
ESTUDIO DE INTERACCIONES ENTRE FÁRMACOS Y PROTEÍNAS DEL SISTEMA MONOAMINÉRGICO Y/O RECEPTORES NICOTÍNICOS UTILIZANDO TÉCNICAS DE MINERÍA DE DATOS

**CAMILA IGNACIA LEÍVA CONTRERAS
INGENIERO EN BIOINFORMÁTICA**

RESUMEN

El estudio de interacciones entre fármacos y proteínas es de gran importancia para el diseño racional de fármacos, en donde la predicción computacional de las mismas es uno de los mayores desafíos. En particular proyectos actuales están buscados nuevos fármacos candidatos para el tratamiento de enfermedades relacionadas con el sistema nervioso central, estudiando las interacciones existentes entre proteínas del sistema monoaminérgico y/o receptores nicotínicos de acetilcolina con diversos fármacos candidatos.

Durante este proceso se comprueba la factibilidad de estos posibles candidatos en una fase modelada a través de computador antes de hacer pruebas experimentales con los fármacos putativos. La etapa de modelado por computador conlleva dos estudios enlazados: el primero es *virtual screening* en donde distintos fármacos son probados con una misma proteína de forma de encontrar aquellos posibles complejos que presenten una buena estabilidad energética en un cierto instante, es decir con una conformación fija tanto para la proteína como para el fármaco. Posterior a esto se debe verificar la estabilidad del posible complejo fármaco-proteína encontrado a través de un estudio de *dinámica molecular*, en donde se comprobará la interacción del complejo a lo largo del tiempo.

En la práctica estos estudios han mostrado que los complejos seleccionados en la etapa de *virtual screening* muestran una baja aprobación en posteriores estudios de *dinámica molecular*. Sólo un 30% de los complejos encontrados en la primera etapa cumplen con el criterio de estabilidad obtenido en el posterior estudio de *dinámica molecular*. Considerando que un estudio de *dinámica molecular* puede durar hasta 1 semana, existe una gran cantidad de tiempo perdido durante esta parte del proceso de diseño de fármacos. La presente memoria de título se hace cargo de esta problemática. Para esto se ha desarrollado un modelo de inferencia, basado en técnicas de minería de datos, capaz de predecir si un complejo fármaco-proteína, de la familia monoaminérgica y/o receptores nicotínicos, seleccionado en una etapa de *virtual screening*, se comportará de forma estable

durante un estudio posterior de *dinámica molecular*. El modelo utiliza para su entrenamiento diversas características estructurales y propiedades fisicoquímicas de los complejos fármaco-proteína estudiados.

Alcanzando una eficiencia medida como tasa de verdaderos positivos de sobre el 80% (esto es identificar complejos estables) y una tasa de falsos positivos cercana al 10% (esto es aceptar erróneamente complejos no-estables).

ABSTRACT

The study of interactions between pharmaceutical drugs and proteins is quite significant for the rational pharmaceutical drug design, where the computational prediction of these is one of the largest challenges. In particular, current projects are seeking new pharmaceutical drug candidates for the treatment of diseases related to the central nervous system, studying the interactions between monoaminergic system proteins and/or nicotinic acetylcholine receptors with various pharmaceutical drug candidates.

During this process the feasibility of these potential candidates is tested in a computerized modeled phase before making experimental tests with the putative pharmaceutical drugs. The stage of computerized modeling involves two studies linked: the first one is a *virtual screening* where different pharmaceutical drugs are tested using the same protein in order to find those potential complexes which have good energy stability at a certain moment, ie with a stable conformation for both the protein and the pharmaceutical drug. After this, it must be verified the stability of possible pharmaceutical drug-protein complex found through a study of *molecular dynamics*, where complex interaction will be checked over time.

In practice, these studies have shown that chosen complex in the *virtual screening* stage show low approval in subsequent *molecular dynamics* studies. Only 30% of the complexes in the first stage meet the stability criterion obtained in subsequent molecular dynamics study. Considering that a study of *molecular dynamics* can last up to 1 week, there is a lot of wasted time during the pharmaceutical drug design part of the process.

The dissertation herein deals with this problem. For this, it has been developed an inference model, based on data mining techniques, capable of predicting whether a pharmaceutical drug-protein, from monoaminergic family and/or nicotinic receptors, selected at a *virtual screening* stage, will behave in a stable way during a subsequent study of *molecular dynamics*. The model uses, for training, various structural and physicochemical properties of the complex pharmaceutical drug-protein studied. Achieving efficiency as a measure of true positive rate of over 80% (ie this is to identify stable complexes) and a false positive rate close to 10% (ie this is to accept non-stable complexes erroneously).