

DISEÑO *IN SILICO* DE PÉPTIDOS CON ACTIVIDAD INHIBITORIA SOBRE  
PROTEÍNAS RAS USANDO ALGORITMOS DE EVOLUCIÓN MOLECULAR

SERGIO ANDRES PINTO RUBIO

UNIVERSIDAD INDUSTRIAL DE SANTANADER  
FACULTAD DE INGENIERÍAS FISICOMECÁNICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA  
2015

DISEÑO *IN SILICO* DE PÉPTIDOS CON ACTIVIDAD INHIBITORIA SOBRE  
PROTEÍNAS RAS USANDO ALGORITMOS DE EVOLUCIÓN MOLECULAR

SERGIO ANDRES PINTO RUBIO

Trabajo de grado presentado para optar por el título de:  
Ingeniero de Sistemas

DIRECTOR:

Ph.D Ingeniería eléctrica y computacional  
Henry Arguello Fuentes

TUTOR:

Ph.D Bioquímica y biología molecular  
Rodrigo Gonzalo Torres Saez

UNIVERSIDAD INDUSTRIAL DE SANTANADER  
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA  
2015

## CONTENIDO

	pág.
INTRODUCCIÓN	12
1. MARCO DE REFERENCIA	14
1.1 DISEÑO DE FÁRMACOS IN SILICO	14
1.2 PÉPTIDOS	15
1.2.1 Péptidos antimicrobianos - PAMs	16
1.2.2 Clasificación	17
1.2.2.1 Según su estructura	17
1.2.2.2 Según actividad biológica	19
1.3 DISEÑO DE PÉPTIDOS	21
1.4 ALINEAMIENTO DE MÚLTIPLES SECUENCIAS	21
1.4.1 Pareado	22
1.4.2 Múltiple	22
1.5 ALGORITMO EVOLUTIVO	24
1.6 DOCKING MOLECULAR	27
2. PLANTEAMIENTO DEL PROBLEMA	31
3. PROCEDIMIENTO	33

3.1	CONJUNTO DE DATOS INICIALES	33
3.2	ALINEAMIENTO DE DATOS	33
3.3	REGIÓN CONSERVADA	34
3.4	SISTEMA DE PUNTUACIÓN	34
3.5	ALGORITMO EVOLUTIVO	35
3.6	DOCKING	37
4.	RESULTADOS Y DISCUSIÓN	41
5.	CONCLUSIONES	50
6.	RECOMENDACIONES	51
	REFERENCIAS BIBLIOGRAFICAS	52
	BIBLIOGRAFÍA	56

## LISTA FIGURAS

	Pág.
Figura 1. Proteína Ras con GTP analog Y Ras con GDP.	13
Figura 2. La estructura general de un alfa-aminoácido.	16
Figura 3. Insulina PDB en 3D.	16
Figura 4. Clasificación de péptidos según su estructura.	17
Figura 5. Diversidad estructural de péptidos antimicrobianos.	18
Figura 6. Clasificación de péptidos según actividad biológica.	20
Figura 7. Secuencias sin alineamiento.	22
Figura 8. Alineamiento Pareado.	22
Figura 9. Alineamiento múltiple.	23
Figura 10. Matriz de castigos.	23
Figura 11. Alineamiento y región conservada en Jview.	24
Figura 12. Esquema general del algoritmo evolutivo en diagrama de flujo.	25
Figura 13. Ejemplo de reproducción.	26
Figura 14. Ejemplo de mutación.	26
Figura 15. Gráfica de función actitud vs. Tiempo.	27
Figura 16. Acoplamiento molecular (docking).	27
Figura 17. Planteamiento del problema.	31
Figura 18. Resultado a la consulta en AntiCP.	35

Figura 19. Grid box en Autodock tools.	38
Figura 20. Algoritmo evolutivo v8.	44
Figura 21. Valores de distancias y energía obtenidos de Autodock vina	45
Figura 22. Resultado en 3D del acople entre receptor y ligando.	46
Figura 23. Comparación del docking entre péptidos de la población inicial contra la población final.	47

## LISTA DE TABLAS

	Pág.
Tabla 1. Funciones Autodock vina.	39
Tabla 2. Péptidos extraídos del APD.	41
Tabla 3. Péptidos alineados y recortados en la región conservada.	42
Tabla 4. Comparación de puntuaciones iniciales vs. Finales.	43
Tabla 5. Resultado del algoritmo evolutivo molecular.	44

## RESUMEN

**TÍTULO:** DISEÑO DE PÉPTIDOS CON ACTIVIDAD INHIBITORIA SOBRE PROTEÍNAS RAS USANDO ALGORITMOS DE EVOLUCIÓN MOLECULAR<sup>1</sup>

**AUTOR:** Sergio Andres Pinto Rubio<sup>2</sup>

**PALABRAS CLAVE:** Proteínas Ras, Péptidos anticancerígenos, Péptidos antitumorales, Algoritmos evolutivos, RMSD.

### DESCRIPCIÓN:

El cáncer es una enfermedad de gran relevancia en el siglo XXI. Además, esta patología es conocida por tener una de las tasas más altas de mortalidad en todo el mundo. Se conocen alrededor de 200 tipos diferentes de esta patología; por esta razón, se han adelantado muchas investigaciones para intentar hallar nuevos fármacos que logren frenar esta enfermedad. Gracias a estos estudios, han surgido nuevas áreas de investigación, como las técnicas de biología computacional. En este marco, se han usado diversos biomarcadores proteicos, como blanco para el diseño de moléculas anticancerígenas. En este trabajo, se lleva a cabo el diseño in silico de péptidos con actividad inhibitoria sobre proteínas Ras, las cuales son las proteínas capaces de controlar fenómenos como el crecimiento celular, la diferenciación y la apoptosis, y que también actúan como interruptores en la proliferación celular. Además, estas proteínas son causantes de aproximadamente 20% a 30% de los diferentes tipos de cáncer. Para resolver este problema, se seleccionaron diferentes péptidos de la base de datos de péptidos antimicrobianos (APD) con actividad anticancerígena. Luego, por medio de técnicas de alineamiento de secuencias, se determinó los dominios conservados de los péptidos, los cuales se usaron para el diseño de los péptidos anticancerígenos con actividad inhibitoria sobre las proteínas Ras. Finalmente, se obtuvo por medio de algoritmos evolutivos, nuevos y mejores péptidos, los cuales fueron evaluados en su capacidad de unión a proteínas Ras por medio de la técnica de docking molecular semiflexible.

---

<sup>1</sup> Trabajo de grado

<sup>2</sup> Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e informática.

Director: Ph.D Henry Arguello Fuentes. Tutor: Ph.D. Rodrigo Gonzalo Torres Saez

## ABSTRACT

TITLE: IN SILICO DESIGN OF PEPTIDES WITH INHIBITORY ACTIVITY ON RAS PROTEINS USING MOLECULAR EVOLUTIONARY ALGORITHMS<sup>1</sup>

AUTHOR: Sergio Andres Pinto Rubio<sup>2</sup>

KEY WORDS: Ras proteins, anticancer peptides, antitumor peptides, evolutionary algorithms, RMSD

### DESCRIPTION:

Cancer is a disease with great relevance in the 21st century. Moreover, this pathology is known for having one of the most high mortality rates thorough the world. Around 200 different types of this pathology are known; for this reason, several investigations have been performed in order to find new pharmaceuticals that achieve to stop this disease. Thanks to these studies, new research areas such as computational biology techniques have appeared. In this context, several protein biomarkers have been used as target for design of anticancer molecules. In this work, we carried out in silico design of peptides with inhibitory activity on Ras proteins, which are proteins able to control phenomena as cell growth, differentiation and apoptosis, and that also act as switches on cell proliferation. Moreover, these proteins are causing of approximately from 20% to 30% of different types of cancer. In order to resolve this problem, we selected different peptides from antimicrobial peptide database (APD) with anticancer activity. Then, through sequence alignment, we determined conserved domains of the peptides, which were used for design of anticancer peptides with inhibitory activity on Ras proteins. Finally, by means of evolutionary algorithms, new and better peptides were obtained, which were evaluated in both their ability to bind Ras proteins through the semiflexible molecular docking technique.

---

<sup>1</sup> Bachelor thesis

<sup>2</sup> Faculty of Physical-Mechanical Engineering. School of Systems and Computing Engineering.  
Director: Ph.D Henry Arguello Fuentes. Advisor: Ph.D. Rodrigo Gonzalo Torres Saez

## INTRODUCCIÓN

Actualmente, una de las mayores causas de muerte en el mundo es el cáncer, siendo considerada esta patología como la enfermedad del siglo XXI. Se conocen más de 200 tipos diferentes de cáncer, siendo los más conocidos: cáncer pulmonar, gástrico, hepático, colorrectal, mamario y cervicouterino, entre otros [1]. En 2008, se reportaron cerca de 7.8 millones de muertes en el mundo, de las cuales un 13% correspondieron a cánceres, estimándose que para el 2030 se producirá un aumento a 13.1 millones o más [2]. A pesar de los incontables avances que se han obtenido en el conocimiento del cáncer y en los tratamientos para combatir esta enfermedad, hoy en día el cáncer sigue siendo uno de los mayores retos a superar para las ciencias biomédicas.

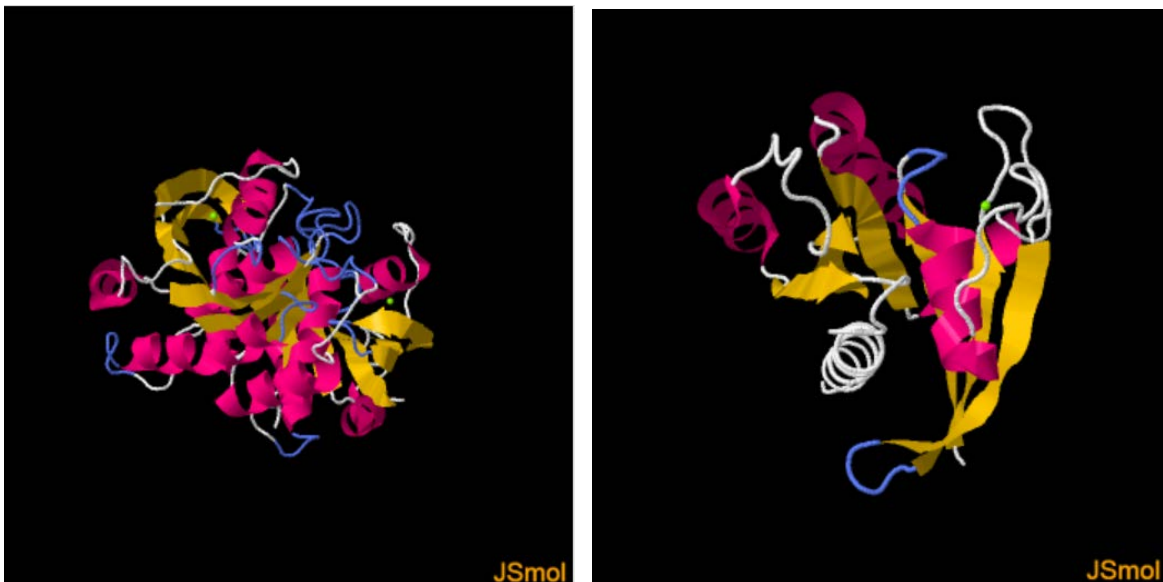
El cáncer es producto de la proliferación o multiplicación sin control de las células, con la capacidad de afectar a tejidos cercanos y distantes del sitio donde se produce el tumor [2]. Esta situación acarrea alteraciones en la fisiología celular, lo cual genera un crecimiento descontrolado de las células por: autoabastecimiento de señales de crecimiento celular, insensibilidad a los factores inhibitorios del crecimiento, evasión de la apoptosis (muerte celular programada) y potencial de replicación ilimitado [3].

Actualmente, existen diferentes tipos de proteínas que son capaces de combatir el cáncer, como es el caso de la proteína SIRT1, conocida como proteína de la longevidad, la cual tiene efectos anticancerígenos, según un estudio publicado recientemente por la Universidad Rockefeller [4]. Uno de los problemas que aumenta el riesgo de cáncer es la inflamación celular; uno de los biomarcadores de inflamación es la Proteína C reactiva (PCR), esta proteína aumenta sus niveles en el plasma sanguíneo como respuesta a una inflamación. Es por ello, que se dice que la Proteína C reactiva es una proteína de fase aguda. El papel de la inflamación en el cáncer, no es bien conocido. Algunos órganos del cuerpo muestran un mayor riesgo de cáncer cuando están inflamados de forma crónica, lo cual afecta la comunicación celular y la señalización entre ellas.

Una de las proteínas más importantes en el proceso de señalización celular son las proteínas G- regulatorias. Estas proteínas poseen un papel importante en la transmisión de señales de las células eucariotas, es decir, las células que tienen su información genética encerrada dentro de una doble membrana, siendo las proteínas Ras un tipo de proteína G monomérica. Dentro de la familia de las proteínas G-regulatorias, se encuentran las proteínas Ras. Las proteínas Ras son un conjunto de proteínas reguladoras con capacidad de controlar fenómenos tales

como el crecimiento celular, la diferenciación y la apoptosis, y funcionan como interruptores moleculares encendiéndose cuando están fosforilados (Ras-GTP) y apagándose al desfosforilarse (Ras-GDP) (ver figura 1). Esta transformación entre las dos formas produce cambios conformacionales de dos regiones flexibles llamadas interruptor I, e interruptor II, las cuales pueden por efectos de mutaciones causar que el mecanismo se quede en la posición "encendido", activando de forma continua la proliferación celular [5]. Por esta razón se requiere del diseño de nuevos fármacos, que permitan controlar la regulación de las proteínas Ras, especialmente en aquellos tejidos, cuyas células aumentan de forma descontrolada su proliferación celular, lo cual puede generar alteraciones celulares como tumores, carcinomas, etc. Una de estas alternativas es el uso de péptidos con actividad inhibitoria sobre proteínas Ras. Para ello es necesario diseñar estos péptidos mediante técnicas de biología computacional. En este sentido, el uso de algoritmos de evolución molecular permitiría obtener péptidos homólogos con posible actividad inhibitoria de proteínas Ras. Estos péptidos se generarían mediante recombinación y selección de nuevas secuencias, con el objeto de buscar un nuevo péptido inhibidor. Finalmente, el diseño de estos péptidos permitiría tener nuevas alternativas terapéuticas contra el cáncer.

Figura 1. Izq. Proteína Ras con GTP analog, Der. Ras con GDP, ambas en su estructura secundaria.



Fuente: Pai, E. F., Krengel, U., Petsko, G. A., Goody, R. S., Kabsch, W., & Wittinghofer, A. (1990). Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *The EMBO Journal*, 9(8), 2351–2359.

## 1. MARCO DE REFERENCIA

### 1.1 DISEÑO DE FÁRMACOS IN SILICO

El diseño de fármacos es un proceso ineficiente y costoso. Estos deben pasar por una serie de eventos:

Estudiar a fondo la enfermedad para poder identificar y validar el blanco (Target en inglés) (duración de 2 a 5 años).

Búsqueda del compuesto efectivo contra el blanco, (duración de 2 a 5 años).

Realización de procesos como la formulación y el escalado para llegar a las pruebas clínicas en humanos (duración de 2 a 10 años).

Aprobación de la FDA (Food and Drug Administration) (duración de 2 a 3 años).

En resumen, un fármaco demora de 8 a 23 años y su valor oscila entre 108 – 109 Dólares [6]. Por otro lado, se encuentra el diseño de fármaco in silico, in vitro, in situ, in vivo, que son más eficientes y menos costosos.

En este estudio, se usó el diseño de fármacos in silico, el cual se realiza por medio de programas computacionales, que permiten visualizar la estructura de la macromolécula previamente registrada en bases de datos públicas, y obtenidas por medio de difracción de rayos X o la espectroscopia de RMN, como es el caso del Protein Data Bank (PDB) [7]. Todo esto con el fin de hacer diferentes cálculos y predicciones de afinidad entre el fármaco y el sitio de unión de la proteína. Existen distintas técnicas computacionales, como lo son el docking molecular semiflexible, el diseño de novo o el cribado virtual de alta eficacia, que son utilizados con el objetivo de optimizar moléculas candidatas a fármacos [8].

El mayor número de aplicaciones exitosas diseñadas por ordenador, han sido las relacionadas en el campo del tratamiento del síndrome de la inmunodeficiencia adquirida (SIDA). Otras aplicaciones, se extienden a enfermedades como el cáncer, la enfermedad de alzheimer, la hipertensión y la artritis, entre muchas otras. Un ejemplo de compuesto diseñado en ordenador, es el Relenza, utilizado en el tratamiento contra el virus de la influenza (el causante de la gripe), el cual se aprobó con el nombre de Zanamivir [6].

## 1.2 PÉPTIDOS

Los péptidos, están conformados por aminoácidos unidos entre sí, mediante enlaces covalentes, los cuales presentan en sus extremos grupos amino y carboxilo libres, con un tamaño inferior a 30-50 aminoácidos. El origen de los péptidos es en su mayoría ribosomal, y también se encuentran los no ribosomales, los cuales son el resultado de rutas metabólicas simples, que cumplen con diferentes funciones biológicas, inmunológicas, estructurales, hormonales, de defensa, entre otras. Por otro lado, se han desarrollado péptidos de origen sintético, que por lo general suelen ser cortos, cíclicos y con residuos modificados, como D- aminoácidos o aminoácidos no naturales [9].

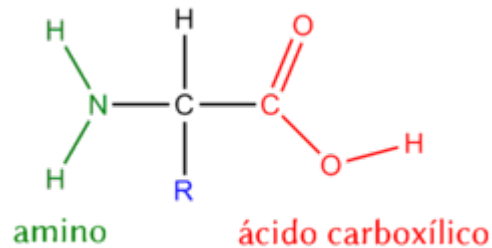
La estructura de los aminoácidos se conforma por un carbono  $\alpha$ , al que están unidos: un grupo carboxilo, el grupo amino, un átomo de hidrogeno y una cadena lateral que diferencia a cada aminoácido de los demás por su alta diversidad (ver figura 2 y figura 3). Por poseer un carbono  $\alpha$ , los aminoácidos poseen una imagen especular y por ende un isómero óptico, por lo cual los aminoácidos pueden ser L o D, dependiendo de la disposición espacial de los cuatro constituyentes. La mayoría de los aminoácidos naturales son de tipo L.

Los aminoácidos están divididos principalmente en dos grupos, polares y no polares, basados en características fisicoquímicas de las cadenas laterales.

Los aminoácidos polares se dividen en: Alifáticos (glicina, alanina, valina, leucina, isoleucina, prolina y metionina) y aromáticos (fenilalanina, tirosina, triptófano). Los aminoácidos no polares se dividen en: Neutros (serina, treonina, cisteína, asparagina, glutamina) y cargados, estos a su vez pueden ser ácidos (ácido glutámico y ácido aspártico) o básicos (lisina, asparagina e histidina) [10].

En cuanto a su disposición espacial, los péptidos pueden adquirir varias conformaciones dependiendo de su secuencia. Las dos disposiciones más frecuentes en cuanto a la estructura secundaria son las hélices  $\alpha$  y las láminas  $\beta$ . Las hélices se forman cuando el péptido se enrolla sobre sí mismo, mediante puentes de hidrógeno entre el grupo CO del enlace peptídico del primer residuo y el grupo NH del enlace peptídico del cuarto residuo, de esta manera, las hélices  $\alpha$  completan una vuelta alrededor de su eje cada 3,6 aminoácidos. Por otra parte, las láminas  $\beta$  están conformadas por dos o más hebras  $\beta$ , que se encuentran una junto a la otra, mediante enlaces de hidrógeno entre los grupos CO y NH, formando una estructura estable. Además de estas dos conformaciones, también se presentan con menor frecuencia giros o bucles [11].

Figura 2. La estructura general de un alfa-aminoácido se establece por la presencia de un carbono central (alfa) unido a un grupo carboxilo (rojo en la figura), un grupo amino (verde), un hidrógeno (en negro) y la cadena lateral (azul).



Fuente: General structure of amino acids - AA-structure-es, [en línea], Wikimedia, Basquetteur, 3 April 2014, [citado 15 de Mayo de 2015], Disponible en: [https://commons.wikimedia.org/wiki/File:General\\_structure\\_of\\_aminoacids\\_-\\_AA-structure-es.png](https://commons.wikimedia.org/wiki/File:General_structure_of_aminoacids_-_AA-structure-es.png)

Figura 3. Insulina PDB (péptido), en su estructura secundaria.



Fuente: Hua, Q. X. . G. S. N. . C. R. E. . H. J. A. . F. B. H. . W. M. A. (1995). Structure of a protein in a kinetic trap. *Nat.Struct.Biol.*, 2, 129–138. doi:7749917

1.2.1 Péptidos antimicrobianos – PAM's Los péptidos antimicrobianos o PAMs son parte esencial del sistema inmune de las bacterias, hongos, plantas, insectos y vertebrados, y para los invertebrados su principal defensa es la inmunología

innata. Estos péptidos se encuentran en casi todos los seres vivos, desde los procariontes, como las bacterianas, y los eucariotas, como: mamíferos, anfibios, insectos y plantas. Siendo estos dos últimos quienes aportan cerca del 85% de los péptidos actualmente conocidos, los cuales son sintetizados a bajo costo metabólico y pueden almacenarse en grandes cantidades en caso de una infección [7].

Estos péptidos conforman una familia de sustancias polifacéticas con complejos mecanismos de acción que se relacionan en la interacción con el patógeno a través de su membrana, afectando blancos internos, e interactuando con el huésped con funciones inmunomoduladoras de la regulación del proceso inflamatorio y de la cicatrización.

1.2.2 Clasificación Actualmente, existe una gran cantidad de clasificaciones debido a la variabilidad de los mismos. Algunos tienen en cuenta el organismo productor, su estructura, función, célula blanco, mecanismos de acción, entre otros. Algunas de las calificaciones más comunes son:

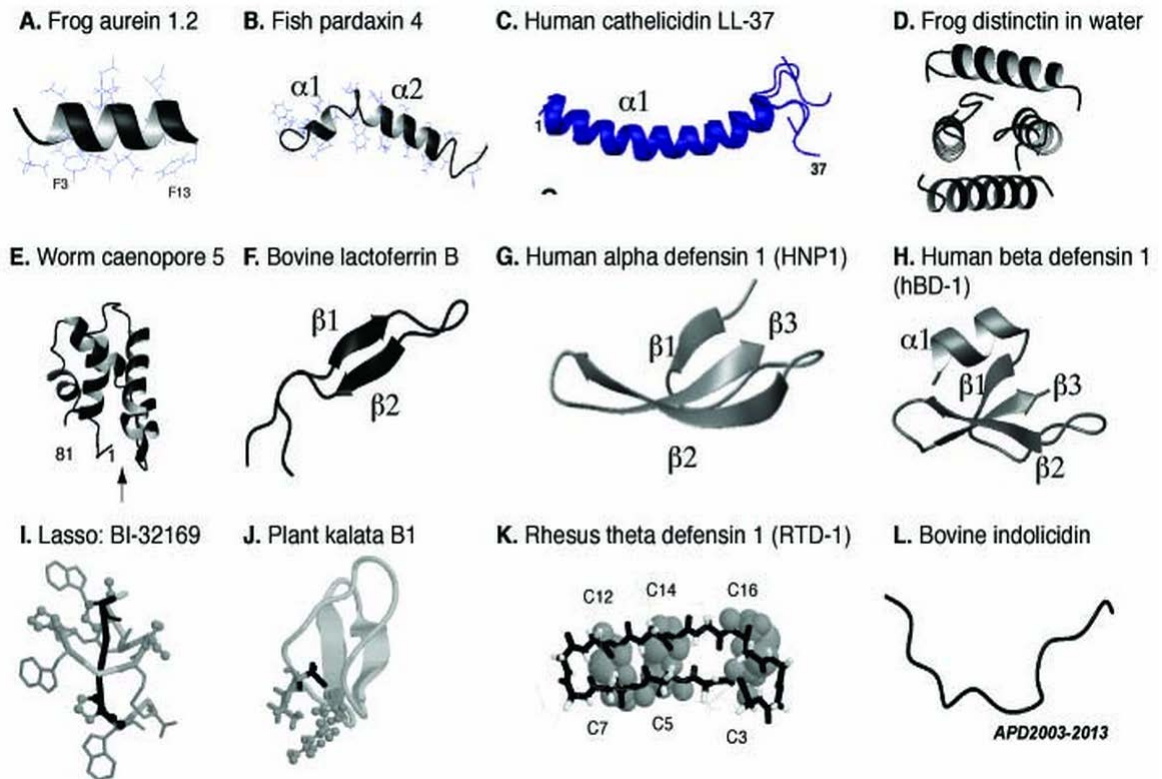
1.2.2.2 Según su estructura Dependiendo de la secuencia de aminoácidos, los PAMs, pueden adoptar diferentes estructuras secundarias o terciarias, las más comunes son las hélices  $\alpha$  y láminas  $\beta$ ; los demás péptidos pueden adquirir conformaciones cíclicas, globulares o irregulares (ver figura 4 y figura 5).

Figura 4. Clasificación de péptidos según su estructura.

ESTRUCTURA	NÚMERO DE PÉPTIDOS	PORCENTAJE
Hélice	358	13,85%
Estructura Beta	102	3,94%
Hélice y Beta desempaquetados	3	0,11%
Combiación Hélice y Beta empaquetados	96	3,71%
Sin estructura Hélice o Beta	9	0,34%
Rica en AA inusual	88	3,40%
Puente de disulfuro (No 3D)	402	15,56%
No desconoce estructura 3D	1510	58,45%

Fuente: Wang, G., Li, X. and Wang, Z. (2009) APD2: the updated antimicrobial peptide database and its application in peptide design. Nucleic Acids Research 37, D933-D937.

Figura 5. Diversidad estructural de péptidos antimicrobianos.



Fuente: Wang, G.\* (2013) Database-guided discovery of potent peptides to combat HIV-1 or superbugs. *Pharmaceuticals* 6, 728-758

- Hélices  $\alpha$  La mayoría de los PAMs están dentro de este grupo. Estos péptidos en fase acuosa no adquieren esta conformación, pero cuando entran en contacto con las membranas celulares blanco, forman la hélice  $\alpha$  para interactuar insertándose o dañando la membrana [12].
- Laminas  $\beta$  Por lo general tienen un peso entre 2 y 8 kDa, y están conformados por cadenas dispuestas en forma antiparalela unidas por uno o varios puentes disulfuro, que permiten formar una estructura estable importante en su actividad antimicrobiana. Esta conformación puede presentarse en solución acuosa, pero cuando interactúan con los lípidos de la membrana blanco presentan una mejor estabilidad [13].

- Otras conformaciones Entre estas los PAMs, conformados por bucles, que por lo general son ricos en prolina y arginina, por ende no forman estructuras anfipáticas [14]. Otros péptidos que recientemente están siendo estudiados, son los que adquieren una conformación cíclica, esta es una estrategia que permite el mejoramiento de la actividad de algunos PAMs, mediante la unión de forma covalente del extremo amida con el extremo carboxilo terminal, formando así una estructura más estable que la lineal, brindando resistencia a proteasas, mayor selectividad y especificidad [15].

1.2.2.2 Según actividad biológica Aunque los PAMs están catalogados con actividad antimicrobiana, sus funciones no están limitadas solo en este ámbito, también pueden tener otros espectros de actividad, así se pueden catalogar y dividir en péptidos con actividades antibacterianas, antifúngicas, antivirales, antitumorales, antiparasitarias, espermicidas e insecticidas [16]. La especificidad de los PAMs frente a diferentes blancos celulares, radica principalmente en las diferencias de conformación de la membrana en cada tipo de célula. Se han descrito gran cantidad de péptidos que tienen selectividad frente a determinadas células y también PAMs con amplio espectro, que actúan sobre muchos tipos de células [14].

La clasificación es dependiente de varios factores: organismo productor, estructura, función, célula blanco, mecanismos de acción, etc. (ver figura 6). Los péptidos antimicrobianos presentan una clasificación definida por factores como: la carga, anfipaticidad, hidrofobicidad y la tendencia estructural del péptido. En esta clasificación se encuentran péptidos con actividad: Antibacteriana, Antifúngica, Antiviral, Antitumoral, Antiparasitaria, Espermicida, Insecticida y Hemolítica [17].

Figura 6. Clasificación de péptidos según actividad biológica.

ACTIVIDAD BIOLÓGICA	NÚMERO DE PÉPTIDOS	PORCENTAJE
Péptidos Antibacteriales	2133	82,57%
Péptidos Antivirales	170	6,58%
Péptidos Antifúngicos	939	36,35%
Péptidos Antiprotozoarios	4	0,15%
Péptidos Antibiopelícula	15	0,58%
Péptidos Antiparasitarios	80	3,09%
Péptidos Insecticidas	27	1,04%
Péptidos Espermicidas	11	0,42%
Péptidos Anti-VIH	105	4,06%
Péptidos Anticancerígenos	183	7,08%
Péptidos Quimiotácticos	53	2,05%
Péptidos Cicatrizantes	10	0,38%
Péptidos Antioxidantes	19	0,73%
Péptidos Inhibitorios de encima/proteaza	12	0,46%

Fuente: Wang, G., Li, X. and Wang, Z. (2009) APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Research* 37, D933-D937.

- Péptidos anticancerígenos / antitumorales Los aminoácidos son compuestos orgánicos, que dan lugar a péptidos y proteínas según el número de aminoácidos que lo conformen. Si están conformados por 2 a 10 aminoácidos se conocen como oligopéptidos, más de 10 y hasta 100 aminoácidos como polipéptidos, y por encima de 100 aminoácidos se constituyen como proteínas. Tanto las proteínas como los péptidos están conformados de aminoácidos; estos son esenciales para el ser humano, pues cumplen con funciones primordiales como es el transporte de oxígeno en la sangre, algo tan elemental para la vida [8].

Por otro lado, también se encuentran los péptidos con actividad anticancerígena o antitumoral, cuya función básicamente es el combatir las células por medio de la creación de anticuerpos, estos péptidos se encuentran en el repositorio de la APD [11] [10].

### 1.3 DISEÑO DE PÉPTIDOS

Teniendo como base los péptidos naturales, de los cuales se conoce su función y estructura, se realizan diferentes variaciones que dan surgimiento a un nuevo péptido diferente al original, el cual fortalece y complementa falencias de los péptidos naturales. Por otro lado, se encuentra el diseño de novo basado en las propiedades individuales de los aminoácidos, el cual busca moléculas más activas y menos citotóxicas [18].

El diseño de péptidos se basa en cuatro diferentes enfoques:

La modificación de la secuencia original de péptidos naturales, eliminando, reemplazando o adicionando algunos residuos; bloqueando ya sea el extremo amino o carboxilo terminal o ensamblando secuencias de diferentes péptidos naturales.

El diseño de péptidos “de novo”, teniendo en cuenta los residuos que favorezcan la formación de hélices  $\alpha$ , tales como lisina o arginina como residuos básicos y alanina, leucina, isoleucina, fenilalanina o triptófano como residuos neutros para brindar hidrofobicidad a la molécula; con esta estrategia se obtienen péptidos muy potentes, pero se ignora el proceso de evolución natural por lo que se puede presentar alta toxicidad.

La utilización de librerías combinatorias, tanto de péptidos naturales como de péptidos sintéticos se obtienen moléculas con alta actividad.

El diseño de péptidos basados en secuencias de amplias colecciones de péptidos naturales, que señalen la distribución de diferentes residuos, e influyan en la variación de las condiciones fisicoquímicas, y puedan ser correlacionadas con la especificidad y la potencia; esta estrategia es muy útil ya que tiene en cuenta la evolución, y disminuye la cantidad de péptidos que se deben sintetizar para ser evaluados [18].

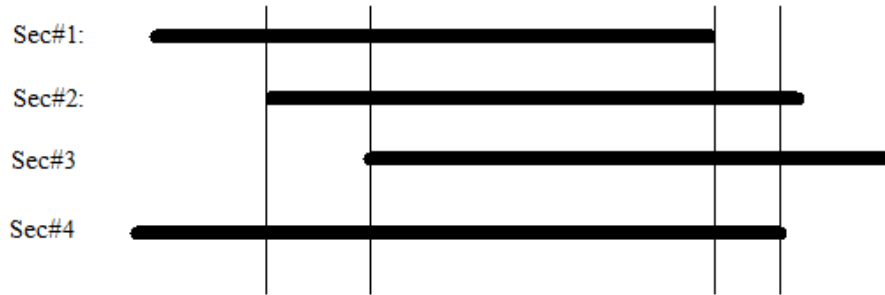
La longitud del péptido es variable y depende de la aplicación en la que se utilizan. Por ejemplo, los péptidos de 10-20 aminoácidos de longitud son ideales para la preparación de anticuerpos, mientras que los péptidos utilizados para estudios de estructura / función pueden ser de longitud variable.

### 1.4 ALINEAMIENTO DE MÚLTIPLES SECUENCIAS

El alineamiento múltiple de secuencias, es un método que ayuda a determinar si una secuencia de proteínas o péptidos está relacionada con otra, de esta forma,



Figura 9. Alineamiento múltiple, para 4 secuencias.



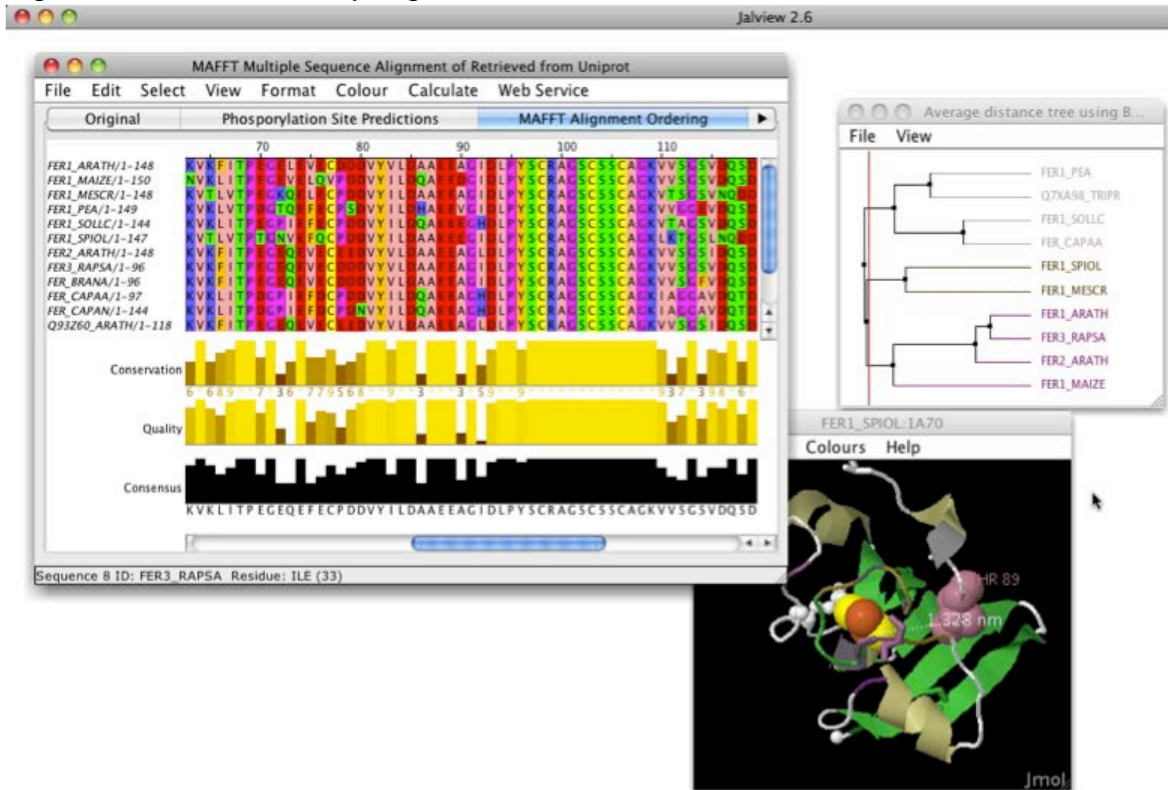
Por otro lado, se encuentra el alineamiento local [19] y el global [20], los cuales son dependientes de la región a alinear. Todas estas técnicas son basadas en matrices de castigo o penalidad (ver figura 10), con el fin de hallar la mayor afinidad posible en el emparejamiento entre las secuencias, por tanto, la ruta con mayor puntuación sería el mejor emparejamiento. También se debe destacar que puede existir más de un emparejamiento viable.

Figura 10. Matriz de castigos para la secuencia ACGT-.

	A	C	G	T	-
A	5	-1	-2	-1	-3
C	-1	5	-3	-2	-4
G	-2	-3	5	-2	-2
T	-1	-2	-2	5	-1
-	-3	-4	-2	-1	*

Por medio de estas técnicas se logra establecer el árbol filogenético de estas secuencias, pero también la región conservada del conjunto de secuencias inicialmente ingresadas (ver figura 11); esta técnica de alineamiento se realizaba de forma manual, lo cual tomaba mucho tiempo, por tal motivo actualmente se realiza con tiempo computacional, usando técnicas de programación dinámica para identificar un alineamiento globalmente óptimo. [21].

Figura 11. Alineamiento y región conservada en Jalview.

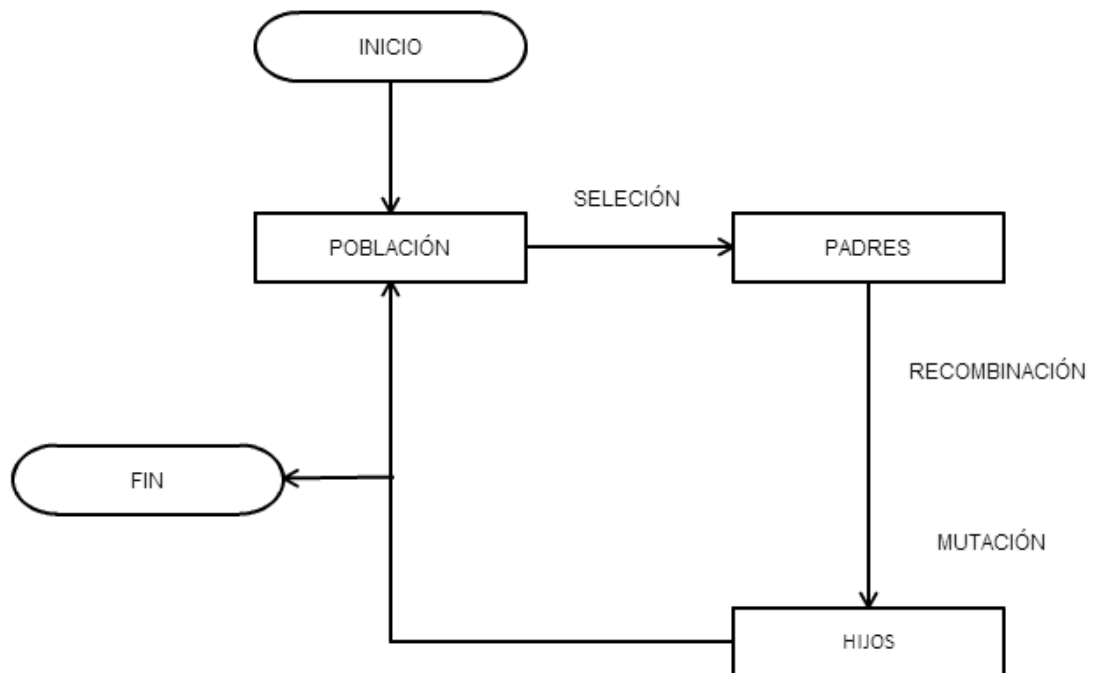


Fuente: Clamp, M., Cuff, J., Searle, S. M., & Barton, G. J. (2004). The jalview java alignment editor. *Bioinformatics*, 20(3), 426-427.

### 1.5 ALGORITMO EVOLUTIVO

El algoritmo evolutivo es una técnica computacional que actualmente es utilizada en el campo de la bioinformática. Esta técnica parte de la teoría de evolución propuesta por Charles Darwin, y desarrollada por John H. Holland en 1975, cuya motivación fue proponer un modelo general de proceso adaptable en el campo de la computación. Estos algoritmos son métodos de optimización y búsqueda de soluciones, basados en los postulados de la evolución biológica. En ellos se mantiene un conjunto de entidades que representan posibles soluciones, las cuales se mezclan y compiten entre sí, de tal manera que las más aptas son capaces de prevalecer a lo largo del tiempo, evolucionando hacia mejores soluciones cada vez. De allí surgen los algoritmos evolutivos, con el fin de solucionar problemas no lineales [22], estos algoritmos son una rama de la inteligencia artificial los cuales básicamente presentan las siguientes características: (ver figura 12)

Figura 12. Esquema general del algoritmo evolutivo en diagrama de flujo.



Población inicial: es determinada por un grupo de individuos con un tamaño específico, para este caso el grupo de péptidos con actividad anticancerígena / antitumoral.

Probabilidad de reproducción: una vez seleccionados los padres (para este caso péptidos) pueden reproducirse entre ellos, para dar lugar a una nueva población de individuos (hijos), o por el contrario, puede no reproducirse y pasar a la siguiente generación, un ejemplo de reproducción es: (ver figura 13).

Figura 13. Ejemplo de reproducción (recombinación).



Probabilidad de mutación: la cual otorga un nivel de diversidad a la población en cada generación, en caso de una mutación se modificara uno de los aminoácidos de forma aleatoria por alguno de los otros 19 restantes, un ejemplo: (ver figura 14).

Figura 14. Ejemplo de mutación.

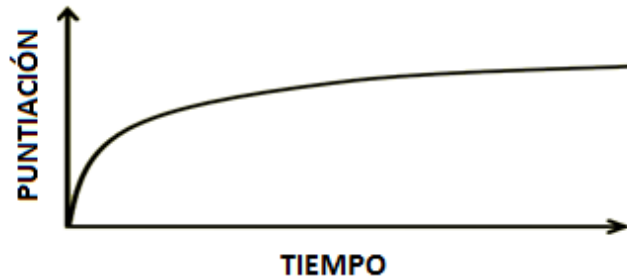


Función de aptitud: clasifica la población según el criterio y patrones definidos en esta, de forma cuantitativa. En este caso es diferida por una máquina de soporte vectorial.

Generación: representa la línea de sucesión de la población, la cual puede llegar a ser un criterio de parada.

Enfoque elitista: consiste en conservar en cada una de las generaciones al mejor individuo, con la finalidad de ver cómo evoluciona o se llega al estancamiento con el paso de las generaciones, como se observa en la figura 15.

Figura 15. Comparación de la función de aptitud vs. tiempo.



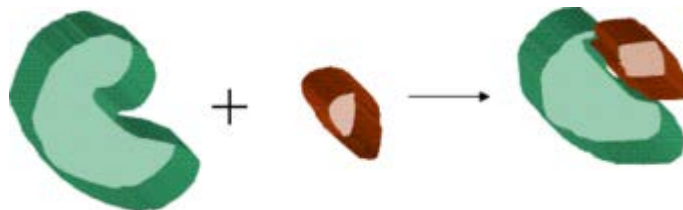
Otros Criterios: el tipo de cruce usado, es cruce en un punto, con la finalidad de no alterar el dominio conservado de las secuencias.

Esta técnica es usada para realizar búsquedas aleatorias, dentro de un espacio de soluciones mucho mayor, representando las posibles soluciones, las cuales por diferentes métodos de selección compiten para dar como resultado los individuos más aptos que conformaran las nuevas poblaciones, cabe destacar que existen otros modelos.

#### 1.6 DOCKING MOLECULAR

El acoplamiento o docking molecular es un método que predice la conformación preferida de una molécula, al estar unida a otra, con el fin de formar un complejo estable. Esta técnica es usada para predecir la orientación del enlace de una molécula pequeña, que será candidata a fármaco (Diana), con la proteína donde ejercerán su acción (Receptor), con lo que se podrá predecir la afinidad y la actividad de la molécula pequeña (ver figura 16).

Figura 16. Acoplamiento molecular (docking), de verde: el receptor y de rojo: el ligando.

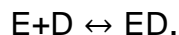


Fuente: Docking (molecular), [en línea], Wikimedia, ArvindRamanathan, 6 April 2013, [citado 15 de Mayo de 2015], Disponible en:  
es.pnghttps://en.wikipedia.org/wiki/Docking\_(molecular)#/media/File:Docking.png

El docking molecular tiene tres enfoques que son: rígido, semiflexible, y flexible [23].

El rígido, se basa en tener tanto el receptor como el ligando rígidos en su estructura química, lo cual limita el problema. Para el caso del semiflexible, se define el receptor rígido, pero el ligando flexible, lo cual permite que su estructura pueda rotar de acuerdo con los enlaces rotables que posea (para los péptidos está definido por la cantidad de carbonos alfa). Por último, el flexible, donde el receptor como el ligando tiene la posibilidad de rotar su estructura, esto se vuelve más completo y costoso computacionalmente, aunque esto puede reducirse si se conoce el sitio activo, de modo contrario sería un docking ciego o global. Para este caso se usó docking molecular semiflexible.

Si se considera la unión no covalente entre el receptor (E) y una diana (D) la ecuación de esa unión es:



Se define una constante de unión  $K_b$  como:

$$K_b = \frac{[ED]}{[E][D]}.$$

La energía libre de unión  $\Delta G_b$  se puede calcular con la reacción.

$$\Delta G_b = RT \ln K_b.$$

Donde R es la constante de los gases ideales ( $8.31 \frac{J}{KMol}$ ) y T es la temperatura absoluta en K. Cuando el valor de la energía libre de la unión se hace menor, aumenta la fuerza de unión D-E.

$$\begin{array}{ll} \Delta G_b > 0 & \text{Proceso no espontaneo, } K_b \text{ pequeña.} \\ \Delta G_b < 0 & \text{Proceso espontaneo, } K_b \text{ grande.} \end{array}$$

El diseño de drogas basado en la estructura, tiene ventajas como lo es: ahorro de tiempo y dinero para la creación de las nuevas dianas. Cada día el protein data bank (PDB), incrementa el número de receptores, cuyas estructuras son determinadas mediante difracción de rayos x, las cuales pueden ser usadas en diferente *software*, para realizar el docking molecular. En este caso se usara el programa Autodock por considerarse el más preciso [16].

La función de puntuación que se trabaja con el programa Autodock vina es:

$$C = \sum_{i>j} f_{t_i t_j}(r_{ij}).$$

Donde la sumatoria es el par de átomos que se pueden mover relativamente a otra molécula. Cada átomo de  $i$  está asignado a un tipo  $t_i$  y se define un set de funciones de interacción  $f_{t_i t_j}$  de una distancia interatómica  $r_{ij}$ , el valor de  $c$  puede verse como la suma de la contribución intramolecular e intermolecular [23].

$$c = c_{inter} + c_{intra}. \text{ [Kcal/mol]}.$$

El algoritmo del programa Autodock vina trata de encontrar el menor valor de  $c$ . La energía libre de la unión se calcula a partir de la conformación con el menor valor de puntuación, se representa como:

$$S_1 = g(c_1 - c_{intra1}) = g(c_{inter1}).$$

Donde  $g$  puede ser una función arbitraria (posiblemente no lineal). En la salida del programa, las otras conformaciones se calculan teniendo en cuenta el  $C_{intra}$  del mejor modo de unión:

$$S_i = g(c_i - c_{intra}).$$

Las funciones de interacción  $f_{t_i t_j}$  son definidas relativas a la distancia de la superficie teniendo en cuenta los radios de Van der Waals  $R_t$ :

$$d_{ij} = r_{ij} - R_{t_i} - R_{t_j}.$$

$$f_{t_i t_j}(r_{ij}) \equiv h_{t_i t_j}(d_{ij}).$$

La función de puntuación  $h$  toma en cuenta los factores: estericos, hidrofóbicos y puentes de hidrógeno. Para comprobar la exactitud de las predicciones de la estructura experimental, se emplea una medida de las distancias entre las estructuras experimentales y predichas que tienen en cuenta la simetría y simetría parcial, se denomina RMSD (desviación de la media cuadrática), para las estructuras  $a$  y  $b$  de una molecular idéntica, RMSD se define como:

$$RMSD_{ab} = \max(RMSD'_{ab}, RMSD'_{ba}).$$

Donde,

$$RMSD'_{ab} = \sqrt{\frac{1}{N} \sum_i \min r_{ij}^2}.$$

Y la sumatoria es: todos los N átomos pesados en la estructura a, el mínimo es: todos los átomos de la estructura b con el mismo tipo de átomo i en la estructura a.

Tanto la energía C [kcal/mol], como la distancia RMSD [Å] son imprescindibles para la comprobación del comportamiento de ligando como el receptor, por esa razón este método tiene un papel muy importante en el diseño racional de fármacos. Dada la importancia biológica y farmacéutica, se han hecho grandes esfuerzos buscando mejorar el método usado para predecir el acoplamiento molecular [24].

## 2. PLANTEAMIENTO DEL PROBLEMA

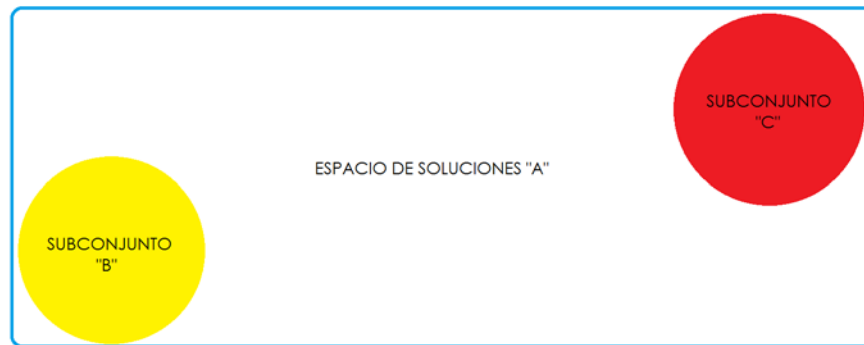
Debido al alto crecimiento de casos de cáncer, se busca con urgencia diferentes soluciones a esta problemática.

Por tal motivo este trabajo se basó en la proteína Ras que es la causante de la proliferación celular descontrolada, buscando nuevos péptidos con actividad inhibitoria sobre esta proteína, partiendo de los péptidos ya conocidos y realizando recombinaciones con el fin de potenciar la actividad de inhibición, visto de otra manera.

Partiendo de:

- Existe un espacio de soluciones "A".
- Se conoce inicialmente un subconjunto "B".
- Se quiere llegar a un subconjunto "C".
- Donde "C" de forma cuantitativa es superior a "B".

Figura 17. Planteamiento del problema.



Donde:

- "A" Es el espacio de todas las posibles soluciones de péptidos con actividad inhibitoria sobre la proteína Ras.
- "B" Es el subconjunto inicial de péptidos extraído de la base de datos del APD.
- "C" Es el subconjunto final de péptido con mejor afinidad a la inhibición antes la proteína Ras.

El problema a resolver:

Partiendo del subconjunto de péptidos "B" se debe llegar al subconjunto de péptidos "C" (ver figura 17).

### 3. PROCEDIMIENTO

#### 3.1 CONJUNTO DE DATOS INICIALES

Para la finalidad de este trabajo, el conjunto de datos con el cual se trabajó está representado por secuencias de péptidos, las cuales son cadenas de texto formadas por aminoácidos, cuya actividad debe estar previamente definida y comprobada científicamente como anticancerígenas o antitumorales. Para ello se acudió a la base de datos en línea APD [9] [10], donde se encuentra una base de péptidos clasificados por su actividad.

En este repositorio se encontraron 166 secuencias de péptidos con actividad anticancerígena (consulta realizada el 2 septiembre de 2014), las cuales son el conjunto de datos iniciales. Se debe tener en cuenta que el número de estos péptidos puede ir aumentando día a día.

Estos datos se deben transformar a formato fasta (formato de fichero informático basado en texto, utilizado para representar secuencias de ácidos nucleicos y péptidos. Donde los pares de bases o los aminoácidos se representan usando códigos de una única letra), como se aprecia a continuación:

>Nombre\_de\_la\_Secuencia no son válidos caracteres especiales  
Secuencia

EJEMPLO:

>AP01809  
GIPCGESCVFIPCISSVIGCSCSSKVCYRN

#### 3.2 ALINEAMIENTO DE DATOS

Por medio de la herramienta en línea ClustalW2 [25] del instituto europeo de bioinformática, se realizó el alineamiento de las secuencias de péptidos, las cuales deben estar previamente en formato fasta. Una vez alineadas las secuencias de péptidos, se pasó a verificar el árbol filogenético, que se obtuvo de la misma herramienta con la finalidad de verificar la diversidad de péptidos para poder realizar recombinación o cruce en los pasos siguientes. Verificado el árbol filogenético, se procedió a identificar la región conservada.

Existe otra versión de Clustal denominada ClustalX la cual es usada de forma local para Unix/Linux, Mac Os y Windows, por el hecho de hacer uso solo una vez del procedimiento para este caso se utilizó ClustalW.

### 3.3 REGIÓN CONSERVADA

La región conservada es el segmento del péptido donde se encuentra explícita la función nativa. En este caso, la función es de inhibición o función anticancerígena. Para identificar la región se hizo uso de la herramienta JALVIEW [26]. De esta, se extrajo la región conservada para realizar el algoritmo molecular evolutivo.

Para poder obtener la región conservada es indispensable realizar el alineamiento de las secuencias, sin el alineamiento no se puede obtener la región donde reside la función de inhibición. Luego del alineamiento, se recortó la región la cual posee una longitud de 11 pares de bases nitrogenadas. Posterior a esto se hizo un refinamiento de datos donde se eliminaron las secuencias idénticas para evitar redundancia. Esta lista de secuencias recortadas, alineadas y refinadas, constituyen la población inicial.

### 3.4 SISTEMA DE PUNTUACIÓN

El sistema de puntuación determinó de forma cuantitativa, que tan buena es la función de inhibición, a mayor puntuación se considera más efectiva la actividad anticancerígena del péptido. Para obtener este puntaje se realizó una consulta externa a la página web del Instituto de la India Anticp (ver figura 18) [27]. Donde el modelo definido para la identificación y clasificación de los péptidos anticancerígenos es por medio del uso de máquinas de soporte vectorial, capaces de identificar y evaluar la actividad los péptidos en un intervalo desde  $-\infty$  hasta  $+\infty$ . Siendo lo péptidos cuantitativamente positivos aquellos con actividad anticancerígena, y los negativos con actividad no anticancerígena.

Figura 18. Resultado luego de una consulta realizada en el portal web del ANTICP [27], donde la puntuación para cada uno de los péptidos es definida por la máquina de soporte vectorial (SVM Score), y el resultado de la predicción sobre su actividad (Prediction).

Query Peptides			
Peptide ID ↕	Peptide Sequence ↕	SVM Score ▾	Prediction ↕
seq4	LTAEHYAAQATS	0.95	Anticp
seq1	FKRIVQRIKDFLR	0.93	Anticp
seq3	GIPCGESC VFIPCISSVIGCSCSSKVCYRN	0.78	Anticp
seq5	PGLGFY	0.47	Anticp
seq2	GRSGDSDEEL	-0.04	Non-Anticp

Fuente: AntiCP:-Designing of Anticancer Peptides, online, Institute of Microbial Technology, Gajendra P. S. Raghava, 18 Octubre 2013, Cita (20 de Mayo de 2015) Disponible en: <http://crdd.osdd.net/raghava/anticp/>

Para poder automatizar este procedimiento se realizan peticiones por método POST desde Python, con el fin de enviar cada una de las generaciones y obtener el puntaje (SVM Score) para poder clasificar los péptidos.

### 3.5 ALGORITMO EVOLUTIVO

Tomando como base el algoritmo evolutivo estándar, se realizan distintos ajustes con la finalidad de optimizar y adaptar el algoritmo a este problema, variando los siguientes parámetros:

El porcentaje de reproducción definido fue del 80%, ya que con porcentajes inferiores el número de interacciones aumenta, lo cual aumenta el tiempo de convergencia del algoritmo, y con porcentajes mayores la puntuación de los péptidos en promedio disminuye.

La probabilidad para escoger el punto de corte previo para la reproducción se define como: Aleatoria, para poder dar un porcentaje igual en cada uno de las posiciones de los aminoácidos a recombinar.

La probabilidad de mutación fue del 1%, se define un porcentaje bajo para poder dar diversidad en la población, ya que al definirse un porcentaje alto se encuentra degeneración en la población.

El número de generaciones del algoritmo fue de 300, al correr 1000, 800 y 600 generaciones, se encuentra un criterio de parada en la evolución del mejor péptido sobre las primeras 200 a 250 generaciones.

Cada 100 generaciones se modifica la población, por todos aquellos péptidos cuyo puntaje sea alto en las 100 generaciones pasadas, de esta forma, el puntaje general y máximo mejora notablemente.

El tipo de cruce usado fue recombinación en un punto, con la finalidad de seguir con la homología de los péptidos este tipo de cruce es el más viable.

El enfoque fue elitista, el cual busca mostrar la evolución del algoritmo y estancamiento del mismo.

La población inicial fue la región conservada de los péptidos.

La función de aptitud fue definida por medio de una máquina de soporte vectorial, ubicada en el portal web del ANTICP [27] para la predicción de la actividad anticancerígena.

El algoritmo queda de la siguiente forma:

---

#### ALGORITMO EVOLUTIVO V8

---

0: Entrada: Población inicial de péptidos.

1: Para 300 iteraciones.

2:     Si Acumula 100 iteraciones.

3:         Se cambia la población inicial por la población mejorada.

4     Para cada péptido.

5:         Consultar la puntuación de la generación en la página web del ANTICP.

6:         Si puntaje es mayor al puntaje del mejor péptido.

7:             Se añade el péptido a la población mejorada.

8:             Se guarda el puntaje como mejor péptido de la generación.

9:             Si el puntaje es diferente al mejor.

10:             Si la probabilidad de reproducción se cumple.

- 11: Aleatoriamente se busca una pareja para el individuo.  
12: Aleatoriamente se define el punto de corte.  
13: Se realiza la recombinación (hijos).  
14: Los hijos pasan a la siguiente generación.  
15: Si la probabilidad de mutación se cumple.  
16: Muta el individuo (hijo) en una única posición.  
17: Sino  
18: El individuo pasa a la siguiente generación.  
19: Salida: Nueva generación.

Al final se obtienen como resultado los péptidos con mejor afinidad a las proteínas Ras para la inhibición.

### 3.6 DOCKING

Luego de obtener los primeros resultados del algoritmo evolutivo se observó una posible mejora en la función de inhibición de los péptidos, pero para corroborar estos resultados se hizo uso del docking molecular semiflexible con la finalidad de demostrar un mejor acople entre la proteína Ras y la nueva población de péptidos por medio de Autodock v 1.5.4 y Autodock vina [28].

Para realizar este docking se necesita:

- La proteína Ras debe estar en formato pdb.

Para poder obtener las proteínas Ras en este formato se consultó la página web del protein data bank PDB, en el cual se encuentra la estructura tridimensional con el código 5P21.

- Modelo tridimensional de las secuencias de péptido obtenidas luego del algoritmo evolutivo molecular.

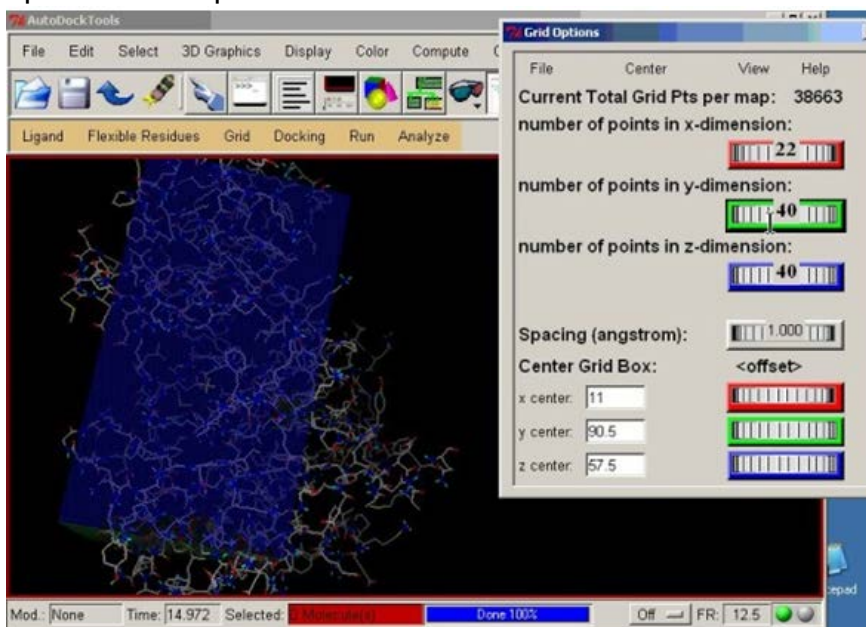
En este caso, como son secuencias nuevas no fue posible obtenerlas en el PDB. Por tal motivo, se hizo uso de la herramienta web pepfold, en la cual se introdujo las secuencias obtenidas y el servidor generó diferentes modelos 3D de la molécula. En este estudio se usó el modelo número 1 en todos los casos. Para este proceso también se pueden usar la herramienta I-Tasser para la visualización, la cual se puede vincular con correos institucionales para obtener los resultados directamente en el correo.

Una vez obtenidos los modelos tridimensionales se usó la herramienta Autodock, para lo cual se realizó modificaciones a la proteína Ras, preparándola para el acople con cada una de las nuevas secuencias obtenidas. Cabe resaltar que también se hizo uso de la herramienta pyMol [29], para visualizar la estructura de los péptidos y de la proteína Ras.

Se cargó el archivo de la proteína Ras en Autodock, allí se seleccionó y eliminó los cofactores, ligandos, solventes, iones, etc. Se guardó como un nuevo archivo pdb. Lo siguiente fue cargar el ligando que es la estructura tridimensional de los nuevos péptidos y se definieron las rotaciones como el mismo número de carbonos alfa presentes en la secuencia de péptidos.

Se identificó el sitio activo de la proteína Ras, anteriormente referenciado en las posiciones 32 a 40 creando un grid box (Área cubica, la cual delimita la región donde se realiza el acople entre el receptor y el ligando) para la interacción entre ligando y receptor ( $x = 0.642$ ,  $y = 38.347$ ,  $z = 18.769$ ,  $size\_x = 40$ ,  $size\_y = 20$ ,  $size\_z = 58$ ), una vez definidos estos patrones, se crea el archivo de extensión: PDBQT que es el formato nativo (ver figura 19).

Figura 19. Grid box en Autodock tools, en la cual se observa de color azul el recuadro que delimita la interacción entre receptor y ligando para el acople molecular, en el sitio activo de la proteína Ras. De esta forma, se limita en esta área la búsqueda del acople.



Fuente: Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2), 455-461.

Luego de los ajustes hechos al receptor y el ligando se procede hacer uso del programa Autodock vina, teniendo en cuenta que es por línea de comandos, el archivo debe estar en la misma carpeta que el ejecutable. Se usan los siguientes parámetros en el archivo \*.confi

Tabla 1. Parámetros utilizados en el programa Autodock vina, para el acople entre los péptidos obtenidos y la proteína Ras.

Función	Parametro	Uso
receptor	5P21.pdbqt	Nombre de Receptor
ligando	péptido.pdbqt	Nombre del Peptido
center_x	0,642	Coordenada en el eje x
center_y	38,347	Coordenada en el eje y
center_z	18,769	Coordenada en el eje z
size_x	40	Tamaño sobre el eje x
size_y	20	Tamaño sobre el eje y
size_z	58	Tamaño sobre el eje z
out	resultado_peptido_1.pdbqt	Nombre Archivo de salida
exhaustiveness	100	Carácter exhaustivo
num_modes	20	numero de modelos

Como resultado se obtienen 20 diferentes modelos con su respectiva energía de interacción y distancia RMSD, el tiempo de corrida varía de acuerdo con el equipo usado, en este caso, se empleó un Intel Core i7 920 (4 núcleos @2.66Ghz) y 6GB de Ram, demorándose aproximadamente para cada péptido un promedio de 13 horas.

Una vez el Autodock vina termina el docking molecular semiflexible, se obtienen los 20 modelos por péptido resultante del algoritmo evolutivo, en donde se encontró la energía expresada en Kcal/mol, que se traduce en la energía de acople entre el ligando y el receptor, considerando un mejor acople cuanto más negativa se presente. También se obtuvo como parte del resultado la distancia de las moléculas más próximas entre el ligando y el receptor, denominada como la desviación cuadrática media RMSD expresada en Å.

Estos resultados fueron comparados con el mejor péptido de la generación inicial, con el fin de saber si existe una mejoría en la actividad de inhibición de los nuevos péptidos sintéticos antes la proteína Ras 5P21 (ver figura 23).

#### 4. RESULTADOS Y DISCUSIÓN

La optimización evolutiva de los ligandos peptídicos es un proceso complejo, en el que numerosos parámetros y diferentes mecanismos evolutivos podrían depender e influir entre sí. Para mantener esas variables en un nivel manejable se definió un marco general en el principio.

Como meta evolutiva se decidió optimizar la lista de ligandos peptídicos que se pueden unir a la proteína Ras. Estas moléculas fueron particularmente elegidas por varias razones: En primer lugar, las moléculas están comprobadas para su actividad anticancerígena. En segundo lugar, a pesar de que sus tamaños son altamente variables, al realizar un alineamiento base poseen un dominio conservado. Debido a este dominio conservado, son buenos candidatos para interactuar con la proteína Ras. Sin embargo, a pesar que se ha comprobado su actividad biológica, se requiere de nuevos candidatos, los cuales se pueden hallar por medio del diseño computacional, debido a que pueden surgir nuevas secuencias de péptidos con actividad anticancerígena mejorada.

Para este trabajo, como población inicial para el proceso de evolución, se utilizaron 166 secuencias de péptidos previamente identificadas (ver Tabla 2) como anticancerígenas.

Tabla 2. Péptidos extraídos del APD, cada nombre corresponde al código definido por el APD para identificar los péptidos con actividad anticancerígena.

AP00001	AP00058	AP00200	AP00345	AP00566	AP01124	AP01413	AP01777	AP01983	AP02297
AP00012	AP00060	AP00214	AP00351	AP00605	AP01161	AP01414	AP01784	AP01984	AP02325
AP00013	AP00061	AP00236	AP00352	AP00640	AP01196	AP01415	AP01785	AP01985	AP02326
AP00014	AP00062	AP00240	AP00353	AP00644	AP01223	AP01416	AP01794	AP01986	AP02327
AP00015	AP00090	AP00242	AP00366	AP00708	AP01239	AP01417	AP01805	AP01987	AP02328
AP00016	AP00101	AP00244	AP00367	AP00770	AP01277	AP01418	AP01806	AP01988	AP02329
AP00017	AP00128	AP00245	AP00396	AP00809	AP01278	AP01465	AP01807	AP01989	AP02340
AP00018	AP00135	AP00246	AP00427	AP00810	AP01279	AP01466	AP01808	AP01990	AP02372
AP00019	AP00139	AP00247	AP00429	AP00832	AP01280	AP01467	AP01809	AP01991	AP02374
AP00020	AP00140	AP00248	AP00451	AP00878	AP01281	AP01512	AP01810	AP01992	AP02375
AP00021	AP00144	AP00260	AP00473	AP00964	AP01282	AP01519	AP01811	AP02111	AP02378
AP00022	AP00146	AP00261	AP00506	AP00973	AP01283	AP01521	AP01812	AP02141	AP02379
AP00023	AP00165	AP00262	AP00513	AP01026	AP01284	AP00494	AP01813	AP02142	AP02380
AP00025	AP00176	AP00263	AP00532	AP01031	AP01328	AP01650	AP01922	AP02228	
AP00026	AP00177	AP00308	AP00541	AP01036	AP01342	AP01774	AP01923	AP02234	
AP00029	AP00178	AP00310	AP00553	AP01121	AP01406	AP01775	AP01956	AP02235	
AP00035	AP00191	AP00339	AP00565	AP01123	AP01412	AP01776	AP01957	AP02296	

Para permitir la identificación del dominio conservado, se realizó un alineamiento en Clustal y se visualizó con Jalview, luego se procedió a recortar el fragmento que posee el dominio conservado (ver Tabla 3). Este dominio es importante, ya que posee la actividad esencial de los péptidos, tal como se puede ver también en el estudio realizado sobre el factor de transcripción del Nrf2, el cual tienen seis dominios conservados asociados a su regulación negativa [30].

Tabla 3. Péptidos alineados y recortados por región conservada, los péptidos de la tabla 2, son alineados mediante Clustal W y luego se extrae la región conservada, como resultado se obtienen estos péptidos, y definidos como la generación inicial del algoritmo evolutivo, puesto que este fragmento conserva la función de inhibición.

DIK-KIAESI	KFLG-----RI	-LLRKILE---	KFKFSPVILG-	GHIGKAIL---	CVGGCGCSWPV	CVWICGCSNKV
DIK-KIAESF	KILG-KIIAK-	GLLSGVSAPHV	RSLGRAWKIIV	SAVGSALSQE-	CVGGCSCSWPV	CVWICSCKNNV
DIVK-KIASSI	KILG-KIIVK-	GLLSGVSVPVH	MSLLK-----	WLIKAIHAHGL	CFGGCTCDWPV	CVFICSCDKV
DIVK-KIAGSI	KVLG-KIIAK-	GLFSGAVAPHV	PIIAAKVVCAI	HIKLFHAHGL	CAGGCSCSWPI	CVYICSCDKV
DIVK-KVVGSL	KILG-HLI---	GLFGGSIAPHV	PIIAAKVLCAI	HIFRIVHVHRL	CTLGCSCSWPV	CVFLCCKSKV
DIK-KVIGSL	KKIG-KIEGQH	GLFKGVSAPHV	PIIAAKFLCTI	NIYNQVCAKI	CTLGCSCSWPI	CYLICYVSRQ
DIVK-KVVAGL	EVGKKAAGKA	GLLGGSVV-HV	PAIVAKFLCAI	NIYNERCAKI	CFTGCSCNWPT	C---CNC----
DIK-KVAGGL	EAAKKAALGLV	GLFVAKVA-HV	PVIAANFLCAI	NIYNEVCAKI	CVQGCSCDWPI	----DGC----
DVIK-KVAGGL	GIGKQAAL-KT	GLFVAKVA-HN	PLLAANFLCKI	NIYNETCAKI	CLLGCTCSRPI	CRTGDVCRRYQ
DFLK-KVAANV	GGVKKGALSTY	-LLDE-----	PILAAKFGCLV	NIYNPTCAKI	CFKFCSCSYPF	CDFTNRCPGKI
DVVKKNVAEQL	SGLKKGAASTY	GLLGKIAA--V	GGL-IKIVCAV	NCYNEICAKI	CVWICCKSKV	NDLGQGLIVNV
DTIKKTVDKLV	SAGKKGLAEKY	-ILGSTIG--V	----IGMTCLI	NLAD--CFC--	CVFICSCSKV	----TGLPRKI
DIKKEFANNL	-GFKKGAAKTV	KLFKHSKAK---	PIL--SLIGK-	NLAD-----	CVFICCKSKV	SAYSGATAKKL
RKLKKSVKKEF	IDWKDAAK---	KLFKHLAK---	PIR--SLI---	CRIPCIYQWAF	CVFICCSNKV	NVLSQLLAARD
----SVAQGM	AVLKTGLPSWI	----HSAK-KA	PLK--KLR---	CLYSTCYRAK-	CVFICCKSKV	DLKKANII-SI
GQHG-----	MALKKVL----	KIFKNGIIAVA	SLIP-SLVSAF	CFPKMDCRWK-	CVFICCSNKV	KAMKKFIAKEH
GQGGEAIQKDF	LA-KKKLA---	KLFKDGIIAVA	GMIP-GLISAF	CYKQ--CV-TY	CVWICCSNKV	RSSRLQFPHRL
EKIGRIVQKDF	FASKKDLADLV	KLFDDGV----	SLLP-SLISAI	CYRG--CY-RR	CVWICCSNKV	PFPPPIPRPF
---GRIVQKDF	LKLKSWAK---	-LLGSAISL--	DMVGDVTQVC	CTPG--CG-VC	CVWICSCDKV	LPFRIRIHYPP
KKLGPSITRRA	EKLESLAKGIQ	----WAIV---	AMVG-T-----	CVGGCACSWPV	CVYICCKSKV	KQLQTPCEEKI

Para medir la actividad anticancerígena de los péptidos, se procedió a consultar las secuencias recortadas en la página web del Instituto de la India Anticp [20] donde obtiene el puntaje para cada uno de los péptidos. Esta página también ha sido usada en estudios como la revisión de literatura sobre las actividades antitumorales de los péptidos, liberados de proteínas de la leche [31], y también

como base de datos en el diseño de péptidos con vida media deseada en el intestino, el cual buscaba conseguir péptidos anticancerígenos poco tóxicos [32].

Después de determinar la lista de péptidos con su respectivo puntaje, se usaron péptidos cuya puntuación fuese mayor a una (1.0) unidad para la población inicial (puntaje obtenido de la máquina de soporte vectorial para calificar los péptidos con mejor actividad), con la finalidad de iniciar con una mejor población y minimizar el tiempo de convergencia, ya que el mayor puntaje de todo el grupo inicial de pedidos fue de 1.35 unidades. Después de este filtrado, se obtuvo una población de 140 péptidos.

Una vez se obtienen estos segmentos, se da inicio al algoritmo evolutivo molecular, en el cual se busca ver una evolución de la población, tanto como grupo, como del mejor individuo. Este algoritmo es ampliamente utilizado para ver la evolución de péptidos en actividades de regulación, como se puede observar en el trabajo de evolución dirigida de péptidos que inhiben el uso de tripsina, el cual emplea un proceso dirigido a la búsqueda de moléculas que presentan ciertas propiedades deseables [33].

Como resultado del algoritmo evolutivo, se observó un incremento en cuanto al puntaje de los péptidos, un total de 40 péptidos con un puntaje mayor 1.7, un puntaje máximo de 1.98 y un puntaje promedio de 1.78 (ver tablas 4, 5, y figura 20).

Tabla 4. Comparación de puntuaciones de población de péptidos iniciales vs. población de péptidos finales, se observa un mejor resultado en la función de inhibición.

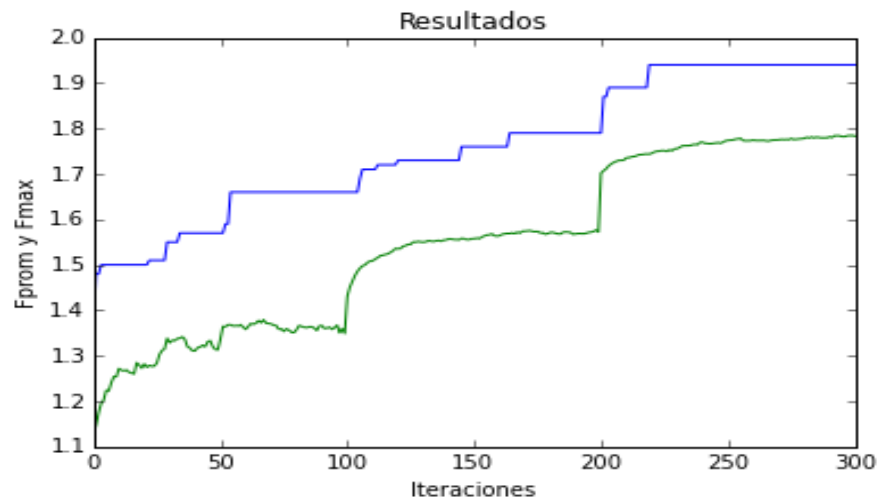
	PUNTAJE POBLACION INICIAL	PUNTAJE POBLACION FINAL	% MEJORIA
<b>MÍNIMO</b>	1.01	1.7	68%
<b>MÁXIMO</b>	1.35	1.98	46%
<b>PROMEDIO</b>	1.12	1.78	58%

La mejora del puntaje de generación en generación fue progresiva, aunque fue lenta comparada con otros estudios [33]; se debió a que la población inicial, al ser un dominio conservado, presenta una naturaleza estructural peptídica y actividad biológica similar. Esto se observó en el estudio realizado por Chittenden sobre el dominio conservado en Bak [34].

Tabla 5. Resultado del algoritmo evolutivo molecular, pasadas 300 generaciones se obtienen como resultado estos péptidos, los cuales tiene una mayor afinidad a la proteína Ras para su inhibición.

KEGCCNESCKV	KLGCCSEECKV	KLPCCSCSCKV	KLG-CNCCKI
KAPCCLESCKV	KEGCCNLECKE	KLGCCNCCKV	KDGCCSCECKI
KLGCCNESCKF	KMGCCSECKI	KLGCCNCCKQ	KAPCCSTCKV
KLGCCNESCKI	KAGCCNEECKE	KYGCCSECKI	KLGLCNCCKV
KLGCCNESCKS	KRGCCNCCKI	KLGCCSCCKF	KDGCCSCCKC
KLGCCSECKD	KLGCCRTECKI	KLGCCSCCKC	KDGCCSCCKS
KLGCCSECKI	KLPCCNTECKI	KLGCCSCCKV	KDGCCSCCKV
KLGCCSECKV	KLG-CNTCKI	KLGCCSCECKI	KAGCCSCCKI
KLGCCSECKY	KLPCCNTCKI	KLGCCSCCKY	KLFCCSCCKV
KLG-CSECKS	KTPCCNCECKS	KLGCCSCCKI	KLGLCSCCKV

Figura 20. Algoritmo evolutivo v8, de color azul el puntaje máximo del mejor péptido para cada generación, el verde el puntaje promedio de la población para cada generación.



Se observa en la figura 20, el comportamiento de la población para cada una de las generaciones (iteraciones), una notable evolución tanto general como del mejor individuo. A través de las generaciones, también se observa un estancamiento entre la generación 200 a 250. En la tabla 5 se observan, como se fijan o establecen determinados aminoácidos en las diferentes posiciones.

Para evaluar la afinidad de estos nuevos péptidos con la proteína Ras 5P21, se realizó un docking molecular semiflexible (ver figura 22). Para ello, se hizo uso del Autodock vina, con el fin de comparar los valores de las distancias RMSD entre los péptidos iniciales y los péptidos resultantes del algoritmo. Como resultado, se obtuvieron 20 diferentes modelos para cada péptido (ver figura 21), esto con el fin de tener mayor precisión. En la figura 21, se logra apreciar una mejoría tanto de la distancia RMSD (la cual disminuye confirmando un mejor acople) como de la energía de acoplamiento. Esto confirma una evolución en los nuevos péptidos obtenidos comparados con los iniciales.

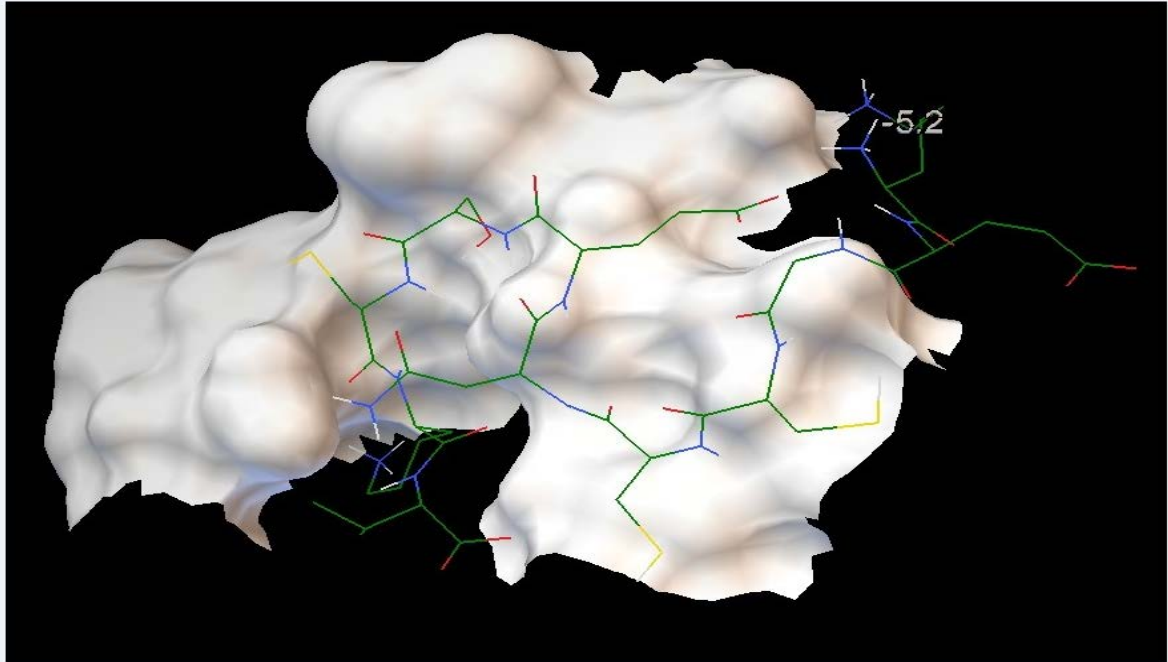
Figura 21. Valores de distancias y energía obtenidos de Autodock Vina luego del acople molecular entre la proteína Ras y el péptido CVFICSCSSKV en la figura a, y entre la proteína Ras y el péptido KLGCCNESCKI en la figura b, se observa una mejora tanto en la distancia RMSD como en la energía de interacción, lo cual se traduce en un mejor acople entre la proteína Ras y el péptido KLGCCNESCKI.

	⟨kcal/mol⟩	rmsd l.b.	rmsd u.b.		⟨kcal/mol⟩	rmsd l.b.	rmsd u.b.
1	-4.6	0.000	0.000	1	-5.6	0.000	0.000
2	-4.6	5.716	13.333	2	-5.6	6.374	13.828
3	-4.5	5.230	9.482	3	-5.5	4.012	13.981
4	-4.5	6.604	10.349	4	-5.5	3.326	14.015
5	-4.5	5.943	10.804	5	-5.5	3.705	14.628
6	-4.5	6.755	11.475	6	-5.4	3.778	14.483
7	-4.5	20.078	24.342	7	-5.4	3.523	12.924
8	-4.5	20.346	25.228	8	-5.4	4.590	7.733
9	-4.4	6.202	10.793	9	-5.4	6.022	11.480
10	-4.4	3.135	8.860	10	-5.4	3.151	8.443
11	-4.4	6.980	13.187	11	-5.4	3.007	5.039
12	-4.4	22.088	26.207	12	-5.4	3.120	5.509
13	-4.4	4.351	8.987	13	-5.4	5.236	9.073
14	-4.4	4.270	10.052	14	-5.4	3.268	5.433
15	-4.4	6.665	12.425	15	-5.3	3.136	6.350
16	-4.3	5.591	9.289	16	-5.3	2.823	5.979
17	-4.3	2.967	7.966	17	-5.3	4.858	7.482
18	-4.3	3.529	9.170	18	-5.3	3.143	6.387
19	-4.3	3.722	8.755	19	-5.3	5.544	12.189
20	-4.3	3.092	8.211	20	-5.3	4.941	12.715

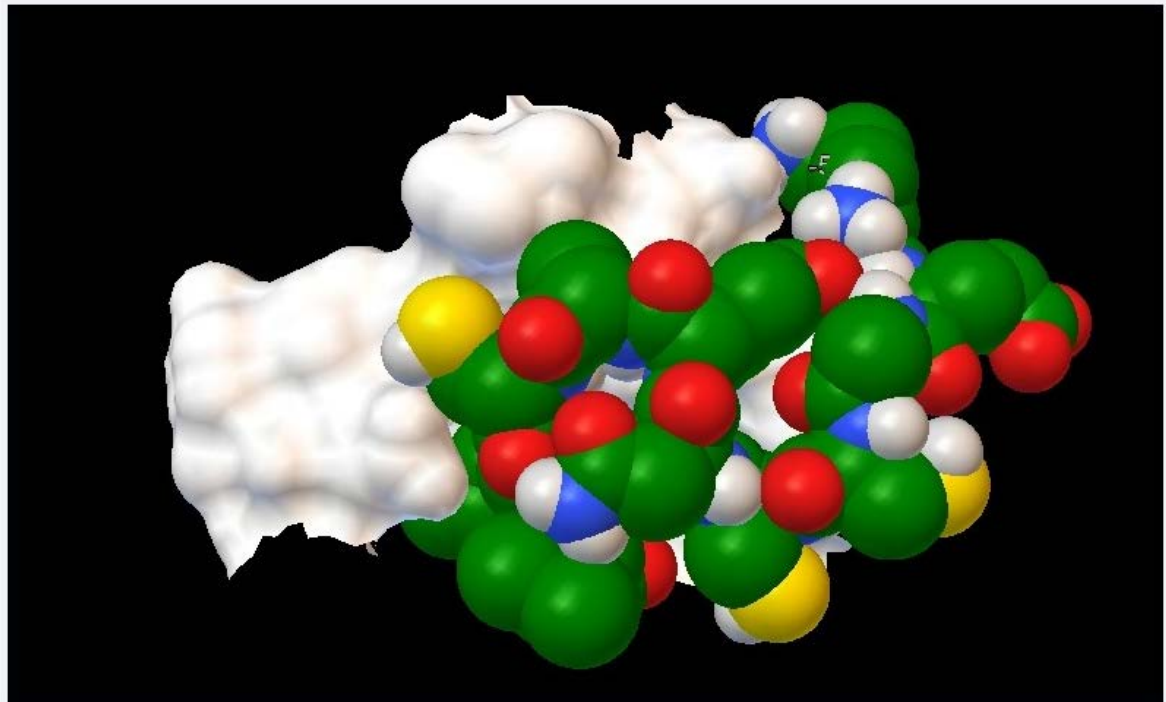
(a)

(b)

Figura 22. Resultado en 3D del acople entre la proteína Ras y el péptido KLGCCNESCKI, en la figura a de color blanco se observa la región activa de la proteína Ras en su estructura terciaria, y el péptido en estructura primaria (conformada de diversos colores). En la figura b se observa el péptido en su estructura terciaria y la proteína Ras en su estructura terciaria. En ambas figuras se observa un buen acople.



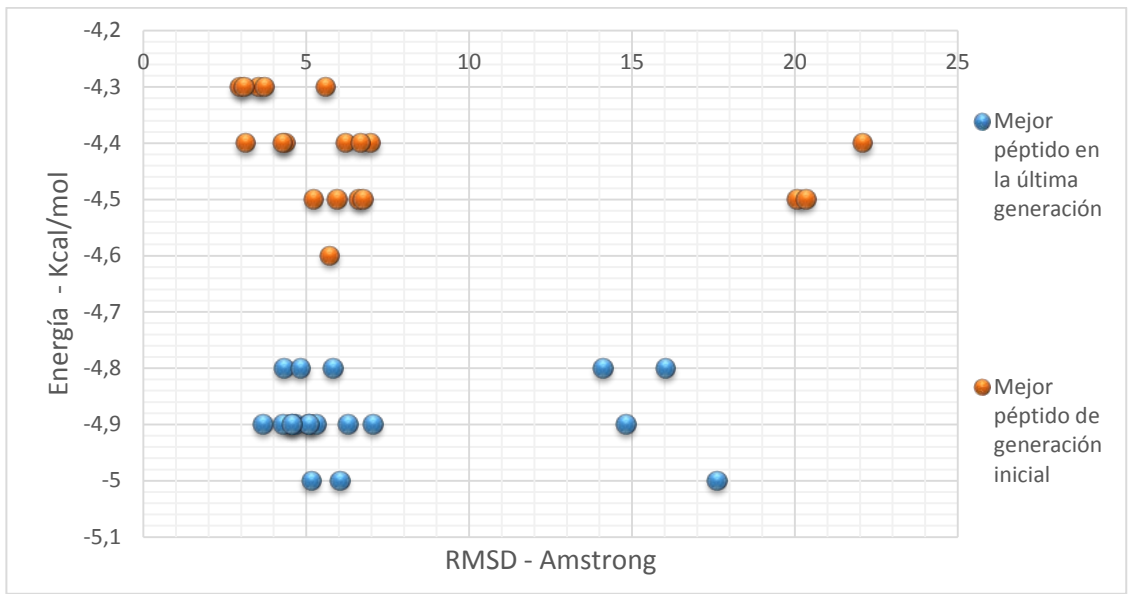
(a)



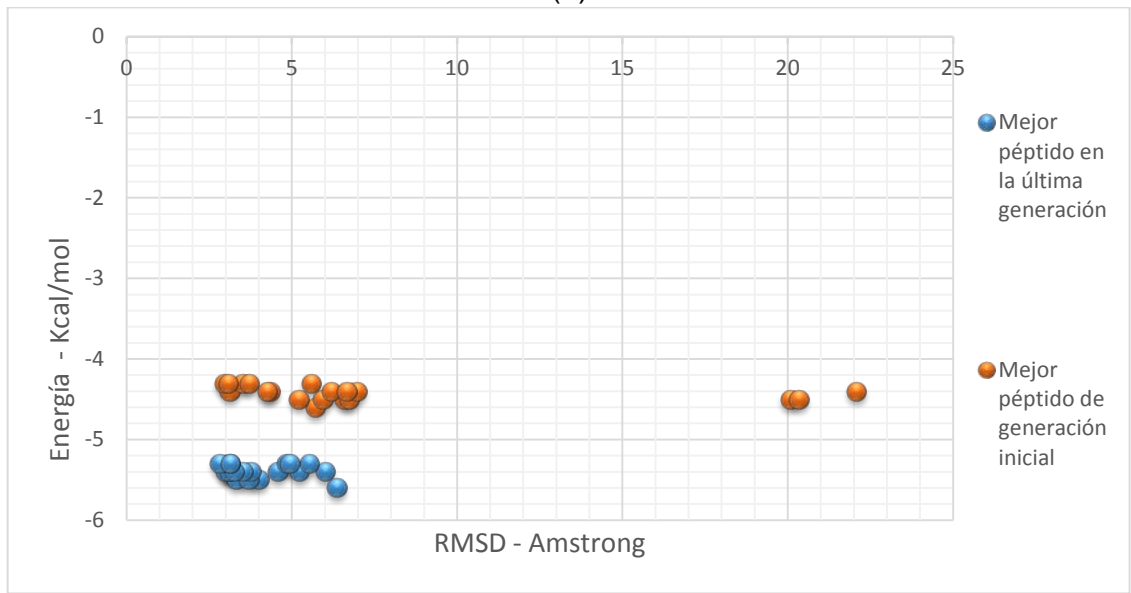
(b)

Claramente se aprecia el acople entre el ligando y el receptor, en donde el ligando cubre gran parte del sitio activo del receptor, lo cual inhabilita la posibilidad fosforilación en la proteína Ras para la proliferación celular. Para poder tener una mejor visión en las diferencias de los resultados obtenidos (ver figura 23).

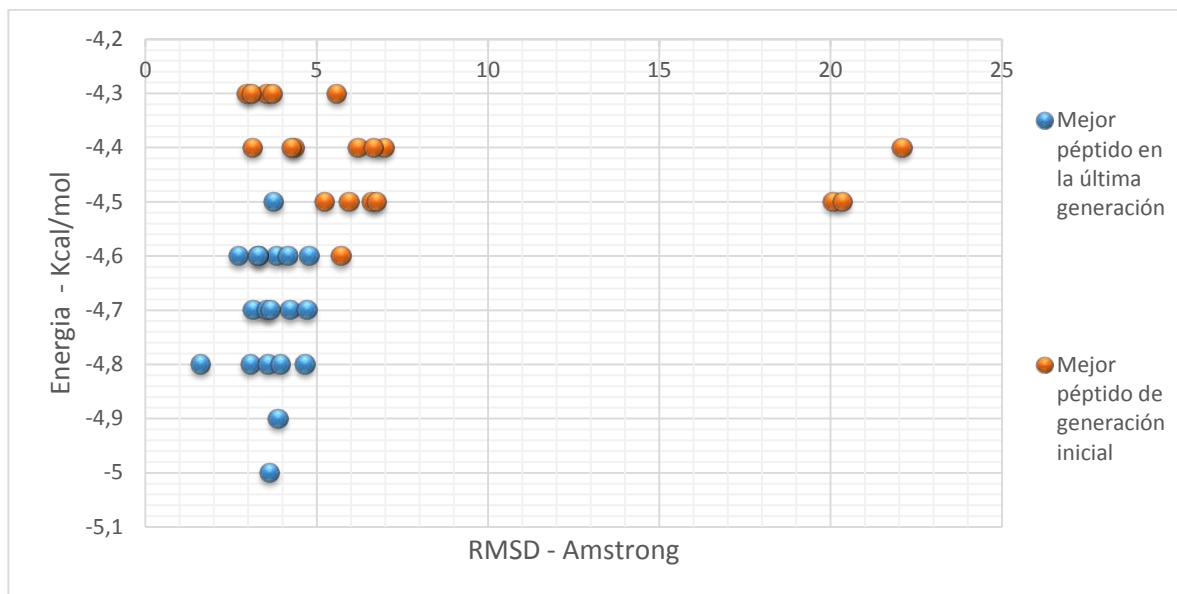
Figura 23. Comparación del acople molecular (docking) entre péptidos de la población inicial contra la población final (Energía -Kcal/mol vs. distancia - Amstrong). Figura a: de color azul el péptido KEGCCNESCKV correspondiente a la población final, de color rojo el péptido CVFICSCSSKV correspondiente a la población inicial. Figura b: de color azul el péptido KLGCCNESCKI correspondiente a la población final, de color rojo el péptido CVFICSCSSKV correspondiente a la población inicial. Figura c: De color azul el péptido KLGCCSECKD correspondiente a la población final, de color rojo el péptido CVFICSCSSKV correspondiente a la población inicial.



(a)



(b)



(c)

En la figura 23, se aprecia la diferencia del acople obtenido entre: Los péptidos iniciales que sirven como población inicial vs. los péptidos resultantes del algoritmo evolutivo. De azul se aprecian los péptidos resultantes del algoritmo, en comparación a los rojos. Se observa una mejoría tanto en la energía de acoplé entre el ligando y receptor, como en la distancia entre ambas, dejando una clara mejoría del acople y confirmando así una evolución.

## 5. CONCLUSIONES

Después del proceso de mejoramiento de los péptidos, se apreció una evolución en los nuevos péptidos, confirmada mediante una simulación de acople efectuada entre los nuevos péptidos y el receptor (proteína Ras).

El método definido para la búsqueda de nuevos péptidos fue exitoso ya que se encontraron nuevos péptidos con actividad potenciada.

La recombinación con la finalidad de buscar nuevos péptidos sin alterar su función aportó resultados positivos.

Con la finalidad de inhibir la reproducción celular descontrolada, se postulan nuevos péptidos como posibles dianas contra el 20% al 30% de los diferentes tipos de cáncer, sin causar efectos secundarios tan fuertes como es el caso de la quimioterapia (ver tabla 4). Estos resultados fueron validados por el experto en el área de diseño de péptidos Rodrigo Gonzalo Torres Saez, quien aprobó los parámetros y avaló su efectividad.

## 6. RECOMENDACIONES.

Sintetizar en laboratorio estos péptidos, puesto que pueden ser una diana potencial contra el cáncer.

El APD [9] [10] es una base de datos que se encuentra actualizándose día a día en búsqueda de nuevos péptidos, lo cual aumenta la población inicial para este método, partiendo de esto, en un futuro se podrá encontrar mayor diversidad en la población inicial lo cual llevaría a mejores resultados.

El algoritmo evolutivo molecular es un modelo de búsqueda guiada hacia nuevos individuos, en este caso péptidos, para guiar esta búsqueda de una mejor forma es fundamental la implementación de patrones de búsqueda. Con la finalidad de mejorar la búsqueda, se recomienda un estudio profundo sobre nuevos patrones.

Una parte muy importante, es el tiempo computacional, lo cual genera una limitación en el número de generación para el algoritmo, lo cual nos puede estar llevando a un resultado parcial (máximo local) que para este trabajo es óptimo, pero al no tener esta limitación en el tiempo computacional, se podría realizar una búsqueda mucho más amplia.

En trabajos futuros, se puede realizar una máquina de soporte vectorial de forma local la cual aceleraría drásticamente el tiempo y predicción de los posibles péptidos, con patrones más afines respecto a los péptidos que se deseen hallar.

Hacer uso de supercomputadoras, con la finalidad de ampliar los patrones que definen la búsqueda, y guiar esta de una mejor forma con una mayor precisión, lo cual podría significar un aumento considerable en los tipos de cáncer a tratar.

## REFERENCIAS BIBLIOGRÁFICAS

- [1] Cairns J. (1981). *CANCER science and society*. Editorial Reverte. pág. 22-23
- [2] Organización Mundial de la Salud (2011). Informe sobre la situación mundial de las enfermedades no transmisibles.
- [3] Giehl K. (2005). Oncogenic Ras in Tumor Progression and Metastasis. *Biological Chemistry* 386(3):193-205, Mar 2005
- [4] Yuan, J., Minter-Dykhouse, K., & Lou, Z. (2009). A c-Myc–SIRT1 feedback loop regulates cell growth and transformation. *The Journal of cell biology*, 185(2), 203-211.
- [5] Shima F., Yoshikawa Y., Ye M., Arakia M., Matsumoto S., Liao J., Hu L., Sugimoto T., Ijiri Y., Takeda A., Nishiyama Y., Sato C., Muraoka S., Tamura A., Osoda T., Tsuda K., Miyakawa T., Fukunishi H., Shimada J., Kumasaka T., Yamamoto M., and Kataoka T.(2013). In silico discovery of small-molecule Ras inhibitors that display antitumor activity by blocking the Ras–effector interaction. Editorial Board April 2.
- [6] Matamoros S. V. (1999) “El diseño de fármacos por ordenador”. Almirall Prodesfarma.
- [7] Palanques, R. M. Capítulo III: El desarrollo de los nuevos fármacos en la era postgenómica. *Genoma y Medicina*, 37.
- [8] Damodaran, S. (2000). Aminoácidos, péptidos y proteínas. *Química de alimentos*, 2, 490.
- [9] Correa Gómez, E. (2013). Evaluación de péptidos sintéticos sobre bacterias, levaduras y mohos (Doctoral dissertation, Universidad Nacional de Colombia, Sede Medellín).
- [10] Wang, Z. and Wang, G. (2004) APD: the antimicrobial peptide database. *Nucleic Acids Research* 32, D590-D592.

- [11] Damodaran, S. (2000). Aminoácidos, péptidos y proteínas. *Química de alimentos*, 2, 490.
- [12] Wimley, W. C. (2010). Describing the Mechanism of Antimicrobial Peptide Action with the Interfacial Activity Model. *ACS Chemical Biology*, 5(10), 905–917
- [13] Yeaman, M. R., & Yount, N. Y. (2003). Mechanisms of antimicrobial peptide action and resistance. *Pharmacological Reviews*, 55(1), 27–55.
- [14] Zhao, H. (2003). Mode of Action of Antimicrobial Peptides. University of Helsinki.
- [15] Marshall, S. H., & Arenas, G. (2003). Antimicrobial peptides: A natural alternative to chemical antibiotics and a potential for applied biotechnology. *Electronic Journal of Biotechnology*, 6(3).
- [16] Wang, G. (2010). Antimicrobial peptides: discovery, design and novel therapeutic strategies. (G. Wang, Ed.) (1st ed.). Wallingford: Cabi.
- [17] Castañeda-Casimiro, J., Ortega-Roque, J. A., Venegas-Medina, A. M., Aquino-Andrade, A., Serafín-López, J., Estrada-Parra, S., & Estrada, I. (2009). Péptidos antimicrobianos: péptidos con múltiples funciones. *Alergia, Asma e Inmunología Pediátricas*, 18(1), 16-29
- [18] Lee, S.-H., Kim, S.-J., Lee, Y.-S., Song, M.-D., Kim, I.-H., & Won, H.-S. (2011). De novo generation of short antimicrobial peptides with simple amino acid composition. *Regulatory Peptides*, 166(1-3)
- [19] Smith T.F., Waterman M.S.(1981) Identification of common molecular subsequences. *J. Mol. Biol.* 147:195–197.
- [20] Needleman S.B., Wunsch C.D.(1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48:443–453.

- [21] Edgar, R. C., & Batzoglou, S. (2006). Multiple sequence alignment. *Current opinion in structural biology*, 16(3), 368-373.
- [22] Castrillón, O. D., Sarache, W. A., & Giraldo, J. A. (2011). Aplicación de un algoritmo evolutivo en la solución de problemas Job Shop-Open Shop. *Información tecnológica*, 22(1), 83-92.
- [23] Zelezetsky, I., & Tossi, A. (2006). Alpha-helical antimicrobial peptides--using a sequence template to guide structure-activity relationship studies. *Biochimica et Biophysica Acta*, 1758(9), 1436-49.
- [24] Sobolev, V., Wade, R. C., Vriend, G., & Edelman, M. (1996). Molecular docking using surface complementarity. *Proteins: Structure, Function, and Bioinformatics*, 25(1), 120-129.
- [25] Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21), 2947-2948.
- [26] Clamp, M., Cuff, J., Searle, S. M., & Barton, G. J. (2004). The jalview java alignment editor. *Bioinformatics*, 20(3), 426-427.
- [27] Tyagi, A., Kapoor, P., Kumar, R., Chaudhary, K., Gautam, A., & Raghava, G. P. S. (2013). In silico models for designing and discovering novel anticancer peptides. *Scientific reports*, 3.
- [28] Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2), 455-461.
- [29] DeLano, W. L. (2002). The PyMOL molecular graphics system.
- [30] Fainstein, M. K. (2007). Nrf2: la historia de un nuevo factor de transcripción que responde a estrés oxidativo. *Revista de Educación Bioquímica*, 26(001), 18-25.)

- [31] Sah, B. N. P., Vasiljevic, T., McKechnie, S., & Donkor, O. N. (2015). Identification of Anticancer Peptides from Bovine Milk Proteins and Their Potential Roles in Management of Cancer: A Critical Review. *Comprehensive Reviews in Food Science and Food Safety*, 14(2), 123-138.
- [32] Sharma, A., Singla, D., Rashid, M., & Raghava, G. P. (2014). Designing of peptides with desired half-life in intestine-like environment. *BMC bioinformatics*, 15(1), 282.
- [33] Yokobayashi, Y., Ikebukuro, K., McNiven, S., & Karube, I. (1996). Directed evolution of trypsin inhibiting peptides using a genetic algorithm. *J. Chem. Soc., Perkin Trans. 1*, (20), 2435-2437.
- [34] Chittenden, T., Flemington, C., Houghton, A. B., Ebb, R. G., Gallo, G. J., Elangovan, B., ... & Lutz, R. J. (1995). A conserved domain in Bak, distinct from BH1 and BH2, mediates cell death and protein binding functions. *The EMBO Journal*, 14(22), 5589.

## BIBLIOGRAFIA

CALZA, S., & Pawitan, Y. (2010). Normalization of gene-expression microarray data. *Methods in Molecular Biology* (Clifton, N.J.), 673, 37–52. doi:10.1007/978-1-60761-842-3

GARBAY, C., Liu, W. Q., Vidal, M., & Roques, B. P. (2000). Inhibitors of RAS signal transduction as antitumor agents. *Biochemical Pharmacology*, 60(8), 1165–1169. doi:10.1016/S0006-2952(00)00428-7

RÖCKENDORF, N., Borschbach, M., & Frey, A. (2012). Molecular Evolution of Peptide Ligands with Custom-Tailored Characteristics for Targeting of Glycostructures. *PLoS Computational Biology*, 8(12). doi:10.1371/journal.pcbi.1002800

SHIMA, F., Yoshikawa, Y., Ye, M., Araki, M., Matsumoto, S., Liao, J., ... Kataoka, T. (2013). In silico discovery of small-molecule Ras inhibitors that display antitumor activity by blocking the Ras-effector interaction. *Proc Natl Acad Sci U S A*, 110(20), 8182–8187. doi:10.1073/pnas.1217730110

ZENG, J., & Treutlein, H. R. (1999). A method for computational combinatorial peptide design of inhibitors of Ras protein. *Protein Engineering*, 12(6), 457–468.

ZENG, J., Nheu, T., Zorzet, a, Catimel, B., Nice, E., Maruta, H., ... Treutlein, H. R. (2001). Design of inhibitors of Ras--Raf interaction using a computational combinatorial algorithm. *Protein Engineering*, 14(1), 39–45.