



## Tópicos em Biotecnologia III

# CONSTRUÇÃO DE UMA VACINA DE MRNA CONTRA SARS-COV-2



Beatriz Oliveira M. de Carvalho, N° USP13658036

Bruno Correia de Oliveira, N° USP 12540930

Cauê Leite Freire, N° USP 13745711

Camilly Vitória Barbosa Pires, N° USP13716826

Carolina Carla Valera Sanchez, N° USP 13658102

Emilly Regina Ramos, N° USP 13658119

Gabriel Henrique Caceres da Silva N° USP 13658099

Gustavo Fortunato Passos N° USP 13727099

Henrique Trevisam de Oliveira, N° USP 13658022

Iara Vilarino Chartouni, N° USP 13658078

Jakeline de Oliveira Carvalho, N° USP 13658123

Jonas Albuquerque Nascimento, N° USP 13827862

# TÓPICOS



1

IMUNIDADE À VÍRUS

2

VACINAS DE MRNA

3

INFORMAÇÕES DE OBTENÇÃO DO VÍRUS

4

OBTENÇÃO E INFORMAÇÕES ACERCA DA SPIKE

5

REGIÕES UTR E PASSO A PASSO DA INSERÇÃO

6

ESTRUTURA SECUNDÁRIA DAS REGIÕES 5'UTR E 3'UTR E VACINA (PRÉ-OTIMIZAÇÃO)

7

OTIMIZAÇÃO DE CÓDONS

8

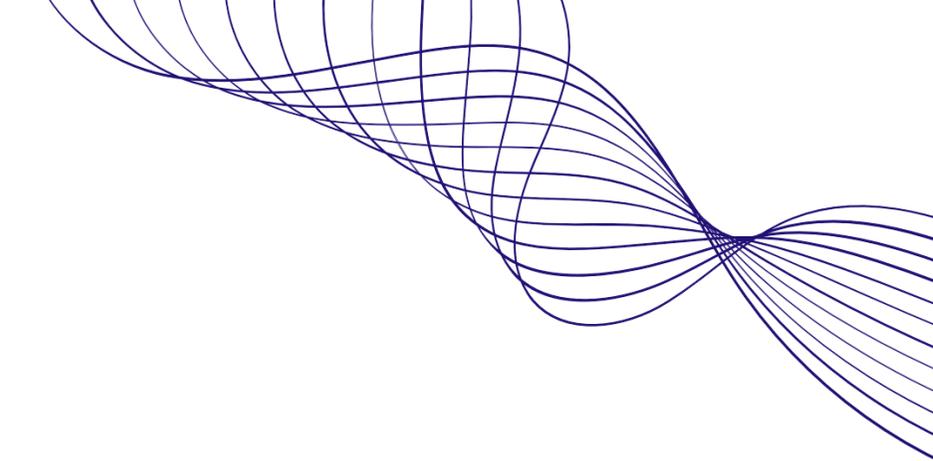
ENRIQUECIMENTO CG

9

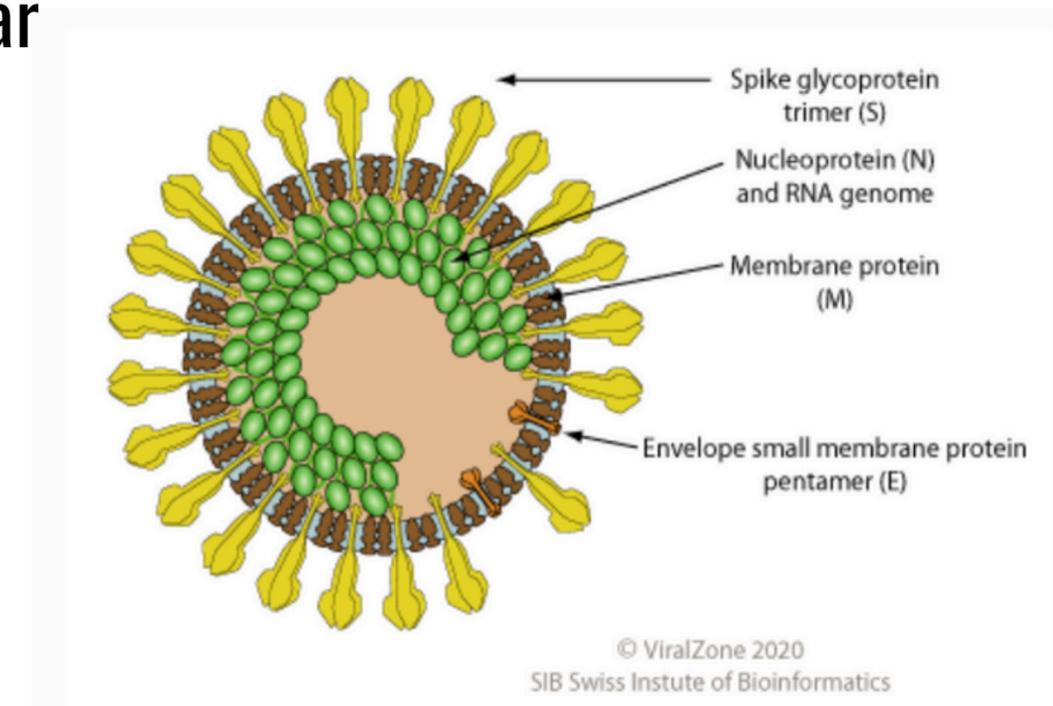
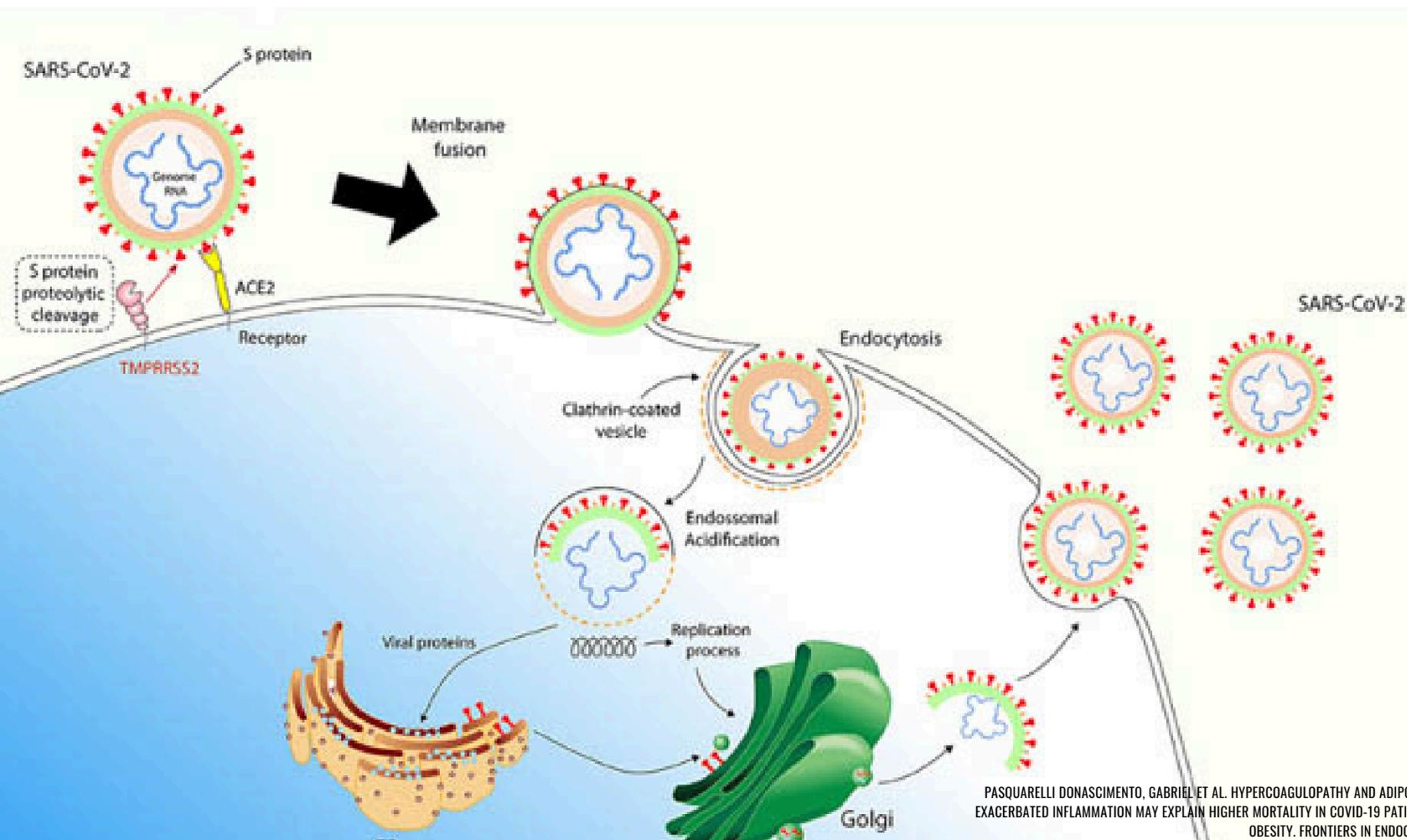
PREDIÇÃO DE EPÍTOPOS DE ANTICORPOS

10

MHC-I E MHC-II



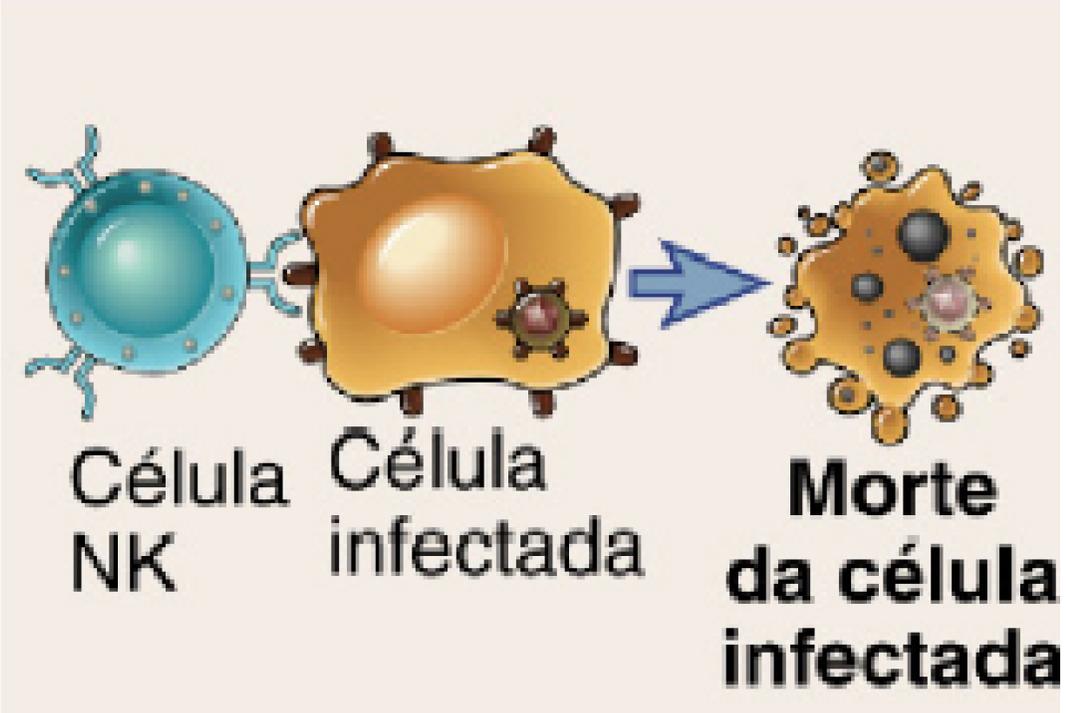
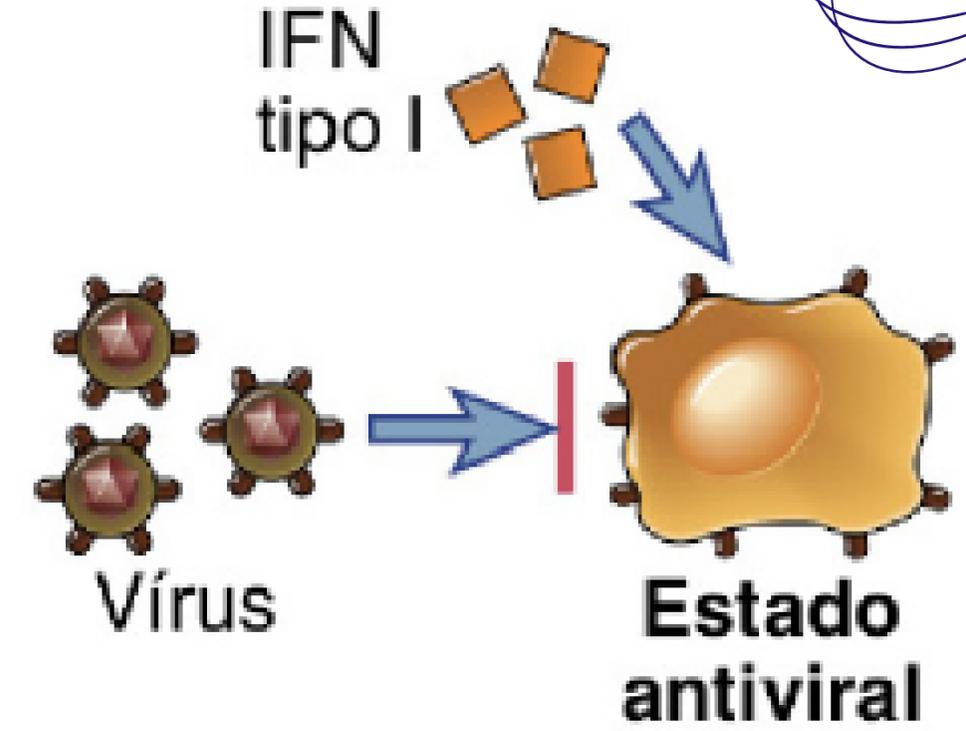
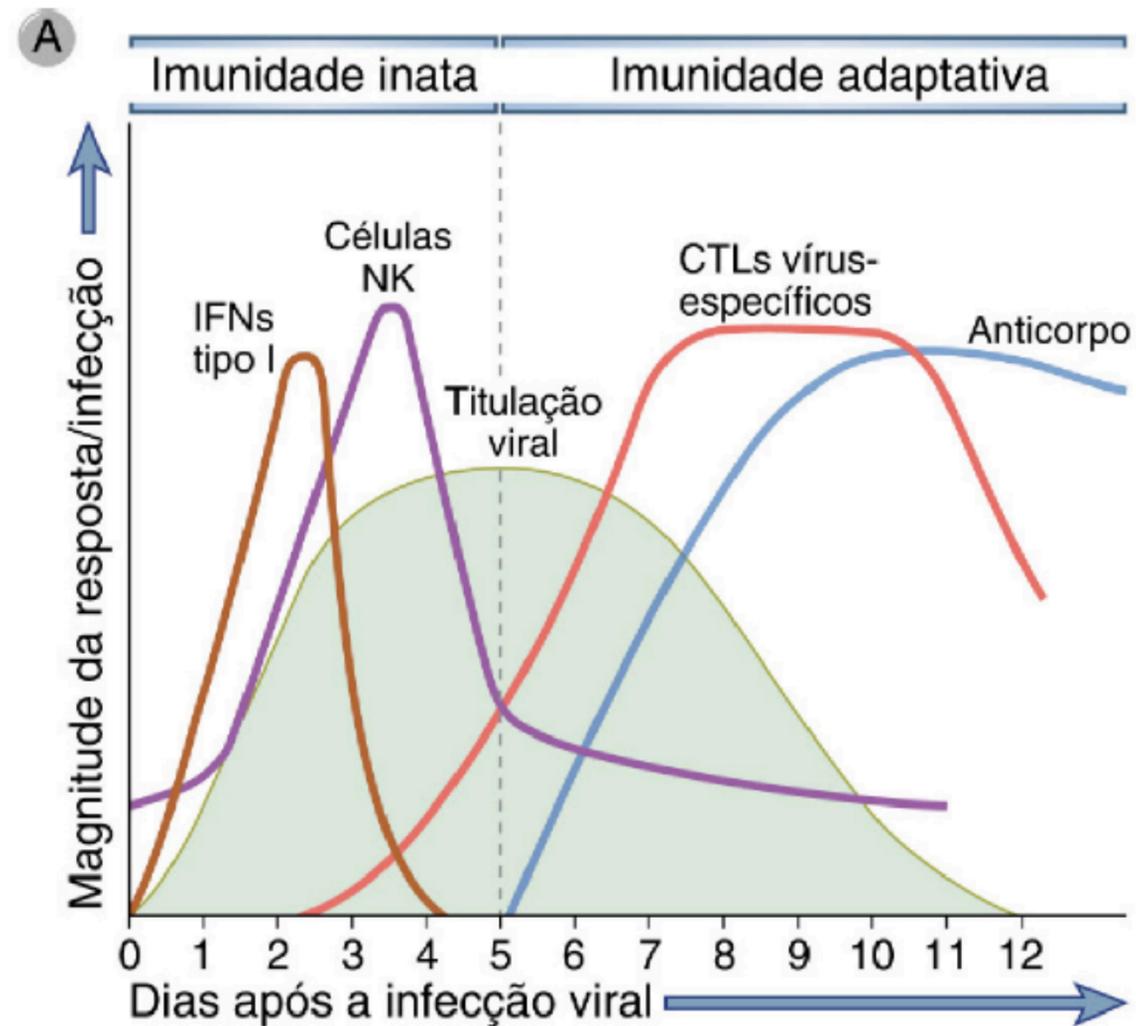
Parasitas intracelulares obrigatórios que usam componentes do ácido nucleico e a maquinaria de síntese proteica do hospedeiro para se replicar



- A replicação viral interfere na síntese proteica e na função celular normal
- Estímulo de respostas inflamatórias -> danos teciduais

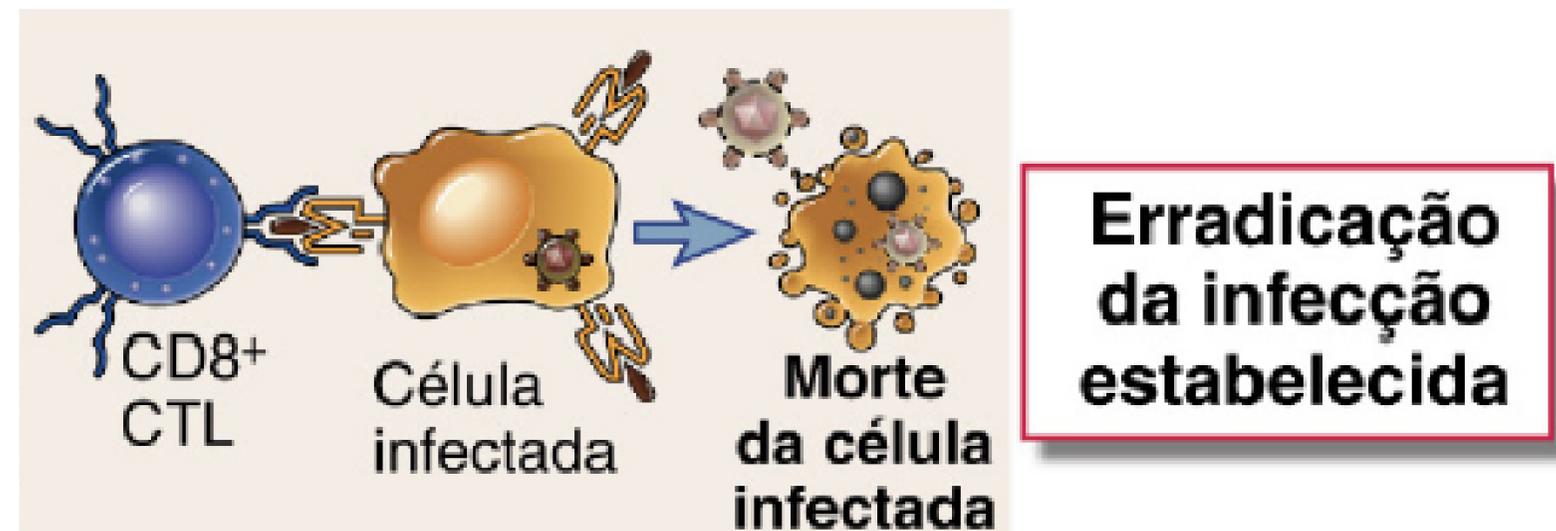
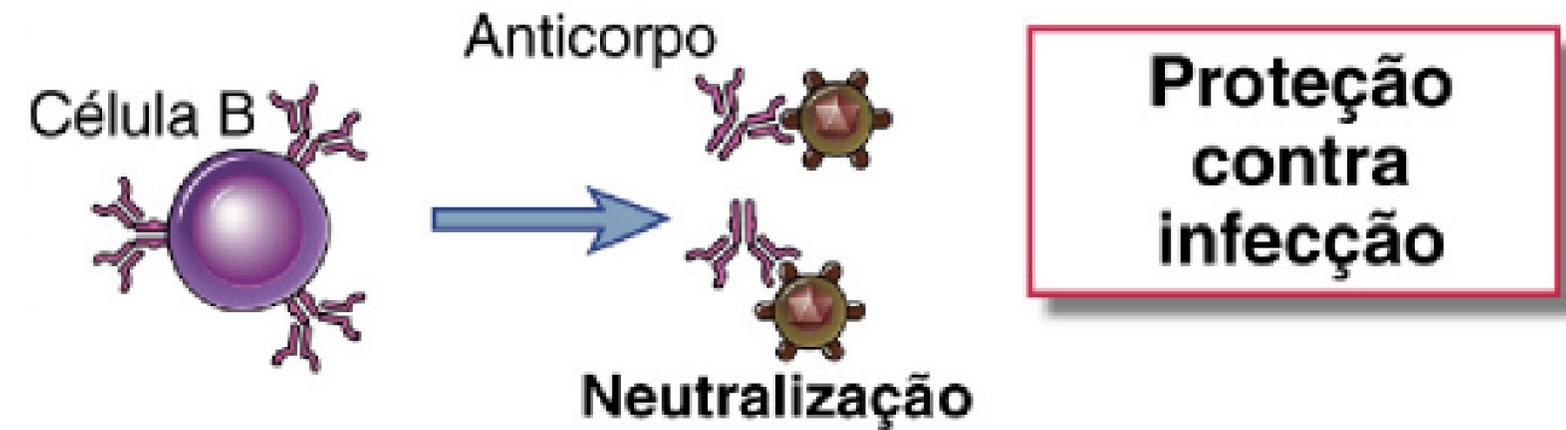
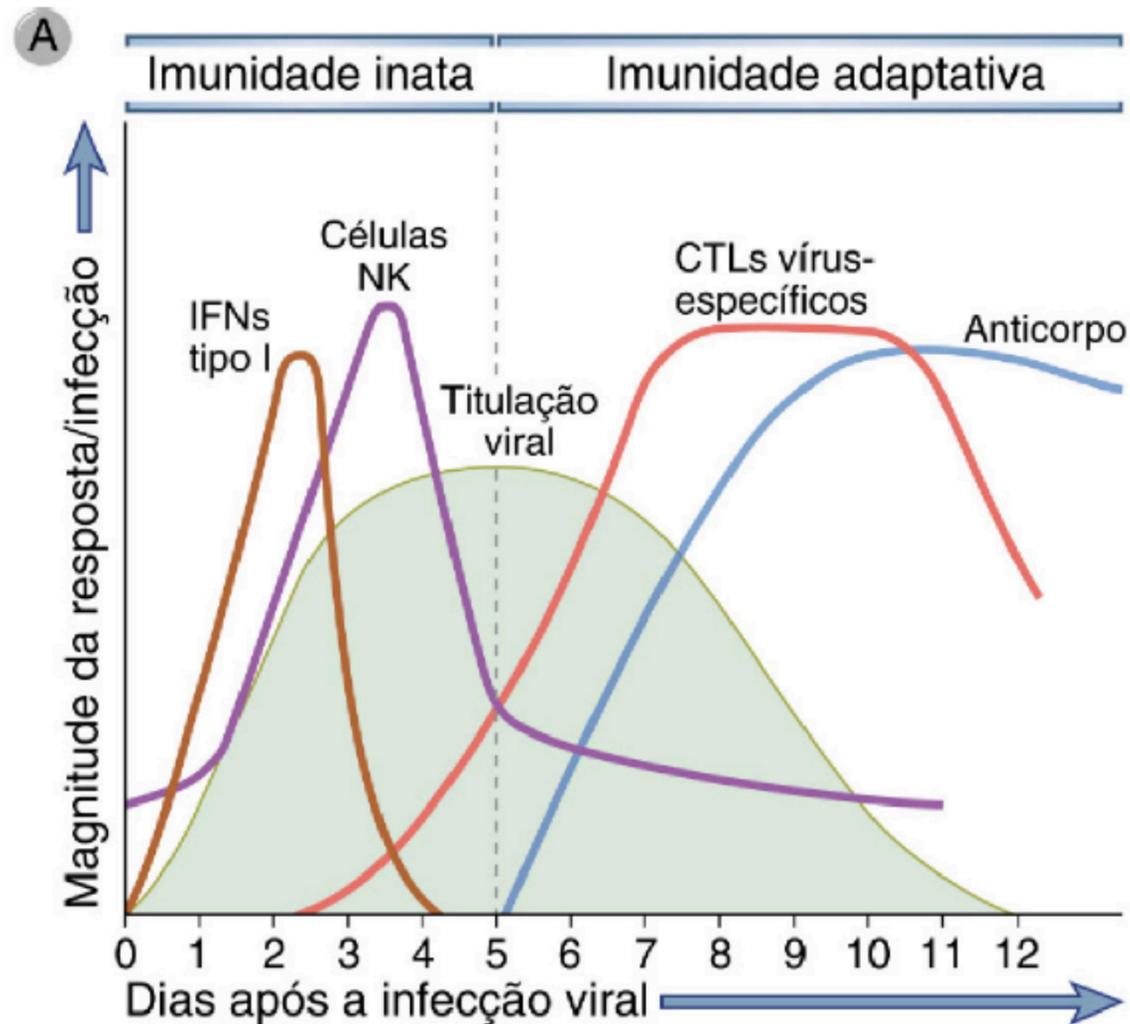
# IMUNIDADE INATA AOS VÍRUS

1. Inibição da infecção por interferons do tipo I e II
2. Killing das células infectadas mediado por células NK

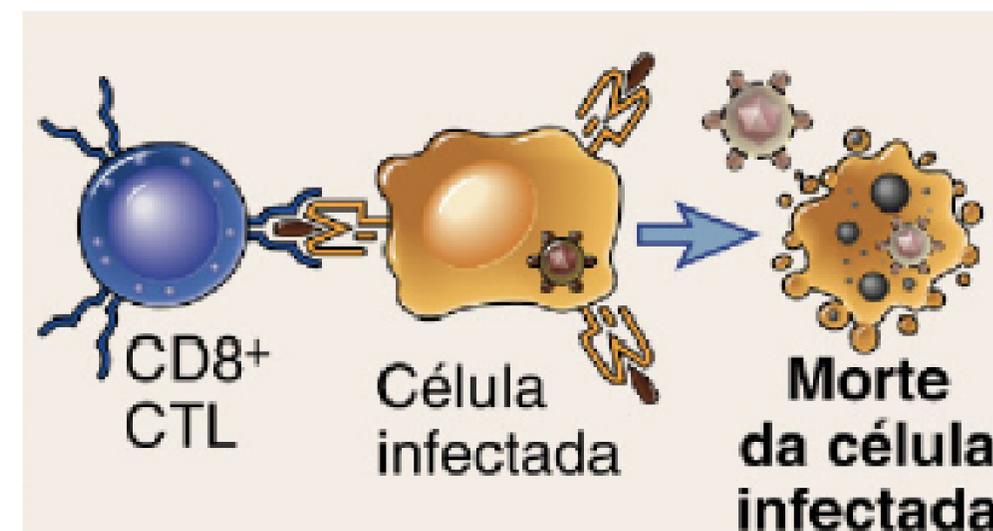
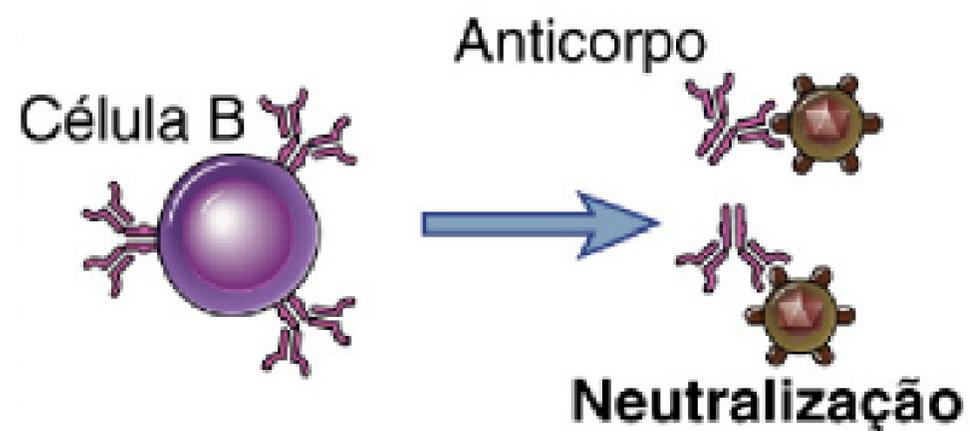


# IMUNIDADE ADAPTATIVA AOS VÍRUS

- 1. Anticorpos: neutralização e opsonização
- 2. Células T CD8+: destruição das células infectadas



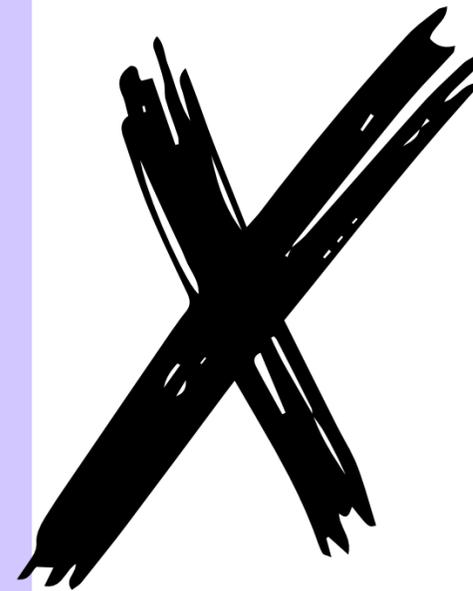
**O princípio fundamental da vacinação é administrar uma forma morta ou atenuada de um agente infeccioso ou um componente de um microrganismo que não causa a doença, mas provoca uma resposta imune que fornece proteção contra a infecção pelo microrganismo patogênico vivo.**



# VACINA DE mRNA

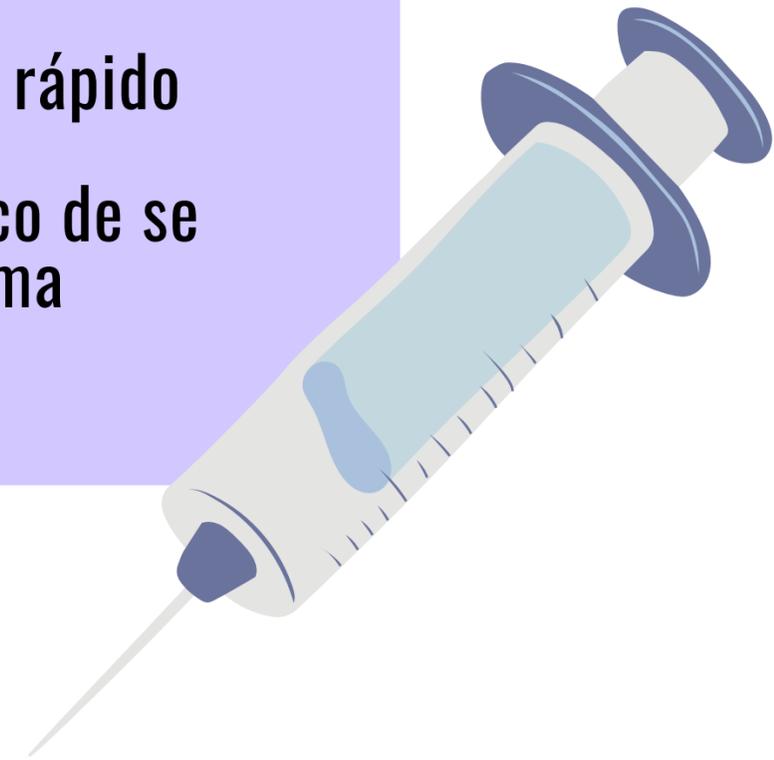
## Vacinas convencionais

- Vírus inativado ou atenuado
- Desenvolvimento lento
- Pode se integrar ao genoma

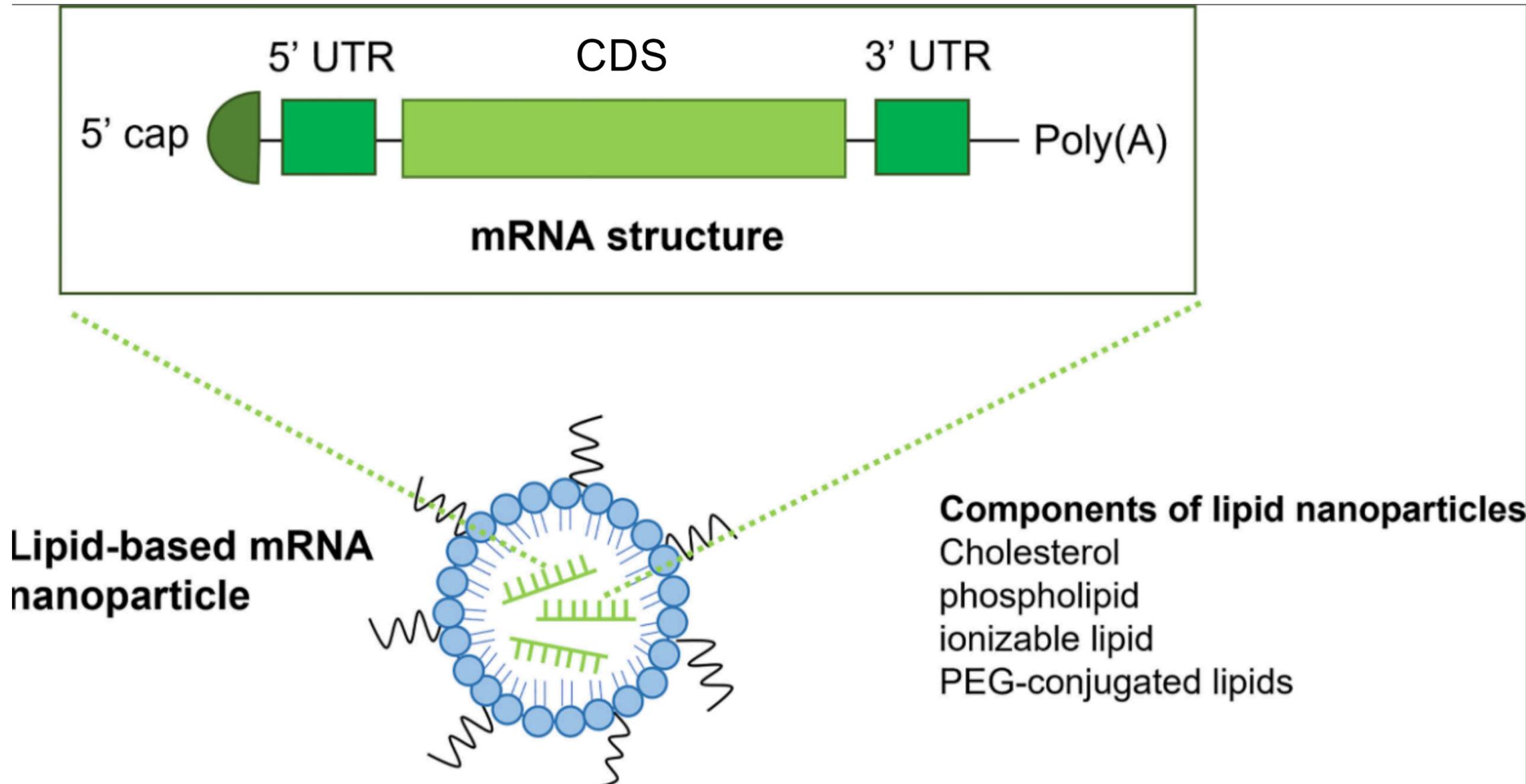


## Vacinas de mRNA

- Diretamente traduzido
- Desenvolvimento rápido
- Não possuem risco de se integrar ao genoma



# VACINA DE mRNA

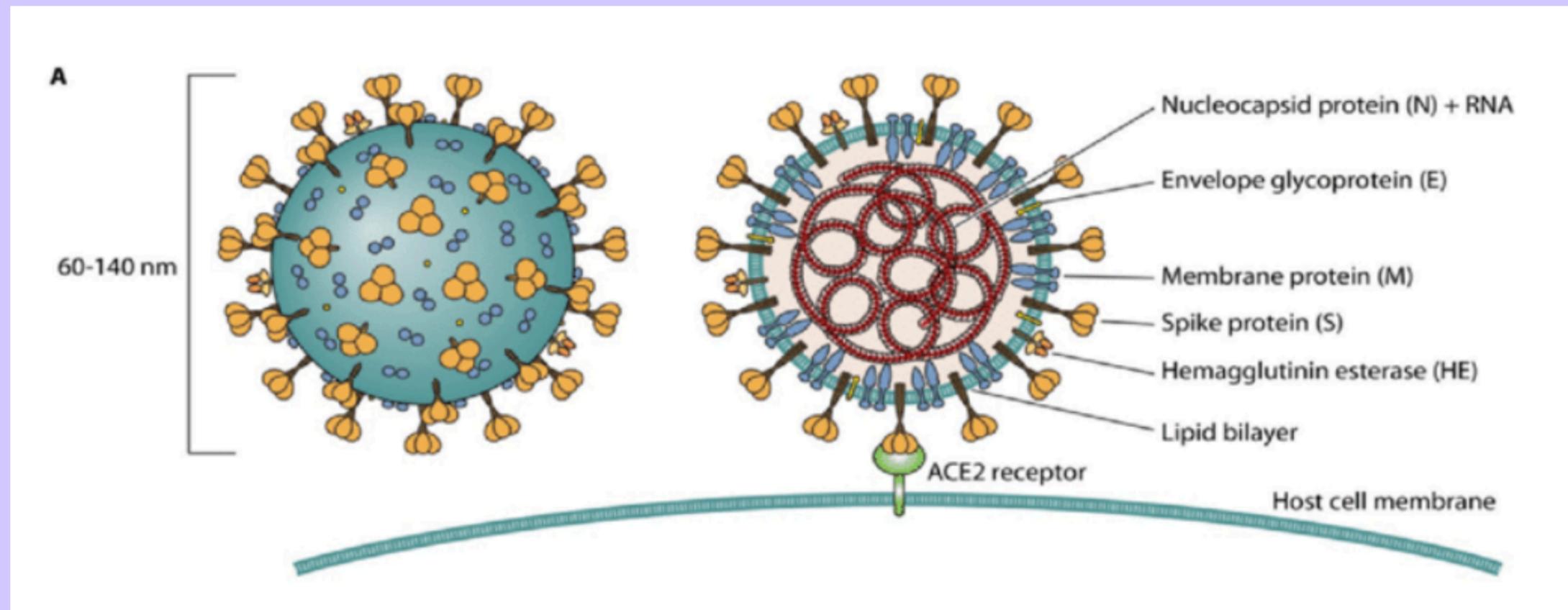


Modificado para aumentar sua durabilidade e reduzir a degradação pelo sistema imunológico.

# VACINA DE mRNA



## Proteína Spike



# Objetivo

**Desenhar uma vacina de mRNA otimizada para o vírus Sars-CoV-2 original (Wuhan) utilizando a proteína Spike e as regiões 5' UTR e 3' UTR da  $\beta$ -globina**

# Ferramentas

Objetivo gerais: Montar uma sequência de nucleotídeos para o desenvolvimento de uma potencial vacina de mRNA contra a Covid-19.

1



SARS-CoV-2 original (Wuhan – depositado no GISAID sob número: EPI\_ISL\_402124);

2

Obtenção da sequência de cDNA da proteína Spike original;

3

Site Primer Show: Tradução da proteína spike do vírus SARS-CoV-2 original. Localização das posições em que foram realizadas as duas mutações missense das vacinas de mRNA da Pfizer e da Moderna;

4

Obtenção do cDNA da proteína spike do vírus original contendo as duas mutações missense nas posições K986P e V987P.

# Ferramentas

Objetivo gerais: Montar uma sequência de nucleotídeos para o desenvolvimento de uma potencial vacina de mRNA contra a Covid-19.

1



Busca de informações sobre a hemoglobin subunit beta (HBB).

3

Determinação do tamanho de molécula de mRNA que codifica a proteína spike e contém as regiões 5'UTR e 3'UTR do gene da beta globina.

2

Adição das respectivas sequências 5'UTR e 3'UTR juntamente ao cDNA contendo o cDNA da proteína spike de interesse com as duas mutações missense 2P.

4



Determinação da estrutura secundária da vacina de mRNA e a energia livre, medida da estabilidade da estrutura secundária da vacina de mRNA.

# Ferramentas

Objetivo gerais: Montar uma sequência de nucleotídeos para o desenvolvimento de uma potencial vacina de mRNA contra a Covid-19.

1



**Codon Optimization Tool da IDT (Integrated DNA Technologies) para a otimização de códons da região codificadora.**

3



**Informações sobre o enriquecimento CG após a otimização de códons.**

2

**ViennaRNA Web Services**  
Institute for Theoretical Chemistry

**Após a otimização dos códons, foi determinada novamente a estrutura secundária e a energia livre da molécula de mRNA relativa à vacina.**

# Ferramentas

Objetivo geral: Análise da Imunogenicidade das respectivas moléculas de mRNA por meio do banco IEDB.

1

**Expasy** 

Conversão de uma sequência de nucleotídeos em sequência de aminoácidos.

3

Predição de epítopos de MHC de classe I;

2

 **IEDB**  
Immune Epitope Database & Tools

Determinação da predição de epítopos de anticorpos da proteína spike codificada pela vacina de mRNA, antes e após a adição das mutações missense 2P.

4

Predição de epítopos de MHC de classe II.

# INFORMAÇÕES E OBTENÇÃO DO VÍRUS

SARS-CoV-2 original (Wuhan – depositado no GISAID sob número: EPI\_ISL\_402124).

A screenshot of the GISAID EpiCoV™ web interface. The top navigation bar includes links for Registered Users, EpiFlu™, EpiCoV™ (selected), EpiPox™, EpiArbo™, and My Profile. Below the navigation bar, there are links for Search, Downloads, and Upload, with a sub-header indicating "targets, phylogenetic trees and many more." A search bar is located on the right side. The main content area features a grid of tool icons: Audacity (a tree with colorful leaves), AudacityInstant (a tree with a DNA helix), BLAST (a colorful DNA helix), CoVizu<sup>e</sup> (a phylogenetic tree with labels like B.1.144, B.1.67, B.1.82, B.1.132, B.1.145, B.1.71, B.1.150, B.1.129, B.1.139, B.1.78, B.1.128), Emerging Variants (three orange arrows pointing up), Lineage Frequency (a colorful area chart), Official GISAID reference sequence (a colorful bar chart with labels like ORF1a, ORF1b, E, M, ORF3, ORF3a, ORF3b), PrimerChecker (a colorful circular pattern), Submission tracker (a world map), Spike glycoprotein mutation surveillance (a phylogenetic tree with a virus particle), and Wastewater (a blue water drop with a trash can icon).

Fonte: Autoral. 2024.

Plataforma que fornece o compartilhamento de informações genômicas e epidemiológicas sobre o vírus da gripe e outros, como o SARS-CoV-2.



# INFORMAÇÕES E OBTENÇÃO DO VÍRUS

SARS-CoV-2 original (Wuhan – depositado no Gisaaid sob número: EPI\_ISL\_402124)

Search | Downloads | Upload

Search

EPI\_ISL ID: **EPI\_ISL\_402124** Virus name:  EPI\_SET ID:   Complete ?

Location:  Host:   High coverage ?

Collection:  to  Submission:  to   Low coverage excluded ?

Grade:  all  Lineage:  Variant:   With patient status ?

AA Substitutions:  Nucl Mutations:   Collection date complete ?

Under investigation

Next Search

Virus name	Passage date	Accession ID	Collection date	Submission date	Length	Host	Location	Originating
hCoV-19/Wuhan/WIV04/2019	Original	<a href="#">EPI_ISL_402124</a>	2019-12-30	2020-01-11	29 891	Human	Asia / China / Hu	Wuhan Ji

1 viruses

Charts EPI\_SET Select Analysis Download

**Virus detail**

**Virus name:** hCoV-19/Wuhan/WIV04/2019

**Accession ID:** EPI\_ISL\_402124

**Type:** betacoronavirus

**Clade:** L

**Pango Lineage:** B (Pango v.4.3.1 PANGO-v1.23)

**AA Substitutions:**

**Variant:**

**Sample information**

**Collection date:** 2019-12-30

**Location:** Asia / China / Hubei / Wuhan

**Host:** Human

**Additional location information:** Retailer at Huanan Seafood Wholesale Market

**Gender:** Female

**Patient age:** 49

**Patient status:** Live

**Specimen source:** Bronchoalveolar lavage fluid

**Additional host information:**

**Sampling strategy:**

**Outbreak:** 2019-12-27

**Last vaccinated:**

**Treatment:**

**Authors:** Peng Zhou, Xing-Lou Yang, Ding-Yu Zhang, Lei Zhang, Yan Zh

**Submitter information**

**Submitter:** Shi, Zhengli

**Submission Date:** 2020-01-11

**Address:** 44 Xiao Hong Shan, 430071 Wuhan, Hubei Province, Peoples

**FASTA**

```
>hCoV-19/Wuhan/WIV04/2019|EPI_ISL_402124|2019-12-30
ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAACCAACTTTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACTTTA/
AATCTGTGTGGCTGTCACCTCGGCTGCATGCTTAGTGCACTCACGCAGTATAATTAATAACTAATTACTGTCGTTGACAGC
ACACGAGTAACCTCGTCTATCTTCTGCAGGCTGCTTACGGTTTCGTCCGTGTTGCAGCCGATCATCAGCACATCTAGGTTT
CGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTCCCTGGTTTTCAACGAGAAAACACACGTCCAACCTCAGTTTGC
CTGTTTTACAGGTTTCGCGACGTGCTCGTACGTGGCTTTGGAGACTCCGTGGAGGAGGTCTTATCAGAGGCACGTCAACAT
CTTAAAGATGGCACTTGTGGCTTAGTAGAAGTTGAAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTTCATCA/
```

**TAMANHO DO GENOMA**  
29891 nt

# INFORMAÇÕES E OBTENÇÃO DO VÍRUS



Linhagem atual mais prevalente JN.1.1 (depositado no Gisaïd sob número: EPI\_ISL\_18492450).

Virus detail	
<b>Virus name:</b>	hCoV-19/Belgium/regA-47745/2023
<b>Accession ID:</b>	EPI_ISL_18492450
<b>Type:</b>	betacoronavirus
<b>Clade:</b>	GRA
<b>Pango Lineage:</b>	JN.1.1 (Pango v.4.3.1 consensus call), Omicron (BA.2-like) (Scorpio)
<b>AA Substitutions:</b>	Spike A27S, Spike A264D, Spike A570V, Spike D405N, Spike D614G, Spike D796Y, Spike E484K, Spike E554K, Spike F157S, Spike F486P, Spike G142D, Spike G339H, Spike G446S, Spike H69del, Spike H245N, Spike H655Y, Spike I332V, Spike ins16MPLF, Spike K356T, Spike K417N, Spike L24del, Spike L212I, Spike L216F, Spike L452W, Spike L455S, Spike N211del, Spike N440K, Spike N450D, Spike N460K, Spike N481K, Spike N501Y, Spike N679K, Spike N764K, Spike N969K, Spike P25del, Spike P26del, Spike P621S, Spike P681R, Spike P1143L, Spike Q498R, Spike Q954H, Spike R21T, Spike R158G, Spike R403K, Spike R408S, Spike S50L, Spike S371F, Spike S373P, Spike S375F, Spike S477N, Spike S939F, Spike T19I, Spike T376A, Spike T478K, Spike V70del, Spike V127F, Spike V213G, Spike V445H, Spike V483del, Spike Y144del, Spike Y505H, E T9I, M A63T, M A104V, M D3H, M Q19E, M T30A, N E31del, N G204R, N P13L, N Q229K, N R32del, N R203K, N S33del, N S413R, NS3 T223I, NS7a R25K, NS7b F19L, NSP1 S135R, NSP2 A31D, NSP2 F319L, NSP3 A1892T, NSP3 G489S, NSP3 K1155R, NSP3 N1708S, NSP3 T24I, NSP3 V238L, NSP4 L264F, NSP4 T327I, NSP4 T492I, NSP5 P132H, NSP6 F108del, NSP6 G107del, NSP6 R252K, NSP6 S106del, NSP6 V24F, NSP9 T35I, NSP12 P323L, NSP13 R392C, NSP14 I42V, NSP15 T112I
<b>Variant:</b>	VOI GRA (JN.1+JN.1.*) first detected in Luxembourg/Iceland

## MUTAÇÕES

Missense: 52 mutações

Indel: 9 mutações

## TAMANHO DO GENOMA

29649 nt

Virus name	Passage de	Accession ID	Collection da	Submission E		Length	Host	Location
hCoV-19/Belgium/regA-47745/2023	Original	EPI_ISL_18492450	2023-11-04	2023-11-10		29 649	Human	Europe / Belgium

# Obtenção e informações da Spike



- PROTEÍNA SPIKE

## TAMANHO

3822 nt

```
ATGTTTGTTTTCTTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCCTGCATA
CACTAATTCTTTCACACGTGGTGTGTTTATTACCCTGACAAAGTTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCT
TACCTTTCTTTTCCAATGTTACTTGGTTCATGCTATAACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTC
CTACCATTTAATGATGGTGTGTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGA
TTCGAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATC
CATTTTTGGGTGTTTATTACCACAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGC
```

```
ACACAGTTTATGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTTAAGAATCATAACATCACCG
ATGTTGATTTAGGTGACATCTCTGGCATTAAATGCTTCAGTTGTAAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTTGCC
AAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACTTGGAAAGTATGAGCAGTATATAAAATGGCCATGGTACATTTGGCT
AGGTTTTATAGCTGGCTTGATTGCCATAGTAATGGTGACAATTATGCTTTGCTGTATGACCAGTTGCTGTAGTTGTCTCAAGG
GCTGTTGTTCTTGTGGATCCTGCTGCAAATTTGATGAAGACGACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACATTACACA
TAA
```

# Obtenção e informações da Spike



- IDENTIFICAÇÃO DAS POSIÇÕES DAS DUAS MUTAÇÕES MISSENSE

Results for 3810 residue sequence "mRNA /gene="Spike atual" (exons in uppercase)" starting "ATGTTTGTTT"

```
901 R F N G I G V T Q N V L Y E N Q K L I A
2701 AGGTTTAATGGTATTGGAGTTACACAGAATGTTCTCTATGAGAACCAAAAATTGATTGCC
921 N Q F N S A I G K I Q D S L F S T A S A
2761 AACCAATTTAATAGTGCTATTGGCAAAAATTCAAGACTCACTTTTTTCCACAGCAAGTGCA
941 L G K L Q D V V N H N A Q A L N T L V K
2821 CTTGGAAAACCTTCAAGATGTGGTCAACCATAATGCACAAGCTTTAAACACGCTTGTTAAA
961 Q L S S K F G A I S S V L N D I L S R L
2881 CAACTTAGCTCCAAATTTGGTGCAATTTCAAGTGTTTTAAATGATATCCTTTCACGTCTT
981 D K V E A E V Q I D R L I T G R L Q S L
2941 GACAAAGTTGAGGCTGAAGTGCAAATTGATAGGTTGATCACAGGCAGACTTCAAAGTTTG
1001 Q T Y V T Q Q L I R A A E I R A S A N L
3001 CAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCTTCTGCTAATCTT
1021 A A T K M S E C V L G Q S K R V D F C G
3061 GCTGCTACTAAAATGTCAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTGTGGA
1041 K G Y H L M S F P Q S A P H G V V F L H
3121 AAGGGCTATCATCTTATGTCCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCAT
1061 V T Y V P A Q E K N F T T A P A I C H D
3181 GTGACTTATGTCCCTGCACAAGAAAAGAACTTCACAACCTGCTCCTGCCATTTGTCATGAT
1081 G K A H F P R E G V F V S N G T H W F V
```

## MUTAÇÕES MISSENSE

K986P

V987P

# Obtenção e informações da Spike

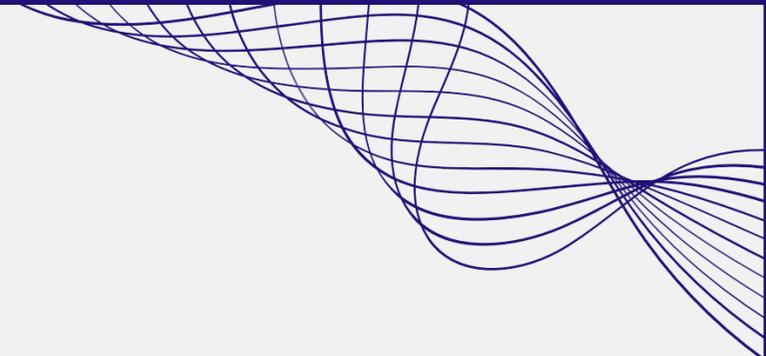


```
961 T L V K Q L S S N F G A I S
2881 ACGCTTGTTAAACAACCTTAGCTCCAATTTTGGTGCAATTTC
981 L S R L D P P E A E V Q I D
2941 CTTTCACGTCTTGACCCUCCGGAGGCTGAAGTGCAAATTGA
1001 L Q S L Q T Y V T Q Q L I R
3001 CTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATTAG
1021 S A N L A A T K M S E C V F
```

**SUBSTITUIÇÃO POR PROLINA**



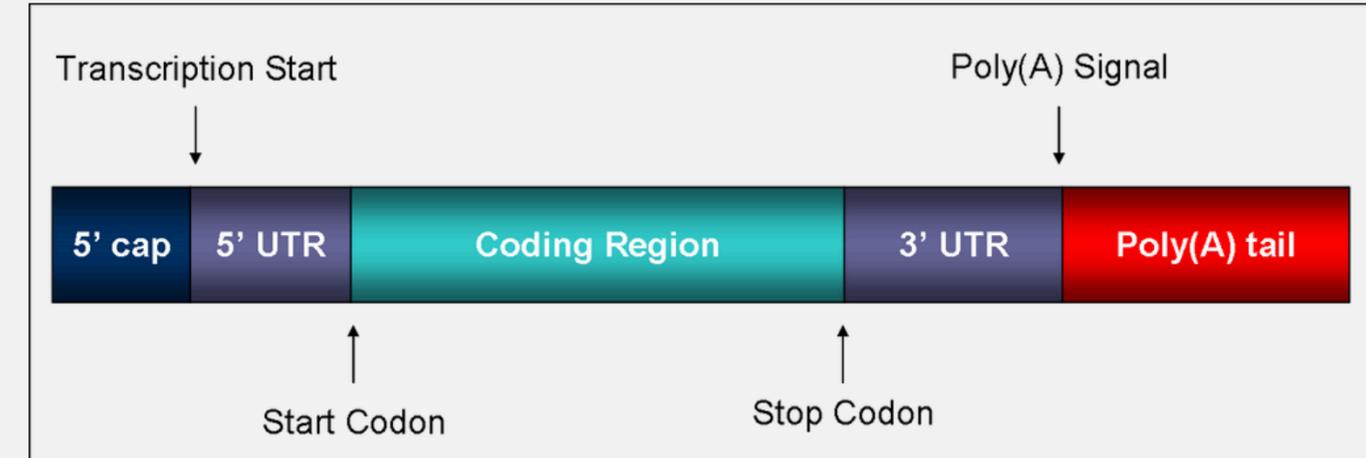
**AO ADOTAR ESTA ESTRATÉGIA, DOIS AMINOÁCIDOS NO TOPO DA POSIÇÃO HELICOIDAL DO CENTRO DA SUBUNIDADE S2 SÃO SUBSTITUÍDOS POR PROLINAS (K986P E V987P), O QUE DEMONSTROU MELHORAR EFICAZMENTE A ESTABILIDADE DA PROTEÍNA**



# Regiões UTR e passo a passo para inserção

# Regiões UTR

- As UTR (UnTranslated Region) são cruciais na regulação da expressão gênica;
- 5'UTR atua na regulação da estabilidade do mRNA e eficiência translacional;
- 3' UTR regula a estabilidade do mRNA por meio de ligação de proteínas reguladoras e microRNAs;
- Em vacinas de mRNA as regiões UTR são projetadas para que se alcance uma imunogenicidade e eficiência da vacina.



# Passo a Passo

- No NCBI digitar o código do gene ***HBB: NM\_000518.5***

The screenshot shows the NCBI Gene database interface. At the top, there is a navigation bar with the NIH logo and the text "National Library of Medicine National Center for Biotechnology Information". A search bar is located below the navigation bar, with a dropdown menu set to "Gene" and a "Search" button. The main content area features a large image of chromosomes on the left and a text box on the right that reads: "Gene integrates information from a wide range of species. A record may include nomenclature, Reference Sequences (RefSeqs), maps, pathways, variations, phenotypes, and links to genome-, phenotype-, and locus-specific resources worldwide." Below this, there are three columns of links: "Using Gene" (Gene Quick Start, FAQ, Download/FTP, RefSeq Mailing List, Gene News, Factsheet), "Gene Tools" (Submit GeneRIFs, Submit Correction, Statistics, BLAST, Splign), and "Other Resources" (OMIM, RefSeq, RefSeqGene, Protein Clusters).

# Passo a Passo

- Encontrar o tamanho do mRNA

## Homo sapiens hemoglobin subunit beta (HBB), mRNA

NCBI Reference Sequence: NM\_000518.5

[FASTA](#) [Graphics](#)

Go to:

LOCUS	NM_000518	628 bp	mRNA	linear	PRI 13-MAY-2024
DEFINITION	Homo sapiens hemoglobin subunit beta (HBB), mRNA.				
ACCESSION	NM_000518				
VERSION	NM_000518.5				
KEYWORDS	RefSeq; MANE Select.				
SOURCE	Homo sapiens (human)				
ORGANISM	<a href="#">Homo sapiens</a> Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 628)				
AUTHORS	Waye,J.S., Hanna,M., Hohenadel,B.A., Nakamura,L., Walker,L., Eng,B. and Nfonam,L.E.				
TITLE	beta0-Thalassemia Caused by a Novel Nonsense Mutation [HBB:c.199A > T]				
JOURNAL	Hemoglobin 48 (1), 69-70 (2024)				
PUBMED	<a href="#">38425097</a>				
REMARK	GeneRIF: beta[0]-Thalassemia Caused by a Novel Nonsense Mutation [HBB:c.199A > T].				

# Passo a Passo

- Identificar a região codificadora para que encontrar as regiões 5'UTR

CDS

```
51..494
/gene="HBB"
/gene_synonym="beta-globin; CD113t-C; ECT6"
/note="beta globin chain; hemoglobin, beta; hemoglobin
beta subunit"
/codon_start=1
/product="hemoglobin subunit beta"
/protein_id="NP_000509.1"
/db_xref="CCDS:CCDS7753.1"
/db_xref="GeneID:3043"
/db_xref="HGNC:HGNC:4827"
/db_xref="MIM:141900"
/translation="MVHLTPEEKSAVTALWGKVNVDVEVGGGEALGRLLVYPWTQRRFE
SFGDLSTPDAMGNPKVKAHGKKVLGAFSDGLAHLNLIKGTFFATLSELHCDKLHVDPE
NFRLLGNVLVCVLAHFGKEFTPPVQAAYQKVVAGVANALAHKYH"
```

# Passo a Passo

- Fazer o download da sequência do mRNA como arquivo FASTA

## Homo sapiens hemoglobin subunit beta (HBB), mRNA

NCBI Reference Sequence: NM\_000518.5

[FASTA](#) [Graphics](#)

Go to:

LOCUS NM\_000518 628 bp mRNA linear PRI 13-MAY-2024  
DEFINITION Homo sapiens hemoglobin subunit beta (HBB), mRNA.  
ACCESSION NM\_000518  
VERSION NM\_000518.5  
KEYWORDS RefSeq; MANE Select.  
SOURCE Homo sapiens (human)  
ORGANISM [Homo sapiens](#)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 628)  
AUTHORS Waje,J.S., Hanna,M., Hohenadel,B.A., Nakamura,L., Walker,L., Eng,B.  
and Nfonsam,L.E.  
TITLE beta<sup>0</sup>-Thalassemia Caused by a Novel Nonsense Mutation [HBB:c.199A >  
T]  
JOURNAL Hemoglobin 48 (1), 69-70 (2024)  
PUBMED [38425097](#)  
REMARK GeneRIF: beta[<sup>0</sup>]-Thalassemia Caused by a Novel Nonsense Mutation  
[HBB:c.199A > T].

# Passo a Passo

- Após realizar o download, colocar a sequência no Word e identificar a região 5'UTR e 3'UTR

```
ACATTTGCTTCTGACACAACCTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCTGACTCCTGA  
GGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGC  
AGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATG  
CTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGC  
TCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCCTGTGACAAGCTGCACGTGGAT  
CCTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCA  
CCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAGTATCA  
CTAAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCTTTGTTCCCTAAGTCCAACACTAAACT  
GGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCAA
```

# Passo a Passo



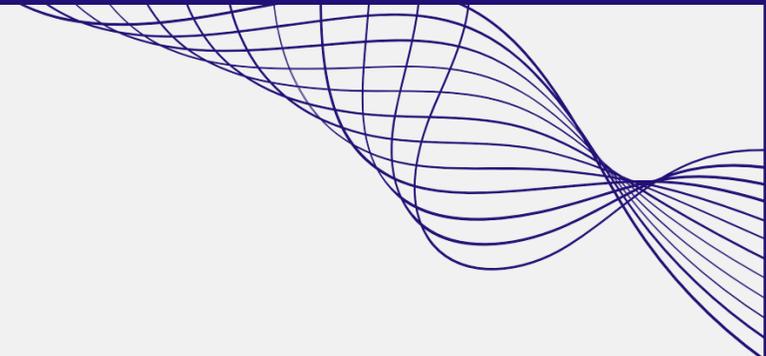
- Com as sequências das regiões UTR, inseri-las na sequência FASTA do cDNA da proteína spike original com as mutações.

## 5'UTR

```
>EPI_ISL_402124-S
ACATTTGCTTCTGACACAACCTGTGTTCACTAGCAACCTCAAACAGACACCATGTTTGTITTTCTTGTTTTATTGC
CACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCTGCATACACTAATTCTTTCACAC
GTGGTGTITATTACCCTGACAAAGTTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTCTTACCTTCT
TTTCCAATGTTACTTGGTTCATGCTATACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCC
TACCATTTAATGATGGTGTITATTTTGCTTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTA
CTTTAGATTGCAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTC
AATTTTGTAATGATCCATTTTTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAAGTGAGTTCAGAG
TTTATTCTAGTGCGAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGG
GTAATTTCAAAAATCTTAGGGAATTTGTGTTTAAAGATATTGATGGTTATTTTAAAATATATTCTAAGCACACGC
CTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTA
ACATCACTAGGTTTTCAAACCTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGA
CAGCTGGTGCTGCAGCTTATTATGTGGGTTATCTTCAACCTAGGACTTTTCTATTAAAATATAATGAAAATGGAA
CCATTACAGATGCTGTAGACTGTGCACCTTGACCCTCTCTCAGAAAACAAAGTGACGTGAAATCCTTCACTGTAG
AAAAAGGAATCTATCAAACCTTCTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCTTAATATTACAA
```

## 3'UTR

```
CTGCCATTTGTCATGATGGAAAAGCACACTTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTG
TAACACAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACACATTTGTGTCTGGTAACTGTGATGTTG
TAATAGGAATTGTCAACAACACAGTTTTATGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATA
AATATTTAAGAATCATACTACCAGATGTTGATTTAGGTGACATCTCTGGCATTAAATGCTTCAGTTGTAACA
TTCAAAAAGAAATGACCGCTCAATGAGGTTGCCAAGAATTTAAATGAATCTCTCATCGATCTCCAAGAATTG
GAAAGTATGAGCAGTATATAAAATGGCCATGGTACATTTGGCTAGGTTTTATAGCTGGCTTGATTGCCATAGTAA
TGGTGACAATTATGCTTTGCTGTATGACCAGTGTCTGTAGTTGTCTCAAGGGCTGTGTTCTTGTGGATCCTGCT
GCAAATTTGATGAAGACGACTCTGAGCCAGTGTCAAAGGAGTCAAATTACATTACACATAACTAACTCGCTTT
CTTGCTGTCCAATTTCTATTAAAGGTTCCCTTTGTTCCCTAAGTCCAACCTACTAACTGGGGATATTATGAAGGG
CCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCAA
```



# **Estrutura Secundária das regiões 5'UTR, 3'UTR e vacina de mRNA (pré-otimização)**

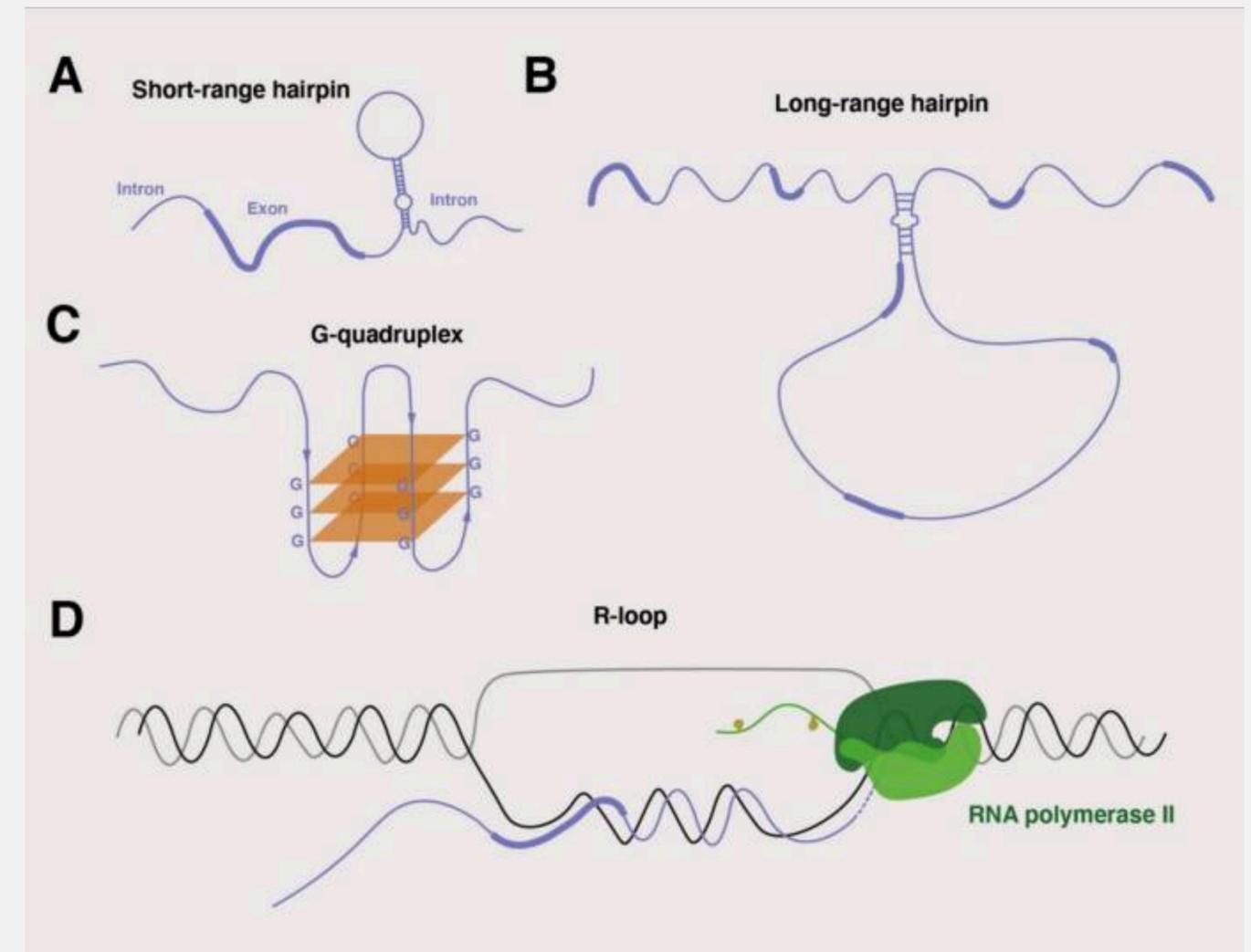
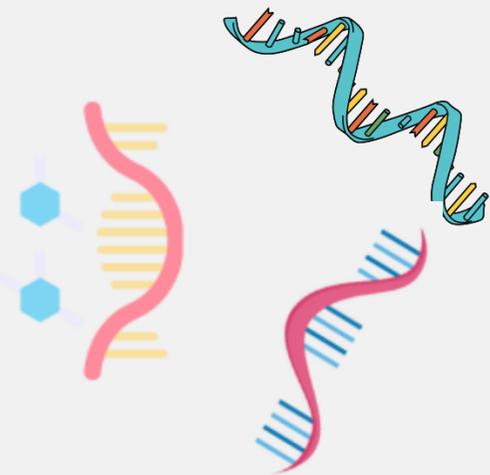
# Estrutura secundária do mRNA

Função do RNA - **predita pela sequência**

Estrutura secundária - **define propriedades importantes**

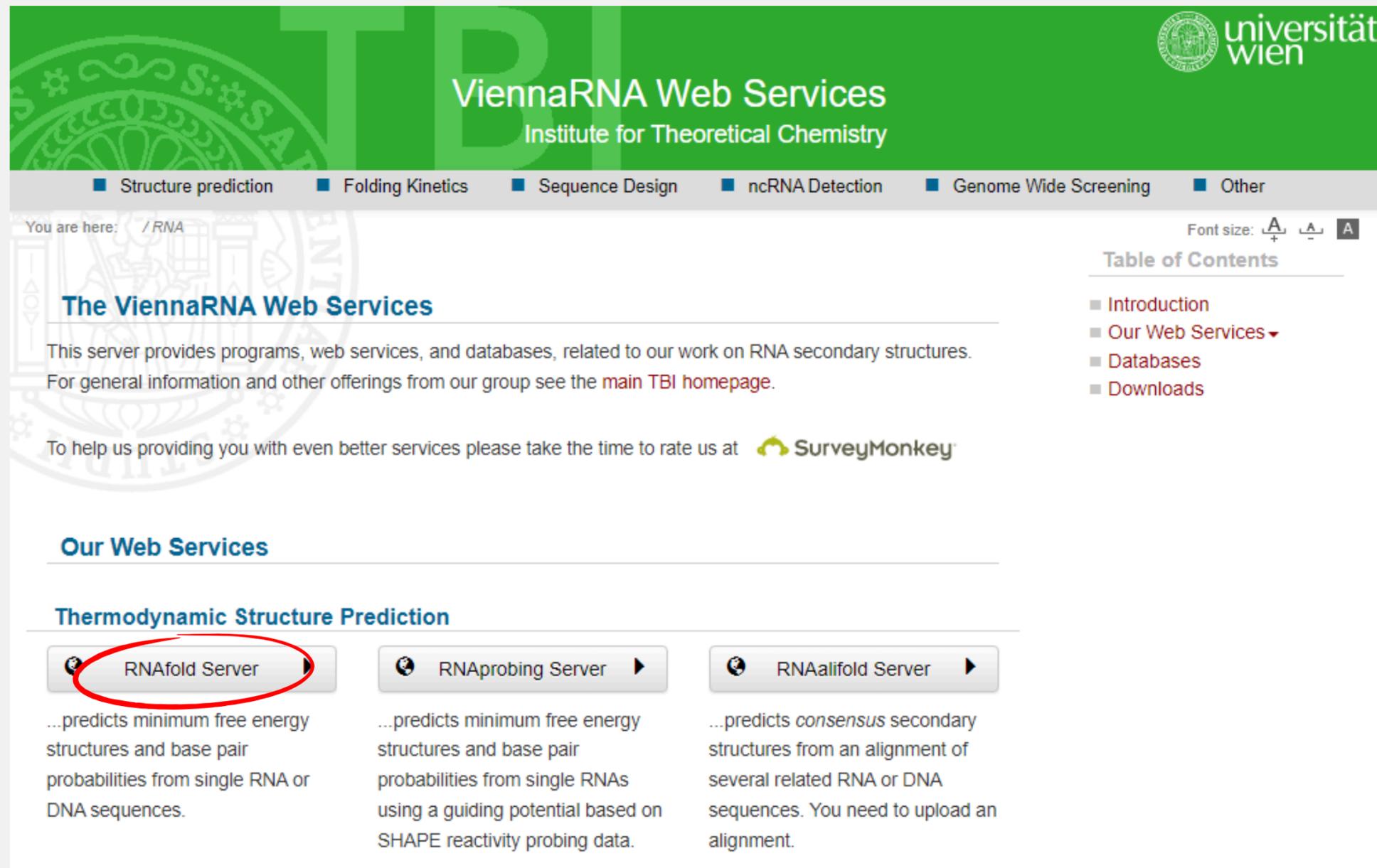
Reshetnikov, et al., 2024

- Splicing Alternativo
- Estabilidade e decaimento
- Iniciação da tradução



# Estrutura secundária do mRNA - ViennaRNA

Software para predição da **estrutura secundária** de RNAs



ViennaRNA Web Services  
Institute for Theoretical Chemistry

■ Structure prediction ■ Folding Kinetics ■ Sequence Design ■ ncRNA Detection ■ Genome Wide Screening ■ Other

You are here: / RNA

Font size:  A

**The ViennaRNA Web Services**

This server provides programs, web services, and databases, related to our work on RNA secondary structures. For general information and other offerings from our group see the [main TBI homepage](#).

To help us providing you with even better services please take the time to rate us at  SurveyMonkey

**Our Web Services**

**Thermodynamic Structure Prediction**

- RNAfold Server**  
...predicts minimum free energy structures and base pair probabilities from single RNA or DNA sequences.
- RNAprobing Server**  
...predicts minimum free energy structures and base pair probabilities from single RNAs using a guiding potential based on SHAPE reactivity probing data.
- RNAalifold Server**  
...predicts *consensus* secondary structures from an alignment of several related RNA or DNA sequences. You need to upload an alignment.

- Função RNAfold Server

# Estrutura secundária do mRNA - ViennaRNA

Software para predição da **estrutura secundária** de RNAs

**RNAfold** WebServer 1 Enter Input Parameters 2 View Results

[Home|New Job|Help]

The **RNAfold web server** will predict secondary structures of single stranded RNA or DNA sequences. Current limits are 7,500 nt for partition function calculations and 10,000 nt for minimum free energy only predictions.

Simply paste or upload your sequence below and click Proceed. To get more information on the meaning of the options click the ? symbols. You can test the server using [this sample sequence](#).

**Sequence Input**

Paste or type your **sequence** here: [clear]

Or upload a file in FASTA format:  Nenhum arquivo escolhido

**Folding Constraints**

**Fold algorithms and basic options**

- minimum free energy (MFE) and partition function ?
- minimum free energy (MFE) only ?
- no GU pairs at the end of helices ?
- avoid isolated base pairs ?
- assume RNA molecule to be circular ?
- Incorporate G-Quadruplex formation into the structure prediction algorithm ?

**Dangling end options** ?

**Energy Parameters** ?

**Modified Bases**

**SHAPE reactivity data** ?

**Output options**

- interactive RNA secondary structure plot ?
- RNA secondary structure plots with reliability annotation (Partition function folding only) ?
- Mountain plot ?

Notification via e-mail upon completion of the job (optional):

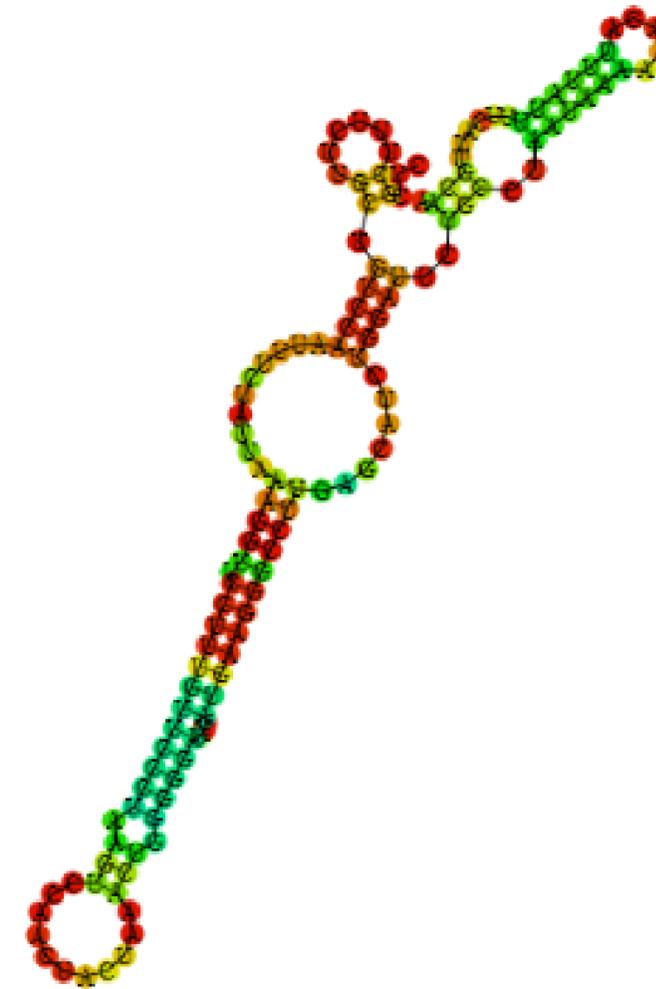
# Estrutura secundária do mRNA - Regiões 5'/3' UTR

Estrutura 5' UTR



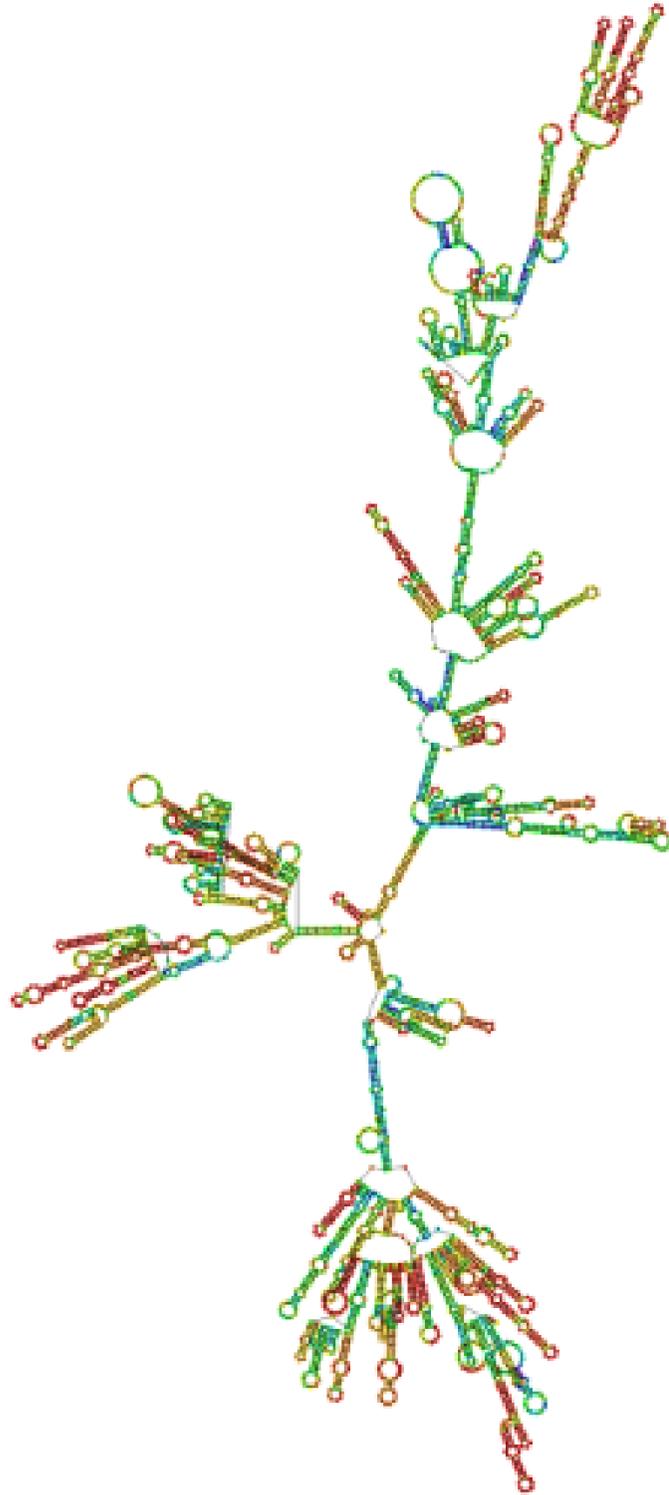
MFE: -10,10 kcal/mol

Estrutura 3' UTR



MFE: -24,00 kcal/mol

# Estrutura secundária do mRNA - vacina mRNA

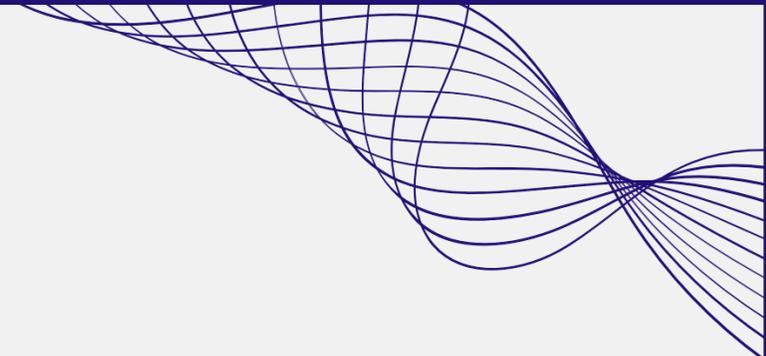


Estrutura predita - antes da otimização dos códons

MFE (Minimum Free Energy)

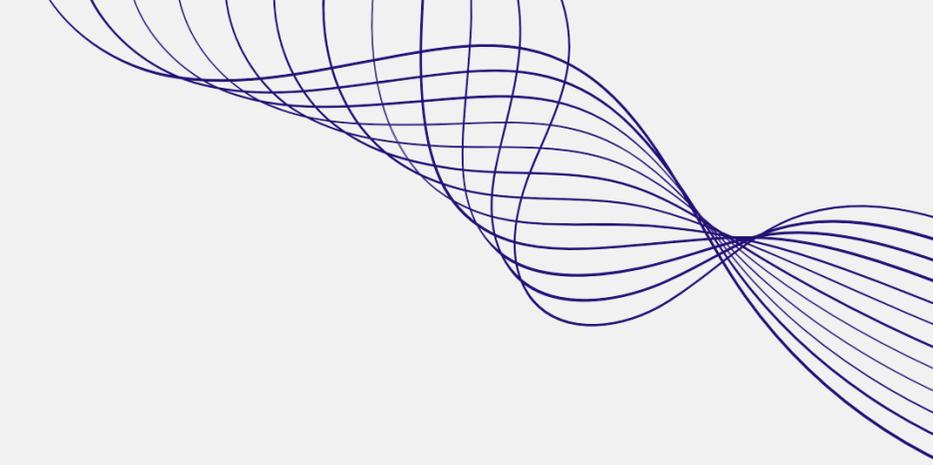
- Indicativo da estabilidade da molécula de RNA
- Tamanho, composição e arranjo

**MFE: -1118,05 kcal/mol**



# Otimização e predição de estrutura secundária

# OTIMIZAÇÃO DE CÓDONS



Diferentes espécies: **Viés no uso de códons.**

“Preferência” de códons para um mesmo aminoácido.

Afeta a estabilidade, processamento e qualidade de tradução do mRNA.



**Otimização:** substituição de códons do mRNA viral por códons preferencialmente utilizados em mRNA humano, SEM ALTERAR SEQUÊNCIA PROTEICA.

# OTIMIZAÇÃO DE CÓDONS: PLATAFORMA IDT

*Integrated DNA technologies:* Plataforma online de otimização de códons.



# OTIMIZAÇÃO DE CÓDONS: PLATAFORMA IDT



Sequence Type:  DNA Bases  Amino Acids

RESTRICTION SITE INFORMATION ⓘ

Product Type: gBlocks® Gene Fragments ▼

Organism: Homo sapiens (human) ▼

Single Entry

Bulk Entry

```
ATGTTTGTTCCTTGTTCCTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCCTGCATACACTAA
TTCTTTCACACGTGGTGTTCCTTATTACCCTGACAAAGTTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCTTACCTTTCTTTT
CCAATGTTACTTGGTTCATGCTATACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGT
GTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTAT
TGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTGGGTGTTTATTACCACAAAAACA
ACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGAC
CTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTTATTTTAAAATATATTCTAAGCACAC
GCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGT
TTCAAACCTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCTTATTATGTG
GGTTATCTTCAACCTAGGACTTTTCTATTAATAATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTC
AGAAACAAAGTGACGTTGAAATCCTTCACTGTAGAAAAGGAATCTATCAAACCTTCTAACTTTAGAGTCCAACCAACAGAATCTATTG
```

# of bases: 3822

HIGHLIGHT CODONS

# OTIMIZAÇÃO DE CÓDONS: PLATAFORMA IDT



Sequence Type:  DNA Bases  Amino Acids

RESTRICTION SITE INFORMATION ⓘ

Product Type: gBlocks® Gene Fragments ▼

Organism: Homo sapiens (human) ▼

Single Entry

Bulk Entry

```
ATGTTTGTTCCTTGTTCCTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCCTGCATACACTAA
TTCTTTCACACGTGGTGTTCCTTATTACCCTGACAAAGTTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCTTACCTTTCTTTT
CCAATGTTACTTGGTTCATGCTATACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGT
GTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTAT
TGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTGGGTGTTTATTACCACAAAAACA
ACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGAC
CTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTTATTTTAAAATATATTCTAAGCACAC
GCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGT
TTCAAACCTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCTTATTATGTG
GGTTATCTTCAACCTAGGACTTTTCTATTAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTC
AGAAACAAAGTGACGTTGAAATCCTTCACTGTAGAAAAGGAATCTATCAAACCTTCTAACTTTAGAGTCCAACCAACAGAATCTATTG
```

# of bases: 3822

HIGHLIGHT CODONS

# OTIMIZAÇÃO DE CÓDONS: PLATAFORMA IDT



Sequence Type:  DNA Bases  Amino Acids

RESTRICTION SITE INFORMATION ⓘ

Product Type: gBlocks® Gene Fragments ▼

Organism: Homo sapiens (human) ▼

Single Entry

Bulk Entry

```
ATGTTTGTTCCTTGTTCCTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCCTGCATACACTAA
TTCTTTCACACGTGGTGTTCCTTATTACCCTGACAAAGTTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCTTACCTTTCTTTT
CCAATGTTACTTGGTTCATGCTATACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGT
GTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTAT
TGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTGGGTGTTTATTACCACAAAAACA
ACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGAC
CTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTTATTTTAAAATATATTCTAAGCACAC
GCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGT
TTCAAACCTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCTTATTATGTG
GGTTATCTTCAACCTAGGACTTTTCTATTAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTC
AGAAACAAAGTGACGTTGAAATCCTTCACTGTAGAAAAGGAATCTATCAAACCTTCTAACTTTAGAGTCCAACCAACAGAATCTATTG
```

# of bases: 3822

HIGHLIGHT CODONS

# OTIMIZAÇÃO DE CÓDONS: PLATAFORMA IDT



Sequence Type:  DNA Bases  Amino Acids

RESTRICTION SITE INFORMATION ⓘ

Product Type: gBlocks® Gene Fragments ▼

Organism: Homo sapiens (human) ▼

Single Entry

Bulk Entry

```
ATGTTTGTTCCTTGTTCCTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCCTGCATACACTAA
TTCTTTCACACGTGGTGTTCCTTATTACCCTGACAAAGTTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCCTTACCTTTCTTTT
CCAATGTTACTTGGTTCATGCTATACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGT
GTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTAT
TGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTGGGTGTTTATTACCACAAAAACA
ACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGAC
CTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTTATTTTAAAATATATTCTAAGCACAC
GCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGT
TTCAAACCTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCTTATTATGTG
GGTTATCTTCAACCTAGGACTTTTCTATTAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTC
AGAAACAAAGTGACGTTGAAATCCTTCACTGTAGAAAAAGGAATCTATCAAACCTTCTAACTTTAGAGTCCAACCAACAGAATCTATTG
```

# of bases: 3822

HIGHLIGHT CODONS

# OTIMIZAÇÃO DE CÓDONS: PLATAFORMA IDT



Sequence Type:  DNA Bases  Amino Acids

RESTRICTION SITE INFORMATION ⓘ

Product Type: gBlocks® Gene Fragments ▼

Organism: Homo sapiens (human) ▼

Single Entry

Bulk Entry

```
ATGTTTGTTCCTTGTTCCTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCCTGCATACACTAA
TTCTTTCACACGTGGTGTTCCTTATTACCCTGACAAAGTTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCTTACCTTTCTTTT
CCAATGTTACTTGGTTCATGCTATACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGT
GTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTAT
TGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTGGGTGTTTATTACCACAAAAACA
ACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGAC
CTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTTATTTTAAAATATATTCTAAGCACAC
GCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGT
TTCAAACCTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCTTATTATGTG
GGTTATCTTCAACCTAGGACTTTTCTATTAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTC
AGAAACAAAGTGACGTTGAAATCCTTCACTGTAGAAAAGGAATCTATCAAACCTTCTAACTTTAGAGTCCAACCAACAGAATCTATTG
```

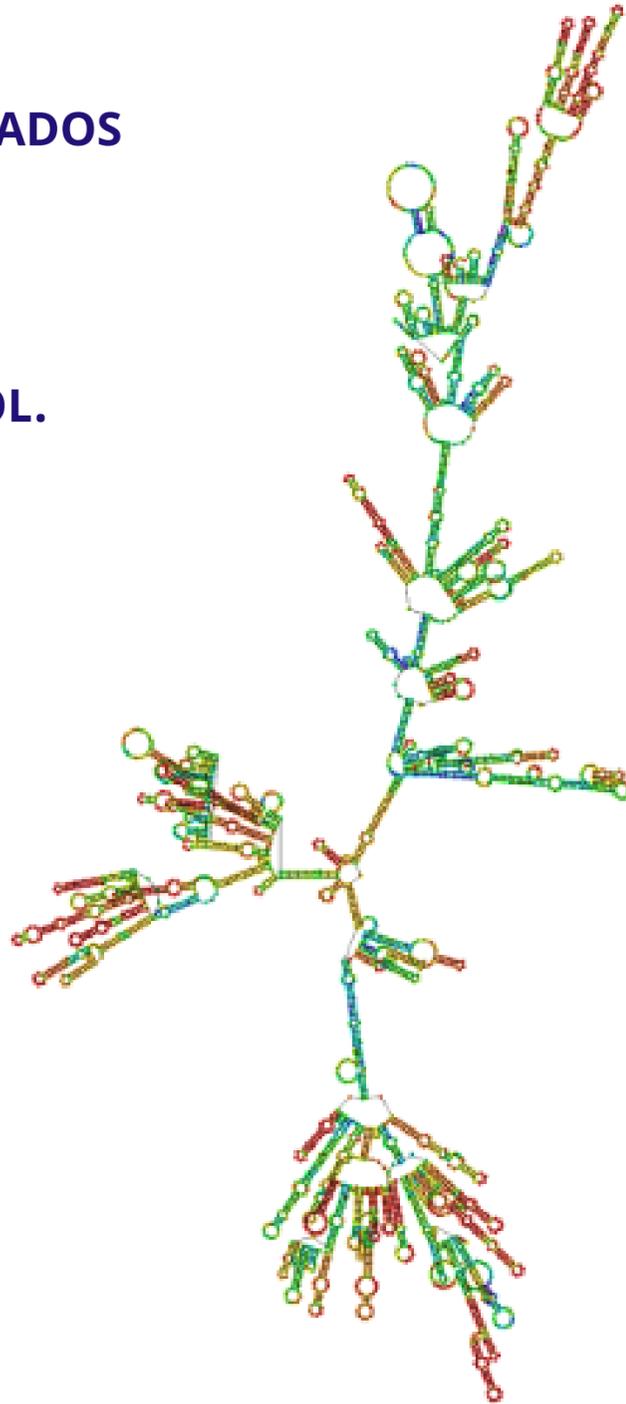
HIGHLIGHT CODONS

# of bases: 3822

# ESTRUTURA SECUNDÁRIA

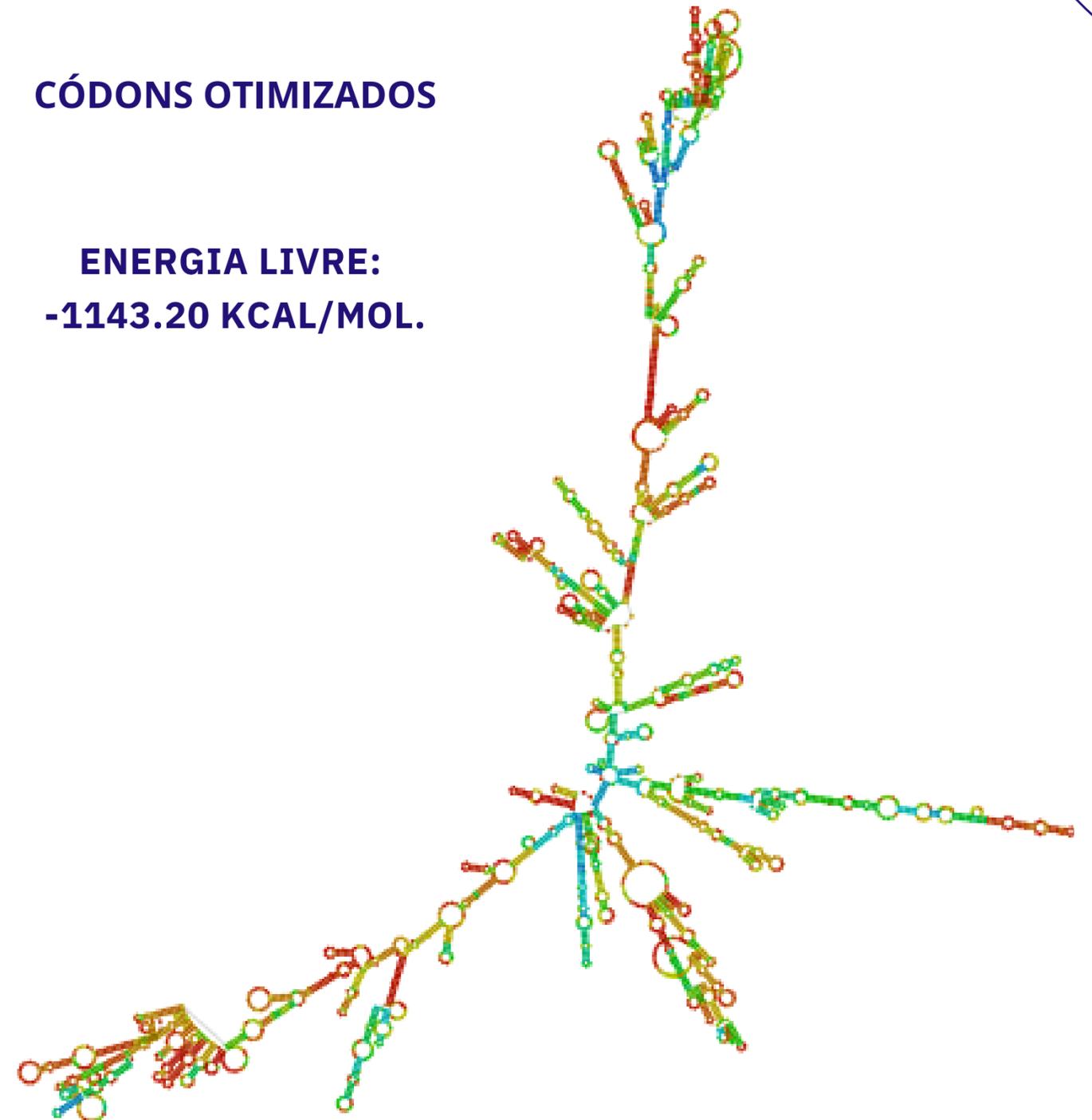
CÓDONS NÃO-OTIMIZADOS

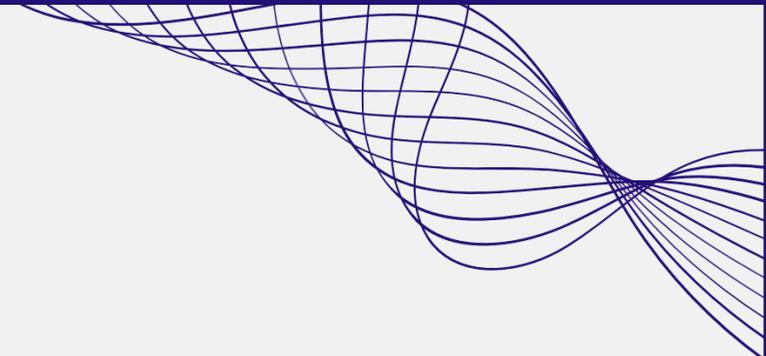
ENERGIA LIVRE:  
-1188.05 KCAL/MOL.



CÓDONS OTIMIZADOS

ENERGIA LIVRE:  
-1143.20 KCAL/MOL.

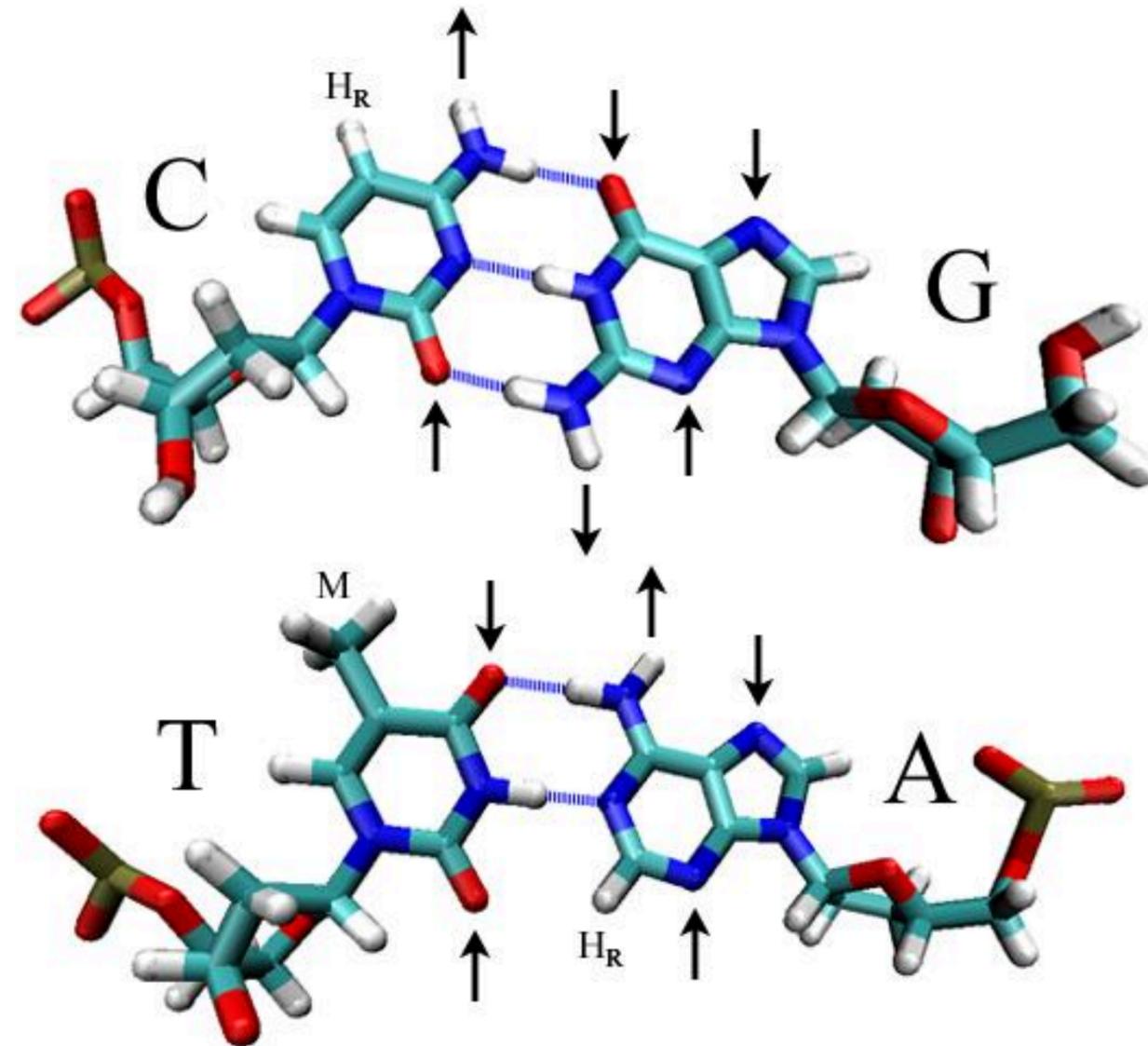




# Enriquecimiento CG

# ENRIQUECIMENTO CG

As sequências CG estão relacionadas a uma maior estabilidade da fita de mRNA;



# ENRIQUECIMENTO CG

As sequênc  
da fita de

## Human SARS-CoV-2 has evolved to reduce CG dinucleotide in its open reading frames

[Yong Wang](#) , [Jun-Ming Mao](#), [Guang-Dong Wang](#), [Zhi-Peng Luo](#), [Liu Yang](#), [Qin Yao](#) & [Ke-Ping Chen](#)

[Scientific Reports](#) **10**, Article number: 12331 (2020) | [Cite this article](#)

## On the Normalization of the Minimum Free Energy of RNAs by Sequence Length

[Edoardo Trotta](#) \*

## GC content shapes mRNA storage and decay in human cells

[Maïté Courel](#),<sup>1</sup> [Yves Clément](#),<sup>2</sup> [Clémentine Bossevain](#),<sup>1</sup> [Dominika Foretek](#),<sup>3</sup> [Olivia Vidal Cruchez](#),<sup>4</sup> [Zhou Yi](#),<sup>5</sup> [Marianne Bénard](#),<sup>1</sup> [Marie-Noëlle Benassy](#),<sup>1</sup> [Michel Kress](#),<sup>1</sup> [Caroline Vindry](#),<sup>6</sup> [Michèle Ernoult-Lange](#),<sup>1</sup> [Christophe Antoniewski](#),<sup>7</sup> [Antonin Morillon](#),<sup>3</sup> [Patrick Brest](#),<sup>4</sup> [Arnaud Hubstenberger](#),<sup>5</sup> [Hugues Roest Crollius](#),<sup>2</sup> [Nancy Standart](#),<sup>6</sup> and [Dominique Weil](#)<sup>1</sup>

Karsten Weis, Reviewing Editor and James L Manley, Senior Editor

Karsten Weis, ETH Zurich, Switzerland;

de

# ENRIQUECIMENTO CG

As sequências CG estão relacionadas a maior estabilidade da fita de mRNA.

**Códon não-otimizado**

166 sequências CG



**Códon otimizado**

227 sequências CG

# Predição de epítomos de anticorpos

# EPÍTOPO

Qualquer substância que pode ser especificamente ligada por uma molécula de anticorpo ou receptor de célula T.



**Fundamental para o estudo  
de uma vacina de mRNA**

### Translate tool

**Translate** is a tool which allows the translation of a nucleotide (DNA/RNA) sequence to a protein sequence.

#### DNA or RNA sequence

```
ATGTTTGTTTTCTTGTTTTATTGCCACTAGTCT
CTAGTCAGTGTGTTAATCTTACAACCAGAACTC
AATTACCCCTGCATACACTAATTCTTTCACAC
GTGGTGTTTATTACCCTGACAAAGTTTTTCAGAT
CCTCAGTTTTACATTCAACTCAGGACTTGTTCT
TACCTTTCTTTTCCAATGTTACTTGGTTCCATGC
TATACATGTCTCTGGGACCAATGGTACTAAGAG
GTTTGATAACCCTGTCCTACCATTTAATGATGGT
GTTTATTTTGCTTCCACTGAGAAGTCTAACATAA
TAAGAGGCTGGATTTTTTGGTACTACTTTAGATT
```

**Genetic codes** - [See NCBI's genetic codes](#)

Standard

reset

**TRANSLATE!**

#### Output format

- Verbose: Met, Stop, spaces between residues
- Compact: M, -, no spaces
- Includes nucleotide sequence
- Includes nucleotide sequence, no spaces

#### DNA strands

- forward
- reverse

# DOMÍNIO RBD (319-537)

## Spike Original

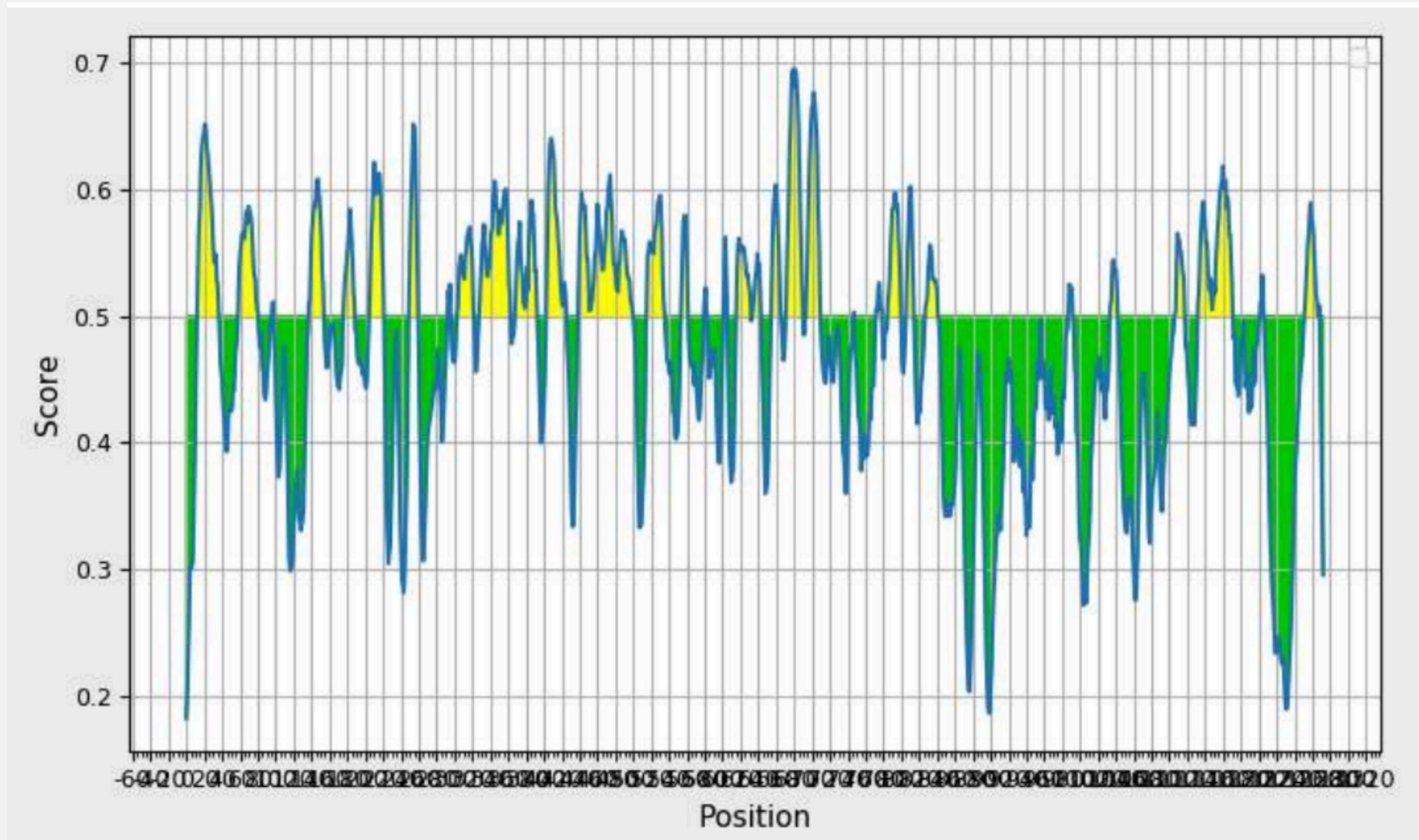
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNG  
TKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNK  
SWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEP  
LVDLPIGINITRFQTLALHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK  
CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASF  
TFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN  
YLYRLFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC  
GPKKSTNLVKNKCVNFNFENGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLILDITPCSFGGVSVITP  
GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQ  
TQTNPRRARSVASQSI IAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTEC  
NLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK  
VTLADAGFIKQYGDCLGDIAARDLICAQKFENGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAM  
QMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISS  
VLNDILSRLD **PP**EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM  
SFPQSAPHGVVFLHVTVPAQEKNFTTAPAICHGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGN  
CDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL  
QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT-

# MUTAÇÕES K986P E V987P -> TROCA DE UMA LISINA E UMA VALINA POR DUAS PROLINAS

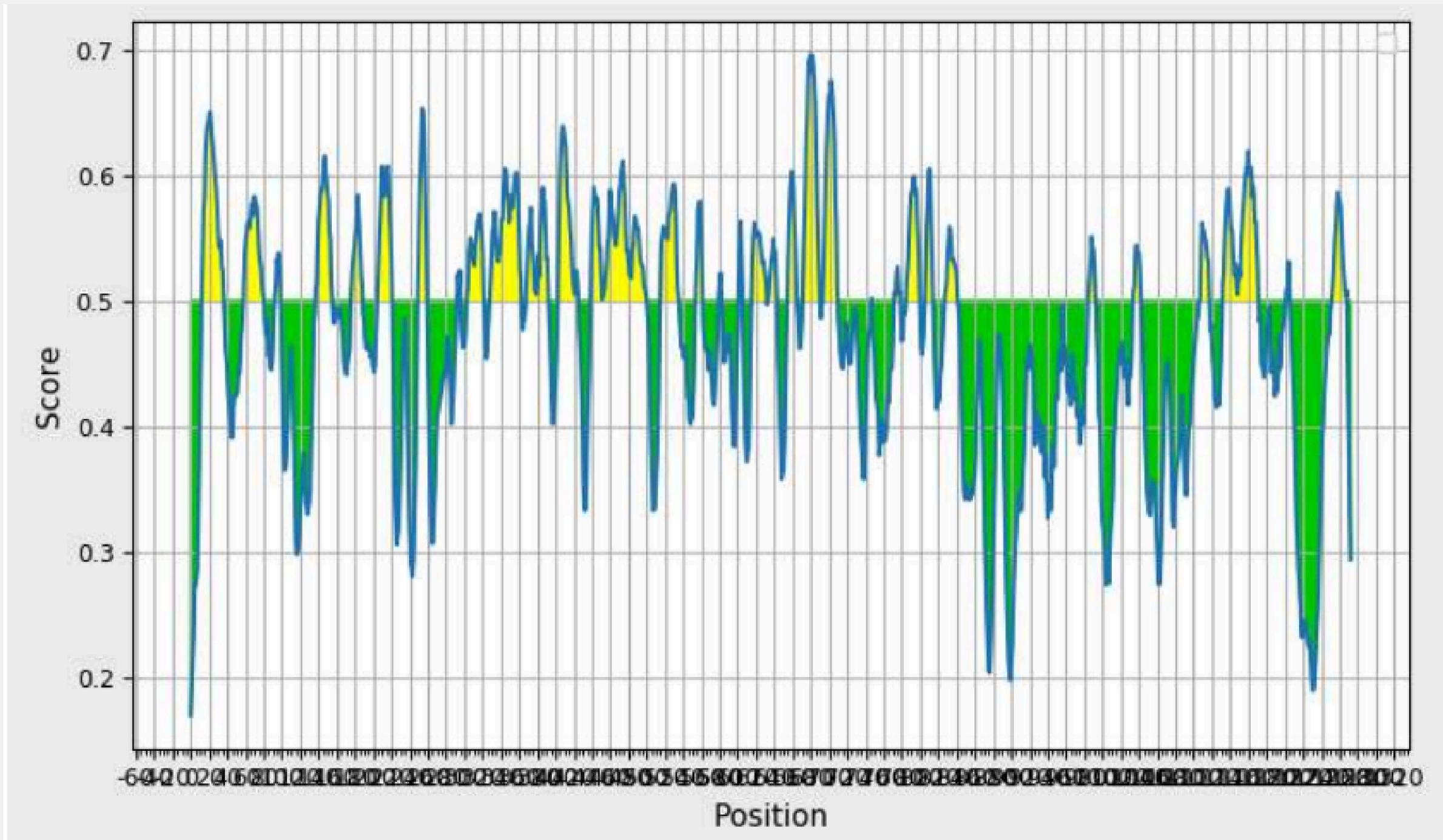
## Spike Original

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNG  
TKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNK  
SWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEP  
LVDLPIGINITRFQTLALHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK  
CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASF  
TFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG  
GNYNYLYRFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC  
GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP  
GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQ  
TQTNSPRRARSVASQSI IAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTEC  
S  
NLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK  
VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAM  
QMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISS  
VLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM  
SFPQSAPHGVVFLHVTVPAQEKNFTTAPAICHGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGN  
CDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL  
QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT-

## ANTES DA MUTAÇÃO MISSENSE 2P



## DEPOIS DA MUTAÇÃO MISSENSE 2P



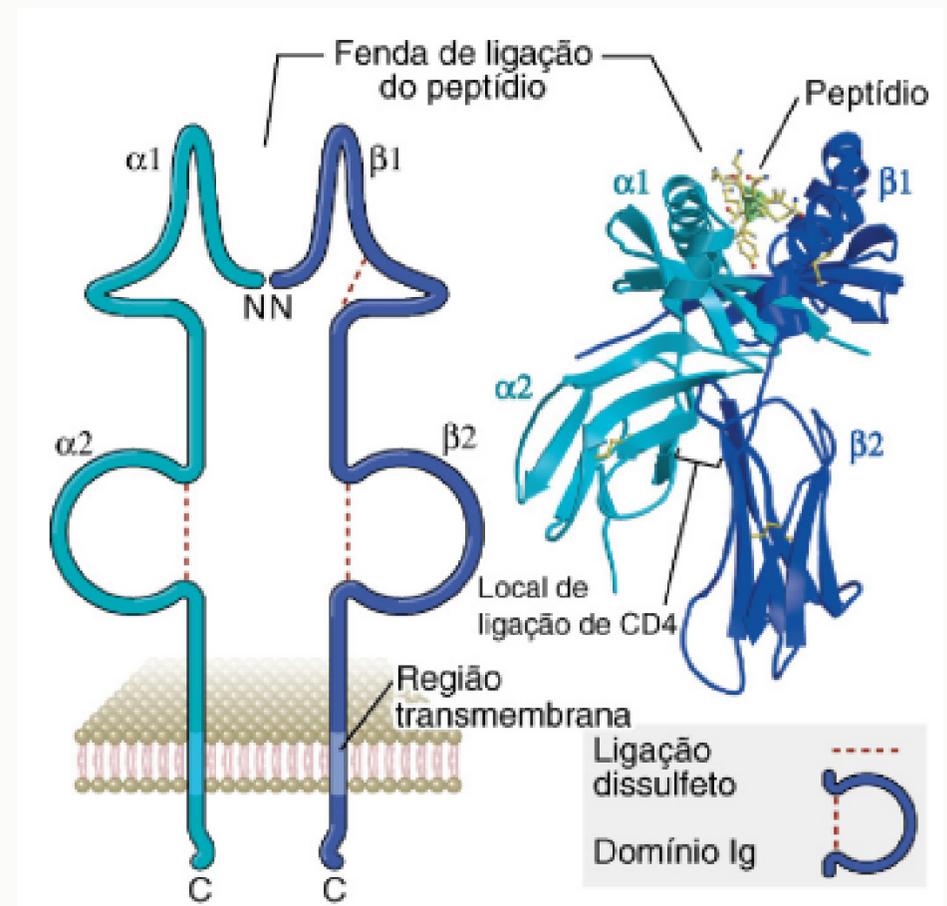
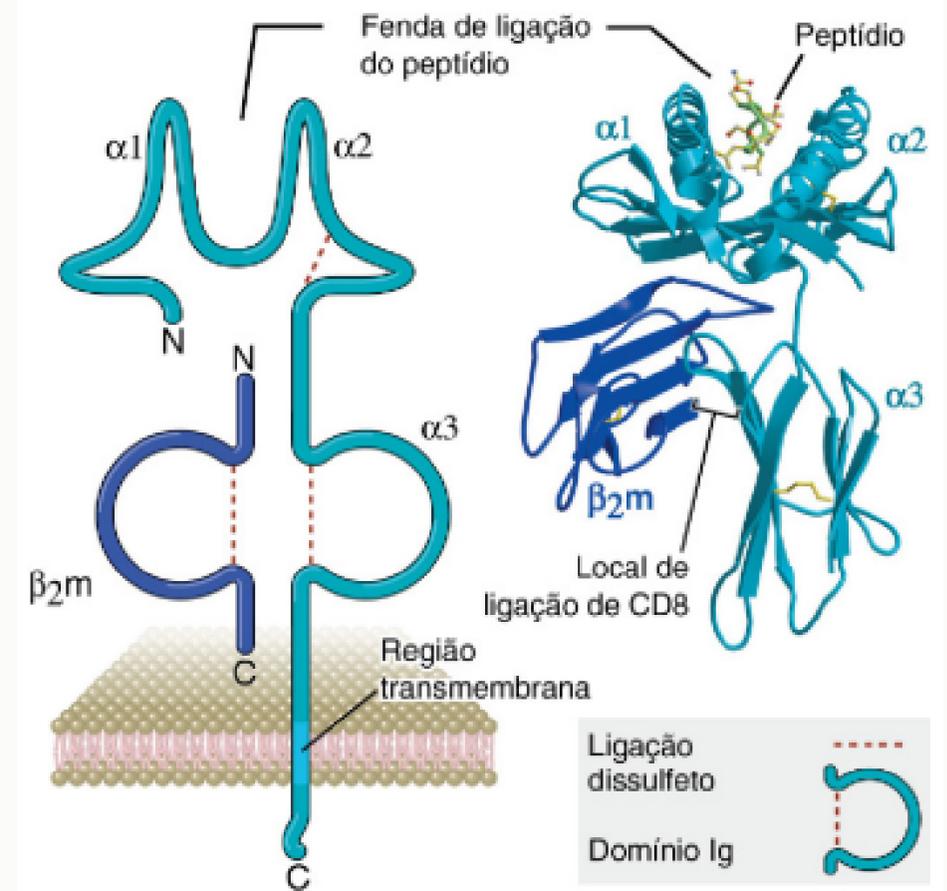
## ANTES DA MUTAÇÃO MISSENSE 2P

980	L	0.407	.
981	S	0.440	.
982	R	0.435	.
983	L	0.453	.
984	D	0.483	.
985	K	0.487	.
986	V	0.490	.
987	E	0.506	E
988	A	0.525	E
989	E	0.522	E
990	V	0.523	E
991	Q	0.520	E
992	I	0.498	.
993	D	0.485	.
994	R	0.462	.
995	L	0.448	.
996	I	0.409	.

## DEPOIS DA MUTAÇÃO MISSENSE 2P

980	L	0.417	.
981	S	0.448	.
982	R	0.450	.
983	L	0.471	.
984	D	0.501	E
985	P	0.510	E
986	P	0.521	E
987	E	0.532	E
988	A	0.551	E
989	E	0.540	E
990	V	0.541	E
991	Q	0.532	E
992	I	0.506	E
993	D	0.493	.
994	R	0.469	.
995	L	0.446	.
996	I	0.408	.

# Predição de epítopos de MHC-I e MHC-II



# Ferramenta

IEDB Analysis Resource

**1**

**MHC-I Binding Predictions**

**2**

**MHC-II Binding Predictions**

# MHC-I Binding Predictions

Iara Vilarino

IEDB Analysis Resource

## MHC-I Binding Predictions

Prediction Method Version: v2.24 [\[Older versions\]](#)

**Specify Sequence(s)**

Enter protein sequence(s) in FASTA format or as whitespace-separated sequences.

```
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSQTNGT
KRFDNPVLPFNDGVYFASTSEKSNIIIRGWIFGTTLDLSDKTSLLIVNNATNWWIKVCEFCNDPFLGVYYHKNNKSW
MESEFRVYSSANNCTFEYVSQPFLLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVD
LPIGINITRFQTLALHRSYLTGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSEKCTLK
SFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSLYNSASFSTFKCY
GVSPTKLNDLCFTMYYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVIANNLSDKVGNNYNYLYRLF
RKSNNLKPFFERDISTEIIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRWWLSFELLHAPATVCGPKKSTN
LVKNNKCVNFMNGLTGTGVLTESNKKFLPFQGFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQV
AVLYQDWNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAHEVNNSECDIPIGAGICASYQTQTNSPRRA
RSVASQSIIAYTMSLGAENSVAYSNNIAIPTNFTISVTEILPVSMTKTSVDCTMYICGDSSTECNLLLLQYGSFC
```

Space\_separated format detected.

Or select file containing sequence(s):  Nenhum arquivo escolhido

**Choose a Prediction Method**

Prediction Method:  Show all the method versions:  NetMHCpan 4.1 EL (recommended epitope predictor-2023.09) [Help on prediction method selections](#)

**Specify what to make binding predictions for**

MHC source species: human

Show only frequently occurring alleles:  [?](#)

Select MHC allele(s):  Allele:  Length:  [Upload allele file](#) [?](#)

Select HLA allele reference set:  [?](#)  
[\(Specify MHC allele sequence\)](#)

**Specify Output**

Sort peptides by: Predicted Score (descend)

Output format: XHTML table

Email address (optional):  [?](#)

FREQUÊNCIAS  
ALÉLICAS HLA E  
CONJUNTOS DE  
REFERÊNCIA COM  
COBERTURA  
POPULACIONAL  
MÁXIMA

SEQUÊNCIA DA  
PROTEÍNA  
SPIKE

# MHC-I Binding Predictions

## MHC-I Binding Prediction Results

-----

Method used: netmhcpa\_n\_e1

allele	seq_num	start	end	length	peptide	score	percentile_rank	
HLA-A*01:01	1	865	873	9	LTDEMIAQY	0.997196	0.01	
HLA-B*35:01	1	896	904	9	IPFAMQMAY	0.994479	0.01	
HLA-B*58:01	1	625	633	9	HADQLTPTW	0.990462	0.01	
HLA-B*35:01	1	84	92	9	LPFNDGVYF	0.984594	0.01	

## MHC-I Binding Prediction Results

-----

Method used: netmhcpa\_n\_e1

allele	seq_num	start	end	length	peptide	score	percentile_rank	
HLA-A*02:01	1	322	331	10	PTESIVRFPN	0.0	100	

- TABELA COM **8 INFORMAÇÕES SOBRE CADA EPÍTOPO**
- NÚMERO DE EPÍTOPOS: 68283
- ÚLTIMA COLUNA: **AFINIDADE DE LIGAÇÃO** DO EPÍTOPO AO ALELO DE MHC-I
- **EXEMPLO:** EPÍTOPO COM 100% DE AFINIDADE, PERTENCENTE À REGIÃO DO RECEPTOR (RBD)
  - **TAMANHO DO EPITOPO:** 10
  - **SEQUÊNCIA DE AMINOÁCIDOS:** PTESIVRFPN
  - **É LIGANTE DE QUAL ALELO DE MHC I?** HLA-A\*02:01

# MHC-II Binding Predictions

Iara Vilarino

IEDB Analysis Resource

## MHC-II Binding Predictions

Specify Sequence(s)

Enter protein sequence(s) in FASTA format

```
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIRGWIF
GTTLDSTKQSLLVNNTATNVIKVFCEPQFCNDPFLGVYHKNKSWMESEFR
VYSSANNCTFEYVVSQPFLLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPIN
LVRDLPGQFSALEPLVDLPIGINITRFQTLALHRSYLTPGDSSSGWTAAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLGPFGEVFNATRFASVYAWNRRKRISNCVADY'SVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFFERDISTEIQAGS
TPCNGVEGFNCFYFPLQSYGFQPTNGVGYQPYRVVWLSFELLHAPATVCGPK
VSTNLLVLLKQVLEMEMLTQSPVLEKLVLEKQVLEKLVLEKQVLEKLVLEKQV
```

Space\_separated format detected.

Or select file containing sequence(s)  Nenhum arquivo escolhido

Choose a Prediction Method

Prediction Method [?](#) NetMHCIIpan 4.1 EL (recommended epitope predictor-2023.09) [Help on prediction method selections](#)  
Show all the method versions:

Specify what to make binding predictions for

Select species/locus Human, HLA-DR [v](#)

Select MHC allele(s)

Select a & p chains separately if applicable:  [?](#)

Select full HLA reference set:  [?](#)

Select 7-allele HLA reference set:  [?](#)

Allele  [Upload allele file](#) [?](#)

Select length(s)  [?](#)

default 12-18 as is

11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30

Specify Output

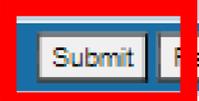
Sort peptides by Percentile Rank [v](#)

Output format XHTML table [v](#)

Email address (optional)  [?](#)

FREQUÊNCIAS ALÉLICAS HLA E CONJUNTOS DE REFERÊNCIA COM COBERTURA POPULACIONAL MÁXIMA

SEQUÊNCIA DA PROTEÍNA SPIKE



# MHC-I Binding Predictions

IEDB Analysis Resource

## MHC-II Binding Prediction Results

Method used: netmhciipan\_el

allele	seq_num	start	end	length	core_peptide	peptide	score	rank
HLA-DPA1*02:01/DPB1*05:01	1	136	150	15	FLGVYYHKN	CNDPFLGVYYHKNNK	0.3122	0.01
HLA-DPA1*02:01/DPB1*05:01	1	137	151	15	FLGVYYHKN	NDPFLGVYYHKNNKS	0.3376	0.01
HLA-DPA1*02:01/DPB1*05:01	1	339	353	15	FNATRFASV	GEVFNATRFASVYAW	0.3427	0.01
HLA-DPA1*02:01/DPB1*05:01	1	338	352	15	FNATRFASV	FGEVFNATRFASVYA	0.3865	0.01

## MHC-II Binding Prediction Results

Method used: netmhciipan\_el

allele	seq_num	start	end	length	core_peptide	peptide	score	rank
HLA-DRB1*11:01	1	478	492	15	CNGVEGFNC	TPCNGVEGFNCYFPL	0.0000	100.0

- TABELA COM 9 **INFORMAÇÕES SOBRE CADA EPÍTOPO**
- NÚMERO DE EPÍTOPOS: 33993
- **EXEMPLO:** EPÍTOPO COM 100% DE AFINIDADE, PERTENCENTE À REGIÃO DO RECEPTOR (RBD)
  - **TAMANHO DO EPITOPO:** 15
  - **SEQUÊNCIA DE AMINOÁCIDOS:** TPCNGVEGFNCYFPL
  - **É LIGANTE DE QUAL ALELO DE MHCII?** HLA-DRB1\*11:01

# Conclusão

# CONCLUSÃO

- Regiões UTR desempenham um papel crucial na regulação da expressão gênica;
- Otimização das sequências UTR é essencial para imunogenicidade e eficácia adequadas;

# CONCLUSÃO

- Estrutura secundária do mRNA, uso de códons e presença de sequências CG auxiliam na estabilidade e eficiência da tradução do mRNA.
- Mutações K986P e V987P
  - ➔ Estabilidade da proteína spike
  - ➔ Immunogenicidade
  - ➔ Eficiência da vacina

A decorative graphic consisting of numerous thin, dark blue wavy lines that flow and curve around the central text, creating a sense of movement and depth. The lines are most dense around the text and taper off towards the edges of the frame.

**Obrigado!**