

Universidad de Sevilla  
Facultad de Biología  
Departamento de Bioquímica Vegetal y Biología Molecular

# **Nanomecánica de la $\beta$ -catenina e interacciones proteína-proteína**

Memoria presentada por **Andrés Manuel Vera Gómez**  
para optar al grado de Doctor

El presente trabajo de tesis doctoral ha sido desarrollado en el  
Departamento de Neurobiología Molecular, Celular y del Desarrollo del Instituto  
Cajal (IC-CSIC), bajo la dirección del  
**Dr. Mariano Sixto Carrión Vázquez.**

*Madrid, Marzo 2015*



Memoria presentada por **ANDRÉS MANUEL VERA GÓMEZ**, licenciado en Biología, para optar al grado de DOCTOR por la UNIVERSIDAD DE SEVILLA.

**MARIANO SIXTO CARRIÓN VÁZQUEZ, Doctor en Biología y Científico Titular del CSIC,**

**AUTORIZA** la presentación de la Tesis Doctoral titulada “**Nanomecánica de la  $\beta$ -catenina e interacciones proteína-proteína**”, de la que es autor el Lcdo. Andrés Manuel Vera Gómez y ha sido realizada bajo mi dirección en el Instituto Cajal del CSIC, Madrid.

Y para que así conste a los efectos oportunos, firmo el presente certificado en Madrid, el día 16 de Marzo de 2015.

Fdo.: Mariano Sixto Carrión Vázquez

Instituto Cajal, CSIC

Madrid

## Agradecimientos

En primer lugar debo agradecer de manera especial a mi director de tesis, Mariano Carrión Vázquez, por darme la oportunidad de hacer la tesis en un tema tan apasionante como es el de la nanomecánica de proteínas, y en especial por confiar en mí y apoyarme todos estos años. Gracias también por todas las cosas que he aprendido de él, sobre todo por enseñarme (no sin mucho sufrimiento de su parte) que es mucho más productivo discutir tus ideas con otros, a quedarte con tus ideas originales por muy seguro que estés de ellas.

Mi agradecimiento también a nuestros colaboradores del Instituto de Química-Física Rocasolano, Margarita Menéndez y Douglas Laurents, por su inestimable ayuda y su asesoramiento en el procesamiento y discusión de los datos. En especial a Douglas, por su paciencia a la hora de explicar todo, y su manera de contagiar su pasión por el trabajo que hace.

Debo agradecer igualmente a los profesores de la Universidad de Sevilla, Juan Arroyo Marín, M<sup>a</sup> del Rosario Espuny Gómez, y en especial a Fernando Publio Molina Heredia, por su asesoramiento, buena disposición, y su ayuda a la hora de lidiar con los “papeleos” de la tesis.

A mis compañeros y amigos del laboratorio y zonas cercanas: al Cabezón, el rulo, Chirri, Bortttt, ojkirr-ojkirr!!, Ànyal, se Pope, Loren, H, Anahimmm, Rosauron, Isa, Manowar, Pacho, la-mary, Bartoll, Laura, a los 20 millones de estudiantes y demás... a todos ellos, por los buenos y los malos momentos, por las risas, las cervezas calientes, el Canasta y el dosMa.

Un especial agradecimiento de todo corazón a Ana, sin tu apoyo todos estos años jamás hubiera podido salir delante de toda esta pesadilla. Mil gracias por su comprensión, y por haberme regalado los mejores momentos de todos estos años.

Un gran gracias a mis amigos de La Línea, en especial a Diego y a José Mari, por las risas y las bromas, por habernos hecho adultos juntos, y porque siempre me hacéis ver las cosas en perspectiva. Espero seguir siendo vuestro amigo otros 1000 años.

Finalmente, el agradecimiento más especial a mi familia, a mis padres y mi hermano Roberto. Por su cariño y amor, y en especial a mis padres por haberme hecho la persona que soy, por enseñarme el valor y la fuerza de la perseverancia. Nunca podría haber llegado a ser la persona que soy sin vosotros. Gracias

## Resumen

Las fuerzas mecánicas juegan un papel fundamental en los sistemas biológicos. Las proteínas y las interacciones intermoleculares entre proteínas sufren estrés mecánico en procesos celulares tan cotidianos como el transporte a través de membrana, los procesos de adhesión celular, la segregación cromosómica o la replicación del ADN. La fuerza es una magnitud vectorial, por lo que la dirección de su aplicación es difícil, si no imposible de controlar mediante experimentos de bioquímica clásica. A nivel de moléculas individuales, la magnitud de estas fuerzas es tan pequeña que ha sido necesario el desarrollo de sofisticadas técnicas experimentales para medirlas y controlarlas. Entre ellas destaca la espectroscopia de fuerzas de molécula individual basada en microscopía de fuerza atómica (AFM-SMFS), que ha permitido un avance sin precedentes en el análisis molecular de las propiedades mecánicas de proteínas y en el estudio directo de las fuerzas que mantienen las interacciones proteína-proteína.

En los metazoos, el sistema cadherina/catenina constituye una pieza clave en la maquinaria de adhesión célula-célula, que no sólo media los contactos directos entre células vecinas, sino que constituye un elemento esencial del proceso de mecanotransducción que se lleva a cabo en estas zonas de contacto. Las cadherinas en la superficie de las células interactúan con las cadherinas de las células vecinas, y simultáneamente mantienen una conexión mecano-funcional con el citoesqueleto contráctil de actomiosina mediante un complejo de proteínas citosólicas. Entre ellas se encuentra la  $\beta$ -catenina, esencial para la mediación de la continuidad mecánica que existe entre el dominio extracelular de las cadherinas y el esqueleto de actomiosina, y en el proceso de mecanotransducción que llevan a cabo estas estructuras. Así pues, conocer las propiedades mecánicas de esta proteína parece clave para poder empezar a entender el funcionamiento interno del sistema cadherina/catenina. Por todo ello, en este trabajo se ha llevado a cabo el estudio de las propiedades nanomecánicas de la  $\beta$ -catenina mediante AFM-SMFS. Nuestros resultados muestran que la región ARM (armadillo) de la  $\beta$ -catenina presenta una baja estabilidad mecánica y sugieren que sus propiedades mecánicas son muy sensibles a pequeños cambios en la estructura. El desplegamiento mecánico de esta estructura puede ocurrir a través de múltiples caminos, con diferentes intermedios, lo que indica que el proceso tiene lugar a través de un paisaje energético rugoso con mínimos energéticos muy superficiales. Este desplegamiento mecánico puede ser modulado por sus terminales desestructurados. Considerando que la fuerza media de desplegamiento de la  $\beta$ -catenina está en el rango de fuerzas que soportan las adhesiones cadherina-cadherina y que algunas de las repeticiones de la región ARM pueden desplegarse a baja fuerza (por



debajo de la resolución del aparato de medida), la  $\beta$ -catenina podría actuar como un tampón mecánico a lo largo de todo el rango de fuerzas al que podría estar sometida, “sacrificando” algunas de sus repeticiones para mantener los contactos entre las uniones cadherinas y el citoesqueleto de actina.

Debido a su sensibilidad en torno a los pocos piconewtons y su precisión subnanométrica, el AFM se ha convertido en una herramienta esencial no sólo en la medida de las propiedades mecánicas de proteínas y en la energética y dinámica de sus interacciones intramoleculares, sino que también se ha impuesto como una poderosa herramienta en la exploración de las fuerzas y la dinámica de las interacciones entre proteínas. La estrategia general en estos estudios reside en unir una de las proteínas a la punta del AFM y la otra a la superficie de un sustrato, midiendo la fuerza de desunión provocada al alejar la superficie de la punta tras establecer contacto. Sin embargo, esta estrategia hace imposible distinguir el evento de ruptura de la interacción de los eventos de despegado de las proteínas de cualquiera de las superficies, o de interacciones inespecíficas punta-superficie. En el trabajo que aquí se presenta proponemos una estrategia autocontrolada para la medida de interacciones intermoleculares, que permita la identificación directa del pico de fuerza provocado por la ruptura de una interacción proteína-proteína. El diseño está basado en una estrategia general y modular, que no depende de las características o la naturaleza de las proteínas estudiadas, y que descansa en la incorporación de dos tipos de marcadores: marcadores de monomolecularidad a las proteínas de interés (basados en poliproteínas); y un marcador específico de la interacción (la exposición a la fuerza de un polipéptido elastomérico que une de forma covalente las dos proteínas interaccionantes). Como demostración a nivel de prueba de concepto de que este diseño es posible, se ha implementado con éxito esta estrategia en una interacción modelo, la interacción del par cohesina/dockerina de *Clostridium thermocellum*.



3.10.- Experimentos de AFM-SMFS .....	53
4. Resultados .....	56
4.1.-Nanomecánica de la $\beta$ -Catenina.....	56
4.1.1.- Mecánica de la región ARM de la $\beta$ -catenina.....	57
4.1.2.- Mecánica de la $\beta$ -catenina completa .....	60
4.1.3.- Mecánica de la $\beta$ -catenina completa sin controles de monomolecularidad.....	62
4.1.4.- Controles estructurales y de unión a ligando.....	65
4.1.5.- Simulaciones de dinámica molecular de la $\beta$ -catenina.....	68
4.2.- Estrategia autocontrolada para el estudio de interacciones intermoleculares por AFM-SMFS. ....	71
4.2.1.- Diseño experimental: elección de los elementos del sistema. ....	71
4.2.2.- Diseño experimental: diseño de dominios de oligomerización. .	73
4.2.3.- Diseño experimental: generación de una batería de vectores.....	76
4.2.4.- Testado de los elementos del sistema.....	78
4.2.5.- La interacción cohesina/dockerina origina un pico de fuerza con un $\Delta L_c$ característico. ....	97
4.2.6.- Exploración de diferentes vías para aumentar el rendimiento en los experimentos de AFM-SMFS.....	101
5. Discusión .....	108
5.1.- Nanomecánica de la $\beta$ -Catenina.....	108
5.2.- Estrategia autocontrolada para el estudio de interacciones intermoleculares por AFM-SMFS. ....	111
6. Conclusiones.....	117
7. Bibliografía.....	119
8. Publicaciones.....	131

## Clave de abreviaturas y siglas

---

- $\beta$ /I27: construcción recombinante formada por la  $\beta$ -catenina humana flanqueada por módulos I27.
- $\beta$ -cat: proteína  $\beta$ -catenina.
- $\Delta L_c$ : incremento de longitud de contorno (*Contour Length Increase*).
- AFM: microscopio de fuerza atómica o microscopía de fuerzas atómicas (*Atomic Force Microscope/Microscopy*)
- AFM-SMFS: espectroscopía de fuerzas de molécula individual basada en AFM.
- Agua MilliQ: agua destilada purificada mediante un intercambio iónico (cartucho) y ultrafiltración (filtro de membrana). La pureza del agua se mide mediante su conductividad, el inverso de la resistividad, que debe ser mayor de 18.2 M $\Omega$ /cm (a 25 °C).
- DC: dicroísmo circular (*Circular Dichroism*).
- ARM: región armadillo de la proteína  $\beta$ -catenina.
- ARM/I27: construcción recombinante formada por la región ARM de la  $\beta$ -catenina humana flanqueada por módulos I27.
- Cistamina: 2,2'-Ditiobis(etilamina). Pequeña molécula portadora de enlaces disulfuro, utilizada como compuesto oxidante en los tampones redox.
- Cisteamina: 2-Aminoetanetiol. Pequeña molécula portadora de un grupo sulfhidrilo, utilizada como compuesto reductor en los tampones redox.
- Co<sup>2+</sup>: catión cobalto (2+).
- Cohesina: Cohesina 7 de la escafoldina CipA de *Clostridium thermocellum*.
- Cu<sup>2+</sup>: catión cobre (2+).
- Dockerina: Dockerina de la celulasa Cel8A de *C. thermocellum*.
- DTT: ditioneitol.
- E. coli*: *Escherichia coli*.
- EDTA: ácido etilen-diamino-tetraacético (*EthyleneDiamine Tetraacetic Acid*).
- F<sub>D</sub>: fuerza de desplegamiento/disociación (*unfolding/unbinding Force*).
- Gb1: Dominio B1 de unión a IgG de la proteína G estreptocócica.
- I27: módulo tipo inmunoglobulina de titina cardiaca humana, siguiendo la nomenclatura de Labeit y Kolmerer.
- IDA: ácido iminodiacético
- IPTG: isopropyl  $\beta$ -D-1-tiogalactopiranosido (*IsoPropyl  $\beta$ -D-1 ThioGalactopyranoside*).
- MCS: sitio de clonaje múltiple (*MultiCloning Site*).
- MOPS: ácido 4-morfolino propano sulfónico (*4-MorpholinePropaneSulfonic acid*).
- Ni<sup>2+</sup>: catión níquel (2+).
- NMR: resonancia magnética nuclear (*Nuclear Magnetic Resonance*).
- NTA: ácido nitrilo triacético (*Nitrilo Triacetic Acid*).
- kD : kiloDalton.
- OD: densidad óptica (*optical density*).
- o/n: durante la noche (*over night*).
- PAGE: electroforesis en gel de poli-acrilamida (*PolyAcrylamide Gel Electrophoresis*). Si se realiza en condiciones desnaturalizantes mediante la adición de SDS, se denomina SDS-PAGE.
- PBS: tampón fosfato salino (*Phosphate Buffer Saline*).
- PCR: reacción en cadena de la polimerasa (*Polymerase Chain Reaction*).
- PDB: banco de datos de estructuras de proteínas (*Protein Data Bank*).
- SDS: dodecil sulfato sódico (*Sodium Dodecyl Sulfate*).
- SDS-PAGE: ver PAGE.
- SM: molécula individual (*single molecule*).
- SMFS: espectroscopía de fuerzas de molécula individual (*Single-Molecule Force Spectroscopy*).

Tampón Redox: tampón con moléculas portadoras de grupos tiol y puentes disulfuro que se usan para mediar reacciones de intercambio de tioles.

Term : terminal

Tris: Tris(hidroximetil)aminometano.

Ubiquitina: repetición de la proteína ubiquitina.

UV: UltraVioleta.

VMD: programa de visualización de dinámica molecular (*Visual Molecular Dynamics*).

WLC: modelo de física de polímeros de la cadena gusaniforme (*Worm-Like Chain*).

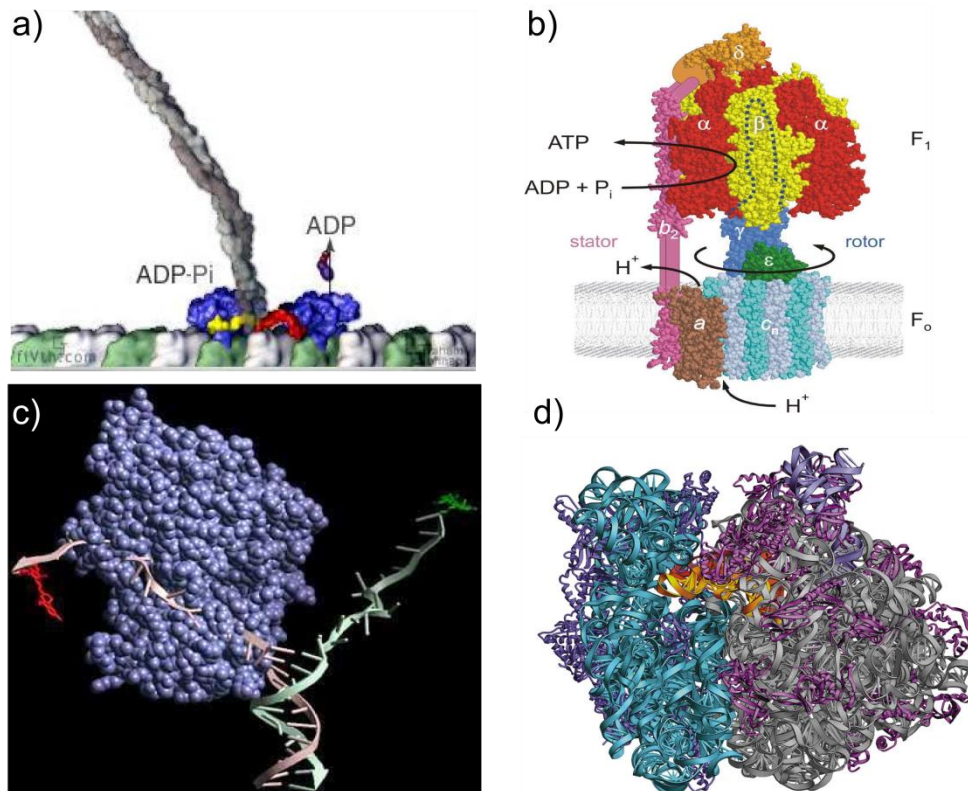
# 1. Introducción

---

## 1.1.- Fuerzas, Biomoléculas y Nanomáquinas.

Los procesos mecánicos están involucrados en casi todas las facetas del ciclo celular. Las denominadas fuerzas mecánicas son generadas y percibidas en la célula durante procesos tan diversos como la segregación de cromosomas, replicación, transcripción, traducción, translocación de proteínas a través de membranas, adhesión celular, motilidad celular, transporte de orgánulos, fusión de membranas, plegamiento y desplegamiento catalizados de proteínas y ácidos nucleicos (Bustamante et al., 2004). En la actualidad la célula es contemplada como una “pequeña factoría” donde pequeñas y complejas estructuras moleculares o nanomáquinas llevan a cabo tareas altamente especializadas de manera coordinada. Muchas de estas bionanomáquinas están formadas por proteínas, compuestas bien por una sola cadena polipeptídica o por complejos moleculares mayores, y a menudo su función implica la conversión de energía química en trabajo mecánico (**figura 1.1**). Así pues, desde el núcleo hasta la matriz extracelular gran parte de esta maquinaria celular está sometida a fuerza durante su funcionamiento fisiológico (Bustamante et al., 2004; Carrión-Vázquez et al., 2006).

Desde un punto de vista químico las proteínas son polímeros lineales de aminoácidos, que en solución, forman estructuras localmente ordenadas (estructuras secundarias) que se organizan espacialmente entre sí para originar estructuras tridimensionales concretas, o plegamientos, que conforman lo que se conoce como estructura nativa de la proteína (Cantor y Schimmel, 1980) (una excepción a este esquema lo representan las proteínas intrínsecamente desordenadas). Para comprender la actividad de las proteínas, es esencial saber cómo estas nanomáquinas mueven sus partes y cambian de forma en respuesta a fuerzas mecánicas y térmicas que se presentan en su nanoambiente. Estas estructuras nativas, ya estén formadas por una sola cadena polipeptídica (estructura terciaria), o por varias de ellas (estructura cuaternaria), son el resultado de muchas interacciones débiles que se suman para originar una estructura estable. Bajo condiciones fisiológicas, todas las interacciones relevantes (electroestáticas, puentes de hidrógeno, de Van der Waals, e interacciones hidrofóbicas), son de corto rango ( $\sim 1$  nm), y resultan en entalpías de unión del orden de la energía térmica  $k_B T$  (4.1 pN·nm), lo que genera una estabilidad global que oscila entre los 8-25  $k_B T$  (Carrión-Vázquez et al., 2006; Merkel, 2001). Estas fuerzas son de la misma naturaleza que las que estabilizan la unión entre pares interactuantes proteína-proteína. Teniendo en cuenta que las energías



**Figura 1.1.- Ejemplos de algunas nanomáquinas. a)** Las quinesinas caminan a lo largo de los microtúbulos transportando vesículas, utilizando para ello la hidrólisis del ATP, transformando la energía química en mecánica. **b)** La F<sub>1</sub>-ATP sintasa es la encargada de la síntesis de ATP en las mitocondrias, actuando como un motor rotatorio que utiliza la energía almacenada en un gradiente de protones para la síntesis de ATP. **c)** Las helicinas se encargan del desplegado mecánico de la doble hélice de ADN en dos hebras de cadena sencilla. **d)** Detalle del ribosoma 70S de procariotas. Esta nanomáquina está constituida por proteínas y moléculas de ARN.

involucradas en cambios conformacionales proteicos están entre los  $1-25 k_B T$ , las proteínas pueden contemplarse como estructuras “plásticas” capaces de sufrir perturbaciones durante su actividad fisiológica mientras mantienen su estabilidad global.

Teniendo en cuenta que los movimientos implicados en los cambios conformacionales proteicos se miden de angstrom a nanómetros, y que las energías involucradas oscilan del  $1-25 k_B T$ , las fuerzas esperadas en estos procesos se esperan en el rango de los picoNewtons. De hecho, experimentalmente se han medido las fuerzas generadas por algunos motores moleculares como el de la ARN polimerasa de *Escherichia coli*, la proteasa ClpXP, la proteína portal del bacteriófago  $\Phi 29$  o el motor de los *pili* bacterianos, que pueden ejercer fuerzas de 25, 33, 60 y 110 pN respectivamente (Wang et al., 1998; Aubin-Tam et al., 2011; Smith et al., 2001; Maier et al., 2002).

Muchas proteínas están implicadas en procesos fisiológicos complejos en los que la fuerza juega un papel primordial. En ellos las fuerzas son generadas, detectadas, transmitidas y resistidas por los tejidos y células, implicando a proteínas que en la mayoría de los casos se encuentran interaccionando las unas con las otras. Por todo ello, el

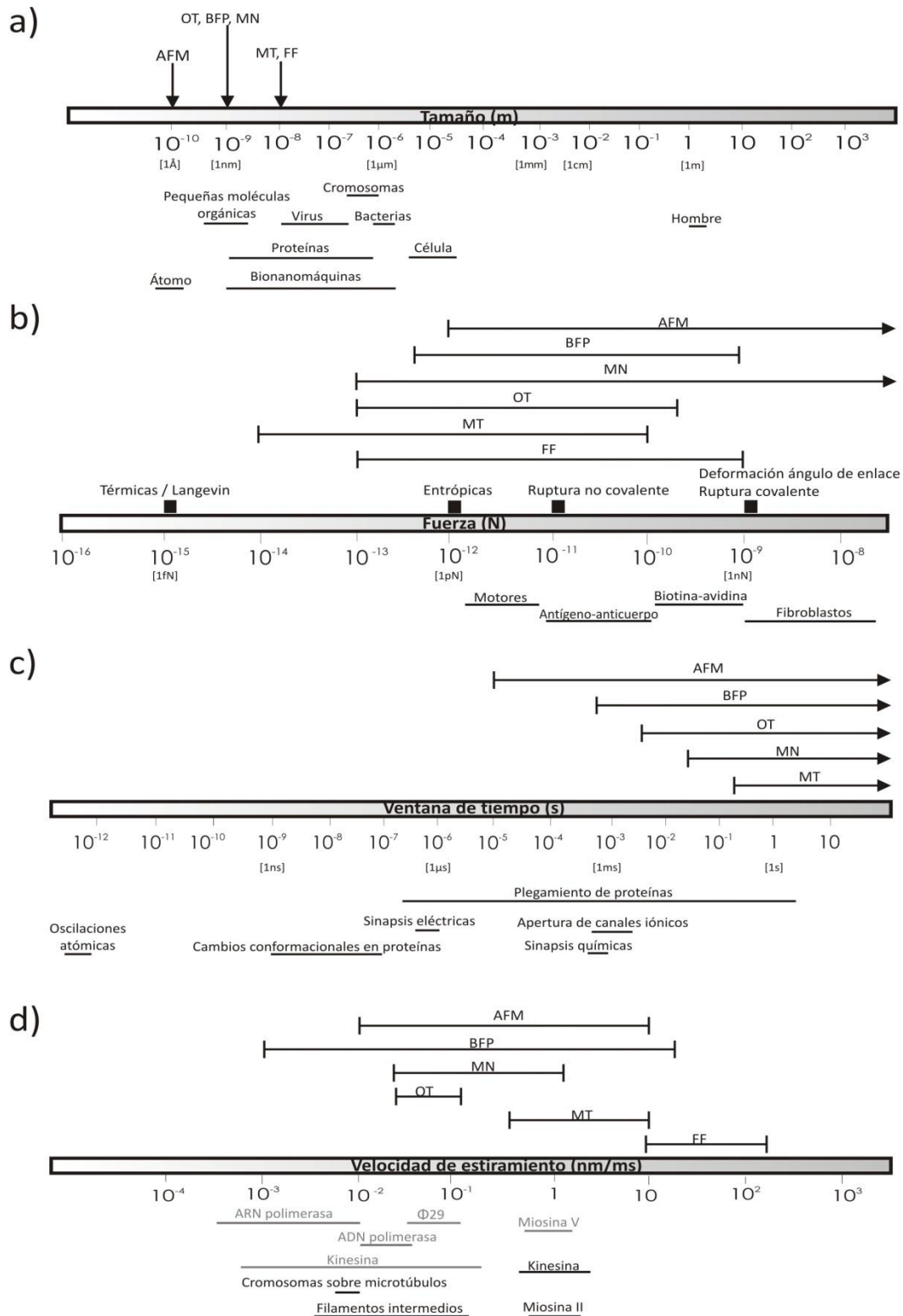
estudio de la respuesta de estas proteínas y de sus interacciones ante la fuerza parece de vital importancia para comprender en profundidad estos procesos. Por ejemplo, la proteína titina es la responsable de la elasticidad pasiva en el sarcómero de los músculos esquelético y cardíaco, y está involucrada en detección de la tensión mecánica en el mismo (Linke y Grutzner, 2008; Bertz et al., 2009). Fibronectina y tenascina son dos proteínas de la matriz extracelular que se unen a gran variedad de componentes de la matriz extracelular y receptores celulares, generando una red de interacciones que provee una resistencia y elasticidad adecuada a los tejidos, y que está relacionada con la migración y la adhesión celulares (Oberhauser et al., 1998; Oberhauser et al., 2002; Jollymore et al., 2009). Los sistemas de adhesión célula-célula o célula-matriz extracelular están formados por cascadas de proteínas que son activadas por estrés mecánico (Marshall et al., 2003; Tzima et al., 2005; Schwartz y DeSimone, 2008; del Rio et al., 2009; Yonemura et al., 2010; Brunet et al., 2013).

Las maquinarias celulares encargadas de la translocación de proteínas a orgánulos (Thoren et al., 2009; Rodriguez-Larrea y Bayley, 2014) y degradación de proteínas (Aubin-Tam et al., 2011; Sen et al., 2013) funcionan también mecánicamente desnaturalizando previamente las mismas, incluso hay evidencias de que el plegamiento asistido de proteínas por medio de chaperonas es un proceso mediado por fuerza en alguna de sus fases (Lin et al., 2008). Esto implica que la mayoría de las proteínas en la célula podrían ser desplegadas mecánicamente en alguna fase de su ciclo funcional, lo que implica de nuevo que las propiedades mecánicas de muchas proteínas del proteoma son críticas para su funcionalidad, aunque no tengan directamente una función mecánica, pudiendo desencadenar incluso procesos patológicos (Hervas et al., 2012).

### **1.2.- Técnicas de manipulación de molécula individual.**

Así pues, conceptos físicos básicos como fuerza, momento de una fuerza, trabajo, conversión energética, etc, se hacen necesarios para entender el funcionamiento fisiológico de estas biomoléculas. Las técnicas de manipulación de molécula individual (SM, del inglés *single molecule*) nos han permitido la manipulación mecánica de moléculas individuales, abriéndonos el camino a investigar estos procesos de manera directa y en tiempo real. Además de acceder a medir fuerzas como un parámetro experimental de relevancia fisiológica, estas técnicas, dado que estudian una sola molécula, permiten eliminar el “ruido poblacional” inherente a las técnicas de la bioquímica clásica. Este ruido es inherente al estudio de un enorme número de moléculas (del rango del número de





**Figura 1.2.- Procesos biológicos y técnicas de manipulación de molécula individual. a)** Resolución espacial y tamaño de estructuras biológicas representativas. **b)** Rango de fuerzas de las técnicas de manipulación de SM y rango de fuerzas que experimentan algunas biomoléculas. **c)** Resolución temporal de las técnicas, y duración de algunos procesos biológicos. **d)** Rango de velocidad de estiramiento de las técnicas, y rango de velocidad de bionanomaquinas representativas. AFM: microscopio de fuerza atómica; BFP: sonda de fuerza de biomembrana; MN: microagujas; OT: pinzas ópticas; MT: pinzas magnéticas; FF: técnicas de campo de flujo. Figura modificada de (Carrión-Vázquez et al., 2006).

Avogadro  $\sim 10^{23}$ ) que proporciona parámetros promediados en los que quedan ocultos eventos o intermedios raros. Por otra parte, con estas técnicas, al evitarse el ruido poblacional, se elimina también la necesidad de mantener la población de moléculas sincronizada (que es imprescindible cuando se estudia con técnicas poblacionales un proceso que implica múltiples pasos), y se posibilita la detección de eventos o intermedios raros, pudiéndose de hecho reconstruir la distribución de probabilidad de cada parámetro directamente (Weisel et al., 2003; Carrión-Vázquez et al., 2006).

En el caso del estudio de interacciones intermoleculares, las ventajas de las técnicas SM son aún más obvias. Los estudios clásicos han estado dominados por un punto de vista estático, en el que el énfasis se hace en moléculas en solución en condiciones de equilibrio, mientras que sus interacciones biológicas reales generalmente ocurren en superficies bajo condiciones de no equilibrio (Weisel et al., 2003). Además, independientemente de si la fuerza es un parámetro fisiológico relevante en el establecimiento y mantenimiento de la interacción, las SM han permitido la estimación de las constantes cinéticas, el mapeado del paisaje energético, extraer datos estructurales de los bolsillos de unión, identificar barreras cinéticas ocultas a las técnicas poblacionales, y la discriminación y caracterización cuantitativa de diferentes modos de unión entre proteínas y sus ligandos (Merkel et al., 1999; Nevo et al., 2003; Perret et al., 2004; Hinterdorfer, 2005; Lee et al., 2007; Morfill et al., 2008; Fuhrmann et al., 2009; Fuhrmann y Ros, 2010; Gao et al., 2012; Schoeler et al., 2014).

Las técnicas de manipulación de molécula individual más representativas incluyen transductores de fuerzas mecánicas tales como el AFM (del inglés *Atomic Force Microscopy*), la sonda de fuerza de biomembrana; y manipuladores de campo externo como las pinzas ópticas y las pinzas magnéticas (Bustamante et al., 2004; Ritort, 2006). Se trata de técnicas complementarias que cubren rangos solapantes de fuerza, y que también se diferencian en su resolución espacial y en su rango dinámico (para detalle de los diferentes rangos de éstas y otras técnicas ver **figura 1.2**). Las pinzas ópticas y magnéticas son ideales para los estudios en el régimen de fuerza de unos pocos picoNewtons e inferior. La espectroscopia de fuerzas de molécula individual basada en AFM (AFM-SMFS, del inglés *AFM-based single molecule force spectroscopy*, **sección 1.3.1**) se ha utilizado ampliamente para investigar las propiedades mecánicas de proteínas y los procesos inducidos por fuerzas, tales como el desplegamiento de proteínas y la ruptura de interacciones moleculares, que típicamente ocurren a fuerzas “elevadas” entre varias decenas y cientos de picoNewtons (a las velocidades de estiramiento que emplea esta técnica) (Puchner y Gaub, 2009).

### **1.3.- Microscopía de Fuerza Atómica.**

Una característica general de todas las técnicas de barrido mediante sonda (SPM, del inglés *Scanning Probe Microscopy*), entre las que se encuentra el AFM, es la presencia de un brazo flexible que interactúa directamente con la superficie de la muestra a través de una punta. Esto nos permite estudiar detalles de la muestra a escala nanoscópica obteniendo información que de otro modo resultaría inaccesible. El AFM es un aparato relativamente simple que emplea un transductor de fuerza mecánica basado en un sensor (micropalanca) que termina en una punta fina, que a su vez contacta directamente con la muestra (**figura 1.3**). En esta técnica, las fuerzas son calculadas en base a la deflexión de la micropalanca, mientras que las distancias se infieren en base al desplazamiento de un posicionador piezoeléctrico sobre el que se encuentra la muestra. Gracias a la gran precisión en la medida de la posición del actuador piezoeléctrico, la resolución espacial en el eje Z, dirección normal a la superficie, puede llegar a valores menores de 1 nm. En cuanto a la resolución en el plano XY, está limitada por el radio mínimo de la punta utilizada para escanear la superficie, que suele oscilar entre 2 y 50 nm aproximadamente. Por tanto, la resolución lateral de un AFM dependerá fundamentalmente del estado de la punta y, en las mejores condiciones, podrá llegar a unos 5 a 10 nm en medio fisiológico.

#### **1.3.1. - AFM: Modo de espectroscopia de fuerzas de molécula individual**

Dado a que el AFM utiliza un transductor de fuerza, una aplicación típica del mismo ha sido la medida de propiedades nanomecánicas de materiales (inmovilizados en superficies) al someterlos a presión o estiramiento. De hecho, como ya se ha mencionado, este instrumento se ha empleado extensivamente para desplegar mecánicamente biopolímeros y para determinar fuerzas de interacción en pares de biomoléculas. La modalidad de SMFS permite medir fuerzas con una sensibilidad de pN y cambios en longitud (distancias) con resolución de nm, y todo ello con una resolución temporal inferior al ms. Esto hace que el AFM-SMFS haya sido utilizado para caracterizar las propiedades nanomecánicas de un gran número de proteínas (Rief et al., 1997; Oberhauser et al., 2001; Li et al., 2002; Carrión-Vázquez et al., 2006; Valbuena et al., 2009; Hervas et al., 2012; Scholl et al., 2014). Además, el AFM-SMFS ha sido profusamente utilizado para estudiar las interacciones proteína-proteína a nivel de molécula individual (Schwesinger et al., 2000; Baumgartner et al., 2003; Ritco-Vonsovici et al., 2007; Kong et al., 2009; Stahl et al., 2012; Rangl et al., 2013); sin embargo a día de hoy es un gran reto identificar, en experimentos de AFM-SMFS, un evento real de disociación de ligando-receptor y distinguirlo de eventos múltiples e interacciones inespecíficas (ver **sección 1.4.2**) (Weisel et al., 2003; Carrión-Vázquez et al., 2006; Han et al., 2012).

Para poder medir la fuerza ejercida sobre nuestro sistema de interés, es necesario transformar en fuerzas la deflexión que sufre la micropalanca. Esto se consigue mediante la medición de la desviación angular de un láser reflejado en la parte superior de la misma, y que incide en un fotodiodo dividido en dos sectores (A y B) (**figura 1.3**). De tal forma que si la fuerza es cero, el punto reflejado del láser generará el mismo voltaje en los dos sectores del fotodiodo, en cambio, si la micropalanca se dobla por efecto de una fuerza, al presionarlo contra la superficie o al estirar una molécula, el haz láser se desviará generando diferentes voltajes en los sectores A y B. Por tanto, la fuerza que experimenta la micropalanca es proporcional a la diferencia de voltajes generados por la luz láser en ambos sectores del fotodiodo. El siguiente paso consiste en convertir las diferencias de voltajes medidas en el fotodiodo en fuerzas ejercidas entre la punta y la muestra. Para ello hemos de considerar que la micropalanca se puede modelizar como un muelle Hookeano:

$$F = -k \Delta z_c \quad \text{Ec. 1}$$

donde  $k$  es la constante elástica de la micropalanca y  $\Delta z_c$  es su desplazamiento en el eje Z.

Para convertir las diferencias de voltaje del fotodiodo,  $\Delta V$ , en  $\Delta z_c$ , se realiza un ciclo de aproximación-alejamiento sobre un sustrato duro no deformable midiendo la pendiente de  $\Delta V$  vs.  $\Delta z_p$  (desplazamiento del posicionador piezoeléctrico). En estas condiciones el desplazamiento de la micropalanca,  $\Delta z_c$ , se iguala al del posicionador,  $\Delta z_p$ . El valor de esta pendiente se denomina factor de sensibilidad  $C$ , y con él podemos transformar la señal de voltaje del fotodiodo en deflexión de la micropalanca,  $\Delta z_c$ :

$$\Delta z_c = \Delta z_p = \Delta V / C \quad \text{Ec. 2}$$

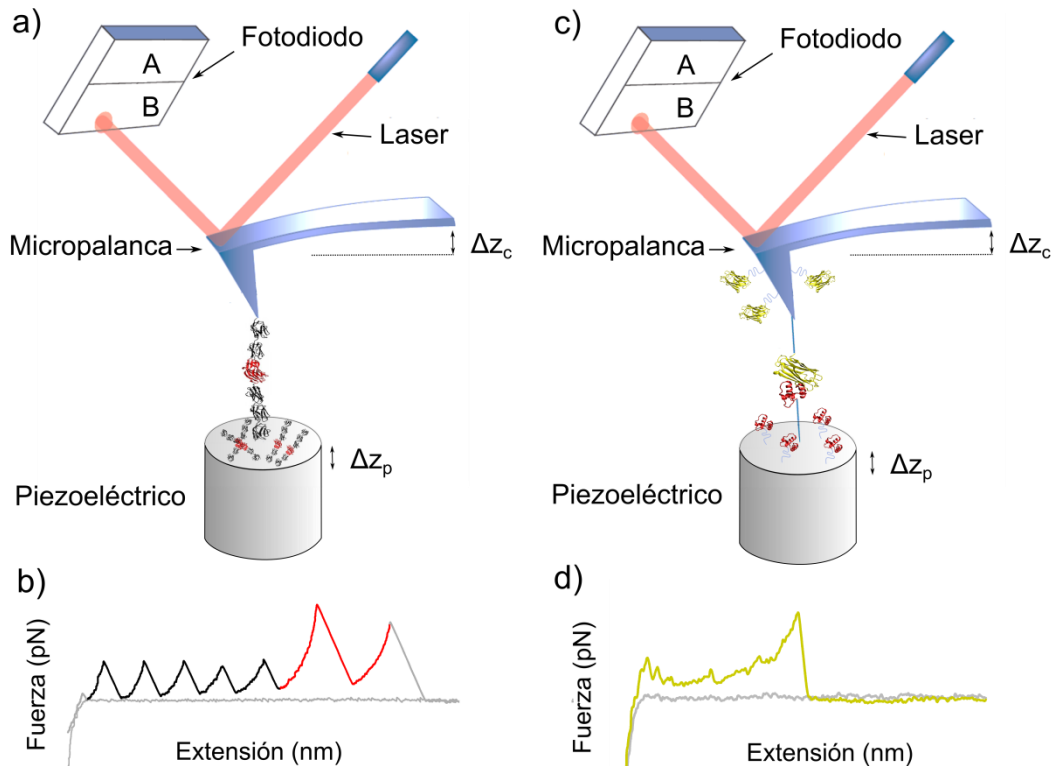
Para calcular la constante de muelle,  $k$ , de la micropalanca el método más empleado es el denominado “método térmico”, en el que la micropalanca se modeliza como un oscilador armónico fluctuando en respuesta al ruido térmico. De tal manera que podemos calcular su constante  $k$  usando el teorema de equipartición:

$$k = k_B T / \langle \Delta x^2 \rangle \quad \text{Ec. 3}$$

donde  $\langle \Delta x^2 \rangle$  representa el desplazamiento cuadrático medio de la micropalanca debido al ruido térmico (Hutter y Bechhoefer, 1993). La constante  $k$  de la micropalanca normalmente se encuentra en el rango de 10-100 pN·nm<sup>-1</sup> para experimentos de SMFS.

### **1.3.2. - AFM: análisis de las propiedades mecánicas de las proteínas y de las interacciones intermoleculares.**

En un experimento típico de caracterización mecánica de proteínas por AFM-SMFS, la muestra se deposita (en el tampón experimental adecuado) sobre un sustrato montado en un posicionador piezoeléctrico. Para poder llevar a cabo el estudio de las propiedades



**Figura 1.3.- Dibujo esquemático de un AFM y configuraciones experimentales en AFM-SMFS.**

**a)** Representación de los elementos clave de un AFM y ejemplo de la caracterización mecánica de una proteína. El sensor de fuerza está constituido por una micropalanca que sufre deflexión cuando se ejerce una fuerza sobre ella. La deflexión de la micropalanca se sigue gracias a reflexión sobre la micropalanca de un láser, que finalmente incide sobre un fotodiodo dividido en dos regiones (A y B). La muestra se deposita sobre un posicionador piezoeléctrico con una resolución espacial subnanométrica. La muestra se aproxima a la punta y, tras sufrir el contacto, algunas moléculas quedan unidas a punta y superficie, creando un circuito mecánico que permite ejercer fuerza sobre la molécula. **b)** Ejemplo de un registro de fuerza-extensión obtenido en una proteína multi-modular al alejar el posicionador a una velocidad constante. Cada pico de fuerza se relaciona con el desplegamiento de un módulo proteico en el ejemplo. **c)** Representación de los elementos clave de un AFM y ejemplo de la caracterización mecánica de una interacción proteína-proteína. A diferencia del ejemplo anterior la punta de la micropalanca porta una de las proteínas, y el sustrato la complementaria. Tras hacer contacto, se produce la interacción entre ambas proteínas y se genera el circuito mecánico. **d)** Ejemplo de un registro de fuerza-extensión obtenido tras aplicar fuerza a una interacción proteína-proteína al alejar el actuador piezoeléctrico a una velocidad constante. El último pico de fuerza observado se asocia a la ruptura de la interacción (ver **sección 1.4.2** para una discusión en profundidad).  $\Delta z_p$  representa el desplazamiento del posicionador en el eje Z (perpendicular a la superficie de la muestra),  $\Delta z_c$  representa el desplazamiento de la micropalanca en el eje Z con respecto a su posición de equilibrio.

mecánicas de la proteína, es necesario que ésta quede anclada tanto al sustrato como a la punta de la micropalanca de forma que se establezca un circuito mecánico (**figura 1.3.a**). Las moléculas sobre el sustrato se acercan a la punta gracias al desplazamiento del posicionador, y tras establecer contacto con la punta algunas moléculas quedan atrapadas entre el sensor de fuerza (micropalanca) y el sustrato (bien por adsorción o por algún tipo de funcionalización específica de la superficie). El circuito mecánico que se establece permite generar fuerza sobre la biomolécula alejando el posicionador del sensor (**figura 1.3.a**). En el modo más utilizado de SMFS, el modo de **longitud controlada**, el

posicionador se retrae a velocidad constante hasta una distancia fijada. La micropalanca reporta las fuerzas que se oponen a la extensión generadas por cualquier material que se hubiera quedado atrapado entre punta y sustrato, idealmente una sola molécula de nuestra proteína. Así pues, cuando por azar se engancha una sola molécula entre punta y sustrato, entonces podemos estudiar su resistencia al estiramiento en condiciones de molécula individual. En este modo de SMFS los registros obtenidos son espectros de fuerza-extensión (**figura 1.3.b** y **1.6.a**), en los que se representa la fuerza medida por la micropalanca (con resolución en el rango de los pN) frente a la extensión recorrida por el posicionador piezoeléctrico (con resolución sub-nanométrica). El otro modo experimental que se utiliza en el campo de la nanomecánica es el de **fuerza controlada**, en el que mediante un mecanismo de retroalimentación, se controla la dependencia temporal de la fuerza variando la posición del actuador piezoeléctrico. En este modo de SMFS los registros obtenidos son espectros de distancia-tiempo (**figura 1.6.b**) (la distancia se varía automáticamente para conseguir la aplicación de la fuerza deseada en cada momento) (Carrión-Vázquez et al., 2006).

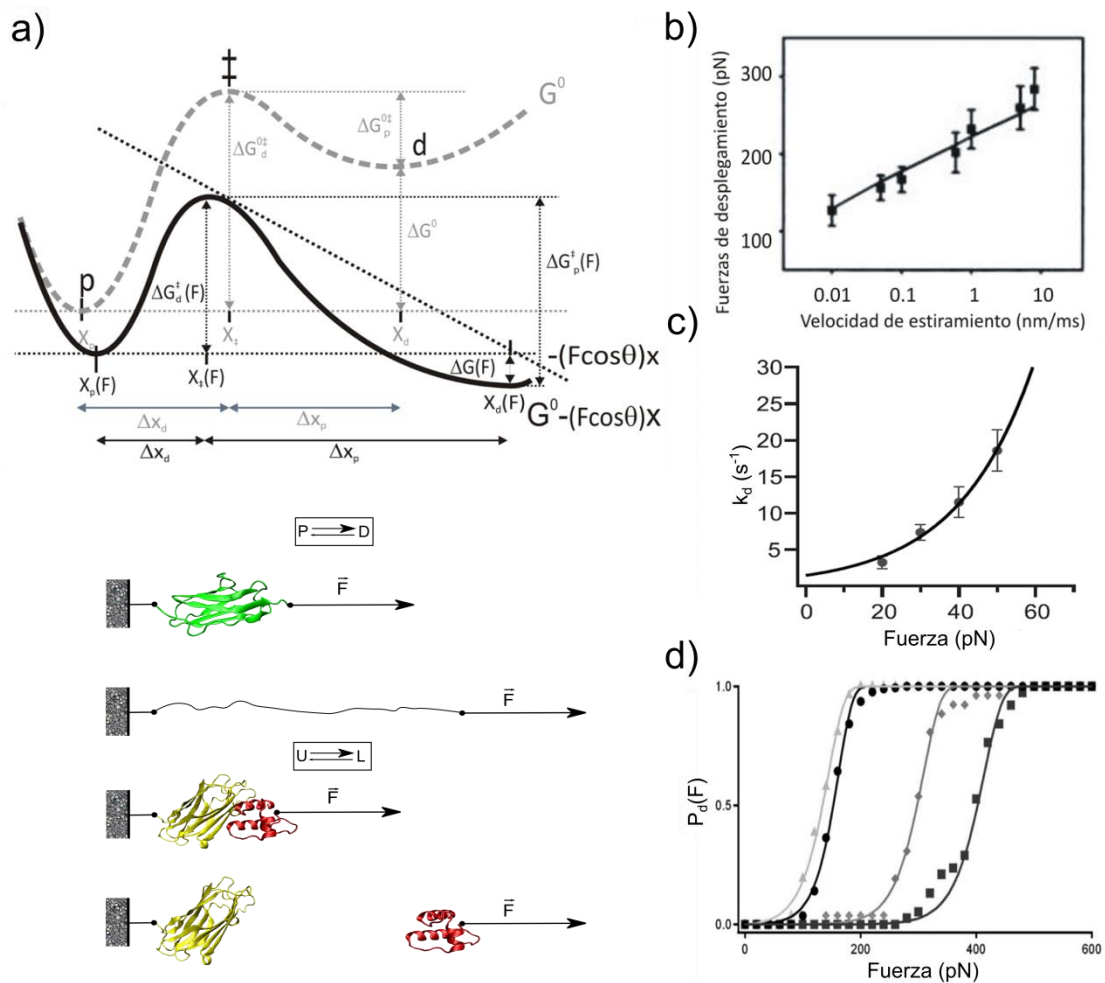
En el caso de medidas de interacciones intermoleculares la configuración experimental clásica es diferente. La punta de la micropalanca porta un elemento del par de moléculas interaccionantes, y el sustrato el otro (**figura 1.3.c**). El procedimiento de medida requiere del contacto de la superficie del sustrato con la punta, para que se forme el complejo proteína-proteína. Cuando se aleja el sustrato de la punta se produce el estiramiento del complejo, midiéndose la fuerza a la que se produce la ruptura del mismo (Hinterdorfer y Dufrene, 2006; Fuhrmann y Ros, 2010) (**figura 1.3.d**). El modo de fuerza controlada también puede implementarse en estos experimentos, midiendo el tiempo de vida del complejo a una fuerza determinada (Marshall et al., 2003; Kong et al., 2009; Rakshit et al., 2012). Hay que tener en cuenta que, sea cual sea el modo experimental utilizado, lo que se está midiendo es la pérdida de adhesión entre la punta y el sustrato (Merkel, 2001), lo que hace difícil discernir entre la señal experimental, el despegado de una de las proteínas de la superficie, y la adhesión inespecífica entre punta-sustrato (**figura 1.5.a**). Otro asunto importante, que es difícil de responder en estos experimentos, es si sólo una o múltiples parejas de moléculas están siendo estudiadas. A día de hoy no se dispone en el campo de un criterio que por sí mismo pueda resolver en cada registro ambas incógnitas: ¿se trata de un pico de fuerza de ruptura proveniente de la interacción, de la pérdida de adhesión de una de las moléculas de su superficie, o de adhesión inespecífica? ¿se estudia una sola pareja de moléculas interaccionantes? (Weisel et al., 2003; Carrión-Vázquez et al., 2006).

### 1.3.3. - Propiedades mecánicas de proteínas y estudio de interacciones intermoleculares: análisis teórico

Los experimentos de desplegamiento mecánico y de interacciones proteína-proteína se analizan asumiendo un modelo basado en la teoría del estado de transición (Bell, 1978; Evans y Ritchie, 1997; Evans y Ritchie, 1999). El proceso de desplegamiento/disociación se modeliza como un proceso activado térmicamente en el que existen dos estados: plegado/unido y desplegado/libre, separados ambos por una barrera energética (**figura 1.4.a**). De tal manera que uno de los principales efectos de la fuerza es disminuir la altura relativa de la barrera energética, aumentando la probabilidad de transición al estado desplegado/libre:

$$k_d(F) = A \exp \left[ -\frac{\Delta G_d^{0\ddagger} - F\Delta x_d}{k_B T} \right] = k_d(0) \exp \left[ \frac{F\Delta x_d}{k_B T} \right] \quad \text{Ec. 4}$$

donde  $k_d(F)$  es la frecuencia de desplegamiento/disociación a una fuerza  $F$  dada,  $A$  es el denominado factor pre-exponencial,  $\Delta G_d^{0\ddagger}$  es la altura de la barrera energética entre el estado plegado/unido y el estado de transición,  $\Delta x_d$  la distancia al estado de transición a lo largo de las coordenadas de reacción y  $k_B T$  es la energía térmica.



**Figura 1.4.- Efecto de la fuerza en el paisaje energético y extracción de parámetros cinéticos en experimentos de AFM-SMFS.** **a)** El proceso de desplegamiento de una proteína, o de disociación ligando-receptor, se puede modelizar como un proceso de dos estados: plegado/unido y desplegado/libre que están separados por una barrera energética ( $\Delta G_d^{0\ddagger}$ ). El efecto de la fuerza es rebajar la altura de dicha barrera energética en una magnitud  $(F\cos\theta)x$ , lo que aumenta la probabilidad de transición de un estado al otro según la **Ec. 4**. De hecho, la fuerza provoca una distorsión del paisaje energético que no sólo disminuye la altura efectiva de la barrera energética en el estado de transición, sino que también desplaza la localización del mismo. Sin embargo, como una muy buena aproximación, para paisajes con mucha curvatura, se puede considerar que la distancia al estado de transición,  $\Delta x_d$ , no se ve afectada por la fuerza. **b), c) y d)** extracción de los parámetros cinéticos en experimentos de AFM-SMFS. **b)** La dependencia de las fuerzas de ruptura con el logaritmo de la tasa de carga en experimentos de LC nos permite calcular  $\Delta x_d$  de la pendiente de la recta, y  $k_d(0)$  de la extrapolación a fuerza cero. **c)** La dependencia de  $k_d(F)$  con la fuerza en experimentos de FC nos permite estimar  $k_d(0)$  por extrapolación a fuerza cero. **d)** en experimentos de FR el ajuste de la  $P_d(F)$  experimental a la expresión **Ec. 7** permite estimar  $\Delta x_d$  y  $k_d(0)$ . Figuras a) y b) adaptadas de (Carrión-Vázquez et al., 2006), figuras d) y e) adaptadas de (del Rio et al., 2009) y (Valbuena et al., 2009), respectivamente.

$k_d(0) = A \exp[-\Delta G_d^{0\ddagger}/k_B T]$  es la tasa de desplegamiento/disociación a fuerza cero.

Los experimentos de SMFS nos permiten extraer información sobre la cinética del proceso y del paisaje energético, ya que  $k_d(0)$  y  $\Delta x_d$  pueden calcularse experimentalmente, y una vez conocido  $k_d(0)$  puede estimarse la altura de la barrera energética (Hinterdorfer, 2005; Carrión-Vázquez et al., 2006; Lee et al., 2007).

Los protocolos experimentales de SMFS se dividen en dos grupos generales, los de **longitud controlada** (LC, del inglés *Length Clamp*) en los que las proteínas se estiran a una velocidad constante hasta una longitud previamente determinada, y los de **fuerza controlada**, en los que se controla la dependencia temporal de la fuerza.

En los experimentos de LC, a medida que se produce el alejamiento entre la punta y la muestra, las proteínas van sufriendo un aumento de la tensión que se ejerce sobre ella. Esto tiene como consecuencia directa un aumento continuado en la frecuencia de desplegamiento/disociación a lo largo del experimento (**Ec. 4**). De hecho cómo la fuerza varía de tal forma que  $dF/dt = r = k_{ef} \cdot v$ , donde  $v$  es la velocidad de alejamiento y  $k_{ef}$  la constante efectiva de muelle (que incluye a la proteína y a la micropalanca), puede calcularse la distribución de fuerzas de ruptura (Evans y Ritchie, 1997; Evans y Ritchie, 1999), siendo el valor de fuerza más probable:

$$F_d = \left(\frac{k_B T}{\Delta x_d}\right) \ln\left(\frac{r \Delta x_d}{k_d(0) k_B T}\right) \quad \text{Ec. 5}$$

Esta ecuación pone de manifiesto un hecho muy importante, y es que la estabilidad mecánica de una proteína o de una interacción proteica, no es un valor absoluto, sino que se trata de un parámetro dinámico que depende de la velocidad a la que llevamos a cabo el experimento. Por otra parte, esta relación entre la  $F_d$  y el logaritmo de  $r$ , nos permite calcular los parámetros  $k_d(0)$  y  $\Delta x_d$  (Rief et al., 1997; Rief et al., 1998; Oberhauser et al., 2002; Morfill et al., 2007; Jollymore et al., 2009; Rangl et al., 2013; Schoeler et al., 2014):



$\Delta x_d$ , se estima de la pendiente de la recta que se obtiene al representar  $F_d$  vs logaritmo de  $r$ , y de la extrapolación a fuerza cero se obtiene la  $k_d(0)$  (**figura 1.4.b**).

El modo más sencillo de los protocolos de fuerza controlada, es el de fuerza constante (FC, del inglés *force clamp*). En FC sometemos a la proteína/interacción a fuerza constante, de manera que esta soporta una fuerza determinada durante un tiempo  $t_i$ , tras el cual se producirá su desnaturalización/disociación. Si representamos la frecuencia de ocurrencia de dichos tiempos, podremos ajustar a una exponencial simple y de esta manera obtener el tiempo de decaimiento de la función, el cual corresponde con la inversa de la constante cinética de desplegamiento/disociación a esa fuerza (Oberhauser et al., 2001; Marshall et al., 2003; Schlierf et al., 2004; Cao et al., 2008; Kong et al., 2009; Rakshit et al., 2012) (**Ec. 4, figura 1.4.c**). El principal problema de este método es que los tiempos necesarios para observar el evento pueden resultar demasiado largos, de tal manera que la medida podría verse afectada por la deriva de la fuerza y de la posición a los tiempos experimentales.

Para solventar este problema, y adquirir una mayor cantidad de eventos en la ventana experimental, es posible realizar rampas de fuerza (FR, del inglés *force ramp*). En FR la fuerza ejercida tiene un comportamiento lineal y creciente a un ritmo fijado (que suele expresarse en pN/s). Podemos formular un modelo cinético simple de dos estados para la transición mecánica, en el que las probabilidades de transición van a venir determinadas por la integración de la siguiente expresión:

$$dP_d(t) = k_d(t)(1 - P_d(t))dt \quad \text{Ec. 6}$$

Teniendo en cuenta que la fuerza varía según  $F=at$ , donde  $a$  es la pendiente de la rampa, podemos integrar la anterior expresión, y cambiar de variable ( $t$  por  $F$ ), obteniendo una expresión analítica de probabilidad de transición en función de la fuerza, a la que podemos ajustar los datos experimentales para obtener  $k_d(0)$  y  $\Delta x_d$  (**figura 1.4.c**) (Oberhauser et al., 2001; Schlierf et al., 2004; Valbuena et al., 2007) :

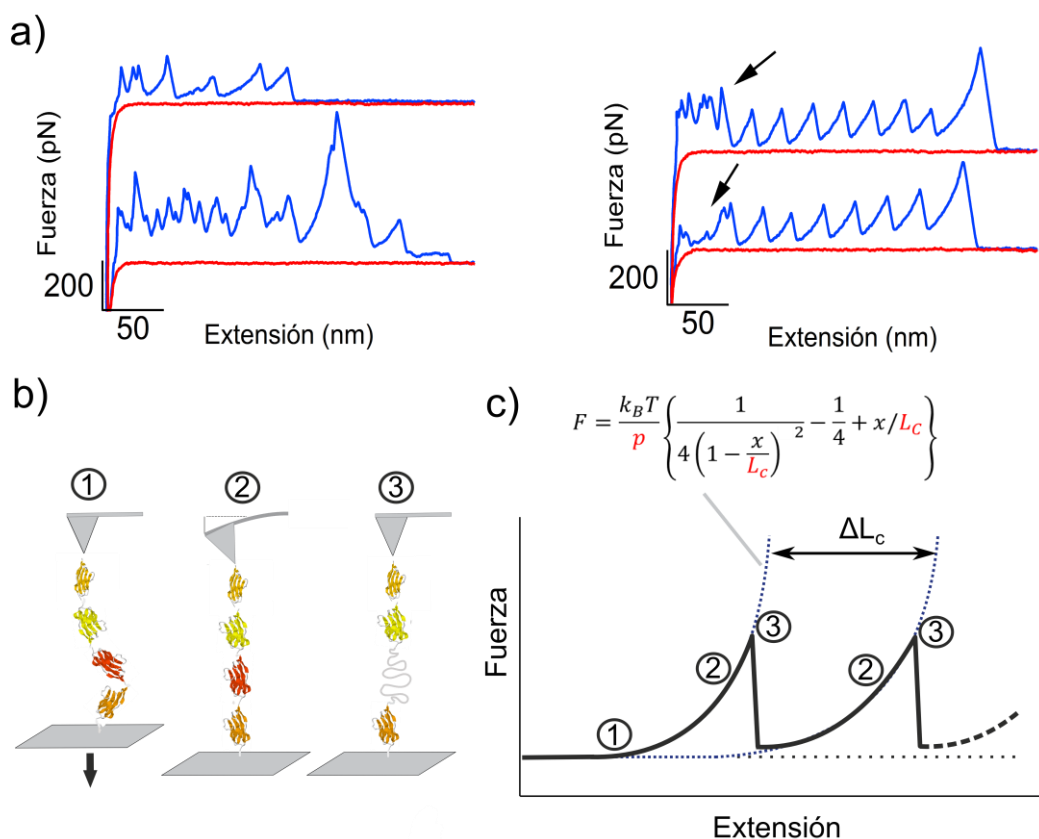
$$P_d(F) = 1 - e^{-(k_d(0)/a) \int_0^F e^{(f\Delta x_d/k_B T)} df} \quad \text{Ec. 7}$$

Es evidente que para obtener toda esta información sobre cinética y energética del proceso, un pre-requisito indispensable, es la correcta identificación del evento de ruptura (de plegamiento o interacción), lo que como se verá más adelante puede llegar a ser problemático y complejo en estudios de AFM-SMFS.

#### 1.3.4. - AFM-SMFS: El uso de poliproteínas.

Uno de los principales problemas de la AFM-SMFS es distinguir de forma inequívoca la señal del ruido experimental. Cuando se produce el contacto punta-muestra y posterior alejamiento de ambos se suelen observar fuerzas de adhesión inespecíficas que

suponen un gran problema: ¿cómo distinguimos el desplegamiento de una proteína de la rotura de una adhesión inespecífica? (**figura 1.5.a izquierda**). La conocida como estrategia de las poliproteínas, ha sido profusamente utilizada en el estudio del desplegamiento de proteínas por AFM-SMFS para solventar este problema. En ella, ya sea mediante técnicas de ingeniería genética (Carrion-Vazquez et al., 1999), o de bioquímica de proteínas (Yang et al., 2000), se generan repeticiones en tándem de la proteína o módulo proteico de interés, originándose en el experimento una señal repetida que es el indicador de molécula individual (**figura 1.5.a derecha**, y **figura 1.6.a**). Además de la clara ventaja de poder distinguir con seguridad la señal experimental, las poliproteínas tienen otras muchas ventajas (que comentaremos más adelante en detalle) entre las cuales cabe destacar que no necesitan de controles de molécula individual adicionales, ya que dichos controles están incluidos en la propia molécula de proteína que estamos estudiando. Aunque las primeras poliproteínas constaban de repeticiones del mismo módulo (homopoliproteínas, **figura 1.5.a derecha**) (Carrion-Vazquez et al., 1999; Carrion-Vazquez et al., 2000; Li et al., 2000), hoy en día es común el uso de heteropoliproteínas (Sharma et al., 2007; Randles et al., 2007; Valbuena et al., 2009; Hervas et al., 2012), en las que uno de los módulos de una homopoliproteína (por ejemplo I27, o ubicuitina) es reemplazado a nivel de ADN (en un vector de expresión) por la secuencia de la proteína de interés (**figura 1.6.a**).



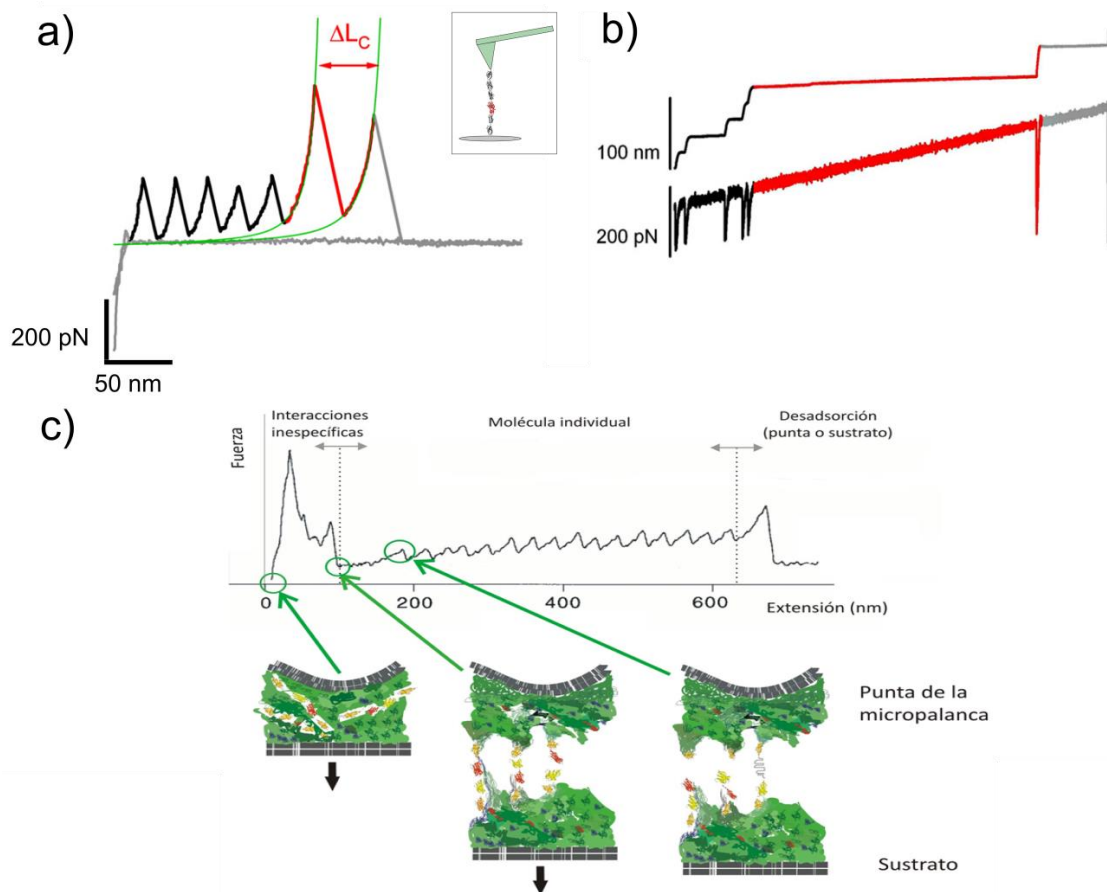
**Figura 1.5.- Uso de poliproteínas en AFM-SMFS. a)** A la izquierda dos registros fuerza-extensión obtenidos en LC correspondientes a adsorción inespecífica entre la punta y la muestra. La presencia de interacciones inespecíficas hace difícil la identificación de la señal experimental proveniente del desplegamiento de una proteína por estiramiento. La estrategia de las poliproteínas (derecha), origina una señal repetida (dientes de sierra en LC) que permite separar inequívocamente la señal originada por la proteína del ruido proveniente de interacciones inespecíficas (flechas). **b)** y **c)** interpretación del patrón en dientes de sierra originado por una poliproteína en LC. Cuando la poliproteína queda atrapada entre la punta y la superficie, cerrando el circuito mecánico, la tensión aumenta a medida que el posicionador se aleja (1 y 2). La fuerza que opone la proteína al estiramiento en función de la distancia  $x$  entre sus extremos puede ser descrita por el modelo WLC (ecuación y línea discontinua en **c)**), que nos permite estimar la longitud de contorno  $L_c$  (distancia máxima que puede alcanzar el polímero con todos sus componentes perfectamente alineados en la misma dirección) y la longitud de persistencia  $p$ , que se puede interpretar como el mínimo componente rígido de la cadena (se puede aproximar a la longitud de un residuo de aminoácido). Si nos fijamos en un módulo cualquiera de la poliproteína, cuando la fuerza aumenta hasta un umbral determinado (fuerza de desplegamiento), la estructura del módulo colapsa y se produce una caída abrupta en la fuerza (3). La caída en la fuerza se debe a que el desplegamiento del módulo libera un grupo de residuos aminoacídicos que estaban ocultos a la fuerza, de tal manera que la longitud de la poliproteína aumenta súbitamente, lo que origina la caída en la fuerza. A medida que el posicionador continúa desplazándose, la fuerza vuelve a aumentar nuevamente, y del ajuste de  $L_c$ , mediante el WLC, podemos calcular la distancia liberada por la proteína como el incremento de  $L_c$  con respecto a la curva anterior. Este  $\Delta L_c$  es un parámetro característico de cada proteína y representa la longitud de la proteína liberada tras un desplegamiento, lo que a su vez está relacionado con la posición de la barrera mecánica que origina la estabilidad mecánica. La repetición de este proceso en varias repeticiones de la poliproteína origina el patrón en dientes de sierra.

Cuando estiramos una proteína mediante SMFS, la primera fuente de resistencia a la extensión que ofrece la molécula tiene origen entrópico, resultante de la tendencia de un polímero a formar un ovillo para maximizar su libertad conformacional (o entropía, la cual se reduce con el estiramiento). Esta elasticidad entrópica sigue una relación no lineal que se puede describir formalmente por el modelo WLC (*Worm like Chain*) (Bustamante et al., 1994) (**figura 1.5.c** y **sección 3.10**). Este modelo predice la fuerza de resistencia entrópica,  $F$ , del polipéptido en función de tres parámetros: la distancia entre sus extremos ( $x$ ), la longitud de contorno ( $L_c$ , longitud total del polipéptido desplegado a fuerza infinita), y la longitud de persistencia ( $p$ , longitud máxima a lo largo de la cual la dirección del polímero permanece correlacionada).  $p$  se puede interpretar como el mínimo componente rígido de la cadena; en proteínas suele utilizarse el valor 0.4 nm, que coincide aproximadamente con la longitud de un residuo de aminoácido en una proteína estirada (Ainavarapu et al., 2007).

En general, mientras una proteína está plegada, la tensión no se distribuye homogéneamente por toda la cadena polipeptídica, sino que existen regiones concretas que soportan la mayor parte de la carga mecánica y regiones que están "ocultas" a la fuerza. Una vez que las interacciones que suponen la resistencia mecánica de la proteína (región conocida como "broche mecánico") se rompen, la proteína generalmente se despliega por completo, liberando una nueva región de la proteína que queda expuesta a la

fuerza (**figura 1.5.b** y **figura 1.5.c**). El tipo de plegamiento, así como el tipo de estructura secundaria (mantenida por interacciones locales) que contiene la proteína, determinan en gran medida la resistencia que la proteína ofrecerá a su estiramiento antes de que colapse la estructura. Se ha observado una menor estabilidad mecánica en proteínas con gran contenido hélice  $\alpha$  en su estructura, mientras que las proteínas mecánicamente más estables contienen normalmente hebras  $\beta$  conectando los extremos N- y C-terminales en una configuración "en cizalla" donde los enlaces de hidrógeno del esqueleto peptídico se disponen ortogonalmente respecto la fuerza aplicada (Carrión-Vázquez et al., 2006).

Si llevamos a cabo la extensión de una proteína modular o una poliproteína, provocaremos el desplegamiento de cada uno de sus módulos o repeticiones (**figura 1.5.a derecha, 1.5.b, y figura 1.5.c**). Cada vez que se produce el desplegamiento de un módulo, la longitud de la proteína oculta a la fuerza se libera ( $\Delta L_c$ ). Esta distancia se puede estimar en los experimentos de longitud controlada ajustando el modelo WLC a dos picos de fuerza consecutivos y restando sus valores de  $L_c$  ( $\Delta L_c$ ). Este parámetro es característico de cada proteína, depende del plegamiento, y es importante para identificar la molécula a estudio usándose como una especie de "huella digital" o reportero molecular.



**Figura 1.6.- Patrones generados por poliproteínas y ventajas generales que ofrecen. a)** Patrón en dientes de sierra en LC generado por una heteropoliproteína. El desplegamiento repetido de los módulos marcadores (en negro) ofrece el control de molécula individual. En rojo el

despliegamiento del módulo a estudio, con su  $\Delta L_c$  característico. **b)** Patrón en escalera obtenido en experimentos de fuerza controlada en la misma heteropoliproteína. En negro los saltos originados por el despliegamiento por el módulo marcador; en rojo el módulo a estudio. En la figura se presenta un experimento de FR donde la fuerza aumenta a una tasa predeterminada gracias a un mecanismo de retroalimentación. Las espigas en la fuerza se deben al tiempo que tarda el mecanismo de control en restaurar la fuerza nuevamente. **c)** Ventajas adicionales que ofrecen las poliproteínas o las proteínas homo-modulares. Al presentar múltiples repeticiones del mismo módulo, algunos módulos pueden sacrificarse en la adsorción de la proteína a las superficies, además el hecho de presentar múltiples módulos hace que algunos de ellos aparezcan alejados de la zona proximal del espectro, que suele estar contaminada con interacciones inespecíficas punta-sustrato. El final del experimento ocurre cuando la proteína se despega de la punta/sustrato, por lo que el último pico de fuerza es el pico de desadsorción.

El despliegamiento secuencial de cada módulo plegado da lugar a los patrones repetidos que se utilizan para identificar una sola molécula: los “dientes de sierra” de los espectros fuerza-extensión en el modo de longitud controlada (**figura 1.6.a**), o las “escaleras” en los espectros extensión-tiempo en los modos fuerza controlada (**figura 1.6.b**). En estos experimentos los módulos son desplegados de manera ordenada de acuerdo a su estabilidad mecánica, por lo que los menos estables se despliegan antes. El experimento finaliza cuando se produce el desanclaje de la proteína bajo estudio, de la punta o del sustrato (así pues el último pico de fuerza en el modo longitud controlada refleja el desanclaje **figura 1.6.c**).

Como se ha mencionado, el uso de poliproteínas tiene ventajas añadidas. Las proteínas tienden a desnaturalizarse total o parcialmente cuando se adsorben en una superficie, por lo que proteínas largas compuestas por múltiples módulos o poliproteínas poseen una ventaja adicional ya que tienen la posibilidad estadística de “sacrificar” parte de sus módulos o repeticiones en el proceso de anclaje (**figura 1.6.c**). Además, como la parte inicial (próxima al sustrato) de los espectros fuerza-extensión suele estar contaminada por interacciones inespecíficas (**figura 1.5.a derecha y figura 1.6.c**), una proteína modular o una poliproteína, al ser relativamente más largas, mostrarán eventos de despliegamiento alejados de esta zona proximal. Por todo ello, no es de extrañar que los estudios iniciales en este campo se centraran en proteínas multimodulares con funciones mecánicas como titina, tenascina, espectrina y fibronectina (Rief et al., 1997; Oberhauser et al., 1998; Rief et al., 1999; Oberhauser et al., 2002).

#### **1.4.- Introducción a las temáticas abordadas.**

En la presente tesis se han abordado dos temáticas completamente diferentes, enmarcadas ambas dentro de la nanomecánica de proteínas. Por una parte se ha utilizado el AFM-SMFS para la caracterización nanomecánica de una proteína implicada en un sistema biológico cuya funcionalidad se encuentra profundamente afectada por las fuerzas mecánicas: la  $\beta$ -catenina del sistema cadherina/catenina en las uniones adherentes

(uniones célula-célula). En segundo lugar, se ha llevado a cabo un desarrollo de nanomecánica básica en el campo de las interacciones moleculares, para cuyo análisis por AFM-SMFS se ha propuesto una estrategia autocontrolada basada en conceptos similares a los de la estrategia de las poliproteínas.

### 1.4.1. - Nanomecánica de la $\beta$ -catenina.

Una de las características más relevantes de los organismos multicelulares, es que las células que forman sus tejidos deben integrar información bioquímica y **mecánica** del entorno para ejercer su función de manera coordinada dentro del mismo (Vogel y Sheetz, 2006) (**figura 1.7.a**). Cada vez resulta más evidente que las fuerzas mecánicas son factores centrales en la regulación de la morfogénesis de tejidos y en su homeostasis. Durante el desarrollo, en la remodelación de tejidos y en la función de muchos órganos y sistemas (corazón, capilares sanguíneos, pulmones, sistema hematopoyético, tracto gastrointestinal y sistema musculoesquelético), las células se encuentran sometidas a fuerzas de magnitud y dirección variable. Es de vital importancia, que las células perciban estas fuerzas mecánicas y las transduzcan en señales intracelulares que finalmente dirijan modificaciones en procesos tan importantes como el crecimiento, forma, diferenciación y migración celular (Schwartz y DeSimone, 2008; Nelson, 2008; Wang et al., 2009; Asparuhova et al., 2009; Mammoto et al., 2012; Huveneers y de Rooij, 2013; Twiss y de Rooij, 2013).

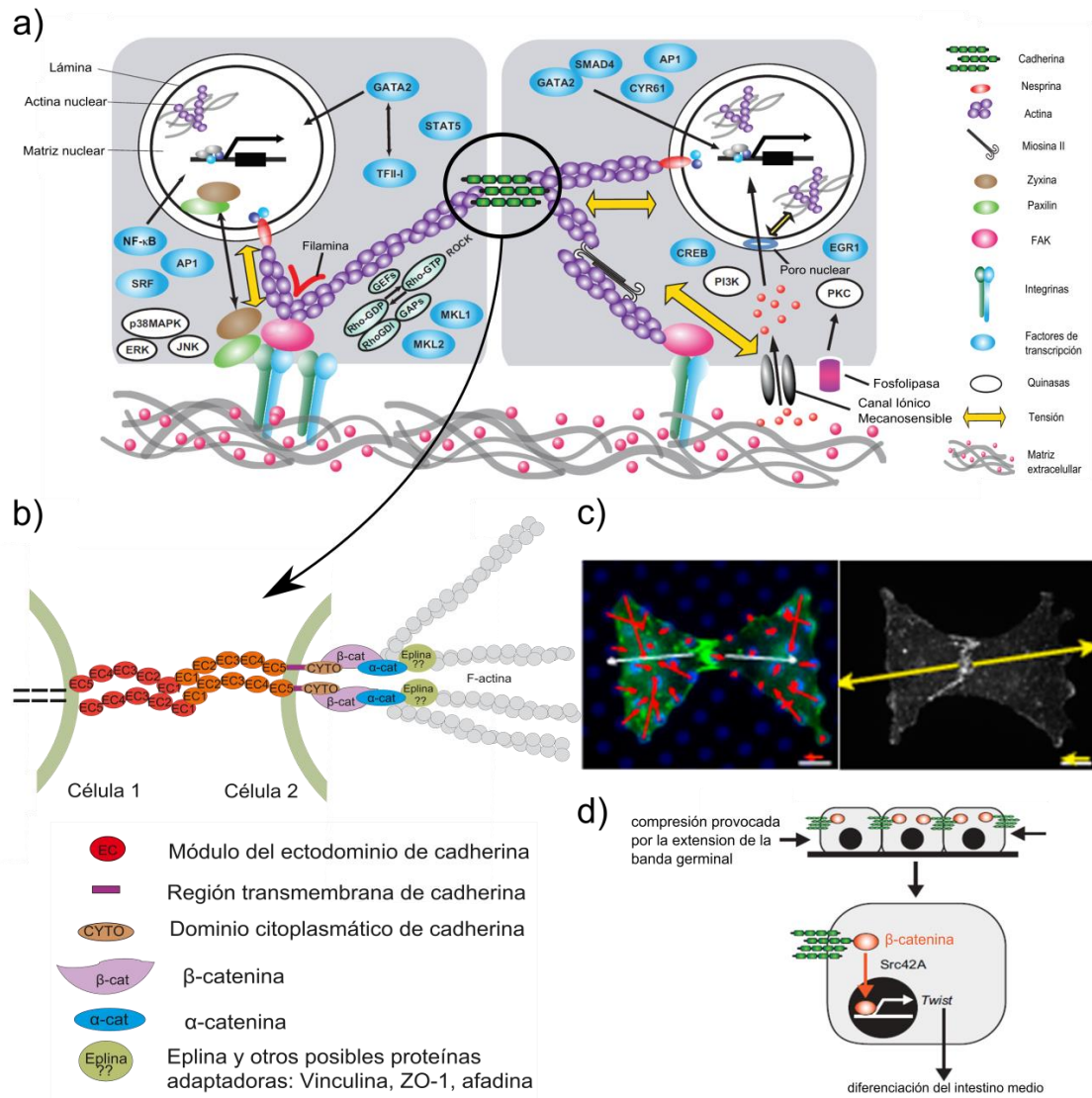
El microambiente celular está formado por la matriz extracelular (ME), y las células vecinas. En las células existen microestructuras especializadas en establecer y mantener contactos con la matriz extracelular (las **adhesiones focales**) y con las células vecinas (las **uniones adherentes**) (Ganz et al., 2006; Schwartz y DeSimone, 2008; Yonemura, 2011; Twiss y de Rooij, 2013). Tanto las adhesiones focales como las uniones adherentes, son zonas íntimamente relacionadas con el citoesqueleto de actomiosina. En ellas, y a través de receptores de superficie (**integrinas** en las adhesiones focales; **cadherinas** y **nectinas** en las uniones adherentes), existe un acoplamiento entre el citoesqueleto y la matriz extracelular, y entre los citoesqueletos de células vecinas (Asparuhova et al., 2009; Goldmann, 2012; Indra et al., 2013; Zaidel-Bar, 2013)(**figura 1.7**).

El sistema cadherina/catenina es el constituyente central de las uniones adherentes (**AJ**, del inglés *Adherent Junctions*) (Ganz et al., 2006; Harris y Tepass, 2010; Desai et al., 2013) (**figura 1.7.a** y **1.7.b**). Las cadherinas clásicas son glicoproteínas transmembrana de paso sencillo, con una región extracelular compuesta por cinco dominios extracelulares cadherina (EC), y una región citosólica conservada (Shapiro y

Weis, 2009; Sivasankar et al., 2009) (**figura 1.7.b**). Cuando se produce un contacto entre células, las cadherinas interaccionan mediante interacciones en *cis* (entre cadherinas de la misma célula) y *trans* (entre cadherinas de células diferentes) por medio de los dominios extracelulares cadherina produciéndose un agrupamiento de las mismas. La unión de cadherinas en superficie induce la organización alrededor de la región citosólica de la “placa” adhesiva (formada principalmente por las cateninas: catenina p120,  $\beta$ -catenina,  $\alpha$ -catenina), y haces de F-actina (actina Filamentosa) que se observan en íntimo contacto con ella (Mege et al., 2006; Shapiro y Weis, 2009; Huveneers y de Rooij, 2013; Zaidel-Bar, 2013). En las fases iniciales de la formación del contacto célula-célula se produce una reorganización del citoesqueleto de actomiosina local, reclutándose a la zona gran cantidad de reguladores de actina como ARP2/3, vinculina, zyxina, ena/VASP, formina-1, y se produce una potenciación de la actividad miosina II. Incluso en los contactos maduros, el citoesqueleto de actina es muy dinámico, y está siendo re-organizado constantemente por reguladores de actina como ARP2/3, N-WASP,  $\alpha$ -actinina y cortactina. Así pues los complejos adhesivos célula-célula regulan la organización de la F-actina y viceversa: la organización del citoesqueleto de actomiosina influencia la composición, maduración y aparición de los complejos adhesivos célula-célula (Mege et al., 2006; Ratheesh y Yap, 2012; Twiss y de Rooij, 2013; Huveneers y de Rooij, 2013).

En los últimos años, se ha hecho evidente la existencia de una conexión mecánica entre el citoesqueleto de actina y el sistema cadherina/catenina, que representa un circuito mecánico directo entre las zonas extracelulares de unión celular y el interior. En cultivos celulares, la contracción del citoesqueleto de actomiosina se traduce en el establecimiento de tensión en la región intracelular de la E-cadherina, y no sólo eso, sino que la aplicación de fuerza directa sobre la célula vecina mediante una micropipeta, origina tensión en dicha región a través de las uniones adherentes (Borghi et al., 2012). Se han medido las fuerzas que las células son capaces de ejercer a través de las uniones de cadherina sobre pilares transductores funcionalizados con N-cadherina, y también las fuerzas que las células pueden ejercer las unas sobre las otras a través de uniones adherentes, observándose que en todos los casos se requiere un esqueleto organizado de actina y actividad miosina II (Ganz et al., 2006; Ladoux et al., 2010; Liu et al., 2010) (**figura 1.7.c**). Pero en las zonas de unión celular mediadas por el sistema cadherina/catenina no sólo se produce una mera transmisión de fuerza desde el exterior celular hacia el interior (y viceversa), sino que se producen claros procesos de mecanotransducción. Mediante estos, las células son capaces de modificar el tamaño de las uniones, la organización del citoesqueleto y la fuerza de tracción que ejercen; todo ello ante estimulación directa por fuerza de las cadherinas, cambios en la actividad miosina II, y cambios en la constante

elástica del medio (Ladoux et al., 2010; le Duc et al., 2010; Liu et al., 2010; Yonemura et al., 2010; Twiss et al., 2012).



**Figura 1.7.- Integración de señales mecánicas, sistema cadherina/catenina y papel de la β-catenina.** **a)** La integración de la información proveniente de estímulos mecánicos en las uniones focales y uniones adherentes, provoca la activación de cascadas de señalización, y alteraciones en la expresión génica, que permiten la correcta homeostasis celular. **b)** El sistema cadherina/catenina de las AJ, en la interfase de unión célula-célula, permite la unión mecánica y funcional de los citoesqueletos de actomiosina de células vecinas. En base a su capacidad de unirse tanto al dominio citoplasmático de las cadherinas como a la proteína de unión a actina α-catenina, la β-catenina constituye una proteína esencial en la interconexión mecánica y funcional entre dos células, conectando los dominios extracitoplasmáticos de las cadherinas (ectodominios, que contactan con cadherinas de las células vecinas) con el citoesqueleto de actomiosina. **c)** Configuración experimental desarrollada en (Liu et al., 2010) que permite calcular la fuerza ejercida por dos células en sus zonas de contacto. En verde zonas de acumulación de β-catenina. **d)** La β-catenina no actúa solamente como elemento estructural que une mecánicamente los contactos célula-célula y el citoesqueleto, sino que participa en sistemas de mecanotransducción que implican su liberación de las uniones adherentes para su translocación al núcleo. Figuras. a) y d) adaptadas de (Mammoto et al., 2012).



La visión clásica de la unión entre el sistema cadherina/catenina y el citoesqueleto de actina, se ha basado en la unión del dominio citoplasmático de las cadherinas a la  **$\beta$ -catenina**, y la unión de esta última a la proteína de unión de actina  $\alpha$ -catenina. Este modelo de unión directa se vio cuestionado en 2005 por la publicación de sendos trabajos en los que no se conseguía unir *in vitro* simultáneamente F-actina y complejo E-cadherina/ $\beta$ -catenina (Drees et al., 2005; Yamada et al., 2005). Sin embargo, existe una gran acumulación de evidencias experimentales que apuntan a que *in vivo*, la  $\alpha$ -catenina resulta esencial para el acoplamiento entre el complejo cadherina/ $\beta$ -catenina y el citoesqueleto de actina (Abe y Takeichi, 2008; Yonemura et al., 2010; Borghi et al., 2012; Twiss et al., 2012; Desai et al., 2013), lo que podría ocurrir a través de interacciones muy dinámicas, de un cambio conformacional de la  $\alpha$ -catenina en las uniones adherentes, o indirectamente, a través de proteínas secundarias como Eplina, vinculina, ZO-1 o afadina (Abe y Takeichi, 2008; Yonemura et al., 2010; Yonemura, 2011; Zaidel-Bar, 2013) (**figura 1.7.b**).

La interacción simultánea de la  $\beta$ -catenina con la región citoplasmática de las cadherinas y la  $\alpha$ -catenina, la convierte en un elemento central clave para la transmisión de la tensión mecánica entre el citoesqueleto y las cadherinas en las uniones adherentes. Cuando se elimina el sitio de unión a  $\beta$ -catenina de las cadherinas, el acoplamiento entre el flujo de actina y las cadherinas se ve seriamente afectado, así como la transmisión de fuerza desde el citoesqueleto a las mismas y la tensión intercelular en las uniones adherentes (Bard et al., 2008; Liu et al., 2010; Borghi et al., 2012). La  $\beta$ -catenina ha sido utilizada como marcador de acumulación de complejo cadherina/ $\beta$ -catenina en AJ sometidas a estrés mecánico e involucradas en procesos de mecanotransducción (Ladoux et al., 2010; le Duc et al., 2010), y la eliminación del sitio de unión a la misma en cadherinas, o a  $\alpha$ -catenina compromete la correcta formación de contactos celulares por AJ y su estabilidad ante estrés mecánico (Abe y Takeichi, 2008; Desai et al., 2013; Ray et al., 2013).

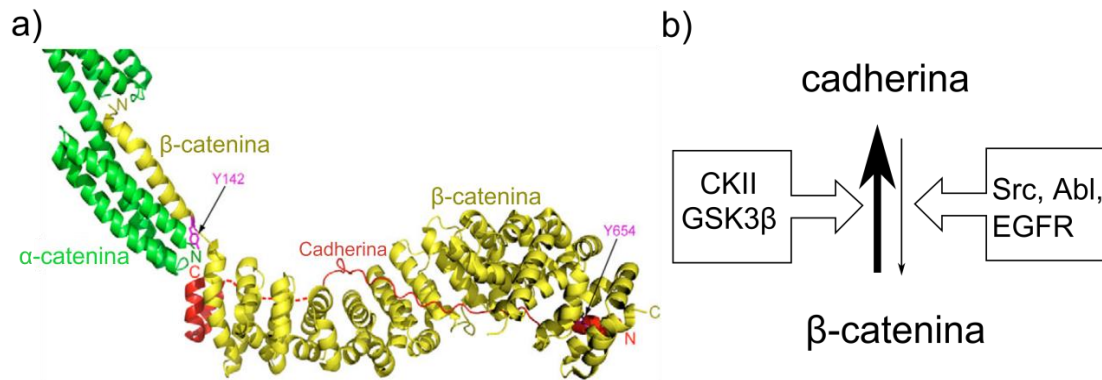
Más allá del papel como mero transmisor de fuerzas en la mecanotransducción local que ocurre en las AJ, la  $\beta$ -catenina está implicada en mecanismos de mecanotransducción derivados de su papel como cofactor de transcripción (**figura 1.7.d**). En las células existen dos reservorios de  $\beta$ -catenina, uno soluble en citosol y otro en membrana formando parte del complejo cadherina/ $\beta$ -catenina. La reserva citosólica se mantiene constitutivamente en niveles bajos, gracias a la fosforilación doble que sufre por la acción de las quinasas CKI y GSK-3 $\beta$  (unida a un complejo de andamiaje axina y APC) que la marca para su destrucción por el complejo ubiquitina-proteasoma. La activación de la vía de señalización Wnt (por unión de ligandos Wnt a los receptores Frizzled o LRP-5/6), provoca la inhibición de GSK-3 $\beta$ , y como consecuencia la estabilización y

acumulación citosólica de  $\beta$ -catenina que finalmente se transloca al núcleo y activa la expresión génica (Nelson y Nusse, 2004). En un primer momento podría parecer que ambos reservorios están funcionalmente separados; sin embargo existen mecanismos de mecanotransducción que originan la liberación de  $\beta$ -catenina del complejo cadherina/ $\beta$ -catenina y su translocación a núcleo. Durante el desarrollo de *Drosophila* y pez cebra, la compresión que sufren las células en el embrión provoca la fosforilación de la tirosina 667/654 de  $\beta$ -catenina, con la consecuente pérdida de interacción con cadherina, translocación al núcleo e inducción de genes relacionados con la determinación de la identidad celular (Desprat et al., 2008; Brunet et al., 2013) (**figura 1.7.d**). En osteoblastos en cultivo, el estrés mecánico origina la estabilización de  $\beta$ -catenina por inhibición de GSK-3 $\beta$  y posiblemente la liberación de la misma de un reservorio (el complejo cadherina/ $\beta$ -catenina de la membrana caveolar), que conduce finalmente a la translocación nuclear de  $\beta$ -catenina y expresión de los genes *Wisp* y *Cox2* (implicados en la diferenciación celular a osteoblastos y osteogénesis) (Case et al., 2008b).

En base a su complejo comportamiento, para poder llegar a tener una comprensión profunda de la fisiología molecular de la  $\beta$ -catenina (cómo elemento pasivo en la transmisión de fuerza en las AJ o en los procesos de mecanotransducción que derivan en su translocación nuclear) es necesario conocer sus propiedades mecánicas. Por ello, en esta tesis doctoral se ha abordado el estudio del comportamiento ante estrés mecánico de la  $\beta$ -catenina mediante AFM-SMFS.

### **1.4.1.1. $\beta$ -catenina: regulación y estructura.**

Como se ha comentado, la  $\beta$ -catenina es una proteína esencial en las uniones adherentes y en la señalización de la vía Wnt. Se trata de una proteína multifuncional, con más de 20 compañeros diferentes de interacción, y que juega papeles esenciales en el desarrollo embrionario y en la homeostasis tisular en adultos (su desregulación está relacionada con varios tipos de cáncer) (Huber et al., 1997; Xu y Kimelman, 2007; Xing et al., 2008). La regulación por fosforilación de la  $\beta$ -catenina no se circunscribe solo al control de su degradación por la vía Wnt, sino que también afecta a los complejos cadherina/cateninas (entendido como el complejo cadherina/ $\beta$ -catenina/ $\alpha$ -catenina) y en consecuencia a la adhesión celular. En general la fosforilación en treonina/serina por GSK3 $\beta$  y CKII en  $\beta$ -catenina o el dominio citosólico de la cadherina se traduce en una estabilización del complejo. Por otra parte, la fosforilación de tirosinas en la  $\beta$ -catenina, en Tyr<sup>489</sup> o Tyr<sup>654</sup> (por las quinasas Abl, Src, EGFR) promueve la pérdida de interacción con las cadherina y en Tyr<sup>142</sup> (por la quinasa Fer) la disminución de afinidad por la  $\alpha$ -catenina (Nelson y Nusse, 2004; Nelson, 2008) (**figura 1.8**).



**Figura 1.8.- Estructura de la  $\beta$ -catenina y regulación por fosforilación.** **a)** Figura adaptada de (Xu y Kimelman, 2007) donde se representa la región central armadillo de la  $\beta$ -catenina (amarillo) unida a la región intracelular de cadherinas (rojo) y a la  $\alpha$ -catenina (verde). **b)** Regulación por fosforilación de la  $\beta$ -catenina. La fosforilación en treonina/serina por GSK3 $\beta$  y CKII aumenta la afinidad de la  $\beta$ -catenina por el dominio citosólico de las cadherinas, mientras que la fosforilación de tirosina 654 en  $\beta$ -catenina por Src, Abl, y EGFR origina la pérdida de afinidad.

La estructura de la  $\beta$ -catenina puede dividirse en las siguientes tres regiones. La región N-terminal (N-term) contiene 150 residuos desestructurados que pueden formar una  $\alpha$ -hélice al final de dicha zona. A continuación, la región armadillo (ARM) contiene 515 residuos aminoacídicos y está compuesta por 12 repeticiones tipo ARM, con  $\sim 42$  residuos cada una, formando tres hélices  $\alpha$  en forma de triángulo (H1, H2 y H3), con la excepción de dos repeticiones (R7, que carece de la primera hélice  $\alpha$  (H1), y R10, que presenta un lazo entre H2 y H3 con unos 20 residuos por encima de la media). Debido a la interacción entre las repeticiones, cada una de ellas se apila encima de la otra, encajando el tercer residuo de una hélice en el cuarto hueco hidrófobo formado por la misma hélice de la siguiente repetición, adoptando una estructura helicoidal (con  $30^\circ$  de rotación promedio y paso de 1 nm por repetición). Finalmente, el comienzo de la región C-terminal de la  $\beta$ -catenina muestra una  $\alpha$ -hélice que interacciona con la región armadillo, protegiendo la parte exterior de R12 (Huber et al., 1997; Huber y Weis, 2001; Xing et al., 2008). Dado que las regiones terminales en gran parte de su extensión no presentan una estructura estable y determinada, existe la posibilidad de que existan interacciones de carácter transitorio entre los terminales y la región ARM debido a la diferencia de carga para un mismo pH ( $pI_{\text{term}} \sim 4$ ,  $pI_{\text{ARM}} \sim 8$ ). Las regiones desestructuradas podrían tener algún papel en la regulación de las interacciones de baja afinidad entre la región ARM y diversas proteínas, pero los substratos que interaccionan fuertemente con el surco de la super-hélice ARM no ven disminuida su afinidad cuando las regiones terminales están presentes (Piedra et al., 2001; Castaño et al., 2002; Choi et al., 2006; Xing et al., 2008). El dominio citosólico de la cadherina interacciona prácticamente con toda la región ARM de la  $\beta$ -catenina, mientras que la región de unión de la  $\alpha$ -catenina se encuentra en la cola amino-terminal, en la

región justamente adyacente a la primera repetición ARM (Shapiro y Weis, 2009) (**figura 1.8.a**).

#### **1.4.2. - Estudio de interacciones intermoleculares entre proteínas por AFM: hacia una estrategia autocontrolada.**

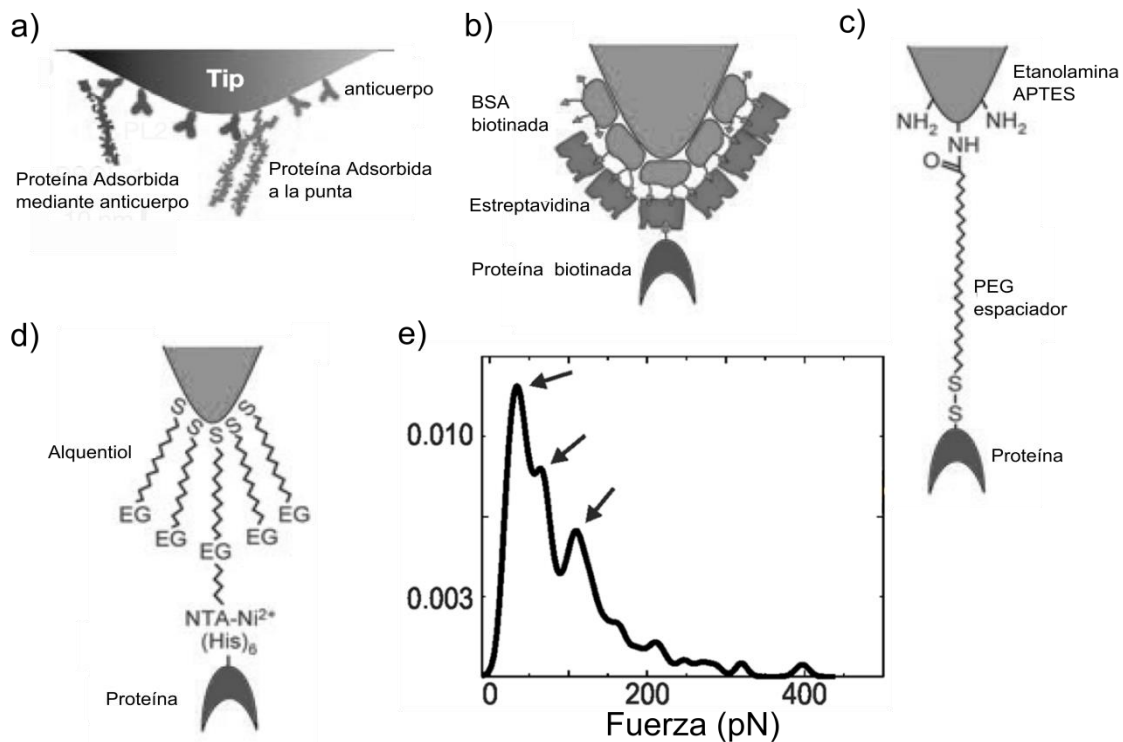
Cuando estudiamos una interacción entre proteínas por AFM, la aplicación de fuerza al enlace termina con la disociación de la interacción, y tras la ruptura de la misma, la punta y la muestra continúan separándose. Así pues, el proceso de disociación inducido por fuerza es, a efectos prácticos, irreversible. El estado unido no puede considerarse como un estado en equilibrio térmico en el momento en que la fuerza está presente, es pues un estado metaestable. El estudio de la disociación de un enlace por fuerza supone en esencia medir el tiempo de vida medio de un estado metaestable ( $k_U(F)$ ) (ver **Ec. 4**).

Cómo ya se ha explicado, podemos encontrar una expresión para  $k_U(F)$  mediante un modelo basado en la teoría del estado de transición (Bell, 1978; Evans y Ritchie, 1997; Evans y Ritchie, 1999) (ver **sección 1.3.3**, y **figura 1.4**). Los cálculos experimentales de  $k_U(F)$  y de  $k_U(0)$  (fuerza cero) requieren obviamente de la correcta identificación de los eventos de ruptura en los experimentos de AFM-SMFS, y también de que aseguren condiciones de molécula individual (de una sola interacción realmente).

Sin embargo, a día de hoy, la identificación directa del evento de ruptura y la certeza de que se trata de la interacción de un solo par de moléculas suponen un auténtico desafío (Merkel, 2001; Zhu et al., 2002; Weisel et al., 2003; Hinterdorfer y Dufrene, 2006; Valle et al., 2008; Fuhrmann y Ros, 2010; Han et al., 2012; Safenkova et al., 2012). La estrategia utilizada comúnmente implica complejos protocolos de funcionalización de las proteínas a las superficies, con el objeto de minimizar la adhesión inespecífica entre la punta y el sustrato, y utilizar bajas concentraciones de proteínas para favorecer estadísticamente la captura de un solo par de moléculas interaccionantes (Hinterdorfer, 2005; Hinterdorfer y Dufrene, 2006; Zimmermann et al., 2010; Safenkova et al., 2012).

Los protocolos de funcionalización tienen que cumplir una serie de requisitos para ser realmente útiles:

1. La unión de las moléculas a las superficies debe ser mucho más estable que la interacción intermolecular a estudiar (Berquand et al., 2005). Esto se requiere para evitar que el pico de ruptura detectado se deba a la pérdida de adhesión de cualquiera de las proteínas de su superficie. La estrategia óptima para asegurar este requisito reside en unir covalentemente las moléculas a las superficies, ya que los enlaces covalentes son del orden de 10 veces más mecanoestables, 1-2 nN (Grandbois et al., 1999), que un enlace ligando-receptor típico. Es importante que el proceso de inmovilización no genere heterogeneidad



**Figura 1.9.- Ejemplos de estrategias de funcionalización en experimentos de medida de interacciones intermoleculares e interpretación de los mismos. a)** Proteínas adsorbidas a la punta directamente o por medio de anticuerpos; no se utiliza ningún bloqueante de la superficie. **b)** adsorción de las proteínas por medio del sistema no covalente biotina-estreptavidina. **c)** y **d)** las superficies son bloqueadas con APTES, etanolamina o alquentioles. Se utiliza una molécula espaciadora que aumenta la libertad del ligando y lo aleja de la superficie (polietilenglicol, PEG). En **c)** la proteína se une covalentemente al espaciador, en **d)** se utiliza el sistema NTA- Ni<sup>2+</sup>-Histag mucho menos estable que un enlace covalente (Valle et al., 2008) **e)** Distribución de fuerzas mostradas en (Rangl et al., 2013) para una interacción proteína-proteína. En la distribución se intuyen tres máximos (flechas) que se relacionan con la ruptura de múltiplos de un evento, que representa la ruptura de una sola interacción. La figura a) adaptada de (Marshall et al., 2003) b), c) y d) de (Hinterdorfer y Dufrene, 2006).

adicional en la población de moléculas, sin embargo los métodos de fijación covalente no son específicos de sitio, utilizándose inmovilización a través de grupos amino de lisinas o sulfhidrilos de cisteínas (Baumgartner et al., 2000; Berquand et al., 2005; Schoeler et al., 2014), que no son por lo tanto de aplicación general. Es por esto que en muchas ocasiones se utilizan procedimientos de inmovilización que se basan en métodos no covalentes, como la mera fisisorción, el uso de anticuerpos, el sistema avidina-biotina o el de NTA-Ni<sup>2+</sup>-Histag (Marshall et al., 2003; Kong et al., 2009; Rakshit et al., 2012; Rangl et al., 2013) (figura 9.a-d).

2. Hay que minimizar el efecto de la inmovilización en las proteínas. Las proteínas deben preservar su estructura nativa y mantener suficiente movilidad para que puedan interactuar libremente con las moléculas complementarias. Esto puede conseguirse uniendo las moléculas a la superficie vía un espaciador molecular flexible (como el polietilenglicol, PEG). El espaciador permite la libre orientación de las moléculas en el

espacio, y aleja a las proteínas de los efectos del material de soporte o la punta. (Baumgartner et al., 2003; Nevo et al., 2003; Hinterdorfer y Dufrene, 2006; Morfill et al., 2007) (**figura 9.c-d**).

3. Minimización de las interacciones no específicas entre la punta y el soporte. Este punto es de vital importancia ya que las interacciones no específicas pueden estar presentes en el 15-85 % de las curvas de fuerza, dependiendo del complejo estudiado y de protocolo de funcionalización (Marshall et al., 2003; Ritco-Vonsovici et al., 2007; Safenkova et al., 2012; Rakshit et al., 2012). Los pasos de funcionalización de ambas superficies, suelen incluir pasos de modificación de las mismas con silanos, etanolamina, o alquencios (Baumgartner et al., 2003; Hinterdorfer, 2005; Berquand et al., 2005; Stahl et al., 2012), que disminuyen las interacciones inespecíficas de adhesión. Los espaciadores moleculares también pueden utilizarse para minimizar la adhesión (Morfill et al., 2007; Morfill et al., 2008) (**figura 9.c-d**).

4. La densidad de moléculas debe ser suficientemente baja para intentar asegurar la captura de SM. El método de funcionalización tiene pues que permitir regular la densidad de moléculas en superficie (Zimmermann et al., 2010). Sin embargo, la baja densidad de moléculas no asegura la situación de SM. Al no existir control de SM en la propia muestra, podríamos estar estudiando agrupaciones de moléculas, lo que es especialmente relevante en moléculas que tienden a dimerizar o agregar, y que pueden mantener este estado pese a utilizar muestras muy diluidas (Zhu et al., 2002). En ocasiones, se pueden observar en los histogramas de fuerzas picos que parecen coincidir con múltiplos de una unidad mínima (Florin et al., 1994; Baumgartner et al., 2000; Berquand et al., 2005; Rangl et al., 2013) (**figura 9.e**). Esto se ha interpretado desde un punto de vista determinista como que la unidad mínima representa la estabilidad de un único enlace, y que  $n$  enlaces en paralelo tendrán  $n$  veces esa estabilidad. En líneas generales esta interpretación puede no mantenerse a tasas de carga menores, debido a la naturaleza estocástica del proceso y a su dependencia con la tasa de carga (Evans y Ritchie, 1997; Evans y Ritchie, 1999; Weisel et al., 2003). En relación a esto, la fuerza de disociación de una sola pareja de moléculas a una velocidad determinada (suceso elemental) podría encontrarse por debajo del umbral de detección del aparato, pero no así el de agrupaciones de las mismas (ya que la fuerza se “reparte” entre varias parejas). Así pues podríamos identificar erróneamente el evento unitario de la distribución de fuerzas como una sola interacción debiéndose en realidad a agrupaciones de moléculas.

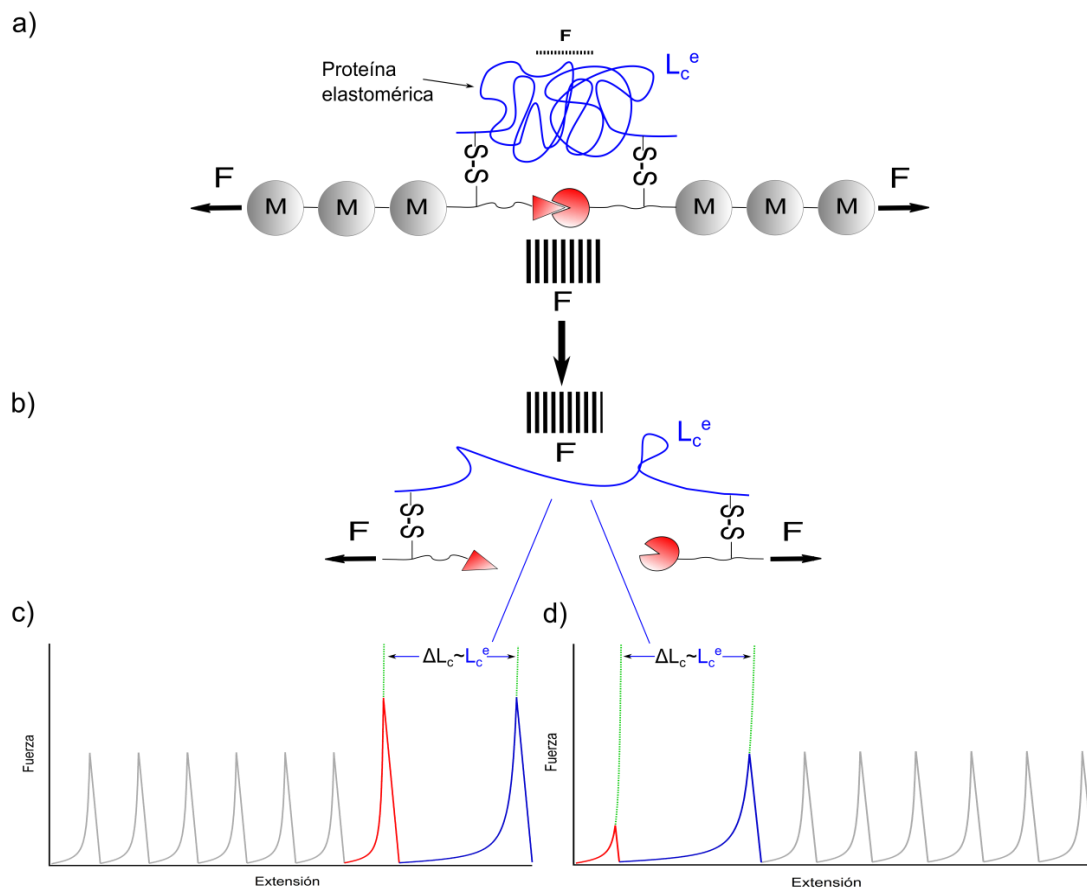
Los métodos de funcionalización presentan además inconvenientes de tipo práctico. La calidad y reproducibilidad de los pasos de preparación de la punta y el sustrato son factores clave que gobiernan el éxito y la fiabilidad del experimento

(Berquand et al., 2005). Para cada nuevo sistema la calidad de las modificaciones de las superficies debe ser asegurada, y es necesario validar la calidad de la punta utilizando habitualmente modelos planos que son tratados con las puntas en paralelo y luego han de ser caracterizados (Hinterdorfer y Dufrene, 2006). Teniendo en cuenta que suelen ser protocolos muy complejos con hasta más de 40 pasos (Zimmermann et al., 2010) solucionar un problema con la funcionalización puede ser una tarea nada sencilla. Una vez funcionalizadas las muestras (puntas y sustratos) han de mantenerse hidratadas y en condiciones donde retengan su funcionalidad.

Un punto clave de estos experimentos, es que precisan la realización obligada de experimentos control adicionales para demostrar la especificidad de las fuerzas medidas. Esto suele hacerse mediante ensayos de bloqueo (usando ligandos/anticuerpos bloqueantes en solución (Baumgartner et al., 2000; Nevo et al., 2003; Berquand et al., 2005; Ritco-Vonsovici et al., 2007; Kong et al., 2009)), o eliminando/sustituyendo por una proteína inerte una de las proteínas interaccionantes (Marshall et al., 2003; Zhang et al., 2004; Morfill et al., 2007; Morfill et al., 2008). Hay que tener en mente que estos experimentos de control nunca muestran una falta total de interacción entre punta y sustrato, sino que simplemente ponen de manifiesto un umbral mínimo inevitable de contaminación en los datos.

Como resumen general, estos métodos de funcionalización son complejos y costosos experimentalmente, y no demuestran de por sí la especificidad de las fuerzas medidas (Han et al., 2012), necesitándose experimentos control o incluso, en ocasiones, marcadores de desplegamiento de la propia proteína para detectar el evento de disociación (Stahl et al., 2012; Kim et al., 2012; Schoeler et al., 2014).

El uso de poliproteínas supuso un hito en los estudios de AFM-SMFS. Éstas, introducían un control interno de monomolecularidad en la muestra, lo que permitía la identificación del desplegado de los módulos proteicos de manera inequívoca. El caso de los estudios de interacciones intermoleculares es muy diferente. En los protocolos al uso no existe ningún criterio que asegure que se está estudiando un solo par de moléculas interaccionantes, ni manera de identificar el evento de disociación en los registros. Uno de los trabajos desarrollados en esta tesis se centra en implementar, en un sistema modelo, un diseño experimental que presente controles internos tanto de monomolecularidad, como de identificación de la ruptura de la interacción, y demostrar así que una estrategia autocontrolada total también es posible en estudios de interacciones intermoleculares (**figura 1.10**).



**Figura 1.10.- Estrategia internamente controlada para el estudio de interacciones intermoleculares. a)** Esquema de la estrategia propuesta. El control de monomolecularidad viene dado por la inclusión de pequeñas poliproteínas en las construcciones a estudio. Como marcador del pico de ruptura de la interacción se utiliza un tercer polipéptido: un polipéptido elastomérico que se estira sin barreras mecánicas por AFM-SMFS (ej.: proteínas desestructuradas que formen ovillos estadísticos). Este polipéptido tiene que ser lo suficientemente flexible como para que la interacción soporte la mayor parte de la tensión generada por el estiramiento del sistema (ver **sección 4.2**). Los módulos marcadores de monomolecularidad (M) se indican en gris, en rojo la interacción a estudio y en azul el polipéptido puente. **b)** Tras la ruptura de la interacción, el elastómero queda expuesto a la fuerza, y se libera un  $\Delta L_c$  proporcional a la longitud de contorno del mismo  $L_c^e$ . **c)** la liberación a la fuerza del polipéptido elastomérico sirve de marcador del pico de fuerza proveniente de la ruptura de la interacción, tanto en el caso de que la estabilidad de la interacción sea mayor que la de los módulos marcadores (izquierda), como en el caso contrario (derecha).

A partir de ahora por claridad y simplicidad cuando nos refiramos a interacciones intermoleculares hablaremos de situación de molécula individual para referirnos al estudio de un solo par de moléculas interaccionantes.

#### 1.4.2.1. - Estrategia autocontrolada propuesta.

La estrategia propuesta en este trabajo, reside en utilizar dos identificadores diferentes para asegurar molécula individual y marcar de manera directa el pico de ruptura de la interacción en los registros experimentales. Para asegurar el criterio de molécula individual, utilizaremos la estrategia de las poliproteínas. Para marcar el pico de



fuerza originado por la ruptura de la interacción se utilizará la liberación de una longitud de contorno que identifique el pico de fuerza, de manera análoga a lo que ocurre en experimentos de SMFS con heteropoliproteínas (en el que el  $\Delta L_c$  identifica el módulo a estudio) (**figura 1.10.a** y **figura 1.6.a**).

La manera más directa de incorporar la estrategia de las poliproteínas al estudio de interacciones intermoleculares por AFM sería incluir mediante ingeniería genética repeticiones de un módulo marcador acompañando a cada uno de los elementos del par proteico a estudio (**figura 1.10.a**). De esta forma las repeticiones actuarían de marcadores o reporteros que identificarían a cada una de las moléculas. Esto aunque aportaría ventajas, sería insuficiente para generar los dos controles necesarios. Si la interacción que estamos estudiando tiene una estabilidad mecánica inferior a la que presenten los módulos marcadores, el circuito mecánico se romperá antes de producirse el desplegamiento de los módulos marcadores, y por lo tanto perderíamos nuestro marcador de monomolecularidad. En el caso de que la interacción a estudio tuviera una estabilidad mecánica superior a la de los módulos marcadores, seríamos capaces de observar el patrón de dientes de sierra del marcador, indicándonos que estamos estirando una molécula. Hay que advertir que al observar un número de picos que fuera mayor que el número de repeticiones de cada molécula por separado (en el caso del ejemplo de la **figura 1.10.a** serían entre cuatro y seis módulos), podríamos afirmar, debido a la geometría del sistema, que estamos sometiendo a fuerza a la interacción. En este caso uno estaría tentado a pensar que el último pico de fuerza observado, que representa el fallo del circuito mecánico que se establece entre la punta-muestra vía las proteínas interaccionantes, representa la ruptura de la interacción. No obstante el último pico de fuerza puede también representar la ruptura del circuito mecánico por la desorción de alguna de las proteínas, bien de la punta o bien el sustrato.

Estos problemas pueden solucionarse con el uso de un tercer polipéptido que se una a las proteínas interaccionantes y asegure el mantenimiento del circuito mecánico hasta que se produzca el desplegamiento de los módulos marcadores de molécula individual. En la **figura 1.10.a** se representa esta idea, que en esencia pretende mantener una continuidad mecánica en el sistema mediante el uso de este tercer polipéptido (en azul en la **figura 1.10**), que se una entre las proteínas interaccionantes a estudio y los marcadores poliproteicos. Es importante tener en cuenta que este polipéptido “puente”, ha de tener una longitud y flexibilidad suficientes para que cuando se ejerza fuerza en el sistema, ésta se transmita principalmente a través de la interacción y no a través de este “puente”. El candidato ideal para actuar de molécula de unión es un polipéptido que se comporte en estudios mecánicos de AFM como un elastómero sin resistencia mecánica

(típico de proteínas desestructuradas no amiloidogénicas (Li et al., 2001; Li et al., 2002; Lv et al., 2010)). Estas moléculas no presentan puntos detectables de resistencia a la fuerza (y por tanto no presentan picos de fuerza en AFM-SMFS), y sus curvas fuerza-extensión ajustan perfectamente al modelo WLC. Una proteína con estas características, si es suficientemente grande, presenta un perfil de aumento de la fuerza prácticamente plano en los primeros nanómetros (ver **sección 4.2**). Al disponer este polipéptido en paralelo con la interacción, la fuerza se transmitiría mayoritariamente a través de la interacción al estirarlos conjuntamente. Esto se puede aprovechar como una ventaja añadida, ya que además de mantener una continuidad en el sistema, el polipéptido quedaría en esencia oculto a la fuerza hasta que no se produjera la ruptura de la interacción. La longitud liberada por el estiramiento del polipéptido (con un  $L_c$  característico) tras la ruptura de la interacción, puede usarse como un marcador o reportero del pico de fuerza (**figura 1.10.b** y **figura 1.10.c**), ya que deberíamos observar un  $\Delta L_c \approx L_c$  polipéptido puente -  $L_c$  de la interacción.

El mantenimiento del circuito mecánico permitiría el desplegamiento de los módulos marcadores de monomolecularidad también en el caso de que la estabilidad mecánica de la interacción fuera menor que la de los módulos marcadores. Por tanto, en este caso, como en el caso de que la interacción tuviera una estabilidad mecánica superior a la de los módulos marcadores de molécula individual, permitiría identificar el pico de fuerza debido a la ruptura de la interacción a estudio (que iría seguido del  $\Delta L_c$  característico debido a la presencia del polipéptido “puente”) (**figura 1.10.c**).

Aunque desde un punto de vista conceptual la utilización del tercer polipéptido es realmente ventajoso, supone retos de tipo práctico: ¿cómo conseguir la unión entre las tres proteínas? ¿qué características debe tener el enlace que mantenga el circuito mecánico entre las tres proteínas? Mientras que la interacción a estudio se mantenga, la fuerza a la que se ve sometido el polipéptido puente (y por tanto los enlaces) es despreciable, solo una vez que la interacción se rompe queda expuesto a la fuerza y a través de él (y sus enlaces con las otras dos proteínas) se mantiene el circuito mecánico. Este hecho impone condiciones de estabilidad en los enlaces entre las tres proteínas. En el caso de que la ruptura de la interacción se haya producido después del desplegamiento de los módulos marcadores de monomolecularidad, los enlaces han de mantenerse estables el tiempo necesario para que pueda ser identificable el perfil de  $\Delta L_c$  que genera la exposición a fuerza del polipéptido puente. Cuando la ruptura de la interacción precede al desplegamiento de los módulos marcadores, los enlaces entre las tres proteínas han de mantener el circuito mecánico hasta que sea observable el patrón repetido originado por los marcadores de monomolecularidad. Teniendo en cuenta que la estabilidad mecánica

de los módulos típicamente utilizados como marcadores poliproteicos oscila entre los 150-200 pN (Sharma et al., 2006; Valbuena et al., 2009; Hervas et al., 2012), la estabilidad de las uniones entre los tres polipéptidos ha de estar por encima de este rango. Es por esto que un enlace de tipo covalente resulta ser el candidato ideal, ya que su estabilidad mecánica está muy por encima del rango de 150-200 pN (1-2 nN (Grandbois et al., 1999)).

Una manera sencilla y controlable de inducir la formación de enlaces covalentes entre proteínas es la formación de enlaces disulfuro entre cisteínas. Los enlaces disulfuro se forman por oxidación del grupo tiol de la cadena lateral de residuos de cisteína que se encuentran espacialmente enfrentados. La formación de enlaces disulfuro puede controlarse modificando las condiciones redox del medio (Gough y Lees, 2005) y la posición de las mismas dentro de las proteínas puede seleccionarse mediante técnicas de ingeniería de proteínas, de tal modo que podemos controlar la formación de enlaces covalentes entre las tres proteínas del sistema, en las posiciones que se deseemos, mediante enlaces disulfuro.

## 2. Objetivos

---

En esta tesis doctoral se ha trabajado en dos aspectos fundamentales de la nanomecánica de proteínas mediante AFM-SMFS. Por una parte se ha llevado a cabo el estudio de las propiedades mecánicas de una proteína que forma parte de un sistema biológico en el que la fuerza mecánica es un parámetro esencial: la  $\beta$ -catenina del sistema cadherina/catenina, un componente central de los sistemas de adhesión célula-célula en metazoos. Por otra parte se ha trabajado en el estudio de interacciones intermoleculares entre proteínas mediante AFM-SMFS. Las estrategias clásicas para el estudio de interacciones intermoleculares mediante AFM-SMFS no permiten la identificación directa del pico de fuerza indicativo de la estabilidad mecánica de la interacción. En este trabajo se presenta la prueba de concepto de la viabilidad de una estrategia general autocontrolada que identifica directamente el pico de fuerza de la interacción.

Los objetivos resumidos son los siguientes:

**Objetivo 1.** Caracterización nanomecánica de la proteína  $\beta$ -catenina del sistema cadherina/catenina. Para ello se utilizará una combinación de ingeniería de proteínas, análisis nanomecánico mediante AFM-SMFS y simulaciones de dinámica molecular.

**Objetivo 2.** Interpretación de la relevancia fisiológica de las propiedades mecánicas de la  $\beta$ -catenina en el sistema biológico en el que normalmente lleva a cabo su función: el sistema de adhesión célula-célula de cadherina/catenina.

**Objetivo 3.** Diseño de una estrategia general autocontrolada para la estudio de interacciones intermoleculares por AFM-SMFS. La estrategia debe ser de aplicación general, sin depender de características propias del par de proteínas a estudiar, y debe permitir la identificación de manera directa del pico de fuerza de ruptura de la interacción, a la vez que proporcione controles de monomolecularidad.

**Objetivo 4.** Demostrar experimentalmente, a nivel de prueba de concepto, que la medición de interacciones intermoleculares mediante AFM-SMFS usando una estrategia general autocontrolada es posible. Para ello el diseño será implementado en una interacción modelo y se medirá su estabilidad mecánica asegurándonos de que presente todos los controles introducidos.

## 3. Materiales y Métodos

---

### 3.1.- Clonaje molecular

Las técnicas de clonaje molecular estándar se basan en (Sambrook y Russel, 2001), las enzimas utilizadas para el clonaje se adquirieron en New England Biolabs (Ipswich, MA) y Fermentas (Thermo Fisher Scientific Inc., Waltham, MA) a excepción de la ADN polimerasa *Phusion polymerase* (Fynnzimes, Thermo Fisher Scientific Inc., Waltham, MA), Los oligonucleótidos (u oligos) de cadena sencilla fueron sintetizados por Sigma-Aldrich (St. Louis, MO), y los dNTPs para PCR fueron adquiridos en Invitrogen (Thermo Fisher Scientific Inc., Waltham, MA). Para todas las etapas de clonación se utilizó la cepa XL1 Blue (Stratagene, Agilent Technologies Inc, Santa Clara, CA) de *Escherichia coli*.

Cada vez que ha sido necesario generar *in vitro* oligonucleótidos de cadena doble a partir de oligonucleótidos complementarios de cadena sencilla, se ha seguido el siguiente procedimiento: los oligonucleótidos eran resuspendidos en agua MilliQ a una concentración de 100  $\mu\text{M}$  y mezclados en un tubo Eppendorf de 1,5 ml. Con ayuda de un flotador, el Eppendorf era introducido en un vaso de precipitados (1 litro) con agua hirviendo durante 5 minutos y finalmente el vaso de precipitados (con el Eppendorf flotando) se dejaba enfriar a temperatura ambiente.

Los clones utilizados en la **sección 4.1** ( $\beta$ -catenina) se generaron de la siguiente forma:

Para generar la construcción ARM/I27, las repeticiones ARM de la  $\beta$ -catenina humana (residuos 134-671) fueron clonadas en el vector pRSETA-(I27)<sub>8</sub> (Steward et al., 2002). Los sitios de restricción usados para clonar fueron KpnI y MluI y el molde usado para la amplificación de la región ARM fue un plásmido pQE32 (Qiagen GmbH, Hilden, Germany) que contenía la  $\beta$ -catenina completa (Behrens et al., 1996). La construcción  $\beta$ /I27 se generó clonando entre los sitios KpnI y MluI de pRSETA-(I27)<sub>8</sub> la  $\beta$ -catenina completa utilizando como molde de PCR el mismo vector.

El vector pRSETA-(I27)<sub>8</sub> fue amablemente cedido por la Dra. Jane Clarke (Universidad de Cambridge), y el plásmido pQE32 que portaba la secuencia de la  $\beta$ -catenina (unida a una etiqueta de poliHistidinas) por la Dra. Amparo Cano (Universidad Autónoma de Madrid). Ambos plásmidos fueron utilizados directamente para producir las proteínas (I27)<sub>8</sub> y  $\beta$ -cat.

El constructo GST\_Ncito fue creado clonando por PCR entre las dianas BamHI y NotI del plásmido pGEX-5x-1 el fragmento citosólico de la N-cadherina murina (residuos 746-906), utilizando como molde el vector pCMV-SPORT6 DCH2 (RZPD - *German Science*

Centre for Genome Research). En el oligonucleótido NotI se incluyó una secuencia codificante para una etiqueta de 6 histidinas. El clon utilizado para producir la proteína GST se originó digiriendo el plásmido pEX-5x-1 entre las dianas BamHI y NotI, rellenando los extremos cohesivos mediante la polimerasa *Phusion Polymerase* (Fynnzymes), y recircularizándolo. Esto origina un codón de parada justo después del gen de la GST.

Los clones utilizados en la **sección 4.2** (interacciones intermoleculares) se generaron de la siguiente forma:

En primer lugar, notar que los dominios de oligomerización ‘Velcro’ utilizados tienen la siguiente secuencia aminoacídica:

**Ácido1:** AQLKELQANEKELAQLEWELQALEKELAQ

**Ácido2:** AQLKELQALEKELAQLEWENQALEKELAQ

**Básico:** AQLKKKLQALKKKLAQLKWK~~N~~QALKKKLAQ

Para generar las construcciones Cohesina Básico (I27)<sub>3</sub> y Dockerina Básico (I27)<sub>3</sub> se siguió el siguiente esquema. La secuencia del dominio Básico fue clonada mediante PCR del molde pVS16N (amablemente cedido por la Dra. Martha Oakley de la Universidad de Indiana (Kim y Oakley, 2002)) entre las dianas SallI y NotI del plásmido pET-24d (Novagen, Merck Millipore, Darmstadt, Alemania). Se incluyó la secuencia aminoacídica GGGCGGG en el oligonucleótido portador la diana NotI, esta secuencia porta la cisteína encargada de formar el puente disulfuro que da continuidad covalente a todo el complejo, y varias glicinas que aportan flexibilidad y evitan la distorsión de la hélice del dominio de oligomerización (O’Shea et al., 1993). Sobre las dianas NotI y XhoI de este vector, se clonaron por PCR las 3 repeticiones de módulos de I27, (I27)<sub>3</sub>, incluidas entre las dianas NheI y MluI del vector pRSETA-(I27)<sub>3</sub>. De hecho, se generaron dos construcciones diferentes. Una construcción como la descrita, sobre la que posteriormente se clonó la secuencia codificante de la Cohesina 7 de la escafoldina de *Clostridium thermocellum*, y una segunda muy similar, en la que se incluyeron dos cisteínas terminales después del último I27 en la diana XhoI (para la unión a las puntas recubiertas de oro del AFM, ver **sección 3.10**). Sobre este segundo plásmido se clonó la Dockerina de la celulasa Cel8A del celulosoma de *C. thermocellum*. Una vez clonados los dominios de oligomerización y los módulos marcadores, se procedió a clonar las secuencias codificantes de Cohesina y Dockerina entre las dianas NcoI y EcoRI para generar las construcciones Cohesina Básico (I27)<sub>3</sub> y Dockerina Básico (I27)<sub>3</sub> respectivamente. Los insertos de Cohesina y Dockerina se clonaron por PCR de los plásmidos p(I27)<sub>3</sub>-C7a-(I27)<sub>2</sub> (Valbuena et al., 2009), y del plásmido pET28-CelA (amablemente cedido por el Prof. Ed Bayer del Instituto Weizmann) respectivamente.

La familia de vectores **poliR\_MBP** está constituida por los vectores pETpolyR\_MBP y pETpolyR\_MBP\_tev. El vector pETpolyR\_MBP es un vector pET-24d (Novagen) modificado, en el que se ha introducido en la región previa a la primera diana de restricción (NcoI) del sitio de clonaje múltiple (MCS, del inglés *multi cloning site*) la etiqueta polyR\_MBP seguida de una diana de la proteasa TEV (GENLYFQ/A). La etiqueta polyR\_MBP está constituida por 5 residuos de arginina seguidos de la proteína de unión a la maltosa MBP (MARRRRRGKIEEGKL...). Como resultado, las secuencias que se inserten en el vector pET modificado se traducirán con la etiqueta polyR\_MBP en su región N-terminal, que puede eliminarse una vez purificada la proteína gracias a la diana de la TEV proteasa. Para llevar a cabo la modificación, se generó mediante rondas de PCR solapante, un inserto de ADN flanqueado por las dianas XbaI NcoI que incluía la región del esqueleto del vector pET-24d (Novagen) desde la diana XbaI hasta el inicio del MCS, seguido de la etiqueta polyR\_MBP y terminando en una diana de la TEV proteasa (GENLYFQ/A). Este inserto se clonó en el vector pET-24d (Novagen) entre las dianas XbaI y NcoI para generar el vector pETpolyR\_MBP. De manera detallada: se generaron por PCR dos fragmentos de ADN con secuencias parcialmente solapantes, uno utilizando como molde el plásmido pET-24d (Novagen) y otro utilizando un plásmido que portaba el gen de la proteína MBP (pRK793 (Kapust et al., 2001)). El primer fragmento incluye la región del esqueleto del vector pET-24d (Novagen) desde la diana XbaI a la diana NcoI, introduciendo una mutación en la diana NcoI (c~~g~~atgg) y terminando la región 3' con la secuencia de la etiqueta polyR y los primeros nucleótidos de la MBP. El segundo fragmento empieza en el 5' con la etiqueta polyR y toda la secuencia de la proteína MBP y termina con una diana de la TEV proteasa seguida de la diana NcoI. Utilizando como molde estos fragmentos se generó, mediante una segunda ronda de PCR, el inserto de ADN descrito flanqueado por las dianas XbaI NcoI. La digestión con proteasa TEV deja incorporada a la proteína de interés un residuo de alanina (A), justo antes de la metionina codificada por el codón atg de la diana NcoI (ccatgg).

Partiendo del vector pETpolyR\_MBP se construyó el vector pETpolyR\_MBP\_tev, clonando en la diana XhoI una segunda diana de la proteasa TEV seguida de un módulo I27. Al incluir la segunda diana TEV en el plásmido se pretende poder eliminar la cola de histidinas del mismo una vez utilizada para su purificación, con el fin de optimizar la estrategia de funcionalización para AFM (ver **sección 3.10**). El módulo de I27 se incluye simplemente como un marcador para testar fácilmente que la segunda diana de la proteasa TEV se procesa adecuadamente, ya que al ser cortada podremos observar un cambio significativo en la movilidad electroforética (de unos 10 kDaltons) de la proteína. El módulo I27 se clonó por PCR de un plásmido portador de un mutante (Y9P) del módulo

I27 (Li et al., 2000). En el oligonucleótido 5' se incluyó la secuencia entre las dianas NotI y XhoI del plásmido pET24-d (Novagen), seguido de la diana TEV rodeada de una serie de residuos aminoácidos que aportan flexibilidad y accesibilidad a la diana (GGSAGENLYFQ/GAGGSGGS), y en el oligonucleótido en 3' una diana Sall. Tras digerirse con las enzimas NotI y Sall, el fragmento se clonó entre las dianas NotI y XhoI del plásmido pETpolyR\_MBP. Las dianas Sall y XhoI son cohesivas entre sí pero no reconstituyen ninguna de las dos dianas, de modo que el inserto queda integrado en el esqueleto plasmídico de pETpolyR\_MBP. La inclusión de las dianas NotI y XhoI en el 5' del inserto, permite conservar las dianas NotI y XhoI del MCS de pETpolyR\_MBP.

La construcción Ácido1<sub>12XR</sub>Ácido2 se generó como sigue. Mediante el ensamblaje *in vitro* de oligonucleótidos complementarios de cadena sencilla, se generó un fragmento de ADN codificante para dos repeticiones del mutante (Y8T) de la repetición consenso de resilina utilizada en (Lv et al., 2010). De la misma manera se ensamblaron insertos que codificaban para los dominios de oligomerización Ácido1 y Ácido2. Utilizando como molde el inserto de dos repeticiones de resilina, se generaron por PCR dos fragmentos con la siguiente estructura: BamHI\_(resilina)<sub>2</sub>BglII; BamHI\_(resilina)<sub>2</sub>Sall. Utilizando como molde los fragmentos de resilina y de Ácido1, se sintetizó mediante PCR solapante un inserto de ADN codificante de BamHI\_Ácido1\_(resilina)<sub>2</sub>BglII, que incluía en el extremo 5' a continuación de la diana BamHI la secuencia aminoácida GGGCGGG. El inserto BamHI\_(resilina)<sub>2</sub>BglII se digirió solamente con la enzima BglII o solamente con BamHI, BamHI\_(resilina)<sub>2</sub>Sall con BamHI, y BamHI\_Ácido1\_(resilina)<sub>2</sub>BglII con BglII, y finalmente todos se fosforilaron con la enzima T4 fosfonucleótido quinasa. Esto deja un extremo romo y otro cohesivo (BamHI y BglII son cohesivos entre sí) en cada fragmento. Los fragmentos se clonaron por parejas en el plásmido T7blue (Novagen) digerido con SmaI (extremos romos) en la siguiente combinación: BamHI\_Ácido1\_(resilina)<sub>2</sub>**BglII** + **BamHI**\_(resilina)<sub>2</sub>BglII; BamHI\_(resilina)<sub>2</sub>**BglII** + **BamHI**\_(resilina)<sub>2</sub>BglII; BamHI\_(resilina)<sub>2</sub>**BglII** + **BamHI**\_(resilina)<sub>2</sub>Sall (en negrita la diana de restricción digerida). Esto genera 3 plásmidos: T7 BamHI\_Ácido1\_(resilina)<sub>4</sub>BglII, T7 BamHI\_(resilina)<sub>4</sub>BglII, y T7 BamHI\_(resilina)<sub>4</sub>Sall. Para generar un fragmento de 12 repeticiones de resilina (12XR), primeramente se subclonó el inserto BamHI\_(resilina)<sub>4</sub>BglII en el plásmido T7 BamHI\_(resilina)<sub>4</sub>Sall, creándose el plásmido T7 BamHI\_(resilina)<sub>8</sub>Sall. El fragmento BamHI\_(resilina)<sub>8</sub>Sall se subclonó en el plásmido T7 BamHI\_Ácido1\_(resilina)<sub>4</sub>BglII utilizando la diana BglII del inserto y la diana Sall del MCS de T7blue, lo que genera el plásmido T7 BamHI\_Ácido1\_(resilina)<sub>12</sub>Sall.

En paralelo, la secuencia nucleotídica entre las dianas NcoI y BamHI del plásmido pETpolyR\_MBP\_tev se modificó a ccatggctagcatgactggtggacagcaaatgggtg**cgggatc**, mediante



un fragmento de ADN que se ensambló a partir de dos oligonucleótidos de cadena sencilla, y que tras digerirlo, se clonó entre las dianas NcoI y BamHI. Esto permite que las regiones cercanas a la diana BamHI se traduzcan en los residuos aminoacídicos GAG, y no en los residuos GR, ya que la arginina es un aminoácido básico y voluminoso que podría interferir en el emparejamiento del dominio de oligomerización. Una vez realizada la modificación se clonó por PCR entre las dianas Sall y NotI el dominio Ácido2, incluyéndose en el oligonucleótido NotI la secuencia aminoacídica GGGCGGG. Finalmente para generar la construcción Ácido1<sub>12XR</sub>Ácido2 se subclonó el inserto BamHI\_Ácido1\_(resilina)<sub>12</sub>\_Sall entre las dianas BamHI y Sall de este plásmido.

Para producir la construcción Ácido1\_(Gb1T18P)<sub>5</sub>Ácido2 se siguió el siguiente esquema. Mediante el ensamblaje *in vitro* de oligonucleótidos complementarios de cadena sencilla, se sintetizó un fragmento de ADN codificante para la región central del mutante Gb1T18P; este fragmento se utilizó como molde para generar la secuencia completa mediante PCR, incluyendo los nucleótidos restantes en los oligonucleótidos 5' y 3'. El inserto, se clonó mediante ligación de extremos romos en el plásmido T7blue (Novagen) digerido con SmaI. Este plásmido y el inserto ensamblado Ácido1 se utilizaron como moldes para generar mediante PCR solapante un fragmento BamHI\_Ácido1\_(Gb1T18P)<sub>1</sub>\_SfoI\_Sall que se clonó entre las dianas BamHI y Sall del plásmido T7blue (Novagen). Una vez obtenido el plásmido T7 BamHI\_Ácido1\_(Gb1T18P)<sub>1</sub>\_SfoI\_Sall, en la diana SfoI, se clonó por PCR un inserto SmaI\_Gb1T18P\_SfoI, lo que genera el plásmido BamHI\_Ácido1\_(Gb1T18P)<sub>2</sub>\_SfoI\_Sall. Este último paso se repitió cíclicamente hasta conseguir el plásmido T7 BamHI\_Ácido1\_(Gb1T18P)<sub>5</sub>\_SfoI\_Sall. El inserto BamHI\_Ácido1\_(Gb1T18P)<sub>5</sub>\_SfoI\_Sall, se subclonó entre las dianas BamHI Sall del vector pETpolyR\_MBP\_tev portador del dominio Ácido2, para producir el constructo Ácido1\_(Gb1T18P)<sub>5</sub>Ácido2.

Se produjeron dos variantes de la construcción Básico (I27)<sub>3</sub>. El inserto completo de Cohesina Básico (I27)<sub>3</sub> (entre dianas NcoI XhoI) se subclonó entre las dianas NcoI XhoI del vector pETpolyR\_MBP. La secuencia de cohesina entre las dianas NcoI y EcoRI, se sustituyó por la secuencia GGSG, utilizando dos oligonucleótidos de cadena sencilla que tras ensamblarlos regeneraban los extremos cohesivos NcoI EcoRI y codificaban esta secuencia. Esta construcción representa la primera variante, que no permite retirar la cola de poliHistidinas de la construcción tras la purificación. La segunda variante se generó subclonando el inserto Dockerina Básico (I27)<sub>3</sub> entre dianas NcoI XhoI, en el vector pETpolyR\_MBP\_tev, y posteriormente sustituyendo el inserto de Dockerina entre las dianas NcoI y EcoRI por la secuencia GGSG. La segunda variante permite pues eliminar la

cola de poliHistidinas después de la purificación, dejando cerca del C-terminal dos cisteínas libres para la unión a las puntas recubiertas de oro de las micropalancas.

La construcción Ácido1 (I27)<sub>3</sub> se generó subclonando el inserto completo de Dockerina Básico (I27)<sub>3</sub> entre dianas NcoI XhoI en el vector pETpolyR\_MBP. El inserto de Dockerina, entre dianas NcoI EcoRI se sustituyó por la secuencia GGSG, y posteriormente el inserto Básico entre las dianas Sall NotI por Ácido1, incluyendo por PCR la secuencia aminoacídica GGGCGGG en el oligo Sall. La construcción Ácido1 (I27)<sub>3</sub> no permite la eliminación de la cola de poliHistidinas tras la purificación, y presenta dos cisteínas libres entre los (I27)<sub>3</sub> y la cola de poliHistidinas.

Las construcciones Cohesina Básico (Gb1)<sub>3</sub> y Dockerina Básico (Gb1)<sub>3</sub>, se obtuvieron subclonando el inserto (Gb1)<sub>3</sub> entre las dianas NotI y XhoI de los vectores pETpolyR\_MBP sobre los que se había subclonado Cohesina Básico (I27)<sub>3</sub> y Dockerina Básico (I27)<sub>3</sub>. Para generar el inserto (Gb1)<sub>3</sub> se produjo mediante PCR (utilizando como molde T7 Gb1T18P) tres insertos diferentes: BamHI\_NotI\_Gb1\_G; G\_Gb1\_G; G\_Gb1\_XhoI\_EcoRI, donde G denota el residuo glicina, e incluyendo en el oligonucleótido 5' un *mismatch* que devolvía la mutación T18P al fenotipo silvestre. Se fosforiló el inserto G\_Gb1\_G, y se digirieron los insertos BamHI\_NotI\_Gb1\_G y G\_Gb1\_XhoI\_EcoRI con las enzimas BamHI y EcoRI respectivamente, para poder así clonar los tres insertos simultáneamente entre las dianas BamHI y EcoRI del plásmido T7blue (Novagen).

La construcción (I27)<sub>3</sub>-12XR-(I27)<sub>3</sub> se generó clonando mediante PCR las 12 repeticiones de resilina entre las dianas KpnI XbaI del plásmido pRSETA-(I27)<sub>8</sub> (Steward et al., 2002). El molde utilizado fue el plásmido T7 BamHI\_Ácido1\_(resilina)<sub>12</sub>\_Sall.

Utilizando el vector pETpoly\_RMBP se diseñó un vector que permitiera la producción de la proteasa TEV unida a la etiqueta polyR\_MBP. Para ello se eliminó la diana de la proteasa TEV dispuesta tras la polyR\_MBP del vector pETpolyR\_MBP, amplificando por PCR la etiqueta sin incluir dicha diana y subclonándola entre las dianas XbaI y NcoI del vector pET-24d (Novagen). Seguidamente se introdujo entre la dianas NcoI y XhoI un mutante de alta solubilidad de la TEV proteasa (Blommel y Fox, 2007) que fue generado mediante PCR solapante utilizando como molde el plásmido pRK793 (Kapust et al., 2001). La TEV proteasa que se produce en este vector, presenta la etiqueta polyR\_MBP en el extremo N-terminal y poliHistidinas en el C-terminal.

### 3.2.- Expresión y purificación de las proteínas recombinantes

Las proteínas se expresaron en las cepas de *E. coli* BLR (DE3), BL21 (DE3) (Novagen, Merck Millipore, Darmstadt, Alemania), BL21 Star (DE3) (Invitrogen) o C41 (DE3) (Miroux y Walker, 1996). Todos los cultivos se crecieron a 37°C hasta que

alcanzaron una  $OD_{595nm}$  de entre 0.5 y 1, y la expresión se indujo durante 4 h a 25 °C /37 °C mediante la adición de 1 mM IPTG.

En el caso de las proteínas utilizadas en los experimentos de la **sección 4.1** ( $\beta$ -catenina) las construcciones ARM/I27,  $\beta$ /I27 y  $\beta$ -cat fueron purificadas en primer lugar mediante cromatografía de afinidad de  $Ni^{2+}$  utilizando columnas HisTrap HP 5 ml (GE Healthcare, Uppsala, Suecia) en un aparato FPLC (ÄKTA Purifier, GE Healthcare). El tampón de unión usado fue 50 mM de fosfato de sodio (NaP)/500 mM NaCl/50 mM imidazol, pH7.4, y fueron eluidas en 50 mM NaP/500 mM NaCl/500 mM imidazol, pH7.4. Después de la purificación por afinidad se llevó a cabo un paso adicional de purificación por cromatografía de exclusión molecular en una columna Hiloal 16/60 (GE Healthcare) en 50 mM NaP/300 mM NaCl/1mM EDTA/1 mM DTT (en el caso de las construcciones  $\beta$ -cat y (I27)<sub>8</sub>) o 100 mM NaP/100mM NaCl/1mM EDTA/1 mM DTT (en el caso de ARM/I27 y  $\beta$ /I27). En el caso de  $\beta$ -cat se incluyó un paso de cromatografía de intercambio aniónico entre las purificaciones de afinidad y exclusión molecular. La cromatografía de intercambio aniónico se llevó a cabo utilizando columnas HiTrap Q (GE Healthcare) en 50 mM NaP/50 mM NaCl/1 mM EDTA/1 mM DTT, pH 7.4, incrementando la fuerza iónica linealmente hasta 1 M de NaCl. Las fracciones purificadas de cada proteína fueron concentradas utilizando concentradores de ultrafiltración Amicon 10K (Merck Millipore, Darmstadt, Alemania), y dializadas mediante tripas de diálisis Spectra/Por Dialysis Membrane 3 (Spectrum Laboratories, Rancho Dominguez, CA), en PBS/0.2 mM EDTA/ 1 mM DTT para dicroísmo circular (DC), o PBS/0.2 mM EDTA/ 5 mM DTT para AFM. La concentración de las proteínas, medidas a  $OD_{280}$  usando sus coeficientes de extinción teóricos, se mantuvo entre 0.3-0.4 mg/ml en el caso de  $\beta$ -cat, ARM/I27 y  $\beta$ /I27, y 1 mg/ml para (I27)<sub>8</sub>. La pureza de la preparación se estimó, en todos los casos, en más del 95 % mediante una electroforesis en gel de poliacrilamida (PAGE) teñido con azul de Coomassie (Coomassie Brilliant Blue R-250, Bio Rad).

Las proteínas GST\_Ncito y GST se purificaron mediante cromatografía de afinidad de GST utilizando la resina Glutathione Sepharose 4B (GE Healthcare) en tampón PBS/5mM EDTA/10 mM DTT. Las proteínas fueron eluidas en 50 mM TrisHCl/5 mM EDTA/10 mM DTT/10 mM glutatión reducido. GST\_Ncito fue sometida a un paso extra de purificación de afinidad de iones  $Ni^{2+}$  utilizando columnas HisTrap HP 5 ml (GE), en 8 M Urea, 50mM NaP/500 mM NaCl/50 mM imidazol, pH7.4, una vez unidas a la columna se eliminó gradualmente la Urea y se eluyó la proteína en 50mM NaP/500 mM NaCl/500 mM imidazol, pH7.4. Finalmente el tampón fue cambiado a PBS/0.2mM EDTA/ 5 mM DTT mediante tripas de diálisis Spectra/Por Dialysis Membrane 3 (Spectrum Laboratories).

En el caso de las proteínas utilizadas en los experimentos de la **sección 4.2**. Para la purificación de las proteínas recombinantes producidas a partir de plásmidos pETpolyR\_MBP y pETpolyR\_MBP\_tev se siguió el siguiente procedimiento general: en primer lugar las proteínas fueron purificadas mediante cromatografía de afinidad de Ni<sup>2+</sup> utilizando columnas HisTrap HP 5 ml (GE Healthcare) en un aparato FPLC (ÄKTA Purifier, GE Healthcare). El tampón de unión usado fue 50mM de fosfato de sodio/ 500 mM NaCl/ 50 mM imidazol, pH7.4. Previo a la elución de las proteínas se cambió el tampón de la columna a TrisHCl 20 mM/100 mM NaCl pH 7.4, y se dispuso tras la columna de afinidad de iones Ni<sup>2+</sup> una columna de intercambio catiónico HiScreen SP HP (GE Healthcare) equilibrada en el mismo tampón. La proteína se eluyó de la columna de afinidad mediante un tampón 20 mM TrisHCl/150 mM NaCl/300 mM imidazol pH 8, que tiene la concentración de imidazol necesaria para la elución de la columna de afinidad, pero una fuerza iónica no muy elevada permitiendo la retención de las proteínas cargadas positivamente por la columna de intercambio catiónico. Como resultado, las proteínas con la etiqueta polyR\_MBP eluyen de la columna de afinidad para ser retenidas por la columna HiScreen SP HP (GE Healthcare). Como paso final se retiró la columna de afinidad y se cambió el tampón de la columna de intercambio catiónico a 20 mM TrisHCl/100 mM NaCl pH 7.4, y se llevó a cabo un gradiente lineal de fuerza iónica creciente hasta un tampón final 20 mM TrisHCl/1M NaCl/1 mM EDTA/1 mM DTT pH 7.4, lo que provoca la elución de la proteína de la columna.

Las alícuotas con la proteína purificada se mezclaron en una unidad concentradora Amicon 10K (Merck Millipore), y fueron concentradas por ultrafiltración hasta un volumen de 2 ml. Las muestras se suplementaron con EDTA y DTT hasta una concentración final de 1 mM EDTA, 5 mM DTT y se incubaron toda la noche a 4 °C con proteasa TEV (a una relación 1  $\mu$ g de enzima:100  $\mu$ g de proteína).

En el caso de las proteínas producidas a partir del plásmido pETpoly\_RMBP, las muestras, una vez digeridas con la proteasa TEV, se purificaron con una columna MBPTrap HP 5ml (GE Healthcare), en un tampón 20 mM TrisHCl /100 mM NaCl pH 7.4, o 50 mM de fosfato de sodio/500 mM NaCl/50 mM imidazol, pH 7.4. Esto permite retirar de la preparación la polyR\_MBP cortada, la TEV proteasa, y las moléculas no procesadas por la TEV proteasa. En el caso de las proteínas producidas a partir de un plásmido pETpoly\_RMBP\_tev, para retirar de la preparación la etiqueta de poliHistidinas, se montó simultáneamente con la columna MBPTrap HP 5ml una columna HisTrap HP 5 ml en serie. Las proteínas fueron finalmente concentradas mediante ultrafiltración con un filtro Amicon 10K (Merck Millipore), y dializadas con una tripa de diálisis Spectra/Por 3 (Spectrum Laboratories) en 20 mM TrisHCl/100 mM NaCl/ 2 mM EDTA, pH 8.5.

Finalmente las proteínas fueron directamente guardadas a  $-80\text{ }^{\circ}\text{C}$ , o a  $-20\text{ }^{\circ}\text{C}$  en 50% glicerol/10 mM TrisHCl/50 mM NaCl/ 1mM EDTA/10 mM DTT, pH 8.5. Todas las proteínas presentaban una pureza superior al 95% cuando se analizaron por SDS-PAGE. En todos los casos, la concentración de proteína fue determinada midiendo absorbancia a 280 nm, utilizando en cada caso el coeficiente de extinción molar y peso molecular teóricos; la concentración final de proteína se mantuvo entre 0.5-2 mg/ml.

Para producir las construcciones Cohesina Básico (I27)<sub>3</sub> y Dockerina Básico (I27)<sub>3</sub>, las proteínas fueron purificadas en primer lugar mediante cromatografía de afinidad de Ni<sup>2+</sup> utilizando columnas HisTrap HP 5 ml (GE Healthcare) en un aparato FPLC (ÄKTA Purifier, GE Healthcare). El tampón de unión usado fue 20 mM TrisHCl /500 mM NaCl/50 mM imidazol, pH 7.4, y fueron eluidas en 20 mM TrisHCl/500 mM NaCl/500 mM imidazol, pH 7.4. Tras la purificación de afinidad se estimó visualmente en un gel SDS-PAGE teñido con azul de Coomassie (Coomassie Brilliant Blue R-250, Bio Rad) la concentración de ambas proteínas, se mezclaron ambas preparaciones en una proporción aproximada de 1:1, y se llevó a cabo un paso adicional de purificación por cromatografía de exclusión molecular en una columna Hiload 16/60 (GE Healthcare) en 20 mM TrisHCl /300 mM NaCl/1mM Cl<sub>2</sub>Ca/1 mM DTT. Las muestras correspondientes al complejo Cohesina/Dockerina Básico (I27)<sub>3</sub> fueron finalmente concentradas mediante ultrafiltración con un filtro Amicon 10K (Merck Millipore), y dializadas con una tripa de diálisis Spectra/Por 3 (SpectrumLabs) en 20 mM TrisHCl, 100 mM NaCl, 1mM Cl<sub>2</sub>Ca, pH 8.5. Finalmente fueron directamente guardadas a  $-80\text{ }^{\circ}\text{C}$ , o a  $-20\text{ }^{\circ}\text{C}$  en 50% glicerol, 10 mM TrisHCl, 50 mM NaCl, 0.5 mM Cl<sub>2</sub>Ca, 10 mM DTT, pH 8.5. En el caso de las construcciones Cohesina Básico (Gb1)<sub>3</sub> y Dockerina Básico (Gb1)<sub>3</sub>, las muestras fueron primero sometidas al esquema de purificación descrito para proteínas producidas a partir del vector pETpoly\_RMBP, y posteriormente se sometió al mismo tipo de purificación por exclusión molecular.

La proteína (I27)<sub>3</sub>-12XR-(I27)<sub>3</sub> fue purificada en primer lugar mediante cromatografía de afinidad de Ni<sup>2+</sup> utilizando columnas HisTrap HP 5 ml (GE Healthcare) en un aparato FPLC (ÄKTA Purifier, GE Healthcare). El tampón de unión usado fue 50 mM de fosfato de sodio (NaP)/500 mM NaCl/50 mM imidazol, pH7.4, y eluida en 50 mM NaP/500 mM NaCl/500 mM imidazol, pH7.4. Después de la purificación de afinidad se llevó a cabo un paso adicional de purificación por cromatografía de exclusión molecular en una columna Hiload 16/60 (GE Healthcare) en 50 mM NaP/300 mM NaCl/1mM EDTA/ 1 mM DTT. Las fracciones purificadas de la proteína fueron concentradas utilizando concentradores de ultrafiltración Amicon 10K (Merck Millipore), y dializada mediante tripas de diálisis Spectra/Por Dialysis Membrane 3 (Spectrum Laboratories) en PBS/ 1

mM EDTA/ 5 mM DTT. La proteína, a 1mg/ml (medida a OD<sub>280</sub> usando su coeficiente de extinción teórico), se guardó a -20 °C.

Las proteasas TEV, bien solamente con una cola de poliHistidinas en el N-terminal (Kapust et al., 2001), o la versión descrita más arriba (con la etiqueta polyR\_MBP en el extremo N-terminal y poliHistidinas en el C-terminal), se purificaron mediante cromatografía de afinidad de Ni<sup>2+</sup> utilizando columnas HisTrap HP 5 ml (GE Healthcare) en un aparato FPLC (ÄKTA Purifier, GE Healthcare). El tampón de unión usado fue 50 mM de fosfato de sodio (NaP)/500 mM NaCl/50 mM imidazol, pH7.4, y eluída en 50 mM NaP/500 mM NaCl/500 mM imidazol, pH7.4. Las fracciones purificadas de cada proteína fueron concentradas utilizando concentradores de ultrafiltración Amicon 10K (Merck Millipore), y dializadas mediante tripas de diálisis Spectra/Por Dialysis Membrane 3 (Spectrum Laboratories) en 50 mM TrisHCl, 100 mM NaCl, 2mM EDTA, pH 8. Cada proteína, a 1mg/ml (medida a OD<sub>280</sub> usando su coeficiente de extinción teórico), se guardó a -20 °C en 50 % glicerol/25 mM TrisHCl/ 50 mM NaCl/ 1mM EDTA/5 mM DTT.

### 3.3.- Protocolo Redox

Previo a su utilización, las proteínas guardadas en 50% glicerol eran dializadas con una tripa de diálisis Spectra/Por 1 (Spectra/Por Dialysis Membrane 1, SpectrumLabs) en 20 mM Tris/100 mM NaCl/2 mM EDTA pH 8.5, o 20 mM Tris/100 mM NaCl/1 mM Cl<sub>2</sub>Ca pH 8.5 durante 1 h. Cistamina/cisteamina (Sigma-Aldrich) (Mukhopadhyay, 2000; Bouvier et al., 2003) es la pareja de moléculas oxidante/reductor que controla el potencial redox del tampón redox. Para que el intercambio de tioles se produzca de manera eficiente el pH del tampón es básico, porque solo las formas deprotonadas de los tioles (tiolatos, -S<sup>-</sup>) son reactivas en una reacción de intercambio de tioles.

El procedimiento para encontrar las condiciones redox óptimas es el siguiente. Para conseguir una relación equimolar de cada proteína en la reacción de ensamblado, en primer lugar se estimaron los volúmenes de cada proteína necesarios para conseguir una banda de la misma intensidad en un gel SDS-PAGE teñido con azul de Coomassie (Coomassie Brilliant Blue R-250, Bio Rad). La intensidad de las bandas se estimó visualmente y se supuso como primera aproximación que intensidades iguales en proteínas diferentes implicaban igual número de moléculas. Una vez determinadas las relaciones volumétricas necesarias para conseguir una disolución equimolar de cada proteína, se ensayaban simultáneamente diferentes condiciones redox en pequeños volúmenes de entre 10-20  $\mu$ l. En un experimento típico las condiciones que se ensayaban eran:

- 1mM oxidante;

- 1mM oxidante/1mM reductor;
- 0.1mM oxidante;
- 0.1mM oxidante/0.1mM reductor;

En algunas ocasiones también se ensayó la relación 1mM oxidante/0.1mM reductor. Además de las cuatro condiciones redox, se incluían como control las proteínas aisladas sometidas a una de las condiciones redox (generalmente 1mM oxidante) para poder monitorizar el efecto del tratamiento sobre las proteínas (y poder identificar así los homodímeros, y otras configuraciones aberrantes no deseadas). Tras incubar las muestras dos horas para que se produjeran las reacciones de intercambio de puentes de disulfuro, se procedía a cargar las muestras en un gel SDS-PAGE en condiciones no reductoras (sin  $\beta$ -mecaptoetanol) para evaluar la condición redox más apropiada. La formación de puentes disulfuro entre las moléculas se identificaba por la aparición de bandas de mayor peso molecular, y la naturaleza de las mismas (si son homodímeros, configuraciones aberrantes o el complejo deseado) por comparación con las muestras control. Si la estimación de las concentraciones no había sido adecuada, se llevaba a cabo una segunda ronda de prueba en la que se elegía la condición redox que había resultado más óptima y se variaban las cantidades de las proteínas en función de lo observado en los experimentos preliminares hasta conseguir el mejor resultado.

Una vez determinada la mejor condición redox y las cantidades de proteínas necesarias para maximizar la producción del complejo deseado, se replicaba la reacción en un volumen mayor de unos 100  $\mu$ l totales, y esta muestra era analizada por AFM. Las proteínas eran dializadas y ensambladas en fresco de nuevo cada día para los experimentos de AFM-SMFS.

### 3.4.- Experimentos de Dicroísmo Circular

Los espectros de dicroísmo circular en el rango del UV lejano fueron adquiridos en un espectropolarímetro JASCO-J810 (JASCO Inc., Tokyo, Japan), como se describe en (Varea et al., 2004). Tras corregir por la contribución del tampón (PBS/0.2mM EDTA/1mM DTT, pH 7.4, en el caso de las proteínas de la **sección 4.1**, y 10 mM de fosfato de sodio (NaP)/0.2 mM EDTA pH 7, en el caso de Ácido1\_12XR\_Ácido2) los datos experimentales fueron convertidos a elipticidades molares ( $\theta$ ) usando como masas moleculares medias por residuo: 110.11 (I27)<sub>8</sub>, 109.65 ( $\beta$ -cat y ARM/I27), 109.72 ( $\beta$ /I27), y 93.06 (Ácido1\_12XR\_Ácido2) daltons.

El contenido de estructura secundaria de los constructos de la **sección 4.1** fue estimado por deconvolución de los espectros usando el programa DCNN (Bohm et al., 1992). La desnaturalización térmica fue seguida midiendo los cambios de elipticidad

inducidos por el aumento de temperatura (60°C/h) a longitudes de onda fijadas: 210 nm para ARM/I27,  $\beta$ /I27 y  $\beta$ -cat; a 224 nm para (I27)<sub>8</sub> (las longitudes de eligen en base a la mayor diferencia registrada entre un espectro de la proteína en estado plegado, y un espectro en estado desnaturalizado).

### 3.5.- Experimentos de <sup>1</sup>H RMN

Los experimentos de RMN se llevaron a cabo en colaboración con el Dr. Douglas Vinson Laurents (Instituto Rocasolano/CSIC), experto en RMN y estructura de proteínas.

Para comprobar el plegamiento de los I27 de las muestras Básico (I27)<sub>3</sub>, Ácido1 (I27)<sub>3</sub>, y de heterodímeros Ácido1-Básico (I27)<sub>3</sub> unidos mediante puentes disulfuro, se llevaron a cabo experimentos monodimensionales de RMN. Las muestras de Básico (I27)<sub>3</sub> + Ácido1 (I27)<sub>3</sub> se incubaron previamente 2 h en un tampón redox con 100  $\mu$ M de cistamina, para generar los heterodímeros Ácido1-Básico (I27)<sub>3</sub>. Muestras concentradas (0.5-1 mM) se dializaron contra 10 mM KH<sub>2</sub>PO<sub>4</sub> pH 4.5 con 10% D<sub>2</sub>O. Las medidas se tomaron a 25 °C usando un espectrómetro Bruker AV 800 (Bruker BioSpin) equipado con una criosonda con un gradiente-z. La temperatura fue calibrada usando una muestra de etanol y la señal del motivo trimetil de sodio 2,2-dimetil-2-silapentano-5-sulfonato [(CH<sub>3</sub>)<sub>3</sub>-Si-CH<sub>2</sub>-CH<sub>2</sub>-CH<sub>2</sub>-SO<sub>3</sub>- Na<sup>+</sup>] (DSS) se usó como referencia interna de los desplazamientos químicos (Markley et al., 1998). Se usó un módulo WATERGATE (Piotto et al., 1992) para reducir la señal del agua presente en la muestra. Los espectros fueron analizados utilizando TopSpin 2.0 (Bruker BioSpin).

### 3.6.- Ensayos de co-precipitación.

Para el ensayo de co-precipitación entre ARM/I27 y GST\_Ncito todas las proteínas, ARM/I27, GST\_Ncito, GST y (I27)<sub>8</sub>, se utilizaron a una concentración inicial de 5  $\mu$ M, en PBS/0.2 mM EDTA/5mM DTT. Para en el ensayo se mezclaron volúmenes iguales de proteínas o tampón (30  $\mu$ l), por lo que la concentración final se mantuvo a 2.5  $\mu$ M. Tras 40 minutos de incubación, se añadieron 10  $\mu$ l de resina Glutathione Sepharose 4B (GE Healthcare) pre-equilibrada según las instrucciones del fabricante. Después de 30 minutos de incubación con la resina en un agitador orbital, se llevaron a cabo tres lavados de la misma con el tampón. Finalmente la resina se hirvió en tampón de electroforesis Laemmli, y tras centrifugar la muestra, una fracción de los sobrenadantes se cargó en un gel SDS-PAGE.



### 3.7.- Ensayos de protección frente a digestión por proteasas.

Para los ensayos de protección ante la digestión por tripsina se prepararon 10 muestras. El volumen total de la reacción fue 14  $\mu$ l en el tampón 20 mM Tris/100 mM NaCl/1mM EDTA pH 8.5. La cantidad de Ácido1 (I27)<sub>3</sub> por reacción fue de 12  $\mu$ g, la de Básico (Gb1)<sub>3</sub> de 8  $\mu$ g, la nomenclatura utilizada fue la siguiente: tripsina = incubación con 50 ng de tripsina, 2 x tripsina = incubación con 100 ng de tripsina, redox = la muestra se incubó previamente 2 h en el tampón descrito suplementado con 1 mM Cistamina. Las muestras se incubaron con la proteasa durante 15 minutos, tras lo cual se cargaron en un gel SDS-PAGE del 15 %, en este orden de izquierda a derecha en la **figura 4.16.b**:

- 1.- Básico (Gb1)<sub>3</sub>
- 2.- Ácido1 (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub>
- 3.- Básico (Gb1)<sub>3</sub>; Redox; Tripsina
- 4.- Básico (Gb1)<sub>3</sub>; Redox; 2x Tripsina
- 5.- Ácido1 (I27)<sub>3</sub>; Redox; Tripsina
- 6.- Ácido1 (I27)<sub>3</sub>; Redox; 2x Tripsina
- 7.- Ácido1 (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub>; Redox; Tripsina
- 8.- Ácido1 (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub>; Redox; 2x Tripsina
- 9.- Ácido1 (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub>; Tripsina
- 10.- Ácido1 (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub>; 2x Tripsina

### 3.8.- Simulaciones de dinámica molecular.

Las simulaciones de dinámica molecular de la región **ARM** de la  $\beta$ -catenina, y la región ARM portando una hélice  $\alpha$  extra en el extremo C-terminal (**ARMHC**) fueron llevadas a cabo por otro miembro del laboratorio (Dr. Alejandro Valbuena, Centro de Biología Molecular/CSIC) mientras que yo participé en el análisis. Los detalles concretos del procedimiento de las simulaciones pueden encontrarse en (Valbuena et al., 2009; Valbuena et al., 2012). En un breve resumen, las simulaciones fueron llevadas a cabo como sigue:

Las simulaciones de dinámica molecular de las estructuras **ARM** y **ARMHC** (código PDB 1qz7 y 2z6h respectivamente (Xing et al., 2003; Xing et al., 2008)) fueron llevadas a cabo utilizando un modelo de solvente implícito (GB/SA, *generalized Born/surface Area* (Tsui y Case, 2000)) que predice las propiedades mecánicas tan precisamente como los modelos explícitos (Valbuena et al., 2009).

Las coordenadas atómicas iniciales fueron tomadas de la estructura cristalográfica de la región ARM de la  $\beta$ -catenina humana (1qz7) porque contenía el esqueleto peptídico

completo (incluyendo el de la repetición R10). Para completar el lazo R10 de 2z6h (que está incompleto en la estructura depositada), se llevó a cabo un alineamiento estructural con 1qz7 y se tomó el esqueleto del lazo R10 de esta última estructura como molde. Las simulaciones de dinámica molecular libre, así como los estiramientos, se realizaron mediante el programa SANDER del paquete AMBER10 (Case et al., 2008a) y se utilizó el campo de fuerza parm-99. El solvente acuoso, con una concentración de 0.1 M de iones monovalentes, fue simulado de forma implícita, realizando correcciones a la energía debido a la exposición al solvente, mediante el generalizado de Born con correcciones de superficie (GB/SA (Tsui y Case, 2000)). La superficie expuesta al solvente se calculó mediante el método LCPO (Hawkins et al., 1996). El potencial electrostático se truncó a 10 Å disminuyendo hasta cero en 12 Å. Primero se llevó a cabo una minimización de las estructuras de 5 ps con los  $C_{\alpha}$  restringidos (20 kcal/mol·Å<sup>2</sup>). Después, con un paso de tiempo de 1 fs, se aumenta la temperatura desde 0 a 298 K a lo largo de 20 ps con los  $C_{\alpha}$  aún restringidos. Finalmente, antes de llevar a cabo la dinámica molecular libre, se eliminan gradualmente las restricciones y se estabiliza el sistema durante 8 ps. Las dinámicas moleculares libres se llevaron a cabo durante 4 ns, con un intervalo de tiempo de 2 fs. Se seleccionaron las coordenadas de las proteínas a 1, 2 y 3 ns de la dinámica libre como puntos de comienzo para llevar a cabo las simulaciones de los estiramientos. Para simular los estiramientos se impuso una restricción ( $X_{restr}$ ) en la distancia de los  $C_{\alpha}$  N- y C-terminales ( $k_{restr} = 5$  kcal/mol·Å<sup>2</sup>). Dicha distancia se aumentó a razón de 1 Å/ps. La visualización de las trayectorias, extracción de las desviaciones cuadráticas medias (RMSD) de los  $C_{\alpha}$ , y la distancia entre los  $C_{\alpha}$  terminales fueron obtenidos con el software VMD (Humphrey et al., 1996). Estos datos fueron cargados en Igor Pro 6 (Wavemetrics) para calcular las fuerzas como:  $F = -k_{restr} \times (X - X_{restr})$ .

### **3.9.- Funcionalización de sustratos de AFM-SMFS**

#### **3.9.1.- Cubreobjetos funcionalizados con oro.**

Los cubreobjetos de vidrio fueron limpiados por sonicación en pasos de 15 min, primero con EtOH, después con 0.1% Hellmanex II (Hellma GmbH & Co. KG, Mülheim, Germany) y finalmente con agua MilliQ. Entre cada paso de sonicación, los cubreobjetos fueron aclarados intensamente con agua MilliQ. Finalmente fueron depositados en una cámara de vacío casera (a  $2 \cdot 10^{-6}$  mbar), y por evaporación térmica se depositó sobre ellos una capa adhesiva de cromo y una capa final de oro.

#### **3.9.2.- Cubreobjetos funcionalizados con NTA-Ni<sup>2+</sup>.**

Para su preparación, los cubreobjetos se sumergieron o/n en primer lugar en una solución 20 N de KOH y luego se colocaron bajo un flujo de agua MilliQ durante 1 h antes de que finalmente fuesen incubados en una solución al 2% de 3-mercapto-propil-trimetoxisilano (MPTS; Sigma-Aldrich)/0,02% CH<sub>3</sub>COOH a 90°C durante 1 h. Posteriormente, se lavaron con agua MilliQ durante 1 h y seguidamente se incubaron 15 min en un horno a 120°C. Después de dejarlos enfriar a temperatura ambiente (RT) durante aproximadamente 10 min, los cubreobjetos se transfirieron a una solución 100 mM DTT durante 15 min y después se lavaron con agua MilliQ durante 1 h. A continuación se depositó sobre cada cubreobjetos (durante 30 min) una gota de aproximadamente 50  $\mu$ l de una solución de 20 mg/ml de maleimida-C3-NTA (Laboratorios Dojindo) disuelta en 10 mM de ácido 3-(N-morfolino) propanosulfónico (MOPS) [pH 7,0]. Después de lavar con agua MilliQ se añadió a cada cubreobjetos una gota (60  $\mu$ l) de 10 mM NiCl<sub>2</sub> y se dejaron incubar durante 10 min. Por último, los cubreobjetos se lavaron brevemente en agua MilliQ antes de su almacenamiento para su uso a temperatura ambiente (RT).

### 3.9.3.- Cubreobjetos funcionalizados con NTA-Ni<sup>2+</sup>. Protocolo modificado

Este otro protocolo de funcionalización es una modificación del descrito en (Hervas et al., 2012), esquematizado arriba, afectando las modificaciones más importantes al procedimiento de lavado de los cubreobjetos y de deposición del silano. El proceso se llevó a cabo en placas Petri de vidrio, y en cada paso de lavado múltiple se incluyó al menos un paso en el que los vidrios fueron traspasados a una nueva placa Petri que contenía la solución de lavado (agua, etanol o tolueno), para asegurar que se eliminaban los restos de solución que quedan entre el cubreobjetos y la placa Petri. Entre el paso de limpieza de vidrios y deposición de silano se incluyeron dos pasos de lavado con etanol y tolueno para eliminar los remanentes de solución acuosa del paso de limpieza. Para su limpieza los cubreobjetos se sumergieron en una solución 1:1 de metanol (Fisher Chemical, HPLC *gradient grade*) y ácido clorhídrico (Merck, 37% *for analysis*) durante 30 min (Cras et al., 1999), tras lo cual fueron lavados con agua MilliQ cuatro veces y dos veces con alcohol absoluto (Merck, *absolute for analysis*); finalmente los vidrios se dejaron secar. Los vidrios fueron traspasados uno a uno a otra placa Petri con tolueno (Panreac, 99.5%, 0.03% agua), y tras retirar el tolueno se procedió al paso de deposición del silano, incubando los vidrios en una solución de 5% de 3-mercapto-propil-trimetoxisilano (MPTS; Sigma-Aldrich, 95%) en tolueno durante 4 h en un agitador a 30 rpm (revoluciones por minuto) (Halliwell y Cass, 2001). Posteriormente los vidrios fueron lavados dos veces en tolueno, y dos veces en etanol absoluto para luego ser curados durante toda la noche en un horno a 120°C. Después de dejarlos enfriar a RT durante aproximadamente 10 min, los

cubreobjetos se transfirieron a una solución 100 mM DTT/50 mM TrisHCl pH 8.4 durante 15 min y después se lavaron cuatro veces con agua MilliQ. A continuación se depositó sobre la mitad de los cubreobjetos una gota de 19  $\mu$ l de una solución de 20 mg/ml de maleimida-C3-NTA (Laboratorios Dojindo) disuelta en 10 mM de ácido 3-(N-morfolino)propanosulfónico (MOPS) pH 7.0, y se utilizó la otra mitad de los cubreobjetos para hacer un sándwich sobre la gota. Tras incubar 30 minutos se procedió a hacer un lavado con agua MilliQ y se añadió nuevamente en sándwich una gota (30  $\mu$ l) de 50 mM NiSO<sub>4</sub> y se dejaron incubar durante 10 minutos. Por último, los cubreobjetos se lavaron brevemente en agua MilliQ antes de su almacenamiento a RT para su uso.

### 3.9.4.- Membranas de celulosa funcionalizadas con IDA.

El esquema general de funcionalización de celulosa viene descrito en (Hu et al., 2006). El protocolo se puso a punto en primer lugar para celulosa particulada Sigmacell Celulose, Type 101 (Sigma-Aldrich), y posteriormente se aplicó a membranas de celulosa regenerada Spectra/Por Dialysis Membrane 3 (Spectrum Laboratories).

La celulosa se incubó en una solución de 5 ml de epiclorohidrina (Fluka, purum  $\geq$  99 %, Sigma-Aldrich) en fase acuosa (20 ml de 1 M NaOH) agitadas a 100 rpm durante 2 h a 60 °C. Tras la incubación, las membranas fueron lavadas con agua MilliQ. Para el acoplamiento del ácido iminodiacético (IDA), la celulosa conjugada a epiclorohidrina se hizo reaccionar con 50 ml de 0.2 M IDA y 1 M Na<sub>2</sub>CO<sub>3</sub>, pH 11, a 80 °C toda la noche. Después de la reacción, la celulosa se limpió con 5 % ácido acético y agua destilada. Las muestras se incubaron durante 1h en soluciones de 50 mM NiSO<sub>4</sub>/CoCl<sub>2</sub>/CuCl<sub>2</sub> en función del metal con el que se deseaba cargar la celulosa: Ni<sup>2+</sup>, Co<sup>2+</sup> o Cu<sup>2+</sup>, respetivamente. Finalmente las membranas se lavaron en agua MilliQ, se impregnaron en una solución de 10 % glicerol y se dejaron secar.

### 3.10.- Experimentos de AFM-SMFS

Nuestro AFM, usado para llevar a cabo las medidas de SMFS pero que tiene capacidad de obtener imágenes, y su modo de funcionamiento han sido descritos con anterioridad (Valbuena et al., 2007). Antes de cada experimento, la micropalanca se limpió durante 15-20 min con una lámpara UV/ozono (UV/Ozone ProCleaner™ Plus, *Bioforce Nanoscience Inc.*). La constante de muelle de cada micropalanca individual se calculó usando el teorema de equipartición. Se utilizaron dos tipos de micropalanca: MLCT-AUHW (Veeco Probes, Camarillo, CA) con una constante de muelle nominal de  $\sim$  40 pN/nm; o BL-RC (Olympus, Tokyo, Japón) con una constante de muelle nominal de  $\sim$  30 pN/nm. Sobre los sustratos se depositó una gota de entre 5-40  $\mu$ l de proteína sobre un

volumen final de 40  $\mu\text{l}$  de tampón, y se incubó durante 15 min., posteriormente la superficie se limpió 2-3 veces con el tampón experimental. Al acercarse el posicionador piezoeléctrico a la muestra en el eje Z, las proteínas, depositadas sobre el sustrato, resultan ancladas por adsorción a la punta de la micropalanca. El actuador piezoeléctrico era retraído varios cientos de nm para estirar las proteínas previamente adsorbidas. Todos los experimentos de SMFS se realizaron a una velocidad de estiramiento constante de 400 nm/s en el denominado modo de longitud controlada y los datos obtenidos se analizaron utilizando el software Igor Pro 6 (*Wavemetrics*). La elasticidad de las moléculas estiradas se analizó utilizando el modelo de elasticidad de polímeros WLC (Bustamante et al., 1994; Marko y Siggia, 1995):

$$F = \frac{k_B T}{p} \left\{ \frac{1}{4 \left(1 - \frac{x}{L_c}\right)^2} - \frac{1}{4} + x/L_c \right\},$$

donde  $F$  es la fuerza,  $p$  es la longitud de persistencia,  $x$  corresponde a la longitud de extremo a extremo y  $L_c$  es la longitud de contorno de la proteína estirada.  $L_c$  y  $p$  son los dos parámetros ajustables de la ecuación.

En la **sección 4.1** los resultados estadísticos están expresados como media  $\pm$  SEM (error estándar de la media). En la **sección 4.2** los resultados estadísticos están expresados como media  $\pm$  SD (desviación estándar)

El tampón experimental utilizado fue PBS, a excepción de los experimentos de construcciones Cohesina/Dockerina, en los que se utilizó 20 mM Tris/100 mM NaCl/ 0.5-1 mM  $\text{Cl}_2\text{Ca}$  pH 8.

En el caso de los experimentos de la **sección 4.1** se utilizaron puntas MLCT-AUHW (Veeco Probes) o BL-RC (Olympus, Tokyo, Japón), y sustratos de oro o de níquel funcionalizados con el procedimiento original.(Hervas et al., 2012)

En el caso de los experimentos de la **sección 4.2**, se utilizaron sustratos funcionalizados con cationes divalentes (principalmente vidrios funcionalizados con  $\text{Ni}^{2+}$ ) y puntas BL-RC (Olympus) recubiertas de oro. La idea de esta configuración es utilizar las poliHistidinas de las proteínas para la unión al sustrato, y cisteínas libres terminales para la unión a las puntas recubiertas de oro. Debido a que muchos de los experimentos de la **sección 4.2** incluyen 2-3 proteínas diferentes dentro de la construcción analizada, la existencia de etiquetas de poliHistidinas terminales en los dos extremos sobre los que se aplica la fuerza, podría llevar al bloqueo de ambos terminales en la superficies (y por lo tanto la imposibilidad de estirar la construcción). Cuando se incluyó en las muestras la construcción el polipéptido elastomérico, la cola de poliHistidinas de esta construcción se retiró siempre de la proteína. Para conocer la posible interferencia de mantener las colas de poliHistidinas en ambos extremos, se probaron las siguientes combinaciones en

diferentes constructos, no encontrándose diferencias entre las diferentes combinaciones, por lo que el bloqueo de las construcciones en la superficie no parece probable:

- manteniendo las dos colas de poliHistidinas y cisteínas terminales en una de ellas: en el caso de los constructos Ácido1-Básico (I27)<sub>3</sub>, Ácido1 (I27)<sub>3</sub> - Básico (Gb1)<sub>3</sub>, construcciones Cohesina/Dockerina Básico (I27)<sub>3</sub>

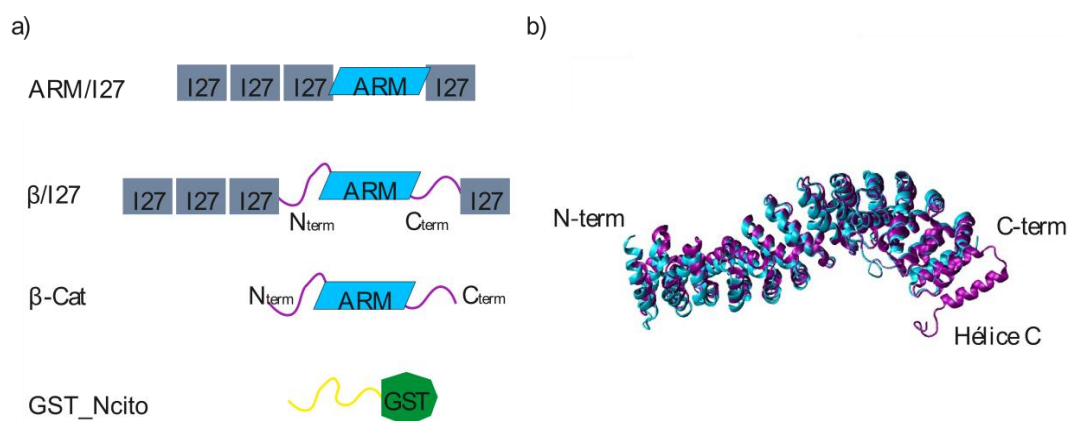
- eliminando una de las dos colas de poliHistidinas manteniendo cisteínas terminales en la construcción con la cola eliminada: Ácido1\_12XR\_Ácido2 + Básico (I27)<sub>3</sub>.

- manteniendo las dos colas de poliHistidinas sin incluir cisteínas terminales: Ácido1\_12XR\_Ácido2 + Básico (I27)<sub>3</sub>, construcciones Cohesina/Dockerina Básico (Gb1)<sub>3</sub>

## 4. Resultados

### 4.1.- Nanomecánica de la $\beta$ -Catenina

Para llevar a cabo la caracterización nanomecánica de la  $\beta$ -catenina adoptamos una estrategia en dos fases. En la primera fase estudiamos las propiedades mecánicas de la región estructurada armadillo (ARM), que representa la región central de la molécula y la zona de unión de un gran número de los ligandos de la  $\beta$ -catenina (incluido el dominio citosólico de la cadherina (Xing et al., 2008)). Para ello generamos una construcción recombinante a la que llamamos ARM/I27, en la que la región ARM (residuos 134-671) está flanqueada por 4 módulos de I27 como controles de monomolecularidad (**figura 4.1.a**). Una vez caracterizado el comportamiento nanomecánico de la región ARM, estudiamos la mecánica de la  $\beta$ -catenina completa para conocer los efectos y la aportación de las colas N- y C-terminales al desplegamiento mecánico de toda la molécula. Este análisis fraccionado está justificado por el hecho de que al estar desestructuradas, no se espera resistencia mecánica directa de las colas terminales (Li et al., 2001; Li et al., 2002; Lv et al., 2010). Para conocer el efecto de las regiones terminales en la región ARM generamos dos construcciones adicionales con la  $\beta$ -catenina completa, tanto con controles de monomolecularidad (constructo  $\beta$ /I27, **figura 4.1.a**) como sin ellos ( $\beta$ -cat, **figura 4.1.a**).



**Figura 4.1.- Constructos recombinantes y estructuras cristalográficas utilizadas para el análisis nanomecánico de la  $\beta$ -catenina.** a) Se muestran esquemáticamente las construcciones recombinantes utilizadas para los estudios por AFM-SMFS: ARM/I27,  $\beta$ /I27 y  $\beta$ -cat. Se representa también la proteína GST\_Ncito utilizada en los controles de unión de ligando natural (ver **sección 4.1.4**). b) Estructuras utilizadas para las simulaciones de estiramiento: ARM en azul (PDB:1qz7) y la región ARM + hélice C de la región C-terminal en morado (PDB:2z6h)

Finalmente, para tener una información más detallada del proceso de desplegamiento mecánico se realizaron simulaciones de dinámica molecular (DM). Para

disponer de una comparación lo más similar posible a los experimentos se utilizaron dos estructuras cristalográficas diferentes (**figura 4.1.b**). La estructura cristalográfica de la región ARM de la  $\beta$ -catenina humana (**ARM** en azul en la **figura 4.1.b**, PDB:1qz7 (Xing et al., 2003)) se utilizó para como equivalente del constructo ARM/I27, y para estudiar los efectos de las regiones terminales en la región armadillo se utilizó una estructura cristalizada con una porción de la región C-terminal (**ARM/HC** en morado en la **figura 4.1.b**, PDB:2z6h (Xing et al., 2008)). Hay que tener en cuenta que existe diferencia entre los residuos de la  $\beta$ -catenina utilizados para los estudios por AFM-SMFS (**ARM/I27** residuos 134-671,  $\beta$ -cat y  $\beta$ /I27 residuos 1-781), y los utilizados para la cristalización (**ARM** residuos 131 al 666, **ARM/HC** 138-781). También hay diferencias entre el número de residuos que se utilizaron para la cristalización y los que finalmente se resuelven en la estructura en el caso de **ARM** (residuos 149-166). Esta diferencia es mayor en el caso de la estructura **ARM/HC** (residuos 150-691), en la que de hecho sólo se resuelve una  $\alpha$ -hélice del extremo C-terminal, que se apila sobre la repetición 12 de la región armadillo (que llamamos hélice C). Aunque existan diferencias entre las estructuras analizadas por AFM-SMFS y por DM, pensamos que sería de gran interés realizar las simulaciones porque nos permiten obtener información de tipo cualitativo sobre el desplegamiento mecánico y sobre el efecto de las regiones terminales.

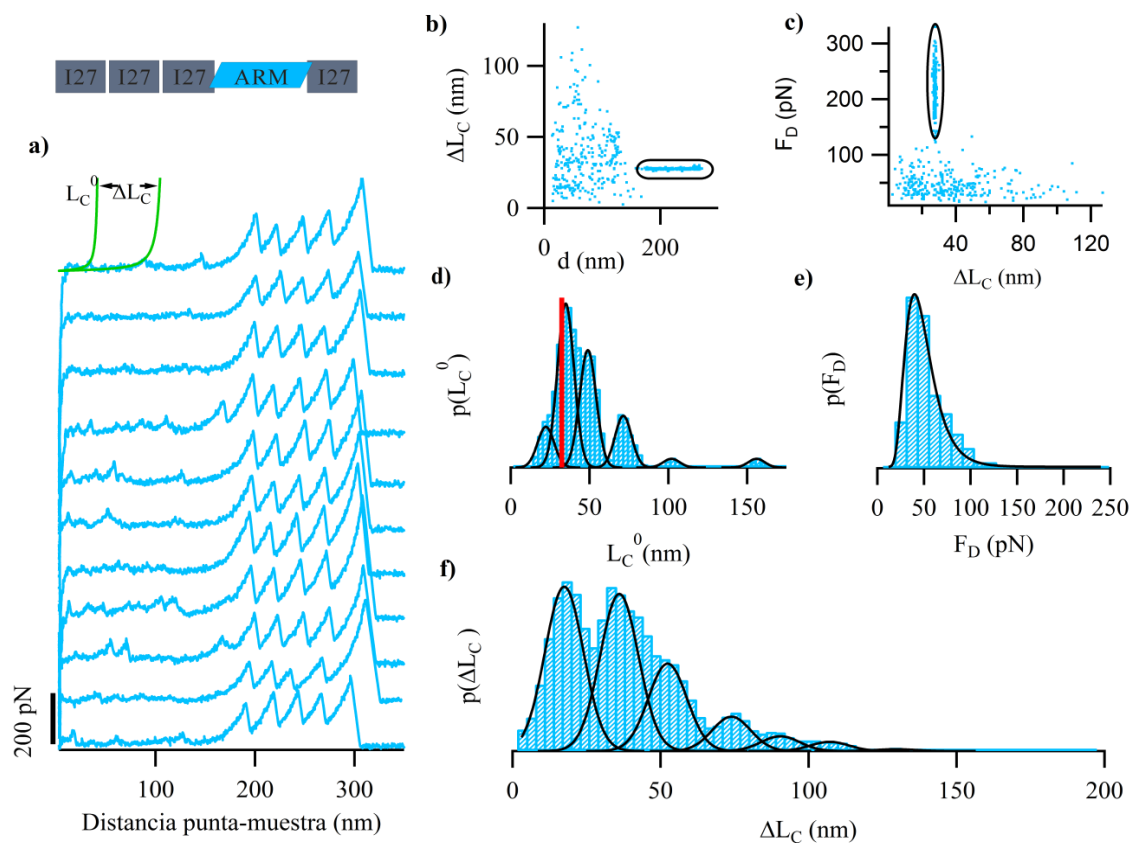
#### 4.1.1.- Mecánica de la región ARM de la $\beta$ -catenina

Para estudiar la nanomecánica de la región armadillo analizamos la construcción ARM/I27 (**figura 4.1.a**). Por la configuración del constructo, es necesario observar el desplegamiento de cuatro módulos I27 para estar totalmente seguros de que se ha estirado la región ARM, por esto sólo los registros con cuatro módulos marcadores de I27 fueron seleccionados en los análisis (**figura 4.1.a** y **figura 4.2.a**).

Cómo se observa en la **figura 4.2.a**, la región ARM de la  $\beta$ -catenina se desplegaba siempre a fuerzas menores que los módulos marcadores, por lo que aparecía siempre en la región proximal del registro, es por esto que se tomó un segundo criterio para la selección de registros, y es que estos deberían estar “limpios” en la zona proximal (particularmente susceptible a ruido, como se ha comentado). Debido a que la región ARM está formada por 12 repeticiones armadillo, era posible que al igual que se ha observado para proteínas con repeticiones estructuralmente relacionadas (ANK, (Li et al., 2006)) obtuviéramos un patrón repetido en dientes de sierra, en el que cada pico de fuerza correspondiese al desplegamiento de una repetición. Sin embargo los registros de fuerza-extensión (**figura 4.2.a**) muestran un número de picos de fuerza variable (de 2-9 picos por curva,  $n^{\circ}$  picos/curva  $\sim 4.5$ ), con una posición y un incremento de longitud de contorno ( $\Delta L_c$ )



variable. En contraposición a esta variabilidad, la distribución de la magnitud de las fuerzas de desplegamiento es mucho más homogénea (**figura 4.2.e** y **figura 4.2.f**). La distribución de fuerzas se ajusta a una distribución log-normal ( $F_0 = 33.7$ ,  $\chi^2 = 1.5 \times 10^{-5}$ ), con una fuerza de desplegamiento media de  $44 \pm 1$  pN ( $n = 278$ ) y un solo evento, separado de la población anterior por un mínimo local en la distribución, con una fuerza de desplegamiento de  $113 \pm 9$  pN (aquí el error es el error de la medida).



**Figura 4.2.- Análisis nanomecánico de la región ARM de la  $\beta$ -catenina.** **a)** Detalle del desplegamiento de varias moléculas ARM/I27, en la región proximal se observa el desplegamiento de la región armadillo, seguido de los 4 módulos marcadores I27. El ajuste por medio del modelo de WLC al primer pico nos permite estimar  $L_c^0$ , y el valor entre picos estimar  $\Delta L_c$ . El pictograma representa el constructo analizado. **b)** y **c)** representaciones de  $\Delta L_c$  vs  $d$  ( $d$ =distancia punta-muestra a la que aparece el pico), y  $F_D$  vs  $\Delta L_c$ . Los eventos provenientes de los marcadores I27 se muestran agrupados (elipses). **d)** Histograma de  $L_c^0$ . La distribución ajusta a varias gaussianas centradas en 21, 33, 47, 70, 100 y 154 nm ( $\sigma = 8.2$ ,  $\chi^2 = 3.7 \times 10^{-5}$ ). La línea roja marca la longitud de la proteína plegada (31 nm =  $\sim 5$  nm  $\times$  4 (I27) + 11 nm ARM). La existencia de eventos muy alejados a la derecha de la línea roja indica que existen repeticiones que se encuentran desplegadas previamente o se despliegan con fuerzas por debajo del límite de detección. **e)** Histograma de fuerzas de la región ARM. El valor medio de la población es  $46 \pm 1$  pN, en línea negra se representa el ajuste a una distribución log-normal ( $F_0 = 33.7$ ,  $\chi^2 = 1.5 \times 10^{-5}$ ). **f)** Histograma de  $\Delta L_c$ , la distribución puede ajustarse a varias gaussianas cuyos máximos 16, 35, 51, 72, 89, 106 nm ( $\sigma = 9.2$  nm,  $\chi^2 = 1.5 \times 10^{-5}$ ) corresponden aproximadamente a la longitud de una o varias repeticiones armadillo.

La gran variabilidad que presentan los registros sugiere que la región ARM puede tomar rutas de desplegamiento variables. Para corroborar esta idea representamos los  $\Delta L_c$

vs distancia (a la que aparecen los picos) (**figura 4.2.b**), y la magnitud del pico de fuerza  $F_D$  vs  $\Delta L_c$  (**figura 4.2.c**) ya que cualquier agrupamiento en estos gráficos sería indicativo de la existencia de un camino de desplegamiento preferido. En ambos gráficos se observa perfectamente el agrupamiento proveniente de los marcadores I27 (círculos, con  $\Delta L_c$  de  $\sim 28$  nm, y  $\langle F_D \rangle \sim 200$  pN), lo que contrasta fuertemente con los eventos provenientes de la  $\beta$ -catenina, que aparecen dispersos no apreciándose correlación en las representaciones.

El análisis del histograma de  $\Delta L_c$  (**figura 4.2.f**) muestra una distribución que puede ser ajustada a siete curvas gaussianas con máximos en: 16, 35, 51, 72, 89, 106 nm ( $\sigma = 9.2$  nm,  $\chi^2 = 1.5 \times 10^{-5}$ ). Estas distancias se corresponden a la liberación de unos 40, 87, 127, 181, 223, 265 y 325 residuos (aproximando 0.4 nm/aa (Ainavarapu et al., 2007)); teniendo en cuenta que la secuencia consenso de las repeticiones ARM es de 42 residuos (Huber et al., 1997) los ajustes parecen reflejar el desplegamiento de 1, 2, 3, 4, 5, 6 y 8 repeticiones ARM. Siendo la fuerza de desplegamiento media  $F_D$  ( $44 \pm 1$  pN, **figura 4.2.e**) mayor que el ruido ( $\sim 9$  pN) podemos asumir que este comportamiento es intrínseco a la molécula. Esta variabilidad sugiere que varias repeticiones se despliegan a una fuerza por debajo de lo que somos capaces de detectar con el AFM, o bien algunas repeticiones se despliegan casi simultáneamente a una fuerza similar, lo que implicaría la existencia cooperatividad en el desplegamiento mecánico de las repeticiones armadillo.

Además, como algunos registros fuerza-extensión mostraban una región inicial bastante larga sin picos de fuerza, estimamos mediante el modelo de WLC (ver **sección 3.10**) la distancia a la que aparecía el primer pico de fuerza ( $L_c^0$ ), para así comprobar si existían regiones de la molécula que estaban previamente desnaturalizadas antes de comenzar el estiramiento (**figura 4.2.d**). Este análisis nos permitió extraer información adicional importante: en primer lugar, aparecen picos de fuerza a una extensión menor que la longitud de la molécula plegada ( $11$  nm ARM +  $5$  nm  $\times 4$  I27 =  $31$  nm, línea roja en la **figura 4.2.d**) que pueden ser interpretados como eventos de despegado (bien de la molécula de la superficie, bien de posibles interacciones espurias de los marcadores con la región ARM), o también pueden ser originados por la ruptura de las fuerzas que mantienen la estructura terciaria (la superhélice ARM). Tras este máximo inicial, aparece otro entorno a los  $33$  nm (muy cerca del valor de  $31$  nm calculados para la molécula plegada), que indica que gran parte de las moléculas empiezan a mostrar resistencia a la fuerza cuando se está produciendo el alineamiento perpendicular a la superficie de la molécula. Finalmente aparecen eventos con una  $L_c^0$  mayor que la de la molécula plegada, con máximos a  $47$ ,  $70$ ,  $100$  y  $154$  nm que tras restarles  $31$  nm (del alineamiento de la molécula) coinciden con 1, 2, 4 y 7 repeticiones ARM. Esto supone que en el momento de comenzar el estiramiento algunas moléculas presentan repeticiones que están

desnaturalizadas o que pueden desplegarse debajo de nuestro límite de detección, lo que es compatible con que se encuentren en un equilibrio entre conformaciones estructuradas y desestructuradas, similar a lo que se ha observado en las repeticiones anquirina 5-6 de  $\text{I}\kappa\text{B}\alpha$  (Lamboy et al., 2011). En cualquier caso los datos estarían en concordancia con la existencia de una plasticidad estructural previa en la  $\beta$ -catenina como ya se ha descrito anteriormente (Ritco-Vonsovici et al., 2007).

#### 4.1.2.- Mecánica de la $\beta$ -catenina completa.

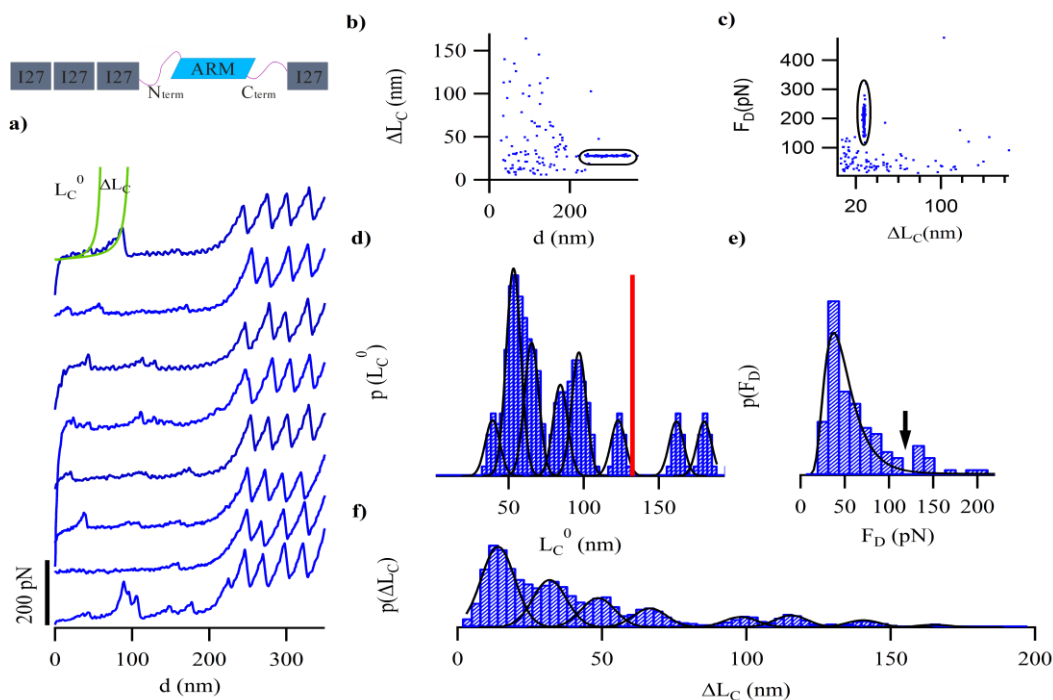
La estrategia de dividir la caracterización nanomecánica de la  $\beta$ -catenina en dos fases nos permite una interpretación de los datos más cómoda y segura. Una vez caracterizada la región ARM, conocer la aportación de las colas terminales desestructuradas a la mecánica de toda la proteína es mucho más sencilla. Para estudiar por SMFS la  $\beta$ -catenina completa utilizamos la construcción recombinante  $\beta/127$  (**figura 4.1.a**), en la que la proteína está rodeada por 4 repeticiones de I27 como marcador de monomolecularidad. Igualmente, el criterio seguido para seleccionar los registros fue la existencia de 4 módulos marcadores de I27. En la **figura 4.3.a** se observa que al igual que ocurría con la región ARM aislada, la  $\beta$ -catenina completa se despliega siempre a una fuerza menor que los módulos I27, siendo los registros muy variables en el número y posición de picos de fuerza, y en el  $\Delta L_c$  de los mismos.

Parece pues, que la  $\beta$ -catenina completa también se podría desplegar cada vez por caminos de desplegamiento diferentes. De nuevo el análisis de  $\Delta L_c$  vs distancia, y  $F_D$  vs  $\Delta L_c$  (**figura 4.3.b** y **figura 4.3.c**) muestra una dispersión total de los eventos provenientes del desplegamiento de la  $\beta$ -catenina, a diferencia del claro agrupamiento de los marcadores I27. De nuevo, dicha dispersión estaría en línea con la existencia de múltiples caminos de desplegamiento en contraposición a un camino bien definido.

Al igual que en el caso de la región ARM aislada, los histogramas de  $\Delta L_c$  muestran una distribución multimodal (**figura 4.3.f**). Esta distribución de nuevo nos permite interpretarla como que el desplegamiento de la molécula ocurre a través de la liberación longitudes correspondientes a múltiplos del tamaño de una repetición armadillo. El histograma puede ajustarse a varias gaussianas, con máximos en 13, 30, 47, 65, 97, 114, 139, 163 nm ( $\sigma = 8.2$ ,  $\chi^2 = 3.7 \times 10^{-5}$ ), que pueden interpretarse aproximadamente como la longitud de 1, 2, 3, 4, 6, 7, 8 y 10 repeticiones armadillo (42 aa por repetición (Huber et al., 1997) considerando 0.4 nm por residuo (Ainavarapu et al., 2007)).

El análisis de la longitud a la que aparece el primer pico de fuerza  $L_c^0$  muestra las primeras diferencias con respecto a la región ARM aislada, además de observarse ciertas peculiaridades (**figura 4.3.d**). En este caso la longitud de la proteína plegada la estimamos

$\sim 131$  nm (5 nm x 4 I27 + 11 nm ARM + terminales desestructurados extendidos ( $aa_N + aa_C$ ) x 0.4 nm/aa). Es muy llamativo en el histograma de  $L_C^0$ , que la mayoría de picos de fuerza aparecen antes de que la molécula se alinee perpendicularmente con la superficie. El origen de estos picos de fuerza puede ser múltiple. La inclusión de los extremos terminales desestructurados podría provocar un aumento en las fuerzas de desorción de



**Figura 4.3.- Análisis nanomecánico de la  $\beta$ -catenina con controles de monomolecularidad. a)** Detalle del desplegamiento de varias moléculas  $\beta$ /I27. En la región proximal se observa el desplegamiento de la  $\beta$ -catenina seguido de los 4 módulos marcadores I27. El ajuste por medio del modelo de WLC nos permite estimar  $L_C^0$  y  $\Delta L_C$ . El pictograma representa el constructo analizado. **b)** y **c)** representaciones de  $\Delta L_C$  vs  $d$ , y  $F_D$  vs  $\Delta L_C$ . Los eventos provenientes de los marcadores I27 se muestran agrupados (elipses). **d)** Histograma de  $L_C^0$ . La distribución se ajusta a varias gaussianas centradas en 38, 52, 64, 83, 95, 122, 160 y 179 nm ( $\sigma = 6.8$ ,  $\chi^2 = 4.1 \times 10^{-5}$ ). La línea roja marca la longitud de la proteína plegada (131 nm =  $\sim 5$  nm x 4 (I27) + 11 nm ARM + longitud de los terminales desplegados). La existencia de eventos distantes a la derecha de la línea roja indica que existen regiones de la molécula que se encuentran desplegadas previamente o se despliegan con fuerzas por debajo del límite de detección. **e)** Histograma de fuerzas de la  $\beta$ -catenina. Los valores medios son  $47 \pm 2$  pN ( $n = 86$ ) y  $140 \pm 8$  pN ( $n = 12$ ), para las poblaciones a la izquierda y a la derecha del mínimo en  $p(F_D)$  (flecha negra). En línea negra se representa el ajuste a una distribución log-normal ( $F_0 = 31$ ,  $\chi^2 = 8.5 \times 10^{-5}$ ). **f)** Histograma de  $\Delta L_C$ , la distribución puede ajustarse a varias gaussianas cuyos máximos 13, 30, 47, 65, 97, 114, 138 y 163 nm ( $\sigma = 8.2$  nm,  $\chi^2 = 3.7 \times 10^{-5}$ ) corresponden aproximadamente a la longitud de una o varias repeticiones armadillo.

la molécula de la superficie, ya que las proteínas desestructuradas tienen una mayor tendencia a adsorberse inespecíficamente a superficies (Ikeda y Kuroda, 2011). Por otra parte no puede descartarse una interacción directa de los terminales con la región ARM, ya que se ha propuesto que los mismos puedan interactuar de manera dinámica con la misma (Xing et al., 2008) e incluso modular la unión del mismo con sus ligandos (Castaño et al., 2002; Choi et al., 2006). El posible mecanismo de interacción de las colas terminales

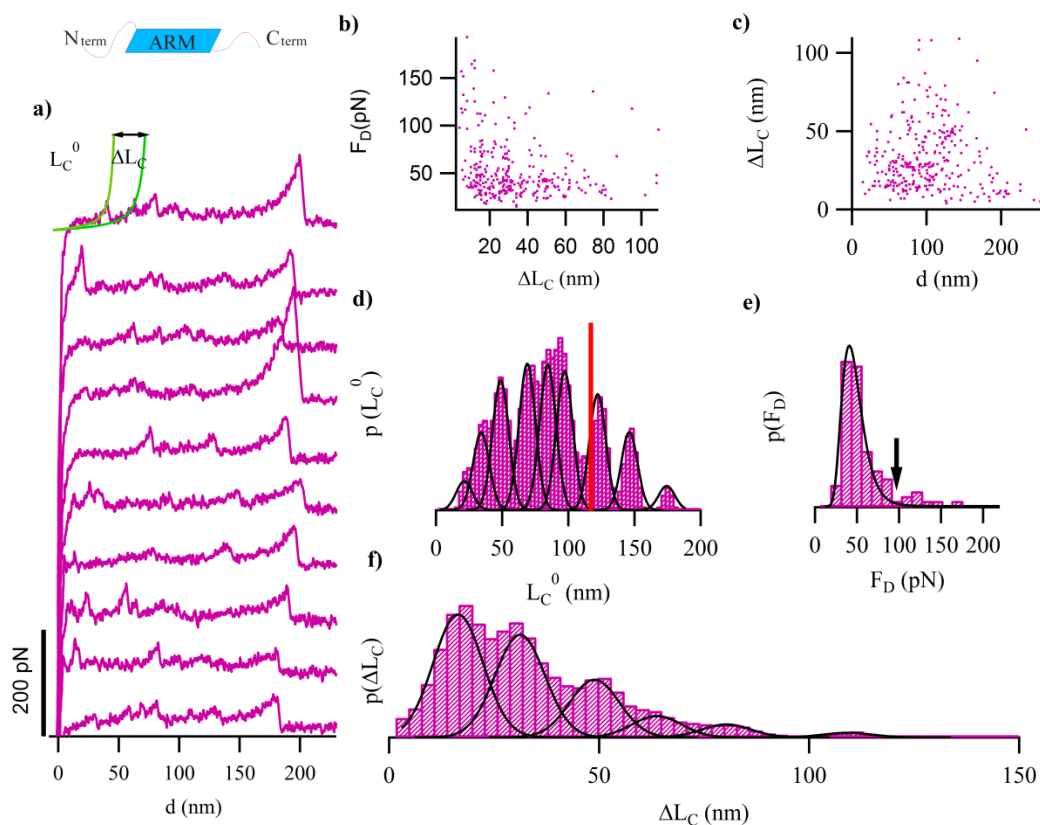
con la región armadillo se piensa que reside en la diferencia de carga que presentan los terminales y la región armadillo a pH fisiológico ( $pI \sim 4$  para los terminales,  $pI \sim 8$  para ARM) (Huber et al., 1997; Xing et al., 2008). Debido a esto, tampoco podemos excluir la posibilidad de interacción de los módulos I27 (ácido) con la región ARM, que además se vería facilitada por la inclusión en la construcción de las colas terminales que aumentan la flexibilidad. Finalmente, los eventos a una  $L_c^0$  menor que 131 nm podrían provenir del estiramiento de los terminales, aunque esta opción es poco probable ya que al ser regiones desestructuradas (Huber et al., 1997; Xing et al., 2008) en principio se espera que solo tengan resistencia de origen entrópico y no generen picos de fuerza (Li et al., 2001; Li et al., 2002; Lv et al., 2010). La información más interesante proviene de la región de los eventos que aparecen después de la longitud plegada de la molécula. De forma similar a lo que ocurre con el histograma de  $L_c^0$  de la construcción ARM/I27 (**figura 4.2.d**), el histograma de  $L_c^0$  de la construcción  $\beta$ -cat/I27 (**figura 4.3.d**) muestra que siguen apareciendo eventos a longitudes significativamente mayores a la de la longitud de la molécula “plegada”, por lo que existen repeticiones ARM que se despliegan a muy bajas fuerzas o sin presentar resistencia mecánica al comenzar el estiramiento.

La distribución de fuerzas de desplegamiento puede ajustarse a una función log-normal ( $F_0 = 31$ ,  $\chi^2 = 8.5 \times 10^{-5}$ ) como en el caso de la región ARM aislada (**figura 4.3.e** y **figura 4.2.e**). La distribución muestra un mínimo (flecha en **figura 4.3.e**) que nos permite separar los datos en dos poblaciones: una de menor estabilidad ( $<110$  pN) que tiene un valor medio de  $47 \pm 2$  pN ( $n=86$ ), y otra población de mayor estabilidad con una media de  $140 \pm 8$  pN ( $n=12$ ). Estos valores medios son muy parecidos a los obtenidos para ARM/I27, sin embargo el porcentaje de datos de mayor estabilidad mecánica ha aumentado desde el 0.4 % al 12 %. El aumento de la población de eventos de mayor estabilidad sugiere que los terminales podrían estar provocando un aumento de la mecanoestabilidad de la molécula.

### 4.1.3.- Mecánica de la $\beta$ -catenina completa sin controles de monomolecularidad.

Para descartar posibles artefactos que pudieran estar originando los marcadores de monomolecularidad en el desplegamiento mecánico de la  $\beta$ -catenina, analizamos la molécula completa aislada, sin incluir dichos marcadores (llamamos a la proteína  $\beta$ -cat **figura 4.1.a**). En principio la no inclusión de los marcadores de monomolecularidad hace más difícil la identificación de los registros provenientes de una sola molécula, sin embargo una vez conocido el patrón de desplegamiento de la proteína, obtenido con marcadores de monomolecularidad, puede utilizarse éste como guía orientativa para identificar los registros (Garcia-Manyes et al., 2007). Pese a esto, el criterio de más peso

seguido para la selección de registros fue que la molécula tuviese aproximadamente el  $L_C$  correspondiente a una molécula completa (287 nm correspondientes a los 710 aa), y que la zona proximal estuviera relativamente limpia.



**Figura 4.4.- Análisis nanomecánico de  $\beta$ -catenina.** **a)** Detalle del desplegamiento de varias moléculas de  $\beta$ -catenina. El ajuste por medio del modelo de WLC nos permite estimar  $L_C^0$  y  $\Delta L_C$ . El pictograma representa el constructo utilizado. **b)** y **c)** representaciones de  $\Delta L_C$  vs  $d$ , y  $F_D$  vs  $\Delta L_C$ . **d)** Histograma de  $L_C^0$ . La distribución puede ajustarse a varias gaussianas con máximos en 20, 33, 47, 68, 83, 96, 121, 145 y 176 nm ( $\sigma = 8.7$ ,  $\chi^2 = 2.6 \times 10^{-5}$ ). La línea roja marca la longitud de la proteína plegada (111 nm = 11 nm ARM + longitud de los terminales desplegados). La existencia de eventos distantes a la derecha de la línea roja indica que existen regiones de la molécula que se encuentran desplegadas previamente o se despliegan con fuerzas por debajo del límite de detección. **e)** Histograma de fuerzas de la  $\beta$ -catenina. Los valores medios son  $43 \pm 1$  pN ( $n = 273$ ) y  $125 \pm 4$  pN ( $n=33$ ), para las poblaciones a la izquierda y a la derecha del mínimo en  $p(F_D)$  (flecha negra). La línea negra representa el ajuste a una distribución log-normal ( $F_0 = 35$ ,  $\chi^2 = 2.3 \times 10^{-5}$ ). **f)** Histograma de  $\Delta L_C$ , la distribución puede ajustarse a varias gaussianas cuyos máximos 15, 30, 47, 62, 79 y 108 nm ( $\sigma = 8.4$  nm,  $\chi^2 = 1.3 \times 10^{-5}$ ) corresponden aproximadamente a la longitud de una o varias repeticiones armadillo.

Los registros obtenidos tienen una gran similitud con los obtenidos con ARM/I27 y  $\beta$ /I27 (figura 4.4.a, figura4.2.a y figura 4.3.a), observándose registros muy heterogéneos, con picos de fuerza que varían en número, posición y  $\Delta L_C$  entre curvas. Las representaciones de  $\Delta L_C$  vs distancia, y  $F_D$  vs  $\Delta L_C$  (figura 4.4.b y figura 4.4.c) vuelven a mostrar una gran dispersión en los datos, lo que apoya nuevamente que la  $\beta$ -catenina se despliega por múltiples rutas diferentes, en lugar de seguir una única ruta bien definida.

En la **figura 4.4.f** se representa el histograma de  $\Delta L_c$  para el constructo  $\beta$ -cat. La distribución puede ajustarse a múltiples gaussianas con máximos en 15, 30, 47, 62, 79 y 108 nm ( $\sigma = 8.4$ ,  $\chi^2 = 1.3 \times 10^{-5}$ ), que pueden adscribirse a la longitud de 1, 2, 3, 4, 5 y 6 repeticiones armadillo respectivamente. Al igual que lo que sucedía con los constructos ARM/I27 y  $\beta$ /I27, la  $\beta$ -catenina completa se despliega liberando cada vez longitudes que corresponden aproximadamente a la longitud de múltiplos de repeticiones ARM.

El análisis de los valores de  $L_c^0$  se muestran en la **figura 4.4.d**. Al igual que en el caso de la construcción  $\beta$ /I27 se pueden observar eventos tanto por encima como por debajo del umbral que representa la proteína “plegada” (línea roja, considerando esta distancia como 11 nm ARM + terminales extendidos  $(aa_N + aa_C) \times 0.4 \text{ nm/aa} = 111 \text{ nm}$ ). La ventaja que ofrece la construcción  $\beta$ -cat es que al no portar módulos marcadores de I27 podemos descartar que los eventos por debajo de la distancia de 111 nm se deban a la interacción de los I27 con el surco cargado positivamente de la región ARM. Sin embargo, y de forma similar a lo que ocurría en el caso de la construcción  $\beta$ /I27, no podemos descartar que estos eventos se deban a interacciones con la superficie, a interacciones de los terminales con el ARM, o a picos de desplegamiento provenientes de las colas terminales. En cualquier caso, nuevamente vuelven a aparecer eventos muy por encima de la longitud de la proteína plegada por lo en algunas moléculas existen regiones que se despliegan sin resistencia mecánica (o desplegarse a fuerzas más bajas que nuestro nivel de detección) al comenzar el estiramiento.

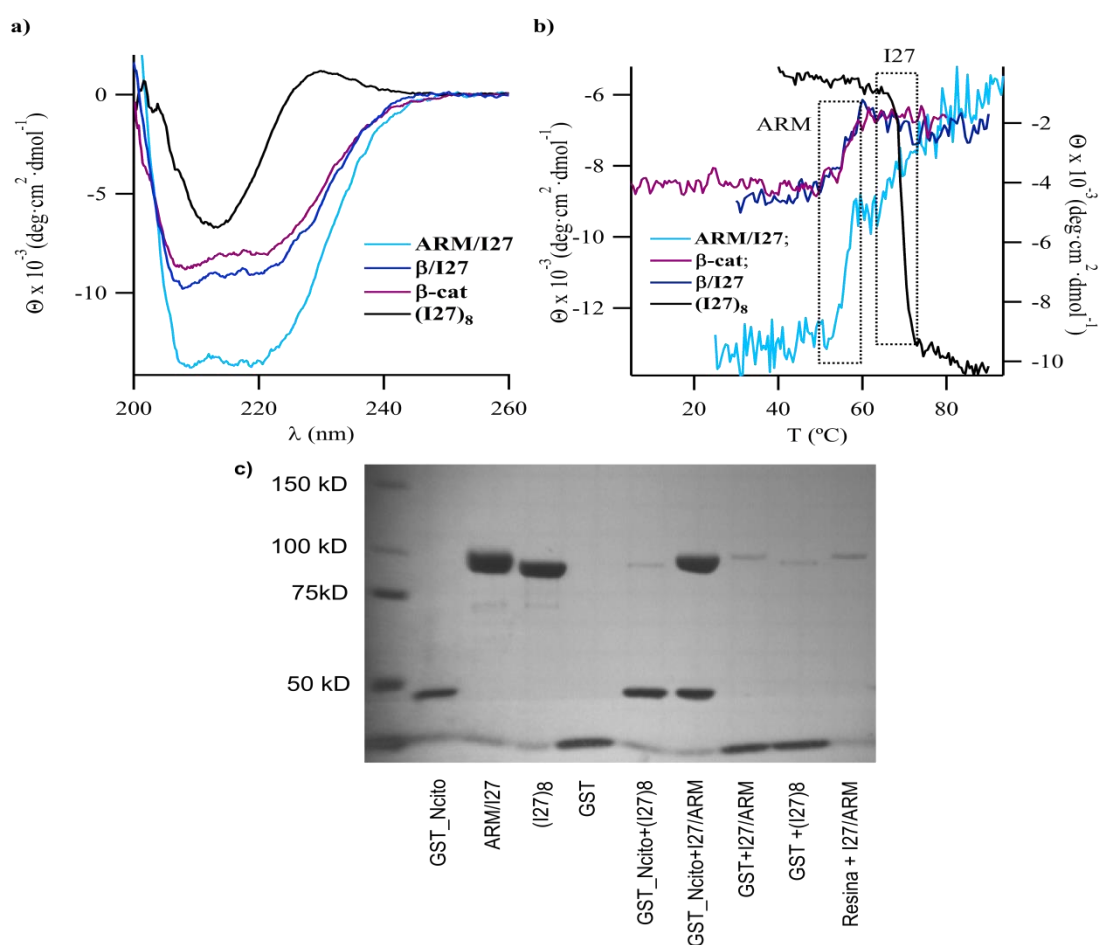
La distribución de fuerzas puede ajustarse a una función log-normal ( $F_0 = 35$ ,  $\chi^2 = 2.3 \times 10^{-5}$ , línea negra en **figura 4.4.e**). Como en el caso de la construcción  $\beta$ /I27, la distribución de fuerzas muestra un mínimo (flecha en **figura 4.4.e**) que nos permite separar los datos en dos poblaciones: una de menor estabilidad que tiene un valor medio de  $43 \pm 1 \text{ pN}$  ( $n = 273$ ), y otra población de mayor estabilidad con una media de  $125 \pm 4 \text{ pN}$  ( $n = 33$ , 11% de la población). Tanto los valores medios, como los porcentajes de las dos poblaciones son muy parecidos a los obtenidos para  $\beta$ /I27, lo que refuerza la interpretación de que las regiones terminales originan una estabilización de la molécula (la ausencia de módulos I27 en este constructo permite asegurar que la población de mayor estabilidad se debe a la presencia de los terminales).

En conclusión, nuestros resultados experimentales indican que el desplegamiento mecánico de la región ARM y de la  $\beta$ -catenina completa ocurre de una manera muy similar, mostrando fuerzas de desplegamiento medio muy parecidas, similares perfiles de  $\Delta L_c$ , y caminos de desplegamiento múltiples con varios intermedios. Además, los terminales N y C de la  $\beta$ -catenina podrían estar modulando la estabilidad mecánica de la proteína, como sugiere el aumento de frecuencia de picos de alta fuerza (superior a 100 pN) en las

distribuciones de  $F_D$  de los dos constructos que portan los terminales N y C (0.4 % para ARM/I27 y 11-12 % para  $\beta$ /I27 y  $\beta$ -cat).

#### 4.1.4.- Controles estructurales y de unión a ligando.

El uso de heteropoliproteínas como marcadores de monomolecularidad, se basa en la premisa de que cada uno de los módulos constituyentes de la poliproteína mantiene su plegamiento de manera autónoma. La premisa del plegamiento autónomo en proteínas modulares se ha verificado experimentalmente (Batey et al., 2008), y el uso de heteropoliproteínas para caracterizar el desplegamiento mecánico de módulos proteicos



**Figura 4.5.- Controles estructurales y ensayos de co-precipitación. a)** Espectros de DC (25 °C) de las construcciones ARM/I27,  $\beta$ /I27,  $\beta$ -cat y (I27)<sub>8</sub>. En los espectros se observa claramente que la  $\beta$ -catenina mantiene su estructura secundaria cuando está formando parte de la heteropoliproteína  $\beta$ /I27, ya que los dos mínimos típicos de la estructura se mantienen en ARM/I27. La deconvolución de los espectros (ver **sección 3.4**) muestra que los porcentajes de estructura secundaria  $\alpha$  y  $\beta$  se aproximan mucho a los valores teóricos esperados, en especial en la construcción ARM/I27 (ver **tabla 4.1**). **b)** Espectros de desnaturalización térmica de las muestras. La desnaturalización se siguió monitorizando los cambios de elipticidad a 210 nm para ARM/I27,  $\beta$ /I27 y 224 nm para (I27)<sub>8</sub> (ver **sección 3.4**), y se aplicó un rampa de temperatura de 60 °C/h). Las heteropoliproteínas



muestran dos transiciones diferentes durante la desnaturalización con dos  $T_{1/2}$  muy similares para ARM/I27 y  $\beta$ /I27:  $T_{1/2}^1 \sim 55.9$  °C;  $T_{1/2}^2 \sim 65.4$  °C para ARM/I27, y  $T_{1/2}^1 \sim 55.7$  °C;  $T_{1/2}^2 \sim 67.8$  °C para  $\beta$ /I27. Las muestras  $\beta$ -cat e  $(I27)_8$  muestran una sola transición:  $T_{1/2} \sim 56.1$  °C para  $\beta$ -cat y  $T_{1/2} \sim 69.2$  °C  $(I27)_8$ . Los recuadros en línea discontinua señalan las desnaturalizaciones de los I27 y de la región ARM. **c)** Ensayo de co-precipitación entre GST\_Ncito y ARM/I27. En los cuatro primeros carriles se cargaron las diferentes proteínas utilizadas en el ensayo a modo de marcadores. En los cinco carriles restantes los resultados de las co-precipitaciones. De la comparación del sexto carril con el resto de carriles experimentales se deduce que I27/ARM interacciona específicamente con el dominio citosólico de la N-cadherina de GST\_Ncito. Los ensayos se realizaron a una concentración de proteína final de  $2.5 \mu\text{M}$  y una relación molar 1:1.

por AFM-SMFS ha sido validada en varias ocasiones (Hervas et al., 2012; Randles et al., 2007; Sharma et al., 2007; Valbuena et al., 2009). El comportamiento tan similar que presentan las tres construcciones recombinantes analizadas (I27/ARM,  $\beta$ /I27 y  $\beta$ -cat) validan también en nuestro caso esta estrategia. Sin embargo, los resultados observados en las proteínas (múltiples caminos de desplegamiento y repeticiones que parecen desplegadas al inicio del experimento) podrían interpretarse como originados por la existencia de algún defecto en el plegamiento de las proteínas recombinantes. Para testar esta posibilidad llevamos a cabo una serie de experimentos para controlar si las proteínas mantenían su estructura. Concretamente llevamos a cabo controles estructurales con técnicas de baja resolución (dicroísmo circular, DC), y ensayos de unión de la región ARM a un ligando natural (el dominio citoplasmático de la N-cadherina).

Los controles estructurales de baja resolución los llevamos a cabo en colaboración con la Dra. Margarita Menéndez (Instituto de Química-Física Rocasolano, CSIC). El DC se basa en medir los cambios de elipticidad que sufre una onda circularmente polarizada en el plano al atravesar una muestra ópticamente activa. En el caso de proteínas, los espectros de DC en el rango del UV lejano se deben principalmente a los enlaces amida peptídicos, y los diferentes tipos de estructura secundaria de las proteínas (hélice  $\alpha$ , lámina  $\beta$ , u ovillo estadístico) originan mínimos y máximos característicos en los espectros. Así pues los experimentos de DC nos permiten hacer una estima del tipo y porcentaje de estructura secundaria de las muestras, y por comparación con la estructura secundaria calculada a partir de las estructuras conocidas, conocer si existen grandes desviaciones respecto a los valores esperados.

En la **figura 4.5.a** se representan los espectros de DC de las tres construcciones recombinantes ARM/I27,  $\beta$ /I27,  $\beta$ -cat y de una poliproteína de I27  $(I27)_8$ , está última para poder comparar la contribución de las repeticiones I27 a los espectros. Los espectros de ARM/I27,  $\beta$ /I27 y  $\beta$ -cat presentan mínimos cercanos a 208 y 222 nm típicos de  $\alpha$ -hélice, de hecho la deconvolución de los espectros para estimar el porcentaje de hélice  $\alpha$  o lámina  $\beta$  muestran una buena concordancia con los porcentajes teóricos, especialmente en el caso del constructo ARM/I27 (**tabla 4.1**). Para extraer información adicional, desnaturalizamos

las muestras mediante rampas de temperatura (60 °C/h) (**figura 4.5.b**). Es interesante destacar que las heteropoliproteínas muestran dos transiciones diferentes durante la desnaturalización, con dos  $T_{1/2}$  (temperatura a la cual la mitad de la población se encuentra en un estado plegado y la otra mitad en el desplegado) muy similares para ARM/I27 y  $\beta$ /I27:  $T_{1/2}^1 \sim 55.9$  °C;  $T_{1/2}^2 \sim 65.4$  °C para ARM/I27, y  $T_{1/2}^1 \sim 55.7$  °C;  $T_{1/2}^2 \sim 67.8$  °C para  $\beta$ /I27. La comparación con los espectros de desnaturalización de las muestras  $\beta$ -cat ( $T_{1/2} = 56.1$  °C) y (I27)<sub>8</sub> ( $T_{1/2} = 69.2$  °C) nos indica que estas transiciones parecen deberse al desplegamiento de la región ARM, y de los módulos I27 respectivamente, por lo que estos resultados confirman que ambas estructuras mantienen su plegamiento independiente dentro de las heteropoliproteínas.

Construcción	$\alpha$ -hélice teórico	$\alpha$ -hélice estimado	Hoja $\beta$ teórico	Hoja $\beta$ estimado
<b>ARM/I27</b>	40.7 %	40-41 %	14.7 %	15 %
<b><math>\beta</math>/I27</b>	33 %	29-23 %	11.9 %	20-22 %
<b><math>\beta</math>-cat</b>	49 %	21-26 %	-	22-26 %
<b>(I27)<sub>8</sub></b>	-	5 %	35.7 %	44 %

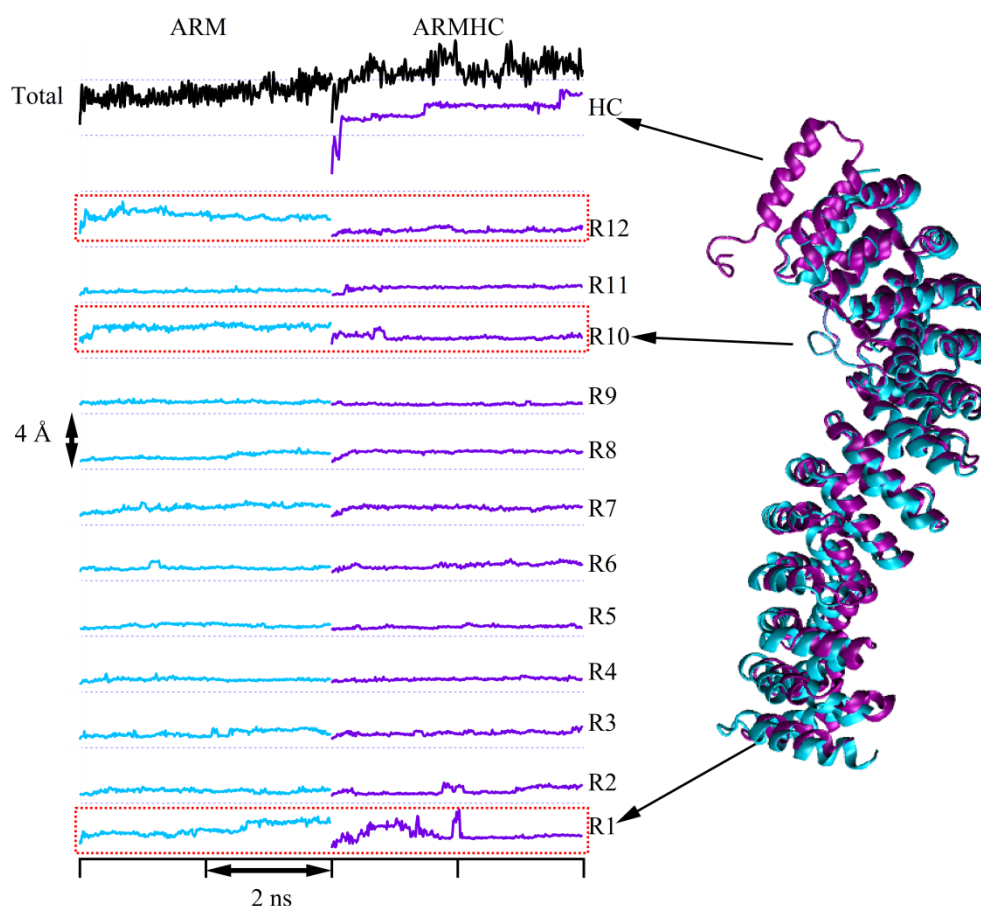
**Tabla 4.1.-** Resumen de los porcentajes de  $\alpha$ -hélice y hoja  $\beta$  estimados a partir de los espectros de DC y comparación con los valores teóricos esperados.

Para obtener otra evidencia de que las proteínas mantienen su estructura nativa y de que su comportamiento por AFM-SMFS no se debe a una conformación anómala, elegimos una vía experimental alternativa: la unión a un ligando. Elegimos una de las construcciones (ARM/I27) para verificar que puede interactuar con uno de los ligandos naturales de la  $\beta$ -catenina: el dominio citoplasmático de la N-cadherina (que interactúa con prácticamente toda la longitud de la región ARM). Para ello diseñamos unos sencillos ensayos de co-precipitación. En estos ensayos el dominio citoplasmático de la N-Cadherina fusionado a GST (*glutathione s transferase*) se incubó con la construcción ARM/I27, tras la incubación se incorporó una resina de afinidad a GST que se precipita y se lava con tampón varias veces. Si las proteínas interactúan, la construcción ARM/I27 co-precipitaría con la resina de afinidad por medio de su interacción con GST\_Ncito (GST fusionado a dominio citosólico de la N-Cadherina). La **figura 4.5.c** muestra que ARM/I27 co-precipita con gran afinidad con GST\_Ncito, y que esa interacción es específica de la región ARM, ya que los controles muestran que la co-precipitación no se debe a interacciones inespecíficas con la resina o entre la GST\_Ncito/(I27), GST/ARM o GST/(I27).

La suma de evidencias estructurales y los ensayos de co-precipitación nos llevan a concluir con confianza que las proteínas utilizadas para los experimentos de desplegamiento mecánico por AFM se encuentran en su conformación nativa.

#### 4.1.5.- Simulaciones de dinámica molecular de la $\beta$ -catenina.

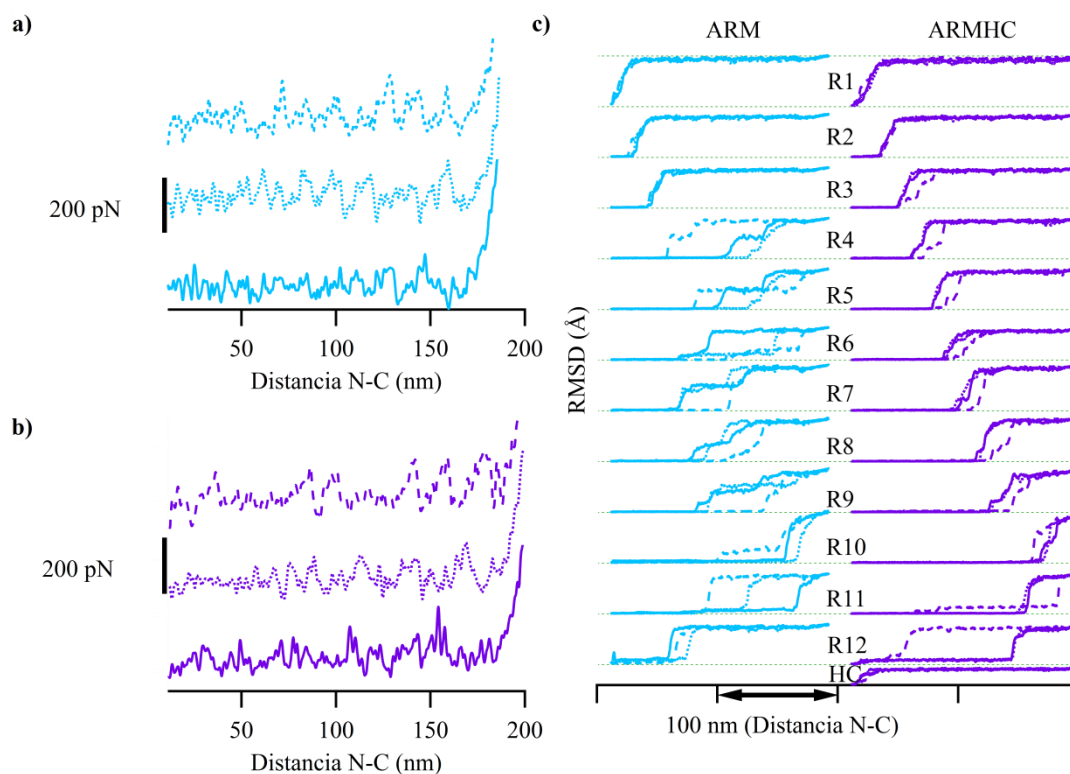
Finalmente, llevamos a cabo simulaciones de dinámica molecular para poder estudiar los detalles del desplegamiento mecánico de la  $\beta$ -catenina. Para estudiar el desplegamiento de la región ARM se utilizó una estructura que comprendía la región ARM humana (código PDB 1qz7, (Xing et al., 2003)), y para intentar emular el efecto de los terminales utilizamos una estructura que comprendía la región ARM más una hélice  $\alpha$  extra de la región c-terminal (código PDB 2z6h, (Xing et al., 2008)). La estructura 1qz7 la llamaremos a partir de ahora **ARM** (en negrita, para distinguirla de la región estructural ARM), y la estructura 2z6h **ARMHC**.



**Figura 4.6.- Resumen de los RMSD del esqueleto peptídico de ARM y ARMHC durante 4 ns de dinámica libre.** En negro, arriba, se muestra el RMSD total de las dos estructuras, observándose una gran desviación en el caso de **ARMHC** por la contribución de los residuos terminales libres que no forman parte de la hélice  $\alpha$  de la hélice C. En los recuadros rojos se resaltan las repeticiones R1,

R10 y R12, que son las que mayores diferencias presentan entre ambas estructuras, y que tienen mayor desviación en el caso de **ARM**. A la derecha, en azul **ARM** y en morado ARMHC.

Las desviaciones (RMSD, de *root mean square deviation*) que sufren los esqueletos peptídicos de **ARM** y **ARMHC** durante una dinámica libre de 4 ns quedan recogidas en la **figura 4.6**. Las desviaciones totales que sufren ambas estructuras no son muy informativas, ya que los residuos C-terminales de **ARMHC** no forman parte de la hélice  $\alpha$  extra (HC) y sufren grandes fluctuaciones, lo que provoca un mayor RMSD de **ARMHC**. Si estudiamos la desviación de cada repetición armadillo, observamos en **ARM** que la mayoría de repeticiones sufren desviaciones muy similares, a excepción de R1, R10 y R12. Como cabría esperar, las repeticiones localizadas en los extremos libres de la región armadillo (repeticiones R1 y R12) muestran mayor movilidad que las que se encuentran en la parte central de la estructura. La mayor desviación de R10 puede atribuirse a que se trata de una repetición de mayor tamaño que presenta un largo bucle (loop) con una estructura secundaria fluctuante. En el caso de **ARMHC**, las desviaciones de R1, R10 y R12 son equiparables a las del resto de repeticiones de la región armadillo. Parece que las interacciones entre la hélice C y R12 (tres residuos hidrofóbicos y dos puentes de hidrógeno [Xing et al., 2008]), llevan a toda la estructura a un estado de menor energía, y R1, R10 y R12 no se desvían mucho.



**Figura 4.7. - Simulación del estiramiento de ARM y ARMHC.** Curvas de fuerza-extensión de los tres estiramientos de **ARM** (a) y **ARMHC** (b). c) RMSD del esqueleto peptídico de cada repetición durante los *steereds*. Con esta representación seguimos el desplegamiento de cada repetición y observamos que las estructuras pueden desplegarse por caminos diferentes. El camino hacia el

estado desplegado es más restrictivo en el caso de **ARMHC**. Las líneas continuas indican el primer estiramiento, la línea de puntos el segundo, y la línea formada por guiones el tercer estiramiento.

En la **figura 4.7.a** se representan 3 estiramientos de dinámica molecular (*steered*) de **ARM**. De manera similar a lo que se observó experimentalmente, los registros de desplegamiento mecánico muestran baja estabilidad mecánica comparado con el desplegamiento de I27 bajo las mismas condiciones (Valbuena et al., 2009). Para seguir la secuencia de desplegamiento durante las simulaciones representamos el RMSD del esqueleto peptídico de cada repetición ARM (**figura 4.7.c**). A excepción de R1, R2 y R3 que se despliegan siempre primero en las tres simulaciones, el resto de repeticiones tienen trayectorias de desplegamiento variable, con la presencia de intermediarios y la particularidad de que el desplegamiento ocurre a diferentes distancias N-C en las tres trayectorias (ver por ejemplo R4, R5 y R11). Es decir, de forma similar a lo que ocurre en los experimentos, la trayectoria de desplegamiento entre *steereds* es diferente.

Los *steereds* de la estructura **ARMHC** muestran la misma jerarquía en el desplegamiento de las repeticiones R1, R2 y R3, con la particularidad añadida de que las tres trayectorias son muy similares para el resto de las repeticiones, y existe una marcada direccionalidad N-C en el camino de desplegamiento. **ARMHC** parece estar en un mínimo de energía más profundo durante el desplegamiento, ya que presenta casi el mismo camino de desplegamiento en las tres trayectorias. Así pues, parece ser que pequeños cambios en la conformación pueden llevar a la estructura a diferentes estados energéticos, conduciendo a diferentes caminos de desplegamiento. Aunque nosotros no tenemos evidencia experimental de la direccionalidad N-C en el desplegamiento, sí observamos cierta estabilización mecánica de la molécula en la presencia de las colas terminales, lo que iría en la misma dirección que los resultados de las simulaciones.

En resumen las simulaciones por DM reproducen cualitativamente, al menos en el caso de la estructura ARM, la variabilidad en el camino de desplegamiento mecánico (y las  $F_D$  de baja magnitud) observadas en los experimentos. Además los resultados obtenidos con **ARMHC** indican una cierta estabilización de las repeticiones ARM por interacciones con la hélice C, un efecto que pudiera estar ocurriendo también en nuestros experimentos en la presencia de las colas N- y C- terminales completas.

#### **4.2.- Estrategia autocontrolada para el estudio de interacciones intermoleculares por AFM-SMFS.**

Como ya se ha introducido en la **sección 1.4.2**, el segundo bloque temático de esta tesis se centra en implementar sobre una interacción modelo, una estrategia autocontrolada para el estudio de interacciones intermoleculares por AFM-SMFS. El objetivo es demostrar, a nivel de prueba de concepto, que este diseño experimental para el estudio de interacciones intermoleculares de proteínas por AFM-SMFS es viable.

La estrategia propuesta descansa en la incorporación en las proteínas a estudio de controles que permitan la identificación directa del pico de ruptura de la interacción, y aseguren condiciones de molécula individual. Esta estrategia haría también innecesario el uso de complejos procesos de funcionalización típicos de los estudios de interacciones por AFM-SMFS (ver **sección 1.4.2**). Para asegurar condiciones de molécula individual se incorporan pequeñas poliproteínas a las proteínas de interés (**figura 1.10 y 4.8**). Para marcar el pico de fuerza de ruptura de la interacción se utiliza un polipéptido que se une en paralelo a las proteínas interaccionantes a estudio y que está protegido de la fuerza por la interacción (**figura 1.10**). De esta forma cuando se produce la ruptura de la interacción entre las moléculas a estudio, este polipéptido queda expuesto a la fuerza y se produce la liberación de un  $\Delta L_c$  dado (ver **sección 1.4.2.1** para una descripción más detallada)

##### **4.2.1.- Diseño experimental: elección de los elementos del sistema.**

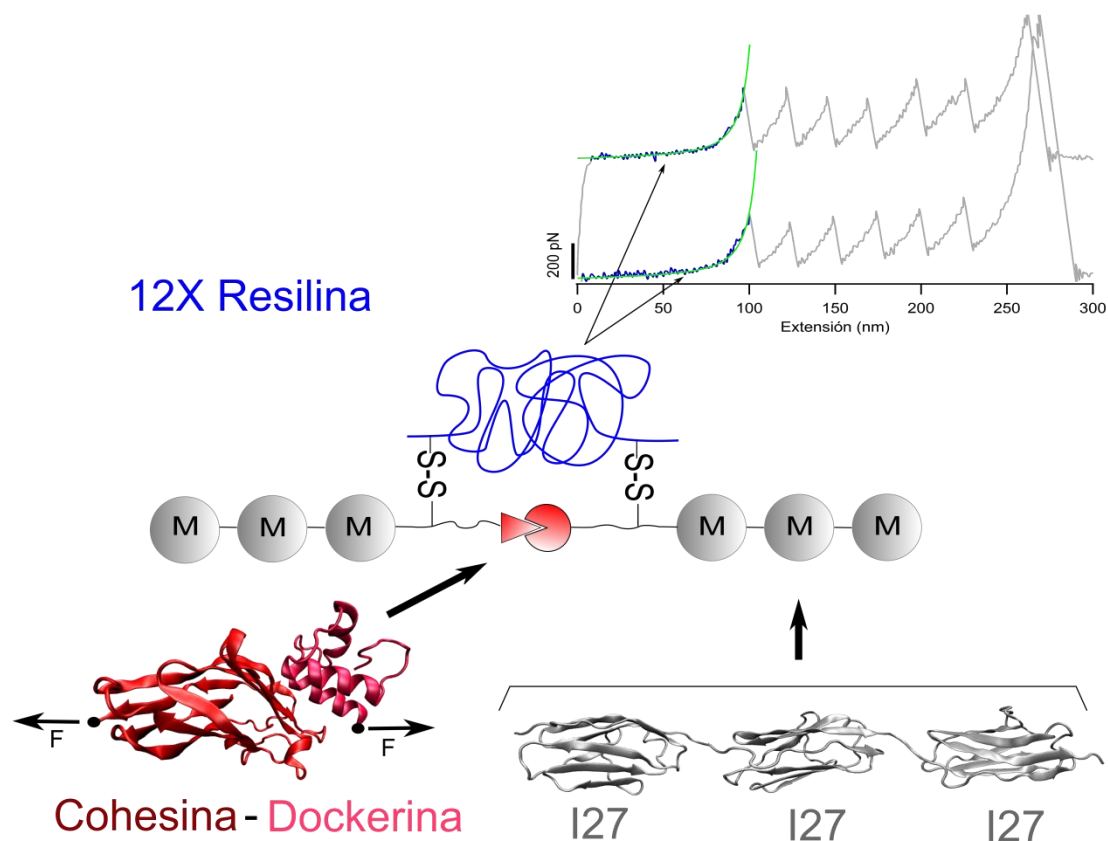
El esquema experimental propuesto queda resumido conceptualmente en la **sección 1.4.2.1**. Una vez determinado el diseño general, es necesario elegir una interacción modelo sobre la que implementar el diseño, utilizar un módulo como marcador de monomolecularidad (mini-poliproteína), y un polipéptido que mantenga la continuidad mecánica en el sistema tras la ruptura de la interacción (y que servirá también de marcador o reportero del pico de ruptura de la interacción) (**figura 1.10 y 4.8**).

La interacción modelo elegida para implementar la estrategia fue la interacción entre la cohesina 7 y la dockerina de la celulasa Cel8A del celulosoma de *Clostridium thermocellum*. Desde el punto de vista de la implementación de la estrategia, la principal ventaja que presenta este sistema es la gran afinidad que presentan estas proteínas entre sí (en el rango nanomolar (Mechaly et al., 2001; Schaeffer et al., 2002)). Además de las ventajas técnicas, el estudio nanomecánico del sistema es interesante desde el punto de vista fisiológico como se describe a continuación.

Los celulosomas son grandes complejos macromoleculares bacterianos capaces de llevar a cabo la degradación eficiente de celulosa recalcitrante mediante la acción

coordinada de diversas enzimas hidrolíticas (Himmel et al., 2010). Los módulos cohesina de la escafoldina (proteína de andamiaje del celulosoma) presentan zonas de unión a los dominios dockerina de las enzimas hidrolíticas. Al constituir en definitiva un sistema de adhesión entre las células bacterianas y las partículas de celulosa, los celulosomas están previsiblemente expuestos a tensión mecánica, de ahí el interés en conocer las propiedades mecánicas de sus componentes (Valbuena et al., 2009; Schoeler et al., 2014).

*A priori* existen cuatro configuraciones posibles en las que se puede estirar el complejo, desde los N-terminales de ambas proteínas, desde sus C-terminales, y desde el C-terminal de una y el N-terminal de la otra (que genera dos posibilidades). Nosotros nos decantamos por estirar las proteínas desde los C-terminales por razones de tipo experimental. Por experiencia previa sabemos que estas proteínas tienden a ser más estables (a la degradación) cuando se producen en el N-terminal de constructos recombinantes.



**Figura 4.8.- Elementos del diseño experimental.** El pictograma del centro de la figura representa conceptualmente los diferentes componentes de la estrategia autocontrolada para la medida de interacciones moleculares. La interacción modelo a estudio (rojo y rosa) es la interacción entre cohesina 7 y la dockerina de la celulasa Cel8A del celulosoma de *C. thermocellum*. Como módulo marcador de monomolecularidad se seleccionó el módulo I27 de titina cardiaca humana (en gris). Finalmente se muestra en azul el elastómero protegido de la fuerza, constituido por 12 repeticiones (12XR) de un mutante de una secuencia consenso de resilina (Lv et al., 2010). En la figura se incluyen dos registros de fuerza-extensión obtenidos con una poliproteína heteromérica de I27 y 12XR donde se observa el comportamiento de muelle entrópico de las repeticiones 12XR (en verde se muestra el ajuste al modelo WLC) (ver sección 4.2.4.3).

Como marcadores de monomolecularidad se pueden utilizar repeticiones de cualquier módulo que haya sido anteriormente empleado en estudios de AFM-SMFS como marcador, como es el caso del módulo I27, Gb1 o Ubicuitina (Sharma et al., 2006; Hervas et al., 2012; Scholl et al., 2014). Nosotros elegimos el módulo I27 de titina, que ha sido extensamente usado como marcador con anterioridad, siendo un sistema modelo de referencia en el campo (Ainavarapu et al., 2005; del Rio et al., 2009; Ma et al., 2009; Valbuena et al., 2009), y del que sabemos por experiencia previa que se expresa eficientemente de manera recombinante y con buenas propiedades de solubilidad.

Uno de los puntos críticos del diseño reside en elegir una proteína que se comporte como un muelle entrópico puro (sin barreras mecánicas a la fuerza), y que sea lo suficientemente grande para que cuando se estire en paralelo con la interacción de interés asuma una cantidad despreciable de la carga de fuerza. Esto provoca que sea la interacción a estudio la que resista principalmente el estiramiento. En nuestro caso decidimos utilizar un mutante (Y8T) de la repetición consenso de resilina utilizada en (Lv et al., 2010). Proteínas formadas por hasta 12 repeticiones de la secuencia utilizada en ese trabajo muestran una conformación de ovillo estadístico por dicroísmo circular, y los resultados presentados demuestran que las repeticiones se comportan como muelles entrópicos perfectos por AFM-SMFS (sin resistencia detectable a la fuerza y que ajustan al modelo de WLC).

Así pues decidimos utilizar como polipéptido elastomérico sin resistencia a la fuerza 12 repeticiones del citado mutante (que a partir de ahora referiremos como 12XR). En el caso de que el polipéptido se ajuste a WLC, y no presente barreras a la fuerza, debería tener una longitud de contorno esperada de  $L_c^e \sim 84$  nm (15 residuos por repetición x 12 repeticiones + residuos extras provenientes de la estrategia de clonaje = 210 residuos, que a 0.4 nm/residuo (Ainavarapu et al., 2007) suponen un  $L_c \sim 84$  nm). En la **sección 4.2.4.2** se presentan los resultados que muestran que el polipéptido seleccionado cumple los requisitos exigidos. Por simplicidad a partir de ahora llamaremos **elastómero** a esta proteína que se comporta como un muelle entrópico (aunque la definición estricta de elastómero es más amplia).

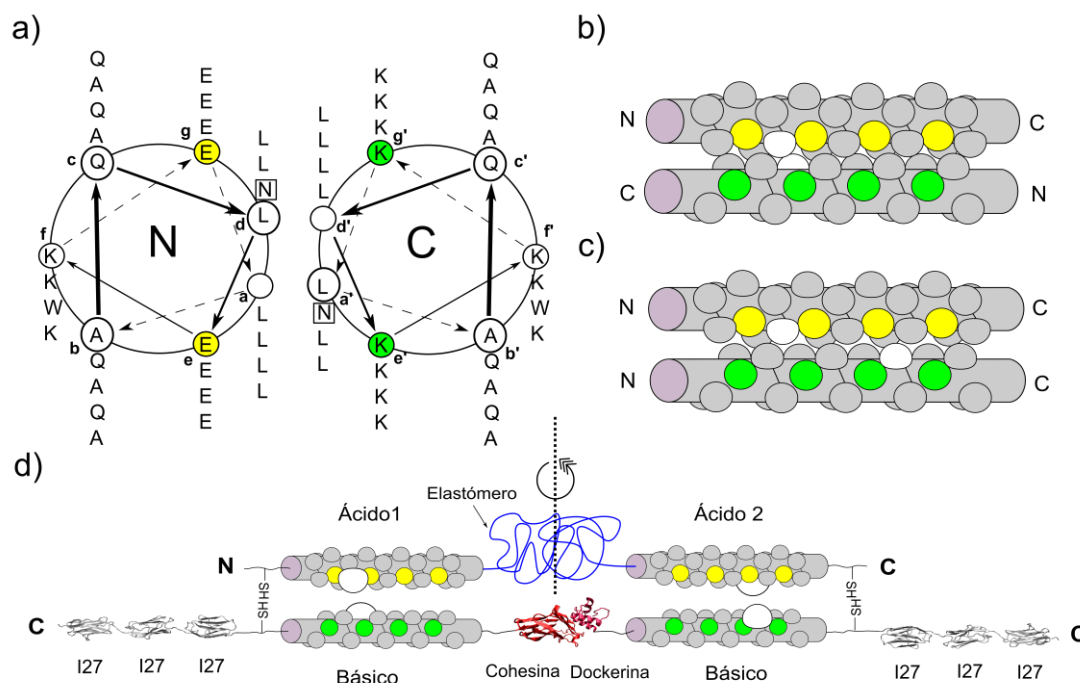
#### **4.2.2.- Diseño experimental: diseño de dominios de oligomerización.**

Como se ha discutido en la **sección 1.4.2.1**, la mejor opción para conseguir la unión entre las proteínas de interés y el polipéptido elastomérico es la utilización de puentes disulfuro. La formación de un enlace disulfuro entre residuos de cisteínas requiere que estos estén espacialmente enfrentados. La formación de estos enlaces al azar



sería muy ineficiente con el agravante de que podrían originarse enlaces disulfuro “aberrantes” con cisteínas de los marcadores, de las proteínas de interés, o en configuraciones no deseadas (como por ejemplo entre las dos cisteínas del polipéptido elastomérico). Por ello se hace necesaria la incorporación al sistema de un último componente: algún tipo de dominio de oligomerización que dirija la formación de los puentes disulfuro aproximando espacialmente los grupos sulfhidrilos de las cisteínas.

La elección de los dominios de oligomerización viene delimitada por una serie de restricciones. Idealmente deben ser estructuralmente pequeños, para que no se produzcan impedimentos estéricos con el resto de componentes del sistema. Para que las cisteínas queden enfrentadas, y se pueda dirigir de manera efectiva la producción de puentes disulfuro, es esencial que los extremos de dichos motivos queden espacialmente cercanos tras la oligomerización. Por todo ello, los dominios de oligomerización heterodiméricos tipo ‘Velcro’ son en nuestro caso la mejor elección. Se trata de cremalleras de leucina sintéticas, estructuralmente pequeñas (de unos 30 aminoácidos) y se conocen perfectamente las causas que hacen que estos péptidos ‘Velcro’ se asocien para formar heterodímeros, y los fundamentos que determinan la orientación paralela/antiparalela de los mismos (O’Shea et al., 1993; Lumb y Kim, 1995; Oakley y Kim, 1998). Este conocimiento profundo nos permite diseñar nuestros propios dominios de oligomerización ‘Velcro’ según nuestras necesidades.



**Figura 4.9.- Diseño de dominios de oligomerización.** a) Representación helicoidal del heterodímero antiparalelo Velcro Ácido1-Básico. b) y c) representación esquemática de complejos antiparalelo y paralelo del heterodímero Ácido1-Básico, en círculos blancos se representan los residuos de asparagina. En la orientación antiparalela estos residuos quedan enfrentados y pueden

interaccionar. Por el contrario, en el complejo paralelo las asparaginas quedan muy separadas, quedando esta configuración muy desfavorecida. **d)** Pictograma con la representación final del diseño experimental propuesto. La localización de los residuos de asparagina en los dominios Ácido1 y Ácido2 promueve que estos dominios interaccionen, respectivamente, de manera antiparalela y paralela con el dominio básico de las construcciones Cohesina/Dockerina Básico (I27)<sub>3</sub>. En la figura, la construcción Ácido1\_12XR\_Ácido2 se orienta de izquierda a derecha desde el N- al C-terminal presentando así orientación antiparalela Ácido2-Básico, y paralela Ácido1-Básico. Hay que tener en cuenta que en la orientación opuesta (rotando 180° Ácido1\_12XR\_Ácido2 perpendicular al eje N-C), seguiría manteniendo una orientación antiparalela Ácido2-Básico, y paralela Ácido1-Básico.

Los péptidos 'Velcro' tienen una estructura denominada de hélices superenrolladas (*coiled coils*). La estructura de hélice superenrollada de las cremalleras de leucina consiste en dos alfa-hélices dextrógiras enrolladas una sobre otra con un giro superhelicoidal ligeramente levógiro. Su secuencia consiste en repeticiones de 7 residuos aminoacídicos, numerados con las letras **a-g**, estando las posiciones **a** y **d** ocupadas por residuos hidrofóbicos y las posiciones **e** y **g** por residuos cargados (O'Shea et al., 1993) (**figura 4.9.a**). Los heterodímeros 'Velcro' difieren en los residuos de las posiciones **e** y **g**, escogiéndose ácido glutámico para esas posiciones en un monómero (dominio Ácido), y lisina en el otro monómero (dominio Básico), de manera que los homodímeros son desestabilizados a pH neutro por interacciones electrostáticas desfavorables, conduciéndose la formación del heterodímero. Las posiciones hidrofóbicas **a** y **d**, que representan la superficie principal de interacción, están ocupadas por leucinas, y se incluye un residuo polar (asparagina) en la posición **a** de una de las repeticiones. Esta posición polar enterrada en el núcleo apolar de la estructura es la que determina la orientación de los dímeros por formación de un enlace de puente de hidrógeno con la asparagina de la posición opuesta (Oakley y Kim, 1998) (**figura 4.9.b y c**). De esta forma si se incluye una asparagina en una de las repeticiones cercanas al N-terminal de cada uno de los monómeros (Ácido y Básico), se está favoreciendo la orientación paralela, ya que solo en esta configuración las asparaginas de ambas hélices quedarían enfrentadas. Si por el contrario la asparagina se incluyera cerca de extremos opuestos de las dos clases de monómeros, el enfrentamiento de las asparaginas solo ocurriría con orientación antiparalela (ver **figura 4.9.b y c**).

Para implementar la configuración de estiramiento seleccionada (desde los C-terminales de ambas proteínas), se diseñaron 3 dominios 'Velcro': 2 dominios ácidos (Ácido1 y Ácido2), y uno básico (**figura 4.9.d**). Los dos dominios ácidos se disponen a ambos lados del elastómero, cerca del N- y del C terminal, y las construcciones que portan la cohesina y la dockerina portan el mismo tipo de dominio básico. Los dominios ácidos se diferencian entre sí en la posición de la asparagina de la posición **a** (**figura 4.9.d**), de tal manera que el dominio básico interacciona de manera antiparalela con el dominio ácido

más próximo al N-terminal, y paralelamente con el del C-terminal (**figura 4.9.b, c y d**, ver **sección 3.1** para consultar secuencia).

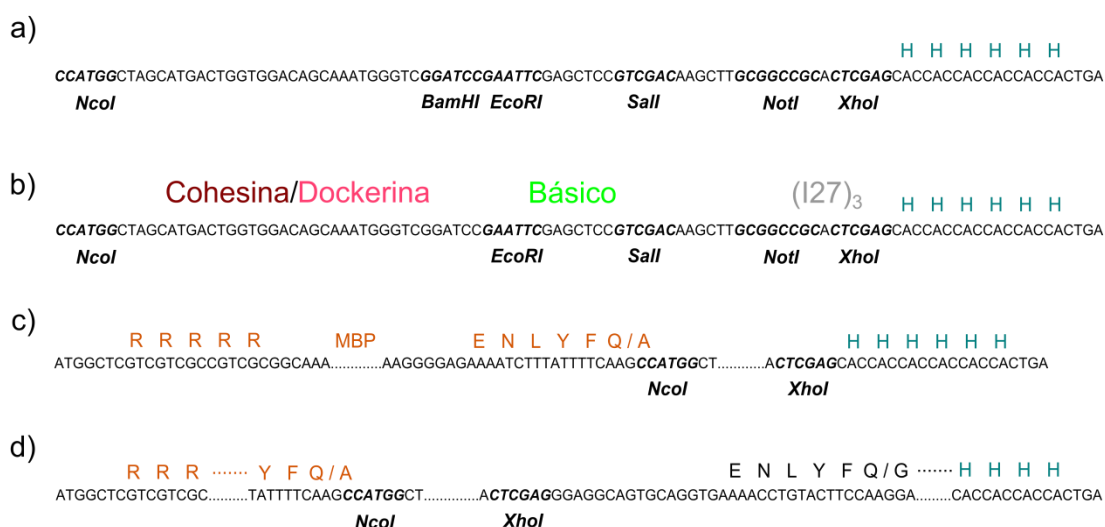
El diseño experimental final requiere pues de 3 construcciones recombinantes:

- Cohesina Básico (I27)<sub>3</sub>: porta de N- a C-terminal el módulo Cohesina 7, el dominio de oligomerización Básico y tres repeticiones del módulo I27 como marcador de monomolecularidad.
- Dockerina Básico (I27)<sub>3</sub>: porta de N- a C-terminal, la Dockerina de la celulasa Cel8A, el dominio de oligomerización Básico y tres repeticiones del módulo I27 como marcador de monomolecularidad.
- Ácido1\_12XR\_Ácido2: porta de N- a C-terminal el dominio de oligomerización Ácido1, 12 repeticiones de resilina, y el dominio de oligomerización Ácido2.

#### 4.2.3.- Diseño experimental: generación de una batería de vectores.

Para implementar el diseño experimental de manera versátil, se implementó una estrategia de clonaje tipo *cassette*. Con esta estrategia, cada uno de los elementos que conforman el sistema, como las secuencias codificantes de la proteína a estudio, los dominios de oligomerización, y las repeticiones de módulos marcadores, se clonan siempre entre dianas de restricción específicas, lo que permite intercambiar y sustituir de manera modular los distintos elementos por otros diferentes que lleven a cabo la misma función.

Para implementar los vectores se tomó como base el vector comercial pET-24d (Novagen). Para construcciones que no presentaban problemas de degradación, solubilidad o expresión, se utilizó directamente el sitio de clonaje múltiple del vector pET24d y sus etiquetas de purificación (**figura 4.10.a y b**, y **sección 3.1**).



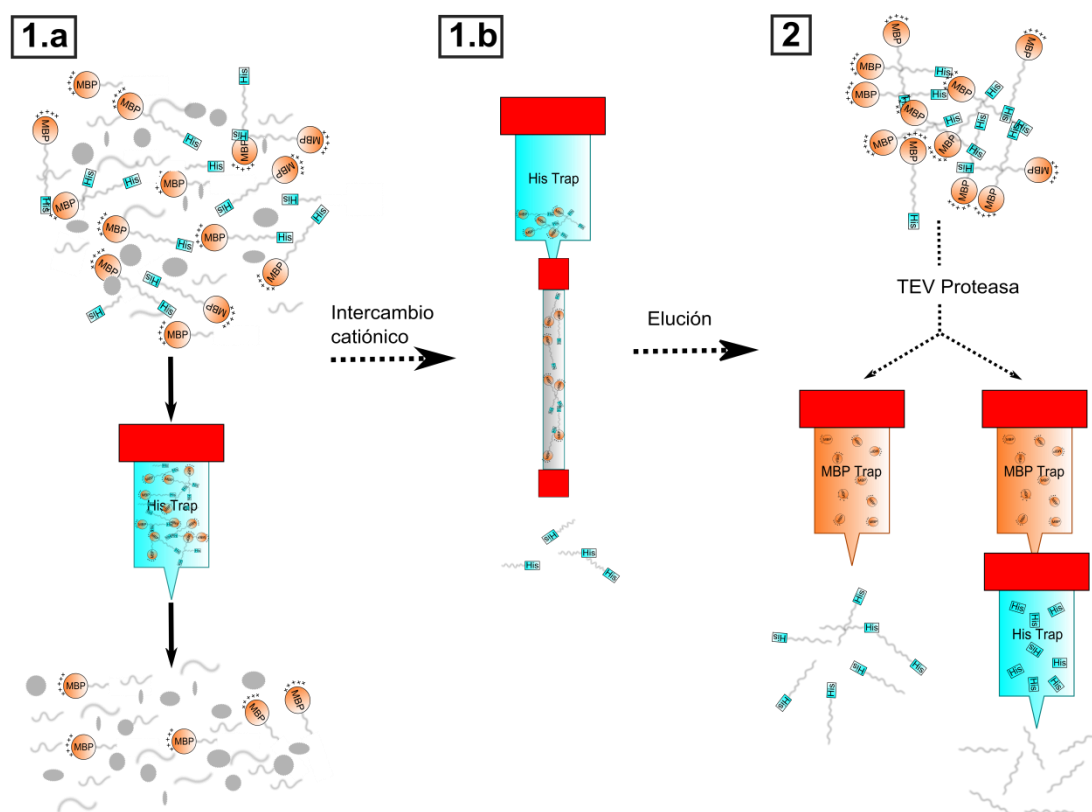
**Figura 4.10.- Detalle del sitio de clonaje múltiple y las etiquetas de purificación de los vectores desarrollados. a)** Detalle de las dianas de restricción del sitio de clonaje múltiple del

vector pET-24d (Novagen). En negrita y cursiva se destacan las secuencias de reconocimiento de las distintas enzimas de restricción. En azul se remarca la etiqueta de poliHistidinas que incorpora el vector para su purificación por cromatografía de afinidad de iones divalentes. **b)** Ejemplo de cómo se ha utilizado la estructura del sitio de clonaje múltiple del vector pET-24d (Novagen) para generar una batería de vectores tipo *cassette*. Su estructura permite sustituir los componentes de manera independiente, por ejemplo podría sustituirse la interacción de interés clonando entre las dianas NcoI-EcoRI las proteínas de interés. **c)** y **d)** vectores poliR\_MBP. Esta familia de vectores incluye, además del sitio de clonaje múltiple y la etiqueta de poliHistidinas del vector pET-24d (Novagen), la etiqueta poliR\_MBP, formada por una cola de poliargininas situadas en el extremo N-terminal de la proteína MBP. Los vectores incluyen una diana de la proteasa TEV (ENLYFQ/A) que permite eliminar la etiqueta poliR\_MBP tras la purificación. En **d)** se representa una variante de los vectores poliR\_MBP con una segunda diana de la proteasa TEV (ENLYFQ/G) que permite eliminar la etiqueta de poliHistidinas tras la purificación.

Sin embargo, para proteínas que presentaban problemas de degradación, solubilidad o expresión, se generó una nueva familia de vectores, los vectores **poliR\_MBP**, que además presentaban otras ventajas. En estos vectores se ha incluido en el esqueleto del vector pET24d una nueva etiqueta de purificación antes del inicio del sitio de clonaje múltiple. Estos vectores mantienen el sitio de clonaje múltiple y las etiquetas de purificación originales de pET24d. La nueva etiqueta, llamada poliR\_MBP, es una etiqueta doble, formada por una cola de poliarginina situada en el extremo N-terminal de la proteína MBP (del inglés *Maltose Binding Protein*), y tiene una función múltiple. Por una parte la proteína MBP garantiza altos niveles de expresión y solubilidad a las proteínas a las que se fusiona (Kapust y Waugh, 1999; Korf et al., 2005), pudiéndose usar además para purificación por cromatografía de afinidad (Riggs, 2000). Por otra parte la cola de poliargininas dota a la proteína de carga positiva, lo que permite su purificación por unión a una matriz de intercambio catiónico (Brewer y Sassenfeld, 1985; Kapust et al., 2001). Tras la etiqueta poliR\_MBP, se incluye una diana para la proteasa TEV (ENLYFQ/A), que permite eliminar la etiqueta tras la purificación (**figura 4.10.c** y **d**). Además se generó un subtipo de este vector que permite eliminar la cola de poliHistidinas de las proteínas una vez purificadas (**figura 4.10.d**), para evitar interferencias con la funcionalización necesaria para los experimentos de AFM-SMFS (ver **sección 3.10**).

Por otra parte, las etiquetas de poliR\_MBP y poliHistidinas permiten el diseño de una estrategia general de purificación de proteínas en 2 pasos (**figura 4.11**). El primer paso de purificación incluye dos cromatografías consecutivas, de afinidad de Ni<sup>2+</sup> y de intercambio catiónico. El lisado celular se hace pasar por una columna de cromatografía de afinidad de Ni<sup>2+</sup> (en azul en la **figura 4.11**), de forma que las proteínas que presentan etiqueta de poliHistidina quedan unidas a esta matriz. El contenido unido a la matriz de afinidad se eluye directamente en una columna de intercambio catiónico (blanco en **figura 4.11**), uniéndose a esta segunda matriz sólo aquellas proteínas que presenten carga positiva (las proteínas con la etiqueta poliR\_MBP tienen alta densidad de carga positiva

gracias a las poliargininas). Tras ser eluidas de la matriz de intercambio catiónico, las etiquetas de purificación son eliminadas de la construcción por digestión con la proteasa TEV. Las etiquetas son finalmente retiradas de la preparación con un último paso de purificación, de afinidad de MBP (columnas marrones en **figura 4.11**) si sólo se ha retirado la etiqueta de poliR\_MBP o doble afinidad de MBP/iones  $\text{Ni}^{2+}$  si además de la etiqueta de poliR\_MBP se ha retirado la de polihistidinas.



**Figura 4.11.- Estrategia de purificación de proteínas de los vectores poliR\_MBP.** Las proteínas producidas en los vectores poliR\_MBP permiten una estrategia de purificación de proteínas en dos fases. La primera fase consiste en dos cromatografías consecutivas. Las proteínas producidas en los vectores poliR\_MBP presentan dos etiquetas de purificación, poliR\_MBP en el extremo N-terminal (en marrón en la figura) y poliHistidinas en el C-terminal (en azul). Gracias a que cada etiqueta se encuentra en un extremo de la construcción, dos cromatografías consecutivas, afinidad de iones  $\text{Ni}^{2+}$  (columnas azules) seguida de intercambio catiónico (columna blanca), permiten seleccionar la población de moléculas con la longitud completa (*full length*). En la segunda fase las proteínas son digeridas con la proteasa TEV para eliminar las etiquetas de purificación no deseadas, que se retiran de la preparación aprovechando la afinidad por maltodextrinas (columnas marrones)/iones  $\text{Ni}^{2+}$  de la proteína MBP y las poliHistidinas respectivamente.

#### 4.2.4.- Testado de los elementos del sistema.

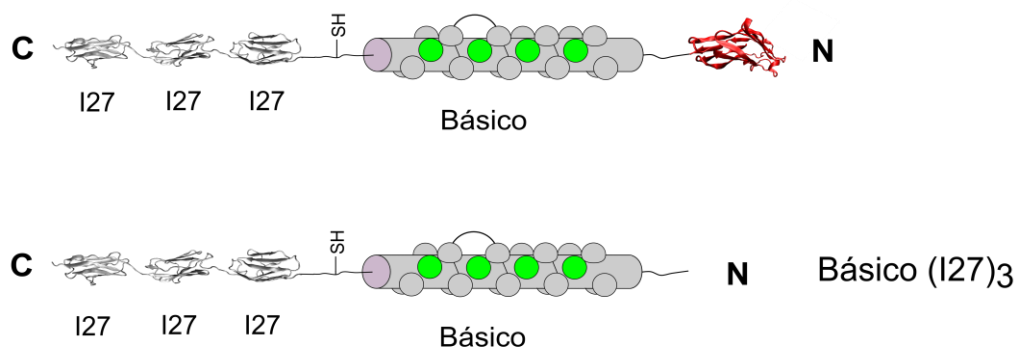
Debido a que la estrategia diseñada es compleja, con varios elementos con funciones específicas, antes de implementar el diseño final se llevaron a cabo una serie de experimentos control para comprobar la funcionalidad de cada uno de ellos. Los objetivos principales de estos experimentos son tres: 1) comprobar que los dominios de oligomerización diseñados son capaces de mediar la formación de enlaces disulfuro en un

sistema supersimplificado; 2) comprobar que el polipéptido utilizado como elastómero se comporta como tal (sin originar picos de fuerza por AFM-SMFS, y que ajusta al WLC), y estimar si la longitud de contorno del mismo es suficiente para que la mayor parte de la fuerza se transmita por la interacción (**figura 1.10**); 3) comprobar el comportamiento conjunto de todos los componentes en un sistema en el que no se incluyan las proteínas interaccionantes de interés (que además supone un control negativo).

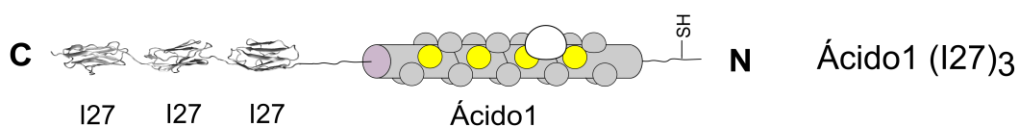
**4.2.4.1.- Los dominios de oligomerización son capaces de guiar la formación de enlaces disulfuro.**

Para testar la orientación y especificidad de los dominios de oligomerización, y su capacidad de mediar la producción de enlaces disulfuro entre cisteínas, se generaron dos construcciones diferentes. Una de ellas, llamada Básico (I27)<sub>3</sub>, tiene la misma estructura que las construcciones Cohesina/Dockerina Básico (I27)<sub>3</sub>, con la diferencia de que la interacción de interés (cohesina/dockerina) ha sido eliminada y sustituida por solo 4 residuos aminoacídicos (GGSG) (ver **figura 4.12.a**). En la segunda construcción, Ácido1 (I27)<sub>3</sub>, el dominio de oligomerización básico de Básico (I27)<sub>3</sub> ha sido reemplazado por el dominio Ácido1 (**figura 4.12.b**). Ambas construcciones incorporan un residuo de cisteína accesible, que quedarían enfrentados solo en el caso de que los dominios de oligomerización interaccionaran de manera antiparalela (**figura 4.12**). El sistema ofrece una manera sencilla de comprobar si los dominios de oligomerización son capaces de guiar la formación de los enlaces disulfuro, y descartar posibles interferencias entre estos y los marcadores de monomolecularidad.

a)

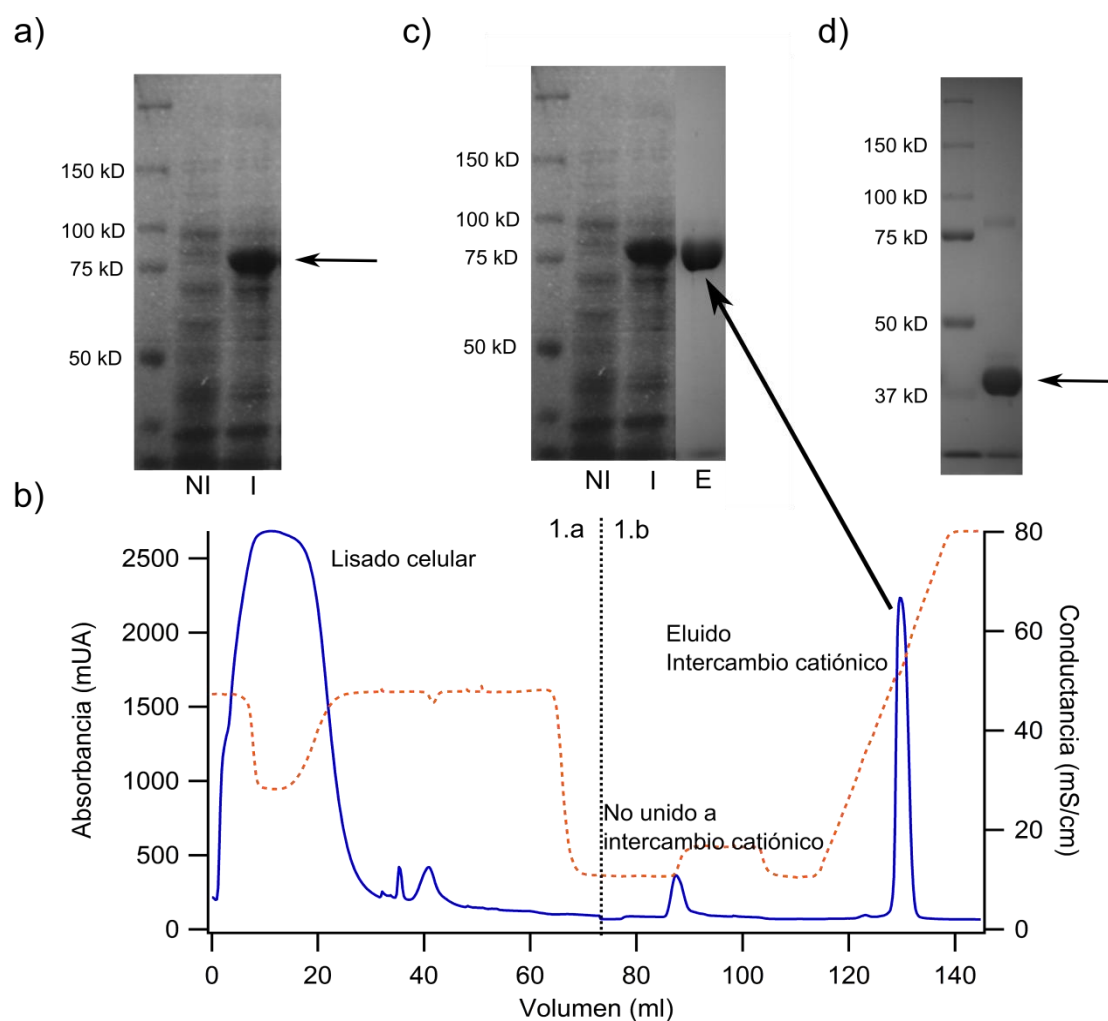


b)



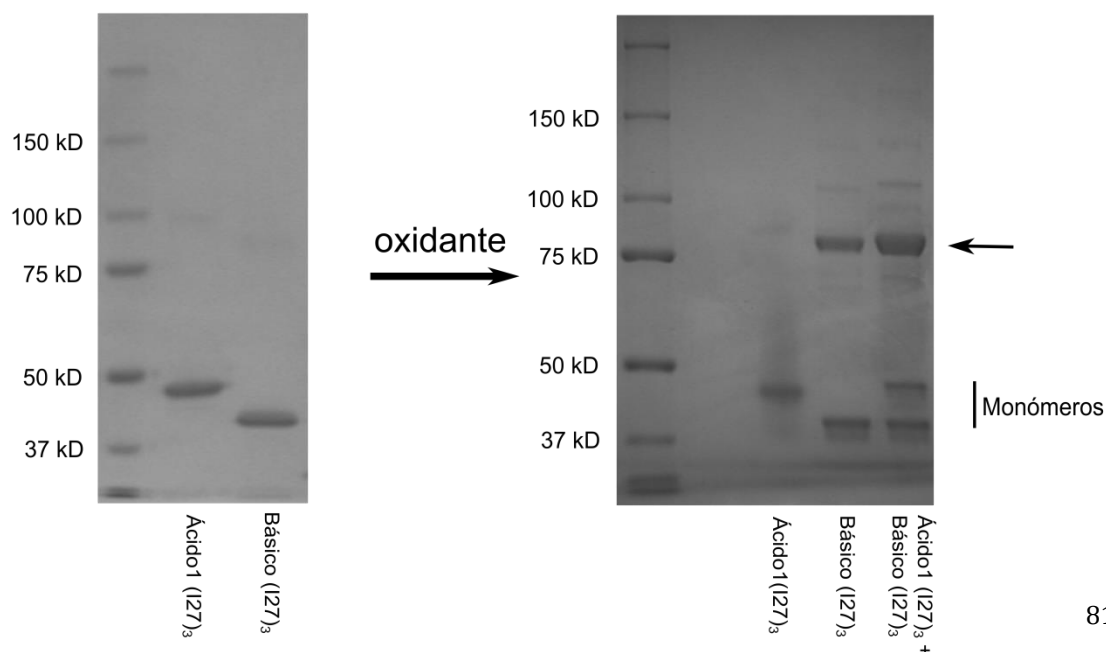
**Figura 4.12.- Construcciones recombinantes utilizadas para testar los dominios de oligomerización.** **a)** construcción Básico (I27)<sub>3</sub> (abajo) que tiene la misma estructura que la construcción Cohesina Básico (I27)<sub>3</sub> (arriba), con la diferencia de que el par de interés ha sido reemplazado por unos pocos residuos "inertes" (GGSG). **b)** construcción Ácido1 (I27)<sub>3</sub>. En Ácido1 (I27)<sub>3</sub>, el dominio de oligomerización básico presente en Básico (I27)<sub>3</sub> ha sido sustituido por el dominio Ácido1.

Las construcciones se clonaron en el sistema de vectores poliR\_MBP, ya que así teníamos la posibilidad de comprobar las ventajas teóricas que ofrecen y la utilidad de la estrategia de purificación combinada que permiten (**figura 4.11**). Tal como es esperable, las proteínas fusionadas a poliR\_MBP se expresan solubles y en gran cantidad (ver expresiones piloto en **figura 4.13.a** y cromatograma de **4.13.b**). Además la estrategia de purificación, que incluye dos cromatografías (de afinidad de Ni<sup>2+</sup> y de intercambio catiónico) en un mismo experimento permite purificar las proteínas unidas a la etiqueta poliR\_MBP con un alto grado de pureza de manera sencilla (**figura 4.13.b** y **4.13.c**). Finalmente la etiqueta poliR\_MBP se elimina por digestión con la proteasa TEV, y la muestra se purifica a homogeneidad (**figura 4.13.d**).



**Figura 4.13.- Expresión y purificación de la construcción Básico (I27)<sub>3</sub>.** **a)** SDS-PAGE de una expresión piloto de la construcción MBP\_Básico (I27)<sub>3</sub>. En el carril en el que se induce la expresión de la construcción (I), se puede observar la aparición de una banda de gran intensidad que corresponde a la construcción MBP\_Básico (I27)<sub>3</sub> (flecha). NI significa no inducido. **b)** Cromatograma de la doble purificación a la que se somete la construcción MBP\_Básico (I27)<sub>3</sub>. En azul y línea continua, absorbancia (en mUA), en marrón y línea discontinua, conductancia (en miliSiemens/cm). A la izquierda de la línea negra discontinua se lleva a cabo la purificación por afinidad de iones Ni<sup>2+</sup> (paso 1.a en **figura 4.11**). Tras la purificación por afinidad se procede a eluir las proteínas retenidas en la columna de afinidad en una columna de intercambio catiónico, para finalmente someterlas a un gradiente de fuerza iónica creciente que provoca la elución de las proteínas unidas a esta segunda matriz (paso 1.b en la **figura 4.11**). **c)** SDS-PAGE en el que se muestra la proteína MBP\_Básico (I27)<sub>3</sub> eluída (E) de la columna de intercambio iónico. **d)** Estado final de la muestra Básico (I27)<sub>3</sub> una vez que ha sido retirada la etiqueta poliR\_MBP.

Aunque la oxidación de cisteínas por el oxígeno molecular atmosférico y la consecuente formación de enlaces disulfuro es una reacción espontánea (Sevier y Kaiser, 2002), *in vitro* la formación de enlaces disulfuro nativos se suele facilitar con el uso de un tampón redox que cataliza la formación de puentes disulfuro vía reacciones de intercambio de tioles (Gough y Lees, 2005). Por su alta solubilidad, precio asequible y por haber sido utilizado previamente para inducir formación de puentes de disulfuro en proteínas (Mukhopadhyay, 2000; Bouvier et al., 2003), se optó por utilizar el sistema cistamina (como oxidante) y cisteamina (como reductor) para generar el tampón redox. De manera estándar, a lo largo de todo el trabajo presentado, se utilizaron diferentes condiciones redox para hallar la condición que maximizaba la formación de puentes disulfuros entre las proteínas (ver **sección 3.3**). Después de dos horas de incubación en el tampón redox, las muestras eran cargadas en un gel SDS-PAGE en condiciones no reductoras. El SDS-PAGE en condiciones no reductoras desnaturaliza las proteínas manteniendo los enlaces disulfuro formados, de tal manera que se puede seguir la aparición de dímeros de proteínas unidos mediante puentes disulfuro por la aparición de bandas de mayor peso molecular que los monómeros (**figura 4.14 y 4.15**).

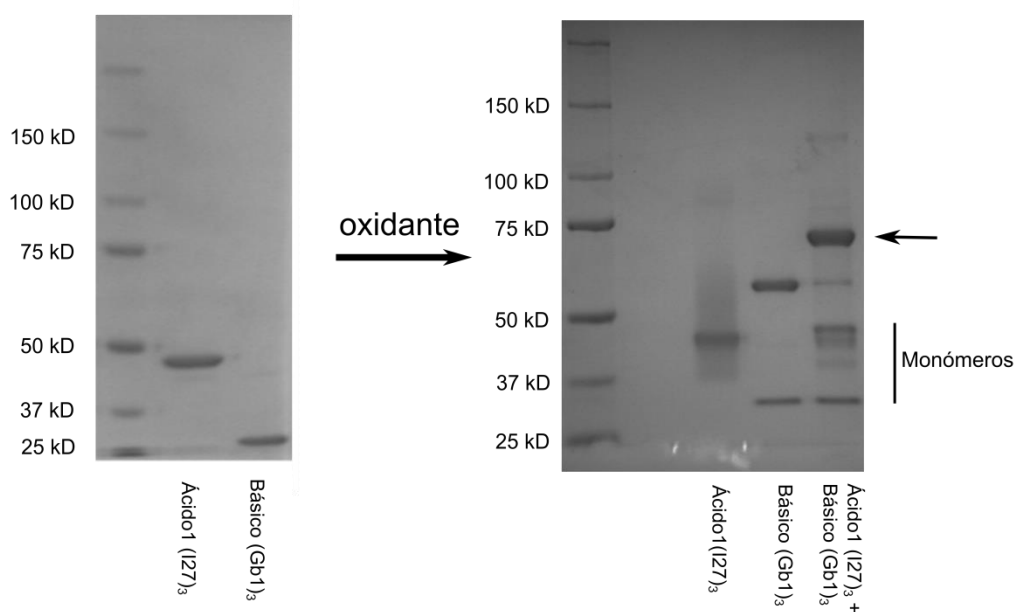




**Figura 4.14.- Formación de enlaces disulfuro entre las construcciones Ácido1 (I27)<sub>3</sub> y Básico (I27)<sub>3</sub>.** A la izquierda se incluye como referencia un gel SDS-PAGE reductor con las muestras Ácido1 (I27)<sub>3</sub> y Básico (I27)<sub>3</sub>, con objeto de poder identificar la movilidad electroforética de las proteínas monoméricas. En la derecha, gel SDS-PAGE no reductor de las muestras (Ácido1 (I27)<sub>3</sub>; Básico (I27)<sub>3</sub>; Ácido1 (I27)<sub>3</sub> + Básico (I27)<sub>3</sub>) tras ser incubadas durante 2 h en un tampón redox oxidante (20 mM Tris, 100 mM NaCl, 1 mM EDTA, **100 μM cistamina**, pH 8.5). La incubación en el tampón oxidante origina la aparición de bandas de mayor peso molecular en alguna de las muestras, que por su peso molecular (aproximadamente la suma del peso molecular de los monómeros) se adscriben a dímeros unidos mediante puentes disulfuro (flecha). Tras la incubación, Ácido1 (I27)<sub>3</sub> se mantiene en forma predominantemente monomérica, mientras que en la muestra Básico (I27)<sub>3</sub> se observa una banda de mayor peso molecular asignada a homodímeros Básico (I27)<sub>3</sub>. La muestra co-incubada Ácido1 (I27)<sub>3</sub> + Básico (I27)<sub>3</sub> presenta una banda mayoritaria con un peso molecular similar a los homodímeros Básico (I27)<sub>3</sub>. Esta banda, de movilidad electroforética ligeramente mayor que la banda homodimérica Básica (I27)<sub>3</sub>, y de mayor intensidad, fue adscrita a heterodímeros Ácido1-Básico (I27)<sub>3</sub>. Los volúmenes de reacción y cantidades de proteínas se mantuvieron constantes entre las muestras en solitario y las muestras co-incubadas.

En la **figura 4.14** se observa el efecto de incubar las muestras Básico (I27)<sub>3</sub> y Ácido1 (I27)<sub>3</sub> en un tampón redox oxidante durante 2 h. La comparación del gel SDS-PAGE reductor (**figura 4.14 izquierda**) con el incubado en el tampón redox (**figura 4.14 derecha**) muestra varias diferencias. La proteína Básico (I27)<sub>3</sub> incubada en el tampón oxidante muestra la aparición de una banda de mayor peso molecular (aproximadamente el doble de tamaño que el monómero, **figura 4.14 derecha**) que puede adscribirse a la formación de homodímeros Básico (I27)<sub>3</sub> unidos por puentes disulfuro. Por el contrario, Ácido1 (I27)<sub>3</sub> se mantiene predominantemente en forma monomérica tras la incubación (**figura 4.14 derecha**), lo que concuerda con la gran desestabilización que sufren los homodímeros 'Velcro' ácidos a pH ligeramente básico (O'Shea et al., 1993). Cuando ambas muestras son incubadas conjuntamente en las condiciones oxidantes, una banda mayoritaria de mayor peso molecular aparece (flecha en **figura 4.14 derecha**). Al comparar con las muestras Ácido1 (I27)<sub>3</sub> y Básico (I27)<sub>3</sub> en solitario, se observa que el porcentaje de monómeros libre es menor en la muestra co-incubada, además la banda de mayor peso molecular es más intensa que la banda de alto peso molecular que aparece en Básico (I27)<sub>3</sub>. Aunque de movilidad electroforética muy similar a los homodímeros Básico (I27)<sub>3</sub> (**figura 4.14 derecha**, carril 2 y 3), la banda de mayor peso molecular de la muestra co-incubada migra como una banda de peso molecular ligeramente mayor, lo que concuerda con que dicha banda represente heterodímeros Ácido1-Básico (I27)<sub>3</sub> (ya que, tal como se observa en los geles, la movilidad electroforética de Ácido1 (I27)<sub>3</sub> es menor que la de Básico (I27)<sub>3</sub>). Tomando todos estos resultados experimentales conjuntamente, podemos inducir que la banda de mayor peso molecular que aparece en las muestras co-incubadas debe estar, al menos en parte, formada por heterodímeros Ácido1-Básico (I27)<sub>3</sub>.

Pese a que la intensidad de la banda de mayor peso molecular que aparece en las muestras Ácido1 (I27)<sub>3</sub> + Básico (I27)<sub>3</sub> es mayor que la de las muestras Básico (I27)<sub>3</sub>, la movilidad electroforética tan similar de ambas bandas hace difícil asegurar que la primera está constituida principalmente por heterodímeros Ácido1-Básico (I27)<sub>3</sub>, y por consiguiente que la formación de heterodímeros está favorecida frente a la de homodímeros. Para soslayar este problema se generó una nueva construcción: Básico (Gb1)<sub>3</sub> en la que las repeticiones de I27 son sustituidas por repeticiones del módulo Gb1, de menor tamaño (Cao et al., 2006) y por consiguiente de mayor movilidad electroforética. En la **figura 4.15** se puede observar que, tras la incubación en el tampón oxidante, la muestra Básico (Gb1)<sub>3</sub> muestra una banda de mayor peso molecular que el monómero, correspondiente por su peso molecular a homodímeros Básico (Gb1)<sub>3</sub>. Esta población de homodímeros Básico (Gb1)<sub>3</sub> es solo vestigial en la muestra Ácido1 (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub>, que presenta una banda mayoritaria (flecha en **figura 4.15 derecha**) de mayor peso molecular que los homodímeros Básico (Gb1)<sub>3</sub>. La banda mayoritaria de las muestras Ácido1 (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub> puede adscribirse sin duda por su peso molecular a heterodímeros Ácido1 (I27)<sub>3</sub>-Básico (Gb1)<sub>3</sub>, lo que apunta claramente a que la formación de heterodímeros está claramente favorecida con respecto a la de homodímeros.

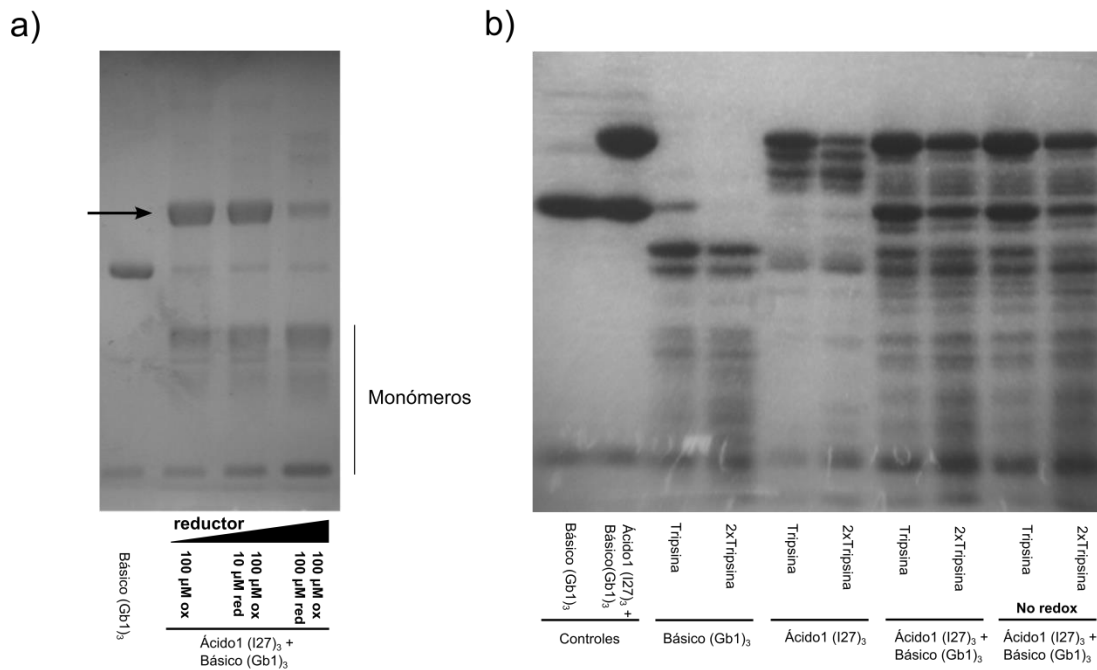


**Figura 4.15.- Formación de enlaces disulfuro entre las construcciones Ácido1 (I27)<sub>3</sub> y Básico (Gb1)<sub>3</sub>.** A la izquierda se incluye como referencia un gel SDS-PAGE reductor con las muestras Ácido1 (I27)<sub>3</sub> y Básico (Gb1)<sub>3</sub>, con objeto de poder identificar la movilidad electroforética de las proteínas monoméricas. A la derecha, gel SDS-PAGE no reductor de las muestras (Ácido1 (I27)<sub>3</sub>; Básico (Gb1)<sub>3</sub>; Ácido1 (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub>) tras ser incubadas durante 2 h en un tampón redox oxidante (20 mM Tris, 100 mM NaCl, 1 mM EDTA, **100  $\mu$ M cistamina**, pH 8.5). La incubación en el tampón oxidante origina la aparición de bandas de mayor peso molecular en alguna de las muestras, que por su peso molecular (aproximadamente la suma del peso molecular de los monómeros) se adscriben a dímeros unidos mediante puentes disulfuro. Ácido1 (I27)<sub>3</sub> se mantiene en forma predominantemente monomérica tras la incubación, y en la muestra Básico (Gb1)<sub>3</sub> se

observa una banda de mayor peso molecular asignada a homodímeros Básico (Gb1)<sub>3</sub>. La muestra co-incubada Ácido1 (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub> presenta una banda mayoritaria de mayor peso molecular que la banda homodimérica Básica (I27)<sub>3</sub>, que representa una población mayoritaria de heterodímeros Ácido1 (I27)<sub>3</sub>-Básico (Gb1)<sub>3</sub>. Los volúmenes de reacción y cantidades de proteínas se mantuvieron constantes entre las muestras en solitario y las muestras co-incubadas.

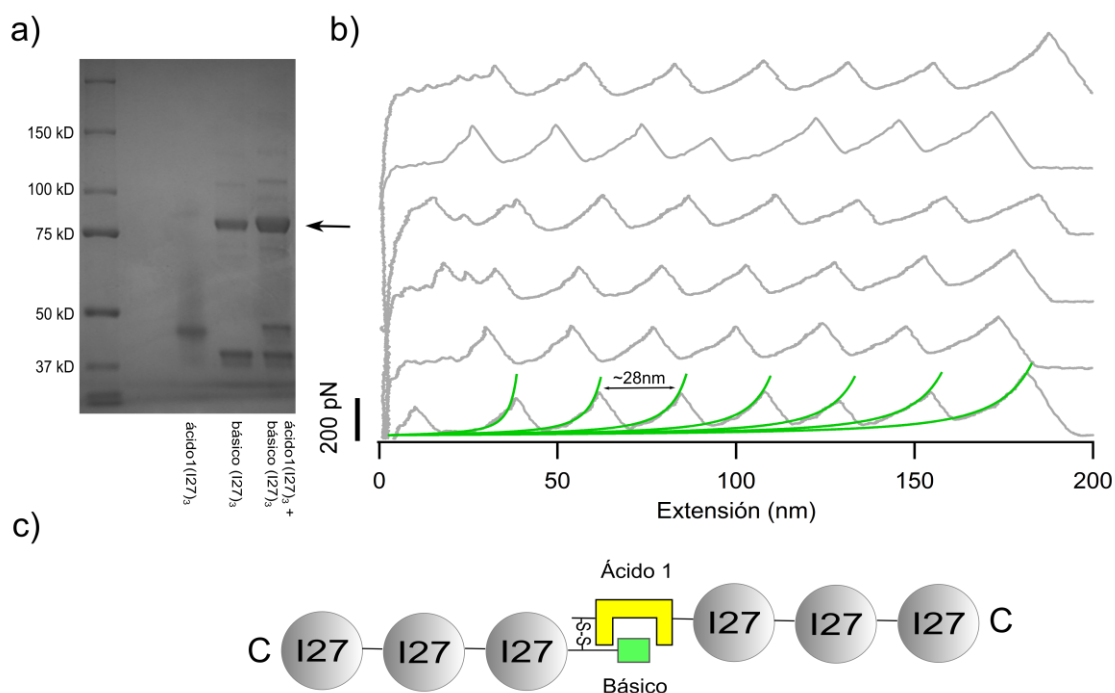
Un hecho relevante que se debe mencionar es que la formación de bandas de alto peso molecular en los geles no reductores está directamente relacionada con el poder oxidante del tampón redox utilizado para la incubación (**figura 4.16.a**). De esta forma, en la **figura 4.16.a**, puede observarse como la banda del heterodímero Ácido (I27)<sub>3</sub>-Básico (Gb1)<sub>3</sub> (flecha en **figura 4.16.a**) va desapareciendo a medida que se incluye un agente reductor en el tampón redox, lo que apoya nuestra interpretación de que la especie heterodimérica que se observa en los geles se mantiene unida gracias a la formación de puentes disulfuro.

Para comprobar que, efectivamente, los dominios de oligomerización heterodimerizan en las construcciones, se llevaron a cabo experimentos de protección frente a digestión por tripsina. Las regiones desestructuradas y accesibles de proteínas son más sensibles a la digestión por proteasas, de ahí que la resistencia a proteólisis de una región se asocia con la existencia de estructura en la misma, o la participación en la unión a un ligando (Huber et al., 1997; Castaño et al., 2002). Los dominios 'Velcro' ácidos y básicos carecen mayoritariamente de estructura definida (forman ovillos estadísticos) cuando están aislados (Oakley y Kim, 1998; O'Shea et al., 1993), de forma que sólo cuando interaccionan en el heterodímero adquieren estructura en hélice  $\alpha$  y forman la estructura en cremallera de leucina. Aprovechando este hecho, se llevaron a cabo ensayos de proteólisis controlada para comprobar si la incubación conjunta de ambas proteínas origina una protección de las mismas frente a proteólisis. En la **figura 4.16.b** se presentan los resultados de la digestión con tripsina. Es evidente, que las muestras tratadas con tripsina en las que Básico (Gb1)<sub>3</sub> y Ácido (I27)<sub>3</sub> están en solitario, muestran una mayor desaparición de la banda correspondiente a la proteína completa comparadas con las muestras co-incubadas (que previsiblemente se encuentran interaccionando). La desaparición de las muestras correlaciona con la cantidad de tripsina utilizada (lo que demuestra que se debe a la degradación por tripsina), y es independiente de la preincubación de las muestras en el tampón oxidante, lo que indica que la protección ante degradación se debe exclusivamente a la interacción entre ambas proteínas. Es necesario notar que la mayor degradación que sufre la muestra Básico (Gb1)<sub>3</sub> en solitario con respecto a la Ácido (I27)<sub>3</sub> en solitario muy probablemente se deba a la gran preferencia de la tripsina por hidrolizar tras residuos de lisina (Olsen et al., 2004).



**Figura 4.16.- Los dominios de oligomerización median la formación de heterodímeros Ácido (I27)<sub>3</sub>-Básico (Gb1)<sub>3</sub> unidos por puentes disulfuro. a)** gel SDS-PAGE no reductor en el que se muestra cómo va desapareciendo el heterodímero Ácido (I27)<sub>3</sub>-Básico (Gb1)<sub>3</sub> (flecha) a medida que aumentamos el poder reductor del tampón redox. Las muestras Ácido (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub> fueron incubadas en 20 mM Tris, 100 mM NaCl, 1mM EDTA pH 8.5 con una concentración de oxidante/reductor variable, de izquierda a derecha: **100 μM cistamina; 100 μM cistamina/10 μM cisteamina; 100 μM cistamina/100 μM cisteamina.** Básico (Gb1)<sub>3</sub> fue incubado en 20 mM Tris, 100 mM NaCl, 1mM EDTA pH 8.5 **100 μM cistamina.** **b)** Ensayo de protección frente a digestión por proteasa de las muestras. El ensayo demuestra que la incubación conjunta de las muestras Ácido (I27)<sub>3</sub> + Básico (Gb1)<sub>3</sub> protege a ambas construcciones frente a la digestión por tripsina, indicando que ambas proteínas se encuentran interaccionando. Ver **sección 3.7** para detalles de la digestión.

Finalmente, las muestras Ácido (I27)<sub>3</sub>+Básico (I27)<sub>3</sub>, Ácido (I27)<sub>3</sub>+Básico (Gb1)<sub>3</sub> se sometieron al protocolo de oxidación (**figura 4.17.a** y **figura 4.18.a**) y se llevó a cabo su análisis mediante AFM-SMFS. En la **figura 4.17.b** se presentan registros fuerza-extensión de Ácido (I27)<sub>3</sub>+Básico (I27)<sub>3</sub>. Los registros presentan el perfil característico de una poliproteína de 6 repeticiones de I27 ( $\Delta L_c \sim 28$  nm,  $F_u \sim 200$  pN, (Carrion-Vazquez et al., 1999)) dispuestas geoméricamente en serie, lo que es esperable si las dos cadenas de Ácido (I27)<sub>3</sub> y Básico (I27)<sub>3</sub> se encuentran unidas mediante puentes disulfuro por las cisteínas incluidas para tal fin (**figura 4.17.c**). Estos resultados refuerzan la conclusión de que los dominios de oligomerización son capaces de orientar las cadenas de acuerdo con el diseño geométrico planeado y facilitar la formación de enlaces disulfuro enfrentando espacialmente las cisteínas. Además, pese a que la eficiencia de adquisición de datos es menor que la obtenida con una poliproteína estándar, estos resultados ponen de manifiesto que no existen interferencias entre los marcadores de AFM y los dominios de oligomerización.



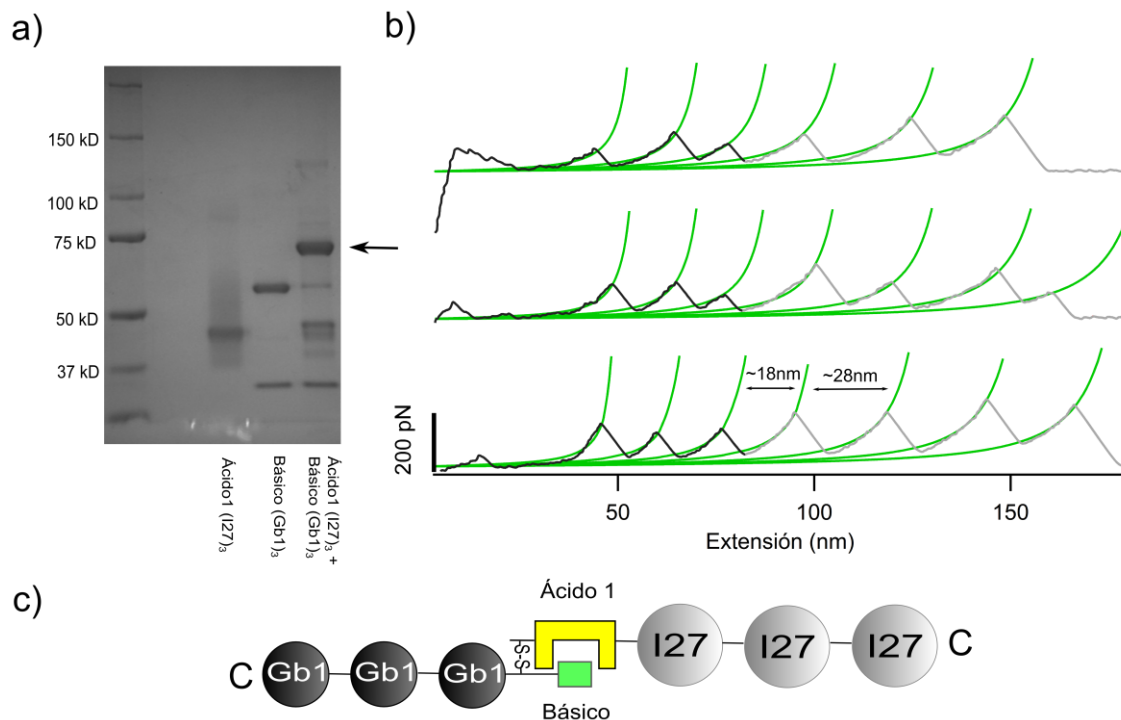
**Figura 4.17.- Análisis mediante AFM-SMFS de los heterodímeros Ácido-Básico (I27)<sub>3</sub>.** **a)** y **b)** Tras su incubación en el tampón redox, las muestras Ácido (I27)<sub>3</sub>+Básico (I27)<sub>3</sub> fueron analizadas mediante AFM-SMFS. En **b)** se presentan algunos registros obtenidos, que presentan el perfil característico de una poliproteína de 6 repeticiones de I27, lo que constituye una evidencia de que los heterodímeros Ácido-Básico (I27)<sub>3</sub> se encuentran unidos mediante puentes disulfuro. La línea continua verde representa los ajustes al modelo WLC. **c)** Pictograma de los heterodímeros Ácido-Básico (I27)<sub>3</sub> unidos mediante enlaces disulfuro.

En la **figura 4.18.b** se presentan registros de AFM-SMFS de Ácido (I27)<sub>3</sub>+Básico (Gb1)<sub>3</sub>. Los espectros fuerza-extensión presentan repeticiones de dos tipos de marcadores, tres picos de fuerza con un  $\Delta Lc$  de  $\sim 28$  nm, y otros tres picos de un módulo más pequeño con  $\Delta Lc$  de  $\sim 18$  nm, lo que ajusta con los patrones de una heteropoliproteína de I27 (Carrion-Vazquez et al., 1999) y Gb1 (Cao et al., 2006). Estos resultados refuerzan nuevamente la conclusión de que los dominios de oligomerización diseñados forman heterodímeros capaces de facilitar la formación de puentes disulfuro.

Aunque en las muestras Ácido (I27)<sub>3</sub>+Básico (I27)<sub>3</sub> y Ácido (I27)<sub>3</sub>+Básico (Gb1)<sub>3</sub>, debe producirse el estiramiento de los dominios de oligomerización debido a la configuración de las construcciones, lo cierto es que en los registros no se aprecian desviaciones al comportamiento de poliproteínas de seis repeticiones de módulos marcadores. El estudio de la estabilidad mecánica de los dominios de oligomerización en las geometrías de estiramiento utilizadas en el diseño final se discuten en profundidad en la **sección 4.2.4.4**.

#### **4.2.4.2.- Las 12 repeticiones de resilina se comportan como un muelle entrópico ideal.**

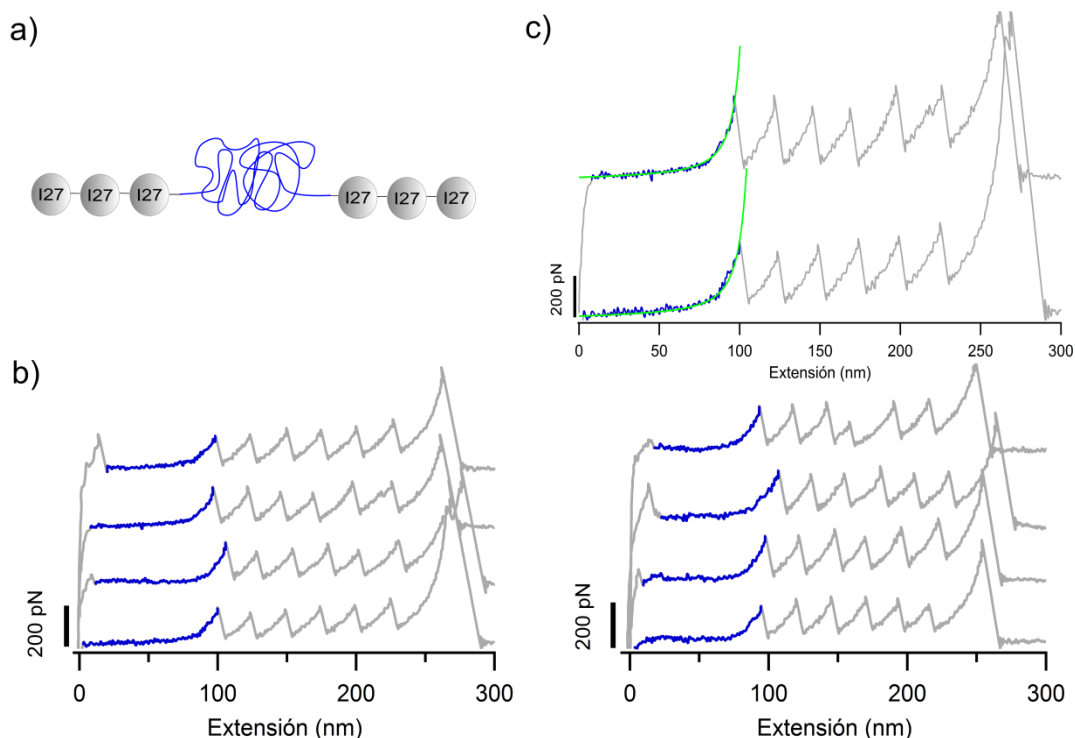
Para comprobar si las 12 repeticiones en serie de resilina (12XR) se comportan como un muelle entrópico ideal, sin mostrar barreras entrálpicas apreciables (en estudios AFM-SMFS), estas fueron clonadas en una poliproteína de módulos I27 para estudiar sus propiedades mecánicas (**figura 4.19.a**). Esta estrategia ha sido utilizada anteriormente para estudiar el comportamiento de polipéptidos desestructurados, en estos estudios la zona inicial de los espectros fuerza-extensión de la poliproteína refleja las propiedades elásticas de la región desestructurada (Li et al., 2001; Li et al., 2002).



**Figura 4.18.- Análisis mediante AFM-SMFS de los heterodímeros Ácido1 (I27)<sub>3</sub>-Básico (Gb1)<sub>3</sub>.** **a)** y **b)** tras ser incubadas en el tampón redox, las muestras Ácido (I27)<sub>3</sub>+Básico (Gb1)<sub>3</sub> fueron analizadas mediante AFM-SMFS. En **b)** se muestran algunos registros obtenidos, que presentan el perfil característico de una heteropoliproteína de 3 repeticiones de I27 (en gris) y otras 3 de Gb1 (en negro), en concordancia con la conclusión de que los heterodímeros Ácido1 (I27)<sub>3</sub>-Básico (Gb1)<sub>3</sub> se encuentran unidos mediante puentes disulfuro. En línea continua verde se muestran los ajustes al modelo WLC. **c)** Pictograma de los heterodímeros Ácido1 (I27)<sub>3</sub>-Básico (Gb1)<sub>3</sub> unidos mediante enlaces disulfuro.

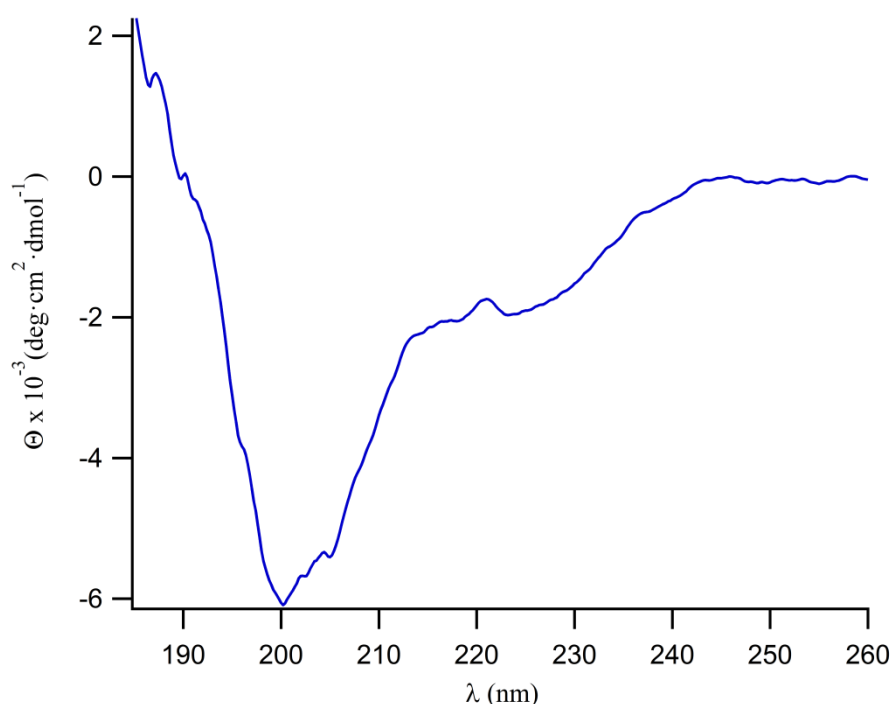
Los espectros fuerza-extensión de la heteropoliproteína (I27)<sub>3</sub>-12XR-(I27)<sub>3</sub> (n=36) muestran una larga región inicial sin picos de fuerza, seguida de 6 picos de desplegamiento de los marcadores I27 (**figura 4.19.b y c**). Teniendo en cuenta que la existencia en el espectro de 6 módulos de I27 implica el estiramiento de la región 12XR, y que la aparición en el espectro es jerárquica (aparecen en primer lugar los eventos de menor estabilidad seguidos por los de mayor estabilidad), los resultados indican que las repeticiones de resilina se despliegan sin mostrar barreras entrálpicas al estiramiento, y con las características típicas de un elemento desestructurado (resistencia entrópica). De hecho la zona inicial del espectro, que recoge principalmente el comportamiento elástico

de las repeticiones 12XR, ajusta perfecta al modelo WLC (**figura 4.19.c**). Así pues las repeticiones 12XR presentan el comportamiento elástico que requerimos para utilizarlo como elastómero en el diseño experimental.



**Figura 4.19.- Propiedades mecánicas de las 12 repeticiones en tándem de resilina. a)** esquema de la heteropoliproteína (I27)<sub>3</sub>-12XR-(I27)<sub>3</sub> utilizada para estudiar las propiedades mecánicas de las 12 repeticiones de resilina. **b)** y **c)** Espectros de fuerza-extensión de la heteropoliproteína (I27)<sub>3</sub>-12XR-(I27)<sub>3</sub>. Los espectros muestran una larga región inicial (azul) sin picos de fuerza en la que se produce el desplegamiento mecánico de las repeticiones 12XR, seguida del desplegamiento de los módulos marcadores de I27 (en gris). En línea verde continua (**c**) se representan ajustes al modelo WLC.

El espectro de dicroísmo circular de la construcción final Ácido1\_12XR\_Ácido2 (**figura 4.20 y 4.9.d**) muestra que la proteína está mayormente desestructurada, con un marcado mínimo cercano a 200 nm característico de los ovillos estadísticos. El espectro muestra un segundo mínimo, de menor intensidad a 222 nm, que probablemente se debe a la aportación de los homodominios ácidos de oligomerización, que presentan una pequeña tendencia a formar estructuras helicoidales (el espectro recuerda mucho a los espectros de DC de homodímeros ácidos presentados en (O'Shea et al., 1993; Oakley y Kim, 1998)). En resumen, los espectros de DC muestran que las repeticiones presentan una configuración desestructurada, lo que está de acuerdo con el comportamiento de muelle entrópico sin barreras entálpicas a la fuerza que se observa por AFM.



**Figura 4.20.- Espectro de DC (25 °C) de la construcción Ácido1\_12XR\_Ácido2.** El espectro de DC (25 °C) de la construcción Ácido1\_12XR\_Ácido2 muestra que la proteína está mayormente desestructurada, con un marcado mínimo cercano a 200 nm característico de ovillo estadístico.

#### 4.2.4.3.- 12 repeticiones de resilina son suficientemente flexibles para actuar de elastómero en el diseño experimental.

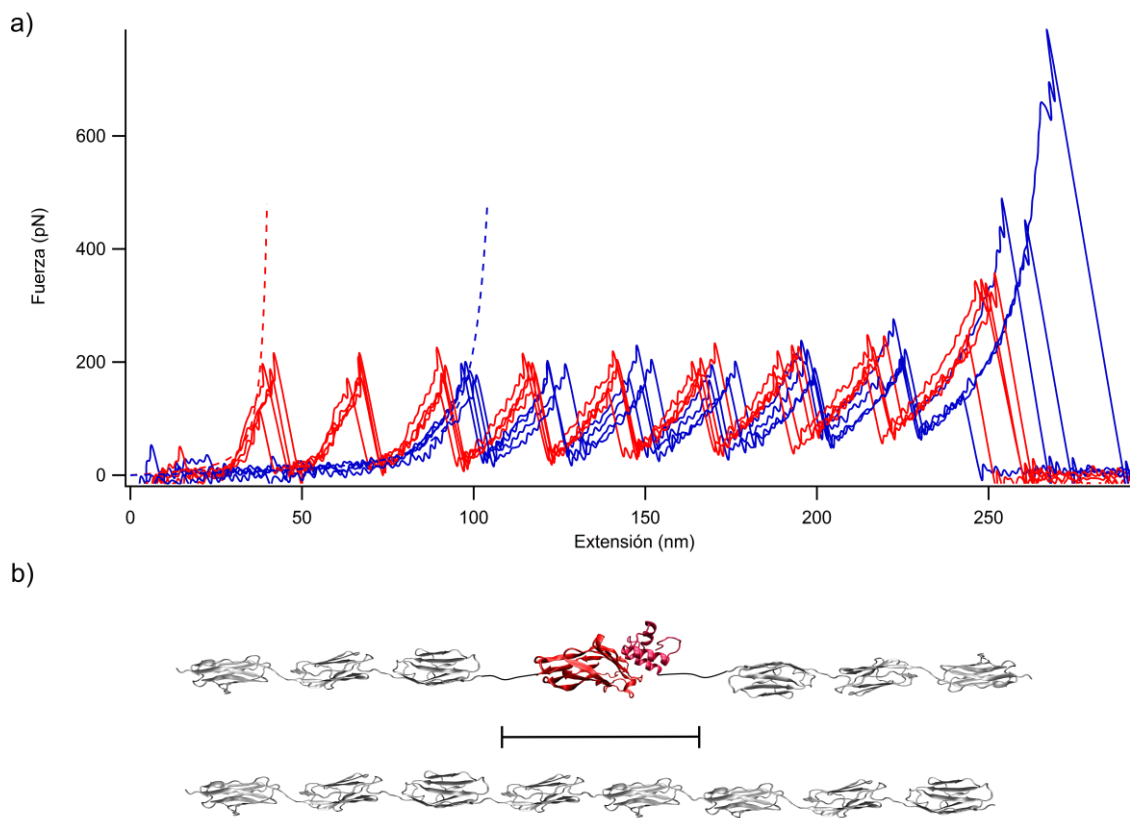
En el diseño experimental final, la interacción a estudio y el elastómero se encuentran dispuestos en paralelo, por lo que la fuerza que se genera al estirarlos solidariamente (que es la fuerza que registraremos en el experimento de AFM-SMFS) se debe a la suma de la fuerza generada al estirar ambos polímeros (Kellermayer et al., 2003). Es decir, la fuerza total es igual a la suma de la fuerza que genera la interacción y la que genera el elastómero al ser estirados,  $F_{\text{total}} = F_{\text{interacción}} + F_{\text{elastómero}}$ . Para estimar, en una primera aproximación, si el elastómero tiene la longitud suficiente para que la interacción soporte la mayor parte de la carga de fuerza durante el estiramiento, podríamos estimar la fuerza total como la suma de  $F_{\text{interacción}}$  y  $F_{\text{elastómero}}$  calculadas a partir del modelo WLC. Esto último requiere conocer los parámetros de longitud de contorno ( $L_c$ ) y longitud de persistencia ( $p$ ) de cada uno (de la interacción y del elastómero).

Como hemos visto, el elastómero se comporta como una proteína desestructurada, así que podríamos utilizar como valor de  $p$  la persistencia de un residuo aminoácido, 0.4 nm (Ainavarapu et al., 2007), y  $L_c = p \times n^{\circ}$  de residuos = 84 nm. Podemos encontrar una validación experimental de estos valores ( $p = 0.4$  nm y  $L_c = 84$  nm) mediante su uso para el cálculo teórico del  $L_c$  inicial ( $L_c^0$ ) de la construcción (I27)<sub>3</sub>-12XR-(I27)<sub>3</sub>, y la posterior



comparación con el  $L_c^0$  de los registros de dicha proteína. El  $L_c^0$  teórico puede calcularse como la suma de los  $L_c$  iniciales de los componentes de la poliproteína (Alegre-Cebollada et al., 2010), utilizando el valor calculado de 84 nm de 12XR, y 30 nm de los 6 x I27 plegados ( $6 \times (\sim 5\text{nm}) = 30\text{ nm}$ ), suman un total de 114 nm de  $L_c^0$ . El  $L_c^0$  de los registro de  $(I27)_3$ -12XR- $(I27)_3$  se estima por ajuste al modelo de WLC del primer pico de fuerza de los registros (ver **sección 4.1**).

En la **figura 4.21.a** se muestra una superposición de registros de  $(I27)_3$ -12XR- $(I27)_3$  (azul), y un ajuste de WLC con los parámetros de  $L_c$  y  $p$  medios provenientes del ajuste del  $L_c^0$  de 36 registros ( $L_c = 113 \pm 12\text{ nm}$ ,  $p = 0.38 \pm 0.02\text{ nm}$ ). La buena concordancia del valor experimental de  $L_c^0$  obtenido con respecto al valor calculado, 113 nm vs 114 nm, justifica el uso de  $p = 0.4\text{ nm}$  y  $L_c = 84\text{ nm}$  como valores para el cálculo de la fuerza generado al estirar el elastómero.

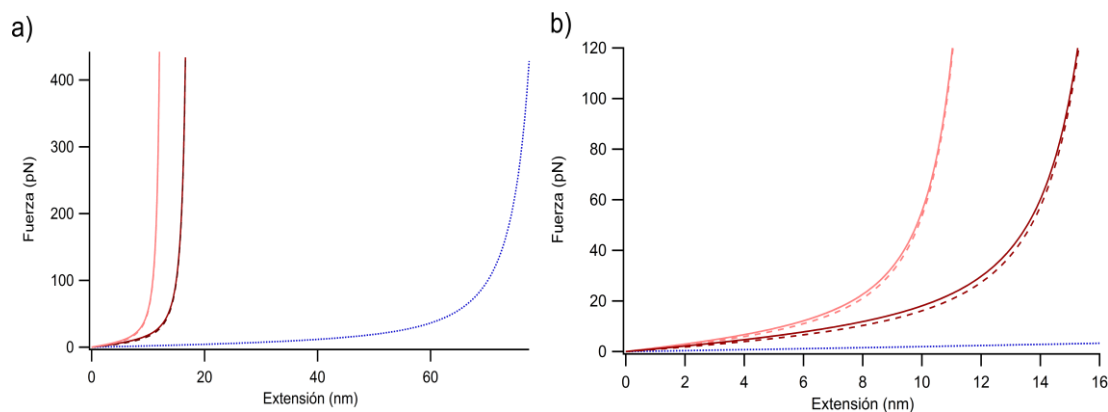


**Figura 4.21.- Flexibilidad del elastómero vs flexibilidad de la interacción. a)** Superposición de 4 registros fuerza-extensión de  $(I27)_8$  (rojo) y 4 registros fuerza-extensión de  $(I27)_3$ -12XR- $(I27)_3$  (azul). En línea discontinua roja  $L_c^0$  generada mediante el modelo WLC con los parámetros  $L_c = 43\text{ nm}$ ,  $p = 0.37\text{ nm}$  (valores medios del ajuste del  $L_c^0$  de 61 registros *full length* de  $(I27)_8$ ). En línea discontinua azul  $L_c^0$  generada mediante el modelo WLC con los parámetros  $L_c = 113\text{ nm}$ ,  $p = 0.38\text{ nm}$  (valores medios del ajuste del  $L_c^0$  de 36 registros *full length* de  $(I27)_3$ -12XR- $(I27)_3$ ). **b)** Similitud entre una poliproteína  $(I27)_8$  (abajo) y la interacción cohesina/dockerina en el diseño experimental (arriba). En gris se representan los módulos I27, en rojo módulo cohesina, y en rosa dockerina. La figura representando la interacción cohesina/dockerina en la construcción experimental no incluye los dominios de oligomerización por simplicidad.

La estimación de los valores de  $L_c$  y  $p$  para la interacción, y su confirmación experimental, es mucho más problemática. Sin embargo basándonos en que cohesina, dockerina e I27 son módulos relativamente compactos, y con tamaños del mismo orden (144, 62 residuos y 89 residuos, respectivamente) (**figura 4.8** y **4.9.d**), la interacción cohesina/dockerina más las 6 repeticiones de I27 que las acompañan en el diseño experimental (tres repeticiones de I27 por molécula, **figura 4.8**, **4.9.d** y **4.21.b**) van a tener a un comportamiento elástico similar a una poliproteína de 8 repeticiones de I27 (**figura 4.21.b**). Esta aproximación supone que dos repeticiones de I27 continuas van a tener, debido a su tamaño similar a cohesina y dockerina, un comportamiento elástico que se acerque aproximadamente al comportamiento de la interacción cohesina/dockerina (**figura 4.21.b**). Así, podemos estimar los valores de  $L_c$  y  $p$  de 2 módulos de I27, y confirmarlos experimentalmente en una poliproteína (I27)<sub>8</sub>. El objeto último de este análisis es validar un valor de  $p$  obtenido experimentalmente en un sistema que va a tener, de manera aproximada, una flexibilidad muy similar a nuestro sistema experimental final (aunque hay que tener en cuenta que los dos módulos I27 tendrán previsiblemente mayor flexibilidad que la interacción cohesina/dockerina debido a los residuos libres entre ambos). Esto nos permitirá finalmente utilizar este valor de  $p$  y un valor de  $L_c$  medido en una estructura atómica resuelta de cohesina/dockerina para hacer la estima.

Los ajustes anteriores de  $L_c^0$  de la poliproteína (I27)<sub>3</sub>-12XR-(I27)<sub>3</sub>, nos indican que un  $L_c$  de  $\sim 5$  nm por I27 plegado y 0.4 nm por residuo de aminoácido que no participe en una estructura es una buena aproximación. De hecho, en la **figura 4.21.a** se muestra una superposición de registros de (I27)<sub>8</sub> (rojo), y un ajuste de WLC con los parámetros de  $L_c$  y  $p$  medios provenientes del ajuste del  $L_c^0$  a 61 curvas *full length* de (I27)<sub>8</sub> ( $L_c = 43 \pm 14$  nm,  $p = 0.37 \pm 0.02$  nm). Teniendo en cuenta un valor de  $\sim 5$  nm por módulo de I27, la poliproteína (I27)<sub>8</sub> debería tener un  $L_c \sim 8 \times 5$  nm = 40 nm, en buena concordancia con el valor experimental. Es llamativo el bajo valor de  $p$  del ajuste, muy similar al de un residuo de aminoácido (0.4 nm), lo que probablemente refleje el significativo aporte de los residuos libres entre módulos de I27 a la flexibilidad de la poliproteína total (el tamaño de estos residuos están incluidos en la estima de 5 nm por I27). Así pues un valor de  $p$  de 0.4 nm parece una buena aproximación.

El cálculo de  $L_c$  de la interacción cohesina/dockerina podría hacerse tomando como referencia la distancia entre los extremos C-terminales de la cohesina y la dockerina (desde donde se estiraría el complejo). Aunque no se tiene estructura cristalográfica del complejo cohesina/dockerina utilizado como modelo aquí (cohesina 7-dockerina de la celulasa Cel8A), si existe estructura del complejo parólogo de *C. thermocellum* cohesina 2-



**Figura 4.22.-Transmisión de fuerzas: interacción vs polipéptido elastomérico. a)** Curvas de WLC generadas para representar el perfil de fuerzas de la interacción, el elastómero y la suma de ambas. Para todas las curvas se utilizó  $p = 0.4$  nm. En azul se representa el elastómero ( $L_c = 84$  nm), en línea discontinua burdeos la interacción con una  $L_c$  estimada de 18 nm, y en línea continua burdeos la suma de ambos valores. En línea discontinua rosa se representa el perfil de fuerzas de la interacción con una estima de  $L_c$  de 13 nm, y en línea continua rosa la suma de estos valores con los del elastómero ( $L_c = 84$  nm). Queda claramente patente que la aportación de la interacción es muy mayoritaria, de hecho la diferencia entre las curvas de interacción y la suma utilizando cualquiera de las dos estimas de  $L_c$  (13 ó 18 nm) es imperceptible en **b)**. **c)** Ampliación de la zona proximal que permite visualizar la diferencia entre las curvas de interacción y la suma. Nuevamente se aprecia que ambas curvas son casi solapantes.

dockerina de xilanas 10B que posee un alto grado de identidad de secuencia (cohesina 75% identidad y 83 % similitud; dockerina 57% identidad y 84 % similitud). Midiendo la distancia entre los extremos C-terminales de la estructura (código pdb 1ohz) podemos utilizar un  $L_c$  de  $\sim 5$  nm para la interacción, lo que sumando los residuos libres provenientes de la estrategia de clonaje suponen un  $L_c$  total de 13 nm. En la figura **4.22.a** y **4.22.b** se representan las curvas de WLC generadas con estos parámetros, para la interacción, el elastómero y la suma de ambas (elastómero en azul, interacción en línea discontinua rosa, suma en línea continua rosa) quedando claramente patente que la aportación de la interacción es muy dominante. Incluso tomando valores de  $L_c$  mayores, como 18 nm, que sería el  $L_c$  si efectivamente la interacción cohesina/dockerina tuviera el mismo comportamiento elástico que dos módulos de I27 ( $18 \text{ nm} = 2 \times (5 \text{ nm por módulo}) + \text{residuos libres de la estrategia de clonaje}$ ) se vuelve a respaldar la conclusión de que la interacción sería el principal camino de transmisión de la fuerza en nuestro diseño experimental. En la **4.22.a** y **4.22.b** (elastómero en azul, interacción en línea discontinua burdeos, suma en línea continua burdeos) se representa la estimación con este valor de  $L_c$ , lo que vuelve a respaldar la conclusión de que la interacción sería el principal camino de transmisión de la fuerza en nuestro diseño experimental, siempre que la diferencia entre el  $L_c$  de la interacción y el elastómero sea grande.

Hay que tener en cuenta que esta estima no pretende ser una aproximación cuantitativa de la fuerza que soportaría la interacción en el diseño experimental

propuesto. De hecho, en el análisis no se incluyen los residuos de aa provenientes de los dominios de oligomerización, ni tampoco se tienen en cuenta consideraciones derivadas del perfil de resistencia a la fuerza de los mismos. El objetivo de este análisis es mostrar cualitativamente como el elastómero (con una  $L_c$  de 84 nm) presenta una diferencia de flexibilidad muy grande con respecto a la interacción (más rígida), y cómo la fuerza se transmite mayoritariamente por el elemento más rígido (la interacción) cuando se estiran en paralelo.

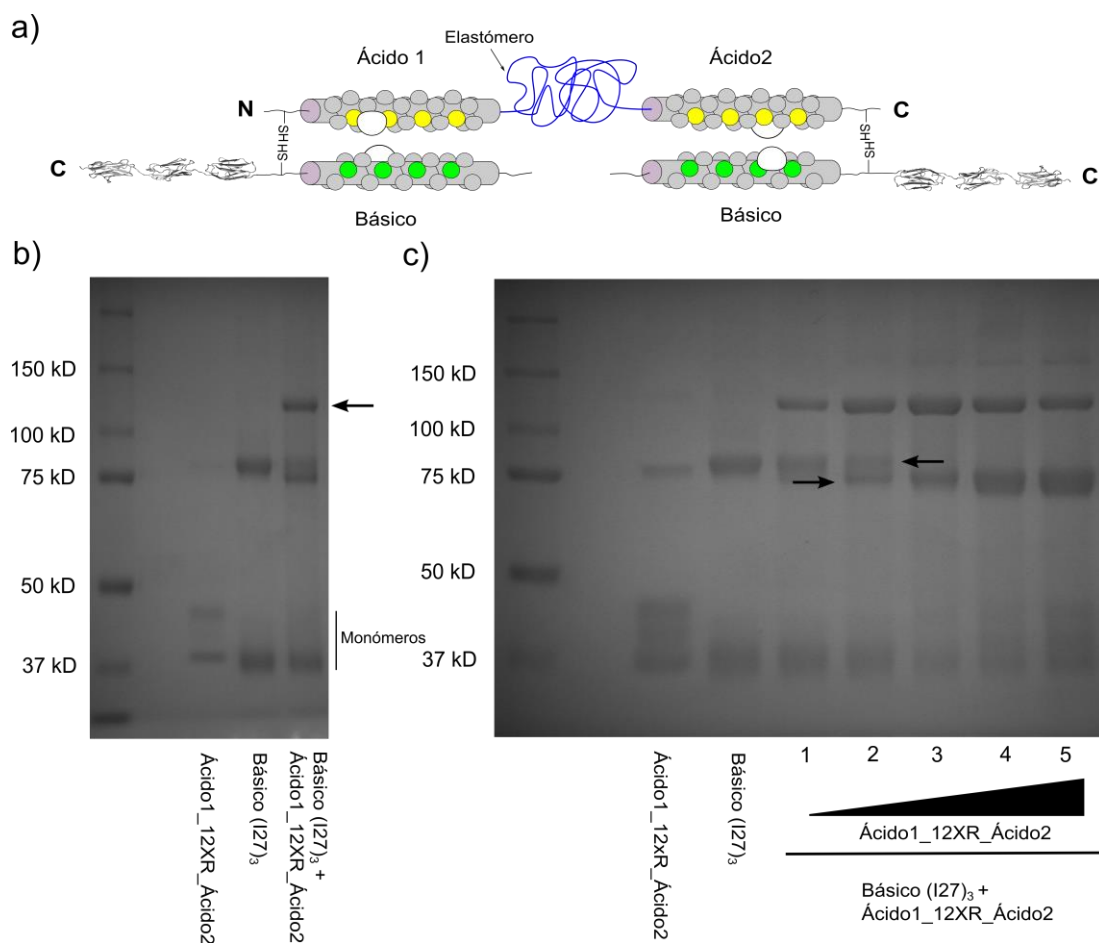
#### **4.2.4.4.- Comportamiento conjunto de todos los componentes del diseño experimental.**

Como paso previo a la implementación de la estrategia experimental diseñada en la interacción modelo cohesina/dockerina, se ideó un modelo “vacío”, sin incluir las proteínas interaccionantes a estudio, que nos permitiera evaluar el comportamiento conjunto de todos los componentes del sistema y comprobar si mantienen su funcionalidad en el diseño final. Además la comparación de los registros de fuerza de la interacción cohesina/dockerina con los de este modelo vacío nos ayudará a identificar los picos de fuerza originados específicamente por la interacción cohesina/dockerina.

El modelo vacío es muy similar al diseño experimental final, con la única diferencia que las construcciones Cohesina Básico (I27)<sub>3</sub> y Dockerina Básico (I27)<sub>3</sub> son sustituidas por la construcción Básico (I27)<sub>3</sub> (**figura 4.9.d, 4.12.a y 4.23.a**). Este modelo mantiene todos los elementos de control de sistema, pero al ser Básico (I27)<sub>3</sub> una construcción en el que las proteínas interaccionantes han sido eliminadas representa también un control negativo.

En la **figura 4.23.b y 4.23.c** se muestran los geles SDS PAGE en condiciones no reductoras después de incubar las muestras Básico (I27)<sub>3</sub> + Ácido1\_12XR\_Ácido2 en un tampón redox oxidante. La incubación de las dos proteínas simultáneamente produce la aparición de tres bandas de mayor peso molecular que los monómeros (flechas en **figura 4.23.b y 4.23.c**). La comparación con el carril de la muestra Básico (I27)<sub>3</sub> incubada en solitario, nos permite identificar entre estas tres bandas, a la banda intermedia como homodímeros Básico (I27)<sub>3</sub>. Las bandas por debajo y por encima de los homodímeros Básico (I27)<sub>3</sub> pueden asignarse por su peso molecular a una molécula de Ácido1\_12XR\_Ácido2 unida a una de Básico (I27)<sub>3</sub>, o a dos moléculas de Básico (I27)<sub>3</sub> respectivamente (flecha en **figura 4.23.b**).

En la **figura 4.23.c** se presenta un gel SDS-PAGE en condiciones no reductoras en el que se estudia el efecto de variar la relación Ácido1\_12XR\_Ácido2/Básico (I27)<sub>3</sub> en las incubaciones. En el experimento se mantiene constante la concentración de Básico (I27)<sub>3</sub> y



**Figura 4.23.- Comportamiento del modelo vacío.** a) Esquema del modelo vacío en la configuración diseñada. El modelo vacío está formado por las construcciones Ácido1\_12XR\_Ácido2 y Básico (I27)<sub>3</sub>. b) SDS-PAGE no reductor de las muestras tras ser incubadas dos horas en tampón redox oxidante (20 mM Tris, 100 mM NaCl, 1 mM EDTA, **100 μM cistamina**, pH 8.5). Tras incubación en estas condiciones aparecen 3 bandas de mayor peso molecular que los monómeros en las muestras Ácido1\_12XR\_Ácido2 y Básico (I27)<sub>3</sub>. Por comparación con las muestras Básico (I27)<sub>3</sub> y por su peso molecular podemos identificar las bandas de arriba hacia abajo como: Ácido1\_12XR\_Ácido2 unida a dos moléculas de Básico (I27)<sub>3</sub> (flecha), homodímeros Básico (I27)<sub>3</sub>, y Ácido1\_12XR\_Ácido2 unida a una molécula de Básico (I27)<sub>3</sub>. c) Efecto del aumento de la relación Ácido1\_12XR\_Ácido2/Básico (I27)<sub>3</sub> en las incubaciones. La flecha que apunta hacia la derecha señala la población Ácido1\_12XR\_Ácido2 unida a una molécula de Básico (I27)<sub>3</sub>, mientras que la que apunta a la izquierda indica los homodímeros Básico (I27)<sub>3</sub>. En las muestras se mantiene la concentración de Básico (I27)<sub>3</sub> constante mientras que la concentración de Ácido1\_12XR\_Ácido2 se va aumentando paulatinamente de izquierda a derecha. A una relación Ácido1\_12XR\_Ácido2/Básico (I27)<sub>3</sub> intermedia entre las utilizadas en los carriles 2 y 3 de las muestras co-incubadas, se consiguen condiciones óptimas que favorecen la formación de la población Ácido1\_12XR\_Ácido2 unida a dos moléculas de Básico (I27)<sub>3</sub>, en detrimento del resto de poblaciones (incluida monómeros libres). Por debajo de esta relación se ven significativamente representados los monómeros Básico (I27)<sub>3</sub> y sus homodímeros, y por encima la población Ácido1\_12XR\_Ácido2 unida a una molécula de Básico (I27)<sub>3</sub> (que va ganando importancia en detrimento de Ácido1\_12XR\_Ácido2 unida a dos moléculas de Básico (I27)<sub>3</sub> a medida que va aumentando la relación). El tampón redox utilizado fue 20 mM Tris, 100 mM NaCl, 1 mM EDTA, **100 μM cistamina**, pH 8.5, el volumen se mantuvo constante en todas las muestras.

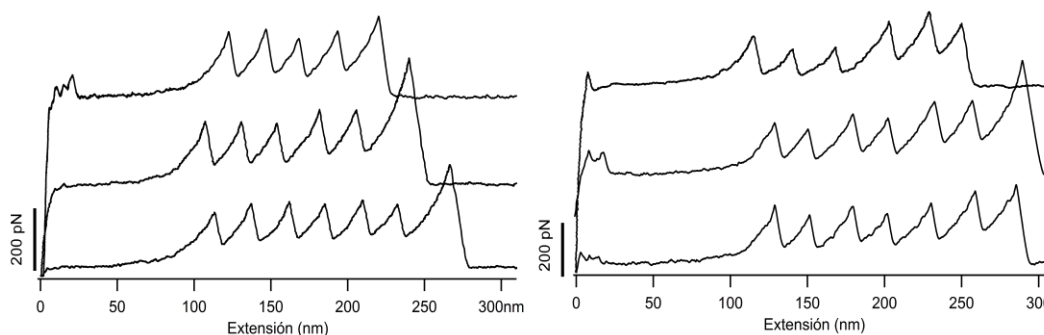
se va aumentando paulatinamente la concentración de Ácido1\_12XR\_Ácido2 (de izquierda a derecha en el gel SDS-PAGE no reductor de la **figura 4.23.c**). A una relación Ácido1\_12XR\_Ácido2/Básico (I27)<sub>3</sub> intermedia entre las utilizadas en los carriles 2 y 3 de

las muestras co-incubadas, se consiguen condiciones óptimas que favorecen la formación de la población Ácido1\_12XR\_Ácido2 unida a dos moléculas de Básico (I27)<sub>3</sub>, en detrimento del resto de poblaciones (incluida la de monómeros libres). Por debajo de esta

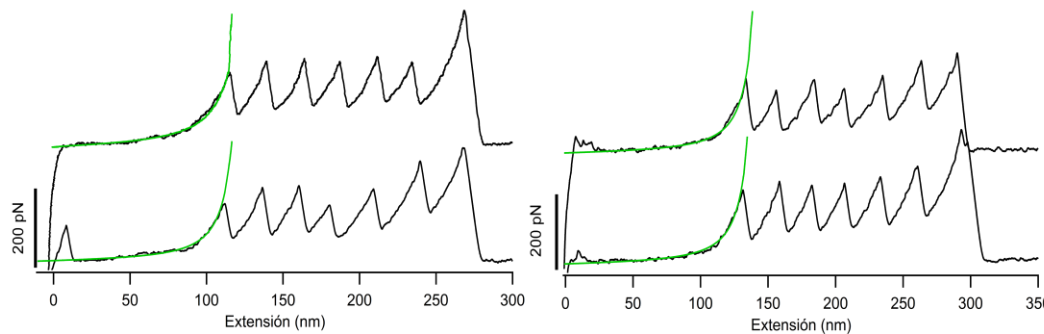
a)



b)



c)



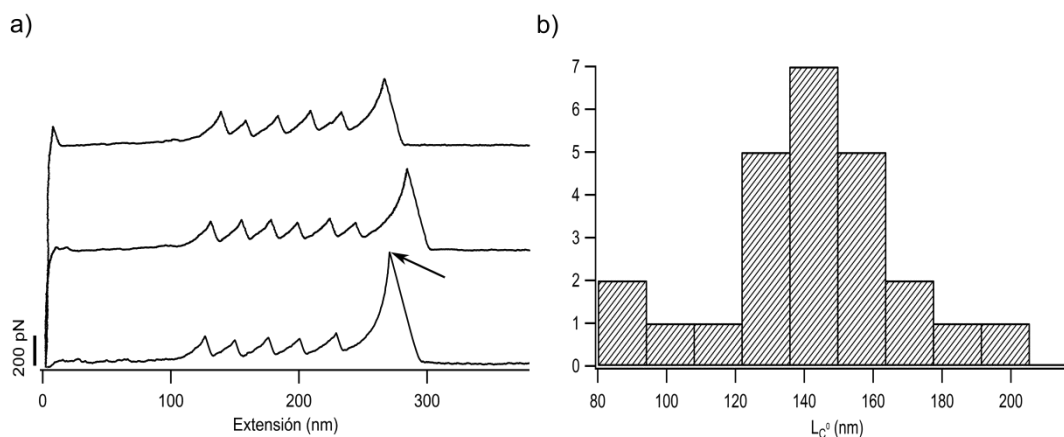
**Figura 4.24.- Nanomecánica del modelo vacío.** a) Pictograma del modelo vacío. Se incluye una indicación de la dirección de la fuerza aplicada en su estudio por AFM. b) y c) registros fuerza-extensión del modelo vacío. Los espectros presentan los dos marcadores esperados, 4-6 picos de fuerza procedentes del desplegamiento de los módulos I27, y una larga región proximal que presenta las características correspondientes al estiramiento del elastómero 12XR (sin barreras entálpicas a la fuerza y que se comporta como un muelle entrópico ajustando al modelo WLC, línea continua verde en c)).

relación los monómeros Básico (I27)<sub>3</sub> y sus homodímeros se ven significativamente representados, y por encima la población Ácido1\_12XR\_Ácido2 unida a una molécula de Básico (I27)<sub>3</sub> (que va ganando importancia en detrimento de Ácido1\_12XR\_Ácido2 unida a dos moléculas de Básico (I27)<sub>3</sub> a medida que va aumentando la relación).

Estos resultados implican que en la construcción final los dominios de oligomerización son capaces de inducir la formación de enlaces disulfuro entre las moléculas en la orientación deseada (figura 4.23.c), y que este comportamiento es además muy próximo al comportamiento ideal ya que la población de moléculas Ácido1\_12XR\_Ácido2 unida a dos moléculas de (Básico (I27)<sub>3</sub>) es muy mayoritaria una vez conseguidas las condiciones óptimas (por ejemplo el carril número 3 en figura 4.23.c). El comportamiento observado al variar la relación Ácido1\_12XR\_Ácido2/Básico (I27)<sub>3</sub> en las

muestras apoya la conclusión del comportamiento cercano al ideal. Las variaciones en las proporciones de las diferentes poblaciones al cambiar esta relación, pueden explicarse asumiendo un comportamiento ideal y utilizando simples argumentos estadísticos. Por ejemplo, el aumento de la relación Ácido1\_12XR\_Ácido2/Básico (I27)<sub>3</sub> una vez sobrepasada la relación óptima, debería provocar, tal como se observa en el gel (**figura 4.23.c**), un aumento en la población Ácido1\_12XR\_Ácido2 unida a una molécula de Básico (I27)<sub>3</sub> debido a que se estaría aumentando su probabilidad de ocurrencia.

Una vez comprobado que las proteínas se comportaban de la manera esperada, se llevó a cabo el análisis nanomecánico de la construcción. El estiramiento por AFM de la construcción desde los extremos C-terminales de las moléculas Básico (I27)<sub>3</sub> (**figura 4.24.a**) debería mostrar dos eventos marcadores en los registros, el estiramiento del elastómero en Ácido1\_12XR\_Ácido2, y el desplegamiento en serie de 4-6 módulos marcadores de I27 (**figura 4.24.b y 4.24.c**). En efecto, en los experimentos se identificaron registros (n= 62) que presentaban el desplegamiento de módulos de I27 (4-6 módulos), con un patrón en dientes de sierra típico de una disposición en serie de los mismos, precedidos por una larga primera región sin picos de fuerza (**figura 4.24.b y 4.24.c**). Estos registros son muy similares a los obtenidos con la poliproteína (I27)<sub>3</sub>-12XR-(I27)<sub>3</sub> (**figura 4.19**). La larga región proximal se comporta como un muelle entrópico, dado que se ajusta muy bien al modelo WLC (**figura 4.24.c**), mostrando las características identificativas del comportamiento elástico de las repeticiones 12XR.



**Figura 4.25.- Los dominios de oligomerización no presentan resistencia mecánica detectable en la configuración estudiada. a)** registros de fuerza-extensión del modelo vacío con elevadas fuerzas de desanclaje. Los valores pueden llegar incluso a más de 600 pN, flecha, sin observarse picos de fuerza que puedan adscribirse a eventos diferentes que el desplegamiento de módulos I27. **b)** Histograma de  $L_c^0$  de 25 registros del modelo vacío con 6 picos de desplegamiento de I27. El histograma presenta un máximo claro en torno a 140 nm, en buena concordancia con la asunción de que los dominios de oligomerización se despliegan sin barreras mecánicas detectables (que originaría un  $L_c^0$  de 142 nm).

Finalmente, debemos tener en cuenta que los dominios de oligomerización son también sometidos a una fuerza paralela a su eje mayor, por lo que los registros también deberían de presentar el perfil de estiramiento característico de los mismos. Los registros de fuerza no muestran ningún patrón de picos de fuerza, a excepción de los módulos marcadores de I27, y el comportamiento entrópico del elastómero 12XR. Esto implica bien que los dominios de oligomerización se despliegan en esta configuración sin mostrar resistencia a la fuerza, o bien con una estabilidad mecánica tan alta que se produce la ruptura del circuito mecánico siempre antes de su desplegamiento, razón por la cual no aparecerían en los registros. Teniendo en cuenta por un lado que dominios de cremallera de leucina muy similares (GCN4) han sido estirados en la misma orientación en (Sarkar et al., 2007) sin reportar ningún tipo de patrón de picos de fuerza, y por otro las altas fuerzas de desanclado que aparecen en alguno de los registros (de hasta 630 pN, ver flecha en **figura 4.25.a**), es más probable que los dominios de oligomerización se desplieguen sin resistencia apreciable a la fuerza. Si aceptamos que los dominios de oligomerización no presentan resistencia a la fuerza, y se comportan como tal formando parte del polipéptido elastomérico, podemos estimar de forma aproximada el  $L_c^0$  de aquellas moléculas que hayan sido estiradas desde los extremos terminales de Básico (I27)<sub>3</sub> como 5 nm x 6 módulos de I27 = 30 nm, más 0.4 nm/aa x 280 residuos (dominios + 12XR) = 142 nm. En la **figura 4.25.b**, se presenta el histograma de  $L_c^0$  estimado en 25 registros con 6 repeticiones de I27. Aunque el número de datos no es suficientemente grande para distinguir perfectamente la forma de la distribución, se observa que aparece un máximo claro en torno a 140 nm siendo el valor medio de la población  $141 \pm 28$  nm, en buena concordancia con el valor calculado de 142 nm. Estos resultados apoyan la idea de que los dominios de oligomerización se despliegan sin resistencia mecánica detectable.

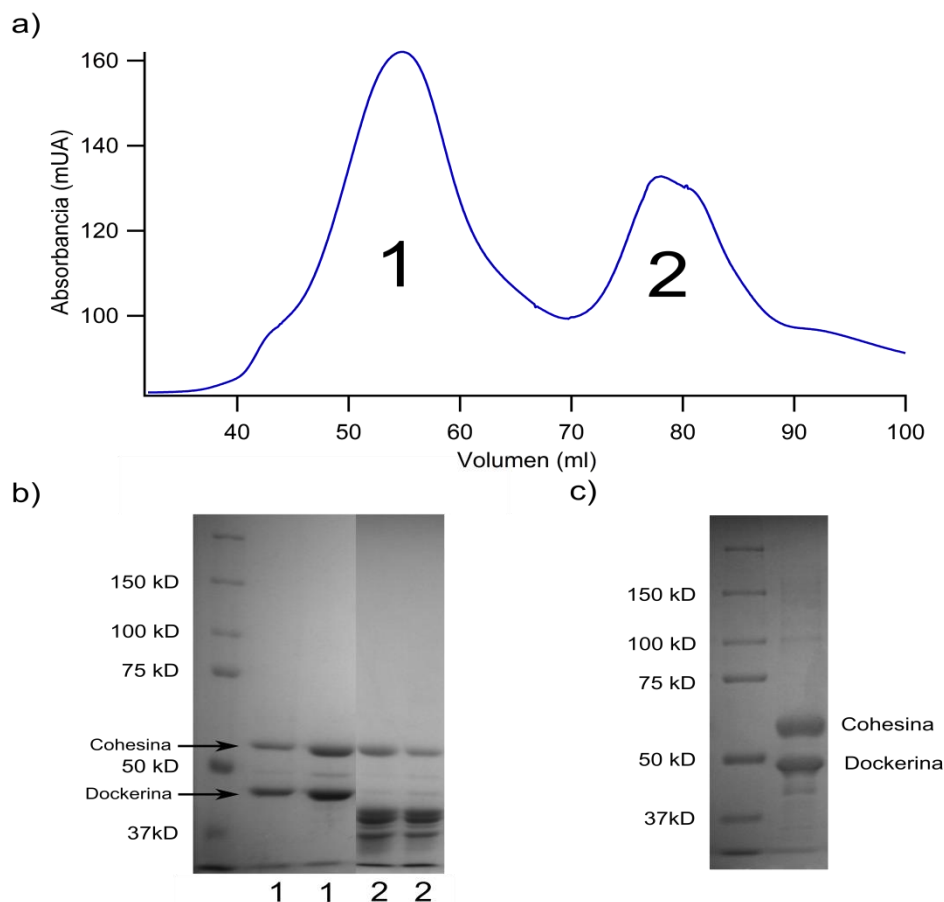
#### **4.2.5.- La interacción cohesina/dockerina origina un pico de fuerza con un $\Delta L_c$ característico.**

Una vez comprobado el correcto funcionamiento de todos los elementos de control del sistema, se llevó a cabo la expresión y purificación de los constructos portadores de los genes cohesina y dockerina (generados sobre el vector pET-24d según la estrategia de *cassette* descrita en la **sección 4.2.3**, y **figura 4.10.b**) y se llevó a cabo su caracterización nanomecánica.

Para asegurar que cohesina y dockerina mantenían la capacidad de interactuar en las construcciones, las proteínas fueron en primer lugar purificadas por afinidad de iones Ni<sup>2+</sup>, incubadas conjuntamente y purificadas finalmente por cromatografía de exclusión molecular. Esta cromatografía muestra dos poblaciones de moléculas

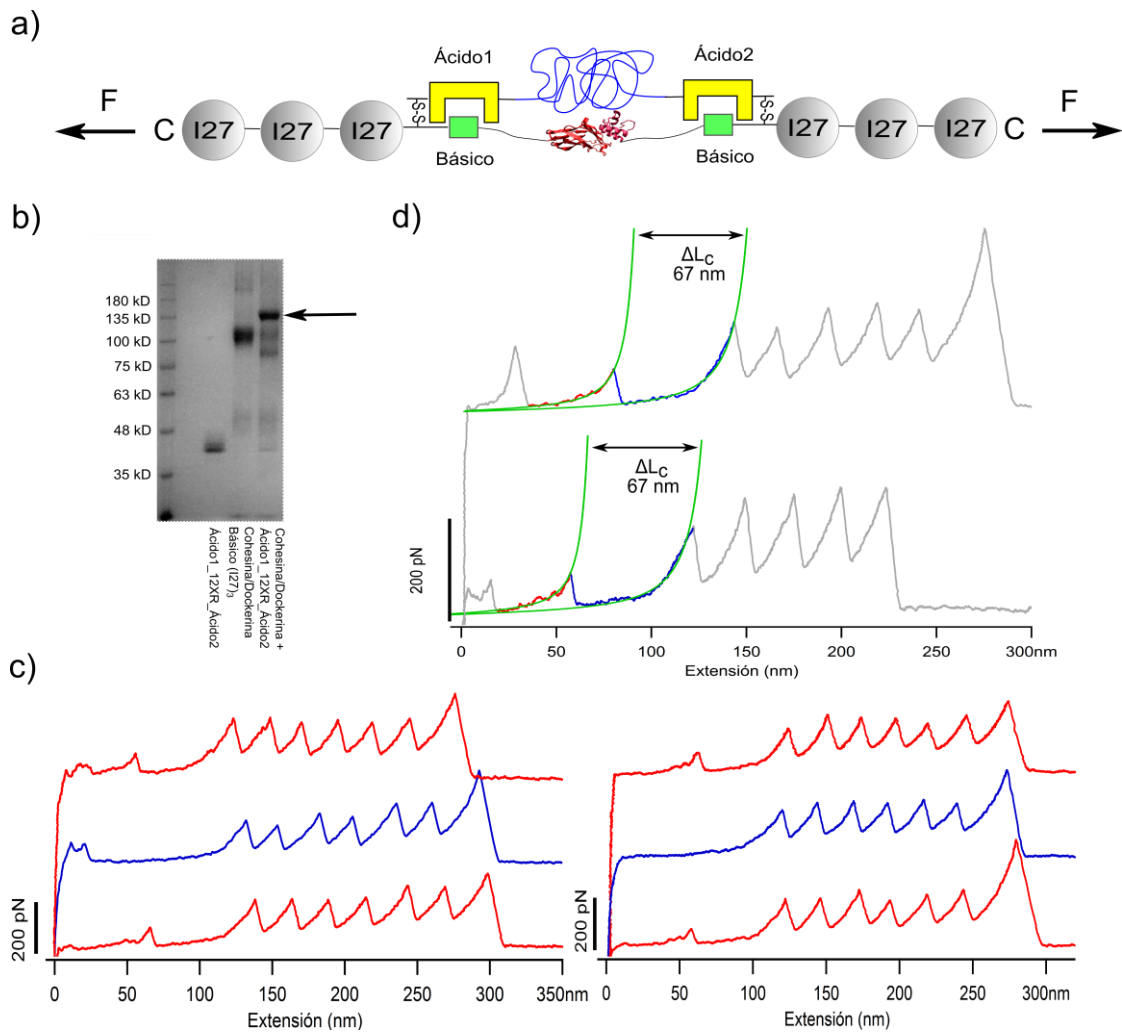


claramente separadas, una constituida por los monómeros de las proteínas y contaminantes de peso molecular similar (pico 2 en cromatograma de **figura 4.26.a** y **4.26.b**), y una población de mayor peso molecular que representa los complejos cohesina/dockerina. La exclusión molecular nos permite pues, separar los complejos cohesina/dockerina de los monómeros y de los contaminantes, purificándose los complejos a homogeneidad (**figura 4.26.c**).



**Figura 4.26.- Purificación de los complejos Cohesina/Dockerina Básico (I27)<sub>3</sub>.** **a)** Cromatograma de purificación por exclusión molecular de los complejos Cohesina/Dockerina Básico (I27)<sub>3</sub>. En el cromatograma se diferencian dos picos (picos 1 y 2). **b)** El pico cromatográfico 1 corresponde a los complejos Cohesina/Dockerina Básico (I27)<sub>3</sub> mientras que el pico 2 recoge los monómeros y contaminantes de peso molecular similar. **c)** La purificación por exclusión molecular permite purificar los complejos Cohesina/Dockerina Básico (I27)<sub>3</sub> a homogeneidad.

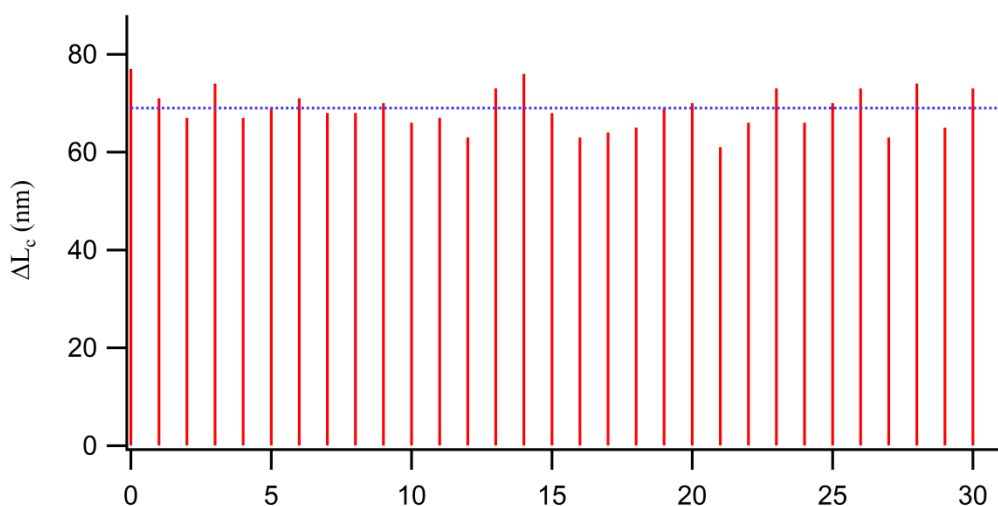
Al igual que el modelo vacío (**sección 4.2.4.3**), las muestras Cohesina/Dockerina Básico (I27)<sub>3</sub> adoptaban mayoritariamente la configuración deseada tras ser incubadas con Ácido1\_12XR\_Ácido2 en el tampón redox (**figura 4.27.a** y flecha en **4.27.b**). Las muestras, incubadas con Ácido1\_12XR\_Ácido2 en el tampón redox, fueron analizadas por AFM-SMFS.



**Figura 4.27.- Implementación final de la estrategia para medir interacciones intermoleculares en la interacción cohesina/dockerina.** **a)** Dibujo esquemático del diseño experimental para medir de manera autocontrolada la interacción modelo cohesina/dockerina. Se incluye un diagrama de la dirección de la fuerza aplicada por AFM. **b)** SDS-PAGE en condiciones no reductoras de las muestras Ácido1\_12XR\_Ácido2 y Cohesina/Dockerina Básico (I27)<sub>3</sub> tras ser incubadas 2 h en un tampón redox oxidante (20 mM Tris, 100 mM NaCl, 1 mM Cl<sub>2</sub>Ca, **1 mM cistamina**, pH 8.5). La población de moléculas Ácido1\_12XR\_Ácido2 unidas mediante puentes disulfuro al complejo Cohesina/Dockerina Básico (I27)<sub>3</sub> (flecha) puede identificarse por su peso molecular y por comparación con la muestra Cohesina/Dockerina Básico (I27)<sub>3</sub> incubada de forma aislada. Al igual que con el modelo vacío, la población de Ácido1\_12XR\_Ácido2 unida al complejo Cohesina/Dockerina Básico (I27)<sub>3</sub> se hace mayoritaria en las condiciones óptimas, las cuales se determinaron experimentalmente (ver **sección 3.3**). **c)** Registros de fuerza-extensión de las muestras Ácido1\_12XR\_Ácido2 + Cohesina/Dockerina Básico (I27)<sub>3</sub> en rojo, en comparación con los obtenidos con el modelo vacío (**sección 4.2.4.4**), en azul. Se identifica un pico de fuerza de menor estabilidad que los marcadores de I27 específico de las muestras Ácido1\_12XR\_Ácido2 + Cohesina/Dockerina Básico (I27)<sub>3</sub>. **d)** El diseño experimental implementado nos permite identificar este pico de fuerza como específico de la ruptura de la interacción cohesina/dockerina. Gracias a la geometría del diseño, a medida que se produce el estiramiento desde los extremos C-terminales de los módulos I27 de la construcción (**a**) la tensión creciente se transmite a través de la interacción cohesina/dockerina. Cuando la tensión sobrepasa la estabilidad mecánica de la interacción se produce la ruptura de la misma (pico en rojo), y el elastómero queda expuesto al circuito principal de fuerza (azul). Finalmente se produce el desplegamiento de los módulos marcadores de I27 (gris). En línea verde continua se muestra el ajuste a la ecuación WLC.

Para cumplir el criterio de monomoleculalidad sólo fueron incluidos en el análisis aquellos registros que presentaban entre 4-6 módulos marcadores de I27. Este criterio

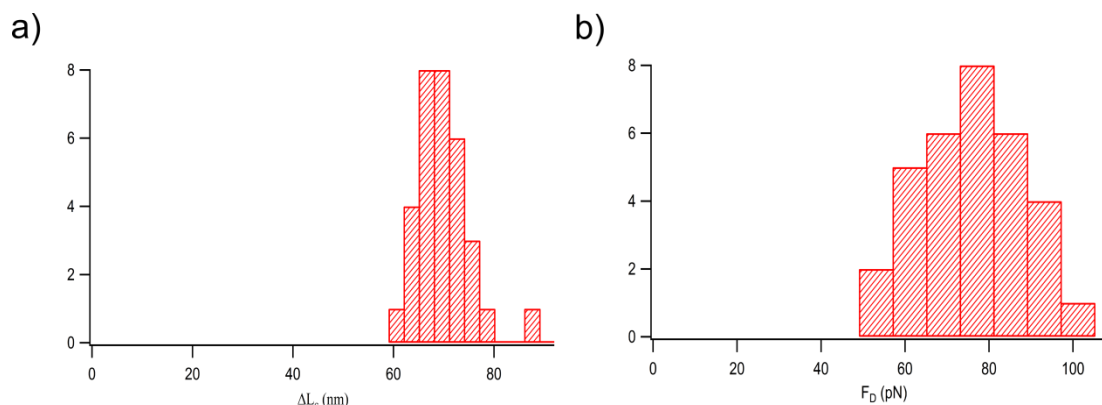
asegura además el estiramiento de la interacción cohesina/dockerina debido a la geometría del diseño. Al igual que en el caso del modelo vacío, los registros analizados presentan la huella digital típica de los 4-6 módulos de I27 en serie, pero a diferencia de aquellos, los registros de cohesina/dockerina muestran un pico de fuerza de menor estabilidad que los módulos marcadores I27 (**figura 4.27.c** y **4.27.d**). En la **figura 4.27.c** se observa este pico de fuerza en los registros (registros en color rojo), y la diferencia con los registros del modelo vacío (registros en azul). Estos picos de fuerza específicos, presentan además un  $\Delta L_c$  (**figura 4.27.d**) con baja variabilidad, tal cómo se muestra en la **figura 4.28**. La diferencia de estos registros con los registros del modelo vacío, y la presencia de los dos marcadores incluidos en el diseño experimental: 1) el marcador de molécula individual representado por las 4-6 repeticiones de I27, y 2) la existencia de pico de fuerza con un  $\Delta L_c$  específico nos hacen identificar este pico de fuerza como proveniente de la ruptura de la interacción cohesina/dockerina.



**Figura 4.28.- Baja variabilidad de los  $\Delta L_c$  del pico de fuerza específico de los registros cohesina/dockerina.** Los  $\Delta L_c$  del pico de fuerza específico de los registros cohesina/dockerina tienen baja variabilidad, oscilando entorno al valor medio de 69 nm (línea azul discontinua).

Aunque el número de moléculas analizadas es pequeño ( $n = 32$ ), debido al bajo rendimiento de la técnica en su primera versión, el histograma de  $\Delta L_c$  presenta un máximo claro en torno a 70 nm, siendo el valor medio de la población  $69 \pm 5$  nm (**figura 4.29.a**). Este valor medio está muy próximo al valor teórico esperado que puede calcularse como la diferencia entre el  $L_c$  de la construcción cuando la cohesina/dockerina se encuentra interaccionando ( $L_c^I$ ) y el  $L_c$  una vez que se ha dissociado la interacción ( $L_c^D$ ).  $L_c^I = 5$  nm x (6 módulos I27) + 5 nm de la interacción cohesina/dockerina (ver **sección 4.2.4.2**) + 0.4 nm x (núm. residuos de los dominios de oligomerización y estrategia clonaje) = 71 nm, y  $L_c^D = L_c^0$  del modelo vacío, que se estimó en la sección anterior como 142 nm. Así pues el  $\Delta L_c$

teórico sería de  $142 \text{ nm} - 71 \text{ nm} = 71 \text{ nm}$ , en buena consonancia con el valor experimental obtenido. En cuanto a la magnitud de fuerzas de disociación, estas también muestran una baja dispersión tal como se observa en el histograma de  $F_D$  (**figura 4.29.b**), con un valor medio de  $75 \pm 12 \text{ pN}$ , que es un valor comparable al valor de  $\sim 100 \text{ pN}$  obtenidos para el par de proteínas parálogas cohesina2/dockerina de celulasa Cel48S a una tasa de carga similar (Stahl et al., 2012).



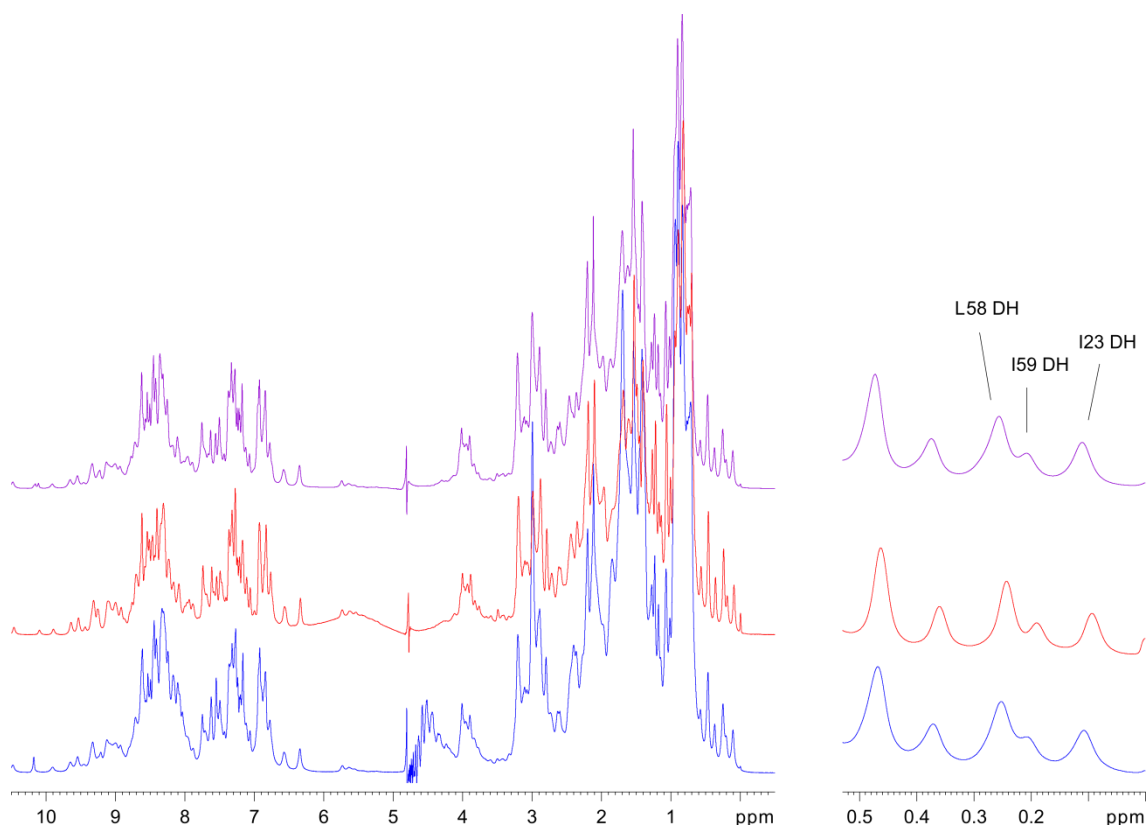
**Figura 4.29.- Histogramas de  $\Delta L_c$  y  $F_D$  de la interacción cohesina/dockerina medida. a)** Histograma de  $\Delta L_c$ . Aunque el número de datos es pequeño ( $n = 32$ ) el histograma de  $\Delta L_c$  muestra un máximo en torno a 70 muy cerca del valor teórico esperado de 71 nm. **b)** Histograma de  $F_D$ . A pesar del reducido número de datos ( $n = 32$ ) el histograma de  $F_D$  muestra poca dispersión, con un máximo en torno a 75 pN.

Un problema que se ha encontrado tanto en esta construcción como en el modelo vacío y las minipoliproteínas Ácido  $(I27)_3$  + Básico  $(I27)_3/(Gb1)_3$ , es la baja eficiencia de adquisición de datos por AFM-SMFS. Aunque los resultados presentados demuestran claramente la viabilidad de esta estrategia general autocontrolada para la medición directa de las interacciones intermoleculares entre proteínas, para que esta sea útil y de aplicación general, sería necesario aumentar la eficiencia de adquisición de datos por AFM-SMFS. En el siguiente apartado, se resumen una serie de experimentos exploratorios llevados a cabo con tal fin, que de momento han tenido escaso éxito.

#### 4.2.6.- Exploración de diferentes vías para aumentar el rendimiento en los experimentos de AFM-SMFS.

Debido a que la adquisición de datos por AFM-SMFS depende directamente de la observación de las señales repetidas de desplegamiento de los módulos de I27, es necesario descartar primero una afectación grave del plegamiento de los módulos marcadores como causa de la baja tasa de adquisición de datos. Para estudiar el plegamiento de los módulos se llevaron a cabo estudios de RMN de las construcciones Básico  $(I27)_3$ , Ácido1  $(I27)_3$  y Básico  $(I27)_3$  + Ácido1  $(I27)_3$  (**figura 4.12**) tras ser sometidas al tratamiento redox (unidas por puentes disulfuro). En la **figura 4.30**, se

presentan espectros unidimensionales de protón de NMR de las tres muestras (Básico (I27)<sub>3</sub> en azul, Ácido1 (I27)<sub>3</sub> en rojo, Básico + Ácido1 (I27)<sub>3</sub> en morado) en los que se observan numerosas señales características del módulo I27 nativo (Improta et al., 1996): por ejemplo K37 HN a 9.87 ppm, G16 y K54 HN a 9.35 ppm, W34 y A82 HN a 9.30 ppm, W34 HZ3 y HH2 a 6.60 ppm, etc. En los paneles ampliados de la misma figura, se presentan algunas señales representativas de grupos metilos empaquetados cerca de residuos aromáticos del núcleo hidrofóbico del módulo I27 plegado. Así pues, los módulos I27 de las construcciones parecen encontrarse correctamente plegados (y no se despliegan después de que se ha llevado a cabo la unión covalente de las construcciones mediante puentes disulfuro). Por tanto, podemos concluir que la afectación del plegamiento no parece ser el origen común de la baja eficiencia en la obtención de registros en las construcciones. Aunque estos resultados no excluyen ciertamente la posibilidad añadida de que en las construcciones del modelo vacío y de cohesina/dockerina el plegamiento de los I27 pudiera estar afectado, el hecho de que los módulos se mantengan plegados en las construcciones Básico (I27)<sub>3</sub>, (I27)<sub>3</sub>-12XR-(I27)<sub>3</sub>, y en poliproteínas heteroméricas de cohesina 7 (Valbuena et al., 2009) hacen poco probable esta posibilidad.



**Figura 4.30.- Espectros 1D 1H de RMN de las construcciones Básico (I27)<sub>3</sub>, Ácido1 (I27)<sub>3</sub> y Básico (I27)<sub>3</sub> + Ácido1 (I27)<sub>3</sub>.** Espectros unidimensionales de protón de RMN de las muestras Básico (I27)<sub>3</sub> en azul, Ácido1 (I27)<sub>3</sub> en rojo, Básico + Ácido1 (I27)<sub>3</sub> en morado a 25 °C, pH 4.5. Se observan numerosas señales características del módulo I27 nativo (Improta et al., 1996), por ejemplo K37 HN a 9.87 ppm, G16 y K54 HN a 9.35 ppm, W34 y A82 HN a 9.30 ppm, W34 HZ3 y HH2

a 6.60 ppm, etc. En los paneles ampliados de la derecha se presentan algunas señales representativas de grupos metilo empaquetados cerca de residuos aromáticos del núcleo hidrofóbico del módulo I27 plegado.

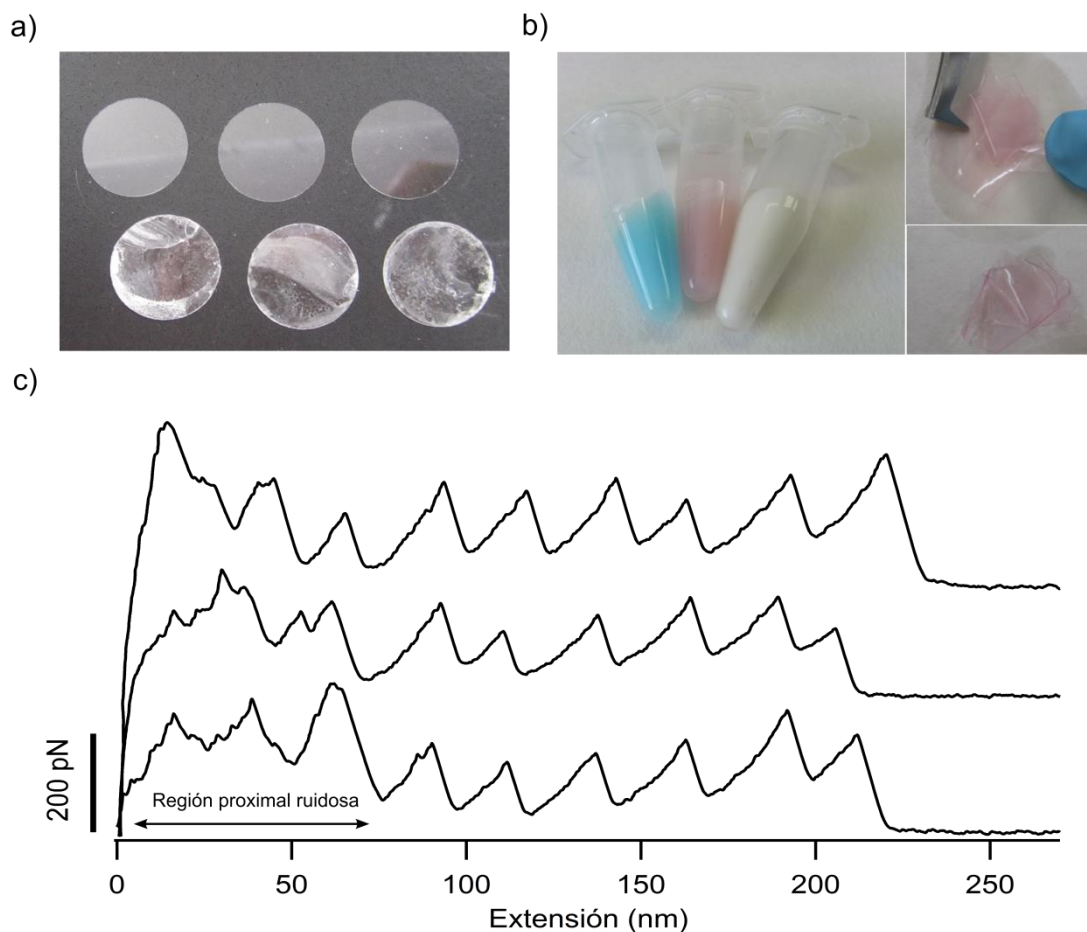
#### **4.2.6.1.- Exploración de diferentes vías para aumentar el rendimiento en los experimentos de AFM-SMFS: funcionalización de sustratos.**

Uno de los factores clave que determina el éxito de un ensayo de AFM-SMFS es la funcionalización del sustrato. Esta debe mantener la proteína adherida a la superficie a la vez que mantiene su actividad. En esta sección se presentan varias estrategias de funcionalización que se han testado. Una vez generados los sustratos, estos eran probados con poliproteínas estándar de I27 ((I27)<sub>8</sub> y (I27)<sub>12</sub>) para evaluar su bondad como sustratos para experimentos de AFM-SMFS.

Los sustratos que utilizamos de manera estándar en el laboratorio son cubreobjetos de vidrio sobre los que depositamos un silano (3-MPTS) que se funcionaliza con un quelante (NTA) de iones Ni<sup>2+</sup> (Hervas et al., 2012) (ver **sección 3.9.2**). El Ni<sup>2+</sup> permite inmovilizar las proteínas en la superficie aprovechando las etiquetas de poliHistidinas que hemos incorporado a las proteínas recombinantes. El protocolo de funcionalización utilizado en el laboratorio (Hervas et al., 2012), producía cubreobjetos con una gran heterogeneidad en el rendimiento en experimentos de AFM-SMFS (con gran cantidad de cubreobjetos que no generaban ningún registro utilizable por experimento), tanto entre diferentes lotes de cubreobjetos, como entre cubreobjetos del mismo lote. Por ello, el primer objetivo fue modificar el protocolo utilizado con el fin de generar cubreobjetos funcionalizados más homogéneos.

Con la excepción de pequeños detalles (ver **sección 3.9.2 y 3.9.3**) la mayor parte de los cambios en el protocolo de funcionalización se centraron en el paso de deposición de silanos, que parecía ser una fuente importante de la heterogeneidad en la producción de los sustratos, ya que los cubreobjetos aparecían macroscópicamente (*de visu*) heterogéneos tras la deposición del mismo (**figura 4.31.a abajo**). Se sustituyó el paso de limpieza de vidrios por un paso más sencillo y robusto (Cras et al., 1999) y se pasó a depositar los silanos mediante deposición en solvente orgánico (Halliwell y Cass, 2001). Estos cambios (y otros cambios menores, ver **sección 3.9.3**), generaban vidrios que eran macroscópicamente más homogéneos (ver **figura 4.31.a** para comparación con el procedimiento anterior), y que también eran mucho más homogéneos en experimentos de AFM-SMFS, y superiores en rendimiento como sustratos. De hecho, este nuevo procedimiento modificado, es el nuevo protocolo estándar de nuestro laboratorio para generar los sustratos funcionalizados para AFM-SMFS. La práctica totalidad de los datos de AFM-SMFS presentados en esta **sección 4.2** fueron adquiridos con estos sustratos. Sin embargo pese a presentar de manera general una mayor eficiencia en la adquisición de

datos (con poliproteínas de I27, homo y heteroméricas), la eficiencia con nuestras construcciones complejas seguía siendo muy baja.



**Figura 4.31.- Funcionalización de sustratos para AFM-SMFS. a)** Comparativa del aspecto macroscópico de los vidrios funcionalizados con 3 MPTS-NTA-Ni<sup>2+</sup> utilizando el nuevo protocolo de funcionalización desarrollado (arriba), y el anterior (abajo). **b)** Funcionalización de celulosa regenerada con quelante de iones divalentes (IDA). El protocolo de funcionalización fue puesto a punto en primer lugar en partículas de celulosa. Izquierda, se muestran tres viales de partículas de celulosa funcionalizadas con Cu<sup>2+</sup> (vial azul), Co<sup>2+</sup> (vial rosa), y sin funcionalizar (blanco). Derecha, ejemplo de las tripas de diálisis funcionalizadas con iones divalentes (en este caso Co<sup>2+</sup>) utilizadas como sustratos de AFM-SMFS. **c)** Registros fuerza-extensión de la poliproteína (I27)<sub>8</sub> obtenidos con los sustratos de celulosa regenerada. Los registros presentaban recurrentemente una larga zona proximal muy ruidosa.

Para intentar disminuir los efectos negativos de la superficie (Ballet et al., 2010; Nakanishi et al., 2001) sobre las construcciones, se decidió usar como sustrato de AFM-SMFS celulosa regenerada. La celulosa regenerada es un material muy utilizado en bioquímica de proteínas (típicamente en filtros de ultrafiltración y tripa de diálisis) por su baja capacidad de unión de proteínas y por tratarse de una superficie hidrofílica. Por su forma plana, se utilizaron tripas de diálisis de celulosa (Spectra/Por Dialysis Membrane 3, SpectrumLabs) que eran modificadas químicamente para incorporar un quelante de iones (IDA, (Hu et al., 2006)). Al incorporar el quelante, la tripa puede cargarse con cualquier ion

divalente como  $\text{Ni}^{2+}$ , (ver **figura 4.31.b**) de tal manera que puede retener por enlace coordinado a las proteínas recombinantes mediante sus colas de poliHistidinas, pero manteniendo intacta su estructura y función gracias a las propiedades de la celulosa. Se produjeron sustratos de celulosa cargados con  $\text{Ni}^{2+}$ ,  $\text{Co}^{2+}$  y  $\text{Cu}^{2+}$ , que se probaron como sustratos de AFM-SMFS. Los registros de poliproteína obtenidos con estos sustratos presentaban un problema recurrente: una larga zona proximal atestada de picos de fuerza inespecíficos (ver **figura 4.31.c**), lo que supone un problema a la hora de detectar picos de fuerza de estabilidad menor que la de los marcadores de I27 (como es el caso de la interacción cohesina/dockerina). Es probable que las altas tasas de funcionalización de la celulosa con el quelante la transformen en una matriz de intercambio iónico fuerte (Wu et al., 2003), quedando las proteínas adheridas fuertemente por carga a su superficie. Si eso es así, la larga región inicial de picos inespecíficos se podría deber al despegamiento de la poliproteína de la superficie. El uso de estos sustratos se abandonó en base a estos problemas, y por el hecho de que al llevarse a cabo experimentos con la construcción de cohesina/dockerina no se obtuvo ni solo registro que cumpliera nuestros requisitos.

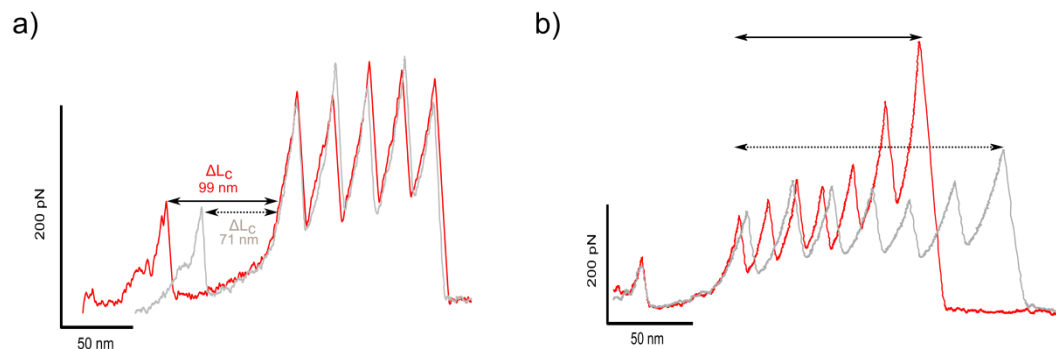
#### **4.2.6.2.- Exploración de diferentes vías para el aumento del rendimiento en experimentos de AFM-SMFS: uso de diferentes marcadores de monomolecularidad.**

Otra línea de trabajo que seguimos para tratar de mejorar la eficiencia del diseño experimental consistió en cambiar las proteínas utilizadas como marcadores de monomolecularidad. Nuevamente, tampoco se consiguió aumentar la eficiencia en la adquisición de datos con esta nueva vía.

El elastómero de repeticiones de resilina fue sustituido por 5 repeticiones del mutante T18P de Gb1. Este mutante se comporta como un polipéptido desestructurado en experimentos de AFM-SMFS, así que parece un buen candidato como elastómero (Cao y Li, 2008). Así pues la construcción Ácido1\_(Gb1T18P)<sub>5</sub>Ácido2, se utilizó como elastómero para estudiar la interacción cohesina/dockerina (Cohesina/Dockerina Básico (I27)<sub>3</sub>). Pese a que estos experimentos presentaban una muy baja eficiencia de adquisición de registros, se obtuvieron 3 registros con los marcadores de monomolecularidad (4-6 módulos de I27). En los espectros de fuerza-extensión se observa un pico de menor estabilidad que los módulos marcadores I27, con una magnitud similar al pico de interacción cohesina/dockerina identificado con el elastómero Ácido1\_(12XR)Ácido2, pero con un  $\Delta L_c$  mayor, entorno a  $\sim 100$  nm (concretamente 96, 99 y 104 nm) (ver **figura 4.32.a**). Pese a que el número de datos es demasiado pequeño, este  $\Delta L_c$  mayor está en concordancia con lo esperado, ya que las 5 repeticiones de Gb1T18P están constituidas por un número mayor de residuos de aminoácidos que las 12 repeticiones de resilina. En concreto posee



88 residuos más, que a 0.4 nm/residuo, suponen  $\sim 35$  nm más de  $\Delta L_c$  con respecto a los 71 nm calculados para las repeticiones 12XR (ver **sección 4.3.5**), es decir un total de 106 nm.



**Figura 4.32.- (Gb1T18P)<sub>5</sub> como elastómero y (Gb1)<sub>3</sub> como marcador de monomolecularidad.** **a)** Comparación de registro obtenido con las muestras Ácido1\_(Gb1T18P)<sub>5</sub>\_Ácido2 + Cohesina/Dockerina Básico (I27)<sub>3</sub> (rojo), con registro obtenido con las muestras Ácido1\_12XR\_Ácido2 + Cohesina/Dockerina Básico (I27)<sub>3</sub> (gris). Ambos registros muestran un pico de fuerza de menor estabilidad que las 4 repeticiones de I27. Este pico de fuerza es muy similar en morfología y magnitud en ambas construcciones, diferenciándose en el  $\Delta L_c$  que presentan, siendo mayor en la construcción que utiliza (Gb1T18P)<sub>5</sub> como elastómero. **b)** Comparación de registro obtenido con las muestras Ácido1\_12XR\_Ácido2 + Cohesina/Dockerina Básico (Gb1)<sub>3</sub> (rojo), con registro obtenido con las muestras Ácido1\_12XR\_Ácido2 + Cohesina/Dockerina Básico (I27)<sub>3</sub> (gris). Ambos registros muestran un pico de fuerza de baja estabilidad muy similar en magnitud y  $\Delta L_c$ . La diferencia entre ambos reside en el  $\Delta L_c$  de los módulos marcadores, cómo se observa en la diferente longitud total de los 6 marcadores (mayor en el caso de los I27, tal cómo es esperable del mayor  $\Delta L_c$  de los módulos I27 con respecto a los marcadores Gb1,  $\sim 28$  nm vs  $\sim 18$  nm respectivamente).

También se llevó a cabo el mismo tipo de experimentos pero manteniendo Ácido1\_(12XR)\_Ácido2 como elastómero y sustituyendo los módulos de I27 en las construcciones de cohesina/dockerina por el módulo silvestre de Gb1 (Cohesina/Dockerina Básico (Gb1)<sub>3</sub>), como marcador de monomolecularidad. Nuevamente, estas construcciones tenían una muy baja eficiencia, pero se consiguió obtener un registro con las repeticiones marcadoras de Gb1 (6 módulos). Cómo se observa en la **figura 4.32.b** el registro presenta un pico de fuerza de una magnitud y  $\Delta L_c$  muy similares a los obtenidos con las construcciones Cohesina/Dockerina Básico (I27)<sub>3</sub> lo que permite identificarlo como la interacción cohesina/dockerina. El mismo  $\Delta L_c$  es esperable, ya que el elastómero utilizado es el mismo. La diferencia entre los registros reside en el menor  $\Delta L_c$  de las repeticiones marcadores, lo que concuerda con el menor tamaño de los marcadores Gb1 (el  $\Delta L_c$  de I27 es  $\sim 28$  nm (Carrion-Vazquez et al., 1999) frente al de Gb1 que es de  $\sim 18$  nm (Cao et al., 2006)).

Aunque estos experimentos fracasaron en nuestro objetivo de aumentar la tasa de adquisición de datos por AFM-SMFS, sirven para validar de forma independiente la estrategia diseñada, basada en el uso de repeticiones de módulos como marcadores de monomolecularidad, y la liberación de un  $\Delta L_c$  específico como marcador del pico de

ruptura de la interacción. El hecho de que se obtuvieran picos identificables de la interacción cohesina/dockerina al utilizar diferentes marcadores de monomolecularidad y diferentes elastómeros, hace improbable que este pico de fuerza se deba a interacciones espurias entre la cohesina/dockerina y los marcadores.

## 5. Discusión

---

### 5.1.- Nanomecánica de la $\beta$ -Catenina

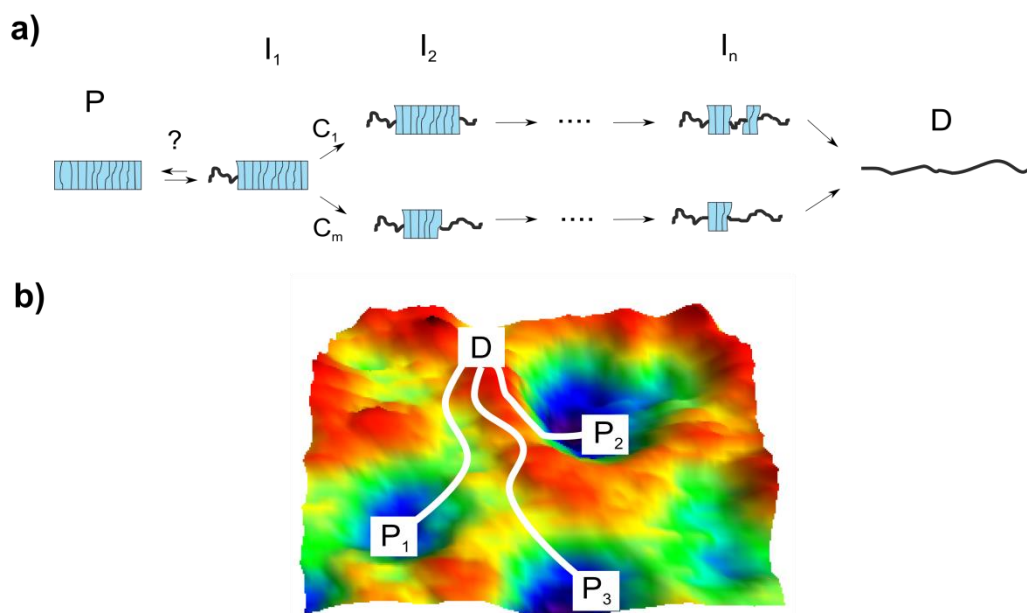
Nuestros resultados muestran que la región ARM de la  $\beta$ -catenina presenta una baja estabilidad mecánica y sugieren que su desnaturalización mecánica ocurre a través del desplegamiento de las repeticiones ARM, tanto a nivel individual como en múltiplos de las mismas. Las propiedades mecánicas de la región ARM de la  $\beta$ -catenina parecen ser muy sensibles a pequeños cambios en la estructura. El desplegamiento mecánico de esta estructura puede ocurrir a través de múltiples caminos de desplegamiento, con diferentes intermedios, lo que indica que el proceso tiene lugar a través de un paisaje energético rugoso con mínimos energéticos muy superficiales (**figura 5.1**).

Es interesante comparar nuestros resultados con los obtenidos para otras proteínas con repeticiones de motivos estructurales  $\alpha$ -hélice naturales, en concreto con aquellas proteínas que contienen repeticiones ANK (anquirina). En D34 (un fragmento de 12 repeticiones ANK) se ha propuesto un paisaje energético muy rugoso, con múltiples mínimos de baja energía que contienen un número variable de módulos plegados. En esta proteína, la desestabilización de repeticiones provoca que un mayor número de repeticiones se desplieguen en una sola transición, debido a que la estabilidad de la repetición es sensible al estado de plegamiento de las vecinas (Werbeck y Itzhaki, 2007). En nuestro caso, los contactos entre las diferentes repeticiones ARM apiladas (ver **sección 1.4.2.1**) ofrecen una explicación al hecho de que una misma fuerza de desplegamiento sea necesaria para desnaturalizar una o varias repeticiones, ya que el desplegamiento de una repetición podría provocar la perturbación y desestabilización de las repeticiones ARM vecinas. En el caso del dominio anquirina de Notch, pequeñas diferencias de estabilidad en las repeticiones ANK provocan grandes cambios en la cooperatividad entre repeticiones en el desplegamiento (Bradley y Barrick, 2002). Esto origina transiciones de desplegamiento de múltiples estados, mecanismo que podría explicar los múltiples caminos de desplegamiento observados en nuestros experimentos. Diferentes estados energéticos en las repeticiones ARM podrían provocar diferentes grados de cooperatividad entre las mismas durante el desplegamiento, originando la variabilidad observada en los espectros de fuerza/distancia y, por extensión, la de los histogramas de  $\Delta L_c$ .

Estudios de AFM-SMFS en la proteína de 7 repeticiones anquirina ganquirina, ponen de manifiesto que el desplegamiento mecánico de las mismas también ocurre por

múltiples caminos de desplegamiento, proponiendo los autores del trabajo que el desplegamiento ocurre a través de un paisaje energético rugoso y complejo. De hecho, en esos experimentos pueden observarse una o dos repeticiones desplegándose simultáneamente, e incluso curvas de fuerza-extensión que no producen picos de fuerza detectables, lo que podría corresponder a ganquirina desplegada previamente al estiramiento. El desplegamiento de una o dos repeticiones en proteínas con repeticiones ANK, ARM y HEAT (proteínas tipo solenoide) también ha sido observado por AFM-SMFS (Kim et al., 2010; Lee et al., 2006). En (Kim et al., 2010) se analizan ciclos de desplegamiento-plegamiento de la  $\beta$ -catenina completa (ARM), aunque las repeticiones presentan una fuerza de desplegamiento comparable a la obtenida en nuestro trabajo, y se observa que en ocasiones dos repeticiones pueden desplegarse simultáneamente, en general las repeticiones parecen desplegarse de una manera más discreta. Hay que tener en cuenta que no se puede hacer una comparación directa de estos resultados en  $\beta$ -catenina con los nuestros, en primer lugar porque esos experimentos se llevaron a cabo sin ningún tipo de control de monomolecularidad, y por otra parte no consiguen estirar más de un tercio de la molécula ( $\sim 100$  nm, sin tener control de qué parte concreta de la molécula se está estirando cada vez). En segundo lugar, y el mayor factor diferencial con nuestro sistema experimental, en (Kim et al., 2010) el desplegamiento de las repeticiones es sólo visible después de re-plegar la molécula tras un gran pico de desorción inespecífico que parece ocurrir concomitantemente a la ruptura de toda la estructura. Teniendo en cuenta que, por todo lo mencionado anteriormente, el desplegamiento de las proteínas con repeticiones de motivos estructurales  $\alpha$ -hélice es muy sensible a las condiciones iniciales y a los estados energéticos de las repeticiones, la comparación de nuestro paradigma experimental con el suyo es muy difícil cuanto no imposible.

La comparación de los resultados de las construcciones ARM/I27,  $\beta$ /I27 y  $\beta$ -cat permite concluir que el aumento de la frecuencia de los picos de alta fuerza se origina exclusivamente por la presencia de los terminales de la  $\beta$ -catenina (0.4 % de los picos de fuerza para ARM/I27 y  $\sim 11$  % para  $\beta$ /I27 y  $\beta$ -cat). Aunque no se puede descartar la posibilidad de que esos picos de fuerza puedan originarse de interacciones de los terminales con el sustrato, el hecho de que los experimentos se llevaron a cabo en sustratos con diferente naturaleza química (ver **sección 3.10**) nos hace pensar que probablemente los picos de alta fuerza puedan deberse a interacciones directas de la región armadillo con los terminales, o bien cómo indican las simulaciones de la región ARM con la  $\alpha$ -hélice en el C-terminal (**ARMHC**), a una estabilización provocada por las regiones terminales en el desplegamiento mecánico de la región ARM.



**Figura 5.1- Representación esquemática del posible escenario para los múltiples caminos de desplegamiento mecánico de la  $\beta$ -catenina. a)** Por simplicidad el modelo se ha restringido a la región armadillo, que se representa como una caja. Basándonos en el análisis de  $L_c^0$  hipotetizamos que algunas repeticiones ARM pueden estar desplegadas o desestabilizadas antes del estiramiento. Dependiendo del estado inicial son posibles múltiples caminos de desplegamiento (**C**) con varios estados intermedios (**I**) que presentan diferentes bloques mecánicos (compuestos de un número diferente de repeticiones que despliegan a la misma fuerza). **D** representa el estado desplegado. **b)** Ilustración que simboliza la rugosidad del paisaje energético de la  $\beta$ -catenina, que contiene varios mínimos con niveles de energía muy similares, y diferentes trayectorias de desplegamiento.  $P_i$  representa los diferentes estados plegados iniciales, mientras que **D** representa el estado desplegado.

En resumen, la región ARM de la  $\beta$ -catenina es muy sensible a pequeños cambios en su estructura y sigue múltiples caminos de desplegamiento ante una fuerza mecánica, que pueden ser modulados por sus terminales desestructurados.

Es interesante notar que la estabilidad mecánica de la  $\beta$ -catenina está en el rango de la determinada para talina (del Rio et al., 2009), una proteína del complejo de unión del sistema integrina-actina de las uniones focales, y lo que resulta todavía más interesante, en el rango de fuerzas que pueden soportar las adhesiones cadherina-cadherina (Baumgartner et al., 2000; Baumgartner et al., 2003; Perret et al., 2004; Zhang et al., 2009), y por debajo de la estabilidad mecánica de la región extracelular de la C-cadherina (Oroz et al., 2011). Teniendo en cuenta que el dominio citosólico de la E-cadherina está sometido a una tensión que oscila entre los 1-3 pN cuando está unida funcionalmente al citoesqueleto de actomiosina (a través de  $\beta$ -catenina (Borghi et al., 2012)), la mecanoestabilidad de la  $\beta$ -catenina que observamos en nuestros experimentos es perfectamente compatible con el papel de la misma como plataforma de conexión mecánica entre el citoesqueleto de actomiosina y las uniones extracelulares cadherina-cadherina. Considerando que la fuerza media de desplegamiento de la  $\beta$ -catenina está en el rango de fuerzas que soportan las

adhesiones cadherina-cadherina (Baumgartner et al., 2000; Baumgartner et al., 2003; Perret et al., 2004; Zhang et al., 2009), y que, cómo muestran los análisis de  $L_c^0$ , parte de las repeticiones pueden desplegarse a baja fuerza, la  $\beta$ -catenina podría actuar como un tampón mecánico a lo largo de todo el rango de fuerzas al que podría estar sometido este sistema (entre 1-3 pN y  $\sim$  50 pN), “sacrificando” algunas de sus repeticiones para mantener los contactos entre las uniones cadherina y el citoesqueleto de actina. Finalmente, cabe señalar que nuestros resultados son compatibles con la pre-existencia de un conjunto de conformaciones diferentes en  $\beta$ -catenina, con niveles energéticos heterogéneos y que podrían tener diferentes afinidades por diversos ligandos, como ya se ha propuesto (Ritco-Vonsovici et al., 2007).

## 5.2.- Estrategia autocontrolada para el estudio de interacciones intermoleculares por AFM-SMFS.

El diseño de un método inequívoco y general de medida directa de la mecanoestabilidad de interacciones intermoleculares entre proteínas por AFM-SMFS, supone un objetivo de enorme envergadura, que supondría un avance metodológico equivalente a la introducción de poliproteínas en nanomecánica de proteínas y de gran transcendencia dada la importancia de las interacciones proteicas en biología. El trabajo presentado aquí, representa un trabajo preliminar en este sentido, pero un paso esencial hacia la consecución de este objetivo ideal. Nuestros resultados muestran que conceptual y prácticamente, es posible diseñar una **estrategia de tipo general** que permita la identificación directa de los picos de ruptura de interacciones entre proteínas en espectros de AFM-SMFS, manteniendo controles de molécula individual.

Se ha de resaltar, que el diseño de esta estrategia no se ideó para medir un sistema en concreto, ni se utilizaron características concretas de la interacción a estudio para implementarlo. Se trata pues de una estrategia de **tipo general**, modular en su diseño, basada en la inclusión en las proteínas de interés (como proteínas de fusión) de poliproteínas como controles de molécula individual y en la utilización de un polipéptido en paralelo a la interacción, que origina la liberación de  $\Delta L_c$  como marcador identificativo del pico de fuerza de la interacción. Por todo ello, la estrategia presenta un gran potencial para el estudio de las interacciones proteína-proteína y tiene un valor heurístico (de hecho se ha procedido al registro de la correspondiente patente).

El éxito de la estrategia experimental propuesta depende de conseguir la unión mediante puentes disulfuro de un polipéptido elastomérico en paralelo a la interacción. Nosotros elegimos como solución a este problema la utilización de dominios de oligomerización que medien la formación de puentes disulfuro entre cisteínas enfrentadas.

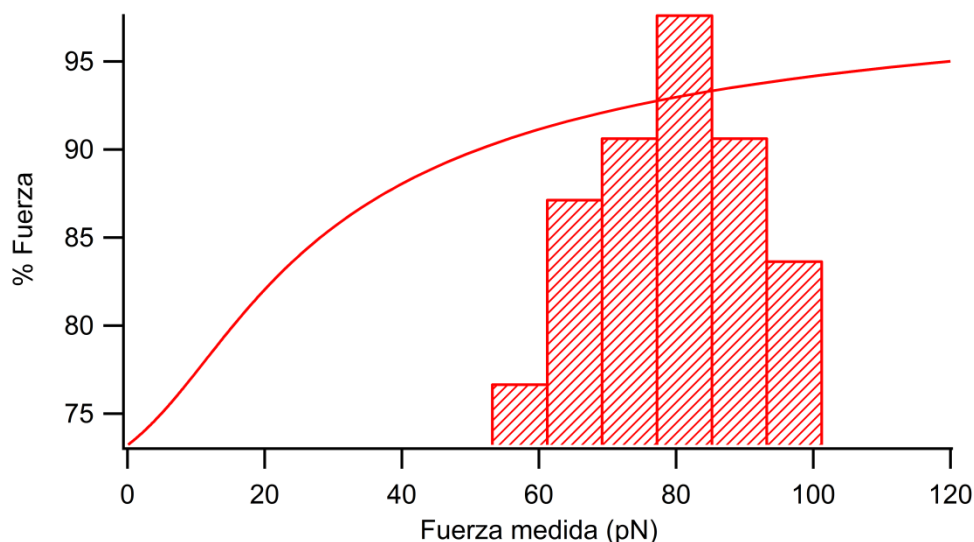
Hay que tener en cuenta que además de la mediación de los puentes disulfuro, estos dominios deben tener la capacidad de orientar las moléculas entre sí de manera controlada, con el fin último de poder controlar la aplicación de estrés mecánico mediante AFM-SMFS. Nuestros resultados (SDS-PAGE en condiciones no reductoras y AFM-SMFS) muestran que los dominios de oligomerización 'Velcro' son capaces de mediar la formación de puentes disulfuro. Nuestros resultados indican también, que los dominios controlan la orientación relativa de las moléculas, ya que las parejas de dominios Ácido1/Básico y Ácido2/Básico han de disponerse de forma antiparalela y paralela respectivamente para que la formación de los enlaces disulfuro sea posible, ya que de otro modo las cisteínas quedarían espacialmente muy separadas para reaccionar. Aunque en el presente trabajo se ha implementado una configuración C-C (es decir, con la aplicación de estrés mecánico desde los C-terminales de ambas proteínas), el diseño general ofrece la posibilidad de implementar las dos configuraciones restantes (N-N y N-C) mediante la utilización de un nuevo dominio de oligomerización Básico (que se diferencia del Básico utilizado aquí en la posición de la asparagina de la superficie hidrofóbica de los dominios).

Con el objetivo de conseguir una mayor estandarización de la estrategia, hemos desarrollado una plataforma de vectores (polyR\_MBP) que nos permite beneficiarnos de los altos niveles de expresión y solubilidad que confiere la MBP a las proteínas a las que se fusiona (Kapust y Waugh, 1999; Korf et al., 2005), y lo que es más importante nos permita implementar una estrategia de purificación general, que sea independiente de las características de las construcciones de interés. En el presente trabajo se han purificado a homogeneidad siguiendo el esquema de purificación diseñado un total de 7 proteínas (ver **materiales y métodos**), lo que claramente indica el potencial de la familia de vectores para conseguir una estandarización en lo referente a la preparación de muestras. Por otra parte, los vectores han sido implementados siguiendo una estrategia modular en *cassette*, lo que nos permitirá en un futuro re-utilizar la misma plataforma de vectores para implementar futuras mejoras o estudiar otras interacciones intermoleculares (por ejemplo sustituyendo los genes de cohesina/dockerina entre las dianas NcoI EcoRI por otro par diferente).

La gran ventaja que presenta el diseño experimental implementado aquí, es que la interpretación de los registros de fuerza-extensión de interacciones intermoleculares se vuelve realmente sencilla, y la identificación del pico de fuerza de la interacción es directa (ver **figura 4.27.c y d**). En el caso de la interacción cohesina/dockerina, el perfil de los registros es igual al esperado para una interacción que tenga una estabilidad mecánica menor que la de los marcadores de I27 (**sección 1.4.2.1**). El pico de fuerza de la interacción cohesina/dockerina aparece primero en el registro fuerza-extensión (rojo en

**figura 4.27.d**), y va acompañado de la liberación del  $\Delta L_c$  correspondiente, debido a que el polipéptido elastomérico queda expuesto al circuito principal de fuerza tras la ruptura de la interacción. A continuación se observa el desplegamiento de los módulos marcadores de monomolecularidad de I27 (4-6 módulos). Esta interpretación está sostenida por varias evidencias experimentales. En primer lugar el modelo vacío (**sección 4.2.4.4**), con la misma arquitectura general que la construcción Cohesina/Dockerina pero en el que las proteínas interactuantes han sido sustituidas por unos pocos residuos “inertes” (entre los que no se espera interacción), no genera pico de fuerza alguno observable en los registros. Esto implica que la aparición del pico de fuerza de baja estabilidad está directamente asociada con la inclusión en los constructos de las proteínas cohesina/dockerina. Si interpretamos los experimentos de AFM-SMFS como que la ruptura de la interacción cohesina/dockerina provoca la liberación del polipéptido elastomérico, podemos estimar el  $\Delta L_c$  que ha de acompañar la ruptura de la interacción como 71 nm (ver **sección 4.2.5**). El valor medio del  $\Delta L_c$  del pico de fuerza de baja estabilidad encontrado es de  $69 \pm 5$  nm, en buena concordancia con el valor teórico calculado suponiendo que este  $\Delta L_c$  proviene de la liberación del polipéptido elastomérico tras la ruptura de la interacción cohesina/dockerina. Los histogramas de  $\Delta L_c$  y  $F_D$  (**figura 4.29**) de este pico de fuerza muestran una baja dispersión, con máximos claros en ambos casos ( $\sim 70$  nm, y  $\sim 75$  pN), lo que hace altamente improbable que este pico de fuerza de deba a interacciones inespecíficas de la cohesina/dockerina con otros elementos de la construcción o con el sustrato. En este sentido, los resultados obtenidos con el elastómero Ácido1\_(Gb1T18P)<sub>5</sub>\_Ácido2 y con la construcción Cohesina/Dockerina Básico (Gb1)<sub>3</sub>, constituyen un control adecuado que apoya la idea de que los picos de fuerza no se deben a interacciones inespecíficas con los componentes de las construcciones, ya que en estos casos son diferentes. Además el elastómero Ácido1\_(Gb1T18P)<sub>5</sub>\_Ácido2 provoca la liberación de un  $\Delta L_c$  mayor tras el pico putativo de la interacción, lo que es esperable si se debe a la liberación del polipéptido elastomérico tras la ruptura de la interacción cohesina/dockerina, ya que el elastómero utilizado tiene una mayor  $L_c$ . Todas estas evidencias experimentales nos hacen concluir que efectivamente el pico de fuerza de baja estabilidad obtenido en los registros se debe a la ruptura de la interacción cohesina/dockerina.





**Figura 5.2- Porcentaje teórico de fuerza que soporta la interacción cohesina/dockerina en el momento de la ruptura en nuestras condiciones experimentales.** En línea continua roja el porcentaje de fuerza teórico que se estima que está realmente soportando la interacción cohesina/dockerina en función de la fuerza medida. Se incluye el histograma de  $F_D$  medido para la interacción cohesina/dockerina, donde se observa que el porcentaje es siempre mayor al 90 %. El cálculo se ha realizado repitiendo el análisis presentado en la **sección 4.2.4.3** con un  $L_c$  de la interacción de 41 nm, y de 112 nm para el elastómero.

Para el análisis de los datos de estabilidad mecánica de la interacción cohesina/dockerina, hay que tener en cuenta el hecho de que aunque la mayor parte de la fuerza registrada en el momento de ruptura está siendo soportado por la interacción, una fracción de la misma está siendo soportada por el polipéptido elastomérico. Esta diferencia entre la fuerza medida y la fuerza que realmente está soportando la interacción en el momento de la media, puede ser relevante, sobre todo en medidas de *Force clamp* ya que la vida media de la interacción depende exponencialmente de la fuerza (ver **sección 1.3.3**). Podemos estimar el porcentaje de fuerza soportado por la interacción en función de la fuerza medida, repitiendo el análisis de la **sección 4.2.4.3** (teniendo en cuenta ahora que los dominios de oligomerización se comportan como un polipéptido desestructurado). En la **figura 5.2** se representa la estimación del porcentaje de fuerza que soporta la interacción vs la fuerza medida, y se incluye el histograma de fuerzas  $F_D$  de la interacción cohesina/dockerina. Tal y cómo se observa en la figura, el porcentaje de fuerza que sufre la interacción es igual o mayor al 90% para las fuerzas observadas, por lo que efectivamente la mayor parte de la carga de fuerza está siendo soportada por la interacción. Debemos aclarar, que el objetivo de esta tesis no era la caracterización de la estabilidad mecánica de la interacción cohesina/dockerina, por lo que futuros experimentos deberán abordar un estudio nanomecánico más exhaustivo del sistema,

mediante medidas cinéticas y el uso de mutantes desestabilizantes de la interacción como los mutantes de cohesina D39A y L83A (Stahl et al., 2012).

Uno de los grandes problemas encontrados en este desarrollo ha sido la baja eficiencia de adquisición de datos por AFM-SMFS. Para que el diseño se convierta en una estrategia estándar en el análisis de interacciones intermoleculares por AFM-SMFS, será de vital importancia aumentar la eficiencia en la adquisición de datos. Esto sería especialmente importante sobre todo para el estudio los parámetros cinéticos ( $k_d(0)$  y  $\Delta x_d$ , **sección 1.3.3**) mediante la determinación de la fuerzas de disociación a diferentes velocidades (Rangl et al., 2013; Schoeler et al., 2014), aunque el problema podría paliarse mediante el uso de simulaciones de Montecarlo que permiten hallar estos parámetros con el histograma de fuerzas a una sola velocidad (Pernigo et al., 2010; Bertz et al., 2013). Una de las razones directas que puede explicar en parte la baja tasa de adquisición de datos, reside en las condiciones tan restrictivas en la selección de datos que debe cumplir un registro para poder ser considerado para su análisis. Los registros deben de tener entre 4-6 repeticiones de módulos marcadores, un  $\Delta L_c$  concreto asociado con el pico de interacción y, en el caso de la interacción cohesina/dockerina dado que es menos mecanoestable que los módulos marcadores, una zona proximal limpia de interacciones inespecíficas contaminantes. Para el desarrollo de futuras versiones de esta estrategia se debería explorar la posibilidad de utilizar marcadores de menor estabilidad mecánica, e incluso utilizar marcadores diferentes en cada una de las proteínas interaccionantes. Esto, dependiendo de la mecanoestabilidad relativa de la interacción de interés, preservaría el pico de fuerza de la contaminación por el ruido de la zona proximal, y bastaría con que los registros tuviesen solamente 2 módulos marcadores (uno de cada tipo). Finalmente hay que tener en cuenta también que en las construcciones las proteínas se ven en un contexto de multimodularidad diferente a su contexto fisiológico, lo que podría ser causa de un plegamiento poco eficiente (Batey et al., 2008) que origine la baja eficiencia.

Pese al desarrollo a lo largo de los años de complejos y depurados protocolos de funcionalización para el estudio de interacciones intermoleculares, todos estos procedimientos tropezaban siempre con el mismo problema: la **imposibilidad** de distinguir en los registros el evento de ruptura de la interacción del desanclaje de alguna de la proteínas de la superficie, o la adhesión inespecífica (Merkel, 2001; Pernigo et al., 2010; Berkemeier et al., 2011). Por ello, en los últimos años han existido intentos de generar huellas moleculares identificativas de los eventos de ruptura de interacciones intermoleculares. Uno de ellos es el desarrollo de asas poliméricas que portan repeticiones equiespaciadas de una de las moléculas interaccionantes, este diseño genera un patrón de picos de fuerza similar al generado con la técnica de las poliproteínas. Sin embargo, pese a

ser un intento prometedor, esta estrategia no se ha generalizado al estudio de interacciones proteína-proteína (Valle et al., 2008; Han et al., 2012). En (Bertz et al., 2009; Pernigo et al., 2010; Bertz et al., 2010; Berkemeier et al., 2011; Kim et al., 2012) se incorporan a las proteínas interaccionantes poliproteínas (de ubiquitinas, dominios filamina, módulos inmunoglobulina, y SNasa) que actúan como asas y como marcadores de monomolecularidad, de manera similar al diseño presentado aquí. En cuatro de estos trabajos (Bertz et al., 2009; Pernigo et al., 2010; Bertz et al., 2010; Berkemeier et al., 2011), los autores van más allá, y utilizan la información de la estructura tridimensional resuelta de las proteínas para unir por puentes disulfuro sus extremos mediante extensiones flexibles. Estas extensiones provocan un incremento apreciable de la longitud de contorno tras la ruptura de la interacción. Esta implementación recuerda a nuestro diseño, con la liberación de un  $\Delta L_c$  tras la ruptura de la interacción. Sin embargo en estos trabajos, la identificación del pico de fuerza de la interacción no es directa. Para ello se requiere un conocimiento profundo de la estructura tridimensional de las proteínas, y el reconocimiento del  $\Delta L_c$  del desplegamiento de los módulos proteicos que interaccionan entre sí (que se despliegan en algún momento del experimento, al seguir expuestos a la fuerza incluso después de la ruptura de la interacción). Sólo un análisis meticuloso de las poblaciones de  $\Delta L_c$ , e información de alta resolución del tamaño y configuración de las proteínas permite identificar el pico de fuerza de la interacción. Por otra parte, hay que tener en cuenta que la estrategia implementada en todos estos trabajos no es de aplicación general, ya que en cada caso particular se utilizan las características estructurales de las moléculas interaccionantes para intentar conseguir la unión covalente de sus extremos (que por otra parte se deja al azar), variando en cada caso la naturaleza y longitud de las extensiones flexibles. Incluso se debe mencionar también que en (Bertz et al., 2010) no se llega a observar el desplegamiento de las ubiquitinas que forman la extensión flexible después de la ruptura de la interacción, lo que se sería esperable debido a la configuración experimental usada. Finalmente, se debe puntualizar que nuestra estrategia fue diseñada de forma independiente con anterioridad a la publicación de estos artículos mencionados.

En resumen, la mayor ventaja que presenta la estrategia presentada en este trabajo con respecto a los trabajos encontrados en la bibliografía y que acabamos de discutir, es que esta estrategia es independiente de características concretas de la interacción de interés, por lo que es potencialmente de aplicación general a cualquier interacción entre proteínas.

## 6. Conclusiones

---

Las conclusiones del trabajo presentado en esta tesis se resumen en los siguientes puntos:

1. La región ARM de la  $\beta$ -catenina presenta una baja mecanoestabilidad, y su desnaturalización mecánica ocurre a través del desplegamiento de las repeticiones ARM, tanto a nivel individual como en múltiplos de las mismas.
2. El desplegamiento mecánico de esta estructura puede ocurrir a través de múltiples caminos de desplegamiento, con diferentes intermedios, lo que indica que el proceso tiene lugar a través de un paisaje energético rugoso con mínimos energéticos muy superficiales.
3. Las propiedades mecánicas de la región ARM de la  $\beta$ -catenina parecen ser muy sensibles a pequeños cambios en la estructura o/y en la energética inicial del sistema, y sigue múltiples caminos de desplegamiento ante fuerza, que parecen resultar modulados por sus terminales desestructurados.
4. En base a estos resultados, proponemos que la  $\beta$ -catenina podría actuar como un tampón mecánico a lo largo de todo el rango de fuerzas al que podría estar sometida, “sacrificando” algunas de sus repeticiones para mantener los contactos entre las uniones cadherina y el citoesqueleto de actina.
5. Se ha ideado una estrategia autocontrolada para el estudio de interacciones intermoleculares por AFM-SMFS que permite la identificación directa del pico de fuerza generado por la ruptura de la interacción. La estrategia se basa en la inclusión en las proteínas de interés de controles de monomolecularidad y en el uso de un marcador que reporte el pico de ruptura de la interacción.
6. La estrategia diseñada es de tipo general, no depende de la naturaleza ni las características de las proteínas de interés. El control de monomolecularidad viene dado por la inclusión de poliproteínas en las proteínas a estudio, mientras que el marcador del pico de fuerza de la interacción consiste en la liberación de un  $\Delta L_c$  característico, debido a la exposición a la fuerza de un polipéptido elastomérico tras la ruptura de la misma.
7. Se ha demostrado a nivel de prueba de concepto que la medida autocontrolada de interacciones intermoleculares es viable. El sistema se implementó sobre la interacción cohesina/dockerina de *Clostridium thermocellum*, determinándose con éxito su estabilidad mecánica. Los registros de fuerza-extensión de la construcción presentaban todos los controles introducidos, y la identificación del pico de fuerza de la interacción es directa.

8. El diseño modular de la estrategia permite generalizar la configuración experimental implementada (aplicando fuerza desde los C-terminales de las proteínas de interés) fácilmente a otras configuraciones con pequeños cambios. Para implementar esta estrategia modular, se ha desarrollado una plataforma de vectores tipo *cassette* que permiten un fácil clonaje y un protocolo de purificación de proteínas estandarizado, y que puede servir de plataforma para implementar el sistema en otras proteínas, y en otras configuraciones.

---

## 7. Bibliografía

---

- Abe, K., and M. Takeichi. 2008. Eplln mediates linkage of the cadherin catenin complex to f-actin and stabilizes the circumferential actin belt. *Proceedings of the National Academy of Sciences of the United States of America*. 105:13-19.
- Ainavarapu, S.R., J. Brujic, H.H. Huang, A.P. Wiita, H. Lu, L. Li, K.A. Walther, M. Carrion-Vazquez, H. Li, and J.M. Fernandez. 2007. Contour length and refolding rate of a small protein controlled by engineered disulfide bonds. *Biophysical journal*. 92:225-233.
- Ainavarapu, S.R.K., L. Li, C.L. Badilla, and J.M. Fernandez. 2005. Ligand binding modulates the mechanical stability of dihydrofolate reductase. *Biophysical journal*. 89:3337-3344.
- Alegre-Cebollada, J., C.L. Badilla, and J.M. Fernandez. 2010. Isopeptide bonds block the mechanical extension of pili in pathogenic streptococcus pyogenes. *The Journal of biological chemistry*. 285:11235-11242.
- Asparuhova, M.B., L. Gelman, and M. Chiquet. 2009. Role of the actin cytoskeleton in tuning cellular responses to external mechanical stress. *Scandinavian journal of medicine & science in sports*. 19:490-499.
- Aubin-Tam, M.E., A.O. Olivares, R.T. Sauer, T.A. Baker, and M.J. Lang. 2011. Single-molecule protein unfolding and translocation by an atp-fueled proteolytic machine. *Cell*. 145:257-267.
- Ballet, T., L. Boulange, Y. Brechet, F. Bruckert, and M. Weidenhaupt. 2010. Protein conformational changes induced by adsorption onto material surfaces: An important issue for biomedical applications of material science. *In Bulletin of the Polish Academy of Sciences: Technical Sciences*. Vol. 58. 303.
- Bard, L., C. Boscher, M. Lambert, R.M. Mege, D. Choquet, and O. Thoumine. 2008. A molecular clutch between the actin flow and n-cadherin adhesions drives growth cone migration. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 28:5879-5890.
- Batey, S., A.A. Nickson, and J. Clarke. 2008. Studying the folding of multidomain proteins. *HFSP Journal*. 2:365-377.
- Baumgartner, W., N. Golenhofen, N. Grundhofer, J. Wiegand, and D. Drenckhahn. 2003. Ca<sup>2+</sup> dependency of n-cadherin function probed by laser tweezer and atomic force microscopy. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 23:11008-11014.
- Baumgartner, W., P. Hinterdorfer, W. Ness, A. Raab, D. Vestweber, H. Schindler, and D. Drenckhahn. 2000. Cadherin interaction probed by atomic force microscopy. *Proceedings of the National Academy of Sciences of the United States of America*. 97:4005-4010.
- Behrens, J., J.P. von Kries, M. Kuhl, L. Bruhn, D. Wedlich, R. Grosschedl, and W. Birchmeier. 1996. Functional interaction of [beta]-catenin with the transcription factor lef-1. *Nature*. 382:638-642.
- Bell, G.I. 1978. Models for the specific adhesion of cells to cells. *Science (New York, N.Y.)*. 200:618-627.
- Berkemeier, F., M. Bertz, S. Xiao, N. Pinotsis, M. Wilmanns, F. Gräter, and M. Rief. 2011. Fast-folding  $\alpha$ -helices as reversible strain absorbers in the muscle protein myomesin. *Proceedings of the National Academy of Sciences*. 108:14139-14144.
- Berquand, A., N. Xia, D.G. Castner, B.H. Clare, N.L. Abbott, V. Dupres, Y. Adriaensen, and Y.F. Dufrene. 2005. Antigen binding forces of single

- antilysozyme fv fragments explored by atomic force microscopy. *Langmuir : the ACS journal of surfaces and colloids*. 21:5517-5523.
- Bertz, M., J. Buchner, and M. Rief. 2013. Mechanical stability of the antibody domain ch3 homodimer in different oxidation states. *Journal of the American Chemical Society*. 135:15085-15091.
- Bertz, M., J. Chen, M.J. Feige, T.M. Franzmann, J. Buchner, and M. Rief. 2010. Structural and mechanical hierarchies in the alpha-crystallin domain dimer of the hyperthermophilic small heat shock protein hsp16.5. *Journal of molecular biology*. 400:1046-1056.
- Bertz, M., M. Wilmanns, and M. Rief. 2009. The titin-telethonin complex is a directed, superstable molecular bond in the muscle z-disk. *Proceedings of the National Academy of Sciences*. 106:13307-13310.
- Blommel, P.G., and B.G. Fox. 2007. A combined approach to improving large-scale production of tobacco etch virus protease. *Protein expression and purification*. 55:53-68.
- Bohm, G., R. Muhr, and R. Jaenicke. 1992. Quantitative analysis of protein far uv circular dichroism spectra by neural networks. *Protein engineering*. 5:191-195.
- Borghi, N., M. Sorokina, O.G. Shcherbakova, W.I. Weis, B.L. Pruitt, W.J. Nelson, and A.R. Dunn. 2012. E-cadherin is under constitutive actomyosin-generated tension that is increased at cell-cell contacts upon externally applied stretch. *Proceedings of the National Academy of Sciences of the United States of America*. 109:12568-12573.
- Bouvier, A., J. Chapline, R. Boerner, S. Jeyarajah, S. Cook, P.S. Acharya, I. Henderson, J.L. Schrimsher, and S.R. Shepard. 2003. Identifying and modulating disulfide formation in the biopharmaceutical production of a recombinant protein vaccine candidate. *Journal of biotechnology*. 103:257-271.
- Bradley, C.M., and D. Barrick. 2002. Limits of cooperativity in a structurally modular protein: Response of the notch ankyrin domain to analogous alanine substitutions in each repeat. *Journal of molecular biology*. 324:373-386.
- Brewer, S.J., and H.M. Sassenfeld. 1985. The purification of recombinant proteins using c-terminal polyarginine fusions. *Trends in Biotechnology*. 3:119-122.
- Brunet, T., A. Bouclet, P. Ahmadi, D. Mitrossilis, B. Driquez, A.-C. Brunet, L. Henry, F. Serman, G. Béalle, C. Ménager, F. Dumas-Bouchiat, D. Givord, C. Yanicostas, D. Le-Roy, N.M. Dempsey, A. Plessis, and E. Farge. 2013. Evolutionary conservation of early mesoderm specification by mechanotransduction in bilateria. *Nat Commun*. 4.
- Bustamante, C., Y.R. Chemla, N.R. Forde, and D. Izhaky. 2004. Mechanical processes in biochemistry. *Annual review of biochemistry*. 73:705-748.
- Bustamante, C., J.F. Marko, E.D. Siggia, and S. Smith. 1994. Entropic elasticity of lambda-phage DNA. *Science (New York, N.Y.)*. 265:1599-1600.
- Cantor, C.R., and P.R. Schimmel. 1980. Biophysical chemistri.Pat i: The conformation of biological macromolecules. W.H. Freeman and company, U.S.A.
- Cao, Y., R. Kuske, and H. Li. 2008. Direct observation of markovian behavior of the mechanical unfolding of individual proteins. *Biophysical journal*. 95:782-788.
- Cao, Y., C. Lam, M. Wang, and H. Li. 2006. Nonmechanical protein can have significant mechanical stability. *Angewandte Chemie*. 45:642-645.
- Cao, Y., and H. Li. 2008. Engineered elastomeric proteins with dual elasticity can be controlled by a molecular regulator. *Nat Nano*. 3:512-516.
- Carrión-Vázquez, M., A.F. Oberhauser, H. Díez, R. Hervás, J. Oroz, J. Fernández, and D. Martínez-Martín. 2006. Protein nanomechanics—as studied by afm single-

- molecule force spectroscopy. *Advanced techniques in biophysics*. Springer, Heidelberg:163–245.
- Carrion-Vazquez, M., A.F. Oberhauser, T.E. Fisher, P.E. Marszalek, H. Li, and J.M. Fernandez. 2000. Mechanical design of proteins studied by single-molecule force spectroscopy and protein engineering. *Progress in biophysics and molecular biology*. 74:63-91.
- Carrion-Vazquez, M., A.F. Oberhauser, S.B. Fowler, P.E. Marszalek, S.E. Broedel, J. Clarke, and J.M. Fernandez. 1999. Mechanical and chemical unfolding of a single protein: A comparison. *Proceedings of the National Academy of Sciences of the United States of America*. 96:3694-3699.
- Case, D.A., T.A. Darden, Cheatham, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, M. Crowley, R.C. Walker, W. Zhang, K.M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, and P.A. Kollman. 2008a. Amber 10. University of California, San Francisco.
- Case, N., M. Ma, B. Sen, Z. Xie, T.S. Gross, and J. Rubin. 2008b. Beta-catenin levels influence rapid mechanical responses in osteoblasts. *The Journal of biological chemistry*. 283:29196-29205.
- Castaño, J., I. Raurell, J.A. Piedra, S. Miravet, M. Dunach, and A. Garcia de Herreros. 2002. Beta-catenin n- and c-terminal tails modulate the coordinated binding of adherens junction proteins to beta-catenin. *The Journal of biological chemistry*. 277:31541-31550.
- Cras, J.J., C.A. Rowe-Taitt, D.A. Nivens, and F.S. Ligler. 1999. Comparison of chemical cleaning methods of glass in preparation for silanization. *Biosensors and Bioelectronics*. 14:683-688.
- Choi, H.J., A.H. Huber, and W.I. Weis. 2006. Thermodynamics of beta-catenin-ligand interactions: The roles of the n- and c-terminal tails in modulating binding affinity. *The Journal of biological chemistry*. 281:1027-1038.
- del Rio, A., R. Perez-Jimenez, R. Liu, P. Roca-Cusachs, J.M. Fernandez, and M.P. Sheetz. 2009. Stretching single talin rod molecules activates vinculin binding. *Science (New York, N.Y.)*. 323:638-641.
- Desai, R., R. Sarpal, N. Ishiyama, M. Pellikka, M. Ikura, and U. Tepass. 2013. Monomeric alpha-catenin links cadherin to the actin cytoskeleton. *Nature cell biology*. 15:261-273.
- Desprat, N., W. Supatto, P.A. Pouille, E. Beaurepaire, and E. Farge. 2008. Tissue deformation modulates twist expression to determine anterior midgut differentiation in drosophila embryos. *Developmental cell*. 15:470-477.
- Drees, F., S. Pokutta, S. Yamada, W.J. Nelson, and W.I. Weis. 2005. Alpha-catenin is a molecular switch that binds e-cadherin-beta-catenin and regulates actin-filament assembly. *Cell*. 123:903-915.
- Evans, E., and K. Ritchie. 1997. Dynamic strength of molecular adhesion bonds. *Biophysical journal*. 72:1541-1555.
- Evans, E., and K. Ritchie. 1999. Strength of a weak bond connecting flexible polymer chains. *Biophysical journal*. 76:2439-2447.
- Florin, E.L., V.T. Moy, and H.E. Gaub. 1994. Adhesion forces between individual ligand-receptor pairs. *Science (New York, N.Y.)*. 264:415-417.
- Fuhrmann, A., and R. Ros. 2010. Single-molecule force spectroscopy: A method for quantitative analysis of ligand-receptor interactions. *Nanomedicine*. 5:657-666.



- Fuhrmann, A., J.C. Schoening, D. Anselmetti, D. Staiger, and R. Ros. 2009. Quantitative analysis of single-molecule rna-protein interaction. *Biophysical journal*. 96:5030-5039.
- Ganz, A., M. Lambert, A. Saez, P. Silberzan, A. Buguin, R.M. Mege, and B. Ladoux. 2006. Traction forces exerted through n-cadherin contacts. *Biology of the cell / under the auspices of the European Cell Biology Organization*. 98:721-730.
- Gao, Y., S. Zorman, G. Gundersen, Z. Xi, L. Ma, G. Sirinakis, J.E. Rothman, and Y. Zhang. 2012. Single reconstituted neuronal snare complexes zipper in three distinct stages. *Science (New York, N.Y.)*. 337:1340-1343.
- Garcia-Manyes, S., J. Brujić, C.L. Badilla, and J.M. Fernández. 2007. Force-clamp spectroscopy of single-protein monomers reveals the individual unfolding and folding pathways of i27 and ubiquitin. *Biophysical journal*. 93:2436-2446.
- Goldmann, W.H. 2012. Mechanotransduction and focal adhesions. *Cell biology international*. 36:649-652.
- Gough, J.D., and W.J. Lees. 2005. Effects of redox buffer properties on the folding of a disulfide-containing protein: Dependence upon ph, thiol pka, and thiol concentration. *Journal of biotechnology*. 115:279-290.
- Grandbois, M., M. Beyer, M. Rief, H. Clausen-Schaumann, and H.E. Gaub. 1999. How strong is a covalent bond? *Science (New York, N.Y.)*. 283:1727-1730.
- Halliwell, C.M., and A.E. Cass. 2001. A factorial analysis of silanization conditions for the immobilization of oligonucleotides on glass surfaces. *Analytical chemistry*. 73:2476-2483.
- Han, X., M. Qin, H. Pan, Y. Cao, and W. Wang. 2012. A versatile "multiple fishhooks" approach for the study of ligand-receptor interactions using single-molecule atomic force microscopy. *Langmuir : the ACS journal of surfaces and colloids*. 28:10020-10025.
- Harris, T.J., and U. Tepass. 2010. Adherens junctions: From molecules to morphogenesis. *Nat Rev Mol Cell Biol*. 11:502-514.
- Hawkins, G.D., C.J. Cramer, and D.G. Truhlar. 1996. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *The Journal of Physical Chemistry*. 100:19824-19839.
- Hervas, R., J. Oroz, A. Galera-Prat, O. Goni, A. Valbuena, A.M. Vera, A. Gomez-Sicilia, F. Losada-Urzaiz, V.N. Uversky, M. Menendez, D.V. Laurents, M. Bruix, and M. Carrion-Vazquez. 2012. Common features at the start of the neurodegeneration cascade. *PLoS biology*. 10:e1001335.
- Himmel, M.E., Q. Xu, Y. Luo, S.-Y. Ding, R. Lamed, and E.A. Bayer. 2010. Microbial enzyme systems for biomass conversion: Emerging paradigms. *Biofuels*. 1:323-341.
- Hinterdorfer, P. 2005. Molecular recognition force microscopy. *In Nanotribology and nanomechanics*. B. Bhushan, editor. Springer Berlin Heidelberg. 283-312.
- Hinterdorfer, P., and Y.F. Dufrene. 2006. Detection and localization of single molecular recognition events using atomic force microscopy. *Nature methods*. 3:347-355.
- Hu, H.-L., M.-Y. Wang, C.-H. Chung, and S.-Y. Suen. 2006. Purification of vp3 protein of infectious bursal disease virus using nickel ion-immobilized regenerated cellulose-based membranes. *Journal of Chromatography B*. 840:76-84.
- Huber, A.H., W.J. Nelson, and W.I. Weis. 1997. Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell*. 90:871-882.

- Huber, A.H., and W.I. Weis. 2001. The structure of the beta-catenin/e-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. *Cell*. 105:391-402.
- Humphrey, W., A. Dalke, and K. Schulten. 1996. Vmd: Visual molecular dynamics. *Journal of molecular graphics*. 14:33-38, 27-38.
- Hutter, J.L., and J. Bechhoefer. 1993. Calibration of atomic-force microscope tips. *Review of Scientific Instruments*. 64:1868-1873.
- Huveneers, S., and J. de Rooij. 2013. Mechanosensitive systems at the cadherin-f-actin interface. *Journal of Cell Science*. 126:403-413.
- Ikeda, T., and A. Kuroda. 2011. Why does the silica-binding protein "si-tag" bind strongly to silica surfaces? Implications of conformational adaptation of the intrinsically disordered polypeptide to solid surfaces. *Colloids and surfaces. B, Biointerfaces*. 86:359-363.
- Improta, S., A.S. Politou, and A. Pastore. 1996. Immunoglobulin-like modules from titin i-band: Extensible components of muscle elasticity. *Structure (London, England : 1993)*. 4:323-337.
- Indra, I., S. Hong, R. Troyanovsky, B. Kormos, and S. Troyanovsky. 2013. The adherens junction: A mosaic of cadherin and nectin clusters bundled by actin filaments. *The Journal of investigative dermatology*. 133:2546-2554.
- Jollymore, A., C. Lethias, Q. Peng, Y. Cao, and H. Li. 2009. Nanomechanical properties of tenascin-x revealed by single-molecule force spectroscopy. *Journal of molecular biology*. 385:1277-1286.
- Kapust, R.B., J. Tozser, J.D. Fox, D.E. Anderson, S. Cherry, T.D. Copeland, and D.S. Waugh. 2001. Tobacco etch virus protease: Mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein engineering*. 14:993-1000.
- Kapust, R.B., and D.S. Waugh. 1999. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein science : a publication of the Protein Society*. 8:1668-1674.
- Kellermayer, M.S.Z., C. Bustamante, and H.L. Granzier. 2003. Mechanics and structure of titin oligomers explored with atomic force microscopy. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. 1604:105-114.
- Kim, B.M., and M.G. Oakley. 2002. A general method for selection and screening of coiled coils on the basis of relative helix orientation. *Journal of the American Chemical Society*. 124:8237-8244.
- Kim, M., K. Abdi, G. Lee, M. Rabbi, W. Lee, M. Yang, C.J. Schofield, V. Bennett, and P.E. Marszalek. 2010. Fast and forceful refolding of stretched alpha-helical solenoid proteins. *Biophysical journal*. 98:3086-3092.
- Kim, M., C.-C. Wang, F. Benedetti, and P.E. Marszalek. 2012. A nanoscale force probe for gauging intermolecular interactions. *Angewandte Chemie International Edition*. 51:1903-1906.
- Kong, F., A.J. Garcia, A.P. Mould, M.J. Humphries, and C. Zhu. 2009. Demonstration of catch bonds between an integrin and its ligand. *The Journal of cell biology*. 185:1275-1284.
- Korf, U., T. Kohl, H. van der Zandt, R. Zahn, S. Schleege, B. Ueberle, S. Wandschneider, S. Bechtel, M. Schnolzer, H. Ottleben, S. Wiemann, and A. Poustka. 2005. Large-scale protein expression for proteome research. *Proteomics*. 5:3571-3580.

- Ladoux, B., E. Anon, M. Lambert, A. Rabodzey, P. Hersen, A. Buguin, P. Silberzan, and R.M. Mege. 2010. Strength dependence of cadherin-mediated adhesions. *Biophysical journal*. 98:534-542.
- Lambooy, J.A., H. Kim, K.S. Lee, T. Ha, and E.A. Komives. 2011. Visualization of the nanospring dynamics of the ikappalpha ankyrin repeat domain in real time. *Proceedings of the National Academy of Sciences of the United States of America*. 108:10178-10183.
- le Duc, Q., Q. Shi, I. Blonk, A. Sonnenberg, N. Wang, D. Leckband, and J. de Rooij. 2010. Vinculin potentiates e-cadherin mechanosensing and is recruited to actin-anchored sites within adherens junctions in a myosin ii-dependent manner. *The Journal of cell biology*. 189:1107-1115.
- Lee, C.K., Y.M. Wang, L.S. Huang, and S. Lin. 2007. Atomic force microscopy: Determination of unbinding force, off rate and energy barrier for protein-ligand interaction. *Micron*. 38:446-461.
- Lee, G., K. Abdi, Y. Jiang, P. Michaely, V. Bennett, and P.E. Marszalek. 2006. Nanospring behaviour of ankyrin repeats. *Nature*. 440:246-249.
- Li, H., M. Carrion-Vazquez, A.F. Oberhauser, P.E. Marszalek, and J.M. Fernandez. 2000. Point mutations alter the mechanical stability of immunoglobulin modules. *Nature structural biology*. 7:1117-1120.
- Li, H., W.A. Linke, A.F. Oberhauser, M. Carrion-Vazquez, J.G. Kerkvliet, H. Lu, P.E. Marszalek, and J.M. Fernandez. 2002. Reverse engineering of the giant muscle protein titin. *Nature*. 418:998-1002.
- Li, H., A.F. Oberhauser, S.D. Redick, M. Carrion-Vazquez, H.P. Erickson, and J.M. Fernandez. 2001. Multiple conformations of pev proteins detected by single-molecule techniques. *Proceedings of the National Academy of Sciences of the United States of America*. 98:10682-10686.
- Li, L., S. Wetzel, A. Plückthun, and J.M. Fernandez. 2006. Stepwise unfolding of ankyrin repeats in a single protein revealed by atomic force microscopy. *Biophysical journal*. 90:L30-L32.
- Lin, Z., D. Madan, and H.S. Rye. 2008. Groel stimulates protein folding through forced unfolding. *Nat Struct Mol Biol*. 15:303-311.
- Linke, W.A., and A. Grutzner. 2008. Pulling single molecules of titin by afm--recent advances and physiological implications. *Pflugers Archiv : European journal of physiology*. 456:101-115.
- Liu, Z., J.L. Tan, D.M. Cohen, M.T. Yang, N.J. Sniadecki, S.A. Ruiz, C.M. Nelson, and C.S. Chen. 2010. Mechanical tugging force regulates the size of cell-cell junctions. *Proceedings of the National Academy of Sciences of the United States of America*. 107:9944-9949.
- Lumb, K.J., and P.S. Kim. 1995. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry*. 34:8642-8648.
- Lv, S., D.M. Dudek, Y. Cao, M.M. Balamurali, J. Gosline, and H. Li. 2010. Designed biomaterials to mimic the mechanical properties of muscles. *Nature*. 465:69-73.
- Ma, L., M. Xu, J.R. Forman, J. Clarke, and A.F. Oberhauser. 2009. Naturally occurring mutations alter the stability of polycystin-1 polycystic kidney disease (pkd) domains. *The Journal of biological chemistry*. 284:32942-32949.
- Maier, B., L. Potter, M. So, C.D. Long, H.S. Seifert, and M.P. Sheetz. 2002. Single pilus motor forces exceed 100 pn. *Proceedings of the National Academy of Sciences of the United States of America*. 99:16012-16017.
- Mammoto, A., T. Mammoto, and D.E. Ingber. 2012. Mechanosensitive mechanisms in transcriptional regulation. *J Cell Sci*. 125:3061-3073.

- Markley, J.L., A. Bax, Y. Arata, C.W. Hilbers, R. Kaptein, B.D. Sykes, P.E. Wright, and K. Wuthrich. 1998. Recommendations for the presentation of nmr structures of proteins and nucleic acids. Iupac-iubmb-iupab inter-union task group on the standardization of data bases of protein and nucleic acid structures determined by nmr spectroscopy. *Journal of biomolecular NMR*. 12:1-23.
- Marko, J.F., and E.D. Siggia. 1995. Stretching DNA. *Macromolecules* 28:8759-8770.
- Marshall, B.T., M. Long, J.W. Piper, T. Yago, R.P. McEver, and C. Zhu. 2003. Direct observation of catch bonds involving cell-adhesion molecules. *Nature*. 423:190-193.
- Mechaly, A., H.P. Fierobe, A. Belaich, J.P. Belaich, R. Lamed, Y. Shoham, and E.A. Bayer. 2001. Cohesin-dockerin interaction in cellulosome assembly: A single hydroxyl group of a dockerin domain distinguishes between nonrecognition and high affinity recognition. *The Journal of biological chemistry*. 276:9883-9888.
- Mege, R.M., J. Gavard, and M. Lambert. 2006. Regulation of cell-cell junctions by the cytoskeleton. *Current opinion in cell biology*. 18:541-548.
- Merkel, R. 2001. Force spectroscopy on single passive biomolecules and single biomolecular bonds. *Physics Reports*. 346:343-385.
- Merkel, R., P. Nassoy, A. Leung, K. Ritchie, and E. Evans. 1999. Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature*. 397:50-53.
- Miroux, B., and J.E. Walker. 1996. Over-production of proteins in escherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of molecular biology*. 260:289-298.
- Morfill, J., K. Blank, C. Zahnd, B. Luginbühl, F. Kühner, K.-E. Gottschalk, A. Plückthun, and H.E. Gaub. 2007. Affinity-matured recombinant antibody fragments analyzed by single-molecule force spectroscopy. *Biophysical journal*. 93:3583-3590.
- Morfill, J., J. Neumann, K. Blank, U. Steinbach, E.M. Puchner, K.E. Gottschalk, and H.E. Gaub. 2008. Force-based analysis of multidimensional energy landscapes: Application of dynamic force spectroscopy and steered molecular dynamics simulations to an antibody fragment-peptide complex. *Journal of molecular biology*. 381:1253-1266.
- Mukhopadhyay, A. 2000. Reversible protection of disulfide bonds followed by oxidative folding render recombinant hcgbeta highly immunogenic. *Vaccine*. 18:1802-1810.
- Nakanishi, K., T. Sakiyama, and K. Imamura. 2001. On the adsorption of proteins on solid surfaces, a common but very complicated phenomenon. *Journal of Bioscience and Bioengineering*. 91:233-244.
- Nelson, W.J. 2008. Regulation of cell-cell adhesion by the cadherin-catenin complex. *Biochemical Society transactions*. 36:149-155.
- Nelson, W.J., and R. Nusse. 2004. Convergence of wnt, beta-catenin, and cadherin pathways. *Science (New York, N.Y.)*. 303:1483-1487.
- Nevo, R., C. Stroh, F. Kienberger, D. Kaftan, V. Brumfeld, M. Elbaum, Z. Reich, and P. Hinterdorfer. 2003. A molecular switch between alternative conformational states in the complex of ran and importin beta1. *Nature structural biology*. 10:553-557.
- O'Shea, E.K., K.J. Lumb, and P.S. Kim. 1993. Peptide 'velcro': Design of a heterodimeric coiled coil. *Current biology : CB*. 3:658-667.
- Oakley, M.G., and P.S. Kim. 1998. A buried polar interaction can direct the relative orientation of helices in a coiled coil. *Biochemistry*. 37:12603-12610.

- Oberhauser, A.F., C. Badilla-Fernandez, M. Carrion-Vazquez, and J.M. Fernandez. 2002. The mechanical hierarchies of fibronectin observed with single-molecule afm. *Journal of molecular biology*. 319:433-447.
- Oberhauser, A.F., P.K. Hansma, M. Carrion-Vazquez, and J.M. Fernandez. 2001. Stepwise unfolding of titin under force-clamp atomic force microscopy. *Proceedings of the National Academy of Sciences of the United States of America*. 98:468-472.
- Oberhauser, A.F., P.E. Marszalek, H.P. Erickson, and J.M. Fernandez. 1998. The molecular elasticity of the extracellular matrix protein tenascin. *Nature*. 393:181-185.
- Olsen, J.V., S.-E. Ong, and M. Mann. 2004. Trypsin cleaves exclusively c-terminal to arginine and lysine residues. *Molecular & Cellular Proteomics*. 3:608-614.
- Oroz, J., A. Valbuena, A.M. Vera, J. Mendieta, P. Gomez-Puertas, and M. Carrion-Vazquez. 2011. Nanomechanics of the cadherin ectodomain: "Canalization" by  $Ca^{2+}$  binding results in a new mechanical element. *The Journal of biological chemistry*. 286:9405-9418.
- Pernigo, S., A. Fukuzawa, M. Bertz, M. Holt, M. Rief, R.A. Steiner, and M. Gautel. 2010. Structural insight into m-band assembly and mechanics from the titin-obscurin-like-1 complex. *Proceedings of the National Academy of Sciences of the United States of America*. 107:2908-2913.
- Perret, E., A. Leung, H. Feracci, and E. Evans. 2004. Trans-bonded pairs of e-cadherin exhibit a remarkable hierarchy of mechanical strengths. *Proceedings of the National Academy of Sciences of the United States of America*. 101:16472-16477.
- Piedra, J., D. Martinez, J. Castano, S. Miravet, M. Dunach, and A.G. de Herreros. 2001. Regulation of beta-catenin structure and activity by tyrosine phosphorylation. *The Journal of biological chemistry*. 276:20436-20443.
- Piotto, M., V. Saudek, and V. Sklenar. 1992. Gradient-tailored excitation for single-quantum nmr spectroscopy of aqueous solutions. *Journal of biomolecular NMR*. 2:661-665.
- Puchner, E.M., and H.E. Gaub. 2009. Force and function: Probing proteins with afm-based force spectroscopy. *Current opinion in structural biology*. 19:605-614.
- Rakshit, S., Y. Zhang, K. Manibog, O. Shafraz, and S. Sivasankar. 2012. Ideal, catch, and slip bonds in cadherin adhesion. *Proceedings of the National Academy of Sciences*. 109:18815-18820.
- Randles, L.G., R.W.S. Rounsevell, and J. Clarke. 2007. Spectrin domains lose cooperativity in forced unfolding. *Biophysical journal*. 92:571-577.
- Rangl, M., A. Ebner, J. Yamada, C. Rankl, R. Tampé, H.J. Gruber, M. Rexach, and P. Hinterdorfer. 2013. Single-molecule analysis of the recognition forces underlying nucleo-cytoplasmic transport. *Angewandte Chemie International Edition*. 52:10356-10359.
- Ratheesh, A., and A.S. Yap. 2012. A bigger picture: Classical cadherins and the dynamic actin cytoskeleton. *Nat Rev Mol Cell Biol*. 13:673-679.
- Ray, S., H.P. Foote, and T. Lechler. 2013. Beta-catenin protects the epidermis from mechanical stresses. *The Journal of cell biology*. 202:45-52.
- Rief, M., J. Fernandez, and H. Gaub. 1998. Elastically coupled two-level systems as a model for biopolymer extensibility. *Phys. Rev. Lett*. 81:4764-4767.
- Rief, M., M. Gautel, F. Oesterhelt, J.M. Fernandez, and H.E. Gaub. 1997. Reversible unfolding of individual titin immunoglobulin domains by afm. *Science (New York, N.Y.)*. 276:1109-1112.

- Rief, M., J. Pascual, M. Saraste, and H.E. Gaub. 1999. Single molecule force spectroscopy of spectrin repeats: Low unfolding forces in helix bundles. *Journal of molecular biology*. 286:553-561.
- Riggs, P. 2000. Expression and purification of recombinant proteins by fusion to maltose-binding protein. *Molecular biotechnology*. 15:51-63.
- Ritco-Vonsovici, M., A. Ababou, and M. Horton. 2007. Molecular plasticity of  $\beta$ -catenin: New insights from single-molecule measurements and md simulation. *Protein science : a publication of the Protein Society*. 16:1984-1998.
- Ritort, F. 2006. Single-molecule experiments in biological physics: Methods and applications. *Journal of physics. Condensed matter : an Institute of Physics journal*. 18:R531-583.
- Rodriguez-Larrea, D., and H. Bayley. 2014. Protein co-translocational unfolding depends on the direction of pulling. *Nat Commun*. 5.
- Safenkova, I.V., A.V. Zherdev, and B.B. Dzantiev. 2012. Application of atomic force microscopy for characteristics of single intermolecular interactions. *Biochemistry. Biokhimiia*. 77:1536-1552.
- Sambrook, J., and D.W. Russel. 2001. Molecular cloning: A laboratory manual. Third Ed., Cold Spring Harbor.
- Sarkar, A., S. Caamano, and J.M. Fernandez. 2007. The mechanical fingerprint of a parallel polyprotein dimer. *Biophysical journal*. 92:L36-38.
- Schaeffer, F., M. Matuschek, G. Guglielmi, I. Miras, P.M. Alzari, and P. Beguin. 2002. Duplicated dockerin subdomains of clostridium thermocellum endoglucanase celd bind to a cohesin domain of the scaffolding protein cipa with distinct thermodynamic parameters and a negative cooperativity. *Biochemistry*. 41:2106-2114.
- Schlierf, M., H. Li, and J.M. Fernandez. 2004. The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *Proceedings of the National Academy of Sciences of the United States of America*. 101:7299-7304.
- Schoeler, C., K.H. Malinowska, R.C. Bernardi, L.F. Milles, M.A. Jobst, E. Durner, W. Ott, D.B. Fried, E.A. Bayer, K. Schulten, H.E. Gaub, and M.A. Nash. 2014. Ultrastable cellulosome-adhesion complex tightens under load. *Nat Commun*. 5.
- Scholl, Z.N., W. Yang, and P.E. Marszalek. 2014. Chaperones rescue luciferase folding by separating its domains. *The Journal of biological chemistry*. 289:28607-28618.
- Schwartz, M.A., and D.W. DeSimone. 2008. Cell adhesion receptors in mechanotransduction. *Current opinion in cell biology*. 20:551-556.
- Schwesinger, F., R. Ros, T. Strunz, D. Anselmetti, H.J. Guntherodt, A. Honegger, L. Jermutus, L. Tiefenauer, and A. Pluckthun. 2000. Unbinding forces of single antibody-antigen complexes correlate with their thermal dissociation rates. *Proceedings of the National Academy of Sciences of the United States of America*. 97:9972-9977.
- Sen, M., R.A. Maillard, K. Nyquist, P. Rodriguez-Aliaga, S. Presse, A. Martin, and C. Bustamante. 2013. The clpxp protease unfolds substrates using a constant rate of pulling but different gears. *Cell*. 155:636-646.
- Sevier, C.S., and C.A. Kaiser. 2002. Formation and transfer of disulphide bonds in living cells. *Nat Rev Mol Cell Biol*. 3:836-847.
- Shapiro, L., and W.I. Weis. 2009. Structure and biochemistry of cadherins and catenins. *Cold Spring Harbor perspectives in biology*. 1:a003053.

- Sharma, D., Y. Cao, and H. Li. 2006. Engineering proteins with novel mechanical properties by recombination of protein fragments. *Angewandte Chemie*. 45:5633-5638.
- Sharma, D., O. Perisic, Q. Peng, Y. Cao, C. Lam, H. Lu, and H. Li. 2007. Single-molecule force spectroscopy reveals a mechanically stable protein fold and the rational tuning of its mechanical stability. *Proceedings of the National Academy of Sciences*. 104:9278-9283.
- Sivasankar, S., Y. Zhang, W.J. Nelson, and S. Chu. 2009. Characterizing the initial encounter complex in cadherin adhesion. *Structure (London, England : 1993)*. 17:1075-1081.
- Smith, D.E., S.J. Tans, S.B. Smith, S. Grimes, D.L. Anderson, and C. Bustamante. 2001. The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature*. 413:748-752.
- Stahl, S.W., M.A. Nash, D.B. Fried, M. Slutzki, Y. Barak, E.A. Bayer, and H.E. Gaub. 2012. Single-molecule dissection of the high-affinity cohesin-dockerin complex. *Proceedings of the National Academy of Sciences of the United States of America*. 109:20431-20436.
- Steward, A., J.L. Toca-Herrera, and J. Clarke. 2002. Versatile cloning system for construction of multimeric proteins for use in atomic force microscopy. *Protein science : a publication of the Protein Society*. 11:2179-2183.
- Thoren, K.L., E.J. Worden, J.M. Yassif, and B.A. Krantz. 2009. Lethal factor unfolding is the most force-dependent step of anthrax toxin translocation. *Proceedings of the National Academy of Sciences*. 106:21555-21560.
- Tsui, V., and D.A. Case. 2000. Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers*. 56:275-291.
- Twiss, F., and J. de Rooij. 2013. Cadherin mechanotransduction in tissue remodeling. *Cellular and molecular life sciences : CMLS*. 70:4101-4116.
- Twiss, F., Q. Le Duc, S. Van Der Horst, H. Tabdili, G. Van Der Krogt, N. Wang, H. Rehmann, S. Huvneers, D.E. Leckband, and J. De Rooij. 2012. Vinculin-dependent cadherin mechanosensing regulates efficient epithelial barrier formation. *Biology open*. 1:1128-1140.
- Tzima, E., M. Irani-Tehrani, W.B. Kiosses, E. Dejana, D.A. Schultz, B. Engelhardt, G. Cao, H. DeLisser, and M.A. Schwartz. 2005. A mechanosensory complex that mediates the endothelial cell response to fluid shear stress. *Nature*. 437:426-431.
- Valbuena, A., J. Oroz, R. Hervas, A.M. Vera, D. Rodriguez, M. Menendez, J.I. Sulkowska, M. Cieplak, and M. Carrion-Vazquez. 2009. On the remarkable mechanostability of scaffoldins and the mechanical clamp motif. *Proceedings of the National Academy of Sciences of the United States of America*. 106:13791-13796.
- Valbuena, A., J. Oroz, A.M. Vera, A. Gimeno, J. Gomez-Herrero, and M. Carrion-Vazquez. 2007. Quasi-simultaneous imaging/pulling analysis of single polyprotein molecules by atomic force microscopy. *The Review of scientific instruments*. 78:113707.
- Valbuena, A., Andrés M. Vera, J. Oroz, M. Menéndez, and M. Carrión-Vázquez. 2012. Mechanical properties of  $\beta$ -catenin revealed by single-molecule experiments. *Biophysical journal*. 103:1744-1752.
- Valle, F., G. Zuccheri, A. Bergia, L. Ayres, A.E. Rowan, R.J. Nolte, and B. Samori. 2008. A polymeric molecular "handle" for multiple afm-based single-molecule force measurements. *Angewandte Chemie*. 47:2431-2434.

- Varea, J., B. Monterroso, J.L. Saiz, C. Lopez-Zumel, J.L. Garcia, J. Laynez, P. Garcia, and M. Menendez. 2004. Structural and thermodynamic characterization of pal, a phage natural chimeric lysin active against pneumococci. *The Journal of biological chemistry*. 279:43697-43707.
- Vogel, V., and M. Sheetz. 2006. Local force and geometry sensing regulate cell functions. *Nat Rev Mol Cell Biol*. 7:265-275.
- Wang, M.D., M.J. Schnitzer, H. Yin, R. Landick, J. Gelles, and S.M. Block. 1998. Force and velocity measured for single molecules of rna polymerase. *Science (New York, N.Y.)*. 282:902-907.
- Wang, N., J.D. Tytell, and D.E. Ingber. 2009. Mechanotransduction at a distance: Mechanically coupling the extracellular matrix with the nucleus. *Nat Rev Mol Cell Biol*. 10:75-82.
- Weisel, J.W., H. Shuman, and R.I. Litvinov. 2003. Protein-protein unbinding induced by force: Single-molecule studies. *Current opinion in structural biology*. 13:227-235.
- Werbeck, N.D., and L.S. Itzhaki. 2007. Probing a moving target with a plastic unfolding intermediate of an ankyrin-repeat protein. *Proceedings of the National Academy of Sciences*. 104:7863-7868.
- Wu, C.Y., S.Y. Suen, S.C. Chen, and J.H. Tzeng. 2003. Analysis of protein adsorption on regenerated cellulose-based immobilized copper ion affinity membranes. *Journal of chromatography. A*. 996:53-70.
- Xing, Y., W.K. Clements, D. Kimelman, and W. Xu. 2003. Crystal structure of a beta-catenin/axin complex suggests a mechanism for the beta-catenin destruction complex. *Genes & development*. 17:2753-2764.
- Xing, Y., K. Takemaru, J. Liu, J.D. Berndt, J.J. Zheng, R.T. Moon, and W. Xu. 2008. Crystal structure of a full-length beta-catenin. *Structure (London, England : 1993)*. 16:478-487.
- Xu, W., and D. Kimelman. 2007. Mechanistic insights from structural studies of beta-catenin and its binding partners. *J Cell Sci*. 120:3337-3344.
- Yamada, S., S. Pokutta, F. Drees, W.I. Weis, and W.J. Nelson. 2005. Deconstructing the cadherin-catenin-actin complex. *Cell*. 123:889-901.
- Yang, G., C. Cecconi, W.A. Baase, I.R. Vetter, W.A. Breyer, J.A. Haack, B.W. Matthews, F.W. Dahlquist, and C. Bustamante. 2000. Solid-state synthesis and mechanical unfolding of polymers of t4 lysozyme. *Proceedings of the National Academy of Sciences*. 97:139-144.
- Yonemura, S. 2011. A mechanism of mechanotransduction at the cell-cell interface: Emergence of alpha-catenin as the center of a force-balancing mechanism for morphogenesis in multicellular organisms. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 33:732-736.
- Yonemura, S., Y. Wada, T. Watanabe, A. Nagafuchi, and M. Shibata. 2010. Alpha-catenin as a tension transducer that induces adherens junction development. *Nature cell biology*. 12:533-542.
- Zaidel-Bar, R. 2013. Cadherin adhesome at a glance. *J Cell Sci*. 126:373-378.
- Zhang, X., D.F. Bogorin, and V.T. Moy. 2004. Molecular basis of the dynamic strength of the sialyl lewis x--selectin interaction. *Chemphyschem : a European journal of chemical physics and physical chemistry*. 5:175-182.
- Zhang, Y., S. Sivasankar, W.J. Nelson, and S. Chu. 2009. Resolving cadherin interactions and binding cooperativity at the single-molecule level. *Proceedings of the National Academy of Sciences of the United States of America*. 106:109-114.



- Zhu, C., M. Long, S.E. Chesla, and P. Bongrand. 2002. Measuring receptor/ligand interaction at the single-bond level: Experimental and interpretative issues. *Annals of biomedical engineering*. 30:305-314.
- Zimmermann, J.L., T. Nicolaus, G. Neuert, and K. Blank. 2010. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nature protocols*. 5:975-985.

REVIEW OF SCIENTIFIC INSTRUMENTS 78, 113707 (2007)

**Quasi-simultaneous imaging/pulling analysis of single polyprotein molecules by atomic force microscopy**

Alejandro Valbuena, Javier Oroz, and Andrés Manuel Vera  
*Instituto Cajal, CSIC and Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), E-28002 Madrid, Spain*

Alejandro Gimeno  
*Nanotec Electrónica S.L., Centro Empresarial Euronova 3, Tres Cantos, E-28760 Madrid, Spain*

Julio Gómez-Herrero  
*Laboratorio Nuevas Microscopias, Departamento Física de la Materia Condensada, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*

Mariano Carrión-Vázquez<sup>a)</sup>  
*Instituto Cajal, CSIC and Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), E-28002 Madrid, Spain*

(Received 23 July 2007; accepted 15 September 2007; published online 13 November 2007)

Most of studies in protein nanomechanics have used the atomic force microscope (AFM) in its force-measuring mode on immobilized protein repeats (polyproteins) as single-molecule markers. Here, we add imaging capabilities to a standard, state-of-the-art AFM “puller” and integrate the most powerful programs of analysis available for both AFM modes. This unique instrument allows high-resolution, quasi-simultaneous imaging/force spectroscopy in aqueous solution. We demonstrate its capabilities using polyproteins of a model system (titin I27 domain). This tool should greatly facilitate the development of a much needed universal functionalization system for AFM, one that should allow better sample control and an improved efficiency of protein immobilization. © 2007 American Institute of Physics. [DOI: 10.1063/1.2794732]

**I. INTRODUCTION**

The atomic force microscope (AFM) is a relatively simple but powerful apparatus that allows the measurement of the mechanical properties of materials with atomic resolution.<sup>1</sup> In biology, the main application of this instrument has been the high-resolution surface imaging of biological structures in their aqueous native environment.<sup>2</sup> In addition to this, important mechanical information from the sample can be obtained using the AFM in the so-called force spectroscopy configuration,<sup>3</sup> which was exclusively designed to record force-extension curves and, accordingly, it tends to be based on a single piezoelectric positioner (i.e., *z* axis). State-of-the-art “force spectrometers” or “pullers,” usually custom-built specialized AFMs, are the ideal choice for studying single biomolecules as they are stretched in one dimension.<sup>4</sup>

In the so-called single-molecule force spectroscopy (SMFS) mode the AFM is capable of measuring forces with a sensitivity of tens of piconewtons and changes in length with nanometer resolution, with a time scale resolution in the millisecond range. These instruments have allowed the analysis of single-molecule unfolding and refolding events from modular proteins<sup>5</sup> and, with the use of recombinant polyproteins (which allow unequivocal single-molecule identification), from single protein modules<sup>6</sup> and nonmodular

proteins. As a result, SMFS has quickly spread so that a new field of inquiry, dubbed as “protein nanomechanics,” has emerged.<sup>4</sup> In contrast to this, in the subfield of force spectroscopy of intermolecular interactions to date there is not a single unambiguous strategy to identify single events, so several criteria have to be fulfilled at the same time in addition to the quantal analysis of the multimodal distribution.<sup>7,8</sup>

After the first pulling experiments on single proteins were done<sup>5,6</sup> SMFS has evolved into a second generation of instruments that incorporates a compact three-dimensional (3D) positioner with subnanometer resolution (~0.5 nm), constant-force mode, and powerful software developments such as semiautomatic mode of operation, automatic recognition of polyprotein periodical patterns, and specific protocols for protein unfolding and refolding. All these advanced features and open source functions have now been integrated into a specific software package based on IGOR PRO software, which is available as public domain.<sup>9</sup> Although the instrument itself is not yet commercial, all its components are so that several other groups have successfully setup this configuration already.<sup>10-13</sup>

Although quickly developing, protein nanomechanics is still in its infancy. Three main constraints limit its potential at present: the lack of knowledge of the initial state of the sample before the experiment, the fact that there is no real choice of the specific system been analyzed at each try (as the pickup of substrate-immobilized molecules is a random process), and the low efficiency of the pickup process. Upon

<sup>a)</sup> Author to whom correspondence should be addressed. Tel.: (91) 585-4830. FAX: (91) 585-4754. Electronic mail: mcarri@cajal.csic.es

# On the remarkable mechanostability of scaffoldins and the mechanical clamp motif

Alejandro Valbuena<sup>a</sup>, Javier Oroz<sup>a</sup>, Rubén Hervás<sup>a</sup>, Andrés Manuel Vera<sup>a</sup>, David Rodríguez<sup>a</sup>, Margarita Menéndez<sup>b</sup>, Joanna I. Sulkowska<sup>c</sup>, Marek Cieplak<sup>c</sup>, and Mariano Carrión-Vázquez<sup>a,1</sup>

<sup>a</sup>Instituto Cajal, Consejo Superior de Investigaciones Científicas and Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Avenida Doctor Arce 37, E-28002 Madrid, Spain; <sup>b</sup>Instituto de Química-Física Rocasolano, Consejo Superior de Investigaciones Científicas and Centro de Investigación Biomédica en Red sobre Enfermedades Respiratorias (CibeRes), Serrano 119, E-28006 Madrid, Spain; and <sup>c</sup>Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, 02668 Warsaw, Poland

Edited by Carlos J. Bustamante, University of California, Berkeley, CA, and approved June 30, 2009 (received for review December 22, 2008)

Protein mechanostability is a fundamental biological property that can only be measured by single-molecule manipulation techniques. Such studies have unveiled a variety of highly mechanostable modules (mainly of the Ig-like,  $\beta$ -sandwich type) in modular proteins subjected to mechanical stress from the cytoskeleton and the metazoan cell–cell interface. Their mechanostability is often attributed to a “mechanical clamp” of secondary structure (a patch of backbone hydrogen bonds) fastening their ends. Here we investigate the nanomechanics of scaffoldins, an important family of scaffolding proteins that assembles a variety of cellulases into the so-called cellulosome, a microbial extracellular nanomachine for cellulose adhesion and degradation. These proteins anchor the microbial cell to cellulose substrates, which makes their connecting region likely to be subjected to mechanical stress. By using single-molecule force spectroscopy based on atomic force microscopy, polyprotein engineering, and computer simulations, here we show that the cohesin I modules from the connecting region of cellulosome scaffoldins are the most robust mechanical proteins studied experimentally or predicted from the entire Protein Data Bank. The mechanostability of the cohesin modules studied correlates well with their mechanical kinetic stability but not with their thermal stability, and it is well predicted by computer simulations, even coarse-grained. This extraordinary mechanical stability is attributed to 2 mechanical clamps in tandem. Our findings provide the current upper limit of protein mechanostability and establish shear mechanical clamps as a general structural/functional motif widespread in proteins putatively subjected to mechanical stress. These data have important implications for the scaffoldin physiology and for protein design in biotechnology and nanotechnology.

cellulosome | cohesin | mechanical stability | protein nanomechanics | single-molecule force spectroscopy

**M**echanical forces are ubiquitous in biological systems. However, we have only recently learned how to experimentally control and measure this new biophysical parameter through single-molecule manipulation. Such studies have revealed that many proteins generate, sense, transmit, or are somehow subjected to mechanical forces during their normal functioning. A subset of these “mechanical proteins” is responsible for maintaining structural integrity in vivo and they display widely different mechanical stabilities, measured in vitro as the most probable unfolding force,  $F_u$  (after having been pulled apart, normally in the N-C direction). Such forces range from the pure entropic elasticity regime (e.g., elastomers like elastin) to a few tens of piconewtons (pN) (e.g., spectrin) and to slightly above 300 pN (titin immunoglobulin-Ig-domains) (1–5).

Mechanical proteins tend to be multimodular and the modules they are made of are often of the same type, although they commonly display distinct mechanical stabilities. These mechanical modules frequently seem to act as shock absorbers that protect critical biological interactions from mechanical stress, particularly in the metazoan cell–cell interface (3–5). Their

mechanostability is often attributed to a “mechanical clamp” of secondary structure that fastens their ends and, to date, Ig modules appear to be among the most mechanically stable ones.

Here we have investigated the nanomechanics of scaffoldins (Fig. 1), a family of microbial modular proteins that are particularly motivating from the mechanical point of view for 3 reasons: (i) Scaffoldins are noncatalytic structural proteins of the cellulosome, a multienzyme, cell-surface complex required for adhesion and degradation of crystalline cellulose, a particularly recalcitrant substrate (Fig. 1A). These scaffolding proteins act as a molecular Lego, binding a number of cellulases through its type I cohesin (cohesin I) modules to spatiotemporally regulate the efficiency of the entire enzymatic cascade (6). Furthermore, they physically anchor the microbial cell (6, 7) to the crystalline cellulose substrate (8), although the quaternary structure of the model cellulosomes, including their linkages to the cell and substrate, is highly polymorphic and not yet well resolved (Fig. 1A; ref. 6). The bonds joining the scaffoldin system are very tenacious, which include several of the highest affinity, noncovalent bonds known in nature [interestingly, one such bond in some scaffoldins is even covalent (9)], suggesting that they may be part of a stable mechanical circuit (Fig. 1A). Like in cell adhesion proteins from metazoans (5), we hypothesize that under physiological conditions, the scaffoldin region located between the 2 key attachment points (cell and substrate) might be exposed to a greater axial mechanical stress. We call this the “connecting region,” which is a segment putatively subjected to a more intense mechanical stress than the rest of the protein, referred to as the “hanging region”. (ii) The cohesin I modules are jellyroll domains with an Ig-like,  $\beta$ -sandwich topology. A simple inspection of their structure reveals a mechanical topology similar to that of titin Ig domains (Fig. 1B). However, the presence of longer putative mechanical clamps of  $\beta$ -sheet secondary structure in cohesins suggests that they may have a higher mechanical stability. (iii) Finally, it was suggested that a linear hydrodynamic flow could mechanically disrupt the integrity of cellulosomes releasing the bound cellulases (10), thus providing indirect evidence that the mechanical integrity of this extracellular complex is important for its function.

## Results and Discussion

We have studied 2 well-characterized model scaffoldins from motile Clostridia: CipA from *Clostridium thermocellum* (a ther-

Author contributions: M.C.-V. designed research; A.V., J.O., R.H., A.M.V., D.R., M.M., J.I.S., and M.C. performed research; A.V., M.M., J.I.S., M.C., and M.C.-V. analyzed data; and M.C.-V. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

<sup>1</sup>To whom correspondence should be addressed. E-mail: mcarri@cajal.csic.es.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0813093106/DCSupplemental](http://www.pnas.org/cgi/content/full/0813093106/DCSupplemental).

## Nanomechanics of the Cadherin Ectodomain "CANALIZATION" BY $Ca^{2+}$ BINDING RESULTS IN A NEW MECHANICAL ELEMENT\*

Received for publication, July 30, 2010, and in revised form, December 13, 2010. Published, JBC Papers in Press, December 22, 2010, DOI 10.1074/jbc.M110.170399

Javier Oroz<sup>1,2</sup>, Alejandro Valbuena<sup>1,1</sup>, Andrés Manuel Vera<sup>1,2</sup>, Jesús Mendieta<sup>1,3</sup>, Paulino Gómez-Puertas<sup>5</sup>, and Mariano Carrón-Vázquez<sup>1,4</sup>

From the <sup>1</sup>Instituto Cajal/Consejo Superior de Investigaciones Científicas (CSIC), Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), and Instituto Madrileño de Estudios Avanzados (IMDEA) Nanociencia, Avenida Doctor Arce 37, E-28002 Madrid, Spain, <sup>2</sup>Centro de Biología Molecular Severo Ochoa, CSIC-Universidad Autónoma de Madrid, E-28049 Madrid, Spain, and <sup>3</sup>Biomol-Informatics SL, E-28049 Madrid, Spain

Cadherins form a large family of calcium-dependent cell-cell adhesion receptors involved in development, morphogenesis, synaptogenesis, differentiation, and carcinogenesis through signal mechanotransduction using an adaptor complex that connects them to the cytoskeleton. However, the molecular mechanisms underlying mechanotransduction through cadherins remain unknown, although their extracellular region (ectodomain) is thought to be critical in this process. By single molecule force spectroscopy, molecular dynamics simulations, and protein engineering, here we have directly examined the nanomechanics of the C-cadherin ectodomain and found it to be strongly dependent on the calcium concentration. In the presence of calcium, the ectodomain extends through a defined ("canalized") pathway that involves two mechanical resistance elements: a mechanical clamp from the cadherin domains and a novel mechanostable component from the interdomain calcium-binding regions ("calcium rivet") that is abolished by magnesium replacement and in a mutant intended to impede calcium coordination. By contrast, in the absence of calcium, the mechanical response of the ectodomain becomes largely "decanalized" and destabilized. The cadherin ectodomain may therefore behave as a calcium-switched "mechanical antenna" with very different mechanical responses depending on calcium concentration (which would affect its mechanical integrity and force transmission capability). The versatile mechanical design of the cadherin ectodomain and its dependence on extracellular calcium facilitate a variety of mechanical responses that, we hypothesize, could influence the various adhesive properties mediated by cadherins in tissue morphogenesis, synaptic plasticity, and disease. Our work represents the first step toward the mechanical characterization of the cadherin system, opening the door to understanding the mechanical bases of its mechanotransduction.

The acquisition of specific machinery for cell adhesion represented a crucial step in the evolution of metazoan organisms, enabling multicellularity and the development of tissues. There are four major classes of cell adhesion receptors: the immunoglobulin (Ig)-like superfamily, the cadherin superfamily, selectins, and integrins (1). Because of their location at the cell-cell interface, these receptors are often subject to mechanical stress (2, 3). However, to date only the mechanical properties of Ig and fibronectin type III domains have been investigated using atomic force microscopy (AFM)<sup>3</sup>-based single molecule force spectroscopy (SMFS) and protein engineering (4).

Cadherins can be defined (with some exceptions) as a superfamily of calcium ( $Ca^{2+}$ )-dependent homophilic cell-cell adhesion proteins that mediate most cell adhesion in vertebrates, playing a crucial role in morphogenesis, synaptogenesis, tissue architecture, tissue repair, and carcinogenesis (5–7). Still, certain cadherins are heterophilic or/and connect two membranes from the same cell. At least in the case of the so-called "classical" cadherins (6, 8), their connection to the cytoskeleton appears to be mediated by several adaptor proteins (Refs. 9–11 and see Fig. 1A). This linkage is responsible for reinforcing cell-cell adhesion through a signaling pathway involving the activation of Rho, Rac, Cdc42, and the Arp2/3 actin nucleator complex (12–15). External mechanical stimuli are among the signals transduced by cadherins, which can induce changes in cell shape and synaptic plasticity (6, 16–19), triggering chemical or electrical (20) signaling. Furthermore, cadherin interactions are known to exert forces that have already been measured (21).

Classical cadherins are single pass transmembrane glycoproteins of 720–750 amino acids with an extracellular region (ectodomain or EC) that has a rodlike structure composed of five autonomously folded tandem domains (extracellular cadherin domains EC<sub>1</sub>–EC<sub>5</sub> from the distal N terminus to the proximal C terminus; see Fig. 1, A and B). Although it is accepted that these cadherins establish homophilic interactions in *cis* and *trans* (22), the quaternary structure of these complexes remains unclear. The resolution of the atomic struc-

\* This work was supported in part by grants from the Ministerio de Ciencia e Innovación, MICINN (Grant BIO2007-671116), Consejería de Educación de la Comunidad de Madrid, CECM (Grant S-0505/MAT/0283), and CSIC (Grant 200620F00) (to M. C.-V.) and from the MICINN (Grant SAF2007-61926), the CECM (Grant S-BIO-0260/2006-COMBACT), and the European Union (Grant FP7-223431) (to P. G.-P.).

<sup>1</sup> Both authors contributed equally to this work.

<sup>2</sup> Recipient of a fellowship from the CECM.

<sup>3</sup> Recipient of a fellowship from the Fundación Areces.

<sup>4</sup> To whom correspondence should be addressed. Tel.: 34-915854830; Fax: 34-915854754; E-mail: mcarron@icaj.csic.es.

<sup>5</sup> The abbreviations used are: AFM, atomic force microscopy; SMFS, single molecule force spectroscopy; EC, ectodomain of cadherins; MD, molecular dynamics; SMD, steered MD;  $\Delta L$ , increase in contour length; WLC, wormlike chain; TM, triple mutant; pN, piconewtons; r.m.s.d., root mean square deviation.





# Common Features at the Start of the Neurodegeneration Cascade

Rubén Hervás<sup>1,2,3</sup>, Javier Oroz<sup>1,2,3</sup>, Albert Galera-Prat<sup>1,2</sup>, Oscar Goñi<sup>1,2</sup>, Alejandro Valbuena<sup>1,2</sup>, Andrés M. Vera<sup>1,2</sup>, Ángel Gómez-Sicilia<sup>1,2</sup>, Fernando Losada-Urzáiz<sup>1,2</sup>, Vladimir N. Uversky<sup>3,4</sup>, Margarita Menéndez<sup>5</sup>, Douglas V. Laurents<sup>6</sup>, Marta Bruix<sup>6</sup>, Mariano Carrión-Vázquez<sup>1,2\*</sup>

**1** Instituto Cajal, IC-CSIC & Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, Spain, **2** Instituto Madrileño de Estudios Avanzados en Nanociencia (IMDEA-Nanociencia), Madrid, Spain, **3** University of South Florida, College of Medicine and Byrd Alzheimer's Research Institute, Tampa, Florida, United States of America, **4** Institute for Biological Instrumentation, Russian Academy of Sciences, Pushchino, Moscow Region, Russia, **5** Instituto de Química-Física Rocasolano, IQFR-CSIC & Centro de Investigación Biomédica en Red sobre Enfermedades Respiratorias (CIBERES), Madrid, Spain, **6** Instituto de Química-Física Rocasolano, IQFR-CSIC, Madrid, Spain

## Abstract

Amyloidogenic neurodegenerative diseases are incurable conditions with high social impact that are typically caused by specific, largely disordered proteins. However, the underlying molecular mechanism remains elusive to established techniques. A favored hypothesis postulates that a critical conformational change in the monomer (an ideal therapeutic target) in these “neurotoxic proteins” triggers the pathogenic cascade. We use force spectroscopy and a novel methodology for unequivocal single-molecule identification to demonstrate a rich conformational polymorphism in the monomer of four representative neurotoxic proteins. This polymorphism strongly correlates with amyloidogenesis and neurotoxicity: it is absent in a fibrillization-incompetent mutant, favored by familial-disease mutations and diminished by a surprisingly promiscuous inhibitor of the critical monomeric  $\beta$ -conformational change, neurotoxicity, and neurodegeneration. Hence, we postulate that specific mechanostable conformers are the cause of these diseases, representing important new early-diagnostic and therapeutic targets. The demonstrated ability to inhibit the conformational heterogeneity of these proteins by a single pharmacological agent reveals common features in the monomer and suggests a common pathway to diagnose, prevent, halt, or reverse multiple neurodegenerative diseases.

**Citation:** Hervás R, Oroz J, Galera-Prat A, Goñi O, Valbuena A, et al. (2012) Common Features at the Start of the Neurodegeneration Cascade. *PLoS Biol* 10(5): e1001335. doi:10.1371/journal.pbio.1001335

**Academic Editor:** Jonathan S. Weissman, University of California San Francisco/Howard Hughes Medical Institute, United States of America

**Received:** October 18, 2011; **Accepted:** April 6, 2012; **Published:** May 29, 2012

**Copyright:** © 2012 Hervás et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was funded by grants from the Ministerio de Ciencia e Innovación (MICINN-BIO2007-67116), the Consejería de Educación de la Comunidad de Madrid (S-0505/MAT/0283), and the Consejo Superior de Investigaciones Científicas (CSIC-200620F00) to M.C.-V., and from the MICINN to D.V.L. (CTQ2010-21567-C02-02), to M.B. (CTQ2008-00080/BQU), and to M.M. (BFU2009-10052). R.H. and A.G.-P. are recipients of fellowships from the Fundación Ferrer (Severo Ochoa fellowships), while J.O., A.M.V., and A.G.-S. received fellowships from the Consejería de Educación de la Comunidad de Madrid, the Fundación Areces and the CSIC, respectively. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors declare a competing financial interest: M.C.-V., J.O., and R.H. are co-inventors on an international patent application (No. P 201031846/PCT/ES2011/070867) covering the results contained in this article. Any potential income generated by exploitation of the patent rights and received by their employer, the CSIC, shall be shared with these authors according to Spanish law and the regulations of the CSIC.

**Abbreviations:**  $\Delta L_c$ , contour length increase; AFM, atomic force microscopy; Arc, arctic; CD, circular dichroism; DSC, differential scanning calorimetry; F, unfolding force; hM, hyper-mechanostable; IDP, intrinsically disordered protein; M, mechanostable; MCS, multicloning site; NM, non-mechanostable; NMR, nuclear magnetic resonance; NP, neurotoxic protein; pFS, plasmid for force spectroscopy; QBP1, polyQ-binding peptide 1; RC, random coil; SCA, spinocerebellar ataxia; SDS, sodium dodecyl sulfate; SMFS, single-molecule force spectroscopy; TEM, transmission electron microscopy; WLC, worm-like chain; wt, wild-type

\* E-mail: mcarrion@cajal.csic.es

These authors contributed equally to this work.

## Introduction

Amyloid-related neurodegenerative diseases are incurable disorders that are currently classified as conformational diseases (a subset of the proteinopathies). These diseases are thought to be primarily caused by an unknown conformational change in specific proteins, termed “neurotoxic proteins” (NPs), which confers them a cytotoxic gain of function. This initial transition in the soluble monomer triggers a series of events that lead to the formation of small soluble oligomers and protofibrils, and ultimately the generation of highly ordered fibrillar insoluble aggregates. More than 20 of these NPs have been found to be intrinsically disordered proteins (IDPs). Despite the fact that the monomers of NPs have different primary and, when folded, tertiary structures, their oligomers and aggregates

share common structural properties [1,2], including recognition by a conformational antibody [3]. Whereas the nature of the cytotoxic species and the molecular mechanisms involved in their misfolding and aggregation remain largely unknown, there is growing evidence that conformational changes preceding aggregation are fundamental in the molecular pathogenesis of these diseases [1]. Furthermore, as the formation of monomeric misfolding intermediates is regarded as the earliest event in the pathogenic cascade, this process represents an ideal therapeutic target [4].

Remarkably, it was recently demonstrated that the key conformational change that endows amyloidogenesis occurs in the monomer, triggering aggregation and fibrillogenesis as well as cytotoxicity [5–7]. Thus, to understand the acquisition of toxic properties by NPs, the three-dimensional structure of this toxic

## Mechanical Properties of $\beta$ -Catenin Revealed by Single-Molecule Experiments

Alejandro Valbuena,<sup>†</sup> Andrés Manuel Vera,<sup>†</sup> Javier Oroz,<sup>†</sup> Margarita Menéndez,<sup>‡</sup> and Mariano Carrión-Vázquez<sup>†\*</sup>

<sup>†</sup>Instituto Cajal/CSIC, Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED) and IMDEA Nanociencia, Madrid, Spain; and <sup>‡</sup>Instituto de Química-Física Rocasolano, CSIC and Centro de Investigación Biomédica en Red sobre Enfermedades Respiratorias (CIBERES), Madrid, Spain

**ABSTRACT**  $\beta$ -catenin is a central component of the adaptor complex that links cadherins to the actin cytoskeleton in adherens junctions and thus, it is a good candidate to sense and transmit mechanical forces to trigger specific changes inside the cell. To fully understand its molecular physiology, we must first investigate its mechanical role in mechanotransduction within the cadherin system. We have studied the mechanical response of  $\beta$ -catenin to stretching using single-molecule force spectroscopy and molecular dynamics. Unlike most proteins analyzed to date, which have a fixed mechanical unfolding pathway, the  $\beta$ -catenin armadillo repeat region (ARM) displays low mechanostability and multiple alternative unfolding pathways that seem to be modulated by its unstructured termini. These results are supported by steered molecular dynamics simulations, which also predict its mechanical stabilization and unfolding pathway restrictions when the contiguous  $\alpha$ -helix of the C-terminal unstructured region is included. Furthermore, simulations of the ARM/E-cadherin cytosolic tail complex emulating the most probable stress geometry occurring *in vivo* show a mechanical stabilization of the interaction whose magnitude correlates with the length of the stretch of the cadherin cytosolic tail that is in contact with the ARM region.

### INTRODUCTION

The importance of mechanical force in regulating intracellular signaling and gene expression during embryonic development, tissue morphogenesis, and cell differentiation in normal physiology as well as in pathological conditions like cancer is well established (1). However, the molecular mechanisms by which cells sense and respond to mechanical stress through intracellular biochemical changes remain unclear. Cell adhesion receptors like cadherins and integrins are key components of this mechanotransduction machinery because they fulfill three different functions: cell-to-cell and cell-to-extracellular matrix adhesion; connecting the membrane to the cytoskeleton inside the cell; and the regulation of intracellular signaling pathways, including gene expression (1). The nanomechanical properties of the talin rod, a key component of the adaptor complex that links the extracellular matrix to the cytoskeleton through integrins, have been recently characterized (2). This structure was found to display complex mechanical features, including the exposure of cryptic binding sites to vinculin upon stretching, advancing our understanding of the mechanisms of force transduction within this pathway.

Initiation of cell-cell contacts requires interactions through cadherins, an important class of cell-cell adhesion receptors (3) linked to the cortical actin cytoskeleton through connecting proteins.  $\beta$ -catenin is a key component of the plasma membrane-cytoskeleton chain of the cadherin adhesion system (Fig. 1 A), being a multifunctional protein that acts both as a structural adaptor protein and as a tran-

scription cofactor translocating to the nucleus in the Wnt signaling pathway (4). Thus,  $\beta$ -catenin has the potential to coordinate changes in gene expression with dynamic changes in cell adhesion and migration. Indeed, its deregulation can lead to alterations in cell fate, adhesion, and migration, provoking the development of different types of cancer (4). Furthermore, this protein has been recently related to neurodegeneration (5).

Mechanotransduction often involves conformational changes in the protein domains that are subjected to tension (1,2). The cadherin-catenin contact is known to be subjected to mechanical stress (3,6), and a molecular clutch model has been proposed as the mechanism to connect the system to the cytoskeleton (7). Thus, it seems likely that catenins, the main component in the adaptor complex of these contacts, are good candidates to sense and/or transmit mechanical forces to the cell interior to trigger biochemical changes. Significantly, it has been proposed that tension may regulate the fraction of  $\beta$ -catenin that is either bound to cell-cell junctions or free in the cytosol and, in doing so, it could stabilize the cell-cell contacts and control downstream signaling (8). Because catenins connect cadherins to the actin cytoskeleton, knowing how these central proteins respond to force is of paramount importance to understand the basis of these mechanotransduction processes.

$\beta$ -catenin binds to the cytoplasmic domain of cadherins (9) and  $\alpha$ -catenin in a ternary complex (10), and the exact chain of components that physically links them to the actin cytoskeleton has been solved (11). Like many proteins with a putative mechanical function (12),  $\beta$ -catenin is modular, containing 12 characteristic armadillo (ARM) repeats in its central region (R1 to R12, from N- to C-termini) that

Submitted December 15, 2011, and accepted for publication July 17, 2012.

\*Correspondence: mcarrion@cajal.csic.es

Editor: Doug Barrick.

© 2012 by the Biophysical Society  
0006-3495/12/10/1744/9 \$2.00

<http://dx.doi.org/10.1016/j.bpj.2012.07.051>