

Capítulo 13

Estudio y bioinformática de Proteínas

José Francisco Morales Domínguez
Cindy Damara Anaya Madrigal
Claudia Samantha Díaz de León Gutiérrez
Nora Lilia Vasco Méndez
*Departamento de Química, Centro de Ciencias Básicas
Universidad Autónoma de Aguascalientes*

Resumen

Hasta ahora se sabe que la actividad y función de las células es dictada por la información genética del organismo en cuestión, de esta forma se sintetizan las proteínas de acuerdo con las necesidades del metabolismo celular. Los genes que serán encendidos van a depender específicamente del momento y las necesidades fisiológicas de la célula, así como del tejido al que pertenezca. De esta forma se sintetizan las proteínas que lleven a cabo los procesos metabólicos requeridos, ya que las proteínas, específicamente las enzimas, son los componentes principales de las rutas metabólicas al realizar las reacciones químicas específicas que sostienen el metabolismo y les dan mantenimiento a las estructuras propias de las células. En este capítulo daremos una vista general de las proteínas con la finalidad de entender-

las, es decir conocer la importancia para las plantas y de cualquier ser vivo. También se abordará un apartado donde se muestre una ruta bioinformática para la búsqueda de proteínas y su modelado hipotético.

Aminoácidos: bloques constructores de las proteínas

Los aminoácidos son los compuestos químicos de los cuales se encuentran constituidas las proteínas, su estructura química general se muestra en la figura 13.1. El aminoácido debe su nombre a que contiene un grupo amino y un grupo ácido, ambos unidos a un carbono central denominado carbono alfa, al cual también se le une un átomo de hidrogeno y un grupo R o cadena lateral que es un comodín, ya que puede ser intercambiado por diferentes Radicales o compuestos químicos que le da identidad al aminoácido, dependiendo del grupo R es el nombre que recibe el aminoácido (Fig.13.1).

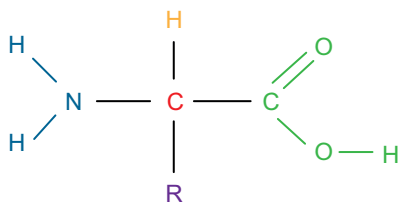


Figura 13.1. Estructura general de un aminoácido. En azul se muestra el grupo amino, en rojo el carbono alfa, en verde el grupo carboxilo y en morado el grupo de cadena lateral (radical R) quien diferencia a los aminoácidos.

Cada aminoácido tiene su propio nombre y su nombre abreviado ya sea por una o por tres letras (Tabla 13.1). Dependiendo de las características químicas del grupo R o radical, los aminoácidos pueden ser clasificados de muy diferentes formas.

Tabla 13.1. Características y código genético de los aminoácidos

Aminoácido	Codón	Características
Arginina Arg R	AGA	Aminoácido esencial para el ser humano, presenta carga positiva debido a sus grupos amino extras en su cadena lateral que le otorgan su carácter básico y es polar. Es glucogénico. Se encuentra en carnes, pescado y lácteos.
	AGG	
	CGA	
	CGC	
	CGG	
	CGU	
Histidina His H	CAC	Aminoácido esencial para el ser humano, presenta carga positiva debido a su grupo imidazol en su cadena lateral que le otorga su carácter básico y es polar. Es glucogénico. Se encuentra en carnes, pescado y lácteos.
	CAU	
Lisina Lys K	AAA	Aminoácido esencial para el ser humano, presenta carga positiva debido a un grupo amino extra en su cadena lateral que le otorga su carácter básico, es polar y cetogénico. Se encuentra en carnes, pescado y lácteos.
	AAG	
Ácido Aspártico Asp D	GAU	Aminoácido no esencial para el ser humano, presenta carga negativa debido a un grupo carboxilo extra en su cadena lateral y es polar. Es glucogénico. Se encuentra en chicharos, lentejas soya y espárragos.
	GAC	
Ácido Glutámico Glu E	GAA	Aminoácido no esencial para el ser humano, presenta carga negativa debido a un grupo carboxilo extra en su cadena lateral y es polar. Es glucogénico. Se encuentra en proteínas vegetales y animales.
	GAG	
Serina Ser S	AGC	Aminoácido no esencial para el ser humano, no presenta carga, pero es polar debido al grupo hidroxilo en su cadena lateral. Es glucogénico. Se obtiene de la hidrólisis de la seda. Se encuentra en proteínas vegetales y animales.
	AGU	
	UCA	
	UCC	
	UCG	
	UCU	
Treonina Thr T	ACA	Aminoácido esencial para el ser humano, no presenta carga, pero es polar debido al grupo hidroxilo en su cadena lateral. Es cetogénico y glucogénico. Se encuentra en proteínas vegetales y animales. Especialmente en el requesón, lentejas y ajonjolí.
	ACC	
	ACG	
	ACU	
Asparagina Asn N	AAC	Aminoácido no esencial, polar sin carga debido a que presenta un grupo carboxamida como cadena lateral. Es glucogénico. Se encuentra en proteínas vegetales y animales.
	AAU	
Glutamina Gln Q	CAA	Aminoácido no esencial con cadena lateral polar sin carga. Es muy importante en la degradación de aminoácidos y síntesis de aminoácidos. Es glucogénico. Se encuentra en vegetales de hoja verde.
	CAG	

Aminoácido	Codón	Características
Alanina Ala A	GCA GCC GCG GCU	Aminoácido no esencial para el ser humano, hidrófobo sin carga, con un grupo metilo como su cadena lateral. Es glucogénico. Se encuentra en proteínas de origen animal como carne, leche y huevo.
Isoleucina Ile I	AUA AUC AUU	Aminoácido esencial para el ser humano, es hidrófobo sin carga debido a que su cadena lateral es alifática. Es cetogénico y glucogénico. Se encuentra en proteína animal, en la soya y amaranto.
Leucina Leu L	CUA CUC CUG CUU UUA UUG	Aminoácido esencial para el ser humano, la cadena lateral es alifática lo que le confiere su carácter hidrófobo, es cetogénico. Se localiza en proteínas animales y en leguminosas.
Metionina Met M	AUG	Aminoácido esencial para el ser humano, es hidrófobo debido a su cadena lateral alifática que presenta azufre, es uno de los dos aminoácidos proteicos azufrados. Es glucogénico. Es uno de los dos únicos aminoácidos con un solo codón que codifica para el aminoácido además del Triptófano, dicho codón indica el inicio de la síntesis de las proteínas. Todas las proteínas “nacen” con la metionina como primer aminoácido.
Fenilalanina Phe F	UUC UUU	Aminoácido esencial para el ser humano, hidrófobo debido a que contiene un anillo bencénico en su cadena lateral por lo que pertenece a los aminoácidos aromáticos, es tanto cetogénico como glucogénico. Se encuentra en proteína animal leche carne y huevo, y en proteína vegetal de leguminosas.
Triptófano Trp W	UGG	Aminoácido esencial para el ser humano, es uno de los dos aminoácidos que es codificado por un solo codón, el otro es la Metionina. Su cadena lateral es un grupo indol aromático lo que le confiere carácter hidrófobo. Es cetogénico y glucogénico. Se encuentra en proteína animal como vegetal.
Tirosina Tyr Y	UAC UAU	Aminoácido no esencial para el ser humano, aromático, su cadena lateral presenta un anillo bencénico hidroxilado lo que le confiere cierta polaridad a la cadena lateral hidrófoba. Es cetogénico y glucogénico. Se encuentra en proteína animal y vegetal.
Valina Val V	GUA GUC GUG GUU	Aminoácido esencial para el ser humano, con cadena alifático lo que le confiere carácter hidrófobo. Es glucogénico. Se encuentra en frutos rojos, chocolate, plátano.

Aminoácido	Codón	Características
Cisteína Cys C	UGC UGU	Aminoácido no esencial para el ser humano, es uno de los dos aminoácidos azufrados, el otro es la metionina, en su cadena lateral se presenta un grupo tiol capaz de formar puentes disulfuro con otra cisteína para formar el residuo Cistina. Es glucogénico. Se encuentra en proteína animal y vegetal.
Prolina Pro P	CCA CCC CCG CCU	Aminoácido no esencial para el ser humano, su cadena lateral alifática se une con el grupo amino formando un anillo, lo que le permite a la cadena polipeptídica dar giros cerrados y cambiar la dirección de los motivos. Es glucogénico.
Glicina Gly G	GGA GGC GGG GGU	Aminoácido no esencial para el ser humano, su cadena lateral es sólo un hidrógeno lo que le hace ser el aminoácido más pequeño, se le considera hidrófobo. Es glucogénico. Se encuentra ampliamente distribuido en proteína vegetal y animal.

Los aminoácidos alifáticos son aminoácidos cuya cadena lateral o grupo R es hidrofóbico constituido exclusivamente por carbono e hidrógeno.

Los aminoácidos aromáticos contienen un anillo de benceno por lo que son hasta cierto punto hidrófobos, son utilizados para sintetizar algunas hormonas en el organismo.

Los aminoácidos ácidos justo tienen la característica de ser ácidos y tienen carga negativa, mientras que los aminoácidos básicos son básicos o alcalinos y tienen carga positiva, en ambos casos por tener carga positiva o negativa son aminoácidos hidrófilos.

Los aminoácidos amidas contienen un segundo grupo amino en su cadena lateral por lo que están cargados positivamente y son hidrófilos.

Los aminoácidos hidroxilados contienen grupos OH en su cadena lateral lo que los hace ser hidrófilos.

Los aminoácidos azufrados contienen azufre en su cadena lateral.

La prolina es el único aminoácido en el que su cadena lateral se une con el grupo amino y forma un anillo.

Enlace Peptídico

Se considera una proteína cuando se han enlazado de forma lineal 100 o más aminoácidos y se considera péptido cuando se han enlazado menos de 100 aminoácidos. El enlace entre aminoácidos se lleva a cabo mediante el denominado enlace peptídico que es un enlace covalente entre el carbono del grupo ácido del primer aminoácido con el nitrógeno del grupo amino del segundo aminoácido, al formarse el enlace peptídico se desprende una molécula de agua como se muestra en la Figura 13.2.

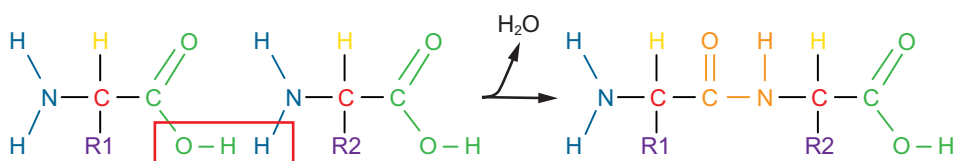


Figura 13.2. Unión peptídica entre dos aminoácidos. El OH del carboxilo del aminoácido uno; se une con el hidrógeno del grupo amino del segundo aminoácido, liberando una molécula de agua durante la reacción de formación del enlace peptídico entre el carbono del grupo ácido del primer aminoácido y el nitrógeno del grupo amino del segundo aminoácido.

El enlace peptídico es un enlace más corto que el enlace covalente simple y también es rígido como un doble enlace covalente, ambas características son debido a que la nube electrónica del doble enlace entre el carbono con el oxígeno del grupo carboxilo también resuena en el enlace peptídico entre el carbono del grupo carboxilo del primer aminoácido y el nitrógeno del grupo amino del siguiente aminoácido, de tal forma que podríamos mencionar que el enlace peptídico es un “semidoble enlace covalente”. Los demás enlaces covalentes simples de los aminoácidos en la cadena de aminoácidos tienen su movilidad normal de 360°.

Imaginemos una proteína como un collar de perlas, en donde cada perla es un aminoácido y el hilo que las une es el enlace peptídico. Recordemos que las cadenas laterales o grupos R de los aminoácidos tienen características químicas que los hacen tener cargas positivas o negativas, ser hidrófilos o hidrófobos, sabemos que las cargas opuestas se atraen, sabemos que cargas iguales se repelen, sabemos que lo hidrófilo atrae a lo hidrófilo, que lo hidrófobo atrae a lo hidrófobo y que lo hidrófobo repele a lo hidrófilo, de esta forma

imaginemos que en ese collar de aminoácidos algunos aminoácidos se atraen entre sí, que otros aminoácidos se repelen uno a otros, y de esta forma sin romperse los enlaces peptídicos, se empiezan a generar formas en la cadena de aminoácidos (proteína) por atracción y repulsión de las cadenas laterales o grupos R de los aminoácidos, un ejemplo sencillo es el que mostramos abajo. Esas diferentes formas van generando estructuras en las proteínas.



Figura 13.3. Ejemplo sencillo de atracciones y repulsiones en la cadena de aminoácidos.

Importancia y funciones de las proteínas.

Aspectos nutricionales de las proteínas

Todos los organismos vivos estamos hechos de proteínas en mayor o menor grado, en general de dos tipos de proteínas, proteínas sin actividad metabólica y proteínas con actividad metabólica (enzimas). Las proteínas sin actividad metabólica, son proteínas estructurales y/o de almacenamiento, es decir forman parte de estructuras como los ribosomas, forman parte del cito esqueleto de las células, y las proteínas de almacenamiento que forman gránulos proteicos en las células del endospermo o cotiledones de las semillas, que le servirán a la semilla para alimentarse y activar el metabolismo para la germinación y establecimiento de la nueva planta en suelo, entre otras muchas funciones de las proteínas.

Las proteínas con actividad metabólica básicamente son enzimas que llevan a cabo las reacciones bioquímicas dentro de nuestras células con lo que se mantienen la homeostasis del organismo.

Por su origen las proteínas pueden clasificarse en proteínas de origen animal o proteínas de origen vegetal. Las proteínas de origen animal son las de alto valor nutricional para el ser humano, estas son las proteínas de la leche, el huevo y todo tipo de carne o pescado.

Las proteínas de intermedio valor nutricional para el hombre, en general son las proteínas de leguminosas y las de cereales. Los vegetarianos mezclan proteínas de leguminosas con proteínas de cereales y de esta forma consumen proteína de mejor calidad nutricional, ya que se complementan en su valor nutricional. Es recomendable que los niños que están en crecimiento consuman proteínas de alta calidad nutricional, mientras que los adultos se encontrarán en buenas condiciones consumiendo la mezcla de proteínas de leguminosas con cereales.

La calidad nutricional de las proteínas depende del número y cantidad de aminoácidos de los cuales se encuentre constituida la proteína. Existen 20 aminoácidos de los cuales están conformadas las proteínas, 10 de ellos son aminoácidos esenciales, lo que significa que deben ser ingeridos de los alimentos ya que nuestro organismo no es capaz de sintetizarlos, de esta forma las proteínas constituidas por esos 10 aminoácidos esenciales y cuyas cantidades en la misma proteína sean las adecuadas para el ser humano, constituyen proteínas de alto valor nutricional (leche, huevo, todo tipo de carne y pescado), mientras que las proteínas de origen vegetal como leguminosas y cereales, presentan de bajo a intermedio valor nutricional, de ahí que deben ser combinadas.

Las proteínas, tanto enzimas como proteínas estructurales y de almacenamiento que se encuentran en las células de animales y vegetales son de las que nos alimentamos. Desde el punto de vista nutricional las proteínas son de gran importancia en la alimentación del ser humano, en general se recomienda 0.99 g de proteína por kilogramo de peso al día. De esta forma, una persona de 60 kg deberá consumir 59.4 g de proteína al día y una persona de 80 kg deberá consumir 79.2 g de proteína al día, para nada son las cantidades exorbitantes que solemos consumir.

Tabla 13.2. Aminoácidos esenciales y no esenciales para el organismo humano

Esenciales	No esenciales
Histidina	Alanina
Isoleucina	Asparagina
Leucina	Ácido aspártico
Lisina	Ácido glutámico
Metionina	Cisteína
Fenilalanina	Glutamina
Treonina	Glicina
Triptófano	Prolina
Valina	Serina
Arginina*	Tirosina

*El aminoácido Arginina es esencial durante la infancia.

Niveles estructurales de las proteínas

Estructura primaria

En este apartado y para fines más didácticos, se analizarán los tres niveles estructurales de la proteína β -expansina de maíz (EXPB1_MAIZE).

La estructura primaria de una proteína se refiere a la secuencia de aminoácidos en la cadena polipeptídica (proteína), dicha secuencia está dada por la información genética almacenada en el ADN, recordemos que los aminoácidos están codificados por tripletes de nucleótidos denominados codones. En la figura 13.4, se observa los primeros treinta nucleótidos y su complementario de la secuencia de la EXPA1 de maíz. Ha de recordar el lector que, en todas las secuencias de ADN, al inicio de cada cadena lleva el 5' y al final el 3'. Debido a que son treinta nucleótidos, entonces se formarán 10 codones o tripletes de nucleótidos y por lo tanto diez aminoácidos; esto se puede observar en la misma figura en el primer marco de lectura abierta (IMLA), donde el codón atg codifica para M, o metionina y el último gtc para V, o valina.

Sin embargo, existe la posibilidad de que la lectura del gen empiece en el segundo nucleótido y de allí se forme otro diferente codón que sería tgg que

sería el 2 MLA. En la figura 13.4, el nucleótido que no se tomaría en cuenta para la nueva lectura esta de color rojo. A este tipo de lectura se le conoce como marcos de lectura abierta (MLA), lo que quiere decir que en nuestra secuencia se tienen seis diferentes MLA, de esta manera se tendrían seis secuencias de aminoácidos diferentes. Para el cuarto MLA, éste empezaría en la cadena complementaria en la dirección de 5' a 3' por lo que el primer codón sería gac que codifica para D o ácido aspártico.

Por lo que el primer codón sería gac que codifica para D, o ácido aspártico. También se observa que en los MLA tres, y seis existen codones de paro que están representados por un guion (taa, tga, que en el ARN serían uaa y uga).

Secuencia de ADN

5' atg gga tcc ctc gtt aat aac atc atg gtc 3'
3' tac cct agg gag caa tta ttg tag tac cag 5'

Marcos de lectura abierta

1 MLA
atg gga tcc ctc gtt aat aac atc atg gtc
M G S L V N N I M V

2 MLA
atgg gat ccc tog tta ata aca tca tgg
W D P S L I T S W

3 MLA
atggg atc cct cgt taa taa cat cat ggt
G I P R - - H H G

4 MLA
gac cat gat gtt att aac gag gga tcc cat
D H D V I N E G S H

5 MLA
gacc atg atg tta tta acg agg gat ccc
T M M L L T R D P

6 MLA
gacca tga tgt tat taa cga ggg atc cca
P - C Y - R G I P

Figura 13.4. Marcos de lectura abierta (MLA). En la figura se observan los seis MLA de la secuencia de ADN. En rojo se representan las bases nitrogenadas que no se toma en cuenta para la formación del codón y el guion es el codón de paro (<https://web.expasy.org/translate/>).

También con la estructura primaria se puede predecir la estructura secundaria de la proteína. Como ejemplo se tiene la estructura primaria de una expansina nombrada en la base de datos como EXPB1_MAIZE (Fig. 13.5); las expansinas son proteínas que modifican la pared celular de plantas. En la misma figura, se muestra, que la secuencia de aminoácidos lleva un orden, en este ejemplo la proteína comienza con el aminoácido uno el cual es M o Metionina y corresponde al grupo Amino terminal (H_3N^+ -, NH_3^+ -) libre, y así sucesivamente hasta el último aminoácido, que en este caso es el aminoácido Y, o tirosina, que se distingue por el grupo carboxilo ($-COO^-$) libre. Los enlaces peptídicos son la fuerza que mantiene la estructura primaria.

```
>sp|P58738|EXPB1_MAIZE  Expansin-B1  OS=Zea  mays  OX=4577  GN=EXPB1  PE=1  SV=2
MGSLANNIMVVGAVLAALVAGGSCGPPKVPPGPNITTNYNGKWLTARATWYGQPNGAGAPDNGGACGIKNVNLPPY
SGMTACGNVPIFKDGGKGCSCYEVRCKEKPECSGNPVTVYITDMNYEPIAPYHFDLSGKAFGSLAKPGLNDKIRHCGI
MDVEFRRVRCKYPAGQKIVFHIEKGCNPNYLAVLVKYVADDGDIVLMEIQDKLSAEWKPMLKSWGAIWRMDTAKALK
GPFSIRLTSESGKKVIKDVIPANWRPDAVYTSNVQFY
```

Figura 13.5. Secuencia de aminoácidos de una expansina de maíz (EXPB1_MAIZE). Sombreado en verde es donde inicia el modelo de la estructura 3D. Sombreado en gris, el dominio EG45 y sombreado en azul el dominio CBD y en magenta el motivo conservado HFD característico de las expansinas (<https://www.ncbi.nlm.nih.gov/>).

Estructura secundaria

Existen dos tipos de estructura secundaria en una proteína, la denominada alfa hélice y la denominada beta hoja plegada, ambas estructuras son estables gracias a la gran cantidad de puentes de hidrógeno (atracciones electrostáticas entre cargas negativas y positivas) que se generan entre las cadenas laterales de los aminoácidos. Ambas estructuras básicamente se generan por la secuencia de aminoácidos, los estudios de estas estructuras muestran aminoácidos con mayor prevalencia en cada una de ellas, como se muestra en la tabla 13.3.

Tabla 13.3. Aminoácidos que muestran mayor presencia en las estructuras secundarias

Aminoácido	Alfa Hélice	Beta Plegada	Vuelta Reversa
Ac. Glutámico	1.58	0.52	1.01
Alanina	1.41	0.72	0.82
Leucina	1.34	1.22	0.57
Metionina	1.30	1.14	0.52
Glutamina	1.27	0.98	0.84
Lisina	1.23	0.69	1.07
Arginina	1.21	0.84	0.90
Histidina	1.05	0.80	0.81
Valina	0.90	1.87	0.41
Isoleucina	1.09	1.67	0.47
Tirosina	0.74	1.45	0.76
Cisteína	0.66	1.40	0.54
Triptófano	1.02	1.35	0.65
Fenilalanina	1.16	1.33	0.59
Treonina	0.76	1.17	0.90
Glicina	0.43	0.58	1.77
Asparagina	0.76	0.48	1.34
Prolina	0.34	0.31	1.32
Serina	0.57	0.96	1.22
Ac. Aspártico	0.99	0.39	1.24

En la figura 13.6, se muestra la estructura secundaria de la EXPB1_MAI-ZE, con número de acceso 2hcz en el PDB. Las letras A mayúsculas en rojo representan la estructura beta, mientras que las letras H mayúsculas en azul son las alfa hélices, los motivos de giro están representados por el símbolo β para las vueltas en beta y γ y vuelta gamma, las uniones entre los números encerrados en círculo amarillo (1, 2 y 3) representa los enlaces de disulfuro (www.ebi.ac.uk)

En resumen, esta proteína tiene tres alfa-hélices y 14 betas con varios giros beta y gamma y tres puentes disulfuro donde participan seis cisteínas (C) conservadas. Otra manera de presentar la estructura secundaria es como se muestra

en la figura 13.7. En la figura 13.7a, las flechas en rosa representan las estructuras betas y los cilindros en rojo las alfa-hélices. La primera estructura en beta empieza con la secuencia de LTARATWY que van de la posición 20-27, seguida de otra beta de la posición 55-59 y dos alfa-hélice de la 60-65 y 66- 68 y así sucesivamente hasta llegar a la posición 131-138 con una beta (Fig. 13.7 a y b).

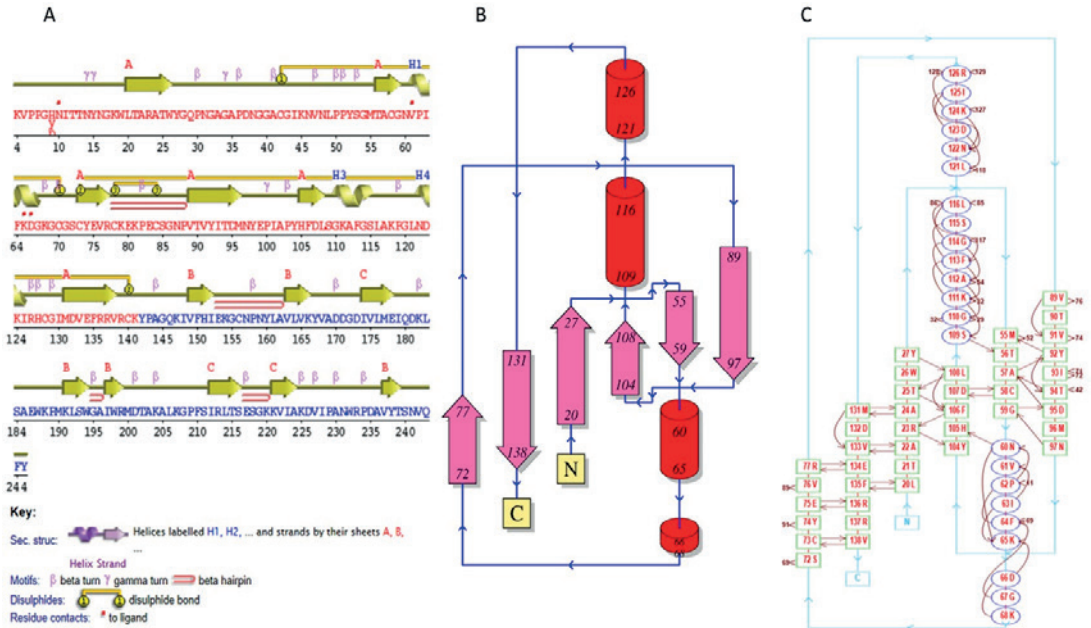


Figura 13.6. Estructura secundaria de la EXPB1_MAIZE. En letras color rojo corresponde al dominio I y en azul al dominio II, la letra A mayúscula en rojo arriba de la flecha indica las estructuras beta, las letras H mayúsculas en azul la estructura alfa-hélice (imagen tomada de PDBSum: www.ebi.ac.uk)

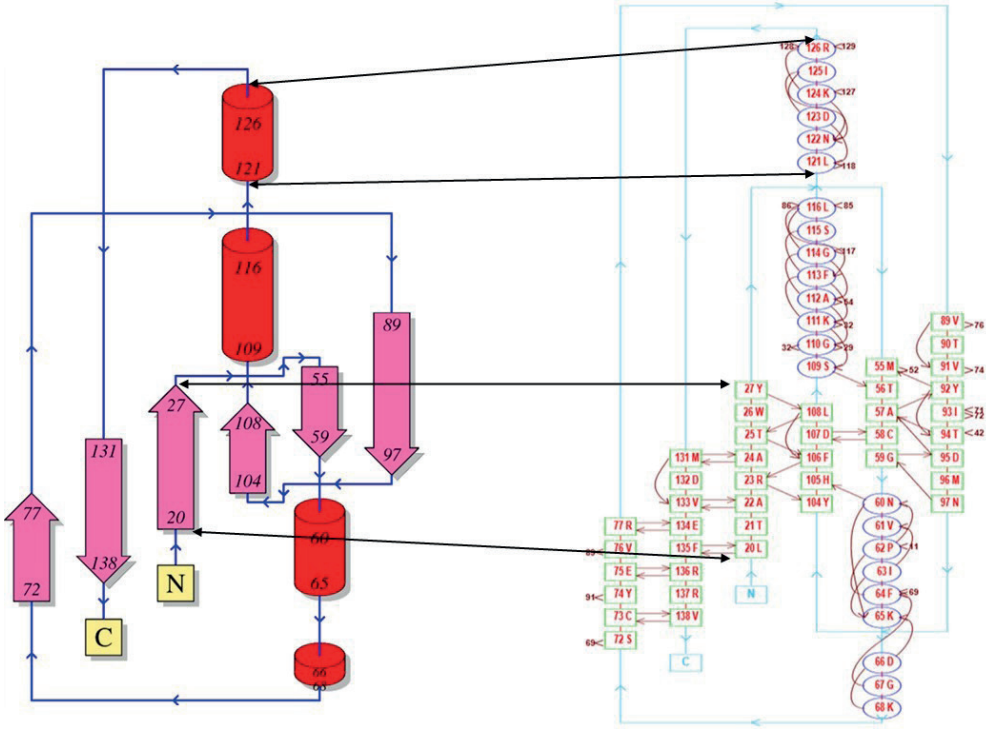


Figura 13.7. Diferentes modelos de la estructura secundaria de la EXPB1_MAIZE. a) Representación de flechas para las estructuras beta y cilindros para las estructura alfa-hélice; cada estructura contiene la posición de los aminoácidos. b) Modelo donde se presenta las estructuras con los aminoácidos correspondientes y su localización. Las flechas negras que unen a las dos figuras muestran la identidad de los dominios (<https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=2hcz>).

Las alfa hélices son estables con vueltas a la derecha (sentido del reloj) e inestables y poco probables a la izquierda, en cada vuelta completa del alfa hélice hay 3.6 aminoácidos (denominados residuos porque cada aminoácido de la cadena perdió un grupo OH del carboxilo y un H del grupo amino al formarse el enlace peptídico), se generan puentes de hidrógeno entre los aminoácidos n y n+4 lo que le da estabilidad a la estructura. En la figura 13.8a, se observa las secuencias y el orden de aminoácidos que están involucrados para la formación de las cuatro alfa-hélices presentes en la estructura de EXPB1_MAIZE.

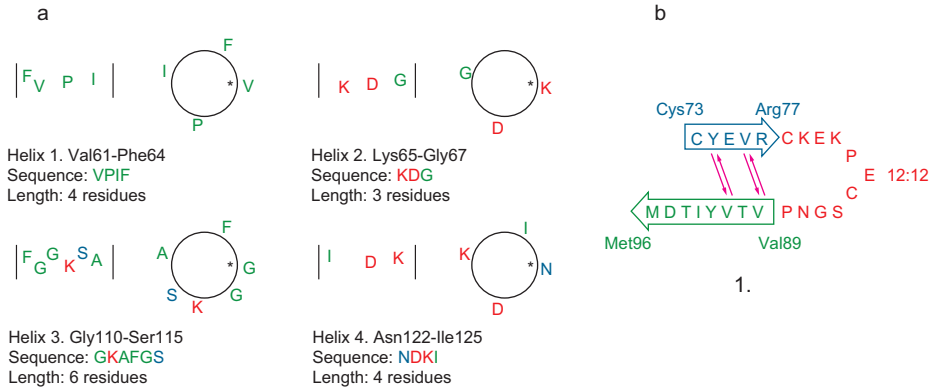


Figura 13.8. Modelos para las estructuras alfa-hélice y beta plegada. a) secuencia y orden de los aminoácidos que forman las alfa-hélices. b) secuencia y orden de los aminoácidos que forman las beta plegada antiparalelas (<https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=2hcz>)

En las hojas plegadas beta, los residuos de aminoácidos se acomodan en zigzag formando una especie de acordeón, las cadenas laterales de los aminoácidos quedan hacia arriba o hacia abajo del plano en zigzag y los puentes de hidrógeno que le dan estabilidad a la estructura se forman al interactuar diferentes zonas de ambas cadenas de aminoácidos, como se observa en la figura 13.8b. Por ejemplo, en la formación de una de las betas plegadas en antiparalelo de la EXPB1_MAIZE se da entre las posiciones 73-77 con la posición 89-96 (Fig. 13.8b).

Dos o más hojas plegadas beta pueden agruparse de forma paralela cuando el sentido de los aminoácidos en ambas cadenas apunta en una misma dirección o bien de forma antiparalela cuando las cadenas apuntan en direcciones opuestas. Existe mayor estabilidad en la estructura antiparalela, debido a que los puentes de hidrógeno se encuentran alineados y a menor distancia que en la forma paralela.

Las fuerzas que mantienen la estructura secundaria además de los enlaces peptídicos son principalmente la gran cantidad de puentes de hidrógeno que se forman entre los aminoácidos contiguos, lo que les da gran estabilidad a ambas estructuras.

En la estructura secundaria existe un elemento de conexión entre alfas hélices y/o laminas plegadas beta denominado giro beta, vuelta reversa o lazo

alzar, que no es tan al azar ya que se presenta como conexión cuando se observa el cambio a otra estructura secundaria o cambio de dirección en la propia cadena. En estos giros beta existe abundancia de Glicina, Asparagina, Prolina, Ácido aspártico y Serina, además se presentan puentes de hidrógeno para darle estabilidad.

Dominios y motivos

En 1973 Wetlaufer, definió dominios como unidades estables de estructura proteica que podían plegarse de forma autónoma, hoy en día un dominio se refiere a las zonas de la proteína que presenta mayor densidad de plegamientos y en general son zonas con actividades específicas que colaboran en equipo con la función de la proteína o presenta una estructura independiente.

Por ejemplo, las expansinas consisten de dos dominios que juntos forman la funcionalidad de la proteína. En el amino terminal se encuentra el dominio I que es rico en cisteínas y seis de estas son altamente conservadas formando tres puentes disulfuro (Fig.13.9a), este dominio tiene una alta similitud con el dominio EG45-like de las endoglucanasas, pero carece de una actividad catalítica neta. El plegamiento de este dominio, consiste de seis betas plegadas flanqueadas por lazos cortos y alfa-hélices. Hacia el extremo carboxilo terminal se encuentra el dominio II que funciona como de unión a celulosa; dominio CBD (por sus siglas en inglés). El plegamiento está conformado por ocho betas plegadas en antiparalelo, forma estructura de sándwich parecida (Fig. 13.9b) ambas figuras fueron tomadas de Prosite Expasy.

Los motivos dentro de las proteínas son secuencias de aminoácidos entre 2 a 20 aminoácidos que representan zonas conservadas que pueden tener o no alguna actividad. Estos motivos pueden estar presentes dentro de los dominios o en otras zonas de la proteína. Por ejemplo, en las expansinas se encuentra el motivo conservado HFD que está dentro del dominio I y es altamente conservados dentro de todas las expansinas; sin este motivo no se consideran expansinas. En la figura 13.5, se puede observar en la zona sombreada de gris el dominio I y en azul el dominio II, además, del motivo HFD en magenta.

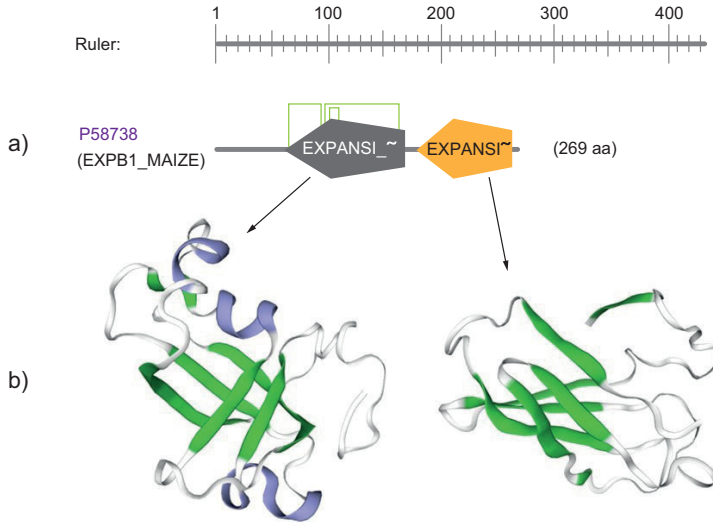


Figura 13.9. Arquitectura y estructura de dominios de la EXPB1_MAIZE. a) arquitectura de los dominios; el dominio I se muestra en gris con tres puentes disulfuro, el dominio II en naranja. Además, se aprecia la longitud de cada uno mediante la regla milimétrica. b) representa la estructura de ambos dominios; en verde las beta-plagadas y en lila las alfa-hélices (imagen tomada de Prosite y SWISS-MODEL: <https://swissmodel.expasy.org/>).

Estructura terciaria

En la estructura terciaria todas las alfas hélices y todas las hojas beta plegadas formadas con la secuencia de aminoácidos, interactúan en el espacio entre ellas de diferentes formas al parecer buscando la menor energía potencial almacenada que le dé estabilidad a la estructura terciaria formada, algunos ejemplos se observan abajo.

La estructura de una proteína dicta la función para la cual fue sintetizada, en otras palabras, la estructura de una proteína está íntimamente relacionada con su función o actividad que desempeña dentro del organismo. A partir de la estructura terciaria las proteínas ya presentan actividad o la función para la que fue sintetizada, mientras que las estructuras primaria y secundaria no tienen actividad o funcionalidad ya que aún no están del todo formadas.

Además de los enlaces peptídicos y los puentes de hidrógeno, en la estructura terciaria intervienen otras fuerzas que le dan estabilidad a la estructura de

las proteínas, como los puentes disulfuro entre cisteínas, las fuerzas hidrofílicas e hidrofóbicas que se generan con la interacción de las cadenas laterales de aminoácidos cercanos debido al plegamiento de la cadena.

En la figura 13.10, se observa la estructura terciaria de la EXPB1_MAIZE, donde se muestran los dominios estructurales y funcionales de las expansinas.

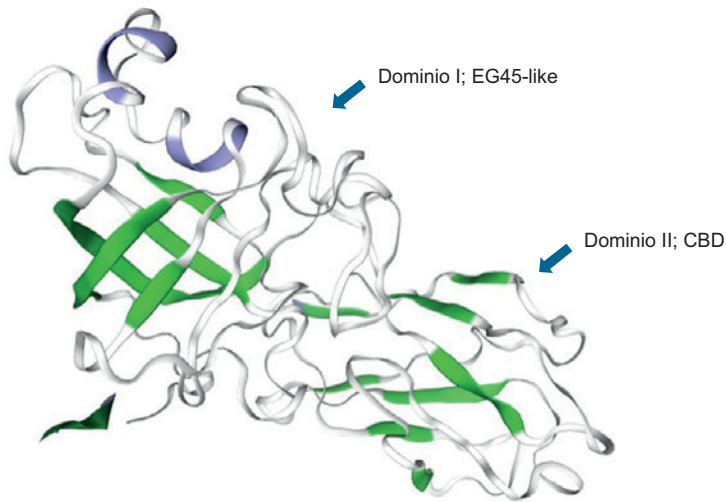


Figura 13.10. Estructura terciaria de la EXPB1_MAIZE. Donde se muestran ambos dominios representes en las expansinas (imagen tomada de SWISS-MODEL).

Estructura cuaternaria

Existen proteínas, por así decirlo, que no trabajan solas, es decir, para que puedan desarrollar la función para la cual fueron sintetizadas, deben trabajar en equipo desde dos hasta varias proteínas unidas, algunas siempre unidas y otras se unen cuando es necesario llevar a cabo su función. En caso de ser separadas no pueden realizar su función. Un ejemplo de estas proteínas es la catalasa formada por cuatro cadenas polipeptídicas. Además de todas las fuerzas antes mencionadas que les dan estabilidad a las proteínas, en la estructura cuaternaria intervienen las fuerzas de Van der Waals, que no son más que momentos dipolares.

Ejemplos funcionales de algunas proteínas

Enzimas

Las proteínas con actividad metabólica básicamente son enzimas que llevan a cabo las reacciones bioquímicas dentro de nuestras células con lo que se mantienen la homeostasis del organismo.

De esta manera las enzimas se clasifican en siete clases dependiendo de la actividad que realicen y en subclases que describen con mayor precisión dicha actividad.

La clasificación enzimática según la EC (Enzyme Classification, siglas en inglés), es de siete clases: EC 1, Oxidorreductasas; EC 2, Transferasas; EC 3, Hidrolasas; EC 4, Liasas; EC 5, Isomerasas; EC 6, Ligasas; EC 7, Translocasas (Enzyme Nomenclature - QM SBCS) Tabla 13.4.

Tabla 13.4 Clasificación de las enzimas según la EC.

Tipo de Enzimas	Función
EC1 Oxidorreductasas	Realizan reacciones de oxidorreducción
EC2 Transferasas	Transfieren grupos funcionales entre moléculas
EC 3 Hidrolasas	Rompen enlaces utilizando la molécula de agua para completar las moléculas resultantes
EC 4 Liasas	Rompen un enlace generando un dobles enlace
EC 5 Isomerasas	Genera isómeros
EC 6 Ligasas	Forman enlaces utilizando ATP
EC 7 Translocasas	Translocan sustancias a través de las membranas

Además, existe la subclasificación que detalla su modo de acción, por ejemplo, en la enzima piruvato deshidrogenasa (quinona) cuya clasificación es EC 1.2.5.1, el primer número indica la clase de enzima, en este caso 1 corresponde a una Oxidorreductasa. El segundo número que es el 2, indica que los donadores de electrones son el grupo aldehído o en el grupo oxo. El tercer número, en este caso el 5 indica que una quinona o un compuesto similar son el aceptor de electrones. El último 1 indica que se trata de Piruvato deshidrogenasa (quinona) (Fig. 13.10a).

Esta información se puede encontrar muy fácilmente en la base de datos BRENDA (<https://www.brenda-enzymes.org/oldstart.php>), además que proporciona la reacción de la enzima (Fig. 13.10b)

a) Information on EC 1.2.5.1 - pyruvate dehydrogenase (quinone)

1.2.5.1

EC Tree

- └ 1 Oxidoreductases
- └ 1.2 Acting on the aldehyde or oxo group of donors
- └ 1.2.5 With a quinone or similar compound as acceptor
- └ 1.2.5.1 pyruvate dehydrogenase (quinone)

b) Reaction Schemes [hide](#)

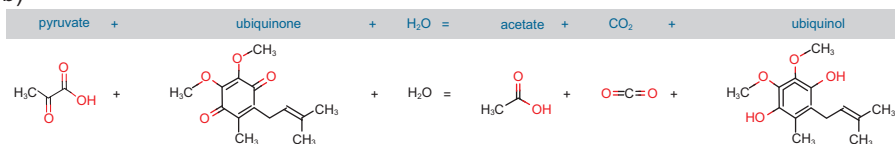


Figura 13.10. Identificación mediante EC de la piruvato-deshidrogenasa. a) clasificación enzimática EC 1.2.3.1, b) reacción que cataliza la 1.2.3.1.

Proteínas de membrana

Las proteínas de membrana consideradas estructurales, tienen múltiples funciones, por ejemplo, proteínas transportadoras como las que constituyen los canales y bombas que permiten o no el paso de diferentes sustancias a través de la membrana dependiendo de las necesidades de la célula.

Un ejemplo es la Palmitoil-transferasa DHHC (Asp-His-His-Cys), enzima de organismos eucariotas, específicamente enzima integral de membrana que catalizan la palmitoilación de proteínas, que es importante en una variedad de procesos fisiológicos, incluida la señalización de la pequeña guanósina trifosfatasa (GTPasa) y la adhesión celular (Fig. 13. 11a).

Proteínas de defensa contra estrés abiótico

Las proteínas de choque térmico o HSP (Heat shock proteins) son proteínas ubicuas que se encuentran en las células vegetales; descritas originalmente con relación al choque térmico. Sin embargo, se ha visto que estas proteínas son

inducidas por una amplia variedad de estrés como: calor, temperaturas altas, frío, luz ultravioleta, remodelación de tejidos entre otros. También actúan en la estabilización de proteínas y ayudan al replegamiento de proteínas en condiciones de estrés. En la figura 13.11b, se observa la estructura de la proteína HSP *Pinus tabuliformis*.

Las proteínas LEA (Late embryogenesis abundant) son grandes grupos de proteínas hidrofílicas que contribuyen en numerosos procesos de desarrollo y se acumulan en relación con la salinidad, la sequía, la congelación y los tratamientos con fitohormonas y ácido abscísico. Estas proteínas se dividen en 8 grupos de los cuales se conoce que el grupo II o dehidrinas (DHN) preservan las macromoléculas contra las lesiones causadas por la sequía, la salinidad y la congelación (Fig. 13.11c)

Proteínas de defensa contra estrés biótico y oxidativo

La heveína es un péptido de 43 aminoácidos, con actividad sobre la quitina. Esta proteína fue identificada en el látex del árbol del caucho, *Hevea brasiliensis*, siendo uno de sus componentes proteicos más abundantes. El mecanismo de acción de este péptido se basa en la inhibición del crecimiento de las hifas de los hongos al unirse a las quitinas de estos mismos, de igual manera, pueden causar filtración del material encontrado en el citoplasma al entrar en contacto con las hifas (Fig. 13.11d).

Las catalasas son proteínas con actividad enzimática que pertenecen a las oxidorreductasas y catalizan la descomposición del peróxido de hidrógeno (H_2O_2) en oxígeno molecular (O_2) y H_2O . Las catalasas protegen a las células de los efectos tóxicos del peróxido de hidrógeno. El peróxido de hidrógeno es una especie reactiva de oxígeno (ERO) y daña a varias macromoléculas tales como el ADN, proteínas, lípidos. El incremento de ERO en las plantas es cuando están bajo cierto tipo de estrés, en la figura 13.11e, se muestra la estructura de la catalasa de *Arabidopsis thaliana*.

Proteínas involucradas en la germinación de semillas

Las lipasas actúan principalmente en semillas oleaginosas, hidrolizan triglicéridos a ácidos grasos y glicerol y bajo ciertas condiciones catalizan la reacción inversa. Algunas de estas enzimas son también capaces de catalizar reacciones

de transesterificación e hidrólisis enantioselectivas. Específicamente la Fosfolipasa A1 específica de Sn-1 que cataliza el paso inicial de la biosíntesis de oxilipinas y jasmonato (JA) (Fig. 13.11f).

Proteína involucrada en la modificación de la pared celular

Las Expansinas son proteínas que participan en la expansión de la pared celular permitiendo que las células vegetales crezcan, participan en la maduración de frutos, en la elongación de la raíz y en el alargamiento celular.

Proteínas involucradas en la fotosíntesis

Las carboxilasas en plantas se relacionan con la fotosíntesis C4 y CAM, como la fosfoenolpiruvato-carboxilasa (PEPC), enzima citosólica que produce oxalacetato a partir de una reacción, es la primera enzima que fija el carbono en las plantas C4 (Fig.13.11g).

Enzimas generadoras de ATP

Enzimas que catalizan reacciones en las membranas como la cadena de transporte de electrones, realizando reacciones de óxido-reducción para generar un gradiente quimio-osmótico y de esta forma generar el ATP que requiere la célula.

La ATP sintasa es una de las maravillas del mundo molecular. La ATP sintasa es una enzima, un motor molecular, una bomba de iones y otro motor molecular, todo envuelto en una increíble máquina a nanoescala. Desempeña un papel indispensable en nuestras células, construyendo la mayor parte del ATP que alimenta nuestros procesos celulares. El mecanismo por el cual realiza esta tarea es una verdadera sorpresa. La síntesis de ATP se compone de dos motores rotativos, cada uno alimentado por un combustible diferente. El motor en la parte superior, denominado F₀, un motor eléctrico, está incrustado en una membrana y funciona con el flujo de iones de hidrógeno a través de la membrana. A medida que los protones fluyen a través del motor, hacen girar un rotor circular. Este rotor está conectado al segundo motor, denominado F₁. El motor F₁ es un motor químico, alimentado por ATP. Los dos motores están conectados entre sí por un estator, de modo que cuando F₀ gira, F₁ también gira (Fig. 13.11 h).

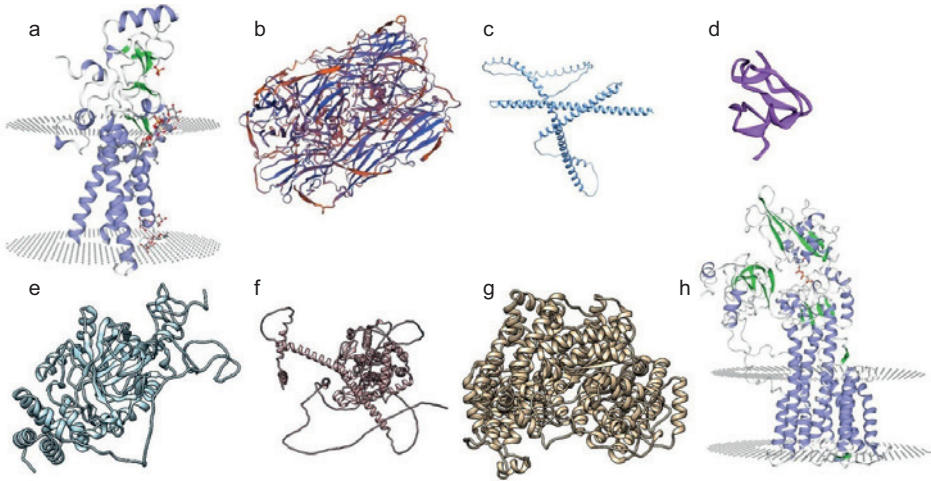


Figura 13.11. Estructuras de diferentes proteínas de plantas. a) Palmitoil transferasa (PDB:6bms.1), b) proteína HSP de *Pinus tabulaeformis* (PDB: 1gme), c) proteína LEA de *Oryza sativa* (Alpha fold: AF-A3AHG5-F1), d) Heveína, *Hevea brasiliensis* (PDB. 1hev), e) Catalasa CAT1 de *Arabidopsis thaliana* (Alpha fold: AF-Q96528-F1), f) Fosfolipasa A1 de *Arabidopsis thaliana* (Alpha fold: AF-F4HXL0-F1), g) Fosfoenolpiruvato carboxilasa de *Arabidopsis thaliana* (Alpha fold: AF-Q5GM68-F1), h) ATPase 10, plasma membrane-type de *Arabidopsis thaliana* (Q43128).

Maduración de proteínas después de la traducción

Las proteínas tienen un periodo de vida que inicia con su síntesis en los ribosomas y hasta su degradación cuando ya no son requeridas o están dañadas, durante este intervalo de vida, las proteínas pueden ser modificadas de diferentes formas para regular su función, su localización, su señalamiento de estabilidad o degradación al final de su periodo de vida.

Al ser sintetizadas en los ribosomas, las proteínas no adquieren su funcionalidad al instante, deben adquirir la estructura específica para realizar su función, es decir deben madurar y ser conducidas al lugar en el cual ejercerán su función y al llegar son activadas, estas modificaciones que se realizan en las proteínas durante y/o después de su traducción se les denomina modificaciones cotraduccionales o postraduccionales de las

proteínas. Las modificaciones que han sido las más estudiadas implican: correcto plegamiento mediante la formación de enlaces disulfuro en las posiciones precisas, fosforilación, hidroxilación, acetilación, glucosilación, adición de lípidos, ubiquitinación y escisión o ruptura proteolítica en posiciones específicas, dichas adiciones o modificaciones pueden ser permanentes o reversibles como una forma de regular su función o actividad.

Las modificaciones co o postraduccionales se pueden observar en las cadenas laterales de los aminoácidos, en los grupos amino terminal o carboxilo terminal de la proteína, con lo que se modifican los grupos funcionales, pudiendo ser como una forma de señalización, bien para modificar las propiedades de los aminoácidos, por ejemplo, haciéndolos más hidrofílicos o más hidrofóbicos y al modificar los grupos funcionales existentes o bien introducir nuevos grupos funcionales, la estructura de la proteína puede variar o modificar su estructura. Las modificaciones co o postraduccionales más estudiadas son las siguientes:

Puentes Disulfuro, denominados también como enlaces disulfuro, se generan entre dos aminoácidos de cisteína, el azufre del grupo Tiol (-SH) de ambos aminoácidos de cisteína se unen covalente (S-S) formando a la Cistina. Los puentes disulfuro tienen la función de generar la estructura adecuada en una misma cadena polipeptídica o bien mantener unidas las dos cadenas polipeptídicas para mantener la estructura y funcionalidad de la o las proteínas.

Fosforilación, es la unión covalente de un grupo fosfato ($-\text{PO}_4$) a la cadena lateral de un aminoácido mediante la acción de una enzima quinasa, tiene la función de modificar la estructura de una proteína para activarla, desactivarla o señalizarla, pudiendo ser la señal para el ensamblaje con otras proteínas. Se fosforilan principalmente los aminoácidos que tienen en su cadena lateral grupos hidroxilo (-OH), grupos amino (NH_3^-) o grupos carboxilo (COO^-).

Glicosilación o glucosilación, se refiere a la unión covalente de un glúcido a una proteína formando una glucoproteína, pudiendo ser desde un monosacárido hasta un oligosacárido, tiene la función de hacer más hidrosoluble a una proteína. Básicamente la glicosilación se realiza en el lumen del retículo endoplásmico rugoso a medida que la proteína va siendo sintetizada, para ser transportada a algún organelo específico, para formar parte de los lisosomas o bien para ser secretada por las vacuolas del retículo endoplásmico.

Hidroxilación, se refiere a la unión de grupos hidroxilo (-OH) a las proteínas por la acción de las enzimas denominadas hidroxilasas, con lo cual zonas de las proteínas aumentan su hidrofiliidad y por ende se pueden generar cambios ligeros en la estructura de la proteína.

Acetilación, es la unión de grupos acetilo ($\text{CH}_3\text{CO}-$) a la proteína por la acción de acetiltransferasas, principalmente a los grupos aminos de las cadenas laterales de la Arginina y Lisina ambos aminoácidos con carga positiva, que al ser acetilados se neutralizan sus cargas positivas y con ello la estructura de la proteína se ve modificada.

Adición de lípidos, se unen lípidos denominados también grupos acilo a las proteínas para generar lipoproteínas, entre otros el ácido palmítico (palmitilo) o el ácido mirístico (miristilo) con ello se les dota a las proteínas de hidrofobicidad, se observa sobre todo en proteínas que conformarán membranas.

Ubiquitinación, la ubiquitina es una proteína pequeña de 76 aminoácidos con peso molecular de 8.5 kDa y cuyo nombre indica que es ubicua, es decir, se encuentra ampliamente localizada en todas partes dentro del organismo, se une a las proteínas de diferentes formas desde la monoubiquitinación hasta la poliubiquitinación, ésta última puede ser lineal o ramificada como forma de señalar a aquellas proteínas que han de ser recicladas por las chaperonas o degradadas en proteosomas. Se han encontrado en eventos celulares como la endocitosis, apoptosis, diferenciación y desarrollo celular, morfogénesis, entre otros eventos celulares, habiendo mucho por esclarecer su participación en dichos eventos.

Relación entre estructura y función (actividad biológica). Estructura nativa y desnaturalización

Se denomina estructura nativa a la conformación estructural de las proteínas en la cual la proteína desempeña su función o actividad para la cual fue sintetizada. Cuando una proteína pierde su estructura funcional se encuentra desnaturalizada, la desnaturalización puede ser provocada por cambios de temperatura, cambios de pH, cambios de la fuerza iónica del disolvente en el cual se encuentra disuelta.

La desnaturalización en algunos casos puede ser reversible, si se elimina la causa que la ocasiona, si la desnaturalización es por alta concentración de sales se puede dializar la proteína para eliminar la alta concentración de sales y la proteína en un alto porcentaje puede regresar a su estado nativo, pero si la desnaturalización fue provocada por altas temperaturas o pH muy ácidos o básicos la proteína no regresa a su estado nativo. Por esta razón al purificar las proteínas se debe tener especial cuidado en no desnaturalizar a la proteína de forma irreversible sobre todo si se trata de una enzima cuya actividad sirve de seguimiento durante su purificación. En general las proteínas desnaturalizadas se precipitan y no pueden ser nuevamente solubilizadas.

Proteómica

Comencemos por entender que el vocablo “proteoma” se refiere al perfil de proteínas expresadas por el genoma de un organismo en estudio bajo condiciones específicas, en un momento dado. Es de entender que el perfil proteico de un organismo cambie de acuerdo con las condiciones bajo las cuales se desarrolla, desde condiciones del medio ambiente que le rodea, de la edad fisiológica en la que se encuentre, si se encuentra o no en estado de estrés, si se encuentra en estado de salud o enfermedad, si se encuentra bajo la presencia o ausencia de medicamentos o tóxicos, entre otros.

Entonces la “proteómica” es el estudio comparativo que permite observar los cambios generados en el proteoma, bajo condiciones específicas que permitan dilucidar los cambios o eventos fisiológicos bajo circunstancias específicas dadas, lo cual permite comprender la forma en la que el organismo responde, por ejemplo, bajo estrés comparado con el mismo organismo no estresado. En otras palabras, conocer que proteínas se expresan por ejemplo en estado sano y compararlas con las proteínas que se expresan en estado de enfermedad, las proteínas diferentes serían las que marcan la diferencia y por lo tanto hay que estudiar la función que desempeñan para comprender el papel que juegan en el estado de enfermedad.

Una gran parte de la investigación en ciencias de la vida se encuentra enfocada al entendimiento de la compleja relación entre el genotipo y el fenotipo. Especialmente la investigación se dirige a preguntas fundamentales del cómo, cuándo y dónde la información codificada en el genoma de un organismo es

expresada y modulada por factores externos (ejemplo, medioambiente) o factores internos (ejemplo, genómica) para generar un fenotipo específico.

Bioinformática de proteínas

El propósito de este tema es dar a conocer y guiar al estudiante o lector con las herramientas bioinformáticas disponibles gratuitamente en internet y que son utilizadas para caracterizar computacionalmente una proteína completa a partir de una secuencia de ADN o de su gen correspondiente. Esta guía implicará buscar en el GenBank la secuencia de interés, el análisis con diferentes herramientas de ExPasy; traducción virtual (proteína), búsqueda de dominios, proteínas homólogas, propiedades fisicoquímicas y un modelo 3D en la base de datos de SWISS-MODEL

La manera o las diferentes herramientas en las bases de datos utilizadas para realizar estos análisis bioinformáticos, depende del estudiante, investigador o persona que lo realiza ya que existe diferentes bases de datos que llevan al mismo resultado. En la figura 13.12, el diagrama de flujo muestra una alternativa de las rutas bioinformáticas que utilizaremos para el análisis de la proteína en estudio. Se inicia con una secuencia de nucleótidos conocida (cuadro en blanco), las figuras en color azul indican que se usó la misma base de datos con diferentes herramientas para obtener los resultados, y en amarillo para el modelado en SWISS-MODEL

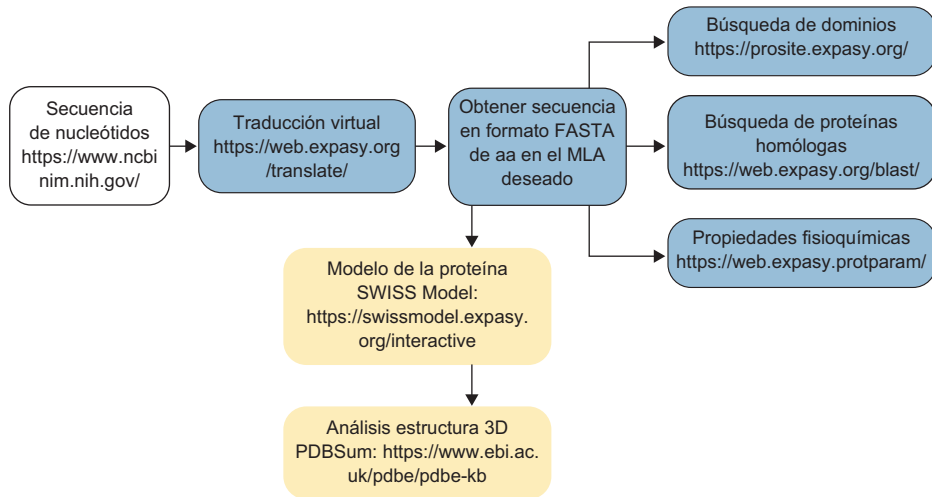


Figura 13.12. Diagrama de flujo de una ruta para el análisis bioinformático de proteínas. La figura en color blanco indica el inicio con la secuencia de nucleótidos, las figuras en azul indican el uso de las herramientas en ExPASy, en amarillo las bases de datos para el modelado.

Análisis bioinformático

Búsqueda del gen de interés

Para este ejercicio, se trabajará con una secuencia de un ARN mensajero que codifica para una alfa expansina de jitomate: *Lycopersicon esculentum expansin*; LeEXP1 (U82123.1). Existen diferentes maneras de obtener la secuencia de nucleótidos en la base de datos de NCBI tal como: nombre del gen, número de acceso, nombre de la proteína y nombre científico del organismo. Hay que mencionar que para obtener cualquier información dentro de NCBI se tiene que escribir en el idioma inglés. Para este caso, utilizaremos la frase “*Lycopersicon esculentum expansin*”.

Paso 1.- Búsqueda del gen en estudio. Ingresar a la base de datos de la NCBI (<https://www.ncbi.nlm.nih.gov/>), y en el cuadro de búsqueda escribir *Lycopersicon esculentum expansin* (Fig. 13.13, círculo rojo) y dejar la opción All Databases posteriormente seleccionar Search.

Al seleccionar All Databases, se obtiene toda la información relacionada con el gen en estudio como: artículos, libros, genes, proteínas, genomas, nucleótidos (Fig.13.14)

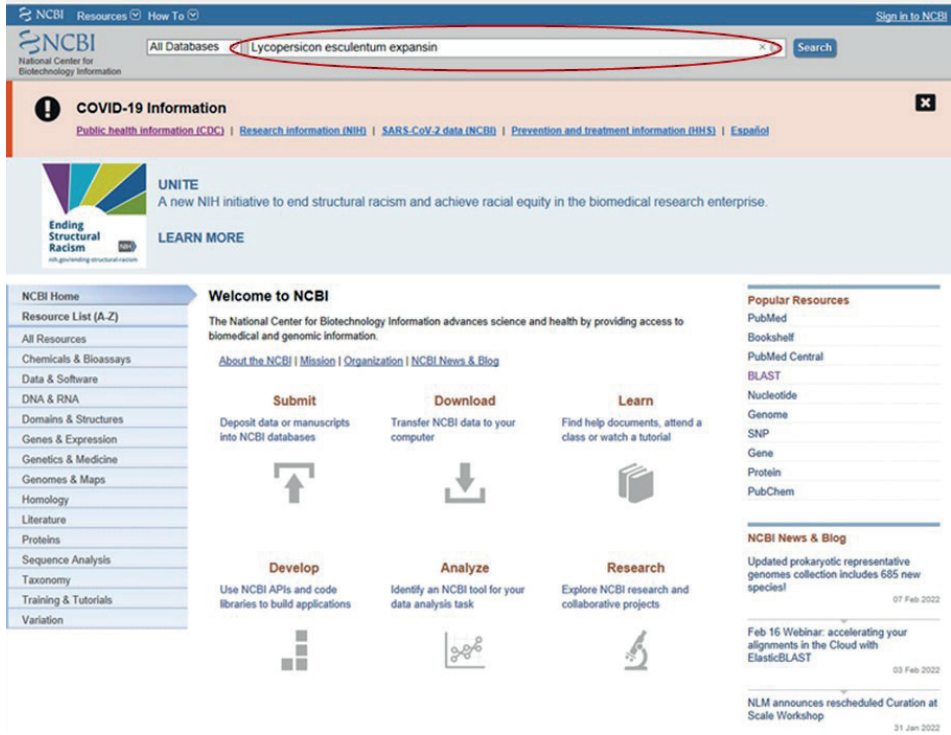


Figura 13.13. Página Principal de NCBI. En el círculo rojo, la frase del gen a buscar.

Paso 2.- Selección del gen. Seleccionar Nucleotide en la sección de genomes (figura 13.14) Nos interesa las secuencias de nucleótidos que pueden ser desde RNAm parciales o completos hasta fragmentos de DNA y este apartado se pueden encontrar 74 secuencias de expansinas relacionadas con jitomate.

NIH National Library of Medicine
National Center for Biotechnology Information

Search NCBI: Lycopersicon esculentum expansin

COVID-19 Information
Public health information (CDC) | Research information (NIH) | SARS-CoV-2 data (NCBI) | Prevention and treatment information (HHS) | Español

Did you mean [Lycopersicon esculentum expansion](#)?

Results found in 13 databases

Literature	Genes	Proteins
Bookshelf: 0	Gene: 40	Conserved Domains: 0
MeSH: 3	GEO DataSets: 0	Identical Protein Groups: 15
NLM Catalog: 10	GEO Profiles: 16	Protein: 85
PubMed: 45	HomoloGene: 0	Protein Family Models: 0
PubMed Central: 818	PopSet: 0	Structure: 0

Genomes	Clinical	PubChem
Assembly: 8	ClinicalTrials.gov: 0	BioAssays: 0
BioCollections: 0	ClinVar: 0	Compounds: 0
BioProject: 0	dbGaP: 0	Pathways: 0
BioSample: 16,046	dbSNP: 0	Substances: 0
Genome: 1	dbVar: 0	
Nucleotide: 74	GTR: 0	
SRA: 16,928	MedGen: 0	
Taxonomy: 0	OMIM: 0	

Figura 13.14. Resultados de la búsqueda en all database. La flecha rojo indica la selección de nucleótidos.

En la figura 13.15, se muestra el resultado de esta selección y se desglosa las 74 secuencias en nucleótidos.

The screenshot shows the NCBI Nucleotide search results for the query "Lycopersicon esculentum expansin". The search results are displayed in a list format, with four entries shown. Entry 4 is selected, and its details are visible. The entry is "Lycopersicon esculentum expansin (LeEXP1) mRNA, complete cds" with accession number U82123.1. The interface includes a search bar at the top, a sidebar with navigation options, and a main content area with filters and search results. A search box highlights the query: "See LOC4335222 (EXPANSIN) expansin-A1-like in the Gene database expansin reference sequences Transcript (1) Protein (1)".

Search Results Summary:

- Species: Plants (73), Bacteria (1), Customize...
- Molecule types: genomic DNA/RNA (29), mRNA (45), Customize...
- Source databases: INSDC (GenBank) (46), RefSeq (28), Customize...
- Sequence Type: Nucleotide (58), EST (16)
- Sequence length: Custom range...
- Release date: Custom range...
- Revision date: Custom range...

Search Results List:

- [Lycopersicon esculentum expansin precursor \(EXPA5\) mRNA, complete cds](#)
1. 1,225 bp linear mRNA
Accession: AF059489.1 GI: 4138915
[PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Lycopersicon esculentum expansin precursor \(EXPA4\) mRNA, complete cds](#)
2. 1,239 bp linear mRNA
Accession: AF059488.1 GI: 4138913
[PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Lycopersicon esculentum expansin precursor \(EXPA3\) mRNA, complete cds](#)
3. 1,046 bp linear mRNA
Accession: AF059487.1 GI: 4138911
[PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Lycopersicon esculentum expansin \(LeEXP1\) mRNA, complete cds](#)
4. 1,070 bp linear mRNA
Accession: U82123.1 GI: 2062420
[PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

Filters: Manage Filters

Results by taxon: Top Organisms [Tree] (44)
Solanum lycopersicum (44)
Arabidopsis thaliana (12)
Coffea arabica (8)
Nicotiana benthamiana (2)
Gossypium hirsutum (2)
All other taxa (6)
More...

Find related data: Database: [Select] Find items

Search details: ("Solanum lycopersicum"[Organism] OR Lycopersicon esculentum[All Fields]) AND expansin[All Fields] Search See more...

Recent activity: Turn Off Clear
Lycopersicon esculentum expansin (74) Nucleotide

Figura 13.15. Resultados de la selección de nucleótidos. Se selecciona la opción 4 que es del gen de interés.

Paso 3. Selección de la secuencia deseada. Como nos interesa la LeEXP1, seleccionamos la secuencia 4 (Fig. 13.15). En la figura 13.16, se presenta la información completa del gen LeEXPA1 y se muestra tanto la secuencia del RNAm en nucleótidos como de la proteína en aminoácidos. Existen dos maneras de copiar la secuencia; 1) copiando directamente la secuencia de nucleótidos y pegarla en un documento Word o 2) ir a la pestaña de FASTA (círculo rojo) y se obtiene esta secuencia en este formato. Como se dará cuenta en la figura 13.15, también aparece el número de acceso del gen que es U82123, y también la secuencia de su proteína con número de acceso: AAC63088.1. Con este número de acceso el lector puede ingresar directamente a la información del gen

o de la proteína en la base de datos de NCBI. Como ya se tiene la secuencia de aminoácidos, se puede trabajar directamente para realizar el modelo 3D de la proteína, pero en este ejercicio se hará por otra vía.

The screenshot displays the GenBank entry for **Lycopersicon esculentum expansin (LeEXP1) mRNA, complete cds**. Key details include:

- LOCUS:** LE182123 1070 bp mRNA linear P11 07-OCT-1998
- DEFINITION:** Lycopersicon esculentum expansin (LeEXP1) mRNA, complete cds.
- ACCESSION:** U82123
- VERSION:** U82123.1
- SOURCE:** Solanum lycopersicum (Lycopersicon esculentum)
- REFERENCE:** 1 (bases 1 to 1070) Rose, J.K., Lee, H.H., and Bennett, A.B. Expression of a divergent expansin gene is fruit-specific and ripening-regulated. Proc. Natl. Acad. Sci. U.S.A. 94 (11), 5955-5960 (1997)
- FEATURES:**
 - source:** 1..1070 /organism="Solanum lycopersicum" /mol_type="mRNA" /cultivar="Castlemar" /db_xref="taxon:4081" /tissue_type="fruit"
 - gene:** 1..1070 /gene="LeEXP1"
 - cds:** 28..813 /gene="LeEXP1" /note="fruit ripening regulated expansin" /codon_start=1 /product="expansin" /protein_id="AAC63088.1" /translation="MGIIIFILVLFVDFCFNIDVEGRIPGVYGGGSHETNATFYGGSDASGTMSGKGVRLYSQGVVNTALSTALFMSLSCGACFLKCTNTPMNAUCLPQNPSSLLITATNFPPNYALPNDGMCNPPPHFDLAWPHLKLQZYRAGLVVYVRI PCRKQGGIRFINGFRYFNLVLTNAGAGDIKVAAGKTKN4IPLSRNAGQPMQSN AVL7GQSLFRVKAASHSS15MNPVSH4QGGIFIGKFK1"
- ORIGIN:**

```

1 gaaattcaat tccattaast cttagaagt ggtatcataa tttctacct tttctctct
81 ttgtgagct catgttcaa cattgtgaa gaaagatcc ctgtgttga cctcaggagt
121 tcatgagaaa ctgacatgc tacatttacc ggcgagtag atgtctctga acaaatgag
181 ggtcgtgtg gttatgaaa ttatacagc caagatagc gatttaaac agcagcactg
241 agtactcctt tttttaaca tggattaat tgtgacctt gtttgtaact taatgtaca
301 aatactccca attggaagt gttctctct gaaacctt ccatttaac cacagtacc
361 aattctgcc caccaaatta cgcgtgcc aagcaaat gttgctgtg taacctctt
421 gcccctcaat tgcctctgc tatgctatg tttctaac tttctcaga cgcgctagg
481 atgtctcag taactatgc cagatcca tgcgcaagc agggagaat cagattacc
541 atcaatgat tccgttact caacttag tgcacaga atgacagc tcaagagat
601 attattaagg tttggtaaa agaaacaa acaattga ttcctatg ccgtaattg
661 gpaacaatt gccaatcaa tgcgcttta actgctcaat cactctctt cagatgtaa
721 gctagtacc atgatcttc tactcatg aatagctt ctctcaatg gcaatttgc
781 caactctca tggaaaga tttcaaaa taanaatg agggatag ttattttaa
841 tttgagaaa aactagata tttcagatg tttctacct taggaaga aatcagctt
901 tcaatgaaa tccagataga taataata aactcaaa attttgat attttgat
961 gttatcaac atttaacct aagtatgtt aatggagag aaggttga gttgctcga
1021 aatcagcag cccgacgtg tttttttt ttatacaat acatcaag
//

```

Figura 13.16. Información completa del gen LeEXP1 en la base de datos de NCBI.

Paso 5.- Traducción virtual. Se realizará con la herramienta Translate de Expsy (<https://web.expsy.org/translate/>). Esta herramienta es muy sencilla de utilizar

ya que solo se pega la secuencia y se selecciona el formato de salida y el código genético deseado. Para este caso, se insertó la secuencia de nucleótidos de LeEXP1, se seleccionó el formato de salida compacto y el código genético estándar. El resultado es la traducción en seis marcos de lectura abierta (Fig.13.17a). El programa por default marca todas las metioninas (M) en color rojo indicando que es el inicio de la proteína u otro aminoácido después del gap indicando un marco de lectura. El lector debe de seleccionar el marco de lectura correcto que en este caso es el primer marco. Si selecciona la primera metionina del primer marco (flecha amarilla en el primer marco, figura 13.17a) y al dar clic se despliega la secuencia en aminoácidos en formato FASTA y tres herramientas como: BLAST, ScanProsite y Squence analysis tools (Fig.13.17b).



Figura 13.17. Traducción virtual de la secuencia de nucleótidos de LeEXP1 (Expassy-Translate; <https://web.expasy.org/translate/>). a) se presenta los seis marcos de lectura abierta (MLA), en el primero se presenta con una flecha la M que es aminoácido de inicio de la proteína. b) la secuencia en formato fasta y las diferentes herramientas como: BLAST, ScanProsite.

Paso 6.- Búsqueda de proteínas homólogas (<https://web.expasy.org/blast/>). En la figura 13.17b, se observa en la parte superior derecha un recuadro en verde con la palabra BLAST, al posicionar en éste y dar clic se hará una búsqueda de proteínas homólogas y un gráfico con la arquitectura de los dominios (Fig.13.18).

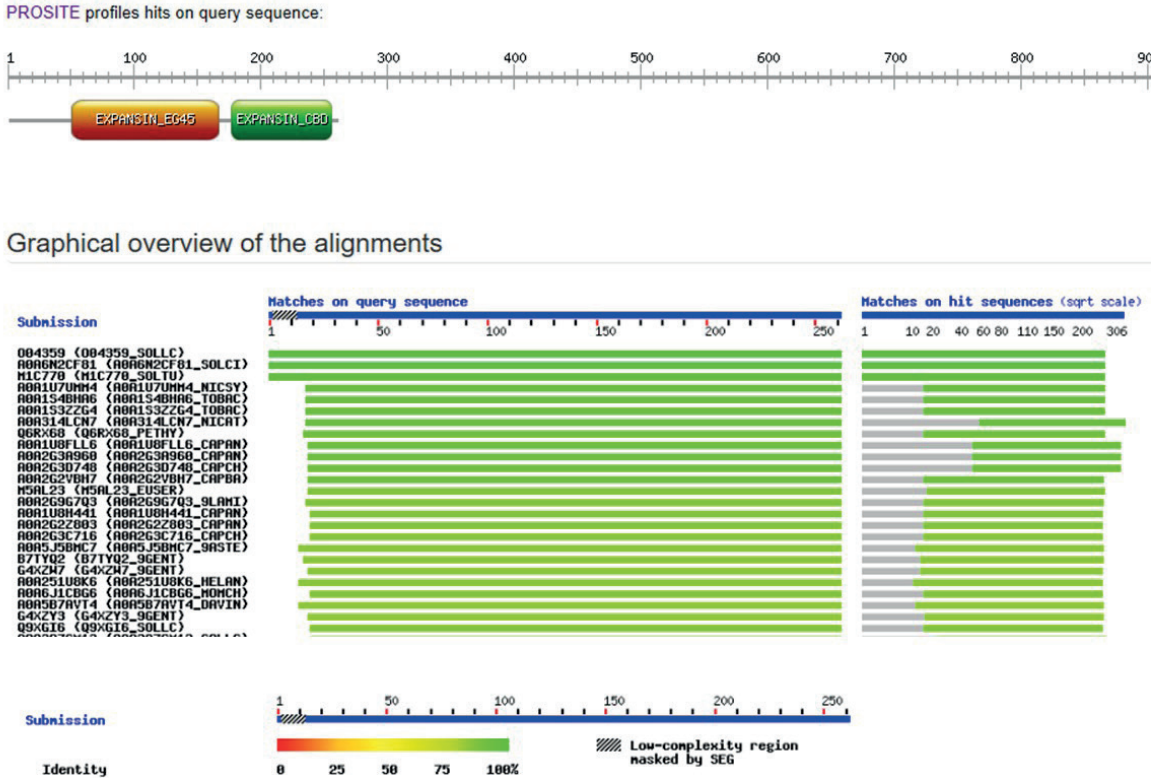


Figura 13.18. Arquitectura de dominios y secuencias homólogas de LeEXP1. de dominios b) secuencias homólogas a LeEXP1.

Paso 7.- Búsqueda de dominios (<https://prosite.expasy.org/>). Con los resultados de la figura 13.17b, en la pestaña de ScanProsite, se realiza la búsqueda de dominios que contiene la proteína. Los resultados es una arquitectura gráfica

de los dominios, así como la secuencia de aminoácidos que los conforman (Fig.13.19). En la secuencia de aminoácido de la figura 13.19a se observa en amarillo el dominio I, y en 13.19b la arquitectura de dominios.



Figura 13.19. Secuencia de dominios y su arquitectura. a) secuencia en aminoácidos de Le-xEXP1, en amarillo el dominio I, b) Arquitectura de dominios.

Paso 8.- Propiedades fisicoquímicas de la proteína (<https://web.expasy.org/protparam/>). Este tipo de análisis es de gran importancia para las características de las proteínas, entre estos están el peso molecular, el punto isoelectrico, GRAVY, índice alifático En la misma herramienta de Translate (Fig.13.17b) se puede observar la pestaña Sequence analysis tools, esta contiene las herramientas ProtParam, ProtScale, Compute pI/Mw, PeptideMass, y PeptideCutter. En la tabla 13.5, se muestra los valores obtenidos en ProtParam

Tabla 13.5 Parámetros fisicoquímicos de la proteína LeEXPA1

Numero de aminoácidos	261
Peso Molecular	28686.94 Daltons
pI teórico	9.46
Total, de aa negativos	(Asp + Glu): 9
Total, de aa positivos	(Arg + Lys): 21
Formula	$C_{1304}H_{1966}N_{352}O_{354}S_{14}$
Índice de inestabilidad	26.23
Índice alifático	77.36
Gran promedio de hidropaticidad (GRAVY)	-0.005

Paso 9.- Modelado de la proteína. Ingresar a la base de dato del SWISS-MODEL (<https://swissmodel.expasy.org/>) y pegar la secuencia de aminoácidos en el recuadro donde dice: “paste your target sequence” (Fig.13.20). Se tienen dos opciones para el modelado: 1) que se realice a partir de moldes o templates existentes en la base de datos de PDB o 2) construir el modelo. Para este caso, se tomó la primera opción y el resultado es con 41 moldes para la proteína que estamos analizando (Fig.13.21). Se seleccionó la primera opción que es con el molde 1n10.1 que es el número en el PDB y es de la proteína alergeno del polen Phl p1. Como resultado, se observa la estructura 3D, así como la secuencia en aa y las zonas que forman las beta plegadas y alfa-hélices.

Esto es un ejemplo para la expansina, pero el lector lo puede hacer para la proteína de interés.

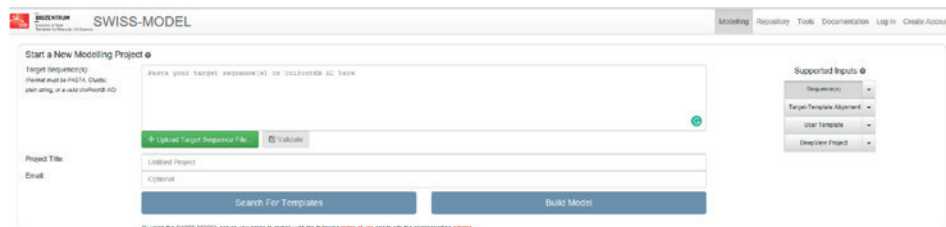


Figura 13.20. Página principal de SWISS-MODEL para el análisis del modelo.

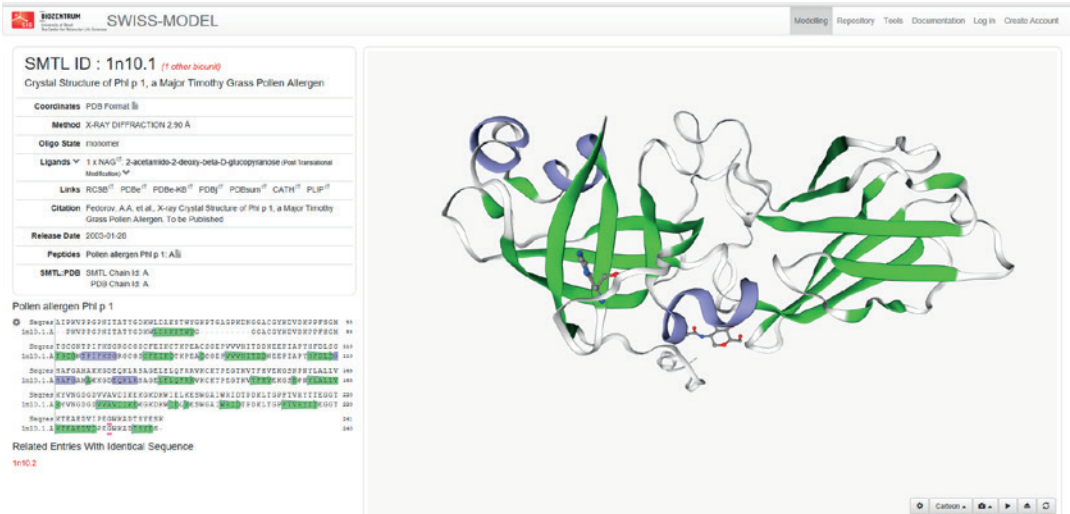


Figura 13.21. Modelado 3d de la proteína LeEXPA1 en SWISS-MODEL (<https://swissmodel.expasy.org/>).

Protocolos

Protocolo de extracción de proteínas de Leguminosas

Las proteínas presentes en las semillas pueden ser divididas en dos grandes grupos, dependiendo de la función biológica de la misma: a) proteínas con actividad biológica como las enzimas y b) proteínas de reserva. Estas últimas llegan a conformar más del 80 % de la proteína total de la semilla.

La metodología que se muestra a continuación es para extraer la proteína total de las semillas o granos. Se recomienda trabajar muestras de leguminosas, tales como frijol, soya, garbanzo, lenteja, chícharo, etc.,

Buffer de extracción

Tris-HCl 0.0625 M a pH 6.8
 SDS 0.1%

Procedimiento

1. Moler algunas semillas en mortero o molino eléctrico hasta formar una harina fina.
2. Pesar 150 mg de harina y transferir a un tubo para microcentrífuga.
3. Agregar 1 ml de buffer de extracción.
4. Homogenizar con vortex. Agitar con vortex cada 5 min durante media hora y dejar en hielo.
5. Centrifugar por 10 min a 10,000 rpm
6. Recuperar el sobrenadante en un tubo limpio.
7. Centrifugar nuevamente el sobrenadante por 10 min a 10,000 rpm
8. Recuperar el sobrenadante en tubo limpio.

Protocolo de Cuantificación de la proteína.

La absorbancia a 280 nm, lee la absorbancia de los aminoácidos aromáticos como la fenilalanina, tirosina y triptófano presentes en las proteínas. Obviamente las proteínas que tengan mayor cantidad de los aminoácidos mencionados tendrán mayor absorbancia que las proteínas que contienen menor cantidad de dichos aminoácidos.

Por la ley de Beer y Lambert, se tiene: $A = KCl$

Donde, A= absorbancia

C= concentración

K= coeficiente de extinción

l= longitud en centímetros de la celda utilizada, las celdas que se utilizan tienen un tamaño estándar de 1 cm

Procedimiento

Medir absorbancia a 280 nm de las muestras, calcular su concentración con un coeficiente de extinción de 1.15.

Cuantificación de proteínas por método DC-Protein assay (Bio-Rad)

Leer el manual del kit, resolver dudas con la profesora, elaborar la curva patrón con seroalbúmina bovina (BSA) y cuantificar la concentración de proteína de sus muestras.

Comparar las dos metodologías para cuantificar proteína, describir las ventajas y desventajas de ambos métodos y sacar sus propias conclusiones.

Protocolo para Electroforesis SDS PAGE (Polyacrylamide Gel Electrophoresis, por sus siglas en inglés)

Soluciones a preparar

Solución de acrilamida /bis-acrilamida 30% (29.2% / 0.8%) Tóxico usar guantes y tapaboca, no inhalar.

Tris-HCl 1M, pH 8.8, 100 ml

Tris-HCl 1M, pH 6.8, 100 ml

SDS 10 %, 10 ml

Persulfato de amonio 10% (prepararlo en el momento de utilizarlo), 1 ml

Solución de agarosa al 1%, 100 ml, calentar para solubilizar

Azul de bromofenol al 1 %, 10 ml

Buffer para muestra

0.5 M Tris HCl a pH de 6.8

1 ml

Glicerol

0.8 ml

SDS al 10 %

1.6 ml

β -mercaptoetanol

0.4 ml

Azul de bromofenol al 1 %

0.4 ml

Mezclar todo y aforar a 8 ml con agua desionizada

Buffer de corrida 5X

Tris-HCl	15 g
Glicina	72 g
SDS	5 g
Aforar a 1 L	

Solución de tñido

40 % de Metanol
10 % de Ácido acético
0.1 % Azul de Coomassie R-250
50 % de Agua
Preparar 250 ml

Solución de desteñido

40 % de Metanol
10 % de Ácido acético
50 % de Agua
Preparar 1 L

Preparación del gel:

1. Limpiar las superficies del vidrio con alcohol y montar la cámara de electroforesis.
2. Preparar el gel separador de poliacrilamida al 12 %

Tris pH 8.8	1.88 ml
Poliacrilamida	1.67 ml
SDS al 10 %	0.025 ml
Agua desionizada	1.39 ml
Mezclar perfectamente y desgasificar	
Persulfato de amonio al 10 %	50 μ l
TEMED	5 μ l

- Mezclar perfectamente sin ingresar burbujas
3. Vaciar el gel separador en la cámara de electroforesis dejando 2 cm antes del borde.
 4. Cubrir la superficie con 1 ml de alcohol y retirarlo cuando polimerice el gel.
 5. Preparar el gel concentrador.

Tris pH 6.8	312.5 μ l
Poliacrilamida	312.5 μ l
SDS al 10 %	25 μ l
Agua desionizada	1.84 ml
Mezclar perfectamente y desgasificar	
Persulfato de Amonio al 10 %	7.5 μ l
TEMED	2 μ l
Mezclar perfectamente sin ingresar burbujas	

6. Vaciar el gel en la cámara sobre el gel separador y colocar el peine. Esperar a que polimerice.

Preparación de las muestras y del estándar de peso molecular.

7. Tomar 20 μ l de su muestra y mezclarlos con 20 μ l de buffer de muestra en un tubo eppendorf, agitar con vortex.
8. Preparar el estándar de peso molecular como se indica en las instrucciones del proveedor.
9. Incubar las muestras bien tapadas y en su caso el estándar por 3 min en agua hirviendo, sacar y dejar enfriar.

Preparación de la cámara de electroforesis para la corrida:

10. Agregar el buffer de corrida en la cámara.
11. Cargar el gel con las muestras y el estándar de peso molecular.
12. Correr el gel a 20 mili Amperes hasta que el frente entre al gel separador.
13. Correr el gel a 40 mili Amperes en el gel separador y dejar correr hasta 0.5 mm antes de llegar al final.
14. Apagar y desconectar la fuente de poder. Desmontar el gel.
15. Teñir el gel durante 2 h, lavarlo con agua corriente.

16. Desteñir hasta que las bandas sean visibles.

A reportar:

1. Número de bandas y abundancia relativa de cada una de ellas.
2. Similitudes y diferencias entre los patrones electroforéticos de sus proteínas y de las proteínas de otros equipos.
3. Peso molecular calculado de cada una de las 5 principales bandas de cada muestra. Graficar en Excel el log 10 del peso molecular de los estándares en el eje de las "Y" y el Rf en el eje de las "X", obtener la ecuación de la curva y calcular los pesos moleculares de las 5 principales bandas de sus dos muestras.

PM de cada proteína del estándar

Eje Y $\log_{10}(\text{PM})$

Eje X Rf de esa proteína estándar

94,000 $\log_{10}(94,000) = 4.973$

El siguiente estándar

36,000 $\log_{10}(36,000) = 4.55$

Se deben ingresar todos los estándares de peso molecular del estándar utilizado en el gel.

Con esto elabora una gráfica como esta, en el eje de las X pones el Rf

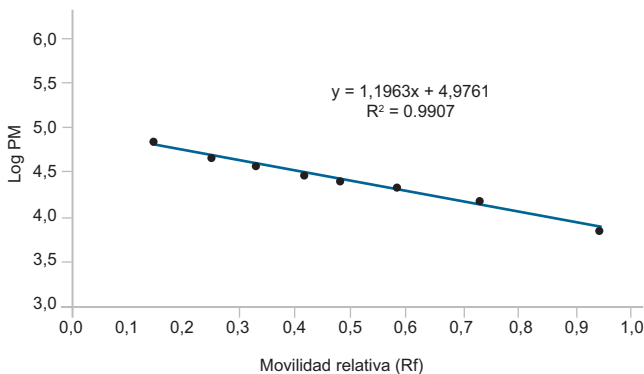


Figura 13.22. Curva de calibración para la determinación de pesos moleculares.

Si lo haces en Excel, puedes obtener la ecuación de la recta y la R cuadrada mientras más cercana a 1 significa que hay menos error.

$Y = mX + b$ en donde Y es el Rf de las bandas que se quiere calcular el peso molecular y la X el Rf obtenido para la banda que se quiere conocer su peso molecular.

Como ya conoces el valor de “m” que es la pendiente y conoces el valor de “b” que es la ordenada al origen, sólo haces la operación y obtienes el peso molecular de la banda de la cual no se conoce su peso molecular.

Ejemplo de la electroforesis realizada con proteínas totales de leguminosas

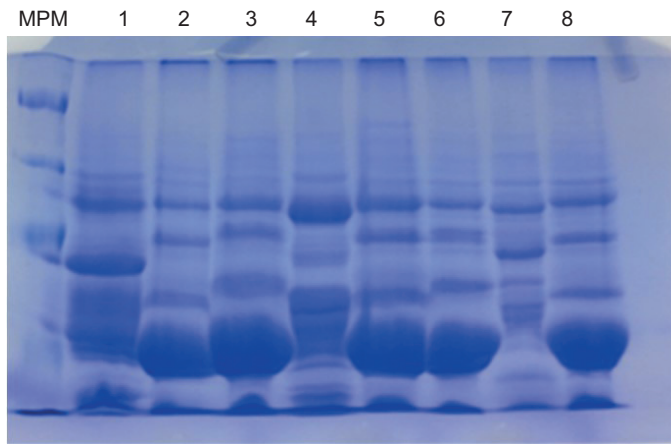


Figura 13.23. Electroforesis de proteínas totales de leguminosas. MPM, Marcado de peso molecular. 1) *Lens culinaris* (Lenteja), 2) *Cicer arietinum* (Garbanzo), 3) *Phaseolus vulgaris* (Frijol bayo), 4) *Phaseolus vulgaris* (Frijol pinto), 5) *Vicia faba* (Haba), 6) *Phaseolus vulgaris* (Alubia), 7) *Phaseolus vulgaris* (Frijol peruano 1), 8) *Phaseolus vulgaris* (Frijol peruano 2).

Bibliografía

- Bensimon, A. Heck, A. A. J. R. and Ruedi Aebersold, R. (2012). Mass Spectrometry–Based Proteomics and Network Biology. *Annu. Rev. Biochem.*, 81:379-405. Downloaded from www.annualreviews.org. Access provided by Universidad Autónoma de Aguascalientes on 02/18/20. For personal use only.
- Chaperoninas: su papel en el plegamiento de proteínas y su relación con enfermedades. <http://webquestcreator2.com/majwq/ver/ver/7919> (28 enero 2022).
- Cothia, C. and Finkelstein, A. V. (1990). The classification and origins of protein folding patterns. *Annu. Rev. Biochem.*, 59:1007-1039.
- Enzyme Nomenclature - QM SBCS. www.sbcqs.qmul.ac.uk/iubmb/enzyme (18 febrero 2020).
- Estructura de proteínas. <http://www.iib.unsam.edu.ar/archivos/docencia/licenciatura/biotecnologia/2018/BioProt/1527182496.pdf> (17 enero 2022).
- FAO,OMS,UNU. (1985). Necesidades de energía y de proteínas, Informe de una Reunión Consultiva Conjunta FAO/OMS/UNU de Expertos, Serie de Informes Técnicos, No.724, 56-68, OMS, Ginebra.
- Herráez, A. Biomodel. UAH. es · LicenciaCC-by-nc-sa. <http://biomodel.uah.es/model1j/prot/aa-props.htm> (12 enero 2022).
- <https://israelmasa.files.wordpress.com/2013/10/tema-4a-proteinas-estructura-2a-farmacia.pdf> (17 enero 2022).
- Las Proteínas. Características generales. <https://docplayer.es/111245317-Caracteristicas-generales.html> (13 enero 2022).
- Leser, S. (2013). The 2013 FAO report on dietary protein quality evaluation in human nutrition: Recommendations and implications. *Nutrition Bulletin*, 38(4), 421–428. <https://doi-org.dibpxy.uaa.mx/10.1111/nbu.12063>
- Luque Gillen, M. V. (Sin fecha) Estructura y propiedades de las proteínas. Master Ingeniería Biomédica. https://www.uv.es/tunon/pdf_doc/proteinas_09.pdf
- Marioni, P. (sin fecha). Tema: Proteínas. Materia: Química aplicada. Escuela Nro. 537. Material producido en el Laboratorio Pedagógico. Localidad: Coronda.
- Millar, A.H., Heazlewood, J.L.,Giglione, C., Holdsworth, M. J., Bachmair, A. Waltraud X. & Schulze, W. X. (2019). The Scope, Functions, and Dynamics

- of Posttranslational Protein Modifications. *Annual Review of Plant Biology*. 70:119-151. <https://doi.org/10.1146/annurev-arplant-050718-100211>
- Ponting, C. P. and Robert R. Russell, R. R. (2002). The Natural History of Protein Domains. *Annual Review of Biophysics and Biomolecular Structure*. Vol. 31:45-71 (Volume publication date June 2002). <https://doi-org.dibpxy.uaa.mx/10.1146/annurev.biophys.31.082901.134314>
- RTECH-El enlace peptídico. <https://scottxrt.wordpress.com/2018/12/09/el-enlace-peptidico/> (12 enero 2022).
- Tandang-Silvas, M. R. G., Tecson-Mendoza, E. M., Mikami, B., Utsumi, S., and Maruyama, N. (2011). Molecular Design of Seed Storage Proteins for Enhanced Food Physicochemical Properties. *Annual Review of Food Science and Technology*. Vol. 2:59-73 (Volume publication date April 2011). First published online as a Review in Advance on October 4, 2010. <https://doi-org.dibpxy.uaa.mx/10.1146/annurev-food-022510-133718>
- Zhang, Y. *et al.* Vista estructural tridimensional de la red metabólica central de *Thermatoga maritima*. *Ciencia* 325, 1544-1549 (2009). (PDB).
<https://web.expasy.org/translate/>, visitada por última vez el 15 de marzo de 2022.
<https://www.ncbi.nlm.nih.gov/>, visitada por última vez el 15 de marzo de 2022.
<https://www.rcsb.org/>, visitada por última vez el 15 de marzo de 2022.
<https://prosite.expasy.org/>, visitada por última vez el 15 de marzo de 2022.
<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html>, visitada por última vez el 15 de marzo de 2022.
<https://www.brenda-enzymes.org/>, visitada por última vez el 15 de marzo de 2022.
<https://www.ebi.ac.uk/thornton-srv/databases/cgibin/pdbsum/GetPage.pl?pdbcode=2hcz>, visitada por última vez el 15 de marzo de 2022.
<https://www.brenda-enzymes.org/oldstart.php>, visitada por última vez el 15 de marzo de 2022.
<https://alphafold.com/>, visitada por última vez el 15 de marzo de 2022.
<https://swissmodel.expasy.org/>, visitada por última vez el 15 de marzo de 2022.

