

**INSTITUTO POLITÉCNICO NACIONAL
CENTRO DE BIOTECNOLOGÍA GENÓMICA**



**“CRIBADO VIRTUAL INTELIGENTE PARA IDENTIFICAR
INHIBIDORES MULTI-BLANCO ENFOCADOS AL TRATAMIENTO DE LA
ASOCIACIÓN ALZHEIMER-DIABETES MELLITUS”**

ANTEPROYECTO DE INVESTIGACIÓN

**PRESENTA
ELIUD ULISES AGUILAR DURÁN**

CD. REYNOSA, TAMAULIPAS, MÉXICO

JUNIO, 2024

**INSTITUTO POLITÉCNICO NACIONAL
CENTRO DE BIOTECNOLOGÍA GENÓMICA**



**“CRIBADO VIRTUAL INTELIGENTE PARA IDENTIFICAR
INHIBIDORES MULTI-BLANCO ENFOCADOS AL TRATAMIENTO DE LA
ASOCIACIÓN ALZHEIMER-DIABETES MELLITUS”**

ANTEPROYECTO DE INVESTIGACIÓN

**PRESENTA
ELIUD ULISES AGUILAR DURÁN**

CD. REYNOSA, TAMAULIPAS, MÉXICO

JUNIO, 2024

ÍNDICE

Sección	Página
LISTA DE FIGURAS.....	V
LISTA DE SÍMBOLOS Y/O NOMENCLATURA.....	VI
INTRODUCCIÓN	1
2. ANTECEDENTES	2
2.1. La demencia	2
2.2. Alzheimer	2
2.2.1. El Alzheimer.....	2
2.2.2. Fisiopatología del Alzheimer	4
2.2.3. Patogénesis del Alzheimer	6
2.3. Diabetes mellitus	7
2.3.1. La Diabetes Mellitus	7
2.3.2. Fisiopatología y patogénesis de la diabetes mellitus.....	8
2.4. Relación Alzheimer-Diabetes Mellitus	9
2.5. Diseño de fármacos	13
2.5.1. El diseño de fármacos	13
2.5.2. La inteligencia artificial y el diseño de fármacos.....	15
2.5.3 Aprendizaje Automático	15
2.5.4. Aprendizaje supervisado	16
2.5.5. Aprendizaje no supervisado	18
2.5.6 Aprendizaje profundo.....	19
2.6. Diseño de fármacos para la enfermedad de Alzheimer	20

2.7. Proteínas objetivo.....	25
2.7.1. STAT3.....	¡Error! Marcador no definido.
2.7.2.....	¡Error! Marcador no definido.
2.7.3. IRS1.....	¡Error! Marcador no definido.
3. JUSTIFICACIÓN	29
4. HIPÓTESIS	30
5. OBJETIVOS	31
6. MATERIALES Y MÉTODOS	32
6.1. Selección de las proteínas diana.....	¡Error! Marcador no definido.
6.2. Obtención de datos.	32
7. BIBLIOGRAFÍA.....	40

LISTA DE FIGURAS

Figura

Página

LISTA DE SÍMBOLOS Y/O NOMENCLATURA

EA	Enfermedad de Alzheimer
(A β)	beta-amiloide
APP	(APP, por sus siglas en inglés "Amyloid Precursor Protein")
FDA	por sus siglas en inglés "Food and Drug Administration"
proteína precursora amiloide	(APP, por sus siglas en inglés "Amyloid Precursor Protein")
(NMDA).	N-metil-aspartato
ADAM	por sus siglas en inglés "A Disintegrin And Metalloproteinase"
APOE	Alipoproteína E
PSEN1	Presenilina 1
PSEN2	Presenilina 2
(DM1)	diabetes tipo 1
(DM2)	diabetes tipo 2
antígeno leucocitario humano	(HLA, por sus siglas en inglés 'human leukocyte antigen')
respuesta de proteínas desplegadas	(UPR, por sus siglas en inglés "Unfolded Protein Responses")
factor de crecimiento similar a la insulina	(IGF-1, por sus siglas en inglés "Insuline Grow Factor")
receptor de insulina	RI
enzima degradadora de insulina	(IDE, por sus siglas en inglés "Insulin Degrading Enzyme")
(CADD)	por sus siglas en inglés "Computer, Aided Drug Design"

(DT, por sus siglas en inglés “Decision Trees”)	los árboles de decisión
diseño de fármacos basado en la estructura	(SBDD, por sus siglas en inglés “Structure Based Drug Design”) y
diseño de fármacos basado en el ligando	(LBDD, por sus siglas en inglés “Ligand Based Drug Design”)
(IA)	inteligencia artificial
aprendizaje automático	(ML, por sus siglas en inglés “Machine Learning”)
Aprendizaje profundo	(DL, por sus siglas en inglés “Deep Learning
máquinas de soporte vectorial	(SVM, por sus siglas en inglés, “Support Vector Machines”)
algoritmo de vecinos más cercanos	(k-NN, por sus siglas en inglés “k-Nearest Neighbors)
DBSCAN	(por sus siglas en inglés “Density-Based Spatial Clustering of Applications with Noise”),
(PCA, por sus siglas en inglés “Principal Component Analysis”)	análisis de componentes principales
t-SNE	(por sus siglas en inglés “Stochastic Neighbor Embedding”)
N-metil-D-aspartato	(NMDA)
modelos QSAR	(Relación Cuantitativa Estructura-Actividad, por sus siglas en inglés “Quantitative Structure-Activity Relationship”)
IC50	
(CPI)	interacciones químico-proteína
nM	
STAT3	(por sus siglas en inglés “Signal Transducer and Activator of Transcription 3”)

EGFR	(Epidermal Growth Factor Receptor)
IRS1	(por sus siglas en inglés “Insulin Receptor Substrate 1”)

1. INTRODUCCIÓN

En la actualidad, la demencia es la séptima causa de defunción a nivel mundial y es una de las principales razones de discapacidad en personas de edad avanzada (OPS, 2023a). Además, la enfermedad de Alzheimer es el tipo más común de demencia, abarcando hasta casi el 70% de los casos (WHO, 2023). Tan solo en México alrededor un millón 300 mil personas padecen la enfermedad de Alzheimer (SSA México, 2021). Las personas afectadas con EA experimentan un deterioro gradual en habilidades cognitivas como la memoria, la orientación, el lenguaje y el aprendizaje (IMSS, 2015).

Existen varios medicamentos aprobados por la Administración de Alimentos y Medicamentos de los Estados Unidos de América para tratar la EA. Algunos de estos medicamentos son la galantamina, la rivastigmina y el donepezilo (Uddin et al., 2020). Sin embargo, estos medicamentos están orientados a disminuir los efectos de la EA en un determinado momento y tienen como efectos secundarios náuseas, vómitos, diarrea, reacciones alérgicas, pérdida de apetito, dolor de cabeza, confusión, mareo y caídas (NIH, 2023). Uno de los principales problemas del diseño de nuevos fármacos es su baja tasa de éxito, pues alta tasa de deserción de los fármacos. Se encontró una tasa de éxito general de solo el 0.4% (Cummings et al., 2014). Para abordar estos problemas, estudios han demostrado que se deberían de diseñar fármacos multidiana debido a la naturaleza multifactorial de la EA, así como que se ha asociado está a varias causas (Lovestone et al., 2015). Recientemente, se ha asociado como una de estas causas a la diabetes mellitus (Yang & Song, 2013). Y múltiples estudios han relacionado a la EA con la diabetes mellitus afectada (Brands et al., 2005; Kopf & Frölich, 2009; Strachan et al., 1997). Asimismo, se han probado diversos medicamentos destinados a la diabetes mellitus para pacientes de EA y estos han resultado en que pueden mejorar la cognición y la memoria en pacientes con EA (Michailidis et al., 2022).

Sin embargo, en la actualidad no existe ningún estudio acerca de diseño de fármacos para la asociación Alzheimer-Diabetes mellitus. Por lo que, en esta investigación propongo emplear modelos de aprendizaje automático sobre blancos relacionados con la asociación Alzheimer-Diabetes mellitus para determinar potenciales inhibidores de este.

2. ANTECEDENTES

2.1. La demencia

La demencia es un término empleado para diversas enfermedades que son crónicas y progresivas y que resultan en deterioros cognitivos que interfieren en la capacidad de realizar actividades de la vida diaria (OPS, 2023a). La demencia suele manifestarse en edades avanzadas, siendo poco común en personas menores de 60 años. (Medline, 2024). En la actualidad, esta es la séptima causa de defunción y una de las principales razones de discapacidad y dependencia entre la población anciana a nivel mundial. En todo el mundo, más de 55 millones de personas viven con demencia, generando un coste anual de \$1 billón de dólares en 2018. En el continente americano, más de 10 millones de personas viven con esta condición. Las proyecciones a futuro indican que el número de personas afectadas por este trastorno se duplicará cada 20 años (OPS, 2023a).

2.2. Alzheimer

2.2.1. El Alzheimer

La enfermedad de **Alzheimer (EA)** es la forma más prevalente de demencia, abarcando entre el 60% y el 70% de los casos (WHO, 2023). Esta fue descrita por primera vez en el año 1906 en Alemania por el psiquiatra Alois Alzheimer (Stelzmann et al., 1995). Esta demencia neurodegenerativa se caracteriza por el deterioro progresivo de las funciones mentales, afectando las neuronas, provocando su degeneración y muerte. Las personas afectadas experimentan un deterioro gradual en habilidades cognitivas como la memoria, la orientación, el lenguaje, el aprendizaje, entre otros (IMSS, 2015). Se estima que en México aproximadamente un millón 300 mil personas padecen la enfermedad de Alzheimer, lo que representa entre el 60 y el 70 por ciento de los casos de demencia diagnosticados. Esta condición afecta principalmente a personas mayores de 65 años (SSA México, 2021).

La Administración de Alimentos y Medicamentos de los Estados Unidos de América (FDA, por sus siglas en inglés “Food and Drug Administration”) ha aprobado varios medicamentos para su uso para la enfermedad de Alzheimer. Estos medicamentos son para controlar los síntomas o tratar la enfermedad y funcionan mejor para personas en etapas tempranas o intermedias, sin embargo, no se conocen medicamentos que curen esta enfermedad. Algunos de estos medicamentos son la galantamina, la rivastigmina y el donepezilo, los cuales son inhibidores de la colinesterasa. Estos medicamentos pueden ayudar a reducir o controlar algunos síntomas cognitivos y conductuales. Otro medicamento es la memantina, un antagonista del *N*-metil-aspartato (NMDA). Este podría permitir al paciente mantener funciones diarias por un poco más de tiempo del que lo harían sin el medicamento (Uddin et al., 2020). Sin embargo, todos estos medicamentos se recetan en dosis bajas para reducir la probabilidad de los efectos secundarios. Tienen posibles efectos secundarios, que incluyen náuseas, vómitos, diarrea, reacciones alérgicas, pérdida de apetito, dolor de cabeza, confusión, mareo y caídas (NIH, 2023).

Las características principales de la EA son la formación de placas amiloides y la acumulación de ovillos neurofibrilares. Las placas amiloides, compuestas principalmente por la proteína beta-amiloide ($A\beta$), se generan a partir del procesamiento secuencial de la proteína precursora amiloide (APP, por sus siglas en inglés “Amyloid Precursor Protein”) y se acumulan entre las neuronas. Estas placas interfieren en la comunicación entre las células nerviosas, provocan inflamación y activan una respuesta inmunitaria que contribuye a la muerte neuronal. (Sun et al., 2012). Por otro lado, los ovillos neurofibrilares son formados por la proteína tau, la cual es un grupo de proteínas asociadas a los microtúbulos neuronales que se forman a través de un proceso de empalme alternativo del ARNm. Estos ovillos neurofibrilares se acumulan en el cerebro durante el desarrollo de la enfermedad de Alzheimer (Johnson & Stoothoff, 2004). Esta acumulación de tau altera el transporte de nutrientes y otras moléculas esenciales dentro de las neuronas, lo que finalmente conduce a la disfunción y muerte de las células neuronales. (Zhang et al., 2021). Existen diversas teorías acerca de la fisiopatología del EA. Entre ellas, se encuentran la hipótesis amiloide, la hipótesis colinérgica, el papel de la proteína Tau y la implicación del estrés oxidativo y el calcio (Sanabria-Castro et al., 2017).

2.2.2. Fisiopatología del Alzheimer

La hipótesis de la proteína amiloide postula que la enfermedad de Alzheimer se debe a anomalías en el procesamiento de la APP. En este contexto, el desequilibrio entre la producción y la eliminación de beta-amiloide es responsable de la formación anormal de placas amiloides, las cuales provocan la muerte de las células neuronales. (Cummings et al., 2007). El péptido beta-amiloide es un péptido compuesto por entre 37 y 43 aminoácidos, conocido por ser altamente resistente a la degradación proteolítica. Las isoformas más comunes son A β 1-40 y A β 1-42 (Deane et al., 2009), siendo la isoforma A β 1-42 la más hidrofóbica y la que tiene la mayor toxicidad (Mohandas et al., 2009).

La beta-amiloide se origina a través del procesamiento de la APP en la membrana plasmática, como se ilustra en la **Figura 1** (Sanabria-Castro et al., 2017). En el proceso normal, la APP es descompuesta principalmente por enzimas con actividad α -secretasa, que pertenecen a las familias de desintegrina y metaloproteasa (ADAM, por sus siglas en inglés “A Disintegrin And Metalloproteinase”). La escisión de la APP por las ADAM resulta en la formación y liberación de APPs α , una forma soluble bajo ciertas condiciones. (Tanzi & Bertram, 2005). Durante el proceso de producción de beta-amiloide, la primera escisión de APP genera un fragmento amino terminal más corto conocido como APPs β . Esta división es realizada por la enzima BACE1, una proteasa transmembrana ubicua con actividad de β -secretasa. Posteriormente, la γ -secretasa, un complejo compuesto por cuatro subunidades: presenilinas, nicastrina, PEN-2 y APH-1, lleva a cabo otra escisión que libera el péptido beta-amiloide (A β). (Sanabria-Castro et al., 2017).

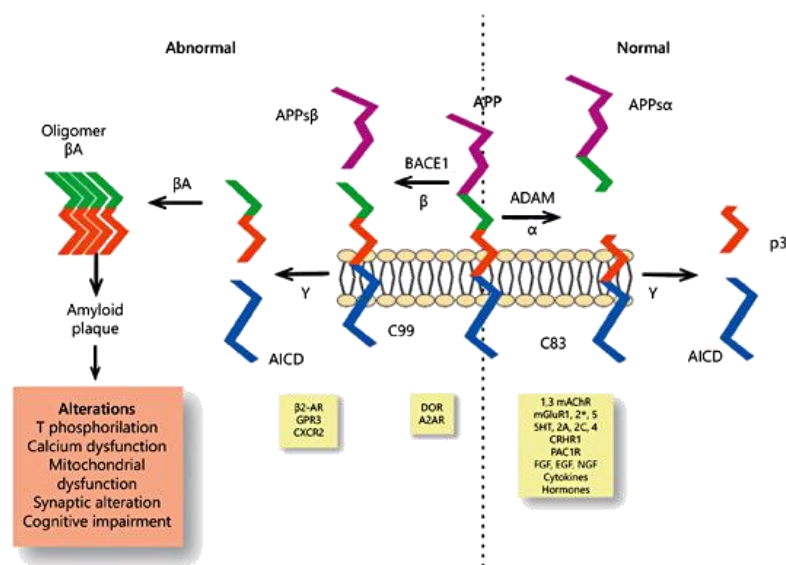


Figura 1. (Sanabria-Castro et al., 2017)

La proteína tau es el componente principal de los filamentos helicoidales que conforman los ovillos neurofibrilares en el cerebro de personas con EA, y en estos, esta proteína se encuentra anormalmente fosforilada (Grundke-Iqbal et al., 1986). Tau es una fosfoproteína, y su fosforilación anormal impide su capacidad para estimular el ensamblaje de microtúbulos. Además, la proteína Tau es una proteína neuronal que presenta hasta seis isoformas distintas. (Johnson & Stoothoff, 2004). Esta proteína contiene 80 residuos de serina o treonina y 5 residuos de tirosina, lo que significa que casi el 20% de la proteína tiene el potencial de ser fosforilada por proteínas quinasas (Goedert et al., 1989). La proteína Tau probablemente desempeña un papel en la regulación de la extensión de las neuritas, ya que estudios han demostrado que la supresión de su expresión inhibe el crecimiento de estas. Además, está implicada en la regulación del transporte axonal (Ishihara et al., 1999). Sin embargo, la fosforilación inapropiada de la proteína Tau está presente en enfermedades neurodegenerativas como la EA, donde resulta en la incapacidad de esta proteína para ensamblar microtúbulos, lo que conduce a la formación de ovillos neurofibrilares (Johnson & Stoothoff, 2004).

La hipótesis colinérgica propone que la disfunción de las neuronas que contienen acetilcolina contribuye al deterioro cognitivo observado tanto en el envejecimiento normal como en la enfermedad de Alzheimer (Terry & Buccafusco, 2003). Los receptores muscarínicos son un

tipo de receptor de acetilcolina presentes en las células del sistema nervioso y en diversos tejidos corporales. Los antagonistas de estos receptores, como la escopolamina, bloquean la acción de la acetilcolina en estos receptores, lo que afecta la vía de señalización de la acetilcolina y, por ende, la memoria y otras funciones cognitivas (Liskowsky & Schliebs, 2006). Además, se ha comprobado que tanto los agonistas muscarínicos selectivos como los no selectivos mejoran el aprendizaje y la memoria. Se han identificado agonistas selectivos para los receptores muscarínicos M1, los cuales han demostrado reducir los niveles de beta-amiloide en el líquido cefalorraquídeo (Fisher, 2008). Actualmente, esta hipótesis ha sido la base de varios tratamientos, y existe un consenso en que la relación entre el deterioro cognitivo y la reducción de la transmisión colinérgica en el cerebro desempeña un papel crucial en la enfermedad de Alzheimer. (Pena et al., 2006).

2.2.3. Patogénesis del Alzheimer

La enfermedad de Alzheimer es una afección multifactorial que involucra tanto factores genéticos como ambientales en su patogenia. Entre estos factores se encuentran principalmente la edad y los antecedentes familiares positivos de demencia (Kandimalla et al., 2017). Los tipos de Alzheimer se dividen en dos grandes grupos según el momento en que aparecen los primeros signos. En el caso del Alzheimer de inicio temprano, los síntomas surgen por primera vez entre los 30 y los 65 años, mientras que, en el Alzheimer de inicio tardío, los signos comienzan a manifestarse alrededor de los 65 años (NIH, 2017). Existen cuatro subtipos de la enfermedad de Alzheimer (EA), clasificados según el gen afectado (Bird, 2008). El primer subtipo, conocido como EA1, implica mutaciones en el gen de la APP. Estas mutaciones pueden llevar a una producción excesiva del péptido beta-amiloide, lo que contribuye a la formación de placas amiloides en el cerebro. (Weggen & Beher, 2012). La apolipoproteína está implicada en el subtipo 2. La **apolipoproteína E (APOE)** es una proteína de unión a lípidos y constituye la principal proteína de transporte de colesterol en el cerebro. Aunque su función principal es el transporte, también interactúa con la beta amiloide, los betaglucanos y los ácidos lipoteicoicos (Mahley & Rall, 2000). En los seres humanos, existen tres isoformas de APOE, conocidas como $\epsilon 2$, $\epsilon 3$ y $\epsilon 4$. Las personas portadoras del alelo $\epsilon 2$ tienen un menor riesgo de desarrollar la enfermedad de Alzheimer (EA), mientras que aquellas con el alelo $\epsilon 4$ tienen un riesgo de 10 a

15 veces mayor. La presencia del alelo $\epsilon 4$ se considera el factor de riesgo genético más fuerte para la EA. (Troutwine et al., 2022). En el subtipo EA3, el gen implicado es el **PSEN1**, también conocido como STM2, como una de las causas del Alzheimer. Este es el subtipo de EA más común. Este gen humano codifica la proteína **presenilina 1**. Se ha identificado que una mutación puntual en el gen PSEN1 resulta en un cambio de isoleucina por una asparagina. Estas mutaciones pueden alterar la actividad de la γ -secretasa, una enzima que participa en el procesamiento de la APP, lo que puede llevar a un aumento en la producción de beta-amiloide y a la formación de placas amiloides características de la enfermedad de Alzheimer (Levy-Lahad et al., 1995). El subtipo 4 de la enfermedad de Alzheimer (EA) está relacionado con mutaciones en el gen **PSEN2**. Al igual que las mutaciones en PSEN1, estas mutaciones también afectan la función de la γ -secretasa. Sin embargo, este subtipo es menos común que el subtipo EA3 (Rogaev et al., 1995). Sin embargo, la causa de la enfermedad de Alzheimer de inicio tardío sigue siendo desconocida, aunque se han identificado muchos factores de riesgo. Entre estos, la diabetes se destaca como un fuerte factor de riesgo para el desarrollo de la EA (Yang & Song, 2013).

2.3. Diabetes mellitus

2.3.1. La Diabetes Mellitus

La diabetes mellitus es una enfermedad caracterizada por niveles elevados de glucosa en la sangre. A largo plazo, esto incrementa el riesgo de daño a los ojos, riñones, páncreas, nervios y corazón. Existen dos principales tipos de diabetes, también conocidas como diabetes mellitus: la **diabetes tipo 1 (DM1)** y la **diabetes tipo 2 (DM2)**. (NIDDK, 2023). La diabetes tipo 1 se caracteriza principalmente por la destrucción autoinmune de las células beta del páncreas, las cuales son responsables de la producción de insulina, mientras que la diabetes tipo 2 se distingue principalmente por dos defectos fisiológicos: la resistencia a la insulina y la disfunción progresiva de las células beta del páncreas. (Medline, 2020). Se estima que 62 millones de personas en el continente americano viven con diabetes tipo 2. Además, se calcula que este número se ha triplicado desde 1980 y podría alcanzar los 109 millones de pacientes para el año

2040. En 2020, la DM2 fue la sexta causa principal de muerte en la región, con aproximadamente 244,084 muertes atribuidas directamente a esta enfermedad (OPS, 2023b).

2.3.2. Fisiopatología y patogénesis de la diabetes mellitus

El páncreas es un órgano complejo que desempeña dos roles principales: por un lado, sus células exocrinas participan en la digestión, mientras que, por otro, sus células endocrinas regulan la homeostasis de la glucosa en el cuerpo. Dentro del páncreas, los islotes de Langerhans albergan células β especializadas en la producción de insulina, un componente que regula los niveles de glucosa en la sangre. (Docherty & Russ, 2019). Después de una ingesta de alimentos, cuando se elevan los niveles de glucosa, las células β pancreáticas responden liberando insulina en la sangre. La insulina permite la utilización de la glucosa como fuente de energía en las células. Además, promueve el almacenamiento de glucosa en forma de glucógeno en el hígado y los músculos (Gillespie, 2006).

En la diabetes tipo 1, las células β productoras de insulina en el páncreas son atacadas y destruidas por el sistema inmunológico. Este tipo de enfermedades son denominadas enfermedades autoinmunes, en las cuales el sistema inmunitario responde de manera anómala a antígenos propios, lo que resulta en daño celular y tisular (Erf, 2014). Del mismo modo que otras enfermedades autoinmunes, la diabetes tipo 1 está asociada con el **antígeno leucocitario humano (HLA, por sus siglas en inglés 'human leukocyte antigen')**. El HLA es un conjunto de genes que codifican proteínas presentes en la superficie de las células del cuerpo, permitiendo que el sistema inmunológico las reconozca como propias y no las ataque. Se ha demostrado que el HLA localizado en el cromosoma 6 humano está asociado con la diabetes tipo 1 y se estima que contribuye a la mitad de la predisposición genética a esta enfermedad. (Speight & Pouwer, 2023). Los haplotipos del gen HLA DR4-DQ8 y DR3-DQ2 son especialmente importantes, ya que están presentes en el 90% de los casos de diabetes tipo 1 en niños. Estos alelos pueden predisponer a un individuo a una respuesta autoinmune contra las células β pancreáticas, provocando la DM1 (Devendra & Eisenbarth, 2003). Asimismo, en las personas en riesgo, se cree que los virus, los factores ambientales (incluidos los dietéticos) y otros factores estresantes

pueden desencadenar la destrucción autoinmune de las células β . Sin embargo, aunque se han identificado varios factores de riesgo y se comprenden algunos de los mecanismos, las causas de la diabetes tipo 1 aún no se conocen completamente (Krischer et al., 2022).

En la diabetes tipo 2 los factores son más complejos. Su desarrollo se debe a una combinación de factores genéticos y ambientales, y se caracteriza principalmente por la resistencia a la insulina y la disfunción de las células beta del páncreas. Esta resistencia ocurre cuando las células del cuerpo no responden adecuadamente a la insulina, Este hecho conduce a que la glucosa se acumule en el torrente sanguíneo en lugar de emplearse como energía. Inicialmente, la resistencia a la insulina es compensada con un aumento en la secreción de insulina, lo que permite mantener los niveles de glucosa dentro de un rango normal. Sin embargo, a medida que la enfermedad progresa, las células beta del páncreas comienzan a fallar y la secreción de insulina ya no es suficiente para mantener la homeostasis de la glucosa, resultando en hiperglucemia (Henson et al., 2023). La resistencia a la insulina puede causar estrés en ciertos orgánulos de las células beta, como el retículo endoplásmico, debido a la sobreproducción de insulina. Este estrés activa la respuesta de proteínas desplegadas (UPR, por sus siglas en inglés “Unfolded Protein Responses”), el cual es un mecanismo que intenta restaurar la función normal del RE. Sin embargo, si el estrés persiste, la UPR puede inducir la apoptosis en las células beta (de Boer et al., 2024).

Diversos factores pueden contribuir al desarrollo de la diabetes mellitus tipo 2. Entre estos se encuentran los altos niveles de ácido úrico sérico, la calidad y cantidad del sueño, el tabaquismo, la depresión, las enfermedades cardiovasculares, la dislipidemia, la hipertensión, el envejecimiento, el origen étnico, los antecedentes familiares de diabetes, la inactividad física y la obesidad (Ismail et al., 2021).

2.4. Relación Alzheimer-Diabetes Mellitus

En las últimas décadas, numerosos estudios han identificado vínculos comunes entre la enfermedad de Alzheimer y la diabetes mellitus tipo 2. La mayoría de estos estudios utilizan

pruebas estandarizadas, como el Mini-Mental State Examination (MMSE), para evaluar el rendimiento cognitivo tanto en pacientes diabéticos como en pacientes no diabéticos. Estos estudios comparan los resultados de las pruebas para observar las diferencias en el desempeño cognitivo entre ambos grupos. Los resultados de estos estudios han revelado vínculos significativos entre la EA y la DM2. Incluso, se ha propuesto el término "diabetes tipo 3" para describir la forma específica en que la diabetes está relacionada con la EA (Kandimalla et al., 2017).

Desde finales del siglo pasado, se han llevado a cabo investigaciones sobre el estado cognitivo de pacientes con diabetes mellitus tipo 2. Strachan y sus colaboradores realizaron una revisión de la literatura sobre este tipo de estudios. Se identificaron 19 estudios de este tipo y en 13 de estos estudios, se demostró que los individuos diabéticos tenían un desempeño más pobre en al menos un aspecto de la función cognitiva, siendo la memoria verbal la capacidad más afectada (Strachan et al., 1997). En otro estudio de revisión de la literatura, se identificaron un total de 33 estudios y se procedió a calcular el tamaño del efecto utilizando el d de Cohen. Se comparó el grupo de pacientes con diabetes mellitus con un grupo de control de sujetos no diabéticos, y se encontró que el grupo diabético mostró un rendimiento significativamente menor en varias áreas cognitivas: inteligencia ($d = -0.7$), velocidad de procesamiento de la información ($d = -0.3$), eficiencia psicomotora ($d = -0.6$), atención visual ($d = -0.4$) y atención sostenida ($d = -0.3$), flexibilidad cognitiva ($d = -0.5$) y percepción visual ($d = -0.4$). Por lo tanto, en pacientes con diabetes mellitus, se observa una disminución en la velocidad mental y la flexibilidad cognitiva, con efectos que pueden variar de leves a moderados (Brands et al., 2005). En otra investigación, se buscaron informes publicados de estudios longitudinales que describieran la evaluación de la función cognitiva en pacientes con diabetes mellitus. Se identificaron 25 estudios en los que se aplicaron las pruebas Mini-Mental State Examination y Digit Symbol Substitution Test. Los resultados indicaron que los individuos con diabetes mellitus experimentaron un deterioro de la función cognitiva entre 1.2 y 1.5 veces mayor que aquellos sin diabetes. Por lo tanto, se concluye que, en comparación con las personas sin diabetes, las personas con diabetes tienen una mayor tasa de deterioro de la función cognitiva y un mayor riesgo de desarrollar deterioro cognitivo (Cukierman et al., 2005). En 2009, Kopf y Frölich realizaron una revisión sistemática de la literatura basada en estudios prospectivos que

examinaron el riesgo de desarrollar enfermedad de Alzheimer en pacientes diabéticos. Identificaron 14 estudios de este tipo. Todos los estudios informaron cocientes de riesgo mayores de uno (mediana de 1.59, rango de 1.15 a 2.7). Además, en cuatro de estos estudios, este exceso de riesgo fue estadísticamente significativo (mediana de 1.73, rango de 1.59 a 1.9). Por lo tanto, es probable que la diabetes mellitus aumente el riesgo de padecer la enfermedad de Alzheimer (Kopf & Frölich, 2009). Los resultados positivos sobre la asociación entre la diabetes mellitus y la enfermedad de Alzheimer llevaron a los investigadores a explorar los posibles vínculos fisiopatológicos que conectan ambas condiciones.

Uno de los vínculos fisiopatológicos que relacionan la EA con DM2 es la resistencia a la insulina con la β -amiloide. Esta implica una serie de interacciones entre la señalización de insulina y los depósitos de β -amiloide. La resistencia a la insulina y los niveles reducidos del factor de crecimiento similar a la insulina (IGF-1, por sus siglas en inglés “Insuline Grow Factor”) pueden contribuir a la acumulación de β -amiloide en el cerebro. Esta acumulación, a su vez, puede antagonizar la unión de la insulina y el receptor IGF-1, desencadenando una cascada de eventos que incluyen la secreción de agentes inflamatorios y la aparición de resistencia a la insulina. Este ciclo de retroalimentación patológica conduce tanto a la resistencia a la insulina como la neurotoxicidad asociada a la β -amiloide (Mullins et al., 2017). Del mismo modo, la desregulación de la señalización de la insulina puede afectar la función de la proteína precursora amiloide, lo que conduce a la acumulación de β -amiloide en las células. (Tumminia et al., 2018). Además, se ha demostrado que la insulina y el receptor de insulina (RI) desempeñan un papel crucial en el aprendizaje asociativo. De hecho, la interrupción de la producción de insulina y de la actividad del RI provoca déficits en el aprendizaje asociativo. En pacientes con EA se han observado niveles anormales de insulina y del RI, y varios estudios han demostrado que la administración de insulina en estos pacientes mejora significativamente su rendimiento cognitivo. (Zhao & Alkon, 2001). Por otro lado, la enzima degradadora de insulina (IDE, por sus siglas en inglés “Insulin Degrading Enzyme”) es la principal peptidasa responsable de la degradación tanto de la insulina como de la β -amiloide. El papel fundamental de la IDE en la resistencia a la insulina está relacionado con su capacidad para regular los niveles de insulina (Kurochkin et al., 2018). La IDE degrada la insulina, lo que ayuda a mantener su equilibrio en el cuerpo. En situaciones de resistencia a la insulina, los niveles

elevados de insulina pueden sobrecargar la capacidad de la IDE para degradarla, lo que conduce a un exceso de insulina circulante. Además, la IDE también degrada la β -amiloide, una proteína que se acumula en el cerebro de los pacientes con Alzheimer. Cuando la IDE está ocupada degradando cantidades excesivas de insulina, su capacidad para descomponer la β -amiloide puede verse comprometida (De Felice, 2013). Finalmente, estudios en ratones han demostrado que la resistencia a la insulina incrementa la actividad de la γ -secretasa, lo que resulta en un aumento de la producción de β -amiloide en el cerebro. Además, se observa una disminución en la actividad de la enzima degradadora de insulina, lo que contribuye a la acumulación de β -amiloide (Starks et al., 2015).

Se han identificado vínculos moleculares entre la resistencia a la insulina y la proteína tau. La proteína tau no solo se expresa en el cerebro, sino también en otros órganos, como el endotelio pulmonar. Además, estudios han demostrado que tanto la proteína tau como la β -amiloide se expresan en el páncreas. Un análisis del tejido pancreático realizado en 21 casos de autopsias de pacientes con DM2 confirmó la presencia de estas proteínas (Balczon et al., 2024). Una investigación reveló la importancia de la tau pancreática en la regulación de la secreción de insulina y la homeostasis de la glucosa. Se observó que los niveles de tau estaban elevados en las células de los islotes β de pacientes con DM2, y la pérdida de tau resultó en un aumento en la secreción de insulina en líneas celulares, drosófila y ratones (Balczon et al., 2024). La tasa de agregación de la proteína tau está fuertemente influenciada por su proceso de fosforilación. De acuerdo con investigaciones realizadas en humanos y animales, se ha observado que la enzima fosfatasa PP2A juega un papel crucial en este proceso. Esta enzima está implicada tanto en la fosforilación de la tau en enfermedades como la EA como en la DM2. Además, se ha descubierto que la expresión de la fosfatasa PP2A disminuye cuando se administra insulina al paciente, lo que sugiere una conexión entre la regulación de la glucosa y la fosforilación de la tau (Clodfelder-Miller et al., 2006).

Se cree que la neuroinflamación desempeña un papel crucial como nexo entre la EA y la DM2. Esta inflamación surge en las etapas iniciales de la EA y contribuye a su patología mediante el daño oxidativo, la hiperfosforilación de la proteína tau, la acumulación de β -amiloide y la disfunción del sistema colinérgico. (Kandimalla et al., 2017). La neuroinflamación

es reconocida por numerosos estudios como una de las principales causas de la resistencia a la insulina y al IGF-1 observadas en el cerebro de pacientes con EA. La resistencia a la insulina inducida por DM2 y la obesidad genera altos niveles de macromoléculas lipídicas citotóxicas, las cuales, al atravesar la barrera hematoencefálica, desencadenan inflamación y resistencia a la insulina en el cerebro. Un proceso fundamental en esta neuroinflamación es la liberación de diversas células microgliales por el β -amiloide, lo que conduce a la secreción de citoquinas inflamatorias como la interleucina-6. Se destaca que ambos fenómenos median la relación entre la DM2 y la EA. (de la Monte, 2009). Mientras que por otro lado se ha demostrado que estas moléculas también provocan una disminución en la función sináptica y otros efectos, lo que respalda la teoría de que la DM2 posiblemente actúa como precursora al activar sustancias inflamatorias que conducen a la neuroinflamación y, en última instancia, a la EA (Rosenberg, 2005).

Otro proceso que establece un vínculo entre la EA y la DM2 es el estrés oxidativo. Esta se refiere a un desequilibrio entre la producción de especies reactivas de oxígeno y la defensa antioxidante. (Sies, 2019). Se ha demostrado que, en condiciones de resistencia a la insulina, se observa un aumento de las reacciones oxidativas en el plasma sanguíneo (Henriksen et al., 2011). Esto es perjudicial para el cerebro humano, ya que este es particularmente vulnerable al daño por el estrés oxidativo (Cobley et al., 2018). Asimismo, algunos estudios han propuesto que el estrés oxidativo y la resistencia a la insulina pueden conducir a la acumulación de la proteína β -amiloide y tau, lo que implica que haya una fuerte correlación con la patología de la EA (Gu et al., 2008).

2.5. Diseño de fármacos

2.5.1. El diseño de fármacos

El diseño de fármacos es el proceso de crear moléculas que tengan propiedades específicas para interactuar con objetivos biológicos, como proteínas o enzimas, con el fin de modular su actividad y producir efectos terapéuticos deseados. El proceso de descubrimiento y desarrollo de fármacos está compuesto de varias etapas. Estas comprenden desde la identificación del

objetivo farmacológico hasta la optimización de compuestos líderes, pasando por investigaciones preclínicas y ensayos clínicos y finalizando con la puesta en el mercado del nuevo fármaco (Salazar & Gormley, 2017).

La etapa inicial del desarrollo de un fármaco, conocida como el descubrimiento de fármacos, se centra en la búsqueda de nuevos compuestos con propiedades químicas específicas destinadas al tratamiento de enfermedades (Carracedo-Reboredo et al., 2021). Este proceso inicia con una investigación sobre las causas de una enfermedad, en algunos casos conduciendo a la identificación de una o varias dianas moleculares asociadas con dicha enfermedad. Posteriormente, se procede a la identificación de compuestos activos que interactúan con la diana molecular y a la optimización de su actividad biológica. La evaluación de estos compuestos activos comprende diversos pasos experimentales, que incluyen ensayos en líneas celulares, estudios en modelos animales, y ensayos clínicos en humanos (FDA, 2019). Finalmente, los compuestos que pasan satisfactoriamente por todas las etapas son aprobadas por un agente regulatorio (Salazar & Gormley, 2017).

Dado el alto costo económico y de tiempo requeridos, junto con la baja eficacia intrínseca de los métodos convencionales, desde finales del siglo pasado se han desarrollado herramientas computacionales para el diseño de fármacos. A los enfoques que emplean estas herramientas computacionales se conocen hoy en día como diseño de fármacos asistido por computadora (CADD, por sus siglas en inglés “Computer, Aided Drug Design”). Estos métodos permiten un descubrimiento de fármacos más preciso, sino que también se destacan por su eficiencia mejorada, constituyendo así una alternativa más efectiva en el proceso de desarrollo de fármacos (Coumar, 2021).

El diseño de fármacos asistido por computadora abarca una diversidad de enfoques teóricos y computacionales que constituyen elementos esenciales en el proceso contemporáneo de descubrimiento de fármacos (Prieto-Martínez et al., 2019). El CADD puede ser dividido en dos enfoques de acuerdo con la disponibilidad de la estructura de la proteína diana: **diseño de fármacos basado en la estructura** (SBDD, por sus siglas en inglés “Structure Based Drug

Design”) y diseño de fármacos basado en el ligando (LBDD, por sus siglas en inglés “Ligand Based Drug Design”) (Coumar, 2021).

2.5.2. La inteligencia artificial y el diseño de fármacos

En la última década, la **inteligencia artificial (IA)** ha experimentado un significativo avance que la ha colocado como el centro de atención de la sociedad. Su rápido desarrollo ha despertado el interés en diversos sectores, convirtiéndola en un tema común de conversación a nivel global y en un nuevo recurso cotidiano para muchas personas. La IA nace en los años 1950 como un “esfuerzo de automatizar tareas normalmente realizadas por humanos” (Chollet, 2018). En la actualidad, la IA es definida como el estudio y la creación de máquinas inteligentes que se asemejan a los humanos, o incluso que superen su inteligencia (Singh & Khanna, 2023). El campo de la IA abarca el **aprendizaje automático (ML, por sus siglas en inglés “Machine Learning”)** y el Machine Learning a su vez tiene un subcampo llamado aprendizaje profundo **(DL, por sus siglas en inglés “Deep Learning”)** (Batoool et al., 2019).

Las tecnologías que integran el aprendizaje automático y la IA se han convertido en herramientas versátiles que pueden ser aplicadas de manera generalizada en diversas etapas del proceso de desarrollo de fármacos. Esto incluye la identificación y validación de objetivos farmacológicos, el diseño de nuevos medicamentos, la reutilización de fármacos existentes, la mejora de la eficiencia en la investigación y desarrollo, la recopilación y análisis de información biomédica, así como la optimización del proceso de toma de decisiones para la selección de pacientes en ensayos clínicos (Huang et al., 2017). Además, se han desarrollado métodos basados en la IA para predecir la estructura tridimensional de una proteína muy prometedores que mejoran la calidad estructural acercándose a la precisión experimental (Bordin et al., 2023). Algunos ejemplos de ello son el software AlphaFold (Jumper et al., 2021), RoseTTafold (Baek et al., 2021) y CollabFold (Mirdita et al., 2022).

2.5.3 Aprendizaje Automático

El aprendizaje automático representa un paradigma nuevo en comparación con la programación clásica. En esta última, se proporcionan una serie de reglas y datos a un programa, y las respuestas se generan conforme a esas instrucciones. En cambio, en el aprendizaje automático, se presentan tanto los datos como las respuestas al programa, con la expectativa de que aprenda "reglas" que luego puedan aplicarse a nuevos datos para obtener resultados (Chollet, 2018). El aprendizaje automático se puede categorizar de acuerdo con la inclusión o no de la variable dependiente (y) en los datos utilizados para entrenar el modelo. Cuando el modelo es entrenado con datos que contienen la variable dependiente (y), se le llama aprendizaje supervisado; mientras que, si carecen de ellos, se le conoce como aprendizaje no supervisado. Además, ambas categorías pueden subdividirse según la naturaleza de los datos, ya sean continuos o categóricos (Batoool et al., 2019). El aprendizaje supervisado se emplea principalmente para tareas de regresión y clasificación, permitiendo la predicción de propiedades farmacocinéticas y toxicidad (Pillai et al., 2022), mientras que el aprendizaje no supervisado, en el contexto del cribado de fármacos, se puede emplear el aprendizaje no supervisado para analizar la similitud de los compuestos a través de métodos de agrupamiento, lo que ayuda a agrupar los compuestos según sus similitudes (Voicu et al., 2020).

2.5.4. Aprendizaje supervisado

En el aprendizaje supervisado, los datos deben ir acompañados de su respectivo resultado y. Un modelo de aprendizaje supervisado tiene como objetivo determinar una función $f(x) = y$ a partir de una lista de pares de entrenamiento $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ (Raghunathan & Priyakumar, 2022). El aprendizaje supervisado se subdivide según la naturaleza de los datos que procesa, ya sea continua o categórica. Para datos continuos, se emplean algoritmos de regresión, mientras que, para datos categóricos, se recurre a algoritmos de clasificación. Aunque estas son las clasificaciones principales, hay otros algoritmos que no se ajustan solo a una de estas categorías y pueden aplicarse para ambos tipos de datos. Los modelos de aprendizaje automático más comúnmente utilizados en el descubrimiento de nuevos fármacos, tanto para clasificación como para regresión, incluyen la regresión lineal, la regresión logística, los árboles de decisión (DT, por sus siglas en inglés "Decision Trees"), los bosques aleatorios (conocidos en inglés como "random forests"), las máquinas de soporte vectorial (SVM, por sus siglas en

inglés, “Support Vector Machines”), el algoritmo de vecinos más cercanos (k-NN, por sus siglas en inglés “k-Nearest Neighbors”), XGBoost y el clasificador Naive Bayes (Schneider et al., 2019).

El modelo de regresión es un algoritmo empleado para comprender la relación entre una variable dependiente (denotada como Y) y una o más variables independientes (denotadas como X). El objetivo principal es modelar esta relación ajustando una ecuación lineal a los datos observados (Shobha & Rangaswamy, 2018). La forma básica del modelo de regresión lineal es:

De este mismo modo, la regresión logística es otro modelo más que puede ser aplicado en el descubrimiento de nuevos fármacos, especialmente para problemas de clasificación. A diferencia de la regresión lineal, que se usa para predecir valores continuos, la regresión logística se emplea para predecir una variable categórica, generalmente binaria, es decir, 0 y 1 (Edgar & Manz, 2017).

Las SVM son algoritmos de aprendizaje supervisado diseñados para encontrar el hiperplano óptimo que mejor separa las clases dentro de un conjunto de datos. Este hiperplano es aquel que maximiza el margen entre las clases. Los puntos de datos más cercanos a este hiperplano de separación se conocen como vectores de soporte, y juegan un papel crucial en la definición de este. Gracias a su estructura, las SVM pueden emplearse tanto para tareas de clasificación como de regresión (Djuris et al., 2013). Los árboles de decisión siguen una estructura análoga a la de los árboles de decisión convencionales, donde diferentes condiciones conducen a una respuesta. En su configuración básica, constan de un nodo raíz, múltiples nodos internos y varias hojas. El nodo raíz y los nodos internos representan condiciones de prueba para entidades o atributos, que son utilizadas para clasificar registros con diferentes características. Por otro lado, las hojas del árbol reflejan la decisión final alcanzada tras seguir las condiciones establecidas en los nodos internos y raíz. Esta estructura jerárquica permite una interpretación intuitiva del proceso de toma de decisiones del modelo (Shang et al., 2022). Los bosques aleatorios siguen una metodología similar a la de los árboles de decisión. En este algoritmo, se generan múltiples árboles de decisión a partir de conjuntos de datos. Cada árbol genera predicciones independientes y, posteriormente, se realiza una votación para seleccionar la mejor predicción entre ellos. Este enfoque combinado y de votación permite obtener una predicción final robusta

y precisa (Chaitanya Kumar et al., 2023). Los algoritmos de XGBoost emplean modelos considerados "débiles" y los mejora significativamente, aumentando su rendimiento y precisión. Los algoritmos XGBoost están basados en árboles de decisión. Estos algoritmos construyen de manera iterativa nuevos árboles con el objetivo de corregir los errores del árbol anterior. Este proceso secuencial es fundamental para su eficacia. Para lograr esto, XGBoost implementa una optimización mediante el cálculo del gradiente de la función de pérdida. Esto permite ajustar los parámetros del modelo de manera eficiente, mejorando continuamente su precisión (T. Chen & Guestrin, 2016a). El algoritmo Naive Bayes también es ampliamente utilizado en el descubrimiento de fármacos y es uno de los algoritmos de clasificación más simples. Se basa en el teorema de Bayes, que asume la independencia de las características dadas las clases. A pesar de esta suposición, el algoritmo funciona sorprendentemente bien en diversos contextos, especialmente en la clasificación de textos y sistemas de recomendación. (H. Chen et al., 2021). Finalmente, otro algoritmo ampliamente utilizado para el descubrimiento de nuevos fármacos es el de vecinos más cercanos. Este es un método de aprendizaje supervisado utilizado tanto para clasificación como para regresión. En lugar de emplear una ecuación matemática específica, K-NN almacena los valores y sus etiquetas correspondientes. En tareas de clasificación, K-NN encuentra los "k" puntos más cercanos y asigna al nuevo punto la clase más común entre ellos. En tareas de regresión, se utiliza el promedio de los valores de sus "k" vecinos más cercanos para realizar la predicción (Subasi, 2020).

2.5.5. Aprendizaje no supervisado

Los enfoques de aprendizaje no supervisado buscan identificar patrones en conjuntos de datos que carecen de etiquetas y variables de salida (Raghunathan & Priyakumar, 2022). Al igual que en el aprendizaje supervisado, este tipo de enfoque puede clasificarse según la naturaleza de los datos: categóricos o continuos (Schneider & Xhafa, 2022). Los algoritmos de aprendizaje no supervisado más utilizados en el descubrimiento de fármacos incluyen K-means clustering, clustering jerárquico, DBSCAN (por sus siglas en inglés "Density-Based Spatial Clustering of Applications with Noise"), análisis de componentes principales (PCA, por sus siglas en inglés "Principal Component Analysis") y t-SNE (por sus siglas en inglés "Stochastic Neighbor Embedding"). Estos algoritmos permiten reducir la dimensionalidad de los datos,

identificar patrones y seleccionar candidatos prometedores para pruebas adicionales. Aplicaciones como estas mejoran la eficiencia en la construcción de modelos QSAR y en la identificación de nuevas moléculas con potencial terapéutico.

2.5.6 Aprendizaje profundo

El aprendizaje profundo constituye un subcampo del aprendizaje automático que transforma el paradigma de adquisición de conocimientos. Mientras que en el aprendizaje automático se busca obtener una representación de los datos, el aprendizaje profundo va más allá. En este nuevo paradigma, se obtienen capas sucesivas de representaciones para los datos. La profundidad en el aprendizaje profundo no implica una comprensión más exhaustiva de los "datos", sino que se refiere a la presencia de múltiples capas "profundas" de representaciones. Por ende, una denominación más apropiada para el aprendizaje profundo sería "aprendizaje de representaciones en capas" (Chollet, 2018).

Los algoritmos de aprendizaje profundo emplean redes neuronales artificiales para lograr su cometido. Las redes neuronales artificiales (ANN, por sus siglas en inglés “Artificial Neural Network”) son algoritmos de aprendizaje que reciben múltiples señales de entrada, procesan la información calculando una suma ponderada de estas entradas mediante una función de activación no lineal, y luego transmiten la señal resultante a las neuronas conectadas en la siguiente capa hasta finalmente dar una respuesta (van Gerven & Bohte, 2017). Si bien, se les llama redes neuronales, no hay prueba alguna de que el cerebro funcione de esta manera (Chollet, 2018).

Las redes neuronales se categorizan según su arquitectura y las conexiones internas que poseen. Entre los ejemplos de estas redes se encuentran las redes neuronales recurrentes y las redes neuronales convolucionales (Paul et al., 2021). Las redes neuronales recurrentes son una red neuronal artificial especializada diseñada específicamente para procesar datos de series temporales o información que incluye secuencias de datos. Estas incorporan el concepto de "memoria", lo que les permite retener estados o información de las entradas previas para construir la salida subsiguiente de la secuencia. Las redes neuronales convolucionales se

emplean para analizar imágenes. Utiliza una técnica llamada convolución, la cual es una operación matemática entre dos funciones que produce una tercera función, describiendo cómo la forma de una afecta la forma de la otra. (Batoool et al., 2019).

2.6. Diseño de fármacos para la enfermedad de Alzheimer

Con los avances tecnológicos de los últimos años, ha aumentado la cantidad de datos farmacéuticos disponibles para ser utilizados por la IA. La aplicación del aprendizaje automático en el diseño de fármacos se está generalizando rápidamente en la última década. El aprendizaje automático permite extraer características estructurales moleculares, analizar interacciones fármaco-diana y establecer vínculos entre fármacos, enfermedades y proteínas. Este enfoque facilita el desarrollo de sistemas de predicción altamente precisos en diversas etapas del proceso, minimizando la incertidumbre en cada fase de la investigación (Geng et al., 2024). Gracias a los avances en los algoritmos de IA, las soluciones impulsadas por estos tienen un potencial sin precedentes para acelerar el proceso de descubrimiento de fármacos, mejorando significativamente la tasa de éxito (Vatansever et al., 2021), siendo este uno de los principales problemas en el desarrollo de fármacos para EA, pues alta tasa de deserción de los fármacos en desarrollo. Durante el período 2002-2012, se encontró una tasa de éxito general de solo el 0.4%. Esta alta tasa de deserción es una de las razones más importantes para enfocar el desarrollo de fármacos en este ámbito (Cummings et al., 2014).

El principal reto en el desarrollo de nuevos fármacos radica en caracterizar las interacciones fármaco-diana y luego aplicar estas caracterizaciones para evaluar la probabilidad de interacción entre una molécula de fármaco y una diana específica (Ye et al., 2021). Las investigaciones actuales en el diseño de fármacos para la EA abordan diversos aspectos, como la cantidad y tipo de dianas terapéuticas, con estudios que se enfocan tanto en dianas únicas como múltiples. Se utilizan diferentes metodologías para identificar nuevos inhibidores potenciales, siendo los modelos QSAR (Relación Cuantitativa Estructura-Actividad, por sus siglas en inglés “Quantitative Structure-Activity Relationship”) una herramienta común (Roy et al., 2015). Estos modelos mapean la relación entre la estructura de una molécula y su actividad biológica o reactividad química. Algunos investigadores utilizan aprendizaje

automático para desarrollar QSAR, mientras que otros emplean métodos convencionales, usando descriptores moleculares o resultados de acoplamiento molecular. Las dianas terapéuticas comunes en las investigaciones de descubrimiento de fármacos contemporáneos para la EA incluyen principalmente inhibidores de la acetilcolinesterasa, que aumentan los niveles de acetilcolina en el cerebro, mejorando la transmisión sináptica. También se investigan inhibidores dirigidos contra la beta-amiloide y antagonistas del receptor **N-metil-D-aspartato (NMDA)**, que modulan la actividad del glutamato al antagonizar estos receptores, aliviando así los síntomas de la enfermedad (Parvathy Dharshini et al., 2022).

Existen diversos estudios enfocados en descubrir nuevos fármacos para el Alzheimer EA centrados en una sola diana terapéutica. En 2022, Mukerjee y su equipo realizaron un estudio para identificar posibles fármacos dirigidos contra Keap1, una proteína implicada en el estrés oxidativo relacionado con la EA. Entrenaron varios modelos de inteligencia artificial utilizando datos de compuestos como el ácido ursólico y compuestos similares, descargando los valores de **IC₅₀**, que miden la concentración necesaria para inhibir un proceso biológico en un 50% (Ugwu & Conradie, 2023). Emplearon descriptores moleculares y seleccionaron el mejor modelo basándose en R² y RMSE, resultando el modelo de bosques aleatorios como el más eficaz. Con este modelo, calcularon el IC₅₀ de nuevas moléculas, identificando cinco compuestos prometedores. Posteriormente, realizaron acoplamiento molecular con Keap1 y luego realizaron análisis de dinámica molecular obteniendo resultados positivos y determinando una nueva molécula prometedora contra la EA (Mukerjee et al., 2022). En el año 2022, Lee y colaboradores se centraron en desarrollar varios modelos de machine learning para identificar inhibidores de S100A9, una proteína relacionada con la enfermedad de Alzheimer debido a su papel en la inflamación y el estrés oxidativo. Obtuvieron datos de bases de datos públicas y calcularon descriptores moleculares 2D. Para las huellas dactilares, consideraron cada bit como un descriptor molecular independiente. Entrenaron y optimizaron modelos de bosque aleatorio mediante selección de características, seleccionando los mejores ocho modelos. A través de una votación consenso de estos modelos, identificaron hasta 46 posibles inhibidores de S100A9 (Lee et al., 2019). En ese mismo año, Cavas y sus colaboradores emplearon redes neuronales artificiales para predecir la actividad inhibitoria contra la acetilcolinesterasa. Estas redes neuronales fueron entrenadas utilizando los resultados experimentales de inhibición de

derivados de oxazolona in vitro (Cavas et al., 2019). Sin embargo, debido a la naturaleza multifactorial del Alzheimer (EA), los tratamientos con fármacos dirigidos a una sola diana pueden no ser exitosos. En 2015, Lovestone y colaboradores realizaron un ensayo de fase II para evaluar la efectividad de tideglusib en EA. Este fármaco es un inhibidor de la glucógeno sintasa quinasa-3. Se administró tideglusib o un placebo a grupos de pacientes con EA, evaluando criterios como la atrofia cerebral en RM y los niveles de tau, amiloide- β y BACE1 en el líquido cefalorraquídeo. Los resultados no mostraron diferencias significativas entre los grupos, sugiriendo que centrarse en una sola diana terapéutica podría haber sido una limitación (Lovestone et al., 2015).

Debido a la complejidad y los múltiples mecanismos implicados en la EA, es crucial descubrir fármacos multidiana capaces de abordar las diversas vías involucradas en su desarrollo (Vatansever et al., 2021). Un fármaco multidiana es aquél que puede actuar en el tratamiento de enfermedades complejas mediante la modulación de múltiples dianas (Lavecchia & Cerchia, 2016). Un estudio señala que debe enfatizarse que el desarrollo y la aparición de la EA es el resultado de los efectos colectivos de múltiples factores, no de un único factor específico. Este hecho justifica la necesidad de dirigir los tratamientos hacia múltiples mecanismos y vías simultáneamente. Probablemente, esta es la razón por la cual algunos de los fármacos desarrollados basándose en una sola diana terapéutica, ha tenido éxito (Gong et al., 2018). El desarrollo de fármacos multidiana ha ganado mucha aceptación en las últimas décadas y ha sido aplicado principalmente en el cáncer. En el año 2019, se desarrollaron fármacos multidiana para tratar el cáncer de pulmón utilizando un enfoque de farmacología de sistemas que incluyó cribado fenotípico, química, fosfoproteómica y secuenciación de ARN. Se seleccionaron los inhibidores de multiquinasa foretinib y cabozantinib, identificando sus dianas moleculares y mecanismos de acción. La validación mediante estudios bioquímicos y celulares confirmó estas dianas. Finalmente, desarrollaron una combinación sinérgica de foretinib con barasertib, produciendo un inhibidor más potente de AURKB, para el cáncer de pulmón de células pequeñas amplificado por MYC (Sumi et al., 2019).

Se han realizado múltiples estudios para descubrir inhibidores multidiana que puedan tratar la enfermedad de Alzheimer. Algunos emplean técnicas de aprendizaje automático, mientras

que otros utilizan enfoques más convencionales. En 2023, Kumar y sus colaboradores desarrollaron modelos QSAR multiobjetivo para identificar nuevos inhibidores contra la EA. Trabajaron con varias dianas terapéuticas, incluyendo AChE, BuChE, BACE1, β -amiloide, 5-HT6, CDK-5, gamma-secretasa, Glutaminil Cicloasa, GSK-3 β , MAO-B, NMDA y PDE10A, obteniendo datos de actividad molecular contra estas dianas. Desarrollaron modelos QSAR para cada diana utilizando regresión por mínimos cuadrados parciales y regresión lineal múltiple. Posteriormente, los modelos se utilizaron para predecir la actividad inhibitoria de aproximadamente 19 millones de compuestos obtenidos de bases de datos químicas como ZINC12, Asinex, NCI y InterBioscreen. Con los modelos desarrollados, identificaron varios compuestos con actividad dual inhibitoria, incluyendo combinaciones como AChE y BACE1, AChE y β -amyloid, y AChE y BuChE, entre otros (Kumar et al., 2023). En el año 2020, Concu y sus colaboradores desarrollaron un modelo multidiana para predecir la actividad de los fármacos monooxidasa A y B. La actividad de las enzimas monooxidasa están relacionadas la EA. Para llevar a cabo el modelo, crearon un modelo QSAR basado en análisis discriminante lineal que se desarrolló en base a 5759 compuestos de una base de datos. El modelo resultante fue capaz de clasificar correctamente a todos esos compuestos por lo que se considera un modelo robusto y reproducible (Concu et al., 2020).

Aunque en los estudios previos se crearon modelos QSAR tradicionales, los modelos QSAR basados en aprendizaje automático se consideran superiores debido a su mayor precisión y robustez. Un estudio demostró que los modelos de aprendizaje automático son más precisos y robustos en comparación con los modelos QSAR tradicionales, lo que los hace más efectivos para predecir la actividad biológica de compuestos en el descubrimiento de fármacos (Bosc et al., 2019). Dhamodharan y Mohan desarrollaron modelos de aprendizaje automático para predecir la actividad de inhibidores duales de AChE y BACE1 para el tratamiento de la enfermedad de Alzheimer (EA). Utilizaron técnicas de machine learning, incluyendo máquinas de vectores de soporte y redes neuronales, para cada proteína diana. Los modelos mostraron alta precisión predictiva en validaciones internas y externas. Además, identificaron descriptores clave para la actividad inhibidora de AChE y BACE1, esenciales para el diseño racional de inhibidores multitarget (Dhamodharan & Mohan, 2022). En 2023, se llevó a cabo un estudio para el diseño de fragmentos químicos con el fin de descubrir inhibidores duales de AChE y

BACE1. Utilizaron datos de 3,524 compuestos con mediciones de AChE y BACE1 obtenidos de la base de datos ChEMBL para entrenar árboles de decisión y construir modelos QSAR. Se obtuvieron buenos rendimientos para los modelos construidos. A partir de estos árboles, se identificaron reglas para descubrir inhibidores duales, resultando en la predicción de 250 nuevos inhibidores basados en fragmentos activos (Bao et al., 2023).

Actualmente, no se han realizado investigaciones sobre el descubrimiento de nuevos fármacos para la EA. La mayoría de estas investigaciones se han centrado en evaluar los efectos de los fármacos antidiabéticos en la EA, aprovechando la interrelación entre ambas enfermedades para identificar tratamientos potenciales. En una revisión sobre la aplicación de fármacos antidiabéticos en el tratamiento de la EA realizada por Michailidis, se determinó que los investigadores están probando medicamentos como la insulina intranasal, la metformina, las incretinas y las tiazolidinedionas. Varios estudios han concluido que estos medicamentos pueden mejorar la cognición y la memoria en pacientes con EA (Michailidis et al., 2022). Una investigación reciente exploró los vínculos moleculares entre la EA y la DM2 utilizando redes de interacción proteína-proteína. En este estudio, se identificaron diez proteínas clave que podrían servir como biomarcadores y objetivos terapéuticos para ambas enfermedades. Estas proteínas destacan por su potencial en el desarrollo de tratamientos combinados, ya que abordan los mecanismos subyacentes compartidos entre EA y DM2. Este enfoque podría conducir a terapias más efectivas y específicas, mejorando los resultados clínicos para pacientes con ambas condiciones. Las proteínas propuestas fueron STAT3, EGFR, IRS1, MAPK1, SRC, HSP90AA1, PIK3R1, UBC, MAPK3 y ESR1 (Castillo-Velázquez et al., 2023). En la presente investigación, se realizará.

2.7. Proteínas objetivo

PROTEÍNA	Descripción	RELACIÓN CON LA EA	RELACIÓN CON LA DM2
STAT3	Es una proteína de la familia de las STAT involucrada en la señalización celular de procesos inflamatorios y de crecimiento celular. Actúa como un factor de transcripción que, en respuesta a diversas citoquinas y factores de crecimiento, se transloca al núcleo de la célula para activar la transcripción de genes específicos (H. Q. Wang et al., 2022).	Implicada precisamente debido a la neuroinflamación y la muerte neuronal inducida por beta-amiloide. Esta es activada en respuesta a la acumulación de A β y contribuye a la inflamación crónica en el cerebro. Un estudio reciente demostró que un inhibidor de STAT3 mitigaba la angiopatía amiloide cerebral y las placas amiloides mejorando las funciones cognitivas y las redes cerebrales en modelos de EA (Mehla et al., 2021).	También implicada en procesos de inflamación, sensibilización a la insulina y reprogramación celular. Una investigación realizada por Wakabayashi demostró que la inhibición de STAT3 puede facilitar la reprogramación de células alfa a células beta, lo cual es crucial dado que en la DM2 las células beta están parcial o completamente destruidas (WAKABAYASHI et al., 2019)
EGFR	Es un miembro de la familia de proteínas quinasas. Actúa como receptor para los factores de crecimiento epidérmico. Al unirse a su ligando, induce la dimerización del receptor y la autofosforilación en residuos de tirosina, lo que desencadena una cascada de señales intracelulares que conducen a la	Una revisión realizada por Choi y colaboradores demostró que los inhibidores de EGFR pueden atenuar la patología de la beta-amiloide y mejorar la función cognitiva en modelos de EA. Estos inhibidores actúan reduciendo la neuroinflamación y el estrés oxidativo (Choi et al., 2023).	En relación con la DM2, numerosos estudios han demostrado que la activación transitoria de la vía del receptor del factor de crecimiento epidérmico es esencial para promover la recuperación renal tras una lesión aguda. Sin embargo, su activación persistente está implicada en la progresión de diversas enfermedades renales crónicas, incluida la enfermedad renal diabética,

	proliferación celular (Gan et al., 2013)		por lo que dirigirse al EGFR también podría tener un potencial terapéutico para la enfermedad renal diabético (Sheng et al., 2020).
IRS1	Es una proteína involucrada en la señalización celular que actúa como un intermedio en la transducción de señales desde el receptor de insulina hasta diversas señales (NIH gene, 2024c).	Inhibir la IRS1 es arriesgado debido a que la disfunción de la IRS1 se ha asociado con la EA. Se ha demostrado que la resistencia a la insulina en el cerebro, provocada por inhibir la IRS1 puede promover la acumulación de beta-amiloide (Arnold et al., 2018).	Debido a que la IRS1 es fundamental en la vía de señalización del receptor de insulina, su disfunción puede causar resistencia a la insulina (Gabbouj et al., 2019).
SRC	Es una proteína perteneciente a las familias de las tirosino quinasas no receptoras. Tiene vínculos con procesos celulares como el crecimiento, la diferenciación y la supervivencia celular (Uniprot, 2024)	La proteína SRC se ha visto implicada en la fosforilación anormal de proteínas tau, que forman ovillos neurofibrilares en la EA. Esta hiperfosforilación es un evento clave en la patogénesis de la EA y la SRC podría contribuir a este evento (Arnold et al., 2018).	SRC está implicado en la modulación de la señalización de insulina. La activación anormal de SRC puede interferir con la transducción de señales de insulina, contribuyendo a la resistencia a la insulina, un factor clave en la patogénesis de la DM2 (Greeny et al., 2024).
MAPK1	Las proteínas MAPK (Mitogen-Activated Protein Kinases) son una familia de quinasas serina/treonina que juegan roles cruciales en la transducción de señales intracelulares. Estas proteínas están involucradas en una amplia variedad de procesos celulares, incluyendo el crecimiento, la diferenciación, la	Se ha demostrado que la señalización aberrante de ERK2 puede llevar a la hiperfosforilación de proteínas tau, resultando en la formación de ovillos neurofibrilares, una característica de la EA (Miao et al., 2019).	Un estudio demostró que la desregulación de la señalización de MAPK contribuye a la disfunción endotelial progresiva en la DM2 (Huang et al., 2012).

	respuesta al estrés, la apoptosis y la regulación de funciones metabólicas (Cargnello & Roux, 2011).		
PIK3R1	Es esencial para la regulación de la señalización celular a través de la vía PI3K/AKT. Sus funciones abarcan desde el control del crecimiento y la supervivencia celular hasta la regulación del metabolismo y la respuesta inmunológica. Las mutaciones en PIK3R1 están implicadas en diversas enfermedades, incluyendo el cáncer y la diabetes (NIH gene, 2024d)	Se ha demostrado que la disfunción de esta vía puede contribuir a la acumulación de beta-amiloide y la fosforilación de tau, dos características patológicas clave del Alzheimer. Se que esto puede afectar la función cognitiva y la memoria (Cho, 2024).	PIK3R1 es fundamental para la señalización de insulina, ya que forma parte de la vía PI3K/AKT. Esta vía es esencial para la regulación de la captación de glucosa en las células, y las mutaciones o la disfunción de PIK3R1 pueden llevar a la resistencia a la insulina, una característica principal de la DM2 (Boukhalifa et al., 2023)-
HSP90AA1	Es un gen que codifica una de las isoformas de la proteína Hsp90, una chaperona molecular esencial involucrada en el plegamiento, estabilidad y función de muchas proteínas clave en la célula. Hsp90 es especialmente importante en situaciones de estrés celular, como temperaturas elevadas y condiciones que causan desnaturalización de proteínas (NIH gene, 2024b).	HSP90AA1 ayuda a estabilizar y plegar correctamente proteínas como tau y beta-amiloide, que son esenciales en la patología de AD. HSP90AA1 impide la agregación patológica de tau, ayudando a mantener la homeostasis celular y la función neurona (Liu et al., 2024).	Esta proteína está involucrada en la regulación de la señalización de la insulina. Las chaperonas HSP90 son necesarias para la funcionalidad de la proteína quinasa, que es crucial en la vía de señalización de la insulina para la captación de glucosa. La disfunción de HSP90AA1 puede contribuir a la resistencia a la insulina, una característica principal de la DM2 (Ding et al., 2022).
ESR1	Es el gen que codifica el receptor de estrógeno alfa, una proteína nuclear	Un estudio demostró que polimorfismos en el gen ESR1, como las variantes PvuII	Un estudio encontró que estos polimorfismos están significativamente

	<p>que actúa como un factor de transcripción regulado por ligando. Esta proteína regula la expresión de muchos genes inducibles por estrógenos los cuales desempeñan un papel en el crecimiento, el metabolismo, el desarrollo sexual, la gestación y otras funciones reproductivas, y se expresa en muchos tejidos no reproductivos (NIH gene, 2024a).</p>	<p>y XbaI, están asociados con un mayor riesgo de desarrollar Alzheimer. Estos polimorfismos pueden influir en la expresión del receptor de estrógeno alfa, afectando la neuroprotección que los estrógenos proporcionan en el cerebro. La disfunción del ESR1 puede desencadenar neuroinflamación y otros mecanismos patológicos en la EA (Ma et al., 2009).</p>	<p>asociados con la resistencia a la insulina y la diabetes tipo 2 en mujeres palestinas. Los resultados mostraron que las variantes genéticas pueden influir en la susceptibilidad a la enfermedad, sugiriendo que ESR1 juega un papel en la patogénesis de la DM2 (Gallagher et al., 2007).</p>
<p>UBC</p>	<p>Este gen codifica la proteína ubiquitina C. Esta proteína juega un papel importante en la regulación de varios procesos celulares a través de la modificación postraduccional de otras proteínas. La ubiquitinación se ha asociado con la degradación de proteínas, la reparación del ADN, la regulación del ciclo celular, la modificación de la quinasa, la endocitosis y la regulación de otras vías de señalización celular (NIH gene, 2024e)</p>	<p>Si bien, no existen estudios donde se relacione de manera directa a la UBC y a la EA, se destaca que la UBC juega un rol fundamental puesta esta es una proteína central en una red grande de genes relacionados con la ubiquitinación, incluida la reparación del ADN y la regulación del ciclo celular, por lo que inhibir la UBC podría tener consecuencias catastróficas para el organismo (X. L. Wang & Li, 2021).</p>	<p>No existen estudios que relacionen de manera directa a la UBC con la DM2.</p>

3. JUSTIFICACIÓN

1. Importancia del problema.
2. Necesidad de hacer esta investigación.
3. Beneficios de la investigación.
4. Producto final de la propuesta

Existen varios medicamentos aprobados por la Administración de Alimentos y Medicamentos de los Estados Unidos de América para tratar la EA. Algunos de estos medicamentos son la galantamina, la rivastigmina y el donepezilo (Uddin et al., 2020). Sin embargo, estos medicamentos están orientados a disminuir los efectos de la EA en un determinado momento y tienen como efectos secundarios náuseas, vómitos, diarrea, reacciones alérgicas, pérdida de apetito, dolor de cabeza, confusión, mareo y caídas (NIH, 2023). Uno de los principales problemas del diseño de nuevos fármacos es su baja tasa de éxito, pues alta tasa de deserción de los fármacos. Se encontró una tasa de éxito general de solo el 0.4% (Cummings et al., 2014). Para abordar estos problemas, estudios han demostrado que se deberían de diseñar fármacos multidiana debido a la naturaleza multifactorial de la EA, así como que se ha asociado está a varias causas (Lovestone et al., 2015). Recientemente, se ha asociado como una de estas causas a la diabetes mellitus (Yang & Song, 2013). Y múltiples estudios han relacionado a la EA con la diabetes mellitus afectada (Brands et al., 2005; Kopf & Frölich, 2009; Strachan et al., 1997). Asimismo, se han probado diversos medicamentos destinados a la diabetes mellitus para pacientes de EA y estos han resultado en que pueden mejorar la cognición y la memoria en pacientes con EA (Michailidis et al., 2022).

4. HIPÓTESIS

El empleo de modelos QSAR combinados con técnicas de aprendizaje automático permitirá identificar potenciales inhibidores que actúen sobre múltiples blancos moleculares relevantes tanto para el Alzheimer como para la Diabetes Mellitus.

5. OBJETIVOS

5.1. Objetivo general

Identificar inhibidores multi-blanco con capacidad para modular múltiples blancos relacionados con la asociación entre la enfermedad de Alzheimer y la diabetes mellitus

5.2. Objetivos específicos

5.2.1. Recopilar y analizar datos experimentales de inhibición de moléculas dirigidas a las proteínas STAT3, EGFR, SRC, MAPK1 y ESR1.

5.2.2. Desarrollar modelos QSAR con aprendizaje automático para predecir la actividad inhibitoria de moléculas dirigidas a las proteínas STAT3, EGFR, SRC, MAPK1 y ESR1.

5.2.3. Identificar potenciales inhibidores aplicando los modelos QSAR desarrollados a una base de datos de compuestos naturales.

5.2.4. Realizar simulaciones de acoplamiento molecular para los inhibidores predichos por los modelos QSAR.

6. MATERIALES Y MÉTODOS

6.1. Recopilación y análisis de datos experimentales de inhibición

6.1.1. Obtención de los datos experimentales de los inhibidores

Se descargarán datos experimentales de inhibición y la información en formato SMILES de moléculas dirigidas a las proteínas STAT3, EGFR, SRC, MAPK1 y ESR1 utilizando la base de datos bioinformática de acceso libre ChEMBL (Davies et al., 2015). Para ello, se empleará el buscador de ChEMBL, introduciendo el nombre de cada proteína para recuperar todos los datos experimentales y la información de actividad inhibitoria correspondiente. Los datos se descargarán en el formato estándar IC50 y se almacenarán en un archivo CSV.

6.1.2. Tratamiento de datos

Se utilizará el lenguaje de programación de alto nivel Python 3.11.0 (Van & Fred, 2009) junto con la biblioteca de análisis de datos pandas (The pandas development team, 2020) para llevar a cabo el análisis y procesamiento de los datos de las moléculas recolectadas para todas las proteínas. Se empleará como entorno de desarrollo de programación el software Visual Studio Code, versión 1.89.1 (Microsoft, 2011). Para cada uno de los cinco archivos CSV descargados, se cargarán a un DataFrame distinto utilizando la función `read_csv` de pandas. A continuación, cada DataFrame se filtrará para conservar solo las columnas de interés: Molecule ChEMBL ID, Molecule Name, Molecular Weight, SMILES, Standard Type, Standard Relation, Standard Value, y Standard Units. Posteriormente, se realizará un análisis de los valores faltantes en cada columna utilizando el método `isna` de pandas para cada DataFrame. Luego, se procederá a limpiar los valores faltantes en las columnas "Standard Value" y "SMILES" en cada DataFrame, ya que estas son esenciales para el entrenamiento de los modelos de aprendizaje automático y el cálculo de descriptores moleculares. A continuación, se eliminarán las moléculas duplicadas basándose en la columna "Molecule Name" utilizando el método `drop_duplicates` de pandas en cada DataFrame. Finalmente, los DataFrames resultantes de la limpieza de los datos se almacenarán en archivos CSV.

6.1.3. Cálculo de pIC₅₀ y establecimiento del umbral de inhibición

Se cargarán los nuevos archivos guardados en formato CSV como DataFrames de pandas. En cada uno de estos nuevos DataFrames, se añadirá una columna denominada "actividad" cuyos valores serán binarios, es decir, 0 o 1. El valor 1 representará la actividad de una molécula respecto a la proteína correspondiente del DataFrame, mientras que el valor 0 indicará inactividad respecto a esa proteína. El umbral para definir la actividad e inactividad se establecerá en 1 μ M. Posteriormente, se creará una nueva columna denominada "pIC₅₀" en cada DataFrame para almacenar el valor de pIC₅₀ de cada molécula. Para calcular el valor de pIC₅₀, se empleará la ecuación $pIC_{50} = -\log_{10}(IC_{50})$. Esta transformación es necesaria para normalizar los datos de actividad inhibitoria y facilitar el análisis comparativo entre diferentes moléculas y proteínas. Como un análisis exploratorio inicial de las moléculas, se calcularán la media, mediana y moda para los grupos inhibitorio y no inhibitorio en cada uno de los DataFrames, utilizando la función groupby de pandas y las respectivas funciones de agregación. Posteriormente, se crearán boxplots empleando la función sns.boxplot de la librería Seaborn en Python (Waskom, 2021). Finalmente, los DataFrames actualizados se guardarán nuevamente en archivos CSV para su uso en etapas posteriores.

6.1.4. Cálculo de los descriptores moleculares

Se cargarán los cinco archivos CSV que contienen los valores de pIC₅₀ como nuevos DataFrames de pandas. A continuación, en cada uno de los cinco DataFrames, se creará una nueva columna que almacenará objetos tipo mol generados a partir de los SMILES de cada molécula, utilizando la función MolFromSmiles del módulo RDKit versión 2023.9.5 (RDKit team, 2024). Estos objetos tipo mol son representaciones internas de moléculas químicas que permiten manipular y analizar estructuras moleculares dentro de la librería RDKit. Posteriormente, se utilizará la función all_descriptors para calcular específicamente descriptores 2D, hasta un total de 110 descriptores bidimensionales ofrecidos por RDKit. Estos descriptores incluyen huellas moleculares, carga y polaridad, topología, anillos y ciclos, y estado electrónico. Se añadirá una nueva columna a cada DataFrame con el nombre del

descriptor respectivo y los valores calculados para cada molécula. Finalmente, los DataFrames con los descriptores calculados se guardarán como archivos CSV.

6.1.5. Selección de características importantes

Se cargarán los archivos CSV, que ahora incluyen los descriptores moleculares bidimensionales calculados con RDKit, en nuevos DataFrames de pandas. Se creará un correlograma que incluya los descriptores moleculares y el valor de pIC₅₀ utilizando el método corr de pandas. Para la visualización gráfica del correlograma, se empleará la función heatmap de seaborn, y las imágenes resultantes de los cinco DataFrames se guardarán. Posteriormente, se descartarán aquellos descriptores moleculares cuya correlación entre sí sea superior a 0.7, así como aquellos con una correlación de 0 o cercana a 0, utilizando máscaras de pandas sobre el correlograma. Para cada DataFrame, se eliminarán las columnas de los descriptores que cumplan con estos criterios, con el fin de evitar el sobreajuste y asegurar la optimización de los modelos de aprendizaje automático.

6.2. Entrenamiento de los modelos de aprendizaje automático

6.2.1. División del conjunto de datos

Se cargarán los archivos CSV que contienen los datos de los descriptores moleculares en nuevos DataFrames. Se creará una variable llamada X para cada uno de los DataFrames, que contendrá únicamente las columnas de los descriptores moleculares de todas las moléculas, correspondientes a su respectiva proteína diana. Asimismo, para cada DataFrame se creará una variable y que contendrá únicamente la columna de pIC₅₀ de las moléculas. Posteriormente, tanto X como y se dividirán en conjuntos de entrenamiento y prueba utilizando la función train_test_split de la librería scikit-learn 1.3.2. de Python (Pedregosa et al., 2011). Los datos se dividirán en una proporción de 80% para el conjunto de entrenamiento y 20% para el conjunto de prueba. Cada una de las nuevas variables X de entrenamiento y prueba, así como y de entrenamiento y prueba, se guardarán en formato CSV para su posterior uso en la validación interna y externa, así como para el entrenamiento de los modelos de aprendizaje automático.

6.2.2. Validación interna

Se emplearán los modelos de aprendizaje automático Random Forest de la librería scikit-learn (Pedregosa et al., 2011) y XGBoost de la librería XGBoost 2.0.3 (T. Chen & Guestrin, 2016). Se cargarán todos los archivos CSV de las variables previamente divididas en conjuntos de entrenamiento y prueba como DataFrames. Se entrenará el modelo XGBoost y el modelo Random Forest por cada uno de los DataFrames divididos en conjunto de datos y de pruebas, dando como resultado 10 modelos. Para cada uno de estos conjuntos, se instanciarán los modelos de XGBoost y Random Forest utilizando las clases XGBClassifier de XGBoost y RandomForestRegressor de scikit-learn, respectivamente. Se realizará una validación interna previa al entrenamiento de los modelos con el objetivo de determinar los mejores hiperparámetros para cada uno de ellos. Para los modelos de Random Forest, se ajustarán los hiperparámetros “n_estimators”, “max_depth” y “max_features”. El hiperparámetro “n_estimators” define el número de árboles en el bosque, “max_depth” la profundidad máxima de cada árbol, y “max_features” controla cuántas características se consideran para dividir en cada nodo. Se utilizará la siguiente rejilla de hiperparámetros para Random Forest: “n_estimators” = [50, 100, 150], “max_depth” = [10, 20, 30] y “max_features” = [“auto”, “sqrt”, “log2”]. Del mismo modo, para el modelo XGBoost se ajustarán los hiperparámetros “n_estimators”, “learning_rate” y “max_depth”. Los hiperparámetros “n_estimators” y “max_depth” funcionan de manera similar a los del Random Forest. El hiperparámetro “learning_rate” define la tasa de aprendizaje. Se utilizará la siguiente rejilla de hiperparámetros para XGBoost: “n_estimators” = [50, 100, 150], “max_depth” = [10, 20, 30] y “learning_rate” = [0.01, 0.05, 0.1]. Para determinar los mejores hiperparámetros de cada modelo, se realizará una búsqueda en cuadrícula con validación cruzada utilizando la función GridSearchCV de scikit-learn, con 5 pliegues de validación cruzada, empleando el error cuadrático medio (MSE, por sus siglas en inglés “Mean Squared Error”) como métrica de evaluación. Finalmente, se entrenarán estos modelos con la búsqueda en cuadrícula y se obtendrán los mejores hiperparámetros para cada modelo. Los mejores hiperparámetros de los 10 modelos construidos se guardarán en un único DataFrame para su posterior uso en el entrenamiento de los modelos.

6.2.3. Entrenamiento de los modelos y validación externa

Se cargarán nuevamente los datos de los archivos CSV que contienen las variables divididas en los conjuntos de entrenamiento y prueba en DataFrames. Se instanciarán los modelos empleando las clases XGBClassifier y RandomForestRegressor, y se configurarán cada uno de los 10 modelos con los mejores hiperparámetros obtenidos durante la validación interna. Posteriormente, cada modelo se entrenará con su respectivo conjunto de datos de entrenamiento, X_train e y_train. Una vez entrenados los modelos, se almacenarán en una variable llamada y_predicciones los resultados de las predicciones al aplicar los modelos a sus respectivos conjuntos de prueba, X_test. Finalmente, se utilizará la variable y_predicciones junto con y_test de cada uno de los modelos para estimar las métricas de precisión, accuracy, recall y F1 score, empleando la función classification_report de scikit-learn. Con el objetivo de evitar el reentrenamiento de los modelos, se empleará la librería de Python joblib, que permite guardar y cargar modelos de aprendizaje automático. Cada modelo se guardará en un archivo con formato “.pkl” y se cargará posteriormente para su aplicación en la predicción de la actividad de moléculas en una base de datos de compuestos naturales.

6.3. Determinación de potenciales inhibidores

6.3.1. Obtención de las moléculas a partir de las bases de datos

Se descargarán datos de moléculas químicas para el reposicionamiento de fármacos de las bases de datos ReframeDB (Janes et al., 2018) y COCONUT (Sorokina et al., 2021). Estas bases de datos se utilizarán debido a las ventajas que ofrecen, como la validación previa de los compuestos y la inclusión de productos naturales. Se descargará la totalidad de los compuestos de ReframeDB, junto con sus representaciones SMILES, resultando en un total de 12,790 compuestos en formato CSV. De manera similar, se descargarán los compuestos y sus SMILES de la base de datos COCONUT, obteniendo un total de 407,270 compuestos en formato SMI.

Estos archivos se emplearán como fuente de datos para los modelos QSAR previamente entrenados.

6.3.2. Tratamiento de los datos descargados

Los datos descargados de las bases de datos se cargarán como DataFrames utilizando la librería pandas de Python. Para la base de datos de Reframe, se utilizará la función `read_csv` de pandas para cargar los datos en un DataFrame. Se creará un nuevo DataFrame que incluya solo las columnas de interés, que son “smile” y “rfm_id”. De manera similar, se cargará el archivo de la base de datos de COCONUT, que está en formato “smi”, leyéndolo con pandas como si fuera un CSV, especificando que la separación se realiza mediante espacios en blanco. Para ambos DataFrames, se empleará el método `dropna` de pandas con el objetivo de limpiar los datos faltantes y el método `drop_duplicates` para eliminar datos repetidos. Los DataFrames resultantes después del tratamiento de datos se guardarán como archivos en formato CSV.

6.3.3. Cálculo de descriptores moleculares

Se cargarán los archivos CSV con los datos tratados de las bases de datos como DataFrames de pandas. Para ambos DataFrames, se creará una nueva columna que almacenará objetos tipo mol generados a partir de los SMILES de cada molécula, utilizando la función `MolFromSmiles` del módulo RDKit, de la misma manera que se hizo para los datos empleados en el entrenamiento y pruebas de los modelos de aprendizaje automático. Posteriormente, se utilizará la función `all_descriptors` para calcular específicamente los descriptores seleccionados resultantes del correlograma. Se añadirá una nueva columna a cada DataFrame con el nombre del descriptor respectivo y los valores calculados para cada molécula. Finalmente, los DataFrames con los descriptores calculados se guardarán como archivos CSV.

6.3.4. Determinación de potenciales inhibidores multiblanco

Se cargarán los archivos CSV que contienen las moléculas de las bases de datos, así como los 10 modelos QSAR entrenados, utilizando la función `load` de la librería `joblib` de Python. Se

creará una variable llamada X que almacenará todos los datos de los descriptores moleculares de las moléculas descargadas. Esta variable X se introducirá en cada uno de los diez modelos QSAR para predecir la actividad de estas moléculas utilizando la función predict de scikit-learn. Los resultados de las predicciones de los 10 modelos se guardarán en un único DataFrame, donde cada columna llevará el nombre del respectivo modelo y la proteína diana. Posteriormente, se realizará un filtrado utilizando la librería pandas para determinar cuáles moléculas presentan inhibición en al menos 3 proteínas dianas. Con base en este criterio, se identificarán los potenciales inhibidores multidiana para la asociación Alzheimer-Diabetes mellitus. El DataFrame resultante se guardará como un archivo CSV para su posterior análisis mediante acoplamiento molecular.

6.4. Acoplamiento molecular

7. CRONOGRAMA DE ACTIVIDADES

8. BIBLIOGRAFÍA

- Arnold, S. E., Arvanitakis, Z., Macauley-Rambach, S. L., Koenig, A. M., Wang, H. Y., Ahima, R. S., Craft, S., Gandy, S., Buettner, C., Stoeckel, L. E., Holtzman, D. M., & Nathan, D. M. (2018). Brain insulin resistance in type 2 diabetes and Alzheimer disease: concepts and conundrums. *Nature Reviews Neurology* 2018 14:3, 14(3), 168–181. <https://doi.org/10.1038/nrneurol.2017.185>
- Balczon, R., Lin, M. T., Voth, S., Nelson, A. R., Schupp, J. C., Wagener, B. M., Pittet, J. F., & Stevens, T. (2024). Lung Endothelium, Tau, and Amyloids in Health and Disease. *Physiological Reviews*, 104(2), 533–587. https://doi.org/10.1152/PHYSREV.00006.2023/ASSET/IMAGES/LARGE/PHYSREV.00006.2023_F016.JPEG
- Bao, L. Q., Baecker, D., Mai Dung, D. T., Phuong Nhung, N., Thi Thuan, N., Nguyen, P. L., Phuong Dung, P. T., Huong, T. T. L., Rasulev, B., Casanola-Martin, G. M., Nam, N. H., & Pham-The, H. (2023). Development of Activity Rules and Chemical Fragment Design for In Silico Discovery of AChE and BACE1 Dual Inhibitors against Alzheimer's Disease. *Molecules*, 28(8), 3588. <https://doi.org/10.3390/MOLECULES28083588/S1>
- Batool, M., Ahmad, B., & Choi, S. (2019). A Structure-Based Drug Discovery Paradigm. *International Journal of Molecular Sciences* 2019, Vol. 20, Page 2783, 20(11), 2783. <https://doi.org/10.3390/IJMS20112783>
- Bird, T. D. (2008). Genetic Aspects of Alzheimer Disease. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 10(4), 231. <https://doi.org/10.1097/GIM.0B013E31816B64DC>
- Bosc, N., Atkinson, F., Felix, E., Gaulton, A., Hersey, A., & Leach, A. R. (2019). Large scale comparison of QSAR and conformal prediction methods and their applications in drug discovery. *Journal of Cheminformatics*, 11(1), 1–16. <https://doi.org/10.1186/S13321-018-0325-4/TABLES/6>
- Boukhalifa, W., Jmel, H., Kheriji, N., Gouiza, I., Dallali, H., Hechmi, M., & Kefi, R. (2023). Decoding the genetic relationship between Alzheimer's disease and type 2 diabetes: potential

- risk variants and future direction for North Africa. *Frontiers in Aging Neuroscience*, 15, 1114810. <https://doi.org/10.3389/FNAGI.2023.1114810/BIBTEX>
- Brands, A. M. A., Biessels, G. J., De Haan, E. H. F., Kappelle, L. J., & Kessels, R. P. C. (2005). The Effects of Type 1 Diabetes on Cognitive PerformanceA meta-analysis. *Diabetes Care*, 28(3), 726–735. <https://doi.org/10.2337/DIACARE.28.3.726>
- Cargnello, M., & Roux, P. P. (2011). Activation and Function of the MAPKs and Their Substrates, the MAPK-Activated Protein Kinases. *Microbiology and Molecular Biology Reviews* : *MMBR*, 75(1), 50. <https://doi.org/10.1128/MMBR.00031-10>
- Carracedo-Reboredo, P., Liñares-Blanco, J., Rodríguez-Fernández, N., Cedrón, F., Novoa, F. J., Carballal, A., Maojo, V., Pazos, A., & Fernandez-Lozano, C. (2021). A review on machine learning approaches and trends in drug discovery. *Computational and Structural Biotechnology Journal*, 19, 4538–4558. <https://doi.org/10.1016/J.CSBJ.2021.08.011>
- Castillo-Velázquez, R., Martínez-Morales, F., Castañeda-Delgado, J. E., García-Hernández, M. H., Herrera-Mayorga, V., Paredes-Sánchez, F. A., Rivera, G., Rivas-Santiago, B., & Lara-Ramírez, E. E. (2023). Bioinformatic prediction of the molecular links between Alzheimer’s disease and diabetes mellitus. *PeerJ*, 11. <https://doi.org/10.7717/PEERJ.14738>
- Cavas, L., Topcam, G., Gundogdu-Hizliates, C., & Ergun, Y. (2019). Neural Network Modeling of AChE Inhibition by New Carbazole-Bearing Oxazolones. *Interdisciplinary Sciences – Computational Life Sciences*, 11(1), 95–107. <https://doi.org/10.1007/S12539-017-0245-4/FIGURES/7>
- Chaitanya Kumar, A., Andrew John, J., Raja, M., & Vijaya, P. (2023). Genetic factor analysis for an early diagnosis of autism through machine learning. *Data Science for Genomics*, 69–84. <https://doi.org/10.1016/B978-0-323-98352-5.00001-X>
- Chen, H., Hu, S., Hua, R., & Zhao, X. (2021). Improved naive Bayes classification algorithm for traffic risk management. *Eurasip Journal on Advances in Signal Processing*, 2021(1), 1–12. <https://doi.org/10.1186/S13634-021-00742-6/TABLES/5>
- Chen, T., & Guestrin, C. (2016a). XGBoost: A Scalable Tree Boosting System. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794. <https://doi.org/10.1145/2939672.2939785>

- Chen, T., & Guestrin, C. (2016b). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Cho, S. B. (2024). Comorbidity Genes of Alzheimer's Disease and Type 2 Diabetes Associated with Memory and Cognitive Function. *International Journal of Molecular Sciences*, 25(4), 2211. <https://doi.org/10.3390/IJMS25042211/S1>
- Choi, H. J., Jeong, Y. J., Kim, J., & Hoe, H. S. (2023). EGFR is a potential dual molecular target for cancer and Alzheimer's disease. *Frontiers in Pharmacology*, 14, 1238639. <https://doi.org/10.3389/FPHAR.2023.1238639/BIBTEX>
- Chollet, F. (2018). *Deep Learning with Python* (Toni Arritola, Ed.). Manning Publications Co.
- Clodfelder-Miller, B. J., Zmijewska, A. A., Johnson, G. V. W., & Jope, R. S. (2006). Tau Is Hyperphosphorylated at Multiple Sites in Mouse Brain In Vivo After Streptozotocin-Induced Insulin Deficiency. *Diabetes*, 55(12), 3320–3325. <https://doi.org/10.2337/DB06-0485>
- Cobley, J. N., Fiorello, M. L., & Bailey, D. M. (2018). 13 reasons why the brain is susceptible to oxidative stress. *Redox Biology*, 15, 490–503. <https://doi.org/10.1016/J.REDOX.2018.01.008>
- Concu, R., González-Durruthy, M., & Cordeiro, M. N. D. S. (2020). Developing a Multi-target Model to Predict the Activity of Monoamine Oxidase A and B Drugs. *Current Topics in Medicinal Chemistry*, 20(18), 1593–1600. <https://doi.org/10.2174/1568026620666200603121224>
- Coumar, M. S. (2021). *Molecular docking for computer-aided drug design: fundamentals, techniques, resources and applications*.
- Cukierman, T., Gerstein, H. C., & Williamson, J. D. (2005). Cognitive decline and dementia in diabetes - Systematic overview of prospective observational studies. *Diabetologia*, 48(12), 2460–2469. <https://doi.org/10.1007/S00125-005-0023-4/FIGURES/3>
- Cummings, J. L., Doody, R., & Clark, C. (2007). Disease-modifying therapies for Alzheimer disease: challenges to early intervention. *Neurology*, 69(16), 1622–1634. <https://doi.org/10.1212/01.WNL.0000295996.54210.69>
- Cummings, J. L., Morstorf, T., & Zhong, K. (2014). Alzheimer's disease drug-development pipeline: Few candidates, frequent failures. *Alzheimer's Research and Therapy*, 6(4), 1–7. <https://doi.org/10.1186/ALZRT269/TABLES/3>

- Davies, M., Nowotka, M., Papadatos, G., Dedman, N., Gaulton, A., Atkinson, F., Bellis, L., & Overington, J. P. (2015). ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic Acids Research*, 43(Web Server issue), W612. <https://doi.org/10.1093/NAR/GKV352>
- de Boer, R. J., van Lidth de Jeude, J. F., & Heijmans, J. (2024). ER stress and the unfolded protein response in gastrointestinal stem cells and carcinogenesis. *Cancer Letters*, 587. <https://doi.org/10.1016/J.CANLET.2024.216678>
- De Felice, F. G. (2013). Alzheimer's disease and insulin resistance: translating basic science into clinical applications. *The Journal of Clinical Investigation*, 123(2), 531–539. <https://doi.org/10.1172/JCI64595>
- de la Monte, S. M. (2009). Insulin resistance and Alzheimer's disease. *BMB Reports*, 42(8), 475. <https://doi.org/10.5483/BMBREP.2009.42.8.475>
- Deane, R., Bell, R., Sagare, A., & Zlokovic, B. (2009). Clearance of amyloid-beta peptide across the blood-brain barrier: implication for therapies in Alzheimer's disease. *CNS & Neurological Disorders Drug Targets*, 8(1), 16–30. <https://doi.org/10.2174/187152709787601867>
- Devendra, D., & Eisenbarth, G. S. (2003). 17. Immunologic endocrine disorders. *Journal of Allergy and Clinical Immunology*, 111(2 SUPPL. 2). <https://doi.org/10.1067/mai.2003.81>
- Dhamodharan, G., & Mohan, C. G. (2022). Machine learning models for predicting the activity of AChE and BACE1 dual inhibitors for the treatment of Alzheimer's disease. *Molecular Diversity*, 26(3), 1501–1517. <https://doi.org/10.1007/S11030-021-10282-8/TABLES/6>
- Ding, X., Meng, C., Dong, H., Zhang, S., Zhou, H., Tan, W., Huang, L., He, A., Li, J., Huang, J., Li, W., Zou, F., & Zou, M. (2022). Extracellular Hsp90 α , which participates in vascular inflammation, is a novel serum predictor of atherosclerosis in type 2 diabetes. *BMJ Open Diabetes Research and Care*, 10(1), e002579. <https://doi.org/10.1136/BMJDRC-2021-002579>
- Djuris, J., Ibric, S., & Djuric, Z. (2013). Chemometric methods application in pharmaceutical products and processes analysis and control. *Computer-Aided Applications in Pharmaceutical Technology*, 57–90. <https://doi.org/10.1533/9781908818324.57>
- Docherty, F. M., & Russ, H. A. (2019). Cell–Cell Interactions Driving Differentiation of Adult Pancreatic Stem Cells. *Encyclopedia of Tissue Engineering and Regenerative Medicine: Volumes 1-3*, 1–3, 367–374. <https://doi.org/10.1016/B978-0-12-801238-3.65615-5>

- Edgar, T. W., & Manz, D. O. (2017). Chapter 2 - Science and Cyber Security. *Research Methods for Cyber Security*, 33–62.
<http://www.sciencedirect.com/science/article/pii/B9780128053492000029>
- Erf, G. F. (2014). Autoimmune Diseases of Poultry. *Avian Immunology: Second Edition*, 315–332.
<https://doi.org/10.1016/B978-0-12-396965-1.00018-2>
- FDA. (2019). *Drugs@FDA Glossary of Terms | FDA*. <https://www.fda.gov/drugs/drug-approvals-and-databases/drugsfda-glossary-terms>
- Fisher, A. (2008). M1 muscarinic agonists target major hallmarks of Alzheimer’s disease--the pivotal role of brain M1 receptors. *Neuro-Degenerative Diseases*, 5(3–4), 237–240.
<https://doi.org/10.1159/000113712>
- Gabbouj, S., Ryhänen, S., Martinen, M., Wittrahm, R., Takalo, M., Kemppainen, S., Martiskaine, H., Tanila, H., Haapasalo, A., Hiltunen, M., & Natunen, T. (2019). Altered insulin signaling in Alzheimer’s disease brain-special emphasis on pi3k-akt pathway. *Frontiers in Neuroscience*, 13(JUN), 456800. <https://doi.org/10.3389/FNINS.2019.00629/BIBTEX>
- Gallagher, C. J., Keene, K. L., Mychaleckyj, J. C., Langefeld, C. D., Hirschhorn, J. N., Henderson, B. E., Gordon, C. J., Freedman, B. I., Rich, S. S., Bowden, D. W., & Sale, M. M. (2007). Investigation of the Estrogen Receptor- α Gene With Type 2 Diabetes and/or Nephropathy in African-American and European-American Populations. *Diabetes*, 56(3), 675–684.
<https://doi.org/10.2337/DB06-0303>
- Gan, H. K., Cvrljevic, A. N., & Johns, T. G. (2013). The epidermal growth factor receptor variant III (EGFRvIII): Where wild things are altered. *FEBS Journal*, 280(21), 5350–5370.
<https://doi.org/10.1111/febs.12393>
- Geng, C., Wang, Z. Bin, & Tang, Y. (2024). Machine learning in Alzheimer’s disease drug discovery and target identification. *Ageing Research Reviews*, 93, 102172.
<https://doi.org/10.1016/J.ARR.2023.102172>
- Gillespie, K. M. (2006). Type 1 diabetes: pathogenesis and prevention. *CMAJ*, 175(2), 165–170.
<https://doi.org/10.1503/CMAJ.060244>
- Goedert, M., Spillantini, M. G., Jakes, R., Rutherford, D., & Crowther, R. A. (1989). Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer’s disease. *Neuron*, 3(4), 519–526.
[https://doi.org/10.1016/0896-6273\(89\)90210-9](https://doi.org/10.1016/0896-6273(89)90210-9)

- Gong, C. X., Liu, F., & Iqbal, K. (2018). Multifactorial Hypothesis and Multi-Targets for Alzheimer's Disease. *Journal of Alzheimer's Disease: JAD*, 64(s1), S107–S117. <https://doi.org/10.3233/JAD-179921>
- Greeny, A., Nair, A., Sadanandan, P., Satarker, S., Famurewa, A. C., & Nampoothiri, M. (2024). Epigenetic Alterations in Alzheimer's Disease: Impact on Insulin Signaling and Advanced Drug Delivery Systems. *Biology 2024*, Vol. 13, Page 157, 13(3), 157. <https://doi.org/10.3390/BIOLOGY13030157>
- Grundke-Iqbal, I., Iqbal, K., Tung, Y. C., Quinlan, M., Wisniewski, H. M., & Binder, L. I. (1986). Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proceedings of the National Academy of Sciences of the United States of America*, 83(13), 4913. <https://doi.org/10.1073/PNAS.83.13.4913>
- Gu, F., Zhu, M., Shi, J., Hu, Y., & Zhao, Z. (2008). Enhanced oxidative stress is an early event during development of Alzheimer-like pathologies in presenilin conditional knock-out mice. *Neuroscience Letters*, 440(1), 44–48. <https://doi.org/10.1016/J.NEULET.2008.05.050>
- Henriksen, E. J., Diamond-Stanic, M. K., & Marchionne, E. M. (2011). Oxidative stress and the etiology of insulin resistance and type 2 diabetes. *Free Radical Biology and Medicine*, 51(5), 993–999. <https://doi.org/10.1016/J.FREERADBIOMED.2010.12.005>
- Henson, J., Anyiam, O., & Vishnubala, D. (2023). Type 2 Diabetes. *Exercise Management for Referred Medical Conditions*, 223–252. <https://doi.org/10.4324/9781315102399-12>
- Huang, A., Yang, Y. M., Yan, C., Kaley, G., Hintze, T. H., & Sun, D. (2012). Altered MAPK Signaling in Progressive Deterioration of Endothelial Function in Diabetic Mice. *Diabetes*, 61(12), 3181. <https://doi.org/10.2337/DB12-0559>
- IMSS. (2015). *Enfermedad de Alzheimer*. <https://www.imss.gob.mx/salud-en-linea/enfermedad-alzheimer>
- Ishihara, T., Hong, M., Zhang, B., Nakagawa, Y., Lee, M. K., Trojanowski, J. Q., & Lee, V. M. Y. (1999). Age-dependent emergence and progression of a tauopathy in transgenic mice overexpressing the shortest human tau isoform. *Neuron*, 24(3), 751–762. [https://doi.org/10.1016/S0896-6273\(00\)81127-7](https://doi.org/10.1016/S0896-6273(00)81127-7)
- Ismail, L., Materwala, H., & Al Kaabi, J. (2021). Association of risk factors with type 2 diabetes: A systematic review. *Computational and Structural Biotechnology Journal*, 19, 1759–1785. <https://doi.org/10.1016/J.CSBJ.2021.03.003>

- Johnson, G. V. W., & Stoothoff, W. H. (2004). Tau phosphorylation in neuronal cell function and dysfunction. *Journal of Cell Science*, 117(24), 5721–5729. <https://doi.org/10.1242/JCS.01558>
- Kandimalla, R., Thirumala, V., & Reddy, P. H. (2017). Is Alzheimer's disease a Type 3 Diabetes? A critical appraisal. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1863(5), 1078–1089. <https://doi.org/10.1016/J.BBADIS.2016.08.018>
- Kopf, D., & Frölich, L. (2009). Risk of Incident Alzheimer's Disease in Diabetic Patients: A Systematic Review of Prospective Trials. *Journal of Alzheimer's Disease*, 16(4), 677–685. <https://doi.org/10.3233/JAD-2009-1011>
- Krischer, J. P., Liu, X., Lernmark, Å., Hagopian, W. A., Rewers, M. J., She, J. X., Toppari, J., Ziegler, A. G., & Akolkar, B. (2022). Predictors of the Initiation of Islet Autoimmunity and Progression to Multiple Autoantibodies and Clinical Diabetes: The TEDDY Study. *Diabetes Care*, 45(10), 2271. <https://doi.org/10.2337/DC21-2612>
- Kumar, V., Saha, A., & Roy, K. (2023). Multi-target QSAR modeling for the identification of novel inhibitors against Alzheimer's disease. *Chemometrics and Intelligent Laboratory Systems*, 233, 104734. <https://doi.org/10.1016/J.CHEMOLAB.2022.104734>
- Kurochkin, I. V., Guarnera, E., & Berezovsky, I. N. (2018). Insulin-Degrading Enzyme in the Fight against Alzheimer's Disease. *Trends in Pharmacological Sciences*, 39(1), 49–58. <https://doi.org/10.1016/J.TIPS.2017.10.008>
- Lavecchia, A., & Cerchia, C. (2016). In silico methods to address polypharmacology: current status, applications and future perspectives. *Drug Discovery Today*, 21(2), 288–298. <https://doi.org/10.1016/J.DRUDIS.2015.12.007>
- Lee, J., Kumar, S., Lee, S. Y., Park, S. J., & Kim, M. H. (2019). Development of Predictive Models for Identifying Potential S100A9 Inhibitors Based on Machine Learning Methods. *Frontiers in Chemistry*, 7, 453573. <https://doi.org/10.3389/FCHEM.2019.00779/BIBTEX>
- Levy-Lahad, E., Wasco, W., Poorkaj, P., Romano, D. M., Oshima, J., Pettingell, W. H., Yu, C. E., Jondro, P. D., Schmidt, S. D., Wang, K., Crowley, A. C., Fu, Y. H., Guenette, S. Y., Galas, D., Nemens, E., Wijsman, E. M., Bird, T. D., Schellenberg, G. D., & Tanzi, R. E. (1995). Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science (New York, N.Y.)*, 269(5226), 973–977. <https://doi.org/10.1126/SCIENCE.7638622>

- Liskowsky, W., & Schliebs, R. (2006). Muscarinic acetylcholine receptor inhibition in transgenic Alzheimer-like Tg2576 mice by scopolamine favours the amyloidogenic route of processing of amyloid precursor protein. *International Journal of Developmental Neuroscience: The Official Journal of the International Society for Developmental Neuroscience*, 24(2–3), 149–156. <https://doi.org/10.1016/J.IJDEVNEU.2005.11.010>
- Liu, A., Fernandes, B. S., Citu, C., & Zhao, Z. (2024). Unraveling the intercellular communication disruption and key pathways in Alzheimer's disease: an integrative study of single-nucleus transcriptomes and genetic association. *Alzheimer's Research and Therapy*, 16(1), 1–17. <https://doi.org/10.1186/S13195-023-01372-W/TABLES/1>
- Lovestone, S., Boada, M., Dubois, B., Hüll, M., Rinne, J. O., Huppertz, H. J., Calero, M., Andrés, M. V., Gómez-Carrillo, B., León, T., & Del Ser, T. (2015). A Phase II Trial of Tideglusib in Alzheimer's Disease. *Journal of Alzheimer's Disease*, 45(1), 75–88. <https://doi.org/10.3233/JAD-141959>
- Ma, S. L., Tang, N. L. S., Tam, C. W. C., Lui, V. W. C., Lau, E. S. S., Zhang, Y. P., Chiu, H. F. K., & Lam, L. C. W. (2009). Polymorphisms of the estrogen receptor α (ESR1) gene and the risk of Alzheimer's disease in a southern Chinese community. *International Psychogeriatrics*, 21(5), 977–986. <https://doi.org/10.1017/S1041610209990068>
- Mahley, R. W., & Rall, S. C. (2000). Apolipoprotein E: Far more than a lipid transport protein. *Annual Review of Genomics and Human Genetics*, 1(2000), 507–537. <https://doi.org/10.1146/ANNUREV.GENOM.1.1.507/CITE/REFWORKS>
- Medline. (2020). *Diabetes: MedlinePlus en español*. <https://medlineplus.gov/spanish/diabetes.html>
- Medline. (2024). *Demencia: MedlinePlus enciclopedia médica*. <https://medlineplus.gov/spanish/ency/article/000739.htm>
- Miao, J., Shi, R., Li, L., Chen, F., Zhou, Y., Tung, Y. C., Hu, W., Gong, C. X., Iqbal, K., & Liu, F. (2019). Pathological Tau From Alzheimer's Brain Induces Site-Specific Hyperphosphorylation and SDS- and Reducing Agent-Resistant Aggregation of Tau in vivo. *Frontiers in Aging Neuroscience*, 11, 447903. <https://doi.org/10.3389/FNAGI.2019.00034/BIBTEX>
- Michailidis, M., Tata, D. A., Moraitou, D., Kavvadas, D., Karachrysafi, S., Papamitsou, T., Vareltzis, P., & Papaliagkas, V. (2022). Antidiabetic Drugs in the Treatment of Alzheimer's

- Disease. *International Journal of Molecular Sciences*, 23(9).
<https://doi.org/10.3390/IJMS23094641>
- Microsoft. (2011). *Visual Studio Code* (1.89.1). Microsoft.
- Mohandas, E., Rajmohan, V., & Raghunath, B. (2009). Neurobiology of Alzheimer's disease. *Indian Journal of Psychiatry*, 51(1), 55. <https://doi.org/10.4103/0019-5545.44908>
- Mukerjee, N., Al-Khafaji, K., Maitra, S., Suhail Wadi, J., Sachdeva, P., Ghosh, A., Buchade, R. S., Chaudhari, S. Y., Jadhav, S. B., Das, P., Hasan, M. M., Rahman, M. H., Albadrani, G. M., Altyar, A. E., Kamel, M., Algahtani, M., Shinan, K., Theyab, A., Abdel-Daim, M. M., ... Sharma, R. (2022). Recognizing novel drugs against Keap1 in Alzheimer's disease using machine learning grounded computational studies. *Frontiers in Molecular Neuroscience*, 15, 1036552. <https://doi.org/10.3389/FNMOL.2022.1036552/BIBTEX>
- Mullins, R. J., Diehl, T. C., Chia, C. W., & Kapogiannis, D. (2017). Insulin resistance as a link between amyloid-beta and tau pathologies in Alzheimer's disease. *Frontiers in Aging Neuroscience*, 9(MAY), 251216. <https://doi.org/10.3389/FNAGI.2017.00118/BIBTEX>
- NIDDK. (2023). *¿Qué es la diabetes?* - NIDDK. <https://www.niddk.nih.gov/health-information/informacion-de-la-salud/diabetes/informacion-general/que-es>
- NIH. (2017). *¿Qué causa la enfermedad de Alzheimer?* | National Institute on Aging. <https://www.nia.nih.gov/espanol/enfermedad-alzheimer/causa-enfermedad-alzheimer>
- NIH. (2023). *¿Cómo se trata la enfermedad de Alzheimer?* | National Institute on Aging. <https://www.nia.nih.gov/espanol/cuidado-medico-enfermedad-alzheimer/como-se-trata-enfermedad-alzheimer>
- NIH gene. (2024a). *ESR1 estrogen receptor 1 [Homo sapiens (human)]* - Gene - NCBI. <https://www.ncbi.nlm.nih.gov/gene/2099>
- NIH gene. (2024b). *HSP90AA1 heat shock protein 90 alpha family class A member 1 [Homo sapiens (human)]* - Gene - NCBI. <https://www.ncbi.nlm.nih.gov/gene/3320>
- NIH gene. (2024c). *IRS1 insulin receptor substrate 1 [Homo sapiens (human)]* - Gene - NCBI. <https://www.ncbi.nlm.nih.gov/gene/3667>
- NIH gene. (2024d). *PIK3R1 phosphoinositide-3-kinase regulatory subunit 1 [Homo sapiens (human)]* - Gene - NCBI. <https://www.ncbi.nlm.nih.gov/gene/5295>
- NIH gene. (2024e). *UBC ubiquitin C [Homo sapiens (human)]* - Gene - NCBI. <https://www.ncbi.nlm.nih.gov/gene/7316>

- OPS. (2023a). *Demencia* - OPS/OMS | Organización Panamericana de la Salud.
<https://www.paho.org/es/temas/demencia>
- OPS. (2023b). *Diabetes* - OPS/OMS | Organización Panamericana de la Salud.
<https://www.paho.org/es/temas/diabetes>
- Parvathy Dharshini, S. A., Sneha, N. P., Yesudhas, D., Kulandaisamy, A., Rangaswamy, U., Shanmugam, A., Taguchi, Y.-H., & Gromiha, M. M. (2022). Exploring Plausible Therapeutic Targets for Alzheimer's Disease using Multi-omics Approach, Machine Learning and Docking. *Current Topics in Medicinal Chemistry*, 22(22), 1868–1879.
<https://doi.org/10.2174/1568026622666220902110115>
- Paul, D., Sanap, G., Shenoy, S., Kalyane, D., Kalia, K., & Tekade, R. K. (2021). Artificial intelligence in drug discovery and development. *Drug Discovery Today*, 26(1), 80.
<https://doi.org/10.1016/J.DRUDIS.2020.10.010>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Pena, F., Gutierrez-Lerma, A., Quiroz-Baez, R., & Arias, C. (2006). The role of beta-amyloid protein in synaptic function: implications for Alzheimer's disease therapy. *Current Neuroparmacology*, 4(2), 149–163. <https://doi.org/10.2174/157015906776359531>
- Pillai, N., Dasgupta, A., Sudsakorn, S., Fretland, J., & Mavroudis, P. D. (2022). Machine Learning guided early drug discovery of small molecules. *Drug Discovery Today*, 27(8), 2209–2215.
<https://doi.org/10.1016/J.DRUDIS.2022.03.017>
- Raghunathan, S., & Priyakumar, U. D. (2022). Molecular representations for machine learning applications in chemistry. *International Journal of Quantum Chemistry*, 122(7), e26870.
<https://doi.org/10.1002/QUA.26870>
- RDKit team. (2024). *RDKit: Open-source cheminformatics* (2023.9.5).
- Rogaev, E. I., Sherrington, R., Rogaeva, E. A., Levesque, G., Ikeda, M., Liang, Y., Chi, H., Lin, C., Holman, K., Tsuda, T., Mar, L., Sorbi, S., Nacmias, B., Piacentini, S., Amaducci, L., Chumakov, I., Cohen, D., Lannfelt, L., Fraser, P. E., ... George-Hyslop, P. H. S. (1995). Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome

- 1 related to the Alzheimer's disease type 3 gene. *Nature* 1995 376:6543, 376(6543), 775–778.
<https://doi.org/10.1038/376775a0>
- Rosenberg, P. B. (2005). Clinical aspects of inflammation in Alzheimer's disease. *International Review of Psychiatry*, 17(6), 503–514. <https://doi.org/10.1080/02646830500382037>
- Roy, K., Kar, S., Das, R. N., Roy, K., Kar, S., & Das, R. N. (2015). Chapter 12 – Future Avenues. In *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*. <http://www.sciencedirect.com:5070/book/9780128015056/understanding-the-basics-of-qsar-for-applications-in-pharmaceutical-sciences-and-risk-assessment>
- Salazar, D. E., & Gormley, G. (2017). Modern Drug Discovery and Development. *Clinical and Translational Science: Principles of Human Research: Second Edition*, 719–743.
<https://doi.org/10.1016/B978-0-12-802101-9.00041-7>
- Sanabria-Castro, A., Alvarado-Echeverría, I., & Monge-Bonilla, C. (2017). Molecular Pathogenesis of Alzheimer's Disease: An Update. *Annals of Neurosciences*, 24(1), 46–54.
<https://doi.org/10.1159/000464422>
- Schneider, P., Walters, W. P., Plowright, A. T., Sieroka, N., Listgarten, J., Goodnow, R. A., Fisher, J., Jansen, J. M., Duca, J. S., Rush, T. S., Zentgraf, M., Hill, J. E., Krutoholow, E., Kohler, M., Blaney, J., Funatsu, K., Luebkekmann, C., & Schneider, G. (2019). Rethinking drug design in the artificial intelligence era. *Nature Reviews Drug Discovery* 2019 19:5, 19(5), 353–364.
<https://doi.org/10.1038/s41573-019-0050-3>
- Schneider, Patrick., & Xhafa, Fatos. (2022). *Anomaly detection and complex event processing over IoT data streams with application to eHealth and patient data monitoring*. Academic Press.
<http://www.sciencedirect.com:5070/book/9780128238189/anomaly-detection-and-complex-event-processing-over-iot-data-streams>
- Shang, W. L., Zhang, H., & Sui, Y. (2022). Data mining technologies for Mobility-as-a-Service (MaaS). *Big Data and Mobility as a Service*, 203–228. <https://doi.org/10.1016/B978-0-323-90169-7.00008-7>
- Sheng, L., Bayliss, G., & Zhuang, S. (2020). Epidermal Growth Factor Receptor: A Potential Therapeutic Target for Diabetic Kidney Disease. *Frontiers in Pharmacology*, 11.
<https://doi.org/10.3389/FPHAR.2020.598910>
- Shobha, G., & Rangaswamy, S. (2018). Machine Learning. *Handbook of Statistics*, 38, 197–228.
<https://doi.org/10.1016/BS.HOST.2018.07.004>

- Sies, H. (2019). Oxidative Stress: Eustress and Distress in Redox Homeostasis. *Stress: Physiology, Biochemistry, and Pathology Handbook of Stress Series, Volume 3*, 153–163. <https://doi.org/10.1016/B978-0-12-813146-6.00013-8>
- Speight, J., & Pouwer, F. (2023). Type 1 Diabetes. *Cambridge Handbook of Psychology, Health and Medicine: Third Edition*, 477–480. <https://doi.org/10.29309/tpmj/2017.24.12.614>
- SSA México. (2021). *Enfermedad de Alzheimer, demencia más común que afecta a personas adultas mayores* | Secretaría de Salud | Gobierno | gob.mx. <https://www.gob.mx/salud/es/articulos/enfermedad-de-alzheimer-demencia-mas-comun-que-afecta-a-personas-adultas-mayores?idiom=es>
- Starks, E. J., Patrick O’grady, J., Hoscheidt, S. M., Racine, A. M., Carlsson, C. M., Zetterberg, H., Blennow, K., Okonkwo, O. C., Puglielli, L., Asthana, S., Dowling, N. M., Gleason, C. E., Anderson, R. M., Davenport-Sis, N. J., Derungs, L. A. M., Sager, M. A., Johnson, S. C., & Bendlin, B. B. (2015). Insulin Resistance is Associated with Higher Cerebrospinal Fluid Tau Levels in Asymptomatic APOE ϵ 4 Carriers. *Journal of Alzheimer’s Disease*, 46(2), 525–533. <https://doi.org/10.3233/JAD-150072>
- Stelzmann, R. A., Norman Schnitzlein, H., & Reed Murtagh, F. (1995). An English translation of Alzheimer’s 1907 paper, “Über eine eigenartige Erkrankung der Hirnrinde.” *Clinical Anatomy (New York, N.Y.)*, 8(6), 429–431. <https://doi.org/10.1002/CA.980080612>
- Strachan, M. W. J., Deary, I. J., Ewing, F. M. E., & Frier, B. M. (1997). Is Type II Diabetes Associated With an Increased Risk of Cognitive Dysfunction?: A critical review of published studies. *Diabetes Care*, 20(3), 438–445. <https://doi.org/10.2337/DIACARE.20.3.438>
- Subasi, A. (2020). Use of artificial intelligence in Alzheimer’s disease detection. *Artificial Intelligence in Precision Health: From Concept to Applications*, 257–278. <https://doi.org/10.1016/B978-0-12-817133-2.00011-2>
- Sumi, N. J., Ctortocka, C., Hu, Q., Bryant, A. T., Fang, B., Remsing Rix, L. L., Ayaz, M., Kinose, F., Welsh, E. A., Eschrich, S. A., Lawrence, H. R., Koomen, J. M., Haura, E. B., & Rix, U. (2019). Divergent Polypharmacology-Driven Cellular Activity of Structurally Similar Multi-Kinase Inhibitors through Cumulative Effects on Individual Targets. *Cell Chemical Biology*, 26(9), 1240-1252.e11. <https://doi.org/10.1016/j.chembiol.2019.06.003>

- Sun, X., Bromley-Brits, K., & Song, W. (2012). Regulation of β -site APP-cleaving enzyme 1 gene expression and its role in Alzheimer's Disease. *Journal of Neurochemistry*, 120(SUPPL. 1), 62–70. <https://doi.org/10.1111/J.1471-4159.2011.07515.X>
- Tanzi, R. E., & Bertram, L. (2005). Twenty years of the Alzheimer's disease amyloid hypothesis: a genetic perspective. *Cell*, 120(4), 545–555. <https://doi.org/10.1016/J.CELL.2005.02.008>
- Terry, A. V., & Buccafusco, J. J. (2003). The cholinergic hypothesis of age and Alzheimer's disease-related cognitive deficits: recent challenges and their implications for novel drug development. *The Journal of Pharmacology and Experimental Therapeutics*, 306(3), 821–827. <https://doi.org/10.1124/JPET.102.041616>
- The pandas development team. (2020). *pandas-dev/pandas: Pandas (2.1.4)*. Zenodo.
- Troutwine, B. R., Hamid, L., Lysaker, C. R., Strope, T. A., & Wilkins, H. M. (2022). Apolipoprotein E and Alzheimer's disease. *Acta Pharmaceutica Sinica B*, 12(2), 496–510. <https://doi.org/10.1016/J.APSB.2021.10.002>
- Tumminia, A., Vinciguerra, F., Parisi, M., & Frittitta, L. (2018). Type 2 Diabetes Mellitus and Alzheimer's Disease: Role of Insulin Signalling and Therapeutic Implications. *International Journal of Molecular Sciences* 2018, Vol. 19, Page 3306, 19(11), 3306. <https://doi.org/10.3390/IJMS19113306>
- Uddin, M. S., Al Mamun, A., Kabir, M. T., Ashraf, G. M., Bin-Jumah, M. N., & Abdel-Daim, M. M. (2020). Multi-Target Drug Candidates for Multifactorial Alzheimer's Disease: AChE and NMDAR as Molecular Targets. *Molecular Neurobiology* 2020 58:1, 58(1), 281–303. <https://doi.org/10.1007/S12035-020-02116-9>
- Ugwu, D. I., & Conradie, J. (2023). Anticancer properties of complexes derived from bidentate ligands. *Journal of Inorganic Biochemistry*, 246. <https://doi.org/10.1016/j.jinorgbio.2023.112268>
- Uniprot. (2024). *SRC - Proto-oncogene tyrosine-protein kinase Src - Homo sapiens (Human)* | UniProtKB | UniProt. <https://www.uniprot.org/uniprotkb/P12931/entry>
- van Gerven, M., & Bohte, S. (2017). Editorial: Artificial neural networks as models of neural information processing. *Frontiers in Computational Neuroscience*, 11, 333898. <https://doi.org/10.3389/FNCOM.2017.00114/BIBTEX>
- Van, R., & Fred, L. (2009). *Python 3 Reference Manual*. CreateSpace.

- Vatansever, S., Schlessinger, A., Wacker, D., Kaniskan, H. Ü., Jin, J., Zhou, M. M., & Zhang, B. (2021). Artificial intelligence and machine learning-aided drug discovery in central nervous system diseases: State-of-the-arts and future directions. *Medicinal Research Reviews*, 41(3), 1427–1473. <https://doi.org/10.1002/MED.21764>
- Voicu, A., Duteanu, N., Voicu, M., Vlad, D., & Dumitrascu, V. (2020). The rcdk and cluster R packages applied to drug candidate selection. *Journal of Cheminformatics*, 12(1), 1–8. <https://doi.org/10.1186/S13321-019-0405-0/FIGURES/7>
- WAKABAYASHI, Y., MIYATSUKA, T., SUZUKI, L., HIMURO, M., KATAHIRA, T., MIURA, M., & WATADA, H. (2019). 346-LB: The Role of STAT3 Signaling toward α -to- β Reprogramming within the Adult Pancreas. *Diabetes*, 68(Supplement_1). <https://doi.org/10.2337/DB19-346-LB>
- Wang, H. Q., Man, Q. W., Huo, F. Y., Gao, X., Lin, H., Li, S. R., Wang, J., Su, F. C., Cai, L., Shi, Y., Liu, B., & Bu, L. L. (2022). STAT3 pathway in cancers: Past, present, and future. *MedComm*, 3(2). <https://doi.org/10.1002/MCO2.124>
- Wang, X. L., & Li, L. (2021). Cell type-specific potential pathogenic genes and functional pathways in Alzheimer's Disease. *BMC Neurology* 2021 21:1, 21(1), 1–18. <https://doi.org/10.1186/S12883-021-02407-1>
- Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60), 3021. <https://doi.org/10.21105/JOSS.03021>
- Weggen, S., & Beher, D. (2012). Molecular consequences of amyloid precursor protein and presenilin mutations causing autosomal-dominant Alzheimer's disease. *Alzheimer's Research and Therapy*, 4(2), 1–14. <https://doi.org/10.1186/ALZRT107/TABLES/3>
- WHO. (2023). *Demencia*. <https://www.who.int/es/news-room/fact-sheets/detail/dementia>
- Yang, Y., & Song, W. (2013). Molecular links between Alzheimer's disease and diabetes mellitus. *Neuroscience*, 250, 140–150. <https://doi.org/10.1016/J.NEUROSCIENCE.2013.07.009>
- Ye, Q., Hsieh, C. Y., Yang, Z., Kang, Y., Chen, J., Cao, D., He, S., & Hou, T. (2021). A unified drug–target interaction prediction framework based on knowledge graph and recommendation system. *Nature Communications* 2021 12:1, 12(1), 1–12. <https://doi.org/10.1038/s41467-021-27137-3>

- Zhang, H., Cao, Y., Ma, L., Wei, Y., & Li, H. (2021). Possible Mechanisms of Tau Spread and Toxicity in Alzheimer's Disease. *Frontiers in Cell and Developmental Biology*, 9. <https://doi.org/10.3389/FCELL.2021.707268>
- Zhao, W. Q., & Alkon, D. L. (2001). Role of insulin and insulin receptor in learning and memory. *Molecular and Cellular Endocrinology*, 177(1–2), 125–134. [https://doi.org/10.1016/S0303-7207\(01\)00455-5](https://doi.org/10.1016/S0303-7207(01)00455-5)