



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS QUÍMICAS

**“DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS
DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”**

TESIS

PARA OPTAR POR EL GRADO DE

MAESTRO EN CIENCIAS

PRESENTA

Q.F.B. NORBERTO SÁNCHEZ CRUZ

TUTOR: DR. RAMÓN GARDUÑO JUÁREZ
INSTITUTO DE CIENCIAS FÍSICAS, UNAM

CIUDAD UNIVERSITARIA, CD. MX., ENERO DE 2017



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo económico brindado durante mis estudios de maestría. Número de becario: 335997.

A la Universidad Nacional Autónoma de México (UNAM) por la educación que me ha brindado a lo largo de tantos años.

Al Instituto de Ciencias Físicas (ICF) por las facilidades otorgadas para la realización de este trabajo de Tesis.

Al Dr. Ramón Garduño Juárez por brindarme su asesoría y apoyo, tanto académico como personal, a lo largo de este tiempo.

Al jurado evaluador por sus valiosas opiniones acerca del presente trabajo.

A mis amigos, los de toda la vida y los que conocí en este camino, por dejarme compartir con ellos esta grata experiencia.

A mi familia, por ser la piedra angular de mi formación.

*A mis padres David y Magdalena
Todo lo que soy es gracias a ustedes*

JURADO ASIGNADO:

Presidente: Dr. Miguel Antonio Costas Basín

Vocal: Dr. Andrés Hernández Arana

Vocal: Dr. Enrique García Hernández

Vocal: Dra. Alejandra Hernández Santoyo

Secretario: Dra. Karina Martínez Mayorga

Lugar donde se realizó la tesis: Instituto de Ciencias Físicas, UNAM.

El contenido de este trabajo fue presentado de manera parcial, en la modalidad de poster, en los siguientes eventos:

- XIV Reunión Mexicana de Físicoquímica Teórica. Desarrollo de un potencial estadístico basado en la geometría del puente de hidrógeno para la identificación de la estructura nativa de proteínas. Jalisco, México: Universidad de Guadalajara, Centro Universitario de Tonalá, 2015.
- Congreso de la Rama de Físicoquímica, Estructura y Diseño de Proteínas de la SMB and 4th International Workshop Frontiers in Protein Folding, Evolution and Function. Development of a statistical potential based on the hydrogen bond geometry for identification of native structures in proteins. Oaxaca, México: Sociedad Mexicana de Bioquímica, 2015.

ÍNDICE

ÍNDICE DE FIGURAS	7
ÍNDICE DE TABLAS	8
ABREVIATURAS	9
I. INTRODUCCIÓN	10
II. ANTECEDENTES	12
1. LA ESTRUCTURA DE LAS PROTEÍNAS	12
1.1 ESTRUCTURA PRIMARIA	13
1.2 ESTRUCTURA SECUNDARIA	16
1.3 ESTRUCTURA TERCIARIA Y CUATERNARIA	21
1.4 FACTORES QUE AFECTAN LA ESTABILIDAD DE UNA PROTEÍNA NATIVA	21
2. PLEGAMIENTO DE PROTEÍNAS	24
2.1 PLEGAMIENTO <i>IN VITRO</i>	24
2.2 PLEGAMIENTO <i>IN VIVO</i>	27
3. DETERMINACIÓN EXPERIMENTAL DE ESTRUCTURAS DE PROTEÍNAS	28
3.1 DIFRACCIÓN DE RAYOS X	29
3.2 MICROSCOPIA ELECTRÓNICA	30
3.3 ESPECTROSCOPIA DE RESONANCIA MAGNÉTICA NUCLEAR	31
4. PREDICCIÓN DE ESTRUCTURA DE PROTEÍNAS	32
4.1 MÉTODOS <i>AB INITIO</i>	34

4.2 MÉTODOS BASADOS EN PLANTILLAS.	34
4.3 POTENCIALES ESTADÍSTICOS	36
III. TRABAJOS RELACIONADOS	40
IV. DESCRIPCIÓN DEL PUENTE DE HIDRÓGENO	42
V. SELECCIÓN DE LA MUESTRA	44
VI. POTENCIALES ESTADÍSTICOS H-ERX Y H-ERMN	52
1. GENERACIÓN DE LOS POTENCIALES	52
2. PRUEBAS DE DISCRIMINACIÓN DE SEÑUELOS	54
VII. POTENCIALES ESTADÍSTICOS H-ERX* Y H-ERMN*	62
1. UN NUEVO ANÁLISIS DE DATOS	62
2. GENERACIÓN DE LOS POTENCIALES	66
3. PRUEBAS DE DISCRIMINACIÓN DE SEÑUELOS	68
VIII. COMPARACIÓN CON OTROS POTENCIALES	71
IX. CONCLUSIONES	73
X. PERSPECTIVAS	75
XI. REFERENCIAS	76

ÍNDICE DE FIGURAS

Figura 1. Niveles de estructura en proteínas.	12
Figura 2. Estructura general de un aminoácido.	14
Figura 3. Formación de un enlace peptídico por condensación.	15
Figura 4. Modelo de bolas y palos de una hélice α.	17
Figura 5. Diagrama de cintas de una hélice α.	17
Figura 6. Esquema de puentes de hidrógeno para las distintas hélices.	18
Figura 7. Conformación β de las cadenas polipeptídicas.	19
Figura 8. Diagrama de cintas de una hoja β con hebras antiparalelas.	20
Figura 9. Estructura del giro β.	20
Figura 10. Ruta de plegado simulada.	25
Figura 11. Termodinámica del proceso de plegado representada como un embudo de energía libre.	26
Figura 12. Versión resumida de un archivo pdb representativo (2CGA).	28
Figura 13. Representación esquemática de los parámetros usados por Kortemme y colaboradores para describir los puentes de hidrógeno.	40
Figura 14. Representación esquemática de los parámetros usados para describir los puentes de hidrógeno.	43
Figura 15. Distribuciones de frecuencias de los parámetros geométricos estudiados en los puentes de hidrógeno formados por la cadena principal de proteínas, en muestras experimentales.	46
Figura 16. Distribuciones de frecuencias de los parámetros geométricos estudiados en los puentes de hidrógeno formados por la cadena principal de proteínas en los grupos ERX + AMBER99SB y ERMN + AMBER99SB.	49
Figura 17. Estructura de las proteínas usadas para generar los conjuntos de señuelos I-TASSER.	54
Figura 18. Contribución por conectividad de los puentes de hidrógeno en las distribuciones de frecuencias de los parámetros geométricos estudiados en los grupos ERX* + AMBER99SB y ERMN + AMBER99SB.	63

ÍNDICE DE TABLAS

Tabla 1. Clasificación, convenciones y polaridad asociada a los veinte aminoácidos comunes.	14
Tabla 2. Puentes de hidrógeno preservados al sustituir los hidrógenos polares en estructuras experimentales, de acuerdo a los parámetros del campo de fuerza AMBER99SB.	48
Tabla 3. Desempeño de los potenciales H-ERX y H-ERMN en las pruebas de discriminación de señuelos en grupos cuya estructura nativa fue resuelta por resonancia magnética nuclear.	59
Tabla 4. Desempeño de los potenciales H-ERX y H-ERMN en las pruebas de discriminación de señuelos en grupos cuya estructura nativa fue resuelta por difracción de rayos X.	60
Tabla 5. Clasificación de los puentes de hidrógeno hallados en los grupos ERX* + AMBER99SB y ERMN + AMBER99SB de acuerdo a la conectividad residuo aceptor (CO)→ residuo donador (NH) que presentan con respecto a su secuencia.	62
Tabla 6. Desempeño de los potenciales H-ERX* y H-ERMN* en las pruebas de discriminación de señuelos en grupos cuya estructura nativa fue resuelta por resonancia magnética nuclear.	68
Tabla 7. Desempeño de los potenciales H-ERX* y H-ERMN* en las pruebas de discriminación de señuelos en grupos cuya estructura nativa fue resuelta por difracción de rayos X.	69
Tabla 8. Desempeño de distintos potenciales en los grupos de señuelos de I-TASSER.	72

ABREVIATURAS

kJ	Kilojoules
°C	Grados Celsius
PDB	Protein Data Bank
Å	Angstroms
RMN	Resonancia Magnética Nuclear
kDa	Kilodaltones
ERX	Estructuras resueltas por difracción de rayos X con una resolución mejor a 1.0 Å
ERMN	Estructuras resueltas por RMN
H-ERX	Potencial estadístico basado en la geometría de los puentes de hidrógeno en ERX
H-ERMN	Potencial estadístico basado en la geometría de los puentes de hidrógeno en ERMN
ERX*	Estructuras resueltas por difracción de rayos X con una resolución mejor a 1.5 Å
H-ERX*	Potencial estadístico basado en la geometría de los puentes de hidrógeno en ERX* que contempla la conectividad aceptor-donador
H-ERMN*	Potencial estadístico basado en la geometría de los puentes de hidrógeno en ERMN que contempla la conectividad aceptor-donador
u.a.	Unidades arbitrarias

I. INTRODUCCIÓN

Las proteínas son el grupo de macromoléculas biológicas que más abunda en cualquier célula y se encuentran en una gran variedad de formas y tamaños, desde péptidos relativamente pequeños hasta polímeros enormes. Además, las proteínas presentan una enorme diversidad en lo que se refiere a su función biológica, ya que actúan como catalizadores, transportan y almacenan otras moléculas como el oxígeno, proporcionan apoyo mecánico y protección inmunológica, generan movimiento, transmiten impulsos nerviosos y controlan el crecimiento y la diferenciación celular¹⁻⁴.

Durante los últimos años, el problema de la predicción de estructura de proteínas ha generado un enorme interés por parte de la comunidad científica, debido en parte al impacto que tendría la resolución de este problema en muchas áreas de la biología, ya que el conocimiento de la estructura terciaria de una proteína es esencial para entender tanto su función biológica como su mecanismo de acción.

El experimento de Anfinsen en la década de los 70, demostró que toda la información que una proteína necesita para plegarse adecuadamente se encuentra codificada en su secuencia de aminoácidos⁵, sugiriendo que, al menos en teoría, es posible predecir la estructura tridimensional de una proteína partiendo únicamente de su secuencia. Este problema ha sido denominado “el santo grial de la biología molecular” y considerado como equivalente a descifrar la segunda mitad del código genético⁶.

La importancia de la resolución computacional de estructuras de proteínas se está incrementando debido al rápido aumento en el número de genomas secuenciados y al crecimiento relativamente lento del número de estructuras de proteínas determinadas experimentalmente. No es de sorprender que la predicción de estructura de proteínas se haya vuelto una parte vital en los proyectos de genómica estructural en todo el mundo⁷.

Se han utilizado una gran variedad de métodos para tratar de resolver este problema. La mayoría de ellos se basa en la hipótesis termodinámica, la cual establece que el estado nativo de una proteína será aquél que posea la menor energía libre bajo condiciones fisiológicas⁵ e implica la necesidad de desarrollar funciones de energía que permitan distinguir la estructura nativa de una proteína dentro de un grupo de señuelos con conformaciones similares, siendo este uno de los problemas más retadores en el campo de la predicción de estructura de proteínas.

Existen dos tipos de funciones de energía: 1) las basadas en principios físicos^{8,9} y 2) los potenciales basados en el conocimiento (PBC)¹⁰⁻¹⁵ o potenciales estadísticos. Los segundos tienen la ventaja de ser más fáciles de calcular y que al ser extraídos a partir de estructuras conocidas, contienen de manera implícita una gran cantidad de información acerca de las interacciones que estabilizan la estructura de las proteínas.

Por otra parte, es bien sabido que las interacciones por puente de hidrógeno son abundantes en las proteínas y que son responsables de muchos fenómenos, como la estabilización de su estructura y el reconocimiento molecular^{16,17}, en el primer caso se trata de puentes de hidrógeno intramoleculares, siendo particularmente abundantes aquellos formados por la cadena principal de las proteínas. Por esta razón, el presente trabajo persigue dos objetivos: (1) construir potenciales estadísticos basados en la geometría de los puentes de hidrógeno intramoleculares de varias muestras de proteínas con estructura conocida y (2) evaluar los potenciales generados de acuerdo a su capacidad para identificar la estructura nativa de una proteína dentro de un conjunto de señuelos.

A continuación se presenta una breve revisión acerca de la estructura y el plegamiento de las proteínas, los métodos empleados para la determinación y predicción de su estructura y el papel de los potenciales estadísticos en estas tareas.

II. ANTECEDENTES

1. La estructura de las proteínas

Para entender y describir la estructura de moléculas tan grandes como las proteínas es necesario estudiarlas en distintos niveles de complejidad, ordenados en una especie de jerarquía conceptual. Comúnmente se definen cuatro niveles de organización estructural en las proteínas (**Fig. 1**):

- **Estructura primaria:** Es una descripción de todos los enlaces covalentes que unen a los residuos de aminoácidos en una cadena polipeptídica.
- **Estructura secundaria:** Se refiere a arreglos particularmente estables de aminoácidos que dan origen a patrones estructurales recurrentes.
- **Estructura terciaria:** Describe todos los aspectos del plegado tridimensional de un polipéptido.
- **Estructura cuaternaria:** Cuando una proteína está constituida por dos o más polipéptidos (llamados subunidades), es la descripción del arreglo espacial de estos.

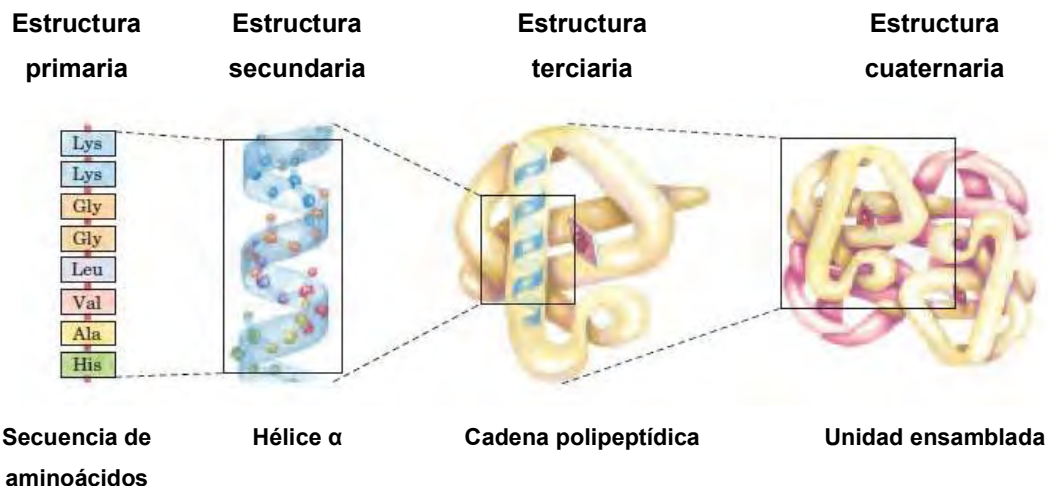


Figura 1. Niveles de estructura en proteínas. La estructura primaria consiste en la secuencia de aminoácidos unidos por enlaces peptídicos. El polipéptido resultante puede ser enrollado en unidades de estructura secundaria, como la hélice α ; un polipéptido plegado consta de varias de estas unidades, dando lugar a lo que se conoce como estructura terciaria, que en ocasiones puede ser una de las subunidades que forman la estructura cuaternaria de una proteína, en este caso la hemoglobina¹.

El arreglo espacial de los residuos de aminoácido en una proteína es denominado conformación. Las posibles conformaciones de una proteína incluyen cualquier arreglo estructural que pueda ser alcanzado sin la ruptura de enlaces covalentes, por ejemplo, con la rotación de un enlace sencillo.

De las numerosas conformaciones teóricamente posibles en una proteína que contiene cientos de enlaces sencillos, existe una o unas cuantas que predominan bajo condiciones biológicas. Las conformaciones predominantes bajo un conjunto de condiciones son usualmente aquellas que son las más estables en términos termodinámicos, es decir, con la energía libre de Gibbs más baja. Las proteínas en cualquiera de estas conformaciones plegadas y funcionales se denominan proteínas nativas.

1.1 Estructura primaria

Las proteínas son polímeros de aminoácidos, con cada residuo de aminoácido unido a su vecino por un tipo específico de enlace covalente (el término residuo se refiere a la pérdida de los elementos que constituyen una molécula de agua cuando un aminoácido es unido a otro). Veinte aminoácidos diferentes son los que se encuentran comúnmente en proteínas: alanina, valina, leucina, isoleucina, glicina, prolina, cisteína, metionina, histidina, fenilalanina, tirosina, triptófano, asparagina, glutamina, serina, treonina, lisina, arginina, aspartato (ácido aspártico) y glutamato (ácido glutámico); siendo todos ellos estereoisómeros L.

Estos 20 aminoácidos comunes son α -aminoácidos, ya que poseen un grupo carboxilo y un grupo amino enlazados al mismo átomo de carbono (el carbono α) (**Fig. 2**). Difieren entre sí por sus cadenas laterales o grupos R, los cuales varían en estructura, tamaño y carga eléctrica, por lo que se pueden clasificar en grupos basados en su comportamiento ácido-base al ser disueltos en agua, o de acuerdo a su solubilidad en ella (**Tabla 1**).

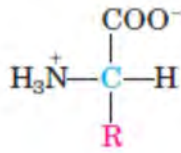


Figura 2. Estructura general de un aminoácido. Esta estructura es común para todos los aminoácidos excepto la prolina, que es un aminoácido cíclico. El grupo R o cadena lateral (rojo) unido al carbono α (azul) es diferente en cada aminoácido¹.

Tabla 1. Clasificación, convenciones y polaridad asociada a los veinte aminoácidos comunes¹.

Aminoácido	Abreviación	Símbolo	Índice hidropático*
Grupos R alifáticos, no polares			
Glicina	Gly	G	-0.4
Alanina	Ala	A	1.8
Prolina	Pro	P	1.6
Valina	Val	V	4.2
Leucina	Leu	L	3.8
Isoleucina	Ile	I	4.5
Metionina	Met	M	1.9
Grupos R aromáticos			
Fenilalanina	Phe	F	2.8
Tirosina	Tyr	Y	-1.3
Triptófano	Trp	W	-0.9
Grupos R no cargados, polares			
Serina	Ser	S	-0.8
Treonina	Thr	T	-0.7
Cisteína	Cys	C	2.5
Asparagina	Asn	N	-3.5
Glutamina	Gln	Q	-3.5
Grupos R cargados positivamente			
Lisina	Lys	K	-3.9
Histidina	His	H	-3.2
Arginina	Arg	R	-4.5
Grupos R cargados negativamente			
Aspartato	Asp	D	-3.5
Glutamato	Glu	E	-3.5

*Escala que combina la hidrofobicidad y la hidrofiliidad de las cadenas laterales, se puede emplear para medir la tendencia de un aminoácido por buscar un ambiente acuoso (valores negativos) o un ambiente hidrofóbico (valores positivos).

Además de estos 20 aminoácidos existen algunos otros no tan comunes, algunos de ellos son residuos modificados después de que una proteína es sintetizada; otros son aminoácidos presentes en los organismos vivos pero no como constituyentes de las proteínas¹⁸⁻²⁴.

Los péptidos son cadenas de aminoácidos. Dos moléculas de aminoácidos pueden unirse covalentemente a través de una reacción de condensación entre el grupo amino de un aminoácido y el grupo carboxilo del otro, liberando una molécula de agua y dando origen a un dipéptido. Este tipo de uniones entre aminoácidos se conoce como enlace peptídico (**Fig. 3**). Tres aminoácidos pueden unirse mediante dos enlaces peptídicos para formar un tripéptido; de manera similar se pueden formar tetrapéptidos, pentapéptidos y demás. Cuando diez o menos aminoácidos se encuentran unidos de esta manera, la estructura es denominada oligopéptido y cuando son más de diez aminoácidos los involucrados, el producto se denomina polipéptido. Las proteínas pueden contener cientos de residuos de aminoácidos.

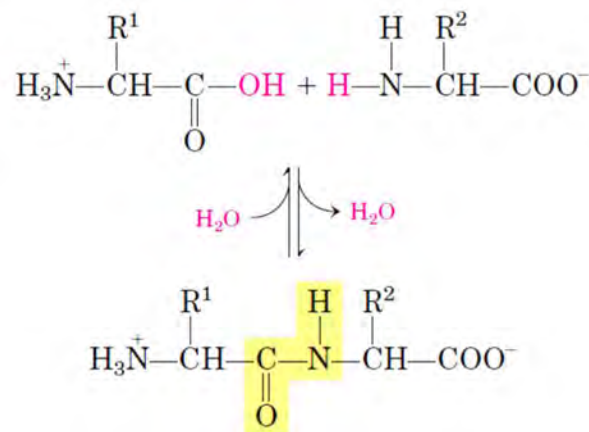


Figura 3. Formación de un enlace peptídico por condensación. El grupo amino de un aminoácido (con cadena lateral R²) actúa como nucleófilo para desplazar el grupo hidroxilo de otro aminoácido (con cadena lateral R¹), formando un enlace peptídico (sombreado en amarillo). Los grupos amino son buenos nucleófilos, pero el grupo hidroxilo no es un buen grupo saliente, por lo que en condiciones fisiológicas esta reacción no ocurre de manera significativa¹.

1.2 Estructura secundaria

La estructura secundaria se refiere a la relación espacial que existe entre aminoácidos que se encuentran cercanos en la estructura primaria, normalmente refiriéndose a los patrones de plegamiento que adopta el esqueleto de una cadena polipeptídica. Las unidades básicas de la estructura secundaria son: las hélices α , las hebras β y los giros; el resto de las unidades de estructura secundaria conocida son variaciones de estos tres tipos. Todas las unidades de estructura secundaria son generadas por la formación de puentes de hidrógeno entre los distintos grupos carboxilo y amino de distintos residuos de aminoácidos.

1.2.1 Hélice α

El arreglo más simple que una cadena polipeptídica puede tomar con sus enlaces peptídicos rígidos es una estructura helicoidal, la cual Pauling y Corey²⁵ denominaron hélice α . En esta estructura, el esqueleto de la cadena polipeptídica se encuentra firmemente enrollada alrededor de un eje imaginario dibujado longitudinalmente a través de la hélice mientras que los grupos R de los residuos de aminoácidos sobresalen hacia la parte exterior de la hélice (**Fig. 4**). En diagramas esquemáticos de proteínas, las hélices α se representan como cintas torcidas o como cilindros (**Fig. 5**). Alrededor de una cuarta parte de los residuos en una cadena polipeptídica se encuentran en esta conformación, aunque la fracción exacta varía de una proteína a otra.

Esta estructura se encuentra estabilizada por puentes de hidrógeno formados entre el átomo de hidrógeno unido covalentemente al átomo de nitrógeno en un enlace peptídico y el átomo de oxígeno carbonílico del cuarto aminoácido del lado amino-terminal del enlace peptídico (**Fig. 6a**), de esta manera prácticamente todos los enlaces peptídicos participan en la formación de puentes de hidrógeno y todas estas interacciones combinadas le dan a la estructura helicoidal una estabilidad considerable.

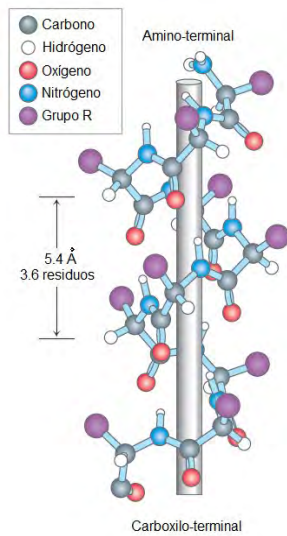


Figura 4. Modelo de bolas y palos de una hélice α . La unidad repetida de este tipo de hélices es de 3.6 residuos¹.

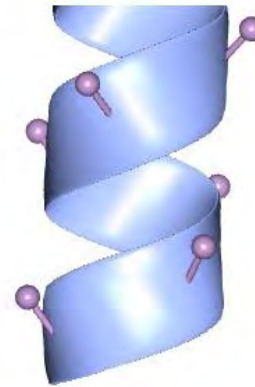


Figura 5. Diagrama de cintas de una hélice α . Los grupos R (morado) sobresalen alrededor del eje de la cinta.

La hélice 3_{10} ²⁶ es una variación estructural de la hélice α encontrada en proteínas. Usualmente se encuentra cuando una hélice α es distorsionada por la presencia de residuos poco favorables, cerca de una región de giro o cuando secuencias cortas se pliegan en una conformación helicoidal. En la hélice 3_{10} los puentes de hidrógeno dominantes se forman entre el átomo de hidrógeno unido covalentemente al átomo de nitrógeno en un enlace peptídico y el átomo de oxígeno carbonílico del tercer aminoácido del lado amino-terminal del enlace peptídico (**Fig. 6b**).

Una tercera posibilidad de hélice es una que se encuentre enrollada de manera más libre que la hélice α , con puentes de hidrógeno formados entre el átomo de hidrógeno unido covalentemente al átomo de nitrógeno en un enlace peptídico y el átomo de oxígeno carbonílico del quinto aminoácido del lado amino-terminal del enlace peptídico (**Fig. 6c**). Esta estructura es la hélice π ²⁷, que en un principio fue considerada como no ocuriente de manera natural, sin embargo en la actualidad se ha descrito esta estructura en algunas proteínas.

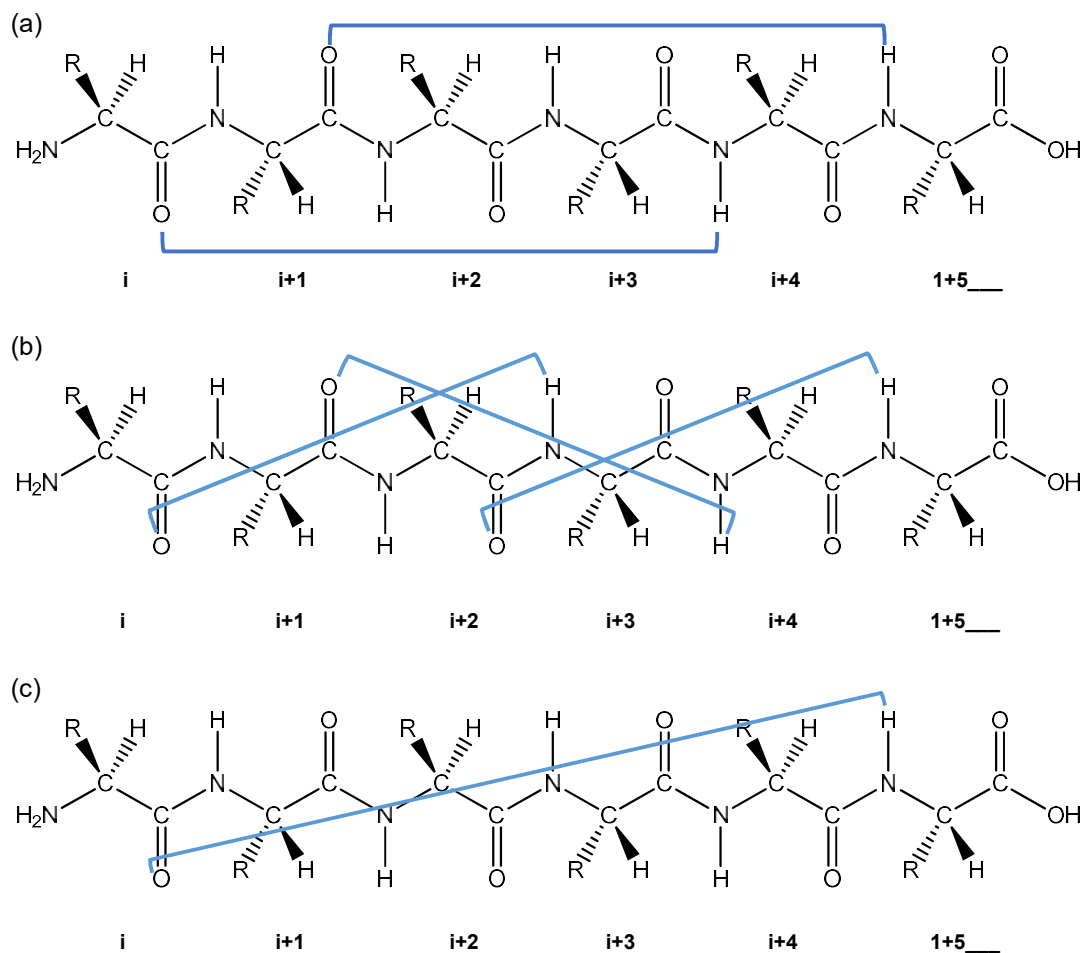


Figura 6. Esquema de puentes de hidrógeno para las distintas hélices. El puente de hidrógeno entre los grupos CO y NH se da de la siguiente manera $i \rightarrow i+4$ para la hélice α (a), $i \rightarrow i+3$ para la hélice 3_{10} (b) e $i \rightarrow i+5$ para la hélice π (c).

1.2.2 Hoja β

Pauling y Corey²⁵ predijeron un segundo tipo de estructura repetitiva, la conformación β o hebra β , esta es una conformación más extendida de las cadenas polipeptídicas. En esta conformación, el esqueleto de una cadena polipeptídica se encuentra extendido en forma de zigzag más que en una forma helicoidal. Las cadenas polipeptídicas en zigzag pueden agruparse lado a lado para formar una estructura conocida como hoja β , en ella se forman puentes de hidrogeno entre segmentos adyacentes de una cadena polipeptídica. Los segmentos individuales que forman una hoja β se encuentran usualmente cerca en una cadena polipeptídica, pero también pueden encontrarse bastante distantes una de otra en

la estructura primaria de un polipéptido e incluso pueden ser segmentos de cadenas polipeptídicas distintas.

Los grupos R de los aminoácidos adyacentes sobresalen de la estructura zigzag, creando un patrón alternado al ver esta estructura de lado. Las cadenas polipeptídicas adyacentes en una hoja β pueden ser paralelas o antiparalelas, teniendo la misma orientación amino-carboxilo o la orientación opuesta, respectivamente (**Fig. 7**).

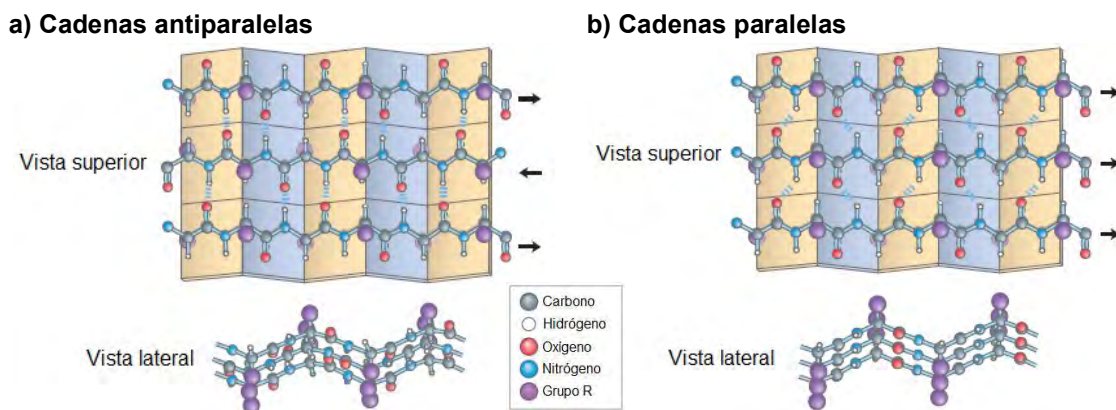


Figura 7. Conformación β de las cadenas polipeptídicas. La forma de las hojas β está definida por la planaridad de los enlaces peptídicos, los grupos R sobresalen de la hoja. Se muestran también los puentes de hidrógeno formados de acuerdo al tipo de cadenas presentes en la hoja: a) antiparalelas y b) paralelas¹.

En diagramas esquemáticos de proteínas, las hebras β se representan normalmente por flechas anchas apuntando hacia el extremo carboxilo-terminal para indicar el tipo de hojas β formadas (**Fig. 8**). Las estructuras son un tanto similares aunque el periodo de repetición es menor para la conformación paralela y los patrones para la conformación de puentes de hidrogeno son diferentes.

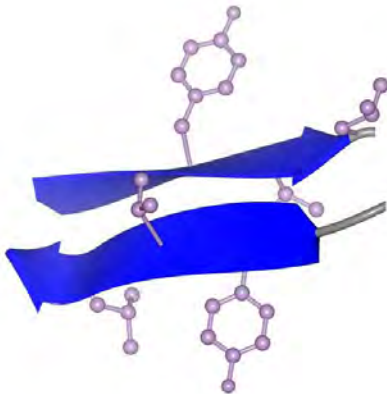
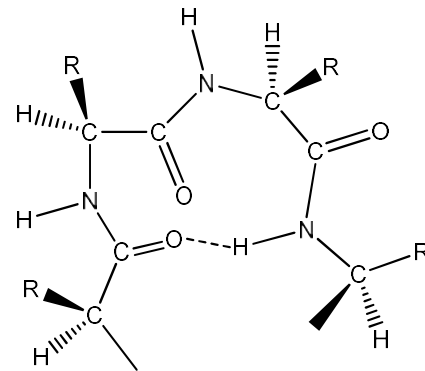


Figura 8. Diagrama de cintas de una hoja β con hebras antiparalelas. Los grupos R (morado) sobresalen de las cintas (4UBY).



Tipo I

Figura 9. Estructura del giro β . El más frecuente de los giros β es el tipo I. Se forma un puente de hidrógeno entre los residuos primero y cuarto del giro.

1.2.3 Giros

En proteínas globulares, que tienen una estructura plegada compacta, cerca de una tercera parte de los residuos de aminoácidos se encuentran en giros o bucles donde la cadena polipeptídica cambia de dirección²⁸. Estos son los elementos que conectan series sucesivas de hélices α o hebras β . Algunos particularmente comunes son los giros β (**Fig. 9**), que conectan dos segmentos adyacentes de una hoja β antiparalela. La estructura es un giro de 180° que involucra cuatro residuos de aminoácidos, con el oxígeno carbonílico del primer residuo formando un puente de hidrógeno con el hidrógeno del grupo amino del cuarto, mientras que los grupos peptídicos de los dos residuos centrales no participan en la formación de ningún puente de hidrógeno. Los residuos de glicina y prolina ocurren de manera frecuente en los giros β , el primero porque es pequeño y flexible y el segundo porque los enlaces peptídicos que involucran el nitrógeno de la prolina adoptan fácilmente la configuración *cis*, una forma que es particularmente favorable a un giro cerrado. Un giro considerablemente menos común es el giro γ , un giro de tres residuos con un puente de hidrógeno entre el primero y el tercero de ellos.

1.3 Estructura Terciaria y Cuaternaria

La estructura terciaria de una proteína se refiere al arreglo tridimensional de todos los átomos en ella. Mientras que el término “estructura secundaria” se refiere al acomodo espacial de residuos de aminoácidos adyacentes en la estructura primaria, la estructura terciaria incluye aspectos de mayor distancia en la secuencia de aminoácidos. Aminoácidos que se encuentran ampliamente separados en la secuencia de un polipéptido y que pertenecen a distintos tipos de estructura secundaria pueden interactuar dentro de la estructura completamente plegada de una proteína. La ubicación de giros en la cadena polipeptídica y la dirección de estos están determinados por el número y ubicación de residuos específicos, segmentos interactuantes de una cadena polipeptídica son mantenidos en su posición terciaria característica por diferentes tipos de interacciones débiles y a veces por enlaces covalentes como un puente disulfuro.

Algunas proteínas contienen dos o más cadenas polipeptídicas separadas o subunidades, las cuales pueden ser idénticas o diferentes, el ordenamiento de estas subunidades en complejos tridimensionales constituyen la estructura cuaternaria.

1.4 Factores que afectan la estabilidad de una proteína nativa

Las proteínas nativas son marginalmente estables, ya que la energía libre de Gibbs que separa los estados plegado y desplegado es de entre 20 y 65 kJ/mol bajo condiciones fisiológicas. Una cadena polipeptídica dada puede adoptar múltiples conformaciones distintas, por lo que el estado desplegado de una proteína se caracteriza por un alto grado de entropía conformacional. Esta entropía y las interacciones por puente de hidrógeno que se presentan entre múltiples grupos de la cadena polipeptídica y las moléculas de agua del entorno tienden a mantener el estado desplegado. Las interacciones químicas que contrarrestan estos efectos y estabilizan la conformación nativa incluye a los puentes disulfuro y las interacciones no covalentes: puentes de hidrógeno intramoleculares, interacciones hidrofóbicas e interacciones coulómbicas. A continuación se describen brevemente cada una de estas interacciones^{1,3,16,29}.

- **Puentes disulfuro:** Los puentes disulfuro dictan el plegado de algunas proteínas por formación de enlaces covalentes entre los átomos de azufre de las cadenas laterales de cisteínas separadas en la secuencia primaria. Un puente de este estilo no puede ser formado entre residuos consecutivos, comúnmente los residuos que participan en la formación de este tipo de enlaces se encuentran separados por al menos 5 residuos. Estos enlaces pueden romperse a alta temperatura, pH ácido o en presencia de agentes reductores.
- **Efecto hidrofóbico:** Las sustancias no polares tienden a formar agregados al encontrarse en disolución acuosa, excluyendo de esta forma a las moléculas de agua de su interior y permitiéndoles formar puentes de hidrógeno entre sí. Como la mayoría de las proteínas se encuentran en un ambiente de este estilo, ocurre la formación de aglomerados de aminoácidos con cadena lateral no polar, entre los cuales el agua queda excluida, dejando en claro que este efecto contribuye significativamente al total de las interacciones intramoleculares presentes en una proteína.
- **Interacciones coulombicas:** Este tipo de interacciones se presentan entre partículas cargadas eléctricamente y son descritas en base a la ley de Coulomb; en el caso de las proteínas, participan en este tipo de interacciones los aminoácidos con cadenas laterales cargadas, así como los grupos amino (NH_3^+) y carboxilo (COO^-) terminales de las cadenas polipeptídicas. Como resultado del efecto hidrofóbico, los aminoácidos con cadenas laterales cargadas se encuentran normalmente en la superficie de las proteínas y su interacción con aminoácidos vecinos se ve debilitada por la presencia de moléculas de agua que provocan un efecto de apantallamiento.

- **Interacciones de van der Waals:** Existen fuerzas de van der Waals tanto atractivas como repulsivas que controlan las interacciones entre átomos sin carga y no enlazados entre sí. Estas fuerzas provienen de la inducción de dipolos en la molécula debido a la fluctuación de las densidades de carga entre los átomos. Las interacciones englobadas en este efecto y ordenadas de mayor a menor fuerza, son: la interacción entre dipolos permanentes, la interacción entre dipolos temporales y las fuerzas de dispersión de London.
- **Puentes de hidrógeno:** Son un tipo de interacción atractiva entre un átomo electronegativo y un átomo de hidrógeno unido covalentemente a otro átomo electronegativo. Este tipo de interacciones contribuyen significativamente a la estabilidad de las hélices α y a la interacción de las hebras β para la formación de hojas β . Como resultado, los puentes de hidrógeno contribuyen significativamente a la estabilidad total de la estructura terciaria de una proteína. La mayor parte de estas interacciones está dada por puentes de hidrógeno formados entre los grupos NH y CO de distintos enlaces peptídicos en la cadena principal, aunque también pueden darse entre las cadenas laterales de distintos aminoácidos, o entre la cadena principal de la proteína y las cadenas laterales de algunos residuos.

Se requieren alrededor de 200 a 460 kJ/mol para romper un enlace sencillo mientras que las interacciones no covalentes se pueden ver modificadas por entre 4 y 30 kJ/mol. Los enlaces covalentes individuales que contribuyen a mantener la conformación nativa de una proteína, como los puentes disulfuro, son claramente más fuertes que las interacciones no covalentes individuales. Sin embargo, las interacciones no covalentes son numerosas, por lo que, en conjunto, son estas las que contribuyen en mayor medida a la estabilización de la conformación nativa en una proteína. En general, la conformación de una proteína con la menor energía libre de Gibbs será aquella con el mayor número de interacciones no covalentes.

2. Plegamiento de proteínas

La estructura terciaria de una proteína está determinada por su secuencia de aminoácidos. La prueba más importante de esto está dada por experimentos que muestran que la desnaturalización de algunas proteínas es reversible. Un experimento clásico es la desnaturalización y renaturalización de la ribonucleasa, llevado a cabo por Anfinsen⁵ en la década de los 70, el cual proporcionó la primera evidencia de que la secuencia de aminoácidos de una cadena polipeptídica contiene toda la información necesaria para plegar dicha cadena en su estructura tridimensional nativa.

En las células, las proteínas son ensambladas a partir de aminoácidos a una velocidad muy alta, por ejemplo, células de *E. coli* pueden hacer una proteína completa de 100 residuos en aproximadamente cinco segundos a 37°C. Suponiendo que cada uno de estos residuos pueda tomar 10 distintas conformaciones en promedio, tendríamos 10^{100} conformaciones diferentes para el polipéptido. Suponiendo también que la proteína se pliega espontáneamente mediante un proceso aleatorio en el que prueba todas las posibles conformaciones hasta encontrar su estructura nativa y que cada conformación fuera muestreada en el tiempo más corto posible ($\sim 10^{-13}$ s, el tiempo requerido para una vibración molecular), tomaría alrededor de 10^{77} años muestrear todas las conformaciones posibles. Por lo tanto el plegado de las proteínas no puede ser un proceso de prueba y error. Este problema fue planteado por primera vez por Cyrus Levinthal³⁰ en 1968.

2.1 Plegamiento in vitro

La vía de plegamiento de una cadena polipeptídica es complicada y no todos los principios que guían este proceso han sido descubiertos. Sin embargo, diversos estudios han llevado al desarrollo de algunos modelos plausibles. En uno de ellos, el proceso de plegado es visto como jerárquico, en él, estructuras secundarias locales se forman primero; ciertas secuencias de aminoácidos se pliegan rápidamente en hélices α u hojas β , esto es seguido por interacciones de largo

alcance entre elementos de estructura secundaria. El proceso continúa hasta que el polipéptido completo se encuentra plegado (**Fig. 10**).

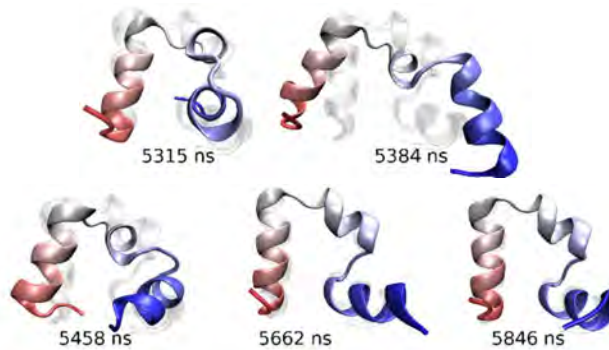


Figura 10. Ruta de plegado simulada. Simulación computacional de la ruta de plegado de un segmento de la villina³¹.

En un modelo alternativo, el plegado es iniciado por un colapso espontáneo de polipéptido en un estado compacto, mediado por interacciones hidrofóbicas entre residuos no polares. El estado resultante de este “colapso hidrofóbico” puede tener un gran contenido de estructura secundaria, pero muchas cadenas laterales de aminoácidos no se encuentran completamente fijadas, este estado es conocido comúnmente como glóbulo fundido. Dentro de este glóbulo fundido se establecen interacciones de largo alcance entre los residuos, permitiendo la estabilización de la estructura terciaria y finalmente la de la estructura secundaria.

La mayoría de las proteínas probablemente se pliegan por un proceso que incorpora elementos de ambos modelos. En lugar de seguir una sola vía, una población de péptidos podría tomar una variedad de rutas hacia el mismo destino (nucleación o condensación), con el número de especies parcialmente plegadas distintas disminuyendo a medida que el plegado alcanza su terminación. Para muchas proteínas, especialmente aquellas con menos de 100 residuos, este modelo se ajusta bien con las evidencias experimentales³²⁻³⁴.

Termodinámicamente el proceso de plegado puede ser visto como una especie de embudo de energía libre (**Fig. 11**). Los estados desplegados están caracterizados por un alto grado de entropía conformacional y energía libre relativamente alta. A medida que el plegado avanza el estrechamiento del embudo representa un decremento en el número de especies conformacionales presentes. Las pequeñas depresiones a lo largo del embudo de energía libre representan intermediarios semiestables que pueden alentar brevemente el proceso de plegado. En el fondo del embudo, un ensamble de intermediarios de plegado ha sido reducido a una única conformación nativa (o a una de un conjunto pequeño de conformaciones nativas)^{35,36}.

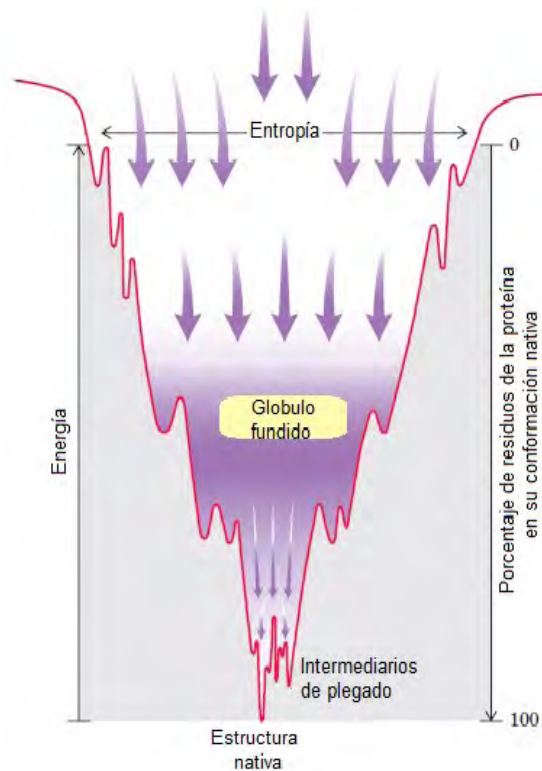


Figura 11. Termodinámica del proceso de plegado representada como un embudo de energía libre. A medida que el proceso de plegado avanza se reduce tanto el número de estados presentes (entropía) como la energía libre e incrementa la proporción de proteína en la conformación nativa.

2.2 Plegamiento in vivo

En una célula, la formación de intermediarios de plegado se dificulta debido a la gran cantidad de moléculas presentes en el medio, lo cual promueve las interacciones entre distintas macromoléculas que conllevarían a la formación de agregados entre ellas. Adicionalmente, el proceso de traducción de una proteína incrementa por sí mismo la probabilidad de un mal plegamiento debido a que cadenas polipeptídicas incompletas no pueden formar los intermediarios estables para alcanzar la conformación nativa de la proteína completa, por lo tanto, sería necesario que se sintetizara toda la proteína, o al menos una subunidad completa, antes de empezar el proceso de plegado. Es en este punto que intervienen las chaperonas moleculares, las cuales son proteínas que interactúan de manera paralela al proceso de traducción de una proteína con la finalidad de inhibir su plegamiento prematuro. Cabe resaltar que este tipo de proteínas no interactúan aportando información al proceso de plegado, simplemente contribuyen a la optimización de dicho proceso. Por otro lado, algunas proteínas poseen bajas eficiencias intrínsecas de plegado y esencialmente no pueden plegarse en ausencia de chaperonas debido a que, sin ayuda de estas, no pueden alcanzar la energía suficiente para modificar una conformación cinéticamente estable^{35,36}.

3. Determinación experimental de estructuras de proteínas

La manera más común de representar a una proteína es mediante el uso de listones planos a pesar de que esta no sea una representación totalmente fiel a su estructura. Esta y todas las formas de representar una proteína parten de archivos que contienen las coordenadas en tres dimensiones de cada átomo en la molécula. Estas coordenadas son obtenidas experimentalmente y almacenadas para su uso en el "Protein Data Bank" con la designación de archivo "pdb" (**Fig. 12**). Estas coordenadas atómicas pueden posteriormente ser representadas de distintas maneras^{3,4}.

```

HEADER      HYDROLASE(ZYMOGEN)                16-JAN-87  2CGA    2CGA  3
COMPND      CHYMOTRYPSINOGEN *A                2CGA    4
SOURCE      BOVINE (BOS STAURUS) PANCREAS      2CGA    5
AUTHOR      D.WANG,W.BODE,R.HUBER             2CGA    6
>>>>>>
>>>>>>
JRNL        AUTH  D.WANG,W.BODE,R.HUBER        2CGA    8
JRNL        TITL  BOVINE CHYMOTRYPSINOGEN *A. X-RAY CRYSTAL STRUCTURE 2CGA    9
>>>>>>
REMARK      1                                     2CGA   14
REMARK      1 REFERENCE 1                       2CGA   15
>>>>>>
SEQRES      1 A  245  CYS GLY VAL PRO ALA ILE GLN PRO VAL LEU SER GLY LEU 2CGA   71
SEQRES      2 A  245  SER ARG ILE VAL ASN GLY GLU GLU ALA VAL PRO GLY SER 2CGA   72
SEQRES      3 A  245  TRP PRO TRP GLN VAL SER LEU GLN ASP LYS THR GLY PHE 2CGA   73
>>>>>>
>>>>>>
SEQRES      17 B 245  LEU VAL GLY ILE VAL SER TRP GLY SER SER THR CYS SER 2CGA  106
SEQRES      18 B 245  THR SER THR PRO GLY VAL TYR ALA ARG VAL THR ALA LEU 2CGA  107
SEQRES      19 B 245  VAL ASN TRP VAL GLN GLN THR LEU ALA ALA ASN      2CGA  108
>>>>>>
CRYST1      59.300  77.100 110.100  90.00  90.00  90.00 P 21 21 21  8 2CGA  113
ORIGX1      1.000000  0.000000  0.000000  0.000000  0.000000  2CGA  114
ORIGX2      0.000000  1.000000  0.000000  0.000000  0.000000  2CGA  115
ORIGX3      0.000000  0.000000  1.000000  0.000000  0.000000  2CGA  116
SCALE1      .016863  0.000000  0.000000  0.000000  0.000000  2CGA  117
SCALE2      0.000000  .012970  0.000000  0.000000  0.000000  2CGA  118
SCALE3      0.000000  0.000000  .009083  0.000000  0.000000  2CGA  119
MTRIX1      1  .987700  -.155000  .017700  6.21700  1  2CGA  120
MTRIX2      1  .022800  -.031400  -.999200  115.61600  1  2CGA  121
MTRIX3      1  -.154300  .987400  -.034600  -3.74800  1  2CGA  122
ATOM        1  N  CYS A  1  -10.656  55.938  41.808  1.00 11.66  2CGA  123
ATOM        2  CA CYS A  1  -10.044  57.246  41.343  1.00 11.66  2CGA  124
ATOM        3  C  CYS A  1  -10.076  58.323  42.431  1.00 11.66  2CGA  125
ATOM        4  O  CYS A  1  -10.772  58.097  43.448  1.00 11.66  2CGA  126
ATOM        5  CB CYS A  1  -10.807  57.718  40.066  1.00 11.66  2CGA  127
>>>>>>
>>>>>>
ATOM        744 N  ASN A 100  -13.152  77.724  22.378  1.00  8.65  2CGA  866
ATOM        745 CA ASN A 100  -14.213  76.940  23.011  1.00  8.65  2CGA  867
ATOM        746 C  ASN A 100  -14.134  75.441  22.693  1.00  8.65  2CGA  868
ATOM        747 O  ASN A 100  -13.706  75.062  21.563  1.00  8.65  2CGA  869
>>>>>>
>>>>>>
ATOM        1461 N  VAL A 200  -9.212  70.793  39.923  1.00  9.30  2CGA1583
ATOM        1462 CA VAL A 200  -9.875  69.689  40.639  1.00  9.30  2CGA1584
ATOM        1463 C  VAL A 200  -10.634  70.148  41.868  1.00  9.30  2CGA1585
ATOM        1464 O  VAL A 200  -10.151  70.985  42.657  1.00  9.30  2CGA1586
>>>>>>
HETATM     3601 O  HOH  601  -20.008  66.224  26.138  1.00 26.69  2CGA3723
HETATM     3602 O  HOH  602  -21.333  66.182  28.756  1.00 18.10  2CGA3724
HETATM     3603 O  HOH  603  -18.000  68.022  22.774  1.00 34.03  2CGA3725
MASTER     60  3  0  0  0  0  0  9 3927  2  0  38 2CGAA  6
END                                                2CGA4053

```

Figura 12. Versión resumida de un archivo pdb representativo (2CGA). Los símbolos >>>>> indican ausencia de varios renglones semejantes³⁷

3.1 Difracción de rayos X

La primera estructura de una proteína, obtenida con resolución atómica, fue la de la mioglobina de cachalote (*Physeter macrocephalus*) en 1957³⁸. John Kendrew y Max Perutz recibieron el premio Nobel de química en 1962, tanto por esta contribución como por la primera estructura de la hemoglobina humana, publicada en 1960. La técnica de difracción de rayos X fue usada en la determinación de estas estructuras y actualmente sigue siendo la técnica más comúnmente usada para resolver nuevas estructuras de proteínas.

Dos grandes problemas que enfrentaron los primeros biólogos estructurales fueron: el cómo resolver estructuras de objetos separados entre sí por la distancia de un enlace covalente (alrededor de 1.5 Å) y cómo prevenir el promediado rotacional de la molécula durante la recolección de datos. El primer problema recae en el hecho de que la longitud de onda de la radiación usada para resolver objetos no puede ser mucho mayor que el tamaño del objeto en cuestión. Por lo tanto, para resolver objetos del orden de 1.5 Å, se necesita radiación con una longitud de onda cercana o menor a este valor. Los rayos X, que son fáciles de producir y relativamente inertes, cubren esta necesidad.

Cuando las moléculas se encuentran en disolución, existen en todas las formas de rotación y traslación posibles en su contenedor. Si esas moléculas fueran expuestas a rayos X, la imagen tomada estaría promediada rotacionalmente, perdiendo información de alta resolución acerca de las posiciones relativas de cada átomo. Para obtener una mejor imagen, las moléculas deben estar fijas en una o en un número pequeño de conformaciones, este requerimiento se cumple al utilizar monocristales en los estudios de difracción de rayos X. Las moléculas en un cristal están dispuestas en un arreglo fijo, regular y repetitivo que facilita la obtención de una imagen con resolución atómica.

En un experimento de difracción de rayos X, los patrones de dispersión de estos rayos, provenientes de un cristal de proteína, son comparados con cálculos de los

patrones que se esperarían para la proteína dentro de la misma red cristalina, una buena correlación entre los patrones de dispersión calculados y observados indica una estructura acertada. Una vez que se tiene disponible una muestra pura con una alta concentración de proteína, el límite para la capacidad de resolver estructuras macromoleculares por cristalografía de rayos X viene dado por la calidad de difracción de los cristales y no por el tamaño o la complejidad de la molécula.

3.2 Microscopía electrónica

Las proteínas que se encuentran naturalmente dentro de membranas biológicas son especialmente difíciles de cristalizar, ya que involucran la disolución de la membrana con detergentes, así como el aislamiento y cristalización del complejo detergente-proteína. Sin embargo, estas proteínas pueden formar arreglos bidimensionales dentro de la membrana, con lo cual pueden ser analizadas mediante el uso de microscopía electrónica. Dado que las moléculas en este arreglo bidimensional tienen la misma orientación, es posible calcular una “imagen promedio” de una molécula de proteína. Desafortunadamente, la resolución de dichas imágenes en la mayoría de los casos está limitada a entre 10 y 20 Å, la cual no es suficiente para determinar el esqueleto de la proteína.

Durante las últimas dos décadas se han realizado grandes avances en el uso de esta técnica, como lo es el desarrollo de la criomicroscopía electrónica, una técnica en la que la muestra es estudiada a temperaturas criogénicas, empleando nitrógeno líquido para congelar rápidamente las proteínas en disolución. Esta nueva técnica ha avanzado a tal punto que han sido determinadas varias estructuras de ribosomas de diferentes organismos, incluido el humano, con una resolución cercana a los 3 Å^{39,40}.

3.3 Espectroscopia de Resonancia Magnética Nuclear

Otra técnica adecuada para la determinación de estructuras de macromoléculas a nivel atómico es la espectroscopia de resonancia magnética nuclear (RMN). Esta se basa en el hecho de que los núcleos atómicos con un número impar de nucleones (^1H , ^{13}C , ^{31}P), o un número impar tanto de protones como de neutrones (^{14}N) poseen un momento angular llamado espín. Este espín de un cuerpo cargado positivamente resulta en la generación de un pequeño campo magnético que puede ser medido por un espectrómetro de RMN.

Dado que el campo magnético de los núcleos es afectado por los electrones que rodean a los átomos, la espectroscopia de RMN puede medir el ambiente molecular único de cada núcleo, obteniendo así, información de los núcleos cercanos entre sí y así generar un conjunto de restricciones de distancia que existen entre los núcleos de la molécula. Resolver una estructura mediante RMN involucra la manipulación computacional de las conformaciones de un conjunto de átomos enlazados hasta encontrar una que se ajuste a los contactos observados. Dentro de las ventajas de la resonancia magnética nuclear para la resolución de estructuras moleculares está la capacidad de trabajar con una muestra en disolución en lugar de con una en estado cristalino, sin embargo, existe un límite de aproximadamente 50 kDa para el tamaño molecular que puede ser estudiado con resolución atómica.

4. Predicción de estructura de proteínas

Los primeros trabajos en el campo de la predicción de estructura terciaria de proteínas estuvieron enfocados en métodos basados en principios físicos, como un intento de entender el proceso de plegado^{41,42}. La idea básica del plegado computacional de una proteína es la de encontrar la estructura de menor energía libre para la secuencia de aminoácidos dada, basándose en la hipótesis termodinámica formulada por Anfinsen⁵, a través de su búsqueda en el exageradamente largo espacio conformacional de la proteína. Aunque esto representa la meta final del modelado de la estructura de una proteína, actualmente es una meta lejana, debido a lo enorme y complejo del espacio conformacional de una proteína, en comparación con los recursos computacionales disponibles actualmente.

Una alternativa es la de predecir únicamente la estructura final del proceso de plegado de una proteína, la cual resulta ser más atractiva ya que en la práctica el problema se vuelve más sencillo de resolver, además de que al enfocarse únicamente en la estructura estática final de un proceso de plegado complejo, se permite tomar ventaja de la gran cantidad de información disponible en las estructuras de proteínas previamente resueltas, lo que convierte al problema de la predicción de estructura de proteínas en un paradigma de predicción de estructura basado en plantillas⁴³.

Los métodos de predicción de estructura basados en plantillas, que van desde los métodos *de novo*, que usan plantillas estructurales relativamente pequeñas, hasta los métodos de modelado por homología, que usan una proteína completa como plantilla, han dado grandes pasos en la predicción de estructuras de proteínas en los últimos años y han sido empleados para realizar muchas predicciones de estructuras aún antes de que los datos experimentales estuvieran disponibles; posteriormente se demostró que estos métodos podían ser usados como guías útiles en el diseño experimental de estructuras de proteínas⁴⁴.

De manera general, las técnicas empleadas en la predicción de estructura de proteínas pueden ser clasificadas en tres categorías: métodos *ab initio*, hilvanado (threading) o reconocimiento del plegado y modelado por homología^{45,46}.

Los métodos *ab initio* predicen la estructura de una proteína sin el uso de ningún tipo de información estructural de proteínas previamente resueltas, en su lugar, se basan en los primeros principios de la física. El modelado por homología se basa en el alineamiento de secuencias entre la proteína en cuestión y una proteína plantilla, cuya estructura es conocida; por esta razón, la precisión de la predicción en esta clase de métodos depende en gran medida de la similitud en secuencia entre las dos proteínas. El hilvanado de proteínas representa una clase más general de técnicas de predicción de estructura que el modelado por homología, ya que usa tanto la similitud en secuencia como el ajuste estructural entre la proteína en cuestión y la proteína plantilla.

El modelado por homología ha sido ampliamente usado para una predicción de estructura con un alto nivel de detalle (todos los átomos pesados, por ejemplo) cuando la proteína problema tiene un homólogo cercano en el Protein Data Bank⁴⁷, mientras que el hilvanado de proteínas es usado frecuentemente para la predicción de la estructura de la cadena principal de una proteína cuando esta presenta remotos homólogos estructurales en el PDB. Los límites entre estas tres clases de técnicas se han vuelto difusos debido a que se han comenzado a integrar sus fortalezas para hacer de los métodos de predicción de estructura de proteínas más efectivos y con una aplicación más general⁴⁸.

4.1 Métodos *ab initio*

Los métodos de modelado *ab initio* tratan de construir modelos proteicos desde cero, basándose esencialmente en principios. De acuerdo a la hipótesis termodinámica, la estructura nativa de una proteína depende solo de su secuencia de aminoácidos, por lo que en principio debería ser posible predecir la estructura de una proteína sobre la base de sus propiedades fisicoquímicas. Existen bastantes procedimientos posibles que, o bien intentan imitar el plegado de proteínas, o bien aplican algún método estocástico para buscar posibles soluciones (por ejemplo, la optimización global de una función de energía apropiada). Un problema importante enfrentado por los métodos *ab initio* radica en que las cadenas polipeptídicas tienen cantidades astronómicas de conformaciones no nativas de baja energía, por lo que por el momento es bastante difícil, aún con las computadoras más rápidas disponibles, determinar la conformación polipeptídica de menor energía. No obstante, muchos algoritmos intensivos computarizados han conseguido buenos resultados en la predicción de la estructura de péptidos pequeños y un éxito esporádico en la predicción de los plegamientos de polipéptidos más grandes⁴⁹⁻⁵². Entre estos métodos están los algoritmos bioinspirados como la colonia de hormigas y el vuelo en enjambre, y los evolutivos como los genéticos. En consecuencia, la capacidad para predecir la estructura nativa de un polipéptido a partir de su secuencia aún es uno de los objetivos no alcanzados más importantes de la bioquímica.

4.2 Métodos basados en plantillas.

Las premisas básicas para la predicción de estructuras de proteínas basada en plantillas son tres:

- 1) Secuencias de aminoácidos similares adoptan estructuras similares⁵³.
- 2) Muchas secuencias no relacionadas se pliegan de manera similar⁵⁴.
- 3) Existe solamente un pequeño número de plegados estructurales únicos, en comparación con el número de proteínas existentes en la naturaleza⁵⁵.

La primera observación es la base del modelado por homología, mientras que las dos siguientes son los fundamentos del hilvanado o reconocimiento del plegado. El primer modelo estructural obtenido utilizando la aproximación basada en plantillas fue construido en 1969 por Browne y colaboradores. El desarrollo computacional posterior ha provisto de herramientas importantes para el modelado comparativo de estructuras de proteínas⁵⁶ y desde el reporte de este primer modelo, la estructura de muchas proteínas importantes ha sido modelada mediante este tipo de métodos⁵⁷⁻⁶².

El desarrollo de la aproximación por hilvanado fue introducido por Jones y colaboradores en 1992⁶³. Este tipo de métodos estima teóricamente el número de plegados únicos presentes en la naturaleza basados en el hecho de que, en los últimos años, menos del 10% de estructuras depositadas en el PDB presentan nuevos tipos de plegado⁶⁴. Se ha encontrado también que muchas proteínas con secuencias no relacionadas se pliegan de manera similar y que especialmente algunos plegados parecen ser muy frecuentes, aún en proteínas con ninguna aparente similitud en secuencia⁶⁵⁻⁶⁸. Con base en estas observaciones, las técnicas de hilvanado de proteínas han sido empleadas para resolver dos preguntas clave: ¿Qué plegado tomará una proteína dada, de los plegados conocidos? y ¿Dónde deben colocarse cada uno de los residuos de la proteína dada, en el plegado identificado?

Los métodos de predicción de estructura basados en plantillas generalmente consisten de cinco etapas:

- 1) La identificación de plantillas estructurales a través de métodos basados tanto en estructura como en secuencia.
- 2) El alineamiento de la secuencia problema con la plantilla identificada.
- 3) La construcción del modelo, basándose en el alineamiento previo.
- 4) La evaluación del modelo.
- 5) El refinamiento del modelo.

La idea básica del hilvanado de proteínas es la de “enhebrar” los aminoácidos de una proteína dada, siguiendo su orden secuencial y permitiendo tanto inserciones como huecos, en las posiciones estructurales de una proteína plantilla, de una manera óptima, medida por una función de evaluación. Este procedimiento es repetido para cada una de las plantillas en una base de datos de estructuras de proteínas. La calidad del alineamiento es usualmente determinada empleando términos de energía, basados ya sea en parámetros estadísticos o en principios físicos.

4.3 Potenciales estadísticos

A diferencia de la energía física usada en los estudios de plegado de proteínas^{8,9}, las funciones empleadas para evaluar un alineamiento particular de estructura-secuencia en el hilvanado de proteínas están basadas principalmente en parámetros estadísticos, también conocidos como basados en el conocimiento. La idea de utilizar potenciales basados en el conocimiento reside en el hecho de que las estructuras determinadas experimentalmente contienen una gran cantidad de información acerca de las fuerzas que estabilizan la estructura de las proteínas. El análisis estadístico de estructuras de proteínas puede, en potencia, capturar las reglas subyacentes que gobiernan la estabilidad estructural de las proteínas y puede ser realizado confiando, de manera implícita o explícita en el principio de Boltzmann: estados frecuentemente observados corresponden a estados de baja energía del sistema⁶⁹.

La idea de generar potenciales basados en el conocimiento a partir de estructuras de proteínas conocidas fue previa a sus primeras aplicaciones en el hilvanado de proteínas. Por ejemplo, Tanaka y Scheraga¹¹ reportaron en 1976 un estudio acerca de los potenciales de interacción de mediano y largo alcance, así como su aplicación en la predicción de estructuras de proteínas; el cual fue seguido por numerosos estudios similares^{10,42,70,71}.

Las primeras aproximaciones de este tipo de potenciales generalmente ignoraban las similitudes en secuencia entre la secuencia problema y la proteína plantilla⁷², en su lugar, consideraban únicamente las preferencias de cada aminoácido en la proteína problema por un ambiente estructural particular. El término de energía en este tipo de potenciales puede ser calculado a partir de una base de datos de estructuras de proteínas utilizando la estadística de Boltzmann. La idea básica puede ser descrita de la siguiente forma: si un aminoácido es observado frecuentemente en un ambiente estructural particular, esto sugiere que ese ambiente es favorable para dicho aminoácido. Este término de energía puede ser escrito de la siguiente manera:

$$E_{(i,j)} = -\ln\left(\frac{f^O_{i,j}}{f^E_{i,j}}\right) \quad (1)$$

Donde $f^O_{i,j}$ es la frecuencia observada del aminoácido i en el ambiente estructural j , mientras que $f^E_{i,j}$ representa la frecuencia esperada del aminoácido i en el ambiente estructural j .

Otro tipo de función de energía es conocido como energía de pares de contactos, la cual describe la interacción entre dos residuos. En ella se mide la preferencia de que dos tipos de aminoácidos se encuentren cercanos entre sí. Jones y colaboradores aplicaron la función de energía desarrollada por Sippl^{10,63} en su estudio de reconocimiento del plegado. La idea básica de esta función de energía proviene de la mecánica estadística y puede ser escrita como sigue:

$$E_{(i,j)} = -kT \ln\left(\frac{f^o_{i,j}}{f^r_{i,j}}\right) \quad (2)$$

Donde k y T son la constante de Boltzmann y la temperatura absoluta, respectivamente. $f^o_{i,j}$ es la frecuencia observada de los residuos i y j a una distancia dada, donde la distancia es medida entre los carbonos β de ambos

residuos; y $f_{i,j}^r$ es la frecuencia con la que se presentan los residuos i y j a una distancia dada en un estado de referencia. Existen dos formas de calcular la energía de pares de contactos, dependientes de distancia e independientes de ella, observando que la primera forma provee de mejores resultados. Uno de los mayores problemas en este tipo de funciones es el de establecer un estado de referencia adecuado para estimar $f_{i,j}^r$, lo que ha dado origen a una gran cantidad de potenciales de este estilo^{10,12-15}.

Adicionalmente a estos potenciales con formas simples, existen funciones más sofisticadas que incluyen términos distintos a los descritos previamente, los cuales han mostrado una mejora en comparación con los métodos más simples, pero dada su complejidad, han vuelto más complicada su aplicación práctica.

La calidad de los potenciales basados en el conocimiento normalmente recae en su capacidad de reconocer la conformación nativa de una proteína dentro de un grupo de señuelos con conformaciones similares. Para medir la similitud de un señuelo dado se compara la posición de los átomos en el señuelo con la posición de los átomos en la estructura nativa. El método más empleado es el de medir la desviación cuadrática media entre las dos estructuras, el cual se calcula con base a los carbonos α de la cadena principal, aunque el uso de todos los átomos es más útil cuando se comparan modelos de muy alta calidad, es decir, con pequeñas diferencias entre sí. No es suficiente con identificar correctamente la estructura nativa de una proteína dentro de un conjunto de señuelos, para darle significancia estadística a la evaluación se recurre a métodos como el z-score, el cual determina la energía de la estructura nativa en unidades de desviación estándar relativas al promedio de energía de los señuelos⁷³. Es decir:

$$z = \frac{E - \bar{E}}{\sigma} \quad (3)$$

Donde \bar{E} y σ son el promedio y la desviación estándar de la distribución de energía resultante del conjunto de señuelos. Esta aproximación es ampliamente usada y efectiva hasta cierto punto, ya que supone una distribución normal para la energía de los señuelos. En la literatura existe una gran variedad de grupos de señuelos empleados para este fin^{74,75}.

III. TRABAJOS RELACIONADOS

El uso de potenciales estadísticos como una herramienta para la resolución del problema de la predicción de la estructura de una proteína fue iniciado por Tanaka y Scheraga¹¹. A partir de ese momento, han surgido una gran cantidad de aproximaciones similares¹²⁻¹⁵ que han sido empleadas en diversas aplicaciones, tales como la caracterización de interacciones proteína-proteína⁷⁶⁻⁸⁰ o proteína-ácido nucleico^{81,82}. Dentro de todos ellos y debido a la naturaleza del presente trabajo, cabe resaltar aquellos en los cuales se emplea la descripción de los puentes de hidrógeno intramoleculares.

Tal es el caso del potencial desarrollado en 2003 por Kortemme y colaboradores⁷⁶, en el cual emplean la descripción de cuatro parámetros geométricos para describir los puentes de hidrógeno encontrados en proteínas, estos son: la distancia donador-aceptor, los ángulos formados tanto en el átomo aceptor como en el hidrógeno involucrado en la formación del puente y el ángulo diedro formado por la rotación del enlace entre el átomo aceptor y su átomo base (**Fig. 13**). Se definió un puente de hidrógeno, como aquel en el que un átomo de hidrógeno polar estuviera a una distancia entre 1.4 y 2.6 Å de un átomo aceptor.

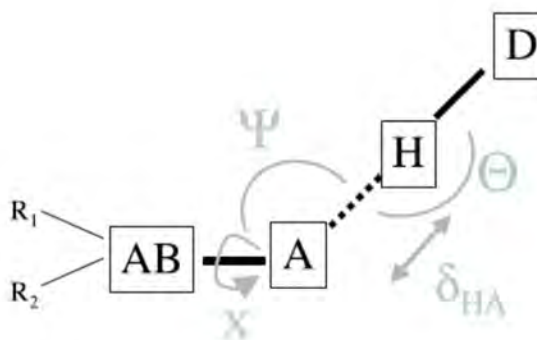


Figura 13. Representación esquemática de los parámetros usados por Kortemme y colaboradores⁷⁶ para describir los puentes de hidrógeno. δ_{HA} , distancia hidrógeno-aceptor; Ψ , ángulo al átomo aceptor; Θ , ángulo al átomo de hidrógeno y X, ángulo diedro en el enlace aceptor-hidrógeno.

En este caso, además de incluir los parámetros mencionados, se clasificaba los puentes de hidrógeno en: cadena principal-cadena principal, cadena lateral-cadena lateral; que a su vez eran distinguidos, en el primer caso de acuerdo a la estructura secundaria a la que pertenecían y en el segundo caso debido a la hibridación del átomo aceptor. Este potencial es acoplado a los cálculos de energía de los campos de fuerza, para sustituir los términos usados en el cálculo de la energía del puente de hidrógeno, mostrando una notable mejoría en el desempeño de los campos de fuerza.

Potenciales como el de Sánchez y colaboradores, basado en Beta-Complex⁸³, que usa la teoría de Beta-Shape y la cuasi-triangulación para la descripción de la estructura de las proteínas, han demostrado que incluso las descripciones geométricas simples del puente de hidrógeno formado por la cadena principal de proteínas (distancia hidrógeno-aceptor y ángulo al aceptor) proveen de mejoras significativas en el desempeño de potenciales complejos.

Por otro lado, los campos de fuerza utilizan seis parámetros geométricos de enlace para describir por completo las interacciones en las que participa un átomo dado, aunque los puentes de hidrógeno raramente son tomados en cuenta en esta descripción, ya que en la mayoría de los casos son tratados ya sea como una interacción coulombica, una interacción de tipo van der Waals o una mezcla de ambas^{8,9}, que a pesar de ser una descripción formal de este tipo de interacciones, se ha visto superada por descripciones en forma de potenciales estadísticos⁷⁶.

Por estas razones el presente trabajo explora la idea de que el estudio de los seis parámetros geométricos que describen completamente las interacciones, en las que participan los átomos de hidrógeno formadores de puente en la cadena principal de proteínas, pueda, por sí solo, ser usado como un potencial estadístico que permita identificar la estructura nativa de una proteína dentro de un conjunto de señuelos.

IV. DESCRIPCIÓN DEL PUENTE DE HIDRÓGENO

El objetivo del presente trabajo es el desarrollo de un potencial estadístico, para lo cual es conveniente que la cantidad de datos a analizar sea lo más grande posible. Dentro de los puentes de hidrógeno intramoleculares presentes en proteínas resulta evidente que los más abundantes son aquellos formados por la cadena principal de las mismas. Aunque se podría pensar que al considerar sólo la cadena principal de estas se está obteniendo información únicamente acerca de su estructura secundaria, es necesario tomar en cuenta el hecho de que estos datos son tomados de estructuras experimentales y que bajo cualquiera de los modelos de plegamiento aceptados se puede establecer una relación entre la estructura secundaria y terciaria de una proteína, ya sea que la primera dicté la segunda o viceversa, por ello es posible suponer que con el análisis de los puentes de hidrógeno formados por la cadena principal se estaría obteniendo, de forma indirecta, información acerca de la estructura terciaria de las proteínas.

Por otra parte, cabe mencionar que el presente trabajo no busca estimar de manera precisa la energía de las interacciones por puente de hidrógeno presentes en proteínas, sino describir el comportamiento de dicha energía con respecto a distintos parámetros geométricos y utilizar esta descripción en la identificación de estructuras correctamente plegadas. Por esta razón no se emplean términos de tipo electrostático, Van der Waals o el efecto del disolvente en la descripción de los puentes de hidrógeno.

De esta manera, se establecieron seis parámetros geométricos para estudiar los puentes de hidrógeno formados por la cadena principal de proteínas, estos son: (a) la distancia hidrógeno-oxígeno, (b) el ángulo al átomo de hidrógeno, formado por el átomo de nitrógeno donador y el átomo de oxígeno aceptor, (c) el ángulo al átomo de oxígeno, formado por el átomo de hidrógeno y el átomo de carbono carbonílico, (d) el ángulo diedro que corresponde a la rotación del enlace carbono-oxígeno, (e) el ángulo diedro que corresponde a la rotación del puente de hidrógeno y (f) el ángulo diedro que corresponde a la rotación del enlace hidrógeno-nitrógeno (**Fig. 14**).

Para la medición de los parámetros establecidos se tiene contemplada la presencia explícita de los átomos de hidrógeno en las estructuras a estudiar, lo cual hace más sencilla la definición de los puentes de hidrógeno en comparación con descripciones que no incluyen estos átomos explícitamente, en las cuáles es necesario tomar en cuenta la orientación relativa tanto de los átomos donador y aceptor como de sus respectivas bases⁸⁴.

Bajo el nivel de descripción utilizado en este trabajo, se considera como un puente de hidrógeno a cualquier caso en el que un átomo de hidrógeno, en un grupo NH, estuviera a una distancia menor a 2.6 Å de un átomo de oxígeno, en un grupo CO.

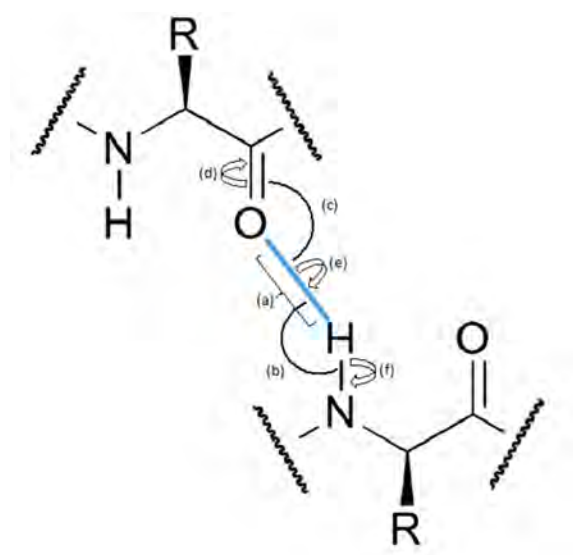


Figura 14. Representación esquemática de los parámetros usados para describir los puentes de hidrógeno. (a) la distancia hidrógeno-oxígeno, (b) el ángulo al átomo de hidrógeno, (c) el ángulo al átomo de oxígeno (d) el ángulo diedro que corresponde a la rotación del enlace carbono-oxígeno, (e) el ángulo diedro que corresponde a la rotación del puente de hidrógeno y (f) el ángulo diedro que corresponde a la rotación del enlace hidrógeno-nitrógeno.

Una vez establecidos estos parámetros, se desarrolló un programa en lenguaje Python (**Fig. 1 – Anexo**) que permitiera medirlos adecuadamente.

V. SELECCIÓN DE LA MUESTRA

Con la finalidad de obtener la muestra con la mayor cantidad de hidrógenos polares incluidos explícitamente, evitar la presencia de información proveniente de interacciones debidas a la estructura cuaternaria o a residuos de aminoácidos poco comunes y obtener una muestra lo más heterogénea posible, se seleccionaron y obtuvieron del Protein Data Bank estructuras de proteínas que cumplieran con las siguientes características: monoméricas, sin residuos modificados y con identidad de secuencia no mayor a 30%, para la posterior formación de los siguientes grupos:

- I. 73 estructuras resueltas por difracción de rayos X con una resolución de 1.0 Å o mejor (**ERX**) (**Tabla 1 - Anexo**).
- II. 4020 conjuntos de estructuras resueltas por resonancia magnética nuclear (**ERMN**), utilizando el primer confórmero reportado en ellos (**Tabla 2 - Anexo**).

Se midieron todos los puentes de hidrógeno presentes en cada una de las muestras y se obtuvieron, para su comparación, las distribuciones de frecuencias (normalizadas) de los seis parámetros geométricos establecidos (**Fig. 15**); estableciendo para su creación intervalos de 0.05 Å para (a), de 5 grados para (b) y (c) y de 10 grados para (d), (e) y (f). Se obtuvieron un total de 9,042 puentes de hidrógeno para el grupo **ERX** y 236,576 para el grupo **ERMN**.

Dado que se desconoce el tipo de distribución de frecuencia que siguen los distintos parámetros analizados, las distribuciones obtenidas pueden ser comparadas utilizando una prueba no paramétrica como el contraste de Kolmogorov-Smirnov⁸⁵, el cual hace uso del siguiente estadístico:

$$D_{n_1, n_2} = \max |F_{n_1}(x) - F_{n_2}(x)|$$

Que se refiere a la mayor diferencia encontrada entre las distribuciones de frecuencia acumulada (F) para el parámetro x de dos muestras de tamaño n_1 y n_2 , dado un determinado número de intervalos.

Si las dos muestras proceden de la misma población, sus distribuciones de frecuencia acumulada no pueden ser muy distintas, por lo que el contraste es siempre de una cola y se rechaza la hipótesis de igual distribución si el estadístico toma un valor suficientemente grande. El valor crítico (λ) de dicho estadístico, con un nivel de confianza del 95%, puede aproximarse por:

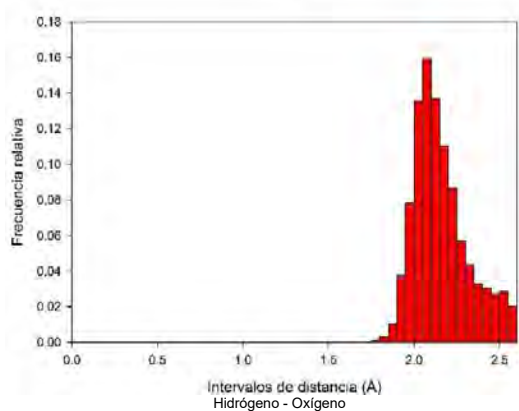
$$\lambda = 1.22 \sqrt{\frac{n_1+n_2}{n_1 n_2}}$$

Como se puede apreciar, este valor crítico se aproxima a cero a medida que aumenta el tamaño de las muestras, para las distribuciones obtenidas en el presente trabajo, el valor crítico de λ es de 0.0131. Con sólo observar las gráficas se puede notar que existen pares de intervalos cuya diferencia por si sola supera fácilmente este valor, por lo que se puede establecer que las distribuciones de frecuencia son distintas para todos los parámetros estudiados en los grupos **ERX** y **ERMN**, aun sin realizar el contraste.

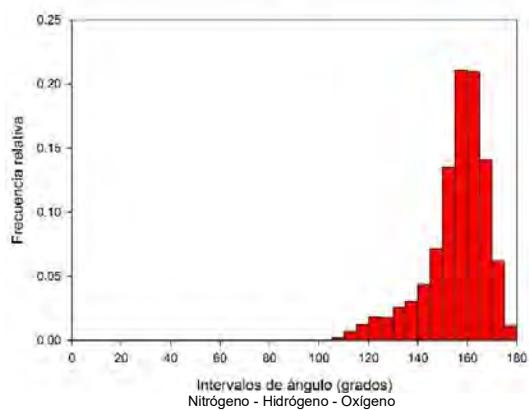
Con un análisis general de estas distribuciones se puede notar que para todos los parámetros estudiados, el comportamiento es semejante entre los dos grupos, aunque existen diferencias importantes, como las posiciones de las frecuencias máximas y el hecho de que las distribuciones de frecuencias del grupo **ERMN** son, en su mayoría, más dispersas que las del grupo **ERX**, razón por la cual los máximos en frecuencia se encuentran a una altura menor en el grupo **ERMN**. Estas diferencias pueden ser explicadas con base en los métodos empleados para la determinación de las estructuras estudiadas en ambos grupos, ya que las estructuras obtenidas por difracción de rayos X provienen de un cristal, en el cuál el movimiento de las moléculas se encuentra restringido, a diferencia de las estructuras determinadas por resonancia magnética nuclear, en donde las moléculas se encuentran en disolución y tienen una mayor libertad de movimiento.

ERX

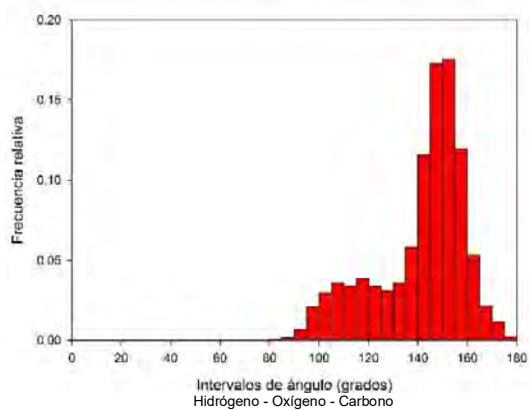
(a)



(b)

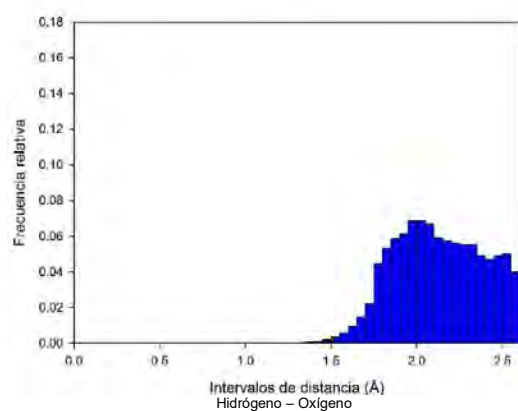


(c)

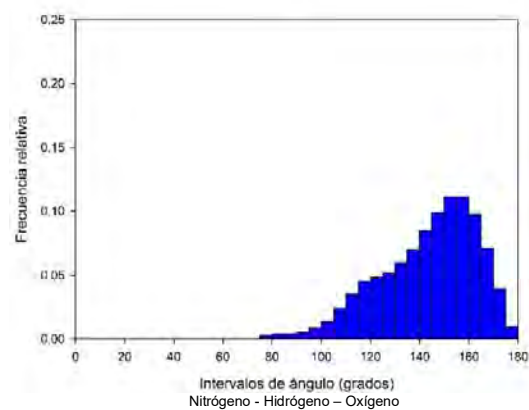


ERMN

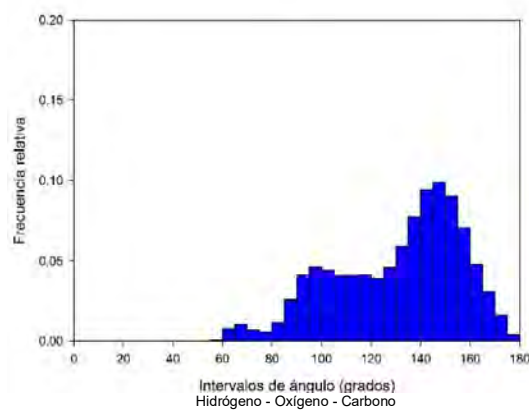
(a)



(b)



(c)



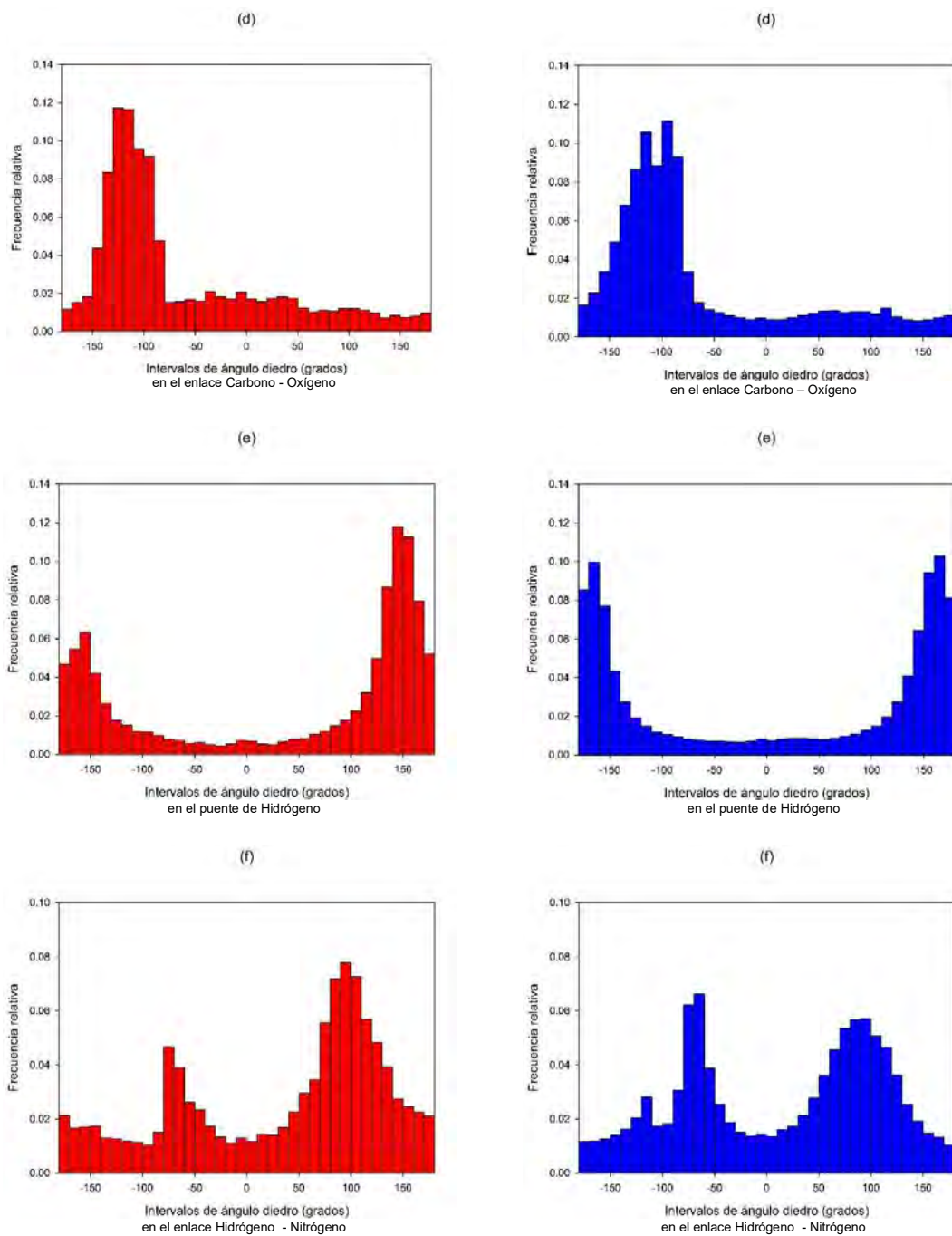


Figura 15. Distribuciones de frecuencias de los parámetros geométricos estudiados en los puentes de hidrógeno formados por la cadena principal de proteínas, en muestras experimentales. (a) la distancia hidrógeno-oxígeno, (b) el ángulo al átomo de hidrógeno, (c) el ángulo al átomo de oxígeno (d) el ángulo diedro que corresponde a la rotación del enlace carbono-oxígeno, (e) el ángulo diedro que corresponde a la rotación del puente de hidrógeno y (f) el ángulo diedro que corresponde a la rotación del enlace hidrógeno-nitrógeno.

Los grupos de confórmeros en los que se prueban los potenciales estadísticos, al igual que la mayoría de las estructuras determinadas por cristalografía de rayos X, presentan el problema de no incluir los hidrógenos polares en las estructuras. Por esta razón, se hace necesario establecer una metodología para añadirlos de tal forma que al utilizar dicha metodología se obtengan datos, para los parámetros geométricos estudiados, que sean congruentes con los obtenidos de estructuras de alta resolución. Para realizar esta comparación y utilizando GROMACS 4.5, se sustituyeron estos hidrógenos en las estructuras de los grupos **ERX** y **ERMN** de acuerdo a los parámetros del campo de fuerza AMBER99SB⁸, se determinó la cantidad de puentes de hidrógeno encontrados y la cantidad de estos que correspondían a los observados en las estructuras obtenidas experimentalmente (**Tabla 2**), así como las distribuciones de frecuencias (normalizadas) de los seis parámetros geométricos establecidos (**Fig. 16**).

Tabla 2. Puentes de hidrógeno preservados al sustituir los hidrógenos polares en estructuras experimentales, de acuerdo a los parámetros del campo de fuerza AMBER99SB.

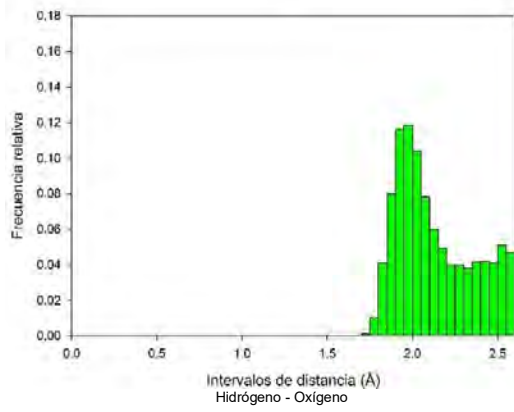
	ERX	ERX + AMBER99SB	ERMN	ERMN + AMBER99SB
Puentes de hidrógeno totales	9042	9153	236,576	237,718
Puentes de hidrógeno preservados (%)*	8308 (91.88)		232,762 (98.39)	
“Nuevos” puentes de hidrógeno (%)*	845 (9.23)		4,956 (2.08)	

*El porcentaje de puentes de hidrógeno preservados es calculado con respecto al total encontrado en estructuras experimentales mientras que el porcentaje de “nuevos” puentes de hidrógeno se calcula con respecto al total encontrado al reemplazar los hidrógenos polares con el campo de fuerza AMBER99SB, por esa razón la suma no es 100.

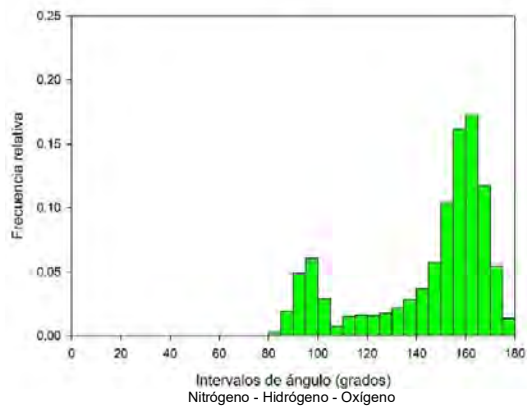
Al reemplazar los hidrógenos polares de las estructuras experimentales de acuerdo a los parámetros del campo de fuerza AMBER99SB se modifican considerablemente ambos grupos de estructuras, viéndose menos afectados los resultados observados en estructuras resueltas por RMN en comparación con aquellas resueltas por difracción de rayos X.

ERX + AMBER99SB

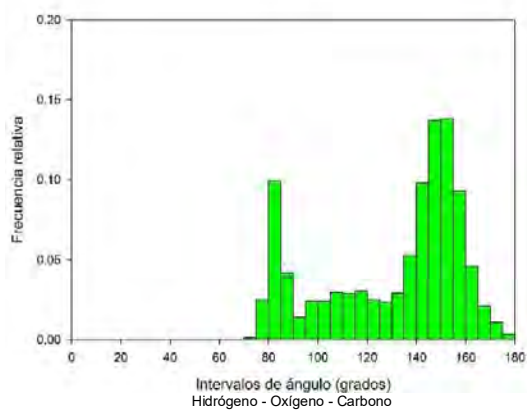
(a)



(b)

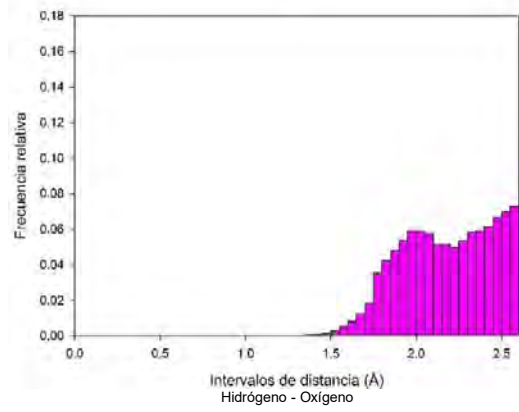


(c)

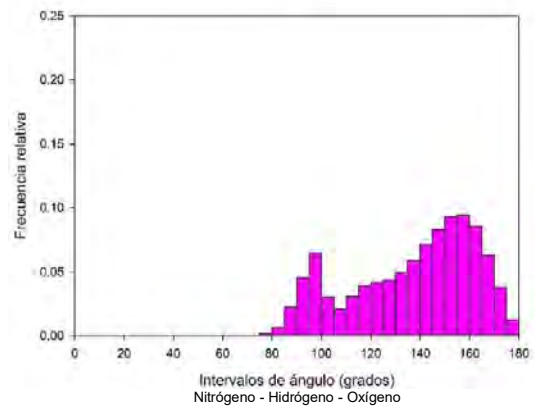


ERMN + AMBER99SB

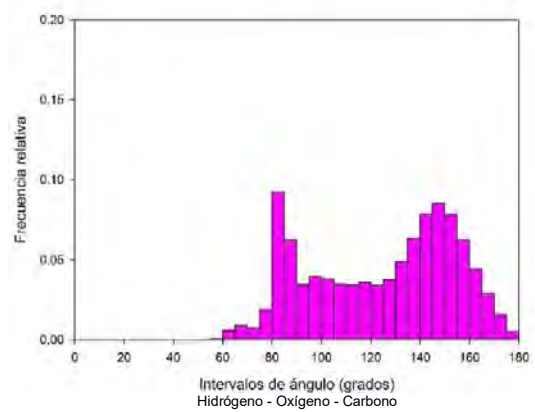
(a)



(b)



(c)



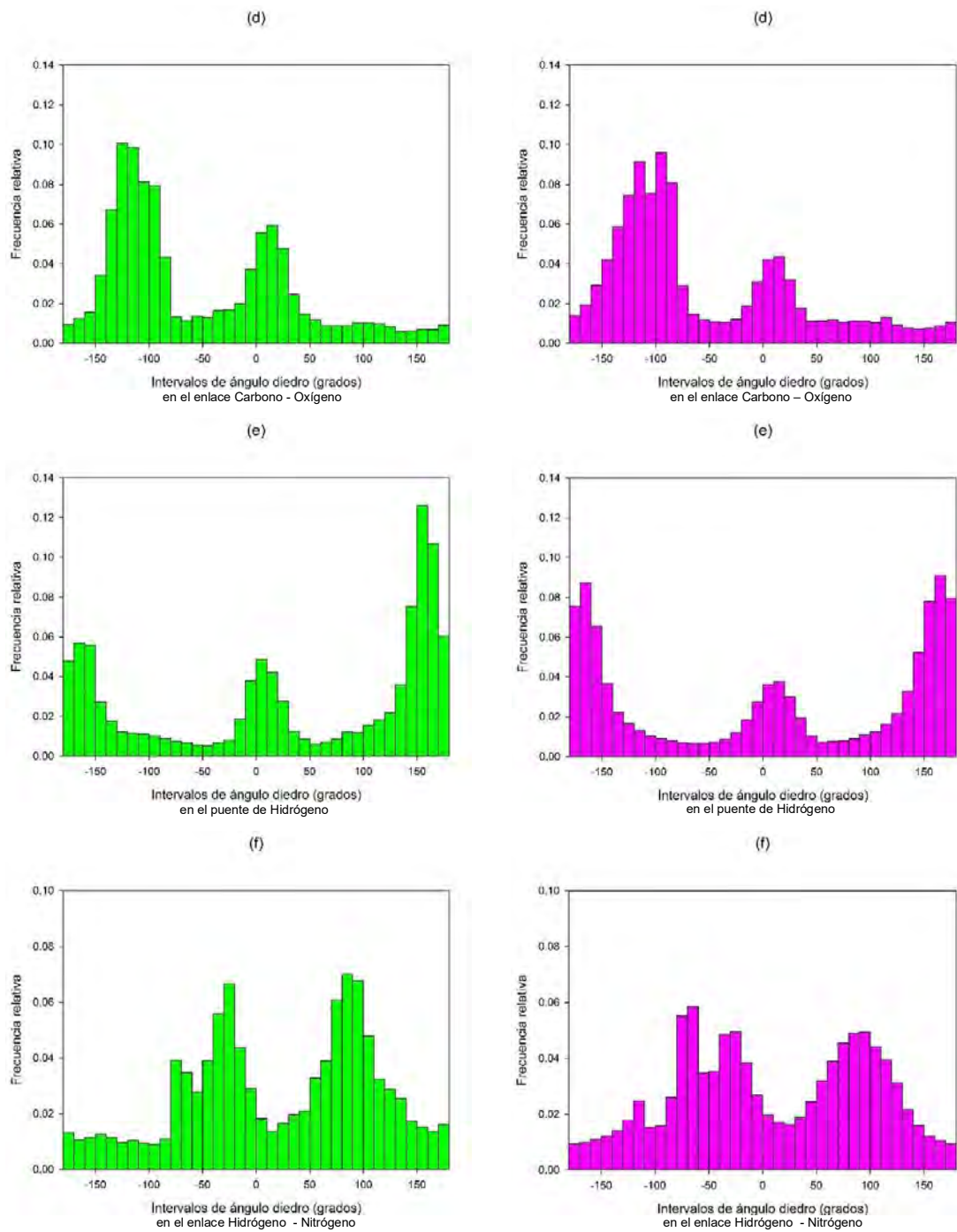


Figura 16. Distribuciones de frecuencias de los parámetros geométricos estudiados en los puentes de hidrógeno formados por la cadena principal de proteínas en los grupos ERX + AMBER99SB y ERMN + AMBER99SB. (a) la distancia hidrógeno-oxígeno, (b) el ángulo al átomo de hidrógeno, (c) el ángulo al átomo de oxígeno (d) el ángulo diedro que corresponde a la rotación del enlace carbono-oxígeno, (e) el ángulo diedro que corresponde a la rotación del puente de hidrógeno y (f) el ángulo diedro que corresponde a la rotación del enlace hidrógeno-nitrógeno.

Al comparar los dos grupos de estructuras en los cuales se reemplazaron los hidrógenos polares, se puede apreciar que existen las mismas diferencias que entre los grupos antes de reemplazar los hidrógenos polares: las posiciones de las frecuencias máximas y la mayor dispersión de las distribuciones en el grupo **ERMN**.

Adicionalmente, comparando los grupos antes y después del reemplazo de los hidrógenos polares, aunque se respeta el comportamiento general de las distribuciones, se puede observar un incremento considerable de frecuencia en los siguientes intervalos: (a) arriba de los 2.3 Å, (b) alrededor de los 95 grados, (c) alrededor de los 85 grados, (d) y (e) alrededor de los 20 grados y (f) alrededor de los -30 grados; efecto debido al empleo del campo de fuerza AMBER99SB en el remplazo de los hidrógenos polares.

Aunque con el reemplazo de los hidrógenos polares se obtiene información que no concuerda completamente con los datos experimentales, más del 90% de estos datos pueden ser mapeados en las nuevas distribuciones, razón que los hace factibles de emplear en la construcción de potenciales estadísticos.

VI. POTENCIALES ESTADÍSTICOS H-ERX Y H-ERMN

1. Generación de los potenciales

Con base en las distribuciones de frecuencias presentadas en el capítulo IV para los grupos **ERX + AMBER99SB** y **ERMN + AMBER99SB** y tomando en cuenta los intervalos con una frecuencia relativa superior a 0.0001, se generaron dos potenciales estadísticos: **H-ERX**, **H-ERMN**, respectivamente (**Fig. 2 - Anexo**). Para ello, se calculó la energía asociada a cada intervalo i , dado un parámetro geométrico x ($E_{(x,i)}$) de acuerdo con la estadística de Boltzmann:

$$E_{(x,i)} = -\ln\left(\frac{f^o_{(x,i)}}{f^r_{(x,i)}}\right) \quad (4)$$

La constante de Boltzmann y la temperatura absoluta no fueron incluidas debido a que únicamente determinan la amplitud de la función de energía, más no su comportamiento; $f^o_{(x,i)}$ es la frecuencia relativa del i -ésimo intervalo asociado al parámetro x , observada en la muestra, y $f^r_{(x,i)}$ es la frecuencia relativa del mismo intervalo en una distribución aleatoria, calculada como:

$$f^r_{(x,i)} = \frac{1}{N_x} \quad (5)$$

Donde N_x es el número de intervalos asociados al parámetro geométrico x : 52 para (a) y 36 para (b), (c), (d), (e) y (f). A los intervalos con frecuencias relativas menores a 0.0001 se les asoció un valor de energía arbitrario de 6.0, el entero superior inmediato a la energía asociada a este valor de frecuencia relativa (**Tabla 3, Fig. 2 – Anexo**).

Bajo este esquema, cada parámetro geométrico tiene una contribución independiente a la energía total de un puente de hidrógeno dado (E_{PH}), por lo que la energía asociada a dicho puente, es la suma de las contribuciones asociadas a cada parámetro, esto es:

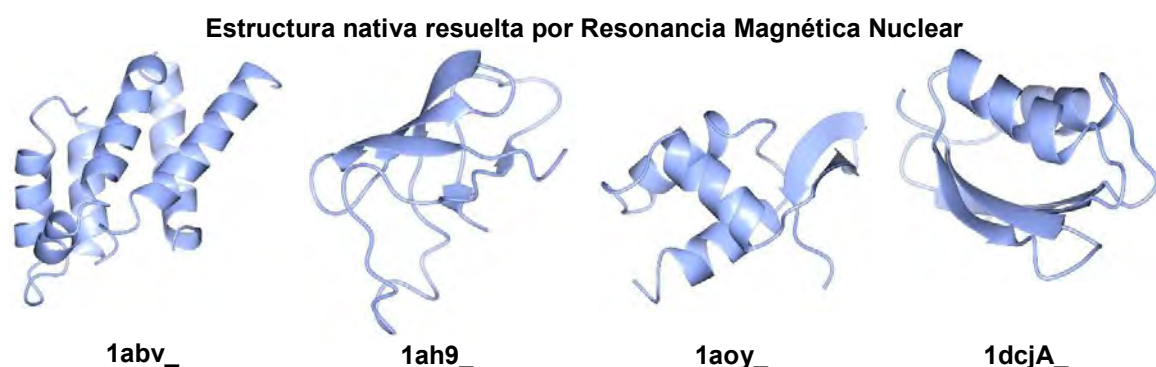
$$E_{PH} = E_{(a)} + E_{(b)} + E_{(c)} + E_{(d)} + E_{(e)} + E_{(f)} \quad (6)$$

Y la energía total asociada a una estructura es la suma de las energías asociadas a cada uno de los puentes de hidrógeno presentes en ella.

2. Pruebas de discriminación de señuelos

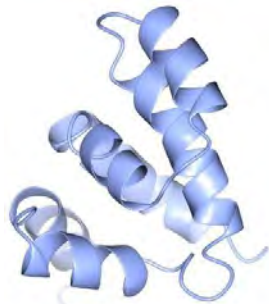
La habilidad de identificar la estructura nativa dentro de un conjunto de confórmers es una de las metas buscadas en el desarrollo de los potenciales estadísticos. Actualmente existen para ello distintos grupos de confórmers que han sido generados de diversas maneras, como simulaciones computacionales, variación manual de la estructura nativa o modelado por homología. Dentro de esos grupos de señuelos, los casos más realistas y retadores son aquellos generados por simulaciones moleculares, como lo es el caso de los señuelos generados por Zhang y colaboradores mediante I-TASSER (Iterative Threading ASSEMBly Refinement), usado para evaluar los potenciales generados en este trabajo.

El grupo de señuelos mencionado tiene como base la estructura de 56 subunidades de proteínas no homologas, variables tanto en longitud como en contenido de estructura secundaria y resueltas ya sea por difracción de rayos X o por resonancia magnética nuclear (**Fig. 17**). Tomando como base estas secuencias, se generaron estructuras cuyo esqueleto inicial fue modelado por hilvanado y sometido posteriormente a simulaciones de mecánica molecular a baja temperatura. Para cada una de las secuencias se obtuvieron un total de entre 12500 y 13200 modelos, de entre todas esas estructuras se seleccionaron por agrupamiento iterativo entre 300 y 500 confórmers para cada grupo, de tal manera que las estructuras resultantes fueran lo más parecidas a la estructura nativa, con una desviación cuadrática promedio menor a los 4 Å^{86,87}.





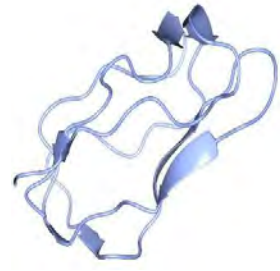
1egxA



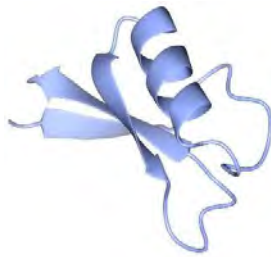
1fadA



1fo5A



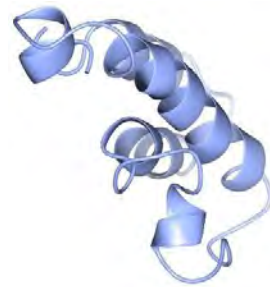
1gixA



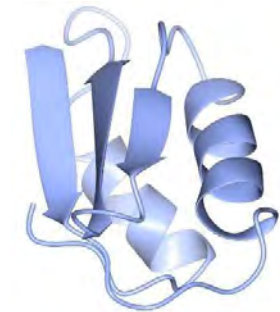
1gpt_



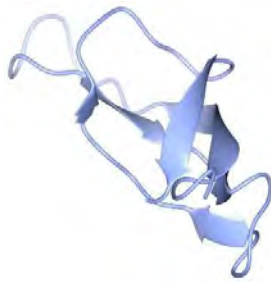
1itpA



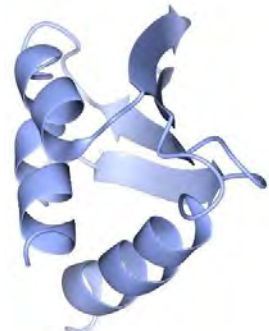
1kjs_



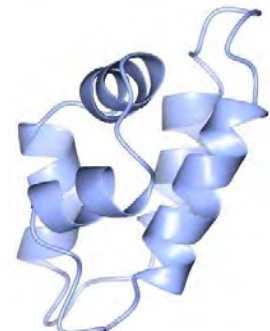
1kviA



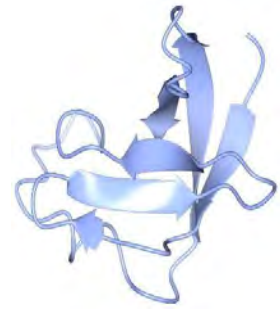
1ne3A



1o2fB_



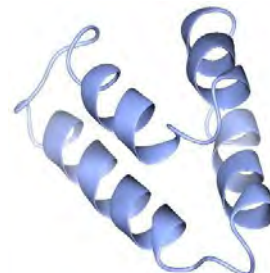
1of9A



1sro_

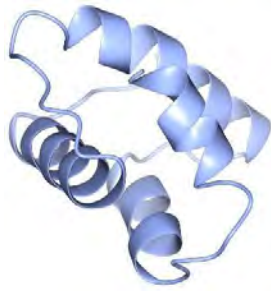


1tfi_

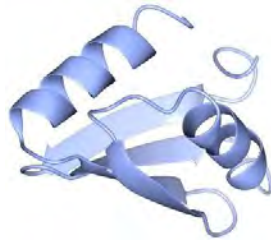


2cr7A

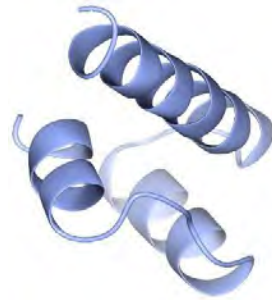
Estructura nativa resuelta por difracción de rayos X



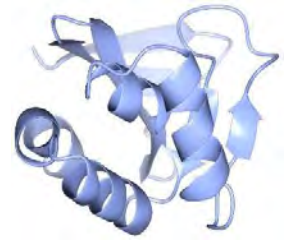
1af7_



1b4bA



1b72A



1bm8



1bq9A



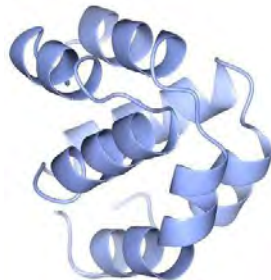
1cewl



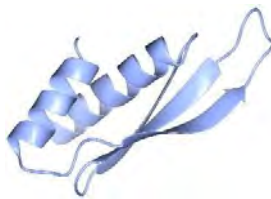
1cqkA



1csp_



1cy5A



1di2A_



1dtjA_



1g1cA



1gnuA



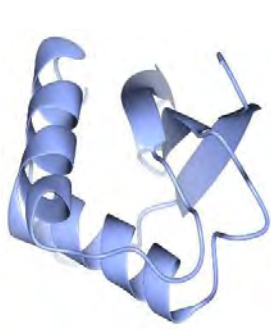
1gyvA



1hbkA



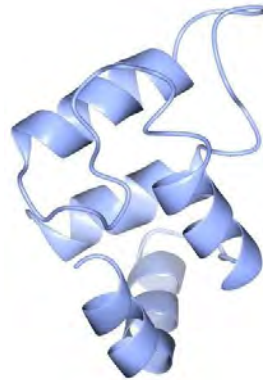
1jnuA



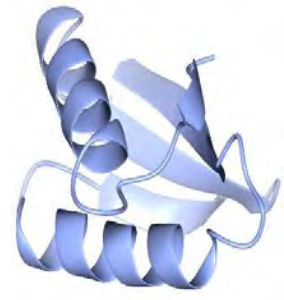
1mkyA3



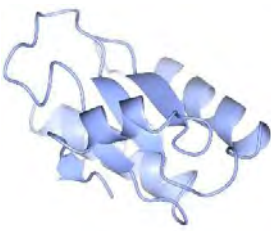
1mla_2



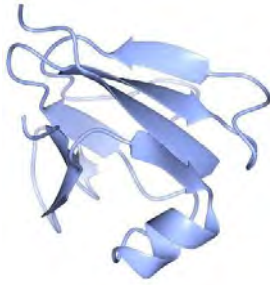
1mn8A



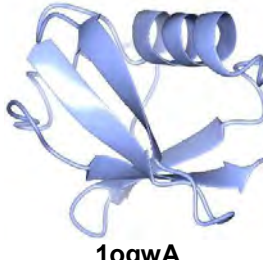
1n0uA4



1no5A



1npsA



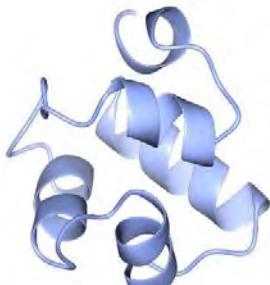
1ogwA



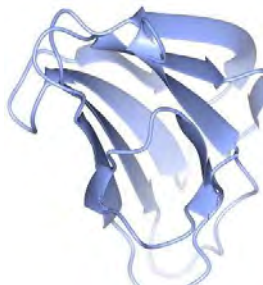
1orgA



1pgx_



1r69_



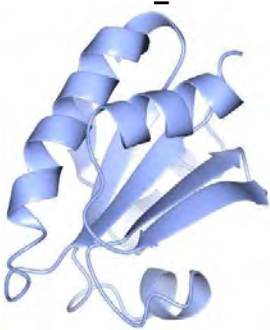
1sfp_



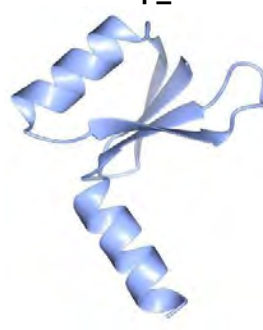
1shfA



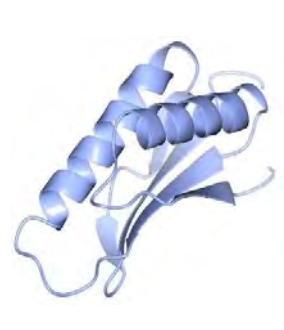
1ten_



1thx_



1tif_



1tig_

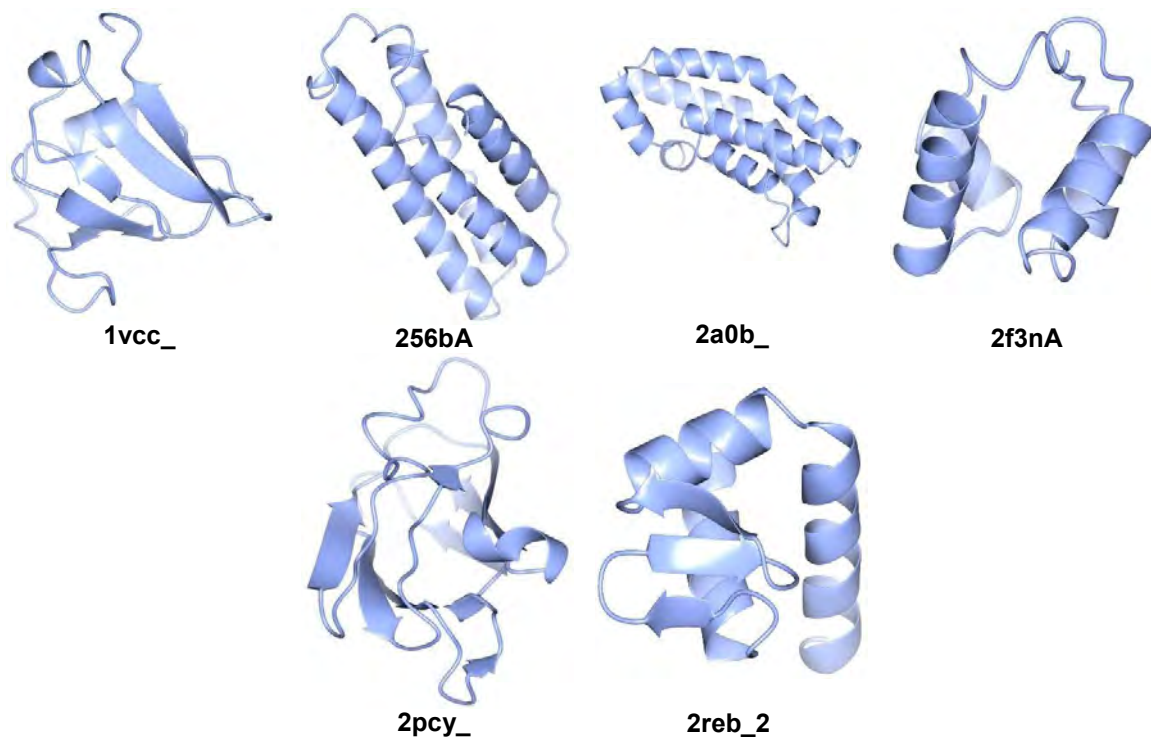


Figura 17. Estructura de las proteínas usadas para generar los conjuntos de señuelos I-TASSER. Representación en forma de listones. Todas las imágenes fueron obtenidas usando el software CCP4MG.

Se evaluó el desempeño de los potenciales **H-ERX** y **H-ERMN** en base a: 1) su capacidad para identificar la estructura nativa en cada uno de los grupos de señuelos y 2) el z-score de la estructura nativa en cada uno de los grupos. La primera acción se realiza asignando un valor de energía a cada estructura de un grupo dado, de manera que al ordenarlas de menor a mayor, la estructura nativa debería quedar en la posición número uno, mientras que la segunda se calcula según la ecuación **(3)** presentada en el capítulo I. Estos resultados se agruparon de acuerdo al método experimental por el cual fue resuelta la estructura nativa en cada uno de los grupos de señuelos: resonancia magnética nuclear (**Tabla 3**) y difracción de rayos X (**Tabla 4**). Los modelos de los señuelos a los cuales se le asoció la energía más baja (cuando no se trataba de la estructura nativa) se pueden consultar en el material anexo a este trabajo (**Fig. 3 y 4 – Anexo**).

Tabla 3. Desempeño de los potenciales H-ERX y H-ERMN en las pruebas de discriminación de señuelos en grupos cuya estructura nativa fue resuelta por resonancia magnética nuclear.

Grupo de señuelos	H-ERX			H-ERMN		
	Posición	Posición relativa	z-score	Posición	Posición relativa	z-score
1abv_	43	0.0814	-1.2853	12	0.0227	-2.1013
1ah9_	33	0.0645	-1.5596	35	0.0684	-1.6520
1aoy_	1	0.0019	-4.0043	2	0.0038	-2.9730
1dcjA_	1	0.0019	-5.2614	1	0.0019	-7.6265
1egxA	39	0.1102	-1.2748	1	0.0028	-3.9255
1fadA	5	0.0097	-3.5554	2	0.0039	-3.6952
1fo5A	26	0.0760	-1.4488	95	0.2778	-0.6555
1gjxA	233	0.4421	-0.1396	12	0.0228	-2.1401
1gpt_	4	0.0085	-3.1982	1	0.0021	-3.6211
1itpA	1	0.0019	-2.8567	1	0.0019	-4.8793
1kjs_	490	0.8909	1.1085	479	0.8709	0.7887
1kviA	17	0.0308	-2.0903	3	0.0054	-3.2674
1ne3A	36	0.0634	-1.6006	1	0.0018	-4.5844
1o2fB_	1	0.0020	-4.8479	1	0.0020	-4.8322
1of9A	504	0.9902	2.4519	301	0.5914	0.1765
1sro_	1	0.0019	-3.0299	1	0.0019	-3.6073
1tfi_	4	0.0117	-2.5449	22	0.0645	-1.6420
2cr7A	30	0.0554	-1.5856	115	0.2122	-0.7624
Estructuras nativas identificadas correctamente	5 / 18			7 / 18		
Estructuras nativas dentro del mejor 5% de estructuras	9 / 18			12 / 18		
z-score promedio	-2.0402			-2.8333		

Tabla 4. Desempeño de los potenciales H-ERX y H-ERMN en las pruebas de discriminación de señuelos en grupos cuya estructura nativa fue resuelta por difracción de rayos X.

Grupo de señuelos	H-ERX			H-ERMN		
	Posición	Posición relativa	z-score	Posición	Posición relativa	z-score
1af7__	1	0.0019	-5.6018	2	0.0038	-3.5436
1b4bA	1	0.0022	-6.4786	1	0.0022	-5.4267
1b72A	1	0.0019	-9.3010	1	0.0019	-6.8176
1bm8_	1	0.0030	-12.0942	1	0.0030	-11.3639
1bq9A	1	0.0017	-6.6952	1	0.0017	-6.8938
1cewl	1	0.0022	-3.8230	1	0.0022	-5.2810
1cqkA	1	0.0035	-5.6906	1	0.0035	-4.1682
1csp_	1	0.0032	-6.1014	1	0.0032	-5.5791
1cy5A	1	0.0036	-7.1602	1	0.0036	-5.5174
1di2A_	1	0.0027	-8.1228	1	0.0027	-6.1699
1dtjA_	1	0.0035	-3.9099	2	0.0070	-2.6128
1g1cA	1	0.0032	-6.8675	1	0.0032	-6.2216
1gnuA	1	0.0018	-8.2476	1	0.0018	-6.3250
1gyvA	1	0.0029	-7.5302	1	0.0029	-5.0581
1hbka	1	0.0033	-4.9776	4	0.0132	-2.6813
1jnuA	1	0.0037	-5.4016	1	0.0037	-4.3867
1mkyA3	1	0.0035	-8.6970	1	0.0035	-6.3974
1mla_2	1	0.0030	-8.6300	1	0.0030	-5.7519
1mn8A	1	0.0018	-9.9541	1	0.0018	-7.5005
1n0uA4	1	0.0033	-5.9186	1	0.0033	-4.5964
1no5A	1	0.0023	-6.1184	1	0.0023	-4.4764
1npsA	1	0.0021	-8.0121	1	0.0021	-6.2740
1ogwA_	1	0.0019	-7.0152	1	0.0019	-4.2688
1orgA	1	0.0023	-5.6018	3	0.0068	-4.8126
1pgx_	1	0.0018	-6.4786	1	0.0018	-6.6559
1r69_	1	0.0034	-9.3010	2	0.0068	-2.7410
1sfp_	1	0.0032	-12.0942	3	0.0097	-2.7941
1shfA	1	0.0019	-6.6952	1	0.0019	-6.1518
1ten_	1	0.0034	-3.8230	1	0.0034	-4.7436
1thx_	1	0.0033	-5.6906	1	0.0033	-4.1558
1tif_	1	0.0018	-6.1014	1	0.0018	-7.2366
1tig_	1	0.0018	-7.1602	1	0.0018	-4.6349
1vcc_	1	0.0018	-8.1228	1	0.0018	-7.0232
256bA	1	0.0020	-3.9099	2	0.0039	-4.6469
2a0b_	1	0.0035	-6.8675	1	0.0035	-5.4844
2f3nA	1	0.0021	-8.2476	1	0.0021	-6.1481
2pcy_	1	0.0023	-7.5302	1	0.0023	-5.9835
2reb_2	1	0.0018	-4.9776	1	0.0018	-4.0648
Estructuras nativas identificadas correctamente			38 / 38			32 / 38
Estructuras nativas dentro del mejor 5% de estructuras			38 / 38			38 / 38
z-score promedio			-7.1266			-5.3839

En los grupos de señuelos donde la estructura nativa fue determinada por resonancia magnética nuclear, ambos potenciales tienen un desempeño pobre, ya que identifican adecuadamente la estructura nativa en menos de la mitad de los casos. Comparando entre ambos potenciales, es claro que, para estos grupos de señuelos, el potencial **H-ERMN** tiene un mejor desempeño que el potencial **H-ERX**, ya que identifica correctamente la estructura nativa en dos grupos más que el potencial **H-ERX**, obtiene un mayor número de predicciones dentro del 5% de mejores estructuras, así como un z-score promedio menor.

En los grupos de señuelos donde la estructura nativa fue determinada por difracción de rayos X, ambos potenciales tienen un buen desempeño, sin embargo, para estos grupos de señuelos, el potencial **H-ERX** es claramente superior, al identificar correctamente la estructura nativa en todos los casos (ocho más que **H-ERMN**) y presentar un z-score promedio menor.


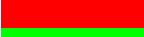



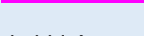
VII. POTENCIALES ESTADÍSTICOS H-ERX* Y H-ERMN*

1. Un nuevo análisis de datos

Con la finalidad de incrementar la cantidad de datos a analizar en el grupo **ERX + AMBER99SB**, se tomó una nueva muestra del PDB, semejante al grupo **ERX** descrito en el capítulo IV y denominado grupo **ERX***. Este grupo fue conformado por 1981 estructuras de proteínas monoméricas sin residuos modificados y con identidad de secuencia no mayor a 30%, resueltas por difracción de rayos X con una resolución de 2.0 Å o mejor (**Tabla 4 - Anexo**). A esta muestra se añadieron los hidrógenos polares de acuerdo a los parámetros del campo de fuerza AMBER99SB y se contabilizaron un total de 274,358 puentes de hidrógeno, cantidad comparable con los 237,718 encontrados en el grupo **ERMN + AMBER99SB**.

Con la finalidad de describir con más detalle los puentes de hidrógeno encontrados en ambos grupos: **ERX* + AMBER99SB** y **ERMN + AMBER99SB**, se decidió clasificarlos de acuerdo a la conectividad residuo aceptor (CO) → residuo donador (NH) que presentan con respecto a su secuencia (**Tabla 5**) y se obtuvieron las contribuciones de cada una de estas conectividades en las distribuciones de frecuencias de los seis parámetros geométricos establecidos (**Fig.18**).

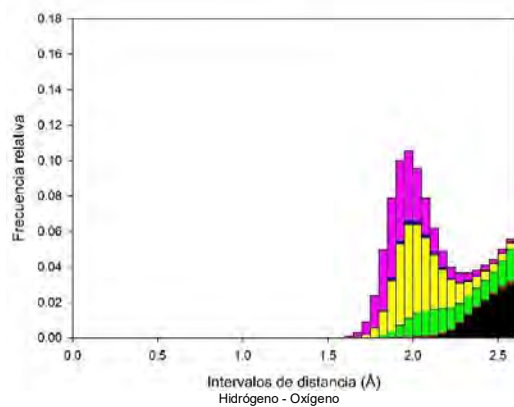
Tabla 5. Clasificación de los puentes de hidrógeno hallados en los grupos ERX* + AMBER99SB y ERMN + AMBER99SB de acuerdo a la conectividad residuo aceptor (CO) → residuo donador (NH) que presentan con respecto a su secuencia.

Conectividad residuo aceptor → residuo donador	Estructura secundaria que representa	Color en Figura 18	Porcentaje de puentes de hidrógeno	
			ERX* + AMBER99SB	ERMN + AMBER99SB
$i \rightarrow i$	+++		14.69	14.73
$i \rightarrow i + 2$	Giros γ		1.49	7.20
$i \rightarrow i + 3$	Giros β y hélices 3_{10}		17.76	20.19
$i \rightarrow i + 4$	Hélices α		35.15	34.36
$i \rightarrow i + 5$	Hélices π		1.59	1.39
Otras	Hojas β y otros		29.32	22.13

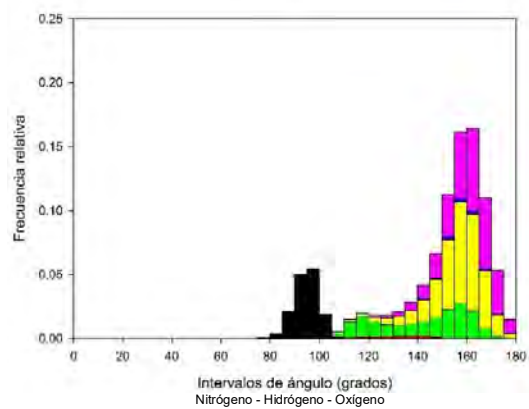
+++ Este tipo de conectividad representa a los puentes de hidrógeno en los que el grupo CO aceptor y NH donador pertenecen al mismo residuo, aunque no representan ningún tipo de estructura secundaria, son un porcentaje considerable del total y corresponden a una conformación posible.

ERX* + AMBER99SB

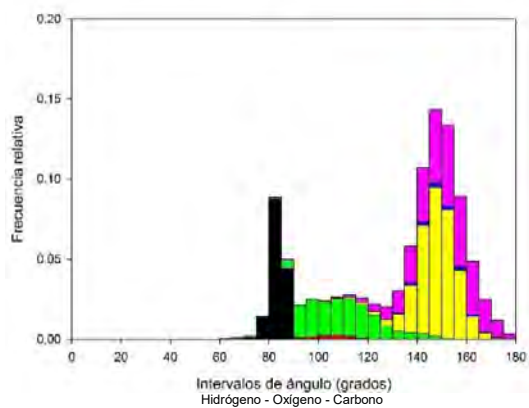
(a)



(b)

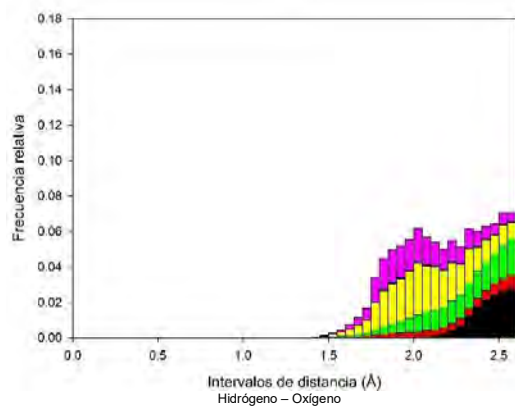


(c)

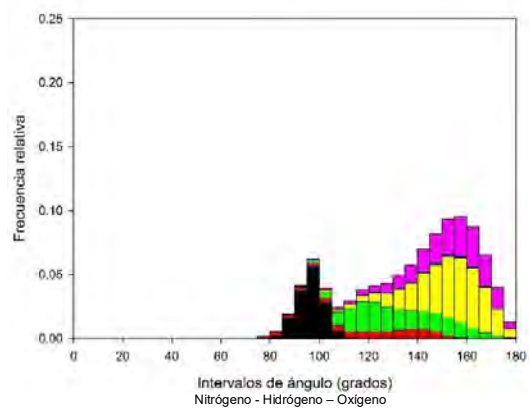


ERMN + AMBER99SB

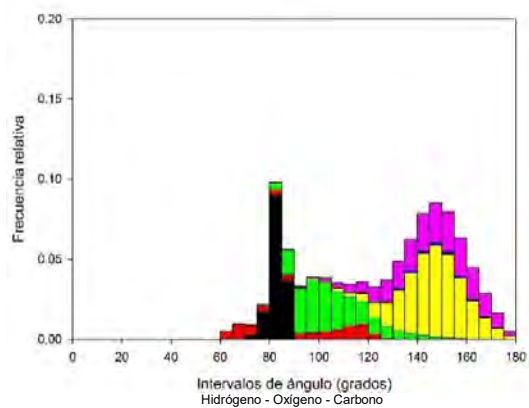
(a)



(b)



(c)



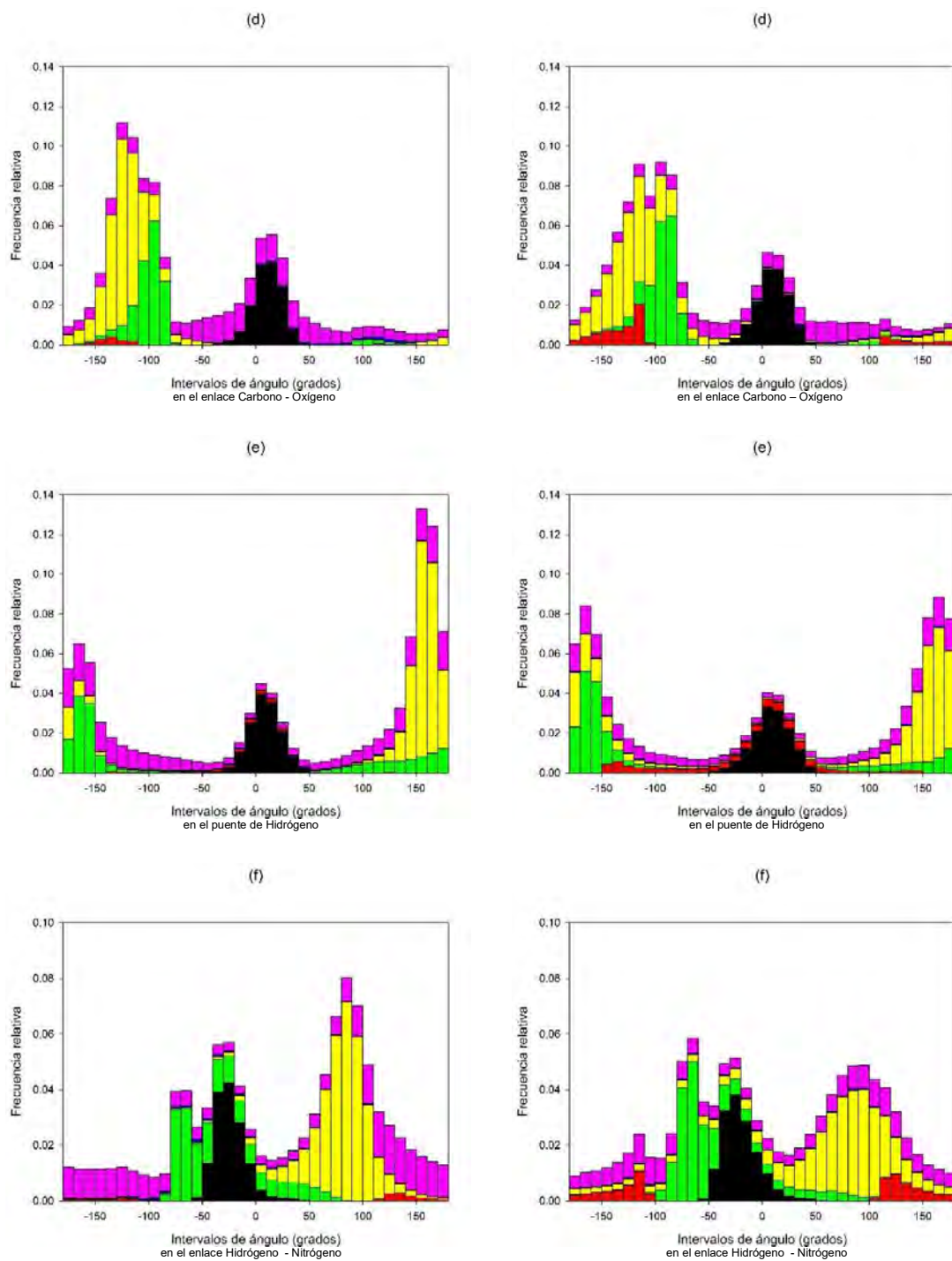


Figura 18. Contribución por conectividad de los puentes de hidrógeno en las distribuciones de frecuencias de los parámetros geométricos estudiados en los grupos ERX* + AMBER99SB y ERMN + AMBER99SB. (a) la distancia hidrógeno-oxígeno, (b) el ángulo al átomo de hidrógeno, (c) el ángulo al átomo de oxígeno (d) el ángulo diedro que corresponde a la rotación del enlace carbono-oxígeno, (e) el ángulo diedro que corresponde a la rotación del puente de hidrógeno y (f) el ángulo diedro que corresponde a la rotación del enlace hidrógeno-nitrógeno. Código de colores según tabla 5.

Cada una de las conectividades presenta una prevalencia distinta en ambos grupos, se puede destacar el hecho de que estructuras cerradas, como los giros y las hélices 3_{10} , son considerablemente más frecuentes en el grupo **ERMN + AMBER99SB**, de manera opuesta a estructuras más libres como las hojas β , las cuales son más frecuentes en el grupo **ERX* + AMBER99SB**, lo cual puede ser una consecuencia directa del muestreo realizado o de las condiciones experimentales en las que se determinaron las estructuras, ya que la mayor libertad de movimiento que tienen las moléculas en solución, permite observar de mejor forma estados que energéticamente pudieran no ser tan favorables, como las estructuras cerradas.

Por otra parte, el análisis de las contribuciones de cada una de las conectividades en las distribuciones de frecuencia de los parámetros geométricos estudiados muestra ligeras diferencias en cuanto a la altura y posiciones de las frecuencias máximas para cada conectividad, por lo que cada una de estas conectividades debería tener una contribución diferente en la estabilización de la estructura nativa de una proteína. Además, el incremento de frecuencia observado en los siguientes intervalos: (a) arriba de los 2.3 Å, (b) alrededor de los 95 grados, (c) alrededor de los 85 grados, (d) y (e) alrededor de los 20 grados y (f) alrededor de los -30 grados, señalados en el capítulo IV, se debe esencialmente a los puentes de hidrógeno cuya conectividad es $i \rightarrow i$, es decir, aquellos en los que el grupo CO aceptor y NH donador pertenecen al mismo residuo, los que a la vez pueden ser asociados al campo de fuerza AMBER99SB, empleado en la sustitución o adición de los hidrógenos polares.

2. Generación de los potenciales

Con base en las distribuciones de frecuencias presentadas en este capítulo para los grupos **ERX* + AMBER99SB** y **ERMN + AMBER99SB** y tomando en cuenta los intervalos con una frecuencia relativa superior a 0.0001, se generaron dos potenciales estadísticos: **H-ERX***, **H-ERMN***, respectivamente. Para ello, se calculó, para cada una de las conectividades (k), la energía asociada a cada intervalo i , dado un parámetro geométrico x ($E_{(k,x,i)}$) de acuerdo con la estadística de Boltzmann:

$$E_{(k,x,i)} = -\ln\left(\frac{f^o_{(k,x,i)}}{f^r_{(k,x,i)}}\right) \quad (7)$$

De nueva cuenta, la constante de Boltzmann y la temperatura absoluta no fueron incluidas; $f^o_{(k,x,i)}$ es la frecuencia relativa del i -ésimo intervalo asociado al parámetro x de la conectividad k , observada en la muestra y $f^r_{(k,x,i)}$ es la frecuencia relativa del mismo intervalo en una distribución aleatoria, calculada como:

$$f^r_{(k,x,i)} = \left(\frac{1}{N_k}\right)\left(\frac{1}{N_x}\right) \quad (8)$$

Donde N_k es el número de conectividades posibles, seis, y N_x es el número de intervalos asociados al parámetro geométrico x : 52 para (a) y 36 para (b), (c), (d), (e) y (f). A los intervalos con frecuencias relativas menores a 0.0001 se les asoció un valor de energía arbitrario de 4.0, el entero superior inmediato a la energía asociada a este valor de frecuencia relativa (**Tabla 5 y 6, Fig. 5 – Anexo**).

Bajo este esquema, cada parámetro geométrico tiene una contribución independiente a la energía total de un puente de hidrógeno, la cual depende a su

vez de la conectividad de dicho puente (E^k_{PH}), por lo que la energía asociada a dicho puente, es la suma de las contribuciones asociadas a cada parámetro, dada la conectividad, esto es:

$$E^k_{PH} = E^k_{(a)} + E^k_{(b)} + E^k_{(c)} + E^k_{(d)} + E^k_{(e)} + E^k_{(f)} \quad (9)$$

Y la energía total asociada a una estructura es la suma de las energías asociadas a cada uno de los puentes de hidrógeno presentes en ella. De esta manera, en lugar de utilizar un potencial estadístico para cada parámetro geométrico, se pueden emplear seis distintos, dependiendo de la conectividad que presente el puente de hidrógeno.

3. Pruebas de discriminación de señuelos

Para evaluar los potenciales **H-ERX*** y **H-ERMN*** se emplearon pruebas de discriminación de confórmeros utilizando los mismos 56 grupos de señuelos generados por I-TASSER descritos en el capítulo V. Al igual que en sus potenciales homólogos, **H-ERX** y **H-ERMN**, se evaluó el desempeño de los potenciales **H-ERX*** y **H-ERMN*** con base en: 1) su capacidad para identificar la estructura nativa en cada uno de los grupos de señuelos y 2) el z-score de la estructura nativa en cada uno de los grupos. Estos resultados se agruparon de acuerdo al método experimental por el cual fue resuelta la estructura nativa en cada uno de los grupos de señuelos: resonancia magnética nuclear (**Tabla 6**) y difracción de rayos X (**Tabla 7**). Los modelos de los señuelos a los cuales se le asoció la energía más baja (cuando no se trataba de la estructura nativa) se pueden consultar en el material anexo (**Fig. 6 y 7 – Anexo**).

Tabla 6. Desempeño de los potenciales H-ERX* y H-ERMN* en las pruebas de discriminación de señuelos en grupos cuya estructura nativa fue resuelta por resonancia magnética nuclear.						
Grupo de señuelos	H-ERX*			H-ERMN*		
	Posición	Posición relativa	z-score	Posición	Posición relativa	z-score
1abv_	12	0.0227	-2.3356	4	0.0076	-2.7606
1ah9_	206	0.4023	-0.2459	164	0.3203	-0.4160
1aoy_	1	0.0019	-3.6890	3	0.0056	-2.6932
1dcjA_	1	0.0019	-4.8945	1	0.0019	-5.6824
1egxA	1	0.0028	-3.9089	1	0.0028	-4.1571
1fadA	9	0.0174	-3.2954	3	0.0058	-3.2449
1fo5A	45	0.1316	-1.0775	54	0.1579	-1.0472
1gjxA	5	0.0095	-2.8298	4	0.0076	-2.7973
1gpt_	1	0.0021	-3.4177	1	0.0021	-3.5241
1itpA	1	0.0019	-4.6766	1	0.0019	-5.9998
1kjs_	515	0.9364	1.7440	500	0.9091	1.0997
1kviA	3	0.0054	-2.6939	2	0.0036	-3.5617
1ne3A	1	0.0018	-3.3533	1	0.0018	-3.7524
1o2fB_	1	0.0020	-5.0830	1	0.0020	-4.1215
1of9A	431	0.8468	0.8819	262	0.5147	-0.0644
1sro_	43	0.0832	-1.4716	6	0.0116	-2.2962
1tfi_	1	0.0029	-3.7173	2	0.0059	-2.5817
2cr7A	33	0.0609	-1.5268	63	0.1162	-1.0989
Estructuras nativas identificadas correctamente	8 / 18			6 / 18		
Estructuras nativas dentro del mejor 5% de estructuras	12 / 18			13 / 18		
z-score promedio	-2.5444			-2.7055		

Tabla 7. Desempeño de los potenciales H-ERX* y H-ERMN* en las pruebas de discriminación de señuelos en grupos cuya estructura nativa fue resuelta por difracción de rayos X.

Grupo de señuelos	H-ERX*			H-ERMN*		
	Posición	Posición relativa	z-score	Posición	Posición relativa	z-score
1af7__	2	0.0038	-4.7835	2	0.0038	-3.0435
1b4bA	1	0.0022	-7.0470	1	0.0022	-5.0350
1b72A	1	0.0019	-7.5167	1	0.0019	-6.5052
1bm8_	1	0.0030	-11.4303	1	0.0030	-10.8764
1bq9A	1	0.0017	-5.3729	1	0.0017	-4.5587
1cewl	1	0.0022	-6.2774	1	0.0022	-5.9580
1cqkA	1	0.0035	-5.9776	1	0.0035	-4.2183
1csp_	1	0.0032	-6.4483	1	0.0032	-5.5857
1cy5A	1	0.0036	-6.5232	1	0.0036	-5.2468
1di2A_	1	0.0027	-7.2699	1	0.0027	-5.7598
1dtjA_	1	0.0035	-3.7275	2	0.0070	-2.4289
1g1cA	1	0.0032	-7.5009	1	0.0032	-5.9756
1gnuA	1	0.0018	-7.0906	1	0.0018	-5.7271
1gyvA	1	0.0029	-7.1393	1	0.0029	-4.6795
1hbka	1	0.0033	-4.9004	1	0.0132	-2.7174
1jnuA	1	0.0037	-4.9058	1	0.0037	-4.4255
1mkyA3	1	0.0035	-6.8813	1	0.0035	-6.0749
1mla_2	1	0.0030	-7.4609	1	0.0030	-5.9195
1mn8A	1	0.0018	-8.3067	1	0.0018	-6.6654
1n0uA4	1	0.0033	-5.6996	1	0.0033	-3.9713
1no5A	1	0.0023	-5.5931	1	0.0023	-4.4861
1npsA	1	0.0021	-7.3013	1	0.0021	-5.4589
1ogwA_	1	0.0019	-4.8114	1	0.0019	-3.6235
1orgA	1	0.0023	-6.2839	3	0.0068	-4.4251
1pgx_	1	0.0018	-8.6061	1	0.0018	-6.8363
1r69_	1	0.0034	-3.4877	3	0.0068	-2.1240
1sfp_	1	0.0032	-6.5764	1	0.0097	-4.2455
1shfA	1	0.0019	-8.2721	1	0.0019	-6.6269
1ten_	1	0.0034	-7.5113	1	0.0034	-5.7563
1thx_	1	0.0033	-5.6133	1	0.0033	-3.5880
1tif_	1	0.0018	-7.7918	1	0.0018	-6.4721
1tig_	1	0.0018	-6.3085	1	0.0018	-3.9623
1vcc_	1	0.0018	-7.8710	1	0.0018	-6.2185
256bA	1	0.0020	-6.3533	3	0.0039	-3.8392
2a0b_	1	0.0035	-6.0424	1	0.0035	-4.7138
2f3nA	1	0.0021	-6.0661	1	0.0021	-4.5661
2pcy_	1	0.0023	-7.5558	1	0.0023	-5.8255
2reb_2	1	0.0018	-6.4901	1	0.0018	-4.7559
Estructuras nativas identificadas correctamente			37 / 38			33 / 38
Estructuras nativas dentro del mejor 5% de estructuras			38 / 38			38 / 38
z-score promedio			-6.5999			-5.1312

Nuevamente, en los grupos de señuelos donde la estructura nativa fue determinada por resonancia magnética nuclear, ambos potenciales tienen un desempeño pobre, ya que la estructura nativa es identificada adecuadamente en menos de la mitad de los casos. Comparando entre ambos potenciales, aunque la diferencia no es tan grande, para estos grupos de señuelos, el potencial **H-ERMN*** tiene un desempeño ligeramente mejor que el potencial **H-ERX***, ya que, aunque éste identifica correctamente la estructura nativa en dos grupos menos que el potencial **H-ERX***, se obtiene un mayor número de predicciones dentro del 5% de mejores estructuras, así como un z-score promedio menor.

En los grupos de señuelos donde la estructura nativa fue determinada por difracción de rayos X, ambos potenciales tienen un buen desempeño, sin embargo, para estos grupos de señuelos, el potencial **H-ERX*** es claramente superior, al identificar correctamente la estructura nativa en 37 casos (cuatro más que **H-ERMN***) y presentar un z-score promedio menor.

Comparando los potenciales **H-ERX*** y **H-ERMN*** con sus predecesores, **H-ERX** y **H-ERMN** respectivamente, se puede notar una ligera mejoría en su desempeño en los grupos de señuelos donde la estructura nativa fue determinada por resonancia magnética nuclear, ya que en ambos casos se incrementó el número de predicciones dentro del 5% de mejores estructuras y se disminuyó el z-score promedio. En los grupos de señuelos donde la estructura nativa fue determinada por difracción de rayos X se observa el comportamiento opuesto en ambos casos, ya que, a pesar de que en todos los casos se obtienen predicciones dentro del 5% de mejores estructuras, el z-score promedio aumenta ligeramente.

VIII. COMPARACIÓN CON OTROS POTENCIALES

Finalmente se compararon los resultados obtenidos por los cuatro potenciales generados con los resultados obtenidos por otros potenciales probados en los mismos grupos de señuelos: como DOPE⁸⁸, DFIRE¹⁵, RWplus⁸⁶, DBNI⁸⁹, OPUS-PSP⁹⁰, Multi_well⁹¹, DOKB⁹², GOAP⁹³ y ROTAS⁹⁴ (**Tabla 8**). Las bases de estos potenciales se describen brevemente a continuación:

DOPE (Discrete Optimized Protein Energy) es un potencial dependiente de la distancia, derivado de la unión de las densidades de probabilidad de las coordenadas cartesianas de los átomos de una proteína y de las distancias entre ellos.

DFIRE (Distance-scaled, Finite Ideal-gas Reference State) es un potencial de pares de contacto de todos los átomos de una proteína, que usa como estado de referencia una distribución uniforme de puntos en una esfera finita, semejante a la distribución de un gas ideal en un espacio finito.

RWplus (Random walk) es un potencial de pares de contacto, dependiente de la orientación de las cadenas laterales de los residuos que constituyen una proteína, que usa como estado de referencia una cadena ideal generada por un movimiento aleatorio.

DBNI (Delaunay-Based Nonlocal Interactions) es un potencial de pares de contacto basado en las interacciones entre 167 tipos de átomos, separados entre sí por más de 5 aminoácidos, y determinadas en base a la triangulación de Delauney.

OPUS-PSP es un potencial de pares de contacto, dependiente de la orientación de los aminoácidos en contacto, representados como bloques.

Multi_well es un potencial de pares de contacto entre los átomos presentes en la estructura secundaria de una proteína que pretende identificar la topología nativa de dicha estructura entre todas las posibilidades existentes.

DOKB (Distance and Orientation dependent energy function of amino acid Key Blocks) es un potencial de pares de contacto, dependiente de la distancia y orientación de los aminoácidos en contacto, representados como bloques.

GOAP (Generalized Orientation-dependent All-atom Potential) es un potencial de pares de contacto que depende de la orientación relativa de los planos asociados a cada átomo pesado en los pares interactuantes.

ROTAS (ROTamer-dependent Atomic Statistical potential) es un potencial que toma en cuenta el papel de los estados rotaméricos de los residuos de aminoácidos en la especificidad de una interacción atómica dada.

Tabla 8. Desempeño de distintos potenciales en los grupos de señuelos de I-TASSER.

Potencial	Estructuras nativas identificadas (z- score promedio)
DOPE	30 (-2.18)
DFIRE	48 (-5.03)
RWplus	56 (-5.77)
DBNI	42 (-3.63)
OPUS-PSP	49 (-5.40)
Multi_well	16 (n.r.)
DOKB	53 (n.r.)
GOAP	48 (-5.81)
ROTAS	49 (-7.31)
H-ERX	43 (-5.57)
H-ERMN	39 (-4.56)
H-ERX*	45 (-5.35)
H-ERMN*	39 (-4.33)

n.r. = no reportada

Se puede notar que el mejor de los potenciales generados en este trabajo es el **H-ERX***, al ser el que identifica correctamente la mayor cantidad de estructuras nativas. Al compararlo con el desempeño de los otros potenciales, supera en el número de estructuras nativas identificadas a tres de los nueve presentados, pero en cuanto al z-score promedio, es comparable con los potenciales de mejor desempeño.

IX. CONCLUSIONES

Los resultados mostrados en el capítulo IV mostraron que, describiendo la estructura de las proteínas en base a la geometría de los puentes de hidrógeno formados por la cadena principal de estas, las estructuras resueltas por difracción de rayos X difieren considerablemente de aquellas resueltas por resonancia magnética nuclear. Se demostró también que el reemplazo de los hidrógenos polares en las estructuras experimentales, mediante el campo de fuerza AMBER99SB, modifica considerablemente las estructuras bajo el nivel de descripción usado en este trabajo, aunque permite mapear los datos experimentales en un 91% para las estructuras resueltas por difracción de rayos X y en un 98% para las estructuras resueltas por resonancia magnética nuclear.

Las pruebas de discriminación de señuelos realizadas con los potenciales estadísticos basados en la geometría de los puentes de hidrógeno presentes en proteínas, presentados en el capítulo V, mostraron que este tipo de potenciales se desempeñan muy bien en grupos de señuelos donde la estructura nativa fue determinada por difracción de rayos X, especialmente el potencial construido a partir de estructuras resueltas por este mismo método experimental, que identifica correctamente la estructura nativa en todos los casos probados. Esto se debe a que los parámetros geométricos estudiados presentan tendencias más claras para este grupo de proteínas en comparación con aquellas determinadas por resonancia magnética nuclear, lo que en cierta medida facilita la tarea de identificar una buena estructura.

En grupos de señuelos donde la estructura nativa fue determinada por resonancia magnética nuclear, el desempeño de este tipo de potenciales es pobre, donde el potencial construido a partir de estructuras resueltas por este mismo método experimental tiene el mejor desempeño al identificar correctamente la estructura nativa en 7 de los 18 casos probados.

Lo anterior permite establecer que el desempeño de este tipo de potenciales estadísticos en un grupo de señuelos dado, está influido por el método experimental con el que fueron determinadas las estructuras usadas para su construcción y que

este tipo de potenciales es especialmente útil en pruebas de discriminación de señuelos en grupos donde la estructura nativa fue determinada por difracción de rayos X.

Las pruebas de discriminación de señuelos realizadas con los potenciales estadísticos basados en la geometría de los puentes de hidrógeno en proteínas que contemplan la conectividad de dichos puentes, presentados en el capítulo VI, confirman lo establecido en el párrafo anterior y por otro lado, demuestran que no existe una gran mejora al comparar su desempeño con los potenciales presentados en el capítulo V. Esto sugiere que la clasificación de los puentes de hidrógeno de acuerdo a su conectividad (o tipo de estructura secundaria que forman) no es crucial para tener una buena descripción al emplear potenciales estadísticos como los presentados en este trabajo, lo cual simplifica considerablemente (6 veces, en este caso particular) el cálculo de la energía para una estructura dada.

De acuerdo con los resultados presentados en el capítulo VII, los potenciales presentados en este trabajo distan de tener el mejor desempeño en cuanto al número de estructuras nativas identificadas correctamente, sin embargo es de resaltar el hecho de que su construcción es más simple que la del resto de potenciales comparados, ya que se emplea un solo tipo de interacción atómica: los puentes de hidrógeno formados por la cadena principal de proteínas. Este hecho limita en cierta forma al potencial para diferenciar cambios conformacionales en secciones sin estructura secundaria de una proteína o modificaciones en las cadenas laterales de los aminoácidos. Aun así, se obtienen z-score promedios bastante buenos, lo cual indica que el nivel de descripción de la estructura de proteínas empleado en este trabajo permite diferenciar ampliamente la estructura nativa de una proteína dentro de un conjunto de confórmeros.

X. PERSPECTIVAS

Los hallazgos presentados en este trabajo permiten sugerir algunos estudios que podrían ayudar a mejorar el desempeño de los potenciales estadísticos, entre ellos se pueden mencionar los siguientes:

Ya que las estructuras de proteínas determinadas por difracción de rayos X difieren considerablemente de aquellas determinadas por resonancia magnética nuclear a nivel de la geometría de los puentes de hidrógeno formados por la cadena principal, sería conveniente buscar y seleccionar un nivel de descripción que se asemeje más entre los dos grupos de estructuras.

Dado que el uso de los parámetros establecidos de acuerdo a un campo de fuerza (AMBER99SB) para la adición de los hidrógenos polares en la estructura de una proteína no es del todo comparable con las observaciones experimentales, sería de gran utilidad el reestablecer los parámetros usados para la adición de estos hidrógenos.

El desempeño de los potenciales estadísticos generados en este trabajo podría verse mejorado al incluir términos que consideren los puentes de hidrógeno en donde participan las cadenas laterales de los residuos de aminoácidos, así como la inclusión de parámetros geométricos para describir otro tipo de interacciones presentes en proteínas, como las interacciones de van der Waals.

Por otro lado, aunque los seis parámetros geométricos usados en la descripción de los puentes de hidrógeno lo describen completamente, sería conveniente determinar si todos ellos son necesarios para un buen desempeño del potencial o si es posible omitir algunos de ellos, para simplificar los cálculos de energía.

XI. REFERENCIAS

1. Boyle, J. *Lehninger principles of biochemistry (4th ed.): Nelson, D., and Cox, M. Biochemistry and Molecular Biology Education* **33**, (2005).
2. Berg, J. M., Tymoczko, J. L. & Stryer, L. *Biochemistry. W H Freeman New York.*, (2002).
3. Fromm, H. & Hargrove, M. *Essentials of Biochemistry. Annual Review of Biochemistry* (2012). doi:10.1146/annurev.bi.74.052705.100003
4. Buxbaum, E. *Fundamentals of Protein Structure and Function.* (Springer US, 2007). doi:10.1007/978-0-387-68480-2
5. Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science* **181**, 223–230 (1973).
6. Kolata, G. Trying to crack the second half of the genetic code. *Science* **233**, 1037–1039 (1986).
7. Burley, S. K. *et al.* Structural genomics: beyond the Human Genome Project. *Nat. Genet.* **23**, 151–157 (1999).
8. Case, D. A. *et al.* The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **26**, 1668–1688 (2005).
9. Brooks, B. R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217 (1983).
10. Sippl, M. J. Calculation of conformational ensembles from potentials of mean force. *J. Mol. Biol.* **213**, 859–883 (1990).
11. Tanaka, S. & Scheraga, H. A. Medium- and Long-Range Interaction Parameters between Amino Acids for Predicting Three-Dimensional Structures of Proteins. *Macromolecules* **9**, 945–950 (1976).
12. DeWitte, R. S. & Shakhnovich, E. I. SMOG: de Novo Design Method Based on Simple, Fast, and Accurate Free Energy Estimates. 1. Methodology and Supporting Evidence. *J. Am. Chem. Soc.* **118**, 11733–11744 (1996).
13. Lu, H. & Skolnick, J. A distance-dependent atomic knowledge-based potential for improved protein structure selection. *Proteins Struct. Funct. Genet.* **44**, 223–232 (2001).
14. Samudrala, R. & Moult, J. An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction. *J. Mol. Biol.* **275**, 895–916 (1998).
15. Zhou, H. & Zhou, Y. Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. *Protein Sci.* **11**, 2714–2726 (2009).
16. Bordo, D. & Argos, P. The Role of Side-chain Hydrogen Bonds in the Formation and Stabilization of Secondary Structure in Soluble Proteins. *J. Mol. Biol.* **243**, 504–519 (1994).
17. Fersht, A. *Enzyme Structure and Mechanism. New York - N.Y. : Freeman* (1985).
18. Gorres, K. L. & Raines, R. T. Prolyl 4-hydroxylase. *Crit. Rev. Biochem. Mol. Biol.* **45**, 106–124 (2010).
19. Van Slyke, D. D. & Hiller, A. An Unidentified Base among the Hydrolytic Products of Gelatin. *Proc. Natl. Acad. Sci.* **7**, 185–186 (1921).

20. Hardy, M. F., Harris, C. I., Perry, S. V & Stone, D. Occurrence and formation of the N(□)-methyl-lysines in myosin and the myofibrillar proteins. *Biochem. J.* **120**, 653–660 (1970).
21. Stenflo, J. & Suttie, J. W. Vitamin K-Dependent Formation of γ -Carboxyglutamic Acid. *Annu. Rev. Biochem.* **46**, 157–172 (1977).
22. Anwar, R. A. & Oda, G. The biosynthesis of desmosine and isodesmosine. *J. Biol. Chem.* **241**, 4638–4641 (1966).
23. Johansson, L., Gafvelin, G. & Arnér, E. S. J. Selenocysteine in proteins—properties and biotechnological use. *Biochim. Biophys. Acta - Gen. Subj.* **1726**, 1–13 (2005).
24. Bronk, J. R. & Fisher, R. B. The role of ornithine and citrulline in urea synthesis. *Biochem. J.* **64**, 111–118 (1956).
25. Pauling, L., Corey, R. B. & Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci.* **37**, 205–211 (1951).
26. Karpen, M. E., De Haseth, P. L. & Neet, K. E. Differences in the amino acid distributions of 3 10⁻helices and α -helices. *Protein Sci.* **1**, 1333–1342 (1992).
27. Cooley, R. B., Arp, D. J. & Karplus, P. A. Evolutionary Origin of a Secondary Structure: π -Helices as Cryptic but Widespread Insertional Variations of α -Helices That Enhance Protein Functionality. *J. Mol. Biol.* **404**, 232–246 (2010).
28. Rose, G. D., Glerasch, L. M. & Smith, J. A. in 1–109 (1985). doi:10.1016/S0065-3233(08)60063-7
29. Whitford, D. *Proteins: Structure and Function*. (2005).
30. Levinthal, C. Are there pathways for protein folding? *J. Chim. Phys. Physico-Chimie Biol.* **65**, 44–45 (1968).
31. Freddolino, P. L., Liu, F., Gruebele, M. & Schulten, K. Ten-microsecond molecular dynamics simulation of a fast-folding WW domain. *Biophys. J.* **94**, L75–L77 (2008).
32. Hausrath, A. C. A kinetic theory of tertiary contact formation coupled to the helix-coil transition in polypeptides. *J. Chem. Phys.* **125**, 84909 (2006).
33. Bartlett, A. I. & Radford, S. E. An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* **16**, 582–588 (2009).
34. Brockwell, D. J. & Radford, S. E. Intermediates: ubiquitous species on folding energy landscapes? *Current Opinion in Structural Biology* **17**, 30–37 (2007).
35. Onuchic, J. N. & Wolynes, P. G. Theory of protein folding. *Curr. Opin. Struct. Biol.* **14**, 70–75 (2004).
36. Hartl, F. U. & Hayer-Hartl, M. Converging concepts of protein folding in vitro and in vivo. *Nat. Struct. Mol. Biol.* **16**, 574–581 (2009).
37. Whitford, D. *Proteins: Structure and Function*. (John Wiley & Sons, 2005).
38. KENDREW, J. C. *et al.* A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **181**, 662–666 (1958).
39. Amunts, A., Brown, A., Toots, J., Scheres, S. H. W. & Ramakrishnan, V. The structure of the human mitochondrial ribosome. *Science* **348**, 95 LP-98 (2015).

40. Khatter, H., Myasnikov, A. G., Natchiar, S. K. & Klaholz, B. P. Structure of the human 80S ribosome. *Nature* **520**, 640–645 (2015).
41. Levitt, M. & Warshel, A. Computer simulation of protein folding. *Nature* **253**, 694–698 (1975).
42. Levitt, M. A simplified representation of protein conformations for rapid simulation of protein folding. *J. Mol. Biol.* **104**, 59–107 (1976).
43. Contreras-Moreira, B., Ezkurdia, I., Tress, M. L. & Valencia, A. Empirical limits for template-based protein structure prediction: the CASP5 example. *FEBS Lett.* **579**, 1203–1207 (2005).
44. Bajorath, J., Seyama, K., Nonoyama, S., Ochs, H. D. & Aruffo, A. Classification of mutations in the human CD40 ligand, gp39, that are associated with X-linked hyper IgM syndrome. *Protein Sci.* **5**, 531–534 (2008).
45. Murzin, A. G. Progress in protein structure prediction. *Nat. Struct. Biol.* **8**, (2001).
46. Schonbrun, J., Wedemeyer, W. J. & Baker, D. Protein structure prediction in 2002. *Curr. Opin. Struct. Biol.* **12**, 348–354 (2002).
47. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Research* **28**, 235–242 (2000).
48. Jones, D. T. Progress in protein structure prediction. *Curr. Opin. Struct. Biol.* **7**, 377–387 (1997).
49. Morales, L. B., Garduno, J. R. & Romero, D. Applications of simulated annealing to the multiple-minima problem in small peptides. *J. Biomol. Struct. Dyn.* **8**, 721–735 (1991).
50. Morales, L. B., Garduño-Juárez, R., Aguilar-Alvarado, J. M. & Riveros-Castro, F. J. A parallel tabu search for conformational energy optimization of oligopeptides. *J. Comput. Chem.* **21**, 147–156 (2000).
51. Garduno-Juarez, R. & Morales, L. B. A genetic algorithm with conformational memories for structure prediction of polypeptides. *J. Biomol. Struct. Dyn.* **21**, 65–87 (2003).
52. Montero, M. R. Métodos Basados en Optimización por Colonia de Hormigas Aplicados al Modelo Hidrofóbico-Polar del Plegamiento de Proteínas., (Instituto Nacional de Astrofísica, Óptica y Electrónica., 2009).
53. Chothia, C. & Lesk, A. M. The relation between the divergence of sequence and structure in proteins. *The EMBO Journal* **5**, 823–826 (1986).
54. Sippl, M. J. & Flöckner, H. Threading thrills and threats. *Structure* **4**, 15–19 (1996).
55. Wang, Z. X. A re-estimation for the total numbers of protein folds and superfamilies. *Protein Eng.* **11**, 621–626 (1998).
56. Ripka, W. C. Computer-assisted model building. *Nature* **321**, 93–94 (1986).
57. Blundell, T. L., Bedarkar, S., Rinderknecht, E. & Humbel, R. E. Insulin-like growth factor: a model for tertiary structure accounting for immunoreactivity and receptor binding. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 180–184 (1978).
58. Greer, J. Comparative model-building of the mammalian serine proteases. *J. Mol. Biol.* **153**, 1027–1042 (1981).

59. Blundell, T., Sibanda, B. L. & Pearl, L. Three-dimensional structure, specificity and catalytic mechanism of renin. *Nature* **304**, 273–275 (1983).
60. Greer, J. Model structure for the inflammatory protein C5a. *Science* **228**, 1055–1060 (1985).
61. Palmer, K. A., Scheraga, H. A., Riordan, J. F. & Vallee, B. L. A preliminary three-dimensional structure of angiogenin. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 1965–1969 (1986).
62. Chothia, C. *et al.* The predicted structure of immunoglobulin D1.3 and its comparison with the crystal structure. *Science* **233**, 755–758 (1986).
63. Jones, D. T., Taylor, W. R. & Thornton, J. M. A new approach to protein fold recognition. *Nature* **358**, 86–89 (1992).
64. Chandonia, J.-M. & Brenner, S. E. The impact of structural genomics: expectations and outcomes. *Science* **311**, 347–351 (2006).
65. Orengo, C. A., Jones, D. T. & Thornton, J. M. Protein superfamilies and domain superfolds. *Nature* **372**, 631–634 (1994).
66. Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536–540 (1995).
67. Govindarajan, S. & Goldstein, R. A. Why are some proteins structures so common? *Proc. Natl. Acad. Sci. U. S. A.* **93**, 3341–3345 (1996).
68. Orengo, C. *et al.* CATH – a hierarchic classification of protein domain structures. *Structure* **5**, 1093–1109 (1997).
69. Sippl, M. J. Knowledge-based potentials for proteins. *Curr. Opin. Struct. Biol.* **5**, 229–235 (1995).
70. Miyazawa, S. & Jernigan, R. L. Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. *Macromolecules* **18**, 534–552 (1985).
71. Eisenberg, D. & McLachlan, A. D. Solvation energy in protein folding and binding. *Nature* **319**, 199–203 (1986).
72. Miller, R. T., Jones, D. T. & Thornton, J. M. Protein fold recognition by sequence threading: tools and assessment techniques. *FASEB J.* **10**, 171–8 (1996).
73. Bryant, S. H. & Altschul, S. F. Statistics of sequence-structure threading. *Curr. Opin. Struct. Biol.* **5**, 236–244 (1995).
74. Samudrala, R. & Levitt, M. Decoys ‘R’ Us: A database of incorrect conformations to improve protein structure prediction. *Protein Sci.* **9**, 1399–1401 (2000).
75. John, B. & Sali, A. Comparative protein structure modeling by iterative alignment, model building and model assessment. *Nucleic Acids Research* **31**, 3982–3992 (2003).
76. Kortemme, T., Morozov, A. V. & Baker, D. An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein-protein complexes. *J. Mol. Biol.* **326**, 1239–1259 (2003).
77. Krüger, D. M., Garzón, J. I., Chacón, P. & Gohlke, H. DrugScorePPI knowledge-based potentials used as scoring and objective function in protein-protein docking. *PLoS One* **9**, (2014).

78. Morozov, A. V. & Kortemme, T. Potential Functions for Hydrogen Bonds in Protein Structure Prediction and Design. *Adv. Protein Chem.* **72**, 1–38 (2005).
79. Ma, X. H., Wang, C. X., Li, C. H. & Chen, W. Z. A fast empirical approach to binding free energy calculations based on protein interface information. *Protein Eng.* **15**, 677–681 (2002).
80. Moal, I. H. & Fernandez-Recio, J. Intermolecular contact potentials for protein-protein interactions extracted from binding free energy changes upon mutation. *J. Chem. Theory Comput.* **9**, 3715–3727 (2013).
81. Pokala, N. & Handel, T. M. Energy functions for protein design: Adjustment with protein-protein complex affinities, models for the unfolded state, and negative design of solubility and specificity. *J. Mol. Biol.* **347**, 203–227 (2005).
82. Chen, Y., Kortemme, T., Robertson, T., Baker, D. & Varani, G. A new hydrogen-bonding potential for the design of protein-RNA interactions predicts specific contacts and discriminates decoys. *Nucleic Acids Res.* **32**, 5147–5162 (2004).
83. Sánchez-González, G., Kim, J.-K., Kim, D.-S. & Garduño-Juárez, R. A beta-complex statistical four body contact potential combined with a hydrogen bond statistical potential recognizes the correct native structure from protein decoy sets. *Proteins Struct. Funct. Bioinforma.* **81**, 1420–1433 (2013).
84. Sticke, D. F., Presta, L. G., Dill, K. A. & Rose, G. D. Hydrogen bonding in globular proteins. *J. Mol. Biol.* **226**, 1143–1159 (1992).
85. DeGroot, M. H. *Probability and Statistics*. (Addison Wesley Publishing Company, 1986).
86. Zhang, J. & Zhang, Y. A Novel Side-Chain Orientation Dependent Potential Derived from Random-Walk Reference State for Protein Fold Selection and Structure Prediction. **5**, (2010).
87. Wu, S., Skolnick, J. & Zhang, Y. Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biol.* **5**, 17 (2007).
88. Shen, M.-Y., Shen, M.-Y., Sali, A. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci. A Publ. Protein Soc.* **15**, 2507–24 (2006).
89. Mirzaie, M. & Sadeghi, M. Delaunay-based nonlocal interactions are sufficient and accurate in protein fold recognition. *Proteins Struct. Funct. Bioinforma.* **82**, 415–423 (2014).
90. Jianpeng, M. Explicit orientation dependence in empirical potentials and its significance to side-chain modeling. *Acc. Chem. Res.* **42**, 1087–1096 (2009).
91. Sun, W. & He, J. Native secondary structure topology has near minimum contact energy among all possible geometrically constrained topologies. *Proteins Struct. Funct. Bioinforma.* **77**, 159–173 (2009).
92. Chen, L. & He, J. A distance- and orientation-dependent energy function of amino acid key blocks. *Biopolymers* **101**, 681–692 (2014).
93. Zhou, H. & Skolnick, J. GOAP: A generalized orientation-dependent, all-atom statistical potential for protein structure prediction. *Biophys. J.* **101**, 2043–2052 (2011).

94. Park, J. & Saitou, K. ROTAS: a rotamer-dependent, atomic statistical potential for assessment and prediction of protein structures. *BMC Bioinformatics* **15**, 307 (2014).

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

```

import Bio.PDB as bp
from sys import argv
import numpy

DB = [] #Átomo base del átomo donador (Carbono alfa)
D = [] #Átomo donador (Nitrogeno del grupo amino)
DAT = [] #Tipo de átomo donador
AB = [] #Átomo base del átomo aceptor (carbono carbonílico)
AR = [] #Átomo de referencia para el átomo aceptor, para el calculo del ángulo diedro (carbono alfa)
A = [] #Átomo aceptor (oxigeno carbonílico)
AAT = [] #tipo del átomo aceptor
HP = [] #Hidrogenos polares (hidrógenos de grupo amino)
Archivo = [] #Todos los átomos del PDB
ResD = [] #Tipo de residuo donador
ResA = [] #Tipo de residuo aceptor
ResDN = [] #Número del residuo donador (para calcular conectividad)
ResAN = [] #Número del residuo aceptor (para calcular conectividad)

#Este programa mide los siguientes parámetros en todos los hidrogenos polares de la cadena principal de proteínas que se encuentren a una
distancia menor a 2.6 Angstroms de un átomo de oxígeno carbonílico de cadena principal: distancia H-A, angulos D-H-A y H-A-AB, y angulos
diedros DB-D-H-A, D-H-A-AB y H-A-AB-AR

#Se genera una salida con líneas escritas de la siguiente manera: archivo de entrada, tipo de residuo donador, tipo de residuo aceptor,
tipo de átomo donador, tipo de átomo aceptor, número del residuo donador, número del residuo aceptor, distancia H-A, ángulo D-H-A, ángulo
H-A-AB, ángulo diedro DB-D-H-A, ángulo diedro D-H-A-AB, ángulo H-A-AB-AR. Donde cada uno de los datos se ubica en una columna diferente.

#Esta parte del código lee un pdb y toma todos los átomos presentes en el.

f = open(argv[1])
for line in f:
    if line[0:4] == "ATOM":
        Archivo.append(line)
f.close()

for i in range(0, len(Archivo)):

#Esta parte del código identifica los hidrógeno polares y átomos donadores de hidrógeno en la cadena principal, así como sus referencias
para el cálculo de los parámetros geométricos.

    if Archivo[i][12:16] == " H " or Archivo[i][12:16] == " HI " or Archivo[i][12:16] == " H2 " or Archivo[i][12:16] == " H3 ":
        for j in Archivo:
            if int(j[22:26]) == int(Archivo[i][22:26]) and j[12:16] == " N ":
                for k in Archivo:
                    if int(k[22:26]) == int(Archivo[i][22:26]) and k[12:16] == " CA ":
                        HP.append([float(Archivo[i][30:38]), float(Archivo[i][38:46]), float(Archivo[i][46:54])])
                        D.append([float(j[30:38]), float(j[38:46]), float(j[46:54])])
                        DB.append([float(k[30:38]), float(k[38:46]), float(k[46:54])])
                        ResD.append(k[17:20])
                        ResDN.append(int(k[22:26]))
                        DAT.append(j[12:16])

#Esta parte del código identifica los átomos aceptores de hidrógeno en la cadena principal, así como sus referencias para el cálculo de
los parámetros geométricos.

        elif Archivo[l][12:16] == " O ":
            for j in Archivo:
                if int(j[22:26]) == int(Archivo[l][22:26]) and j[12:16] == " C ":
                    for k in Archivo:
                        if int(k[22:26]) == int(Archivo[l][22:26]) and k[12:16] == " CA ":
                            A.append([float(Archivo[l][30:38]), float(Archivo[l][38:46]), float(Archivo[l][46:54])])
                            AB.append([float(j[30:38]), float(j[38:46]), float(j[46:54])])
                            AR.append([float(k[30:38]), float(k[38:46]), float(k[46:54])])
                            ResA.append(k[17:20])
                            ResAN.append(int(k[22:26]))
                            AAT.append(Archivo[l][12:16])

#Esta parte del código selecciona los hidrógenos polares que se encuentran a menos de 2.6 Angstroms de un átomo de oxígeno carbonílico de
cadena principal, calcula los parámetros geométricos y genera la salida.
for i in range(0, len(HP)):

    h = bp.Vector(HP[i])
    d = bp.Vector(D[i])
    db = bp.Vector(DB[i])
    resd = ResD[i]
    resdn = ResDN[i]
    dat = DAT[i]
    for j in range(0, len(A)):
        resan = ResAN[j]
        aat = AAT[j]
        if dat == " N " and resan != resdn-1:
            a = bp.Vector(A[j])
            ab = bp.Vector(AB[j])
            ar = bp.Vector(AR[j])
            resa = ResA[j]
            Dif = bp.Vector.norm(bp.Vector._sub_(h,a))
            if Dif <= 2.6:
                AngDHA = bp.calc_angle(d,h,a)*180/numpy.pi
                AngHAAB = bp.calc_angle(h,a,ab)*180/numpy.pi
                DieDRDHA = bp.calc_dihedral(db,d,h,a)*180/numpy.pi
                DieDHAAB = bp.calc_dihedral(d,h,a,ab)*180/numpy.pi
                DieHAABAR = bp.calc_dihedral(h,a,ab,ar)*180/numpy.pi
                print argv[1], "\t", resd, "\t", resa, "\t", dat, "\t", aat, "\t", str(resdn), "\t", str(resan), "\t", Dif,
                "\t", AngDHA, "\t", AngHAAB, "\t", DieDRDHA, "\t", DieDHAAB, "\t", DieHAABAR

```

Figura 1. Código del programa empleado para identificar y caracterizar geoméricamente los puentes de hidrógeno formados por la cadena principal de proteínas en un archivo pdb.

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 1. Códigos PDB de las 73 proteínas que formaron parte del grupo ERX. La letra que se encuentra frente al código PDB es el identificador de la cadena estudiada en el caso de que este fuese distinto de “A”.

1BXO	1CEX	1EJG	1G66	1GCI	1IXH	1JFB	1L9L
1M40	1MJ5	1MUW	1NQJ	1O7J	1RTQ	1TQG	1TT8
1UCS	1VYR	1X6Z	1YRI	2B97	2ERL	2FVY	2P5K
2RH2	2VB1	2XJP	2XU3	2Y78	3A38	3AJ4	3F1L
3GOE	3HYD	3IP0	3O4P	3QR7	3RWN	3TEU	3UI4
3V1A	3VOR	3W5H	3WCQ	3ZOJ	3ZUC	4ACJ	4AXO
4E0K	4E3Y	4EGU	4F1V	4G9S	4G9S_B	4GA2	4HNO
4HS1	4KQP	4LFS	4MZC	4NDS	4NPD	4NSV	4PNO
4R2X	4REK	4TXR	4TXR_B	4TXR_C	4U9H_L	4U9H_S	4WEE
4X5P							

Tabla 2. Códigos PDB de las 4020 proteínas que formaron parte del grupo ERMN. La letra que se encuentra frente al código PDB es el identificador de la cadena estudiada en el caso de que este fuese distinto de “A”.

1A1W	1A63	1A6S	1A7M	1A90	1AA3	1AB7	1ADR
1ADX	1ADZ	1AG4	1AGG	1AH2	1AH9	1AIW	1AJ3
1AJW	1ALE	1ALF	1ALG	1ANP	1ANS	1AOY	1AP0
1APF	1APJ	1APQ	1AQ5	1AQG	1AUU	1AUZ	1AWE
1AXH	1B03	1B22	1B2T	1B4R	1B64	1B75	1B8Q
1B8Q_B	1B8W	1B9P	1BA5	1BAK	1BAL	1BBA	1BBG
1BBI	1BC9	1BCT	1BDS	1BEG	1BF8	1BFM	1BGK
1BH7	1BHA	1BHI	1BHU	1BI6_H	1BI6_L	1BIP	1BJ8
1BL1	1BM4	1BMW	1BMX	1BO9	1BOE	1BPR	1BQ0
1BQZ	1BR0	1BRV	1BUY	1BV8	1BVM	1BW3	1BW6
1BY0	1BY1	1BY6	1BYM	1BYY	1BZG	1C01	1C05
1C3Y	1C54	1C89	1C8P	1CDB	1CDQ	1CE3	1CE4
1CEK	1CEU	1CFE	1CHL	1CI5	1CIX	1CKV	1CKW
1CMR	1CN7	1CO1	1COU	1CQ0	1CQU	1CS9	1CT6
1CW5	1CWW	1CWX	1CX1	1CXW	1CZ4	1CZ6	1D1D
1D1N	1D1R	1D2D	1D4B	1D5Q	1D6T	1D6X	1D7Q
1D7Q_B	1D8B	1D8J	1D8V	1D9A	1D9N	1DBD	1DCJ
1DDB	1DDF	1DDM	1DDM_B	1DE1	1DEC	1DEP	1DG4
1DGN	1DGQ	1DGU	1DJF	1DKC	1DL0	1DN3	1DNG
1DOQ	1DP3	1DSK	1DTC	1DU1	1DU6	1DV9	1DX0
1DZ7	1E0G	1E0N	1E0Q	1E17	1E52	1E5U	1E68
1E8P	1E8R	1E9K	1ECI	1ECI_B	1ED7	1EDP	1EDS
1EDV	1EDW	1EDX	1EF5	1EGT	1EHS	1EHX	1EI0
1EIJ	1EIK	1EIW	1EJ5	1EJQ	1EL0	1EMW	1EMX
1EMZ	1ENW	1EO0	1EOQ	1EQ1	1EQK	1EQX	1ERC
1ERD	1ERP	1ERY	1EV0	1EWI	1EWS	1EXE	1EXG
1EZE	1EZY	1F0Z	1F2H	1F3K	1F43	1F4I	1F53
1F5X	1F6G	1F6V	1F7E	1F7W	1FAC	1FAF	1FCT
1FDF	1FES	1FEX	1FGD	1FGE	1FGP	1FHO	1FHQ
1FHT	1FJ7	1FLI	1FME	1FMF	1FMY	1FOV	1FOY
1FQQ	1FR0	1FRY	1FSB	1FSH	1FU3	1FUL	1FUV
1FVY	1FW5	1FWO	1FWP	1FYB	1FYC	1FYJ	1FZT
1G03	1G10	1G2H	1G4F	1G5V	1G6E	1G6Z	1G7E
1G7O	1G84	1G91	1G92	1G9L	1G9P	1GCF	1GD5
1GE9	1GH8	1GH9	1GHC	1GHH	1GHT	1GJJ	1GJS
1GJZ	1GM2	1GNA	1GNB	1GNC	1GO0	1GO5	1GP8
1GW3	1GWP	1GXE	1GXI	1GYZ	1H0Z	1H20	1H3Z
1H5P	1H67	1H7D	1H7J	1H8B	1H8B_B	1H8C	1H95
1H9E	1H9F	1HA8	1HA9	1HAE	1HBW	1HCE	1HD6

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 2 (continuación).

1HEH	1HF9	1HFF	1HFI	1HFN	1HHN	1HI7	1HIC
1HJ0	1HK6	1HKS	1HKY	1HN3	1HN6	1HO7	1HO9
1HP3	1HP8	1HP9	1HPW	1HQB	1HS7	1HTX	1HU5
1HU6	1HU7	1HVW	1HVZ	1HX2	1HY9	1HYK	1HZ3
1HZM	1I16	1I17	1I1S	1I26	1I35	1I4V	1I5H_B
1I5H_W	1I6Z	1I8C	1I8X	1I88	1I8A	1IBX	1IBX_B
1ICH	1ID6	1ID7	1IDY	1IE5	1IE6	1IEZ	1IFY
1IG6	1IIE	1IIJ	1IIO	1IJA	1IJZ	1ILF	1ILO
1ILY	1IM1	1IMO	1INZ	1IOJ	1IOU	1IPG	1IQS
1IRG	1IRL	1IRY	1IRZ	1ISK	1IT5	1ITP	1ITY
1IUF	1IUR	1IUY	1IV0	1IVZ	1IW4	1IWF	1IX5
1IXT	1IXU	1IYC	1IYG	1IYR	1IYY	1J0S	1J0T
1J1H	1J26	1J2M	1J4M	1J4O	1J57	1J6Q	1J6Y
1J7H	1J7M	1J7Q	1J8C	1J9I	1JBI	1JBL	1JBN
1JCU	1JDM	1JEG	1JEG_B	1JEI	1JFJ	1JFN	1JH3
1JH4	1JH4_B	1JI8	1JJG	1JJR	1JJS	1JLI	1JLZ
1JMQ	1JMQ_P	1JNS	1JO5	1JO6	1JOY	1JQ1	1JRJ
1JRM	1JRU	1JT8	1JVR	1JW3	1JWE	1JXC	1K0S
1K0X	1K18	1K1V	1K1Z	1K36	1K3K	1K42	1K43
1K5K	1K5O	1K7B	1K8B	1K8H	1K8M	1K91	1K99
1K9C	1KA7	1KA7_B	1KAL	1KAT_V	1KAT_X	1KDL	1KG1
1KJ0	1KJ5	1KJK	1KJS	1KKD	1KKG	1KLA	1KLP
1KMA	1KMD	1KMR	1KMX	1KN5	1KN6	1KN7	1KOY
1KQI	1KRI	1KRS	1KS0	1KSQ	1KSR	1KUL	1KUP
1KUP_B	1KUW	1KVN	1KWD	1KWE	1KX6	1KXL	1KZ0
1KZ2	1KZ5	1L0M	1L1I	1L1K	1L1P	1L2Z	1L2Z_B
1L3G	1L3H	1L3Q	1L3Y	1L5I	1L6H	1L6N	1L6T
1L7B	1L8Y	1LB0	1LB7	1LBJ	1LCX	1LDL	1LDR
1LEA	1LFC	1LG4	1LL8	1LMR	1LMZ	1LQC	1LR1
1LS4	1LUP	1LV4	1LVQ	1LVR	1LVZ	1LWM	1LWR
1LX8	1LY7	1LYP	1M02	1M23	1M2E	1M3G	1M4E
1M4P	1M4P_B	1M5Z	1M62	1M7K	1M7L	1M8L	1M9G
1M9L	1M9W	1MA2	1MA4	1MA5	1MA6	1MEQ	1MF6
1MH6	1MIT	1MJD	1MKC	1MKN	1MM0	1MM4	1MNT
1MO7	1MOT	1MP1	1MPH	1MSZ	1MUZ	1MV3	1MVZ
1MWY	1MYN	1MYU	1MZI	1MZK	1N2Y	1N3J	1N3K
1N4C	1N4I	1N5G	1N5P	1N6T	1N6Z	1N87	1N88
1N9J	1N9U	1N9V	1NCT	1ND9	1NE3	1NEI	1NEQ
1NFA	1NG7	1NHO	1NI7	1NI8	1NIZ	1NJ0	1NKL
1NLA	1NM4	1NMJ	1NMR	1NMV	1NNV	1NO8	1NQ4
1NR3	1NS1	1NSO	1NTC	1NWB	1NXI	1NY4	1NY8
1NY9	1NYJ	1NYN	1NYO	1NZ8	1NZP	1O1W	1O2F
1O2F_B	1O53	1O6W	1O6X	1O78	1O7B	1O8R	1O8Y
1O8Z	1OCP	1ODP	1OEF	1OEG	1OEH	1OEI	1OF9
1OH1	1OIG	1OJG	1OKD	1OM2	1OM2_B	1OMQ	1ON4
1ONB	1ONV	1ONV_B	1OO9	1OO9_B	1OP1	1OP4	1OQA
1OQK	1ORX	1OSX	1OV2	1OVQ	1OVY	1OW5	1OWA
1OWT	1OWX	1OY2	1OYI	1OZZ	1P1D	1P23	1P4S
1P4W	1P6R	1P6S	1P82	1P88	1P8B	1P94	1P9F
1P9K	1PA4	1PAV	1PBA	1PBU	1PC0	1PC2	1PCE
1PCP	1PD6	1PDC	1PE3	1PFL	1PFS	1PGY	1PJD
1PJV	1PJZ	1PLO	1PLP	1PLS	1PLW	1PLX	1PM6
1PN5	1PNB	1PNB_B	1POG	1POQ	1POU	1PPQ	1PQN
1PRB	1PSM	1PSY	1PU1	1PUL	1PUZ	1PV0	1PVE
1PVZ	1PX9	1PYV	1PZQ	1PZR	1Q01	1Q0V	1Q1V
1Q27	1Q2F	1Q2K	1Q2Z	1Q38	1Q3J	1Q3T	1Q53
1Q56	1Q59	1Q5F	1Q5L	1Q5L_B	1Q80	1Q8K	1Q8L

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 2 (continuación).

1Q9P	1QCE	1QCM	1QEY	1QFN	1QFN_B	1QG9	1QGP
1QH2	1QH2_B	1QHK	1QJT	1QK7	1QK9	1QKH	1QKL
1QLO	1QMC	1QND	1QO6	1QP6	1QPM	1QRJ	1QSZ
1QTG	1QTT	1QU5	1QU6	1QVK	1QVL	1QVP	1QW1
1QWP	1QXC	1QXF	1QYT	1QZE	1R02	1R05	1R1F
1R21	1R36	1R3B	1R48	1R4G	1R4K	1R4T	1R4Y
1R57	1R5E	1R5S	1R6R	1R73	1R7C	1R8T	1R9K
1RAX	1RCH	1RDU	1RES	1RFA	1RFL	1RG3	1RG4
1RG6	1RHW	1RHX	1RIJ	1RIK	1RIM	1RIP	1RJH
1RJJ	1RJT	1RKL	1RL1	1RMJ	1RMK	1RO4	1ROO
1ROT	1RPB	1RPR	1RQ6	1RQ8	1RQS	1RQT	1RQU
1RRB	1RRZ	1RSF	1RSO	1RSO_B	1RSW	1RSX	1RT0
1RVS	1RW2	1RW5	1RWU	1RY3	1RYJ	1RYK	1RYU
1RZS	1RZW	1S1N	1S2H	1S3A	1S4H	1S4J	1S4T
1S4W	1S4Z	1S4Z_C	1S62	1S6D	1S6L	1S6W	1S79
1S7A	1S7E	1S7P	1S7P_B	1S6B	1SBO	1SE7	1SE9
1SF0	1SG5	1SGG	1SGO	1SJQ	1SJR	1SKH	1SKI
1SKK	1SLJ	1SM7	1SMZ	1SN6	1SNL	1SOH	1SOL
1SOU	1SOY	1SP0	1SP7	1SPF	1SQ8	1SR2	1SR3
1SRO	1SRZ	1SS3	1SS6	1SSE	1SSE_B	1SSF	1SSK
1SSL	1SSN	1SSU	1STU	1SUH	1SUT	1SUY	1SUY_C
1SVQ	1SX0	1SXD	1SXL	1SZV	1T0C	1T0G	1T0V
1T0Y	1T17	1T1H	1T2Y	1T30	1T3V	1T4N	1T4Y
1T50	1T5Q	1TBA	1TBA_B	1TCP	1TDP	1TE7	1TH5
1TIZ	1TK7	1TKN	1TKV	1TLE	1TM9	1TMR	1TNS
1TOR	1TOS	1TOZ	1TP4	1TPG	1TPM	1TQ1	1TQZ
1TR4	1TRL	1TTE	1TTN	1TTY	1TUZ	1TVI	1TVM
1TVS	1TXP	1TYK	1TZ1	1U0I	1U0I_B	1U2F	1U37
1U3B	1U3N	1U3O	1U57	1U5M	1U6F	1U6U	1U6V
1U7Q	1U89	1UAO	1UAP	1UAW	1UB1	1UC6	1UDK
1UEO	1UF0	1UFG	1UFM	1UFN	1UFZ	1UG2	1UG7
1UG8	1UGJ	1UGL	1UHM	1UHS	1UHT	1UHU	1UJL
1UJR	1UJS	1UJX	1UK5	1UL7	1ULO	1UMQ	1URF
1UST	1UT3	1UTA	1UUC	1UVF	1UVG	1UXC	1UYA
1UYB	1V06	1V1C	1V1D	1V2Y	1V31	1V32	1V46
1V4R	1V4Z	1V5K	1V5M	1V5P	1V5S	1V61	1V65
1V66	1V6B	1V6R	1V88	1V92	1V95	1V9V	1V9W
1VAZ	1VCS	1VD0	1VD2	1VD7	1VDI	1VDL	1VEH
1VFI	1VIG	1VKR	1VRY	1VTP	1VYN	1VZS	1W2Q
1W3D	1W4I	1W4J	1W4M	1W6V	1W9R	1WA8	1WA8_B
1WAZ	1WCJ	1WCL	1WEY	1WF6	1WF9	1WFD	1WFR
1WFS	1WFY	1WGH	1WGK	1WGM	1WGO	1WGP	1WGR
1WGS	1WGW	1WH0	1WH2	1WH4	1WH9	1WHB	1WHN
1WHR	1WHU	1WHV	1WI0	1WI5	1WI9	1WIB	1WIC
1WID	1WIH	1WIJ	1WIK	1WIN	1WIT	1WIX	1WJ1
1WJ4	1WJ5	1WJ6	1WJ7	1WJJ	1WJK	1WJN	1WJQ
1WJR	1WJT	1WJU	1WJW	1WK1	1WKI	1WKT	1WLM
1WLO	1WLP	1WLP_B	1WLX	1WM7	1WM8	1WMT	1WMV
1WN4	1WN8	1WNJ	1WNK	1WNM	1WNN	1WO9	1WOT
1WPI	1WQB	1WQD	1WQE	1WRF	1WRG	1WT8	1WU0
1WUZ	1WVK	1WVO	1WWQ	1WWT	1WWY	1WXA	1WXM
1WXP	1WXS	1WYJ	1WYO	1WZ4	1X0H	1X1F	1X32
1X37	1X3A	1X3U	1X4O	1X4Q	1X4R	1X4T	1X51
1X52	1X53	1X57	1X58	1X5M	1X60	1X67	1X7K
1X93	1X9B	1X9V	1X9X	1XAX	1XC5	1XDX	1XEE
1XG1	1XHS	1XJ1	1XKE	1XKM	1XKM_B	1XNA	1XNE
1XO8	1XOY	1XPN	1XPV	1XPW	1XQ8	1XR0	1XR0_B

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 2 (continuación).

1XRD	1XS3	1XSA	1XSF	1XSX	1XU6	1XUT	1XV3
1XWE	1XX0	1XX3	1XZY	1Y00	1Y03	1Y1B	1Y29
1Y49	1Y5C	1Y6D	1Y6U	1Y76	1Y76_B	1Y7Q	1Y7X
1Y8M	1Y9J	1Y9O	1YBJ	1YDU	1YEL	1YEZ	1YGM
1YGO	1YH5	1YHP	1YKA	1YNX	1YO4	1YQA	1YS5
1YSE	1YSF	1YSM	1YT6	1YTR	1YUA	1YUB	1YVC
1YWI	1YWI_B	1YWU	1YWW	1YWY	1YX0	1YX3	1YX4
1YX7	1YXE	1YYB	1YYC	1YZA	1YZS	1Z1D	1Z1D_B
1Z1M	1Z1V	1Z1Z	1Z23	1Z2F	1Z2G	1Z2K	1Z2Q
1Z2T	1Z3K	1Z4H	1Z6V	1Z7P	1Z87	1Z8S	1Z99
1Z9B	1Z9E	1Z9I	1Z9Q	1Z9V	1ZAE	1ZAQ	1ZC1
1ZDA	1ZDV	1ZDX	1ZFI	1ZG2	1ZHC	1ZKH	1ZL8
1ZL8_B	1ZO0	1ZRI	1ZRV	1ZRX	1ZTN	1ZTS	1ZU2
1ZUF	1ZUG	1ZUV	1ZWT	1ZWV	1ZXA	1ZXF	1ZYI
1ZZA	1ZZF	1ZZP	1ZZV	2A02	2A05	2A1C	2A2B
2A2P	2A2Y	2A3D	2A3J	2A4H	2A63	2A7O	2A7U
2A7U_B	2A7Y	2AB9	2ABO	2ABY	2ACM	2ACM_B	2ADL
2ADZ	2AFD	2AFJ	2AGM	2AHQ	2AIV	2AJ0	2AJE
2AJJ	2AJW	2AKK	2AL3	2ALJ	2AMI	2AMN	2AN7
2AP7	2APN	2AQ0	2AQA	2AQE	2ARI	2ARW	2ASY
2ATG	2AXD	2AXL	2AYA	2AYX	2AYY	2B0G	2B0H
2B0Y	2B19	2B1W	2B38	2B3A	2B3W	2B5B	2B5X
2B7E	2B7T	2B7V	2B8F	2B9K	2B9Z	2BA3	2BAF
2BB8	2BBL	2BBX	2BBY	2BEY	2BGO	2BIC	2BJX
2BL5	2BN5	2BN5_B	2BN8	2BP4	2BTT	2BUD	2BW2
2BYE	2BYF	2BZT	2C0S	2C34	2C55	2C5Z	2C7H
2CEH	2CH0	2CJN	2CK5	2CKA	2CKC	2CKN	2CKU
2COB	2COE	2COK	2COM	2COO	2COU	2COW	2CP5
2CP6	2CP9	2CPM	2CPR	2CPT	2CQ7	2CQ8	2CQ9
2CQA	2CQJ	2CQK	2CQL	2CQM	2CQU	2CQW	2CQY
2CR0	2CR2	2CR5	2CR7	2CRB	2CRF	2CRI	2CRL
2CRQ	2CRU	2CRV	2CRY	2CS4	2CSA	2CSF	2CSK
2CSO	2CSW	2CT6	2CU1	2CUJ	2CVR	2CW1	2CYK
2CZN	2CZO	2CZY	2CZY_B	2D1U	2D35	2D3J	2D46
2D49	2D56	2D5U	2D85	2D86	2D8B	2D8M	2D93
2D9D	2D9F	2D9I	2D9O	2D9T	2D9W	2DAD	2DAE
2DAF	2DAJ	2DAL	2DAM	2DAO	2DAV	2DAW	2DB2
2DB9	2DBA	2DBC	2DBD	2DBF	2DBH	2DCE	2DCI
2DCO	2DCP	2DCQ	2DCV	2DGR	2DGY	2DGZ	2DHM
2DHS	2DHX	2DHZ	2DI0	2DIG	2DII	2DIR	2DIU
2DIW	2DJM	2DJP	2DJV	2DK1	2DK3	2DK4	2DK6
2DK7	2DK8	2DKS	2DKX	2DKY	2DKZ	2DL1	2DL6
2DLL	2DLW	2DLX	2DME	2DMH	2DMW	2DNF	2DNT
2DNW	2DNX	2DO1	2DO3	2DO5	2DO7	2DO9	2DOG
2DRN	2DRN_C	2DSM	2DT6	2DT7	2DT7_B	2DUN	2DUW
2DW3	2DX2	2DX3	2DX4	2DY7	2DY8	2DYF	2DYF_B
2DZJ	2DZK	2DZL	2DZM	2E0G	2E19	2E2F	2E2W
2E44	2E45	2E4E	2E4H	2E4H_B	2E5I	2E5J	2E5N
2E5O	2E5P	2E5T	2E5U	2E5Z	2E60	2E62	2E63
2E6J	2E6N	2E6Z	2E70	2E71	2E7C	2E7G	2E7M
2E8D	2E8J	2E8M	2E8O	2E8P	2E9G	2EAM	2EAP
2EBI	2EBU	2EC1	2EC3	2EC4	2EDN	2EDO	2EDU
2EE1	2EE4	2EE7	2EEF	2EEL	2EEM	2EFI	2EFZ
2EH0	2EJE	2EJM	2EJS	2EJY	2EJY_B	2EKF	2EKI
2EKK	2EKO	2EL8	2ELH	2ELJ	2ENJ	2ENK	2ENQ
2EO2	2EOC	2EP8	2EPB	2EQH	2EQJ	2EQK	2EQM
2EQN	2EQO	2EQQ	2EQR	2EQS	2EQU	2EQX	2ERS

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 2 (continuación).

2ES6	2ESX	2ESZ	2EVN	2EVQ	2EW9	2EXD	2EXN
2EYA	2EYV	2EYX	2EYY	2EZH	2EZK	2EZW	2F05
2F1E	2F3I	2F3J	2F40	2F63	2F76	2FBU	2FCD
2FCE	2FCG	2FE0	2FE9	2FEB	2FFT	2FGX	2FH0
2FHO	2FHO_B	2FHW	2FHW_B	2FI2	2FIN	2FIN_B	2FJ6
2FJL	2FKX	2FM4	2FMR	2FO8	2FQA	2FQH	2FS1
2FTU	2FUI	2FV4	2FV4_B	2FXP	2FXY	2FXZ	2FYH
2FYJ	2FZ0	2G0K	2G0Q	2G0U	2G1D	2G2B	2G2K
2G6U	2G7H	2G7J	2G9B	2G9L	2G9P	2GA5	2GAQ
2GBS	2GCC	2GCX	2GD3	2GD7	2GDL	2GF5	2GFU
2GGR	2GJF	2GJH	2GJI	2GLE	2GLW	2GM2	2GMG
2GMO	2GO9	2GOW	2GPQ	2GQB	2GQC	2GRG	2GTJ
2GUT	2GV1	2GVA	2GVP	2GW6	2GX1	2GYT	2GZO
2H1Z	2H2M	2H3J	2H3K	2H41	2H7A	2H7T	2H95
2HA1	2HAC	2HAJ	2HBP	2HC5	2HCC	2HD7	2HDE
2HDL	2HEQ	2HF6	2HFH	2HFI	2HFQ	2HFR	2HFV
2HG6	2HG7	2HGA	2HGC	2HGF	2HGK	2HGN	2HH2
2HH8	2HHI	2HI3	2HI6	2HJJ	2HJQ	2HLG	2HLU
2HLW	2HM2	2HM3	2HN8	2HO9	2HQ3	2HQR	2HRJ
2HST	2HSX	2HTF	2HTJ	2HUG	2HUG_B	2HVZ	2HW0
2HWT	2HX6	2HZD	2I0N	2I3B	2I4K	2I7U	2I83
2I85	2I8L	2I9M	2I9N	2I9S	2I9Y	2IB1	2IDY
2IEM	2IF1	2IFJ	2IFS	2IGU	2IIJ	2IIJ_B	2IKD
2IKE	2ILX	2IMU	2IN2	2ITA	2IUE	2IV4	2IV5
2IVW	2IZ3	2J0Z	2J48	2J4M	2J5D	2J5O	2J5P
2J8J	2J8P	2JGW	2JM2	2JMB	2JMC	2JML	2JMM
2JMP	2JMR	2JMU	2JMV	2JMX	2JMX_B	2JMY	2JN4
2JN5	2JN6	2JN7	2JN8	2JN9	2JNA	2JNC	2JNH
2JNJ	2JNK	2JNQ	2JNS	2JNT	2JNV	2JNW	2JNW_B
2JNZ	2JO1	2JO6	2JO8	2JOA	2JOA_B	2JOB	2JOD_B
2JOD_B	2JOE	2JOG	2JOG_B	2JOI	2JOK	2JON	2JOO
2JOQ	2JOR	2JOS	2JOU	2JOV	2JOZ	2JP0	2JP1
2JP2	2JP3	2JP5	2JP6	2JP8	2JPB	2JPC	2JPD
2JPE	2JPF	2JPH	2JPI	2JPJ	2JPK	2JPN	2JPS
2JPU	2JPX	2JQ2	2JQ3	2JQ8	2JQA	2JQE	2JQF
2JQJ	2JQN	2JQO	2JQQ	2JQV	2JQW	2JQX	2JQY
2JQZ	2JR0	2JR1	2JR3	2JR5	2JR8	2JRA	2JRB
2JRF	2JRH	2JRM	2JRO	2JRR	2JRT	2JRV	2JRW
2JRZ	2JS1	2JS3	2JS5	2JS9	2JSN	2JSP	2JSS
2JSS_B	2JSW	2JSX	2JSY	2JT1	2JTA	2JTC	2JTD
2JTF	2JTK	2JTM	2JTQ	2JTV	2JTX	2JTY	2JU1
2JU5	2JUA	2JUB	2JUC	2JUF	2JUG	2JUH	2JUI
2JUJ	2JUO	2JUQ	2JUR	2JUS	2JUT	2JUZ	2JV2
2JV3	2JV4	2JV5	2JV6	2JV7	2JV8	2JVA	2JVB
2JVD	2JVE	2JVF	2JVG	2JVL	2JVM	2JVR	2JVU
2JVW	2JVZ	2JW2	2JW5	2JW8	2JWE	2JWG	2JWH
2JWK	2JWN	2JWP	2JWY	2JX0	2JX3	2JX5	2JX8
2JXD	2JXF	2JXG	2JXJ	2JXP	2JXT	2JXU	2JXW
2JY0	2JY5	2JY9	2JYA	2JYB	2JYL	2JYN	2JYO
2JYP	2JYS	2JYT	2JZ0	2JZ1	2JZ2	2JZ4	2JZ5
2JZ6	2JZ7	2JZA	2JZB	2JZB_B	2JZC	2JZD	2JZM
2JZN	2JZN_C	2JZT	2JZX	2JZY	2K02	2K06	2K07
2K0L	2K0M	2K0N	2K0Q	2K0S	2K0Z	2K10	2K13
2K14	2K18	2K19	2K1G	2K1L	2K1M	2K1O	2K1S
2K1X	2K21	2K24	2K27	2K29	2K2B	2K2E	2K2J
2K2O	2K2P	2K2R	2K2R_B	2K2S	2K2S_B	2K2W	2K2X
2K2Y	2K2Z	2K32	2K35	2K36	2K37	2K3A	2K3D

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUEBLOS DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 2 (continuación).

2K3F	2K3G	2K3I	2K3K	2K3N	2K3O	2K3P	2K3Q
2K3T	2K3W	2K3W_B	2K3X	2K40	2K42	2K42_B	2K44
2K47	2K48	2K49	2K4B	2K4E	2K4F	2K4K	2K4M
2K4N	2K4Q	2K4T	2K4V	2K4Z	2K50	2K51	2K52
2K54	2K57	2K58	2K59	2K5B	2K5B_B	2K5D	2K5E
2K5G	2K5H	2K5I	2K5J	2K5K	2K5L	2K5R	2K5S
2K5V	2K5X	2K5X_B	2K6B	2K6G	2K6H	2K6I	2K6L
2K6M	2K6O	2K6P	2K6S	2K6U	2K6U_B	2K6V	2K6W
2K6X	2K72	2K73	2K75	2K76	2K77	2K7I	2K7K
2K7M	2K7N	2K7P	2K7Q	2K7V	2K7Y	2K7Z	2K84
2K85	2K87	2K88	2K89	2K8E	2K8H	2K8I	2K8J
2K8O	2K8P	2K8Q	2K8S	2K8V	2K8X	2K8Y	2K9A
2K9D	2K9F	2K9F_B	2K9I	2K9K	2K9L	2K9M	2K9N
2K9O	2K9P	2K9Q	2K9S	2K9U	2K9U_B	2K9X	2K9Y
2K9Z	2KA0	2KA1	2KA3	2KA5	2KA7	2KA9	2KA9_B
2KAA	2KAC	2KAF	2KAP	2KAT	2KAV	2KB1	2KB4
2KB8	2KB9	2KBE	2KBF	2KBG	2KBI	2KBL	2KBS
2KBS_B	2KBT	2KBY	2KBZ	2KC1	2KC2	2KC3	2KC5
2KC6	2KC7	2KC8	2KC8_B	2KCA	2KCC	2KCD	2KCK
2KCM	2KCO	2KCT	2KCU	2KCV	2KCW	2KCX	2KCY
2KCZ	2KD3	2KD7	2KDC	2KDD	2KDI	2KDK	2KDR
2KDS	2KDV	2KDY	2KE3	2KE4	2KE7	2KEA	2KEB
2KEF	2KEG	2KEL	2KEN	2KEO	2KEP	2KEQ	2KER
2KES	2KEY	2KF2	2KF3	2KFD	2KFE	2KFK	2KFK_B
2KFP	2KFQ	2KFS	2KFV	2KFW	2KG4	2KG5	2KG7
2KG7_B	2KGF	2KGH	2KGJ	2KGL	2KGR	2KGT	2KGU
2KGY	2KH2	2KH2_B	2KHA	2KHC	2KHD	2KHE	2KHG
2KHH	2KHH_B	2KHI	2KHJ	2KHK	2KHL	2KHN	2KHO
2KHQ	2KHR	2KHS	2KHS_B	2KHT	2KHV	2KHX	2KHZ
2KI0	2KI3	2KI8	2KI9	2KIA	2KIB	2KIC	2KIE
2KIG	2KIJ	2KIM	2KIQ	2KIS	2KIT	2KIV	2KIW
2KIX	2KJ1	2KJ3	2KJ5	2KJ6	2KJB	2KJF	2KJG
2KJI	2KJM	2KJN	2KJQ	2KJR	2KJV	2KJW	2KJX
2KJY	2KJZ	2KK0	2KK1	2KK2	2KK4	2KK6	2KK7
2KKC	2KKE	2KKG	2KKL	2KKM	2KKN	2KKO	2KKP
2KKQ	2KKS	2KKU	2KKV	2KKX	2KKZ	2KL1	2KL2
2KL3	2KL4	2KL5	2KL7	2KLA	2KLB	2KLD	2KLL
2KLN	2KLO	2KLR	2KLU	2KLV	2K LX	2KLZ	2KM1
2KM2	2KM4	2KM6	2KM9	2KMA	2KMC	2KMG	2KML
2KMM	2KMO	2KMS	2KMT	2KMU	2KMW	2KMZ	2KN0
2KN4	2KN6	2KN8	2KNA	2KNC	2KNC_B	2KND	2KNG
2KNH	2KNH_B	2KNI	2KNJ	2KNO	2KNP	2KNQ	2KNR
2KNS	2KNU	2KNZ	2KO1	2KO2	2KO6	2KOB	2KOD
2KOE	2KOG	2KOK	2KOL	2KON	2KOU	2KOY	2KOZ
2KP2	2KP5	2KP6	2KP7	2KPA	2KPE	2KPF	2KPH
2KPJ	2KPK	2KPM	2KPP	2KPQ	2KPT	2KPY	2KPZ
2KPZ_B	2KQ1	2KQ2	2KQ5	2KQ6	2KQ8	2KQA	2KQR
2KQU	2KQV	2KQX	2KR0	2KR4	2KR6	2KR7	2KR9
2KRA	2KRB	2KRB_B	2KRC	2KRE	2KRF	2KRG	2KRH
2KRK	2KRR	2KRT	2KRX	2KS6	2KS9	2KS9_B	2KSD
2KSE	2KSF	2KSG	2KSH	2KSK	2KSL	2KSM	2KSN
2KSR	2KSV	2KSW	2KT0	2KT6	2KT7	2KT8	2KT9
2KTA	2KTC	2KTE	2KTL	2KTM	2KTS	2KTU	2KU1
2KU7	2KUA	2KUB	2KUD	2KUE	2KUF	2KUI	2KUK
2KUM	2KUP	2KUP_B	2KUQ	2KUS	2KUT	2KUX	2KV2
2KV4	2KV5	2KV9	2KVE	2KVI	2KVO	2KVP	2KVR
2KVS	2KVT	2KVV	2KVX	2KVZ	2KW0	2KW1	2KW3

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUEENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 2 (continuación).

2KW3_C	2KW5	2KW6	2KW7	2KW8	2KW9	2KWA	2KWH
2KWT	2KWX	2KWY	2KWZ	2KX2	2KX4	2KX7	2KXD
2KXE	2KXG	2KXH	2KXH_B	2KXI	2KXJ	2KXL	2KXO
2KXQ	2KXQ_B	2KXR	2KXS	2KXT	2KXX	2KXY	2KY5
2KY9	2KYA	2KYB	2KYC	2KYG	2KYG_C	2KYH	2KYI
2KYM	2KYM_B	2KYR	2KYS	2KYW	2KYY	2KYZ	2KZ3
2KZ4	2KZ6	2KZ7	2KZ7_B	2KZ7_C	2KZ8	2KZ9	2KZA
2KZC	2KZF	2KZG	2KZH	2KZK	2KZN	2KZQ	2KZR
2KZT	2KZT_B	2KZU	2KZU_B	2KZV	2KZW	2KZX	2L01
2L03	2L04	2L05	2L07	2L08	2L09	2L0C	2L0G
2L0J	2L0K	2L0L	2L0N	2L0R	2L0S	2L10	2L14
2L14_B	2L16	2L17	2L1A	2L1I	2L1J	2L1L	2L1L_B
2L1N	2L1P	2L1Q	2L1S	2L1T	2L1X	2L22	2L25
2L26	2L27	2L27_B	2L28	2L2C	2L2D	2L2L	2L2L_B
2L2M	2L2N	2L2O	2L2Q	2L2R	2L34	2L36	2L37
2L38	2L3D	2L3F	2L3G	2L3I	2L3L	2L3N	2L3O
2L3P	2L3S	2L3T	2L3U	2L3X	2L3Y	2L40	2L42
2L48	2L49	2L4A	2L4C	2L4E	2L4J	2L4M	2L4N
2L4O	2L4R	2L4U	2L4V	2L4W	2L4X	2L54	2L55
2L57	2L5A	2L5C	2L5F	2L5G	2L5G_B	2L5I	2L5L
2L5N	2L5P	2L5Q	2L5R	2L5T	2L63	2L66	2L67
2L6A	2L6B	2L6C	2L6F	2L6K	2L6N	2L6O	2L6P
2L6Q	2L6S	2L6U	2L6W	2L70	2L72	2L73	2L76
2L77	2L7A	2L7B	2L7C	2L7E	2L7H	2L7J	2L7K
2L7N	2L7Q	2L7T	2L7W	2L7Y	2L81	2L82	2L83
2L89	2L8A	2L8B	2L8D	2L8K	2L8L	2L8N	2L8S
2L8T	2L8V	2L8X	2L8Y	2L91	2L92	2L93	2L95
2L97	2L99	2L9B	2L9B_B	2L9D	2L9F	2L9G	2L9J
2L9L	2L9M	2L9N	2L9R	2L9U	2L9W	2L9Y	2LA3
2LA4	2LA7	2LA8	2LAA	2LAE	2LAG	2LAG_B	2LAH
2LAI	2LAK	2LAT	2LAV	2LB6	2LB7	2LBB	2LBC
2LBF	2LBF_B	2LBG	2LBN	2LBO	2LBW	2LC0	2LC1
2LC2	2LC3	2LC4	2LC6	2LC9	2LCC	2LCH	2LCJ
2LCK	2LCL	2LCM	2LCR	2LCS	2LCS_B	2LCU	2LCV
2LCW	2LCX	2LD3	2LD4	2LD6	2LD7	2LD7_B	2LDE
2LDF	2LDS	2LDU	2LDY	2LE0	2LE1	2LE2	2LE3
2LE7	2LE8	2LE8_B	2LE9	2LE9_B	2LEH	2LEH_B	2LEK
2LEM	2LEN	2LEO	2LEQ	2LER	2LEY	2LEZ	2LF0
2LF2	2LF3	2LF6	2LF8	2LFB	2LFC	2LFE	2LFG
2LFH	2LFI	2LFJ	2LFK	2LFN	2LFP	2LFR	2LFU
2LFV	2LFW	2LFW_B	2LG1	2LG4	2LG5	2LGC	2LGE
2LGH	2LGN	2LGO	2LGQ	2LGR	2LGW	2LGX	2LGY
2LGZ	2LH0	2LH9	2LHF	2LHR	2LHS	2LHT	2LHU
2LI5	2LI5_B	2LI6	2LIE	2LIF	2LIO	2LIU	2LIX
2LIY	2LIZ	2LJ0	2LJ2	2LJ6	2LJ9	2LJB	2LJH
2LJI	2LJK	2LJM	2LJP	2LJT	2LJU	2LJV	2LJW
2LK2	2LK4	2LK9	2LKB	2LKC	2LKE	2LKG	2LKL
2LKM	2LKM_B	2LKN	2LKP	2LKQ	2LKS	2LKT	2LKY
2LKZ	2LL1	2LL2	2LL3	2LL5	2LL6	2LL6_B	2LLD
2LLE	2LLF	2LLG	2LLH	2LLK	2LLL	2LLN	2LLP
2LLR	2LLV	2LLW	2LLX	2LLZ	2LM0	2LM1	2LM3
2LM4	2LM7	2LM8	2LM9	2LMA	2LMB	2LMC	2LMC_B
2LMD	2LME	2LMF	2LMG	2LMI	2LMK	2LMR	2LMZ
2LN3	2LN7	2LN8	2LNA	2LNB	2LNC	2LNE	2LNF
2LNG	2LNI	2LNJ	2LNK	2LNK_C	2LNL	2LNM	2LNS
2LNV	2LNX	2LNY	2LNZ	2LO0	2LO1	2LO7	2LOE
2LOH	2LOJ	2LOK	2LOL	2LOM	2LON	2LOO	2LOQ

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 2 (continuación).

2LOR	2LOT	2LOY	2LOZ	2LOZ_B	2LP1	2LP4	2LP4_Y
2LP6	2LPB	2LPB_B	2LPD	2LPE	2LPF	2LPK	2LPM
2LPN	2LPU	2LPX	2LQ0	2LQ1	2LQ2	2LQ3	2LQ4
2LQ5	2LQ7	2LQ8	2LQ9	2LQB	2LQG	2LQJ	2LQL
2LQM	2LQN	2LQQ	2LQT	2LQV	2LQX	2LQY	2LR1
2LR1_B	2LR2	2LR3	2LR4	2LR6	2LR7	2LR8	2LRA
2LRC	2LRD	2LRG	2LRJ	2LRK	2LRK_D	2LRQ	2LRS
2LRT	2LRU	2LRV	2LRX	2LS0	2LS1	2LS2	2LS3
2LS4	2LS6	2LS7	2LS8	2LS9	2LSA	2LSE	2LSH
2LSJ	2LSJ_B	2LSL	2LSM	2LSN	2LSO	2LSQ	2LSR
2LSR_B	2LST	2LSU	2LSW	2LSY	2LT3	2LT4	2LT5
2LT9	2LTB	2LTD	2LTF	2LTH	2LTI	2LTJ	2LTK
2LTL	2LTP	2LTR	2LTS	2LTU	2LU1	2LU2	2LU3
2LU6	2LU7	2LUD	2LUF	2LUG	2LUH	2LUH_B	2LUI
2LUM	2LUO	2LUQ	2LUR	2LUS	2LUT	2LUV	2LUW
2LUZ	2LV4	2LV5	2LVA	2LVC	2LVF	2LVG	2LVL
2LVN	2LVS	2LVV	2LVW	2LVX	2LW1	2LW3	2LW4
2LW5	2LW6	2LW7	2LW8	2LW9	2LWA	2LWB	2LWC
2LWD	2LWF	2LWJ	2LWL	2LWQ	2LWS	2LWT	2LWU
2LWV	2LWX	2LWY	2LWZ	2LX0	2LX2	2LX4	2LX5
2LX6	2LX9	2LXB	2LXE	2LXI	2LXL	2LXN	2LXO
2LXR	2LXS	2LXS_B	2LXU	2LY1	2LY3	2LY5	2LY7
2LY8	2LY9	2LYC	2LYD	2LYD_B	2LYE	2LYF	2LYH
2LYI	2LYJ	2LYV	2LYX	2LYY	2LZ0	2LZ1	2LZ3
2LZF	2LZI	2LZJ	2LZL	2LZN	2LZO	2LZP	2LZQ
2M00	2M02	2M03	2M05	2M09	2M0A	2M0B	2M0C
2M0M	2M0N	2M0Q	2M0R	2M0S	2M0T	2M0V	2M0V_B
2M0W	2M0Y	2M19	2M1A	2M1B	2M1C	2M1H	2M1J
2M1L	2M1M	2M1N	2M1T	2M1U	2M1W	2M1Z	2M20
2M25	2M2A	2M2B	2M2D	2M2E	2M2F	2M2I	2M2J
2M2K	2M2L	2M2Q	2M2T	2M2Y	2M35	2M36	2M37
2M38	2M3A	2M3D	2M3E	2M3F	2M3J	2M3K	2M3N
2M3T	2M3V	2M3X	2M41	2M41_B	2M45	2M46	2M47
2M4E	2M4F	2M4G	2M4H	2M4I	2M4K	2M4L	2M4M
2M4N	2M4V	2M4Y	2M5G	2M5H	2M5I	2M5J	2M5L
2M5N	2M5O	2M5S	2M5V	2M5W	2M5X	2M5Y	2M63
2M64	2M66	2M6A	2M6I	2M6K	2M6O	2M6R	2M6T
2M6U	2M6Y	2M70	2M71	2M72	2M73	2M74	2M76
2M77	2M78	2M79	2M7E	2M7F	2M7G	2M7K	2M7L
2M7O	2M7S	2M7T	2M7W	2M7X	2M7Z	2M80	2M83
2M86	2M86_B	2M88	2M8C	2M8E	2M8F	2M8G	2M8H
2M8I	2M8N	2M8R	2M8S	2M8S_B	2M8T	2M8W	2M8X
2M94	2M99	2M9H	2M9K	2M9L	2M9M	2M9N	2M9O
2M9U	2M9V	2M9X	2MA0	2MA1	2MA2	2MA3	2MA4
2MA8	2MA9	2MA9_B	2MA9_C	2MAA	2MAB	2MAH	2MAJ
2MAL	2MAM	2MAO	2MAR	2MAU	2MAW	2MAX	2MAZ
2MB1	2MB7	2MB9	2MBC	2MBD	2MBF	2MBG	2MBK
2MBS	2MBY	2MC2	2MC3	2MC4	2MC6	2MC6_B	2MC7
2MC8	2MC9	2MCA	2MCE	2MCF	2MCH	2MCQ	2MCR
2MCT	2MCY	2MD0	2MD1	2MD2	2MD3	2MD4	2MD5
2MD6	2MDA	2MDC	2MDF	2MDJ	2MDK	2MDL	2MDN
2MDR	2MDT	2MDU	2MDV	2MDW	2MDX	2MDZ	2ME1
2ME8	2MEK	2MEL	2MEM	2MET	2MEW	2MEX	2MEY
2MEZ	2MF2	2MF3	2MF4	2MF6	2MF7	2MF9	2MFI
2MFJ	2MFK	2MFL	2MFn	2MFR	2MFS	2MFV	2MFZ
2MG1	2MG2	2MGO	2MGQ	2MGR	2MGS	2MGV	2MGX
2MH1	2MH2	2MH3	2MH9	2MHC	2MHD	2MHE	2MHF

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 2 (continuación).

2MHG	2MHH	2MHJ	2MHK	2MHL	2MHO	2MHO_B	2MHP
2MHS	2MHV	2MHW	2MHY	2MI1	2MI2	2MI5	2MI6
2MI7	2MID	2MIE	2MIF	2MIG	2MIH	2MII	2MIM
2MIT	2MIX	2MIZ	2MJ1	2MJ2	2MJ3	2MJ4	2MJ6
2MJ7	2MJ8	2MJA	2MJF	2MJF_B	2MJG	2MJK	2MJL
2MJM	2MJN	2MJO	2MJU	2MJW	2MJW_B	2MJY	2MJZ
2MK0	2MK2	2MK3	2MK4	2MK5	2MK6	2MK9	2MKB
2MKC	2MKC_B	2MKC_C	2MKF	2MKH	2MKL	2MKR	2MKR_B
2MKV	2MKX	2MKY	2MKZ	2ML5	2ML7	2ML8	2ML9
2MLB	2MLD	2MLG	2MLH	2MLJ	2MLK	2MLO	2MLQ
2MLU	2MLW	2MLX	2MLX_B	2MM0	2MM4	2MM5	2MM8
2MM9	2MMA	2MMA_B	2MMB	2MMG	2MMM	2MMP	2MMU
2MMW	2MMZ	2MN3	2MN4	2MNG	2MNH	2MNI	2MNU
2MNU_B	2MNW	2MO0	2MOD	2MOM	2MOQ	2MOU	2MOW
2MOW_B	2MOX	2MOZ	2MP4	2MP5	2MPB	2MPC	2MPE
2MPF	2MPL	2MPN	2MPO	2MPQ	2MPU	2MPV	2MPW
2MQ2	2MQ3	2MQ4	2MQ5	2MQB	2MQC	2MQD	2MQE
2MQG	2MQU	2MQW	2MR3	2MR5	2MR7	2MRA	2MRC
2MRI	2MRL	2MRM	2MRN	2MRP	2MRW	2MSR	2MSR_B
2MSU	2MSV	2MSW	2MSX	2MT4	2MT6	2MTC	2MTF
2MTG	2MTL	2MTN	2MTS	2MTW	2MTY	2MU0	2MU1
2MU2	2MU6	2MU7	2MU8	2MUB	2MUE	2MUH	2MUJ
2MUK	2MUY	2MV0	2MV3	2MV6	2MVB	2MVC	2MVC_B
2MVF	2MVM	2MVX	2MW1	2MW3	2MW5	2MW9	2MWD
2MWF	2MWI	2MWL	2MWM	2MWN	2MWN_B	2MWR	2MXG
2MXH	2MXR	2MXT	2MY7	2MZ6	2MZD	2MZD_B	2MZY
2NCM	2NEF	2NLW	2NOU	2NPB	2NPL	2NPU	2NR1
2NRG	2NS4	2NS5	2NVJ	2NWM	2NWT	2NX6	2NX7
2NXU	2NZZ	2O0S	2O2O	2O3D	2O4E	2ODC	2OFG
2OFQ	2OFQ_B	2OGH	2OP7	2OPU	2OPV	2OQ3	2OQ9
2OQP	2ORU	2OS6	2OS6_B	2OSR	2OT2	2OTQ	2OTR
2OUT	2OV6	2OVN	2OWI	2OX2	2OZX	2P01	2P0Q
2P3M	2P5H	2P5J	2P6J	2P7R	2P81	2PCO	2PDD
2PFU	2PHE	2PHE_C	2PJH	2PJH_B	2PJV	2PKU	2PKU_B
2PNG	2PO8	2PPZ	2PTL	2PV6	2PXG	2QMV	2REL
2RLG	2RLH	2RLJ	2RLO	2RLW	2RM4	2RM6	2RM8
2RMG	2RMH	2RMJ	2RML	2RMR	2RMW	2RN4	2RN7
2RND	2RNG	2RNK	2RNL	2RNN	2RNO	2RNQ	2RNZ
2RO0	2RO3	2ROH	2ROO	2ROP	2ROQ	2ROV	2ROZ
2ROZ_B	2RP4	2RP5	2RPA	2RPB	2RPI	2RPJ	2RPS
2RPW	2RQ0	2RQ1	2RQ2	2RQ5	2RQ6	2RQ8	2RQE
2RQG	2RQG_B	2RQL	2RQM	2RQP	2RQQ	2RQR	2RQX
2RR0	2RR6	2RR7	2RR8	2RRD	2RRE	2RRF	2RRI
2RRK	2RRL	2RRN	2RRS	2RRU	2RS6	2RSC	2RSG
2RSM	2RSO	2RST	2RSV	2RSW	2RSX	2RT3	2RT4
2RT6	2RTS	2RTT	2RTU	2RTX	2RTY	2RTZ	2RU0
2RU1	2RU5	2RU8	2RU9	2RUC	2RUH	2RUM	2RUN
2RUO	2RUP	2UWQ	2UZ5	2V0E	2V0F	2V1N	2V37
2V6Z	2V9H	2VDA	2VDA_B	2VIK	2VKC	2VXD	2VXE
2VXF	2W0N	2W9U	2WBR	2WCY	2WGN	2WGO	2WH9
2WNM	2WXC	2X43	2X8N	2XA6	2XDI	2XK0	2XKS
2XL1	2XV9	2XXS	2Y1S	2Y4W	2Y9T	2YH0	2YKA
2YKA_B	2YMJ	2YQF	2YQG	2YQK	2YQR	2YRA	2YRB
2YRL	2YRN	2YRP	2YRQ	2YRU	2YRV	2YRY	2YS0
2YS3	2YS4	2YSR	2YST	2YSZ	2YT0	2YT8	2YTU
2YTV	2YTW	2YU0	2YU3	2YUA	2YUD	2YUE	2YUF
2YUG	2YUH	2YUJ	2YUK	2YUM	2YW5	2YZ0	2Z4D

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUEENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 2 (continuación).

2Z5V	2ZAJ	3AIT	3CRD	3HSF	3LRI	3MRA	3MSP
3NCM	3NLA	3SAK	3ZBE	3ZD0	3ZEH	3ZG4	3ZGK
3ZPD	3ZPM	3ZQD	3ZUA	4A1M	4A53	4A54	4A54_B
4A5V	4AAI	4AKA	4AQZ	4AR0	4ASV	4AXP	4B19
4B2R	4B2U	4B2V	4B6V	4BF8	4BHP	4BIT	4BMF
4BWH	4BXL	4BXL_C	4BXU	4BXU_B	4C26	4CA9	4CH0
4CPG	4CRP	4CSQ	4ULL	4UQT	4UQT_B	4UZM	4UZW
4UZX	4V10	7HSC	8TFV				

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 3. Valores de energía asociados a los parámetros establecidos en los potenciales H-ERX y H-ERMN. (a) la distancia hidrógeno-oxígeno, (b) el ángulo al átomo de hidrógeno, (c) el ángulo al átomo de oxígeno (d) el ángulo diedro que corresponde a la rotación del enlace carbono-oxígeno, (e) el ángulo diedro que corresponde a la rotación del puente de hidrógeno y (f) el ángulo diedro que corresponde a la rotación del enlace hidrógeno-nitrógeno.

Intervalo de distancia (Å)			Energía (u.a.)	
			H-ERX (a)	H-ERMN (a)
0.00	-	1.00	6.0000	6.0000
1.00	-	1.05	6.0000	6.0000
1.05	-	1.10	6.0000	6.0000
1.10	-	1.15	6.0000	6.0000
1.15	-	1.20	6.0000	6.0000
1.20	-	1.25	6.0000	6.0000
1.25	-	1.30	6.0000	5.1695
1.30	-	1.35	6.0000	4.4958
1.35	-	1.40	6.0000	3.9058
1.40	-	1.45	6.0000	3.1701
1.45	-	1.50	6.0000	2.5443
1.50	-	1.55	6.0000	1.8332
1.55	-	1.60	6.0000	1.3526
1.60	-	1.65	5.1706	0.8495
1.65	-	1.70	4.4774	0.4544
1.70	-	1.75	2.6857	0.0319
1.75	-	1.80	0.6062	-0.6273
1.80	-	1.85	-0.7563	-0.7904
1.85	-	1.90	-1.4252	-0.9150
1.90	-	1.95	-1.7983	-1.0257
1.95	-	2.00	-1.8187	-1.1221
2.00	-	2.05	-1.6890	-1.1236
2.05	-	2.10	-1.4045	-1.0988
2.10	-	2.15	-1.1393	-0.9899
2.15	-	2.20	-0.9431	-0.9869
2.20	-	2.25	-0.7348	-0.9469
2.25	-	2.30	-0.7293	-1.0210
2.30	-	2.35	-0.6816	-1.1032
2.35	-	2.40	-0.7722	-1.1204
2.40	-	2.45	-0.7774	-1.1615
2.45	-	2.50	-0.7510	-1.2399
2.50	-	2.55	-0.9800	-1.2886
2.55	-	2.60	-0.8909	-1.3311

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 3 (continuación).

Intervalo de ángulo (grados)			Energía (u.a.)			
			H-ERX (b)	H-ERMN (b)	H-ERX (c)	H-ERMN (c)
0	-	5	6.0000	6.0000	6.0000	6.0000
5	-	10	6.0000	6.0000	6.0000	6.0000
10	-	15	6.0000	6.0000	6.0000	6.0000
15	-	20	6.0000	6.0000	6.0000	6.0000
20	-	25	6.0000	6.0000	6.0000	6.0000
25	-	30	6.0000	6.0000	6.0000	6.0000
30	-	35	6.0000	6.0000	6.0000	6.0000
35	-	40	6.0000	6.0000	6.0000	6.0000
40	-	45	6.0000	6.0000	6.0000	6.0000
45	-	50	6.0000	6.0000	6.0000	6.0000
50	-	55	6.0000	6.0000	6.0000	6.0000
55	-	60	6.0000	6.0000	6.0000	4.0078
60	-	65	6.0000	6.0000	6.0000	1.5306
65	-	70	6.0000	5.6173	4.4397	1.1282
70	-	75	6.0000	4.5758	2.9734	1.3325
75	-	80	6.0000	2.5648	0.0959	0.3711
80	-	85	2.3603	1.4270	-1.2740	-1.2027
85	-	90	0.3622	0.2049	-0.4123	-0.8098
90	-	95	-0.5709	-0.4967	0.6785	-0.2291
95	-	100	-0.7825	-0.8418	0.1402	-0.3599
100	-	105	-0.0564	-0.0867	0.1311	-0.3178
105	-	110	1.3794	0.2690	-0.0712	-0.2299
110	-	115	0.6330	-0.1055	-0.0414	-0.2247
115	-	120	0.5547	-0.3457	-0.1036	-0.2613
120	-	125	0.5616	-0.4121	0.0959	-0.2026
125	-	130	0.4631	-0.4606	0.1492	-0.3101
130	-	135	0.2301	-0.5861	-0.0527	-0.5724
135	-	140	-0.0262	-0.7478	-0.6396	-0.8287
140	-	145	-0.2847	-0.9520	-1.2596	-1.0377
145	-	150	-0.7270	-1.0985	-1.5982	-1.1215
150	-	155	-1.3171	-1.2059	-1.6045	-1.0362
155	-	160	-1.7601	-1.2234	-1.2069	-0.8110
160	-	165	-1.8281	-1.1256	-0.5043	-0.4755
165	-	170	-1.4399	-0.8214	0.2602	-0.0383
170	-	175	-0.6662	-0.3165	0.9232	0.5431
175	-	180	0.7261	0.8088	2.0726	1.6438

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUEENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

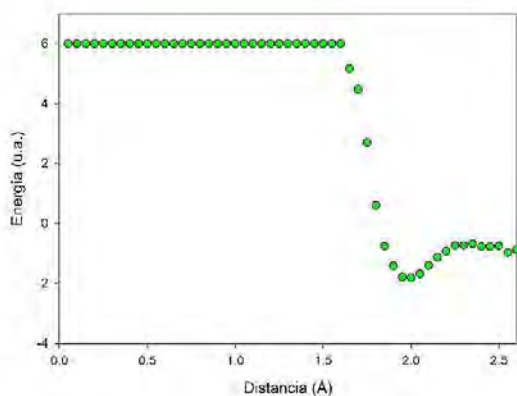
Tabla 3 (continuación).

Intervalo de ángulo diedro (grados)			Energía (u.a.)					
			H-ERX (d)	H-ERMN (d)	H-ERX (e)	H-ERMN (e)	H-ERX (f)	H-ERMN (f)
-180	-	-170	1.0840	0.6644	-0.5485	-0.9987	0.7508	1.0851
-170	-	-160	0.7847	0.3505	-0.7174	-1.1485	0.9534	1.0438
-160	-	-150	0.5616	-0.0494	-0.6981	-0.8592	0.8844	0.9437
-150	-	-140	-0.2143	-0.4193	0.0249	-0.2764	0.7934	0.8329
-140	-	-130	-0.8849	-0.7448	0.4694	0.2069	0.8939	0.6819
-130	-	-120	-1.2882	-0.9858	0.8198	0.5220	1.0610	0.4437
-120	-	-110	-1.2674	-1.1923	0.8844	0.7610	0.9844	0.1114
-110	-	-100	-1.0724	-1.0021	0.9232	0.9854	1.0840	0.5964
-100	-	-90	-1.0506	-1.2420	1.0057	1.1092	1.1316	0.5669
-90	-	-80	-0.4481	-1.0677	1.1316	1.2468	0.9331	0.0653
-80	-	-70	0.7180	-0.0432	1.3042	1.3898	-0.3450	-0.6895
-70	-	-60	0.8655	0.6133	1.4112	1.4169	-0.2237	-0.7460
-60	-	-50	0.6941	0.8326	1.6263	1.4213	0.0049	-0.2260
-50	-	-40	0.7508	0.9297	1.6465	1.3412	-0.3394	-0.2419
-40	-	-30	0.5210	0.9733	1.4274	1.1712	-0.7000	-0.5581
-30	-	-20	0.4821	0.8263	1.2208	0.8452	-0.8751	-0.5806
-20	-	-10	0.3234	0.3810	0.4025	0.4085	-0.4506	-0.3270
-10	-	0	-0.2906	-0.1162	-0.3110	0.0049	-0.0452	0.0335
0	-	10	-0.6981	-0.4189	-0.5598	-0.2611	0.4203	0.3430
10	-	20	-0.7606	-0.4498	-0.4123	-0.3006	0.7100	0.4908
20	-	30	-0.5439	-0.1445	0.0089	-0.0759	0.5144	0.5274
30	-	40	0.1134	0.4425	0.8021	0.3460	0.3509	0.3824
40	-	50	0.6257	0.9168	1.1689	0.9574	0.2808	0.1275
50	-	60	0.8378	0.8866	1.5493	1.3540	-0.1621	-0.1431
60	-	70	1.1689	0.8505	1.3639	1.3115	-0.3366	-0.3389
70	-	80	1.1689	0.9534	1.1689	1.2384	-0.7842	-0.4950
80	-	90	1.1689	0.9070	0.8378	1.1333	-0.9232	-0.5706
90	-	100	1.0057	0.9036	0.8749	0.9359	-0.8882	-0.5756
100	-	110	1.0165	0.9813	0.5896	0.7973	-0.5439	-0.4620
110	-	120	1.0385	0.7578	0.4203	0.5315	-0.1487	-0.3523
120	-	130	1.1945	1.1032	0.2450	0.2433	-0.0338	-0.1205
130	-	140	1.5680	1.2833	-0.2547	-0.1731	0.0873	0.2492
140	-	150	1.4953	1.3389	-0.9984	-0.6334	0.4757	0.5518
150	-	160	1.3952	1.3065	-1.5118	-1.0332	0.5967	0.8225
160	-	170	1.4112	1.1586	-1.3472	-1.1864	0.7020	0.9629
170	-	180	1.1316	0.9737	-0.7788	-1.0517	0.5411	1.0833

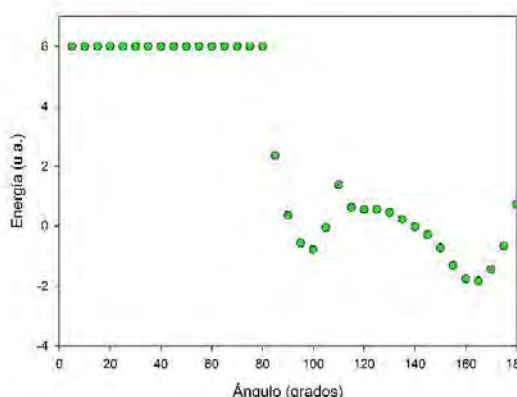
ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

H-ERX

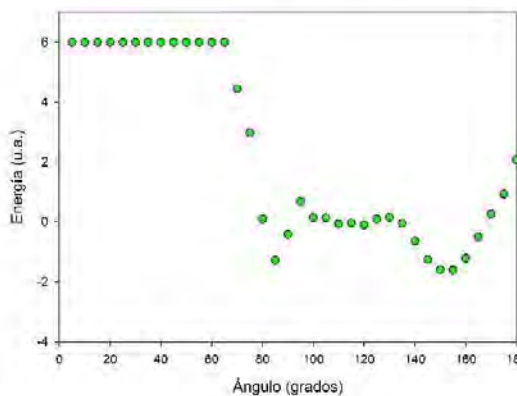
(a)



(b)

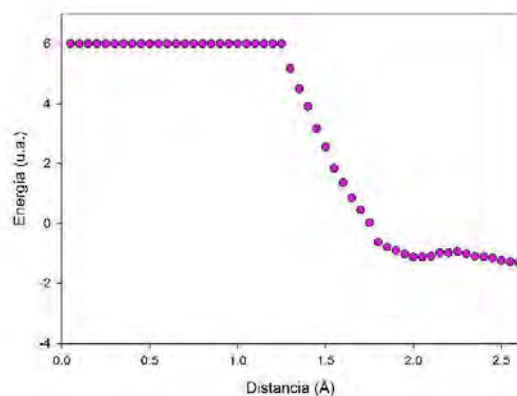


(c)

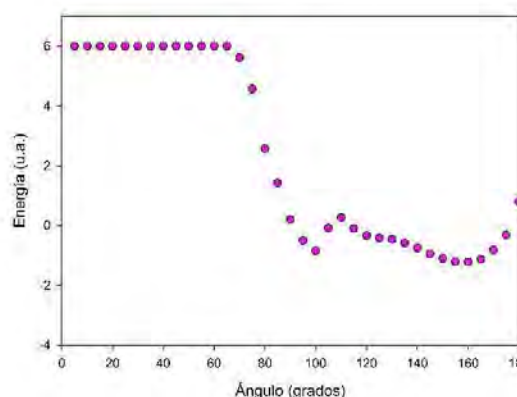


H-ERMN

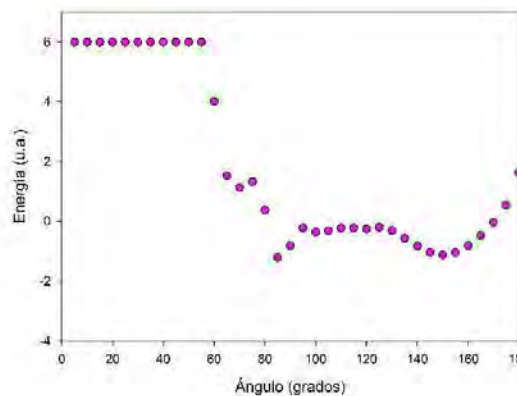
(a)



(b)



(c)



ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

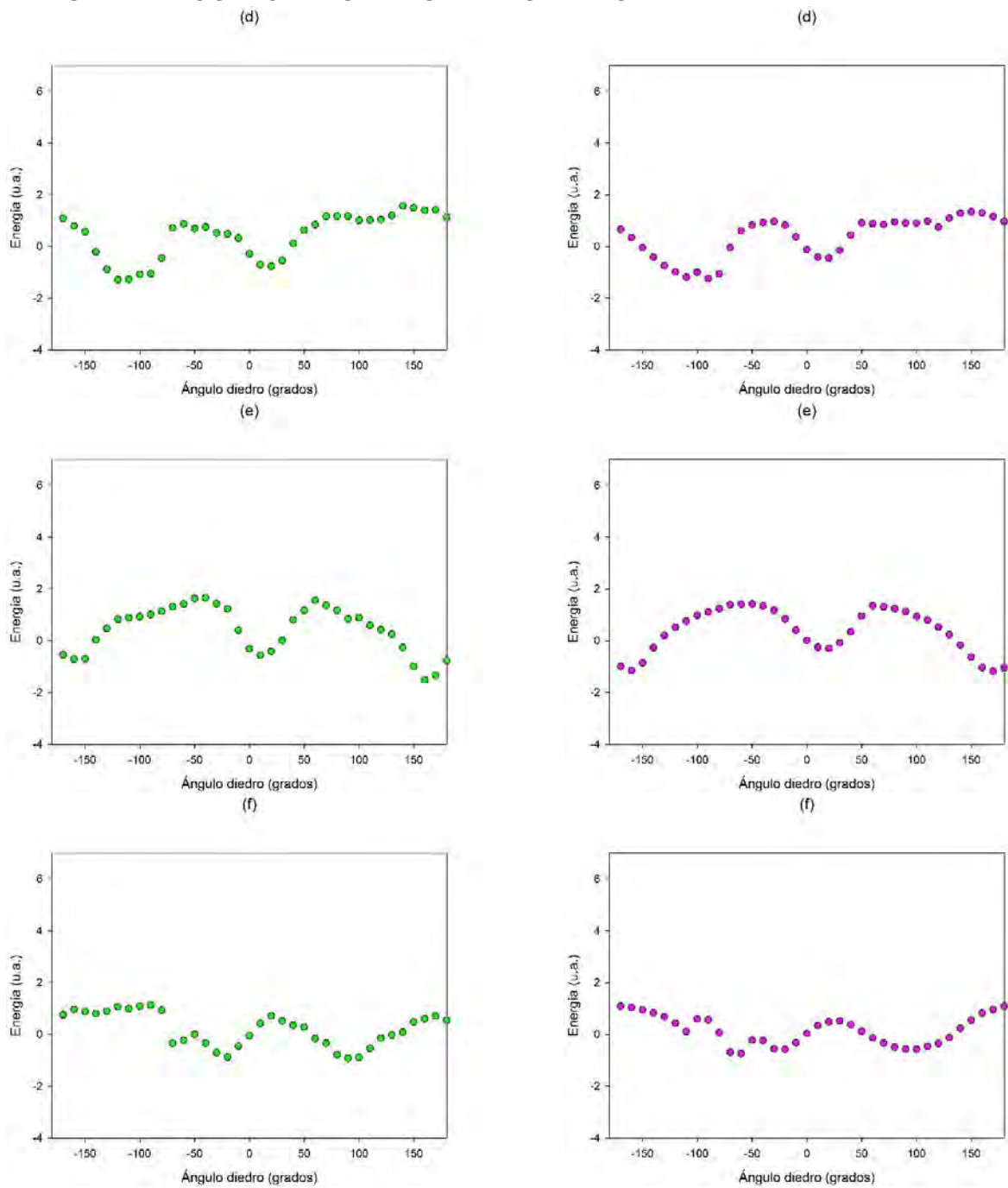
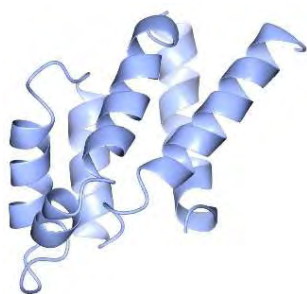


Figura 2. Componentes de energía asociados a los parámetros establecidos en los potenciales H-ERX y H-ERMN. (a) la distancia hidrógeno-oxígeno, (b) el ángulo al átomo de hidrógeno, (c) el ángulo al átomo de oxígeno (d) el ángulo diedro que corresponde a la rotación del enlace carbono-oxígeno, (e) el ángulo diedro que corresponde a la rotación del puente de hidrógeno y (f) el ángulo diedro que corresponde a la rotación del enlace hidrógeno-nitrógeno.

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”



1abv_



13.088



1ah9_



7.692



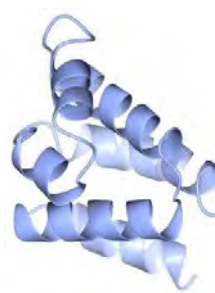
1egxA



2.727



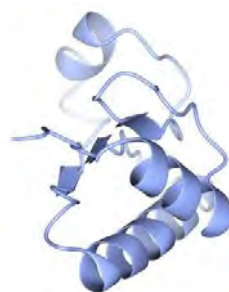
1fadA



3.611

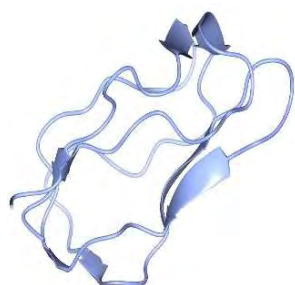


1fo5A



3.97

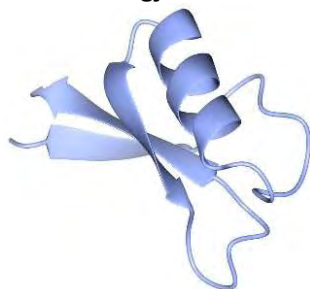
ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"



1gxA



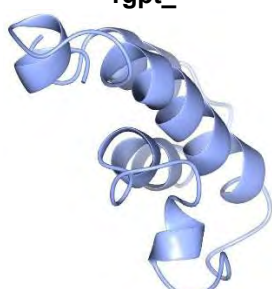
6.286



1gpt_



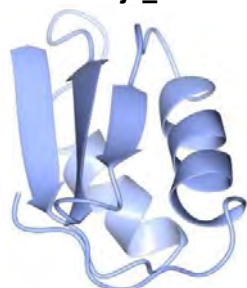
7.824



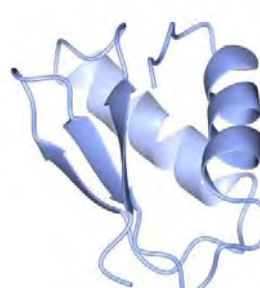
1kjs_



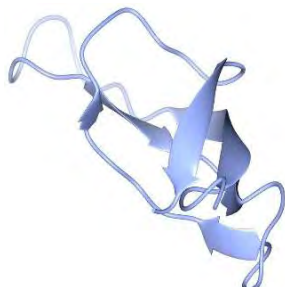
6.056



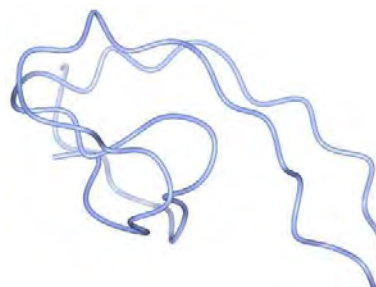
1kviA



2.63



1ne3A



9.361

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUEENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

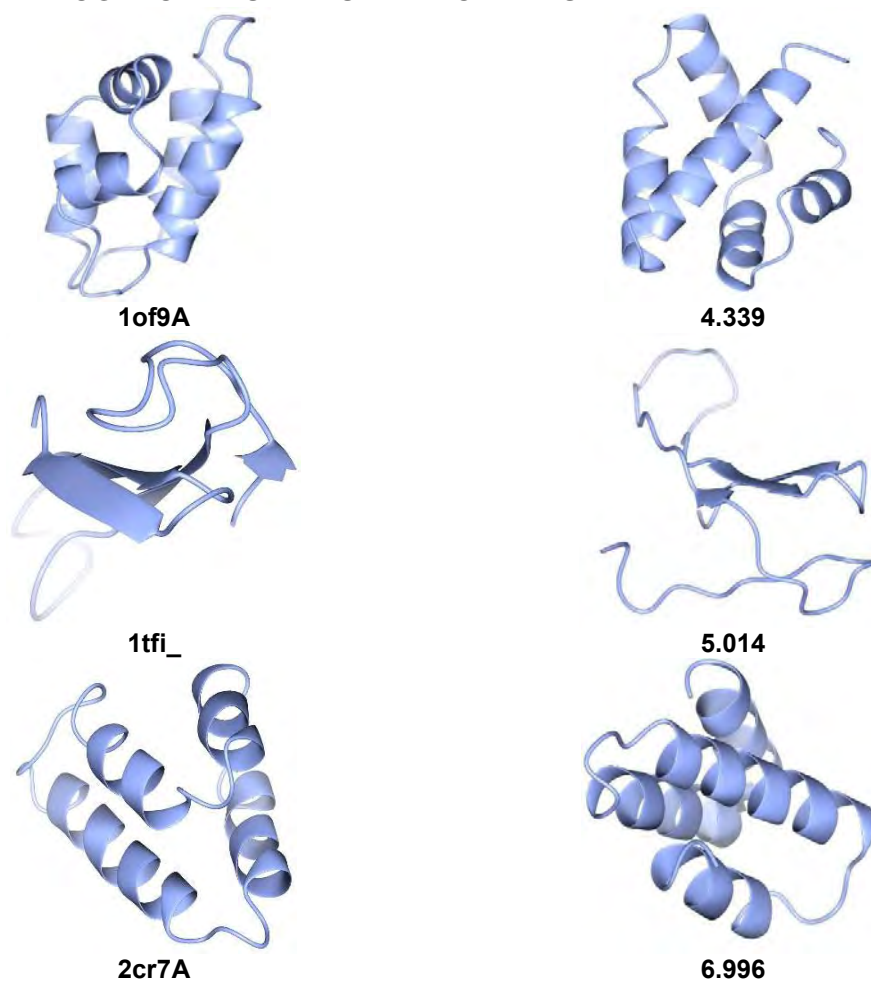
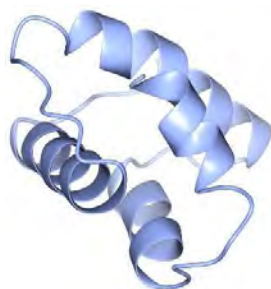
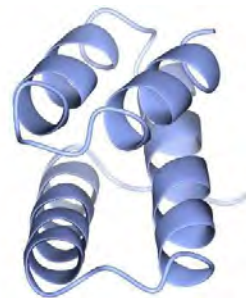


Figura 3. Modelos identificados como estructura nativa por el potencial H-ERX. El número debajo del modelo indica su desviación cuadrática media con respecto a la estructura nativa.

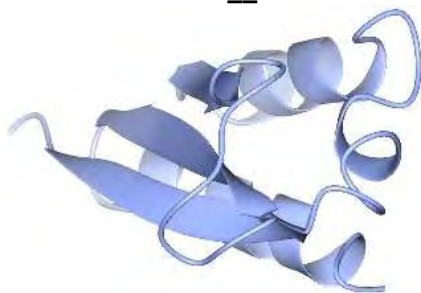
ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”



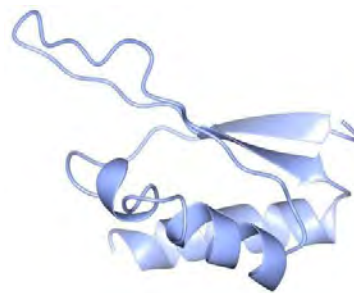
1af7_



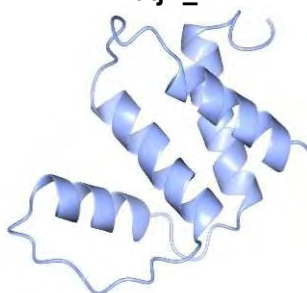
11.021



1dtjA_



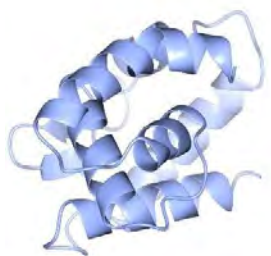
2.985



1hbka



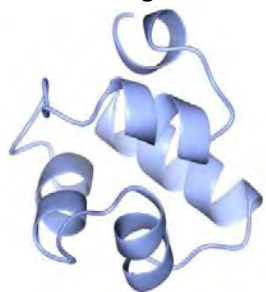
12.220



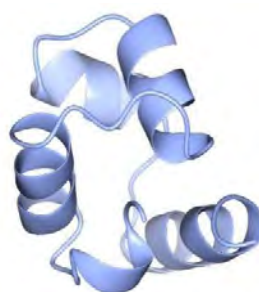
1orgA



2.613



1r69_



2.399

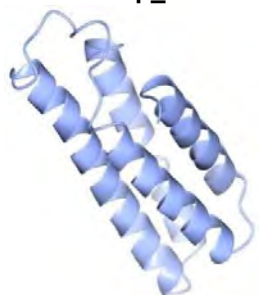
ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"



1sfp_



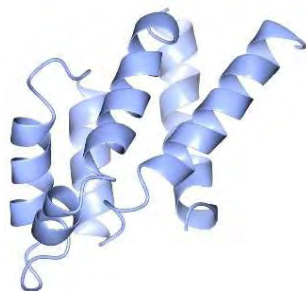
5.134



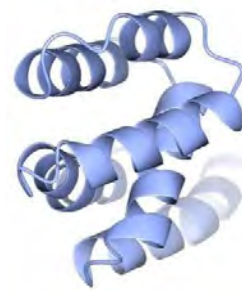
256bA



3.172



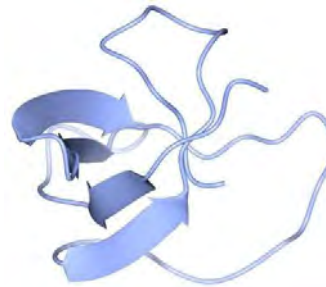
1abv_



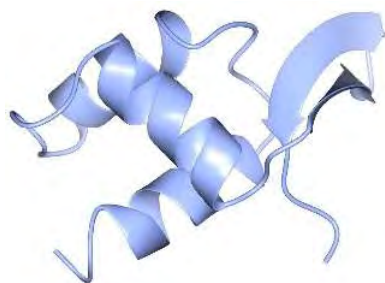
12.78



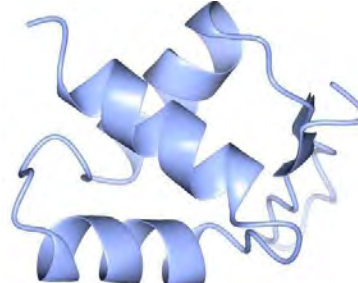
1ah9_



3.271



1aoy_

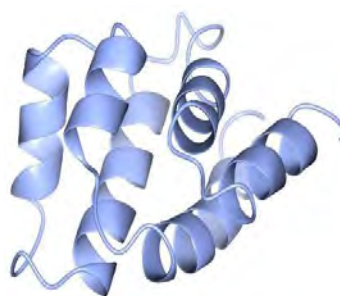


5.127

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"



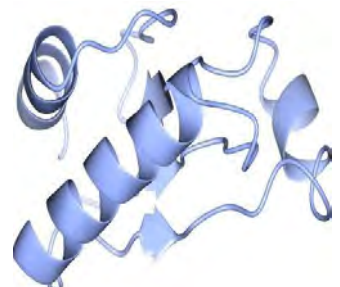
1fadA



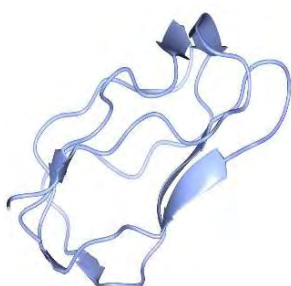
3.705



1fo5A



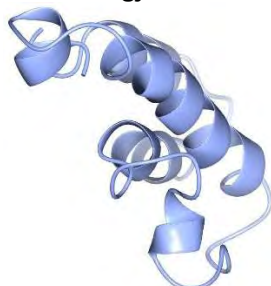
3.97



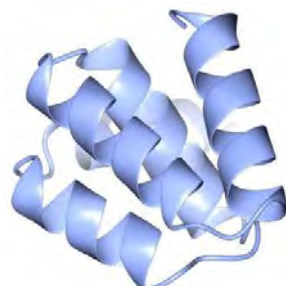
1gjxA



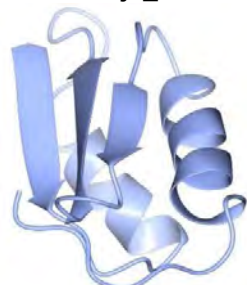
6.286



1kjs_



11.019



1kviA



2.63

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

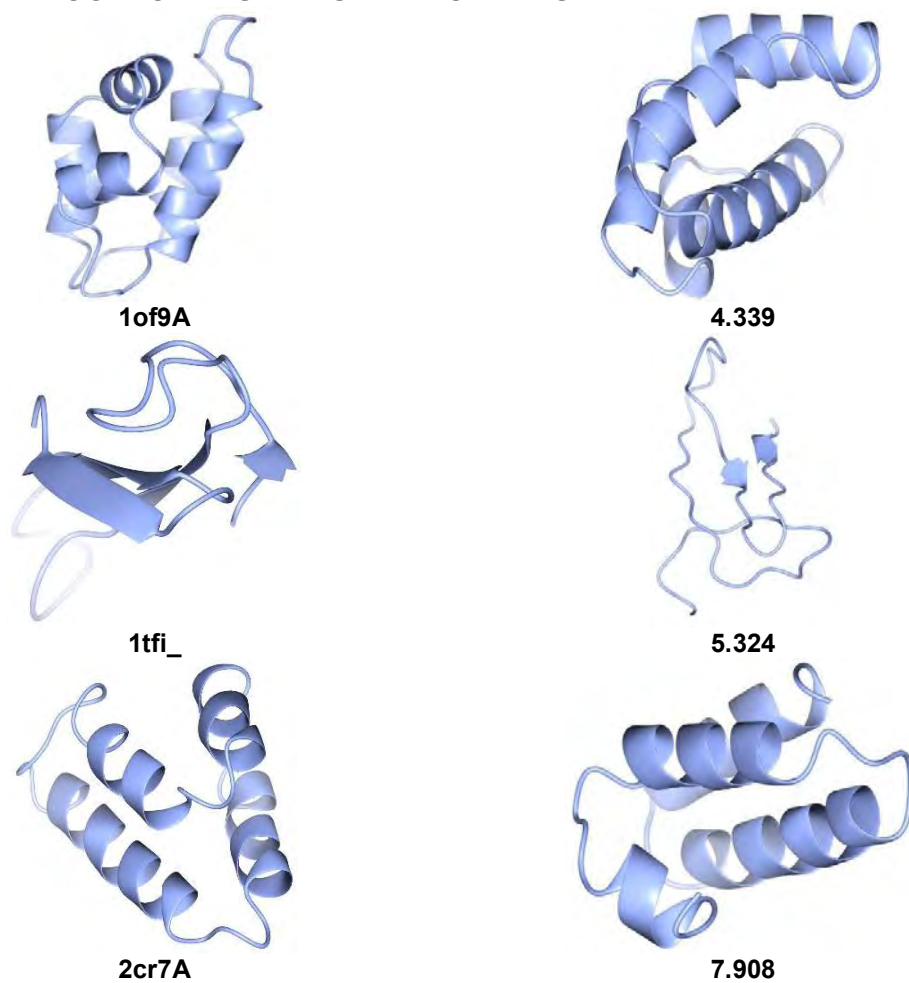


Figura 4. Modelos identificados como estructura nativa por el potencial H-ERMN. El número debajo del modelo indica su desviación cuadrática media con respecto a la estructura nativa.

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 4. Códigos PDB de las 1981 proteínas que formaron parte del grupo ERX*. La letra que se encuentra frente al código PDB es el identificador de la cadena estudiada en el caso de que este fuese distinto de "A".

1A12	1A3A	1A4U	1A7U	1AE9	1AGJ	1AGQ	1AIL
1AK1	1AKO	1AMX	1ATZ	1AYI	1B2P	1B65	1BGF
1BHE	1BJ7	1BK7	1BKP	1BKR	1BQC	1BYR	1BZ4
1C02	1C3C	1C3P	1C44	1C48	1CEM	1CEO	1CEX
1CHD	1CJD	1CQ3	1CQY	1CRW_G	1CWY	1CY9	1D4T
1D4T_B	1DD3	1DD3_C	1DEU	1DF4	1DQ0	1DTS	1DUN
1DVO	1E15	1E4F_T	1E5M	1EDG	1EDQ	1EG4	1EG4_P
1EL6	1ELK	1ERZ	1ES5	1ETX	1EW4	1EYH	1EZG
1F2T	1F2T_B	1F41	1F46	1F4Q	1F5Z	1F60	1F60_B
1FAZ	1FCG	1FCQ	1FE6	1FL0	1FO9	1FPO	1FSF
1FTR	1G0S	1G2Q	1G4E	1G61	1G8A	1G8Q	1G9W
1G9W_B	1GAK	1GCU	1GK6	1GO3_E	1GO3_F	1GOU	1GPP
1GQN	1GQZ	1GS9	1GSM	1GSO	1GVJ	1GXN	1GY7
1GYU	1GYV	1H03_P	1H13	1H1N	1H4A_X	1H4Y	1H6K
1H6K_X	1H6T	1H7E	1H99	1HM5	1HQK	1HXH	1I07
1I1J	1I2T	1I7K	1I9Y	1IDP	1IGQ	1IHJ	1IHJ_C
1IIB	1IJB	1IJQ	1IJY	1INL	1IRQ	1IU1	1IU8
1J1I	1J23	1J27	1J3A	1J48	1J7G	1JB3	1JCD
1JDH_B	1JFL	1JHS	1JVV	1JY2_N	1JY2_O	1JY2_P	1JYH
1K1B	1K30	1K6A	1KHI	1KLX	1KMT	1KMZ	1KNG
1KOE	1KPT	1KT9	1KU8	1KU8_B	1KWB_B	1KYF	1KYF_P
1KZF	1KZQ	1L2H	1L3K	1L6P	1L7A	1L7R	1LBW
1LM5	1LN4	1LOU	1LTU	1LU4	1LU9	1LZL	1M0Z
1M1Z	1M3S	1M45	1M45_B	1M4V	1MBY	1MF7	1MG6
1MG7	1MHN	1MIX	1MJS	1MK4	1MKB	1MKK	1MN8
1MR7	1MS3	1MSI	1MTP	1MTP_B	1MTZ	1MY7	1MZL
1N1J	1N1J_B	1N7O	1N93_X	1NC5	1NG2	1NG5	1NG6
1NJ4	1NKD	1NOG	1NRV	1NTY	1NWA	1NXM	1O0W
1O8X	1O9Z	1OA4	1OAI	1OAI_B	1OBQ	1OCK	1OGH
1OGM_X	1OI7	1OJQ	1OJX	1OK7	1OK7_C	1ONR	1OQW
1OR7_C	1OTK	1OUV	1OW1	1OX3	1OXJ	1OZ9	1P3C
1P9H	1P9I	1PBW	1PCF	1PDO	1PE9	1PEF	1PEV
1PGV	1PM4	1PQE	1PXZ	1PZC	1Q5Z	1Q8I	1QB5_D
1QCX	1QNT	1QO7	1QTO	1QU1	1QWD	1QWK	1QYA
1QZ1	1QZN	1R0M	1R0V	1R1T	1R29	1R5Z	1R62
1R77	1R7J	1R8O	1R8O_B	1R9H	1R9W	1RC9	1REG_X
1RFY	1RGX	1RI6	1RJ1	1RL0	1RW0	1RWR	1RWZ
1S2T	1SAU	1SDO	1SGV	1SGZ	1SH8	1SJY	1SLL
1SMX	1SNT	1SQE	1SUR	1SUU	1T06	1T1D	1T1J
1T1U	1T2I	1T3Y	1T4D	1T5O	1T5R	1T6F	1T7H
1TC5	1TCZ	1TGR	1TH7	1TJ6	1TJE	1TJV	1TK1
1TKI	1TKS	1TLU	1TM2	1TP5	1TP5_B	1TP6	1TQJ
1TUA	1TUD	1TUO	1TVN	1TW9	1TWU	1TY0	1TZV
1U07	1U09	1U1S	1U2K	1U53	1U79	1UAI	1UAJ
1UAN	1UEB	1UEK	1UFB	1UGN	1UGQ	1UGQ_B	1UH9
1UII	1UIR	1UIU	1UJ8	1UJN	1UKF	1ULK	1ULN
1ULR	1UNP	1UOK	1UOY	1USE	1USG	1USM	1UXZ
1UZ3	1V05	1V0S	1V1H	1V6T	1V8E	1V8H	1V8I
1VAV	1VAX	1VC1	1VDH	1VE2	1VF8	1VH5	1VJK
1VJS	1VKA	1VL2	1VL4	1VPK	1VPN	1VQ3	1VR6
1W30	1W41	1W5R	1W7B	1W8Z	1WDJ	1WJ9	1WKA
1WKO	1WKU	1WKX	1WLG	1WLT	1WLY	1WLZ	1WMH
1WMH_B	1MMW	1WNH	1WNY	1WOH	1WOS	1WOU	1WTJ
1WU9	1WVH	1WYC	1WZ3	1X13	1X10	1X2I	1X56

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 4 (continuación).

1X6Q	1X91	1XDZ	1XFK	1XIX	1XKR	1XLY	1XQO
1XSZ	1XTE	1XX2	1XXO	1XYF	1Y2T	1Y7Y	1Y9U
1YAC	1YBI	1YDI	1YDI_B	1YGA	1YHH	1YIF	1YJP
1YKS	1YN3	1YNB	1YO3	1YP5	1YPF	1YPY	1YRW
1YTL	1YTQ	1YU5_X	1YW5	1YXY	1YYH	1YZ1	1YZF
1YZM	1Z21	1Z3Y	1Z4E	1Z96	1ZCJ	1ZCU	1ZD8
1ZEQ_X	1ZHV	1ZKR	1ZMM	1ZRS	1ZVA	1ZVB	1ZVG
1ZVT	1ZXT	1ZZG	2A10	2A35	2A4A	2A6Z	2A8F
2AAG	2AHF	2AHN	2AKA	2AKA_B	2AKA_L	2AN1	2AO9
2B0A	2B0J	2B1E	2B1K	2B1L	2B2H	2B3M	2B8I
2B9V	2BAY	2BCE	2BCM_B	2BF5	2BJQ	2BK8	2BN3
2BN3_B	2BOE_X	2BPT	2BPT_B	2BQX	2BT2	2BV9	2BZV
2C2X	2C46	2C61	2CAR	2CG7	2CGH	2CGQ	2CHC
2CI3	2CIU	2CJJ	2CKX	2CL3	2CO3	2COV_D	2CU5
2CUO	2CWC	2CWL	2CWR	2CWX	2CX5	2CXC	2D4P
2D59	2D5W	2D5W_C	2D5W_D	2D68	2D7J	2D8E	2DBB
2DKA	2DP9	2DPL	2DQW	2DST	2DVK	2DYJ	2DYO_B
2DYU	2E01	2E0T	2E10	2E3D	2E3H	2E3Z	2E5B
2E7A	2E7V	2E8F	2E8G	2E9Y	2EBB	2EBE	2ECE
2ECR	2EEN	2EEY	2EGJ	2EGU	2EHG	2EIF	2EJ8
2EKC	2ENG	2EPI	2ERF	2ERW	2ETX	2EX0	2F1N
2F1S	2F51	2F5G	2F6E	2F6L	2FB5	2FBN	2FBQ
2FC3	2FD5	2FE7	2FEZ	2FH7	2FHZ	2FHZ_B	2FI9
2FJ8	2FJR	2FJZ	2FK9	2FL4	2FL7	2FLU_P	2FLU_X
2FN9	2FQ3	2FR2	2FRG_P	2FUK	2FVH	2FW7	2FXQ
2FZP	2G30	2G30_P	2G30_S	2G40	2G5X	2G69	2G7O
2GAI	2GAS	2GDG	2GEC	2GKG	2GKT_I	2GMY	2GOM
2GQ0	2GQV	2GR8	2GRR	2GRR_B	2GT1	2GTD	2GUB
2GXG	2GZ4	2GZ6	2H0Q	2H2R	2H2Z	2H3L	2H7O
2H8E	2H8O	2HC8	2HDW	2HE7	2HJV	2HLR	2HLY
2HPL	2HPL_B	2HQ4	2HQE	2HVM	2HWX	2HY5	2HY5_B
2HY5_C	2HY7	2HYX	2I49	2I4L	2I6V	2I9I	2IB0
2IBL	2IC6	2IG8	2IGD	2IIA	2ILR	2IM9	2IP2
2IPR	2IQ5	2IRU	2IU1	2IU5	2IUM	2IVY	2IXM
2IY9	2IZ6	2J2J	2J5Y	2J6B	2J71	2J7V	2J8B
2J9W	2JCP	2JEM	2JJF	2JK9	2JK9_B	2JLJ	2LIS
2NML	2NNU	2NNU_B	2NPT_B	2NRR	2NV0	2NWD_X	2NX2
2NX4	2NXC	2NYC	2O2K	2O4X	2O4X_B	2O5U	2O6X
2O70	2O7M	2O9V	2O9V_B	2OCH	2OCT	2OEB	2OEI
2OEI_B	2OF3	2OG3	2OG4	2OGQ	2OHW	2OKT	2OLX
2OMP	2OMZ	2OMZ_B	2ON9	2ONW_X	2ONX	2OR2	2OU1
2OVA	2OY7	2P1G	2P2O	2P3E	2P3W	2P4H_X	2P4X
2P52	2P5D	2P5K	2P65	2P6O	2PA6	2PB7	2PBO
2PBQ	2PE3	2PET	2PFB	2PGE	2PLC	2PN8	2PND
2PO4	2PPO	2PSF	2PST_X	2PTH	2PV2	2PV2_E	2Q5X
2Q80	2Q8R_E	2Q9V	2QBW	2QBW_B	2QEV	2QHT	2QIY
2QIY_C	2QN4	2QOL	2QPW	2QQ5	2QT4	2QVK	2QXV
2QXV_B	2R0F	2R2Y	2R6Q	2R6U	2R77	2RB8	2RCI
2RER	2RF1	2RFA	2RH3	2RIK	2RJD	2RK3	2RK5
2RKQ	2UWB	2UWI	2UYO	2V75	2V78	2VC8	2VE8
2VFY	2VGA	2VH7	2VOK	2VQ4	2VT1	2VT1_B	2VWV
2VXW	2VY8	2W0G	2W0I	2W1R	2W1V	2W2A	2W4E
2W61	2W6A	2WB3	2WFV	2WFV_B	2WFW	2WJ5	2WMF
2WN4	2WNK	2WP7	2WPG	2WUJ	2WUX	2WWE	2WWX
2WWX_B	2WZ9	2X1Q	2X35	2X36	2X3M	2X4J	2X4L
2X5P	2X5Y	2X8X_X	2XBG	2XF7	2XHF	2XJ4	2XMZ
2XPP	2XPP_B	2XTY	2XWS	2XWX	2XX6	2XXN	2XXN_B

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 4 (continuación).

2XZE	2XZE_Q	2Y0T	2Y2C	2Y2Z	2Y32	2Y3K	2Y3V
2Y3W	2Y72	2Y7S	2Y7Y	2Y9F	2YGS	2YLB	2YLE
2YLE_B	2YML	2YMU	2YNY	2YVA	2YVS	2YYV	2Z0M
2Z0T	2Z14	2Z1E	2Z37	2Z43	2Z84	2ZA7	2ZCN
2ZCO	2ZF3	2ZFY	2ZGL	2ZJD	2ZJD_B	2ZQ5	2ZQE
2ZU1	2ZVY	2ZXJ	3A0Y	3A2Z	3A4C	3A54	3A57
3A5P	3A7L	3AAP	3ABD	3ABD_X	3ADO	3ADY	3AEH
3AFF	3AG7	3AGY	3AGY_C	3AKJ	3AMC	3APQ	3AQ2
3AQY	3AU4	3AU4_B	3AUB	3B02	3B4D	3B7H	3B8F
3BA1	3BB7	3BBC	3BCI	3BF7	3BHS	3BOI	3BOK
3BOR	3BQE	3BQO	3BQO_B	3BQQ_C	3BS4	3BS4_B	3BZG
3BZT	3BZZ	3BZZ_B	3C1D	3C4S	3C57	3C5V	3C8I
3C8P	3C8X	3CA7	3CAI	3CE7	3CGI	3CI9	3CIV
3CJ1	3CJW	3CKF	3CL6	3CNU	3CO1	3COO	3COU
3CQT	3CRM	3CSG	3CSP	3CSR	3CT6	3CTB	3CWV
3CX2	3CYP_B	3CZZ	3D1B	3D2A	3D79	3D9X	3DA0
3DAC_M	3DAC_P	3DAL	3DEL_B	3DFG	3DFQ	3DG1	3DGJ
3DGP	3DGP_B	3DKM	3DLM	3DLQ_I	3DLQ_R	3DMS	3DO8
3DQG	3DRF	3DRF_B	3DRZ	3DS4	3DS4_T	3DSH	3DU1_X
3DWV	3DXE	3DXE_B	3E17	3E1R	3E1R_C	3E21	3E3M
3E4H	3E57	3E7H	3E7P	3E7R_L	3E96	3ECH	3ECH_C
3ECY	3EFY	3EG4	3EJ9	3EJ9_B	3EJF	3EJG	3EMF
3EMI	3ENB	3ENU	3EO5	3EOI	3ER6	3ERB	3ERJ
3ETZ	3EW1	3EWI	3EXR	3EXV	3EYE	3F2G	3F6F
3FAU	3FB9	3FFV	3FG7	3FH2	3FKE	3FLG	3FN7
3FPN	3FPN_B	3FPO	3FRR	3FSO	3FTD	3FTJ	3FTK
3FTL	3FVA	3FX7	3FY3	3FZE	3G19	3G19_C	3G1S
3G39	3G40	3G98	3GAX	3GD0	3GDM	3GGY	3GMG
3GMS	3GPG	3GRE	3GRH	3GSZ	3GV3	3GW3	3H04
3H2G	3H5L	3H6J	3HA9	3HAK	3HC7	3HD4	3HDE
3HE5	3HE5_B	3HFO	3HGU	3HH7	3HHP	3HJ2	3HKM
3HKX	3HMT	3HNY_M	3HO7	3HOL	3HR0	3HR8	3HRO
3HS3	3HS8	3HS8_P	3HYD	3HZ8	3I0C	3I2Z_B	3I40
3I8S	3ICH	3ID1	3ID4	3IDW	3ILC	3ILS	3IM1
3IOY	3IPF	3ITY	3IU0	3IU5	3IVV	3IVV_D	3IXR
3JRR	3JSN	3JSY	3JU0	3JXG	3JXO	3JZZ	3K1H
3K3V	3K6U	3K6Y	3K86	3K8U	3K8W	3KBE	3KCI
3KCW	3KDG	3KF6	3KF6_B	3KG4	3KG9	3KGK	3KH8
3KJT	3KP8	3KQ5	3KSN	3KT2	3KT9	3KUG	3KZD
3L18	3L32	3L3B	3L40	3L78	3L7O	3L8U	3L9U
3LAA	3LE4	3LET	3LFG	3LFH	3LFJ	3LFP	3LIG
3LLB	3LOZ	3LPZ	3LS0	3LTJ	3LWG	3LWT_X	3LY7
3M5B	3M5R	3M66	3M6N	3M7D	3M8J	3M91	3M91_B
3M9J	3M9Q	3MAB	3MD4	3MD5	3ME8	3MEW	3MGK
3MGM	3ML3	3MM4	3MQ0	3MRE_B	3MRE_P	3MX7	3N0K
3N11	3N2T	3N3E	3N77	3NDD	3NDD_B	3NE0	3NFL
3NFL_E	3NFT	3NGG	3NJD	3NJE	3NJJ	3NJJ_B	3NPH_B
3NPV_B	3NR5	3NRP	3NSO	3NTK	3NUI	3NVG	3O0A
3O1Q	3O1Z	3O2H	3O2H_B	3O48	3O5E	3O5P	3O7I
3O81	3O90	3OBQ	3OBQ_B	3OGG	3OI9	3OL4	3ON9
3ONH	3ONJ	3OSX	3OT7	3P1W	3P4L	3P6J	3P8T
3P9A	3PBN	3PC6	3PD7	3PG4	3PH9	3PHG	3PHS
3PID	3PIW	3PKV	3PM2	3PNO	3PS0	3PST	3PTE
3PTL	3PUB	3PVE	3PXM	3PZ9	3PZF	3PZZ	3Q0H
3Q1C	3Q49_B	3Q49_C	3Q6B	3Q6S	3Q6S_E	3Q8T	3Q9D
3Q9V	3QAS_B	3QBP	3QC7	3QF2	3QH4	3QHP	3QIT
3QMQ	3QNE	3QSQ	3QT5	3QUW	3QWG	3QYJ	3R1W

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUEENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 4 (continuación).

3R42	3R42_B	3R4Y	3R87	3RDJ	3RDY	3RFF	3RFI
3RGI	3RGR	3RJP	3RJS	3RK6	3RKC	3RLS	3RMH
3RNQ	3RNQ_B	3RPP	3RQ9	3RRS	3RT2	3RV1	3RVC
3RZY	3S0A	3S60	3SD4	3SDB	3SGS	3SH4	3SHT
3SIH	3SK9	3SO6	3SO6_Q	3SQF	3SR3	3SRI	3SRI_B
3SSW_R	3SUK	3SUM	3SWY	3SXU	3SXU_B	3SXU_C	3SXW
3SXY	3SZ7	3SZH	3T0H	3T30_B	3T3K	3T43	3T47
3T4C	3T7H	3TC2	3TCH	3TCQ	3TCV	3TDN	3TE8
3TEF	3TG7	3TIP	3TOW	3TP3	3TPA	3TQT	3TU3
3TU3_B	3TUA	3TVT	3TVT_B	3TXS	3U0V	3U28	3U28_B
3U28_C	3U3B	3U80	3UAH	3UC9	3UCI	3UCS	3UCS_C
3UEK	3UFB	3UGU	3UID	3UN7	3UR6	3UR8	3US6
3UTK	3UTM	3UTM_C	3UV0	3UV9	3V1Q	3V32_B	3V46
3V4G	3V55	3V6G	3VBA	3VBC	3VDJ	3VDM	3VFI
3VHO	3VJ8	3VJZ	3VK0	3VL1	3VN0	3VN5	3VOQ
3VPZ	3VQL	3VSR	3VVM	3VVV	3VZ9_B	3VZ9_D	3VZH
3W4Q	3W90	3W9U	3WA1	3WAI	3WDF	3WFI	3WHO
3WHR	3WHR_B	3WIO	3WJE	3WMI	3WMI_B	3WP5	3WP8
3WP9	3WPQ	3WVZ	3WWF	3WX4	3WYD	3ZBD	3ZCD
3ZHI	3ZIB	3ZIH	3ZIT	3ZK0	3ZLC	3ZO8	3ZRD
3ZRXX	3ZSL	3ZSS	3ZVQ	3ZVQ_B	3ZXY	3ZYL	3ZZY
3ZZY_C	4A5U	4A5U_B	4A8X	4A8X_B	4A8X_C	4ACO	4AD1
4AE7	4AFV	4AGH	4AGK	4AK9	4AM1	4ANN	4AOM
4AOM_T	4AVR	4AY0	4AZZ	4B2F	4B4H	4B6G	4B6I
4B89	4B91	4B9G	4B9I	4BFH	4BI8	4BI8_B	4BK6
4BL7	4BL7_B	4BMW	4BOU	4BT9	4BT9_C	4BVK	4BXP
4C2E	4C5E_E	4C8X	4CAY	4CAY_B	4CAY_C	4CBE	4CFI
4CIL	4CKK	4CP6	4CQ4	4CRH	4CVD	4CYA	4CZ5
4CZI	4D0K	4D0K_B	4D5R	4D8L	4D90	4D9S	4DAM
4DB6	4DD5	4DDP	4DEX	4DEX_B	4DJG	4DK2	4DLH
4DM7	4DMV	4DOT	4DQ7	4DQN	4DT4	4DYQ	4E1P
4E2A	4E2U	4E40	4E5R	4E6S	4E9E	4E9L	4EFO
4EGT	4EJQ	4EJR	4EKX_B	4EL6	4ELL	4ELO	4EM8
4EO0	4EQ6	4EQA	4EQA_C	4EQY	4ERN	4ERY	4ERY_D
4ES1	4ES9	4EVM	4EYC	4EZA	4F01	4F3J	4F80
4F8K	4FBR	4FCC	4FD6	4FD9	4FGQ	4FH3	4FHR
4FHR_B	4FML	4FNV	4FQN	4FW1	4FZO	4G08	4G2E
4G3N	4G30	4G54	4G5X	4G6C	4G6T	4G6T_B	4G7X
4G7X_B	4G94_B	4G9M	4GA0	4GCO	4GDO	4GEI	4GF3
4GF3_B	4GFK	4GGO	4GLF	4GMQ	4GQM	4GRD	4GVB
4GVB_B	4GVR	4H0P	4H6X	4H7Y	4H9J	4HAH	4HCS
4HDJ	4HFJ	4HH3	4HH3_C	4HJP	4HRG	4HRG_C	4HS2
4HTI	4HYE	4HZA	4I2T	4I61	4I6X	4I84	4IBN
4IC4	4ID3	4IDH	4IGI	4IHE	4IKN	4IKN_B	4IL7
4ILF	4ILH	4ILH_B	4ILY	4IMH	4INK	4INO	4IPC
4IPV	4IQH	4IQM	4IRF	4IRG	4IUJ	4IZB	4IZO
4J0U	4J0W	4J11	4J2C	4J2C_B	4J2P	4J4R	4J5O
4J5Q	4J78	4J78_B	4J7A	4J87	4J8S	4J8S_B	4JAK
4JCQ	4JG2	4JHN	4JIF	4JIF_B	4JIS	4JJO	4JM7
4JMP	4JNU	4JO6	4JO6_Y	4JP0	4JQF	4JTM	4JVU
4JVW	4JYP	4JZ5	4JZC	4JZQ	4K02	4K12	4K12_B
4K5A	4K5A_B	4K5Q	4K82	4K8L	4K8Y	4K8Y_B	4KDU
4KEF	4KG4	4KN8	4KT6	4KT6_B	4KTI	4KUE	4KV2
4L00	4L0N	4L4Q	4L4Y	4L6W	4L9E	4L9R	4LA5
4LD8	4LDM	4LE3	4LEB	4LEB_B	4LF0	4LFQ	4LJ1
4LKU	4LLD	4LLD_B	4LLO	4LLO_B	4LN2	4LN2_B	4LTT
4M1F	4M1H	4M23	4M67	4M6B	4M6B_C	4M85	4M91

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 4 (continuación).

4M91_B	4M98	4M9K	4M9P	4ME2	4MFI	4MGP	4MGS
4MH4	4MHP	4MKX	4ML1	4MLS	4MLS_B	4MNO	4MOA
4MOD	4MQ3	4MQV	4MQV_B	4MT8	4MUO	4MUP	4MYL
4MYZ	4MZ2	4MZ6	4MZ6_E	4MZD	4MZZ	4N1D	4N1L
4N30	4N3X	4N5Q	4N6L	4N6T	4N77	4N7F	4N8K
4NAC	4NCU	4NFB	4NG0	4NI6	4NIO	4NIP	4NJ8
4NMI	4NP8	4NPF_Y	4NPN	4NUU	4NUU_C	4NZZ	4O0Q
4O5S	4O6G	4O7Q	4OD6	4OGD	4OHJ	4OIX	4ONK
4OQ1	4OQZ	4ORD	4OS3	4OSN	4OUH	4OUQ	4OWI
4OX6	4OX8	4OZW	4OZX	4P09	4P2I	4P2K	4P47
4P5U	4P61	4P6B	4P7T	4P93	4P9I	4PA1	4PAS
4PAS_B	4PBO	4PD0	4PEK	4PGR	4PK9	4PQD	4PRS
4PSD	4PSF	4PSJ	4PU7_B	4PUH	4PUI	4PWE	4PWQ_B
4PYU	4PYU_C	4PZ9	4Q2Q	4Q2S	4Q4N	4Q9B	4Q9P
4QB4	4QCJ	4QFT	4QMI	4QS8	4QXX_Z	4QYX	4R0P
4R0U	4R3Q	4R6F	4R8R	4R9P	4RDJ	4RF6	4RJV
4RL1	4RMB	4RTH	4RVC	4RWH	4RWU	4RWZ	4RXV
4TJX	4TJX_B	4TQ1	4TQ1_B	4TRK	4TTL	4TVE	4TX1
4U0C_B	4U1E_B	4U1E_G	4U1E_I	4U3V	4U5H	4U7E	4U7E_B
4U9C	4UED	4UED_B	4ULW	4UMG	4UMI	4UML	4UOS
4UQY	4UQY_B	4UUC	4UWW	4UYI	4W4R	4W5X_B	4WDC
4WE2	4WFI	4WIL	4WJS	4WUM	4WY4	4WY4_B	4WY4_C
4WY4_D	4X0J	4XEH	4Y65	4YUD			

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 5. Valores de energía asociados a los parámetros de cada conectividad establecidos en el potencial H-ERX*. (a) la distancia hidrógeno-oxígeno, (b) el ángulo al átomo de hidrógeno, (c) el ángulo al átomo de oxígeno (d) el ángulo diedro que corresponde a la rotación del enlace carbono-oxígeno, (e) el ángulo diedro que corresponde a la rotación del puente de hidrógeno y (f) el ángulo diedro que corresponde a la rotación del enlace hidrógeno-nitrógeno.

Intervalo de distancia (Å)			Energía (u.a.) H-ERX* (a)					Otras
			$i \rightarrow i$	$i \rightarrow i + 2$	$i \rightarrow i + 3$	$i \rightarrow i + 4$	$i \rightarrow i + 5$	
0.00	-	1.00	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.00	-	1.05	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.05	-	1.10	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.10	-	1.15	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.15	-	1.20	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.20	-	1.25	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.25	-	1.30	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.30	-	1.35	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.35	-	1.40	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.40	-	1.45	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.45	-	1.50	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.50	-	1.55	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.55	-	1.60	4.0000	4.0000	4.0000	4.0000	4.0000	2.7902
1.60	-	1.65	4.0000	4.0000	4.0000	3.1416	4.0000	1.5643
1.65	-	1.70	4.0000	4.0000	3.4470	1.8665	4.0000	0.3240
1.70	-	1.75	4.0000	4.0000	2.6048	0.8130	3.1416	-0.7969
1.75	-	1.80	4.0000	3.1957	1.6916	-0.4405	2.3018	-1.7117
1.80	-	1.85	4.0000	2.6203	0.8259	-1.4415	1.5481	-2.3523
1.85	-	1.90	4.0000	2.0343	0.0084	-2.2030	0.9651	-2.6447
1.90	-	1.95	4.0000	1.8163	-0.7258	-2.6570	0.6567	-2.6583
1.95	-	2.00	4.0000	1.5427	-1.1681	-2.8076	0.5507	-2.5127
2.00	-	2.05	4.0000	1.3196	-1.4249	-2.7452	0.6159	-2.2219
2.05	-	2.10	2.2794	1.2537	-1.4888	-2.5503	0.5706	-1.8645
2.10	-	2.15	1.4960	1.1771	-1.5189	-2.2575	0.8738	-1.4651
2.15	-	2.20	0.4477	1.1197	-1.4554	-1.9200	1.0556	-1.0500
2.20	-	2.25	-0.1831	1.1624	-1.3426	-1.5795	1.4085	-0.7208
2.25	-	2.30	-0.8653	1.2340	-1.2401	-1.2734	1.5062	-0.4316
2.30	-	2.35	-1.3643	1.2070	-1.1543	-0.9671	1.6259	-0.1878
2.35	-	2.40	-1.6500	1.1128	-1.1710	-0.7113	1.9429	0.0015
2.40	-	2.45	-1.8606	1.2458	-1.2208	-0.5063	2.0256	0.1591
2.45	-	2.50	-2.0368	1.0888	-1.2883	-0.3461	2.3248	0.3462
2.50	-	2.55	-2.1586	0.9473	-1.5073	-0.2428	2.1942	0.4162
2.55	-	2.60	-2.2367	0.6415	-1.7456	-0.1266	2.3603	0.4951

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 5 (continuación).

Intervalo de ángulo (Å)			Energía (u.a.) H-ERX* (b)					Otras
			$i \rightarrow i$	$i \rightarrow i + 2$	$i \rightarrow i + 3$	$i \rightarrow i + 4$	$i \rightarrow i + 5$	
0	-	5	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
5	-	10	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
10	-	15	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
15	-	20	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
20	-	25	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
25	-	30	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
30	-	35	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
35	-	40	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
40	-	45	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
45	-	50	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
50	-	55	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
55	-	60	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
60	-	65	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
65	-	70	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
70	-	75	5.0675	4.0000	4.0000	4.0000	4.0000	4.0000
75	-	80	2.5219	4.0000	4.0000	4.0000	4.0000	4.0000
80	-	85	0.2862	4.0000	4.0000	4.0000	4.0000	4.0000
85	-	90	-1.5109	4.0000	4.0000	4.0000	4.0000	4.0000
90	-	95	-2.3733	3.2151	4.0000	5.0675	4.0000	4.0000
95	-	100	-2.4547	2.4374	4.0000	3.2551	4.0000	4.0000
100	-	105	-1.3278	2.2794	3.0865	2.2871	4.0000	3.7796
105	-	110	1.7951	1.8946	0.5566	1.4103	4.0000	3.0865
110	-	115	3.5360	1.6295	-0.9517	0.9464	3.7129	2.2124
115	-	120	4.0000	1.3268	-1.2220	0.7351	3.7129	1.7763
120	-	125	4.0000	1.1782	-0.9389	0.4159	3.0198	1.3600
125	-	130	4.0000	1.0006	-0.7098	-0.0286	2.5722	0.8628
130	-	135	4.0000	0.9125	-0.5863	-0.4032	2.3186	0.2462
135	-	140	4.0000	0.7839	-0.6707	-0.8120	1.9540	-0.2640
140	-	145	4.0000	0.7220	-0.8780	-1.2828	1.4770	-0.8311
145	-	150	4.0000	1.3121	-1.1820	-1.8678	1.0354	-1.3922
150	-	155	4.0000	2.8984	-1.5727	-2.4686	0.5802	-1.9588
155	-	160	4.0000	4.0000	-1.7838	-2.8411	0.3935	-2.4044
160	-	165	4.0000	4.0000	-1.5588	-2.7878	0.5831	-2.6385
165	-	170	4.0000	4.0000	-0.5624	-2.2751	0.9752	-2.4773
170	-	175	4.0000	4.0000	0.8969	-1.2616	1.6254	-2.0025
175	-	180	4.0000	4.0000	2.8565	0.3579	2.9573	-0.8815

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 5 (continuación).

Intervalo de ángulo (Å)			Energía (u.a.) H-ERX* (c)					Otras
			$i \rightarrow i$	$i \rightarrow i + 2$	$i \rightarrow i + 3$	$i \rightarrow i + 4$	$i \rightarrow i + 5$	
0	-	5	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
5	-	10	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
10	-	15	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
15	-	20	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
20	-	25	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
25	-	30	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
30	-	35	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
35	-	40	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
40	-	45	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
45	-	50	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
50	-	55	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
55	-	60	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
60	-	65	4.0000	2.2949	4.0000	4.0000	4.0000	4.0000
65	-	70	4.0000	1.4840	4.0000	4.0000	4.0000	4.0000
70	-	75	1.4909	1.6746	4.0000	4.0000	4.0000	4.0000
75	-	80	-1.0932	2.2416	4.0000	4.0000	4.0000	4.0000
80	-	85	-2.9389	2.1565	2.3429	4.0000	4.0000	4.0000
85	-	90	-2.2519	2.0655	-0.1101	4.0000	4.0000	4.0000
90	-	95	2.4374	1.6457	-1.4699	4.0000	4.0000	4.0000
95	-	100	4.0000	0.9628	-1.6129	4.0000	4.0000	3.1039
100	-	105	4.0000	0.5566	-1.5139	3.0038	4.0000	2.6583
105	-	110	4.0000	0.5175	-1.5910	2.4284	4.0000	1.8486
110	-	115	4.0000	0.8226	-1.6234	1.8739	4.0000	1.2361
115	-	120	4.0000	1.9540	-1.5278	1.4071	4.0000	0.6131
120	-	125	4.0000	4.0000	-1.2047	0.9047	3.6206	-0.0054
125	-	130	4.0000	4.0000	-0.6208	0.2543	2.5518	-0.5457
130	-	135	4.0000	4.0000	-0.1157	-0.8207	1.6538	-1.0987
135	-	140	4.0000	4.0000	0.0925	-1.8568	1.0244	-1.5771
140	-	145	4.0000	4.0000	0.2737	-2.6801	0.7044	-1.9889
145	-	150	4.0000	4.0000	0.7787	-2.9973	0.5242	-2.2870
150	-	155	4.0000	4.0000	1.8637	-2.8501	0.5416	-2.3776
155	-	160	4.0000	4.0000	3.2151	-2.2319	0.7156	-2.2462
160	-	165	4.0000	4.0000	4.0000	-1.1249	1.1186	-1.9605
165	-	170	4.0000	4.0000	4.0000	0.0951	1.8288	-1.4681
170	-	175	4.0000	4.0000	4.0000	1.4874	2.3847	-0.7721
175	-	180	4.0000	4.0000	4.0000	2.8842	3.7796	0.4661

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 5 (continuación).

Intervalo de ángulo diedro (Å)			Energía (u.a.) H-ERX* (d)					Otras
			i → i	i → i + 2	i → i + 3	i → i + 4	i → i + 5	
-180	-	-170	4.0000	4.0000	3.4092	-0.0720	3.5093	0.1205
-170	-	-160	4.0000	2.8702	2.8565	-0.4001	3.5916	0.0176
-160	-	-150	4.0000	1.4136	2.0843	-0.8880	4.0000	-0.1809
-150	-	-140	4.0000	0.4207	1.1356	-1.6782	4.0000	-0.3857
-140	-	-130	4.0000	0.0977	0.3557	-2.5275	4.0000	-0.5615
-130	-	-120	4.0000	0.6684	-0.4450	-3.0095	4.0000	-0.5985
-120	-	-110	4.0000	0.8516	-1.3394	-2.8102	4.0000	-0.5221
-110	-	-100	4.0000	4.0000	-2.2115	-2.0154	3.7129	-0.3836
-100	-	-90	4.0000	4.0000	-2.6026	-1.0534	3.6206	-0.2586
-90	-	-80	4.0000	4.0000	-1.9293	-0.3061	4.0000	-0.2371
-80	-	-70	4.0000	4.0000	1.7045	0.0822	3.6504	-0.3216
-70	-	-60	4.0000	4.0000	4.0000	0.3842	4.0000	-0.5123
-60	-	-50	4.0000	4.0000	4.0000	0.7600	4.0000	-0.7887
-50	-	-40	4.0000	4.0000	4.0000	1.4168	3.3402	-0.9922
-40	-	-30	2.4835	4.0000	4.0000	2.4556	3.4333	-1.0906
-30	-	-20	0.7822	4.0000	4.0000	3.2349	2.9725	-1.1074
-20	-	-10	-0.3049	4.0000	4.0000	3.3857	2.5722	-1.1110
-10	-	0	-1.4282	4.0000	4.0000	4.0000	1.9821	-1.0861
0	-	10	-2.1493	4.0000	4.0000	3.5634	1.7353	-1.0275
10	-	20	-2.1894	4.0000	4.0000	4.0000	1.7308	-1.0648
20	-	30	-1.8372	4.0000	4.0000	4.0000	1.7533	-1.0922
30	-	40	-0.5277	4.0000	4.0000	4.0000	1.7998	-1.0802
40	-	50	1.8486	4.0000	4.0000	4.0000	1.8894	-1.0082
50	-	60	4.0000	4.0000	4.0000	4.0000	1.8337	-0.7706
60	-	70	4.0000	4.0000	4.0000	4.0000	1.9936	-0.5198
70	-	80	4.0000	4.0000	4.0000	4.0000	1.8142	-0.2956
80	-	90	4.0000	4.0000	3.3627	3.1039	1.8688	-0.1616
90	-	100	4.0000	4.0000	0.8030	2.7775	1.7624	-0.1400
100	-	110	4.0000	4.0000	0.5005	3.1766	1.6295	-0.1276
110	-	120	4.0000	1.9105	0.8969	3.2151	1.3912	-0.1352
120	-	130	4.0000	2.3762	1.5522	2.7902	1.3880	-0.0844
130	-	140	4.0000	2.9274	2.0780	2.6251	1.6704	0.0192
140	-	150	4.0000	3.1579	2.4464	2.1771	2.1100	0.1438
150	-	160	4.0000	4.0000	3.0865	1.4431	2.6361	0.1724
160	-	170	4.0000	4.0000	3.0526	0.8778	2.7161	0.2342
170	-	180	4.0000	4.0000	3.2151	0.3523	2.9573	0.1837

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 5 (continuación).

Intervalo de ángulo diedro (Å)			Energía (u.a.) H-ERX* (e)					Otras
			i → i	i → i + 2	i → i + 3	i → i + 4	i → i + 5	
-180	-	-170	4.0000	4.0000	-1.2949	-1.2472	2.7775	-1.4263
-170	-	-160	4.0000	4.0000	-2.1155	-0.5217	2.8294	-1.3792
-160	-	-150	4.0000	4.0000	-2.0210	0.1923	2.8565	-1.2660
-150	-	-140	4.0000	4.0000	-0.6097	0.8388	2.8294	-1.1551
-140	-	-130	4.0000	1.6916	0.4838	1.4299	2.9880	-1.0024
-130	-	-120	4.0000	2.0780	1.2861	1.8486	2.8842	-0.8601
-120	-	-110	4.0000	2.7525	1.6375	2.3511	2.8162	-0.7377
-110	-	-100	4.0000	2.8031	2.0052	2.3594	2.6361	-0.5907
-100	-	-90	4.0000	2.9422	1.9821	2.8294	2.4835	-0.4840
-90	-	-80	4.0000	2.8162	2.1565	2.8842	2.4020	-0.3447
-80	-	-70	4.0000	3.0198	2.2717	3.3857	2.2794	-0.2628
-70	-	-60	4.0000	2.5619	2.5122	3.4580	2.2717	-0.1080
-60	-	-50	3.7457	2.5219	2.9880	3.5634	2.3847	0.0602
-50	-	-40	2.5722	2.3594	3.4092	4.0000	2.2053	0.2737
-40	-	-30	1.5748	1.8386	4.0000	4.0000	2.3347	0.4291
-30	-	-20	0.4303	1.5978	4.0000	4.0000	2.1363	0.4219
-20	-	-10	-0.8297	1.2361	4.0000	4.0000	2.1430	0.5268
-10	-	0	-1.7043	1.1259	4.0000	4.0000	2.0971	0.6146
0	-	10	-2.1484	1.0135	4.0000	4.0000	2.2343	0.5268
10	-	20	-2.0358	1.1210	4.0000	4.0000	2.1841	0.6030
20	-	30	-1.5237	1.4365	4.0000	4.0000	2.4464	0.5215
30	-	40	-0.5337	2.2416	3.2349	3.8147	2.5619	0.4099
40	-	50	0.9066	3.0038	2.4195	3.5634	2.4284	0.3182
50	-	60	3.3627	3.5634	1.8894	3.5093	2.5219	0.3010
60	-	70	4.0000	3.5634	1.2776	3.0526	2.4835	0.2472
70	-	80	4.0000	3.8888	0.7500	2.8565	2.4374	0.1169
80	-	90	4.0000	3.7796	0.4425	2.3594	2.3847	-0.0316
90	-	100	4.0000	4.0114	0.1521	1.7624	2.2343	-0.1913
100	-	110	4.0000	3.2151	-0.0247	1.1630	2.2124	-0.3569
110	-	120	4.0000	3.0360	-0.1630	0.4953	1.8288	-0.5420
120	-	130	4.0000	2.7043	-0.1900	-0.2414	1.9484	-0.7086
130	-	140	4.0000	4.0000	-0.2457	-1.1456	1.8637	-0.9044
140	-	150	4.0000	4.0000	-0.3836	-2.3139	2.0532	-1.1244
150	-	160	4.0000	4.0000	-0.5646	-3.1528	2.0170	-1.2388
160	-	170	4.0000	4.0000	-0.7578	-3.0310	2.3186	-1.3527
170	-	180	4.0000	4.0000	-0.9656	-2.1446	2.5826	-1.4049

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 5 (continuación).

Intervalo de ángulo diedro (Å)			Energía (u.a.) H-ERX* (f)					Otras
			i → i	i → i + 2	i → i + 3	i → i + 4	i → i + 5	
-180	-	-170	4.0000	2.3847	4.0000	2.8031	3.4092	-0.8854
-170	-	-160	4.0000	2.4195	4.0000	3.3402	3.0198	-0.8103
-160	-	-150	4.0000	2.4741	4.0000	3.2349	2.7161	-0.8200
-150	-	-140	4.0000	2.3678	4.0000	3.1039	2.5122	-0.8068
-140	-	-130	4.0000	2.1497	4.0000	3.7129	2.4195	-0.8085
-130	-	-120	4.0000	1.6580	4.0000	3.1766	2.1982	-0.8127
-120	-	-110	4.0000	1.8842	4.0000	3.2551	2.2124	-0.7125
-110	-	-100	4.0000	4.0000	4.0000	2.9128	2.2641	-0.6187
-100	-	-90	4.0000	4.0000	2.8428	3.0526	2.0655	-0.4625
-90	-	-80	4.0000	4.0000	0.6715	2.6471	1.8142	-0.3026
-80	-	-70	4.0000	4.0000	-1.9588	2.7043	1.7763	-0.1926
-70	-	-60	4.0000	4.0000	-1.9702	2.5826	1.7998	-0.1262
-60	-	-50	1.1731	4.0000	-1.4225	2.1911	1.5901	0.0461
-50	-	-40	-1.0829	4.0000	-1.1378	1.9879	1.6096	0.2154
-40	-	-30	-2.1396	4.0000	-0.9101	1.6704	1.6416	0.2706
-30	-	-20	-2.2237	4.0000	-0.6981	1.3631	1.8586	0.4486
-20	-	-10	-1.8093	4.0000	-0.5461	1.0762	1.9266	0.5268
-10	-	0	-1.0646	4.0000	-0.4011	0.7483	2.1100	0.5677
0	-	10	0.1932	4.0000	-0.3147	0.5110	2.3594	0.5175
10	-	20	1.0354	4.0000	-0.2470	0.1999	2.1363	0.4940
20	-	30	1.4736	4.0000	-0.2252	-0.1400	2.3594	0.4661
30	-	40	1.8046	4.0000	-0.2371	-0.5291	2.5518	0.3490
40	-	50	2.3678	4.0000	-0.2055	-0.9849	2.5826	0.2342
50	-	60	3.3402	4.0000	-0.0224	-1.5297	2.9128	0.0249
60	-	70	4.0000	4.0000	0.3000	-2.0629	2.9725	-0.1115
70	-	80	4.0000	4.0000	1.3388	-2.5341	2.8984	-0.3436
80	-	90	4.0000	4.0000	2.7649	-2.7329	2.7043	-0.6028
90	-	100	4.0000	4.0000	4.0000	-2.5445	2.6926	-0.8615
100	-	110	4.0000	4.0000	4.0000	-2.0123	2.5122	-1.1079
110	-	120	4.0000	1.8436	4.0000	-1.1647	2.6361	-1.2459
120	-	130	4.0000	0.5242	4.0000	-0.3663	2.8565	-1.3278
130	-	140	4.0000	0.4851	4.0000	0.3193	3.1766	-1.2591
140	-	150	4.0000	0.8703	4.0000	0.9484	3.2757	-1.1449
150	-	160	4.0000	1.3755	4.0000	1.5121	3.5360	-1.0869
160	-	170	4.0000	1.6789	4.0000	2.0532	3.4333	-0.9911
170	-	180	4.0000	1.9821	4.0000	2.6251	3.1396	-0.9224

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 6. Valores de energía asociados a los parámetros de cada conectividad establecidos en el potencial H-ERMN*. (a) la distancia hidrógeno-oxígeno, (b) el ángulo al átomo de hidrógeno, (c) el ángulo al átomo de oxígeno (d) el ángulo diedro que corresponde a la rotación del enlace carbono-oxígeno, (e) el ángulo diedro que corresponde a la rotación del puente de hidrógeno y (f) el ángulo diedro que corresponde a la rotación del enlace hidrógeno-nitrógeno.

Intervalo de distancia (Å)			Energía (u.a.) H-ERMN* (a)					Otras
			$i \rightarrow i$	$i \rightarrow i + 2$	$i \rightarrow i + 3$	$i \rightarrow i + 4$	$i \rightarrow i + 5$	
0.00	-	1.00	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.00	-	1.05	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.05	-	1.10	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.10	-	1.15	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.15	-	1.20	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.20	-	1.25	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.25	-	1.30	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.30	-	1.35	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
1.35	-	1.40	4.0000	4.0000	4.0000	3.0757	4.0000	4.0000
1.40	-	1.45	4.0000	4.0000	4.0000	2.3969	4.0000	3.2988
1.45	-	1.50	4.0000	3.1345	2.8468	1.5373	4.0000	2.0461
1.50	-	1.55	4.0000	2.7392	2.2366	0.7421	4.0000	1.5684
1.55	-	1.60	4.0000	2.4263	1.7038	0.3927	4.0000	1.0438
1.60	-	1.65	4.0000	1.9125	1.3787	-0.1012	3.1345	0.3085
1.65	-	1.70	4.0000	1.3787	0.9407	-0.4532	2.7809	-0.2275
1.70	-	1.75	4.0000	0.9820	0.7933	-0.8306	2.3270	-0.7137
1.75	-	1.80	4.0000	0.5129	0.2241	-1.5660	1.8107	-1.4652
1.80	-	1.85	4.0000	0.3734	-0.2697	-1.8532	1.3277	-1.6670
1.85	-	1.90	4.0000	0.1743	-0.4698	-1.9583	1.2329	-1.7241
1.90	-	1.95	4.0000	0.1312	-0.7092	-2.0191	1.2465	-1.6962
1.95	-	2.00	3.0201	0.0190	-0.8854	-2.1146	1.2792	-1.6456
2.00	-	2.05	2.2243	0.0686	-1.1000	-2.2139	1.1588	-1.7494
2.05	-	2.10	1.6543	-0.1142	-1.2206	-2.0773	1.2604	-1.5532
2.10	-	2.15	0.9042	-0.0321	-1.2599	-2.0200	1.3377	-1.4081
2.15	-	2.20	0.2907	-0.0988	-1.2681	-1.8625	1.4161	-1.2741
2.20	-	2.25	-0.3303	-0.0940	-1.3987	-1.9015	1.3787	-1.2766
2.25	-	2.30	-0.7561	-0.2390	-1.4308	-1.6678	1.5434	-1.1044
2.30	-	2.35	-1.3541	-0.2677	-1.4970	-1.7788	1.4215	-1.2372
2.35	-	2.40	-1.7175	-0.3811	-1.5479	-1.4470	1.6071	-0.9165
2.40	-	2.45	-1.9136	-0.4624	-1.6020	-1.3596	1.5190	-0.7878
2.45	-	2.50	-2.0387	-0.6688	-1.6542	-1.1683	1.7257	-0.6175
2.50	-	2.55	-2.1100	-0.7696	-1.7647	-1.2543	1.6406	-0.7233
2.55	-	2.60	-2.1653	-0.8789	-1.8576	-0.9980	1.5941	-0.4490

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 6 (continuación).

Intervalo de ángulo (Å)			Energía (u.a.) H-ERMN* (b)					Otras
			$i \rightarrow i$	$i \rightarrow i + 2$	$i \rightarrow i + 3$	$i \rightarrow i + 4$	$i \rightarrow i + 5$	
0	-	5	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
5	-	10	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
10	-	15	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
15	-	20	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
20	-	25	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
25	-	30	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
30	-	35	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
35	-	40	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
40	-	45	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
45	-	50	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
50	-	55	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
55	-	60	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
60	-	65	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
65	-	70	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
70	-	75	4.0000	2.8716	4.0000	4.0000	4.0000	4.0000
75	-	80	3.0285	1.0879	4.0000	4.0000	4.0000	4.0000
80	-	85	0.6961	0.4434	4.0000	3.2146	4.0000	3.6315
85	-	90	-1.2514	0.5626	4.0000	2.8244	4.0000	3.4434
90	-	95	-2.0856	0.5404	1.9881	2.1546	4.0000	2.9383
95	-	100	-2.4938	0.3934	0.9029	1.5916	4.0000	2.0359
100	-	105	-1.8160	0.3166	0.0271	1.1757	3.4434	1.5349
105	-	110	-0.2224	0.0094	-0.7969	0.8333	2.9383	1.0907
110	-	115	1.8340	-0.0250	-1.3510	0.2943	2.5101	0.7315
115	-	120	3.4152	-0.1119	-1.6318	-0.0170	2.5328	0.2606
120	-	125	4.0000	-0.1882	-1.5898	-0.4284	2.0220	-0.0938
125	-	130	4.0000	-0.1836	-1.4104	-0.8185	1.8748	-0.4513
130	-	135	4.0000	-0.3434	-1.2646	-1.2057	1.7054	-0.8225
135	-	140	4.0000	-0.4635	-1.1471	-1.5240	1.5392	-1.0520
140	-	145	4.0000	-0.4583	-1.1351	-1.8535	1.3050	-1.3479
145	-	150	4.0000	-0.1866	-1.1174	-2.1061	1.1581	-1.5950
150	-	155	4.0000	0.7663	-1.1334	-2.3327	1.0665	-1.7914
155	-	160	4.0000	2.1625	-0.9549	-2.3946	1.1669	-1.8932
160	-	165	4.0000	3.4434	-0.5274	-2.3202	1.1182	-1.8884
165	-	170	4.0000	4.0000	0.0196	-2.0257	1.4115	-1.6674
170	-	175	4.0000	4.0000	0.7353	-1.4966	1.8226	-1.2326
175	-	180	4.0000	4.0000	1.9050	-0.3658	3.0669	-0.1527

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

Tabla 6 (continuación).

Intervalo de ángulo (Å)			Energía (u.a.) H-ERMN* (c)					Otras
			$i \rightarrow i$	$i \rightarrow i + 2$	$i \rightarrow i + 3$	$i \rightarrow i + 4$	$i \rightarrow i + 5$	
0	-	5	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
5	-	10	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
10	-	15	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
15	-	20	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
20	-	25	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
25	-	30	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
30	-	35	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
35	-	40	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
40	-	45	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
45	-	50	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
50	-	55	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
55	-	60	4.0000	2.8091	4.0000	4.0000	4.0000	4.0000
60	-	65	4.0000	-0.0737	4.0000	4.0000	4.0000	4.0000
65	-	70	3.1486	-0.6911	4.0000	4.0000	4.0000	4.0000
70	-	75	0.6374	-0.2833	3.5975	4.0000	4.0000	3.7406
75	-	80	-1.3253	0.1853	1.8571	4.0000	4.0000	3.4724
80	-	85	-2.9657	0.3485	0.0403	4.0000	4.0000	2.8879
85	-	90	-2.0791	0.3331	-1.1791	3.1921	4.0000	2.3448
90	-	95	3.0867	0.3166	-1.8206	2.6680	4.0000	1.7105
95	-	100	4.0000	0.1081	-1.9724	2.2196	4.0000	1.2462
100	-	105	4.0000	-0.0617	-1.8818	1.5607	4.0000	0.8375
105	-	110	4.0000	-0.2730	-1.6460	1.0098	3.2375	0.3628
110	-	115	4.0000	-0.6180	-1.3653	0.5065	2.8091	-0.0488
115	-	120	4.0000	-0.7345	-1.1248	-0.0660	2.3936	-0.3964
120	-	125	4.0000	0.3079	-0.7981	-0.6993	2.0151	-0.7141
125	-	130	4.0000	1.9050	-0.4738	-1.1788	1.6422	-1.0343
130	-	135	4.0000	4.0000	-0.1297	-1.7072	1.3428	-1.2853
135	-	140	4.0000	4.0000	0.2570	-2.1052	1.0879	-1.4281
140	-	145	4.0000	4.0000	0.6220	-2.4019	1.0934	-1.6152
145	-	150	4.0000	4.0000	1.0249	-2.5088	0.9681	-1.6688
150	-	155	4.0000	4.0000	1.5477	-2.4215	1.1043	-1.6840
155	-	160	4.0000	4.0000	2.0643	-2.1037	1.1816	-1.6226
160	-	165	4.0000	4.0000	2.7361	-1.6407	1.4266	-1.4398
165	-	170	4.0000	4.0000	3.7799	-1.0589	1.9815	-1.1552
170	-	175	4.0000	4.0000	4.2262	-0.3746	2.4555	-0.6777
175	-	180	4.0000	4.0000	4.6009	0.7881	3.6665	0.4434

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 6 (continuación).

Intervalo de ángulo diedro (Å)			Energía (u.a.) H-ERMN* (d)					Otras
			i → i	i → i + 2	i → i + 3	i → i + 4	i → i + 5	
-180	-	-170	4.0000	0.6943	2.9914	-0.4828	2.9557	0.7585
-170	-	-160	4.0000	0.2086	2.7646	-0.9532	3.0867	0.5310
-160	-	-150	4.0000	-0.2289	2.0571	-1.3602	2.8556	0.4167
-150	-	-140	4.0000	-0.4152	1.4190	-1.7882	3.0285	0.1722
-140	-	-130	4.0000	-0.4343	0.8397	-2.2195	3.1069	-0.0188
-130	-	-120	4.0000	-0.7092	0.0261	-2.4303	3.2375	-0.1306
-120	-	-110	4.0000	-1.4880	-0.8779	-2.4415	3.3352	-0.2632
-110	-	-100	4.0000	1.5016	-1.8317	-2.1218	3.2375	-0.3038
-100	-	-90	4.0000	4.0000	-2.5960	-1.6070	3.2375	-0.3834
-90	-	-80	4.0000	4.0000	-2.6423	-1.0589	3.7029	-0.4128
-80	-	-70	4.0000	4.0000	-1.2168	-0.5582	3.4724	-0.4951
-70	-	-60	4.0000	4.0000	0.5279	-0.1425	3.6315	-0.5122
-60	-	-50	4.0000	4.0000	3.5648	0.1391	3.4434	-0.5258
-50	-	-40	3.7029	4.0000	4.0000	0.4349	3.1921	-0.5144
-40	-	-30	1.5694	4.0000	4.0000	0.7146	3.0098	-0.4490
-30	-	-20	0.3343	4.0000	4.0000	0.9924	2.5561	-0.3865
-20	-	-10	-0.7928	4.0000	4.0000	1.3711	2.4989	-0.3421
-10	-	0	-1.5563	4.0000	4.0000	1.7784	2.2113	-0.3317
0	-	10	-2.1004	4.0000	4.0000	1.6281	2.2030	-0.4349
10	-	20	-2.0790	4.0000	4.0000	2.1704	2.3166	-0.4037
20	-	30	-1.6961	4.0000	4.0000	2.2625	2.1235	-0.5095
30	-	40	-0.6947	4.0000	4.0000	2.2452	2.1704	-0.6307
40	-	50	1.3220	4.0000	4.0000	1.8868	2.1546	-0.6902
50	-	60	4.0000	4.0000	4.0000	2.0643	2.0571	-0.7940
60	-	70	4.0000	4.0000	4.0000	1.9362	2.1389	-0.8229
70	-	80	4.0000	4.0000	2.7792	1.8513	2.0935	-0.6961
80	-	90	4.0000	4.0000	1.6006	1.8748	1.9362	-0.6562
90	-	100	4.0000	4.0000	1.0665	1.6660	1.9425	-0.5535
100	-	110	4.0000	3.1486	0.9467	1.4575	1.8989	-0.3362
110	-	120	4.0000	0.0233	1.4894	1.2654	2.0083	-0.1480
120	-	130	4.0000	0.6726	2.2030	1.1639	2.0289	0.0910
130	-	140	4.0000	0.9420	2.5920	0.9467	2.1235	0.3104
140	-	150	4.0000	1.1997	2.7792	0.6780	2.1467	0.5886
150	-	160	4.0000	1.3084	2.9914	0.3280	2.4770	0.6426
160	-	170	4.0000	1.0586	3.4434	0.0432	2.3936	0.7922
170	-	180	4.0000	0.9326	3.4152	-0.2743	2.5679	0.7702

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 6 (continuación).

Intervalo de ángulo diedro (Å)			Energía (u.a.) H-ERMN* (e)					Otras
			i → i	i → i + 2	i → i + 3	i → i + 4	i → i + 5	
-180	-	-170	4.0000	4.0000	-1.6128	-1.7759	2.2892	-1.0981
-170	-	-160	4.0000	4.0000	-2.4040	-1.3845	2.3544	-1.1117
-160	-	-150	4.0000	4.0000	-2.2932	-0.9137	2.1785	-0.9271
-150	-	-140	4.0000	0.0634	-1.2682	-0.5001	2.1235	-0.6856
-140	-	-130	4.0000	-0.2398	-0.1683	-0.0600	2.1159	-0.5226
-130	-	-120	4.0000	0.2547	0.6068	0.3017	2.1467	-0.4260
-120	-	-110	4.0000	0.5357	1.1846	0.6655	1.9881	-0.2547
-110	-	-100	4.0000	0.7053	1.8002	0.9633	2.1625	-0.0754
-100	-	-90	4.0000	0.7526	2.1084	1.2057	2.5444	-0.0312
-90	-	-80	4.0000	0.7861	2.1947	1.3533	2.6549	0.0990
-80	-	-70	4.0000	0.9777	2.4770	1.5307	2.8399	0.1339
-70	-	-60	4.0000	0.8917	2.9734	1.6143	2.7503	0.3104
-60	-	-50	2.6167	0.8829	3.1069	1.9236	2.8556	0.3420
-50	-	-40	1.5434	0.7392	3.1486	1.8571	2.8879	0.4520
-40	-	-30	0.5172	0.6426	3.1069	2.0571	3.0669	0.6002
-30	-	-20	-0.3051	0.6514	3.1921	2.0289	3.2146	0.6816
-20	-	-10	-0.9864	0.4870	3.2611	2.1947	2.9383	0.6514
-10	-	0	-1.5212	0.4057	3.3099	2.1009	3.0475	0.6409
0	-	10	-1.9714	0.1986	3.1486	2.0500	3.2146	0.6708
10	-	20	-1.9167	0.1454	2.9557	2.0500	3.2852	0.6708
20	-	30	-1.5788	0.1657	2.9557	1.9618	3.1486	0.6426
30	-	40	-0.9472	0.1657	2.4138	1.7206	3.1275	0.6136
40	-	50	0.2943	0.4826	2.0429	1.6097	2.7503	0.5562
50	-	60	1.9881	0.7546	1.8868	1.6143	2.8556	0.4885
60	-	70	4.0000	0.9561	1.4497	1.2850	2.6167	0.4335
70	-	80	4.0000	1.2883	1.1552	1.0665	2.5328	0.2430
80	-	90	4.0000	1.5520	0.7881	0.8895	2.4661	0.2030
90	-	100	4.0000	1.7309	0.5968	0.5357	2.4879	0.0683
100	-	110	4.0000	1.9683	0.4796	0.2211	2.4989	-0.0788
110	-	120	4.0000	1.8688	0.3280	-0.2296	2.2625	-0.2820
120	-	130	4.0000	1.7004	0.3615	-0.7467	2.0429	-0.5193
130	-	140	4.0000	1.2590	0.2349	-1.4037	2.1312	-0.6647
140	-	150	4.0000	1.3463	0.1329	-2.0249	2.0500	-0.9253
150	-	160	4.0000	4.0000	-0.1714	-2.5326	2.1235	-1.0938
160	-	170	4.0000	4.0000	-0.4918	-2.6482	2.1785	-1.1745
170	-	180	4.0000	4.0000	-0.9723	-2.3603	2.2802	-1.2356

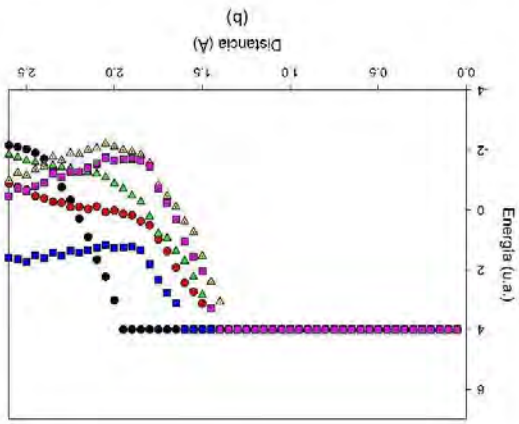
ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

Tabla 6 (continuación).

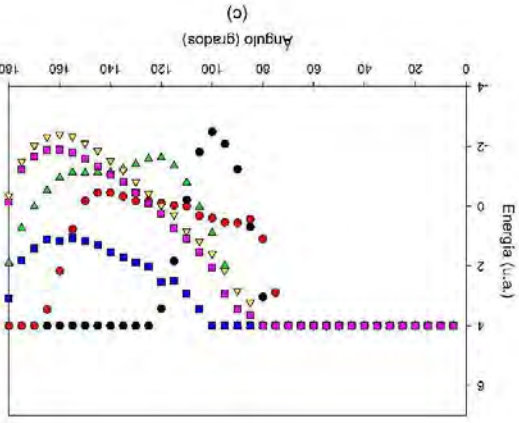
Intervalo de ángulo diedro (Å)			Energía (u.a.) H-ERMN* (f)					Otras
			i → i	i → i + 2	i → i + 3	i → i + 4	i → i + 5	
-180	-	-170	4.0000	0.8084	3.8207	0.7902	3.1921	0.0168
-170	-	-160	4.0000	0.6136	4.0000	0.7742	3.0285	-0.1322
-160	-	-150	4.0000	0.4462	4.0000	0.9120	2.8556	-0.2121
-150	-	-140	4.0000	0.2980	4.0000	0.8653	2.6043	-0.2840
-140	-	-130	4.0000	0.1809	3.6665	0.8229	2.4138	-0.4600
-130	-	-120	4.0000	-0.1904	3.4434	0.8187	2.1312	-0.6428
-120	-	-110	4.0000	-0.8269	2.7221	0.7565	2.0935	-0.7965
-110	-	-100	4.0000	0.6409	1.9489	0.7278	1.9815	-0.7384
-100	-	-90	4.0000	3.4434	0.2500	0.6690	2.0571	-0.6190
-90	-	-80	4.0000	4.0000	-1.1011	0.6237	2.1467	-0.4420
-80	-	-70	4.0000	4.0000	-2.1722	0.5467	2.0788	-0.2929
-70	-	-60	4.0000	4.0000	-2.3775	0.5126	2.2452	-0.1322
-60	-	-50	1.6954	4.0000	-1.7433	0.4420	2.5328	-0.0522
-50	-	-40	-0.8974	4.0000	-1.1491	0.3537	2.4879	0.0196
-40	-	-30	-1.9537	4.0000	-0.6628	0.3498	2.5920	0.1690
-30	-	-20	-2.1121	4.0000	-0.1904	0.2665	2.6420	0.2761
-20	-	-10	-1.8233	4.0000	0.0556	0.2337	2.9557	0.2846
-10	-	0	-1.3302	4.0000	0.3017	0.1071	2.7646	0.3104
0	-	10	-0.7441	4.0000	0.3356	-0.1193	2.5920	0.1941
10	-	20	0.0980	4.0000	0.3694	-0.2956	2.7221	0.2109
20	-	30	0.8784	4.0000	0.4195	-0.5078	2.9383	0.2761
30	-	40	1.4813	4.0000	0.4635	-0.8384	2.7083	0.1497
40	-	50	1.7947	4.0000	0.4650	-1.1872	2.6549	-0.0206
50	-	60	2.9212	4.0000	0.4112	-1.5276	2.7941	-0.1566
60	-	70	4.0000	4.0000	0.2943	-1.8122	2.3738	-0.2561
70	-	80	4.0000	4.0000	0.5594	-2.0133	2.3544	-0.4508
80	-	90	4.0000	4.0000	0.8851	-2.1003	2.1159	-0.5566
90	-	100	4.0000	3.3612	1.5434	-2.1231	2.2113	-0.6032
100	-	110	4.0000	1.4773	2.3448	-1.9340	2.1704	-0.7213
110	-	120	4.0000	-0.5850	2.9557	-1.5484	2.3166	-0.7583
120	-	130	4.0000	-0.7292	3.3352	-1.0349	2.4036	-0.6249
130	-	140	4.0000	-0.3205	3.3878	-0.5809	2.4240	-0.4589
140	-	150	4.0000	0.0076	4.0000	-0.1813	2.8716	-0.3251
150	-	160	4.0000	0.3217	3.6315	0.1820	2.9212	-0.1377
160	-	170	4.0000	0.5886	4.0000	0.4448	2.8879	-0.1551
170	-	180	4.0000	0.7072	4.0000	0.5853	3.2852	-0.0821

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"

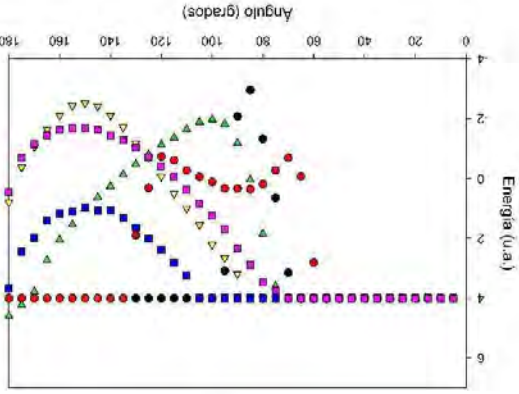
H-ERMN*



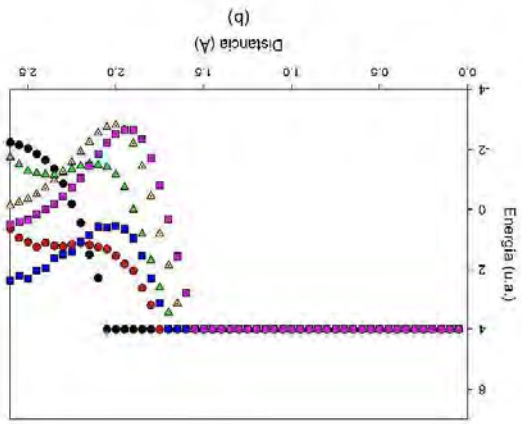
(b)



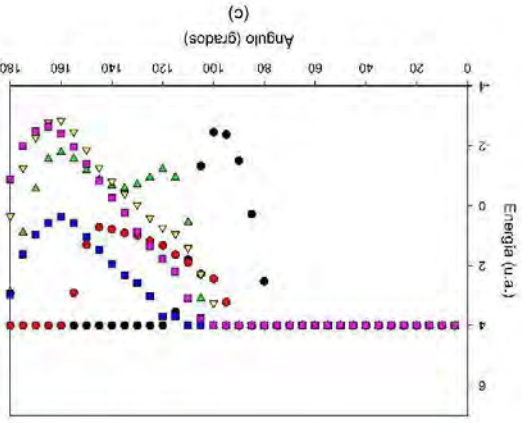
(c)



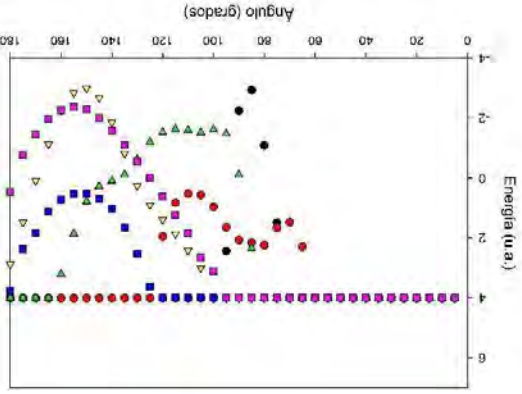
H-ERX*



(b)



(c)



ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUENTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

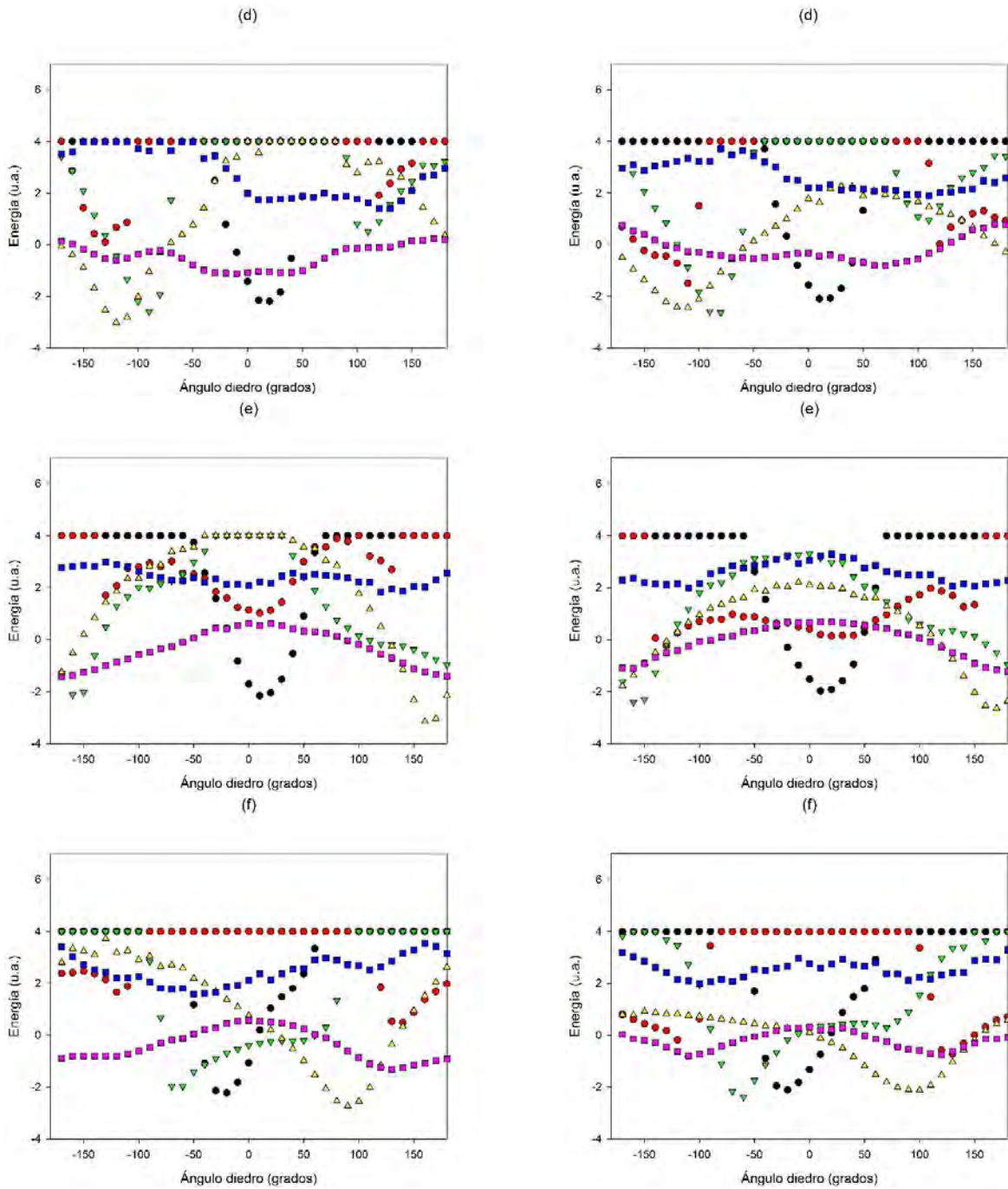


Figura 5. Componentes de energía asociados a los parámetros establecidos en los potenciales H-ERX* y H-ERMN*. (a) la distancia hidrógeno-oxígeno, (b) el ángulo al átomo de hidrógeno, (c) el ángulo al átomo de oxígeno (d) el ángulo diedro que corresponde a la rotación del enlace carbono-oxígeno, (e) el ángulo diedro que corresponde a la rotación del puente de hidrógeno y (f) el ángulo diedro que corresponde a la rotación del enlace hidrógeno-nitrógeno.

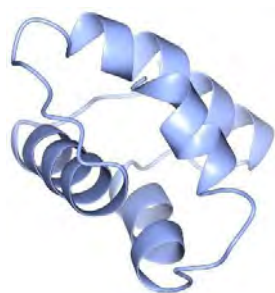
Conectividad residuo aceptor → residuo donador

- $i \rightarrow i$
- $i \rightarrow i + 2$
- $i \rightarrow i + 3$
- $i \rightarrow i + 4$
- $i \rightarrow i + 5$
- Otras

Color



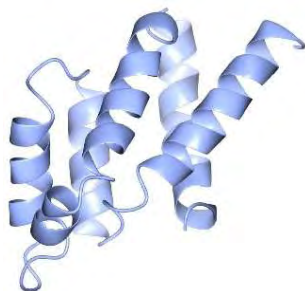
ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"



1af7_



11.021



1abv_



13.561



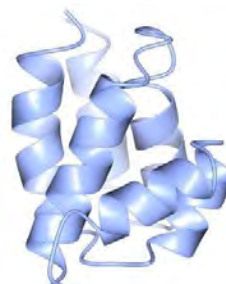
1ah9_



10.875



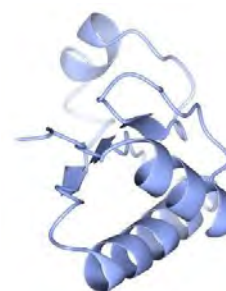
1fadA



3.705

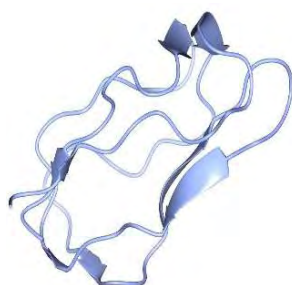


1fo5A



3.97

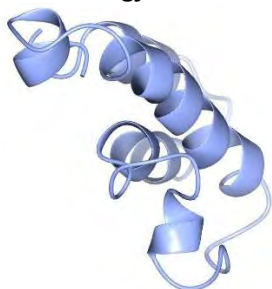
ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”



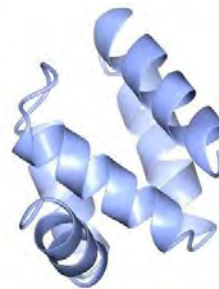
1gixA



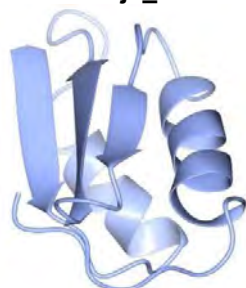
6.286



1kjs_



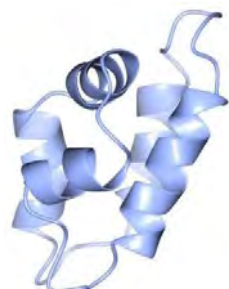
11.019



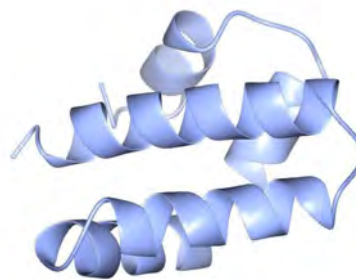
1kviA



2.243



1of9A



3.591



1sro_



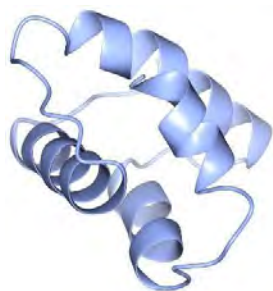
4.393

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

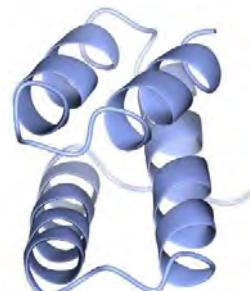


Figura 6. Modelos identificados como estructura nativa por el potencial H-ERX*. El número debajo del modelo indica su desviación cuadrática media con respecto a la estructura nativa.

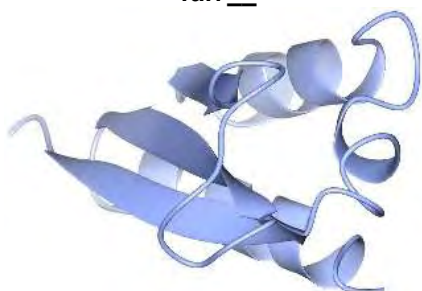
ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"



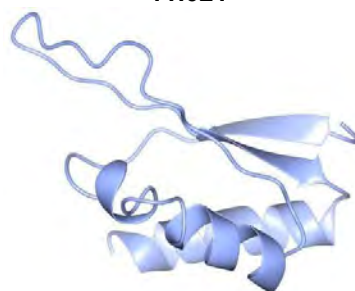
1af7_



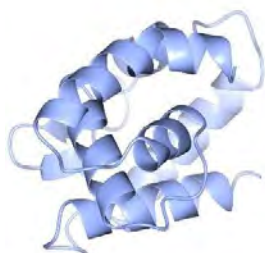
11.021



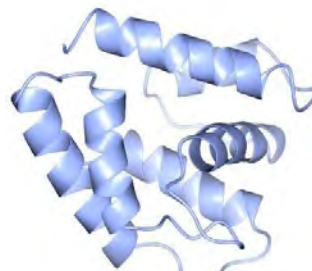
1dtjA_



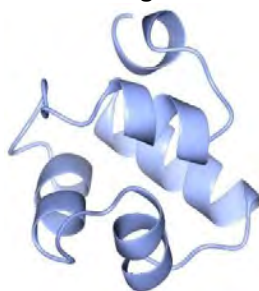
2.985



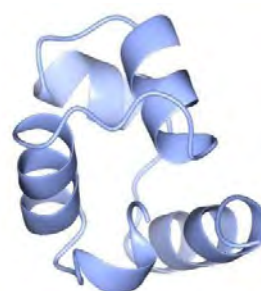
1orgA



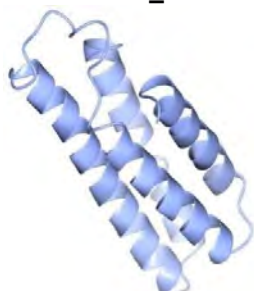
2.613



1r69_



2.027

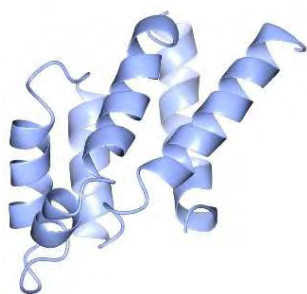


256bA

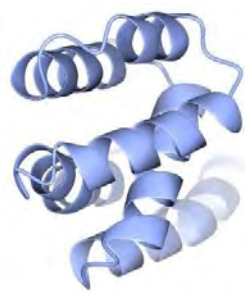


3.172

ANEXO - "DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS"



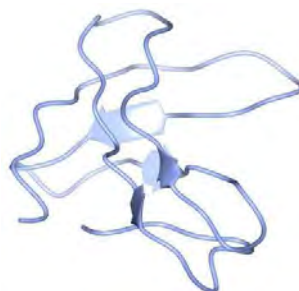
1abv_



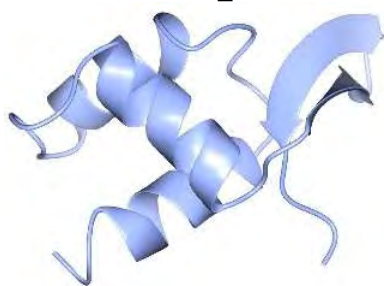
12.780



1ah9_



10.875



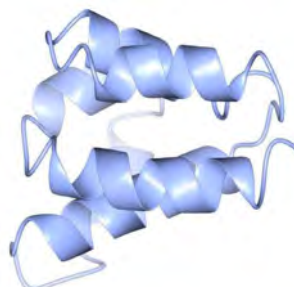
1aoy_



5.127



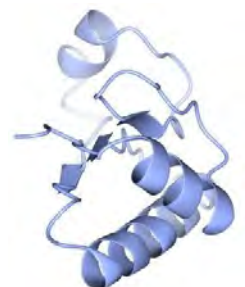
1fadA



3.705

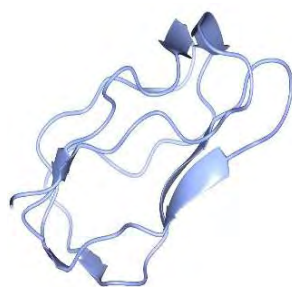


1fo5A

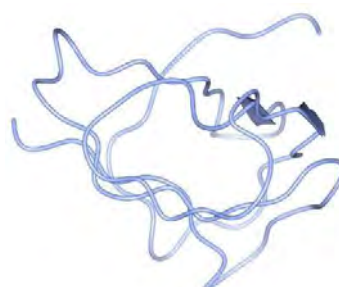


3.97

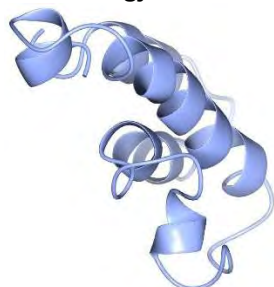
ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”



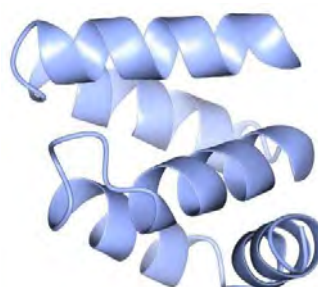
1gjaA



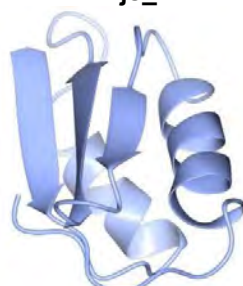
8.723



1kjs_



11.019



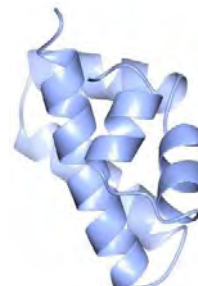
1kviA



2.243



1of9A



3.591



1sro_



4.188

ANEXO - “DESARROLLO DE UN POTENCIAL ESTADÍSTICO BASADO EN EL ANÁLISIS DE LOS PUNTES DE HIDRÓGENO PRESENTES EN PROTEÍNAS”

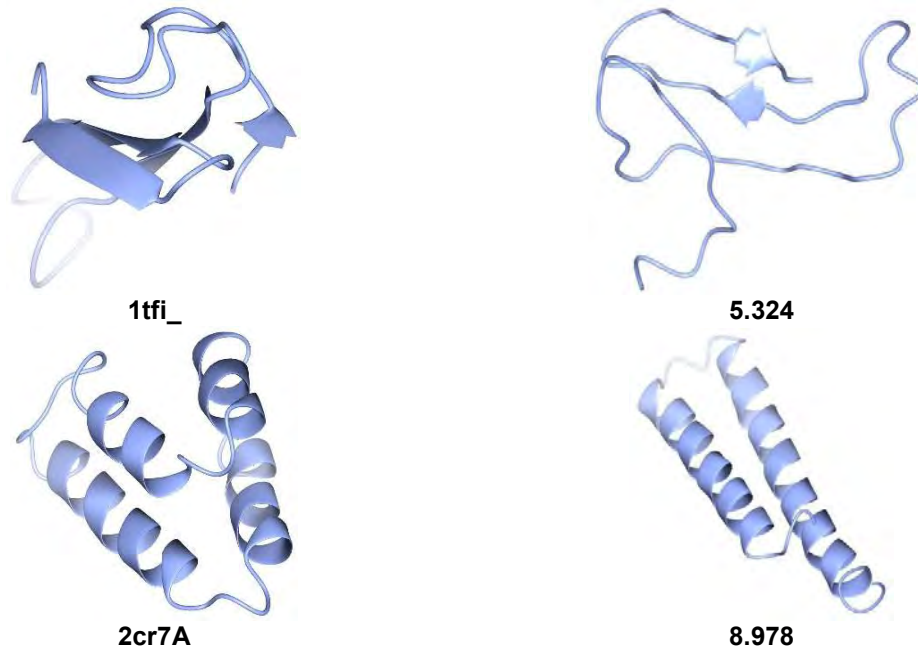


Figura 7. Modelos identificados como estructura nativa por el potencial H-ERMN*. El número debajo del modelo indica su desviación cuadrática media con respecto a la estructura nativa.