

Estrutura Primária de Proteínas

Ignez Caracelli



Julio Zukerman Schpector



1

Bioinformática I



Ignez Caracelli

Funções biológicas das Proteínas

O grupo de biomoléculas mais diversificado quanto à estrutura e função.

- Biocatalizadores (enzimas: anidrase carbônica, proteases, etc.)
- Receptores de sinais químicos
- Transportadores
- Estruturais (citoesqueleto, colágeno, queratina)
- Defesa (sistema imunológico, restrição bacteriana, etc.)
- Mobilidade (motores moleculares)
- Armazenamento
- Transdução
- Aderência celular e organização tissular
- Enovelamento correto de outras proteínas
- Outras

2

Bioinformática I



Ignez Caracelli

Por que estudar proteínas?

Permite entender, propor mecanismos, fazer previsões, propor alterações em:

1. mecanismos de funcionamento
2. enovelamento (folding)
3. evolução molecular
4. modelagem molecular
5. engenharia de proteínas
6. desenho de fármacos

3

Bioinformática I




Ignez
Caracelli

vai onde?

o que vai fazer?

de onde veio?

4

Bioinformática I




Ignez
Caracelli

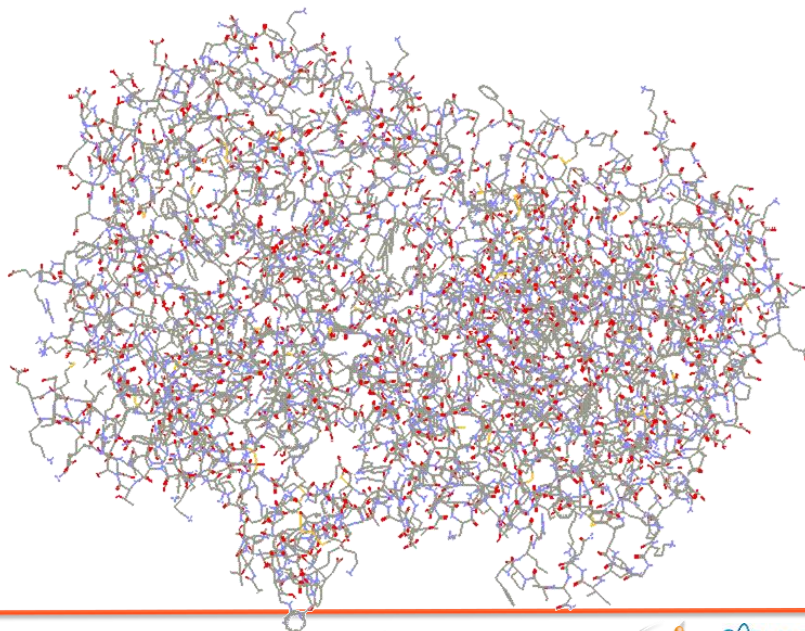
onde foi
parar?



5



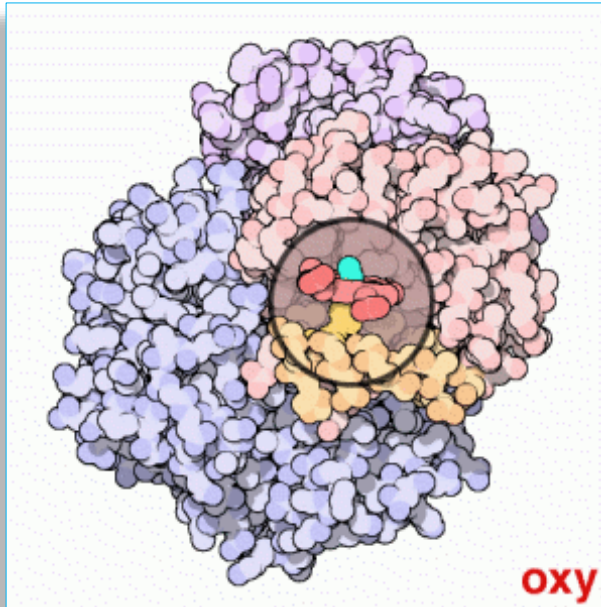
Bioinformática I

Ignez
Caracelli

6

Bioinformática I

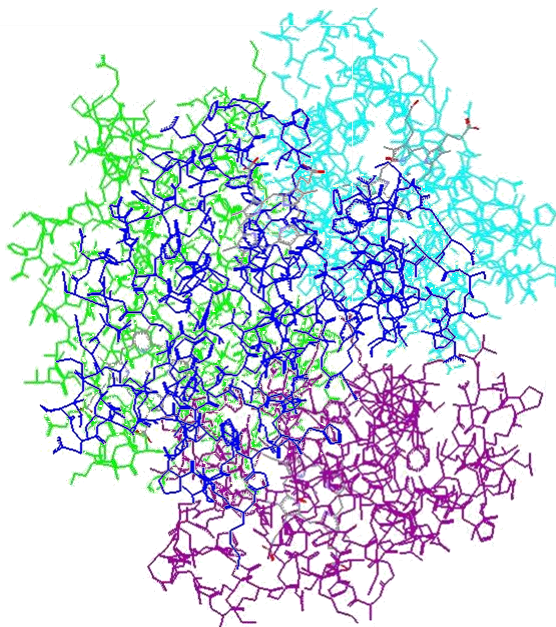
Ignez
Caracelli



Hemoglobina

7

Bioinformática I

Ignez
Caracelli

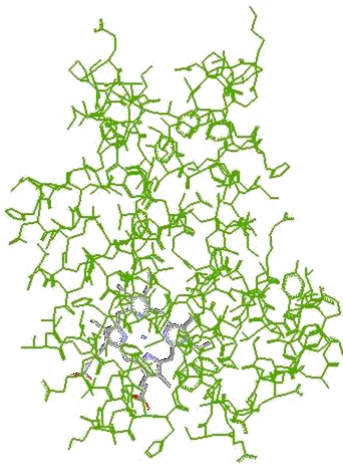
Hemoglobina

8

Bioinformática I

Ignez
Caracelli

Hemoglobina



parte protéica

parte não-protéica

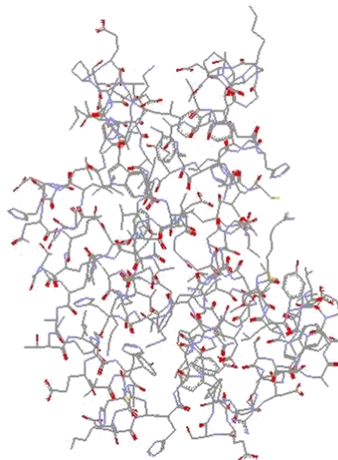
9

Bioinformática I



Ignez
Caracelli

Aminoácidos



parte protéica

átomos : C, N, H, O, S



aminoácidos

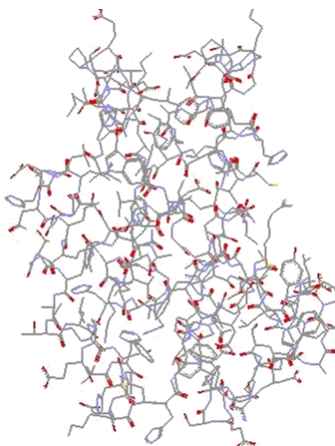
10

Bioinformática I



Ignez
Caracelli

Aminoácidos



11

Bioinformática I

Ignez
Caracelli

Pauling, aminoácidos e hélices



aminoácidos: *building blocks*
 ligação entre aminoácidos
 hélice alfa
 hélice DNA



12

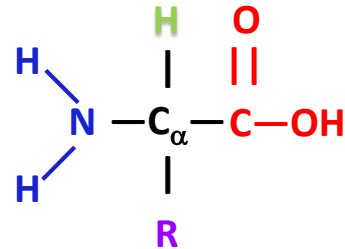
Bioinformática I

Ignez
Caracelli

O aminoácido

■ Estrutura:

- ◆ carbono central C_{α}
- ◆ um hidrogênio
- ◆ grupo amino
- ◆ grupo carboxílico (ácido)
- ◆ grupo R (cadeia lateral)
 - grupo variável
 - confere as propriedades químicas individuais do aminoácido



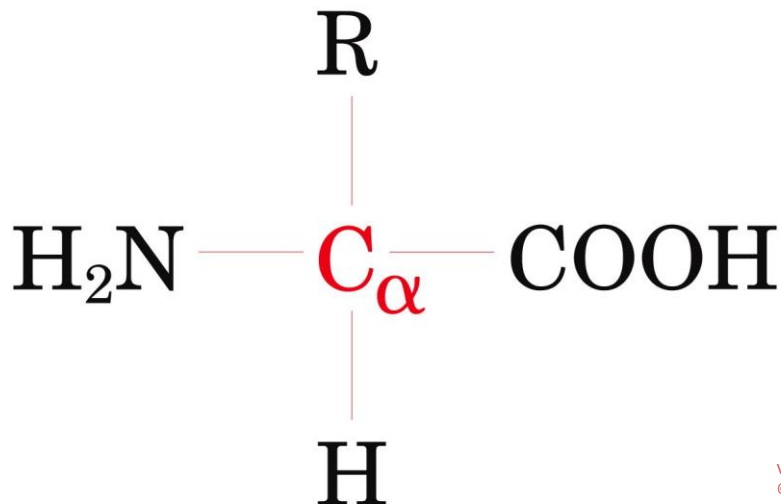
13

Bioinformática I



Ignez Caracelli

O aminoácido



Voet Biochemistry 3e
© 2004 John Wiley & Sons, Inc.

14

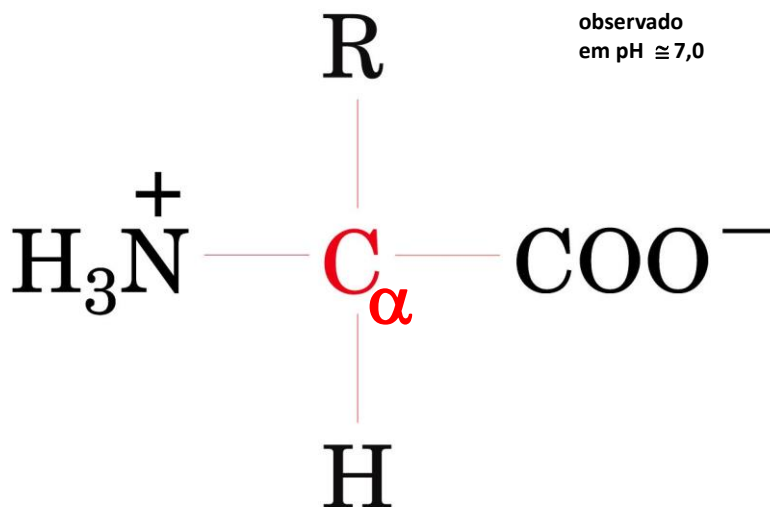
Bioinformática I



Ignez Caracelli

O aminoácido e sua forma ionizada

observado
em pH \cong 7,0



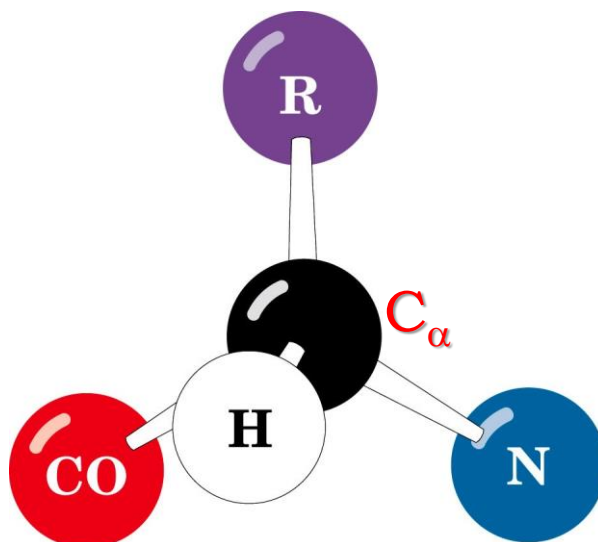
1st *Biochemistry* 3e
© 2004 John Wiley & Sons, Inc.

15

Bioinformática I

Ignez
Caracelli

O aminoácido e sua forma 3D



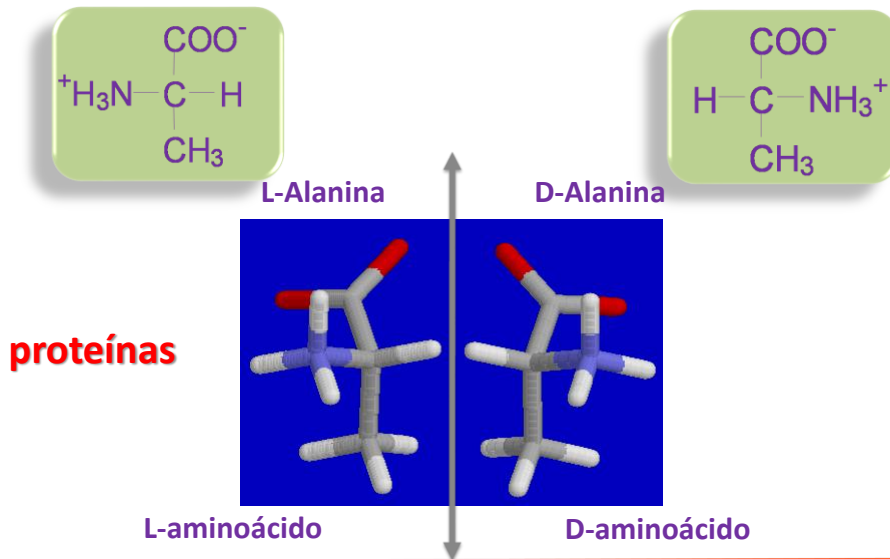
Voet *Biochemistry* 3e
© 2004 John Wiley & Sons, Inc.

16

Bioinformática I

Ignez
Caracelli

L-aminoácido e D-aminoácido



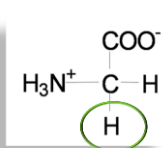
17

Bioinformática I

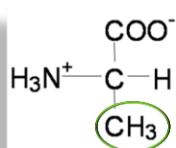


Ignez Caracelli

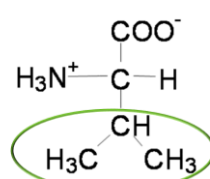
Cadeias laterais não-polares



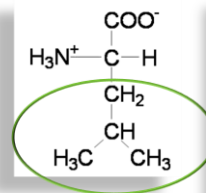
Glicina
Gly, G



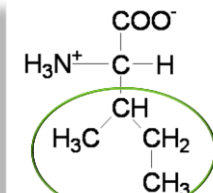
Alanina
Ala, A



Valina
Val, V



Leucina
Leu, L



Isoleucina
Ile, I

G, A, V, L, I, F, W, M, P

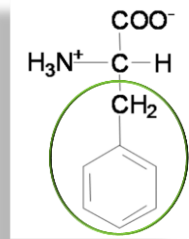
18

Bioinformática I

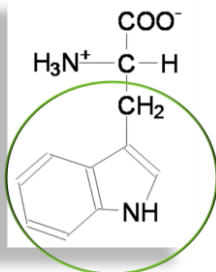


Ignez Caracelli

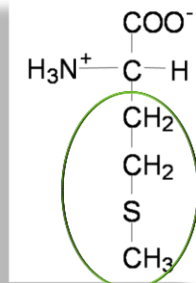
Cadeias laterais não-polares



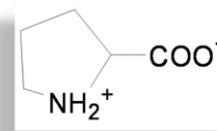
Fenilalanina
Phe, F



Triptofano
Trp, W



Metionina
Met, M



Prolina
Pro, P

G, A, V, L, I, F, W, M, P

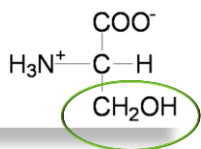
19

Bioinformática I

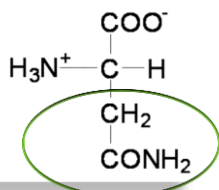


Ignez
Caracelli

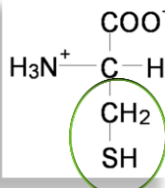
Cadeias laterais polares não-carregadas



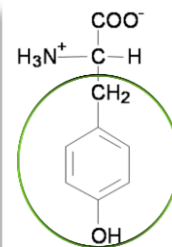
Serina
Ser, S



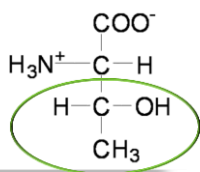
Asparagina
Asn, N



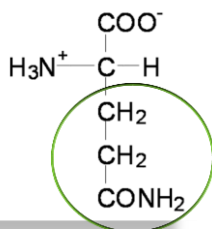
Cisteína
Cys, C



Tirosina
Tyr, Y



Treonina
Thr, T



Glutamina
Gln, Q

S, T, N, Q, Y, C

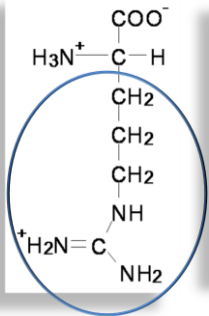
20

Bioinformática I

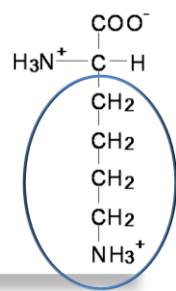


Ignez
Caracelli

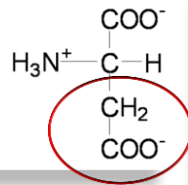
Cadeias laterais polares carregadas



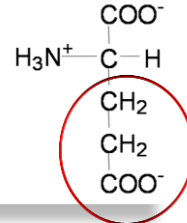
Arginina
Arg, R



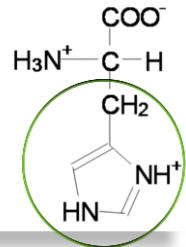
Lisina
Lys, K



Ácido Aspártico
Asp, D



Ácido Glutâmico
Glu, E



Histidina
His, H

R, K, D, E, H

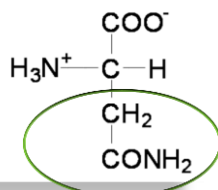
21

Bioinformática I



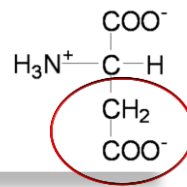
Ignez
Caracelli

Cadeias laterais
polares
não-carregadas



Asparagina
Asn, N

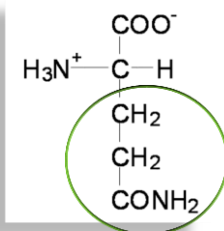
Asx



Ácido Aspártico
Asp, D

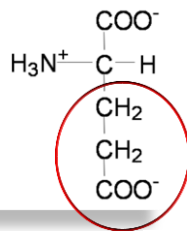


Cadeias laterais
polares
carregadas



Glutamina
Gln, Q

Glx



Ácido Glutâmico
Glu, E



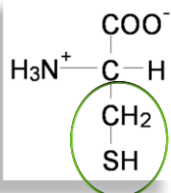
22

Bioinformática I

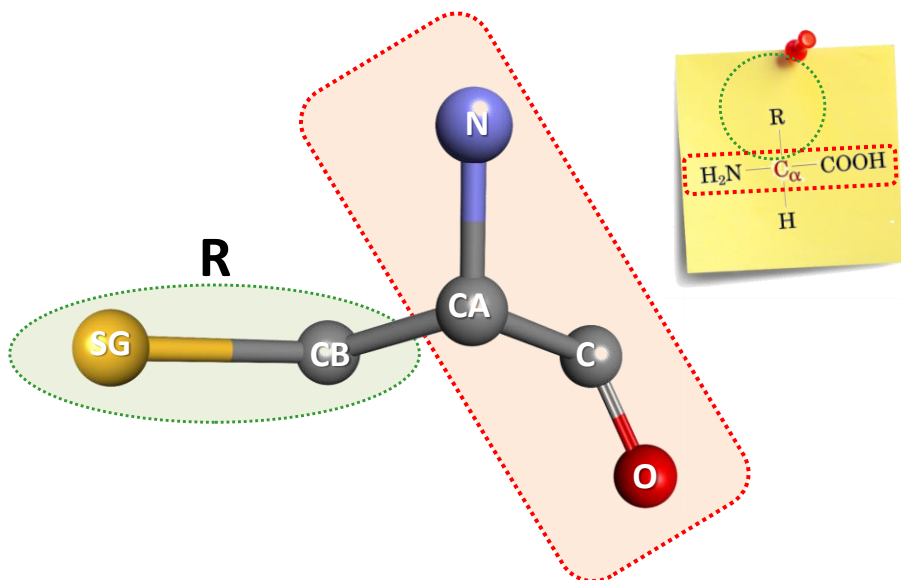


Ignez
Caracelli

Cadeias laterais com S



Cisteína
Cys, C



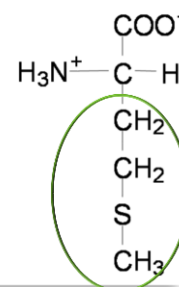
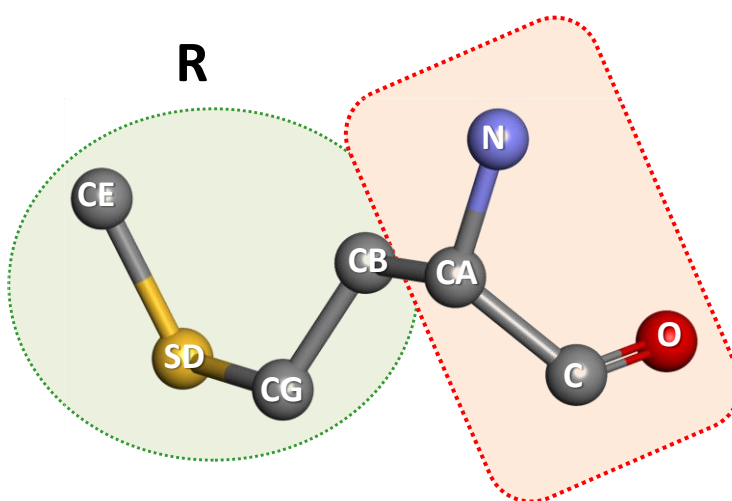
23

Bioinformática I



Ignez
Caracelli

Cadeias laterais com S



Metionina
Met, M

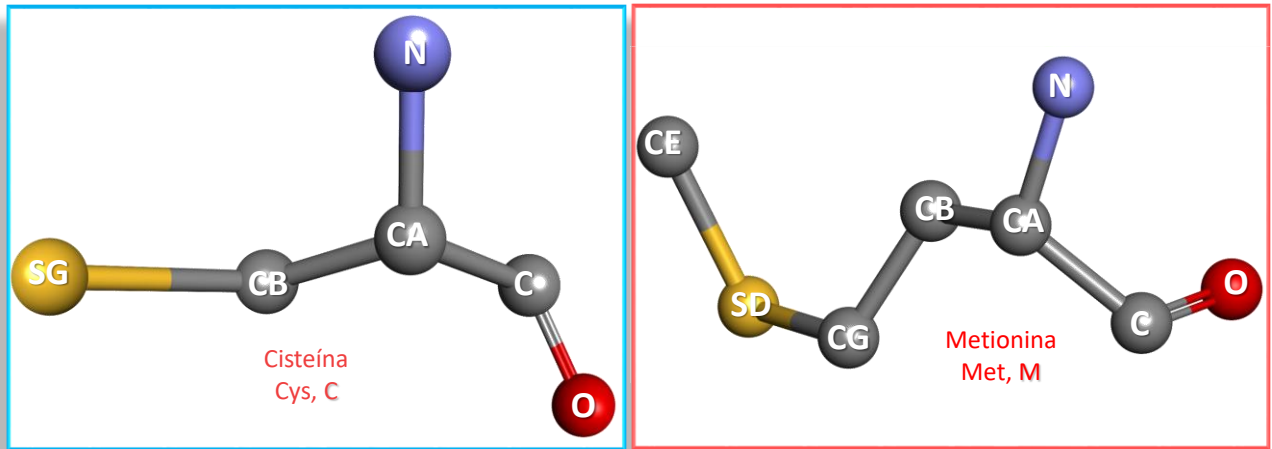
24

Bioinformática I



Ignez
Caracelli

Cadeias laterais com S



25

Bioinformática I

Ignez
Caracelli

Aminoácidos

Quantos são?

até 1986 eram

20 aminoácidos proteínogênicos (constituintes das proteínas),

mas...

26

Bioinformática I

Ignez
Caracelli

ano: 1985

Selênio em Moléculas Biológicas



1a. Estrutura Cristalográfica de uma enzima contendo um aminoácido com Se

Eur. J. Biochem. 133, 51 – 69 (1983)

© FEBS 1983

The Refined Structure of the Selenoenzyme Glutathione Peroxidase at 0.2-nm Resolution

Otto EPP, Rudolf LADENSTEIN, and Albrecht WENDEL

Max-Planck-Institut für Biochemie, Abteilung Strukturforschung II, Martinsried bei München;
and Institut für Physiologische Chemie der Universität Tübingen

(Received November 22, 1982/February 2, 1983) – EJB 6242

27

Bioinformática I



Ignez Caracelli

ano: 1985

Selênio em Moléculas Biológicas



1a. Estrutura Cristalográfica de uma enzima contendo um aminoácido com Se

1GP1

THE REFINED STRUCTURE OF THE SELENOENZYME **GLUTATHIONE PEROXIDASE** AT 0.2-NM RESOLUTION

- DOI: [10.2210/pdb1GP1/pdb](https://doi.org/10.2210/pdb1GP1/pdb)
- Classification: [OXIDOREDUCTASE\(H2O2\(A\)\)](#)
- Organism(s): [Bos taurus](#)
- Deposited: 1985-06-11 Released: 1985-11-08
- Deposition Author(s): [Epp, O.](#), [Ladenstein, R.](#)

28

<http://www.ebi.ac.uk/pdbsum/>

Bioinformática I



Ignez Caracelli

ano: 1985

Selênio em Moléculas Biológicas

PDB
sumPDB id: 1gp1 [Links](#)

Name: Oxidoreductase(h2o2(a))

Title: The refined structure of the selenoenzyme glutathione peroxidase at 0.2-nm resolution

Structure: Glutathione peroxidase. Chain: a, b. Engineered: yes

Source: Bos taurus. Bovine. Organism_taxid: 9913. Tissue: erythrocyte

Biol. unit: Tetramer (from PQS)

Resolution: 2.00Å R-factor: 0.171

Authors: O.Epp,R.Ladenstein

Key ref: O.Epp et al. (1983). The refined structure of the selenoenzyme glutathione peroxidase at 0.2-nm resolution. *Eur J Biochem*, 133, 51-69. PubMed id: [6852035](#)

Date: 11-Jun-85 Release date: 08-Nov-85

29

<http://www.ebi.ac.uk/pdbsum/>

Bioinformática I

PPG biotec

UFPE-IT

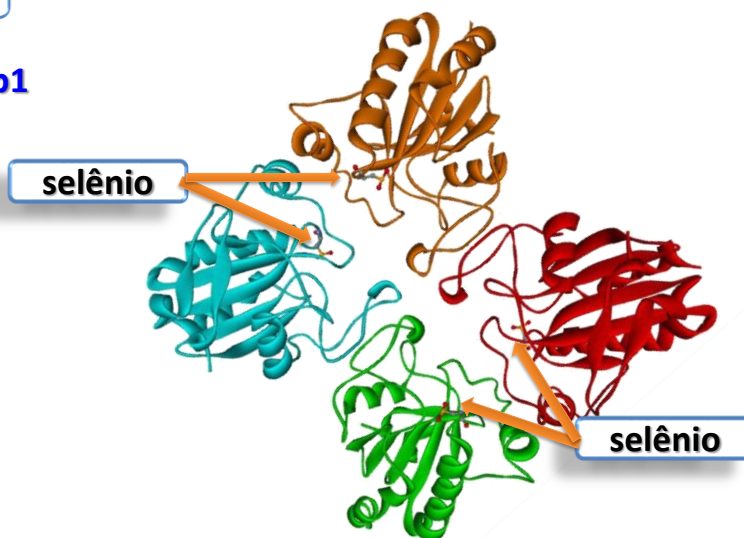
toMax

Ignez Caracelli

ano: 1985

Selênio em Moléculas Biológicas

pdb code 1gp1



30

Bioinformática I

PPG biotec

UFPE-IT

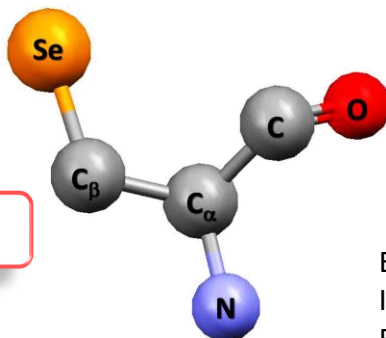
toMax

Ignez Caracelli

pdbe code 1gp1

Selênio: em um aminoácido?

selenocisteína



Eur. J. Biochem. 264, 607-609 (1999)
IUPAC-IUBMB Joint Commission on
Biochemical Nomenclature (JCBN)
and
Nomenclature Committee of IUBMB
(NC-IUBMB)

31

<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl>

Bioinformática I

UFPA

SCM

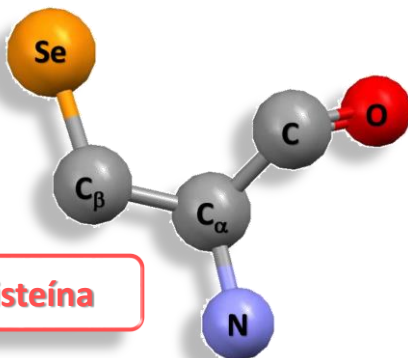
Ignez Caracelli

ano: 1985

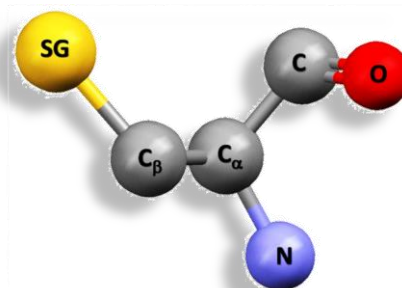
Selênio em Moléculas Biológicas

pdbe code 1gp1

selenocisteína



cisteína



32

Bioinformática I

PPG biotec

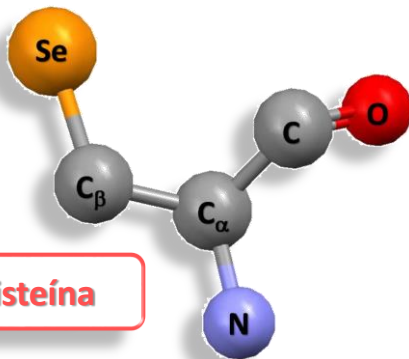
UFPA

SCM

Ignez Caracelli

ano: 1985

pdb code 1gp1



selenocisteína

Selênio em Moléculas Biológicas

Perguntas:

- ocorreu um erro?
- uma mutação?
- ou é um novo aminoácido?

33

Bioinformática I

Ignez
Caracelli

Códons para aminoácidos

Códon:

Sequência de três nucleotídeos de RNA mensageiro (mRNA) que codificam um determinado aminoácido

Exemplo: UUU → fenilalanina

ou

start

stop

que indicam o ponto de início ou fim de tradução da cadeia de mRNA.

Exemplo:



34

34

Bioinformática I

Ignez
Caracelli

Códons para aminoácidos

Alanine	Ala	A	GCA GCC GCG GCU
Cysteine	Cys	C	UGC UGU
Aspartic acid	Asp	D	GAC GAU
Glutamic acid	Glu	E	GAA GAG
Phenylalanine	Phe	F	UUC UUU
Glycine	Gly	G	GGA GGC GGG GGU
Histidine	His	H	CAC CAU
Isoleucine	Ile	I	AUA AUC AUU
Lysine	Lys	K	AAA AAG
Leucine	Leu	L	UUA UUG CUA CUC CUG CUU

Sequência de três nucleotídeos de RNA mensageiro (mRNA) que codificam um determinado aminoácido

35

Bioinformática I



Ignez Caracelli

Códons para aminoácidos

Methionine	Met	M	AUG
Asparagine	Asn	N	AAC AAU
Proline	Pro	P	CCA CCC CCG CCU
Glutamine	Gln	Q	CAA CAG
Arginine	Arg	R	AGA AGG CGA CGC CGG CGU
Serine	Ser	S	AGC AGU UCA UCC UCG UCU
Threonine	Thr	T	ACA ACC ACG ACU
Valine	Val	V	GUA GUC GUG GUU
Tryptophan	Trp	W	UGG
Tyrosine	Tyr	Y	UAC UAU

Sequência de três nucleotídeos de RNA mensageiro (mRNA) que codificam um determinado aminoácido

36

Bioinformática I



Ignez Caracelli

Start códon: AUG

Stop códon: UAA, UAG, UGA

37

Bioinformática I



Ignez Caracelli

Middle Base 5' Base	U	C	A	*G	Middle Base 3' Base
*U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	*Sec Stop	*A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	Hís	Arg	U
	Leu	Pro	Hís	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met ▲ Initiator	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Códons para aminoácidos

A tabela mostra os códons e os aminoácidos. A direção do mRNA é de 5' → 3'.



38

Longtin, R. J. Natl. Cancer Inst. 2004 96:504-505; doi:10.1093/jnci/96.7.504

Bioinformática I



Ignez Caracelli

Middle Base 5' Base	U	C	A	*G	Middle Base 3' Base
*U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	*A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	▲ Met (Initiator)	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Códons para aminoácidos

exemplos:

UUU → fenilalanina

UUC → fenilalanina

AAG → lisina

39

Longtin, R. J. Natl. Cancer Inst. 2004 96:504-505; doi:10.1093/jnci/96.7.504

Bioinformática I

PPG biotec

UFPR

loMax

Ignez Caracelli

Middle Base 5' Base	U	C	A	*G	Middle Base 3' Base
*U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	*A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	▲ Met (Initiator)	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Códons para aminoácidos

Stop códon:

UAA,

UAG,

UGA

40

Longtin, R. J. Natl. Cancer Inst. 2004 96:504-505; doi:10.1093/jnci/96.7.504

Bioinformática I

PPG biotec

UFPR

loMax

Ignez Caracelli

5' Base	Middle Base	U	C	A	*G	Middle Base	3' Base
*U		Phe	Ser	Tyr	Cys	U	
		Phe	Ser	Tyr	Cys	C	
		Leu	Ser	Stop	* Sec	*A	
		Leu	Ser	Stop	Trp	G	
C		Leu	Pro	His	Arg	U	
		Leu	Pro	His	Arg	C	
		Leu	Pro	Gln	Arg	A	
		Leu	Pro	Gln	Arg	G	
A		Ile	Thr	Asn	Ser	U	
		Ile	Thr	Asn	Ser	C	
		Ile	Thr	Lys	Arg	A	
		▲ Met <i>(initiator)</i>	Thr	Lys	Arg	G	
G		Val	Ala	Asp	Gly	U	
		Val	Ala	Asp	Gly	C	
		Val	Ala	Glu	Gly	A	
		Val	Ala	Glu	Gly	G	

Códons para aminoácidos

Stop códon:

UAA,
UAG,
UGA



Selenocisteina (Sec):

UGA

41

Longtin, R. J. Natl. Cancer Inst. 2004 96:504-505; doi:10.1093/jnci/96.7.504

Bioinformática I

PPG biotec

UFPA

toMax

Ignez Caracelli

5' Base	Middle Base	U	C	A	*G	Middle Base	3' Base
*U		Phe	Ser	Tyr	Cys	U	
		Phe	Ser	Tyr	Cys	C	
		Leu	Ser	Stop	* Sec	*A	
		Leu	Ser	Stop	Trp	G	
C		Leu	Pro	His	Arg	U	
		Leu	Pro	His	Arg	C	
		Leu	Pro	Gln	Arg	A	
		Leu	Pro	Gln	Arg	G	
A		Ile	Thr	Asn	Ser	U	
		Ile	Thr	Asn	Ser	C	
		Ile	Thr	Lys	Arg	A	
		▲ Met <i>(initiator)</i>	Thr	Lys	Arg	G	
G		Val	Ala	Asp	Gly	U	
		Val	Ala	Asp	Gly	C	
		Val	Ala	Glu	Gly	A	
		Val	Ala	Glu	Gly	G	

Códons para aminoácidos

Start códon: AUG

Já havia sido observado

Metionina (Met): AUG

42

Longtin, R. J. Natl. Cancer Inst. 2004 96:504-505; doi:10.1093/jnci/96.7.504

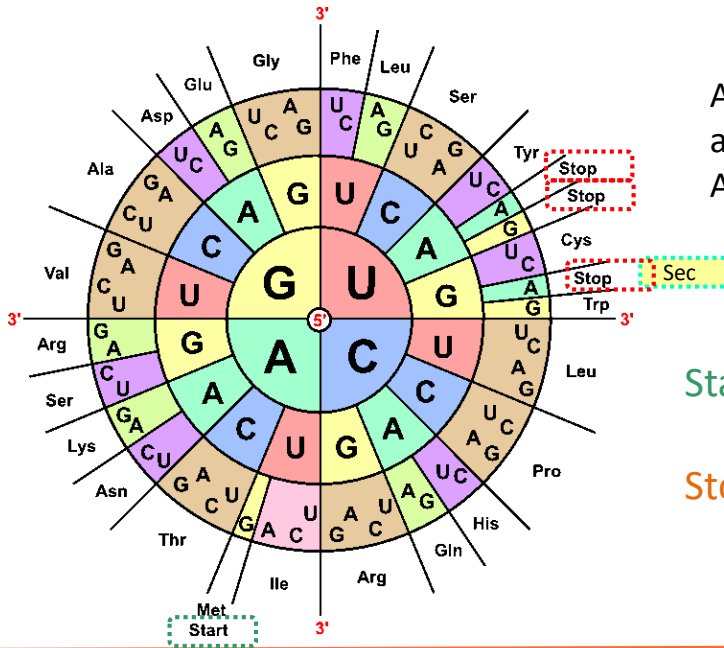
Bioinformática I

PPG biotec

UFPA

toMax

Ignez Caracelli



A tabela mostra os códons e os aminoácidos.
A direção do mRNA é de 5' → 3'.

Start códon: AUG

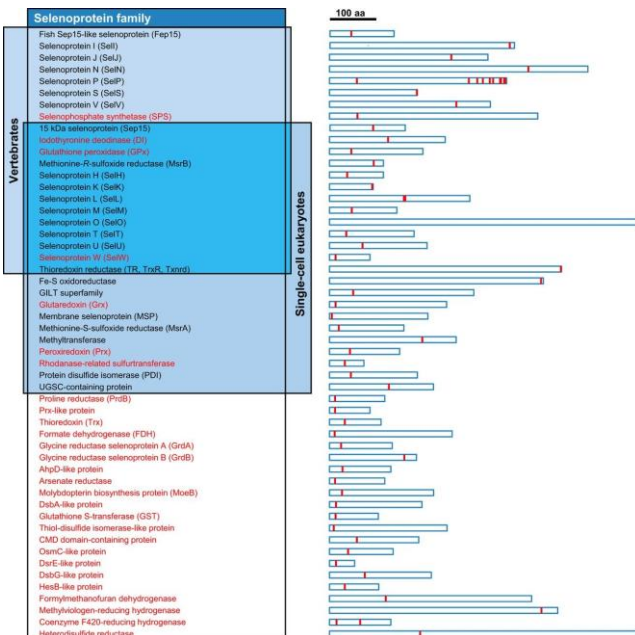
Stop códon: UAA, UAG, UGA

Selênio em proteínas

Selenoprotein families. Selenoproteins that are found in vertebrates or single-cell eukaryotes are indicated by shaded boxes, whereas selenoproteins that occur in prokaryotic organisms are highlighted in red. The relative size of each selenoprotein is shown on the right, and the location of Sec residues is indicated by red lines.

Labunsky, V. M., Hatfield, D. L., & Gladyshev, V. N. (2014). Selenoproteins: molecular pathways and physiological roles. *Physiological reviews*, 94(3), 739-777.

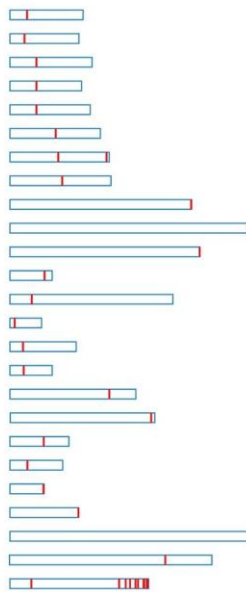
<https://www.physiology.org/doi/full/10.1152/physrev.00039.2013>



<https://www.physiology.org/doi/full/10.1152/physrev.00039.2013>

Selênio em proteínas

Selenoprotein	Abbreviation	Sec location (protein size)	Function
Glutathione peroxidase 1	GPx1, GPX1	47 (201)	Cytosolic glutathione peroxidase
Glutathione peroxidase 2	GPx2, GPX2	40 (190)	Gastrointestinal glutathione peroxidase
Glutathione peroxidase 3	GPx3, GPX3	73 (226)	Plasma glutathione peroxidase
Glutathione peroxidase 4	GPx4, GPX4	73 (197)	Phospholipid hydroperoxide glutathione peroxidase
Glutathione peroxidase 6	GPx6, GPX6	73 (221)	Olfactory glutathione peroxidase
Iodothyronine deiodinase 1	DI1, D1, DIO1	126 (249)	Thyroid hormone-activating iodothyronine deiodinase
Iodothyronine deiodinase 2	DI2, D2, DIO2	133, 266 (273)	Tissue-specific thyroid hormone-activating iodothyronine deiodinase
Iodothyronine deiodinase 3	DI3, D3, DIO3	144 (278)	Tissue-specific thyroid hormone-deactivating iodothyronine deiodinase
Thioredoxin reductase 1	TR1, TrxR1, TXNRD1	498 (499)	Reduction of cytosolic thioredoxin
Thioredoxin/glutathione reductase	TGR, TR2, TrxR3, TXNRD3	655 (656)	Testis-specific thioredoxin reductase
Thioredoxin reductase 3	TR3, TrxR2, TXNRD2	522 (523)	Reduction of mitochondrial thioredoxin and glutaredoxin
Methionine-S-adenosylmethionine transferase	MetB1, SelR, SelX, MSR1	95 (116)	Reduction of oxidized methionine residues
Selenophosphate synthetase 2	SPS2, SEPHS2	60 (448)	Involved in synthesis of selenoproteins
Selenoprotein W	SelW, SEPW1	13 (87)	Unknown
Selenoprotein T	SelT	36 (182)	Unknown
Selenoprotein H	SelH	38 (116)	Unknown
Selenoprotein V	SelV	273 (346)	Unknown
Selenoprotein I	SelI, SEPI, EPT1	387 (397)	Unknown
15 kDa selenoprotein	Sep15	93 (162)	Putative role in quality control of protein folding in the ER
Selenoprotein M	SelM, SEPM	48 (145)	Unknown
Selenoprotein K	SelK	92 (94)	Putative role in ER-associated degradation
Selenoprotein S	SelS, SEPS1, VIMP	188 (189)	Putative role in ER-associated degradation
Selenoprotein O	SelO	667 (669)	Unknown
Selenoprotein N	SelN, SepN, SEPN1	428 (556)	Putative role during muscle development
Selenoprotein P	SelP, SEPP1	59, 300, 318, 330, 345, 352, 367, 369, 376, 378 (381)	Se transport



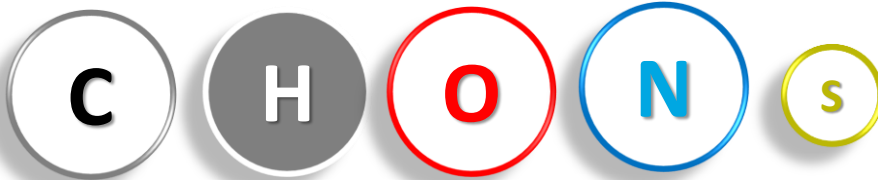
Human selenoproteome. The relative length of selenoproteins and location of Sec within different proteins are shown on the right.

Labunskyy, V. M., Hatfield, D. L., & Gladyshev, V. N. (2014). Selenoproteins: molecular pathways and physiological roles. *Physiological reviews*, 94(3), 739-777.

Aminoácidos

até 1986

20 aminoácidos



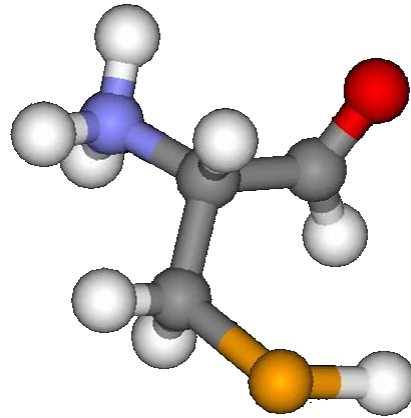
ano: 1986

Aminoácidos

21º aminoácido
SELENOCISTEÍNA
SEC, U

UGA códon
também é código de STOP

Molecular Microbiology (1991), 5(3), 515-520
Eur. J. Biochem. 264, 607-609 (1999)



47

Bioinformática I

Ignez Caracelli

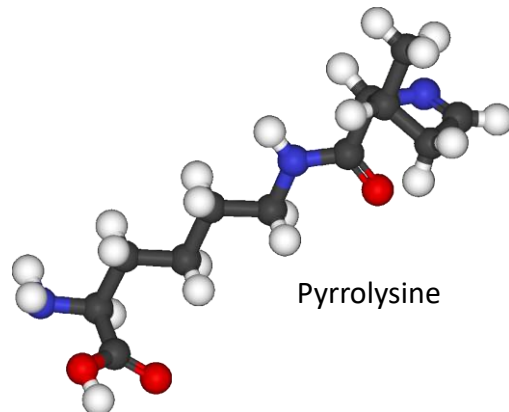
ano: 2002

Aminoácidos

22º aminoácido
PIRROLISINA
PYL, O

UAG códon
também é código de STOP

A **Pirrolisina** é um dos aminoácidos de ocorrência natural, codificados pelo código genético, usado por algumas archaeametanogênicas e uma bactéria *Desulfitobacterium hafniense*.



Pyrrolysine

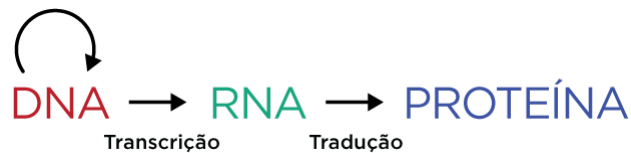
The Journal of Biological Chemistry, **280**, 20740–20751, 2005

48

Bioinformática I

Ignez Caracelli

Síntese Proteica



Dogma central da biologia

49

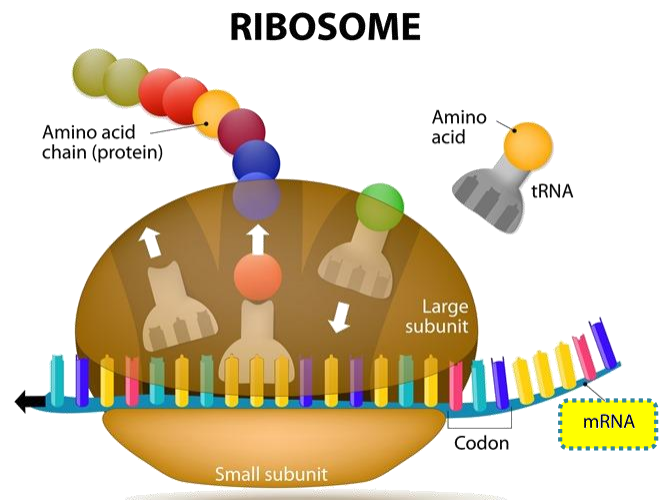
Bioinformática I

Ignez
Caracelli

Síntese Proteica

3 RNAs são necessários para efetuar a síntese protéica:

mRNA (RNA mensageiro):
carrega a “informação” (ou
seja, a sequência de bases)
para a síntese da proteína



50

Bioinformática I

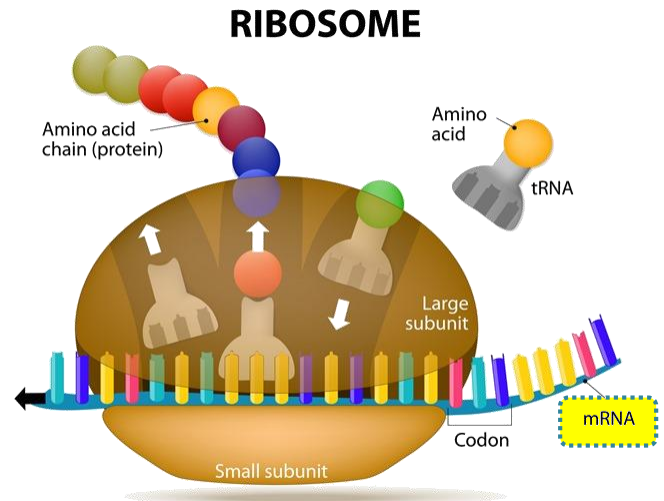
Ignez
Caracelli

Síntese Proteica

3 RNAs são necessários para efetuar a síntese protéica:

rRNA (RNA ribossômico):

é um constituinte estrutural e funcional dos ribossomas, onde a vai acontecer síntese proteica



51

Bioinformática I

PPG biotec

UFPA

SciMax

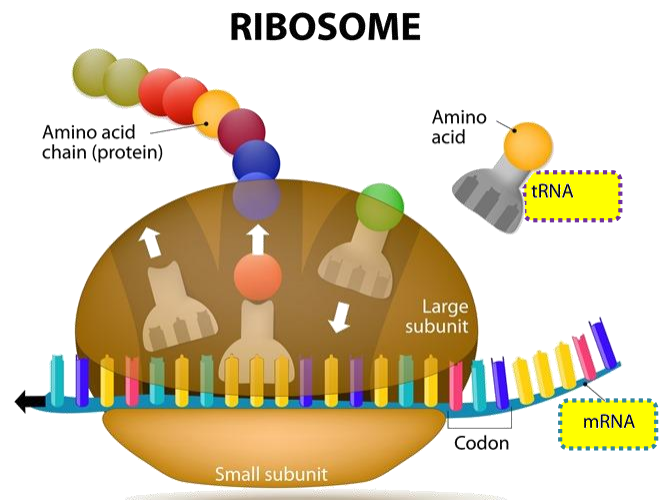
Ignez Caracelli

Síntese Proteica

3 RNAs são necessários para efetuar a síntese protéica:

tRNA (RNA transportador):

carrega os aminoácidos que serão adicionados à proteína nascente, e faz a “leitura” da sequência de bases do mRNA. Isso quer dizer que o tRNA é a molécula que decodifica o código genético



52

Bioinformática I

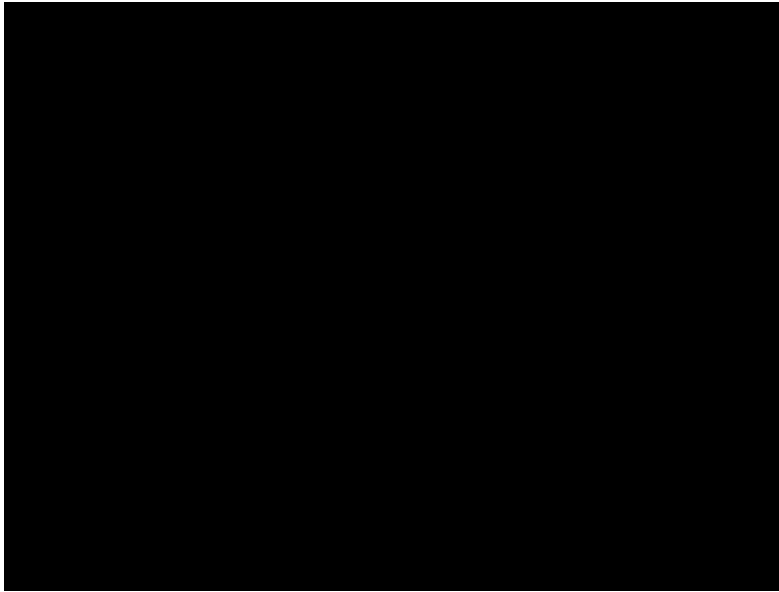
PPG biotec

UFPA

SciMax

Ignez Caracelli

Translação e Síntese



<https://youtu.be/5bLEDd-PSTQ?t=3>

53

Bioinformática I

Ignez
Caracelli

Estrutura Primária

Ordem dos aminoácidos na cadeia

a sequência de aminoácidos é determinada pelo gene (DNA).

cadeia polipeptídica

54

Bioinformática I

Ignez
Caracelli



Estrutura Primária

pequenas mudanças na sequência podem afetar a estrutura e função da proteína.

apenas um aminoácido pode fazer a diferença!

Este é o caso a seguir: **anemia falciforme.**

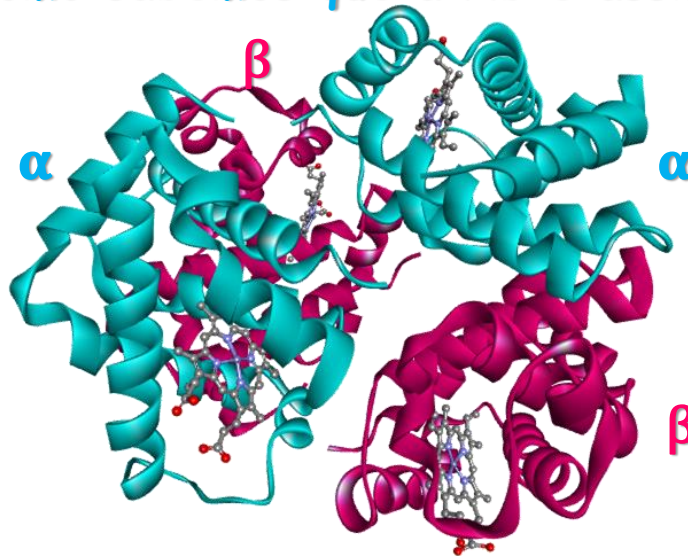
cadeia polipeptídica

55

Bioinformática I

Ignez Caracelli

como sabemos que a Hb é assim?



56 56

Bioinformática I

Ignez Caracelli

Eritrócitos

normais



anemia falciforme



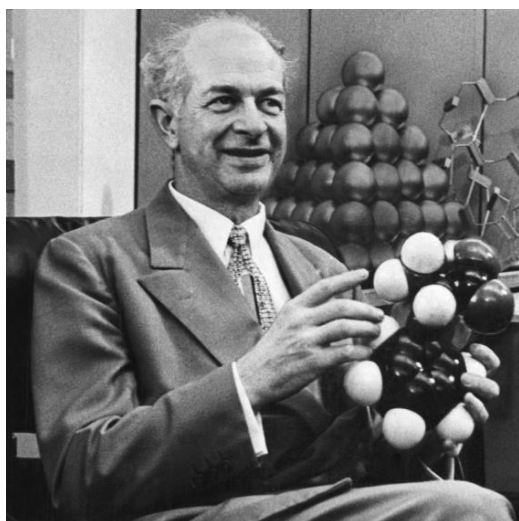
57

Bioinformática I



Ignez
Caracelli

Linus Pauling



Campo(s)

[Química quântica](#), [bioquímica](#),
[cristalógrafo](#), [biólogo molecular](#) e
pesquisador médico

Instituições

[Instituto de Tecnologia da Califórnia](#),
[Universidade da Califórnia em San Diego](#),
[Universidade de Stanford](#)

Alma mater

[Universidade do Estado do Oregon](#),
[Instituto de Tecnologia da Califórnia](#)

Tese

*The Determination with X-Rays of the
Structures of Crystals* (1925)

Conhecido(a) por

elucidar a natureza das [ligações químicas](#) e
da [estrutura das moléculas](#)

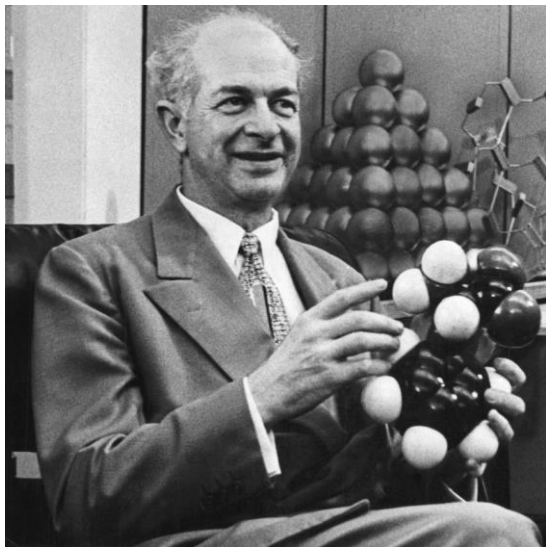
58

Bioinformática I



Ignez
Caracelli

Linus Pauling



Prêmio(s)

[Prêmio ACS de Química Pura](#) (1931),
[Medalha William H. Nichols](#) (1941),
[Medalha Davy](#) (1947),
 1 [Nobel de Química](#) (1954),
 2 [Nobel da Paz](#) (1962),
[Prêmio Linus Pauling](#) (1966),
[Medalha Nacional de Ciências](#) (1974),
[Medalha de Ouro Lomonossov](#) (1977),
[Prêmio em Ciências Químicas NAS](#) (1979),
[Medalha Priestley](#) (1984),
[Prêmio George C. Pimentel de Química](#) (1987),
[Prêmio Vannevar Bush](#) (1989),
[Medalha Benjamin Franklin](#) (1994)

59

Bioinformática I



Ignez Caracelli

Linus Pauling

Scientific contributions

Pauling's discoveries led to decisive contributions in a diverse array of areas including around 350 publications in the fields of

- quantum mechanics,
- inorganic chemistry,
- organic chemistry,
- protein structure,
- molecular biology,
- and medicine.

His work on chemical bonding marks him as one of the founders of modern quantum chemistry.

https://en.wikipedia.org/wiki/Linus_Pauling

60

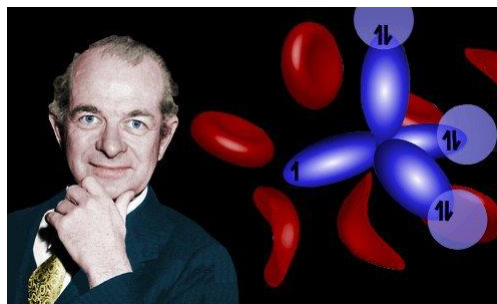
Bioinformática I



Ignez Caracelli

Scientific contributions

The Nature of the Chemical Bond was the standard work for many years, and concepts like hybridization and electronegativity remain part of standard chemistry textbooks.



https://en.wikipedia.org/wiki/Linus_Pauling

61

Bioinformática I

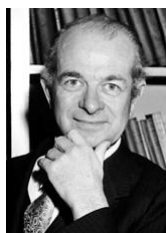
PPG biotec

UFPA

3toMax

Ignez Caracelli

Nobelprize.org



1949 – Linus Pauling mostrou que a anemia falciforme é uma doença molecular, resultante de uma mutação na molécula protéica Hemoglobina

Reprinted from SCIENCE, November 25, 1949, Vol. 110, No. 2865, pages 543-548.

Sickle Cell Anemia, a Molecular Disease¹

Linus Pauling, Harvey A. Itano,² S. J. Singer,² and Ibert C. Wells²

Gates and Crellin Laboratories of Chemistry,
California Institute of Technology, Pasadena, California¹

THE ERYTHROCYTES of certain individuals possess the capacity to undergo reversible changes in shape in response to changes in the partial pressure of oxygen. When the oxygen pressure is lowered, these cells change their forms from the normal biconcave disk to crescent, holly wreath, and other forms. This process is known as sickling. About 8 percent of American Negroes possess this characteristic; usually they exhibit no pathological consequences ascribable to it. These people are said to have sickle-cell anemia, or sickle cell trait. However, about 1 in 40 (4) of these individuals whose cells are capable of sickling suffer from a severe chronic anemia resulting from excessive destruction of their erythrocytes; the term sickle cell anemia is applied to their condition.

The main observable difference between the erythrocytes of sickle cell trait and sickle cell anemia has been that a considerably greater reduction in the partial pressure of oxygen is required for a major fraction of the trait cells to sickle than for the anemia cells (11). Tests *in vivo* have demonstrated that between 30 and 60 percent of the erythrocytes in the venous circulation of sickle cell anemia individuals, but less than 1 percent of those in the venous circulation of sickle cell trait individuals, are normally sickled. Experiments *in vitro* indicate that under sufficiently low oxygen pressure, however, all the cells of both types assume the sickled form.

The evidence available at the time that our investigation was begun indicated that the process of sickling might be intimately associated with the state and the nature of the hemoglobin within the erythrocyte. Sickled cell erythrocytes in which the hemoglobin is combined with oxygen or carbon monoxide have the biconcave disk contour and are indistinguishable in appearance from normal erythrocytes. In this condition they are termed pronormoerythrocytes. The hemoglobin appears to be uniformly distributed and randomly oriented within normal cells and pronormoerythrocytes, and no birefringence is observed. Both types of cells are very flexible. If the oxygen or carbon monoxide is removed, however, transforming the hemoglobin to the uncombined state, the pronormoerythrocytes undergo sickling. The hemoglobin within the sickled cells appears to aggregate into one or more foci, and the cell membranes collapse. The cells become birefringent (12) and quite rigid. The addition of oxygen or carbon monoxide to these cells reverses these phenomena. Thus the physical effects just described depend on the state of combination of the hemoglobin, and only secondarily, if at all, on the cell membrane. This conclusion is supported by the observation that sickled cells when lysed with water produce discs, rather than sickle-shaped, ghosts (10).

It was decided, therefore, to examine the physical and chemical properties of the hemoglobins of individuals with sickle cell anemia and sickle cell trait, and to compare them with the hemoglobin of normal individuals to determine whether any significant differences might be observed.

EXPERIMENTAL METHODS

The experimental work reported in this paper deals largely with an electrophoretic study of these hemoglobins. In the first phase of the investigation, which concerned the comparison of normal and sickle cell anemia hemoglobins, three types of experiments were performed: 1) with carbonmonoxyhemoglobins; 2) with uncombined ferrohemoglobins in the presence of dithionite ion, to prevent oxidation to methemoglobins; and 3) with carbonmonoxyhemoglobins in the presence of dithionite ion. The experiments of type 3 were performed and compared with those of type 1 in order to ascertain whether the dithionite ion itself causes any specific electrophoretic effect.

Samples of blood were obtained from sickle cell anemia individuals who had not been transfused within three months prior to the time of sampling. Strama-

Doença Molecular

62

UFPA

3toMax

Ignez Caracelli

Doença Molecular

1949 – Linus Pauling mostrou que a anemia falciforme é uma doença molecular, resultante de uma mutação na molécula protéica Hemoglobina

...sickle-cell hemoglobin carried more positive electric charges on its surface

A slight change in the electrical charge of a single type of molecule in the body could spell the difference between life and death. Never before had the cause of a disease been traced to a molecule. This discovery--to which Pauling attached the memorable title "molecular disease"--received widespread attention. Itano and Singer's subsequent work demonstrated a pattern of inheritance for the disease, firmly wedding molecular medicine to genetics.

63

Bioinformática I



Ignez Caracelli

Hemoglobina & Anemia Falciforme



Val	His	Leu	Thr	Pro	Glu	Glu	...
1	2	3	4	5	6	7	

eritrócitos com
hemoglobina normal



Val	His	Leu	Thr	Pro	Val	Glu	...
1	2	3	4	5	6	7	

eritrócitos com
hemoglobina S

64

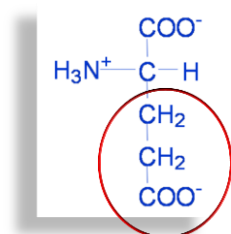
Bioinformática I



Ignez Caracelli

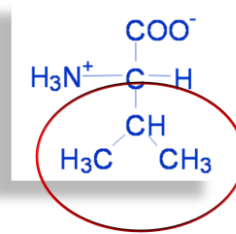
O resultado da mutação

o aminoácido valina substitui o ácido glutâmico



Ácido Glutâmico
Glu, E

aminoácido **polar carregado**
negativamente



Valina
Val, V

aminoácido **não-polar**

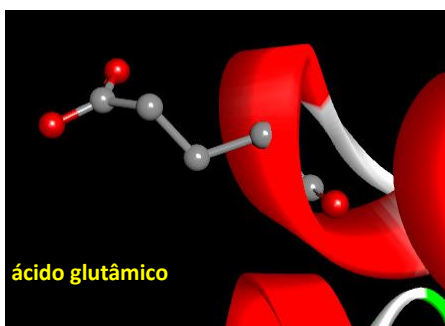
65

Bioinformática I

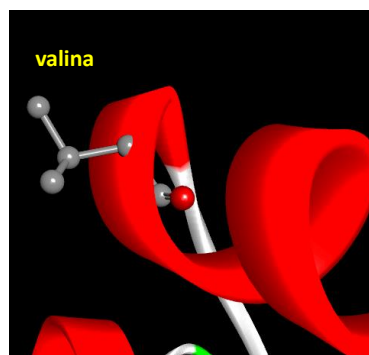
Ignez
Caracelli

O resultado da mutação

o aminoácido valina substitui o ácido glutâmico



eritrócitos com
hemoglobina normal



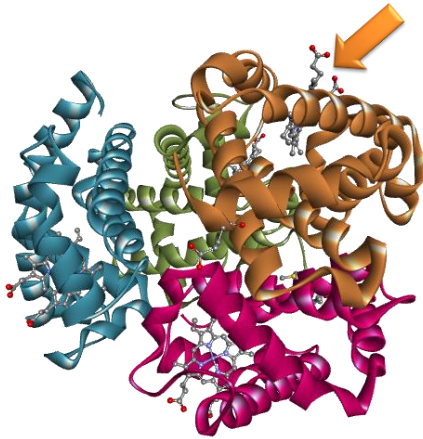
eritrócitos com
hemoglobina S

66

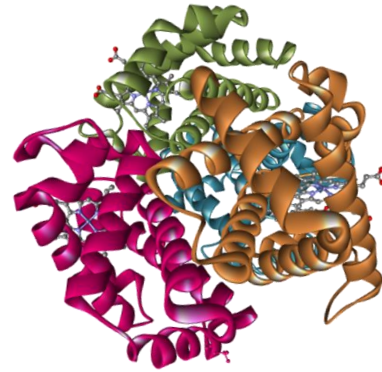
Bioinformática I

Ignez
Caracelli

hemoglobina normal



tetrâmero

O resultado da mutação
hemoglobina S

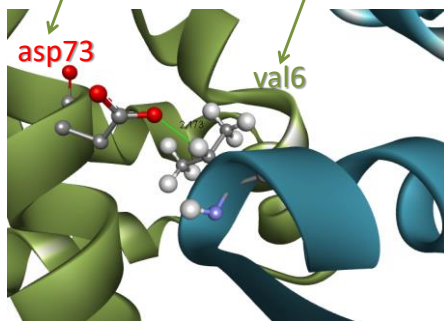
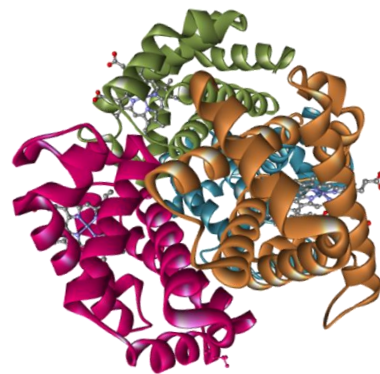
tetrâmero

67

Bioinformática I

Ignez
Caracelli

polar carregado

não-polar
mutaçãohemoglobina normal
glu 6 (polar carregado)O resultado da mutação
hemoglobina S

tetrâmero

68

Bioinformática I

Ignez
Caracelli

Distorção das células sanguíneas vermelhas

As moléculas de hemoglobina S repelem a água e são atraídas uma pela outra.

Formam-se filamentos.

Os filamentos distorcem a forma das células sanguíneas.

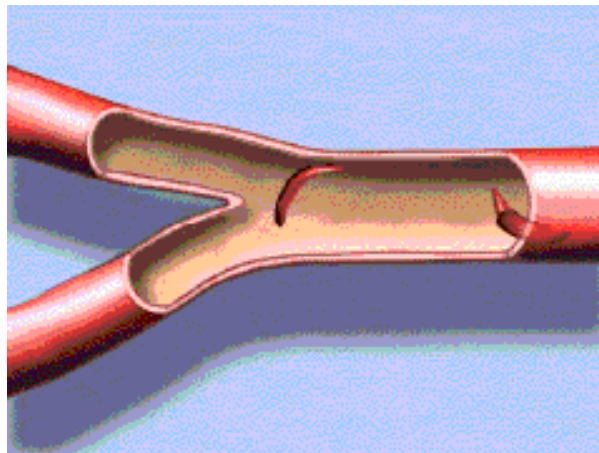
As células com HbS são rígidas e isso não facilita o transporte de O_2 .

69

Bioinformática I

Ignez
Caracelli

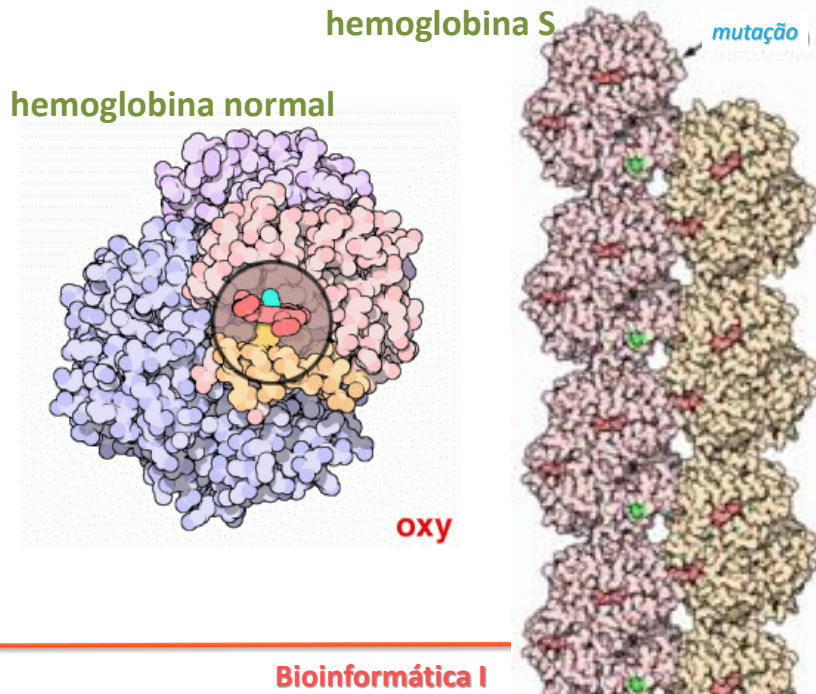
Bloqueio Capilar



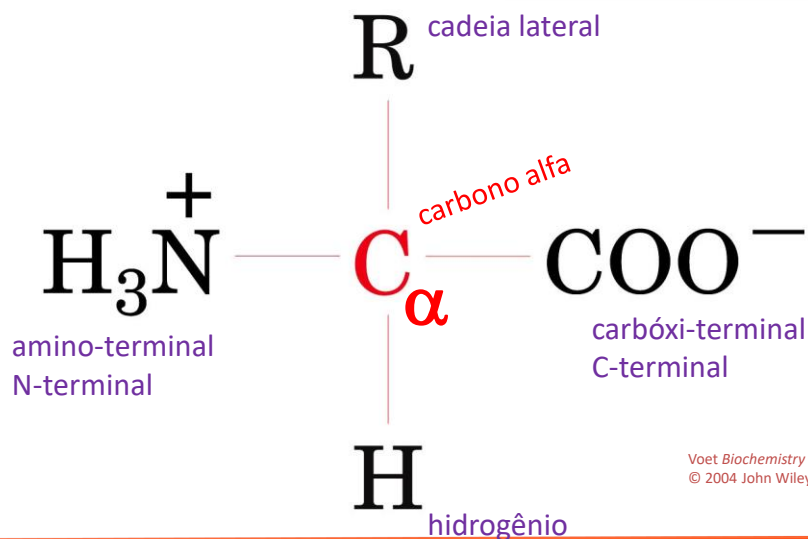
70

Bioinformática I

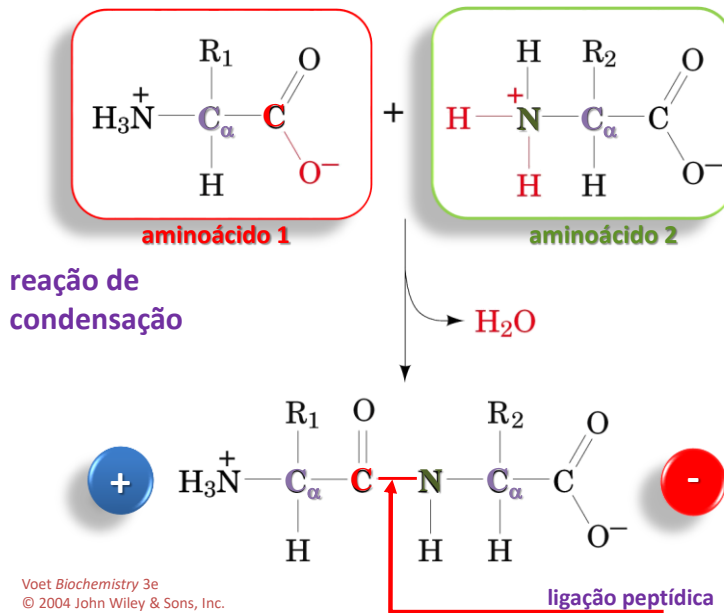
Ignez
Caracelli



O aminoácido em sua forma ionizada
forma zwitteriônica



Ligação peptídica



73

Bioinformática I

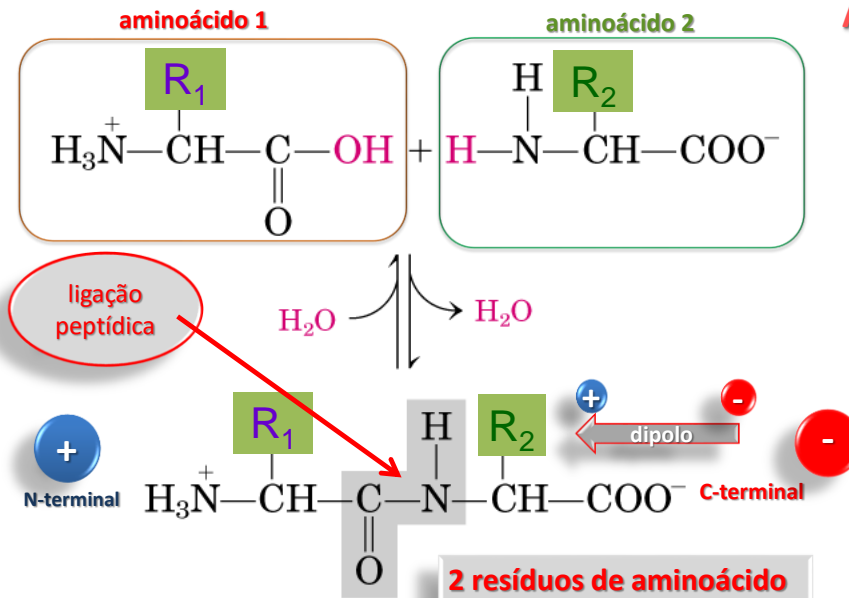
PPG biotec

UFPA

SciMax

Ignez Caracelli

A ligação peptídica



74

Bioinformática I

PPG biotec

UFPA

SciMax

Ignez Caracelli



Estrutura Primária

Os aminoácidos estão conectados por **ligações peptídicas** e formam a cadeia polipeptídica.

cadeia polipeptídica

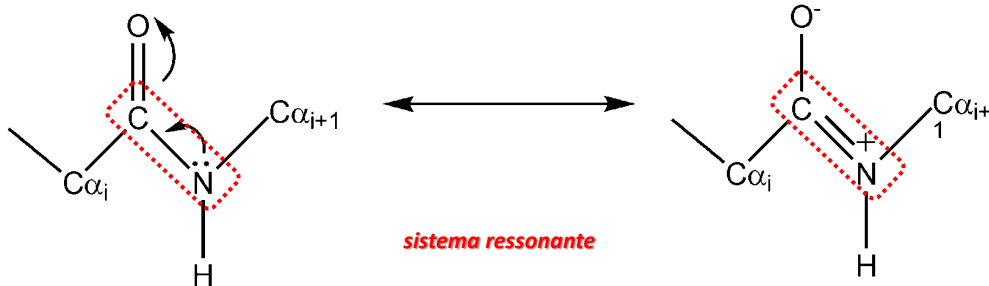
75

Bioinformática I



Ignez Caracelli

A ligação peptídica



ligação peptídica



76

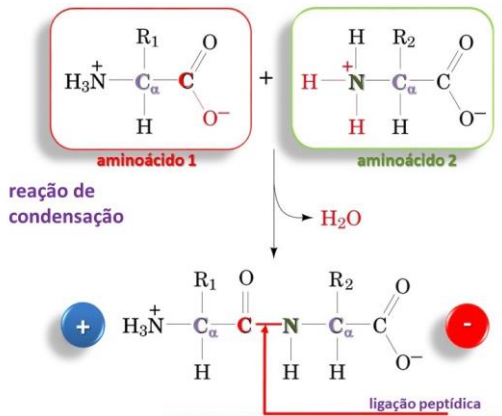
Bioinformática I



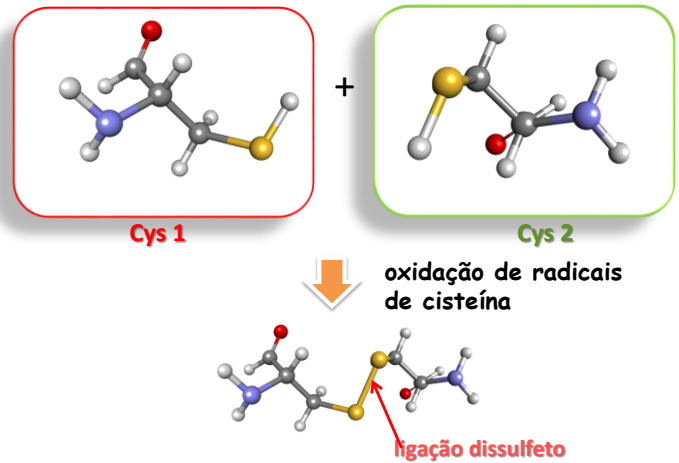
Ignez Caracelli

Ligações covalentes

ligação peptídica



ligação dissulfeto



77

Bioinformática I

PPG biotec

UFPE-IT

SciMax

Ignez Caracelli

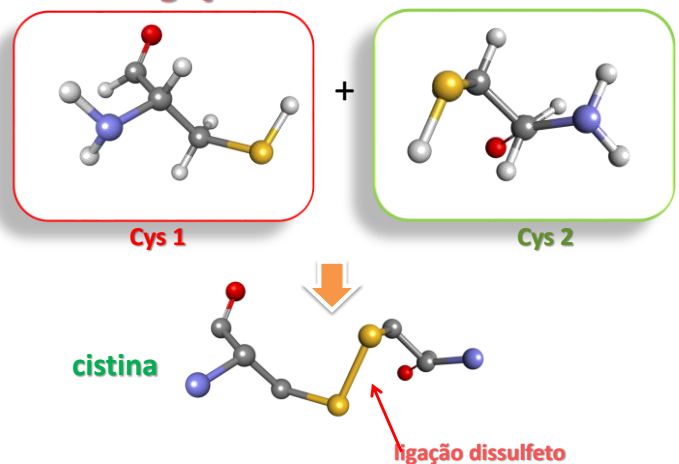
As ligações dissulfeto têm um papel importante na manutenção da estrutura 3D de algumas proteínas



Cistina: composta por duas cisteínas unidas por uma **ligação dissulfeto**.

Ligações covalentes

ligação dissulfeto



78

Bioinformática I

PPG biotec

UFPE-IT

SciMax

Ignez Caracelli

Arquivo texto de proteínas

arquivo texto das coordenadas atômicas de um aminoácido

					x	y	z
ATOM	1084	N	XXX A 138		3.800	20.200	9.600
ATOM	1085	CA	XXX A 138		4.500	21.400	9.300
ATOM	1086	C	XXX A 138		4.200	21.700	7.800
ATOM	1087	O	XXX A 138		3.900	22.800	7.400
ATOM	1088	CB	XXX A 138		5.900	21.100	9.900
ATOM	1089	CG	XXX A 138		7.000	21.900	9.300
ATOM	1090	CD1	XXX A 138		7.500	21.600	8.100
ATOM	1091	CD2	XXX A 138		7.400	23.100	9.900
ATOM	1092	CE1	XXX A 138		8.600	22.300	7.500
ATOM	1093	CE2	XXX A 138		8.400	23.900	9.200
ATOM	1094	CZ	XXX A 138		9.000	23.500	8.100

79

Bioinformática I

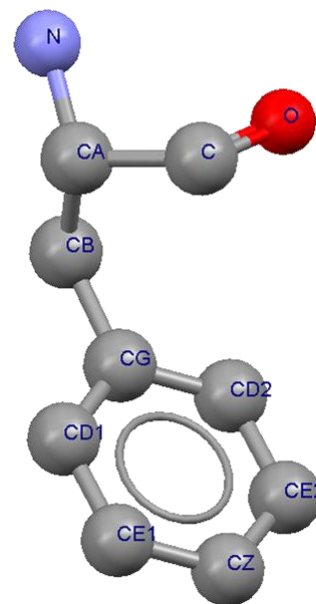


Ignez Caracelli

Arquivo texto de proteínas

arquivo texto das coordenadas atômicas de um aminoácido

					x	y	z
ATOM	1084	N	XXX A 138		3.800	20.200	9.600
ATOM	1085	CA	XXX A 138		4.500	21.400	9.300
ATOM	1086	C	XXX A 138		4.200	21.700	7.800
ATOM	1087	O	XXX A 138		3.900	22.800	7.400
ATOM	1088	CB	XXX A 138		5.900	21.100	9.900
ATOM	1089	CG	XXX A 138		7.000	21.900	9.300
ATOM	1090	CD1	XXX A 138		7.500	21.600	8.100
ATOM	1091	CD2	XXX A 138		7.400	23.100	9.900
ATOM	1092	CE1	XXX A 138		8.600	22.300	7.500
ATOM	1093	CE2	XXX A 138		8.400	23.900	9.200
ATOM	1094	CZ	XXX A 138		9.000	23.500	8.100



80

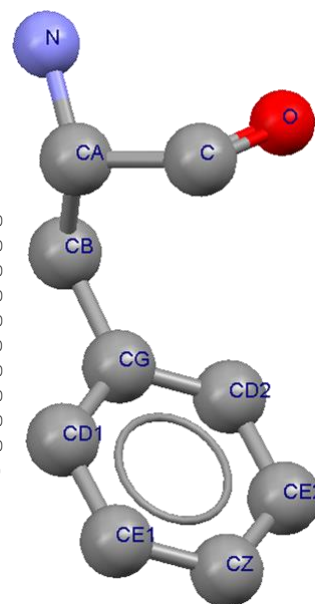
Bioinformática I



Ignez Caracelli

Qual o aminoácido?

				x	y	z	
ATOM	1084	N	PHE A 138	3.800	20.200	9.600	1.00
ATOM	1085	CA	PHE A 138	4.500	21.400	9.300	1.00
ATOM	1086	C	PHE A 138	4.200	21.700	7.800	1.00
ATOM	1087	O	PHE A 138	3.900	22.800	7.400	1.00
ATOM	1088	CB	PHE A 138	5.900	21.100	9.900	1.00
ATOM	1089	CG	PHE A 138	7.000	21.900	9.300	1.00
ATOM	1090	CD1	PHE A 138	7.500	21.600	8.100	1.00
ATOM	1091	CD2	PHE A 138	7.400	23.100	9.900	1.00
ATOM	1092	CE1	PHE A 138	8.600	22.300	7.500	1.00
ATOM	1093	CE2	PHE A 138	8.400	23.900	9.200	1.00
ATOM	1094	CZ	PHE A 138	9.000	23.500	8.100	1.00



81

81

Bioinformática I

Ignez
Caracelli