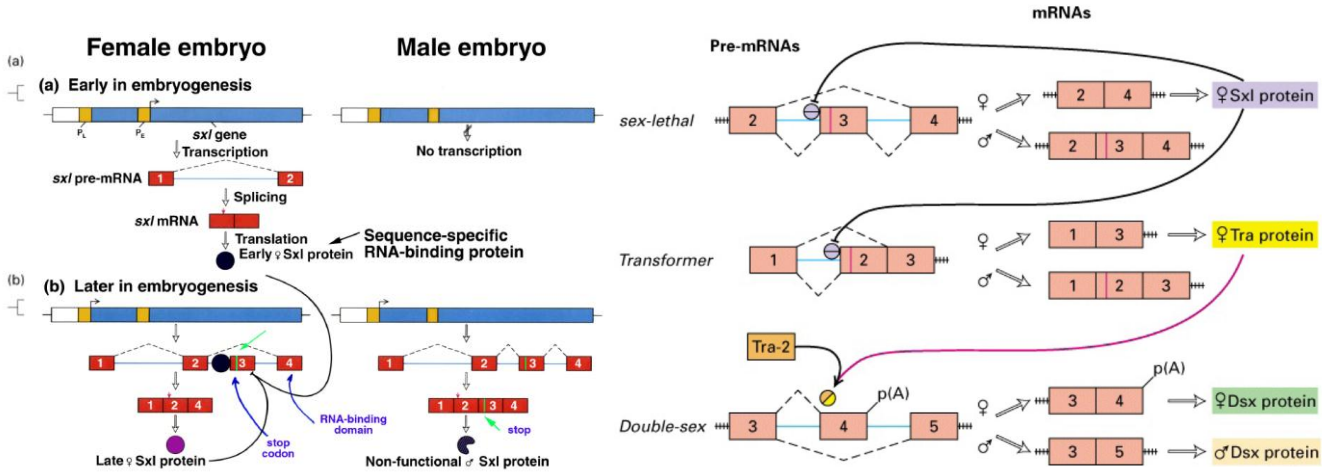
Negative regulation:

- ESS: Exonic splicing silencers
- ISS: Intronic splicing silencers
- Specific or general factors binding to regulatory elements within introns or exons or 
- Inhibitory secondary structure 

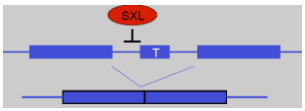


Example alternative splicing in *Drosophila* sexual differentiation: Sex in flies is determined by the ration of X chromosomes to autosomes, 2X gives raise to females, 1X to males. Gene dosage compensation is accomplished by enhancement of transcription of genes on the one X-chromosome in males. The female-specific protein Sex-lethal (Sxl) coordinately controls sex determination and dosage compensation.

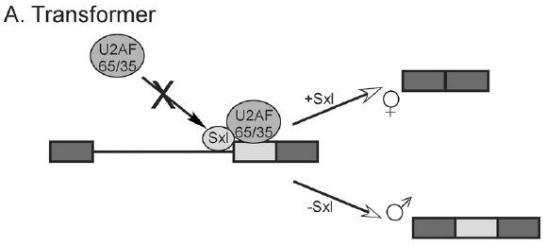
Sxl is a negative regulator of splicing in female flies (Sxl binds RNA and blocks binding of U2AF), in males, a premature stop codon prevents expression of Sxl in males. So in male embryogenesis, no Sxl is transcribed. As Sxl is there in females (expression is triggered in early development by the 2 X-chromosomes), Tra is expressed and this further leads to the expression of the female version of the Dsx protein.

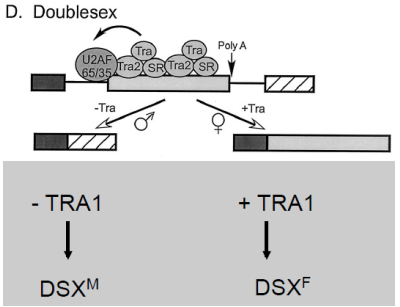


Sxl negatively regulates three known splicing events:

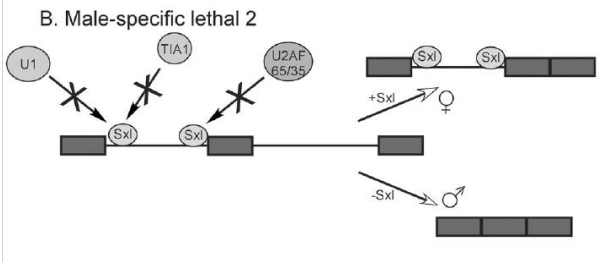


- Repression of inclusion of an exon which would hinder its complete translation (positive feed-back loop) in females
- Sxl binds the pyrimidine-rich region of Tra thereby blocking the binding of U2AF and U2 snRNP and preventing splicing of exon 2 to exon 3 that contains a stop codon (splicing at exon 4 instead, producing a functional TRA1 protein in females); Tra is not an RNA-binding protein, forms a complex with Tra2 (RNA-binding SR protein); the complex activates splicing of doublesex (activation of an upstream 3' splice site otherwise not used because of a poor sequence context); DSXF and DSXM are transcriptional repressors, establish appropriate expression patterns





- Male-specific-lethal 2 mRNA (msl2): MSL proteins are required for dosage compensation (hyperactivation of X-linked transcription) and are expressed only in males; in female flies, Sxl blocks splicing of an intron which then interferes with translation of Msl2

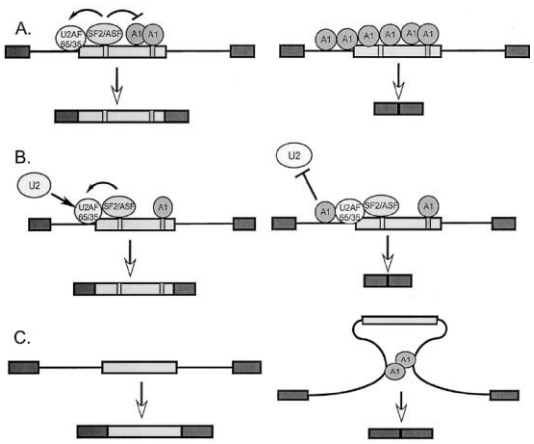


In vertebrates, no highly specific alternative splicing factors have been identified. Alternative splicing arises from variations in the relative concentrations or activity of splicing factors or from the strength and arrangement of binding sites for regulators and constitutive splicing factors. Again, SR proteins (positive) and hnRNPs (negative) are key players in alternative splicing:

- hnRNP A1: Antagonist of SR proteins, often exonic splicing repressor
- hnRNP F and H: Binding of exonic splicing silencer
- PTB (hnRNP I): Mostly intronic splicing repressor; often blocks poly-pyrimidine tract

Model for splicing repression by hnRNP A1:

A. A1 binding nucleates the assembly of additional A1 molecules along the RNA, creating a zone of RNA where spliceosome assembly is repressed
B. An additional A1 binding site adjacent to the branch point blocks splicing, A1 binding to this intronic element does block U2 snRNP binding to the branch point
C. Intronic binding sites for A1 allow it to multimerize, thus looping out the exon and causing exon skipping



Summary:

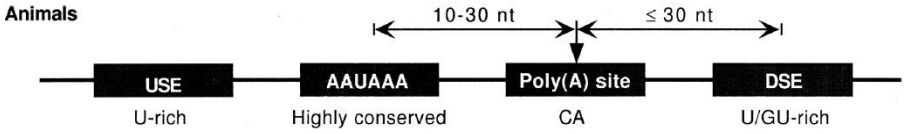
- Alternative splicing is a mechanism that increases the coding capacity of a single gene
- Alternative exons often have suboptimal splice sites
- Positive and negative regulation of splicing requires transacting factors like SR proteins (often positive regulators) and certain hnRNP proteins (antagonists)
- Specific splicing factors and SR proteins function as negative regulators of splicing in *Drosophila*; these influence sex determination and dosage compensation

1.2. mRNA 3'-end Formation

Roles of polyadenylation:

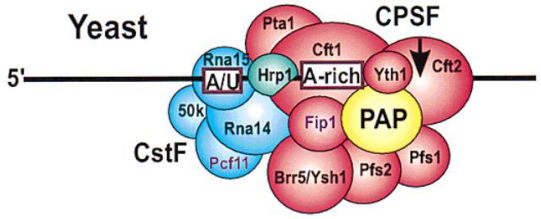
- Stimulation of transcription termination
- Facilitation of splicing of the most 3' intron
- mRNA export (stimulatory role)
- Generation of different mRNAs (alternative polyadenylation, can confer different characteristics to certain mRNAs, different localization, stability etc.)
- mRNA stability
- Translational initiation

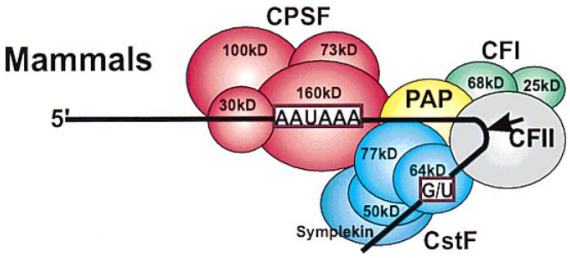
The poly(A)-signal is a conserved hexamer, AAUAAA, additionally with a poly(A) site, a U/GU rich downstream element and an upstream element (USE) which is positive or negative.

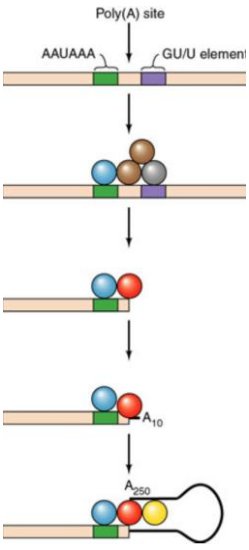


Polyadenylation is a two-step reaction in that the 3' end is cut and the remaining end polyadenylated and the part that has been cut is degraded. Some key players:

- CF: Cleavage factor 1
- CstF: Cleavage stimulatory factor
- CPSF: Cleavage and polyadenylation specificity factor
- PAP: Poly(A) polymerase
- PABII: Poly(A)-binding protein 2



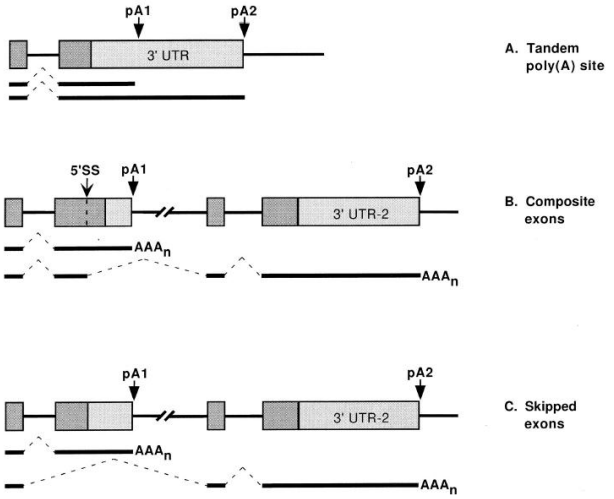




Mechanism: A complex of CPSF (blue), CFI and II (brown) and CstF (grey) bind to the AAUAAA and GU/U rich regions. Cleavage occurs and CPSF remains and is joined by PAP (red). PAP begins synthesis of poly[A], resulting in the addition of the first 10 A's. PAB II joins the reaction, stimulating the synthesis of poly [A], extending the tail to about 200 A residues.

Regulation: Yes/no decision, either a transcript is processed or not, e.g. U1 snRNP component U1a blocks polyadenylation of its own pre-mRNA. If there are two poly(A) sites, there is a question on where to process and place the 3' end:

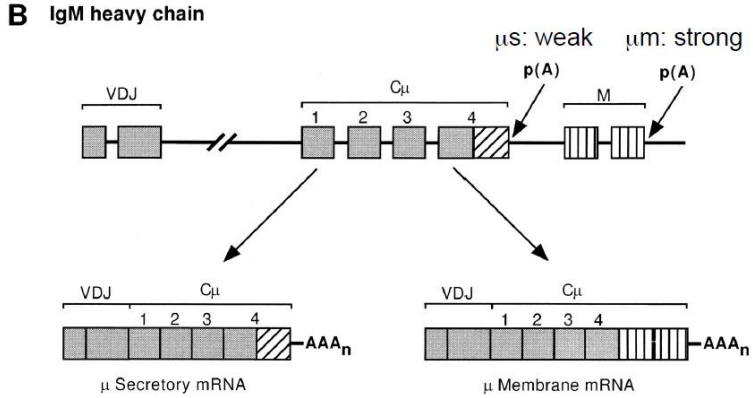
- Tandem arrays of poly(A) sites within a single 3'UTR
- Composite exons whose 3' end is formed by either a 5' splice site or a poly(A) site (splice or polyadenylate?)
- Alternative 3' exons



How are these decisions made?

- Regulation of the efficiency with which the processing complex assembles:
  - Many regulated poly(A) sites are weak
  - Effect of the concentration of polyadenylation factors
  - Additional stabilizing influence of other factors recruited onto polyadenylated enhancers
- Activity of processing enzymes may be regulates by post-translational modifications or by interacting repressor proteins
- So far, those tissue- or transcript-specific factors have not yet been discovered!

Example regulation of the IgM heavy chain synthesis during B cell differentiation: IgM can exist in two functionally and structurally different forms, membrane bound (mIgM) and secreted (sIgM). mIgM acts as an antigen receptor in pre B-cells, whereas sIgM is the first type of antibody found in the blood when an organism is exposed to an antigen for the first time. sIgM and mIgM differ from each other at their carboxyl ends, the molecules are identical except that the membrane-bound form contains a hydrophobic "tail" that keeps it inserted in the membrane. The mRNAs for membrane and secreted forms also differ. Both of them are found to be transcribed from the same gene, but differently processed.

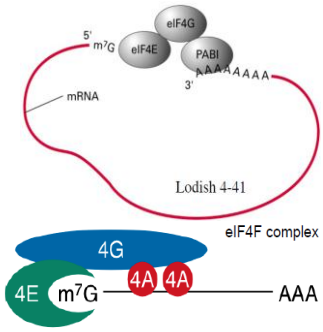


Two different polyadenylation sites, CstF-64 has a higher affinity for the second polyadenylation site (μm). In plasma cells, CstF-64 is there in high concentrations and the use of the weak affinity poly(A) site is favored whereas in pre-B cells, CstF-64 is the limiting factor and therefore the strong poly(A) site μm is preferred.

Summary:

- Polyadenylation is performed in a two-step reaction by a complex machinery of processing factors
- Polyadenylation can be controlled by the abundance of processing factors and the strength of the polyadenylation signal
- Competition between splicing and polyadenylation might also influence choice of poly(A) site

2. mRNA Stability



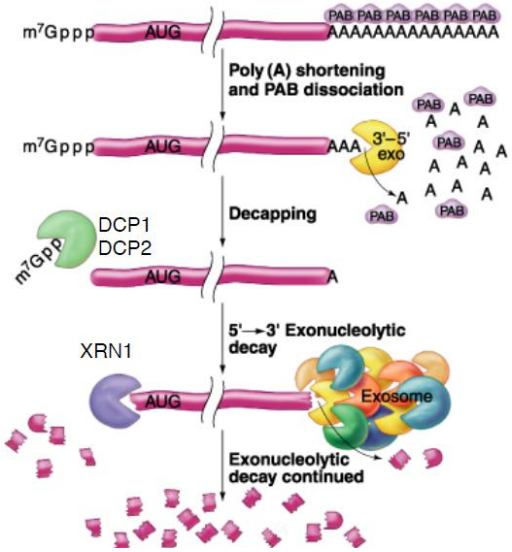
The steady-state concentration of an mRNA is determined by the rates of synthesis and decay. The half-life of individual mRNAs within a given cell may vary by several orders of magnitude, from a few minutes to many hours or even days. The half-life of an mRNA determines how rapidly protein synthesis of the respective protein can be shut off after transcription ceases (therefore half-life of mRNAs coding for signaling molecules like lymphokines or certaint TK factors (e.g. involved in cell cycle control) is very short). mRNAs are protected from degradation by the 5' cap and the 3' poly(A) tail.

Poly(A) mRNAs form circles by interactions between the 5' and 3' ends. This interaction is important for initiation of protein synthesis and mRNA stability. PAB1 binds the poly(A) in the cytoplasm and at the same time it binds to a subunit of the cap bound translation initiation complex eIF4 called eIF4G.

2.1. Determinants

Cis-acting elements determine mRNA stability.

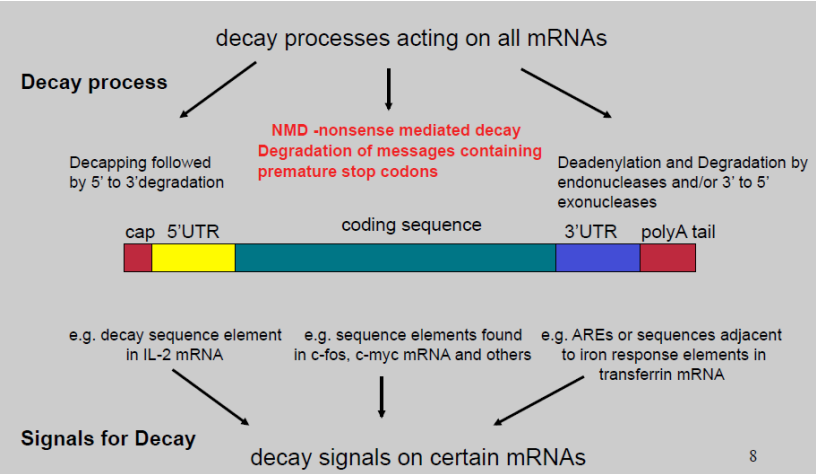
- Removal of the 5' cap followed by degradation from 5' to 3' by endonucleases, e.g. decay sequence element in IL-2 mRNA
- Deadenylation and degradation by endonucleases and/or 3' to 5' exonucleases, e.g. AREs or sequences adjacent to iron response elements in transferrin mRNA
- NMD nonsense mediated decay, degradation of messages containing premature stop codons, e.g. sequence elements found in c-fos, c-myc and others



2.2. Global Decay and Nonsense-Mediated Decay (NMD)

As the mRNA ages, the poly(A) tail gets gradually shorter, which releases poly(A) binding proteins, which then leads to the degradation of poly(A) by 3' to 5' exonucleases (CCR4/NOT complex). As the RNA can no longer be circularized, the RNA also loses its 5' cap (DCP 1/2). Without this physical interactions between 5' and 3' ends and the protecting proteins, the RNA is rapidly degraded by 5' to 3' exonucleases (XRN1) and by 3' to 5' exonucleases (exosome).

P bodies (cytoplasmic foci) contain mRNA decay factors (all proteins that function in the 5' to 3' decay have been localized to P bodies) and seem to be sites for mRNA decay, as the number and size of P bodies has been correlated to a rise in mRNA decapping, and intermediates can be found. Some factors found are Dcp1,2p, Lsm1-7 (decapping factors), Ccr4 (deadenylation factors); 5' to 3' exonuclease (Xrn1).



Nonsense-mediated decay: Cellular processes that eliminate aberrant mRNAs that contain nonsense mutation withing the protein coding region. Transcripts harbouring premature termination codons (PTC) are degraded rapidly, preventing the synthesis of incomplete and potentially deleterious proteins that could dominantly interfere with the normal functioning of the cell, this is the case for:

- Transcription errors
- Mutant genes
- Improper pre-mRNA processing events, such as splicing and polyadenylation

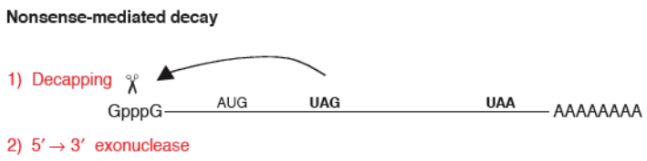
The NMD pathway has been observed in all eukaryotic cells examinde, often also described as mRNA surveillance.

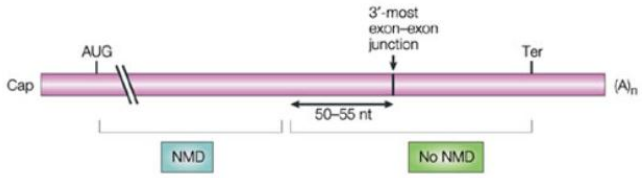
Ex.: PGK1 mRNA, Northern blotting of RNAs isolated at different times after inhibition of transcription (shift of a yeast strain harbouring a temperature-sensitive RNA polymerase II to 37 °C). Presence of PTCs enhance the decay of PGK1 mRNA ($t_{1/2}$=45 vs. $T_{1/2}$=3 min).

Translation is the basis of discrimination between normal and premature termination codons, and the activity of the NMD pathway depends on the recognition of the premature stop codon by the translation machinery:

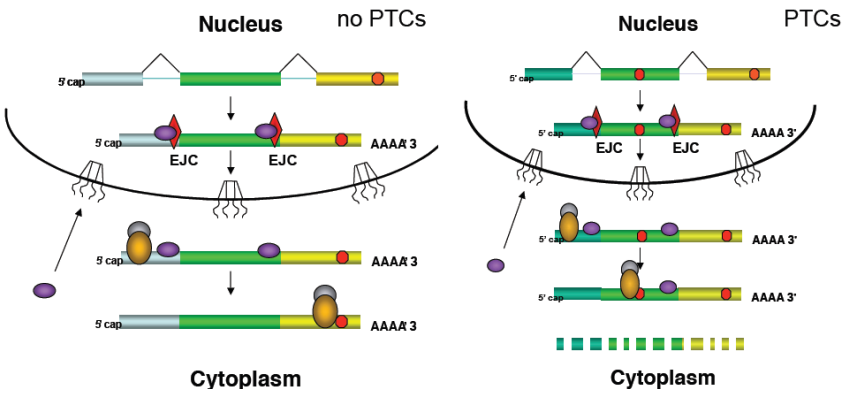
- NMD is inhibited by drugs (cycloheximidine) and mutations that block translation initiation and elongation (the rapid turnover of nonsense-containing mRNAs resumes immediately after cycloheximide is washed out from the growth medium)
- The NMD pathway is functional only after at least one translation initiation has been completed
- Nonsense-containing mRNAs are associated with polysomes whose size reflect the position of the premature termination codon within the open reading frame
- NMD can be prevented by nonsense-suppressing tRNAs

The turnover of nonsense-containing mRNAs can occur in deadenylation-independent manner. Nonsense-containing mRNAs are decapped, followed by a 5' to 3' degradation of the body of the transcript by the Xrn1p exoribonuclease. Once these mRNAs are decapped, deadenylation and 3' to 5' decay by the exosome also contribute to NMD.



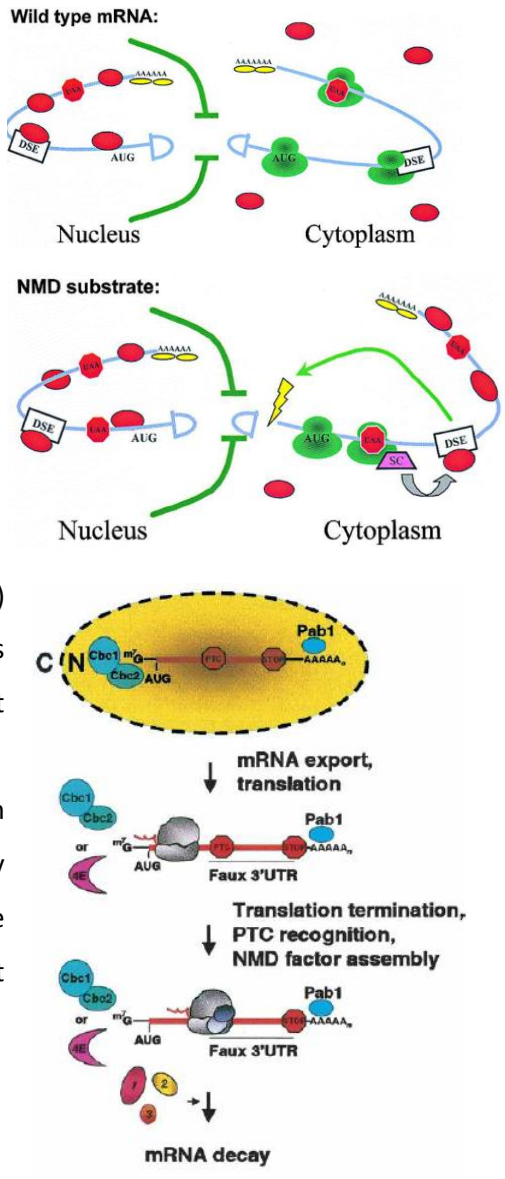


Position-of-an-exon-exon-junction rule: In spliced mRNAs, the last exon-exon junction is important for the definition of a PTC. Mammalian mRNAs generally have an average of 7-8 splicing-generated exon-exon junctions. A premature termination codon (PTC) that is located in the region indicated in blue, which is followed by an exon-exon junction more than 50-55 nucleotides downstream, elicits nonsense-mediated mRNA decay whereas a PTC that is located in the region indicated in green fails to elicit NMD. The



normal termination codon usually resides within the 3'-most exon. The 3'-most exon-exon junction determines where codons are recognized. But how? Factors involved in NMD are UPFs, SMGs, Y14 etc. During translation of a PTC-containing transcript, the downstream EJC signals to the terminating ribosome via the EJC-associated factor Upf3. Upf3 together with Upf1 and 2, may signal the presence of the PTC to the 5' end of the transcript, resulting in decapping and rapid exonucleolytic digestion of the mRNA. In yeast, there are no EJC, and there are two different models for NMD:

1. Intronless mRNAs contain a downstream element (DSE) defining PTCs by interaction with Hrp1 (in red). Nuclear RNPs are remodeled upon entry into the cytoplasm. An mRNA is exported to the cytoplasm as an RNP complex with nuclear RNA-binding proteins attaches, such as Hrp1/Nab4. During the initial rounds of translation, the attached nuclear proteins are displaced by the ribosomes, and complete translation to a cytoplasmic RNP is achieved. Incomplete RNP remodeling activates NMD. The premature termination codon prevents the ribosome from displacing Hrp1 from the DSE and following termination, the surveillance complex (SC) recognizes Hrp1p as a signal that RNP remodeling is incomplete and subsequently, the mRNA is decapped at the 5' end.
2. Measuring the distance between the termination codon and the poly(A) tail (Pab1 localization). An abnormally long distance between the termination site and the poly(A)-bound Pab1 is thought to result in inefficient translation termination.



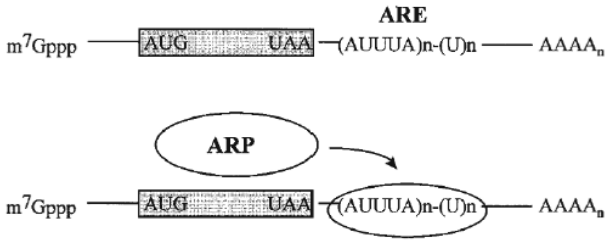
2.3. Specific Decay

AREs control stability of specific mRNAs. AREs are AU-rich elements, AUUUA, found in 3' UTRs of cytokines, proto-oncogenes and growth factors and mostly are destabilizing (but can become stabilizing when associated to certein proteins). Elements may be complex and composed of different types:

- AUUUA with couples nearby U-rich region or stretch
- At least two overlapping nonamers in a U-rich region [UUAUUUA(U/A)(U/A)]
- U-rich regions

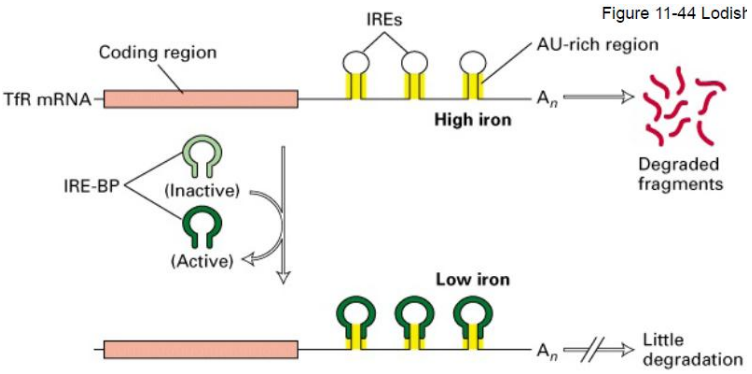
Group	Motif	Examples
I	WAUUUAW and a U-rich region	c-fos, c-myc
IIA	AUUUAUUUAUUUAUUUAUUUA	GM-CSF, TNF-α
IIB	AUUUAUUUAUUUAUUUA	Interferon-α
IIC	WAUUUAUUUAUUUAW	cox-2, IL-2, VEGF
IID	WWAUUUAUUUAWW	FGF2
IIE	WWWWAUUUAWWWW	u-PA receptor
III	U-rich, non-AUUUA	c-jun

When AREs of short-lived proteins are inserted into stable mRNAs, these mRNAs get short-lived as well. ARE-binding proteins (ARPs) control the stability of ARE-mRNAs. There is a good correlation between ARE-binding protein abundance or activity and increase or decrease in mRNA decay rates. Pathway of decay of these rapidly degraded mRNAs is thought to be associated with rapid and complete deadenylation, followed by degradation of the body via endo- and/or exonucleases. Binding of ARPs can have positive or negative effects on stability. AUF/hnRNP D: Binding to ARE correlates with instability, HuR/HuA: binding to ARE correlates with stabilization.



Interaction of ARE with a destabilizing factor such as AUF1 might promote rapid deadenylation by reducing the affinity of the poly(A) binding protein (PABP) for the poly(A) tail. Conversely, stabilizing factors such as HuR might enhance binding of PABP to the poly(A) tail.

Ex.: Transferrin receptor mediates uptake of iron from the blood. IRE is the iron response element (stem-loop structure) and the IRE-BP is an IRE binding protein, bifunctional (aconitase in the presence of iron). Binding of IRE-BP to IREs in 3' UTR prevents nucleolytic attack.



Summary:

- The stability of different mRNAs in the cytoplasm varies widely
- Generally, mRNAs are degraded by exonucleases after poly(A) tail shortening and decapping

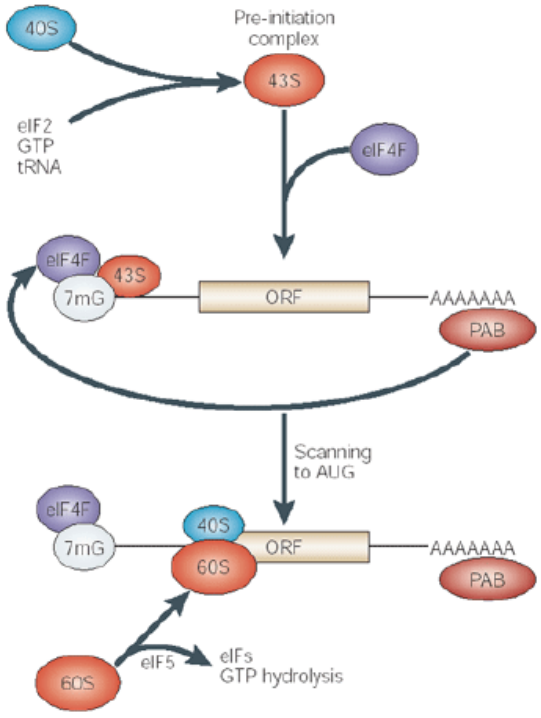
- RNAs containing premature termination codons are funneled into the NMD pathway; recognition of PTCs requires translation and the EJC for spliced mRNA, whereas DSEs are involved for intronless mRNAs
- Some cellular mRNA have short half-lives because they contain AU-rich elements (ARE) which bind ARPs that can have either destabilizing or stabilizing functions
- The rate of degradation of some specific mRNAs can be regulated; the transferrin receptor mRNA contains IREs that can protect the mRNA from degradation if iron uptake is required

3. Translational Initiation

Translation is mostly regulated at the level of initiation. Global regulation takes place by the modification of eukaryotic initiation factors (eIFs) and for individual mRNAs, structural featers of the mRNA itself (e.g. 5' or 3' UTR inhibit or act as receptors for regulatory RNA-binding proteins. Global control of protein synthesis is generally achieved by changes in the phosphorylation state of initiation factors or the regulators that interact with them. Two well-characterized examples are discussed. Translation initiation can also be regulated by RNA-binding proteins that modulate the translation of specific RNAs.

3.1. Initiation

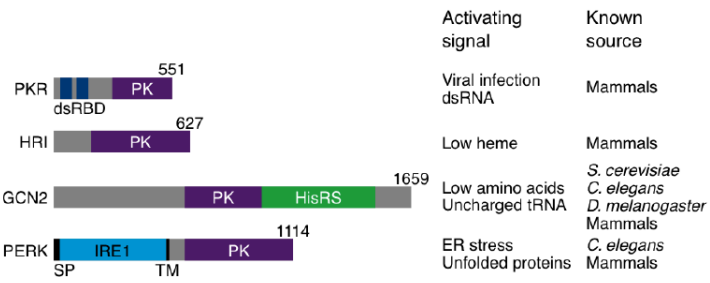
- 40S assembles with eIF2, GTP and initiator tRNA to form the 43S pre-initiation complex
- 5' cap of the mRNA attracts the eIF4F complex (composed of eIF4E (cap-binding protein), eIF4A (RNA-dependent ATPase) and eIF4G (recruits the 40S subunit))
- The 43S complex then scans the mRNA until the AUG codon is recognized
- This triggers eIF5 to hydrolyze GTP thereby promoting the other eIFs to dissociate and the 60S subunit to join resulting in a fully functional ribosome



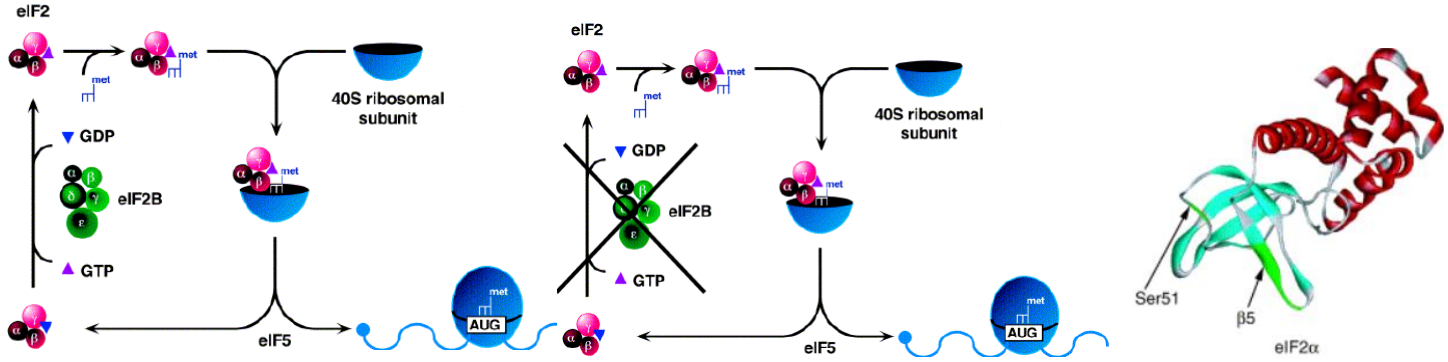
3.2. Global Regulation of Translation Initiation Factors

Regulation of eIF2: EIF2 is made up of 3 subunits, α, β and γ. In mammals, phosphorylation of the α-subunit (eIF2α) is induced in response to a number of different stress conditions, e.g. iron (haem) deprivation, heat-shock, viral infection of interferon. Four different kinases have been identified:

- Haem-controlled repressor (HCR)
- PKR (protein kinase activated by dsRNA)
- GCN2 (monitors uncharged tRNAs - starvation for amino acids)
- PERK (PKR-like kinase activated by ER stress)



Phosphorylation of eIF2α inhibits guanine nucleotide exchange by eIF2B which is the responsible GEF. The result is a general down-regulation of protein synthesis under various stresses. EIF2 is a key target in translational regulation. Phosphorylation of eIF2α on Ser51 converts it from a substrate to a competitive inhibitor of its exchange factor, eIF2B.



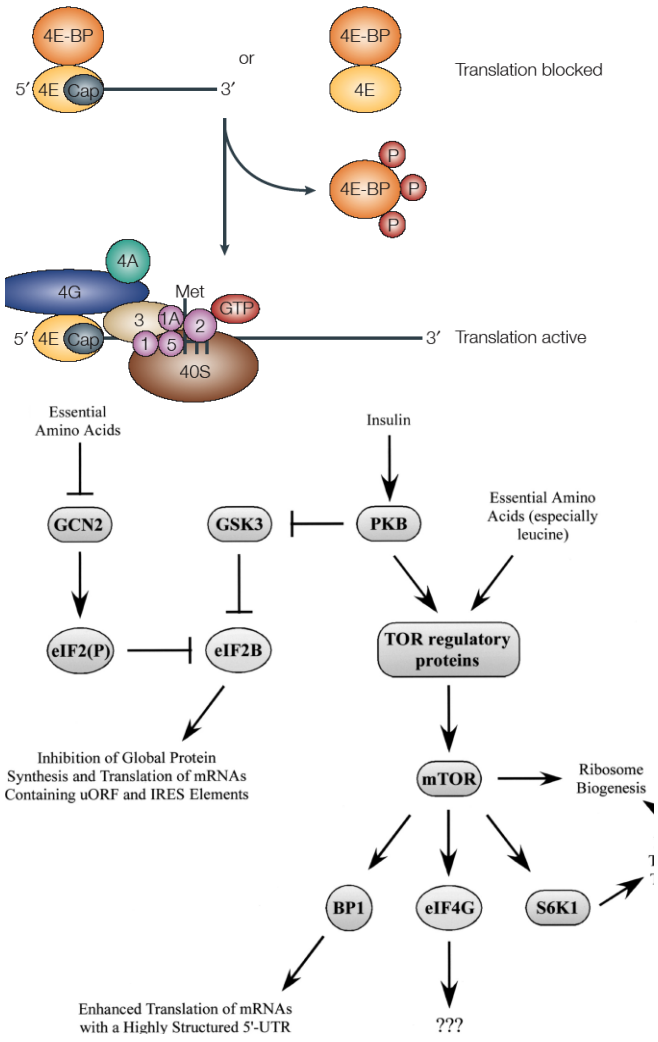
Regulation of eIF4E: eIF4E is the cap-binding subunit of eIF4F that facilitates 43S complex binding to the mRNA. There are 3 different ways of regulation:

1. Limiting concentrations of eIF4E
2. Phosphorylation of eIF4E (activation): Phosphorylated eIF4E has a 3-fold greater affinity for both the cap structure and eIF4G. Phorphorylation takes place in response to hormones and growth factors and dephosphorylation in response to heat-shock and viral infection (exeption: HSP mRNAs are efficiently translated after heat-shock as initiation of their synthesis is not strictly dependent on eIF4E)

3. Two translational repressors (4E-BP1 and 2) regulate eIF4E function, 4E-BP1 becomes phosphorylated in cells treated with hormones or growth factors and subsequently dissociates from eIF4E

eIF4E binding proteins (eIF4B-BPs) bind eIF4E, preventing its interaction with eIF4G and so inhibiting translation. Phosphorylation of BPs is triggered by hormones or growth factors releases the 4E-BPs from eIF4E which allows interaction with eIF4G and allows translation to proceed.

TOR (target of rapamycin) kinase mediates 4E-BP phosphorylation. The kinase is activated in response to amino acids and growth factors and phosphorylates 4E-BP1.



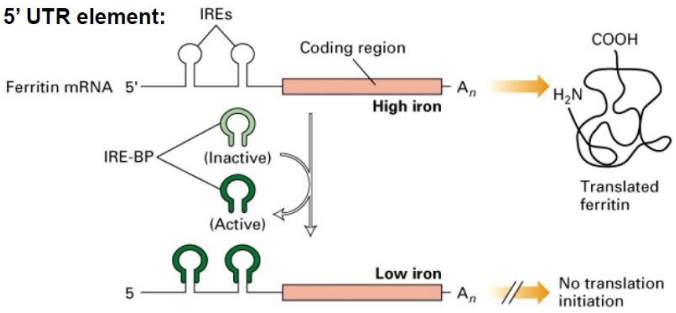
3.3. mRNA specific Regulation

Principles and examples:

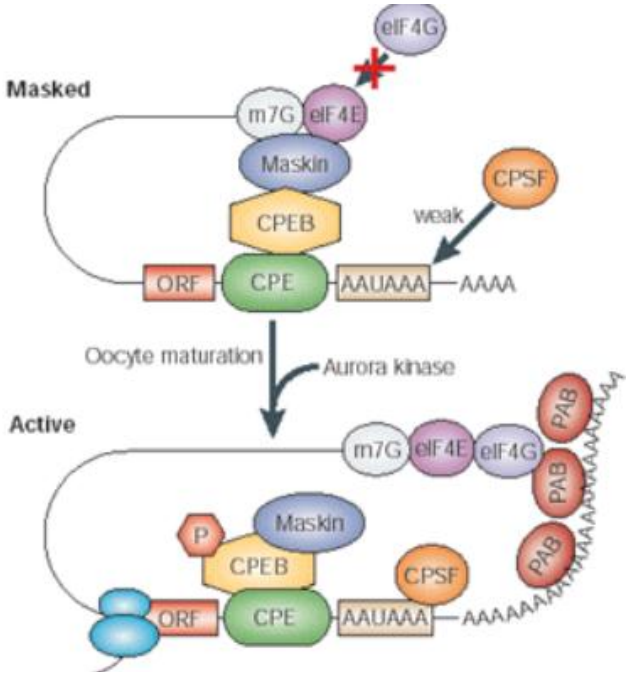
- Steric hindrance by elements in the 5' UTR (ex. IRE in ferritin mRNA)
- Interference with eIF4F complex assembly (ex. binding of maskin/CPEB or Bicoid to certain maternal mRNAs during vertebrate oocyte maturation and early development)
- Interference with 60S subunit joining by hnRNP K and hnRNP E1 (ex. 15-lipoxygenase mRNA, element in 3' UTR called DICE)
- Translocational control by micro RNAs (ex. bantam miRNA control of Hid)

Example ferritin: Ferritin is an iron storage protein and needed if there is too much iron. IRE located in te 5' UTR bound by IRP at low iron concentration and blocks initiation.

Example early development: In animal development, some maternally inherited mRNAs, which encode



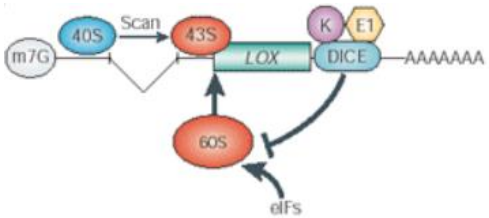
proteins that are crucial for embryogenesis, are translationally silent, or masked. When masked, these mRNAs typically have short poly(A) tails. Masking is controlled by the cytoplasmic polyadenylaion element (CPE) that is located in the 3' UTR. CPE binds CPEB and Maskin. Maskin associates directly with eIF4E , preventing its interaction with eIF4G. CPEB also prevents the stable association of cleavage and polyadenylation specificity factor (CPSF) with the AAUAAA sequence resulting in typically short poly(A) tails. The result is a translationally silenced, or masked, mRNA. Induction of oocyte maturation by, for example, exposure to progesterone, causes phosphorylation of CPEB and results in two significant events:



1. Maskin dissociates from eIF4E and CPSF binds to the AAUAAA
2. Binding of CPSF recruits poly(A) polymerase that extends the poly(A) tail

This allows the association of the 5' and 3' ends of the mRNA by the interaction of poly(A) binding protein (PAB) with eIF4G, which in turn, binds to the newly liberated eIF4E. This allows the recruitment of the 40S suunit.

Example erythropoiesis: Durin red blood cell differentiation, the 15-lipoxygenase (r15-LOX) mRNA is synthesized at an early stage of erythorpoiesis, but is translationally repressed until the mature reticulocyte stage is reached in the peripheral blood. This translational control ensures that r15-LOX, an enzyme that catalyzes the dioxygenation of phospholipids in mitochondrial membranes and thereby induces the breakdown of the mitochondria, is expressed at the correct time. Before maturation into peripheral blood reticulocytes, hnRNPs K and E1 associate with the differentiation control element (DICE) in the LOX 3' UTR and inhibit the joining of the 60S ribosomal subunit with the 43S pre-initiation complex that is primed for translation at the AUG. More specifically, hnRNPs K and E1 probably inhibit one or more of the eukaryotic initiation factors (eIFs) involved in 60S subunit joining.



Summary:

- Translational regulation can be glocal or mRNA-specific, and most examples of translational regulation that have been described so far affect the rate-limiting initiation step
- Global control of translation is frequently exerted by regulating the phosphorylation or availability of initiation factors. Two of the most well-known examples are the regulation of eukaryotic initiation factor (eIF)4E availability by 4E-binding proteins (4E-BP), and the modulation of the levels of active ternary complex by eIF2 phosphorylation
- mRNA-specific translational control is driven by RNA sequences and/or structures that are commonly located in the untranslated regions of the transcript; these features are usually recognized by regulatory proteins or micro RNAs (miRNAs)
- Many regulatory proteins target the stable association of the small ribosomal subunit with the mRNA; these factors function by steric hindrance (for example iron-regulatory proein IRP), by interfering with the eIF4F complex (ex. Maskin, bicoid) or others

Post-Transcriptional Gene Silencing by siRNAs and miRNAs

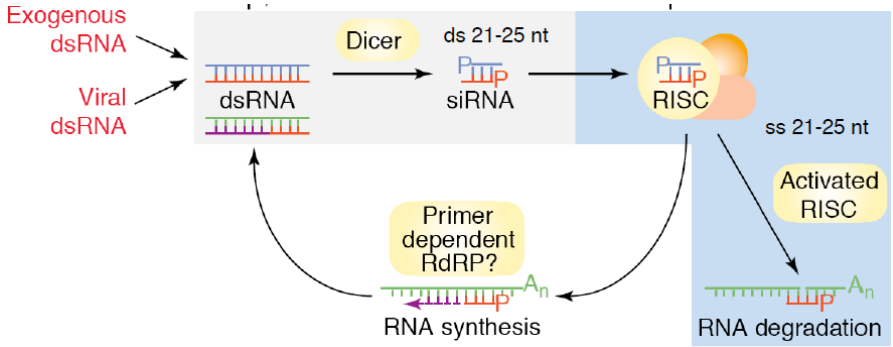
RNA interference is the process of sequence-specific post-transcriptional gene silencing initiated by double-stranded RNA that is homologous to the silenced gene. RNAi is a protecting mechanism against invasion by foreign genes. RNA silencing by micro RNAs ia a gene-regulatory mechanism in multicellular organisms.

Discovery: Studies in plants and fungi uncovered gene silencing pathways that were triggered by transgene expression or viral replication. *C. elegans* developmental studies identified lin-4, a small RNA that controls the expression of lin-14 mRNA. dsRNA induces gene silencing in *C. Elegans*.

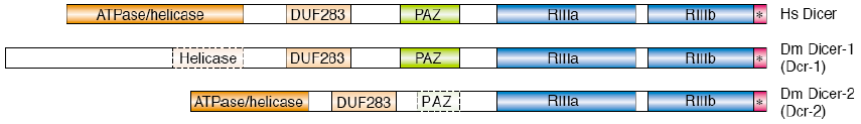
RNAi Mechanism

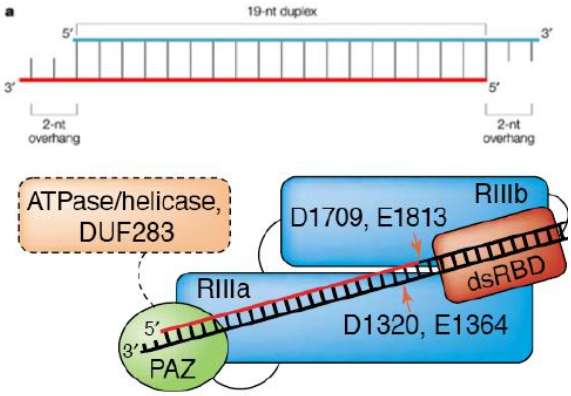
In plants, fungi and worms (not in mammals and flies), RNA-dependent RNA polymerase RdRP leads to a signal amplification (secondary siRNA production).

Step by step:

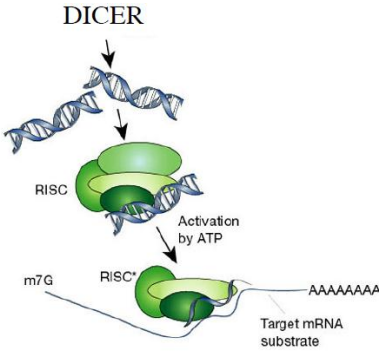


- Initiator step: Dicer, the siRNA producing enzyme belongs to the Rnase III-family. Loss of dicer leads to early embryonic lethality. Dicer is 220 kDa large and a multi-domain protein present in all eukaryotes studied to date, exception is budding yeast. Dicer is made up of ifferent domains -> DexH/DEAH RNA helicase/ATPase domain, DUF283 domain of unknown function, PAZ domain, RNase III-like domains, dsRNA-binding domain (RBD). The produced siRNA duplexes conatain a 5' phosphate and a 3' OH with 2 non-paired nucleotides on 3' end (3' overhang). How is this achieved? The Rnase III domains of Dicer are active when dimerized, and Dicer then cuts the dsRNA preferentially at their termini. Dicer contains a signle dsRNA cleavage center with two independent catalytic sites. The center is formed by the RIIIa and RIIIb domains of the same Dicer molecule. It processes the dsRNA 20bp from its terminus. The PAZ domain recognizes the 3' overhang end created by Dicer (dsRNA) or Drosha processing (miRNA pathway). DmDcr-2 and CeDcr use ATP, hsDcr does not.



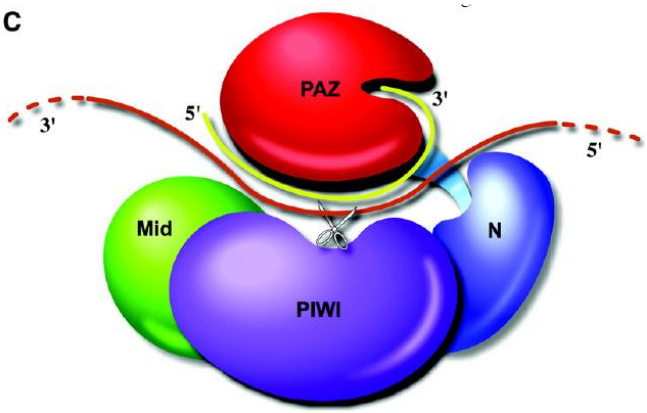


- For RNAi, only the antisense strand is needed. The effector step includes the RISC (RNA-induced silencing complex). It mediates ds siRNA unwinding and is responsible for sequence specific mRNA degradation mediated by siRNAs. RISC is a 500 kDa multiprotein complex (with a 140 kDa core).
  - RISC induces ATP-dependent unwinding of dsRNA
  - RISC targets the ssRNA to specific, homologous mRNAs
  - Slicer cleaves the mRNAs



The argonaute protein family (PPD) are the key components of RISC complexes. It is a large and evolutionary conserved family and is required for post-transcriptional gene silencing. The proteins are about 100 kDa in size and highly basic and contain two common domains, PAZ and Piwi. The PAZ domains is 130 aa long and is names after three proteins (Piwi, Argonaute, Zwille/Pinhead), is also present in Dicer and has an oligonucleotide binding fold. The Piwi domain is 300 aa long, was first identified within the C-terminal part of Piwi, is highly conserved (also found in prokaryotes). The fold of the Piwi domain

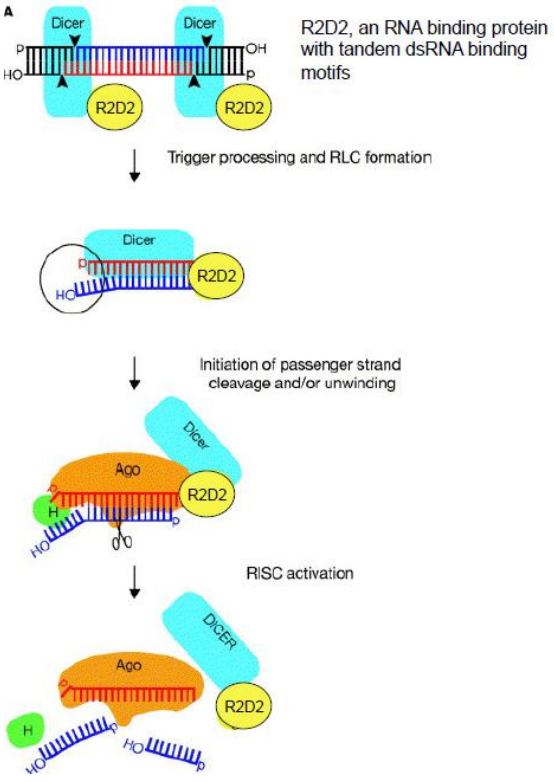
resembles RNase H which cuts RNA in RNA/DNA hybrids. The model for siRNA-guided cleavage of mRNA suggests, that the siRNA (yellow) first binds to the PAZ cleft with its 3' end. The mRNA (brown) comes in between the N-terminal and PAZ doains and out between the PAZ and middle domain. The active site in the PIWI domain (scissors) cleaves the mRNA opposite the middle of the siRNA guide.



The siRNA guide strand is bound at the 5' end by the PIWI and at the 3' end by the PAZ. The 5' phosphate is coordinated by conserved basic residues. mRNA targets are initially bound by the seed region of the siRNA and pairing is extended to the 3' end. The RNaseH fold hydrolyzes the target in a cation dependent manner. Slicer cleavage is measured from the 5' end of the siRNA. The product is released by an unknown mechanism and the enzyme recycles.

RISC Loading

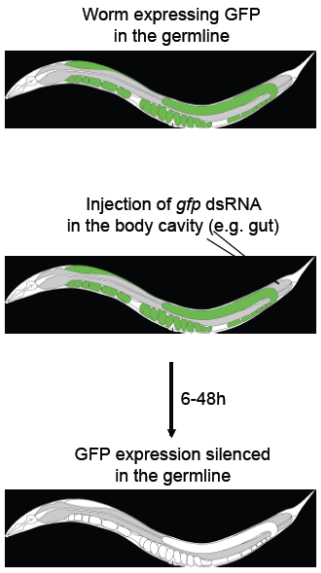
How are siRNA duplexes converted to single-chain forms and how is a correct (antisense) strand selected for loading onto the RISC? Asymmetric RISC formation is defined by the relative thermodynamic strength of the first four nucleotide-pairs of the 5' termini of an siRNA (a very GC-rich region of an siRNA is more stable than an AU-rich region, therefore, the more unstable region is unwound first).



Model: Dicer and R2D2 form a heterodimer and bind to ds siRNAs. This complex senses the free energy differences between the two 5' ends. R2D2 binds to the thermodynamically more stableend of an siRNA, and its binding is faciliated by the characteristic 5' phosphate. By blocking the non-preferred strand of the siRNA, R2D2 positions Dicer at the opposite end of the duplex. The

dsRNA binding domain of human Dicer recruits Ago2 via the PIWI domain. The thermodynamically non-preferred strand is eliminated after unwinding by Argonaute or an associated RNA helicase.

RNAi as a Genetic Tool

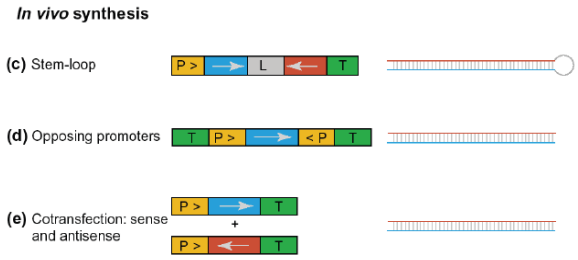


- Plants
- *C. Elegans*
- HeLa cells
- *Drosophila*

Silencing RNA can either be introduced as a long double strand or in already small siRNAs (21 nt). In *C. elegans*, there are 3 different methods to introduce dsRNA, either through feeding with bacteria expressing the RNA, by just soak the worms in a solution containing the RNA or by manual injection.

In cultured mammalian cells, long dsRNA does not work, as it induces an immune reaction (interferon response, protein kinase R (PKR) and oligo-adenylate synthesis (OAS)). Short, 21 nt long siRNAs however can be used (either by directly transfect synthetic siRNA or by transfecting plasmids expressing miRNA precursor-like stem-loops to generate siRNA in cells).

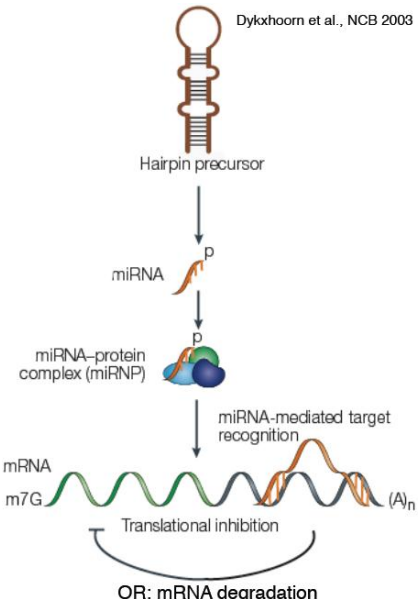
In vivo synthesis of siRNA: P-promoter, L-loop and T-Terminator.

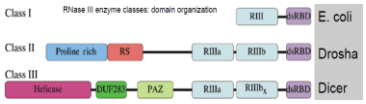


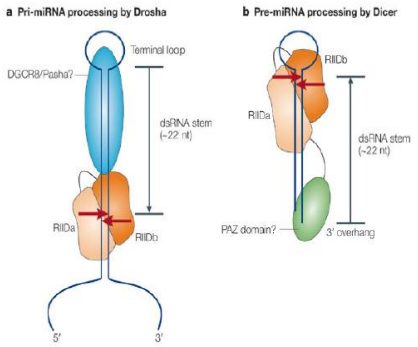
Summary:

- dsRNA is an important regulator of gene expression in many eukaryotes; it triggers a process referred to as RNA silencing or RNA interference
- RNA silencing mechanisms were first recognized as antiviral mechanisms that protect organisms from RNA viruses
- A key step is the processing of dsRNAs into short RNA duplexes of about 21 nt by RNase III-like enzyme Dicer
- An siRNA-containing effector complex is commonly referred to as a RISC, which directs the siRNA to its cognate target mRNA
- The ss siRNA in RISC guides sequence-specific degradation of complementary or near-complementary target mRNAs
- RISC contains a member of the Argonaute (Ago) protein family involved in mRNA cleavage

- RNAi is now being widely used as a genetic tool and also applied in genome-wide screens in many higher eukaryotic organisms





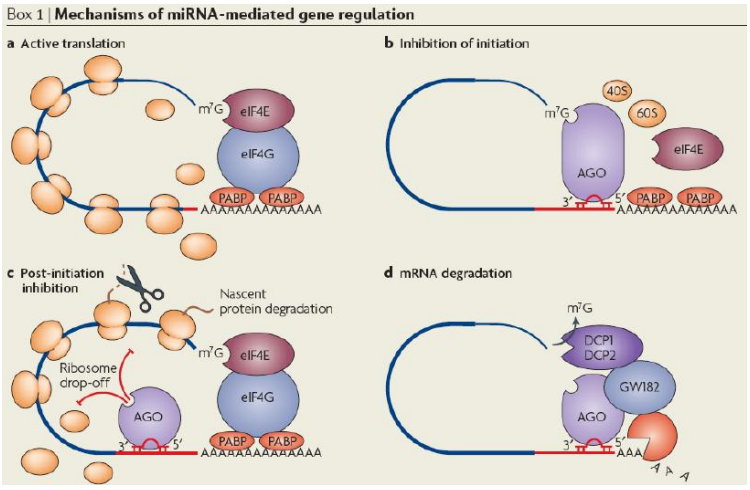


microRNAs-Novel Regulators of Gene Expression

miRNAs are encoded in the genome of many organisms. They constitute a class of noncoding, small RNAs that are phylogenetically widespread in invertebrates, vertebrates and plants. They control post-transcriptional gene expression in eukaryotes through inactivation or degradation of mRNAs is the cytoplasm.

Biogenesis: Transcription by Pol II leads to primary miRNAs (100s-1000s nt) which are further processed by Drosha to get pre-miRNAs (60 nt). These are then exported by Exp5 to get to the cytoplasm. There, the pre-miRNAs are processed by Dicer to yield ds miRNAs (22 nt, no hairpin anymore) which are then loaded into RISC to get single-stranded miRNAs (22 nt).

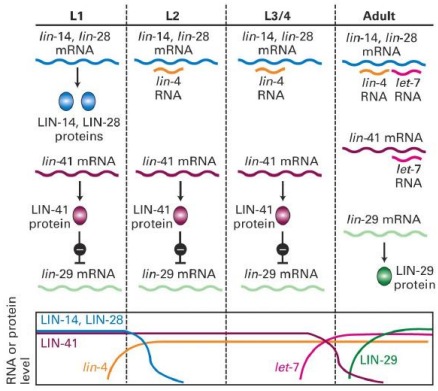
Today, several hundred miRNAs have been identified in nematodes, flies, plants, fish, mice and humans. Some miRNA genes are present in multiple copies on the same or on different chromosomes. The human genome may encode over 1000 miRNAs. Recent bio-informatic predictions suggest they may regulate the expression of up to one third or more of the expressed genes which makes miRNAs the biggest player in eukaryotic post-transcriptional gene regulation. Many are highly conserved between organisms. Some have a tissue specific expression and others are temporarilly regulated.

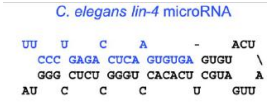


Mechanisms: Either the mRNA is cleaved speficically by a miRNA or translational repression takes place (specified by miRNAs). The mechanism of miRNA-mediated translational repression is still a

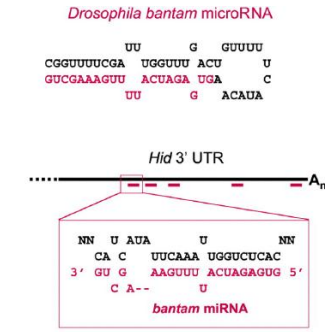
matter of controversy.

Example miRNAs in *C. elegans*: The first discovered miRNAs were let-7 and lin-4. They are expressed in a timely regualted fashion, stRNAs (small temporal). They control the expression of genes which are required for proper developmental timing.

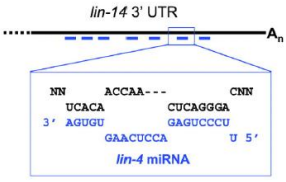


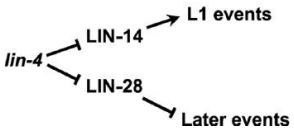


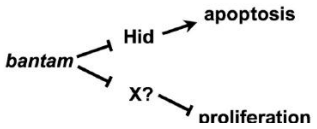
lin-4 controls lin-14 expression: lin-4 is partially complementary to 7 sites in the 3' UTR of lin-14 mRNA. lin-4 expressionat the end of the L1 larval stage results in downregulation of LIN-14 (and also LIN-28) protein synthesis. This controls the transition from L1 to later developmental events (LIN-14 is a nuclear protein required for L1 events, but not later).



Example bantam miRNA in *Drosophila*: bantam depletion produces small flies, bantam overexpression organ overgrowth. bantam functions to block apoptosis and has growth promoting activity, Hid is a proapoptotic gene.

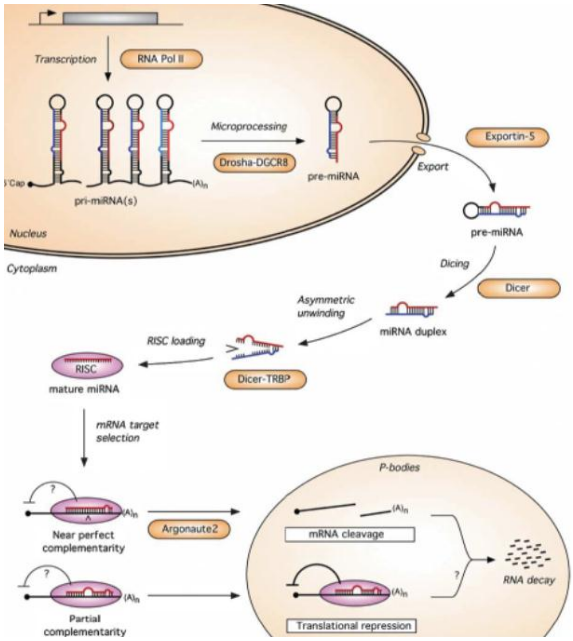






Summary:

- miRNAs are small RNAs that regulate the expression of complementary mRNAs; hundreds of miRNA genes have been found in diverse animals and many of these are phylogenetically conserved
- The maturation of small RNAs is a stepwise process catalysed by dsRNA-specific RNase III-type endonucleases, termed Drosha and Dicer



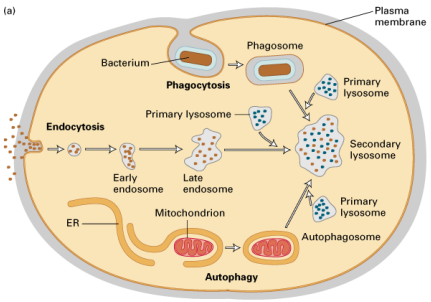
- miRNAs are transcribed as long primary transcripts, which are first processed by Drosha in the nucleus and the exported by Exp5 for cleavage by Dicer in the cytoplasm
- miRNAs can work by two models:
    - In plants, miRNAs base pair with mRNA targets by percise or nearly precise complementarity and direct cleavage and destruction of the target mRNA through a mechanism involving the RNAi machinery
    - Most animal miRNAs are imprecisely complementary to their mRNA targets and they inhibit protein synthesis through an unknown mechanism that preserves the stability of the mRNA target
- Roles for miRNAs have been identified in developmental timing, cell death, cell proliferation, haematopoiesis and patterning of the nervous system and many more are anticipated.
- Short notice: there is another way for RNA-mediated silencing, namely transcriptional silencing (chromatin remodelling, active chromatin, RNAs, histone methylation -> silent chromatin)

Table 1
Biological functions of miRNAs in animals and disease

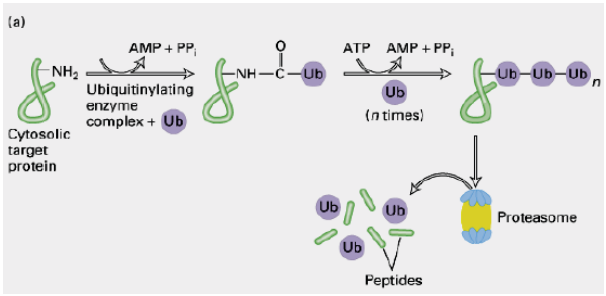
miRNA	Target(s)	Function(s)
Ceanorhabditis elegans		
lin-4	*lin-14*, *lin-28*	Early developmental timing
let-7	*lin-41*, *hbl-1*, *daf-12 pha-4*, *ras*	Late developmental timing
lsy-6	*cog-1*	Left/right neuronal asymmetry
miR-273	*die-1*	Left/right neuronal asymmetry
Drosophila melanogaster		
bantam	*hid*	Programmed cell death
miR-14	*Drice?*	Programmed cell death and fat metabolism
miR-7	Notch targets?	Notch signaling
Danio rerio		
miR-430	?	Brain morphogenesis
Mus musculus		
miR-196	*Hoxb8*	Developmental patterning
miR-181	?	Hematopoietic lineage differentiation
miR-1	*Hand2*	Cardiomyocyte differentiation and proliferation
miR-375	*Mtpn*	Insulin secretion
Human and other vertebrate cell lines		
miR-16	Several	AU-rich element-mediated mRNA instability
miR-32	Retrovirus *PFV-1*	Anitviral defense
miR-143	*Erk5?*	Adipocyte differentiation
SVmiRNAs	*SV40* viral mRNAs	Susceptibility to cytotoxic T cells
Cancer in humans		
miR-15-miR-16	?	Downregulated in B-cell chronic lymphocyte leukemia
miR-143, *miR-145*	?	Downregulated in colonic adenocarcinoma
miR-155/BIC	?	Upregulated in diffuse large B-cell lymphoma
let-7	*Ras?*	Downregulated in lung cell carcinoma
miR-17-92	?	Upregulated in B-cell lymphoma

Protein Degradation

Definition: Hydrolysis of peptide bonds catalyzed by proteases or peptidases. The biological roles are degradation of misfolded proteins, the adaption of the cell to environmental changes, cell division and the elimination of invaders. Proteases are highly specific or regulated or they are compartementalized (and unspecific).



- Proteolysis by vacuolar/lysosomal proteases: Serves the degradation of extracellular and membrane proteins, the degradation of organelles and ribosomes and is the primary site for constitutive protein degradation
- Ubiqiutin/proteasome pathway: Serves mainly the degradation of endogenous proteins, e.g. transcription factors, cyclins, proteins encoded by viruses and other intracelular parasites, proteins that are folded incorrectly because of translation errors, because they are encoded by faulty genes or because they have been damaged by other molecules in the cytosol



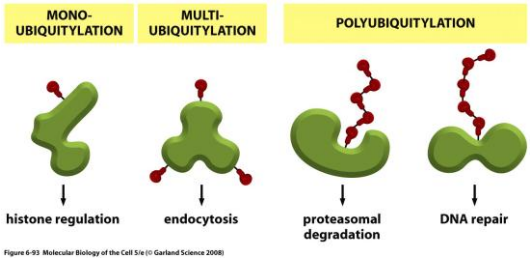


Figure 6-93 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Ubiquitin/proteasome pathway: Ubiquitin chains serve as degradation signal. Ubiquitin is 76 aa long and a highly conserved globular protein. The α-carboxyl group of the C-terminal Gly is attached to the ε-amino group of internal Lys of target proteins.

The C-terminus of ubiquitin is initially activated through its high-energy thioester linkage to a cysteine side chain on the E1 protein. This reaction requires ATP, and it proceeds via a covalent AMP-ubiquitin intermediate. The activated ubiquitin on E1, also known as the ubiquitin-activating enzyme, is then transferred to the cysteines on a set of E2 molecules. These exist as complexes with an even larger family of E3 molecules. In a mammalian cell there are several hundred distinct E2-E3 complexes, many of which recognize a

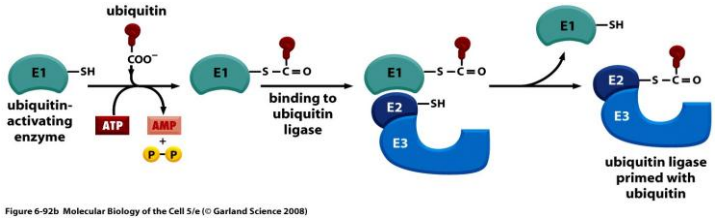


Figure 6-92b Molecular Biology of the Cell 5/e (© Garland Science 2008)

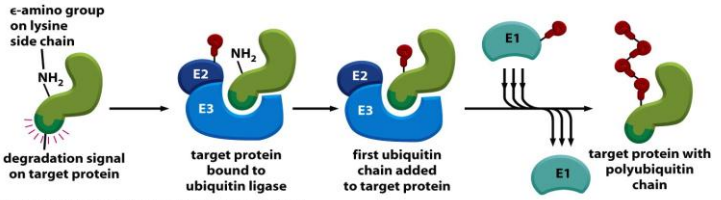
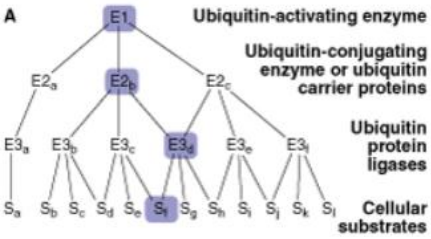
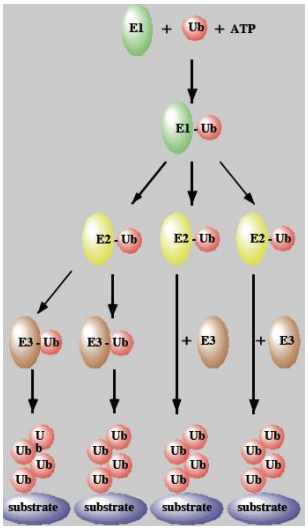


Figure 6-92c Molecular Biology of the Cell 5/e (© Garland Science 2008)

specific degradation signal on target proteins by means of the E3 component. The E2s are called ubiquitin-conjugating enzymes. The E3s have been referred to traditionally as ubiquitin ligases, but it is more accurate to reserve this name for the functional E2-E3 complex. Substrate recognition: Constitutive substrates have degradation signals and occasianal substrates are e.g. misfolded proteins. Structural motifs that target proteins for ubiquitination are:



- N-terminal residue (F, L, W, Y, R, K and H), N-end rule
- Phosphorylation (e.g. transcription factors like β-catenin)
- PEST elements (regions rich in P, S, T and E) often linked to phosphorylation
- Destruction boxes (e.g. cyclins) R(A/T)(A)L(G)x(I/V)(G/T)(N)

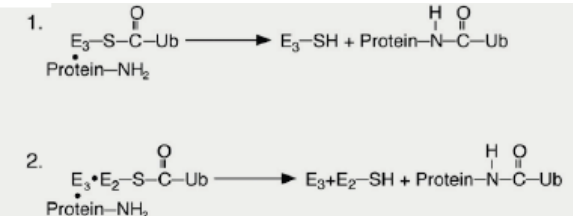


Activation reaction:

- E1 + ATP + Ub -> E1.Ub-AMP + Pp_i
- E1.Ub-AMP + Ub -> E1-s-co-Ub.AMP-Ub

Transfer reaction:

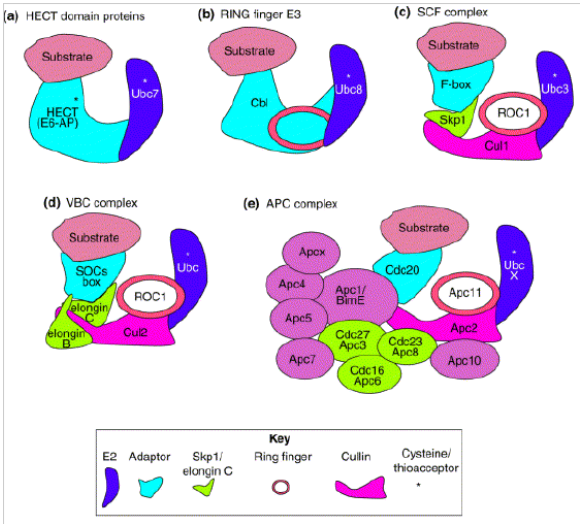
- E1-s-co-Ub.AMP-Ub + E2-SH -> E2-s-co-Ub + E1.AMP-Ub
- E2-s-co-Ub + Protein-NH_2 -> E2-SH + Protein-NH-CO-Ub

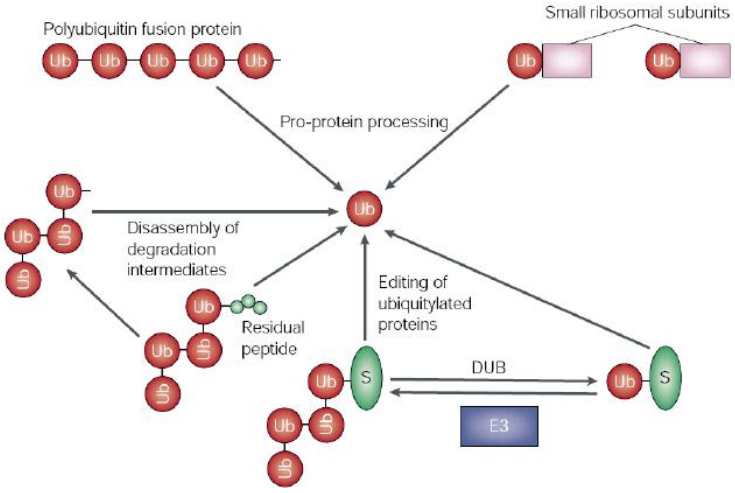


Comparison of E3 ubiquitin ligases:

- HECT domain proteins (homologous to E6-AP C terminus): within the HECT domain of E6, a conserved cysteine forms a thioester with ubiquitin
- RING (really interesting new gene) finger E3 ubiquitin ligases: single-polypeptide RING-finger E3 enzymes, SCF, VBC and APC, RING motif consists of a defined pattern of Cys and His residues that coordinate two zinc ions; it is engaged in ubiquitin ligation through the recruitment and positioning of the E2 enzyme

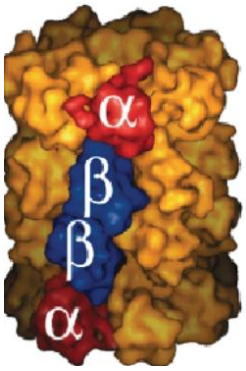
DUB: Deubiquitinylating enzymes, counteracting ubiquitinylation. After the degradation of protein substrates, ubiquitin mus be freed from residual peptides and disassembled. DUBs also reverse the activity of E3s, sequentially removing ubiquitin from substrates. This might occur in specific cellular location where ubiquitylation is occuring and at the proteasome.

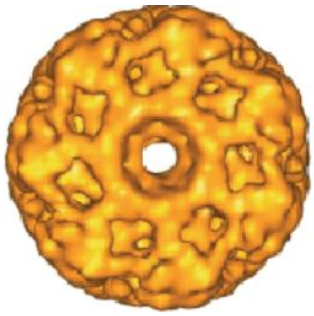




Proteasome: Structure and Function

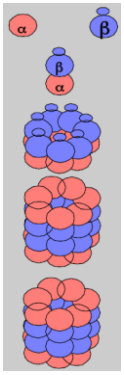
20S Proteasome: In pro- and eukaryotes, in higher eukaryotes up to 1% of total cellular protein, conserved, 700 kDa, consists of 4 rings, each ring having 7 subunits, cylinder form. The outer rings are α-subunits and the outer ones β-subunits. 11.5 x 15nm.





The β-subunits contain the active site, Thr1, Glu17 and Lys33. The fact that the active site is shielded (because it is in the middle of the complex) guarantees that only proteins that really should be degraded are processed.

Lactacystin, peptidylvinylsulfonsäure and peptidylvinylborsäure are covalent and irreversible proteasome inhibitors, whereas peptide aldehydes are reversible and noncovalent inhibitors.



Assembly: Propeptide -> heterodimer -> half proteasome -> preholoproteasome (autoproteolytic maturation) -> holoproteasome.

The 20S core structure is conserved from Archaea to Eukaryotes. In higher eukaryotes, there are 7 different catalytic β-subunits, and 3 have catalytic activity. β1 cleaves behind acidic aa like Glu (caspase-like), β2 cleaves behind basic aa like Lys or Arg (trypsin-like) and β5 cleaves behind bulky hydrophobic aa like Phe, Tyr or Trp (chymotrypsin-like). Specificity is determined

by aa 45 of active β-subunit and by the inactive neighbouring β-subunits.

Proteasomal degradation leads to an average of 7-8 aa long peptides (but peptides can be between 4-30 aa long). Cleavage is processive, the active sites are under allosteric regulation (bite and chew model).

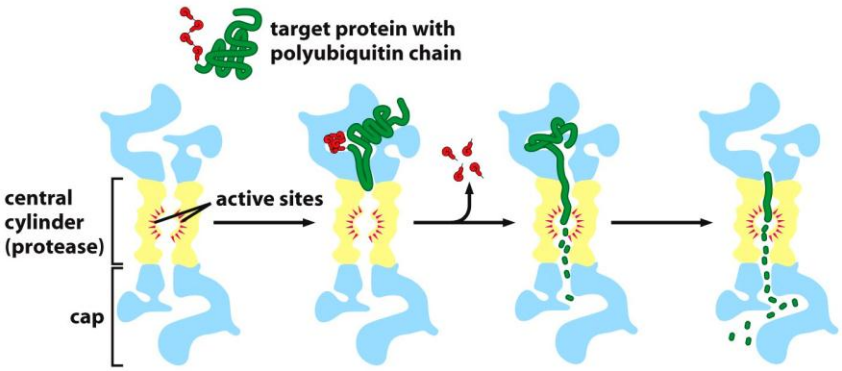


Figure 6-90 Molecular Biology of the Cell 5/e (© Garland Science 2008)

The proteasome cap recognizes a substrate protein, in this case marked by a polyubiquitin chain, and subsequently translocates it into the proteasome core, where it is digested. At an early stage, the ubiquitin is cleaved from the substrate protein and is recycled. Translocation into the core of the proteasome is mediated by a ring of ATP-dependent proteins that unfold the substrate protein as it is threaded through the ring and into the proteasome core. Access to the central cavitity is restricted to unfolded proteins.

Delivery to the central cavity is achieved by AAA-family (ATPases associated with various cellular activities). They have a conserved ATPase domain of about 250aa, chaperone activity for unfolding proteins and are responsible for assembly/disassembly of protein complexes. They form characteristic ring like oligomers. AAA-family ATPases are components of the regulatory complex of the proteasome. Their proposed function is the one of reverse chaperones, to unfold and resolve proteins and protein aggregates, respectively, to allow their entrance into the 20S proteasome.

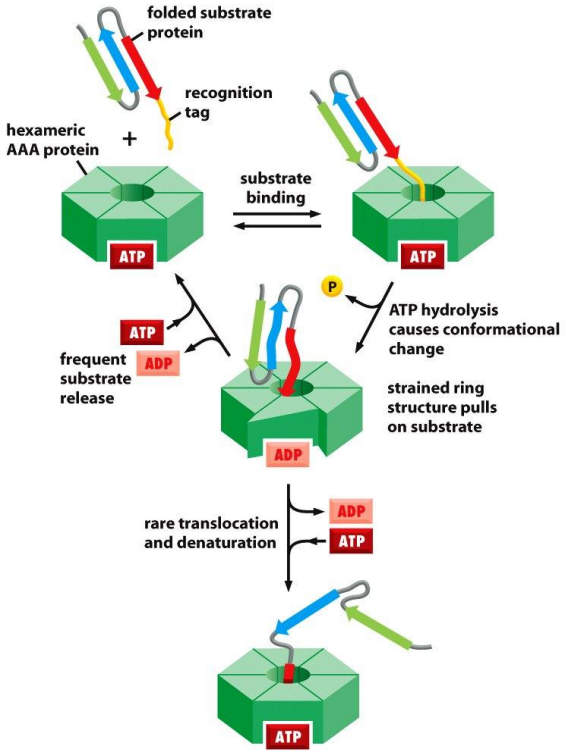
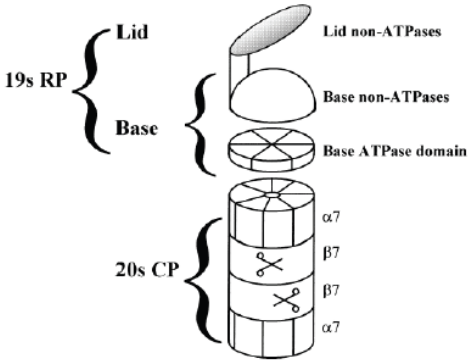


Figure 6-91b Molecular Biology of the Cell 5/e (© Garland Science 2008)

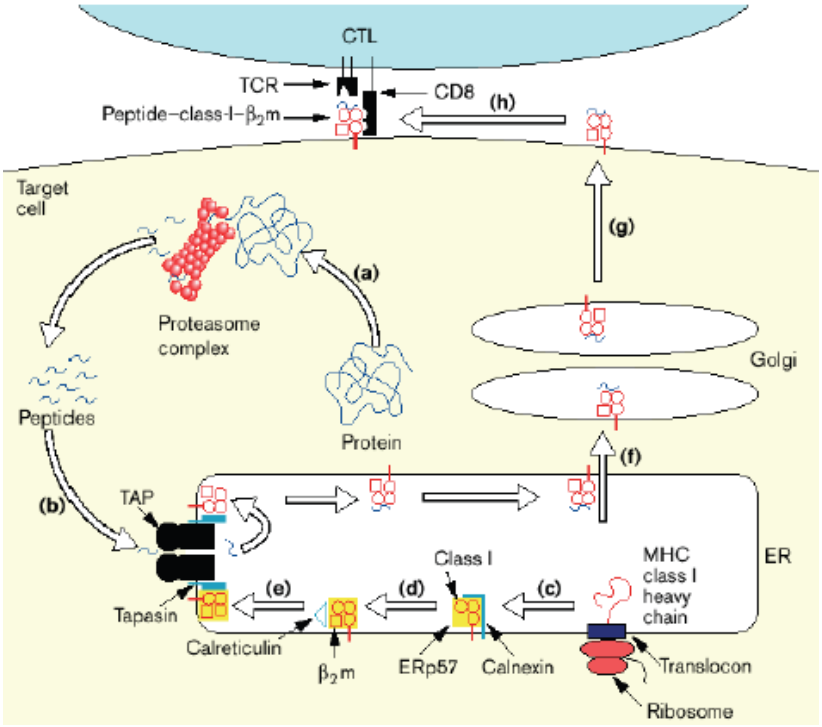
Mechanism: The ATP-bound form of a hexameric ring of AAA proteins binds a folded substrate protein that has been marked for unfolding by a recognition tag such as a polyubiquitin chain or the peptide added to mark incompletely synthesized proteins. A conformational change, made irreversible by ATP hydrolysis, pulls the substrate, which is being tuggedupon, can partially unfold and enter further into the pore or it can maintain its structure and dissociate. Very stable

protein substrates may require hundreds of cycles of ATP hydrolysis and dissociation before they are successfully pulled into the AAA ring. Onde unfolded, the substrate protein moves relatively quickly through the pore by successive rounds of ATP hydrolysis.

The 19S regulatory particle: Recognizes poly-ubiquitinated degradation substrates, has isopeptidase activity (recycling of ubiquitin), unfolds substrates and gates the 20S proteasome.

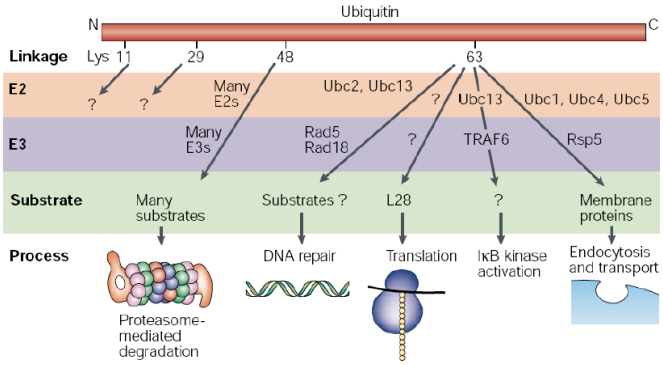


The proteasome is also involved in antigene presentation:



Interferon-γ activates both the core (LMP2, LMP7 and MECL-1, the catalytic β-subunits, production of longer peptides better suited for presentation) and the lid (PA28 activator, 11S regulatory complex, PA28α and PA28β, 200kDA hexamer, does not recognize ubiqiutin proteins). This leads to the production of antigenic peptides.

Ubiquitin has several lysine residues and can form multi-ubiquitin chains linked through positions 11, 29, 48 and 63. The functions of Lys11 and Lys29-linked chains are unknown, Lys48-linked chains target proteins to the proteasome but might have other functions, and Lys63-linked chains have a range of fates.



Signalling from the ER to the Cell Nucleus

ER: Serves as the entry for secretory and membrane proteins into the secretory pathway, synthesizes cholesterol and other lipids, folds and glycosylates secretory and membrane proteins. ER stress can be induced by glucose starvation, cholesterol starvation or viral infection.

1. Unfolded Protein Response (UPR)

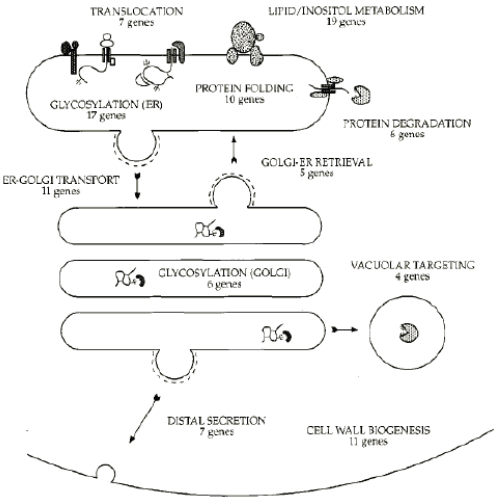
Induced by a variety of cellular insults that result in the accumulation of proteins in the ER, e.g. glycosylation inhibitors, reducing agents, misfolded mutant proteins in the ER, overproduction of normal secretory proteins, overproduction of ER membrane proteins or block of ER exit.

Table 1 ER chaperones regulated by the UPR

Mammalian	Yeast	Function
BiP (GRP78)	KAR2	hsp70 homologue
PDI (Erp59)	PDI	protein disulfide isomerase
GRP170		hsp70-like
GRP94		hsp90-like
ERP72		PDI-like
GRP58		PDI-like
	FKBP12	prolyl-*cis*-*trans*-isomerase
	LHS1	BiP-like
	EUG1	PDI-like
	ERO1	oxidoreductase

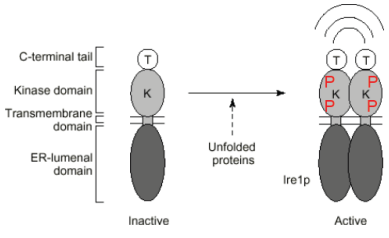
GPRs (glucose-regulated proteins): Polypeptides that are highly induced upon glucose starvation:

UPR-induced genes can be identified by a functional genomics approach in yeast. UPR is induced by either DTT or tunicamycin, RNA is prepared and hybridized and analysed (microarray, 65'000 different oligonucleotides, yeast about 6'200 genes, poly(A) RNA extracted, labeled and hybridized). Results did show that more than 380 genes are upregulated by the UPR.



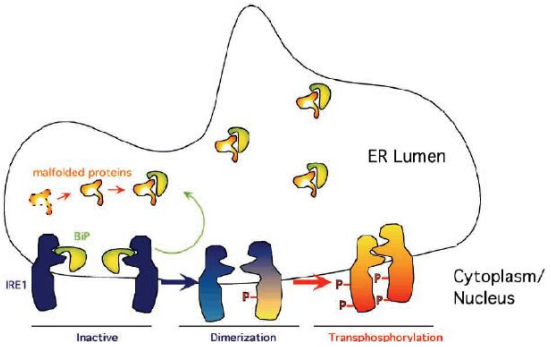
Components mediating the UPR are IRE1 (transmembrane kinase/endonuclease), HAC1 (transcription factor) and RLG1 (tRNA ligase). These components lead to a signal transduction cascade.

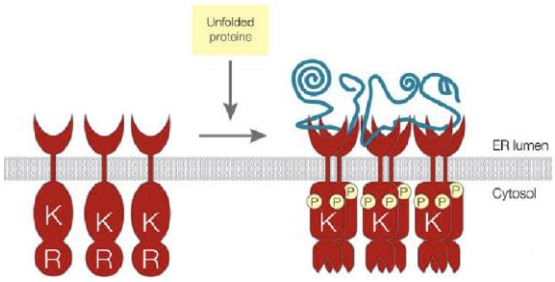
Ire1p: This transmembrane kinase is activated by phosphorylation and oligomerization. The C-terminus is cytoplasmic (nuclear) and has a Ser/Thr kinase domain. It resembles ribonucleases. The N-terminus is in the ER lumen and a sensor to detect the accumulation of unfolded proteins.



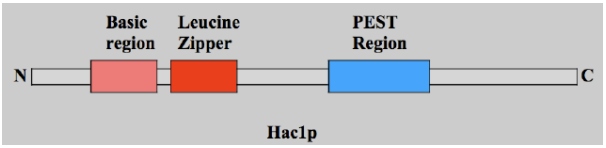
There are two models for the activation of Ire1p by unfolded proteins. The first model is BiP. BiP binds to the lumenal domain of IRE1 and keeps it in an inactive, monomeric conformation in the presence of little unfolded proteins. If malfolded proteins accumulate, BiP is released from the lumenal domain, IRE1 undergoes dimerization and transautophosphorylation and its effector functions are activated.

The second model suggests that binding of unfolded proteins need to the phosphorylation and activation of Ire1p.

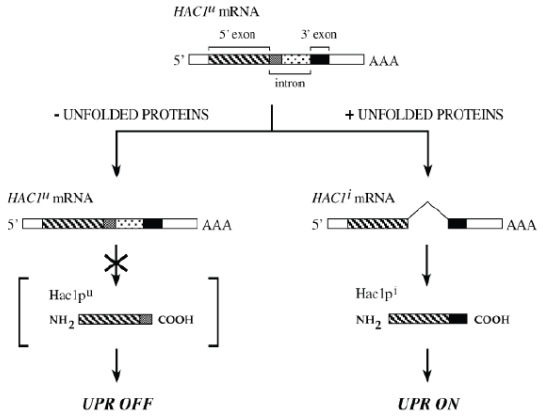




Transcriptional regulation of UPR-induced genes: The UPRE is a 22bp element that is conserved in the promoters of all UPR target genes. Insertion of the UPRE into genes normally not controlled by the UPR suddenly get controlled by the UPR. The UPRE can confer induction by UPR to a heterologous promoter. The UPRE is characterized by a short E-boc like palindromic sequences (CANCNTG). An E-box is a binding site of bHLH transcription factors in the promoters of many developmentally controlled genes. The Hac1p is a bZIP transcription factor. It also contains a PEST-region, which is rich in the amino acids Pro, Glu, Ser, Thr and Asp. These regions destabilize proteins by targeting them into the ubiquitin-dependent proteolysis pathway. The half-life of Hac1p is 2 min.



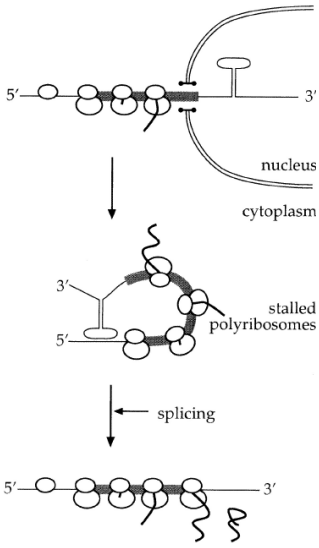
UPRE-dependent transcription is regulated by the abundance of Hac1p. In the absence of unfolded proteins, no Hac1p is detected in yeast cells. However, Hac1 mRNA is always transcribed and present.



This means that Hac1p expression is regulated post-transcriptionally. This is because of the presence of two different mRNA forms, $HAC1^u$ and $HAC1^i$. $HAC1^u$ is present in the absence of unfolded proteins and cannot be translated. $HAC1^i$ is a splicing product of $HAC1^u$ (splicing is tightly regulated, processing occurs only when misfolded proteins accumulate in the ER lumen) and can be translated into Hac1p.

Interestingly, unlike other pre-mRNAs, $HAC1^u$ mRNA is stable and exported from the nucleus. Although no protein is produced from this message, it co-sediments with polyribosomes. The intron in the $HAC1^u$

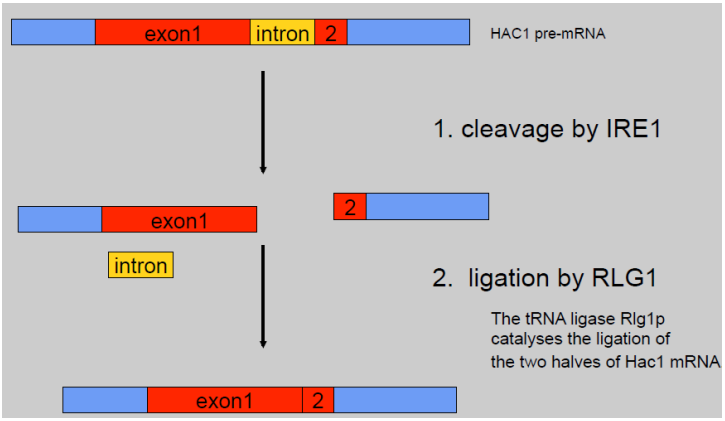
mRNA was shown to cause the translational block. There are two different models, the first one proposes that the intron can also work from some distance (indirect) whereas the direct model states that the intron stops the ribosome when it arrives there. This intron can still block translation when transplanted to the 3' UPR of an mRNA, more pointing to the indirect model.



The current model suggests that base-pairing interactions between the intron and the 5' untranslated region is required and sufficient to block mRNA translation. Unspliced HAC1 mRNA is stable and associated with polyribosomes, yet does not produce protein. This polysomal, cytoplasmic pool of HAC1 mRNA is a substrate for splicing.

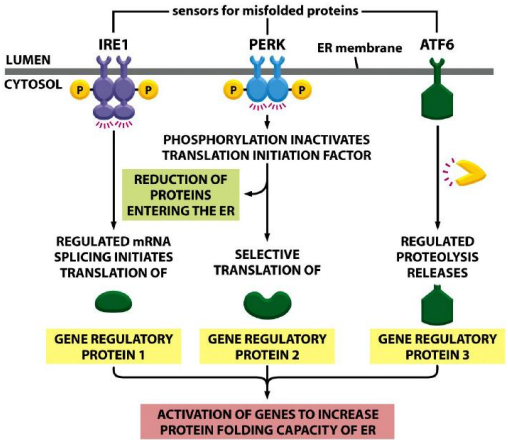
Splicing of HAC1 mRNA is not catalyzed by the spliceosome bus uses a unique machinery composed of two proteins (non-conventional splicing):

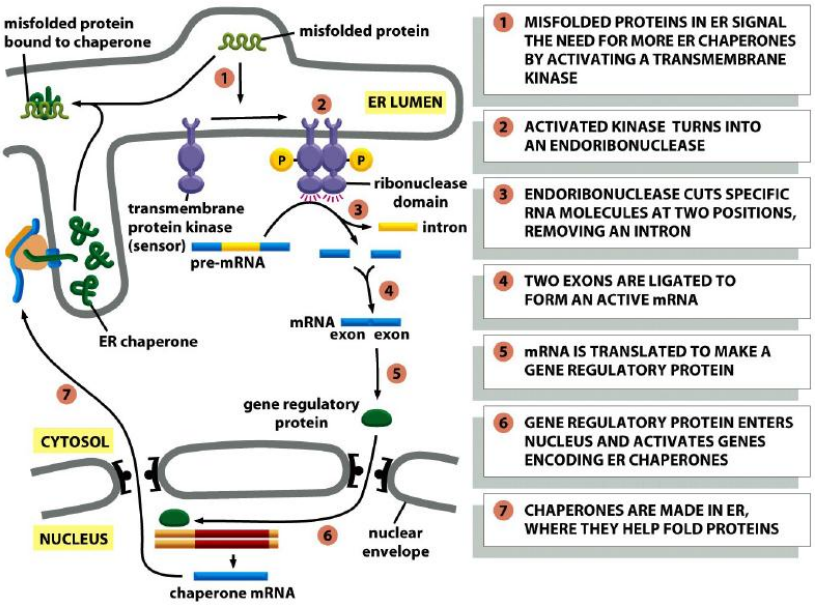
- Bifunctional Ire1p transmembrane kinase/endonuclease (cleavage of mRNA, directly demonstrated in an experiment, where a recombinant fusion protein consisting of the C-terminal half of Ire1p (contains kinase and Rnase domain and GST (glutathione-transferase, dimer) was expressed; GST-Ire1p autophosphorylates and cleaves $HAC1^{u}$ mRNA *in vitro,* it is assumed that Ire1p binds to Hac mRNA as a dimer with each monomer recognizing one of the very similar stem-loop RNA structures present at the 5' and 3' splice sites of the RNA): Resembles RnaseL, is a soluble enzyme, activated by 2'-5' oligoadenylates, oligomerization, function in cellular defence against viruses



- tRNA ligase Rlg1p (ligation)

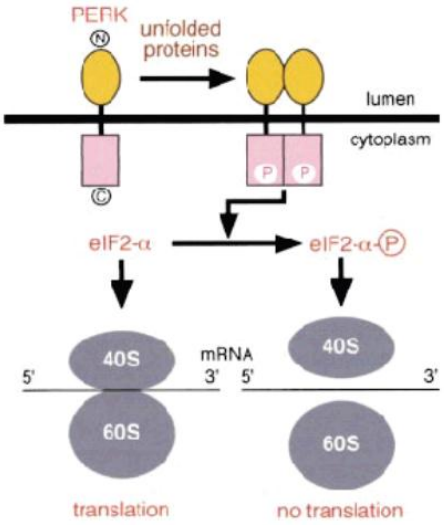
Summary:



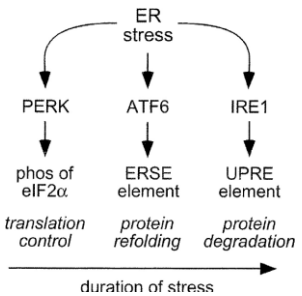


In higher eukaryotes, ER stress can be handled in three different ways. After short exposure to stress, the PERK pathway is activated. Next the ATF6 pathway starts and after long stress the IRE1 pathway is activated.

PERK (PKR-like ER kinase) is a transmembrane kinase with homology to IRE in its lumenal domain. It only exists in metazoans and phosphorylates eIF2-α thereby inhibiting translational initiation. This inhibition of translation is transient.



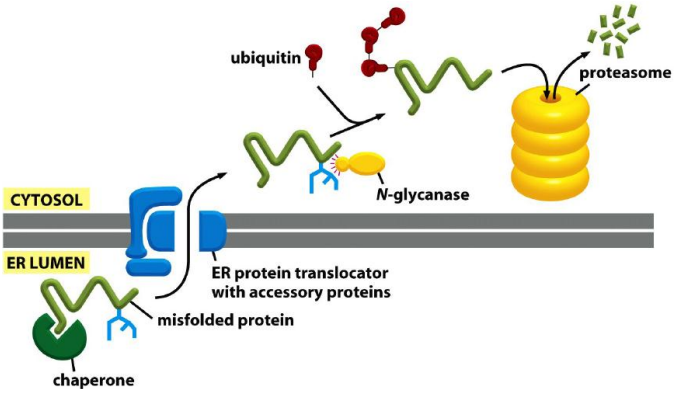
The three braches of the metazoan UPR sense the levels of unfolded protein in the lumen of the ER. PERK reduces general translation in cells, thereby reducing the protein influx into the ER. All three activate cognate bZip transcription factor via regulation of translational control, regulated proteolysis and regulated mRNA splicing. In mammalian cells, ATF6 (also activated by ATF4) upregulates expression of XBP1 mRNA (by binding ERSE, refolding proteins, chaperone etc.). The output of the transcription factors is integrated through their combinatorial action on UPR target genes, whose products increase the protein folding capacity of the cell and hence help the system to reestablish homeostasis. If homeostasis in ER protein folding connot be reached, cells undergo apoptosis.



Tripartite management of unfolded proteins in the ER:

1. Translational attenuation

2. Transcriptional activation (UPR)
3. ERAD (ER-associated degradation, degradation of proteins that cannot get fixed anymore)

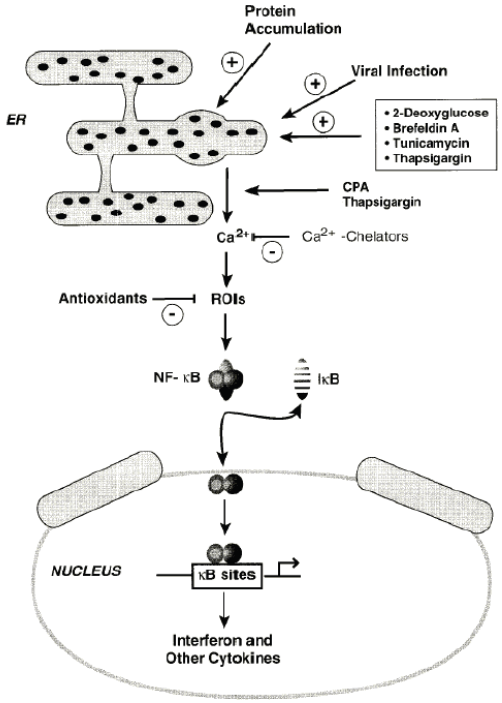


2. ER-overload Response (EOR)

The EOR is activated by accumulation of membrane proteins in the ER in response to viral infection or genetic diseases, e.g. cystic fibrosis. There, a mutation in a gene prevents the protein to be transported to the cytosol, this leads to ER overload response and activation of NF-κB. Ca^{2+} can no longer be kept (because of too many proteins in the ER) which affects the mitochondria, reactive oxygen species etc. Summarized, ER overload activates the transcription factor NF-κB.

TABLE 2. *ER stress conditions that activate nuclear factor κB*

ER overload by protein overexpression
Hepatitis B virus truncated middle HB surface antigen
Influenza virus hemagglutinin
Immunoglobulin μ-heavy chain
MHC class I
Adenovirus E3/19K
EPO receptor
Drugs perturbing ER function
Tunicamycin
2-Deoxyglucose
Monensin
Brefeldin A
Thapsigargin
Cyclopiazonic acid



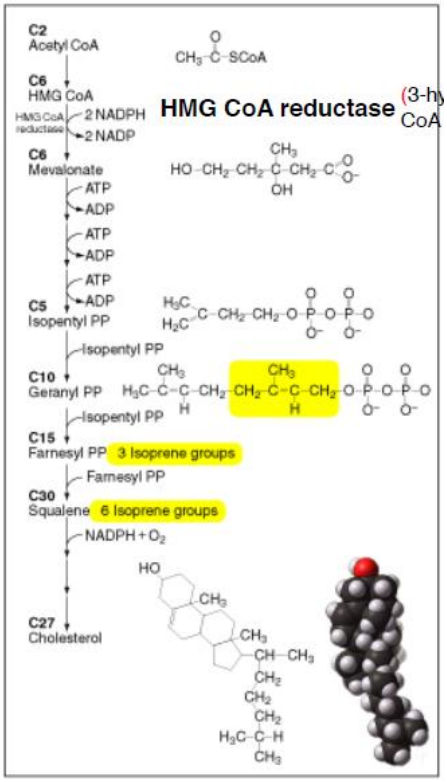
Summary:

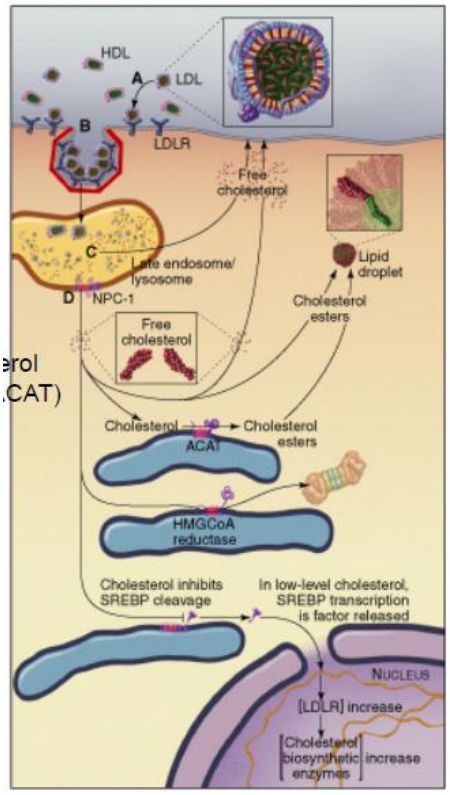
- Accumulation of misfolded proteins in the ER trigger two important signaling pathways, the unfolded protein response (UPR) and ER overload response (EOR)
- UPR uses a novel signaling pathway involving a transmembrane kinase/endonuclease Ire1p and the TK factor Hac1p
- Activation of Ire1p is triggered by the binding of misfolded proteins, Ire1p multimerizes, auto-phosphorylates and becomes active as an endonuclease to splice Hac1 mRNA
- Hac1p is expressed and induces genes to handle ER stress
- In mammalian cells, UPR is supported by PERK and ATF6
- The ERAD pathway helps to clear unfolded proteins from the ER
- EOR initiaties activation of NF-κB which in turn leads to the expression of e.g. cytokine genes to trigger an immune response (defense against viral infection)

3. ER-nucleus Signalling by SREBPs - Cholesterol Homeostasis

- *De novo* synthesis
- Through nutrition: Uptake of cholesterol-rich low-density lipoproteins (LDLs) by LDL receptors

The biogenesis of cholesterol involves acetyl-CoA, a C27 steroid-like molecule and the HMG CoA reductase (rate-limiting step). In LDL, cholesterol is stored as ester, receptors bind cholesterol. The ER controls cholesterol homeostasis.

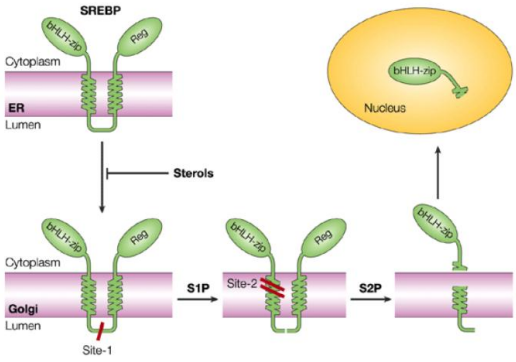


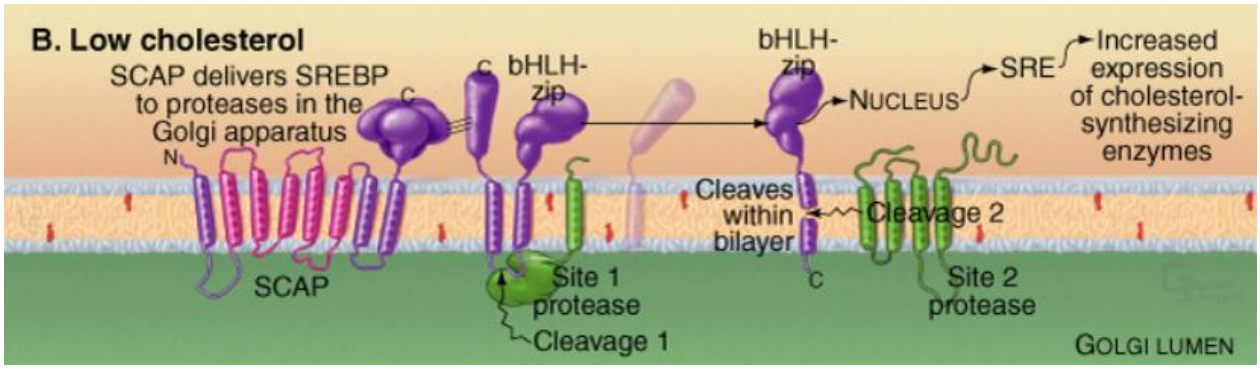


Both cholesterol uptake and synthesis are regulated at the transcriptional level (e.g. HMG-CoA and LDL receptor). The important element is the SRE (sterol regulatory element). It is 10bp long and in promoters of sterol-regulated genes. SREBP-1/SREBP-2 (sterol regulatory element binding protein) are bHLH transcription factors, transmembrane proteins of the ER/nuclear membrane (two transmembrane segments). The N-terminal portion is the transcriptional activation domain and the C-terminus interacts with the SCAP protein (SREBP activating protein).

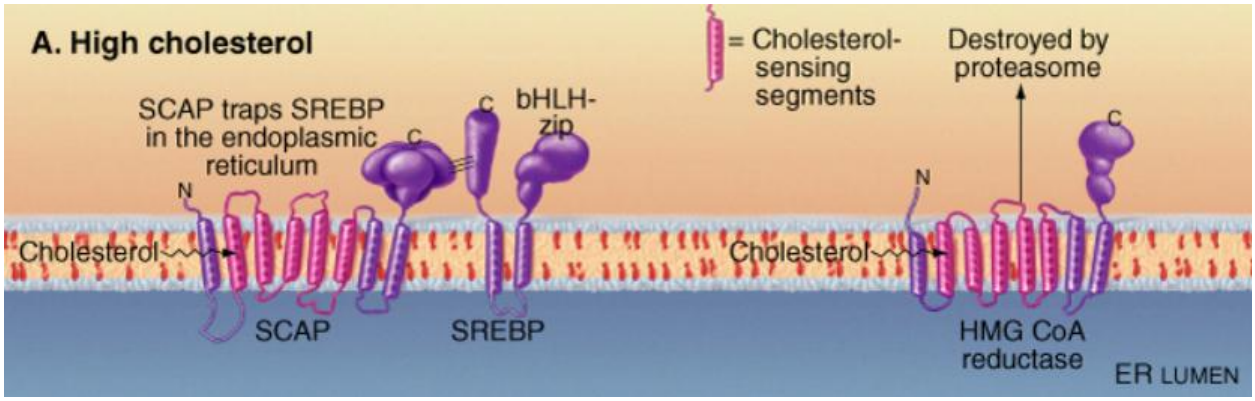
In low cholesterol conditions, upon sterol depletion, SREBP is delivered to the Golgi, where it is cleaved by two sequential proteolytic steps:

1. In the middle of the lumenal domain (regulated by sterol) by Site 1 protease (S1P)
2. Middle of the first transmembrane region by Site 2 protease (S2P)





In high cholesterol environment, SREBP is retained in the ER by SCAP and HMG CoA reductase is destroyed by the proteasome.



The SCAP protein may serve as the sterol sensor and is retained in the ER by itsinteraction with INSIG-1. In need for sterols, SCAP is retained from INSIG-1 and shuttles with SREBP from the ER to the Golgi.

Summary:

- Cholesterol homeostasis is critical to human health
- Levels of cholesterol are tightly controlled at the level of synthesis (HMG-CoA reductase) and esterification by the acyl CoA, cholesterol acyltransferase (ACAT)
- If the levels of free cholesterol increase substantially, secondary responses are initiated such as degradation of HMG-CoA reductase
- If the levels of free cholesterol decrease, SREBP is activated by intramembrane proteolysis controlled by SCAP
- SREBP is a TK factor that controls synthesis of HMG-CoA reductase and LDL receptor