



Identificación de sitios alostéricos en proteínas del *Homo sapiens* que interactúan con la molécula de ATP mediante la herramienta Deacon Active Site Profiler (DASP3)

Adriana Camila González Rosas

Facultad de Ingeniería Mecánica, Electrónica y Biomédica
Programa de Ingeniería Biomédica
Universidad Antonio Nariño
Bogotá, Colombia
2020

Identificación de sitios alostéricos en proteínas del Homo sapiens que interactúan con la molécula de ATP mediante la herramienta Deacon Active Site Profiler (DASP3)

Adriana Camila Gonzalez Rosas

Trabajo Integral de Grado presentado como requisito para optar al título de:
Ingeniera Biomédica

Director:
Ing. Mario Enrique Duarte González PhD

Línea de investigación:
Bioingeniería

Universidad Antonio Nariño
Facultad de Ingeniería Mecánica, Electrónica y Biomédica
Programa de Ingeniería Biomédica
Bogotá, Colombia
2020

Nota de Aceptación

Firma presidente del Jurado

Firma Jurado 1

Firma Jurado 2

Bogotá D.C. _____

Todo empieza por el deseo de hacer algo bueno por el mundo, al principio el camino puede parecer difícil tal vez largo o imposible, pero “Un viaje de mil millas comienza con un primer paso” Lao-Tse.

No dejes que el miedo se apodere de tus sueños, no dejes de luchar, ni de creer en ti, recuerda: “Nuestra mayor debilidad radica en renunciar, la forma más segura de tener éxito es siempre intentarlo una vez más” Thomas A. Edison.

Agradecimientos

Primeramente agradecida con Dios, siempre con mi fé en él, tome la decisión de estudiar ingeniería biomédica como el camino para mi vida. Al principio parecía algo lejano, pero de a poco se ha llegado a la culminación de este proceso, cumpliéndose uno de mis mayores sueños, que costó tiempo, lágrimas, dedicación y esfuerzo, pero ha dejado consigo grandes aprendizajes.

Agradezco a mi familia, pues ha llenado mis días, proponiéndome ser mejor cada día, no caer ante los pequeños o grandes desafíos, pues estamos para aportar algo a nuestros hermanos. Especial agradecimiento a mi hermana Yulieth G.R, por su compañía, por cada día escucharme, y a mi madre Nancy R. por sus enseñanzas, valores, esperanza y fé en mi.

Agradezco a todos y cada uno de los docentes e ingenieros que estuvieron durante este proceso.

En especial al ingeniero Mario E. Duarte, él me ha regalado grandes aprendizajes, desde mis laboriosas clases de circuitos eléctricos hasta señales, analógicos y el asesoramiento de este trabajo, mi total agradecimiento, ha sido una persona con gran disposición, con mucho conocimiento por aportar.

A la ingeniera Alejandra V. y al ingeniero Leonardo T., personas maravillosas, con gran energía y entusiasmo por su vocación.

Al ingeniero Andrés S., por su visión, por hacer saber que siempre se puede llegar más allá. Al ingeniero Sebastián J. y el ingeniero Andrés R., por brindarme sus conocimientos, reflexiones y por su disposición durante las clases.

A los biólogos Olga E., Johana G. y Alexander R., por su tiempo, disposición y aportes al desarrollo del presente trabajo.

A mis amigos y amigas de quienes he recibido apoyo sincero y con quienes he tenido la suerte de compartir experiencias tanto en el ámbito personal como estudiantil.

Por último doy gracias por la vida, por llegar hasta este momento, permitiéndome crecer como persona, dando fin a una etapa para el comienzo de una nueva.

Resumen

En las proteínas se distinguen los sitios activos y los sitios alostéricos. Estos últimos poseen su sitio de unión dentro de la enzima en un lugar diferente al del sitio activo. Su importancia radica en la contribución que realiza en la inhibición y/o activación de su función biológica, esto, mediante moléculas (ligandos) que actúan como moduladores alostéricos, siendo una base para el diseño de fármacos, pues proporcionan menos efectos adversos que los tradicionales (reguladores activos).

Una de las principales moléculas que puede actuar como regulador alostérico es el adenosín trifosfato (ATP), ya que es fundamental para la obtención de energía celular.

Entre las bases de datos públicas que cuentan con información sobre la estructura de las proteínas, se encuentra el *National Center for Biotechnology Information* (NCBI), allí existen 1'364.990 proteínas en *Homo sapiens*, algunas de estas no han sido estudiadas y por tanto no se conoce si existe un sitio alostérico, ni su posición molecular.

En el trabajo que se desarrolla a continuación, se presenta la creación de un perfil de sitio alostérico (ASP), a partir de tres proteínas y con la ayuda de la herramienta *Deacon Active Site Profiler 3* (DASP3), que permite identificar sitios activos a partir de la creación de un perfil de sitio activo; a dichas proteínas, por medio de la literatura revisada, se les conoce sitio alostérico en interacción con el ligando ATP.

El ASP creado fue comparado con un archivo FASTA, descargado del NCBI, que contiene las secuencias de aminoácidos en *Homo sapiens*, permitiendo la localización de proteínas a las que no se les ha identificado su sitio alostérico en interacción con la molécula de ATP. De las proteínas identificadas, se eligieron tres (3), fueron verificadas virtualmente, a las proteínas que no contaban con estructura tridimensional (PDB), se les realizó una reconstrucción geométrica en I-TASSER, finalmente se les realizó docking (herramienta computacional para realizar acoplamiento proteína-ligando) mediante AutoDocktools (ADT) con el ejecutable de AutoDock4 y Autogrid, para comparar sus energías de unión y sus valores de RMSD, con los Docking de las proteínas que se seleccionaron para realización de ASP.

Gracias a lo anteriormente mencionado se identificaron tres proteínas: 1) *BCL2-Like-protein 12*, 2) *Galactine-9* y 3) *Fumarate hydratase, mitochondrial*, que posiblemente tienen sitio alostérico e interacción con el ligando ATP.

Se propuso una metodología para identificar y verificar virtualmente ligandos que pueden actuar como moduladores alostéricos en proteínas, llegando a la conclusión de que al simular las proteínas identificadas mediante la herramienta DASP3 con el software AutoDockTools, muy posiblemente se identificaron nuevos sitios alostéricos en interacción con ATP.

Palabras clave: Proteínas, Sitios alostéricos, NCBI, ATP, DASP3, Docking, ADT

Abstract

Active proteins and allosteric sites are distinguished in proteins. The latter have their binding site within the enzyme in a different place from the active site. Its importance lies in the contribution it makes in the inhibition and / or activation of its biological function, this, through molecules (ligands) that act as allosteric modulators, being a basis for the design of drugs, since they provide fewer adverse effects than traditional ones. (active regulators).

One of the main molecules that can act as an allosteric regulator is adenosine triphosphate (ATP), since it is essential for obtaining cellular energy.

Among the public databases that have information on the structure of proteins, is the *National Center for Biotechnology Information* (NCBI), there are 1,364,990 proteins in *Homo sapiens*, some of which have not been studied and therefore it is not known if there is an allosteric site, nor its molecular position.

In the work that follows, the creation of an allosteric site profile (ASP) is presented, from three proteins and with the help of the *Deacon Active Site Profiler 3* tool (DASP3), which allows identify active sites by creating an active site profile; these proteins, through the reviewed literature, are known allosteric site in interaction with the ATP ligand.

The created ASP was compared to a FASTA file, downloaded from NCBI, that contains the amino acid sequences in *Homo sapiens*, allowing the localization of proteins that have not been identified their allosteric site in interaction with the ATP molecule. Of the identified proteins, three (3) were chosen, they were virtually verified, the proteins that did not have a three-dimensional structure (PDB), a geometric reconstruction was performed in I-TASSER, finally they were docked (computational tool to perform protein-ligand coupling) using AutoDocktools (ADT) with the executable of AutoDock4 and Autogrid, to compare their binding energies and their RMSD values, with the Docking of the proteins that were selected for ASP.

Thanks to the aforementioned, three proteins were identified: 1) *BCL2-Like-protein 12*, 2) *Galactine-9* and 3) *Fumarate hydratase, mitochondrial*, which possibly have an allosteric site and interaction with the ATP ligand.

A methodology was proposed to identify and virtually verify ligands that can act as allosteric modulators in proteins, concluding that by simulating the proteins identified by the DASP3 tool with AutoDockTools software, it was very possible to identify new allosteric sites in interaction with ATP .

Keywords : Proteins, NCBI, ATP, DASP3, Docking, ADT

Tabla de contenidos

Agradecimientos	ix
Resumen	xi
Abstract	xiii
Lista de símbolos	1
Lista de Figuras	5
Lista de Tablas	5
1. Introducción	7
1.1. Estado del arte	8
1.2. Planteamiento del problema	11
1.3. Justificación	12
1.4. Objetivos	13
1.4.1. Objetivo general	13
1.4.2. Objetivos específicos	13
1.5. Alcance y limitaciones	13
1.5.1. Ubicación dentro de las líneas de trabajo del programa	13
1.5.2. Usuarios directos y formas de utilización de los resultados del proyecto	14
1.6. Estructura del documento	14
2. Marco Teórico	15
2.1. Proteína	15
2.1.1. Estructura de las proteínas	16
2.1.2. Propiedades y funciones de las proteínas	18
2.1.3. Enzimas	19
2.1.4. Sitio Alostérico	20
2.1.5. Adenosin Triphosphate (ATP)	21
2.2. Bioinformática	22
2.2.1. DASP3	23
2.2.2. Docking	24
2.2.3. I-TASSER	26

2.2.4.	Swiss-PdbViewer (también conocido como DeepView)	26
2.2.5.	UCSF Chimera	27
2.2.6.	Avogadro	27
2.3.	Desviación cuadrática media (RMSD)	28
3.	Identificación de proteínas mediante realización de perfil alostérico	31
3.1.	Revisión de proteínas para elaboración de perfil alostérico	32
3.2.	Residuos que interactúan con el ligando ATP	33
3.3.	Creación del perfil alostérico e identificación de Proteínas mediante DASP3	34
3.3.1.	Selección de proteínas identificadas mediante el DASP3	35
4.	Verificación de proteínas identificadas en NCBI	39
4.1.	Preparación de proteínas y ligandos	40
4.1.1.	Preparación proteínas	40
4.1.2.	Preparación ligando	42
4.2.	Elección rejilla para elaboración de <i>Docking</i>	43
4.3.	<i>Docking</i>	44
5.	Resultados y discusión de proteínas identificadas mediante DASP3	47
5.1.	Acoplamiento molecular de los complejos proteína-ligando	47
5.2.	Análisis de interacciones	48
5.3.	Energías de unión y valores RMSD de proteínas elegidas para creación de perfil alostérico	52
5.3.1.	Endoplasmic reticulum chaperone BiP	53
5.3.2.	Cytosolic purine 5'-nucleotidase	54
5.3.3.	ATP-depent 6-phosphofructokinase	54
5.4.	Energías de unión y RMSD para proteínas identificadas por el DASP3	55
5.4.1.	<i>Docking</i> BCL2- like-protein 12	55
5.4.2.	<i>Docking</i> Galactine-9	58
5.4.3.	<i>Docking</i> Fumarate hydratase, mitochondrial	60
5.5.	Comparación energías de unión de proteínas identificadas del NCBI con proteínas de entrenamiento en el DASP3	62
6.	Conclusiones	63
6.1.	Recomendaciones para trabajos futuros	63
A.	Lista de proteínas identificadas mediante el DASP3	65
A.0.1.	pdb con exponente -5	65
A.0.2.	NP con exponente -5	66
A.0.3.	SP con exponente -5	68
A.0.4.	XP con exponente -5	68

A.0.5. Diferentes con exponente -5	70
A.0.6. Proteínas con exponente -6	72
A.0.7. XP con exponente -7	73
A.0.8. NP con exponente -7	73
A.0.9. Diferente con exponente -7	74
A.0.10. Pdb con exponente -8	74
A.0.11. XP con exponente -8	75
A.0.12. NP con exponente -8	76
A.0.13. Diferentes con exponente -8	76
A.0.14. NP con exponente -10	76
A.0.15. XP con exponente -12	76
A.0.16. NP con exponente -12	77
A.0.17. XP con exponente -13	77
A.0.18. NP con exponente -13	77
A.0.19. Diferentes con exponente -13	77
A.0.20. NP con exponente -15	77
A.0.21. XP con exponente -16	78
A.0.22. XP con exponente -17	78
A.0.23. NP con exponente -17	78
A.0.24. Pdb con exponente -17	79
A.0.25. Diferentes con exponente -17	79
A.0.26. NP con exponente -19	79
A.0.27. Diferentes con exponente -19	79
A.0.28. Pdb con exponente -22	79
A.0.29. NP con exponente -22	80
A.0.30. XP con exponente -22	80
A.0.31. Diferentes con exponente -22	80

Bibliografía**81**

Lista de símbolos

Abreviaturas

Aminoácido	Abreviatura	Símbolo
Glicina	Gly	G
Alanina	Ala	A
Valina	Val	V
Leucina	Leu	L
Isoleucina	Ile	I
Fenilalanina	Phe	F
Tirosina	Tyr	Y
Triptófano	Trp	W
Serina	Ser	S
Treonina	Thr	T
Cisteína	Cys	C
Metionina	Met	M
Ácido Aspártico	Asp	D
Ácido Glutámico	Glu	E
Histidina	His	H
Lisina	Lys	K
Arginina	Arg	R
Asparagina	Asn	N
Glutamina	Gln	Q
Prolina	Pro	P

Lista de Figuras

2-1. Esquema general de la expresión genética	15
2-2. Niveles de estructura de una proteína	16
2-3. Unión de la enzima con el sustrato	20
2-4. Unión de la enzima con el sustrato en el sitio alostérico	21
2-5. Estructura molecular Adenosin Triphosphate (ATP)	22
2-6. Carpeta Docking con archivos necesarios para correr AutoDockTools	25
2-7. Icono Swiss-PdbViewer	27
2-8. Icono UCSF Chimera	27
2-9. Icono Avogadro	28
3-1. Diagrama bloques con procedimiento para identificar proteínas mediante DASP3, Imagen de Autor	31
3-2. Aminoácidos cytosolic purine 5'-nucleotidase con PDB 2XCW	33
3-3. Comando en linux para comparación de ASP con archivo FASTA	35
4-1. Diagrama de bloques con procedimiento simplificado para realización de <i>Doc- king</i> en AutoDockTools, imagen de Autor	39
4-2. Proteína en AutoDockTools, imagen de Autor	42
4-3. ATP 2D en Avogadro, Imagen de Autor	43
4-4. ATP minimizada en Avogadro, Imagen de Autor	43
4-5. Selección de puntos de la proteína en AutoDockTools, imagen de Autor . . .	44
4-6. Tamaño rejilla en AutoDockTools, imagen de Autor	44
4-7. Parámetros para ejecutar Autogrid en AutoDockTools, imagen de Autor . .	45
4-8. Parámetros para ejecutar Autodock4 en AutoDockTools, imagen de Autor .	46
5-1. Bolsillo proteínas entrenamiento DASP3 obtenidas mediante ADT, imagen de autor	48
5-2. Bolsillo proteínas elegidas del NCBI a partir de DASP3 obtenidas mediante ADT, imagen de autor	48
5-3. Interacciones BCL2-Like-Protein 12 obtenidas mediante AutoDockTools, ima- gen de Autor	50
5-4. Interacciones Fumarate obtenidas mediante AutoDockTools, imagen de Autor	51
5-5. Interacciones Galactine-9 obtenidas mediante AutoDockTools, imagen de Autor	52

Lista de Tablas

2-1. Estructuras de las proteínas	17
2-2. Herramientas en el campo de la bioinformática para predecir sitios alostéricos y/o activos	23
3-1. Proteínas utilizadas para realización de ASP mediante DASP3	34
4-1. Estructuras de proteínas listas para realización de <i>Docking</i>	41
5-1. Conformaciones con residuos conocidos en interacción con ATP, <i>Endoplasmic reticulum chaperone BiP</i>	53
5-2. Conformaciones con residuos conocidos en interacción con ATP, <i>Cytosolic purine 5'-nucleotidase</i>	54
5-3. Conformaciones con residuos conocidos en interacción con ATP, <i>ATP-depent 6-phosphofructokinase</i>	55
5-4. Conformaciones de <i>Docking</i> cubriendo toda la <i>BCL2-Protein 12</i> en interacción con ATP	56
5-5. Conformaciones del ATP con la proteína <i>BCL2-Protein 12</i> en el sitio 1 estimado por el DASP3	57
5-6. Conformaciones del ATP con la proteína <i>BCL2-Protein 12</i> en el sitio 2 estimado por el DASP3	57
5-7. Conformaciones <i>Docking</i> cubriendo toda la <i>Galactine-9</i> en interacción con ATP	58
5-8. Conformaciones <i>Galactine-9</i> con ATP en el sitio 1 predicho por el DASP3	59
5-9. Conformaciones <i>Galactine-9</i> con ATP en el sitio 2 predicho por el DASP3	59
5-10. Conformaciones <i>Docking</i> cubriendo toda la <i>Fumarate hydratase, mitochondrial</i> en interacción con ATP	60
5-11. Conformaciones ATP con <i>Fumarate hydratase, mitochondrial</i> en el sitio 1 predicho por el DASP3	61
5-12. Conformaciones ATP con <i>Fumarate hydratase, mitochondrial</i> en el sitio 2 predicho por el DASP3	61

1. Introducción

Las proteínas son cadenas de aminoácidos que se unen por medio de enlaces péptidos formando un grupo amida entre un grupo amino y uno carboxilo [1].

La proteína es estudiada de acuerdo a su estructura. Estructura primaria (forma de organización más básica): contiene la información necesaria para determinar el orden, número y tipo de aminoácidos que componen la proteína [2, 3]; estructura secundaria: codifica la manera en que los aminoácidos se encuentran enlazados en el espacio (alfa hélice y beta plegada) y el plegamiento entre residuos de aminoácidos de la cadena polipéptida (formación de enlaces de hidrógeno) [4]; estructura terciaria: contiene la información de la interacción de las estructuras secundarias en el espacio (unión de alfa hélice con beta plegada), se determina mediante enlaces covalentes (puentes disulfuro) y no covalentes (puentes de hidrógeno, interacciones hidrófobas, fuerza de Vander Waals, interacciones iónicas), pueden ser fibrosas (insolubles en agua) o globulares (Solubles en agua) [5]; y estructura cuaternaria: varias estructuras terciarias forman la proteína, la cual está formada por dos o más cadenas polipéptidas (se unen por enlaces de diferentes tipos, similares a los de la estructura terciaria). Cada una de las cadenas terciarias son conocidas como una subunidad o protómero y el complejo cuaternario es entonces una proteína oligomérica o multimérica; sin embargo no todas la proteínas alcanzan esta estructura [6].

Dependiendo de la estructura de la proteína se puede desarrollar una función: estructural, de transporte y almacenamiento, movimiento, protección y/o reguladora [7, 8, 9].

Otra de las funciones más importantes de las proteínas, es la catálisis de reacciones biológicas, las proteínas que tienen esta función son las que actúan como enzimas, las cuales aceleran las reacciones que se pueden producir espontáneamente [10].

Las enzimas presentan sitios activos y alostéricos. En los sitios activos la enzima se une con el sustrato, por lo que quedan orientados de tal manera que se favorece la catálisis (aumenta la velocidad de una reacción química), permitiendo la formación de productos [11].

Cuando estas enzimas se unen en un lugar diferente al sitio activo, el lugar de unión del regulador se conoce como sitio alostérico. En el sitio alostérico la actividad catalítica puede modularse (aumentar o disminuir) mediante unión de moléculas efectoras; estos ligandos inducen rápidos cambios de conformación que pueden incrementar o reducir su velocidad de unión al sustrato [11].

La molécula de Adenosín Trifosfato (ATP) genera un suministro de energía para la realización del trabajo celular, al ser un ligando, actúa como regulador alostérico, potenciando las reacciones celulares que requieren energía [12]. El ATP en mayores concentraciones puede

comportarse como inhibidor y en menores proporciones como activador de la actividad enzimática [13].

Usando la bioinformática, es posible diseñar sustratos (productos), los cuales permiten activar o inhibir la actividad enzimática tanto en el sitio activo como en el sitio alósterico mediante la ruptura y/o formación de determinadas uniones químicas [14].

En las bases de datos de National Center for Biotechnology Information (NCBI), hasta el momento, se reportan 1'364.990 proteínas en *Homo sapiens* [15]. En muchas de estas proteínas no se conoce si hay y en qué posiciones están los sitios alostéricos; por lo tanto, en este trabajo se propone una metodología para identificar sitios alostéricos en proteínas que interactúan con el ligando ATP.

El DASP3 permite identificar secuencias de proteínas con similitud de sitio activo, a partir de un perfil creado con características de proteínas similares a las que se desean identificar, mediante métodos iterativos y automatizables, generando como resultado, grupos de proteínas funcionalmente relevantes [16].

A partir de tres proteínas con sitio alostérico de unión conocido con el ATP, se creó un perfil de sitio alostérico en el DASP3 y se identificaron proteínas, de las cuales se eligieron tres, para realizar validación mediante Docking.

El Docking es una técnica que permite predecir el acoplamiento de una proteína con un ligando, con el fin de conocer la afinidad de enlace entre las dos moléculas, mediante funciones de puntuación [17], por ejemplo, en el presente trabajo se realizó Docking en AutoDockTools mediante el ejecutable de AutoDock4 y Autogrid, en donde se generaron diversas conformaciones con valores RMSD y energías unión.

Se realizó Docking tanto a las proteínas elegidas para creación de perfil alostérico como a las proteínas seleccionadas del NCBI, a partir de este perfil. Esto con el fin de utilizar como modelo de Docking las proteínas de entrenamiento del DASP3 para comparar sus energías y valores de RMSD con las de proteínas seleccionadas a partir del ASP creado. Concluyendo que mediante la metodología propuesta posiblemente se identificaron tres proteínas con sitio alostérico en interacción con el ligando ATP.

1.1. Estado del arte

El Adenosín trifosfato (ATP), genera un aporte energético al interior de la célula por lo que su activación y/o inhibición, permite regular procesos bioquímicos en la misma [12].

Al unir el ATP como modulador alostérico con la estructura cristalina de una Hsp70, es posible estimular la unión a nucleótidos y la hidrólisis del mismo [18].

Así como también la información estructural extraída de uno de los fragmentos de unión a sitios alostéricos y del fragmento que se une a la molécula del ATP sirvió de base para el diseño de FAK (Focal Adhesion Kinase) más potentes y específicos en terapias contra el cáncer [19].

De igual manera, partir del gen GLUT1 se sintetizan proteínas con sitios de unión para

moléculas pequeñas accesibles por su cara exofacial y endofacial que incluyen secuencias con homología a sitios de unión de ATP que cumplen una función alostérica, inhibiendo la función del transporte de glucosa de GLUT1 [20], responsable de la Enfermedad de De Vivo (trastorno genético que implica una deficiencia en el transportador cerebral de glucosa tipo 1) [21].

La bioinformática es una disciplina cuyos orígenes se remontan al año 1953, en la universidad de Cambridge, en donde Watson y Crick proponen la estructura de doble hélice del ADN; en el año 1962, L. Pauling elabora su teoría sobre evolución molecular y en el año 1965 Margaret Dayhoff, publica el primer *Atlas of Protein Sequences*, ya en los años 1968 y 1970 surgen los primeros análisis con ordenador de las secuencias de ADN y proteínas[22].

Mediante la bioinformática se busca identificar nuevos sitios alostéricos en proteínas por lo que se han desarrollado diversas herramientas como son el software *fpocket*, el cual está basado en la teselación de Voronoi y esferas alfa construidas sobre la herramienta Qhull (software de código abierto para la detección de bolsillos en las proteínas) [23]; *Exploration of Protein Structural Ensembles (ExProSe)* basado en el enfoque de la geometría y la distancia euclidiana para generar conjuntos de proteínas a partir de dos estructuras (código abierto, accesible y efectivo) [24]; *STRucturally identified ESSential residues (STRESS)* de proteínas que identifican residuos superficiales e interiores que pueden actuar como puntos de acceso alostérico [25].

Por su parte el *Deacon Active Site Profiler (DASP)* permite extraer información existente en bases de datos de proteínas caracterizadas funcionalmente y algún conocimiento sobre su sitio activo [16, 26]. Mediante la herramienta DASP, se describe la creación de un ASP (active site profiler), a partir de la estructura cuaternaria cristalizada (información almacenada en archivos de extensión PDB) de las proteínas, teniendo en cuenta dos protocolos: Protocolo básico 1 (describe la creación de una firma de sitio funcional para cada proteína de estructura conocida seguida de una alineación de esas firmas para crear el perfil de sitio activo para el sitio funcional de interés) y Protocolo básico 2 (Se crea una matriz de puntuación específica de posición (PSSM) a partir del perfil se busca en una base de datos de secuencias de aminoácidos (estructura primaria de las proteínas) para proteínas con firmas de sitio funcional relacionadas.

De esta manera, la información sobre un sitio funcional de las estructuras tridimensionales conocidas se puede utilizar para buscar en bases de datos de secuencias donde la estructura secundaria o terciaria o cuaternaria de la proteína son desconocidas[27].

La validación demostrada por DASP3 proporciona una mayor separación de puntaje entre positivos verdaderos y falsos positivos que las versiones anteriores. Además, DASP3 reúne y agrupa secuencias de proteínas en grupos isofuncionales de manera más eficiente y precisa en la identificación de proteínas que contienen características de sitio activo similares a las del conjunto de consultas, con respecto a DASP y DASP2 [16].

Utilizando la herramienta de bioinformática DASP se identificaron las peroxiredoxinas (familia diversa y ubicua de peroxidases dependientes de cisteína altamente expresadas que

son importantes para la defensa antioxidante y para la regulación de las vías de señalización celular en eucariotas) de la base de datos PREX (cuenta con 3516 secuencias de proteínas de peroxiredoxina (Prx o PRDX)) [28].

Mediante el cálculo de firmas y perfiles de sitios activos se ha comparado la secuencia y la química de proteínas ciclooxigenasa (COX) con estructuras conocidas; mediante la estructura tridimensional tomada de *Protein Data Bank* (PDB), a partir de la base de datos Psi-BLAST se identificaron diez residuos de ratones y ovejas que son determinantes entre las proteínas COX-1 y COX-2. En el descubrimiento de fármacos, las proteínas de organismos, como el ratón y la oveja, se utilizan como modelos, con el objetivo de diseñar fármacos que sean selectivos para las proteínas humanas [29].

Observando el buen desempeño del DASP3, en identificación de sitios activos, se verificó la posibilidad de utilizar la herramienta DASP3 para encontrar sitios alostéricos a partir de la creación de un perfil de sitio alostérico (ASP) [30].

Para validación virtual de sitios alostéricos y/o activos de proteínas con un determinado ligando es posible realizar docking de proteína-ligando mediante diversas aplicaciones de software como se describe seguidamente.

La inhibición computacional de las proteínas Ras (enzimas de unión a trifosfato de guanosa (GTP) que regulan una serie de vías de señalización clave involucradas en el control de la división, proliferación y desarrollo celular), en cuatro sitios de unión al ligando alostérico, basado principalmente en docking ciego mediante conjuntos y FTMap (técnica computacional útil para acelerar la identificación del sitio de unión al ligando en la estructura de las proteínas diana), utiliza un enfoque continuo para el acoplamiento de fragmentos moleculares en la superficie de un receptor rígido. Los enfoques en esta revisión son un objetivo contra el cáncer aplicables a muchos miembros de la gran superfamilia Ras [31].

También se presentan tres sitios de unión alostéricos de la integrasa, una enzima que se encuentra en el virus de la inmunodeficiencia humana (VIH), para la cual la Administración de Alimentos y Medicamentos (FDA) aprobó dos medicamentos (que se dirigen al sitio activo), identificados mediante (1) AutoDock Vina más inspección visual; (2) un nuevo motor de farmacóforo común; (3) Simulaciones de energía libre de intercambio de réplica BEDAM y un enfoque de consenso que combina las predicciones de las tres estrategias. Resaltando la utilidad de las herramientas y estrategias recientemente descubiertas en un objetivo farmacológico crítico involucrado en el tratamiento de infecciones por VIH, los laboratorios Olson y Levy colaboraron en el desafío SAMPL4 que se llevó a cabo en la Universidad de Stanford, proporcionando artículos computacionales relacionados con proteínas [32].

La acetilcolinesterasa (AChE), conocida por la hidrólisis del neurotransmisor acetilcolina, con un subsitio alostérico periférico responsable de la amiloidosis en la enfermedad de Alzheimer a través de la interacción con el péptido β amiloide. Mediante el software Fpocket se identificaron 4 cavidades (sitios alostéricos y sus residuos clave) en la enzima AChE. Posteriormente se simuló virtualmente mediante docking en el software Glide. La puntuación del docking para las moléculas presentó una variación entre -7.7 y 1.8 kcal/mol. Después

se validaron los resultados pronosticados con estudios experimentales (in vitro y estudios cinéticos). Ampliando de esta manera el conocimiento de los sitios farmacológicos de AChE [33].

AutoDock4 con AutoDockTools4, es útil como herramienta para predecir la manera en que una molécula, sustratos o candidatos a fármacos, se unen a un receptor, en el análisis de ligandos unidos covalentemente, mediante el método de acoplamiento basado en la cuadrícula y la técnica de cadena lateral flexible y rígida [34].

1.2. Planteamiento del problema

Comprender como funcionan las enzimas es un principio clave para el desarrollo de muchos medicamentos farmacéuticos, pues, es en ellas donde se evidencian tanto los sitios activos como alostéricos. El conocimiento del plegamiento (estructura cuaternaria) y la identificación de los sitios ya mencionados, permite activar o inhibir la función de las proteínas [12].

Para que la proteína pueda realizar estas funciones, en los sitios activos, el sustrato debe acoplarse en su sitio de unión con la enzima. Por otra parte, en los sitios alostéricos, el ligando, se une al correspondiente sitio alostérico y modula la actividad enzimática de la proteína. Un ejemplo, es el ligando ATP que puede actuar como modulador alostérico o como sustrato en múltiples proteínas [11, 12].

La base de datos de NCBI al ser parte de la Biblioteca Nacional de Medicina de Estados Unidos (National Library of Medicine), una rama de los Institutos Nacionales de Salud (National Institutes of Health o NIH), almacena y constantemente actualiza la información referente a secuencias genómicas, además de contar con diversas bases de datos disponibles en línea de manera gratuita, entre las que incluye proteínas con secuencias y estructuras de Homo sapiens, que no han sido estudiadas, por lo que no se sabe si tiene sitio alostérico, ni su posición de interacción con la molécula ATP [15].

La bioinformática ha desarrollado software y bases de datos biológicas para la manipulación de información genómica y estructural en las células, permitiendo predecir la estructura de las proteínas a partir de su secuencia y la identificación de sitios activos [35].

El Deacon Active Site Profiler (DASP) es una herramienta de búsqueda de perfiles que utiliza mapeo de estructuras finas y perfiles para identificar motivos de importancia funcional. El DASP requiere la selección de residuos clave que definen un sitio funcional dentro de una familia de proteínas y luego identifica fragmentos de secuencia cercanos [16].

A pesar que el DASP3 es una herramienta elaborada en principio para la identificación de sitios activos, es posible utilizar esta misma herramienta para hallar sitios alostéricos mediante la creación de un perfil alostérico [30]. Son varios los trabajos realizados en el DASP para la identificación de sitios activos en proteínas mediante la creación de un ASP, demostrando un buen desempeño, aún así, solamente en [30] se evidencia la identificación de proteínas con sitio alostérico en esta herramienta, sin embargo, no hay validación rigurosa (Docking o in-vitro).

Teniendo en cuenta que la base de datos de NCBI cuenta con proteínas en *Homo sapiens* que no han sido estudiadas, por lo que no se sabe si tiene sitio alostérico, ni su posición de interacción con la molécula ATP [15], surge la siguiente pregunta: ¿Es posible identificar y validar sitios alostéricos en las proteínas que interactúan con la molécula de ATP en *homo sapiens* del NCBI mediante la herramienta DASP3?

1.3. Justificación

El diseño de fármacos ha crecido exponencialmente, en gran parte gracias a la bioinformática pues esta aplica modelos y algoritmos matemáticos para la determinación, simulación y entendimiento de proteínas [35].

Se conocen dos tipos de medicamentos: ortostéricos, se unen en el sitio activo; y alostéricos, se unen a otras partes de la superficie de la proteína y cambian alostéricamente la conformación del sitio de unión en la misma [36].

La identificación de sitios alostéricos ha demostrado ser apropiada para la inhibición de la actividad enzimática, además de reducir el riesgo de toxicidad, por lo que espera tener menos efectos negativos en la salud humana, que la inhibición en sitios activos [37, 38].

El descubrimiento de sitios alostéricos proporciona una base a los biólogos, químicos y/o profesionales que trabajan en este campo en colaboración con otros científicos, para el diseño y la mejora de medicamentos [39, 40], además amplía el desarrollo de medicamentos disponibles, brindando una oportunidad para descubrimiento de medicamentos alostéricos [41, 42].

En algunos casos, el ATP, al ser un modelo conocido permite obtener aproximaciones iniciales en la identificación de sitios alostéricos, pues puede comportarse como modulador alostérico [36, 12].

Por tanto, descubrir similitudes en los sitios alostéricos de *Homo sapiens* mediante la herramienta DASP3 a partir de la secuencia de aminoácidos de tres proteínas, a las cuales se le conoce su actividad alostérica, con interacción del ligando de ATP y compararlas con las de la base de datos de NCBI, ayudará a identificar proteínas en las que no se conoce su sitio alostérico ni su interacción con la molécula de ATP, lo que ampliará el conocimiento de nuevas proteínas que cuentan con el ATP como modulador alostérico, permitiendo activar o inhibir las proteína que realizan diversas funciones, algunas ayudan al ensamblado de estructuras celulares como la membrana plasmática, el citoesqueleto y otros orgánulos. Otras funcionan como hormonas, anticuerpos y elementos contráctiles en el tejido muscular. Otras actúan como enzimas y regulan el ritmo de numerosas reacciones químicas en las células, o como transportadores y llevan diversos materiales en la sangre [43]; lo que podría tener gran importancia en el estudio de enfermedades como son: la diabetes, la obesidad, el cáncer, las enfermedades del tiroides, el envejecimiento, enfermedades mitocondriales, entre otras, que afectan principalmente a los tejidos que exigen energía, tales como el sistema nervioso, el músculo esquelético y el corazón [44].

1.4. Objetivos

1.4.1. Objetivo general

Identificar proteínas en Homo sapiens que tienen un sitio alostérico que interactúa con la molécula de ATP haciendo uso de la herramienta DASP3.

1.4.2. Objetivos específicos

1. Seleccionar tres (3) proteínas en Homo sapiens en las que se conoce en la literatura su actividad alostérica con el ligando Adenosín Trifosfato (ATP).
2. Identificar, por medio de un perfil alostérico (herramienta DASP3), proteínas candidatas a tener actividad alostérica regulada por el ATP.
3. Estimar la estructura geométrica de las proteínas elegidas identificadas con el DASP3 que no posean estructura tridimensional (PDB).
4. Verificar, por medio de Docking, a partir del PDB de las tres (3) proteínas elegidas, que los sitios alostéricos identificados con la herramienta DASP3 efectivamente forman un bolsillo alostérico para el ligando ATP.

1.5. Alcance y limitaciones

Este proyecto se realizó con estructuras tridimensionales de las proteínas, obtenidas mediante cristalografía de rayos X, Resonancia Magnética Nuclear (RMN) y/o microscopía electrónica, adquiridas de la base de datos *PDB (Protein Data Bank)*.

La comprobación de los sitios alostéricos que interactúan con la molécula de ATP se realizó de manera virtual, por medio de simulaciones. Por tanto si se desea saber si la estructura 3D que se está generando a partir de la estructura primaria de la proteína es auténtica en la variación del tiempo sería necesario realizar ensayos de laboratorio (*in-vivo*) o dinámica molecular.

Se identificaron sitios alostéricos en interacción con el ligando ATP, en las proteínas de Homo sapiens tomadas de la base de datos de NCBI, además se propuso una metodología para encontrar proteínas con sitios alostéricos que interactúen con un determinado ligando. A pesar de haber más sustratos que actúen como inhibidores o activadores alostéricos, sólo se realizó para la molécula de ATP pues se conoce como un buen modulador alostérico.

1.5.1. Ubicación dentro de las líneas de trabajo del programa

Dentro de las líneas de trabajo del programa de ingeniería Biomédica encontramos:

- Bioinformática.
- Bioquímica
- Procesamiento de señales.

1.5.2. Usuarios directos y formas de utilización de los resultados del proyecto

Este proyecto va dirigido a los biólogos y/o científicos como complemento para la identificación de proteínas que tienen sitio alostérico e interactúan con el ligando ATP. Así mismo va dirigido a personas interesadas en el diseño y/o elaboración de medicamentos cimentados en el sitio alostérico.

1.6. Estructura del documento

El presente trabajo integral de grado se encuentra dividido en 5 capítulos de la siguiente manera:

- Capítulo 1: se presenta contextualización de lo que abarca el trabajo, teniendo en cuenta el por qué y para qué, así como el objetivo, el alcance y las limitaciones del trabajo.
- Capítulo 2: se realiza una revisión de los conceptos expuestos en el trabajo, como son, la importancia de las proteínas, dentro de ellas, las enzimas, con el sitio activo y sitio alostérico, además de la presentación de los software empleados como son el DASP3, AutoDockTools, I-TASSER, Swiss-PdbViewer, UCSF Chimera y Avogadro.
- Capítulo 3: se presenta explicación respecto a la creación de perfil alostérico con respectivos protocolos en el DASP3 para la identificación de secuencias y estructuras de proteínas en el NCBI.
- Capítulo 4: se presenta verificación de proteínas (identificadas del NCBI), mediante realización de docking con AutoDock4 en la interfaz de AutoDockTools.
- Capítulo 5: se presentan conclusiones del trabajo, y recomendaciones para futuros trabajos.

2. Marco Teórico

En este capítulo se presenta una revisión de los conceptos fundamentales durante el desarrollo del presente trabajo.

2.1. Proteína

Las proteínas determinan las características físicas y químicas de las células y, por ende, de los organismos que están constituidos por éstas [43]. Las proteínas son el principal componente estructural de las células y, están formadas fundamentalmente por carbono, hidrógeno, oxígeno y nitrógeno [7].

En la figura 2-1, se observa la síntesis de una proteína específica que requiere la transcripción del ADN de un gen en una molécula de ARN y la traducción del ARN en su correspondiente cadena de aminoácidos.

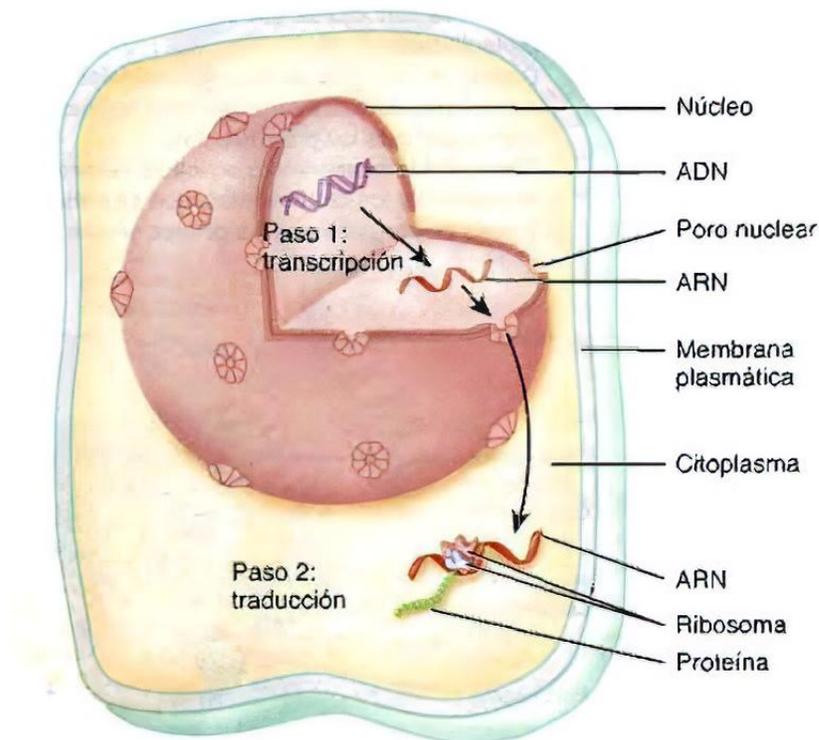


Figura 2-1.: Esquema general de la expresión genética

Las proteínas son cadenas de aminoácidos unidas que forman una estructura tridimensional, esta forma les permite llevar a cabo diferentes funciones dentro del organismo que van desde su papel catalítico (enzimas) hasta su función en la motilidad corporal (actina, miosina), pasando por su papel mecánico (elastina, colágeno), de transporte y almacén (hemoglobina, mioglobina, citocromos), protección (anticuerpos), reguladora (hormonas) y, están codificadas en el material genético de cada organismo, donde se especifica su secuencia de aminoácidos, y luego son sintetizadas por los ribosomas [8].

2.1.1. Estructura de las proteínas

Todas las proteínas poseen una misma estructura química central, que consiste en una cadena lineal de aminoácidos. Lo que hace distinta a una proteína de otra es la secuencia de aminoácidos de que está hecha, a tal secuencia se conoce como estructura primaria de la proteína [5]. A continuación, en la figura 2-2 se evidencian los cuatro niveles de la estructura de una proteína, así mismo, en la tabla 2-1, adaptada de [2, 3, 5, 6], se resumen algunas de sus características.

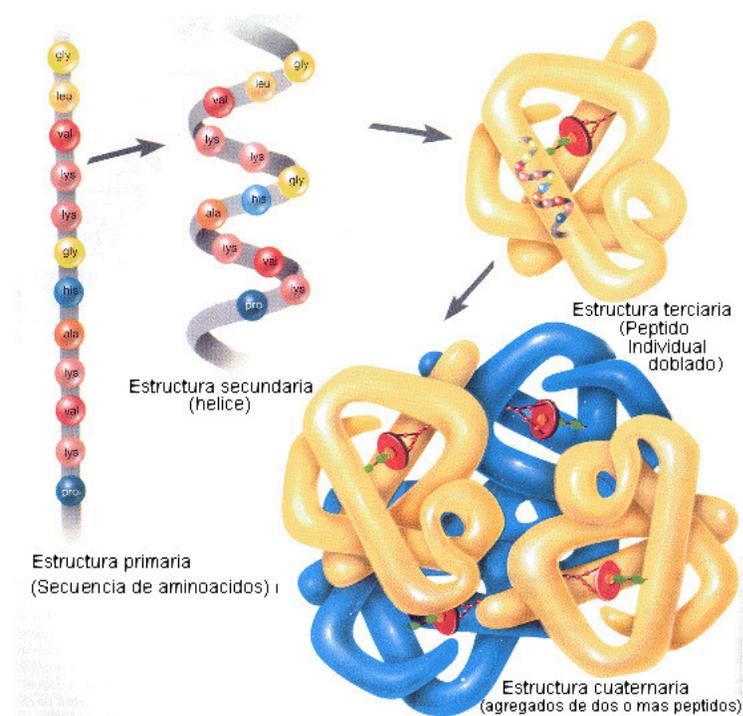


Figura 2-2: Niveles de estructura de una proteína [2]

Tabla 2-1.: Estructuras de las proteínas

Estructuras			
Primaria	Secundaria	Terciaria	Cuaternaria
Determinada por la secuencia de aminoácidos en la cadena proteica. Las posibilidades de estructuración a nivel primario son prácticamente ilimitadas. Los enlaces peptídicos forman el esqueleto de la proteína, del que emergen las cadenas laterales de los aminoácidos.	Adopta espacialmente una parte del polipéptido. Ocurre cuando los hidrógenos de la secuencia interactúan mediante puentes de hidrógeno. Los puentes de hidrógeno se establecen entre los grupos -CO- y -NH- del enlace peptídico (el primero como aceptor de H, y el segundo como donador de H).	Disposición tridimensional de todos los átomos que componen la proteína, concepto equiparable al de conformación absoluta en otras moléculas. directa de sus propiedades biológicas, pues la disposición espacial de los distintos grupos funcionales determina su interacción con los diversos ligandos.	Cuando una proteína consta de más de una cadena polipeptídica, es decir, cuando se trata de una proteína oligomérica, decimos que tiene estructura cuaternaria.
Conocer la estructura primaria de una proteína no solo es importante para el estudio de enfermedades genéticas. Es posible que el origen de una enfermedad genética radique en una secuencia anormal; si es severa, podría resultar en que la función de la proteína no se ejecute de manera adecuada o que no se ejecute.	Alfa helice Estructura helicoidal, que se mantiene mediante puentes de hidrógeno, con los grupos R orientados hacia el exterior. Para formar esta estructura, el grupo carboxilo de cada aminoácido (n) se une mediante un puente de hidrógeno al grupo amino de otro aminoácido (n+4). Es una estructura estable porque da lugar a un máximo número de interacciones.	Fibrosas Estructuras con forma de fibra o lámina. Insolubles en el agua. Las proteínas que dan forma y protección a los organismos suelen ser fibrosas. Las proteínas fibrosas se forman por repetición de estructuras secundarias simples.	Globulares Estructuras globulares. Solubles en el agua. Muchas enzimas y proteínas reguladoras tienen esta forma. Las proteínas globulares tienen una estructura terciaria más compleja formada a partir de varias estructuras secundarias diferentes.
	Conformación beta La cadena queda estirada y la estructura se dispon espacialmente en zigzag formando láminas (hojas plegadas β). La disposición puede ser paralela o antiparalela. Puede darse entre regiones próximas o distantes del polipéptido. Los grupos R sobresalen de la lámina en ambos sentidos, de forma alterna.		

2.1.2. Propiedades y funciones de las proteínas

Las proteínas son sustancias de elevado peso molecular, las propiedades conocidas dependen, en su mayoría, de radicales R libres y la relación de estos con el medio, entre estas tenemos [5].

- Especificidad. Las proteínas son específicas en cada especie y, en algunos casos, en cada individuo. Esto se debe a su estructura primaria determinada y una conformación espacial propia; un cambio de estructura significaría una pérdida de la función. Esta propiedad es la responsable de rechazos en trasplantes e injertos. Al introducir una nueva proteína el organismo la reconoce como extraña y fabrica anticuerpos para su destrucción.
- Desnaturalización. Es un cambio en la estructura tridimensional de la proteína cuando se la somete a cambios bruscos de pH, de temperatura, acción de rayos ultravioletas, detergentes, ácidos o bases fuertes, agitación fuerte que provoquen ruptura de los enlaces, pérdida de su estructura y de su función. La desnaturalización provoca generalmente una disminución de la solubilidad y la proteína precipita. Si la desnaturalización es suave, al restablecer las condiciones iniciales puede recuperar su estructura y a esto se le llama renaturalización.
- Solubilidad. Por su tamaño, las proteínas son insolubles en agua o se disuelven formando coloides. Este es el caso de las proteínas globulares cuyos radicales se ionizan y establecen puentes de hidrógeno con el agua formando coloides. Las fibrilares son insolubles.

Por otro lado, las funciones de las proteínas son de gran importancia pues determinan la forma y estructura de las células, guían procesos vitales, defiende a las células de agentes externos. Las funciones dependen de cada tipo de proteína, aunque todas realizan su función por unión selectiva a moléculas. Las principales funciones son [5]:

- Estructural. Las proteínas forman tejidos de sostén y relleno que otorgan elasticidad y resistencia a órganos y tejidos (ejemplos: colágeno, gelatina, glucoproteínas, queratinas)
- Enzimática (catalítica). Las proteínas actúan como biocatalizadores acelerando las reacciones químicas del metabolismo. A las sustancias que se transforman por medio de una reacción enzimática se les llama sustratos. Los sustratos reconocen un sitio específico en la superficie de la proteína que se denomina sitio activo.
- Hormonal. Muchas proteínas son hormonas, a diferencia de las enzimas, su función no es local sino en todo el organismo (ejemplos: insulina, FSH, tiroxina)
- Defensiva. Las proteínas crean anticuerpos y regulan factores contra agentes extraños o infecciones.

- Transporte. Las proteínas pueden transportar distintos tipos de sustancias (ejemplos: hemoglobina, citocromos, lipoproteínas).
- Reserva. Las proteínas cumplen una función energética para el organismo. (ejemplo: las albúminas y la caseína que actúan como almacén de aminoácidos)
- Contráctil. La contracción de los músculos través de la miosina y actina es una función de las proteínas contráctiles que facilitan el movimiento de las células constituyendo las miofibrillas que son responsables de la contracción de los músculos.
- Homeostática. Las proteínas funcionan como amortiguador para mantener estables las constantes del medio, desde el pH hasta el equilibrio osmótico.

2.1.3. Enzimas

Las enzimas son biomoléculas de naturaleza proteica que buscan alcanzar el equilibrio acelerando la velocidad de reacción. Son catalizadores excepcionales porque [45]:

- Son muy eficientes pues, en condiciones óptimas las reacciones enzimáticas ocurren más rápidamente que sin catalizar.
- Son específicas tanto respecto al tipo de reacción que catalizan como respecto al sustrato que transforma.
- Catalizan diferentes reacciones (hidrólisis, polimerizaciones, redox, transferencias de grupos, condensaciones, etc.)
- Su concentración final es objeto de control genético.

Estructura de las enzimas

Las enzimas son proteínas globulares formadas por una o más cadenas polipeptídicas plegadas creando una “hondonada” donde encaja el sustrato y tiene lugar la reacción. Esta zona de la enzima se denomina centro activo y sólo unos pocos aminoácidos están implicados en él [11]. La proximidad de los aminoácidos en el centro activo está determinada por la estructura terciaria, aunque también pueden ocupar posiciones adyacentes en la estructura primaria. En una enzima con estructura cuaternaria, los aminoácidos del centro activo pueden encontrarse incluso en diferentes cadenas. En esta estructura tridimensional se forman cavidades, llamadas “sitio activo”, las cuales muestran afinidad por las moléculas específicas (sustratos) que se convertirán en productos[46].

Sitio Activo

Región específica de la enzima que es ocupada por el sustrato, este interacciona con la enzima mediante enlaces débiles; en este sitio ocurre la activación del sustrato y se realiza la reacción [10]. Como cualquier catalizador, al finalizar la transformación del sustrato y liberarse el producto del sitio activo, la enzima regresa a su estado original y puede involucrarse en un nuevo ciclo de catálisis [46]. Dentro de todo el conjunto de enzimas las hay que presentan una alta especificidad, aceptando tan sólo un tipo de moléculas sobre las que realizar la catalización, y siendo capaces de discriminar incluso entre moléculas isoméricas; por otro lado, otras enzimas con un menor nivel de especificidad catalizan reacciones utilizando como sustratos moléculas que presenten una cierta similitud.

La interacción entre enzima y sustrato se realiza a través de enlaces de naturaleza débil entre la molécula de sustrato y el centro activo. Cuanto mayor sea el número de estos enlaces, mayor será la especificidad de la enzima, y mayor también su capacidad de discriminar entre dos sustratos estructuralmente próximos. Lo anteriormente referenciado puede evidenciarse en la figura 2-3.

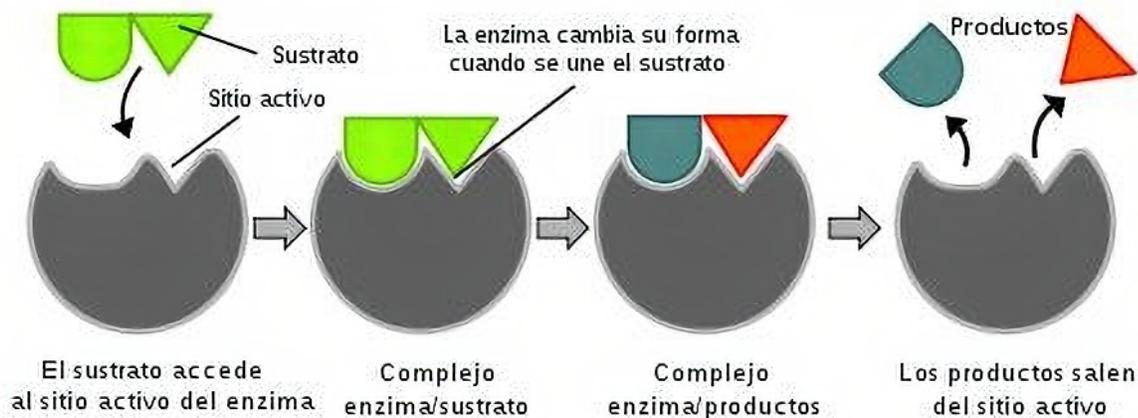


Figura 2-3.: Unión de la enzima con el sustrato
[10]

2.1.4. Sitio Alostérico

Lugar o punto de una enzima que al unirse con efector o modulador desencadena un cambio que puede alterar propiedades catalíticas o de unión. A su vez, existen bolsas o bolsillos los cuales tienen contenidos residuos identificados que al unirse con un modulador alostérico desencadena cambios. La importancia del alosterismo es que constituye un medio a través de la cual la célula realiza la regulación de la actividad de enzimas o proteínas que dependen de condiciones variables o alterables que son necesarias en cada momento [30].

El sitio alostérico carece de actividad catalítica, pero se une a una molécula efectora o modu-

ladora que puede inhibir (modulador o efector negativo o inhibidor) o estimular (modulador o efector positivo o activador) la actividad enzimática, mediante modificaciones en la conformación del sitio activo haciéndolo menos o más afín al sustrato, respectivamente; cuando ingresa la primera molécula de sustrato aumenta la afinidad del sitio activo por el sustrato de las otras subunidades, este efecto dirigido por cambios conformacionales favorables se denomina cooperatividad [10].

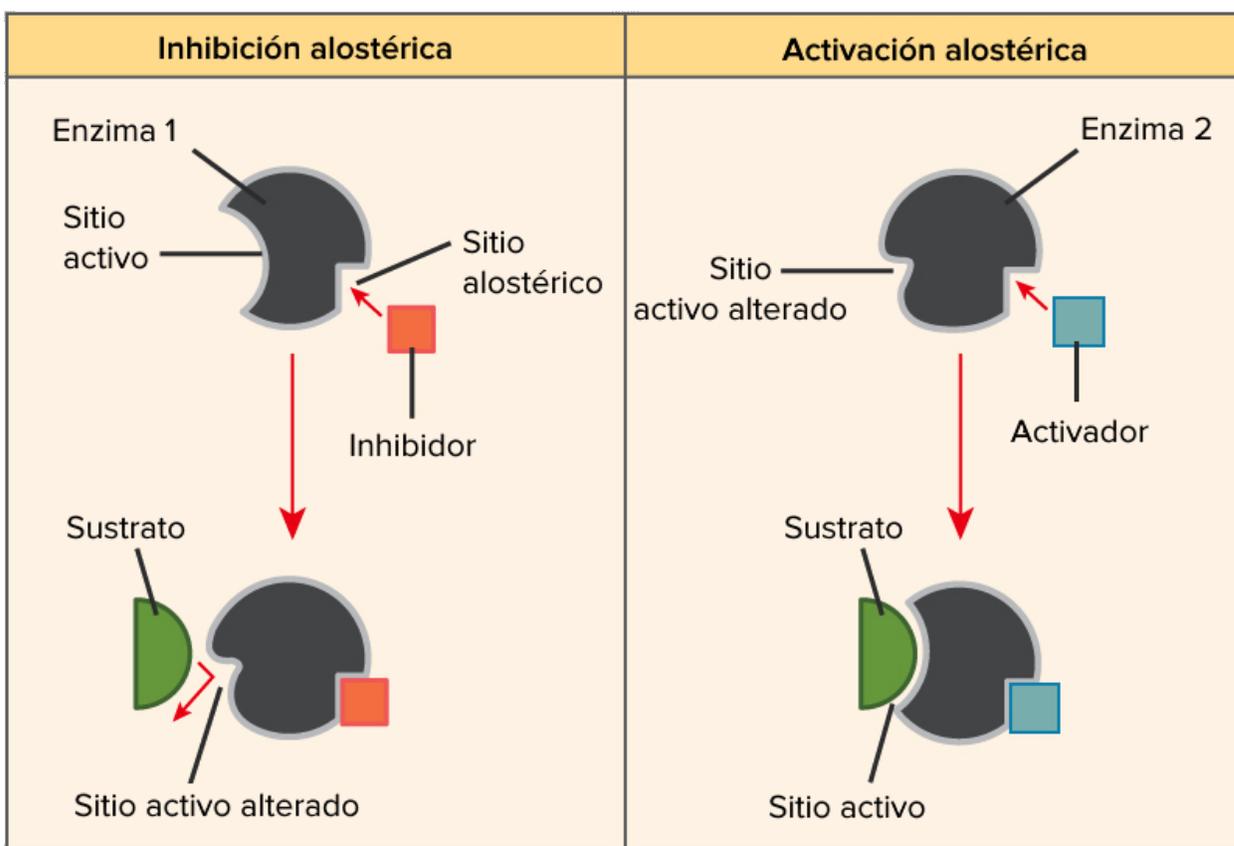


Figura 2-4.: Unión de la enzima con el sustrato en el sitio alostérico [12]

2.1.5. Adenosin Triphosphate (ATP)

Es un nucleótido fundamental en la obtención de energía celular. Está formado por una base nitrogenada (adenina) unida al carbono 1 (C.1) de un azúcar de tipo pentosa, la ribosa, que en su carbono 5 (C.5) tiene enlazados tres grupos fosfato [47], figura 2-5. Se produce durante la fotorrespiración y la respiración celular, y es consumido por muchas enzimas en la catálisis de numerosos procesos químicos. Cada proceso dentro de una especie, incluyendo la síntesis de ADN y proteínas, contracción muscular, transporte activo de nutrientes, actividad nerviosa, mantenimiento de osmosis, o fijación de carbono, requiere una fuente de ATP [48]. La molécula de ATP, al ser un ligando, puede actuar como regulador alostérico, comportándose-

se como inhibidor de la actividad enzimática en mayores concentraciones y como activador en menores proporciones [11, 13].

En la figura 2-5, la imagen superior fue tomada de: [49] y la imagen inferior de: [50].

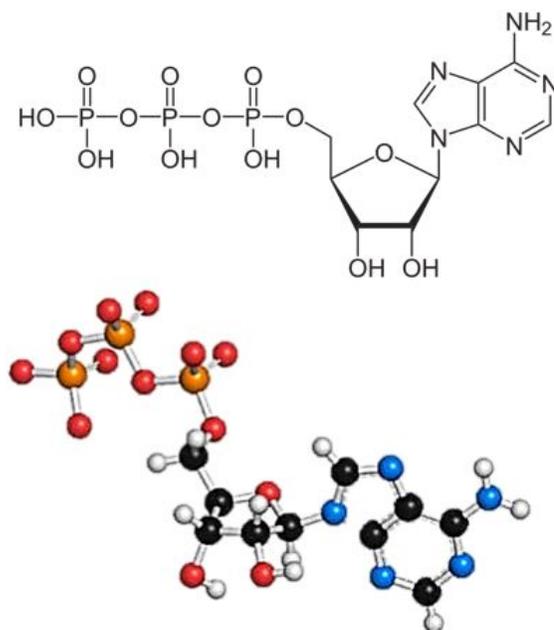


Figura 2-5.: Estructura molecular Adenosin Triphosphate (ATP)

2.2. Bioinformática

La bioinformática surge debido a la problemática dada en fenómenos tan complejos como la genética, la simulación del efecto de medicinas, la predicción de enfermedades, etc. Todas estas actividades manejan gran cantidad de información y variables, de allí surge la necesidad de apoyarse con las nuevas tecnologías [51].

La bioinformática estructural consiste, entre otras cosas, en realizar una simulación del comportamiento de biomoléculas, principalmente proteínas y sus entornos, en diferentes situaciones.

La simulación con las biomoléculas puede establecer la capacidad de una droga para inhibir algunas enzimas o la posibilidad que dos proteínas puedan interactuar entre ellas. De este modo la bioinformática contribuirá a un cambio en el paradigma del diseño de drogas permitiendo una mayor rapidez en su descubrimiento y en el proceso de optimización [35].

El paradigma del descubrimiento de nuevas drogas o fármacos consiste en encontrar una molécula pequeña capaz de unirse e inhibir/activar eficientemente un blanco macromolecular específico, generalmente una proteína. La característica principal que debe poseer una droga es la de unirse fuertemente a la proteína blanco. Para las tareas mencionadas anteriormente, se han desarrollado diversas herramientas.

A continuación, en la tabla **2-2**, adaptada de [30], se describen algunas herramientas conocidas en el campo de la bioinformática para predecir sitios activos y/o alostéricos.

Tabla 2-2.: Herramientas en el campo de la bioinformática para predecir sitios alostéricos y/o activos

Sotware	Características
ExProSE	Capaz de generar conjuntos de estructuras de proteínas a partir de dos estructuras de entrada y predecir bolsas alostéricas en proteínas. Se muestra usando un conjunto de datos de pares estructurales que dos métodos ampliamente utilizados para generar conjuntos no pueden abarcar grandes cambios conformacionales. Al agregar restricciones adicionales que representan moduladores potenciales, ExProSE puede predecir sitios alostéricos.
Fpocket	Extrae información de la estructura tridimensional de la proteína, basándose en parámetros geométricos, contenidos en tres pasos: el primero identifica grupos de esferas α determinando la estructura de la proteína, el segundo reconoce las bolsas alostéricas y elimina las que no son de interés y, finalmente se da puntuación y se categorizan las bolsas identificadas.
ESTRESS	Identificación de residuos críticos, empleando el modelo conformacional de proteínas, que identifican los residuos superficiales e interiores que pueden actuar como puntos de acceso alostérico, permitiendo un análisis que incorpore directamente la información de la estructura y dinámica de las proteínas 3D
DASP	Deacon Active Site Profiler es un software para la creación de perfiles de sitios activos, fue desarrollado previamente para reconocer secuencias de proteínas con una similitud de sitio activo a un conjunto de consultas, permitiendo la extracción de características de la secuencia estructural de un sitio funcional y la conservación o variación de esas características en la familia de proteínas. El DASP está conformado por dos métodos iterativos y automatizables para agrupar proteínas en grupos funcionalmente relevantes han demostrado una gran capacidad en la identificación de estudios de las superfamilias de proteínas.

2.2.1. DASP3

Es una herramienta diseñada para la creación de perfiles de sitios activos para cualquier familia de proteínas y, cabe destacar que no requiere información 3D de la proteína, sino que, emplea dos protocolos para el análisis de sitios funcionales, el primero crea firmas de sitios funcionales o perfiles de proteínas de estructura conocida (Active Site Profiler - ASP), el segundo, usa el perfil de sitio activo para buscar secuencias que contienen fragmentos similares [16]. Considerando el buen desempeño del DASP3 para la identificación de sitios activos y la similitud de los problemas de identificación de sitios activos y de sitios alostéricos, en

este trabajo se evalúa y se presenta una metodología para utilizar el DASP3 para identificar sitios alostéricos [30]. Durante el desarrollo de procesos de búsqueda iterativos que agrupan proteínas en grupos isofuncionales, se identificaron tres limitaciones del algoritmo DASP. Se exploraron tres mejoras para resolver estas limitaciones de búsqueda iterativa: longitud mínima de motivo utilizable, proceso de alineación de motivo dentro del perfil y longitud de motivo utilizada en la búsqueda [16].

2.2.2. Docking

Docking es una técnica para realizar el acoplamiento molecular de proteína-proteína o de proteína-ligando. Cuando se trata de un acoplamiento proteína-ligando, el objetivo es predecir la posición y orientación de un ligando cuando está unido en el sitio activo y/o alostérico de una proteína o enzima [17]. Se emplea en la investigación farmacéutica, permitiendo el diseño de medicamentos por computadora [52, 17], incluida la identificación de sitios de unión crípticos o alostéricos, la mejora de las metodologías tradicionales de detección virtual, y la predicción directa de las energías de unión de moléculas pequeñas[53].

Para la realización del acoplamiento de proteínas y ligandos se encuentran disponibles varias aplicaciones de software que calculan el sitio, la geometría y la energía de moléculas pequeñas o péptidos que interactúan con proteínas, como son: AutoDock4 y AutoDock Vina , rDock , FlexAID , FMap, Molecular Operating Environment y Glide .

AutoDock4

AutoDock es software libre, de código abierto, con un conjunto de herramientas de acoplamiento automatizadas. Está diseñado para predecir cómo las moléculas pequeñas, como sustratos o candidatos a fármacos, se unen a un receptor de estructura 3D conocida [54]. Las distribuciones actuales de AutoDock constan en dos generaciones de software: AutoDock 4 y AutoDock Vina .

AutoDock4 consta principalmente de dos programas [54]:

- Autodock realiza el acoplamiento del ligando a un conjunto de cuadrículas que describen la proteína objetivo.
- Autogrid calcula previamente estas cuadrículas o rejillas (lugar donde se realizará el docking en la proteína).

AutoDock 4 comprende tres mejoras significativas [54]:

1. Los resultados de acoplamiento son más precisos y confiables.
2. Opcionalmente, puede modelar la flexibilidad en la macromolécula objetivo.
3. Permite el uso de AutoDock en la evaluación de interacciones proteína-proteína.

AutoDock 4.0 no solo es más rápido que las versiones anteriores, sino que permite que las cadenas laterales en la macromolécula sean flexibles. El acoplamiento rígido es más rápido, y el acoplamiento flexible de alta calidad se puede realizar en aproximadamente un minuto. Se pueden realizar hasta 40,000 acoplamientos rígidos en un día en una CPU [54].

AutoDockTools (ADT)

Es una interfaz gráfica de usuario (GUI) gratuita, para el empleo de AutoDock, posee compatibilidad con los formatos de AutoDock4 y Autogrid. Puede usarse para configurar, ejecutar y analizar acoplamientos AutoDock e isocontour de mapas de afinidad AutoGrid, así como calcular superficies moleculares, mostrar cintas de estructura secundaria, calcular enlaces de hidrógeno, entre otros [55, 54].

Instalación AutoDock4 con AutoDockTools

A partir de la estructura tridimensional se procedió a realizar el Docking. Para la realización del docking fue necesario preparar el entorno para correr Autodock4:

1. Instalar python
2. Instalar MGL Tool 1.5.6 (interfaz gráfica de usuario a utilizar AutoDockTools/ADT)
3. Instalar AutoDock 4.2.3
4. Instalar Cygwin (provee un entorno linux en windows)
5. Después se descargaron los PDB de la molécula y el ligando ATP
6. Se guardaron en una misma carpeta llamada Docking1, junto con autogrid4.exe, autodock4.exe, AD4_parameters.dat y AD4.1_bound.dat como se observa en la figura 2-6.

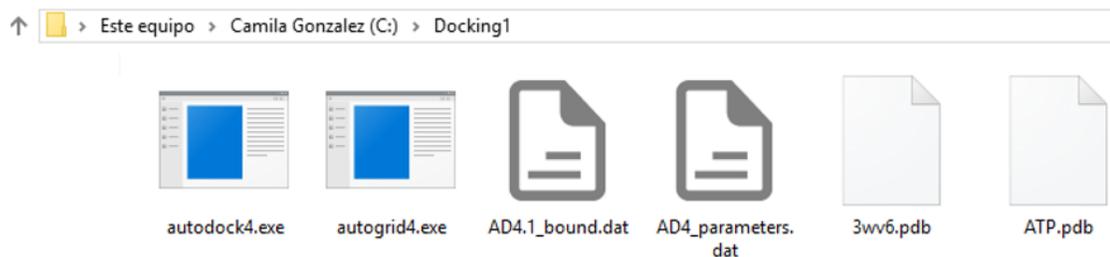


Figura 2-6.: Carpeta Docking con archivos necesarios para correr AutoDockTools (Autor)

2.2.3. I-TASSER

I-TASSER es un método informático para predecir la estructura terciaria de la proteína a partir de su estructura primaria (secuencia de aminoácidos) [56]. Cuenta con un servidor en línea construido en el Laboratorio Yang Zhang de la Universidad de Michigan , Ann Arbor , que permite a los usuarios enviar secuencias y obtener predicciones de estructura y función. Se puede descargar un paquete independiente de I-TASSER en el sitio web de I-TASSER <http://zhang.bioinformatics.ku.edu/I-TASSER>. Constantemente ha sido calificado como el mejor método de predicción de CASP (Critical Assessment of protein Structure Prediction): experimento mundial para la predicción de la estructura de la proteína, se realiza cada dos años desde 1994 [57]. Además de ser capaz de predecir pliegues correctos y modelos de alta resolución [58].

Funcionamiento I-TASSER

Teniendo en cuenta una secuencia de aminoácidos con longitud de entre 10 a 1.500 residuos I-TASSER [56]:

1. Genera modelos atómicos tridimensionales a partir de múltiples alineamientos de rosado y simulaciones iterativas de ensamblaje estructural.
2. La función de la proteína se infiere haciendo coincidir estructuralmente los modelos 3D con otras proteínas conocidas.
3. El resultado de una ejecución del servidor contiene predicciones de estructura secundaria y terciaria completas y anotaciones funcionales en los sitios de unión a ligandos, números de la Comisión de enzimas y términos de ontología génica.

2.2.4. Swiss-PdbViewer (también conocido como DeepView)

Ha sido desarrollado desde 1994 por Nicolas Guex. Es una aplicación con una interfaz amigable con el usuario, permite analizar varias proteínas al mismo tiempo. Las proteínas pueden superponerse para deducir alineamientos estructurales y comparar sus sitios activos o cualquier otra parte relevante. Las mutaciones de aminoácidos, los enlaces H, los ángulos y las distancias entre los átomos son fáciles de obtener gracias a la interfaz gráfica y de menú intuitiva. Además en el panel de control se cuenta con distintas técnicas de Modelado Molecular y Minimización de la Energía. Está estrechamente vinculado a SWISS-MODEL, un servidor de modelado de homología automatizado desarrollado dentro del Instituto Suizo de Bioinformática (SIB) en el Grupo de Bioinformática Estructural en el Biozentrum en Basilea. Trabajar con estos dos programas reduce en gran medida la cantidad de trabajo necesario para generar modelos, ya que es posible enhebrar una secuencia primaria de proteína en una plantilla 3D y obtener una respuesta inmediata de qué tan bien la estructura de referencia aceptará la proteína enhebrada antes de enviarla [59].



Figura 2-7.: Icono Swiss-PdbViewer
[59]

2.2.5. UCSF Chimera

Es un programa para la visualización interactiva y el análisis de estructuras moleculares y datos relacionados, incluidos mapas de densidad, trayectorias y alineaciones de secuencias. Se pueden generar imágenes y animaciones de alta calidad. Mediante las herramientas de Chimera es posible comparar diferentes conformaciones e incluso diferentes proteínas mediante la transformación de una estructura a otra, muestra secuencias individuales y alineaciones de secuencias múltiples, permite preparar estructuras para el acoplamiento u otros cálculos agregando hidrógenos, asignando cargas parciales entre otras tareas relacionadas [60].



Figura 2-8.: Icono UCSF Chimera
[60]

2.2.6. Avogadro

Es un software libre, de código abierto e intuitivo, que permite editar y visualizar moléculas. Se utiliza en química computacional, la bioinformática, las ciencias de los materiales y otras áreas relacionadas [61]. Por otro lado, mediante el software de Avogadro, es posible realizar minimización de energía a las moléculas para que estas conserven una distancia óptima entre sus átomos, mejorando la estructura mediante un minimizador de energía mecánica molecular. Una extensión del programa permite preparar archivos de estímulo para realizar cálculos. La principal ventaja del software es permitir visualizar las moléculas desde diferentes ángulos y también desplazarlos en el espacio usando los comandos del programa.

Desde el punto de vista didáctico, al facilitar la visualización de átomos en tercera dimensión conduce a un mejor aprendizaje del mundo a nivel molecular. La gran capacidad interactiva del software fomenta un aprendizaje activo ya que se puede ver desde diferentes perspectivas la forma de cada compuesto, rotar las moléculas y acomodar cada enlace, además de poder realizar y comprobar sus propios enlaces de manera mucho más atractiva y dinámica [62].



Figura 2-9.: Icono Avogadro
[61]

2.3. Desviación cuadrática media (RMSD)

En bioinformática es la medida entre dos poses (describe el modo de unión de un ligando en un sitio de unión a proteínas) de la misma molécula con el ligando acoplada en la misma estructura de la proteína. El cálculo de RMSD se puede aplicar a otras moléculas no proteicas, como las moléculas orgánicas pequeñas. Esto es particularmente relevante para comparar ligandos con la estructura de la proteína, pues el RMSD mide la distancia promedio entre átomos de 2 estructuras de ligando a través de la ecuación simple [63]:

$$RMSD_{(a,b)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (a_{ix} - b_{ix})^2 + (a_{iy} - b_{iy})^2 + (a_{iz} - b_{iz})^2} \quad (2-1)$$

donde a_{ix} se refiere a la posición en el eje x del i -ésimo átomo de la molécula a , y b_{ix} se refiere a la posición en el eje x del i -ésimo átomo de la molécula b . La variación de los subíndices $\{x, y, z\}$ denota las coordenadas cada átomo.

La medida del valor RMSD con la primera pose como referencia es bastante útil para entrenar, probar y comparar diferentes funciones de puntuación. Por ejemplo, en docking, es posible determinar la deformación que ha sufrido el ligando para acoplarse a la proteína, en el mismo sitio de unión [63].

Para el análisis del valor RMSD en la herramienta de AutoDockTools (ADT), se obtienen de entre 3 y 10 conformaciones (posiciones en la que se acomoda el ligando en la proteína), en donde cada conformación posee un valor RMSD que es promediado con la primera conformación, tomando como referencia que un valor menor a 2 Å, es bueno, es decir, el ligando es estable dentro de las conformaciones (no tuvo que esforzarse para acomodarse dentro de la proteína).

*En el presente trabajo la unidad de medida para el valor RMSD es tomada en ángstrom (Å).

3. Identificación de proteínas mediante realización de perfil alostérico

En este capítulo se identificaron 3 proteínas que poseen actividad alostérica con el ATP, para elaboración de perfil alostérico (ASP) mediante la herramienta DASP3 y se comparó con las proteínas de homo sapiens localizadas en la base de datos NCBI. En el diagrama de bloques de la figura 3-1, se presentan de manera general los pasos para la elaboración del perfil alostérico.

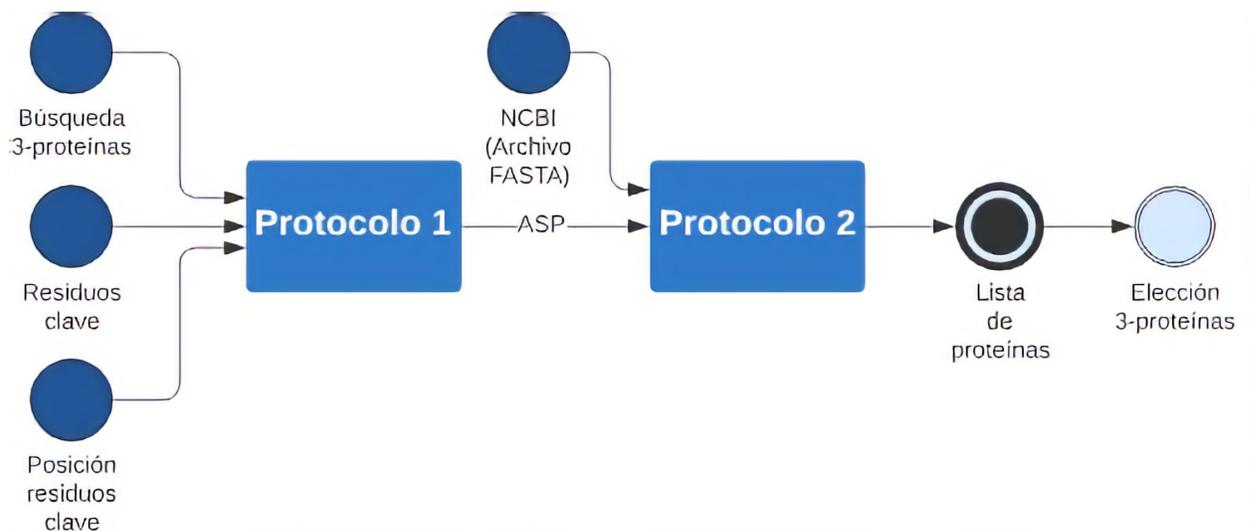


Figura 3-1.: Diagrama bloques con procedimiento para identificar proteínas mediante DASP3, Imagen de Autor

Donde:

- **Protocolo 1:** describe la creación de una firma de sitio funcional para cada proteína a la que se le conoce su estructura y su sitio alostérico en interacción con ATP, seguidamente se genera una alineación de los residuos para crear el perfil de sitio alostérico (ASP).
- **Protocolo 2:** la herramienta DASP3 crea una matriz de puntuación específica de posición (PSSM) a partir del ASP y busca en la estructura primaria de las proteínas

(archivo FASTA) descargada de la base de datos del NCBI las secuencias de aminoácidos del homopiens (estructura primaria de las proteínas) para proteínas con firmas de sitio funcional relacionadas.

A continuación se presentan de manera más detallada el procedimiento para la elaboración del perfil alostérico en el DASP3.

3.1. Revisión de proteínas para elaboración de perfil alostérico

Para creación de perfil alostérico en el DASP3 se requieren 3 proteínas para posteriormente identificar los residuos clave de cada proteína, por lo que inicialmente se realizó una búsqueda en la literatura de las proteínas en Homo sapiens que interactúan con el ligando ATP en el sitio alostérico. Para ello se ha recurrido al repositorio UniprotKB [64], donde es posible identificar proteínas de Homo sapiens que interactúan con ATP, aún así, se revisan cada una de ellas en diversos artículos para realizar la selección de 3 proteínas, que según la literatura, efectivamente cuenten con sitio alostérico e interactúen con la molécula ATP. A continuación, se presenta descripción de las proteínas (todas con una función biológica diferente), elegidas para la elaboración de perfil alostérico en el DASP:

Cytosolic 5'-nucleotidase: Activada alostéricamente por varios compuestos, incluido el ATP. Las estructuras en esta proteína proporcionan datos estructurales para el diseño de fármacos dirigidos tanto al sitio activo como al sitio alostérico del gen cN-II. Los inhibidores de nucleotidasas podrían, por ejemplo, usarse como complemento a los análogos de nucleósidos en terapias contra el cáncer para reducir la desfosforilación no deseada de los análogos de nucleósidos activados (fosforilados) en tumores resistentes. También podrían usarse para minimizar la degradación de los fármacos antivirales e inmunosupresores basados en nucleósidos [65]. Posee un papel fundamental en el metabolismo celular de las purinas y actúa sobre análogos de nucleósidos antitumorales y antivirales, por lo que es de importancia terapéutica potencial [66].

ATP-dependent 6-phosphofructokinase, platelet type: Activada alostéricamente por ADP, AMP o fructosa 2,6-bisfosfato e inhibida alostéricamente por ATP o citrato. Actúa como el principal regulador de la glucólisis y cataliza la formación de fructosa-1,6-bisfosfato (F16BP) dependiente de Mg-ATP a partir de fructosa-6-fosfato (F6P) en células procariotas y eucariotas. La estructura de Pfk está involucrado en la regulación alostérica; puede servir como guía para desarrollar moduladores de la actividad de Pfk como posibles fármacos antitumorales, ya que la inhibición de esta reduce la formación de tumores [67, 68].

Endoplasmic reticulum chaperone BiP: La actividad de esta chaperona está regulada por el acoplamiento alostérico inducido por ATP de los dominios de unión a nucleótidos (NBD) y unión a sustrato (SBD) [69]. Es una chaperona del retículo endoplásmico que desempeña un papel clave en el plegamiento de proteínas y el control de calidad en la luz del retículo endoplásmico [70].

Seguidamente se efectuó una búsqueda en el repositorio *Protein Data Bank* (PDB) con el fin de descargar la estructura terciaria de las proteína identificadas. El PDB es una Base de datos de la estructura tridimensional de las proteínas y ácidos nucleicos. Se tienen datos, obtenidos regularmente por cristalografía de rayos X o Resonancia Magnética Nuclear (RMN), enviados por biólogos y bioquímicos de todo el mundo [71].

3.2. Residuos que interactúan con el ligando ATP

Para la creación del perfil alostérico el algoritmo DASP3 requiere 3 residuos clave de cada proteína que interactúan con el ligando.

En la base de datos de UniProtKB, se cuenta con la herramienta PDBe, la cual permite conocer las posiciones de los residuos que interactúan con el ligando ATP.

En la figura 3-2, se observan los aminoácidos de la proteína cytosolic purine 5'-nucleotidase, que interactúan con el ligando ATP.

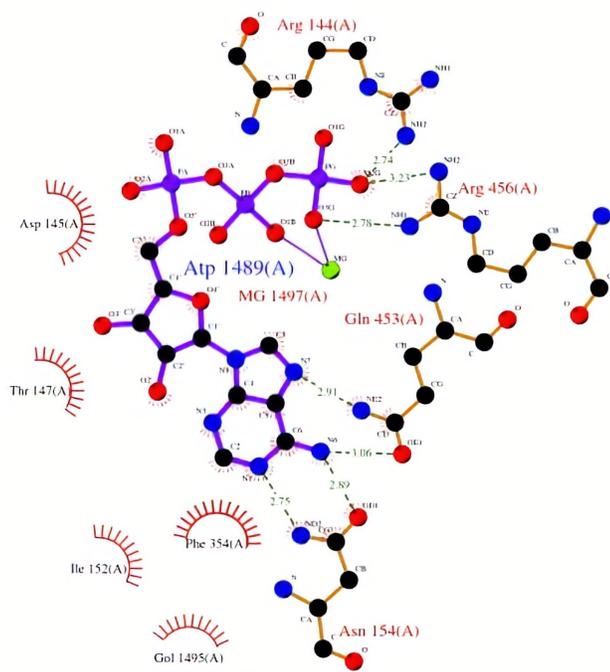


Figura 3-2.: Aminoácidos cytosolic purine 5'-nucleotidase con PDB 2XCW [64]

A partir de los aminoácidos identificados mediante PDBe, se eligieron en las secuencias los tres residuos que más se conservan (es decir, la inclusión de todos los aminoácidos, permitiendo la integración de cada uno de ellos, en cada intervalo, esto mediante diversas iteraciones combinando de manera alterna los residuos en el DASP3, hasta lograr crear ASP), en las proteínas elegidas, para llevarlos al DASP.

La tabla **3-1** muestra las posiciones y los residuos elegidos, que se llevaron al DASP3, de las proteínas encontradas en la revisión bibliográfica.

Tabla 3-1.: Proteínas utilizadas para realización de ASP mediante DASP3

Proteínas	PDB	Residuos relevantes
Cytosolic purine 5'-nucleotidase	2XCW	ARG 144, ASN 154, ARG 456
ATP-dependent 6-phosphofructokinase	4XYJ	GLY 127, SER 130, SER 173
Endoplasmic reticulum chaperone BiP	6ASY	THR 37, LYS 96, ARG 367

3.3. Creación del perfil alostérico e identificación de Proteínas mediante DASP3

Teniendo en cuenta el procedimiento realizado en [30], junto con los protocolos 1 y 2:

- Se instaló la herramienta de DASP3, se descargaron los archivos PDB (Estructura terciaria) de las proteínas elegidas para realización de perfil alostérico y la secuencia FASTA (Estructura primaria) de las proteínas de homo sapiens encontradas en NCBI.
- Se guardaron los archivos PDB, en una carpeta donde el nombre serán las dos letras del medio de la proteína. Por ejemplo, el archivo pdb1fa9.ent, correspondiente a la proteína 1FA9, debe estar almacenado en la carpeta “fa”.
- Se ingresaron al DASP3 los PDB y los residuos clave de las proteínas seleccionadas.
- Como resultado, el DASP3 indico que las posiciones de los residuos ingresados fueron aceptados y se pudo generar un ASP (Allosteric Site Profiler).
- Mediante el ASP, es decir, con la información sobre el sitio alostérico de las 3 estructuras tridimensionales conocidas, se utilizo el archivo FASTA descargado del NCBI, para identificar secuencias donde la estructura secundaria o terciaria o cuaternaria de las proteínas son desconocidas mediante el comando de la figura **3-3**, en el sistema operativo de linux.

```
$ ./dasp3.sh.-i Prueba/ATPinput8.txt -p/home/agonzalez73/mario.duarte@uan.edu.co  
/CarpetasTIG/ Software/dasp3.master/PDBFiles/ -r 10 -o Prueba/outputHomo.res -d  
/home/agonzalez73/mario.duarte@uan.edu.co/Software/dasp3.master/FastaFiles/Homos  
apiens.fasta -c 0.0001 -v > Prueba/Homo_terminal.txt
```

Figura 3-3.: Comando en linux para comparación de ASP con archivo FASTA

Donde:

- -i: Es el nombre del archivo de entrada.
- -r: El radio para su inclusión en la firma del sitio activo, este argumento se puede variar acorde al requerimiento del análisis.
- -c: Es el valor de corte para buscar en la base de datos de secuencias.
- -p: Es la ruta a la base de datos pdb, es decir, pdbDatabase.
- -o: Nombre del archivo para incluir el perfil del sitio alostérico.
- -d: La base de datos a utilizar para la búsqueda de secuencia.
- -v: Lugar donde se almacenara la salida de manera detallada desde el DASP3.

3.3.1. Selección de proteínas identificadas mediante el DASP3

El DASP3 es determinista, por lo que al realizar el perfil alostérico con las mismas proteínas y pasarlo por el mismo archivo FASTA, se obtuvieron los mismos resultados, esto se verifico comparando los diferentes archivos de texto obtenidos tras la realización del protocolo 2.

El DASP3 proporcionó los resultados en un archivo de texto, con un P-valor(entre más negativo, más similar es la proteína al perfil creado de acuerdo a los criterios de la matriz PSSM) y códigos por cada proteína, por lo que se buscó cada código en el NCBI, se organizó y se eliminaron las secuencias repetidas, es decir, la misma proteína, bajo diferentes códigos y/o cadenas.

A partir del Anexo A, donde se describen proteínas del NCBI organizadas por exponente y nombre encontradas mediante el DASP3, y a base de los siguientes criterios se eligieron 3 de las proteínas identificadas mediante perfil alostérico:

- En la literatura no se reportan sitio alostérico en interacción con el ATP.
- Elección de menor P-valor (exponente más negativo).
- Asesoramiento por parte del grupo de investigación instituto de errores innatos del metabolismo (IEIM) de la universidad Javeriana.

Se eligieron las estructuras primarias (FASTA) de las siguientes proteínas:

1. *Galactine-9*: Es un miembro de la familia galactina de unión a β -galactósido, induce la agregación de ciertos tipos de células. En [72], evaluaron la contribución de la *galactine-9* a la agregación de células de cáncer de mama, así como la relación entre la expresión de *galactine-9* en el tejido tumoral y la metástasis a distancia en pacientes con cáncer de mama. Deduciendo que *galactine-9* es un posible factor pronóstico con potencial antimetastásico en cáncer de mama. En [73], sugieren que *Galactine-9* induce apoptosis a través del Ca^{2+} -calpain-caspase-1, y juega un papel en la inmunomodulación de las respuestas inmunes mediadas por células T. En [74], sugieren que la misma proteína *Galactine-9*, pero en ratones, regula la inflamación inducida por LPS (lipopolisacáridos) y los protege de la reacción de Shwartzman al atraer PMN (polimorfonucleares) que produce PGE (prostaglandina E2). Se ha demostrado que esta proteína desempeña una actividad antiinflamatoria además de que puede actuar como un agente anticancerígeno induciendo apoptosis en especies de cáncer epitelial y tumores malignos hematológicos.
2. *BCL-2 like protein 12*: Es una proteína que en humanos está codificada por el gen BCL2L12. La expresión de esta proteína produce resistencia y/o detención de la apoptosis. Contribuye a las características biológicas tumorales clásicas de glioblastoma, como la resistencia a la apoptosis intensa y la necrosis florida, y puede proporcionar un objetivo para mejorar la respuesta terapéutica de este cáncer letal [75]. En [76], se prueba la hipótesis de que el factor de necrosis tumoral aumenta BCL2L12 para suprimir la expresión de interleucina 10 en las células B periféricas de pacientes con enfermedad inflamatoria intestinal (EII). Los datos sugieren que BCL2L12 puede ser un objetivo terapéutico para el tratamiento de la EII. En estudios recientes demostraron que *Bcl2-like protein 12 (Bcl2L12)*, se asoció con la patogénesis de T-H2-sesgado inflamación en el intestino, facilitando la polarización de T-H-2. Por lo tanto, se planteo la hipótesis de que Bcl2L12 también podría estar asociado con la patogénesis del asma alérgica [77]. La *BCL2 like protein 12* juega un papel en la supresión del desarrollo y la función de las células T reguladoras Foxp3 + (función de supresión inmune) en pacientes con artritis rematoide. Por tanto, se sugiere que la inhibición de Bcl2L12 puede tener potencial terapéutico en el tratamiento del artritis rematoide [78].
3. *Fumarate hidratase, mitochondrial*: Esta localizada en la matriz mitocondrial y existen en formas mitocondriales y citosólicas [79]. En las mitocondrias, la *Fumarate hidratase* es esencial para la respiración celular como parte del ciclo del ácido tricarboxílico (TCA), mientras en el citosol, en virtud de su actividad enzimática, participa en la respuesta al daño del ADN [80]. Están presentes en una amplia variedad de organismos, incluidas bacterias, levaduras, mohos, plantas, invertebrados y mamíferos [81]. Las mutaciones en la *Fumarate hidratase*, predisponen a los individuos a la leiomiomatosis hereditaria y al cáncer de células renales [82]. Se distribuyen sobre todos los organismos y juegan papeles importantes en la producción de energía, la reparación del ADN y como supresores de tumores [79, 82, 80]. Además de ser objetivos importantes tanto

en el estudio de los trastornos metabólicos humanos como como posibles objetivos terapéuticos en enfermedades tropicales desatendidas y tuberculosis [83].

4. Verificación de proteínas identificadas en NCBI

En el presente capítulo se describe la metodología para verificar si efectivamente las proteínas elegidas del NCBI, mediante creación de perfil alostérico en el DASP3, tienen acoplamiento en el sitio alostérico con el ligando ATP.

Para la realización de las simulaciones, inicialmente se realizó búsqueda en la base de datos de protein Data Bank, para verificar si las proteínas elegidas contaban con un PDB, además se tuvo en cuenta que la estructura estuviera completa y que la resolución con la que se obtuvo la estructura tridimensional de la proteína fuera buena (resolución menor a 3 Ångstrom (Å)). Con respecto a la BCL2-Like-protein 12, únicamente se encontró la estructura primaria de la proteína, por lo que fue necesario estimar la estructura terciaria por medio de herramientas computacionales. Para estimar la estructura geométrica, se emplea el software I-TASSER, el cual predice la estructura terciaria y la función de la proteína.

Al tener elegidas las proteínas con el PDB correspondiente, se elaboró un procedimiento para realización del *Docking* (tomando como referencia [55, 84]), este se presenta de manera general en el diagrama de bloques de la figura 4-1.

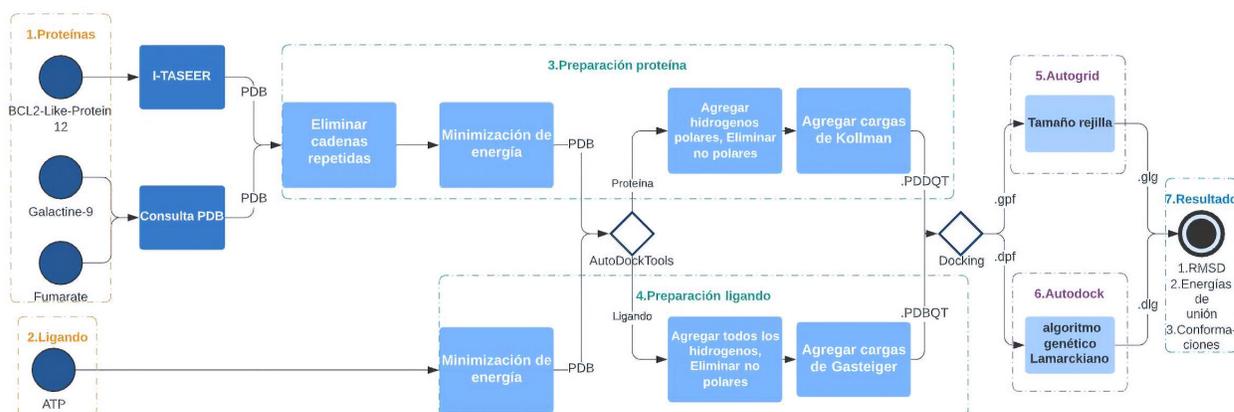


Figura 4-1.: Diagrama de bloques con procedimiento simplificado para realización de *Docking* en AutoDockTools, imagen de Autor

A continuación se detalla el procedimiento descrito en la figura 4-1.

4.1. Preparación de proteínas y ligandos

Teniendo preparado el entorno para correr Autodock4 (procedimiento presentado en el segundo capítulo del presente trabajo) y los PDB de las proteínas se realiza la preparación de las mismas: *Endoplasmic reticulum chaperone BiP*, *Cytosolic purine 5'-nucleotidase*, *ATP-dependent 6-phosphofructokinase*, *BCL2-Like-protein 12*, *Galactine-9*, *Fumarate hydratase*, *mitochondrial* y del ligando ATP.

4.1.1. Preparación proteínas

Mediante USCF Chimera se eliminan las cadenas repetidas, las moléculas de agua y los compuestos que no conforman la molécula, dejando únicamente la proteína. Se guardaron como PDB, y se dispusieron en Swiss-PdbViewer para una minimización de energía a los PDB de las moléculas, esto con el fin de que la distancia entre un átomo y el otro sea el óptimo, consiguiendo evitar choques y deformaciones en cada estructura.

En la tabla **4-1**, se presentan los nombres de las proteínas tomados de la base de datos Protein Data Bank (PDB) preparadas mediante USCF Chimera para la realización de *Docking*.

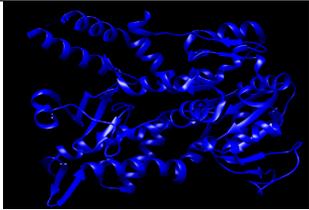
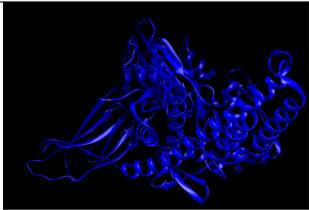
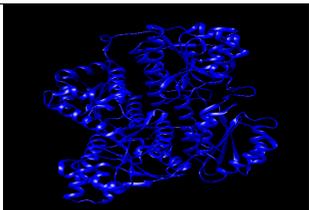
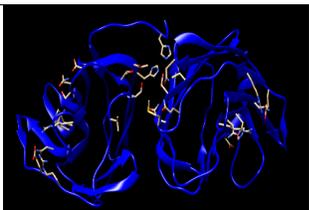
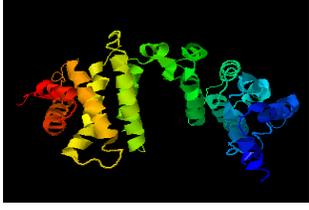
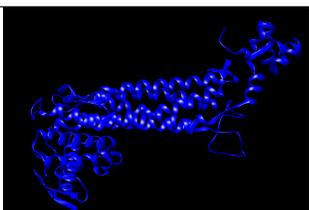
Se realizó reconstrucción geométrica a la proteína *BCL2-like-protein 12*, mediante I-TASSER, donde se obtuvo un C-score: -2.26.

Cabe aclarar que el puntaje C está establecido en I-TASSER entre -5 y 2, donde un puntaje C de mayor valor significa un modelo con una alta confianza y viceversa.

El método para obtener las estructuras fue por difracción de rayos x, con una resolución menor a 3 Ångstrom (Å).

- Mediante la entrada *read molecule* del menú de ADT es posible cargar el archivo PDB de la proteína anteriormente minimizada. En la figura **4-2**, se observa la proteína en el entorno de AutoDockTools. Es posible observar la proteína representada en líneas, tubos, esferas, cintas y en superficie. Se utilizó en cintas como se observa en la figura **4-2**, esto por comodidad de usuario.

Tabla 4-1.: Estructuras de proteínas listas para realización de *Docking*

Código PDB	Nombre proteína	Estructura
2XCW	<i>Cytosolic purine 5'-nucleotidase</i>	
6ASY	<i>Endoplasmatic reticulum Chaperone bip</i>	
4XYJ	<i>ATP-depent-6-phospofructokinase</i>	
3WV6	<i>Galactine-9</i>	
Reconstrucción geométrica	<i>BCL2-like-protein 12</i>	
5D6B	<i>Fumarate hydratase, mitochondrial</i>	

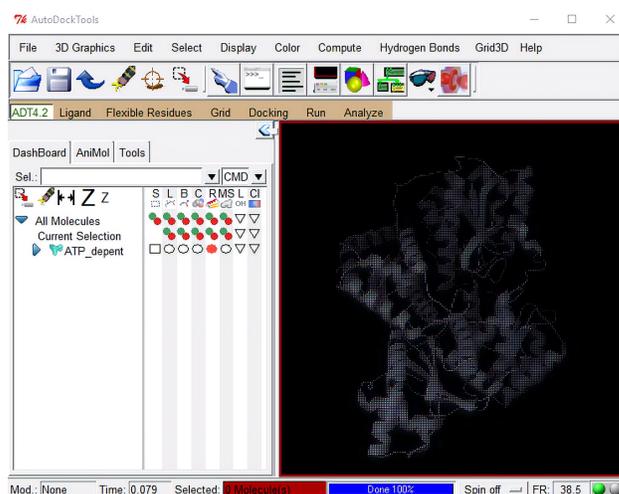


Figura 4-2.: Proteína en AutoDockTools, imagen de Autor

- Después de tener la proteína predispuesta en el entorno de desarrollo ADT, se procede a agregar los hidrógenos polares y se eliminan los no polares, para disminuir el gasto computacional y conservar las cargas calculadas de la proteína. Cabe aclarar que los hidrógenos polares son hidrógenos que están unidos a átomos electronegativos como oxígeno y nitrógeno mientras los no polares son hidrógenos unidos a los átomos de carbono.
- Para equilibrar el modelo enzimático se agregan las cargas de Kollman¹ y se procede a guardar la molécula, permitiendo el ordenamiento de los nodos, permitiendo que el algoritmo genere un arreglo de los elementos con la extensión .PDBQT.

4.1.2. Preparación ligando

Para la preparación del ligando se descargó la estructura 2D del banco de datos de proteínas *RCSB*.

- En el software de Avogadro se carga el PDB del ligando ATP (Adenosín Trifosfato) en dos dimensiones como se observa en la figura 4-3.
- Se realizó una minimización de energía (figura 4-4), hasta que las fuerzas sobre cada átomo fueron cero (0).

¹Cargas derivadas del potencial molecular electrostático que se ajustan en un conjunto de puntos alrededor de la molécula.

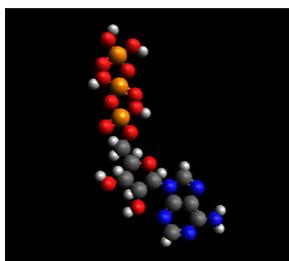


Figura 4-3.: ATP 2D en Avogadro, Imagen de Autor

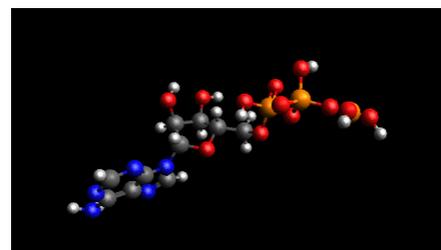


Figura 4-4.: ATP minimizada en Avogadro, Imagen de Autor

- En el entorno de desarrollo de AutoDockTools se carga el PDB correspondiente al ATP y se agregan todos los hidrógenos para agregar correctamente las cargas.
- Se computan las cargas de Gasteinger², pues son las más favorables a utilizar con el ligando, para equilibrar el modelo enzimático.
- Se eliminan los hidrógenos no polares y se procede a guardar el ligando con extensión .PDBQT

4.2. Elección rejilla para elaboración de *Docking*

Para definir el espacio para la realización del *Docking* fue necesario la generación de mapas con la información correspondiente a esta delimitación para así ser ingresado al ejecutable de *AutoGrid*.

- En el menú de ADT se selecciona *Grid*, para la elección de la macromolécula, y se seleccionan como tipo de búsqueda el ligando ATP seleccionando *grid*, *Set Map Types* 4-5.
- Se buscaron los residuos (aminoácidos) donde se presenta interacción alostérica de las proteínas predicho por el DASP. Para seleccionar el sitio donde se desea elaborar el *Docking*, se cuenta con un recuadro con el nombre de la proteína, la cadena y los aminoácidos que la componen 4-5.

²Permiten el equilibrio de electronegatividad para el sistema molecular, en caso de que no tener cargas parciales disponibles.

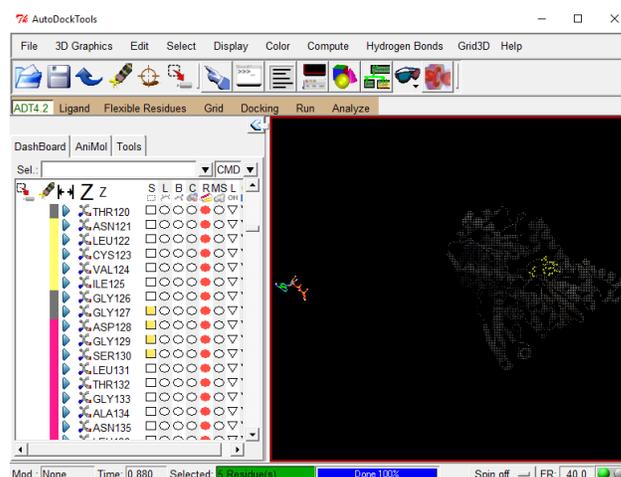


Figura 4-5.: Selección de puntos de la proteína en AutoDockTools, imagen de Autor

- Se definió el tamaño del *gridbox* para la realización del *Docking*, encerrando el sitio alostérico. Se guardo de la siguiente manera: *Grid, grid box, file, close saving* Dimensiones aproximadas de: 40 x 40 x 40 Ångstrom (Å) alrededor del sitio alostérico.

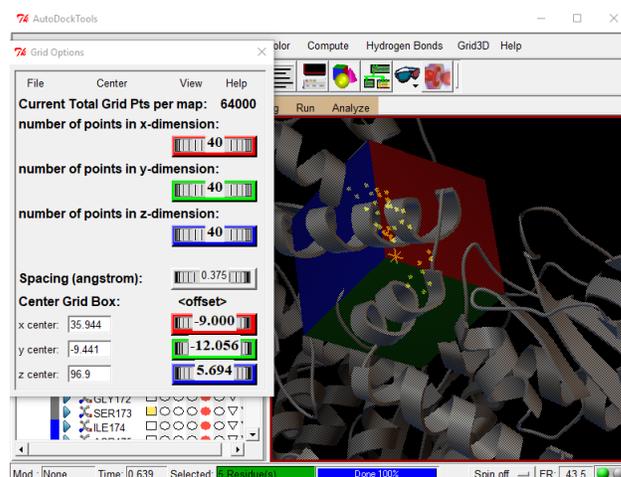


Figura 4-6.: Tamaño rejilla en AutoDockTools, imagen de Autor

- En *Grid output*, se procede a guardar la delimitación propuesta con la extensión *.gpf*, para ser ingresado al correr el ejecutable *Autogrid*.

4.3. Docking

El estudio de *Docking* Molecular utilizó el algoritmo genético Lamarckiano (LGA), este permite descartar las conformaciones menos favorables y adecuar las que sí lo son para la

generación del cálculo de los resultados. El protocolo LGA aplica un tamaño de población de 150 conformeros, mientras que un máximo de 2.500.000 evaluaciones de energía se utilizaron para las 100. Las mejores conformaciones fueron escogidas de las soluciones de energía más bajas, es decir las energías de unión más negativas.

- Mediante el panel de control de ADT se selecciona *Docking, macromolecule, set rigid filename*, para elegir la proteína, es decir, la molécula que va permanecer rígida (no se van a mover los enlaces de la proteína).
- Para elegir el ligando se dirige al panel de control de ADT y selecciona: *Docking, ligand, choose* y elige el ligando.
- Para elegir los parámetros de búsqueda se utilizo un algoritmo genético lamarckiano (LGA) en: *Docking, search Parameters, genetic algorit, acceppt*. Aquí es posible elegir la exhaustividad con la que se va a realizar la búsqueda del *Docking*, por tanto se eligió un algoritmo largo para obtener mejores resultados, al elegir un algoritmo largo en algunos *Docking* se tardo hasta dos horas, un algoritmo de búsqueda corto o medio se tardara un menor tiempo, es decir de entre 5 y 15 minutos.
- En *Docking, Docking parameters*, es posible visualizar los parámetros ingresados para el *Docking*, estos se dejaron de manera preseleccionada por el programa.
- Se procede a guardar lo realizado mediante: *Docking, output, Lamarckian, save* con la extensión *.dpf*
- Para que el programa eliga el lugar donde se realizará el *Docking* se ingresa: *Run, autogrid* y se ingresan los siguientes parámetros (archivo con extensión *.gpf*, generado al realizar la delimitación de rejilla) **4-7**:

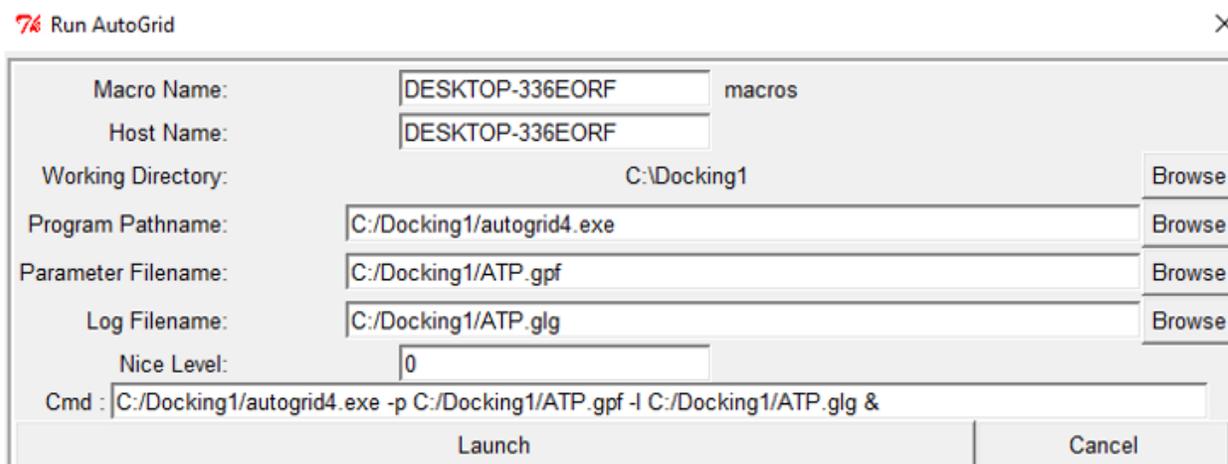


Figura 4-7.: Parámetros para ejecutar Autogrid en AutoDockTools, imagen de Autor

- Click en *launch* y esperar, esto generara un archivo con extensión .glg
- Al tener preseleccionada el área donde se realizara el *Docking* se procede a correr autodock seleccionando: *Run, Autodock 4-8*.

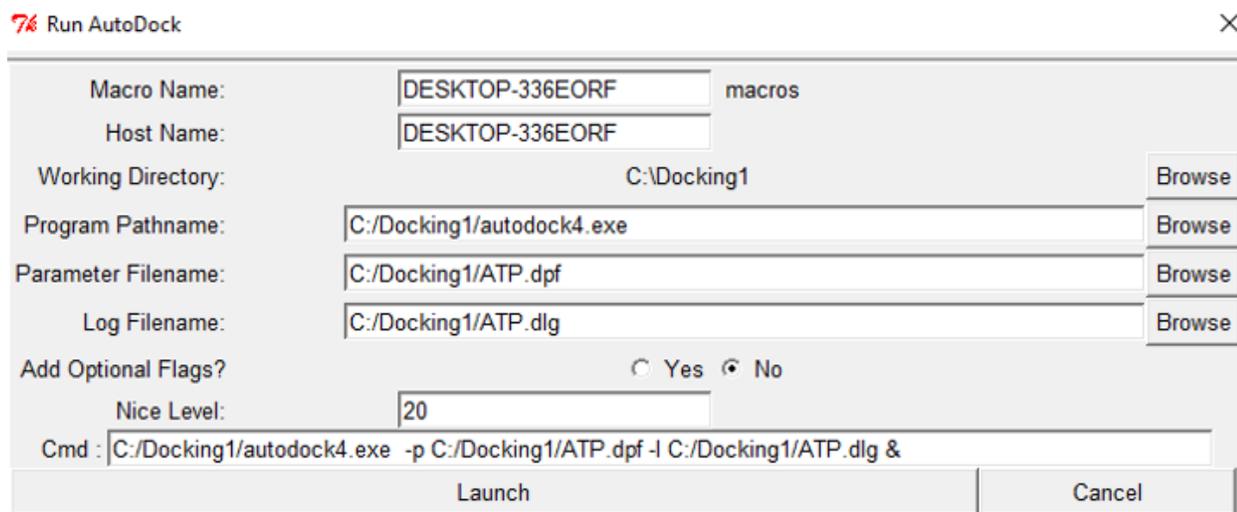


Figura 4-8.: Parámetros para ejecutar Autodock4 en AutoDockTools, imagen de Autor

- Click en *launch* y esperar, esto generara un archivo con extensión .dlg, donde será posible observar los resultados del *Docking* realizado.

5. Resultados y discusión de proteínas identificadas mediante DASP3

Se realizó el *Docking* tanto a las proteínas empleadas para realizar el perfil alostérico mediante el DASP3 (*Endoplasmic reticulum chaperone BiP*, *Cytosolic purine 5'-nucleotidase y ATP- depend 6- phosphofructokinase*) como a las identificadas del NCBI por medio del DASP3 (*BCL2-Like-protein 12*, *Galactine-9 y Fumarate hydratase, mitochondrial*). Lo anterior se realizó para poder comparar las energías de unión (Kcal/mol) y valores RMSD (Å) en cada caso.

Se presentan acoplamientos moleculares y un análisis de interacciones de los complejos proteína-ligando, donde es posible observar los aminoácidos en el bolsillo donde interactúan en el sitio de unión. También se presentan los valores RMSD correspondientes a cada acoplamiento proteína-ligando.

5.1. Acoplamientos moleculares de los complejos proteína-ligando

El acoplamiento molecular permite predecir la energía de afinidad en el sitio de unión de una proteína con el ligando, donde podrá actuar como activador y/o inhibidor.

A continuación se presentan algunos acoplamientos en forma sólida de las proteínas en ADT. En la figura 5-1, se observan las proteínas con algunos de sus respectivos bolsillos, es decir, el lugar de unión de la proteína con el ligando ATP, con las que se realizó entrenamiento al DASP3.

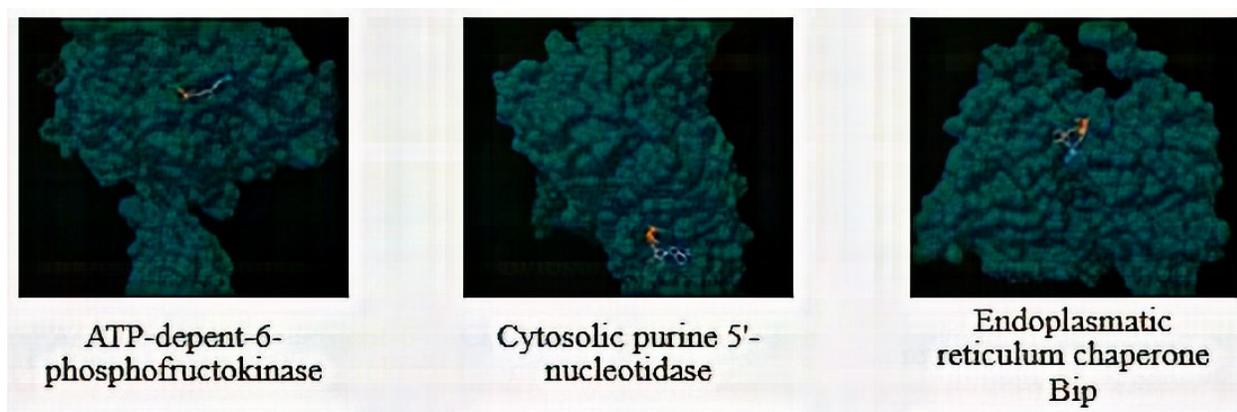


Figura 5-1.: Bolsillo proteínas entrenamiento DASP3 obtenidas mediante ADT, imagen de autor

De igual manera en la figura 5-2, se observan algunos bolsillos de las proteínas que se eligieron del NCBI mediante la realización de ASP en el DASP3.

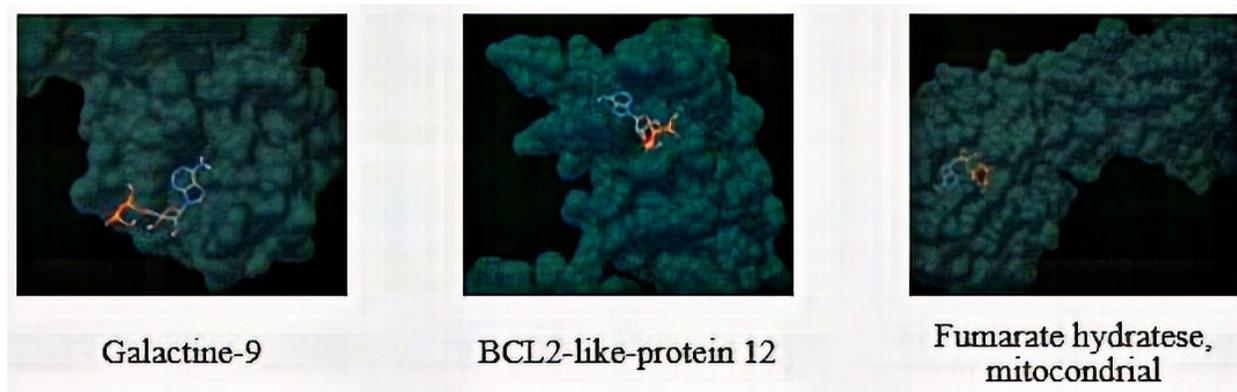


Figura 5-2.: Bolsillo proteínas elegidas del NCBI a partir de DASP3 obtenidas mediante ADT, imagen de autor

En las figuras 5-1, 5-2, se observan las proteínas a las que se le realizó *Docking* en representación sólida con el ligando ATP. En cada proteína se obtuvieron 10 conformaciones y en cada conformación se distingue un bolsillo donde el ligando se acomoda dentro de la proteína.

5.2. Análisis de interacciones

Mediante un análisis de interacciones (función de *Docking* en AutoDockTools) se pueden observar los aminoácidos en el sitio alostérico con el ligando, estos se describen en los *Docking* correspondiente a cada proteína mediante tablas, descritas en las dos siguientes subsecciones, donde se presentan, el sitio de unión con su correspondiente energía (cuánto más negativa

es la energía de unión mayor es la afinidad de unión de la proteína con el ligando).

En las figuras **5-3,5-4,5-5**, se presentan algunas de las interacciones obtenidas tras el *Docking* molecular.

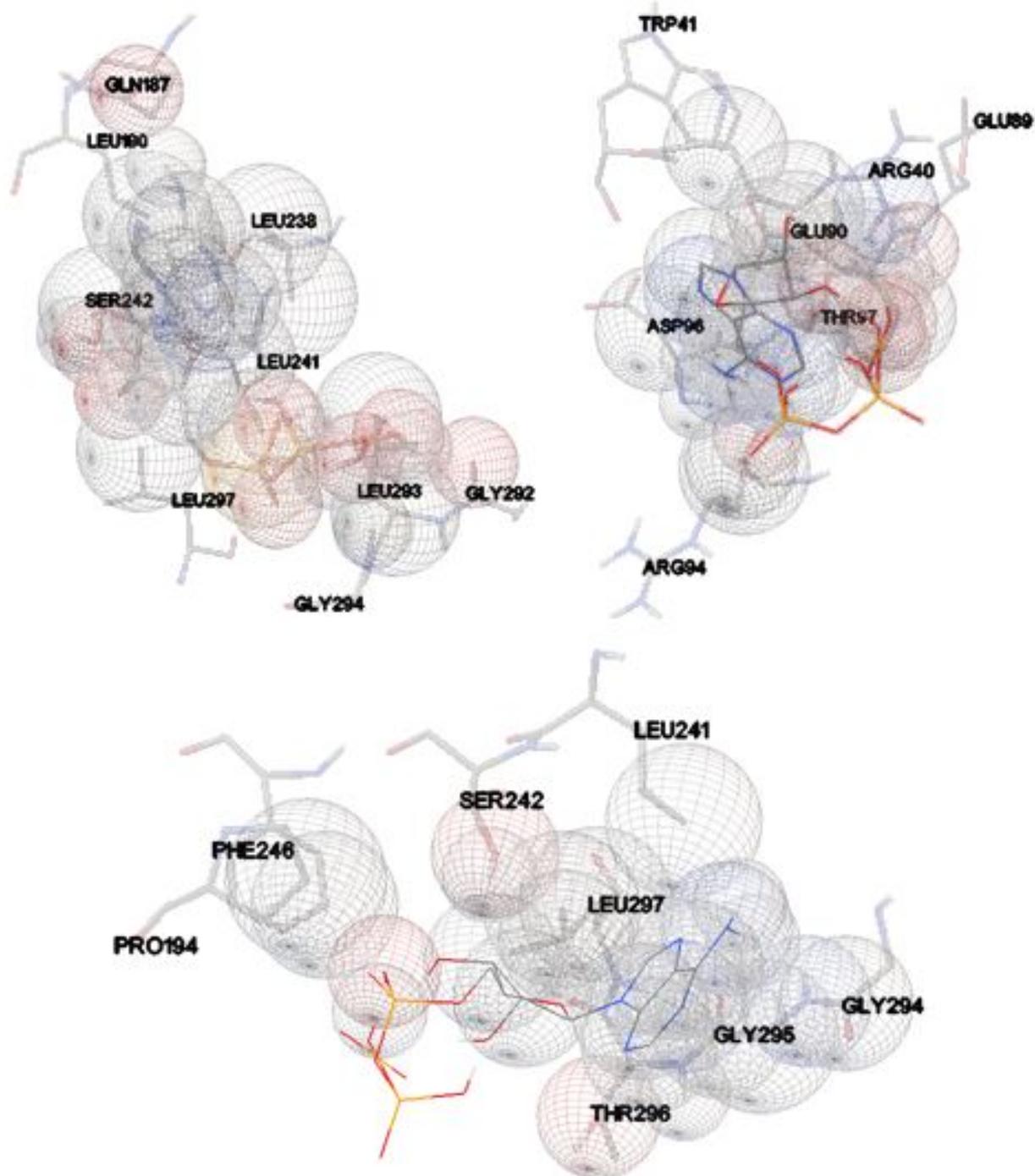


Figura 5-3.: Interacciones BCL2-Like-Protein 12 obtenidas mediante AutoDockTools, imagen de Autor

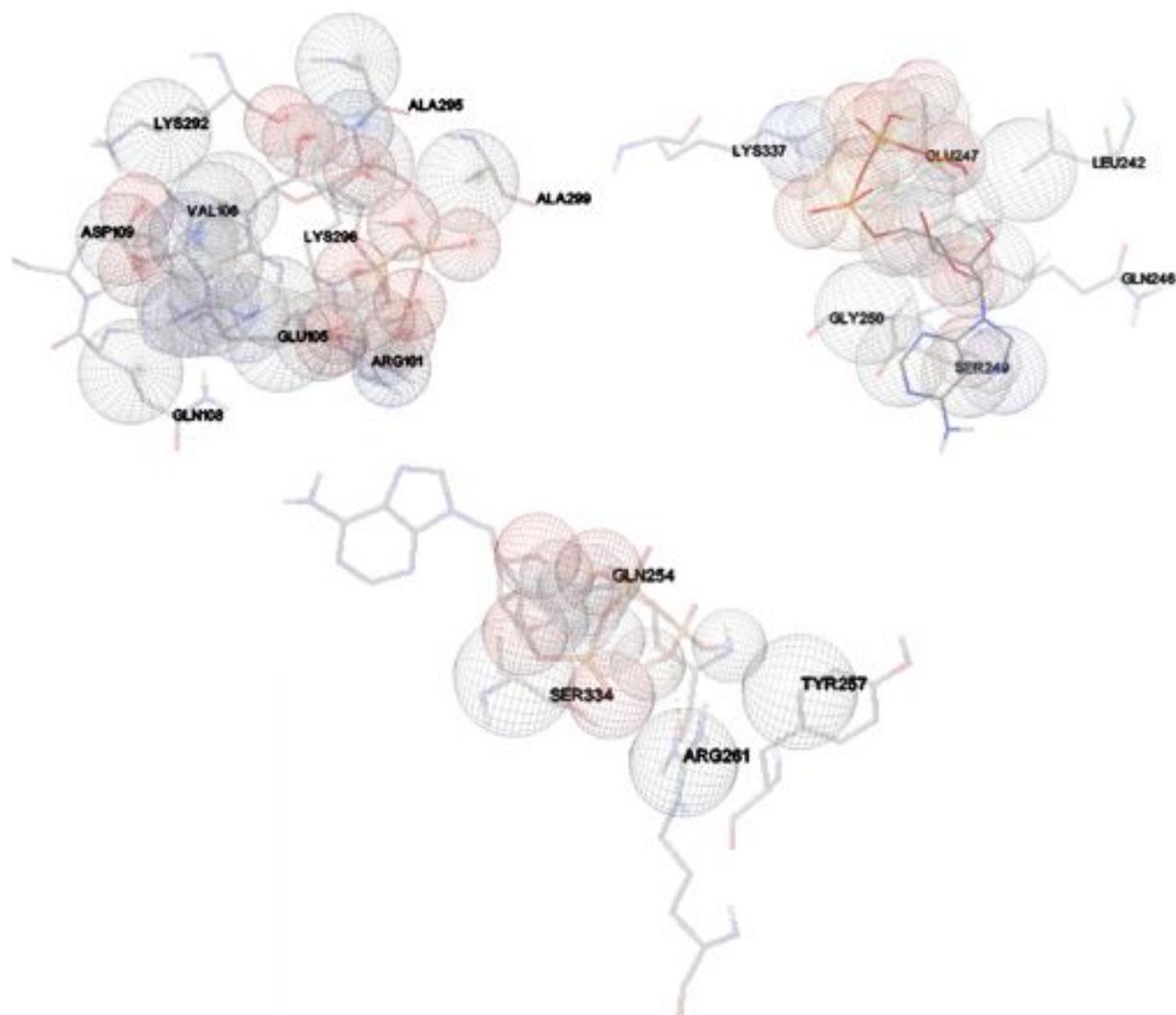


Figura 5-4.: Interacciones Fumarate obtenidas mediante AutoDockTools, imagen de Autor

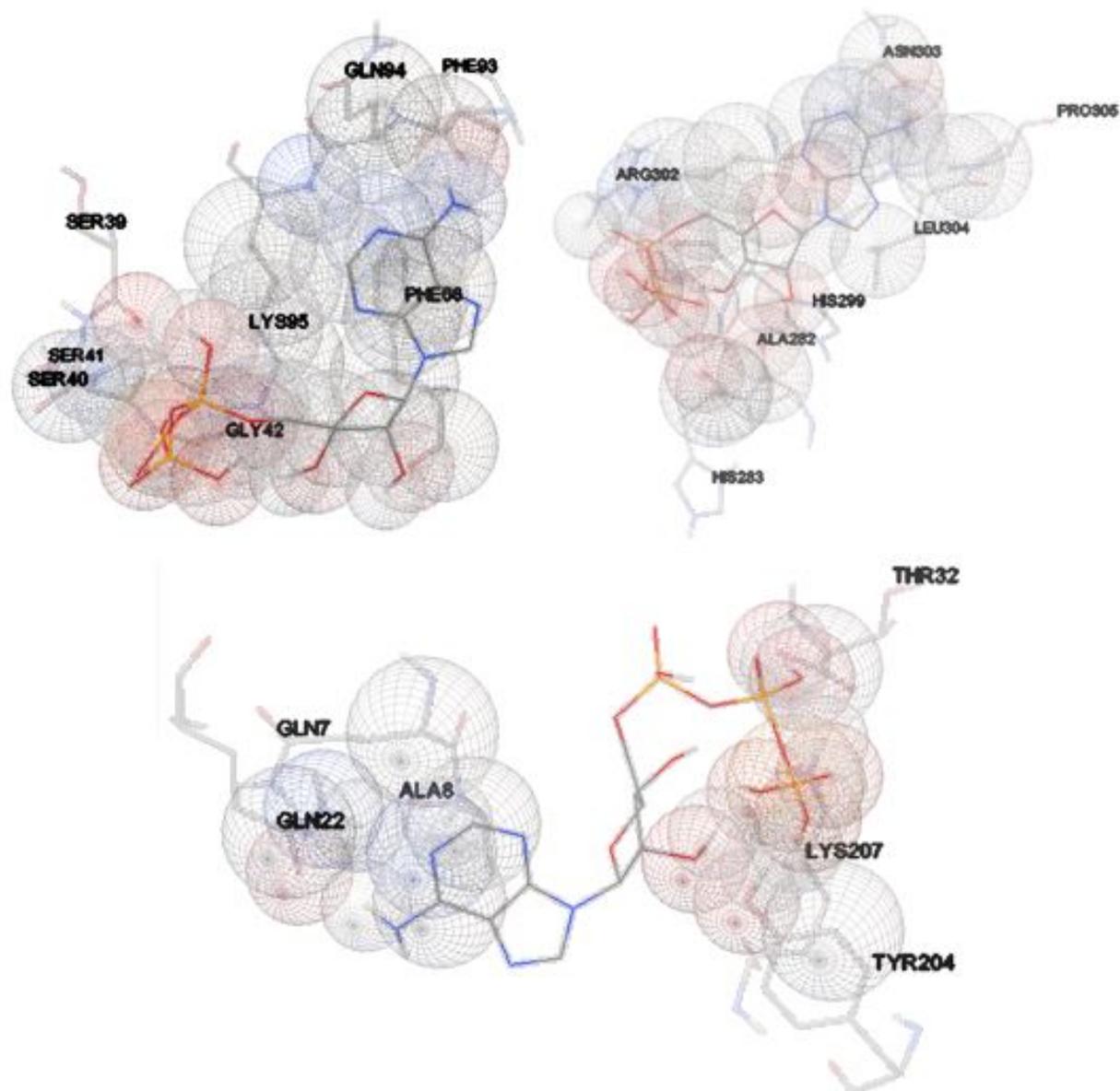


Figura 5-5.: Interacciones Galactine-9 obtenidas mediante AutoDockTools, imagen de Autor

5.3. Energías de unión y valores RMSD de proteínas elegidas para creación de perfil alostérico

Para la creación de perfil alostérico en el DASP3 se eligieron las siguientes proteínas: *Endoplasmic reticulum chaperone BiP*, tabla 5-1, *Cytosolic purine 5'-nucleotidase*, tabla 5-2 y *ATP- depend 6- phosphofructokinase* tabla 5-3, pues presentan un sitio de unión conocido

con el ligando ATP.

Considerando que al realizar el *Docking* se obtienen las energías de unión y valores RMSD, se realizó el *Docking* a las proteínas empleadas para el entrenamiento del DASP3, de tal manera que las energías y valores RMSD computados se desempeñan como un patrón de referencia para la existencia de un sitio alostérico que interactúa con el ATP.

5.3.1. Endoplasmic reticulum chaperone BiP

Se realizó *Docking* a la proteína *Endoplasmic reticulum chaperone BiP*, con PDB 6ASY, realizando rejilla mediante *Autogrid*, en los aminoácidos seleccionados para ser ingresados al DASP3.

Tabla 5-1.: Conformaciones con residuos conocidos en interacción con ATP, *Endoplasmic reticulum chaperone BiP*

Proteína	Energía de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
Endoplasmic reticulum chaperone BiP	-5.77	0.0	THR 38, TYR 39, ILE 61, GLY 227, GLU 293, LYS 296, ARG 297, GLY 364, SER 365, ARG 367, ASP 391
	-5.64	0.0	GLY 36, THR 37, THR 38, TYR 39, LYS 96, GLU 201, GLY 227, ARG 297, SER 300, GLY 364, ARG 367, ASP 391
	-4.76	0.0	ASP 34, GLY 36, THR 38, TYR 39, GLY 227, LYS 296, ARG 297, SER 300, GLY 364, ARG 367, ASP 391
	-4.39	1.93	ASP 34, THR 38, TYR 39, ILE 61, GLY 227, LYS 296, ARG 297, GLY 364, SER 365, ARG 367, ILE 368, ASP 391
	-4.25	0.0	ASP 34, THR 38, TYR 39, GLY 227, ASP 224, GLY 225, GLU 256, LYS 296, ARG 297, GLY 363, GLY 364, SER 365, ARG 367, PRO 390, ASP 391, VAL 394
	-4.04	1.58	ASP 34, THR 38, TYR 39, ILE 61, GLY 227, ARG 297, SER 300, GLY 364, ARG 367, ILE 368, ASP 391
	-3.52	0.0	ASP 34, GLY 36, THR 37, THR 38, TRY 39, LYS 96, GLU 201, GLY 224, GLY 226, GLY 227, GLU 293, ARG 297, GLY 363, GLY 364, ASP 391, VAL 394
	-3.5	1.85	ASP 34, GLY 36, THR 37, THR 38, TRY 39, ILE 61, GLY 227, GLY 255, GLU 256, LYS 296, ARG 297, GLY 363, GLY 364, SER 365, ARG 367, ASP 391
	-3.34	1.47	THR 38, TYR 39, GLY 227, GLU 293, LYS 296, ARG 297, GLY 364, SER 365, ARG 367, ASP 391
PDB 6ASY	-1.92	0.0	LEV 35, GLY 36, THR 37, THR 38, TYR 39, LYS 96, GLU 201, GLY 226, GLY 227, GLY 228, GLY 364, ARG 367, ASP 391

En la tabla 5-1, se presentan energías entre -1.92 y -5.77. Es de notar que, para la conformación del ligando con mayor energía de unión (-5.64), los residuos con los que se obtuvo el perfil alostérico (posiciones: 37,96,367) también fueron estimados, al realizar el *Docking*, como los residuos que interactúan con el ligando.

5.3.2. Cytosolic purine 5'-nucleotidase

Se realizó *Docking* a la proteína *Cytosolic purine 5'-nucleotidase*, con PDB 2XCW, realizando rejilla mediante Autogrid, en los aminoácidos seleccionados para ser ingresados al DASP3.

Tabla 5-2.: Conformaciones con residuos conocidos en interacción con ATP, *Cytosolic purine 5'-nucleotidase*

Proteína	Energías de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
Cytosolic purine 5'-nucleotidase	-5.13	0.0	ARG 144, ILE 152, ASN 154, PHE 354, LEU 358, LYS 362, GLN 453, ARG 456, TYR 457
	-4.48	1.61	ARG 144, ASN 154, PHE 354, LEU 358, LYS 359, LYS 362, GLN 453
	-4.5	1.67	ARG 144, ILE 152, ASN 154, PHE 354, LEU 358, LYS 362, GLN 453, TYR 457
	-4.14	0.0	ARG 144, ILE 152, PHE 354, LEU 358, LYS 362, GLN 453, ARG 456, TYR 457
	-3.6	0.0	ARG 144, ASN 154, ILE 152, LEU 358, LYS 359, LYS 362, GLN 453
	-3.03	1.76	ARG 144, THR 147, ILE 152, ASN 154, PHE 354, LEU 358, LYS 362, GLN 453
	-2.38	1.61	ASP 145, ARG 144, ILE 152, ASN 154, PHE 354, LEU 358, LYS 359, LYS 362, GLN 453
PDB 2XCW	-1.5	0.0	ARG 144, ASP 145, ILE 162, PHE 354, LEU 358, LYS 362, GLN 453, TYR 457
	-1.25	0.0	ARG 144, ILE 152, PHE 354, LEU 358, LYS 362, GLN 453, ARG 456, TYR 457
	-0.93	0.0	ARG 144, ILE 152, PHE 354, LEU 358, LYS 362, GLN 453, ARG 456, TYR 457

En la tabla 5-2, se presentan energías entre -0.93 y -5.13. Es de notar que, para la conformación del ligando con mayor energía de unión (-5.13), los residuos con los que se obtuvo el perfil alostérico (posiciones: 144,154,456) también fueron estimados, al realizar el *Docking*, como los residuos que interactúan con el ligando.

5.3.3. ATP-depent 6-phosphofructokinase

Se realizo *Docking* a la proteína *ATP-depent 6-phosphofructokinase*, con PDB 4XYJ, realizando rejilla mediante Autogrid, en los aminoácidos seleccionados para ser ingresados al DASP3.

5.4 Energías de unión y RMSD para proteínas identificadas por el DASP3

Tabla 5-3.: Conformaciones con residuos conocidos en interacción con ATP, *ATP-depend 6-phosphofructokinase*

Proteína	Energía de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
ATP-depend-6-phosphofructokinase	-3.46	0.0	TYR 64, CYS 98, PHE 101, ARG 102, GLY 129, SER 130, THR 132, GLY 133
	-2.86	0.0	SER 32, GLY 33, GLY 34, ARG 102, GLY 127, ASP 128, GLY 129, SER 130, THR 132, SER 173, ASP 175, ASP 177, ARG 219
	-2.28	0.0	GLY 34, ARG 97, ARG 102, ASP 126, GLY 127, GLY 129, SER 130, THR 132, SER 173, ASP 177, ASP 175, ARG 219, ARG 310
PDB 4XYJ	-1.69	1.43	GLY 33, ARG 97, ARG 102, GLY 127, ASP 128, GLY 129, SER 130, THR 132, SER 173, ASP 175, ASP 177, ARG 219

En la tabla 5-3, se presentan energías entre -1.69 y -3.46. Es de notar que, para la conformación del ligando con la energía de unión (-2.28), los residuos con los que se obtuvo el perfil alostérico (posiciones: 127,130,173) también fueron estimados, al realizar el *Docking*, como los residuos que interactúan con el ligando.

5.4. Energías de unión y RMSD para proteínas identificadas por el DASP3

Se presentan energías de unión y valores RMSD de las proteínas: *BCL2-Like-protein 12*, *Galactine-9* y *Fumarate hydratase, mitochondrial*, identificadas del NCBI por medio de creación de ASP en el DASP3.

5.4.1. Docking BCL2- like-protein 12

Se realizó un *Docking* con la rejilla cubriendo toda la proteína (*Docking* “ciego”) *BCL2- like protein 12*, tabla 5-4, para conocer los lugares de la proteína donde, de acuerdo al *Docking*, se presenta mayor energía de unión con el ligando ATP.

Se observa que la conformación con mayor energía de unión (mayor afinidad) se presenta en los residuos con posiciones: 51, 53, 54, 98, 99, y 102. Lo anterior evidencia la correspondencia entre el sitio estimado por el *Docking* y el sitio 1 estimado por el DASP3.

Los sitios de unión predichos por el DASP3 son:

1. **Sitio 1:** ELGLREDT correspondiente a las posiciones 90 a la 97 de la secuencia FASTA.
2. **Sitio 2:** KLVRLSSDSF correspondiente a las posiciones 237 a 247 de la secuencia FASTA.

Tabla 5-4.: Conformaciones de *Docking* cubriendo toda la *BCL2-Protein 12* en interacción con ATP

Proteína	Energías de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
BCL2-like protein 12	-2.23	0.0	ARG 51, GLY 53, LYS 54, LEU 98, ARG 99, ALA 102
	-1.50	0.0	ARG 261, PRO 262, ARG 264, GLY 268
	-1.47	0.0	ARG 51, LEU 98, ARG 99, LEU 134
	-1.44	0.0	ARG 51, LEU 98, ARG 99, ALA 102, CYS 137, LEU 138, SER 141
	-1.31	0.0	ARG 3, ARG 50, ARG 52
	-1.09	0.0	ARG 120, LEU 142, GLY 164, PRO 167, GLY 168
Docking ciego	-0.93	0.0	ARG 51, LYS 54, LEU 98, ARG 99
	-0.38	0.0	ARG 73, SER 113, PRO 152, ARG 153
	0.50	0.0	GLU 222, ASN 225, GLN 226, ALA 283, LEU 286, SER 287, ARG 288
	0.89	0.0	PRO 194, LUE 241, SER 242, PHE 246, GLY 294, GLY 295, THR 296, LEU 297

En consecuencia, se realizó un *Docking* dirigido únicamente a la región del sitio 1, con los residuos estimados por el DASP3. En la tabla 5-5, se muestran los sitios de unión con sus respectivas energías y valores RMSD, para las posibles conformaciones del ligando. Algunas conformaciones presentan mayor energía de unión al compararlas con las energías de unión obtenidas para las proteínas de entrenamiento del DASP3 (tablas 5-1, 5-2 y 5-3). Adicionalmente, el valor RMSD es cero para todas las conformaciones, lo cual indica que la estructura geométrica del ligando es muy estable dentro de todas las posibles interacciones estimadas. Por lo tanto, se puede asegurar que el sitio 1, identificado por el DASP3, muy posiblemente es un sitio alostérico para el ligando ATP con la proteína *BCL2-Like-protein 12*.

Seguidamente se realizó *Docking* dirigido únicamente a la región del sitio 2 con los residuos estimados por el DASP3. En la tabla 5-6, se muestran los sitios de unión con sus respectivas energías y valores RMSD, para las posibles conformaciones del ligando.

La tabla 5-6, presentó una menor energía de unión tanto en el *Docking* completo de la proteína como en la delimitación propuesta, con una máxima energía de unión de -2.21.

Teniendo en cuenta las energías de unión de las proteínas de entrenamiento (tablas 5-1, 5-2 y 5-3), posiblemente el sitio 2 no sea un sitio de unión para el ligando ATP en la proteína *BCL2-Like-protein 12*.

5.4 Energías de unión y RMSD para proteínas identificadas por el DASP3

Tabla 5-5.: Conformaciones del ATP con la proteína *BCL2-Protein 12* en el sitio 1 estimado por el DASP3

Proteína	Energías de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
BCL2-like protein 12	-4.13	0.0	ALA 42, GLY 43, GLU 90, LEU 91, GLY 92, LEU 93, ARG 94, ASP 96, THR 97
	-3.75	0.0	ARG 40, TRP 41, GLY 43, GLU 89, GLU 90, LEU 91, GLY 92, LEU 93, ARG 94
	-3.39	0.0	LEU 39, ARG 40, GLU 89, GLU 90, LEU 91, GLY 92, ARG 94
	-2.14	0.0	ARG 40, TRP 41, GLU 89, GLU 90, LEU 91, ARG 94
	-2.09	0.0	ARG 40, TRP 41, GLU 89, GLU 90, ARG 94, ASP 96, THR 97
	-2.04	0.0	GLU 89, GLU 90, LEU 91, GLY 92, ARG 94
	-1.79	0.0	ARG 40, GLU 89, GLU 90, ARG 94, GLU 95, ASP 96
	-1.33	0.0	GLU 89, GLU 90, LEU 91, GLY 92, LEU 93, ARG 94
	-1.08	0.0	TRP 41, ALA 42, GLY 43, GLU 90, LEU 91, GLY 92, ARG 94
Docking 90-97	-0.82	0.0	ARG 40, ALA 42, GLU 89, GLU 90, ARG 94, GLU 95

Tabla 5-6.: Conformaciones del ATP con la proteína *BCL2-Protein 12* en el sitio 2 estimado por el DASP3

Proteína	Energía de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
BCL2 like Protein 12	-2.21	0.0	GLN 187, LEU 190, LEU 238, LEU 241, SER 242, GLY 292, LEU 293, GLY 294, LEU 297
	-1.74	0.0	GLN 187, LEU 190, LYS 191, LEU 238, SER 242, GLY 292, LEU 293, GLY 294
	-1.29	0.0	GLN 187, LEU 190, LEU 241, SER 242, LEU 297, THR 296
	-0.13	0.0	GLN 187, LEU 190, LYS 191, LEU 238, LEU 241, SER 242
	0.01	0.0	LEU 190, LEU 238, LEU 241, SER 242, GLY 292, LEU 293, GLY 294
	0.01	0.0	LEU 238, LEU 241, SER 242, GLY 292, LEU 293, GLY 294, LEU 297
	0.41	1.7	LEU 190, LEU 238, LEU 241, SER242, LEU 297
	0.44	0.0	LEU 190, LEU 241, SER 242, LEU 238, LEU 293
Docking 237-247	0.97	0.0	LYS 237, LEU 238, LEU 241, GLY 292, LEU 293, GLY 294
	1.36	0.0	LEU 190, LEU 241, SER 242, LEU 238, LEU 297

5.4.2. Docking Galactine-9

Se realizó un *Docking* con la rejilla cubriendo toda la proteína (*Docking* “ciego”) *Galactine-9*, tabla 5-7, para conocer los lugares de la proteína donde, de acuerdo al *Docking*, se presenta mayor energía de unión con el ligando ATP. Presentando mayor afinidad en los residuos como son: 7,8,22,204,207,323, 36, 265, 309. Evidenciando mayor energía de unión en algunos residuos del sitio 1 y del sitio 2 predichos por el DASP.

Los sitios de unión predichos por el DASP3 son:

1. **Sitio 1:** TVLSSSGTRFA correspondiente a las posiciones 36 a la 46 de la secuencia FASTA.
2. **Sitio 2:** HLFYYHRLRNLP correspondiente a las posiciones 293 a la 305 de la secuencia FASTA.

En consecuencia, se realizó un *Docking* dirigido únicamente a la región del sitio 1 con los residuos estimados por el DASP3.

En la tabla 5-8, se muestran los sitios de unión con sus respectivas energías y valores RMSD, para las posibles conformaciones del ligando.

Tabla 5-7.: Conformaciones *Docking* cubriendo toda la *Galactine-9* en interacción con ATP

Proteína	Energías de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
Galactin-9	-1.74	0.0	GLN 7, ALA 8, GLN 22, THR 204, LYS 207, THR 323
	-1.23	0.0	THR 36, LYS 265, MET 266, VAL 269, ASP 290, GLN 292
	-1.09	0.0	PHE 240, LYS 265, MET 266, PHE 268, VAL 269
	0.79	0.0	GLY 229, ASN 230, ILE 232, GLN 250, TRP 255, ARG 309, GLU 311
	-0.31	0.0	PHE 66, GLN 94, LYS 95
	-0.4	0.0	PRO 192, MET 193, PRO 194, ILE 196
Docking ciego	-0.05	0.0	PHE 240, LYS 265, MET 266, VAL 269, ARG 270
	-0.02	0.0	GLN 148, TYR 191, PHE 195, ILE 196, THR 197, THR 198, VAL 321
	-0.58	0.0	TYR 191, THR 197, THR 323
	0.07	0.0	SER 12, GLN 141, GLY 291, GLN 292, ARG 264

En la tabla 5-8, algunas conformaciones presentan una energía de unión (-5.61), aún más negativa que las energías de unión obtenidas para las proteínas de entrenamiento del DASP3 (tablas 5-1, 5-2 y 5-3). Adicionalmente, el valor RMSD es menor a 2 Å para todas las conformaciones, lo cual indica que la estructura geométrica del ligando es muy estable dentro de todas las posibles interacciones estimadas.

Por tanto, el DASP3 muy posiblemente identificó el sitio 1 como un posible sitio alostérico

5.4 Energías de unión y RMSD para proteínas identificadas por el DASP3

Tabla 5-8.: Conformaciones *Galactine-9* con ATP en el sitio 1 predicho por el DASP3

Proteína	Energía de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
Galactine-9	-5.61	0.0	SER 39, SER 40, SER 41, GLY 42, PHE 66, PHE 93, GIN 94, LYS 95
	-3.13	0.0	PHE 66, GLY 69, MET 91, PHE 93, GIN 94, LYS 95
	-2.79	1.50	PHE 66, GLU 67, ASP 68, GLY 69, PHE 93, GIN 94, LYS 95
	-2.66	0.0	SER 40, SER 41, GLY 42, THR 43, GLU 66, GLU 67, ASP 68, GLY 69, LYS 96
	-2.34	0.0	SER 39, SER 40, GLY 42, LYS 95
	-1.99	1.61	PHE 66, GLU 67, GLY 69, GLY 70, MET 91, PHE 93, GIN 94, LYS 95
Docking 36-46	-1.61	0.0	GLN 94, LYS 95
	-1.47	0.0	PHE 66, MET 91, GIN 94, LYS 95
	-1.16	0.0	PHE 66, GLY 69, GLY 70, PHE 93, GIN 94, LYS 95
	-1.01	0.0	SER 39, SER 40, GLY 42, LYS 95

para el ligando ATP con la proteína *Galactine-9*.

Seguidamente se realizó *Docking* dirigido únicamente a la región del sitio 2 con los residuos estimados por el DASP3. En la tabla 5-9, se muestran los sitios de unión con sus respectivas energías y valores RMSD, para las posibles conformaciones del ligando.

Tabla 5-9.: Conformaciones *Galactine-9* con ATP en el sitio 2 predicho por el DASP3

Proteína	Energías de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
Galactine-9	-3.37	0.0	ALA 282, HIS 283, HIS 299, ARG 302, ASN 303, LEU 304, PRO 305
	-2.26	0.0	PRO 9, LEU 11, PRO 205, SER 206, LEU 279, GLU 281, ALA 282
	-1.33	1.51	ALA 282, HIS 283, HIS 299, ARG 302, ASN 303, PRO 305
	-1.41	0.0	PRO 9, LEU 11, TRP 277, LEU 279, LYS 286
	-1.05	0.0	SER 228, HIS 231, LEU 301, ARG 302, ASN 303, THR 306
	-0.96	0.0	LEU 11, SER 206, LEU 279, GLU 281, ALA 282
	-0.97	0.0	ALA 282, HIS 283, ARG 302, ASN 303, LEU 304
Docking 293-305	-0.9	0.0	TYR 10, LEU 11, SER 206, TRP 277, LEU 279
	-0.56	0.0	PRO 9, TYR 10, LEU 11, TRP 277, LEU 279, LYS 286
	0.13	0.0	PRO 205, SER 206, GLU 281, ALA 282, HIS 283, LYS 286

En la tabla 5-9, se presenta la mayor energía de unión de -3.37, con un valor RMSD menor a 2 Å, lo que indica, que el DASP3 posiblemente identificó otro sitio alostérico de unión para

el ligando ATP con la proteína *Galactine-9*.

5.4.3. Docking Fumarate hydratase, mitochondrial

Se realizó un *Docking* con la rejilla cubriendo toda la proteína (*Docking* “ciego”) *Fumarate hydratase, mitochondrial*, tabla 5-10, para conocer los lugares de la proteína donde, de acuerdo al *Docking*, se presenta mayor energía de unión con el ligando ATP.

En la tabla 5-10 se presentan bajas energías de unión, es decir, son más positivas en comparación con el *docking* “ciego” realizado en la *Galactine-9* (Tabla 5-7), a pesar de esto, entre sus mejores energías de unión, se encuentran algunos residuos del sitio 1 y del sitio 2, predichos por el DASP3.

Los sitios de unión predichos por el DASP3 son:

1. **Sitio 1:** TLGQEFSG correspondiente a las posiciones 243 a la 250.
2. **Sitio 2:** KVAAKVAALTGLP correspondiente a las posiciones 292 a la 304.

En consecuencia, se realizó un *Docking* dirigido únicamente a la región del sitio 1 con los residuos estimados por el DASP3. En la tabla 5-11, se muestran los sitios de unión con sus respectivas energías y valores RMSD, para las posibles conformaciones del ligando.

Tabla 5-10.: Conformaciones *Docking* cubriendo toda la *Fumarate hydratase, mitochondrial* en interacción con ATP

Proteína	Energía de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
Fumarate hydratase, mitochondrial	-0.60	0.0	THR 85, ARG 87, THR 90, ASP 134, LYS 414
	-0.11	0.0	LEV 507, GLY 508, PRO 509, LYS 510
	0.48	0.0	GLN 253, GLN 254, TYR 257, ARG 261, SER 334
	0.67	0.0	GLN 254, TYR 257, ARG 261, SER 334
	0.71	0.0	GLN 253, TYR 257, ARG 261
	1.08	0.0	ARG 261, ALA 264
Docking ciego	1.09	0.0	GLN 254, ARG 261, SER 334, CYS 333, LYS 337
	1.45	0.0	TYR 257, ARG 261
	1.73	0.0	THR 260, ARG 261
	2.68	0.0	HIS 391, VAL 392

La tabla 5-11, presenta la mayor energía de unión en -2.19, siendo más positiva que las energías de unión obtenidas para las proteínas de entrenamiento del DASP3 (tablas 5-1, 5-2 y 5-3) con un valor RMSD de 0.0 en todas las conformaciones. A pesar de que el DASP3 identificó el sitio 1 como un sitio alostérico para el ligando ATP con la proteína, es menos probable, que el sitio 1 de la *Fumarate hidratase, mitochondrial*, sea un sitio de unión.

Seguidamente se realizó *Docking* dirigido únicamente a la región del sitio 2 con los residuos estimados por el DASP3. En la tabla 5-12, se muestran los sitios de unión con sus respectivas energías y valores RMSD, para las posibles conformaciones del ligando.

5.4 Energías de unión y RMSD para proteínas identificadas por el DASP3

Tabla 5-11.: Conformaciones ATP con *Fumarate hydratase, mitochondrial* en el sitio 1 predicho por el DASP3

Proteína	Energía de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
Fumarate hydratase, mitochondrial	-2.19	0.0	LEV 242, GLN 246, GLU 247, GLY 250, TYR 251, LYS 337, ASP 341
	-1.76	0.0	LEV 242, GLN 246, GLU 247, SER 249, GLY 250, LYS 337
	-1.67	0.0	GLU 247, SER 249, GLY 250, TYR 251, GLN 253, GLN 254, LYS 337
	-1.51	0.0	LEV 242, GLU 247, GLY 250, TYR 251, GLN 253, GLN 254, LYS 337
	-1.24	0.0	LEV 242, GLN 246, GLU 247, SER 249, GLY 250, TYR 251, GLN 253, LYS 337, ASP 341
	Docking 243-250	-1.15	0.0
-1.13		0.0	LEV 242, GLN 246, GLU 247, SER 249, GLY 250, GLN 253, LYS 337

Tabla 5-12.: Conformaciones ATP con *Fumarate hydratase, mitochondrial* en el sitio 2 predicho por el DASP3

Proteína	Energía de unión (Kcal/mol)	RMSD (Å)	Sitio de unión
Fumarate hydratase, mitochondrial	-3.35	0.0	ARG 101, GLU 105, VAL 106, GLN 108, ASP 109, LYS 292, ALA 295, LYS 296, ALA 299
	-1.20	0.0	GLU 291, ALA 294, ALA 295, ALA 298, ALA 299, PHE 305, VAL 306
	-0.87	0.0	ALA 294, ALA 295, ALA 298, LEV 303, PRO 304, PHE 305, VAL 306, THR 307
	-0.69	0.0	GLU 105, VAL 106, ASP 109, LYS 292, LYS 296, ALA 299
	-0.45	0.0	GLU 291, ALA 294, ALA 295, ALA 298, PHE 305, VAL 306
	-0.25	0.0	ALA 295, ALA 298, ALA 299, GLY 302
	Docking 292-304	-0.16	0.0
-0.05		0.0	ARG 101, GLU 105, LYS 292, ALA 295, LYS 296, ALA 299, LEU 300
-0.03		0.0	GLU 291, ALA 294, ALA 295, ALA 298, PRO 304, PHE 305, VAL 306, THR 307

La tabla 5-12, presenta la mayor energía en -3.5, con un valor RMSD de 0.0 en todas las conformaciones, por lo que, el DASP3 posiblemente identificó un sitio alostérico de unión con el ligando ATP con la proteína *Fumarate hydratase, mitochondrial*.

5.5. Comparación energías de unión de proteínas identificadas del NCBI con proteínas de entrenamiento en el DASP3

Las mayores energías de unión de las proteínas del NCBI, anteriormente descritas e identificadas mediante el perfil alostérico en interacción con el ligando ATP, se encuentran con valores entre -5.61 Kcal/mol y -2.19 Kcal/mol, con un valor RMSD menor a 2 Å.

A diferencia, las mayores energías en el presente análisis de las proteínas de entrenamiento para el DASP3 están entre -5.77 Kcal/mol y -3.46 Kcal/mol.

Considerando que mientras más negativa es la energía de unión y menor es el valor RMSD, mayor certeza hay que el ligando sea un activador o inhibidor alostérico en ese sitio de unión de la proteína y comparando las energías de interacción para las proteínas de entrenamiento (*Endoplasmic reticulum chaperone BiP*, tabla **5-1**, *Cytosolic purine 5'-nucleotidase*, tabla **5-2** y *ATP- depend 6- phosphofructokinase* tabla **5-3**) con las energías de unión de las proteínas obtenidas mediante los algoritmos del DASP3 (*BCL2-Like-protein 12*, tablas: **5-4**, **5-5**, **5-6**, *Galactine-9*, tablas: **5-7**,**5-8**,**5-9**, *Fumarate hydratase, mitochondrial*, tablas: **5-10**, **5-11**, **5-12**), es posible afirmar que los sitios de interacción, identificados por el DASP3, son un potencial sitio alostérico de la proteína con el ligando ATP.

6. Conclusiones

En el presente trabajo se abordó la identificación de sitios alostéricos en interacción con el ligando ATP en proteínas del Homo sapiens. Se planteó y verificó un procedimiento, utilizando herramientas computacionales, para encontrar sitios alostéricos que interactúan con un determinado ligando a partir de la estructura primaria de las proteínas.

Dando cumplimiento a los objetivos planteados al inicio del presente documento se seleccionaron tres (3) proteínas de Homo sapiens a las que se le conoce su actividad alostérica e interacción con el ligando ATP, para elaborar un perfil alostérico (ASP) empleando la herramienta DASP3. A partir de la información de la estructura primaria de las proteínas de Homo sapiens, descargado del NCBI, el DASP3 identificó proteínas con sus respectivos sitios de unión donde posiblemente existe un sitio alostérico con el ligando ATP. Posteriormente, se realizó *Docking* para confirmar, computacionalmente, que las proteínas identificadas por el DASP3 cuentan con un sitio alostérico modulado por el ATP.

El protocolo propuesto mediante el DASP3 esta encontrando sitios alostéricos que pueden interactuar virtualmente con el ligando ATP, por lo que podría actuar como modulador, lo que en un futuro podría ser integrado en tratamientos terapéuticos como el cáncer.

De acuerdo al trabajo realizado en el DASP se concluye que es posible emplear algoritmos computacionales, creados para identificación de sitios activos, para identificar sitios alostéricos.

El programa de AutoDockTools cuenta con más de un criterio de selección y proporciona más de una conformación, lo que lleva a elegir de entre las mejores energías y el mejor RMSD de cada una de ellas, además de presentar residuos similares pero no iguales en cada pose del ligando, por tanto, se requiere realizar una verificación en laboratorio.

6.1. Recomendaciones para trabajos futuros

Las recomendaciones, basadas en los resultados obtenidos son:

- Indagar sobre nuevas proteínas de Homo sapiens (archivo FASTA) para realizar el protocolo 2 con el DASP3, que cuenten con información suficiente a la hora de conocer su funcionalidad en procesos vitales del ser humano.
- Realización de dinámica molecular con el fin de visualizar y analizar en el tiempo las conformaciones obtenidas mediante AutoDockTools.

- Verificación, a proteínas del NCBI, a las que no se le realizó *docking*, para conocer cuales de ellas pueden contar con un sitio alostérico en interacción con el ligando ATP.
- Realización del procedimiento propuesto con otros ligandos para conocer nuevos moduladores alostéricos en proteínas.

A. Lista de proteínas identificadas mediante el DASP3

A.0.1. pdb con exponente -5

Submitted name: 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase

Seq pdb—6Q19—L PASSED. Final pValue = 3.017228564675582E-5

Submitted name: Alpha/beta hydrolase

Seq pdb—6Q0E—L PASSED. Final pValue = 2.690435849927383E-5

Submitted name: UDP-glycosyltransferase 79A6

Seq pdb—6FG2—E PASSED. Final pValue = 5.348127578040633E-5 ;E,G

Fumarate hydratase, mitochondrial

Seq pdb—5UPP—A PASSED. Final pValue = 7.311472156930086E-5 ;AB

Seq pdb—6EBT—A PASSED. Final pValue = 7.714058322177067E-5 ;AB

Seq pdb—5D6B—A PASSED. Final pValue = 7.265035171450406E-5

Seq pdb—3E04—A PASSED. Final pValue = 7.166859894670971E-5 ;AD

Histone-lysine N-methyltransferase EHMT1

Seq pdb—5VSF—A PASSED. Final pValue = 1.6421081048321003E-5 ;AB

Seq pdb—5VSD—A PASSED. Final pValue = 1.6421081048321003E-5 ;AB

Seq pdb—5TUZ—A PASSED. Final pValue = 1.6421081048321003E-5 ;AB

Seq pdb—3SW9—A PASSED. Final pValue = 2.1339852751578386E-5 ;AB

Seq pdb—3HNA—A PASSED. Final pValue = 2.4254309569480813E-5 ;AB

Seq pdb—5TTG—A PASSED. Final pValue = 2.4254309569480813E-5 ;AB

Seq pdb—5V9J—A PASSED. Final pValue = 2.4254309569480813E-5 ;AB

Seq pdb—6MBP—A PASSED. Final pValue = 1.6421081048321003E-5 ;AB

Seq pdb—6MBO—A PASSED. Final pValue = 1.6421081048321003E-5 ;AB

Tumor necrosis factor ligand superfamily member 13B

Seq pdb—5Y9J—L PASSED. Final pValue = 8.359797426216899E-5

Seq pdb—6FXN—I PASSED. Final pValue = 8.359797426216899E-5 ;GERPN

PTB domain-containing engulfment adapter protein 1

Seq pdb—6ITU—A PASSED. Final pValue = 4.988050180301146E-5

Phosphatidylcholine-sterol acyltransferase

Seq pdb—5BV7—L PASSED. Final pValue = 6.275669533945528E-5

Guanylate cyclase soluble subunit alpha-1

Seq pdb—4NI2—B PASSED. Final pValue = 6.078299781209105E-5

Platelet-derived growth factor subunit B

Seq pdb—4QCI—A PASSED. Final pValue = 4.574670510024468E-5 ;A,L

Beta-lactamase

Seq pdb—6AZZ—B PASSED. Final pValue = 8.902280498783742E-5;B,E

Galectin-9

Seq pdb—3WV6—A PASSED. Final pValue = 8.081078453906514E-5 ;AB

Immunoglobulin heavy constant gamma 1

Seq pdb—4DAG—L PASSED. Final pValue = 3.215742986200315E-5

Heat shock-related 70 kDa protein 2

Seq pdb—5FPN—A PASSED. Final pValue = 3.814714015870882E-5;A-B

A.0.2. NP con exponente -5

eukaryotic peptide chain release factor subunit 1 isoform 2

Seq NP_001351089.1 PASSED. Final pValue = 8.954932752946462E-5

Seq NP_001278904.1 PASSED. Final pValue = 8.954932752946462E-5

Seq NP_001243231.1 PASSED. Final pValue = 8.954932752946462E-5

Seq NP_001278903.1 PASSED. Final pValue = 8.954932752946462E-5

Seq NP_001269114.1 PASSED. Final pValue = 9.897835774781682E-5

bcl-2-like protein 12 isoform 3

Seq NP_001035758.1 PASSED. Final pValue = 2.6365167659979937E-5

Seq NP_619580.1 PASSED. Final pValue = 2.725832778274344E-5

1-phosphatidylinositol 3-phosphate 5-kinase isoform 2 [Homo sapiens]

Seq NP_055855.2 PASSED. Final pValue = 1.685269072898643E-5

dual 3',5'-cyclic-AMP and -GMP phosphodiesterase 11A isoform 4

Seq NP_058649.3 PASSED. Final pValue = 4.287432780204796E-5

leucine-rich repeat-containing protein 10

Seq NP_963844.2 PASSED. Final pValue = 1.0338992741430656E-5

zinc transporter ZIP10 precursor

Seq NP_001120729.1 PASSED. Final pValue = 3.8511335192856176E-5

Seq NP_065075.1 PASSED. Final pValue = 3.8511335192856176E-5

lysosomal acid glucosylceramidase isoform 1 precursor

Seq NP_001005741.1 PASSED. Final pValue = 3.358247584374615E-5

Seq NP_001165283.1 PASSED. Final pValue = 2.3987705523696532E-5

Seq NP_001005742.1 PASSED. Final pValue = 3.358247584374615E-5

Seq NP_000148.2 PASSED. Final pValue = 3.358247584374615E-5

PTB domain-containing engulfment adapter protein 1 isoform a

Seq NP_001362862.1 PASSED. Final pValue = 9.259392788615386E-5

Seq NP_001362866.1 PASSED. Final pValue = 9.259392788615386E-5
Seq NP_001362855.1 PASSED. Final pValue = 9.500021584219442E-5
Seq NP_001362872.1 PASSED. Final pValue = 9.471048765999192E-5
Seq NP_001362863.1 PASSED. Final pValue = 9.259392788615386E-5
Seq NP_001362869.1 PASSED. Final pValue = 9.471048765999192E-5
Seq NP_001362864.1 PASSED. Final pValue = 9.259392788615386E-5
Seq NP_001362881.1 PASSED. Final pValue = 6.37606891810356E-5
Seq NP_001362870.1 PASSED. Final pValue = 9.471048765999192E-5
Seq NP_001362880.1 PASSED. Final pValue = 8.503799526508495E-5
Seq NP_001362858.1 PASSED. Final pValue = 9.500021584219442E-5
Seq NP_001362857.1 PASSED. Final pValue = 9.500021584219442E-5
Seq NP_001362865.1 PASSED. Final pValue = 9.259392788615386E-5
Seq NP_001362856.1 PASSED. Final pValue = 9.500021584219442E-5
Seq NP_001362861.1 PASSED. Final pValue = 9.259392788615386E-5
Seq NP_001362871.1 PASSED. Final pValue = 9.471048765999192E-5
Seq NP_001362859.1 PASSED. Final pValue = 9.500021584219442E-5
Seq NP_001362868.1 PASSED. Final pValue = 9.259392788615386E-5
Seq NP_001362860.1 PASSED. Final pValue = 9.500021584219442E-5
Seq NP_001362854.1 PASSED. Final pValue = 9.500021584219442E-5
Seq NP_001362867.1 PASSED. Final pValue = 9.259392788615386E-5
Seq NP_001239598.1 PASSED. Final pValue = 4.7601595756768124E-5
Seq NP_001239597.1 PASSED. Final pValue = 5.268672388966307E-5
Seq NP_057399.1 PASSED. Final pValue = 9.500021584219442E-5 **cell division cycle-associated protein 3 isoform 4**
Seq NP_001317948.1 PASSED. Final pValue = 4.2124772871557096E-5
Seq NP_001284533.1 PASSED. Final pValue = 6.34764310935441E-5
heat shock-related 70 kDa protein 2
Seq NP_068814.2 PASSED. Final pValue = 3.839415911000301E-5
heat shock cognate 71 kDa protein isoform 1,2
Seq NP_006588.1 PASSED. Final pValue = 3.150760260410084E-5
Seq NP_694881.1 PASSED. Final pValue = 2.5511230970706225E-5
keratin, type I cytoskeletal 23 isoform 2
Seq NP_001269362.1 PASSED. Final pValue = 1.6452220944923376E-5
immediate early response gene 2 protein
Seq NP_004898.2 PASSED. Final pValue = 5.386411271332673E-5
galectin-9B
Seq NP_001036150.1 PASSED. Final pValue = 1.3380030368604777E-5
Seq NP_001354221.1 PASSED. Final pValue = 6.421104751228931E-5
Seq NP_002299.2 PASSED. Final pValue = 3.0801036207096144E-5
Seq NP_033665.1 PASSED. Final pValue = 3.517353089803895E-5

A.0.3. SP con exponente -5

1-phosphatidylinositol 3-phosphate 5-kinase

Seq sp-Q9Y2I7.3—FYV1_HUMAN PASSED. Final pValue = 1.685269072898643E-5

Dual 3',5'-cyclic-AMP and -GMP phosphodiesterase 11A

Seq sp-Q9HCR9.2—PDE11_HUMAN PASSED. Final pValue = 4.287432780204796E-5

Galectin-9B

Seq sp-Q3B8N2.3—LEG9B_HUMAN PASSED. Final pValue = 6.421104751228931E-5

Seq sp-O00182.2—LEG9_HUMAN PASSED. Final pValue = 3.517353089803895E-5

Leucine-rich repeat-containing protein 10

Seq sp-Q5BKY1.1—LRC10_HUMAN PASSED. Final pValue = 1.0338992741430656E-5

PTB domain-containing engulfment adapter protein 1

Seq sp-Q9UBP9.1—GULP1_HUMAN PASSED. Final pValue = 9.500021584219442E-5

Immediate early response gene 2 protein

Seq sp-Q9BTL4.1—IER2_HUMAN PASSED. Final pValue = 5.386411271332673E-5

Lysosomal acid glucosylceramidase

Seq sp-P04062.3—GLCM_HUMAN PASSED. Final pValue = 3.358247584374615E-5

Bcl-2-like protein 12

Seq sp-Q9HB09.1—B2L12_HUMAN PASSED. Final pValue = 2.725832778274344E-5

Heat shock-related 70 kDa protein 2

Seq sp-P54652.1—HSP72_HUMAN PASSED. Final pValue = 3.839415911000301E-5

Seq sp-P11142.1—HSP7C_HUMAN PASSED. Final pValue = 3.150760260410084E-5

Zinc transporter ZIP10

Seq sp-Q9ULF5.2—S39AA_HUMAN PASSED. Final pValue = 3.8511335192856176E-5

A.0.4. XP con exponente -5

1-phosphatidylinositol 3-phosphate 5-kinase isoform X1,X2,X3,X5,X6,X7,X10,X12,X16

Seq XP_011509081.1 PASSED. Final pValue = 1.806782122375991E-5

Seq XP_011509080.1 PASSED. Final pValue = 1.806782122375991E-5

Seq XP_011509082.1 PASSED. Final pValue = 1.802956016742831E-5

Seq XP_011509083.1 PASSED. Final pValue = 1.751775584329631E-5

Seq XP_011509084.1 PASSED. Final pValue = 1.4880529289473295E-5

Seq XP_011509086.1 PASSED. Final pValue = 1.8119801706319804E-5

Seq XP_011509088.1 PASSED. Final pValue = 1.600383853310069E-5

Seq XP_011509085.1 PASSED. Final pValue = 1.8158318489133522E-5

Seq XP_011509090.1 PASSED. Final pValue = 1.6218595472734882E-5

Seq XP_011509092.1 PASSED. Final pValue = 1.0134816553518826E-5
Seq XP_016860011.1 PASSED. Final pValue = 1.2514296621935248E-5
Seq XP_016859796.1 PASSED. Final pValue = 7.812442145702932E-5
Seq XP_016859795.1 PASSED. Final pValue = 9.500021584219442E-5
Seq XP_016859731.1 PASSED. Final pValue = 9.777017113011406E-5
Seq XP_016859730.1 PASSED. Final pValue = 9.777017113011406E-5
Seq XP_016859062.1 PASSED. Final pValue = 1.0134816553518826E-5
Seq XP_016859063.1 PASSED. Final pValue = 1.0134816553518826E-5
Seq XP_016859060.1 PASSED. Final pValue = 1.5062284322722307E-5
Seq XP_016859061.1 PASSED. Final pValue = 1.0134816553518826E-5
Seq XP_016859058.1 PASSED. Final pValue = 1.6144876624657342E-5
Seq XP_016859059.1 PASSED. Final pValue = 1.6335882948010914E-5
Seq XP_016859057.1 PASSED. Final pValue = 1.6362015468994565E-5
Seq XP_016880112.1 PASSED. Final pValue = 3.806605119775409E-5
Seq XP_011522897.1 PASSED. Final pValue = 1.6452220944923376E-5
Seq XP_011522298.1 PASSED. Final pValue = 2.2073967521280148E-5
Seq XP_011522299.1 PASSED. Final pValue = 2.646230613855295E-5
Seq XP_011541100.1 PASSED. Final pValue = 3.150760260410084E-5
Seq XP_011513349.1 PASSED. Final pValue = 1.1284933148490655E-5
Seq XP_011509808.1 PASSED. Final pValue = 3.8511335192856176E-5
Seq XP_011509806.1 PASSED. Final pValue = 3.8511335192856176E-5
Seq XP_011509809.1 PASSED. Final pValue = 1.2514296621935248E-5
Seq XP_011509807.1 PASSED. Final pValue = 3.8511335192856176E-5
Seq XP_005246746.2 PASSED. Final pValue = 3.8511335192856176E-5
Seq XP_011509637.1 PASSED. Final pValue = 7.812442145702932E-5
Seq XP_011509091.1 PASSED. Final pValue = 1.494277466095815E-5
Seq XP_011509089.1 PASSED. Final pValue = 1.5967904347712528E-5
Seq XP_011509087.1 PASSED. Final pValue = 1.7476883027172704E-5

glucokinase regulatory protein isoform X2

Seq XP_011531065.1 PASSED. Final pValue = 4.760307687795723E-5

fumarate hydratase, mitochondrial isoform X1

Seq XP_011542434.1 PASSED. Final pValue = 5.3789643151161664E-5

PTB domain-containing engulfment adapter protein 1 isoform X7

Seq XP_006712652.1 PASSED. Final pValue = 7.740689245327022E-5

keratin, type I cytoskeletal 23 isoform X1

Seq XP_005257257.1 PASSED. Final pValue = 1.6452220944923376E-5

galectin-9C isoform X3,X2,X1

Seq XP_005256827.1 PASSED. Final pValue = 4.054753494130673E-5

Seq XP_006721955.1 PASSED. Final pValue = 3.3582448375834235E-5

Seq XP_005256826.1 PASSED. Final pValue = 4.44607820600521E-5

A.0.5. Diferentes con exponente -5

GBA

Seq ASU47376.1 PASSED. Final pValue = 3.438503109209894E-5

Adenomatous Polyposis Coli, partial

Seq SMR98740.1 PASSED. Final pValue = 4.2556569571945425E-5

unknown, partial

Seq AAX93242.1 PASSED. Final pValue = 5.797239708004366E-5

human factor FVIII, partial

Seq ABR22588.1 PASSED. Final pValue = 6.57881229162231E-5

Seq ABR22587.1 PASSED. Final pValue = 6.57881229162231E-5

glucocerebrosidase

Seq AAC51820.1 PASSED. Final pValue = 3.358247584374615E-5

Seq AAC63056.1 PASSED. Final pValue = 3.358247584374615E-5

U15

Seq AMD82145.1 PASSED. Final pValue = 3.6264228253928864E-5

immunoglobulin light chain variable region, partial

Seq AMB38645.1 PASSED. Final pValue = 8.670913898843688E-5

Seq AAB36577.1 PASSED. Final pValue = 2.0178993322097413E-5

Seq AAD16668.1 PASSED. Final pValue = 8.152996382212678E-5

Seq AAD16665.1 PASSED. Final pValue = 2.0892576205147034E-5

Seq AAK94814.1 PASSED. Final pValue = 8.52869121765942E-5

Seq ACJ71719.1 PASSED. Final pValue = 5.791549855004366E-5

Seq ACF34422.1 PASSED. Final pValue = 2.2913912925383106E-5

Seq ACF34441.2 PASSED. Final pValue = 9.777694633872572E-5

Seq ABU90630.2 PASSED. Final pValue = 4.710652420778833E-5

Seq ABU90629.2 PASSED. Final pValue = 8.420854263934053E-5

Seq ABU90558.2 PASSED. Final pValue = 9.530583484702515E-5

Seq ABU90561.1 PASSED. Final pValue = 7.585424678214293E-5

Seq ASL06614.1 PASSED. Final pValue = 3.49527689404245E-5

Seq ASL06612.1 PASSED. Final pValue = 8.592190196322964E-5

Seq AVN89174.1 PASSED. Final pValue = 4.5356718258406416E-5

Seq AVN88957.1 PASSED. Final pValue = 5.7180344894876843E-5

Seq AVN88922.1 PASSED. Final pValue = 1.5345796166902273E-5

Seq QBK47554.1 PASSED. Final pValue = 4.7777356138888433E-5

Seq QBK47550.1 PASSED. Final pValue = 4.422037041909841E-5

Seq QBK47546.1 PASSED. Final pValue = 2.974042964221615E-5

Seq QBK47536.1 PASSED. Final pValue = 4.9943576319290616E-5

Seq QBK47559.1 PASSED. Final pValue = 4.188450826806323E-5
Seq QBK47513.1 PASSED. Final pValue = 5.6861831990731426E-5
Seq QDF62259.1 PASSED. Final pValue = 5.505513007941506E-5
Seq QDF62268.1 PASSED. Final pValue = 2.8288118099338076E-5
Seq QDF61469.1 PASSED. Final pValue = 9.391791165865672E-5

KIAA1265 protein

Seq BAA86579.2 PASSED. Final pValue = 3.8511335192856176E-5

phosphodiesterase 11A, isoform CRA_b

Seq EAX11051.1 PASSED. Final pValue = 6.0006588260334786E-5
Seq EAX11049.1 PASSED. Final pValue = 5.949879464406791E-5
Seq EAX11047.1 PASSED. Final pValue = 5.9517657622120557E-5
Seq EAX11050.1 PASSED. Final pValue = 4.2864382346976056E-5
Seq EAX11048.1 PASSED. Final pValue = 4.2978213530852194E-5

GULP, engulfment adaptor PTB domain containing 1, isoform CRA_a

Seq EAX10914.1 PASSED. Final pValue = 9.500021584219442E-5
Seq EAX10917.1 PASSED. Final pValue = 9.500021584219442E-5
Seq EAX10915.1 PASSED. Final pValue = 9.500021584219442E-5
Seq EAX10913.1 PASSED. Final pValue = 9.500021584219442E-5

mitochondrial ribosomal protein L28, isoform CRA_d

Seq EAW85831.1 PASSED. Final pValue = 2.649229050503486E-5

immediate early response 2, isoform CRA_a

Seq EAW84360.1 PASSED. Final pValue = 5.386411271332673E-5
Seq EAW84358.1 PASSED. Final pValue = 5.386411271332673E-5
Seq EAW84359.1 PASSED. Final pValue = 5.386411271332673E-5

TATA box binding protein like 2, isoform CRA_b

Seq EAW80676.1 PASSED. Final pValue = 1.1206068898569054E-5

hCG2038865, partial

Seq EAW75594.1 PASSED. Final pValue = 7.426238411525289E-5

phosphatidylinositol-3-phosphate/phosphatidylinositol 5-kinase, type III, isoform CRA_d

Seq EAW70444.1 PASSED. Final pValue = 1.618565376740841E-5
Seq EAW70449.1 PASSED. Final pValue = 1.0106443272647293E-5
Seq EAW70447.1 PASSED. Final pValue = 1.676719731416803E-5

solute carrier family 39 (zinc transporter), member 10, isoform CRA_a

Seq EAW70116.1 PASSED. Final pValue = 3.8511335192856176E-5
Seq EAW70118.1 PASSED. Final pValue = 3.8511335192856176E-5
Seq EAW70119.1 PASSED. Final pValue = 3.8511335192856176E-5

fumarate hydratase, isoform CRA_d

Seq EAW70093.1 PASSED. Final pValue = 5.3789643151161664E-5
Seq EAW70090.1 PASSED. Final pValue = 8.835128397805136E-5

heat shock 70kDa protein 8, isoform CRA_a

Seq EAW67540.1 PASSED. Final pValue = 3.150760260410084E-5

Seq EAW67541.1 PASSED. Final pValue = 2.5511230970706225E-5

Seq EAW67539.1 PASSED. Final pValue = 3.150760260410084E-5

Seq EAW80864.1 PASSED. Final pValue = 3.839415911000301E-5

Seq EAW80865.1 PASSED. Final pValue = 3.839415911000301E-5

eukaryotic translation termination factor 1, isoform CRA_b

Seq EAW62132.1 PASSED. Final pValue = 8.954932752946462E-5

centaurin, gamma 3, isoform CRA_d, partial

Seq EAW54027.1 PASSED. Final pValue = 2.4584930738759115E-5

Seq EAW53103.1 PASSED. Final pValue = 5.899107685216637E-5

glucosidase, beta; acid (includes glucosylceramidase), isoform CRA_a

Seq EAW53101.1 PASSED. Final pValue = 5.899107685216637E-5

Seq EAW53102.1 PASSED. Final pValue = 5.899107685216637E-5

Seq EAW53100.1 PASSED. Final pValue = 5.899107685216637E-5

BCL2-like 12 (proline rich), isoform CRA_d

Seq EAW52517.1 PASSED. Final pValue = 2.725832778274344E-5

lectin, galactoside-binding, soluble, 9 (galectin 9), isoform CRA_b

Seq EAW51045.1 PASSED. Final pValue = 3.517353089803895E-5

Seq EAW51043.1 PASSED. Final pValue = 3.0801036207096144E-5

Seq EAW51039.1 PASSED. Final pValue = 3.3582448375834235E-5

Seq EAW51037.1 PASSED. Final pValue = 3.0801036207096144E-5

Seq EAW51044.1 PASSED. Final pValue = 3.517353089803895E-5

Seq EAW51042.1 PASSED. Final pValue = 3.806605119775409E-5

Seq EAW51040.1 PASSED. Final pValue = 3.712407456216204E-5

Seq EAW51038.1 PASSED. Final pValue = 3.3582448375834235E-5

epididymis secretory sperm binding protein

Seq ADO22260.1 PASSED. Final pValue = 3.150760260410084E-5

Seq ADO22186.1 PASSED. Final pValue = 3.150760260410084E-5

A.0.6. Proteínas con exponente -6**Endoplasmic reticulum chaperone BiP**

Seq pdb—6ASY—A PASSED. Final pValue = 4.191388541439982E-6 ;AB

Seq pdb—5E84—A PASSED. Final pValue = 6.444205151809317E-6 ;A-F

ras GTPase-activating protein 1 isoform X1

Seq NP_002881.1 PASSED. Final pValue = 6.833919766438436E-6

Seq XP_011541827.1 PASSED. Final pValue = 6.285236229648092E-6

Endoplasmic reticulum chaperone BiP

Seq sp—P11021.2—BIP_HUMAN PASSED. Final pValue = 4.389587090197981E-6

Seq NP_005338.1 PASSED. Final pValue = 4.389587090197981E-6

Ras GTPase-activating protein 1; Short=GAP

Seq sp—P20936.1—RASA1_HUMAN PASSED. Final pValue = 6.833919766438436E-6

immunoglobulin lambda 3 light chain, partial

Seq ABU90710.1 PASSED. Final pValue = 8.996601584183286E-6

leucine rich repeat containing 10, partial

Seq EAW97231.1 PASSED. Final pValue = 9.593302196159977E-6

RAS p21 protein activator (GTPase activating protein) 1, isoform CRA_b

Seq EAW95932.1 PASSED. Final pValue = 6.91532862316812E-6

Seq EAW95933.1 PASSED. Final pValue = 6.91532862316812E-6

heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa), isoform CRA_a

Seq EAW87620.1 PASSED. Final pValue = 4.389587090197981E-6

thyroid hormone receptor associated protein 1, isoform CRA_a, partial

Seq EAW51437.1 PASSED. Final pValue = 4.123649990276631E-6

A.0.7. XP con exponente -7

Cytosolic purine 5'-nucleotidase Seq XP_024303678.1 PASSED. Final pValue = 2.484018155478276E-7

Seq XP_024303676.1 PASSED. Final pValue = 3.067738116476186E-7

Seq XP_024303673.1 PASSED. Final pValue = 3.067738116476186E-7

Seq XP_024303679.1 PASSED. Final pValue = 3.067738116476186E-7

Seq XP_024303677.1 PASSED. Final pValue = 2.484018155478276E-7

Seq XP_024303675.1 PASSED. Final pValue = 3.067738116476186E-7

Seq XP_024303672.1 PASSED. Final pValue = 3.293361137911282E-7

Seq XP_016871456.1 PASSED. Final pValue = 3.293361137911282E-7

Seq XP_016871463.1 PASSED. Final pValue = 2.484018155478276E-7

Seq XP_016871455.1 PASSED. Final pValue = 3.293361137911282E-7

Seq XP_006717786.1 PASSED. Final pValue = 3.293361137911282E-7

Seq XP_006717784.1 PASSED. Final pValue = 3.293361137911282E-7

Seq XP_005269703.1 PASSED. Final pValue = 3.293361137911282E-7

Seq XP_005269702.1 PASSED. Final pValue = 3.293361137911282E-7

A.0.8. NP con exponente -7

Cytosolic purine 5'-nucleotidase

Seq NP_001338126.1 PASSED. Final pValue = 3.293361137911282E-7

Seq NP_001338124.1 PASSED. Final pValue = 2.484018155478276E-7
Seq NP_001338122.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338120.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338118.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338116.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338114.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338112.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338110.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338108.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338106.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338123.1 PASSED. Final pValue = 2.484018155478276E-7
Seq NP_001338121.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338119.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338117.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338115.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338113.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338111.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338109.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338107.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338105.1 PASSED. Final pValue = 3.293361137911282E-7
Seq NP_001338125.1 PASSED. Final pValue = 2.484018155478276E-7

A.0.9. Diferente con exponente -7

Phosphofructokinase, muscle type

Seq ALQ33712.1 PASSED. Final pValue = 5.173211151585745E-7

A.0.10. Pdb con exponente -8

Cytosolic purine 5'-nucleotidase

Seq pdb—5OPP—A PASSED. Final pValue = 9.392245081743423E-8
Seq pdb—5OPN—A PASSED. Final pValue = 9.233242971031568E-8
Seq pdb—5OPK—A PASSED. Final pValue = 9.079084849781615E-8
Seq pdb—5OPO—A PASSED. Final pValue = 9.809520055863562E-8
Seq pdb—5OPM—A PASSED. Final pValue = 8.594773560829318E-8
Seq pdb—5OPL—A PASSED. Final pValue = 4.7409133608626E-8
Seq pdb—5CQZ—A PASSED. Final pValue = 4.7868477861627694E-8 ;AB
Seq pdb—5L4Z—A PASSED. Final pValue = 4.511819953439645E-8

Seq pdb—5L50—A PASSED. Final pValue = 4.8532874996255676E-8
Seq pdb—5K7Y—A PASSED. Final pValue = 4.749364024616417E-8
Seq pdb—2XCV—A PASSED. Final pValue = 4.695083748945954E-8
Seq pdb—2XJF—A PASSED. Final pValue = 4.694928913738752E-8
Seq pdb—5CR7—A PASSED. Final pValue = 4.803279949099185E-8;A-B
Seq pdb—6DE2—A PASSED. Final pValue = 3.764154914551416E-8
Seq pdb—6DE0—A PASSED. Final pValue = 4.501125117013394E-8
Seq pdb—6DDY—A PASSED. Final pValue = 4.85545685987217E-8
Seq pdb—6DDQ—B PASSED. Final pValue = 3.6838721226642753E-8
Seq pdb—6DDO—B PASSED. Final pValue = 3.715976104347154E-8
Seq pdb—6DDL—B PASSED. Final pValue = 4.501125117013394E-8
Seq pdb—6DDK—B PASSED. Final pValue = 3.6838721226642753E-8
Seq pdb—6DE3—A PASSED. Final pValue = 3.6838721226642753E-8
Seq pdb—6DDH—A PASSED. Final pValue = 4.657806140888786E-8
Seq pdb—6DE1—A PASSED. Final pValue = 3.715976104347154E-8
Seq pdb—6DDC—A PASSED. Final pValue = 4.695083748945954E-8
Seq pdb—6DDZ—A PASSED. Final pValue = 4.424364029175766E-8
Seq pdb—6DDB—A PASSED. Final pValue = 4.657806140888786E-8
Seq pdb—6DDX—A PASSED. Final pValue = 4.760257057366274E-8
Seq pdb—6DDQ—A PASSED. Final pValue = 3.6838721226642753E-8
Seq pdb—6DDO—A PASSED. Final pValue = 3.715976104347154E-8
Seq pdb—6DDL—A PASSED. Final pValue = 4.501125117013394E-8
Seq pdb—6DDK—A PASSED. Final pValue = 3.6838721226642753E-8
Seq pdb—6DDC—B PASSED. Final pValue = 4.695083748945954E-8
Seq pdb—6DDB—B PASSED. Final pValue = 4.657806140888786E-8
Seq pdb—6DD3—A PASSED. Final pValue = 4.805667251796739E-8
Seq pdb—6FXH—A PASSED. Final pValue = 3.793922380371164E-8
Seq pdb—6FIU—A PASSED. Final pValue = 4.884889374858706E-8
Seq pdb—6FIR—A PASSED. Final pValue = 4.7868477861627694E-8
Seq pdb—6FIW—A PASSED. Final pValue = 4.884889374858706E-8
Seq pdb—6FIS—A PASSED. Final pValue = 4.884889374858706E-8

A.0.11. XP con exponente -8

Cytosolic purine 5'-nucleotidase

Seq XP_024303671.1 PASSED. Final pValue = 3.434509659007881E-8
Seq XP_024303669.1 PASSED. Final pValue = 3.979098950120248E-8
Seq XP_024303670.1 PASSED. Final pValue = 3.793922380371164E-8
Seq XP_016871436.1 PASSED. Final pValue = 3.374767980686374E-8

Seq XP_011537839.1 PASSED. Final pValue = 2.415686658394379E-8

Seq XP_005269694.1 PASSED. Final pValue = 3.7072240722463464E-8

A.0.12. NP con exponente -8

Cytosolic purine 5'-nucleotidase

Seq NP_001338101.1 PASSED. Final pValue = 3.979098950120248E-8

Seq NP_036361.1 PASSED. Final pValue = 3.793922380371164E-8 Seq NP_001338104.1 PASSED. Final pValue = 3.21550009726979E-8

Seq NP_001338102.1 PASSED. Final pValue = 3.979098950120248E-8

Seq NP_001338099.1 PASSED. Final pValue = 3.979098950120248E-8

Seq NP_001338103.1 PASSED. Final pValue = 3.536366286184223E-8

Seq NP_001338100.1 PASSED. Final pValue = 3.979098950120248E-8

Seq NP_001338098.1 PASSED. Final pValue = 3.793922380371164E-8

Seq NP_001127845.1 PASSED. Final pValue = 3.793922380371164E-8

A.0.13. Diferentes con exponente -8

Cytosolic purine 5'-nucleotidase

Seq sp—P49902.1—5NTC_HUMAN PASSED. Final pValue = 3.793922380371164E-8

5'-nucleotidase, cytosolic II

Seq EAW49657.1 PASSED. Final pValue = 3.7277148333186105E-8

Seq EAW49656.1 PASSED. Final pValue = 3.7277148333186105E-8

epididymis secretory sperm binding protein

Seq ADO22399.1 PASSED. Final pValue = 3.793922380371164E-8

A.0.14. NP con exponente -10

ATP-dependent 6-phosphofructokinase, muscle type

Seq NP_001341677.1 PASSED. Final pValue = 2.547475093908801E-10

Seq NP_001341676.1 PASSED. Final pValue = 2.547475093908801E-10

Seq NP_001341674.1 PASSED. Final pValue = 3.1993154099066494E-10

A.0.15. XP con exponente -12

ATP-dependent 6-phosphofructokinase, platelet type

Seq XP_024303806.1 PASSED. Final pValue = 8.32111590220752E-12

Seq XP_006717512.1 PASSED. Final pValue = 9.309036324916555E-12

A.0.16. NP con exponente -12

ATP-dependent 6-phosphofruktokinase, platelet type

Seq NP_001310000.1 PASSED. Final pValue = 8.32111590220752E-12

Seq NP_001332873.1 PASSED. Final pValue = 8.32111590220752E-12

Seq NP_001310001.1 PASSED. Final pValue = 8.32111590220752E-12

Seq NP_001309996.1 PASSED. Final pValue = 8.948962769572405E-12

A.0.17. XP con exponente -13

ATP-dependent 6-phosphofruktokinase, liver type

Seq XP_016883861.1 PASSED. Final pValue = 1.7049700549276855E-13

Seq XP_016883858.1 PASSED. Final pValue = 1.8850577817101274E-13

Seq XP_016883859.1 PASSED. Final pValue = 1.7049700549276855E-13

Seq XP_016883857.1 PASSED. Final pValue = 2.8418938195606875E-13

Seq XP_024307853.1 PASSED. Final pValue = 2.9191490986984033E-13

Seq XP_006724074.1 PASSED. Final pValue = 2.8418938195606875E-13

Seq XP_005261194.1 PASSED. Final pValue = 1.7049700549276855E-13

A.0.18. NP con exponente -13

ATP-dependent 6-phosphofruktokinase, liver type

Seq NP_001002021.2 PASSED. Final pValue = 2.9191490986984033E-13

A.0.19. Diferentes con exponente -13

phosphofruktokinase, liver type

Seq EAX09439.1 PASSED. Final pValue = 2.9191490986984033E-13

A.0.20. NP con exponente -15

ATP-dependent 6-phosphofruktokinase, platelet type

Seq NP_001229268.1 PASSED. Final pValue = 3.3865502132049093E-15

A.0.21. XP con exponente -16**ATP-dependent 6-phosphofructokinase, muscle type**

Seq XP_024304788.1 PASSED. Final pValue = 1.1189882351887075E-16

A.0.22. XP con exponente -17**ATP-dependent 6-phosphofructokinase, muscle type**

Seq XP_024304790.1 PASSED. Final pValue = 8.88067739606044E-17

Seq XP_024304789.1 PASSED. Final pValue = 9.471854340563292E-17

Seq XP_016874958.1 PASSED. Final pValue = 6.914907396511316E-17

Seq XP_011536789.1 PASSED. Final pValue = 6.279814725953219E-17

Seq XP_005269033.1 PASSED. Final pValue = 8.927914714316217E-17

Seq XP_005269036.1 PASSED. Final pValue = 9.890708347609185E-17

Seq XP_005269031.1 PASSED. Final pValue = 8.927914714316217E-17

A.0.23. NP con exponente -17**ATP-dependent 6-phosphofructokinase, muscle type**

Seq NP_000280.1 PASSED. Final pValue = 8.88067739606044E-17

Seq NP_001341671.1 PASSED. Final pValue = 8.88067739606044E-17

Seq NP_001341673.1 PASSED. Final pValue = 8.88067739606044E-17

Seq NP_001341672.1 PASSED. Final pValue = 8.88067739606044E-17

Seq NP_001160159.1 PASSED. Final pValue = 8.88067739606044E-17

Seq NP_001160160.1 PASSED. Final pValue = 8.88067739606044E-17

Seq NP_001350548.1 PASSED. Final pValue = 6.054902261571896E-17

Seq NP_001160158.1 PASSED. Final pValue = 9.890708347609185E-17

Seq NP_001341669.1 PASSED. Final pValue = 5.44009418386607E-17

Seq NP_001341675.1 PASSED. Final pValue = 7.417485168749869E-17

Seq NP_001341670.1 PASSED. Final pValue = 7.86933778260881E-17

Seq NP_001341665.1 PASSED. Final pValue = 8.927914714316217E-17

Seq NP_001341664.1 PASSED. Final pValue = 8.927914714316217E-17

Seq NP_001341666.1 PASSED. Final pValue = 9.890708347609185E-17

Seq NP_001341668.1 PASSED. Final pValue = 9.890708347609185E-17

Seq NP_001341667.1 PASSED. Final pValue = 9.890708347609185E-17

A.0.24. Pdb con exponente -17**3'-hydroxy-N-methyl-(S)-coclaurine 4'-O-methyltransferase**

Seq pdb—4OMT—A PASSED. Final pValue = 8.88067739606044E-17

A.0.25. Diferentes con exponente -17**phosphofructokinase, muscle type**

Seq sp—P08237.2—PFKAM_HUMAN PASSED. Final pValue = 8.88067739606044E-17

Seq AAA82938.1 PASSED. Final pValue = 8.88067739606044E-17

Seq ALQ33710.1 PASSED. Final pValue = 6.054902261571896E-17

Seq ALQ33711.1 PASSED. Final pValue = 5.705384831708226E-17

Seq ALQ33709.1 PASSED. Final pValue = 8.88067739606044E-17

Seq EAW57975.1 PASSED. Final pValue = 8.88067739606044E-17

Seq EAW57973.1 PASSED. Final pValue = 6.054902261571896E-17

Seq EAW57974.1 PASSED. Final pValue = 8.88067739606044E-17

A.0.26. NP con exponente -19**ATP-dependent 6-phosphofructokinase, liver type**

Seq NP_002617.3 PASSED. Final pValue = 1.370194177255313E-19

Seq XP_005261192.1 PASSED. Final pValue = 1.3684043402506835E-19

A.0.27. Diferentes con exponente -19**ATP-dependent 6-phosphofructokinase, liver type**

Seq sp—P17858.6—PFKAL_HUMAN PASSED. Final pValue = 1.370194177255313E-19

Seq EAX09440.1 PASSED. Final pValue = 1.370194177255313E-19

A.0.28. Pdb con exponente -22**ATP-dependent 6-phosphofructokinase, platelet type**

Seq pdb—4XYJ—A PASSED. Final pValue = 3.4297395794338135E-22

A-H Seq pdb—4WL0—A PASSED. Final pValue = 4.031039527962828E-22;A-D

Seq pdb—4XYK—A PASSED. Final pValue = 3.4297395794338135E-22;AD

Seq pdb—4RH3—A PASSED. Final pValue = 4.048324653014709E-22;A-D

Seq pdb—4XZ2—A PASSED. Final pValue = 4.031039527962828E-22;A-D

Seq pdb—4U1R—A PASSED. Final pValue = 4.048324653014709E-22;A-D

A.0.29. NP con exponente -22

ATP-dependent 6-phosphofructokinase, platelet type

Seq NP_002618.1 PASSED. Final pValue = 3.826864806085976E-22

Seq NP_001309997.1 PASSED. Final pValue = 3.243140616242221E-22

A.0.30. XP con exponente -22

ATP-dependent 6-phosphofructokinase, platelet type

Seq XP_005252523.1 PASSED. Final pValue = 3.4398784209562515E-22

Seq XP_005252522.1 PASSED. Final pValue = 4.02094988605949E-22

A.0.31. Diferentes con exponente -22

ATP-dependent 6-phosphofructokinase, platelet type

Seq sp—Q01813.2—PFKAP_HUMAN PASSED. Final pValue = 3.826864806085976E-22

Seq EAW86497.1 PASSED. Final pValue = 4.622209507155341E-22

Seq EAW86496.1 PASSED. Final pValue = 3.7576810422708697E-22

Seq EAW86495.1 PASSED. Final pValue = 5.187340604601118E-22

Bibliografía

- [1] O Flores Herrera, E Rendón Huerta, H Riveros Rosas, A Sosa Peinado, E Vázquez Contreras, and I Velázquez López. La estructura y la visualización molecular de proteínas. *Mensaje bioquímico*, 29, 2005.
- [2] David L Nelson, Michael M Cox, and Albert L Lehninger. *Principles of biochemistry*. Freeman New York, 2008.
- [3] HN Curtis. *Textbook of biology*. Editorial Médica Panamericana, 2002.
- [4] Homero Sáenz-Suárez, Leonardo René Lareo, Carlos Oribio-Quinto, Juan Martínez-Mendoza, and Aura Chávez-Zobel. Predicción computacional de estructura terciaria de las proteínas humanas hsp27, ab-cristalina y hspb8. *Universitas Scientiarum*, 16(1):15–28, 2011.
- [5] Victoria Luque Guillén. *Estructura y propiedades de las proteínas*, 2009.
- [6] Luis A Chel Guerrero, Luis Corzo Ríos, and David A Betancur Ancona. Estructura y propiedades funcionales de proteínas de leguminosas. *Revista de la Universidad Autónoma de Yucatán*, pages 34–43, 2003.
- [7] O Martínez Augustin and E Martínez de Victoria. Proteínas y péptidos en nutrición enteral. *Nutrición Hospitalaria*, 21:01–14, 2006.
- [8] Ángel Gil and Fermín Sánchez de Medina Contreras. *Tratado de Nutrición: Bases fisiológicas y bioquímicas de la nutrición*. Acción Médica, 2005.
- [9] Micaela Anahí Santucho Cordoba. *Proteínas*. Monografía, 2014.
- [10] Jesús Merino Pérez and Maria José Noriega Borge. *Enzimas*. universidad de cantabria, 2011.
- [11] Trudy McKee and James R Mckee. Enzimas. In *Bioquímica: la base molecular de la vida*, chapter 6, pages 184–226. McGraw-Hill/Interamericana, 5 edition, 2003.
- [12] Yael Avissar, Jung Choi, Jean DeSaix, Vladimir Jurukovski, Robert Wise, Connie Rye, et al. Atp: Adenosine triphosphate. In *Biology*. OpenStax, 2018.

- [13] Elizabeth Lira Silva, Ricardo Jasso Chávez, and Juan Pablo Pardo Vázquez. Respuestas al problema bioquímico. *REB. Revista de educación bioquímica*, 33(2):68–72, 2014.
- [14] Ismael Lares-Asseff and Francisca Trujillo-Jiménez. La farmacogenética y su importancia en la clínica. *Gaceta medica de Mexico*, 137(3), 2001.
- [15] Claude Denson Pepper. National center for biotechnology information (ncbi). [citado 17 marzo 2020]. Disponible en: <https://www.ncbi.nlm.nih.gov/>, 1988.
- [16] Janelle B Leuthaeuser, John H Morris, Angela F Harper, Thomas E Ferrin, Patricia C Babbitt, and Jacquelyn S Fetrow. Dasp3: identification of protein sequences belonging to functionally relevant groups. *BMC bioinformatics*, 17(1):458, 2016.
- [17] Richard D Taylor, Philip J Jewsbury, and Jonathan W Essex. A review of protein-small molecule docking methods. *Journal of computer-aided molecular design*, 16(3):151–166, 2002.
- [18] Ruifeng Qi, Evans Boateng Sarbeng, Qun Liu, Katherine Quynh Le, Xiping Xu, Hongya Xu, Jiao Yang, Jennifer Li Wong, Christina Vorvis, Wayne A Hendrickson, et al. Allosteric opening of the polypeptide-binding site when an hsp70 binds atp. *Nature structural & molecular biology*, 20(7):900, 2013.
- [19] Marta Acebrón García de Eulate. *Fragment based ligand discovery on Focal Adhesion Kinase*. PhD thesis, Universidad Autónoma de Madrid, 2018.
- [20] Alejandro Reyes. *ANÁLISIS FUNCIONAL DE MUTANTES PUNTUALES EN SITIOS ALOSTÉRICOS ENDO Y EXOFACIALES EN EL TRANSPORTADOR DE GLUCOSA GLUT1*. PhD thesis, Universidad de Concepción, 2009.
- [21] Joerg Klepper and Baerbel Leiendecker. Glut1 deficiency syndrome–2007 update. *Developmental Medicine & Child Neurology*, 49(9):707–716, 2007.
- [22] Rafael Lahoz-Beltrá. *Bioinformática: Simulación, vida artificial e inteligencia artificial*. Ediciones Díaz de Santos, 2010.
- [23] Vincent Le Guilloux, Peter Schmidtke, and Pierre Tuffery. Fpocket: an open source platform for ligand pocket detection. *BMC bioinformatics*, 10(1):168, 2009.
- [24] Joe G Greener, Ioannis Filippis, and Michael JE Sternberg. Predicting protein dynamics and allostery using multi-protein atomic distance constraints. *Structure*, 25(3):546–558, 2017.
- [25] Declan Clarke, Anurag Sethi, Shantao Li, Sushant Kumar, Richard WF Chang, Jieming Chen, and Mark Gerstein. Identifying allosteric hotspots with dynamics: Application to inter-and intra-species conservation. *Structure*, 24(5):826–837, 2016.

- [26] Leslie B Poole and Kimberly J Nelson. Distribution and features of the six classes of peroxiredoxins. *Molecules and cells*, 39(1):53, 2016.
- [27] Jacquelyn S Fetrow. Active site profiling to identify protein functional sites in sequences and structures using the deacon active site profiler (dasp). *Current protocols in bioinformatics*, 14(1):8–10, 2006.
- [28] Laura Soito, Chris Williamson, Stacy T Knutson, Jacquelyn S Fetrow, Leslie B Poole, and Kimberly J Nelson. Prex: Peroxiredoxin classification index, a database of subfamily assignments across the diverse peroxiredoxin family. *Nucleic acids research*, 39(suppl_1):D332–D337, 2011.
- [29] Ryan G Huff, Ersin Bayram, Huan Tan, Stacy T Knutson, Michael H Knaggs, Allen B Richon, Peter Santago, and Jacquelyn S Fetrow. Chemical and structural diversity in cyclooxygenase protein active sites. *Chemistry & biodiversity*, 2(11):1533–1552, 2005.
- [30] Maritza Rodríguez Charry. Identificación automática de sitios alostéricos en proteínas mediante la herramienta deacon active site profiler (dasp3). Trabajo integral de grado, Antonio Nariño, 2019.
- [31] Michael McCarthy, Priyanka Prakash, and Alemayehu A Gorfe. Computational allosteric ligand binding site identification on ras proteins. *Acta biochimica et biophysica Sinica*, 48(1):3–10, 2016.
- [32] Alexander L Perryman, Daniel N Santiago, Stefano Forli, Diogo Santos-Martins, and Arthur J Olson. Virtual screening with autodock vina and the common pharmacophore engine of a low diversity library of fragments and hits against the three allosteric sites of hiv integrase: participation in the sampl4 protein–ligand binding challenge. *Journal of computer-aided molecular design*, 28(4):429–441, 2014.
- [33] Carlos Roca, Carlos Requena, Víctor Sebastián-Pérez, Sony Malhotra, Chris Radoux, Concepción Pérez, Ana Martínez, Juan Antonio Paez, Tom L Blundell, and Nuria E Campillo. Identification of new allosteric sites and modulators of ache through computational and experimental tools. *Journal of enzyme inhibition and medicinal chemistry*, 33(1):1034–1047, 2018.
- [34] Garrett M Morris, Ruth Huey, William Lindstrom, Michel F Sanner, Richard K Belew, David S Goodsell, and Arthur J Olson. Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, 30(16):2785–2791, 2009.
- [35] Marcelo Adrian Marti and Adrian Turjanski. La bioinformática estructural o la realidad virtual de los medicamentos. *Química Viva*, 2009.

- [36] Ruth Nussinov and Chung-Jung Tsai. The different ways through which specificity works in orthosteric and allosteric drugs. *Current pharmaceutical design*, 18(9):1311–1316, 2012.
- [37] Shaoyong Lu, Shuai Li, and Jian Zhang. Harnessing allostery: a novel approach to drug discovery. *Medicinal research reviews*, 34(6):1242–1285, 2014.
- [38] Alejandro Panjkovich and Xavier Daura. Exploiting protein flexibility to predict the location of allosteric sites. *BMC bioinformatics*, 13(1):273, 2012.
- [39] Fernanda Saldívar-González, Fernando D Prieto-Martínez, and José L Medina-Franco. Descubrimiento y desarrollo de fármacos: un enfoque computacional. *Educación química*, 28(1):51–58, 2017.
- [40] Carlos Roca Magadán. *Estrategias computacionales en el desarrollo de neurofármacos: una tecnología de éxito*. PhD thesis, Universidad Complutense de Madrid, 2018.
- [41] Shaoyong Lu, Mingfei Ji, Duan Ni, and Jian Zhang. Discovery of hidden allosteric sites as novel targets for allosteric drug design. *Drug discovery today*, 23(2):359–365, 2018.
- [42] Ruth Nussinov and Chung-Jung Tsai. The design of covalent allosteric drugs. *Annual review of pharmacology and toxicology*, 55:249–267, 2015.
- [43] Gerard J Tortora and Bryan Derrickson. *Principios de anatomía y fisiología*. Médica Panamericana,, 2013.
- [44] W Sperl, P Ješina, J Zeman, JA Mayr, L Demeirleir, Rudy VanCoster, A Pickova, H Hansikova, H Houšt'ková, Z Krejčík, et al. Deficiency of mitochondrial atp synthase of nuclear genetic origin. *Neuromuscular Disorders*, 16(12):821–829, 2006.
- [45] Paolo Cremonesi. *L'ambiente acquoso per il trattamento di opere policrome*. il prato publishing house srl, 2012.
- [46] Marcela Ayala Aceves. Enzimas:¿ qué son y cómo funcionan? *Revista Digital Universitaria*, 15(12), 2017.
- [47] Frank Bradley Armstrong, Frank Bradley Armstrong, and Thomas Peter Bennett. *Bioquímica*. Reverte, 1982.
- [48] Laurence A Cole. *Biology of Life: Biochemistry, Physiology and Philosophy*. Academic Press, 2016.
- [49] Dominio público NEUROtiker. Estructura del trifosfato de adenosina (atp), protonado. [citado 15 julio 2020]. Disponible en: https://upload.wikimedia.org/wikipedia/commons/3/31/Adenosintriphosphat_protoniert.svg, 2007.

- [50] Greelane. El atp en el metabolismo. [citado 15 julio 2020]. Disponible en: <https://www.greelane.com/es/ciencia-tecnolog%C3%ADa-matem%C3%A1ticaciencia/phosphorylation-definition-4140732/>, 2019.
- [51] Simon Orozco Arias and Jeferson Arango López. Aplicación de la inteligencia artificial en la bioinformática, avances, definiciones y herramientas. *UGCiencia*, 22(1):159–171, 2016.
- [52] Dinler A Antunes, Didier Devaurs, and Lydia E Kavraki. Understanding the challenges of protein flexibility in drug design. *Expert opinion on drug discovery*, 10(12):1301–1313, 2015.
- [53] Jacob D Durrant and J Andrew McCammon. Molecular dynamics simulations and drug discovery. *BMC biology*, 9(1):71, 2011.
- [54] Morris. Adt / autodocktools. [citado 05 de mayo 2020]. Disponible en: <http://autodock.scripps.edu>, 2007.
- [55] Ruth Huey, Garrett M Morris, and Stefano Forli. Using autodock 4 and autodock vina with autodocktools: A tutorial. *The Scripps Research Institute Molecular Graphics Laboratory*, 2012.
- [56] Ambrish Roy, Alper Kucukural, and Yang Zhang. I-tasser: a unified platform for automated protein structure and function prediction. *Nature protocols*, 5(4):725, 2010.
- [57] John Moult, Jan T Pedersen, Richard Judson, and Krzysztof Fidelis. A large-scale experiment to assess protein structure prediction methods. *Proteins: Structure, Function, and Bioinformatics*, 23(3):ii–iv, 1995.
- [58] Sitao Wu, Jeffrey Skolnick, and Yang Zhang. Ab initio modeling of small proteins by iterative tasser simulations. *BMC biology*, 5(1):17, 2007.
- [59] Nicolas Gue. Swiss-pdbviewer. [citado 16 de junio 2020]. Disponible en: <https://spdbv.vital-it.ch/>.
- [60] Visualización e Informática (RBVI) Recurso para Biocomputación. Ucsf chimera. [citado 16 de junio 2020]. Disponible en: <https://www.cgl.ucsf.edu/chimera/>, 2019.
- [61] Version 1.XX. Software Libre. Avogadro: an open-source molecular builder and visualization tool. [citado 10 de junio 2020]. Disponible en: <https://avogadro.cc/>, 2006.
- [62] Celia Torres Quezada, Patricia Varela Gangas, María Verónica Frías, and Patricio Flores-Morales. Implementación de avogadro como visualizador y constructor de moléculas para alumnos de primer año de odontología en la asignatura química general y orgánica. *Educación química*, 28(1):22–29, 2017.

- [63] Sebastian Raschka. Molecular docking, estimating free energies of binding, and autodock's semi-empirical force field, 2014.
- [64] Instituto Europeo de Bioinformática (EBI), Instituto Suizo de Bioinformática (SIB), and Protein Information Resource (PIR). The UniProt Knowledgebase (UniProtKB). [citado 17 marzo 2020]. Disponible en: <https://www.uniprot.org/>, 2002.
- [65] Karin Walldén and Pär Nordlund. Structural basis for the allosteric regulation and substrate recognition of human cytosolic 5-nucleotidase ii. *Journal of molecular biology*, 408(4):684–696, 2011.
- [66] AS Bretonnet, LP Jordheim, C Dumontet, and JM Lancelin. Regulation and activity of cytosolic 5-nucleotidase ii: A bifunctional allosteric enzyme of the haloacid dehalogenase superfamily involved in cellular metabolism. *FEBS letters*, 579(16):3363–3368, 2005.
- [67] Marco Kloos, Antje Brüser, Jürgen Kirchberger, Torsten Schöneberg, and Norbert Sträter. Crystal structure of human platelet phosphofructokinase-1 locked in an activated conformation. *Biochemical Journal*, 469(3):421–432, 2015.
- [68] Wen Yi, Peter M Clark, Daniel E Mason, Marie C Keenan, Collin Hill, William A Goddard, Eric C Peters, Edward M Driggers, and Linda C Hsieh-Wilson. Phosphofructokinase 1 glycosylation regulates cell growth and metabolism. *Science*, 337(6097):975–980, 2012.
- [69] Lukasz Wieteska, Saeid Shahidi, and Anastasia Zhuravleva. Allosteric fine-tuning of the conformational equilibrium poises the chaperone bip for post-translational regulation. *Elife*, 6:e29430, 2017.
- [70] Jiao Yang, Melesse Nune, Yinong Zong, Lei Zhou, and Qinglian Liu. Close and allosteric opening of the polypeptide-binding site in a human hsp70 chaperone bip. *Structure*, 23(12):2191–2203, 2015.
- [71] Edgar Meyer and Walter Hamilton. PDB (Protein Data Bank). [citado 17 marzo 2020]. Disponible en: <https://www.rcsb.org/>, 1971.
- [72] Akemi Irie, Akira Yamauchi, Keiichi Kontani, Minoru Kihara, Dage Liu, Yukako Shirato, Masako Seki, Nozomu Nishi, Takanori Nakamura, Hiroyasu Yokomise, et al. Galectin-9 as a prognostic factor with antimetastatic potential in breast cancer. *Clinical cancer research*, 11(8):2962–2968, 2005.
- [73] Yumiko Kashio, Kazuhiro Nakamura, Mohammad J Abedin, Masako Seki, Nozomu Nishi, Naoko Yoshida, Takanori Nakamura, and Mitsuomi Hirashima. Galectin-9 induces apoptosis through the calcium-calpain-caspase-1 pathway. *The Journal of Immunology*, 170(7):3631–3636, 2003.

- [74] Yuka Tsuboi, Hiroko Abe, Ryusuke Nakagawa, Souichi Oomizu, Kota Watanabe, Nozomu Nishi, Takanori Nakamura, Akira Yamauchi, and Mitsuomi Hirashima. Galectin-9 protects mice from the shwartzman reaction by attracting prostaglandin e2-producing polymorphonuclear leukocytes. *Clinical immunology*, 124(2):221–233, 2007.
- [75] Alexander H Stegh, Hyunggee Kim, Robert M Bachoo, Kristin L Forloney, Jean Zhang, Harald Schulze, Kevin Park, Gregory J Hannon, Junying Yuan, David N Louis, et al. Bcl2l12 inhibits post-mitochondrial apoptosis signaling in glioblastoma. *Genes & development*, 21(1):98–111, 2007.
- [76] Xiutian Guo, Mao-Gang Li, Shan-Shan Li, Feng-Hua Liu, Zhan-Ju Liu, and Ping-Chang Yang. Tumor necrosis factor suppresses interleukin 10 in peripheral b cells via upregulating bcl2-like protein 12 in patients with inflammatory bowel disease. *Cell biochemistry and function*, 35(2):77–82, 2017.
- [77] Zhi-Qiang Liu, Ying Feng, Li-Hua Mo, Xian-Hai Zeng, Jiang-Qi Liu, Rui-Di Xie, Zhi-Gang Liu, Ping-Chang Yang, Guang-Ji Zhang, and Shan-Dong Wu. Bcl2-like protein 12 plays a critical role in development of airway allergy through inducing aberrant th2 polarization. *Journal of Allergy and Clinical Immunology*, 143(1):427–430, 2019.
- [78] Hongyan Li, Dongbai Yang, and Zhifeng Tang. Bcl2 like protein-12 suppresses foxp3+ regulatory t cells in patients with rheumatoid arthritis. *American Journal of Translational Research*, 11(5):3048, 2019.
- [79] Ekaterina Dik, Adi Naamati, Hadar Asraf, Norbert Lehming, and Ophry Pines. Human fumarate hydratase is dual localized by an alternative transcription initiation mechanism. *Traffic*, 17(7):720–732, 2016.
- [80] Ohad Yogev, Adi Naamati, and Ophry Pines. Fumarase: a paradigm of dual targeting and dual localized functions. *The FEBS journal*, 278(22):4230–4242, 2011.
- [81] Julie Adam, Ming Yang, Christina Bauerschmidt, Mitsuhiro Kitagawa, Linda O’Flaherty, Pratheesh Maheswaran, Gizem Özkan, Natasha Sahgal, Dilair Baban, Keiko Kato, et al. A role for cytosolic fumarate hydratase in urea cycle metabolism and renal neoplasia. *Cell reports*, 3(5):1440–1448, 2013.
- [82] IP Tomlinson, NA Alam, AJ Rowan, E Barclay, EE Jaeger, D Kelsell, I Leigh, P Gorman, H Lamlum, S Rahman, et al. Multiple leiomyoma consortium: Germline mutations in fh predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nat Genet*, 30(4):406–410, 2002.
- [83] Mariana A Ajalla Aleixo, Victor L Rangel, Joane K Rustiguel, Ricardo AP de Pádua, and Maria Cristina Nonato. Structural, biochemical and biophysical characterization of recombinant human fumarate hydratase. *The FEBS journal*, 286(10):1925–1940, 2019.

-
- [84] Margarita Velásquez, Juan Drosos, Carlos Gueto, Johana Márquez, and Ricardo Vivas-Reyes. Método acoplado autodock-pm6 para seleccionar la mejor pose en estudios de acoplamiento molecular. *Revista Colombiana de Química*, 42(1), 2013.