

Identificación sistemática de motivos lineales de interacción con la familia de proteínas *pocket*.

Tesista: Carla Lorenze

Directora: Dra. Lucía B. Chemes

Co-Directora: Dra. Juliana Glavina

Los motivos lineales (SLiMs) son elementos modulares cortos (~6 residuos) localizados en regiones intrínsecamente desordenadas de las proteínas y median interacciones proteína-proteína. La familia de proteínas *pocket* incluye a Retinoblastoma, p107 y p130; involucradas en la regulación del ciclo celular y su inactivación puede dar lugar a procesos oncogénicos. A través de su dominio *pocket* conservado interactúan con los SLiMs E2F y LxCxE presentes en sus blancos proteicos. Se realizó un ensayo ProP-PD (Proteomic Peptide Phage Display) utilizando los dominios *pocket* como carnada y una biblioteca de péptidos de 16 residuos provenientes de regiones desordenadas del proteoma humano (HD2) obteniendo una lista de más de 1000 péptidos *hit*. El objetivo general de este trabajo es caracterizar estructuralmente péptidos *hit* utilizando herramientas bioinformáticas que permitan identificar y priorizar nuevos SLiMs E2F y LxCxE para validar experimentalmente.

En primer lugar, se evaluó la calidad del ensayo de Pro-PD. Para esto se calculó el *recall*, que permite evaluar la recuperación de interactores previamente validados que poseen los SLiMs E2F o LxCxE (True Positives, TP) e interactores conocidos pero que se desconoce si poseen o no un SLiM E2F o LxCxE reportados en la base de datos IntAct. Luego, se detectaron los SLiMs utilizando sus expresiones regulares y se construyeron logos de secuencia para evaluar la variabilidad de secuencia en los péptidos recuperados. La caracterización estructural se realizó utilizando el predictor de desorden IUPred, modelos AlphaFold2 para evaluar el grado de exposición al solvente (RSA) y se analizó el solapamiento con dominios Pfam. Con estos tres parámetros estructurales, se estableció una estrategia de filtrado para priorizar candidatos. Por último, se evaluó la estabilidad energética de péptidos *hit* utilizando matrices FoldX construidas a partir de una estructura resuelta del dominio *pocket* de Rb y péptidos con el motivo LxCxE o E2F. Para la priorización además, se colectó la localización celular de los péptidos.

En primer lugar, se obtuvieron valores de *recall* para los TP (35%) y para interactores reportados en IntAct (9%) mayores a los reportados para este tipo de ensayos. En segundo lugar, el 48% de péptidos *hit* poseía al menos uno de los SLiMs LxCxE o E2F. Estos resultados indican que el ensayo de ProP-PD es de buena calidad. En cuanto a los parámetros estructurales se observa que más del 55% de los péptidos *hit* pertenecen a proteínas con un alto grado de desorden predicho por IUPred, más del 70% se encuentran en regiones accesibles de la proteína y el 70% no se superpone con un dominio Pfam. Por otro lado, FoldX permitió identificar péptidos energéticamente estables. Finalmente, todos los parámetros fueron evaluados en conjunto con la localización celular para obtener una lista priorizada de péptidos candidatos a futuros ensayos experimentales que se realizarán en el laboratorio de trabajo.

Palabras clave: proteínas *pocket*, motivos lineales, interacciones proteína-proteína, ProP-PD.

Keywords: pocket proteins, SLiMs, protein-protein interaction, ProP-PD