

**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL**

CINVESTAV IRAPUATO

UNIDAD DE GENÓMICA AVANZADA (LANGEBIO)

**“Diseño de la interfaz p53-MDM2 sobre un andamio proteico del tipo
inmunoglobulina”**

Tesis que presenta

M. C. Pedro Jiménez Sandoval

Para obtener el grado de

Doctor en Ciencias

En la especialidad de

Biología de Plantas

Director de Tesis

Dr. Luis Gabriel Briebe de Castro

Irapuato, Guanajuato

Agosto 2018

El presente trabajo de investigación se desarrolló en el Laboratorio de Bioquímica Estructural de la Unidad de Genómica Avanzada (UGA-Langebío) del Cinvestav bajo la dirección del Dr. Luis Gabriel Briebe de Castro, Profesor Titular del Langebio y con asesoría de la Dra. Laila Pamela Partida Martínez, Profesora Titular del Departamento de Ingeniería Genética de la Unidad Irapuato, la Dra. Silvia Edith Valdés Rodríguez y el Dr. Alejandro Blanco Labra, ambos Profesores del Departamento de Biotecnología y Bioquímica de la Unidad Irapuato, el Dr. José Juan Ordaz Ortiz, Profesor Adjunto del Langebio y el Dr. Samuel Lara González, Investigador Titular de la División de Biología Molecular del IPICYT.

Agradecimientos

Agradezco al CONACyT por la beca número 282042 otorgada durante la realización de mis estudios de posgrado en el programa de Doctorado en Ciencias en la especialidad de Biotecnología de Plantas en el Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional Sede Irapuato, en el periodo comprendido entre septiembre del 2014 a agosto del 2018.

Este trabajo fue financiado por el proyecto “Fronteras de la Ciencia CONACyT #11” otorgado al Dr. Luis Gabriel Brieba de Castro.

Agradezco al Dr. Lino Sánchez por su ayuda en los análisis de microscopía.

Al laboratorio del Dr. Alfredo Torres Larios por el apoyo en los ensayos de interferometría de biocapa.

Al Dr. Daniel Maturana de NanoTemper, por el apoyo en los ensayos de termoforesis.

A la M.C. Itzel Rentería y a la Dra. Claudia Benítez por el apoyo en los ensayos de dicroísmo circular.

Así mismo agradezco a la I. A. Corina Elizabeth Díaz Quezada, auxiliar del Laboratorio de Bioquímica Estructural, por su asistencia y apoyo durante el desarrollo de este proyecto.

También agradezco a los compañeros de administración, apoyo, mantenimiento, seguridad, transporte, limpieza y cafetería, por su enorme trabajo, dedicación y siempre tener una sonrisa para mí.

Agradezco especialmente:

A mi jefe el Dr. Luis Brieba, por la libertad que me dio para desarrollar el trabajo en el laboratorio, por impulsar mi desarrollo profesional con tanto ahínco, por la confianza, por sus enseñanzas y por su amistad.

A mi comité tutorial, por mostrarme otros puntos de vista en la realización del proyecto, por sus excelentes consejos y sugerencias tanto académicas como personales y por compartir la emoción de los logros obtenidos durante esta etapa en que ayudaron a mi formación. Muchas gracias.

A todos mis compañeros del Laboratorio de Bioquímica Estructural, tuve la suerte de convivir desde la maestría con varias generaciones de estudiantes, de todos me llevo algo, se les estima.

A mis cuates de aquí y de allá, compañeros fieles de viajes, fiestas y reuniones. Tan alegres que contagian.

A todos mis roomies, la familia en casa.

A mi familia que es grande no sólo en número, abuelos, tíos, primos. Siempre que los veo me siento el favorito, así de buenos son. Gracias por todo.

A mis hermanas Mariana y Diana, a Fernando y Nataly, por las risas, apoyo y cariño.

A mis padres Martha y Virgilio, por enseñarme a ser curioso, por enseñarme a trabajar, por enseñarme a compartir, por darme todo y cuidar que nada me falte.

Índice General

Resumen.....	xi
Abstract	xii
I. Introducción	1
1.1 La ingeniería de proteínas.....	1
1.2 Uso de andamios proteicos para el diseño racional de proteínas	2
II. Hipótesis.....	7
III. Objetivos.....	8
3.1 Objetivo general	8
3.2 Objetivos específicos.....	8
IV. Materiales y Métodos.....	9
4.1 Obtención del modelo de homología, diseño racional y predicción del plegamiento <i>in silico</i> . 9	
4.2 Expresión y purificación de las proteínas recombinantes	9
4.3 Determinación y separación de los complejos formados	10
4.4 Cristalización y colección de datos de difracción.	10
4.5 Determinación de las constantes de disociación de los complejos.	11
4.6 Determinación de la estabilidad térmica de las quimeras.....	12
4.7 Solución estructural y refinamiento del complejo.	12
4.8 Formación de fibras en EhICP1 silvestre y análisis por microscopía electrónica de transmisión	13
V. Resultados.....	14
5.1 Los ICPs tienen un asa conservada y dos asas variables expuestas al solvente.....	14
5.2 La versión mutante de EhICP1 libre de cisteínas no forma agregados moleculares y es más estable que el EhICP2 y chagasina	15

5.3 Las asas expuestas al solvente de EhICP1 tienen regiones que pueden ser modificadas sin comprometer el plegado global.....	18
5.3.1 EhICP1 tolera modificaciones en las asas DE y FG pero no en BC.....	20
5.4 Las quimeras EhCIP1-p53(DE) y EhICP1-p53(FG) forman complejo con MDM2	22
5.5 Cristalización del complejo EhICP1-p53(FG)/MDM2	24
5.6 Detalle del modelo cristalográfico y densidad electrónica del complejo EhICP1-p53(FG)/MDM2	25
5.7 Los complejos formados entre las quimeras de EhICP1-p53 y MDM2 muestran una afinidad cercana a la de la interacción canónica p53-MDM2	27
5.8 La estructura cristalográfica del complejo EhICP1-p53(FG)/MDM2 muestra un heterotetrámero formado por 2 moléculas de EhICP1-p53(FG) interconectadas, y 2 moléculas de MDM2.....	29
5.9 El EhICP1 forma agregados de alto peso molecular y fibrillas del tipo amiloide.....	34
VI. Discusión.....	35
VII. Conclusiones y Perspectivas.....	38
VIII. Apéndice A	39
IX. Bibliografía.....	43

Índice de Tablas y Figuras

Fig. 1. Representación del dominio tipo Ig.....	3
Fig. 2. Proteínas utilizadas como andamios para el diseño racional y utilizando técnicas de evolución dirigida para la obtención de proteínas con funciones de reconocimiento.....	5
Fig. 3. Detalle estructural de las asas en los ICPs.....	15
Fig. 4. Formación de oligómeros y agregados de alto peso molecular en EhICP1 en condiciones no reductoras.....	16
Fig. 5. Análisis estructural del EhICP1 como andamio proteico.....	19
Fig. 6. Modelos de homología de las quimeras de EhICP1-p53.....	20
Fig. 7. Comparación de la estabilidad térmica de las quimeras DE y FG con EhICP1.....	21
Fig. 8. Interacción de las quimeras EhICP1-p53 con MDM2.....	23
Fig. 9. Mejoramiento de la cristalización del complejo EhICP1-p53(FG)/MDM2.....	25
Fig. 10. Patrón de difracción y características de la colecta de datos.....	26
Fig. 11. Detalle del modelo estructural y mapa de densidad electrónica del complejo EhICP1-p53(FG)/MDM2.....	27
Fig. 12. Afinidad de las quimeras de EhICP1-p53 contra MDM2.....	28
Fig. 13. Estructura cristalográfica del complejo EhICP1-p53(FG)/MDM2.....	31
Fig. 14. Comparación estructural entre los complejos EhICP1-p53(FG)/MDM2 y un anticuerpo contra MDM4.....	33
Fig. 15. Análisis de agregados moleculares y formación de fibras en EhICP1 silvestre por TEM.....	35
Tabla 1. Estabilidad térmica de ICPs por dicroísmo circular.....	17
Tabla 2. Correlación entre la probabilidad del modelo estructural con la estabilidad térmica y el cambio en la energía libre de Gibbs para EhICP1 y las quimeras BC, DE y FG.....	22
Tabla 3. Constantes de disociación para las quimeras de EhICP1-p53 determinadas por interferometría de biocapa.....	29
Tabla 5. Tabla de aminoácidos y nomenclatura.....	42

Abreviaturas

Å	ångström, unidad de longitud
AP	aptámero proteico
APs	aptámeros peptídicos
BLI	interferometría de biocapa
CD	dicroísmo celular
CD2	cúmulo de diferenciación 2
CD4	cúmulo de diferenciación 4
CP	cisteín preoteasa
DSF	fluorimetría diferencial de barrido
DTT	ditiotreitól
EDTA	ácido etilendiaminotetraacético
EhICP1	inhibidor de cisteín proteasas 1 de <i>Entamoeba histolytica</i>
EhICP2	inhibidor de cisteín proteasas 2 de <i>Entamoeba histolytica</i>
ICP	inhibidor de cisteín proteasas
ICPs	inhibidores de cisteín proteasas
Ig	inmunoglobulina
IgG1	inmunoglobulina G subclase 1
Igs	inmunoglobulinas
IgSF	súper familia de las inmunoglobulinas
IMAC	cromatografía de afinidad a metales

IPP	interacciones proteína-proteína
IPTG	isopropil-β-D-1-tiogalactopiranosido
MDM2	proteína doble minuto murino 2
MHC	complejo mayor de histocompatibilidad
MR	reemplazo molecular
MST	termoforesis de microescala
PDB	Protein Data Bank
pI	punto isoeléctrico
PMI	péptido sintético de p53, inhibidor de alta afinidad
SDS	dodecilsulfato sódico
TCEP	Tris(2-carboxietil)fosfina
TCR	receptores de linfocitos T
TEM	microscopía electrónica de transmisión
Tris	tris(hidroximetil)aminometano

Resumen

Las interacciones proteína-proteína juegan un papel crucial en las funciones biológicas y en la actualidad existe abundante información estructural de diversos sistemas. Dichas interacciones se dan mediante contactos físicos de alta especificidad entre dos o más moléculas proteicas, como resultado de eventos bioquímicos dirigidos por fuerzas electrostáticas. La habilidad para reproducir o imitar estas interfaces representa una poderosa herramienta biotecnológica, con repercusiones industriales, de investigación y terapéuticas. En los últimos años los esfuerzos por mimetizar interfaces moleculares han resultado en el uso de proteínas con características especiales que las hacen aptas para el desarrollo de proteínas con características novedosas, estas proteínas se conocen como andamios proteicos. La naturaleza ha desarrollado una gran diversidad de proteínas basadas en sus propios andamios, un ejemplo claro son los anticuerpos.

El dominio del tipo inmunoglobulina es uno de los plegamientos estructurales de proteínas más abundantes en la naturaleza. La robustez de este plegamiento ha permitido la diversificación funcional en proteínas tales como co-receptores, resortes moleculares, anticuerpos, inhibidores, etc. Un ejemplo son los inhibidores de cisteína proteasas (ICP) de la familia MEROPS I42. En este trabajo se utilizó a un miembro de esta familia (EhICP1) como andamio para el diseño de una interfaz que mimetiza la interacción entre p53 y el biomarcador MDM2. Los ensayos de interacción proteína-proteína y la estructura cristalográfica demuestran que los ICPs pueden ser utilizados para el diseño racional de proteínas con características deseadas.

Los detalles estructurales del complejo generado sugieren que el EhICP1 puede formar agregados estructurados de alto peso molecular, que fueron confirmados por microscopía electrónica de transmisión.

Abstract

Protein-protein interactions play a pivotal role in many biological functions. During the last decade, abundantly structural information about protein interactions from various biological systems has become available. These interactions occur through high specific physical contacts between two or more protein molecules, as a result of biochemical events driven by electrostatic forces. The ability to reproduce or mimic these interfaces represents a powerful biotechnological tool, with outcomes in industry, research and therapeutic use. In recent years, the efforts to mimic molecular interfaces have resulted in the use of proteins with special features that make them suitable for their use as a protein scaffolds, for the design of proteins with novel properties. Nature has developed an overwhelming diversity of proteins based on its own scaffolds, a clear example are antibodies.

Immunoglobulin-like domain is one of the most abundant protein folds in nature. The robustness of this fold has allowed functional diversification in proteins such as co-receptors, molecular springs, antibodies, inhibitors, etc. An example is the cysteine protease inhibitors (ICP) of the MEROPS I42 family. In this work a member of this family (EhICP1) was used as a scaffold for the design of an interface that mimics the interaction between p53 and MDM2. The protein-protein interaction results and the crystallographic structure showed that the ICPs can be used for the rational design of proteins with customized properties.

The structural details of the crystal complex suggest that EhICP1 can form structured aggregates of high molecular weight as confirmed by electron transmission microscopy.

I. Introducción

Las interacciones proteína-proteína (IPP) son contactos físicos de alta especificidad entre dos o más moléculas proteicas, como resultado de eventos bioquímicos dirigidos por fuerzas electrostáticas. Estas interacciones tienen una dinámica e intrincada fisicoquímica molecular para llevar a cabo las funciones celulares [1]. El mapa completo de las interacciones proteicas que pueden ocurrir en un organismo vivo se conoce como el interactoma [2]. Descifrar el interactoma de un organismo se ha convertido en uno de los principales objetivos de las actuales investigaciones biológicas, donde la bioquímica estructural ha hecho importantes contribuciones, permitiéndonos entender la relación entre la estructura y la función de las proteínas.

1.1 La ingeniería de proteínas

La ingeniería de proteínas es un método para la producción de proteínas con mejores o nuevas propiedades, y que pueden ser utilizadas con potenciales aplicaciones biotecnológicas [3]. Se basa en el conocimiento sobre plegamiento de proteínas, en la información de interfaces proteicas, en los principios del diseño racional o en la generación de bibliotecas, combinando evolución dirigida con los correspondientes mapeos o escrutinios para la selección de la propiedad deseada [4].

En el diseño racional se utiliza el conocimiento detallado de la estructura y función de la proteína para hacer las modificaciones deseadas. A pesar de que las proteínas son marginalmente estables, éstas pueden tolerar la sustitución de múltiples aminoácidos y mantener su estructura nativa o plegamiento [5]. Los primeros trabajos en el diseño racional de proteínas se hicieron modificando las asas de la tioredoxina de *Escherichia coli*, para mimetizar o imitar la función de reconocimiento de la región determinante de la complementariedad de las inmunoglobulinas (Ig) y llamaron a esas proteínas modificadas “aptámeros peptídicos” (APs). Los APs son una clase de proteínas de reconocimiento basadas en un “andamio proteico” que restringe la libertad conformacional y exponen una secuencia peptídica al solvente, que es la responsable de la función de reconocimiento hacia la proteína interactora o blanco [6].

1.2 Uso de andamios proteicos para el diseño racional de proteínas

Los andamios proteicos o APs se asemejan a los anticuerpos, en el sentido de que la secuencia de reconocimiento se encuentra restringida conformacionalmente por el resto de la proteína o dominio. Los anticuerpos son proteínas modificadas por la naturaleza a partir de un andamio conservado, son las proteínas de alta especificidad históricamente usadas para aplicaciones de reconocimiento, sin embargo, los anticuerpos tienen ciertas limitaciones, como el tamaño (>100 kDa), la tendencia a formar agregados y requerir la formación de puentes disulfuro para su plegamiento [7]. A pesar de que los anticuerpos son difíciles de expresar y purificar de bacteria [8], la obtención de algunas versiones se ha logrado mediante métodos de selección *in vitro* y mediante la manipulación de las cepas bacterianas [9, 10].

En los últimos años se ha utilizado un número de proteínas como andamios para el diseño racional de interfaces entre IPP, entre las que se incluyen el dominio III de fibronectina, el dominio A (*affibody*), los dominios PDZ y SH3, el dominio repetido de ankirina (*darpins*), tioredoxina, la proteína verde fluorescente, los receptores de las células T, entre otros [6, 11-16]. Numerosos estudios usan diversas proteínas como APs para introducir secuencias peptídicas, generalmente con un tamaño de 4 a 20 aminoácidos [17, 18]. Las características deseadas en los APs son: tolerancia a la sustitución de aminoácidos, termoestabilidad, alta expresión en bacteria y baja o nula toxicidad.

Como se mencionó con anterioridad, el diseño racional utiliza en detalle el conocimiento sobre el plegamiento de proteínas. El plegamiento del tipo Ig es uno de los motivos estructurales más comunes en la naturaleza [19, 20]. Este plegamiento consiste de dos hojas β formadas de 7 a 8 hebras β antiparalelas [21], **Fig. 1**, numerosos estudios de plegamiento de proteínas se han realizado en este dominio, indicando los residuos que son importantes para adoptar la conformación nativa, así como aquellos que no son tan relevantes para adoptar dicha conformación [22-24].

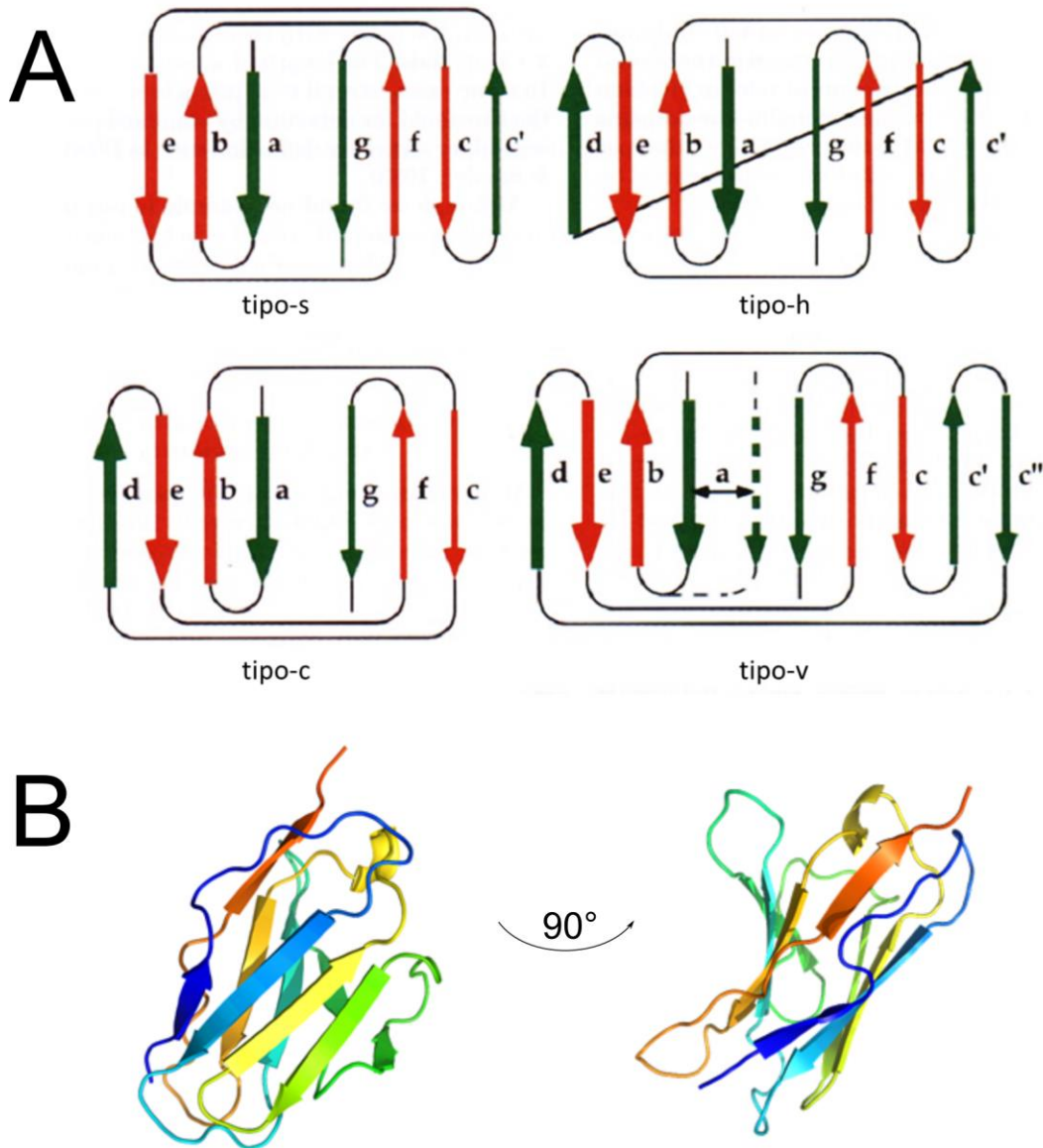


Fig. 1. Representación del dominio tipo Ig. **A** Diagramas de topología de los patrones de los puentes de hidrógeno observados para los subtipos del plegamiento tipo inmunoglobulina. Las 7-9 hebras (*a, b, c, c', c'', d, e, f, g*) forman un sándwich de 2 hojas β (hoja inferior, flechas delgadas; hoja superior, flechas gruesas). Las hebras que contienen los residuos del núcleo conservado se muestran en rojo. El dominio tipo inmunoglobulina consta invariablemente de 7 hebras β (tipo c). Los dominios tipo inmunoglobulina variables tienen adicionalmente un asa (*c'-c''*) entre las hebras *c* y *d*, con un total de 9 hebras (tipo v). En el tipo s, la cuarta hebra β se ha cambiado de hoja. El tipo h, es un híbrido entre el tipo c y el s, tiene una hebra extra interconectada, haciendo que ambas hojas tengan 4 hebras β . **B** Diagramas de cintas del dominio tipo Ig de la cadena ligera del IgG1. La estructura es semejante a un sándwich constituido por 2 hojas β , formadas por hebras β antiparalelas (tipo v) [19].

El dominio Ig se encuentra en otras proteínas que no son anticuerpos, por ejemplo, la proteína de la membrana externa del género *Leptospira* (LigB) [21], el dominio tipo III de fibronectina (FN3) [25] y los inhibidores de cisteína proteasas (ICP) de la familia MEROPS I42 [26-29], entre muchos otros. El ejemplo más notable de un AP con un plegamiento del tipo Ig es el FN3, este dominio ha sido utilizado como andamio para el desarrollo de proteínas de unión a una variedad de blancos como integrinas, lisozima, Src SH3 y el péptido ARVCF [13, 30-34]. En la **Fig. 2** se ilustran andamios del tipo Ig y andamios con un plegado diferente al de Ig, cuyas modificaciones dieron como resultado la formación de complejos moleculares y de los cuales se ha resuelto la estructura cristalográfica.

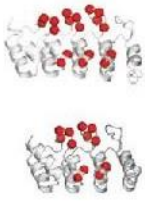
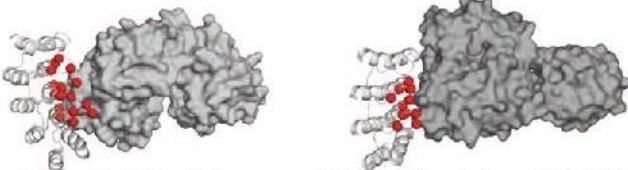

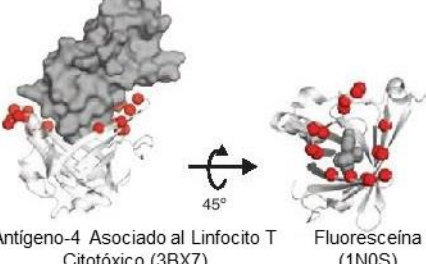

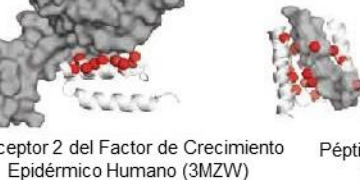

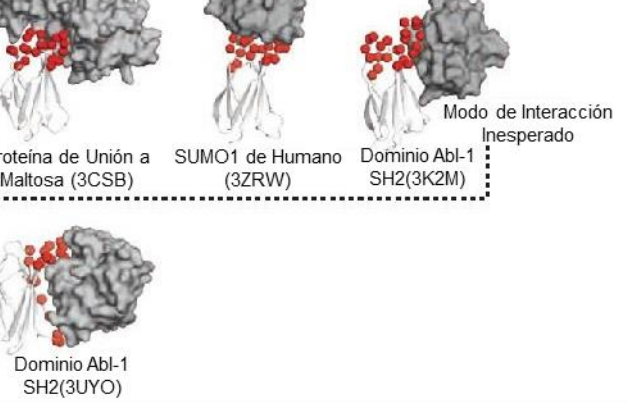
Biblioteca de diseño	Modos de interacción observados	
<p>DARPin</p> 	 <p>Proteína de Unión a Maltosa (1SVX)</p> <p>DARPin en Complejo con ORF49 del Fago Lactococcal TP901-1 (3HG0)</p>	
<p>Anticualinas</p> 	 <p>Antígeno-4 Asociado al Linfocito T Citotóxico (3BX7)</p> <p>Fluoresceína (1N0S)</p>	
<p>Affibodies</p> 	 <p>Receptor 2 del Factor de Crecimiento Epidérmico Humano (3MZW)</p> <p>Péptido Aβ (1-40) (2OTK)</p> <p>Modo de Interacción Inesperado</p>	
<p>Monobodies</p>  <p>Biblioteca de diseño alternativa</p>	 <p>Proteína de Unión a Maltosa (3CSB)</p> <p>SUMO1 de Humano (3ZRW)</p> <p>Dominio Abl-1 SH2(3K2M)</p> <p>Modo de Interacción Inesperado</p> <p>Dominio Abl-1 SH2(3UYO)</p>	

Fig. 2. Proteínas utilizadas como andamios para el diseño racional y utilizando técnicas de evolución dirigida para la obtención de proteínas con funciones de reconocimiento. En rojo, a la izquierda se ilustran las zonas de los dominios modificadas mediante evolución dirigida, a la derecha se muestran las estructuras cristalográficas obtenidas, los códigos de acceso al PDB se indican debajo de cada complejo, la proteína objetivo se indica en color gris y con representación de superficie. Imagen tomada y modificada de Gilbreth y Koide, 2012 [35].

Las estructuras cristalográficas de diversos ICPs indican que miembros de esta familia de proteínas presentan ventajas estructurales para ser utilizados como APs. Estas proteínas son termoestables, tienen un plegamiento del tipo Ig, tienen 3 asas expuestas al solvente responsables de la función de reconocimiento de proteasas [28, 29, 36-38]. Chagasina (*chagasin*) es el miembro modelo de la familia de ICP, esta proteína ha sido cristalizada sola o en complejo con algunas cisteín proteasas (CPs) [28, 37, 38]. El genoma del protozooario *Entamoeba histolytica* codifica para dos ICPs (EhICP1 y EhICP2) y la estructura cristalográfica de uno de ellos ha sido resuelta [20, 36]. Los ICPs incluso se asemejan estructuralmente al correceptor CD8 α de las células T de humano que reconocen una gran variedad de péptidos [39, 40]. Basados en estas propiedades estructurales, postulamos que los ICPs pueden ser usados como andamios para el diseño racional de proteínas [26, 29].

II. Hipótesis

El EhICP1 tiene un dominio estable tipo Ig, el cual hace posible la introducción de secuencias peptídicas en las asas sin alterar el plegamiento, por lo que puede ser utilizado como andamio proteico para el diseño racional de proteínas con funciones de reconocimiento.

III. Objetivos

3.1 Objetivo general

Generar mediante diseño racional una proteína con función de reconocimiento hacia MDM2, utilizando como andamio un ICP de la familia MEROPS I42.

3.2 Objetivos específicos

- I. Seleccionar un ICP para ser utilizado como andamio proteico.
- II. Identificar las regiones a modificar en el ICP con base en un analisis estructural.
- III. Expresar y purificar las versiones generadas de ICP y la proteína blanco.
- IV. Evaluar la formacion de complejos proteicos entre las contrucciones generadas y la proteína blanco.
- V. Determinar la fuerza de unión entre los complejos formados.
- VI. Evaluar la estabilidad térmica de las contrucciones generadas.
- VII. Determinar las condiciones óptimas para la cristalización de los complejos.
- VIII. Resolver la estructura de los cristales de complejo obtenidos.
- IX. Analizar las estructuras cristalográficas.

IV. Materiales y Métodos

4.1 Obtención del modelo de homología, diseño racional y predicción del plegamiento *in silico*.

El modelo de homología para el EhCIP1 se obtuvo utilizando el software MOE, siguiendo el protocolo predeterminado para el modelado de homología, el modelo final se generó usando el campo de fuerza CHARMM. Para mimetizar la interfaz entre p53 y MDM2, seleccionamos un péptido sintético análogo a la zona de transactivación de p53 que interactúa con MDM2 con alta afinidad [41]. La secuencia (N-TSFAEYWNLLSP) fue independientemente “injertada” en cada una de las asas de EhCIP1 para generar 3 posibles quimeras con afinidad por MDM2. Utilizando el software PyMOL de visualización molecular, se realizó un análisis estructural para la selección de los residuos de las asas que fueron modificados, el criterio fue: la conservación de los residuos del núcleo hidrofóbico, importantes para el mecanismo de plegamiento del dominio Ig [22], así como la cercanía de los residuos que rodean el núcleo y que potencialmente podrían formar otras interacciones estabilizantes, como puentes de hidrógeno, puentes salinos y otras interacciones, en un radio de 6 Å, donde las fuerzas de van der Waals son relativamente débiles. La predicción del plegamiento se realizó en la plataforma I-Tasser, siguiendo el protocolo predeterminado.

4.2 Expresión y purificación de las proteínas recombinantes

Las quimeras de EhCIP1-p53 y el dominio de MDM2 clonados en los sitios NdeI y BamHI del vector pET19b (cuya secuencia polipeptídica está indicada en el **Apéndice A**) fueron transformadas en la cepa *E. coli* BL21 (DE3) de forma separada. Partiendo de un preinóculo para cada construcción, 1 L de medio de cultivo LB adicionado con 100 mg mL⁻¹ de ampicilina fue inoculado y crecido a 37 °C con agitación a 200 rpm. Una vez alcanzada una O.D.₆₀₀ = 0.7, los cultivos fueron inducidos utilizando IPTG a una concentración final de 0.5 mM e incubados a 16 °C durante la noche. Las proteínas fueron purificadas mediante cromatografía de afinidad a metales (IMAC), utilizando

columnas pre-empacadas de Niquel-Sefarosa. Las células fueron resuspendidas en 30 mL de la solución de lisis (25 mM Tris pH 8.0, 150 mM NaCl y 0.5 mg mL⁻¹ de lisozima) e incubadas en hielo durante 1 hora. Se procedió a la sonicación y la centrifugación a 21,000 g por 45 minutos a 4 °C. El sobrenadante fue filtrado utilizando un poro de 0.45 µm y pasado por la columna de níquel. Se realizó un lavado de 30 volúmenes de columna con solución de lavado 1 (25 mM Tris pH 8.0, 150 mM NaCl y 50 mM imidazol) y un lavado con 30 volúmenes con solución de lavado 2 (25 mM Tris pH 8.0, 500 mM NaCl y 50 mM imidazol). Las proteínas se eluyeron con 5 volúmenes de solución de elución (25 mM Tris pH 8.0, 150 mM NaCl y 500 mM de imidazol). La etiqueta de histidinas de las proteínas purificadas fue removida, según se requirió, en la solución de diálisis (25 mM Tris pH 8.0, 150 mM NaCl, 1mM EDTA y 1mM DTT) utilizando la enzima PreScission Protease (GE Healthcare Life Sciences). Se realizó un segundo paso de purificación mediante filtración en gel (exclusión molecular) utilizando la columna Superdex 75 10/300 GL (GE Healthcare Life Sciences), en una solución con 25 mM Tris pH 8.0 y 150 mM NaCl.

4.3 Determinación y separación de los complejos formados

Las quimeras de EhICP1-p53 purificadas, fueron incubadas (de forma individual) con la proteína MDM2 (esta última en exceso) durante 1 hora a 4 °C y con agitación suave. La determinación de los complejos fue mediante filtración en gel en una columna Superdex 75 10/300 GL (GE Healthcare Life Sciences), 25 mM Tris pH 8.0 y 150 mM NaCl. Las fracciones con los complejos entre las quimeras de EhICP1-p53 y MDM2 fueron analizadas por SDS-PAGE. El peso molecular de los complejos fue calculado utilizando la curva de calibración para la columna mencionada.

4.4 Cristalización y colección de datos de difracción.

Los ensayos de cristalización fueron realizados utilizando la técnica de la gota colgante – difusión de vapor, mezclando diferentes relaciones del complejo purificado y

la solución de cristalización (proteína/precipitante, con una concentración de proteína de 5 mg mL⁻¹. Se probaron todas las condiciones de los siguientes kits comerciales: Crystal Screen 1 y 2 (Hampton Research) y Wizard I, II, III y IV (Rigaku Reagents). Se lograron obtener cristales adecuados para la difracción de rayos X en dos condiciones, condición 1 consistió de 0.1 M NaCl, 0.1 M Hepes pH 7.5, 1.6 M (NH₄)₂SO₄, y la condición 2 de 1.5 M (NH₄)₂SO₄, 0.1 M Tris pH 8.5, 12 % v/v glicerol. Los cristales fueron crioprotectados en una solución adicionada con 20 % v/v de glicerol e inmediatamente sumergidos en nitrógeno líquido para su almacenamiento hasta su transportación y difracción. Un único cristal de ambas condiciones fue utilizado para coleccionar datos en sincrotrón (APS, Chicago) a 100 K, con un detector CCD MarMosaic 225, en la *beam line* LS-CAT 21-ID-F, con 0.5 ° de oscilación por imagen y una exposición de 1 segundo.

4.5 Determinación de las constantes de disociación de los complejos.

La determinación de las constantes de disociación de los complejos se realizó utilizando dos técnicas diferentes, la termoforesis a microescala (MST) y la interferometría de biocapa (BLI). Para el análisis mediante MST, las proteínas se purificaron por IMAC, se les removió la etiqueta de histidinas y fueron de nuevo purificadas mediante filtración en gel. Los ensayos MST se realizaron a temperatura ambiente (Monolith NT.115 Nano Temper) con una concentración de proteína fija de las quimeras de EhICP1-p53 (1.5×10^{-7} M) y con una curva de MDM2 marcada (Monolith Protein Labeling Kit RED-NHS-amino reactiva) en un rango de 2.67×10^{-9} a 8.75×10^{-5} M en una solución amortiguadora de fosfatos (PBS) pH 7.4.

Para la BLI, los experimentos fueron realizados en un equipo Octet Red96 System (ForteBio) a 30 °C en PBS pH 7.4, Triton X-100 0.05% v/v y 5mM TCEP. Como se realizó en los experimentos de termoforesis, la proteína MDM2 fue marcada (biotinilada, siguiendo el protocolo del fabricante, Sigma, B2643, y dializada para remover el exceso de biotina). La proteína biotinilada MDM2 fue inmovilizada en biosensores de estreptavidina (SA Biosensors, ForteBio). Las muestras fueron agitadas a 1000 rpm, los

biosensores fueron saturados usando una concentración de 4.7×10^{-7} M de MDM2 biotinilada durante 10 minutos. Las quimeras de EhICP1-p53 fueron usadas en un rango de 4.16 a 24.9×10^{-6} M. La asociación se monitoreo por 500 segundos y la disociación por 400 segundos en solución PBS. El cambio en la distancia espectral de la luz ($\Delta\lambda$), el cual está directamente correlacionado con el cambio en el espesor (nm) en la capa biológica en la punta del biosensor. La asociación y disociación fueron corregidas de acuerdo a la línea base.

4.6 Determinación de la estabilidad térmica de las quimeras.

La estabilidad térmica fue determinada mediante fluorometría diferencial de barrido (DSF), en un equipo Prometheus NT.48 (NanoTemper Technologies), el gradiente de temperatura fue de 20–95 °C con un incremento de 1.5 ° por minuto. El análisis de los datos se realizó utilizando el software NT Melting Control. Las T_{ms} (temperatura a la que la fracción de proteína desnaturalizada es igual a 0.5) fueron determinadas experimentalmente utilizando una función polinomial, las T_{ms} se obtuvieron mediante la determinación de la primera derivada de los datos.

4.7 Solución estructural y refinamiento del complejo.

Las reflexiones obtenidas en la difracción de rayos X fueron indexadas e integradas utilizando el software iMosflm y Scala [42], respectivamente. Se encontró una solución por reemplazo molecular (MR) mediante Phaser-MR [43], utilizando para la búsqueda un modelo truncado a alaninas del EhICP2 (PDB ID: 3M86 [36]) y la estructura de la MDM2 de humano (PDB ID: 3EQS [41]), y una búsqueda de múltiples componentes como estrategia. El refinamiento se realizó en phenix.refine [44]. El modelo fue ajustado manualmente para mejorar el ajuste al mapa electrónico en Coot [45]. La calidad del modelo se evaluó utilizando las herramientas de validación de Coot, Molproivity y el servidor de validación del wwPDB [46].

4.8 Formación de fibras en EhICP1 silvestre y análisis por microscopía electrónica de transmisión

La proteína silvestre EhICP1 fue purificada y se le removió la etiqueta de histidinas y se llevó a una concentración de 1 mg mL^{-1} . Fue incubada a $42 \text{ }^{\circ}\text{C}$ durante 24 horas en una solución amortiguadora con 25 mM Tris pH 7.6, 150 mM NaCl. La confirmación de oligómeros se analizó por SDS-PAGE.

Para el análisis por microscopía electrónica de transmisión (TEM), la muestra se centrifugó a $21,000 \text{ g}$ por 10 minutos y el sobrenadante se descartó por decantación, el remanente se lavó 3 veces con agua mili Q y finalmente resuspendida en agua, $7 \text{ }\mu\text{L}$ de muestra fueron colocados en una grilla (Silicon Monoxide on Formvar, Ted Pella, Inc.) e incubados durante 20 minutos para permitir la sedimentación de las fibras. La gota fue secada con cuidado. Para el contraste se utilizaron $10 \text{ }\mu\text{L}$ de acetato de uranilo (Electron Microscopy Sciences) al 2.5% en agua y se incubó la grilla durante 10 minutos más. Se utilizó un equipo Philips / FEI (Morgagni 268), operado a 80 kV con ampliaciones entre 89 y 180×10^3 aumentos.

V. Resultados

5.1 Los ICPs tienen un asa conservada y dos asas variables expuestas al solvente

El plegamiento de todos los ICPs de esta familia MEROPS I42 es del tipo Ig, es decir, su plegamiento está constituido por dos hojas β formando una especie de sándwich. La comparación estructural y el alineamiento de las secuencias entre los ICPs del género *Entamoeba* y la proteína modelo chagasina, muestran que los inhibidores de esta familia presentan 3 asas expuestas al solvente, un asa como sitio reactivo a proteasas, asa BC con una conformación de hélice 3_{10} , y dos asas más largas y flexibles que son importantes en el reconocimiento y unión hacia las proteasas blanco, además tienen una similitud estructural con otras proteínas de reconocimiento de la superfamilia de las Ig [36] [29-34, 37]. **La Fig. 3** muestra las similitudes estructurales entre EhICP1, EhICP2 y el ICP modelo chagasina, así como la variabilidad en las asas denominadas BC, DE y FG de estos ICPs.

reductoras. La **Fig. 4** muestra la formación de agregados de alto peso molecular, asistidos por puentes disulfuro en el EhICP1 en condiciones no reductoras. La figura muestra que el EhICP1 silvestre puede formar multímeros en condiciones no reductoras aún cuando no es incubado con un agente caotrópico, **Fig.4**, carriles 6 y 10.

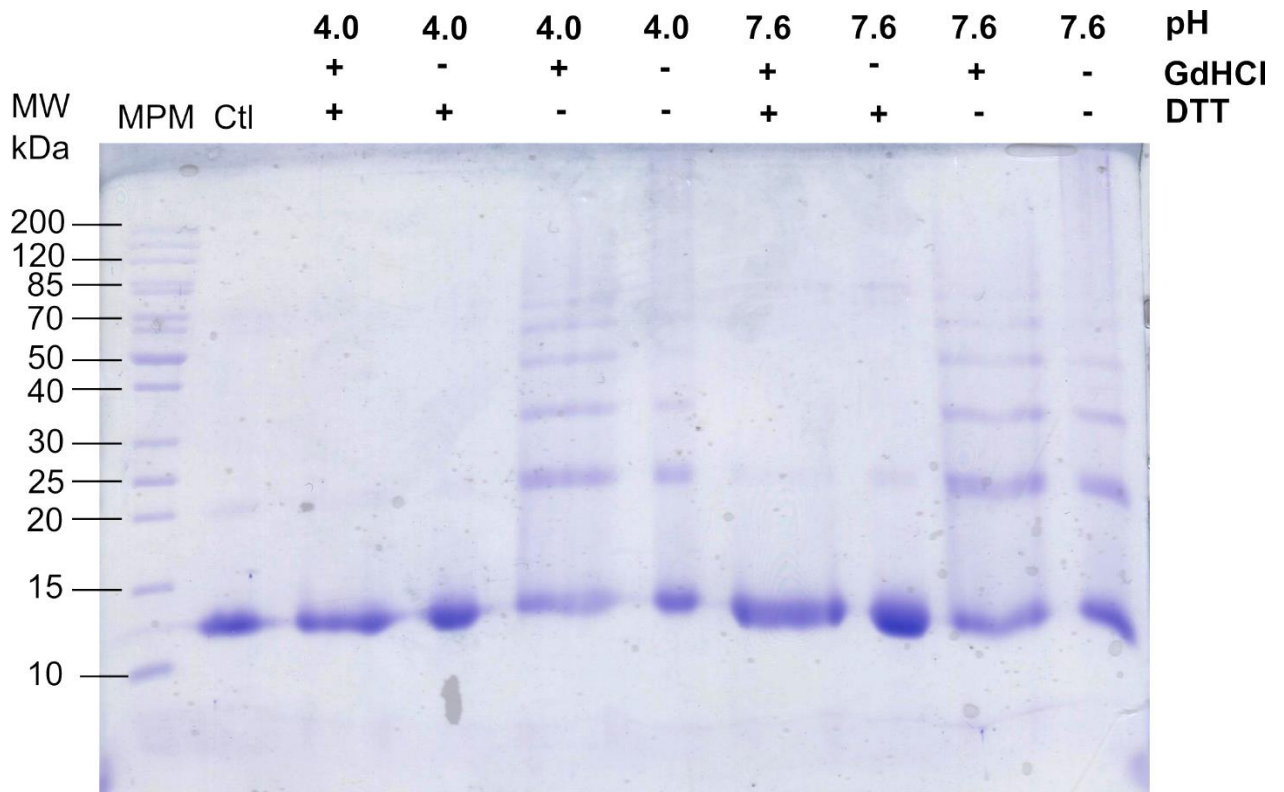


Fig. 4. Formación de oligómeros y agregados de alto peso molecular en EhICP1 en condiciones no reductoras. EhICP1 purificado e incubado por 16 horas a 42 °C en presencia o ausencia del agente reductor DTT y/o cloruro de guanidinio como agente caotrópico. La formación de oligómeros y agregados de alto peso molecular se promueve en condiciones no reductoras (carriles 5, 6, 9 y 10) en pH ácido y cercano al fisiológico.

Debido a la posible interferencia en los procesos de purificación y en los ensayos de interacción, decidimos trabajar con una versión de EhICP1 libre de cisteínas, la triple mutante por serinas (C36S, C58S, C60S). En todo este trabajo nos referiremos a esta mutante como EhICP1, ya que es el fondo en el que realizamos las modificaciones de las asas.

La modificación de la secuencia en cualquier proteína, ya sea la mutación de un simple residuo o la modificación de regiones más grandes (como asas), da generalmente como resultado una desestabilización, que se ve reflejada en la estabilidad térmica de la proteína [47], debido a lo anterior, una alta termoestabilidad favorece las probabilidades de que una proteína modificada adopte un plegamiento global similar al nativo.

Para evaluar la termoestabilidad de los ICPs de *Entamoeba*, se llevaron a cabo experimentos de dicroísmo circular (CD) a 218 y 232 nm, y se determinaron las temperaturas de desplegamiento térmico. Las T_{ms} se muestran en la **Tabla 1** y van de un rango de 69.9 a 72.9 °C y de 81.7 a 83.3 °C cuando son evaluadas a 218 y 232 nm, respectivamente. A pesar de que las temperaturas son similares, decidimos utilizar a la triple mutante de cisteínas (EhICP1-cys en **Tabla 1**) debido a que presenta T_{ms} ligeramente más elevadas que el EhICP2, y además presenta reversibilidad por la estructura nativa en múltiples eventos de desplegamiento/plegamiento (datos no mostrados). La T_m del EhICP1 es relativamente alta y es bastante similar a la del dominio tipo III de fibronectina (FN3, T_m de 84 °C) que también tiene un plegamiento del tipo Ig y ha sido utilizado como andamio proteico [34] [48].

Tabla 1. Estabilidad térmica de ICPs por dicroísmo circular.

Proteína	Longitud de onda, nm	T_m (°C)
EhICP2	218	71.2 ± 0.18
EhICP1-cys	218	72.9 ± 0.21
Chagasina	218	69.9 ± 0.16
EhICP2	232	81.74 ± 0.17
EhICP1-cys	232	83.34 ± 0.12
Chagasina	232	82.34 ± 0.29

*Las curvas de desplegamiento fueron obtenidas monitoreando el cambio en la elipticidad a 218 y 232 nm. Las T_m representan el punto medio en la curva de desnaturalización.

5.3 Las asas expuestas al solvente de EhICP1 tienen regiones que pueden ser modificadas sin comprometer el plegado global

Como prueba de concepto y para evaluar la capacidad de los ICPs de ser usados como andamios proteicos para el diseño racional, seleccionamos la interfaz entre la proteína supresora de tumores p53 y su principal antagonista, la proteína MDM2 ubiquitina ligasa E3. La selección de la interfaz mencionada tiene como bases la relevancia de la interacción, MDM2 ha sido sugerido como un biomarcador de cáncer [49], y además existe suficiente información estructural detallada de la interacción de estas proteínas, específicamente de la interfaz entre ambas [41] [50]. Para mimetizar la interfaz entre p53 y MDM2, seleccionamos un péptido sintético denominado PMI, que corresponde con la región N-terminal de transactivación de p53, dicho péptido se une a MDM2 con alta afinidad e incluso su estructura cristalográfica ha sido resuelta [41]. Con el fin de generar proteínas basadas en EhICP1 con modificaciones en sus tres asas, y que puedan unirse a MDM2, buscamos producir 3 versiones o quimeras del EhICP1-p53, cada una conteniendo el péptido (PMI) de unión a MDM2 en una de sus asas. Denominamos a esas proteínas EhICP1-p53(BC), EhICP1-p53(DE) y EhICP1-p53(FG).

Para seleccionar las posibles regiones de EhICP1 que pueden ser susceptibles de mutaciones sin comprometer el plegado global y ser sustituidas por el péptido de p53, realizamos un modelo de homología del EhICP1. A partir de un análisis estructural para delimitar los residuos que en teoría pueden ser modificados sin alterar al dominio Ig, seleccionamos una serie de residuos en las asas del ICP para ser reemplazados por el péptido PMI. Para cada quimera, seleccionamos una serie de residuos de un asa, es decir, para la generación de la quimera EhICP1-p53(BC) se modificaron los residuos correspondientes al asa BC indicados en color verde en la **Fig. 5**, lo anterior aplica para las quimeras en las asas DE y FG. Los residuos seleccionados para los cambios, interesantemente son los que presentan los factores B (factores de desplazamiento térmicos) más elevados y generalmente se asocian con una alta flexibilidad, como se aprecia en las dos diferentes estructuras del EhICP2 [36].

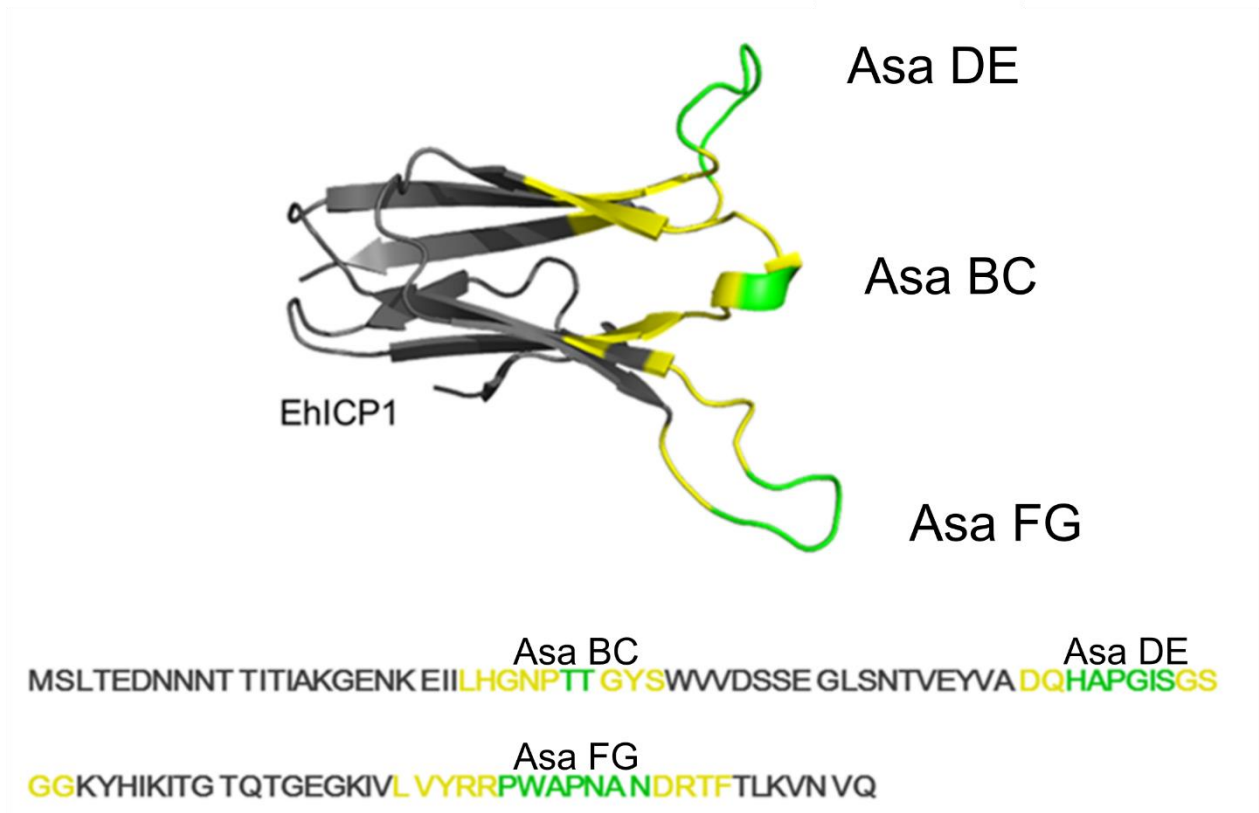


Fig. 5. Análisis estructural del EhICP1 como andamio proteico. La parte superior de la imagen muestra una representación de listones del EhICP1, la zona coloreada de amarillo representa el núcleo hidrofóbico conservado en los ICPs, en verde se indican las zonas sugeridas de las asas para ser modificadas con la inserción del péptido de p53. Dichos residuos (verdes), aparentemente no forman interacciones con el resto del dominio Ig. Los aminoácidos para la modificación por el péptido N-TSFAEYWNLSP se indican en la secuencia en la parte inferior de la imagen y corresponden a los residuos 29-30 del asa BC, 53-59 del asa DE y 85-91 del asa FG.

Para mimetizar la interfaz entre MDM2 y p53, el péptido N-TSFAEYWNLSP se insertó en las regiones marcadas en verde en la **Fig.5** para cada una de las 3 quimeras. La predicción del plegamiento de las construcciones *in silico*, utilizando el servidor I-Tasser [51] muestra una alta probabilidad (valor de confianza cercano a 0.5 C-score, **Tabla 2**) de que las construcciones adopten el plegamiento del tipo Ig. Los modelos de las quimeras se muestran en la **Fig. 6**, las quimeras de las asas DE y FG son las que muestran el péptido de p53 más expuesto al solvente.

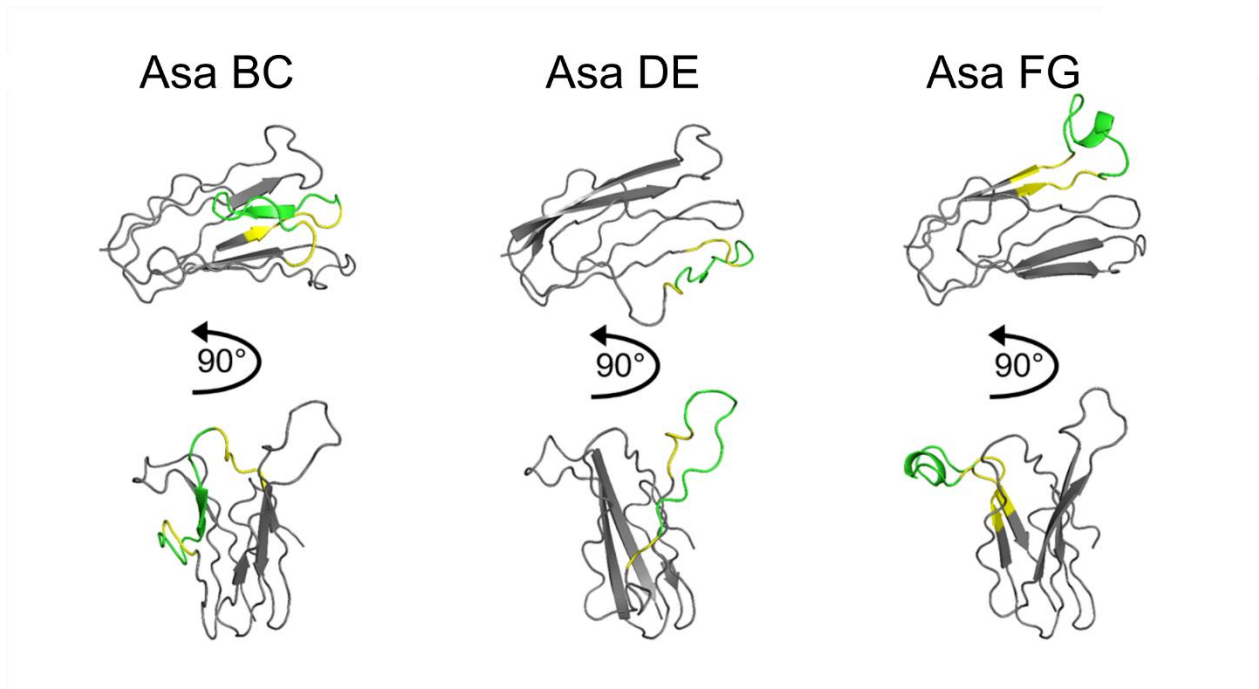


Fig. 6. Modelos de homología de las quimeras de EhICP1-p53. Diferentes vistas de los modelos de homología para las tres quimeras de EhICP1-p53 en las asas BC, DE y FG. La secuencia correspondiente al péptido de p53 está representada en color verde y los residuos que la restringen cercanos al núcleo hidrofóbico se muestran en amarillo. La predicción muestra que las quimeras adoptan un plegamiento global del tipo Ig, las quimeras en las asas DE y FG muestran el péptido insertado preferentemente expuesto al solvente.

5.3.1 EhICP1 tolera modificaciones en las asas DE y FG pero no en BC.

Las proteínas recombinantes modificadas con el péptido de p53 en las asas DE y FG se expresaron de forma soluble en bacteria, sin embargo la quimera BC no se logró obtener de forma soluble, dando como resultado la formación de cuerpos de inclusión (datos no mostrados). La inserción del péptido de p53 en el asa BC altera el plegamiento y promueve la agregación. Para evaluar el efecto en la termoestabilidad de las proteínas quimeras EhICP1-p53(DE) y EhICP1-p53(FG) realizamos ensayos de desplegamiento térmico por dicroísmo circular (CD) y fluorimetría diferencial de barrido (DSF). En CD, la estabilidad térmica de las quimeras fue medida a 218 y 232 nm. La T_m a 218 nm fue 49 y 59 °C para las quimeras DE y FG, respectivamente, mientras que a 232 nm fueron de 51 y 60 °C, **Fig. 7**. Por lo que la inserción del péptido de p53 tiene un

efecto detrimental en la T_m de EhICP1 de 12 y 14 °C para la quimera FG y DE, respectivamente. Los cambios fueron similares al comparar con la técnica DSF. Las curvas de desplegamiento térmico del DSF muestran 2 picos, el primero puede corresponder al triptófano en el péptido p53, mientras el segundo puede deberse a la contribución del triptófano en el núcleo hidrofóbico, que indica el desplegamiento total de la proteína. Los valores obtenidos fueron T_{m1} de 56.6 °C y T_{m2} de 83 °C para el EhICP1, para la quimera DE los valores fueron de T_{m1} de 43 °C y T_{m2} de 82.7 °C, mientras que para la quimera FG fueron de T_{m1} de 45 °C y T_{m2} de 82.4 °C. La comparación se hace en la **Fig. 7** y en la **Tabla 2**.

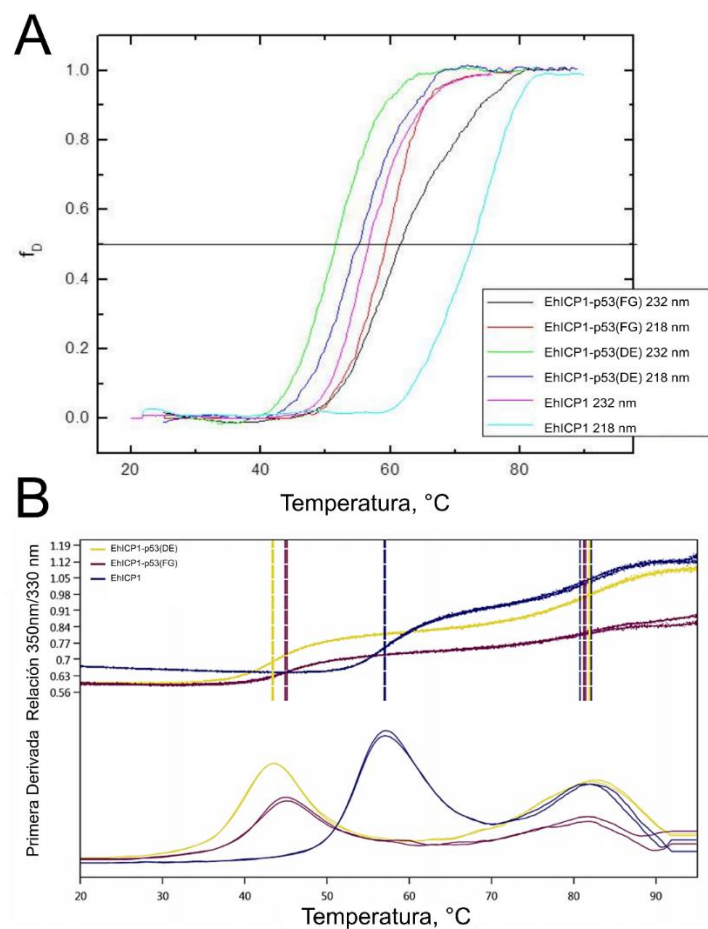


Fig. 7. Comparación de la estabilidad térmica de las quimeras en las asas DE y FG con EhICP1. A Fracción de desplegamiento normalizada para EhICP1, EhICP1-p53(DE) y EhICP1-p53(FG) en la curva de temperatura. Las lecturas de CD fueron realizadas a 218 y 232 nm. **B** Desplegamiento térmico monitoreado por DSF en gradiente de temperatura, la curva de la parte inferior representa la primera derivada de los datos.

La información obtenida indica que la introducción del péptido p53 en las asas DE y FG del EhICP1 altera la estabilidad de la proteína en 12 y 11 °C, respectivamente. El análisis de la energía libre de Gibbs, para el desplegamiento de EhICP1-p53(DE) y EhICP1-p53(FG) usando los datos obtenidos por DSF [52], indica que la introducción del péptido PMI derivado de p53 reduce la estabilidad en 23.85 kJ mol⁻¹ y 19.62 kJ mol⁻¹, respectivamente, **Tabla 2**.

Tabla 2. Correlación entre la probabilidad del modelo estructural con la estabilidad térmica y el cambio en la energía libre de Gibbs para EhICP1 y las quimeras BC, DE y FG.

Proteína	EhICP1	EhICP1-p53(BC)	EhICP1-p53(DE)	EhICP1-p53(FG)
C-score	0.54	0.40	0.65	0.67
TM-score	0.79 ± 0.09	0.77 ± 0.10	0.80 ± 0.09	0.80 ± 0.09
Tm, °C	56.50 ± 0.00	No data	43.62 ± 0.07	45.10 ± 0.00
$\Delta G_{N \rightarrow U}$, kJ mol ⁻¹	42.47 ± 0.65	No data	18.62 ± 0.34	22.85 ± 0.33

*C-score es el valor de confianza para la predicción de homología que realiza el servidor I-Tasser, va de -5 a 2, los valores cercanos a 2 tienen mayor probabilidad de adoptar la conformación predicha. TM-score es un valor de similitud estructural entre la predicción y un modelo real, el valor es de 0 a 1, donde 1 indica total similitud con el modelo comparado. $\Delta G_{N \rightarrow U}$ es el cambio en la energía libre de Gibbs de la proteína plegada (N) a desplegada (U), calculada de acuerdo a la metodología descrita por Wright y colaboradores [52].

5.4 Las quimeras EhICP1-p53(DE) y EhICP1-p53(FG) forman complejo con MDM2

Para evaluar la formación de complejos entre las quimeras de EhICP1-p53 y MDM2, las proteínas purificadas y el EhICP1 como control fueron incubadas con un exceso de MDM2 de forma separada. Los complejos moleculares fueron monitoreados por filtración en gel. Como era de esperarse, la proteína control EhICP1 no forma complejos moleculares con MDM2, ya que se obtienen dos picos correspondientes a las proteínas EhICP1 y MDM2, con pesos moleculares aproximados de 15 y 35 kDa, respectivamente, **Fig. 8A** (curva azul) y **Fig. 8B** (carril 7 y 8). Cabe mencionar que

MDM2 es un dímero en solución, y por eso el peso aproximado de 35 kDa. En cambio, las proteínas quimeras DE y FG muestran la formación de un complejo con MDM2. Para EhICP1-p53(DE) se observa la formación de un complejo de un peso aproximado de 36 kDa, lo que coincide con el peso esperado de un complejo en relación 1:1 entre EhICP1-p253(DE) y MDM2, **Fig. 8B** carril 5 y 6. Para la quimera del asa FG, se tiene la formación de un pico que corresponde con un peso molecular aproximado de 67 kDa. Lo que indica la formación de un complejo en relación 2:2, es decir, un heterotetrámero, **Fig. 8**, carril 2 al 4. Las fracciones fueron analizadas por SDS-PAGE para confirmar la formación de complejos y se muestran en la **Fig. 8B**.

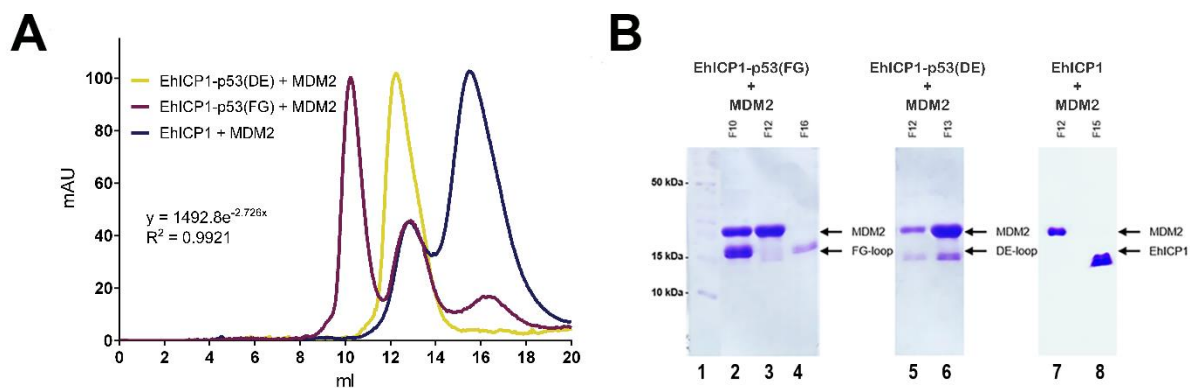


Fig. 8. Interacción de las quimeras EhICP1-p53 con MDM2. A Cromatograma de filtración en gel para las muestras de las quimeras de EhICP1-p53 y EhICP1 incubadas con MDM2 e inyectadas a la columna S75. La línea azul corresponde al control, EhICP1 incubado con MDM2. Los picos que corresponden a 15 y 35 kDa corresponden a EhICP1 y el dímero MDM2, respectivamente. La línea magenta que corresponde a la muestra incubada con EhICP1-p53(FG) y MDM2 muestra un pico de alto peso molecular de aproximadamente 67 kDa, mientras que la línea amarilla correspondiente a la muestra de EhICP1-p53(DE) incubada con MDM2 muestra un pico de aproximadamente 36 kDa. **B** Análisis por SDS-PAGE de las fracciones correspondientes a los picos en el cromatograma. El carril 1 corresponde al marcador de peso molecular, el carril 7 y 8 corresponden a los dos picos obtenidos de la incubación del control EhICP1 y MDM2. El carril 2 corresponde al pico de 67 kDa formado en la incubación de EhICP1-p53(FG) con MDM2, el carril 3 corresponde a MDM2 y el carril 4 corresponde a EhICP1-p53(FG). Los carriles 5 y 6 corresponden al pico de la muestra incubada con EhICP1-p53(DE) con MDM2. La fracción de elución está indicada en la parte superior, las fracciones fueron de 1 mL.

5.5 Cristalización del complejo EhICP1-p53(FG)/MDM2

Se realizaron múltiples intentos de cristalización de los complejos formados entre las quimeras de las asas DE y FG con MDM2. Debido a la baja concentración obtenida de proteína en los complejos purificados por filtración (la recomendación mínima es 5 mg mL⁻¹), se realizó una co-purificación de las proteínas para promover la formación del complejo antes de las etapas de purificación y así maximizar los rendimientos. De esta manera se logró conseguir el complejo purificado a una concentración aproximada de 8.0 mg mL⁻¹. Los ensayos de cristalización se realizaron siguiendo el protocolo descrito en la sección de metodología, las muestras incubadas a 21 °C se secaron después de un par de meses y no aparecieron cristales. Se encontró que la proteína MDM2 forma multímeros asistidos por puentes disulfuro en condiciones poco o no reductoras (datos no mostrados). Para prevenir la posible formación de agregados de alto peso molecular, se procedió a modificar las condiciones de la solución de cristalización, añadiendo el agente reductor TCEP a una concentración final de 10 mM en la proteína. Con lo anterior, se lograron obtener cristales adecuados para su difracción luego de una ligera optimización manual en base a una matriz casera de las condiciones de cristalización. La **Fig. 9** muestra el proceso de mejoramiento de los cristales del complejo EhICP1-p53(FG)/MDM2. No se lograron obtener cristales para el complejo entre la quimera del asa DE y MDM2.

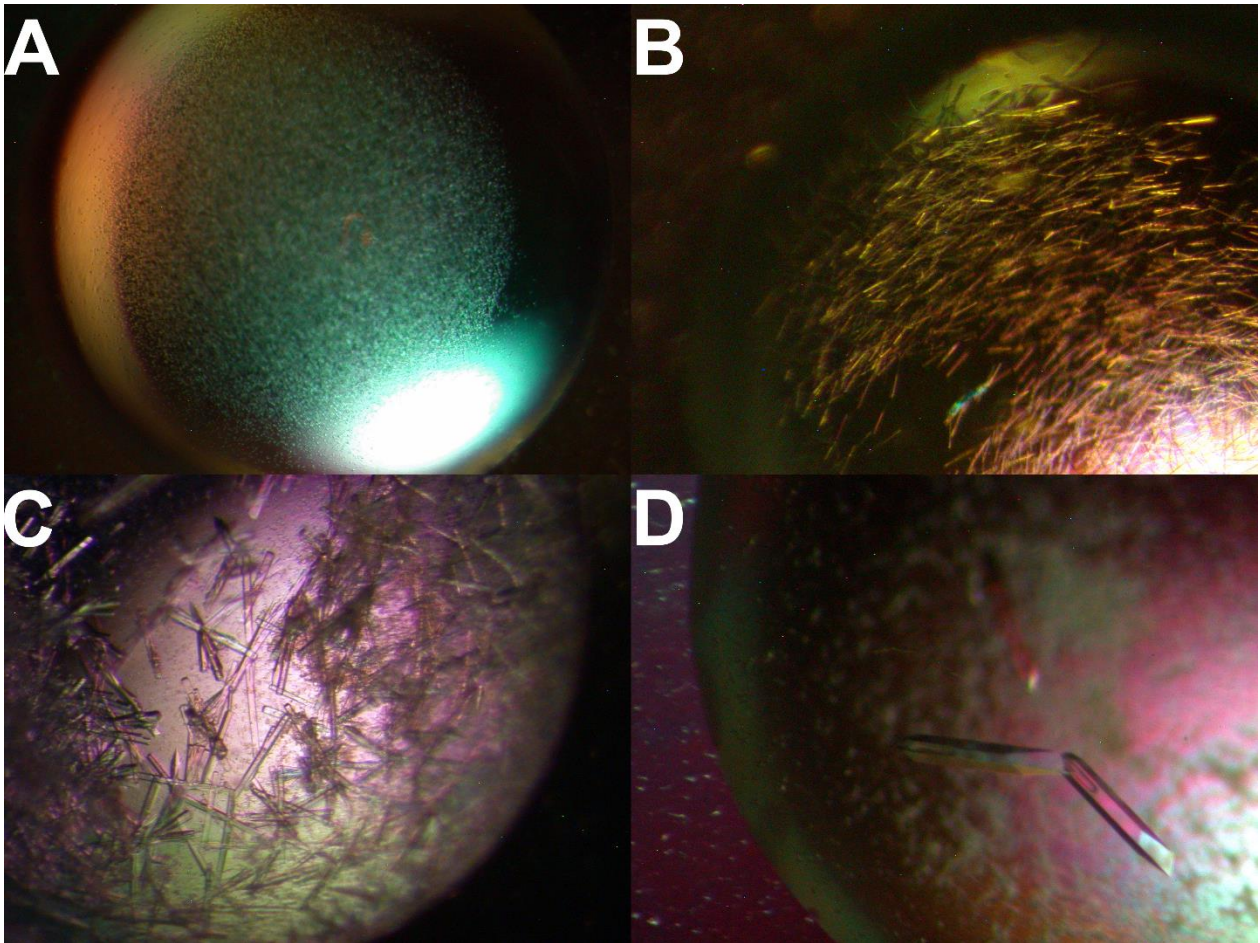


Fig. 9. Mejoramiento de la cristalización del complejo EhICP1-p53(FG)/MDM2. Las fotografías muestran el proceso de mejoramiento de los cristales del complejo formado entre la quimera del asa FG y MDM2, en **A** se muestran la formación de microcristales. Mediante la modificación de las condiciones de cristalización, modificando el porcentaje del agente precipitante en la gota, se pudieron desarrollar cristales en forma de barras delgadas o agujas, mostrados en **B**, y una segunda ronda de optimización de la concentración de proteína en **C**. Los mejores cristales para ser difractados se lograron en una solución de cristalización que contenía 0.1 M NaCl, 0.1 M Hepes pH 7.5, 1.6 M $(\text{NH}_4)_2\text{SO}_4$, en una relación 2:1 proteína:precipitante y son mostrados en **D**.

5.6 Detalle del modelo cristalográfico y densidad electrónica del complejo EhICP1-p53(FG)/MDM2

Los cristales del complejo EhICP1-p53(FG) difractaron a mediana-alta resolución (1.9 angstroms, Å), la **Fig. 10** Muestra los detalles de la colección de datos y un ejemplo del patrón de difracción.

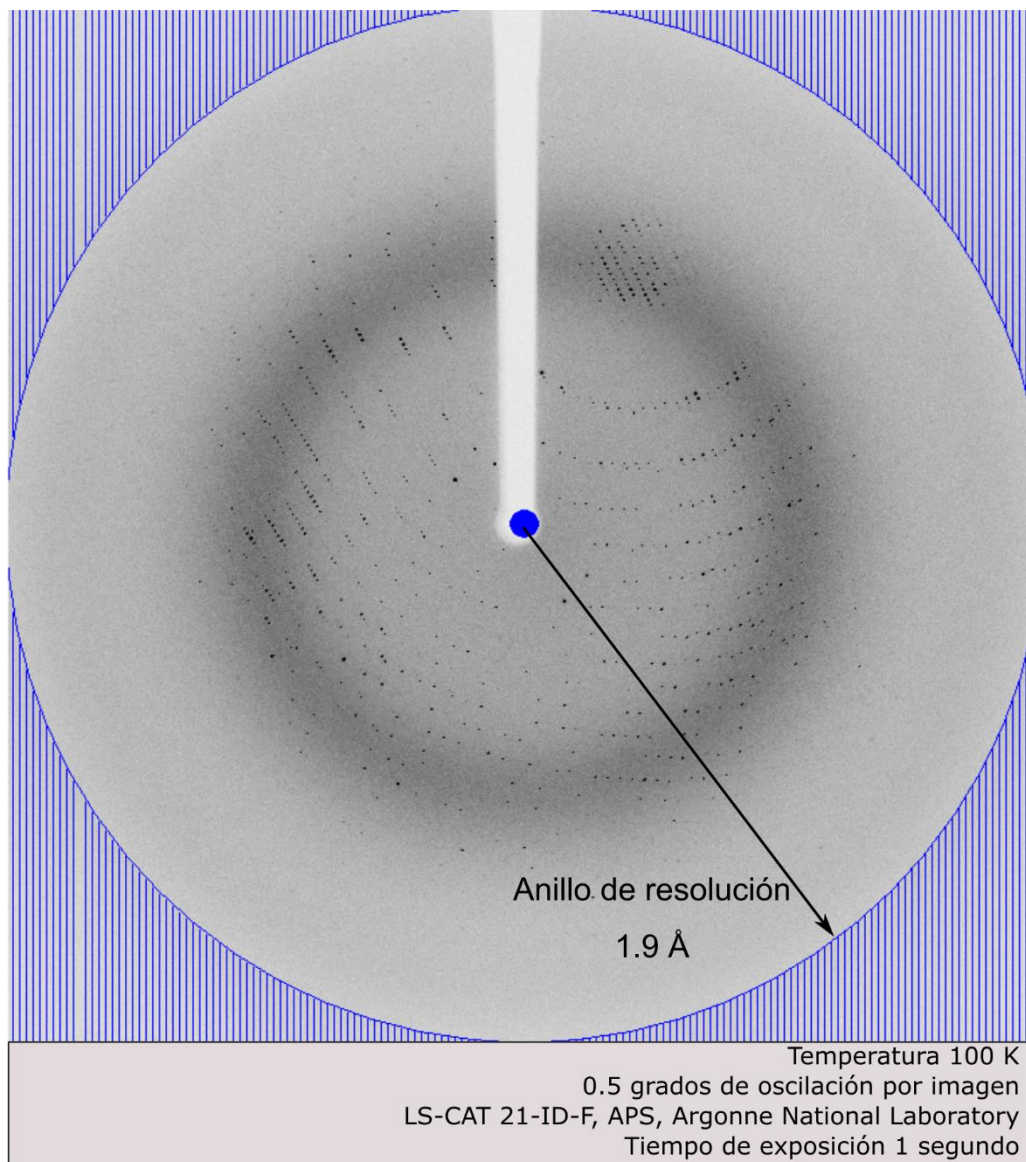


Fig. 10. Patrón de difracción y características de la colecta de datos. Los cristales se difractaron en el sincrotrón, *Advanced Photon Source* de Chicago, a una resolución de 1.9 Å. El patrón de difracción mostrado en esta figura es uno de las 400 imágenes colectadas para el cristal del cual se obtuvo la estructura del complejo EhICP1-p53(FG)/MDM2.

Los datos del cristal que difractó a mayor resolución fueron procesados y se logró obtener una solución por remplazo molecular, el modelo final fue refinado a una resolución de 1.9 Å, el nivel de detalle del mapa de densidad electrónica obtenido a esta resolución se muestra en la **Fig. 11**.

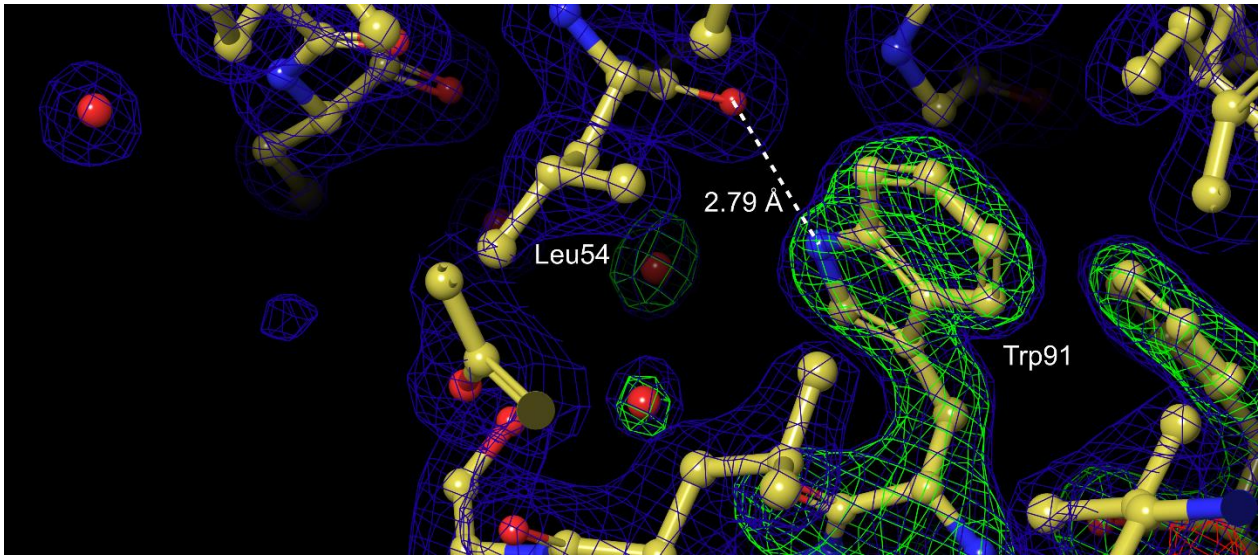


Fig. 11. Detalle del modelo estructural y mapa de densidad electrónica del complejo EhICP1-p53(FG)/MDM2. Se muestran en detalle los principales residuos de la interfaz en color amarillo. El mapa de contorno 2Fo - Fc que se muestra en color verde a un nivel de sigma 5 (como resultado de un *Omitmap*) alrededor de los residuos triptófano y fenilalanina que se encuentran en la interfaz. El triptófano 91 del EhICP1-p53(FG) además de encontrarse en una caviad hidrofóbica, forma un puente de hidrógeno con la leucina 54 de MDM2. En la vista se muestran además tres moléculas de agua (esferas rojas sin conecciones).

5.7 Los complejos formados entre las quimeras de EhICP1-p53 y MDM2 muestran una afinidad cercana a la de la interacción canónica p53-MDM2

Para evaluar la fuerza de la interacción entre las quimeras de EhICP1-p53 y MDM2, utilizamos dos técnicas, en la primera, la termoforesis de microescala (MST) evaluamos las interacciones entre las quimeras de EhICP1-p53 y MDM2. De nuevo, como era de esperarse, no se observó interacción entre el control EhICP1 y MDM2, **Fig. 12**, curva verde. En el caso de las quimeras de EhICP1-p53 en el asa DE y FG se observó un efecto dosis-dependiente, **Fig. 12**, curvas azul y rojas, respectivamente.

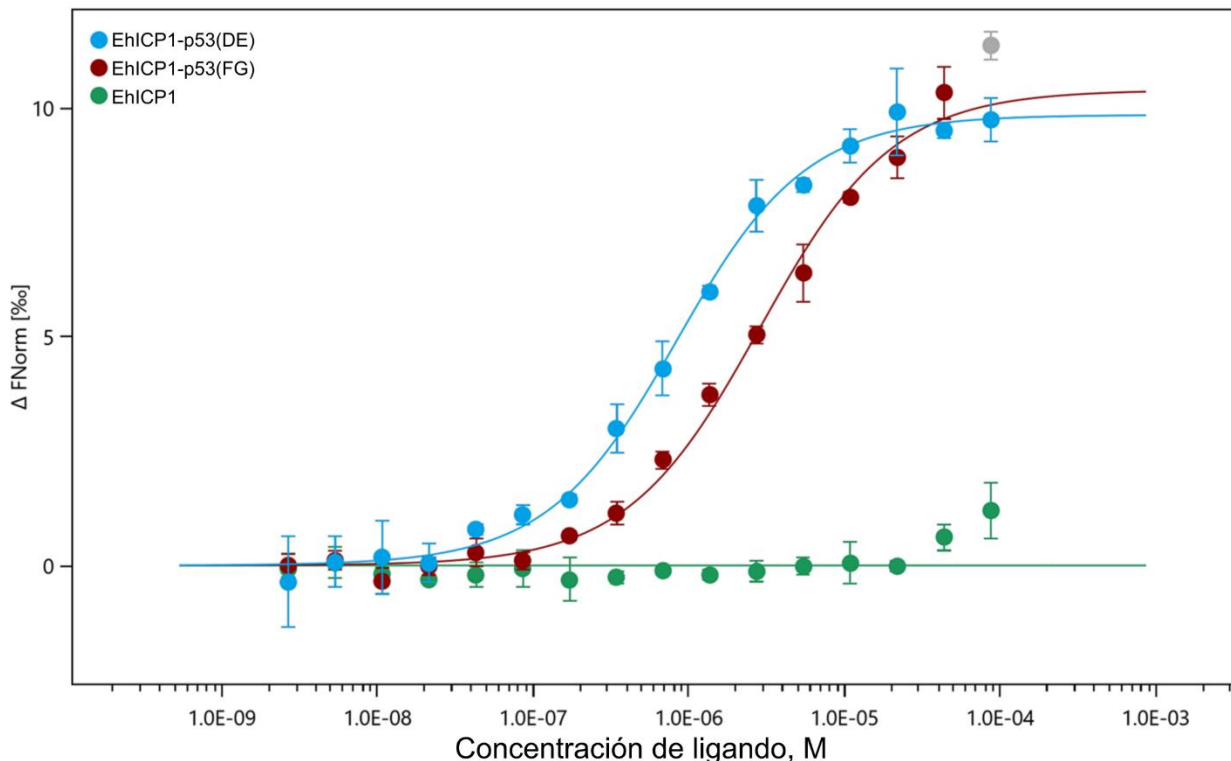


Fig. 12. Afinidad de las quimeras de EhICP1-p53 contra MDM2. Curvas de unión isotermales por MST, la curva de titulación para el control EhICP1 + MDM2 se muestra en verde, no hay efecto dosis-respuesta. Las curvas de unión para las quimeras DE y FG se muestran en color azul y rojo, respectivamente. Los datos indican que EhICP1-p53(DE) tiene una mayor afinidad por MDM2 (K_D más pequeña) que EhICP1-p53(FG).

La K_D entre MDM2 y EhICP1-p53(DE) obtenida por MST fue de 0.76×10^{-6} M, mientras que la obtenida para la interacción EhICP1-p53(FG) fue de 2.79×10^{-6} M, indicando que la inserción del péptido de p53 en el asa DE en lugar del asa FG resulta en una afinidad que es aproximadamente 3.7 veces mayor. Con el fin de corroborar los valores de afinidad obtenidos entre las quimeras de EhICP1-p53 y MDM2, realizamos la medición de las constantes utilizando la técnica de interferometría de biocapa, **Tabla 3**.

Tabla 3. Constantes de disociación para las quimeras de EhICP1-p53 determinadas por interferometría de biocapa.

	$K_D(M)$	K_{Derror}	$k_{on}(1/MS)$	$k_{onError}$	$k_{off}(1/s)$	$k_{offError}$	R^2
FG 2:1 HL fit	2.57×10^{-6}	5.89×10^{-8}	1.83×10^2	2.07	4.7×10^{-4}	9.37×10^{-6}	0.9956
DE 2:1 HL fit	1.44×10^{-6}	2.37×10^{-8}	1.12×10^3	1.63	1.62×10^{-3}	1.28×10^{-5}	0.9961

*La K_D fue calculada mediante la relación k_{off}/k_{on} , donde k_{off} es la constante de disociación y k_{on} es la constante de asociación, determinadas en el instrumento Octet (ForteBio)

5.8 La estructura cristalográfica del complejo EhICP1-p53(FG)/MDM2 muestra un heterotetrámero formado por 2 moléculas de EhICP1-p53(FG) interconectadas, y 2 moléculas de MDM2

La solución por remplazo molecular del complejo EhICP1-p53(FG)/MDM2 revela un heterotetrámero formado por 2 moléculas de la quimera en el asa FG y dos moléculas de MDM2, **Fig. 13**. La información del procesamiento de los datos de difracción, del refinamiento estructural así como el contenido de la unidad asimétrica se muestra en la **Tabla 4** de cristalografía.

Tabla 4. Datos cristalográficos y estadísticas del refinamiento del complejo EhICP1-p53(FG)/MDM2.

PDB ID	5SWK
Colección de datos:	
Longitud de onda (Å)	0.97872
Grupo espacial	P4 ₃ 2 ₁ 2
Dimensiones de la celda unitaria (Å)	a,b = 87.361, c = 160.84
Rango de resolución (Å)	61.77 - 1.92
Número total de observaciones	350,170
Multiplicidad	7.3
Completeness (%) ^a	99.99 (100.00)
R _{sym} (%)	16.0
I/σ (I) ^a	9.78 (1.07)
Refinamiento:	
Rango de resolución (Å)	48.99 - 1.923 (1.992 - 1.923)
Reflexiones totales	48,093
Reflecciones (<i>working/test</i>)	45,619 / 2,474
Wilson factor B	29
R _{work}	0.2213 (0.3177)
R _{free}	0.2487 (0.3518)
Número de átomos no-hidrógenos	3067
macromoléculas	2882
Ligandos	28
Agua	157
Residuos de proteína	370
RMS(enlaces)	0.005
RMS(ángulos)	0.83
Favorecidos Ramachandran (%)	99
Clashscore	0.35
Promedio de factores B	32.30
macromoléculas	32.00
ligandos	51.50
solvente	34.70

^a Los valores para la capa de mayor resolución se muestran en paréntesis. Valor de Rfree test del 5.14%

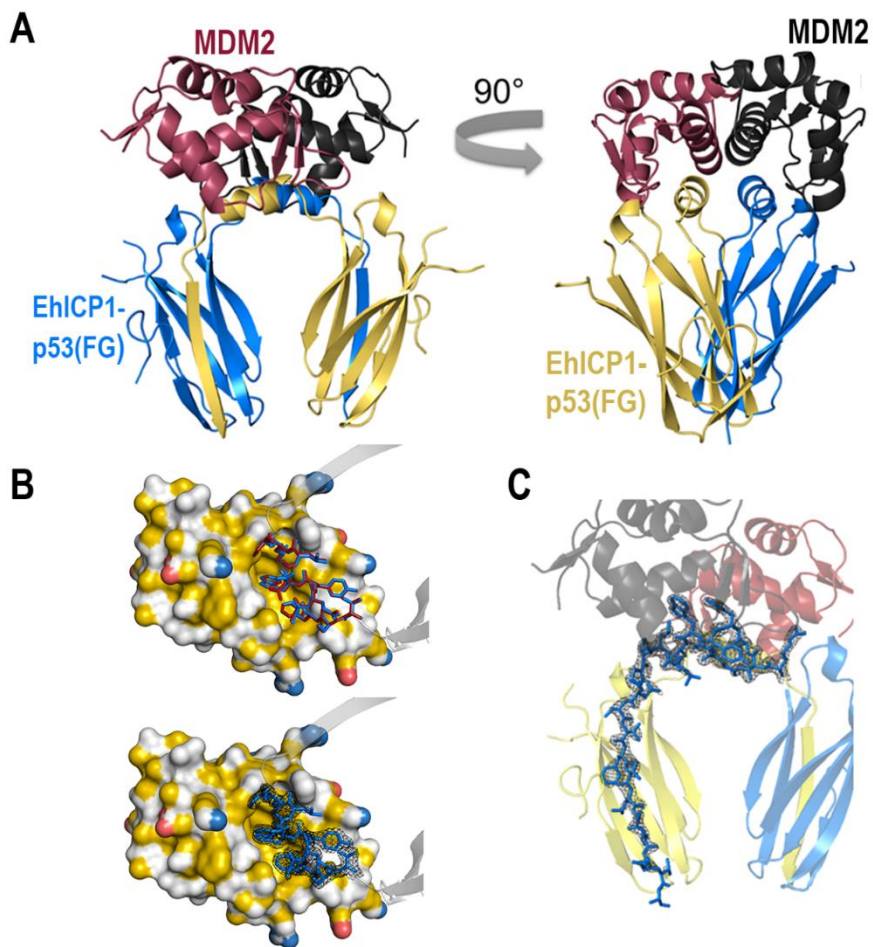


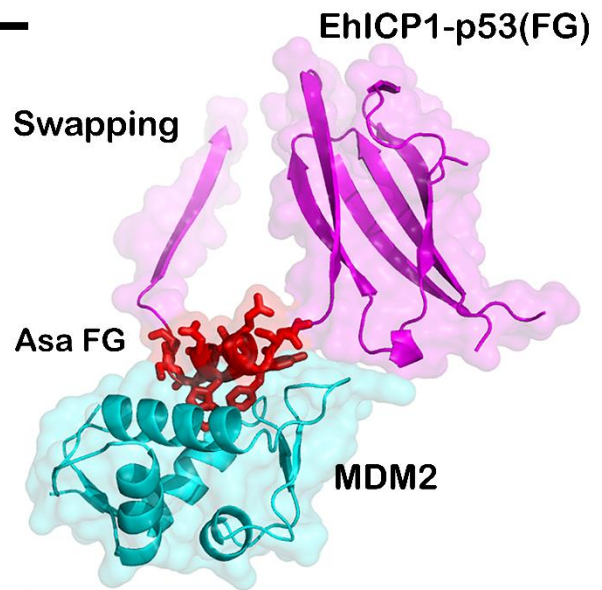
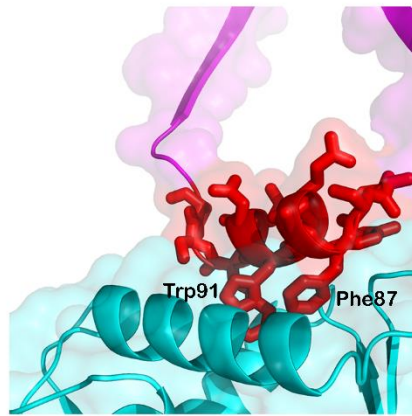
Fig. 13. Estructura cristalográfica del complejo EhICP1-p53(FG)/MDM2. **A**, diagrama de listones del heterotetrámero formado por 2 moléculas de EhICP1-p53(FG) y dos moléculas de MDM2. El tetrámero es estabilizado por una interconexión de dominios o *swapping* entre dos moléculas de EhICP1-p53(FG), en azul y amarillo, se muestra la vista de 90 °. Cada EhICP1-p53(FG) interacciona con una molécula de MDM2, negro y magenta. **B**, el péptido de p53 introducido en el asa FG de EhICP1 adopta una conformación de α -hélice y forma una interacción con MDM2 en su cavidad hidrofóbica (los residuos de MDM2 están coloreados de amarillo (hidrofóbicos), magenta (negativos) y cian (positivos), de acuerdo al script YRB de Pymol [53], la parte superior muestra un alineamiento estructural del complejo EhICP1-p53(FG)/MDM2 con el dodecapéptido p53/MDM2 (PDB ID: 3EQS). La región correspondiente al péptido en EhICP1-p53(FG) está representada en azul, mientras que el péptido de p53 está en color rojo. La parte inferior de B muestra un Omitmap de la secuencia del péptido de p53 en la quimera del asa FG en la cavidad de MDM2. **C**, Omitmap de la α -hélice de EhICP1-p53(FG) que interactúa con MDM2 y de la hebra β que se interconecta con la segunda molécula de EhICP1-p53(FG).

La estructura cristalográfica del complejo EhICP1-p53(FG)/MDM2 muestra que el péptido insertado, correspondiente a la zona de transactivación de p53 adopta una conformación correcta de α -hélice restringida conformacionalmente por el andamio EhICP1, y además se interna en la cavidad de MDM2 formando una interfaz muy similar a la observada en el complejo canónico p53/MDM2 [41], **Fig.13B**, parte inferior. La estructura cristalográfica también muestra una interconexión de dominios entre las dos moléculas de EhICP1-p53(FG) contenidas en la estructura. El *swapping* ocurre por un intercambio de una hebra β con la subunidad vecina, **Fig. 13C**. En la estructura heterotetramérica hay 2 sitios de interacción, uno para cada par EhICP1-p53(FG)/MDM2. El detalle estructural se muestra en la **Fig. 14**.

A modo de comparación en la interfaz p53/MDM2 y para verificar la capacidad de los ICPs para ser utilizados como andamios para el diseño racional de otras proteínas de unión, realizamos una comparación estructural entre la quimera aquí generada y cristalizada con MDM2 y un anticuerpo contra MDM4 (un ortólogo de MDM2), **Fig. 14**. Esta comparación estructural remarca la convergencia estructural entre ambos complejos.

En el anticuerpo (de un solo dominio) se muestran dos residuos aromáticos conservados, Trp100 y Tyr101, éstos interactúan de la misma forma en que p53 y MDM2 lo hacen. EhICP1-p53(FG) interactúa con MDM2 únicamente con los aminoácidos que le fueron introducidos, del péptido de p53, mientras que el anticuerpo presenta además interacciones adicionales, como puentes de hidrógeno entre los residuos Glu110 y Tyr99 de MDM4, o Phe111 que se interna en la cavidad de MDM4 (**Fig. 14B**).

A EhICP1-p53(FG)/MDM2



B Anticuerpo/MDM4

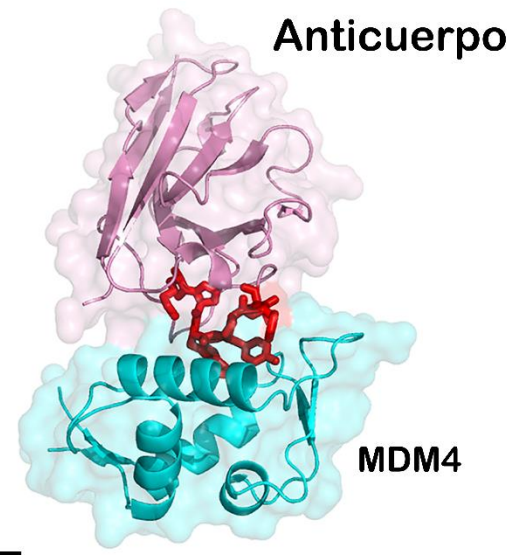
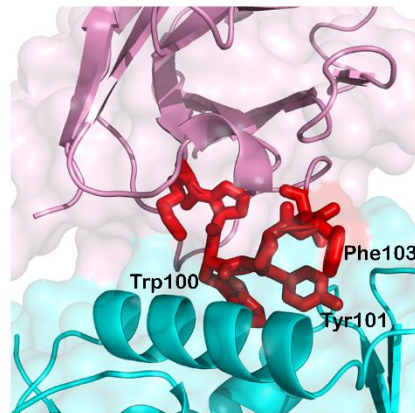


Fig. 14. Comparación estructural entre los complejos EhICP1-p53(FG)/MDM2 y un anticuerpo contra MDM4. **A**, Estructura cristalográfica del complejo EhICP1-p53(FG) en complejo con MDM2. Los residuos Phe87 y Trp91 forman interacciones principalmente hidrofóbicas con MDM2. **B**, Estructura cristalográfica del complejo anticuerpo/MDM4. El asa de reconocimiento del anticuerpo, en rojo, tiene los residuos Trp100, Tyr101 y Phe103 que establecen interacciones con MDM4. En ambas estructuras, las principales interacciones están mediadas por un triptófano. La parte izquierda de la imagen es un acercamiento de la parte derecha.

5.9 El EhICP1 forma agregados de alto peso molecular y fibrillas del tipo amiloide

La estructura cristalográfica muestra un arreglo de intercambio de dominios entre las construcciones del EhICP1, lo que sugiere que el EhICP1 podría, bajo ciertas circunstancias, formar agregados moleculares mediados por este mecanismo de *swapping*. Debido a la tendencia de ciertos ICPs de formar agregados de alto peso molecular [54] y a la formación de fibras amiloides de otros miembros de la familia de las Igs [55-57], además a la capacidad del EhICP1 de formar oligómeros y otros agregados moleculares (**Fig.4**), decidimos evaluar la formación de fibras del tipo amiloide constituidas por la proteína silvestre EhICP1. La incubación de la proteína purificada a 42 °C durante 24 horas resultó en la formación de estructuras fibrilares y otros agregados, como se muestra en la **Fig. 15**. Las fibras observadas por TEM tienen tamaños que van de los 200 a cerca de 1000 nm de longitud y un diámetro aproximado de 40 nm, sin embargo también se aprecian abundantes agregados amorfos. Lo anterior confirma la capacidad del EhICP1 silvestre de formar estructuras del tipo amiloide al menos *in vitro*.

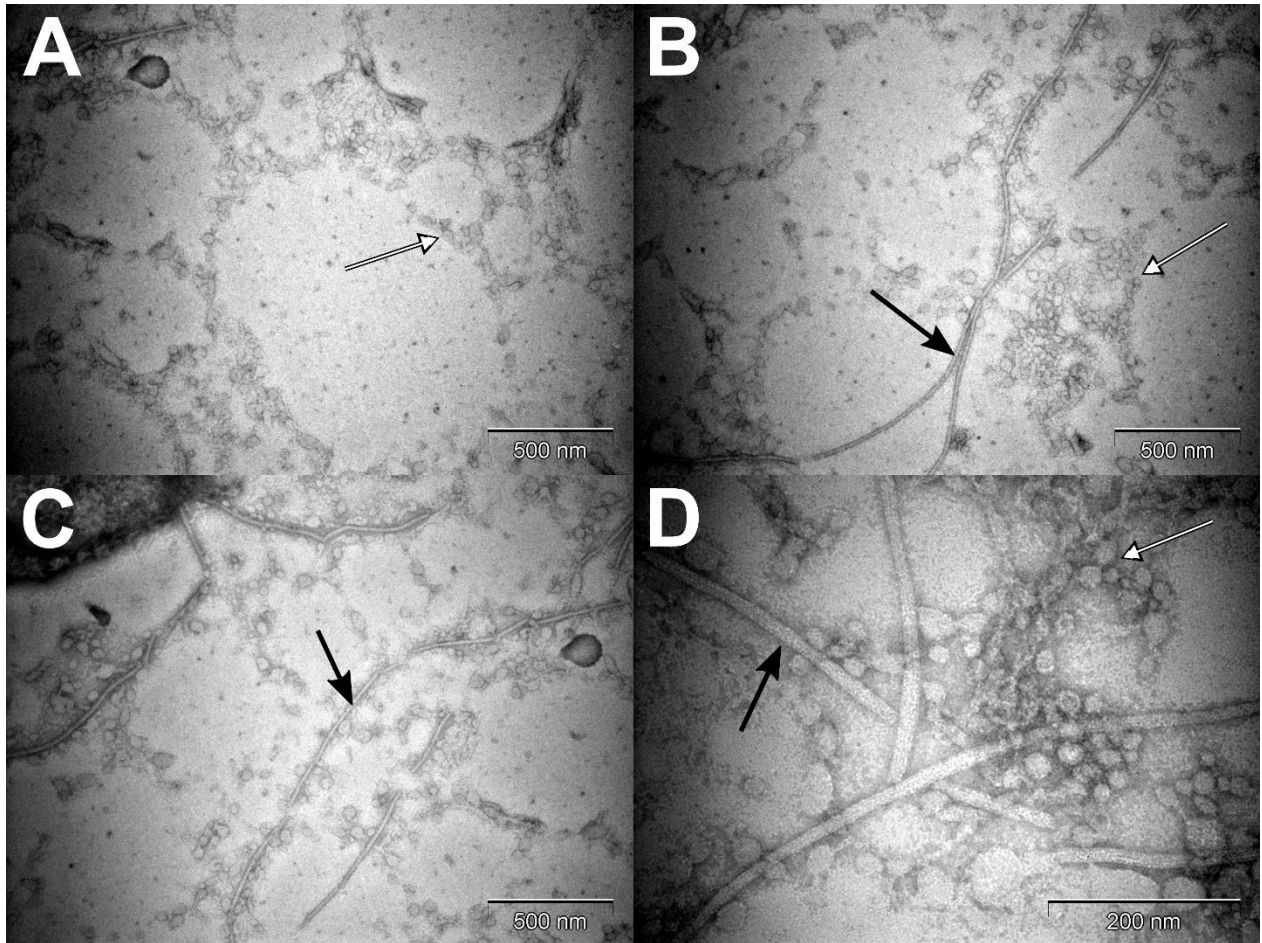


Fig. 15. Análisis de agregados moleculares y formación de fibras en EhICP1 silvestre por TEM.

Diferentes imágenes de microscopía de los agregados producidos por EhICP1 silvestre, en **A** se observan múltiples agregados amorfos, en **B** y en **C** se aprecian estructuras fibrilares que van de los 200 nm a los 800 nm. En **D** se muestra una ampliación en la región con las estructuras fibrilares obtenidas. Las flechas negras señalan las fibrillas, los agregados se indican con una flecha blanca. La escala se encuentra indicada en la parte inferior derecha de cada fotografía.

VI. Discusión

Los andamios proteicos con un plegamiento del tipo inmunoglobulina representan una valiosa plataforma para el diseño racional de proteínas, no en vano la naturaleza ha explotado el dominio Ig para la generación de un número muy grande de anticuerpos y otras proteínas de unión. El diseño racional sobre anticuerpos como andamios tiene importantes retos, principalmente por el contenido de puentes disulfuro y los

consiguientes problemas de purificación y agregación cuando son obtenidos de forma recombinante en bacteria, a pesar de que esas limitaciones han sido parcialmente solucionadas en ciertos casos mediante el uso de cepas especializadas y distintos métodos de selección [9, 10]. En este trabajo presentamos una estrategia para el diseño racional de proteínas, que no tiene las limitaciones de las Igs con puentes disulfuro; el uso del EhICP1 como andamio proteico, que puede tolerar la sustitución de dos de sus asas, DE y FG, y permanecer correctamente plegado. Debido a la similitud estructural con otros ICPs, es altamente probable que otros miembros de la familia MEROPS I42 puedan ser utilizados como andamios proteicos, en una forma muy similar a la que aquí desarrollamos con el ICP1 de *E. histolytica*.

MDM2 ha sido blanco de múltiples procesos de diseño de interfaces proteicas, por ejemplo con el uso de tioredoxina y la proteína no estructurada 4EBP1 como andamios [47] [58]. El péptido sintético N-MPRFMDYWEGLN se une a MDM2 y tiene una afinidad 2 órdenes de magnitud mayor a la de la secuencia canónica p53. [59]. Utilizando a tioredoxina y a la proteína 4EBP1 como andamios, mediante la inserción del péptido N-MPRFMDYWEGLN resultó en una afinidad aún mayor a la del péptido solo, en concreto 2 órdenes de magnitud mayor [47]. En los casos anteriores el uso de un andamio para exponer una secuencia peptídica incrementó la afinidad de la interacción con MDM2. Este incremento en afinidad usando los andamios puede estar relacionado con la baja libertad conformacional ya que el péptido p53 se encuentra restringido en el andamio. Como mencionamos, el péptido optimizado de p53 se une con una afinidad de 3.36×10^{-9} M, aproximadamente 100 veces mayor que la del péptido canónico de p53 (N-ETFSDLWKLLPE) [41]. La afinidad observada utilizando al EhICP1 como andamio fue más baja, en el orden micromolar. La alta afinidad observada en el péptido libre se atribuye a una red de puentes de hidrógeno más extendida con MDM2 [41]. La formación del dominio de interconexión o *swapping* entre EhICP1-p53(FG) podría explicar la disminución en la afinidad, comparado con el dodecapéptido. Las limitaciones espaciales podrían ser responsables de la formación del *swapping* en la quimera de EhICP1-p53, **Fig.13**.

A pesar de que la afinidad obtenida para las quimeras de EhICP1-p53 con MDM2 fue menor que la observada para el péptido libre, la afinidad que logramos está en el mismo rango de magnitud que la observada para la interacción canónica p53/MDM2, por lo que además de mimetizar la interfaz en nuestro andamio, logramos también una afinidad similar [50]. En el caso de que fuera deseable conseguir quimeras de EhICP1-p53 con una afinidad mayor, serían necesarios otros experimentos para conseguir construcciones que permitieran ligeras variaciones conformacionales en el dodecapéptido insertado, como, por ejemplo, insertar un par de residuos que permitan adoptar una conformación más estable al péptido insertado. Nos referimos a la selección y mejoramiento por técnicas como la evolución dirigida, que es utilizada típicamente en combinación con el diseño racional [23]. Sin embargo, el objetivo esencial de este trabajo no es la generación de una quimera EhICP1-p53 con una altísima afinidad a MDM2, sino demostrar que es factible el uso de éste y otros ICPs como andamios proteicos para el diseño racional, debido al bajo peso molecular, fácil de expresar y purificar en bacteria y a las características bioquímicas y estructurales ya mencionadas y que resultan ventajosas frente a otros andamios sugeridos a la fecha.

La formación del dominio de intercambio o *swapping* observada en nuestro complejo resultó sorprendente y por demás interesante y pudiera explicar algunos fenómenos observados en otros sistemas. Ya que la formación del complejo EhICP1-p53(FG)/MDM2 es independiente de condiciones especiales de incubación y/o tratamiento, es posible que otros ICPs como chagasina o el EhICP2 puedan formar ensamblajes o agregados de alto peso molecular originados por *swapping* [60]. Está reportado que los ICPs del género *Plasmodium* forman oligómeros y agregados de peso molecular mayor a 100 kDa y que además tienen actividad inhibitoria. Lo anterior sugiere que los ICPs tienen una disposición a la formación de agregados de alto peso molecular [54]. Otras moléculas con un plegamiento del tipo Ig y que tienen una tendencia a la agregación y a la formación de fibras amiloides (los amiloides están constituidos principalmente por hebras β) son las cadenas ligeras de los anticuerpos [55-57]. Recientemente se ha reportado la formación de fibras amiloides en proteínas que no son del sistema inmune, como la proteína muscular titina [61]. Aparentemente las Igs tienen una tendencia natural a la agregación y posiblemente a la formación de

fibras. A partir de la evidencia anterior, con el dominio de intercambio o *swapping* observado en la quimera EhICP1-p53 y con la confirmación de formación de fibras de la proteína silvestre EhICP1, postulamos que el mecanismo de *swapping* puede ser un mecanismo en la formación de amiloides en otros sistemas.

VII. Conclusiones y Perspectivas

Los andamios proteicos o péptido aptámeros son una importante herramienta biotecnológica que permite la generación de proteínas novedosas y no existentes en la naturaleza, lo que permite hacer uso de la imaginación para el diseño racional de proteínas con ciertas características deseables. El EhICP1, al igual que otros ICPs tiene un plegamiento del tipo Ig estable y robusto que permite la modificación estructural de sus asas DE y FG para ser modificadas mediante el diseño racional en la ingeniería de proteínas. Nuestra aportación con el andamio EhICP1 en el área del diseño de proteínas con funciones predichas, permitirá el diseño de otras proteínas de unión, cuyos requerimientos para la interacción son ya conocidos, así como su aplicación como alternativa a los anticuerpos monoclonales (con menor costo), su uso como biosensores e incluso a largo plazo su posible uso como agentes terapéuticos o de diagnóstico.

La inesperada interconexión de dominios, *swapping*, en la quimera EhICP1-p53 en el complejo cristalizado tiene una relevancia estructural importante para los procesos de oligomerización de otras proteínas similares y aporta pistas en el fenómeno de la formación de fibras del tipo amiloide, que ocurre con cierta frecuencia en otras proteínas con el mismo plegamiento y algunos miembros de la superfamilia de las Igs. El EhICP1 además puede ser utilizado como sistema modelo para el estudio de formación de estas estructuras oligoméricas y fibrilares. Los datos aquí presentados sobre la capacidad del EhICP1 de formar fibras del tipo amiloide dejan muchas puertas abiertas para la investigación del fenómeno en éste y en otros ICP, así como las posibles repercusiones biológicas.

VIII. Apéndice A

Genes correspondientes a las quimeras del EhICP1 y el dominio N terminal de Mdm2

Todas las variantes de EhICP1 fueron expresadas utilizando el sistema de expresión pET19b, el cual añade una cola de 10 histidinas seguido de una zona de corte para la proteasa PPS, dicha región comprende los siguientes aminoácidos:

MGHHHHHHHHHSSGHIDDDDKHTSLEVLFGQPH, donde el sitio de corte de la PPS está después de Q, es decir, deja como cicatriz los tres residuos GPH, por ejemplo, la proteína MDM2 digerida quedaría como sigue:

Proteína codificada y digerida con PPS

GPHMCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYLGQYI
MTKRLYDEKQQHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLLVVVNQQESSDSGTSVSEN
RCHLEGGSDQKDLVQELQEEKPSSS

Número de aminoácidos: 153
Peso molecular: 17.218 kDa
pl teórico: 5.80

EhICP1-p53(BC)

GACGACAAGCATATGTCCCTGACCGAAGATAACAATAACACGACCATTACCATCGCAAAG
GTGAAAACAAGAAATCATCCTGCACGGCAATCCGACGACGAGCTTTGCGGAATATTGGAA
CCTGCTGAGTCCGACGGGCTATAGCTGGGTGGTTGATTCTTCAGAAGGTCTGAGTAACACG
GTTGAATACGTCCGACAGACCATGCTCCGGGTATTTACGGTTCAGGCGGTAAATATCACA
TTAAAATCACCGGCACGCAAACCGGCGAAGGTAATCGTCCTGGTGTACCGTCGTGGTG
CAGCACCGAACGAAATGACCGCACCTTTACGCTGAAAGTGAATGTCCAATGAGGATCCGG
CTGCTAA

Proteína codificada

MGHHHHHHHHHSSGHIDDDDKHTSLEVLFGQPHMSLTEDNNNTTITIAKGENKEIILHGNPTT
SFAEYWNLLSPTGYSWVVDSSSEGLSNTVEYVADQHAPGISGSGGKYHIKITGTQTGEGKIVLVY
RRGAAPNANDRTFTLKVNVQ

Número de aminoácidos: 148
Peso molecular: 16.306 kDa
pl teórico: 6.21

EhICP1-p53(DE)

GACGACAAGCATATGTCCCTGACCGAAGATAACAATAACACGACCATTACCATCGCAAAG
GTGAAAACAAAGAAATCATCCTGCACGGCAATCCGACGACGGGCTATAGCTGGGTGGTTG
ATTCTTCAGAAGGTCTGAGTAACACGGTTGAATACGTCGCAGACCAGACCAGCTTTGCGGA
ATATTGGAACCTGCTGAGCCCGGGTTCAGGCGGTAAATATCACATTAATCACCGGCACG
CAAACCGGCGAAGGTAAATCGTCCTGGTGTACCGTCGTGGTGCCGCACCGAACGCAAT
GACCGCACCTTTACGCTGAAAGTGAATGTCCAATGAGGATCCGGCTGCTAA

Proteína codificada

MGHHHHHHHHHSSGHIDDDDKHTSLEVLFGQPHMSLTEDNNNTTITIAKGENKEIILHGNPTT
GYSWVVDSSSEGLSNTVEYVADQTSFAEYWNLLSPGSSGKYHIKITGTQTGEGKIVLVYRRGAA
PNANDRTFTLKVNVQ

Número de aminoácidos: 142
Peso molecular: 15.743 kDa
pI teórico: 6.17

EhICP1-p53(FG)

GACGACAAGCATATGTCCCTGACCGAAGATAACAATAACACGACCATTACCATCGCAAAG
GTGAAAACAAAGAAATCATCCTGCACGGCAATCCGACGACGGGCTATAGCTGGGTGGTTG
ATTCTTCAGAAGGTCTGAGTAACACGGTTGAATACGTCGCAGACCAGCATGCTCCGGGTAT
TTCAGGTTCAAGCGGTAAATATCACATTAATCACCGGCACGCAAACCGGCGAAGGTAA
ATCGTCCTGGTGTACCGTCGTACCAGCTTTGCGGAATATTGGAACCTGCTGAGCCCGGACC
GCACCTTTACGCTGAAAGTGAATGTCCAATGAGGATCCGGCTGCTAA

Proteína codificada

MGHHHHHHHHHSSGHIDDDDKHTSLEVLFGQPHMSLTEDNNNTTITIAKGENKEIILHGNPTT
GYSWVVDSSSEGLSNTVEYVADQHAPGISGSSGKYHIKITGTQTGEGKIVLVYRRTSFAEYWNLL
SPDRFTFTLKVNVQ

Número de aminoácidos: 141
Peso molecular: 15.710 kDa
pI teórico: 6.21

MDM2

GACGACAAGCATATGTGCAACACCAACATGAGCGTGCCGACCGATGGCGCGGTGACCACC
AGCCAGATTCCGGCGAGCGAACAGGAAACCCTGGTGCGCCCGAAACCGCTGCTGCTGAAA
CTGCTGAAAAGCGTGGGCGCGCAGAAAGATACCTATAACCATGAAAGAAGTGCTGTTTTATC
TGGGCCAGTATATTATGACCAAACGCCTGTATGATGAAAAACAGCAGCATATTGTGTATTGC
AGCAACGATCTGCTGGGCGATCTGTTTGGCGTGCCGAGCTTTAGCGTGAAAGAACATCGC
AAAATTTATACCATGATTTATCGCAACCTGGTGGTGGTGAACCAGCAGGAAAGCAGTGATA
GCGGCACCAGCGTGAGCGAAAACCGCTGCCATCTGGAAGGTGGAAGCGATCAGAAAGATC
TGGTGCAGGAAGTGCAGGAAGAAAAACCGAGCAGCAGCTGAGGATCCGGCTGCTAA

Proteína codificada

MGHHHHHHHHHSSGHIDDDDKHTSLEVLFGQPHMCNTNMSVPTDGAVTTSQIPASEQETLV
RPKPLLLKLLKSVGAKDYTMKEVLFYLGQYIMTKRLYDEKQQHIVYCSNDLLGDLFGVPSFSV
KEHRKIYTMIRNLVVVNQQESSDSGTSVSENRCHEGGSDQKDLVQELQEEKPSSS

Número de aminoácidos: 153
Peso molecular: 17.218 kDa
pI teórico: 5.80

Nota: Los sitios de restricción BamHI y NdeI están subrayados.

Tabla 5. Tabla de aminoácidos y nomenclatura.

Aminoácido	Código de tres letras	Código de una letra
Alanina	Ala	A
Arginina	Arg	R
Asparagina	Asn	N
Ácido aspártico	Asp	D
Cisteína	Cys	C
Glutamina	Gln	Q
Ácido glutámico	Glu	E
Glicina	Gly	G
Histidina	His	H
Isoleucina	Ile	I
Leucina	Leu	L
Lisina	Lys	K
Metionina	Met	M
Fenilalanina	Phe	F
Prolina	Pro	P
Serina	Ser	S
Treonina	Thr	T
Triptófano	Trp	W
Tirosina	Tyr	Y
Valina	Val	V

IX. Bibliografía

1. De Las Rivas, J. and C. Fontanillo, *Protein–protein interactions essentials: key concepts to building and analyzing interactome networks*. PLoS computational biology, 2010. **6**(6): p. e1000807.
2. Cusick, M.E., et al., *Interactome: gateway into systems biology*. Human molecular genetics, 2005. **14**(suppl_2): p. R171-R181.
3. Patel, A.K., R.R. Singhanian, and A. Pandey, *Production, purification, and application of microbial enzymes*, in *Biotechnology of Microbial Enzymes*. 2017, Elsevier. p. 13-41.
4. Lutz, S., *Beyond directed evolution—semi-rational protein engineering and design*. Current opinion in biotechnology, 2010. **21**(6): p. 734-743.
5. Dill, K.A., *Dominant forces in protein folding*. Biochemistry, 1990. **29**(31): p. 7133-7155.
6. Colas, P., et al., *Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2*. Nature, 1996. **380**(6574): p. 548.
7. Honegger, A., *Engineering antibodies for stability and efficient folding*, in *Therapeutic Antibodies*. 2008, Springer. p. 47-68.
8. Skerra, A., *Bacterial expression of immunoglobulin fragments*. Current opinion in immunology, 1993. **5**(2): p. 256-262.
9. Mazor, Y., et al., *Isolation of engineered, full-length antibodies from libraries expressed in Escherichia coli*. Nature biotechnology, 2007. **25**(5): p. 563.
10. Makino, T., et al., *Comprehensive engineering of Escherichia coli for enhanced expression of IgG antibodies*. Metabolic engineering, 2011. **13**(2): p. 241-251.
11. Chlewicki, L.K., et al., *High-affinity, peptide-specific T cell receptors can be generated by mutations in CDR1, CDR2 or CDR3*. Journal of molecular biology, 2005. **346**(1): p. 223-239.
12. Hiipakka, M., K. Poikonen, and K. Saksela, *SH3 domains with high affinity and engineered ligand specificities targeted to HIV-1 Nef1*. Journal of molecular biology, 1999. **293**(5): p. 1097-1106.
13. Koide, A., et al., *The fibronectin type III domain as a scaffold for novel binding proteins1*. Journal of molecular biology, 1998. **284**(4): p. 1141-1151.
14. Reina, J., et al., *Computer-aided design of a PDZ domain to recognize new target sequences*. Nature Structural and Molecular Biology, 2002. **9**(8): p. 621.
15. Binz, H.K., et al., *High-affinity binders selected from designed ankyrin repeat protein libraries*. Nature biotechnology, 2004. **22**(5): p. 575.
16. Abedi, M.R., G. Caponigro, and A. Kamb, *Green fluorescent protein as a scaffold for intracellular presentation of peptides*. Nucleic acids research, 1998. **26**(2): p. 623-630.
17. Reverdatto, S., D.S. Burz, and A. Shekhtman, *Peptide aptamers: development and applications*. Current topics in medicinal chemistry, 2015. **15**(12): p. 1082.
18. Skerra, A., *Alternative non-antibody scaffolds for molecular recognition*. Current opinion in biotechnology, 2007. **18**(4): p. 295-304.
19. Bork, P., L. Holm, and C. Sander, *The immunoglobulin fold: structural classification, sequence patterns and common core*. Journal of molecular biology, 1994. **242**(4): p. 309-320.
20. Bork, P. and R.F. Doolittle, *Proposed acquisition of an animal protein domain by bacteria*. Proceedings of the National Academy of Sciences, 1992. **89**(19): p. 8990-8994.
21. Mei, S., et al., *Solution structure of a bacterial immunoglobulin-like domain of the outer membrane protein (L ig B) from Leptospira*. Proteins: Structure, Function, and Bioinformatics, 2015. **83**(1): p. 195-200.
22. Fowler, S.B. and J. Clarke, *Mapping the folding pathway of an immunoglobulin domain: structural detail from phi value analysis and movement of the transition state*. Structure, 2001. **9**(5): p. 355-366.

23. Wright, C.F., et al., *The importance of loop length in the folding of an immunoglobulin domain*. Protein Engineering Design and Selection, 2004. **17**(5): p. 443-453.
24. Tian, P., et al., *The Folding Pathway of an Ig Domain is Conserved On and Off the Ribosome*. bioRxiv, 2018: p. 253013.
25. Batori, V., A. Koide, and S. Koide, *Exploring the potential of the monobody scaffold: effects of loop elongation on the stability of a fibronectin type III domain*. Protein engineering, 2002. **15**(12): p. 1015-1020.
26. Jimenez-Sandoval, P., et al., *Cysteine Proteases Inhibitors with Immunoglobulin-Like Fold in Protozoan Parasites and their Role in Pathogenesis*. Current Protein and Peptide Science, 2017. **18**(10): p. 1035-1042.
27. Rawlings, N.D., A.J. Barrett, and R. Finn, *Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors*. Nucleic acids research, 2015. **44**(D1): p. D343-D350.
28. Wang, S.X., et al., *The structure of chagasin in complex with a cysteine protease clarifies the binding mode and evolution of an inhibitor family*. Structure, 2007. **15**(5): p. 535-543.
29. Costa, T.F. and A.P.C. Lima, *Natural cysteine protease inhibitors in protozoa: Fifteen years of the chagasin family*. Biochimie, 2016. **122**: p. 197-207.
30. Gilbreth, R.N., et al., *A dominant conformational role for amino acid diversity in minimalist protein–protein interfaces*. Journal of molecular biology, 2008. **381**(2): p. 407-418.
31. Lipovšek, D., et al., *Evolution of an interloop disulfide bond in high-affinity antibody mimics based on fibronectin type III domain and selected by yeast surface display: molecular convergence with single-domain camelid and shark antibodies*. Journal of molecular biology, 2007. **368**(4): p. 1024-1041.
32. Karatan, E., et al., *Molecular recognition properties of FN3 monobodies that bind the Src SH3 domain*. Chemistry & biology, 2004. **11**(6): p. 835-844.
33. Huang, J., et al., *Design of protein function leaps by directed domain interface evolution*. Proceedings of the National Academy of Sciences, 2008. **105**(18): p. 6578-6583.
34. Hackel, B.J., A. Kapila, and K.D. Wittrup, *Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling*. Journal of molecular biology, 2008. **381**(5): p. 1238-1252.
35. Gilbreth, R.N. and S. Koide, *Structural insights for engineering binding proteins based on non-antibody scaffolds*. Current opinion in structural biology, 2012. **22**(4): p. 413-420.
36. Casados-Vázquez, L.E., S. Lara-González, and L.G. Brieba, *Crystal structure of the cysteine protease inhibitor 2 from Entamoeba histolytica: functional convergence of a common protein fold*. Gene, 2011. **471**(1): p. 45-52.
37. da Silva, A.A.F., et al., *Crystal structure of chagasin, the endogenous cysteine-protease inhibitor from Trypanosoma cruzi*. Journal of structural biology, 2007. **157**(2): p. 416-423.
38. Redzynia, I., et al., *Crystal structure of the parasite inhibitor chagasin in complex with papain allows identification of structural requirements for broad reactivity and specificity determinants for target proteases*. The FEBS journal, 2009. **276**(3): p. 793-806.
39. Leahy, D.J., R. Axel, and W.A. Hendrickson, *Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2.6 Å resolution*. Cell, 1992. **68**(6): p. 1145-1162.
40. Garcia, K.C., et al., *Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen*. Science, 1998. **279**(5354): p. 1166-1172.
41. Pazgier, M., et al., *Structural basis for high-affinity peptide inhibition of p53 interactions with MDM2 and MDMX*. Proceedings of the National Academy of Sciences, 2009. **106**(12): p. 4665-4670.
42. Battye, T.G.G., et al., *iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM*. Acta Crystallographica Section D: Biological Crystallography, 2011. **67**(4): p. 271-281.

43. McCoy, A.J., et al., *Phaser crystallographic software*. Journal of applied crystallography, 2007. **40**(4): p. 658-674.
44. Afonine, P.V., et al., *Towards automated crystallographic structure refinement with phenix.refine*. Acta Crystallographica Section D: Biological Crystallography, 2012. **68**(4): p. 352-367.
45. Vagin, A.A., et al., *REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use*. Acta Crystallographica Section D: Biological Crystallography, 2004. **60**(12): p. 2184-2195.
46. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography*. Acta Crystallographica Section D: Biological Crystallography, 2010. **66**(1): p. 12-21.
47. Brown, C.J., et al., *Rational design and biophysical characterization of thioredoxin-based aptamers: insights into peptide grafting*. Journal of molecular biology, 2010. **395**(4): p. 871-883.
48. Parker, M., et al., *Antibody mimics based on human fibronectin type three domain engineered for thermostability and high-affinity binding to vascular endothelial growth factor receptor two*. Protein Engineering Design and Selection, 2005. **18**(9): p. 435-444.
49. Chen, H., L. Xie, and B. Liu, *Clinical significance of MDM2 as a tumor biomarker*. The Chinese-German Journal of Clinical Oncology, 2012. **11**(6): p. 356-360.
50. Grace, W.Y., et al., *The central region of HDM2 provides a second binding site for p53*. Proceedings of the National Academy of Sciences, 2006. **103**(5): p. 1227-1232.
51. Yang, J., et al., *The I-TASSER Suite: protein structure and function prediction*. Nature methods, 2015. **12**(1): p. 7.
52. Wright, T.A., et al., *Extraction of Thermodynamic Parameters of Protein Unfolding Using Parallelized Differential Scanning Fluorimetry*. The journal of physical chemistry letters, 2017. **8**(3): p. 553-558.
53. Hagemans, D., et al., *A script to highlight hydrophobicity and charge on protein surfaces*. Frontiers in molecular biosciences, 2015. **2**: p. 56.
54. Sundararaj, S., et al., *Cross-talk between malarial cysteine proteases and falciparum: the BC loop as a hot-spot target*. PloS one, 2014. **9**(4): p. e93008.
55. Khurana, R., et al., *A model for amyloid fibril formation in immunoglobulin light chains based on comparison of amyloidogenic and benign proteins and specific antibody binding*. Amyloid, 2003. **10**(2): p. 97-109.
56. Ramirez-Alvarado, M., *Amyloid formation in light chain amyloidosis*. Current topics in medicinal chemistry, 2012. **12**(22): p. 2523-2533.
57. Oberti, L., et al., *Concurrent structural and biophysical traits link with immunoglobulin light chains amyloid propensity*. Scientific reports, 2017. **7**(1): p. 16809.
58. See, H.Y. and D.P. Lane, *A novel unstructured scaffold based on 4EBP1 enables the functional display of a wide range of bioactive peptides*. Journal of molecular biology, 2010. **404**(5): p. 819-831.
59. Böttger, V., et al., *Identification of novel mdm2 binding peptides by phage display*. Oncogene, 1996. **13**(10): p. 2141-2147.
60. Blancas-Mejía, L.M. and M. Ramirez-Alvarado, *Systemic amyloidoses*. Annual review of biochemistry, 2013. **82**: p. 745-774.
61. Bobylev, A.G., et al., *Smooth muscle titin forms in vitro amyloid aggregates*. Bioscience reports, 2016: p. BSR20160066.