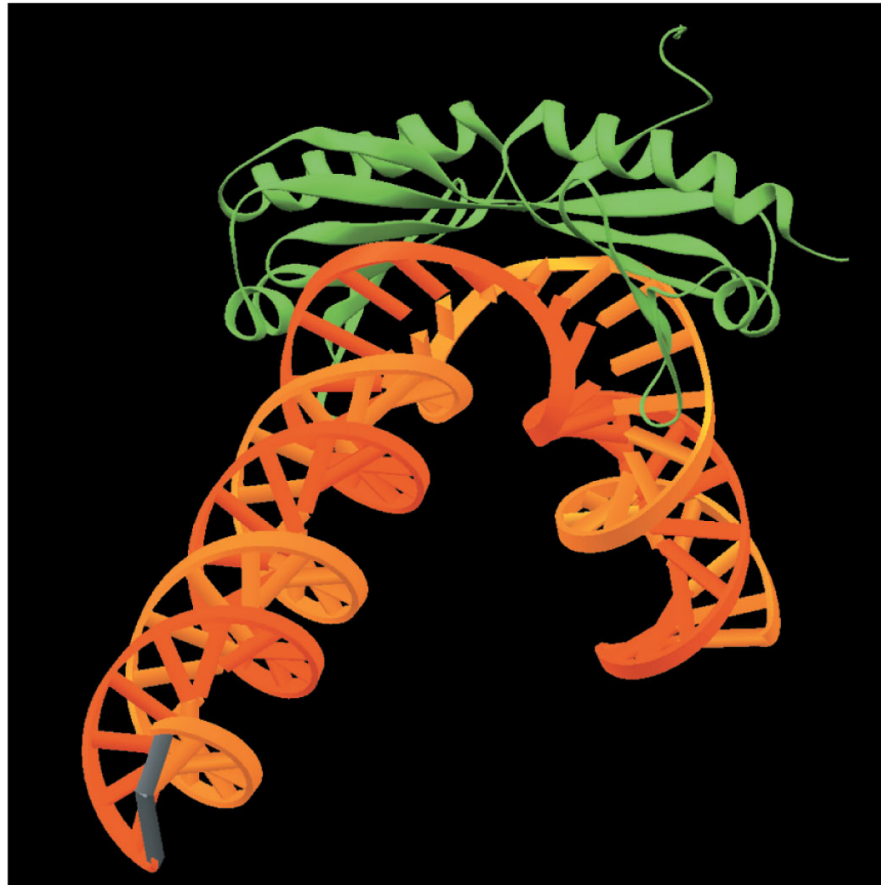


# Gene Expression at the molecular level: Transcription

Wie wird die genetische Information in Polypeptidketten umgesetzt?

- 1: Bei der Transkription wird DNA in RNA umgeschrieben.
- 2: Bei der Translation wird die in RNA gespeicherte Information in Proteine übersetzt



Bei der Transkription wird ein DNA Strang in RNA umgeschrieben. Da RNA (wie DNA) nur in 5' → 3' Richtung synthetisiert werden kann, muss die DNA in 3' – 5' Richtung gelesen werden.

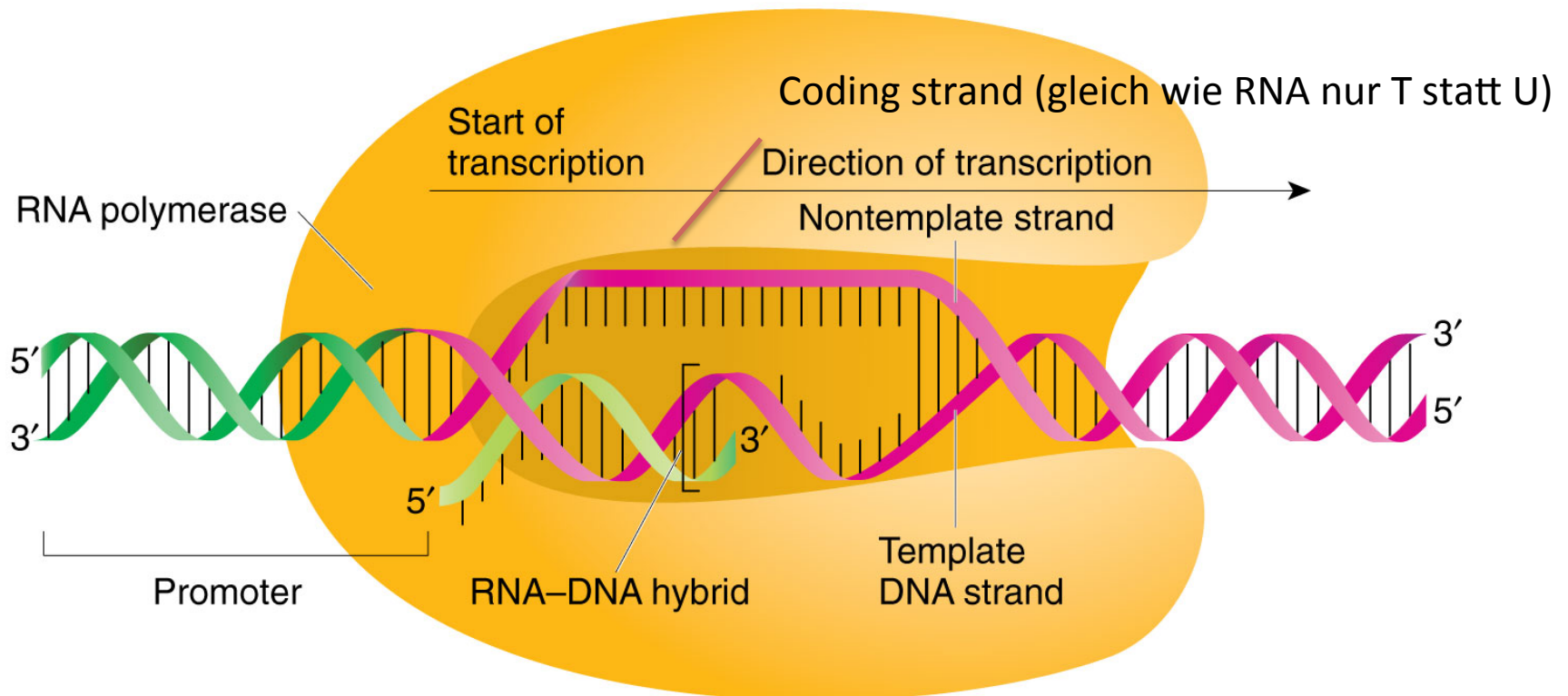
D.h. **der 3'-5' Strang ist der Template Strang!**

Der **5'-3' Strang ist dem RNA Strang nahezu ident (Coding strand).**

**Sequenzinformation lässt sich daher von diesem Strang leicht ablesen.**

Die Synthese neuer RNA wird durch eine RNA Polymerase katalysiert.

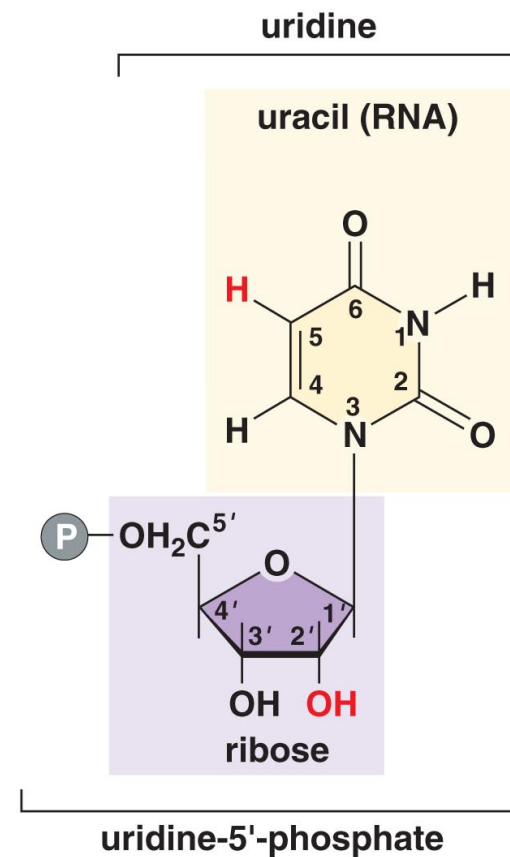
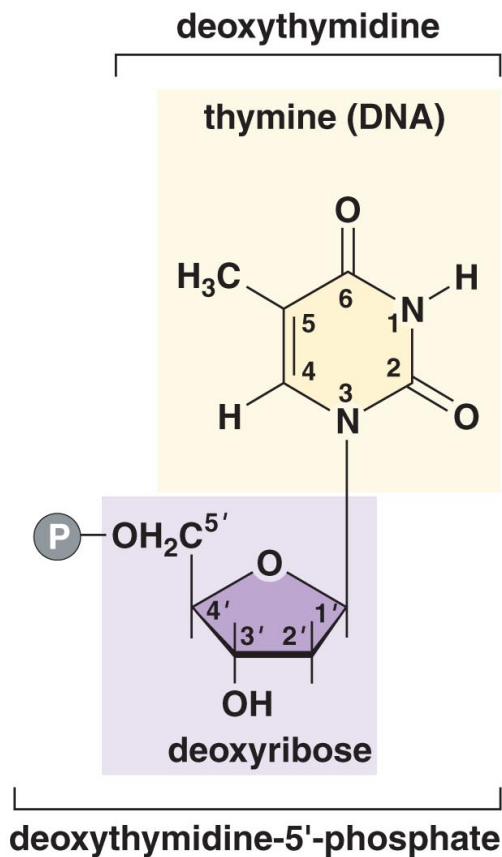
Um den Startpunkt der RNA Synthese zu definieren, werden regulatorische Sequenzen benötigt.



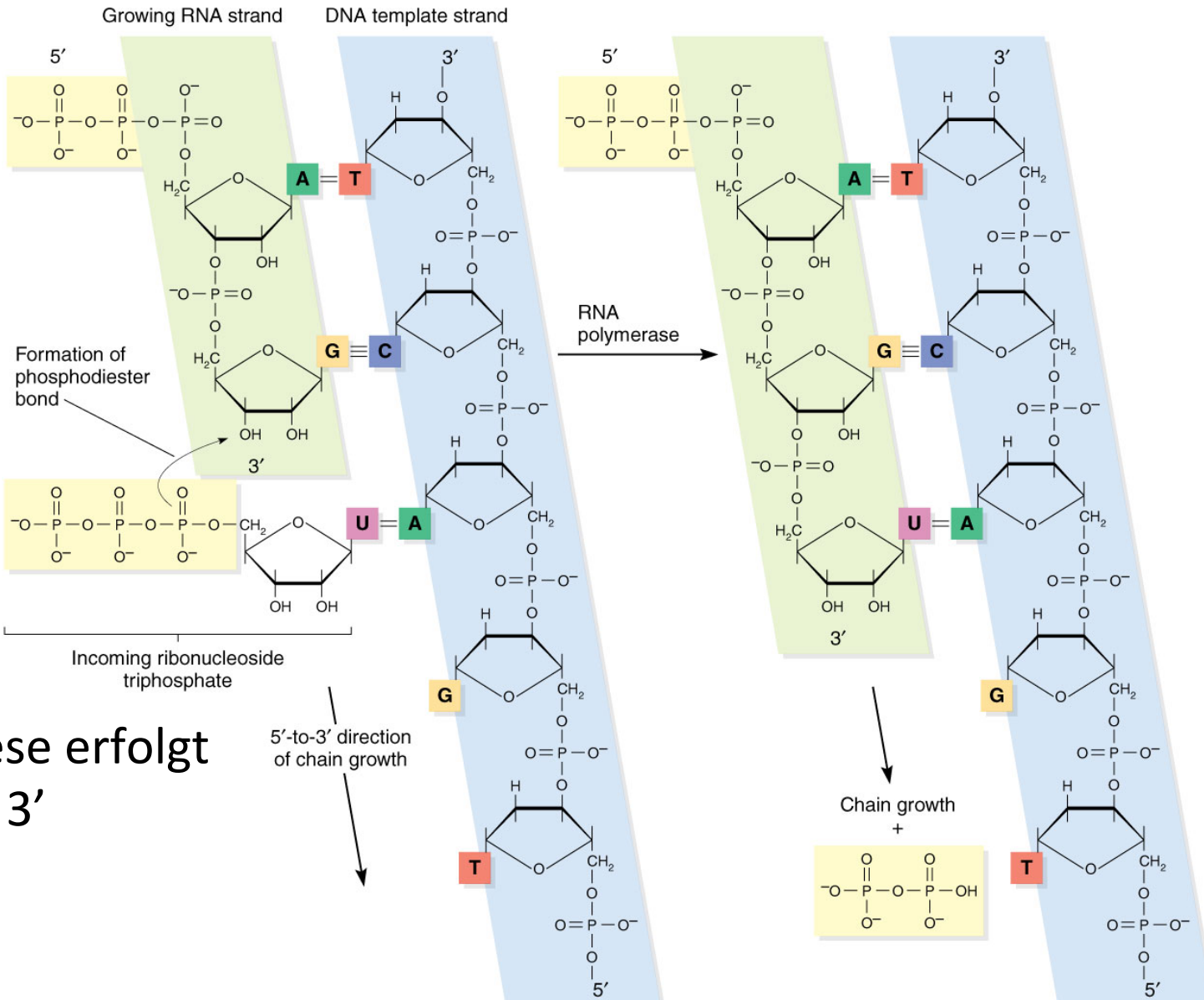
# Unterschiede zwischen RNA und DNA

## Ribose statt Desoxyribose

## Uracil statt Thymin



# DNA beinhaltet Desoxyribose, RNA Ribose. Die Basen der RNA sind Adenosin, Guanosin, Cytosin und Uracil (statt Thymidin)



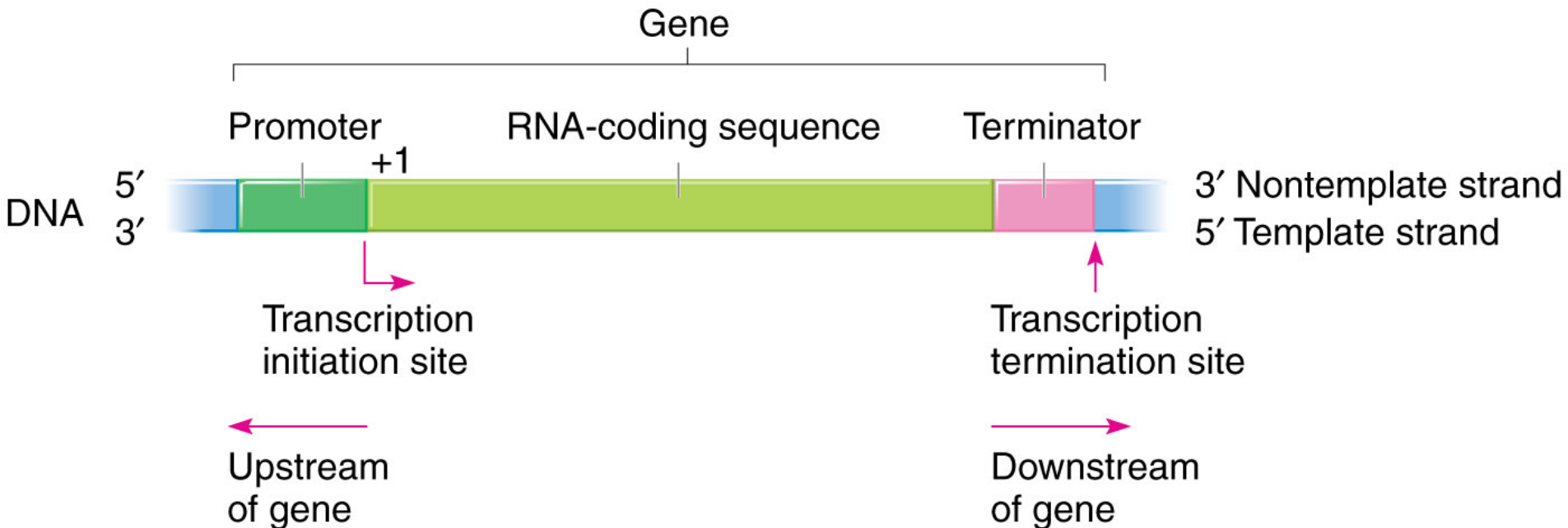
RNA-Synthese erfolgt von 5' nach 3'

Gene auf Chromosomen können in jede Richtung orientiert sein



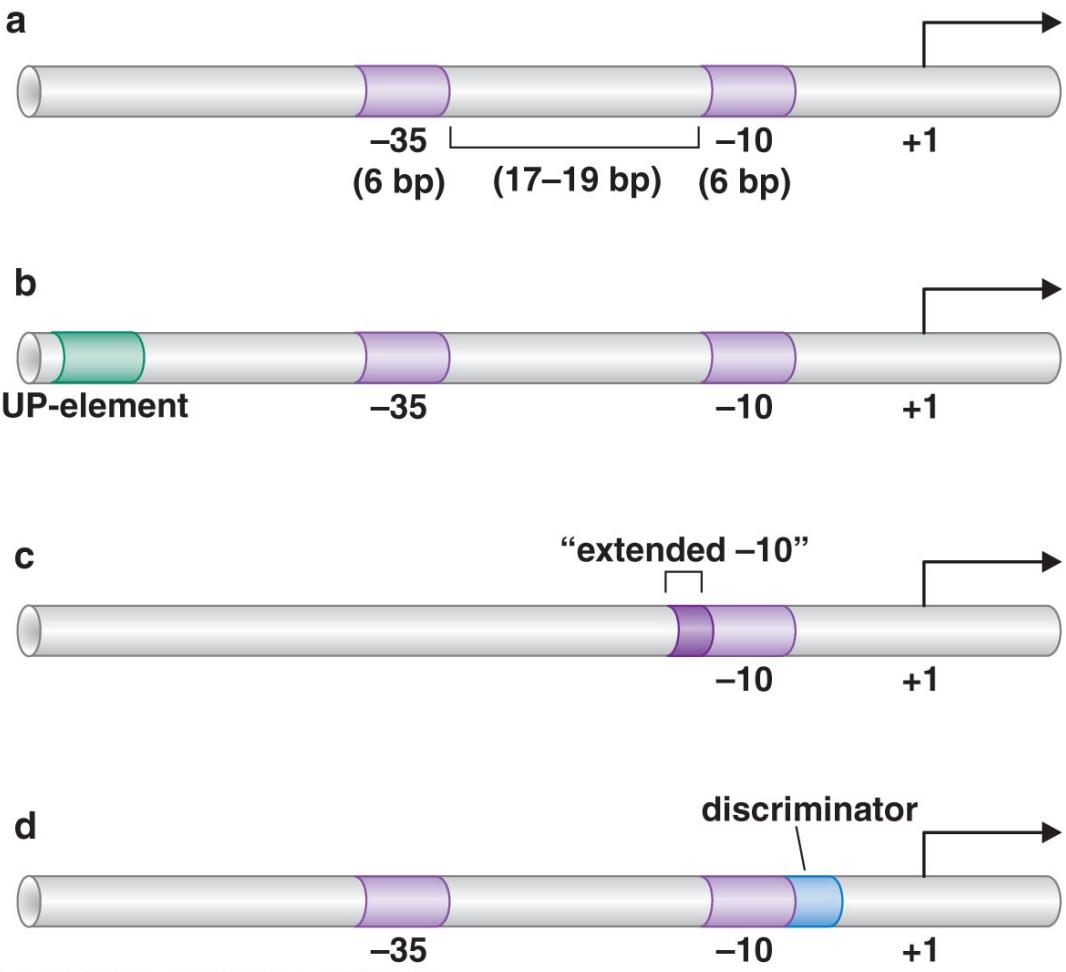
# Transkription in Bakterien erfordert **Initiation**, **Elongation** und **Termination**

Initiation der Transkription erfordert einen Promoter. Der Promoter bestimmt ORT und RICHTUNG der RNA-Neusynthese. Der Promotor und die kodierende Sequenz sind Teil eines Gens, d.h. bilden eine funktionelle Einheit.



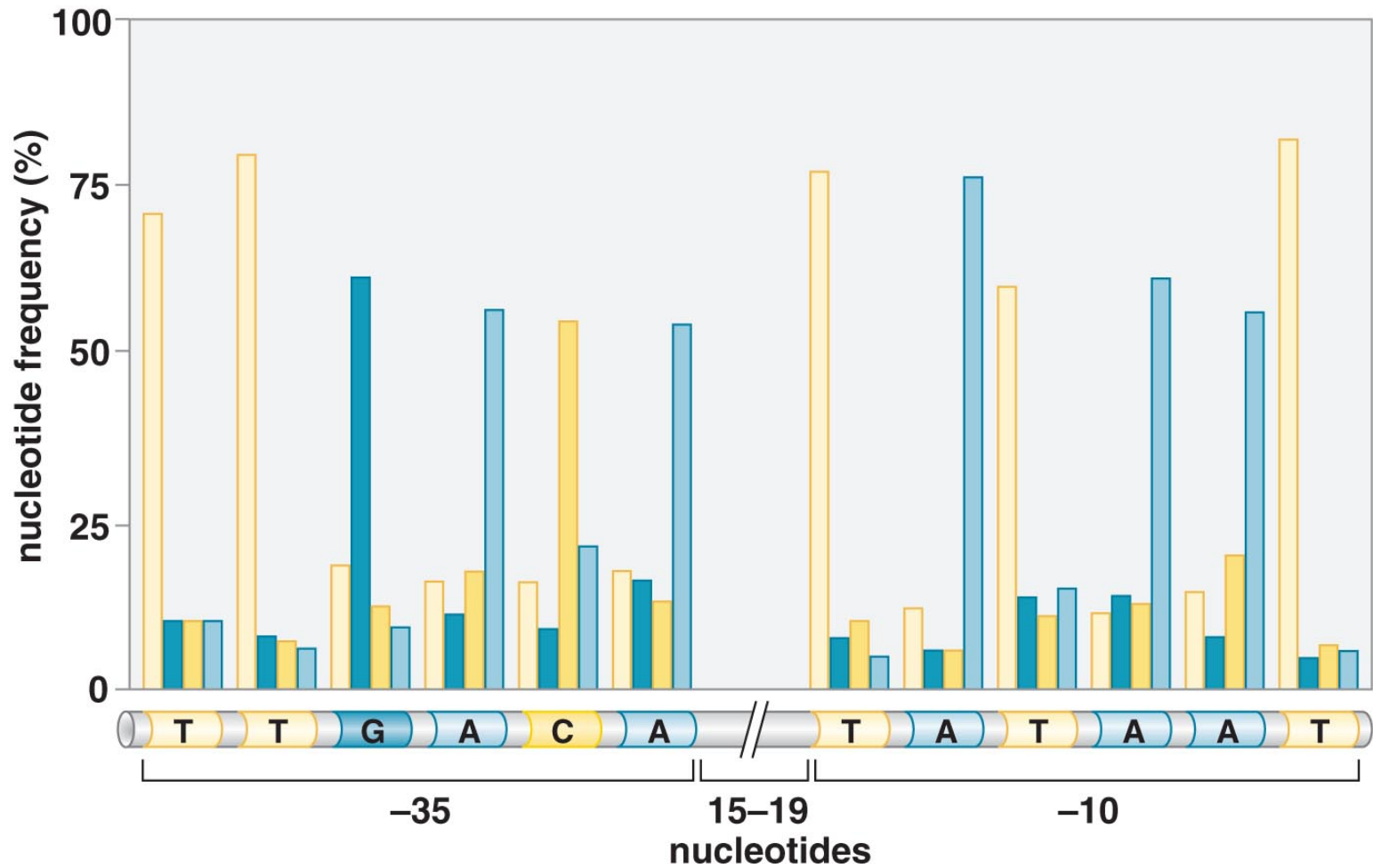
Bestimmte Sequenzen sind in den Promotoren bakterieller Gene konserviert.

Position -10 und -35 sind hoch konserviert.



Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.

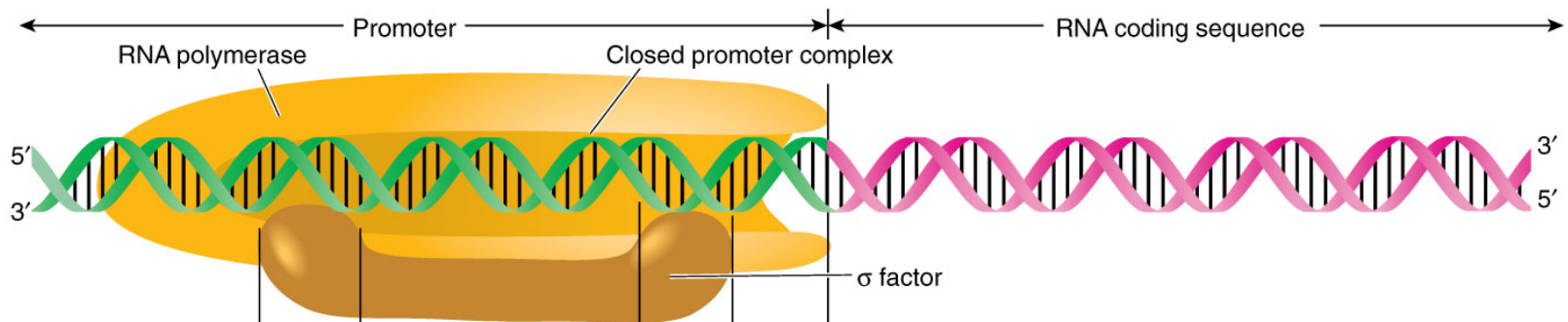
# Promotor Element Konsensus:



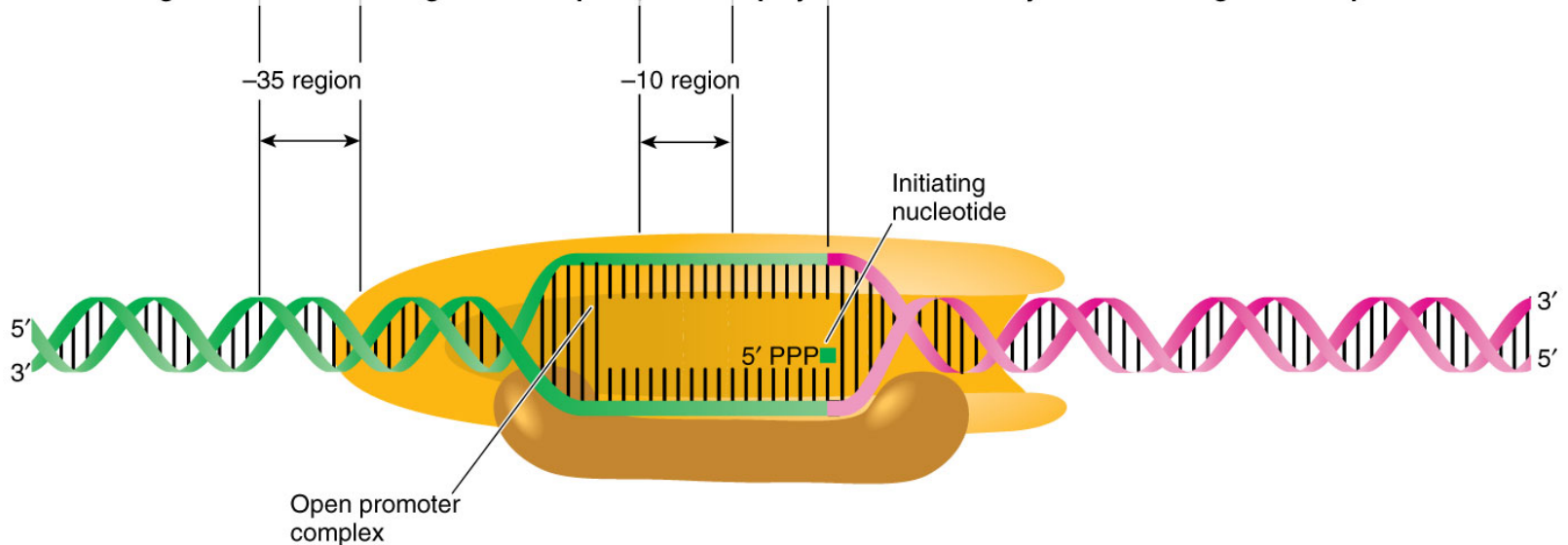


In vitro kann bakterielle RNA Polymerase von jeder Stelle Transkription initiieren. Assoziation mit **Sigma Faktoren (oft Sigma 70)** beschränkt die Initiation auf Promoterregionen. Pol besteht aus  $2\alpha$   $1\beta$   $1\beta'$   $1\sigma$ . Verschiedene  $\sigma$  Faktoren sind für unterschiedliche Promotoren spezifisch

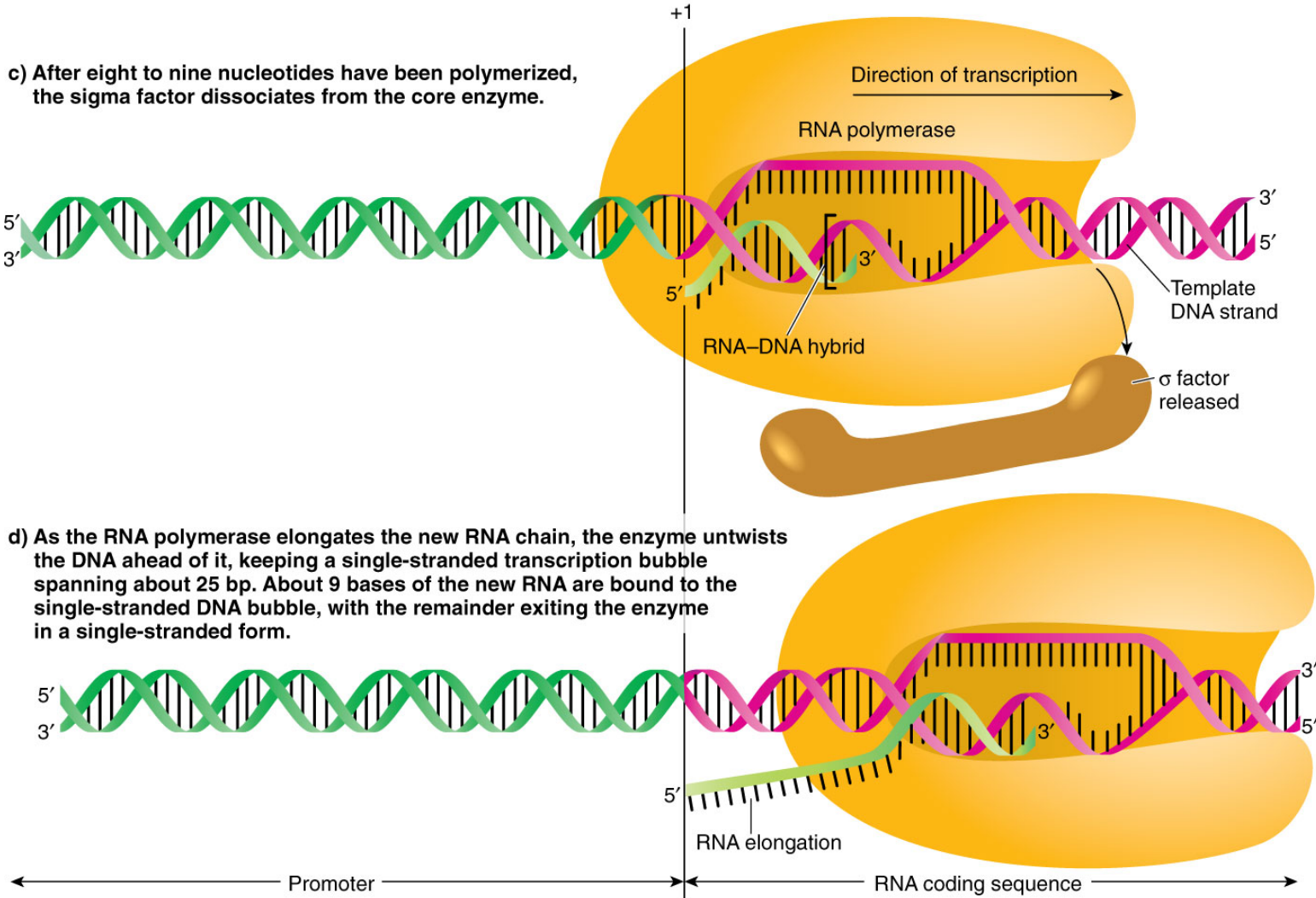
a) In initiation, the RNA polymerase holoenzyme first recognizes the promoter at the -35 region and binds to the full promoter.



b) As initiation continues, RNA polymerase binds more tightly to the promoter at the -10 region, accompanied by a local untwisting of the DNA in that region. At this point, the RNA polymerase is correctly oriented to begin transcription at +1.



Nach dem unwinding der DNA dissoziiert der Sigma Faktor und die RNA Polymerase kann Transkription beginnen. Während der Elongation verändert die Polymerase Ihre Form (wird kompakter). Bakterielle RNA Polymerase hat eine eigene unwinding Aktivität.



# Termination der Transkription in Prokaryoten:

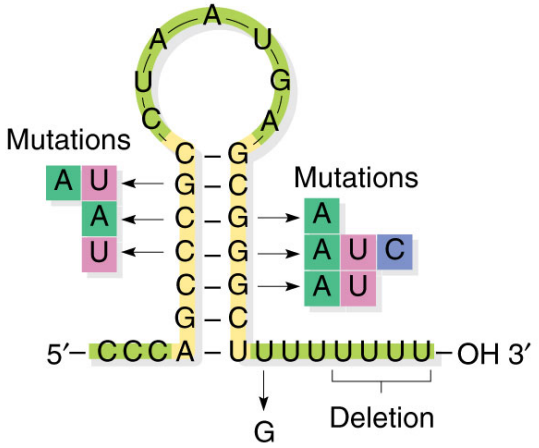
1. Durch terminale Hairpins (Haarnadeln) reguliert. Die Struktur verlangsamt die RNA Pol. Die folgenden U-reichen Sequenzen destabilisieren die RNA Interaktion mit dem Template Strang und die Polymerase dissoziiert.
2. Durch Rho-abhängige Termination. Eine C-reiche Sequenz in der RNA erlaubt das Binden von Rho. Rho ist eine Helikase und folgt der RNA-Polymerase. Durch ATP-abhängiges Entwinden der RNA vom Template Strang dissoziiert die RNA.

← Two fold symmetry →

Template (DNA) 5' CCCAGCCCGCCTAATGAGCGGGCTTTTTTTTGAACAAA 3'  
 3' GGGTCGGGCGGATTACTCGCCCGA AAAAAAACTTGT TTT 5'

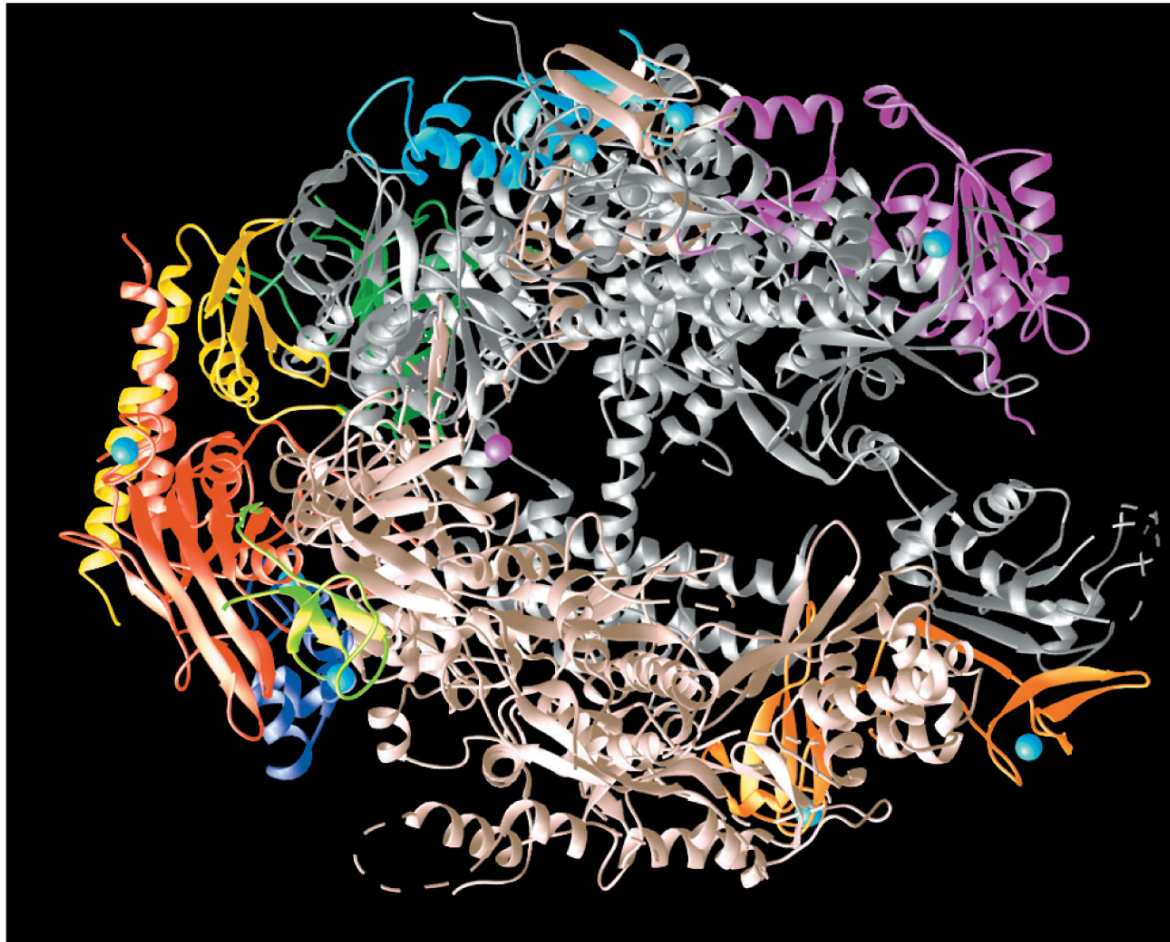
Transcript (RNA) 5' CCCAGCCCGC CUA AUGAGCGGGCU UUUUUUUU -OH 3'

Transcript folded to form termination hairpin



# Transkription in Eukaryoten:

## Modell der eukaryotischen Pol II



## In Eukaryoten werden verschiedene RNAs durch verschiedene RNA Polymerasen gebildet:

Ribosomale RNA (**rRNA**) macht den Großteil der zellulären RNA aus > 90%! Aus rRNA werden Ribosomen gebildet, wobei die RNA den aktiven Teil des Ribosoms ausmacht. rRNA wird durch **RNA-Polymerase I** synthetisiert.

Messenger RNA (**mRNA**) kodiert für Proteine. Enthält die Information, die am Ribosom in ein Protein übersetzt wird. mRNA wird von **RNA-Polymerase II** synthetisiert.

Kleine nicht kodierende RNAs (**snRNAs, tRNAs, snoRNAs**) haben regulatorische oder katalytische Funktion. Werden von **RNA-Polymerase III** synthetisiert.



# Regulation der Proteinkodierenden Gene:

Pol II. Benötigt **Promotoren** und **Enhancer**.

Wie wurden diese definiert:

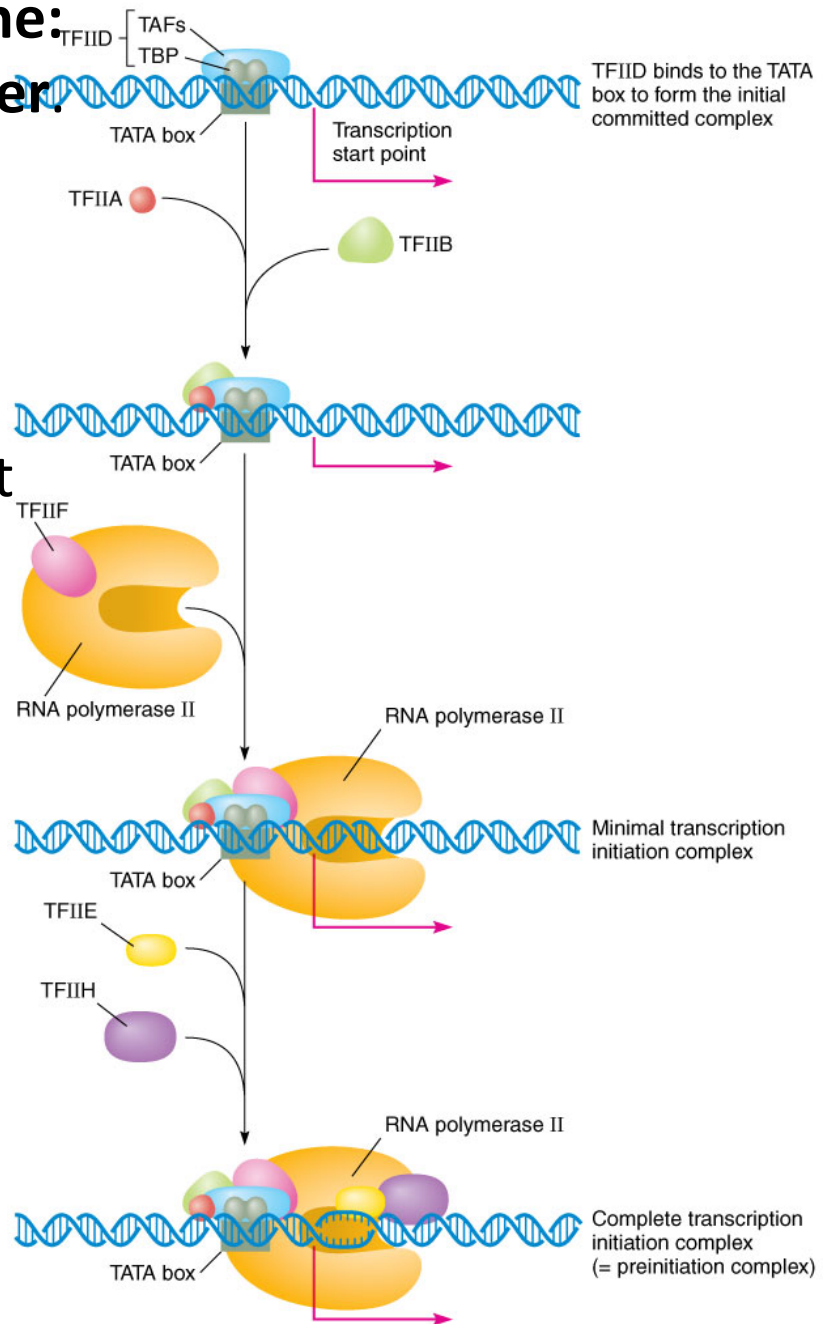
\* Durch Mutationen wird die Aktivität verringert.

*Oder:*

\* Transplantation des Promoters führt zur Aktivierung von "downstream" Sequenzen.

Promotoren können bis zu 200bp upstream der kodierenden Sequenz reichen.

Assembly of preinitiation complex



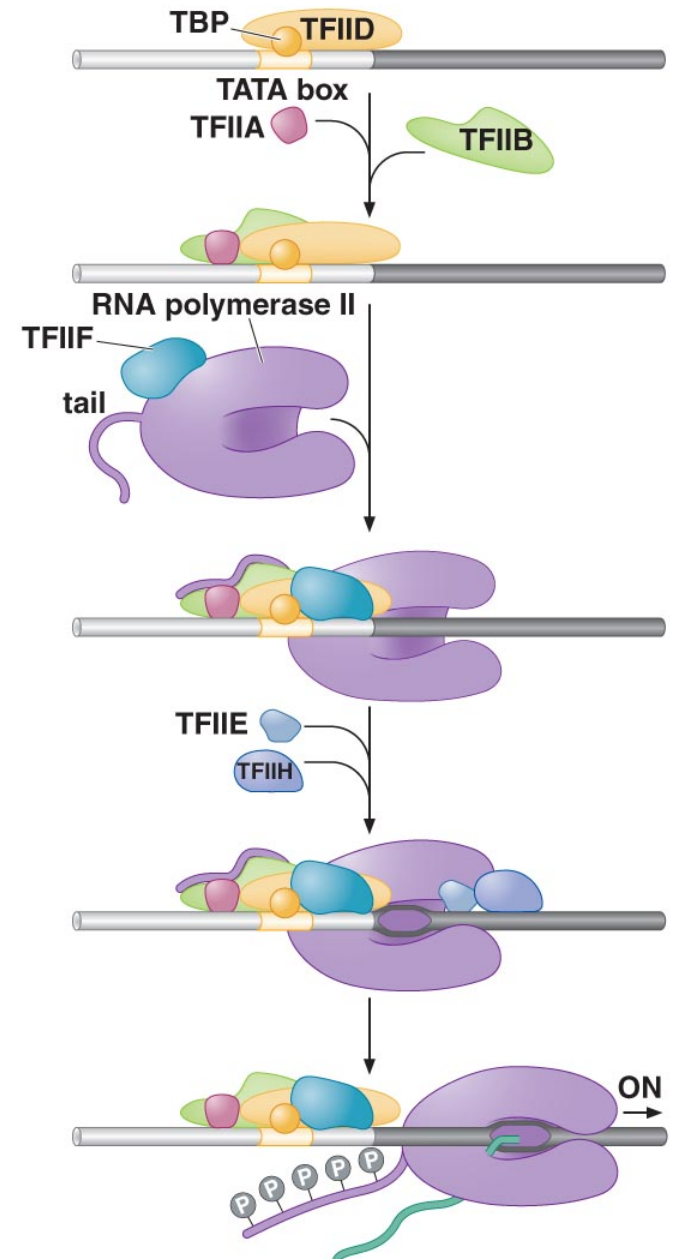
# TBP und TFIIs assemblieren an der TATA box

**Enhancer** können mehrere 1000 bp upstream oder downstream der kodierenden Sequenz liegen.

Der minimale core-Promoter umfasst ca 50 bp upstream des Transkriptionsstarts.

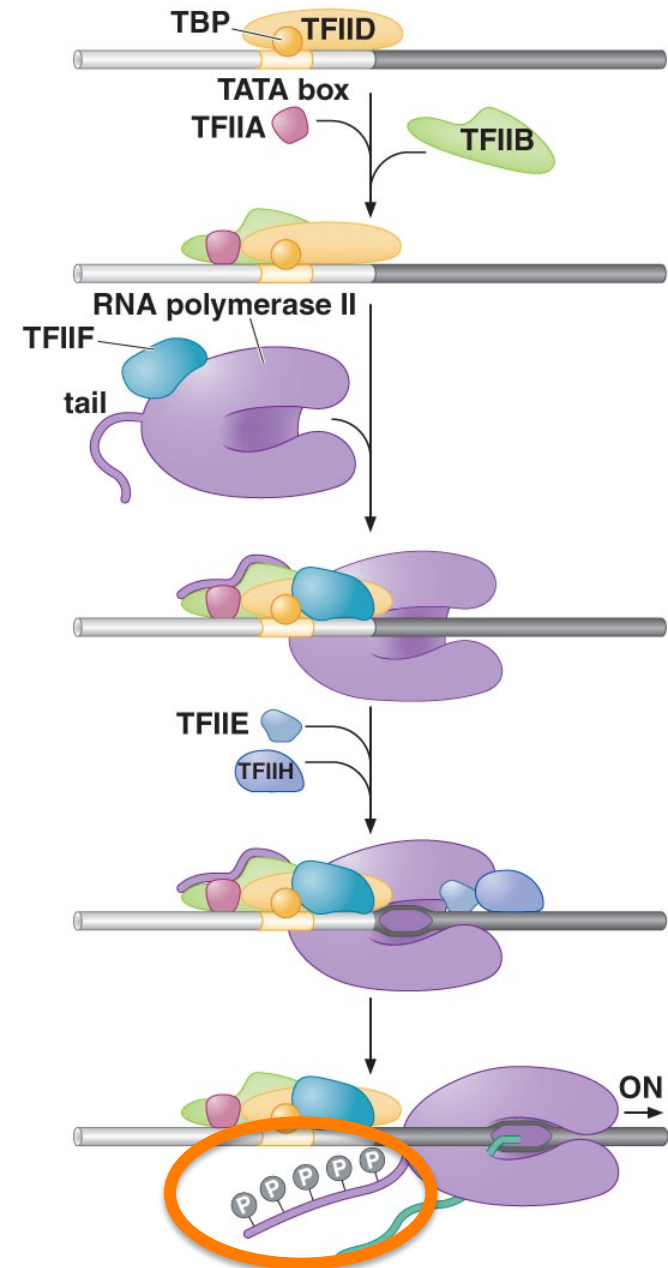
Bei +1 liegt die Initiator Sequenz, hier beginnt Transkription.

Bei -25 liegt die TATA Sequenz (Goldberg-Hogness box), dient als Bindungsstelle für TBP (Tata binding protein) was wiederum als Grundlage für die Bindung weiterer TFII (Transcription Factor II Proteine) dient.



# TBP und TFIIIs assemblieren an der TATA box

**TFII H** hat sowohl **Helikase** als auch **Kinase** Aktivität. Die Helikase entwindet die DNA. Die Kinase **Phosphoryliert** den C-Terminus von Pol II

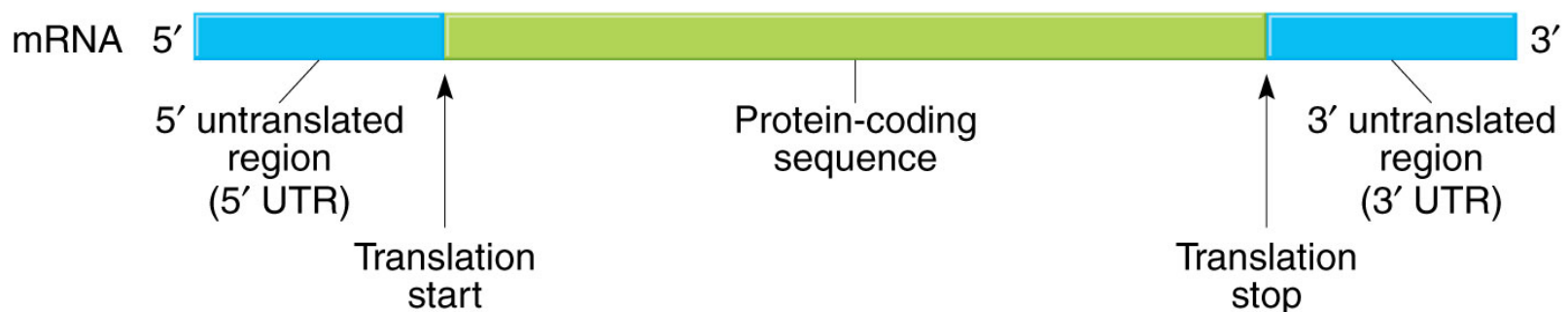




5' (upstream) der Promotor Region liegen "promotor proximale Elemente" -50 bis -200.

z.B **CAAT** box, **GC box**..... Diese Sequenzen werden von Transkriptions-Aktivatoren gebunden. Diese Aktivator Proteine erleichtern die Assemblierung des zentralen Transkriptionsapparates und kontrollieren so die Genexpression. Transkriptionsfaktoren sind Gewebs- und Zelltypspezifisch expremmiert.

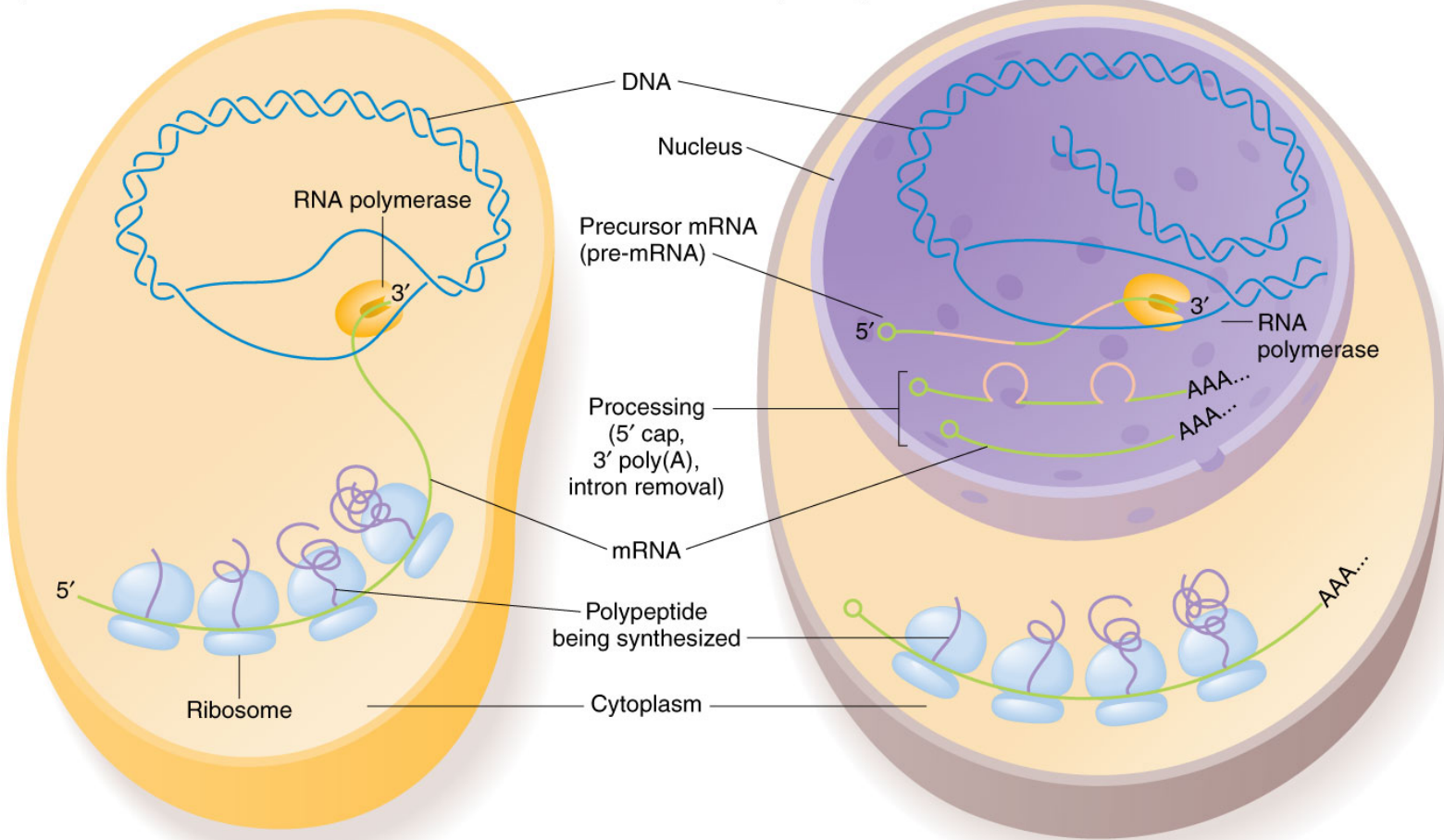
Transkribierte RNA besteht aus einer 5' untranslatierten Region, der Proteinkodierenden Sequenz, und einer 3' untranslatierten Region.  
**5' und 3' Untranslatierte Region bestimmen die Stabilität und Translatierbarkeit einer RNA.**



**Prokaryotische RNA** wird bereits **kotranskriptionell** translatiert.  
**Eukaryotische RNA** wird modifiziert, prozessiert, aus dem Kern exportiert und erst **im Zytoplasma** translatiert.

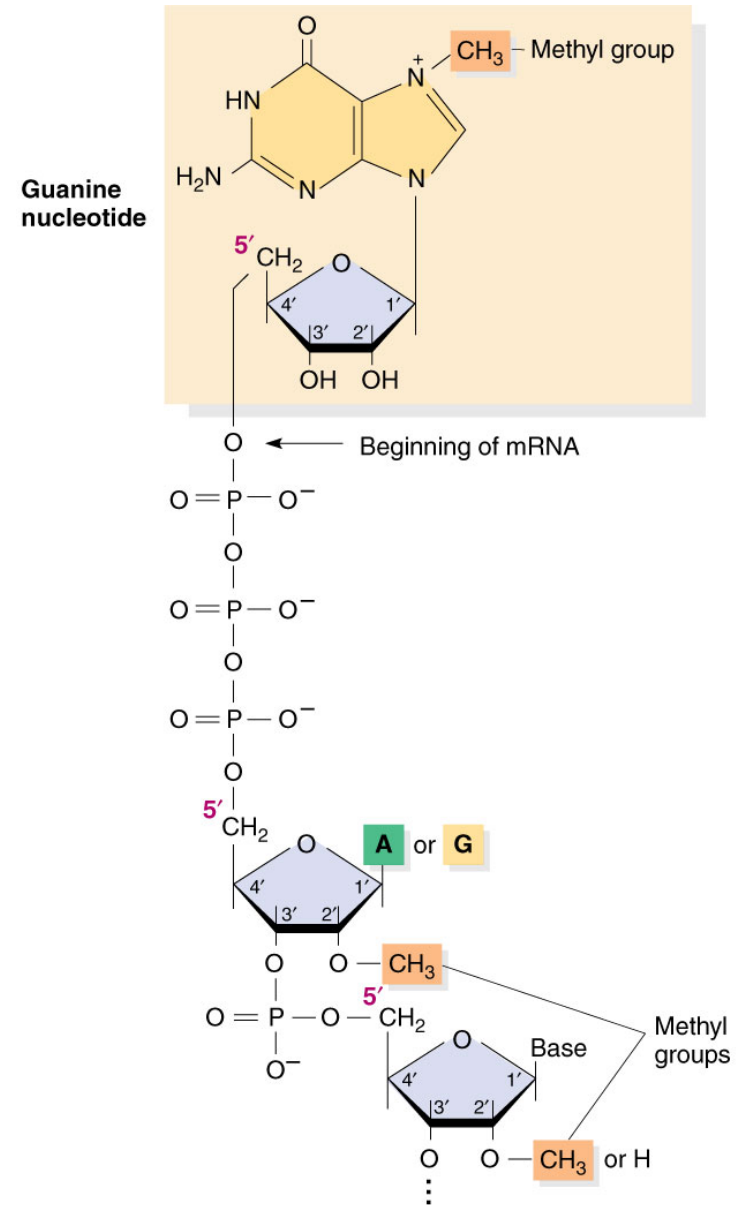
a) Bacterium

b) Eukaryote



# Modifikationen und Reifungen eukaryotischer mRNAs

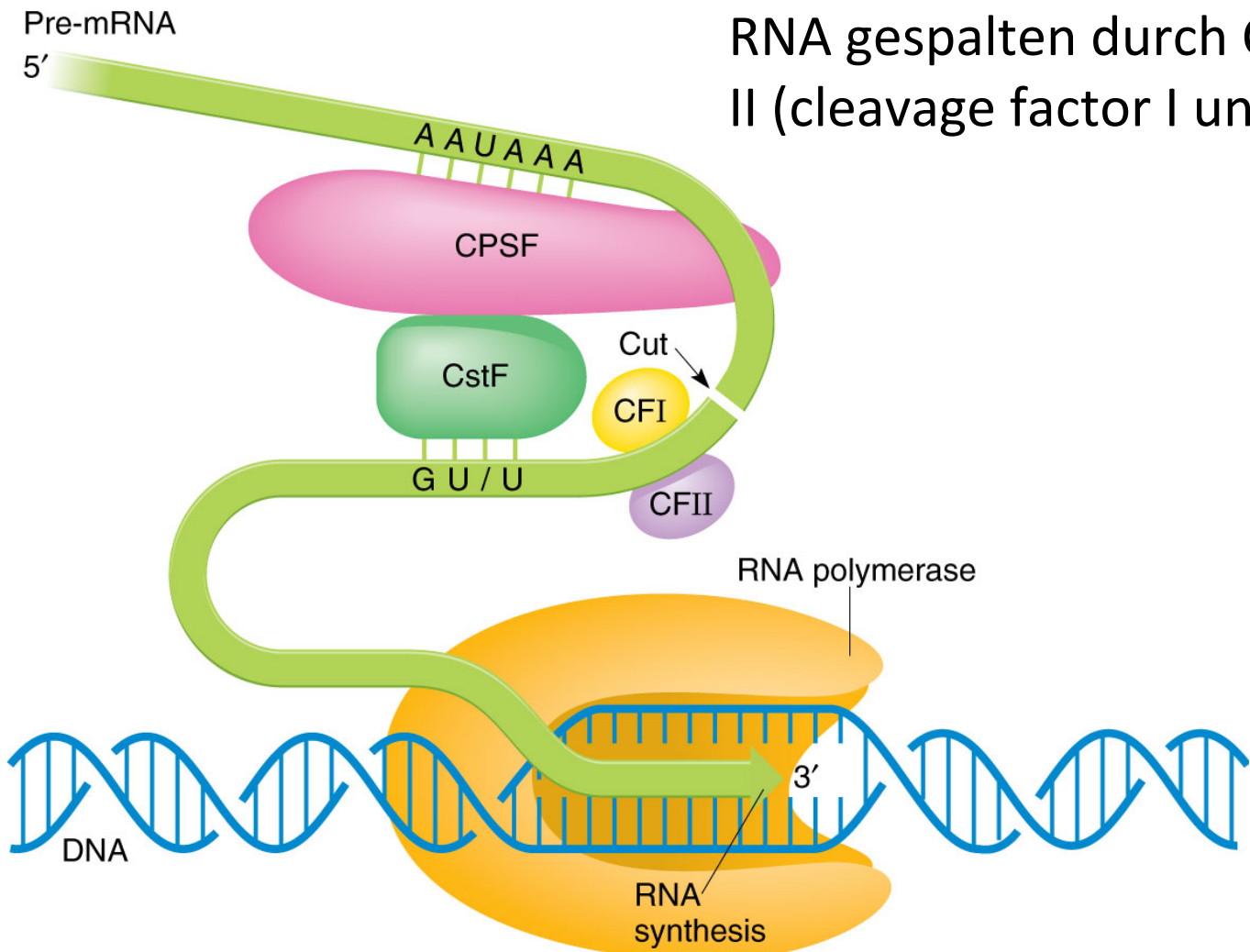
Wenn ca 20 Nukleotide der RNA Synthetisiert sind, wird ein 5' invertiertes methyl-Guanosin Cap durch ein "capping-Enzyme" angefügt. Die Zucker von Nukleotid 1 und 2 werden ebenfalls methyliert.



# Eukaryotische mRNAs werden an deren 3' Ende Polyadenyliert

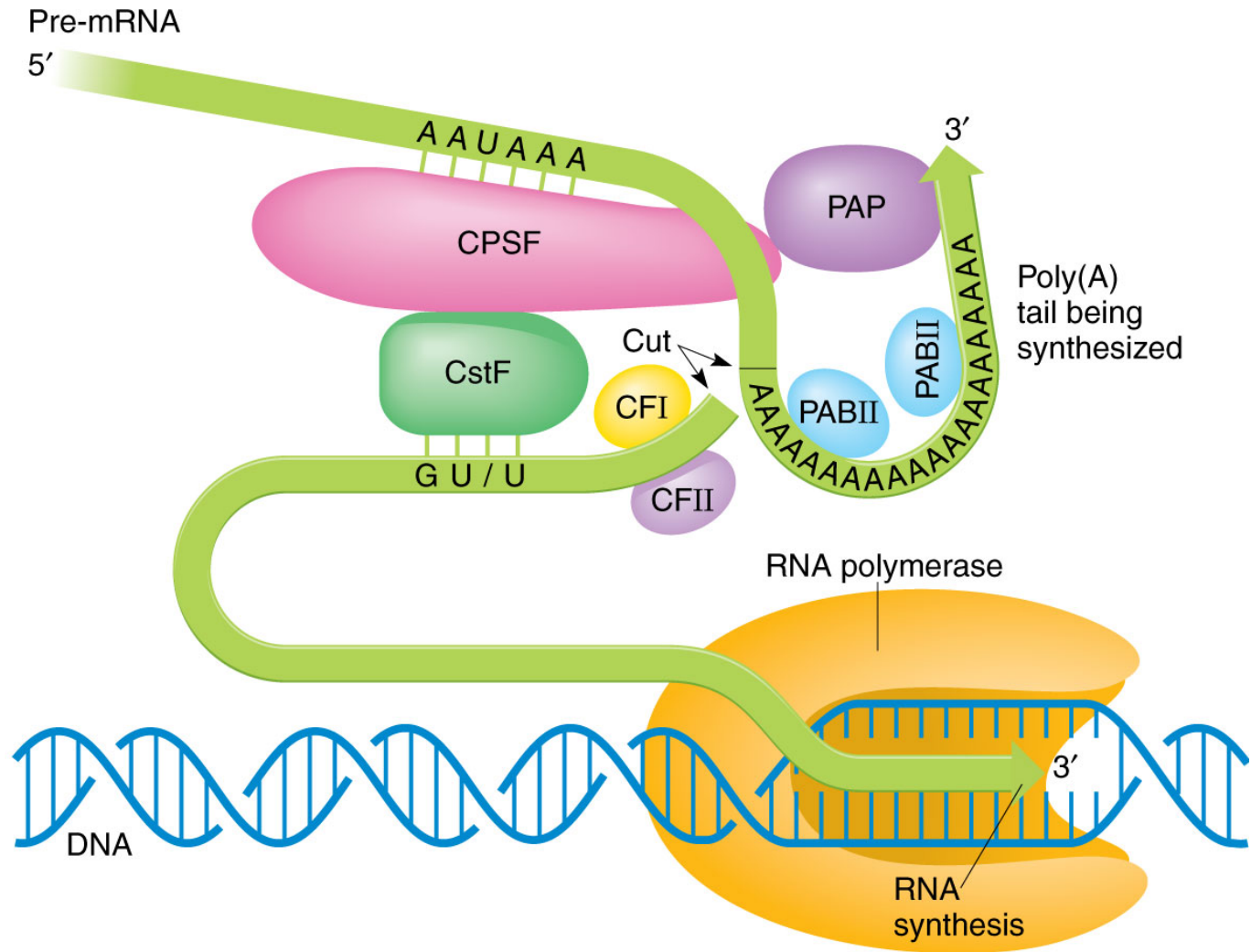
Sequenzen am 3' Ende der RNA werden erkannt (CPSF, CstF). In unmittelbarer Nähe wird die RNA gespalten durch CF I und CF II (cleavage factor I und II).

a) Cleavage of the pre-mRNA

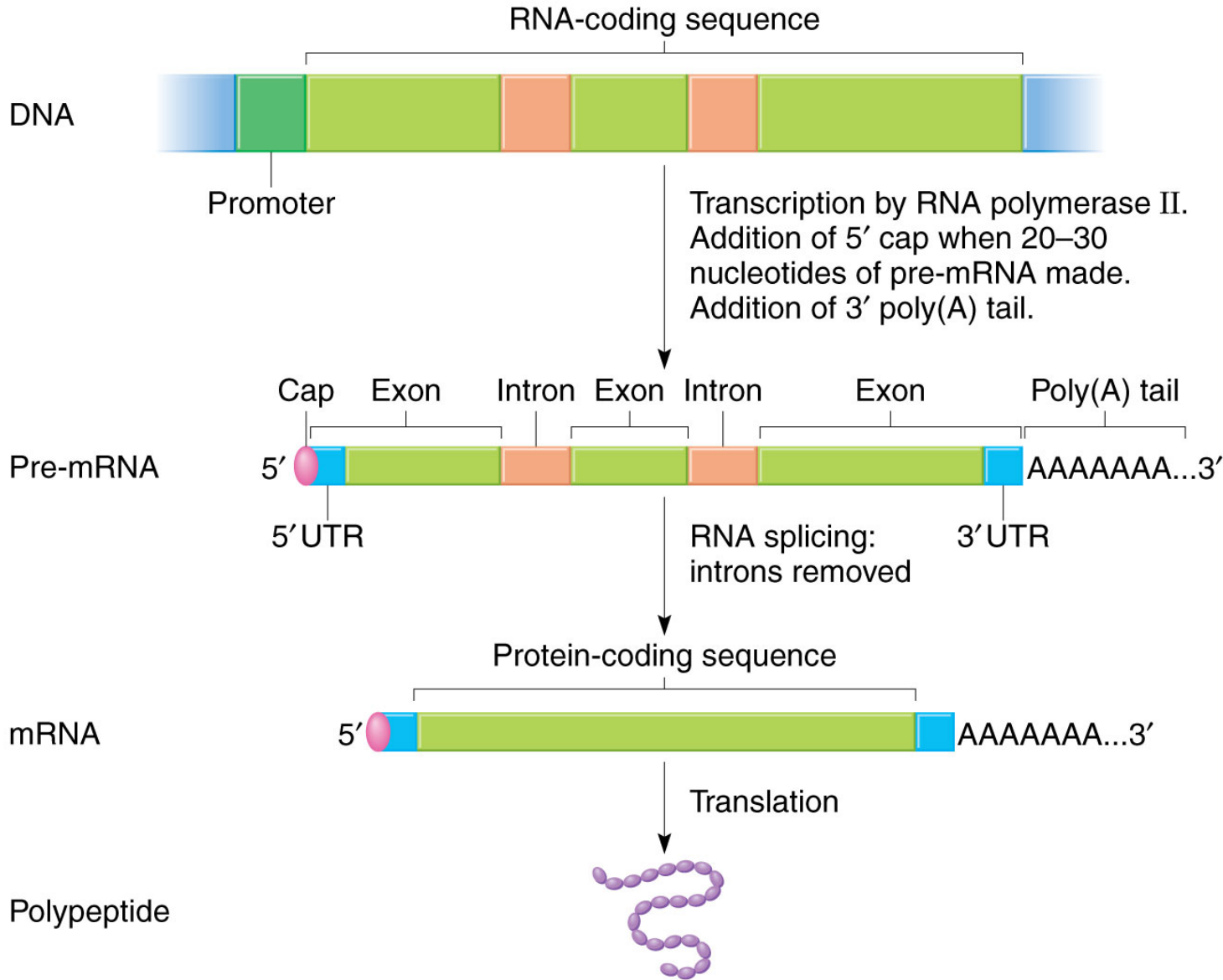


# Polyadenylation erfolgt durch PAP (Poly A polymerase)

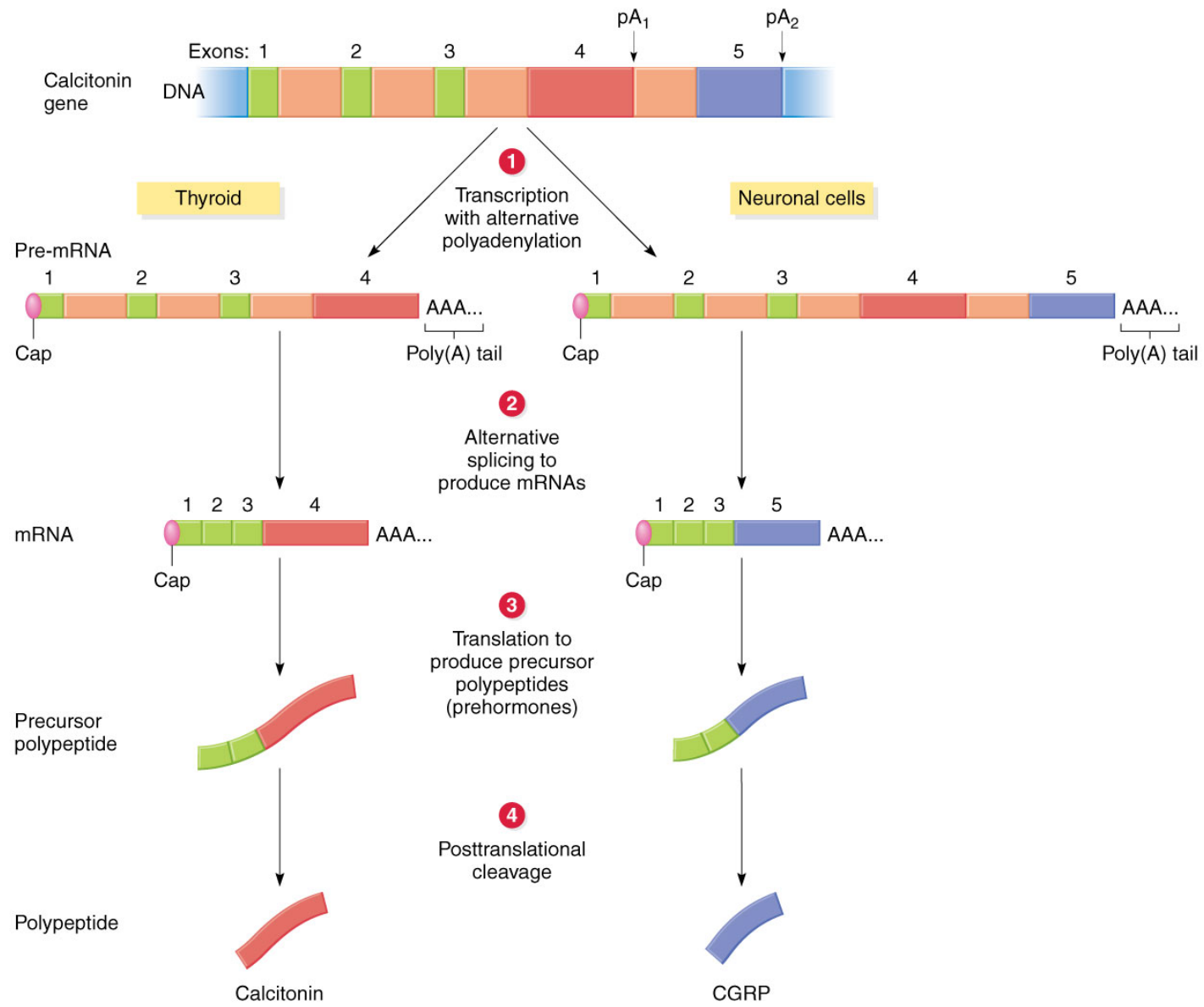
## b) Addition of the poly(A) tail



# Eukaryotische Gene sind nahezu immer von nicht kodierenden Sequenzen (Introns) unterbrochen. Introns werden im Zellkern entfernt: prä-mRNA Spleissen



# Alternatives Spleissen führt zur Diversifizierung des Proteoms (Menge aller in einer Zelle bzw. Organismus vorhandenen Proteine)

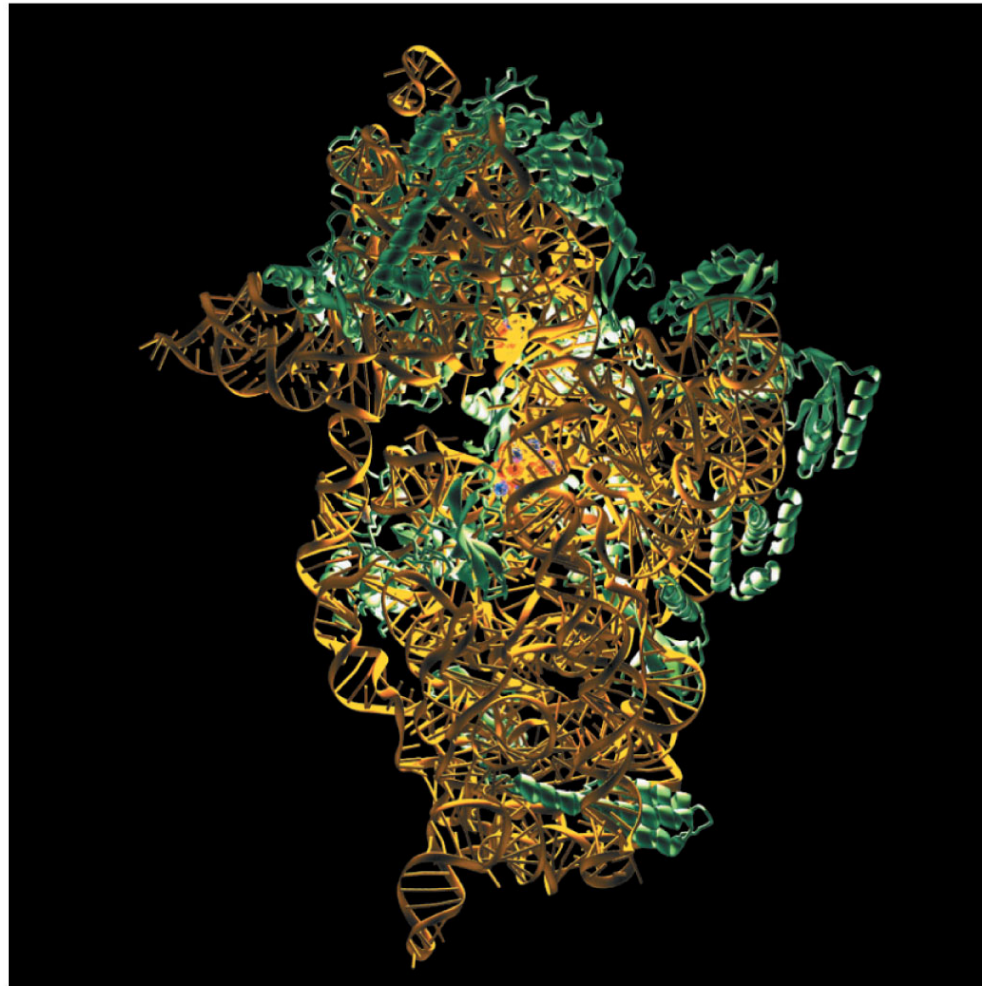


# Translation

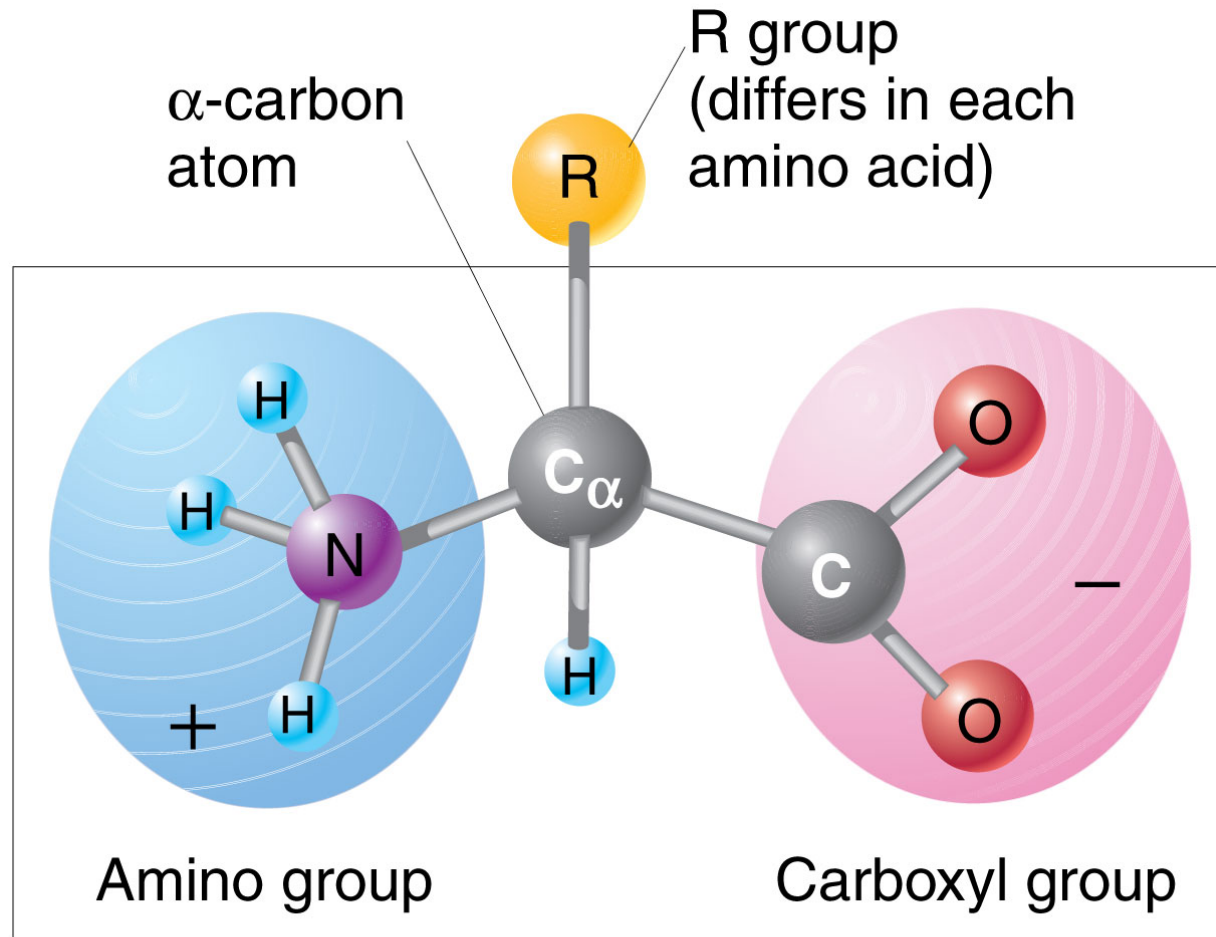


In RNA gespeicherte Information wird am Ribosom in Proteine übersetzt.

Das Ribosom besteht selbst aus RNA und Proteinen (Struktur des Ribosoms: Nobelpreis 2009)



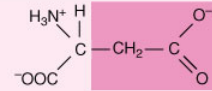
# Struktur einer Aminosäure.



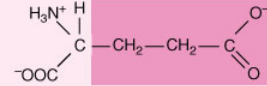
Structures common to all amino acids

# In Proteinen finden sich 20 verschiedene Aminosäuren

## Acidic

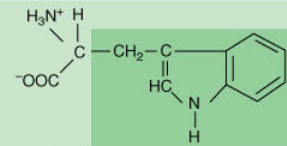


Aspartic acid  
(Asp) (D)

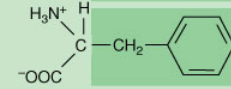


Glutamic acid  
(Glu) (E)

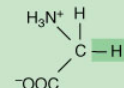
## Neutral, nonpolar



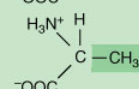
Tryptophan  
(Trp) (W)



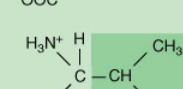
Phenylalanine  
(Phe) (F)



Glycine  
(Gly) (G)



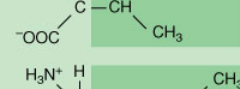
Alanine  
(Ala) (A)



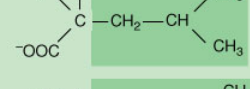
Valine  
(Val) (V)



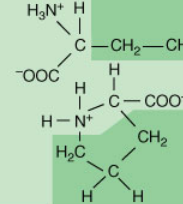
Isoleucine  
(Ile) (I)



Leucine  
(Leu) (L)

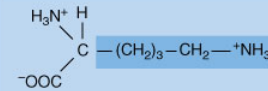


Methionine  
(Met) (M)

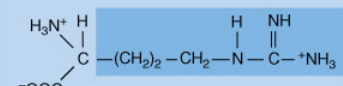


Proline  
(Pro) (P)

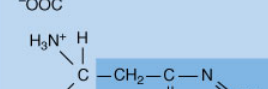
## Basic



Lysine  
(Lys) (K)

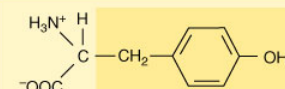


Arginine  
(Arg) (R)

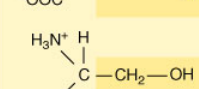


Histidine  
(His) (H)

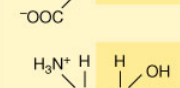
## Neutral, polar



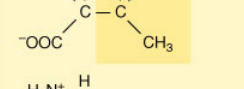
Tyrosine  
(Tyr) (Y)



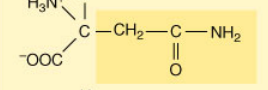
Serine  
(Ser) (S)



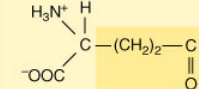
Threonine  
(Thr) (T)



Asparagine  
(Asn) (N)

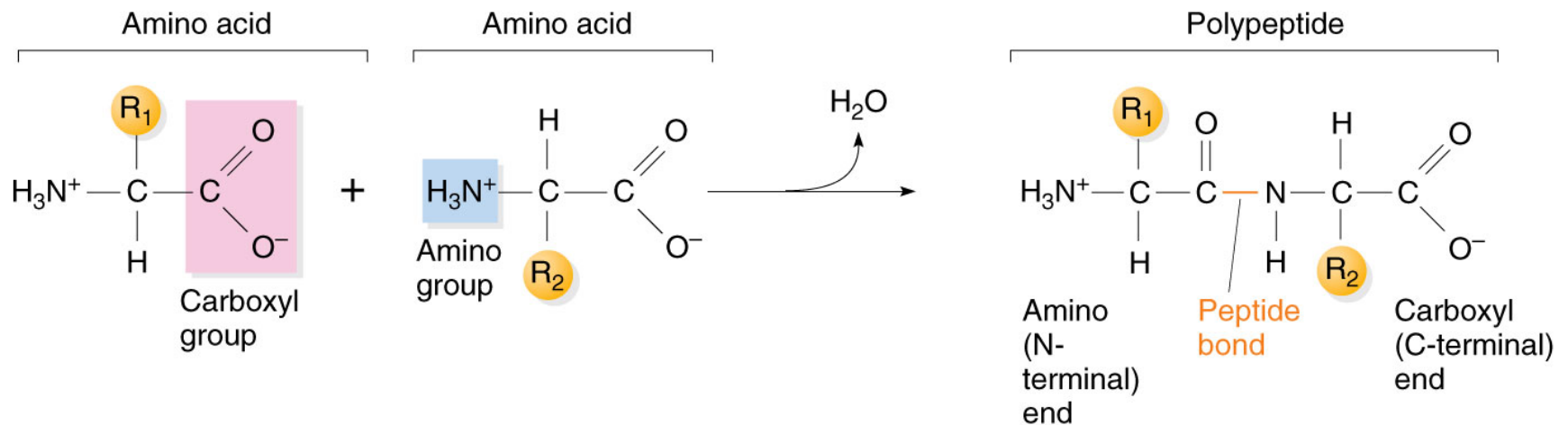


Glutamine  
(Gln) (Q)



Cysteine  
(Cys) (C)

Einzelne Aminosäuren werden durch eine Peptidbindung miteinander verknüpft. Diese Verknüpfung erfolgt am Ribosom



# Der genetische code ist ein Tripletcode.

Jede Position einer RNA kann 4 verschiedene Informationen beinhalten G, A, U, C.

2 aufeinander folgende Nukleotide könnten 16 ( $4 \times 4$ ) verschiedene Informationen beschreiben

3 aufeinanderfolgende Nukleotide können  $4 \times 4 \times 4 = 64$  Aminosäuren beschreiben

Da es nur 20 natürliche Aminosäuren gibt, ist der genetische Code redundant.

Experimente am Bakteriophagen T4 bestätigten den Triplet code.

2 Arten von Plaques auf Bakterienrasen:

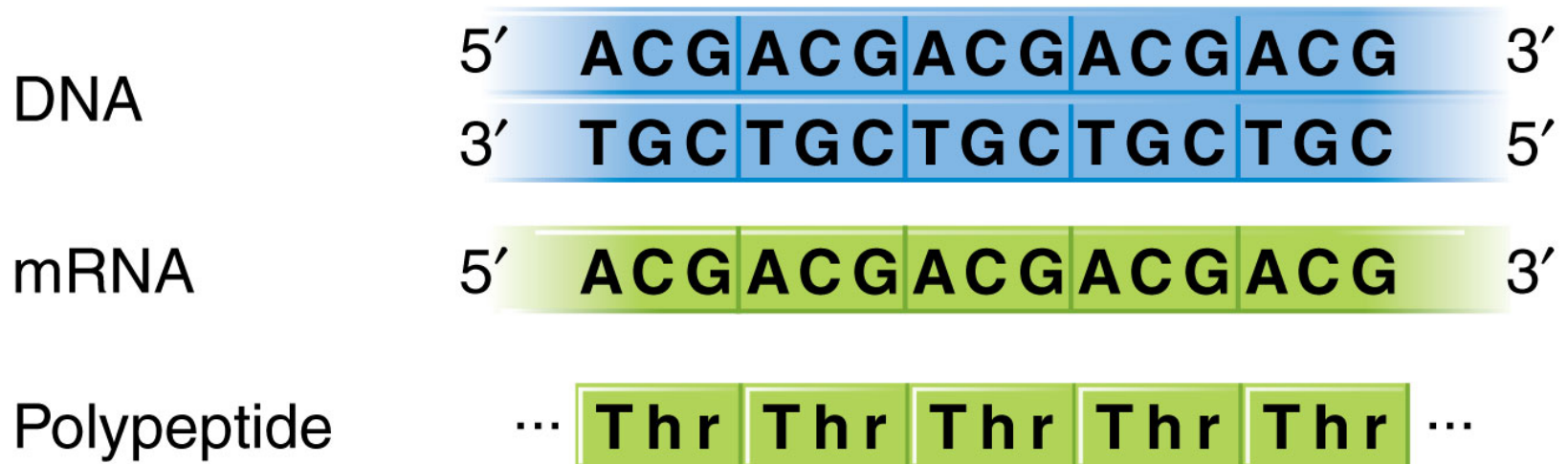
***r+* Stämme: trübe Plaques**

***rII* Stämme: klare Plaques**

Proflavin erzeugt frameshifts: Entfernung oder Hinzufügen einzelner Basen

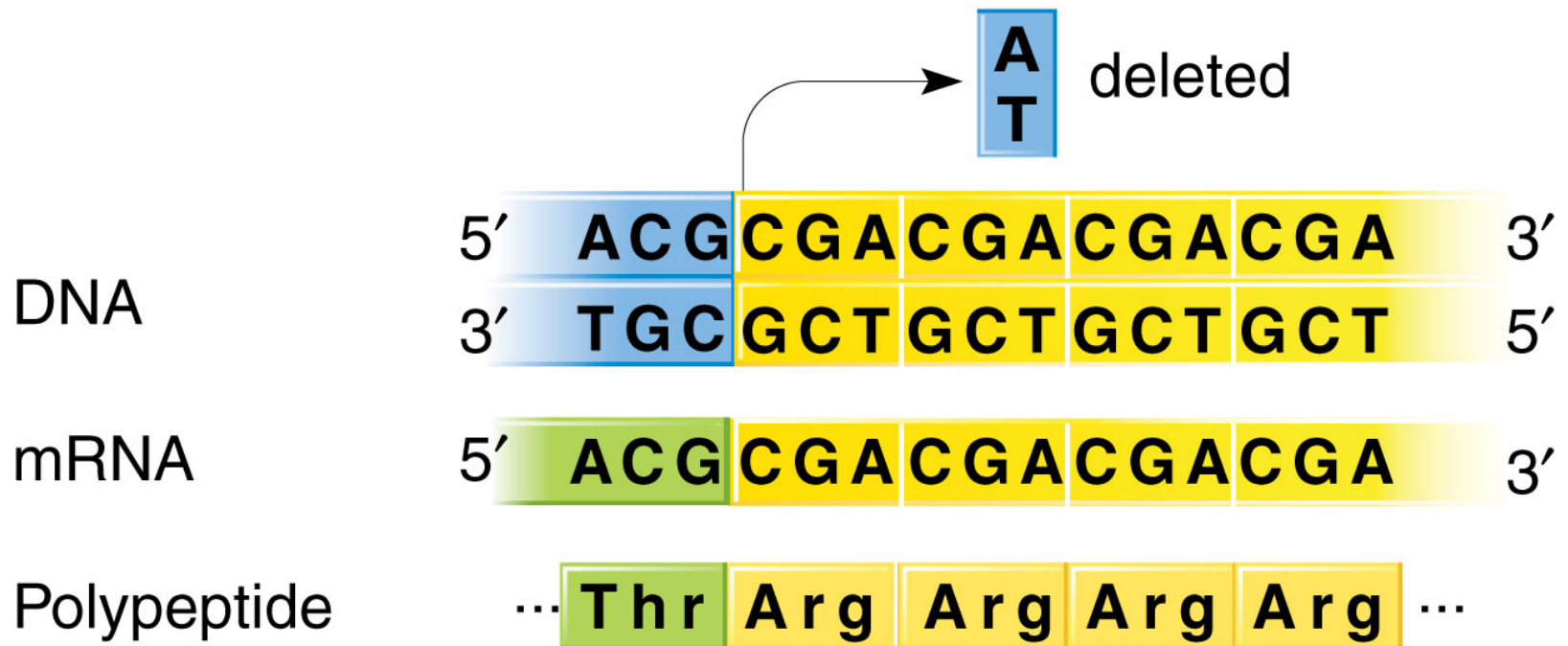
Mutationen führen zur Bildung von Phagen, welche klare Plaques erzeugen

**a) Wild type**

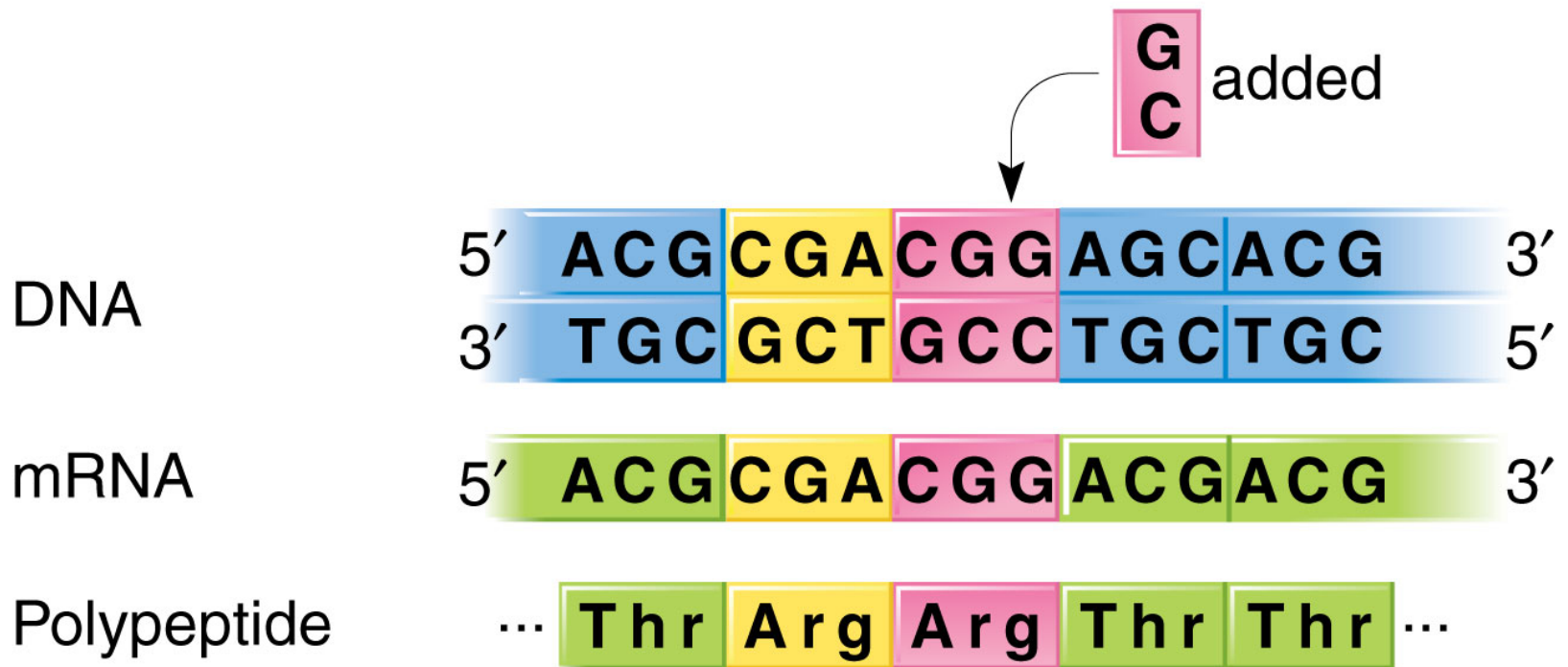


Proflavin erzeugt frameshifts: Entfernung oder Hinzufügen einzelner Basen

## b) Frameshift mutation by deletion



### c) Reversion of deletion mutation by addition



© 2010 Pearson Education, Inc.

Bis auf 2 falsche Aminosäuren, wird wieder das richtige Protein gebildet (wieder trübe Plaques)



Durch Hinzufügen von **3x je 1 Base** an benachbarten Positionen kann die Proteinfunktion wieder hergestellt werden

Normal mRNA

AUG ACA CAU AAC GGC UUC GUA UGG UGU GAA

Amino acids

Met Thr His Asn Gly Phe Val Trp Cys Glu

3 + mutations

+U

+C

+A

mRNA

AUG AUC ACA UAC ACG GCA UUC GUA UGG UGU GAA

Amino acids

Met Ile Thr Tyr Thr Ala Phe Val Trp Cys Glu

Incorrect amino acids  
in polypeptide

# Aber: wie lautet der genetische Code?

*Nierenberg und Khorana (Nobelpreis 1968)*

Verwendeten Zell-freies in vitro Translationssystem (Zelllysat) in welches synthetische RNA zugefügt wurde. Die so synthetisierten Aminosäuren wurden analysiert.

Beispiele der verwendeten RNAs:

AAA AAA AAA

UCU CUC UCU C

GAG GAG GAG GAG

usw.

So konnten fast alle möglichen Aminosäureketten synthetisiert werden

Der genetische code ist redundant.

Ein Triplet code.

Ohne Unterbrechungen



Nahezu Universell

(Ausnahmen sind Organelle oder Protozoen)

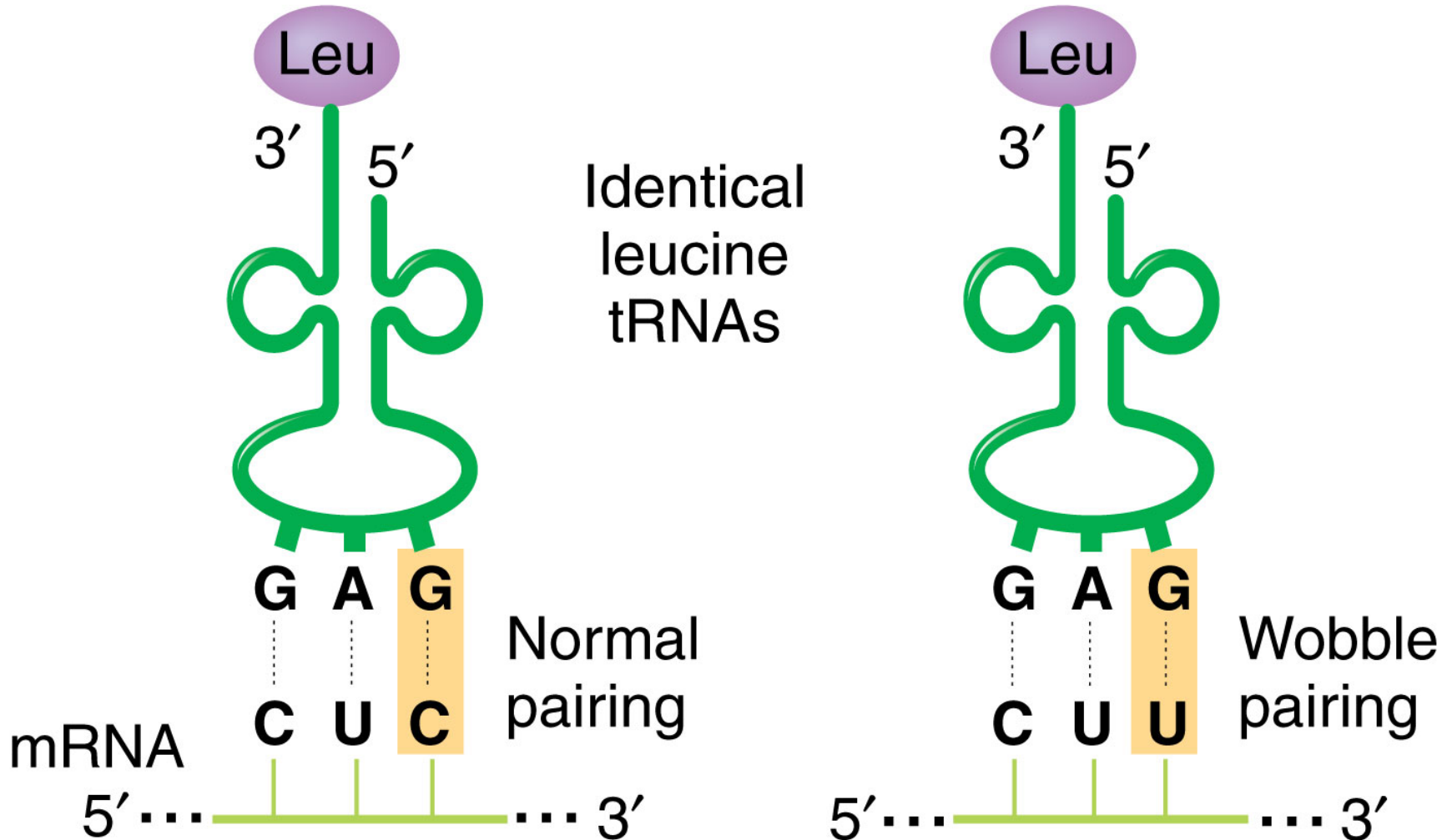
Im code gibt es **Start und Stop** Sequenzen.

Die dritte Position kann eine ungenaue Basenpaarung eingehen (wobble Hypothese)

		Second letter							
		U	C	A	G				
U	UUU	Phe (F)	UCU	Ser (S)	UAU	Tyr (Y)	UGU	Cys (C)	U C A G
	UUC								
	UUA	Leu (L)	UCA		UAA	Stop	UGA	Stop	
	UUG		UCG		UAG	Stop	UGG	Trp (W)	
C	CUU	Leu (L)	CCU	Pro (P)	CAU	His (H)	CGU	Arg (R)	U C A G
	CUC								
	CUA		CCA		CAA	Gln (Q)	CGA		
	CUG		CCG		CAG		CGG		
A	AUU	Ile (I)	ACU	Thr (T)	AAU	Asn (N)	AGU	Ser (S)	U C A G
	AUC								
	AUA		ACA		AAA	Lys (K)	AGA	Arg (R)	
	AUG	Met (M)	ACG		AAG		AGG		
G	GUU	Val (V)	GCU	Ala (A)	GAU	Asp (D)	GGU	Gly (G)	U C A G
	GUC								
	GUA		GCA		GAA	Glu (E)	GGA		
	GUG		GCG		GAG		GGG		

 = Chain termination codon (stop)  
 = Initiation codon

Durch wobble-basepairs kann eine t-RNA verschiedene codons decodieren



# Mögliche Basenpaarungen der wobble Position:

**Table 6.1 Wobble in the Genetic Code**

<b>Nucleotide at 5' End of Anticodon</b>		<b>Nucleotide at 3' End of Codon</b>
G	can pair with	U or C
C	can pair with	G
A	can pair with	U
U	can pair with	A or G
I (inosine)	can pair with	A, U, or C

# Jede Sequenz hat drei mögliche Codes

ctgtggcagagaatccggcggaatctgagccatccgagccgccaccatgacggtgggcaagagcagcaagatgctgcagcacattgactacaggatgaggtgcatcctgcaggacggcc  
L W Q R I R R E S E P S E P P P \* R W A R A A R C C S T L T T G \* G A S C R T A  
C G R E S G G N L S H P S R H H D G G Q E Q Q D A A A H \* L Q D E V H P A G R P  
V A E N P A G I \* A I R A A T M T V G K S S K M L Q H I D Y R M R C I L Q D G R

gtatcttcatcgggaccttcaaagcctttgacaagcacatgaacttgatcctgtgtgactgtgatgagttcaggaagatcaagccaaagaactccaacaagcagaaaggaagagaagc  
V S S S G P S K P L T S T \* T \* S C V T V M S S G R S S Q R T P N K Q K G K R S  
Y L H R D L Q S L \* Q A H E L D P V \* L \* \* V Q E D Q A K E L Q T S R K G R E A  
I F I G T F K A F D K H M N L I L C D C D E F R K I K P K N S K Q A E R E E K R

gagtccttggtctggtgcttcgaggggagaacctggtctccatgacagtagaaggacctcctccaaagatactggcattgccgagtgccacttgcaggagctgctggggggccag  
E S L V W C C F E G R T W S P \* Q \* K D L L P K I L A L P E C H L L E L L G G Q  
S P W S G V A S R G E P G L H D S R R T S S Q R Y W H C P S A T C W S C W G A R  
V L G L V L L R G E N L V S M T V E G P P P K D T G I A R V P L A G A A G G P G

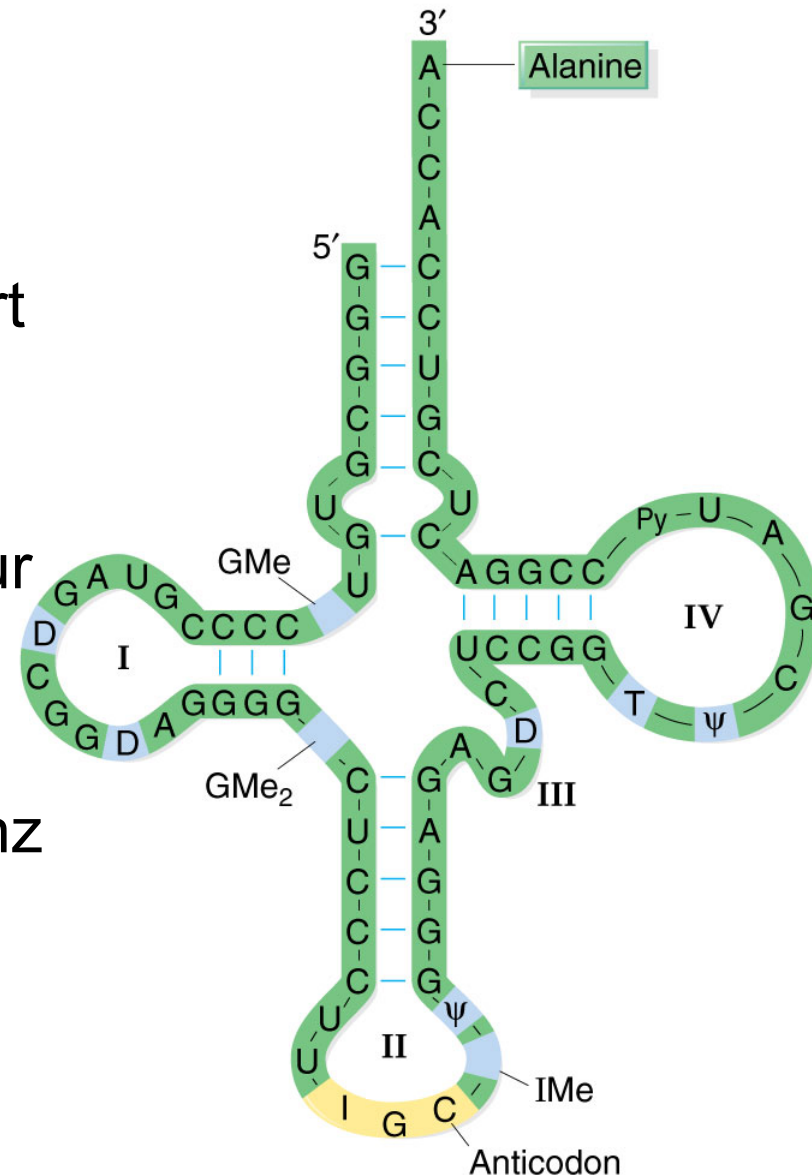
caggcatgcgaggtctgctttgatcttcacacagcattatgtcagtagctcagaggcgatgtggaccctgttccccaggggtcaggttaccacagacctgtttgtttgttatgcagttcat  
Q A C E V C F D L H T A L C Q \* L R G D V D P V P Q G Q V T T D L F V C Y A V H  
R H A R S A L I F T Q H Y V S S S E A M W T L F P R V R L P Q T C L F V M Q F M  
G M R G L L \* S S H S I M S V A Q R R C G P C S P G S G Y H R P V C L L C S S W

ggagtctcctggattgtctggtttccctttcagggctccctccctgggaaatgtgcctccaagggcctagactcatcttggttctcctcagctccttgctgtttccttcccaaaa  
G V S L D C L V S L S G L P P W E M C P P R P \* T H L G S P Q L L A C F L P K  
E S P W I V W F P F Q G S L P G K C A L Q G P R L I L V L L S S L P V S F P K  
S L P G L S G F P F R A P S L G N V P S K A L D S S W F S S A P C L F P S Q

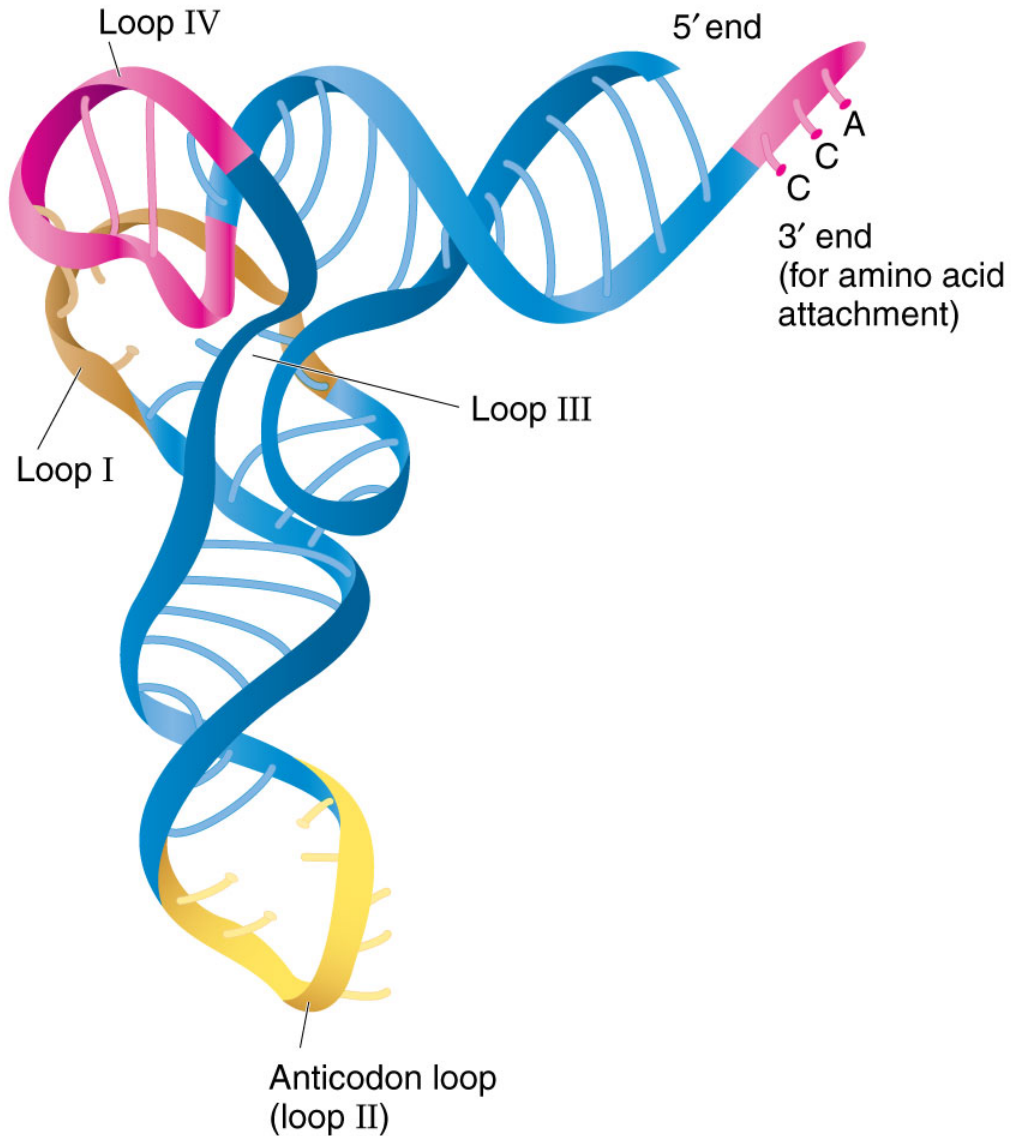
Nur eine Sequenz ergibt einen längeren offenen Leserahmen (reading frame) welcher durch ein Start und Stopp Kodon definiert ist

Translation erfolgt vom 5' zum 3' Ende der mRNA. tRNAs (transfer RNAs) bringen Aminosäuren zum Ribosom  
a) Cloverleaf model of tRNA

Jede tRNA wird mit einer spezifischen Aminosäure beladen.  
Das Anticodon der tRNA basenpaart mit dem Codon der mRNA.  
tRNAs nehmen eine Kleeblattstruktur (durch Basenpaarungen) ein.  
Alle tRNAs haben eine CCA Sequenz an deren 3' Ende. Diese Sequenz wird post-transkriptionell angefügt.



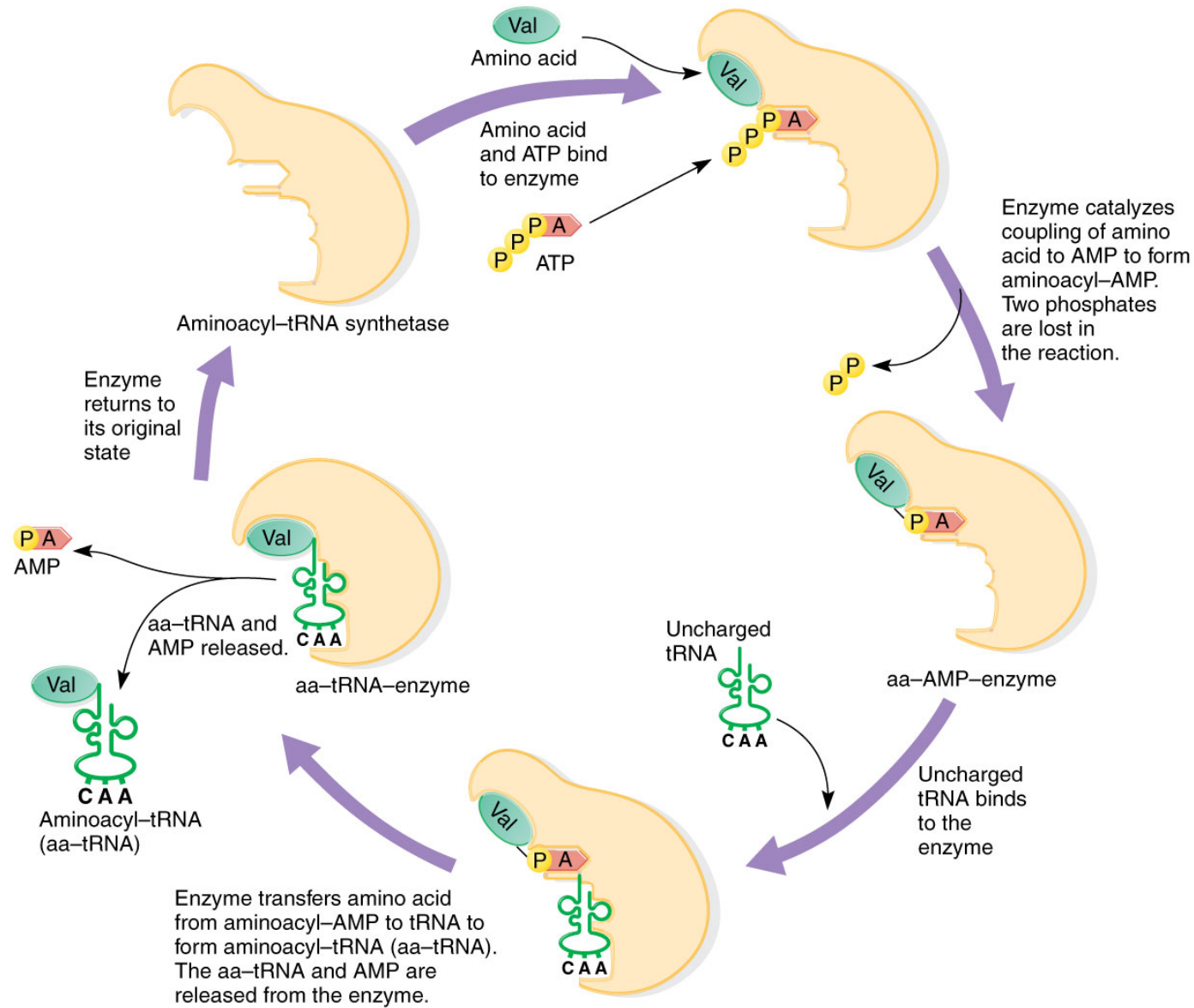
**b) Schematic of the three-dimensional L-shaped structure of a tRNA, here yeast phenylalanine tRNA**



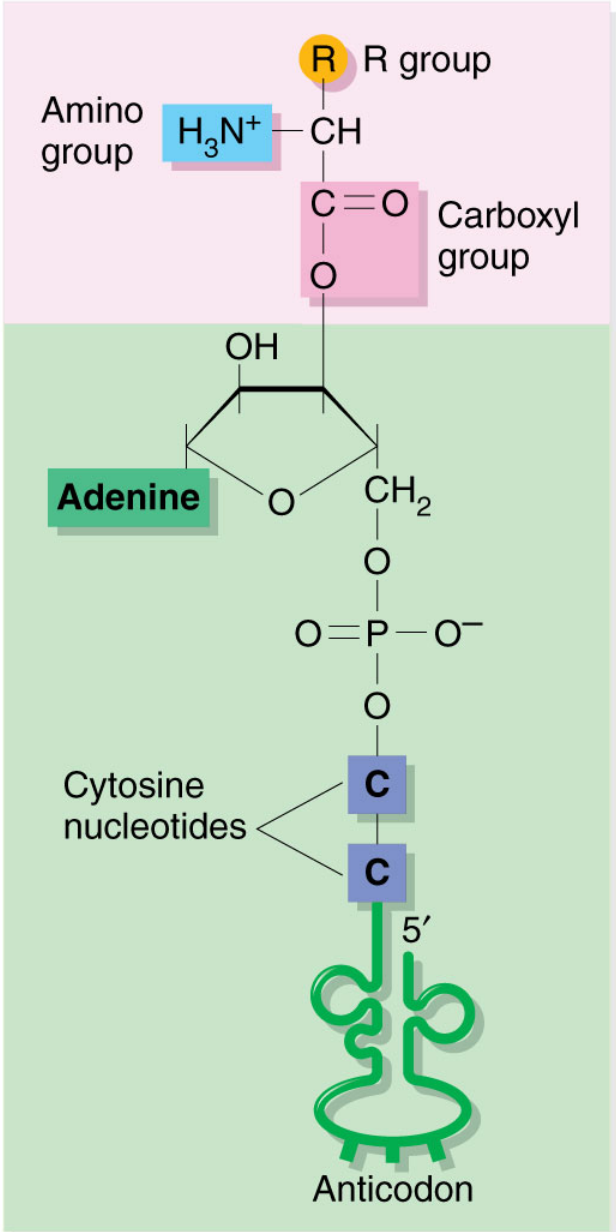
In gefaltetem Zustand nimmt die tRNA eine L-Form ein



# Aminoacyl-tRNA Synthetase belädt tRNAs unter ATP Verbrauch



# Die Carboxyl Gruppe der Aminosäure bindet an das 3' oder 2' OH der tRNA



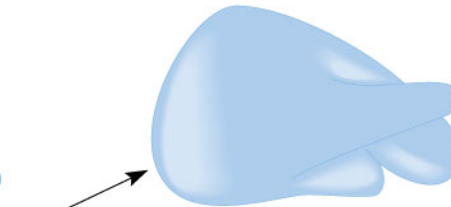
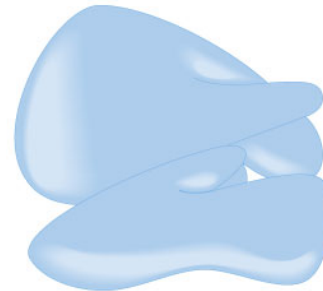
Amino acid attached by carboxyl group to ribose of last ribonucleotide of tRNA chain

Last 3 nucleotides of all tRNAs are -CCA-3'

Bakterielle und eukaryotische Ribosomen sind einander sehr ähnlich. Die RNAs im Ribosom haben strukturelle aber auch katalytische Funktion

a)

Bacterial ribosome (70S)  
( $2.5 \times 10^6$  daltons)



50S subunit



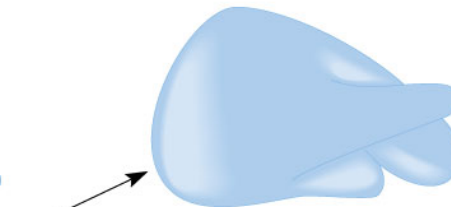
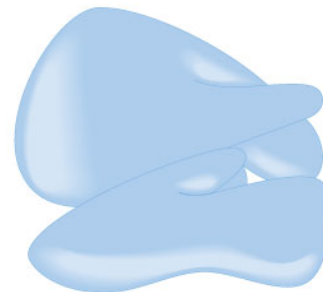
30S subunit

23S rRNA (2,904 nt)  
+  
5S rRNA (120 nt)  
+  
31 proteins

16S rRNA (1,542 nt)  
+  
21 proteins

b)

Mammalian ribosome (80S)  
( $4.2 \times 10^6$  daltons)



60S subunit



40S subunit

28S rRNA (4,718 nt)  
+  
5.8S rRNA (160 nt)  
+  
5S rRNA (120 nt)  
+  
49 proteins

18S rRNA (1,874 nt)  
+  
33 proteins

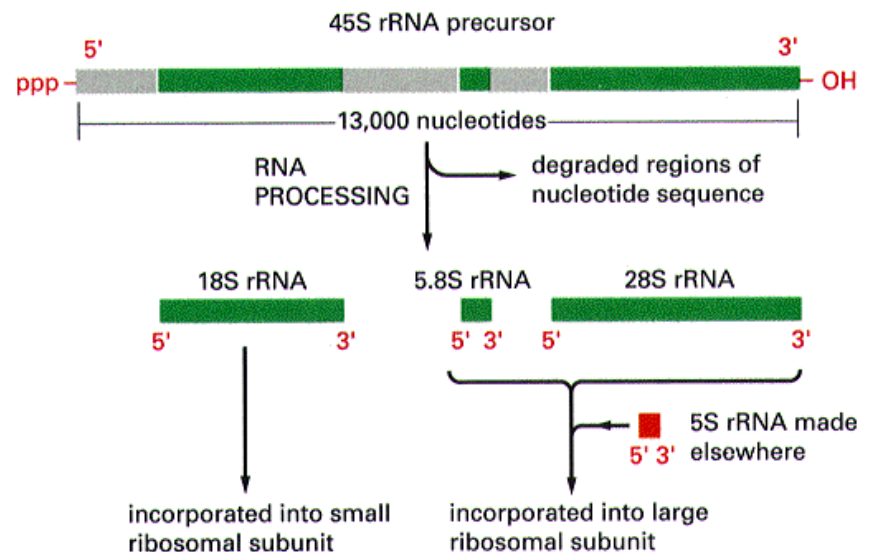
nt = nucleotides

Ribosomale RNA wird als ein precursor-Molekül transkribiert:

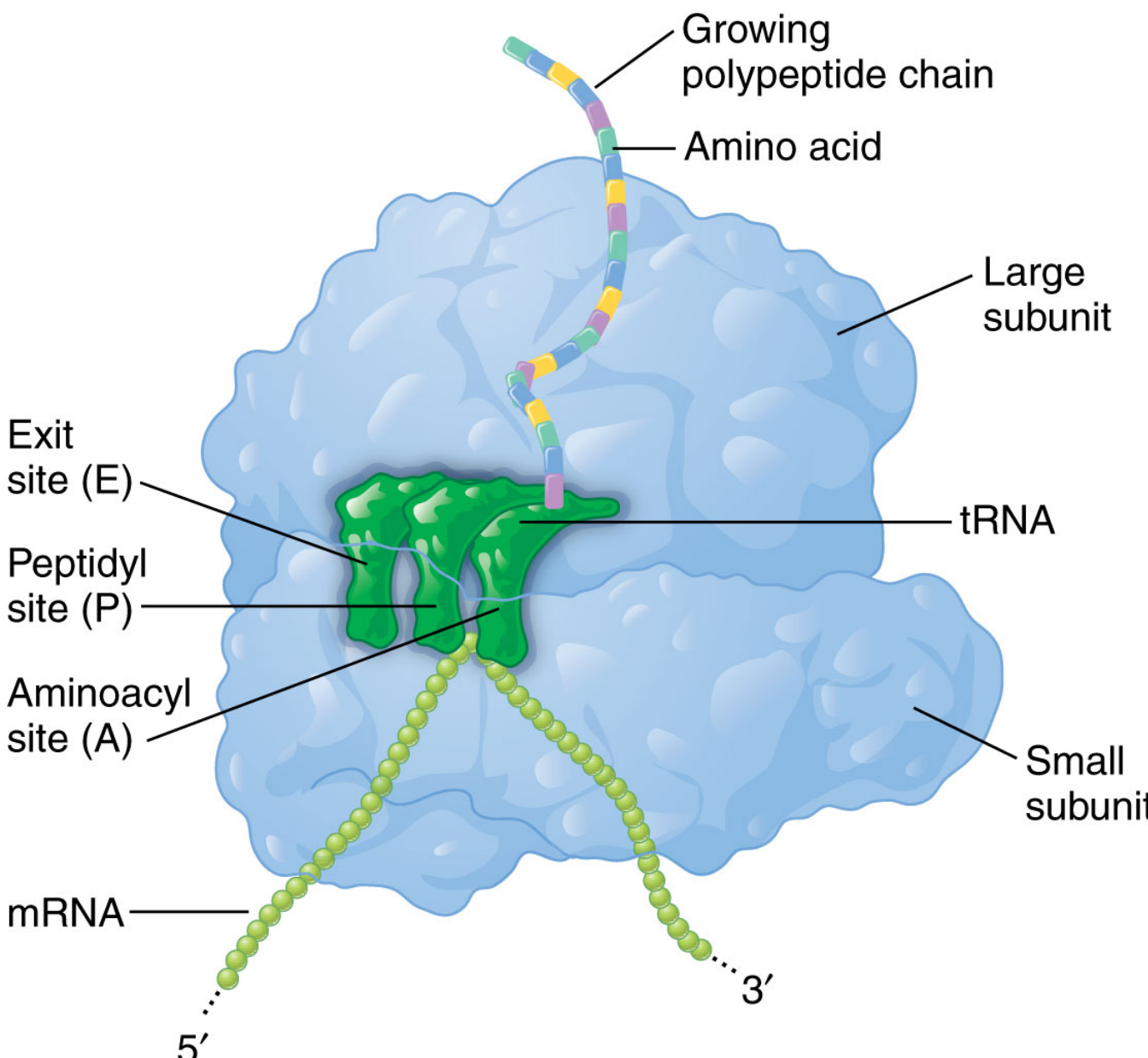
In Bakterien ist dies als 16S 23S 5S Transkript organisiert

In Eukaryoten als 18S-5.8S-28S (bzw. 25S bzw. 23 S (Pflanzen, Tiere, Pilze))

Das primäre Transkript wird in seine Untereinheiten zerschnitten (prozessiert).



# tRNA interagiert mit dem Ribosom an der A, P, und E site

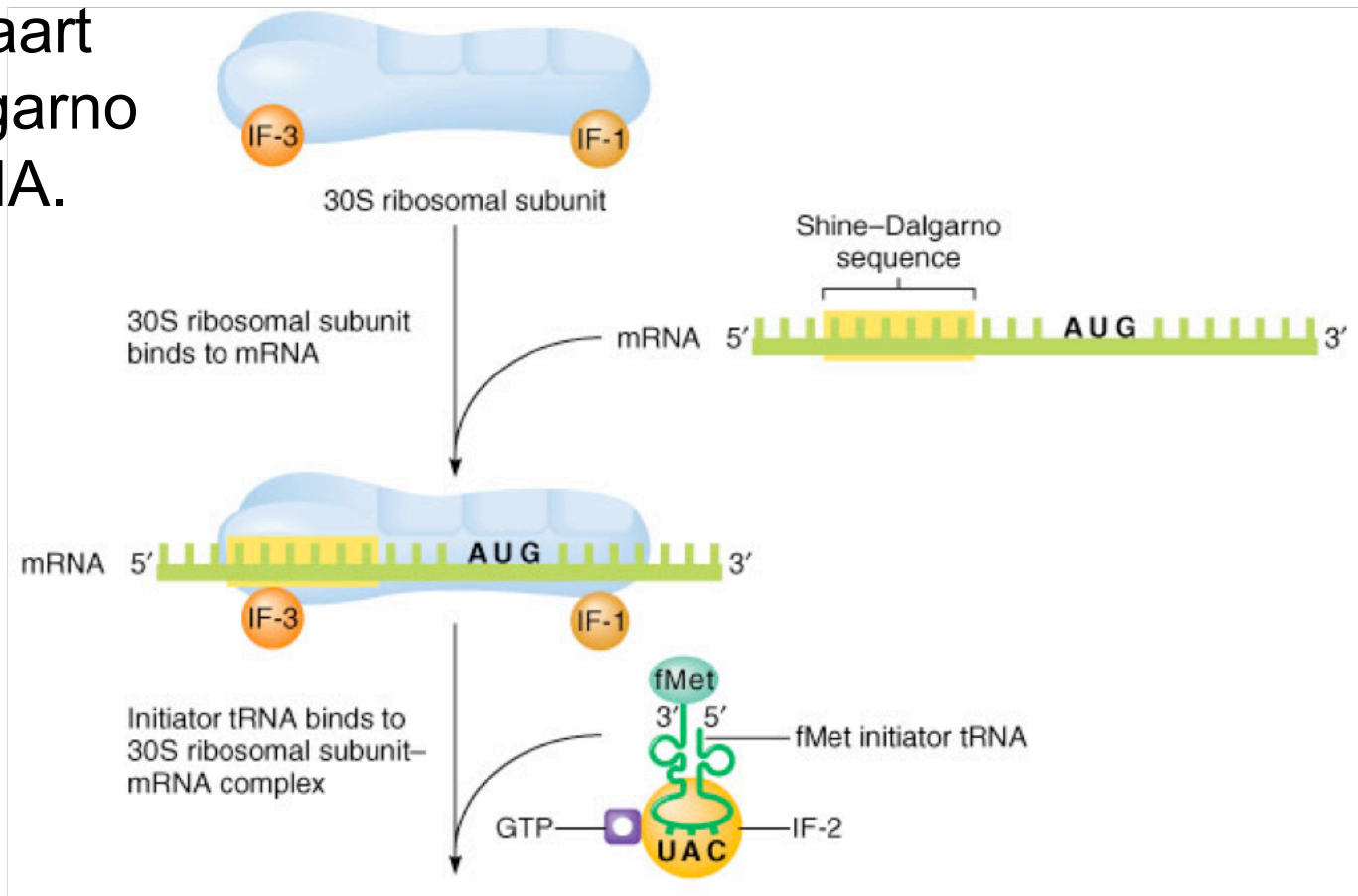


Translation wird in 3 Phasen reguliert: Initiation, Elongation, Termination

## In Bakterien:

Initiation wird von Initiationsfaktoren reguliert. Initiation Factor 1 und 3 binden die kleine ribosomale Untereinheit. Diese bindet das 5' -Ende der RNA

Die rRNA Basenpaart mit der Shine-Dalgarno Sequenz der mRNA.

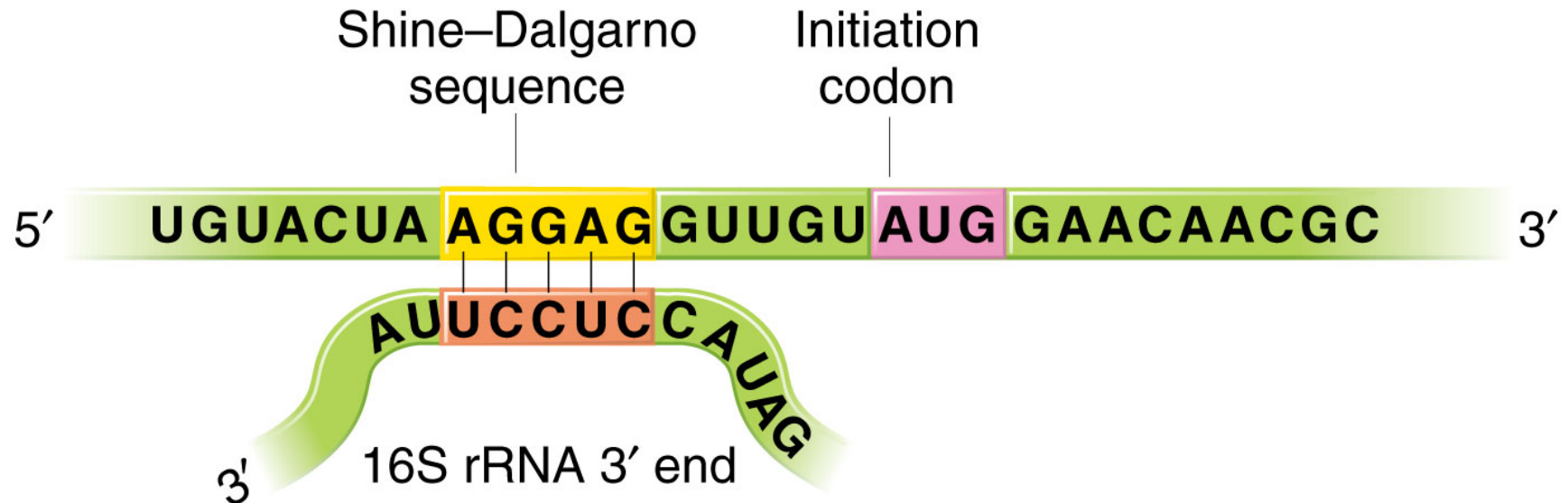


Diese Basenpaarung wurde mittels Mutationsanalyse verifiziert.

a) Sequence at 3' end of 16S rRNA

3' **AUUCCUCCA**UAG 5'

b) Example of sequence upstream of the AUG codon in an mRNA pairing with the 3' end of 16S rRNA





In Pro- und Eukaryoten beginnt die Translation mit einem Methionin (AUG) In Bakterien ist die erste Aminosäure modifiziert **fMET** (= **Formylmethionin**). fMET tRNA wird durch IF2 und GTP von normaler MET tRNA unterschieden. IF1 blockiert die A site. fMET wird über die P-site geladen. Binden der grossen Untereinheit entlässt die drei Initiationsfaktoren.

