



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA
Unidad Cuajimalpa



UNIVERSIDAD AUTÓNOMA METROPOLITANA

UNIDAD CUAJIMALPA

POSGRADO EN CIENCIAS NATURALES E INGENIERÍA

Tema de ICR de Maestría:

*“ESTUDIO BIOINFORMÁTICO DE LAS CARACTERÍSTICAS
ESTRUCTURALES DE LA INTERACCIÓN PROTEÍNA-PROTEÍNA
ENTRE LAS PROTEÍNAS DE CHOQUE TÉRMICO 70 DE HUMANO
(HSP70) Y PROTEÍNAS BLANCO”*

PRESENTA:

BIOL. MOL. LUIS ÁNGEL RODRÍGUEZ GARCÍA

CO DIRECTORES: DRA. CYNTHIA GABRIELA SÁMANO SALAZAR

DR. GERARDO PÉREZ HERNÁNDEZ

CUAJIMALPA DE MORELOS, CIUDAD DE MÉXICO, 9 DE MAYO DE 2022



Esta idónea comunicación de resultados fue realizada dentro del Programa de Maestría del **Posgrado en Ciencias Naturales e Ingeniería** de la División de Ciencias Naturales e Ingeniería (DCNI) de la Universidad Autónoma Metropolitana- Unidad Cuajimalpa. El trabajo teórico fue realizado de diciembre-2019 a mayo-2022 en el DCN de la DCNI.

DECLARACIÓN DE CESIÓN DE DERECHOS

En la Ciudad de México, D. F. el día 9 del mes de mayo del año 2022, el que suscribe **Luis Ángel Rodríguez García** alumno del Programa de Maestría en Ciencias Naturales e Ingeniería de la División de Ciencias Naturales e Ingeniería de la Universidad Autónoma Metropolitana-Unidad Cuajimalpa, manifiesta que es autor intelectual de la presente idónea comunicación de resultados titulada; "Estudio bioinformático de las características estructurales de la interacción proteína-proteína entre las proteínas de choque térmico 70 de humano (Hsp70) y proteínas blanco" realizada bajo la dirección de la Dra. Cynthia Gabriela Sámano Salazar y el Dr. Gerardo Pérez Hernández y cede los derechos de este trabajo a la Universidad Autónoma Metropolitana (UAM) para su difusión con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráfico o de datos del trabajo, sin el permiso expreso de los directores del trabajo como representantes de la UAM. Este puede ser obtenido escribiendo a la siguiente dirección: (csamano@cua.uam.mx o gperezh@cua.uam.mx)

Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.



Biol. Mol. Luis Ángel Rodríguez García

DECLARACIÓN DE ORIGINALIDAD

“El que suscribe, Luis Ángel Rodríguez García, alumno del Programa de Maestría en Ciencias Naturales e Ingeniería declara que los resultados reportados en esta idónea comunicación de resultados, son producto de mi trabajo con el apoyo permitido de terceros en cuanto a su concepción y análisis. Así mismo, declaro que hasta donde es de mi conocimiento no contiene material previamente publicado o escrito por otra (s) persona (s) excepto donde se reconoce como tal a través de citas y que este fue usado con propósitos exclusivos de ilustración o comparación. En este sentido, afirmo que cualquier información sin citar a un tercero es de mi propia autoría. Declaro, finalmente, que la redacción de este trabajo es producto de mi propia labor con la dirección y apoyo de mi director y de mi comité tutorial, en cuanto a la concepción del proyecto, al estilo de la presentación y a la expresión escrita.”



Biol. Mol. Luis Ángel Rodríguez García

DECLARACIÓN DE NO LUCRO:

El que suscribe, Luis Ángel Rodríguez García alumno del Programa de Maestría en Ciencias Naturales e Ingeniería, manifiesta su compromiso de no utilizar con fines de difusión, publicación, protección legal por cualquier medio, licenciamiento, venta, cesión de derechos parcial o total o de proporcionar ventajas comerciales o lucrativas a terceros, con respecto a los materiales, datos analíticos o información de toda índole, relacionada con las actividades e intercambios de información derivados de la relación de investigación académica y tecnológica desarrollada entre la Universidad Autónoma Metropolitana (UAM) y Luis Ángel Rodríguez García.



Biol. Mol. Luis Ángel Rodríguez García

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada durante los dos años de la realización de este proyecto. CVU: 1042616.

A la Dra. Cynthia Gabriela Sámano Salazar por la confianza en mi depositada, por su orientación, por todo su apoyo y paciencia desde la licenciatura y a lo largo del posgrado.

Al Dr. Gerardo Pérez Hernández por todo su apoyo y el tiempo dedicado en el proyecto, sin su participación no hubiera sido posible.

A los revisores: Dra. Elizabeth Ortiz Gutiérrez, Dr. Rafael Arturo Zubillaga Luna y Dr. Ponciano García Gutiérrez quiénes con su apoyo y valioso conocimiento se puedo concluir el proyecto.

Agradezco a la planta académica del Posgrado en Ciencias Naturales e Ingeniería por sus enseñanzas que tanto aportaron a mi formación.

A mis padres, por motivarme a seguir preparándome, por su constante apoyo y porque su sacrificio me permitió llegar a esta instancia.

A mis hermanas por apoyarme en todo momento y mostrarme que la perseverancia y el esfuerzo son la clave para lograr objetivos.

A mi esposa, por impulsarme constantemente a ser mejor persona, por su apoyo y amor incondicional que me permiten seguir adelante.

MUCHAS GRACIAS

Contenido

Palabras clave.....	8
Abreviaturas	8
Resumen.....	9
1. Introducción	10
2 Marco teórico.....	12
2.1 Chaperonas moleculares	12
2.2 Proteína de choque térmico 70 (Hsp70) como modelo de estudio	15
2.3 Unión de la Hsp70 con diversas proteínas.....	19
2.4 Proteínas intrínsecamente desordenadas.....	20
2.5 La Hsp70 y proteínas intrínsecamente desordenadas	22
2.6 Herramientas bioinformáticas para el análisis de las interacciones entre proteínas.....	22
3. OBJETIVOS	24
4. HIPÓTESIS	25
5. METODOLOGÍA.....	26
5.1 Búsqueda de proteínas blanco o “clientes” de la Hsp70	27
5.2 Elaboración de la base de datos	27
5.3 Predicción de sitios de unión de la Hsp70 en posibles zonas de estructura desplegada de los clientes.....	28
5.4 Identificación de regiones intrínsecamente desordenadas (IDPR).....	28
5.5 Modelado de proteínas	29
5.6 Docking	30
5.7 identificación de contactos: <i>Fingerprint</i>	32
5.8 Análisis de los contactos.....	33
6. Resultados	34

6.1 Identificación de proteínas clientes de la Hsp70	34
6.2 Determinación de sitios de unión con la Hsp70 e identificación de IDPRs ..	36
6.3 Modelado de proteínas	38
6.4 Residuos de referencia del SBD de la Hsp70	40
6.6 Energía electrostática en la interacción Hsp70-cliente.....	45
6.7 Determinación del <i>fingerprint</i> de la interacción Hsp70-sustrato	47
7. Conclusiones.....	53
8. Perspectivas.....	54
Referencias bibliográficas	55

Palabras clave

Chaperonas moleculares, cochaperonas, docking, estrés oxidativo, estrés térmico, fingerprint, proteínas de choque térmico (Hsp).

Abreviaturas

Dominio de unión a nucleótidos (NBD)

Dominio de unión a sustrato (SBD)

Proteína de choque térmico de 70 kDa (Hsp70)

Proteínas intrínsecamente desordenadas (PIDs)

Regiones proteicas intrínsecamente desordenadas (IDPRs)

Sistema de ubiquitina-proteasoma (UPS)

Resumen

Las proteínas desempeñan diversas funciones en diversos procesos biológicos. Para que estas funcionen adecuadamente es necesario que se plieguen y mantengan una estructura tridimensional, sin embargo, diferentes condiciones de estrés pueden poner en riesgo a las proteínas y llevarlas a perder parcial o totalmente su estructura. Las proteínas de choque térmico (Hsp) pertenecen a la familia de las chaperonas moleculares, las cuales se encargan de mantener la homeostasis de las proteínas. La Hsp70 es una de las proteínas más estudiadas de esta subfamilia, se caracteriza por ser una proteína muy conservada en los organismos, cuyas actividades están relacionadas con el plegamiento, replegamiento y desagregación de diversas proteínas. En este proyecto de investigación nos enfocamos en el estudio de la Hsp70 de humano y su capacidad de interacción con otras proteínas blanco también denominadas clientes. Por medio de herramientas bioinformáticas, analizamos los residuos de algunas proteínas cliente de la Hsp70, con la finalidad de identificar si hay similitud de los sitios de interacción de la Hsp70 con ciertas regiones de las proteínas cliente. A partir de una búsqueda en UniProt, identificamos diferentes proteínas que se han reportado que interactúan con la Hsp70, los resultados de evaluar la interacción proteína-proteínas al acoplar la Hsp70 con diferentes clientes, muestra que los residuos involucrados en la interacción son principalmente aminoácidos con carga positiva y negativa. Las regiones de las proteínas cliente a las que se une la Hsp70 son segmentos que no forman α -hélices ni hojas β , es decir se unen a loops o regiones poco estructuradas. Tomando en cuenta lo anterior, se modelaron las regiones intrínsecamente desordenadas (IDPRs) en las estructuras tridimensionales de las proteínas cliente previamente reportadas, ya que estas no estaban presentes en las estructuras depositadas en el Protein Data Bank. Finalmente, observamos que la interacción de la Hsp70 con las IDPRs de los clientes no correlaciona cómo se esperaba. Lo anterior nos lleva a considerar que la interacción de la Hsp70 con el conjunto de proteínas estudiadas no se debe al mal plegamiento, sino que puede que esté involucrada la función de holdasa, sin embargo, no podemos asegurarlo. Además, encontramos que el tipo de aminoácidos cargados y por ende las interacciones electrostáticas son determinantes en la interacción de la Hsp70 con sus diferentes sustratos.

1. Introducción

Las proteínas son macromoléculas que desempeñan un papel central en todos los seres vivos, ya que están involucradas en casi todos los procesos biológicos de los organismos. En el cuerpo humano se han identificado aproximadamente 50,000 clases diferentes de proteínas sintetizadas en los ribosomas con longitudes de hasta miles de aminoácidos (Tunyasuvunakool et al., 2021). La mayoría de las proteínas recién sintetizadas deben plegarse en una estructura tridimensional necesaria para realizar su función. Por naturaleza las proteínas son estructuralmente dinámicas, esto es que pueden adoptar diferentes conformaciones tridimensionales lo que les permite realizar múltiples funciones. Aunque su entorno comúnmente es adecuado para que puedan desempeñar su función, ciertas condiciones pueden suponer un riesgo, por ejemplo, un exceso de la concentración de las proteínas puede inducir agregación de las mismas; cambios de pH, temperatura y la radiación UV, pueden comprometer la estructura tridimensional funcional de las proteínas a tal grado que pueden llegar a desplegarse parcial o totalmente, lo que conlleva a que deban ser procesadas por los sistemas de reciclaje celular. Aquellas proteínas que pierden su estructura tienden a entrelazarse con otras proteínas, formando agregados inespecíficos (Buchner, 2019; Ma et al., 2020).

La adaptación molecular de las proteínas para que cumplan su función ocurre de diversas maneras. Por una lado se debe considerar la parte intrínseca en la que juega un papel fundamental el balance entre la hidrofobicidad y la distribución de cargas en su secuencia de aminoácidos, de tal manera que permita cierta flexibilidad durante su función. Por otro lado, existen otros factores extrínsecos en los que intervienen la interacción de dos o más proteínas o subunidades, ya sea para cumplir un proceso de comunicación entre ellas o para coordinar un control de calidad en su papel funcional dentro de la célula. Por ejemplo, las chaperonas moleculares que actúan en diversos mecanismos moleculares que contribuyen al mantenimiento de la homeostasis de las proteínas o proteostasis (Buchner, 2019; Macošek et al., 2021).

Las chaperonas se asocian a diferentes factores estresantes y las más conocidas son las proteínas de choque térmico (Hsp, por sus siglas en inglés) las cuales fueron descritas inicialmente tras notar que su expresión aumentaba después de inducir un aumento de temperatura en células de *Drosophila* (Ritossa, 1962). Las Hsp se dividen en diferentes subfamilias según su tamaño que va de 20 a 110 kDa. Entre las Hsp más estudiadas se encuentran las Hsp70, una subfamilia muy conservada en todos los organismos, y de las más versátiles. En la célula eucariota en condiciones fisiológicas, se expresan niveles basales de Hsp70, ya que participan en una amplia gama de actividades, incluyendo el plegamiento de proteínas

nacientes, la translocación de polipéptidos en mitocondrias, cloroplastos y el retículo endoplásmico, entre otros procesos. En general, las Hsp70 evitan que diversos factores proteotóxicos tengan efectos perjudiciales en las células, ya que provocarían un desequilibrio de la homeostasis proteica (Rosenzweig et al., 2019).

Bajo estrés celular aumentan los niveles de expresión de las Hsp70, lo cual es necesario para disuadir la agregación de proteínas dañadas por temperatura u otro tipo de estrés proteotóxico. Además, cooperan con mecanismos de degradación celular para eliminar proteínas aberrantes y agregados de proteínas. Las Hsp70 se han relacionado con diversas enfermedades neurodegenerativas, tales la enfermedad de Alzheimer, la enfermedad de Parkinson, la enfermedad de Huntington, la esclerosis lateral amiotrófica y las enfermedades priónicas, donde el plegamiento incorrecto de una o más proteínas específicas es la característica común en estas patologías (Christopher G. Evans et al., 2010; Martín-Peña et al., 2018).

La expresión y sobreexpresión de las Hsp70 es fundamental. Esto lo han demostrado algunos experimentos *knockout* en los genes que codifican a las Hsp, correlacionando con la acumulación de proteínas mal plegadas y que caracteriza a enfermedades neurodegenerativas como el Parkinson, Alzheimer o incluso diabetes tipo 2 (Radons, 2016; Tittelmeier et al., 2020). Por lo tanto, la actividad biológica de las chaperonas resulta extremadamente relevante en este contexto. Se han desarrollado diversos estudios *in silico*, *in vitro* e *in vivo* para dilucidar el mecanismo del plegamiento incorrecto y el papel de las Hsp70 en estas enfermedades, con el propósito de desarrollar terapias potenciales para enfrentar estas patologías humanas (Gupta et al., 2020; Pratt et al., 2015).

Ha sido claro que las Hsp70 ejercen un papel fundamental en la homeostasis proteica y actúan sobre una gran cantidad de sustratos. No obstante, aspectos en su funcionamiento e interacción molecular, así como el papel del dominio de unión a nucleótidos (NBD) en la unión y liberación de sustratos, los mecanismos de selectividad de proteínas blanco o por qué se requiere a las cochaperonas, no son del todo claros. Por lo que comprender los aspectos funcionales específicos de la interacción de las Hsp70 humanas con la gama de sustratos es complejo pero indispensable estudiar. En este proyecto de investigación nos enfocamos al estudio de la interacción de la Hsp70 humana con diferentes proteínas que experimentalmente se ha demostrado que interactúa, centrándonos en el análisis de los residuos interaccionantes desde un enfoque bioinformático, mediante el uso de herramientas como docking, modelado molecular, entre otros.

2 Marco teórico

Las proteínas son macromoléculas biológicas que realizan y/o median prácticamente todos los procesos biológicos centrales en los organismos vivos. Para realizar sus funciones biológicas adecuadamente, las proteínas recién sintetizadas deben plegarse en una estructura tridimensional adecuada, la cual es dinámica por naturaleza, lo que les permite realizar múltiples funciones en diferentes ambientes celulares.

El interior de la célula está repleto de proteínas y en ese entorno las proteínas chocan unas con otras constantemente, lo que podría tener consecuencias desfavorables que alteren su estructura funcional. Además, las proteínas mal plegadas o aquellas que han perdido parte de su estructura debido a factores externos pueden entrelazarse fácilmente con otras proteínas, formando agregados inespecíficos de proteínas debido a la exposición de las superficies hidrófobas. Esto conduce al deterioro de las proteínas y puede provocar una variedad de enfermedades humanas (Buchner, 2019; Ma et al., 2020).

Por lo tanto, para asegurar la funcionalidad óptima de las proteínas, tanto en condiciones fisiológicas como bajo condiciones ambientales cambiantes, se requieren factores que disminuyan el impacto negativo o que funcionen como control de calidad para enfrentar estos desafíos. En este contexto, las chaperonas moleculares resultan ser una solución para afrontar los distintos desafíos que pudieran alterar la funcionalidad de las proteínas.

En 1987 John Ellis propuso el término “chaperona molecular” para referirse a una clase específica de proteínas cuya función es asegurar el correcto plegamiento de otras cadenas de polipéptidos, además de su ensamblaje en estructuras oligoméricas (Ellis, 1988). Hoy se ha demostrado que la actividad de las chaperonas moleculares no sólo se limita a facilitar el plegamiento de otras proteínas, sino que también tienen un papel fundamental en el plegamiento de un gran número de proteínas de *novo*, y están involucradas en diferentes aspectos relacionados con el mantenimiento del proteoma, incluyendo la corrección del plegamiento de proteínas, disociación de agregados de proteínas, transporte y degradación proteica (Camberg et al., 2013; Nishimura & Akiyoshi, 2020).

2.1 Chaperonas moleculares

Las proteínas contienen una buena parte de la información requerida para adoptar su conformación nativa codificada en su secuencia de aminoácidos, sin embargo,

no todas logran plegarse. En la síntesis de proteínas se pueden generar polipéptidos desplegados, además, una vez que las proteínas alcanzan su conformación nativa existen condiciones estresantes que las pueden desplegar, tales como calor, especies reactivas de oxígeno y distintas condiciones patológicas. Las chaperonas moleculares son capaces de plegar proteínas desplegadas y/o mal plegadas, previenen la formación de agregados de proteínas y pueden reactivar estas proteínas agregadas, además de involucrarse en la degradación de proteínas aberrantes al interactuar con proteasas específicas (Camberg et al., 2013; Macošek et al., 2021).

Las diferentes proteínas de las células deben ser controladas cuidadosamente para mantener el estado de homeostasis proteica. De tal forma, que una compleja red de proteostasis ha evolucionado en todos los organismos vivos para asegurar que las proteínas se sinteticen en el momento y lugar adecuado. También evita que las proteínas se plieguen de forma incorrecta, y aseguran que se eliminen las proteínas mal plegadas, ya sea por autofagia o por degradación mediada por el proteasoma. Los componentes principales de la red de proteostasis son la presencia de diferentes chaperonas moleculares, relacionadas con el sistema de ubiquitina-proteasoma (UPS) y el sistema de autofagia que, en su conjunto, evitan la acumulación de agregados de proteínas que son potencialmente tóxicos (Hipp et al., 2019).

En la red de proteostasis, las chaperonas se pueden describir a partir de sus funciones efectivas en el contexto celular ya que pueden actuar como holdasa, foldasa, translocasa o disgregasa (Figura 1). Además, una chaperona no está limitada a una única actividad, pues en algunos casos una misma chaperona es capaz de llevar a cabo más de una de estas funciones (Macošek et al., 2021).

- **Las chaperonas holdasas**, son chaperonas usualmente independientes de ATP que se asocian con proteínas no nativas durante un tiempo prolongado para estabilizarlas y prevenir su agregación. Aunque no participan en el plegamiento directo de las proteínas, su actividad es indispensable ya que protegen a las proteínas que son vulnerables a la agregación. Un ejemplo de holdasas, son las proteínas de choque térmico pequeñas (sHsp, por sus siglas en inglés) (Hall, 2020; Haslbeck et al., 2019).
- **Las chaperonas foldasas**, participan en la remodelación estructural activa de las proteínas recién sintetizadas no nativas en un proceso de ciclado entre estados unidos a ADP y ATP. Los sistemas principales de foldasas incluyen a las familias de proteínas Hsp70/DnaK, Hsp60/GroEL y Hsp90 (Hayer-Hartl et al., 2016; Rosenzweig et al., 2019).

- **Las chaperonas translocasas**, transportan proteínas nacientes a través de las membranas. Son muy importantes en procariontes donde aproximadamente un tercio de todas las proteínas se exportan del citoplasma y, por lo tanto, deben translocarse a través de la membrana interna (Bechtluft et al., 2010; Vrontou & Economou, 2004).
- **Las chaperonas disgregasas**, permiten desenredar agregados insolubles y replegar las proteínas, o en su caso fungen como guía para la degradación proteica al interactuar con componentes proteolíticos. La familia de proteínas dependientes de ATP Clp/Hsp100 son uno de los principales sistemas de chaperonas con actividad disgregasa (Rosano et al., 2011; Sousa, 2014).

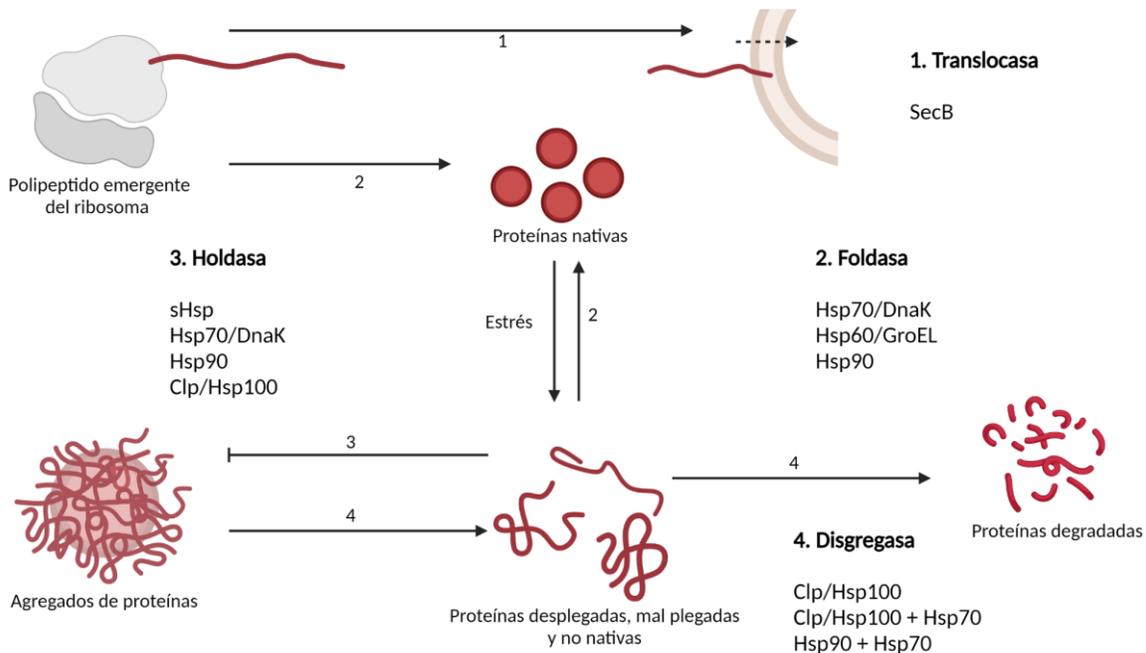


Figura 1. Funciones de las chaperonas en el control de calidad del proteoma. 1. Las proteínas recién sintetizadas pueden ser liberadas y transportadas como proteínas desplegadas (función de translocasa), 2. Pueden ser plegadas por chaperonas asociadas al ribosoma (función foldasa). Las proteínas desplegadas y mal plegadas pueden surgir como consecuencia de diversas condiciones de estrés celular. Algunas de ellas pueden ser remodeladas en proteínas nativas activas con ayuda de chaperonas. 3. Algunas chaperonas pueden brindar protección contra la agregación a la que son propensas las proteínas no nativas que se acumulan (función holdasa). 4. Algunas chaperonas o sistemas de chaperonas en conjunto pueden revertir el proceso de agregación al solubilizar los agregados (función disgregasa). Cuando las proteínas no pueden ser rescatadas por las chaperonas, son dirigidas por chaperonas a proteasas compartimentadas o al proteasoma. Imagen de autoría propia (Camberg et al., 2013). **Creado con BioRender.com**

2.2 Proteína de choque térmico 70 (Hsp70) como modelo de estudio

Se eligió a las proteínas de choque térmico 70 (Hsp70) como modelo de estudio debido a que sobresalen de entre las demás chaperonas moleculares por ser de las mejor conservadas en todos los reinos de la vida y cruciales en el mantenimiento de gran parte del proteoma. Además de ser unas de las chaperonas que más han llamado la atención en los últimos 50 años.

Las Hsp son una gran familia de chaperonas moleculares muy conservadas evolutivamente que se expresan constitutivamente en células procariotas y eucariotas de forma ubicua en los diferentes compartimentos celulares. En condiciones fisiológicas su expresión se incrementa cuando las células se someten a condiciones de estrés ambiental tales como: choque térmico, radiación, estrés oxidativo, hipoxia, entre otros. El descubrimiento de estas proteínas surgió como resultado de generar estrés térmico en *Drosophila* y observar un incremento de la expresión de genes que codifican estas proteínas (Ritossa, 1962). La mayoría son chaperonas dependientes de ATP y se clasifican de acuerdo a su peso molecular, e incluyen las Hsp pequeñas Hsp10 (10kD), Hsp27 (20-30kD) y las Hsp grandes Hsp 40 (40kD), Hsp 60 (60kD), Hsp 70 (70kD), Hsp 90 (90kD) y Hsp 110 (100 kD) (Dubrez et al., 2020; Guerrero-Rojas & Guerrero-Fonseca, 2018; Hirsh & Junger, 2008).

La subfamilia de chaperonas moleculares Hsp70 juega un papel relevante en el sistema de control de calidad de las proteínas celulares, su capacidad de unión a una amplia gama de proteínas sustrato le permite ayudar en distintos procesos celulares. Los miembros de la subfamilia Hsp70 se expresan en prácticamente todos los seres vivos, desde arqueobacterias, procariotas y eucariotas, incluyendo plantas y mamíferos, y están altamente conservadas con alrededor del 40-60% de identidad entre los homólogos procariotas y eucariotas (Havalová et al., 2021; Karunanayake & Page, 2021). Las Hsp70 están codificadas en un número variable de genes en procariotas y eucariotas, por ejemplo, *E. Coli* expresa tres isoformas de Hsp70 mientras que el genoma humano contiene 13 homólogos de Hsp70 (Tabla1) las cuales se expresan en citosol, núcleo, retículo endoplásmico y mitocondrias. Además, la abundancia de cada isoforma está regulada por los requerimientos celulares (Rosenzweig et al., 2019).

Se ha considerado que las distintas isoformas de Hsp70 pueden desempeñar funciones redundantes, como aquellas funciones básicas de mantenimiento de la homeostasis proteica relacionadas con el plegamiento de proteínas o la cooperación con maquinarias de degradación que son llevadas a cabo por cada una de las Hsp70, no obstante, también poseen un alto nivel de especialización, debido principalmente a diferencias en su expresión espacio temporal (Daugaard et al.,

2007). Algunas de estas Hsp70 son específicas de un compartimento celular, donde desempeñan actividades únicas y esenciales en el retículo endoplásmico, mitocondrias o plastidios (Figura 2), además de contar con cierta especificidad por distintas proteínas blanco.

Particularmente, las Hsp70 tienen un gran número de funciones de mantenimiento celular (conocido como *housekeeping*), tales como el plegamiento de proteínas recién sintetizadas, la translocación de polipéptidos a diferentes compartimentos celulares, montaje y desmontaje de complejos de proteínas, entre otros. También tienen funciones en respuesta a condiciones de estrés, previenen la agregación de proteínas, solubilizan proteínas agregadas, repliegan proteínas desplegadas o mal plegadas y cooperan con mecanismos de degradación celular como el sistema de ubiquitina-proteasoma (UPS) para eliminar proteínas aberrantes y agregados de proteínas (Kabani & Martineau, 2008; Rosenzweig et al., 2019).

Tabla 1. Proteínas de choque térmico 70 (Hsp70) humanas.

Proteína	Gen	Inducible por estrés	Localización celular	ID UniProt	Longitud (aa)
HspA1A	HSPA1A	Si	Citosol, núcleo, membrana celular, exosoma extracelular	P0DMV8	641
HspA1B	HSPA1B	Si	Citosol, núcleo, exosoma extracelular	P0DMV9	641
HspA1L	HSPA1L	No	Citosol, núcleo	P34931	641
HspA2	HSPA2	No	Citosol, núcleo, membrana celular, exosoma extracelular	P54652	639
HspA5	HSPA5	No	Retículo endoplasmático, exosoma extracelular	P11021	654
HspA6	HSPA6	Si	Citosol, exosoma extracelular	P17066	643
HspA7	HSPA7	Si	Micropartículas de sangre, exosoma extracelular	P48741	367
HspA8	HSPA8	No	Citosol, núcleo, membrana celular, exosoma extracelular	P11142	646
HspA9	HSPA9	No	Mitocondria, núcleo	P38646	679
HspA12A	HSPA12A	No	Intracelular, exosoma extracelular	O43301	675
HspA12B	HSPA12B	No	Células endoteliales, intracelular, plasma sanguíneo	B7ZLP2	685
HspA13	HSPA13	No	Retículo endoplasmático, exosoma extracelular, microsomas	P48723	471
HspA14	HSPA14	Si	Citosol, membrana celular	Q0VDF9	509

Tabla 1. Se muestran los 13 miembros de la familia Hsp70 expresadas en humano. Para cada una de las proteínas se menciona el nombre, el gen que la codifica, si es que su expresión puede ser inducida por factores de estrés tales como: calor, radiación UV, infecciones virales, hipoxia, estrés oxidativo, entre otros. Se menciona también la localización celular donde han sido identificadas y la longitud de su secuencia de aminoácidos.

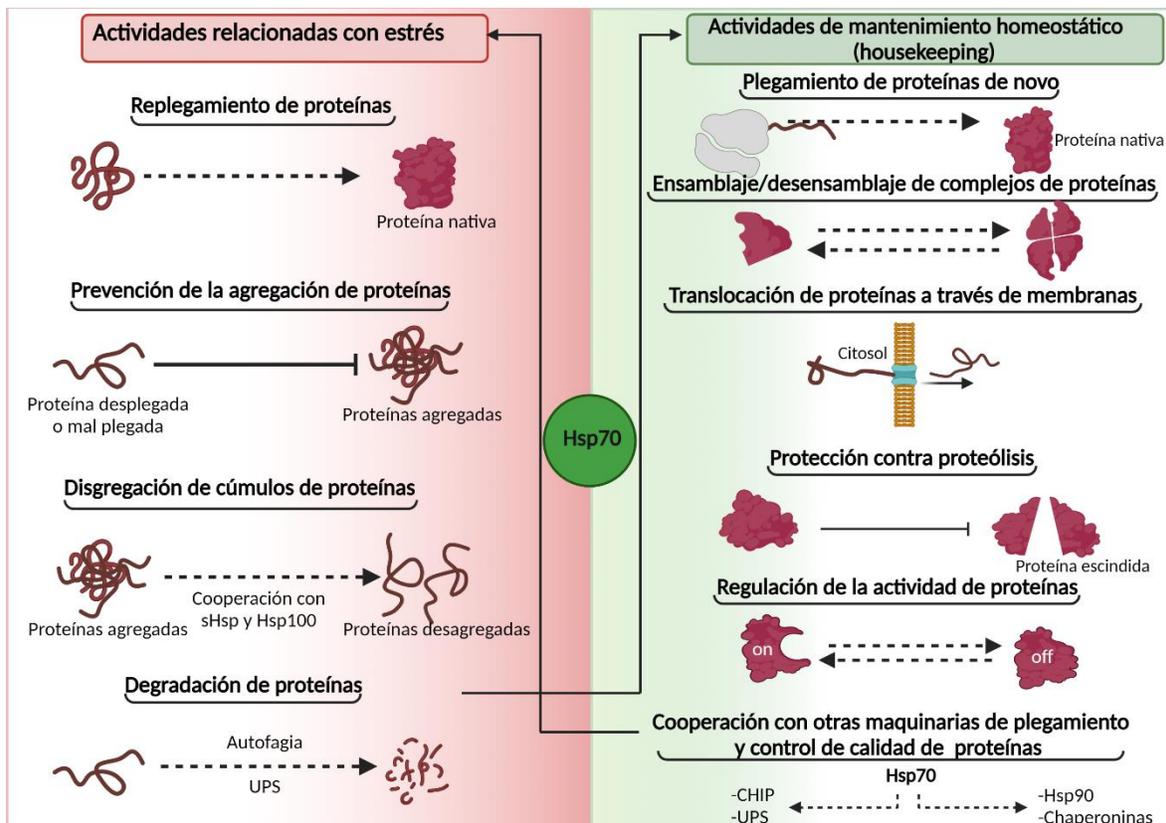


Figura 2. Funciones de la proteína de choque térmico 70 (Hsp70). Las Hsp70 participan en una amplia gama de funciones de mantenimiento homeostático (*housekeeping*), incluido el plegamiento de proteínas recién sintetizadas, el ensamblaje y desensamblaje de complejos de proteínas, la translocación de polipéptidos en mitocondrias, cloroplastos y el retículo endoplásmico, protección de las proteínas contra la proteólisis, la regulación de actividad proteica y apoyan con el traspaso de proteínas a otras chaperonas, como Hsp90 y las chaperoninas. Adicionalmente, las Hsp70 participan también en actividades relacionadas con estrés celular, incluyendo el replegamiento de proteínas mal plegadas o desplegadas, la prevención de la agregación de proteínas, la solubilización de proteínas agregadas y la cooperación con mecanismos de degradación celular, como la autofagia y el sistema de ubiquitina-proteasoma (UPS), para eliminar proteínas aberrantes y agregados de proteínas. Imagen de autoría propia (basado en Rosenzweig et al., 2019). **Creado con BioRender.com**

La estructura 3D de las Hsp70 es muy similar al homólogo de *E. coli*, denominado DnaK, el cual se compone de 4 regiones altamente conservadas. 1. Un dominio de unión de nucleótidos N-terminal (NBD) en forma de V integrado por dos subdominios o lóbulos que encierran el sitio donde se une e hidroliza ATP. 2. Un enlazador hidrofóbico rico en glicina/fenilalanina que conecta el NBD con el dominio de unión a sustrato (SBD). 3. Un SBD que se une a polipéptidos extendidos formado por un subdominio de hoja beta (SBD β) y un subdominio de hélice alfa (SBD α) que actúa como una tapa flexible y 4. Una cola C-terminal no estructurada de longitud variable (Figura 3). El NBD está bien conservado en todos los miembros de la familia Hsp70,

con la excepción de *HSPA12* que tiene un NBD más divergente. Las Hsp70 citosólicas contienen una región C-terminal rica en glicina/prolina que alberga un motivo altamente conservado EEVD que está implicado en la unión de cochaperonas y otras Hsp. El motivo EEVD está ausente en las Hsp70 específicas de algún organelo, como la Hsp70-5 residente en retículo endoplásmico o la Hsp70-9 mitocondrial, ambas con una señal de localización N-terminal (Burgess et al., 2016; Mayer & Gierasch, 2019; Radons, 2016; Zuiderweg et al., 2017).

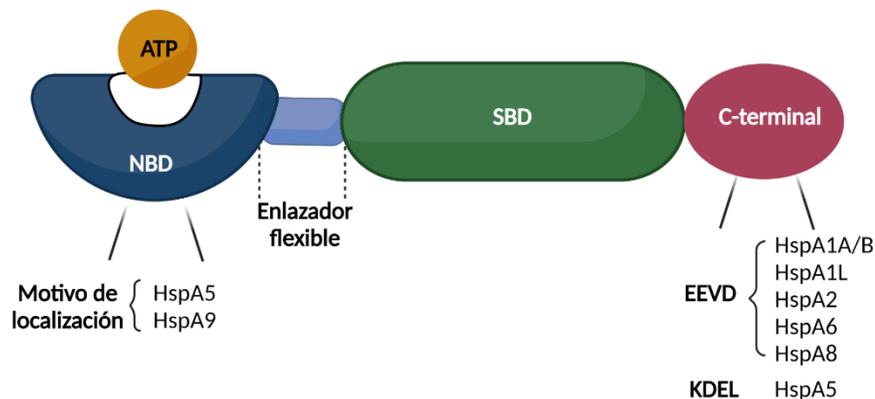


Figura 3. Representación esquemática de la estructura de los dominios conservados en la familia Hsp70. Las Hsp70 se componen de un dominio de unión a nucleótidos N-terminal (NBD, denominado también dominio de unión a ATP), una región flexible rica en glicina/fenilalanina, un dominio de unión a sustrato (SBD) seguido de una región C-terminal rica en glicina/prolina. Imagen de autoría propia (basado en Guerrero-Rojas & Guerrero-Fonseca, 2018; Radons, 2016). **Creado con BioRender.com**

6.4 Unión de la Hsp70 con diversas proteínas

Las Hsp70 se unen a diferentes sustratos de una manera que aparenta ser simple: el SBD se une a secuencias cortas de polipéptidos, mismos que no están del todo definidos estructuralmente. Evidencia basada en DnaK ha demostrado que las Hsp70 se unen a una variedad de péptidos de siete residuos de longitud. El uso de bibliotecas de péptidos y diversos ensayos han mostrado que existe una preferencia por los residuos hidrófobos y algunos aminoácidos de carga positiva, pero aún no se sabe de un motivo de unión claro (Clerico et al., 2015). Además, la capacidad de controlar la gran variedad de procesos celulares en eucariotas es impulsada por la capacidad de cooperar con más de 30 cochaperonas de dominio J, las cuales son

esenciales en la discriminación de sustratos; es así que, la Hsp70 y cochaperonas tienen un papel crucial en el control de la calidad de las proteínas en las células (Goloubinoff, 2017).

Por su diversidad o especificidad funcional las chaperonas moleculares, como la Hsp70 también se han propuesto como blancos terapéuticos. Lo anterior, debido a que cuando las proteínas escapan de los mecanismos de control de calidad de las proteínas, las formas mal plegadas se acumulan y posteriormente se agregan. Muchas enfermedades son consecuencia de la agregación de proteínas celulares, las más estudiadas son las enfermedades neurodegenerativas como la enfermedad de Parkinson, la enfermedad de Huntington, la esclerosis lateral amiotrófica y la enfermedad de Alzheimer (Tittelmeier et al., 2020), además de enfermedades musculares como la miopatía por cuerpos de inclusión y las miopatías por agregados de proteínas (Carlisle et al., 2018). Diversos estudios *in-silico*, *in-vitro* e *in-vivo* se han desarrollado para estudiar el papel de la Hsp70 sobre el control y prevención del plegamiento incorrecto y la agregación de proteínas, con la finalidad de desarrollar terapias de inducción farmacológica para prevenir la progresión de alguna patología (Gupta et al., 2020).

El análisis de la interacción de las Hsp70 de humano con otras proteínas podría contribuir a una mejor comprensión de los mecanismos que subyacen a las patologías, y podrían ayudar al desarrollo de terapias farmacológicas dirigidas a la actividad biológica de la proteína Hsp70.

2.4 Proteínas intrínsecamente desordenadas

Las proteínas intrínsecamente desordenadas (IDP, por sus siglas en inglés) y las proteínas que contienen regiones intrínsecamente desordenadas (IDPRs, por sus siglas en inglés) son de gran relevancia debido a que carecen de una estructura secundaria o terciaria estable, sin embargo, son capaces de llevar a cabo una amplia gama de funciones biológicas (Bondos et al., 2021). De modo que sugieren cambios en el paradigma clásico (secuencia, estructura, función) en la función biológica de las proteínas (Oldfield & Dunker, 2014) ya que las IDP/IDPRs se encuentran abundantemente en todos los proteomas de diferentes especies, y su presencia aumenta proporcionalmente con la complejidad de los organismo (Uversky, 2019), las IDP/IDPRs han llamado la atención ya que permiten tener cierto grado de adaptación conformacional debido a que incrementan su flexibilidad, aunque también aumentan la probabilidad de que dichas zonas sean puntos que promuevan la desnaturalización o agregación (Ayyadevara et al., 2021).

Hasta ahora, se ha estimado que una tercera parte de las proteínas eucariotas poseen IDPRs de más de 30 residuos de longitud, aunque esto no significa no

tengan una estructura. Más bien significa que muchas proteínas que poseen IDPRs tienen una estructura mayormente plegada con una región de desorden local o que presentan enlazadores desordenados que unen distintos dominios proteicos (Uemura et al., 2018; Wallin, 2017). Además, se puede considerar que las IDP pueden poseer un cierto “desorden intrínseco” determinado por su secuencia, pues tienen baja proporción de residuos que promueven el orden (baja hidrofobicidad) y alta proporción de residuos que promueven el desorden (cargas netas altas) (Bhattacharai & Emerson, 2020). Las IDP también pueden tener “sensibilidad ambiental” y es que condiciones como el pH, la temperatura, el estado redox y la alta osmolaridad inducen cambios conformacionales en algunas IDP (Cuevas-Velazquez et al., 2021).

Aunque estas proteínas carecen de una estructura estable, las IDP/IDPRs participan en distintas funciones biológicas importantes relacionadas con regulación, reconocimiento y señalización, donde las interacciones con otras proteínas tienen un papel crucial. Debido a que regulan distintos procesos celulares, las IDP/IDPR normalmente funcionan adecuadamente. Si estas se expresan, procesan o modifican incorrectamente son propensas a formar interacciones promiscuas que a menudo se asocian con el desarrollo de diferentes patologías. De hecho, la presencia de algunas IDP's e IDPR largos se ha asociado con algunos tipos de cáncer, neurodegeneración, diabetes, amiloidosis y enfermedades genéticas (Kulkarni & Uversky, 2019). Algunos ejemplos de IDPs, son la proteína alfa-sinucleína y la proteína Tau, las cuales son características de algunas enfermedades neurodegenerativas, donde estas proteínas constituyen los ovillos neurofibrilares observados en el cerebro de los pacientes con enfermedad de Parkinson y Alzheimer, respectivamente (Lermyte, 2020).

No obstante, no todas las IDPs se asocian a patologías. Particularmente la p53 que consta de regiones no estructuradas y regiones estructuradas se clasifica como una proteína nativa parcialmente desestructuradas o ligeramente plegada (Bell et al., 2002), que al interactuar con diversas proteínas de unión desempeña diferentes funciones vitales para mantener la estabilidad genómica y promover la reparación del ADN reduciendo la frecuencia de mutaciones que podrían ser cancerígenas (Uversky, 2016). Paradójicamente, algunas chaperonas moleculares que garantizan el correcto plegamiento de otras proteínas presentan algún nivel de desorden intrínseco. Son ejemplos los miembros de la familia de proteínas de choque térmico de 90 kDa y la misma Hsp70 que presentan un enlazador de estructura flexible o variable que vincula dos de sus dominios (Alderson et al., 2014; Gong et al., 2018; Hoter et al., 2018; V. Sudnitsyna et al., 2012).

2.5 La Hsp70 y proteínas intrínsecamente desordenadas

Se ha propuesto que las IDP son propensas a formar agregados y se sabe que el desorden está relacionado con la susceptibilidad a la proteólisis, por tal motivo su presencia y funcionamiento *in vivo* resulta de interés (Galea et al., 2006). Se sabe que las chaperonas moleculares interactúan con distintas IDP's, particularmente Hsp70 actúa sobre la proteína Tau inhibiendo su agregación e incluso es capaz de disminuir estos agregados que juegan un papel importante en la enfermedad de Alzheimer (Dou et al., 2003; Kundel et al., 2018). Otro ejemplo es la capacidad de Hsp70 para disminuir la agregación tóxica de la α -sinucleína relacionada con la enfermedad de Parkinson (Burmam et al., 2020).

Son distintos los casos que señalan que algunas IDP pueden necesitar la acción de las chaperonas moleculares para su correcto funcionamiento. De ser cierto, esto podría explicar cómo es que las IDP son estables en condiciones fisiológicas. No es claro si las IDP dependen de la asistencia de la Hsp70 o si la interacción se lleva a cabo con la IDPR, por lo que sigue siendo tema de estudio.

2.6 Herramientas bioinformáticas para el análisis de las interacciones entre proteínas

La bioinformática se define como la aplicación de herramientas de computación para el análisis e interpretación de datos biológicos. Es un área interdisciplinaria donde convergen la informática, las matemáticas, la física y la biología. Los alcances de la bioinformática son muchos ya que: nos permite analizar y alinear secuencias del genoma de cualquier organismo, analizar las variaciones y expresión de genes, predecir la estructura y función de proteínas, analizar las interacciones proteína-proteína, entre otras funciones (Kanehisa & Bork, 2003; Petrey & Honig, 2014). Cabe mencionar que los métodos y herramientas bioinformáticas involucran cierto grado de incertidumbre por lo que no reemplazan a las técnicas de laboratorio tradicionales, sin embargo, tienen la ventaja de ser más rápidas y menos costosas. De hecho, cada vez surgen más estudios *in silico* (en referencia al sílice del cual se compone la estructura interna de las computadoras) debido a que la tecnología actual permite que las simulaciones se acerquen cada vez más a los sistemas reales (Ekins et al., 2007; Fermeiglia et al., 2003; Liang et al., 2019).

Entre las herramientas de bioinformática más útiles para estudiar las interacciones entre distintas proteínas, se encuentran una gran variedad de programas y

servidores en línea dedicados al modelado estructural de proteínas. Estas herramientas facilitan la determinación de estructuras tridimensionales de proteínas para las que no se cuenta con información experimental proveniente de técnicas como cristalografía de rayos X o Resonancia Magnética Nuclear (NMR, por sus siglas en inglés) (Schwede, 2013). Además, existen diversos servidores dedicados a simular el acoplamiento molecular (*docking*) para predecir la estructura de complejos proteína-proteína, o proteína-ligando. A partir de las estructuras tridimensionales individuales. Estas herramientas son útiles para comprender aspectos fundamentales sobre la red de interacciones entre proteínas involucradas en el metabolismo celular (Petrey & Honig, 2014; Rodrigues & Bonvin, 2014).

De esta forma, diversas herramientas bioinformáticas se han usado para estudiar y mejorar la comprensión de los mecanismos moleculares involucrados con la Hsp70, su papel en el plegamiento proteico, la translocación y el desmontaje de agregados de proteínas (Ko et al., 2015; Kothandan et al., 2017; Siam et al., 2020; Singh et al., 2019). Muchos de los esfuerzos en investigación han sido para establecer un perfil de reconocimiento del sustrato por la Hsp70 a partir de datos experimentales obtenidos de matrices de péptidos, en los cuales se determinan cuáles son las secuencias que se unen a la chaperona. Esto ha permitido desarrollar algoritmos predictivos basados en secuencias para identificar probables sitios de unión de la Hsp70 presentes en otras proteínas (Rüdiger et al., 1997; Van Durme et al., 2009).

Por otro lado, se han realizado otros estudios experimentales sobre la comprensión en la interacción de la Hsp70 con sustratos biológicamente relevantes, así como las mutaciones que pueden afectar la unión de la chaperona, y su relevancia en distintas enfermedades denominadas *chaperonopatías* (Bross & Fernandez-Guerra, 2016; Clerico et al., 2015). Este tipo de patologías son causadas por chaperonas consideradas anormales (en estructura, localización, incluso concentración celular) que pueden estar afectadas por mutaciones o modificaciones postraduccionales tales como adenilación o AMPilación, que mediante estudios en modelos de *Saccharomyces cerevisiae* se ha observado que ocasiona alteraciones en su funcionamiento, dando lugar a la disminución del crecimiento celular (donde la raíz de este defecto es la interrupción de la proteostasis); plegamiento incorrecto de proteínas y al aumento de la presencia de agregados de proteínas tóxicas. Debido a esto, la deficiencia de la actividad de la Hsp70 y de otras chaperonas moleculares ha llamado la atención en años recientes, particularmente aquellas deficiencias relacionadas con trastornos neurodegenerativos y neuromusculares (Macario & de Macario, 2020; Scalia et al., 2021; Truttmann et al., 2017).

Aunque existen diversos avances experimentales en el estudio de la Hsp70, aún hay mecanismos que siguen sin ser claros. Uno de ellos, se relaciona con la falta de disponibilidad de la estructura completa de la Hsp70 humana, y es que debido a

la naturaleza flexible del enlazador que une el NBD y el SBD no se ha logrado la caracterización de alta resolución mediante cristalografía de rayos X de la proteína completa o de complejos Hsp70-sustrato. Además, las concentraciones de proteína requeridas para realizar su análisis estructural permiten a la Hsp70 formar dímeros y oligómeros (Fernández-Fernández & Valpuesta, 2018). Es por ello que el uso de las diferentes herramientas bioinformáticas tiene un gran potencial para enfrentar estos retos en el estudio de la Hsp70 que aún no han sido descritos.

6. OBJETIVOS

Objetivo general

Caracterizar la interacción y la relación funcional de proteínas blanco o cliente con la proteína de choque térmico (Hsp70) del humano.

Objetivos particulares:

- 1) Elaborar una base de datos sobre la interacción que tiene la Hsp70 con proteínas blanco o cliente realizando una búsqueda en UniProt.
- 2) Predecir la interacción de secuencias con la Hsp70, a partir de las secuencias de aminoácidos de sus proteínas cliente, usando el programa ChaperISM.
- 3) Identificar las regiones intrínsecamente desordenadas (IDPR) de los clientes mediante el servidor DisResClusteredDB.
- 4) Modelar las IDPR con el servidor SWISS-MODEL que no se encuentren presentes en la estructura tridimensional de las proteínas cliente.
- 5) Evaluar las posibles regiones de interacción estructural entre la Hsp70 y sus clientes mediante el servidor HADDOCK 2.4 usando de referencia las IDPR modeladas y las regiones que se predijeron en las proteínas cliente.

4. HIPÓTESIS

Si las proteínas de choque térmico 70 (Hsp70) se unen a sus proteínas cliente en regiones específicas, como las regiones proteicas intrínsecamente desordenadas (IDPRs), entonces, es posible establecer un patrón de secuencia-estructura que permita dilucidar la actividad de holdasa con nuevas proteínas no reportadas.

5. METODOLOGÍA

Resumen gráfico

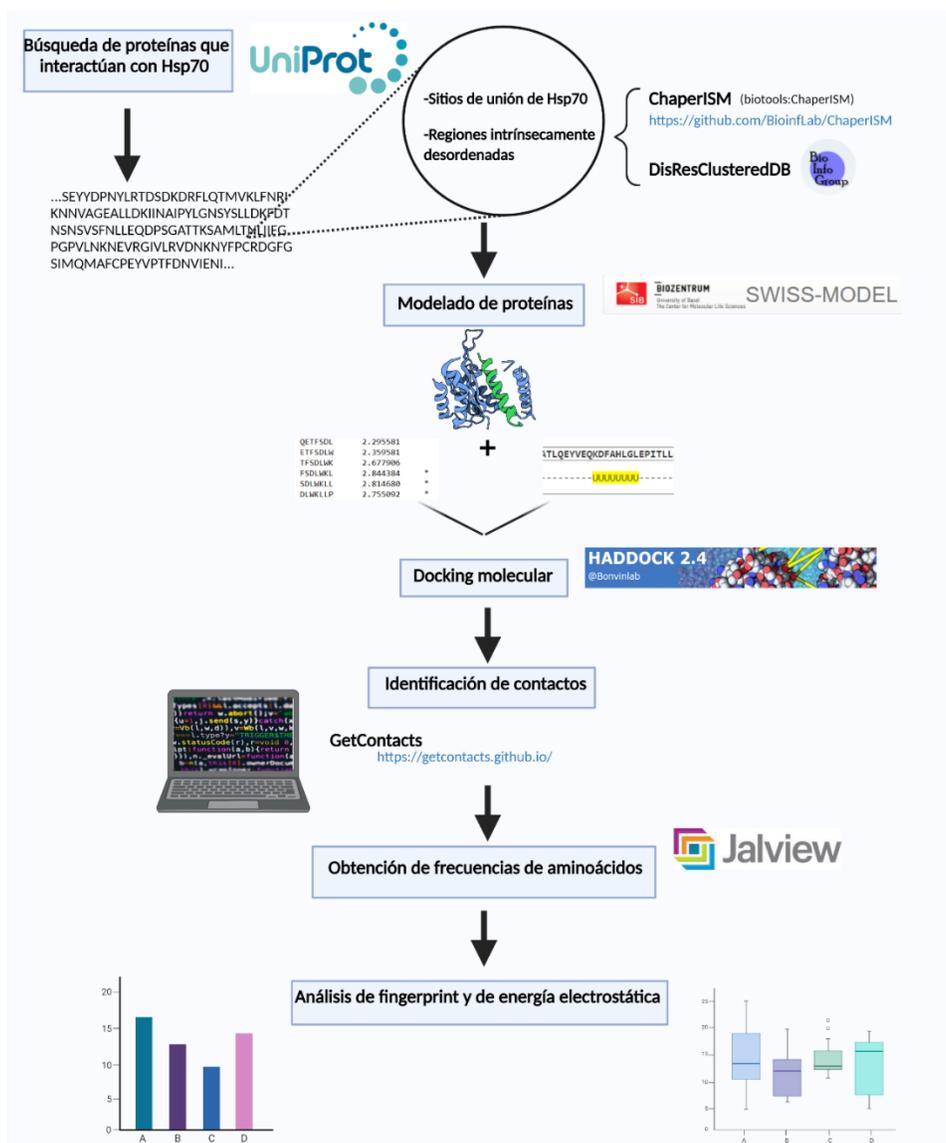


Figura 4. Resumen de la metodología empleada. El trabajo se dividió en los siguientes pasos: 1) elaboración de una base de datos que consistió en la búsqueda de proteínas blanco o cliente de la Hsp70, empleando la base de datos Uniprot. 2) identificación de regiones en las secuencias de los clientes con probabilidad de unirse a la Hsp70, empleando el predictor de unión ChaperISM. 3) identificación de regiones intrínsecamente desordenadas en los clientes, usando la base de datos DisResClusteredDB. 4) modelado de las proteínas cliente para incluir las regiones desordenadas. 5) elaboración de docking proteína-proteína con el SBD de la Hsp70 con distintas regiones específicas de las proteínas cliente, mediante HADDOCK 2.4. 6) identificación de los aminoácidos que hacen contactos en los resultados del docking usando la herramienta *getcontacts*. 7) realizar un análisis conocido como *fingerprint* de la interacción y un análisis de la energía electrostática involucrada en la interacción.

5.1 Búsqueda de proteínas blanco o “cliente” de la Hsp70

Con el propósito de identificar aquellas proteínas previamente reportadas que interactúan con las Hsp70 de humano, se realizó una búsqueda en la base de datos UniProt (<https://www.uniprot.org/>). Para ello nos enfocamos en las “interacciones binarias” donde se incluye información sobre diferentes reportes experimentales que confirman dicha interacción. Para ello se realizaron los siguientes pasos:

- i) Búsqueda del ID UniProt de cada una de las 13 Hsp70 humanas.
- ii) Recopilación de la secuencia e información relevante de cada Hsp70.
- iii) Identificación de las distintas proteínas que interactúan con cada Hsp70, para posteriormente organizarlas en una tabla de datos que incluían diferentes parámetros: Nombre, organismo, datos del experimento de interacción, secuencia, estructura 3D si existe y el ID Uniprot.

5.2 Elaboración de la base de datos

Para identificar la información que se recabó de cada una de las 76 proteínas identificadas y determinar las que serían incluidas en el estudio, elaboramos una base de datos en la que incorporamos los siguientes parámetros: Secuencia completa, archivos de estructura 3D disponibles, localización celular y datos de experimentos de interacción. La base de datos final incluye 30 proteínas de las cuales 11 se conservaron hasta el final del estudio. La razón por la que muchas proteínas se excluyeron fue que la información era escasa.

Habiendo determinado las proteínas cliente que se estudiarían el siguiente paso fue determinar las regiones donde la Hsp70 podría unirse, utilizando dos criterios principales: 1) secuencias de unión y 2) regiones intrínsecamente desordenadas (IDPRs).

5.3 Predicción de sitios de unión de la Hsp70 en posibles zonas de estructura desplegada de los clientes

Los predictores de unión de chaperonas se han desarrollado a partir de experimentos de unión de DnaK con diferentes péptidos, anotando las secuencias a las que la chaperona se une o no para construir una matriz de puntuación. Esto permite que al proporcionar una secuencia problema, sus aminoácidos sean evaluados para determinar si se unirá o no a la DnaK (Rüdiger et al., 1997). Otros métodos de predicción combinan cálculos de energía de interacción mejorando el poder de predicción, sin embargo, el modelo en que se basan es la DnaK de *E. coli* y no de la Hsp70 humana (Zhu et al., 1996).

El predictor ChaperISM (<https://github.com/BioinfLab/ChaperISM>) es un algoritmo desarrollado para identificar regiones de unión con la Hsp70 humana. ChaperISM presenta una mejora respecto a otros predictores debido a que no solo considera la frecuencia de aminoácidos en una secuencia para determinar si se une o no a la Hsp70, el algoritmo se desarrolló considerando la energía de interacción, además de datos de quimioluminiscencia provenientes de experimentos de interacción de la Hsp70 humana con distintos sustratos (Gutierrez et al., 2020).

ChaperISM está implementado en *Python* y su uso requirió de las secuencias de las proteínas cliente en un archivo previamente organizado en formato FASTA. Una vez ejecutado el algoritmo se obtiene un archivo de texto que contiene cada una de las secuencias de las proteínas cliente divididas en ventanas de heptámeros. Aquellos heptámeros con alta probabilidad de interacción con la Hsp70 se marcan con un asterisco. Finalmente se registró la información de las regiones probables de unión de cada una de las proteínas cliente.

5.4 Identificación de regiones intrínsecamente desordenadas (IDPR)

Con el propósito de identificar si las proteínas cliente de la Hsp70 presentaban una o más IDPRs, se realizó una búsqueda en la base de datos en línea “DisResClusteredDB” (http://bioproteom.protres.ru/cluster_pdb/), la cual contiene información de regiones desordenadas dentro de un gran número de proteínas cuya estructura 3D se encuentra disponible en el Protein Data Bank. Además, es posible entre otras cosas, encontrar homólogos en el banco de datos, realizar análisis estadísticos de residuos desordenados en estructuras de proteínas, buscar patrones desordenados y homo-repeticiones e identificar regiones desordenadas en distintas cadenas una proteína (Lobanov et al., 2020).

Para realizar la búsqueda se indicó el código PDB de la estructura 3D de las distintas proteínas cliente. DisResClusteredDB mostró la secuencia de la proteína y destacó las regiones desordenadas con una letra “U” contenidas en las distintas cadenas de la proteína. Finalmente se identificaron las IDPRs de todas las proteínas cliente, resaltando su ubicación dentro de la secuencia.

5.5 Modelado de proteínas

Para las proteínas de las que no se dispone de estructura 3D depositadas en el PDB o su estructura está incompleta, realizamos el modelado de su estructura a partir de la secuencia de aminoácidos con el servidor SWISS-MODEL.

El servidor web de modelado SWISS-MODEL (<https://swissmodel.expasy.org/>), es uno de los servidores de modelado de estructuras de proteínas más usado en el mundo, es pionero en el modelado automatizado de proteínas y cuenta con una buena precisión de modelado debido a la constante mejora de su motor y a su método de estimación de calidad. SWISS-MODEL genera un modelo de proteína en 3D a partir de una secuencia, lo cual se logra extrapolando información de una plantilla que puede ser la estructura de una proteína relacionada evolutivamente. Al final proporciona una estimación de la precisión del modelo que utiliza potenciales estadísticos de fuerza media para generar estimaciones de calidad global y por residuo (Waterhouse et al., 2018).

Procedimiento de modelado

El uso del servidor SWISS-MODEL consistió en:

- i) Crear un nuevo proyecto e identificar con el nombre de la proteína a modelar.
- ii) Subir la secuencia de la proteína a modelar.
- iii) En los parámetros de entrada se seleccionó “plantilla de usuario”, y se subió el archivo PDB de la plantilla elegida.
- iv) Para las proteínas que contienen IDPR y por lo tanto su estructura contiene regiones incompletas, la plantilla fue la misma proteína para que el modelo fuera lo más cercano a la estructura original, pero incluyendo las regiones ausentes.

- v) Generar un modelo de proteína 3D y se evaluó la estimación de la calidad del modelo global (GMQE) y el valor de QMEAN.
- vi) Finalmente se descargó el modelo de la proteína, se nombró el archivo y se registraron los valores de la estimación de la calidad.

Este procedimiento se repitió para obtener un modelo para cada proteína cliente. Al final se descartaron de los análisis posteriores a aquellas proteínas cuyos modelos tenían valores de GMQE menores a 0.7 y valores de QMEAN superiores a 1.5, además de los modelos que SWISS-MODEL directamente señala como malos modelos.

5.6 Docking

Una vez que se tenían las propuestas de los sitios de las proteínas donde se lleva a cabo la interacción con la Hsp70, se realizó un análisis de la interacción entre la Hsp70 y sus proteínas cliente mediante docking rígido utilizando el servidor HADDOCK 2.4.

Existen diversas herramientas para realizar docking molecular, la elección de un método depende de la información disponible acerca de la interacción y del tipo de moléculas. Para este trabajo se utilizó el servidor “HADDOCK 2.4” (<https://bianca.science.uu.nl/haddock2.4/>) debido a que una de las claves de su desarrollo es que parte de datos derivados de métodos bioquímicos, biofísicos o bioinformáticos para mejorar su potencial de acoplamiento. El método de HADDOCK intenta establecer las mejores conformaciones (o poses) entre las moléculas involucradas, tomando en cuenta una combinación de términos energéticos incluyendo la energía electrostática y la energía de Van der Waals, que junto a otros parámetros conforman la puntuación (score) HADDOCK. El método también incluye un análisis de *clusters* o grupos, donde las poses finales se organizan en *clusters* o grupos asociados con el score HADDOCK y otros parámetros energéticos, de modo que los miembros o poses más similares se incluyan en un mismo cluster (Van Zundert et al., 2016; Vangone et al., 2017).

Preparación de los datos de entrada

Para usar el servidor HADDOCK 2.4 se requieren al menos dos estructuras PDB (que se identifican como molécula A y B), además de información que debe ser proporcionada por el usuario, tal como indicar los residuos “activos” o de referencia, es decir, aquellos residuos específicos que son clave para el acoplamiento de

ambas moléculas. Para definir como “activo” un aminoácido en la molécula receptora se debe tener evidencia clara de interacción con un ligando, dicha evidencia fue obtenida con el uso de la herramienta *getcontacts* donde se identificaron los aminoácidos de la molécula A que interactúan con un sustrato peptídico, y mientras que para la molécula B, los aminoácidos activos se definieron usando el predictor ChaperISM y las IDPRs.

Molécula A:

La molécula receptora en todos los acoplamientos realizados fue la Hsp70, se consideró el archivo PDB ID 4PO2 que contiene la estructura de un dímero del dominio de unión al sustrato de la Hsp70 en complejo con un sustrato peptídico. Se modificó el archivo para conservar un monómero, posteriormente se identificaron aquellos residuos que están en contacto directo con el sustrato peptídico usando la herramienta *getcontacts* (se detalla más adelante). Por último, se eliminó el péptido del complejo SBD-sustrato de modo que las proteínas cliente pudieran ser acopladas en la misma región donde se encontraba el sustrato peptídico.

Molécula B:

Esta molécula es la proteína cliente o sustrato que corresponde a cada uno de los modelos elaborados previamente para todas las proteínas que se incluyeron en la base de datos.

Residuos de referencia:

Los residuos de referencia de la molécula A son los mismos en cada uno de los distintos acoplamientos realizados, a diferencia de los residuos de referencia para la molécula B que son distintos en cada acoplamiento. Estos residuos provienen de la información obtenida previamente en 5.3 y 5.4, correspondiente a la predicción de sitios de unión con la Hsp70 e IDPRs respectivamente.

Procedimiento para hacer el docking

El procedimiento para hacer el docking rígido proteína-proteína con HADDOCK 2.4 consta de los siguientes pasos:

- i) Se crea un proyecto nuevo, se nombra y se define el número de moléculas involucradas.

- ii) Se suben las estructuras correspondientes a las moléculas A y B, además se indica el tipo de estructura que se está proporcionando (proteína, molécula pequeña, ARN o ADN).
- iii) Se indican los residuos activos correspondientes a ambas moléculas y se sube el proyecto.
- iv) El tiempo que tarda HADDOCK en completar el docking varía alrededor de un par de horas, se puede monitorear el avance del proyecto e incluso se notifica por correo electrónico una vez que se finalizó un proyecto.
- v) Finalmente, se descargan los resultados en un archivo comprimido (estructuras de los modelos, gráficos y estadísticas) y se organizan en subcarpetas según la región estructural y la proteína evaluada.

5.7 identificación de contactos: *Fingerprint*

Para identificar aquellos residuos clave o *fingerprint* de la interacción entre las proteínas acopladas con HADDOCK 2.4 se utilizó la herramienta *getcontacts* (<https://getcontacts.github.io/>). La herramienta *getcontacts* está implementada en *Python* y permite calcular redes dinámicas de contacto no covalente de un archivo de un complejo PDB proteína-cliente.

El procedimiento consistió en:

- i) Se organizaron los archivos PDB resultantes de docking proteína-proteína.
- ii) Se filtraron los resultados del docking para conservar aquellos *clusters* que presentaban una energía electrostática por debajo de -200 kcal/mol. Además, se conservó sólo la primera pose de cada uno.
- iii) Se organizaron los *clusters* en directorios de acuerdo al nombre de la proteína cliente y región estructural evaluada.
- iv) Se utilizó la herramienta *getcontacts* partiendo de los archivos PDB seleccionados.

- v) Se obtuvieron diferentes archivos con valores separados por tabulaciones (.tsv) donde se identifican los elementos involucrados en la interacción.

5.8 Análisis de los contactos

A partir de los archivos tsv obtenidos con la herramienta *getcontacts* y mediante programación en *Python*, se obtuvieron únicamente las columnas correspondientes a aminoácidos de la proteínas cliente involucrados en el contacto con la Hsp70.

Se utilizó la herramienta Jalview con el propósito de obtener las frecuencias de aminoácidos que aparecían en los contactos obtenidos. El procedimiento consistió en:

- i) Se cambia el formato de las secuencias de aminoácidos al código de una letra y se compilan en archivos de texto identificando en nombre de la proteína cliente, la región evaluada y el cluster.
- ii) Se cargaron los archivos de texto a Jalview y se alinearon con el método “muscle”.
- iii) Una vez realizado el alineamiento, se importaron los resultados de la frecuencia de aminoácidos a una hoja de cálculo y se realizaron gráficos de barras.
- iv) Se compilaron los datos de la energía electrostática y se incorporaron a la misma hoja de cálculo y se realizaron gráficas de caja y bigotes.

6. Resultados

Como se mencionó previamente, hasta ahora no hay estructura tridimensional completa reportada de algún miembro de la familia de la Hsp70 de humano. Lo que está disponible es la estructura del SBD de la **HSPA1A** (PDB ID 4PO2), la principal Hsp70 que puede ser inducida por estrés. No obstante, el dominio SBD es suficiente para establecer posibles relaciones de la unión con los clientes. Además, el SBD está altamente conservado en todas las isoformas de la Hsp70, por lo que se decidió realizar el docking molecular proteína-proteína usando esta estructura y así evaluar la interacción con las proteínas de la base datos que se construyó de los clientes de las 13 Hsp70.

6.1 Identificación de proteínas clientes de la Hsp70

Las proteínas blanco de la Hsp70 son abundantes, la interacción ocurre en distintas regiones dentro de la célula, por lo que en este proyecto recopilamos datos de las proteínas que se ha demostrado tienen interacción con las diferentes Hsp70 humanas. Para el análisis se obtuvo una serie de 11 proteínas clientes ([Tabla 2](#)) las cuales contaban con la descripción de datos experimentales de la interacción biológica con proteínas, localización celular, secuencia de aminoácidos completa y estructura 3D reportada.

Tabla 2. Proteínas cliente de la Hsp70 evaluadas.

Nombre	ID Uniprot	Gen	Localización celular	Longitud (aa)
Serina / treonina-proteína quinasa/ endoribonucleasa IRE1	O75460	ERN1	Membrana del retículo endoplásmico	977
Cadena A de L-lactato deshidrogenasa	P00338	LDHA	Citoplasma	332
Cinasa rica en repetidos de leucina 2	Q5S007	LRRK2	Membrana del aparato de Golgi, lisosoma, endosoma, membrana del retículo endoplásmico, membrana externa de la mitocondria	2,527
Proteína de reparación de mismatch de ADN Mlh1	P40692	MLH1	Núcleo, cromosoma	756
Antígeno tumoral celular p53	P04637	TP53	Núcleo, centrosoma, retículo endoplásmico, citoplasma, matriz mitocondrial	393
Proteína tumoral p73	O15350	TP73	Núcleo, citoplasma	636
Proteína disulfuro-isomerasa A6	Q15084	PDIA6	Membrana celular, luz del retículo endoplásmico, melanosoma	440
Proteína de protección de los telómeros 1	Q9NUX5	POT1	Núcleo, telómeros	634
Proteína similar a la S-adenosilhomocisteína hidrolasa 1	O43865	AHCYL1	Membrana celular apical, retículo endoplásmico, citosol	530
Proteína tumoral controlada traslacionalmente	P13693	TPT1	Citoplasma	172
Toxina tetánica (<i>Clostridium tetani</i>)	P04958	tetX	Citosol, región extracelular, membrana plasmática	1,315

Tabla 2. 11 proteínas cliente de la Hsp70 seleccionadas. Se especifica el nombre, el ID en Uniprot, el gen que las codifica, la localización celular donde se han identificado y la longitud de su secuencia completa. Información recopilada de la base de datos Uniprot (Bateman, 2019).

La versatilidad funcional de la Hsp70 de participar biológicamente en diferentes regiones celulares queda evidenciada con el tipo de cliente de la Tabla 2. Por

ejemplo, algunas de estas proteínas cliente como la proteína MLH1, desempeñan funciones tan importantes como la reparación de errores durante la replicación del DNA (Pećina-Šlaus et al., 2020). La proteína p53 conocida como “guardián del genoma” la cual destaca su función durante la apoptosis, ya que controla el ciclo celular en caso de presentarse una mutación (Chen, 2016). No se explica si la Hsp70 participa en el buen plegamiento de las proteínas, en el acompañamiento hasta su lugar de acción o para evitar la agregación de proteínas.

6.2 Determinación de sitios de unión con la Hsp70 e identificación de IDPRs

Una vez identificadas las proteínas que son asistidas por alguna de las Hsp70 se identificaron las regiones de las proteínas cliente donde se podía llevar a cabo la interacción. Así, para cada proteína cliente se obtuvieron distintas regiones con la probabilidad de unirse a la Hsp70, además se identificaron los IDPRs en la estructura de cada proteína (Tabla 3). Cabe mencionar que estos dos tipos de regiones no concuerdan entre sí.

Tabla 3. Regiones de las proteínas cliente con probabilidad de unirse a Hsp70.

Proteína Cliente	Heptámeros	IDPRs
Serina / treonina-proteína quinasa / endoribonucleasa IRE1	14	3
Cadena A de L-lactato deshidrogenasa	10	1
Cinasa rica en repetidos de leucina 2	7	2
Proteína de reparación de mismatch de ADN Mlh1	9	3
Antígeno tumoral celular p53	7	3
Proteína tumoral p73	8	1
Proteína disulfuro-isomerasa A6	5	2
Proteína de protección de los telómeros 1	13	2
Proteína similar a la S-adenosilhomocisteína hidrolasa 1	15	2
Proteína tumoral controlada traslacionalmente	6	1
Toxina tetánica (<i>Clostridium tetani</i>)	15	2

Tabla 3. Se muestra la cantidad de sitios predichos de unión (heptámeros) y de IDPRs identificados para cada proteína cliente.

Mediante ChaperISM, usando la secuencias de las proteínas cliente, se obtuvieron diferentes sitios probables de unión para cada proteína. Ya que los heptámeros de unión que predice ChaperISM se van a analizar para ver si presentan correlación con la estructura 3D de la proteína cliente, entonces, se descartaron aquellos sitios predichos que no se podían relacionar con la estructura del PDB disponible, ya que la estructura estaba incompleta.

Por otro lado, se predicen menos regiones de IDPRs que los heptámeros obtenidos por ChaperISM, esto sugiere que la estabilidad marginal de las proteínas cliente dificulta que las proteínas se desplieguen y que las Hsp70 actúan como chaperonas de plegamiento, entonces, si esto es cierto, la participación de la Hsp70 debería actuar como holdasa o disgregasas.

No obstante, los héptameros y las IDPRs sí pueden participar como puntos de reconocimiento entre las proteínas cliente y la Hsp70. Para probar esta idea, a continuación, se explora el docking proteína-proteína usando los heptámeros y las IDPRs predichas.

6.3 Modelado de proteínas

Debido a que las IDPRs no se encuentran como una única conformación estructural, sino que se trata de regiones con una estructura flexible, pudiendo pasar de una conformación estructural a otra distinta en poco tiempo, son difíciles de analizar y cristalizar. Por lo tanto, en algunas proteínas cliente usadas no se lograron resolver dichas regiones en las estructuras tridimensionales reportadas de las proteínas (Figura 5. a, c). Para afrontar este inconveniente se modelaron las regiones desordenadas presentes en las proteínas cliente mediante el servidor SWISS-MODEL (ejemplos representativos se muestran en la figura 5. b, d) y se eligieron sólo aquellos modelos con una calidad de modelado aceptable determinada por su valor de **GMQE** (un valor de 1 es ideal) y **QMEAN** (el valor óptimo es 0 pero se considera bueno un valor en el rango de -1 a 1) (Tabla 4).



Figura 5. Modelos representativos de las regiones proteicas intrínsecamente desordenadas (IDPRs). **a.** Se muestra la estructura de la proteína MLH1 resaltando una región ausente en la estructura correspondiente a una IDPR, **b.** Se muestra el modelo de la proteína MLH1 que incluye la IDPR (amarillo) ausente en la estructura original, **c.** Se muestra la estructura de la proteína p53 con una región ausente, **d.** Se muestra el modelo de la proteína p53 que incluye la IDPR (amarillo) la cual no incluye la estructura original. Imágenes obtenidas con VMD.

Tabla 4. Modelos elaborados de las 11 proteínas cliente.

Nombre proteína modelada	PDB	GMQE	QMEAN
Serina / treonina-proteína quinasa / endoribonucleasa IRE1	4YZC	0.90	-0.73
Cadena A de L-lactato deshidrogenasa	4JNK	0.93	-0.38
Cinasa rica en repetidos de leucina 2	2ZEJ	0.75	-1.11
Proteína de reparación de mismatch de ADN Mlh1	3RBN	0.79	-1.54
Antígeno tumoral celular p53	1GZH	0.87	0.52
Proteína tumoral p73	2XWC	0.86	-0.23
Proteína disulfuro-isomerasa A6	4GWR	0.84	0.90
Proteína de protección de los telómeros 1	3KJO	0.94	-0.76
Proteína similar a la S-adenosilhomocisteína hidrolasa 1	3MTG	0.91	-0.52
Proteína tumoral controlada traslacionalmente	1YZ1	0.76	-1.42
Toxina tetánica (<i>Clostridium tetani</i>)	3HMY	0.96	-0.83

Tabla 4. Se indican las proteínas modeladas, el archivo PDB del cual partió del modelado de las IDPRs y los valores de estimación de la calidad de modelado GMQE y QMEAN.

6.4 Residuos de referencia del SBD de la Hsp70

A partir del archivo PDB (4PO2), que contiene la estructura tridimensional del SBD de la Hsp70 en complejo con un péptido representativo de una estructura desplegada de una proteína, *getcontacts* identificó los residuos del SBD que están en contacto con el péptido. A continuación, se muestra la secuencia de la HSPA1A resaltando el SBD y se señalan los residuos de referencia.

```

MAKAAAIGID LGTTYSCVGV FQHGKVEIIA NDQGNRTTPS YVAFDTERL 50
IGDAAKNQVA LNPQNTVFDA KRLIGRKFGD PVVQSDMKHW PFQVINDGDK 100
PKVQVSYKGE TKAFYPEEIS SMVLTKMKEI AEAYLGYPVT NAVITVPAYF 150
NDSQRQATKD AGVIAGLNVL RIINEPTAAA IAYGLDRTGK GERNVLIFDL 200
GGGTFDVSIL TIDDGIFEVK ATAGDTHLGG EDFDNRLVNH FVEEFKRKHK 250
KDISQNKRAV RRLRTACERA KRTLSSSTQA SLEIDSLFEG IDFYTSITRA 300
RFEELCSDLF RSTLEPVEKA LRDAKLDKAQ IHDLVLVGGG TRIPKVQKLL 350
QDFFNGRDLN KSINPDEAVA YGAAVQAAIL MGDKSENVQD LLLLDVAPLS 400
**** * ***** * ****
LGLLETAGGVM TALIKRNSTI PTKQTQIFTT YSDNQPGVLI QVYEGERAMT 450
* * *
KDNLLGRFE LSGIPPAPRG VPQIEVTFDI DANGILNVTA TDKSTGKANK 500
ITITNDKGR L SKEEIERMVQ EAEKYKAEDE VQRERVSASN ALESYAFNMK 550
SAVEDEGLKG KISEADKKV LDKCQEVISW LDANTLAEKD EFEHKRKELE 600
QVCNPIISGL YQGAGGPGPG GFQAQGPCKG SSGPTIEEV D 641

```

6.5 Interacción de proteínas con la Hsp70 a través del SBD

La idea inicial para realizar el docking era determinar si la Hsp70 interactúa con las proteínas cliente mediante el SBD, de manera similar a como se encuentra capturado el pequeño péptido en la estructura cristalizada (Figura 6. a, c).

Para evaluar si HADDOCK es adecuado para los objetivos de este proyecto se realizó un experimento control en el cual se separó la estructura del péptido del complejo, posteriormente se realizó el docking como estructuras independientes dando como referencia los aminoácidos del SBD de Hsp70. El resultado obtenido fue el esperado ya que el acoplamiento se lleva a cabo en el mismo sitio, solo con mínimas diferencias en la posición de las cadenas laterales del péptido (Figura 6. b, d).

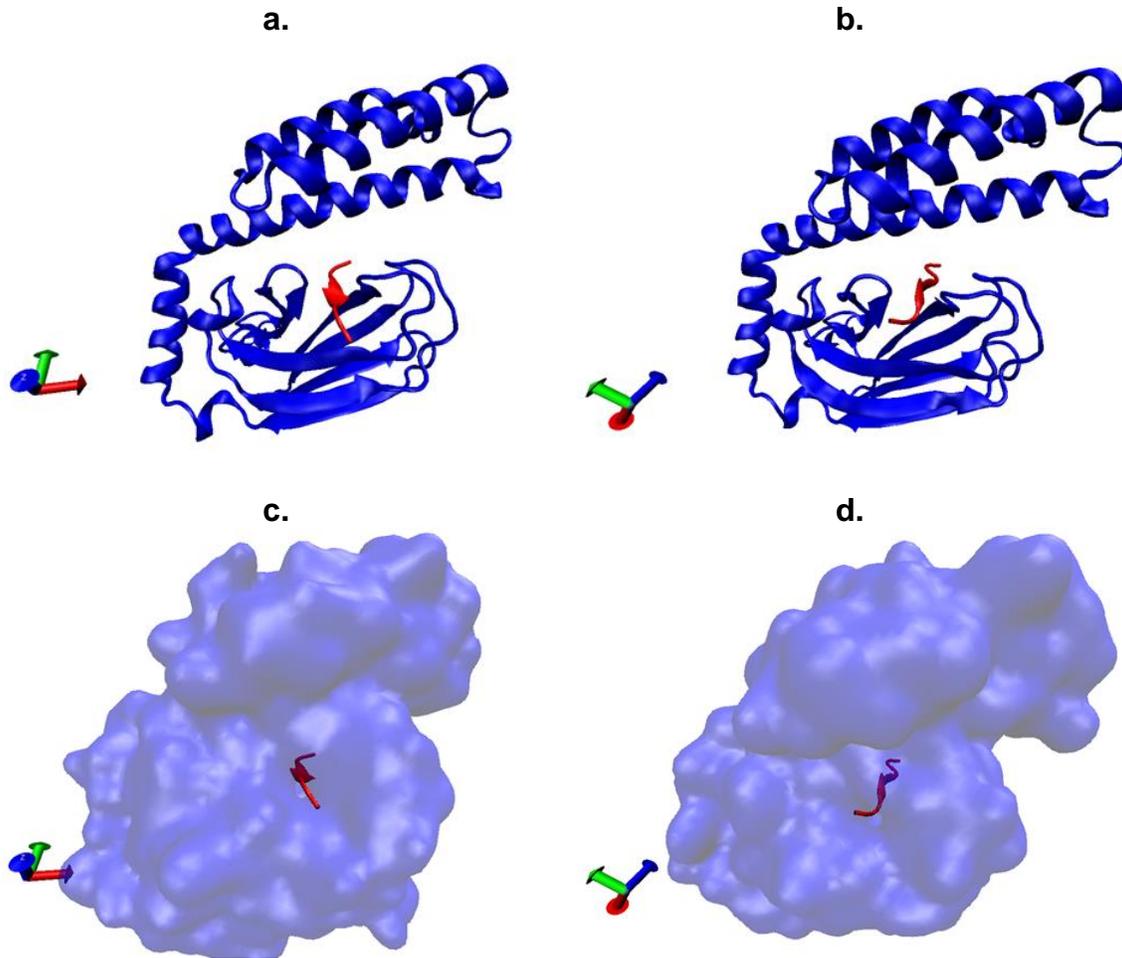


Figura 6. Interacción del SBD de la Hsp70 con un péptido. **a, c.** Interacción de la Hsp70 con un péptido incluidos en la estructura tridimensional reportada (archivo pdb: 4PO2). **b, d.** Interacción Hsp70 con un péptido obtenida con HADDOCK 2.4. Imágenes obtenidas con VMD.

El SBD de la Hsp70 forma una “cavidad” que reconoce un pequeño segmento de las proteínas cliente (Figura 7. a, c, e, g), para la mayor parte de los acoplamientos elaborados, los segmentos de las proteínas cliente que interactúan con el SBD de la Hsp70 son loops o segmentos que no forman estructuras definidas como α -hélices u hojas β (Figura 7. b, d, f, h) que deben acoplarse en dicho bolsillo.

El tamaño de los polipéptidos acoplados con la Hsp70 puede ser un factor importante para la interacción. Ya que las proteínas cliente se acoplan en una zona muy específica de la Hsp70, el tamaño de la superficie de las proteínas clientes podrían facilitar o dificultar la interacción ya que sólo un segmento de la proteína pudiera acoplarse en la cavidad de la Hsp70. En cualquier caso, se debe considerar que hasta ahora muchos de los ensayos representativos de interacción con la Hsp70, así como de otras chaperonas, se llevan a cabo empleando péptidos cortos

y extendidos brindando información muy útil sobre los mecanismos de interacción, pero resulta no ser suficiente para comprender cómo se capturan las proteínas completas.

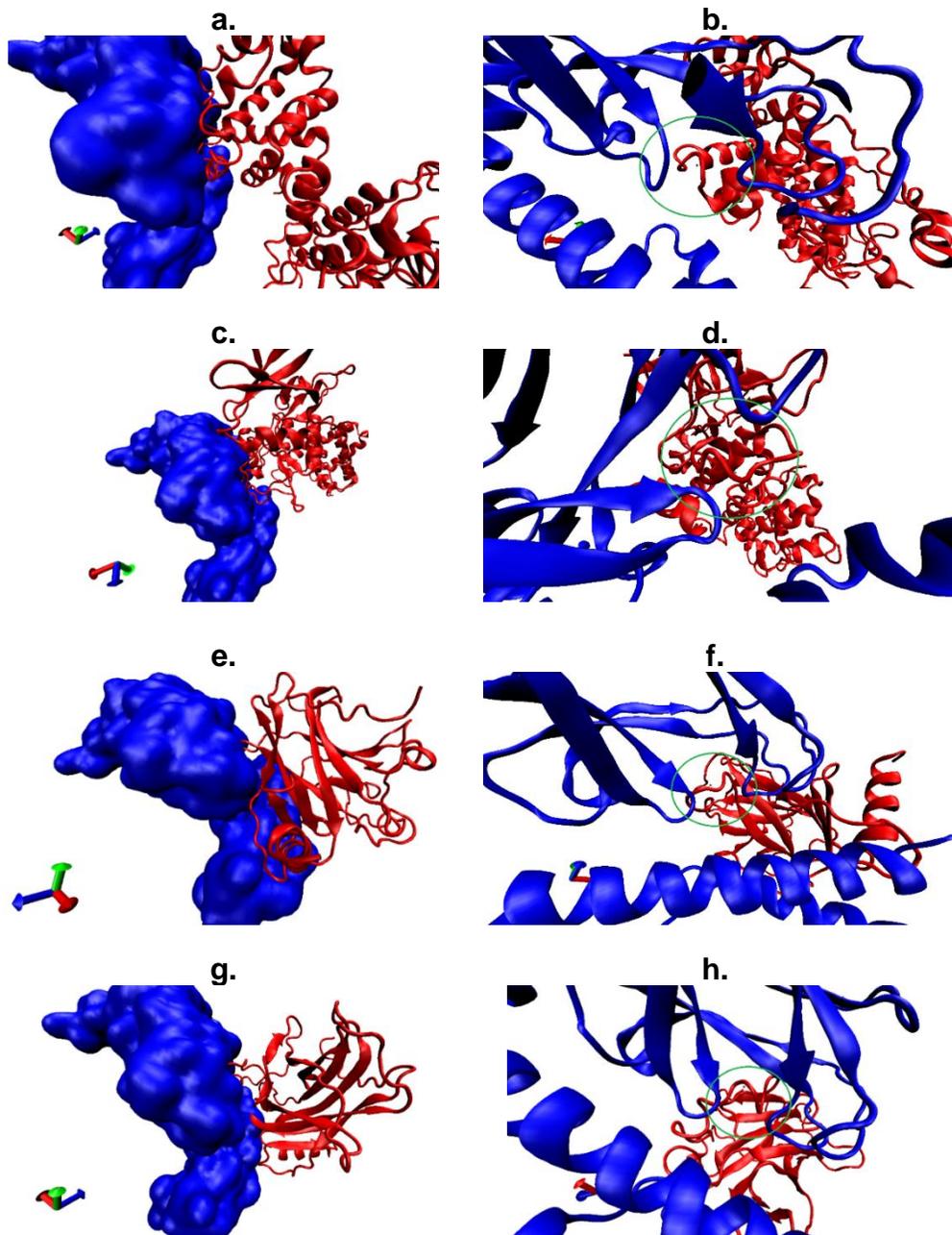


Figura 7. Figuras representativas del docking proteína-proteína hecho con el servidor HADDOCK 2.4. Se muestran distintos acoplamientos entre el SBD de la HSP70 (azul) y distintas proteínas cliente (rojo). Cada par de figuras pertenecen a un mismo plano. **a. b.** Hsp70 y una región de unión heptamérica de la proteína ENR1 (rojo). **c. d.** Hsp70 una IDPR de la proteína ENR1 (rojo). **e. f.** Hsp70 y una región de unión heptamérica de la proteína p53 (rojo). **g. h.** Hsp70 una IDPR de la proteína P53 (rojo). Se resalta con un círculo verde el segmento de la proteína cliente que interactúa con el SBD de la Hsp70.

Mediante la herramienta HADDOCK 2.4 acoplamos la Hsp70 con diferentes proteínas cliente, las cuales son de mayor tamaño que los péptidos empleados en diferentes ensayos de interacción. En los resultados de docking podemos identificar distintos parámetros entre los que se encuentran: 1) la energía electrostática, 2) energía de Van der Waals, 3) energía de desolvatación y 4) la energía de restricciones de distancia. Se tomó la energía electrostática como principal criterio para analizar los mejores resultados del docking. En nuestro caso se consideró que un valor menor a -200 kcal/mol de energía electrostática aumenta la probabilidad que se trate de un buen acoplamiento. Se realizó un filtrado y se descartaron los acoplamientos donde había una interacción electrostática con valor mayor al corte.

Se realizaron en total 131 acoplamientos con HADDOCK 2.4, este servidor arroja los resultados en *clusters* que contienen distintas poses para cada acoplamiento. En la [tabla 5](#) se muestran los *clusters* obtenidos para cada proteína cliente y los que se seleccionaron para posteriormente incluirlos en el análisis de contactos, también llamado *fingerprint*, proteína-proteína.

Tabla 5. Resultados del docking proteína-proteína obtenidos con HADDOCK 2.4.

Nombre proteína	Clusters iniciales (Heptámeros)	Clusters finales (Heptámeros)	Clusters iniciales (IDPRs)	Clusters finales (IDPRs)
Serina / treonina-proteína quinasa / endoribonucleasa IRE1	131	81	28	9
Cadena A de L-lactato deshidrogenasa	96	40	10	7
Cinasa rica en repetidos de leucina 2	60	16	17	8
Proteína de reparación de mismatch de ADN Mlh1	87	29	23	8
Antígeno tumoral celular p53	57	28	15	3
Proteína tumoral p73	73	7	9	3
Proteína disulfuro-isomerasa A6	39	13	16	9
Proteína de protección de los telómeros 1	114	34	15	2
Proteína similar a la S- adenosilhomocisteína hidrolasa 1	130	55	7	3
Proteína tumoral controlada traslacionalmente	46	34	9	5
Toxina tetánica (<i>Clostridium tetani</i>)	137	49	19	4

Tabla 5. Se muestra la cantidad de *clusters* resultantes de los acoplamientos realizados con HADDOCK (iniciales) y la cantidad de *clusters* que cumplieron con los valores de energía electrostática previamente definidos (finales).

6.6 Energía electrostática en la interacción Hsp70-cliente

HADDOCK evalúa el puntaje de interacción como la suma de fuerzas electrostáticas y de Van der Waals, y la energía de desolvatación estimados. Debido a la presencia de aminoácidos con carga en las regiones de las proteínas cliente evaluadas, la energía electrostática predomina sobre los otros parámetros evaluados. Para comparar si la estimación de dicha energía difería entre las regiones heptámicas predichas y las IDPRs, se realizó el análisis de dispersión sobre cada clúster seleccionado (Figura 8.).

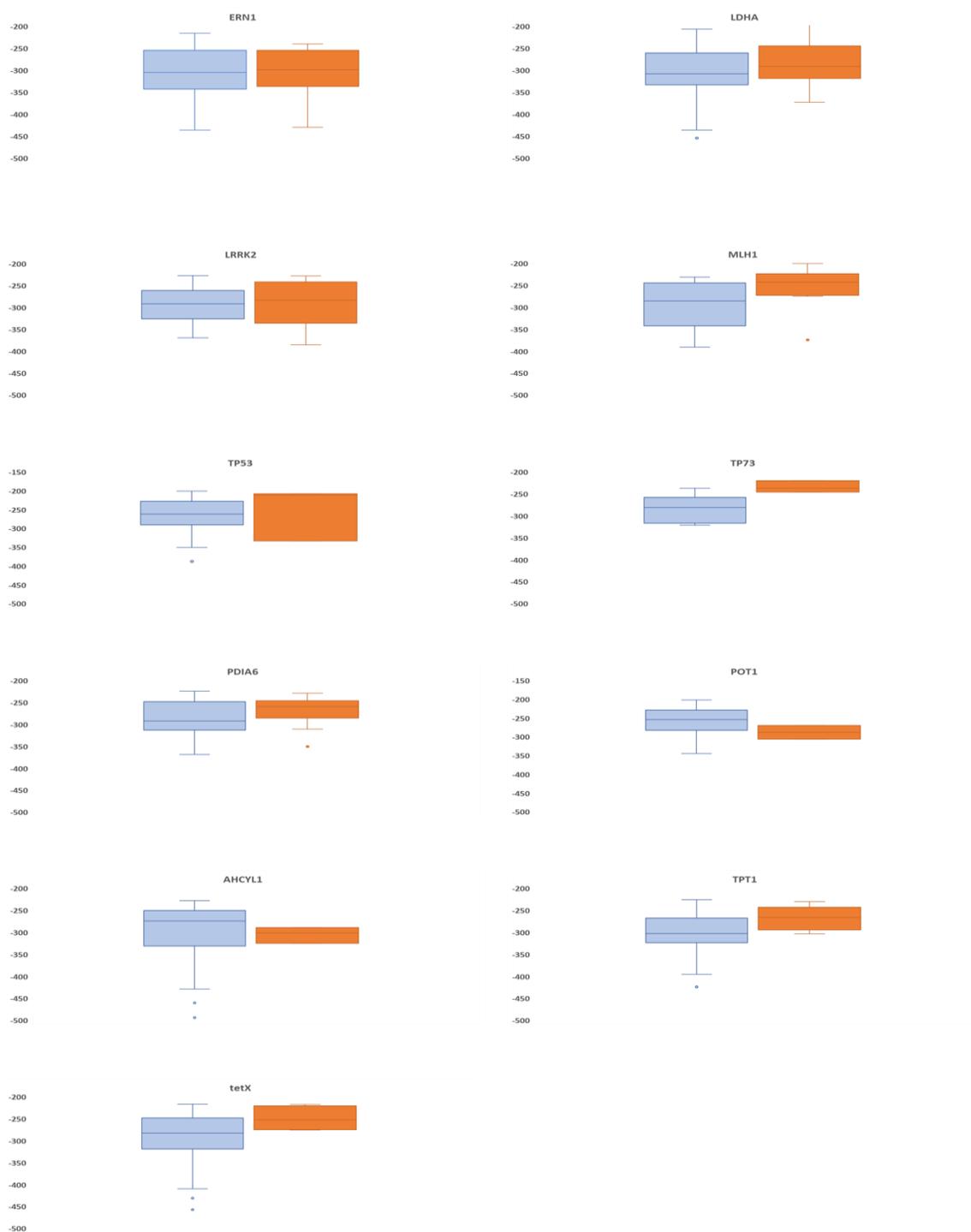


Figura 8. Gráficas de caja y bigotes de la energía electrostática de los resultados de docking. Se muestra la dispersión de la energía electrostática en los resultados de HADDOCK, en azul la energía de los acoplamientos entre la Hsp70 y los heptámeros de unión predichos en las proteínas cliente, en naranja la energía de los acoplamientos entre la Hsp70 y las IDPRs de las proteínas cliente.

Los rangos de los valores de energía electrostática son similares en ambos tipos de regiones evaluadas, sin embargo, el origen de las magnitudes de dichas energías puede variar, ya que pueden estar involucrados diferentes aminoácidos de contacto que favorezcan las interacciones electrostáticas. El origen de estos resultados se puede analizar mediante la obtención de los *fingerprint* de los contactos Hsp70-cliente.

6.7 Determinación del *fingerprint* de la interacción Hsp70-sustrato

A partir de los resultados de HADDOCK, se analizaron los contactos proteína-proteína con la herramienta *getcontacts* para cada uno de los *clusters* seleccionados acuerdo con el criterio mencionado previamente, de modo que para cada cluster de las regiones heptaméricas e IDPRs evaluadas, se obtuvo una lista de los aminoácidos involucrados en la interacción proteína-proteína en formato de texto (Figura 9. a).

Para evaluar si las regiones de contacto son similares, se optó por alinear las secuencias de aminoácidos de cada región evaluada, aunque estos no eran consecutivos. Este procedimiento permitió distinguir qué aminoácidos eran comunes dentro de las listas de contactos (Figura 9. b). Finalmente se obtuvo la frecuencia de aminoácidos a partir del consenso (Figura 9. c).

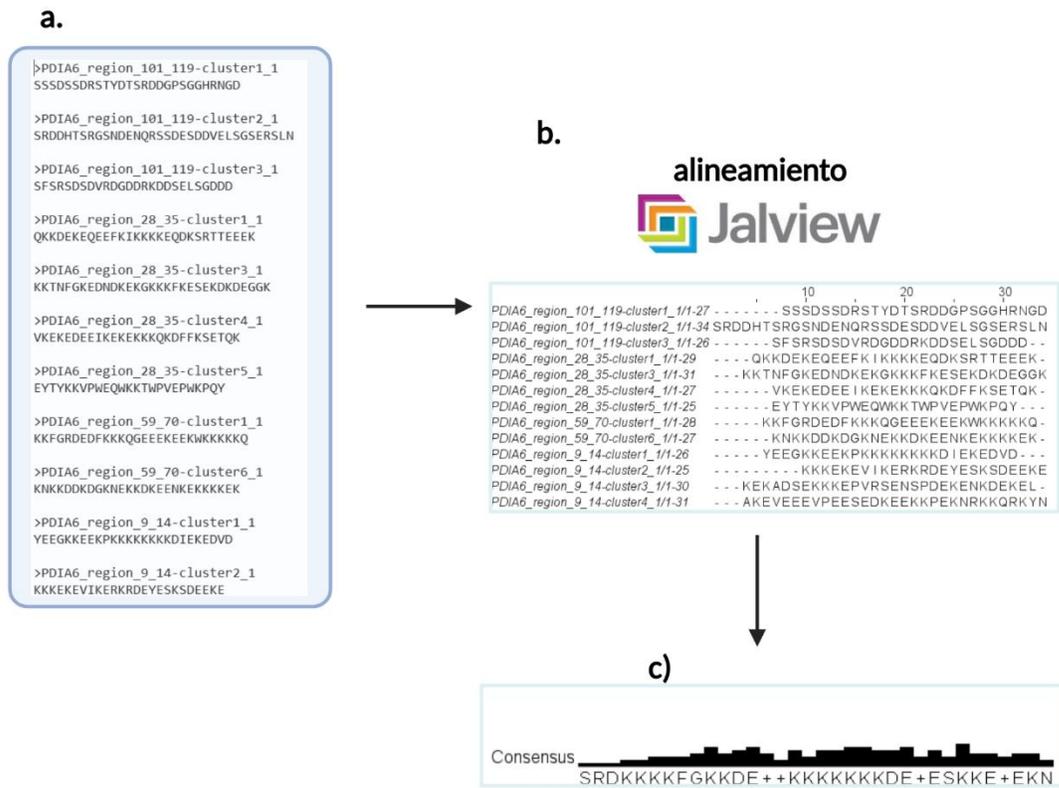
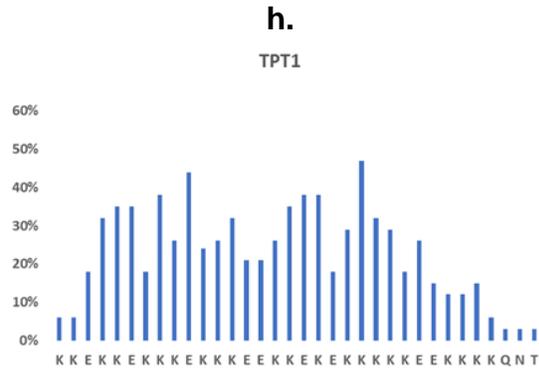
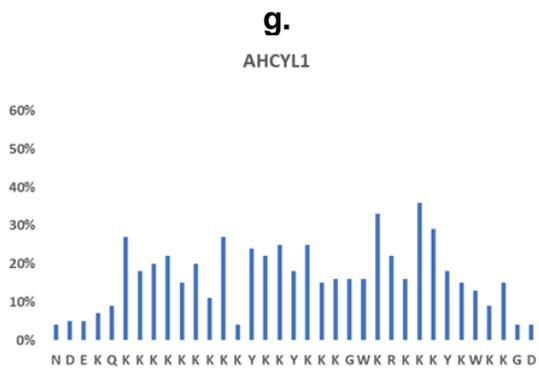
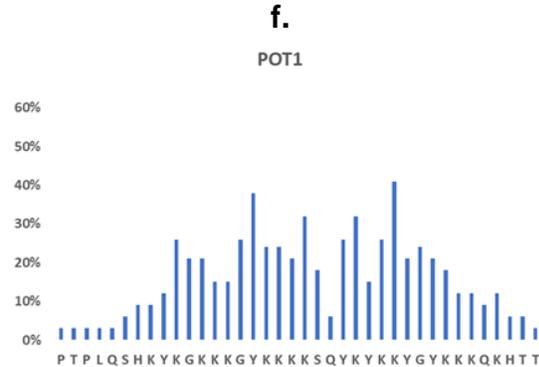
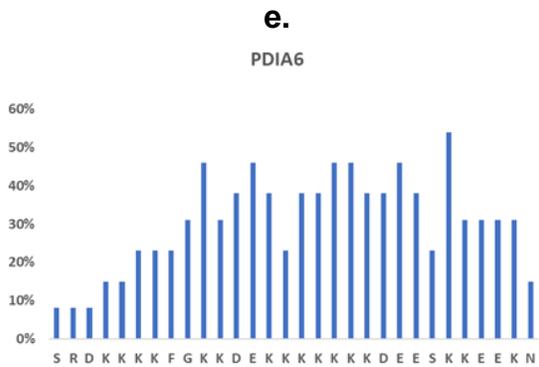
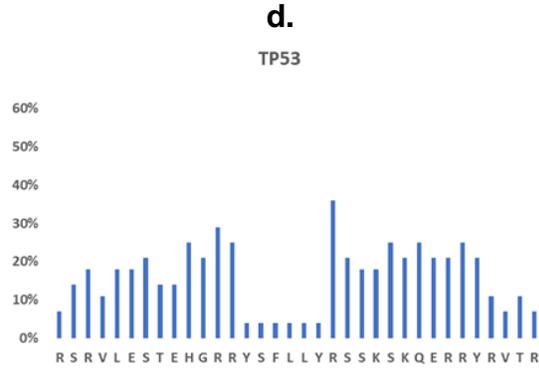
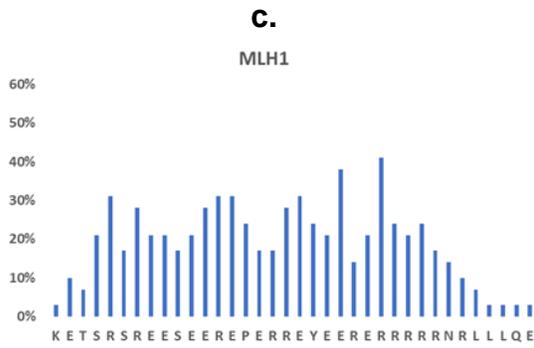
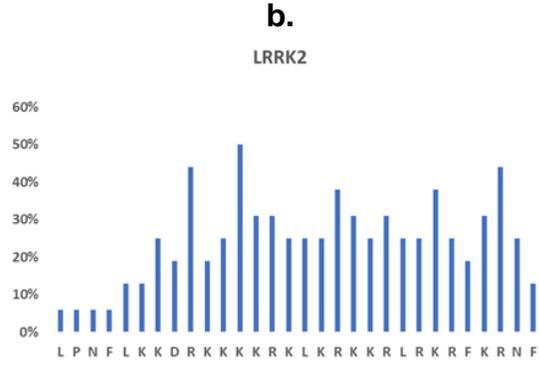
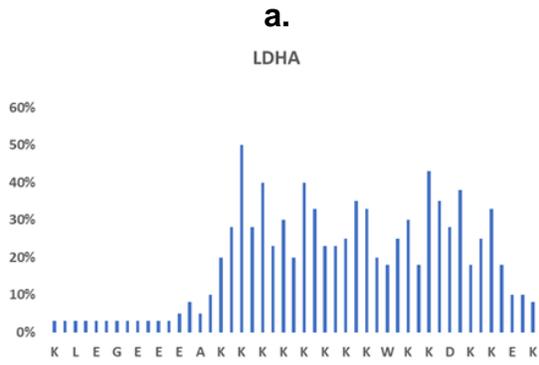
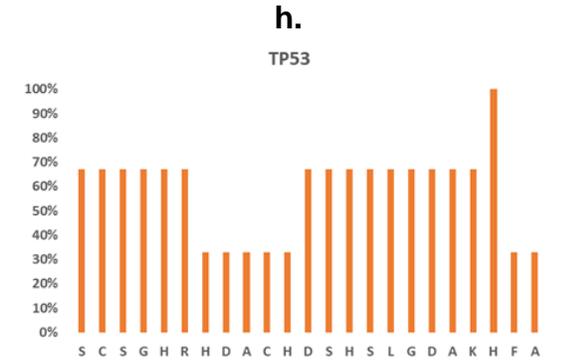
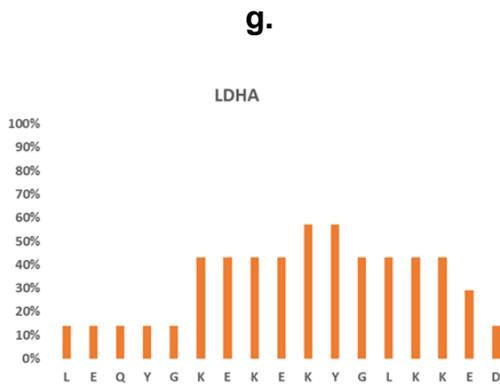
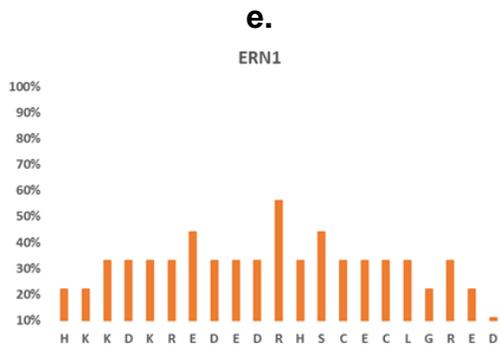
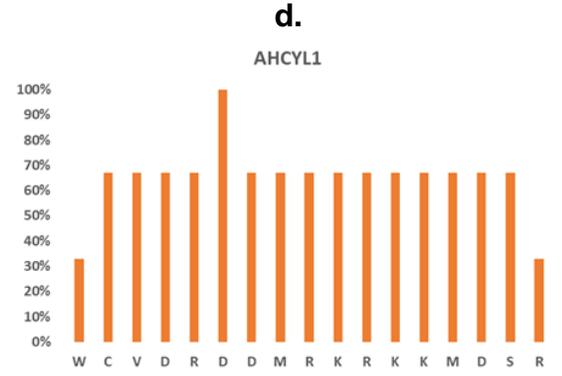
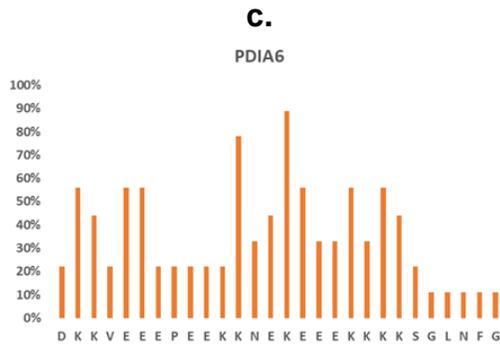
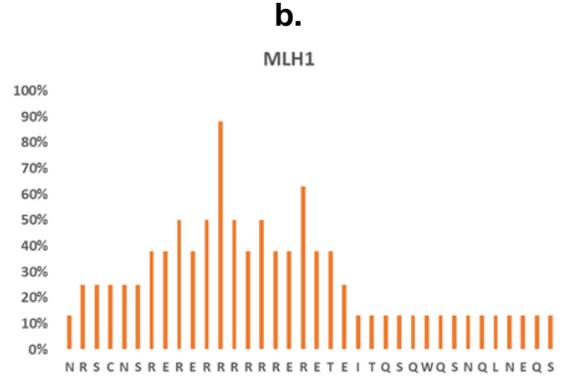
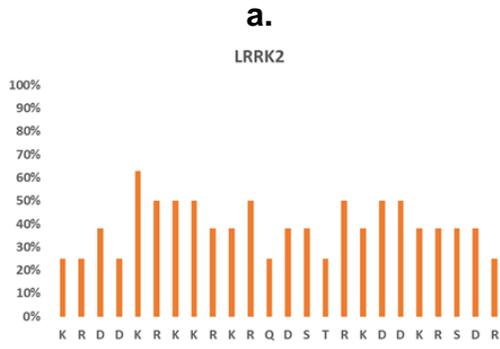


Figura 9. Determinación del *fingerprint* de la interacción. a. Lista de aminoácidos de las diferentes regiones de las proteínas cliente que hacen contacto con Hsp70. b. Alineamiento de secuencias con jalview. c. Secuencia consenso de aminoácidos de contacto de las proteínas cliente. **Creado con BioRender.com**

Para el conjunto de resultados de los heptámeros de interacción, la frecuencia de aminoácidos obtenida como una lista de contactos para los sitios probables de unión, muestra que los aminoácidos que predominan en varias posiciones de contacto son lisinas (K) y argininas (R) que en algunos casos representa de un 40 a 50% de la frecuencia en dicha posición, por ejemplo, en la [figura 10. a-h](#). En la [figura 10. i-j](#) aunque participan K y R no son predominantes y tienen bajas frecuencias (30% o menos), mientras que en la [figura 10. k](#) los aminoácidos de contacto son principalmente polares (K, E, Q, R) e hidrófobos (W, S, T), cuya frecuencia es semejante entre ellos. Esto demuestra que la característica predominante de la interacción de un cliente con Hsp70 es mediante contactos polares debido a una alta presencia de K y R, que podrían favorecer puentes salinos con sus contrapartes E404 y E460 o interacciones polares con aminoácidos de la región de unión de Hsp70. Lo anterior explica el origen de la magnitud de la energía electrostática en los resultados de HADDOCK.





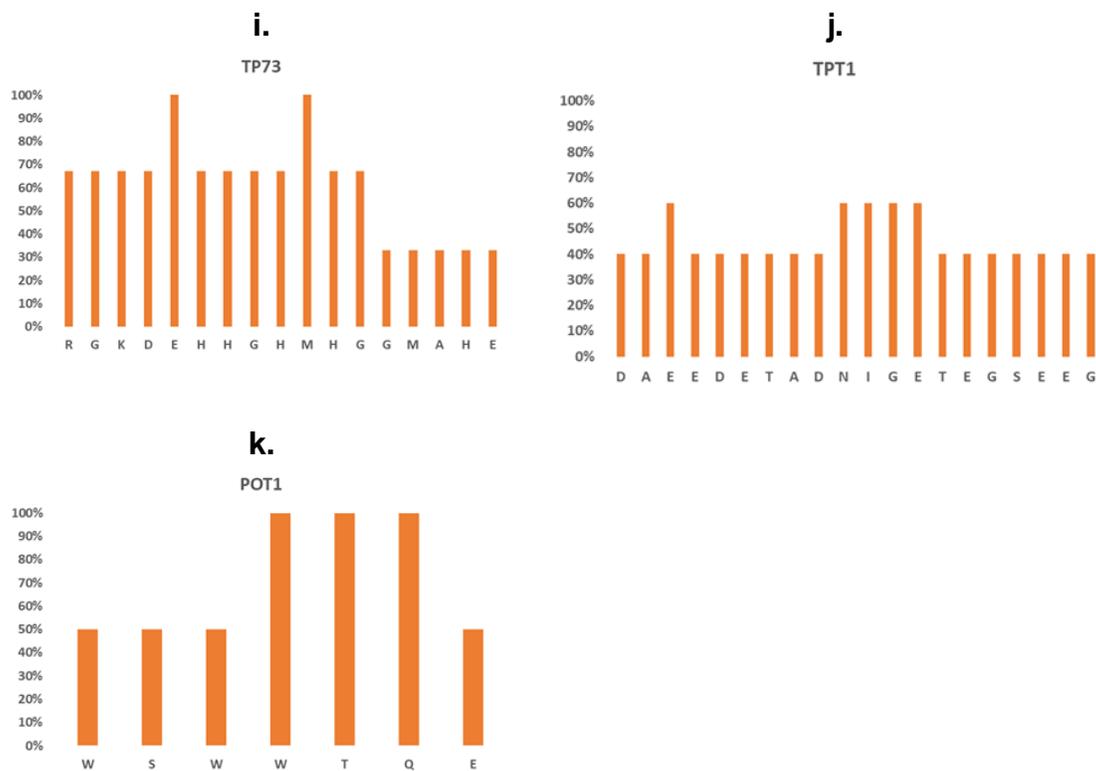


Figura 11. Fingerprint de la interacción Hsp70-cliente. Se muestran las gráficas de frecuencia de aminoácidos en las IDPRs. Se muestra la frecuencia de aminoácidos para las 11 proteínas cliente evaluadas en este estudio. El porcentaje mostrado corresponde al total de *clusters* evaluados.

En este punto podemos observar una diferencia entre las regiones evaluadas. Considerando que las interacciones electrostáticas tienen mucho peso en la interacción Hsp70-cliente y que las IDPRs no contienen tantos aminoácidos con carga, se podría pensar que la Hsp70 se uniría a estas regiones desordenadas con menor afinidad con respecto a otros sitios probables de unión conteniendo mayor número de aminoácidos cargados.

7. Conclusiones

Diversos estudios han tratado de esclarecer los mecanismos moleculares involucrados en la actividad biológica que tiene la Hsp70, tales como su participación en la homeostasis proteica y los mecanismos involucrados en la interacción con sus clientes. Sin embargo, aún hay diversos aspectos que no son del todo comprendidos. Uno de los avances más importantes en este trabajo fue aplicar una metodología que toma en cuenta la estructura tridimensional de las proteínas, ya que otros métodos que analizan la interacción están limitados solo por estudiar relaciones de secuencias entre las proteínas cliente y la Hsp70, lo cual no es suficiente.

La unión de Hsp70 con las IDPRs de las proteínas cliente no parece tener prioridad en comparación con otras regiones predichas. Es importante mencionar que nuestro estudio está limitado a la función de holdasa de la Hsp70. Por tal motivo no podemos determinar si las IDPRs no requieren de la asistencia de chaperonas o en qué condiciones puede presentarse esta interacción.

Entre los diferentes aspectos evaluados en este proyecto, destacan las características que presentan las regiones de interacción entre las proteínas que mejor se acoplan con el SDB de la Hsp70. La complementariedad geométrica muestra en su mayoría la participación de loops y regiones poco estructuradas, características que resultan compatibles para hacer interacción en la zona de la cavidad del SDB de la Hsp70. A la luz de nuestra estrategia de análisis, los resultados indican que las proteínas cliente deben presentar regiones poco estructuradas para interactuar con Hsp70.

Los análisis de contactos entre aminoácidos mediante técnicas de *fingerprint* presentaron un alto contenido de aminoácidos con carga o polares, principalmente K y R, con algunos casos en los que participan E y D, esto explica el origen de una fuerte interacción de tipo electrostática. Aunque no obtuvimos un patrón más detallado que permita identificar nuevos blancos potenciales de Hsp70.

8. Perspectivas

Un limitante de nuestro proyecto fue la poca información de las IDPRs, esto se debe a que la identificación de estas regiones es compleja mediante técnicas de cristalografía de rayos X. Por lo que sería necesario mejorar la estrategia de predicción de regiones desordenadas con el objetivo de aumentar la confianza de que una región identificada como IDPR lo sea realmente.

El estudio de las IDP/IDPR cada vez es más relevante, por tal motivo se están desarrollando diferentes herramientas dedicadas a analizar sus aspectos funcionales. Sería interesante realizar un estudio que involucre la determinación del orden intrínseco de estas proteínas, así como la susceptibilidad de estas regiones desordenadas ante distintos ambientes. Esto podría ayudar en la comprensión del papel que juega la interacción de Hsp70 con estas IDP/IDPR y en qué condiciones pudiera ser necesaria la función de foldasa u holdasa.

Muchas proteínas cliente quedaron fuera de los análisis realizados debido a que no se cuenta con estructuras 3D disponibles, algunas con gran relevancia en diversas patologías, sería interesante estudiar estas proteínas, evaluar la interacción con Hsp70 y asociar esta información con su papel biológico.

Referencias bibliográficas

- Alderson, T. R., Kim, J. H., Cai, K., Frederick, R. O., Tonelli, M., & Markley, J. L. (2014). Erratum: The specialized Hsp70 (HscA) interdomain linker binds to its nucleotide-binding domain and stimulates ATP hydrolysis in both cis and trans configurations (Biochemistry (2014) 53:46 (7148-7159) DOI: 10.1021/bi5010552). *Biochemistry*, 53(50), 7999. <https://doi.org/10.1021/bi501490v>
- Ayyadevara, S., Ganne, A., Balasubramaniam, M., & Shmookler Reis, R. J. (2021). Intrinsically disordered proteins identified in the aggregate proteome serve as biomarkers of neurodegeneration. *Metabolic Brain Disease*, 147–152. <https://doi.org/10.1007/s11011-021-00791-8>
- Bateman, A. (2019). UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Research*, 47(D1), D506–D515. <https://doi.org/10.1093/nar/gky1049>
- Bechtluft, P., Nouwen, N., Tans, S. J., & Driessen, A. J. M. (2010). SecB - A chaperone dedicated to protein translocation. *Molecular BioSystems*, 6(4), 620–627. <https://doi.org/10.1039/b915435c>
- Bell, S., Klein, C., Müller, L., Hansen, S., & Buchner, J. (2002). P53 Contains Large Unstructured Regions in Its Native State. *Journal of Molecular Biology*, 322(5), 917–927. [https://doi.org/10.1016/S0022-2836\(02\)00848-3](https://doi.org/10.1016/S0022-2836(02)00848-3)
- Bhattarai, A., & Emerson, I. A. (2020). Dynamic conformational flexibility and molecular interactions of intrinsically disordered proteins. *Journal of Biosciences*, 45(1), 1–17. <https://doi.org/10.1007/s12038-020-0010-4>
- Bondos, S. E., Dunker, A. K., & Uversky, V. N. (2021). On the roles of intrinsically disordered proteins and regions in cell communication and signaling. *Cell Communication and Signaling*, 19(1), 1–9. <https://doi.org/10.1186/s12964-021-00774-3>
- Bross, P., & Fernandez-Guerra, P. (2016). Disease-associated mutations in the HSPD1 gene encoding the large subunit of the mitochondrial HSP60/HSP10 chaperonin complex. *Frontiers in Molecular Biosciences*, 3(AUG), 1–7. <https://doi.org/10.3389/fmolb.2016.00049>
- Buchner, J. (2019). Molecular chaperones and protein quality control: An introduction to the JBC Reviews thematic series. *Journal of Biological Chemistry*, 294(6), 2074–2075. <https://doi.org/10.1074/jbc.REV118.006739>
- Burgess, A., Shah, K., Hough, O., & Hynynen, K. (2016). Specific binding of tetratricopeptide repeat (TPR) proteins to heat shock protein 70 (Hsp70) and heat shock protein 90 (Hsp90) is. *Biochemistry*, 15(5), 477–491. <https://doi.org/10.1021/acs.biochem.5b00801>. Specific
- Burmann, B. M., Gerez, J. A., Matečko-Burmann, I., Campioni, S., Kumari, P., Ghosh, D., Mazur, A., Aspholm, E. E., Šulskis, D., Wawrzyniuk, M., Bock, T.,

- Schmidt, A., Rüdiger, S. G. D., Riek, R., & Hiller, S. (2020). Regulation of α -synuclein by chaperones in mammalian cells. *Nature*, 577(7788), 127–132. <https://doi.org/10.1038/s41586-019-1808-9>
- Camberg, J. L., Doyle, S. M., Johnston, D. M., & Wickner, S. (2013). Molecular Chaperones. In *Brenner's Encyclopedia of Genetics: Second Edition* (Vol. 4). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-374984-0.00221-7>
- Carlisle, C., Prill, K., & Pilgrim, D. (2018). Chaperones and the proteasome system: Regulating the construction and demolition of striated muscle. *International Journal of Molecular Sciences*, 19(1), 1–22. <https://doi.org/10.3390/ijms19010032>
- Chen, J. (2016). The cell-cycle arrest and apoptotic and progression. *Cold Spring Harbor Perspectives in Biology*, 1–16.
- Christopher G. Evans, Lyra Chang, & Gestwicki, J. E. (2010). Heat Shock Protein 70 (Hsp70) as an Emerging Drug Target Introduction to Hsp70 Structure and Function. *Journal of Medicine Chemistry*, 53(12), 4585–4602. <https://doi.org/10.1021/jm100054f.Heat>
- Clerico, E. M., Tilitky, J. M., Meng, W., & Gierasch, L. M. (2015). How Hsp70 molecular machines interact with their substrates to mediate diverse physiological functions. *Journal of Molecular Biology*, 427(7), 1575–1588. <https://doi.org/10.1016/j.jmb.2015.02.004>
- Cuevas-Velazquez, C. L., Velloso, T., Guadalupe, K., Schmidt, H. B., Yu, F., Moses, D., Brophy, J. A. N., Cosio-Acosta, D., Das, A., Wang, L., Jones, A. M., Covarrubias, A. A., Sukenik, S., & Dinneny, J. R. (2021). Intrinsically disordered protein biosensor tracks the physical-chemical effects of osmotic stress on cells. *Nature Communications*, 12(1). <https://doi.org/10.1038/s41467-021-25736-8>
- Daugaard, M., Rohde, M., & Jäättelä, M. (2007). The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions. *FEBS Letters*, 581(19), 3702–3710. <https://doi.org/10.1016/j.febslet.2007.05.039>
- Dou, F., Netzer, W. J., Tanemura, K., Li, F., Hartl, F. U., Takashima, A., Gouras, G. K., Greengard, P., & Xu, H. (2003). Chaperones increase association of tau protein with microtubules. *Proceedings of the National Academy of Sciences of the United States of America*, 100(2), 721–726. <https://doi.org/10.1073/pnas.242720499>
- Dubrez, L., Causse, S., Borges Bonan, N., Dumétier, B., & Garrido, C. (2020). Heat-shock proteins: chaperoning DNA repair. *Oncogene*, 39(3), 516–529. <https://doi.org/10.1038/s41388-019-1016-y>
- Ekins, S., Mestres, J., & Testa, B. (2007). In silico pharmacology for drug discovery: Methods for virtual ligand screening and profiling. *British Journal of Pharmacology*, 152(1), 9–20. <https://doi.org/10.1038/sj.bjp.0707305>
- Ellis, J. (1988). Proteins as molecular chaperones. *Nature*, 328(6129), 378–379.

<https://doi.org/10.1038/328378a0>

- Fermeglia, M., Pricl, S., & Longo, G. (2003). Molecular modeling and process simulation: Real possibilities and challenges. *Chemical and Biochemical Engineering Quarterly*, 17(1), 19–29.
- Fernández-Fernández, M. R., & Valpuesta, J. M. (2018). Hsp70 chaperone: A master player in protein homeostasis [version 1; peer review: 3 approved]. *F1000Research*, 7, 1–10. <https://doi.org/10.12688/F1000RESEARCH.15528.1>
- Galea, C. A., Pagala, V. R., Obenauer, J. C., Park, C. G., Slaughter, C. A., & Kriwacki, R. W. (2006). Proteomic studies of the intrinsically unstructured mammalian proteome. *Journal of Proteome Research*, 5(10), 2839–2848. <https://doi.org/10.1021/pr060328c>
- Goloubinoff, P. (2017). Editorial: The HSP70 molecular chaperone machines. *Frontiers in Molecular Biosciences*, 4(JAN), 1–4. <https://doi.org/10.3389/fmolb.2017.00001>
- Gong, W., Hu, W., Xu, L., Wu, H., Wu, S., Zhang, H., Wang, J., Jones, G. W., & Perrett, S. (2018). The C-terminal GGAP motif of Hsp70 mediates substrate recognition and stress response in yeast. *Journal of Biological Chemistry*, 293(46), 17663–17675. <https://doi.org/10.1074/jbc.RA118.002691>
- Guerrero-Rojas, R., & Guerrero-Fonseca, C. (2018). Mecanismos moleculares de las proteínas de choque térmico (HSPs) implicados en el desarrollo neoplásico TT - Molecular mechanisms of heat shock proteins (HSPs) involved in neoplasm development. *Revista Salud Uninorte*, 34(2), 455–474. <https://doi.org/10.14482/sun.34.2.616.98>
- Gupta, A., Bansal, A., & Hashimoto-Torii, K. (2020). HSP70 and HSP90 in neurodegenerative diseases. *Neuroscience Letters*, 716(3), 139–148. <https://doi.org/10.1016/j.neulet.2019.134678>
- Gutierrez, M. B. B., Bonorino, C. B. C., & Rigo, M. M. (2020). ChaperISM: Improved chaperone binding prediction using position-independent scoring matrices. *Bioinformatics*, 36(3), 735–741. <https://doi.org/10.1093/bioinformatics/btz670>
- Hall, D. (2020). On the nature of the optimal form of the holdase-type chaperone stress response. *FEBS Letters*, 594(1), 43–66. <https://doi.org/10.1002/1873-3468.13580>
- Haslbeck, M., Weinkauff, S., & Buchner, J. (2019). Small heat shock proteins: Simplicity meets complexity. *Journal of Biological Chemistry*, 294(6), 2121–2132. <https://doi.org/10.1074/jbc.REV118.002809>
- Havalová, H., Ondrovičová, G., Keresztesová, B., Bauer, J. A., Pevala, V., Kutejová, E., & Kunová, N. (2021). Mitochondrial HSP70 chaperone system-the influence of post-translational modifications and involvement in human diseases. *International Journal of Molecular Sciences*, 22(15). <https://doi.org/10.3390/ijms22158077>

- Hayer-Hartl, M., Bracher, A., & Hartl, F. U. (2016). The GroEL-GroES Chaperonin Machine: A Nano-Cage for Protein Folding. *Trends in Biochemical Sciences*, 41(1), 62–76. <https://doi.org/10.1016/j.tibs.2015.07.009>
- Hipp, M. S., Kasturi, P., & Hartl, F. U. (2019). The proteostasis network and its decline in ageing. *Nature Reviews Molecular Cell Biology*, 20(7), 421–435. <https://doi.org/10.1038/s41580-019-0101-y>
- Hirsh, M. I., & Junger, W. G. (2008). Roles of heat shock proteins and $\gamma\delta$ T cells in inflammation. *American Journal of Respiratory Cell and Molecular Biology*, 39(5), 509–513. <https://doi.org/10.1165/rcmb.2008-0090TR>
- Hoter, A., El-Sabban, M. E., & Naim, H. Y. (2018). The HSP90 family: Structure, regulation, function, and implications in health and disease. *International Journal of Molecular Sciences*, 19(9). <https://doi.org/10.3390/ijms19092560>
- Kabani, M., & Martineau, C. (2008). Multiple Hsp70 Isoforms in the Eukaryotic Cytosol: Mere Redundancy or Functional Specificity? *Current Genomics*, 9(5), 338–348. <https://doi.org/10.2174/138920208785133280>
- Kanehisa, M., & Bork, P. (2003). Bioinformatics in the post-sequence era. *Nature Genetics*, 33(3S), 305–310. <https://doi.org/10.1038/ng1109>
- Karunanayake, C., & Page, R. C. (2021). Cytosolic protein quality control machinery: Interactions of Hsp70 with a network of co-chaperones and substrates. *Experimental Biology and Medicine*, 246(12), 1419–1434. <https://doi.org/10.1177/1535370221999812>
- Ko, S. K., Kim, J., Na, D. C., Park, S., Park, S. H., Hyun, J. Y., Baek, K. H., Kim, N. D., Kim, N. K., Park, Y. N., Song, K., & Shin, I. (2015). A small molecule inhibitor of ATPase activity of HSP70 induces apoptosis and has antitumor activities. *Chemistry and Biology*, 22(3), 391–403. <https://doi.org/10.1016/j.chembiol.2015.02.004>
- Kothandan, S., Sasikala, R. P., & Meena, K. S. (2017). Structure based Pharmacophore modeling, Virtual screening and Molecular Docking of Potential Phytochemicals against HSP70. *Journal of Applied Pharmaceutical Science*, 7(2), 137–141. <https://doi.org/10.7324/JAPS.2017.70218>
- Kulkarni, P., & Uversky, V. N. (2019). Intrinsically disordered proteins in chronic diseases. *Biomolecules*, 9(4), 1–6. <https://doi.org/10.3390/biom9040147>
- Kundel, F., De, S., Flagmeier, P., Horrocks, M. H., Kjaergaard, M., Shammass, S. L., Jackson, S. E., Dobson, C. M., & Klenerman, D. (2018). Hsp70 Inhibits the Nucleation and Elongation of Tau and Sequesters Tau Aggregates with High Affinity. *ACS Chemical Biology*, 13(3), 636–646. <https://doi.org/10.1021/acscchembio.7b01039>
- Lermyte, F. (2020). Roles, characteristics, and analysis of intrinsically disordered proteins: A minireview. *Life*, 10(12), 1–19. <https://doi.org/10.3390/life10120320>
- Liang, X., Zhu, W., Lv, Z., & Zou, Q. (2019). Molecular computing and bioinformatics.

- Molecules*, 24(13), 1–7. <https://doi.org/10.3390/molecules24132358>
- Lobanov, M. Y., Likhachev, I. V., & Galzitskaya, O. V. (2020). Disordered residues and patterns in the protein data bank. *Molecules*, 25(7), 1–14. <https://doi.org/10.3390/molecules25071522>
- Ma, F. H., Li, C., Liu, Y., & Shi, L. (2020). Mimicking Molecular Chaperones to Regulate Protein Folding. *Advanced Materials*, 32(3), 1–14. <https://doi.org/10.1002/adma.201805945>
- Macario, A. J. L., & de Macario, E. C. (2020). Molecular mechanisms in chaperonopathies: clues to understanding the histopathological abnormalities and developing novel therapies. *Journal of Pathology*, 250(1), 9–18. <https://doi.org/10.1002/path.5349>
- Macošek, J., Mas, G., & Hiller, S. (2021). Redefining Molecular Chaperones as Chaotropes. *Frontiers in Molecular Biosciences*, 8(June), 1–11. <https://doi.org/10.3389/fmolb.2021.683132>
- Martín-Peña, A., Rincón-Limas, D. E., & Fernandez-Fúnez, P. (2018). Engineered Hsp70 chaperones prevent A β 42-induced memory impairments in a *Drosophila* model of Alzheimer's disease. *Scientific Reports*, 8(1), 1–14. <https://doi.org/10.1038/s41598-018-28341-w>
- Mayer, M. P., & Gierasch, L. M. (2019). Recent advances in the structural and mechanistic aspects of Hsp70 molecular chaperones. *Journal of Biological Chemistry*, 294(6), 2085–2097. <https://doi.org/10.1074/jbc.REV118.002810>
- Nishimura, T., & Akiyoshi, K. (2020). Artificial Molecular Chaperone Systems for Proteins, Nucleic Acids, and Synthetic Molecules. *Bioconjugate Chemistry*, 31(5), 1259–1267. <https://doi.org/10.1021/acs.bioconjchem.0c00133>
- Oldfield, C. J., & Dunker, A. K. (2014). Intrinsically disordered proteins and intrinsically disordered protein regions. *Annual Review of Biochemistry*, 83, 553–584. <https://doi.org/10.1146/annurev-biochem-072711-164947>
- Pecina-Šlaus, N., Kafka, A., Salamon, I., & Bukovac, A. (2020). Mismatch Repair Pathway, Genome Stability and Cancer. *Frontiers in Molecular Biosciences*, 7(June), 1–12. <https://doi.org/10.3389/fmolb.2020.00122>
- Petrey, D., & Honig, B. (2014). Structural bioinformatics of the interactome. *Annual Review of Biophysics*, 43(1), 193–210. <https://doi.org/10.1146/annurev-biophys-051013-022726>
- Pratt, W. B., Gestwicki, J. E., Osawa, Y., & Lieberman, A. P. (2015). Targeting Proteostasis Through the Protein Quality Control Function of the Hsp90/Hsp70-based Chaperone Machinery for Treatment of Adult Onset Neurodegenerative Diseases. *Annual Review of Pharmacology and Toxicology*, 55, 353–371. <https://doi.org/10.1146/annurev-pharmtox-010814-124332>. Targeting
- Radons, J. (2016). The human HSP70 family of chaperones: where do we stand? *Cell Stress and Chaperones*, 21(3), 379–404. <https://doi.org/10.1007/s12192->

- Ritossa, F. (1962). A new puffing pattern induced by temperature shock and DNP in drosophila. *Experientia*, 18(12), 571–573. <https://doi.org/10.1007/BF02172188>
- Rodrigues, J. P. G. L. M., & Bonvin, A. M. J. J. (2014). Integrative computational modeling of protein interactions. *FEBS Journal*, 281(8), 1988–2003. <https://doi.org/10.1111/febs.12771>
- Rosano, G. L., Bruch, E. M., & Ceccarelli, E. A. (2011). Insights into the CLP/HSP100 chaperone system from chloroplasts of *Arabidopsis thaliana*. *Journal of Biological Chemistry*, 286(34), 29671–29680. <https://doi.org/10.1074/jbc.M110.211946>
- Rosenzweig, R., Nillegoda, N. B., Mayer, M. P., & Bukau, B. (2019). The Hsp70 chaperone network. *Nature Reviews Molecular Cell Biology*, 20(11), 665–680. <https://doi.org/10.1038/s41580-019-0133-3>
- Rüdiger, S., Germeroth, L., Schneider-Mergener, J., & Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries respect to the bound folding conformer are only partly. *The EMBO Journal*, 16(7), 1501–1507. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1169754/pdf/001501.pdf>
- Scalia, F., Vitale, A. M., Santonocito, R., de Macario, E. C., Macario, A. J. L., & Cappello, F. (2021). The neurochaperonopathies: Anomalies of the chaperone system with pathogenic effects in neurodegenerative and neuromuscular disorders. *Applied Sciences (Switzerland)*, 11(3), 1–22. <https://doi.org/10.3390/app11030898>
- Schwede, T. (2013). Protein modeling: What happened to the “protein structure gap”? *Structure*, 21(9), 1531–1540. <https://doi.org/10.1016/j.str.2013.08.007>
- Siam, M. K. S., Karim, A., & Shohan, M. U. S. (2020). In-Silico Study for Potential Inhibitors of Both HSP72 and HSC70 Proteins in the Treatment of Cancer. *PervasiveHealth: Pervasive Computing Technologies for Healthcare*, 61–67. <https://doi.org/10.1145/3429210.3429226>
- Singh, R., Gurao, A., Rajesh, C., Mishra, S. K., Rani, S., Behl, A., Kumar, V., & Kataria, R. S. (2019). Comparative modeling and mutual docking of structurally uncharacterized heat shock protein 70 and heat shock factor-1 proteins in water buffalo. *Veterinary World*, 12(12), 2036–2045. <https://doi.org/10.14202/vetworld.2019.2036-2045>
- Sousa, R. (2014). Structural mechanisms of chaperone mediated protein disaggregation. *Frontiers in Molecular Biosciences*, 1(SEP), 1–18. <https://doi.org/10.3389/fmolb.2014.00012>
- Tittelmeier, J., Nachman, E., & Nussbaum-Krammer, C. (2020). Molecular Chaperones: A Double-Edged Sword in Neurodegenerative Diseases. *Frontiers in Aging Neuroscience*, 12(October), 1–21.

<https://doi.org/10.3389/fnagi.2020.581374>

- Truttmann, M. C., Zheng, X., Hanke, L., Damon, J. R., Grootveld, M., Krakowiak, J., Pincus, D., & Ploegh, H. L. (2017). Unrestrained AMPylation targets cytosolic chaperones and activates the heat shock response. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(2), E152–E160. <https://doi.org/10.1073/pnas.1619234114>
- Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Židek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G. J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S. A. A., ... Hassabis, D. (2021). Highly accurate protein structure prediction for the human proteome. *Nature*, *596*(7873), 590–596. <https://doi.org/10.1038/s41586-021-03828-1>
- Uemura, E., Niwa, T., Minami, S., Takemoto, K., Fukuchi, S., Machida, K., Imataka, H., Ueda, T., Ota, M., & Taguchi, H. (2018). Large-scale aggregation analysis of eukaryotic proteins reveals an involvement of intrinsically disordered regions in protein folding. *Scientific Reports*, *8*(1), 1–11. <https://doi.org/10.1038/s41598-017-18977-5>
- Uversky, V. N. (2016). p53 proteoforms and intrinsic disorder: An illustration of the protein structure–function continuum concept. *International Journal of Molecular Sciences*, *17*(11). <https://doi.org/10.3390/ijms17111874>
- Uversky, V. N. (2019). Intrinsically disordered proteins and their “Mysterious” (meta)physics. *Frontiers in Physics*, *7*(FEB), 8–23. <https://doi.org/10.3389/fphy.2019.00010>
- V. Sudnitsyna, M., V. Mymrikov, E., S. Seit-Nebi, A., & B. Gusev, N. (2012). The Role of Intrinsically Disordered Regions in the Structure and Functioning of Small Heat Shock Proteins. *Current Protein & Peptide Science*, *13*(1), 76–85. <https://doi.org/10.2174/138920312799277875>
- Van Durme, J., Maurer-Stroh, S., Gallardo, R., Wilkinson, H., Rousseau, F., & Schymkowitz, J. (2009). Accurate prediction of DnaK-peptide binding via homology modelling and experimental data. *PLoS Computational Biology*, *5*(8). <https://doi.org/10.1371/journal.pcbi.1000475>
- Van Zundert, G. C. P., Rodrigues, J. P. G. L. M., Trellet, M., Schmitz, C., Kastitis, P. L., Karaca, E., Melquiond, A. S. J., Van Dijk, M., De Vries, S. J., & Bonvin, A. M. J. J. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *Journal of Molecular Biology*, *428*(4), 720–725. <https://doi.org/10.1016/j.jmb.2015.09.014>
- Vangone, A., Rodrigues, J. P. G. L. M., Xue, L. C., van Zundert, G. C. P., Geng, C., Kurkcuoglu, Z., Nellen, M., Narasimhan, S., Karaca, E., van Dijk, M., Melquiond, A. S. J., Visscher, K. M., Trellet, M., Kastitis, P. L., & Bonvin, A. M. J. J. (2017). Sense and simplicity in HADDOCK scoring: Lessons from CASP-CAPRI round 1. *Proteins: Structure, Function and Bioinformatics*, *85*(3), 417–423.

<https://doi.org/10.1002/prot.25198>

- Vrontou, E., & Economou, A. (2004). Structure and function of SecA, the preprotein translocase nanomotor. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1694(1-3 SPEC.ISS.), 67–80. <https://doi.org/10.1016/j.bbamcr.2004.06.003>
- Wallin, S. (2017). Intrinsically disordered proteins: structural and functional dynamics. *Research and Reports in Biology*, Volume 8, 7–16. <https://doi.org/10.2147/rrb.s57282>
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., De Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1), W296–W303. <https://doi.org/10.1093/nar/gky427>
- Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., & Hendrickson, W. A. (1996). Structural Analysis of Substrate Binding by the Molecular Chaperone DnaK. *Science*, 272(5268), 1606–1614. <https://doi.org/10.1126/science.272.5268.1606>
- Zuiderweg, E. R. P., Hightower, L. E., & Gestwicki, J. E. (2017). The remarkable multivalency of the Hsp70 chaperones. *Cell Stress and Chaperones*, 22(2), 173–189. <https://doi.org/10.1007/s12192-017-0776-y>