



UNIVERSIDAD AUTÓNOMA DE MADRID

## 33082 - BIOINFORMÁTICA ESTRUCTURAL

### Información de la asignatura

**Código - Nombre:** 33082 - BIOINFORMÁTICA ESTRUCTURAL

**Titulación:** 702 - Máster en Bioinformática y Biología Computacional

**Centro:** 106 - Facultad de Medicina

**Curso Académico:** 2023/24

### 1. Detalles de la asignatura

#### 1.1. Materia

Análisis de secuencias

#### 1.2. Carácter

Obligatoria

#### 1.3. Nivel

Máster (MECES 3)

#### 1.4. Curso

1

#### 1.5. Semestre

Primer semestre

#### 1.6. Número de créditos ECTS

3.0

#### 1.7. Idioma

Español y/o Inglés en función de las necesidades de los alumnos.

#### 1.8. Requisitos previos

-

#### 1.9. Recomendaciones

- Conocimientos generales de estructura de ácidos nucleicos y proteínas.
- Conocimientos básicos de evolución molecular.

Estos conocimientos se deben haber adquirido en cualquiera de los grados en ciencias de la vida/salud o en la asignatura “Complementos Docentes Bioquímica y Biología Molecular”

|                                       |  |                |            |     |
|---------------------------------------|--|----------------|------------|-----|
| <b>Código Seguro de Verificación:</b> |  | <b>Fecha:</b>  | 22/08/2023 | 1/7 |
| <b>Firmado por:</b>                   | <i>Esta guía docente no estará firmada mediante CSV hasta el cierre de actas</i> |                |            |     |
| <b>Url de Verificación:</b>           |  | <b>Página:</b> | 1/7        |     |

Para garantizar la asimilación de los contenidos y la adquisición de habilidades se recomienda:

- Lectura crítica de los textos de la bibliografía.
- Uso del material electrónico disponible on-line en la plataforma Moodle de Posgrado de la UAM (<http://posgrado.uam.es>).
- Revisión de material complementario en la red.

Es recomendable disponer de un dominio de inglés para leer la bibliografía de consulta.

Asimismo, se requiere:

- Trabajo individual antes de las sesiones teóricas y prácticas: Lectura del material a cubrir con el objetivo de familiarizarse con los conceptos que se van a manejar y los problemas a resolver.
- Trabajo individual posterior a las sesiones:
  - Revisión de las notas tomadas en clase.
  - Consulta de la bibliografía.
  - Realización de ejercicios.
  - Evaluación por pares de los ejercicios de otros alumnos.

### 1.10. Requisitos mínimos de asistencia

Tal y como marca la normativa de la UAM es obligatoria la asistencia a un mínimo del 33% de horas lectivas. Además, la realización de las prácticas es obligatoria y se recomienda la realización de las mismas en el aula bajo la supervisión de los profesores.

### 1.11. Coordinador/a de la asignatura

Modesto Redrejo Rodriguez

<https://autoservicio.uam.es/paginas-blancas/>

### 1.12. Competencias y resultados del aprendizaje

---

#### 1.12.1. Competencias

CE1 - Capacidad de aplicar los conocimientos de biología, matemáticas, física y estadística a la bioinformática.

CE2 - Capacidad de aplicar métodos computacionales a la resolución de problemas en el ámbito de la investigación biomédica básica y traslacional.

CE3 - Capacidad de analizar, explotar e investigar secuencias biológicas comparativa, evolutiva y funcionalmente.

CE4 - Capacidad de predecir estructuras, funciones, y dinámica de biomoléculas.

CE9 - Capacidad de explotación de la información de las bases de datos biológicas sobre secuencias, estructuras, transcriptomas, genomas, proteomas, etc.

#### 1.12.2. Resultados de aprendizaje

-

#### 1.12.3. Objetivos de la asignatura

En este curso se tratan las aproximaciones algorítmicas utilizadas para predecir la estructura primaria, secundaria y terciaria de la proteínas, así como detectar su expresión (proteómica). Además de entender las bases teóricas, se estudiará cómo aplicar estos métodos en la práctica para predicción de estructuras y análisis de expresión. Puesto que la bioinformática es un área en rápida evolución, más que centrarnos en el uso concreto de una herramienta específica de cada aplicación se hará énfasis a los formatos de entrada, parámetros e interpretación de resultados comunes a muchas de ellas.

Al final del semestre (objetivos generales), y de cada unidad (objetivos por tema) el estudiante deberá ser capaz de:

Comprender las bases teóricas, tanto conceptuales como algorítmicas, de la predicción y análisis de estructura de macromoléculas y su dinámica.

Aplicar herramientas bioinformáticas de visualización, predicción y análisis de estructura e interacciones

|                                       |  |                |            |     |
|---------------------------------------|--|----------------|------------|-----|
| <b>Código Seguro de Verificación:</b> |  | <b>Fecha:</b>  | 22/08/2023 | 2/7 |
| <b>Firmado por:</b>                   | <i>Esta guía docente no estará firmada mediante CSV hasta el cierre de actas</i> |                |            |     |
| <b>Url de Verificación:</b>           |  | <b>Página:</b> | 2/7        |     |

de macromoléculas, e interpretar los resultados de estos programas.

Integrar los resultados de modelado de estructuras para resolver problemas biológicos.

| <b>OBJETIVOS GENERALES</b> |  |
|----------------------------|--|
| G1                         | Identificar las principales bases de datos de estructuras biológicas, así como acceder y extraer información de ellas.   |
| G2                         | Comprender las bases teóricas, tanto conceptuales como algorítmicas, de la predicción y análisis de estructura de macromoléculas   |
| G3                         | Aplicar herramientas bioinformáticas de visualización, predicción y análisis de estructura e interacciones de macromoléculas, e interpretar los resultados de estos programas. |
| G4                         | Integrar los resultados de modelado de estructuras para resolver problemas biológicos.   |
| G5                         | Comprender la base experimental y computacional de la caracterización de proteomas.  |

| <b>OBJETIVOS ESPECIFICOS POR TEMA</b>                    |  |
|--|--|
| <b>TEMA 1.- Bases de datos de estructuras</b>            |  |
| 1.1.   | Conocer las principales bases de datos de estructuras de macromoléculas biológicas y de clasificación de proteínas |
| 1.2.   | Ser capaz de hacer búsquedas y extraer información de las bases de datos de estructuras                            |
| <b>TEMA 2.- Visualización de estructuras 3D</b>          |  |
| 2.1.   | Conocer y ser capaz de usar programas de visualización de estructuras  |
| <b>TEMA 3.- Predicción de estructura secundaria y 3D</b> |  |
| 3.1  | Comprender las principales estrategias algorítmicas para la predicción de estructura 2D y 3D de proteínas          |
| 3.2  | Interpretar los resultados de los métodos de predicción y entender sus limitaciones                                |
| 3.3  | Ser capaz de emplear programas específicos para predecir estructuras 2D y 3D                                       |
| <b>TEMA 4.- Proteómica</b>                               |  |
| 4.1  | Comprender las bases del análisis de proteomas y la identificación de proteínas por huella peptídica               |
| 4.2  | Entender y aplicar las técnicas de identificación de proteínas por fragmentación (MS/MS)                           |
| 4.2  | Conocer las bases de la proteómica cuantitativa  |

### 1.13. Contenidos del programa

#### 1. Bases de datos de estructuras

- 1.1. Protein Data Bank
- 1.2. SCOP y CATH

#### 2. Introducción a bioinformática en biología estructural.

- 2.1. Predicción de estructura de proteínas: aplicaciones y limitaciones.
- 2.2. Predicción de estructura 3D mediante modelado por homología

#### 3. Software de visualización de estructuras 2D y 3D

- 3.1. Pymol para el análisis de estructuras y la realización de figuras.
- 3.2. Otros programas especialmente indicados para alternativas específicas.

|                                       |  |                |            |            |
|---------------------------------------|--|----------------|------------|------------|
| <b>Código Seguro de Verificación:</b> |  | <b>Fecha:</b>  | 22/08/2023 | <b>3/7</b> |
| <b>Firmado por:</b>                   | <i>Esta guía docente no estará firmada mediante CSV hasta el cierre de actas</i> |                |            |            |
| <b>Url de Verificación:</b>           |  | <b>Página:</b> | 3/7        |            |

#### 4. Predicción de estructura secundaria y 3D

- 4.1. Predicción de estructura secundaria y otras características 1D
- 4.2. Predicción de estructura 3D mediante métodos mas complejos: la revolución de Alphafold2.
- 4.3. Refinamiento y evaluación de modelos

#### 5. Proteómica

- 5.1. Proteómica por espectrometría de masas
- 5.2. Identificación de proteínas utilizando bases de datos, por huella peptídica (peptide mass fingerprinting), e identificación de péptidos por fragmentación (MS/MS, o MS2)
- 5.3. Fundamentos de proteómica cuantitativa

#### 1.14. Referencias de consulta

##### 1. Libros

- Structural Bioinformatics 2 by Jenny Gu (Redactor), Philip E. Bourne (Redactor) Wiley-Blackwell 2nd edition 2011
- A basic introduction to protein structure and function: Protein Structure and Function Gregory A. Petsko (Autor), Dagmar Ringe (Autor) OUP Oxford 2nd edition (2008)
- Exploring Protein Structure: Principles and Practice (2018) Skern, T. (Ed) SpringerBooks.  
<https://doi.org/10.1007/978-3-319-76858-8>
- Protein Structure Prediction. (Methods in Molecular Biology, Vol. 1137, (2014) Kihara, Daisuke (Ed.)  
<https://doi.org/10.1007/978-1-4939-0366-5>

##### 2. Artículos

- **Artículos sobre modelado:**
  - Baker D, Sali A. Protein structure prediction and structural genomics. *Science*, 2001 <https://science.sciencemag.org/content/294/5540/93.long>
  - Ginalski K, Grishin NV, Godzik A, Rychlewski L. Practical lessons from protein structure prediction. *Nucleic Acids Res*, 2005 <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gki327>
  - Zhang Y. Progress and challenges in protein structure prediction. *Curr Opin Struct Biol*, 2008 [https://linkinghub.elsevier.com/retrieve/pii/S0959-440X\(08\)00034-1](https://linkinghub.elsevier.com/retrieve/pii/S0959-440X(08)00034-1)
  - Kc DB. Recent advances in sequence-based protein structure prediction. *Brief Bioinform.*, 2017, <https://academic.oup.com/bib/article-lookup/doi/10.1093/bib/bbw070>
  - Martinez, X. Molecular Graphics: Bridging Structural Biologists and Computer Scientists. *Structure*, 2019  
[https://linkinghub.elsevier.com/retrieve/pii/S0969-2126\(19\)30307-7](https://linkinghub.elsevier.com/retrieve/pii/S0969-2126(19)30307-7)
  - Goddard, TD. UCSF **ChimeraX**: Meeting modern challenges in visualization and analysis. *Protein Science*, 2018. <https://onlinelibrary.wiley.com/doi/full/10.1002/pro.3235>
- **Artículos sobre Proteómica:**
  - **Artículo donde se presenta el programa Comet (véase más abajo):** Eng, J, Jahan, TA, and Hoopmann, MR, «Comet: An open-source MS/MS sequence database search tool», *Proteomics*, 2012 <http://onlinelibrary.wiley.com/doi/10.1002/pmic.201200439/full>
  - **Artículo en el que se presenta el XCorr, el parámetro principal utilizado por Comet para identificar péptidos:** Eng, J, McCormack, AL, and Yates III, JR, «An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database», *Journal of the American Society for Mass Spectrometry*, 1994 <http://www.sciencedirect.com/science/article/pii/1044030594800162>
  - **Artículo donde se habla de la principal estrategia para identificar proteínas,**

|                                |   |         |            |     |
|--------------------------------|---|---------|------------|-----|
| Código Seguro de Verificación: |   | Fecha:  | 22/08/2023 | 4/7 |
| Firmado por:                   | Esta guía docente no estará firmada mediante CSV hasta el cierre de actas |         |            |     |
| Url de Verificación:           |   | Página: | 4/7        |     |

- utilizando bases de datos señuelo (*target-decoy approach*):** Elias, JE, and Gygi, SP, «Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry», *Nature Methods*, 2007 <http://www.nature.com/nmeth/journal/v4/n3/abs/nmeth1019.html>
- **Revisión sobre proteómica basada en espectrometría de masas:** Aebersold, R, and Mann, M, «Mass spectrometry-based proteomics», *Nature*, 2003 <http://www.nature.com/nature/journal/v422/n6928/abs/nature01511.html>

## Recursos adicionales

- **Comet:** programa de código abierto para la identificación de secuencias peptídicas. Descargable gratuitamente desde aquí: <http://comet-ms.sourceforge.net>
- **Mascot:** programa (de código no abierto) para la identificación de secuencias peptídicas. Se puede utilizar en red sin necesidad de instalar nada. La versión gratuita tiene ciertas limitaciones, pero no afectan a la práctica individual. Se puede encontrar aquí: [http://www.matrixscience.com/search\\_form\\_select.html](http://www.matrixscience.com/search_form_select.html)
- **«Interpreting MS/MS Proteomics Results», de Brian Searle:** Una presentación con diapositivas amena y didáctica que cuenta bastantes detalles sobre cómo se identifican proteínas, [http://www-bac.esi.umontreal.ca/~dbcm2003/proteomique/docs/Interpreting\\_MSMS\\_results\\_cartoon.ppt](http://www-bac.esi.umontreal.ca/~dbcm2003/proteomique/docs/Interpreting_MSMS_results_cartoon.ppt)

**Material electrónico de trabajo :** los documentos electrónicos de trabajo (material adicional, ejercicios del curso y ejemplos de exámenes, etc.) se publican en la plataforma Moodle ( <http://posgrado.uam.es/> )

## 2. Metodologías docentes y tiempo de trabajo del estudiante

### 2.1. Presencialidad

Tal y como marca la normativa es obligatoria la asistencia a un mínimo del 33% de horas lectivas.

### 2.2. Relación de actividades formativas

| Actividades presenciales          | Nº horas |
|-----------------------------------|----------|
| Clases teóricas en aula           | 16       |
| Seminarios                        |          |
| Clases prácticas en aula          | 8        |
| Prácticas clínicas                |          |
| Prácticas con medios informáticos |          |
| Prácticas de campo                |          |
| Prácticas de laboratorio          |          |
| Prácticas externas y/o practicum  |          |
| Trabajos académicamente dirigidos |          |
| Tutorías                          |          |
| Actividades de evaluación         | 3        |
| Otras                             | 48       |

La metodología utilizada en el desarrollo de la actividad docente incluye los siguientes tipos de actividades:

\*Clases de teoría:

|                                |   |         |            |     |
|--------------------------------|---|---------|------------|-----|
| Código Seguro de Verificación: |   | Fecha:  | 22/08/2023 | 5/7 |
| Firmado por:                   | Esta guía docente no estará firmada mediante CSV hasta el cierre de actas |         |            |     |
| Url de Verificación:           |   | Página: | 5/7        |     |

### Actividad del profesor

Clases expositivas breves en las que se realizarán programas y ejercicios ilustrativos. Se combinarán con cuestiones y preguntas a los alumnos para estimular el debate y asentamiento de los conceptos presentados.

### Actividad del estudiante:

*Actividad presencial:* Toma de apuntes, participación en clase respondiendo a las cuestiones planteadas. Resolución de los ejercicios propuestos.

*Actividad no presencial:* lectura del material bibliográfico y de apoyo, estudio de la materia y realización de ejercicios propuestos y de los cuestionarios planteados en la plataforma Moodle.

### \*Tutorías en aula:

#### Actividad del profesor:

Tutorización a toda la clase o en grupo reducidos con el objetivo de resolver dudas comunes plantadas por los alumnos a nivel individual o en grupo, surgidas a partir de cuestiones/ejercicios señalados en clase para tal fin y orientarlos en la realización de los mismos.

#### Actividad del estudiante:

*Actividad presencial:* Planteamiento de dudas individuales o en grupo y enfoque de posibles soluciones a las tareas planteadas.

*Actividad no presencial:* Estudio de las tareas marcadas y debate de las soluciones planteadas en el seno del grupo.

### \*Prácticas:

#### Actividad del profesor:

Asignar una práctica/proyecto a cada grupo de trabajo y explicar la práctica asignada a cada grupo de trabajo al comienzo de la sesión de prácticas. Los medios utilizados son los entornos de programación, *software* específico de uso local y en servidores remotos y los ordenadores del propio laboratorio para la ejecución y análisis de los programas realizados.

En la segunda parte identificaremos proteínas utilizando herramientas accesibles en red y archivos procedentes de espectrometría de masas proporcionados por el profesor. Identificaremos proteínas tanto por huella peptídica como por fragmentación. Veremos las diferencias entre ambas técnicas, y cómo mejoran o empeoran los resultados en función de los parámetros utilizados.

Las prácticas se realizarán en grupos de 2-4 estudiantes, entregando a final del curso un *Cuaderno de laboratorio electrónico* en el que se incluyan los resultados obtenidos. La evaluación de las prácticas será fundamentalmente formativa e incluirá evaluación por pares.

#### Actividad del estudiante:

*Actividad presencial:* Planteamiento inicial, previo al desarrollo de la práctica, sobre información contenida en el enunciado. Debate en el seno del grupo sobre el planteamiento de la solución óptima. Ejecución de la práctica y entrega de resultados. Los alumnos participarán activamente en la identificación de las proteínas con los espectros proporcionados.

*Actividad no presencial:* Profundizar en el enunciado de la práctica, plantear el diagrama de flujo óptimo para la resolución de la misma e implementar los ejercicios propuestos. Los alumnos dispondrán de archivos adicionales para hacer más búsquedas fuera de clase si lo desean, así como interpretar los resultados obtenidos. Se realizará evaluación por pares de las prácticas de los demás grupos.

## 3. Sistemas de evaluación y porcentaje en la calificación final

### 3.1. Convocatoria ordinaria

Todas las partes de la asignatura, teoría y prácticas puntúan sobre 10.

La calificación final de la asignatura en la convocatoria ordinaria se obtiene a partir de las notas de examen(es) de teoría (Teoría) y calificación obtenido en los ejercicios y/o prácticas y participación en clase (Prácticas) mediante la ecuación:

$$\text{Calificación: } 0.6 * \text{Prácticas} + 0.4 * \text{Teoría}$$

|                                |   |         |            |     |
|--------------------------------|---|---------|------------|-----|
| Código Seguro de Verificación: |   | Fecha:  | 22/08/2023 | 6/7 |
| Firmado por:                   | Esta guía docente no estará firmada mediante CSV hasta el cierre de actas |         |            |     |
| Url de Verificación:           |   | Página: | 6/7        |     |

Para aprobar la asignatura es obligatorio obtener una nota mayor o igual a 5 puntos tanto en la parte de teoría como en las prácticas. En caso contrario, la nota final en actas será

Calificación:  $\text{Mín}(4.9, 0.6 * \text{Prácticas} + 0.4 * \text{Teoría})$

La calificación de teoría será media aritmética ponderada (por cantidad contenido incluido en el examen) de los exámenes que se realicen. Las pruebas no realizadas por causas sin justificar computan como cero en la media.

La calificación de prácticas será media aritmética ponderada (por cantidad contenido incluido en el examen) de los ejercicios que se realicen. Los retrasos en las entregas de ejercicios/prácticas serán penalizados con el 25% de la calificación inicial por día de retraso.

La calificación de la parte práctica tendrá en cuenta la calidad de los diseños realizados, los resultados obtenidos y de su presentación, según las indicaciones del profesorado. Se tendrá en cuenta también la participación activa en clase.

**ATENCIÓN: Cualquier copia descubierta que se haya realizado a lo largo del curso, tanto en cualquiera de las actividades de teoría desarrolladas, ejercicios, y debate como en cualquiera de los apartados de las prácticas, serán penalizadas con rigurosidad. La penalización por copia implica la aplicación de la normativa interna de la UAM, que supone suspender la convocatoria actual.**

### 3.1.1. Relación actividades de evaluación

Las prácticas de la asignatura podrán sumar hasta un 60% de la nota final. Se realizará un examen tipo test que contará el 40% restante.

### 3.2. Convocatoria extraordinaria

Para la convocatoria extraordinaria

- Los estudiantes deben revisar las entregas realizadas, completándolas y corrigiendo sus errores sobre la versión entregada durante el curso, sobre un problema de modelado alternativo.
- En el caso de que no se realizara una o más entregas en la convocatoria ordinaria se podrá acordar con los profesores el enunciado de las prácticas para la convocatoria extraordinaria.
- Realizarán un examen escrito o/y oral.
- Calificación:  $\text{Mín}(4.9, 0.6 * \text{Prácticas} + 0.4 * \text{Teoría})$

### 3.2.1. Relación actividades de evaluación

-

## 4. Cronograma orientativo

| Semana | Contenido  |
|--------|--|
| 1      | Introducción a las bases de datos de estructuras<br>Visualización de estructuras |
| 2      | Predicción de estructuras  |
| 3      | Predicción de estructuras<br>Proteómica  |
| 3      | Proteómica   |
|        | Examen Final convocatoria ordinaria  |
|        | Examen Final convocatoria extraordinaria   |

|                                |   |         |            |     |
|--------------------------------|---|---------|------------|-----|
| Código Seguro de Verificación: |   | Fecha:  | 22/08/2023 | 7/7 |
| Firmado por:                   | Esta guía docente no estará firmada mediante CSV hasta el cierre de actas |         |            |     |
| Url de Verificación:           |   | Página: | 7/7        |     |