

Evaluación computacional del potencial de modulación por fármacos (drogabilidad) de oncoproteínas del virus de Epstein-Barr para el tratamiento del cáncer.

Tipo de apoyo: Investigador Joven

Area: II. Biología y Química

Responsable técnico: Diego Prada Gracia

Institución de adscripción: Hospital Infantil de México Federico Gómez

1. Resumen

El oncovirus de Epstein-Barr (EBV) es uno de los virus más extendidos entre los humanos. Se estima que más del 90 % de los adultos en el planeta están infectados por EBV. Este virus tiene un ciclo de vida bifásico con una etapa latente y otra lítica. Durante la etapa latente el virus permanece en forma de plásmido, doble hélice de ADN cerrada circular, replicándose junto con la célula hospedera que mayoritariamente es un linfocito B. En esta fase el virus expresa proteínas con potencial oncogénico, inflamatorio y autoinmune; pudiendo desencadenar la transformación del linfocito B en célula tumoral como ocurre en cánceres como el linfoma de Burkitt, el linfoma de Hodgkin, el carcinoma nasofaríngeo o el carcinoma gástrico. Los productos expresados por el virus en esta fase están bien determinados, especialmente las proteínas virales. Sin embargo no existe a la fecha tratamiento antiviral capaz de inhibir a EBV latente. En la última década, tras el esclarecimiento del rol de algunas de las proteínas de EBV, se han comenzado a sofisticar los proyectos de descubrimiento de potenciales inhibidores que tradicionalmente se basaban en campañas de cribado masivo experimental; con la implementación del diseño racional asistido por computadora. Sin embargo estos acercamientos, reportados tras la resolución de la estructura de varias de estas proteínas, se han realizado únicamente mediante el diseño basado en estructura. No aparece en la literatura ningún estudio que, por medio de modelos físicos de sistemas biomoleculares integrados por computadora, evalúe la potencial inhibición o modulación de la función de estas

proteínas mediante la interacción con una molécula diseñada ad hoc (drogabilidad). El presente proyecto que aquí se presenta pretende ser pionero en la caracterización de los mecanismos que determinan la drogabilidad de dichas proteínas con el propósito de ser capaces de guiar el diseño racional de antivirales de EBV.

2. Antecedentes

Diversos factores, genéticos y epigenéticos, pueden desencadenar la proliferación anómala de células transformando su ciclo de vida. Uno de estos agentes desencadenantes son los virus oncogénicos [?, ?]. Estos virus, si bien desarrollan diferentes estrategias, tienen en común que producen elementos ajenos a la célula como proteínas o fragmentos de RNA con la capacidad de transformarla en tumoral. Alrededor del 20 % de los cánceres humanos tienen alguno de los ADN oncovirus desempeñando un papel decisivo en su patogénesis [?, ?]. Este es el caso del virus ADN oncogénico herpesvirus humano 4, o como comúnmente se denomina honrando a sus descubridores: virus de Epstein-Barr (EBV).

EBV es un gamma herpesvirus B linfotrópico del que somos portadores la gran mayoría de nosotros. Se estima que más del 90 % de la población adulta mundial está infectada por este oncovirus [?]. Si atendemos a la población infantil existen diferencias geográficas debido a factores socio-económicos, demográficos y culturales. En países en vías de desarrollo el contagio suele producirse a edades muy tempranas, en la niñez, mientras que en sociedades desarrolladas dicho contagio se retrasa mayoritariamente a la adolescencia [?].

El mecanismo de infección y alojamiento en el hospedero adoptado por EBV tiene particularidades que lo hacen especialmente exitoso y que explican que sea uno de los virus más frecuentemente encontrado en humanos [?, ?, ?, ?, ?]. Este oncovirus se transmite entre individuos a través de la saliva, siendo las células epiteliales del tubo digestivo superior la puerta de entrada más frecuente. Una vez en el hospedero, EBV infecta preferentemente a linfocitos B en los que puede “residir” replicándose con ellos durante toda la vida del individuo infectado sin manifestar ni desencadenar ninguna patología, de manera asintomática. El ADN viral se conserva como episoma expresando una cantidad mínima de productos cuya función principal es desarrollar los mecanismos que aseguran su conservación, transcripción y replicación. Replicación que se produce una vez por ciclo celular mediante el acoplamiento físico del episoma al cromosoma de la célula hospedera en su metafase. Esta etapa del ciclo de infección del virus es conocida como “fase de latencia” y en ella se pueden distinguir a su vez varias subfases (0, I, II y III) caracterizadas por la presencia de algunas de las proteínas virales. Tras el contagio el virus adopta la fase de latencia III en la que se expresan los antígenos nucleares EBNA (del inglés Epstein-Barr Nuclear Antigen): EBNA1, EBNA2, EBNA3A, EBNA3B, EBNA3C y EBNA-LP; además de las proteínas de membrana LMP1, LMP2a y LMP2b, y varios miRNAs [?]. En la latencia II, I y 0 únicamente algunas de estas proteínas se expresan, siendo EBNA1 la única presente en todas las fases. Es por este motivo que EBNA1 ha atraído la atención de todo el que pretende desentrañar el mecanis-

mo molecular que posibilita que EBV permanezca tanto tiempo de manera silente en el hospedero sin levantar ninguna alarma antigénica [?, ?, ?, ?, ?, ?]. La importancia clínica de este ciclo latente reside en que los productos expresados por los genes de EBV tienen capacidades oncogénicas, inflamatorias y autoinmunes pudiendo transformar la célula hospedera infectada en tumoral [?, ?]. Desafortunadamente en la actualidad no existen fármacos con potencial inhibidor antiviral para EBV latente [?].

Tras la fase latente, y sin que haya sido esclarecido todavía el mecanismo que desencadena esta transición, el virus entra en fase lítica. En esta nueva fase, a diferencia de la anterior, EBV produce viriones con la participación de ADN polimerasa del hospedero con el objeto de infectar otras células o ser transmitidos por saliva para la infección de nuevos individuos. El mecanismo desarrollado por el virus en esta fase es completamente distinto al anterior implicando la expresión de nuevos productos. Las primeras proteínas expresadas, y a las que se les adjudica un papel decisivo en el paso latente-lítico por su participación en la expresión de posteriores productos, son las denominadas BZLF1 y BRLF1. Posteriormente se expresan otras 42 proteínas líticas [?]. Para esta fase sí contamos con antivirales capaces de atacar al virus con cierta eficacia. Es por este motivo que una de las estrategias mas exploradas para tratar la infección de EBV es la inducción de su fase lítica [?].

Se estima que EBV se encuentra asociado al origen del 2 % de los cánceres humanos [?, ?, ?]. Este oncovirus se encuentra en el origen a varios tipos de linfomas de linfocitos B y T como los linfomas de Burkitt o Hodgkin, así como los carcinomas nasofaríngeo y de estómago, y sarcomas de músculo liso. También se encuentra asociado a enfermedades autoinmunes e inflamatorias como mononucleosis infecciosa, síndrome crónico de EBV y síndrome histiocítico, esclerosis múltiple o lupus eritematoso sistémico entre otras [?, ?, ?, ?, ?]. Todo esto convierte al virus EBV en una de las principales infecciones a controlar en pacientes inmunodeprimidos evidenciando también la importancia clínica de encontrar moléculas inhibitoras de sus oncoproteínas preferencialmente en su fase latente.

Desde antes de resolver la estructura cristalizada, sola o en su interacción con ADN [?, ?], la proteína EBNA1 ha sido la principal diana del diseño de fármacos [?, ?]. Además de ser la única proteína expresada en cualquier punto del ciclo de vida bifásico del virus, la polivalencia de sus mecanismos moleculares la convierte en la piedra angular sobre la que descansa la supervivencia de EBV en latencia: realiza funciones de persistencia, replicación y transcripción del genoma. Para realizar estas funciones EBNA1 reparte sus 641 aminoácidos en distintos dominios con distinto rol. Así por ejemplo, este antígeno nuclear es capaz de unirse en distintos puntos al episoma vírico en forma de homodímero mediante su dominio de interacción a ADN, bien para contribuir en su interacción con la maquinaria proteica del hospedero a su transcripción, bien para hacer de enlace físico entre el episoma y el material cromosómico de la célula [?]. Otro dominio cuyo rol es reseñable es el dominio de repetición de pares de aminoácidos arginina-alanina. Este dominio ha demostrado ser uno de los responsables de que la célula infectada en fase latente pase desapercibida frente a linfocitos T citotóxicos CD8 [?]. Además, interacciona con proteosomas e inhibe la proteólisis de dicha proteína [?].

Tras EBNA1, otra proteína que destaca por acumular numerosos ensayos para

determinar la naturaleza de sus funciones es BZLF1. La proteína BZLF1, cuya estructura fue resuelta en su interacción con ADN en 2006 [?], es un regulador transcripcional que aparece como el primer producto que EBV expresa en su fase lítica. Este hecho la relaciona de manera reconocida en la literatura como el posible agente interruptor que desencadena el resto de acontecimientos de la fase lítica [?, ?, ?, ?].

En las últimas décadas se han llevado a cabo campañas de cribado masivo de moléculas, inicialmente experimentales, en la búsqueda de un antiviral efectivo especialmente para la fase latente [?, ?, ?, ?, ?]. Aparece en la literatura reciente un no despreciable número de moléculas encontradas *de novo* al que se unen aquellos compuestos como antivirales de alto espectro cuyo efecto ha sido probado en EBV [?, ?, ?, ?]. Casi todas estas moléculas se reportan con problemas de biodisponibilidad, toxicidad o ineficiencia. Sobre muchas de ellas se desconoce su molécula blanco y por lo tanto las características del sitio de unión como para poder emprender proyectos de optimización. El resultado de esta situación es que actualmente no contamos con moléculas capaces de actuar como antivirales en la fase latente, y los intentos experimentales de descubrimiento de moléculas inhibitoras no han contribuido a acumular conocimiento sobre los mecanismos de drogabilidad de las proteínas de EBV [?]. En cuanto a los intentos de diseño racional computacional que aparecen en la literatura, estos están todavía en fase incipiente y no abundan. La mayoría de estos intentos reportados hacen uso del diseño basado en estructura partiendo directamente de la caracterización obtenida del cristal: obviando las modificaciones de su estructura en un proceso de relajación o ignorando cualquier cambio conformacional que pueda estar en el origen del mecanismo a escala atómica de su función [?, ?, ?, ?, ?, ?, ?]. Solamente un estudio aparece apoyado en la simulación de dinámica molecular para el diseño de un inhibidor de EBNA1. Se trata de un experimento de simulación “fuerza bruta” durante un tiempo de 200 ns [?], a todas luces insuficiente para determinar la estabilidad de la unión y la naturaleza de las interacciones atómicas que más contribuyen entálpica y entrópicamente a la energía libre de unión. Otro elemento característico y común a los intentos de diseño racional de inhibidores sobre las proteínas de EBV es el abordaje del filtrado de quimiotecas únicamente mediante su interacción o docking de proteína [?, ?]. Estos métodos, lejos de evaluar la física de todos los grados de libertad del sistema obteniendo magnitudes físicas relativas a la interacción, se apoyan en argumentos heurísticos para la rápida puntuación de la idoneidad de la interacción. En algunos casos, esta puntuación está basada en la evaluación de una energía potencial de interacción de un modelo de biomoléculas reducido, introduciendo algunas interacciones electrostáticas de cargas puntuales, interacción estérica o parametrización sencilla de puentes de hidrógeno, por ejemplo. Sin embargo, estas aproximaciones nunca consideran ninguna contribución entrópica y por lo tanto no estiman las energías libres de unión de las distintas moléculas candidato, que son las que determinan cuantitativamente las probabilidades de estancia de un sistema físico en un determinado estado (en este caso “acoplados”) [?, ?]. Tampoco de esta manera se pueden validar sus predicciones dado que no hay magnitudes físico-químicas termodinámicas -relativas o absolutas- y/o cinéticas de sus modelos **de biomoléculas**

que poder comparar con experimentos in vitro [?].

Es notable que Existe una ausencia importante de estudios teórico computacionales que aborden, termodinámica y cinéticamente, el conjunto de interacciones y cambios estructurales a escala atómica que permitan determinar el potencial que tienen ciertas proteínas **de EBV, potenciales blancos para su inhibición como EBNA1 o BZLF1**, para ver alguna de sus funciones **inhibida interrumpida** o modulada mediante la interacción con moléculas diseñadas ad-hoc (drogabilidad [?]).

3. Planteamiento del problema

El oncovirus EBV se encuentra en el origen del mecanismo que desencadena diversos tipos de cáncer como el linfoma de Hodgkin o el linfoma de Burkitt, así como carcinomas nasofaríngeos o diversas enfermedades autoinmunes [?, ?, ?, ?, ?, ?]. EBV tiene la capacidad de infectar linfocitos B y en su fase latente hospedarse en la célula de manera asintomática sin más trascendencia hasta la muerte del hospedero. Pero su carga viral, **según su dependiendo del** estadio de **la** fase latente, puede presentar actividad autoinmune, inflamatoria y oncogénica. El virus tiene la capacidad de transformar la **célula** infectada tornándola inmortal [?, ?, ?, ?, ?].

Las proteínas que EBV expresa en los distintos estadios de su ciclo de vida bifásico (latente-lítico) están bien caracterizadas. Algunas de ellas, como es el caso del antígeno nuclear de Epstein-Barr 1 (EBNA1) o la proteína BZLF1, cuentan con la estructura resuelta **de sus dominios de interacción con y sin ADN** en ambos casos por difracción de rayos X **en su interacción con ADN** [?, ?, ?]. **Esta información** Esto es esencial para poder determinar, haciendo uso de modelos físicos computacionales, los mecanismos a escala atómica que describen cómo estas proteínas llevan a cabo su función mediante cambios conformacionales e interacciones con otras biomoléculas. Esta información, no **solamente** estructural sino también dinámica, permite revelar sitios de la proteína que convenientemente alterados en sus propiedades físicas pueden modular su comportamiento [?, ?]. Este efecto puede producirse por la interacción de dicha proteína con una molécula o péptido diseñados *ad hoc* con las propiedades adecuadas para inhibir su función, desestabilizar su estructura, evitar la formación de una interfase, bloquear el sitio activo en el caso de una enzima, u ocupar el sitio de unión en un receptor para desencadenar un efecto agonista u antagonista modulado en ocasiones de manera alostérica. A esta potencial modulación de la actividad de una proteína mediante moléculas diseñadas para tal efecto se le conoce en el argot como “drogabilidad” [?]. Y su evaluación es el paso previo indispensable para iniciar un proyecto de diseño y/u optimización de fármacos.

Este proyecto pretende centrarse en la evaluación de la drogabilidad de proteínas del virus EBV con objeto de determinar regiones específicas en la estructura de los blancos moleculares sobre las que comprobar la factibilidad del diseño para su inhibición. De manera más concreta, y como hipótesis inicial, se trabajará en la búsqueda de mecanismos de inhibición de EBNA1 y BZLF1 dejando abierta la posibilidad de determinar la drogabilidad de cualquier otra proteína con posible papel

oncogénico.

4. Justificación

Se estima que el 2 % de todos los cánceres se deben a la infección por el herpesvirus de Epstein-Barr (EBV) [?, ?, ?]. Está reportado que anualmente aparecen 200.000 casos de cáncer (Linfoma de Burkitt, Linfoma de Hodgkin, carcinoma nasofaríngeo, carcinoma gástrico) con evidencias de su asociación con EBV [?]. Pero las dimensiones de la relevancia clínica de este oncovirus son mayores si se tiene en cuenta que origina anualmente 125.000 nuevos casos de mononucleosis infecciosa sólo en Estados Unidos, o que aparecen en el mundo aproximadamente un millón de nuevos casos de carcinoma gástrico cada año [?]. ~~Específicamente, en el Hospital Infantil de México Federico Gómez institución vinculada al presente proyecto la Unidad de Virología y Cáncer liderada por el Dr. Ezequiel Fuentes Pananá colaborador del presente proyecto tiene consolidada una línea de investigación sobre papel de dicho virus los linfomas pediátricos atendidos en este hospital tiene contabilizado que el 70 % de los linfomas pediátricos que se atienden en el hospital son causados por este virus.~~ Muchas ~~de las patologías desencadenadas por EBV, como el caso de dichos linfomas pediátricos, de estas patologías~~ se originan con el virus en su fase de latencia. ~~Y~~ y no existen a día de hoy tratamiento antiviral o fármaco inhibidor de EBV en dicha fase [?, ?, ?]. Por contra, sí existe tratamiento para la fase lítica donde el virus requiere de la participación de la ADN polimerasa del hospedero ~~o de su cinasas~~ [?]. Pero dicho tratamiento sólo se puede aplicar cuando la infección se encuentra en pleno apogeo, lo cual es inconveniente cuando tratamos por ejemplo con pacientes inmunodeprimidos.

Adicionalmente es pertinente ~~reseñar señalar~~ que ~~este~~ proyecto de investigación es necesario y oportuno dado que no se encuentra en la literatura iniciativas de diseño racional de nuevas moléculas ~~potenciales inhibidoras de EBV~~ basadas en el estudio computacional de la termodinámica y cinética de los procesos de interacción entre biomoléculas y sus cambios conformacionales. ~~La relevancia de este hecho se puede explicar de manera sencilla con el siguiente ejemplo. La proteína EBNA1 es un blanco molecular muy atractivo cuando se trata de intentar inhibir a EBV en su fase latente, y la estructura de esta proteína ha sido resuelta hace más de 10 años. Sin embargo, el diseño racional reportado hasta ahora se ha limitado a considerar únicamente la estructura rígida cristalizada para la que se reporta, ya desde el artículo original, que en el complejo proteína-ADN se advierten posibles interacciones estabilizantes todavía no formadas. Como si el complejo hubiera encontrado una estructura estable para cristalizar previa a su completa estabilización en su unión. Sin embargo, no aparecen trabajos en los que se haya tratado de predecir cuáles son las conformaciones más estables ni cómo es el proceso de acercamiento y unión para determinar qué interacciones juegan un papel más relevante, y en qué momento, para determinar más acertadamente su posible drogabilidad. Este hecho justifica el importante componente que esta propuesta tiene de explorativa. Dada la~~

ausencia casi completa de estudios sistemáticos sobre la dinámica de estas proteínas este proyecto debe tener como una de sus principales vocaciones la acotación y distinción de qué blancos y en qué puntos son susceptibles ser drogados [?]. Esto servirá para enfocar más acertadamente posteriores proyectos de diseño y optimización racional de moléculas con potencial farmacológico como antivirales para EBV, además de darnos la oportunidad de ubicar nuestro trabajo y resultados como referencia en dicho campo.

~~Este hecho nos da la oportunidad de aprovechar el presente proyecto, además de para iniciar una prometedora línea de investigación en la nueva Unidad de Investigación en Biología Computacional y Diseño de Fármacos del Hospital Infantil de México, para ubicar nuestro trabajo y resultados como referencia en el campo del diseño de antivirales de EBV.~~

~~Por último se hace necesario justificar el componente importante que este proyecto tiene de explorativo. Dado el vacío que hay en estudios sistemáticos sobre la dinámica de estas proteínas, sus cambios conformacionales o sus propiedades termodinámicas y cinéticas, este proyecto debe tener como una de sus principales vocaciones la acotación y distinción de qué blancos y en qué puntos son susceptibles ser drogados. Esto servirá para enfocar más acertadamente posteriores proyectos de diseño y optimización racional de moléculas con potencial farmacológico como antivirales para EBV.~~

5. Hipótesis

El conjunto de oncoproteínas expresadas en las fases de latencia de EBV, como EBNA1, o aquellas necesarias para su tránsito de la fase latente a la forma lítica, como BZLF1, pueden constituir unos blancos moleculares idóneos para el diseño de inhibidores con potencial farmacológico sobre aquellos tipos de cáncer inducidos por EBV. Los modelos físicos de biomoléculas y su simulación pueden aportar las claves termodinámicas y cinéticas necesarias para evaluar eficientemente su potencial drogabilidad y avanzar en el diseño de antivirales para EBV.

6. Objetivos

El objetivo general de este proyecto está descrito con claridad desde su hipótesis principal.

6.1. Objetivo general

Evaluar el potencial de modulación que mediante **la** interacción con pequeños ligandos o péptidos puede presentar la función de proteínas constitutivas del virus EVB como EBNA1 o BZFL1.

6.2. Objetivos particulares

- Detectar y evaluar nuevos sitios drogables **principalmente** sobre proteínas cuya inhibición puede jugar un papel decisivo en el diseño de antivirales para EBV.
- Validar hipótesis de mecanismos de acción a escala atómica de proteínas relevantes cuya función es decisiva para EBV tanto en la fase latente como lítica. ~~Ejemplo: describir el proceso de unión de EBNA1 al episoma de EBV calculando la diferencia de energía libre entre EBNA1 en solvatación y EBNA1 en su interacción con DNA; ratificar o rechazar las hipótesis que aparecen en la literatura sobre el papel del homodímero, los mecanismos de selectividad sobre segmentos de pares de bases específicos y el efecto colectivo de unión en sitios próximos como los presentados en la región denominada “Dyad Symmetry”.~~
- Recopilar de la literatura **la** quimioteca de moléculas cuya posible actividad antiviral sobre EBV ya ha**ya** sido reportada aun sin conocer su posible mecanismo de acción. ~~Construir el mínimo conjunto de farmacóforos que represente dicha quimioteca para comprobar su posible encaje sobre los sitios drogables detectados. Identificar además, mediante el mínimo conjunto de farmacóforos que la representen, el posible encaje de estos ligandos sobre los sitios drogables detectados.~~
- Diseñar, sin atender a criterios de reglas de Lipinski como la apropiada solubilidad, biodisponibilidad o la alta selectividad o baja toxicidad, un conjunto de moléculas o péptidos **cuya teórica** alta afinidad por los nuevos sitios drogables detectados **pueda en el futuro servir como base para el posible diseño de un inhibidor.** ~~a lo largo del proyecto cuya constante de disociación, entropía y energía libre de unión calculados en los modelos físicos puedan ser en el futuro corroborados y validados mediante el uso de la microcalorimetría de titulación isoterma en una posible extensión de este proyecto.~~
- ~~Diseñar, sin atender a criterios de reglas de Lipinski como la apropiada solubilidad, biodisponibilidad o la alta selectividad o baja toxicidad, un conjunto de moléculas o péptidos con alta afinidad por los nuevos sitios drogables detectados a lo largo del proyecto cuyo efecto inhibitorio pueda ser comprobado y caracterizado de manera experimental mediante el uso de la microcalorimetría de titulación isoterma en una posible extensión de este proyecto.~~
- ~~Consolidar el arranque de una importante línea de investigación definiendo sobre criterios sólidos el comienzo del camino hacia el diseño y optimización de antivirales de EBV para el tratamiento de los distintos tipos de cáncer asociados a este virus.~~
- ~~Diseñar mediante la evaluación de la posible drogabilidad de proteínas de EBV experimentos, como por ejemplo la inhibición de interacciones~~

~~con proteínas del hospedero, que arrojen evidencias sobre los mecanismos oncogénicos, inflamatorios o autoinmunes que la infección puede desencadenar.~~

7. Metodología

Tres herramientas computacionales se emplearán adecuadamente según la naturaleza y función de la proteína blanco para la evaluación de nuevos sitios drogables: el docking sobre proteína rígida de librerías de compuestos para el cribado virtual masivo inicial (High Throughput Screening) [?, ?], la simulación de dinámica molecular a todos los átomos con solvente explícito o implícito (según convenga) [?] y los métodos de evaluación de energía libre, determinación del paisaje conformacional de energía libre y métodos de perturbación de energía libre [?]. Estos abordajes serán aplicados sobre distintas proteínas constitutivas de EBV teniendo como principales candidatas el antígeno nuclear 1, EBNA1, única proteína presente en todas las fases de latencia del virus, y BZLF1, regulador transcripcional decisivo en el paso de la forma latente a lítica. Por ejemplo, en el caso de proteínas cuya drogabilidad no ha sido previamente explorada, como BZLF1, se comenzará aplicando abordajes clásicos heurísticos de filtrado de quimiotecas mediante docking sobre proteína rígida para revelar posibles sitios de unión [?]; a la vez que paralelamente se lleva a cabo una búsqueda racional de regiones de interfase con aquellas moléculas con las que interaccionen en su actividad, como p53 o ADN en este caso, susceptibles de ser empleadas como posibles sitios drogables [?]. Por contra, en el caso de una proteína como EBNA1 cuya interacción en su unión con ADN ha sido recientemente empleada sin éxito para el diseño de inhibidores, se llevarán a cabo estudios de dinámica molecular para por primera vez caracterizar los cambios conformacionales en el proceso de unión y las interacciones inter-atómicas más relevantes [?], así como estudios de perturbación de energía libre cuya aplicación por definición resulta más precisa, aunque más costosa computacionalmente, para la predicción de constantes de afinidad molécula blanco-ligando [?, ?].

Se hace necesario puntualizar que aunque estas metodologías emplean el uso masivo de computadoras para la simulación de la física del sistema biomolecular, estas no han de ser confundidas con las metodologías propias del campo de la bioinformática. ~~Es cierto que~~ Ambas disciplinas comparten el uso de herramientas estadísticas para el análisis de datos como el análisis bayesiano [?, ?], los modelos de Markov [?], el análisis multivariante [?] o la potente batería de herramientas propias de la física de sistemas complejos como el uso de teoría de redes complejas o teoría de la información [?]. La metodología a usar aquí está fundamentada en las leyes ~~de la~~ física que ~~dominan la realidad del~~ describen y predicen el comportamiento de las interacciones de biomoléculas a escala de energía del orden de $K_B T$, tiempos del orden de 10^{-12} segundos y distancias en el rango de 10^{-10} metros, posibilitando su simulación. Dicha simulación, ~~o integración,~~ se hace mediante el uso de computadoras debido a la ~~alta dimensionalidad del sistema:~~ alta complejidad de las ecuaciones a resolver. ~~Pero dichas ecuaciones pertene-~~

~~cen a modelos físicos fundamentados para representar y predecir fenómenos naturales a escala de energía del orden de $k_B T$, tiempos del orden de femtosegundos y distancias en el rango de 10^{-10} metros.~~ Esta física está bien consolidada, ~~en especial la que a procesos clásicos en equilibrio se refiere,~~ desde los trabajos de Newton, Einstein (descripción browniana), Fokker y Planck o Boltzman, ~~etc...~~, **en especial la que a procesos clásicos en equilibrio se refiere.** Así por ejemplo los modelos que se emplearán de biomoléculas en este proyecto cumplen con la descripción de procesos estocásticos, la física atómica, la física cuántica, la ley de fluctuación-disipación, la termodinámica clásica y la mecánica estadística, entre otros. **Es posible que** Debido a la dimensión de los sistemas a estudiar, **cuando** las magnitudes de equilibrio termodinámico no puedan ser muestreadas convergentemente mediante las técnicas de simulación habituales ~~Para determinar de manera económica las propiedades de altas barreras de energía libre~~ se hará uso de técnicas de simulación de procesos de no equilibrio para inferir información de equilibrio (desigualdad de Jarzinsky y teorema de Crooks) [?], así como de estrategias de reescalado del sistema como la simulación con campos de fuerza coarse-grained [?].

Las herramientas de bioinformática (minería de datos) y química computacional serán usadas también para dar soporte al proyecto en aspectos como ~~Por ejemplo, en~~ la búsqueda de interacciones proteína-proteína posiblemente descritas ya en la literatura [?].

8. Plan de análisis de los datos

Dado que se trata de un proyecto teórico el plan de análisis de datos, que no se puede anticipar en su totalidad, está descrito implícitamente en la metodología. En el caso de las simulaciones de dinámica molecular se llevarán acabo análisis basados en redes cinéticas [?] y modelos de estados de markov [?]. En las experiencias de docking [?] o perturbación de energía libre [?] se emplearán técnicas de clustering para la estimación de la contribución entrópica además del cálculo entálpico [?]. Además, se hará uso de la batería de herramientas de análisis mencionadas en la sección anterior como el análisis bayesiano [?], el análisis multivariante [?] o las redes complejas [?]. Adicionalmente, como en todo proyecto físico-teórico cabe la posibilidad de diseñar, según el rumbo del proyecto, análisis con fundamento físico y estadístico adecuados para responder a las distintas preguntas puntuales que surjan en el transcurso **del proyecto de la investigación.**

9. Limitaciones del estudio

La principal limitación con la que cuenta el proyecto se debe a los insuficientes recursos computacionales con los que se cuenta en la actualidad en la Unidad de Biología Computacional y Diseño de Fármacos del Hospital Infantil de México. La simulación de un sistema que incluye un segmento de ADN además de una proteína

de más de 600 aminoácidos, como EBNA1, debidamente solvatados en presencia de cationes y aniones requiere de un poder computacional mayor al del cluster de cinco nodos **del que dispone nuestra unidad**. Lo mismo sucede con el cribado **HTS** de una gran quimioteca considerando la proteína rígida, o el cribado **MTS** de pocos ligandos mediante evaluación de energía libre de unión. Se requiere de la participación de un volumen mayor de recursos computacionales. Además, la dotación económica que supuestamente se destina para soportar el desarrollo de este proyecto de investigación no puede ser destinada a la compra del equipo de laboratorio computacional necesario **sin duda** para **trabajar** este tipo de proyectos de investigación.

Para superar esta limitación se buscará la manera de acceder a tiempo de computación en alguno de los tres nodos de supercomputación del Laboratorio Nacional de Computo de Alto Desempeño, LANCAD, de México.

En caso de no disponer del poder computacional adecuado se comprobará la validez de ciertas aproximaciones al modelo físico de sistema biomolecular con el fin de reducir su complejidad a costa de perder detalle fino en la descripción y poder ser utilizadas con el propósito de alcanzar los objetivos descritos en esta propuesta. Estas técnicas pueden ser el uso de solvente implícito **[?]**, **el uso de estrategias de escalado dimensional**, métodos de integración de dinámica estocástica *biased* sobre el espacio reducido de grados de libertad mas lentos **[?]**, el diseño de experimentos computacionales de no equilibrio **[?]** o el uso de campos de fuerza coarse-grained **[?]**.

10. Consideraciones éticas

Dentro de las distintas categorías de riesgo según los agentes, organismos o población manejada en el estudio podemos catalogar el presente proyecto de investigación de “Sin Riesgo”. El proyecto es puramente teórico. No se realizarán en ningún momento experimentos físicos -que no sean computacionales-, químicos o biológicos. Tampoco el objeto de estudio es un segmento de población humana. No procede por tanto para este proyecto hacer ninguna consideración ética relevante o incluir carta de confidencialidad, compromiso de confidencialidad de datos personales o protocolo de actuación frente a efectos adversos. *Ver Formatos Adjuntos*.

11. Consideraciones de bioseguridad

Al igual que en la sección anterior. Se puntualiza de nuevo que este proyecto de investigación es puramente teórico. No se realizarán en ningún momento experimentos físicos -que no sean computacionales-, químicos o biológicos. No se manejarán muestras biológicas, agentes patógenos, material radiactivo o productos tóxicos. No es pertinente entonces incluir ninguna consideración específica relativa a bioseguridad. *Ver Formatos Adjuntos*.

12. Anexos

Se adjuntan al final de la propuesta y como anexos fuera de índice los siguientes documentos: “Solicitud de evaluación de proyecto”, “Formato del comité de ética en investigación”, “Formato del comité de bioseguridad”, “Formato PE-04”, “Formato de distribución de asignación anual” y “Justificación de distribución de asignación anual”.

13. Cronograma

El desarrollo del proyecto de investigación aquí propuesto se llevará a cabo en el transcurso de 8 trimestres de la siguiente manera:

- *RevLit*: Revisión de literatura, recolección de ligandos reportados con potencial antiviral y ejercicio de discusión, diseño y evaluación de las estrategias a adoptar.
- *RepRev*: Escritura de artículo de revisión enfocado hacia el diseño racional asistido por computadora de inhibidores de EBV.
- *EvalProts*: Evaluación de idoneidad de proteínas expresadas por EBV distintas a EBNA1 y BZLF1 para la evaluación de su drogabilidad.
- *SimuEBNA1*: Experimentos computacionales de EBNA1.
- *AnalEBNA1*: Análisis de experimentos computacionales de EBNA1.
- *SimuBZLF1*: Experimentos computacionales de BZLF1.
- *AnalBZLF1*: Análisis de experimentos computacionales de BZLF1.
- *SimuOtras*: Simulación de posibles proteínas alternativas seleccionadas en la etapa *EvalProts*.
- *AnalOtras*: Análisis de experimentos computacionales de *SimuOtras*.
- *EvalLigsLit*: Evaluación sobre las proteínas abordadas y sus sitios drogables de la colección de ligandos obtenida en la etapa *RevLit*.
- *RepRes1*: Escritura de comunicación de lo aprendido con EBNA1.
- *RepRes2*: Escritura de comunicación de lo aprendido con BZLF1.
- *RepRes3*: Escritura de comunicación de lo aprendido con las posibles proteínas seleccionadas en la etapa *EvalProts*.

El programa es ambicioso pero su estructura permite el alcance de resultados intermedios independientes. El cumplimiento de 2/3 de los resultados planteados se puede considerar un desarrollo exitoso.

	Año 1				Año 2			
	1	2	3	4	5	6	7	8
<i>RevLit</i>	■							
<i>RepRev</i>		■						
<i>EvalProts</i>		■						
<i>SimuEBNA1</i>			■	■				
<i>AnalEBNA1</i>				■	■			
<i>SimuBZLF1</i>				■	■			
<i>AnalBZLF1</i>					■	■		
<i>SimuOtras</i>					■	■		
<i>AnalOtras</i>						■	■	
<i>EvalLigsLit</i>								■
<i>RepRes1</i>					■	■		
<i>RepRes2</i>						■	■	
<i>RepRes3</i>							■	■

14. Productos a obtener del estudio

14.1. Formación de recursos humanos

Se graduará durante el segundo año a un alumno de licenciatura.

14.2. Comunicación de resultados

14.2.1. Comunicación en congresos

Se buscará participar a lo largo del segundo año en un congreso nacional con una contribución de tipo oral más una de tipo poster además de en un congreso internacional con una contribución de tipo oral más una de tipo poster.

14.2.2. Comunicación en artículos

Se publicarán a lo largo de los dos años, y como está programado en el cronograma, 4 artículos en revistas internacionales indexadas de alto impacto: 1 artículo de revisión y 3 artículos de investigación original.

14.3. Utilidad del trabajo de investigación

El proyecto está dirigido a obtener la información relevante que nos permita **determinar las regiones más apropiadas de las proteínas constitutivas de EBV susceptibles de ser “drogadas” con un posible efecto final inhibitorio. Estos resultados son esenciales para sentar las bases** de un proyecto de diseño y optimización de moléculas con potencial farmacológico como antivirales de EBV.

15. Desglose de presupuesto

Ver Formatos Adjuntos.