

PÉPTIDOS INMUNOGÉNICOS CONTRA EL VIRUS DISTEMPER CANINO (CDV)

CAMPO TÉCNICO

5 El presente desarrollo se encuentra en el campo de la inmunología y medicina veterinaria, en particular se relaciona con el diseño de péptidos inmunogénicos contra el Virus Distemper Canino y composiciones que los comprenden, útiles para la generación de inmunidad frente a dicho virus en fauna silvestre y doméstica.

10 DESCRIPCIÓN DEL ESTADO DE LA TÉCNICA

El Virus del Distemper Canino (CDV por su sigla en inglés Canine Distemper Virus), es el agente etiológico de una enfermedad altamente contagiosa conocida como enfermedad de Carré o moquillo, que afecta perros domésticos, así como una gran variedad de especies silvestres. El distemper es una enfermedad de distribución mundial que produce
15 síntomas respiratorios, digestivos, cutáneos y neurológicos entre otros. No existe un tratamiento específico para la enfermedad y la mayoría de los esfuerzos se concentran en la prevención mediante la administración de dos o más dosis de una vacuna entre la sexta o séptima semana de edad hasta los tres o cuatro meses, seguida de re-vacunación anual durante toda la vida del animal.

20

El CDV tiene una alta tasa de sustitución genómica, lo que hace que las variantes circulantes en campo difieran en más del 10% frente a las cepas utilizadas en vacunas. Esto implica consecuencias en el desarrollo de la inmunidad y la aparición de la enfermedad incluso en animales previamente vacunados, además de la re-emergencia
25 mundial de la infección en fauna silvestre, para la cual no se ha comprobado la seguridad de las vacunas disponibles comercialmente.

El CDV pertenece al género *Morbillivirus* de la familia paramyxovirus y contiene un genoma RNA de cadena sencilla con polaridad negativa que codifica 8 proteínas,
30 incluidas la Hemaglutinina (H) y la proteína de Fusión (F). La glicoproteína H posee gran variación genética, donde se estima que las variantes circulantes en campo difieren en un 10% con las utilizadas en las vacunas convencionales. Estas variaciones

nucleotídicas pueden llevar a cambios en la secuencia de aminoácidos de la proteína H y, por ende, implicar modificaciones en su estructura, los cuales pueden traer consecuencias en la antigenicidad de esta, ya que es considerada uno de los principales determinantes antigénicos de CDV. El efecto de la variabilidad de esta proteína entre las diferentes
5 variantes de CDV distribuidas a nivel mundial, trae consigo la necesidad de desarrollar estrategias vacunales que permitan una cobertura universal. Por tal razón, es preponderante la determinación de secuencias consenso de los epítopes inmunodominantes, con el fin de obtener dichas coberturas.

- 10 Por lo anterior, se considera que el diseño de péptidos mediante herramientas de predicción de epítopes inmunogénicos con base en secuencias consenso, que tome como suministro la información genética y antigénica de las variantes circulantes alrededor del mundo, puede dar como resultado una alternativa inmunogénica efectiva y segura para el desarrollo de una vacuna universal que genere inmunidad contra los diferentes linajes del
15 virus, que pueda ser aplicada a las especies susceptibles tanto de fauna doméstica como silvestre.

- En el estado de la técnica se han descrito composiciones inmunogénicas contra el CDV que incluyen aislamientos atenuados del CDV, ácidos nucleicos que codifican proteínas
20 antigénicas y vectores que expresan antígenos del CDV, entre otras estrategias. A la fecha, se cuenta con algunas alternativas vacunales. Se conocen vacunas de tipo de virus vivo modificado (MLV, del inglés, *modified live virus*), la cual fue propuesta en la década de 1960, con base en una cepa ancestral, denominada Onderstepoort. Por tanto, considerando que esta variante incluso ya no circula en campo, la inmunidad generada
25 por esta vacuna pueda disminuir para los demás linajes distribuidos a nivel mundial, además que no pueden ser utilizadas en ninguna especie para la cual no ha sido probada la seguridad, por tener la atenuación como plataforma de obtención. Existe una vacuna recombinante, que utiliza un esqueleto de virus de la viruela del canario (Canary pox), que expresa las proteínas H y F de CDV, la cual protege contra el desarrollo de distemper
30 sintomático. Además, se cuenta con una plataforma de nueva generación, considerada una vacuna bivalente recombinante, que emplea el virus de la rabia, que expresa las proteínas H y F, la cual ha sido probada en caninos domésticos y hurones. Uno de los

problemas que pueden enfrentar todas las plataformas vacunales disponibles, es la emergencia de variantes distribuidas geográficamente, las cuales poseen diferencias de aminoácidos en los principales determinantes antigénicos, lo cual puede llevar a la generación de una respuesta inmunológica insuficiente frente a las variantes de campo.

5

En el estado de la técnica se describen vacunas contra el CDV. Por ejemplo, la patente US8647637 describe composiciones inmunogénicas y vacunas de amplio espectro que contienen aislamientos recientemente identificados del CDV, recolectados en un área geográfica determinada. Estos aislamientos exhiben atributos tanto del linaje silvestre europeo como de uno o ambos de los linajes ártico y americano-2 del CDV. Adicionalmente, esta patente se relaciona con vacunas basadas en ácidos nucleicos que codifican péptidos antigénicos o proteínas, o epítopes antigénicos o regiones de péptidos o proteínas.

10

15 Por otra parte, la patente US10076566 describe vectores que contienen y expresan polipéptidos o antígenos CDV *in vivo* o *in vitro*, los cuales generan una respuesta inmune en un animal contra el CDV. También se ilustran composiciones que comprenden dichos vectores y/o polipéptidos y métodos de vacunación contra el CDV, así como métodos para inducir una respuesta inmunogénica o protectora contra el CDV y otros virus caninos.

20

Sin embargo, persiste la necesidad de una alternativa para generar inmunidad contra la mayor cantidad de linajes del virus en una vacuna universal que pueda ser aplicada tanto en fauna silvestre como doméstica y sea eficaz y segura, pues los péptidos sintéticos de tamaño inferior a 20 aminoácidos han demostrado ser seguros en diferentes modelos *in vitro* e *in vivo*. Además, una vacuna que tenga como inmunógeno una subunidad (péptidos inmunogénicos), no posee la capacidad replicativa que, si tienen las vacunas de virus vivo modificado o recombinante, para las cuales es necesario probarles su seguridad en cada una de las especies.

25

30

BREVE DESCRIPCIÓN

El presente desarrollo se refiere a péptidos inmunogénicos contra el CDV, diseñados mediante herramientas computacionales y composiciones inmunogénicas que comprenden dichos péptidos, útiles como una alternativa de vacuna universal para prevenir el contagio por CDV en fauna silvestre y doméstica.

BREVE DESCRIPCIÓN DE LAS FIGURAS

FIG. 1 Representación esquemática del procedimiento para el diseño de péptidos inmunogénicos contra el CDV.

FIG. 2 **A.** Epítopes obtenidos de la proteína H de CDV consenso; **B.** Epítopes obtenidos de la proteína F de CDV consenso.

FIG. 3 Tamaño de péptidos predichos de la secuencia consenso de las proteínas de CDV. **A.** Proteína H; **B.** Proteína F

DESCRIPCIÓN DETALLADA

Para propósitos de interpretar los términos usados a lo largo del presente documento se debe tener en cuenta su significado usual en el campo técnico, a menos que se incorpore una definición particular o el contexto indique claramente lo contrario. Adicionalmente, los términos utilizados en forma singular también incluirán la forma plural.

El uso de herramientas computacionales se ha convertido en una pieza clave para la investigación y el desarrollo de vacunas para diversos modelos de infección, entre los cuales se incluyen las terapias basadas en péptidos.

El presente desarrollo se basa en la evaluación del potencial inmunogénico de péptidos diseñados a partir de la información genética de cepas de CDV circulantes en campo, mediante herramientas computacionales. Para el diseño de los péptidos inmunogénicos contra el CDV de la presente divulgación, inicialmente se generaron secuencias consenso

para las proteínas H y F del CDV, las cuales son consideradas, los principales determinantes antigénicos del virus, sumado a su localización en la cápsula viral, las cuales son blanco de la inmunidad humoral, mediada por los anticuerpos, con base en la secuencia de todos los linajes reportados para este virus a nivel mundial.

5

Posteriormente, estas secuencias se sometieron a varias herramientas de predicción de epítopes inmunogénicos. De un total de 1402 péptidos, predichos para moléculas del complejo mayor de histocompatibilidad (CMH) clase 1, clase 2 y células B, con diferentes herramientas computacionales como MHC2Pred, SVMTriP, CTLPred e IEDB de la Jolla se eligieron los que mostraron mejor valor de predicción positivo como potenciales antígenos y sus características fisicoquímicas favorables.

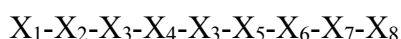
10

Luego, se evaluó la capacidad que estos tienen para interactuar con moléculas del CMH de caninos, así como con receptores tipo toll, mediante acoplamiento molecular y finalmente, se evaluó la seguridad *in silico*, mediante diferentes herramientas para determinar si estos eran buenos antígenos, podían ser tóxicos o alérgenos, además, si existían péptidos propios de los caninos que fueran homólogos a estos.

15

De acuerdo con el procedimiento de diseño descrito anteriormente, se seleccionaron péptidos inmunogénicos contra el CDV que tienen la secuencia

20



en donde,

X_1 es un aminoácido seleccionado entre un aminoácido cargado positivamente, un aminoácido no cargado, un aminoácido especial y un aminoácido hidrofóbico;

25

X_2 es tres aminoácidos iguales o diferentes, seleccionados entre un aminoácido cargado positivamente, un aminoácido cargado negativamente, un aminoácido no cargado, un aminoácido especial, un aminoácido hidrofóbico y un aminoácido aromático;

X_3 es un aminoácido, seleccionado entre un aminoácido cargado positivamente, un aminoácido no cargado, un aminoácido especial, un aminoácido hidrofóbico y un aminoácido aromático;

30

X₄ es dos aminoácidos iguales o diferentes, seleccionados entre un aminoácido cargado positivamente, un aminoácido cargado negativamente, un aminoácido no cargado, un aminoácido hidrofóbico y un aminoácido aromático;

5 X₅ es un aminoácido seleccionado entre un aminoácido cargado positivamente, un aminoácido cargado negativamente, un aminoácido especial, un aminoácido hidrofóbico y un aminoácido aromático;

X₆ está ausente o es un aminoácido seleccionado entre un aminoácido cargado positivamente, un aminoácido no cargado, un aminoácido hidrofóbico y un aminoácido aromático;

10 X₇ es X₁ o está ausente; y

X₈ está ausente o es uno o dos aminoácidos iguales o diferentes seleccionados entre un aminoácido cargado positivamente y un aminoácido hidrofóbico,

en donde

15 el aminoácido cargado positivamente se selecciona del grupo que consiste en R, H y K;

el aminoácido cargado negativamente se selecciona del grupo que consiste en D y E;

el aminoácido no cargado se selecciona del grupo que consiste en S, T, N y Q;

el aminoácido especial se selecciona del grupo que consiste en C, G y P;

20 el aminoácido hidrofóbico se selecciona del grupo que consiste en A, I, L, M y V, y

el aminoácido aromático se selecciona del grupo que consiste en W, Y y F.

25 En una modalidad, los péptidos de la presente divulgación incluyen modificaciones amino terminal tales como biotinilación, acetilación, ácidos grasos para la obtención de lipopéptidos y PEGilación.

En otra modalidad, los péptidos incluyen modificaciones carboxilo terminal tales como la amidación.

30

En un aspecto particular, la presente divulgación se refiere a los péptidos inmunogénicos contra el CDV que se enseñan en la Tabla 1:

Tabla 1. Secuencias de aminoácidos de péptidos inmunogénicos diseñados

PÉPTIDO	SECUENCIA	SEC ID NO.
PCI-H1-004	PYLLFVLLI	1
PCI-H1-030	QVIDVLTPLFK	2
PCII-H1-008	VENLVRIRF	3
PCI-F3-053	LTGTSSSYVRSL	4
PCII-F43-012/PCI-F3-038	LKLLRYYTE	5
PCI-I6EDB-H1	TVIPHVLTf	6
LINB-7IEDB-H	VSEKQEEQKNC	7
PCI-F83-054	PAFKPDLTGTSK	8
PCII-IE9DB-H2	PPYLLFVLLILLV	9
PCI-IED10B-F1	MVYESKVAL	10
PCII-IEDB11-F2	GNKFILSKGNIVA	11
LINB-IEDB12-F	KAQIHWNNL	12

5

La presente divulgación también se refiere a las moléculas de ácidos nucleicos que codifican los péptidos inmunogénicos contra el CDV que se enseñan en la Tabla 2.

Tabla 2. Secuencias de nucleótidos que codifican los péptidos inmunogénicos diseñados

10

PÉPTIDO	ÁCIDO NUCLEICO	SEC ID NO.
PCI-H1-004	CCTTATCTTCTTTTTGTTCTTCTTATT	13
PCI-H1-030	CAACTTATTGATGTTCTTACTCCTCTTTTAAA	14

PCII-H1-008	GTTGAAAATCTTGTTCGTATTCGTTTT	15
PCI-F3-053	CTTACTGGTACTTCTAAATCTTATGTTCGTTCTCTT	16
PCII-F43-012/PCI-F3-038	CTTAAACTTCTTCGTTATTATACTGAA	17
PCI-I6EDB-H1	ACTGTTATTCCTCATGTTCTTACTTTT	18
LINB-7IEDB-H	GTTTCTGAAAAACAAGAACAACAAAAAATTGT	19
PCI-F83-054	CCTGCTTTTAAACCTGATCTTACTGGTACTTCTAAA	20
PCII-IE9DB-H2	CCTCCTTATCTTCTTTTTGTTCTTCTTATTCTTCTTGTT	21
PCI-IED10B-F1	ATGGTTTATGAATCTAAAGTTGCTCTT	22
PCII-IEDB11-F2	GGTAATAAATTTATTCTTTCTAAAGGTAATATTGTTGCT	23
LINB-IEDB12-F	AAACGTCAAATTCATTGGAATAATCTT	24

En un aspecto adicional, la presente divulgación se refiere a una composición inmunogénica que comprende uno o más péptidos inmunogénicos contra el CDV y excipientes farmacéuticamente aceptables. En una modalidad particular, la composición inmunogénica es una vacuna universal que puede ser aplicada a especies de fauna doméstica y silvestre susceptibles al CDV de manera eficaz y segura.

Los excipientes de la composición inmunogénica se seleccionan, sin limitarse al grupo que consiste en preservantes, amortiguadores, estabilizantes, sales y surfactantes.

En una modalidad particular, como preservante se emplea, sin limitarse a alcohol bencílico. En otra modalidad particular, el preservante está en una concentración entre 20 y 30 µg/ml.

En una modalidad particular, como amortiguador se emplea, sin limitarse a buffer de fosfatos. En otra modalidad particular el amortiguador está en una concentración entre 5 y 100 nM.

En una modalidad particular, como estabilizante se emplea, sin limitarse a glicerol. En otra modalidad particular el estabilizante está en una concentración entre 0 y 10%.

En una modalidad particular, como sal se emplea, sin limitarse a hidróxido de aluminio.

5 En otra modalidad particular la sal está en una concentración entre 0 y 300 nM.

En una modalidad particular, como surfactante se emplea, sin limitarse a polisorbato 20-80. En otra modalidad particular el surfactante está en una concentración entre 0,01 y 0,1 % p/v.

10

La presente invención será presentada en detalle a través de los siguientes ejemplos, los cuales son suministrados solamente con propósitos ilustrativos y no con el objetivo de limitar su alcance.

15 **EJEMPLOS**

Ejemplo 1. Secuencias para las proteínas H y F del CDV

Las secuencias de las proteínas H y F de CDV de todos los linajes, reportados por Duque-Valencia y colaboradores (Duque-Valencia, J., Forero-Muñoz, N.R., Díaz, F.J. *et al.* Phylogenetic evidence of the intercontinental circulation of a Canine distemper virus lineage in the Americas. *Sci Rep* **9**, 15747 (2019). <https://doi.org/10.1038/s41598-019-52345-9>), se descargaron en formato fasta traducido a los aminoácidos, con el fin de construir archivos multifasta, alineamientos múltiples y secuencias consenso, con el fin de determinar el potencial inmunogénico de las proteínas H y F del CDV. Posteriormente todas las secuencias de las proteínas H y F del CDV descargadas, pertenecientes a los linajes reportados a nivel mundial, se convirtieron en único archivo multifasta. Adicionalmente, se realizó un archivo multifasta de las secuencias que circulan en Colombia.

30

Por otra parte, se descargaron secuencias de moléculas del CMH, clase I y II de canino y estructuras molde de humano de UniProt y Protein Data Bank (PDB), para ser modeladas con base en la estructura cristalográfica de una proteína de humano.

5 Ejemplo 2. Generación de secuencias consenso

Se realizaron alineamientos múltiples de las secuencias de la proteína H y F de CDV mediante la herramienta en línea Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>), con el propósito de establecer las diferencias entre cada una de las secuencias de esta proteína y los linajes que circulan a nivel mundial. Además, la generación de la matriz p para poder cuantificar las diferencias mediante Mega 10. Por otra parte, se generó un alineamiento para la proteína H y F de CDV de las variantes de los linajes que co-circulan en Colombia. Finalmente, se generaron 4 secuencias consenso mediante la herramienta en línea https://www.ebi.ac.uk/Tools/msa/emboss_cons/, con las cuales se realizaron los análisis de predicción.

Las secuencias obtenidas son:

Con-1. Consenso de la secuencia de la proteína H tomando como base todas las secuencias de los linajes circulantes a nivel mundial (SEC ID NO: 25):

>CDV H protein consensus all linages

MSLYQDKVGAFYKDNARANSSKLSLVTEEQGGRRPPYLLFVLLILLVGIMALLAITGVRFHQVST
SNMEFSRLLKEDMEKSEAVHHQVIDVLTPLFKIIGDEIGLRLPQKLNEIKQFILQKTNFFNPNNREFD
FRDLHWCINPPSKIKVNFTNYCDTIGIRKSIASAANPILLSALSGGRGDIFPPYRCSGATTSVGRVFP
LSVSLSMSLISRTSEIINMLTAISDGVYGKTYLLVPDYIEGEFDTQKIRVFEIGFIKRWLNDMPLLQ
TTNYMVLPEKSKAKVCTIAGELTLASLCVDESTVLLYHDSNGSQDGILVVTLGIFGATPMDQV
EEVIPVAHPSVEKIHITNHRGFIKDSIATWMVPALVSEKQEEQKNCLESACQRKSYPMCNQTSWE
PFGGGQLPSYGRLLPLDPSIDLQLNISFTYGPVILNGDGMDYYESPLLDGWLTPPKNGTVLGLI
NKASRGDQFTVIPHVLTAFRESSGNCYLPIQTSQIMDKDVLTESNLVVLPTQNFRYVIATYDISR
DDHAIVYYVYDPIRTISYTPFRLTTKGRPDFLRIECFVWDDDLWCHQFYRFEADITNSTTSVENL
VRIRFSCNRSKP

Con-3. Consenso de la secuencia de la proteína F tomando como base todas las secuencias de los linajes circulantes a nivel mundial (SEC ID NO: 26):

>CDV F protein consensus all linages

MHNKIPKRSKTQTHTQQDLPQQHSTKSAETKTSQARHSTTSAQRSTHHGPRTSDRPVHYIMNRT
RSCKQASHRSDNIPAHGDHEGIIHHTPESVSQGARSRPKRRQSNATNSGSQCTWLVLWCIGIASLF
LCSKAQIHWNNLSTIGIIGTDSVHYKIMTRPSHQYLVIKLMPNVSLIDNCTKAELGEYEKLLNSVL
5 EPINQALTLMTKNVKPLQSVGSGRRQRRFAGVVLAGAALGVATAAQITAGIALHQSNLNAQAIQ
SLRTSLEQSNKAIEEIREATQETVIAVQGVQDYVNNELVPAMQHMSCELVGQRLGLKLLRYYTE
LLSIFGPSLRDPISAEISIQALSALGGEIHKILEKLGYSNDMIAILESRIKTKITHVDLPGLIILSI
SYPTLSEVKGIVVHRLEAVSYNIGSQEWYTTVPRYVATNGYLISNFDDESSCVFVSESAICSQNSLY
PMSPLLQQCIRGDTSSCARTLVSGTMGNKFILSKGNIVANCASILCKCYSTSTIINQSPDKLLTFIAS
10 DTCPLVEIDGVTIQVGGRRQYPDMVYESKVALGPAISLERLDVGTNLGNALKKLDDAKVLIDSSN
QILETVRRSSFNFGSLLSVPILICTALALLLIYCKRRYQQTLKQNTKVDPTFKPDLTGTSKSYVR
SL

Se seleccionaron 168 péptidos para ambas secuencias con las secuencias consenso de
15 todos los linajes para análisis posterior.

Ejemplo 3. Construcción de librería de péptidos potencialmente inmunogénicos de secuencias consenso.

20 a- Predicción para linfocitos T citotóxicos (CTL) con la herramienta CTL
pred: <http://crdd.osdd.net/raghava/ctlpred/index.html>

La herramienta para CTL integra diferentes herramientas de predicción, con bases de
datos que reportan epítopes para diferentes alelos de antígenos de leucocitos humanos
25 (HLA por su sigla en inglés – human leukocyte antigens) clase I. Esta herramienta, genera
los péptidos con base en diferentes aproximaciones, que toman en consideración
diferentes modelos matemáticos, Support Vector Machine (SVM) con un cut off de 1.2
(máximo 1.5), Artificial Neural Network (ANN), con cut off de 0.95 (valor entre 0 y 1) y
una combinación de estos modelos, lo cual permite la obtención de los potenciales
30 inmunógenos.

Se obtuvieron 22 péptidos ANN, 6 péptidos SVM y en combinación, 28 péptidos para
Con-1 y 43 péptidos ANN, 8 péptidos SVM y en combinación, 50 péptidos para Con-3.

b- Determinación de péptidos lineales para linfocitos B con la herramienta <http://sysbio.unl.edu/prediction.php>

5 Esta herramienta permite determinar péptidos lineales para linfocitos B, los cuales permiten la detección directa de antígenos. Esta herramienta tiene un límite en la búsqueda, puesto que tiene una ventana de péptidos específica. En este caso se predijeron los de 10 aminoácidos. Esta herramienta puede predecir péptidos de 12, 14, 16, 18 y 20 aminoácidos.

10 Con esta herramienta, se encontró un set de péptidos para linfocitos B de ambas proteínas, H y F de CDV, los cuales pueden hacer parte de la librería de péptidos. Con-1 tiene 4 secuencias de 10 aminoácidos, que tienen el potencial de ser epítopes para linfocitos B. Con-3 tiene 3 secuencias de 10 aminoácidos, que tienen el potencial de ser epítopes para linfocitos B.

15

c- Epítopes para los diversos alelos humanos de las moléculas del CMH clase II, con la herramienta <http://crdd.osdd.net/raghava/mhc2pred/index.html>

20 Esta herramienta permite determinar epítopes para los diversos alelos humanos de las moléculas del CMH-II. Para cada alelo se marcó el top 10, sin embargo, el valor umbral mayor a 1.5 se consideró para péptidos potenciales, con la intención de ser más astringentes.

25 Para algunos de los alelos ninguno de los péptidos superó el valor umbral, por lo cual se hizo una depuración y extracción de los péptidos potenciales.

d- Epítopes para linfocitos B, con la herramienta <https://webs.iitd.edu.in/raghava/lbtope/protein.php>

30 Esta herramienta permite predecir epítopes para linfocitos B mediante una escala de colores en donde se establece la probabilidad de los péptidos y los resultados obtenidos. Los epítopes de cada una de las secuencias consenso se basan en este código de colores.

Los epítopes obtenidos de las proteínas H y F de CDV consenso se ilustran en la Figura 2.

e- Recuperación de péptidos de DLA clase I

Se presentaron 39 diferentes péptidos recuperados de la molécula DLA-88*508:01, 2 de la proteína H y 4 para la proteína F. Los péptidos para H y F de CDV se muestran en la Tabla 3:

Tabla 3. Péptidos para las proteínas H y F

Proteína	Secuencia	Tamaño	Posición
Proteína H			
1	AISDGVYGK	9	222
2	QVIDVLTPLFK	11	87
Proteína F			
1	LTGTSKSYVR	10	651
2	PAFKPDLTGTS	11	645
3	LTGTSKSYVRSL	12	651
4	PAFKPDLTGTSK	12	645

Todos estos péptidos fueron recuperados de células infectadas con la variante Ondertepoort.

Ejemplo 4. Depuración y análisis de péptidos para cálculo de propiedades fisicoquímicas

Con el fin de establecer el potencial de cada uno de los péptidos obtenidos y continuar con el estudio *in silico*, se realizó una base de datos con toda la información de los péptidos obtenidos mediante las diferentes herramientas antes mencionadas y así poder hacer una depuración sistemática de la librería para seleccionar los mejores candidatos.

Allí se consigna: 1. Código de péptido. 2. Secuencia. 3. Tamaño. 4. Región del genoma. 5. Tipo de epítipo. 6. Herramienta de obtención. 7. En adelante, propiedades físicoquímicas como: carga, peso molecular, punto isoelectrico, índice de estabilidad, GRAVY, índice alifático y vida media estimada, calculadas con ProtParam (<https://web.expasy.org/protparam/>).

Ejemplo 5. Modelación de algunas moléculas inmunes

Se modeló la cadena alpha de la secuencia de una CMH-I de canino (DLA-I-88), (modelo_1), puesto que esta es la única que interactúa con los péptidos con potencial inmunogénico. El molde utilizado es una CMH-I de canino depositado en el PDB, 5F1N, la cual proviene de una cristalización de moléculas de canino. Por otra parte, se tiene la secuencia de la una CMH-II de canino del alelo DR (DLA-DR), tanto las cadenas alpha y beta (Modelo_2 y Modelo_3, respectivamente). Para esta estructura, se utilizó como molde, la estructura 3D reportada en el PDB, 4FQX, tanto para la cadena alpha como la beta. Esta estructura, es de humanos, del mismo alelo de la secuencia que se tiene para caninos (HLA-DR). Además, se realizó el cálculo de la identidad de los moldes con respecto a cada una de las moléculas modeladas con base en su secuencia. Se realizaron modelaciones por homología mediante el software Modeller 9.24.

Luego de la corrida de cada una de las tres moléculas diferentes, se generaron 100 modelos de cada una, donde se selecciona el mejor modelo con base en tres puntajes que entrega el software: molpdf, DOPE y GA341.

Mejores modelos:

Modelo_1: Modelo 57, el cual obtuvo el mejor puntaje.

Modelo_2: Modelo 76, el cual obtuvo el mejor puntaje.

Modelo_3: Modelo 7, el cual obtuvo el mejor puntaje.

El Modelo_2 y Modelo_3 se pusieron en un archivo como complejo con base en la estructura molde, con el fin de evaluar el docking molecular. Esta estrategia permite

determinar la capacidad de dos moléculas para interactuar, mediante diferentes enlaces no covalentes como los puentes de hidrogeno y las fuerzas de Van de Waals. Para el caso de las herramientas en línea utilizadas, la predicción se hace sin delimitar un sitio de interacción, en cambio, las herramientas presentan la región de interacción más probable, con base en algoritmos que pueden predecir esa pose molecular.

Tabla 4. Identidades

Secuencia modelo y molde	Identidad (%)
Modelo_1 y 5F1N	95.62
Modelo_2 y 4FQXA	88.7
Modelo_3 y 4FQXB	79.7

Los resultados de las corridas para cada una de las proteínas, así como los mejores modelos para cada una de las proteínas fueron validadas mediante diferentes herramientas computacionales.

Los mejores modelos para cada una las tres proteínas (Modelo_1, Modelo_2 y Modelo_3), se validaron mediante tres herramientas computacionales diferentes, ProsaWeb, Swissmodel y TM-Align, con el fin de saber si los modelos podían ser utilizados en análisis posteriores de acoplamiento molecular con los péptidos predichos. A continuación, se muestra una síntesis de los datos de la validación de las tres proteínas. En esta tabla también se incluyen los datos de los moldes para poder establecer una comparación con los datos obtenidos para los modelos.

Tabla 5. Resumen de validación

Protein	Z Value	Favorable region (%)	TM Value	Align AA
Modelo_1	-9.4	98.16	0.99066	274
Modelo_2	-5.34	96.00	0.88193	177
Modelo_3	-5.27	94.33	0.89979	194
5F1N	-9.19	95.60	-	-
4FQXA	-5.47	99.44	-	-

4FQXB	-5.33	96.17	-	-
-------	-------	-------	---	---

Ejemplo 6. Consolidación de datos obtenidos para las propiedades de los péptidos

Posterior al cálculo de cada una de las propiedades fisicoquímicas y la consolidación de todos los datos de los péptidos predichos mediante las diferentes herramientas, se realizó un consolidado de algunas variables importantes para la elección de los mejores candidatos. En la Figura 3 A y B, se observa el tamaño de los péptidos predichos de la secuencia consenso de las proteínas H y F de CDV, respectivamente.

Ejemplo 7. Modelación de algunos péptidos en su estructura tridimensional mediante PEPFOLD3.

Los péptidos que obtuvieron los mejores puntajes fueron modelados mediante la herramienta PEPFOLD3 con el fin de realizar acoples moleculares de estos péptidos con las moléculas del CMH de canino modeladas en el Ejemplo 5.

Se determinaron péptidos con potencial inmunogénico para células B, o para ser presentados en el contexto clase 1 o clase 2 de las moléculas del CMH mediante la herramienta del IEDB. Luego de someter las secuencias, se obtuvieron más de mil péptidos. A esta herramienta, se le suministra la secuencia consenso de todos los linajes tanto de la proteína H como de F.

Ejemplo 8. Selección de péptidos con base en score y propiedades

Luego de tener todos los datos de predicción mediante las diferentes herramientas, se seleccionaron péptidos tanto de la proteína H como de la proteína F, que tuvieran propiedades fisicoquímicas de estabilidad y vida media apropiadas, así como los mejores valores de predicción de cada una de las herramientas, con el fin de validar *in silico*, un grupo más pequeño de péptidos y poder llevar estos a los siguientes pasos del desarrollo. Los péptidos seleccionados inicialmente se muestran en la Tabla 1.

Ejemplo 9. Acoplamiento molecular de los péptidos seleccionados con la MHC de canino correspondiente a la molécula que lo presenta (clase I, II) mediante CABSDOCK.

5 Cada uno de los 12 péptidos, fue sometido a análisis de acoplamiento molecular mediante CABSDOCK, con la molécula blanco respectiva (sea clase I o clase II, según su predicción), en la plataforma en línea. Se suben ambos archivos por separado y posterior al tiempo de análisis, la herramienta entrega la pose más probable en la que interactúan ambas moléculas, teniendo en cuenta un análisis de clustering o agrupamiento.

10 **Acoplamiento molecular de los péptidos seleccionados con la MHC de canino correspondiente a la molécula que lo presenta (clase I, II) mediante MDOCKPEP.**

Cada uno de los 12 péptidos, fue sometido a análisis de acoplamiento molecular mediante HPEPDOCK, con la molécula blanco respectiva (sea clase I o clase II, según su predicción), en la plataforma en línea. Se suben ambos archivos por separado y posterior
15 al tiempo de análisis, la herramienta entrega la pose más probable en la que interactúan ambas moléculas, teniendo en cuenta un análisis de la pose más probable.

Acoplamiento molecular de los péptidos seleccionados con la MHC de canino correspondiente a la molécula que lo presenta (clase I, II) mediante HPEPDOCK

20

Cada uno de los 12 péptidos, fue sometido a análisis de acoplamiento molecular mediante CABSDOCK, con la molécula blanco respectiva (sea clase I o clase II, según su predicción), en la plataforma en línea. Se suben ambos archivos por separado y posterior al tiempo de análisis, la herramienta entrega la pose más probable en la que interactúan
25 ambas moléculas, teniendo en cuenta un análisis de la pose más probable.

Acoplamiento molecular de los péptidos seleccionados con la TLR-2 y TLR-4 mediante HPEPDOCK.

30 Cada uno de los 12 péptidos, fue sometido a análisis de acoplamiento molecular mediante HPEPDOCK, con la molécula blanco (TLR-2 o TLR-4), en la plataforma en línea. Se suben ambos archivos por separado y posterior al tiempo de análisis, la herramienta

entrega la pose más probable en la que interactúan ambas moléculas, teniendo en cuenta un análisis de la pose más probable.

Ejemplo 10. Pruebas de seguridad *in silico*

5

a- Análisis de Blastp de los péptidos seleccionados para determinar péptidos homólogos propios.

10 Con este análisis se busca determinar si los péptidos diseñados, tenían péptidos homólogos en proteínas propias de los caninos, lo cual sería una situación que se debe evitar, puesto que únicamente se desea generar una respuesta inmunológica frente al CDV, pero no frente a las proteínas propias del canino. Para ningún péptido, se tuvo un valor del 100 % tanto en la identidad como en la cobertura, lo que indica que no hay péptidos idénticos en las proteínas de los caninos.

15

b- Análisis de VaxiJenv2.0 de los péptidos seleccionados para determinar si los péptidos son potenciales antígenos

20 Se determinó si los péptidos eran potenciales antígenos virales, mediante un análisis de scoring de máquinas de aprendizaje, comparado con otros péptidos de los cuales hay evidencia experimental de su antigenicidad.

c- Análisis con ToxinPred de los péptidos seleccionados para determinar si los péptidos son compuestos tóxicos.

25

Se determinó si los péptidos eran potenciales compuestos tóxicos, mediante un análisis de scoring de máquinas de aprendizaje, comparado con otros péptidos de los cuales hay evidencia experimental de su toxicidad.

30 d- Análisis con AllergenFPv1.0 de los péptidos seleccionados para determinar si los péptidos son potenciales alérgenos

Se determinó si los péptidos eran potenciales compuestos alérgenos, mediante un análisis de scoring de máquinas de aprendizaje, comparado con otros péptidos de los cuales hay evidencia experimental de su potencial alérgenico.

5 Ejemplo 11. Cálculo de las propiedades fisicoquímicas de los péptidos seleccionados inicialmente

Se determinaron los datos de la naturaleza de los péptidos mediante la herramienta Protparam. Los resultados se encuentran en la Tabla 6.

10 Tabla 6. Datos de la naturaleza de los péptidos

Peptide	Sequence	Length (aa)	Position	Charge	Molecular weight	Stability	Half-life (hours)
CI-H1-004	PYLLFVLLI	9	36	0	1090,41	+	>20
PCI-H1-030	QVIDVLTPLFK	11	97	0	1272,55	+	1
PCII-H1-008	VENLVRIRF	9	592	+1	1145,37	+	100
PCI-F3-053	LTGTSKSYVRSL	12	651	+2	1311,50	+	5.5
PCII-F3-012*	LKLLRYYTE	9	316	+1	1198,43	+	5.5
PCI-IEDB-H1	TVIPHVLTF	9	473	0	1026,24	+	7.2
LINB-IEDB-H	VSEKQEEQKNC	11	367	-1	1321,42	+/-	100
PCI-F3-054	PAFKPDLTGTSK	12	645	+1	1261,44	+	>20
PCII-IEDB-H2	PPYLLFVLLILLV	13	32	0	1512,98	+	>20
PCI-IEDB-F1	MVYESKVAL	9	552	0	1039,26	+	30
PCII-IEDB-F2	GNKFILSKGNIVA	13	488	+2	1360,62	+/-	30
LINB-IEDB-F	KAQIHWNNL	9	134	+1	1123,28	+	1.3