



Universidad Católica de Santa María

Facultad de Odontología Escuela Profesional de Odontología

**Búsqueda computacional de metabolitos naturales de latino america con
potencial para inhibir la proteína come del sistema comcde en streptococcus
sobrinus.**

Tesis presentada por:

Nieves Loyola, Valeria Norma

ORCID: 0009-0008-2856-8120

para optar el Título Profesional de Cirujano Dentista

Asesor:

Dr. Figueroa Banda, Rufo Alberto

ORCID: 0000-0001-7249-0270

Arequipa - Perú

2026

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

ODONTOLOGIA

TITULACIÓN CON TESIS

DICTAMEN APROBACIÓN DE BORRADOR

Arequipa, 19 de Marzo del 2026

Dictamen: 014491-C-EPO-2026

Visto el borrador del expediente 014491, presentado por:

2018246702 - NIEVES LOYOLA VALERIA NORMA

Titulado:

**BÚSQUEDA COMPUTACIONAL DE METABOLITOS NATURALES DE LATINO AMERICA CON
POTENCIAL PARA INHIBIR LA PROTEÍNA COME DEL SISTEMA COMCDE EN STREPTOCOCCUS
SOBRINUS.**

Nuestro dictamen es:

APROBADO

Titulo Profesional/Titulo de Segunda Especialidad/Grado Académico a optar:

CIRUJANO DENTISTA

**29238358 - SALAS ROJAS MONICA HILDA CLEOFE
DICTAMINADOR**



**29649041 - ZEVALLOS CHAVEZ MARCO ANTONIO
DICTAMINADOR**

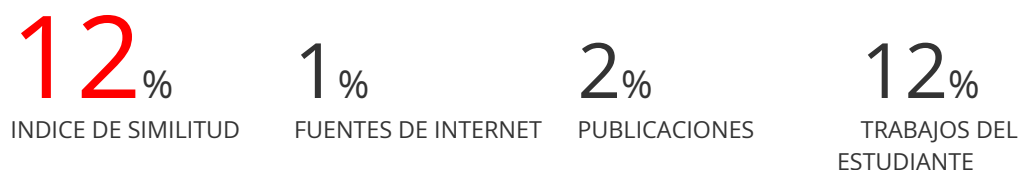


**40563951 - BARRIGA FLORES MARIA DEL SOCORRO
DICTAMINADOR**



BÚSQUEDA COMPUTACIONAL DE METABOLITOS NATURALES DE LATINO AMERICA CON POTENCIAL PARA INHIBIR LA PROTEÍNA COME DEL SISTEMA COMCDE EN STREPTOCOCCUS SOBRINUS.

INFORME DE ORIGINALIDAD



FUENTES PRIMARIAS



Excluir citas

Apagado

Excluir coincidencias < 1%

Excluir bibliografía

Apagado

DEDICATORIA

A Dios por ser mi guía constante en mi vida y en todo este proceso para poder culminar de su mano esta etapa, al brindarme salud, fuerza y resiliencia en momentos difíciles.

A mi madre Sara Loyola y a mi padre Sergio Nieves y a mi familia por haberme acompañado desde siempre pero mas aun cuando hubieron largas jornadas de estudio, momentos de estrés al igual que alegrías y nunca faltó una palabra de aliento y su apoyo incondicional. Formándome como la persona y profesional que soy hoy en día sin su compañía nada de esto estaría sucediendo.

A mi prima Gabriela, por siempre escucharme hablar de mis casos clínicos y el proceso de mi tesis y apoyarme en cada paso que di a pesar de estudiar cosas diferentes, por ser un pilar vital en todo este viaje que es estudiar odontología.

A los que hoy no pueden acompañarme físicamente, pero que esperaron tanto como yo este momento, sé que me acompañaron en cada paso y cada sesión de estudio, este logro también es suyo por ser un ejemplo de superación, por que como siempre me decían “ el que estudia triunfa” cada esfuerzo realizado ha dado resultados y su memoria siempre me animo a seguir en días difíciles

A mis amigos y colegas que supieron apoyarme y guiarme dentro de sus experiencias para poder seguir adelante, cada uno de sus consejos ha sido clave para poder salir adelante.

A mi asesor de tesis, el doctor Figueroa que supo guiarme en todo este proyecto brindándome todo su apoyo y las correcciones necesarias para que el proyecto salga de manera deseada

A la facultad de odontología por haberme dado las mejores experiencias y haberme acogido en sus clases y haber cultivado en mí el amor a la carrera y las ganas de siempre destacar para poder hacer que nuestra rama se impulse en áreas que podemos abordar

AGRADECIMIENTOS

Quiero expresar mi más profundo agradecimiento a todas las personas que hicieron posible la realización de este trabajo.

En primer lugar, agradezco a mi familia, por su amor incondicional, apoyo constante y por ser mi mayor motivación a lo largo de este proceso.

A mis docentes y asesores, por su guía, paciencia y valiosos conocimientos brindados, los cuales fueron fundamentales para el desarrollo de esta tesis.

A mis amigos, por su compañía, ánimo y comprensión en los momentos más difíciles, así como por celebrar cada logro conmigo.

Finalmente, agradezco a todas aquellas personas que, de una u otra forma, contribuyeron a la culminación de esta etapa tan importante en mi vida profesional.

Con gratitud,

Valeria

RESUMEN

La caries dental es una enfermedad crónica que afecta a más del 90% de la población mundial, impulsada en gran medida por la bacteria *Streptococcus sobrinus*. Este microorganismo tiene la capacidad de formar biopelículas resistentes y producir ácidos que destruyen el esmalte dental. Su agresividad está regulada por el sistema de señalización ComCDE, un mecanismo que controla cómo la bacteria se adhiere a los dientes y se adapta al entorno. El objetivo de este estudio fue encontrar una forma de interrumpir este sistema para reducir la virulencia de la bacteria.

Para lograrlo, se realizó una investigación computacional utilizando una base de datos de más de 13,000 metabolitos naturales de origen latinoamericano. El proceso de selección fue riguroso: primero se filtraron moléculas que tuvieran características físicas ideales para actuar localmente en la boca sin ser absorbidas por el cuerpo, y luego se descartaron aquellas que presentaran riesgos de toxicidad, como efectos cancerígenos o mutagénicos. Al final del proceso, solo tres compuestos resultaron ser candidatos seguros y eficaces.

A través de modelos tridimensionales avanzados y simulaciones de "docking" molecular, los investigadores evaluaron qué tan bien se unían estos metabolitos a la proteína ComD de la bacteria. Uno de los compuestos destacó por tener una alta afinidad de unión, lo que sugiere que puede bloquear eficazmente la comunicación bacteriana. Este hallazgo es fundamental, ya que sienta las bases para crear nuevos tratamientos preventivos, como enjuagues o geles, basados en productos naturales para combatir la caries de forma más segura.

Palabras clave: Caries dental, *Streptococcus sobrinus*, Drug Discovery.

ABSTRACT

Dental caries is a chronic disease affecting more than 90% of the world's population, largely driven by the bacterium *Streptococcus sobrinus*. This microorganism has the ability to form resistant biofilms and produce acids that destroy dental enamel. Its aggressiveness is regulated by the ComCDE signaling system, a mechanism that controls how the bacteria adhere to teeth and adapt to the environment. The objective of this study was to find a way to disrupt this system to reduce the bacteria's virulence.

To achieve this, a computational investigation was conducted using a database of over 13,000 natural metabolites of Latin American origin. The selection process was rigorous: first, molecules with ideal physical characteristics for local action in the mouth without systemic absorption were filtered; then, those presenting toxicity risks, such as carcinogenic or mutagenic effects, were discarded. At the end of the process, only three compounds emerged as safe and effective candidates.

Through advanced three-dimensional models and molecular docking simulations, researchers evaluated how well these metabolites bound to the bacteria's ComD protein. One of the compounds stood out for having a high binding affinity, suggesting it can effectively block bacterial communication. This finding is fundamental, as it lays the groundwork for developing new preventive treatments, such as mouthwashes or gels, based on natural products to combat cavities more safely.

Keywords: Dental caries, *Streptococcus sobrinus*, Drug Discovery.

ÍNDICE

DEDICATORIA	
AGRADECIMIENTOS	
RESUMEN	
ABSTRACT	
INTRODUCCIÓN.....	1
CAPÍTULO I.....	3
PLANTEAMIENTO TEÓRICO	3
1.1. DETERMINACIÓN DEL PROBLEMA	4
1.2. PREGUNTA DE INVESTIGACIÓN.....	5
1.3. JUSTIFICACIÓN	5
1.4. OBJETIVOS	6
1.5. MARCO CONCEPTUAL Y ANTECEDENTES INVESTIGATIVOS.	7
1.5.1. Marco conceptual.....	7
1.5.2. Análisis de antecedentes investigativos.....	18
1.6. HIPÓTESIS	26
CAPÍTULO II.....	27
PLANTEAMIENTO OPERACIONAL	27
2.1. DISEÑO METODOLÓGICO.....	28
2.2. POBLACIÓN Y MUESTRA.....	28
2.3. TABLA DE VARIABLES.....	29
2.4. TÉCNICAS Y PROCEDIMIENTOS.....	29
2.5. PLAN DE ANALISIS.....	34
2.6. CONSIDERACIONES ÉTICAS	34
CAPÍTULO III	35
RESULTADOS.....	35

3.1 RESULTADOS.....36

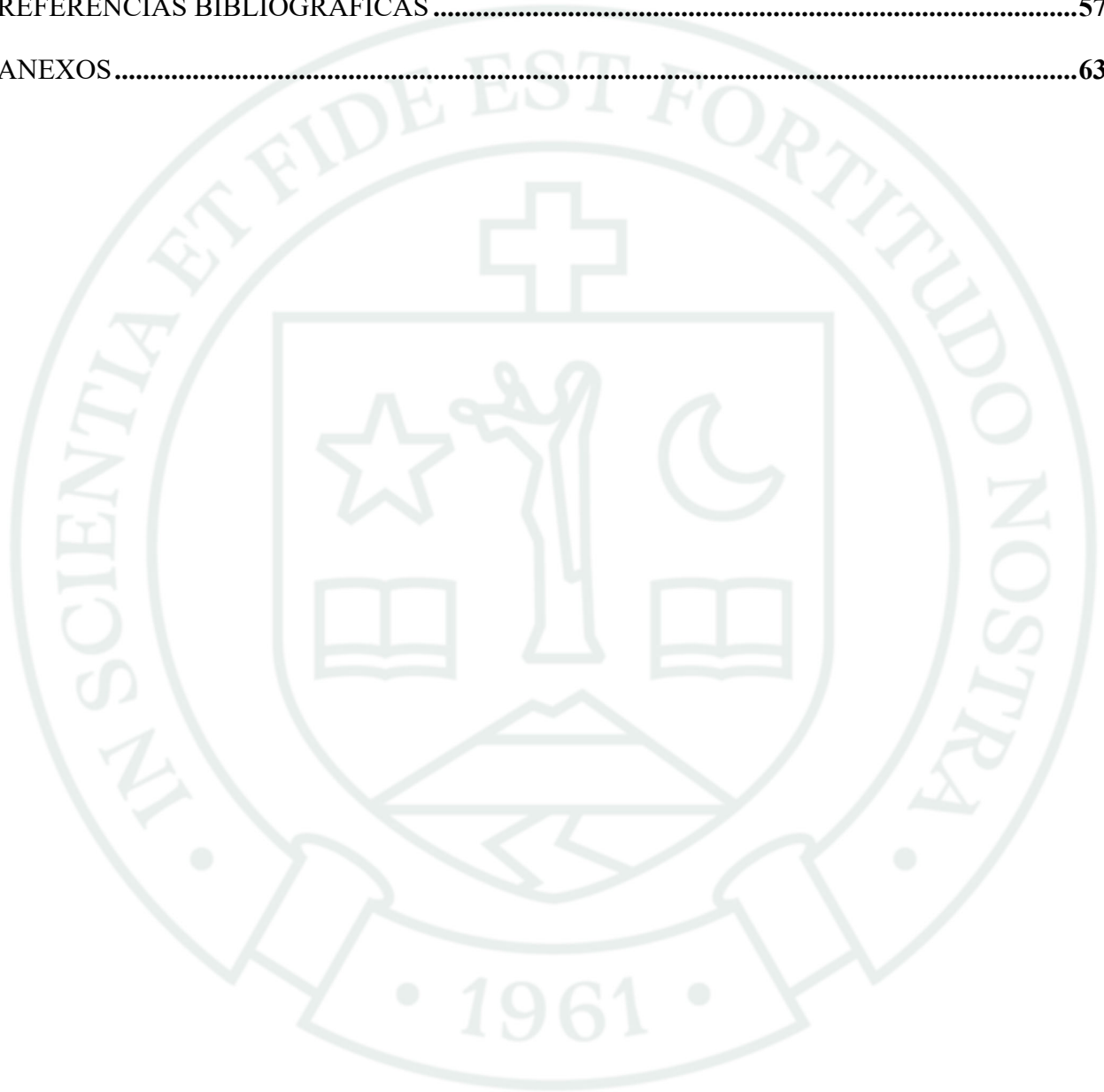
3.1 DISCUSIÓN.....53

3.3. CONCLUSIONES.....55

3.4. RECOMENDACIONES56

REFERENCIAS BIBLIOGRÁFICAS57

ANEXOS.....63



ÍNDICE DE TABLAS

Tabla 1. Tabla de variables	29
Tabla 2. Tratamiento estadístico	34
Tabla 3. Predicción de propiedades toxicológicas mediante DataWarrior.	41
Tabla 4. Resultados del análisis estructural de la proteína ComD.	45
Tabla 5. Parámetros estructurales y probabilidad de drugabilidad de los bolsillos.....	48
Tabla 6. Resultados del acoplamiento molecular	49



ÍNDICE DE FIGURAS

Figura 1. ID de la secuencia de la proteína ComD.....	31
Figura 2. Configuración de la caja de búsqueda en la región activa de ComD.....	33
Figura 3.El mapa de calor.....	36
Figura 4. Relación entre NumHBD y TPSA	38
Figura 5.Distribución de NumHBD, NumHBA y NumRotatableBonds.....	39
Figura 6. Plot de estabilidad de pH	40
Figura 7.Alineamiento entre la secuencia de la proteína ComD.....	42
Figura 8.Estructura tridimensional predicha de la proteína ComD.....	43
Figura 9. Diagrama de Ramachandran	44
Figura 10.Comparación del Z-score del modelo de ComD.....	46
Figura 11.Perfil energético por residuo	47
Figura 12.Localización estructural de los bolsillos.....	48
Figura 13.Mapa de calor de la afinidad de unión.....	50
Figura 14.Representación 3D del acoplamiento de LANaPDB10455	51
Figura 15.Diagrama 2D de interacciones proteína-ligando generado con LigPlot.....	52

INTRODUCCIÓN

La caries dental es una enfermedad infecciosa crónica que afecta aproximadamente al 60-90% de los niños en edad escolar y a casi el 100% de los adultos a nivel mundial, según la Organización Mundial de la Salud (OMS). Esta condición se desarrolla cuando ciertas bacterias presentes en la cavidad oral, especialmente *Streptococcus sobrinus* y *Streptococcus mutans*, se adhieren a las superficies dentales formando biopelículas. Estas biopelículas son comunidades bacterianas encapsuladas en una matriz de sustancias producidas por las propias bacterias, que protegen a los microorganismos de factores externos y facilitan la producción de ácidos que degradan el esmalte dental, lo que finalmente conduce a la formación de caries (1).

Un componente clave en la regulación de la virulencia de *Streptococcus sobrinus* es el sistema ComCDE, un mecanismo de señalización bacteriana que funciona como un sistema de dos componentes. Dentro de este sistema, la proteína ComD actúa como una quinasa sensora ubicada en la membrana celular que detecta señales ambientales y activa a la proteína ComE. Esta activación regula genes responsables de procesos fundamentales como la adhesión bacteriana, la síntesis de exopolisacáridos que conforman la matriz extracelular de la biopelícula, y la adaptación a condiciones adversas, tales como el estrés ácido y oxidativo. La importancia del sistema ComCDE radica en que, al controlar la formación y estabilidad de las biopelículas, influye directamente en la capacidad de *S. sobrinus* para colonizar la cavidad oral y provocar caries dental. Por ejemplo, se ha demostrado que la inhibición de ComD puede reducir significativamente la síntesis de glucanos insolubles, componentes clave para la estructura de la biopelícula, disminuyendo la virulencia bacteriana. Por ello, este sistema representa un objetivo terapéutico estratégico para el control de la caries, que afecta a más del 90% de la población mundial en alguna etapa de sus vidas (2).

Actualmente, los métodos preventivos y tratamientos disponibles, como la higiene bucal, la aplicación de fluoruro y el uso de agentes antimicrobianos, no son específicos para bloquear mecanismos moleculares implicados en la formación de biopelículas. En consecuencia, existe un creciente interés en desarrollar nuevas moléculas capaces de interferir a nivel molecular con procesos bacterianos clave, reduciendo la resistencia microbiana y efectos secundarios. Los compuestos naturales constituyen una fuente valiosa de moléculas bioactivas. América Latina, reconocida por su megadiversidad, alberga una amplia variedad de metabolitos naturales. Por

ejemplo, la base de datos LANaPDB reúne más de 13,500 metabolitos originarios de esta región, lo que ofrece un vasto repertorio para exploración. Sin embargo, debido a la gran cantidad de compuestos, es imprescindible el uso de herramientas computacionales que permitan filtrar, seleccionar y evaluar eficientemente estas moléculas, reduciendo así tiempo y costos asociados a la investigación experimental tradicional. Las técnicas de modelado molecular y acoplamiento molecular (docking) posibilitan la predicción precisa de la estructura tridimensional de proteínas y la simulación de interacciones con ligandos, lo que facilita la identificación de compuestos con alta afinidad de unión y potencial inhibidor. En el contexto de este estudio, se modeló la estructura de la proteína ComD, que cuenta con 448 residuos, y se identificaron sitios funcionales viables para la interacción con metabolitos naturales. Posteriormente, se evaluaron las interacciones de los compuestos seleccionados mediante docking, aplicando criterios rigurosos de filtrado basados en propiedades fisicoquímicas como peso molecular inferior a 500 Dalton, área superficial polar menor a 140 Å² y número de donadores de hidrógeno no mayor a 6 así como perfiles toxicológicos seguros (3).

Este enfoque permitió reducir inicialmente más de 13,500 compuestos a un grupo selecto de siete metabolitos con características ideales para actuar localmente en la cavidad oral, y finalmente identificar tres moléculas con perfil toxicológico adecuado y afinidad significativa hacia ComD, con valores de afinidad de unión cercanos a -8.16 kcal/mol. De este modo, se propone una metodología que integra la riqueza natural latinoamericana con tecnologías computacionales avanzadas para avanzar en el desarrollo de nuevos inhibidores específicos del sistema ComCDE, contribuyendo al control efectivo de la caries dental .



CAPÍTULO I
PLANTEAMIENTO TEÓRICO

1.1. Determinación del problema

La caries dental es una de las enfermedades infecciosas crónicas más comunes en todo el mundo y afecta a personas de todas las edades. Según la Organización Mundial de la Salud (OMS),(1) aproximadamente el 60-90% de los niños en edad escolar y casi el 100% de los adultos han experimentado caries dental en algún momento de sus vidas (2).

La patogénesis de la caries está asociada con la presencia de *Streptococcus mutans*, una bacteria altamente cariogénica que coloniza la superficie dental y contribuye a la formación de biopelículas ácidas que erosionan el esmalte dental. En los últimos años, la prevalencia de caries ha ido en aumento en algunas regiones, impulsada por el alto consumo de azúcares y prácticas de higiene bucal insuficientes (4, 5, 6).

El sistema ComCDE es un mecanismo de transducción de señales presente en *Streptococcus sobrinus*, una de las principales bacterias asociadas al desarrollo de caries dental. Este sistema regula la expresión de genes clave involucrados en la formación de biopelículas, la producción de exopolisacáridos (EPS) y la resistencia a factores ambientales adversos. En particular, la quinasa sensora ComD modula la fosforilación de la proteína ComE del sistema ComCDE, controlando procesos esenciales para la virulencia bacteriana, como la adhesión y la formación de matrices extracelulares altamente resistentes. Estas matrices no solo favorecen la agregación bacteriana, sino que también incrementan la acidogenicidad y la tolerancia a fluctuaciones del pH, lo que promueve la desmineralización del esmalte dental. Dado que la caries es la enfermedad no transmisible más prevalente en el mundo, el sistema ComCDE se perfila como un objetivo terapéutico prometedor para el desarrollo de estrategias encaminadas a reducir la patogenicidad de *S. sobrinus*. A pesar de los avances en tratamientos preventivos, persiste la necesidad de explorar nuevas estrategias que limiten la colonización y la patogenicidad de *S. sobrinus* sin alterar negativamente la microbiota oral en su conjunto. En este contexto, los compuestos naturales constituyen una fuente prometedora de inhibidores potenciales que podrían modular la interacción de *S. sobrinus* con la proteína ComE del sistema ComCDE, disminuyendo así su adhesión y Virulencia (7, 8).

Además, gracias al progreso constante en el ámbito de la computación científica y las técnicas de modelado molecular asistidas por computadora, es posible analizar detalladamente compuestos naturales registrados en bases de datos confiables. Esto permite evaluar sus propiedades

fisicoquímicas, perfiles ADMET y su afinidad de unión con la proteína ComE del sistema ComCDE en *S. sobrinus*. La implementación de estas aproximaciones computacionales facilita la identificación de inhibidores naturales que puedan interferir en la formación de biopelículas, contribuyendo así al desarrollo de nuevas estrategias preventivas para el control de la caries y abordando una prioridad creciente en salud pública.

1.2. Pregunta de investigación.

Pregunta general.

¿Es posible identificar metabolitos naturales en la base de datos LaNaPDB que posean un perfil de seguridad toxicológica *in silico* y una alta afinidad por el sitio activo de la proteína ComD del sistema ComCDE para actuar como agentes anticaries de acción local?

Preguntas específicas.

2. ¿Qué propiedades fisicoquímicas y perfiles de seguridad toxicológica *in silico* definen a los metabolitos naturales latinoamericanos como candidatos idóneos para un tratamiento de acción localizada en la cavidad oral?
3. ¿Cuáles son las características estructurales y los sitios de unión funcionales de la proteína ComD de *Streptococcus sobrinus* que permiten su validación como blanco terapéutico para ligandos inhibidores?
4. ¿Qué grado de afinidad y qué modos de interacción molecular presentan los metabolitos seleccionados frente a la proteína ComD, y cuáles demuestran mayor potencial para interrumpir la virulencia bacteriana?

1.3. Justificación

Económica

Reducir los elevados costos asociados al tratamiento de la caries dental, que representan una carga significativa para los sistemas de salud pública y los pacientes, especialmente en América Latina, donde el acceso a servicios odontológicos preventivos es limitado. Al identificar productos naturales autóctonos con capacidad para inhibir el sistema ComCDE de *Streptococcus sobrinus*, se promueve el desarrollo de agentes preventivos más accesibles, disminuyendo la necesidad de tratamientos invasivos y costosos. Además, el uso de herramientas computacionales para la selección de compuestos optimiza el proceso de identificación, reduciendo tiempos y costos en comparación con los métodos experimentales tradicionales.

Científica

El proyecto se centra en la necesidad de innovar en el área odontológica, históricamente rezagada en cuanto a la aplicación de nuevas técnicas y métodos modernos para el descubrimiento de compuestos terapéuticos. En este contexto, el empleo de enfoques computacionales, como el "virtual screening" y el modelado molecular, ofrece una perspectiva novedosa que puede abrir nuevas vías en la investigación odontológica, proporcionando estrategias avanzadas para la identificación de agentes antimicrobianos naturales. Estas técnicas permiten optimizar la fase inicial del desarrollo de compuestos mediante métodos *in silico*, reduciendo el tiempo y los recursos invertidos antes de las fases experimentales más costosas y complejas. De este modo, el estudio busca sentar bases sólidas que puedan ser aprovechadas en investigaciones posteriores de carácter experimental.

Salud

La elevada prevalencia de la caries dental, que afecta entre el 60% y el 90% de los niños en edad escolar y casi el 100% de los adultos, según la OMS, constituye un problema de salud pública de gran envergadura. Sus consecuencias incluyen dolor crónico, pérdida de piezas dentales, dificultades para masticar, infecciones recurrentes y, en casos extremos, hospitalización. Estos efectos impactan negativamente en la calidad de vida y suponen un gasto considerable para los sistemas de salud, que invierten miles de millones de dólares anualmente en el tratamiento de enfermedades bucales. La alta prevalencia de la caries, asociada a un consumo excesivo de azúcares y prácticas deficientes de higiene bucal, subraya la urgencia de desarrollar nuevas estrategias preventivas que mitiguen la incidencia de esta patología.

1.4. Objetivos

Objetivo general

Identificar metabolitos naturales latinoamericanos con potencial inhibidor de la proteína ComD del sistema ComCDE en *Streptococcus sobrinus* mediante métodos computacionales

Objetivo específico

OE1. Seleccionar metabolitos naturales latinoamericanos con propiedades fisicoquímicas óptimas para la acción localizada en la cavidad oral y evaluar su perfil de seguridad mediante un análisis toxicológico *in silico*.

OE2. Modelar y validar la estructura tridimensional de la proteína ComD de *Streptococcus sobrinus* para identificar los sitios de unión funcionales con potencial para interactuar con ligandos inhibidores.

OE3. Determinar la afinidad y los modos de interacción molecular de los metabolitos seleccionados mediante *docking* molecular para identificar candidatos con capacidad efectiva de inhibir la proteína ComD.

1.5. Marco conceptual y antecedentes investigativos.

1.5.1. Marco conceptual

Sistema ComCDE

El sistema ComCDE es un sistema de dos componentes (Two-Component System, TCS) ampliamente conservado en bacterias grampositivas. Está compuesto por una histidina cinasa transmembrana (ComD) y un regulador de respuesta citoplasmático (ComE). Este sistema permite a las bacterias adaptarse a diversos estímulos ambientales al regular la expresión de genes clave relacionados con la virulencia, la formación de biopelículas y la supervivencia celular (9).

El complejo ComCDE es un sistema de señalización de dos componentes crucial en *Streptococcus sobrinus*, compuesto por la quinasa sensora ComD y el regulador de respuesta ComE, que controla una amplia gama de procesos esenciales para la virulencia y la supervivencia de esta bacteria en el entorno oral. Este sistema influye en la producción de bacteriocinas, otorgando a *S. sobrinus* una ventaja competitiva al inhibir el crecimiento de otros microorganismos en el biofilm dental. Además, regula la producción del péptido estimulante de competencia (CSP), contribuyendo así a la liberación de ADN extracelular que favorece la estructuración y mantenimiento del biofilm (10). Asimismo, el sistema ComCDE participa en la respuesta a condiciones ambientales adversas, como el estrés ácido y oxidativo, incrementando la adaptabilidad de la bacteria en la compleja y hostil cavidad oral. A través de la modulación de genes implicados en la integridad de la pared celular, la formación de biopelículas y la producción de bacteriocinas, ComCDE ejerce una influencia directa sobre la virulencia de *S. sobrinus*, posicionándose como un objetivo terapéutico atractivo para el desarrollo de nuevas estrategias encaminadas a prevenir la progresión de la caries dental sin recurrir a antibióticos tradicionales (11, 12).

Expresión y función del sistema ComCDE

El sistema ComCDE es un regulador central en *Streptococcus* para controlar la competencia bacteriana, un proceso que permite a estas bacterias captar ADN del ambiente y así adquirir nuevas características genéticas, mejorar su adaptabilidad y reparar daños en su propio ADN. Este sistema es un ejemplo clásico de un sistema de dos componentes, que combina un sensor de membrana (ComD) y un regulador de respuesta (ComE), junto con la señal extracelular representada por el péptido ComC, conocido como el péptido estimulante de competencia (CSP).

En *Streptococcus pneumoniae*, la expresión del operón comCDE está cuidadosamente regulada para permitir tanto una expresión basal como una inducida. La expresión basal se mantiene gracias a la capacidad de ComD para autoiniciarse y fosforilar a ComE, incluso en ausencia del CSP. Este nivel basal asegura que las bacterias tengan siempre disponibles los componentes clave del sistema para poder responder rápidamente a condiciones ambientales favorables o a señales de la comunidad bacteriana. Cuando la concentración de CSP alcanza un umbral crítico, se une a ComD, que aumenta su actividad quinasa y fosforila ComE con mayor eficacia, activando un bucle de retroalimentación positiva. Esto incrementa la expresión del operón comCDE y otros genes relacionados con la competencia, sincronizando la respuesta en la población bacteriana. Este mecanismo no es exclusivo de *S. pneumoniae*. En especies como *Streptococcus mutans* y *Streptococcus sanguinis*, se han identificado sistemas similares que regulan la competencia mediante variantes del sistema ComCDE. En *S. mutans*, por ejemplo, la señalización a través de ComCDE está estrechamente relacionada con la formación de biofilms y la producción de bacteriocinas, moléculas que permiten a la bacteria competir con otras especies en el mismo nicho. En *S. sanguinis*, se ha descrito un sistema ComCDE no canónico que integra señales ambientales para ajustar la competencia de manera adaptativa, mostrando la flexibilidad del sistema en distintas especies para responder a diferentes estímulos. Además, la expresión basal de comCDE puede ser modulada por la lectura continua a través de terminadores transcripcionales situados cerca del operón, como el del tRNA^{Arg5} en *S. pneumoniae*, que permite una expresión mantenida y evita que existan células que carezcan de ComD o ComE. Esta estrategia asegura la uniformidad en la capacidad de competencia en la población, evitando la aparición de subgrupos que podrían quedar fuera de la sincronización social que representa la competencia (11).

Rol de ComCDE en *Streptococcus sobrinus*

El sistema ComCDE cumple un papel clave en la regulación de factores de virulencia. Por ejemplo, modula la producción de glucosiltransferasas (Gtfs), responsables de la síntesis de exopolisacáridos que conforman la matriz extracelular de las biopelículas. Esta capacidad de formar biopelículas cariogénicas está directamente relacionada con la patogenicidad de *S. sobrinus* en la cavidad oral. Además, ComCDE regula genes implicados en la tolerancia al estrés ácido, lo que permite a *S. sobrinus* sobrevivir en el ambiente ácido de la placa dental. Asimismo, interviene en la regulación de proteínas adhesinas como GbpB, que contribuyen a la formación y estabilidad de las biopelículas (12, 13).

El sistema ComCDE es, por tanto, un mecanismo de transducción de señales de dos componentes que desempeña un papel esencial en la patogenicidad de *Streptococcus sobrinus*, sobre todo en la regulación de la formación de biopelículas y la expresión de factores de virulencia fundamentales en el contexto de la cariogénesis. Compuesto por la histidina quinasa ComD y el regulador de respuesta ComE, este sistema responde a estímulos ambientales como el pH ácido, el estrés oxidativo y la disponibilidad de nutrientes, activando la transcripción de genes involucrados en la síntesis de exopolisacáridos (EPS). Estos polímeros son componentes clave de la matriz extracelular de las biopelículas, facilitando la adherencia a las superficies dentales y la resistencia a agentes antimicrobianos, características esenciales para la supervivencia y la patogenicidad de *S. sobrinus* en el entorno oral. La función reguladora de ComCDE se extiende a la modulación de enzimas como las glucosiltransferasas (GtfB y GtfC), indispensables para la síntesis de glucanos insolubles que aseguran la estructura y estabilidad de las biopelículas. Además, el sistema influye en la producción de bacteriocinas, péptidos antimicrobianos que brindan a *S. sobrinus* una ventaja competitiva en el ecosistema oral, al inhibir el crecimiento de otras especies bacterianas. La actividad fosfatasa de ComD es crítica para mantener el estado de fosforilación de ComE, asegurando así una regulación precisa de estos factores de virulencia y reforzando la capacidad adaptativa del microorganismo en entornos adversos como la placa dental (14).

Desde una perspectiva terapéutica, el sistema ComCDE representa un objetivo prometedor para el desarrollo de estrategias innovadoras en odontología preventiva. Inhibidores dirigidos contra la histidina quinasa ComD, como algunos compuestos derivados de catecoles, han demostrado potencial para interrumpir la síntesis de EPS, debilitando así la formación de biopelículas. Por otro

lado, tecnologías basadas en ARN antisentido dirigidas a ComE han mostrado eficacia en la reducción de la cariogenicidad mediante la modulación de la matriz extracelular.

Modo de adhesión de *Streptococcus sobrinus*

El mecanismo de adhesión de *Streptococcus sobrinus* a las superficies dentales depende en gran medida de la interacción de adhesinas específicas con receptores salivales, así como de factores ambientales que influyen en la expresión génica bacteriana. La adhesina P1, también conocida como antígeno I/II o proteína SpaP, se une específicamente a la aglutinina salival (SAG) mediante interacciones cooperativas que generan fuerzas de adhesión de hasta 500 pN, incrementando la persistencia bacteriana en el biofilm dental. Paralelamente, las adhesinas AgI/II pueden variar su expresión en función del fenotipo del receptor gp340 presente en la saliva del huésped, lo que afecta la capacidad de adhesión de la bacteria (17).

Estas adhesinas no solo facilitan la unión inicial a la película adquirida de la superficie dental, sino que también estabilizan el ensamblaje estructural del biofilm mediante interacciones con glucanos insolubles sintetizados por las glucosiltransferasas (GtfB y GtfC). Estas enzimas son reguladas, en parte, por sistemas de señalización como el ComCDE, que controla la expresión de genes críticos para la formación y maduración del biofilm (16).

Factores ambientales como el pH, el estrés oxidativo y metabolitos secundarios influyen significativamente en la adhesión de *S. sobrinus*. La producción de ciertos surfactantes bacterianos, análogos a la mutanofactina, incrementa la hidrofobicidad de la superficie celular, favoreciendo la interacción con la matriz salival y promoviendo la cohesión de las células dentro del biofilm (17).

El sistema de transducción de señales ComCDE desempeña un papel central en la regulación de estos procesos, ya que la fosforilación de ComE modula directamente la expresión de genes relacionados con la síntesis de exopolisacáridos (EPS) y adhesinas. Modificar la actividad de ComCDE puede inducir cambios estructurales en la biopelícula, como un aumento de la agregación y la formación de microcolonias más cohesivas, reforzando así la resistencia del biofilm ante agentes antimicrobianos y mecanismos de defensa del huésped. Este conocimiento técnico subraya la complejidad de los mecanismos de adhesión de *S. sobrinus* y su estrecha relación con la formación de caries, señalando posibles dianas terapéuticas para la prevención de la enfermedad (18).

Patogenicidad y relevancia clínica

Streptococcus sobrinus representa un riesgo clínico significativo en el ámbito odontológico debido a su alta capacidad cariogénica y su contribución a la progresión rápida y severa de la caries dental, especialmente en poblaciones pediátricas y pacientes con condiciones predisponentes como discapacidades intelectuales o higiene bucal deficiente. Su metabolismo ácido, generado por la fermentación de carbohidratos, provoca una acidificación localizada que desmineraliza el esmalte dental, facilitando la formación de lesiones cavitadas y aumentando la dificultad en el tratamiento restaurador. La coexistencia de *S. sobrinus* con *S. mutans* potencia la virulencia bacteriana, creando un entorno oral altamente cariogénico que puede conducir a la pérdida prematura de dientes y afectar la función masticatoria y la calidad de vida del paciente. Clínicamente, la presencia predominante de *S. sobrinus* se asocia con caries agresivas y recurrentes, complicando la prevención y control mediante métodos convencionales. Además, la capacidad de esta bacteria para formar biopelículas resistentes incrementa la persistencia del agente infeccioso y la dificultad para su erradicación mediante higiene oral o tratamientos antimicrobianos tópicos. Este peligro clínico se amplía al considerar que las infecciones orales causadas por *S. sobrinus* pueden actuar como reservorios para la diseminación sistémica, incrementando el riesgo de complicaciones como endocarditis bacteriana y eventos cerebrovasculares, especialmente en pacientes con comorbilidades predisponentes. En consecuencia, el manejo odontológico de infecciones asociadas a *S. sobrinus* requiere estrategias integrales que incluyan diagnóstico microbiológico preciso, intervención temprana y control riguroso de factores de riesgo para minimizar tanto el impacto local como las posibles repercusiones sistémicas. La reciente identificación del sistema ComRS en *S. sobrinus* abre nuevas perspectivas para el desarrollo de terapias dirigidas y la mejora en la manipulación genética que permitan contrarrestar su virulencia y reducir su peligro clínico en la práctica dental (19).

Bases de datos de compuestos naturales

Las bases de datos de compuestos químicos y metabolitos son herramientas indispensables en la investigación científica moderna, especialmente en campos como la quimioinformática, bioinformática, farmacología y diseño de fármacos. Estas bases contienen millones de registros que incluyen estructuras químicas en formatos estándares como SMILES y InChI, datos sobre propiedades físico-químicas, actividades biológicas, perfiles toxicológicos y datos espectrales. Por ejemplo, PubChem, mantenida por el National Institutes of Health (NIH), es una de las bases

públicas más extensas, con más de 111 millones de compuestos químicos y más de 300 millones de relaciones químico-biológicas. PubChem soporta búsquedas avanzadas por similitud molecular, subestructura y permite análisis de docking virtual, siendo una fuente clave para el cribado virtual y descubrimiento de fármacos (20).

LIGAND, integrado en la plataforma KEGG, contiene aproximadamente 20,000 compuestos catalogados con información detallada sobre rutas metabólicas, reacciones enzimáticas y mecanismos de acción. Esta base es especialmente útil para investigaciones que requieren vincular datos químicos con vías bioquímicas, facilitando la interpretación funcional de compuestos en contextos metabólicos. Además, LIGAND proporciona capacidad para realizar búsquedas por subestructura química y descargar conjuntos de datos completos para análisis off-line.

El Human Metabolome Database (HMDB) se especializa en metabolitos humanos, reuniendo más de 114,000 metabolitos con información estructural, concentraciones fisiológicas, rutas metabólicas y asociaciones clínicas, siendo vital para estudios metabolómicos y biomarcadores clínicos. Por otro lado, bases especializadas como DrugBank ofrecen datos integrales sobre más de 14,000 fármacos aprobados y en investigación, con sus correspondientes dianas farmacológicas, mecanismos de acción y efectos adversos, facilitando el diseño racional de fármacos y estudios farmacocinéticos (21).

ChEMBL es una base fundamental, que almacena más de 2 millones de compuestos bioactivos con datos experimentales sobre afinidades y actividades biológicas, ampliamente utilizada para modelado molecular y cribado virtual. Complementariamente, la Comparative Toxicogenomics Database (CTD) integra datos de toxicidad química, genes y enfermedades, con más de 1.2 millones de interacciones documentadas, apoyando la evaluación de riesgos toxicológicos y estudios epidemiológicos. Estas bases de datos juegan un papel crítico en aplicaciones avanzadas como el descubrimiento de agentes anticancerígenos, permitiendo identificar compuestos naturales y sintéticos con actividad contra rutas metabólicas tumorales, y en la nutrición, donde facilitan la caracterización de compuestos bioactivos como polifenoles y péptidos con efectos antioxidantes y antiinflamatorios. Sin embargo, la heterogeneidad en formatos y nomenclaturas sigue siendo un reto para la integración de datos. Por ello, iniciativas como la estandarización de ontologías químicas y el desarrollo de APIs interoperables son prioritarias para maximizar el potencial de estos recursos en la investigación traslacional y farmacéutica (21).

LANaNP (Latin American Natural Products Database)

La Latin American Natural Products Database (LANaNP) es una base de datos especializada que compila una amplia colección de metabolitos naturales originarios de América Latina, una región reconocida por su megadiversidad biológica y química. LANaNP se ha consolidado como un recurso fundamental para la investigación en química medicinal, biotecnología y descubrimiento de fármacos, especialmente orientado a la exploración de compuestos con potencial bioactivo en el contexto de la biodiversidad regional.

Desde el punto de vista técnico, LANaNP contiene más de 13,500 estructuras químicas únicas derivadas de fuentes naturales como plantas medicinales, microorganismos, hongos y organismos marinos característicos de los ecosistemas latinoamericanos. Cada entrada incluye información detallada sobre la estructura molecular, presentada en formatos estándar como SMILES (Simplified Molecular Input Line Entry System), InChI (International Chemical Identifier) y archivos tridimensionales (SDF, MOL), facilitando su integración en herramientas de modelado molecular y cribado virtual. LANaNP también integra datos preliminares de actividad biológica reportada en literatura, agrupando compuestos con efectos antimicrobianos, antioxidantes, anticancerígenos, antiinflamatorios y otros perfiles farmacológicos relevantes. Esta anotación es fundamental para priorizar compuestos durante las etapas de cribado virtual. Una ventaja tecnológica importante de LANaNP es su compatibilidad con plataformas computacionales como RDKit, DataWarrior, AutoDock y SwissADME, que permiten el análisis de propiedades ADMET (absorción, distribución, metabolismo, excreción y toxicidad), evaluación de perfiles toxicológicos in silico, y simulaciones de docking (20).

~~PeruNPDB (Peruvian Natural Products Database)~~

La PeruNPDB es una base de datos especializada en productos naturales originarios del Perú, un país reconocido por su extraordinaria biodiversidad, que incluye ecosistemas como la Amazonía, los Andes y la costa del Pacífico. Este repositorio ha sido desarrollado con el objetivo de consolidar y organizar información química, estructural y biológica de metabolitos naturales peruanos, facilitando su uso en estudios de quimioinformática, cribado virtual y diseño racional de fármacos. Desde el punto de vista técnico, PeruNPDB contiene aproximadamente 280 compuestos, representados mediante formatos estándares ampliamente aceptados en química computacional, tales como SMILES, InChI, archivos SDF y MOL2. Esta representación estructural permite su integración directa en plataformas de modelado molecular y análisis estructural, como RDKit,

AutoDock, SwissADME, entre otros. Además, PeruNPDB incluye información taxonómica sobre el origen biológico de los compuestos (plantas, hongos, microorganismos), así como datos experimentales sobre actividades biológicas reportadas, tales como actividad antimicrobiana, antioxidante, antiinflamatoria y otras propiedades farmacológicas relevantes.

Una característica destacada de PeruNPDB es la incorporación de análisis quimioinformáticos avanzados, como análisis de diversidad química mediante técnicas multivariantes (PCA, t-SNE), que permiten visualizar la distribución de compuestos en el espacio químico global y comparar la diversidad de la base con otras bases de datos de referencia, como BIOFACQUIM y NuBBEDB. Esto es fundamental para identificar compuestos novedosos y estructuras químicas únicas.

PeruNPDB ha sido diseñada para su uso en cribado virtual *in silico*, donde los compuestos son sometidos a filtros basados en parámetros ADMET (absorción, distribución, metabolismo, excreción y toxicidad) para seleccionar candidatos con perfiles farmacocinéticos seguros y eficaces. Estos análisis incluyen la predicción de permeabilidad intestinal, solubilidad acuosa, toxicidad hepática, y posible irritación, entre otros (21).

BUBEDB (Biodiversity-Based Bioactive Database)

BUBEDB es una base de datos especializada que recopila compuestos naturales bioactivos derivados de una amplia diversidad biológica, con un enfoque particular en la biodiversidad global pero también incluye un énfasis en especies latinoamericanas. Su propósito principal es servir como un puente entre la riqueza natural y la biotecnología, facilitando el acceso a datos estructurales, funcionales y farmacológicos de metabolitos naturales con potencial terapéutico. Desde un punto de vista técnico, BUBEDB contiene miles de compuestos organizados con anotaciones detalladas que incluyen no solo la estructura química en formatos compatibles con la mayoría de softwares de modelado molecular, sino también información sobre el origen taxonómico, actividad biológica comprobada y propiedades ADMET predichas. Esta integración multidimensional permite un análisis exhaustivo de cada compuesto en el contexto de su aplicabilidad farmacéutica y biotecnológica. Una de las fortalezas de BUBEDB radica en su capacidad para proporcionar datos curados sobre la farmacocinética y toxicología de los compuestos, lo que facilita la identificación temprana de moléculas con perfil seguro para su posterior desarrollo. Esto es especialmente útil para estudios que buscan inhibidores específicos, antimicrobianos o agentes moduladores de procesos biológicos complejos. Adicionalmente, BUBEDB está diseñada para integrarse en pipelines computacionales avanzados que incluyen

cribado virtual, modelado molecular y análisis de interacción proteína-ligando, permitiendo la selección racional de candidatos para ensayos experimentales. Su enfoque en compuestos bioactivos la hace valiosa para proyectos orientados a enfermedades infecciosas, cáncer y trastornos inflamatorios, donde la diversidad química es crucial para superar resistencias y efectos secundarios (22).

Relevancia en la investigación de *Streptococcus sobrinus*

Desde una perspectiva molecular, *S. sobrinus* exhibe un sistema de señalización complejo, como el sistema ComCDE, que regula la competencia bacteriana, la adhesión y la síntesis de exopolisacáridos, factores críticos para la formación y estabilidad de biopelículas cariogénicas. La proteína ComD, una quinasa sensora dentro de este sistema, actúa como una diana prometedor para inhibidores que busquen reducir la virulencia y la formación de biofilms. Clínicamente, la presencia de *S. sobrinus* se ha correlacionado con índices más elevados de caries dental, especialmente en niños y adolescentes, siendo un marcador de caries severa y recurrente. Estudios epidemiológicos muestran que la detección temprana y el control específico de esta bacteria podrían mejorar significativamente los resultados en salud bucal. La capacidad de *S. sobrinus* para coexistir y competir con otras especies orales también influye en la dinámica microbiana de la cavidad bucal, afectando el equilibrio de la microbiota y la susceptibilidad a la enfermedad. En la investigación actual, el enfoque computacional y el cribado virtual de metabolitos naturales específicos de regiones como Latinoamérica ofrecen nuevas oportunidades para descubrir inhibidores dirigidos a proteínas clave de *S. sobrinus*. Estos enfoques aceleran la identificación de compuestos con potencial terapéutico, permitiendo el desarrollo de agentes antimicrobianos más seguros y efectivos, que actúen localmente en la cavidad oral con menor riesgo de resistencia bacteriana y efectos secundarios sistémicos (23).

Química y biología computacional

La química y la biología computacional constituyen campos interdisciplinarios que aplican herramientas matemáticas, estadística y algoritmos computacionales para modelar, simular y analizar sistemas químicos y biológicos a diferentes escalas. Estas disciplinas permiten abordar la complejidad inherente de procesos moleculares y celulares, proporcionando un entendimiento detallado que complementa y potencia la experimentación tradicional. En química computacional, se emplean métodos ab initio, mecánica cuántica y dinámica molecular para describir con precisión la estructura electrónica, la termodinámica y la cinética de sistemas moleculares. Técnicas como

la teoría del funcional de la densidad (DFT) y los métodos post-Hartree-Fock facilitan el estudio de interacciones intermoleculares y reacciones químicas en un rango amplio de sistemas, desde pequeñas moléculas hasta macromoléculas biológicas. La química computacional es fundamental para predecir propiedades fisicoquímicas, optimizar rutas sintéticas y diseñar compuestos con propiedades específicas, especialmente en la industria farmacéutica y de materiales.

Por su parte, la biología computacional integra enfoques computacionales para analizar datos biológicos masivos generados por tecnologías ómicas, como secuenciación genómica, transcriptómica y proteómica. El desarrollo de modelos matemáticos y algoritmos para la simulación de redes metabólicas, señalización celular y dinámica de proteínas permite estudiar mecanismos moleculares en contextos fisiológicos y patológicos (23).

Herramientas como el modelado estructural de proteínas, la predicción de interacciones proteína-ligando y el análisis de rutas metabólicas son fundamentales para la identificación de blancos terapéuticos y el desarrollo de fármacos dirigidos. El avance en la capacidad computacional y el desarrollo de algoritmos eficientes han posibilitado la simulación multiescala, que integra modelos desde la mecánica cuántica hasta la dinámica molecular clásica y modelos de sistemas, permitiendo la exploración de fenómenos biológicos en escalas temporales y espaciales relevantes. Asimismo, la inteligencia artificial y el aprendizaje automático están transformando la biología computacional, facilitando el análisis predictivo y la interpretación de datos complejos mediante técnicas como redes neuronales y aprendizaje profundo (24).

Drug discovery

El descubrimiento de fármacos es un proceso complejo, multidisciplinario y costoso orientado a identificar compuestos bioactivos que puedan modular de forma específica dianas biológicas implicadas en enfermedades. Este proceso integra etapas que van desde la identificación y validación de blancos terapéuticos hasta el desarrollo preclínico y clínico de candidatos farmacológicos. En la fase inicial, la identificación de dianas moleculares es fundamental para el diseño racional de fármacos. Estas dianas pueden ser proteínas, ácidos nucleicos o estructuras celulares específicas, cuyo papel patológico ha sido validado mediante estudios genómicos, proteómicos y farmacológicos (25).

La validación de blancos se realiza mediante técnicas experimentales y computacionales que demuestran la relevancia terapéutica y la posibilidad de modulabilidad. La química medicinal computacional juega un papel crucial en la optimización de compuestos líderes (lead

optimization), empleando técnicas como el modelado molecular, el acoplamiento molecular (docking) y la simulación de dinámica molecular para predecir afinidades de unión, selectividad y propiedades farmacocinéticas. Además, la integración de métodos de cribado virtual permite evaluar grandes bibliotecas químicas de forma eficiente, identificando moléculas con potencial actividad biológica antes de su síntesis y evaluación experimental. El diseño de fármacos asistido por computadora (CADD, por sus siglas en inglés) combina enfoques basados en la estructura (structure-based drug design) y en ligandos (ligand-based drug design), facilitando la generación de nuevos compuestos con propiedades optimizadas y menor toxicidad. La modelación farmacocinética y farmacodinámica in silico contribuye a prever la absorción, distribución, metabolismo, excreción y toxicidad (ADMET) de los candidatos, permitiendo una mejor selección de moléculas para avanzar en el desarrollo (23).

Virtual screening.

El cribado virtual es una metodología computacional esencial en el proceso de descubrimiento de fármacos, cuyo objetivo es identificar compuestos bioactivos potenciales a partir de grandes bibliotecas químicas mediante simulaciones in silico. Esta técnica permite reducir significativamente el tiempo y costo asociados a los ensayos experimentales tradicionales, facilitando la selección de candidatos con alta probabilidad de interacción con una diana biológica específica. Existen dos enfoques principales en el cribado virtual: el basado en estructura (structure-based virtual screening, SBVS) y el basado en ligandos (ligand-based virtual screening, LBVS). El SBVS utiliza la estructura tridimensional de la proteína diana para evaluar la afinidad y el modo de unión de los compuestos mediante técnicas de acoplamiento molecular (docking), dinámica molecular y cálculos de energía libre. Por otro lado, el LBVS se fundamenta en las propiedades químicas y estructurales de ligandos conocidos activos, empleando métodos como la farmacofórica, los modelos cuantitativos de relaciones estructura-actividad (QSAR) y técnicas de similitud molecular para identificar compuestos con perfiles semejantes (3, 12).

La precisión del cribado virtual depende en gran medida de la calidad estructural de la proteína diana, la precisión del algoritmo de acoplamiento y la parametrización de las interacciones moleculares. Para mejorar la predictividad, es común combinar diferentes métodos y validar los resultados con datos experimentales o mediante técnicas de re-docking y validación cruzada. Además, el cribado virtual se ha beneficiado de la integración de inteligencia artificial y

aprendizaje automático, que permiten analizar patrones complejos en grandes bases de datos químicas y biológicas, mejorando la selección y clasificación de compuestos candidatos (3, 16).

1.5.2. Análisis de antecedentes investigativos

Una nueva vía de competencia en el patógeno oral *Streptococcus sobrinus*

Autores : Li J.W., Wyllie R.M., Jensen P.A. (2021)

Streptococcus sobrinus es una de las principales bacterias asociadas a la caries dental en humanos, pero su estudio ha sido limitado debido a la falta de sistemas de competencia natural que faciliten su manipulación genética, a diferencia de otros estreptococos. Sin embargo, se ha descubierto que la mayoría de las cepas de *S. sobrinus* contienen una nueva variante del sistema de competencia ComRS. Este sistema presenta características únicas, como la presencia de dos copias del regulador transcripcional ComR y un péptido feromona (XIP) sin aminoácidos aromáticos.

El uso de XIP sintético permitió transformar genéticamente a *S. sobrinus* con ADN plasmídico o lineal, lo que confirmó que solo una de las copias de ComR es necesaria para la competencia, mientras que la otra parece suprimirla. Además, la exposición al XIP exógeno incrementa la expresión de genes relacionados con bacteriocinas, generando una respuesta antimicrobiana que inhibe el crecimiento de *Streptococcus mutans*, otra bacteria asociada a la caries y competidora ecológica de *S. sobrinus*. Asimismo, se identificaron cepas de *S. sobrinus* que no responden o no producen XIP, consideradas "tramposas" desde una perspectiva ecológica. Un evento de recombinación genética podría restaurar la funcionalidad del sistema ComRS en estas cepas, pero eliminaría el gen que codifica el XIP. Este estudio proporciona nuevas herramientas para investigar a *S. sobrinus* y arroja luz sobre la evolución de sistemas de competencia en bacterias, destacando el papel de cepas "tramposas" en su ecología y evolución (24).

Una vía de competencia ComRS en el patógeno oral *Streptococcus sobrinus* A ComRS competence pathway in the oral pathogen *Streptococcus sobrinus*

Autores : Walden Li, Ryan M. Wyllie, Paul A. Jensen (2020)

Streptococcus sobrinus es una de las dos especies de bacterias responsables de la caries dental en humanos. Sin embargo, el conocimiento sobre esta bacteria es limitado, en parte debido a la creencia de que carece de sistemas de competencia natural, esenciales en otros estreptococos para regular el crecimiento, la virulencia y el quorum sensing. La falta de esta competencia también ha

impedido su manipulación genética, dificultando el estudio de su patogenicidad. Se descubrió que *S. sobrinus* posee un sistema funcional de competencia ComRS con una estructura única. Este sistema incluye dos copias del regulador transcripcional ComR y un péptido feromona (XIP) que carece de aminoácidos aromáticos. El uso de XIP sintético permite transformar genéticamente a *S. sobrinus* con ADN plasmídico o lineal, confirmando que solo una de las copias de ComR es necesaria para inducir la competencia.

Aunque *S. sobrinus* generalmente se clasifica entre los estreptococos del grupo mutans, el sistema ComRS de esta bacteria muestra similitudes estructurales y funcionales con los sistemas de competencia del grupo salivarius. Al igual que en *S. salivarius*, el grupo génico ComRS de *S. sobrinus* incluye un gen para la exportación/escisión de péptidos y está relacionado con un sistema de respuesta bacteriocina. Estos hallazgos plantean interrogantes sobre la ubicación filogenética real de *S. sobrinus*. Por lo tanto, el sistema ComRS en *S. sobrinus* ofrece nuevas herramientas para su estudio y una mejor comprensión de su biología y ecología (25).

Hacia la comprensión de las señales de producción de bacteriocinas por *Streptococcus* spp. y su importancia en las aplicaciones actuales

Toward understanding the signals of bacteriocin production by *Streptococcus* spp. and their importance in current applications

Autores : Laura García-Curiel, Ma. del Rocío López-Cuellar, Adriana Inés Rodríguez-Hernández, Norberto Chavarría-Hernández (2021)

Los microorganismos han desarrollado sistemas de quorum sensing (QS) para detectar pequeñas moléculas señalizadoras que les ayudan a controlar el acceso a nutrientes adicionales y espacio en nichos polimicrobianos altamente competitivos. Muchos procesos bacterianos están regulados por QS; dos ejemplos son la competencia genética natural y la producción de péptidos antimicrobianos como las bacteriocinas. El género *Streptococcus* es ampliamente estudiado por su competencia y su capacidad para producir bacteriocinas, ya que estos péptidos antimicrobianos tienen un potencial significativo en el tratamiento de infecciones causadas por patógenos multirresistentes, un grave problema de salud pública.

La competencia en estreptococos está controlada por la transducción de un sistema de dos componentes: ComD/E, que regula la competencia en los grupos Mitis y Anginosus, y ComR/S, que desempeña la misma función en los grupos Bovis, Mutans, Salivarius y Pyogenic. La comunicación célula a célula necesaria para la producción de bacteriocinas en los grupos de

Streptococcus está controlada principalmente por un parólogo del sistema ComD/E. Las relaciones entre las señales de feromonas y las vías de inducción están relacionadas con los sistemas de producción de bacteriocinas. En esta revisión, se discuten los avances recientes en la comprensión de la señalización y la inducción de la biosíntesis de bacteriocinas mediante la regulación por QS en estreptococos. Esta información podría ayudar en el diseño de mejores métodos para el desarrollo y producción de estos péptidos antimicrobianos. También podría contribuir al análisis y aplicaciones emergentes de las bacteriocinas en términos de su seguridad, calidad y beneficios para la salud humana (10).

Different Patterns of Virulence Genes in Streptococcus mutans and Streptococcus sobrinus Originating from Estonian Toddlers Mothers Cohort

Autores: Greete Maasi, Jelena Štšepetova, Merike Jõesaa, Jana Olak, Reet Mändar (2022)

El estudio comparó los genes de virulencia de Streptococcus mutans y Streptococcus sobrinus, bacterias responsables de la caries dental, en cepas provenientes de niños y madres, correlacionándolos con datos clínicos como índices de caries y niveles de estreptococos mutans en saliva (puntaje SM). Se encontró que S. mutans posee una mayor cantidad de genes de virulencia en comparación con S. sobrinus. En S. sobrinus, los genes gbpA, gbpB, wapA y ftf estuvieron presentes en todas las cepas, mientras que spaP, gtfB, vicR, SMU.1037c y SMU.105 aparecieron en un 41–88% de las muestras, y los genes gtfD y SMU.104 estuvieron ausentes. Además, se observó que en S. sobrinus existe una correlación entre el gen relacionado con la formación de biofilm (vicR) y el gen relacionado con la producción de polisacáridos (gtfD), aunque algunas cepas que contenían el gen gtfB no presentaban vicR ni gtfD, lo que sugiere una posible variabilidad en los mecanismos de virulencia. Particularmente, en cepas de S. sobrinus derivadas de niños, se encontró una asociación entre el gen de adhesión spaP y un mayor puntaje SM, indicando un posible rol de este gen en la colonización y patogenicidad. Estas diferencias genéticas en S. sobrinus, junto con su menor carga de genes de virulencia en comparación con S. mutans, resaltan la necesidad de profundizar en los mecanismos que contribuyen a su papel en la cariogenicidad, especialmente en contextos de cohabitación con otras bacterias (27).

Streptococcus sobrinus as a Predominant Oral Bacteria Related to the Occurrence of Dental Caries in Polish Children at 12 Years Old

Autores: Korona-Glowniak I, Skawinska-Bednarczyk A, Wrobel R, Pietrak J, Tkacz-Ciebiera I, Maslanko-Switlala M, Krawczyk D, Bakiera A, Borek A, Malm A, Mielnik-Blaszczak M. (2022)

La caries dental, reconocida por la Organización Mundial de la Salud (OMS) como una de las principales enfermedades no transmisibles, requiere estrategias efectivas de prevención y tratamiento. Este estudio se centró en evaluar la prevalencia y severidad de la caries dental en niños polacos de 12 años mediante el índice DMFT (dientes permanentes cariados, perdidos y obturados). Se analizaron muestras de saliva de 118 niños utilizando PCR en tiempo real para detectar 8 especies bacterianas cariogénicas y periodontopatógenas. Los resultados revelaron un puntaje promedio bajo de DMFT (1.58 ± 1.98) y una prevalencia de caries relativamente baja (53.4%), mostrando una tendencia a la disminución en comparación con estudios previos. Aunque no se observaron diferencias significativas en la abundancia general de especies bacterianas entre los grupos con y sin caries, en los niños con caries activas se encontraron correlaciones significativas con *Prevotella pallens*, *Fusobacterium nucleatum*, *Streptococcus mutans* y *Lactobacillus fermentum*. Estas especies podrían estar relacionadas con el mantenimiento del proceso cariogénico en un entorno bucal desfavorable. Los hallazgos refuerzan la importancia de entender las interacciones microbianas en la caries para diseñar estrategias de prevención basadas en el equilibrio de la microbiota oral. En particular, la prevalencia de *Streptococcus sobrinus* fue significativamente mayor en niños con caries ($p = 0.023$) y se correlacionó con puntajes más altos de DMFT. Esto sugiere que *S. sobrinus* podría desempeñar un papel temporal importante en la iniciación o potenciación del proceso cariogénico, posiblemente debido a un desequilibrio ecológico en la microbiota dental. Estos resultados subrayan la necesidad de investigaciones adicionales para entender el papel específico de *S. sobrinus* en la caries y desarrollar enfoques preventivos más efectivos (28).

Identificación de inhibidores potenciales contra la deshidrogenasa del shikimato de *Staphylococcus aureus* mediante cribado virtual y prueba de susceptibilidad

Autores : Mengfan Zhu, Jinfeng Qu & Qi Deng

El estudio utiliza cribado virtual para identificar inhibidores de la deshidrogenasa de shikimato (SaSDH) de *Staphylococcus aureus*, una proteína crucial para la biosíntesis de aminoácidos aromáticos y que no está presente en mamíferos, lo que la convierte en un objetivo atractivo para el desarrollo de nuevos antimicrobianos. Para modelar SaSDH, se utilizó la estructura cristalina de la deshidrogenasa de shikimato de *Staphylococcus epidermidis* como plantilla, logrando un modelo tridimensional de alta calidad validado mediante criterios estructurales. Se evaluaron 720,000 compuestos provenientes de tres bibliotecas comerciales usando técnicas de acoplamiento molecular (docking), lo que permitió seleccionar los 200 compuestos con mayores puntuaciones de acoplamiento. Finalmente, se probaron experimentalmente 12 compuestos, destacándose (-)-galocatechin gallate y rhodiosin por su potente actividad inhibitoria contra SaSDH (29).

El cribado virtual y los estudios de acoplamiento molecular revelaron interacciones clave de los inhibidores con los residuos activos de SaSDH, incluyendo la formación de enlaces de hidrógeno con aminoácidos como Asn58, Asn85 y Glu239. (-)-Galocatechin gallate mostró un modo de inhibición no competitivo con un K_i de 2.47 μM , mientras que rhodiosin actuó como inhibidor competitivo con un K_i de 73.38 μM . Ambos compuestos demostraron actividad antibacteriana contra *S. aureus*, con concentraciones inhibitorias mínimas (MIC) de 50 $\mu\text{g/mL}$ y 250 $\mu\text{g/mL}$, respectivamente. Aunque los inhibidores también afectaron la integridad de la membrana bacteriana, su mecanismo principal radica en la inhibición de la SaSDH. Estos resultados destacan la utilidad del cribado virtual como una estrategia eficaz para identificar inhibidores novedosos con potencial terapéutico frente a infecciones resistentes a los medicamentos actuales (29).

PeruNPDB: la Base de Datos de Productos Naturales Peruanos para el Cribado Virtual de Fármacos PeruNPDB: The Peruvian natural products database for in silico drug screening
Autores : Haruna L. Barazorda-Ccahuana, LenaGálvez Ranilla , MayronAntonio Candia-Puma, EymiGladys Cárcamo-Rodriguez, Angela Emperatriz Centeno-Lopez,Gonzalo Davila-Del-Carpio, José L. Medina-Franco⁵ & MiguelAngel Chávez-Fumagalli.

El artículo presenta el desarrollo de la Peruvian Natural Products Database (PeruNPDB) como una herramienta para el cribado virtual en el diseño de fármacos. Esta base de datos contiene 280 compuestos derivados de productos naturales peruanos, principalmente de plantas, y es la primera iniciativa de este tipo en Perú. Los compuestos se recopilaron mediante un análisis sistemático de la literatura científica, seguido de la creación de representaciones químicas mediante herramientas como SMILES y DataWarrior. Posteriormente, se llevó a cabo la caracterización quimioinformática de la base de datos, evaluando propiedades fisicoquímicas y la diversidad química, así como su distribución en el espacio químico global. Se emplearon técnicas como análisis de componentes principales (PCA) y t-SNE para visualizar las similitudes y diferencias de los compuestos frente a bases de datos de referencia como BIOFACQUIM y NuBBEDB. El enfoque de cribado virtual se destaca como una herramienta central del proyecto, ya que la base de datos está diseñada para identificar compuestos líderes frente a enfermedades específicas mediante simulaciones computacionales. Este enfoque permite acelerar el desarrollo de fármacos, reduciendo costos y tiempo en comparación con métodos experimentales tradicionales. Además, la evaluación de "drug-likeness" mostró que los compuestos de PeruNPDB tienen propiedades químicas similares a medicamentos aprobados por la FDA, lo que refuerza su potencial terapéutico. La base de datos está disponible en línea y busca facilitar el diseño de fármacos basados en la biodiversidad peruana, al mismo tiempo que fomenta la conservación y el uso sostenible de los recursos naturales (21).

COCONUT Online: Collection of Open Natural Products Databases

Autores : María Sorokina ,Peter Merseburger ,Kohulan Rajan ,Mehmet Aziz Yirik yCristoph Steinbeck

El artículo "COCONUT online: Collection of Open Natural Products database" presenta la base de datos COCONUT, una colección amplia y abierta de productos naturales (PN) diseñada para facilitar la investigación en descubrimiento de fármacos y aplicaciones relacionadas. La base de datos integra más de 400,000 productos naturales únicos, recopilados de 53 fuentes diferentes, incluyendo bases de datos públicas y literatura científica. Se caracteriza por su accesibilidad y versatilidad, permitiendo búsquedas basadas en subestructuras, similitudes moleculares, fórmulas químicas y descriptores moleculares. Además, su diseño utiliza tecnologías modernas como MongoDB, React.js y APIs REST para soportar búsquedas avanzadas y consultas programáticas. El cribado virtual es un componente esencial de la funcionalidad de COCONUT. La base de datos está diseñada para integrar los PN en pipelines de diseño de fármacos, utilizando descriptores químicos precomputados y características como SMILES y deepSMILES, las cuales facilitan el modelado computacional (30).

Evaluación de la actividad antibacteriana de extractos de plantas peruanas sobre Streptococcus mutans y Streptococcus sobrinus

Autores: Jackeline Luciana Checalla-Collatupa, Marco Antonio Sánchez-Tito (2021).

Este estudio investigó las propiedades antibacterianas del extracto etanólico de propóleo recolectado en regiones de Perú, enfocándose principalmente en Streptococcus mutans, bacteria estrechamente relacionada con Streptococcus sobrinus en el contexto de la caries dental. Mediante ensayos in vitro, se evaluó la capacidad del extracto para inhibir el crecimiento bacteriano y se realizó un análisis fitoquímico para identificar los compuestos bioactivos presentes. Los resultados indicaron una alta concentración de terpenos y diterpenos, metabolitos conocidos por su actividad antimicrobiana y su capacidad para interferir en procesos de señalización bacteriana como el quorum sensing, que regula sistemas como ComCDE (24).

Presencia de Streptococcus mutans genotipo C en niños y adolescentes peruanos y su relación con la prevalencia de caries dental

Autores: Noé Salazar, María del Carmen Gamboa, et al. (2018)

Mediante técnicas moleculares como PCR-Multiplex, este estudio identificó y caracterizó la prevalencia del genotipo C de Streptococcus mutans en una población pediátrica peruana. Dado que diferentes genotipos pueden presentar variaciones en sus mecanismos de virulencia y resistencia, el conocimiento detallado de estos perfiles es fundamental para diseñar estrategias terapéuticas dirigidas. Aunque el estudio no aborda directamente la inhibición de la proteína ComD, la identificación de genotipos asociados a mayor virulencia proporciona un marco para investigar cómo metabolitos naturales podrían modular la función de proteínas reguladoras de quorum sensing en cepas locales de S. sobrinus y S. mutans, apoyando el desarrollo de fármacos específicos para la realidad epidemiológica peruana (25).

Efecto antibacteriano in vitro del extracto hidroetanólico de hojas de Sambucus peruviana sobre Streptococcus mutans

Autor: Jean Pierre Martín León Silva (2018)

Esta tesis de grado evaluó la actividad antibacteriana del extracto hidroetanólico de hojas de Sambucus peruviana, planta nativa de Perú, frente a Streptococcus mutans, uno de los principales agentes causantes de caries dental. Los ensayos in vitro mostraron una inhibición significativa del crecimiento bacteriano, atribuida a la presencia de metabolitos fenólicos, flavonoides y otros compuestos bioactivos. Estos metabolitos son reconocidos por su capacidad para interferir en mecanismos celulares bacterianos, incluyendo rutas de señalización como el sistema ComCDE, responsable de la regulación de la competencia genética y producción de bacteriocinas. Los resultados sugieren que compuestos derivados de Sambucus peruviana podrían ser estudiados más a fondo para su aplicación en inhibición computacional dirigida a la proteína ComD de Streptococcus sobrinus, proponiendo un recurso natural con potencial terapéutico en la prevención de enfermedades bucales (26).

Exploración de metabolitos naturales peruanos para la inhibición de proteínas reguladoras en bacterias orales

Autores: José Luis Huamán, María Elena Quispe, Ricardo Salas (2022)

Este trabajo revisa y compila investigaciones recientes desarrolladas en Perú que emplean la biodiversidad local como fuente de metabolitos naturales con potencial actividad antibacteriana, enfocándose en inhibidores de sistemas reguladores de quorum sensing, específicamente la proteína ComD del sistema ComCDE en bacterias orales como *Streptococcus sobrinus*. Se describen métodos computacionales utilizados para el cribado virtual y modelado molecular, incluyendo docking y simulaciones dinámicas, que permiten predecir la afinidad de unión y la eficacia potencial de los metabolitos naturales. Además, se enfatiza la importancia de validar estos compuestos en contextos biológicos y clínicos, destacando el valor del enfoque interdisciplinario entre bioinformática, química farmacéutica y microbiología para el desarrollo de nuevos agentes antimicrobianos adaptados a las cepas bacterianas prevalentes en Perú (30).

1.6. Hipótesis

Hipótesis Nula : No es posible identificar metabolitos naturales en la base de datos LANA-PDB que posean un perfil de seguridad toxicológica *in silico* y una alta afinidad por el sitio activo de la proteína ComD de *Streptococcus sobrinus* para actuar como agentes inhibidores.

Hipótesis Alternativa : Existen metabolitos naturales en la base de datos LANA-PDB que presentan un perfil de seguridad toxicológica *in silico* adecuado y una alta afinidad de unión molecular por la proteína ComD, permitiendo su identificación como potenciales agentes inhibidores de la virulencia en *Streptococcus sobrinus*



CAPÍTULO II
PLANTEAMIENTO OPERACIONAL

2.1. Diseño metodológico

Esta investigación se define bajo un abordaje cuantitativo de tipo teórico-computacional, caracterizada por ser un estudio prospectivo, observacional y multivariable. Su diseño se centra en el análisis de datos y modelado computacional para realizar una evaluación comparativa, alcanzando un nivel de investigación que integra lo descriptivo con lo comparativo. Todo el proceso se desarrolla en un ámbito de recolección de datos digital, orientado a medir y contrastar las interacciones moleculares de forma sistemática.

2.2. Población y muestra.

La población de este estudio está constituida por el conjunto total de compuestos químicos registrados en la base de datos de productos naturales de América Latina, específicamente la LANaPDB (Latin American Natural Products Database). Esta base de datos alberga un universo de 13,579 metabolitos naturales de origen regional, que incluyen estructuras como flavonoides, alcaloides y terpenoides. Por su parte, la muestra seleccionada de forma no probabilística corresponde a los siete metabolitos que lograron superar los filtros fisicoquímicos iniciales.

Los criterios de inclusión se basaron en parámetros fisicoquímicos diseñados para asegurar una acción localizada en la cavidad oral y minimizar la absorción sistémica. Se incluyeron moléculas con un peso molecular inferior a 500 Da, un área superficial polar topológica (TPSA) menor a 140 Å², y una flexibilidad molecular moderada con un máximo de 12 enlaces rotables. Además, se incluyeron compuestos con una polaridad específica, definida por tener más de 5 donadores de enlaces de hidrógeno para favorecer su permanencia en el entorno bucal y reducir la permeabilidad celular.

Los criterios de exclusión se aplicaron rigurosamente para garantizar la seguridad de los metabolitos seleccionados. Se descartaron todos los compuestos que presentaron cualquier nivel de alerta (Low o High) en las pruebas de toxicidad *in silico*, específicamente en parámetros de mutagenicidad, carcinogenicidad (tumorigénesis), toxicidad reproductiva e irritación cutánea. Asimismo, se excluyeron aquellos metabolitos que, tras las simulaciones de acoplamiento

molecular (*docking*), no demostraron una afinidad de unión significativa o interacciones estables con los residuos clave del sitio activo de la proteína ComD.

2.3. Tabla de variables.

Tabla 1. Tabla de variables

Variable	Definición	Indicadores	Sub indicadores
Características	Características físico-químicas	· Peso molecular.	· Peso molecular (<500 Da).
Seguridad y	Evaluación de las propiedades de ADMET	· Permeabilidad intestinal.	· Permeabilidad (alta/baja).
Capacidad antimicrobiana	Capacidad del compuesto para	Energía de interacción en docking molecular.	· Energía de interacción (<-10 kcal/mol).

2.4. Técnicas y procedimientos.

Obtención y Filtrado de Metabolitos Naturales.

Se utilizó la base de datos LaNaPDB, que contiene un total de 13,579 metabolitos naturales de origen latinoamericano. Las moléculas fueron descargadas en formato SMILES y posteriormente convertidas a estructuras 2D y 3D utilizando la librería RDKit en Python. Esta base de datos incluye compuestos orgánicos como flavonoides, alcaloides y terpenoides, con diversas características estructurales y funcionales (23).

Selección de Moléculas por Absorción

Se utilizó un conjunto de datos de metabolitos naturales extraídos de una base de datos, en el que se identificaron dos columnas principales: el peso molecular (MW) y el área superficial polar topológica (TPSA) de cada compuesto. Los datos fueron filtrados eliminando las moléculas con valores de MW superiores a 500 y TPSA mayores a 140, ya que estas características se asocian con una menor permeabilidad y, por lo tanto, una menor capacidad de absorción. Posteriormente, se asignó una clasificación a las moléculas en función de su capacidad de absorción, clasificando aquellas que cumplieran con los criterios de absorción como "Se absorbe" y las que no como "No se absorbe (Cumple)". Para visualizar la distribución de los compuestos según el MW y TPSA, se

generó un gráfico de dispersión que diferenciaba las moléculas absorbibles y no absorbibles con distintos colores. El proceso fue implementado utilizando bibliotecas de Python como Pandas, Matplotlib y Seaborn para la carga de datos, aplicación de filtros y creación del gráfico.

Selección por polaridad e interacción con el Agua

Esta selección consiste en la aplicación de un criterio basado en la cantidad de donadores de enlaces de hidrógeno (NumHBD) para seleccionar los compuestos más adecuados. Primero, se cargó el conjunto de datos desde un archivo de Excel que contiene las columnas "NumHBD" y "TPSA". El filtro consistió en seleccionar las moléculas cuyo valor de NumHBD fuera mayor a 5, ya que estas tienen una mayor polaridad y, por lo tanto, una menor permeabilidad celular, lo que puede favorecer su acción localizada en la cavidad bucal. Para visualización, se generó un gráfico de dispersión utilizando las bibliotecas Seaborn y Matplotlib, donde se representaron todos los compuestos con valores de $\text{NumHBD} \leq 5$ en un color, y los compuestos que superaron el umbral de $\text{NumHBD} > 5$ fueron destacados con un color diferente y marcados con un símbolo distinto para facilitar su identificación.

Selección por baja permeabilidad y acción local.

El criterio para la selección en este filtro se basó en las propiedades fisicoquímicas de los compuestos para garantizar su estabilidad en la cavidad bucal a un pH de 6, que refleja las condiciones del entorno bucal. Se aplicaron tres condiciones clave: el número de donantes de enlaces de hidrógeno (NumHBD) fue limitado a 6, para evitar interacciones excesivas con el agua y otros componentes del entorno bucal; el número de aceptores de enlaces de hidrógeno (NumHBA) se restringió a 12, con el fin de mantener la polaridad adecuada y evitar la absorción sistémica; y el número de enlaces rotables (NumRotatableBonds) se estableció en un máximo de 12, para asegurar que los compuestos mantuvieran una conformación estable y predecible, favoreciendo su interacción eficaz en el ambiente bucal.

Evaluación de Propiedades Toxicológicas

Tras el filtrado de los Metabolitos Naturales, se seleccionaron compuestos con mayor potencial inhibitorio sobre la proteína ComD del sistema ComCDE en *Streptococcus sobrinus*. Una vez seleccionados se procedió a la evaluación de sus propiedades toxicológicas mediante herramientas computacionales. Esta etapa resulta fundamental para proponer compuestos candidatos con un perfil de seguridad adecuado, especialmente en estudios preliminares *in silico*, como el desarrollado en este trabajo. La evaluación toxicológica se llevó a cabo utilizando el software

DataWarrior, una herramienta de código abierto que permite predecir propiedades toxicológicas relevantes a partir de la estructura química de las moléculas. Las predicciones generadas por DataWarrior están basadas en modelos estadísticos entrenados con grandes bases de datos experimentales, las propiedades evaluadas fueron:

- Mutagenicidad
- Carcinogenicidad (Tumorigénico)
- Toxicidad reproductiva
- Irritación (cutánea)

Las predicciones fueron cualitativas, categorizadas como “None”, “Low” o “High”. Para continuar con el estudio, se aplicó un criterio de exclusión: todos los compuestos con valores distintos de “None” en cualquiera de las propiedades fueron descartados.

Obtención y Preparación de la proteína ComD

La proteína ComD, es una histidina quinasa asociada al sistema de regulación por quorum sensing ComCDE en *Streptococcus sobrinus*. La secuencia aminoacídica correspondiente fue obtenida desde el portal del NCBI bajo el código GenBank: SQG19849.1, en formato FASTA (ver Figura 1).

Figura 1. ID de la secuencia de la proteína ComD

histidine kinase of the competence regulon, ComD [*Streptococcus sobrinus*]

GenBank: SQG19849.1

[GenPept](#) [Identical Proteins](#) [Graphics](#)

```
>SQG19849.1 histidine kinase of the competence regulon, ComD [Streptococcus sobrinus]  
MLSLVYLNRFTQMLLWSFVELFIIRRLKIRIPWAYQLAFLVSSLPYLWLEFFYANQFFEFALICLLI  
FKFSKKVTNTWLKLFYAFYVKLGVDIARFTLFVIVVPLFPHFVAVQKLPVSIILQVLLTLFGNEVIIRY  
FKMDFLRINEVEDKSIQSVILFHNIFFLAYFLVLWLVTPFGLINATEAANFRINTSTLFMLVFLS  
FISMLNFKVATIYEERLVSQKESELKNITDYSHQIENLYAKIRSFHRDYSNILTSLRYSADHDDLASVRQ  
TLDEITRESQILGTKEFEIANLANLKDDALKSIFSSKLSKAIDLGLDIKVEILNPIGLPVFSFSSLDFIR  
LVSNNLDNAIEAANESEDEQLLISYFKHDDTYVFILANSTKEKTDNIAQLFQTGVSSKGGKDRGVGLAIVK  
KILSTYPNVRLNTSNQNGLFTQHLEVQE
```

Dado que su estructura tridimensional no está disponible en el Protein Data Bank (PDB), se utilizó el servidor SWISS-MODEL para llevar a cabo el modelamiento por homología, la cual se ingresó directamente la secuencia sin especificar ninguna plantilla. El servidor selecciona automáticamente una estructura generada por AlphaFold como modelo de referencia: la proteína A0A380JEM7.1.A, correspondiente a *Streptococcus downei*. Una vez obtenido el modelo

tridimensional de la proteína ComD, se procedió a su validación estructural utilizando dos servidores complementarios. MolProbity permitió evaluar la calidad estereoquímica del modelo mediante parámetros como el diagrama de Ramachandran, clash score, desviaciones en enlaces y ángulos. Por otro lado, ProSA-web se empleó para analizar el perfil energético global de la estructura y generar un puntaje Z-score, comparando el modelo con estructuras experimentales de alta resolución depositadas en el PDB. La combinación de ambas herramientas permitió confirmar que el modelo era estructuralmente confiable para los análisis posteriores de predicción de bolsillos y acoplamiento molecular.

Predicción de Bolsillos

Con la estructura tridimensional validada de la proteína ComD, se procedió a identificar potenciales sitios de unión o "bolsillos" que pudieran servir como blanco para los metabolitos seleccionados.

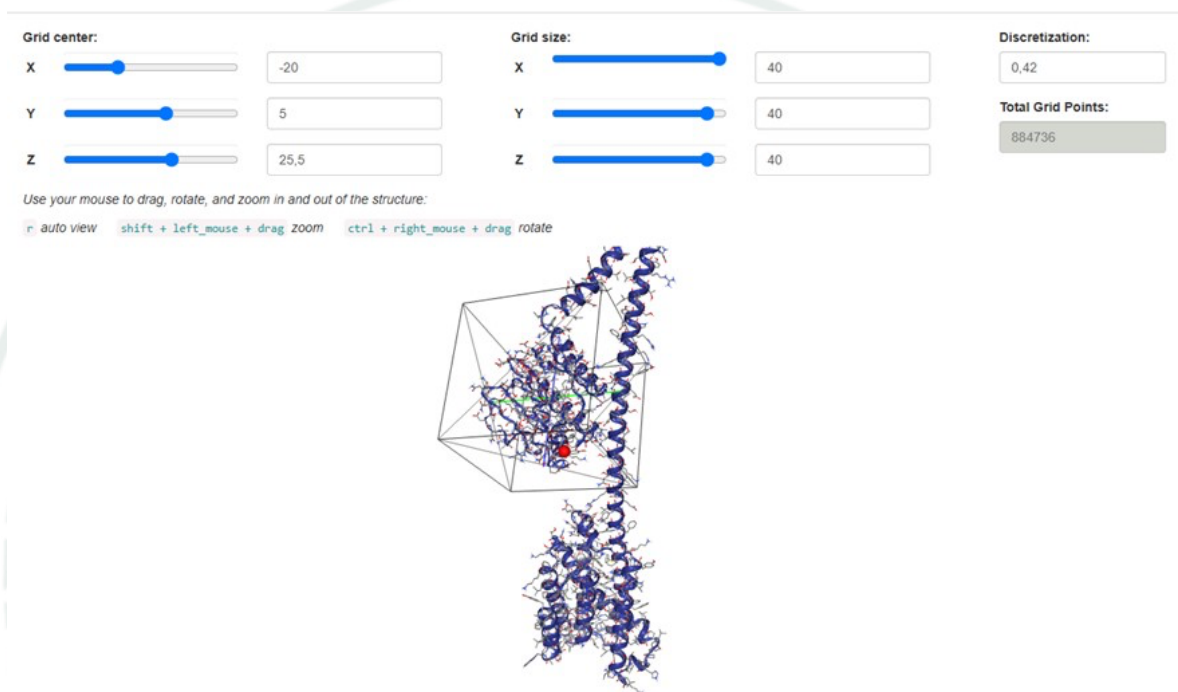
Para esta tarea, se emplearon las herramientas PockDrug, Fpocket y mPocket. Cada una fue aplicada sobre el modelo estructural en formato .pdb con el objetivo de detectar y analizar las cavidades presentes en la superficie de la proteína. Los resultados obtenidos fueron comparados para reforzar la fiabilidad en la selección del sitio de unión. Se prioriza el bolsillo que presenta mayores valores de volumen, accesibilidad superficial y puntajes de drugabilidad, el cual fue posteriormente utilizado en los estudios de acoplamiento molecular.

Acoplamiento Molecular

Antes del proceso de acoplamiento molecular, los metabolitos seleccionados fueron preparados adecuadamente para garantizar una conformación estable y energéticamente favorable. Las estructuras de los compuestos fueron inicialmente optimizadas mediante minimización energética utilizando el software Chem3D, empleando el campo de fuerza MM2, y posteriormente se guardaron en formato tridimensional (.pdb). Este procedimiento permitió reducir tensiones conformacionales y asegurar una geometría adecuada de los ligandos antes del docking. Una vez preparados, los metabolitos fueron sometidos a estudios de acoplamiento molecular con la proteína ComD utilizando la plataforma en línea DockThor. Este servidor permite evaluar las interacciones entre ligandos flexibles y receptores rígidos, generando múltiples poses de acoplamiento y calculando sus respectivas energías de unión (afinidad) expresadas en kcal/mol. El modelo tridimensional validado de la proteína ComD fue cargado en formato .pdb en la plataforma DockThor, donde se definió una caja de búsqueda personalizada (User Defined) para enfocar el

análisis en el bolsillo activo previamente identificado. La caja de búsqueda (grid) se definió manualmente centrada en las coordenadas X = -20, Y = 5, y Z = 25.5, con un tamaño de 40 × 40 × 40 Å en cada eje. La discretización del espacio fue establecida en 0.42 Å, lo que resultó en un total de 884,736 puntos de grid generados por DockThor (ver Figura 2):

Figura 2. Configuración de la caja de búsqueda en la región activa de ComD.



El acoplamiento molecular se realizó utilizando el modo Estándar del servidor DockThor, con una configuración específica del algoritmo genético. Se estableció un número de evaluaciones de 1,000,000, un tamaño de población de 750 individuos, y se realizaron 24 ejecuciones independientes para cada ligando, garantizando una exploración adecuada del espacio conformacional. Para asegurar la reproducibilidad y robustez del experimento, se realizaron tres repeticiones independientes del procedimiento completo, utilizando diferentes semillas aleatorias: -1985, -1990 y -1995. Además, se activó la opción de soft docking.

Cada metabolito fue acoplado individualmente, generando múltiples poses ordenadas por afinidad. Las poses con menor energía libre de unión (kcal/mol) fueron seleccionadas para el análisis posterior de interacciones con residuos del bolsillo activo de ComD.

2.5. Plan de analisis.

Organización de la información

La información fue organizada en forma digital, almacenándose en bases de datos locales y servidores en la nube . Los datos estructurales de metabolitos y proteínas se catalogaron y clasificaron mediante herramientas de gestión de datos químico-biológicos, asegurando la trazabilidad y control de versiones durante las diferentes etapas del estudio.

Validación de instrumento.

Tipo de procesamiento

El procesamiento fue de tipo computacional, cuantitativo , integrando técnicas de análisis estadístico descriptivo, pruebas de hipótesis y modelado computacional para evaluar la interacción proteína-ligando y las propiedades fisicoquímicas y toxicológicas de los metabolitos.

Operaciones del procesamiento

Filtrado y selección de moléculas mediante parámetros fisicoquímicos y toxicológicos. Optimización energética de estructuras químicas para asegurar geometrías estables.

Tabla 2. Tratamiento estadístico

Variable	Tipo	Prueba estadística	Estadística	Pruebas estadísticas
Propiedades Físico-Químicas	Cuantitativo	Screening	Frecuencia absoluta y porcentual	Prueba de hipótesis
Propiedades	Cuantitativo	Screening	Valores de absorción, toxicidad	Análisis de regresión
Actividad	Cuantitativo	Docking	Energía afinidad	GLM/Regresión múltiple
Eficiencia de	Cuantitativo	Modelado computacional	Frecuencia de índices calculados	ANOVA

2.6. Consideraciones éticas

No involucra sujetos humanos ni animales, por lo cual no requiere aprobación previa de un comité de ética.



CAPÍTULO III

RESULTADOS

3.1 Resultados.

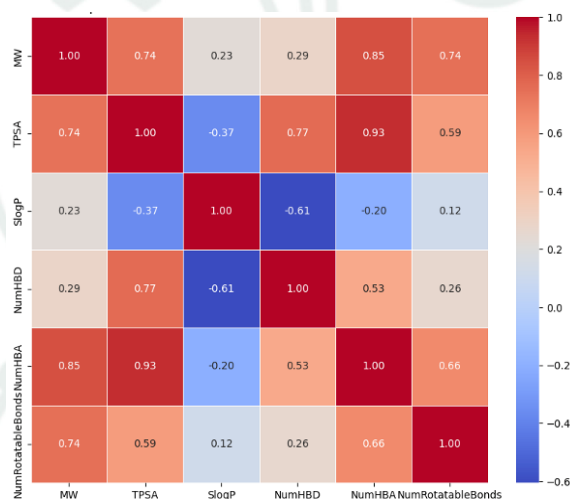
Filtrado de propiedades Físicoquímicas.

En el estudio, se utilizó la base de datos LaNaPDB, que contiene un total de 13,579 metabolitos naturales de origen latinoamericano. Estos metabolitos fueron representados mediante estructuras generadas a partir de SMILES utilizando RDKit, lo que permitió obtener sus estructuras en 2D y 3D.

Correlación de Características Físicoquímicas

Se realizó un análisis de las correlaciones (Pearson) entre diversas propiedades físicoquímicas de los compuestos para visualizar y analizar las correlaciones entre diversas propiedades físicoquímicas, con el objetivo de identificar qué propiedades tienden a estar asociadas entre sí. Esto facilita la selección y el emparejamiento de características clave para filtrar los metabolitos de interés. Se encontró una correlación fuerte y positiva entre TPSA - MW, LogP - TPSA, NumHBD - NumHBA - NumRotatableBonds.

Figura 3. El mapa de calor

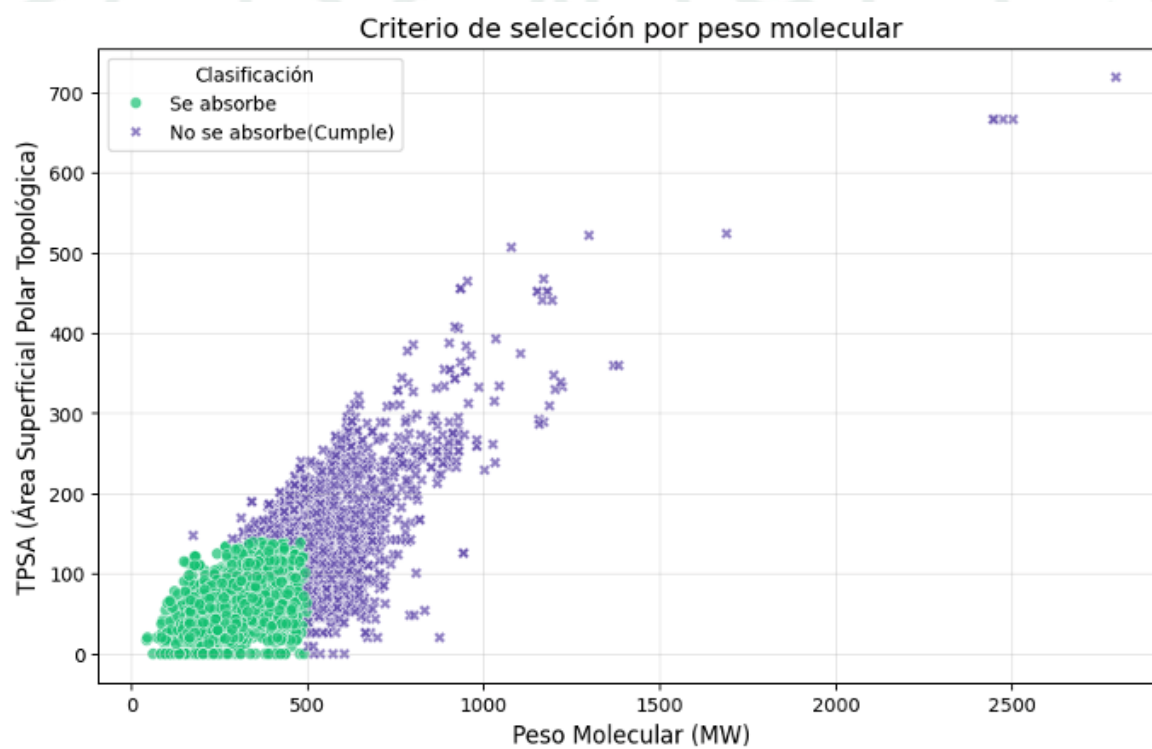


* muestra la correlación entre varias propiedades físicoquímicas de los metabolitos naturales, como el peso molecular (MW), el área superficial polar topológica (TPSA), la hidrofobicidad (SlogP), y el número de donadores y aceptores de puentes de hidrógeno Este análisis facilita la selección de compuestos con características deseadas para su estudio.

Selección de Moléculas por Absorción

Se aplicó un filtro para clasificar los compuestos según su capacidad de absorción. Los compuestos que no se absorbían cumplían con las condiciones de tener un peso molecular (MW) mayor a 500 y un Área Superficial Polar Topológica (TPSA) mayor a 140, lo que los hacía menos permeables y, por ende, menos propensos a atravesar las membranas celulares. Esta característica era ideal para los compuestos que interactuarían en el medio oral, como con las bacterias bucales. En total, se encontraron 945 moléculas que no se absorbían, lo que las convirtió en candidatas ideales para interactuar localmente con las bacterias en la boca. Por otro lado, los compuestos que sí se absorbían no cumplían con las condiciones anteriores, lo que implicaba que tenían una mayor capacidad de penetrar las membranas celulares y distribuirse en el organismo. De un total de 13,578 moléculas en la base de datos, 12,633 moléculas fueron clasificadas como absorbibles. Este filtro permitió seleccionar las moléculas con un perfil adecuado para su acción localizada en el entorno oral.

Figura 4. Criterio de selección de moléculas según MW y TPSA

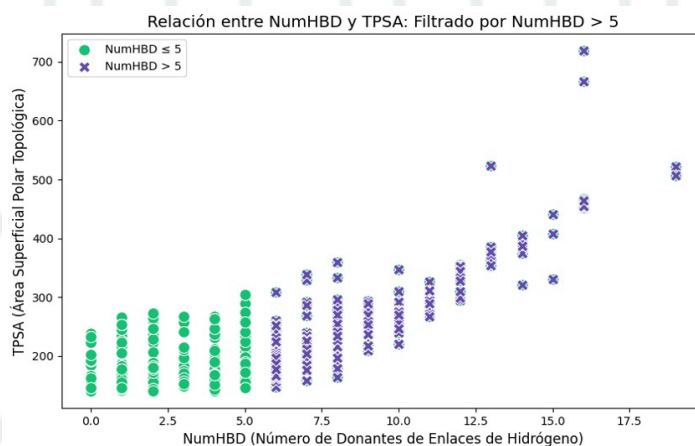


* Aquellas que se absorben (representadas por puntos verdes), las que no se absorben aquellas que cumplen con la condición de selección (puntos morados).

Selección por polaridad e interacción con el Agua

Es importante seleccionar moléculas que no sean altamente permeables en la mucosa oral para asegurar que permanezcan por mayor tiempo en la cavidad bucal y, así, interactúen eficazmente con las bacterias orales implicadas en patologías como la caries dental o la periodontitis. Desde un punto de vista científico, compuestos con baja permeabilidad celular tienen mayor probabilidad de retenerse en la superficie mucosa, lo que les permite interactuar de forma más directa y prolongada con las biopelículas bacterianas presentes en la cavidad oral. Además, una menor permeabilidad reduce el riesgo de absorción sistémica, minimizando posibles efectos secundarios o toxicidades relacionadas con la circulación sistémica de estas moléculas. Las moléculas polares con altos valores de TPSA y múltiples donantes de enlaces de hidrógeno ($\text{NumHBD} > 5$) suelen tener baja permeabilidad transmembrana, lo que favorece su retención en saliva y mucosas, permitiendo una acción antimicrobiana prolongada y localizada en el ambiente bucal. Esta estrategia aumenta la eficacia terapéutica frente a las bacterias orales, potenciando un efecto preventivo o terapéutico focalizado y disminuyendo considerablemente los efectos adversos asociados a tratamientos sistémicos. Se filtraron un total de 427 moléculas que cumplen con los requisitos.

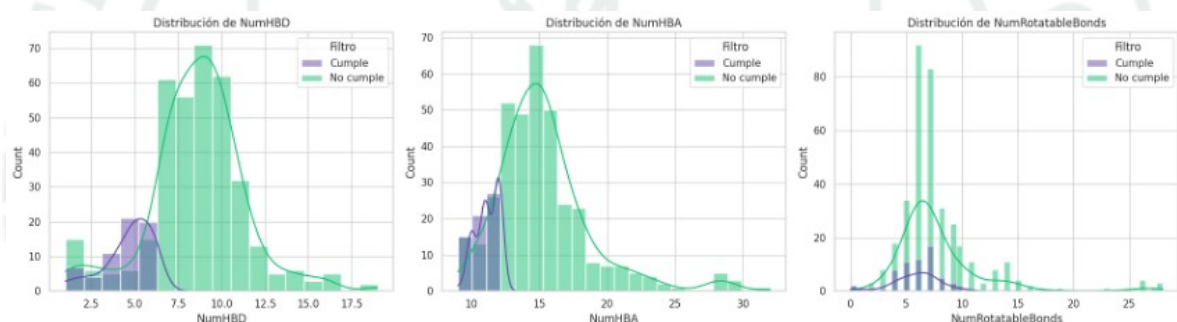
Figura 5. Relación entre NumHBD y TPSA



Selección por baja permeabilidad y acción local

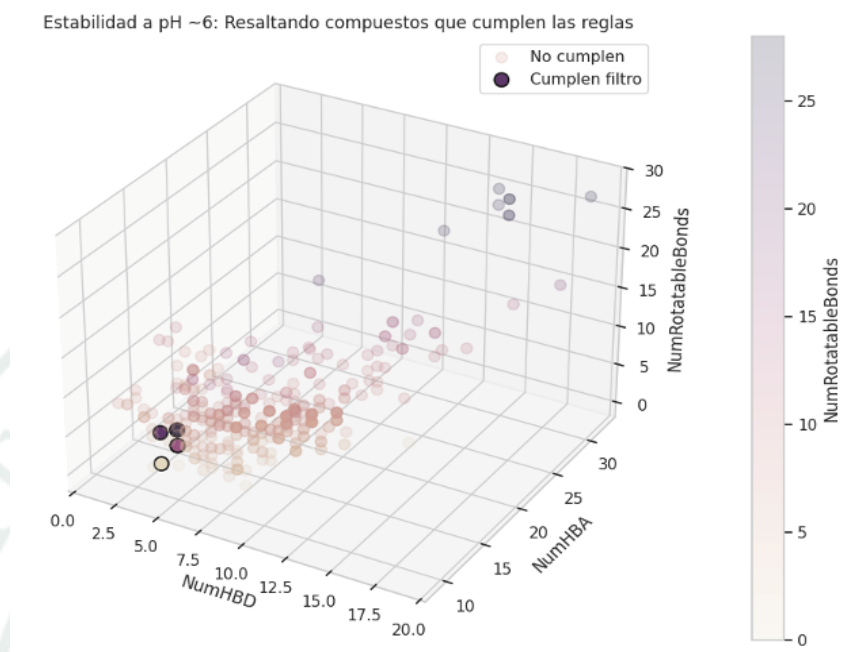
Los resultados muestran que, tras aplicar los criterios de selección, se filtraron un total de 427 moléculas inicialmente, pero después de considerar los requisitos más estrictos, solo 7 moléculas cumplieron todos los criterios establecidos. Las gráficas muestran la distribución de las propiedades clave, NumHBD, NumHBA y NumRotatableBonds, antes y después del filtro. En los gráficos de distribución, se observa cómo las moléculas que cumplen con el filtro se agrupan en rangos específicos, mientras que aquellas que no cumplen quedan distribuidas fuera de estos rangos. Este filtro se diseñó para asegurarnos de que los compuestos también funcionen bien a pH oral, la combinación de estos criterios asegura que las moléculas seleccionadas sean lo suficientemente grandes y polares para evitar la absorción sistémica, pero no tan grandes o flexibles como para interferir con su capacidad de interactuar localmente con las bacterias en la cavidad bucal.

Figura 6. Distribución de NumHBD, NumHBA y NumRotatableBonds.



*En cada gráfico, las barras representan la distribución de las moléculas que cumplen con el filtro (en color morado), mientras que la curva de densidad muestra la distribución general, incluyendo las moléculas que no cumplen con los criterios del filtro (en color verde).

Figura 7. Plot de estabilidad de pH



* Los puntos en color morado corresponden a las moléculas que cumplen con las reglas de filtrado, mientras que los puntos de color más claro representan las moléculas que no cumplen.

Propiedades Toxicológicas

Una vez aplicados los criterios de filtrado estructural y funcional a los 13,579 metabolitos naturales de la base de datos LANaPDB, se seleccionaron siete compuestos candidatos con potencial inhibitorio sobre la proteína ComD de *Streptococcus sobrinus*. Estos compuestos fueron posteriormente evaluados mediante el software DataWarrior para identificar posibles riesgos toxicológicos.

La predicción incluyó la evaluación de cuatro propiedades clave: mutagenicidad, carcinogenicidad, toxicidad reproductiva e irritación. De los siete compuestos evaluados, cuatro presentaron alguna advertencia toxicológica y fueron descartados para el análisis posterior. Solo tres compuestos no presentaron ninguna alerta en las propiedades analizadas, siendo seleccionados como propuestas preliminares con perfil seguro.

Tabla 3. Predicción de propiedades toxicológicas mediante DataWarrior.

ID	País	Mutagénico	Tumorigénico	Reproductivo	Irritante
LANaPDB731	Costa Rica	Low	None	None	None
LANaPDB12	México	None	None	None	High
LANaPDB4562	Brasil	None	None	None	None
LANaPDB6265	Brasil	None	Low	None	None
LANaPDB7944	Brasil	None	None	None	High
LANaPDB10455	Brasil	None	None	None	None
LANaPDB12628	México	None	None	None	None

Como se puede observar en la Tabla 4, los compuestos LANA PDB4562, LANA PDB10455 y LANA PDB12628 fueron los únicos que no presentaron ninguna alerta toxicológica. Estos resultados respaldan su selección como moléculas candidatas seguras dentro del enfoque propuesto.

Proteína ComD

Modelamiento por homología

La secuencia aminoacídica de la proteína ComD de *Streptococcus sobrinus* fue modelada en el servidor SWISS-MODEL, ingresando directamente la secuencia en formato FASTA. Como resultado, se generó un modelo estructural denominado “Model_01”, con una puntuación GMQE de 0.80, lo que indica una alta confiabilidad predictiva. El modelo presenta un estado oligomérico monomérico y cubre la totalidad de la secuencia objetivo (rango de residuos 1–448). El alineamiento entre la proteína ComD de *Streptococcus sobrinus* y la plantilla seleccionada por SWISS-MODEL (estructura A0A380JEM7.1.A) mostró una identidad del 100% y una cobertura completa, lo que indica que toda la proteína fue modelada de manera continua sin regiones faltantes (ver Figura 7.)

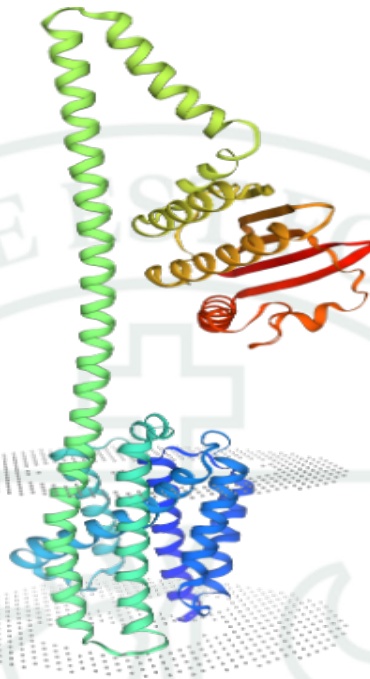
Figura 8. Alineamiento entre la secuencia de la proteína ComD

Model_01	MLSLVYLNRTFQMLLWSFVELFIIR RS SLKI	30
A0A380JEM7.1.A	MLSLVYLNRTFQMLLWSFVELFIIR RS SLKI	30
Model_01	RIPWAYQLAFLVVSSLPYLWLTEFFYANQF	60
A0A380JEM7.1.A	RIPWAYQLAFLVVSSLPYLWLTEFFYANQF	60
Model_01	FEFALICLLIFKFSK KKVTN TWLKLF FAFYV	90
A0A380JEM7.1.A	FEFALICLLIFKFSK KKVTN TWLKLF FAFYV	90
Model_01	KLGV DI IARFTLFVIVVPLFPHFAVQKLP I	120
A0A380JEM7.1.A	KLGV DI IARFTLFVIVVPLFPHFAVQKLP I	120
Model_01	VSII LQV LLTLFGNEVIIRYFKMD FLRINE	150
A0A380JEM7.1.A	VSII LQV LLTLFGNEVIIRYFKMD FLRINE	150
Model_01	VVEDKSIQSVILFHNIFFLAYFLVLWLT YV	180
A0A380JEM7.1.A	VVEDKSIQSVILFHNIFFLAYFLVLWLT YV	180
Model_01	LVTPFGLINATEAANFRINTSTLFMLVFLS	210
A0A380JEM7.1.A	LVTPFGLINATEAANFRINTSTLFMLVFLS	210
Model_01	FISMLNFKVATIEERLVSQKESELKNITD	240
A0A380JEM7.1.A	FISMLNFKVATIEERLVSQKESELKNITD	240
Model_01	YSHQIENLYAKIRSF RHDYSN ILTS LR YSA	270
A0A380JEM7.1.A	YSHQIENLYAKIRSF RHDYSN ILTS LR YSA	270
Model_01	DHDDLASVRQTLDEITRES DQIL GTKEF EI	300
A0A380JEM7.1.A	DHDDLASVRQTLDEITRES DQIL GTKEF EI	300
Model_01	ANLANX DD ALKSIFSSKLSKAIDLGL DI K	330
A0A380JEM7.1.A	ANLANX DD ALKSIFSSKLSKAIDLGL DI K	330
Model_01	VE IL NP D GLPV SFS SLDFIRLVS NLLD NAI	360
A0A380JEM7.1.A	VE IL NP D GLPV SFS SLDFIRLVS NLLD NAI	360
Model_01	EAANESEDEQLLISYFK DD TYVFILAN S T	390
A0A380JEM7.1.A	EAANESEDEQLLISYFK DD TYVFILAN S T	390
Model_01	KEKTDNIAQLFQ TGV SSK GKDRG VGLAIVK	420
A0A380JEM7.1.A	KEKTDNIAQLFQ TGV SSK GKDRG VGLAIVK	420
Model_01	KILSTYPN VRL NTSN QNG LFTQHLE VQ E	448
A0A380JEM7.1.A	KILSTYPN VRL NTSN QNG LFTQHLE VQ E	448

* Alineamiento entre la secuencia de la proteína ComD (Model_01) y la plantilla seleccionada automáticamente por SWISS-MODEL (A0A380JEM7.1.A).

Asimismo, el modelo presenta un estado monomérico y se identificó la presencia de un segmento transmembranal, coherente con la naturaleza de esta histidina quinasa como receptor sensorial en la membrana. En la Figura x, se visualiza la conformación tridimensional de la proteína ComD. La estructura se representa en forma de hélices y láminas, coloreadas en gradiente desde el extremo N-terminal (rojo) hasta el C-terminal (azul). El segmento transmembranal puede observarse inmerso en una representación esquemática de la bicapa lipídica, indicada por líneas punteadas.

Figura 9. Estructura tridimensional predicha de la proteína ComD



*Se resalta el dominio transmembranal en la región inferior.

Validación de estructura proteica ComD

La calidad estructural del modelo tridimensional de la proteína ComD fue evaluada mediante el servidor MolProbity (versión 4.4), y ProSA-web con el fin de verificar su adecuación estereoquímica para estudios posteriores de interacción proteína-ligando.

En la Figura x se muestra el diagrama de Ramachandran, el cual representa la distribución de los ángulos de torsión ϕ (phi) y ψ (psi) de cada residuo aminoacídico del modelo, indicando cuáles adoptan conformaciones geométricamente favorables según datos experimentales.

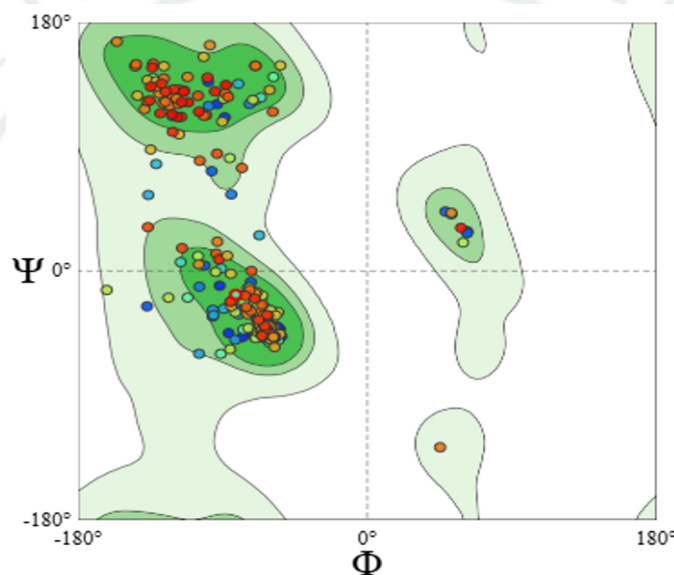
El gráfico se divide en zonas codificadas por colores:

- Las regiones verde oscuro corresponden a conformaciones altamente favorecidas, observadas comúnmente en proteínas bien plegadas.
- Las zonas verde claro representan regiones permitidas, que aunque menos frecuentes, son geométricamente aceptables.

- Las áreas blancas indican regiones no permitidas o energéticamente desfavorables, donde es poco probable que se encuentren residuos en estructuras naturales.

Cada punto en el gráfico representa un residuo, y como se observa, la gran mayoría se concentra dentro de las regiones verdes, lo cual indica que el modelo adopta una conformación estructural estable y estereoquímicamente correcta. Solo se observan unos pocos outliers fuera de estas regiones, lo cual es esperable incluso en modelos de alta calidad.

Figura 10. Diagrama de Ramachandran



*Diagrama de Ramachandran generado por MolProbity para la proteína ComD. Se aprecia la concentración de residuos en zonas verdes, lo que indica estabilidad conformacional.

Los resultados cuantitativos refuerzan la alta calidad estructural del modelo generado (ver Tabla 8). El puntaje de MolProbity (0.98) refleja una excelente calidad global, comparable a estructuras experimentales bien refinadas. El índice de colisiones fue muy bajo (0.67), indicando la casi total ausencia de colisiones atómicas significativas, con solo una interacción leve entre los residuos A167 PHE y A204 PHE. El 97.98% de los residuos se localizó en regiones favorecidas del diagrama de Ramachandran, mientras que solo un 0.45% (dos residuos) fueron clasificados como regiones esperadas, lo cual es aceptable en modelos computacionales. La presencia de 2.18% de rotámeros atípicos y una única desviación C-beta (A183 THR) se consideran tolerables y no comprometen la integridad global del modelo. Además, no se detectaron enlaces defectuosos

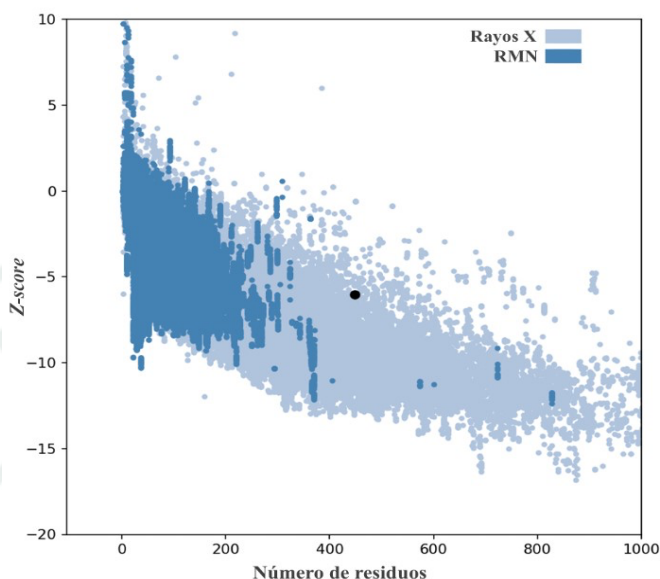
(0/3752) y solo 17 ángulos mostraron desviaciones leves, representando un 0.33% del total, dentro de los límites aceptables. En conjunto, estos resultados respaldan que la estructura predicha de ComD presenta una conformación tridimensional estereoquímicamente correcta y apta para su uso en estudios posteriores como la predicción de sitios activos y acoplamiento molecular.

Tabla 4. Resultados del análisis estructural de la proteína ComD.

Parámetro	Resultado
Puntaje MolProbity	0.98
Índice de Colisiones (A167 PHE-A204 PHE)	0.67
Residuos en regiones favorecidas	97.98%
Residuos fuera de las regiones esperadas (A295 THR, A116 GLN)	0.45%
Rotámeros Atípicos (A52 THR, A416 LEU, A179 TYR, A297 GLU.)	2.18%
Desviaciones en átomos C-Beta (A183 THR)	1
Enlaces defectuosos	0 / 3752
Ángulos defectuosos (A183 THR-A184 PRO, A107 VAL-A108 PRO, A32 ILE-A33 PRO)	17 / 5088

Como complemento al análisis estereoquímico, se utilizó ProSA-web para evaluar la calidad de la proteína modelada. En la Figura 10, se muestra el valor del *Z-score* obtenido para el modelo de la proteína ComD, el cual fue de -6.02 . Este valor se representa como un punto negro sobre un gráfico que incluye los *Z-scores* de todas las cadenas proteicas determinadas experimentalmente depositadas en el Protein Data Bank (PDB). Las estructuras determinadas por cristalografía de rayos X se representan en color azul claro, mientras que aquellas obtenidas por resonancia magnética nuclear (RMN) aparecen en azul oscuro. El *Z-score* obtenido indica la calidad global del modelo y permite verificar si dicho valor se encuentra dentro del rango típico de proteínas nativas de tamaño similar. En este caso, el *Z-score* de ComD se encuentra dentro del grupo de estructuras cristalográficas, lo que confirma que la energía total del modelo es coherente con proteínas reales, reforzando la validez estructural del mismo.

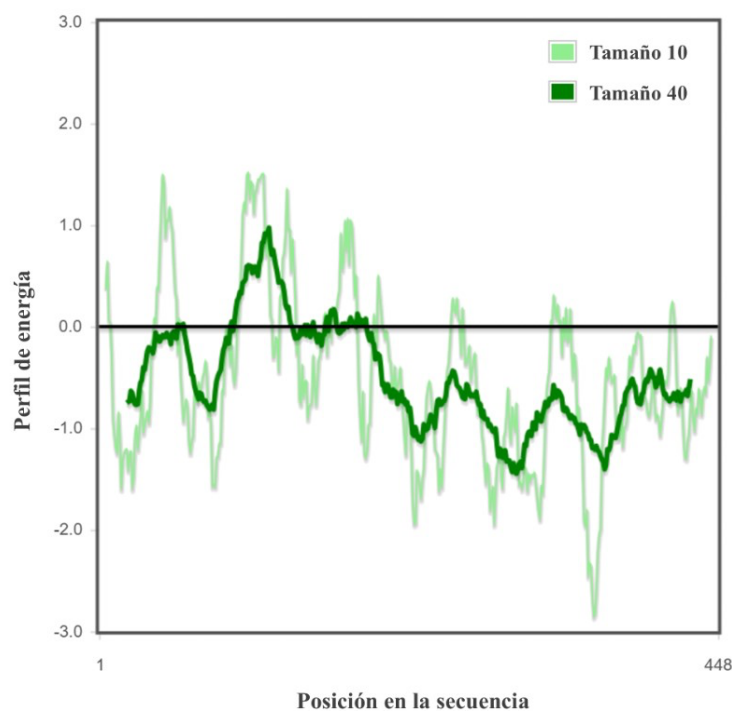
Figura 11. Comparación del Z-score del modelo de ComD



*Comparación del *Z-score* del modelo de ComD (punto negro) con estructuras determinadas experimentalmente mediante cristalografía de rayos X y RMN (resonancia magnética nuclear). El valor obtenido (-6.02) se encuentra dentro del rango esperado para proteínas nativas.

Además, se evaluó la calidad local del modelo mediante el perfil de energía por residuo, el cual permite evaluar la calidad local del modelo estructural a lo largo de su secuencia, mostrado en la Figura 4. Este gráfico representa la energía estimada de cada fragmento estructural, suavizada con ventanas móviles de 40 (línea gruesa, más estable) y 10 residuos (línea delgada, detallada pero ruidosa). Los valores positivos indican regiones potencialmente problemáticas, mientras que energías negativas sugieren zonas estables. En el modelo de ComD, se observa que la mayoría de los segmentos presentan energías negativas, lo que indica una estabilidad estructural local adecuada a lo largo de la secuencia.

Figura 12.Perfil energético por residuo



* Las energías negativas predominantes reflejan estabilidad local del modelo a lo largo de su secuencia.

La combinación de los resultados obtenidos demuestra que el modelo tridimensional de la proteína ComD presenta una alta calidad estructural, tanto a nivel estereoquímico como energético.

Predicción de Bolsillos

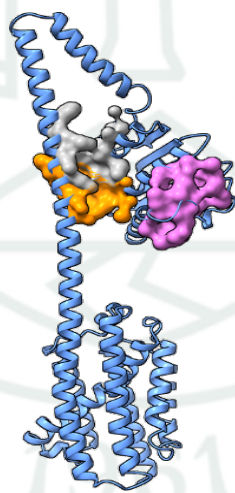
La predicción de cavidades potencialmente funcionales en la superficie de la proteína ComD se llevó a cabo utilizando el servidor PockDrug. Esta herramienta identifica bolsillos estructurales y evalúa su potencial de drugabilidad combinando descriptores estructurales con modelos estadísticos. En total, se identificaron seis bolsillos (P0 a P5), cuyos parámetros se detallan en la Tabla 6.

Tabla 5. Parámetros estructurales y probabilidad de drugabilidad de los bolsillos.

Pocket	Vol.	Hidrofob	Resid. Polares	Resid.	Átomos	Nº de residuos	drugabilidad	Desv. estándar
P0	2356.5	-0.96	0.63	0.07	0	27	0.27	0.02
P1	1950	-0.56	0.71	0.06	0	17	0.46	0.03
P2	2330.9	-0.15	0.57	0.43	0.02	23	0.94	0.05
P3	3035.9	0.08	0.52	0.22	0.03	27	0.94	0.01
P4	1685.5	-0.16	0.63	0.11	0.03	19	0.79	0.06
P5	1579.3	0.33	0.47	0.11	0	19	0.9	0.01

De los bolsillos identificados, P2, P3 y P5 presentaron los valores más altos de probabilidad de drugabilidad (0.94, 0.94 y 0.90, respectivamente), acompañados de volúmenes amplios y proporciones significativas de residuos polares y aromáticos, lo que sugiere una alta capacidad para interactuar con compuestos bioactivos. En consecuencia, los bolsillos P2, P3 y P5 fueron descartados como sitios de acoplamiento.

Figura 13. Localización estructural de los bolsillos



* Localización estructural de los bolsillos P0 (fuxia), P1 (gris) y P4 (anarajado) sobre el modelo tridimensional de la proteína ComD. Se aprecia que estos bolsillos se sitúan fuera del dominio transmembranal, en regiones potencialmente accesibles a ligandos.

Asimismo, con el fin de respaldar visualmente la selección de los bolsillos extracelulares, se generó una representación tridimensional del modelo proteico indicando la localización espacial de los bolsillos P0, P1 y P4. En la Figura x, se observa que estos tres bolsillos se encuentran fuera de la región transmembranal, en zonas accesibles al medio extracelular, lo cual refuerza su viabilidad como sitios funcionales para la interacción con ligandos.

Análisis de Afinidad Proteína-Ligando

Una vez definidos los bolsillos viables de unión en la proteína ComD, se procedió a realizar simulaciones de acoplamiento molecular mediante el servidor DockThor utilizando los mejores metabolitos seleccionados previamente. Para cada ligando se generaron tres repeticiones (poses) con el fin de evaluar la estabilidad de las interacciones y obtener un perfil energético más confiable. Los resultados se pueden observar en la Tabla 7.

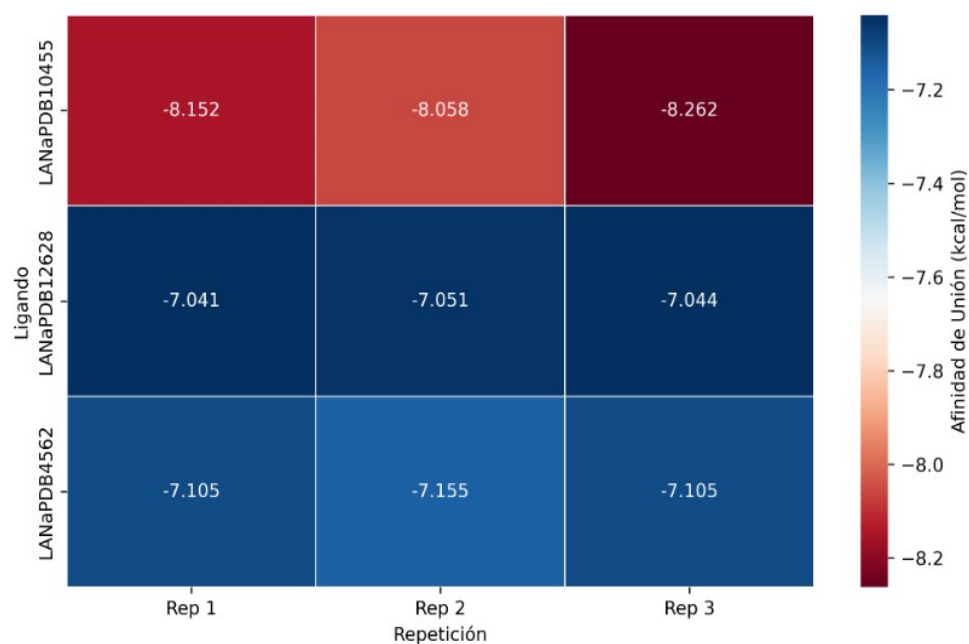
Tabla 6. Resultados del acoplamiento molecular

ID (Name)	Rep.	Score (kcal/mol)	Energía Total	Energía Interna	Energía Van der Waals	Energía Electrostática
LANaPDB4562	1	-7.105	50.129	-45.668	23.375	-69.043
LANaPDB4562	2	-7.155	50.324	-44.717	16.055	-60.772
LANaPDB4562	3	-7.105	50.129	-45.668	23.375	-69.043
LANaPDB10455	1	-8.152	73.574	-36.995	-21.44	-15.555
LANaPDB10455	2	-8.058	73.743	-34.859	-20.152	-14.707
LANaPDB10455	3	-8.262	73.312	-33.582	-22.73	-10.852
LANaPDB12628	1	-7.041	72.873	-33.141	-9.549	-23.592
LANaPDB12628	2	-7.051	72.644	-33.362	-9.603	-23.759
LANaPDB12628	3	-7.044	72.754	-33.233	-8.742	-24.491

En términos del score de acoplamiento, que representa la afinidad estimada entre el ligando y la proteína (siendo valores más negativos indicativos de una mejor interacción), el metabolito LANA PDB10455 obtuvo los mejores resultados, con un valor promedio aproximado de -8.16 kcal/mol. Esta afinidad fue superior a la de LANA PDB4562 (-7.12 kcal/mol) y LANA PDB12628 (-7.05 kcal/mol), lo cual sugiere una interacción más favorable con el sitio de unión seleccionado (Ver Figura 13). Además, LANA PDB10455 mostró energías de Van der Waals negativas y baja

energía electrostática, lo que podría indicar una interacción más estable y específica con los residuos de la proteína ComD. Por otro lado, LANaPDB4562 presentó interacciones fuertemente influenciadas por fuerzas electrostáticas, mientras que LANaPDB12628 mostró valores intermedios en todos los términos energéticos, con menor contribución de Van der Waals.

Figura 14. Mapa de calor de la afinidad de unión



* Mapa de calor de la afinidad de unión (Score en kcal/mol) para cada metabolito en tres repeticiones. Colores más cálidos (rojos) indican afinidades más altas.

El análisis espacial mostró que el ligando LANaPDB10455 se posicionó dentro del Pocket 4, una cavidad previamente identificada con un valor de probabilidad de drogabilidad de 0.79 (siendo una la mayor puntuación en comparación con la P0 y P1), y que se localiza en la región extracelular de la proteína. Este sitio fue considerado viable, tanto por su accesibilidad como por la coincidencia entre la predicción estructural y el docking realizado.

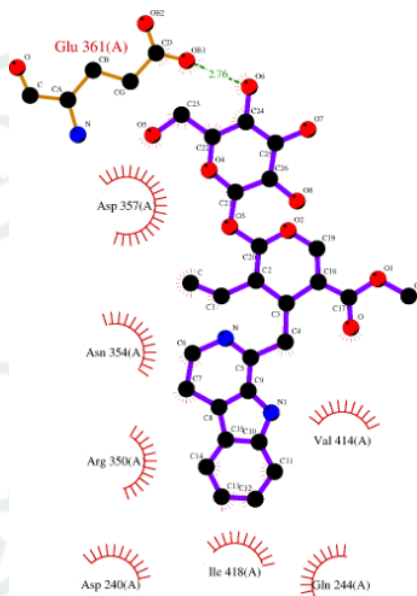
Figura 15.Representación 3D del acoplamiento de LANA_{PDB10455}



*Dentro del Pocket 4 de la proteína ComD. El bolsillo se muestra como una superficie naranja y la proteína en forma de cinta azul. El ligando se representa en forma de esferas y palos. Se observa su ubicación extracelular y su ajuste al bolsillo predicho.

Para evaluar la compatibilidad química y estérica entre el ligando y los residuos del sitio activo, se realizó un análisis 2D de interacciones con LigPlot. En la Figura 14, se observa que el ligando forma una interacción por puente de hidrógeno con el residuo Glu361(A), con una distancia de 2.76 Å. Además, se identificaron 7 contactos no enlazantes (Asp240, Gln244, Arg350, Asn354, Asp357, Val414 e Ile418), todos ellos cercanos al núcleo del bolsillo.

Figura 16.Diagrama 2D de interacciones proteína-ligando generado con LigPlot



*Se observa una interacción por puente de hidrógeno (línea punteada verde) entre el grupo OE1 de Glu361(A) y el átomo O6 del ligando. Las interacciones no enlazantes se representan con líneas rojas semicirculares. El perfil de interacción sugiere una buena compatibilidad química y geométrica entre el ligando y el bolsillo.

En conjunto, los resultados energéticos, espaciales y químicos permiten concluir que LANaPDB10455 exhibe una afinidad de interacción elevada hacia la proteína ComD, sustentada por su adecuada localización, la afinidad de acoplamiento y perfil de interacción molecular. Este compuesto es un fuerte candidato prometedor para estudios posteriores como dinámica molecular o evaluación de propiedades farmacológicas.

3.1 Discusión

El sistema ComCDE, y en particular la histidina quinasa ComD, es fundamental para la regulación de la virulencia y la capacidad de formación de biofilm de *S. sobrinus*, bacteria clave en la etiología de la caries dental. Este sistema regula la expresión de genes responsables de la producción de exopolisacáridos (EPS), adhesinas y bacteriocinas, elementos que contribuyen a la estabilidad, cohesión y resistencia del biofilm en la superficie dental. Estudios previos han corroborado que la inhibición del sistema ComCDE disminuye la capacidad cariogénica de la bacteria, reduciendo la adhesión y la producción de glucanos insolubles (51).

Por tanto, el enfoque en esta diana molecular es coherente con la necesidad de nuevas estrategias preventivas que eviten el uso indiscriminado de antibióticos y preserven la microbiota oral. La modelación por homología de la proteína ComD que presentas es una base sólida para continuar esta línea de investigación. El uso de filtros basados en peso molecular, TPSA, número de donantes y aceptores de enlaces de hidrógeno, así como la consideración de la permeabilidad celular, es una estrategia reconocida para identificar compuestos que permanezcan activos localmente en la cavidad oral sin absorción sistémica. La elección de moléculas con baja permeabilidad promueve la acción directa sobre la biofilm bacteriana sin riesgo de toxicidad sistémica. Tu filtrado que redujo inicialmente 13,579 compuestos a 7 candidatos demuestra rigurosidad y está en línea con estudios similares en farmacología oral. Sin embargo, es importante señalar que en condiciones reales, la dinámica salival y la interacción con la matriz orgánica oral podrían modificar la biodisponibilidad local, un aspecto que aún requiere validación experimental *in vitro* e *in vivo*.

La predicción toxicológica computacional es una herramienta cada vez más empleada en la selección preliminar de compuestos bioactivos (34).

Los resultados muestran que de siete compuestos candidatos, solo tres carecían de alertas toxicológicas relevantes, lo cual es crucial para futuros desarrollos en odontología preventiva. Estudios previos han señalado que algunos metabolitos naturales poseen efectos secundarios inesperados, incluyendo toxicidad reproductiva o mutagenicidad (24).

La integración de esta evaluación asegura un avance responsable hacia la formulación de agentes terapéuticos orales seguros, lo que es esencial para evitar efectos adversos sistémicos o locales en la mucosa oral. La construcción de un modelo 3D confiable de la proteína ComD es crítica, dado que no se dispone de estructura cristalográfica experimental para *S. sobrinus*. La validación mediante análisis estereoquímicos (Ramachandran) y energéticos (ProSA-web) con puntajes

excelentes respalda la utilidad del modelo para la simulación computacional. En la literatura, se ha reconocido que modelos con $GMQE \geq 0.7$ y más del 90% de residuos en regiones favorecidas son suficientemente fiables para estudios de docking y dinámica molecular (34).

La identificación y selección de bolsillos extracelulares accesibles para ligandos potencia la relevancia biológica del acoplamiento y aumenta la probabilidad de éxito en la inhibición funcional de ComD. La selección de los bolsillos P0, P1 y P4 basados en parámetros como volumen, polaridad y drugabilidad, con descarte de bolsillos transmembranales inaccesibles, es coherente con la estrategia de buscar sitios inhibitorios funcionalmente relevantes y accesibles. Estudios similares en sistemas de señalización bacteriana destacan que la inhibición de dominios extracelulares sensoriales puede interrumpir la cascada de fosforilación y la regulación de virulencia (36). No obstante, es necesario considerar que la flexibilidad dinámica de la proteína en condiciones fisiológicas podría modificar la accesibilidad y la conformación de estos bolsillos, por lo que simulaciones complementarias de dinámica molecular serían recomendables.

Los valores de afinidad de unión (scores) obtenidos por docking sugieren que LANaPDB10455 es el mejor candidato, con un score promedio de -8.16 kcal/mol, indicativo de una interacción favorable y estable. En comparación con LANaPDB4562 y LANaPDB12628, que mostraron interacciones menos fuertes, este compuesto puede ofrecer una inhibición más efectiva de ComD. Resultados similares en otras investigaciones han mostrado que valores de docking menores a -7.0 kcal/mol suelen correlacionar con inhibidores biológicamente activos (Morris et al., 2009). Sin embargo, el docking es una aproximación estática y debería complementarse con estudios dinámicos y ensayos funcionales para confirmar la inhibición biológica.

El análisis de las interacciones por LigPlot revela que el ligando LANaPDB10455 forma un puente de hidrógeno con Glu361 y múltiples contactos no enlazantes con residuos clave en el bolsillo activo. Estas interacciones son fundamentales para la estabilidad del complejo proteína-ligando y para interferir en la función quinasa de ComD. Estudios previos en histidina quinasa bacterianas muestran que la interferencia con residuos catalíticos o de unión a ligandos puede bloquear la fosforilación y la señalización, reduciendo la expresión de genes de virulencia (Casino et al., 2010). Esta evidencia computacional respalda la hipótesis de que estos metabolitos pueden modular la capacidad de *S. sobrinus* para formar biopelículas cariogénicas. Desde una perspectiva odontológica, la identificación de inhibidores naturales dirigidos al sistema ComCDE representa una innovación terapéutica con alto potencial para el control localizado de la caries dental. A

diferencia de los antimicrobianos clásicos, que pueden alterar la microbiota oral y generar resistencias, estos metabolitos podrían reducir la virulencia bacteriana sin afectar la supervivencia de bacterias beneficiosas. Además, la aplicación tópica en dentífricos, enjuagues o geles podría mejorar la prevención y el tratamiento, especialmente en poblaciones con alta incidencia de caries. No obstante, para avanzar hacia la clínica, es indispensable validar estos resultados mediante estudios in vitro con cepas clínicas de *S. sobrinus*, evaluar la estabilidad y liberación en formulaciones orales, y confirmar la ausencia de efectos adversos en modelos animales y humanos.

3.3. Conclusiones

Primero. El análisis exhaustivo de las propiedades fisicoquímicas de los 13,579 metabolitos naturales latinoamericanos permitió identificar compuestos con características compatibles para su acción localizada en la cavidad oral. La aplicación de filtros rigurosos basados en peso molecular, área superficial polar topológica (TPSA), número de donadores y aceptores de enlaces de hidrógeno, y flexibilidad molecular, resultó en la selección de siete moléculas con perfil óptimo para evitar la absorción sistémica y favorecer su permanencia en el entorno bucal. Este enfoque permitió enfocar el estudio en compuestos que presentan la mejor relación entre tamaño, polaridad y estabilidad estructural para interactuar eficazmente con la microbiota oral, minimizando riesgos de toxicidad sistémica. La correlación positiva observada entre propiedades como TPSA y MW, así como entre donadores y aceptores de enlaces de hidrógeno, facilitó una selección más precisa, evidenciando la importancia de estas características en la determinación del comportamiento farmacocinético esperado de los metabolitos.

Segundo. De los 7 metabolitos seleccionados tras el filtrado fisicoquímico, la evaluación toxicológica computacional mediante DataWarrior identificó que 4 compuestos presentaban alguna alerta en parámetros de mutagenicidad, carcinogenicidad o irritación, descartándose para análisis posteriores. Los 3 compuestos restantes (identificados como LANaPDB4562, LANaPDB10455 y LANaPDB12628) no presentaron ninguna alerta toxicológica en mutagenicidad, tumorigénesis, toxicidad reproductiva ni irritación cutánea, confirmando su perfil seguro. Esta selección representa menos del 0.03% del total inicial, demostrando la rigurosidad del proceso para asegurar seguridad potencial antes de avanzar a pruebas experimentales.

Tercero. El modelo de la proteína ComD fue generado con un valor GMQE de 0.80, lo que indica una alta confianza predictiva. El alineamiento con la plantilla mostró un 100% de identidad y cobertura completa (residuos 1-448). La validación estereoquímica arrojó que el 97.98% de los residuos se encuentran en regiones favorecidas del diagrama de Ramachandran, mientras solo el 0.45% se ubicó en regiones permitidas, y un mínimo 2.18% correspondió a rotámeros atípicos. El puntaje MolProbity fue de 0.98, y el índice de colisiones fue de apenas 0.67, confirmando la calidad estructural. Además, el Z-score obtenido con ProSA-web fue de -6.02, situándose dentro del rango de proteínas nativas cristalográficas, lo que respalda la estabilidad y confiabilidad del modelo para los análisis de interacción molecular. Los estudios de docking, el compuesto LANaPDB10455 mostró la mejor afinidad de unión con la proteína ComD, con un score promedio de -8.16 kcal/mol, destacando sobre LANaPDB4562 (-7.12 kcal/mol) y LANaPDB12628 (-7.05 kcal/mol). Este metabolito se ubicó en el bolsillo P4, identificado con una probabilidad de drugabilidad de 0.79, accesible y localizado en la región extracelular.

3.4. Recomendaciones

Primera: Fomentar la integración de herramientas computacionales avanzadas, como el modelado molecular y el docking, en la investigación odontológica para la identificación y diseño de nuevos inhibidores dirigidos a proteínas bacterianas clave, como ComD, que contribuyen a la patogenicidad de *Streptococcus sobrinus*.

Segunda: Promover la ampliación y actualización constante de bases de datos de metabolitos naturales latinoamericanos, incorporando datos experimentales y perfiles toxicológicos, para facilitar la identificación de compuestos con potencial terapéutico específico para la cavidad oral.

Tercera: Incentivar la realización de estudios experimentales in vitro e in vivo que validen la actividad inhibidora de los metabolitos identificados mediante métodos computacionales, especialmente aquellos con afinidad elevada y perfil toxicológico seguro, para avanzar en el desarrollo de nuevos agentes preventivos contra la caries dental. Profundizar en el análisis estructural y funcional del sistema ComCDE en *Streptococcus sobrinus*, incluyendo la dinámica molecular y la interacción con múltiples ligandos, para optimizar el diseño de inhibidores específicos que modulen la formación de biopelículas y la virulencia bacteriana.

REFERENCIAS BIBLIOGRÁFICAS

1. Uribe SE, Innes N, Maldupa I. The global prevalence of early childhood caries: A systematic review with meta-analysis using the WHO diagnostic criteria. *Int J Paediatr Dent.* 2021;31(6):817-30. Disponible en: <http://dx.doi.org/10.1111/ipd.12783>
2. Frencken JE, Sharma P, Stenhouse L, Green D, Laverty D, Dietrich T. Global epidemiology of dental caries and severe periodontitis - a comprehensive review. *J Clin Periodontol.* 2017;44 Suppl 18:S94-105. Disponible en: <http://dx.doi.org/10.1111/jcpe.12677>
3. Kale SS, Kakodkar P, Shetiya SH, Rizwan SA. Dental caries prevalence among 5- to 15-year-old children from SEAR countries of WHO: A systematic review and meta-analysis. *Indian J Dent Res.* 2019;30(6):937-47. Disponible en: http://dx.doi.org/10.4103/ijdr.IJDR_654_17
4. Márquez-Pérez K, Zúñiga-López CM, Torres-Rosas R, Argueta-Figueroa L. Reported prevalence of dental caries in Mexican children and teenagers. *Rev Med Inst Mex Seguro Soc.* 2023;61(5):653-60. Disponible en: <http://dx.doi.org/10.5281/zenodo.8316465>
5. Bordoni NE, Salgado PA, Squassi AF. Comparison between indexes for diagnosis and guidance for treatment of dental caries. *Acta Odontol Latinoam.* 2021;34(3):289-97. Disponible en: <http://dx.doi.org/10.54589/aol.34/3/289>
6. Schulze A, Busse M. Sports diet and oral health in athletes: A comprehensive review. *Medicina* (Kaunas). 2024;60(2):319. Disponible en: <http://dx.doi.org/10.3390/medicina60020319>
7. Askar H, Krois J, Göstemeyer G, Bottenberg P, Zero D, Banerjee A, et al. Secondary caries: what is it, and how it can be controlled, detected, and managed? *Clin Oral Investig.* 2020;24(5):1869-76. Disponible en: <http://dx.doi.org/10.1007/s00784-020-03268-7>
8. Khan S, Ul Haq I, Ali I, Rehman A, Almehmadi M, Alsuwat MA, et al. Antibacterial and antibiofilm potential of chlorophyllin against streptococcus mutans in vitro and in silico. *Antibiotics* (Basel). 2024;13(9):899. Disponible en: <http://dx.doi.org/10.3390/antibiotics13090899>
9. Cheng L, Zhang L, Yue L, Ling J, Fan M, Yang D, et al. Expert consensus on dental caries management. *Int J Oral Sci.* 2022;14(1):17. Disponible en: <http://dx.doi.org/10.1038/s41368-022-00167-3>

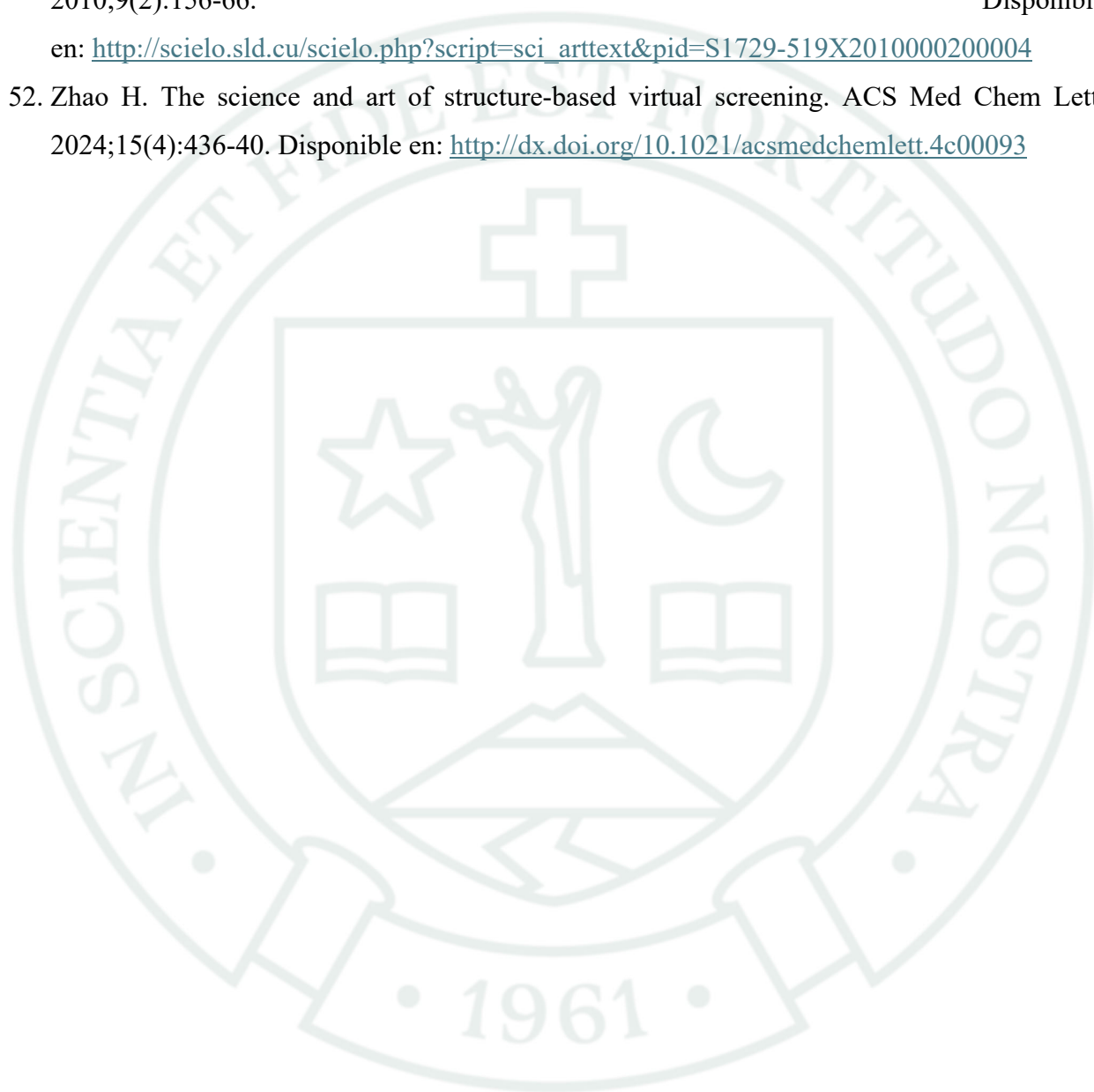
10. Head D, Marsh PD, Devine DA, Tenuta LMA. In silico modeling of hyposalivation and biofilm dysbiosis in root caries. *J Dent Res.* 2021;100(9):977-82. Disponible en: <http://dx.doi.org/10.1177/00220345211000655>
11. Zaidi S, Aswal M, Kumar M, Rashid F, Khan AU. Protein expression profiling, in silico classification and pathway analysis of cariogenic bacteria *Streptococcus mutans* under bacitracin stress conditions. *J Med Microbiol.* 2022;71(8). Disponible en: <http://dx.doi.org/10.1099/jmm.0.001572>
12. Zhang Y, Wang X, Li H, Ni C, Du Z, Yan F. Human oral microbiota and its modulation for oral health. *Biomed Pharmacother.* 2018;99:883-93. Disponible en: <http://dx.doi.org/10.1016/j.biopha.2018.01.146>
13. Homayouni Rad A, Pourjafar H, Mirzakhani E. A comprehensive review of the application of probiotics and postbiotics in oral health. *Front Cell Infect Microbiol.* 2023;13:1120995. Disponible en: <http://dx.doi.org/10.3389/fcimb.2023.1120995>
14. Mullins JGL. Drug repurposing in silico screening platforms. *Biochem Soc Trans.* 2022;50(2):747-58. Disponible en: <http://dx.doi.org/10.1042/BST20200967>
15. Jalil V, Khan M, Haider SZ, Shamim S. Investigation of the antibacterial, anti-biofilm, and antioxidative effect of Piper betle leaf extract against *Bacillus gaemokensis* MW067143 isolated from dental caries, an in vitro-in silico approach. *Microorganisms.* 2022;10(12):2485. Disponible en: <http://dx.doi.org/10.3390/microorganisms10122485>
16. Budama-Kilinc Y, Kurtur OB, Gok B, Kecel-Gunduz S, Alpay-Karaoglu S, Yilmaz Atalı P, et al. Production of prophylactic nanoformulation for dental caries and investigation of its effectiveness by in vitro and in silico methods. *Pharmaceutics.* 2025;17(2):167. Disponible en: <http://dx.doi.org/10.3390/pharmaceutics17020167>
17. van der Donk WA. Introduction: Unusual enzymology in natural product synthesis. *Chem Rev.* 2017;117(8):5223-5. Disponible en: <http://dx.doi.org/10.1021/acs.chemrev.7b00124>
18. Fan G, Li F, Wang P, Jin X, Liu R. Natural-product-mediated autophagy in the treatment of various liver diseases. *Int J Mol Sci.* 2022;23(23):15109. Disponible en: <http://dx.doi.org/10.3390/ijms232315109>
19. Wilson BAP, Thornburg CC, Henrich CJ, Grkovic T, O'Keefe BR. Creating and screening natural product libraries. *Nat Prod Rep.* 2020;37(7):893-918. Disponible en: <http://dx.doi.org/10.1039/c9np00068b>

20. Walker AS, Clardy J. Primed for discovery. *Biochemistry*. 2024;63(21):2705-13. Disponible en: <http://dx.doi.org/10.1021/acs.biochem.4c00464>
21. Zhang Q, Lei X, Yin H, Yue J. Editorial: Natural product chemical biology. *Bioorg Med Chem Lett*. 2024;109:129817. Disponible en: <http://dx.doi.org/10.1016/j.bmcl.2024.129817>
22. Atanasov AG, Waltenberger B, Pferschy-Wenzig EM, Linder T, Wawrosch C, Uhrin P, et al. Discovery and resupply of pharmacologically active plant-derived natural products: A review. *Biotechnol Adv*. 2015;33(8):1582-614. Disponible en: <http://dx.doi.org/10.1016/j.biotechadv.2015.08.001>
23. Xiaokaiti Y, Li X. Natural product regulates autophagy in cancer. *Adv Exp Med Biol*. 2020;1207:709-24. Disponible en: http://dx.doi.org/10.1007/978-981-15-4272-5_53
24. Lavecchia A. Advancing drug discovery with deep attention neural networks. *Drug Discov Today*. 2024;29(8):104067. Disponible en: <http://dx.doi.org/10.1016/j.drudis.2024.104067>
25. Wooller SK, Benstead-Hume G, Chen X, Ali Y, Pearl FMG. Bioinformatics in translational drug discovery. *Biosci Rep*. 2017;37(4):BSR20160180. Disponible en: <http://dx.doi.org/10.1042/BSR20160180>
26. Guedeney N, Cornu M, Schwalen F, Kieffer C, Voisin-Chiret AS. PROTAC technology: A new drug design for chemical biology with many challenges in drug discovery. *Drug Discov Today*. 2023;28(1):103395. Disponible en: <http://dx.doi.org/10.1016/j.drudis.2022.103395>
27. Elbadawi M, Gaisford S, Basit AW. Advanced machine-learning techniques in drug discovery. *Drug Discov Today*. 2021;26(3):769-77. Disponible en: <http://dx.doi.org/10.1016/j.drudis.2020.12.003>
28. Sadybekov AV, Katritch V. Computational approaches streamlining drug discovery. *Nature*. 2023;616(7958):673-85. Disponible en: <http://dx.doi.org/10.1038/s41586-023-05905-z>
29. Jha A, Gupta V, Adinarayan R. LANAP, periodontics and beyond: A review. *J Lasers Med Sci*. 2018;9(2):76-81. Disponible en: <http://dx.doi.org/10.15171/jlms.2018.16>
30. Barazorda-Ccahuana HL, Ranilla LG, Candia-Puma MA, Cárcamo-Rodríguez EG, Centeno-Lopez AE, Davila-Del-Carpio G, et al. PeruNPDB: the Peruvian Natural Products Database for in silico drug screening. *Sci Rep*. 2023;13(1):7577. Disponible en: <https://www.nature.com/articles/s41598-023-34729-0>

31. Lawless J, Keeler J, Gaillard A, Hazel A, Juel A. The unpredictable nature of bubble evolution. *Sci Rep*. 2022;12(1):20752. Disponible en: <https://www.nature.com/articles/s41598-022-23231-8>
32. Korona-Glowniak I, Skawinska-Bednarczyk A, Wrobel R, Pietrak J, Tkacz-Ciebiera I, Maslanko-Switala M, et al. Streptococcus sobrinus as a predominant oral bacteria related to the occurrence of dental caries in Polish children at 12 years old. *Int J Environ Res Public Health*. 2022;19(22):15005. Disponible en: <http://dx.doi.org/10.3390/ijerph192215005>
33. Li JW, Wyllie RM, Jensen PA. A novel competence pathway in the oral pathogen streptococcus sobrinus. *J Dent Res*. 2021;100(5):542-8. Disponible en: <http://dx.doi.org/10.1177/0022034520979150>
34. Veena RL, Nagarathna C. Correlation of streptococcus mutans and streptococcus sobrinus colonization with and without caries experience in preschool children. *Indian J Dent Res*. 2020;31(1):73-9. Disponible en: http://dx.doi.org/10.4103/ijdr.IJDR_432_18
35. Elyassi M, Babaeekhou L, Ghane M. Streptococcus mutans and Streptococcus sobrinus contributions in dental caries in Iranian and Afghan children: A report from serotype distribution and novel STs. *Arch Oral Biol*. 2022;139:105431. Disponible en: <http://dx.doi.org/10.1016/j.archoralbio.2022.105431>
36. Babaeekhou L, Ghane M, Ezatzade F, Eftekhari Toroghi S. Streptococcus mutans and Streptococcus sobrinus distribution in the saliva and plaque of Iranian population: Higher prevalence of S. mutans serotypes f and k. *Int J Dent Hyg*. 2021;19(2):193-200. Disponible en: <http://dx.doi.org/10.1111/idh.12485>
37. Torres MDT, Cao J, Franco OL, Lu TK, de la Fuente-Nunez C. Synthetic biology and computer-based frameworks for antimicrobial peptide discovery. *ACS Nano*. 2021;15(2):2143-64. Disponible en: <http://dx.doi.org/10.1021/acsnano.0c09509>
38. Yang F, Wang C. Profiling of post-translational modifications by chemical and computational proteomics. *Chem Commun (Camb)*. 2020;56(88):13506-19. Disponible en: <http://dx.doi.org/10.1039/d0cc05447j>
39. Schroeder WL, Baber AS, Saha R. Using EuGeneCiD and EuGeneCiM computational tools for synthetic biology. *STAR Protoc*. 2021;2(4):100820. Disponible en: <http://dx.doi.org/10.1016/j.xpro.2021.100820>

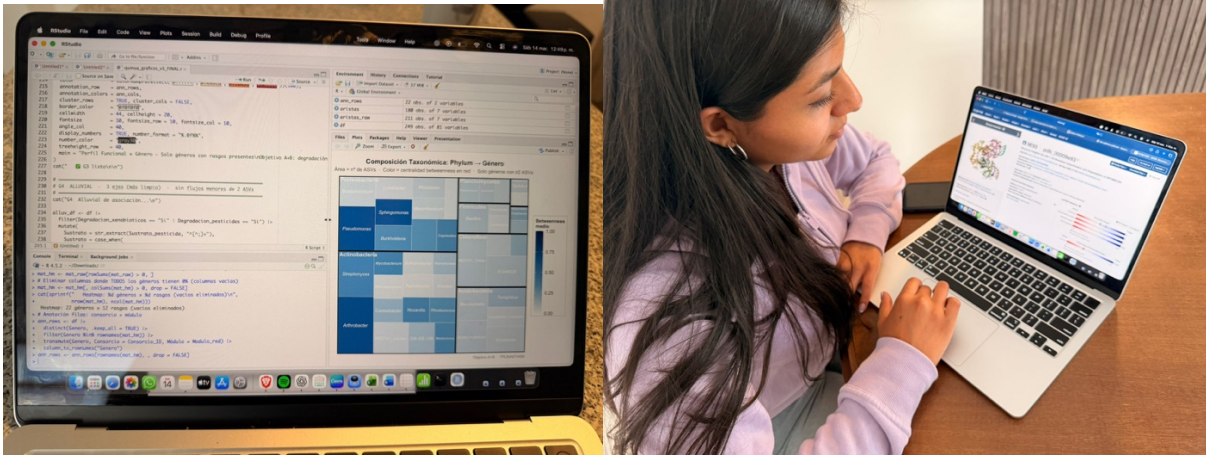
40. Oprea TI, May EE, Leitão A, Tropsha A. Computational systems chemical biology. *Methods Mol Biol.* 2011;672:459-88. Disponible en: http://dx.doi.org/10.1007/978-1-60761-839-3_18
41. Pinto GP, Hendrikse NM, Stourac J, Damborsky J, Bednar D. Virtual screening of potential anticancer drugs based on microbial products. *Semin Cancer Biol.* 2022;86(Pt 2):1207-17. Disponible en: <http://dx.doi.org/10.1016/j.semcancer.2021.07.012>
42. Thomford NE, Senthebane DA, Rowe A, Munro D, Seele P, Maroyi A, et al. Natural products for drug discovery in the 21st century: Innovations for novel drug discovery. *Int J Mol Sci.* 2018;19(6):1578. Disponible en: <http://dx.doi.org/10.3390/ijms19061578>
43. Callaway E. Major AlphaFold upgrade offers boost for drug discovery. *Nature.* 2024;629(8012):509-10. Disponible en: <http://dx.doi.org/10.1038/d41586-024-01383-z>
44. Gainza P, Sverrisson F, Monti F, Rodolà E, Boscaini D, Bronstein MM, et al. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nat Methods.* 2020;17(2):184-92. Disponible en: <http://dx.doi.org/10.1038/s41592-019-0666-6>
45. Fernández-Torras A, Comajuncosa-Creus A, Duran-Frigola M, Aloy P. Connecting chemistry and biology through molecular descriptors. *Curr Opin Chem Biol.* 2022;66:102090. Disponible en: <http://dx.doi.org/10.1016/j.cbpa.2021.09.001>
46. García-Curiel L, Del Rocío López-Cuellar M, Rodríguez-Hernández AI, Chavarría-Hernández N. Toward understanding the signals of bacteriocin production by *Streptococcus* spp. and their importance in current applications. *World J Microbiol Biotechnol.* 2021;37(1):15. Disponible en: <http://dx.doi.org/10.1007/s11274-020-02973-5>
47. Maasi G, Štšepetova J, Jõesaar M, Olak J, Mändar R. Different Patterns of Virulence Genes in *Streptococcus mutans* and *Streptococcus sobrinus* Originating from Estonian Toddlers—Mothers Cohort. *Microbiol Res (Pavia).* 2022;13(4):928-36. Disponible en: <http://dx.doi.org/10.3390/microbiolres13040065>
48. Li W, Wyllie RM, Jensen PA. A ComRS competence pathway in the oral pathogen *Streptococcus sobrinus*. *bioRxiv* [Preprint]. 2020. Disponible en: <http://dx.doi.org/10.1101/2020.03.15.992891>
49. Mendoza Ticona CA, Velasquez Talavera R, Mercado Diaz L, Ballon Echegaray J, Maguiña Vargas C. Susceptibilidad antimicrobiana de *Staphylococcus aureus* sensible, con sensibilidad “BORDERLINE” y resistentes a la meticilina. *Rev Med Hered.* 2003;14(4):181-5. Disponible en: http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1018-130X2003000400006

50. Sorokina M, Merseburger P, Rajan K, Yirik MA, Steinbeck C. COCONUT online: Collection of Open Natural Products database. *J Cheminform.* 2021;13(1):2. Disponible en: <http://dx.doi.org/10.1186/s13321-020-00478-9>
51. Núñez DP, Bacallao LG. Bioquímica de la caries dental. *Rev Habanera Cienc Méd.* 2010;9(2):156-66. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1729-519X2010000200004
52. Zhao H. The science and art of structure-based virtual screening. *ACS Med Chem Lett.* 2024;15(4):436-40. Disponible en: <http://dx.doi.org/10.1021/acsmchemlett.4c00093>



ANEXOS

Anexo 1. Trabajo de laboratorio.



Anexo 2. Base de datos utilizada.

<https://github.com/alexgoga21/LANaPDB-version-2/blob/main/LANaPDB%20version%202.xlsx>

Cuadro de nombres		Reference										Taxonomic classification					Structural Classification					Physicoch									
ID	Signal del	Count	Original	Smile	Inchi	ke	Name	Referen	DOI/Url	Year	Journal	Kingdo	Genu	Specie	Pathwa	Pathwa	Pathwa	uperCla	uperCla	uperCla	Class	Class	Class	Class	Slog	TSA	MW				
LANaP081	BIOFACQJL	Mexico	FQNP1	COc1ccc(-c)DWZAFZES	Santin	J_Ethnoph	10.1016/	2017	Journal_of_Plant	Tanacetum	Tanacetum	Shikimates and Phenylpropanoids	Flavonoids	Flavonols											2.897	98.36	344.3				
LANaP082	BIOFACQJL	Mexico	FQNP2	C=C1C=OJK PLSEPHAM	Santamarir	J_Ethnoph	10.1016/	2017	Journal_of_Plant	Tanacetum	Tanacetum	Terpenoids	Sequiterpenoids	Eudesmane sequiterpenoids											2.2114	46.53	248.3				
LANaP083	BIOFACQJL	Mexico	FQNP7	COCl=CC=ANVQZNTI	Contioxime	J_Agric_Foi	10.1021/	2016	Journal_of_Fungus	Xylaria	Xylaria	See Polyketides	Menosterpenoids	not_classified											-0.3822	59.06	170.1				
LANaP084	BIOFACQJL	Mexico	FQNP8	COCl=CC=KGGCKPCG	2-hydroxy	J_Agric_Foi	10.1021/	2016	Journal_of_Fungus	Xylaria	Xylaria	See Polyketides	Aromatic polyketides	Benzoquinones											0.5005	83.6	168.1				
LANaP085	BIOFACQJL	Mexico	FQNP9	COCl=CO)OXXPFLZ	Fumiquino	J_Agric_Foi	10.1021/	2016	Journal_of_Fungus	Xylaria	Xylaria	See Polyketides	Aromatic polyketides	Benzoquinones											0.3862	83.83	184.1				
LANaP086	BIOFACQJL	Mexico	FQNP10	COCl=CC=KUZGXBH	(4R,5S,6R)	J_Agric_Foi	10.1021/	2016	Journal_of_Fungus	Xylaria	Xylaria	See Polyketides	Miscellaneous polyketides	Miscellaneous polyketides											-0.1813	66.76	206.6				
LANaP087	BIOFACQJL	Mexico	FQNP11	NC1=CC=CZGITUQDP1	2-hydroxy	J_Agric_Foi	10.1021/	2016	Journal_of_Fungus	Xylaria	Xylaria	See Polyketides	Naphthalenes	Naphthoquinones											0.0735	66.4	167.1				
LANaP088	BIOFACQJL	Mexico	FQNP12	COCl=CC=GUACACIR	(4R,5R,6R)	biobolane	10.1021/	2016	Journal_of_Fungus	Xylaria	Xylaria	See not_classified	not_classified	not_classified											0.692	78.79	263.3				
LANaP089	BIOFACQJL	Mexico	FQNP13	CCCC1cccRFYQKUT	2-(4-Butyl)	J_Agric_Foi	10.1021/	2016	Journal_of_Fungus	Xylaria	Xylaria	See not_classified	not_classified	not_classified											3.3972	55.4	229				
LANaP0810	BIOFACQJL	Mexico	FQNP14	COc1ccc(-c)LLSDPKY6	6,6'-3"-trj	J_Nat_Proc	10.1021/	2017	Journal_of_Plant	Salvia	Salvia	See Shikimates and Phenylpropanoids	Flavonoids	Flavones											5.8825	187.49	642.5				
LANaP0811	BIOFACQJL	Mexico	FQNP15	CC1=O)CC)OYHYNA	25,5S,5R	J_Nat_Proc	10.1021/	2017	Journal_of_Plant	Salvia	Salvia	See Terpenoids	Diterpenoids	Colensane and Clerodane diterpenoids												0.4664	142.75	468.5			
LANaP0812	BIOFACQJL	Mexico	FQNP16	CC1@H)ZVLAQNYZL	25,5S,5R	J_Nat_Proc	10.1021/	2017	Journal_of_Plant	Salvia	Salvia	See Terpenoids	Diterpenoids	Colensane and Clerodane diterpenoids												0.3606	151.98	508.5			
LANaP0813	BIOFACQJL	Mexico	FQNP17	COClO)C)MHNZPW	25,5S,5R	J_Nat_Proc	10.1021/	2017	Journal_of_Plant	Salvia	Salvia	See Terpenoids	Diterpenoids	Colensane and Clerodane diterpenoids												0.3331	161.21	538			
LANaP0814	BIOFACQJL	Mexico	FQNP18	COCl@H)CCTCHSCO	25,5S,5R	J_Nat_Proc	10.1021/	2017	Journal_of_Plant	Salvia	Salvia	See Terpenoids	Diterpenoids	Colensane and Clerodane diterpenoids												0.3331	161.21	538			
LANaP0815	BIOFACQJL	Mexico	FQNP19	CC1@H)ROSSVNH	Amarisolid	J_Nat_Proc	10.1021/	2017	Journal_of_Plant	Salvia	Salvia	See Terpenoids	Diterpenoids	Colensane and Clerodane diterpenoids												1.323	138.82	492.5			
LANaP0816	BIOFACQJL	Mexico	FQNP20	COc1cc2ccQWUHUD	Pedalitin	J_Nat_Proc	10.1021/	2017	Journal_of_Plant	Salvia	Salvia	See Shikimates and Phenylpropanoids	Flavonoids	Flavones												2.291	120.36	316.2			
LANaP0817	BIOFACQJL	Mexico	FQNP21	OnC1cc(-c)KMOUOKE	7-(b-D-Gluc)	J_Nat_Proc	10.1021/	2017	Journal_of_Plant	Salvia	Salvia	See Shikimates and Phenylpropanoids	Flavonoids	Flavones												0.0499	170.05	432.3			
LANaP0818	BIOFACQJL	Mexico	FQNP22	COc1ccc(-c)QENMGKG	2-(b-D-Gluc)	J_Nat_Proc	10.1021/	2017	Journal_of_Plant	Salvia	Salvia	See Shikimates and Phenylpropanoids	Flavonoids	Flavones												2.897	98.36	344.3			
LANaP0819	BIOFACQJL	Mexico	FQNP23	COCl=CC)NWZOCF	3R	(6d)trj	J_Nat_Proc	10.1021/	2015	Journal_of_Fungus	Penicillium	Penicillium	Terpenoids	Sequiterpenoids	Eremophilane sequiterpenoids												1.9899	57.53	250.3		
LANaP0820	BIOFACQJL	Mexico	FQNP24	CC1C)C1YDMLUSV	Isopetasol	J_Nat_Proc	10.1021/	2015	Journal_of_Fungus	Penicillium	Penicillium	Terpenoids	Sequiterpenoids	Eremophilane sequiterpenoids												3.0191	37.3	234.3			
LANaP0821	BIOFACQJL	Mexico	FQNP25	CC1C)C1ZPCBDOXN	Sporogon	J_Nat_Proc	10.1021/	2015	Journal_of_Fungus	Penicillium	Penicillium	Terpenoids	Sequiterpenoids	Eremophilane sequiterpenoids												2.0063	49.83	248.3			
LANaP0822	BIOFACQJL	Mexico	FQNP26	CC1C)C1ZCHWPMFV	Hydrospog	J_Nat_Proc	10.1021/	2015	Journal_of_Fungus	Penicillium	Penicillium	Terpenoids	Sequiterpenoids	Eremophilane sequiterpenoids												1.7981	52.99	250.3			
LANaP0823	BIOFACQJL	Mexico	FQNP27	COCl=CC)BEAOSMH	Minimoidi	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Preussia	Preussia	See Polyketides	not_classified	not_classified												2.0882	102.29	384.3			
LANaP0824	BIOFACQJL	Mexico	FQNP28	COCl=CC)HJVLQSDC	Minimoidi	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Preussia	Preussia	See Polyketides	not_classified	Naphthoquinones													3.4457	111.52	462.4		
LANaP0825	BIOFACQJL	Mexico	FQNP29	COCl=CC)FHWNDNK	Preussioch	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Preussia	Preussia	See Polyketides	Chromanes	Chromanes													0.4854	102.29	292.2		
LANaP0826	BIOFACQJL	Mexico	FQNP30	COCl=CC)OZDYAFI	Corymbifer	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Preussia	Preussia	See Shikimates and Phenylpropanoids	Coumarins	Isooumarins													1.6908	85.97	274.2		
LANaP0827	BIOFACQJL	Mexico	FQNP31	COCl=CC)ZNXZBPCI	8-chloroxy	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Preussia	Preussia	See Polyketides	Naphthalenes	Naphthoquinones													1.61852	72.83	288.2		
LANaP0828	BIOFACQJL	Mexico	FQNP32	OnC1=CC=2WHBSVFD	8-chloroxy	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Malbranch	Malbranch	Polyketides	not_classified	not_classified													2.1156	46.53	210.6		
LANaP0829	BIOFACQJL	Mexico	FQNP33	CC1@H)GUNXKKU	Flavorosio	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Malbranch	Malbranch	Shikimates and Phenylpropanoids	Coumarins	Isooumarins													-0.983	145.91	34		
LANaP0830	BIOFACQJL	Mexico	FQNP34	OnC1=CC=2OMBMYVY	Xylarinol	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Malbranch	Malbranch	Polyketides	Cyclic polyketides	Phthalide-derivatives														1.4622	46.53	176.1	
LANaP0831	BIOFACQJL	Mexico	FQNP35	CC1@H)O)VVNBHMY	Xylarinol	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Malbranch	Malbranch	Polyketides	Aromatic polyketides	Simple aromatic polyketides														1.0953	69.92	224.2	
LANaP0832	BIOFACQJL	Mexico	FQNP36	CC1=CC)O)QVNHAFD	Massarigen	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Malbranch	Malbranch	Polyketides	Cyclic polyketides	Simple cyclic polyketides														-0.2359	83.83	224.2	
LANaP0833	BIOFACQJL	Mexico	FQNP37	CC1=CC)O)QVNHAFD	Massarigen	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Malbranch	Malbranch	Polyketides	not_classified	not_classified														-0.4599	83.83	224.2	
LANaP0834	BIOFACQJL	Mexico	FQNP38	CC1=O)C1ccAMZVYFV	Clavatos	J_Nat_Proc	10.1021/	2017	Journal_of_Fungus	Malbranch	Malbranch	not_classified	not_classified	not_classified														1.91724	57.53	180.2	
LANaP0835	BIOFACQJL	Mexico	FQNP39	CCCC1@H)GFARNGPI	Petalotin	J_Nat_Proc	10.1021/	2015	Journal_of_Fungus	Xylaria	Xylaria	See Polyketides	not_classified	not_classified														-0.1384	123.91	390.4	
LANaP0836	BIOFACQJL	Mexico	FQNP40	CC1@H)STSHADG	15,4b-(4)-	J_Nat_Proc	10.1021/	2015	Journal_of_Fungus	Xylaria	Xylaria	See Shikimates and Phenylpropanoids	Coumarins	Isooumarins														0.3845	66.76	194.1	
LANaP0837	BIOFACQJL	Mexico	FQNP42	CC1=CC)O)QVNHAFD	Massarigen	J_Nat_Proc	10.1021/	2015	Journal_of_Fungus	Xylaria	Xylaria	See Shikimates and Phenylpropanoids	Coumarins	Isooumarins															1.7966	35.53	192.2