

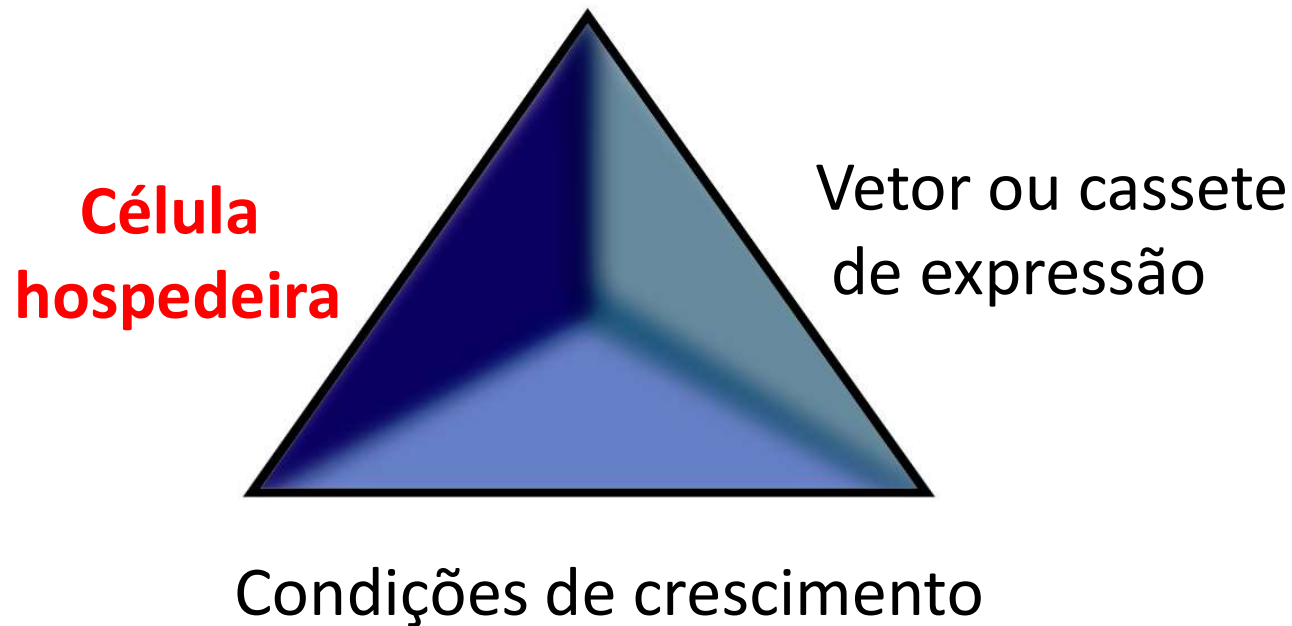


Proteínas Recombinantes – A Escolha do Sistema de Expressão

Aula 2

A escolha do sistema de expressão

O sucesso na expressão de uma proteína recombinante depende, em especial, de três fatores:



Visão geral sobre os principais sistemas de expressão de rProteínas

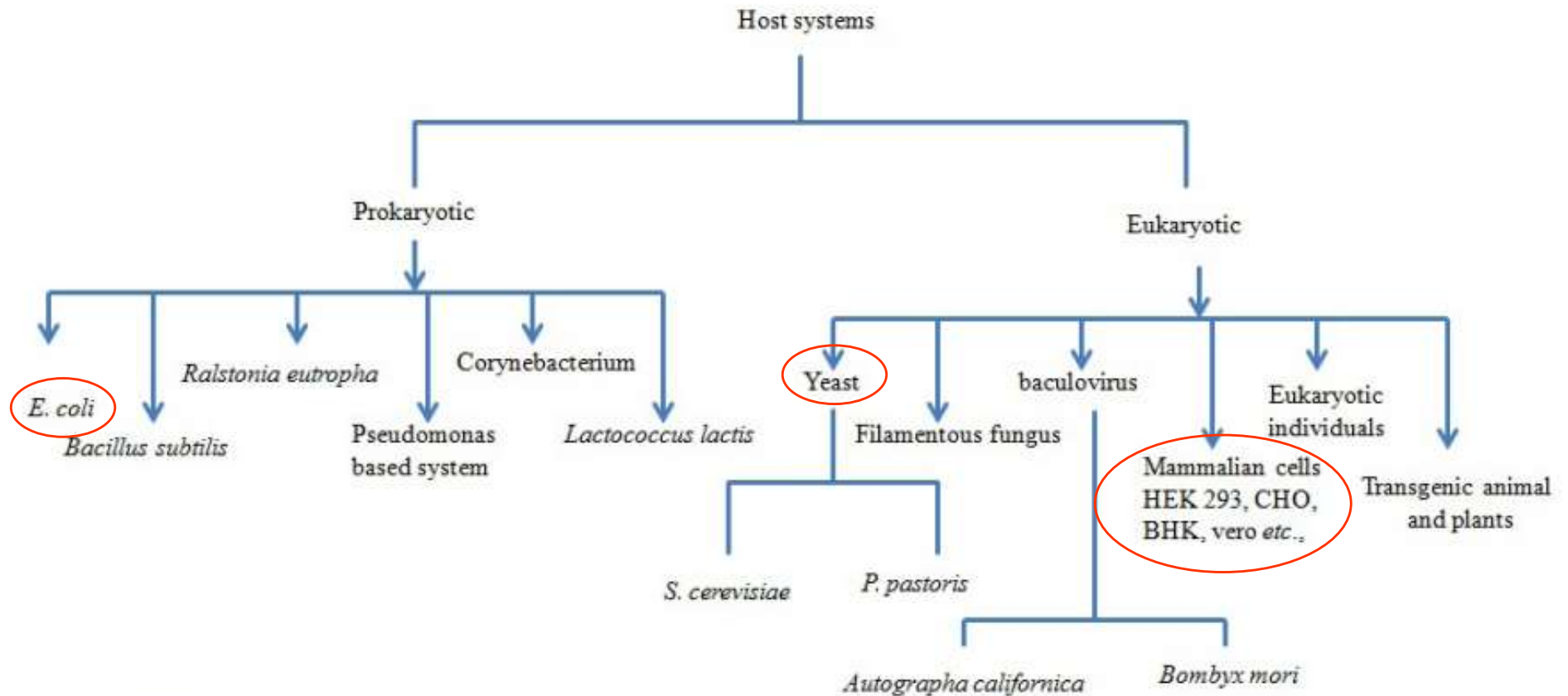
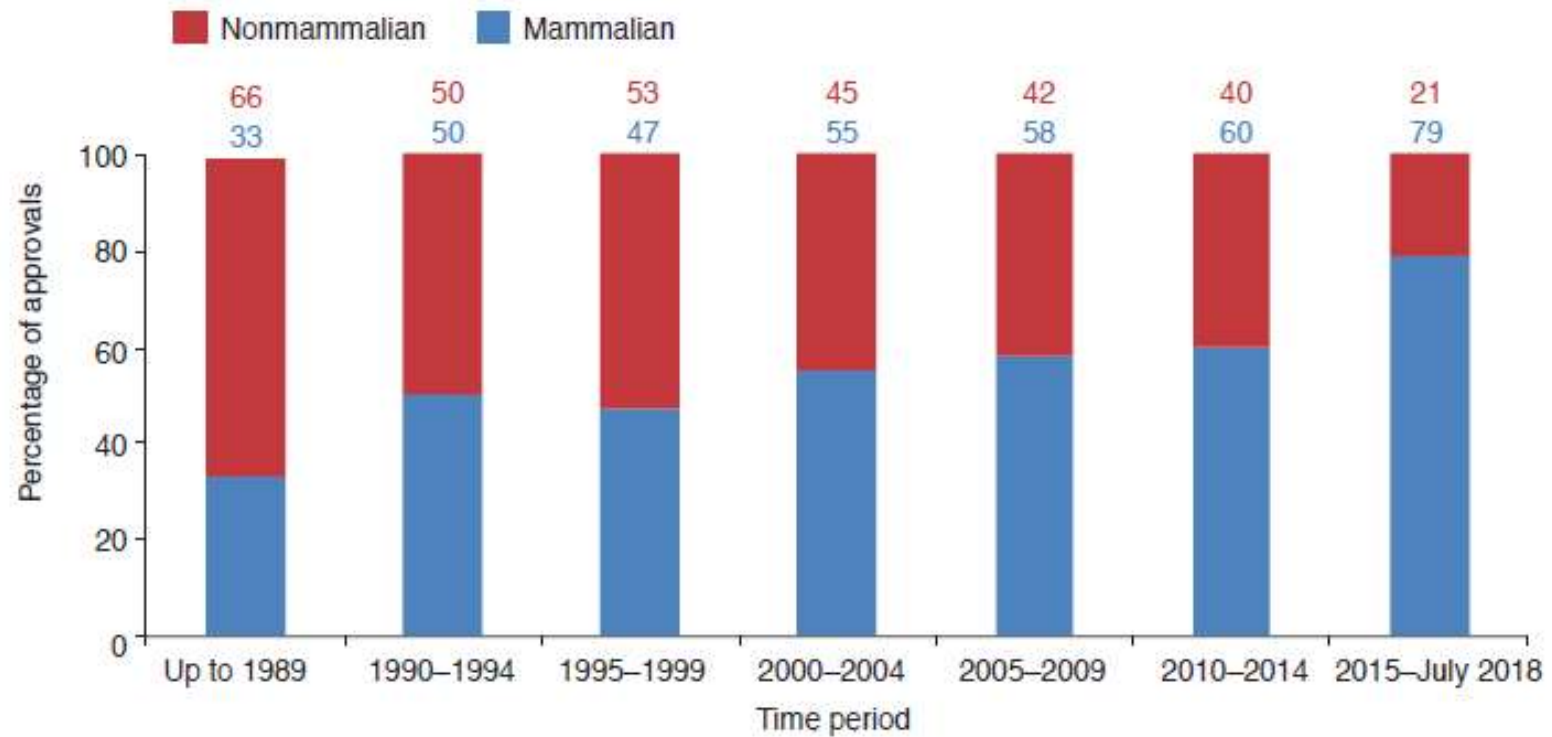


Figure 1: Different host systems available for the production of recombinant Proteins

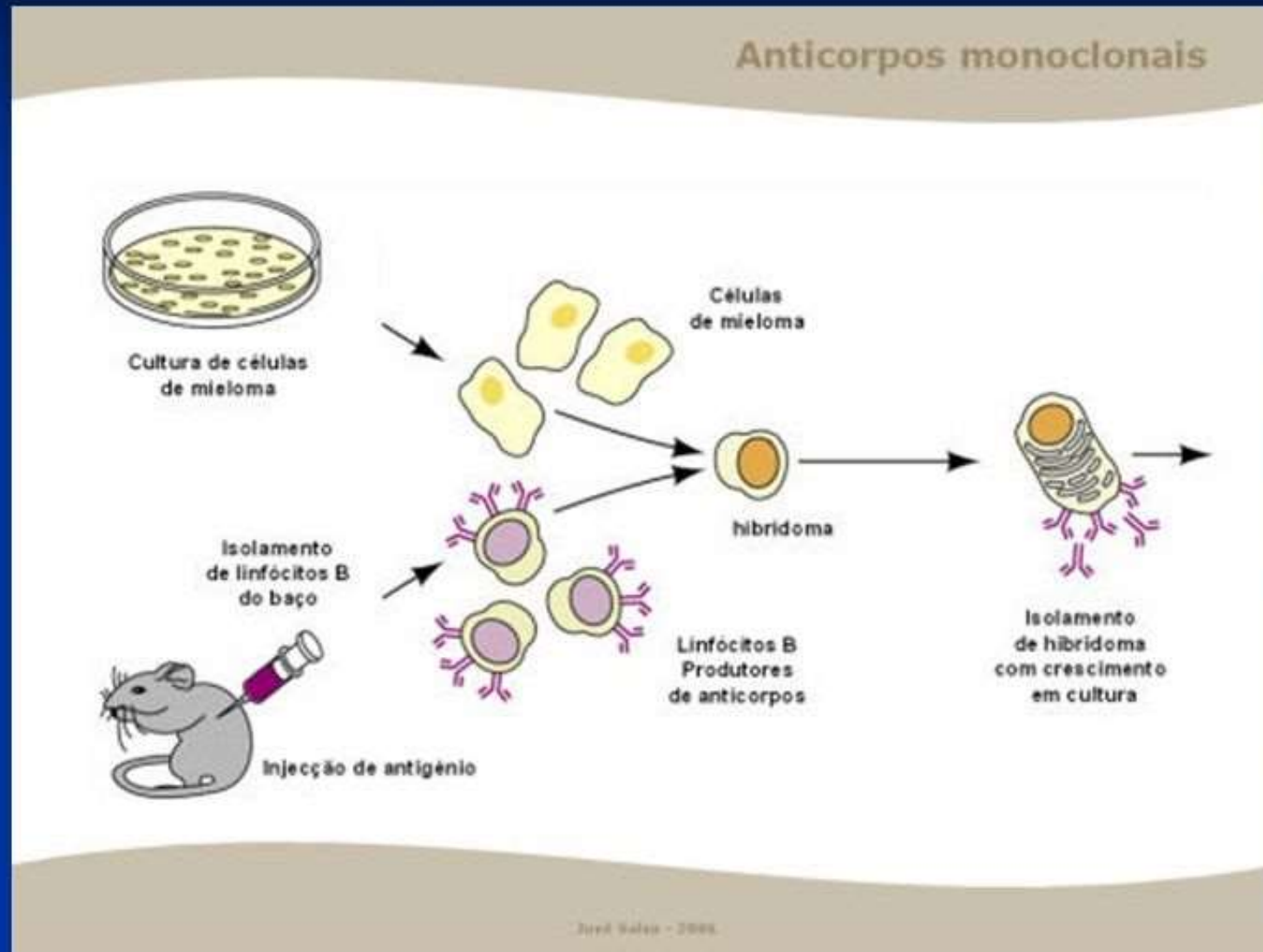
Fonte: Gomes et al., 2016

A preferência pelas células de mamíferos na produção de biofármacos...



Fonte: Walsh G., 2018

Anticorpos Monoclonais



Mieloma: câncer que se desenvolve na medula óssea, devido ao crescimento descontrolado de células plasmáticas .

Principais fatores a serem levados em conta na escolha do sistema de expressão de rProteínas:

- Conhecimento já disponível sobre o sistema na literatura
- Facilidade de manipulação genética (vetores, protocolos, etc)
- Capacidade de realizar a síntese protéica e modificações pós-traducionais
- Taxas de crescimento e nível de expressão da rProteína
- Capacidade de secretar proteínas para o meio de cultura
- Aprovação pelas agências reguladoras
- Potencial de uso comercial (custos, riscos, tecnologia disponível, etc)
- Histórico do grupo de P&D, empresa, etc.

A Informação Genética e a Síntese de Proteínas

O fluxo de informação (Dogma Central da Genética Molecular)

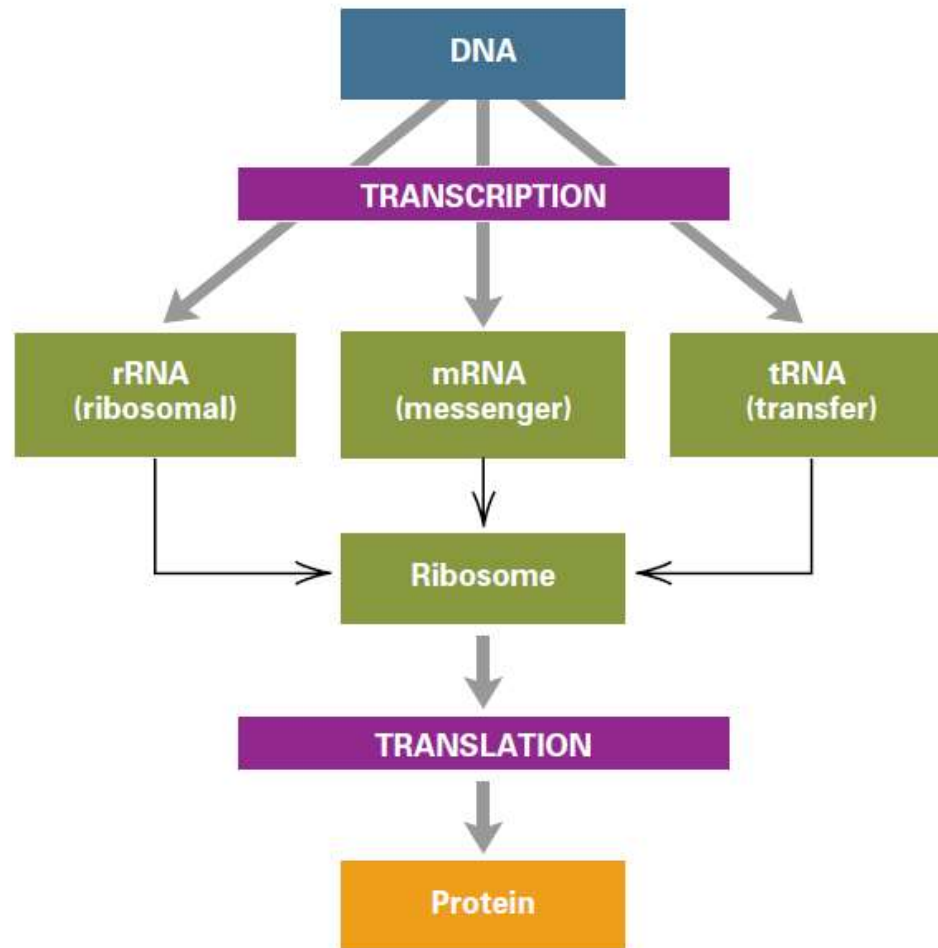
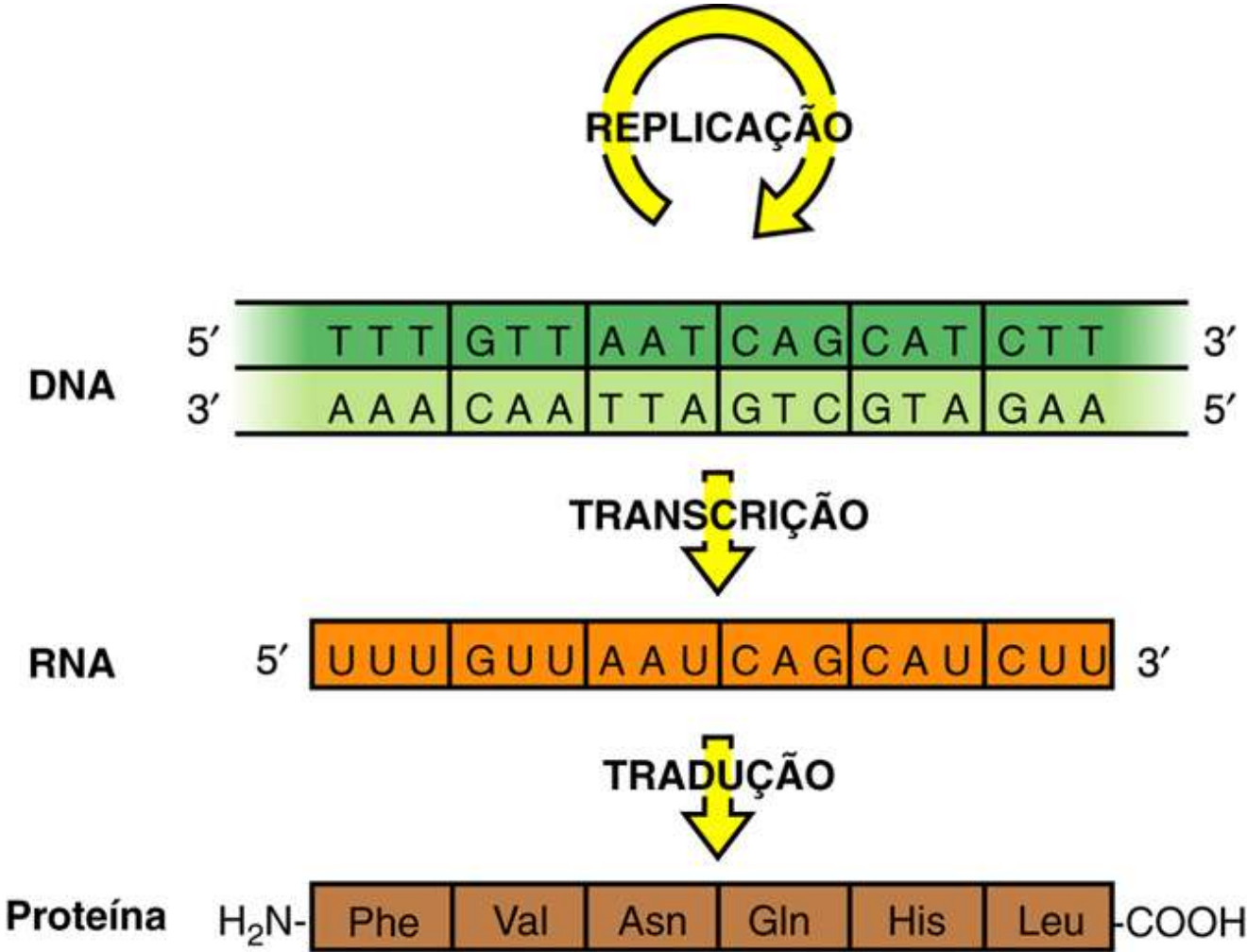


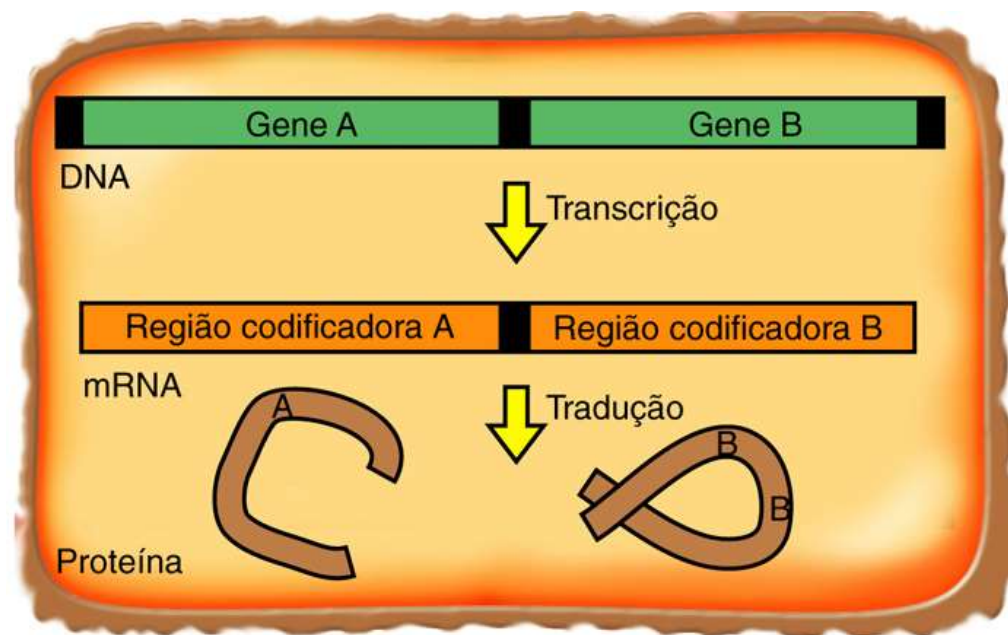
Figure 1.12 The “central dogma” of molecular genetics: DNA codes for RNA, and RNA codes for protein. The DNA → RNA step is transcription, and the RNA → protein step is translation.

O fluxo da informação genética

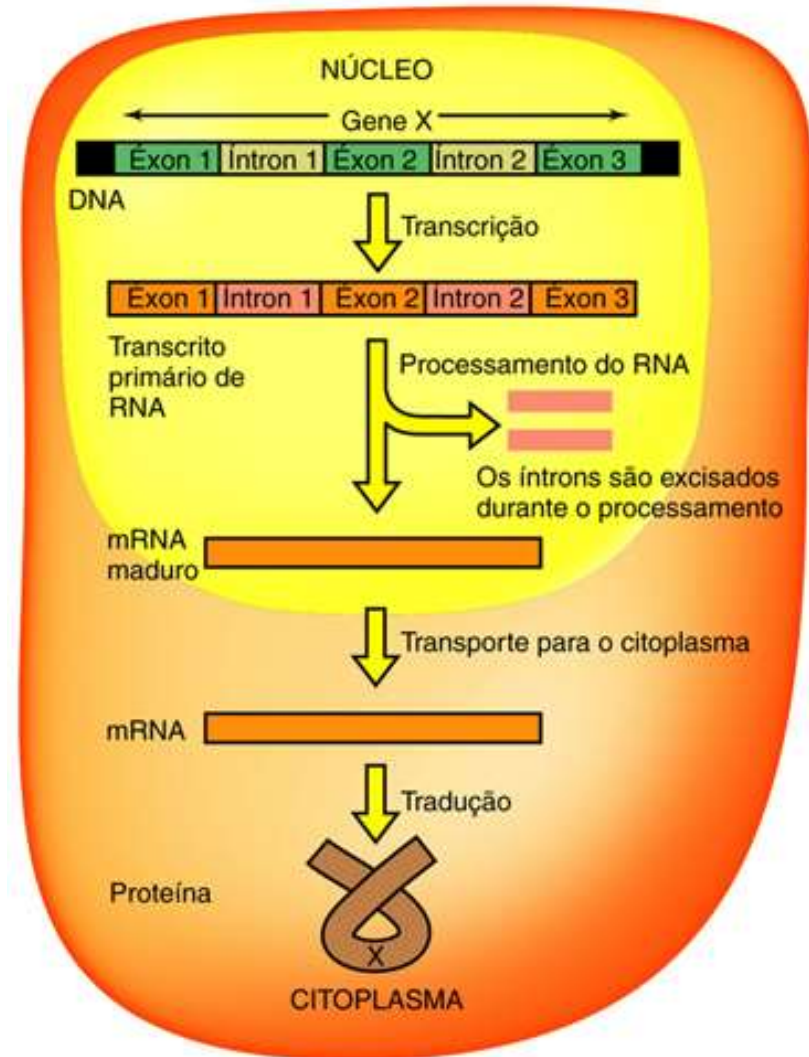


Informação Genética e Proteínas

Transferência de informação: Procariotos vs Eucariotos



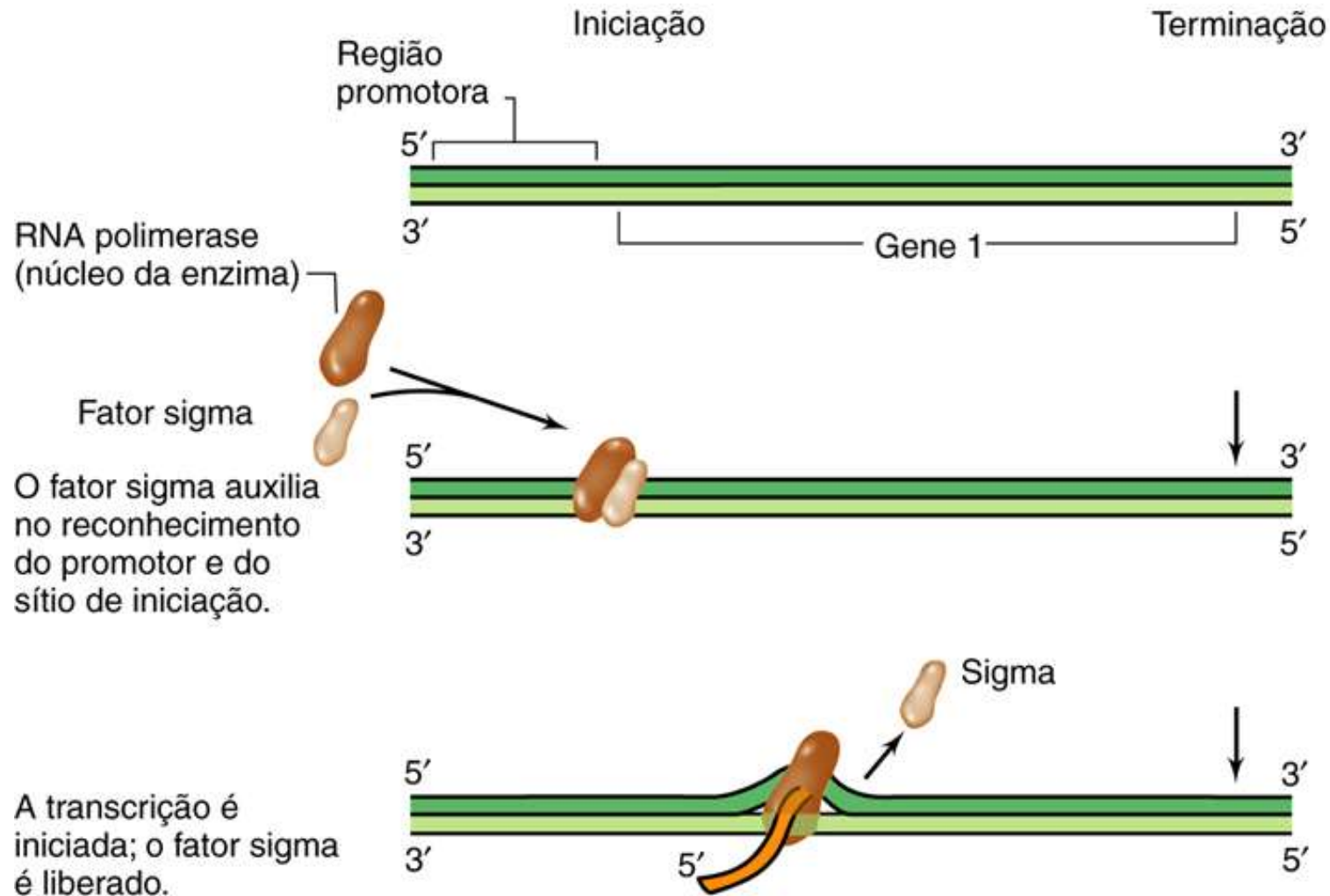
(a) Procariotos



(b) Eucariotos

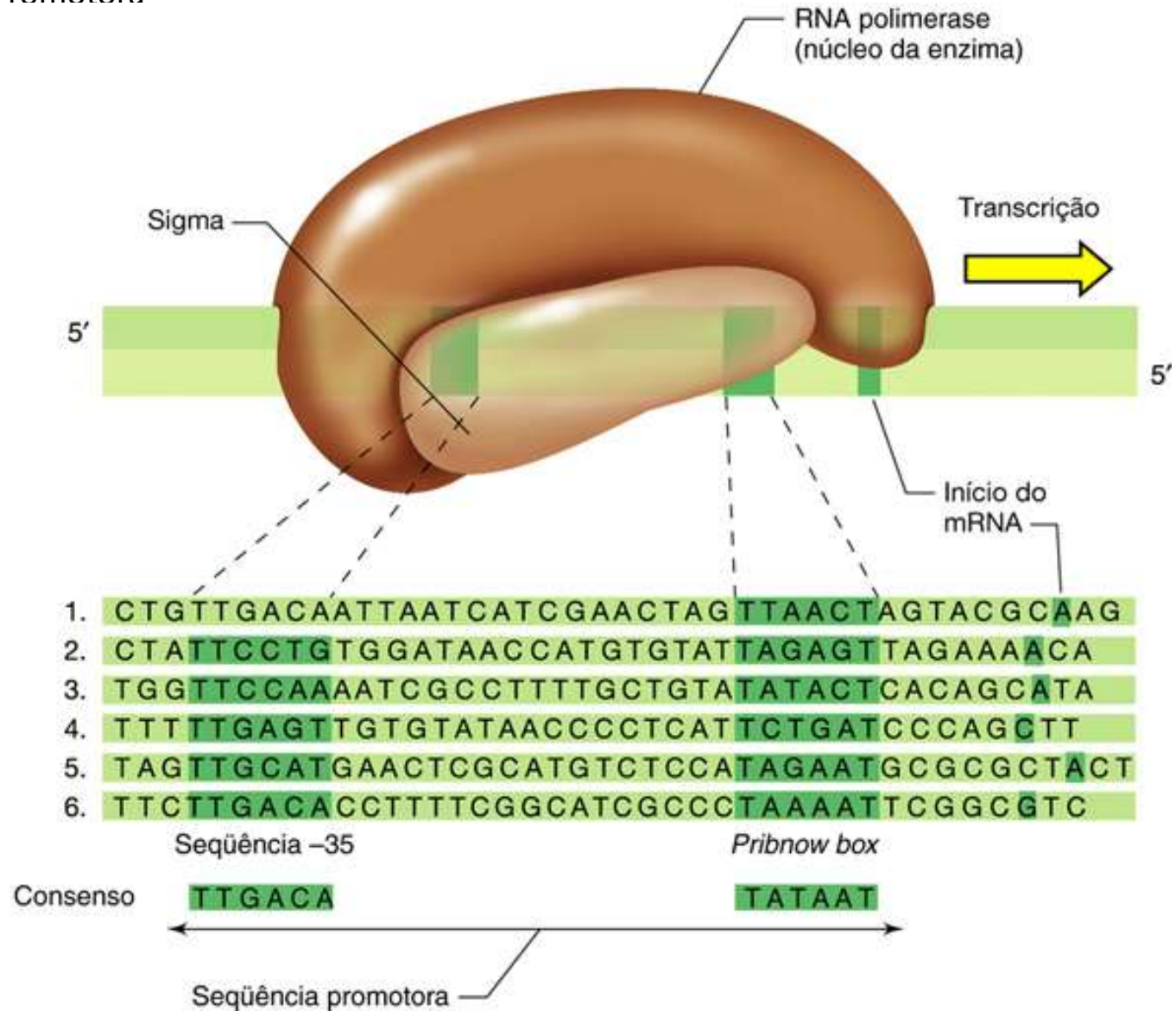
Transcrição

A transcrição de RNA a partir do DNA envolve a enzima RNA polimerase. Essa é formada por várias subunidades protéicas, sendo o “fator sigma” a subunidade que auxilia a enzima a reconhecer a sequência de DNA onde a transcrição terá início (região promotora, localizada perto do início do gene).



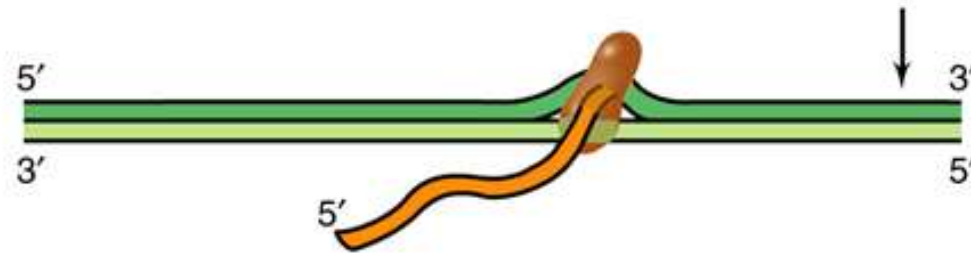
Transcrição

Região Promotora

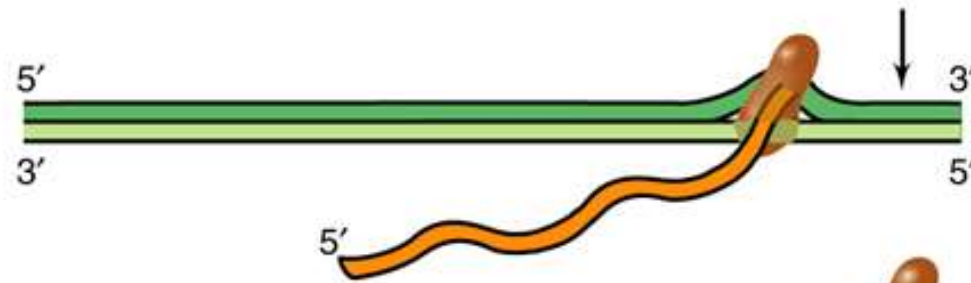


Continuação...

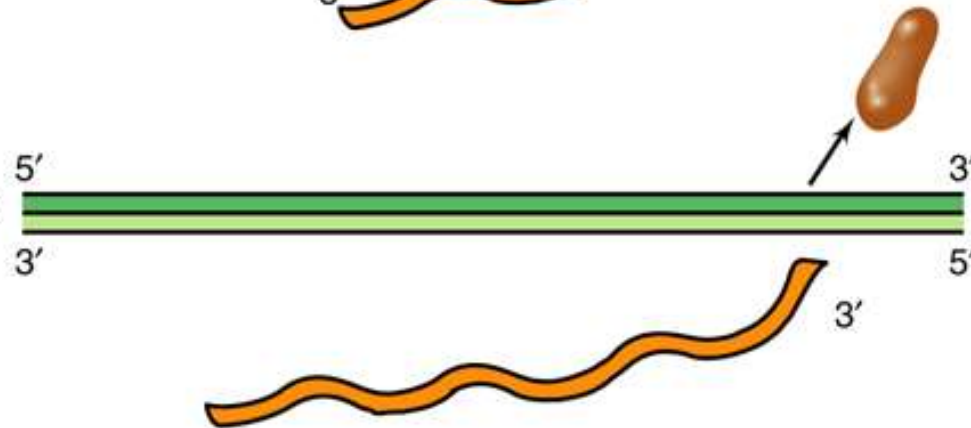
Crescimento da fita de RNA.



O sítio de terminação é alcançado; cessa o crescimento da fita.

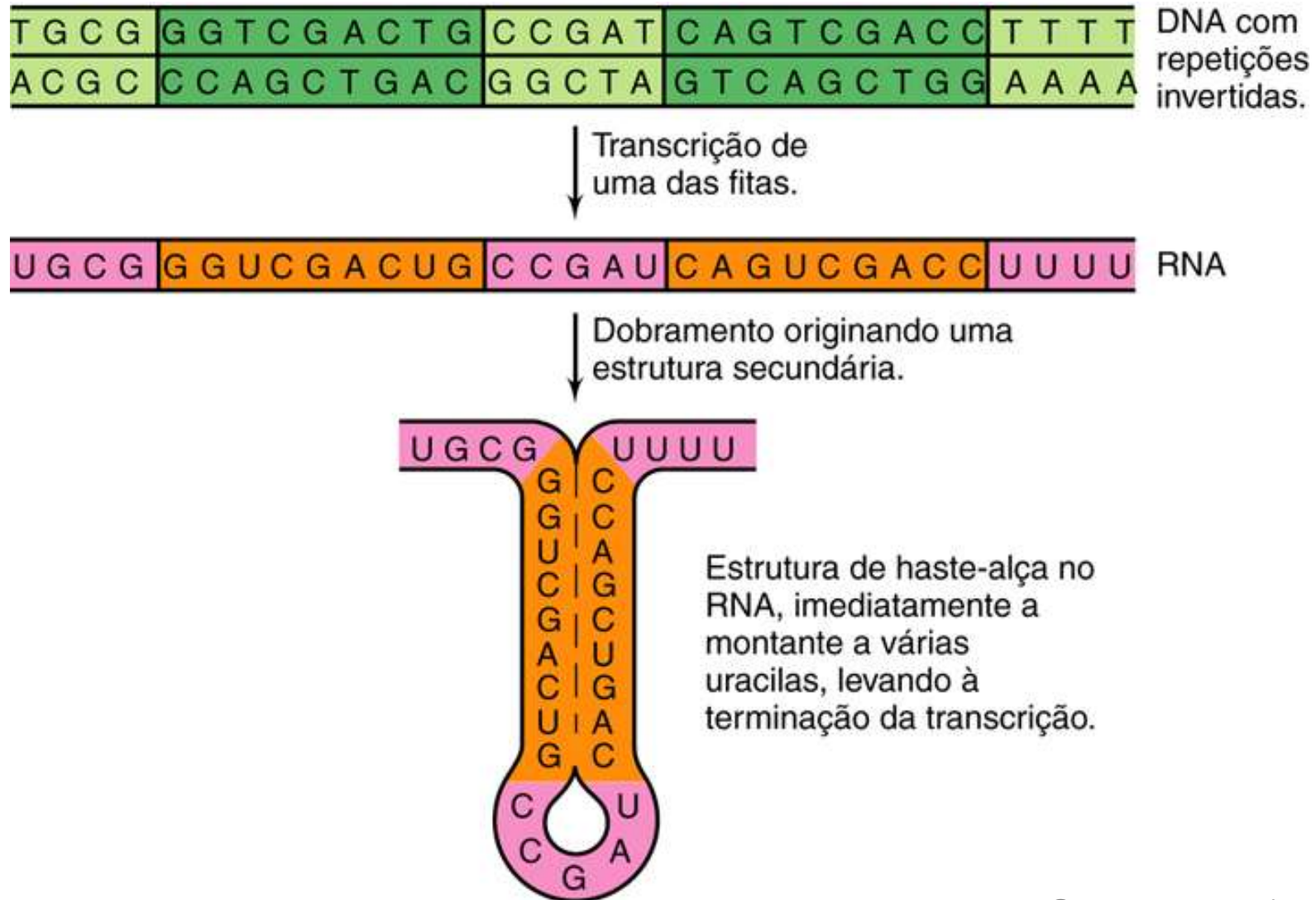


Liberação da polimerase e do RNA.

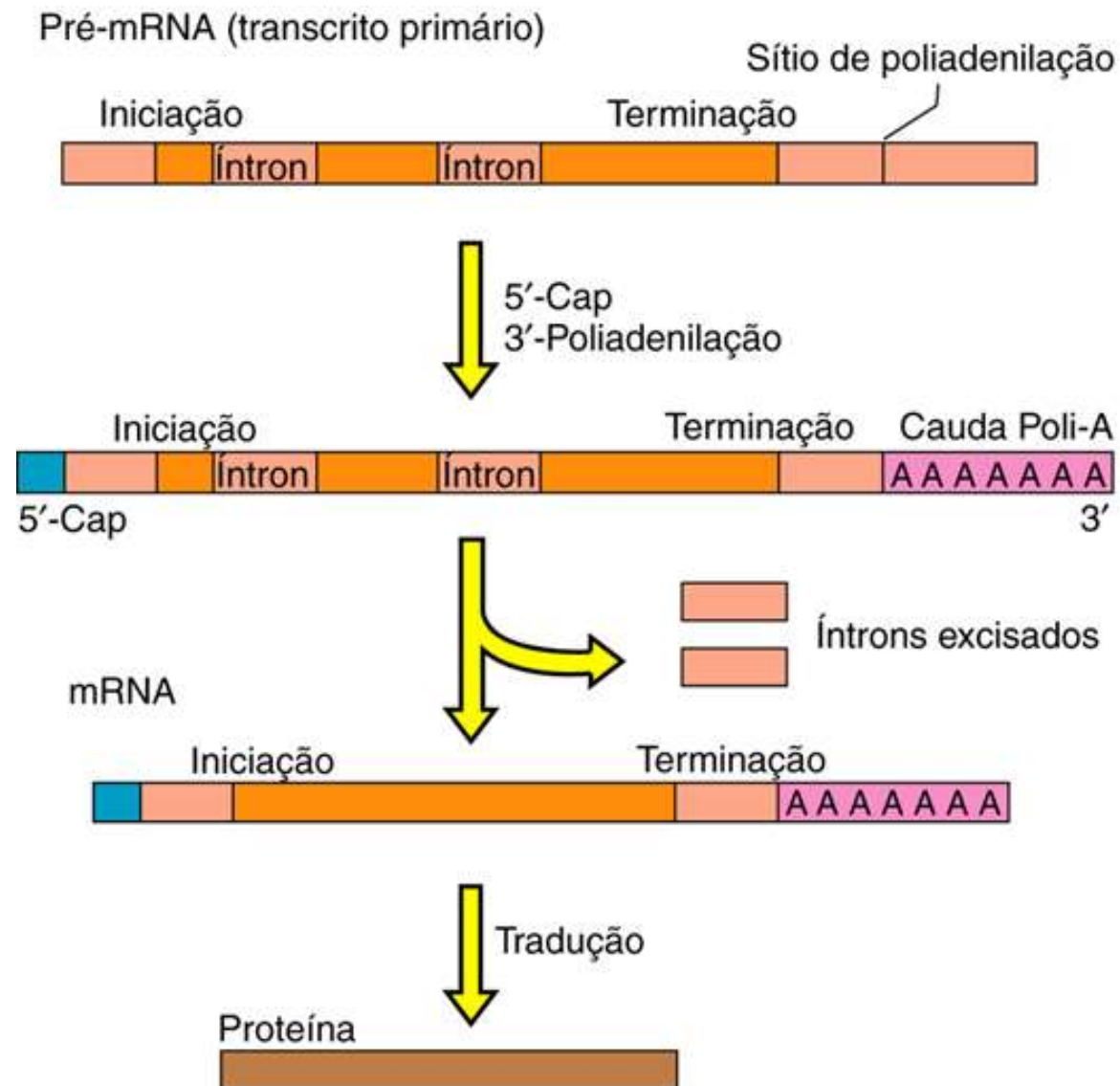


(a)

A terminação da Transcrição



O processamento de RNAs (continuação)



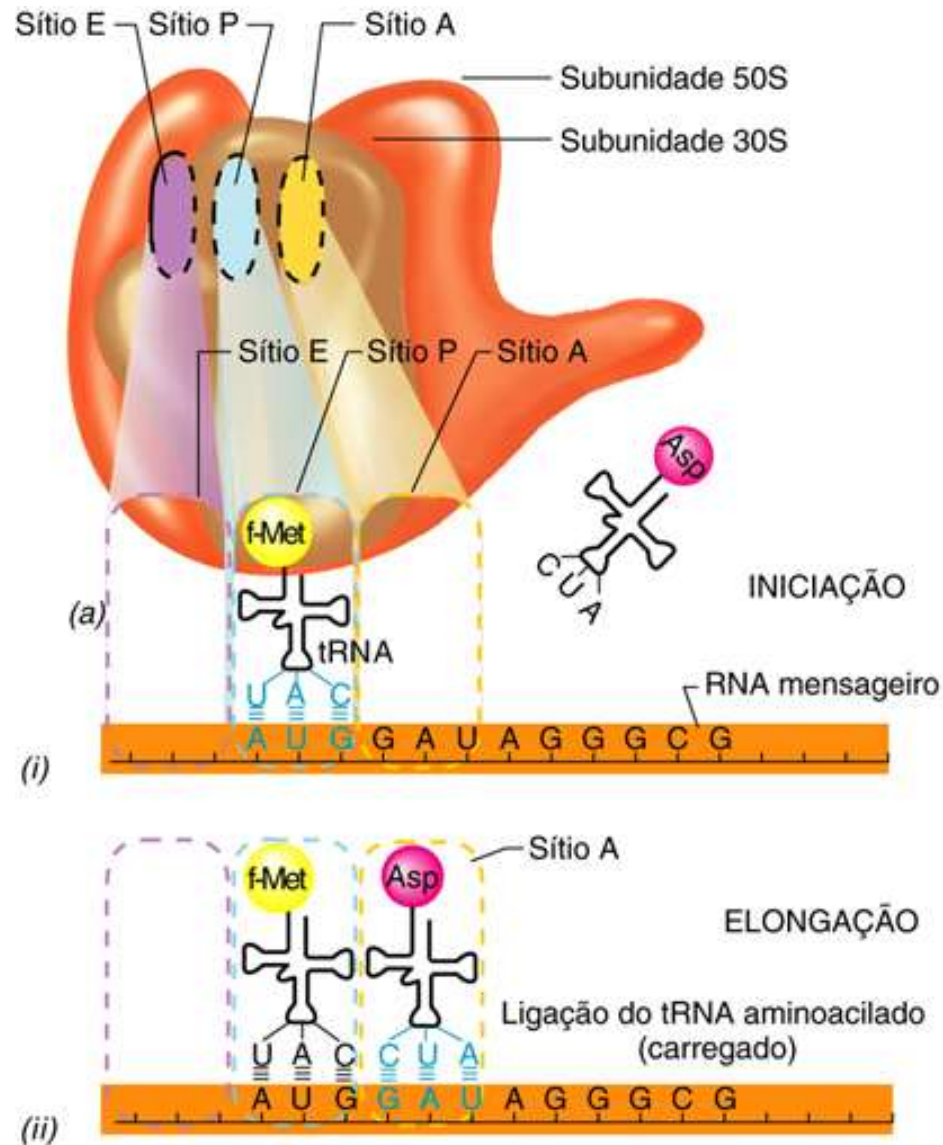
Tradução

Código Genético

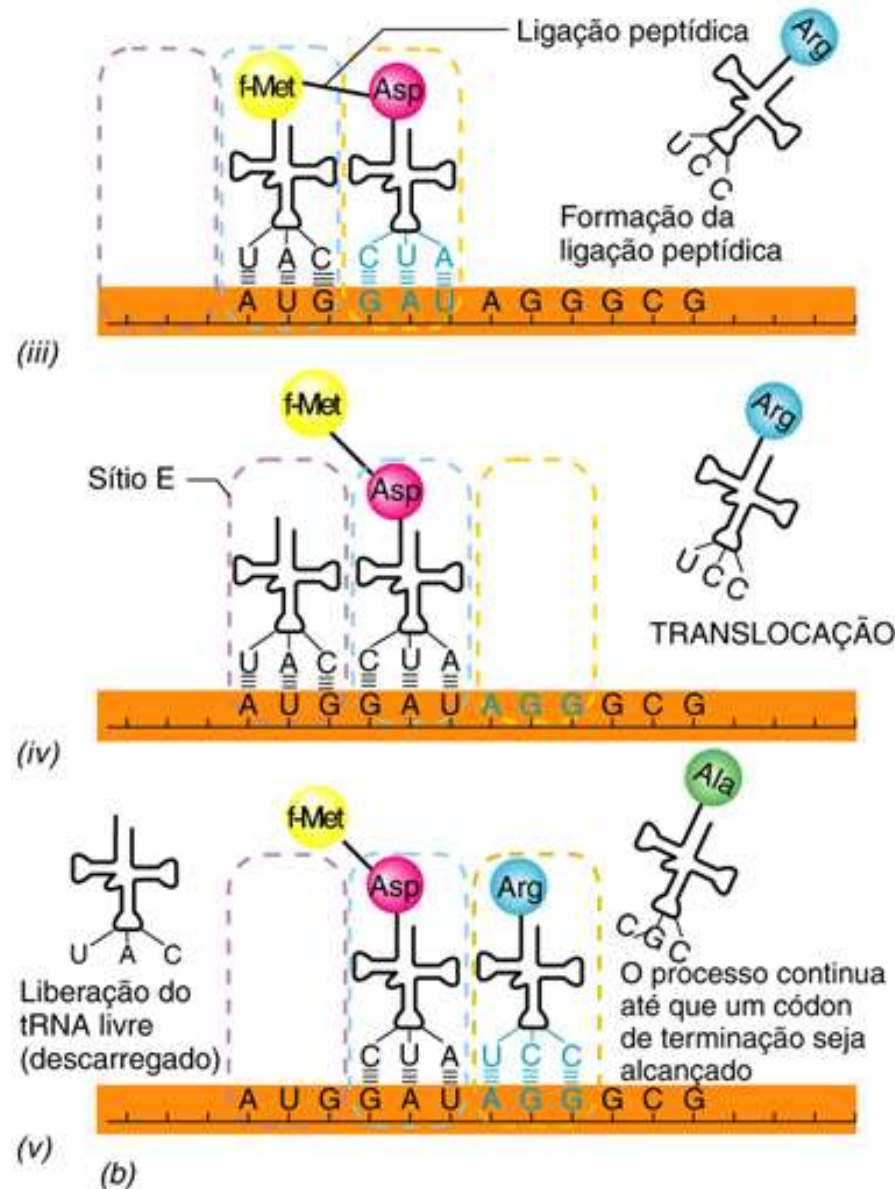
É dito degenerado ou redundante pelo fato de existir, para um mesmo aminoácido, mais de uma trinca de bases para codificá-lo.

		Second Letter					
		T	C	A	G		
First Letter	T	TTT } Phe TTC } TTA } Leu TTG }	TCT } Ser TCC } TCA } TCG }	TAT } Tyr TAC } TAA Stop TAG Stop	TGT } Cys TGC } TGA Stop TGG Trp	T C A G	
	C	CTT } Leu CTC } CTA } CTG }	CCT } Pro CCC } CCA } CCG }	CAT } His CAC } CAA } Gln CAG }	CGT } Arg CGC } CGA } CGG }	T C A G	
	A	ATT } Ile ATC } ATA } ATG Met	ACT } Thr ACC } ACA } ACG }	AAT } Asn AAC } AAA } Lys AAG }	AGT } Ser AGC } AGA } Arg AGG }	T C A G	
	G	GTT } Val GTC } GTA } GTG }	GCT } Ala GCC } GCA } GCG }	GAT } Asp GAC } GAA } Glu GAG }	GGT } Gly GGC } GGA } GGG }	T C A G	

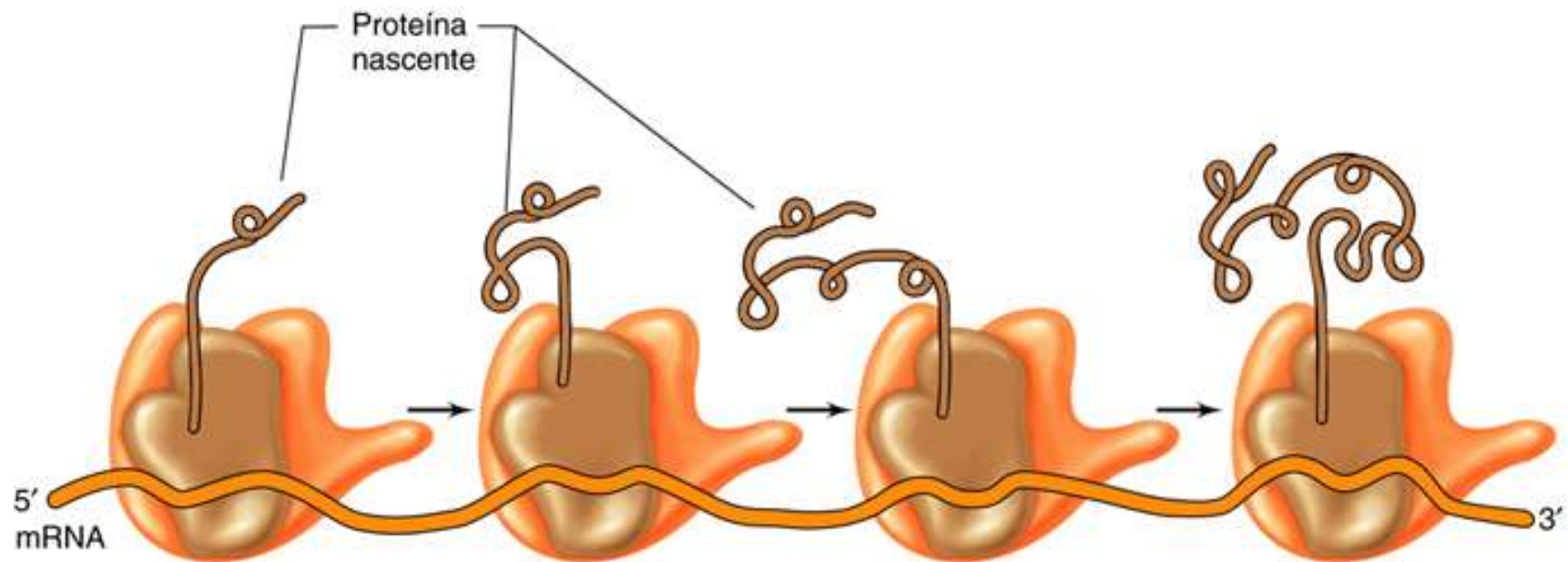
O processo de síntese de proteínas nos ribossomos



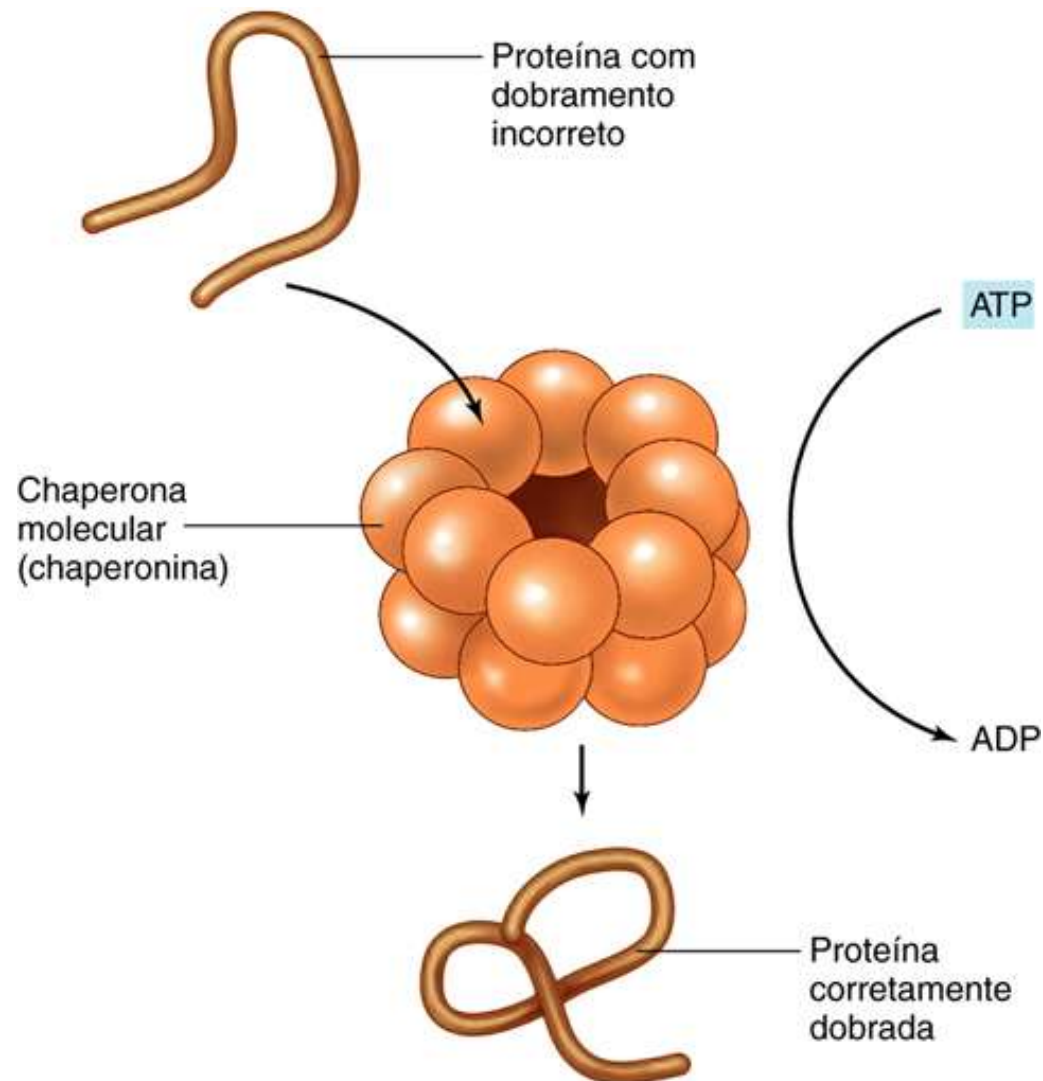
O processo de síntese de proteínas no ribossomo (continuação)



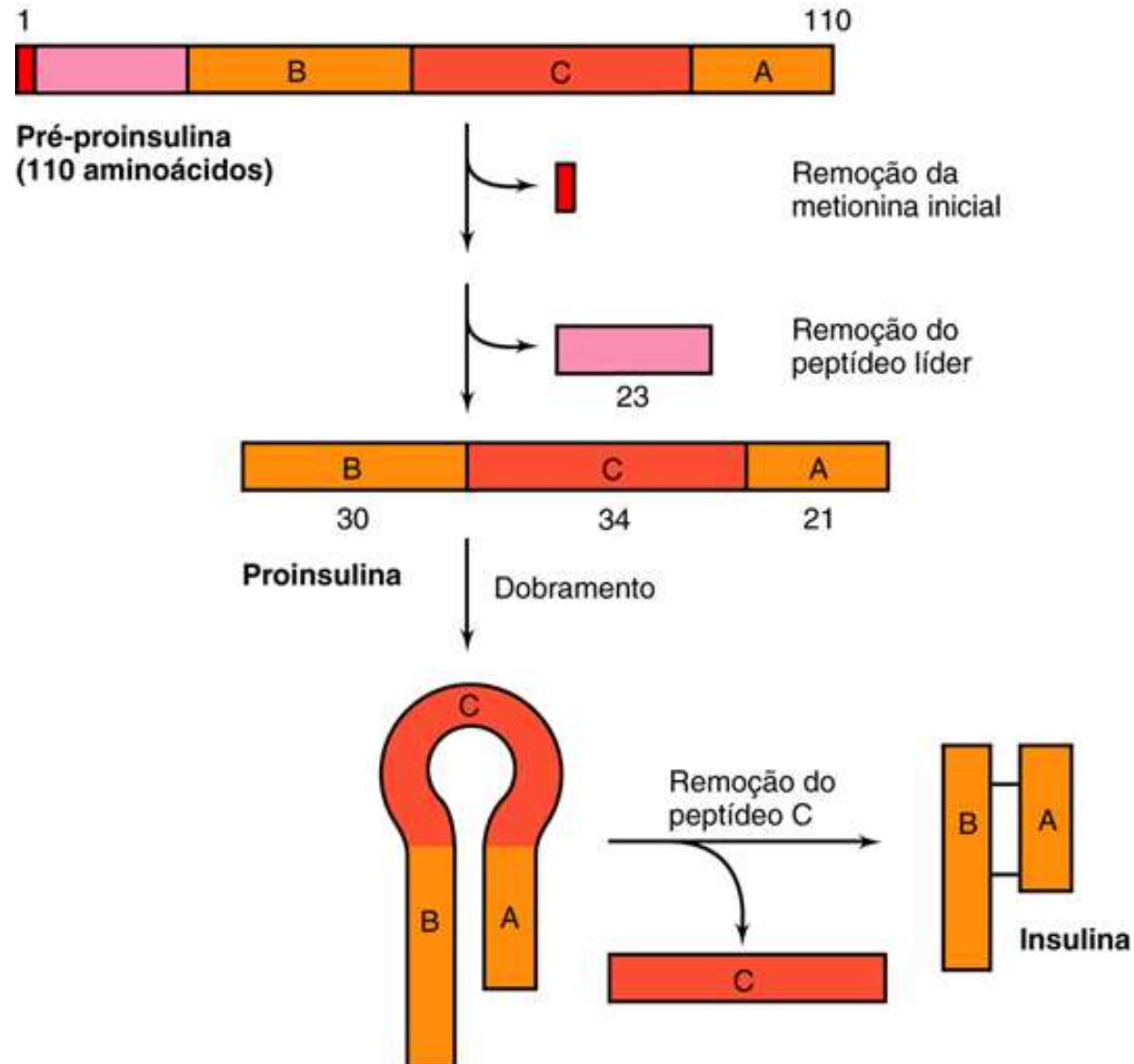
O processo de síntese de proteínas (continuação)



O enovelamento (*folding*) de uma proteína pode ser “finalizado” por chaperonas (complexo protéico com a função de auxiliar o enovelamento).



Muitas vezes, modificações pós-traducionais ainda são necessárias para a obtenção da nova proteína na sua forma ativa (Exemplo: insulina)



Modificações Pós-Traducionais

Table 2.7 Types of PTM that polypeptides may undergo. Refer to text for additional details

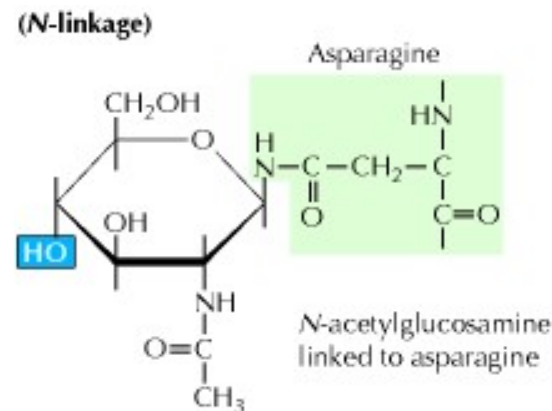
Modification	Example
Proteolytic processing	Various proteins become biologically active only upon their proteolytic cleavage (e.g. some blood factors)
Glycosylation	For some proteins glycosylation can increase solubility, influence biological half-life and/or biological activity
Phosphorylation	Influences/regulates biological activity of various polypeptide hormones
Acetylation	Function unclear
Acylation	May help some polypeptides interact with/anchor in biological membranes
Amidation	Influences biological activity/stability of some polypeptides
Sulfation	Influences biological activity of some neuropeptides and the proteolytic processing of some polypeptides
Hydroxylation	Important to the structural assembly of certain proteins
γ -Carboxyglutamate formation	Important in allowing some blood proteins to bind calcium
ADP-ribosylation	Regulates biological activity of various proteins
Disulfide bond formation	Helps stabilize conformation of some proteins

Modificações Pós-Traducionais - Glicosilação

A glicosilação é um processo enzimático complexo (glicosil transferases) que leva a ligação de um polissacarídeo (glicano) a uma molécula de proteína, lipídeo ou outra molécula orgânica. Estima-se que 50% das proteínas secretadas sejam glicosiladas. Proteínas glicosiladas estão envolvidas em uma grande variedade de processos biológicos críticos, como direcionamento, adesão celular, ligações proteína-proteína, etc.

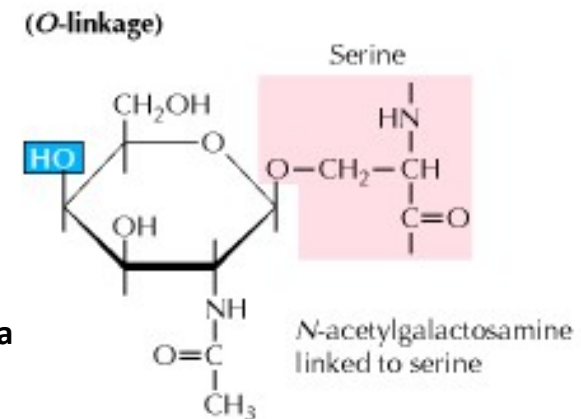
- Glicosilação *N-Linked*:

Sequência de reconhecimento
Asn-X-Thr/Ser



- Glicosilação *O-Linked*:

Resíduos de **serina** ou **treonina**



O papel das glicosilações nas Proteínas

Table 2.8 The potential roles and effects of the glycocomponent of glycoproteins. Reproduced from Walsh, G. and Jefferies, R. (2006). *Nature Biotechnology* **24**, 1241–1252

Role/effect	Comment
Protein folding	Glycosylation can effect local protein secondary structure and help direct folding of the polypeptide chain
Protein targeting/trafficking	The glycocomponent can participate in the sorting/ directing of a protein to its final destination
Ligand recognition/binding	The carbohydrate content of antibodies, for example, plays a role in antibody binding to monocyte F _c receptors and interaction with complement component C1 _q
Biological activity	The carbohydrate side chain of gonadotrophins is essential to the activation of gonadotrophin signal transduction
Stability	Sugar side chains can potentially stabilize a glycoprotein in a number of ways, including enhancing its solubility, shielding hydrophobic patches on its surface, protection from proteolysis and by direct participation in intrachain stabilizing interactions
Regulates protein half-life	High levels of sialic acid (a family of acidic sugars that often caps sugar side chains) can increase a glycoprotein's plasma half-life. Exposure of galactose residues can decrease plasma half-life by promoting uptake via hepatic galactose residues. Yeast glycosylation is of a 'high mannose' type, which can also drive rapid removal from circulation via specific cell-surface mannose receptors
Immunogenicity	Some glycosylation motifs characteristic of plant-derived glycoproteins (often containing fucose and xylose residues) are highly immunogenic in mammals

A questão da glicosilação na escolha do sistema de expressão de rProteínas

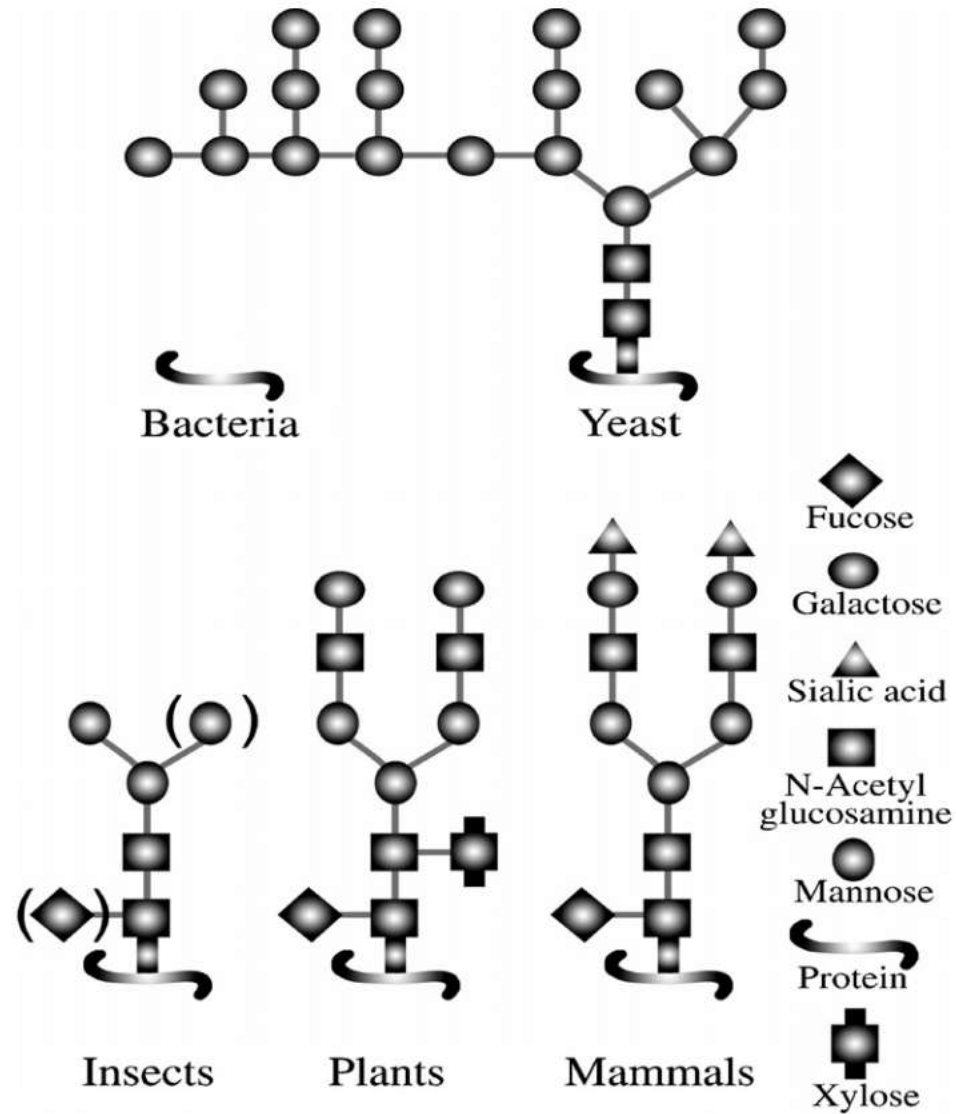
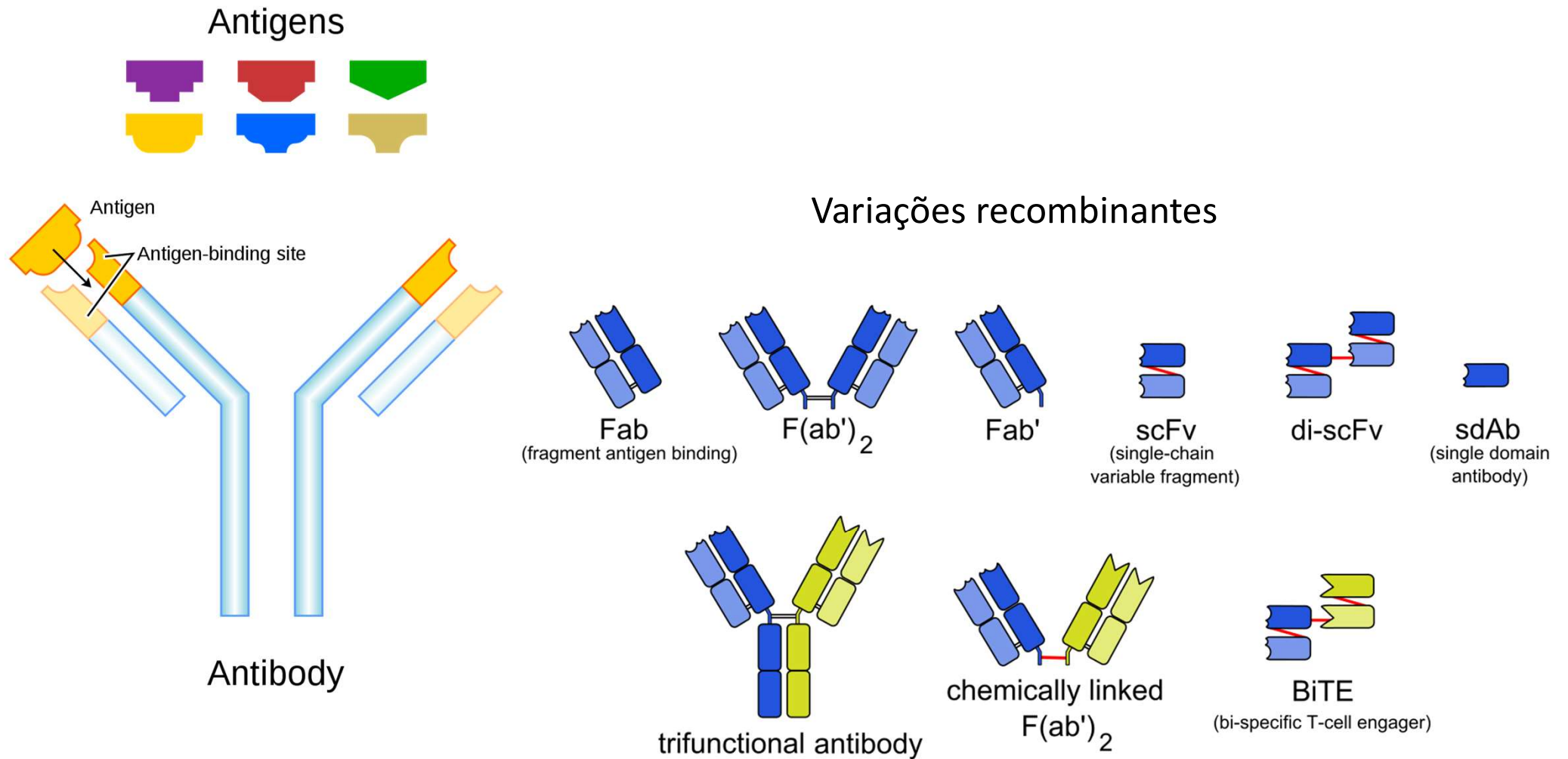


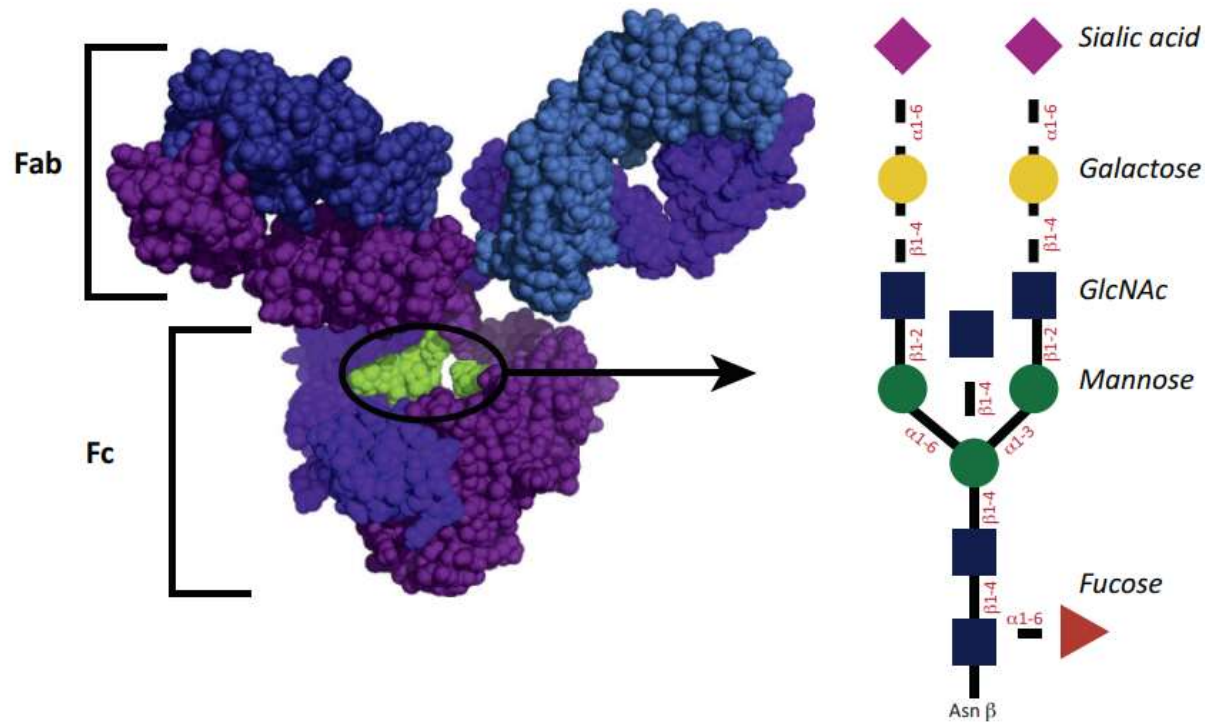
Fig. 1. Comparison of the glycosylation of recombinant proteins produced by different systems.

Exemplo de biofármaco glicosilado (ou não): **Anticorpos monoclonais**



Glicosilação em anticorpos monoclonais

Trends in Immunology



Trends in Immunology

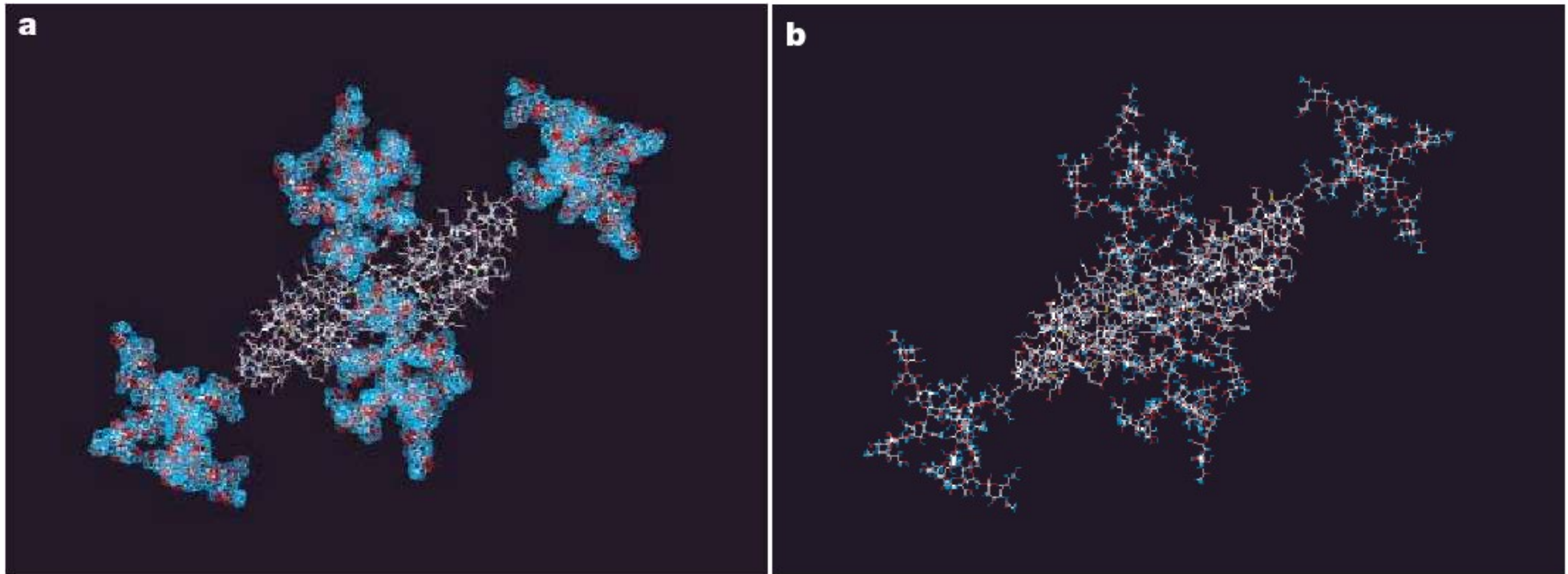
Figure 1. Antibody and Glycan Structure. IgG antibodies are composed of two heavy chains (purple) and two light chains (blue), which together form two functional domains: the antigen-binding (Fab) and crystallizable/constant domains (Fc). Each heavy chain Fc contains a glycan at Asn297. The glycan contains up to 13 monomers of N-acetylglucosamine (GlcNAc, blue squares), fucose (red triangle), mannose (green circles), galactose (yellow circles), and sialic acid (pink diamonds). The linkages orientation of the monomers are indicated in red. PDB accession code 1IGY.

Fonte: Jennewein and Alter (2017)

Exemplo de glicoproteína - Glicocerebrosidase

Com glicosilação

Sem glicosilação



Fonte: Stefan Wildt, Tillman U. Gerngross. **The humanization of N-glycosylation pathways in yeast.** Nature Reviews Microbiology 2005.

Videos: A Informação Genética e a Síntese de Proteínas

Organelas Celulares, Transcrição e Tradução de Proteínas

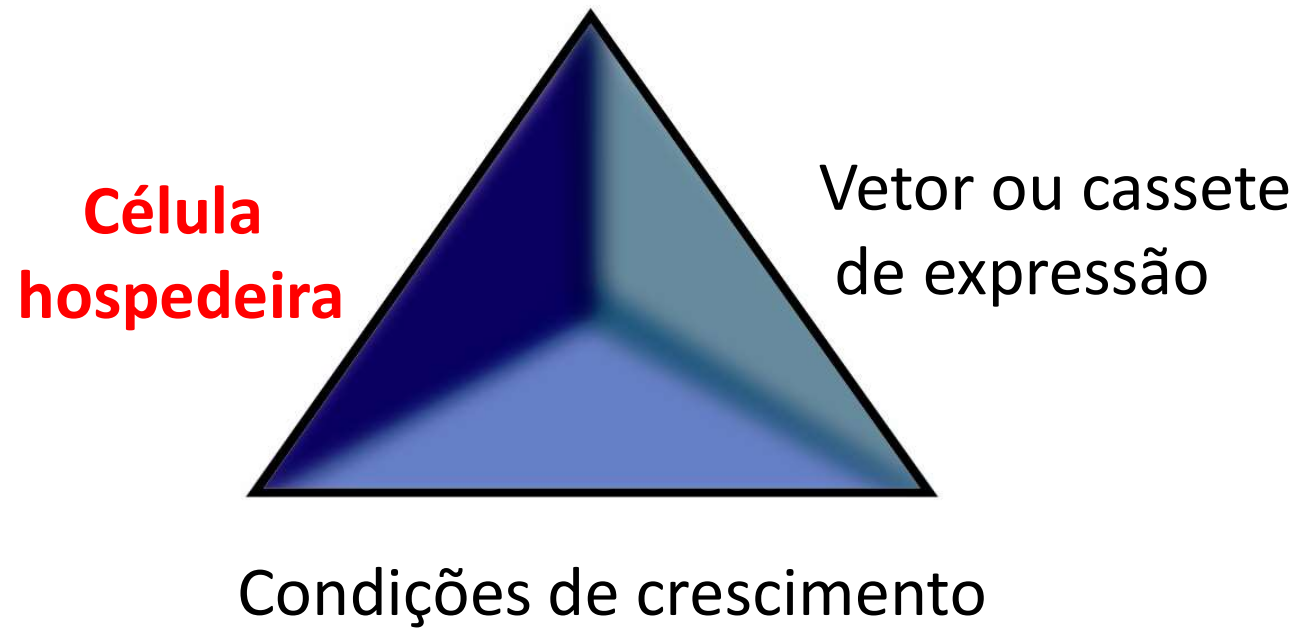
<https://www.youtube.com/watch?v=gG7uCskUOrA>

(Do ADN à proteína 3D)

<https://www.youtube.com/watch?v=URUJD5NEXC8>

(Biology: cell structure medical media)

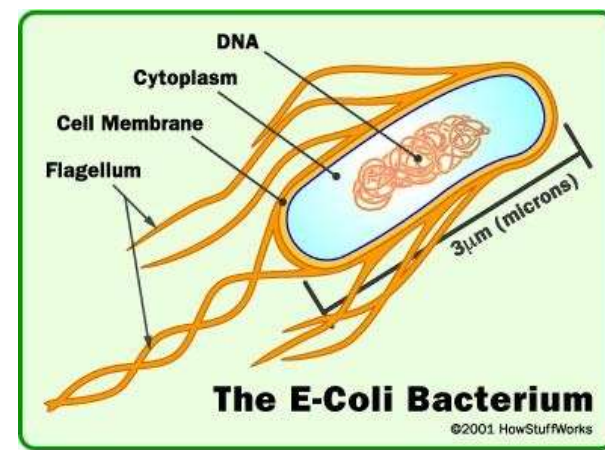
A escolha do sistema de expressão



Escherichia coli como sistema de expressão de rProteínas

Principais vantagens

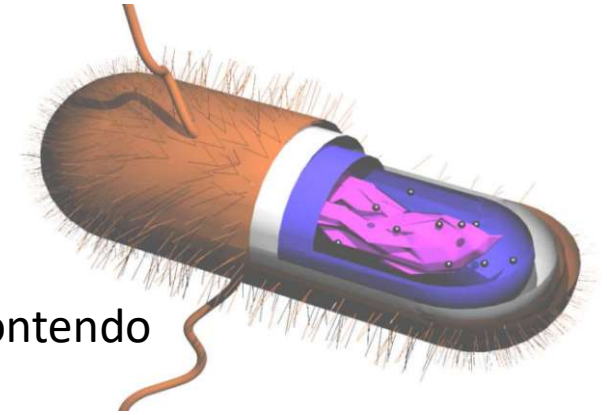
- Amplamente estudada do ponto de vista biológico e de processo
- Rápido crescimento, elevada expressão (20-30% da proteína total)
- Linhagens seguras (não patogênicas)
- Mutantes próprias para clonagem
- Mutantes próprias para expressão de rProteínas
 - Deficientes em proteases
 - Próprias para códons raros (tRNAs)
 - Sistema redox modificado
 - Controle rígido sobre a expressão
 - Crescimento sem antibióticos



Principais vantagens (continuação)

- Amplamente usada para produção comercial (geralmente pequenas proteínas como a insulina e o glucagon ou fragmentos de anticorpos, entre outros)
- Vários processos e produtos aprovados pelo FDA e EMEA (desde 1982)
- Até meados da década de 90 era o principal sistema de expressão da indústria farmacêutica





Principais desvantagens

- Limitações quanto à expressão de proteínas complexas contendo modificações pós-traducionais
- Ambiente redutor do citoplasma freqüentemente leva a enovelamentos incorretos
- Presença de códons raros pode afetar expressão, mesmo em linhagens específicas (ex. BL21(DE3)Rosetta)
- Limitada eficiência em secretar as proteínas produzidas (alguns exemplos de sucesso existem)
- Manutenção dos vetores (plasmídeos) exige presença de um gene de seleção (ex.: resistência à antibióticos)

Leveduras como sistema de expressão de rProteínas (*Saccharomyces cerevisiae* e *Pichia pastoris*)



S. cerevisiae

Principais vantagens

- Bem caracterizadas do ponto de vista da fisiologia celular e genética (principalmente *S. cerevisiae*)
- Frequentemente usadas na indústria (*S. cerevisiae*) e centros de pesquisa (*P. pastoris*) para a expressão de proteínas recombinantes
- *S. cerevisiae* possui status de GRAS (“generally recognized as safe”) dado pelo FDA
- Vetores comercialmente disponíveis e de simples manipulação
- Mutantes próprios para expressão de rProteínas, com mutações que permitem seleção e manutenção de plasmídeos
- Possibilidade de produção intracelular ou extracelular
- Introdução de diferentes tipos de modificações pós-traducionais (glicosilação *N-linked* e *O-linked*, pontes dissulfeto, clivagem, etc)

Principais vantagens (continuação)

- Não patogênicas e não pirogênicas
- Meio de cultura de relativo baixo custo, quando comparado à células de mamífero
- Crescimento em altas densidades utilizando-se biorreatores



Principais desvantagens

- Ainda que capaz de realizar modificações pós-traducionais, o padrão (tipo e número) pode diferir daquele feito por mamíferos
- *P. pastoris*: preocupação com “explosion-proof equipments”
- Menores taxas de crescimento quando comparadas à *E. coli*
- Relativa baixa capacidade de expressão quando comparadas à expressão em *E. coli* (em geral, mas nem sempre)
- Degradação da rProteína por proteases quando secretada para o meio
- Secreção de rProteínas maiores que 30 kDa é pouco eficiente (principalmente *S. cerevisiae*)
- Maiores necessidades nutricionais se comparado à *E. coli*



P. pastoris

Comparação entre algumas leveduras já aprovadas para uso industrial

Table 3 Comparison of some yeast characteristics (according to Swinkels et al. 1993, modified and extended)

Characteristic	<i>Sac. cerevisiae</i>	<i>P. pastoris</i>	<i>K. lactis</i>	<i>H. polymorpha</i>
Industrial application	+ ^a	+ ^b	+ ^c	+ ^d
Requirement of explosion-proofed equipment	–	+	–	+
Food grade (GRAS)	+	–	+	–
Secretion efficiency	–	+	+	+
Hyperglycosylation	+	–	–	–
Episomal vector stability	+	–	+	–
Protease activity in secretion vesicles	High	Low	Low	Low

^aFor examples, see Table 1

^bRecombinant proteomics-grade trypsin production at Roche (Müller et al. 2002)

^cBovine prochymosin (MAXIREN) in a 40 m³ unit at Gist–Brocades (van den Berg et al. 1990)

^dHepatitis B vaccine at Rhein-Biotech (Schult-Linkholt et al. 2002)

Fonte: Schmidt (2004)

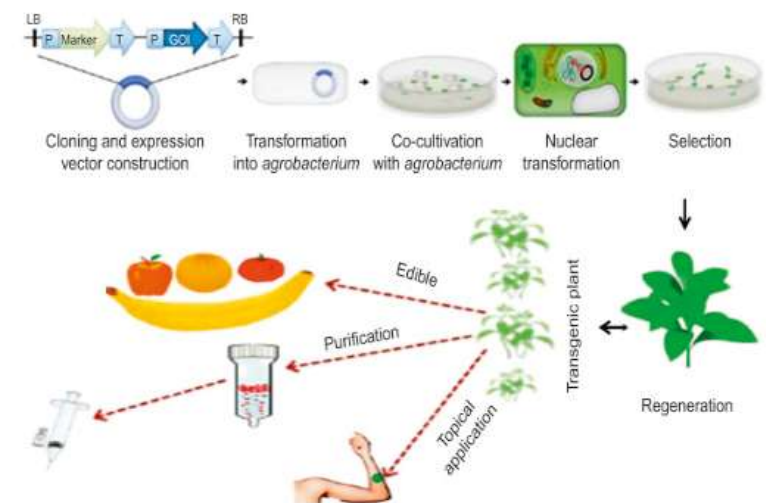
Curiosidade:

http://wolfson.huji.ac.il/purification/PDF/Expression_Systems/NEB_Klactiskit_manual.pdf

Plantas transgênicas como sistema de expressão de rProteínas

Principais vantagens

- Modificações pós-traducionais similares às de mamíferos
- Expressão da rProteína em níveis relativamente altos (até 10% TSP) a baixos custos
- Sem risco de contaminação por agentes patogênicos
- Facilidade de ampliação de escala (basta aumentar a área plantada)
- Facilidade de recuperação e purificação
- Elevada estabilidade no armazenamento, se rProteína for produzida em sementes
- “Edible vaccines”



Principais desvantagens

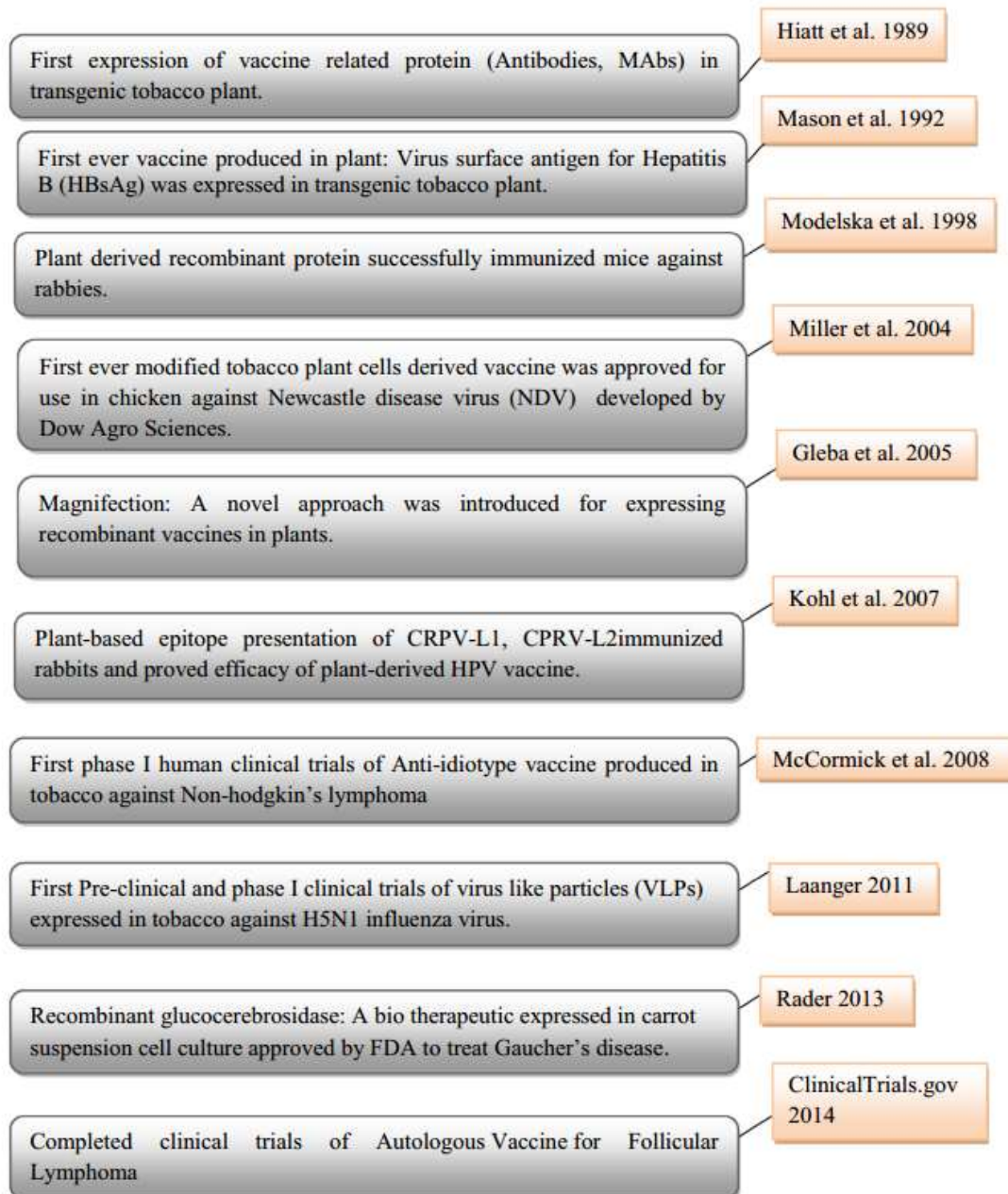
- Padrão de glicosilação diferente do encontrado em humanos, apesar de esforços de “humanização”
- Manipulação genética ainda não muito bem estabelecida
- Longo tempo de transformação, identificação e estabilização das linhagens de produção
- Longo tempo para produção, em algumas culturas
- Riscos ambientais e riscos de contaminação de cadeia alimentar (confinamento)
- Menor aceitação por consumidores e ambientalistas
- Problemas quanto a falta de práticas industriais e incertezas quanto a regulação de biofármacos produzidos em plantas transgênicas ainda persiste



Fig. 1 Key events in the development of plant-derived biopharmaceuticals



Plant cell bioreactors at the Protalix Biotherapeutics production plant.



Produção de rProteínas em células de mamífero

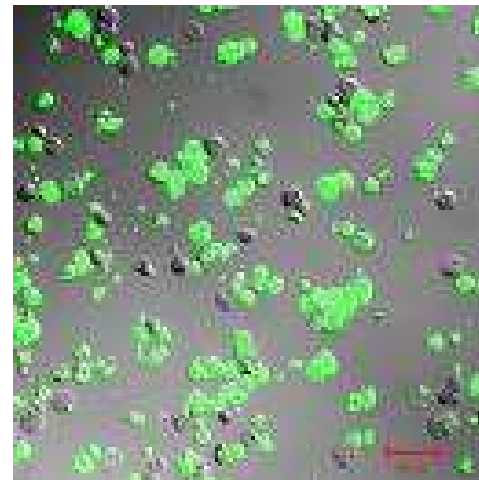
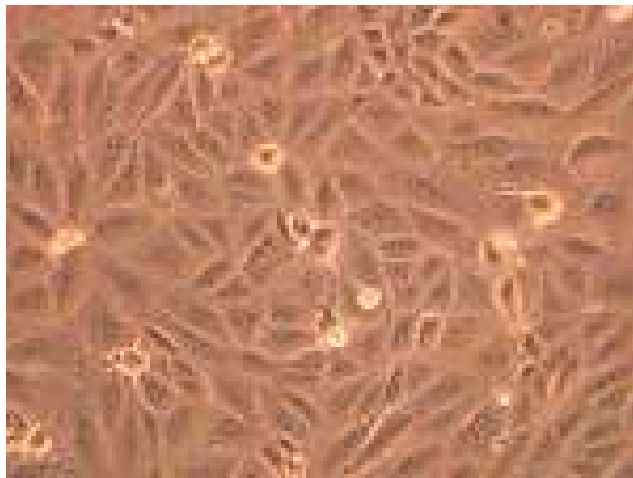
Principais vantagens

- Modificações pós-traducionais exigidas para proteínas terapêuticas (proteínas funcionais e de baixa imunogenicidade)
- Técnicas de manipulação genética bem estabelecidas
- Variedade de vetores comerciais disponíveis
- Níveis de expressão da rProteína relativamente sendo aprimorados
- Produção extracelular (capacidade de secretar para o meio)
- Equipamentos e meios de cultura desenvolvidos para produção em larga escala



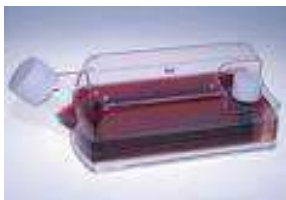
Principais vantagens (continuação)

- Amplamente usada para produção comercial, geralmente (mas não apenas) de proteínas complexas, cujas modificações pós-traducionais são importantes para a função/atividade
- Vários processos e produtos aprovados pelo FDA e EMEA (desde 1986)
- Hoje é o principal sistema de expressão da indústria farmacêutica (~65%)

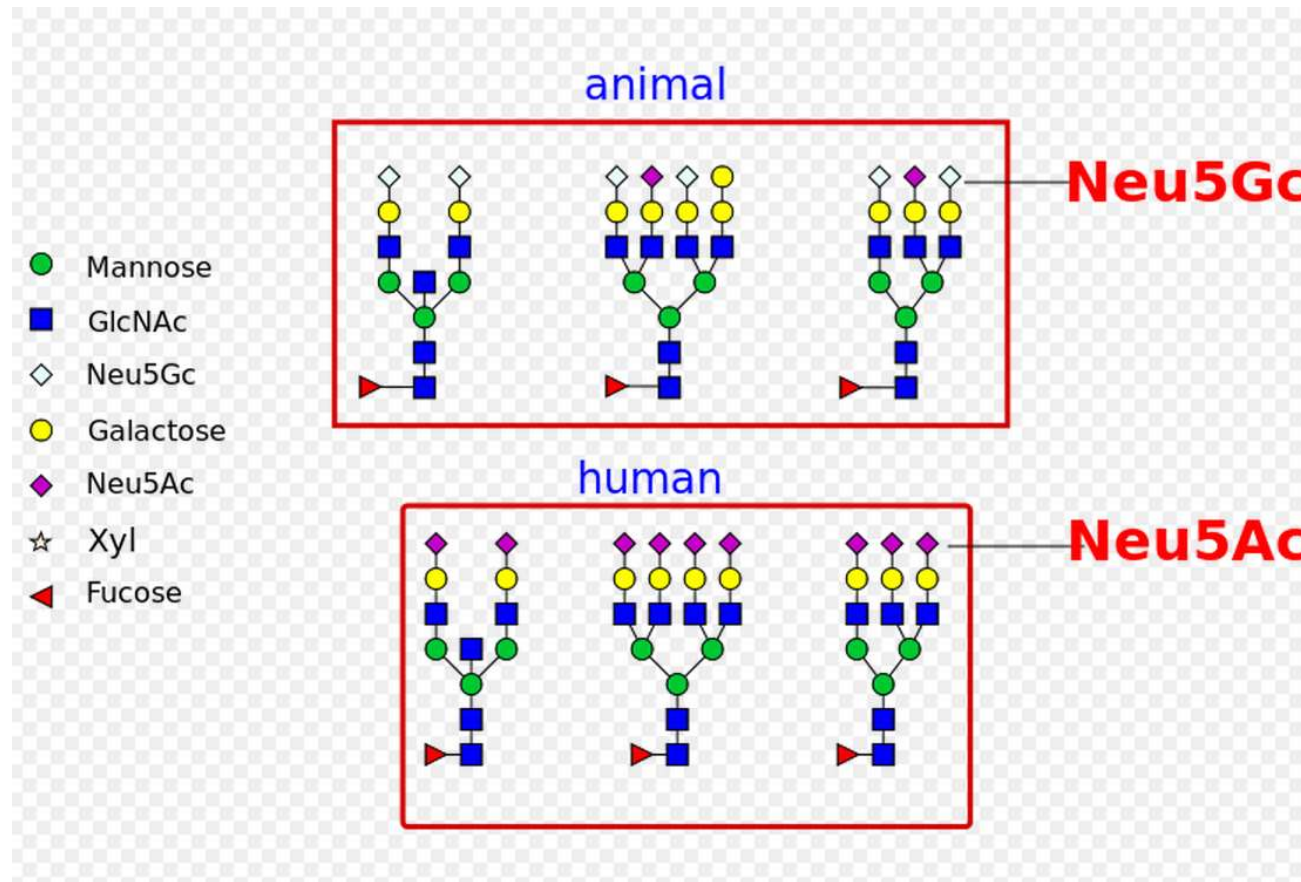


Principais desvantagens

- Maior tempo para seleção dos clones mais produtivos
- Manipulação delicada e maiores riscos de contaminação
- Menores taxas de crescimento
- Maiores custos associados à produção
- Riscos associados à presença/contaminação por agentes patogênicos



Diferenças na glicosilação de células animais e humanas



[N-glycolylneuraminic acid \(Neu5Gc\)](#) vs [N-acetylneuraminic acid \(Neu5Ac\)](#)

Expressão de rProteínas por animais transgênicos

Principais vantagens

- Modificações pós-traducionais necessárias à proteínas terapêuticas (proteínas funcionais e de baixa imunogenicidade)
- Capacidade de secretar a rProteína para fluidos selecionados, como o leite (produção em glândulas mamárias)
- Guidelines já disponibilizados por agências reguladoras para produção em leite de animais transgênicos (produto comercial disponível)



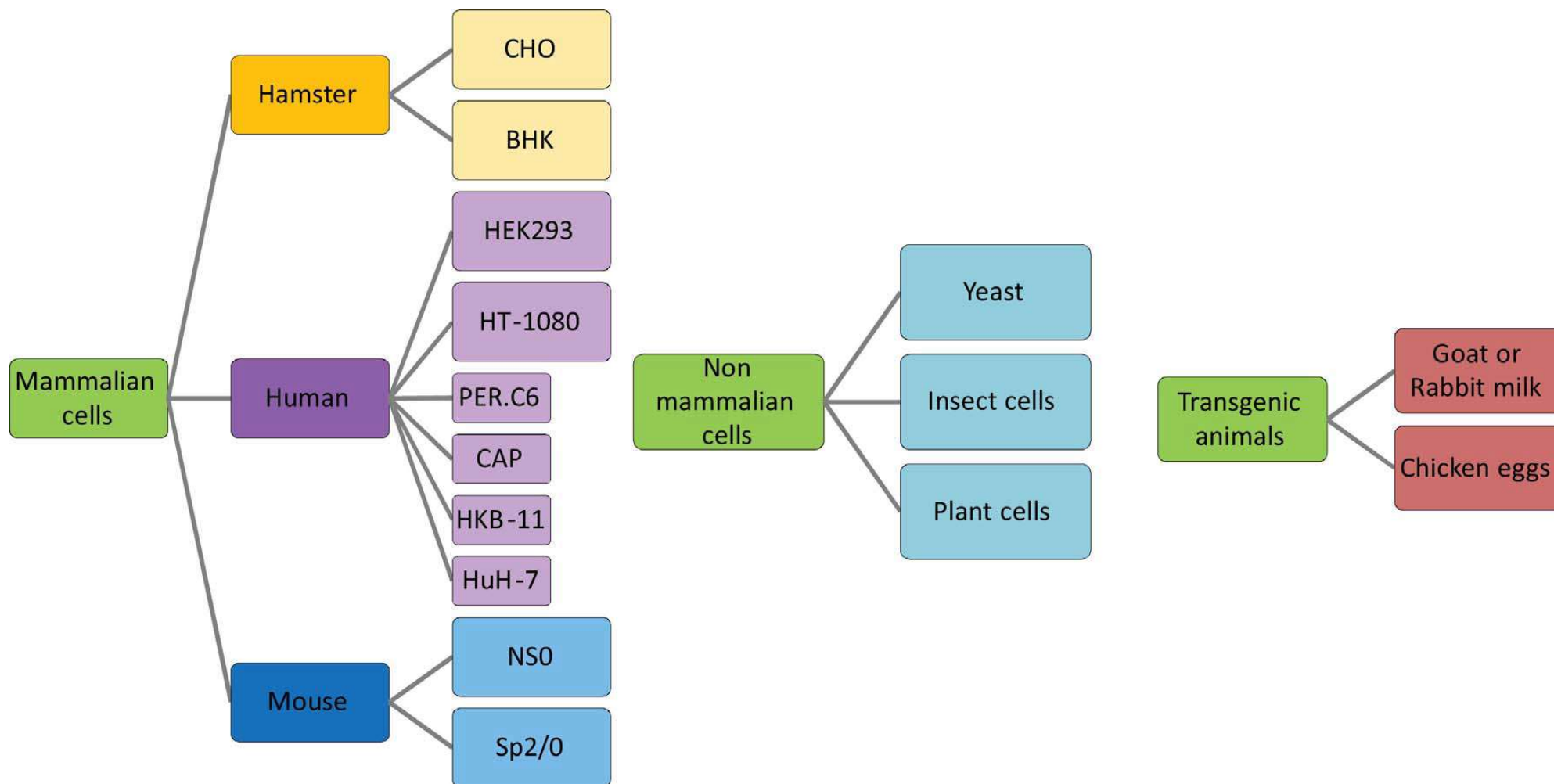
Table 5.7 Typical annual milk yields as well as time lapse between generation of the transgene embryo and first product harvest (first lactation) of indicated species

Species	Annual milk yield (l)	Time to first production batch (months)
Cow	6000–9000	33–36
Goat	700–800	18–20
Sheep	400–500	18–20
Pig	250–300	16–17
Rabbit	4–5	7

Principais desvantagens

- Manipulação genética delicada e técnicas pouco estabelecidas
- Longo tempo para seleção de animais produtivos
- Altos custos associados à segregação e manutenção dos animais
- Riscos associados à presença/contaminação por agentes patogênicos
- Baixa aceitação por parte dos consumidores

Sistemas de expressão usados na produção industrial de biofármacos glicosilados



Principais características, vantagens e desvantagens de diferentes sistemas de expressão

Expression system	Advantages	Limitations	Uses	Expressed proteins
<i>Escherichia coli</i>	Large number of genetic tools available High expression levels	No post-translational modifications	Research Commercial production	Non-complex proteins Antibody fragments
<i>Bacillus subtilis</i>	Secretion capacity Non-pathogenic organism	No post-translational modifications Protein degradation	Research Good potential for commercial production	Non-complex proteins Antibody fragments
<i>Saccharomyces cerevisiae</i>	Non-pathogenic organism Potential to scale-up	Glycosylation pattern Hiperglycosylation Low secretion capacity	Research Commercial production	Complex proteins
<i>Pichia pastoris</i>	High expression and secretion capacity	Glycosylation pattern Grown in explosion-proofed equipment	Research Good potential for commercial production	Complex proteins
Baculovirus – Insect cells	High expression levels Low risk of pathogens	Glycosylation pattern	Research Commercial production	Complex proteins
Plant cells	Simple media Low risk of pathogens	Glycosylation pattern	Potential for commercial production	Complex proteins
Animal cells	Expression of complex proteins Secretion	High production costs Low yields	Research Commercial production	Complex therapeutic proteins Monoclonal antibodies
Transgenic plants	Protein storage in seeds Low risk of pathogens Scale-up	Time consuming Glycosylation pattern	Commercial production	Complex proteins Edible vaccines
<i>Physcomitrella patens</i>	Secretion capacity Low risk of pathogens	Low expression levels	Good potential for commercial production	Complex proteins
Transgenic animals	Expression of complex proteins Secretion in body fluids (i.g. milk)	Time consuming Expensive research and development	Good potential for commercial production	Complex therapeutic proteins Monoclonal antibodies
Cell-free translation	Toxic proteins Suitable for automation	Post-translational modifications	Research/small scale	Non-complex proteins Proteins with unnatural aminoacids