



# Dando Vida a los Datos: Python en Bioinformática

Mercedes Didier Garnham

¿Reconocen esta escena?







¿Qué es la bioinformática?



# ¿Qué es la bioinformática?



Mentimeter

¿Qué es la bioinformática?

Biología con informática

Crear dinosaurios a partir de mosquitos

Ni idea

Análisis de datos biológicos

Es la informática aplicada a la biología

Programar dinosaurios

La informática aplicada a la biología la naturaleza los seres vivos

Biología mezclada con informática



# ¿Qué es la bioinformática?



Mentimeter

¿Qué es la bioinformática?

Base de datos de biológicas

El arte de combinar informática con biología?

Dinosaurios ciborgs (?)

Algo relacionado a los cyborgs

Informatica aplicada a la biología

Es la programación que tiene que ver con la parte biologica

La informatica aplicada a la biologia.

Representación de la biología con bits y bytes



# ¿Qué es la bioinformática?



Mentimeter

¿Qué es la bioinformática?

La información de los datos biológicos

el cuerpo de una ia

Procesamiento relacionado con cosas vivas o que estuvieron vivas

Uso de informática para el análisis de datos biológicos

Aguante python!

Trabajar sistemas biológicos y analizarlos con herramientas informáticas.

Para mí la bioinformática sería la mezcla de bio "vida" e informática.

Transformers



# ¿Qué es la bioinformática?



Mentimeter

¿Qué es la bioinformática?

Primera vez que escucho el termino

Informática aplicada a la biología

Tratamiento informatico de datos biologicos

Digimones

Algo muy complejo

Informática aplicada a la biología.

La bioinformática es una disciplina científica interdisciplinaria que combina la biología, la informática y las matemáticas para analizar, interpretar y utilizar datos biológicos.

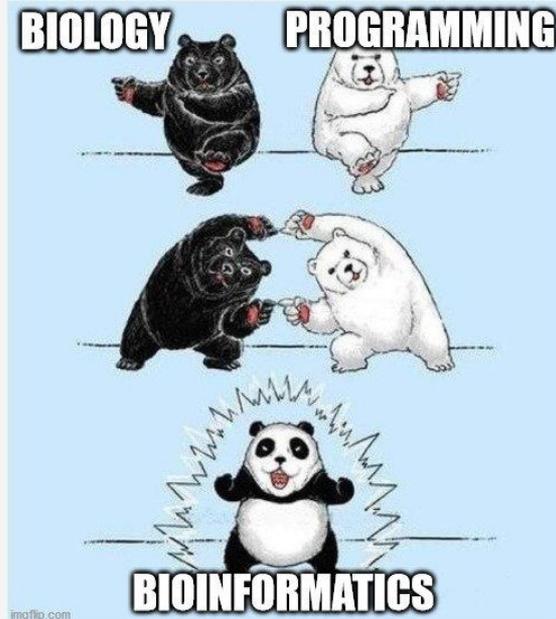
1



# ¿Qué es la bioinformática?



Análisis de datos biológicos mediante software.





UNIVERSIDAD NACIONAL  
DE HURLINGHAM



# ¿Cómo se hace el análisis de datos?



- Bases de datos especializadas
- Software open access y pago
- Programando

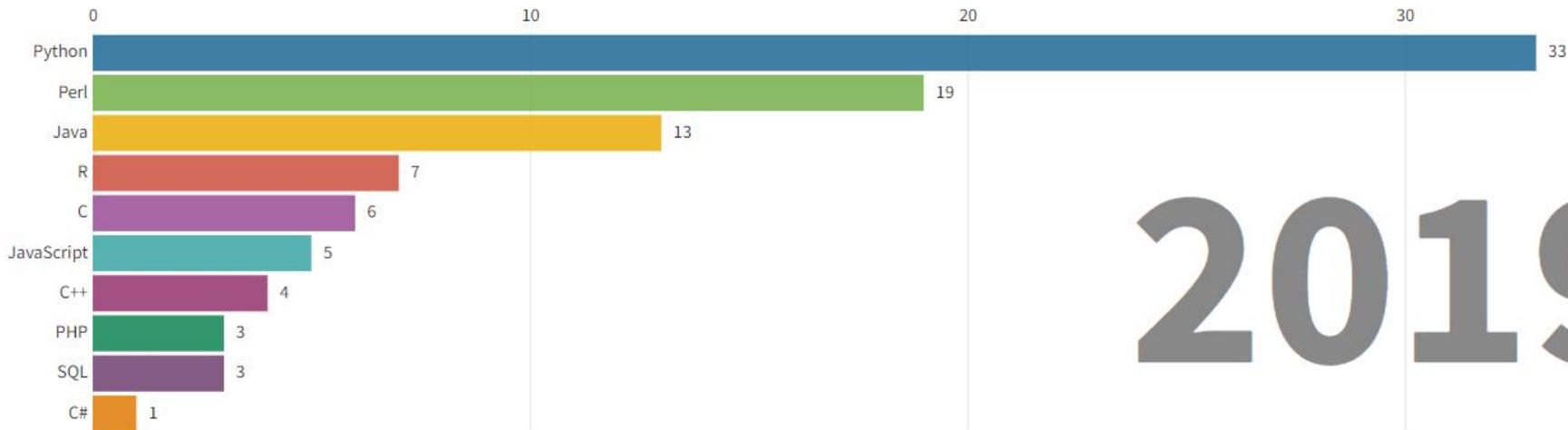
# ¿Cómo se hace el análisis de datos?



## Most popular programming languages for Bioinformatics (1998-2019)

Based on citations in the abstract

Replay



# 2019

<http://bioinfo.dcc.ufmg.br/history/home/languages>

\* A Flourish data visualization

# Ejemplos de la bioinformática



Ejemplo 1



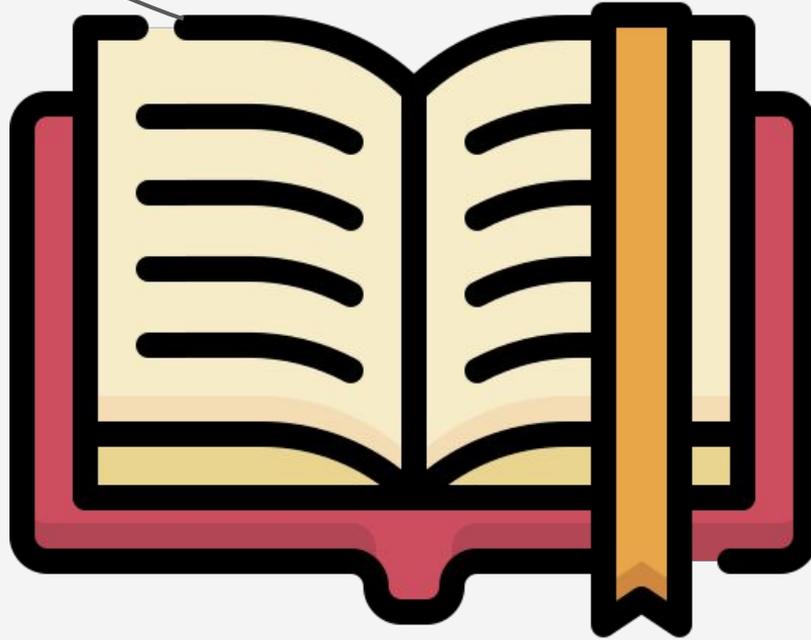
Ejemplo 2



Ejemplo 3

# Ejemplo 1: Estudio del ADN

Receta = Gen

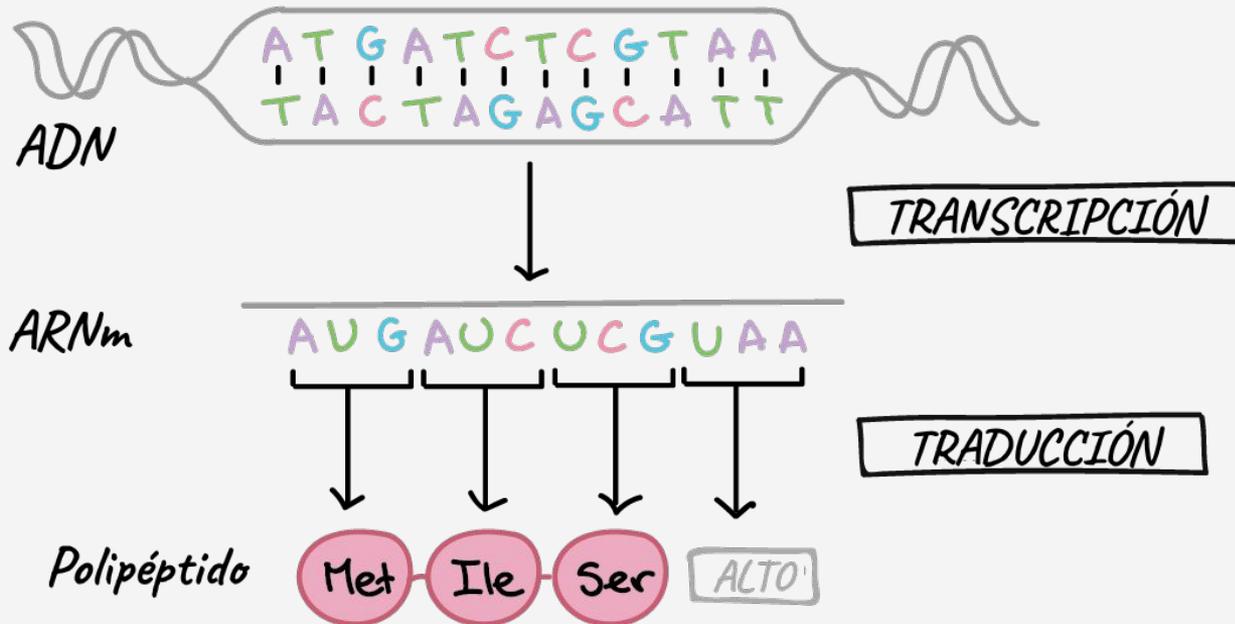


Libro de recetas = Genoma



# Ejemplo 1: Estudio del ADN

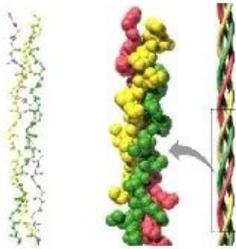
## EL DOGMA CENTRAL



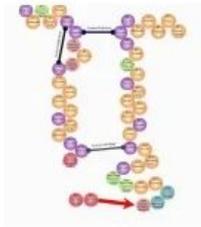
# Ejemplo 1: Estudio del ADN



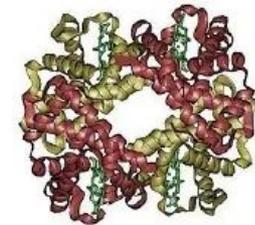
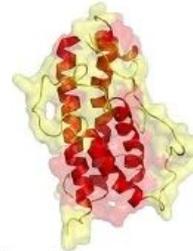
## PROTEÍNAS: Funciones



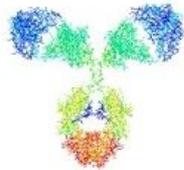
•Estructural (colágeno y queratina)



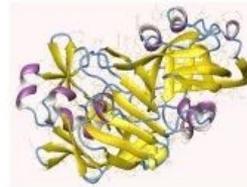
Reguladora (insulina y  
hormona del crecimiento)



•Transportadora  
(hemoglobina),

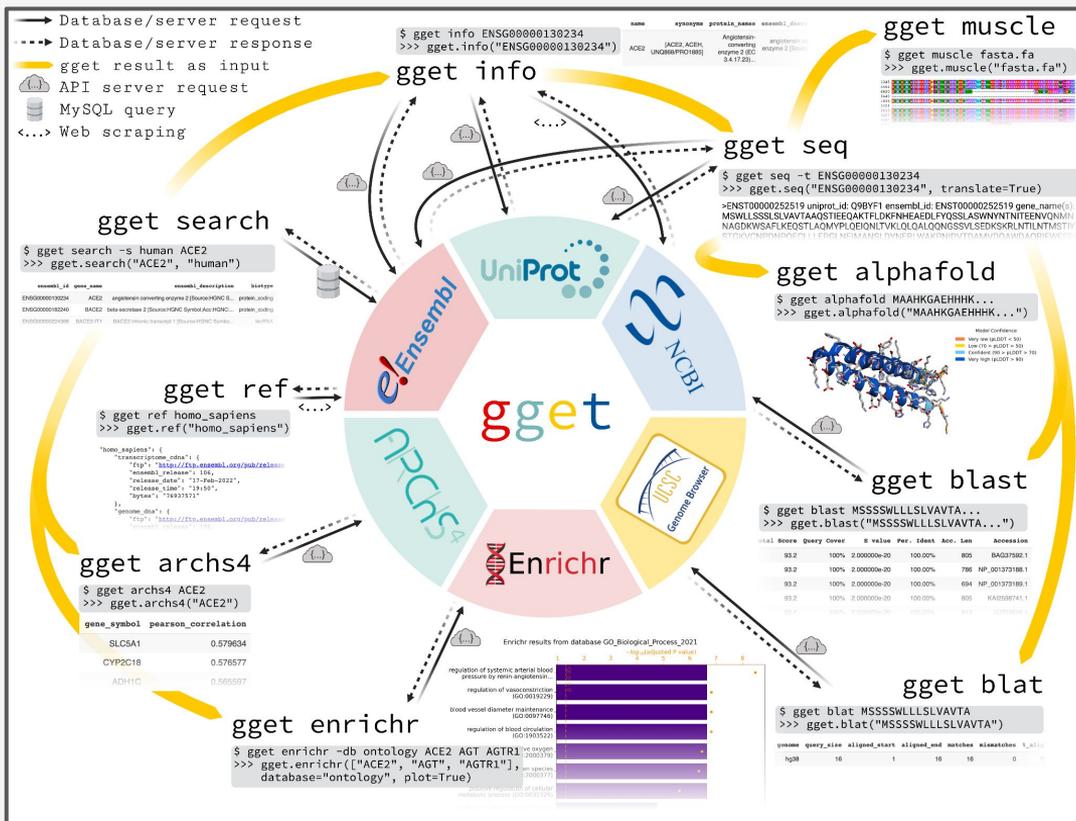


•Inmunológica  
(anticuerpos),



•Enzimática (pepsina),

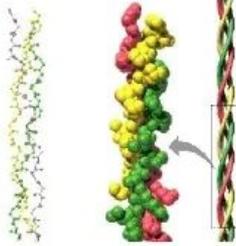
# Ejemplo 1: Estudio del ADN



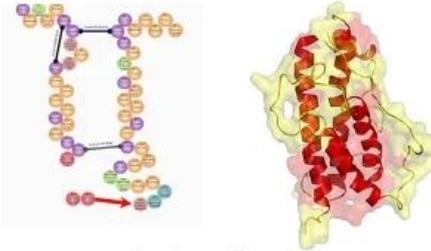
# Ejemplo 1: Estudio del ADN



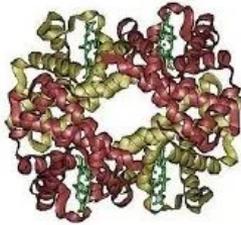
## PROTEÍNAS: Funciones



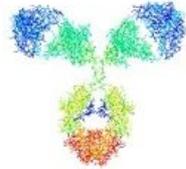
•Estructural (colágeno y queratina)



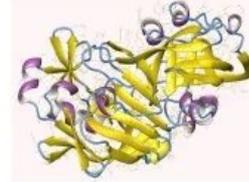
Reguladora (insulina y  
hormona del crecimiento)



•Transportadora  
(hemoglobina),



•Inmunológica  
(anticuerpos),

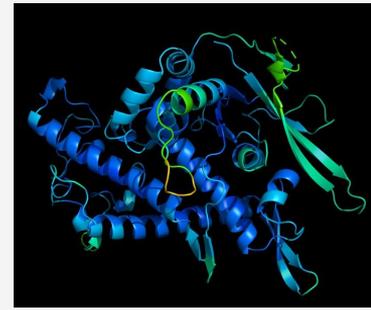


•Enzimática (pepsina),

# Ejemplo 1: Estudio del ADN



# Ejemplo 2: Estructura de proteínas

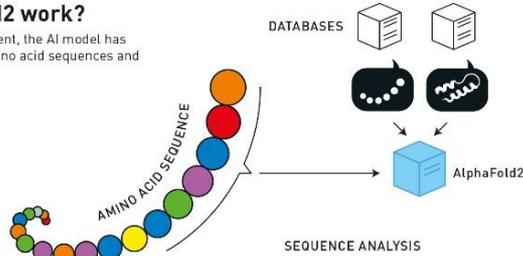


## How does AlphaFold2 work?

As part of AlphaFold2's development, the AI model has been trained on all the known amino acid sequences and determined protein structures.

### 1. DATA ENTRY AND DATABASE SEARCHES

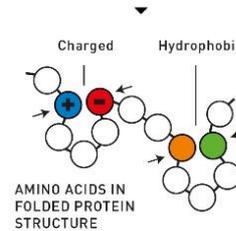
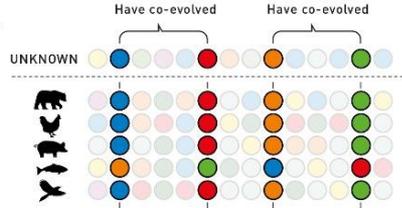
An amino acid sequence with unknown structure is fed into AlphaFold2, which searches databases for similar amino acid sequences and protein structures.



### 2. SEQUENCE ANALYSIS

The AI model aligns all the similar amino acid sequences – often from different species – and investigates which parts have been preserved during evolution.

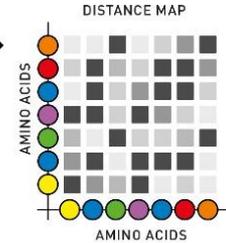
In the next step, AlphaFold2 explores which amino acids could interact with each other in the three-dimensional protein structure. Interacting amino acids co-evolve. If one is charged, the other has the opposite charge, so they are attracted to each other. If one is replaced by a water-repellent (hydrophobic) amino acid, the other also becomes hydrophobic.



AMINO ACIDS IN FOLDED PROTEIN STRUCTURE

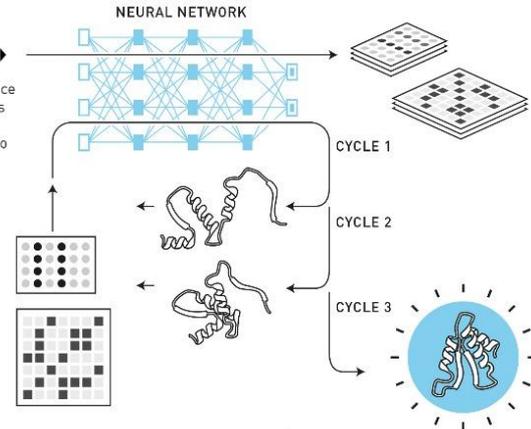
Using this analysis, AlphaFold2 produces a distance map that estimates how close amino acids are to each other in the structure.

← Furthest apart      Closest →



### 3. AI ANALYSIS

Using an iterative process, AlphaFold2 refines the sequence analysis and distance map. The AI model uses neural networks called transformers, which have a great capacity to identify important elements to focus on. Data about other protein structures – if they were found in step 1 – is also utilised.



### 4. HYPOTHETICAL STRUCTURE

AlphaFold2 puts together a puzzle of all the amino acids and tests pathways to produce a hypothetical protein structure. This is re-run through step 3. After three cycles, AlphaFold2 arrives at a particular structure. The AI model calculates the probability that different parts of this structure correspond to reality.



# Ejemplo 2: Estructura de proteínas



AlphaFold2.ipynb

Archivo Editar Ver Insertar Entorno de ejecución Herramientas Ayuda

+ Código + Texto Copiar en Drive

Conectar T4 Alta capacidad de RAM Gemini

## ColabFold v1.5.5: AlphaFold2 using MMseqs2

Easy to use protein structure and complex prediction using [AlphaFold2](#) and [Alphafold2-multimer](#). Sequence alignments/templates are generated through [MMseqs2](#) and [HHsearch](#). For more details, see [bottom](#) of the notebook, checkout the [ColabFold GitHub](#) and [Nature Protocols](#).

Old versions: [v1.4](#), [v1.5.1](#), [v1.5.2](#), [v1.5.3-patch](#)

[Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: Making protein folding accessible to all. Nature Methods, 2022](#)

Input protein sequence(s), then hit **Runtime** -> **Run all**

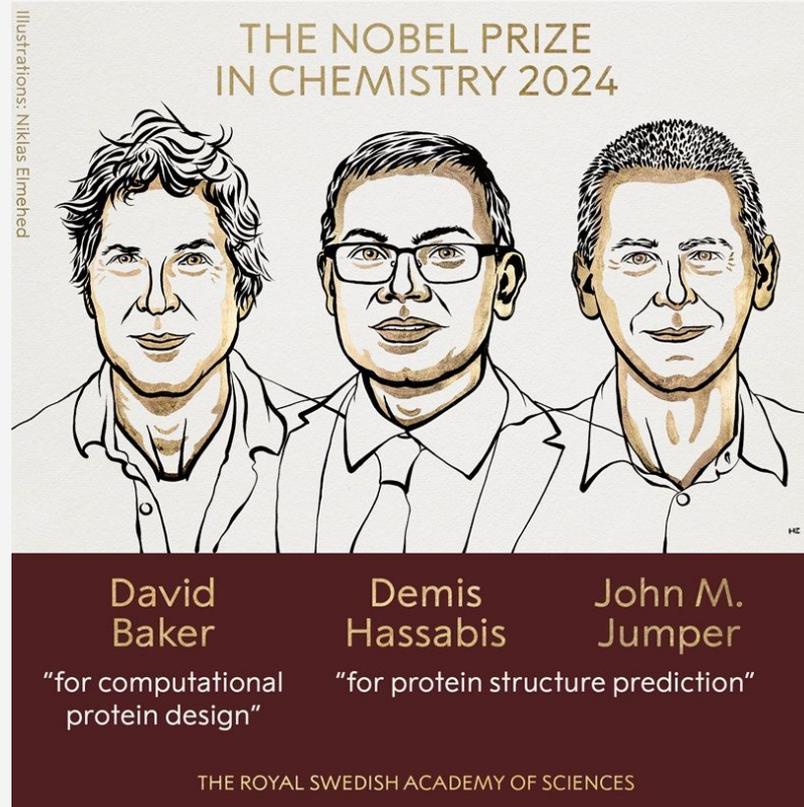
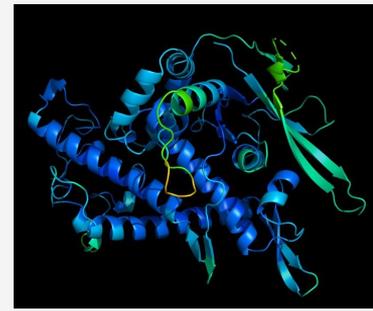
```
query_sequence: "PIAQIHILEGRSDEQKETLIREVSEAIRSLDAPLTSVRVIITEMAKGHFGIGGELASK"
```

- Use `:` to specify inter-protein chainbreaks for **modeling complexes** (supports homo- and hetro-oligomers). For example `PI...SK:PI...SK` for a homodimer

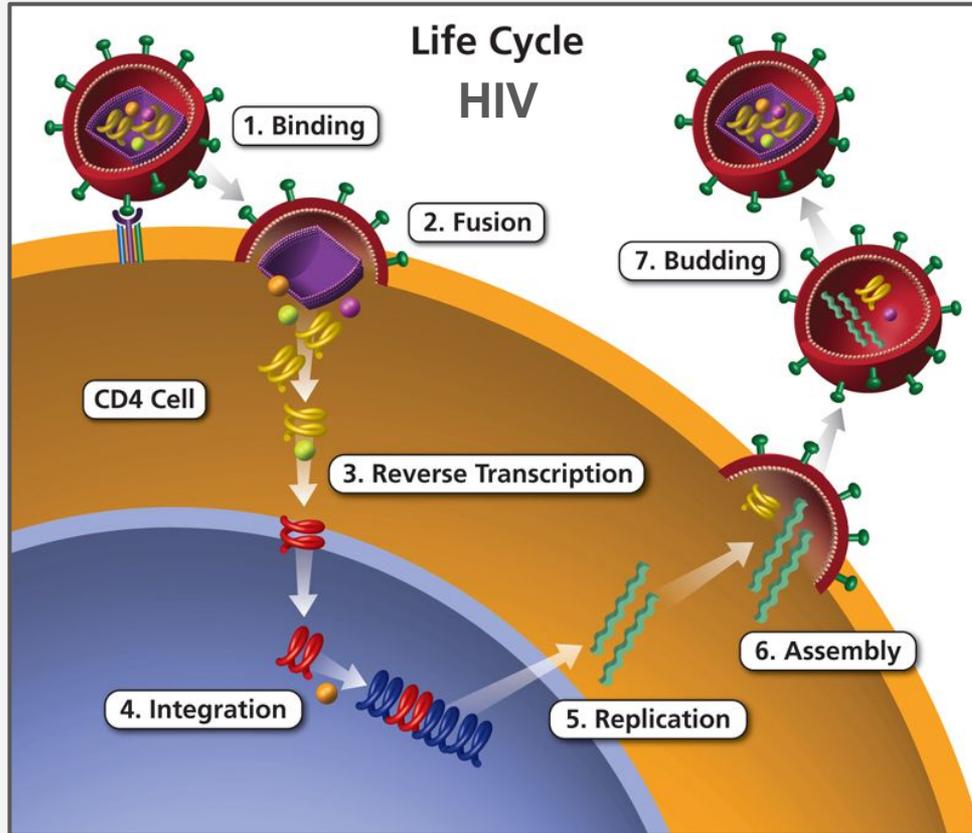
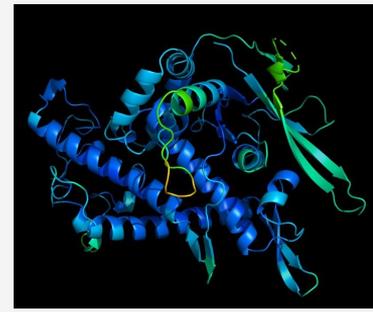
```
jobname: "test"
```



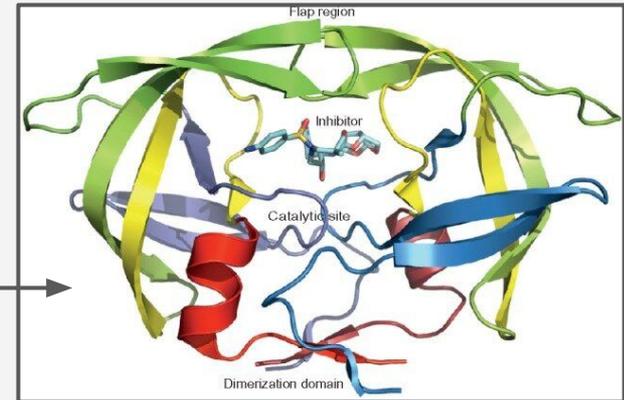
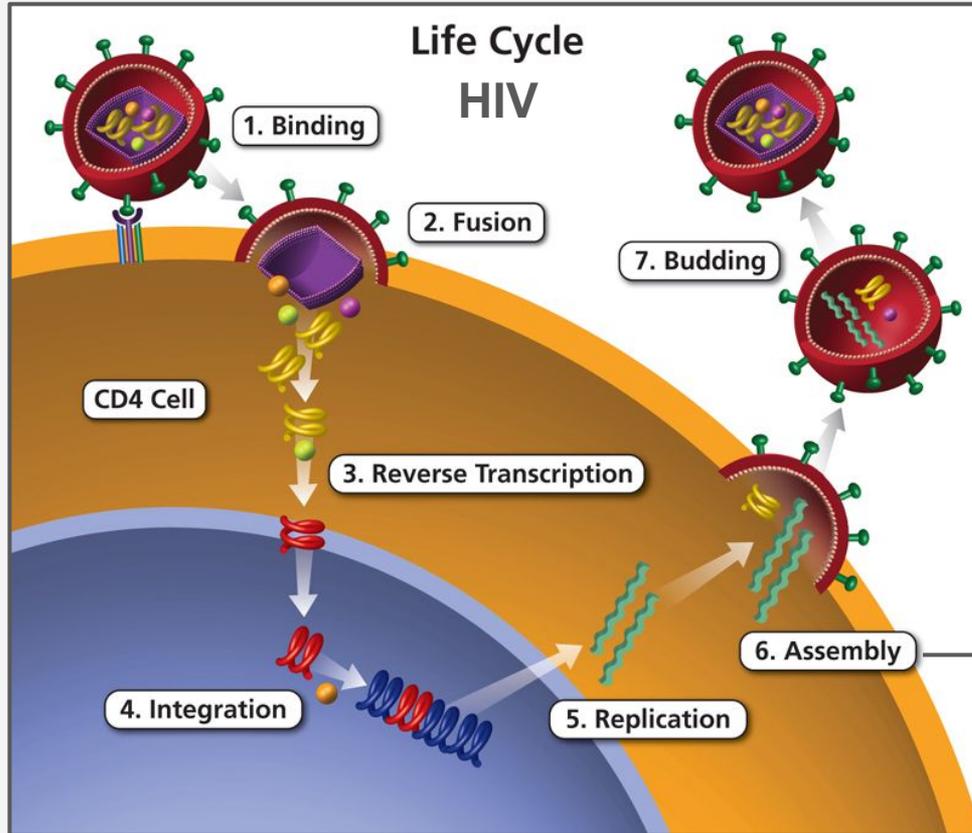
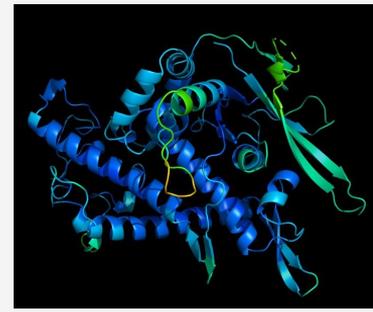
# Ejemplo 2: Estructura de proteínas



# Ejemplo 2: Estructura de proteínas

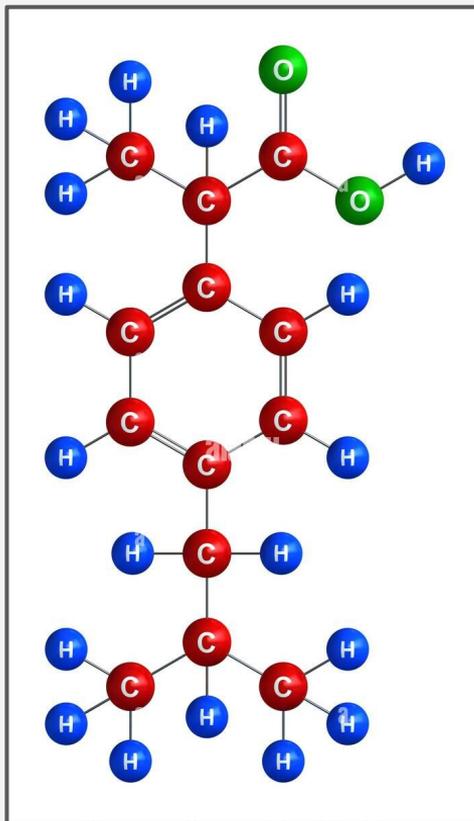


# Ejemplo 2: Estructura de proteínas

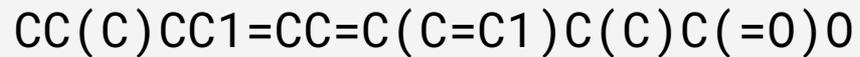
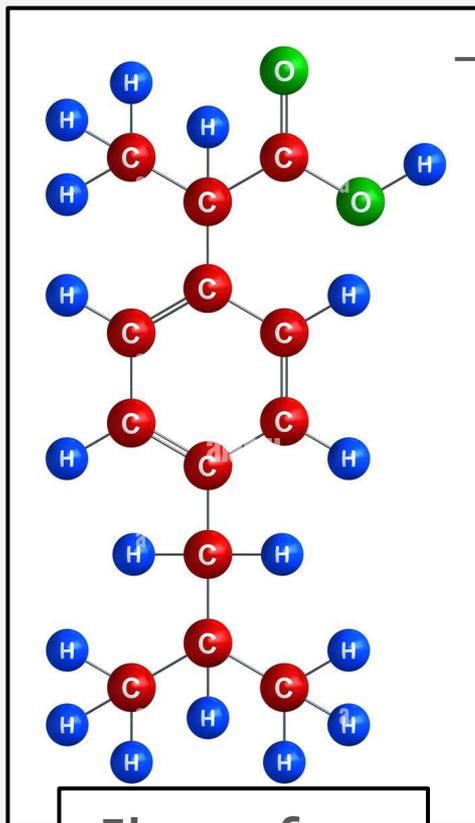


Saquinavir

# Ejemplo 3: Búsqueda de tratamientos

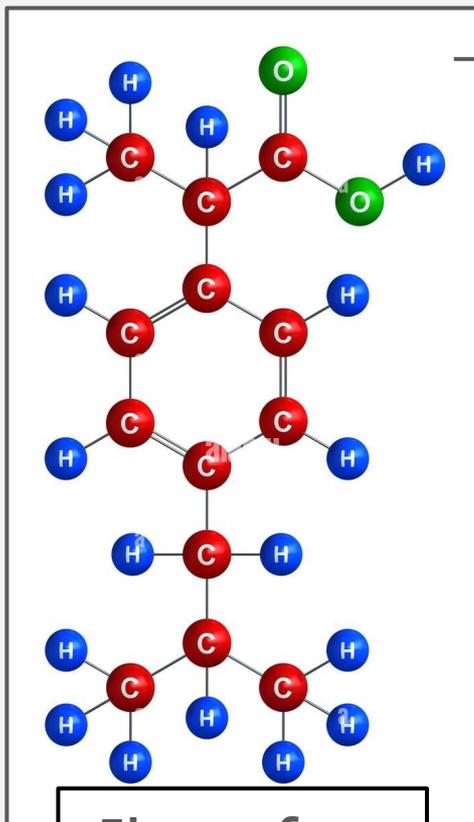


# Ejemplo 3: Búsqueda de tratamientos



Ibuprofeno

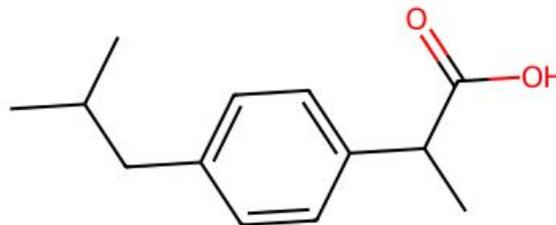
# Ejemplo 3: Búsqueda de tratamientos



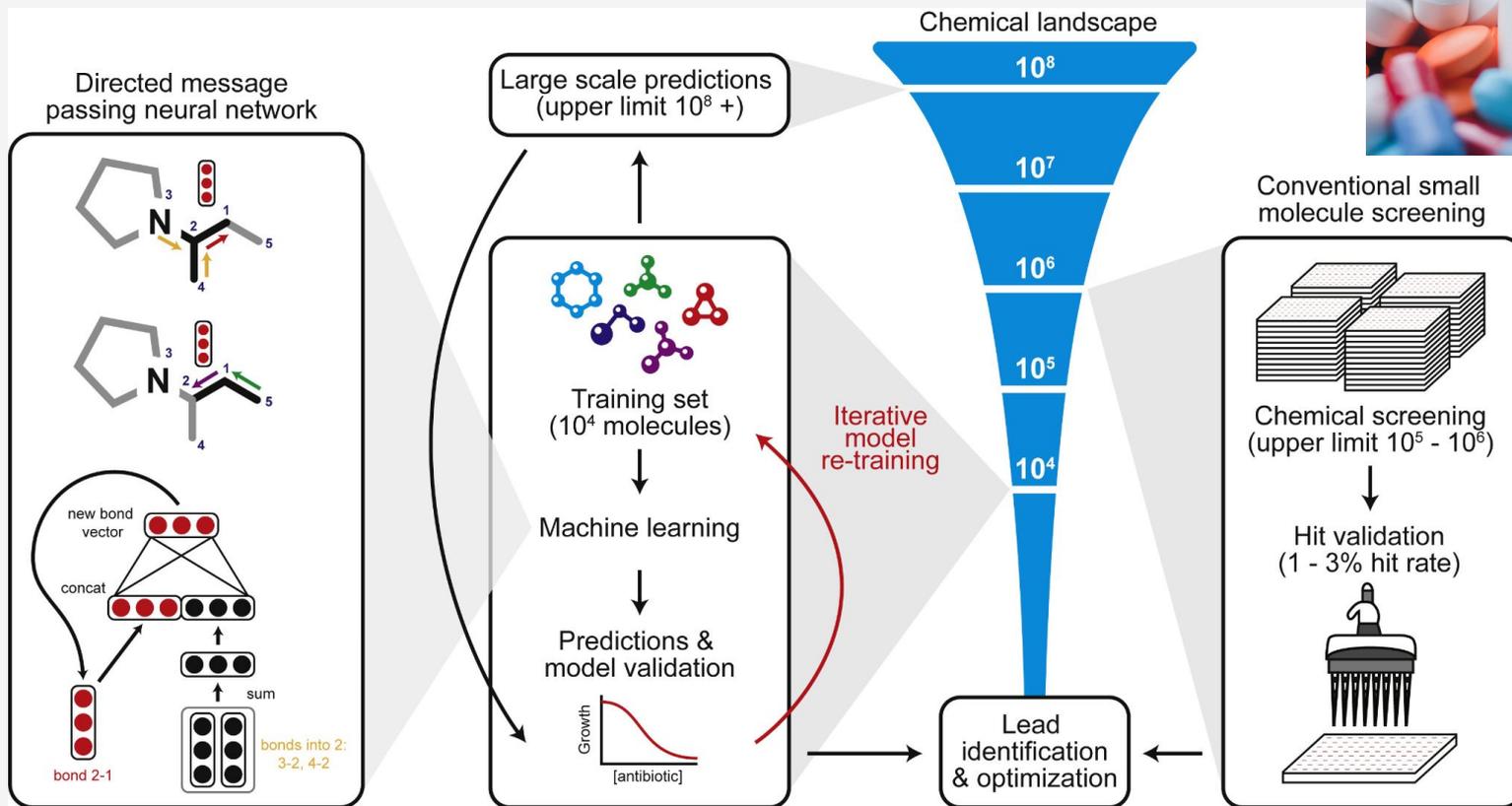
Ibuprofeno

CC(C)CC1=CC=C(C=C1)C(C)C(=O)O

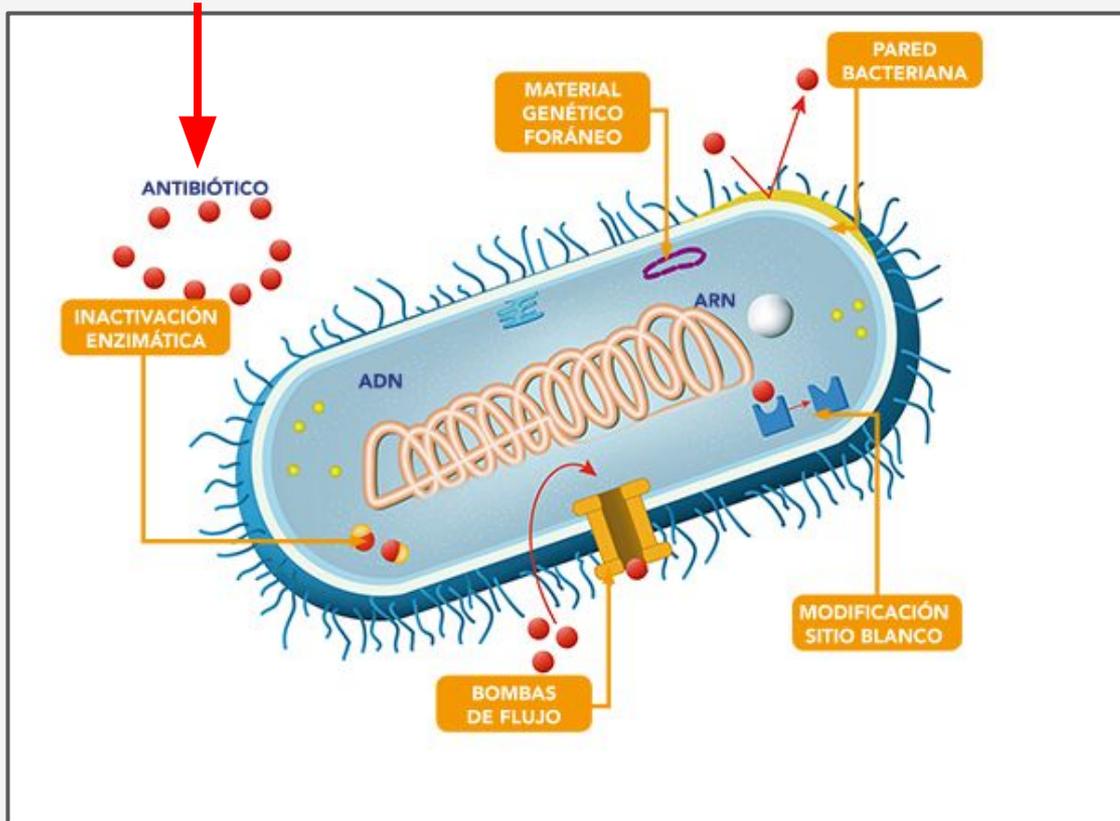
```
1 from rdkit import Chem
2 from rdkit.Chem import Draw
3 from rdkit.Chem import AllChem
4
5 # Define una cadena SMILES
6 smiles = "CC(C)CC1=CC=C(C=C1)C(C)C(=O)O"
7
8 # Crea un objeto molécula a partir de la cadena SMILES
9 mol = Chem.MolFromSmiles(smiles)
10
11 # Dibuja la molécula
12 mol
13
```



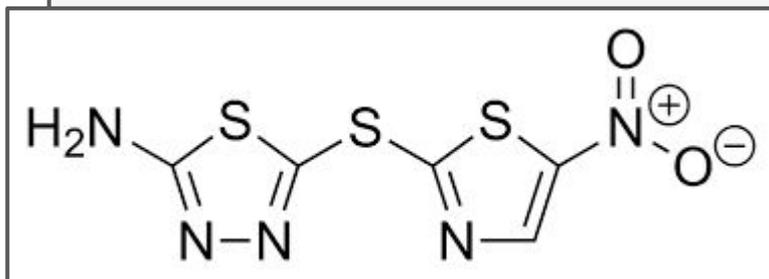
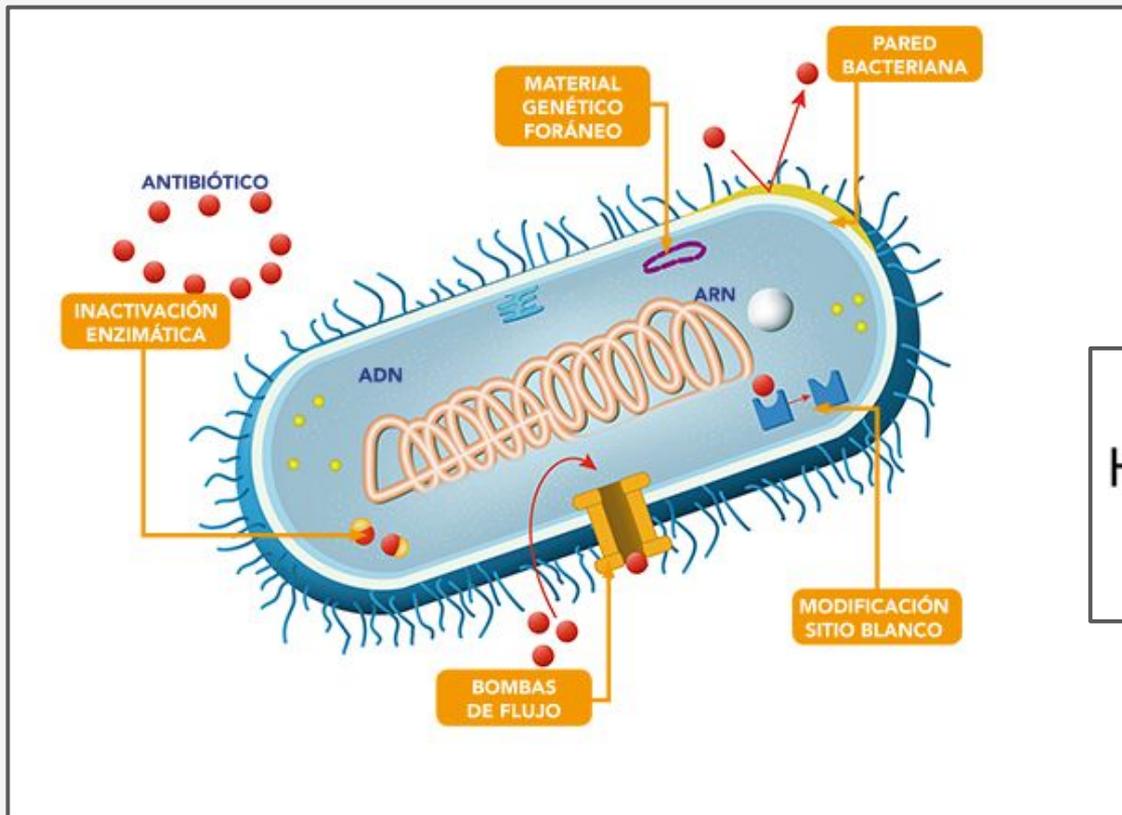
# Ejemplo 3: Búsqueda de tratamientos



# Ejemplo 3: Búsqueda de tratamientos



# Ejemplo 3: Búsqueda de tratamientos

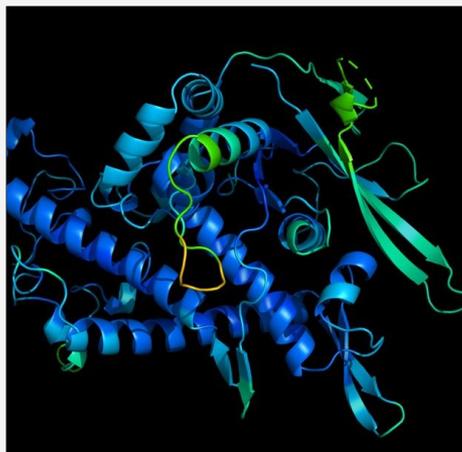


Halicina

# Ejemplos de la bioinformática



Ejemplo 1:  
Estudio del  
ADN



Ejemplo 2:  
Estructura de  
proteínas



Ejemplo 3:  
Búsqueda de  
tratamientos

# Ejemplos de la bioinformática

¿Cuál ejemplo te gustó más?

17



Ejemplo 1: Estudio del ADN

7



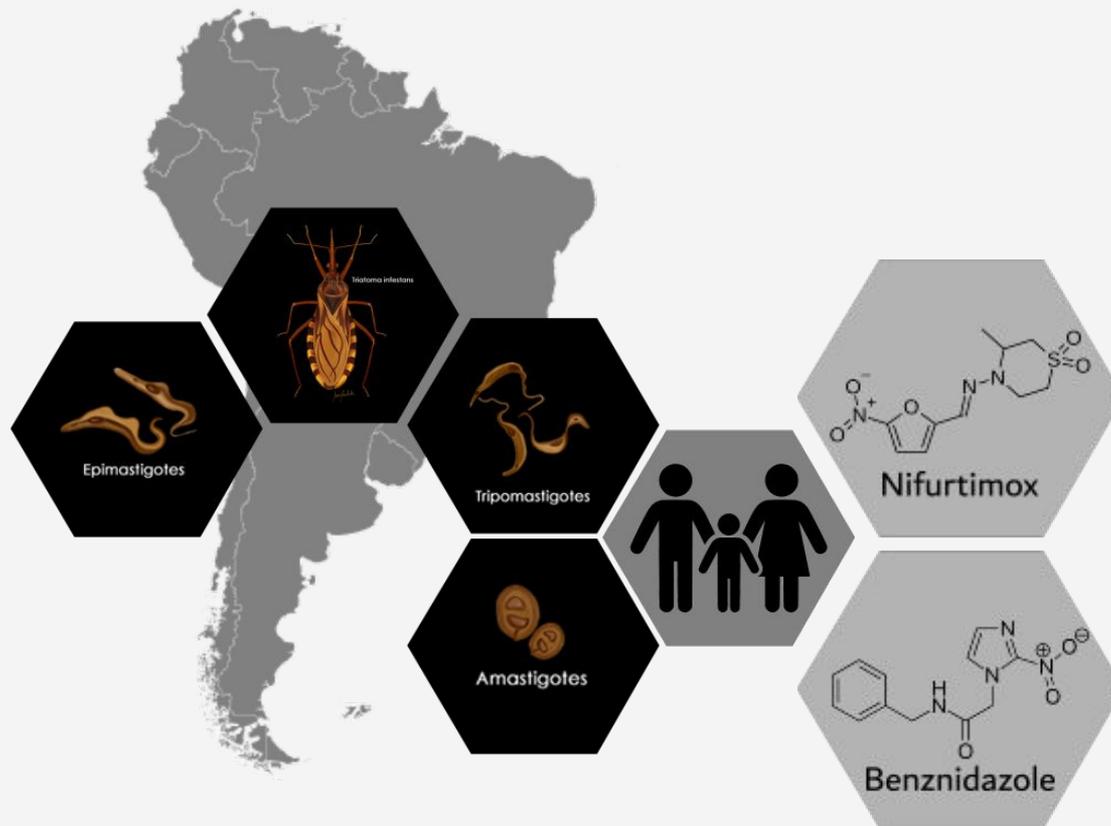
Ejemplo 2: Estructura de proteínas

22



Ejemplo 3: Búsqueda de tratamientos

# Enfermedad de Chagas





**¡Muchas gracias!**



Link a esta presentación

[http://tiny.cc/Python\\_en\\_bioinformatica](http://tiny.cc/Python_en_bioinformatica)

## Links de interés



Comunidad de Bioinformáticos de Argentina  
[https://linktr.ee/rsg\\_arg](https://linktr.ee/rsg_arg)



Ejemplo del uso de Python para analizar moléculas  
[http://tiny.cc/Ejemplo\\_Quimioinformatica](http://tiny.cc/Ejemplo_Quimioinformatica)



# Redes de contacto

Mail:

[mercedesdidiergarnham@gmail.com](mailto:mercedesdidiergarnham@gmail.com)

Github:

<https://github.com/mercedesgarnham>

LinkedIn:

<https://www.linkedin.com/in/mercedes-didier-garnham/>

Web:

<https://mercedesgarnham.github.io/>

