

Bioinformática estructural

TAREA 1 *Foldit*

Carina Paola Cornejo Páramo

Claudia Saraí Reyes Ávila

INTRODUCCIÓN

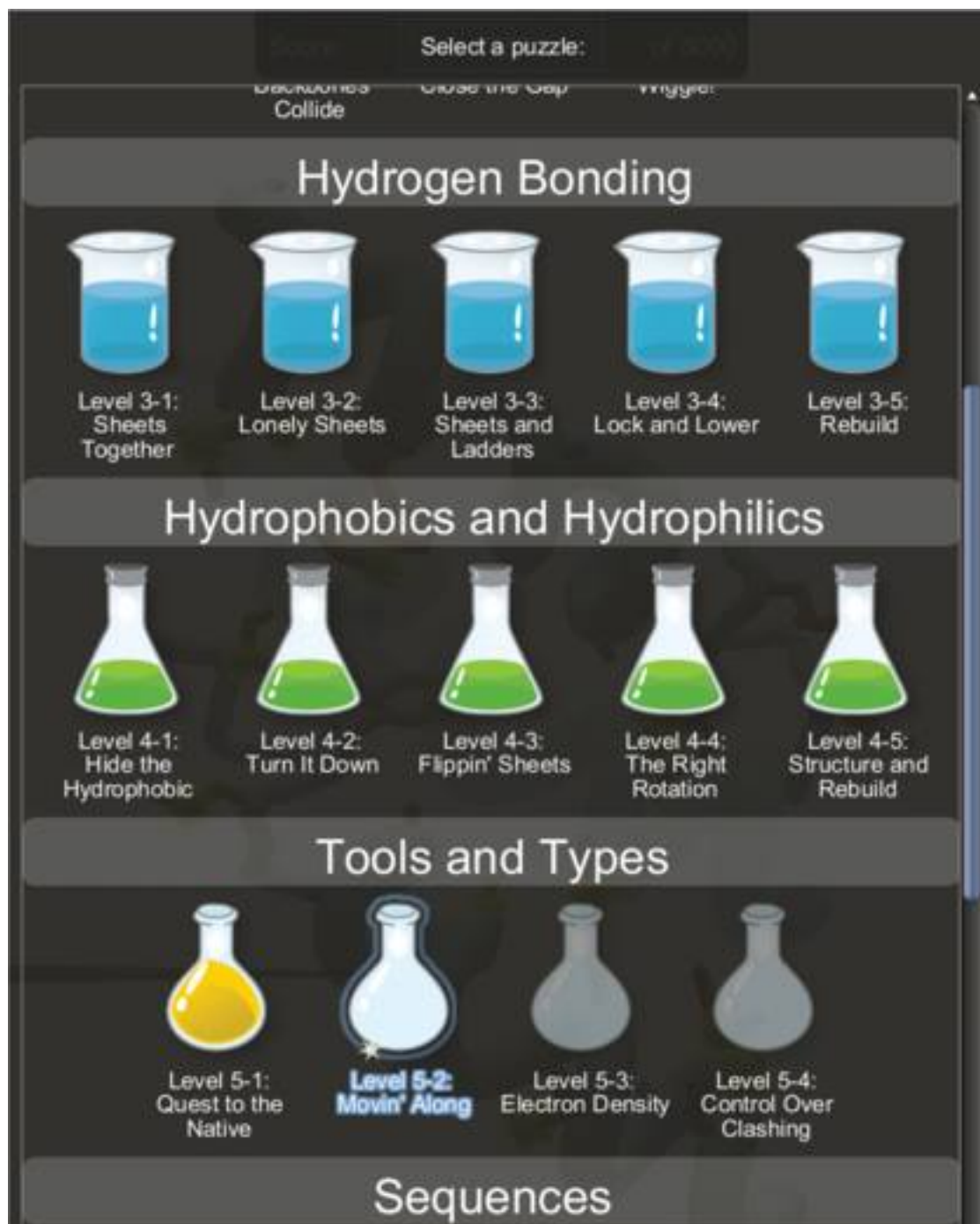
Los ácidos nucleicos y las proteínas son las dos macromoléculas biológicas más importante y su función está ligada a su estructura. Estas macromoléculas naturales deben plegarse, es decir, deben tomar una determinada conformación tridimensional relativamente estable para desempeñar su función biológica.

La estructura de ácidos nucleicos y proteínas puede analizarse de forma jerárquica en 4 niveles. En la estructura de las proteínas se pueden considerar cuatro niveles de organización: primario, secundario, terciario y cuaternario. Cada uno de los cuales resalta un aspecto diferente y depende de distintos tipos de interacciones. Mientras la estructura primaria es simplemente la secuencia lineal de aminoácidos de una cadena polipeptídica las demás establecen su organización tridimensional de péptidos o conjunto de ellos.

DESARROLLO

- 1) Creen un repositorio *git* en tepeu, por ejemplo en su home/algoritmos3D, para ir añadiendo ahí las tareas de los 4 días y sus respectivos informes. El repositorio fue creado en GitHub, los informes se encontrarán en la siguiente liga: https://github.com/carinapaola/Bioinformatica_estructural
- 2) Que completen los *puzzles* de *Foldit* al menos hasta el nivel 4-5, y por el camino vayan tomando capturas de pantalla que reflejen algunos de los conceptos teóricos que hablamos por la mañana:

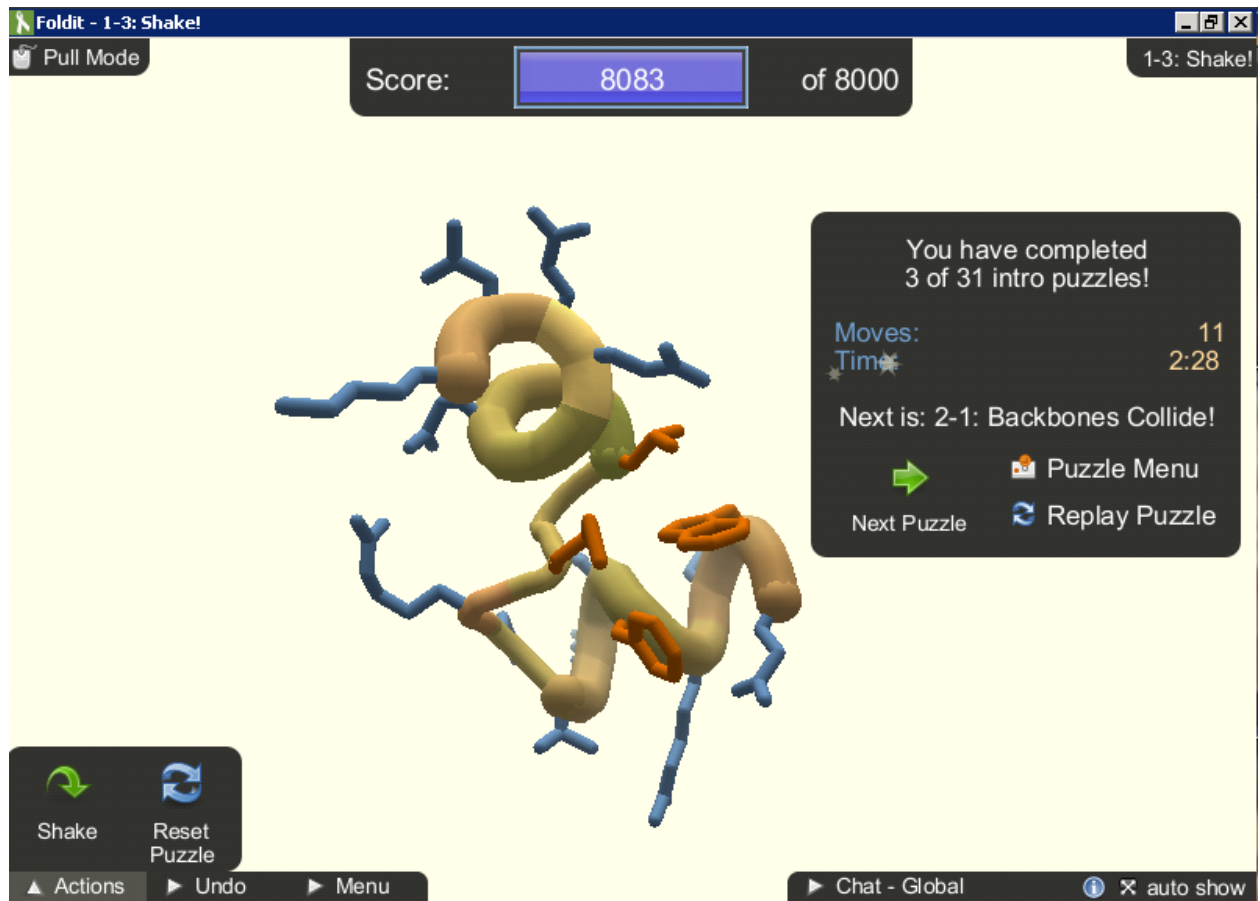
En el juego de Foldit hemos llegado hasta la sección Tools and Types del nivel 5-1, a continuación explicaremos en general algunos conceptos teóricos que vimos en clase y reafirmamos con el juego.



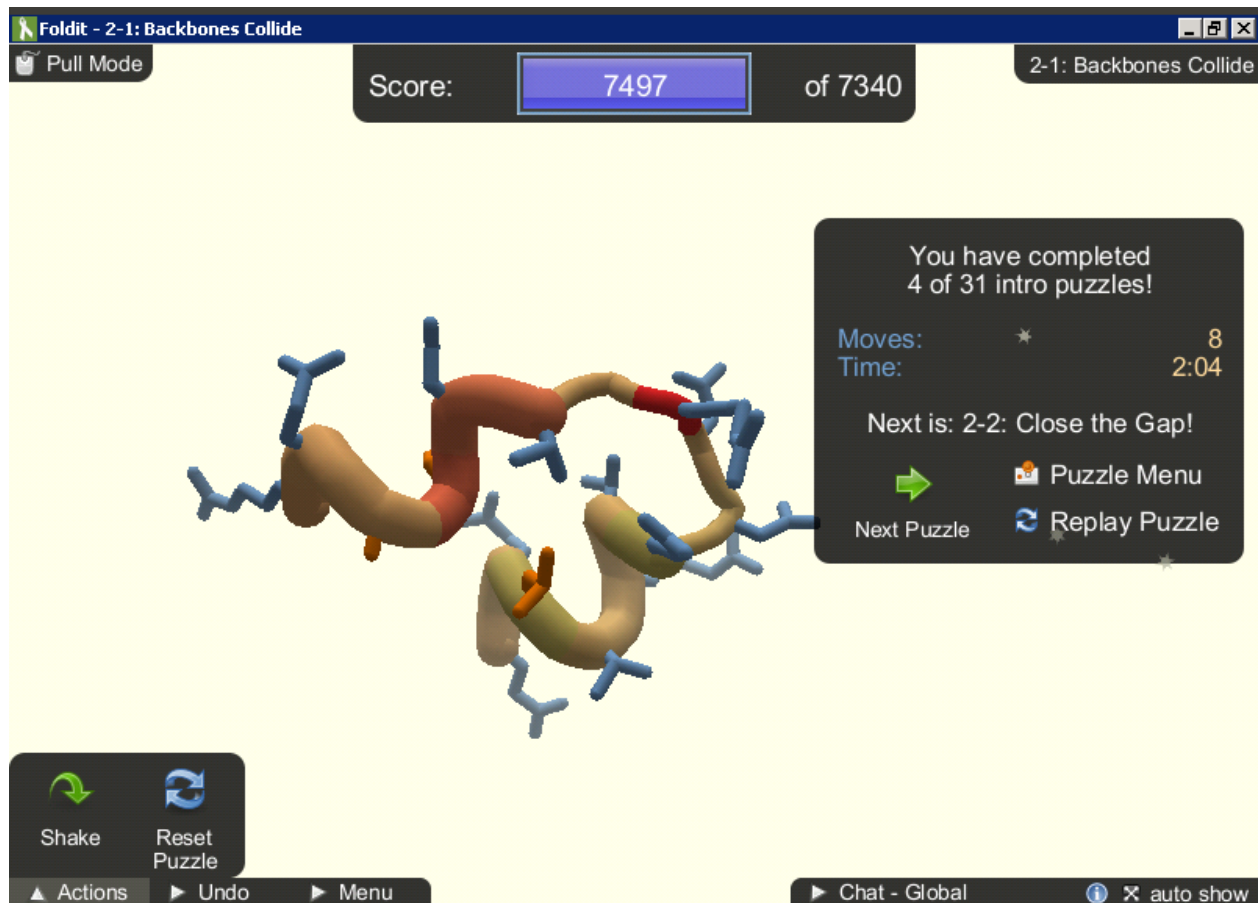
Nivel 1-1: en este nivel encontramos un par de cadenas laterales que estaban chocando en la conformación inicial de la proteína, al separar las cadenas laterales adquiere una conformación correcta.



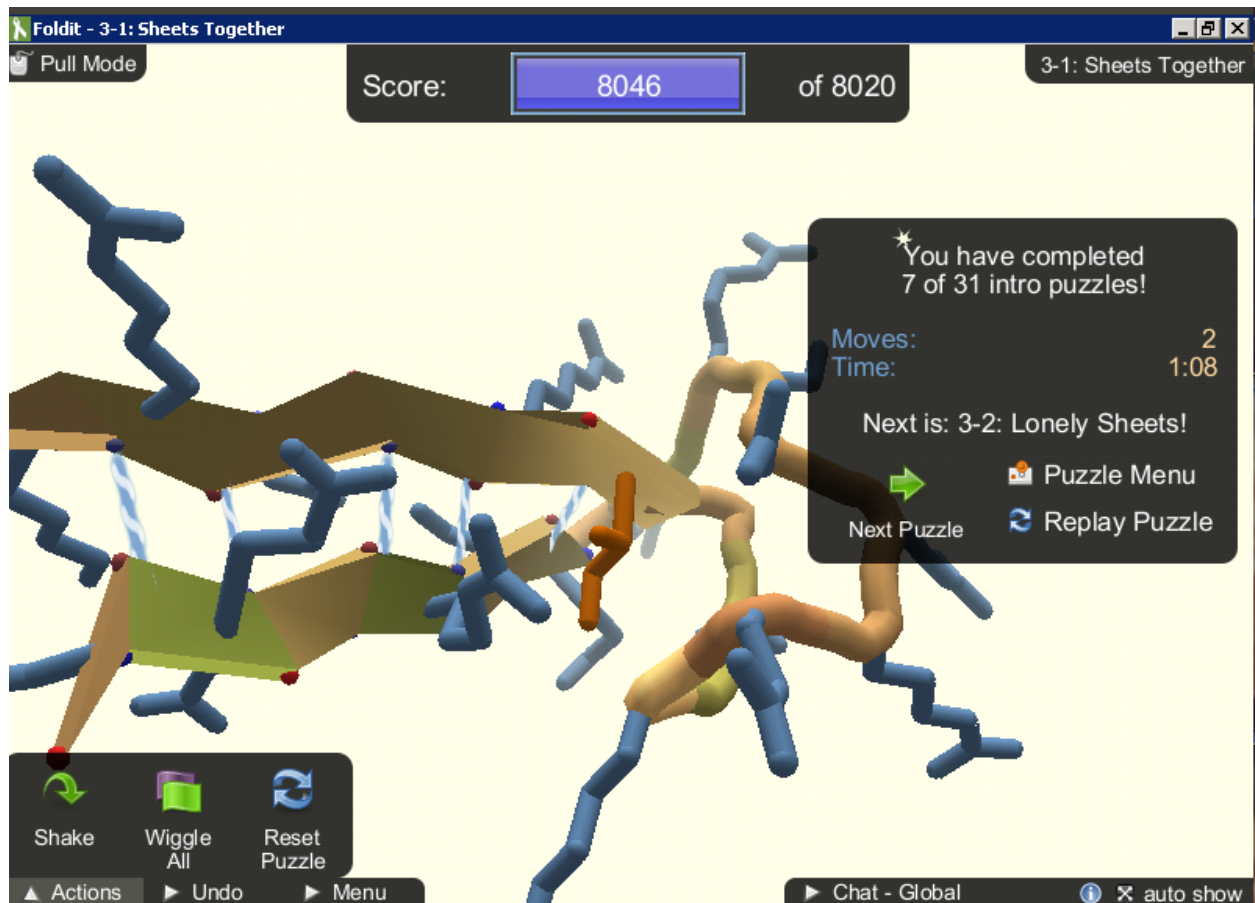
Nivel 1-3: En la siguiente imagen observamos una proteína pequeña con un par de hélices alfa cortas conectadas por una región linker. En anaranjado se muestran residuos hidrofóbicos hacia la parte central de la proteína, en azul los residuos hidrofílicos. Los puentes de hidrógeno que conectan a los aminoácidos de las hélices alfa no se aprecian en esta imagen.



Nivel 2-1: en este nivel también observamos una proteína pequeña, con dos hélices alfa conectadas por un *linker*; observamos también sus cadenas laterales. En este nivel encontramos que las hélices alfa estaban muy cerca entre sí y esto hacía que chocaran sus cadenas laterales, separando las hélices alfa y agitando un poco la proteína (herramienta *SHAKE*) conseguimos el *score* necesario para este nivel.



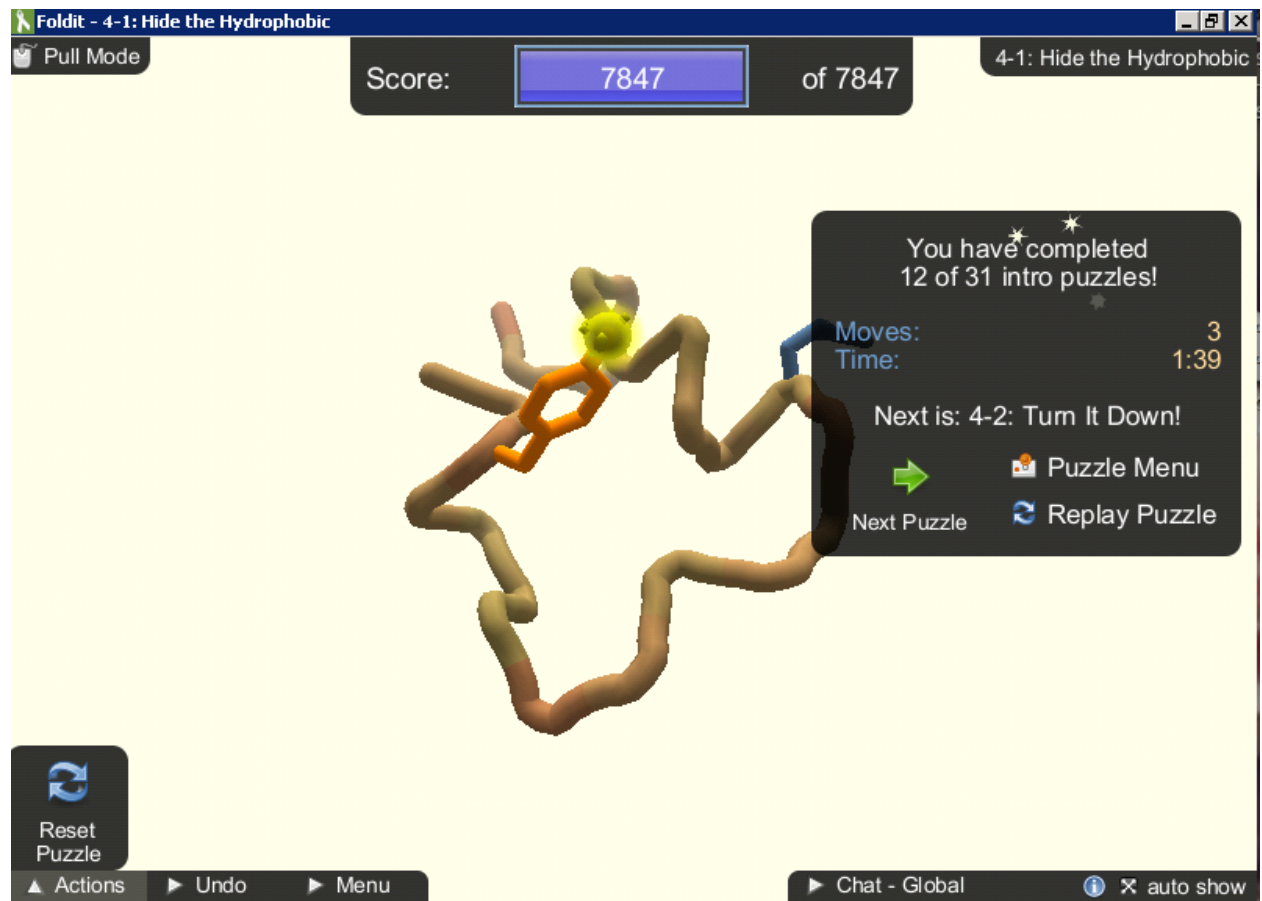
Nivel 3-1: En este nivel encontramos hojas plegadas beta, estructuras en forma de zig-zag paralelas entre sí, éstas están unidas por puentes de hidrógeno, representados por barras de azules. Estos enlaces deben formarse para alcanzar la conformación correcta. Las hojas están conectadas por una parte de la proteína a manera de *linker*. Se aprecian también los grupos R tanto en el *linker* como en las hojas plegadas.



En el nivel 3-4 observamos una proteína grande con hojas plegadas beta y hélices alfa, para alcanzar el score para pasar el nivel se requiere formar los puentes de hidrógeno que unen a las hojas.



Nivel 4-1: en este nivel encontramos un residuo hidrofóbico (en color naranja) que estaba muy expuesto en su configuración inicial, basta con enterrar este residuo para mejorar el score, así como dar más espacio al grupo R en azul.

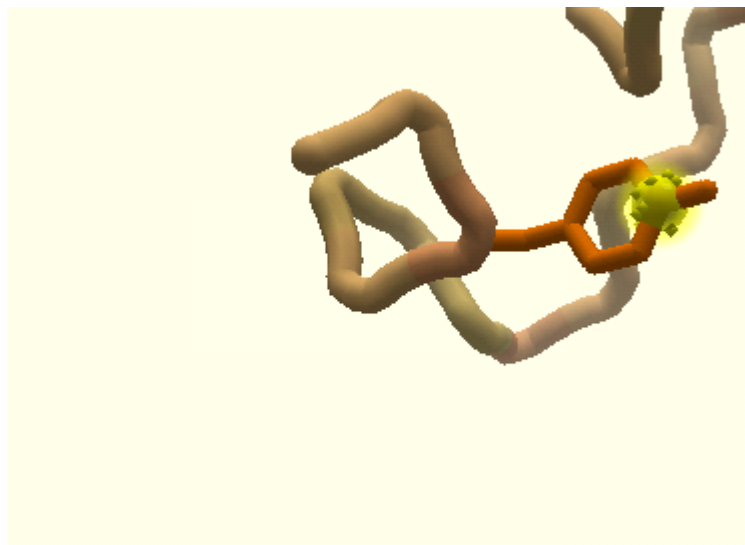


El último nivel que pasamos es el 5-1, en este encontramos una proteína con una hélice alfa, una hoja plegada, regiones que las conectan, y muchos aminoácidos hidrofóbicos.

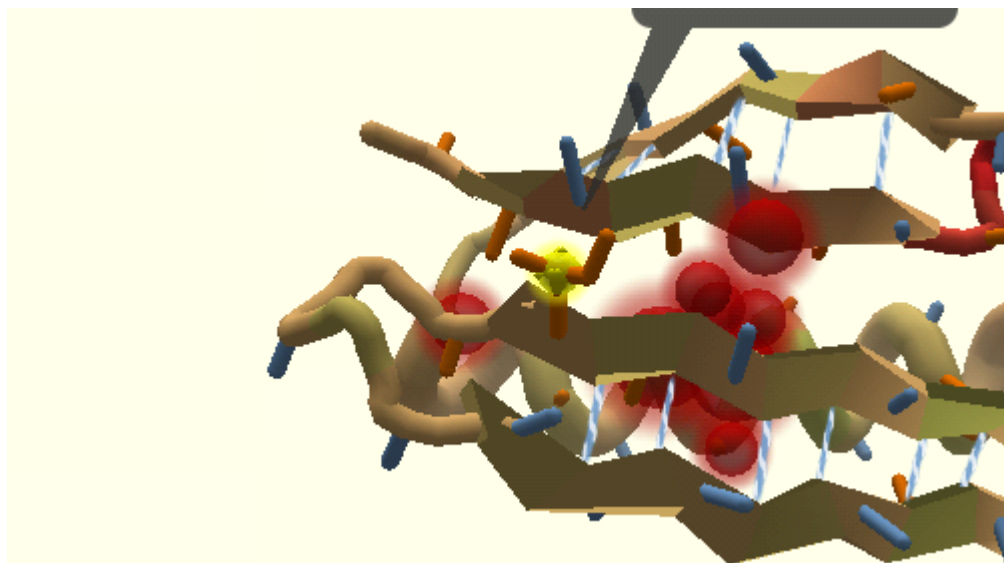


2.1) Ejemplo de aminoácido con cadena lateral aromática

*Nivel 4-1



2.2) Ejemplo de aminoácido con cadena lateral chica



*Nivel 4-3

2.3) Ejemplo de giro en torno a los ángulos phi/psi de un residuo seleccionado, que pasa cuando si sus vecinos tienen cadenas laterales voluminosas?

En las siguientes imágenes podemos observar el giro de un residuo aromático en la parte superior de la proteína.



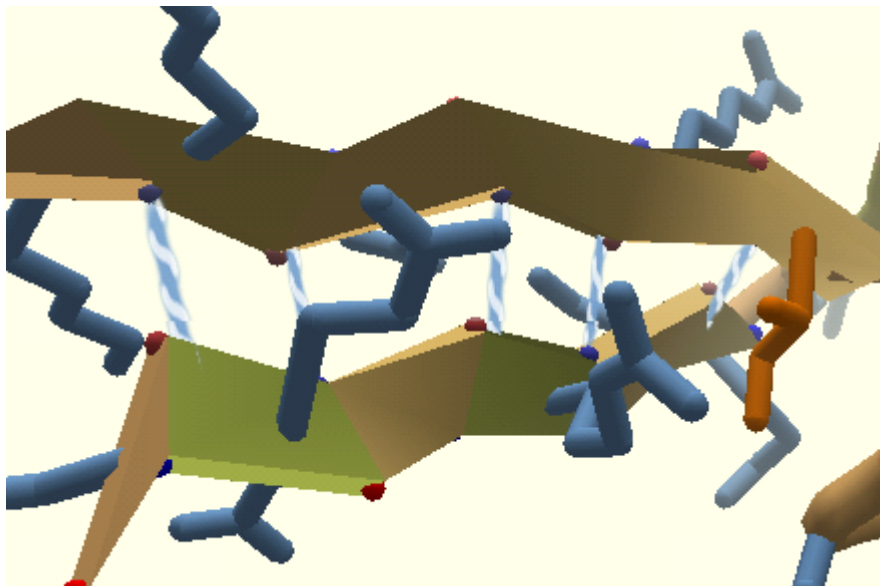


El hecho de que una cadena lateral sea muy voluminosa limita el número de conformaciones que son estables, pues en algunos ángulos esta cadena lateral puede chocar con las que están cerca.

2.4) Ejemplo de puentes de hidrógeno entre residuos de una alfa-hélice y entre hojas de una lámina beta. Desde el punto de vista algorítmico, cuál de los estados de estructura secundaria les parece más difícil de programar?

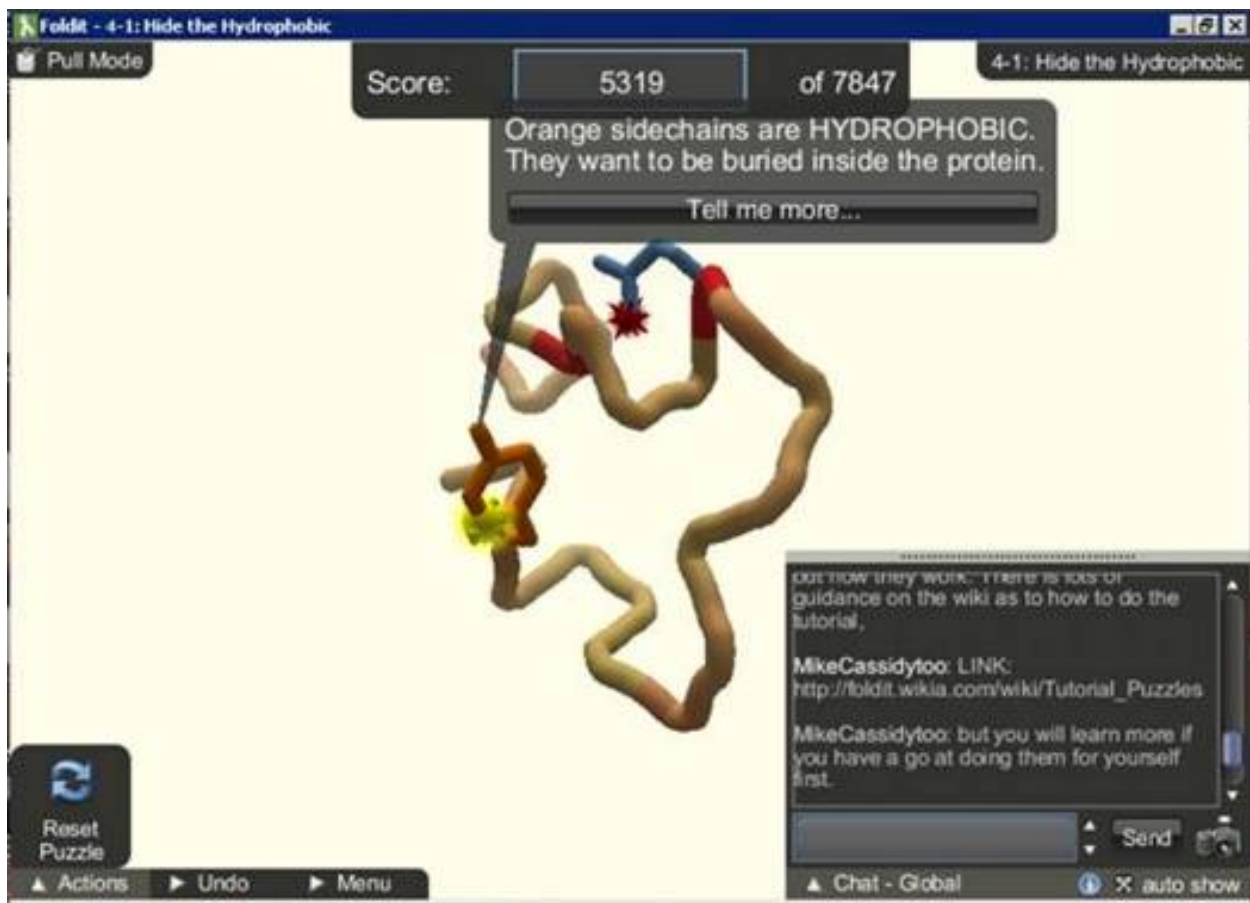
Las hélices alfa dan giros cada determinado número de aminoácidos, además de que se forman entre aminoácidos continuos, y aminoácidos como la prolina no se observan en ellas o muy raramente, por el contrario las hojas plegadas pueden formarse entre partes distantes de la proteína, además de que los posibles puentes de hidrógeno entre ellas son más, por lo que creemos que éstas son más difíciles de programar.

*Nivel 3-1



2.5) Ejemplo de residuo hidrofóbico expuesto y luego correctamente “enterrado” tras operaciones con los vecinos.

*Nivel 4-1





2.6) Ejemplo de conformaciones distintas con puntuaciones similares, para hacer patente el problema de evaluar lo correcto de una conformación.





2.7) De acuerdo con http://eead-csic-compbio.github.io/bioinformatica_estructural/node17.html calcula el tiempo que llevaría explorar todas las conformaciones posibles de uno de los péptidos o proteínas que utilicen en los puzzles.

La fórmula para calcular el tiempo que tardaría probar todas las conformaciones posibles es :

$$t \cdot X^n$$

donde t es el tiempo que tarda cada cambio de estado. X es el número de estados posibles en los que puede estar un aminoácido de la proteína. n representa el número de aminoácidos de la proteína.

Suponiendo que $X=10$ y que $t=10^{13}$ segundos

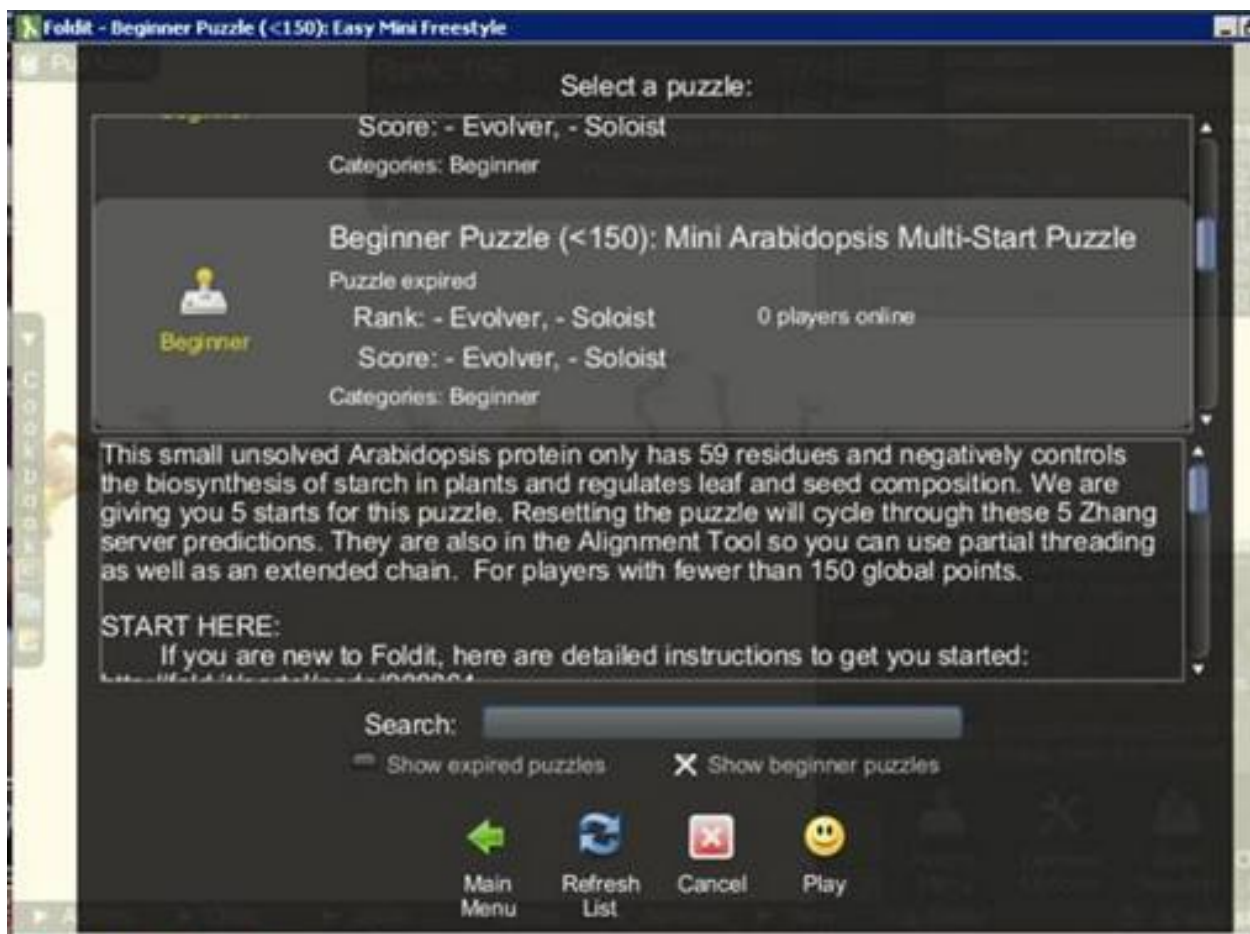
En un ejemplo de foldit de una proteína con 59 aminoácidos:

$$(10^{13})^{(10)}(59)=1 \cdot 10^{72} \text{ segundos}$$

que corresponde con $2.7 \cdot 10^{68}$ horas

Obviamente una célula no podría probar todas las conformaciones, por lo que hay variables que aun no conocemos del plegamiento de las proteínas pues nuestros modelos están muy lejos de imitar el plegamiento de proteínas en una célula, además de que incluso probar un subconjunto de las posibles conformaciones requeriría mucho tiempo de cómputo.

Proteína ejemplo:



Foldit - Beginner Puzzle (<150): Mini Arabidopsis Multi-Start Puzzle

Pull Mode

Rank: 154 Score: 0

Soloist Beginner Puzzle (<150) Multi-Start Puzzle

Wiki Help for this Puzzle

Puzzle expired

► Conditions satisfied: 1 of 1, No bonuses.

Competition

Soloist Competition

#	Player Name	Current	Best
1	Virgo	-	8734
2	theoracle94	-	8672
3	benrh	-	8669
4	Sam256	-	8651
5	LuapNor	-	8613
6	BackBuffer	-	8612
7	hvalnic2	-	8521

Shake Mutate Wiggle All Wiggle Backbone Wiggle Sidechains Help Glossary

Freeze Protein Remove Bands Disable Bands Reset Structures Auto Structures Reset Puzzle Show Alignment

Chat - Global

Send

auto show

auto show

Actions Undo Social Modes Behavior View Menu

out how they work. There is lots of guidance on the wiki as to how to do the tutorial.

likeCassidytoo: LINK: http://foldit.wikia.com/wiki/Tutorial_Puzzles

likeCassidytoo: but you will learn more if you have a go at doing them for yourself first.