



# Instituto Politécnico Nacional Escuela Superior de Cómputo



## Bioinformatics

### Lab Session 8. Protein-protein Docking

González Bocio Erik Alexander 2020630163

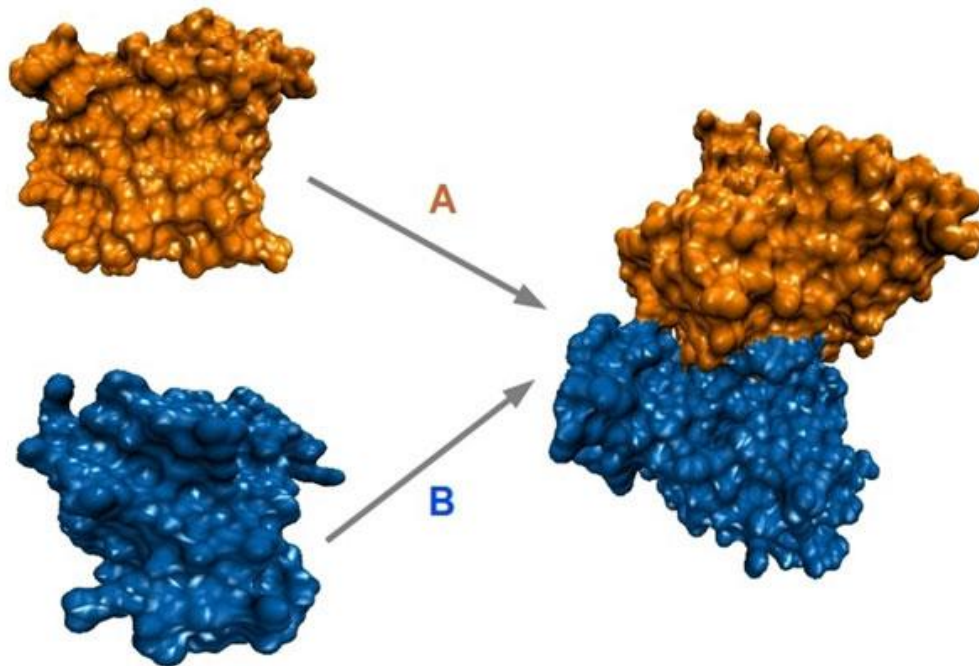
Jorge Luis Rosas Trigueros

Fecha de realización: 25 de abril de 2022

Fecha de entrega: 9 de mayo de 2022

## Marco Teórico:

Predecir la estructura de los complejos proteína-proteína utilizando enfoques de acoplamiento es un problema difícil cuyos principales desafíos incluyen la identificación de soluciones correctas y el tratamiento adecuado de la flexibilidad molecular y los cambios conformacionales. La estructura tridimensional de un complejo proteína-proteína, generalmente, es más difícil de determinar experimentalmente que la estructura de una proteína individual. Las técnicas computacionales adecuadas para modelar las interacciones de proteínas son importantes debido al creciente número de estructuras proteicas conocidas, particularmente en el contexto de la genómica estructural. El Docking ofrece herramientas para estudios fundamentales de interacciones de proteínas y proporciona una base estructural para el diseño de fármacos. El acoplamiento proteína-proteína es la predicción de la estructura del complejo, dadas las estructuras de las proteínas individuales. Originalmente, se consideraban para el acoplamiento estructuras proteicas en su mayoría de alta resolución, determinadas experimentalmente (principalmente mediante cristalografía de rayos X). Sin embargo, más recientemente, el enfoque se ha desplazado hacia estructuras modeladas de menor resolución. Los enfoques de acoplamiento tienen que lidiar con los cambios conformacionales entre las estructuras unidas y no unidas, así como con las imprecisiones de las estructuras modeladas que interactúan, a menudo en un modo de alto rendimiento necesario para el modelado de grandes redes de interacciones de proteínas. El creciente número de desarrolladores de acoplamiento participa en las evaluaciones comunitarias de metodologías predictivas. El desarrollo de enfoques de acoplamiento más potentes y adecuados se ve facilitado por la rápida expansión de los recursos de información y datos, el aumento de las capacidades computacionales y una comprensión más profunda de los principios fundamentales de las interacciones de proteínas.



*Fig. 1. Ejemplo de docking proteína-proteína.*

La flexibilidad se puede abordar en varios niveles: implícitamente, suavizando las superficies de las proteínas o permitiendo cierto grado de interpenetración (acoplamiento suave) o realizando múltiples ejecuciones de acoplamiento desde varias conformaciones (acoplamiento cruzado o conjunto); o explícitamente, al permitir la flexibilidad de la cadena lateral y/o la red troncal. Aunque se han logrado mejoras significativas en el modelado de cadenas laterales, todavía se están desarrollando métodos para la inclusión explícita de la flexibilidad de la red troncal en el acoplamiento. Han surgido algunos enfoques novedosos que involucran grados colectivos de movimiento, representaciones multicopia y acoplamiento multicuerpo, lo que debería permitir modelar cambios conformacionales más grandes.

La predicción de alta resolución del acoplamiento proteína-proteína ahora puede crear estructuras con una precisión de nivel atómico. Este progreso surge tanto de las mejoras en el muestreo rápido de conformaciones como de la mayor precisión de los cálculos de energía libre vinculante. En consecuencia, la calidad de los modelos presentados al desafío de predicción ciega CAPRI (Evaluación crítica de interacciones predichas) ha aumentado constantemente, incluidos los complejos predichos a partir de estructuras de homología de un socio de unión y complejos con precisión atómica en la interfaz. Al explotar la información experimental, el acoplamiento ha creado estructuras modelo para aplicaciones reales, incluso cuando se enfrenta a desafíos como estructuras de monómeros inciertas y estructuras de monómeros en movimiento. Queda trabajo por hacer para acoplar proteínas grandes o flexibles, clasificar los modelos de manera consistente y producir modelos lo suficientemente precisos para permitir el diseño computacional de afinidades o especificidades más altas.

### Material y Equipo:

- Equipo de Computo
- Red de Internet
- Correo Electrónico

### Desarrollo de la Práctica:

En esta práctica se empezó a trabajar con el acoplamiento proteína-proteína o protein-protein docking. Para comenzar la práctica se llevó a cabo un acoplamiento de proteína-proteína usando las siguientes entradas del PDB:

1AMY (chain A) y 2IWT (chain B).

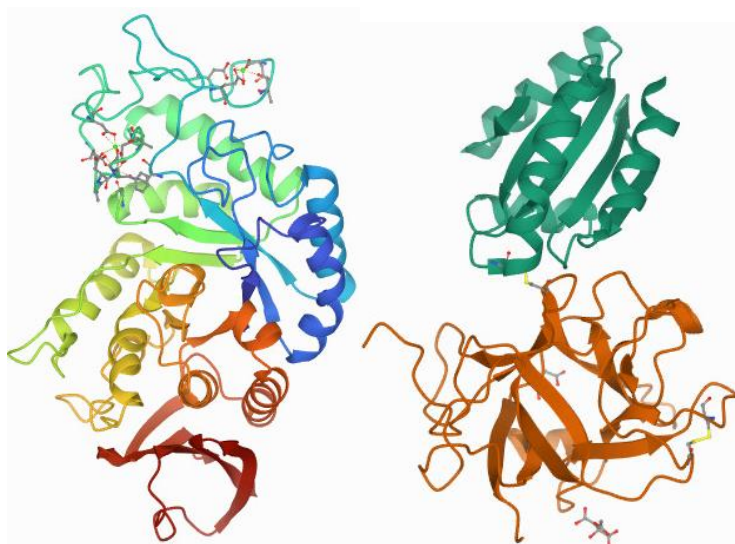


Fig. 2. Estructura de 1AMY y 2IWT, respectivamente

Luego, el acoplamiento se llevo a cabo en diferentes servidores, el servidor con el cual se trabajaría sería el marcado por el ultimo número de nuestra boleta institucional, el cual correspondía a <http://hdock.phys.hust.edu.cn/>

Fig. 3. Servidor con el que se trabajó, correspondiente a HDOCK SERVER

Entonces se colocaron las dos proteínas con las que se debía trabajar y su cadena correspondiente, al igual que en practicas anteriores esta se llevo un tiempo para poder finalizar el trabajo, por lo que es necesario colocar un correo electrónico con el cual puedan comunicarse contigo al finalizar el trabajo, de igual manera se proporciona una liga con el que puedes entrar.

Al final el trabajo se obtuvo una lista de 10 posibles modelos, entre ellos están los mostrados en la siguiente figura:

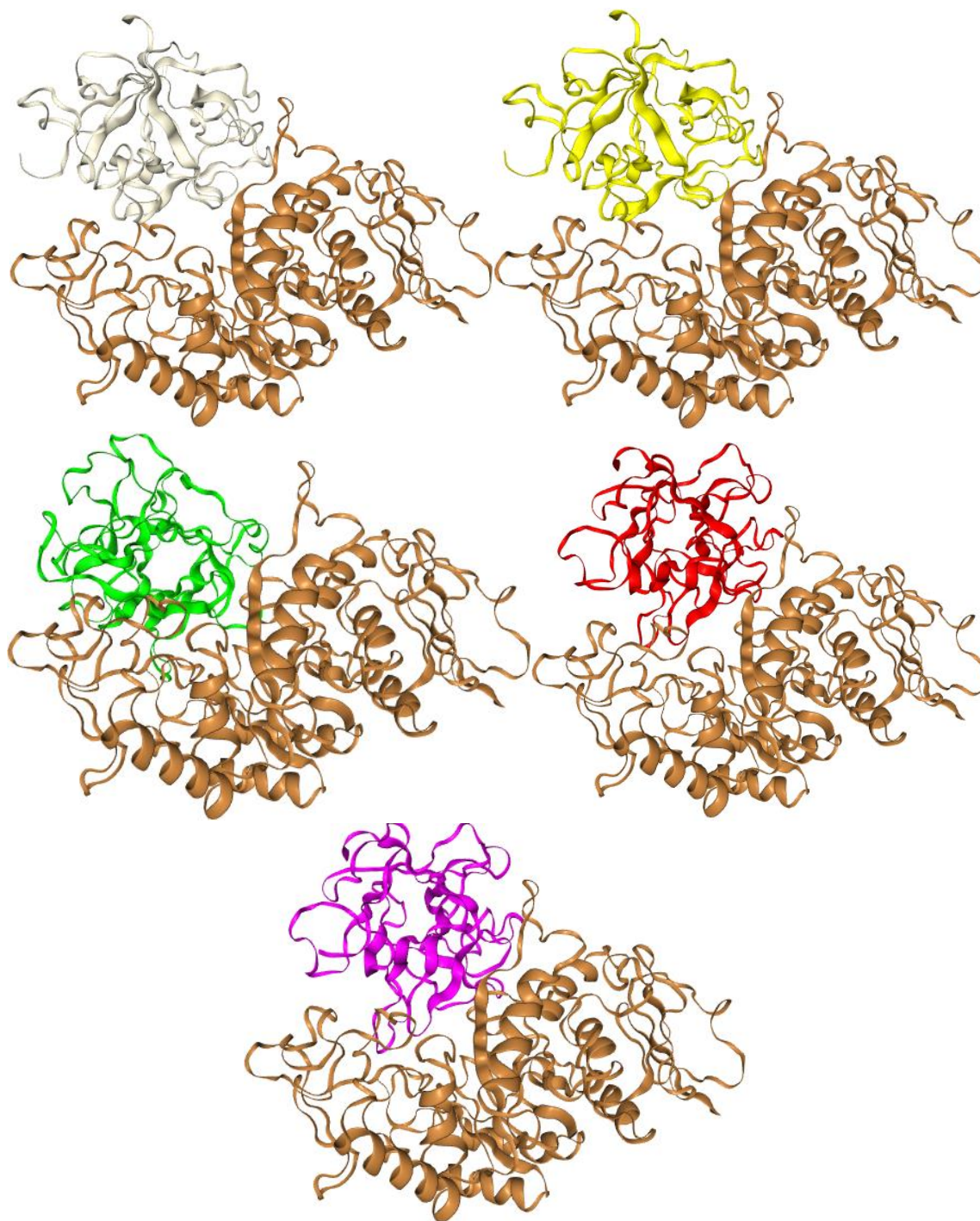


Fig. 4. Resultados de la alineación proteína-péptido, podemos ver desde el modelo 0(blanco) hasta el 4(morado)

De igual forma se nos muestran dos tablas que no ayudan a interpretar los resultados

Molecule	PDB ID	Chain ID	Align_length	Coverage	Seq_ID (%)
Receptor	<a href="#">1AVA</a>	A	403	1.000	100.0
Ligand	<a href="#">1AVA</a>	C	181	0.995	99.4

Fig. 5. Información de plantilla compleja



Rank	1	2	3	4	5	6	7	8	9	10
Docking Score	-375.48	-286.64	-277.75	-258.19	-257.41	-255.63	-251.22	-243.28	-241.36	-239.16
Ligand rmsd (Å)	0.50	30.61	8.06	12.13	27.85	32.09	16.44	27.21	27.93	22.63
Interface residues	<a href="#">model 1</a>	<a href="#">model 2</a>	<a href="#">model 3</a>	<a href="#">model 4</a>	<a href="#">model 5</a>	<a href="#">model 6</a>	<a href="#">model 7</a>	<a href="#">model 8</a>	<a href="#">model 9</a>	<a href="#">model 10</a>

Fig. 6. Resumen de los 10 mejores modelos

Este acoplamiento proteína-proteína está basado en un algoritmo híbrido de modelado basado en plantillas y acoplamiento libre ab initio, el primero como lo dice su nombre pues se basa en varias plantillas para llevar a cabo el acoplamiento, mientras que el segundo se centra en la observación de que las proteínas usualmente adoptan pliegues similares sin importar su similitud de secuencia.

Ahora seguimos con el segundo paso en la práctica, el cual era realizar un acoplamiento de proteínas/péptidos utilizando

6DDF chain D y el péptido:

>beta\_endorphine

YGGFMTSEKSQTPLVTLFKNAIIKKNAYKKGE

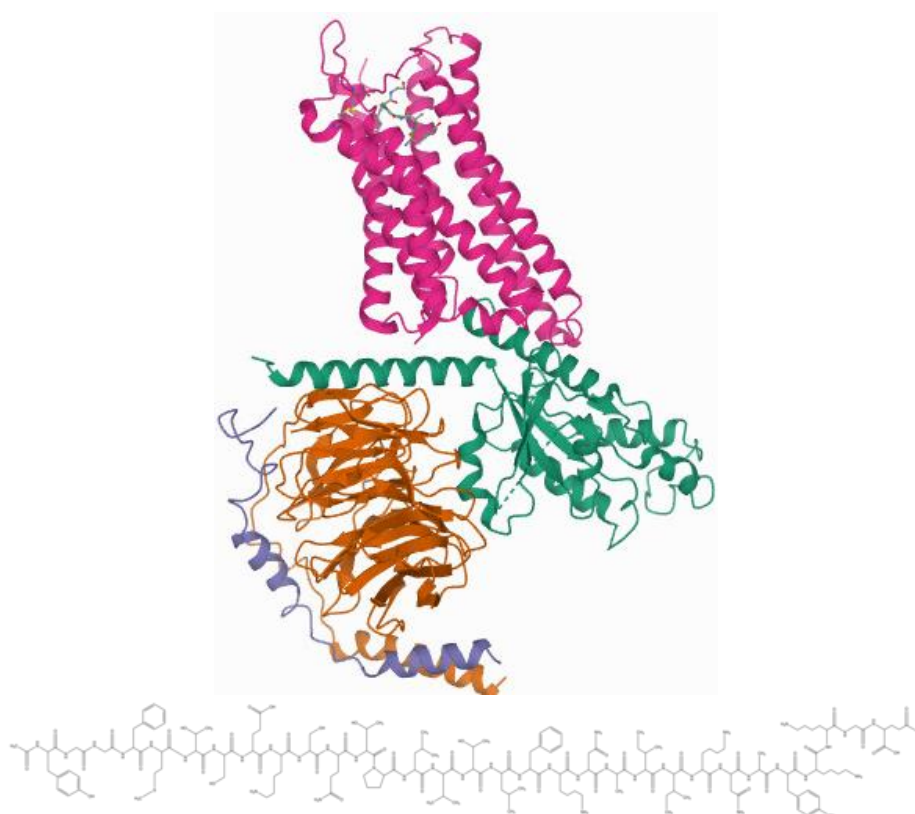


Fig. 7. Proteína 6DDF y péptido seleccionado

Al igual que en la actividad anterior el servidor que usaríamos variaría dependiendo de nuestra boleta, en este caso el servidor fue el de la página <http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=PEPDOCK>

## GalaxyPepDock

Protein-peptide docking based on interaction similarity

### User Information

Job name

E-mail address (Optional)

### Protein-peptide Docking

PROTEIN structure  
 (≤ 900 AA)
 
 No se eligió ningún archivo  
 (allowed file extensions: pdb, txt)

PEPTIDE sequence  
 (≤ 30 AA)
 
 No se eligió ningún archivo  
 (allowed file extensions: fa, fasta, txt, seq)

### Submit

### Help

- Protein structure: [PDB Format](#)
- Peptide sequence: [FASTA format](#)
- E-mail: [Average run time is ~12h. If e-mail address is given, the server sends notifications automatically. If not, the user has to bookmark the report page address.](#)
- [More Information](#)

### Example

- Protein structure: [protein.pdb](#)
- Peptide sequence: [peptide.fa](#)
- Report: [\[View\]](#)

### Software

- You can download a standalone

Fig. 8. Interfaz del servidor GalaxyPepDock

Luego de llenar el formulario el servidor comienza a trabajar, esto puede llevar algun tiempo, en este caso fue en 2 horas y el resultado fueron 10 modelos, los cuales son los siguientes:

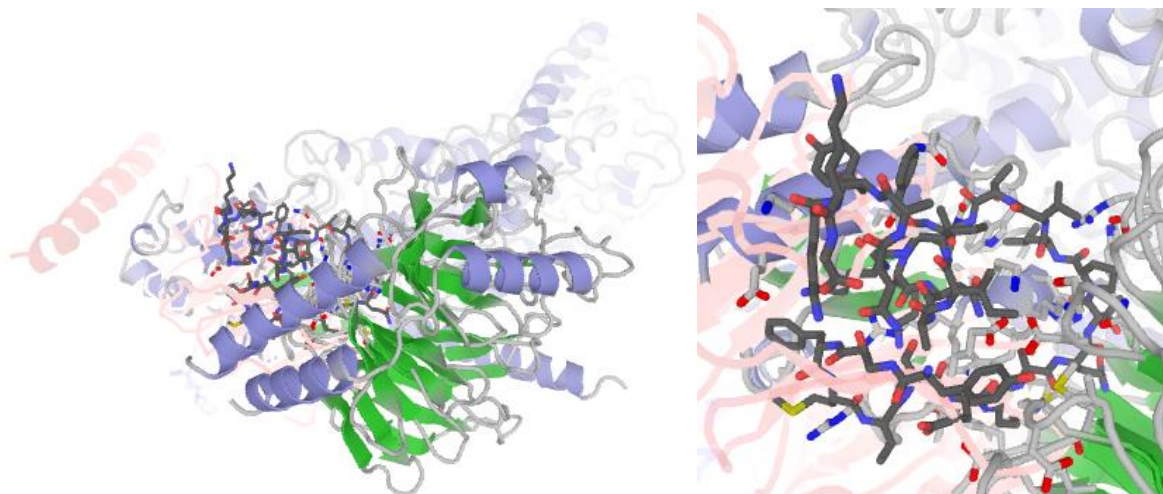


Fig. 9. Estructura compleja proteína-péptido predicha (Modelo 1)

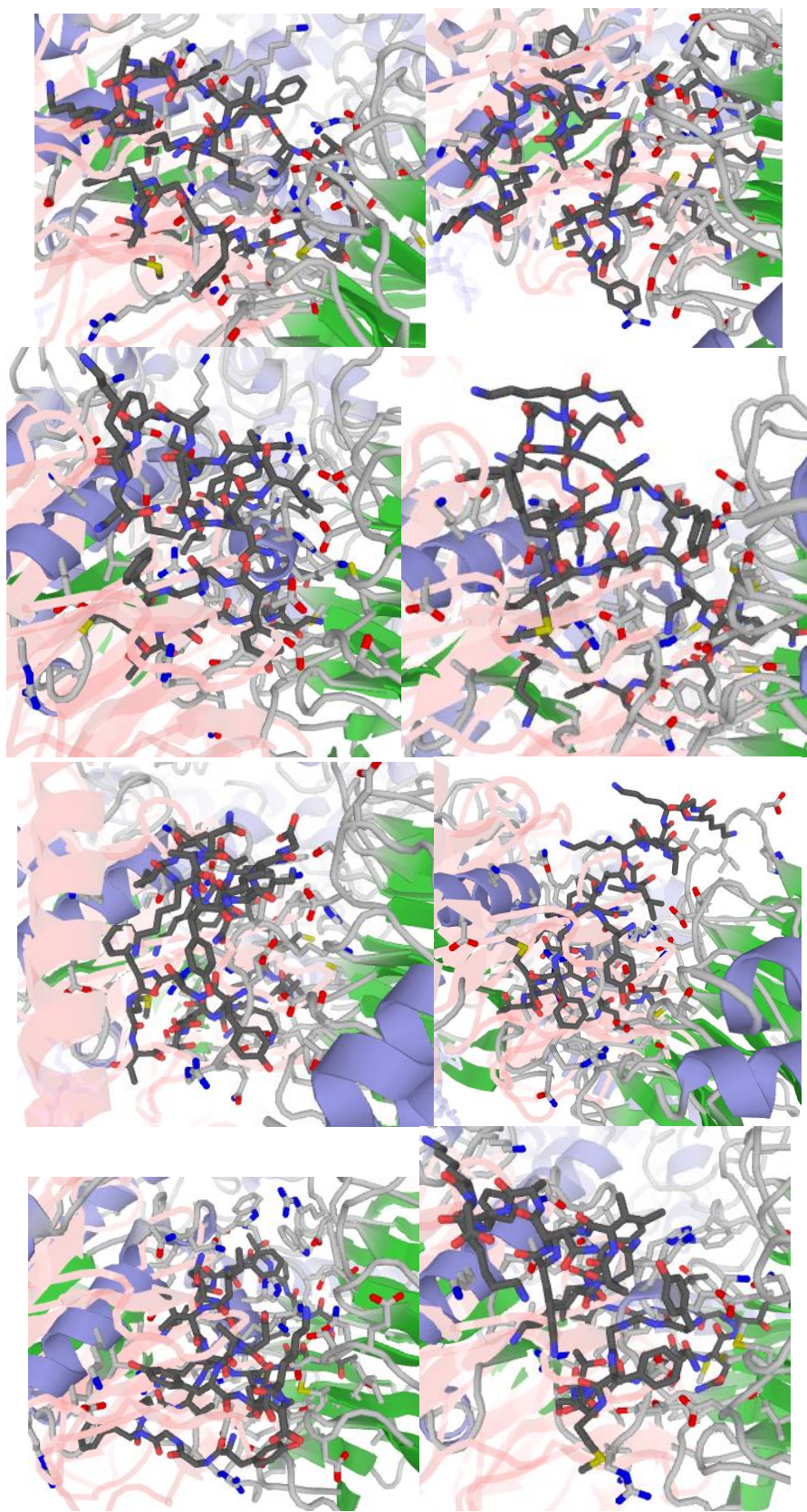


Fig. 10. Modelos 2 al 10, respectivamente.



Model	Protein template	Peptide template	Protein structure similarity (TM-score)	Interaction similarity score	Estimated accuracy
1	1XHM_A	1XHM_C	0.980	148.0	0.938
2	1XHM_A	1XHM_C	0.980	148.0	0.938
3	1XHM_A	1XHM_C	0.980	148.0	0.938
4	1XHM_A	1XHM_C	0.980	148.0	0.938
5	1XHM_A	1XHM_C	0.980	148.0	0.938
6	1XHM_A	1XHM_C	0.980	148.0	0.938
7	1XHM_A	1XHM_C	0.980	148.0	0.938
8	1XHM_A	1XHM_C	0.980	148.0	0.938
9	1XHM_A	1XHM_C	0.980	148.0	0.938
10	1XHM_A	1XHM_C	0.980	148.0	0.938

Fig. 11. Información adicional por modelo

Este segundo servidor lleva a cabo el acoplamiento proteína-péptido basado en la similitud de interacción, realiza un acoplamiento basado en la similitud al encontrar plantillas de la base de datos de estructuras determinadas experimentalmente y modelos de construcción utilizando la optimización basada en la energía que permite la flexibilidad estructural. Por lo tanto, el servidor puede modelar eficazmente las diferencias estructurales entre la plantilla y los complejos proteína-péptido objetivo.

## Conclusiones y Recomendaciones:

Con esta práctica pudimos observar dos tipos de acoplamiento, como lo son el proteína-proteína y el proteína-péptido, estos dos en dos servidores distintos los cuales hacen uso de diferentes algoritmos, por lo que también fue posible ver como es que trabajaban estos dos algoritmos con ayuda de los resultados.

Me pareció una buena práctica, fue un poco más de analizar los algoritmos con los que trabaja cada servidor y sus resultados pero fue interactiva en el punto de solo trabajar con ciertas cadenas de cada proteína.

## Bibliografía:

- Bonvin, A. (2006). Flexible protein-protein docking. *Current Opinion In Structural Biology*, 16(2), 194-200. <https://doi.org/10.1016/j.sbi.2006.02.002>
- Camacho, C., & Vajda, S. (2002). Protein-protein association kinetics and protein docking. *Current Opinion In Structural Biology*, 12(1), 36-40. [https://doi.org/10.1016/s0959-440x\(02\)00286-5](https://doi.org/10.1016/s0959-440x(02)00286-5)
- Gray, J. (2006). High-resolution protein-protein docking. *Current Opinion In Structural Biology*, 16(2), 183-193. <https://doi.org/10.1016/j.sbi.2006.03.003>
- Vakser, I. (2014). Protein-Protein Docking: From Interaction to Interactome. *Biophysical Journal*, 107(8), 1785-1793. <https://doi.org/10.1016/j.bpj.2014.08.033>