

**ANÁLISIS DE HOMOLOGÍA ESTRUCTURAL DE LAS REGIONES CARBOXILO
TERMINAL DE LOS CRIPTOCROMOS EN DIFERENTES ESPECIES**

CARLOS ARTURO OSORIO ISAZA

**Trabajo de grado presentado como requisito parcial para optar el título de
Químico**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE QUÍMICA
BUCARAMANGA
2008**

**ANÁLISIS DE HOMOLOGÍA ESTRUCTURAL DE LAS REGIONES CARBOXILO
TERMINAL DE LOS CRIPTOCROMOS EN DIFERENTES ESPECIES**

CARLOS ARTURO OSORIO ISAZA

**Trabajo de grado presentado como requisito parcial para optar el título de
Químico**

Directora del proyecto de grado:

Dra. Martha Cecilia Daza Espinosa

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE CIENCIAS

ESCUELA DE QUÍMICA

BUCARAMANGA

2008

DEDICATORIA

A Dios le presento este fruto de mi trabajo para demostrarle la gran obra que el esta haciendo en mí.

A mi papás por ellos los directos responsables en este logro, por todo aquello que me enseñaron y que me hicieron ser una persona honesta de la que ellos se puedan sentirse orgullosos. Por brindarme su tiempo, comprensión, y lo más importante, su muestra de amor hacía mí.

A Katha, la inspiración de mi vida, por ser mi mas grande motivación, ya que hace que cada día sea mejor, solo espero que Dios me permita estar contigo también compartiendo tus triunfos así como tu has estado en los míos. Te amo, y espero estar contigo el resto de mis días.

AGRADECIMIENTOS

A la Doctora Martha Cecilia Daza, directora del grupo de bioquímica teórica por abrirme las puertas de su grupo de investigación y por darme enseñanzas que van mucho más allá de una simple formación científica ya que en ella encontré además a una amiga que aprecio mucho.

A mis compañeros del GBQT que de cierta forma me orientaron en partes del proyecto y me colaboraron desinteresadamente, les deseo éxitos en su vida profesional, en especial al químico Ciro Rozo y al estudiante Asdrubal Lozada.

A mis compañeros de carrera, no los nombro porque cometería el error de dejar a alguien por fuera, ya que este logro no hubiera sido posible sin el apoyo y compañía durante el tiempo para poder sacar este reto adelante.

TABLA DE CONTENIDO

	Pág.
INTRODUCCIÓN	1
ABREVIATURAS	4
I. MARCO TEÓRICO	5
1. CRIPTOCROMOS Y FOTOLIASAS	5
1.1 Familia de reparadores de ADN/fotorreceptores de luz azul	6
1.2 Relaciones evolutivas	8
1.3 Diversidad en la función de la familia de reparadores de ADN/fotorreceptores de luz azul	9
1.4 Estructura de las fotoliasas/criptocromos	11
1.5 Región carboxilo terminal de los criptocromos	13
2. ORGANIZACIÓN ESTRUCTURAL DE LAS PROTEÍNAS	16
2.1 Arquitectura en proteínas	16
2.2 Clasificación estructural de proteínas	17
2.3 Predicción de la estructura de proteínas	17
II. METODOLOGÍA	20
1. Selección de las secuencias de los criptocromos	20
2. Determinación de la extensión carboxilo terminal	20
3. Determinación de la homología estructural y de los elementos de estructura secundaria de las regiones carboxilo terminal de los criptocromos	21
4. Modelamiento tridimensional de la homología encontrada en las regiones carboxilo terminal de los criptocromos	23
III. RESULTADOS	24
1. Secuencias seleccionadas y extensión carboxilo terminal de los criptocromos	24

2.	Determinación de la homología estructural de las regiones carboxilo terminal de los criptocromos	27
3.	Determinación de la estructura secundaria de las regiones carboxilo terminal de los criptocromos	41
4.	Modelamiento tridimensional de la homología encontrada en las regiones carboxilo terminal de los criptocromos	49
	CONCLUSIONES	73
	ANEXOS	75
	BIBLIOGRAFÍA	89

LISTA DE TABLAS

	Pág.
Tabla 1	Clasificación de las proteínas según su estructura secundaria y terciaria 18
Tabla 2	Ejemplos de agrupamientos hidrofóbicos encontrados en HCA 22
Tabla 3	Secuencias de criptocromos de insectos seleccionadas 24
Tabla 4	Secuencias de criptocromos de vertebrados seleccionadas 25
Tabla 5	Secuencias de criptocromos de plantas seleccionadas 25
Tabla 6	Subgrupo 1 de regiones carboxilo terminal de insectos 27
Tabla 7	Subgrupo 2 de regiones carboxilo terminal de insectos 27
Tabla 8	Resultados de homología HCA (%) de las regiones CT del subgrupo 1 en insectos. 28
Tabla 9	Relación para encontrar la homología estructural (%) entre todas las regiones CT en insectos. 28
Tabla 10	Subgrupo A en vertebrados 29
Tabla 11	Subgrupo B en vertebrados 29
Tabla 12	Subgrupo C en vertebrados 29
Tabla 13	Subgrupo D en vertebrados 29
Tabla 14	Subgrupo E en vertebrados 29
Tabla 15	Subgrupo F en vertebrados 29
Tabla 16	Valores de homología HCA (%) subgrupo A en vertebrados 30
Tabla 17	Valores de homología HCA (%) subgrupo B en vertebrados 30
Tabla 18	Valores de homología HCA (%) subgrupo C en vertebrados 30
Tabla 19	Valores de homología HCA (%) subgrupo D en vertebrados 30
Tabla 20	Valores de homología HCA (%) subgrupo E en vertebrados 31
Tabla 21	Valores de homología HCA (%) subgrupo F en vertebrados 31
Tabla 22	Resultados de homología HCA (%) para relacionar las secuencias CT de los criptocromos en vertebrados 33
Tabla 23	Conjuntos de secuencias CT de criptocromos con relación estructural en vertebrados basados en los valores de homología de HCA 34
Tabla 24	Subgrupo A en plantas 34
Tabla 25	Subgrupo B en plantas 34
Tabla 26	Subgrupo C en plantas 34
Tabla 27	Subgrupo D en plantas 34
Tabla 28	Subgrupo E en plantas 35
Tabla 29	Subgrupo F en plantas 35
Tabla 30	Subgrupo G en plantas 35

Tabla 31	Subgrupo H en plantas	35
Tabla 32	Valores de homología HCA (%) subgrupo A en plantas	35
Tabla 33	Valores de homología HCA (%) subgrupo B en plantas	35
Tabla 34	Valores de homología HCA (%) subgrupo C en plantas	36
Tabla 35	Valores de homología HCA (%) subgrupo D en plantas	36
Tabla 36	Valores de homología HCA (%) subgrupo E en plantas	36
Tabla 37	Valores de homología HCA (%) subgrupo F en plantas	36
Tabla 38	Valores de homología HCA (%) subgrupo G en plantas	36
Tabla 39	Valores de homología HCA (%) subgrupo H en plantas	37
Tabla 40	Valores de homología HCA (%) para encontrar la relación estructural entre todas las regiones C-T de los criptocromos en plantas	38
Tabla 41	Conjuntos de secuencias con relación estructural en plantas basados en los valores de homología de HCA	39
Tabla 42	Valores de homología HCA (%) para encontrar la relación estructural entre todas las secuencias C-t de los criptocromos seleccionadas	40
Tabla 43	Relación entre las secuencias CT de los criptocromos seleccionadas y sus elementos para determinar la estructura secundaria por HCA	42
Tabla 44	Clasificación estructural de las regiones carboxilo terminal de las secuencias a las cuales se le detectaron sus elementos de estructura secundaria.	49
Tabla 45	Estructuras relacionadas con 1GP	50
Tabla 46	Estructuras relacionadas con 5CV	51
Tabla 47	Estructuras relacionadas con 2CP	52
Tabla 48	Estructuras relacionadas con 1M	54
Tabla 49	Estructuras relacionadas con 2EP	55
Tabla 50	Estructuras relacionadas con 2FP	56
Tabla 51	Estructuras relacionadas con 4FV	58
Tabla 52	Estructuras relacionadas con 3BV	60
Tabla 53	Estructuras relacionadas con 3M	61
Tabla 54	Estructuras relacionadas con 10I	64
Tabla 55	Estructuras relacionadas con 12I	66
Tabla 56	Estructuras relacionadas con 2M	69

LISTA DE FIGURAS

		Pág.
Figura 1	Daños producidos en el ADN a causa de la radiación UV	5
Figura 2	Cofactores que actúan como cromóforos en la familia de reparadores de ADN/fotorreceptores de luz azul.	6
Figura 3	Clasificación de la familia de reparadores de ADN/fotorreceptores de luz azul según la funcionalidad de de las proteínas.	7
Figura 4	Árbol filogenético construido con base en múltiples alineamientos de las secuencias de diversas proteínas fotoliasas y criptocromos.	8
Figura 5	Mecanismo de reparación del ADN por medio de la CPD fotoliasa	10
Figura 6	Estructura tridimensional de la CPD fotoliasa de <i>E.coli</i> . Código PDB: 1DNP	12
Figura 7	Estructura cristalina del dominio similar a la fotoliasa del criptocromo 1 de <i>A. thaliana</i> . Código PDB: 1U3C	13
Figura 8	Esquema general de la estructura de la familia de proteínas reparadoras de ADN/Fotorreceptoras de luz azul	15
Figura 9	Extensión de la región carboxilo terminal de los criptocromos. La zona amarilla muestra aproximadamente la extensión de la región carboxilo terminal. Las fotoliasas no poseen dicha zona y en los criptocromos depende el organismo.	15
Figura 10	Posibles funciones de la región carboxilo terminal en los criptocromos	15
Figura 11	Predicción de la estructura secundaria de las regiones C-terminal en: D: <i>Drosophila melanogaster</i> (insecto), H: <i>Homo sapiens</i> (Vertebrado) y A.t.: <i>Arabidopsis thaliana</i> (planta). Las estructuras α -hélice se muestran en cajas y las hojas β en forma de flechas.	16
Figura 12	Formas de estructura secundaria: a) α -hélice b) hojas β Estas se estabilizan por puentes de hidrógeno (líneas punteadas en azul)	17
Figura 13	Determinación de los elementos de estructura secundaria de las regiones carboxilo terminal de los criptocromos escogidos como representativos. Las flechas indican hojas β y las cajas indican α -hélices. Entre la representación y secuencia bidimensional se muestra el agrupamiento hidrofóbico detectado por HCA	47
Figura 14	Evaluación de 1GP con respecto a 1K0E	50
Figura 15	Evaluación de 1GP con respecto a 1OGM	51
Figura 16	Evaluación de 5CV con respecto a 1AZO	52
Figura 17	Evaluación de 2CP con respecto a 1V3J	53
Figura 18	Evaluación de 2CP con respecto a 1K8O	53
Figura 19	Evaluación de 1M con respecto a 1BOG	54
Figura 20	Evaluación de 1M con respecto a 1WYI	54
Figura 21	Evaluación de 2EP con respecto a 2JE8	55
Figura 22	Evaluación de 2EP con respecto a 1YMV	56
Figura 23	Evaluación de 2FP con respecto a 1I84	57
Figura 24	Evaluación de 2FP con respecto a 1JR7	57
Figura 25	Evaluación de 4FV con respecto a 1XWN	58
Figura 26	Evaluación de 4FV con respecto a 1JF3	59
Figura 27	Evaluación de 4FV con respecto a 2IX5	59
Figura 28	Evaluación de 3BV con respecto a 2PFE	60
Figura 29	Evaluación de 3BV con respecto a 1DI6	60
Figura 30	Evaluación de 3M con respecto a 2BIX	61

Figura 31	Evaluación de 3M con respecto a 2EBS	62
Figura 32	Evaluación de 3M con respecto a 1KXJ	62
Figura 33	Evaluación de 3M con respecto a 2PUA	63
Figura 34	Evaluación de 3M con respecto a 3BIC	63
Figura 35	Evaluación de 10I con respecto a 1C5B	64
Figura 36	Evaluación de 10I con respecto a 1CCW	65
Figura 37	Evaluación de 10I con respecto a 1XSI	65
Figura 38	Evaluación de 10I con respecto a 1PGV	66
Figura 39	Evaluación de 12I con respecto a 1X5H	67
Figura 40	Evaluación de 12I con respecto a 1BWV	67
Figura 41	Evaluación de 12I con respecto a 2QV3	68
Figura 42	Evaluación de 12I con respecto a 1CC1	68
Figura 43	Evaluación de 12I con respecto a 2OEE	69
Figura 44	Evaluación de 2M con respecto a 1C4K	70
Figura 45	Evaluación de 2M con respecto a 1BGP	70
Figura 46	Evaluación de 2M con respecto a 1PMS	71
Figura 47	Evaluación de 2M con respecto a 1Y8Q	71

LISTA DE ANEXOS

		Pág.
Anexo 1	Matriz de sustitución, basada en el principio de HCA, en donde los aminoácidos hidrofóbicos aportan el mayor valor en el alineamiento para obtener correspondencias en los agrupamientos hidrofóbicos	76
Anexo 2	Perfiles de HCA de los conjuntos en insectos, vertebrados y plantas detectados por homología HCA	77
Anexo 3	Representación tridimensional de los modelos obtenidos de los fragmentos de las regiones carboxilo terminal de los criptocromos indicados.	99

TÍTULO*:

ANÁLISIS DE HOMOLOGÍA ESTRUCTURAL DE LAS REGIONES CARBOXILO TERMINAL DE LOS CRIPTOCROMOS EN DIFERENTES ESPECIES

AUTOR:**

Osorio Isaza, Carlos Arturo

PALABRAS CLAVES

Criptocromo, región carboxilo terminal, homología estructural, estructura secundaria, HCA, modelamiento de homología.

RESUMEN:

Los criptocromos son proteínas fotorreceptoras de luz azul que se encuentran en todos los reinos de la naturaleza y son componentes esenciales de los ritmos circadianos. Comparten homología de estructura primaria con las fotoliasas (proteínas reparadoras de ADN) y conforman la familia de proteínas reparadoras de ADN/fotorreceptores de luz azul.

Los criptocromos poseen una extensión carboxilo terminal adicional que no está presente en las fotoliasas, varía en longitud de aminoácidos según la especie, donde las plantas poseen la extensión más larga y los insectos la más corta. Estas regiones carboxilo terminal no presentan homología de estructura primaria entre ellas y no se conoce organización estructural secundaria o superior de las mismas, a excepción de tres de ellas donde se utilizaron métodos de predicción basados en alineamientos unidimensionales.

* Proyecto de Grado.

** Facultad: Ciencias. Escuela: Química. Directora: Dra. Martha Cecilia Daza Espinosa. Grupo de Bioquímica Teórica.

En este trabajo se seleccionaron 76 secuencias de criptocromos, a las cuales se determinaron su extensión carboxilo terminal. Luego por medio del principio de homología secundaria de HCA se clasificaron las regiones con similitud estructural. A continuación se le hallaron sus elementos de estructura secundaria y por último, se hizo un modelamiento de la homología detectada.

Como resultado se clasificaron las regiones carboxilo terminal en 14 tipos diferentes de ordenamiento estructural por HCA que al ser comparados con otros autores muestran diferencias en el arreglo conformacional secundario. En el modelamiento de la homología solo pequeños fragmentos de las regiones sometidas a predicción fueron obtenidas, regiones que no muestran relación funcional detectada hasta el momento.

TITLE*:

ANALYSIS OF STRUCTURAL HOMOLOGY OF THE REGIONS CARBOXY TERMINAL OF THE CRYPTOCHROMES IN DIFFERENT SPECIES

AUTHOR:**

Osorio Isaza, Carlos Arturo

KEYWORDS

Cryptochrome, carboxy terminal region, structural homology, secondary structure, HCA, Homology modelling.

ABSTRACT

The cryptochromes are photoreceptor proteins of blue light that they find in all the kingdoms of the nature and are essential components of the circadian rhythm. They share homology of primary structure with the photolyases (proteins repairs of DNA) and shape the family of proteins repairs of ADN/photoreceptors of blue light.

The cryptochromes possess an extension carboxy terminal additional that is not present in the photolyases, it changes in length of amino acids according to the species, where the plants possess the longest extension and the insects the shortest. These regions carboxy terminal do not present homology of primary structure between them and there is not known structural secondary or top organization of the same ones, with the exception of three of them where there were in use methods of prediction based on alignments one-dimensional.

In this work 76 sequences were selected of cryptochromes, to which they determined your extension carboxy terminal. Based on of the principle of homology

* Work of degree.

** Science Faculty. Chemistry Department. Director: Dr. Martha Cecilia Daza Espinosa. Group of theoretical biochemistry.

secondary of HCA the regions qualified with structural similarity. Later he was located by your elements of secondary structure and finally, he became a homology modeling detected.

As result the regions qualified carboxy terminal in 14 different types from structural classification for HCA that on have been compared by other authors show differences in the arrangement conformational secondary. In the homology modeling alone small fragments of the regions submitted to prediction were obtained, regions that do not show functional relation detected up to the moment.

INTRODUCCIÓN

Los criptocromos son proteínas receptoras de luz azul, se encuentran en todos los seres vivos¹ y se identificaron por primera vez en 1993². Estas proteínas comparten homología de estructura primaria con las fotoliasas, proteínas reparadoras del ADN³. Los criptocromos en animales y plantas son componentes esenciales de los complejos proteínicos responsables de los ritmos circadianos (ritmos que regulan las actividades del reloj biológico). En las plantas, además, regulan el crecimiento del tallo, debido a su hipersensibilidad a la luz azul³.

Estas proteínas (fotoliasas y criptocromos) conforman la familia de proteínas fotoliasas reparadoras de ADN/fotorreceptores de luz azul debido a su homología en estructura primaria y su sensibilidad a la radiación comprendida entre los 320 y 500 nm⁴.

La estructura tridimensional de los criptocromos ha sido determinada, únicamente, para de tres especies: dos plantas (Criptocromo 1 *A. thaliana* PDB: 1U3C⁵, Criptocromo 3 *A. thaliana* PDB: 2IJG⁶ y la bacteria *Synechocystis* PDB: 1NP7¹). La estructura tridimensional de estas proteínas es similar a la de las fotoliasas⁵.

Las fotoliasas y los criptocromos se diferencian en función² y en su estructura. Los criptocromos poseen una región carboxilo terminal característica, adicional al dominio que comparte con las fotoliasas³ (Dominio PHR, reparador del ADN en fotoliasas⁵).

Esta región carboxilo terminal de los criptocromos varía en longitud y en estructura primaria en diferentes organismos⁷, las regiones carboxilo terminal más extensas están en las plantas (entre 150 y 240 residuos de aminoácido), aunque encontramos tres excepciones, dos criptocromos en insectos que poseen una longitud de 270 y 473 residuos (criptocromo 2 de *Antheraea pernyi* y criptocromo 2 de *Anopheles gambiae* respectivamente) y uno en una planta (criptocromo de *Chlamydomonas reinhardtii*) que tiene un criptocromo con una región carboxilo terminal de 520 aminoácidos.

Resultados de predicción de estructura secundaria de la región carboxilo terminal de los criptocromos, de CRY1 *Arabidopsis thaliana* y CRY2 *Homo sapiens* y CRY *Drosophila melanogaster*, utilizando el PredictProtein service (<http://cubic.bioc.columbia.edu/predictprotein>) indican que estas regiones poseen una estructura secundaria carente de orden. Es decir, tienen elementos de estructura secundaria diferentes y contienen principalmente aminoácidos característicos de las zonas flexibles de una proteína⁷.

En este trabajo seleccionamos 78 criptocromos de insectos, vertebrados y plantas. Aplicamos la técnica HCA (Hydrophobic cluster analysis) para la determinación de la homología de estructura secundaria, donde hallamos entre todas las regiones carboxilo terminal 14 tipos de organizaciones estructurales a nivel secundario.

Comparando nuestros resultados con trabajos anteriores, encontramos diferencias en la predicción de la estructura secundaria de la región carboxilo terminal del criptocromo en *Drosophila melanogaster*, en donde encontramos una estructura tipo β , a diferencia de la hélice α que se reportó anteriormente⁷

También intentamos modelar las secuencias carboxilo terminal, primero utilizando la técnica BLAST²⁹, (Basic Local Alignment Search Tool) para detectar posibles homólogos de estructura terciaria. No encontramos proteínas de estructura

terciaria conocida que fueran homólogas con la extensión total de aminoácidos de las regiones carboxilo terminal analizadas, salvo unos cortos fragmentos (entre 25 y 50 aminoácidos); por tal razón, no es posible sugerir una función para la región carboxilo terminal en estas proteínas.

Por último modelamos homología de estos cortos fragmentos por medio del programa MODELLER³⁰ que muestra tridimensionalmente los elementos de estructura secundaria hallados por HCA.

ABREVIATURAS

PDB: Protein Data Bank

CRY: Criptocromo

HCA: Hydrophobic cluster analyses

ADN: Ácido desoxirribonucleico

UV: ultravioleta

CPD: dímeros de pirimidina de ciclobutano

FAD: Flavin adenin dinucleótido

MTHF: Meteniltetrahidrofolato

8-HDF: 7,8-didemetil-8-hidroxi-5-dezaflavina

NCBI: National Center for Biotechnology Information

Blast: Basic Local Aligment Search Tool

aa: Aminoácidos

CT: Región carboxilo terminal

I. MARCO TEÓRICO

1 CRIPTOCROMOS Y FOTOLIASAS:

Las fotoliasas son proteínas fotorreceptoras de luz azul que se encargan de reparar los daños en el ADN ocasionados por la radiación ultra violeta en diferentes organismos⁸, excepto en mamíferos⁹. Estas proteínas se clasifican en dos tipos, dependiendo del daño que reparan en el ADN: Las CPD fotoliasas que reparan los dímeros de ciclobutil pirimidina y las 6-4 fotoliasas que reparan los pirimidin (6-4) pirimidona (Figura 1)³.

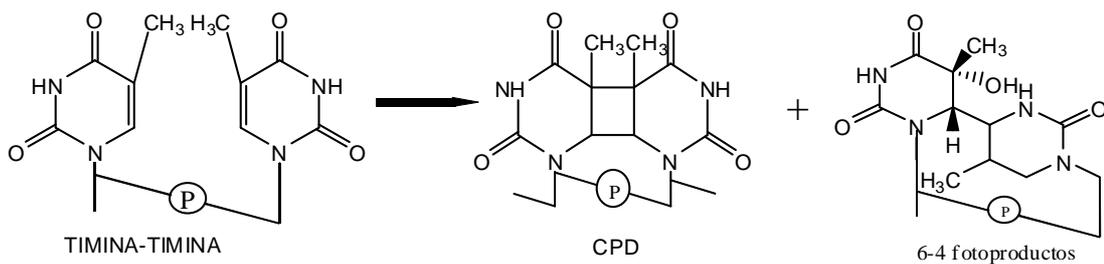


Figura 1: Daños producidos en el ADN a causa de la radiación UV

Durante la investigación de las fotoliasas, a mediados de 1993 se encontró en plantas un nuevo fotorreceptor sensible a la luz azul que era responsable del crecimiento del tallo³, al cual se le llamó criptocromo¹⁰. Estas nuevas proteínas están presentes en todos los seres vivos¹, pero no tienen la capacidad de reparar el ADN dañado por la radiación UV. Estos dos tipos de proteínas (fotoliasas y criptocromos) conforman la familia de reparadores de ADN/Fotorreceptores de luz azul².

1.1 Familia de reparadores de ADN/fotorreceptores de luz azul:

Una característica principal de esta familia de fotorreceptores son los cofactores presentes en ellos. Cada proteína posee dos cofactores, de los cuales todos poseen el FAD, este hace el papel de cofactor catalítico en las fotoliasas y un segundo cofactor que puede ser el MTHF o el 8-HDF (7,8-didimetil-8-hidroxi-5-dezaflavina) (Figura 2), este último cofactor funciona como el cromóforo “antena” que capta el fotón, y por medio de una energía de excitación le transfiere el electrón al cromóforo catalítico³.

De acuerdo con su función, las proteínas pertenecientes a esta familia se dividen en: los reparadores de ADN (fotoliasas) y los fotorreceptores de luz azul (criptocromo).¹¹. (Figura 3²)

Entre los criptocromos existen dos tipos que son los criptocromos de animales, que participan como componentes del reloj biológico, y los criptocromos de plantas que además de participar en los relojes biológicos (ritmos circadianos), participan en la expansión del cotiledón y en la producción de antocianinas³.

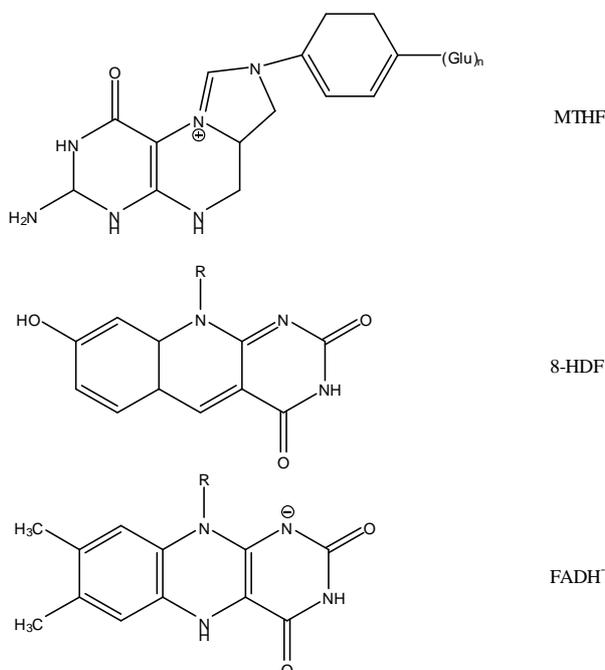


Figura 2: Cofactores que actúan como cromóforos en la familia de reparadores de ADN/fotorreceptores de luz azul.

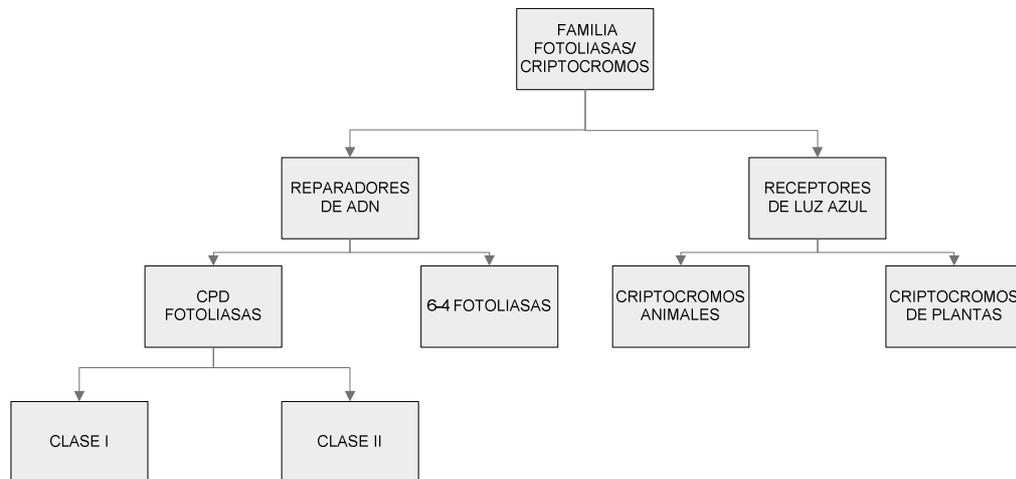


Figura 3: Clasificación de la familia de reparadores de ADN/fotorreceptores de luz azul según la funcionalidad de de las proteínas.

Las clases I y II en las CPD fotoliasas se refiere al tipo de organismos donde han sido obtenidas: la clase I se ha aislado de procarióticas y eucarióticas unicelulares; En cambio la clase II se ha encontrado en eubacteria, archaeobacteria y eucariotas superiores².

En esta clasificación anteriormente mencionada no se encuentra relacionado un tipo de criptocromo llamado DASH¹. Este es llamado así para distinguirse de las demás proteínas pertenecientes a esta familia, y tiene la capacidad de enlazar ADN no específicamente¹², mostrando actividad reparadora tipo CDP fotoliasa pero de manera muy débil, por lo que sugieren que es un regulador transcripcional, pero su función específica no ha sido aún encontrada¹³.

Esta familia de proteínas presenta homología de estructura primaria. (Figura 4)².



Figura 4: Árbol filogenético construido con base en múltiples alineamientos de las secuencias de diversas proteínas fotoliasas y criptocromos. Las abreviaturas son: *N. crassa* (Nc), *S. cerevisiae* (Sc), *E. coli* (Ec), *S. typhimurium* (St), *A. nidulans* (An), *H. halobium* (Hh), *S. griseus* (Sg), *A. thaliana* (At), *S. alba* (Sa), *C. reinhardtii* (Cr), *M. domestica* (Md), *C. auratus* (Ca), *D. melanogaster* (Dm), *M. thermoautotrophicum* (Mt), *Zebrafish* (Z), *H. sapiens* (h), *X. leavis* (Xl).

Con base en la figura anterior, la familia de reparadores de ADN/Fotorreceptores de luz azul se puede clasificar con respecto a su estructura primaria en:

1. Clase I Fotoliasas
2. Clase II fotoliasas
3. Criptocromos de plantas
4. 6-4 fotoliasas/ criptocromos animales

La homología de estructura primaria dentro de los miembros en cada uno de los cuatro grupos varía entre 25% y 80% excepto entre los pertenecientes al grupo 4 donde esta homología varía entre 40% y 60%. Y la homología entre cada grupo es relativamente baja (10% al 17%)².

1.2 Relaciones evolutivas:

La característica evolutiva de esta familia, basada en análisis filogenéticos, llamada evolución repetida, ha sugerido que el primer ancestro de esta familia de proteínas procede de las CPD fotoliasas antes de la divergencia de procariotas y eucariotas, de este punto se derivó la evolución de esta familia pasando por la clase I y II de las CPD fotoliasas en el linaje procariótico, hasta la aparición de los criptocromos en animales superiores².

Inicialmente se sugería que los criptocromos aparecieron después de la divergencia de las procariotas y eucariotas, ya que no se tenía evidencia de la existencia de criptocromos en organismos bacteriales³, pero con la aparición del primer criptocromo DASH, que fue encontrado en una bacteria, llegó a la conclusión que los criptocromos evolucionaron antes de dicha divergencia¹.

Una característica interesante de esta familia es, que a pesar de la similitud entre los criptocromos y las fotoliasas, estos genes parecen no ser ortólogos, lo cual es un ejemplo de evolución repetida³, la cual plantea que los genes de dichas proteínas evolucionaron a diferentes tiempos del ancestro común fotoliasa. Esto se ve claramente en el siguiente ejemplo: En *Arabidopsis thaliana* (una planta), existen dos criptocromos (CRY1 y CRY2) y comparados con los criptocromos de animales, son más parecidos en secuencia los criptocromos animales a las fotoliasas de la planta que los criptocromos entre sí. Esto plantea que los criptocromos evolucionaron de las fotoliasas independientemente⁹.

1.3 Diversidad en la función de la familia de reparadores de ADN/fotorreceptores de luz azul:

Ya que el mecanismo de las CPD fotoliasas es el único que ha sido bien elucidado, ha sido usado para ejemplificar el proceso de fotorrecepción de una forma sencilla y breve. Las fotoliasas (En *E. coli*, con cofactores MTHF y FAD) enlazan al ADN selectivamente en el sitio del daño de manera independiente de la

luz, el cofactor antena capta el electrón y por medio de un proceso de transferencia de energía pasa el electrón del MTHF* al FADH⁻.

El FADH* por medio de un ciclo catalítico parte el dímero y produce ADN intacto (Figura 5). El electrón se regresa al FADH para así completar dicho ciclo¹⁴.

El mecanismo en los criptocromos no es conocido aún, claro que los criptocromos dependiendo el organismo, realizan los siguientes tipo de funciones: En plantas como *Arabidopsis thaliana* existen más de un criptocromo, ambos funcionan como fotorreceptores que regula los tiempos de florecimiento y desarrollo de la planta³, el Criptocromo 1 trabaja en luz azul de alta intensidad y el criptocromo 2 en condiciones de baja intensidad ya que este se degrada rápidamente por la exposición de luz azul intensa¹⁵. En *Drosophila melanogaster* (mosca de la fruta) el criptocromo se acopla al ritmo circadiano y en mamíferos actúa como componente del mecanismo reloj biológico³.

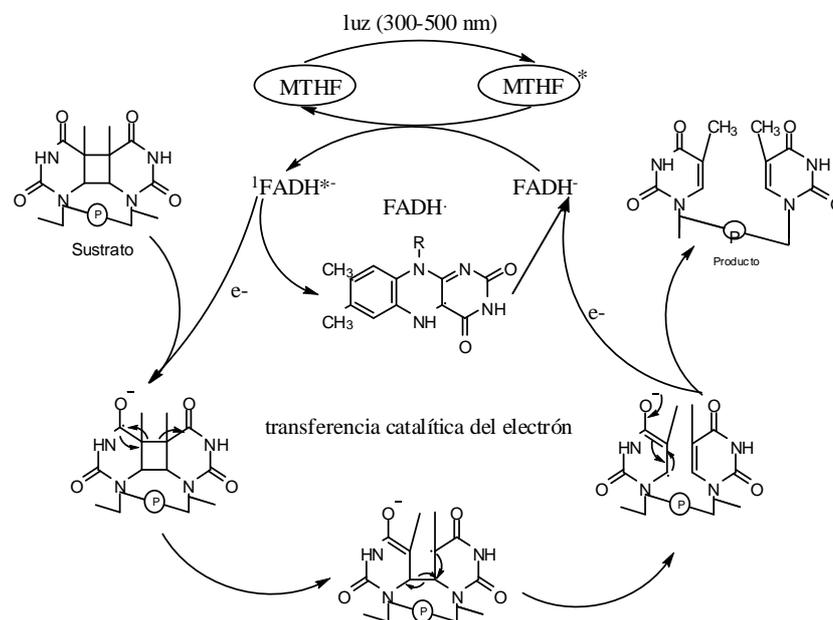


Figura 5: Mecanismo de reparación del ADN por medio de la CPD fotoliasa

1.4 Estructura de las fotoliasas/criptocromos:

La primer fotoliasa en ser caracterizada tridimensionalmente fue la fotoliasa de *E.coli* (figura 6)¹⁶ con las siguientes características: Es una CPD fotoliasa, consta de una cadena polipeptídica de 471 residuos de aminoácidos que se separan en dos dominios, el primer dominio α/β (residuos 1-131) conformada por 5 hojas β paralelas recubierta en ambos lados por hélices α . El segundo dominio es solo α (residuos 204 a 469) conectadas por loops. Entre ambos dominios se encuentra una región loop conectora (132-203 residuos). El dominio α/β es el sitio donde se aloja el cofactor antena y en el dominio α se enlaza el FAD⁸.

Entre las fotoliasas a las cuales se les ha determinado su estructura tridimensional tenemos: *Thermus thermophilus* (1IQR) y *Anacystis nidulans* (1QNF)¹⁶.

Entre las estructuras de los criptocromos, solo se ha determinado la estructura de tres de ellos: Criptocromo 1⁵ y 3⁶ de *Arabidopsis thaliana* (planta) y el criptocromo DASH de la bacteria *Synechocystis*¹. Describiremos la estructura del criptocromo 1 de *Arabidopsis thaliana* (Figura 7)¹⁶:

Posee un dominio inicial α/β entre los residuos 13 y 139, compuesto por cinco láminas β paralelas entrecruzadas por 4 α hélices y una hélice 3_{10} (tres residuos por giro y conformada por una cadena de diez átomos). Un segundo dominio α/α entre los residuos 217 y 495, en donde están presentes 4 α -hélices y dos hélices 3_{10} , en este segundo dominio es donde se enlaza el FAD.

Entre los dos dominios hay 77 aminoácidos que muestran solo una estructura regular secundaria (3 α hélices y una 3_{10}), a la que llamaron zona conectora (140 a 216 residuos). Un puente bisulfuro entre Cys-190 y Cys-80, que no afecta la estructura tridimensional. Utilizaron como modelo molecular la fotoliasa de *E.coli*

descrita anteriormente y únicamente se descifra la zona homologa a dicha fotoliasa (Dominio PHR) ya que la longitud total de la proteína es de 681 aminoácidos, pero los intentos de expresar la longitud total de la proteína fueron infructuosos por cambios inducidos en la fisiología de la célula donde se obtuvo la proteína¹⁷. Además se encontró que estas regiones son carentes de estructura en solución⁷.

Dominio α/β

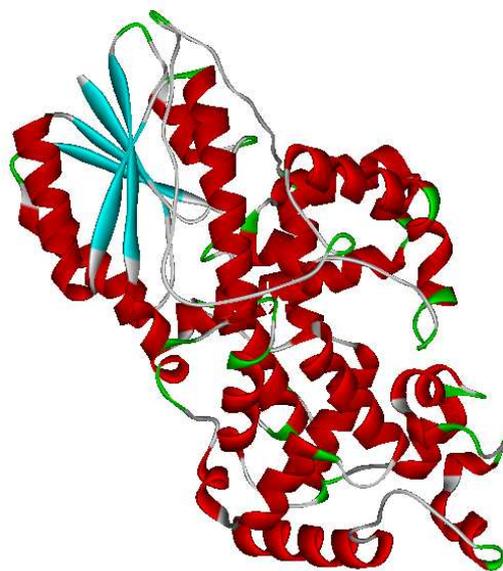


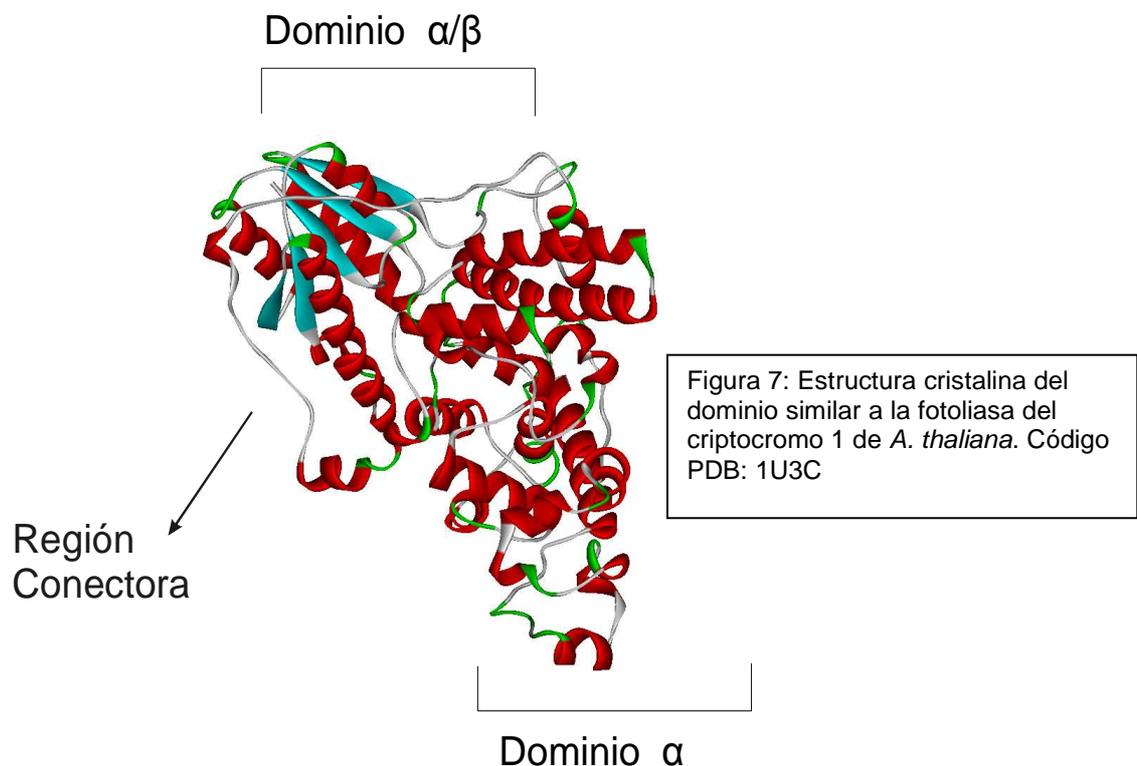
Figura 6: Estructura tridimensional de la CPD fotoliasa de *E. coli*. Código PDB: 1DNP

Dominio α

A pesar de las similitudes estructurales, las comparaciones de los rasgos superficiales de los miembros de la superfamilia de proteínas fotoliasas/criptocromos muestran una diferencia clave entre CRY1-PHR y otras proteínas en la superfamilia: La superficie del CRY1-PHR predominantemente cargada negativamente, con una pequeña concentración de carga positiva cerca de la cavidad de acceso del FAD. Esto contrasta con el fotoliasas y los Criptocromos DASH, que tiene un surco positivamente cargado sobre sus

superficies cerca de la cavidad de acceso del FAD. En fotoliasas, este surco es postulado para ligar los fosfatos del ADN que rodea una lesión inducida por radiación UV. También, esta región de la superficie del criptocromo DASH contribuye a su capacidad de ligar el ADN. La carencia de un surco positivamente cargado sobre la superficie del CRY1-PHR probablemente explica su carencia de actividad de enlazar el ADN⁵.

En los criptocromos se presentan las mismas características estructurales que en las fotoliasas, la presencia de dos dominios: Un dominio α/β y otro dominio α , aunque la región conectora entre los dominios varía en su longitud, en general esta zona similar (dominio PHR) tiene aproximadamente en longitud 490 aminoácidos⁵. Además los criptocromos poseen una región carboxilo terminal adicional que no se presenta en las fotoliasas. (Figura 8²).



1.5 Región carboxilo terminal de los criptocromos:

Los criptocromos poseen una secuencia de aminoácidos adicional en la región carboxilo terminal que no está presente en las fotoliasas³ a la que no se le ha determinado experimentalmente su estructura⁵. Esta región varía en longitud dependiendo el organismo³ (Figura 9). La longitud de esta región en plantas es de 150 a 250 residuos y en insectos de 30 a 80 residuos. Y no presentan homología de estructura primaria entre ellas⁷.

Con respecto a la función de la región carboxilo terminal se han propuesto tres roles⁹ Figura 10:

- En *Arabidopsis thaliana* (una planta) la región carboxilo terminal es un dominio de señalización que en ausencia de luz está inactivo. Cuando es activado el dominio PHR, este activa a la región carboxilo terminal. La activación consiste en un cambio conformacional que facilita la interacción con componentes de los relojes biológicos.
- En *Drosophila melanogaster* (un insecto) la región carboxilo terminal inhibe al dominio PHR (región similar a las fotoliasas) en ausencia de luz y es activado por la interacción con un fotón, la zona terminal sufre un cambio conformacional y permite al dominio PHR realizar sus funciones circadianas.
- En vertebrados la cola carboxilo terminal es necesaria para la ubicación intracelular (núcleo o citoplasma) de la proteína

La estructura terciaria de estas regiones carboxilo terminal no son conocidas aún y solo a algunas zonas carboxilo terminal se les ha predicho su estructura secundaria (regiones CT del CRY1 de *A. thaliana* CRY2 de *H. sapiens* y CRY *D. melanogaster*) que muestran diferencias de organización estructural. Esta predicción fue hecha por medio del PredictProtein

(<http://cubic.bioc.columbia.edu/predictprotein>), el cual es un servicio para el análisis de secuencia, predicción de la estructura y función de una proteína¹⁸. Los resultados obtenidos se muestran en la Figura 11⁷.

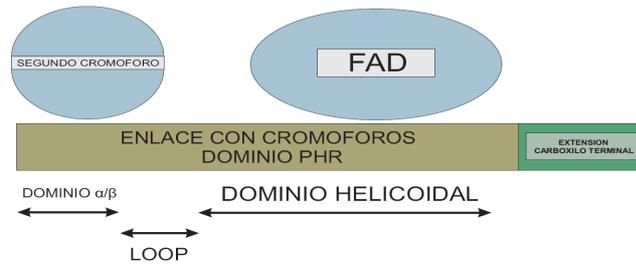


Figura 8: Esquema general de la estructura de la familia de proteínas reparadoras de ADN/Fotorreceptoras de luz azul

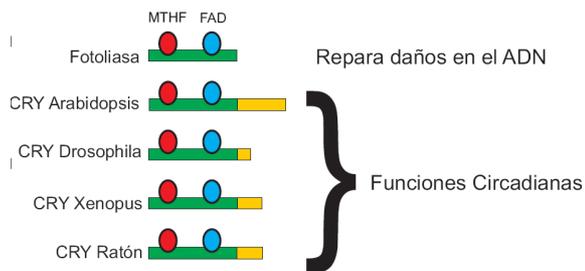


Figura 9: Extensión de la región carboxilo terminal de los criptocromos. La zona amarilla muestra aproximadamente la extensión de la región carboxilo terminal. Las fotoliasas no poseen dicha zona y en los criptocromos depende el organismo.

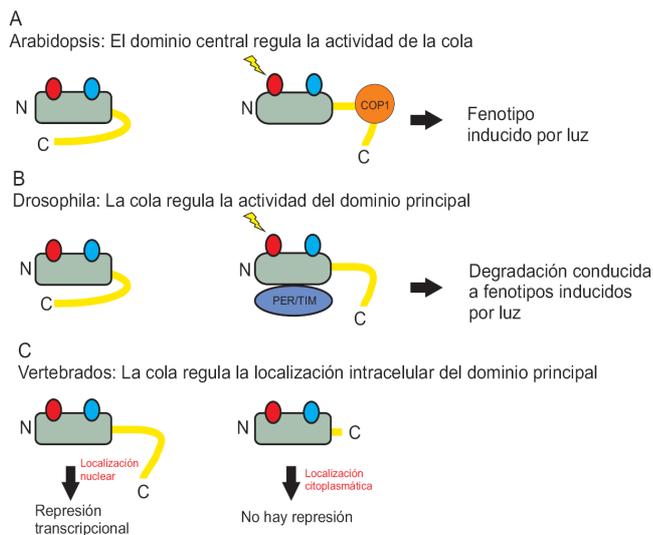


Figura 10: Posibles funciones de la región carboxilo Terminal en los criptocromos

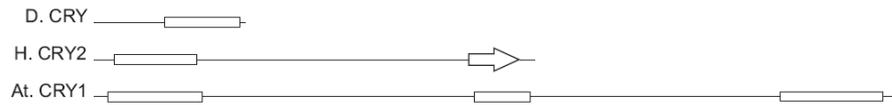


Figura 11: Predicción de la estructura secundaria de las regiones C-terminal en: D: *Drosophila melanogaster* (insecto), H: *Homo sapiens* (Vertebrado) y A.t.: *Arabidopsis thaliana* (planta). Las estructuras α -hélice se muestran en cajas y las hojas β en forma de flechas.

2. ORGANIZACIÓN ESTRUCTURAL DE LAS PROTEÍNAS

2.1. Arquitectura en proteínas:

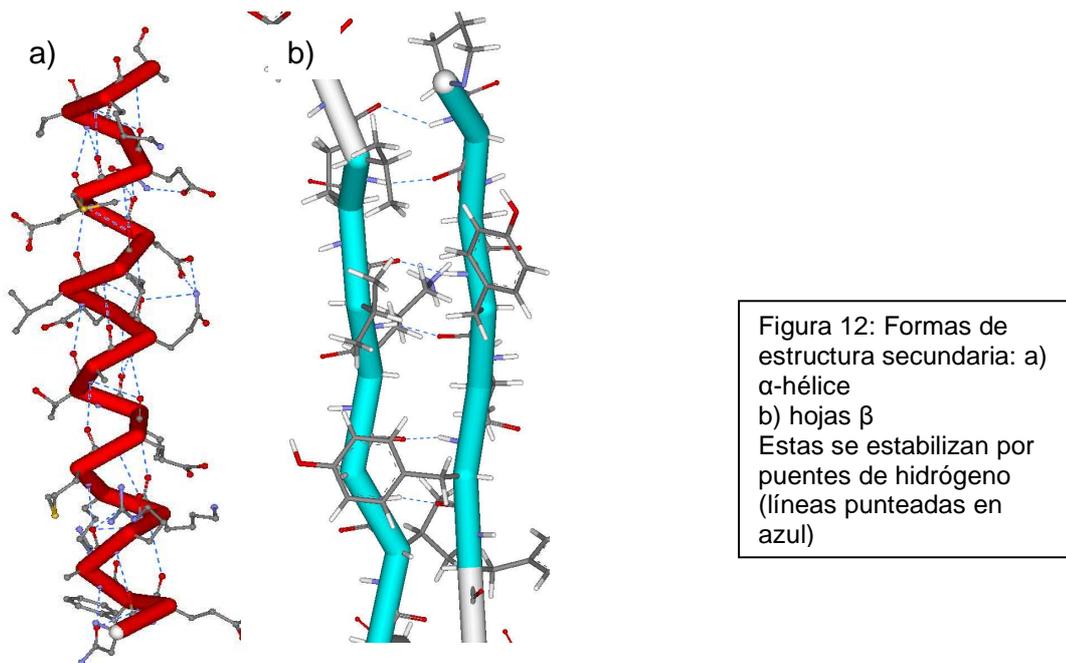
Las proteínas presentan 4 niveles estructurales: el primero es la secuencia de aminoácidos llamado estructura primaria de la cadena polipeptídica (monómero). La estructura secundaria, hace alusión al plegamiento, de la estructura primaria en hélices y hojas beta, estabilizado por puentes de hidrógeno (Figura 12). El ensamble e interacciones entre las α -hélices y las hojas β se denomina estructura terciaria. Estos tres niveles de estructura se presentan en una cadena polipeptídica. Cuando la proteína estructuralmente activa esta integrada por más de una cadena polipeptídica, el nivel estructural relacionado con la interacción de estas subunidades se denomina estructura cuaternaria¹⁹:

Además de estos niveles estructurales, existen otros dos tipos de niveles estructurales en proteínas: Los motivos y dominios.

- Los motivos son pequeñas subestructuras, generalmente constituidas por unos pocos elementos de estructura secundaria. Son, además, elementos integrales en el plegamiento de las proteínas y a menudo tienen significado funcional, en una proteína puede representar la mínima unidad funcional.

- Los Dominios son secciones compactas que representan regiones independientes estructuralmente, en cada dominio de una proteína se mantiene su estructura característica.

Varios motivos se pueden combinar y formar un dominio específico²⁰.



2.2 Clasificación de las estructuras de proteínas¹⁹:

La clasificación de las estructuras está basada en las estructuras secundaria y terciaria (Tabla 1).

2.3 Predicción de la estructura de proteínas:

La predicción de la estructura tridimensional de una proteína a partir de su secuencia de aminoácidos es el área más debatida en la bioinformática, por existir tantos métodos para realizarlos. Estas predicciones involucran un tipo de plegamiento que una secuencia de aminoácidos dada puede adoptar²¹.

Tabla 1: Clasificación de las proteínas según su estructura secundaria y terciaria	
Clase	Característica
α -helicoidal	Estructura secundaria casi exclusivamente α -hélices
β -hoja	Estructura secundaria casi exclusivamente hojas β
$\alpha + \beta$	α -hélices y hojas β separadas en diferentes partes de la proteína. Ausencia del motivo β - α - β
α/β	Hélices y hojas ensambladas por unidades β - α - β

Algunos se basan en principios físicos para reproducir interacciones interatómicas para definir una energía asociada con la conformación. El problema en este tipo de predicción es encontrar un mínimo global energéticamente estable debido a la complejidad de las superficies de energía potencial de una macromolécula y a que los algoritmos de minimización tienden a encontrar mínimos locales. La otra alternativa está basada en encontrar similitudes de estructura primaria entre dos secuencias, en donde una de ellas tiene su estructura tridimensional conocida¹⁹. Estos métodos son los usados actualmente.

Los métodos de predicción de estructura de proteínas a partir de la secuencia de aminoácidos incluyen¹⁹:

- Predicción de estructura secundaria sin ensamblarlos en tres dimensiones. Los resultados son listas de regiones predichas en forma de α -hélices y hojas β .
- Modelado de homología: Predicción de la estructura tridimensional a partir de una estructura conocida de una o varias proteínas relacionadas. Los

resultados son un conjunto de coordenadas para la cadena principal y para las cadenas laterales.

- Reconocimiento del plegamiento: Dada una base de datos de estructuras conocidas, se determina cuáles de ellas comparten un plegamiento similar con la proteína de secuencia conocida y estructura desconocida. Los resultados son una lista de estructuras conocidas que tienen el mismo plegamiento que la proteína problema
- Predicción de nuevos plegamientos: Los resultados son un conjunto completo de coordenadas de por lo menos la cadena principal y algunas cadenas laterales. El modelo está destinado a obtener el plegamiento correcto, ya que no puede ser comparado con estructura conocida, salvo se pueda determinar experimentalmente.

II. METODOLOGÍA

1. Selección de las secuencias de los criptocromos:

La búsqueda de las secuencias de los criptocromos las hicimos en la base de datos de la NCBI²² con la palabra clave *cryptochrome*. Para seleccionar las secuencias tuvimos en cuenta:

- Secuencias que estuvieran reportadas más de una vez.
- Secuencias de fotoliasa (photolyase)
- Secuencias de proteínas que no tienen nombre definido. (se reporta en la página de resultados como “unnamed protein product”).
- Proteínas que por homología las muestre como resultados de la búsqueda pero no son criptocromos.
- Proteínas que se relacionan con los criptocromos en ritmos circadianos que se muestran como resultados de la búsqueda pero que no son criptocromos (proteína TIMELESS, proteína PERIOD).

Además tuvimos en cuenta que en la descripción del criptocromo estuviera presente el código PF03441. Este código es dado por Pfam²³, una base de datos de familias de proteínas. Y para este caso nos indica el dominio enlazante FAD de las fotoliasas.

2. Determinación de la extensión carboxilo terminal:

Para la determinación de la extensión carboxilo terminal observamos hasta donde llegaba la zona FAD enlazante, región hasta donde se tiene homología con las fotoliasas, para poder deducir la cantidad de residuos presentes en su región carboxilo terminal.

3. Determinación de la homología estructural y de los elementos de estructura secundaria de las regiones carboxilo terminal de los criptocromos:

El cálculo de la homología estructural lo realizamos basados en el principio de agrupamientos hidrofóbicos relacionados con los elementos de estructura secundaria presentes en una proteína globular llamado HCA²⁴ por medio de la expresión:

$$\text{Valor de homología HCA} = \frac{2CR*100}{RC_1 + RC_2} \%$$

Donde alineamos dos proteínas (secuencia 1 y secuencia 2) y RC₁ (o RC₂) es el número de residuos hidrofóbicos en el agrupamiento 1 de la secuencia 1 (o agrupamiento 2 de la secuencia 2) en cada una de las secuencias. CR es el número de residuos hidrofóbicos en el agrupamiento 1 que están en correspondencia con los residuos hidrofóbicos en el agrupamiento 2.

Y se siguieron los siguientes pasos:

- Agrupamiento de las regiones carboxilo terminal de los criptocromos escogidos en tres grupos: insectos, vertebrados y plantas.
- Organización de las regiones carboxilo terminal en subgrupos teniendo en cuenta que la diferencia de aminoácidos fuese de 10 a 20 residuos a modo de facilitar los alineamientos.
- Alineamiento de los elementos de cada subgrupo entre sí utilizando el programa de línea de comando CLUSTALW²⁵, y una matriz de sustitución que permita alinear preferiblemente las regiones hidrofóbicas relacionadas con el principio de HCA (Anexo 1).
- Preparación de las entradas para el programa HCA_analyze²⁶ por medio de archivos con extensión /*bat. para obtener los valores de homología a nivel secundario.

- Los pasos tres y cuatro los repetimos para obtener la comparación entre todas las secuencias de insectos, vertebrados y plantas para encontrar la relación estructural en todas las regiones carboxilo terminal escogidas de las diferentes especies.
- De esta forma hicimos conjuntos de secuencias con relación estructural basados en su valor de homología.

Las relaciones estructurales de homología secundaria pueden verse por medio de los gráficos de HCA (perfiles) obtenidos en por medio del programa Drawhca²⁷ y comparados manualmente, por la forma y tamaño de los agrupamientos hidrofóbicos a lo largo de las secuencias (Anexo 2).

Los elementos de estructura secundaria los hallamos manualmente por medio de los gráficos de HCA, ya que forma y tamaño de los agrupamientos hidrofóbicos vistos en ellos determina el tipo de estructura secundaria sea hélice- α u hoja β ²⁸(Tabla 2).

Hélices α			Láminas β			
						
10011001	10001001	100010001	110101	1000111	101011	10101
						
100010011	110011	10110001	111	1111	11101	10111
			<p>Tabla 2: Ejemplos de agrupamientos hidrofóbicos encontrados en HCA en donde se muestra la asociación con α hélices o láminas β. Los diamantes indican la posición de los aminoácidos y el código indica el tipo de aminoácidos involucrados (1 hidrofóbico y 0 hidrofílico).</p>			
10010011	10011	11001				

4. Modelamiento tridimensional de la homología encontrada en las regiones carboxilo terminal de los criptocromos.

Para el modelamiento tridimensional de la región carboxilo terminal. Tomamos un elemento de un conjunto de secuencias con homología de estructura secundaria alta (Mayor o igual a 75 %) y realizamos los siguientes pasos:

- Hallamos homólogos de cada secuencia a modelar: Esto lo hicimos utilizando el programa Blast²⁹ para búsqueda de homólogos remotos, el cual alinea la secuencia desconocida (query) con una base de datos seleccionada. Para este caso la base que utilizamos fue la de la PDB.
- Para escoger la secuencia o secuencias de estructura tridimensional conocida para ser usada como plantilla para modelar la región carboxilo terminal, tuvimos en cuenta: identidad de las secuencias, valores positivos (indican el porcentaje de aminoácidos alineados que tienen las mismas propiedades de hidrofobicidad e hidrofiliidad) y los gap en el alineamiento (desplazamientos en las secuencias para obtener un alineamiento mas óptimo).
- Luego utilizamos el programa para modelamiento de proteínas MODELLER^{30,31} versión 9v2, realizando la respectiva evaluación del modelo tridimensionalmente obtenido de la región carboxilo terminal con respecto al de la estructura conocida.
- Analizamos las zonas de las proteínas de estructura conocida que fueron usadas como plantilla para encontrar dominios o motivos conservados tanto en la estructura tomada del PDB como en la estructura modelada, para poder clasificar estructuralmente las regiones carboxilo terminal.

III RESULTADOS

1. Secuencias seleccionadas y extensión carboxilo terminal:

Las secuencias del criptocromo las seleccionamos teniendo en cuenta lo planteado en la página 20, y se distribuyeron en tres grupos:

- Insectos: 12 secuencias en 8 especies
- Vertebrados: 32 secuencias en 13 especies
- Plantas: 32 secuencias en 16 especies

El motivo por el cual conformamos grupos y subgrupos de secuencias fue la gran cantidad de alineamientos que tendríamos que hacer al compararlos todos al mismo tiempo (aproximadamente 2850 alineamientos), situación que dificultaría el análisis de datos.

En total se tomaron 37 especies y 76 secuencias de criptocromos (Tablas 3, 4, 5)

a. Insectos

Tabla 3: Secuencias de insectos seleccionadas

Número de acceso NCBI	Proteína	Organismo	Longitud aa	Longitud C-T aa	Abreviatura
AAK11644	Criptocromo	<i>Antheraea pernyi</i>	525	26	4I
AAC83828	Criptocromo	<i>Drosophila melanogaster</i>	542	32	1I
BAF45421	Criptocromo	<i>Dianemobius nigrofasciatus</i>	539	34	5I
AAX58599	Criptocromo	<i>Danaus plexippus</i>	534	35	3I
ABB29886	Criptocromo 1	<i>Anopheles gambiae</i>	545	35	9I
NP_001076794	Criptocromo 2	<i>Tribolium castaneum</i>	535	35	6I
AAU14170	Criptocromo	<i>Bactrocera tryoni</i>	547	37	2I
AAY23345	Criptocromo	<i>Mamestra brassicae</i>	548	49	7I
ABO38437	Criptocromo 2	<i>Apis mellifera</i>	570	65	8I
ABA62409	Criptocromo 2	<i>Danaus plexippus</i>	742	229	11I
ABO38435	Criptocromo 2	<i>Antheraea pernyi</i>	788	270	10I
ABB29887	Criptocromo 2	<i>Anopheles gambiae</i>	961	473	12I

b. Vertebrados

Tabla 4: Secuencias de vertebrados seleccionadas

Número de acceso NCBI	Proteína	Organismo	Longitud aa	Longitud C-T aa	Abreviatura
NP_571863	Criptocromo 5	<i>Danio rerio</i>	519	32	1AV
AAP22940	Criptocromo 1	<i>Arvicanthus ansorgei</i>	155	34	2AV
NP_001034685	Criptocromo 4	<i>Gallus gallus</i>	529	43	3AV
AAP13561	Criptocromo 2	<i>Rana catesbeiana</i>	543	50	4AV
NP_571864	Criptocromo 1a	<i>Danio rerio</i>	556	68	1BV
NP_571862	Criptocromo 4	<i>Danio rerio</i>	558	73	2BV
NP_001082139	Criptocromo 2a	<i>Xenopus laevis</i>	568	76	3BV
NP_001083936	Criptocromo 2b	<i>Xenopus laevis</i>	570	78	4BV
AAU86901	Criptocromo 2	<i>Sylvia borin</i>	244	85	1CV
ABL14178	Criptocromo 2	<i>Meleagris gallopavo</i>	144	85	2CV
NP_989575	Criptocromo 2	<i>Gallus gallus</i>	582	85	3CV
NP_034093	Criptocromo 2	<i>Mus musculus</i>	592	86	4CV
NP_066940	Criptocromo 2	<i>Homo sapiens</i>	593	86	5CV
NP_596896	Criptocromo 2	<i>Rattus norvegicus</i>	594	88	6CV
ABL14181	Criptocromo 1	<i>Meleagris gallopavo</i>	126	89	7CV
BAB72089	Criptocromo 1	<i>Macaca fascicularis</i>	586	98	1DV
NP_004066	Criptocromo 1	<i>Homo sapiens</i>	586	98	2DV
AAW48291	Criptocromo 1b	<i>Erithacus rubecula</i>	587	99	3DV
ABH03083	Criptocromo 1b	<i>Sylvia borin</i>	587	99	4DV
Q70AD6	Criptocromo 1	<i>Spalax judaei</i>	587	99	5DV
NP_942045	Criptocromo 1	<i>Rattus norvegicus</i>	588	100	6DV
NP_571861	Criptocromo 3	<i>Danio rerio</i>	598	110	7DV
NP_031797	Criptocromo 1	<i>Mus musculus</i>	606	118	1EV
NP_571865	Criptocromo 1b	<i>Danio rerio</i>	606	118	2PV
NP_001081129	Criptocromo 1	<i>Xenopus laevis</i>	616	128	3EV
ABD36589	Criptocromo 1	<i>Podarcis sicula</i>	618	130	1FV
AAW48290*	Criptocromo 1a	<i>Erithacus rubecula</i>	620	132	2FV
CAG14931	Criptocromo 1	<i>Sylvia borin</i>	620	132	3FV
Q5IZC5*	Criptocromo 1	<i>Erithacus rubecula</i>	620	132	Igual 3FV
NP_989576	Criptocromo 1	<i>Gallus gallus</i>	621	133	4FV
NP_571867	Criptocromo 2b	<i>Danio rerio</i>	638	148	2M
NP_571866	Criptocromo 2a	<i>Danio rerio</i>	655	167	1M

c. Plantas

Tabla 5: Secuencias de plantas seleccionadas

Número de acceso NCBI	Proteína	Organismo	Longitud aa	Longitud C-T aa	Abreviatura
ABB01166	Criptocromo 3	<i>Lycopersicon esculentum</i>	577	26	1AP
NP_568461	Criptocromo 3	<i>Arabidopsis thaliana</i>	526	41	2AP
AAL02092	Criptocromo 1b	<i>Lycopersicon esculentum</i>	583	98	1BP
EDO99639	Criptocromo	<i>Chlamydomonas reinhardtii</i>	595	105	2BP
AAO23972	Criptocromo 2b	<i>Pisum sativum</i>	597	112	3BP
NP_171935	Criptocromo 2	<i>Arabidopsis thaliana</i>	612	123	1CP
BAC67176	Criptocromo 2	<i>Armoracia rusticana</i>	617	128	2CP
CAI63876	Criptocromo 2	<i>Brassica napus</i>	618	130	3CP
AAS79665	Criptocromo 2a	<i>Pisum sativum</i>	629	144	1DP
AAF72557	Criptocromo 2	<i>Lycopersicon esculentum</i>	635	149	2DP
ABB36797	Criptocromo 2	<i>Nicotiana glauca</i>	641	160	1EP
BAC56984	Criptocromo 2	<i>Oryza sativa (japonica cultivar-group)</i>	651	162	2EP
CAC82538	Criptocromo 2	<i>Oryza sativa (indica cultivar-group)</i>	651	162	3EP
ABB13334	Criptocromo 2	<i>Hordeum vulgare subsp. vulgare</i>	350	167	4EP
BAA32807	Criptocromo 1	<i>Adiantum capillus-veneris</i>	637	171	5EP
NP_567341	Criptocromo 1	<i>Arabidopsis thaliana</i>	681	189	1FP
ABI31780*	Criptocromo 1	<i>Brassica rapa</i>	704	190	2FP
CAC87903*	Criptocromo 1	<i>Brassica napus</i>	680	190	Igual a 2FV
AAR08429	Criptocromo 1	<i>Orobancha minor</i>	679	192	3FP
AAN37909	Criptocromo 2	<i>Sorghum bicolor</i>	691	193	4FP
AAD44161	Criptocromo 1	<i>Lycopersicon esculentum</i>	679	194	5FP
ABB36796	Criptocromo 1	<i>Nicotiana glauca</i>	681	196	6FP
AAO23970	Criptocromo 1	<i>Pisum sativum</i>	682	197	7FP
ABB13328	Criptocromo 1a	<i>Hordeum vulgare subsp. vulgare</i>	694	197	8FP
BAB70688	Criptocromo 1b	<i>Oryza sativa (japonica cultivar-group)</i>	700	200	1GP
BAA32808	Criptocromo 2	<i>Adiantum capillus-veneris</i>	679	206	2GP
BAB70686	Criptocromo 1a	<i>Oryza sativa (japonica cultivar-group)</i>	710	209	3GP
BAE54523	Criptocromo	<i>Marchantia paleacea var. diptera</i>	696	210	4GP

ABB13331	Criptocromo 1b	<i>Hordeum vulgare subsp. vulgare</i>	712	215	5GP
BAA32809	Criptocromo 3	<i>Adiantum capillus-veneris</i>	718	231	1HP
BAA83338	Criptocromo	<i>Physcomitrella patens</i>	727	239	2HP
EDO97311	Criptocromo	<i>Chlamydomonas reinhardtii</i>	1008	520	3M

2. Determinación de la homología estructural de las regiones carboxilo terminal de los criptocromos:

Secuencias con un valor de homología HCA mayores o iguales a 75% nos indican proteínas con un plegamiento estructural similar, y valores mayores o iguales a 60% señalan similares elementos de estructura secundaria a lo largo de las secuencias alineadas, pero sin ensamble tridimensional.³²³³

a. Insectos: En los insectos se conformaron dos subgrupos de la siguiente manera (Tablas 6 y 7):

Tabla 6: Subgrupo 1 de insectos

Secuencia	Longitud CT (a.a)
4I	26
1I	32
5I	34
3I	35
9I	35
6I	35
2I	37
7I	49
8I	65

Tabla 7: Subgrupo 2 de insectos

Secuencia	Longitud CT (a.a)
11I	229
10I	270

La secuencia de criptocromo 12I no fue asignada por su longitud (473 aa) ningún subgrupo.

La homología obtenida en cada uno de los subgrupos se presenta en la tabla 8.

Tabla 8: Resultados de homología (%) del subgrupo 1 en insectos. Los valores en rojo indican regiones estructuralmente relacionadas

Secuencia	1l	2l	3l	4l	5l	l6	l7	8l	9l
1l	100								
2l	89	100							
3l	60	88	100						
4l	86	88	100	100					
5l	60	75	75	71	100				
6l	56	50	57	50	57	100			
7l	88	78	100	100	71	55	100		
8l	45	45	63	71	63	56	60	100	
9l	70	88	88	88	75	63	78	50	100

En estos primeros alineamientos mostramos que existe una estrecha relación estructural de plegamiento similar entre los elementos del subgrupo 1, con excepción de 6l y 8l. En el subgrupo 2, el valor de homología fue de 77%, entre los elementos 10l y 11l. Después calculamos la homología tomando las secuencias de criptocromo 2l como representativo con las secuencias 6l, 8l, 10l y 12l para determinar la relación estructural entre todas las secuencias de insectos. (Tabla 9)

Tabla 9: Relación para encontrar la homología (%) entre todas las regiones CT en insectos. Las X indica que ya fueron alineadas entre sí.

Secuencia	2l	6l	8l	10l	12l
2l	100				
6l	X	100			
8l	X	X	100		
10l	67	64	50	100	
12l	56	64	56	54	100

A partir de estos resultados las secuencias entre insectos que no tienen similitud estructural entre sí fueron: 2l (representando a 1l, 3l; 4l, 5l, 7l y 9l), 6l, 8l 10l (representando a 11l) y 12l, las cuales se usaron para buscar relaciones estructurales con las demás especies de los otros dos grupo (vertebrados y plantas).

b. Vertebrados: Entre las secuencias de vertebrados, los subgrupos quedaron de la siguiente manera: (Tablas 10, 11, 12, 13, 14 y 15)

Tabla 10: Subgrupo A en vertebrados

Secuencia	Longitud CT (a.a)
1AV	32
2AV	34
3AV	43
4AV	50

Tabla 11: Subgrupo B en vertebrados

Secuencia	Longitud CT (a.a)
1BV	68
2BV	73
3BV	76
4BV	78

Tabla 12: Subgrupo C en vertebrados

Secuencia	Longitud CT (a.a)
1CV	85
2CV	85
3CV	85
4CV	86
5CV	86
6CV	88
7CV	89

Tabla 13: Subgrupo D en vertebrados

Secuencia	Longitud CT (a.a)
1DV	98
2DV	98
3DV	99
4DV	99
5DV	99
6DV	100
7DV	110

Tabla 14: Subgrupo E en vertebrados

Secuencia	Longitud CT (a.a)
1EV	118
2EV	118
3EV	128

Tabla 15: Subgrupo F en vertebrados

Secuencia	Longitud CT (a.a)
1FV	130
2FV	132
3FV	132
4FV	133

Entre las secuencias de los vertebrados no agrupamos las secuencias de criptocromo 1M y 2M, ya que tienen una longitud mucho mayor a la de una región carboxilo terminal de un vertebrado (167 aa y 148, respectivamente), por lo tanto las seleccionamos para ser comparadas con las secuencias de criptocromo de las demás especies (Plantas e insectos).

Para estos subgrupos los valores de homología HCA fueron: (Valores en rojo indican estrecha relación de tipo estructural) (Tablas 16, 17, 18, 19, 20 y 21)

Tabla 16: Valores de homología (%) subgrupo A en vertebrados

Secuencia	1AV	2AV	3AV	4AV
1AV	100			
2AV	25	100		
3AV	20	63	100	
4AV	33	67	33	100

Tabla 17: Valores de homología (%) subgrupo B en vertebrados

Secuencia	1BV	2BV	3BV	4BV
1BV	100			
2BV	50	100		
3BV	56	65	100	
4BV	60	56	78	100

Tabla 18: Valores de homología (%) subgrupo C en vertebrados

Secuencia	1CV	2CV	3CV	4CV	5CV	6CV	7CV
1CV	100						
2CV	89	100					
3CV	94	94	100				
4CV	71	72	71	100			
5CV	76	78	76	94	100		
6CV	71	72	71	100	94	100	
7CV	63	69	75	68	61	56	100

Tabla 19: Valores de homología (%) subgrupo D en vertebrados

Secuencia	1DV	2DV	3DV	4DV	5DV	6DV	7DV
1DV	100						
2DV	100	100					
3DV	74	74	100				
4DV	74	74	100	100			
5DV	100	100	79	74	100		
6DV	95	95	75	70	95	100	
7DV	59	59	50	55	58	67	100

Tabla 20: Valores de homología (%) subgrupo E en vertebrados

Secuencia	1EV	2EV	3EV
1EV	100		
2EV	68	100	
3EV	71	69	100

Tabla 21: Valores de homología (%) subgrupo F en vertebrados

Secuencia	1FV	2FV	3FV	4FV
1FV	100			
2FV	71	100		
3FV	71	100	100	
4FV	71	100	100	100

Según los resultados anteriores, podemos deducir que:

- En el subgrupo A solo se encuentra una relación con respecto a la estructura secundaria entre las secuencias 2AV con 3AV y 4AV (63% y 67%, valor de homología HCA respectivamente), por lo tanto todas estas secuencias de criptocromos las tomamos para un nuevo alineamiento entre las demás regiones carboxilo terminal entre vertebrados para encontrar posibles relaciones de tipo estructural.
- El grupo B, solo nos mostró estrecha relación estructural entre las secuencias de criptocromo 3BV y 4BV (78% valor de homología), por lo tanto para el próximo alineamiento se tomaron las secuencias de criptocromo 1BV, 2BV y 3BV (por 3BV y 4BV).
- Los subgrupos C y D muestran entre si el mayor número de secuencias con plegamiento similar (6 en cada subgrupo) con valores de homología HCA entre 75% y 100%. Para encontrar más relaciones entre vertebrados tomamos las siguientes secuencias del subgrupo C: 7CV y 5CV (5CV por 1CV, 2CV, 3CV, 4CV y 6CV) y del subgrupo D: 7DV y 6DV (6DV por 1DV, 2DV, 3DV, 4DV y 5DV).
- En el subgrupo E encontramos relaciones de estructura secundaria (entre 68% y 71%), y las usamos en su totalidad para los próximos alineamientos.
- Para el último subgrupo de vertebrados (F) los valores de homología son de 100% con excepción de 1FV (71%), por esto las secuencias de los criptocromos 1FV y 4FV fueron tomadas para los próximos alineamientos.

Para encontrar la totalidad de relaciones estructurales en vertebrados seleccionamos entonces las siguientes secuencias de criptocromo: 1AV, 2AV, 3AV, 4AV, 1BV, 2BV, 3BV, 5CV, 7CV, 6DV, 7DV, 1EV, 2EV, 3EV, 1FV y 4FV. Los valores de homología se muestran a continuación: (Tabla 22)

Tabla 22: Resultados de homología (%) para relacionar las secuencias en vertebrados

Secuencia	1AV	1BV	1EV	1FV	2AV	2BV	2EV	3AV	3BV	3EV	4AV	4FV	5CV	6DV	7CV	7DV
1AV	100															
1BV	43	100														
1EV	50	80	100													
1FV	40	57	55	100												
2AV	X	70	90	53	100											
2BV	29	X	56	48	67	100										
2EV	50	67	X	58	90	50	100									
3AV	X	43	57	38	X	57	44	100								
3BV	17	X	67	58	64	X	60	33	100							
3EV	43	76	X	65	90	67	X	56	68	100						
4AV	X	62	69	55	X	58	77	X	62	67	100					
4FV	50	80	83	X	90	54	67	50	67	79	85	100				
5CV	25	60	72	54	78	64	58	29	67	63	55	72	100			
6DV	50	73	94	58	90	40	65	57	67	75	54	90	73	100		
7CV	25	65	88	67	90	64	72	50	75	80	77	100	X	78	100	
7DV	0	60	58	46	78	43	57	67	67	48	42	53	63	X	65	100

Después de obtenidos todos los valores de homología HCA en vertebrados podemos definir tres conjuntos con secuencias relacionadas estructuralmente (tanto de plegamiento, como de estructura secundaria) :(Tabla 23)

Tabla 23: Conjuntos de secuencias con relación estructural en vertebrados basados en los valores de homología de HCA

	Mismo plegamiento estructural	Conservación de estructura secundaria
Conjunto 1	3BV; 4BV	2BV
Conjunto 2	1CV; 2CV; 3CV; 4CV; 5CV; 6CV	
Conjunto 3	1DV; 2DV; 3DV; 4DV; 5DV; 6DV; 2FV; 3FV; 4FV; 2AV; 4AV; 1BV; 7CV; 1EV; 3EV.	2EV; 1FV; 7DV; 3AV

La secuencia de criptocromo 1AV fue la única secuencia en la que no se detectó homología en vertebrados (entre 25% a 50% entre las demás secuencias de vertebrados). Y a partir de los conjuntos obtenidos, las secuencias escogidas para buscar relaciones con las demás especies son: 1AV, 3BV, 5CV, 6DV, 4FV Y 7CV.

c. Plantas: Los subgrupos conformados en plantas quedaron de la siguiente manera: (Tablas 24, 25, 26, 27, 28, 29, 30 y 31)

Tabla 24: Subgrupo A en plantas

Secuencia	Longitud CT (a.a)
1AP	26
2AP	41

Tabla 25: Subgrupo B en plantas

Secuencia	Longitud CT (a.a)
1BP	98
2BP	105
3BP	112

Tabla 26: Subgrupo C en plantas

Secuencia	Longitud CT (a.a)
1CP	123
2CP	128
3CP	130

Tabla 27: Subgrupo D en plantas

Secuencia	Longitud CT (a.a)
1DP	144
2DP	149

Tabla 28: Subgrupo E en plantas

Secuencia	Longitud CT (a.a)
1EP	160
2EP	162
3EP	162
4EP	167
5EP	171

Tabla 29: Subgrupo F en plantas

Secuencia	Longitud CT (a.a)
1FP	189
2FP	190
3FP	192
4FP	193
5FP	194
6FP	196
7FP	197
8FP	197

Tabla 30: Subgrupo G en plantas

Secuencia	Longitud CT (a.a)
1GP	200
2GP	206
3GP	209
4GP	210
5GP	215

Tabla 31: Subgrupo H en plantas

Secuencia	Longitud CT (a.a)
1HP	231
2HP	239

Entre estos subgrupos no agrupamos la secuencia de criptocromo 3M debido a su gran longitud (520aa) la cual no permitió acomodarse a ninguno de ellos. Los valores de homología obtenidos para estos subgrupos fueron: (Tablas 32, 33, 34, 35, 36, 37, 38 y 39)

Tabla 32: Valores de homología (%) subgrupo A en plantas

Secuencia	1AP	2AP
1AP	100	
2AP	0	100

Tabla 33: Valores de homología (%) subgrupo B en plantas

Secuencia	1BP	2BP	3BP
1BP	100		
2BP	23	100	
3BP	37	18	100

Tabla 34: Valores de homología (%) subgrupo C en plantas

Secuencia	1CP	2CP	3CP
1CP	100		
2CP	87	100	
3CP	93	91	100

Tabla 35: Valores de homología (%) subgrupo D en plantas

Secuencia	1DP	2DP
1DP	100	
2DP	55	100

Tabla 36: Valores de homología (%) subgrupo E en plantas

Secuencia	1EP	2EP	3EP	4EP	5EP
1EP	100				
2EP	59	100			
3EP	59	98	100		
4EP	60	77	79	100	
5EP	59	52	52	44	100

Tabla 37: Valores de homología (%) subgrupo F en plantas

Secuencia	1F	2F	3F	4F	5F	6F	7F	8F
1F	100							
2F	93	100						
3F	71	73	100					
4F	64	69	67	100				
5F	76	78	72	70	100			
6F	71	76	76	71	91	100		
7F	77	82	61	58	64	68	100	
8F	70	68	62	79	72	74	56	100

Tabla 38: Valores de homología (%) subgrupo G en Plantas

Secuencia	1GP	2GP	3GP	4GP	5GP
1GP	100				
2GP	63	100			
3GP	83	59	100		
4GP	63	65	62	100	
5GP	92	63	78	60	100

Tabla 39: Valores de homología (%) subgrupo H en plantas

Secuencia	1HP	2HP
1HP	100	
2HP	52	100

Entre los subgrupos de plantas encontramos menos relaciones de homología en comparación a los vertebrados e insectos.

- Los subgrupos A, B, D y H no presentan entre sí ningún tipo de relación de homología
- El subgrupo C se relaciona entre si con altos valores de homología (entre 87% y 91%).
- En el subgrupo E las relaciones estructurales solo están entre las secuencias de los criptocromos 2EP, 3EP y 4EP.
- En el subgrupo donde existen mayor número de secuencias de criptocromo relacionadas es el F: 1FP, 2FP, 5FP, 6FP y 7FP. Además de la alta conservación de la estructura secundaria (entre 61% y 74%)
- Al igual que el subgrupo E, las secuencias de los criptocromos del subgrupo G presentan tres secuencias con similitud estructural y conservación de la estructura secundaria.

Para encontrar las relaciones totales entre el grupo de las plantas tomamos las secuencias de criptocromo: 1AP, 2AP, 1BP, 2BP, 3BP, 2CP, 1DP, 2DP, 1EP, 2EP, 5EP, 2FP, 3FP, 4FP, 8FP, 1GP, 2GP, 4GP, 1HP y 2HP. (Tabla 40)

Tabla 40: Valores de homología HCA (%) para encontrar la relación estructural entre todas las regiones C-T de los criptocromos en plantas

Secuencia	1AP	1BP	1DP	1EP	1GP	1HP	2AP	2BP	2CP	2DP	2EP	2FP	2GP	2HP	3BP	3FP	4FP	4GP	5EP	8FP
1AP	100																			
1BP	22	100																		
1DP	40	50	100																	
1EP	20	57	55	100																
1GP	29	70	53	58	100															
1HP	29	69	50	58	64	100														
2AP	X	X	67	60	64	58	100													
2BP	0	X	36	24	18	22	22	100												
2CP	25	66	55	58	70	68	54	20	100											
2DP	29	56	X	72	50	56	55	17	47	100										
2EP	22	64	47	X	68	65	50	14	64	58	100									
2FP	33	70	45	52	74	54	56	17	52	49	61	100								
2GP	29	70	46	43	X	57	47	20	62	44	59	60	100							
2HP	17	66	40	47	58	X	54	22	60	52	59	59	73	100						
3BP	50	X	55	50	46	44	55	X	37	52	50	46	43	47	100					
3FP	25	71	40	39	65	59	62	20	63	51	55	X	61	58	43	100				
4FP	40	78	47	40	82	64	46	21	66	42	63	X	67	68	32	X	100			
4GP	50	57	50	54	X	55	67	22	56	50	65	55	X	63	38	61	62	100		
5EP	40	65	52	X	63	67	38	16	51	54	X	53	67	67	42	60	57	58	100	
8FP	29	69	48	47	78	56	64	28	59	50	63	X	63	60	38	X	X	64	63	100

Con los valores de homología HCA ya definidos para plantas, podemos definir en ellas los siguientes conjuntos: (Tabla 41)

Tabla 41: Conjuntos de secuencias con relación estructural en plantas basados en los valores de homología de HCA

	Mismo plegamiento estructural	Conservación de estructura secundaria
Conjunto 1	1CP; 2CP; 3CP	
Conjunto 2	2EP; 3EP; 4EP;	
Conjunto 3	1FP; 2FP; 5FP; 6FP; 7FP.	3FP
Conjunto 4	1GP; 3GP; 5GP; 4FP; 8FP.	2AP; 1BP; 1DP; 5EP; 3FP; 4FP; 2GP; 4GP; 1HP; 2HP; 2DP; 1EP

Las secuencias carboxilo terminal de los criptocromos 1AP, 2BP y 3BP, no les detectamos homología de tipo estructural entre plantas y junto con las secuencias carboxilo terminal 2CP, 2EP, 2FP, 1GP y 2DP fueron comparadas para hallar una posible homología estructural secundaria entre las demás especies (vertebrados e insectos).

d. Entre secuencias: Para esta parte tomamos las secuencias que anteriormente se habían determinado para tal fin, a modo de comparar la totalidad de las secuencias carboxilo terminal de los criptocromos que se escogieron. (Tabla 42)

En estos resultados, las relaciones de homología que encontramos son solo unas pocas en comparación con las secuencias de nuestro estudio (Tabla 42). Entre especies las relaciones de valor de homología altos (mayores o iguales a 75%) están entre insectos y vertebrados: la secuencia 8I está estrechamente relacionada con 1AV y 5CV, asimismo con una zona de la secuencia 2M y otra zona de 1M (vertebrados). La otra secuencia de insecto relacionada con vertebrados es la 6I, además se encontró que la 6I se relaciona con una secuencia de planta (2EP), que a su vez se relacionan en estructura secundaria con dos tipos de secuencias en vertebrados (4FV y 7CV), lo cual muestra la única relación entre vertebrados y plantas.

Tabla 42: Valores de homología HCA (%) para encontrar relación entre todas las secuencias C-t de criptocromo seleccionadas

Secuencia	2I	6I	8I	10I	12I	1AV	3BV	5CV	6DV	4FV	7CV	1AP	2BP	3BP	2CP	2EP	2FP	1GP	2DP	1M	2M	3M	
2I	100																						
6I	X	100																					
8I	X	X	100																				
10I	X	X	X	100																			
12I	X	X	X	X	100																		
1AV	44	43	80	38	44	100																	
3BV	58	67	71	57	62	X	100																
5CV	50	67	82	46	65	X	X	100															
6DV	50	78	62	52	67	X	X	X	100														
4FV	56	80	71	56	59	X	X	X	X	100													
7CV	44	80	69	47	68	X	X	X	X	X	100												
1AP	20	33	50	25	40	50	25	67	67	17	33	100											
2BP	44	33	23	17	25	40	19	20	6	17	17	X	100										
3BP	56	46	62	50	55	33	44	56	47	47	58	X	X	100									
2CP	57	67	45	61	67	50	50	48	59	45	55	X	X	X	100								
2EP	67	78	50	52	73	44	61	62	43	62	64	X	X	X	X	100							
2FP	36	70	69	65	56	27	67	52	55	44	62	X	X	X	X	X	100						
1GP	67	80	47	56	64	33	61	57	50	46	52	X	X	X	X	X	X	100					
2DP	67	89	60	60	63	50	62	40	38	46	60	X	X	X	X	X	X	X	100				
1M	56	67	75	53	56	33	78	67	80	62	72	33	14	40	58	46	59	51	56	100			
2M	56	70	79	44	47	25	61	76	56	61	53	67	6	50	50	44	42	44	46	43	100		
3M	55	67	69	51	47	43	50	63	55	59	59	29	23	60	62	66	64	65	58	56	54	100	

Esto indica en general que las regiones carboxilo terminal en criptocromos de diferentes especies pueden divergir en su función, es decir que estas regiones en las proteínas son parálogos (Proteínas homologas, es decir criptocromos, pero que divergen en su función).

Esta divergencia no solo se demostró entre especies sino también entre grupos (insectos, vertebrados y plantas), en donde la longitud de secuencias de las regiones carboxilo terminal de criptocromos fue el primer indicio de diferencia entre ellas, pasando por la disposición de los agrupamientos hidrofóbicos para la búsqueda de homólogos, situación nos lleva a proponer que las regiones carboxilo terminal en criptocromos cumplen funciones diversas dependiendo del organismo donde se encuentra, las cuales pretendemos encontrar con el modelamiento de la homología de estructura terciaria.

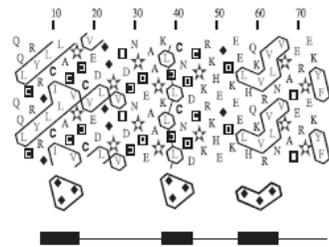
En este análisis hubo secuencias las cuales no presentaron homología de tipo estructural, las cuales no pudimos agrupar: 12I (región C-t del Cry.2 de *Anopheles gambiae*), 2BP (región C-t del Cry. *Chlamydomonas reinhardtii*), 3BP (Región carboxilo terminal del Cry. 2b de *Pisum sativum*) y 3M (Cry. de *Chlamydomonas reinhardtii*).

3. Determinación de la estructura secundaria de las regiones carboxilo terminal de los criptocromos

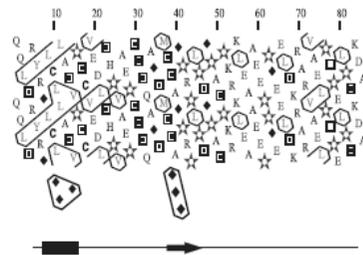
A partir de los resultados de homología, obtenidos anteriormente, pudimos relacionar las regiones carboxilo terminal de los criptocromos de diferentes especies. Para determinar la estructura secundaria de estas regiones tomamos una secuencia de cada agrupamiento de secuencias con similitud estructural (Tabla 43). La determinación de todos los elementos de estructura secundaria se pueden observar en el anexo 2. Las flechas indican elementos de hoja β y las cajas hélices α .

Tabla 43: Relación de las secuencias de criptocromo seleccionadas y elementos para determinar su estructura secundaria por HCA		
Agrupamiento	Elemento escogido	Estructuras Relacionadas
1	2I	1I, 2I, 3I, 4I, 5I, 7I, 9I
2	10I	10I, 11I.
3	3BV	2BV, 3BV, 4BV
4	5CV	1AV, 1CV; 2CV; 3CV; 4CV; 5CV; 6CV, 8I
5	4FV	1DV; 2DV; 3DV; 4DV; 5DV; 6DV; 2FV; 3FV; 4FV; 2AV; 4AV; 1BV; 7CV; 1EV; 3EV, 6I,2EV; 1FV; 7DV; 3AV
6	2CP	1CP; 2CP; 3CP
7	2EP	2EP; 3EP; 4EP;
8	2FP	1FP; 2FP; 5FP; 6FP; 7FP, 3FP
9	1GP	1GP; 3GP; 5GP; 4FP; 8FP., 2AP; 1BP; 1DP; 5EP; 3FP; 4FP; 2GP; 4GP; 1HP; 2HP; 2DP; 1EP
10	1M	1M
11	2M	2M
12	3M	3M
13	12I	12I
14	1AP	1AP

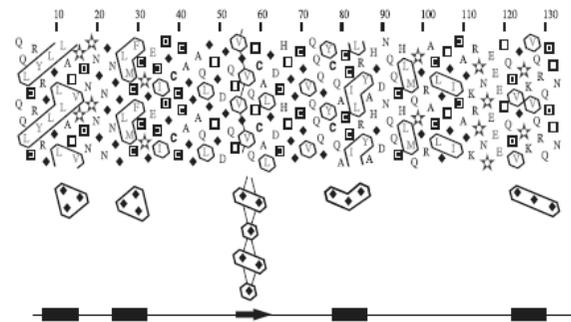
Para determinar los elementos de estructura secundaria, obtuvimos los perfiles de HCA de las secuencias de criptocromo escogidas, luego debajo de ellas colocamos la representación de los agrupamientos hidrofóbicos que corresponden a un elemento de estructura secundaria y por último la representación de la estructura secundaria. Figura 13



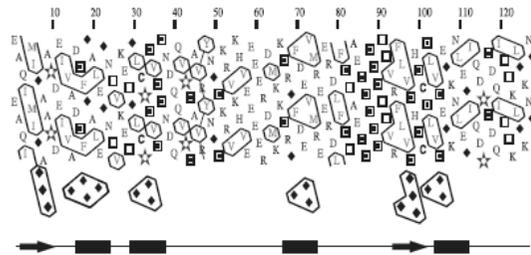
3BV-*Xenopus leavis*



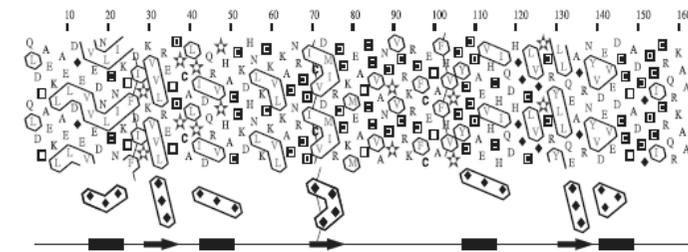
5CV-*Homo sapiens*



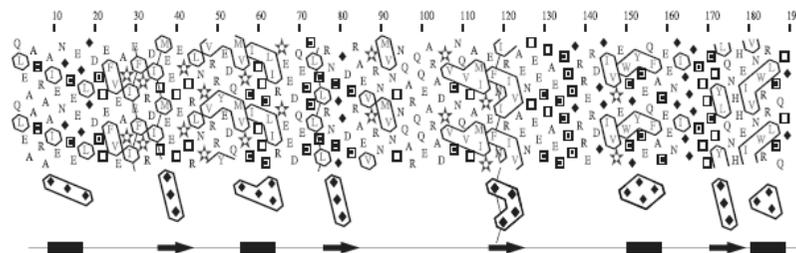
4FV-*Gallus gallus*



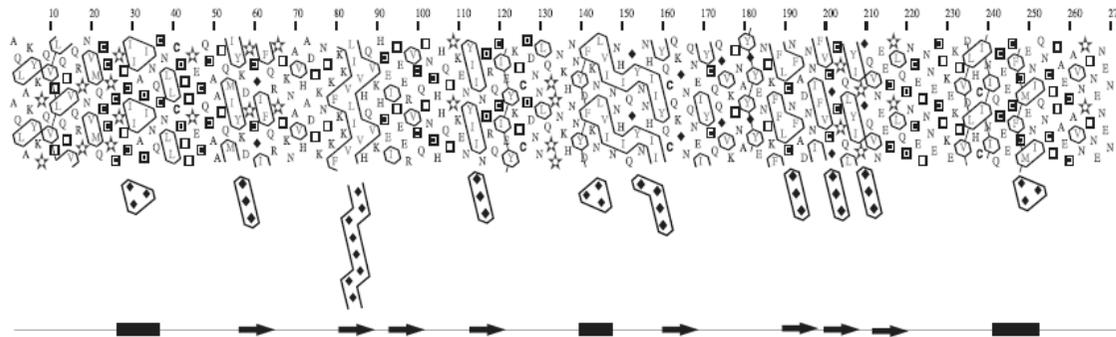
2CP-*Armoracia rusticana*



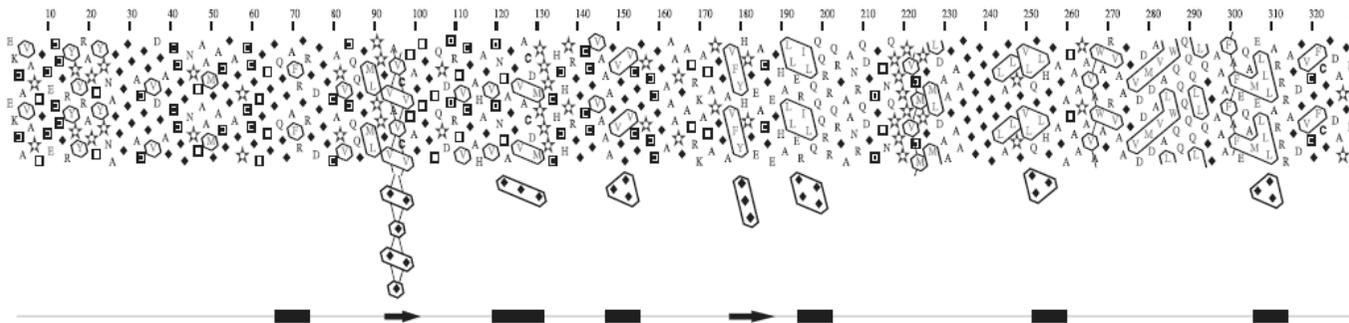
2EP-*Oryza sativa*



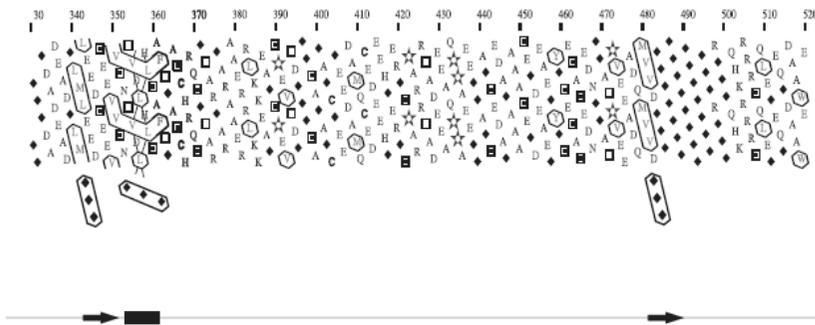
1FP-*Brassica rapa*



10I-*Antheraea pernyi*



3M-*Chlamydomonas reinhardtii*



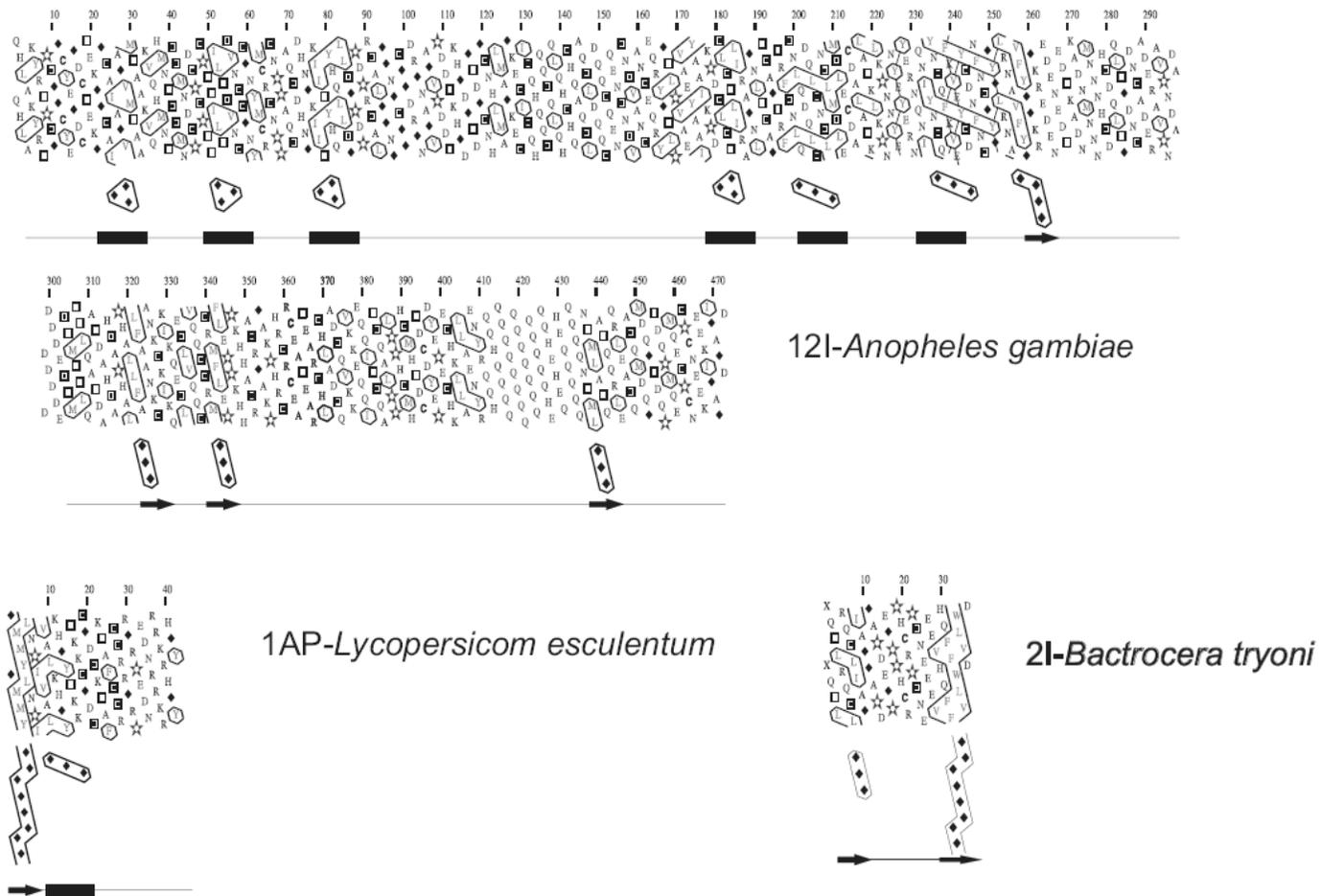


Figura 13: Determinación de los elementos de estructura secundaria de las regiones carboxilo terminal de los criptocromos escogidos como representativos. Las flechas indican hojas β y las cajas indican α -hélices. Entre la representación y secuencia bidimensional se muestra el agrupamiento hidrofóbico detectado por HCA

Comparando estos elementos de estructura determinados por HCA con otros resultados obtenidos anteriormente (Figura 11)⁷:

- La secuencia C-t del criptocromo de *D. melanogaster* por homología de HCA está relacionada con 2l (criptocromo de *B. tryoni*), la cual determinamos sus los elementos de estructura secundaria. Comparando con anteriores trabajos donde mostraron un elemento tipo α -hélice, encontramos una estructura de hoja β . Esta diferencia se debió seguramente a que dentro del agrupamiento hidrofóbico que muestra la hoja β , se puede observar implícitamente un agrupamiento característico de las α -hélices, pero por la cantidad de residuos hidrofóbicos involucrados y la forma del mismo, corresponden a una hoja β .
- La otra secuencia a la cual predijimos anteriormente su estructura secundaria fue a la región c-t del criptocromo de *A. thaliana*, muestra solo elementos de α -hélice, pero en nuestra determinación por HCA en la secuencia relacionada con ella (2FP. C-t de *B. rapa*), se encontró no solo elementos de hélice sino también de hoja β , los cuales observamos claramente en el fragmento analizado.
- Por último la zona c-t del criptocromo 2 de *H. sapiens* la comparamos con resultados publicados con anterioridad, esta predicción si corresponde a los elementos encontrados por medio de HCA para este fragmento (α -hélice y luego hoja β).

Esto demuestra la sensibilidad de HCA para hallar elementos de estructura secundaria por inspección de sus gráficos y que por otros métodos no son fácilmente detectables.

Adicionalmente podemos definir un tipo de clasificación estructural de estas regiones carboxilo terminal según sus elementos de estructura secundaria. (Tabla 44)

Tabla 44: Clasificación de las regiones carboxilo terminal de las secuencias a las cuales se le detectaron sus elementos de estructura secundaria.

Agrupamiento	Elemento escogido	Clasificación estructural
1	2I	β : Estructura exclusivamente hojas β
2	10I	$\alpha + \beta$: presencia de hélices y hojas de manera aleatoria. Predominio de hojas β
3	3BV	α -helicoidal: Solo presencia de α -hélices
4	5CV	α/β : presencia del motivo α/β
5	4FV	$\alpha + \beta$: Con predominio de hélices
6	2CP	$\alpha + \beta$: Con predominio de hélices
7	2EP	α/β : se muestran unidades tipo $\alpha-\beta-\alpha$
8	2FP	α/β : Presencia de dos motivos tipo $\alpha-\beta-\alpha$
9	1GP	$\alpha + \beta$: casi en misma proporción la cantidad de hélices y de hojas
10	1M	α / β : Con predominio de hélices
11	2M	Solo un elemento α . Sin carente de estructura
12	3M	$\alpha + \beta$: Con predominio de hélices
13	12I	Presencia de dos motivos α y un motivo β
14	1AP	Presencia de un motivo α/β

4. Modelamiento tridimensional de la homología encontrada en las regiones carboxilo terminal de los criptocromos.

Para esta parte final del trabajo el método más sensible para hallar proteínas de estructura conocida que puedan relacionarse con nuestros fragmentos fue Blast²⁹, con su variante PsiBlast que es utilizado solo para aminoácidos en la búsqueda de posibles homólogos, el cual compara cada una de nuestras secuencias de criptocromos con una base de datos, que para este caso fue la base de la protein data bank. Las estructuras se identifican con su código PDB.

Para las regiones 2I y 1AP no fueron encontradas similaridades significantes (No significant matches found), por lo tanto no se puede comparar con proteínas con una estructura tridimensional conocida. Para las demás regiones carboxilo terminal no pudimos determinar su homología de estructura terciaria debido a que se encontraban similaridades solo en partes pequeñas de las regiones carboxilo

terminal (entre 20 a 96 residuos de aminoácidos) y no en la extensión de todas las zonas C-T. Los modelos tridimensionales obtenidos se visualizaron en Discovery Studio³⁴ con licencia académica. Los modelos se mostrarán en el Anexo 3. La evaluación de los modelos se hizo por medio de un perfil de energía DOPE³⁵ (Discrete Optimized Protein Energy), es un potencial estadístico optimizado para evaluar modelos, en donde toma una función de energía estándar del programa modelamiento de homología MODELLER. En los perfiles DOPE la escala de energía es arbitraria.

- **1GP**

Tabla 45: Estructuras relacionadas con 1GP			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
1GP	1-96	1K0E	232-333
1GP	70-137	1OGM	5-73

Obtuvimos las siguientes evaluaciones de los modelos de estructura terciaria construidos a partir de las estructuras conocidas, en donde eje x representa la posición del residuo y el eje y el valor arbitrario DOPE. (Figuras 14 y 15).

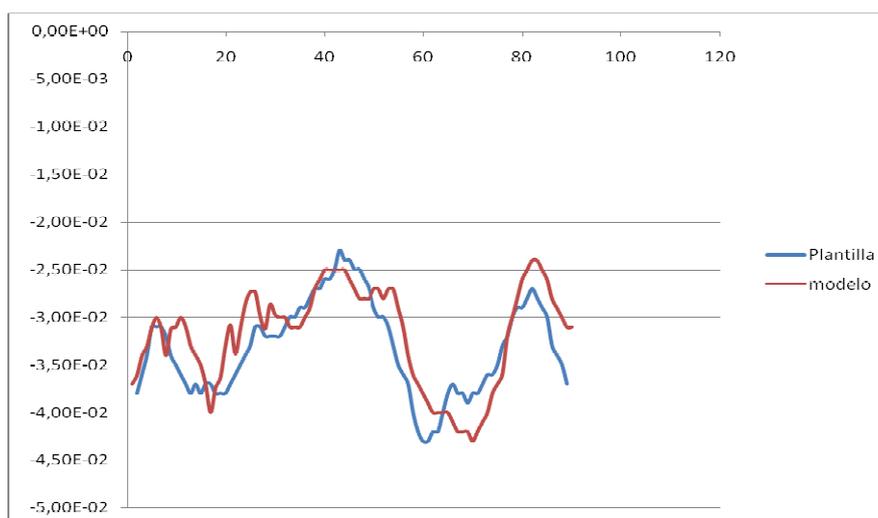


Figura 14: Evaluación de 1GP con respecto a 1K0E

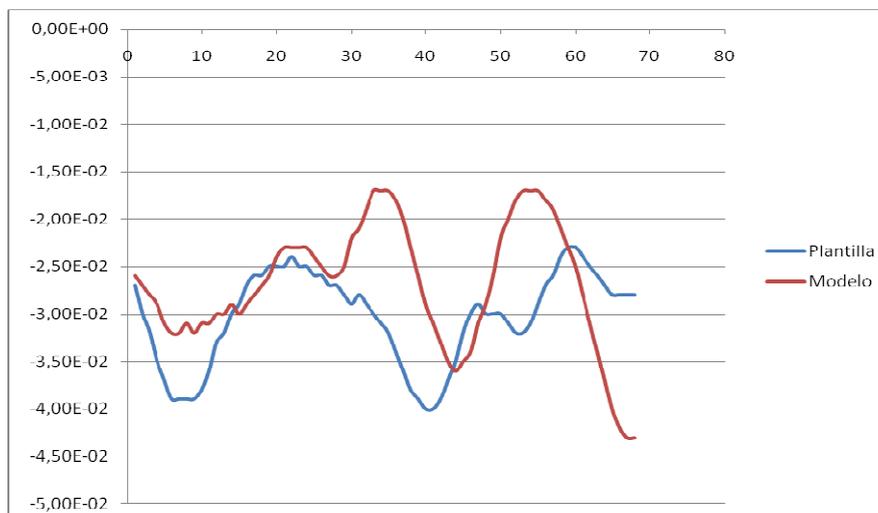


Figura 15: Evaluación de 1GP con respecto a 1OGM

De este modo solo pudimos realizar un modelo de 1GP entre los aminoácidos 1 al 96. La proteína que se utilizó como plantilla 1OGM (Template) tiene función liasa. La clasificación estructural de la proteína 1OGM es $\alpha+\beta$ al igual que lo determinado para 1GP por los perfiles de HCA.

- **5CV**

Tabla 46: Estructuras relacionadas con 5CV			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
5CV	42-85	1AZO	7-50

La evaluación del modelo fue: Figura 16

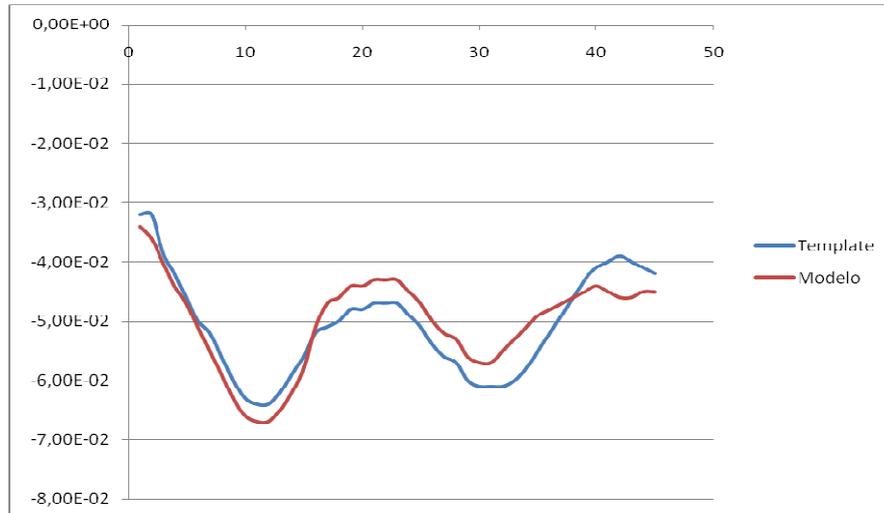


Figura 16: Evaluación de 5CV con respecto a 1AZO

La proteína 1AZO al igual que la 5CV es de clase α/β . La proteína 1AZO funciona como endonucleasa, aunque 5CV no tiene la extensión suficiente del plegamiento tipo $\alpha-\beta-\alpha$, ya que tiene solo dos elementos de estructura secundaria.

- **2CP**

Tabla 47: Estructuras relacionadas con 2CP			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
2CP	35-110	1V3J	76-151
2CP	6-50	1K8O	12-56

La evaluación de los modelos son: Figuras 17 y 18.

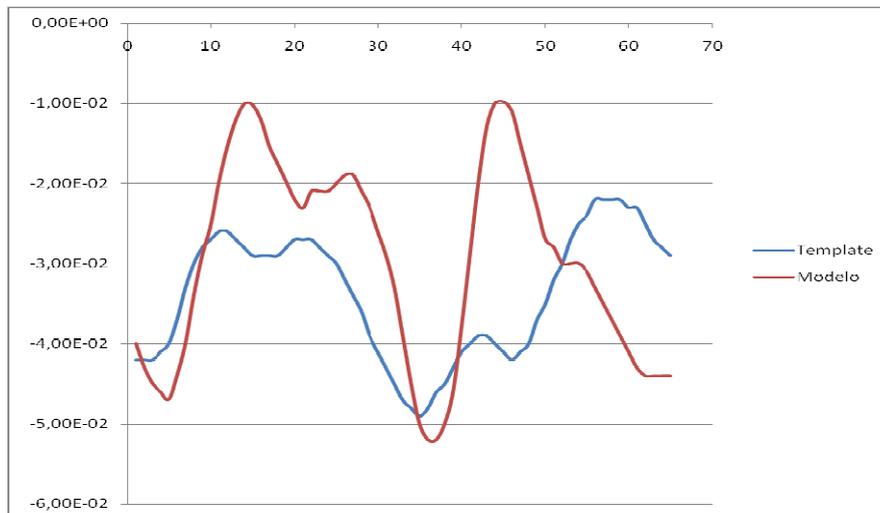


Figura 17: Evaluación de 2CP con respecto a 1V3J

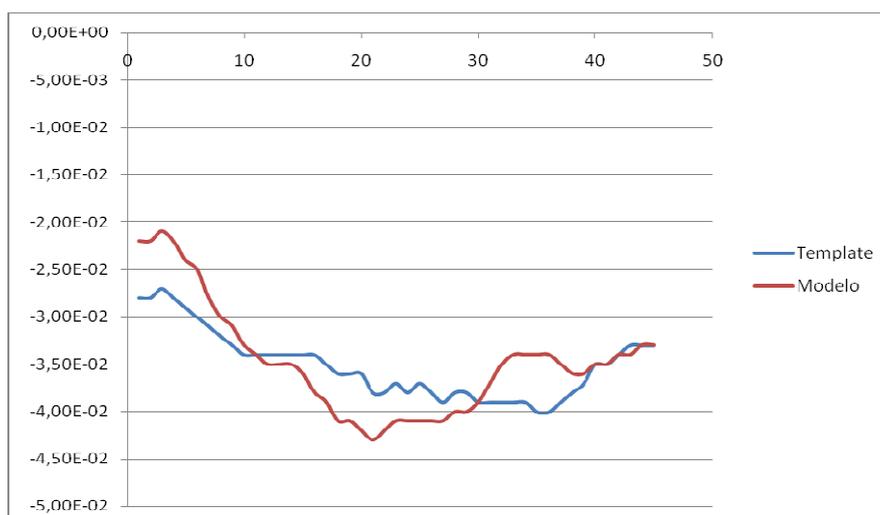


Figura 18: Evaluación de 2CP con respecto a 1K8O

Ninguno de estos dos modelos fueron aceptados para 2CP ya que las evaluaciones de los modelos nos muestran una diferencia notoria entre los valores de energía DOPE con respecto a la posición de los aminoácidos para las estructuras modeladas y las proteínas utilizadas como plantillas en ambas predicciones.

- 1M

Tabla 48: Estructuras relacionadas con 1M			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
1M	98-133	1WYI	193-230
1M	5-28	1BOG	49-72

La evaluación de estos dos modelos son:

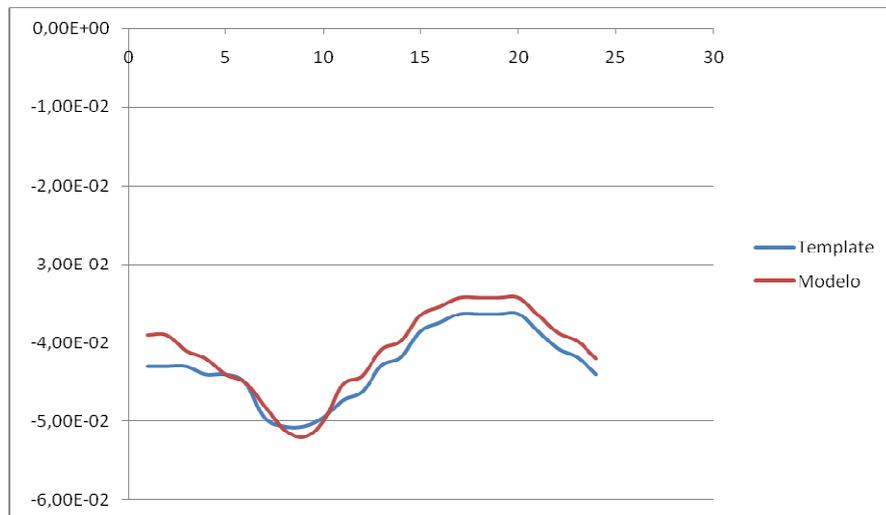


Figura 19: Evaluación de 1M con respecto a 1BOG

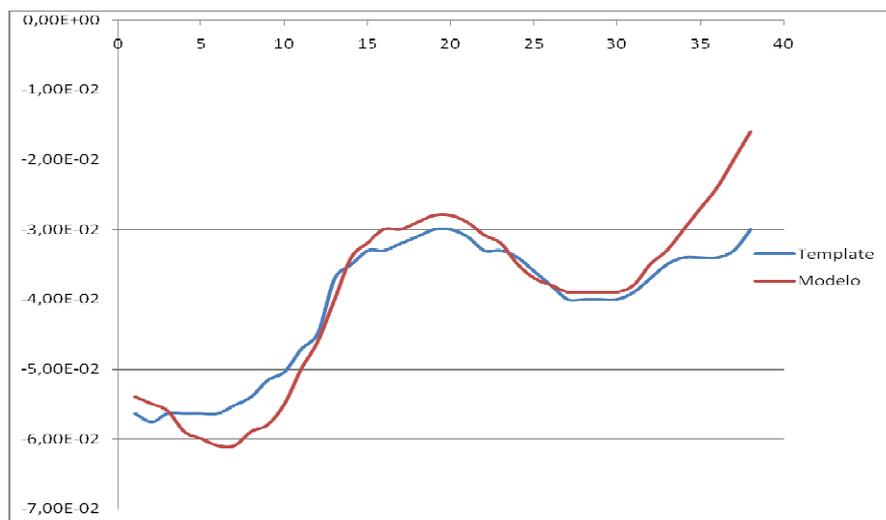


Figura 20: Evaluación de 1M con respecto a 1WYI

A pesar de la sus evaluación, se descarta 1BOG, ya que esta proteína es de clase toda β , y la zona de 1M alineada no evidencia un elemento tipo β . La estructura 1WYI cumple una función de isomerasa y es tipo α/β , al igual que 1M.

- **2EP**

Tabla 49: Estructuras relacionadas con 2EP			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
2EP	49-76	2JE8	548-575
2EP	108-148	1YMV	2-42

Las evaluaciones de los modelos fueron: Figuras 21 y 22.

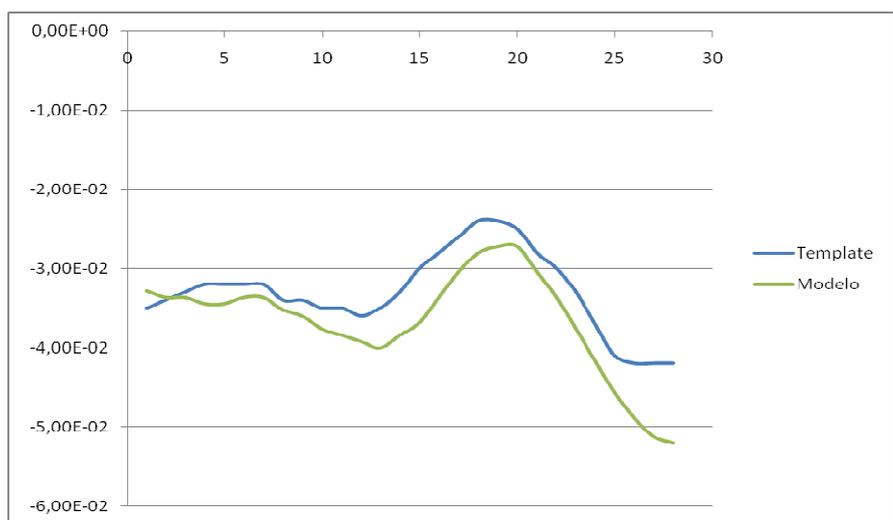


Figura 21: Evaluación de 2EP con respecto a 2JE8

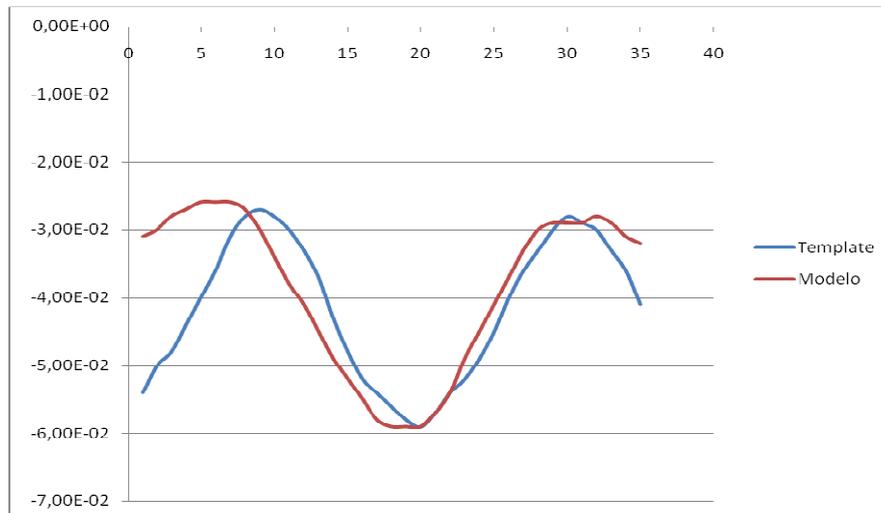


Figura 22: Evaluación de 2EP con respecto a 1YMV

Con ambas estructuras las predicciones coinciden con los elementos de estructura secundaria detectados en 2EP, no obstante 2JE8 es tipo $\alpha+\beta$ y tiene función de quimiotaxis (proteínas que responden a señales químicas dentro de la célula) y 1YMV es tipo α/β y su función es hidrolasa.

- **2FP:**

Tabla 50: Estructuras relacionadas con 2FP			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
2FP	34-100	1I84	991-1057
2FP	113-185	1JR7	115-190

Las evaluaciones de los modelos son: Figuras 23 y 24

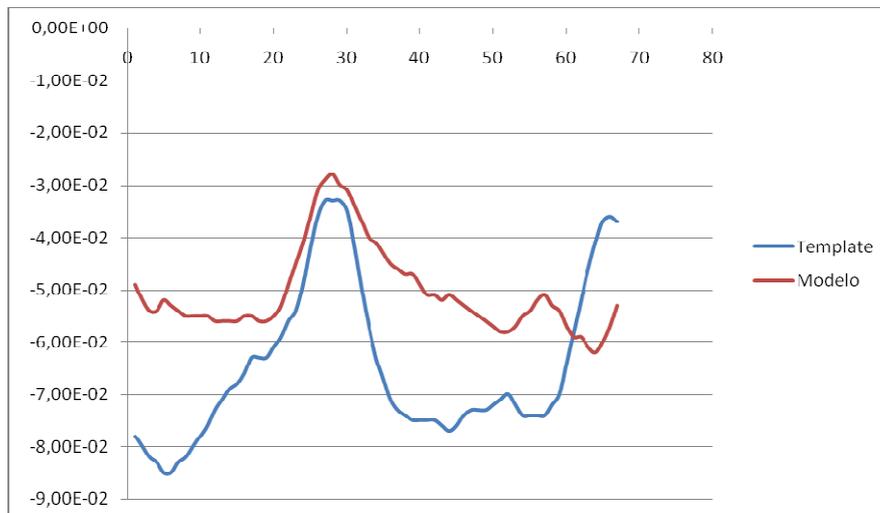


Figura 23: Evaluación de 2FP con respecto a 1184

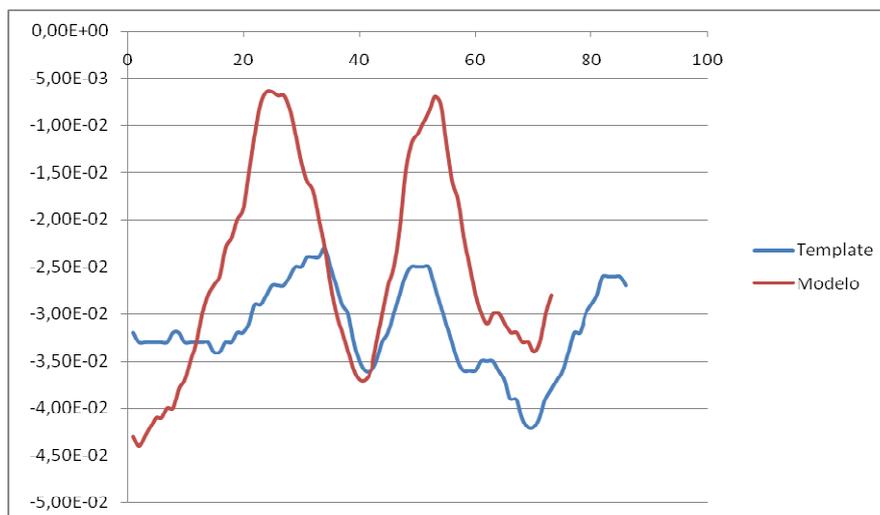


Figura 24: Evaluación de 2FP con respecto a 1JR7

Ninguno de los dos modelos son útiles para 2FP ya que los perfiles DOPE no muestran tendencias similares en la disposición de los aminoácidos a lo largo de la secuencia.

- **4FV:**

Tabla 51: Estructuras relacionadas con 4FV			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
4FV	5-44	1XWN	91-130
4FV	37-77	1JF3	48-93
4FV	72-124	2IX5	331-379

La evaluación de los modelos fue: Figuras 25, 26 y 27.

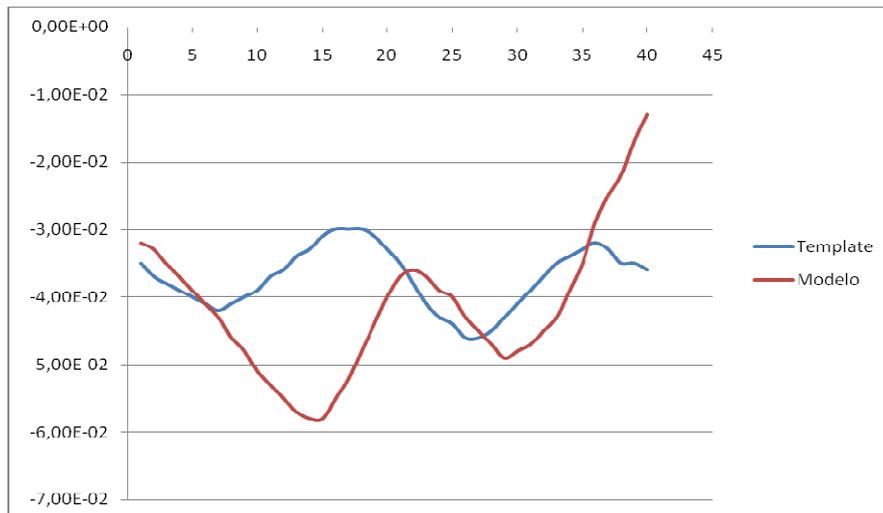


Figura 25: Evaluación de 4FV con respecto a 1XWN

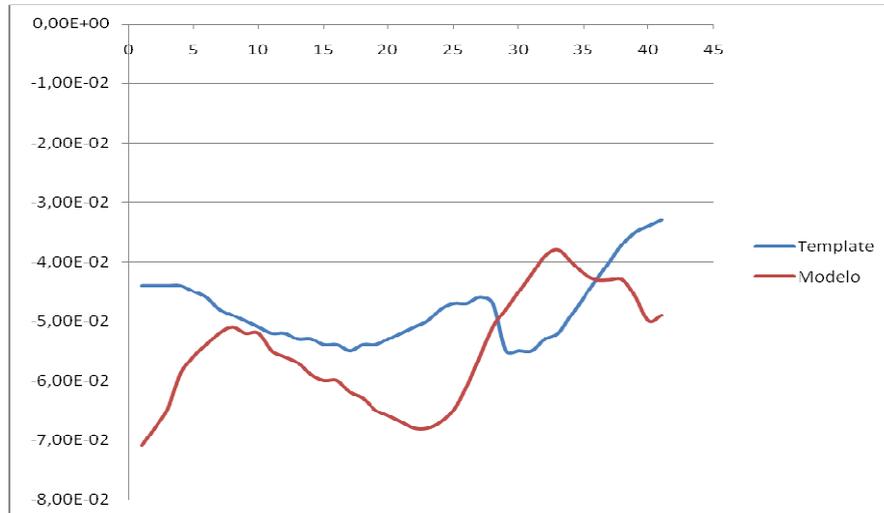


Figura 26: Evaluación de 4FV con respecto a 1JF3

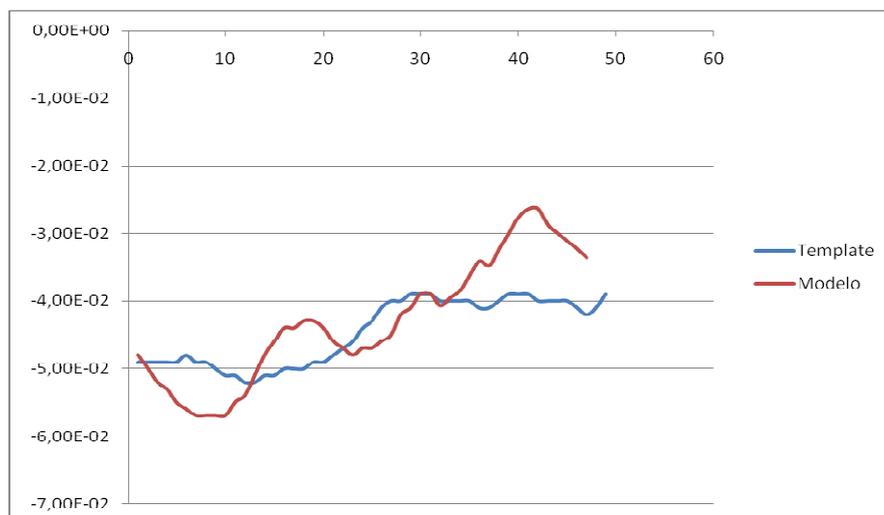


Figura 27: Evaluación de 4FV con respecto a 2IX5

Estos tres fragmentos que modelamos se descartan ya que los perfiles DOPE muestran grandes diferencias a lo largo de los fragmentos de las secuencias de criptocromos comparadas.

- **3BV:**

Tabla 52: Estructuras relacionadas con 3BV			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
3BV	23-62	2PFE	80-119
3BV	11-35	1DI6	158-182

La evaluación de los modelos son: Figuras 26 y 27.

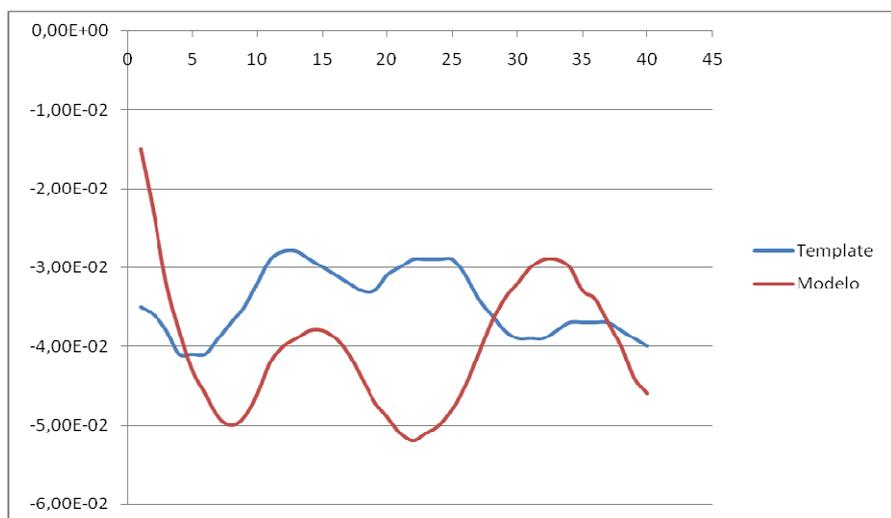


Figura 28: Evaluación de 3BV con respecto a 2PFE

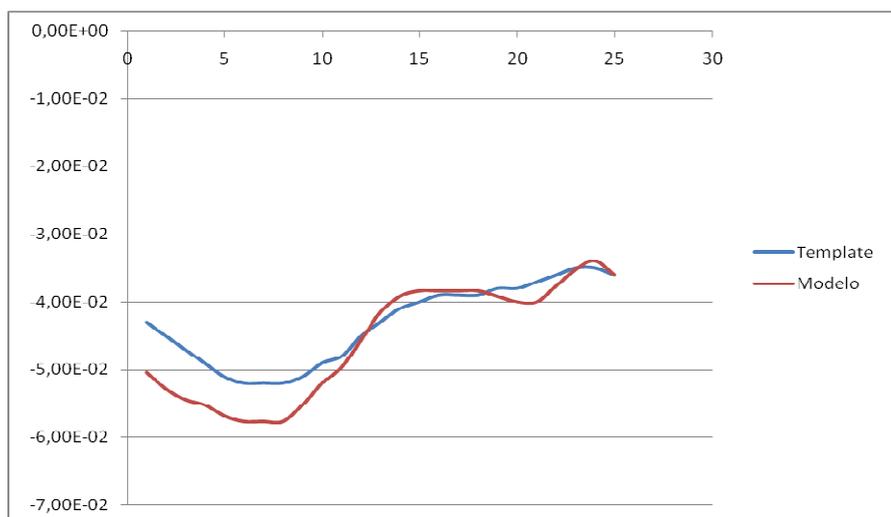


Figura 29: Evaluación de 3BV con respecto a 1DI6

El modelo que construimos usando como plantilla la estructura 2PFE, no se toma en cuenta por la incompatibilidad en los perfiles DOPE. El modelo construido a partir de 1DI6 relaciona una región de α -hélice a pesar de que la estructura de la plantilla es α/β . La función de 1DI6 es desconocida.

- **3M:**

Tabla 53: Estructuras relacionadas con 3M			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
3M	6-41	2BIX	98-133
3M	96-149	2EBS	501-555
3M	153-206	1KXJ	154-205
3M	241-346	2PUA	164-267
3M	350-408	3BIC	368-427

Las evaluaciones que se obtuvieron de los fragmentos de 3M fueron: Figuras 28, 29, 30, 31 y 31

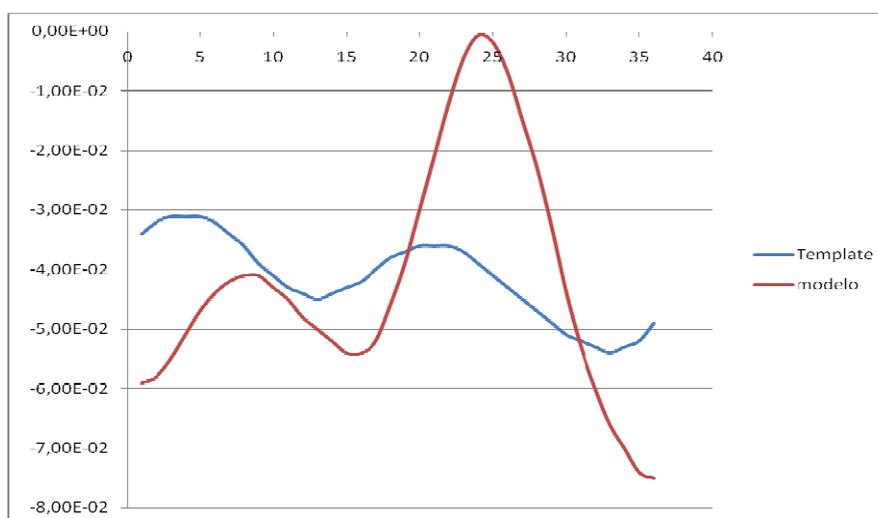


Figura 30: Evaluación de 3M con respecto a 2BIX

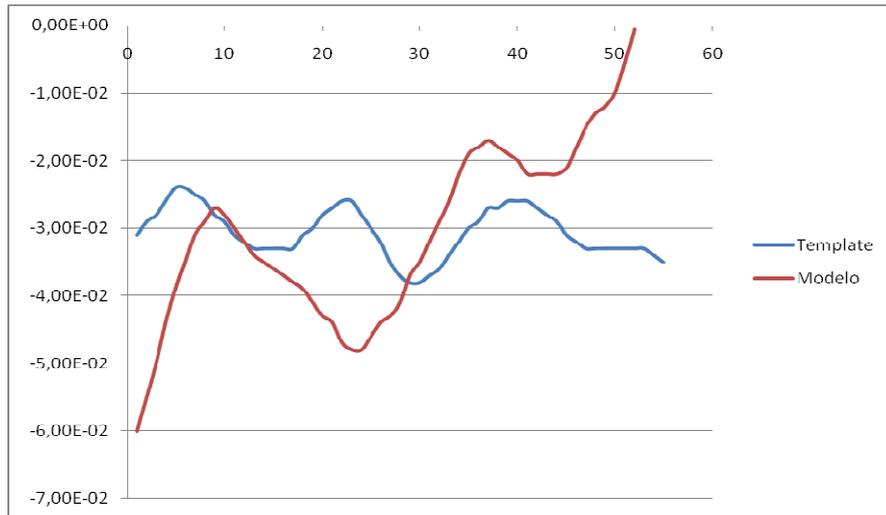


Figura 31: Evaluación de 3M con respecto a 2EBS

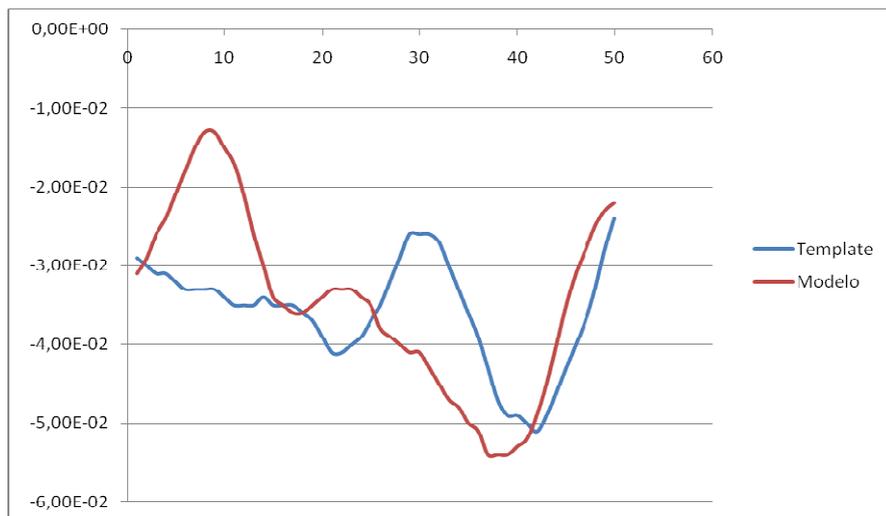


Figura 32: Evaluación de 3M con respecto a 1KXJ

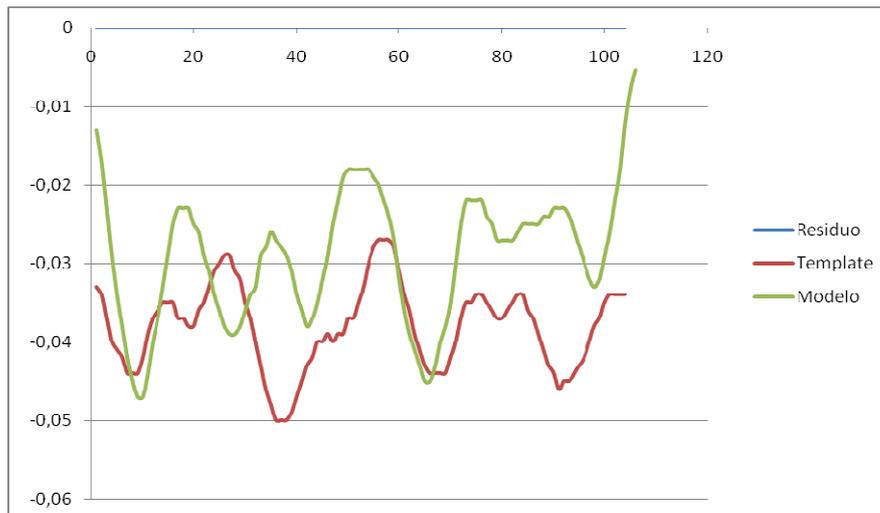


Figura 33: Evaluación de 3M con respecto a 2PUA

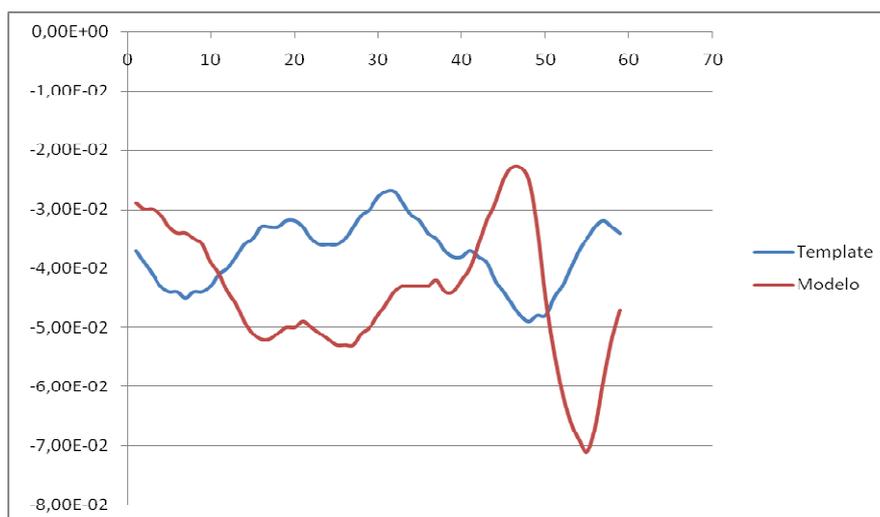


Figura 34: Evaluación de 3M con respecto a 3BIC

En todas las fragmentaciones que se le hicieron a la secuencia de criptocromo 3M, no pudimos encontrar un modelo apropiado para mostrar su homología a nivel tridimensional.

- 101

Tabla 54: Estructuras relacionadas con 10I			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
10I	20-78	1C5B	2-60
10I	38-79	1CCW	308-349
10I	62-191	1XSI	72-203
10I	200-262	1PGV	8-68

Las evaluaciones para estos fragmentos de la secuencia de criptocromo 10I fueron: Figuras 33, 34, 35 y 36.

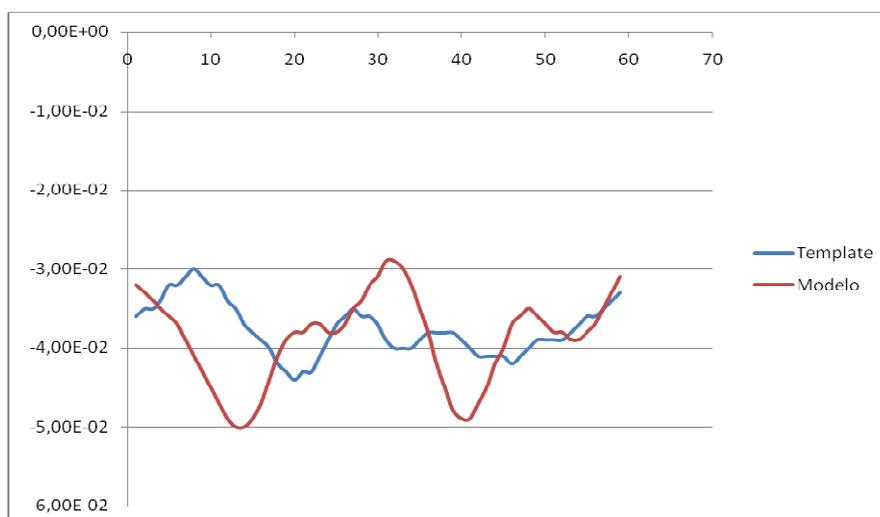


Figura 35: Evaluación de 10I con respecto a 1C5B

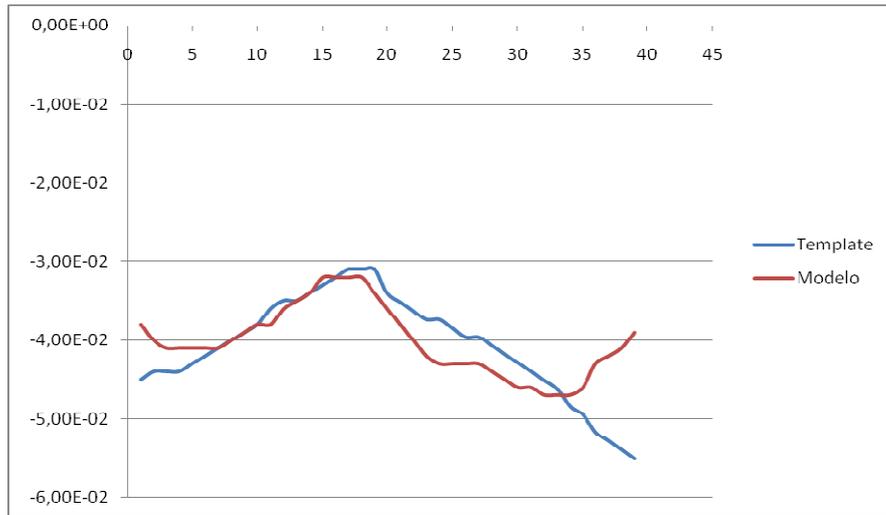


Figura 36: Evaluación de 10I con respecto a 1CCW

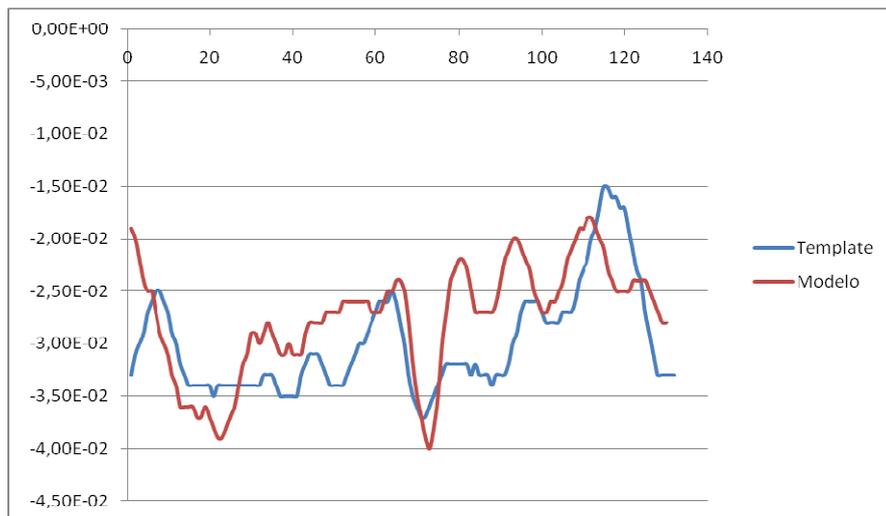


Figura 37: Evaluación de 10I con respecto a 1XSI

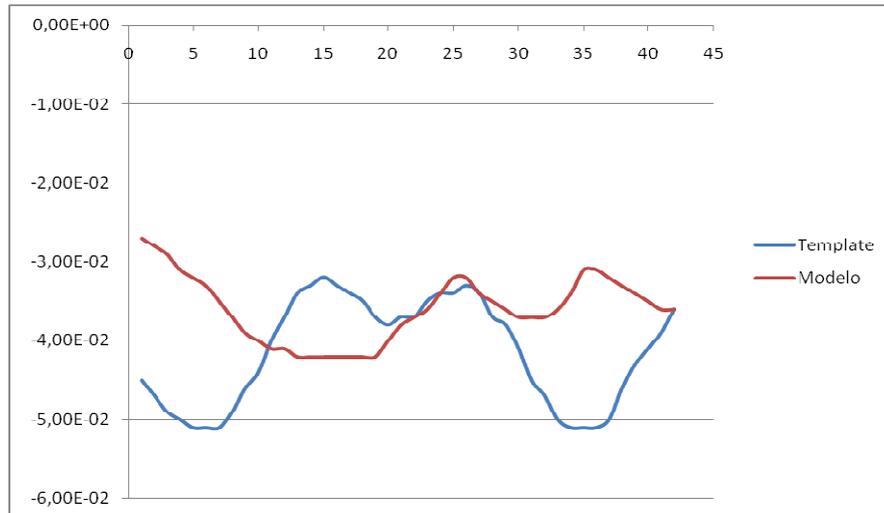


Figura 38: Evaluación de 10l con respecto a 1PGV

El único fragmento modelado aceptado según el perfil DOPE es el comprendido entre 38 a 79 residuos en 10l a partir de la estructura 1CCW, la cual es una estructura tipo α/β , en donde los elementos de estructura secundaria corresponden en ambas secuencias. La función de 1CCW es isomerasa.

- **12l**

Tabla 55: Estructuras relacionadas con 12l			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
12l	42-81	1X5H	13-54
12l	145-180	1BWV	2-37
12l	219-277	2QV3	252-319
12l	243-270	1CC1	235-262
12l	401-454	2OEE	36-319

La evaluación de los modelos fueron: Figuras 37, 38, 39, 40 y 41.

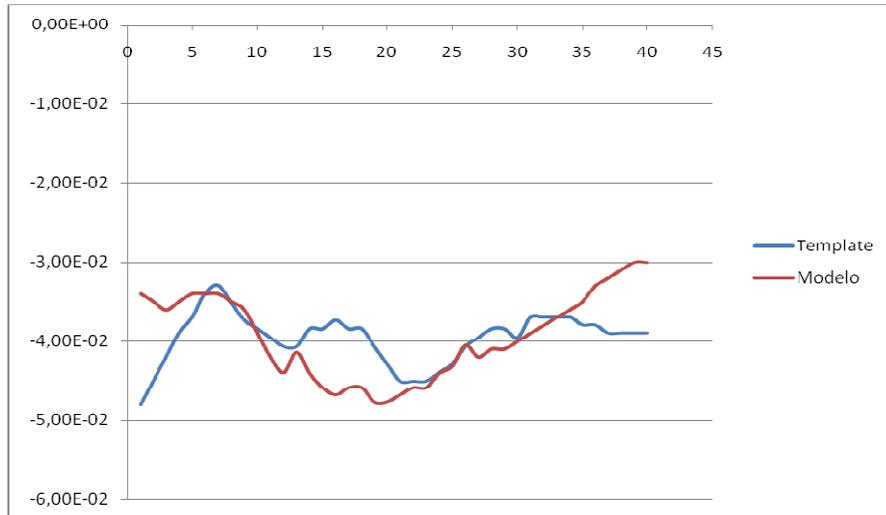


Figura 39: Evaluación de 12I con respecto a 1X5H

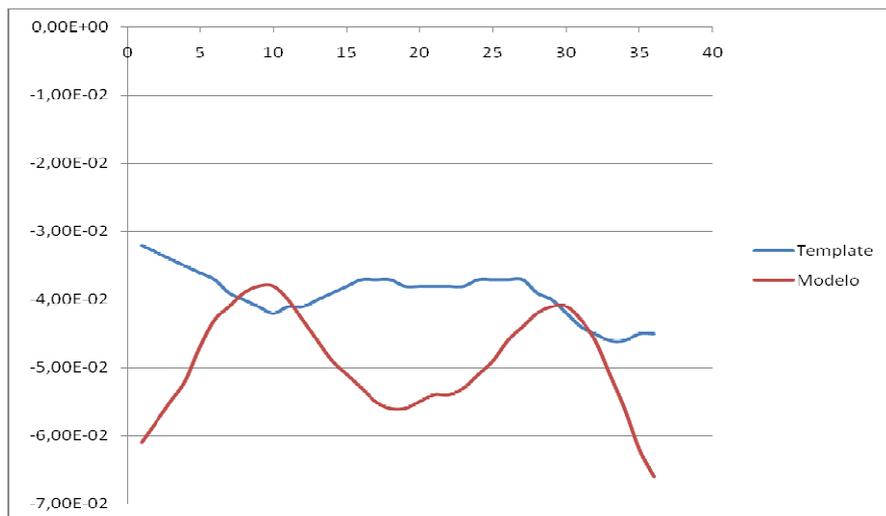


Figura 40: Evaluación de 12I con respecto a 1BWV

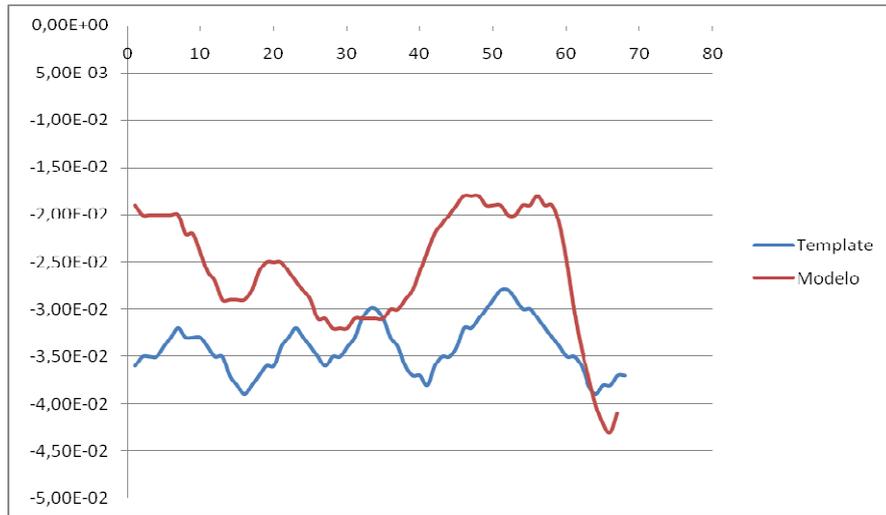


Figura 41: Evaluación de 12I con respecto a 2QV3

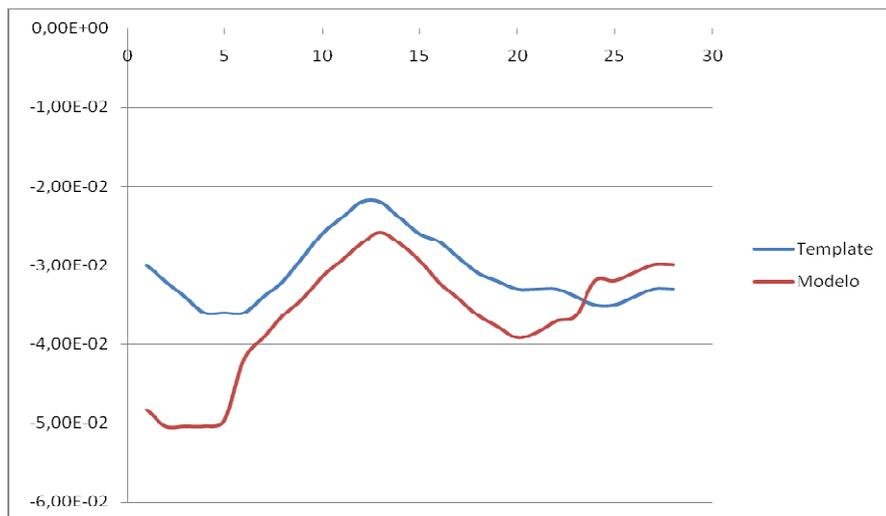


Figura 42: Evaluación de 12I con respecto a 1CC1

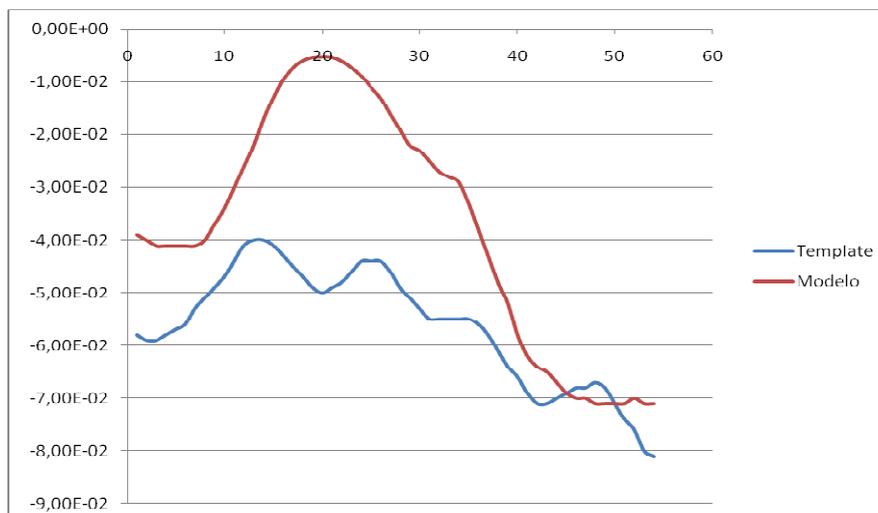


Figura 43: Evaluación de 12I con respecto a 2OEE

De los 5 fragmentos de 12I el único que aceptamos por su evaluación es el construido con 1CC1, en ambas secuencias se muestran correspondencia con los dos elementos de estructura detectadas α - β . La función de la estructura 1CC1 es oxidoreductasa.

- **2M:**

Tabla 56: Estructuras relacionadas con 2M			
Secuencia C-t	Zona C-t (aa)	Estructura (PDB)	Región (aa)
2M	13-65	1C4K	88-139
2M	34-57	1BGP	105-122
2M	54-94	1PMS	69-109
2M	80-139	1Y8Q	63-120

Los modelos construidos mostraron los siguiente perfiles de evaluación: Figuras 42, 43, 44 y 45

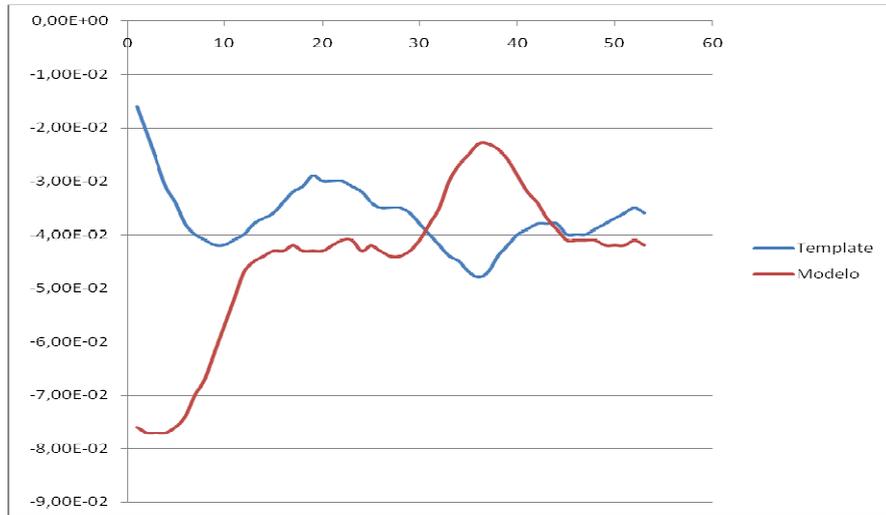


Figura 44: Evaluación de 2M con respecto a 1C4K

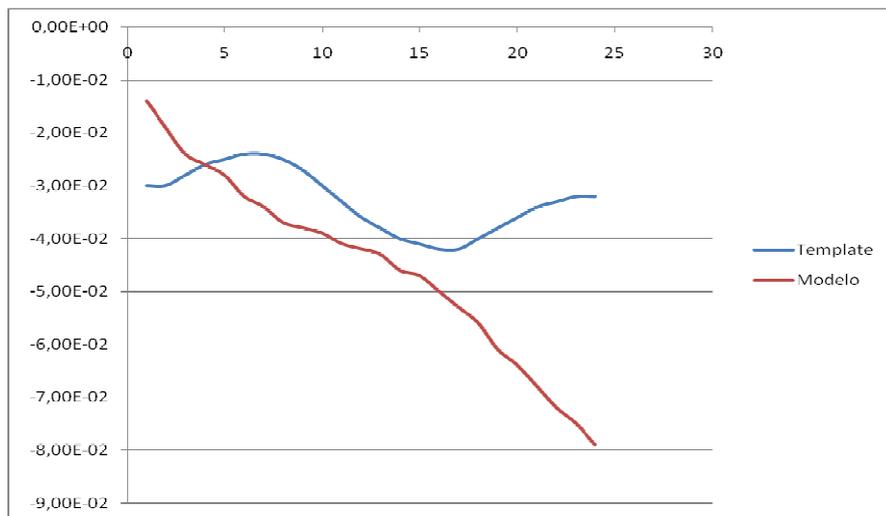


Figura 45: Evaluación de 2M con respecto a 1BGP

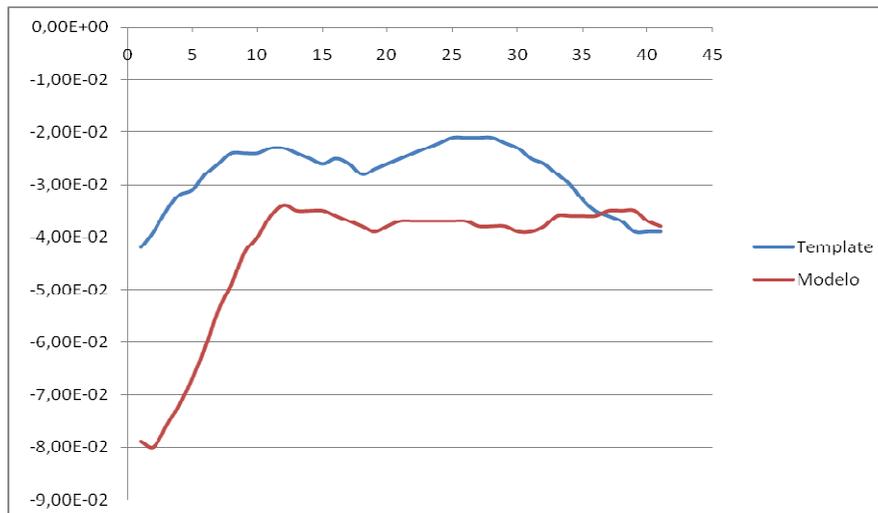


Figura 46: Evaluación de 2M con respecto a 1PMS

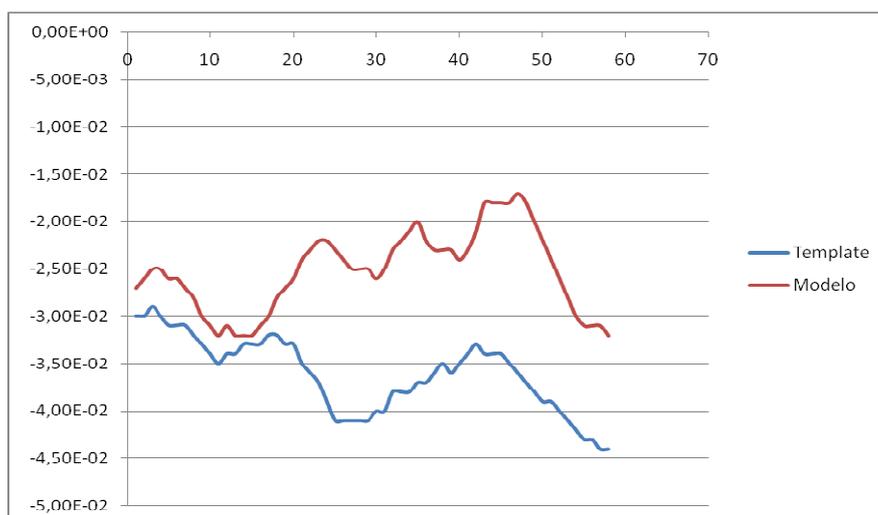


Figura 47: Evaluación de 2M con respecto a 1Y8Q

De estos modelos construidos, no se pudo obtener tendencias equivalentes en los perfiles DOPE para esta predicción.

Después de hacer el modelamiento de las regiones carboxilo terminal de criptocromos de las especies escogidas evidenciamos que estas regiones no presentan una homología total de estructura, se puede afirmar que estas regiones no tienen una estructura total conocida , es decir tienen un plegamiento estructural

no conocido (new folding), además estos fragmentos que pudimos determinar correctamente no pueden establecer la función de dichas regiones en los criptocromos a ciencia cierta por corresponder a fragmentos muy pequeños. Además se evidenció la diversidad de elementos estructurales presentes en estas regiones carboxilo terminal, en donde no muestran un patrón definido de estructura secundaria (Anexo 2).

A partir de los resultados de homología de HCA se da un punto de inicio para poder llegar a definir la organización estructural a nivel terciario de estas regiones que hasta el momento que consideran carentes de orden estructural⁷.

Otros métodos teóricos pueden ser usados para la construcción de modelos de estructura terciaria para estas regiones carboxilo terminal, pero al no tener estructuras conocidas para ser usadas como estructuras de referencia, dichos modelos se deben tomar como reales, pero dichos modelos no nos garantizan que en realidad lo sean²¹.

CONCLUSIONES

- Entre las secuencias de criptocromos que analizamos encontramos tres excepciones en la longitud de la secuencia. Dos en insectos, en donde la tendencia esta en que la región carboxilo terminal de los criptocromos posea entre 26 y 65 residuos de aminoácidos y hallamos dos secuencias con 270 y 473 aminoácidos (criptocromo 2 de *Antheraea pernyi* y criptocromo 2 de *Anopheles gambiae* respectivamente) y en plantas, una región carboxilo terminal con extensión de 520 aminoácidos (criptocromo de *Chlamydomonas reinhardtii*). Esto nos hace suponer algunos tipos de funciones diferentes a los demás criptocromos, ya que estas secuencias no mostraron homología en estructura secundaria con otros criptocromos analizados.
- Las regiones carboxilo terminal de los criptocromos son una diversidad en longitud de aminoácidos y en estructura secundaria mostrando 14 tipos diferentes de ordenamiento en este nivel estructural hallados por HCA (Tabla 44, anexo 2).
- Entre plantas y vertebrados, estas regiones carboxilo terminal de los criptocromos no presentan relaciones estructurales estrechas (con valor de homología HCA mayor o igual que 75%).
- En insectos, las secuencias carboxilo terminal de criptocromo 2 de *Apis mellifera* y el criptocromo 2 de *Tribolium castaneum* están más relacionadas con las secuencias de los criptocromos de vertebrados (Tabla 42).
- HCA reveló un tipo de estructura secundaria para la región carboxilo terminal del criptocromo de *Drosophila melanogaster* diferente a lo encontrado por otros autores⁷, ya que en este trabajo detectamos un

elemento de estructura secundaria tipo β que difiere con la α -hélice encontrada en el trabajo ya mencionado. Esto muestra la capacidad de la técnica para encontrar por inspección de los agrupamientos hidrofóbicos elementos de estructura secundaria de manera diferentes a los métodos que se basan en comparaciones con secuencias que ya han sido determinadas experimentalmente.

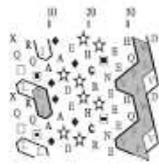
- La predicción de las estructuras terciarias de las regiones carboxilo terminal de los criptocromos escogidos como representativos no se pudieron descifrar totalmente salvo unos cortos fragmentos (entre 20 - 90 residuos de aminoácidos) de la totalidad de las proteínas sometidas a predicción, y al ser comparados dichos fragmentos predichos con las estructuras tridimensionales usadas como modelos, obtenidos de la página de la PDB (protein data bank), mostraba que las regiones carboxilo terminal de los criptocromos no estaban directamente involucradas con una función específica conocida.

ANEXOS

ANEXO 2

Perfiles de HCA de los conjuntos de secuencias detectados por homología HCA (tablas 9, 23, 41)

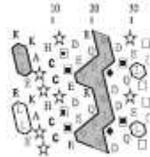
Conjunto 1 de insectos



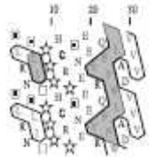
Bactrocera tryoni



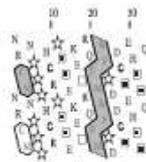
Antheraea pernyi



Danaus plexippus



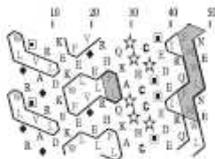
Drosophila melanogaster



Dianemobius nigrofasciatus

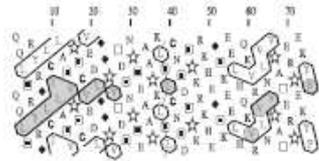


Criptocromo 1 Anopheles gambiae

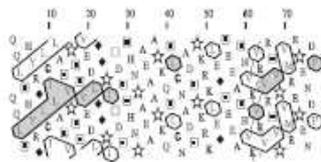


Mamestra brassicae

Conjunto 1 en vertebrados

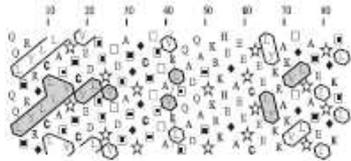


Criptocromo 2a Xenopus laevis

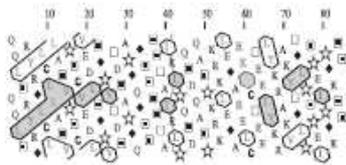


Criptocromo 2b Xenopus laevis

Conjunto 2 en vertebrados



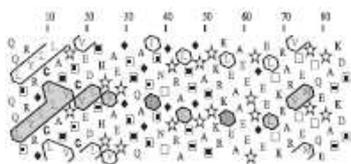
Criptocromo 2 Sylvia borin



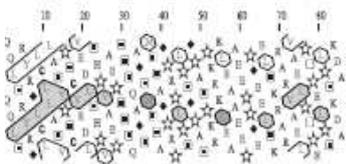
Criptocromo 2 Meleagris gallopavo



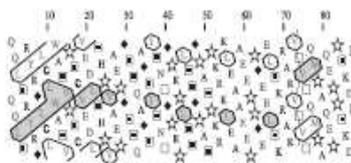
Criptocromo 2 Gallus gallus



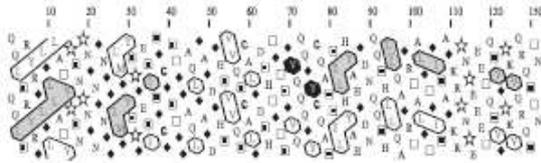
Criptocromo 2 Mus musculus



Criptocromo 2 Homo sapiens



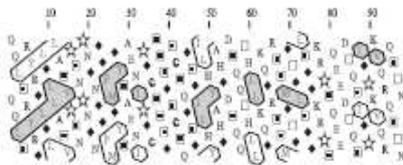
Criptocromo 2 Rattus norvegicus



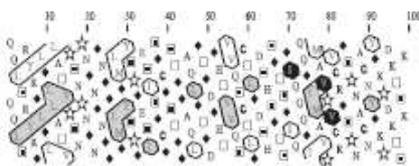
criptocromo 1 Sylvia borin



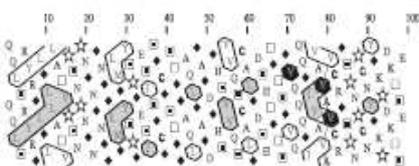
criptocromo 1 Macaca fascicularis



criptocromo 1 Homo sapiens



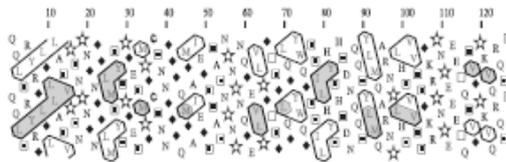
criptocromo-1b Erithacus rubecula



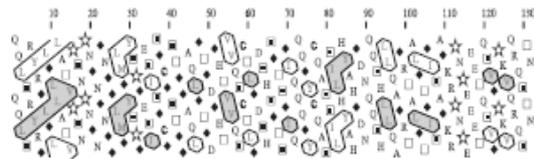
criptocromo 1b Sylvia borin



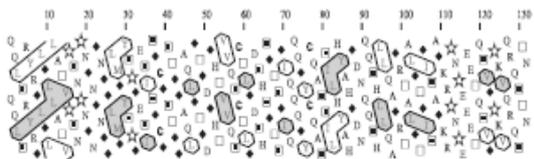
Criptocromo-1. Spalax judaei



criptocromo 1 Xenopus laevis

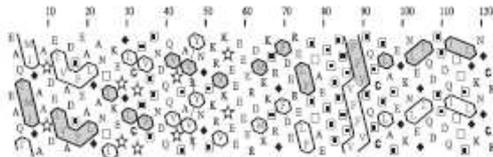


criptocromo-1a Erithacus rubecula

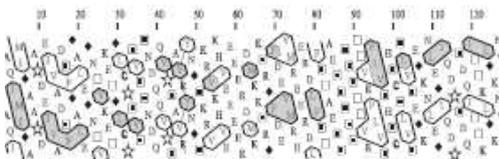


Criptocromo 1 Erithacus rubecula

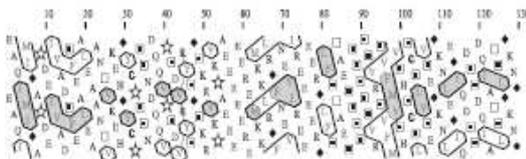
Conjunto 1 en plantas



Criptocromo 2 A.thaliana

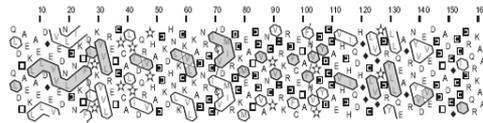


Criptocromo 2 A.rusticana

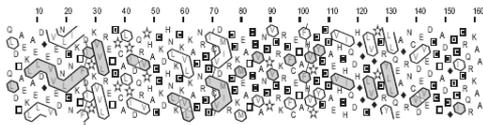


Criptocromo 2 B.napus

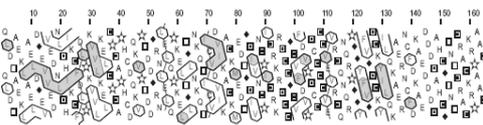
Conjunto 2 en plantas



CRY2 O. Sativa japonica

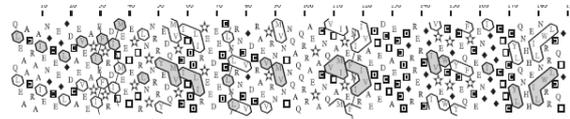


CRY2 O. sativa indica

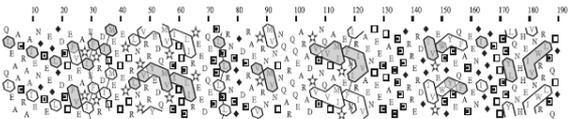


CRY2 H. Vulgare sub. Vulg.

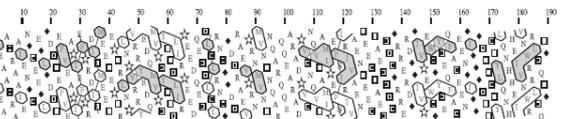
Conjunto 3 en plantas



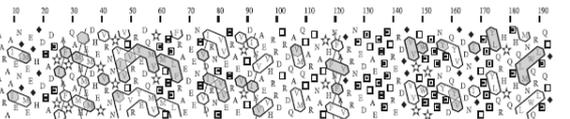
CRY1 A. thaliana



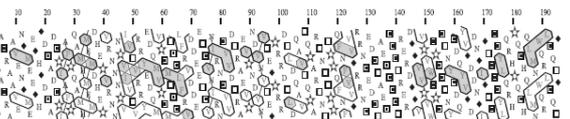
CRY1 B. napus



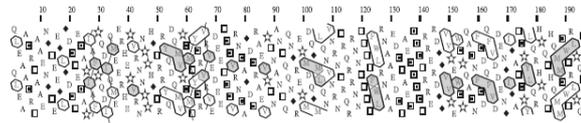
CRY1 B. rapa



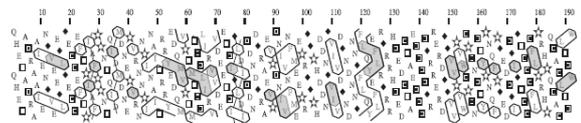
CRY1 L. esculentum



CRY1 N. sylvestris

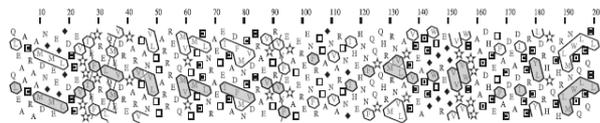


CRY1 *P.sativum*

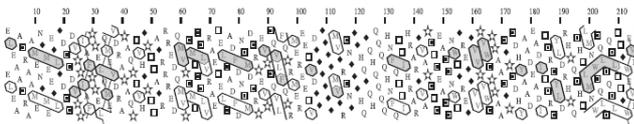


CRY1 *O.minor*

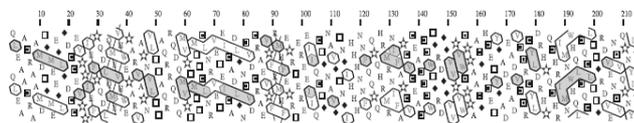
Conjunto 4 en plantas



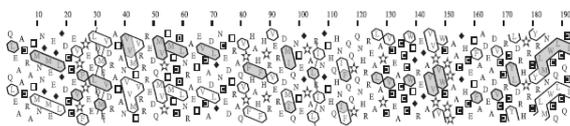
CRY1b *O.sativa-japonica*



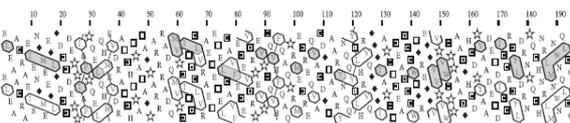
CRY1a *O.sativa-japonica*



CRY1b *H.vulgare.sub.vul*



CRY2 *S.bicolor*

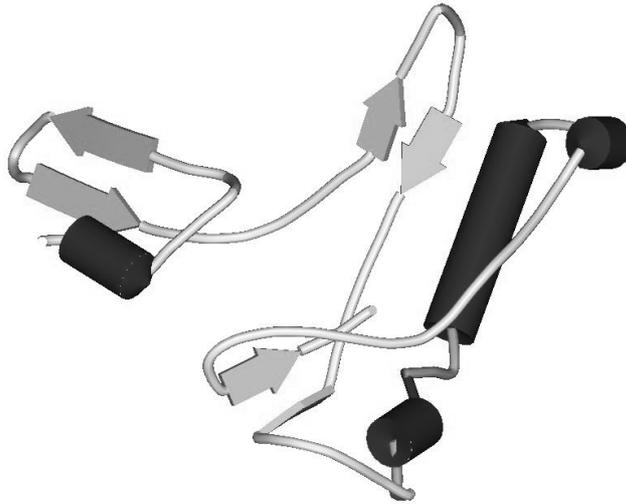


CRY1a *H.vulgare.sub.vul*

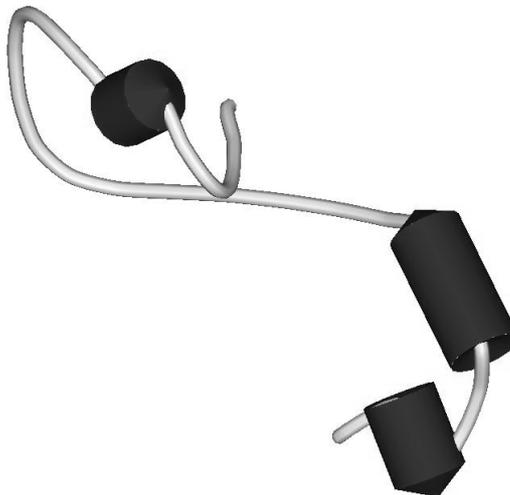
ANEXO 3

Modelos obtenidos correctamente de ciertos fragmentos de las regiones carboxilo terminal

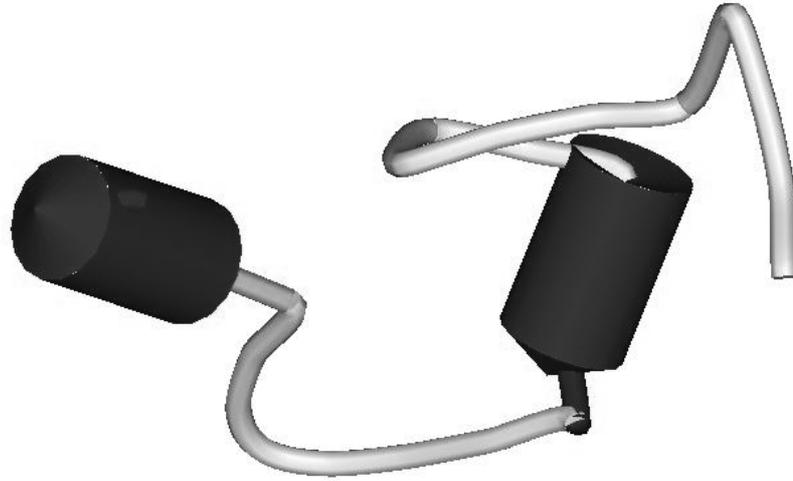
1. 1GP: fragmento entre 1 hasta 96 aminoácidos



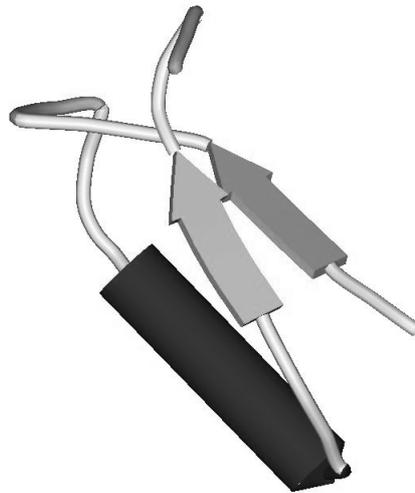
2. 1M: fragmento entre 98 hasta 133 aminoácidos



3. 2EP: Fragmente entre 49 hasta 76 aminoácidos



Fragmento entre 108 y 148 aminoácidos



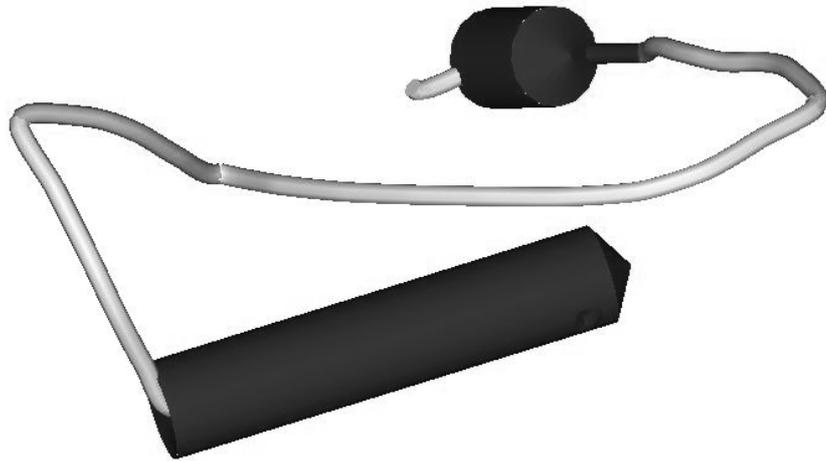
4. 3BV: Fragmento entre 11 hasta 35



5. 5CV: fragmento entre 42 hasta 85 aminoácidos



6. 10I: fragmento desde 38 hasta 79 aminoácidos



7. 12I: fragmento entre 243 hasta 270 aminoácidos



BIBLIOGRAFÍA

- ¹ Brudler, R., Hitomi, K., Daiyasu, H., Toh, H., Kucho, K., Ishiura, M., Kanehisa, M., Roberts, V. A., Todo, T., Tainer, J. A. and Getzoff, E. D. (2003). *Mol. Cell* 11, 59–67
- ² Todo, T. (1999). *Mutat. Res.* 434, 89-97.
- ³ Cashmore, A.R., Jarillo, J.A., Wu, Y.J., and Liu, D. (1999). *Science* 284, 760-765
- ⁴ Essen, L. (2006). *Curr. Opin. In Plant Biol.* 16:51-59.
- ⁵ Brautigam, C.A., Smith, B. S., Ma, Z., Palnitkar, M., Tomchick, D.R., Machius, M. and Deisenhofer, J. 2004. *Proc. Natl. Acad. Sci. USA.* 101, 12142-12147
- ⁶ Huang Y., Baxter R., Smith B., Partch C., Colbert C. and Deisenhofer J. 2006. *Proc. Natl. Acad. Sci. USA.* 103,17701–17706.
- ⁷ Partch, C., Clarkson, M., Özgür, S., Lee, A., Sancar, A. (2005). *Biochemistry* 44, 3795-3805.
- ⁸ Park, H., Kim, S., Sancar, A., (1995) *Science* 268, 1866-1872.
- ⁹ Green, C. (2004). *Curr. Biol.* 14, R847-R849.
- ¹⁰ Ahmad, M., Cashmore A. (1996) *Plant Mol. Biol.* 30, 851–861.
- ¹¹ Deisenhofer, J., (2000). *Mut. Res.* 460, 143-149.
- ¹² Kleine, T., Lockhart, P. & Batschauer, A. (2003). *Plant J.* 35, 93–103.
- ¹³ Daiyasu, H., Ishikawa T., Kuma, K., Iwai S., Todo T. and Toh H. (2004) *Genes to Cells* 9, 479–495
- ¹⁴ Sancar, A., *Chem. Rev.* 2003, 103, 2203–2237
- ¹⁵ Gyula, P., Schäfer, E., Nagy, F. (2003) *Curr. Opin. In plant biol.* 6: 446-452.
- ¹⁶ <http://www.rcsb.org/pdb/home/home.do>
- ¹⁷ Malhotra, K., Kim, S.-T., Batschauer, A., Dawut, L. & Sancar, A. (1995). *Biochemistry* 34, 6892–6899.
- ¹⁸ Rost B., Yachdav G. and Liu J. (2004). *Nucleic Acids Research.* 32, Web Server issue W321–W326
- ¹⁹ LESK, Arthur. *Introduction to bioinformatics.* Second Edition. New York: Oxford University Press, 2005, p.39-50.
- ²⁰ BOURNE, Philip and WEISSIG Helge. *Structural Bioinformatics.* New Jersey: John Wiley & Sons, 2003, p. 15-39
- ²¹ BOURNE, Philip and WEISSIG Helge. *Structural Bioinformatics.* New Jersey: John Wiley & Sons, 2003, p. 501.
- ²² <http://www.ncbi.nlm.nih.gov>
- ²³ R.D. Finn, J. Mistry, B. Schuster-Böckler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S.R. Eddy, E.L.L. Sonnhammer and A. Bateman *Nucleic Acids Research* (2006) .Database Issue 34:D247-D251
- ²⁴ Gaboriaud C., Bissery V., Benchetrit T. and Mornon J.P. (1987) *FEBS Lett.* 224:149–155.
- ²⁵ Thompson JD, Higgins D and Gibson TJ. (1994). *Nucl. Acid Res.* 22,4673-4680.

-
- ²⁶ SILVA, P. (2007). Proteins: Structure, function and bioinformatics. In press
- ²⁷ <http://bioserv.impmc.jussieu.fr/hca-seq.html>
- ²⁸ Lemesle-Varloot L., Henrissat B., Gaboriaud C., Bissery V., Morgat A. And Mornon J.P. (1990). *Biochimie* 72: 555–574.
- ²⁹ Altschul, S. F., Thomas L. Madden, A. Schäffer, Jinghui Z., Zheng Z., Webb M. and Lipman D. (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.
- ³⁰ Šali A. and Blundell TL (1993). *J. Mol. Biol.* 234, 779-815.
- ³¹ www.salilab.org/modeller
- ³² Callebaut, I. Labesse, G. Durand, P. Poupon, A. Canard, L. Chomilier, J. Henrissat, B. And Mornon, J. P. (1997) *Cell. mol. life sci.* 53: 621–645
- ³³ Henrissat, B., Saloheimo, M., Lavaitte, S. and Knowles, J. (1990). *PROTEINS: Structure, function and genetics.* 8, 251-257.
- ³⁴ <http://www.accelrys.com/products/dstudio/>
- ³⁵ A. Šali, R. S_anchez, A. Y. Badretdinov, A. Fiser, F. Melo, J. P. Overington, E. Feyfant, and M. A. Mart__-Renom. *Modeller, A Protein Structure Modeling Program, Release 6.* URL www.salilab.org/Modeller/manual