

UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS QUÍMICAS Y FARMACÉUTICAS



***Caracterización del mecanismo de plegamiento de una
versión monomérica del represor ARC, un miembro de la
superfamilia RHH con topología anudada, mediante
espectroscopia de fuerza de trampa óptica***

**Tesis presentada a la Universidad de Chile para optar al grado de
Magíster en Bioquímica área de Especialización en Proteínas
Recombinantes y Biotecnología y Memoria para optar al Título de
Bioquímico por:**

Andrés Ignacio Bustamante Gatica

Director de Tesis: Dr. Mauricio Báez Larach

Santiago-CHILE

Abril 2014

UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS QUÍMICAS Y FARMACÉUTICAS

INFORME DE APROBACIÓN DE TESIS DE MAGÍSTER

Se informa a la Dirección de la Escuela de Graduados de la Facultad de Ciencias Químicas y Farmacéuticas que la Tesis de Magíster y Memoria de Título presentada por el candidato

ANDRÉS IGNACIO BUSTAMANTE GATICA

Ha sido aprobada por la Comisión evaluadora de Tesis como requisito para optar al grado de Magíster en Bioquímica, Área de Especialización: Proteínas Recombinantes y Biotecnología y Título de Bioquímico, en el examen público rendido el día: _____

Director de Tesis:

Dr. Mauricio Báez Larach

Comisión Evaluadora de Tesis:

Dr. Mario Chiong

Dra. Victoria Guixé

Dra. María Antonieta Valenzuela

*“...A mi viejo y heroico Iquique...
Tu ambiente y la nobleza de tu gente
cautivan el corazón”*

Agradecimientos

En primer lugar, deseo agradecer a Dios y a mi Chinita por darme la fuerza para enfrentar todos los desafíos que se me presentaron a lo largo del desarrollo de esta tesis y, en general, de toda la carrera.

De la misma forma deseo agradecer a mi familia, a mi mamita, Frida, a mi papito, Fernando y a mi hermanita, Andrea. Sin el apoyo incondicional de ellos esto jamás hubiese sido posible. Ellos siempre encontraron la forma de estar cerca (a pesar de la distancia) y su amor nutrió el esfuerzo de todos estos años. Mamá, gracias por alimentar mi curiosidad y mostrarme el hermoso mundo de la ciencia, gracias por compartir tu pasión por enseñar conmigo. Papá, gracias por mostrarme el valor de la responsabilidad y de la perseverancia, gracias por fomentar la creatividad en mí y las ganas de aprender. Andrea, gracias por ser esa compañera incansable y preocupada, por acompañar mis noches de estudio y por regalarme siempre esos abrazos tiernos y cariñosos sin mayor razón.

No puedo dejar de mencionar a Robinson Herrera y Reynaldo García, mis profes de química, quienes alimentaron mi gustito por la química y me dieron la oportunidad de encontrar nuevas fronteras en el área. También agradecer a don Henrique Muñoz, mi profe de biología.

Tampoco puedo dejar pasar la oportunidad de agradecer a mis eternas amigas, Clau, Dani y Mathy. Niñas, conocerlas fue una de las cosas más hermosas de mi vida. Gracias por todas esas juntas locas en Iquique y en Santiago, por ayudarme a escapar de la rutina y por todas las risas regaladas en esas innumerables conversaciones, reuniones, whats'apps, comilonas, etc.

A aquellos amigos que el tiempo convirtió en hermanos, Crsitiancito, Martín, Jony, Juany, Seba, Hector, David. Todos ustedes me hicieron sentir en casa, ya sea aquí en Santiago o de regreso en Iquique. Qué hubiese sido de mí sin los asados, paseos, tarreos, changoseos, dados, caminatas, conversaciones, etc. Gracias por recordarme siempre que hay más vida fuera del laboratorio, por secuestrarme cada verano y darme un espacio para jugar y reír como cuando éramos niños. Gracias por alimentar mi autoestima con comentarios como "...es que tú no eres humano..." o "...bueno, nosotros los normales..." y por controlar (y a veces aprovecharse) de mi consumismo compulsivo.

Cómo dejar de mencionar a aquellos amigos que este paso por la casa de Bello me entregó. Don Gabo y Don Nacho, ustedes tienen un serio problema pero afortunadamente es el mismo

que el mío. Gracias por aquellas largas discusiones (científicas y no-científicas), por la misantropía y el leseo (y el HBH).

También están mis amigos del lab, Javi, Alejo, Mati, Seba, Cristian (profesor), Mae, Feña, AnBar, Mauricio, Maira. Gracias por todos los momentos, por esas juntas en la casa de la Javi y las celebraciones en el lab. Por las risas y las bromas, por la convivencia que hizo de ir al laboratorio día a día fuera algo placentero (incluso cuando todo parecía estar en mi contra). Gracias por los concejos y la sabiduría transmitidas (ya sea durante los almuerzos en vicuña, en el Bella o en algún congreso).

Quiero agradecer de manera muy especial al Dr. Maurico Baéz (mi director de tesis) por la confianza depositada y todas las oportunidades dadas, por enseñarme una forma de hacer ciencia. Del mismo modo, quiero agradecer al Dr. Christian Wilson por todo el apoyo, discusiones y consejos entregados durante el desarrollo de la tesis. Así mismo, agradecer a nuestros colaboradores peruanos ya que sin su ayuda esta tesis no sería una realidad, al Dr. Daniel Guerra, Dr. Carlos Bustamante, Pamela Obando y Robert Sosa, muchas gracias.

También quiero agradecer a la Dra. María Antonieta Valenzuela, no sólo por permitirme dar mis primeros pasos de investigación en su laboratorio, también por su amistad, su consejo y su cariño. Profe, desde que la conocí ha sido un gran pilar para mi, muchas gracias.

Finalmente, quiero agradecer a la persona que me ha acompañado desde el inicio de este viaje llamado Universidad, mi angelito de otoño, Stefania Iannuzzi. Estar lejos del hogar es difícil, sobre todo siendo tan apegado a la tierra como soy. Pero estando contigo siempre estoy en mi hogar, gracias por acompañarme en todos los momentos, por mostrarme cada día que puedo ser mejor. Gracias por cuidarme y por enseñarme el valor de la familia, gracias por amarme. Extiendo este agradecimiento también a tu familia, quienes me han cobijado durante estos siete años y me han hecho sentir uno más.

A todos aquellos que he mencionado, y también todos aquellos que pude haber olvidado, quiero que sepan que sin su apoyo, consejos, cariño y amistad no podría haber logrado esto. Nuevamente, MUCHAS GRACIAS.

Presentaciones a congresos

Los resultados de esta tesis dieron origen total o parcialmente a las siguientes presentaciones a congresos.

Congresos Internacionales:

Bustamante, A., Reyes, J., Obando, P., Guerra, D., Wilson, C.A.M, Bustamante, C, Baez, M. **The energetic cost of knotting a polypeptide chain in its unfolded state determined by optical tweezers.** XII PABMB Congress; XLIX Reunión Anual de la Sociedad Argentina de Investigación Bioquímica y Biología Molecular; XXXVI Reunión Anual de la Sociedad de Bioquímica y Biología Molecular de Chile; 4th Meeting of the Latin-American Protein Society; LVI Reunión Anual de la Sociedad de Biología de Chile. 9 -14 de Noviembre de 2013. Puerto Varas, Chile.

Congresos Nacionales:

Bustamante, A., Reyes, J., Obando, P., Guerra, D., Wilson, C.A.M, Bustamante, C, Baez, M. **Knotted and untied topologies of single-chain ARC repressor characterized by optical trap force spectroscopy.** XXXV Reunión anual de la Sociedad de Bioquímica y Biología Molecular de Chile. 2 -5 de Octubre de 2012. Puerto Varas, Chile.

Financiamiento

Esta tesis fue desarrollada en el Laboratorio de Bioquímica de la Facultad de Ciencias Químicas y Farmacéuticas de la Universidad de Chile (Santiago de Chile) y en el Laboratorio de Visualización y Manipulación de Moléculas Individuales de la Facultad de Ciencias y Filosofía de la Universidad Peruana Cayetano Heredia (Lima – Perú). Contó con el financiamiento del proyecto FONDECYT n°11110534 titulado “Folding mechanism of knotted proteins: insights from single molecule and classic biophysical studies” y de la Beca de CONICYT para realización de Magíster en Chile folio n°22121199.

Índice general

Agradecimientos	i
Presentaciones a congresos	iii
Congresos Internacionales	iii
Congresos Nacionales	iii
Financiamiento	iv
Índice general	v
Índice de figuras y tablas	vii
Abreviaturas	viii
Abstract	ix
Resumen	xi
Introducción	1
1. <i>Topologías anudadas y su efecto en el mecanismo de plegamiento de proteínas</i>	2
2. <i>Superfamilia de factores de transcripción RHH como modelo de estudio de proteínas anudadas</i> ..	5
3. <i>Estudios de plegamiento a nivel de moléculas individuales mediante pinzas ópticas</i>	8
Hipótesis	10
Objetivo general	10
Objetivos específicos	10
Materiales y métodos	11
1. Lista de reactivos	11
2. Métodos	12
2.1. <i>Expresión y purificación de proteínas</i>	12
2.2. <i>Síntesis de moléculas de ADN modificadas</i>	13
2.3. <i>Modificación química de proteínas con hebras de ADN</i>	14
2.4. <i>Purificación del complejo ADN-proteína-ADN</i>	15
2.5. <i>Electroforesis en gel de agarosa</i>	15
2.6. <i>Electroforesis en gel de poliacrilamida</i>	15
2.7. <i>Espectrometría de masas</i>	16
2.8. <i>Experimentos de pinzas ópticas</i>	16
2.9. <i>Procesamiento de los datos obtenidos en pinzas ópticas</i>	16
2.10. <i>Análisis de datos de los experimentos de estiramiento a velocidad constante</i>	17
2.11 <i>Calculo de parámetros termodinámicos</i>	18

2.12. Análisis de datos cinéticos.....	18
Resultados.....	20
<i>Purificación y modificación química de las proteínas mARC y pARC para su utilización en experimentos de pinzas ópticas.</i>	<i>20</i>
<i>Purificación de mARC, mARC_OT y pARC_OT.....</i>	<i>20</i>
<i>Síntesis de moléculas de ADN modificado.....</i>	<i>21</i>
<i>Modificación química de proteínas.</i>	<i>23</i>
<i>Caracterización del mecanismo de plegamiento de mARC y pARC mediante metodologías de moléculas individuales</i>	<i>25</i>
<i>Estiramiento de las proteínas a velocidad constante.</i>	<i>26</i>
<i>Cálculo del ΔG al equilibrio a partir de trabajo de desplegamiento/replegamiento</i>	<i>35</i>
<i>Caracterización cinética del desplegamiento de mARC</i>	<i>37</i>
Discusión	42
<i>mARC se encuentra en solución en dos conformaciones, una anudada y otra no-anudada</i>	<i>42</i>
<i>La conformación anudada de mARC es más estable que su conformación no-anudada</i>	<i>43</i>
<i>Los estudios cinéticos fueron realizados sólo con mARC no-anudada.....</i>	<i>46</i>
<i>Comparación entre el mecanismo de plegamiento de mARC y pARC</i>	<i>47</i>
Conclusiones	49
Referencias.....	50

Índice de figuras y tablas

Índice de figuras

Figura 1. Relación entre el orden de contacto y la velocidad de plegamiento de proteínas pequeñas.	2
Figura 2. Ejemplo de topologías anudadas presentes en proteínas.	3
Figura 3. Esquema de equilibrio y diagrama de energía libre entre estados plegados anudados y no-anudados y estados desplegados anudados y no-anudados.	5
Figura 4. Arquitectura RHH ₂	7
Figura 5. Estructura y representación topológica de las diferentes versiones monoméricas del represor ARC.	7
Figura 6. Diseño experimental de un experimento de pinzas ópticas.	9
Figura 7. Purificación de mARC, mARC_OT y pARC_OT.	21
Figura 8. Síntesis y prueba funcional de moléculas de ADN modificado.	23
Figura 9. Espectros de masas para proteínas en presencia y ausencia de DTDP.	24
Figura 10. Purificación de complejo ADN-Proteína.	26
Figura 11. Curva de fuerza v/s extensión, obtenida de los experimentos de rampa de fuerza, para una molécula de mARC_OT y sobreestiramiento del ADN.	27
Figura 12. Análisis de datos de fuerza-extensión de mARC_OT.	29
Figura 13. Análisis de ANOVA de los eventos de desplegamiento y replegamiento de mARC_OT.	31
Figura 14. Histograma de extensión fraccional de mARC_OT.	33
Figura 15. Análisis de datos de fuerza-extensión de pARC_OT.	34
Figura 16. Histogramas de trabajo de desplegamiento y replegamiento para mARC_OT y pARC_OT.	36
Figura 17. Trayectoria de plegamiento para mARC_OT obtenida de experimentos de pinza de fuerza e histograma de tiempos de permanencia.	38
Figura 18. Chevrón de fuerzas para moléculas de mARC_OT.	39
Figura 19. Diagrama de energía libre v/s coordenada de reacción.	40
Figura 20. Ciclo termodinámico para moléculas de mARC_OT.	44

Índice de tablas

Tabla 1. Secuencia de partidores utilizados para la síntesis de hebras de ADN para manipulación de proteínas en pinzas ópticas.	13
Tabla 2. Protocolo de PCR para la síntesis de las moléculas de ADN.	14
Tabla 3. Rendimiento de pasos de síntesis y purificación de moléculas de ADN modificado	22
Tabla 4. Resultados de modificación de proteínas con DTDP	24
Tabla 5. Parámetros termodinámicos y cinéticos calculados a partir de experimentos de pinza de fuerza.	41

Abreviaturas

aa	Aminoácido
ADN	Ácido Desoxiribonucleico
ANOVA	Análisis de la Varianza
BIO	Biotina
BSA	Seroalbúmina Bovina
DIG	Digoxigenina
dNTP	Desoxiribonucleótidos
DO	Densidad óptica
DTDP	2,2'-Ditiodipiridina
DTT	Ditiotreitol
EDTA	Ácido etilendiaminotetraacético
GdnHCl	Cloruro de Guanidinio
IPTG	Isopropil- β -D-1-tiogalactopiranosido
MALDI-TOF	<i>Matrix Assited Laser Desorption/Ionization – Time Of Flight</i>
OC	Orden de contacto
OT	<i>Optical Tweezers</i>
pb	Pares de bases
PCR	Reacción en cadena de la polimerasa
PDB	<i>Protein Data Bank</i>
RHH	<i>Ribon Helix Helix</i>
rpm	Revoluciones por minuto
SCOP	<i>Structural Classification of Proteins</i>
SDS	Dodecil sulfato de sodio
WLC	<i>Wolrm Like Chain model</i>
ΔG	Cambio de energía libre de Gibbs

Abstract

“Characterization of the folding mechanism of a monomeric version of the ARC repressor, a RHH superfamily member with knotted topology, by light trapping force spectroscopy”

The topology of a protein, defined as the tridimensional path of the polypeptide chain in space, restricts the moves that the chain can perform for searching its native state. For this reason this property defines the folding mechanism of a protein. However, in recent years has been discovered some proteins that seek their native states through complex and intricate movements. Such is the case of knotted proteins which topology challenge current theories of protein folding and have opened new questions about protein folding mechanisms.

The first knotted structure was identified in 2000, since then many others have been discovered reaching the 2% of the structures in the Protein data bank (PDB). The low relative abundance of the knotted structures contrast with simulation of polymer models showing that knotted structures are easy to exist in those systems. This suggests that proteins somehow have avoided knotted topologies.

Another interesting thing to investigate about knotted proteins is the function that may have the knotted topology on this proteins and how that structures may fold. It has been determined that some knotted motifs are conserved in families and superfamilies despite of the low sequence identity. These results suggest that knotted motifs could provide some comparative advantage with respect to normal proteins. In this regard some theoretical and experimental evidence show that knotted topologies could enhance the stability of proteins. Nevertheless, none of these studies explain how this topologies does it.

In this work we seek to elucidate the energetic consequence that brings the knotting of a polypeptide chain to its folding mechanism. For this, we compare the folding process of a knotted protein with their unknotted counterpart. In terms of experimental approach this problem impose two challenges, the first one is that the knot formation should don't disturb the structure of the protein. The second one is that we must control de topology of the unfolded state in order to ensure the homogeneity of the thermodynamic species under experimentation. The first challenge was archived using the mARC protein from the RHH superfamily of transcription factors. This protein is a monomeric version of the ARC repressor from the p22 phage and it was constructed connecting the C terminus of the first subunit to the

N terminus of the next one using a 15 residue glycine-rich polypeptide linker. Structural models mARC shows that the linker forms a flexible loop that doesn't form contacts with hydrophobic core of the protein. Moreover, homology models indicate that the loop can move around the protein generating two conformations: a knotted and an unknotted one.

The second challenge was addressed by characterizing the folding mechanism of mARC using optical tweezers. This approximation allows us to pull the protein from its C and N-terminal ends to control the topology of the unfolded state and thus to compare the energy differences between the knotted unfolded state and the unknotted unfolded state of mARC.

In the other hand, by a rearrangement of the secondary structure of mARC it is possible to generate a non-cyclic permutant, pARC. This protein is reported to have the same architecture and folding stability of mARC but, its topology prevents knotting. So, pARC was used as a topological control of the effect of knotting the mARC protein.

We found that mARC can be captured in two conformations characterized by different contour length (L_0 , total extension of a polymer chain). One group of molecules has a L_0 of 43 ± 7 nm and corresponds to the expected for the unknotted conformation of mARC. The other group of molecules shows a L_0 of 37 ± 7 nm and corresponds to the expected for the knotted conformation. On the contrary pARC only shows one type of molecules characterized by a L_0 of 22 ± 7 nm. All together these results support the hypothesis that mARC is present in two conformations, one knotted and other unknotted. Moreover, applying Crook's theorem to the irreversible work associated to the unfolding and refolding transitions of mARC and pARC allowed to determine the free energy (ΔG) at equilibrium required to the unfold the knotted ($\Delta G_1 = 21,8 \pm 0,4$ kcal/mol), unknotted ($\Delta G_2 = 12,3 \pm 0,5$ kcal/mol) conformations of mARC and pARC ($\Delta G_3 = 7 \pm 0,4$ kcal/mol). These results indicate that knotted mARC is more stable than their unknotted counterpart by $8,9 \pm 0,9$ kcal/mol mainly due to the energetic cost to tie the unfolded state of the protein. This value corresponds to the energetic cost of knotting a polypeptide chain of 120 residues in its unfolded state and explains the source of the stabilizing effect of the knotted topology in mARC.

For the kinetic characterization we performed constant force experiments. It was only possible to study two molecules of mARC. Despite of the differences between the kinetic parameters calculated for them, both corresponds to the unknotted conformation of the protein.

Resumen

La topología de una proteína, definida como el camino tridimensional que sigue la cadena polipeptídica en el espacio, impone restricciones naturales al conjunto de movimientos que esta puede realizar para alcanzar su estado nativo. Por esta razón esta propiedad define el mecanismo de plegamiento de una proteína. Sin embargo, en los últimos años se han descrito proteínas que obligan a la cadena polipeptídica a buscar su camino de plegamiento a través de intrincados y complejos movimientos, es el caso de las proteínas anudadas. Estas proteínas suponen un desafío para las teorías de plegamiento actuales y han abierto nuevas interrogantes respecto al mecanismo de plegamiento de las proteínas.

La primera proteína anudada fue identificada en el año 2000, desde entonces se han descrito muchas otras llegando a representar del 2% del *Protein data bank* (PDB). Esta baja abundancia relativa de proteínas anudadas contrasta con lo observado en simulaciones de modelos de polímeros que indican que la probabilidad de formar este tipo de estructuras es alta. Esto hace pensar que las proteínas evitan la formación de nudos. No obstante no se conoce la tendencia natural de una proteína a anudarse.

Otra interrogante que nace del estudio de proteínas anudadas tiene relación con la función que cumplen estas estructuras en las proteínas que las presentan y cómo estas proteínas logran plegarse. Se ha demostrado que algunos motivos anudados se conservan en distintas familias y superfamilias a pesar de su baja identidad de secuencia. Esta evidencia sugiere que dichos motivos podrían aportar alguna ventaja comparativa a las proteínas que los presentan. En este aspecto, estudios teóricos y evidencias experimentales indican que la topología anudada podría aumentar la estabilidad de las proteínas. Sin embargo, ninguno de dichos estudios explica cómo la topología anudada confiere estabilidad.

En este trabajo tratamos de determinar las consecuencias que tiene sobre el mecanismo de plegamiento de una proteína el obligar a su cadena polipeptídica a anudarse. Para ello se comparó el proceso de plegamiento de una proteína anuda con la de su contraparte no-anudada. Experimentalmente esto sugiere dos desafíos importantes: primero, es necesario que la formación del nudo no altere significativamente la estructura de la proteína. Segundo, es indispensable poder controlar la topología del estado desplegado para asegurar la homogeneidad de los estados termodinámicos estudiados. El primer desafío fue abordado utilizando como modelo de estudio la proteína mARC de la superfamilia de factores de

transcripción RHH. Esta es una versión monomérica del represor ARC del fago p22 generada al unir el extremo carboxilo de una subunidad con el extremo amino de la siguiente utilizando un péptido conector de 15 aminoácidos rico en glicina. Modelos por homología de mARC muestran que el péptido conector puede formar un lazo flexible que no genera contactos con el centro hidrofóbico de la proteína. De esta manera, dependiendo de la posición de dicho lazo se pueden generar dos conformaciones de mARC, una conformación anudada y otra no-anudada.

El segundo desafío experimental fue abordado caracterizando el mecanismo de plegamiento de mARC utilizando pinzas ópticas. Esta aproximación permite estirar la proteína desde sus extremos y aplicar una fuerza sobre ella, de esta forma se puede controlar la topología del estado desplegado y con ello poder comparar la diferencia de energía libre entre el estado desplegado anudado y no-anudado de mARC.

Por otra parte, mediante un rearrreglo de los elementos de estructura secundaria de mARC puede generarse una permutante no circular de ella, pARC. Esta proteína presenta la misma arquitectura y estabilidad al equilibrio que mARC pero su topología le impide formar nudos. Así, esta permutante fue utilizada como control del efecto topológico del anudamiento de mARC.

Los resultados obtenidos al estirar mARC a velocidad constante permitieron identificar dos tipos de moléculas con largos de contorno (L_0 largo de una molécula completamente estirada) diferentes. Un grupo de moléculas presentó un L_0 de 37 ± 7 nm, que coincide con lo esperado para la conformación anudada de mARC. El otro grupo mostró un L_0 de 43 ± 7 nm, esperado para la conformación no-anudada de esta proteína. Al contrario, pARC sólo mostró un tipo de moléculas con un L_0 de 22 ± 7 nm. De estos resultados se concluye que mARC está presente en dos conformaciones, una anudada y otra no anudada.

De los experimentos de estiramiento a velocidad constante se obtuvo el trabajo irreversible, realizado fuera del equilibrio, asociado a las transiciones de desplegamiento y replegamiento de mARC y pARC. Utilizando el teorema de Crooks fue posible calcular, a partir de estos datos, el ΔG en condición de equilibrio para el desplegamiento de las conformaciones anudada ($\Delta G_1 = 21,8 \pm 0,4$ kcal/mol) y no-anudada ($\Delta G_2 = 12,3 \pm 0,5$ kcal/mol) de mARC así como también de pARC ($\Delta G_3 = 7 \pm 0,4$ kcal/mol). Estos resultados indican que mARC anudado es más estable que mARC no-anudado por $8,9 \pm 0,9$ kcal/mol, este valor corresponde al costo energético de anudar una cadena polipeptídica de 120 residuos en su estado desplegado. Esta es la primera vez que es posible determinar experimentalmente este valor. Además, esta

diferencia de energía explica, cómo es que la topología anudada confiere estabilidad en esta proteína.

Para caracterizar cinéticamente las proteínas se realizaron experimentos a fuerza constante. Sólo fue posible estudiar dos moléculas de mARC las cuales, a pesar de las diferencias que muestran en sus parámetros cinéticos calculados, corresponden a la conformación no-anudada de la proteína. No fue posible realizar estudios cinéticos con pARC.

Introducción

La topología de una proteína, definida como el camino tridimensional que sigue una cadena polipeptídica en el espacio, es el factor principal que determina el mecanismo de plegamiento de las proteínas (Onuchic & Wolynes 2004) ya que impone restricciones naturales al conjunto de movimientos que una cadena polipeptídica puede realizar para alcanzar su estado nativo, debido a la imposibilidad de entrecruzar la cadena principal. La importancia de la topología se apoya en evidencias experimentales que han demostrado que las proteínas con topologías más complejas, es decir, estructuras estabilizadas por residuos que se encuentran distantes en su secuencia (Figura 1A), presentan cinéticas de plegamiento más lentas que proteínas con topologías más simples (Figura 1B y C) (Baker 2000, Plaxco y cols 1998). Adicionalmente estudios teóricos que consideran casi exclusivamente la topología para describir el proceso de plegamiento, tales como los modelos simplificados de proteínas (Dill y cols 1995, Fersht 1997, Kouza y cols 2006, Onuchic y cols 1997) y simulaciones moleculares con potenciales basados en la estructura nativa ($G\ddot{o}$), reproducen adecuadamente observaciones experimentales de una amplia gama de proteínas con diferentes topologías (Hills & Brooks 2009, Koga & Takada 2001, Paci y cols 2002). No obstante, desde hace algunos años se han reportado estructuras con topologías anudadas que desafían estos preceptos y han abierto nuevas interrogantes respecto al mecanismo de plegamiento de las proteínas.

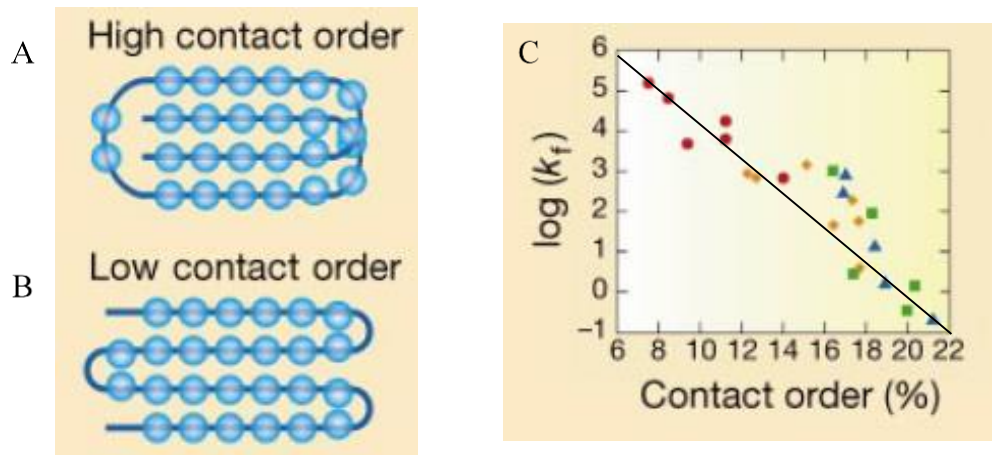


Figura 1. Relación entre el orden de contacto y la velocidad de plegamiento de proteínas pequeñas. El orden de contacto es una medida de la complejidad topológica de una proteína. **A.** La topología de una cadena polipeptídica es más compleja cuando la estructura de una proteína es estabilizada por contactos entre residuos distantes en la secuencia, esto lleva a un alto orden de contacto. **B** En cambio, la topología de una proteína es sencilla si dos residuos cercanos en la secuencia forman un contacto en el estado nativo, esto implica un menor orden de contacto. **C.** Gráfico de $\log k_f$ v/s OC para proteínas pequeñas que se pliegan en dos estados. Se observa una relación lineal inversa lo cual indica que a menor complejidad topológica mayor es la velocidad de plegamiento. Adaptado de Baker 2000.

1. Topologías anudadas y su efecto en el mecanismo de plegamiento de proteínas. La idea de la existencia de nudos en las proteínas fue abordada por primera vez por Masfield en 1994. El concluyó que las estructuras anudadas eran muy complicadas debido a que su topología obliga a la cadena polipeptídica a buscar su camino de plegamiento a través de movimientos muy complejos lo cual debería impedir un plegamiento eficiente y descartó por completo su existencia (Mansfield 1994). Sin embargo, Taylor en 2000 identificó un nudo complejo en la estructura de la enzima ácido hidroxiaacético isomeroeductasa de plantas (Figura 2A, Taylor 2000). Desde entonces se han descubierto sostenidamente más proteínas que presentan topologías anudadas (Figura 2), por lo que los motivos anudados han sido reconocidos como motivos estructurales significativos llegando a representar el 2% de las estructuras depositadas en el PDB (Khatib y cols 2006, Kolesov y cols 2007, Lua & Grosberg 2006, Virnau y cols 2006). La baja abundancia de estructuras anudadas en proteínas contrasta con la tendencia natural de formar nudos observado en simulaciones realizadas con modelos de homopolímeros globulares (Lua y cols 2004, Tubiana y cols 2013, Virnau y cols 2005) las cuales indican que la probabilidad de encontrar estructuras globulares anudadas es de alrededor del 80% (para polímeros de largo de cadena de 400 monómeros en adelante). Experimentalmente se ha abordado un modelo macroscópico de cadenas poliméricas utilizando cuerdas de material

flexible agitadas en ambientes confinados (Raymer & Smith 2007). En este estudio la cuerda en sí no puede generar interacciones de ningún tipo consigo misma por lo que la información obtenida da cuenta sólo de la posibilidad intrínseca de cualquier cadena (cuerda) flexible a formar nudos por efectos del ambiente (agitación y confinamiento). Los resultados arrojan una probabilidad máxima de formar nudos de un 50%, la cual se alcanza con un largo de cuerda de 1,5 m en adelante (largo comparable con polímeros grandes, como los estudiados en las simulaciones antes mencionadas). En conjunto los antecedentes mostrados sugieren que las proteínas actuales evitan de algún modo la formación de nudos debido a que este tipo de estructuras no son difíciles de formar en modelos de polímero o cuerdas. Sin embargo, aún no se conoce cuál es la tendencia intrínseca de una cadena polipeptídica a formar un nudo.

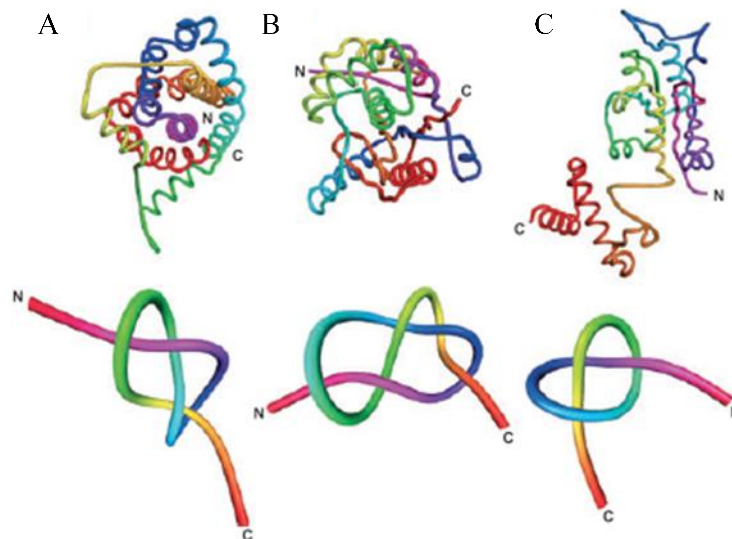


Figura 2. Ejemplo de topologías anudadas presentes en proteínas. Las proteínas anudadas están presentes en todos los dominios de la vida y pueden tener diferente complejidad topológica. A. Enzima ácido hidroxiaacético isomeroreductasa de plantas presenta un complejo nudo 4_1 o en “forma de ocho”. B. Ubiquitin-hidrolasa humana, contiene uno de los nudos más complejos reportado (5_2), con 5 cruces. C. tRNA-metiltransferasa de *H. influenzae* presenta un nudo 3_1 o “en forma de trébol” el nudo más común y menos complejo que pueden presentar estas proteínas. Extraído de Mallam 2009

Otra pregunta que ha surgido a través del estudio de este tipo de proteínas tiene relación con la función que cumplen las topologías anudadas y como estas estructuras logran plegarse. Se ha demostrado que algunos motivos anudados se conservan en familias y superfamilias a pesar de su baja identidad de secuencia (Sulkowska y cols 2012) por lo que algunos autores proponen que dichos motivos anudados podrían aportar alguna ventaja comparativa a las proteínas que los presentan. En este sentido, se ha propuesto que el anudamiento puede aumentar la estabilidad de una proteína (Dzubiella 2009, Sułkowska y cols 2008) e

incrementar su estabilidad cinética disminuyendo la velocidad de desplegamiento (Soler & Faísca 2013). En este aspecto, un estudio interesante es el realizado por King y cols (2010) en el cual se compararon las propiedades de plegamiento de proteínas similares diferenciadas sólo por la presencia o ausencia de la topología anudada. En dicho estudio se observó que la proteína anudada presenta una mayor estabilidad comparada con el control no anudado (King y cols 2010). Sin embargo, en este trabajo la metodología utilizada no permite controlar la topología del estado desplegado por lo que no se tiene certeza de las especies termodinámicas que se comparan. Por esta razón no es posible, a nuestro juicio, atribuir de manera concluyente el aumento de estabilidad observado sólo a la presencia de la topología anudada. Además, este estudio no explica el motivo por el cual se produce un incremento de la estabilidad en estas proteínas.

En este trabajo intentamos conocer cuáles son las consecuencias termodinámicas y cinéticas para el mecanismo de plegamiento de una proteína cuando se obliga a que su cadena polipeptídica se anude. Esto requiere esencialmente poder comparar la diferencia de energía libre entre las conformaciones nativas y desplegadas del polipéptido anudado con la diferencia de energía libre entre las conformaciones nativas y desplegadas de la cadena polipeptídica no-anudada. De esta manera, el $\Delta\Delta G$ entre ambos equilibrios nos permitiría conocer el costo (o ganancia) energético que resulta al restringir la cadena polipeptídica en una topología anudada para el proceso de plegamiento. La figura 3 esquematiza esta situación.

Desde el punto de vista experimental este esquema trae dos desafíos importantes. Primero, se requiere que la formación del nudo no altere de manera significativa la estructura de la proteína para que la diferencia de energía libre se deba sólo al efecto de la topología y no a otros factores, como la formación de interacciones entre la zona del nudo y el resto de la estructura de la proteína. Segundo, se debe controlar la topología del estado desplegado de la proteína para asegurar la homogeneidad de los estados termodinámicos bajo experimentación. El primer requerimiento fue abordado utilizando como modelo de estudio la proteína mARC (Robinson & Sauer 1996b) de la familia de factores de transcripción RHH. El segundo requerimiento se abordó mediante la caracterización de su mecanismo de plegamiento utilizando pinzas ópticas. La ventaja de esta metodología es que permite estudiar el mecanismo de plegamiento de una proteína estirándola desde puntos específicos de su estructura y extraer parámetros termodinámicos sobre su estabilidad y también información sobre las barreras cinéticas que limitan este proceso (Bustamante y cols 2005, Cecconi y cols

2005, Shank y cols 2010). Como se describe a continuación, el estudio de este modelo con la metodología de pinzas ópticas nos permitió abordar experimentalmente el esquema propuesto en la figura 3

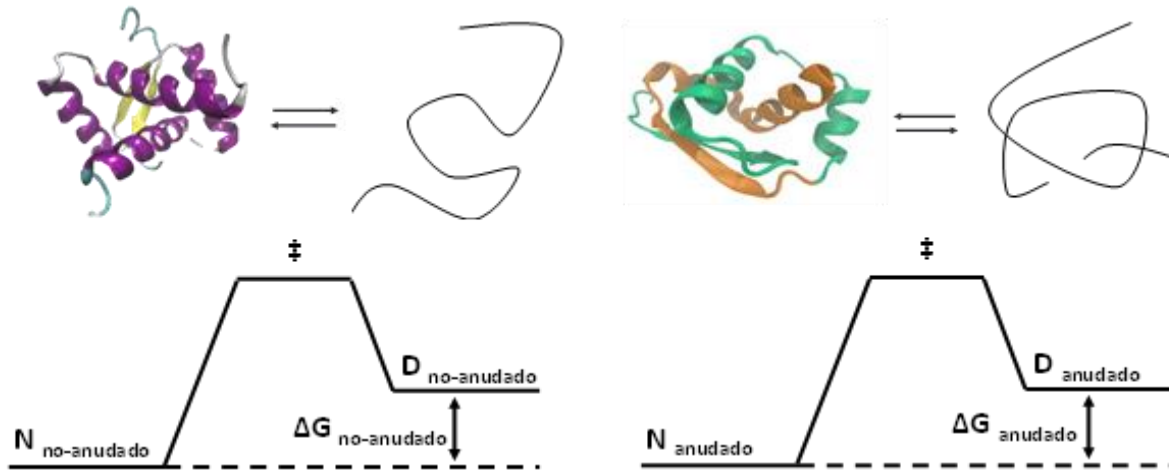


Figura 3. Esquema de equilibrio y diagrama de energía libre entre estados plegados anudados y no-anudados y estados desplegados anudados y no-anudados. El modelo a estudiar debe permitir comparar los equilibrios entre estados nativos anudados y no-anudados y sus estados desplegados respectivos. La diferencia de energía libre entre ambos equilibrios corresponde al costo (o ganancia) energética de anudar una cadena polipeptídica.

2. Superfamilia de factores de transcripción RHH como modelo de estudio de proteínas anudadas. La familia de factores de transcripción RHH se caracteriza por una arquitectura generada por dos motivos de estructura secundaria denominados RHH (β lámina - α hélice - α hélice) que se entrelazan para formar un arreglo de 4 hélices empacadas en contra de una pequeña hoja β (simetría RHH₂, Figura 4) (Schreiter & Drennan 2007). En general los miembros de esta familia han sido descritos como homodímeros cuya arquitectura funcional se genera por la asociación de dos motivos RHH aportados por dos subunidades idénticas. Miembros homodiméricos de esta familia son el represor ARC y Mnt del bacteriófago p22 (SCOP, Murzin y cols 1995). Dentro de esta familia, el represor ARC es la proteína mejor caracterizada en términos de su mecanismo de plegamiento, el cual ha sido descrito mediante un modelo de dos estados (Waldburger & Jonsson 1996; Brown & Sauer 1999; C. R. Robinson & Sauer 2000). De manera interesante, el homodímero de ARC ha sido convertido en un monómero al unir el extremo carboxilo de una subunidad con el extremo amino de la siguiente utilizando un péptido conector de 15 aminoácidos rico en glicina (Robinson & Sauer 1996b). La unión de los dos motivos RHH en una cadena polipeptídica resulta en una proteína monomérica denominada mARC (del inglés monomeric ARC). Estudios cinéticos y de

estabilidad de mARC, realizados en condiciones de equilibrio con agentes caotrópicos, indican que el péptido conector no afecta el mecanismo de plegamiento respecto al dímero silvestre (Robinson & Sauer 1996a). Sin embargo, modelos de la estructura de mARC generados en nuestro laboratorio muestran que dependiendo de la posición del lazo flexible formado por el péptido conector de 15 aminoácidos es posible generar dos conformaciones para la proteína mARC, una conformación anudada (Figura 5A) y una conformación no-anudada (Figura 5B). Ambas conformaciones se logran sin alterar el centro hidrofóbico de la proteína. Debido a que el péptido conector utilizado para unir ambos motivos RHH es artificial y rico en glicina se postula que el lazo es flexible y no establece contactos específicos con el resto de la estructura de la proteína por lo que puede dar origen a ambas conformaciones.

Por otra parte, mediante un re-arreglo de los elementos de estructura secundaria de mARC es posible generar otra versión monomérica del dímero ARC. Esta permutante, pARC, no presenta el péptido conector de 15 aminoácidos rico en glicinas y por lo tanto no puede generar topologías anudadas (Figura 5C). Sin embargo, pARC mantiene la arquitectura del dímero ARC y presenta la misma estabilidad al equilibrio y cinética de plegamiento que mARC (Tabtiang y cols 2005). Dichas características permiten proponer a pARC como control del efecto topológico del anudamiento en mARC.

Para comprobar la hipótesis de que mARC se encuentra en dos conformaciones, se requiere capturar cada conformación de la proteína y realizar estudios de desplegamiento controlando la topología del estado desplegado. Esto es posible utilizando técnicas de manipulación de moléculas individuales como pinzas ópticas.

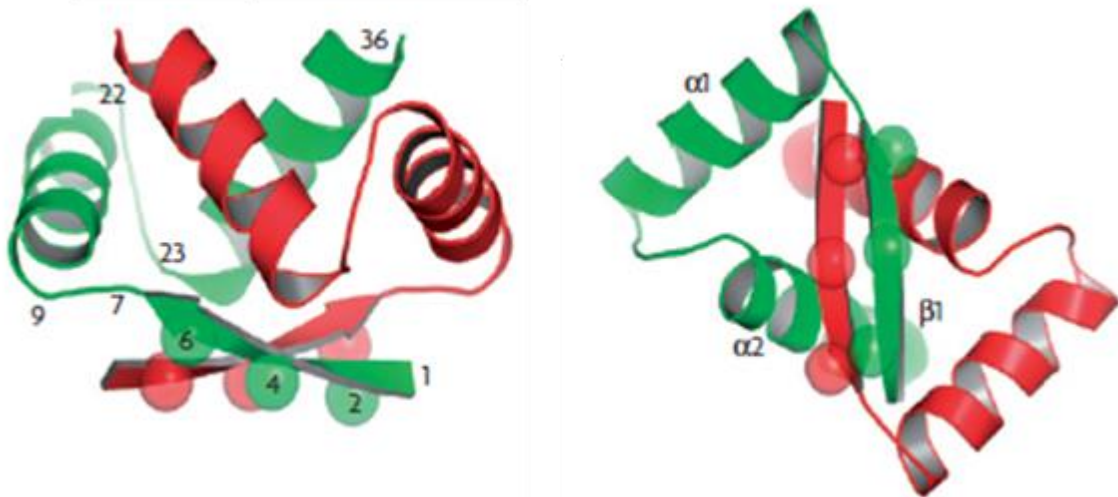


Figura 4. Arquitectura RHH2. Simetría RHH_2 , la arquitectura de los factores de transcripción RHH es generada por la asociación de dos motivos RHH y consiste en cuatro hélices α empaçadas contra una pequeña hoja β antiparalela. Adaptado de Schreiter & Drennan 2007.

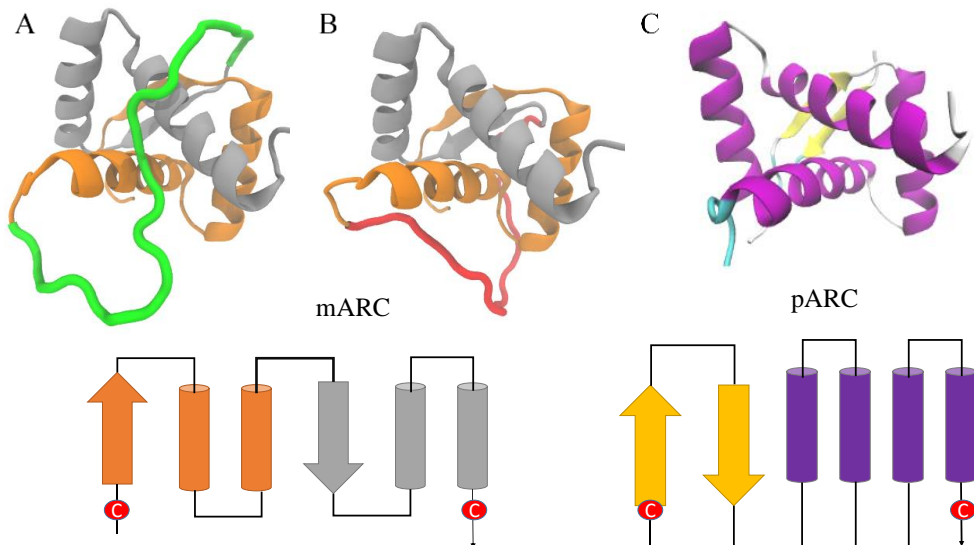


Figura 5. Estructura y representación topológica de las diferentes versiones monoméricas del represor ARC. El modelo de la estructura generado para mARC predice la existencia de dos conformaciones dependiendo de la posición del lazo formado por el conector flexible de 15 aminoácidos utilizado para generar el monómero. Si el conector se presenta en la posición indicada en verde (A) la proteína forma un nudo al ser estirada desde sus extremos, por el contrario, cuando el conector se encuentra en la posición indicada en rojo (B) la proteína no forma el nudo. En el caso de pARC (C) los elementos de estructura secundaria se han reordenado de tal forma de conservar la arquitectura pero al cambiar la topología la proteína no puede formar nudos.

3. Estudios de plegamiento a nivel de moléculas individuales mediante pinzas ópticas. Las ventajas que ofrecen los estudios de molécula individual por sobre los experimentos clásicos (en múltiplo) son variadas, entre ellas podemos mencionar: caracterización individual de poblaciones que pueden no ser homogéneas (como el caso de la proteína en estudio en este trabajo), los estudios de plegamiento se realizan utilizando una propiedad física de la molécula como coordenada de reacción (distancia extremo a extremo, por ejemplo), utilizan poca muestra y permiten controlar la topología del estado desplegado de la proteína en estudio (Bustamante 2008). Además, permiten la obtención de parámetros termodinámicos al equilibrio a partir de datos obtenidos fuera del equilibrio mediante la utilización de teoremas de fluctuaciones como el de Crooks (Crooks 2000). Así también, estas metodologías permiten realizar estudios cinéticos en condiciones de equilibrio siguiendo el cambio de extensión de la molécula ante una fuerza externa constante en tiempo real (Liphardt y cols 2001).

En los experimentos de pinzas ópticas la manipulación de las moléculas se realiza a través de hebras de ADN de doble cadena que se unen a la proteína mediante puente disulfuro en puntos específicos de la cadena polipeptídica (Figura 6, Cecconi y cols 2008). De esta manera es posible estirar la proteína desde puntos específicos de la estructura y así elegir el eje de aplicación de la fuerza. En el caso de proteínas anudadas la única forma de caracterizar el efecto del nudo es tirando desde los extremos amino y carboxilo ya que sólo así la proteína permanece anudada en su estado desplegado (Taylor 2000). De acuerdo a la hipótesis de las dos conformaciones de mARC, la unión de las hebras de ADN en los extremos de la cadena polipeptídica podría capturar la proteína en su conformación anudada (Figura 6B derecha) y no-anudada (Figura 6B izquierda). El estiramiento de ambas conformaciones permitiría controlar la topología de su estado desplegado y por ende acceder a los parámetros termodinámicos asociados a los equilibrios de la figura 3. Vale decir, partiendo desde una conformación nativa anudada hacia una conformación desplegada anudada (Figura 3 derecha) y desde una conformación nativa no-anudada hacia una conformación desplegada no-anudada (Figura 3 izquierda). Dichos datos harían posible calcular la diferencia de energía entre los estados desplegados anudado y no-anudado ($\Delta\Delta G$) y con ello determinar el costo (o ganancia) energético de anudamiento de la proteína mARC.

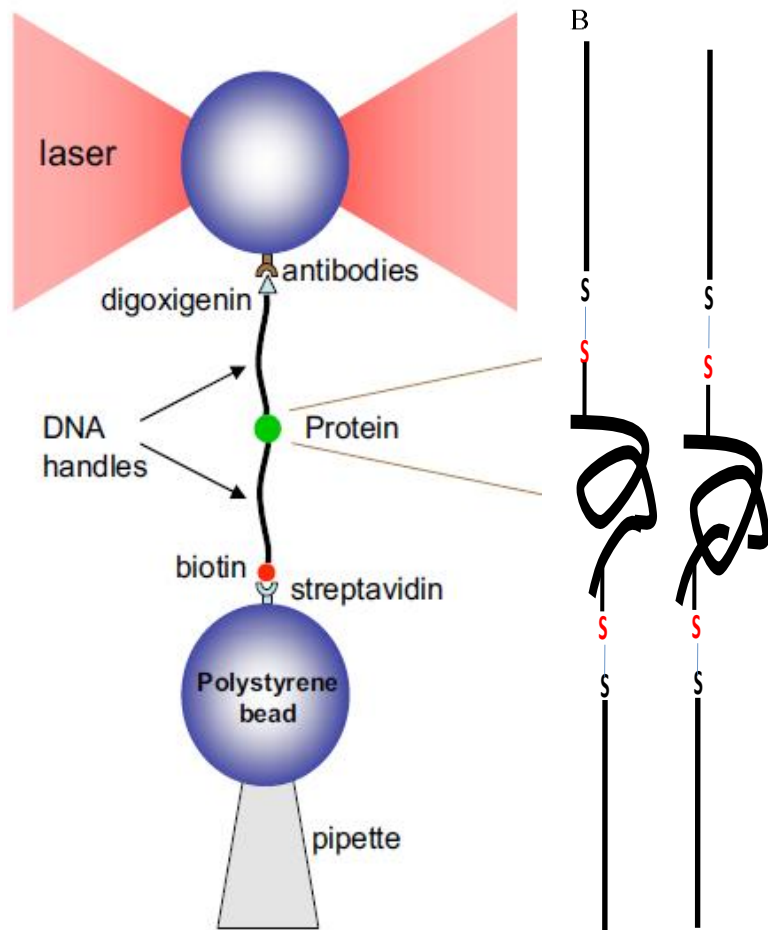


Figura 6. Diseño experimental de un experimento de pinzas ópticas. A. Diagrama general de un experimento de pinzas ópticas, la proteína modificada es mantenida entre dos microesferas de poliestireno, una de estas es mantenida fija sobre la punta de una pipeta microscópica, por succión. La otra es mantenida en la trampa óptica y puede ser movida para manipular la molécula. B. La modificación con ADN permite la captura de cada una de las conformaciones esperadas para mARC, a la derecha se muestra la captura de la conformación anudada y a la izquierda la captura de la conformación no-anudada. Adaptado de Cecconi y cols 2008.

Hipótesis

El estado desplegado anudado de la proteína mARC es más inestable que su estado desplegado no-anudado.

Objetivo general

Determinar la diferencia de estabilidad entre el estado desplegado anudado y no-anudado de la proteína mARC.

Objetivos específicos

- 1) Purificar y modificar químicamente las proteínas mARC y pARC para su utilización en experimentos de pinzas ópticas.
- 2) Comparar el estiramiento mecánico de mARC y pARC para determinar si mARC puede presentarse en diferentes conformaciones y comparar sus parámetros termodinámicos y cinéticos.

Materiales y métodos

1. Lista de reactivos

n°	Reactivo	Marca	Número de catálogo
1.	2-Propanol	Merck	1.01040.4000
2.	Ácido acético glacial	Merck	1.00063.2500
3.	Ácido Clorhídrico (32%)	Merck	1.00317.2500
4.	Acrilamida	Amresco	341
5.	Agarosa	Bioline	BIO-41026
6.	Amortiguador de carga de ADN 6X	Thermo Scientific	R061
7.	Ampicilina	USBiological	A2260
8.	Bis-acrilamida	Amresco	1725
9.	Bis-Tris	Calbiochem	391335
10.	Bisulfito de sodio	Merck	1066570500
11.	Cloruro de Guanidinio	Winkler	BM-0845
12.	Cloruro de Potasio	Merck	1.04936.1000
13.	Dihidrogenofosfato de potasio	Merck	1.04873.0250
14.	dNTP	Bioline	BIO-39025
15.	DreamTaq DNA Polymerase	Thermo Scientific	EP0702
16.	DTDP	Sigma	43791-1G
17.	DTT	Calbiochem	3860
18.	EDTA	Merck	1.08418.1000
19.	Estándar de peso molecular de ADN (1 kb)	Thermo Scientific	SM0313
20.	Estándar de peso molecular de proteínas, preteñido.	Thermo Scientific	26616
21.	Etanol (100%)	Merck	1.00983.2500
22.	GelRed	Biotium	BI.41002
23.	Glicerol (85%)	Merck	1.04094.1000
24.	Glicina	USBiological	C13042598
25.	Hidrogenofosfato de sodio	Merck	1.05104.1000
26.	Imidazol	USBiological	C9101551
27.	IPTG	Bioline	BIO-37036
28.	Medio de cultivo LB Broth	MO BIO	12106-1
29.	MES 1-Hidrato	Winkler	712030
30.	MOPS	Calbiochem	6310
31.	Persulfato de amonio	Amresco	0486-25G
32.	Reactivo de Bradford	Bio-Rad	500-0006
33.	SDS	Merck	1.13760.0100
34.	Tris base	Calbiochem	9210

35.	Triton X-100	Sigma	X100-1L
36.	Tween-20	USBiological	P4379

2. Métodos

2.1. Expresión y purificación de proteínas. Las secuencias de proteínas se diseñaron con un marcador de afinidad de hexa-histidina (His x6) para facilitar su purificación. Además, en el caso de las mutantes utilizadas en experimentos de pinzas ópticas (llamadas OT, por Optical Tweezers) se incluyeron residuos de cisteína expuestos a solvente en los extremos de la cadena polipeptídica para su modificación con hebras de ADN (Cecconi y cols 2008). En el caso de mARC_OT se mutaron los residuos G3C y A122C (Figura 5A y B, panel inferior), y en pARC_OT los residuos mutados fueron G3C y G105C (Figura 5C, panel inferior).

Las proteínas se obtuvieron a partir de cultivos de bacterias *E.coli* BL21 (DE3) hechos en medio LB Broth suplementado con ampicilina 100 µg/ml (como marcador de selección), a una temperatura de 37 °C y agitación constante de 220 rpm. Su producción se indujo durante una hora utilizando 0,5 mM IPTG a una DO_{600nm}=1,0 en el caso de mARC y de mARC_OT; y a DO_{600nm}= 0,5 para el caso de pARC_OT. Finalmente las bacterias se colectaron por centrifugación a 5000 rpm durante 15 minutos y a 4 °C.

Las bacterias colectadas se lisaron utilizando el **amortiguador A** (10 mM Tris-HCL pH 8,0, 50 mM Fosfato pH 8,0, 40 mM Imidazol pH 8,0, 5M GdnHCl) durante 1 hora a 37 °C con agitación constante de 220 rpm. En seguida el lisado se centrifugó a 15000 x g durante 15 minutos a temperatura ambiente y el sobrenadante se inyectó en una columna de afinidad a Ni²⁺ (HisTrap FF Crude, GE Healthcare) previamente equilibrada en amortiguador A. Posterior a un lavado de 10 volúmenes con amortiguador A, las proteínas se eluyeron de la columna utilizando una solución de 0,2 M ácido acético y 5 M GdnHCl. La fracción más concentrada se incubó con 50 mM DTT durante 12 horas, a temperatura ambiente, y luego se inyectó en una columna de exclusión molecular SuperdexTM 75/300 GL (GE Healthcare), previamente equilibrada con **amortiguador B** (25 mM Tris-HCl pH 7,6, 0,2 M KCl, 0,2 mM EDTA). Se colectaron las especies con un volumen de exclusión de alrededor de 13 ml, las cuales corresponden a las proteínas monoméricas según la calibración de la columna.

2.2. Síntesis de moléculas de ADN modificadas. En un extremo 5' las moléculas de ADN presentan un grupo tiol (-SH) libre para formar un puente disulfuro con las cisteínas de la proteína y en el otro extremo 5' las moléculas de ADN presentan biotina o digoxigenina para la unión direccional a las microesferas mediante interacciones específicas antígeno-anticuerpo o biotina-estreptavidina (Cecconi y cols 2008).

Se sintetizaron hebras de ADN de doble cadena de 558 pb mediante PCR utilizando DreamTaq DNA Polymerase (Thermo Scientific) y el plásmido pGEMEX como templado. Para poder generar ambos tipos de moléculas de ADN fue necesaria la utilización de partidores modificados (Tabla 1). El volumen de reacción de PCR fue de 9,6 ml por cada tipo de molécula de ADN a sintetizar. En la tabla 2 se presenta el protocolo de PCR utilizado.

Tabla 1. Secuencia de partidores utilizados para la síntesis de hebras de ADN para manipulación de proteínas en pinzas ópticas.

Nombre	Secuencia
F_Bio	5'-Biotina-CAA-AAA-ACC-CCT-CAA-GAC-CC-3'
F_Dig	5'-Digoxigenina-AA-AAA-ACC-CCT-CAA-GAC-CC-3'
R_SH	5'-thioMC6-D-GCT-ACC-GTA-ATT-GAG-ACC-AC-3'

Tabla 2. Protocolo de PCR para la síntesis de las moléculas de ADN.

Mezcla de reacción de PCR		Protocolo de PCR			
Reactivos	Concentración final en el medio de reacción				
DTT	16 mM	Etapa	Número ciclos	Temperatura	Duración
Amortiguador DreamTaq	1X				
BSA	0,02 mg/ml	1	1	94°C	0:03:00
dNTP	0,2 mM				
Glicerol	2,5%	2	34	94°C	0:00:45
Partidor Bio/Dig	0,5 µM			55°C	0:00:40
Partidor SH	0,5 µM			72°C	0:01:30
Vector pGEMEX	0,08 ng/µl	3	1	72°C	0:05:00
DreamTaq	0,025 U/µl				
H ₂ O	c.s.p 9,6 ml	Al finalizar se mantiene a 4°C			

Los productos de PCR se purificaron utilizando el Plasmid Midi Kit (QIAGEN) según las indicaciones del fabricante, suplementando todos los amortiguadores con 3 mM DTT.

2.3. *Modificación química de proteínas con hebras de ADN.* Inmediatamente después de colectar las proteínas desde la columna de exclusión molecular, las proteínas purificadas se incubaron con un exceso de DTDP (50X respecto a la cantidad de proteína) durante dos horas a temperatura ambiente. La modificación de los residuos de cisteína de las proteínas con la sonda se determinó mediante un método espectrofotométrico. La reacción de DTDP con grupos tiol libres forma 2-tiopiridona la cual presenta un máximo de absorción a 343 nm permitiendo su seguimiento por cambios en la absorbancia a esa longitud de onda (Grassetti & Murray 1967).

El DTDP en exceso se retiró utilizando dos columnas Micro Bio-Spin (Bio-Rad) en serie equilibradas en **amortiguador C** (25 mM Tris-HCl pH 7,6, 0,2 M KCl, 0,1 mM EDTA, 0,001% Tween-20).

En paralelo el ADN se incubó durante dos horas con 10 mM DTT. El exceso de reductor se retiró utilizando dos columnas Micro Bio-Spin (Bio-Rad) en serie, equilibradas con amortiguador C.

Finalmente las proteínas se incubaron en presencia de ADN en una relación en masa de 8:1 (ADN:Proteína) 6 horas a temperatura ambiente y otras 12 horas a 4°C.

2.4. Purificación del complejo ADN-proteína-ADN. El complejo proteína-ADN se purificó utilizando resina Ni-NTA (QUIAGEN). En un tubo de microcentrífuga (de 1,5 ml) se depositaron 400 µl de resina, la cual se centrifugó a 5000 rpm durante 2 minutos para descartar la solución de almacenamiento. El ADN no unido a la proteína se retiró lavando la resina 8 veces mediante sedimentación (5000 rpm por 2 minutos) utilizando **amortiguador D** (25 mM Tris-HCl pH 7,6, 0,2 M KCl). La muestra se eluyó de la resina utilizando el **amortiguador E** (25 mM Tris-HCl pH 7,6, 0,2 M KCl, 1M Imidazol pH 7,6), agitando suavemente por 10 minutos y centrifugando a 15000 xg durante 3 minutos. El sobrenadante corresponde al complejo ADN-proteína-ADN purificado.

Posterior a su purificación, el complejo proteína-ADN se suplementó con 50% glicerol y 0,1 mM EDTA para su almacenamiento a -80°C hasta su utilización en experimentos de pinzas ópticas.

2.5. Electroforesis en gel de agarosa. La síntesis de las hebras de ADN mediante PCR se analizó utilizando electroforesis en gel de agarosa al 1%. La electroforesis se realizó en amortiguador TAE (40 mM Tris, 20 mM ácido acético y 1 mM EDTA, pH 8,0) a voltaje constante (100 volts). Los geles se tiñeron usando GelRed (Biotium) y se revelaron usando un transiluminador UV.

2.6. Electroforesis en gel de poliacrilamida. La purificación de proteínas se analizó por electroforesis en gel de poliacrilamida (Bis-Tris) al 10% en condiciones desnaturantes y reductoras en presencia de SDS. Los geles se cargaron con 10 µg de proteínas totales usando amortiguador de carga 4X (Laemmli 1970). La electroforesis se llevó a cabo usando el siguiente amortiguador de corrida: 50 mM MES, 50 mM Tris, 0,1% SDS, 1 mM EDTA pH 7,3 y 5 µM Na₂SO₃, a voltaje constante, primero a 80 volts para permitir el ingreso de las proteínas al gel separador y luego a 100 volts. Finalmente los geles se tiñeron con azul de Coomassie G-250 y se destiñeron usando una solución de ácido acético 10% v/v y etanol 30 % v/v.

Para el análisis de las proteínas modificadas con ADN se siguió el mismo protocolo anterior con las siguientes modificaciones: se utilizó el amortiguador de carga para ADN 6X (Thermo Scientific) sin reductor, se preparó amortiguador de corrida sin SDS y sin bisulfito. La electroforesis se realizó durante dos horas y media a voltaje constante (100volts). Los geles

se tiñeron con GelRed (Biotium) para detectar ADN y se revelaron utilizando un equipo G:BOX F3 (SYGENE).

2.7. Espectrometría de masas. Los experimentos de espectrometría de masas se realizaron en la unidad de espectrometría de masas del Centro de Estudios para el Desarrollo de la Química (CEPEDEQ) de la Facultad de Ciencias Químicas y Farmacéuticas de la Universidad de Chile, con la ayuda del Dr. Andrés Barriga. La técnica MALDI-TOF se utilizó para determinar los pesos moleculares de las diferentes proteínas y sus mutantes, en presencia y ausencia de DTDP.

2.8. Experimentos de pinzas ópticas. Los experimentos de pinzas ópticas se realizaron utilizando un instrumento de trampa simple llamado “miniTweezers” debido a su tamaño compacto. Corresponde a una trampa óptica analítica capaz de medir fuerzas en el orden los picoNewtons (pN), distancias en la escala de Åmströngs (Å) y eventos temporales del orden de los milisegundos (ms).

Se realizaron dos tipos de experimentos, estiramiento a velocidad constante (rampa de fuerza) y a fuerza constante (pinza de fuerza). Los experimentos de rampa de fuerza consistieron en tirar la proteína desde sus extremos a una velocidad constante alejando la esfera atrapada en la trampa óptica de aquella fija en la pipeta, la velocidad utilizada en todos los experimentos fue de 100 nm/s y los extremos de fuerza utilizados fueron de dos a treinta pN con un tiempo de replegamiento de cinco segundos. En los experimentos de pinza de fuerza se siguió el cambio de extensión de las moléculas estudiadas como función del tiempo manteniendo la fuerza constante durante cuatro minutos utilizando el modo automático del equipo, de esta manera se estudió cada molécula en un rango de fuerzas que fueron desde los 5,4 hasta los 6,8 pN aumentando en intervalos de 0,2 pN cada vez. Los experimentos se realizaron utilizando el amortiguador B.

2.9. Procesamiento de los datos obtenidos en pinzas ópticas. El procesamiento de los datos de los experimentos de estiramiento a velocidad constante se realizó en el programa “TweezerAnalysis” (Jesse Dill) que opera en MATLAB (MathWorks®). Los datos obtenidos de los experimentos hechos a fuerza constante se procesaron utilizando el programa anteriormente mencionado junto con el programa HaMMY (McKinney y cols, 2006) para

determinar las constantes cinéticas asociadas al desplegamiento y replegamiento de las proteínas estudiadas.

2.10. Análisis de datos de los experimentos de estiramiento a velocidad constante. Para los datos de estiramiento a velocidad constante se determinó la fuerza a la cual los eventos de desplegamiento y replegamiento ocurrieron. Además, se determinó la distancia involucrada en cada evento, para ello la curva se ajustó a una recta antes y después del evento y se determinó la distancia entre las rectas. Adicionalmente se calculó el trabajo asociado a los eventos de desplegamiento y replegamiento como el área bajo la curva correspondiente al cambio de extensión asociado a cada evento ($W = F \cdot \Delta x$).

Los datos de fuerza y extensión se utilizaron para calcular el **largo de contorno** (largo teórico de un polímero completamente estirado) de las proteínas estudiadas. Para ello se aplicó el **modelo de la cadena vermiforme (WLC, por su sigla en inglés)** (ecuación 1) que relaciona la fuerza aplicada a un polímero flexible con la extensión que experimenta en un baño térmico (Bustamante y cols 1994).

$$\frac{FP}{k_B T} = \frac{1}{4} \left(1 - \frac{(F_d + x)}{L_0} \right)^{-2} - \frac{1}{4} + \frac{(F_d + x)}{L_0}$$

Ecuación 1. Modelo de la cadena vermiforme.

F: Fuerza (pN), F_d: Distancia entre cisteínas en el estado plegado, la cual corresponde a 4 nm para mARC_OT y 1,7 nm para pARC (nm), k_B: constante de Boltzman (pNm/K), T: temperatura absoluta (K), P: largo de persistencia (nm) se utiliza un valor de 0,65 nm (Cecconi y cols 2005), L₀: largo de contorno (nm), x: extensión (nm).

El largo de contorno teórico de una proteína puede calcularse de acuerdo al número de aminoácidos de su cadena polipeptídica (ecuación 2). Este valor debiera tener un límite máximo correspondiente al número de aminoácidos entre cisteínas utilizadas para estirar la molécula. Como se muestra en la figura 5 pARC posee 103 y mARC 120 aminoácidos entre las cisteínas por lo que según la ecuación 2 se esperan valores máximos de L₀ de 37 y 43 nm respectivamente.

$$L_0 = N \cdot 0,35 \text{ nm/aa}$$

Ecuación 2. Cálculo del largo de contorno teórico de una proteína

L₀: largo de contorno (nm), N: número de aminoácidos.

2.11 Calculo de parámetros termodinámicos. El teorema de fluctuaciones de Crooks (ecuación 3) permite calcular la energía libre (ΔG) en condiciones de equilibrio asociada al cambio de extensión de una molécula a partir de las distribuciones de trabajo de desplegamiento y replegamiento obtenidos fuera del equilibrio (Crooks 2000). En el caso del estiramiento de una proteína, este valor corresponde a la diferencia de energía libre entre el estado compacto y extendido y por lo tanto se le ha asociado a la estabilidad de la proteína. De acuerdo a este teorema, el ΔG al equilibrio corresponde al trabajo en el que se entrecruzan las distribuciones de trabajo de desplegamiento y replegamiento obtenidas al estirar y relajar repetidas veces la proteína.

$$\frac{P_u(W)}{P_f(-W)} = e^{[\beta(w-\Delta G)]}$$

Ecuación 3. Teorema de fluctuaciones de Crooks.

$P_u(W)$: distribución de trabajo de desplegamiento, $P_f(-W)$: distribución de trabajo de replegamiento, β : $(k_B T)^{-1}$, con k_B constante de Boltzman (pNm/K) y T temperatura absoluta (K), W : trabajo (pNm), ΔG : energía libre de Gibbs (pNm).

2.12. Análisis de datos cinéticos. Las constantes cinéticas (k) obtenidas a partir de los datos de los experimentos realizados a fuerza constante se analizaron utilizando el modelo de Bell (Bell y cols 1984) el cual relaciona la constante cinética de un proceso con la fuerza a la cual este ocurre (ecuación 4). Como se desprende de la ecuación esta es una relación lineal entre el logaritmo natural de k y la fuerza a la cual ocurre el proceso cuya tasa de cambio corresponde a la distancia hacia el estado de transición y su intercepto en el eje Y corresponde al logaritmo natural de k a cero fuerza.

$$k(F) = k_0 e^{\frac{F \Delta x^\ddagger}{k_B T}}$$

$$\ln(k) = \ln(k_0) + \frac{\Delta x^\ddagger}{(k_B T)} F$$

Ecuación 4. Modelo de Bell.

$k(F)$: constante cinética en función de la fuerza (s^{-1}), F : fuerza (pN), k_0 : constante cinética a cero fuerza (s^{-1}), k_B : constante de Boltzman (pNm/K), T : temperatura absoluta (K), Δx^\ddagger : distancia al estado de transición (nm).

Los datos obtenidos de los experimentos cinéticos ofrecen una manera alternativa de calcular los ΔG de desplegamiento mediante dos aproximaciones. Una es mediante el cálculo de la constante de equilibrio a partir de las constantes cinéticas extrapoladas a cero fuerza, y la segunda corresponde al cálculo del trabajo realizado a fuerza media.

La fuerza a la cual la velocidad de desplegamiento iguala a la velocidad de replegamiento es conocida como fuerza media ($F_{1/2}$). A temperatura y presión constante el trabajo reversible está dado por: $W_{rev}(F) = \Delta G^\circ + k_B T \ln[K_{eq}(F)] + \Delta G(F)_{estiramiento}$ (Liphardt y cols 2001), a fuerza media la expresión $k_B T \ln[K_{eq}(F)] = 0$ ya que $K_{eq} = 1$ por lo que $W_{rev}(F_{1/2}) = \Delta G^\circ + \Delta G_{estiramiento}$, considerando que $W_{rev}(F_{1/2})$ es aproximadamente $F_{1/2} \cdot \Delta x(F_{1/2})$ se puede obtener una expresión para calcular el ΔG de desplegamiento a fuerza cero a partir del trabajo realizado a fuerza media (ecuación 5).

Ecuación 5. Cálculo de ΔG a partir de trabajo realizado a fuerza media.

ΔG° : energía libre de Gibbs estándar a fuerza cero (pNnm), $F_{1/2}$: fuerza media (pN), $\Delta x_{F_{1/2}}$: diferencia de extensión entre el estado plegado y desplegado a fuerza media (nm), el ΔG de estiramiento se calcula del área bajo la curva de WLC entre 0 y la fuerza media (pNnm).

$$\Delta G^\circ = \left[F_{1/2} \cdot \Delta x(F_{1/2}) \right] - \Delta G_{estiramiento}$$

Resultados.

Purificación y modificación química de las proteínas mARC y pARC para su utilización en experimentos de pinzas ópticas. Utilizando el método expuesto por Cecconi y cols (2008) se puede estudiar el mecanismo de plegamiento de proteínas únicas, evitando el uso de poliproteínas (naturales o artificiales) como lo requieren otros métodos (Rief 1997). Esto se logra manipulando la proteína a través moléculas de ADN unidas a residuos de cisteína expuestos a solvente (Figura 6). Las moléculas de ADN unidas a la proteína funcionan como espaciadores que impiden interacciones inespecíficas entre las microesferas (Cecconi y cols 2005). La manipulación se puede realizar desde cualquier punto de la estructura de la proteína, permitiendo la elección del eje de aplicación de la fuerza. En el equipo de pinzas ópticas, el complejo proteína-ADN se une a microesferas por unión antígeno – anticuerpo o por interacción biotina – estreptavidina. Una esfera es atrapada en la trampa óptica y corresponde a la que presenta anticuerpos contra digoxigenina en su superficie, mientras que la otra esfera (que presenta estreptavidina en su superficie) es mantenida fija en la punta de una pipeta microscópica por succión. Este diseño experimental permite estudiar la proteína en el rango de fuerzas bajas (entre 0 y 140 pN) lo cual da la posibilidad de explorar cambios conformacionales fuera del equilibrio y también al equilibrio (Cecconi y cols 2008).

Purificación de mARC, mARC_OT y pARC_OT. El protocolo de purificación de las proteínas se basó en uno previamente descrito por Robinson y Sauer (1996b), sin embargo, fueron requeridas algunas modificaciones que se explican a continuación. El primer paso consistió en una cromatografía de afinidad a Ni^{2+} en condiciones desnaturantes de tal manera de inactivar las proteasas celulares y evitar al máximo la degradación de la proteína a purificar. Sin embargo, al dializar las mutantes OT para replegarlas se observó formación de una elevada cantidad de agregados. Esto fue solucionado utilizando una columna de exclusión molecular, de esta manera las proteínas se replegaron en la columna y al mismo tiempo se separaron de oligómeros de alto peso molecular y otras impurezas remanentes de la cromatografía de afinidad a Ni^{2+} . La figura 7A muestra los cromatogramas de la purificación de mARC, mARC_OT y pARC_OT obtenidos mediante el protocolo previamente descrito. De acuerdo a la calibración de la columna, los picos mayoritarios corresponden a las proteínas con los pesos moleculares esperados para su conformación monomérica globular. Estas especies fueron

colectadas para su marcación con las hebras de ADN. Nótese que las mutaciones OT no afectan el perfil de elusión respecto a la proteína silvestre, lo que indica que dichas mutaciones no interfieren en el correcto plegamiento de las proteínas. El rendimiento de recuperación de la proteína luego de la cromatografía fue 33% para mARC, 22% para mARC_OT y 63 % para pARC_OT.

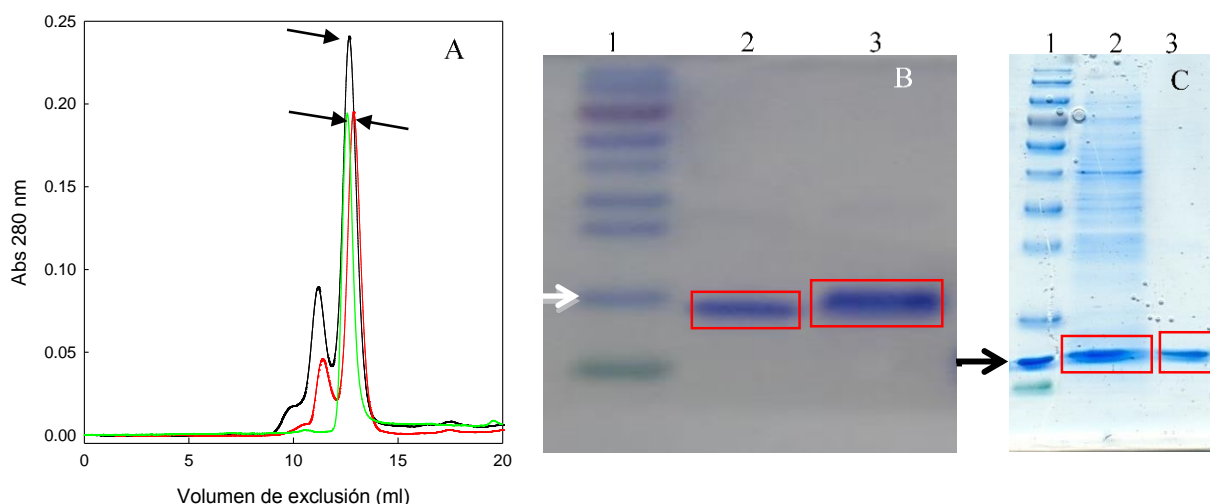


Figura 7. Purificación de mARC, mARC_OT y pARC_OT. Las proteínas se purificaron en dos pasos cromatográficos, luego se analizaron las muestras por SDS-PAGE (Bis-Tris). A. Cromatografía de exclusión molecular, en negro mARC_OT, en verde mARC y en rojo pARC_OT, los picos mayoritarios (indicados por las flechas negras) corresponden al monómero de estas proteínas, los picos minoritarios corresponden a impurezas u oligómeros. B. SDS-PAGE en geles Bis-Tris para mARC y mARC_OT, el carril 2 corresponde a mARC y el carril 3 corresponde a mARC_OT, ambas proteínas presentan un peso molecular de alrededor de 15 kDa de acuerdo al estándar de peso molecular (carril 1, flecha blanca). C. SDS-PAGE en geles Bis-Tris para pARC_OT y mARC, al igual que en B ambas proteínas (mARC en 2 y pARC_OT en 3) presentan un peso molecular de alrededor de 15 kDa de acuerdo al estándar (carril 1, flecha negra).

Síntesis de moléculas de ADN modificado. Se sintetizaron moléculas de ADN de 558 pb mediante PCR utilizando los partidores modificados que presentan en un extremo 5' un grupo sulfhidrido (SH) libre y en el otro extremo 5' biotina (BIO) o digoxigenina (DIG) (materiales y métodos). En la figura 8A se muestra un gel de agarosa de los productos de PCR para la síntesis de las moléculas de ADN. Se pueden observar (carril 2 y 3) bandas de alrededor de 500 pb que corresponden con el tamaño esperado para las moléculas de ADN sintetizadas. En la tabla 3 se muestran los rendimientos de los diferentes pasos de síntesis y purificación de las moléculas de ADN, el rendimiento de la purificación fue de un 90 % y no varió mayormente de una preparación a otra. En cambio, la cantidad de ADN obtenido de la reacción de PCR

varió dependiendo de la polimerasa y la calidad del agua utilizada. En consecuencia, el paso crítico en la generación de las hebras de ADN fue la reacción de PCR.

Tabla 3. Rendimiento de pasos de síntesis y purificación de moléculas de ADN modificado

Etapa	Concentración (µg/µl)	Volumen (µl)	Masa (µg)	% de rendimiento
PCR	0.05	9600	484	100
Purificación	0,09	5000	435,6	90
Precipitación	10,9	40	435,6	90

La unión de las proteínas al ADN se realiza mediante puentes disulfuro entre el grupo sulfhidrilo de su extremo 5' y los residuos de cisteína expuestos al solvente de las proteínas. Por lo tanto, se determinó la reactividad de los grupos sulfhidrilo presentes en las moléculas de ADN mediante la formación de dímeros de ADN. Como se observa en la figura 8B los productos de PCR incubados durante doce horas en ausencia de DTT muestran dos bandas, una mayoritaria de 500 pb (tamaño del monómero) y una minoritaria de alrededor de 1000 pb correspondiente al tamaño del dímero esperado. En cambio, los productos de PCR incubados en 10 mM de DTT durante doce horas sólo presentan una banda de 500 pb correspondiente al monómero de ADN. De la intensidad relativa de estas bandas se puede concluir que aproximadamente del 25% (en algunos casos 50%) de las moléculas de ADN forman el dímero esperado.

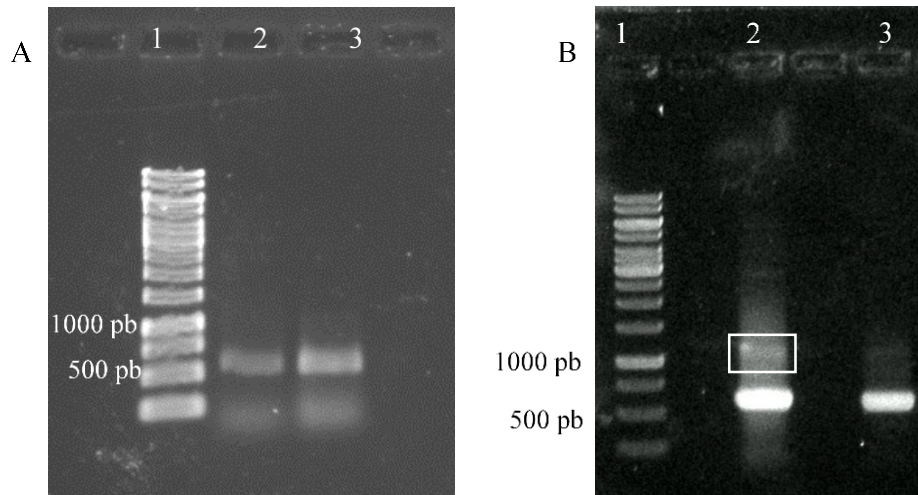


Figura 8. Síntesis y prueba funcional de moléculas de ADN modificado. Geles de agarosa al 1% utilizados para el análisis de las moléculas de ADN. A. Síntesis de moléculas de ADN modificado, el carril 1 corresponde al estándar de peso molecular, el carril 2 corresponde al producto de PCR para la síntesis de ADN modificado con Biotina (BIO) y el carril 3 corresponde al producto de PCR para la síntesis de la molécula modificada con digoxigenina (DIG). B. Prueba funcional de grupos sulfhidrilo de moléculas de ADN, carril 1 estándar de peso molecular, carril 2 productos de PCR en ausencia de DTT, carril 3 productos de PCR incubados por 12 horas en 10 mM DTT.

Modificación química de proteínas. Para modificar las proteínas con las hebras de ADN se requiere un paso previo de activación de las cisteínas expuestas a solvente, para ello se utilizó la sonda DTDP (Cecconi y cols 2008). Considerando que cada proteína posee 2 residuos de cisteína se espera una estequiometría de reacción de 2:1 (DTDP: Proteína). La modificación de las proteínas con DTDP se siguió por un método espectrofotométrico (Grassetti & Murray 1967) lo que permitió calcular la estequiometría de reacción. Como control negativo se utilizó mARC ya que esta proteína no presenta residuos de cisteína. Como se observa en la tabla 4 los resultados muestra una estequiometría de alrededor de 2:1 (DTDP: proteína) lo cual está de acuerdo con lo esperado para esta reacción. Adicionalmente, la modificación de las cisteínas por DTDP se determinó mediante espectrometría de masas (MALDI-TOF). En la figura 9 se comparan los espectros de masas obtenidos para las diferentes proteínas, en presencia y ausencia del modificador. Se observa que sólo mARC_OT y pARC_OT presentan un aumento de 220 kDa lo cual corresponde al peso molecular de dos moléculas de DTDP por proteína. Estos resultados apoyan los resultados obtenidos por el método espectrofotométrico y permiten concluir que la reacción de DTDP con las cisteínas es satisfactoria y presenta la estequiometría adecuada.

Tabla 4. Resultados de modificación de proteínas con DTDP

Proteína	μmol proteína	μmol SH	μmol SH/ μmol proteína
mARC (12 μM)	$2,4 \times 10^{-3}$	~ 0	0
mARC_OT (23 μM)	$4,6 \times 10^{-3}$	$8,4 \times 10^{-3}$	1,8
pARC_OT (20 μM)	$4,3 \times 10^{-3}$	7×10^{-3}	1,6

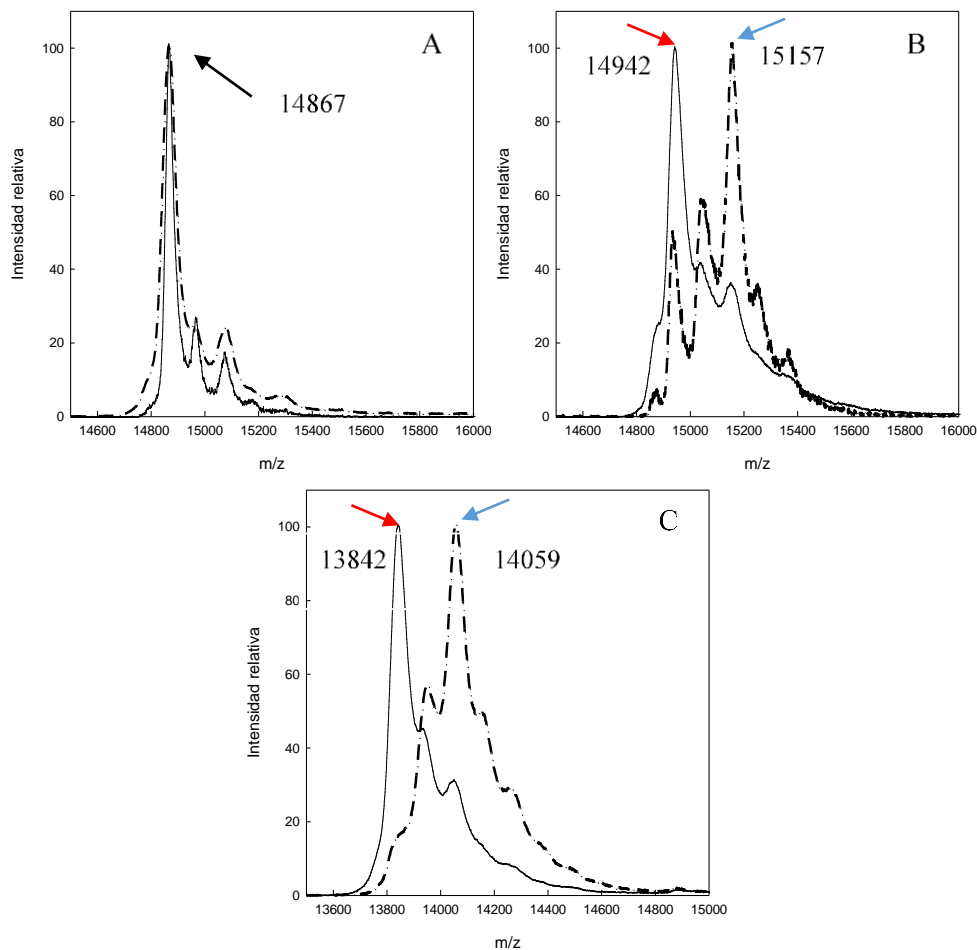


Figura 9. Espectros de masas para proteínas en presencia y ausencia de DTDP. En la figura se muestra los espectros de masas obtenidos para cada una de las proteínas estudiadas, la línea continua (—) corresponde a la proteína sin modificar y la línea discontinua (-•-) corresponde a la proteína incubada con DTDP. A. Espectro de masas para mARC, la proteína no se modificó con la sonda (mARC no presenta cisteínas). B. Espectro de masas de mARC_OT, la proteína sin modificación muestra un pico a un peso molecular correspondiente al esperado para esta proteína sin modificación (flecha roja), en cambio, la proteína incubada con DTDP presenta un pico a un peso molecular correspondiente a la proteína completamente modificada (flecha azul). C. Espectro de masas de pARC_OT, de la condición sin modificador se obtuvo un pico a un peso molecular que da cuenta de la proteína sin modificación (flecha roja). Por el contrario, la proteína modificada con DTDP muestra un pico a un peso molecular que da cuenta de la proteína completamente modificada (flecha azul).

Luego de la modificación de las cisteínas con DTDP la sonda en exceso se retiró utilizando dos columnas Micro Bio-Spin (Bio-Rad) en serie para luego ser mezcladas con ADN en una proporción de 8:1 (ADN: Proteína). Al finalizar la incubación, las proteínas modificadas se purificaron y analizaron por electroforesis en gel de poliacrilamida (Bis-Tris) como se describe en materiales y métodos. En la figura 10 se muestran los geles de la purificación de los complejos proteína-ADN hechos por modificación de pARC_OT (Figura 10A) y mARC_OT (Figura 10B). En ambos casos se pudo observar una banda sobre los 1000 pb en la fracción purificada que corresponde al complejo ADN-proteína-ADN. Al incubar con 20 mM DTT las muestras correspondientes a los complejos la banda de alto peso molecular desapareció y sólo se observó una banda de 500 pb correspondiente al monómero de ADN.

Caracterización del mecanismo de plegamiento de mARC y pARC mediante metodologías de moléculas individuales. Modelos por homología generados para mARC predicen que la proteína puede encontrarse en dos conformaciones dependiendo de la posición del lazo formado por el péptido conector utilizado para unir las subunidades del dímero de ARC. Una conformación anudada (Figura 5A) o una conformación no-anudada (Figura 5B). Debido a que los nudos están matemáticamente definidos en cuerdas cerradas (Zinn-Justin & Zuber 2010), una manera de determinar si la cadena polipeptídica está anudada o no es tirar desde sus extremos ya que de esta forma cerramos el sistema y se hace imposible desatar el nudo mientras estiramos la cadena (Taylor 2000). De esta manera, al unir las hebras de ADN a ambos extremos de la cadena polipeptídica podemos capturar cada una de las conformaciones posibles para mARC (anudada o no-anudada, Figura 6B). Esto permitiría que al estirar mARC_OT desde sus extremos se pueda caracterizar cada una de las conformaciones esperadas para esta proteína. Siendo así, se espera observar dos comportamientos; uno que dé cuenta de la conformación anudada y otro que dé cuenta de la conformación no-anudada. Como control se utilizó la permutante pARC (Tabtiang y cols 2005), la cual presenta la misma arquitectura que mARC pero con una topología que le impide anudarse. En este caso se espera observar un comportamiento único que dé cuenta del completo estiramiento mecánico de pARC.

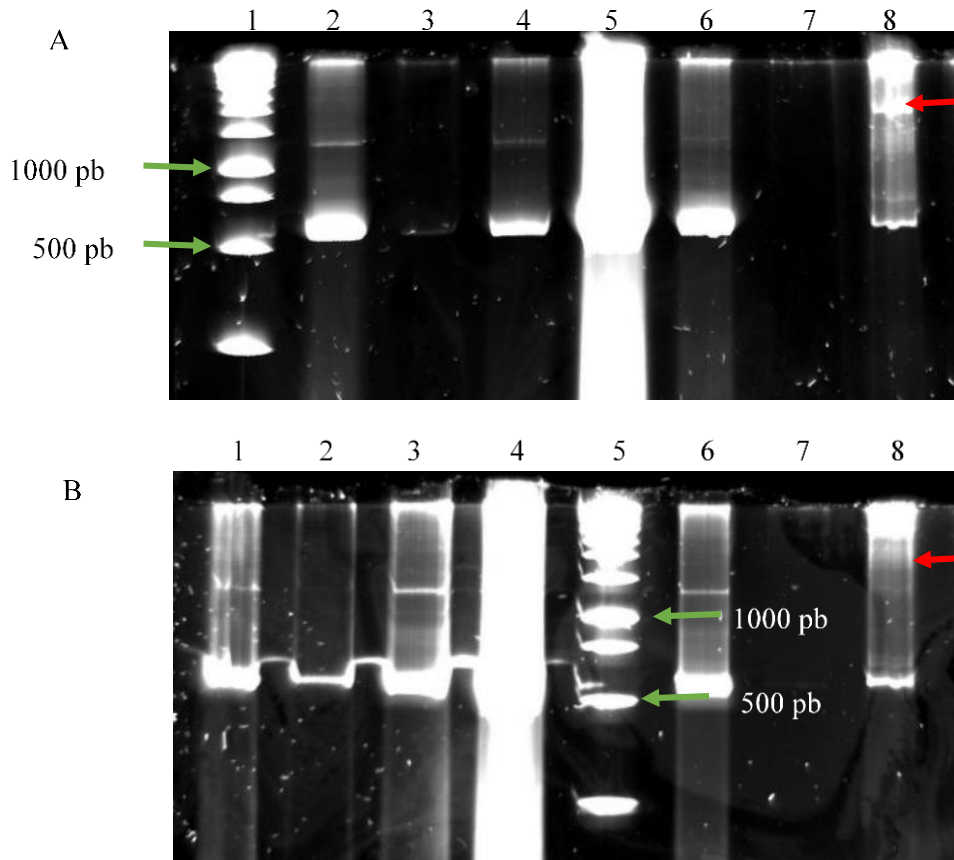


Figura 10. Purificación de complejo ADN-Proteína. Geles de poliacrilamida (Bis-Tris) al 10% teñido con GelRed para visualizar ADN. A. Modificación química de pARC_OT, el carril 1 es estándar de peso molecular, el carril 2 corresponde al ADN utilizado en la reacción de modificación, el carril 4 es la mezcla de reacción previo a su purificación, en el carril 5 está el frente de cargar, el carril 6 son los lavados y el carril 8 corresponde al complejo ADN-proteína (flecha roja), de 2 a 8 se puede observar el enriquecimiento de la banda por sobre los 1000 pb y la disminución de la banda de 500 pb, dando cuenta de la purificación del complejo ADN-proteína. B. Modificación química de mARC_OT, el carril 1 corresponde al ADN utilizado en la reacción de modificación, el carril 3 es la mezcla de reacción previo a su purificación, en el carril 4 está el frente de cargar, el carril 5 corresponde a estándar de peso molecular, el carril 6 son los lavados y el carril 8 corresponde al complejo ADN-proteína (flecha roja), al igual que en A de 1 a 8 se puede observar el enriquecimiento de la banda sobre los 1000 pb y la disminución de la banda a 500 pb.

Estiramiento de las proteínas a velocidad constante. Este experimento consiste en estirar la proteína desde sus extremos a una velocidad de estiramiento constante para aplicar fuerza sobre las moléculas como si fuese un agente desnaturante y así estudiar las transiciones de desplegamiento y replegamiento de la proteína (Shank y cols 2010). Los datos obtenidos de estos experimentos se representan en curvas de fuerza vs extensión como se muestra en la figura 11A. En esta figura se puede observar que al aumentar la distancia entre las microesferas ocurre un incremento no lineal de la fuerza aplicada. Este incremento en la extensión y la fuerza (curva roja, Figura 11A), es debido a la resistencia que oponen las

moléculas de ADN al ser estiradas. Ya que la proteína en su estado nativo es más rígida que el ADN (Bustamante y cols 2000, Rief 1997) presenta una mayor resistencia a la deformación y no aporta a la extensión del ADN. Sin embargo, cuando la tensión es lo suficientemente elevada para romper las interacciones que estabilizan el estado nativo se observa un repentino incremento de la distancia entre las microesferas. **Este evento corresponde al desplegamiento de la proteína** y la fuerza a la cual ocurre este evento corresponde a la fuerza de desplegamiento para dicha proteína (flecha roja en la Figura 11A) (Cecconi y cols 2005). De manera análoga al desplegamiento, al disminuir la fuerza (curva azul en la Figura 11A) se incrementa la probabilidad de formar las interacciones del estado nativo permitiendo **el replegamiento de la proteína**. Este cambio se observa como **una disminución abrupta de la distancia entre las microesferas** debido a la compactación de la proteína. La fuerza a la que esto ocurre corresponde a la fuerza de replegamiento (flecha azul en la Figura 11A).

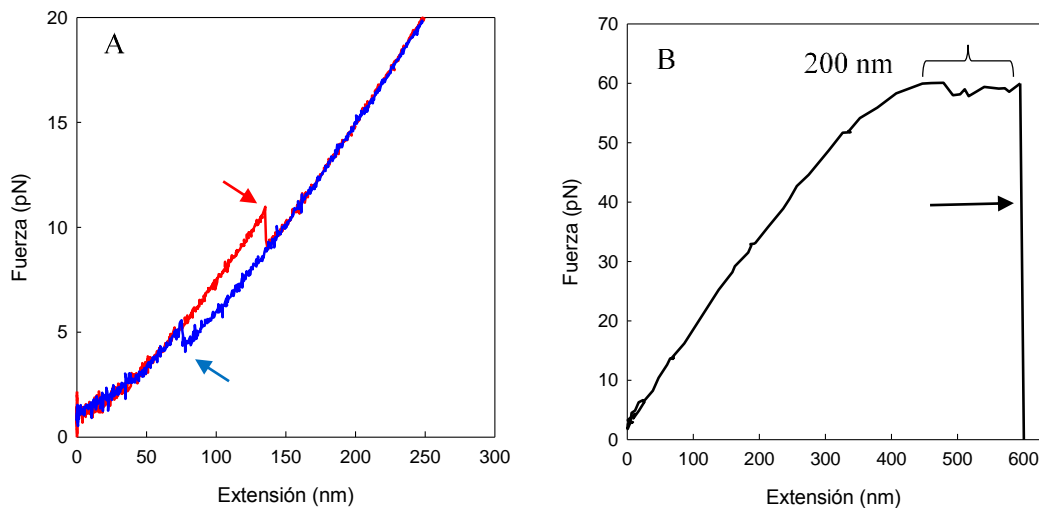


Figura 11. Curva de fuerza v/s extensión, obtenida de los experimentos de rampa de fuerza, para una molécula de mARC_OT y sobreestiramiento del ADN. A. Curva de fuerza vs extensión, la curva roja corresponde al aumento de la fuerza a medida que la extensión aumenta, la fuerza indicada por la flecha roja corresponde a la transición de desplegamiento. La curva azul muestra la caída de la fuerza a medida que la extensión disminuye, la fuerza indicada por la flecha azul corresponde a la transición de replegamiento. B. Sobreestiramiento del ADN, cuando las moléculas modificadas con ADN son estiradas por sobre los 65 pN se observa el fenómeno de sobreestiramiento del ADN en donde se puede observar un *plateau* de aproximadamente 200 nm (70% de la extensión total del ADN) y luego una caída en la fuerza (flecha negra).

Para determinar que los cambios descritos anteriormente correspondan a una sola molécula se realizó un control que consistió en estirar el complejo ADN-proteína hasta 65 pN. A esta fuerza se ha descrito que ocurre el sobreestiramiento del ADN lo cual se observa como un *plateau* en la curva de fuerza vs extensión que da cuenta del 70% de la extensión total de ambas hebras de ADN (Smith y cols 1996). Se espera, entonces, que la presencia de un único complejo ADN-Proteína entre las microesferas de poliestireno genere una curva con un *plateau* que corresponda al 70% del total de la extensión de ambas hebras de ADN. Dado que el largo total de ambas hebras de ADN es de alrededor de 340 nm (1100 pb) se espera observar un *plateau* de entre 200 a 250 nm, aproximadamente. En efecto, al estirar los complejos estudiados a fuerzas de 65 pN se observa el comportamiento descrito anteriormente (Figura 11B). Además, al aplicar fuerzas superiores a 65 pN la unión de la estreptavidina con la biotina se rompe, observándose una caída de la fuerza hasta cero pN sin cambio alguno en la extensión (flecha negra en la Figura 11B). Algunas moléculas no alcanzan a ser estiradas hasta a observar el sobreestiramiento del ADN pero si se observa una caída única, este criterio también puede ser utilizado para seleccionar las moléculas a analizar. En este trabajo se utilizaron ambos criterios para seleccionar los datos que provenían de moléculas únicas (8 moléculas con sobreestiramiento y 19 moléculas con sólo caída única).

Los experimentos de estiramiento a velocidad constante se realizaron con 27 moléculas diferentes y en total se obtuvieron 2719 transiciones de desplegamiento y replegamiento observadas entre dos y treinta pN. En este intervalo de fuerzas, la molécula muestra transiciones únicas tanto para el desplegamiento como para el replegamiento de mARC_OT lo que sugiere que la proteína se despliega en una sola etapa sin la presencia de intermediarios. Esto concuerda con estudios hechos mediante experimentos de estabilidad en condiciones de equilibrio utilizando agentes caotrópicos (Robinson & Sauer 1996a), los cuales han descrito el mecanismo de plegamiento de mARC utilizando un modelo de dos estados, sin intermediarios.

No obstante, al observar las fuerzas de replegamiento de las moléculas estudiadas se distinguieron dos tipos de transiciones de replegamiento. Esto se desprende de la distribución bimodal que muestra el histograma de las fuerzas de replegamiento realizado de todas las transiciones observadas. Como se observa en la figura 12A existen dos tipos de transiciones de replegamiento, unas que ocurren a bajas fuerzas con un promedio de **5,3 ± 0,6 pN** y otras que ocurren a fuerzas elevadas con un promedio de **8,4 ± 1 pN**.

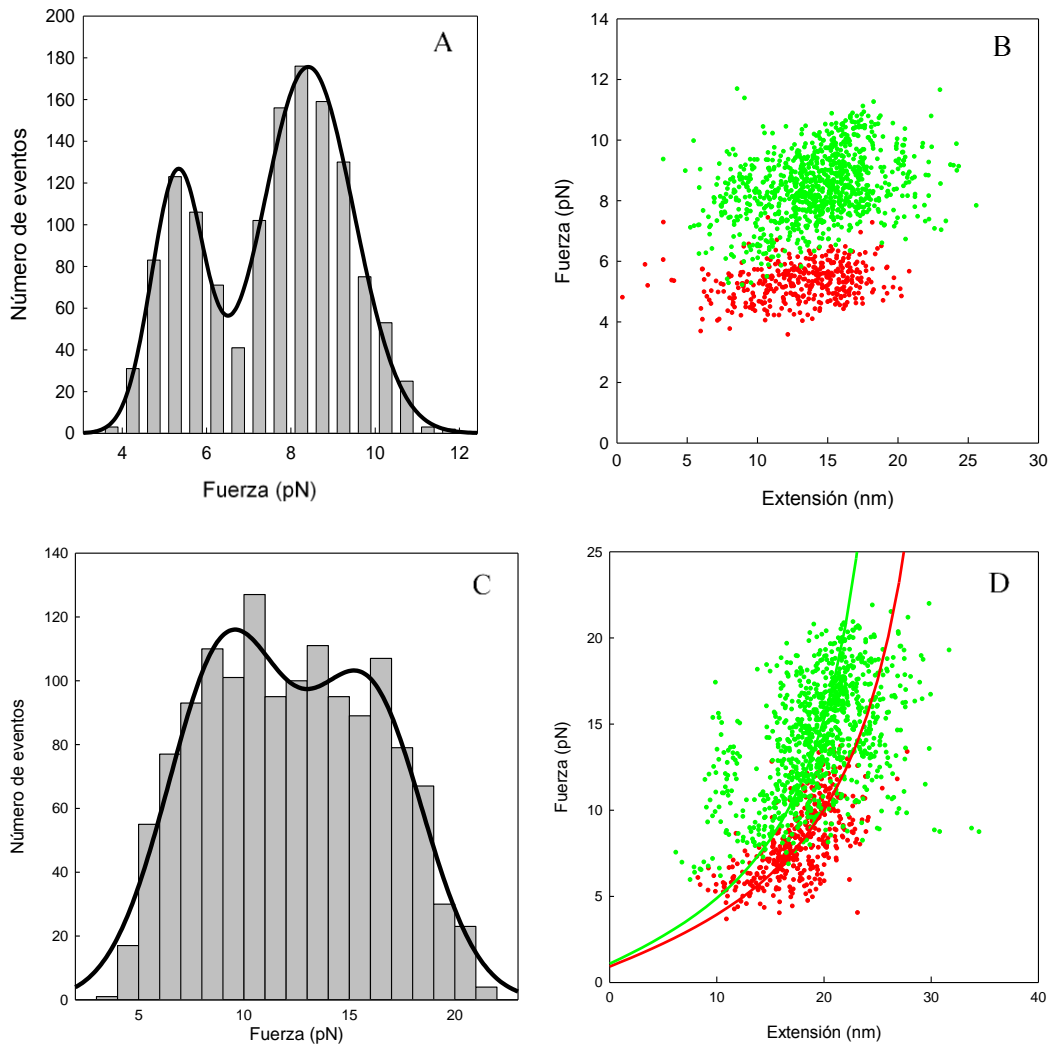


Figura 12. Análisis de datos de fuerza-extensión de mARC_OT. A. Histograma de fuerzas de repliegamiento para mARC_OT, muestra claramente dos tipos de transiciones de repliegamiento B. Fuerza en función de la extensión de los eventos de repliegamiento de para mARC_OT, se observan dos poblaciones de moléculas diferenciadas por sus fuerzas de repliegamiento, a fuerzas bajas (símbolos rojos) y a fuerzas altas (símbolos verde). C. Histograma de fuerzas de desplegamiento para mARC_OT, no es posible detectar dos comportamientos de manera clara. D. Fuerza vs extensión para el desplegamiento de mARC_OT, las moléculas que despliegan a fuerzas altas (símbolos verdes) están representadas por una curva de WLC con L_0 de 37 nm (curva verde), por otro lado, las moléculas que despliegan a fuerzas bajas (símbolos rojo) están representadas por una curva WLC con L_0 de 43 nm (curva roja).

Debido a que de una molécula de proteína se pueden obtener múltiples transiciones de desplegamiento y repliegamiento (en promedio se obtuvieron 100 transiciones por molécula) es importante determinar si ambos tipos de transiciones indicadas por el histograma de fuerzas pueden ser originadas por una molécula o dependen de la proteína atrapada entre las microesferas. Dicho de otra manera, es necesario determinar si es posible clasificar las

moléculas atrapadas entre las esferas de acuerdo a las fuerzas de replegamiento que ellas experimentan; baja o alta. Para evaluar esta posibilidad se compararon los eventos de replegamiento de cada una de las moléculas estudiadas con todas las demás, sin generar grupos a priori, mediante un análisis de varianza (ANOVA). Este análisis agrupó las moléculas que presentan la misma varianza de sus eventos de replegamiento. Los resultados de dicho análisis permitieron generar dos grupos de moléculas (Figura 13A) con diferencias significativas en el promedio de sus fuerzas de replegamiento ($p < 0.05$) según la prueba estadística de Turkey. Estos resultados indicaron que las transiciones observadas dependen de la molécula analizada. Es decir, aquellas moléculas que repliegan a bajas fuerzas no pueden replegar también a altas fuerzas, y viceversa. Se encontró que un grupo de 20 moléculas (958 eventos de replegamiento) mostró el patrón de fuerza-extensión de elevada fuerza (símbolos verdes, Figura 12B). El otro grupo de 7 moléculas (380 eventos de replegamiento) mostró el patrón de fuerza-extensión de baja fuerza (símbolos rojos, Figura 12B). Este comportamiento “molécula dependiente” de las fuerzas de replegamiento soporta la hipótesis de que la proteína mARC puede estar en dos conformaciones que se mantienen una vez que estas han sido atrapadas en la trampa óptica.

Habiendo clasificado las moléculas de acuerdo a su comportamiento respecto a las fuerzas de replegamiento, se procedió a determinar el número de residuos involucrados en el desplegamiento y replegamiento de ambos tipos de moléculas. Para determinar el número de aminoácidos involucrados en estas transiciones se determinó el **largo de contorno (L_0)**, que corresponde a la extensión máxima de un polímero cuando está completamente estirado. El L_0 de un polímero puede ser determinado a partir del **modelo de la cadena vermiforme (WLC, por su sigla en inglés)**. Sin embargo, este modelo pierde poder predictivo a bajas fuerzas (Bouchiat y cols 1999), es decir, en el intervalo de fuerzas en donde se observa normalmente el replegamiento de la proteína. Por esta razón, normalmente los estudios utilizan las transiciones de desplegamiento para calcular el largo de contorno de las moléculas. El histograma de fuerzas de desplegamiento de todas las moléculas (Figura 12C) no mostró dos tipos de conformaciones tan claramente como ocurrió con los eventos de replegamiento. No obstante, un análisis estadístico de los datos, similar al realizado para los datos de replegamiento, pudo separar dos comportamientos como los observados para las fuerzas de replegamiento (Figura 14B). Un grupo de 20 moléculas (960 eventos de desplegamiento) presentaron el patrón de fuerza-extensión de elevada fuerza (en promedio $14,1 \pm 3,3$ pN,

símbolos verdes en la Figura 12D). El otro grupo de 7 moléculas (421 eventos de desplegamiento) presentó el patrón de fuerza-extensión de baja fuerza (en promedio $8,1 \pm 2,2$ pN, símbolos rojos en la Figura 12D). Consistentemente, estos grupos coincidieron con los grupos obtenidos a partir del análisis de los datos de replegamiento los cuales muestran una segregación mucho más evidente. En conclusión aquellas moléculas que se despliegan a fuerzas altas también se repliegan a fuerzas altas y las moléculas que despliegan a fuerzas bajas también se repliegan a fuerzas bajas.

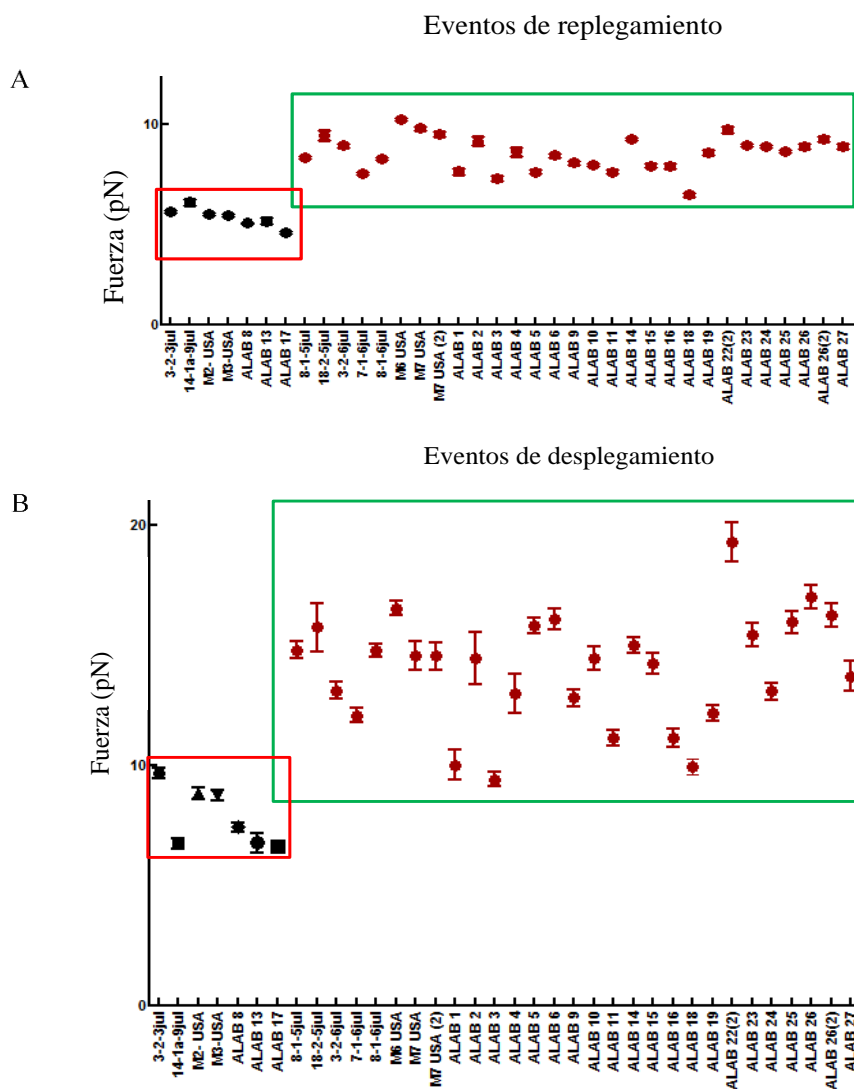


Figura 13. Análisis de ANOVA de los eventos de desplegamiento y replegamiento de mARC_OT. A. ANOVA de los eventos de replegamiento de mARC_OT, el análisis agrupa las moléculas que presentan la misma varianza generando dos grupos de moléculas (cuadros rojo y verde). B. ANOVA de los eventos de desplegamiento, se puede observar una mayor dispersión de los datos pero aun así es posible segregar los datos en dos grupos que coinciden con los grupos generados en A (cuadros rojo y verde).

La variación de la fuerza versus extensión de cada uno de los dos tipos de moléculas puede ser representado, aproximadamente, por el modelo WLC utilizando largos de contorno diferentes. La curva del modelo WLC que representa mejor el comportamiento de las moléculas que despliegan a fuerzas bajas puede obtenerse con un valor de L_0 de **43 nm** (curva roja en la Figura 12D) el cual corresponde al **L_0 esperado para mARC no-anudada**. Por otra parte, la curva del modelo WLC que representa mejor el comportamiento de las moléculas que se despliegan a fuerzas altas correspondió al de una molécula con un L_0 de **37 nm** (curva verde en la Figura 12D). De acuerdo a cálculos teóricos (Dzubiella 2009) el largo de contorno de una proteína debería disminuir en 5-6 nm debido a la presencia de un nudo del tipo 3_1 (como el esperado para mARC) por lo que el L_0 determinado para las moléculas que despliegan a altas fuerzas (**37 nm**) correspondió al **L_0 esperado para mARC en conformación anudada**. Sin embargo, por causa de la gran dispersión de los datos este método para determinar el largo de contorno fue muy aproximado y poco confiable. Para calcular el L_0 de cada grupo de moléculas, sin trazar una curva aproximada, se compararon las extensiones observadas con aquellas predichas por el modelo WLC. Para ello a partir de las fuerzas de desplegamiento se calcularon las extensiones teóricas predichas por el modelo para una molécula de L_0 de 43 nm y luego se obtuvo la razón entre las extensiones experimentales y las calculadas ($\frac{d_x(\text{experimental})}{d_x(\text{calculado})}$). Los histogramas de las extensiones fraccionales presentados en la figura 14 muestran dos tipos de moléculas con extensiones fraccionales de **$1 \pm 0,2$** y **$0,86 \pm 0,17$** lo que equivale a valores de L_0 de **43 ± 7 nm** y **37 ± 7 nm**, respectivamente. Estos resultados apoyan la hipótesis de que mARC puede ser atrapado en dos conformaciones ya se observan los largos de contorno esperados para cada conformación predicha para esta proteína.

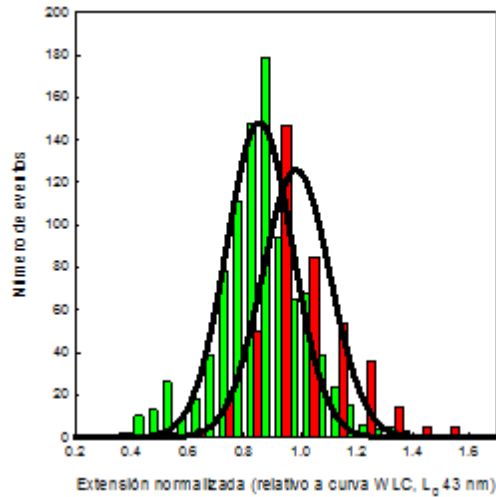


Figura 14. Histograma de extensión fraccional de mARC_OT. Histograma de extensiones fraccionales de desplegamiento de mARC_OT, las barras en verde corresponden a las moléculas que despliega a fuerzas altas las que presentan una extensión fraccional de 0,86, las barras en rojo representan a las moléculas que despliega a bajas fuerzas y presenta una extensión fraccional de 1 aproximadamente.

Para asegurar que los resultados obtenidos con mARC_OT se debieron a una diferencia topológica entre los dos tipos de moléculas identificados, se utilizó la proteína pARC ya que su topología impide la formación de un nudo al ser estirada desde los extremos amino y carboxilo (Figura 5C). Como se desprende del histograma de fuerzas de replegamiento generado a partir de las 483 transiciones de replegamiento obtenidas de 13 moléculas diferentes (Figura 15A), la permutante pARC mostró una distribución unimodal con un promedio de $5,9 \pm 0,8$ pN. Una situación similar se observó para los eventos de desplegamiento de pARC_OT (573 transiciones, Figura 15C) lo que sugiere que pARC no fue capaz de originar conformaciones con diferentes largos. Las transiciones de desplegamiento fueron analizadas de la misma forma que las de mARC_OT. Dicho análisis arrojó un único largo fraccional de $0,6 \pm 0,2$ el cual equivale a un largo de contorno de 22 ± 7 nm.

En resumen, los resultados obtenidos para mARC y su comparación con pARC permiten postular que **mARC_OT en solución existe en dos conformaciones**: una **anudada**, caracterizada por un L_0 de 37 ± 7 nm y una **no-anudada**, caracterizada por un L_0 de 43 ± 7 nm.

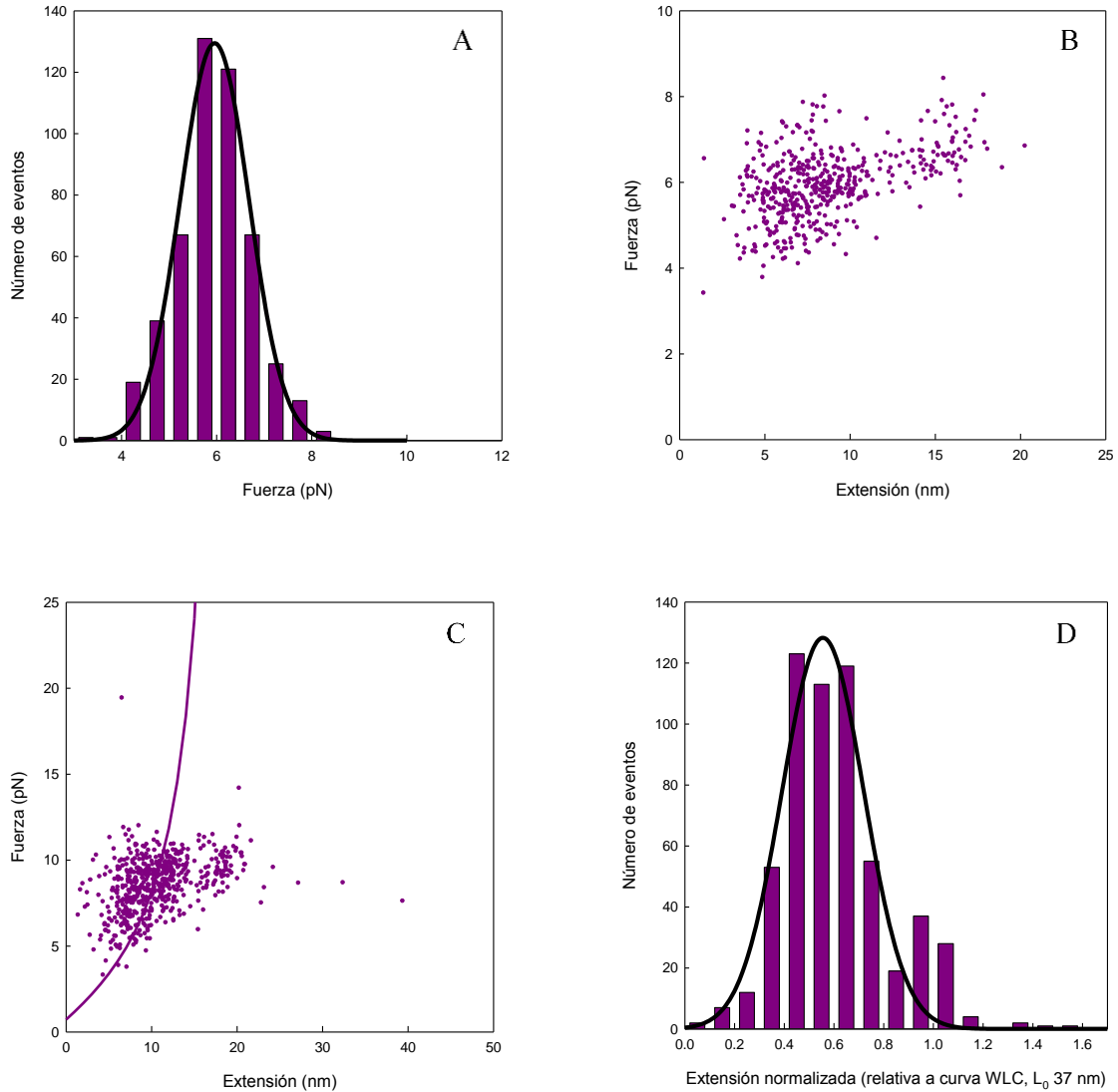


Figura 15. Análisis de datos de fuerza-extensión de pARC_OT. A. Histograma de fuerzas de replegamiento para pARC_OT, se observa una distribución unimodal con una fuerza promedio de alrededor de 6 pN. B. Fuerza en función de la extensión de los eventos de replegamiento de moléculas de pARC_OT, muestra un comportamiento centrado en replegamientos a bajas fuerzas (alrededor de 6 pN). C. Fuerza vs extensión para el desplegamiento de pARC, se observa comportamiento único caracterizado por un WLC con L_0 de 23 nm. D. Histograma de extensiones fraccionales de pARC, la distribución está centrada en 0,6 lo cual indica que pARC se comporta como una proteína más pequeña de lo esperado de acuerdo a su secuencia.

Cálculo del ΔG al equilibrio a partir de trabajo de desplegamiento/replegamiento. Para determinar la estabilidad asociada al desplegamiento de las proteínas estudiadas en condiciones de equilibrio se utilizó el teorema de Crooks (Crooks 2000). Este teorema permite determinar el ΔG en condiciones de equilibrio a partir del trabajo irreversible realizado en condiciones fuera del equilibrio. El trabajo realizado durante el desplegamiento y replegamiento se calculó a partir del área bajo las curvas de fuerza vs extensión. De esta manera, para cada transición de replegamiento o desplegamiento se calculó un trabajo asociado a este proceso (ver materiales y métodos). El ΔG al equilibrio fue calculado como el trabajo en el que se interceptan las gaussianas ajustadas a los histogramas de trabajo de desplegamiento y replegamiento de las proteínas estudiadas (Flecha amarilla en la Figura 16). En el caso de mARC_OT se analizaron los datos provenientes de cada grupo de moléculas de manera independiente. Para aquellas moléculas identificadas como **anudadas** o de menor largo de contorno se calculó un ΔG de **$21,8 \pm 0,4$ kcal/mol** (ΔG_1 , Figura 16A) y para las moléculas identificadas como **no-anudadas** o de mayor largo de contorno se calculó una estabilidad de **$12,3 \pm 0,5$ kcal/mol** (ΔG_2 , figura 16B). En el caso de pARC_OT, el análisis de los datos de trabajo arrojó un ΔG de **$7 \pm 0,4$ kcal/mol** (ΔG_3 , Figura 16C).

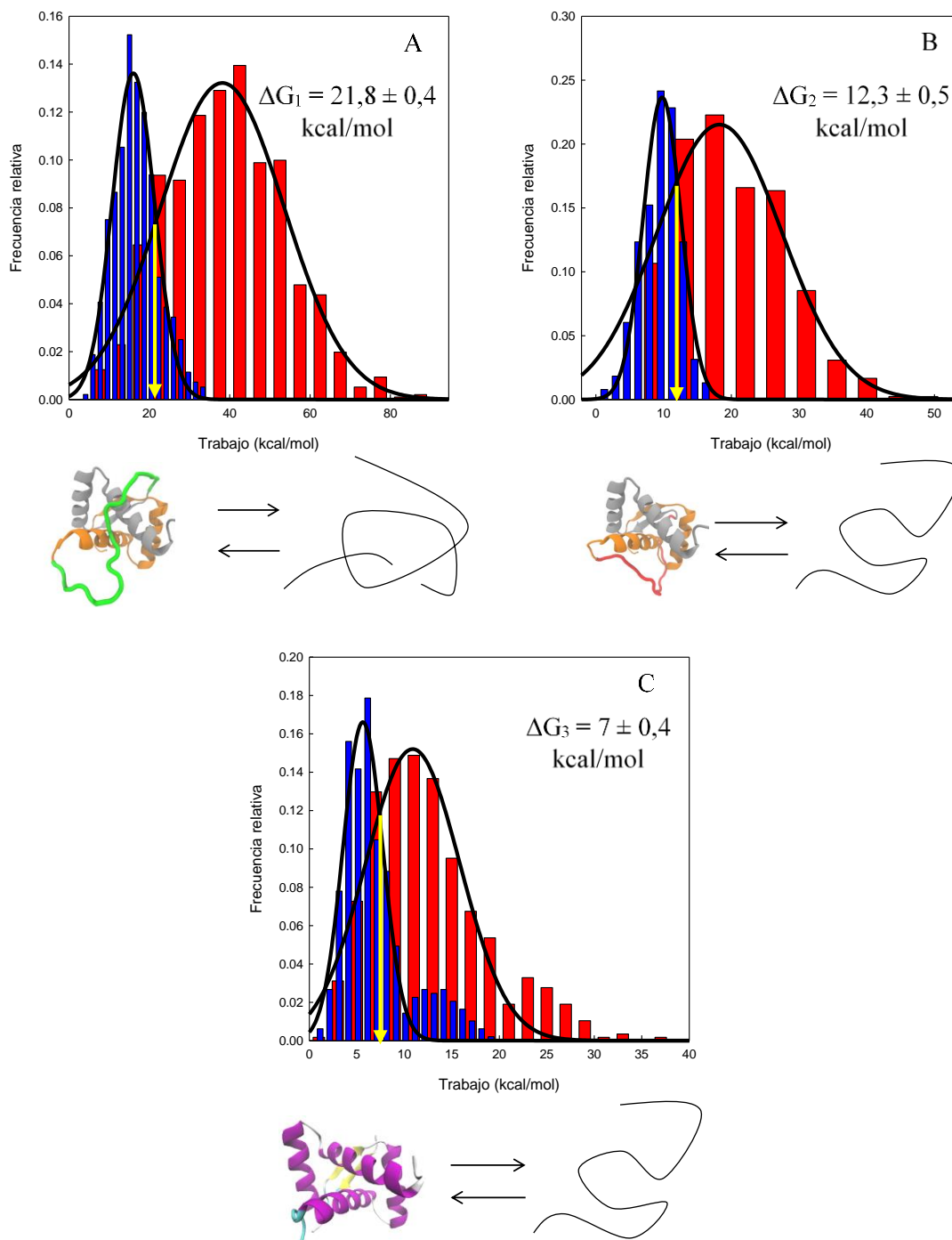


Figura 16. Histogramas de trabajo de desplegamiento y repliegamiento para mARC_OT y pARC_OT. Histogramas de trabajo de desplegamiento (rojo) y repliegamiento (azul), las flechas amarillas indican el valor al cual las gaussianas ajustadas a cada histograma se entrecruzan, el cual corresponde al valor de la estabilidad en condición de equilibrio según el teorema de Crooks. A. mARC_OT anudada. B. mARC_OT no-anudada. C. pARC_OT.

Caracterización cinética del desplegamiento de mARC. Los estudios cinéticos son importantes ya que permiten observar intermediarios metaestables que pueden no ser detectados en los experimentos de estiramiento a velocidad constante. Además permiten explorar características estructurales importantes del estado de transición y dan la posibilidad de caracterizar las barreras energéticas que separan los diferentes estados termodinámicos (Cecconi y cols. 2005, Ozkan y cols. 2001).

La caracterización cinética consiste en mantener las proteínas a una fuerza constante durante un tiempo determinado. La fuerza en el equipo de pinzas ópticas es medida por el cambio del momentum de la luz producido por el desplazamiento del objeto atrapado fuera del centro de la trampa (Smith y cols. 2003). Durante los experimentos de pinza de fuerza la trampa es mantenida fija a una fuerza determinada utilizando el sistema de retroalimentación del equipo. Este sistema monitorea continuamente la posición del objeto atrapado entre los láseres y ajusta la posición de la trampa de tal manera de mantener una fuerza constante aplicada a la microesfera. La mantención de una fuerza constante se logra debido a que el equipo ajusta la distancia entre ambas microesferas de manera tal de que la microesfera atrapada entre los láseres quede siempre a la misma distancia del centro de la trampa (Liphardt y cols. 2001). A modo de ejemplo, en la figura 17A se muestran los cambios de extensión de una molécula de mARC obtenidos en función del tiempo a una fuerza constante de 5,8 pN. Como se observa en el histograma de posición asociado a la trayectoria se pueden distinguir sólo dos estados separados por una distancia de 13 nm. Lo anterior indica que mARC_OT no visita estados intermedios en el régimen de fuerza constante debido a que esta distancia corresponde a la observada para el desplegamiento y replegamiento de estas proteínas en los experimentos de estiramiento a velocidad constante. El tiempo que transcurre la molécula en cada estado puede ser utilizado para determinar las constantes cinéticas de desplegamiento y replegamiento a cada fuerza estudiada (Garcia-Manyes y cols. 2007, Liphardt y cols. 2001, Schlierf y cols. 2004). Por ejemplo, para determinar la constante cinética para el paso del estado nativo al desplegado (N->D; k_{D1}) se genera un histograma acumulativo con los tiempos de residencia observados para el estado nativo (compacto) a lo largo de la trayectoria a determinada fuerza (Figura 17B). Dado que estos histogramas son acumulativos muestran como varía la probabilidad de encontrar el estado nativo en función del tiempo y pueden ser utilizados para determinar la constante cinética suponiendo un decaimiento exponencial simple, como se muestra en la figura 17B. Este procedimiento se repitió para las trayectorias

estudiadas a diferentes fuerzas con lo cual se pudo determinar cómo variaba la constante de desplegamiento en función de la fuerza. De la misma manera se pudo determinar cómo variaba la constante de replegamiento ($U \rightarrow N$, k_f) considerando los tiempos de residencia del estado extendido (desplegado). En la figura 18 se muestra la variación de la constante de replegamiento y desplegamiento de mARC obtenida de trayectorias de dos moléculas diferentes estudiadas a varias fuerzas. Como se observa la variación del logaritmo natural de las constantes en función de la fuerza fue aproximadamente lineal lo cual permite utilizar el modelo de Bell (Bell y cols 1984) para calcular las constantes cinéticas a cero fuerza.

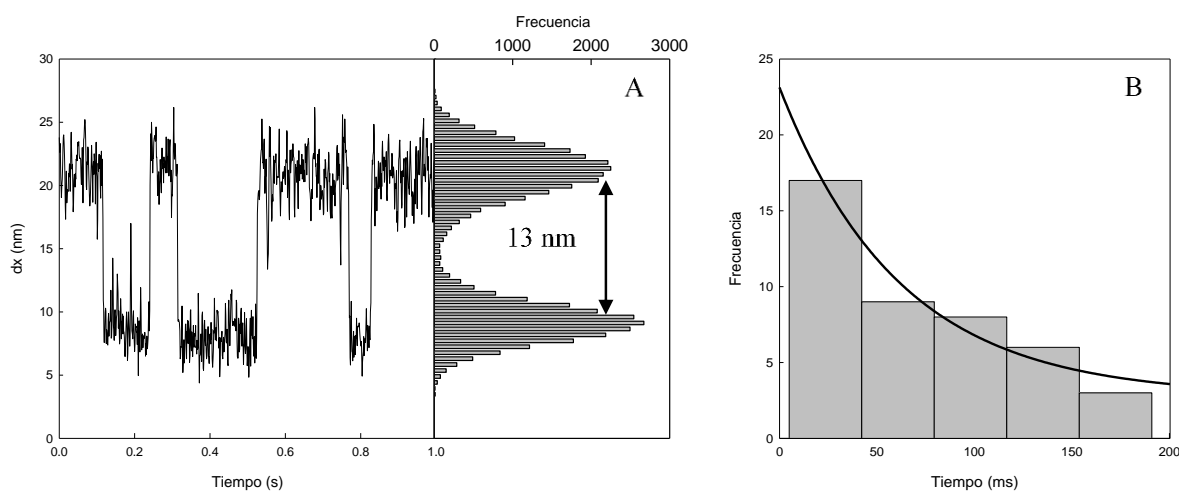


Figura 17. Trayectoria de plegamiento para mARC_OT obtenida de experimentos de pinza de fuerza e histograma de tiempos de permanencia. A. Cuando la proteína es mantenida a una fuerza constante puede “saltar” entre un estado compacto (plegado) y un estado extendido (desplegado), estos estados se caracterizan por una extensión determinada como se observa en el histograma de extensiones asociado a la trayectoria. El tiempo promedio de permanencia en un estado u otro depende de la fuerza aplicada. B. Histograma de tiempos de permanencia para las transiciones de desplegamiento para mARC_OT a fuerza media (5,8 pN), sobre ese histograma se ajusta una exponencial simple (curva negra) y se calcula la constante cinética.

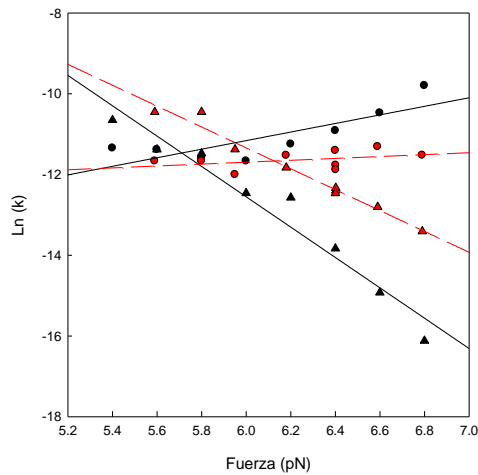


Figura 18. Chevrón de fuerzas para moléculas de mARC_OT. Las constantes cinéticas de desplegamiento y replegamiento obtenidas a partir de los experimentos a fuerza constante se analizaron mediante un gráfico de chevrón y se extrapolaron a fuerza cero utilizando un ajuste de la ecuación Bell. En negro y rojo se indican las moléculas de mARC_OT estudiadas, los puntos (●) corresponden a las constantes cinéticas de desplegamiento y los triángulos (▲) a las constantes cinéticas de replegamiento.

En el caso de los estudios realizados utilizando pinzas ópticas la coordenada de reacción corresponde a su extensión. En consecuencia, fue posible determinar cuánto debe deformarse el estado nativo o el estado desplegado para alcanzar el estado de transición. Este parámetro corresponde la distancia al estado de transición (Figura 19) y su valor se calcula a partir de la pendiente del ajuste al modelo de Bell realizado en el gráfico de logaritmo natural de la constante cinética en función de la fuerza (Liphardt y cols. 2001). En el caso de las moléculas de mARC_OT estudiadas se observó que la distancia al estado de transición de ambas moléculas fue similar tanto desde el estado nativo como desde el estado desplegado, por lo que estaría indicando que ambas proteínas siguen la misma ruta de plegamiento (ver tabla 5).

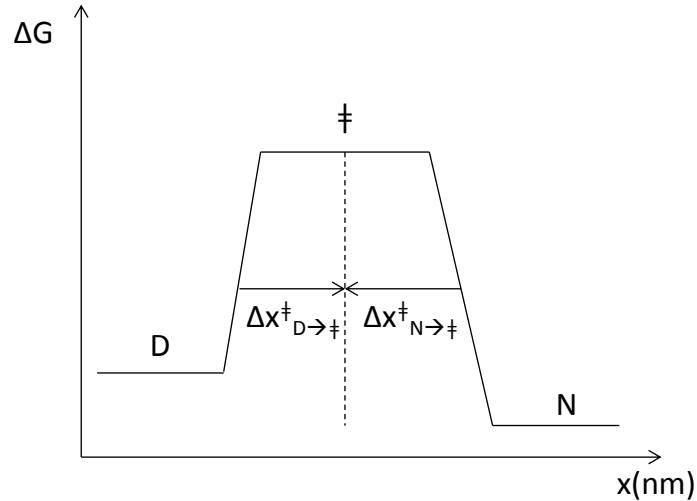


Figura 19. Diagrama de energía libre v/s coordenada de reacción. En el caso de los experimentos de pinzas ópticas la coordenada de reacción es la extensión de la molécula estudiada, así se puede calcular el cambio de extensión desde el estado plegado y desplegado hasta el estado de transición directamente. La suma de $\Delta x_{N \rightarrow \ddagger}^\ddagger$ con $\Delta x_{D \rightarrow \ddagger}^\ddagger$ debe ser igual a la diferencia de extensión entre el estado plegado y desplegado.

Dado que los estudios cinéticos permitieron obtener las constantes cinéticas de plegamiento, estos también permitieron calcular la estabilidad en condiciones de equilibrio para las moléculas estudiadas de manera alternativa a los estudios termodinámicos realizados en condiciones fuera de equilibrio mediante el teorema de Crooks. En este trabajo se utilizaron dos aproximaciones para calcular el ΔG de plegamiento a partir de datos cinéticos. Un método utilizado fue el cálculo de la constante de equilibrio a partir de las constantes cinéticas de desplegamiento y replegamiento extrapoladas a cero fuerza. La otra aproximación fue el cálculo del ΔG a fuerza media ($F_{1/2}$) la cual corresponde a la fuerza en la cual la velocidad de desplegamiento iguala a la velocidad de replegamiento.

Sin embargo, antes de asignar los valores cinéticos, se debe especificar qué tipo de molécula está atrapada en la trampa óptica debido a que como se mostró anteriormente existen dos poblaciones de proteínas que podemos capturar. En la tabla 5 se muestran los diferentes parámetros cinéticos calculados para las moléculas de mARC_OT estudiadas, se puede observar que existen diferencias entre los parámetros obtenidos para cada una de las moléculas lo que podría dar indicios de que cada una corresponde a una conformación distinta de mARC_OT (anudada y no-anudada). Sin embargo, el cálculo del largo de contorno para ambas moléculas arrojó un valor consistente con la conformación no-anudada de mARC_OT (43nm).

En el caso de pARC_OT esta molécula no mostró transiciones a fuerza constante durante el desarrollo de este trabajo.

Tabla 5. Parámetros termodinámicos y cinéticos calculados a partir de experimentos de pinza de fuerza.

Molécula estudiada	$k_0(N \rightarrow D)$ (s^{-1})	$k_0(D \rightarrow N)$ (s^{-1})	$\Delta x_{N \rightarrow \ddagger}$ ‡ (nm)	$\Delta x_{D \rightarrow \ddagger}$ ‡ (nm)	ΔG (kcal/mol)	$F_{1/2}$ (pN)	$dx_{F_{1/2}}$ (nm)	$\Delta G_{F_{1/2}}$ (kcal/mol)
mARC_O T 1	$2,5 \times 10^{-8} \pm 6 \times 10^{-8}$	22027 ± 51771	$4,4 \pm 1$	$15,6 \pm 2$	16 ± 10	5,8	13	4,6
mARC_O T 2	$2,2 \times 10^{-6} \pm 2,4 \times 10^{-6}$	55 ± 57	$0,97 \pm 0,8$	$10,7 \pm 1$	10 ± 1	6,2	13	4,8

$k_{0(N \rightarrow D)}$: constante cinética de desplegamiento a fuerza cero, $k_{0(D \rightarrow N)}$: constante cinética de replegamiento a fuerza cero, $\Delta x_{N \rightarrow \ddagger}$ y $\Delta x_{D \rightarrow \ddagger}$: distancias al estado de transición desde el estado plegado y desplegado, respectivamente; ΔG : estabilidad al equilibrio calculado a partir de las constantes cinéticas de desplegamiento y replegamiento, $dx_{F_{1/2}}$: diferencia de extensión entre el estado plegado y desplegado, $\Delta G_{F_{1/2}}$: cambio de energía libre de Gibbs calculado a fuerza media.

Discusión

mARC se encuentra en solución en dos conformaciones, una anudada y otra no-anudada.

El objetivo de este trabajo fue determinar las consecuencias de la topología anudada sobre el mecanismo de plegamiento de las proteínas. Para ello se utilizó la proteína mARC como modelo de estudio basados en la hipótesis de que la proteína puede estar presente en una conformación anudada o no-anudada dependiendo de la posición del lazo formado por el péptido conector utilizado para generar la versión monomérica (Figura 5A y B). La posibilidad de que dos motivos RHH fusionados en la misma cadena polipeptídica formen estructuras anudadas se apoya en la evidencia de que miembros monoméricos naturales de la familia RHH, como VirC2 de *A. tumefaciens* (PDB id. 2RH3) y MJ0366 de *M. jannaschii* (PDB id 2EFV) presentan topologías anudadas en su estructura (Lu y cols. 2009, Bölinger y cols. 2010).

Los resultados presentados apoyan la hipótesis por dos observaciones principales: primero, la observación de moléculas caracterizadas por largos de contorno de 37 nm que son 6 nm más cortas de lo esperado para la molécula completamente estirada (43 nm). Segundo, la utilización de un control topológico generó exclusivamente una población indicando que la heterogeneidad de mARC no se debió a artefactos experimentales debido que pARC, una proteína con idéntica arquitectura pero imposibilitada de formar nudos, mostró un único tipo de moléculas como se desprende de los histogramas de fuerza y de extensión fraccional. Con respecto a este último gráfico, en la figura 15D se pudo observar que la mayoría de los eventos pueden describirse con una extensión fraccional de $0,6 \pm 0,2$ ($L_0 = 22 \pm 7$ nm). No obstante, la misma figura también muestra un segundo pico más pequeño centrado en 1 ($L_0 = 37$ nm). Esta segunda población no corresponde a un tipo diferente de moléculas debido a que dichas transiciones se originaron de moléculas que también presentaron transiciones de L_0 más cortos. Es decir, al estirar repetidas veces una misma molécula de pARC se pudo observar ambos tipos de transiciones, a diferencia de lo que ocurre con mARC que el tipo de transición observada dependió del tipo de molécula atrapada.

El hecho de que al estirar pARC se observen dos largos de contorno distintos para una misma molécula podría deberse a que luego del primer ciclo de desplegamiento/replegamiento la proteína quede parcialmente plegada, de esta forma en el próximo estiramiento el evento ocurrirá a fuerzas similares pero con un menor cambio de extensión.

La conformación anudada de mARC es más estable que su conformación no-anudada. En general las estabilidades reportadas para los cambios conformacionales en proteínas son diferencias de energías libres entre dos estados termodinámicos, normalmente entre el estado nativo y desplegado de una proteína. Por esta razón es muy importante definir tanto las características del estado nativo como el desplegado. Este requerimiento no es trivial para el caso de las proteínas anudadas debido a que no se tiene claridad si en el estado desplegado estas proteínas permanecen anudadas en presencia de agentes caotrópicos (King y cols 2010, Mallam & Jackson 2006, 2007; Mallam y cols 2008, 2010; Wang y cols 2013). Sin embargo nuestra aproximación no posee estos problemas debido a que nos permite controlar la topología del estado desplegado de las especies involucradas.

Los valores de ΔG calculados mediante el uso del teorema de Crooks indican que la conformación anudada de mARC es más estable ($\Delta G_1 = 21,8 \pm 0,4$ kcal/mol, Figura 17A) que su conformación no-anudada ($\Delta G_2 = 12,3 \pm 0,5$, Figura 17B). Sin embargo, como se desprende de los equilibrios planteados en la figura 3 esta diferencia de estabilidad es calculada respecto a estados desplegados con diferentes topologías. Debido a que los ΔG calculados mediante el teorema de Crooks corresponde a valores en condiciones de equilibrio, pueden ser utilizados para generar un ciclo termodinámico que nos permita calcular la diferencia de energía libre entre las demás especies involucradas (Figura 20). Por ejemplo, experimentalmente no es posible acceder a la diferencia de energía libre entre el estado desplegado anudado y desanudado ($\Delta\Delta G_{1,2}$) pero este valor podría ser calculado si conocemos la diferencia de energía libre entre la conformación nativa anudada y no anudada. Lo notable de la aproximación de moléculas individuales es que permite contar el número de moléculas que se encuentran en cada conformación y así calcular la diferencia de energía libre a partir de la ley de distribución de Boltzman (Boltzmann 1964). Debido a que se encontraron 20 moléculas que originan estados desplegados anudados y 7 moléculas que originan estados desplegados no-anudados, se calculó una diferencia de energía libre de **0,6 kcal/mol**. Este valor es cercano a cero e indica que ambas conformaciones son energéticamente muy similares entre si lo cual concuerda con la hipótesis de que lazo de mARC es una estructura flexible que puede ser interconvertida fácilmente. Utilizando este valor experimental podemos cerrar el ciclo y determinar el costo de anudar esta proteína. Este valor es **de 8,9 \pm 0,9 kcal/mol** lo que indica que el estado desplegado anudado es más inestable que el estado desplegado no-anudado. Este valor es muy elevado e indica que la formación del nudo en el estado desplegado es altamente

desfavorable. Debido a que el lazo que forma el nudo en mARC es rico en glicinas se asume que no genera contactos con el centro hidrofóbico de la proteína. Debido a esto podemos considerar que el efecto del nudo sobre el mecanismo de plegamiento de la proteína es puramente topológico. Entonces, el valor calculado para el costo de anudamiento de mARC corresponde al valor general de obligar a un polipéptido de 120 residuos a anudarse en su estado desplegado. **A nuestro conocimiento esta es la primera vez que es posible determinar experimentalmente este parámetro.** El cálculo de este valor tiene varias implicancias.

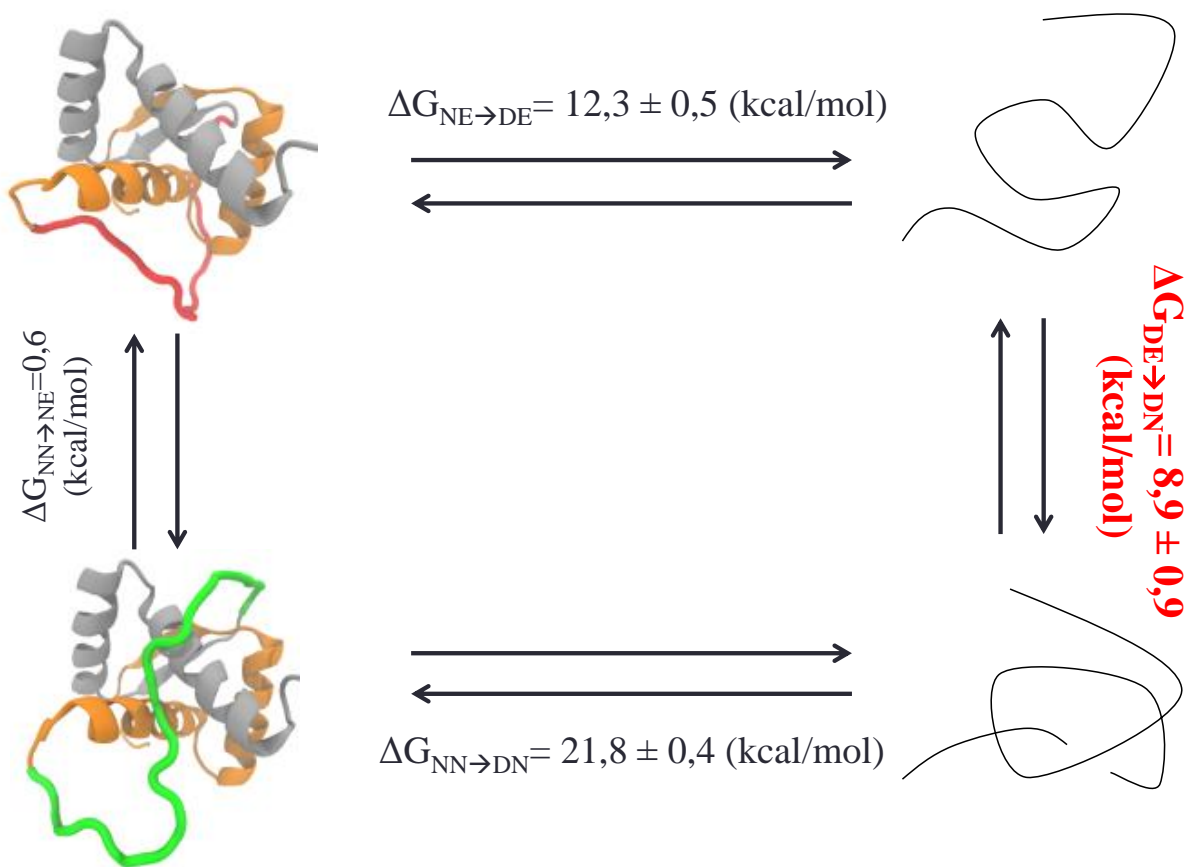


Figura 20. Ciclo termodinámico para moléculas de mARC_OT. Conociendo la abundancia relativa de cada tipo de estado nativo (anudado y no-anudado) se puede calcular el valor de ΔG asociado al equilibrio entre estos estados. Dicho valor junto con las estabilidades calculadas para el desplegamiento de cada conformación permiten calcular el costo energético de anudamiento de una proteína de 120 residuos en su estado desplegado.

Primero, simulaciones de diferentes modelos de polímeros globulares muestran que la probabilidad de formar estructuras anudadas en cadenas poliméricas grandes (de 1000 monómeros en adelante) es de alrededor del 80% (Lua y cols 2004, Tubiana y cols 2013, Virnau y cols 2005). Estos modelos son simplificaciones de la cadena polimérica en las que cada monómero corresponde a una esfera que está conectada a las demás mediante varas. Aun cuando pueden ser muy informativos, esos modelos no permiten relacionar bien el tamaño de la cadena utilizada en la simulación con el tamaño de una cadena polipeptídica real. Esta ambigüedad en los modelos puede llevar a interpretaciones erradas, por lo que el valor experimental obtenido en esta tesis puede ser muy útil para calibrar los modelos simplificados con la situación real de polímeros.

Segundo, aunque no se ha cuantificado como en nuestro caso y las evidencias son indirectas, se ha observado que en proteínas naturalmente anudadas la topología anudada se mantiene en el estado desplegado, indicando que la formación del nudo en este estado puede ser energéticamente favorable (Mallam y cols. 2010, Wang y cols. 2013). Nuestros resultados, obtenidos para el anudamiento de una proteína que no ha evolucionado para hacerlo, permitirán comparar los costos energéticos de anudamiento para proteínas naturalmente anudadas y con ello determinar qué factores influyen para hacer más favorable el proceso de anudamiento.

Tercero, bastantes trabajos teóricos y experimentales han indicado que la topología anudada puede aumentar la estabilidad de las proteínas. Sin embargo, ninguno de estos estudios ha logrado explicar la razón del dicho aumento de estabilidad (Dzubiella 2009, King y cols. 2010, Sayre y cols. 2011, Soler & Faísca 2013, Sułkowska y cols. 2008). Esto se debe a que los experimentos realizados en esta área se han hecho con metodologías que no permiten especificar a ciencia cierta los estados termodinámicos que se comparan. No obstante, la aproximación de moléculas individuales utilizada en este trabajo no tiene este problema ya que podemos controlar la topología del estado desplegado y con ello definir sin ambigüedades los estados termodinámicos que se comparan. Nuestros resultados muestran que el estado desplegado anudado de mARC es más inestable que el estado desplegado no-anudado de mARC. Esta desestabilización del estado desplegado debida al efecto topológico del nudo es la razón del incremento de la estabilidad de una proteína al anudarse.

Los estudios cinéticos fueron realizados sólo con mARC no-anudada. Los estudios cinéticos sólo se pudieron realizar en dos moléculas de mARC, y a pesar de que los parámetros cinéticos calculados para cada molécula difieren entre sí y podrían sugerir que cada una es diferente de la otra (quizás porque tengan conformaciones diferentes, anudada o no-anudada), las evidencias sugieren que ambas moléculas pertenecen al mismo tipo. En primer lugar, el cambio de extensión entre los estados plegado y desplegado a fuerza media para ambas moléculas es similar (ver tabla 5). En segundo lugar, ambas moléculas presentan eventos de desplegamiento y replegamiento a bajas fuerzas por lo que corresponden a moléculas no-anudadas de acuerdo al análisis de los experimentos de estiramiento a fuerza constante. En tercer lugar, los valores de ΔG calculados a partir de las constantes cinéticas (16 ± 10 y 10 ± 1 kcal/mol) para ambas moléculas están muy cercanos los valores de estabilidad calculados por el teorema de Crooks para las moléculas no-anudadas ($12,3 \pm 0,5$ kcal/mol) y los ΔG calculados a fuerza media también son parecidos entre ambas moléculas. En cuarto lugar, las distancias al estado de transición calculadas para ambas proteínas son similares indicando que la ruta de plegamiento de ambas moléculas es la misma o similar. Además, la baja cantidad de datos obtenidos impide asegurar que las diferencias observadas se deban a diferencias topológicas de las moléculas. Se requiere comparar una mayor cantidad de datos. Las diferencias observadas en estas moléculas pueden deberse a fluctuaciones generadas por el baño térmico lo cual es normal en este tipo de estudios (Bustamante 2008).

Estudios teóricos y resultados experimentales sugieren que la topología anudada puede disminuir la velocidad de plegamiento y/o desplegamiento de una proteína (Soler & Faísca 2013, Sułkowska y cols. 2008, Virnau y cols. 2006). Para evaluar estas posibilidades es indispensable poder comparar los parámetros cinéticos de la conformación no-anudada de mARC con los de la conformación anudada y de los de pARC. Sin embargo, no fue posible obtener datos cinéticos ni de pARC ni de la conformación anudada de mARC durante el desarrollo de este trabajo por lo que es indispensable obtener información cinética de dichas proteínas.

Una alternativa para obtener los datos cinéticos requeridos podría ser utilizar el análisis de Dudko (Dudko y cols. 2008). Este análisis permite calcular las constantes cinéticas dependientes de la fuerza a partir de los histogramas de fuerza de desplegamiento y replegamiento obtenidos en los experimentos de estiramiento a fuerza constante.

Comparación entre el mecanismo de plegamiento de mARC y pARC. Debido a que la proteína pARC mantiene la arquitectura de mARC pero con una conexión diferente de elementos de estructura secundaria es interesante realizar algunas comparaciones para revisar el efecto de la topología en cuanto a su mecanismo de plegamiento. El comportamiento de pARC es muy similar al comportamiento observado en las moléculas que presentan un L_0 esperado para una molécula no-anudada de mARC. En este contexto, se pudo observar que el promedio de fuerzas de replegamiento de pARC ($5,9 \pm 0,8$ pN) fue muy similar al promedio de fuerzas de replegamiento de mARC no-anudado ($5,3 \pm 0,6$ pN). Lo mismo ocurrió con los promedios de fuerza de desplegamiento ($8,6 \pm 2$ pN para pARC y $8,1 \pm 2$ pN para mARC no-anudada). No obstante, aparte de la imposibilidad de formar nudos por parte de pARC, esta proteína se escogió debido a que pARC presenta la misma estabilidad termodinámica que mARC reportada en experimentos realizados en presencia de agentes caotrópicos (Tabtiang et al. 2005). En consecuencia, también se esperaba que ambas proteínas presentaran la misma estabilidad calculada mediante el teorema de Crooks.

Sin embargo, la estabilidad más baja calculada para mARC mediante mediaciones de moléculas individual es $12,3 \pm 0,5$ kcal/mol. Este valor es $5,9$ kcal/mol mayor que la reportada en experimentos con agentes caotrópicos y $5,3$ kcal/mol mayor al valor de pARC obtenido en este estudio mediante pinzas ópticas. Dicho de otra manera, el valor de estabilidad calculado mediante moléculas individuales para pARC es más cercano al reportado en experimentos con agentes caotrópicos que el obtenido para mARC. Se especula que esta diferencia puede deberse a dos factores.

Se ha demostrado que al desplegar térmica o químicamente una proteína esta puede mantener una estructura residual (Aune & Salahuddin 1967, Neri y cols. 1992, Robic y cols. 2003, Russell y cols. 2000, Smith y cols. 1996a, Yi y cols. 2000) cuya presencia puede no ser considerada al determinar la diferencia de energía mediante agentes caotrópicos tales como urea o cloruro de guanidinio. En cambio, en la aproximación de moléculas individuales de pinzas ópticas la destrucción de la estructura residual debería estar contenida en los cálculos energéticos debido al estiramiento de la molécula. En nuestro caso sabemos que las proteínas están completamente estiradas (desplegadas) porque los largos de contorno calculados corresponden a la extensión completa de la proteína de 120 residuos (43 nm) en el caso de la conformación no-anudada y 103 residuos (37 nm) para la conformación anudada. De esta manera los valores de estabilidad calculados mediante métodos de desplegamiento mecánico

podrían sobrestimar el valor obtenido en solución debido a que estos últimos no consideran el componente energético asociado al desplegamiento de la estructura residual. Basándose en esta hipótesis, y considerando que la estabilidad de pARC a través del teorema de Crooks ($7 \pm 0,4$ kcal/mol) es cercana al valor determinado en experimentos realizados con agentes caotrópicos (5,9 kcal/mol, Tabtiang y cols. 2005) se predice un mayor contenido de estructura residual para mARC. Esta diferencia es interesante, ya que pARC es una permutante lo cual podría afectar, de alguna manera la cantidad de estructura residual respecto a mARC. Se requieren estudios de NMR con pARC y mARC para comprobar las suposiciones sobre las diferencias en estructura residual originadas por un cambio topológico.

Otra posibilidad que puede explicar las diferencias observadas entre los ΔG obtenidos por el teorema de Crooks y los obtenidos por experimentos hechos con agentes caotrópicos es el error asociado al cálculo del trabajo de desplegamiento y replegamiento a partir de las curvas de fuerza vs extensión obtenidas de los experimentos de estiramiento a velocidad constante. Existen antecedentes que indican que dicho error podría ser de hasta un 100% (Mossa y cols. 2009) dependiendo del método utilizado para integrar el trabajo a partir de la curva de fuerza vs extensión. Considerando esto, es posible que los datos de estabilidad calculados en este trabajo estén alterados debido a que el método utilizado para el cálculo del trabajo no ha sido probado respecto a su exactitud.

Conclusiones

La proteína mARC se encuentra en solución en dos conformaciones, una anudada con largo de contorno de 37 ± 7 nm y una no-anudada con un largo de contorno de 43 ± 7 nm.

La conformación anudada de mARC es más estable que la conformación no-anudada. Esto se debe a que el estado desplegado anudado de la proteína es más inestable que su estado desplegado no-anudado.

La diferencia de energía ($\Delta\Delta G$) ente el estado desplegado no-anudado y el estado desplegado anudado de mARC corresponde al costo de anudamiento de una cadena polipeptídica de 120 residuos en su estado desplegado y tiene un valor de $8,9 \pm 0,9$ kcal/mol.

Referencias

- Aune K, Salahuddin A. 1967. Evidence for residual structure in acid-and heat-denatured proteins. *J. Biol. Chem.* 242(19):4486–89
- Baker D. 2000. A surprising simplicity to protein folding. *Nature.* 405(6782):39–42
- Bell GI, Dembo M, Bongrand P. 1984. Cell adhesion. competition between nonspecific repulsion and specific bonding. *Biophys. J.* 45(6):1051–64
- Bölinger D, Sułkowska JI, Hsu H-P, Mirny LA, Kardar M, et al. 2010. A stevedore's protein knot. *PLoS Comput. Biol.* 6(4):6
- Boltzmann L. 1964. *Vorlesungen über Gastheorie.* University of California Press
- Bouchiat C, Wang MD, Allemand J, Strick T, Block SM, Croquette V. 1999. Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys. J.* 76(1 Pt 1):409–13
- Brown BM, Sauer RT. 1999. Tolerance of arc repressor to multiple-alanine substitutions. *PNAS.* 96(5):1983–88
- Bustamante C. 2008. In singulo biochemistry: when less is more. *Annu. Rev. Biochem.* 77:45–50
- Bustamante C, Liphardt J, Ritort F. 2005. The nonequilibrium thermodynamics of small systems. *Phys. Today.* 58(7):43
- Bustamante C, Marko J, Siggia E, Smith S. 1994. Entropic elasticity of lambda-phage dna. *Science.* 265(5178):1599–1600
- Bustamante C, Smith S, Liphardt J, Smith D. 2000. Single-molecule studies of dna mechanics. *Curr. Opin. Struct. Biol.* 10(3):279–85
- Cecconi C, Shank E a, Bustamante C, Marqusee S. 2005. Direct observation of the three-state folding of a single protein molecule. *Science.* 309(5743):2057–60
- Cecconi C, Shank EA, Dahlquist FW, Marqusee S, Bustamante C. 2008. Protein-dan chimeras for single molecule mechanical folding studies with optical tweezers. *Eur. Biophys. J.* 37(6):729–38
- Crooks G. 2000. Path-ensemble averages in systems driven far from equilibrium. *Phys. Rev. Lett.* 61(3):2361–66
- Dill K a, Bromberg S, Yue K, Fiebig KM, Yee DP, et al. 1995. Principles of protein folding: a perspective from simple exact models. *Protein Sci.* 4(4):561–602
- Dudko OK, Hummer G, Szabo A. 2008. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *PNAS.* 105(41):15755–60

- Dzubiella J. 2009. Sequence-specific size, structure, and stability of tight protein knots. *Biophys. J.* 96(3):831–39
- Fersht A. 1997. Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.* 7(1):3–9
- Garcia-Manyes S, Brujić J, Badilla CL, Fernández JM. 2007. Force-clamp spectroscopy of single-protein monomers reveals the individual unfolding and folding pathways of i27 and ubiquitin. *Biophys. J.* 93(7):2436–46
- Grassetti DR, Murray JF. 1967. Determination of sulfhydryl groups with 2,2'- or 4,4'-dithiodipyridine. *Arch. Biochem. Biophys.* 119(1):41–49
- Hills RD, Brooks CL. 2009. Insights from coarse-grained gō models for protein folding and dynamics. *Int. J. Mol. Sci.* 10(3):889–905
- Khatib F, Weirauch MT, Rohl C a. 2006. Rapid knot detection and application to protein structure prediction. *Bioinformatics.* 22(14):e252–9
- King NP, Jacobitz AW, Sawaya MR, Goldschmidt L, Yeates TO. 2010. Structure and folding of a designed knotted protein. *PNAS.* 107(48):20732–37
- Koga N, Takada S. 2001. Roles of native topology and chain-length scaling in protein folding: a simulation study with a go-like model. *J. Mol. Biol.* 313(1):171–80
- Kolesov G, Virnau P, Kardar M, Mirny LA. 2007. Protein knot server: detection of knots in protein structures. *Nucleic Acids Res.* 35(Web Server issue):W425–W428
- Kouza M, Li MS, O'Brien EP, Hu C-K, Thirumalai D. 2006. Effect of finite size on cooperativity and rates of protein folding. *J. Phys. Chem. A.* 110(2):17
- Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage t4. *Nature.* 227(5259):680–85
- Liphardt J, Onoa B, Smith SB, Tinoco I, Bustamante C. 2001. Reversible unfolding of single rna molecules by mechanical force. *Science.* 292(5517):733–37
- Lua R, Borovinskiy A, Grosberg A. 2004. Fractal and statistical properties of large compact polymers: a computational study. *Polymer (Guildf).* 45(2):717–35
- Lua RC, Grosberg AY. 2006. Statistics of knots, geometry of conformations, and evolution of proteins. *PLoS Comput. Biol.* 2(5):e45
- Mallam AL. 2009. How does a knotted protein fold? *FEBS J.* 276(2):365–75
- Mallam AL, Jackson SE. 2006. Probing nature's knots: the folding pathway of a knotted homodimeric protein. *J. Mol. Biol.* 359(5):1420–36
- Mallam AL, Jackson SE. 2007. A comparison of the folding of two knotted proteins: ybea and yibk. *J. Mol. Biol.* 366(2):650–65

- Mallam AL, Onuoha SC, Grossmann JG, Jackson SE. 2008. Knotted fusion proteins reveal unexpected possibilities in protein folding. *Mol. Cell.* 30(5):642–48
- Mallam AL, Rogers JM, Jackson SE. 2010. Experimental detection of knotted conformations in denatured proteins. *PNAS.* 107(18):8189–94
- Mansfield M. 1994. Are there knots in proteins? *Nat. Struct. Mol. Biol.*
- McKinney S a S, Joo C, Ha T. 2006. Analysis of single-molecule fret trajectories using hidden markov modeling. *Biophys. J.* 91(5):1941–51
- Mossa A, de Lorenzo S, Huguet JM, Ritort F. 2009. Measurement of work in single-molecule pulling experiments. *J. Chem. Phys.* 130(23):234116
- Murzin a G, Brenner SE, Hubbard T, Chothia C. 1995. Scop: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247(4):536–40
- Neri D, Billeter M, Wider G, Wüthrich K. 1992. Nmr determination of residual structure in a urea-denatured protein, the 434-repressor. *Science.* 257:1559–15563
- Onuchic JN, Luthey-Schulten Z, Wolynes PG. 1997. Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* 48(1):545–600
- Onuchic JN, Wolynes PG. 2004. Theory of protein folding. *Curr. Opin. Struct. Biol.* 14(1):70–75
- Ozkan SB, Bahar I, Dill KA. 2001. Transition states and the meaning of ϕ -values in protein folding kinetics. *Nat. Struct. Mol. Biol.* 8(9):16–18
- Paci E, Vendruscolo M, Karplus M. 2002. Validity of g_{ϕ} models: comparison with a solvent-shielded empirical energy decomposition. *Biophys. J.* 83(6):3032–38
- Plaxco KW, Simons KT, Baker D. 1998. Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* 277(4):985–94
- Raymer DM, Smith DE. 2007. Spontaneous knotting of an agitated string. *PNAS.* 104(42):16432–37
- Rief M. 1997. Reversible unfolding of individual titin immunoglobulin domains by afm. *Science.* 276(5315):1109–12
- Robic S, Guzman-Casado M, Sanchez-Ruiz JM, Marqusee S. 2003. Role of residual structure in the unfolded state of a thermophilic protein. *PNAS.* 100(20):11345–49
- Robinson C, Sauer RT. 1996a. Equilibrium stability and sub-millisecond refolding of a designed single-chain arc repressor. *Biochemistry.* 2960(96):13878–84
- Robinson CR, Sauer RT. 1996b. Covalent attachment of arc repressor subunits by a peptide linker enhances affinity for operator dna. *Biochemistry.* 35(1):109–16
- Robinson CR, Sauer RT. 2000. Striking stabilization of arc repressor by an engineered disulfide bond. *Biochemistry.* 39(40):12494–502

- Russell BS, Melenkivitz R, Bren KL. 2000. Nmr investigation of ferricytochrome c unfolding: detection of an equilibrium unfolding intermediate and residual structure in the denatured state. *PNAS*. 97(15):8312–17
- Sayre TC, Lee TM, King NP, Yeates TO. 2011. Protein stabilization in a highly knotted protein polymer. *Protein Eng. Des. Sel.* 24(8):627–30
- Schlierf M, Li H, Fernandez JM. 2004. The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *PNAS*. 101(19):7299–7304
- Schreiter ER, Drennan CL. 2007. Ribbon-helix-helix transcription factors: variations on a theme. *Nat. Rev. Microbiol.* 5(9):710–20
- Shank E a, Cecconi C, Dill JW, Marqusee S, Bustamante C. 2010. The folding cooperativity of a protein is controlled by its chain topology. *Nature*. 465(7298):637–40
- Smith LJ, Fiebig KM, Schwalbe H, Dobson CM. 1996a. The concept of a random coil. residual structure in peptides and denatured proteins. *Fold. Des.* 1(5):R95–106
- Smith S, Cui Y, Bustamante C. 2003. Optical-trap force transducer that operates by direct measurement of light momentum. *Methods Enzymol.* 361(1994):134–62
- Smith SB, Cui Y, Bustamante C. 1996b. B-dna : the elastic response of overstretching individual double-stranded and single-stranded. *Science*. 271(5250):795–99
- Soler M a, Faísca PFN. 2013. Effects of knots on protein folding properties. *PLoS One*. 8(9):e74755
- Sulkowska JI, Rawdon EJ, Millett KC, Onuchic JN, Stasiak A. 2012. Conservation of complex knotting and slipknotting patterns in proteins. *PNAS*. 109(26):E1715–23
- Sulkowska JI, Sulkowski P, Szymczak P, Cieplak M. 2008. Stabilizing effect of knots on proteins. *PNAS*. 105(50):19714–19
- Tabtiang RK, Cezairliyan BO, Grant R a, Cochrane JC, Sauer RT. 2005. Consolidating critical binding determinants by noncyclic rearrangement of protein secondary structure. *PNAS*. 102(7):2305–9
- Taylor WR. 2000. A deeply knotted protein structure and how it might fold. *Nature*. 406(6798):916–19
- Tubiana L, Rosa A, Fragiaco F, Micheletti C. 2013. Spontaneous knotting and unknotting of flexible linear polymers: equilibrium and kinetic aspects. *Macromolecules*. 2(1):1–14
- Virnao P, Kantor Y, Kardar M. 2005. Knots in globule and coil phases of a model polyethylene. *J. Am. Chem. Soc.* 127(43):15102–6
- Virnao P, Mirny L a, Kardar M. 2006. Intricate knots in proteins: function and evolution. *PLoS Comput. Biol.* 2(9):e122
- Waldburger C, Jonsson T. 1996. Barriers to protein folding: formation of buried polar interactions is a slow step in acquisition of structure. *PNAS*. 93(April):2629–34

- Wang P, Yang L, Liu P, Gao YQ, Zhao XS. 2013. Single-molecule detection reveals knot sliding in trnd denaturation. *Chemistry*. 19(19):5909–16
- Yi Q, Scalley-Kim ML, Alm EJ, Baker D. 2000. Nmr characterization of residual structure in the denatured state of protein l. *J. Mol. Biol.* 299(5):1341–51
- Zinn-Justin P, Zuber JB. 2010. Knot theory and matrix integrals. In *Random Matrix Theory*. 9904019(May):1–22