

O processo da Expressão Gênica

O DNA contém a informação necessária para a produção de proteínas e está localizado quase que totalmente no núcleo das células, no entanto a síntese protéica ocorre no citoplasma. Por isso a informação genética, como é conhecida, precisa ser copiada do DNA e levada ao citoplasma através de moléculas que atuam como “mensageiras”, os RNA mensageiros (**RNAm**).

Os genes são as seqüências de nucleotídeos do DNA a partir das quais os RNA são produzidos. Dizemos que um determinado gene se expressa quando ele confere um certo fenótipo resultante da ação de uma dada proteína, a qual foi sintetizada a partir do RNAm produzido pelo próprio gene em questão.

A **expressão gênica** tem, portanto, duas etapas bem distintas: a **produção do RNAm** (chamada de **transcrição**) e a **síntese da proteína** (chamada de **tradução**).

Transcrição é o nome que se aplica ao processo de síntese de moléculas de RNA a partir do DNA. Isto engloba todos os tipos de RNA, e não só o RNAm.

Neste ponto uma importante diferença estrutural entre as moléculas de DNA e RNA precisa ficar clara: o RNA se apresenta constituído por apenas uma cadeia de nucleotídeos, além de ser constituído de uma base pirimídica chamada uracila (U) e um açúcar denominado ribose.

Visto que o DNA e o RNA possuem diferentes açúcares na sua constituição, as peças CFA utilizadas para a construção dos ribonucleotídeos foram padronizadas na cor **roxa** diferenciando daquelas utilizadas para a formação de moléculas de DNA.

A transcrição: DNA → RNA

A transcrição se inicia pelo rompimento temporário das pontes de hidrogênio numa determinada região do DNA correspondente a um gene, onde ocorre uma seqüência inicial **TAC** (local que a enzima RNAPolimerase irá se ligar). O processo transcorre com o pareamento de ribonucleotídeos à fita molde através da formação de novas pontes de hidrogênio. A relação de complementaridade entre as bases deve ser mantida, com a ressalva de que quando houver um nucleotídeo com a base A no DNA, encaixa-se um nucleotídeo com a base U no RNA. Quando encontrar uma seqüência **ATT** (finalizadora) a RNAPolimerase se desliga do DNA e a síntese é interrompida. O **RNAm** é liberado e migra para o citoplasma.

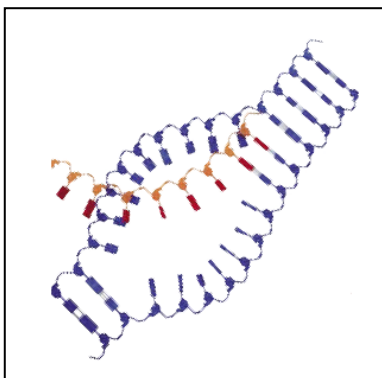


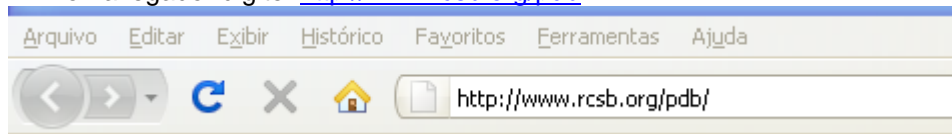
Figura 1 - Os novos nucleotídeos são sempre encaixados no sentido antiparalelo à fita molde, na região do carbono 3' do açúcar do último nucleotídeo incorporado.

A fita do DNA a servir como molde para a montagem do RNA, assim como a quantidade de ribonucleotídeos incorporados devem ser previamente determinadas. Isto indica a localização do gene na molécula de DNA e o seu tamanho.

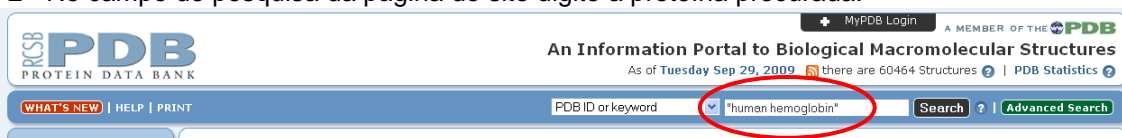
O professor deve orientar as seqüências gênicas a serem montadas para que ao final seja traduzida parte da seqüência de um gene que corresponda a uma proteína real. Exemplos de parte da seqüência gênica de algumas proteínas conhecidas (ou melhor, seguimentos destas proteínas), baixadas de um banco de dados (PDB, protein data bank) estão abaixo, assim como um roteiro para acesso ao PDB.

Sequências de proteínas no “Protein Data Bank” – PDB

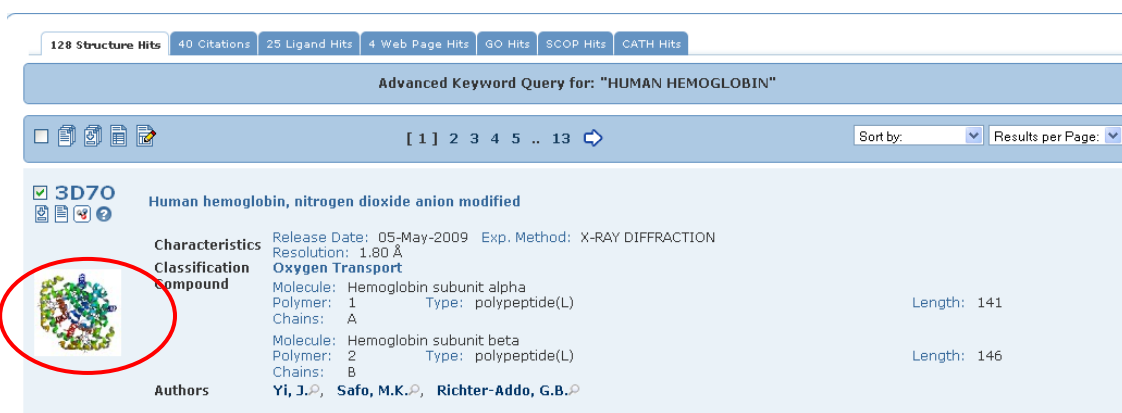
1 - No navegador digite: <http://www.rcsb.org/pdb>



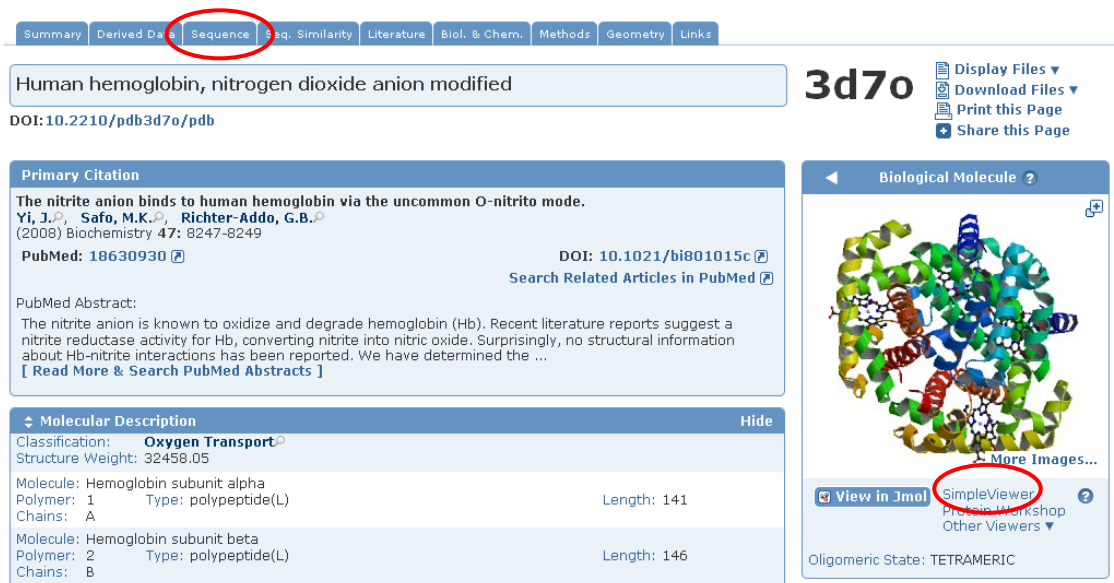
2 - No campo de pesquisa da página do site digite a proteína procurada:



3 - Selecione uma das opções que resultaram da pesquisa.



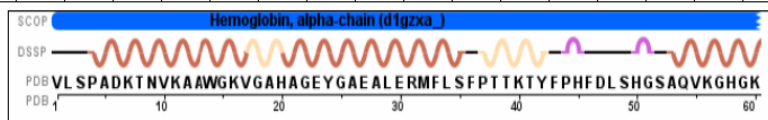
4 - Selecione “sequence” para ver a sequência de aminoácidos ou “SimpleViewer” para visualizar o modelo tridimensional.



Segmentos de sequência de algumas proteínas:

Hemoglobina - Cadeia Alfa

DNA	ATG	GTT	AAA	GCG	GCG	TGG	GGG	AAA	GTT	GGG	GCG	CAT	GCG	GGG	GAA	TAT	GGG	GCG	GAA	GCG	CTA	TAA
DNA -Template	TAC	CAA	TTT	CGC	CGC	ACC	CCC	TTT	CAA	CCC	CGC	GTA	CGC	CCC	CTT	ATA	CCC	CGC	CTT	CGC	GAT	ATT
mRNA	AUG	GUU	AAA	GCG	GCG	UGG	GGG	AAA	GUU	GGG	GCG	CAU	GCG	GGG	GAA	UAU	GGG	GCG	GAA	GCG	CUA	UAA
Proteína		V	K	A	A	W	G	K	V	G	A	H	A	G	E	Y	G	A	E	A	L	



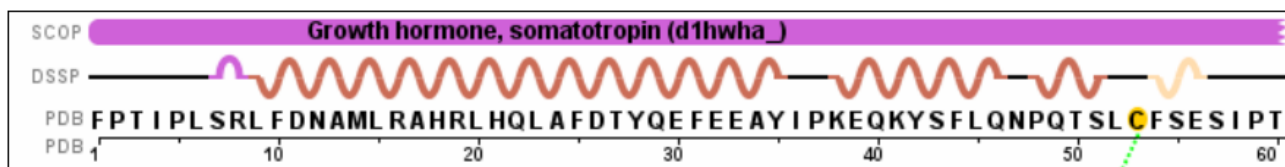
Insulina

DNA	ATG	TTC	GTA	AAC	CAA	CAC	CTC	TGT	GGT	TCC	CAC	CTC	GTA	GAA	GCT	CTC	TAT	CTC	GTA	TGT	GGT	TAA
DNA -Template	TAC	AAG	CAT	TTG	GTT	GTG	GAG	ACA	CCA	AGG	GTG	GAG	CAT	CTT	CGA	GAG	ATA	GAG	CAT	ACA	CCA	ATT
mRNA	AUG	UUC	GUA	AAC	CAA	CAC	CUC	UGU	GGU	UCC	CAC	CUC	GUA	GAA	GCU	CUC	UAU	CUC	GUA	UGU	GGU	UAA
Proteína		F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	



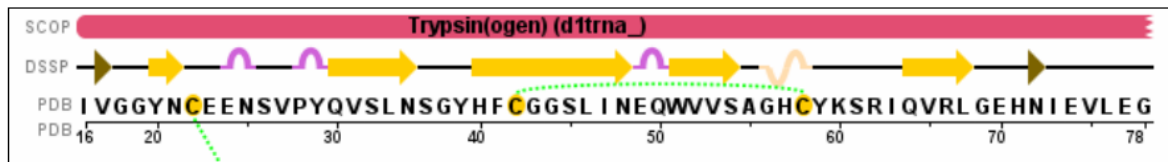
Hormônio de crescimento

DNA	ATG	TTC	CCG	ACA	ATT	CCG	CTT	TCT	AGA	CTT	TTC	GAT	AAC	GCC	ATG	CTT	AGA	GCC	CAC	AGA	CTT	TAA
DNA -Template	TAC	AAG	GGC	TGT	TAA	GGC	GAA	AGA	TCT	GAA	AAG	CTA	TTG	CGG	TAC	GAA	TCT	CGG	GTG	TCT	GAA	ATT
mRNA	AUG	UUC	CCG	ACA	AUU	CCG	CUU	UCU	AGA	CUU	UUC	GAU	AAC	GCC	AUG	CUU	AGA	GCC	CAC	AGA	CUU	UAA
Proteína		F	P	T	I	P	L	S	R	L	F	D	N	A	M	L	R	A	H	R	L	



Tripsina

DNA	ATG	ATC	GTC	GGG	GGA	TAC	AAC	TGC	GAG	GAA	AAC	AGT	GTG	CCC	TAC	CAA	GTG	TCA	CTA	AAC	AGT	TAA
DNA -Template	TAC	TAG	CAG	CCC	CCT	ATG	TTG	ACA	CTC	CTT	TTG	TCA	CAC	GGG	ATG	GTT	CAC	AGT	GAT	TTG	TCA	ATT
mRNA	AUG	AUC	GUC	GGG	GGA	UAC	AAC	UGC	GAG	GAA	AAC	AGU	GUG	CCC	UAC	CAA	GUG	UCA	CUA	AAC	AGU	UAA
Proteína		I	V	G	G	Y	N	C	E	E	N	S	V	P	Y	Q	V	S	L	N	S	



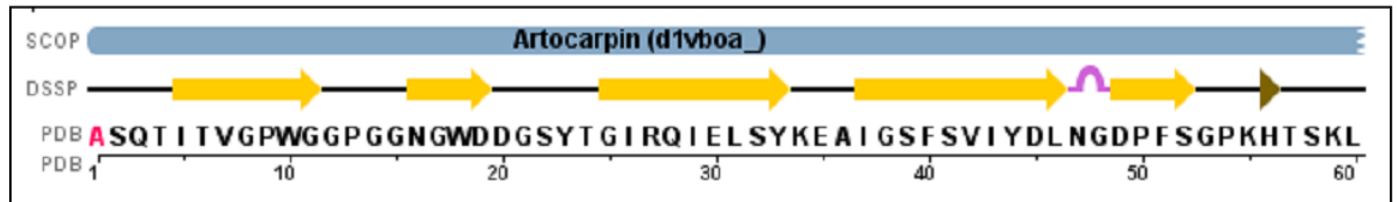
Lactase

DNA:	ATG	GGC	TCC	CAC	ATG	CTA	GAG	GAC	CCC	GTA	GTT	CTA	CAA	CGG	CGA	GAC	TGG	GAA	AAC	CCT	GGA	TAA
DNA -Template:	TAC	CCG	AGG	GTG	TAC	GAT	CTC	CTG	GGG	CAT	CAA	GAT	GTT	GCC	GCT	CTG	ACC	CTT	TTG	GGA	CCT	ATT
mRNA:	AUG	GGC	UCC	CAC	AUG	CUA	GAG	GAC	CCC	GUA	GUU	CUA	CAA	CGG	CGA	GAC	UGG	GAA	AAC	CCU	GGA	UAA
Proteína:		G	S	H	M	L	E	D	P	V	V	L	Q	R	R	D	W	E	N	P	G	



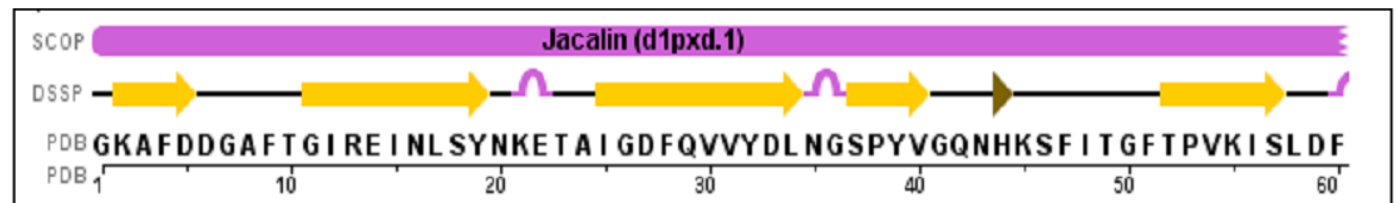
Artocarpina

DNA	ATG	ACA	GGG	ATT	AGA	CAA	ATT	GAA	CTT	TCT	TAC	AAA	GAA	GCC	ATT	GGG	TCT	TTC	TCT	GTG	ATT	TAC	GAT	CTT	AAC	GGG	GAT	CCG	TTC	TCT	GGG
DNA - Template	TAC	TGT	CCC	TAA	TCT	GTT	TAA	CTT	GAA	AGA	ATG	TTT	CTT	CGG	TAA	CCC	AGA	AAG	AGA	CAC	TAA	ATG	CTA	GAA	TTG	CCC	CTA	GGC	AAG	AGA	CCC
mRNA	AUG	ACA	GGG	AUU	AGA	CAA	AUU	GAA	CUU	UCU	UAC	AAA	GAA	GCC	AUU	GGG	UCU	UUC	UCU	GUG	AUU	UAC	GAU	CUU	AAC	GGG	GAU	CCG	UUC	UCU	GGG
Proteína		T	G	I	R	Q	I	E	L	S	Y	K	E	A	I	G	S	F	S	V	I	Y	D	L	N	G	D	P	F	S	G



Jacalina

DNA	ATG	GCC	ATT	GGG	GAT	TTC	CAA	GTG	GTG	TAC	GAT	CTT	AAC	GGG	TCT	CCG	TAC	GTG	GGG	CAA	AAC	CAC	AAA	TCT	TTC	ATT	ACA	GGG	TTC	ACA	CCG
DNA - Template	TAC	CGG	TAA	CCC	CTA	AAG	GTT	CAC	CAC	ATG	CTA	GAA	TTG	CCC	AGA	GGC	ATG	CAC	CCC	GTT	TTG	GTG	TTT	AGA	AAG	TAA	TGT	CCC	AAG	TGT	GCC
mRNA	AUG	GCC	AUU	GGG	GAU	UUC	CAA	GUG	GUG	UAC	GAU	CUU	AAC	GGG	UCU	CCG	UAC	GUG	GGG	CAA	AAC	CAC	AAA	UCU	UUC	AUU	ACA	GGG	UUC	ACA	CCG
Proteína		A	I	G	D	F	Q	V	V	Y	D	L	N	G	S	P	Y	V	G	Q	N	H	K	S	F	I	T	G	F	T	P



A tradução: RNAm → Proteína

As proteínas são longas cadeias formadas pela união de subunidades chamadas aminoácidos e representam uma classe de macromoléculas com características, propriedades e funções próprias.

Após o final da transcrição, a fita simples de RNAm sintetizada migra para o citoplasma, associando-se aos ribossomos, corpúsculos formados por proteínas e um outro tipo de RNA, o RNA ribossômico (RNAr).

A seqüência de bases do RNAm contém uma informação em código, trazida do núcleo para orientar a síntese de uma certa proteína. Os ribossomos, livres no citoplasma ou aderidos à superfície da membrana do retículo endoplasmático rugoso, têm a capacidade de decifrar esse código, conhecido como **código genético**, e fabricar a proteína especificada. Neste processo os ribossomos se valem da informação contida no RNAm, pois esta indica quais aminoácidos, quantos e em qual seqüência devem ser ligados. Cada três nucleotídeos do RNAm corresponde a o que se chama de **códon** e determina a incorporação de um determinado aminoácido. Entretanto, os ribossomos, por si só, não conseguem obter os aminoácidos livres no citoplasma, por isso são auxiliados pelas moléculas de **RNA transportador (RNAt)**.

Existem pelo menos vinte tipos diferentes de RNAt, um para cada aminoácido. Nas moléculas de cada RNAt há uma seqüência de três bases, denominada de **anticódon**, e um local onde se liga o aminoácido específico. O RNAt carrega este aminoácido até o ribossomo e caso seu anticódon se pareie com as bases do **códon** da molécula de RNAm (conforme a regra da complementaridade), o ribossomo incorpora o referido aminoácido à molécula de proteína que está sendo formada e passa a ler o próximo códon, esperando por outro RNAt. Desta forma, o ribossomo vai processando a tradução do código do RNAm à medida que a seqüência dos códons se complementa com os anticódons apresentados pelos RNAt, até que aparece a seqüência **UAA** (códon finalizador).

Tanto o Polígono de Nucleotídeos quanto o disco de Aminoácidos podem ser consultados, pois ambos contêm no verso a tabela do código genético. A partir do centro do disco e em direção à extremidade estão todas as 64 combinações possíveis para os códons do RNAm, com a abreviatura do aminoácido codificado por cada trinca na borda (livros de biologia, bioquímica e biologia molecular também trazem a tabela do código genético).

Os estudantes farão o papel dos ribossomos e RNAt e a seqüência de aminoácidos da proteína a ser sintetizada poderá ser montada utilizando o kit: **Construindo Moléculas de Aminoácidos e Proteínas**.

KIT: Construindo Moléculas de Aminoácidos e Proteínas (Leia o Manual do Professor sobre este kit.)

O professor deve relacionar estas atividades com o processo de síntese de proteínas (também chamado de expressão gênica). Além disso, as proteínas, após serem sintetizadas devem ser corretamente enoveladas, ou seja, adquirem estrutura espacial para exercerem suas funções fisiológicas. Atualmente muitas doenças degenerativas (doença da vaca louca, doença de Alzheimer) estão relacionadas com o enovelamento errado e/ou agregação de proteínas que foram corretamente expressas, mas “algo” as fez enovelar-se errado.

As peças que compõem o kit representam: o carbono alfa ($C\alpha$), com os ângulos de um tetraedro definido pelos pinos de conexão permitindo um encaixe rígido o bastante para manter a estrutura montada e estável, ao mesmo tempo em que permite a livre rotação com um leve esforço representando assim a liberdade rotacional dos grupos funcionais a ele ligados; grupo amino (NH_2), representado pela peça N-H, com um orifício que permite a conexão com o $C\alpha$ e outro que permitirá a formação da ligação peptídica; grupo carboxil ($COOH$), representado pela peça C=O, com o orifício que permite a conexão com o $C\alpha$ e uma haste que representará a ligação peptídica ao ser conectado ao grupo amino de um aminoácido subsequente; esfera branca representando o Hidrogênio; diferentes peças que representam as cadeias laterais de cada aminoácido, em suas formas geométricas; ligação de H e ligação dissulfeto (S-S), que ocorre entre os resíduos de cisteína. A ligação H e ligação dissulfeto permitem a construção de estruturas secundárias de proteínas tais como, folhas-beta e alfa-hélices.

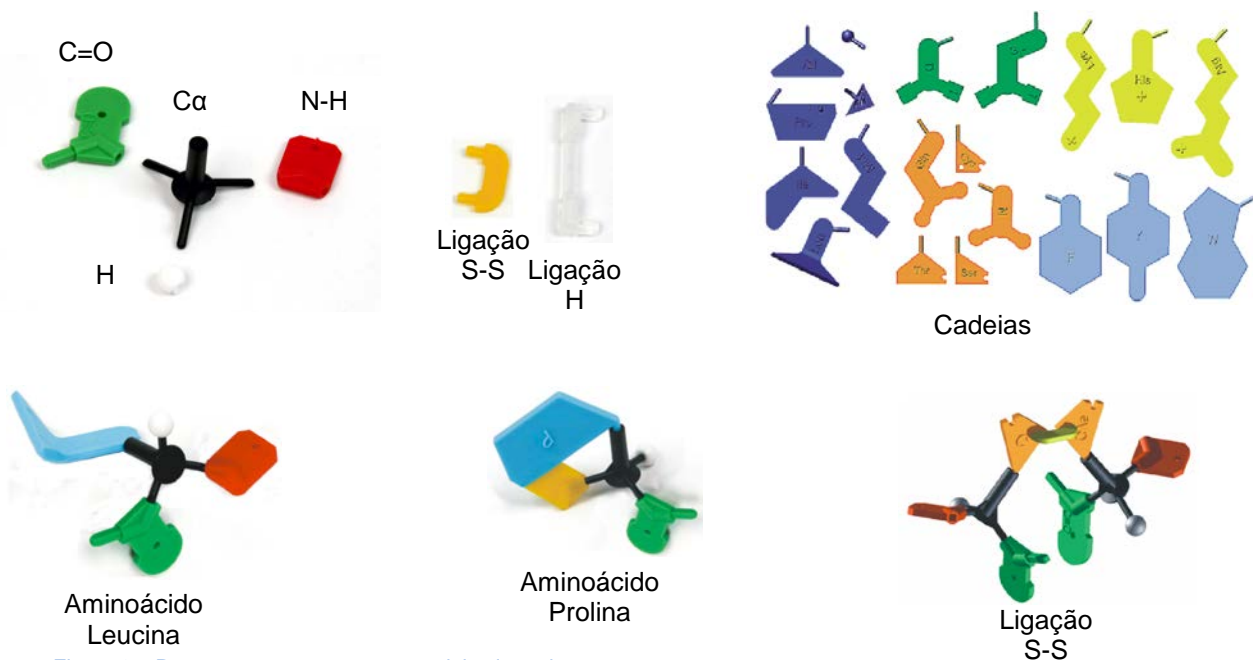


Figura 2 – Peças que representam as cadeias laterais.

Montando aminoácidos

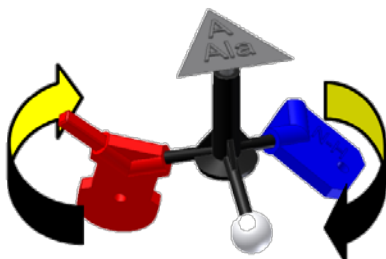


Figura 3 - Representação de um L-aminoácido mostrando a posição dos diferentes grupos.

Iniciar a estrutura primária montando as ligações peptídicas, que ocorrem através da conexão da haste do grupo carboxil com o “orifício quadrado” do grupo amino, conforme demonstrado na figura 4 (a).

Posicione a cadeia lateral para cima, no caso a peça que representa o aminoácido Alanina (ALA ou A), conecte à direita a peça que representa o aminogruppo (N-H), siga no sentido horário e conecte a peça que representa o H (esfera branca), à esquerda conecte a peça da carboxila (C=O), conforme a figura 3.

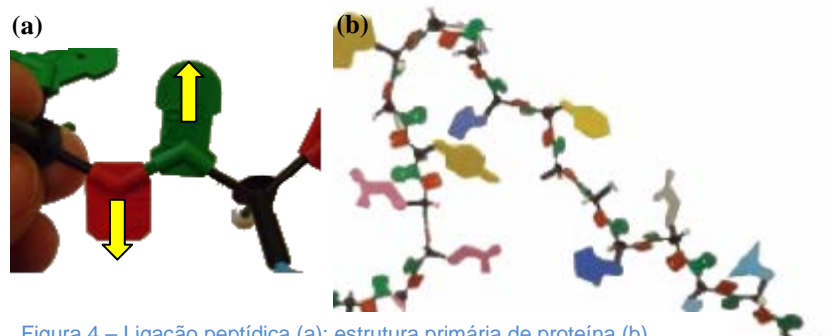


Figura 4 – Ligação peptídica (a); estrutura primária de proteína (b).

Montando a estrutura espacial das proteínas

DICA: Inicie a montagem da estrutura espacial das proteínas com as fitas e folhas beta. Elas são mais fáceis de montar por quem ainda não tem familiaridade com o modelo.

I) Fitas e folhas beta: para montar fitas e folhas- β , após construir uma estrutura primária, deve-se seguir as instruções e os esquemas mostrados abaixo:

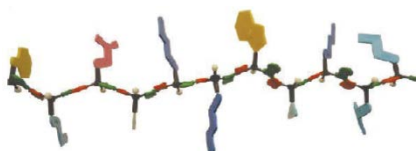


Figura 5 - Fita- β .

A sequência de resíduos é organizada alternando a direção das cadeias laterais de forma que os eixos de ligação com o $C\alpha$ estejam 180° entre os resíduos vizinhos (uma cadeia lateral para cima e outra para baixo, alternadamente). Isso faz com que a sequência polipeptídica fique quase linear e a estrutura é conhecida como **fita- β** .

Em seguida a fita- β é dobrada (volta- β) pareando e fazendo ligações de H entre os grupos carboxil e amino, figura 6, formando a **folha- β antiparalela**. Observar que as ligações de H estão a 90° em relação ao alinhamento dos resíduos na cadeia.

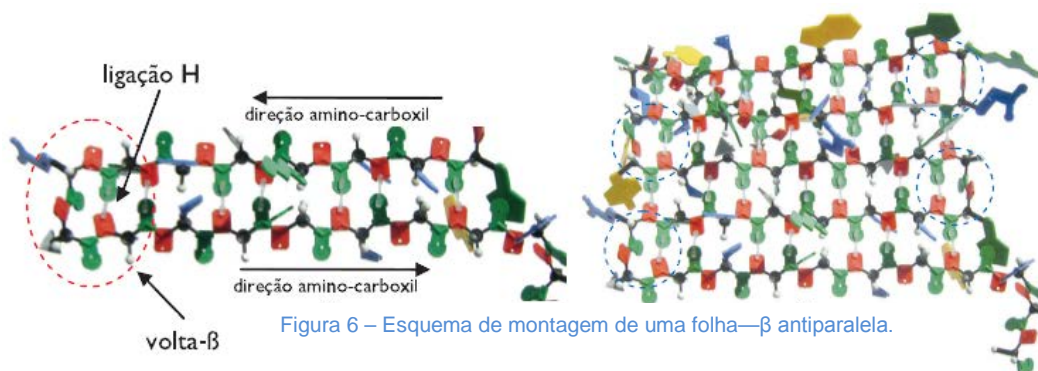
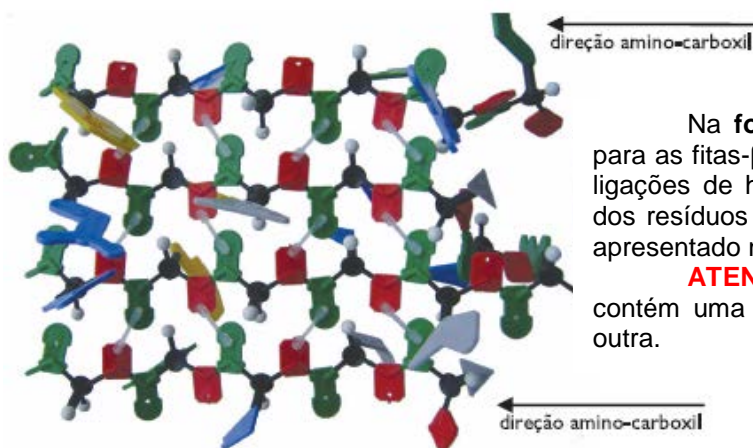


Figura 6 – Esquema de montagem de uma folha- β antiparalela.



Na **folha- β paralela** a direção amino-carboxil é a mesma para as fitas- β que estão lado a lado e não ocorrem as voltas- β . As ligações de hidrogênio são inclinadas em relação ao alinhamento dos resíduos na cadeia. A montagem deve ser realizada conforme apresentado na figura 7.

ATENÇÃO: As estruturas em folha- β paralelas, em geral, contém uma alça, uma fita- β ou uma hélice entre uma direção e outra.

Figura 7 – Folha- β paralela.

II) Hélice Alfa: a partir da estrutura primária, deve-se seguir as instruções abaixo

1. Imaginar um eixo central e iniciar a montagem no **sentido amino-terminal, carboxil terminal**;
2. Girar os resíduos de forma que as faces dos grupamentos amina e carboxila fiquem paralelas a esse eixo;
3. As cadeias laterais devem ser posicionadas para o lado externo da hélice;
4. Grupos amino são direcionados para o **terminal amino da hélice**;
5. Grupos carboxil são direcionados para o **terminal carboxil da hélice**;
6. Colocar uma ligação de hidrogênio entre a carboxila do **resíduo n (R1)** e a amina do **resíduo n+4 (R5)**.
7. Seguindo com a montagem, acrescenta-se uma ligação de hidrogênio a cada pareamento carboxila-amina. Em geral, a cada volta da hélice são encontradas duas ligações de hidrogênio.
8. O resultado será uma hélice com perfil triangular e ligações peptídicas em cada face do triângulo.

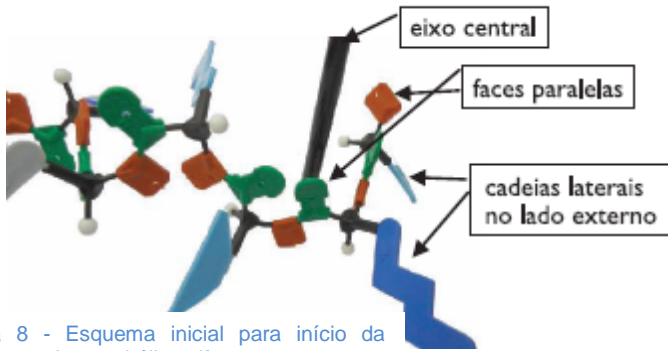


Figura 8 - Esquema inicial para início da montagem de uma hélice alfa.

DICA: Após montada a cadeia primária, começar a montar a hélice pelo terminal-amino, crescendo a cadeia para baixo em sentido horário.

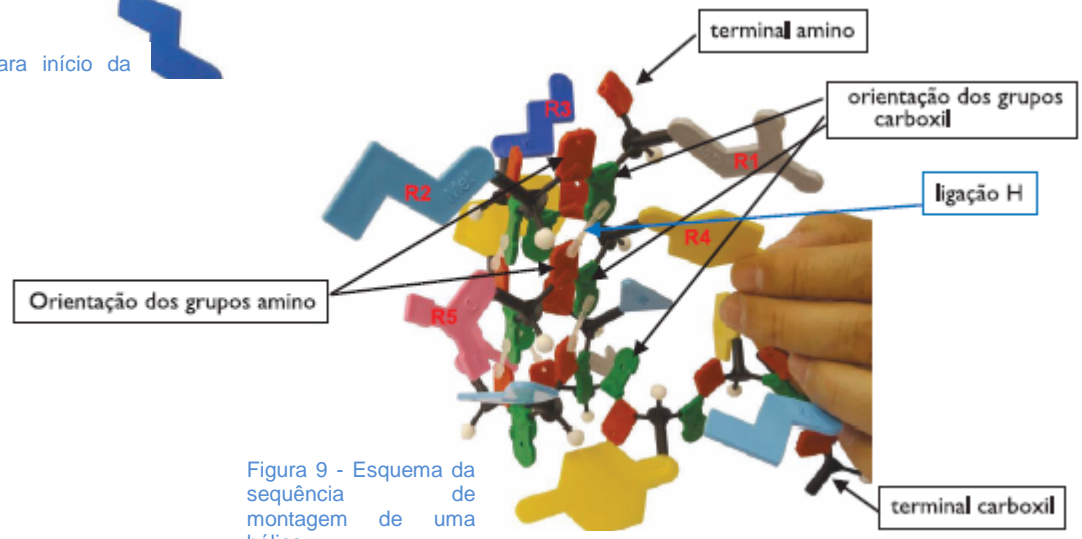


Figura 9 - Esquema da sequência de montagem de uma hélice.

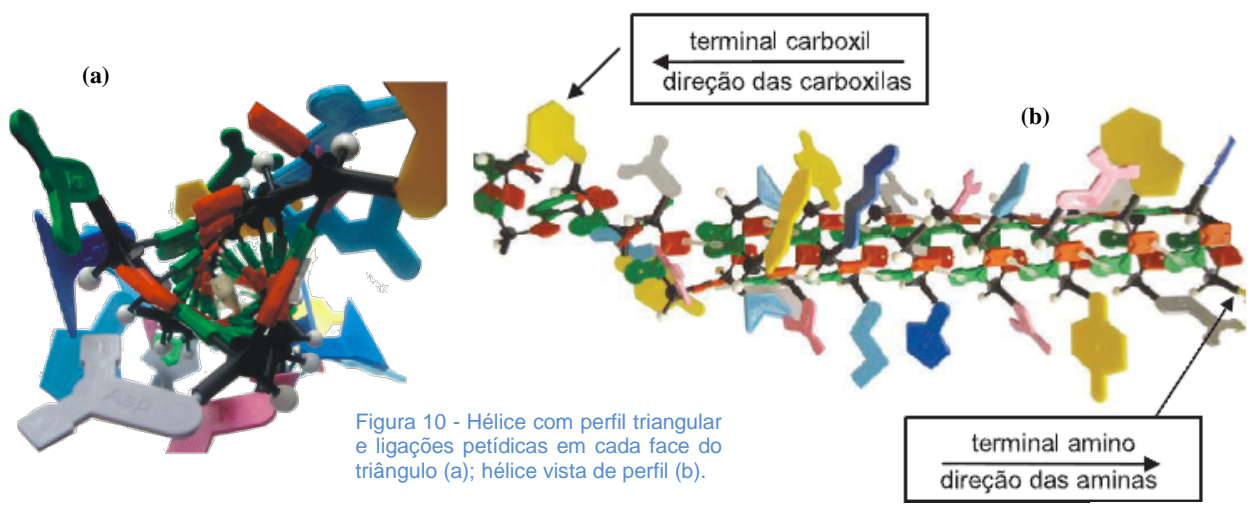


Figura 10 - Hélice com perfil triangular e ligações peptídicas em cada face do triângulo (a); hélice vista de perfil (b).

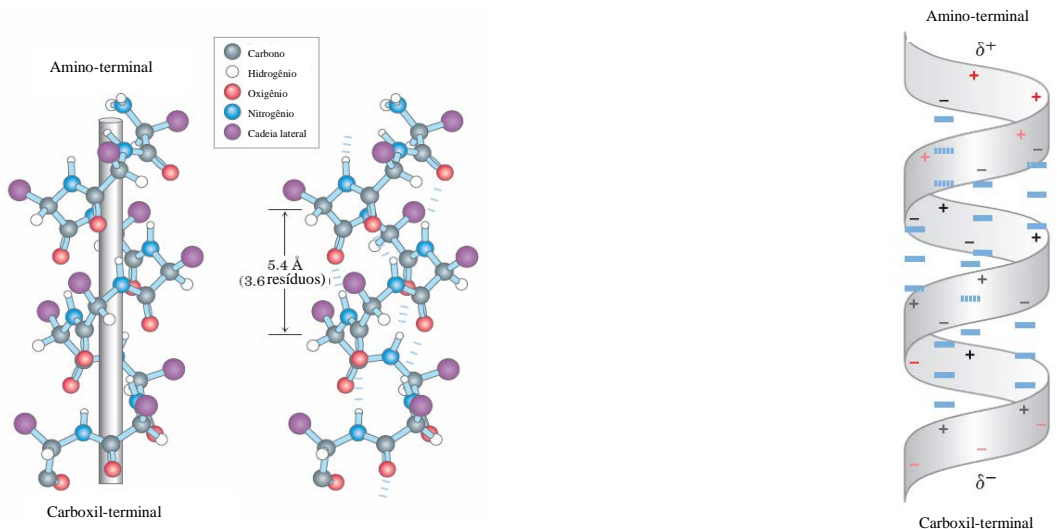


Figura 11 - Sentido da hélice, considerando a direção da carboxila terminal (Adaptado de Lehninger, Nelson & Cox, 3ª ed).

Autores (Os autores agradecem as sugestões, no sentido de tornar este roteiro mais didático, tenham certeza que serão sempre bem vindas.):

Profa. Dra. Leila Maria Beltramini, IFSC-USP: leila@ifsc.usp.br
 Dr. Aparecido Rodrigues da Silva, doutor em Física Biomolecular pelo IFSC-USP, LAM Educacional: aparecido@lameducacional.com.br
 Ms Gislaine Costa, educadora CIBFar/INBEQMeDI, IFSC-USP: gislainecosta@ifsc.usp.br ou cbme@ifsc.usp.br