



CYSBOND

PREDICTOR

MANUAL PARA EL USUARIO

SamuelLozanoJuarez/CysBond_Predictor

Samuel Lozano Juárez

13 de junio de 2024

Bioinformática Estructural | Máster en Bioinformática

Índice de contenidos

1. Qué es CysBond Predictor.....	1
2. Cómo utilizar CysBond Predictor.....	1
3. Salidas generadas.....	2
4. Dependencias necesarias	4
5. Ejemplos de uso	4

1. Qué es CysBond Predictor

Se trata de un script de Python que permite identificar potenciales puentes disulfuro en un archivo PDB que contenga la estructura de una proteína. Además de reportar estos puentes disulfuro también genera un fichero `.py` que permite la visualización de la estructura tridimensional de la proteína y los puentes disulfuro utilizando la herramienta PyMOL.

El programa permite al usuario personalizar los parámetros que determinan los potenciales puentes disulfuro (distancia entre cisteínas y ángulo diedro formado). Para que un par de cisteínas sean consideradas para un potencial puente disulfuro deben cumplir por defecto con un ángulo de entre 84° y 96° y una distancia de entre 1.5 \AA y 2.5 \AA . Por motivos de calidad de la estructura, las cisteínas con un B-factor superior a 30 \AA^2 si la estructura se ha obtenido experimentalmente, o un pLDDT inferior a 50 si la estructura ha obtenido por predicción, serán descartadas.

La información del programa, así como este Manual y los ficheros `.pdb` de ejemplo pueden encontrarse en el repositorio

https://github.com/SamuelLozanoJuarez/CysBond_Predictor.

2. Cómo utilizar CysBond Predictor

Las dependencias necesarias para la ejecución del script se describen en la Sección 4. El script debe ejecutarse empleando Python mediante la línea de comandos del Sistema Operativo. Los argumentos que deben indicarse son los siguientes:

- `-i (--input)`: ruta relativa o absoluta hasta el fichero PDB con la estructura de la proteína en la que se quieren detectar los puentes disulfuro. Es recomendable introducir la ruta relativa, pero también es aceptada la absoluta. **Argumento obligatorio.**
- `-o (--output)`: ruta relativa que indica el directorio en que se desea guardar el script `.py` generado para la posterior visualización usando PyMOL. Si el directorio no existe, será creado. **Argumento opcional.** Por defecto almacena el script generado en el directorio actual.
- `-a (--angle)`: valores mínimo y máximo del ángulo diedro que debe formarse entre cisteínas para considerar el potencial de puente disulfuro. Los valores deben ser enteros, ir entre paréntesis y separados por una coma. **Argumento opcional.** Por defecto el rango de valores para el ángulo diedro es (84,96).
- `-d (--distance)`: valores mínimo y máximo de la distancia a la que deben encontrarse dos cisteínas para considerar el potencial de puente disulfuro. Los valores deben ser decimales, ir entre paréntesis y separados por una coma. **Argumento opcional.** Por defecto el rango de valores para la distancia es (1.5,2.5).
- `-aci (--angle_ci)`: valor entero que representa la cantidad en que se va a ampliar el rango de valores del ángulo tanto hacia el límite superior como inferior para la búsqueda de puentes disulfuro. Un valor de `aci=5` implica que se considerarán como potenciales los puentes disulfuro con un ángulo comprendido

dentro del rango definido en el parámetro *angle* \pm 5. **Argumento opcional.** Por defecto el valor es 5.

- **-dci** (**--distance_ci**): valor decimal que representa la cantidad en que se va a ampliar el rango de valores de la distancia tanto hacia el límite superior como inferior para la búsqueda de puentes disulfuro. Un valor de *aci*=0.1 implica que se considerarán como potenciales los puentes disulfuro entre cisteínas a una distancia comprendida dentro del rango definido en el parámetro *distance* \pm 0.1. **Argumento opcional.** Por defecto el valor es 0.1.
- **-sci** (**--show_ci**): booleano que determina si deben mostrarse en PyMOL los potenciales puentes disulfuros que no cumplen con los valores de *angle* y *distance*, pero sí con estos valores ampliados en su intervalo de confianza correspondiente (*angle* \pm *aci* y *distance* \pm *dci* respectivamente). **Argumento opcional.** Por defecto el valor es False.

A continuación se muestra un ejemplo de ejecución incluyendo todos los parámetros:

```
python cysbond_pred.py -i ejemplos/1jl9.pdb -o resultados -a (84,96) -d (1.5,2.5) -aci 5 -dci 0.1 -sci False
```

3. Salidas generadas

Como consecuencia de la ejecución del script vamos a obtener dos tipos de salidas: información que se nos muestra por pantalla y un archivo *.py* con el código para visualizar la proteína y los puentes disulfuro en PyMOL.

Por pantalla se muestran, en formato de tabla, los potenciales puentes disulfuro que se han detectado considerando los valores de ángulos diedros y distancia entre cisteínas introducidos como parámetros (**-a** y **-d** respectivamente). Para cada potencial puente se indica la información de las cisteínas que lo forman (posición y cadena), y los valores de distancia y ángulo diedro:

The following potential disulfide bonds have been detected:				
Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 148 A	CYS 191 A	2.039062261581421	91.12013043292747
2	CYS 182 A	CYS 224 A	2.036055088043213	95.60815227597449
3	CYS 282 A	CYS 332 A	2.034346342086792	85.45929775519888

Además, se muestran por pantalla aquellos puentes disulfuro que no cumplen alguno de los valores de distancia o ángulo introducidos como parámetros pero que deberían ser revisados, ya que sí que se encuentran dentro del rango “de confianza” introducido como parámetro (**-dci** y **-aci**). Es decir, si por ejemplo el rango de ángulos y distancias que hemos introducido para considerar un puente disulfuro es (86,94) y (1.5,2.5) respectivamente, y los parámetros **-dci** y **-aci** son 5 y 0.1 respectivamente, un par de cisteínas que formen un ángulo de 84° y se encuentren a 1.4 Å no serán mostradas en la tabla anterior, pero sí que se incluirán en esta segunda tabla:

The following disulfide bonds do not meet the requirements but should be reviewed:

Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 14 B	CYS 31 B	2.0282211303710938	96.57688257030237
2	CYS 33 B	CYS 42 B	2.029945135116577	81.76715236872248

En el caso de que no se encuentre ningún potencial puente disulfuro en alguna de las dos tablas, se mostrará en el lugar de la tabla un mensaje avisando de este hecho:

```

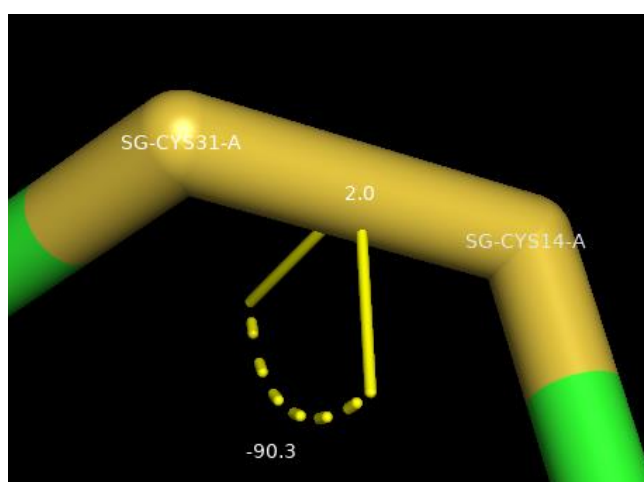
-----
No potential disulfide bond has been detected.
-----

-----
There are no additional disulphide bonds that need to be checked.
-----

```

El fichero *.py* generado tiene el mismo nombre que el archivo *.PDB* introducido como parámetro en la ejecución del script. Este fichero *.py* contiene el código necesario para visualizar en PyMOL la proteína y los puentes disulfuro encontrados. La distancia entre cisteínas y el ángulo formado también se muestran en PyMOL.

Para ejecutar este archivo necesitaremos tener instalado PyMOL en nuestro equipo. Después podremos abrir el fichero *.py* bien por línea de comandos (`pymol nombre.py`) o bien ejecutando primero la aplicación y seleccionando la opción File/Run Script, donde seleccionaremos nuestro fichero *.py*. En caso de que se haya encontrado algún puente disulfuro, al ejecutarse el fichero *.py* en PyMOL se nos cargará la proteína, y se mostrará el enlace entre las cisteínas involucradas, así como la distancia entre ellas y el ángulo diedro:



Si no se ha encontrado ningún puente disulfuro simplemente se cargará la proteína.

En la Sección 5 se incluyen más ejemplos de uso del script, con sus correspondientes salidas.

4. Dependencias necesarias

Para poder ejecutar el script es necesario tener instalados los siguientes programas y paquetes. Se incluyen las versiones utilizadas en el desarrollo del script, con las que no hay problemas de compatibilidades:

- Python versión 3.11.8
- PyMOL versión 3.0.2
- Paquete *Biopython* versión 1.83
- Paquete *NumPy* versión 1.26.4
- Paquete *tabulate* versión 0.9.0

5. Ejemplos de uso

En esta sección se incluyen varios ejemplos de uso del programa con distintos ficheros PDB correspondientes a proteínas con estructuras predichas y también obtenidas experimentalmente. Los ficheros PDB utilizados pueden encontrarse en la carpeta *ejemplos*.

7rax

- Estructura obtenida experimentalmente (difracción de rayos X)
- 1 cadena
- Con parámetros por defecto se encuentran 0 puentes potenciales.
- Tampoco se encuentran otros posibles puentes para chequear.

```
python cysbond_pred.py -i ejemplos/7rax.pdb
```

```
-----  
No potential disulfide bond has been detected.  
-----  
  
-----  
There are no additional disulphide bonds that need to be checked.  
-----
```

2qwo

- Estructura obtenida experimentalmente (difracción de rayos X)
- 2 cadenas
- Con parámetros por defecto se encuentran 0 puentes potenciales.
- Se encuentra 1 puente para chequear.

```
python cysbond_pred.py -i ejemplos/2qwo.pdb
```

```
-----
No potential disulfide bond has been detected.
-----

The following disulfide bonds do not meet the requirements but should be reviewed:
+-----+-----+-----+-----+
| Bond | Cys1 | Cys2 | Distance | Dihedral Angle |
+-----+-----+-----+-----+
| 1 | CYS 171 A | CYS 876 B | 2.050916910171509 | 80.46914958615396 |
+-----+-----+-----+-----+
```

Si cambiamos los parámetros por defecto y establecemos como rango de ángulos (80,85) en vez de (84,96), vemos que este enlace pasa a considerarse potencial puente disulfuro.

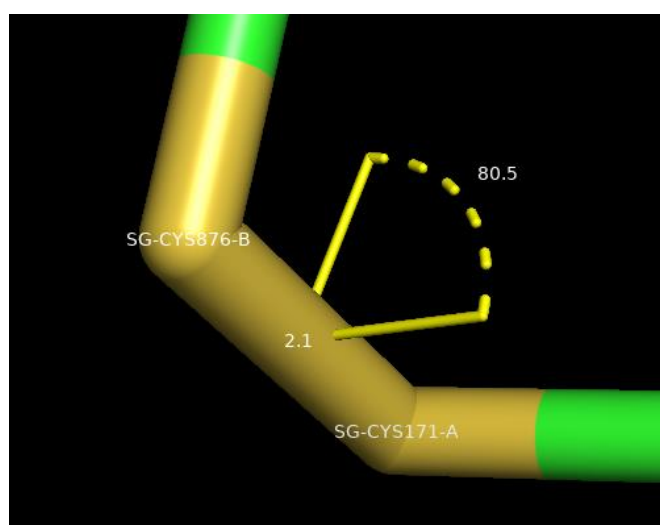
```
python cysbond_pred.py -i ejemplos/2qwo.pdb -a (80,85)
```

```
-----
The following potential disulfide bonds have been detected:
+-----+-----+-----+-----+
| Bond | Cys1 | Cys2 | Distance | Dihedral Angle |
+-----+-----+-----+-----+
| 1 | CYS 171 A | CYS 876 B | 2.050916910171509 | 80.46914958615396 |
+-----+-----+-----+-----+

-----
There are no additional disulphide bonds that need to be checked.
-----
```

Si visualizamos el script generado en Pymol podremos ver este puente disulfuro que ahora sí que se considera como potencial:

```
pymol 2qwo.py
```



1fs3

- Estructura obtenida experimentalmente (difracción de rayos X)
- 1 cadena
- Con parámetros por defecto se encuentran 2 puentes potenciales.
- No se encuentran puentes alternativos para chequear.

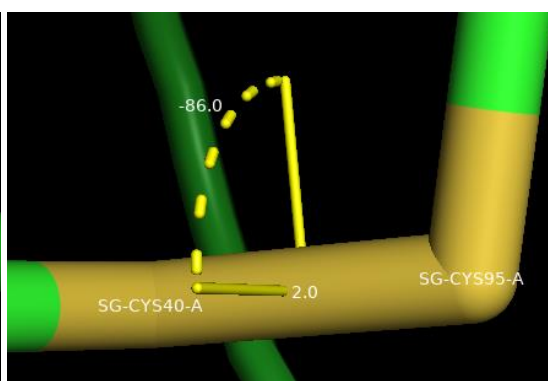
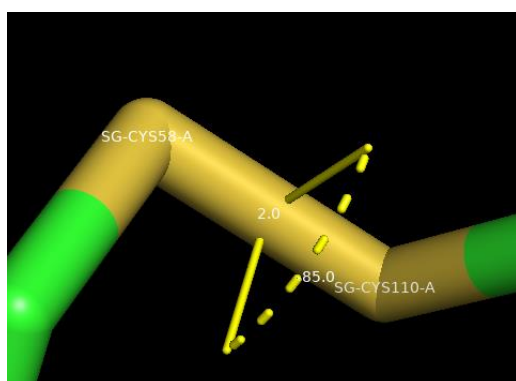
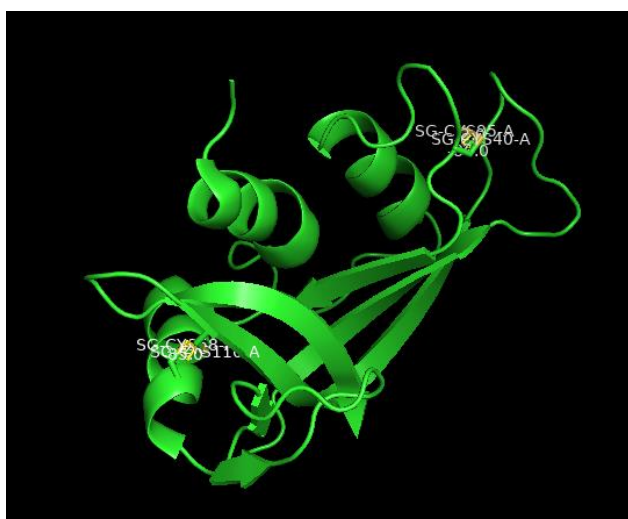
```
python cysbond_pred.py -i ejemplos/1fs3.pdb
```

```
The following potential disulfide bonds have been detected:
+-----+
| Bond | Cys1 | Cys2 | Distance | Dihedral Angle |
+-----+
| 1 | CYS 40 A | CYS 95 A | 2.0255184173583984 | 86.03217950656176 |
| 2 | CYS 58 A | CYS 110 A | 2.0314524173736572 | 84.9913555136457 |
+-----+

-----
There are no additional disulphide bonds that need to be checked.
-----
```

Podemos visualizar los resultados en PyMOL:

```
pymol 1fs3.py
```



6uqc

- Estructura obtenida experimentalmente (difracción de rayos X)
- 8 cadenas
- Con parámetros por defecto se encuentra 1 puente potencial.
- Se encuentra 1 puente adicional para chequear.

```
python cysbond_pred.py -i ejemplos/6uqc.pdb
```

```
The following potential disulfide bonds have been detected:
```

Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 367 D	CYS 425 D	2.0208852291107178	84.99430470706584

```
The following disulfide bonds do not meet the requirements but should be reviewed:
```

Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 367 C	CYS 425 C	2.0224432945251465	82.2088572767831

Podemos ampliar el intervalo de confianza para el ángulo, estableciendo el parámetro `-dci` a 21, para que considere como puentes para chequear aquellos con un ángulo entre (63,117). Al hacer este ajuste el programa encuentra un puente más para chequear.

Además, podemos modificar el parámetro `-sci` y establecerlo como True, para que muestre en PyMOL también los puentes a chequear.

```
python cysbond_pred.py -i ejemplos/6uqc.pdb -aci 21 -sci True
```

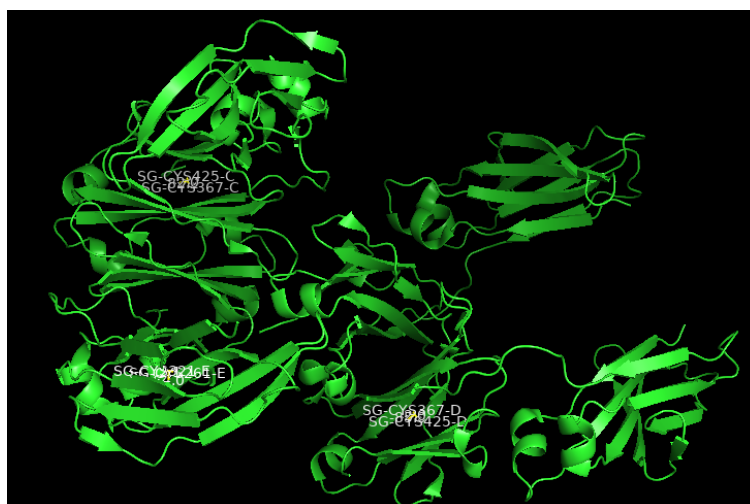
```
The following potential disulfide bonds have been detected:
```

Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 367 D	CYS 425 D	2.0208852291107178	84.99430470706584

```
The following disulfide bonds do not meet the requirements but should be reviewed:
```

Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 367 C	CYS 425 C	2.0224432945251465	82.2088572767831
2	CYS 261 E	CYS 321 E	2.092224597930908	63.96621506973581

Ahora en PyMOL nos mostrará los 3 puentes disulfuro, y no únicamente el potencial.



1jl9

- Estructura obtenida experimentalmente (difracción de rayos X)
- 2 cadenas
- Con parámetros por defecto se encuentra 1 puente potencial.
- Se encuentran 2 puentes adicionales para chequear.

```
python cysbond_pred.py -i ejemplos/1jl9.pdb
```

```
The following potential disulfide bonds have been detected:
```

Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 14 A	CYS 31 A	2.0206663608551025	90.2746789051136

```
The following disulfide bonds do not meet the requirements but should be reviewed:
```

Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 14 B	CYS 31 B	2.0282211303710938	96.57688257030237
2	CYS 33 B	CYS 42 B	2.029945135116577	81.76715236872248

2ltq

- Estructura obtenida experimentalmente (NMR)
- 6 cadenas
- Con parámetros por defecto se encuentran 6 puentes potenciales.
- Se encuentran 6 puentes adicionales para chequear.

```
python cysbond_pred.py -i ejemplos/2ltq.pdb
```

The following potential disulfide bonds have been detected:

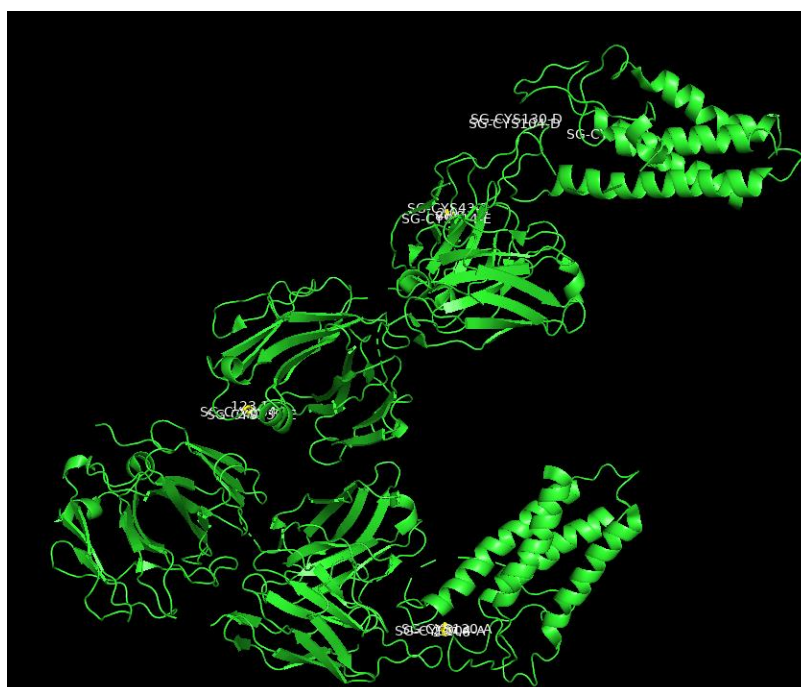
Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 43 E	CYS 114 E	2.044865846633911	86.78958149466047
2	CYS 44 D	CYS 44 D	1.9796417951583862	91.75977704840425
3	CYS 130 D	CYS 130 D	1.6037118434906006	91.29715592298174
4	CYS 104 A	CYS 130 A	2.0203959941864014	88.73616135634899
5	CYS 104 D	CYS 104 D	2.1019585132598877	84.10646917752163
6	CYS 134 F	CYS 239 E	1.5342936515808105	94.3178433354795

The following disulfide bonds do not meet the requirements but should be reviewed:

Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 104 A	CYS 104 A	1.452910304069519	79.69687540909062
2	CYS 96 C	CYS 22 C	1.7086142301559448	80.24744158881812
3	CYS 44 D	CYS 44 D	2.5850794315338135	87.34680458759294
4	CYS 114 E	CYS 43 E	2.485166311264038	83.05744182089052
5	CYS 130 A	CYS 104 A	1.475786805152893	94.28725919545514
6	CYS 130 A	CYS 130 A	2.556990385055542	79.03179383826587
7	CYS 134 F	CYS 239 E	1.8699662685394287	96.06784466414831

Si deseamos podemos ver los 6 puentes potenciales en PyMOL (podríamos hacer zoom sobre cada uno de ellos y veríamos el ángulo y las distancias entre cisteínas).

```
pymol 2ltq.py
```



AF-P00785-F1-model_v4

- Estructura predicha (AlphaFold 2)
- 1 cadena
- Con parámetros por defecto se encuentran 3 puentes potenciales.
- Se encuentran 0 puentes adicionales para chequear.

```
python cysbond_pred.py -i ejemplos/AF-P00785-F1-model_v4.pdb
```

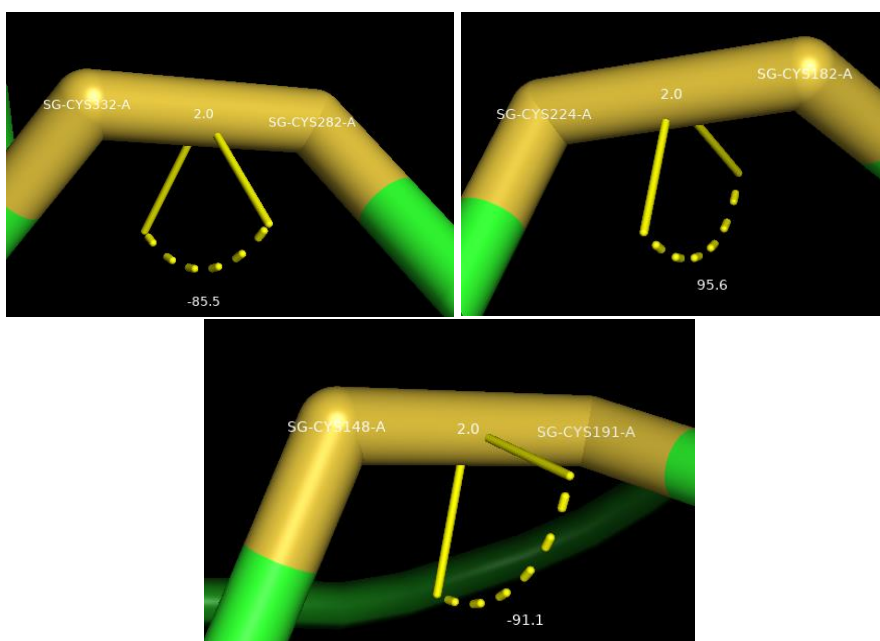
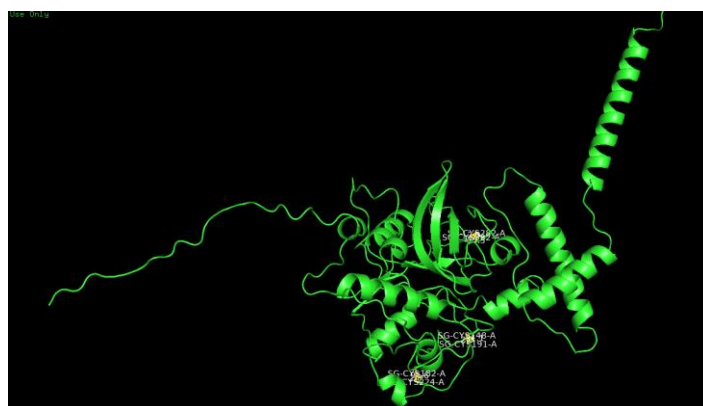
```
The following potential disulfide bonds have been detected:
```

Bond	Cys1	Cys2	Distance	Dihedral Angle
1	CYS 148 A	CYS 191 A	2.039062261581421	91.12013043292747
2	CYS 182 A	CYS 224 A	2.036055088043213	95.60815227597449
3	CYS 282 A	CYS 332 A	2.034346342086792	85.45929775519888

```
There are no additional disulphide bonds that need to be checked.
```

Podemos visualizar la estructura y puentes disulfuro en PyMOL.

```
pymol AF-P00785-F1-model_v4.py
```



AF-Q869Z0-F1-model_v4

- Estructura predicha (AlphaFold 2)
- 1 cadena
- Con parámetros por defecto no se encuentran puentes potenciales.
- Se encuentra 1 puente adicional para chequear.

```
python cysbond_pred.py -i ejemplos/AF-Q869Z0-F1-model_v4.pdb
```

```
-----
No potential disulfide bond has been detected.
-----

The following disulfide bonds do not meet the requirements but should be reviewed:
+-----+-----+-----+-----+-----+
| Bond | Cys1 | Cys2 | Distance | Dihedral Angle |
+-----+-----+-----+-----+-----+
| 1 | CYS 85 A | CYS 92 A | 2.0316555500030518 | 83.0475536779071 |
+-----+-----+-----+-----+-----+
```

ma-t3vr3-037

- Estructura predicha
- 2 cadenas
- Con parámetros por defecto no se encuentran puentes potenciales.
- Tampoco se encuentran puentes adicionales para chequear.

```
python cysbond_pred.py -i ejemplos/ma-t3vr3-037.pdb
```

```
-----
No potential disulfide bond has been detected.
-----

-----
There are no additional disulphide bonds that need to be checked.
-----
```

AF-Q5I0H9-F1-model_v4

- Estructura predicha (AlphaFold 2)
- 1 cadena
- Con parámetros por defecto se encuentra 1 puente potencial.
- Se encuentran 2 puentes adicionales para chequear.

```
python cysbond_pred.py -i ejemplos/AF-Q5I0H9-F1-model_v4.pdb
```

```
The following potential disulfide bonds have been detected:
+-----+-----+-----+-----+-----+
| Bond | Cys1 | Cys2 | Distance | Dihedral Angle |
+-----+-----+-----+-----+-----+
| 1 | CYS 83 A | CYS 92 A | 2.0319597721099854 | 85.7581536451896 |
+-----+-----+-----+-----+-----+

The following disulfide bonds do not meet the requirements but should be reviewed:
+-----+-----+-----+-----+-----+
| Bond | Cys1 | Cys2 | Distance | Dihedral Angle |
+-----+-----+-----+-----+-----+
| 1 | CYS 180 A | CYS 183 A | 2.065502643585205 | 83.35509215162566 |
| 2 | CYS 454 A | CYS 463 A | 2.0274598598480225 | 82.50338110024451 |
+-----+-----+-----+-----+-----+
```