



Universitat Autònoma de Barcelona

Institut de Biotecnologia i Biomedicina

Parc científic de Barcelona

Instituto de Biología Molecular de Barcelona - CSIC

Institut de Recerca Biomèdica de Barcelona

**Estructura tridimensional
del factor de terminación mitocondrial
humano, mTERF en complejo con DNA**

TESI DOCTORAL

Nereida Jiménez Menéndez

Directora i codirector de tesi

Maria Solà Vilarrubia i Miquel Coll

Barcelona, Desembre de 2010

Para Andrés y Rosa,
Estela y Mimi.

Pel Pep.

El Secreto

Kalámi dondára. Mindrádon gundára.

Los dioses poseían el secreto del agua y los humanos se lo robaron.
Los dioses guardaban el secreto del aire y los humanos se lo robaron.
Los dioses tenían escondido el secreto de la tierra y los humanos se lo robaron.
Los dioses mantenían oculto el secreto del fuego y los humanos se lo robaron.
Y fueron creciendo, los humanos, y se fueron acercando a los dioses y comenzaron a desafiarlos.
A los dioses sólo les quedaba un secreto, el pequeño secreto, el secreto insignificante, el secreto de lo humano.
Comenzaron los humanos a investigar, a buscar su secreto, el secreto de sus vidas. Los dioses, temiendo que pudieran robárselo, decidieron esconderlo donde los humanos no pudieran encontrarlo. Le confiaron la tarea a Kala... una diosa traviesa y astuta que gustaba de jugar con los humanos; los conocía en su corazón y en su inteligencia.
Kala pensó primero esconder el secreto en el cielo:
–*Nos cognetum ka antra nemur e nade undo no maresatran.* (Los humanos no pueden volar y allá nunca lo encontrarán.)
Pero luego se dijo:
–*Skara nos cognetum dor-o coroe nos kaerte na kestoá e der o nimar o nos extormis kon tomedio.* (Un día los humanos van a develar los misterios del espacio y van a llegar a los confines del universo.)
Tuvo la idea de esconder el secreto en las profundidades del océano, pero de nuevo su intuición le advirtió:
–*Skara nos cognetum dor e corangar na akéadam e dor o nasupar nos cártane na-konesfastedal.*
(Un día los humanos van a agotar el océano y van a superar los límites de la profundidad.)
Decidió esconder el secreto de la vida en el corazón de la tierra. En el momento de hacerlo, una vez más, la asaltaron las dudas:
–*Na skodenu na sun tarkta, na dentonkétos, na casfintélos den maren nanka ki minde minde na la maira, nanka en kar máнкеle e lenfentema en ko tapelía...* (Los humanos son tan tercos, tan obstinados, tan perspicaces que irán hasta el centro mismo de la tierra, hasta lo más ínfimo y diminuto de la materia.)
La diosa Kala pensó y dudó, dudó y pensó; y un día, siempre hay un día, encontró la solución; visitó a los humanos, los exploró y escondió el secreto allá, donde nunca podrían encontrarlo, donde nunca se atreverían a buscarlo: en el interior mismo de lo humano. Tomó el secreto, lo partió y lo escondió, repartido, pedacito a pedacito, en cada uno de los humanos...
Y los humanos seguimos buscando nuestro secreto.

(Inspirado en un motivo mítico de la tradición oral de la India)
Nicolás Buenaventura Vidal

Lo importante en ciencia no es tanto obtener nuevos hechos como descubrir nuevas formas de pensar sobre ellos.

William Lawrence Bragg

Prefacio

La mitocondria es, a mi entender, un orgánulo fascinante y muy importante energética y metabólicamente del que, a pesar de los grandes avances en el conocimiento sobre su morfología, dinámica y composición y su función metabólica, energética y apoptótica, aún quedan muchas incógnitas por resolver.

El trabajo que se muestra a continuación consiste en la determinación estructural, en complejo con DNA, de mTERF, un factor de terminación de la transcripción mitocondrial, cuyo papel en la mitocondria es probablemente más complicado, ya que parece estar también implicando en la estimulación de la transcripción y en la regulación de la replicación.

En la introducción he querido recoger alguna información general sobre mitocondria a partir de diversas fuentes para poder tener una visión general del marco en que se sitúa la proteína objeto de estudio. Esta pequeña incursión en la mitocondria muestra sólo una pequeña parte del notable esfuerzo científico que se está haciendo a muchos niveles para entender este orgánulo.

Por tratarse de una tesis en cristalografía, los materiales y métodos están básicamente enfocados a describir esta parte, explicando con detalle todos los puntos relacionados con la técnica y con las estrategias que hemos seguido para resolver los problemas cristalográficos. Son esencialmente para poder revisar ciertos aspectos, sobre todo de la parte cristalográfica más teórica, que han quedado relegados a fuerza de las prisas que exigen las publicaciones.

Debido a los problemas que han surgido durante el análisis cristalográfico del complejo proteína-DNA han sido necesarias otras técnicas experimentales, lo que muestra la importancia de combinar la información obtenida mediante diversas técnicas en el estudio de cualquier sistema.

Índice

RESUMEN.....	7
INTRODUCCIÓN.....	9
1. LA MITOCONDRIA.....	11
1.1. Morfología y dinámica de la mitocondria.....	13
1.2. Origen evolutivo de la mitocondria.....	16
1.3. Funciones mitocondriales.....	19
1.3.1. Vías metabólicas.....	19
1.3.2. Cadena de transporte electrónico. Fosforilación oxidativa.....	21
1.3.3. Implicación de la mitocondria en muerte celular.....	25
1.4. Disfunciones y enfermedades mitocondriales.....	27
1.4.1. Enfermedades debidas a alteraciones en el DNAn.....	28
1.4.2. Enfermedades debidas a alteraciones en el DNAm.....	28
1.5. El genoma mitocondrial humano.....	30
1.5.1. Replicación del DNAm humano.....	34
1.5.2. Transcripción del DNAm humano.....	38
1.5.3. Traducción del DNAm humano.....	48
2. TERMINACIÓN DE LA TRANSCRIPCIÓN.....	51
2.1. Factores terminadores de la transcripción mitocondrial.....	53
2.1.1. mTERF.....	53
2.1.2. Factores terminadores de la transcripción mitocondrial en invertebrados.....	58
3. FAMILIA mTERF.....	63
4. PROTEÍNAS HELICAL TANDEM REPEAT.....	66
OBJETIVOS.....	71

RESULTADOS Y DISCUSIÓN	73
1. OBTENCIÓN DE LA PROTEÍNA. PURIFICACIÓN Y CARACTERIZACIÓN	75
1.1. Análisis informático y clonaje de la proteína.....	75
1.2. Pruebas de expresión y solubilidad.....	76
1.3. Purificación de mTERF.....	78
1.4. Optimización del tampón proteico por thermofluor.....	80
1.5. Degradación de mTERF.....	83
1.6. Ensayos de proteólisis limitada.....	85
2. OBTENCIÓN Y CARACTERIZACIÓN DEL COMPLEJO mTERF-DNA	87
2.1. Ensayos de unión a DNA.....	87
2.2. Obtención del complejo para la cristalización.....	90
3. CRISTALIZACIÓN DEL COMPLEJO	91
3.1. Confirmación de la presencia de DNA en los cristales.....	94
4. PROCESAMIENTO DE DATOS Y RESOLUCIÓN DE LA ESTRUCTURA	96
5. ESTRUCTURA TRIDIMENSIONAL	101
5.1. Estructura del complejo mTERF y DNA.....	101
5.2. Estructura del complejo mTERF- Δ N y DNA.....	108
6. ESTUDIO EN SOLUCIÓN DEL COMPLEJO mTERF-DNA	111
6.1. Complejo mTERF y DNA _{TER} en solución.....	113
6.2. Complejo mTERF- Δ N y DNA _{TER} en solución.....	115
7. ENSAYOS DE UNIÓN Y ESPECIFICIDAD DE mTERF	116
8. COMPARACIÓN DE ESTRUCTURAS	120
8.1. Comparación de las estructuras mTERF-DNA15pb y mTERF-DNA22pb.....	120
 CONCLUSIONES	 129
 MATERIALES, MÉTODOS Y FUNDAMENTOS	 133
1. CLONAJE	135

2. EXPRESIÓN Y PURIFICACIÓN DE LOS COSTRUCTOS	
DE mTERF	138
2.1. Optimización de la expresión bacteriana y del tampón de solubilización....	138
2.2. Métodos de purificación.....	144
2.3. Expresión y purificación de proteína a gran escala.....	146
2.4. Caracterización de la proteína.....	147
2.5. Expresión y purificación con selenometionina.....	152
2.6. Mutaciones puntuales por Met.....	155
3. FORMACIÓN DEL COMPLEJO CON DNA	156
3.1. Ensayos de unión a DNA para la cristalización.....	156
3.2. Obtención del complejo para la cristalización.....	157
3.3. Análisis y caracterización del complejo.....	158
3.4. Ensayos de especificidad.....	159
4. CRISTALIZACIÓN	160
4.1. Optimización de los cristales.....	163
4.2. Análisis de la presencia de DNA en los cristales.....	167
4.3. Derivatización de los cristales con átomos pesados.....	169
4.4. Crioprotección y montaje de los cristales.....	171
5. DIFRACCIÓN DE RAYOS X Y RESOLUCIÓN DE LA ESTRUCTURA	173
5.1. Difracción de rayos X.....	173
5.2. Resolución de la estructura.....	177
5.3. Construcción del modelo atómico y refinado.....	195
5.4. Validación del modelo estructural.....	197
6. DISPERSIÓN A BAJO ÁNGULO	200
6.1. Preparación y medida de las muestras de mTERF y el complejo con DNA.	204
6.2. Procesamiento.....	205
BIBLIOGRAFIA	209
Siglas y acrónimos	225

Resumen

El genoma mitocondrial humano es una molécula circular de doble cadena localizada en la matriz mitocondrial. La información genética se encuentra muy compactada y el DNAm_t está completamente saturado de genes sin intrones. La transcripción del DNAm_t se inicia a partir de tres promotores diferentes localizados en el D-loop, uno para la cadena ligera (L) y dos para la cadena pesada (H₁ y H₂) a partir de los cuales se transcriben la mayoría de genes. Las unidades de transcripción que se inician en H₂ y en L producen una especie policistrónica gigante que cubre casi toda la cadena. En cambio, para la cadena pesada la transcripción iniciada en H₁ se termina en el extremo 3' del rRNA 16S, y es la responsable de la síntesis de los rRNAs, y de los tRNA^{Phe} y tRNA^{Val}. El papel principal de esta atenuación se debe al factor de terminación de la transcripción, mTERF. mTERF, es una proteína codificada en núcleo, cuya forma madura tiene 342 aminoácidos. Se ha descrito que se une a varios lugares estratégicos del DNA mitocondrial, donde controla procesos como la terminación de la transcripción o la pausa de la maquinaria de replicación mitocondrial. De estos lugares, el que presenta mayor afinidad es la secuencia de 28 pares de bases (DNA_{28pb}) en el gen del tRNA^{Leu}(UUR) mitocondrial. Ensayos *in vitro* han demostrado que la unión de mTERF a esta región provoca la terminación de la transcripción del DNA mitocondrial en ambos sentidos.

En este trabajo se estudia, desde un punto de vista estructural y mediante técnicas de cristalografía de proteínas, el complejo de mTERF con el DNA que contiene la secuencia de terminación. Durante la producción de mTERF recombinante, la proteólisis espontánea de la forma entera de la proteína (Arg56-Ala399) dio lugar a una variante corta, mTERF-ΔN (Arg99-Ala399) que mantenía la unión al DNA. Se determinó la estructura tridimensional de mTERF-ΔN a 2.4 Å de resolución a partir de cristales derivados con Se-Met y ésta se usó para resolver la estructura de mTERF entera a 3.1 Å de resolución. Ambas cristalizaron únicamente en presencia de DNA.

La disposición del DNA en el eje z del cristal provocó que los cristales de mTERF- Δ N en complejo con el DNA29pb, con un parámetro c de la celda unidad demasiado pequeño para albergar las 29pb del DNA que contenían, presentaran graves problemas de anisotropía y desorden, dando lugar a mapas de densidad electrónica muy distorsionados. Se realizaron entonces ensayos de cristalización con oligonucleótidos de menor longitud, pudiéndose trazar y refinar el DNA y la proteína completa con los cristales de mTERF- Δ N en complejo con el DNA12pb. En los cristales de mTERF entera en complejo con el DNA15pb, la proteína interacciona mediante los subdominio N-terminal y C-terminal con dos moléculas de DNA simétricas que no interaccionan entre ellas y cuya posición relativa viene dada por un ángulo de $\sim 36^\circ$. mTERF se une a la largo del surco mayor del DNA, estableciendo contactos inespecíficos con los fosfatos mediante 19 residuos y contactos con bases nitrogenadas mediante 5 residuos. La estructura cristalográfica de mTERF consiste en nueve repeticiones estructurales, de TERF-I a IX, flanqueadas por un segmento N-terminal y una hélice α C-terminal. Cada motivo TERF está formado por unos 35 aminoácidos que se estructuran en tres hélices α (H1, H2 y H3) formando una superhélice trigonal levógira con un núcleo hidrofóbico central. La combinación de la conectividad en superhélice levógira con la propagación en tándem del motivo para formar un solenoide es única de las proteínas de la familia MTERF.

En base a la estructura cristalográfica se generaron modelos para estudiar los complejos de mTERF y mTERF- Δ N con el oligonucleótido de DNA28pb en solución, mediante la técnica de SAXS. La representación gráfica del error asociado al ajuste entre la curva teórica de dispersión para cada modelo y la curva medida experimentalmente, presenta un mínimo, indicando que la unión de mTERF al DNA28pb se da preferentemente sobre una zona del DNA. Los ensayos de unión y especificidad de mTERF y las formas truncadas mTERF- Δ N, mTERF- Δ C y Δ N-mTERF- Δ C con el DNA28pb mediante geles nativos de retardo muestran que mTERF- Δ C y Δ N-mTERF- Δ C no unen DNA y que mTERF y mTERF- Δ N lo unen específicamente.

INTRODUCCIÓN



1_ LA MITOCONDRIA

Historia

Las mitocondrias se observaron de forma independiente por varios citólogos en el período comprendido entre 1850 y 1880. Probablemente las primeras observaciones se deben al botánico suizo Kolliker quien anotó la presencia de unos gránulos entre las miofibrillas del músculo estriado a los que denominó *sarcosomas*. En 1882, el alemán Walther Flemming descubrió unas inclusiones a las que denominó *fila*. En 1894, Richard Altmann en su obra *Die Elementarorganismen* describe una serie de corpúsculos que observa mediante una tinción especial con fucsina en casi todos los tipos de células. Especula que se trata de un tipo de “partículas elementales de vida” y los denomina *bioblastos*. El nombre de "mitocondria" (del griego mítos: hilo, y kóndros: gránulo), se debe a Carl Benda, quien en 1898 denominó así a unos gránulos que, en tinciones de violeta cristal y alizarina, se distinguían por su brillo. Casi una década y media después, los biólogos celulares BF Kingsbury y Otto Warburg independientemente, sugieren que las mitocondrias juegan un papel en la respiración celular, en la que, mediante reacciones de oxidación-reducción, la energía de los alimentos pasa a estar disponible para el metabolismo celular. En los primeros estudios con células vivas en cultivo de tejidos, Lewis en 1914 y Strangeways y Canti en 1921, observaron que las mitocondrias eran estructuras muy plásticas que continuamente ejercían movimientos lentos y a veces, experimentaban notables cambios de forma. La separación de las mitocondrias por centrifugación diferencial de homogeneizados de células se intentó por primera vez con cierto éxito por Bensley y Hoerr en 1934. El método fue perfeccionado aún más por Claude principios de los años 1940 y Hogeboom, Schneider, y Palade en 1948. En los años siguientes, los estudios de mitocondrias aisladas realizados, entre otros investigadores, por Kennedy y Lehninger (1949) y por Green y col., permitieron establecer su composición química, y confirmaron que la mitocondria era el lugar donde se daban mayoritariamente las reacciones de oxidación-reducción. La presencia del DNA mitocondrial fue descubierta por Margit M. K. Nass y Sylvan Nass en 1963.

La mitocondria

La mitocondria es un orgánulo de doble membrana que se encuentra en casi la totalidad de las células eucariotas. Su principal actividad es la fosforilación oxidativa que da lugar a la generación de ATP, usado por la célula como fuente de energía química. Además, la mitocondria está implicada en otros procesos tales como la síntesis de esteroides y lípidos, la descarboxilación oxidativa del piruvato, el ciclo de los ácidos tricarboxílicos y la oxidación de ácidos grasos. Ciertas enzimas del ciclo de la urea y de la gluconeogénesis están localizadas en la matriz mitocondrial. La mitocondria está también involucrada en la regeneración de NAD^+ y en la homeostasis intracelular de iones como el calcio y el fosfato. En los últimos años, la mitocondria ha sido el objeto central de numerosos estudios debido a su papel como sensor y ejecutor de la apoptosis (Green & Reed, 1998; Ravagnan *et al.*, 2002).

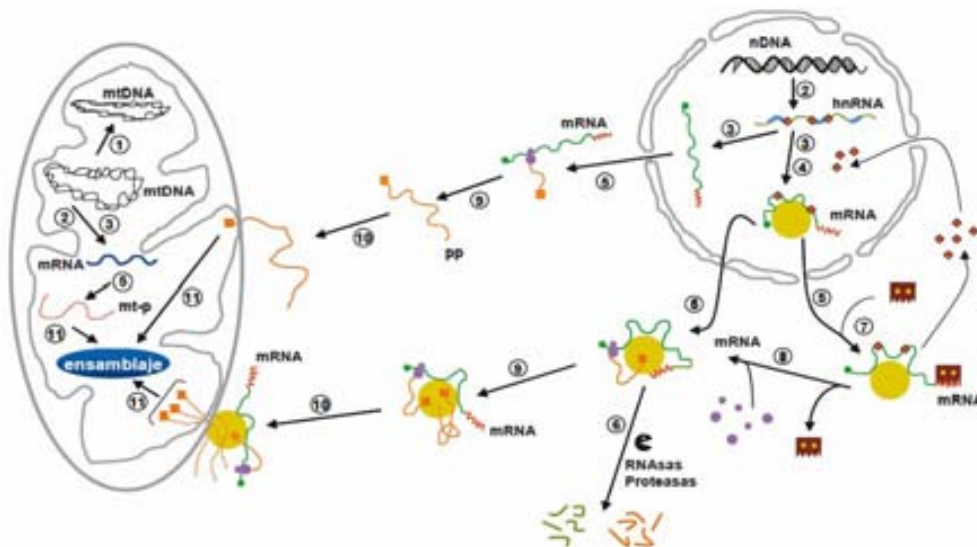


Figura 1.1_ **Procesos moleculares que controlan la expresión de los genes implicados en la biogénesis mitocondrial.** 1) Replicación del DNA mitocondrial. 2) Transcripción de los genes nucleares (DNA_n) y mitocondriales (DNA_{mt}). 3) Procesado y maduración de los transcritos primarios (hnRNA en el núcleo). 4) En el núcleo, antes de salir al citoplasma, algunos mRNAs nucleares se ensamblan en una estructura compleja. Otros, en cambio, son exportados sin un ensamblaje aparente en estructuras definidas involucradas en la localización del mRNA. 5) El mRNA, una vez maduro, y unido a proteínas formando complejos ribonucleoproteicos (mRNPs), se exporta al citoplasma donde se traduce con mayor o menor eficiencia y 6) se degrada con mayor o menor rapidez. 7) En el caso de ciertos mRNAs codificados en el núcleo la ribonucleoproteína exportada puede sufrir procesos de enmascaramiento traduccional debidos a la unión de proteínas reguladoras al 3'UTR del mRNA. 8) Este enmascaramiento puede cesar mediante la acción de determinadas señales que provocan la pérdida de la actividad de unión de estas proteínas al mRNA. 9) Se completa la síntesis de precusores proteicos mitocondriales (pp). 10) Los precusores proteicos que se sintetizan deslocalizados pueden ser importados a la mitocondria con la ayuda de las proteínas involucradas en el proceso de importación mediante difusión aleatoria por la membrana. Por el contrario, los precusores proteicos sintetizados en estructuras localizadas pueden migrar unidas a la ribonucleopartícula a través del citoesqueleto hasta la mitocondria. Las chaperonas moleculares colaboran en la traducción y direccionamiento de los precusores proteicos. 11) Las proteínas recién importadas son dirigidas y, a veces, ensambladas con las proteínas sintetizadas por la propia mitocondria. (Imagen adaptada de (Cuezva *et al.*, 1997).

En la formación de las mitocondrias contribuye mayoritariamente el genoma nuclear y en menor proporción, pero de forma esencial, el genoma mitocondrial. Son importantes también los mecanismos de importación y ordenamiento intramitocondrial de las proteínas que se sintetizan fuera de la mitocondria, los lípidos y, en ciertas especies, algunos RNAs (Fig1.1).

Aunque no tienen una única forma y medida, las mitocondrias son representadas generalmente en forma ovalada, que es como se encuentran en hepatocitos y fibroblastos, con unas dimensiones aproximadas de 3-4 μm de largo por 1 μm de diámetro. Son orgánulos dinámicos que responden al entorno y a señales de desarrollo de acuerdo con la demanda de energía celular. Debido a que las necesidades de energía de la célula varían, el contenido de mitocondrias es variable y puede ser ajustado según la situación (Lowell & Spiegelman, 2000; Hood, 2001).

1.1 Morfología y dinámica de la mitocondria

Gracias a la tomografía con microscopía electrónica, que permitió la reconstrucción tridimensional de la mitocondria (Mannella *et al.*, 1994), se observó que las mitocondrias contienen dos membranas de aproximadamente 7 nm de grosor, una exterior y la otra interior, que compartimentan el organelo en dos espacios, el espacio intermembrana y la matriz. Existen regiones puntuales dinámicas donde las membranas exterior e interior están en contacto directo y que tienen un papel importante en el transporte de proteínas a mitocondria (Pfanner *et al.*, 1990; Rassow *et al.*, 1990; Pfanner *et al.*, 1992). La membrana interna invagina formando estrechos túbulos llamados *cristae*, mayormente conectados a dicha membrana mediante las *cristae junctions* (Mannella *et al.*, 1997; Perkins *et al.*, 1997; Mannella *et al.*, 2001). Los *cristae* pueden presentarse como largos túbulos o se pueden fusionar para formar compartimentos lamelares (Fig1.2). El hecho de que el espacio intracristae comunique con el espacio entre las membranas externa e interna (espacio intermembrana) mediante estrechas *cristae junctions* sugiere que estos espacios intracristae podrían constituir compartimentos funcionalmente separados (Mannella *et al.*, 1997). Las *cristae junctions* podrían representar barreras de difusión entre el espacio intermembrana y el espacio intracristae, así como entre la membrana interna y la

membrana del cristae, previniendo o limitando la difusión de metabolitos como los protones o el ATP entre los espacios o de complejos proteicos en la membrana (Perkins *et al.*, 1997; Perkins & Frey, 2000; Frey *et al.*, 2002). Estudios recientes parecen demostrar que en la membrana interna se pueden diferenciar la membrana interna límite y la membrana cristae como dos subcompartimentos morfológicamente distintos, con diferente composición proteica, dinámica, biogénesis y mantenimiento que puede variar según el estado fisiológico de la célula (Vogel *et al.*, 2006; Wurm & Jakobs, 2006).

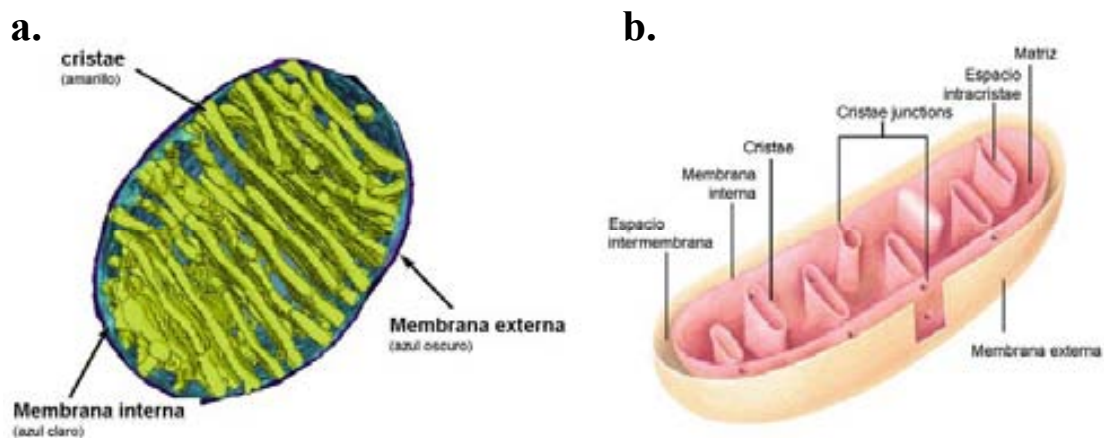


Figura 1.2_ (a) Modelo tridimensional reconstruido por tomografía electrónica (Perkins *et al.*, 1997). Esta técnica ha proporcionado imágenes in situ de mitocondrias de calidad en 3D. Los cristae mitocondriales se presentan como largos túbulos que pueden fusionarse para dar lugar a compartimentos lamelares. El espacio intracristae está conectado al espacio intermembrana mediante cristae junctions. (b) Modelo esquemático de la arquitectura de la mitocondria de M. Bobik y M. Martone de la universidad de California, teniendo en cuenta los datos observados.

Existe una estrecha relación entre el estado energético de la célula y la ultraestructura de la mitocondria, de manera que la estimulación de la respiración induce una conformación condensada con cambios en la matriz, el espacio intermembrana y los cristae (Hackenbrock, 1968; Hackenbrock *et al.*, 1971; Scalettar *et al.*, 1991). La mayoría de estudios realizados con levaduras muestran que en condiciones de represión de glucosa, las mitocondrias tienen mayor tendencia a formar filamentos. El tamaño del filamento, la interconectividad, así como el número de cristae aumentan cuando se termina la glucosa del medio y cuando las células crecen con azúcares no represivos o en presencia de sustratos de la cadena de transporte electrónico (CTE). Al alcanzar la saturación, en los cultivos se puede observar que los filamentos de mitocondrias se separan en pequeñas partículas (Sauvanet *et al.*, 2009).

Las mitocondrias son orgánulos altamente dinámicos. Pueden permanecer en una posición preferente, moverse o ser movidas donde sean necesarias e incluso cambiar de forma (Bereiter-Hahn & Voth, 1994), tamaño y número durante el desarrollo, el ciclo celular (Martinez-Diez *et al.*, 2006), en la apoptosis y senescencia (Yoon *et al.*, 2006; Arnoult, 2007) o en respuesta a diferentes estímulos (Rojo *et al.*, 1998). Además, la distribución y morfología de la mitocondria varía según el tipo celular (Collins & Bootman, 2003). La distribución de las mitocondrias en el citosol no es aleatoria. Se ha sugerido que el gradiente de ATP/ADP podría ser el responsable de la distribución de las mitocondrias en el citosol, de manera que los organelos se anclarían en lugares del citoplasma donde se necesita un suministro de ATP. Además, el movimiento de las mitocondrias puede ser debido a la interacción de éstas con el citoesqueleto y los motores moleculares (Morris & Hollenbeck, 1993, 1995; Hurd & Saxton, 1996).

La morfología y la actividad de las mitocondrias pueden ser reguladas también por procesos de fisión y fusión. En las células de mamífero, la mitocondria puede adaptar una amplia variedad de formas, desde redes tubulares a pequeñas esferas redondas. La morfología depende del balance entre estos procesos. Una fisión elevada da lugar a la fragmentación, mientras que una fusión elevada da lugar a formas alargadas. Estos eventos han sido estudiados ampliamente desde levadura hasta mamíferos (Okamoto & Shaw, 2005; Detmer & Chan, 2007a, b; Hoppins *et al.*, 2007; Suen *et al.*, 2008), corroborando los primeros estudios realizados mediante microscopía electrónica (Bereiter-Hahn & Voth, 1994). Se han desarrollado varios ensayos para medir de manera directa la fusión mitocondrial. Los métodos más comunes son la fusión artificial de células con polietilenglicol (PEG) o mediante virus. En estos experimentos, se detecta una elevada fusión mitocondrial unas horas después de la fusión celular y normalmente se completa al cabo de 12h. Si dos células parentales que contienen mitocondrias marcadas con diferentes fluorescentes se fusionan, la fusión mitocondrial puede cuantificarse en las células híbridas mediante el seguimiento por microscopía de fluorescencia de la mezcla resultante de ambos fluoróforos (Nunnari *et al.*, 1997; Legros *et al.*, 2002; Ishihara *et al.*, 2003; Mattenberger *et al.*, 2003). Estos estudios revelan que la fusión es frecuente en células de mamífero y que depende del potencial de la membrana interna. Otros estudios utilizan la expresión celular de GFP fotoactivable importada a mitocondria para

cuantificar la fusión, activando la GFP mediante un láser y observando la difusión de ésta a las mitocondrias vecinas (Karbowski *et al.*, 2004). En células de mamífero, existen tres GTPasas importantes para la fusión mitocondrial, que requiere la fusión coordinada de las membranas exterior e interior. Las mitofusinas se localizan en la membrana externa y están implicadas en las primeras etapas de la fusión (Koshihara *et al.*, 2004; Meeusen *et al.*, 2004; Song *et al.*, 2009). La proteína OPA1 está asociada con la membrana interna y es esencial para la fusión de ésta (Detmer & Chan, 2007a). Perturbaciones en la fusión mitocondrial, resultan en defectos en el potencial de membrana y la respiración, en un crecimiento pobre de la célula y una mayor susceptibilidad a la apoptosis (Chen *et al.*, 2005a). Estos resultados sugieren además, un modelo en que la fusión protege la función mitocondrial mediante la mezcla del material genético de las mitocondrias fusionadas (Chen *et al.*, 2007; Detmer & Chan, 2007a). Se ha hipotetizado que, debido a esta mezcla del contenido mitocondrial, la fusión puede estar relacionada con la habilidad de las células humanas para tolerar altos niveles de DNA mitocondrial patógeno (Nakada *et al.*, 2001; Nakada *et al.*, 2009). Dependiendo de la mutación, las células heteroplasmáticas pueden acumular hasta 60-90% de DNA mitocondrial patógeno sin una notable declinación de la actividad respiratoria (Rossignol *et al.*, 2003).

1.2 Origen evolutivo de la mitocondria

Después de que, por primera vez, las mitocondrias fuesen observadas en células de organismos superiores, varios autores como R. Altmann (1890), K.C. Mereschkowsky (1910), P. Portier (1918), J.E. Wallin (1927) y J. Lederberg (1952), de manera más o menos explícita sugirieron que podían estar relacionadas en cierto modo con las bacterias. La hipótesis no fue bien recibida por los darwinistas estrictos o partidarios de la teoría de la selección natural, y no fue hasta el descubrimiento del DNA mitocondrial que esta idea fue seriamente considerada. La teoría del origen procariótico de la mitocondria, hoy ampliamente aceptada, fue popularizada por Lynn Margulis en su libro “*Symbiosis In Cell Evolution*” (Margulis, 1981) y explica la endosimbiosis en las etapas iniciales de la vida en la Tierra bajo el dominio de las cianobacterias. Estas algas producían gran cantidad de oxígeno como subproducto de la fotosíntesis aumentando los niveles de oxígeno en la

atmósfera. Frente a este cambio, algunas bacterias desarrollaron capacidades aeróbicas. Los protoeucariotas anaeróbicos heterotróficos incorporaron mediante endocitosis estas bacterias aeróbicas, las α -proteobacterias, desarrollando una relación beneficiosa para ambos. Los primeros presentaban el DNA compartimentado en el núcleo y metabólicamente utilizaban la glicólisis anaeróbica y la fermentación, mientras los procariotas sin núcleo definido, contribuían en el sistema de transporte electrónico.

Además de la presencia del DNA mitocondrial y que éste se transcribe de forma continua, esta hipótesis tiene, entre sus fundamentos, el recubrimiento de la mitocondria por una doble membrana de composición lipídica parecida a las bacteriales y una secuencia de RNA ribosomal más parecida a bacterias que a los ribosomas eucariotas citoplasmáticos (Raven, 1970; Schwartz & Dayhoff, 1978). Otra evidencia que sostiene esta hipótesis es que el código genético del DNA mitocondrial no suele ser el mismo que el código genético del DNA nuclear (Barrell *et al.*, 1979).

En las últimas dos décadas se han secuenciado varios genomas mitocondriales, el análisis de los cuales resume varios aspectos de la estructura, el contenido, la organización y la expresión del genoma mitocondrial (Gray *et al.*, 1998; Boore, 1999; Lang *et al.*, 1999; Feagin, 2000). Actualmente, el contenido genético mitocondrial varía ampliamente, desde el genoma mitocondrial de *Reclinomonas americana* que contiene 67 genes que codifican para proteínas, hasta el genoma mitocondrial de *apicomplexanos*, un grupo de protistas parásitos estrictos, que codifican solo para 3 genes (Vaidya *et al.*, 1989; Lang *et al.*, 1997). Este contenido genético tan dispar es atribuido mayoritariamente a la transferencia de información genética de la mitocondria al núcleo (Martin & Herrmann, 1998; Berg & Kurland, 2000). Parte de la información del genoma mitocondrial se perdió también al substituirse algunos genes mitocondriales por otros nucleares relacionados de función similar. En la secuenciación del genoma mitocondrial del protozoo *Reclinomonas americana* (Palmer, 1997) se descubrió que codificaba para 4 subunidades de una RNA polimerasa parecida a las polimerasas de eubacteria. Este descubrimiento sugiere que durante la evolución se perdió la RNA polimerasa eubacteriana formada por varias subunidades y fue reemplazada por una polimerasa codificada por núcleo de una sola subunidad parecida a la RNA polimerasa del bacteriófago T3/T7.

Seguramente se dio también la pérdida de material genético mitocondrial sin ser funcionalmente cubierto por genes nucleares. Un ejemplo es el complejo I (CTE) en *Saccharomyces cerevisiae*. Los genes clásicos de este complejo no se encuentran codificados ni en el genoma mitocondrial ni en el nuclear resultando en la ausencia del primer complejo de la cadena respiratoria (Kurland & Andersson, 2000).

El DNA mitocondrial presenta una extraordinaria diversidad en estructura, organización y contenido de genes, modo de expresión y replicación en los diferentes organismos.

Actualmente los genes que permanecen en la mitocondria pueden presentarse en estructuras tan diversas como el DNA mitocondrial lineal de 46 Kbases de *Tetrahymena* (Goldbach, 1977), el DNA circular de 16-20 Kbases de los metazoos (Attardi, 1985) o el gigantesco DNA mitocondrial circular de 570 Kbases del maíz (Lonsdale *et al.*, 1984).

Se han realizado estudios en levadura de las proteínas mitocondriales codificadas en el núcleo (Karlberg *et al.*, 2000; Marcotte, 2000; Marcotte *et al.*, 2000). Ambos estudios, basándose en la búsqueda de similitudes de estas proteínas, llegan a conclusiones similares sobre el origen del proteosoma mitocondrial en levaduras: 50-60% de origen procariota (homólogas a proteínas de genomas procariotas), 20-30% de origen eucariota (homólogas a proteínas de genomas eucariotas) y un 20% específicas de organismo. Las proteínas de origen procariota intervienen mayoritariamente en procesos de biosíntesis, bioenergética y síntesis de proteínas, mientras las de origen eucariota son mayoritariamente proteínas de regulación, transporte o componentes de la membrana.

Debido al origen eubacteriano de la mitocondria, podríamos decir que la fracción de genes de origen procariota es una aproximación al primer genoma mitocondrial, mas no podemos asegurar que todos sean genes transferidos de la α -proteobacteria al genoma nuclear porque no se conocen homólogos codificados en el DNA mitocondrial y no se pueden descartar fenómenos de transferencia horizontal de genes en etapas posteriores a la endosimbiosis (Karlin & Brocchieri, 2000).

En lo referente a los genes de origen eucariota, no podemos afirmar que sean genes que la mitocondria reclutó posteriormente del genoma nuclear para complementar el genoma bacteriano antecesor pues algunos de ellos pueden ser genes transferidos al núcleo que hayan divergido tanto en secuencia que los homólogos más directos sean genes eucariotas.

¿Por qué ha permanecido una molécula de DNA en la mitocondria? Se han sugerido una serie de causas (Montoya, 2005):

- (1) Las proteínas codificadas por DNAm_t presentan un alto grado de hidrofobicidad, lo que dificultaría la importación a mitocondria,
- (2) Estas proteínas podrían ser tóxicas en el citoplasma,
- (3) La existencia de un genoma mitocondrial permitiría una regulación local más rápida y fina de dichas proteínas
- (4) Es posible que no se haya terminado todavía de transferir toda la información genética al núcleo.

1.3 Funciones mitocondriales

1.3.1 Vías metabólicas

La mitocondria es la productora de la principal fuente de energía, el ATP y además, contiene muchas vías metabólicas incluyendo el complejo piruvato deshidrogenada, el ciclo de la urea, la biosíntesis del grupo hemo, el metabolismo y biosíntesis de la cardiolipina y otros lípidos, la biosíntesis de ubiquinol, el sistema de la β -oxidación y el ciclo de Krebs.

El ciclo de Krebs

Fue el segundo ciclo, después del ciclo de la urea, postulado por Krebs en 1937 y formulado según la siguiente reacción:



Fue el descubrimiento del coenzima A por N.O. Kaplan y F. Lipmann en 1945 lo que permitió a S. Ochoa y F. Lynen demostrar, seis años después, que el acetil-CoA era el intermediario en la reacción con el oxalacetato para formar citrato.

Actualmente las reacciones del ácido cítrico están totalmente establecidas. En organismos aeróbicos, el ciclo de Krebs es parte de la vía catabólica que realiza la oxidación de glúcidos, ácidos grasos y aminoácidos hasta producir CO_2 , liberando energía en forma utilizable (poder reductor y

GTP). El metabolismo oxidativo de glúcidos, grasas y proteínas frecuentemente se divide en tres etapas, de las cuales, el ciclo de Krebs supone la segunda (Fig 1.3). En la primera etapa, los carbonos de estas macromoléculas dan lugar a moléculas de acetil-CoA de dos carbonos, e incluye las vías catabólicas de aminoácidos, la beta oxidación de ácidos grasos y la glucólisis. La tercera etapa es la fosforilación oxidativa. El ciclo de Krebs también proporciona precursores para muchas biomoléculas, como ciertos aminoácidos. Por ello se considera una vía anfibólica, es decir, catabólica y anabólica al mismo tiempo.

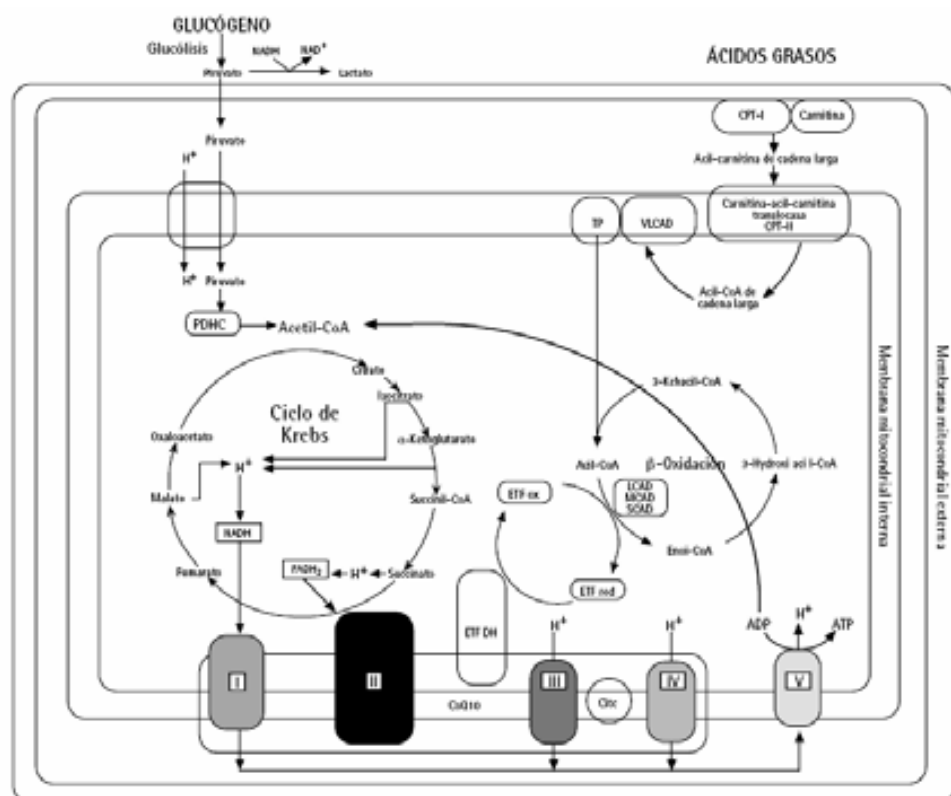
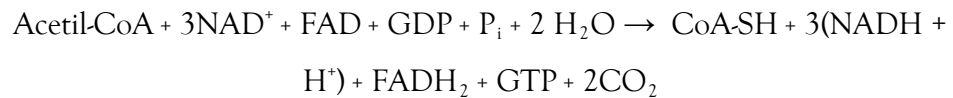


Fig 1.3_ Esquema de las principales reacciones enzimáticas que se dan en mitocondria. Las moléculas de NAD y FAD reducidas, producto del ciclo de Krebs (metabolismo de los glúcidos) y la β -oxidación (metabolismo de los ácidos grasos), son transferidas a la cadena de transporte electrónico. Esta cadena consiste en cuatro complejos enzimáticos multiméricos (I al IV) y dos pequeños transportadores, la coenzima Q o ubiquinona y el citocromo C. La energía producida es usada para bombear protones al espacio intermembrana generando un gradiente electroquímico que es utilizado por la ATP sintasa para producir ATP (Andreu & Gonzalo-Sanz, 2004).

El acetil-CoA (Acetil Coenzima A) es el principal precursor del ciclo. El ácido cítrico (6 carbonos) o citrato se regenera en cada ciclo por

condensación de un acetil-CoA (2 carbonos) con una molécula de oxaloacetato (4 carbonos). El citrato produce en cada ciclo una molécula de oxaloacetato y dos CO₂, por lo que el balance neto del ciclo es:



Los dos carbonos del Acetil-CoA son oxidados a CO₂, y la energía del proceso es liberada en forma de energía química, GTP y poder reductor, NADH y FADH₂, coenzimas utilizados posteriormente en la fosforilación oxidativa para la síntesis final de ATP.

Metabolismo de ácidos grasos

La oxidación de los ácidos grasos fue una de las primeras vías metabólicas localizadas en mitocondria gracias a los estudios pioneros de A.L. Lehninger y E.P. Kennedy a finales de los 1940s.

La beta oxidación (β-oxidación) es un proceso catabólico de los ácidos grasos en el cual, mediante la oxidación, el ácido graso pierde un par de átomos de carbono sucesivamente en cada ciclo del proceso hasta que se descompone por completo (Fig 1.3). La β-oxidación de ácidos grasos consta de cuatro reacciones recurrentes; oxidación por FAD, hidratación, oxidación por NAD⁺ y tiólisis. El resultado de dichas reacciones son unidades de dos carbonos en forma de acetil-CoA, molécula precursora del ciclo de Krebs, y coenzimas reducidos (NADH y FADH₂) que son necesarios en la cadena respiratoria. No obstante, antes de producirse la oxidación, los ácidos grasos deben activarse con coenzima A y atravesar la membrana mitocondrial interna, que es impermeable a ellos.

1.3.2 Cadena de transporte electrónico. Fosforilación oxidativa.

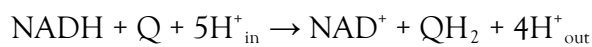
La principal función de la mitocondria es suministrar energía a la célula mediante la fosforilación oxidativa. Durante la fosforilación oxidativa, los electrones son transferidos a través de reacciones redox hasta el aceptor final que es el oxígeno. Estas reacciones bombean protones desde la matriz mitocondrial al espacio intermembrana, generando un gradiente electroquímico que es utilizado para

producir ATP. La teoría quimiosmótica de la fosforilación oxidativa fue formulada por primera vez en 1961 (Mitchell, 1961).

En eucariotas, esta función la realiza la cadena de transporte electrónico (CTE) compuesta por cinco complejos proteicos anclados en la membrana mitocondrial interna, además de la ubiquinona y el citocromo c que actúan como transportadores móviles entre complejos. Cada complejo está formado por varias subunidades peptídicas, 13 de las cuales están codificadas en mitocondria.

Complejo I (NADH deshidrogenasa)

Cataliza la reacción:



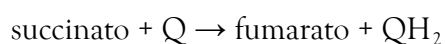
donde Q y QH₂ representan las formas oxidada y reducida de la ubiquinona. En esta reacción se bombean 4 protones de la matriz al espacio intermembrana.

En mamíferos, este complejo está formado por 41 péptidos de los cuales 7 están codificados en mitocondria. Mediante microscopía electrónica de muestras bovinas (Grigorieff, 1998) se observó que el complejo tiene una forma de L igual que en los estudios previos con *Neurospora crassa*. El brazo largo se sumerge en la membrana mitocondrial y presenta actividad ubiquinona hidrogenasa interna. El brazo corto presenta actividad NADH deshidrogenasa y se extiende en la matriz. Contiene el cofactor FMN (flavin mononucleótido), hasta cuatro clústers de Fe-S y el lugar de unión de NADH.

Aunque la estructura del complejo eucariota no está caracterizada del todo, en 2006 se publicó la estructura cristalográfica del dominio hidrofóbico del complejo I de *Thermus thermophilus* (Sazanov & Hinchliffe, 2006).

Complejo II (succinato deshidrogenasa)

Cataliza la reacción:



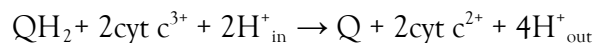
comunicando directamente el ciclo de Krebs y la cadena de transporte electrónico (Davis & Hatefi, 1972).

Está formado por 4 péptidos codificados en núcleo y está anclado a la membrana interna mediante dos proteínas, C_{II3} y C_{II4}. La subunidad flavoproteínica Fp (Kenney *et al.*, 1972) está formada por el péptido mayor del complejo y el FAD (flavina adenina dinucleótido) y se encuentra unida a la subunidad Ip (*iron-protein*) formada por el péptido de 27kDa y 3 centros no-hemo de Fe-S. El subcomplejo Fp-Ip contiene el lugar de unión del succinato, el aceptor de hidrógeno (FAD) y los centros de Fe-S para la transferencia de electrones a las proteínas ancladas en la membrana, mientras éstas últimas contienen el lugar de unión de la ubiquinona y de un citocromo tipo b (Merli *et al.*, 1979).

Se ha resuelto la estructura de la fumarato deshidrogenasa de *E. coli* y *Wolinella* a 2.7 y 2.2Å respectivamente (Iverson *et al.*, 1999; Lancaster *et al.*, 1999) y, más recientemente, la estructura de la succinato reductasa de *E. coli* a 2.6 Å (Yankovskaya *et al.*, 2003).

Complejo III (Q-citocromo c oxidorreductasa)

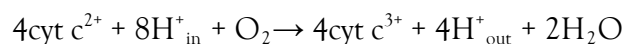
Cataliza la reacción:



De los 11 péptidos que forman este complejo, sólo el citocromo b está codificado por el genoma mitocondrial. Funcionalmente las subunidades más importantes encargadas de la transferencia de electrones y el bombeo de protones son los citocromos b y c₁ y la proteína Fe-S Rieske. La estructura cristalográfica completa del complejo III bovino a 2.9 Å de resolución fue publicada en 1997 (Xia *et al.*, 1997).

Complejo IV (citocromo c oxidasa)

Cataliza la reacción:



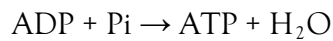
el oxígeno es el aceptor final de electrones.

El complejo en mamíferos está formado por 13 subunidades, 3 de las cuales la I, II y III (las tres mayores) están codificadas en el genoma mitocondrial. Éstas forman el núcleo funcional del complejo: la subunidad I une a los grupos prostéticos hemo (a y a₃) y contiene Cu_B

(centro redox de cobre); la subunidad II contiene Cu_A y la III está involucrada en el bombeo de protones y en la modulación del transporte de electrones a través de los centros metálicos (Schlerf *et al.*, 1988; Schillace *et al.*, 1994). La primera estructura cristalográfica a alta resolución fue la del complejo IV bovino (Tsukihara *et al.*, 1996).

Complejo V (F_1F_0 -ATPasa)

Es el responsable de la síntesis de ATP a partir de ADP y fósforo inorgánico:



Esta reacción se produce gracias al gradiente de protones transmembrana generado por los complejos I, III y IV de la cadena de transporte electrónico. El complejo puede ser dividido en dos subcomplejos funcionalmente distintos. Uno hidrosoluble, denominado F_1 , que se encuentra expuesto a la matriz mitocondrial (Boyer, 1997; Karrasch & Walker, 1999) y que posee la actividad catalítica, y un subcomplejo liposoluble F_0 que se encuentra embebido en la membrana interna mitocondrial y que actúa como canal de protones (Lutter *et al.*, 1993; Karrasch & Walker, 1999). El subcomplejo F_0 es un motor giratorio y está unido a F_1 por dos tallos que hace que la rotación esté directamente acoplada. La orientación de ambos subcomplejos es tal que tienden a girar en direcciones opuestas (Yoshida *et al.*, 2001; Capaldi & Aggeler, 2002). En condiciones fisiológicas, el transporte de protones a través del canal F_0 hace que su fuerza rotatoria sea mayor por lo que F_1 gira en sentido opuesto. Esto origina cambios conformacionales en el subcomplejo F_1 que conducen a la síntesis de ATP dirigida por el gradiente electroquímico ($\Delta\mu\text{H}^+$) (Capaldi & Aggeler, 2002).

El genoma mitocondrial codifica dos de las subunidades de F_0 . La estructura cristalográfica de la ATPsintasa bovina a 2.8 Å de resolución fue publicada en 1994 (Abrahams *et al.*, 1994).

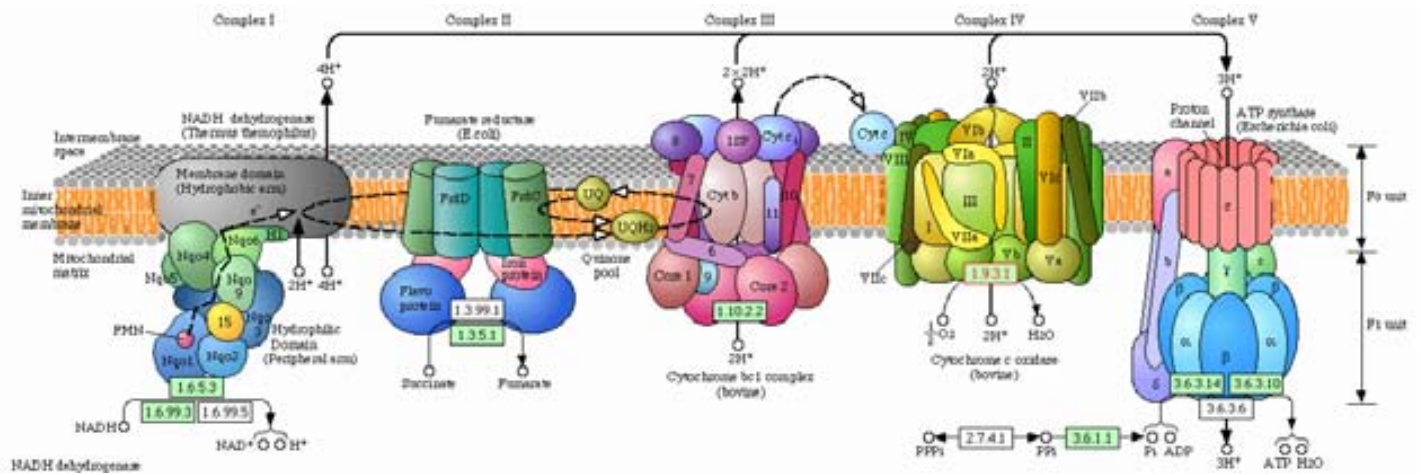


Figura 1.4_ Sistema de fosforilación oxidativa OXPHOS. Compuesto por la cadena respiratoria acoplada a la fosforilación oxidativa. La oxidación de sustratos metabólicos produce la reducción de las coenzimas NAD^+ y FAD que actúan como donadores de electrones (e^-) a los complejos de la cadena respiratoria que se encuentran en la membrana interna de la mitocondria (M.I.). El último receptor de e^- es el oxígeno que da lugar a una molécula de agua. La transferencia de e^- está acoplada a un bombeo de protones desde la matriz al espacio intermembrana por los complejos C-I, C-III y C-IV que genera un gradiente electroquímico a través de la membrana interna. El gradiente electroquímico es la fuerza que utiliza la H^+ -ATP sintasa (C-V) para generar ATP. Figura tomada de (Kanehisa_Laboratories, 2008)

1.3.3 Implicación de la mitocondria en muerte celular.

La clasificación de las vías de muerte celular se ha basado principalmente en criterios morfológicos. De esta manera, podemos diferenciar entre la apoptosis, en la que se produce una condensación del DNA visible; la muerte autofágica, en la que se observa la presencia de vacuolas autofágicas en el citosol; la catástrofe mitótica, que se caracteriza por fallos en la ejecución de la mitosis; y la necrosis, en la que se da un hinchamiento específico de los orgánulos en el citosol (Jaattela, 2004). De todos estos tipos de muerte celular, la apoptosis es la más frecuente y, en consecuencia, la más estudiada. La apoptosis está mediada por un patrón molecular que culmina en la activación de una familia de cisteín-proteasas, conocidas como caspasas, que son las responsables de controlar la destrucción celular a través de cortes proteolíticos muy específicos (Zimmermann *et al.*, 2001). Una vez que las células han llevado a cabo el proceso apoptótico son eliminadas por fagocitos sin que se produzca la liberación del contenido celular, por lo que no se produce una respuesta inflamatoria.

Existen dos cascadas de activación de la apoptosis, la vía extrínseca (Kaltschmidt *et al.*, 2000) inducida por estímulos fisiológicos y la vía intrínseca (Newmeyer &

Ferguson-Miller, 2003), inducida por estímulos de estrés. En la vía intrínseca se integran señales de muerte producidas por alteraciones en el estado metabólico de la célula, daño en el DNA u otras situaciones de estrés a través de la mitocondria. La mitocondria contiene una gran variedad de proteínas pro-apoptóticas y, en este contexto, la salida de estas proteínas de la mitocondria al citosol induce la apoptosis. Sin embargo, los mecanismos de la permeabilización de la membrana mitocondrial generan todavía mucha controversia. Debido a su importancia, la gran mayoría de los estudios realizados se han focalizado en la salida de citocromo c. Se ha propuesto que la mitocondria puede liberar proteínas del espacio intermembrana a través dos mecanismos diferentes. El primero, a través de la permeabilización transitoria de la mitocondria (MPT: *Mitochondrial Permeability Transition*) que ocurre principalmente en respuesta a la salida de Ca^{2+} del retículo endoplásmico (Petit *et al.*, 1998; Yang & Cortopassi, 1998) y que implica la participación de transportadores mitocondriales de la membrana externa, VDAC (*Voltage Dependent Anion Channel*) (Vander Heiden *et al.*, 2000) y de la membrana interna, ANT (*Adenine Nucleotide Translocators*) (Vieira *et al.*, 2000). El segundo mecanismo ocurre a través de la permeabilización de la membrana externa producida por la activación de proteínas pro-apoptóticas de la familia de Bcl-2. Se han descrito varios modelos de permeabilización de la mitocondria mediada por estas proteínas. Una hipótesis es la generación de poros en la membrana externa de la mitocondria mediada por canales formados por miembros pro-apoptóticos de esta familia al integrarse en la membrana mitocondrial (Minn *et al.*, 1997; Schendel *et al.*, 1997).

La existencia de diferentes mecanismos explica, en gran medida, la gran variedad de respuestas de la mitocondria a numerosos estímulos apoptóticos en diferentes líneas celulares. La permeabilización de la membrana mitocondrial es un evento crítico en la muerte celular patológica por isquemia, intoxicación con compuestos xenobióticos, enfermedades neurodegenerativas o infecciones virales. La inhibición de la permeabilización constituye una importante estrategia para la prevención farmacéutica de la muerte celular no deseada. En cambio, la inducción de este fenómeno en las células tumorales constituye una parte importante de la quimioterapia anticancerígena (Kroemer *et al.*, 2007)

Existen evidencias que indican que las células que han sido tratadas con estímulos apoptóticos pueden iniciar un programa de muerte que no requiere la activación de las caspasas, la muerte celular independiente de caspasa (CICD: *Caspase-independent Cell Death*) (Leist & Jaattela, 2001; Chipuk & Green, 2005). Este tipo de muerte ocurre cuando los inhibidores de las caspasas están presentes, o cuando el patrón apoptótico se encuentra genéticamente alterado. Como en la apoptosis clásica, los mecanismos de muerte independientes de caspasas también están mediados por receptores de muerte y por alteraciones en la mitocondria.

Ambas vías, tanto la caspasa-dependiente como la caspasa-independiente, parece que son activadas en muchos casos de forma simultánea (Leist & Jaattela, 2001).

1.4 Disfunciones y enfermedades mitocondriales

Las enfermedades mitocondriales (también conocidas como encefalomiopatías mitocondriales) constituyen un amplio grupo de patologías que se identificaron a partir de la década de 1960, cuando se introdujeron las técnicas morfológicas que permitieron el estudio ultraestructural e histoquímico del tejido muscular. La mitocondria contiene muchas vías metabólicas y, defectos en cualquiera de estas vías causan enfermedades mitocondriales. Pero el término encefalomiopatía mitocondrial –en muchos de los pacientes con miopatías éstas estaban acompañadas por diversos síntomas que implicaban el sistema nervioso central- se ha utilizado para enfermedades relacionadas con la cadena de transporte electrónico (Andreu & Gonzalo-Sanz, 2004).

Las características clínicas de estas enfermedades son muy heterogéneas afectando, en muchos casos, a diferentes órganos y tejidos y su correcto diagnóstico requiere datos clínicos, morfológicos, bioquímicos y genéticos muy precisos (Solano *et al.*, 2001).

La característica más remarcable de este sistema es que está codificado por dos genomas distintos, el nuclear y el mitocondrial. Así pues, desde un punto de vista genético, podemos clasificar las enfermedades mitocondriales en enfermedades debidas a alteraciones del DNAm y enfermedades debidas a alteraciones en el DNAn.

1.4.1 Enfermedades debidas a alteraciones en el DNAn

Todas ellas son transmitidas por herencia mendeliana e incluyen dos grupos principales (Shoubridge, 2001):

Mutaciones en genes que afectan a la cadena respiratoria. Existen mutaciones que afectan directamente a los genes que codifican para los componentes estructurales de los complejos de la fosforilación oxidativa codificados en núcleo y que dan lugar a patologías como el síndrome de Leigh o paraganglioma, o a desórdenes mitocondriales que tienen efectos secundarios sobre el sistema de fosforilación oxidativa, como la ataxia de Friedreich o la paraplejía espástica hereditaria.

Mutaciones que afectan al ensamblaje o el mantenimiento de la integridad del DNA mitocondrial (como oftalmoplegia externa progresiva PEO). Para su correcto funcionamiento y replicación, el DNAm_t es altamente dependiente de numerosos factores codificados por genes nucleares. Mutaciones en estos genes son la causa de enfermedades mitocondriales caracterizadas por alteraciones en el DNAm_t que pueden ser cuantitativas como depleciones, o cualitativas como deleciones múltiples (Hirano & Dimauro, 2001).

1.4.2 Enfermedades debidas a alteraciones en el DNAm_t

Las características físicas y el patrón de herencia del DNAm_t definen algunas reglas de la “genética mitocondrial” distintas de la genética mendeliana:

Heteroplasmia y efecto umbral. Cada célula contiene cientos o miles de copias de DNAm_t dependiendo del tejido, del tipo de célula o del estado metabólico. Durante la división celular las moléculas de DNAm_t se distribuyen de forma aleatoria. En tejidos sanos todas las copias son idénticas (homoplasmia). Las mutaciones patogénicas normalmente afectan a alguna pero no a todas las moléculas de DNAm_t, lo que da lugar a una mezcla de genomas mutados y no mutados (heteroplasmia). El fenotipo clínico de una mutación puntual en el DNAm_t viene principalmente determinada por la proporción relativa de genomas sanos

y mutados en los diferentes tejidos. Este umbral es diferente en tejidos distintos dependiendo de la demanda energética.

Segregación mitótica. En la división celular la proporción de mitocondrias con DNAm_t mutante en las nuevas células puede variar, y con ello su fenotipo.

Herencia materna. En la fertilización, todo el DNA proviene del oocito. En consecuencia la transmisión del DNAm_t (y de las posibles mutaciones) es únicamente materna, a diferencia de la herencia mendeliana.

Alta velocidad de mutación. El DNAm_t muestra una tasa de mutación diez veces superior a la del DNA nuclear. Este fenómeno puede ser causado por la producción de radicales de oxígeno en la mitocondria, como consecuencia de la oxidación de los compuestos de carbono, que dañan el DNA desprotegido. Este hecho da lugar a enormes variaciones en la secuencia del DNA mitocondrial entre individuos de la misma especie e incluso a cierta heterogeneidad a lo largo de la vida de cada individuo. Se ha propuesto que la disminución de la capacidad respiratoria de los tejidos que ocurre con el envejecimiento podría ser debida a la acumulación de daño en el DNA mitocondrial (Miquel, 1998; Michikawa *et al.*, 1999).

Las enfermedades causadas por daño en el genoma mitocondrial tienen en común que presentan una deficiencia en la síntesis de ATP, ya que el DNAm_t codifica básicamente para algunos de los componentes de la CTE.

El grupo de síndromes clínicos debidos a alteraciones primarias del DNAm_t incluye los reordenamientos simples y complejos del DNAm_t (delecciones y duplicaciones) y las mutaciones puntuales. Hasta la fecha, más de 150 mutaciones puntuales han sido identificadas en el DNAm_t de pacientes con una variedad de enfermedades, muchas de las cuales son de herencia materna y multisistémicas, aunque algunas son esporádicas y específicas de tejido (Figura 1.5).

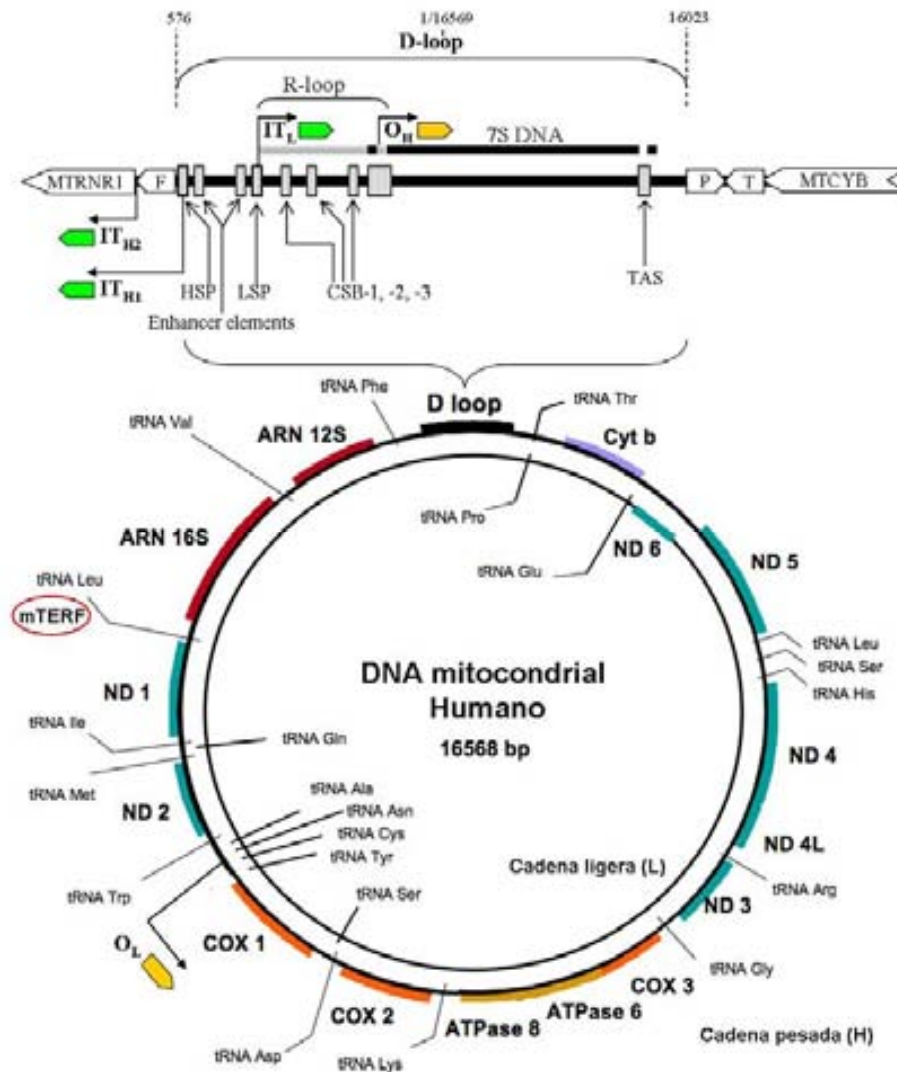


Fig 1.6_ Mapa genético del DNA mitocondrial humano. En la parte superior se muestra el mapa de la región D-loop con la posible región R-loop RNA. Los genes marcados con F, P y T corresponden a los tRNA^{Phe}, tRNA^{Pro} y tRNA^{Thr} respectivamente. Los cuadros sombreados representan varias secuencias involucradas en la transcripción y la replicación del DNA mitocondrial. Los lugares de iniciación de la transcripción y su dirección están señalizados con flechas verdes, se muestran los dos lugares de iniciación de la transcripción de la cadena H (HSP), IT_{H1} e IT_{H2}, y el lugar de iniciación de la transcripción de la cadena L (LSP), IT_L. Se indican también los correspondientes elementos potenciadores *enhancer*, para cada lugar de inicio de la transcripción, donde se une TFAM. O_H y O_L marcados con una flecha amarilla, son los inicios de de replicación de las cadenas H y L respectivamente según el modelo de desplazamiento de hebra (Ver texto). Figura adaptada de (Berdanier, 2005) y (Bellance *et al.*, 2009).

En los años 80, estudios con *Saccharomyces cerevisiae* revelaron que el genoma mitocondrial se organiza en complejos multiproteicos llamados nucleoides (Miyakawa *et al.*, 1984; Miyakawa *et al.*, 1987). Estos complejos fueron visualizados con sondas fluorescentes específicas de DNA en cultivos celulares de mamíferos (Satoh & Kuroiwa, 1991). Así, el DNAmt dentro de la mitocondria no se encuentra como moléculas aisladas sino formando nucleoides que contienen entre 2-10 moléculas de DNA en asociación con

proteínas (Legros *et al.*, 2004) como el factor de transcripción mitocondrial A (TFAM), la proteína *single-stranded DNA binding protein* de mitocondria (mtSSB), la helicasa Twinkle y la polimerasa gamma (POLG) entre otras, algunas de las cuales todavía no han sido identificadas (Garrido *et al.*, 2003). Un análisis cuantitativo del tamaño y el contenido de los nucleoides en cultivos celulares sugiere que éstos contienen una media de 5-7 moléculas de DNA empaquetadas en una estructura de 70 nm de diámetro (Iborra *et al.*, 2004), con una densidad de empaquetado similar a la de los nucleoides bacterianos. La microscopía de fluorescencia muestra que los nucleoides son estructuras dinámicas capaces de dividirse y distribuirse en la red mitocondrial para asegurar la transmisión del DNA mitocondrial durante el crecimiento y la división.

Parece ser que el DNA mitocondrial está anclado en la membrana mitocondrial interna por el D-loop mediante una interacción proteica (Albring *et al.*, 1977; Bogenhagen *et al.*, 2003). Estudios inmunocitoquímicos en células humanas sugieren además, que los nucleoides están sujetos, directa o indirectamente, mediante las membranas mitocondriales al motor de kinesina (factor KIF5B) implicado en el movimiento de las mitocondrias a lo largo de los microtúbulos. Experimentos cinéticos de replicación y transcripción revelan que cada genoma del nucleoide se replica independientemente y que los RNAs nacientes permanecen en las proximidades del nucleoide. Además, éstos RNAs mitocondriales se localizaron -con respecto a los componentes extramitocondriales- próximos a dos componentes de la maquinaria de traducción citoplasmática y de uno de los componentes de la maquinaria que importa a la mitocondria proteínas codificadas en núcleo (Iborra *et al.*, 2004) (Fig 1.7b). Este hecho sugiere que los clústers de genomas mitocondriales influyen de alguna manera la organización de las maquinarias de traducción a ambos lados de la membrana mitocondrial, de manera que la proximidad facilitaría el correcto ensamblaje de los complejos mitocondriales formados por proteínas codificadas en el núcleo y en mitocondria. Recientes estudios de purificación y caracterización de los nucleoides han identificado varias proteínas ribosomales mitocondriales, factores de traducción así como chaperonas y otros enzimas metabólicos (Chen & Butow, 2005; Wang & Bogenhagen, 2006; Bogenhagen *et al.*, 2008) (Fig 1.7c).

El DNAm_t humano codifica para dos RNAs ribosomales (12S y 16S), 22 tRNAs y trece péptidos todos ellos componentes de la cadena respiratoria. Las subunidades 1, 2, 3, 4, 4L, 5 y 6 del complejo I; la subunidad b (citocromo b) del complejo III; las subunidades I, II y III del complejo IV y las subunidades 6 y 8 del complejo V (Fig1.6). Un análisis comparativo de los genomas mitocondriales de diferentes organismos indica que la organización genética está altamente conservada en mamíferos (Wolstenholme, 1992; Jameson *et al.*, 2003). La información genética del DNAm_t se encuentra muy compactada, a diferencia de la mayoría de organismos excepto en los virus (Anderson *et al.*, 1981). Aparte de un corto fragmento en el origen de replicación, el DNAm_t de mamíferos está completamente saturado de genes sin intrones. La mayoría de los genes se transcriben a partir de la cadena H, incluyendo los dos rRNAs, 14 tRNAs y 12 genes que codifican para proteínas (Ojala *et al.*, 1980). En la mayoría de casos, estos genes se encuentran unidos o separados tan solo por algunos nucleótidos, e incluso en algunos casos los genes se superponen. Es el caso de las subunidades 6 y 8 de la ATPasa y de ND4 y ND4L que se superponen en 46 y 7 nucleótidos respectivamente. Además, los mRNAs mitocondriales comienzan directamente con el codón de iniciación, carecen por tanto, de un tramo no codificante en el extremo 5' característico de otros sistemas (Montoya *et al.*, 1981). Aproximadamente la mitad de genes no presentan el codón de terminación completo, en estos casos, la terminación U o UA del gen se completa postranscripcionalmente a UAA (codón de terminación) mediante la poliadenilación de los mRNAs (Ojala *et al.*, 1981). Además de la compactación genética, el DNAm_t sólo requiere de 22 tRNAs de estructura no ortodoxa para leer todos los codones. Otra característica del genoma mitocondrial de mamíferos es la distribución de los genes que codifican para tRNAs ya que se encuentran separando los genes que codifican para proteínas y rRNAs. Parece ser que estructuralmente funcionan como señales para el procesamiento enzimático del RNA dando lugar a los diferentes mRNA maduros (Ojala *et al.*, 1980; Ojala *et al.*, 1981).

La región no codificante del DNAm_t (1122 nucleótidos), conocida como región de control, se encuentra entre el tRNA^{Pro} y el tRNA^{Phe} y contiene el origen de replicación para la cadena H, los promotores para la transcripción de las dos cadenas y el D-loop. Esta región contiene también tres bloques de secuencia conservada (CSBs I, II y III) que se

crea juegan un importante papel en la regulación de la replicación del DNAm_t, y una zona híbrida RNA-DNA (Fig 1.6) (Xu & Clayton, 1996).

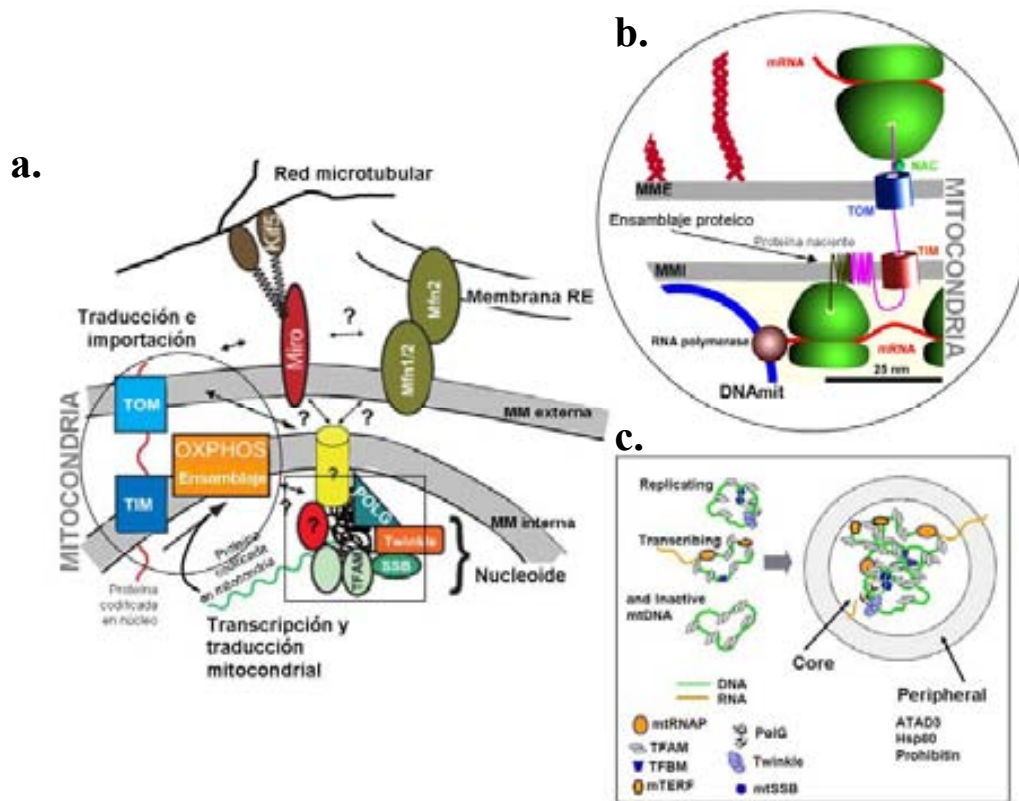


Figura 1.7_ (a) Esta figura ilustra los recientes avances en el estudio de la organización de los nucleoides en mamíferos, en el contexto de la biología mitocondrial, concretamente en la dinámica mitocondrial y la biogénesis de los complejos de la cadena respiratoria (Spelbrink)2010). (b) Los resultados experimentales parecen indicar que los clústers de genomas mitocondriales organizan las maquinarias de traducción a ambos lados de la membrana de manera que las proteínas codificadas en núcleo y destinadas a mitocondria se transcriben cerca de las proteínas codificadas en mitocondria para que su ensamblaje en los complejos mitocondriales sea eficiente. (Iborra *et al.*, 2004). (c) Detalle del modelo estructural de los nucleoides de DNA mitocondrial. Las moléculas de DNA mitocondrial que forman el nucleoide pueden encontrarse en diferentes estadios, de replicación, de transcripción o inactivas. Estas moléculas se encuentran agregadas en el centro del nucleoide que contiene los factores mtRNAP, TFAM, TFBM, mTERF, PolG, Twinkle y mtSSB. Estos cores están rodeados por una zona periférica que contiene otras proteínas que no muestran un cross-linking eficiente con el DNA mitocondrial pero que están presentes en los nucleoides nativos. (Bogehagen *et al.*, 2008).

1.5.1 Replicación del DNAm_t humano

En los últimos 20 años, ha sido ampliamente aceptada la teoría del modelo de desplazamiento de hebra para explicar la replicación del DNAm_t. Estudios recientes, sin embargo, ponen en duda este modelo y proponen un modelo de replicación bidireccional y simétrico desde un único origen de replicación.

Modelo de desplazamiento de hebra.

Este modelo fue propuesto para explicar los datos obtenidos a partir de estudios de microscopía electrónica (Robberson *et al.*, 1972) y sugiere que las dos hebras del mtDNA se replican, de forma continua, a partir de dos orígenes de replicación (O_H y O_L) ampliamente separados.

El inicio de la replicación empieza con la síntesis de un RNA cebador en LSP por la RNapolimerasa (POLRMT) (Chang & Clayton, 1985; Chang *et al.*, 1985), ayudada por TFAM que induce el desenrollamiento de la doble hebra de DNA, la topoisomerasa tipo I que relaja el DNA, la helicasa que desenrolla la doble hebra y la mtSSB que mantiene la integridad de los intermediarios replicativos de DNA de cadena simple. Los factores de transcripción TFMB1 y 2 forman heterodímeros con la POLRMT activándola, de manera que ésta transcribe un cebador empezando por el promotor de la hebra L y dando lugar a un híbrido estable de RNA-DNA llamado R-loop (Fig 1.6) (Lee & Clayton, 1996, 1997). La transición de RNA a DNA tiene lugar en el origen de replicación de la hebra H (O_H) donde el cebador precursor se escinde por una endoribonucleasa específica, la RNasa MRP (Fig 1.9a). Este enzima requiere como substrato una estructura de triple hélice como la anteriormente descrita y está formado, además de los componentes proteicos, por un RNA esencial para su actividad (Chang & Clayton, 1987, 1989; Topper & Clayton, 1990). Se ha descrito también la terminación prematura de la transcripción desde LSP causada por una secuencia conservada (CSBII) rica en G-Cs, lo que podría dar lugar al cebador para la replicación de la cadena H del DNAm_t sin necesidad de la RNasa (Pham *et al.*, 2006).

La elongación de la hebra naciente de DNA se lleva a cabo por la DNA polimerasa γ (POLG), que consta de dos subunidades, una catalítica (A) con actividad polimerasa 5'-3' y exonucleasa 3'-5' responsable probablemente de la alta afinidad descrita para esta polimerasa (Kunkel, 1985; Kunkel & Mosbaugh, 1989) y una subunidad accesoria (B) implicada en el reconocimiento del cebador y que aumenta la procesividad del enzima acelerando la velocidad de polimerización y suprimiendo la actividad exonucleasa, además de incrementar la afinidad (Fan *et al.*, 1999; Johnson & Johnson, 2001). La estructura cristalográfica del holoenzima -heterotrímico formado por una subunidad A y una subunidad B dimérica- ha

sido resuelta por Lee y col. (Lee *et al.*, 2009). Mayoritariamente, el inicio de la replicación de la hebra H termina tras una secuencia asociada a la terminación prematura (TAS, *termination-associated sequences*) (Fig 1.9a), dando lugar a una molécula de 700 nucleótidos sintetizados que permanecen asociados a la molécula circular parental formando el D-loop. Cuando la hebra H naciente es capaz de pasar a través de la región TAS, su elongación continúa de forma unidireccional hasta el origen de replicación de la hebra L (O_L) (Fig 1.6). O_L sólo se activa cuando la hebra H parental es desplazada por la hebra creciente exponiendo ésta en forma de cadena simple, de manera que adopta una estructura de loop (Fig 1.8a). En este punto, POLRMT inicia la síntesis del cebador desde el tramo de poly-dT en la región del loop de cadena simple. Después de aproximadamente 25nt, la POLRMT es desplazada por la POLG que replica la cadena L de forma unidireccional y en sentido contrario al de la hebra H naciente (Fuste *et al.*; Kasamatsu & Vinograd, 1974; Clayton, 1982). Varios métodos, incluyendo EM, protección frente a S1 y *ligation-mediated PCR* confirman la existencia de O_L (Boghenhagen *et al.*, 1979; Nass, 1980).

Modelo de replicación bidireccional y simétrico.

Este modelo se basa en geles bidimensionales neutros de DNA (2D-AGE) ampliamente utilizados para definir mecanismos de replicación (Brewer & Fangman, 1991), y propone que el DNAm_t se replica de un modo bidireccional y simétrico desde un único origen de replicación (Fig 1.8b).

En el primer trabajo publicado en el 2000 se encontraron en células humanas y de ratón, la existencia de dos tipos de intermediarios de replicación (RIs) (Holt *et al.*, 2000). Además de los RIs sensibles a la nucleasa S1, compatibles con el modelo clásico de desplazamiento de hebra, encontraron también RIs de doble cadena que sugerían un mecanismo de replicación *strand-coupled* -donde ambas hebras se replican a la vez- análoga a la observada en el DNA nuclear. Postularon una coexistencia de ambos mecanismos de replicación siendo mayoritario el mecanismo de desplazamiento de hebra en estados basales y el mecanismo bidireccional simétrico en estados de estrés.

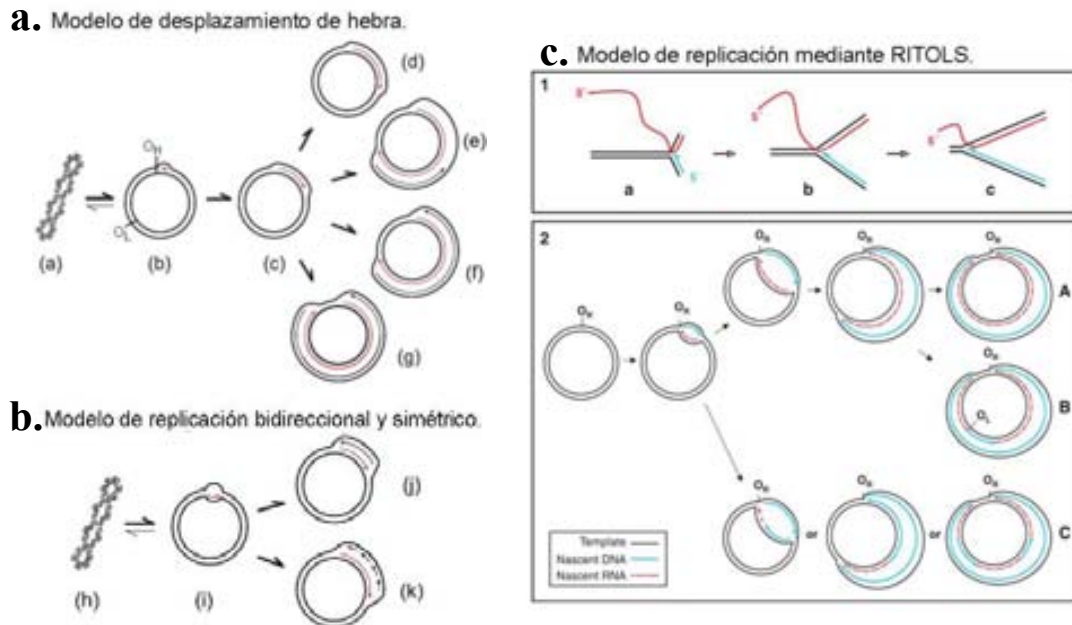


Figura 1.8_ Modelos simplificados de la replicación del DNAm. Diagrama de los intermediarios replicativos del DNAm. (a) En el modelo de desplazamiento de hebra, el DNAm circular cerrado (a) está en equilibrio con el DNA que contiene el D-loop (b). La replicación se lleva a cabo mediante la elongación de la cadena representada en rojo (c) a lo largo del genoma (d) hasta O_L (e) donde se inicia la replicación de la cadena L en sentido contrario (e, f, g). (b) En el modelo de replicación bidireccional, la replicación se inicia de una manera unidireccional en el origen O_R (localizado en O_H o en una zona distal del NCR) y comienza la síntesis de la hebra principal (rojo) (i). La hebra secundaria se establece inicialmente como RNA, que facilita la síntesis de la hebra complementaria (j), la replicación de la cual, se da a lo largo del genoma mediante los conocidos segmentos okazaki (k). (Clayton, 2003). (c) El modelo de replicación mediante RITOLS puede ser atribuido a la síntesis del RNA por una primasa o a la incorporación del RNA preformado (1). En el panel 1, a medida que avanza la horquilla (a, b, c) el RNA preformado se enhebra por el complejo de replicación en dirección 3'-5', hibridando a la vez que se da el desplazamiento de la hebra parental. En el panel 2, se muestra cómo la replicación se inicia de manera unidireccional en un origen O_R (localizado en el punto O_H o en la zona distal de NCR). La hebra *lagging* se establece inicialmente como RNA (rojo). En algunas moléculas, la hebra *lagging* entera se puede incorporar como RNA antes de que empiece la conversión a DNA (A). En otras moléculas, la maduración se puede iniciar a las 2/3 partes del genoma (B) o en lugares dispersos (C) (Yasukawa *et al.*, 2006).

Más tarde, el mismo grupo, concluía que los RIs de cadena simple observados anteriormente eran un artefacto del proceso de extracción del DNA debido a la presencia de RNasa H que degrada el RNA unido a DNA dando lugar a zonas de DNA de cadena simple interpretadas como RIs del modelo de desplazamiento de hebra. Los RIs altamente purificados que obtuvieron de mitocondria mediante gradiente de sacarosa, microscopía electrónica de transmisión e inmunopurificación, parecen ser esencialmente dúplex con una elevada presencia de híbridos DNA-RNA (Pohjoismaki *et al.*; Yang *et al.*, 2002). Este resultado parecería indicar pues, que la replicación en aves y mamíferos frecuentemente implica la incorporación de ribonucleótidos a lo largo de la cadena rezagada,

proponiendo un modelo llamado RITOLS (Ribonucleotide Incorporation Throughout the Lagging Strand), en el que la cadena lagging se establece inicialmente mayoritariamente como RNA (Yasukawa *et al.*, 2006) (Fig 1.8c).

Los experimentos en tejidos para determinar el lugar de inicio de la replicación, delimitaron como origen de replicación una amplia zona que incluía la región no codificante y también la región codificante para citocromo b y las subunidades 5 y 6 de la NADH deshidrogenada (Bowmaker *et al.*, 2003; Reyes *et al.*, 2005). Estudios con cultivos celulares, en cambio, parecen indicar mayoritariamente una replicación bidireccional que se inicia en la región no codificante (NCR), presentando una terminación temprana de una de las horquillas de replicación cerca del O_H . (Yasukawa *et al.*, 2005).

1.5.2 Transcripción del DNAm_t humano

Los primeros informes mostraron que una vez se inicia la transcripción, ambas cadenas se transcriben completamente (Aloni & Attardi, 1971; Murphy *et al.*, 1975). La cadena L se transcribe unas 2-3 veces más rápido que la cadena H (Cantatore & Attardi, 1980), aunque la mayoría de genes están codificados en esta última. La transcripción del DNAm_t se inicia a partir de tres promotores diferentes localizados en el D-loop, uno para la cadena ligera (L) y dos para la cadena pesada (H_1 y H_2) (Montoya *et al.*, 1982; Montoya *et al.*, 1983) (Fig 1.6).

H_1 está localizado 19 nucleótidos por encima del tRNA^{Phe} y H_2 está muy próximo al extremo 5' del 12S rRNA. Ambos transcritos de la cadena H se solapan en la región de los rRNAs. La transcripción iniciada en H_1 se termina en el extremo 3' del rRNA 16S, y es la responsable de la síntesis de ambos rRNAs, y de los tRNA^{Phe} y tRNA^{Val} (Montoya *et al.*, 1983; Kruse *et al.*, 1989). La otra unidad de transcripción de la cadena H, que se inicia en H_2 , se sintetiza unas 20-50 veces menos (Gelfand & Attardi, 1981) y produce una especie policistrónica gigante que cubre casi toda la cadena. El lugar de iniciación para la cadena L se encuentra en la posición extremo 5' del 7S RNA en la posición 407, aproximadamente a 150pb de H_1 , y es responsable de la transcripción de ocho tRNAs y del mRNA ND6.

Iniciación de la transcripción

Elementos cis

Ciertos estudios han permitido determinar cuáles son los requisitos para la iniciación de la transcripción. Los promotores mitocondriales que contienen los puntos de iniciación de la transcripción en H₁ y L, llamados HSP y LSP respectivamente, son funcionalmente independientes. Contienen un elemento promotor esencial para la transcripción con una secuencia consenso de 15pb (Chang & Clayton, 1984; Hixson & Clayton, 1985) y un segundo elemento localizado justo antes de las regiones promotoras (-12 a -39pb) necesario para una transcripción óptima y que puede ser considerado como un *enhancer*. Este elemento regulador contiene el lugar de unión para mtTFA y si se invierte un elemento con respecto al otro, los *enhancer* de H₁ y L presentan similitud secuencial (Fisher *et al.*, 1987). A pesar de la proximidad de ambas regiones promotoras, éstas son funcionalmente independientes (Topper & Clayton, 1989). El segundo punto de inicio de la transcripción H₂ tiene un promotor con cierta similitud a la secuencia consenso de 15pb y aparentemente no tiene el lugar de unión de mtTFA.

Elementos trans

La iniciación de la transcripción del DNA mitocondrial requiere una RNA polimerasa específica de organelo POLRMT y, como mínimo, dos factores de transcripción, mtTFA (también llamada TFAM) y uno de los dos factores B, TFB1M o TFB2M (Fig 1.9b). El fraccionamiento biofísico de extractos de transcripción mitocondrial permitió en la década de los 80 la identificación de dos proteínas codificadas en núcleo, requeridas en la iniciación de la transcripción. La RNA polimerasa POLRMT (Shuey & Attardi, 1985) y un factor de transcripción mtTFA que promueve un incremento de la actividad y confiere selectividad de promotor a la polimerasa (Fisher & Clayton, 1985, 1988).

El dominio C-terminal de la POLRMT presenta una gran homología con las polimerasas de los bacteriófagos T3, T7 y SP6 (Masters *et al.*, 1987; Tiranti *et al.*, 1997). Estudios estructurales y mutacionales indican que esta

región conservada es importante para la selectividad del promotor y para la actividad polimerasa (Gardner *et al.*, 1997). Experimentos en levadura sugieren que la parte N-terminal de la polimerasa interactúa con factores involucrados en eventos post-transcripcionales y asociados a la membrana interna (Rodeheffer & Shadel, 2003), lo que estaría de acuerdo con la disposición de los nucleoides mencionados anteriormente. Se ha encontrado en núcleo una forma de splicing alternativo de la POLRMT en que le faltan 262 aminoácidos del Nterminal (Kravchenko *et al.*, 2005). Esta polimerasa también conocida como spRNAP-VI, transcribe ciertos genes nucleares, pero parece ser que éstos no forman ningún cluster funcional. Así pues, la importancia de esta forma nuclear de la POLRMT sobre la regulación o la dinámica mitocondrial permanece poco clara.

mtTFA está formada por dos dominios HMG box (high mobility group) separados por un *linker* de 27 aminoácidos y una cola C-terminal de 25 residuos. Como el resto de la familia de las HMG, puede envolver, curvar y desenrollar el DNA *in vitro* con un bajo grado de especificidad (Parisi & Clayton, 1991; Fisher *et al.*, 1992). Esta propiedad, además de su abundancia en mitocondria, sugiere que TFAM juega un papel importante en la estabilización y el mantenimiento del cromosoma mitocondrial. Además, la caracterización bioquímica ha revelado que la cola C-terminal es importante en el reconocimiento específico del DNA y esencial en la activación de la transcripción (Dairaghi *et al.*, 1995a). Los lugares específicos de unión de mtTFA al DNAm_t se encuentran antes de los puntos de iniciación de la transcripción H₁ y L, pero la unión y la estimulación de la transcripción son mayores para el promotor L que para el H₁ (Fisher *et al.*, 1987; Fisher & Clayton, 1988; Ghivizzani *et al.*, 1994). Se sugiere que mtTFA forma un complejo con el DNA mediante sus dos dominios HMGbox induciendo un cambio estructural específico en la región del promotor de manera que permite la iniciación de la transcripción por la polimerasa (Shadel & Clayton, 1997).

En levadura, el factor de iniciación específico mtTF1 (Xu & Clayton, 1992) era diferente de mtTFA y tenía un homólogo en xenopus llamado

mtTFB, lo que apuntaba la existencia de factores de transcripción adicionales en vertebrados. Falkenberg y col. identificaron en humanos y ratones el factor de transcripción TFB1M y un segundo factor relacionado, TFB2M (Falkenberg *et al.*, 2002) ambas homólogas a RNA metiltransferasas (McCulloch *et al.*, 2002). TFB2M muestra una estimulación de la actividad transcripcional específica de casi dos órdenes de magnitud superior a TFB1M. Ambos factores parecen interactuar con la polimerasa formando un heterodímero y, se requiere uno u otro factor para la iniciación de la transcripción *in vitro* en los promotores H₁ y L aunque parecen no ser necesarios en la elongación.

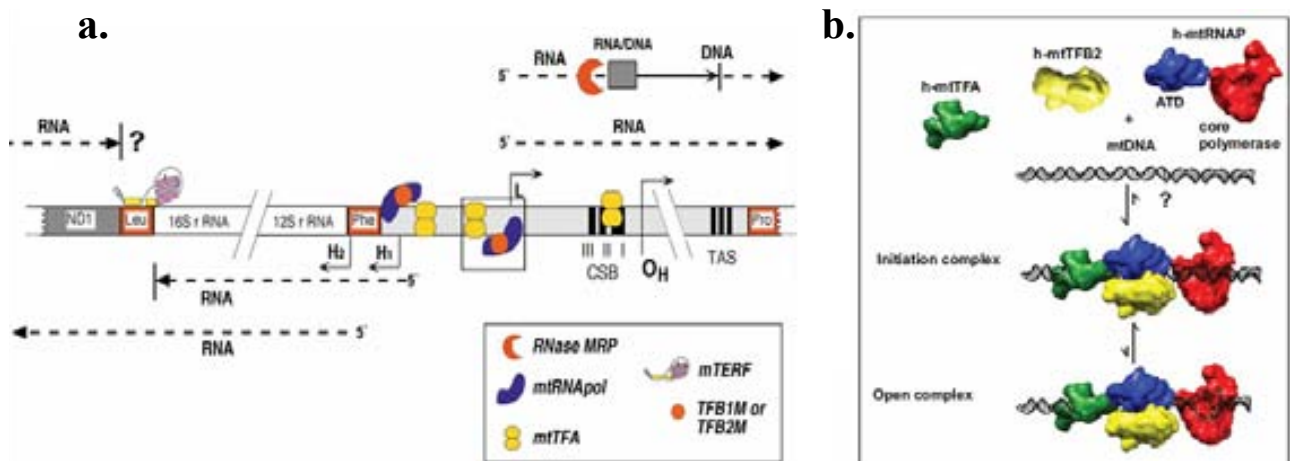


Figura 1.9_ (a) Representación esquemática del D-loop (en gris claro se muestra la región no codificante) y las regiones de terminación en el DNA mitocondrial de mamíferos. Se muestran los principales elementos y factores implicados en la transcripción y la iniciación de la replicación. El punto de iniciación H1 dirige la transcripción de la región rRNA y su actividad está unida a la terminación de la transcripción en la región del tRNA^{Leu} mediante mTERF. El punto de iniciación H2, dirige la transcripción de toda la cadena H. A partir del punto de iniciación L, se transcribe la cadena L y el RNA precursor del inicio de la replicación de la cadena H. La RNasa MRP es una ribonucleoproteína involucrada en el procesamiento de los transcritos primarios de la cadena L, dando lugar al cebador para la iniciación de la replicación de la cadena H. La replicación de la hebra H se termina mayoritariamente en la región TAS, dando lugar a la estructura del D-loop. (Fernandez-Silva *et al.*, 2003) (b) Modelo del inicio de transcripción mitocondrial. Las interacciones entre h-mtTFA y h-mtTFB2 y entre mtTFB2 y h-mtRNAP (ATD Amino Terminal Domain y core polimerase) se dan a través de los extremos carboxi-terminales de h-mtTFA y h-mtTFB2. Los factores de transcripción y la polimerasa se unen al DNA formando un complejo de iniciación que, una vez formado es bastante estable. La formación del primer enlace fosfodiéster requiere grandes concentraciones de ATP, por tanto la eficiencia de este paso está controlada por la disponibilidad de ATP. Una vez formado el dinucleótido, la dependencia de la concentración de ATP en la reacción disminuye en un factor de 10. La transición al complejo de elongación ocurre después de la formación del trinucleótido y es diferente en los dos promotores (Lodeiro *et al.*).

Terminación de la transcripción

Se han realizado ensayos con DNA mitocondrial de rata para aislar regiones que puedan actuar como terminadores de la transcripción *in vivo* en células bacterianas. Se aislaron tres regiones, pRMT1 entre el D-loop y el 12S rRNA, pRMT3 en el citocromo b y pRMT5 en el extremo 3' del gen RNA ribosomal 16S. Parece ser que pRMT3 actúa como elemento terminador independiente de Rho mientras los otros dos necesitan la proteína bacteriana para una terminación eficiente de la transcripción (Staub & Castora, 1993). El lugar de terminación entre los genes 16SRNA y tRNA^{Leu(URR)} había sido descrito ya previamente mediante ensayos de transcripción *in vitro* (Christianson & Clayton, 1988). Además, el análisis cinético de los transcritos en células HeLa muestra que la velocidad de transcripción es de 50-100 veces mayor para los rRNA en comparación con el resto de mRNAs de la cadena H (Gelfand & Attardi, 1981).

El papel principal de esta atenuación se debe al factor de terminación de la transcripción mTERF (Kruse *et al.*, 1989), del cual se hablará en detalle más adelante. Las evidencias de terminación en la zona entre 16SrRNA y tRNA^{Leu(URR)} se debe a la existencia en células HeLa de dos unidades de transcripción que se superponen con propiedades cinéticas distintas (Montoya *et al.*, 1983), al hecho de que el extremo de las moléculas de 3' de 16SrRNA muestran heterogeneidad (Dubin *et al.*, 1982) contrastando con la precisión del extremo de los transcritos procesados y a numerosos experimentos *in vitro* que muestran una terminación de la transcripción específica en ese punto (Kruse *et al.*, 1989; Micol *et al.*, 1996). Curiosamente, los ensayos de transcripción *in vitro* con mTERF muestran una mayor eficiencia de terminación en el sentido opuesto al lugar de iniciación H₁ (Christianson & Clayton, 1986; Asin-Cayuela *et al.*, 2005) (Fig 1.9a). Este fenómeno podría implicar que mTERF, además de atenuar la transcripción de la cadena H, podría terminar la transcripción de la cadena L en un lugar a partir del cual no hay codificados más genes. Además, se ha descrito un nuevo lugar de unión de mTERF a una región cercana al HSP1 (lugar de inicio de la transcripción) del DNAm (Prieto-

Martin *et al.*, 2004). La unión de mTERF a este lugar activa la transcripción *in vitro* y experimentos de microscopía electrónica parecen indicar que el DNA forma un bucle entre los dos lugares de unión de mTERF. Este bucle acerca a ambos extremos de la transcripción de HSP1 y podría suponer el control de ésta atrapando la maquinaria de transcripción en su interior y facilitando el reinicio de la misma (Martin *et al.*, 2005) (Fig 1.15).

Se ha identificado también un lugar de terminación para el transcrito H₂ justo después de la región control y antes del gen que codifica para el tRNA^{Phe} (que coincidiría con pRMT1). Se han asociado con esta región dos proteínas de 45 y 70 KDa, y se ha identificado una de ellas (Sondheimer *et al.*; Camasamudram *et al.*, 2003).

Procesado y maduración del RNA

La transcripción del DNA mitocondrial da lugar a largas moléculas policistrónicas, que requieren de endonucleasas para liberar las diferentes especies de RNA del transcrito primario con el objetivo que éstas puedan ser funcionales. Los transcritos primarios se caracterizan por contener tRNAs entre las secuencias de mRNAs o rRNAs, así pues se postuló el modelo denominado por puntuación de tRNAs (Ojala *et al.*, 1981). En este modelo, las largas moléculas policistrónicas se procesan por cortes endonucleolíticos en los tRNAs que actúan como señales de reconocimiento al adquirir la conformación en hoja de trébol. En los pocos casos en que no hay un tRNA flanqueando la terminación del mRNA, probablemente existe una estructura secundaria parecida al tRNA que puede ser reconocida por la maquinaria de procesamiento. Así pues, la escisión de los tRNAs del transcrito primario policistrónico da lugar no sólo a tRNAs maduros sino también a mRNAs y rRNAs maduros.

El procesado requiere como mínimo tres pasos; el corte endonucleolítico 5' y 3' del tRNA, la poliadenilación de los rRNAs y los mRNAs y la adición de CCA al extremo 3' del tRNA, ya que éste no está codificado (Fig 1.10).

En referencia al procesado de los tRNAs, se ha caracterizado en células HeLa actividad RNasa P, un enzima que también existe en núcleo y es responsable del corte endonucleolítico en el extremo 5' (Doersen *et al.*, 1985; Rossmannith *et al.*, 1995; Puranam & Attardi, 2001). A diferencia de el resto de RNasas P caracterizadas, la RNasa P mitocondrial humana no requiere el RNA para la catálisis y, además, parece ser el resultado de la combinación de varias proteínas multifuncionales (Holzmann *et al.*, 2008). En mitocondria de hígado de rata se identificó actividad en 5' y 3' de procesado de tRNA (Manam & Van Tuyle, 1987). Parece ser que la tRNasa Z (familia ELAC) sería el enzima que realiza el corte endonucleolítico en el extremo 3' (Schiffer *et al.*, 2002), aun así esta actividad no ha sido identificada inequívocamente. El enzima que cataliza la adición de CCA al extremo 3' del tRNA es la ATP(CTP):tRNA nucleotidil-transferasa (Nagaike *et al.*, 2001). El acoplamiento del aminoácido al extremo CCA del tRNA maduro se produce mediante el enzima aminoacil-tRNA sintetasa. Se han investigado otras actividades para la modificación y maduración de los tRNAs (Helm *et al.*, 2000; Nagaike *et al.*, 2001).

Mediante una cromatografía de oligo dT-celulosa se pueden separar los mtRNAs en dos fracciones; los no poliadenilados y los poliadenilados que contienen una cola poli-A en el extremo 3' de aproximadamente 55 nucleótidos (Hirsch & Penman, 1974; Amalric *et al.*, 1978; Montoya *et al.*, 1981). Esta cola es menor que la observada en los mRNAs citoplasmáticos y no está codificada en el DNA, sino que es añadida después de la transcripción. Hay evidencias de que la poliadenilación de los rRNAs y mRNAs, catalizada por una poli(A) polimerasa mitocondrial, ocurre inmediatamente después del corte (Amalric *et al.*, 1978). La poliadenilación en mitocondria es importante ya que en algunos casos contribuye a la generación del codón de terminación. Existe cierta controversia sobre la función de la poliadenilación en la estabilización/degradación del RNA (Slomovic *et al.*, 2005).

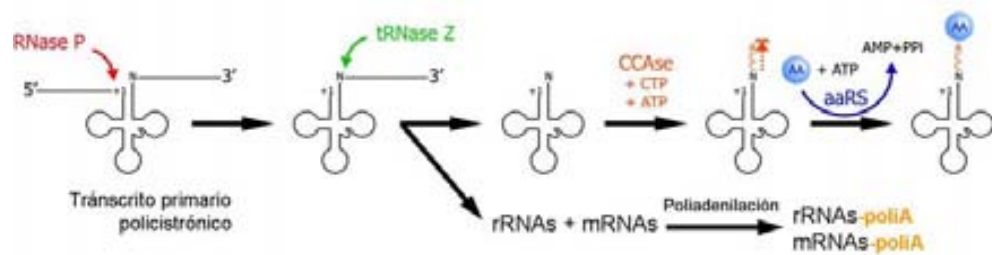


Fig 1.10_ Esquema del procesado del transcrito primario policistrónico. Éste se caracteriza por la presencia de tRNAs entre las secuencias de mRNAs o rRNAs, de manera que las largas moléculas policistrónicas se procesan por cortes endonucleolíticos en los tRNAs mediante la Rnasa P que corta en el extremo 5' (+1) y la tRNasa Z que corta en el extremo 3' (N; +73), liberando las diferentes especies de RNA. El enzima CCAse cataliza la adición de CCA en el extremo N de las moléculas de tRNA, donde posteriormente se acopla el aminoácido. Figura adaptada de (Levinger *et al.*, 2004).

Regulación de la transcripción

En mamíferos se han descrito cambios en la actividad transcripcional y en los niveles de RNAs mitocondriales en diversas situaciones tales como los diferentes estadios de desarrollo, la edad, los cambios en la demanda energética celular o los estados hormonales. Sin embargo, se sabe poco sobre los mecanismos moleculares involucrados en la regulación de la transcripción mitocondrial responsables de estos cambios.

La proliferación mitocondrial requiere la regulación coordinada de la expresión de los genes nucleares y mitocondriales. Así pues, la regulación se puede dar a nivel nuclear y/o mitocondrial.

Parece ser que existen dos tipos de factores nucleares, los *activadores* que son factores de transcripción (factores nucleares de respiración NRF-1 y NRF-2, Sp-1...) y los *coactivadores* que son factores nucleares (PGC-1, PRC) que interaccionan con los factores de transcripción unidos a DNA para regular la transcripción génica en respuesta a señales fisiológicas de termogénesis y proliferación celular (Wu *et al.*, 1999; Andersson & Scarpulla, 2001; Carmona *et al.*, 2002; Kelly & Scarpulla, 2004).

Son comunes los lugares de reconocimiento para NRF-1, NRF-2 y Sp-1 en varios genes nucleares que codifican para subunidades de la cadena respiratoria, en genes que codifican para factores mitocondriales de transcripción y replicación así como en ciertos enzimas y componentes de la maquinaria de transporte proteico (Scarpulla, 1997, 2008).

Un mecanismo potencial de regulación es mediante la diferente estimulación transcripcional de los factores h-mtTFB1 y h-mtTFB2 (Falkenberg *et al.*, 2002; Gaspari *et al.*, 2004; Montoya *et al.*, 2006). Los niveles de TFAM también juegan un papel importante, por ejemplo, se necesitan concentraciones mayores de TFAM para activar HSP que LSP (Shutt *et al.*), y elevadas concentraciones de TFAM inhiben la transcripción *in vitro* (Fisher & Clayton, 1988; Dairaghi *et al.*, 1995b).

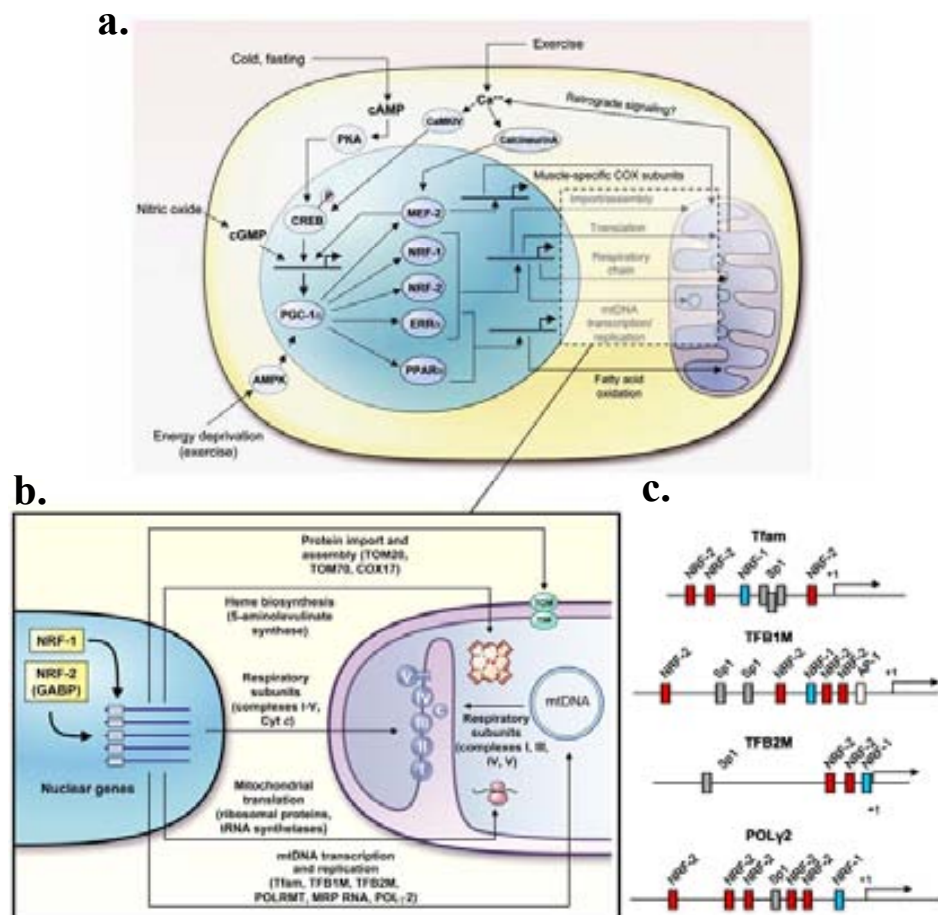


Fig 1.11_ (a) Ilustración resumen de las vías mitocondriales mediadas por el coactivador PGC-1 α . En el núcleo se representan los factores de transcripción (NRF-1, NRF-2, ERR α , PPAR α , y MEF-2) que son diana de PGC-1 α y que actúan en genes nucleares involucrados en funciones mitocondriales. (b) Diagrama esquemático del control de expresión génica nuclear mediante NRF-1 y NRF-2 (GABP). Los NRFs contribuyen directa e indirectamente a la expresión de genes importantes en el mantenimiento y la funcionalidad del aparato respiratorio mitocondrial; como los genes que codifican para el citocromo c, la mayoría de los genes que codifican para las subunidades de los complejos respiratorios I-V, o los genes que codifican para componentes del transporte proteico y de la maquinaria de ensamblaje. Además, promueven la expresión de componentes claves de la maquinaria de transcripción y traducción necesaria para la expresión de las subunidades de los complejos respiratorios codificados en mitocondria. (c) Disposición de los lugares de reconocimiento de NRF-1, NRF-2 y Sp1 en los promotores humanos de TFAM, TFB1, TFB2 y la polimerasa (Scarpulla, 2008).

Usando mitocondrias aisladas como modelo se ha demostrado que puede existir cierto grado de autonomía en la regulación de la transcripción del DNAm. Así, en ausencia de expresión nuclear de genes, la síntesis y la maduración de los mRNAs mitocondriales se puede mantener durante largos periodos de tiempo. Además, algunas señales externas como las concentraciones de ATP o de hormona tiroidea pueden inducir cambios en la transcripción mitocondrial (Enriquez *et al.*, 1996; Enriquez *et al.*, 1999).

Como la expresión genética mitocondrial es esencial para la OXPHOS, existe un gran interés para determinar cómo la transcripción mitocondrial está coordinada a las necesidades energéticas de la célula. Se ha demostrado que el ATP juega un papel importante en la transcripción, ya que se requiere de 15-20 veces más de ATP que de los demás NTPs (Shuey & Attardi, 1985). Parece ser que la transcripción desde LSP o HSP es diferente según la disponibilidad de ATP (Gaines *et al.*, 1987; Narasimhan & Attardi, 1987). Además, experimentos con ATP no hidrolizable muestran que durante la primera etapa de la iniciación de la transcripción, se requiere la hidrólisis de ATP (Fig1.9b).

El papel de las hormonas tiroideas es importante porque son clave en el metabolismo energético. El hipotiroidismo induce una disminución de la concentración de todos los RNAs mitocondriales, pero mayoritariamente de los mRNAs, de manera que produce un cambio también en la relación mRNA/rRNA (Mutvei *et al.*, 1989). Mediante el estudio *in organello*, se ha encontrado un efecto directo de la concentración de la hormona tiroidea sobre los niveles de RNA. Los patrones de *footprinting* del DNA mitocondrial en orgánulos son diferentes en los lugares de iniciación de la transcripción según la concentración de hormona tiroidea, pero se mantienen iguales en el lugar de terminación. La modulación selectiva de los lugares alternativos de inicio de transcripción de la cadena H podría explicar la influencia de la hormona en la relación mRNA/rRNA (Enriquez *et al.*, 1999). Además, existe en mitocondria un receptor T3 (p43) que es una forma truncada del receptor nuclear c-Erb A α .

1(Wrutniak *et al.*, 1995; Wrutniak-Cabello *et al.*, 2001), y que se une a ciertas secuencias de DNA mitocondrial localizadas en el D-loop, muy similares a los elementos nucleares de respuesta a T3 estimulando la transcripción mitocondrial en *organello* e *in vivo* cuando se sobreexpresa en las células (Casas *et al.* 1999 (Casas *et al.*, 1999; Casas *et al.*, 2008).

Proteínas dirigidas a mitocondria	Referencia
Receptores nucleares	
Receptores de estrógenos (α y β)	Yang <i>et al.</i> 2004; Chen <i>et al.</i> 2004
c-ErbA α 1 (receptor de hormona tiroidea)	Wrutniak <i>et al.</i> 1995 and 2001
Receptor glucocorticoide (GR)	Ioannou <i>et al.</i> 1988; Demonacos <i>et al.</i> 1996
Nur 77	Liu <i>et al.</i> 1994; Li <i>et al.</i> 2000; Lin <i>et al.</i> 2004
PPAR γ 2	Casas <i>et al.</i> 2000, Smith and Muscat 2005
Factores de transcripción	
AP1 (c-fos/c-jun)	Ogita <i>et al.</i> 2002 and 2003
CREB	Cammarota <i>et al.</i> 1999, Lee <i>et al.</i> 2005; Ryu <i>et al.</i> 2005
NF-kB	Cogswell <i>et al.</i> 2003
p53	Marchenko <i>et al.</i> 2000
TFAM (mtTFA)	Parisi <i>et al.</i> 1991
TFB1M	Falkenberg <i>et al.</i> 2002
TFB2M	Falkenberg <i>et al.</i> 2002

Fig 1.12_ **Receptores nucleares y factores de transcripción que se localizan en mitocondria.** Se han detectado en mitocondrias de diferentes tipos de células, receptores de glucocorticoides, estrógenos, andrógenos y hormonas tiroideas mediante western blot, inmunofluorescencia, microscopía confocal y microscopía electrónica con anticuerpos con oro. Evidencias experimentales confirman que los factores de transcripción localizados en mitocondria, actúan en la transcripción mitocondrial, el rendimiento energético y la apoptosis. La coordinación de la activación de la transcripción en núcleo y mitocondria por los respectivos receptores se realiza en parte por su unión a elementos trans comunes en los dos genomas. Así pues, los mismos receptores nucleares localizados en los dos compartimentos (mitocondria y núcleo), regulan la transcripción de genes para un fin común. (Lee *et al.*, 2008; Psarra & Sekeris, 2008a, b).

1.5.3 Traducción del DNAm_t humano

La principal función de la maquinaria de traducción mitocondrial es proveer algunas de las proteínas del sistema de fosforilación oxidativa (OXPHOS). Esta maquinaria está compuesta por componentes codificados por el DNAm_t y por el

nuclear, de manera que la biogénesis de este sistema depende de la expresión coordinada de ambos genomas. En el caso de la traducción algunos aspectos continúan sin resolverse, y hasta hoy, todos los intentos de traducción in vitro de mRNAs mitocondriales con los correspondientes tRNAs, ribosomas y factores de traducción han sido fallidos.

El código genético mitocondrial presenta algunas diferencias respecto al código genético universal.

	Universal	DNAMt humano
UGA	STOP	Triptófano
AUA	Isoleucina	Metionina
AUU	Isoleucina	Metionina
AGG, AGA	Arginina	STOP

El genoma mitocondrial humano presenta un inusual sistema de reconocimiento de codón que permite la lectura propicia de todo el código genético con sólo 22 tRNAs. Este sistema se basa en el reconocimiento de las dos primeras bases del codón (Lagerkvist, 1978). Para ocho de los aminoácidos la tercera base no es discriminante y para el resto, el aminoácido varía según si la tercera base es A-G o U-C, de manera que sólo existe un tRNA para cada aminoácido.

Los ribosomas mitocondriales fueron originalmente aislados en la segunda mitad de los años 60 (O'Brien & Kalf, 1967). Están formados por los correspondientes componentes proteicos (80 aprox.) codificados en el DNA nuclear y los rRNAs codificados en el DNAMt. Estos rRNAs son más pequeños que los citosólicos y procarióticos de manera que los ribosomas mitocondriales, además de las correspondientes proteínas homólogas a procariotas, presentan nuevas proteínas para sustituir las funciones perdidas por esta disminución de tamaño del rRNA (O'Brien & Denslow, 1996).

Como se ha mencionado anteriormente, dada la alta compactación del DNAMt, los mRNAs mitocondriales no presentan regiones en 5' o 3' que no se traducen (UTRs) ni caps 5' (Grohmann *et al.*, 1978; Montoya *et al.*, 1981). Estos elementos son cruciales en la determinación de la localización intracelular del mRNA, su estabilidad y la eficiencia con la cual se transcribe y por tanto su ausencia crea cierta controversia. Así pues, los RNAs mensajeros mitocondriales no tienen secuencias líder que faciliten la unión a ribosomas lo que podría explicar la baja

eficiencia de la traducción mitocondrial (Cantatore *et al.*, 1987). En experimentos *in vitro* con ribosomas mitocondriales bobinos, la subunidad pequeña (28S) parece que une fuertemente mRNAs independientemente de la secuencia y en ausencia de factores auxiliares de iniciación. La unión se produce en fragmentos a partir de 30-80 nucleótidos, pero son necesarios unos 400 nucleótidos para que la unión sea estable (Denslow *et al.*, 1989; Liao & Spremulli, 1990a, 1991). En el caso de levadura, en cambio, se ha demostrado la importancia de los 5'UTRs mitocondriales.

Numerosas evidencias indican que la traducción es primordialmente un proceso asociado a la membrana interna (Spithill *et al.*, 1978; Marzuki & Hibbs, 1986; Sanchirico *et al.*, 1998; Jia *et al.*, 2003; Szyrach *et al.*, 2003) facilitando así la inserción en la membrana de los péptidos nacientes involucrados en los complejos de OXPHOS. Todas estas observaciones sugieren que la transcripción, la traducción y la inserción en la membrana están coordinadas (Fig 1.7).

La síntesis de proteínas hidrofóbicas de membrana en un entorno hidrofílico corre el riesgo de la formación de agregados proteicos. Este problema se evita mediante un acoplamiento estricto de la síntesis y la integración en la membrana de la proteína. Así, de un modo análogo a procariontes, los mRNAs mitocondriales parece que se traducen en ribosomas unidos a la membrana mitocondrial. En levaduras, se ha demostrado que el dominio C-terminal de la proteína Oxa1 se une al ribosoma mitocondrial y acopla físicamente el aparato de traducción mitocondrial con la membrana interna (Jia *et al.*, 2003; Szyrach *et al.*, 2003).

Se han purificado varios factores de traducción mitocondriales animales. Dos factores de iniciación MTIF2 (Liao & Spremulli, 1990b) y MTIF3 que estaría implicado en la discriminación del codón de iniciación y podría actuar como factor de disociación ribosomal (Koc & Spremulli, 2002). Y tres factores de elongación (EF-Tu_{mt}, EF-Ts_{mt} y EF-G_{mt}) homólogos a los factores de elongación en *E. coli*. En 1999 y 2004 se resolvieron respectivamente las estructuras cristalográficas del factor de elongación mitocondrial bovino Tu y del complejo Tu-Ts (Andersen *et al.*, 2000; Jeppesen *et al.*, 2005).

2_ TERMINACIÓN de la TRANSCRIPCIÓN

La terminación de la transcripción incluye normalmente tres etapas, la parada del complejo de elongación, la liberación del transcrito nascente de RNA y la disociación de la RNA polimerasa del DNA. Se pueden distinguir tres mecanismos generales de terminación de la transcripción (Richardson, 1993) (Fig 1.13).

El primero se debe a la secuencia intrínseca del RNA sintetizado que normalmente consiste en dos segmentos simétricos ricos en GC con capacidad de aparearse seguidos de una región rica en ATs (Fig 1.13a). Los segmentos simétricos favorecen la formación de un hairpin que desestabiliza la interacción de la molécula nascente de RNA con la RNA polimerasa, lo que da como resultado, gracias a la zona posterior de unión $A_{DNA}-U_{RNA}$ más débil, la disociación de la polimerasa y el RNA (Wilson & Von Hippel, 1995).

En el segundo mecanismo, una proteína que une RNA es la responsable de la disociación del RNA nascente (Fig 1.13b). La proteína que ejemplifica este mecanismo y que ha sido más ampliamente estudiada es la proteína bacteriana rho. Rho, es una proteína hexamérica en forma de anillo con actividad ATPasa y helicasa. En la terminación de la transcripción, Rho une el mRNA libre, preferentemente en zonas ricas en citosina. La actividad ATPasa se activa con la unión a mRNA y proporciona la energía para la translocación de Rho a lo largo del RNA. Cuando se produce el estancamiento del complejo de elongación debido a las estructuras secundarias que forma el RNA, la actividad helicasa de Rho separa el híbrido RNA-DNA y produce la liberación del transcrito (Banerjee *et al.*, 2006).

El tercer mecanismo requiere la acción de una proteína de unión específica a DNA, cuyo lugar de unión se encuentra justo por delante del extremo 3' del transcrito nascente (Fig 1.13c). Un ejemplo de este mecanismo es el factor de terminación para la polimerasa I, TTF-I (Bartsch *et al.*, 1988; Langst *et al.*, 1997). La secuencia específica de unión a DNA contiene 18 pb (Grummt *et al.*, 1986a). TTF-I se une como monómero al DNA y su interacción induce una torsión al DNA de aprox 40° (Smid *et al.*, 1992). Además, se ha descrito que el factor de terminación aumenta la transcripción *in vivo* e *in vitro* desde un promotor adyacente (Henderson & Sollner-Webb, 1986; Mcstay & Reeder, 1986). Las

mutaciones en la zona de terminación que impiden la correcta terminación también provocan una reducción de la cantidad de transcritos desde el promotor adyacente (Grummt *et al.*, 1986b; Mcstay & Reeder, 1986). Parece ser que la activación de la transcripción mediante TTF-I es específica de cromatina y requiere el posicionamiento preciso de la secuencia terminadora en relación al promotor (Langst *et al.*, 1998).

Se ha sugerido que TTF-I está implicado en la prevención de la colisión entre la maquinaria de replicación y de transcripción, ya que existen estudios que muestran que en mamíferos, la terminación específica de secuencia de la replicación del DNA está catalizada mediante un complejo que incluye TTF-I que actúa como contrahelicasa (Gerber *et al.*, 1997; Putter & Grummt, 2002).

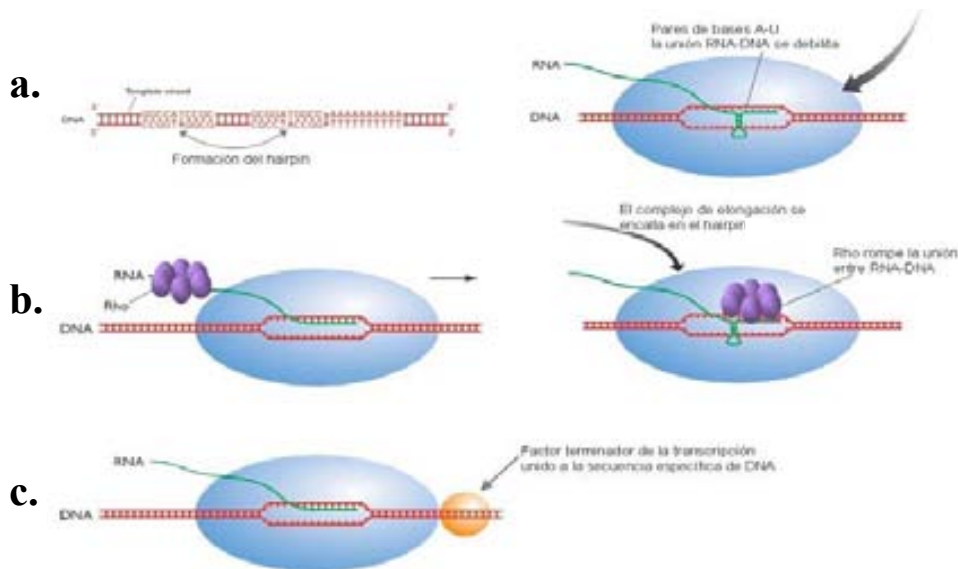


Fig 1.13_ Mecanismos generales de terminación de la transcripción. (a) En la terminación intrínseca, las secuencias palindrómicas del RNA sintetizado forman un *hairpin* que promueve la desestabilización de la unión de la polimerasa al transcrito naciente favorecida además por la zona posterior rica en A-U que provoca la liberación del RNA y por tanto la terminación de la transcripción. Existen ciertas proteínas como la NusA, que favorecen la terminación mediante la estabilización del *hairpin* de RNA. (b) En el segundo mecanismo, el transcrito suele presentar también secuencias palindrómicas, pero el *hairpin* formado es menos estable. Así pues, es necesaria la actividad de una proteína, Rho en este caso, que se une al transcrito y se mueve a lo largo de él hacia la polimerasa. Cuando el complejo de elongación de la transcripción se encalla por la presencia del *hairpin*, la actividad helicasa de Rho interfiere en el complejo DNA-RNA y permite la liberación del transcrito del DNA. (c) El tercer mecanismo de terminación de la transcripción implica la unión de un factor a una secuencia específica de DNA justo después del extremo 3' del transcrito que impide la elongación de la transcripción. Figura adaptada de (Brown, 2006).

La terminación de la transcripción del DNA mitocondrial en mamíferos parece seguir el tercer mecanismo. La unión de mTERF induce una curvatura en el DNA y la terminación se puede producir en ambos sentidos (Christianson & Clayton, 1986; Shang & Clayton, 1994). Se podría suponer, por tanto, que la terminación de mTERF se debe a

una barrera física más que a una interacción específica con el enzima. Aunque existen otras observaciones que no coinciden con esta suposición, como que la eficiencia de terminación es diferente dependiendo de la RNAPolimerasa (Shang & Clayton, 1994) y que en experimentos *in vitro* existen ciertas condiciones (como la concentración de KCl) que no afectan la unión al DNA pero sí a la terminación (Fernandez-Silva *et al.*, 1997).

2.1 Factores terminadores de la transcripción mitocondrial

2.1.1 mTERF

En 1989 fue aislada una fracción proteica de lisados de mitocondria de células HeLa que promovía la terminación específica de la transcripción de la cadena pesada. Mediante estudios de *footprinting* con DNAsa se observó que protegía una región de 28pb en el extremo 3' del 16S rRNA del DNA mitocondrial (Kruse *et al.*, 1989). Esta región protegida contenía la secuencia crítica de las 13 pares de bases necesarias para la terminación de la transcripción (Christianson & Clayton, 1988) (Fig 1.15). Cuatro años más tarde se identifican en la fracción tres polipéptidos que pertenecen a la misma proteína, dos de ~34KDa (electroforesis SDS) asociados con la actividad terminadora y uno de ~31KDa (electroforesis SDS) sin actividad terminadora y que presenta una menor afinidad por la secuencia de reconocimiento (Daga *et al.*, 1993).

En los ensayos de transcripción *in vitro* se mostró que mTERF unida al lugar de terminación en el DNA mitocondrial presentaba actividad terminadora en ambos sentidos e incluso mostraba una mayor eficacia de terminación en el sentido inverso relativo al inicio de transcripción H1 con ambos promotores HSP y LSP (Christianson & Clayton, 1986; Asin-Cayuela *et al.*, 2005). Este hecho, unido a la ausencia de genes en la cadena L en zona posterior al lugar de unión de mTERF podría sugerir una posible función de terminación de la transcripción de la cadena L. mTERF, además, podía terminar la transcripción con RNA polimerasas heterogéneas (Shang & Clayton, 1994). Estudios de *footprinting in vivo* en la región de terminación del DNAm humano revelaron un fuerte patrón de protección,

indicando una elevada ocupación (aproximadamente 80%) del lugar de terminación por mTERF (Micol *et al.*, 1997).

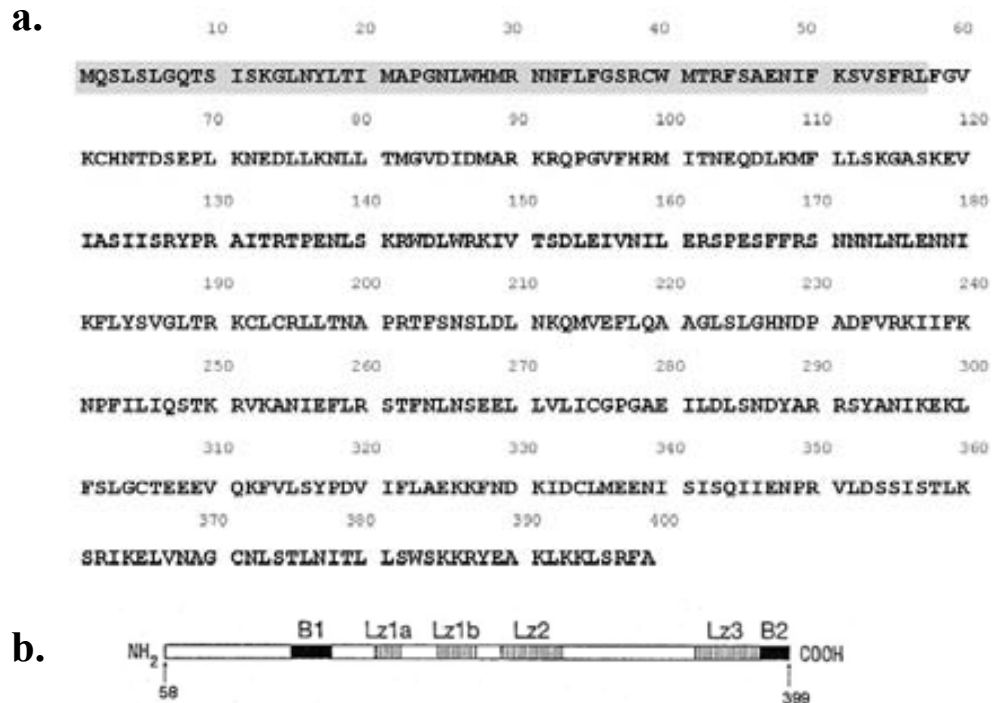


Fig 1.14_ (a) Secuencia de mTERF. En sombreado se indica la secuencia de reconocimiento de mitocondria (MLS) que es proteolizada una vez la proteína entra en el orgánulo. (b) Esquema de la disposición de las tres potenciales cremalleras de leucina y los dos dominios básicos necesarios para la unión a DNA predichos mediante el estudio computacional de mTERF madura (sin MLS), llevado a cabo por Fernández-Silva y col. (Fernández-Silva *et al.*, 1997).

En 1997 se clonó mTERF -la forma madura tiene 342 aminoácidos- y se caracterizó su unión al DNA. Un estudio computacional predijo tres potenciales cremalleras de leucina y dos dominios básicos (Fig 1.14). Los ensayos de mutagénesis demostraron que estos rasgos estructurales eran críticos para la unión específica al DNA, al que se unía como monómero (Fernández-Silva *et al.*, 1997). Experimentos de permutación circular indicaron que el complejo inducía una curvatura en el DNA de aproximadamente 35° (Shang & Clayton, 1994). Otros estudios demostraron que mTERF unía fuertemente la secuencia específica de terminación DNA_{TER} con una vida media de 40-50min (Chomyn *et al.*, 1992), y que además presentaba una fuerte actividad de unión a secuencias no específicas de DNA con una K_d aparente de unos 16nM aproximadamente (Park *et al.*, 2007).

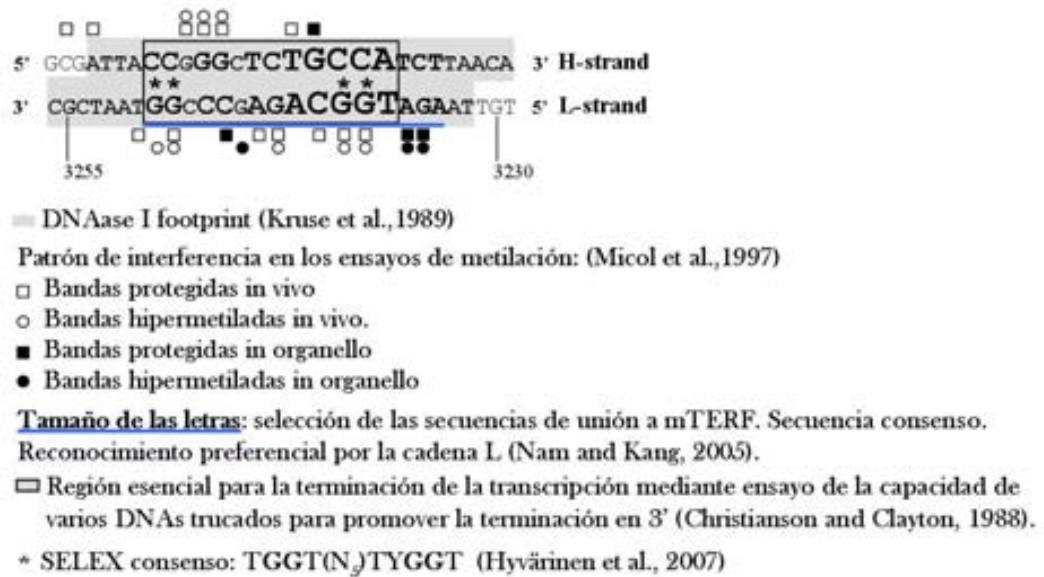


Fig 1.15_ Secuencia de DNA_{TER} mitocondrial humano y los diferentes estudios de la interacción con mTERF.

El mecanismo de regulación de mTERF sigue siendo una incógnita, se ha propuesto la oligomerización como mecanismo de regulación de la unión a DNA, un equilibrio trímero-monómero en el que únicamente el monómero une DNA (Asin-Cayuela *et al.*, 2004). Además, se ha descrito que mTERF en rata, aunque mantiene su capacidad de unión a DNA, sólo tiene actividad terminadora cuando está fosforilada (Prieto-Martin *et al.*, 2004), mientras que parece ser que mTERF humana *in vitro*, es activa en la forma no fosforilada (Asin-Cayuela *et al.*, 2005).

Evidencias recientes muestran que el rol de mTERF es probablemente más complicado. En varios ensayos de transcripción *in vitro* se observó un aumento en los niveles de transcripción al aumentar la cantidad de la fracción purificada de mitocondria de mTERF (Kruse *et al.*, 1989; Asin-Cayuela *et al.*, 2004). Pero únicamente se observa esta estimulación cuando la región promotora y terminadora están en la misma orientación (Asin-Cayuela *et al.*, 2005). En 2004, se advirtió un nuevo lugar de unión para la proteína recombinante madura mTERF de rata, que muestra capacidad de unión al correspondiente promotor HSP. Aun así, los resultados de los geles de retardo no contemplaban la formación de un complejo triple mTERF_{Frata}, DNA_{Ter} y HSP (Prieto-Martin *et al.*, 2004). Martin y col. confirmaron que mTERF nativa provocaba la terminación y la activación de la transcripción, mediante ensayos de transcripción *in vitro* donde

se añadía mTERF a una mezcla de lisado mitocondrial de células HeLa que contenía un rDNA artificial con los lugares de terminación e iniciación (H1). Su conclusión fue que una molécula de mTERF tenía la capacidad de unir simultáneamente las dos regiones, terminadora y promotora, del DNA mitocondrial. Esta unión daría lugar a la formación de un bucle de DNA que promovería el recirculado de la maquinaria de la transcripción (Martin *et al.*, 2005). Este modelo explicaría por qué la velocidad de transcripción es mayor para los rRNA en comparación con el resto de mRNAs (Gelfand & Attardi, 1981) (Fig 1.16).

A pesar de la gran cantidad de información obtenida por ensayos *in vitro*, no hay evidencias experimentales que confirmen el papel de terminación de la transcripción de mTERF *in vivo*. Por ejemplo, el estudio de *footprinting in vivo* de las células que presentan la mutación para MELAS A3243G, que *in vitro* causa una disminución de la eficiencia de terminación relacionada con una disminución de la afinidad de mTERF por el lugar de terminación mutado (Hess *et al.*, 1991), no revela ninguna diferencia en la ocupación del lugar de unión de mTERF (Chomyn *et al.*, 2000). Esto explicaría por qué en las células que presentan la mutación de MELAS el ratio de los transcritos rRNA/mRNA no varía (Chomyn *et al.*, 1992). Se ha descrito también la influencia de la hormona tiroidea (Enriquez *et al.*, 1999) o de la variación de suministro de ATP (Micol *et al.*, 1997) sobre la velocidad relativa de la iniciación de la transcripción en H1 y H2, sin que esto tenga ningún efecto en el lugar de unión de mTERF en el patrón de *footprinting*.

En experimentos con ratones *knockout* en Mpv17 –la ausencia de esta proteína reduce el contenido de copias de DNAm_t, el incremento global de la transcripción mitocondrial observado está asociado a la disminución de los niveles de mTERF, sugiriendo que *in vivo* mTERF tendría una función ‘general’ de regulador negativo de la transcripción mitocondrial (Viscomi *et al.*, 2009).

Para estudiar el efecto sobre la transcripción de los niveles de mTERF *in vivo*, se realizaron experimentos en cultivos celulares humanos de sobreexpresión y

2.1.2 Factores terminadores de la transcripción mitocondrial en invertebrados

Se han realizado varios estudios en invertebrados, en particular con dos sistemas experimentales: el erizo de mar y en *Drosophila melanogaster*.

mtDBP, el factor de terminación de la transcripción del erizo de mar.

El DNA mitocondrial del erizo de mar *Paracentrotus lividus*, contiene los mismos genes que el de los vertebrados, pero con una disposición diferente (Cantatore *et al.*, 1989). Los genes ribosomales están separados por una región de 3.3 Kbp que contiene un clúster de 15 genes tRNAs y los genes ND1 y ND2. La región no codificante (NCR) es muy pequeña y está localizada en el clúster de tRNAs después del gen 12S rRNA (Fig 1.17a). En este organismo, la transcripción mitocondrial parece dar lugar a múltiples unidades parciales de transcripción que se superponen (Cantatore *et al.*, 1990), pero el lugar preciso de iniciación aún no ha sido definido.

mtDBP es una proteína de 348 aminoácidos que une con elevada afinidad dos lugares del DNA mitocondrial, uno localizado en la región NCR en el extremo 3' del D-loop y el otro, en el extremo 3' de los genes ND5 y ND6 (Fig 1.17) (Roberti *et al.*, 1991).

La elevada homología entre mtDBP y mTERF sugería que mtDBP pudiese tener una función de terminación mitocondrial (Loguercio Polosa *et al.*, 1999). Esta hipótesis fue demostrada mediante ensayos de transcripción *in vitro* en presencia de las proteínas recombinantes mtDBP y mtRNAPol de erizo (Fernandez-Silva *et al.*, 2001; Polosa *et al.*, 2007). La terminación de la transcripción en NCR es dependiente de mtDBP cuando el lugar de unión de la proteína está en la dirección de la transcripción de la cadena L y es independiente cuando la polimerasa se encuentra el lugar de unión de la proteína en dirección opuesta. Parece ser, que mtDBP actúa como un factor polar de terminación de la transcripción y que la terminación puede

tener lugar mediante dos mecanismos alternativos: dependiente de secuencia y dependiente de proteína.

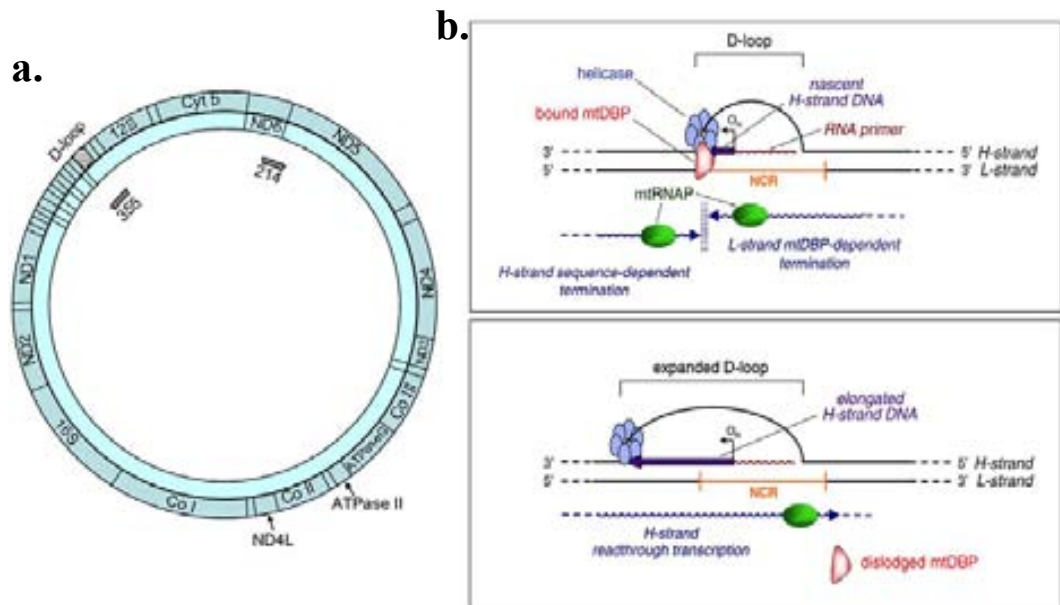


Fig 1.17_ (a) Mapa genético del DNA mitocondrial de erizo de mar. Los recuadros sombreados indican las zonas de interacción de mtDBP (355 y 214) con el DNAm. (b) Diagrama esquemático que representa el papel de mtDBP en la transcripción y replicación en la zona no codificante NCR del DNA mitocondrial del erizo de mar. **(arriba)** Las flechas azules representan los transcritos cuya terminación es dependiente de secuencia o de mtDBP. El esquema muestra también que la síntesis de la cadena pesada está inhibida por la actividad contrahelicasa de mtDBP unida a DNA, formando la estructura de triple hebra del D-loop. La cadena H nascente está compuesta por un cebador de RNA (70-80nt) y un trozo de DNA (20nt aprox.). **(abajo)** La transcripción de la cadena pesada a través del lugar de unión de mtDBP provoca el desplazamiento de la mtDBP, lo que posibilita la reanudación de la replicación de la cadena H del DNA mitocondrial. Figura tomada de (Roberti *et al.*, 2009).

Como el lugar de unión de mtDBP en NCR coincide con el extremo 3' del D-loop -punto clave donde se da la terminación prematura de la síntesis de DNA de la cadena pesada mitocondrial cuando la replicación no es completa- puede plantear la duda de si mtDBP también regula la replicación del DNA mitocondrial controlando la expansión del D-loop. Loguercio Polosa y col. (Polosa *et al.*, 2005) demostraron que mtDBP tenía actividad contrahelicasa bidireccional e independiente de la orientación del lugar de unión de la proteína, sugiriendo, por tanto, un papel de regulador negativo en la replicación. Además, se observó que la transcripción de la cadena H a través del lugar de unión de mtDBP que

coexiste con la terminación dependiente de secuencia en este punto, provoca el desplazamiento de la mtDBP de su lugar de unión, lo que posibilita la reanudación de la replicación del DNA mitocondrial (Fig 1.17b). Así, mtDBP sería el dispositivo que regula la interacción entre transcripción y replicación en el DNA mitocondrial de erizo.

DmTTF, el factor de terminación de la transcripción de *Drosophila*.

La organización genética del DNA mitocondrial de *Drosophila melanogaster* es bastante diferente a la de los genomas humano y de erizo de mar. Los genes están distribuidos equitativamente entre las dos cadenas y forman cuatro bloques localizados alternativamente en las dos cadenas. Parece ser que la transcripción del DNA mitocondrial de *Drosophila* se da mediante múltiples unidades de transcripción que tienen su inicio en el extremo 5' de cada uno de los bloques e, hipotéticamente, el lugar de terminación al final de éstos bloques (Berthier *et al.*, 1986).

DmTTF es la homóloga a mTERF y mtDBP en *Drosophila* (Roberti *et al.*, 2003) y une dos secuencias homólogas no codificantes al final de los bloques de genes que se transcriben en direcciones opuestas (Fig 1.18). Estos lugares de unión coinciden con los dos lugares previamente predichos de terminación de la transcripción (Berthier *et al.*, 1986). Su función como terminador de la transcripción se confirmó mediante ensayos de transcripción *in vitro* (Roberti *et al.*, 2005).

Se ha estudiado la función *in vivo* de DmTTF mediante el estudio del fenotipo de células knock-down (Roberti *et al.*, 2006b). La ausencia de DmTTF no afecta a la formación del extremo 3' del transcrito tRNA^{Ser}(UCN) que se encuentra justo antes del lugar de unión de la proteína a DNA. DmTTF por tanto, no está implicado en la generación del extremo 3' de los transcritos. Estudios mediante PCR a tiempo real (RT-PCR) muestran que el nivel de transcritos localizados después del lugar de unión de DmTTF aumenta en comparación con el control, lo que corrobora la función terminadora de DmTTF. Sorprendentemente,

en cambio, se produce una disminución del nivel de transcritos de los genes que se encuentran entre la región rica en A-T y los lugares de unión de la proteína. Como explicación a esta disminución, se podría suponer que DmTTF estimula la transcripción en la zona promotora de la región rica en A-T -como mTERF-, o que la disponibilidad de RNA polimerasa mitocondrial es menor a causa de la transcripción aberrante que se extiende más allá de los lugares de terminación. Sobre su posible papel de regulación de la replicación, la ausencia de DmTTF no altera el número de copias del DNA mitocondrial. Aun así, a causa del tiempo de evaluación con la técnica de *knock-down*, no se puede descartar que DmTTF esté involucrada en la replicación. La electroforesis SDS de la fracción proteica muestra que la falta de DmTTF no causa alteraciones cualitativas en el perfil de los polipéptidos mitocondriales aunque sí en la cantidad de éstos.

El estudio de los efectos de la sobreexpresión de DmTTF en el nivel de transcritos mitocondriales de ND5 y Cyt b muestra que éstos disminuyen como cabe esperar por la función terminadora de DmTTF. Además, disminuye también la síntesis *de novo* de casi todos los polipéptidos.

Se han propuesto dos modelos para la transcripción de DNA mitocondrial de *Drosophila* (Roberti *et al.*, 2006b). Un modelo asume que la transcripción consta de un promotor para cada cadena localizado en la región rica en A-T (P1 y P3), de manera que DmTTF actuaría mayoritariamente de atenuador. En el otro modelo la transcripción de cada cadena requiere dos promotores localizados en el extremo 5' de cada bloque de genes (P1+P2 y P3+P4). En este caso, DmTTF unido a tRNA^{Ser}(UCN)/ND1 actuaría como terminador en ambas direcciones, mientras que unido a tRNA^{Glu}/tRNA^{Phe} podría actuar como terminador en una dirección y como atenuador en la otra.

Los resultados de los estudios de sobreexpresión de DmTTF que provocan una disminución los transcritos ND5 y cyt b respaldan el primer modelo (Roberti *et al.*, 2009).

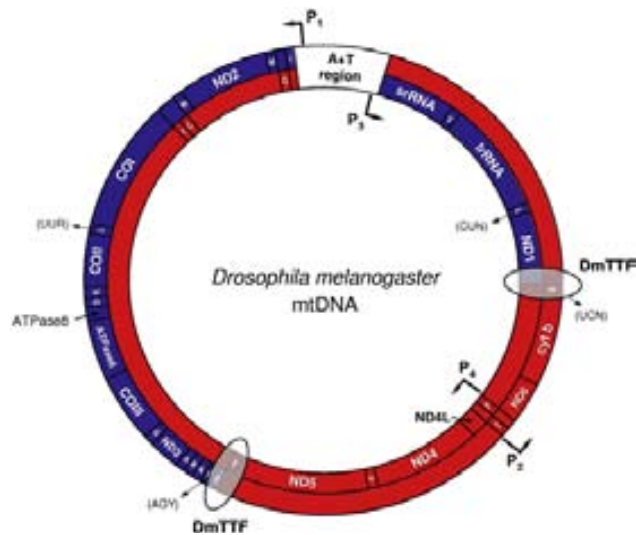


Fig 1.18_ Mapa genético del DNA mitocondrial de *Drosophila melanogaster*. Las dos elipses indican los lugares de unión de DmTTF. Los posibles promotores de la transcripción están indicados como P₁, P₂, P₃ y P₄, cuya dirección viene señalizada por la flecha. En azul se muestran las regiones cuyo nivel de transcritos decrece en los knock-down de DmTTF. En rojo, las regiones cuyo nivel de transcritos incrementa en los knock-down de DmTTF. Figura tomada de (Roberti *et al.*, 2009).

3_ FAMILIA mTERF

La familia de mTERF, identificada en metazoos y plantas, consiste en cuatro subfamilias llamadas MTERF1-4 (Fig 1.19). Las proteínas que pertenecen a esta familia se localizan en mitocondria y muestran una arquitectura modular basada en la repetición de un motivo de 30 aminoácidos, el motivo mTERF (Linder *et al.*, 2005; Roberti *et al.*, 2009).

La familia MTERF1 incluye el factor de terminación humano mTERF, el de erizo de mar mtDBP y el de *Drosophila* DmTTF. Los estudios *in vivo* e *in vitro* muestran que estos factores tienen funciones en la iniciación de la transcripción y en el control de la replicación del DNA, además de la terminación de la transcripción. La multiplicidad de funciones entre los miembros de una misma familia puede estar relacionada con las diferencias en la organización de los genomas mitocondriales.

Los primeros estudios en humanos muestran que mTERF1 y mTERF2 (mTERFL o MTERFD3) presentes sólo en vertebrados, tienen patrones de expresión opuestos en respuesta a sérum. Mientras la expresión de mTERF se induce por la adición de sérum a células privadas de éste, la expresión de MTERF2 se inhibe drásticamente. La sobreexpresión de MTERF2 suprime el crecimiento celular manteniendo las células en el estadio G1. Ensayos de *northern blot* revelan tres transcritos de MTERF2 de 1.7, 3.2 y 3.5Kb. Mientras el transcrito de 3.2kb es exclusivo de músculo esquelético, los otros dos se expresan predominantemente en corazón, hígado, páncreas y músculo esquelético (Chen *et al.*, 2005b). Experimentos de cuantificación demuestran que MTERF2 es relativamente abundante, con aproximadamente un monómero por cada 265pb de DNA mitocondrial y el *crosslinking* con formaldehído parece demostrar que MTERF2 está presente en los nucleoides (Pellegrini *et al.*, 2009). Existen ciertas controversias sobre la unión a DNA, Pellegrini y col. observan que mTERF2 tiene una actividad de unión no específica de secuencia, mientras que Wenz y col. (Wenz *et al.*, 2009) muestran que se une a la región promotora de la transcripción, sugiriendo que interviene en el inicio de ésta. Ratones *knock-out* de mTERF2 resultan en miopatía y déficit en la memoria asociado a un descenso de los transcritos mitocondriales y un desequilibrio en los tRNA. Estas

aberraciones están asociadas con un descenso en los niveles del estado estacionario de las proteínas de la cadena de fosforilación oxidativa causando una disminución de la función respiratoria. Estudios de interacción in vitro sugieren que el DNA mitocondrial media la interacción entre MTERF2 y MTERF3.

MTERF3 (MTERFD1) es el grupo más conservado de la familia. El gen de MTERF3 es esencial porque los embriones de ratón knock-out mueren en la gestación. Los mismos estudios indican que la inactivación específica de MTERF3 en tejido cardíaco causa aberraciones en la transcripción del DNA mitocondrial y severas deficiencias en la cadena respiratoria. MTERF3 se une al DNA mitocondrial en la región promotora y la reducción de MTERF3 da lugar a un incremento en la iniciación de la transcripción en ambas cadenas del DNA (HSP y LSP) (Park *et al.*, 2007). Estudios de sobreexpresión de D-MTERF3 en *Drosophila* mediante un análisis de RT PCR muestran que el número de transcritos decrece en un 0.2-0.6%, confirmando el papel de regulador negativo de la transcripción mitocondrial también en *Drosophila* (Roberti *et al.*, 2009). La estructura cristalográfica de MTERF3 humana ha sido resuelta recientemente y tiene forma de medio donut formado por repeticiones en tándem del motivo triangular formado por tres hélices alfa. En el centro de la cavidad cóncava del donut se aprecia una zona cargada positivamente que podría ser la zona de interacción con el DNA (Spahr *et al.*).

En resumen, MTERF1, 2 y 3 comparten un mismo lugar de unión a DNA en la región promotora. MTERF1 y 2 promueven la iniciación de la transcripción mientras MTERF3 la inhibe. Los tres factores son necesarios para mantener niveles de transcripción óptimos y asegurar el buen funcionamiento de la cadena de transporte electrónico y la fosforilación oxidativa. La función de MTERF4 no se ha investigado en detalle, pero parece ser que une el DNA mitocondrial en la región del D-loop y que forma un homodímero con una RNAmetiltransferasa putativa (Asin-Cayuela, 2008). Estudios recientes en referencia a la implicación de estos factores en la replicación muestran que en experimentos en líneas celulares humanas donde se estudia el efecto de la sobreexpresión y knockdown de MTERF2 y 3 sobre el número de copias del DNAmitocondrial, la sobreexpresión, sobretodo de MTERF3, causa una ligera disminución en el número de copias de DNA mitocondrial y una acumulación de

intermediarios específicos de la replicación del DNA mitocondrial. Este hecho indica que existe un impedimento en los pasos finales de la replicación. Estos experimentos implican a la familia mTERF en el control de la progresión de la horquilla de replicación y parecen apuntar a la idea de la implicación de estas proteínas en la supervisión del paso ordenado de las maquinarias de replicación y transcripción (Hyvarinen *et al.*).

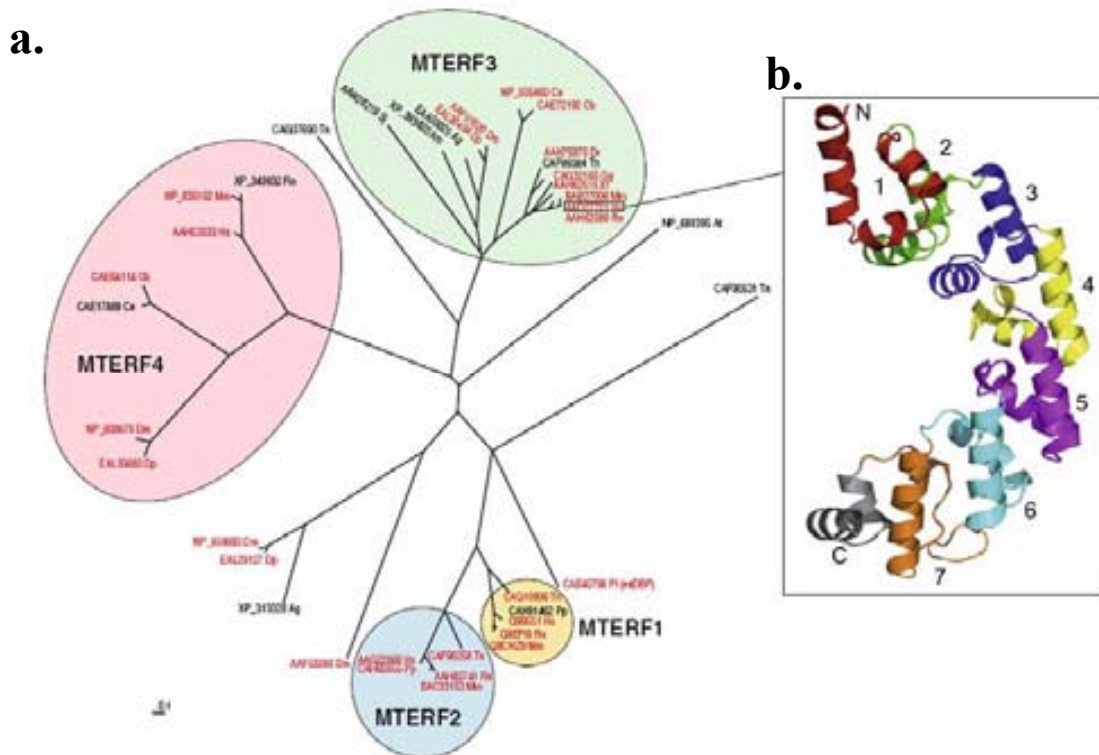


Fig 1.19_ (a) Árbol filogenético de las proteínas MTERF de metazoos. Las proteínas están identificadas con los números de acceso de GenPept. Las abreviaciones corresponden a Ag (*Anopheles gambiae*), Am (*Apis mellifera*), At (*Arabidopsis thaliana*), Cb (*Caenorhabditis briggsae*), Ce (*Caenorhabditis elegans*), Dm (*Drosophila melanogaster*), Dp (*Drosophila pseudoobscura*), Dr (*Danio rerio*), Gg (*Gallus gallus*), Hs (*Homo sapiens*), Mm (*Mus musculus*), Pl (*Paracentrotus lividus*), Pp (*Pongo pygmaeus*), Rn (*Rattus norvegicus*), Sj (*Schistosoma japonicum*), Tn (*Tetraodon nigroviridis*) y Xt (*Xenopus tropicalis*). Las proteínas en rojo son aquellas que se localizan en mitocondria según la predicción con el programa TargetP. (b) Estructura de MTERF3, los motivos triangulares formados por tres hélices alfa están numerados y coloreados.

4_ PROTEÍNAS HELICAL TANDEM REPEAT

Las proteínas *helical tandem repeat* son proteínas eucarióticas involucradas en varios procesos celulares, incluyendo la señalización, el transporte y la regulación citoesquelética. Como su nombre indica, se caracterizan por presentar un dominio estructural compuesto por un tandem de repeticiones de hélices α . Hasta hoy se pueden distinguir las proteínas ARM-repeat, PUF-repeat y HEAT-repeat.

Dentro de este amplio grupo, las proteínas *armadillo repeat* presentan estructuralmente una repetición en tándem de motivos de aproximadamente 42 aminoácidos (Peifer *et al.*, 1994; Huber *et al.*, 1997). El nombre de armadillo se debe a la proteína Armadillo de *Drosophila*, que fue la primera proteína Arm resuelta y que es homóloga de la β -catenina en mamíferos e interviene en la adhesión celular y en la regulación de la expresión genética durante el desarrollo. Desde entonces se han resuelto varias estructuras de proteínas Arm-repeat, demostrando que, aunque no necesariamente presentan una elevada identidad de secuencia, comparten una estructura similar.

Cada una de las repeticiones o motivos consiste en tres hélices α . Estas repeticiones se pliegan en tándem interaccionando entre ellas para formar una superhélice dextrógira de hélices creando una superficie para la interacción proteína-proteína (Conti *et al.*, 1998; Conti & Kuriyan, 2000).

Aunque han sido originalmente caracterizadas en animales, las proteínas que contienen *Arm repeats* existen también fuera de este reino. La estructura del dominio armadillo permite a dichas proteínas tener varias funciones en la célula, como la regulación del transporte de proteínas a núcleo (Importin- α), funciones de unión con el citoesqueleto y la regulación de la expresión genética en respuesta a señales extracelulares (β -catenin y homólogos), entre otras. Las proteínas armadillo de función conocida se puede dividir en varias categorías con diferentes subfamilias que poseen secuencias características fuera del dominio armadillo que contribuyen a su función (Tewari *et al.*; Coates, 2003) (Fig 1.20). Esta familia parece tener mecanismos de regulación similares, como por ejemplo la

acumulación nuclear en respuesta a señales extracelulares y la regulación por fosforilación.

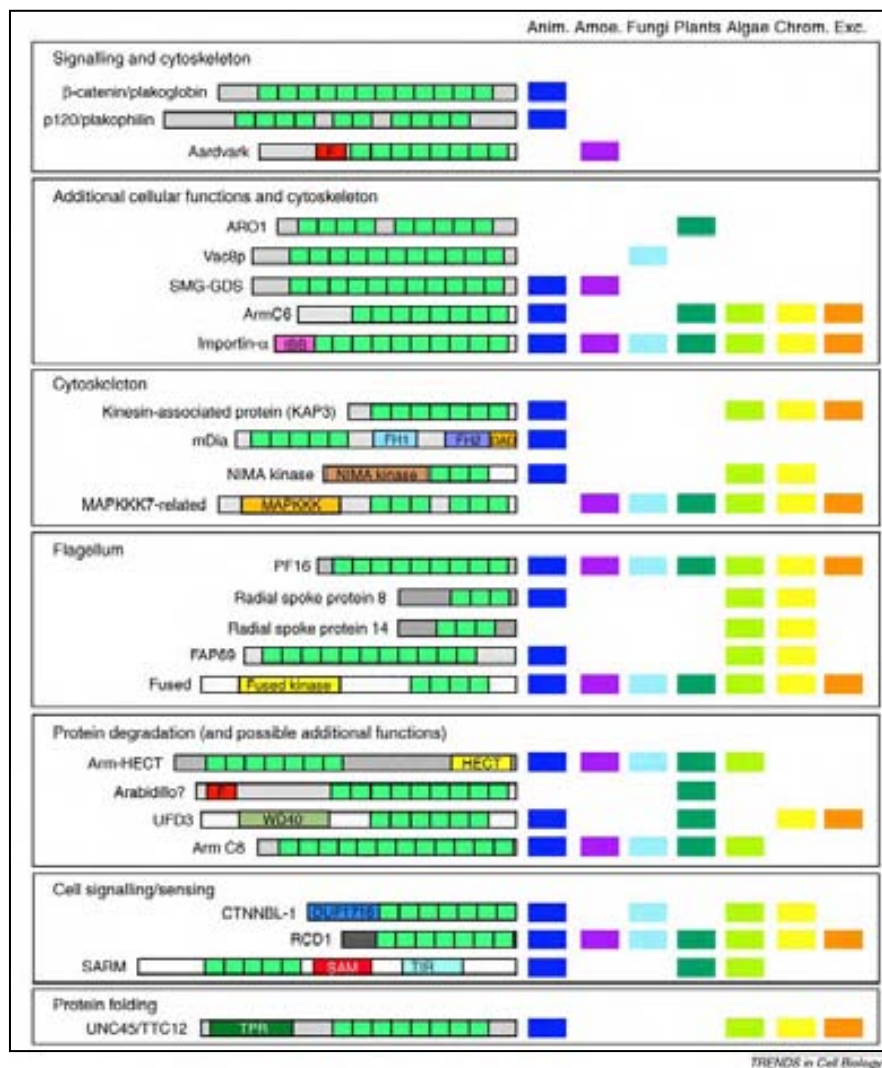


Fig 1.20_ Conservación y función de las proteínas Arm-repeat. La presencia de cada tipo de proteína en las diferentes ramas del árbol de los eucariotas, se indica mediante los rectángulos de colores: animales (Anim., azul oscuro), amebas (Amoe., morado), hongos (Fungi, cian), plantas (Plants, verde oscuro), algas verdes/rojas (Algae, verde claro), chromoalveolates (Chrom., amarillo), excavates (Exc., naranja). Las proteínas se agrupan según su función. Las cajas verdes indican los Arm repeats mientras las otras representan secuencias específicas de cada subfamilia de proteínas. A la derecha de la tabla hay una breve descripción de la función de cada proteína. Figura tomada de (Tewari *et al.*)

En el año 2002 se determinó la estructura cristalográfica del dominio de unión a RNA de la proteína Pumilio1, una proteína *helical tandem repeat*, en complejo con el RNA (Wang *et al.*, 2002). Pum presenta una superficie cóncava pero sin el giro a lo largo del eje de rotación observado en las proteínas ARM-repeat. Cada motivo Puf está formado por tres hélices α , dos largas (H1 y H3) y una corta (H2), y en general los motivos Puf son bastante

regulares (rmsd 0.85-1.6 Å) y se empaquetan uniformemente (Fig 1.22). Estas proteínas (Puf) son reguladoras del desarrollo y controlan la estabilidad y la traducción del mRNA mediante secuencias de unión en la región 3' no codificable. El RNA se une a la superficie cóncava de la molécula, donde cada uno de los ocho motivos contacta con una base diferente del RNA mediante tres cadenas laterales amioácidas de posiciones conservadas.

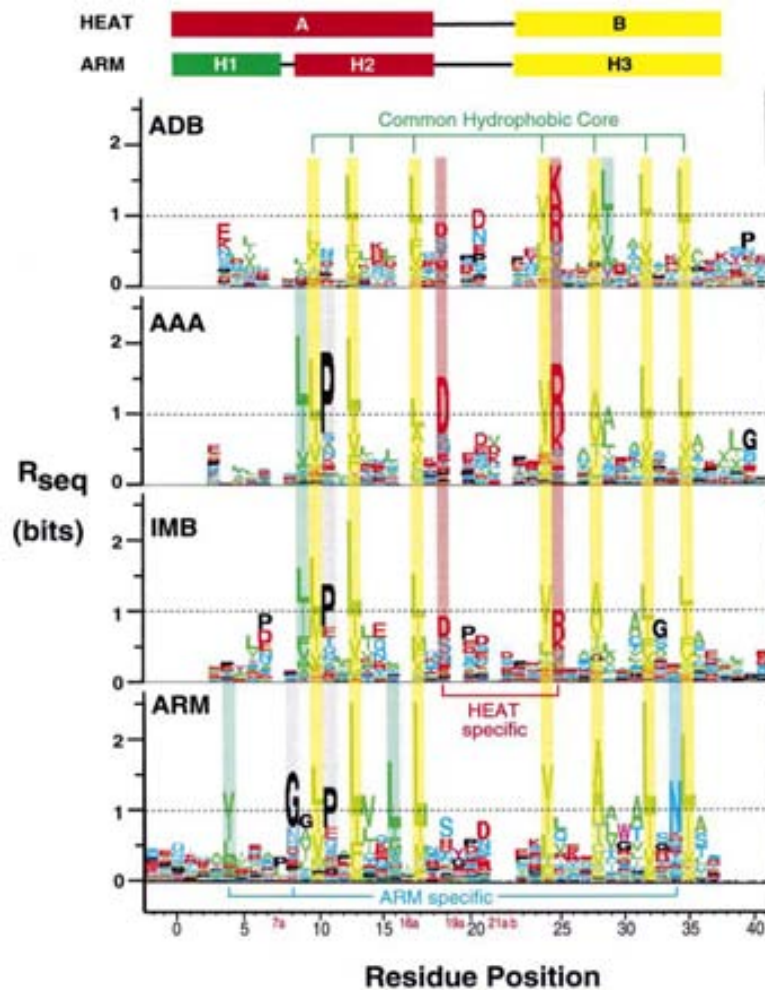


Fig 1.21_ Logos de secuencia generados a partir de alineamientos de ARM y HEAT repeats. Los logos han sido generados usando el servidor WebLogo (<http://www.bio.cam.ac.uk/seqlogo/>). Las posiciones con información significativa han sido resaltadas. Código de color: en verde los aminoácidos hidrofóbicos; en rojo los aminoácidos cargados; en negro los que provocan el giro; en magenta los aromáticos y en azul los polares. Figura tomada de (Andrade *et al.*, 2001).

Las proteínas armadillo repeat están estructuralmente relacionadas con las proteínas HEAT repeat y probablemente tienen un origen filogenético común (Andrade *et al.*, 2001). Una búsqueda exhaustiva de secuencias en las bases de datos revela que como mínimo una de cada 500 proteínas eucariotas contiene dichas repeticiones. Todas las

proteínas identificadas en las bases de datos pueden clasificarse en cuatro grupos según la similitud de secuencia: las ARM-repeat canónicas y tres grupos más divergentes de proteínas HEAT-repeat. Todas ellas comparten un core común de siete residuos hidrofóbicos altamente conservados, que se encuentran enterrados dentro del dominio o entre dominios en las estructuras cristalográficas de proteínas resueltas (Fig 1.21). Aun así, los motivos difieren en la posición de varios residuos específicos, sugiriendo importantes diferencias estructurales o funcionales entre las clases.

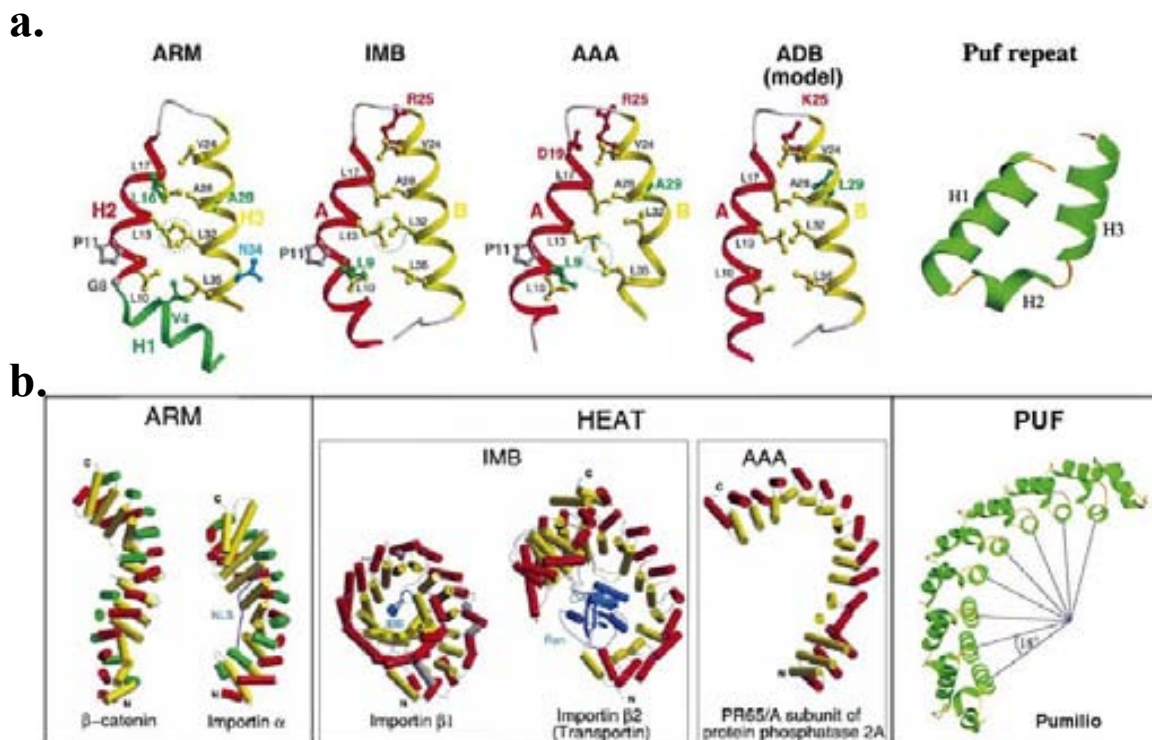


Fig 1.22_ (a) Estructuras representativas de ARM, HEAT y PUF repeats. Los residuos que se muestran en los motivos ARM y HEAT son los que están resaltados en la figura anterior, nombrados y coloreados según el mismo esquema. (b) Proteínas que contienen HEAT o ARM repeats. Las hélices H1, H2 y H3 de las proteínas ARM están señalizadas en verde, rojo y amarillo respectivamente, las hélices A y B de las proteínas HEAT están señalizadas en rojo y amarillo y los ligandos en azul. Las estructuras mostradas son: β -catenin de ratón (PDB 2bct; (Huber *et al.*, 1997); *Importin- α* de levadura (PDB 1ee5; (Conti & Kuriyan, 2000); *Transportin* (PDB 1qbk; (Chook & Blobel, 1999); *human Importin- β* (PDB 1qgk; (Cingolani *et al.*, 1999), la subunidad PR65/A de la fosfatasa humana 2A (PDB 1b3u; (Groves *et al.*, 1999) y el dominio Puf de la proteína Pumilio. Figura adaptada de (Andrade *et al.*, 2001) y (Edwards *et al.*, 2001).

Objetivos

El objetivo principal de esta tesis consistió en la determinación de la estructura tridimensional, mediante cristalografía de proteínas, del factor de terminación de la transcripción humano mTERF, en complejo con el DNA que contiene la secuencia específica de terminación.

En el transcurso de esta tesis, para conseguir el objetivo principal, se plantearon los objetivos parciales que se describen a continuación:

- Producir cantidades elevadas de proteína mTERF soluble mediante la sobreexpresión heteróloga en cepas de *Escherichia coli*.
- Purificar la proteína recombinante para obtenerla en un grado de estabilidad y pureza tales que cumpliera con los requisitos necesarios para llevar a cabo experimentos de cristalización.
- Caracterizar y formar el complejo de mTERF con el DNA que contiene la secuencia específica de terminación.
- Obtener cristales de complejo mTERF-DNA de calidad suficiente para generar datos de difracción de rayos X de resolución y calidad adecuada para la resolución de la estructura.
- Procesar los datos de difracción y utilizarlos para resolver la estructura tridimensional del complejo.
- Estudiar mediante SAXS el complejo en solución de mTERF con el DNA de 28pb que contiene la secuencia específica de terminación resultante de los ensayos de *footprinting*.
- Caracterizar la unión a DNA y la especificidad de mTERF utilizando formas truncadas de la proteína, mediante la estructura cristalográfica, experimentos de SAXS y ensayos con geles nativos de retardo.



RESULTADOS
Y
DISCUSIÓN

1_ Obtención de la proteína. Purificación y caracterización

1.1 Análisis bioinformático y clonaje de la proteína

Un estudio bioinformático previo de predicción de la estructura a partir de la secuencia de mTERF llevado a cabo por Fernández-Silva y col. predijo, para la cadena polipeptídica, tres potenciales cremalleras de leucina y dos dominios básicos necesarios para la unión a DNA (Fernandez-Silva *et al.*, 1997).

La predicción de estructura secundaria de la proteína madura con el programa PsiPred (Jones, 1999) indica que es una proteína formada en su mayor parte por hélices α (Fig 2.1). Se observa además, a lo largo de la secuencia, un patrón de repetición de tres hélices (de aproximadamente 30 residuos), una un poco mayor (H1), una mediana (H2) y otra pequeña (H3) precedida mayoritariamente por una prolina.

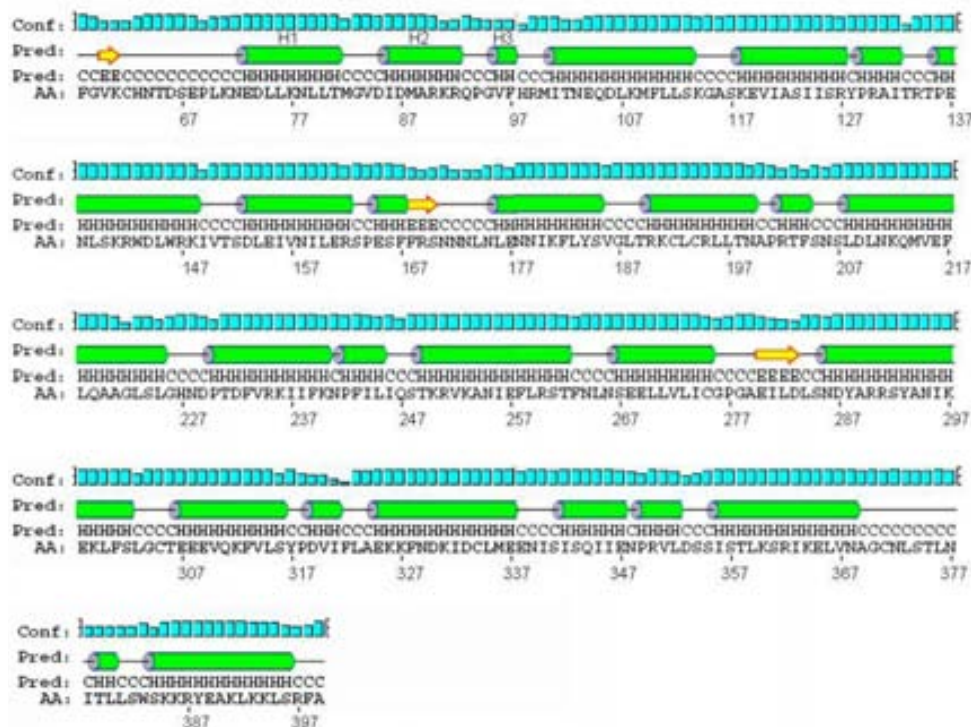


Fig 2.1_ Predicción de la estructura secundaria de la proteína madura mTERF (residuos del 58 al 399, (Fernandez-Silva *et al.*, 1997) mediante el programa PsiPred. Los cilindros verdes indican la predicción de hélices α (H), las flechas amarillas, hebra β (E) y las líneas indican zonas de *random coil* (C). Las barras azules indican la probabilidad de estas predicciones (Jones, 1999). Se indican H1, H2 y H3 para el primer motivo que se repite a lo largo de la secuencia.

Esta arquitectura modular para mTERF fue sugerida por Roberti y col. teniendo en cuenta el resultado del análisis de la secuencia de las proteínas de la familia MTERF, en el que se identificó la repetición de un motivo de unos 30 aa llamado “mTERF-motif” (Roberti *et al.*, 2006a). Con el programa SMART (*Simple Modular Architecture Research Tool*) (Schultz *et al.*, 1998; Letunic *et al.*, 2009) confirmamos que establece para mTERF seis ‘mTERF-motif’ que corresponden a fragmentos entre los aminoácidos: 122-152, 157-189, 194-224, 235-266, 311-341 y 342-372.

Para abordar el clonaje de la proteína es recomendable no cortar ningún elemento de estructura secundaria regular en el diseño de los constructos, ya que podría interferir en la solubilización, el estado de agregación, la purificación y la estabilidad de la muestra, especialmente en las condiciones de alta concentración y pureza necesarias para los estudios cristalográficos. Dado que de antemano no se sabe cómo puede interferir el plegamiento de la proteína en la accesibilidad de la cola de histidinas, se empezaron clonando dos constructos de la proteína madura, uno con una cola de 6 histidinas en el N-terminal (añadiendo un lugar de corte para TEV) y otro con la cola en el C-terminal. Posteriormente se clonaron también versiones truncadas de mTERF en función de la proteólisis espontánea, mTERF- Δ N (residuos 99-399), y del estudio de proteólisis limitada de la proteína, mTERF- Δ C (residuos 56-360) (Fig 3.1).

1.2 Pruebas de expresión y solubilidad

Para establecer el protocolo de expresión, se realizaron varias pruebas en medio LB variando la cepa bacteriana, la densidad óptica del cultivo (OD_{600nm}) en el momento de la inducción, la concentración del agente inductor de la expresión (IPTG), la temperatura y el tiempo de expresión tras añadir el inductor (ver materiales y métodos). En referencia a la concentración de agente inductor y a las cepas bacterianas, las diferencias en la expresión no fueron significativas, así pues, en los siguientes estudios se utilizó una concentración de inductor de 1mM de IPTG y la cepa BL21(DE3) porque es la que presentaba mayor velocidad de crecimiento.

A continuación se muestran los resultados de la electroforesis en geles de poliacrilamida-SDS de algunas pruebas de expresión realizadas con la cepa bacteriana BL21(DE3)

transformada con el plásmido que contenía el constructo mTERF-CHis, a diferentes densidades ópticas (OD) de inducción y a diferentes temperaturas y tiempos de expresión (Fig 2.2). Como se puede observar, la inducción a ODs mayores (alrededor de 1) daba lugar a un mayor rendimiento. Los ensayos de expresión a diferentes temperaturas mostraron que la proporción de proteína soluble aumentaba a medida que disminuía la temperatura y aumentaba el tiempo de expresión (Fig 2.2b). A partir de estos resultados, la expresión a gran escala de mTERF se realizó a 17°C durante 60 horas aproximadamente.

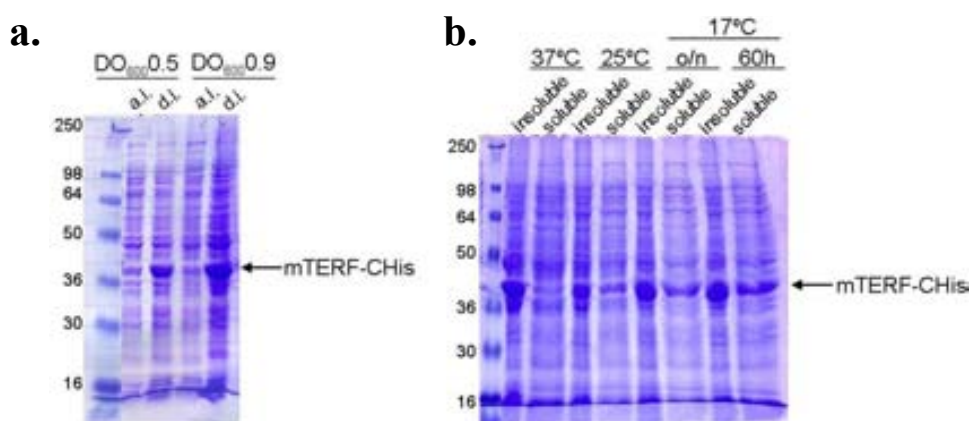


Fig 2.2_ (a) Resultados de la electroforesis en geles de poliacrilamida-SDS de las fracciones antes de inducir (*a.i.*) y después de inducir (*d.i.*) de cultivos en medio LB de células *E. coli* BL21(DE3) transformadas con el plásmido que contiene mTERF-CHis e inducidos a una OD_{600nm} de 0.5 y de 0.9 respectivamente. (b) Resultados de la electroforesis en geles de poliacrilamida-SDS de las fracciones insoluble y soluble de cultivos en medio LB de cepas BL21(DE3) transformadas con el plásmido pET28a que contiene mTERF-CHis y expresados a 37°C durante 4h, a 25°C o/n y a 17°C o/n y durante 60h. Para todos los casos el tampón de lisis era 100 mM HEPES pH7.4, 0,5 M NaCl, 20 mM BME, 1mM PMSF.

Los parámetros físico-químicos de la proteína son cruciales para decidir la composición del tampón de solubilización. Éste se formuló teniendo en cuenta que mTERF es una proteína de unión al DNA y en función del número de cisteínas y del pI teórico (Tabla 2.3) resultante del análisis de la proteína con el programa ProtParam (Gasteiger E., 2005). En la siguiente tabla se muestran algunas de las características físico-químicas de los constructos clonados durante la tesis.

mTERF-NHis	mTERF-CHis	mTERF-ΔN	mTERF-ΔC
Núm aminoácidos 360	Núm aminoácidos 353	Núm aminoácidos 310	Núm aminoácidos 314
PM 41469.9	PM 40618.0	PM 35684.3	PM 36154.7
pI teórico 9.25	pI teórico 9.33	pI teórico 9.36	pI teórico 8.80
Cys 7	Cys 7	Cys 6	Cys 6

Tabla 2.3_ Resultado del análisis de los parámetros físico-químicos de los diferentes constructos mediante el programa ProtParam (Gasteiger E., 2005). Nótese la disminución del pI teórico para el constructo mTERF-ΔC comparado con el resto de constructos.

Así, las primeras pruebas de solubilidad con los constructos mTERF-CHis y mTERF-NHis, se realizaron con un tampón a pH 7.4, al que se añadió un agente reductor para evitar la posible formación de puentes disulfuro. Al ser una proteína de unión a DNA, las pruebas de solubilización se realizaron con una batería de tampones a concentraciones relativamente elevadas de sal, entre 0.25 y 1 M.

En base a estas pruebas y al tampón de lisis mitocondrial descrito para la purificación de mTERF a partir de células HeLa (Daga *et al.*, 1993), la composición del tampón de lisis en las primeras purificaciones de mTERF a gran escala era de 50mM HEPES pH 7.4, 500 mM NaCl y 5 mM BME.

Posteriormente, una vez establecido el protocolo de purificación, se realizaron pruebas de expresión utilizando medio de cultivo SB y empleando el método de autoinducción (ver Materiales y Métodos). En ambos casos, el rendimiento era mayor que utilizando el medio de cultivo LB con el que obteníamos un rendimiento de <2 mg proteína/L de cultivo.

La utilización de medio SB, junto con la optimización del tampón proteico mediante thermofluor (ver más adelante) y la purificación a 4°C ayudaron a mejorar el rendimiento de la producción final (~3 mg proteína/L de cultivo).

1.3 Purificación de mTERF

Para todos los constructos la purificación se llevó a cabo en dos etapas: una columna de afinidad de níquel –todos los constructos presentan una cola de 6 histidinas– y, después de concentrar las fracciones del pico de elución que contenían la proteína más pura, una cromatografía de exclusión molecular.

En la columna de afinidad para la purificación de mTERF- Δ N (Fig 2.4a), mTERF-NHis y mTERF-CHis, una vez cargada la muestra y aplicando un escalón a 12% de tampón B (60mM imidazol) lográbamos eliminar la mayoría de contaminantes, de manera que el pico de muestra, eluido aproximadamente a 30% de tampón B (150 mM imidazol), era bastante puro. En el caso de la purificación de mTERF- Δ C, la elución se realizaba mediante un gradiente lineal.

En la cromatografía de exclusión molecular de la Fig 2.4b, se compara el volumen de elución (V_e) de la proteína con el V_e de las proteínas estándar con las que se calibró la

columna en el mismo tampón de purificación. La representación lineal de K_{av} ($(V_e - V_o) / (V_t - V_o)$) vs el log del peso molecular (pM) de los patrones, nos permitió concluir que la proteína tenía un pM aparente de 44 ± 10 KDa y, por tanto, se comportaba como monómero en solución.

En los casos en que la muestra se inyectó muy concentrada ($\sim 20 \text{ mg ml}^{-1}$) en la columna de exclusión molecular, se pudo distinguir sistemáticamente en el cromatograma un pequeño pico estable de peso molecular aparente de ~ 450 KDa.

Para todos los constructos, mTERF-NHis, mTERF-CHis, mTERF- Δ N y mTERF- Δ C, el perfil de elución de la cromatografía de exclusión molecular mostraba un pico de proteína soluble y estable de pM aparente correspondiente a un monómero.

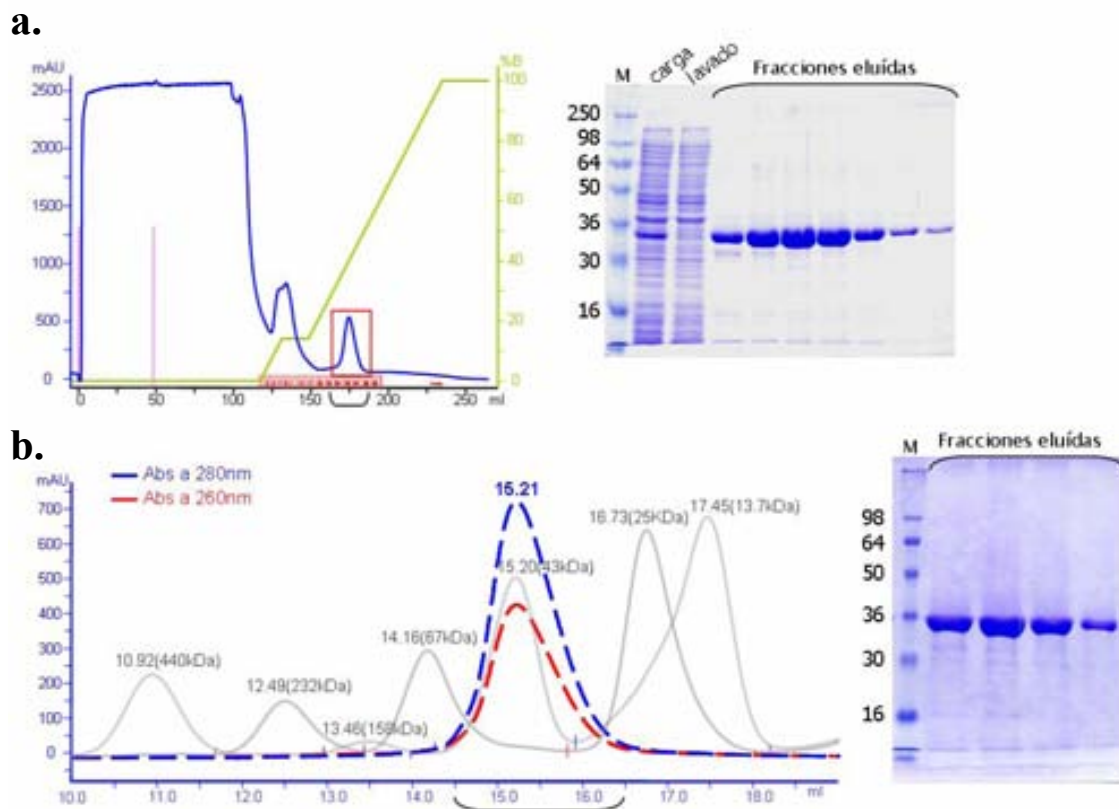


Fig 2.4_ Purificación de mTERF- Δ N (a) Cromatograma (izq) y resultados de la electroforesis en gels de poliacrilamida-SDS (dcha) correspondientes a la columna de afinidad His-Trap. En azul se indica la absorbancia (Abs) a 280nm y en verde el gradiente de % del tampón B (0.5M imidazol). En el margen izquierdo de los gels se muestran los pesos moleculares en kDa de los marcadores (M). Las fracciones cargadas en el gel SDS corresponden al pico de elución de mTERF que se muestra en el cromatograma con un rectángulo rojo. (b) Cromatograma (izq) y resultados de la electroforesis en gels de poliacrilamida-SDS (dcha) correspondientes a la columna de exclusión molecular (Superdex 200). El perfil en gris corresponde a la Abs a 280nm de los patrones de la calibración, en el máximo de los picos se muestra el volumen de elución (V_e) y en paréntesis, el peso molecular. En azul se muestra la abs a 280nm y en rojo la abs a 260nm de la muestra. Las fracciones del pico señaladas con una llave son las que se muestran en el gel SDS (dcha). La banda correspondiente a mTERF- Δ N (35,7 KDa) se encuentra justo debajo del marcador de 36KDa.

1.4 Optimización del tampón proteico por *thermofluor*

La homogeneidad, estabilidad y solubilidad de la proteína son importantes tanto en cristalografía, como en los ensayos de dispersión a bajo ángulo (*small-angle scattering*, SAS). En cristalografía, porque estos parámetros están directamente relacionados con la probabilidad de obtener cristales y, en SAS, porque nos interesa una muestra homogénea y lo más monodispersa posible.

Después de varias purificaciones, e incluso tras obtener los primeros cristales, se realizaron experimentos de estabilidad térmica de la proteína en función de la composición del tampón (técnica de *thermofluor*) para la optimización del mismo. Los ensayos se realizaron midiendo la variación de fluorescencia *vs* la temperatura de la muestra resultante de añadir 5 μ l de proteína a una serie de condiciones diferentes en presencia de un colorante sensible a la hidrofobicidad del entorno.

En el primer ensayo se analizó la influencia del tipo de tampón (fosfato sódico, HEPES, citrato sódico, MES y acetato sódico) y del pH, además de algunas sales (KCl y NaCl) y otros aditivos. La composición y concentración de las condiciones que se probaron en cada pocillo de la placa de *thermofluor* se especifican en la Fig 2.5a. Lo más destacable de este ensayo es que, en las nuevas condiciones con los diferentes tampones no se añadió sal, de manera que la mayoría dieron resultados desfavorables en comparación al control (50 mM HEPES pH7.4, 500 mM NaCl, 10 mM BME) excepto para el citrato, que además era el tampón en el que se habían obtenido los primeros cristales de complejo.

a.

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
100mM Fosfato Na pH5.5	100mM Fosfato Na pH6.0	100mM Fosfato Na pH6.5	100mM Fosfato Na pH7.0	100mM Fosfato Na pH7.5	100mM HEPES pH8.5	100mM HEPES pH7.0	100mM HEPES pH7.4	100mM Citrato sódico pH5.0	100mM Citrato sódico pH5.5	100mM Citrato sódico pH6.0	100mM Citrato sódico pH6.5
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
100mM MES pH5.5	100mM MES pH6.0	100mM MES pH6.5	100mM Acetato sódico pH 5.0	100mM Acetato sódico pH 5.5	10% Glicerol	250mM KCl	100mMKCl	250mM NaCl	100mM NaCl	5mM Glucosa	10mM Glicina
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
10mM Spermine tetrahidrocloride	10mM Trimethylamine	15mM MgCl ₂	15mM CaCl ₂	15mM CuCl ₂	15mM MnCl ₂	15mM NiCl ₂	15mM CoCl ₂	15mM FeCl ₂	15mM BaCl ₂	15mMSrCl ₂	Tampón CONTROL

Fig 2.5_ (a) Ensayo 1. Tabla de pruebas. Se especifica en cada recuadro la composición y concentración de cada nueva condición. En el caso de los reactivos indicados en C1-C12, la condición se preparó añadiendo los aditivos al tampón proteico.

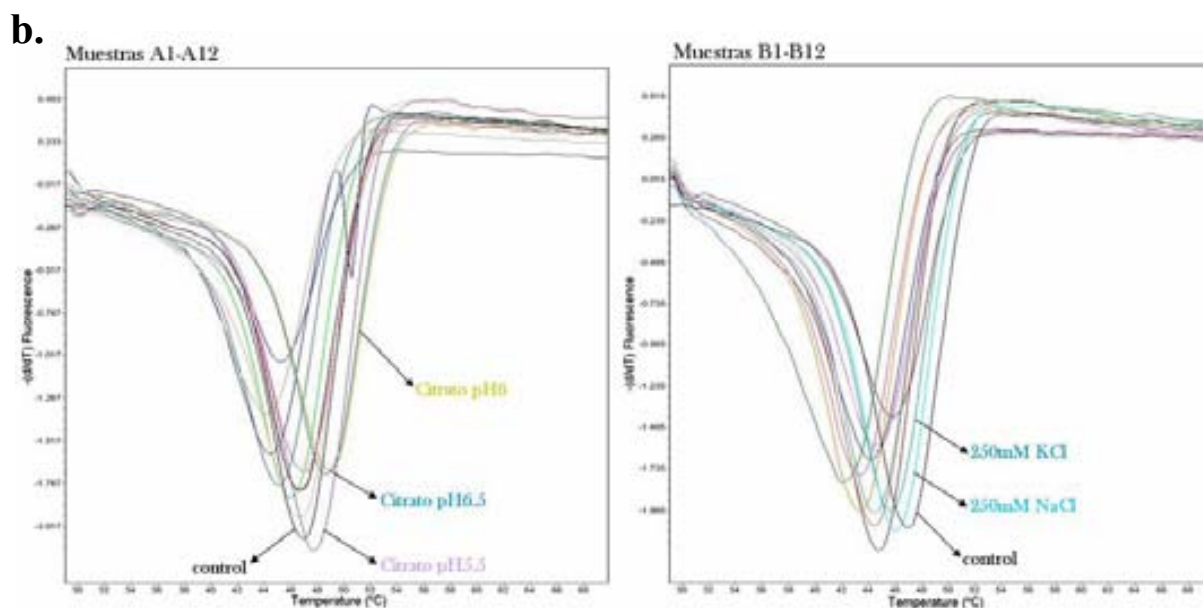


Fig 2.5_ (b) Ensayo 1. Esquema de la primera derivada de las curvas de fusión para A1-A12 y B1-B12. En el eje y se muestra la primera derivada de la variación de la fluorescencia ($-d/dT$) y en el eje x la temperatura (T). El mínimo de la curva indica la temperatura de fusión (T_m). Cuanto mayor sea T_m , más estabilizada se encuentra la proteína en solución. En algunos casos se indica, con el mismo color que la curva resultante, la composición del tampón.

El siguiente ensayo (Fig 2.6) consistió en probar la estabilidad de la proteína con tampón citrato de sodio y diferentes concentraciones de sal (A1-B6), en presencia de glicerol (B7), del DNA_{TER} de 29pb (B8) y de 15pb (B9-B11) (Fig 2.14). Los resultados evidenciaron que mTERF es más estable en presencia de concentraciones elevadas de NaCl, de glicerol (B7) o del oligómero de 29pb (B8). El resultado de este ensayo muestra que el oligómero de 15 no es suficiente para estabilizar de manera notable la proteína. Este hecho se podría relacionar con la precipitación del complejo que se produce con moléculas de DNA menores de 23 pb al dializar la muestra a concentraciones de sal inferiores (~200 mM) para su posterior cristalización.

a.

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
50mM Citrato Na pH5.5 NaCl 500mM	50mM Citrato Na pH5.5 NaCl 750mM	50mM Citrato Na pH5.5 NaCl 1M	50mM Citrato Na pH5.5 KCl 500mM	50mM Citrato Na pH5.5 KCl 750mM	50mM Citrato Na pH5.5 KCl 1M	50mM Citrato Na pH6.0 NaCl 500mM	50mM Citrato Na pH6.0 NaCl 750mM	50mM Citrato Na pH6.0 NaCl 1M	50mM Citrato Na pH6.0 KCl 500mM	50mM Citrato Na pH6.0 KCl 750mM	50mM Citrato Na pH6.0 KCl 1M
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
50mM Citrato Na pH6.5 NaCl 500mM	50mM Citrato Na pH6.5 NaCl 750mM	50mM Citrato Na pH6.5 NaCl 1M	50mM Citrato Na pH6.5 KCl 500mM	50mM Citrato Na pH6.5 KCl 750mM	50mM Citrato Na pH6.5 KCl 1M	50mM Citrato Na pH6.0 NaCl 750mM 10% Glicerol	50mM Citrato Na pH6.0 NaCl 200mM DNA29	50mM Citrato Na pH6.0 NaCl 700mM DNA15	50mM Citrato Na pH6.0 NaCl 500mM DNA15	50mM Citrato Na pH6.0 NaCl 200mM DNA15	CONTROL

Fig 2.6_ (a) Ensayo 2. Tabla que indica la composición y concentración de cada condición. Las que contienen DNA, éste se añadió en proporción equimolar (1:1) a la proteína.

b.

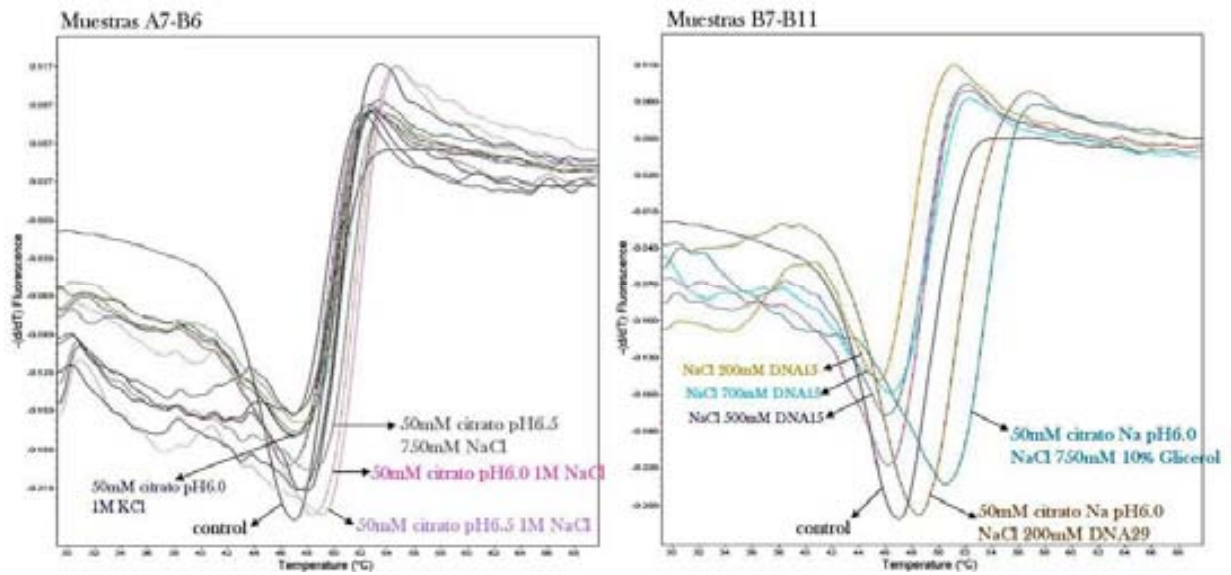


Fig 2.6_ (b) Ensayo 2. Esquema de la primera derivada de las curvas de fusión para A7-B6 y B7-B11.

a.

A1	A2	A3	A4	A5	A6	A7	A8	A9
50mM Citrato Na pH6.0 NaCl 200mM DNA29	50mM Citrato Na pH6.0 NaCl 100mM DNA29	50mM Citrato Na pH6.0 NaCl 50mM DNA29	50mM Citrato Na pH6.5 NaCl 200mM DNA29	50mM Citrato Na pH6.5 NaCl 100mM DNA29	50mM Citrato Na pH6.5 NaCl 50mM DNA29	50mM Citrato Na pH7.0 NaCl 200mM DNA29	50mM Citrato Na pH7.0 NaCl 100mM DNA29	50mM Citrato Na pH7.0 NaCl 50mM DNA29
B1	B2	B3	B4	B5	B6	B7	B8	B9
50mM Citrato Na pH6.0 NaCl 200mM DNA29 10% Glicerol	50mM Citrato Na pH6.0 NaCl 100mM DNA29 10% Glicerol	50mM Citrato Na pH6.0 NaCl 50mM DNA29 10% Glicerol	50mM Citrato Na pH6.5 NaCl 200mM DNA29 10% Glicerol	50mM Citrato Na pH6.5 NaCl 100mM DNA29 10% Glicerol	50mM Citrato Na pH6.5 NaCl 50mM DNA29 10% Glicerol	50mM Citrato Na pH7.0 NaCl 200mM DNA29 10% Glicerol	50mM Citrato Na pH7.0 NaCl 100mM DNA29 10% Glicerol	50mM Citrato Na pH7.0 NaCl 50mM DNA29 10% Glicerol

b.

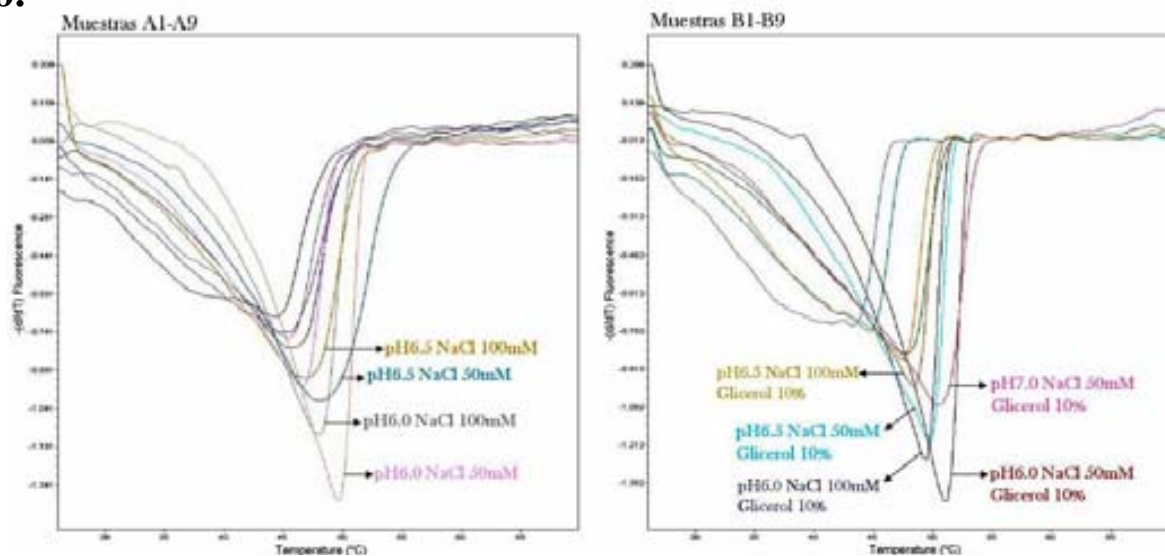


Fig 2.7_ (a) Ensayo 3. Tabla que indica la composición y concentración de cada condición. En todas las condiciones el DNA se añadió en proporción equimolar (1:1) con respecto a la proteína. (b) Esquema de la primera derivada de las curvas de fusión para A1-A9 y B1-B9.

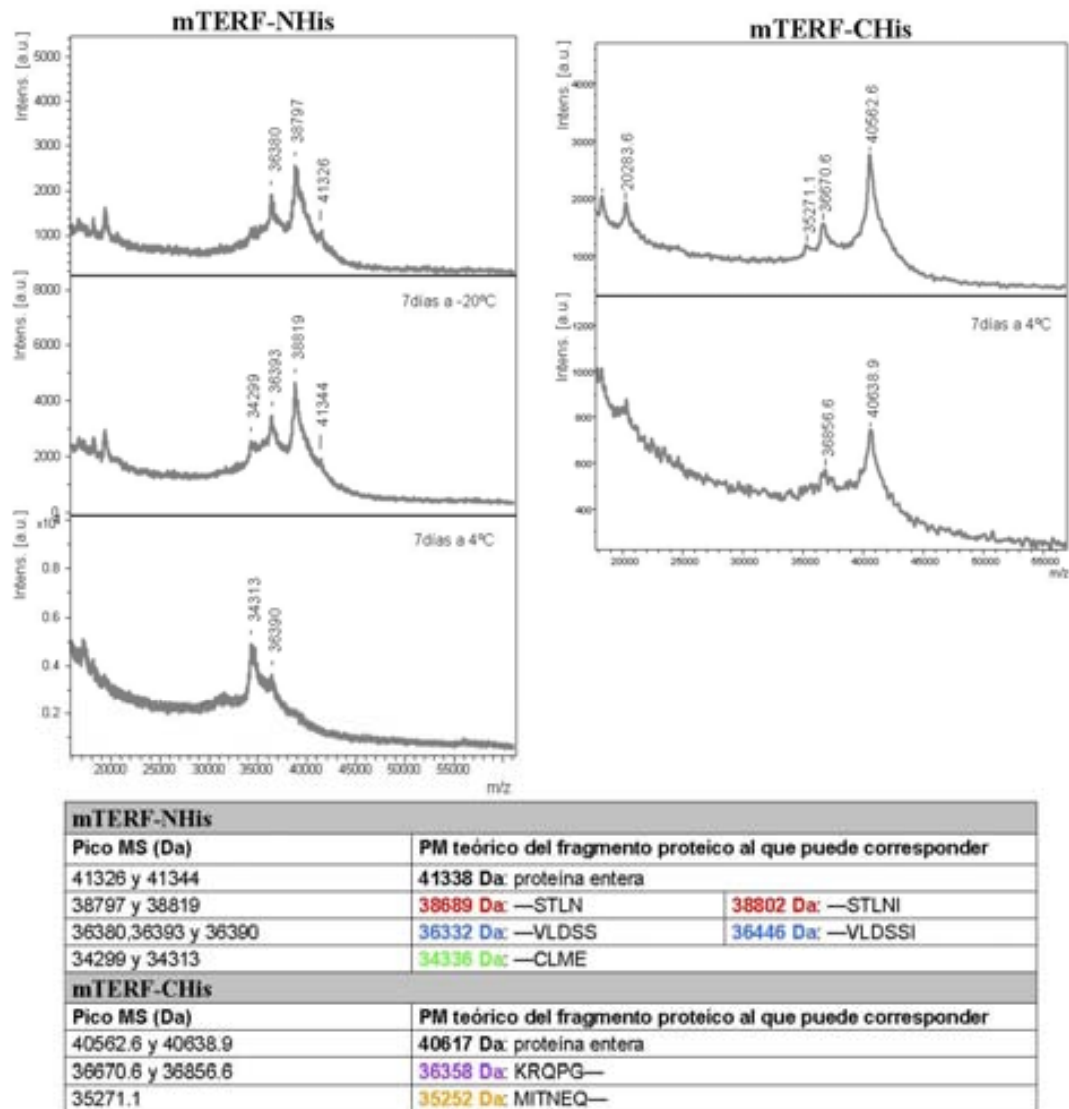
Por último, para optimizar la homogeneidad y monodispersidad de la muestra del complejo proteína-DNA_{TER} (29pb) que sería analizado posteriormente mediante experimentos de SAS, se realizaron algunas pruebas de estabilidad del complejo a distintos valores de pH, concentraciones de sal y en presencia de glicerol (Fig 2.7). Como se puede observar, la condición más estable para el complejo proteína-DNA se dio a pH6, en 50 mM NaCl y 10% de glicerol.

La técnica de *thermofluor* fue importante sobretodo para mejorar el rendimiento final en la producción de la proteína y para la optimización de la monodispersidad de la muestra para los ensayos de SAS de la proteína y del complejo proteína-DNA. La polidispersidad (Pdi) indica el grado de variación, o amplitud de una campana gaussiana que representa los pesos moleculares de una muestra. A partir de los resultados de los experimentos de *thermofluor* se modificaron los tampones de purificación para la proteína y el complejo utilizando el citrato como tampón y concentraciones mayores de NaCl (véase materiales y métodos). En nuestro caso, esta variación del tampón permitió bajar la Pdi de 50-35% a una Pdi por debajo del 20%. La composición del tampón proteico de la muestra a cristalizar pasó a ser de 50 mM citrato de sodio pH6.0, 700 mM NaCl, 10 mM BME, y el tampón de las muestras para los experimentos de SAS fue, para la proteína: 50 mM citrato sódico pH 6.0, 1 M NaCl, 10% glicerol, 10 mM BME, y para el complejo: 50 mM citrato sódico pH 6.0, 25 mM NaCl, 50 mM MgCl₂, 10 mM BME.

1.5 Degradación de mTERF

En la producción de mTERF-NHis y mTERF-CHis se observó cierta degradación. Esta degradación se hizo más evidente dejando las muestras purificadas tres meses a 4°C. En el caso de mTERF-NHis la degradación resultaba en la aparición de varios fragmentos de peso molecular diferente, mientras que en mTERF-CHis daba lugar a un fragmento estable de pM menor y probablemente más compacto, característica óptima para la cristalización (Fig 2.9). El análisis por espectrofotometría de masas (*mass spectrometry*, MS) de mTERF-NHis y mTERF-CHis nos indicó los lugares preferentes de degradación para cada uno de los constructos (Fig 2.8). Y el análisis mediante degradación de Edman del fragmento estable resultante de la degradación de mTERF-CHis nos permitió determinar

el lugar de proteólisis, entre la posición 98 y 99. Según este resultado se diseñó el tercer constructo mTERF-ΔN, en el que faltan los primeros 42 aa de la proteína madura.



Secuencia de la proteína madura

```

63      73      83      93      103
|      |      |      |      |
RLFGVK HNTDSEPLKN EDLLKLLTM GVDIIMARRKR QPGVFRRMIT NEQDLKMFLL 112
113 SNGASKEVIA SIISRYPRAI TRTPENLSKR WDLWRKIPTS DLEIVNILER SPESFFRSIN 172
173 NINLENNIKF LYSVGLTRK LCRLTNAPR TFSNSLDLNK QMVEFLQAG LSLGHNDPAD 232
233 FVRKIIFNFP FILIQSTKRK KANIEFLRST FNINSEELIV LIOGPGAEIL DLSNDYARRS 292
293 YANIEKLFPS LGCTEEVQK FVLSYEDVIF LAEKKFNDKI DCLMEENEISI SQIIEENPVL 352
353 DSSSITLSKSR IKELVNAGCN LSTINTLLS WSKRYEAKL KKLSRFA

```

Fig 2.8_ Análisis por espectrometría de masas (MS) de los lugares de degradación de mTERF. Ambas muestras, mTERF-CHis y mTERF-NHis, se analizaron por MS en el momento y al cabo de 7 días a 4°C y a -20°C en el caso de mTERF-NHis. El tampón proteico tenía una concentración de 0.5M de NaCl, lo que dificulta el análisis de la muestra mediante esta técnica. Es por esa razón que los picos presentan mucho ruido de fondo. En la tabla se muestran los posibles lugares de degradación en la proteína que corresponderían a los picos obtenidos experimentalmente por MS, deducidos a partir del PM teórico de los fragmentos resultantes teniendo en cuenta que mantienen intacto la cola de histidinas en el extremo N-terminal o C-terminal. En la parte inferior de la figura se especifica el lugar de degradación de cada uno de estos fragmentos sobre la secuencia de mTERF manteniendo el código de colores.

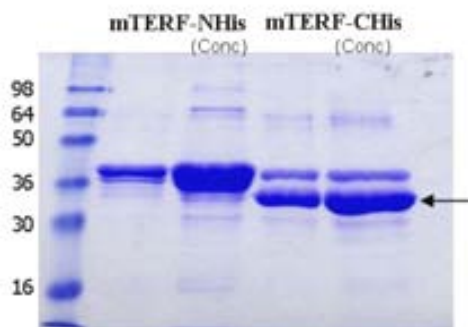


Fig 2.9_ Electroforesis en gel de poliacrilamida-SDS de las muestras de mTERF-NHis y mTERF-CHis al cabo de 3 meses de ser purificadas y guardadas a 4°C. Para cada muestra se cargó una concentrada (Conc) y una diluida. En el margen izquierdo del gel se muestran los pesos moleculares en kDa de los marcadores. La degradación de mTERF-NHis da lugar a varios fragmentos, mientras que la degradación de mTERF-CHis da lugar preferentemente a un fragmento estable. El análisis mediante degradación de Edman del segmento estable, señalado con una flecha, producto de la degradación de mTERF-CHis, muestra que los 4 primeros aminoácidos son **RMIT**, que corresponde a la degradación (en naranja) observada en el análisis por espectrometría de masas de la figura anterior.

1.6 Ensayos de proteólisis limitada de mTERF-ΔN

En cristalografía, cuanto más compacta y estable sea la especie a cristalizar, más probable será obtener cristales que difracten a una resolución aceptable gracias a un empaquetamiento más efectivo. Mediante la proteólisis limitada, en la que la degradación se da principalmente en los segmentos expuestos y flexibles de una proteína, se pueden aislar subdominios estables. Se realizaron estudios de proteólisis limitada del constructo mTERF-ΔN con tripsina, quimiotripsina, proteinasa K y papaína. A 37°C se observaba precipitación en todos los ensayos, mientras que a 20°C sólo hubo precipitación con proteinasa K.

La proteólisis de mTERF-ΔN con papaína, a una concentración de 20 μg ml⁻¹, daba lugar a un fragmento de aproximadamente 33 KDa, de manera que al cabo de 1h-2h de incubación a 20°C la proteólisis era completa (Fig 2.10a). Mediante la degradación de Edman se verificó que el extremo N-terminal del fragmento era MITN-, y mediante espectrometría de masas se determinó que el pM de dicho fragmento era 30 KDa. Así pues, la proteólisis correspondía a una degradación de 39 aa de la parte C-term de la proteína madura (sin contar la cola de His) (Fig 2.10c).

Para estudios posteriores, teniendo en cuenta los resultados de la proteólisis limitada de mTERF- Δ N, se diseñó el constructo mTERF- Δ C, que incluía los residuos desde Ser56 hasta Lys360.

También es posible, mediante la técnica de proteólisis limitada, obtener evidencias de la interacción de la proteína con el DNA, de manera que si la velocidad de digestión de la proteína de interés varía en presencia del DNA_{TER} 28pb, se puede concluir que existe interacción. Se realizaron dos experimentos de proteólisis idénticos, uno con mTERF- Δ N y el DNA y el otro únicamente con la proteína. La velocidad de digestión de la proteína era menor en presencia del DNA (Fig 2.10a). Este ensayo evidenciaba, por tanto, que existía interacción de mTERF- Δ N con el DNA_{TER}.

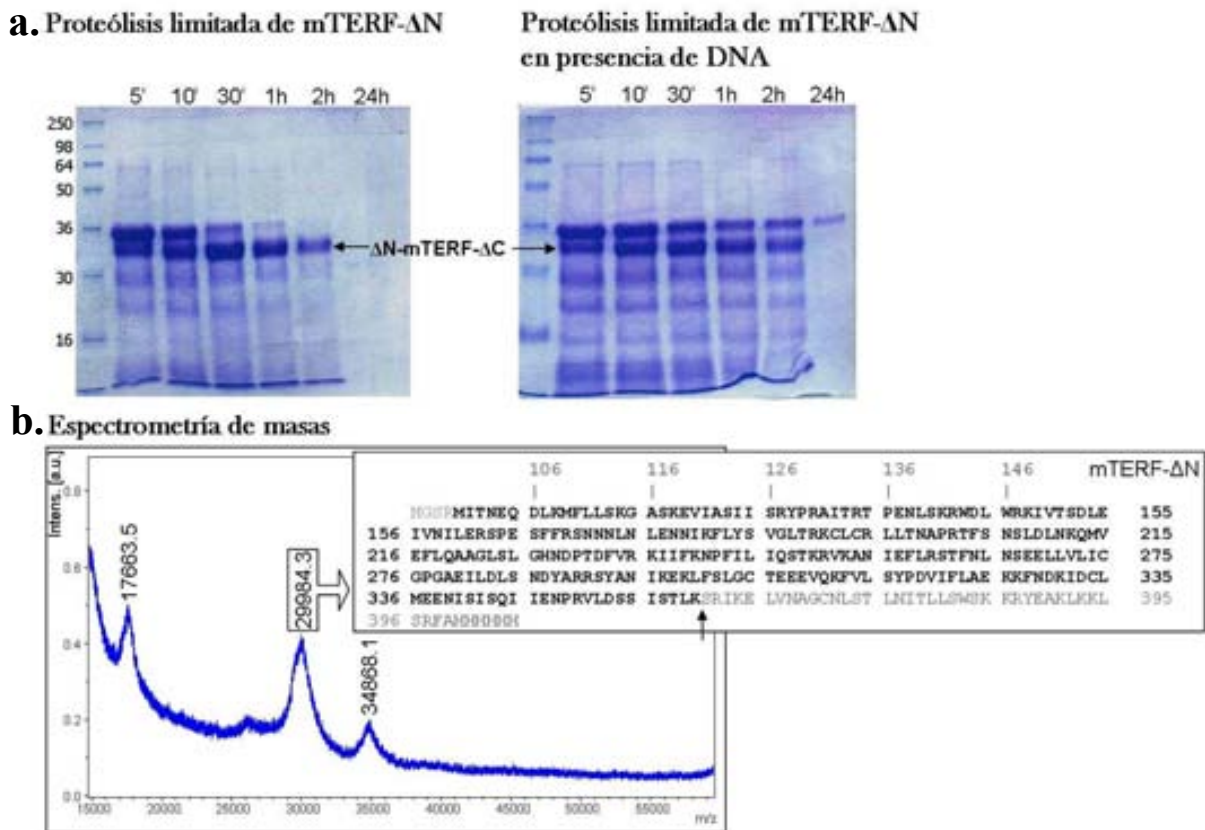


Fig 2.10_ (a) Resultados de la electroforesis en geles de poliacrilamida-SDS del ensayo de proteólisis limitada de mTERF- Δ N con papaína en ausencia de DNA (izq.) y en presencia de DNA (dcha.). La flecha indica el fragmento Δ N-mTERF- Δ C de ~33KDa resultante de la proteólisis. Comparando ambos ensayos de proteólisis, en ausencia y presencia de DNA, observamos que la velocidad de proteólisis es menor en presencia de DNA, lo que indica que existe una interacción entre la proteína y el DNA que dificulta la proteólisis de ésta. (b) El análisis por MS y degradación de Edman nos permitió determinar que la proteólisis de mTERF- Δ N se produce en la parte C-term y corresponde a unos 40 aa (gris). La degradación de Edman muestra una variedad en los aminoácidos del N-terminal (GSRM-/MITNE-/RMIT-). Según la masa observada la secuencia del fragmento proteolizado sería el que está indicado en negro.

2_ Obtención y caracterización del complejo mTERF-DNA

2.1 Ensayos de unión a DNA de los constructos de mTERF

Los primeros ensayos de unión de mTERF a DNA se realizaron únicamente con el DNA_{TER} (28pb) (Fig 2.14). Los gels nativos de retardo de mTERF-NHis, mTERF-CHis y mTERF-ΔN con el DNA en diferentes proporciones proteína:DNA -donde para una cantidad constante de proteína se incrementaba el DNA o viceversa-, muestran sistemáticamente una banda de retardo del DNA (Fig 2.11), indicando que existía interacción entre la proteína recombinante y el DNA.

Para determinar la capacidad de unión al DNA de ΔN-mTERF-ΔC, resultante de la proteólisis limitada de mTERF-ΔN, el fragmento se purificó mediante una columna de gel filtración, se mezcló con el DNA_{TER} de 28pb (Fig 2.14), se dializó y cargó nuevamente en una columna de exclusión molecular. Las fracciones del único pico de elución resultante se cargaron en un gel nativo de retardo que mostró que la proteína no unía el DNA (Fig 2.10b).

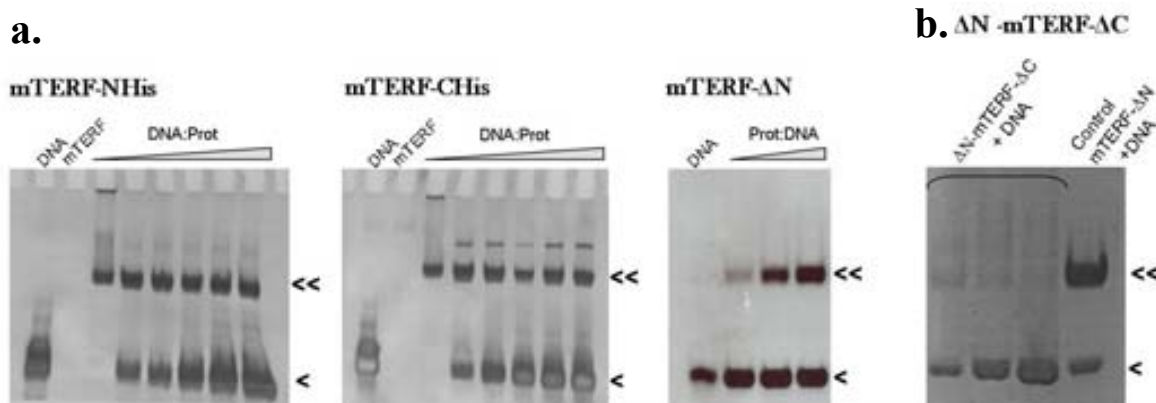


Fig 2.11_ (a) Electroforesis de retardo en gel de poliacrilamida en condiciones nativas de mTERF-NHis, mTERF-CHis y mTERF-ΔN con DNA_{TER}. En los gels de mTERF-NHis y mTERF-CHis se cargaron, además de los controles de DNA y proteína, muestras con relaciones crecientes de DNA:proteína, es decir, para una cantidad constante de proteína se incrementó la cantidad de DNA. En el gel de mTERF-ΔN se mantiene constante la cantidad de DNA y se aumenta la cantidad de proteína. En todos los casos la tinción se realizó con bromuro de etidio. La doble flecha << señala la banda de retardo del DNA resultado de la formación del complejo y la flecha < señala la banda que corresponde al DNAlibre. (b) Ensayo de unión a DNA del fragmento ΔN-mTERF-ΔC resultante de la proteólisis de mTERF-ΔN. En el gel nativo se cargó el pico de elución resultante de la columna de exclusión molecular de la muestra que contenía ΔN-mTERF-ΔC y DNA. Como control se cargó también el complejo mTERF-ΔN-DNA. Como muestra el gel nativo, ΔN-mTERF-ΔC no une DNA ya que no se observa la banda que corresponde al retardo del mismo.

Para la cristalización del complejo proteína-DNA, es importante conocer los parámetros que influyen en la formación de éste y en qué medida. Con el fin de determinar las condiciones óptimas de formación del complejo mTERF-DNA, se estudió el efecto de la longitud del DNA, de la concentración de sal y del pH. Para ello, utilizando la proteína recombinante mTERF- Δ N, se realizaron ensayos mediante geles nativos de retardo que se detallan a continuación.

Estudio de la formación del complejo en función de la longitud de DNA

Los primeros cristales obtenidos fueron con mTERF- Δ N y el DNA de 29 pb. Como se explica más adelante, los graves problemas de anisotropía nos llevaron a intentar la cristalización con oligonucleótidos más cortos. En este ensayo, se comprobó la unión de mTERF- Δ N a los oligonucleótidos de 12, 13, 15, 17, 19, 21, 27, 28 y 31 pb que se indican en la Tabla 2.8 y que contienen, excepto el oligonucleótido de 12 pb, la región esencial para la terminación de la transcripción (Christianson & Clayton, 1988), Fig 1.15), en condiciones tamponadas a pH 7.4 y 200 mM NaCl. Como se observa en la Fig 2.12a, el retardo empieza a ser consistente a partir de un oligonucleótido de 19pb, y para longitudes de 27, 28 y 31pb, el retardo es sistemático.

Estudio de la formación del complejo en función de la concentración de sal

A partir de los primeros cristales obtenidos de mTERF- Δ N-DNA29pb y, debido a problemas cristalográficos que se detallan más adelante, se optó por cristalizar la proteína con oligonucleótidos más cortos. Dado que la diálisis de la mezcla de la proteína con oligonucleótidos menores de 23 pb para la posterior cristalización daba lugar a la precipitación, se optó por mezclar directamente la proteína y el DNA en la muestra de cristalización. En este caso, el tampón resultante de la mezcla de la proteína purificada con el DNA tenía una concentración final de aproximadamente 500 mM de NaCl. Al realizar los ensayos de cristalización, se mezclaba la muestra que contenía la proteína y el DNA con la condición de cristalización 1:1 μ l, de manera que dependiendo de la condición, la concentración de sal quedaba afectada, pero en general eran concentraciones de sal entre 250 y 750 mM NaCl. Teniendo en cuenta que la mayoría de interacciones que sostienen el complejo proteína-DNA son electrostáticas, a

mayor concentración de sal se espera que la cantidad de complejo proteína-DNA formado disminuya debido al apantallamiento de cargas electrostáticas que median la interacción.

En el gel nativo de la Fig 2.12b se muestra la dependencia de la unión de la proteína con el DNA en función de la sal, de manera que al aumentar la concentración de ésta, disminuye la cantidad de complejo formado. Aun así, a 500mM NaCl aun existe formación de complejo.

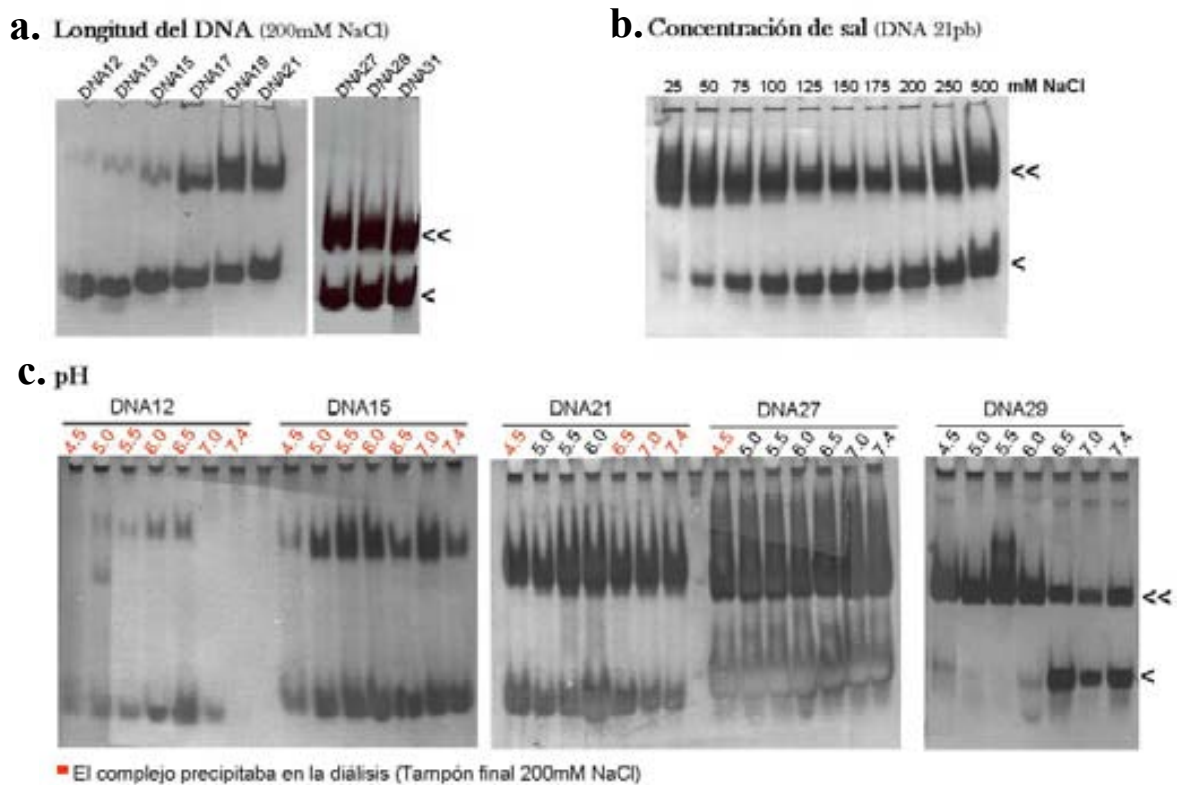


Fig. 2.12_ Estudio de la formación del complejo mTERF-AN con el DNA. (a) Estudio de la formación del complejo en función de la longitud del DNA (12, 13, 15, 17, 19, 21, 27, 28 y 31pb) mediante retardo en geles nativos y tinción con un intercalador de DNA. La existencia de una banda de retardo indica que se forma el complejo. La doble flecha << señala la banda de retardo del DNA resultado de la formación del complejo y la flecha < señala la banda que corresponde al DNAlibre. (b) Estudio del efecto de la sal. A mayor concentración de sal, la cantidad de complejo proteína-DNA formado disminuye y se observa una mayor cantidad de DNA libre. (c) Estudio del efecto del pH en los complejos con DNA de 12, 15, 21, 27 y 29pb. En el ensayo con DNA27pb hay un problema de exceso de DNA. En rojo se indican las muestras donde se observa precipitación. Para todo el rango de pHs, las muestras con DNA de 12 y 15pb precipitaban en el botón de diálisis. Con el DNA de 21pb, se observó precipitación a pHs 4.5, 6.5, 7.0 y 7.5 y con el DNA de 27pb sólo se observó precipitación a pH 4.5.

Estudio del pH

Como las interacciones electrostáticas entre la proteína y el DNA se dan entre especies cuya protonación o desprotonación depende del pH, se realizaron ensayos de retardo con soluciones tamponadas a diferentes pHs. En principio,

debido a este fenómeno, a pHs relativamente bajos la interacción está favorecida por la protonación de grupos amino de la proteína, ya que ésta dispone de mayor carga para interactuar con el DNA.

Para el DNA de 12 y 15 pb, el complejo precipitaba en la diálisis por tanto los resultados no fueron concluyentes. Con el DNA de 21pb el complejo precipitaba a pHs 4.5, 6.5, 7.0 y 7.4, de manera que parecía que el complejo era más estable a pHs 5.0-6.0. En el ensayo del DNA de 27 pb se observó precipitación únicamente a pH 4.5, pero hay un problema de exceso de DNA y las bandas no se resolvieron bien. Para el DNA de 29 pb se observa que a pHs bajos (entre 4.6 y 6) está favorecida la formación del complejo (Fig 2.12c).

2.2 Obtención del complejo para la cristalización

Para los ensayos de cristalización de mTERF- Δ N con los DNAs de 27 pb, 28 pb, 29 pb, 30 pb y 31 pb, una vez mezclado el DNA con la proteína (500 mM NaCl) y después de la correspondiente diálisis hasta una concentración de 100 a 200 mM NaCl, la muestra se inyectaba en una columna de exclusión molecular para purificar el complejo.

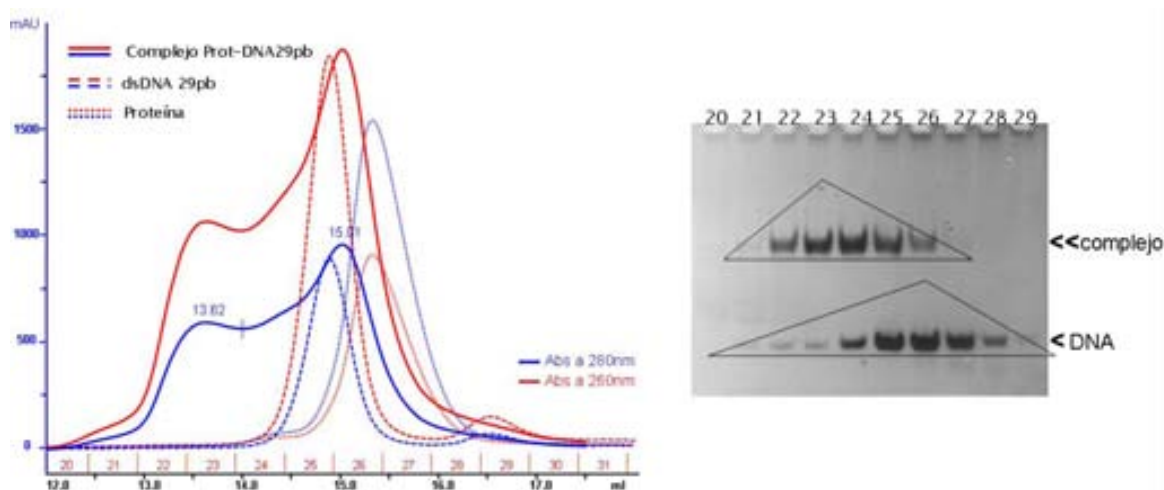


Fig 2.13_ (Izq.) En el cromatograma se muestra el perfil de elución de la proteína mTERF- Δ N (punteado), del DNA de 29 pb (rayado intermitente) y del complejo (liso). En azul se indica la abs a 280 nm y en rojo la abs a 260 nm. (Dcha.) El gel nativo teñido con un intercalante de DNA (bromuro de etidio o *sybr safe*), permite distinguir qué contiene cada fracción eluida. Se indican sobre el gel (triángulo negro) los dos picos que se observan superpuestos en el perfil de elución del complejo y que corresponden al complejo y al DNA en exceso.

El perfil de elución de la mezcla mostraba dos picos superpuestos correspondientes al complejo y al exceso de DNA, como se puede ver en el gel nativo de la Fig 2.13. El Ve del complejo está ligeramente desplazado respecto al de la proteína sola (~1.6 ml), lo que es un indicativo de la formación de una especie con un radio hidrodinámico mayor. Las fracciones que contienen mayoritariamente el complejo (las fracciones de la 22 a la 25 del perfil de elución) son las que se utilizaban en la cristalización.

Para la cristalización de mTERF-ΔN con oligonucleótidos menores de 23 pb, el complejo precipitaba al intentar dializarlo a concentraciones menores de sal. Así pues, la muestra utilizada para la cristalización se obtenía mezclando directamente la proteína y el DNA en una relación molar 1:2 respectivamente.

3_Cristalización del complejo

Los primeros cristales obtenidos fueron del complejo de mTERF truncada en N-terminal (mTERF-ΔN) y el DNA de 29 pb (cohesivo, Fig 2.14), a una concentración de proteína de 2.5 mg ml⁻¹ (tampón del complejo a 200 mM NaCl) y 3.0 mg ml⁻¹ (tampón del complejo a 100 mM NaCl) y en las condiciones de cristalización: 10% PEG 8K, 0.1 M MgCl₂ y 50mM citrato sódico pH5.5. Estos cristales fueron obtenidos a 20°C al cabo de una semana mediante la técnica de difusión de vapor (gota sentada). Se realizó la optimización de las condiciones químicas de cristalización, consiguiendo los mayores cristales a una concentración de proteína de 7mg ml⁻¹ (tampón de purificación del complejo a 200mM NaCl) y una condición de cristalización entre 7-11% PEG 8000, 130-160 mM MgCl₂ y 0.05 M MES pH 5.6. La aplicación de la técnica de *macroseeding* permitió el crecimiento de los cristales hasta un tamaño suficiente para que los datos de la difracción de rayos X fueran procesables (Fig 2.15).

No obstante, estos cristales presentaron graves problemas de anisotropía y desorden. Nos permitieron resolver la estructura, pero dieron lugar a unos mapas de densidad electrónica muy distorsionados donde sólo se pudo trazar parcialmente la estructura de la

proteína y, además, no se visualizaba densidad para el DNA. Aun así, éste estaba presente ya que en las imágenes de difracción de uno de los cristales pudimos distinguir las reflexiones difusas típicas del *stacking* de las bases del B-DNA ($\sim 3.4 \text{ \AA}$) a lo largo del eje c^* . Paradójicamente, el parámetro de celda de este eje (35 \AA) era demasiado pequeño para albergar un oligonucleótido de 29 pb (Fig 2.17). Por tanto, una misma molécula de DNA ocuparía más de una celda unidad provocando un desorden que impedía observar la densidad electrónica correspondiente al DNA. Se intentó cristalizar el complejo de mTERF- Δ N con oligonucleótidos de menor tamaño (10, 11, 12, 13, 14, 15, 19 y 21 pb) que pudiesen encajar en el parámetro de celda unidad c . Se obtuvieron cristales con oligonucleótidos de 19 (romo), 15 (cohesivo), 12 (romo) y 11 pb (romo) (Tabla 3.8) en las condiciones de cristalización 3-10% PEG 8000, 10-120mM MgCl_2 y 0.05M citrato de sodio pH 5.5 (Fig 2.15).

Por último, a causa de la proteólisis espontánea que presentaban mTERF-NHis y mTERF-CHis, se intentó la cristalización del complejo mTERF (entera) con DNA realizando en un mismo día la purificación de la proteína y los ensayos de cristalización con oligonucleótidos de distintas longitudes (10, 11, 12, 13, 14, 15, 16, 19, 21 y 28 pb) sin un paso de diálisis previo. Tras ajustar los parámetros químicos de la condición de cristalización, los mejores cristales se obtuvieron con el DNA de 15 pb (cohesivo) en las condiciones de cristalización 1-3% PEG8000, 0.2-0.5 M KCl, 0.05 M cacodilato de sodio, pH 7 (Fig 2.15).

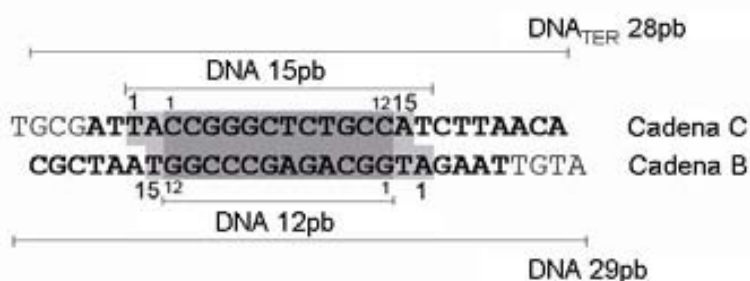


Fig 2.14_Esquema del DNA_{TER} (28pb) resultante de los ensayos de *footprint* con DNAasa (Kruse *et al.*, 1989) donde se indican los oligonucleótidos presentes en los cristales con los que se resolvió la estructura de mTERF (DNA 15pb, cohesivo, en gris claro) y mTERF- Δ N (DNA 12pb, romo, en gris oscuro), el oligonucleótido con el que se obtuvieron los primeros cristales (DNA 29pb) y el DNA que se usó en los experimentos de SAXS (28 pb). Cadena B y C se refiere a la nomenclatura utilizada en el pdb y corresponden a la hebra L y H respectivamente.

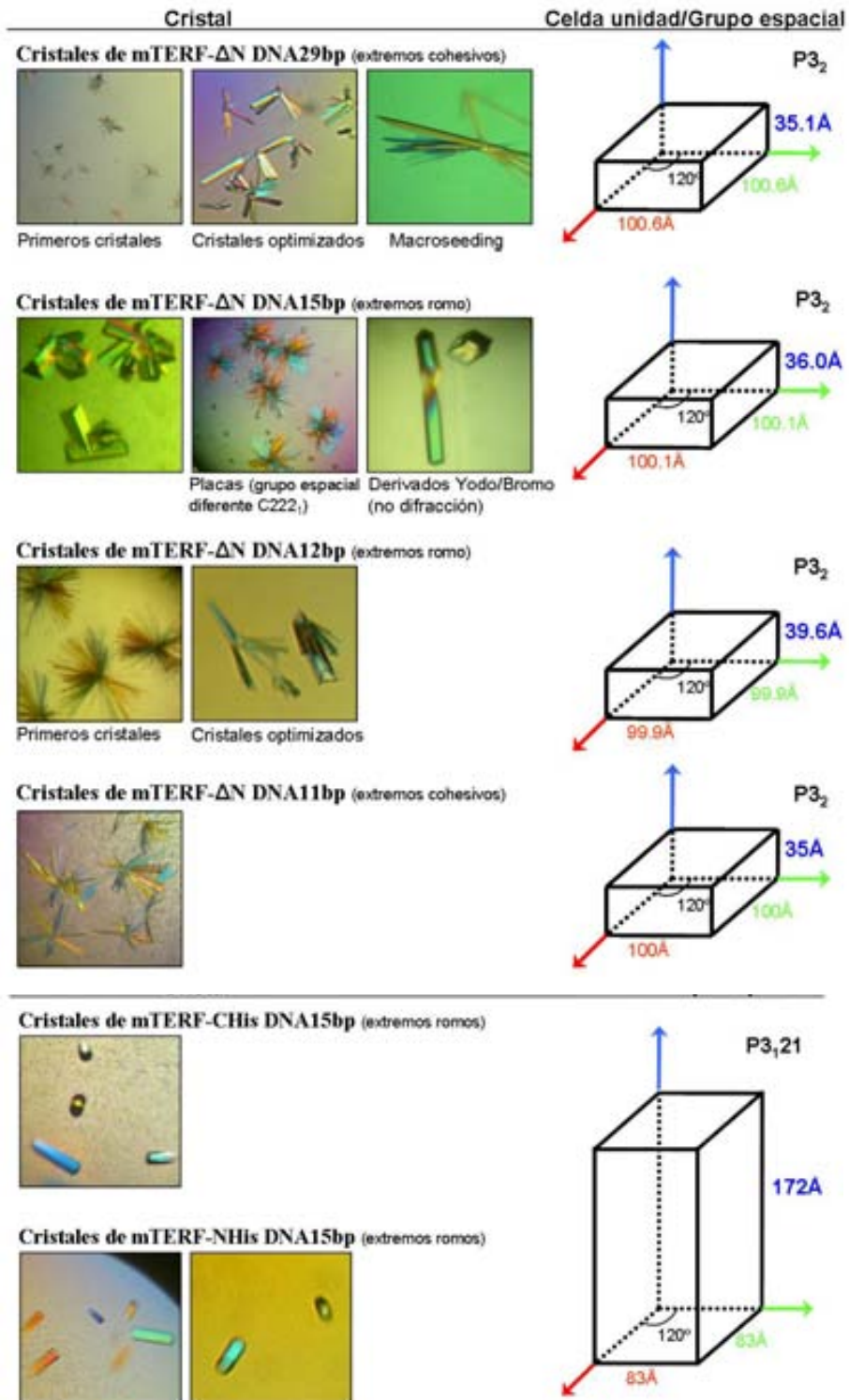


Fig 2.15_ Imágenes de los cristales obtenidos de complejo mTERF- Δ N-DNA y mTERF-DNA. A la derecha se muestra la celda unidad y el grupo espacial de cada uno. La variación de la longitud del DNA en los cristales de mTERF- Δ N-DNA resultaba en una variación del parámetro de celda en el eje z, lo que sugería que la molécula de DNA se encuentra alineada con dicho eje, hipótesis confirmada por el patrón de difracción (ver texto).

Todos los cristales de complejo de mTERF- Δ N tenían una forma de prisma trigonal, mientras que los cristales de complejo mTERF (entera), tenían forma de prisma hexagonal. En el caso del complejo de mTERF- Δ N con el DNA 15pb se obtuvieron también unos cristales en forma de placas con un grupo espacial diferente, C222₁.

3.1 Confirmación de la presencia de DNA en los cristales

Inicialmente, una vez resuelta la estructura de los cristales de mTERF- Δ N-DNA 29pb se podía apreciar densidad para la proteína, pero no para el DNA. Se comprobó por métodos bioquímicos la presencia de DNA en los cristales mediante tres técnicas.

Electroforesis en geles de poliacrilamida en SDS y en condiciones nativas

En este caso se cargaron en un gel varios cristales previamente lavados con la solución de cristalización. Para la electroforesis en geles de poliacrilamida-SDS la muestra se desnaturalizó antes de cargarla y el gel se tiñó con plata, que detecta la proteína y el DNA. El gel nativo de retardo se tiñó con el intercalante *Sybr Green*. En el primer caso se observó una banda correspondiente a la proteína y una al DNA. En el segundo caso la banda de retardo confirmó la presencia del complejo proteína-DNA (Fig 2.16a).

Cristalización con DNA marcado con un cromóforo

Esta técnica consiste en cristalizar el complejo con un oligonucleótido marcado con un cromóforo, Cy5 en este caso. El color claramente azul de los cristales obtenidos indicó que los cristales contenían DNA (Fig 2.16b).

Detección de ácidos nucleicos por fluorescencia

Los cristales se incubaron con *Sybr Green*, que une específicamente ácidos nucleicos, y se observaron mediante un microscopio de fluorescencia. En las primeras imágenes, se utilizó como control cristales de insulina. En el segundo ensayo, el control se realizó con cristales de lisozima crecidos en presencia de DNA. En la Fig 2.16c se muestran las fotos con luz blanca para visualizar los cristales de lisozima y de complejo mTERF- Δ N-DNA, y con luz ultravioleta para distinguir la presencia de DNA en los cristales de complejo. La fluorescencia

observada en los cristales del complejo mTERF- Δ N-DNA29pb confirma la presencia de DNA en el cristal.

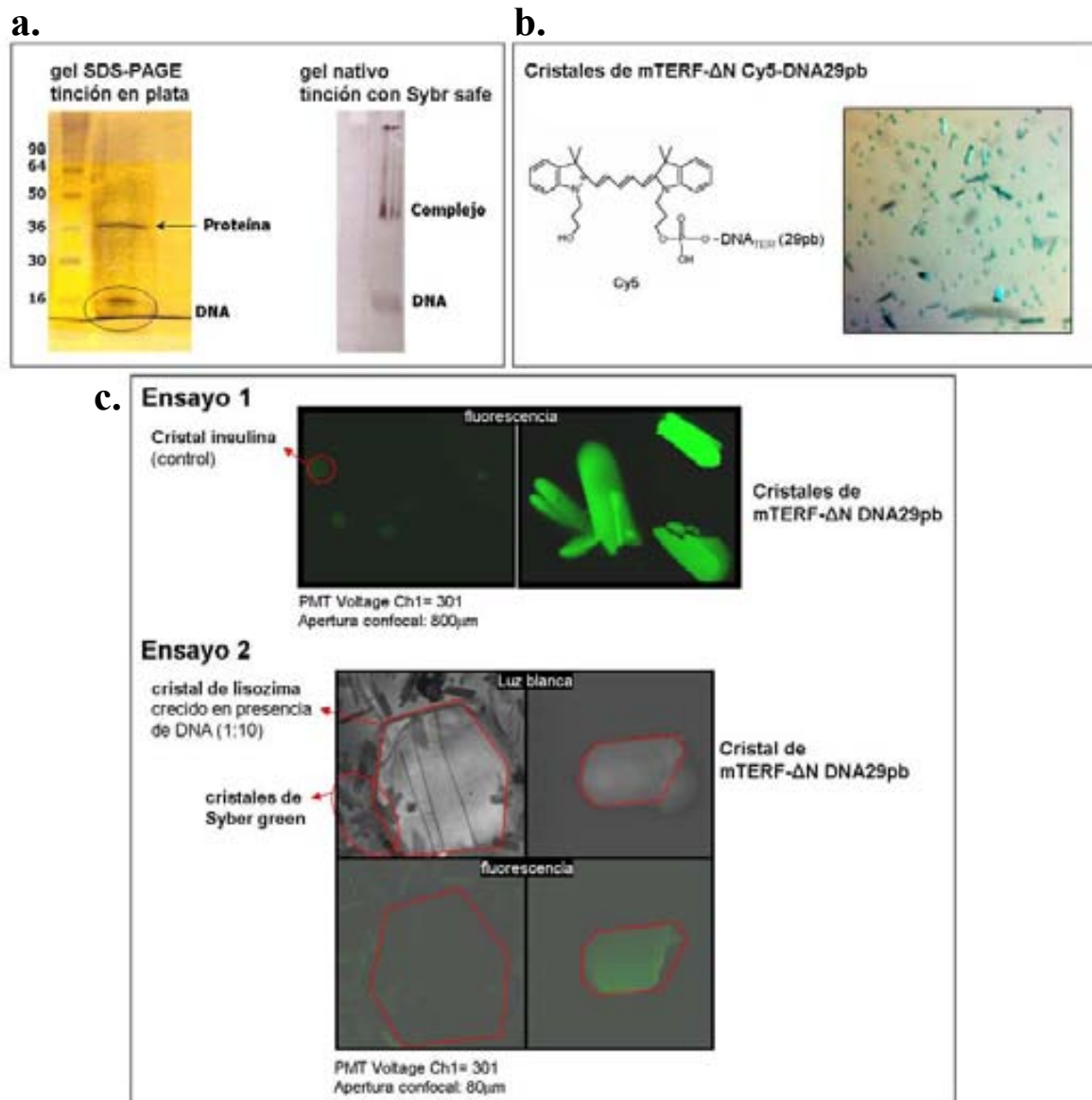


Fig 2.16_ (a) (Izq.) Resultado de la electroforesis en gel de poliacrilamida-SDS y (dcha.) gel nativo de retardo de varios cristales previamente lavados con la solución de cristalización. En el caso del gel nativo, la presencia de PEG en la condición de cristalización y en el tampón de lavado de la muestra provoca que ésta no corra correctamente en el gel. (b) Cristales azules del complejo mTERF- Δ N con un oligonucleótido que contiene el cromóforo Cy5 cuya fórmula se muestra en la figura. (c) Imágenes del microscopio de fluorescencia de cristales control (insulina y lisozima para el ensayo 1 y 2 respectivamente) y de cristales de mTERF- Δ N-DNA29pb. En el segundo ensayo se muestra también la imagen con luz blanca para la visualización de los cristales. La existencia de cristales de *Sybr Green* se debe a que se añadió un exceso de éste, aun así la fluorescencia de éstos es mucho menor comparada con los cristales de complejo. Los cristales de mTERF del primer ensayo muestran una mayor intensidad de fluorescencia debido a que la apertura confocal es diez veces mayor que en el segundo ensayo.

4 _ Procesamiento de datos y resolución de la estructura

Dado que para resolver la estructura cristalográfica no había un modelo inicial para poder utilizar la técnica de reemplazo molecular, se procedió a la obtención de fases cristalográficas experimentales mediante el método de dispersión anómala a múltiples longitudes de onda (*Multiple-wavelength Anomalous Dispersion*, MAD) a partir de cristales de proteína modificada con Se-Met. Con los cristales del complejo mTERF- Δ N (con 5SeMet) y DNA29 pb se obtuvo un conjunto de datos recogidos en el pico de absorción del Se (Fig 3.22a). En el procesamiento se localizaron 3 Se para 310 aa (normalmente se asume que con 1 Se se obtiene información de fase para cada 80-100 aa) (Hendrickson & Ogata, 1997), pero tras asignar las fases a todos los datos, el mapa de densidad electrónica resultante aparecía sin rasgos distintivos (Fig 2.17b_{arriba}), impidiendo el trazado de la estructura. Pensando que podía ser un problema de faseado se recurrió a mutantes para aumentar el número de metioninas. Para escoger los lugares de mutación, se alineó la secuencia de mTERF con las secuencias proteicas de los homólogos en rata, ratón y cerdo y, teniendo en cuenta la posición de las metioninas en las secuencias homólogas, se eligieron tres lugares de mutación para mTERF repartidos por la proteína. Se clonaron, purificaron y cristalizaron con el oligonucleótido de 29pb cuatro mutantes, Mut1 (Val150Met), Mut2 (Leu323Met), Mut3 (Val323Met) y Mut1-3. Esta estrategia, aunque no resolvió el problema de la calidad de los mapas, fue muy útil posteriormente a la hora de asignar la secuencia.

El análisis de los datos y la inspección del aspecto de los mapas indicaron que en realidad había un problema de anisotropía. La difracción anisotrópica se da cuando existe una dependencia direccional en la calidad de la difracción. Se utilizó el servidor <http://www.doe-mpi.ucla.edu/~sawaya/anisyscale/> para evaluar la anisotropía de los cristales de mTERF- Δ N-DNA29pb, y resultó que la dirección de menor difracción coincidía con el eje c^* con una resolución de 3.3 Å frente a una resolución de 2.7 Å en las otras dos direcciones (Fig 2.17a). Una consecuencia de la anisotropía, si ésta no se corrige adecuadamente, es la deformación de los mapas de densidad electrónica tal y

como se puede observar en la Fig 2.17b. En el eje c, de menor resolución, los mapas aparecían como alargados. Así, la corrección anisotrópica con el programa XPREP (Sheldrick, 2008) de los datos de los cristales de mTERF- Δ N SeMet con el DNA 29 pb junto con el método ‘*ab initio*’ de trazado de hélices α (Rodríguez *et al.*, 2009) (Materiales y Métodos), permitieron obtener mapas de densidad interpretables (Fig 2.17b), aunque sólo fue posible el trazado de una parte de la proteína, sin poderse distinguir ninguna densidad para el DNA, ni poderse refinar la estructura.

Como se ha mencionado anteriormente, si observamos detenidamente una de las imágenes de difracción del cristal, podemos distinguir dos manchas a una resolución de $\sim 3.4\text{\AA}$ que corresponden al *stacking* o apilamiento de las bases del B-DNA (Fig 2.17c). Este resultado sugería que el DNA se encontraba en el eje z del cristal. Los primeros cristales obtenidos con el oligonucleótido de 29 pb mostraban $c=35.1\text{\AA}$, un parámetro de celda unidad demasiado pequeño para el DNA que contenían (29 pb serían $\sim 98\text{\AA}$). Una posibilidad era que el DNA estuviera formando ‘fibras’ a lo largo del eje z de manera que éstas quedan deslocalizadas si se toma como referencia la celda unidad. Por esa razón, se realizaron ensayos de cristalización con oligonucleótidos de menor longitud. La variación de longitud del DNA resultó en una variación en el eje z del parámetro de celda c entre valores de 35 a 39.6\AA (Fig 2.15).

La medición de los cristales de mTERF- Δ N con el oligonucleótido de 15pb en el sincrotrón y en el laboratorio de G. Sheldrick en Göttinghen (Alemania) posibilitó la obtención de datos de mayor resolución y de elevada redundancia que nos permitieron acabar de trazar la proteína y empezar a ver parte de densidad electrónica correspondiente al DNA (Fig 2.18c).

El valor de c de la celda unidad que más se aproxima a la longitud del DNA que contiene el cristal es en el caso del oligonucleótido de 12 pb, donde $c= 39.6\text{\AA}$, ya que considerando una distancia entre las bases de 3.4\AA , tal y como se observa en el patrón de difracción, entonces $12\text{ pb} \times 3.4\text{\AA} (\text{stacking}) = 40.8\text{\AA}$. Precisamente éstos son los cristales con los que finalmente se pudo trazar completamente mTERF- Δ N y el DNA (Fig 2.18c) y que fueron utilizados para el refinado final (Tabla 2.19).

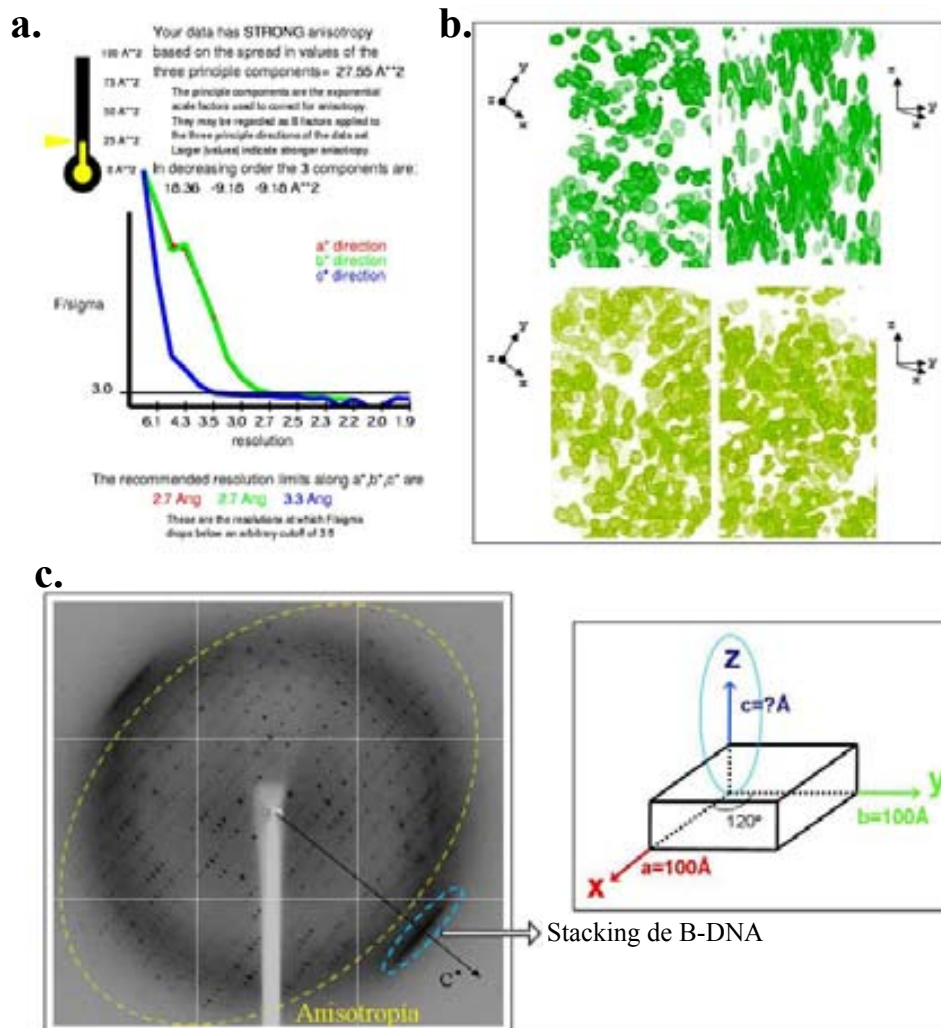


Fig 2.17_ (a) Evaluación de la anisotropía. En abscisas se representa el factor de estructura de los datos respecto al error de los mismos y, en ordenadas, la resolución. El eje de menor resolución corresponde al eje c^* (azul). (b) Mapas de densidad electrónica en dos orientaciones, sin corrección anisotrópica (arriba en verde) y con corrección anisotrópica (abajo en amarillo). Se puede apreciar que antes de la corrección los mapas tienen un aspecto alargado en la dirección z –compárese paneles izq. y dcha.–. En cambio, tras la corrección, la forma de los mapas es coherente en todas las direcciones. (c) (Izq.) Imagen de difracción de un cristal de mTERF- Δ N-DNA15pb. La línea discontinua amarilla indica la anisotropía de la difracción y como se observa en la figura, la dirección de menor resolución corresponde al eje c^* . La línea discontinua azul muestra la reflexión correspondiente al stacking de las bases del DNA en forma B. (Dcha.) Celda unidad. El círculo azul muestra el eje z a lo largo del cual se localiza el DNA. Según la longitud del DNA cristalizado el parámetro de celda c varía (ver Fig 2.15).

A la vista que la interacción de mTERF- Δ N con el DNA se daba en la parte C-terminal (ver más adelante), para los ensayos de cristalización de la proteína entera se decidió expresar y purificar el constructo mTERF-NHis, con la cola de his N-terminal. La difracción de los cristales de mTERF-NHis en complejo con el DNA de 15pb no presentó problemas de anisotropía. Para resolver la estructura de los cristales de la proteína entera, se utilizó la estructura de mTERF- Δ N para el reemplazo molecular. El mapa de densidad resultante permitió trazar el DNA y casi toda la estructura de la proteína excepto el

extremo N-terminal. Se realizaron derivados de Hg y Au para complementar las fases del modelo con fases experimentales, pero el problema parecía ser que dicho extremo se encontraba desordenado en el cristal, quizá como consecuencia de la cola de histidinas.

Así pues, se realizaron ensayos de cristalización con el constructo mTERF-CHis, cuyos cristales con el oligonucleótido de 15pb presentaron unos parámetros de celda y grupo espacial casi idénticos a los anteriores. La estructura se resolvió mediante reemplazo molecular. Como la parte N-terminal estaba más definida en estos cristales del complejo mTERF-CHis con el DNA 15pb, se utilizaron estos datos en el refinado final (Tabla 2.19). En la Fig 2.18a y b se muestran los mapas de densidad electrónica de la estructura de este complejo.

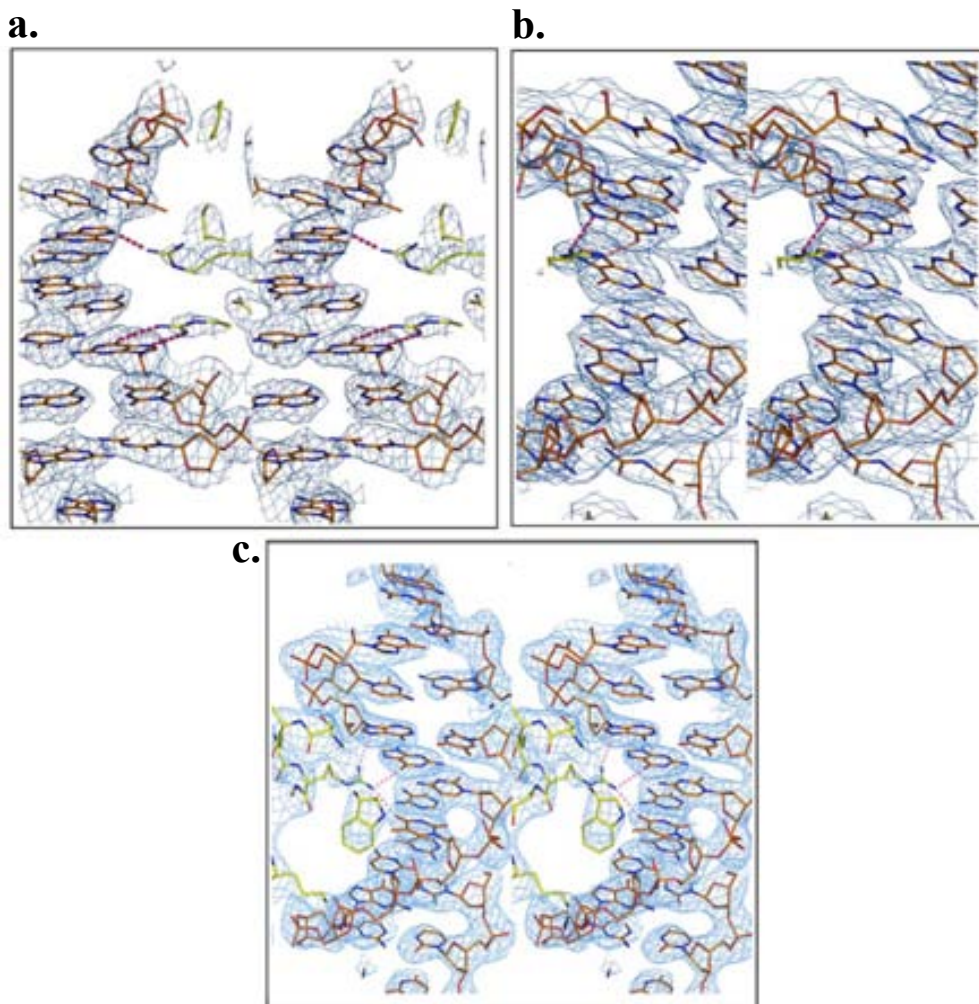


Fig 2.18_ (a) Detalle del mapa de densidad electrónica (1σ) en la interfase entre el subdominio N-terminal de mTERF y el DNA1. Se pueden apreciar las interacciones de las argininas 202 y 169 con las bases del DNA (b) Detalle del mapa de densidad electrónica del subdominio C-terminal de mTERF (Arg387) y el DNA2. (c) Detalle del mapa de densidad electrónica (1σ) del subdominio c-terminal mTERF- Δ N y el DNA2. Se muestran las interacciones de la Arg387 con las correspondientes bases nitrogenadas.

Conjunto de datos mTERF	mTERF-ΔN / DNA29bp (SeMet)			mTERF-ΔN / DNA15bp	mTERF-ΔN / DNA12bp	mTERF / DNA15bp
Recolección de datos						
Grupo Espacial	P3 ₂			P3 ₂	P3 ₂	P3 ₁ 21
Parámetros de celda unidad						
$a=b, c$ (Å)	100.6, 35.1			100.1, 36.0	99.85, 39.6	83.24, 171.86
α, β, γ (°)	90, 90, 120			90, 90, 120	90, 90, 120	90, 90, 120
	Pico	Inflexión	Remota			
Longitud de onda (Å)	0.97951	0.97971	0.97564	0.9755	0.9755	
Rango resolución (Å)	33.0 - 2.6	33.0 - 2.6	33.0 - 2.6	43.0 - 2.3	39.5 - 2.3	66.5 - 3.1
R _{sym} (última capa) (%)	7.0 (39.1)	7.8 (50.9)	8.3 (51.7)	2.2 (42.7)	4.7 (45.7)	9.3 (59.8)
Intensidad promedio (última capa) <[<I> / σ(<I>)]>	9.2 (2.5)	9.2 (1.9)	8.7 (1.9)	33.7 (1.8)	12.5 (2.1)	5.7 (1.3)
Completitud (%) (última capa)	99.8 (99.6)	99.7 (99.5)	99.8 (99.5)	95.6 (50.0)	99.9 (99.8)	97.9 (97.9)
Redundancia	4.5 (4)	4.5 (4)	4.5 (4)	31.0 (8.0)	9.0 (4.5)	10.6 (11.0)
Refinamiento:						
Resolución (Å)					31.0 - 2.4	44.9 - 3.1
No. reflexiones					17251	12704
R _{factor} / R _{free} (%)					22.8/27.8	23.7/28.7
No. átomos:						
Proteína					2402	2774
DNA					486	609
Agua					32	-
Ligandos					13 (citrato)	-
Factores B ¹ :						
Proteína					66.1	88.3
Ligando/ión					99.1	85.1
Agua+ligandos					61.6	-
Desviación cuadrática media:						
Distancias enlace (Å)					0.01	0.005
Ángulos enlace (°)					1.342	0.902

$$R = \frac{\sum \|F_o\| - k\|F_c\|}{\sum F_o}$$

Fig 2.19_ **Tabla de cristalografía.** Procesamiento y estadísticas de refinado. Para la resolución de la estructura del complejo mTERF-ΔN-DNA, la asignación de fases se realizó mediante un experimento de dispersión anómala (MAD) de los cristales mTERF-ΔN/DNA29bp (SeMet). La calidad de los mapas de densidad electrónica mejoró notablemente con la recogida de datos de mayor resolución (2.3Å) y de elevada intensidad (33.7) y redundancia (31.0) de cristales de mTERF-ΔN/DNA15pb. Los cristales de mTERF-ΔN/DNA12pb nos permitieron trazar el DNA. Se utilizaron, por tanto, éstos últimos para el refinado de la estructura. Para la resolución y refinado final de la estructura del complejo mTERF-DNA se utilizaron los datos recogidos de los cristales mTERF-CHis/DNA15pb.

Los parámetros del conjunto de datos indican la calidad de éstos. En paréntesis se indica el valor del parámetro para la última capa de resolución.

5_Estructura tridimensional

5.1 Estructura del complejo mTERF y DNA

La estructura cristalográfica de mTERF consiste en nueve repeticiones estructurales, de TERF-I a IX, que equivalen a las repeticiones de secuencia mencionadas anteriormente como 'mTERF motif' que son generales para toda la familia MTERF (Roberti *et al.*, 2006a), pero con el inicio y final de los motivos definido en distintos puntos. Estos motivos repetitivos están flanqueados por un segmento N-terminal (Arg56-Glu73) y una hélice α C-terminal, la hélice X (Lys385-Ala399) (Fig 2.21a).

Cada motivo TERF está formado por unos 35 aminoácidos que se estructuran en tres hélices α (H1, H2 y H3) formando una superhélice trigonal levógira con un núcleo hidrofóbico central (Fig 2.20a). Esta superhélice se propaga en tándem con una rotación dextrógira entre motivos de 7 a 20° construyendo un solenoide (Fig 2.20b), que a su vez muestra una torsión hacia la derecha dando lugar a una cara convexa y otra cóncava (Fig 2.20c). H1 es la hélice más larga y forma un ángulo de 60° con la hélice de tamaño medio, H2. Las hélices H1 y H2 forman la superficie convexa del solenoide, mientras que H3 configura la cavidad cóncava interna.

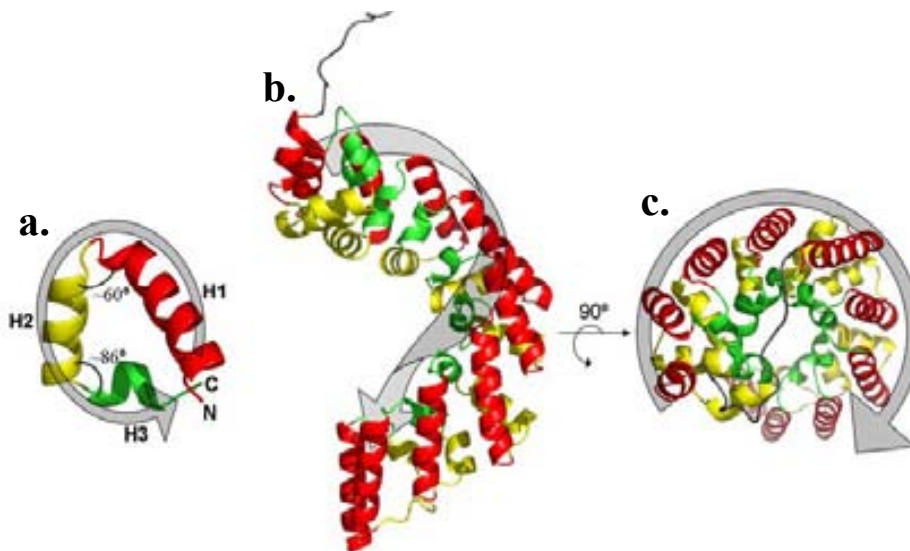


Fig 2.20_ (a) Esquema del motivo TERF. En rojo se muestra la hélice H1, en amarillo, H2 y en verde, H3. La flecha indica el carácter levógira del motivo TERF. (b) Los motivos TERF se propagan en tándem con una leve torsión dextrógira. Véase como varía la disposición de las hélices (H1, H2 y H3) en los diferentes motivos a causa de la rotación de cada motivo TERF con respecto al anterior. (c) El solenoide resultante, formado por los 9 motivos TERF, presenta una curvatura dextrógira que permite diferenciar una cara convexa formada por las hélices H1 en rojo y una cóncava formada por las hélices H3 en verde.

En la introducción, se describen los dominios constituidos por repeticiones en tándem de estructuras α como el Armadillo, HEAT y Pumilio, que han sido previamente caracterizados estructuralmente y que consisten respectivamente en repeticiones ARM, HEAT y PUM con carácter dextrógiro. La repetición TERF es estructuralmente parecida a ARM ya que ambas están compuestas por tres hélices siendo, en ARM, la hélice H1 la menor y H3 la mayor, de manera que ésta última es la que forma parte de la superficie cóncava de la estructura. No obstante, las hélices α tanto de ARM como de HEAT y PUM forman una superhélice dextrógira, mientras que en TERF la relación entre las hélices es levógira. La Fig 2.21b compara ambos motivos explicando detalladamente el carácter levógiro de TERF y dextrógiro de ARM.

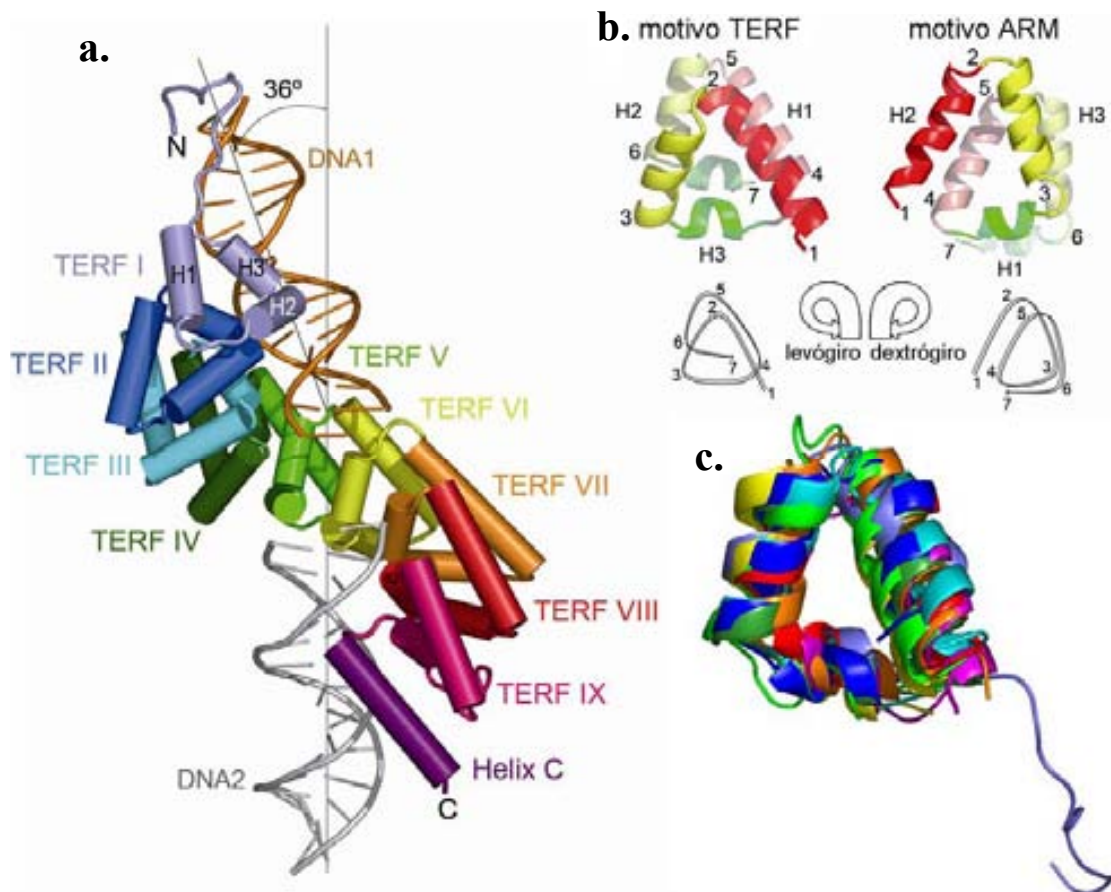


Fig 2.21_ (a) La estructura de mTERF consiste en nueve repeticiones del motivo TERF, que consta de tres hélices H1, H2 y H3 que se indican en el motivo TERF-I. En el cristal, una molécula de proteína une dos DNAs simétricos, DNA1 (naranja) y DNA2 (gris) relacionados por un ángulo de $\sim 36^\circ$. N y C indican los extremos N- y C-terminal de la proteína. (b) Explicación de la superhélice levógira y dextrógira que forman las tres hélices α (H1, H2 y H3) del motivo TERF y ARM respectivamente. Empezando por el punto 1 y siguiendo los números, la superhélice TERF se curva hacia la izquierda mientras el *armadillo repeat* (ARM) lo hace hacia la derecha. (c) Resultado de superposición estructural de los 9 motivos TERF con respecto a TERF-IV -manteniendo los colores usados en (a)-.

El motivo estructural de superhélice levógira se encuentra únicamente en pequeños dominios de cuatro hélices presentes en ATPasas AAA+ como la proteasa FtsH (PDB 2dhr), en la DNA polimerasa III (PDB 1xxh), o en el cargador/inhibidor de helicasa G39P del fago SPP1 de *B. Subtilis*. Así, la combinación de la conectividad en superhélice levógira con la propagación en tándem del motivo para formar un solenoide es única de las proteínas de la familia MTERF.

La desviación media cuadrática (*root mean square deviation*, RMSD) resultante de la superposición estructural de los nueve motivos TERF (Fig 2.21c) presenta valores entre 1.1 y 2.6 Å e indica que TERF-I es el motivo más distinto y que los motivos TERF-II, IV y V por una parte, y los motivos VI y VII por otra, son los que más se parecen entre ellos.

En el alineamiento estructural de los nueve motivos de mTERF se advirtió una pauta en su secuencia (Fig 2.22a). Para el análisis de estos rasgos repetidos se generó un logo a partir del alineamiento estructural de los nueve motivos de mTERF, con los correspondientes motivos de los ortólogos de orangután, rata, ratón, pez globo moteado, erizo de mar y mosca, mediante el servidor <http://weblogo.berkeley.edu/logo.cgi> (Fig 2.22b). Para identificar los motivos TERF en los ortólogos se realizó un análisis informático para la predicción de la estructura secundaria mediante el programa PsiPred (Jones, 1999), ya que en el caso de mTERF éste identifica bastante bien las hélices H1, H2 y H3 (Fig 2.1). En el logo, se puede distinguir el grado de conservación de secuencia para cada posición concreta del motivo (Fig 2.22b), de un total de 53 posiciones. En la Fig 2.22c se muestra la representación sobre un motivo TERF de los residuos conservados.

Así pues, el alineamiento de los motivos TERF muestra que: la hélice H1 de los motivos II-V y VIII presenta 4 residuos (14-17) que siguen el patrón BOBB (donde B es hidrofóbico y O, hidrofílico); la posición 16 de un motivo TERF contacta con la 14 del siguiente motivo; la posición 15 queda expuesta al solvente y la 17 está enterrada en un *core* hidrofóbico en el que contribuyen las posiciones 34, 37 y 38 de la hélice H2 y las posiciones 46 y 48 de la hélice H3 (Fig 2.22c). Este núcleo hidrofóbico se propaga a lo largo del interior de toda la proteína.

Los residuos polares conservados en las posiciones 23, 28 y 51 (Fig 2.22a) están involucrados en el *capping* N o C-terminal de las hélices H1, H2 y H3 respectivamente (Fig 2.22c).

En todos los motivos, excepto en el primero, el residuo de la posición 26 entre las hélices H1 y H2, que mayoritariamente es una glicina, muestra ángulos de conformación de la cadena principal que corresponden a una hélice levógira. Además, la prolina de la posición 44 al principio de la hélice H3 curva aproximadamente 90° la cadena principal. Así pues, la forma triangular de los motivos TERF seguramente se debe a la prolina y la glicina, ya que son residuos que tienden a romper la regularidad de los elementos de estructura secundaria.

Para intentar encontrar una explicación al carácter levógiro del motivo TERF, comparamos este patrón con el motivo ARM de varias proteínas. Podemos ver que en este último la glicina que presenta conformación levógira (G8) y la prolina (P11) se encuentran entre las hélices H1 y H2, justo antes de H2 (Fig 2.22d). En ambos motivos TERF y ARM, la prolina curva la siguiente hélice hacia el centro del motivo, pero la disposición diferente de la glicina (G8 para ARM y G26 para TERF, Fig 2.22c y d) no parece determinar la dirección de las hélices vecinas, sobretodo porque la longitud del *loop* entre H1 y H2 (en el caso de TERF) y H2 y H3 (ARM) puede ser variable y relativamente flexible. Además, en algunos motivos de otras proteínas ARM repeat, en el *loop* entre la hélice H2 y H3 podemos encontrar una glicina u otro aminoácido con ángulos de conformación levógiros a pesar que la superhélice del motivo es dextrógira, con lo cual este rasgo no parece ser la clave para la conformación levógira del motivo TERF. Más bien parece que el carácter levógiro del motivo TERF podría estar determinado por la orientación de los residuos hidrofóbicos (que forman el *core*) respecto de los ejes de las hélices- α , en combinación con la libertad de rotación de los vértices del triángulo que forman las tres hélices en el motivo.

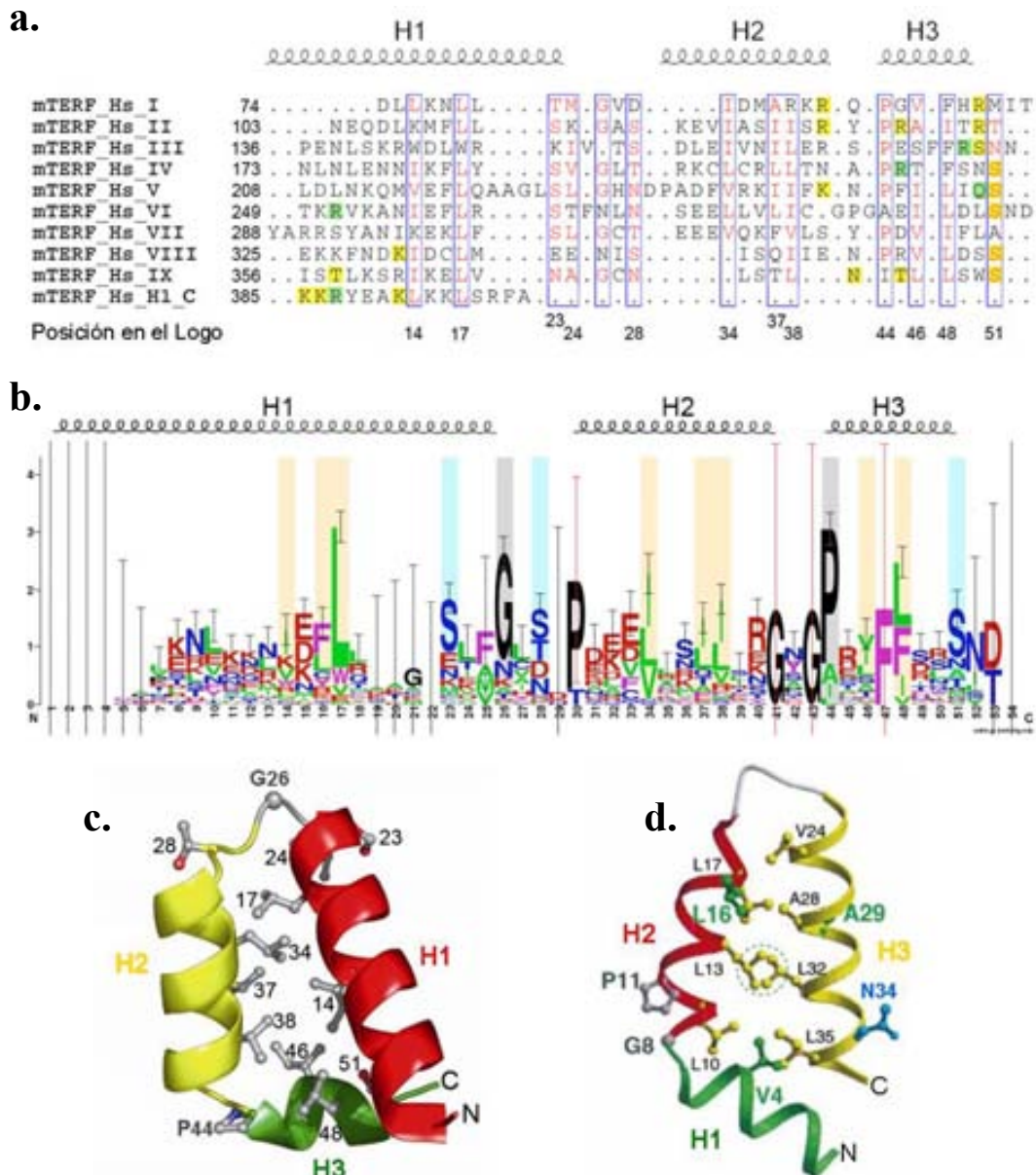


Fig 2.22_ (a) Alineamiento secuencial de los motivos TERF en base a la estructura de mTERF. En la parte superior se indican las hélices H1, H2 y H3 que forman el motivo TERF y en la parte inferior se indica el número que corresponde a la posición en el logo (b). En amarillo se muestran los aminoácidos que establecen interacciones con el esqueleto de fosfatos del DNA, y en verde los aminoácidos que interaccionan con las bases. (b) Logo de TERF generado usando el alineamiento de secuencia de los motivos TERF y añadiendo 53 motivos más identificados mediante la predicción de estructura secundaria del programa PsiPred de *P. pygmaeus*, CAH91462; *R. norvegicus*, Q9EP18; *M. musculus*, Q8CHZ9, *T. nigroviridis*, CAG10806; *P. lividus*, CAB40796 y *D. melanogaster* AAF53386 (Uniprot o GenBank). En el eje X se muestra la posición relativa de cada residuo en el alineamiento, y en el eje Y, el Rseq (bit), la cantidad de información presente en cada posición en la secuencia. Para cada posición, el logo nos da información del orden de predominancia de los residuos y de la frecuencia relativa de cada residuo. La barra de error indica el error en la frecuencia de conservación. Las barras de error rojas corresponden a inserciones puntuales en 5 o 6 secuencias de de los 54 motivos estudiados, e indican que el tamaño del símbolo no corresponde a su grado de conservación. Los rectángulos de color indican las posiciones conservadas en el logo. En azul se muestran los residuos de *capping*; en amarillo, los residuos hidrofóbicos del *core* y en gris, los residuos que suelen perturbar la estructura. Estos aminoácidos llevan asociadas barras de error pequeñas. (c) Representación sobre el motivo TERF-VII de las posiciones destacadas del logo ocupadas por los residuos más conservados. Se puede apreciar el *core* hidrofóbico del motivo TERF. (d) Representación de los residuos conservados en el motivo ARM de la importina- α (Andrade *et al.*, 2001).

Se realizó un análisis de la interfase entre los motivos TERF mediante el programa PISA (*Protein Interfaces Surfaces and Assemblies*) (Krissinel & Henrick, 2007). Δ^iG indica el aumento de energía libre de solvatación que supone la formación de la interfase (sin tener en cuenta los enlaces por puentes de hidrógeno y los puentes salinos), de manera que valores negativos corresponden generalmente a interfases hidrofóbicas. Para la interfase TERF-I y TERF-II se estima para Δ^iG un valor de -5 Kcal/mol que contrasta con las aproximadamente -10 Kcal/mol estimadas para el resto de las interfases entre los motivos TERF. Además, el parámetro CSS (*Complexation Significant Score*) que indica cuán significativa es la interfase para el ensamblaje, estima un valor de 0.1 para la interfase entre TERFI y II vs un valor de 1 para las interfases entre el resto de motivos. Esta diferencia indica que este motivo I queda más accesible al solvente y por tanto puede ser proteolizado más fácilmente como ocurrió en el caso de mTERF- Δ N.

Interacción con el DNA

En el cristal una misma proteína une dos moléculas de DNA, DNA1 y DNA2 (ó DNA1 simétrico), ya que ambos están relacionados por simetría y, a su vez, cada molécula de DNA está unida por dos proteínas relacionadas por simetría. Por tanto, la relación estequiométrica del cristal es una molécula de proteína mTERF por una molécula de DNA. En la Fig 2.21a se puede observar cómo mTERF contacta el DNA de 15 pb. En concreto, los motivos I-VI establecen contactos con una mitad de la molécula de DNA1, mientras los motivos VIII, IX y la hélice X lo hacen con la otra mitad de la molécula DNA1 simétrica o DNA2. Si observamos la representación de la carga electrostática sobre la superficie de la proteína mTERF vemos que las zonas cargadas positivamente de los subdominios N y C-terminal siguen las hebras de DNA (Fig 2.23a). Las moléculas de DNA1 y 2 no interaccionan entre ellas y la posición relativa entre ambas viene dada por un ángulo de $\sim 36^\circ$ (Fig 2.21a), muy parecida a la torsión estimada mediante ensayos previos de permutación circular en los que se determinó experimentalmente que la unión de mTERF a DNA_{TER} daba lugar a una torsión de éste de 35° (Shang & Clayton, 1994). Al único motivo TERF que no contacta el DNA, TERF-VII, le falta el residuo polar en la posición 50/51 del logo que, en los otros motivos, está involucrado en el contacto con el DNA. Este hecho, junto con que el patrón de interacción de H1 con las hélices de los motivos vecinos es diferente, le confiere un carácter excepcional.

En la estructura cristalográfica de mTERF con el oligonucleótido de 15 pb, hay 5 residuos implicados en interacciones con las bases nitrogenadas del DNA y 19 residuos involucrados en interacciones no específicas con los fosfatos del DNA (Fig 2.22a), ocho de los cuales además desempeñan el papel de *C-capping* de las hélices α en la interfase proteína-DNA, disminuyendo la repulsión entre los fosfatos del DNA y la polaridad negativa de la hélice debido a los carbonilos del enlace peptídico.

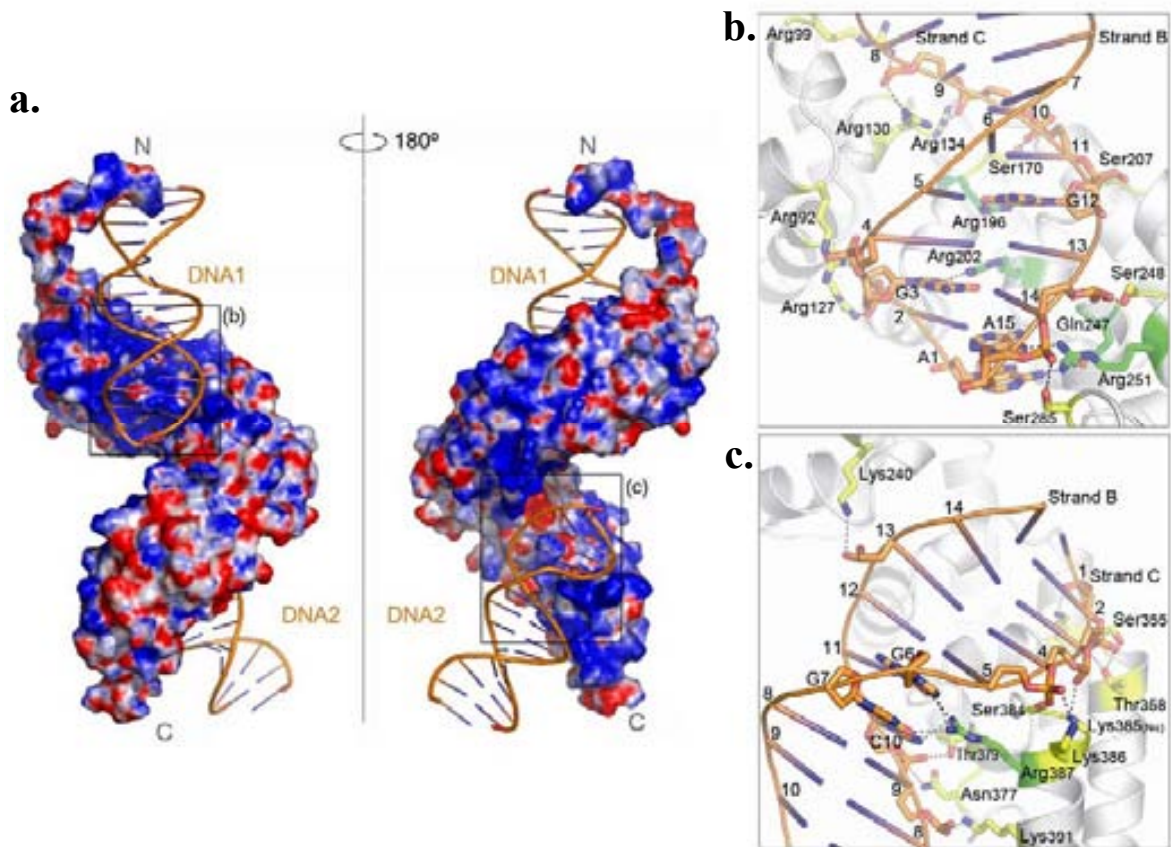


Fig 2.23 (a) Representación del potencial electrostático sobre la superficie de Connolly de mTERF. La superficie se muestra en dos orientaciones relacionadas por un giro de 180° . En el cristal, una molécula de proteína une, mediante dos áreas diferentes, los oligonucleótidos DNA1 y DNA2 que corresponden a un mismo oligonucleótido relacionado simétricamente. La representación de la carga electrostática calculada mediante APBS (Baker *et al.*, 2001) sobre la superficie de la proteína muestra la carga negativa en rojo y la carga positiva en azul ($\pm 10 k_b T/e_c$). Las zonas de carga electrostática positiva que interactúan con el DNA se detallan en (b) y (c). (b) En amarillo se representan los residuos del subdominio N-terminal que interactúan con los fosfatos del DNA y en verde, los residuos que interactúan con las bases nitrogenadas del DNA. La numeración del DNA corresponde a la usada en el PDB (Fig 2.14). (c) Contactos entre los residuos de la parte C-terminal y el DNA, con el código de colores y numeración igual que en (b).

Los motivos de la parte N-terminal, TERF I-VI, contribuyen con los residuos de las posiciones 39, 45/49, 50/51 del logo a la interacción inespecífica con los fosfatos del esqueleto del DNA, posicionando los residuos Arg169 y Arg202 para que contacten con las guaninas 12 (cadena C) y 3 (cadena B) respectivamente, y los residuos Gln247 y

Arg251, con las adeninas 1 (cadena B) y 15 (cadena C), respectivamente. Además, los residuos Gln247 y Arg251 contribuyen a la estabilización de uno de los extremos del oligonucleótido DNA1 (Fig 2.23b). En la parte C-terminal, las posiciones conservadas de los motivos TERF VIII y IX (posiciones 9 y 13 de H1 y 45 y 51 de H3 en el logo) interaccionan con los fosfatos del esqueleto del DNA2 y rodean a la Arg387 de la hélice X, que interacciona con el DNA mediante las guaninas 6 y 7 de la cadena C y la citosina 10 de la cadena B (Fig 2.23c).

En general, los residuos de mTERF involucrados en los contactos con el DNA están conservados en los miembros de la subfamilia de MTERF1 en vertebrados, sugiriendo la formación de complejos similares.

5.2 Estructura del complejo mTERF- Δ N y DNA

A mTERF- Δ N le faltan los 40 primeros aminoácidos, que corresponden al segmento N-terminal (Arg55-Glu73) y al primer motivo TERF-I. La estructura de mTERF- Δ N presenta el mismo patrón de repeticiones que mTERF entera e, incluso, la posición relativa de las tres hélices de un mismo motivo permanece inalterada.

Interacción con el DNA

La resolución de la estructura fue posible gracias a la cristalización de mTERF- Δ N en complejo con el DNA12pb (Fig 2.14). En el cristal, la proteína une dos moléculas de DNA simétricas que interaccionan mediante *stacking* o apilamiento de bases dando lugar a una fibra pseudo-continua de DNA a lo largo del eje z (Fig 2.24a). En esta estructura, el subdominio N-terminal de mTERF únicamente presenta algunas interacciones inespecíficas con los fosfatos, mientras el C-terminal, además de las interacciones con los fosfatos, mantiene la interacción de Arg387 con las bases nitrogenadas, en este caso con la tirosina 7 (cadena C) y la guanina 5 (cadena B) según la nomenclatura del pdb (Fig 2.24b).

En los cristales con forma de placa obtenidos con mTERF- Δ N y DNA 15 pb (grupo espacial C222₁, Fig 2.15) dos dominios C-terminales de dos moléculas simétricas - relacionadas por un eje binario que pasa a través del DNA-, contactan una misma molécula de DNA que en el cristal se dispone en ambos sentidos (Fig 2.25).

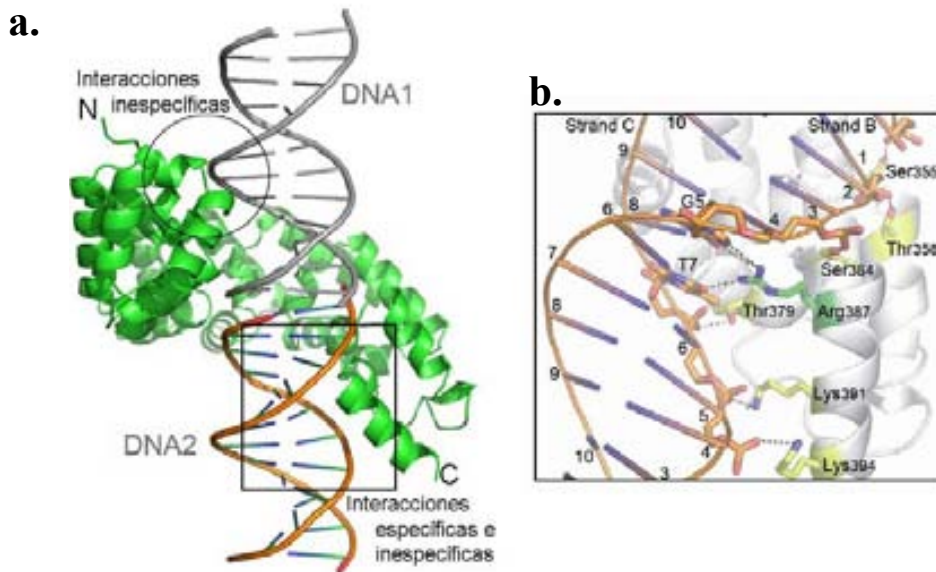


Fig 2.24_ (a) Estructura del complejo de mTERF-ΔN con el oligonucleótido de 12 pb. En el cristal, una molécula de proteína une dos moléculas de DNA simétricas, DNA1 (gris) y DNA2 (naranja), que interactúan mediante *stacking* entre las bases dando lugar a una fibra pseudo-continua de DNA en el eje z del cristal. El recuadro y el círculo muestran las zonas de contacto con el DNA. En la parte N-terminal se dan interacciones con los fosfatos (Fig 2.26b) y en la parte C-terminal, con los fosfatos y las bases. N y C indican los extremos N- y C-terminal de la proteína. (b) Ampliación del recuadro en (a) donde se muestran los contactos entre el subdominio C-terminal de mTERF-ΔN y el DNA2. En amarillo se muestran los residuos que interactúan con los fosfatos y en verde, las interacciones con las bases nitrogenadas. Arg387 contacta con la guanina 5 (cadenaB) y la timina 7 (cadenaC).



Fig 2.25_ Estructura de los cristales en forma de placa de mTERF-ΔN con un oligonucleótido de 15pb. El grupo espacial es C222₁. Dos moléculas simétricas de proteína unen mediante los dominios C-terminal una única molécula de DNA. El eje binario que pasa a través del DNA nos indica que se encuentra desordenado en el cristal pudiéndose disponer en ambos sentidos.

Comparación de las estructuras mTERF y mTERF-ΔN

La superposición de las estructuras de la proteína entera y la troncada indica que los subdominios N y C-terminal están desplazados el uno respecto al otro. Si se superponen las dos estructuras por la parte N-terminal (motivos II, III y IV), queda desenchajada la parte C-terminal. Y viceversa, si superponemos las dos estructuras por la parte C-terminal (motivos VII, VIII, IX y hélice X), la parte N-terminal de mTERF-ΔN (TERFIL-IV) presenta un desplazamiento (hasta 13 Å en el C α N103) y un reordenamiento, de manera

los residuos de la parte N-terminal de la superficie cóncava de mTERF, en la forma corta se abren hacia el exterior respecto al eje central de la proteína (Fig 2.26c). Este reordenamiento de la parte N-terminal con respecto a la proteína entera mTERF, provoca la desaparición de las interacciones con las bases de la molécula DNA1 de los residuos Arg251 y Gln247 y la alteración de la interacción de los residuos Arg169 y Arg202, que pasan de interactuar con las bases a interactuar con los fosfatos (Fig 2.26a y b).

En referencia al DNA, observamos que en la superposición de las dos estructuras por la parte C-terminal los esqueletos de la molécula DNA2 con el que interactúan coinciden, pero con la secuencia en sentido opuesto (Fig 2.26c). Esta inespecificidad de secuencia la observamos también en los cristales C222₁ (3.2Å), cuya estructura fue resuelta pero no refinada, en que la proteína une el DNA en ambos sentidos. Aun así, en todos los casos coincide el emplazamiento del esqueleto de la molécula DNA2 porque se mantiene la interacción entre las mismas regiones de la proteína y el surco mayor del DNA.

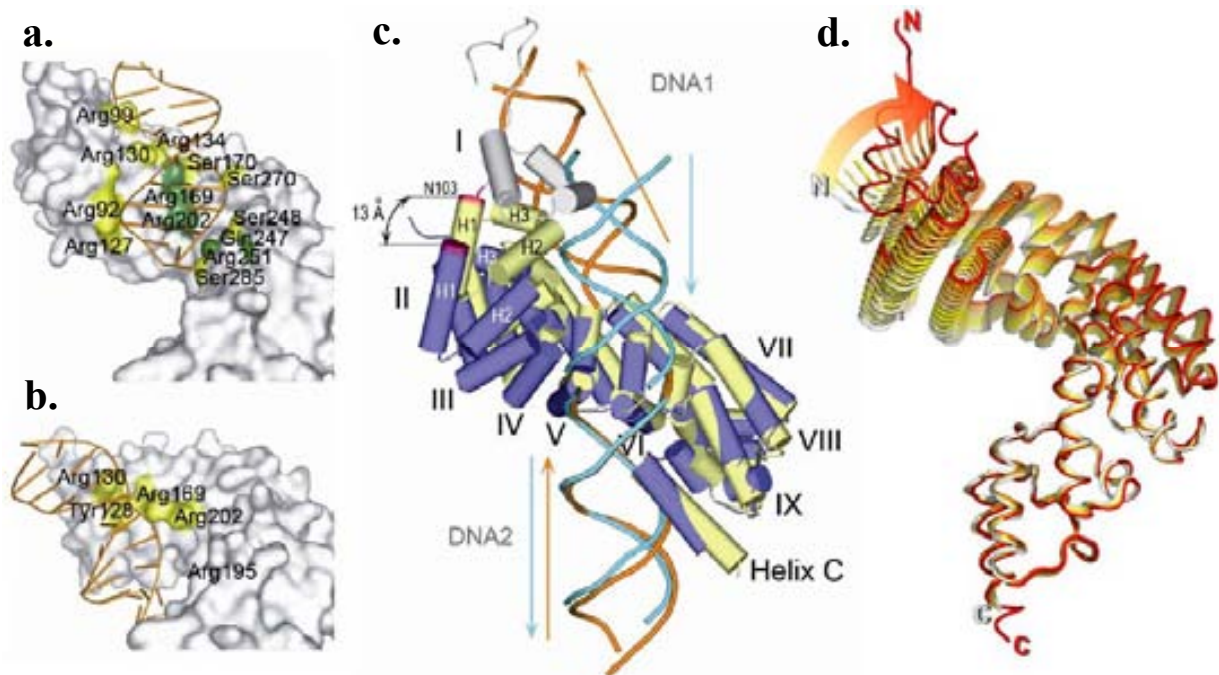


Fig 2.26_ (a) Superficie de contacto del subdominio N-terminal de mTERF con el DNA1. (b) Superficie del mismo subdominio en mTERF- Δ N en contacto con el DNA1. Ambas estructuras están en la misma orientación. En amarillo se muestran las interacciones con los fosfatos y en verde las interacciones con las bases nitrogenadas. (c) Superposición del subdominio C-terminal de mTERF-DNA (en amarillo y naranja respectivamente) y mTERF- Δ N-DNA (en azul y cian respectivamente). La escisión de los primeros 40 aminoácidos que incluyen el segmento N-terminal y TERF-I (en gris), resulta en una reorganización de la parte N-terminal y una disposición diferente del DNA. Las hélices H1-H3 del motivo TERF-II están señalizadas en ambas estructuras y se muestra el desplazamiento máximo en Asn103 (magenta). Las flechas naranja y cian muestran la dirección y sentido del DNA en cada una de las estructuras. (d) Representación del movimiento de la parte N-terminal mediante los pDBs intermedios generados mediante MovieMaker (Maiti *et al.*, 2005) entre mTERF- Δ N (blanco) y mTERF (en rojo).

6_ Estudio en solución del complejo mTERF-DNA

En algunos estudios realizados anteriormente (mediante *thermo*fluor y ensayos de unión a DNA mediante geles de retardo) ha quedado demostrado que el complejo de mTERF con el DNA es más estable cuando la longitud del oligonucleótido es mayor, pero no fue posible obtener cristales de mTERF entera con oligonucleótidos mayores de 15 pb. En ambas estructuras cristalográficas de mTERF y mTERF-ΔN, una molécula de proteína se une a dos oligonucleótidos idénticos. Y este hecho no acababa de explicar la interacción descrita de la proteína con el DNA mitocondrial continuo.

Por tanto, se decidió estudiar los complejos de mTERF y mTERF-ΔN con un oligonucleótido de 28 pb -que incluye toda la secuencia de *footprint* del DNA (Fig 2.14)- en solución, mediante la técnica de SAXS.

Los datos de SAXS que se muestran a continuación fueron tomados en el sincrotrón ESRF de Grenoble (*beamline* ID14.3). Se midieron muestras de proteína sola y en complejo con DNA a diferentes concentraciones (Fig 2.27). Todas las medidas se tomaron a 10° C. Para cada medición se tomaron 2 blancos, antes y después de la medida de la muestra, el promedio de los cuales se debe sustraer a la curva de la muestra. Una vez obtenidas las curvas corregidas, los valores de la intensidad de dispersión a ángulo 0 ($I(0)$) y el radio de giro (R_g) fueron calculados mediante la aproximación de Guinier. El peso molecular de las muestras se estimó a partir de la dispersión de éstas en referencia a la dispersión de una solución estándar de BSA medida en las mismas condiciones, mediante la fórmula:

$$Pm_{muestra} = \frac{I(0)_{muestra}}{Conc_{muestra}} \cdot \frac{Pm_{BSA}}{\frac{I(0)_{BSA}}{Conc_{BSA}}};$$

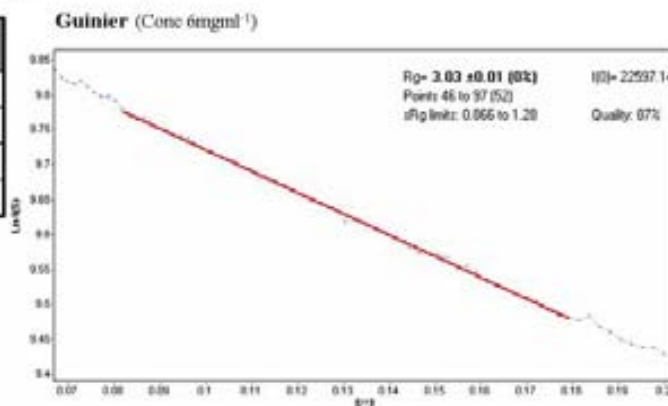
donde para la BSA, $I_0/Conc_{real}=20803,24$.

Como se puede observar en la tabla (Fig 2.27), la masa molecular aparente de la muestra aumenta con la concentración, lo que significa que existen fenómenos de oligomerización. Así pues, como nos interesaba el estudio del complejo proteína-DNA en solución (1:1 de proteína:DNA), únicamente tomamos la curva de menor concentración,

6mg ml⁻¹ para el complejo mTERF-DNA_{TER} y 1.7 mg ml⁻¹ para el complejo mTERF-ΔN-DNA_{TER}, como referencia en los estudios que pasamos a describir.

mTERF-DNA_{TER} (pM teórico del complejo 58.66KDa)

Concentración (mg ml ⁻¹)	Rg (Å)	Masa molecular (KDa)
6	30,3	72,11
8,8	31,8	73,01
15	33,9	87,93
22,8	36,7	131,10



mTERF-ΔN-DNA_{TER} (pM teórico del complejo 52.87KDa)

Concentración (mg ml ⁻¹)	Rg (Å)	Masa molecular (KDa)
1,7	29,5	51,03
3	40,1	70,06
4,5	42,6	82,57
9,92	55,5	152,39

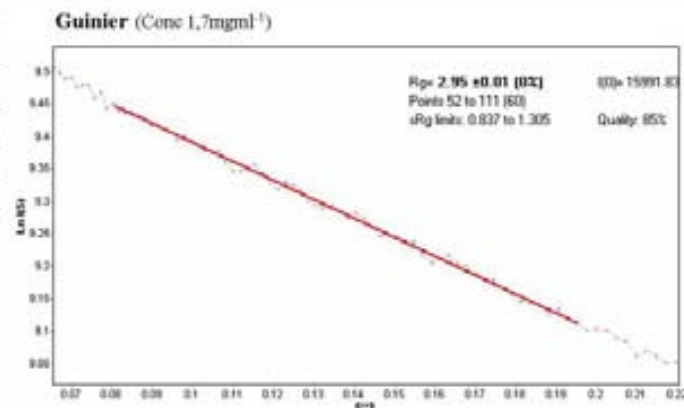


Fig 2.27_ (Izq) Tabla de resultados de las mediciones de SAXS realizadas en el sincrotrón ESRF de Grenoble para las muestras mTERF-DNA_{TER} (arriba) y mTERF-ΔN-DNA_{TER} (abajo). Podemos observar que el Rg y la masa molecular calculada aumentan con la concentración, lo que significa que existen ciertos fenómenos que afectan a la muestra, como el posible equilibrio con especies de mayor tamaño. (Dcha) Aproximación de Guinier para el cálculo de los Rg de las muestras mTERF-DNA_{TER} y mTERF-ΔN-DNA_{TER}. Para valores bajos de s, se encuentra que la representación de Guinier (log (I(s)) vs s²) muestra comportamientos lineales, cuya pendiente corresponde al radio de giro (Rg) de la partícula. Para la muestra mTERF-DNA_{TER} el valor de Rg es de 30.3 Å, y para la muestra mTERF-ΔN-DNA_{TER} el valor de Rg es 29.5 Å.

Para el análisis de los datos, se generaron modelos de mTERF-DNA_{TER} deslizando una molécula de DNA de 28 pb base a base, curvado en el caso de mTERF-NHis y recto en el caso de mTERF-ΔN, a lo largo del eje que forman las dos moléculas de DNA que une cada molécula de mTERF en el cristal, y se evaluó el ajuste entre la curva teórica para cada modelo y la curva medida experimentalmente.

6.1 Complejo mTERF y DNA_{TER} en solución

Para el complejo de mTERF con el oligonucleótido DNA_{TER} (28pb), si representamos gráficamente el error (χ_i) asociado a la curva de dispersión teórica de cada uno de los 30 modelos generados, calculado mediante el programa CRY SOL (Svergun, 1995), con respecto a la curva de dispersión experimental, obtenemos una gráfica con un mínimo (Fig 2.28b y c) que corresponde al modelo generado cuya curva teórica se ajusta mejor a los datos experimentales. Así, los modelos 10 y 11 son los que mejor se asemejarían a la especie que mayoritariamente existe en solución.

Esta representación del valor de χ_i de los modelos nos indica que parece existir una especificidad dentro de la secuencia de DNA_{TER} 28pb, ya que en el ajuste a la curva experimental existen algunos modelos (del 9 al 12 en la Fig 2.28c) que encajan mejor que otros (con un χ_i menor). Es decir, que la unión de mTERF al DNA no es aleatoria, sino que se da preferentemente sobre una región de esta molécula de DNA. Aun así, debido a la baja resolución de la técnica, no es posible delimitar exactamente la zona de unión.

En el modelo 10, que muestra un valor de χ_i prácticamente igual al modelo 11, las interacciones de los aminoácidos de la parte N-terminal de mTERF con el DNA coinciden con las interacciones del modelo cristalográfico. No obstante, en este modelo, el extremo final de la molécula de DNA queda, en la parte C-terminal de la proteína, justo por encima de la Arg387, de modo que ésta no llegaría a interactuar con el DNA. Debemos tener en cuenta que, al modelar una sola molécula de DNA curvada siguiendo los ejes de las dos moléculas de DNA que une una molécula de proteína en la estructura cristalográfica, nos encontramos, además de impedimentos estéricos en el centro de la proteína, que las cadenas de DNA que quedan conectadas en el modelo, tienen direcciones 5'–3' opuestas en la estructura cristalográfica (Fig 2.28a). De manera que si uniéramos las cadenas de las dos moléculas de DNA que tienen el mismo sentido, éste debería sufrir una distorsión desenroscándose en el centro, lo que quizá permitiría una longitud final mayor y, por tanto, una posible interacción de la Arg387 con el DNA.

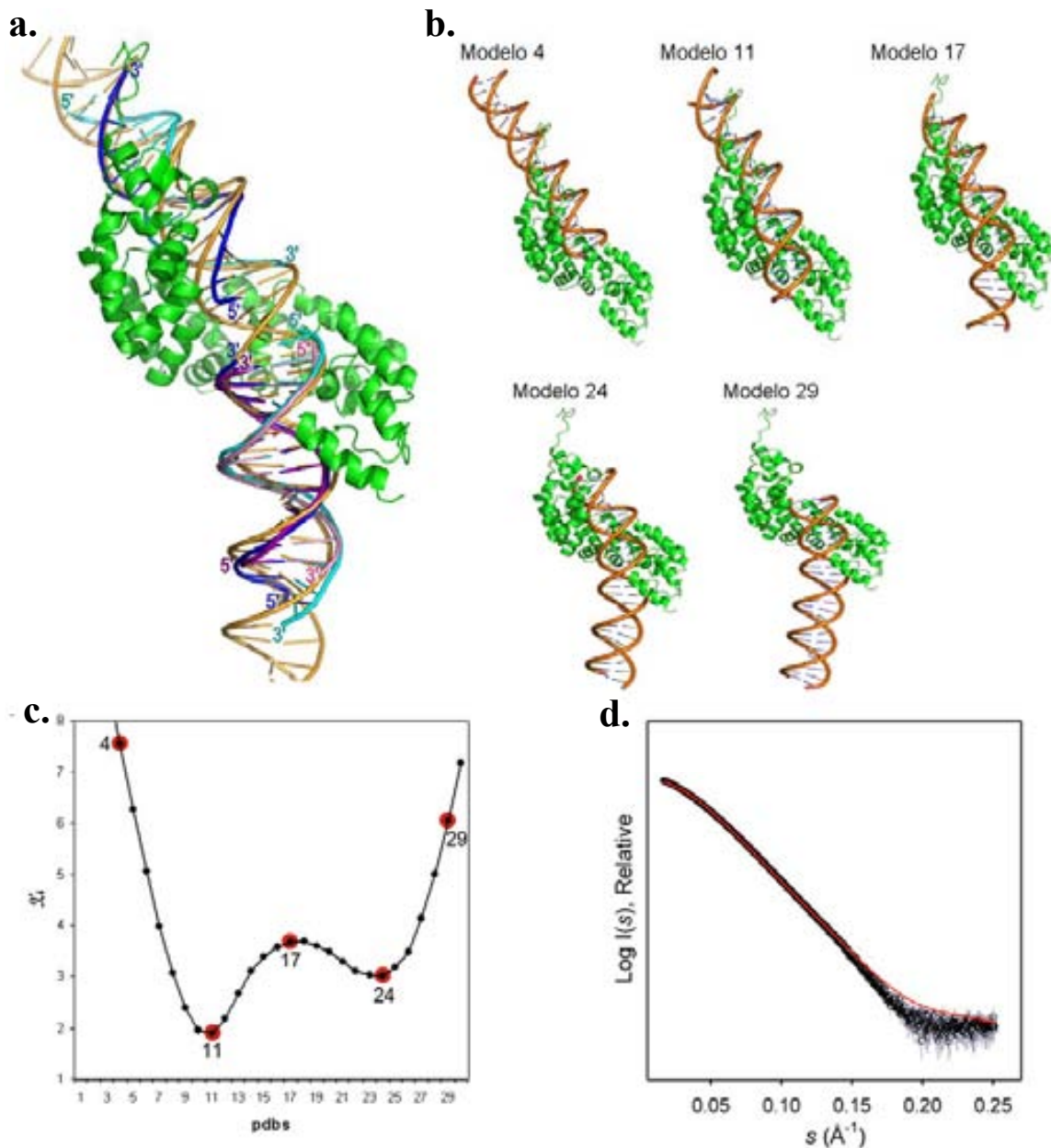


Fig 2.28_ Estudio estructural mediante SAXS del complejo en solución mTERF-DNA28pb. (a) En naranja se muestra el DNA continuo (28 pb) modelado a lo largo del eje que forman los dos DNAs (15 pb) que une cada molécula de mTERF en el cristal (azul marino y cian para cada cadena). En rosa y púrpura se muestran las cadenas del DNA (12 pb) cristalizado con la forma mTERF- Δ N. En ambos casos, coinciden los extremos 5' y 3' del DNA. Al modelar una molécula de DNA continuo añadiendo únicamente un ángulo de torsión, quedan "unidas" dos cadenas que van en sentido contrario. (b) Se generaron 30 modelos del complejo a partir de la estructura cristalográfica, de los que se muestran cinco modelos representativos. (c) Acuerdo (χ^2) entre cada una de las curvas teóricas de los modelos generados y la curva experimental medida por SAXS. El error asociado a los modelos que se muestran en (a) está indicado en rojo. El modelo 10 y 11 son prácticamente equivalentes, y son los que encajan mejor con la curva experimental de SAXS. En el modelo 10, además, la posición del DNA encaja con el DNA1 de la estructura cristalográfica. (d) Intensidad de dispersión (círculos negros) en escala logarítmica en función del momentum transfer $s=4\pi \sin(\theta)/\lambda$, donde λ es la longitud de onda de los rayos X (1.5 \AA) de la muestra mTERF-DNA_{TER} (6 mg ml⁻¹). La línea continua roja representa el encaje de CRYSOLE del modelo 11 con la curva de dispersión.

6.2 Complejo mTERF- Δ N y DNA_{TER} en solución

Para este estudio se generaron 50 modelos del complejo de mTERF con el oligonucleótido DNA_{TER} de 28pb, en los que el DNA se desliza base a base a través de la proteína siguiendo el eje del DNA de la estructura cristalográfica. En la representación del error asociado (χ_i) a cada uno de los modelos con respecto a la curva de dispersión experimental, se aprecian dos mínimos indistinguibles, el modelo 18 y el 34 (Fig 2.29) debido a la ‘simetría’ intrínseca de la estructura cristalográfica de la proteína mTERF- Δ N (Fig 2.30). Teniendo en cuenta la representación de χ_i de los modelos parece que esta forma proteolizada de mTERF también presentaría especificidad para una zona de la molécula de DNA_{TER}. Así, la unión de mTERF- Δ N a DNA_{TER} no es aleatoria sino que se da preferentemente sobre una franja determinada del DNA, ya que en los modelos que mejor se ajustan a la curva experimental (valores menores de χ_i), mTERF- Δ N se une a una zona próxima al extremo del DNA.

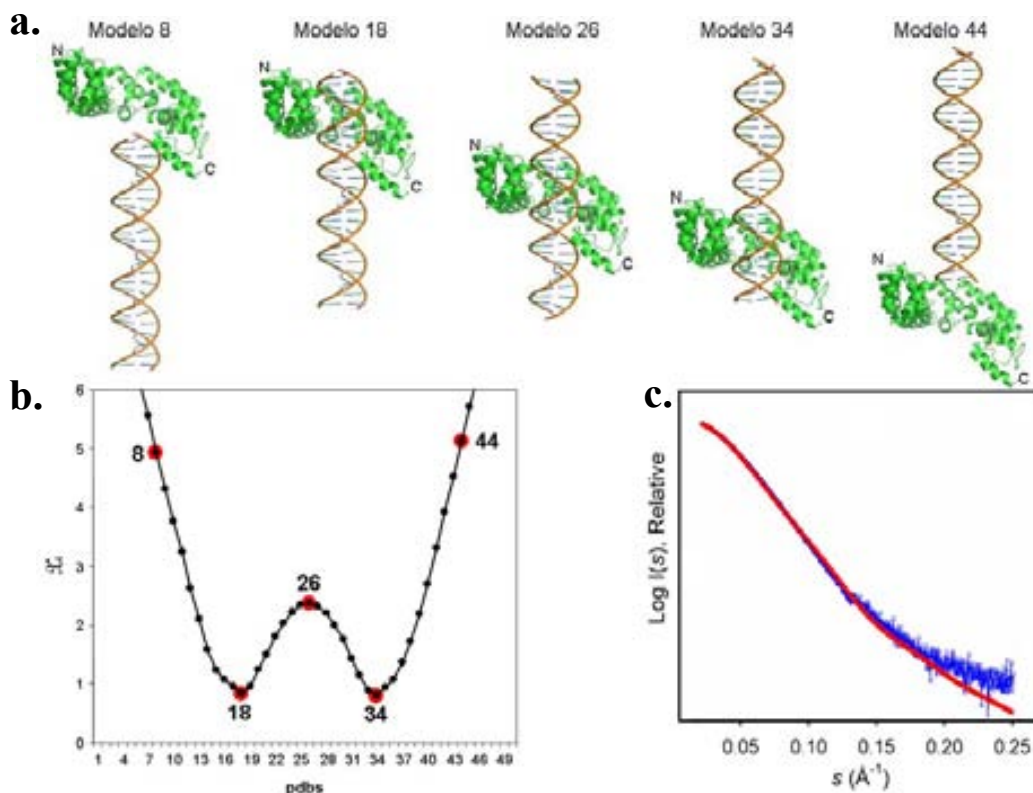


Fig 2.29_ Estudio estructural mediante SAXS del complejo en solución mTERF- Δ N-DNA28pb. (a) Se muestran cinco modelos representativos de los 50 modelos generados. (b) Acuerdo (χ_i) entre cada uno de los modelos generados y la curva experimental medida por SAXS. El error asociado a los modelos que se muestran en (a) está indicado en rojo. Los modelos 18 y 34 son los que encajan mejor con la curva experimental de SAXS (Fig 2.30). (c) Intensidad de dispersión (azul) en escala logarítmica en función del momentum transfer de la muestra mTERF- Δ N-DNA_{TER} (1.7mg ml⁻¹). La línea continua roja representa el encaje de CRYSOLE del modelo 18 con la curva de dispersión.

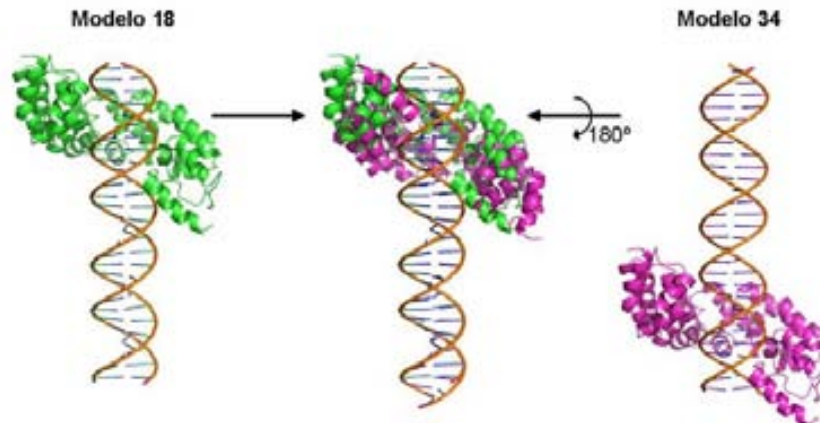


Fig 2.30_ La superposición de los modelos 18 (verde) y 34 (rosa) muestra que, debido a la 'simetría' intrínseca de la proteína, estos dos modelos son muy parecidos, lo que los hace indistinguibles para este análisis del ajuste entre la curva experimental y los modelos mediante CRYSOLOG.

7_ Ensayos de unión y especificidad de mTERF

En ambas estructuras cristalográficas de mTERF y mTERF- Δ N, una proteína une dos oligonucleótidos relacionados por simetría mediante los subdominios N y C-terminal, ambos cargados positivamente. Para analizar la unión a DNA_{TER} y la especificidad de mTERF, debido a la cantidad de interacciones inespecíficas que presentan ambos subdominios en la unión al DNA, se utilizaron como primera aproximación, además de mTERF, las formas truncadas de ésta, mTERF- Δ N, mTERF- Δ C y Δ N-mTERF- Δ C (Fig 2.31a) en vez de diseñar mutantes puntuales de los residuos implicados en la unión. Todos estos constructos son estables en solución y el perfil cromatográfico de la cromatografía de exclusión molecular corresponde a un monómero. MTERF- Δ C se clonó y expresó en las mismas condiciones que las dos anteriores, y la purificación se realizó en dos pasos, mediante una columna de afinidad y otra de exclusión molecular. Δ N-mTERF- Δ C se obtuvo a partir de la proteólisis limitada de mTERF- Δ N con papaína.

El ensayo de unión a DNA se realizó con un oligonucleótido que contenía la secuencia de *footprint*, la molécula de DNA_{TER} 28pb, marcado con un fluoróforo en el extremo 5' y en presencia de cantidades crecientes de *polydIdC* que es un competidor inespecífico. En este ensayo se determinó que la cantidad de *polydIdC* adecuada para los ensayos era de 500 ng. Además, se hizo patente que únicamente unen DNA las formas mTERF y mTERF- Δ N

(Fig 2.31b). Así, mTERF- Δ C, aun manteniendo el dominio N-terminal que en la estructura de mTERF presenta unión con los fosfatos y las bases nitrogenadas, en ausencia de la parte C-terminal, no es capaz de unir DNA.

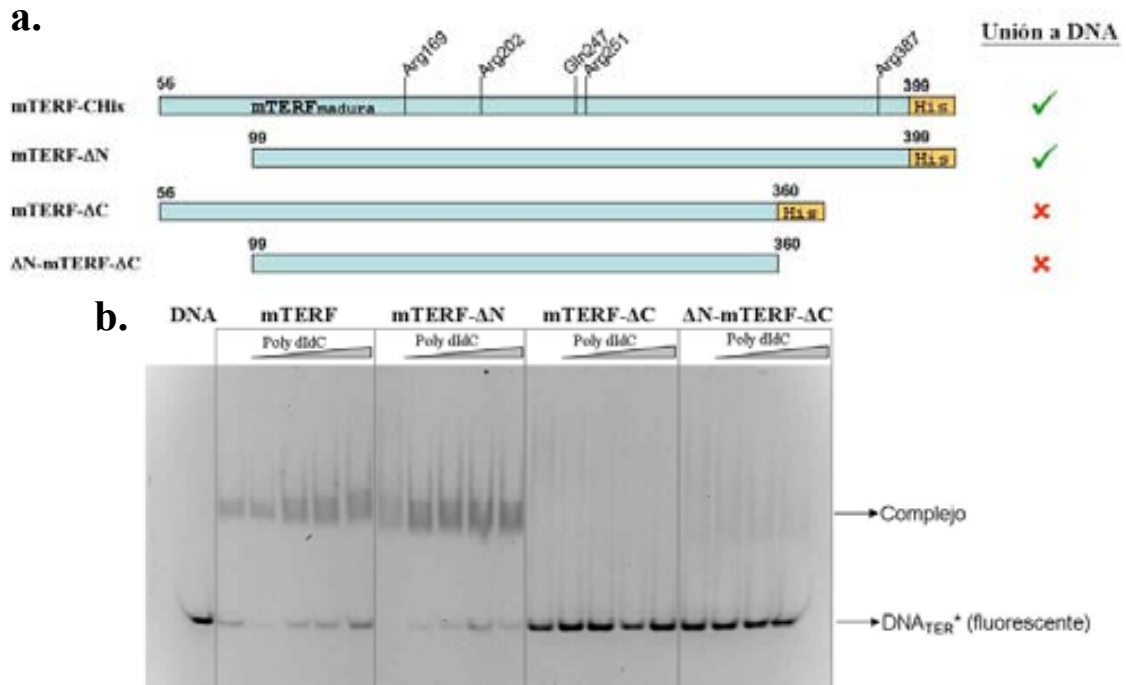


Fig 2.31 (a) Esquema de los constructos utilizados en el ensayo y de su capacidad de unión a DNA. Sobre mTERF-CHis se indican los aminoácidos que contactan con las bases nitrogenadas en la estructura cristalográfica de mTERF-DNA15pb. Recuérdese que en mTERF- Δ N la orientación de la parte N-terminal da lugar a un cambio en el patrón de los contactos de los aminoácidos Arg 169, Arg 202 y Arg251 con el DNA (Fig 2.26a y b) (b) Estudio mediante de la unión de cada uno de los constructos de mTERF con la molécula DNA_{TER} marcada con un fluoróforo. Las muestras se cargaron en el gel nativo en presencia de cantidades crecientes de *polydIdC*. La banda de retardo indica que mTERF y mTERF- Δ N se unen aDNA_{TER}.

Teniendo en cuenta que ni mTERF- Δ C ni Δ N-mTERF- Δ C se unían al DNA, se prosiguieron los ensayos de especificidad únicamente con mTERF y mTERF- Δ N. El ensayo de especificidad se realizó en presencia de DNA_{TER} 28pb marcado, *polydIdC* y cantidades crecientes de DNA_{TER} 28pb o de un oligonucleótido inespecífico (30 pb) sin marcar. En este caso esperamos que DNA_{TER} compita contra sí mismo por la unión a mTERF, mientras el DNA inespecífico sólo competiría si la proteína une cualquier DNA con prácticamente la misma afinidad. Se trataría entonces de una unión inespecífica. El resultado muestra que ambas proteínas se unen a DNA_{TER} específicamente, ya que concentraciones crecientes de DNA inespecífico no fueron capaces de deshacer la unión (Fig 2.32).

El hecho que mTERF- Δ N uniese específicamente DNA_{TER} nos sorprendió, teniendo en cuenta que la estructura cristalográfica del complejo de mTERF- Δ N con el

oligonucleótido DNA12pb muestra que el único aminoácido que tiene un contacto con las bases nitrogenadas es Arg387. Además, en las estructuras resueltas la interacción del subdominio C-terminal con el DNA parece ser inespecífica ya que el DNA se dispone en ambos sentidos y la interacción de la Arg387 se da con nucleótidos diferentes. Y, aunque el fragmento mTERF- Δ N contiene los aminoácidos Arg169, Arg202 y Arg251 (Fig 2.31a) que en la proteína entera contactan las bases del DNA, según la estructura cristalográfica de mTERF- Δ N éstos estarían dispuestos de manera que no interaccionan con las bases del DNA. No obstante, debemos tener en cuenta que en la estructura cristalográfica del complejo el DNA es de 12 pb, mientras que estos experimentos se realizaron con el oligonucleótido DNA_{TER} de 28pb. En este sentido, los resultados de los ensayos de SAXS con el oligonucleótido DNA_{TER}28pb sugieren que en solución, la especie mayoritaria tiene una disposición proteína-DNA muy parecida a la cristalográfica ya que en el modelo que mejor se ajusta a los datos experimentales la proteína no curva el DNA. Así pues, parece ser que la Arg387, la única que contacta con las bases del DNA, podría tener un papel importante en el reconocimiento específico de la molécula de DNA de 28pb.

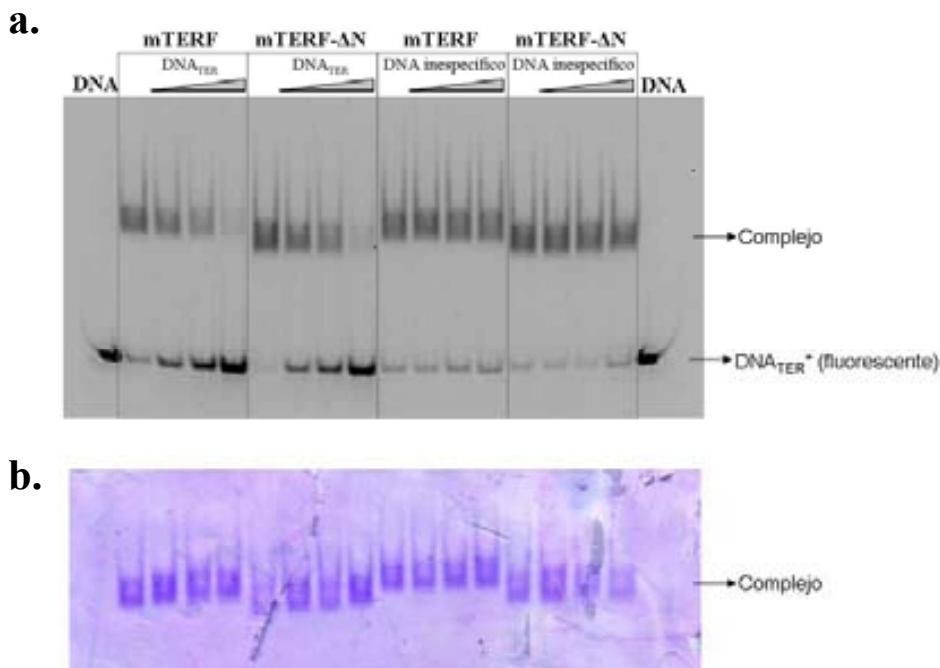


Fig 2.32_ **Ensayo de especificidad de mTERF** (a) Gel nativo en presencia de mTERF-CHis y mTERF- Δ N con DNA_{TER} marcado y cantidades crecientes de DNA_{TER} no marcado (carril 2-9) o DNA inespecífico no marcado. El DNA inespecífico no desplaza la unión de mTERF (pocillos 10-13) o mTERF- Δ N (pocillos 14-17) con el DNA_{TER}, lo que implica que la unión es específica. Como control, se realizó también el ensayo de competencia en presencia de DNA_{TER} no marcado. En este caso sí existe competición, y a medida que se aumenta la concentración de DNA_{TER} no marcado va disminuyendo la intensidad de la banda retardada que corresponde al complejo con DNA_{TER} marcado y aumentando la que corresponde al DNA_{TER}* libre (pocillos 2-9) (b) Tinción en azul de comassie del gel nativo. La tinción con Comassie de la banda de retardo indica que contiene proteína.

En estudios anteriores sobre la unión de mTERF a DNA, Daga y col. (Daga *et al.*, 1993) identificaron en la fracción purificada de mTERF tres polipéptidos, dos de peso molecular aparente de 34 KDa y uno, de 31 KDa que unían específicamente DNA, presentando el de 31 KDa una menor afinidad por la secuencia de reconocimiento que el de 34 KDa. Concluían, además, que la actividad terminadora de la transcripción estaba asociada sólo a los componentes de 34 KDa.

Silva y col. (Fernandez-Silva *et al.*, 1997) realizaron estudios de unión a DNA_{TER} mediante geles nativos de retardo, de versiones truncadas de mTERF (Δ N, residuos 88-399 y Δ C, residuos 58-365). Observaron que, de los dos fragmentos, sólo Δ N mantenía la capacidad de unión a DNA aunque disminuía cinco veces su afinidad con respecto a la proteína entera.

Considerando estos ensayos, y dada la capacidad de unión a DNA del polipéptido de 31KDa que copurificaba en la fracción de mTERF (Daga *et al.*, 1993), se puede deducir que éste correspondería a un fragmento que presenta degradación de la parte N-terminal, ya que mantiene la capacidad de unión a DNA.

Los estudios realizados en esta tesis muestran que mTERF- Δ N (residuos 99-399) es un fragmento estable de mTERF y tiene capacidad de unión a DNA. Así pues, podría coincidir con la forma más corta que se purifica sistemáticamente en los lisados mitocondriales. Además, el hecho que esta forma más corta no tenga actividad terminadora podría ser debido a que no es capaz de curvar del DNA como se observa en la estructura cristalográfica. Queda por resolver si este hecho es verdad y si tiene significado biológico, es decir, si esta forma más corta se encuentra en la mitocondria o si es el resultado de la proteólisis en el proceso de purificación.

8_ Comparación de estructuras

8.1 Comparación de las estructuras mTERF-DNA15pb y mTERF-DNA22pb

Paralelamente a la publicación de los resultados mencionados en esta tesis, se publicó un artículo (Yakubovskaya *et al.*) en el que se describe la estructura de mTERF (residuos 73-396) con un oligonucleótido de 22 pb a 2.2 Å de resolución.

Esta estructura cristalográfica también muestra que mTERF se une a la largo del surco mayor del DNA, aunque en este caso, cada molécula de proteína se une a una sola molécula de DNA de doble cadena provocando en éste una curvatura de aproximadamente 25° y un despliegue del DNA en la parte central de la secuencia. Esta parte central se encuentra altamente alterada ya que mTERF promueve además, la exposición hacia el exterior de tres nucleótidos estabilizados por la proteína mediante *stacking* y puentes de hidrógeno con la base y el fosfato. En la Fig 2.34 se comparan las interacciones entre mTERF y el DNA de las dos estructuras cristalográficas mTERF-DNA22pb y mTERF-DNA15pb. Observamos que a parte de las numerosas interacciones con el esqueleto de fosfatos, en ambas estructuras las argininas Arg169, Arg 202, Arg251 y Arg387 interaccionan con las bases nitrogenadas del DNA.

En el artículo, Yakubovskaya y col. demuestran mediante ensayos de calorimetría (*isothermal titration calorimetry*, ITC) que mTERF se une específicamente a la secuencia de reconocimiento (estequiometría 1:1). También, aunque con menor afinidad, se une a secuencias arbitrarias de DNA con un ratio DNA:proteína relativamente bajo, indicando que, en estos casos, mTERF no se asocia preferentemente al DNA en una conformación particular, sino que se puede unir en diferentes regiones del dúplex y en consecuencia, más de una molécula de mTERF puede unirse simultáneamente a una misma molécula de DNA. Así, una interacción inespecífica con el DNA explicaría la unión de dos moléculas de proteína sobre un mismo DNA y la disposición en sentido inverso de los DNAs en las estructuras cristalográficas de los complejos de mTERF resueltas en esta tesis.

de DNA (Fig 2.35c). En este sentido, la estructura resuelta en esta tesis explicaría el modelo de interacción inespecífica, en que dos proteínas, una mediante el subdominio N-terminal y la otra mediante el C-terminal unen un mismo DNA (Fig 2.35b). Aun así, si imaginamos una molécula continua más larga de DNA, ésta chocaría estéricamente con los subdominios N y C-terminal de las moléculas de proteína que lo unen.

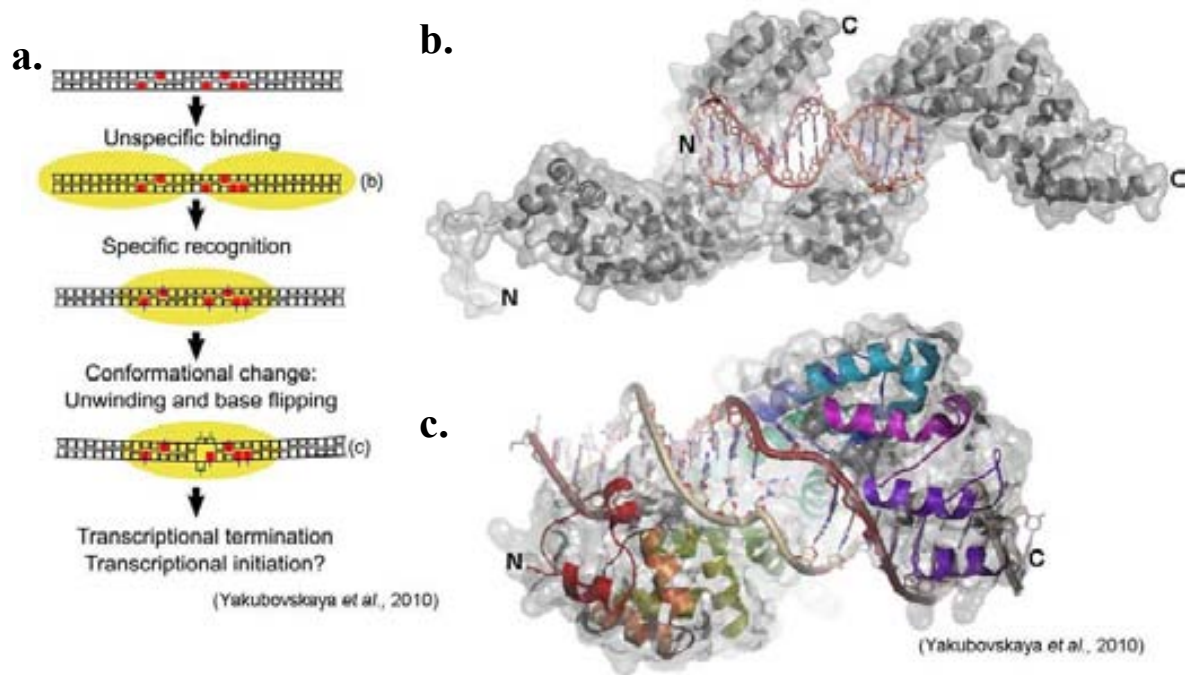


Fig 2.35_ (a) Modelo que describe el proceso de la unión específica de mTERF (óvalo amarillo) al DNA. En rojo se indican las bases nitrogenadas que interactúan específicamente con las argininas. (b) Estructura cristalográfica del complejo mTERF-DNA15pb. Detalle del empaquetamiento cristalográfico en que dos moléculas distintas de proteína unen una misma molécula de DNA de 15 pb, mediante la parte c-terminal y N-terminal de una y otra respectivamente. (c) Estructura cristalográfica del complejo de mTERF con el DNA (22 pb) que contiene la secuencia de terminación (Yakubovskaya et al.). Cada motivo mterf está coloreado con un color diferente. En gris transparente se muestra la superficie proteica. La cadena gris del DNA corresponde a la cadena ligera y la marrón, a la pesada. La unión de mTERF provoca una curvatura y una disminución del twist del DNA que se traduce en un despliegue de éste en la parte central. Así, mientras los extremos conservan una conformación de B-DNA, la parte central se encuentra altamente alterada.

Superposición de las estructuras cristalográficas

Si superponemos las estructuras de mTERF-DNA15pb y mTERF-DNA22pb, el rmsd es de 1.64Å. Como podemos observar en la superposición SSM Superposition realizada mediante el programa coot (Krissinel & Henrick, 2004) de la figura Fig 2.36, la parte central de mTERF es prácticamente idéntica (rmsd de 0.7 Å para los residuos 200 a 300), mientras los subdominios N y C-terminal están ligeramente desplazados.

En el artículo de Yakubovskaya y col. se describe que las interacciones que confieren especificidad son aquellas que se dan entre las argininas Arg169, Arg 202, Arg251, Arg350 y Arg387 y las bases nitrogenadas del DNA. En el mismo trabajo, los experimentos mediante ITC con los mutantes simples para cada arginina por alanina, mostraron que únicamente el mutante R387A perdía completamente la especificidad de unión. El resto de mutantes parecían conservar cierta especificidad de secuencia, aunque todos (excepto R350A) mostraban una menor afinidad de unión por la secuencia de terminación con respecto a la proteína WT. Si miramos la disposición de las cadenas laterales de estos residuos en ambas estructuras, ésta es muy similar - la distancia máxima entre C_{α} es 2.4 Å.

La estructura cristalográfica del triple mutante R162A/F243A/Y288A (Yakubovskaya *et al.*), en el que se eliminan las interacciones por *stacking* con los tres nucleótidos que se exponen hacia el exterior, dio lugar a una estructura prácticamente idéntica a la del complejo con la proteína WT, demostrando que la distorsión de la doble hélice del DNA se produce en la unión y es independiente de la exposición de los tres nucleótidos. En nuestro caso, mTERF une por cada extremo, N y C-terminal, una molécula distinta de DNA, dejando el centro de la proteína libre. En esta estructura no se produce ninguna distorsión del DNA y por tanto, con respecto a ese punto, la proteína no soporta tensión alguna. Aun así, la disposición en ambas estructuras de las argininas que intervienen en el contacto con las bases del DNA y que confieren la especificidad, es muy similar (Fig 2.36b y c). Así pues, parece que el conjunto de las interacciones (todas o algunas) de las argininas de la proteína con el DNA provocarían la tensión que da lugar a la distorsión del DNA. Es decir, dada la interacción inicial, el DNA no podría disipar la tensión porque se encuentra fijado por los extremos y por eso la libera en el centro a través del *untwisting* o desenrollamiento de la doble hélice. La estructura de mTERF-DNA22pb indica que la unión proteína-DNA se estabilizaría mediante las interacciones de *stacking* que los residuos R162, F243 y Y288 establecen con las bases que se exponen hacia el exterior.

Comparando las dos estructuras de mTERF-DNA resueltas, podríamos proponer que el cambio conformacional que observamos entre ambas proteínas, se explicaría por la

acomodación de la proteína, en el caso de la estructura de Yakubovskaya y col., a la tensión generada por el DNA.

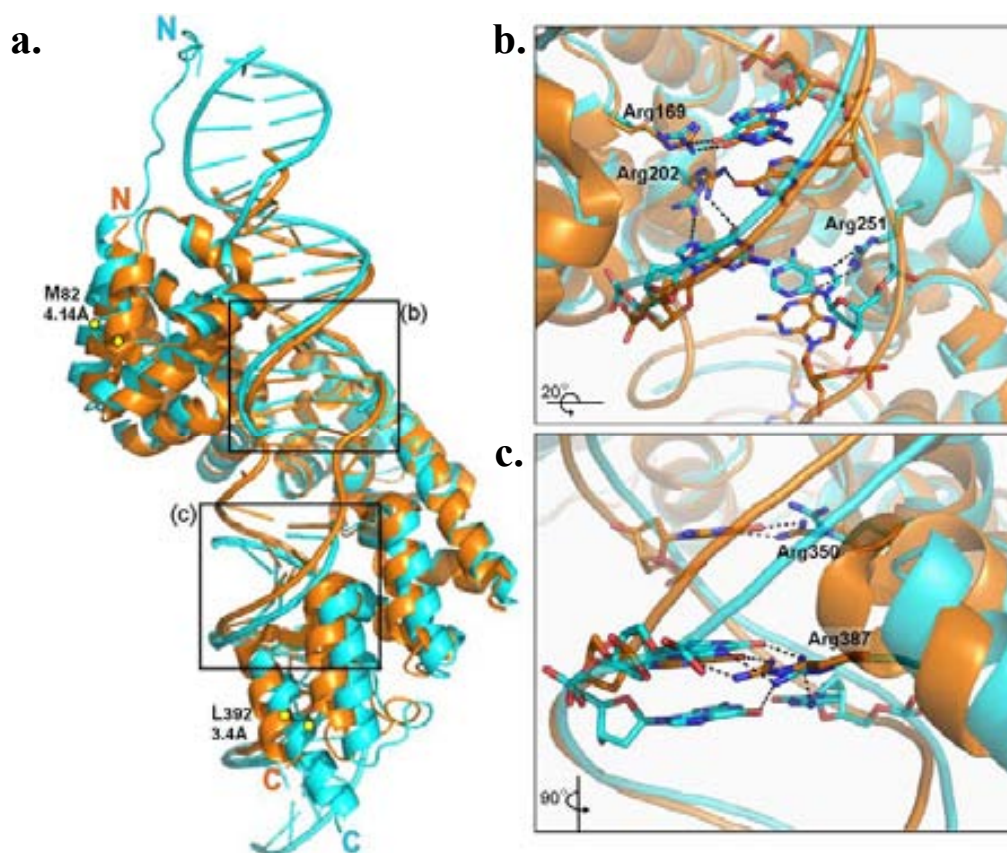


Fig 2.36_ (a) Superposición de las dos estructuras de mTERF-DNA resueltas. En azul se muestra la estructura resuelta en esta tesis y en naranja la estructura resuelta por Yakubovskaya et al. Se detallan las distancias entre el residuo M82 (subdominio N-terminal) y el residuo L392 (subdominio C-terminal) de las dos estructuras. (b) Detalle de la posición de Arg169, Arg202 y Arg 251 y de las interacciones que establecen con las bases nitrogenadas. (c) Detalle de la posición de Arg350 y Arg387 y de las interacciones que establecen con las respectivas bases nitrogenadas. En el caso de Arg350, en la estructura que resolvimos, no llega a interactuar con la base nitrogenada.

Para representar gráficamente el movimiento de una estructura a otra (Fig 2.37), se generaron modelos intermedios mediante el programa MovieMaker (Maiti *et al.*, 2005). Este programa no realiza cálculos de dinámica molecular, sino que es una herramienta de animación que usa algoritmos de superposicionamiento conjuntamente con la interpolación de coordenadas cartesianas para calcular rápida y automáticamente los intermediarios necesarios para una animación de estructuras de macromoléculas. Si representamos estos intermedios para las estructuras de mTERF-DNA15pb y mTERF-DNA22pb, observamos que el movimiento parece consistir en un cierre hacia el interior de los extremos N y C-terminal de mTERF-DNA15pb con respecto a mTERF-DNA22pb.

En la Fig 2.37a y b se muestra el movimiento superponiendo las estructuras mediante SSM y superponiéndolas por el subdominio C-terminal (los 100 últimos residuos), respectivamente.

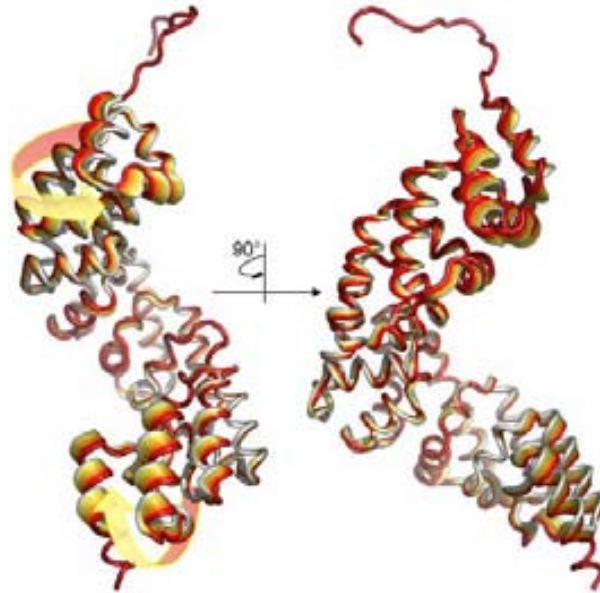


Fig 2.37_ Representación del movimiento que describe el cambio de conformación observado entre la estructura de mTERF resuelta en esta tesis y la resuelta por Yakubovskaya et al., si superponemos las dos estructuras mediante SSM (coot, (Krissinel & Henrick, 2004)). La estructura de mTERF resuelta en esta tesis está coloreada en rojo y la estructura resuelta por Yakubovskaya et al. en blanco. Las flechas indican la dirección del movimiento de los extremos N y C-terminal.

Comparación por modos normales

Si se dispone de dos conformaciones de una misma proteína, como el caso que aquí se discute, se pueden identificar los modos normales que más contribuyen al correspondiente movimiento estructural proteico (Suhre & Sanejouand, 2004). Un modo normal de un sistema oscilatorio es un patrón de movimiento en el cual todas las partes del sistema se mueven sinusoidalmente con la misma frecuencia y en fase. La frecuencia natural o de resonancia es la frecuencia a la cual la estructura deformable oscilará al ser perturbada. Recientemente se ha demostrado que un 50% de los casos donde la estructura de una proteína está descrita mediante dos conformaciones diferentes, el movimiento que las relaciona puede ser descrito aplicando una perturbación en la dirección de uno o dos modos normales de baja frecuencia de la proteína considerada. El programa ElNémo (*Elastic Network Model*) permite visualizar y analizar los modos normales de baja frecuencia de macromoléculas grandes.

a.

<R2>	perturbación	frecuencia	RMSD	overlap acumulativo	amplitud (dq)
modo 7	-100 (20) 100	1.00	1.289	0.291	-59.2878
modo 8	-100 (20) 100	1.40	1.572	0.439	-42.3207
modo 9	-100 (20) 100	1.66	1.746	0.456	14.0386
modo 10	-100 (20) 100	2.80	1.749	0.465	-10.8438
modo 11	-100 (20) 100	2.95	1.749	0.482	-13.9816
modo 12	-100 (20) 100	3.81	1.760	0.482	-0.1148
modo 13	-100 (20) 100	4.39	1.731	0.510	18.2625
modo 14	-100 (20) 100	5.26	1.760	0.514	6.7363
modo 15	-100 (20) 100	5.47	1.760	0.516	5.1794
modo 16	-100 (20) 100	5.96	1.760	0.521	7.4366

b. Modo Normal 7

Parámetros del modelo	dist < 3Å	RMSD
modelo 1 MODO=7 DQ=-100	306	1.325
modelo 2 MODO=7 DQ=-80	315	1.248
modelo 3 MODO=7 DQ=-60	316	1.185
modelo 4 MODO=7 DQ=-40	315	1.225
modelo 5 MODO=7 DQ=-20	309	1.325
modelo 6 MODO=7 DQ=0	297	1.447
modelo 7 MODO=7 DQ=20	275	1.502
modelo 8 MODO=7 DQ=40	251	1.492
modelo 9 MODO=7 DQ=60	240	1.582
modelo 10 MODO=7 DQ=80	228	1.697
modelo 11 MODO=7 DQ=100	206	1.698

Verde Residuos más cerca de 3Å de la estructura de referencia
Azul RMSD entre residuos más cercanos de 3Å
Rojo RMSD entre todos los residuos



Fig 2.38_ (a) En la tabla se muestran los 10 primeros modos normales. A cada modo normal le corresponde una frecuencia de movimiento. La perturbación indica las amplitudes que han sido aplicadas en las perturbaciones del modo normal, desde DQMIN=-100 a DQMAX=100 en escalas de 20. El RMSD indica la desviación media cuadrática mínima entre los modelos del modo normal perturbado y la conformación de la segunda estructura. El overlap mide el grado de similitud entre la dirección de un cambio conformacional dado -en este caso entre la estructura resuelta en esta tesis y la estructura resuelta por Yakubovskaya et al. superpuestas mediante SSM -, y la dirección dada por el modo normal considerado. Un valor de 1 para el overlap significa que la dirección dada por el modo normal considerado es idéntico al cambio conformacional. En este caso, en la tabla se muestra el overlap acumulativo de los modo normales con respecto al cambio conformacional. Así pues, el modo normal que mejor superpone la estructura de Yakubovskaya y col. sería el modo normal 7. (b) Para este modo se muestra la comparación entre cada uno de los modelos del modo normal perturbado y la segunda conformación, en este caso la estructura resuelta por Yakubovskaya y col. (c) Representación esquemática del movimiento de los modos normales 7 y 8.

Introducimos en el programa las dos conformaciones de mTERF, la proteína resuelta en esta tesis y la proteína resuelta por Yakubovskaya y col. (Yakubovskaya *et al.*), superpuestas

mediante SSM Superposition (Krissinel & Henrick, 2004). Como se muestra en la Fig 2.38a, ElNémo calculó el grado de movimiento colectivo para todos los modos normales y mostró la contribución (amplitud dq) de cada uno de los 100 modos de más baja frecuencia al cambio conformacional para las estructuras superpuestas. Calculó también el RMSD entre los modelos generados de cada modo normal perturbado y la conformación de la segunda estructura (Fig 2.38c) para identificar las perturbaciones del modo normal que mejor describen el movimiento asociado a la proteína.

Para la estructura de mTERF resuelta en esta tesis, el modo normal que da lugar al modelo que mejor superpone la estructura de Yakubovskaya y col. sería el modo normal 7, con un rmsd de 1.289Å. Con respecto al movimiento que relacionaría las dos conformaciones, los modos normales que más contribuyen son el 7 y el 8 (59.2 y 42.3 dq), aun así, éstos únicamente podrían describir un 44% del movimiento. En este caso, este hecho podría ser debido a que el cambio conformacional observado en la estructura de mTERF-DNA22pb se deba a la tensión que ejerce el DNA deformado sobre la proteína y por tanto no coincida con el movimiento intrínseco de la proteína sola que definen los modos normales. Aun así, hay que tener en cuenta que los contactos cristalográficos de la proteína en el cristal podrían ser causantes de parte de la deformación proteica. Además, al ser un desplazamiento relativamente pequeño, el ruido estructural puede ser significativo.

Análisis de los datos de SAXS con la estructura de mTERF-DNA22pb

Dado que se disponía de la estructura unida a un DNA continuo resuelta por Yakubovskaya y col. (Yakubovskaya *et al.*), se decidió contrastarla con los datos de SAXS. Se generó un modelo del complejo partiendo de la estructura cristalográfica de mTERF-DNA 22pb resuelta por Yakubovskaya *et al.* con un oligonucleótido de 28pb -el DNA_{TER} utilizado en los ensayos de SAXS- (Fig 2.39a), manteniendo la perturbación que se da en el centro de la proteína, para ver cómo se ajustaba a la curva de dispersión medida por SAXS. Se utilizó el programa CRY SOL para comparar la curva teórica del modelo con respecto a la curva de dispersión experimental. Si tomamos como modelo la proteína (residuos 73-396) y el DNA de 28pb el error (χ^2) asociado es de 3.566 (Fig 2.39b_{arriba}), y si a la proteína le añadimos los residuos del 58 al 73 de la estructura cristalográfica de mTERF-DNA15pb, el error disminuye significativamente a 2.801 (Fig 2.39b_{abajo}).

Si por el contrario, tomamos como modelo la proteína resuelta en esta tesis y el DNA de 28 pb resultante de añadir las 6pb que le faltan al DNA de la estructura resuelta por Yakubovskaya y col. y comparamos la curva teórica del modelo con respecto a la curva de dispersión experimental, el error (χ_i) asociado es de 2.24.

En este sentido, los modelos del 9 al 12 generados anteriormente (χ_i asociado de aproximadamente 2.0) describen mejor la curva de SAXS que el modelo construido a partir de la estructura de mTERF con el DNA de 22pb.

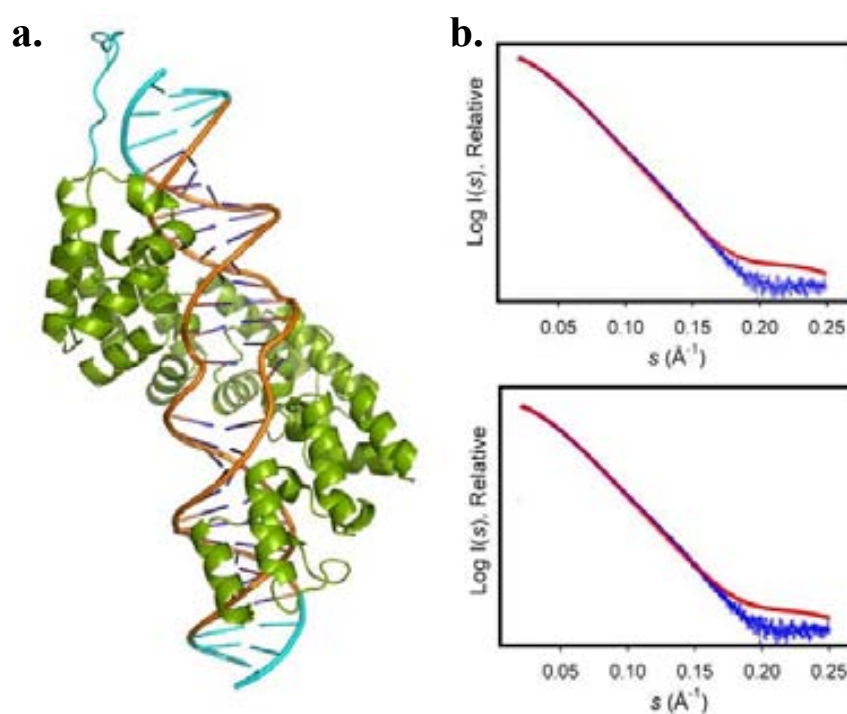


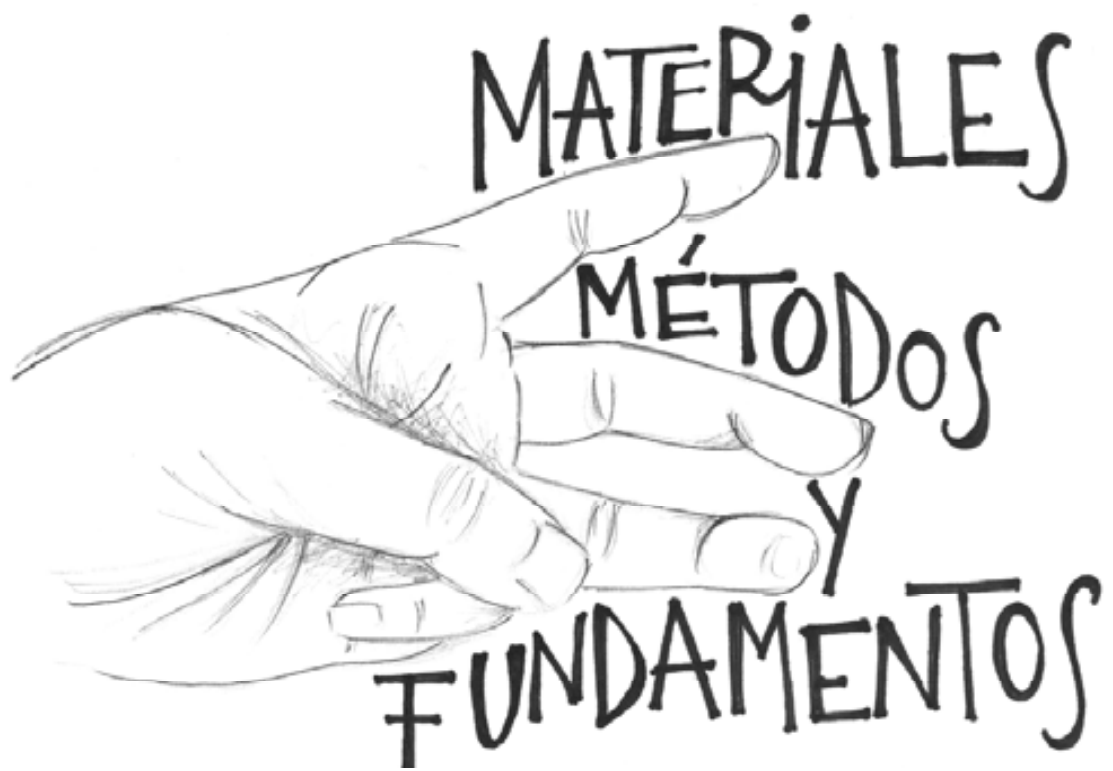
Fig 2.39_ (a) Modelo para los estudios de SAXS, generado a partir de la estructura del complejo de mTERF (residuo 73-396 en verde) y el DNA de 22pb (en naranja), resuelta por Yakubovskaya y col., y añadiendo al DNA 3pb en cada extremo y a la proteína los residuos 58-72 de la estructura de mTERF-DNA15pb (en la figura se muestran en cian). (b) Intensidad de dispersión (azul) en escala logarítmica en función del momentum transfer de la muestra mTERF-NHis-DNA_{TER} (6mg ml⁻¹). La línea continua roja representa el encaje de CRY SOL del modelo -basado en la estructura resuelta por Yakubovskaya y col.- con la curva de dispersión. (arriba) El modelo contiene los residuos del 73-396. (abajo) El modelo contiene los residuos del 58-396.

Conclusiones

1. El factor de terminación de la transcripción mitocondrial humano (mTERF) entero, así como las formas truncadas mTERF- Δ N (residuos del 99 al 399) y mTERF- Δ C (residuos del 58 al 360) se expresan en forma soluble en *E. coli*.
2. Se ha establecido un protocolo de purificación para cada una de las cuatro formas proteicas estudiadas, incluyendo, a parte de las mencionadas en la conclusión anterior, la forma obtenida mediante proteólisis limitada Δ N-mTERF- Δ C (residuos del 99 al 360).
3. Se ha cristalizado y determinado la estructura a 3.1 Å de resolución del factor de terminación de la transcripción mitocondrial, mTERF, en complejo con un DNA de 15 pb que contiene la secuencia de terminación.
4. La estructura cristalográfica de mTERF consiste en nueve repeticiones estructurales, TERF-I a IX, flanqueadas por un segmento N-terminal y una hélice α C-terminal, la hélice X.
5. Cada motivo TERF está formado por unos 35 aminoácidos que se estructuran en tres hélices α (H1, H2 y H3) formando una superhélice trigonal levógira con un núcleo hidrofóbico central. Esta superhélice se propaga en tándem con una rotación dextrógira de 7 a 20° construyendo un solenoide, que a su vez muestra una torsión hacia la derecha dando lugar a una cara convexa y otra cóncava.
6. El motivo estructural de superhélice levógira se encuentra únicamente en pequeños dominios de cuatro hélices. La combinación de la conectividad en superhélice levógira con la propagación en tándem del motivo para formar un solenoide es única de las proteínas de la familia MTERF.

7. A partir del alineamiento estructural de los nueve motivos de mTERF con los correspondientes motivos de los ortólogos, se generó un logo en el que se muestra el grado de conservación de secuencia para cada posición del motivo. Los residuos que forman el núcleo hidrofóbico del motivo; los que están involucrados en el *capping* de las hélices y los que rompen la regularidad de los elementos de estructura secundaria dando lugar a la forma trigonal característica del motivo TERF, son los más conservados.
8. En el cristal de mTERF con el DNA de 15 pb una misma proteína une dos moléculas de DNA, DNA1 y DNA2 (relacionado por simetría con DNA1), y cada molécula de DNA está unida por dos proteínas simétricas. Por tanto, la relación estequiométrica en el cristal es de una molécula de DNA por cada molécula de proteína.
9. mTERF se une a la largo del surco mayor del DNA, estableciendo contactos con la molécula DNA1 mediante el subdominio N-terminal y con DNA2 mediante el subdominio C-terminal. Las moléculas DNA1 y 2 no interaccionan entre ellas y la posición relativa entre ambas viene dada por un ángulo de $\sim 36^\circ$. En la unión de mTERF al DNA están involucrados 19 residuos que interaccionan con los fosfatos y 5 residuos que interaccionan con las bases nitrogenadas del DNA.
10. Estos residuos coinciden, en su mayoría, con los aminoácidos que interaccionan con las bases nitrogenadas en la estructura de mTERF en complejo con el DNA22pb resuelta por Yakubovskaya y col., Arg169, Arg 202, Arg251, Arg350 y Arg387, y que son responsables de la especificidad de la interacción. La superposición de las dos estructuras (rmsd 1.64 Å) muestra que la disposición espacial de estos residuos es muy similar, indicando que el conjunto de las interacciones de las argininas de la proteína con el DNA provocarían la tensión que da lugar a la distorsión del DNA que se observa en la estructura de mTERF con el DNA22pb.
11. Se ha cristalizado y determinado la estructura a 2.4 Å de resolución del fragmento estable producto de la degradación proteolítica, al que le faltan 40 aminoácidos de la parte N-terminal, del factor de terminación de la transcripción mitocondrial,

- mTERF- Δ N, en complejo con un DNA de 12 pb que contiene la secuencia de terminación. La proteína presenta el mismo patrón de repeticiones TERF.
12. En el cristal de mTERF- Δ N en complejo con el DNA de 12 pb, la proteína une dos moléculas de DNA simétricas que interactúan mediante *stacking* o apilamiento de las bases dando lugar a una fibra pseudo-continua de DNA en el eje z. Esta disposición del DNA causó graves problemas de anisotropía en los primeros cristales obtenidos de mTERF- Δ N con un oligonucleótido de 29pb y después de varias pruebas de cristalización y difracción con DNAs menores, finalmente se pudo trazar completamente la proteína y el DNA en los cristales del complejo con un DNA de 12pb.
 13. Los residuos del subdominio N-terminal de mTERF- Δ N que intervienen en la unión con el DNA, establecen interacciones con los fosfatos de la molécula DNA1, mientras que los del subdominio C-terminal, establecen interacciones con los fosfatos y, mediante la Arg387, con las bases nitrogenadas T7 (cadena C) y G5 (cadena B) de la molécula DNA2.
 14. En la superposición de las estructuras de la proteína entera y la truncada por la parte C-terminal coinciden los esqueletos del DNA2 pero en sentidos opuestos y la parte N-terminal de mTERF- Δ N (TERFII-IV) presenta un desplazamiento, de manera que los residuos de la superficie cóncava de mTERF se disponen abriéndose hacia el exterior en la forma corta.
 15. El estudio mediante la técnica de SAXS de los complejos de mTERF y mTERF- Δ N con el oligonucleótido DNA_{TER}28pb en solución, indicó que la unión de mTERF al DNA no es aleatoria, sino que se da preferentemente sobre una zona del DNA.
 16. Los ensayos mediante gels nativos de retardo con el oligonucleótido DNA_{TER}28pb muestran que las formas truncadas mTERF- Δ C y Δ N-mTERF- Δ C no se unen al DNA, mientras mTERF y mTERF- Δ N unen la secuencia DNA_{TER} de forma específica. Así pues los 40 residuos terminales son esenciales en la unión al DNA.



1_Clonaje

Diseño de las construcciones

Para la expresión óptima de la proteína de estudio, mTERF, en *Escherichia coli* (*E. coli*), los constructos se diseñaron siguiendo la regla “N-end” (Tobias *et al.*, 1991), donde se relaciona la identidad de los residuos del N-terminal con la vida media de la proteína y se describe la importancia de que estos residuos no tengan una estructura secundaria regular predicha por análisis informáticos. mTERF es una proteína codificada en núcleo e importada a mitocondria mediante una secuencia de reconocimiento de 57aa (Fernandez-Silva *et al.*, 1997). Partiendo de estas premisas fueron diseñados dos constructos de la proteína madura (Phe58-Ala399) que comprendían del residuo Arg56 al Ala399 (UniProt database entry Q99551) y que incorporaban en el extremo N-terminal los aminoácidos metionina y alanina: **mTERF-NHis** con una cola de 6 histidinas N-terminal y una diana para TEV y **mTERF-CHis** con una cola de 6 histidinas C-term.

Se diseñó un tercer constructo **mTERF-ΔN** que comprende el fragmento Arg99-Ala399, donde faltan los primeros 41 residuos de la secuencia madura y que presenta una cola 6His C-term. Éste fue amplificado y clonado en el plásmido pET28a previamente digerido con NcoI/BamHI. **ΔN-mTERF-ΔC** es el producto de la proteólisis limitada de **mTERF-ΔN** con papaína. Por último, se diseñó, según los ensayos de proteólisis, un cuarto constructo **mTERF-ΔC** que incluye del residuo Ser54 al Leu303.

Diseño de cebadores para la reacción en cadena de la polimerasa (PCR)

En el diseño de los cebadores se deben tener en cuenta varias consideraciones.

- No debe existir ningún sitio alternativo en el *template* donde el cebador pueda hibridarse. La longitud del oligo determina la especificidad, por tanto debe ser de 20 a 24 bases para reducir al mínimo la posibilidad de hibridación en sitios secundarios en el vector o inserto. Aun así, los cebadores no pueden ser muy largos, porque esto afecta a la velocidad de hibridación y la eficiencia del

proceso. En consecuencia, los cebadores nunca deben exceder los 30 nucleótidos.

- Cada oligonucleótido debe tener el extremo 3' bien estabilizado para asegurar un buen anclaje en el punto de inicio de la amplificación.
- Se debe evitar la formación de horquillas intracatenarias y de dímeros cruzados entre un mismo oligonucleótido o entre ambos. Este punto se puede analizar mediante programas informáticos.
- Es importante que los cebadores tengan un contenido de C/G entre 40 y 60%. La temperatura de fusión (T_m) del híbrido debería ser de 55°C o más, de manera que esté 1-2 grados por encima de la temperatura de hibridación durante la reacción de PCR. La T_m de los dos oligos utilizados en la PCR debe ser parecida.

La temperatura de hibridación óptima tiene que ser lo suficientemente baja para permitir la hibridación específica entre el cebador y su secuencia complementaria, pero lo suficientemente alta como para impedir hibridaciones no específicas.

Los respectivos cebadores *forward* y *reverse* (Sigma Aldrich) utilizados en las distintas PCR para amplificar los constructos mencionados están detallados en la Fig 3.1a; F1 y R1 para mTERF-NHis, F2 y R2 para mTERF-CHis, F3 y R2 para mTERF- Δ N y F2 y R3 para mTERF- Δ C.

PCR y clonaje

Se amplificó el cDNA de mTERF (GenBank: BC000965; comprado a ATCC) con los cebadores correspondientes para mTERF-NHis y mTERF-CHis mediante PCR usando la polimerasa Vent (Biolabs). El producto de PCR se purificó con el kit GFX (GE Healthcare). Una vez digeridos el inserto y el vector de expresión pET28a (Novagen) con las enzimas de restricción NcoI y BamHI (Fermentas), se purificaron y se ligaron *o/n* mediante DNA ligasa. Se transformaron mediante choque térmico células competentes D5 α con el producto de la ligación y las colonias positivas se mandaron a secuenciar una vez preparadas las minipreps (quiagen).

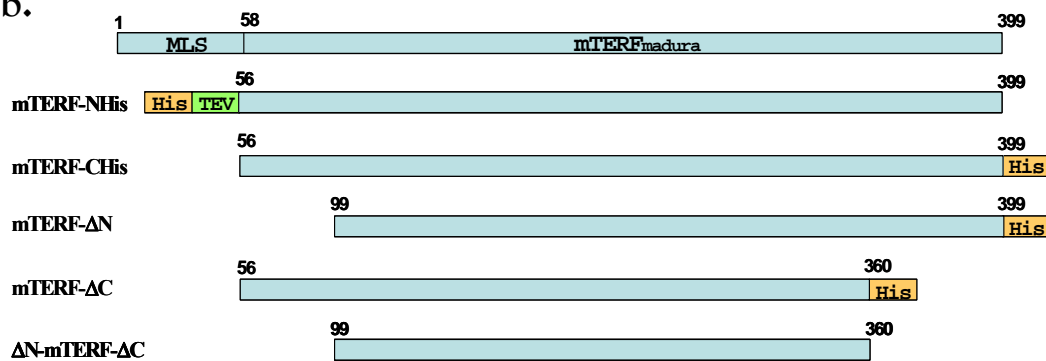
Para los constructos mTERF-ΔN y mTERF-ΔC se siguió el mismo protocolo pero partiendo del clon de mTERF-CHis como *template*.

Todos los constructos fueron confirmados por secuenciación de DNA (Macrogen y unidad de genómica de la universidad de Barcelona).

a.

Nombre	Secuencia	Diana
F1	5'-CATGCCATGGCCCATCATCATCATCAT GAAAACCTTTACTTCCAGGGCAGCAGGCTTTTTGGT-3'	NcoI
R1	5'-CGGGATCCCTTAGGCAAATCTGCTTAACTTTTTCAATTTAGC-3'	BamHI
F2	5'-CATGCCATGGCCAGCAGGCTTTTTGGTGTGAAG-3'	NcoI
R2	5'-CGGGATCCCTTAATGATGATGATGATGATG GGCAAATCTGCTTAACTTTTTCAATTTAGC-3'	BamHI
F3	5'-CATGCCATGGGTAGCAGGATGATTACCAATGAGCAGG-3'	NcoI
R3	5'-CGGGATCCCTTAATGATGATGATGATGATG TTTTAAAGTACTTATGCTTGAATCCAGAACCCG-3'	BamHI

b.



c.

Vector	Medida (pb)	Resistencia	Promotor	C-term	Diana	N-term
pET28a	5,369	Kan	T7lac	His(x6)	Trombina	His(x6)

Fig. 3.1_(a) Cebadores utilizados para la amplificación por PCR de los diferentes constructos. (b) Esquema de los constructos. La secuencia de transporte a mitocondria está indicada por MLS (mitochondrial leader sequence). ΔN-mTERF-ΔC es el producto de la proteólisis limitada de mTERF-ΔN con papaína. (c) Características del vector de clonación.

2_ Expresión y purificación de los constructos de mTERF

2.1 Optimización de la expresión bacteriana y del tampón de solubilización.

Ensayos de expresión

La bacteria gram negativa *Escherichia Coli* ofrece un método de producción de proteínas recombinantes que no requieran modificaciones posttraduccionales rápido, económico y de alto rendimiento. En el caso de mTERF existe cierta controversia sobre su posible fosforilación, pero como parece que mTERF humana es activa en la forma no fosforilada (Asin-Cayuela *et al.*, 2005) decidimos utilizar el sistema de expresión de *E. coli*.

Sin embargo, la producción a gran escala de proteínas eucariotas no es evidente (Sivashanmugam *et al.*, 2009). Existen muchos parámetros a tener en cuenta en la optimización de la sobreexpresión de una proteína heteróloga en *E. coli*.

Para establecer el protocolo de expresión, se realizaron pruebas con volúmenes de cultivo de 20-50 ml; variando la cepa bacteriana, la OD_{600nm} de inducción (densidad óptica del cultivo medida a $\lambda=600$ nm en la que se añade el inductor), la concentración del agente inductor de la expresión IPTG (Isopropil- β -D-1-tiogaláctopiranosido, compuesto no hidrolizable análogo de la lactosa), la temperatura y el tiempo de expresión tras añadir el inductor, y el tampón de lisis.

Parámetros de expresión:

ODinducción	[IPTG]*	Tiempo expresión	T ° expresión*
0.5-0.6 / 0.8-1	0.1 / 1 mM	4 h	37° C
		o/n	25° C
		o/n y 60 h	17° C

*Una de las estrategias para mejorar la solubilidad de la proteína es reducir la velocidad de síntesis de ésta mediante la reducción de la concentración del inductor o mediante la expresión a temperaturas menores.

Cepas bacterianas:

BL21(DE3)

Esta cepa es deficiente en las proteasas *lon* y *omp T*, y contiene el lisógeno DE3, que contiene el gen de la RNA polimerasa de T7 (bajo el promotor *lacUV5*) y el *lacI^q*. Adecuada para la expresión de proteínas clonadas en vectores tipo pET.

RosettaTM (DE3)

Derivada de la cepa anterior. Ha sido diseñada para mejorar la expresión de proteínas eucariotas que contienen “codones raros” para *E. coli*. Contiene los tRNAs para los codones AGG, AGA, AUA, CUA, CCC y GGA en un plasmidio que confiere resistencia a cloramfenicol bajo el control de los promotores nativos.

BL21(DE3)[pT-groE]

La coexpresión de las chaperoninas GroESL codificadas en el plasmidio pT-groE que confiere resistencia a cloranfenicol, ayudan al plegamiento proteico.

Medios de cultivo:

LB

Medio enriquecido. Formulación por litro: 10 g de tristona, 5 g de extracto de levadura, 10 g de NaCl. Se ajusta el pH a 7.0 y se autoclava.

SB

Medio de formulación comercial (Athena Enzyme Systems) diseñado para aumentar la expresión de proteínas recombinantes en *E. coli*. Está tamponado a 7.2 ± 0.2 para evitar la acidificación del medio debida a los metabolitos producidos en el metabolismo bacteriano de la glucosa. Como este medio produce densidades de crecimiento mucho más elevadas que el medio LB, se debe inducir la expresión a densidades ópticas de aproximadamente el doble.

Autoinducción

Se realizaron también algunos batch de expresión siguiendo el método de autoinducción (Studier, 2005). En las cepas utilizadas, la polimerasa T7 está bajo el control del promotor *Lac-UV5*. Cuando las células crecen en un medio sin lactosa, el represor *lac* (*lacI*) se une al operón *lac* y previene la transcripción. Cuando la lactosa es la única fuente de carbono o se añade al medio un análogo como el IPTG, la lactosa o el IPTG se unen al represor e induce la disociación del operón permitiendo así la transcripción. Además, la adición de glucosa al medio contribuye a la represión de la RNA polimerasa T7 mediante la vía de represión catabólica.

En el protocolo clásico para la inducción de la expresión de proteínas en sistemas T7, las bacterias se crecen en un medio rico (como el LB) hasta una fase media-avanzada y la expresión se induce mediante la adición de IPTG. Este protocolo requiere un seguimiento del crecimiento de los cultivos ya que deben ser inducidos antes que alcancen la saturación. Además, una vez inducidos, se detiene la proliferación bacteriana ya que las células se dedican con avidez a producir las proteínas del plásmido que contienen el promotor T7.

La ventaja de la autoinducción es que el crecimiento bacteriano se produce en un medio donde la expresión de la polimerasa T7 se induce automáticamente en una fase avanzada de crecimiento debido al agotamiento de las fuentes de carbono alternativas a la lactosa. A medida que estas fuentes son consumidas, las células son forzadas a utilizar lactosa, punto en el que se expresa la plimerasa T7. Además, una vez que la glucosa se agota, los niveles de cAMP aumentan y se libera la represión catabólica.

Las ventajas de este protocolo son:

- No es necesario seguir el crecimiento del cultivo antes de la inducción.
- El rendimiento de los cultivos es mayor en masa celular y como el medio está bien tamponado y la polimerasa T7 se induce en etapas de crecimiento avanzadas, normalmente es también mayor el rendimiento en proteína recombinante.

- Los cultivos se pueden incubar overnight ya que no necesitan ser manipulados.

Soluciones stock:

1M MgSO₄

Solución de metales (1000x)

Se disuelven en 1L de H₂O, 8 ml de HCl 5 M, 5 g de FeCl₂·4H₂O, 184 mg de CaCl₂·2H₂O, 64 mg de H₃BO₃, 18 mg de CoCl₂·6H₂O, 4 mg de CuCl₂·2H₂O, 340 mg de ZnCl₂, 605 mg de Na₂MoO₄·2H₂O, 40 mg de MnCl₂·4H₂O y se filtra.

Solución 5052 (50x) (200ml)

Solución de 25% (v/v) glicerol, 2.5 % (w/v) glucosa y 10% (w/v) lactosa. Se esteriliza por autoclave.

Solución M (50x) (250ml)

Solución de 1.25 M Na₂HPO₄, 1.25 M KH₂PO₄, 2.5 M NH₄Cl y 0.25 M Na₂SO₄. Se esteriliza por filtración.

Medio de auto-inducción ZYM 5052:

En un erlemeyer de 2L se mezcla 380 ml de H₂O, 4 g de triptona y 2 g de extracto de levadura. Una vez esterilizados por autoclave se añade 800 µL de 1 M MgSO₄, 800 µL de la solución de metales, 8ml de solución 5052, 8ml de solución M, antibiótico y 4 ml de precultivo.

Protocolo:

Los precultivos se deben crecer o/n a 37° C, se inoculan y los cultivos se dejan crecer durante 3 h a 37° C, momento en el cual empezamos a bajar la temperatura hasta 20° C. Se dejan agitando a 20° C durante 36 h.

En este caso, en vez de 20° C, la temperatura de expresión fue 17° C.

Tampón de lisis

A la hora de escoger un tampón de lisis, es importante tener en cuenta los parámetros físico químicos de la proteína: el pH utilizado que debe ser pH ±1-1.5 del punto isoeléctrico (pI); las cisteínas que contenga la proteína, ya que se debe

añadir algún agente reductor para evitar la formación de puentes disulfuro y las características funcionales de la proteína, por ejemplo con proteínas unión a DNA se debe añadir concentraciones elevadas de sal para neutralizar las cargas. Es importante también tener en cuenta los compuestos y las concentraciones máximas toleradas en la columna que será utilizada a continuación.

En la tabla se muestran los diferentes compuestos y concentraciones utilizados en las primeras pruebas de lisis, que fueron realizadas en base al tampón de lisis mitocondrial descrito para la purificación de mTERF a partir de células HeLa; 25 mM HEPES pH 7.6, 100 mM KCl, 5mM MgCl₂, 0.5 mM EDTA, 1 mM DTT, 0.2 mM PMSF, 0.1 μM pepstatina, 10% glicerol y 0.1% Nonidet P-40. Se varió la concentración de NaCl según el gradiente de sal utilizado en la posterior columna de heparina-agarosa (0.3, 0.5 y 0.8 M KCl) (Daga *et al.*, 1993).

Tampón	Sal	antiproteasas	reductor	otros
50 mM Hepes pH 7.4	0.25 M NaCl	1mM PMSF ⁺	1mM DTT	glicerol
50 mM TRIS pH 7.4	0.5 M NaCl		5-10 mM BME	sacarosa
	1 M NaCl			urea*
	0.5 M KCl			detergentes*

* para purificar en condiciones nativas las concentraciones deben ser muy bajas, 0.1% en detergentes y 0.5M de urea.

Tabla 3.2_ Componentes utilizados en los tampones de lisis de los primeros ensayos de expresión y solubilidad de mTERF.

ThermoFluor: Optimización del tampón proteico

Una vez obtenemos suficiente cantidad de proteína pura, podemos hacer ensayos mediante la técnica de thermofluor para determinar el tampón en el que la proteína es más estable térmicamente (Ericsson *et al.*, 2006). Para monitorizar la estabilidad térmica de una proteína y testar qué factores (tampones, aditivos, ligandos...) afectan a dicha estabilidad, este ensayo (*fluorescence-based thermal shift*) utiliza un colorante sensible a la hidrofobicidad del entorno, como el Sypro Orange (Molecular Probes). Durante el proceso de desnaturalización las superficies hidrofóbicas van quedando expuestas al medio incrementando así, la fluorescencia del colorante. La desnaturalización térmica de una proteína es un proceso que sigue un modelo típico de dos estados con una marcada transición

entre los estados nativos y los desnaturalizados, donde la temperatura de fusión (melting, T_m) se define como la temperatura en el punto medio de la transición en la desnaturalización de la proteína que corresponde al máximo en la primera derivada de la curva de fusión (Fig 3.3). Las T_m obtenidas mediante thermofluor, para la mayoría de proteínas, semejantes a las obtenidas por otros métodos biofísicos como el dicroísmo circular, las medidas turbidimétricas y la calorimetría de barrido diferencial. Las medidas se realizaron en una máquina de PCR real-time y los cambios de fluorescencia en las muestras se monitorizan simultáneamente con una cámara CCD. Las longitudes de onda de excitación y emisión del Sypro son 490 y 575 nm, respectivamente.

Este método fue utilizado para la selección de los tampones óptimos en la purificación.

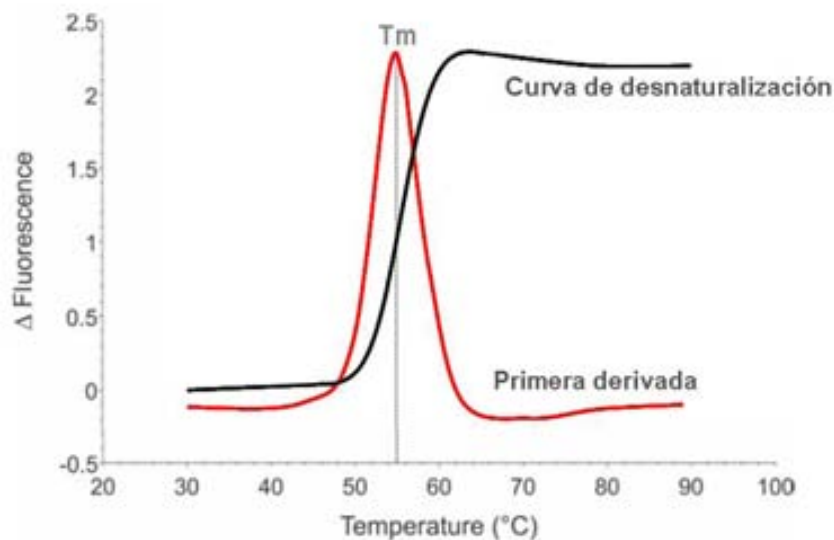


Fig 3.3_ Esquema de una curva típica de thermofluor (negro) y su primera derivada (rojo). Podemos observar cómo incrementa la fluorescencia al aumentar la temperatura debido al proceso de desnaturalización de la proteína. La temperatura de fusión (T_m), en este caso $\sim 55^\circ$ se define como el punto de inflexión de la curva o el máximo de la primera derivada. La estabilidad térmica de la proteína en las diferentes condiciones testadas se determina según el valor de T_m , de manera que cuanto mayor es, más estable.

Primero se debe hacer un ensayo para determinar la concentración óptima de la proteína de interés a la que se realizarán los experimentos. Interesa que sea la mínima cantidad que dé una señal aceptable.

Se utilizan placas especiales de 96 pocillos (placa PCR Bio-Rad). El volumen final de la muestra en cada pocillo depende de la calibración del aparato que se utilice para realizar el experimento. En este caso eran 50 μl .

Se llevaron a cabo ensayos de tampones, sales y aditivos con mTERF- ΔN :

Tampones o sales	Aditivos
25 μl condición nueva (2x)	5 μl aditivo (10x)
13 μl de H_2O	33 μl de tampón
5 μl de proteína (2.5 mg ml^{-1})	5 μl de proteína (2.5 mg ml^{-1})
7.5 μl Sypro Orange (300x)	7.5 μl Sypro Orange (300x)

*En las muestras control se añadió tampón en vez de H_2O y de la condición.

2.2 Métodos de purificación

La cromatografía se basa en la separación de los componentes de una muestra debido a las características físicas o químicas de éstos, que determinan la diferente interacción con la fase móvil y la fase estacionara (matriz o resina) de la columna.

Cromatografía de afinidad

Este tipo de cromatografía permite separar un componente, una proteína en este caso, de una mezcla gracias a la interacción específica y reversible con un ligando unido a la matriz.

Columna de níquel

Este tipo de cromatografía permite separar del crudo la proteína de interés mediante la interacción de la cola de histidinas con el níquel que se encuentra quelado en la resina. La proteína retenida se eluye con un gradiente creciente de imidazol. Es importante ajustar el tipo de gradiente y la pendiente de éste para poder sacar el máximo rendimiento a la purificación. Hay que tener en cuenta que no se debe utilizar EDTA (agente quelante) como inhibidor de proteasas y que la concentración de agentes reductores como el DTT y el BME debe ser mínima. En el caso de mTERF, se utilizaron columnas de 5ml de HisTrap HP (Amersham) para la purificación de todos los constructos.

Columna de Heparina

La heparina es un polímero de tipo glicosaminoglicano que imita el esqueleto azúcar-fosfato del DNA y permite la unión con alta afinidad de proteínas de unión a ácidos nucleicos. Además debido a su carga negativa, crea un efecto de intercambio catiónico y se unen también todo tipo de proteínas básicas. El problema básico a la hora de utilizar esta columna para la purificación de mTERF es la condición de una concentración de sal baja en la muestra inyectada, ya que mTERF precipita en estas condiciones.

Cromatografía de exclusión molecular

La cromatografía de exclusión molecular es una técnica que permite separar moléculas en función de su radio hidrodinámico, en general proporcional a su tamaño molecular (Fig 3.4).

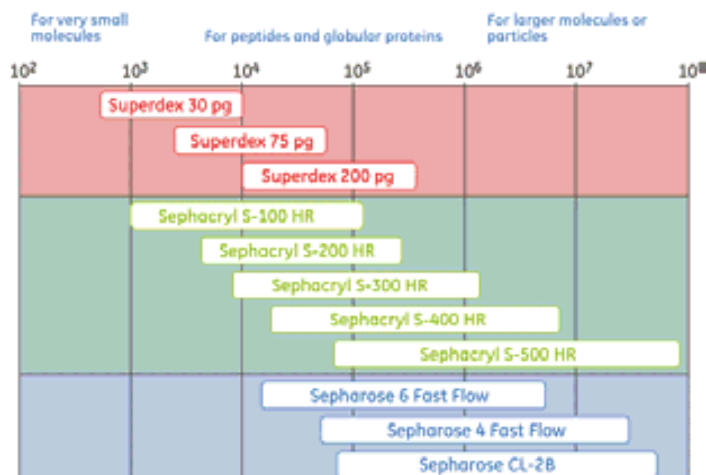


Fig 3.4_ La capacidad separadora de la cromatografía de exclusión molecular reside fundamentalmente en el gel cuya matriz consta de un gran número de esferas porosas microscópicas. Cada gel se caracteriza por un rango de fraccionamiento que depende del tamaño de sus poros.

La cromatografía de exclusión molecular es una técnica especialmente adecuada para el último paso de purificación. Debemos tener en cuenta que sólo eliminará contaminantes de peso molecular diferente a la proteína de interés y que la muestra purificada se eluye diluida. Aun así, permite eliminar si existe parte de proteína agregada, diferenciar entre diferentes estados oligoméricos y obtener la proteína en el tampón final para la cristalización

Las columnas utilizadas según la cantidad de proteína purificada fueron: para volúmenes 0.25 ml de muestra, las Superdex 75 10/300 GL y Superdex 200 10/300 GL; y para volúmenes de 2 ml, la HiLoad 16/60 Superdex200 (GE Healthcare). El tampón de purificación (fase móvil) utilizado fue aquél en el que la proteína presenta una mayor estabilidad.

2.3 Expresión y purificación de proteína a gran escala

Para la expresión se utilizó la cepa de *E. coli* DL21 (DE3). Las células transformadas con el correspondiente plásmido mediante choque térmico, fueron crecidas durante 4 h en precultivos de 4 ml de LB a 37° C. Seguidamente se inocularon los precultivos en cultivos de 500 ml de medio SB previamente autoclavados en erlenmeyers de 2L (para asegurar una buena oxigenación) y se crecieron a 37° C con agitación hasta una OD₆₀₀ de 0.8-1. Una vez enfriados los cultivos a 17° C, la expresión se indujo añadiendo 1mM IPTG. A las 62 horas aproximadamente, los cultivos se centrifugaron y el pellet se resuspendió en tampón de lisis A (50 mM citrato de sodio pH6.5, 1.5 M NaCl, 5 mM imidazol, 5 mM β-mercaptoetanol (BME) que contenía inhibidores de proteasas (Complete EDTAfree; Roche), se sonicó en hielo y se centrifugó a 4° C, 20.000g durante 30min.

Todos los pasos posteriores de purificación se realizaron a 4° C, un parámetro importante para el rendimiento final. El sobrenadante, una vez filtrado, se cargó en la columna de afinidad de níquel (HisTrap HP; GE Healthcare) y se lavó con el tampón B (50 mM citrato de sodio pH 6.5, 1 M NaCl, 20 mM imidazol, 5 mM BME), que contiene una baja concentración de imidazol para disminuir la interacción inespecífica de contaminantes con la columna. La proteína se eluyó con un gradiente de tampón C (igual que el tampón B pero con 500 mM de imidazol) que podía ser lineal o en escala según el constructo expresado. El gradiente en escala se utilizó en la purificación de mTERF-NHis, mTERF-CHis y mTERF-ΔN de manera que las impurezas que quedaban retenidas después del lavado se eliminaban con un escalón de 12% de tampón C.

Las fracciones que contenían la proteína de interés se juntaban, se concentraban mediante Centricon YM-10 filters (Millipore) hasta un volumen final de 2ml y se cargaban en una columna de exclusión molecular HiLoad 16/60 Superdex200 (GE Healthcare), previamente equilibrada con el tampón D (50 mM citrato de sodio pH6.0, 700 mM NaCl, 10 mM BME).

El seguimiento de la purificación se realizó mediante electroforesis, cargando entre 5-15 μ l de las fracciones eluidas en geles de poliacrilamida-SDS.

2.4 Caracterización de la proteína

Técnicas electroforéticas

Electroforesis en gel de poliacrilamida-SDS

La electroforesis desnaturizante en geles de poliacrilamida-SDS (Laemmli, 1970) permite la separación de proteínas según su masa molecular. Además, ajustando el % de acrilamida se puede variar el rango de peso molecular donde la separación es más efectiva. En general, esta técnica permite, mediante el uso de marcadores (proteínas estándares de diferente peso molecular), determinar la masa molecular de la proteína de interés ya que existe una relación lineal entre el desplazamiento de las proteínas en el gel y el logaritmo del peso molecular (PM) de éstas. No obstante, las proteínas con pls extremos presentan movilidades anómalas.

Esta técnica se utilizó de forma rutinaria para analizar la expresión de la proteína de interés, para el seguimiento de la purificación así como para determinar su grado de pureza con respecto a la presencia de otras proteínas contaminantes o a la presencia de degradaciones de la proteína de interés. Los geles utilizados fueron del 12% de acrilamida.

Para la visualización de las proteínas se realizó una tinción con azul de comassie o tinción en plata, cuyo protocolo se describe a continuación.

Tinción en plata

Esta tinción es mucho más sensible que la tinción con comassie y permite visualizar también el DNA.

Una vez corrido el gel, se fija en metanol 50% durante 1 h como mínimo. Se decanta el metanol y se añade la mezcla de tinción. La mezcla de tinción se prepara previamente mezclando en un vaso de precipitados 40ml de H₂O, 1.4 ml de amoníaco al 39% y 190 µl de 10 M NaOH. En otro vaso se mezclan 0.8g de nitrato de plata y 4ml de H₂O y se añade muy lentamente a la mezcla anterior en agitación. Se tiñe 15 min en agitación.

Después de retirar la solución de tinción y lavar el gel con abundante agua miliQ se procede al revelado. La solución de revelado se prepara mezclando 75 µl de formaldehído al 36-38%, 35.7 µl de ácido cítrico 1M y agua hasta 150 ml. Añadimos un poco de solución de revelado, agitamos y la desechamos. Añadimos el resto y manteniendo el gel en agitación observamos el revelado hasta que se alcance la señal adecuada. Para parar la reacción se añade una solución de 50% metanol y 12% acético.

Una vez parada la reacción, para poder secar el gel, se debe incubar en agua aproximadamente 1h.

Así pues, al cargar en el gel las fracciones eluidas de la columna de purificación podemos observar si estas contienen proteína y DNA: la proteína desnaturalizada aparece como una banda con movilidad según su PM y el DNA aparece como una mancha marrón-amarilla en el frente del gel.

Cromatografía de exclusión molecular

Como las técnicas electroforéticas, la cromatografía de exclusión molecular proporciona también un método para estimar el peso molecular o el tamaño y el estado de oligomerización de la proteína de interés bajo unas condiciones concretas de pH, fuerza iónica y temperatura. El parámetro que rige la elución en GF de una sustancia depende de su radio hidrodinámico. Mediante la calibración de la columna con estándares (proteínas globulares), se puede relacionar el

Identificación de proteínas por huella peptídica (Digestión-MALDI/fingerprinting).

Dada su simplicidad, la exactitud de la masa, su alta resolución y gran sensibilidad, la espectrometría de masas MALDI-TOF es la más empleada para identificar proteínas. El espectro de masas obtenido tras la digestión de una muestra, no demasiado compleja, permite obtener lo que se denomina “mapa peptídico”.

Mediante el método denominado “peptide mass fingerprinting” las proteínas se identifican por comparación de una lista de masas peptídicas experimentales con la lista de masas peptídicas calculadas de todas las entradas de una base de datos.

Esta técnica se utilizó para comprobar que la proteína sobrepresada correspondía a mTERF.

Determinación de peso molecular (MALDI-TOF).

El peso molecular de la proteína o de ciertos fragmentos proteolizados fue determinado mediante esta técnica en el servicio de proteómica del parque científico de Barcelona (PCB) y de la Vall d’Hebron. Este método permite determinar el peso molecular de proteínas enteras o péptidos con bastante precisión y sensibilidad.

Degradación de Edman

La degradación de Edman, desarrollada por Pehr Edman, es un método de secuenciación de aminoácidos en un péptido. En este método, el residuo amino-terminal se etiqueta y se separa del péptido sin afectar a los enlaces peptídicos entre los otros residuos. Esta técnica fue utilizada para determinar exactamente el extremo N-terminal de los constructos expresados y en ciertos casos degradados. Se utilizó el servicio de proteómica del PCB y de la Vall d’Hebron.

Espectrometría ESI-MS/MS (trampa iónica) Análisis de secuencia

En esta técnica, los péptidos obtenidos de la digestión de proteínas (generalmente procedentes de muestras líquidas) son separados por cromatografía.

Una vez separados los péptidos se analizan en un espectrómetro de masas tipo Trampa de Iones con fuente de ionización de electrospray (ESI) o electronebulización. En este caso, se obtienen espectros de fragmentación de cada una de las masas que detecta el espectrómetro de masas.

Estas técnicas permiten el análisis en tiempo real de péptidos individuales presentes en la mezcla mediante la fragmentación de los mismos y la obtención de espectros de fragmentación o espectros MS/MS de manera mucho mas efectiva que en el análisis por MALDI-TOF/TOF.

Este tipo de análisis fue realizado por la unidad de proteómica de la Vall d'Hebron con las diferentes formas proteolizadas de la proteína de interés.

Proteólisis limitada

En cristalografía, cuanto más compacta y estable sea la especie a cristalizar, más probable será obtener cristales que difracten a una resolución aceptable gracias a un empaquetamiento más efectivo. Mediante la proteólisis limitada se pueden aislar subdominios estables. Es posible también, obtener evidencias de la interacción de una proteína con otra corriendo dos experimentos de proteólisis idénticos, uno con la proteína y el compuesto con el que interacciona y el otro sin. Si la velocidad de digestión de la proteína de interés es menor en presencia de otra proteína o compuesto, se puede concluir que existe interacción.

En estos ensayos se deben tener en cuenta las variables siguientes:

- la cantidad de proteína para hacer el ensayo
- la cantidad y el tipo de proteasa: la concentración de proteasa debe ser en masa, unas 1000 veces menor que la de proteína. Dependerá también del tiempo de reacción. Es interesante la primera vez utilizar una batería de proteasas para probar un amplio rango de especificidad.
- el volumen de reacción: según las alícuotas que queramos tomar para analizar la proteólisis según el tiempo.
- el tiempo,
- el tamaño del gel de poliacrilamida y el método de tinción, comassie o plata según la cantidad de proteína.

- La temperatura, que influye en la actividad enzimática. Normalmente se hacen ensayos a temperatura ambiente y a 4°C.

Se llevaron a cabo para este fin, estudios de proteólisis limitada con mTERF-ΔN utilizando diferentes enzimas; proteinasa K, papaína, tripsina y quimiotripsina (Tabla 3.6), y en el caso de la papaína se realizaron experimentos con el complejo mTERF-ΔN y DNA.

Los ensayos se realizaron en un volumen final de 100μl con una concentración final de mTERF-ΔN de 3 mg ml⁻¹ a 37° y 20° C, ajustándose las condiciones de concentración y tiempo para cada enzima.

Se probaron dos concentraciones finales de 0.02 mg ml⁻¹ y 2 μg ml⁻¹ para cada enzima, se tomaron muestras a diferentes tiempos (5 min, 30 min, 1 h, 2 h, 5 h y o/n) y se cargaron en un gel desnaturante de SDS con el objetivo de poder seguir la proteólisis.

Una vez establecidas las condiciones, al final de la proteólisis se debe añadir el inhibidor correspondiente (Tabla 2.6).

Enzima	Tampón (10x)	Inhibidor
Proteinasa K	0.3 M TRIS-HCl pH 8.0	
Papaína	1 M fosfato potásico pH 7.5, 25 mM EDTA, 30 mM BME	Iodoacetamida 5.4mM
Tripsina	0.45 M TRIS-HCl pH 8.0	Pefablock 1mM
Quimiotripsina	0.8 M TRIS-HCl pH 7.8, 1 M CaCl ₂	Pefablock 1mM

Tabla 3.6_ Condiciones de la proteólisis limitada para cada tipo de enzima.

2.5 Expresión y purificación con selenometionina

Los derivados de SeMet de los constructos mTERF-NHis, mTERF-CHis y mTERF-ΔN fueron producidos en células no autótrofas de *E. coli* BL21 (DE3) creciéndolas en medio mínimo de cultivo al que se le añade selenometionina. Para la purificación y cristalización se siguió el mismo protocolo que para la proteína nativa.

Soluciones utilizadas

Solución de aminoácidos 250 mg ml⁻¹ (5x)

Se disuelven en 80 ml de H₂O miliQ 100 mg de cada uno de los aminoácidos Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Phe, Pro, Ser, Thr y Val.

A la vez se disuelven 100 mg de Tyr y Trp en 300 ml de H₂O tamponada con fosfato. Se debe añadir lentamente HCl 1M mientras se mantiene la solución en agitación hasta la disolución completa de los aminoácidos.

Una vez preparadas las dos soluciones se mezclan y se enrasan a 400 ml.

La solución se debe esterilizar mediante filtración (0.22mm).

Solución de aminoácidos represores 250 mg ml⁻¹ (5x)

Para 6 ml de solución, se añaden 100 mg de cada uno de los aminoácidos Thr, Lys y Phe; y 50 mg de Ile, Leu y Val. Se esteriliza por filtración.

Solución 20% (p/v) de glucosa

Esterilizada por filtración

Solución 1 M de MgSO₄

Esterilizada mediante autoclave.

Solución 1 M de CaCl₂

Esterilizada mediante autoclave.

Solución de tiamina 10 mg ml⁻¹

Esterilizada por filtración.

Solución biotina 5 mg ml⁻¹

Para que la biotina se disuelva se debe mantener la solución en agitación y añadir lentamente NaOH 1 M hasta la completa disolución. Enrasar y esterilizar por filtración.

Solución de SeMet 60 mg ml⁻¹

Esterilizada por filtración. En caso de no conseguir una disolución completa de la SeMet, añadir alguna gota de 10 M NaOH.

Medio M9 (10x)

Medio de cultivo mínimo. Para 1L de cultivo se deben pesar 80 g de Na₂HPO₄, 40 g de KH₂PO₄, 5 g de NH₄Cl y 5 g de NaCl. Se esteriliza mediante autoclave.

Medio TES (100x)

Para preparar 1L de medio se deben disolver 5 g de EDTA en 800 ml de H₂O miliQ. Una vez ajustado el pH a 7.5 (para garantizar la disolución de las demás sales) se añaden 0.83 g de FeCl₃·6H₂O, 84 mg de ZnCl₂, 13 mg de CuCl₂·2H₂O, 10 mg de CoCl₂·6H₂O, 10 mg de H₃BO₃ y 1.6 mg de MnCl₂·6H₂O. Se esteriliza por filtración.

Protocolo

Preparación del medio de cultivo (600ml)

Se añade en condiciones estériles a un erlenmeyer de 2L previamente autoclavado con 400 ml de H₂O miliQ:

- 120 ml de la solución de aminoácidos (5x)
- 60 ml de medio M9 (10x)
- 6 ml de medio TES (100x)
- 12 ml de solución glucosa 20%
- 0.6 ml MgSO₄ 1M
- 0.18 ml de CaCl₂ 1M
- 1.2 ml de biotina 5 mg ml⁻¹
- 0.6 ml de tiamina 10 mg ml⁻¹
- 0.12 ml SeMet 50 mg ml⁻¹
- Antibiótico

Una vez preparado, se inocula cada cultivo con 4 ml de un precultivo o/n crecido en LB. Los cultivos se crecen a 37° C con agitación hasta una OD₆₀₀=0.5-0.6. Se dejan enfriar y se añade en condiciones de esterilidad:

- 3.6 ml de la solución de aminoácidos represores
- 0.6 ml de la solución de SeMet a 50 mg ml⁻¹
- 6 ml de glucosa 20%
- 60 ml de la solución de aminoácidos (5x)
- 0.6 ml de biotina 5mg ml⁻¹
- 0.3 ml de tiamina 10 mg ml⁻¹

Los cultivos se incuban 30 min más a 37° C, se dejan enfriar y se inducen añadiendo 1mM IPTG a 17° C. El tiempo de expresión es aproximadamente 62 horas.

2.6 Mutaciones puntuales por Met

Como se explica posteriormente, en cristalografía, para poder utilizar los métodos de faseado experimental en la resolución de la estructura de la proteína de interés, es necesario tener cristales con átomos pesados que tengan difracción anómala en una longitud de onda medible. En este caso se utilizó la estrategia de las mutaciones puntuales para incrementar el número de metioninas y aumentar así el poder de faseado.

Para escoger los lugares de mutación, se alinearon las secuencias proteicas de los homólogos en rata, ratón y cerdo y según la presencia de Met en la secuencia de éstos, se eligieron tres lugares de mutación repartidos en la proteína.

Para llevar a cabo las mutaciones, se utilizó el QuikChange Site-Directed Mutagenesis Kit (Esquema Fig 3.7).

a.

Mutación	Secuencia primers
Mut_Met1	5' GAT CTG TGG AGA AAG ATT ATG ACA TCA GAC CTT GAA ATT G 3' 5' C AAT TTC AAG GTC TGA TGT CAT AAT CTT TCT CCA CAG ATC 3'
Mut_Met2	5' GAA TTT TTG CAG GCA GCC GGT ATG TCA TTG GGT CAC AAT G 3' 5' C ATT GTG ACC CAA TGA CAT ACC GGC TGC CTG CAA AAA TTC 3'
Mut_Met3	5' GTC TTA AGC TAT CCA GAT ATG ATC TTC TTG GCA GAG 3' 5' CTC TGC CAA GAA GAT CAT ATC TGG ATA GCT TAA GTC 3'

b.

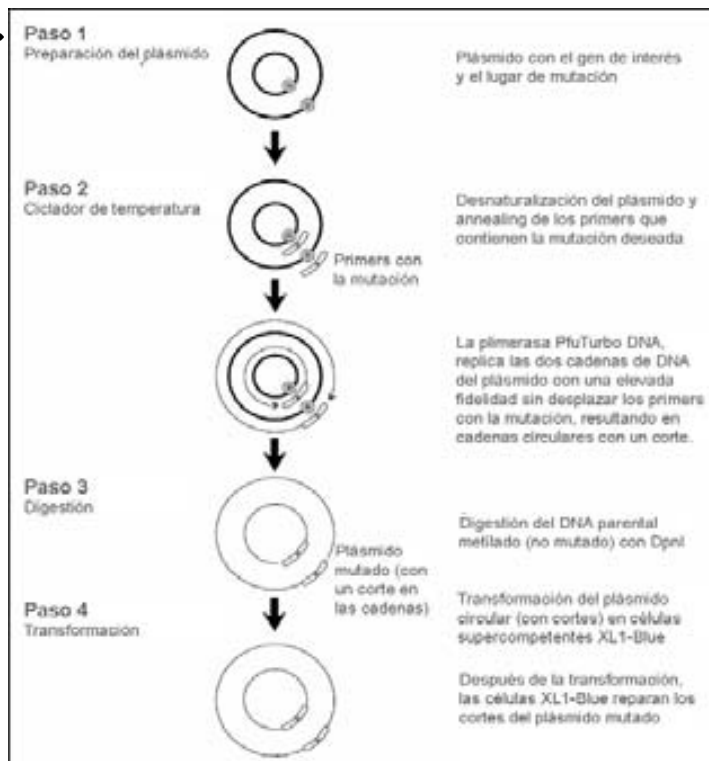


Fig 3.7_ (a) Secuencia de los primers (forward y reverse) diseñados para los tres lugares de mutación. (b) Esquema del protocolo de mutagénesis.

3_ Formación del complejo con DNA

3.1 Ensayos de unión a DNA para la cristalización

Para comprobar la unión al DNA de las proteínas recombinantes se utilizó la técnica de *electrophoretic mobility shift assay* (EMSA) o gel de retardo, en el que la muestra se corre en un gel en condiciones nativas y se visualiza el DNA. La capacidad de unión de la proteína se confirma si se observa un retardo en la banda del DNA al añadir la proteína a la mezcla. Se realizaron geles de retardo con un oligonucleótido de 28pb que contenía la secuencia de terminación (Kruse *et al.*, 1989) y cantidades crecientes de mTERF-NHis, mTERF-CHis y mTERF- Δ N. Además, se realizaron varios estudios con el constructo mTERF- Δ N mediante geles de retardo para determinar las condiciones de pH, sal y longitud de DNA óptimas para la formación del complejo.

Las muestras que contenían proteína y DNA se dializaron según las condiciones de estudio en botones de diálisis de 15-20 μ l (Hampton). Una vez dializadas, se cargan en un gel nativo para analizar el retardo de los complejos (se detalla un poco más adelante).

Estudio de la longitud del DNA

Los complejos de mTERF con los oligonucleótidos de 12, 13, 15, 17, 19, 21, 27, 28, 31 pb (Tabla 3.8) se dializaron contra una solución tamponada de HEPES pH 7.4, 200 mM NaCl, 10 mM BME.

Estudio de la concentración de sal

El complejo de mTERF con el oligonucleótido de 21 pb fue dializado contra diferentes soluciones tamponadas con HEPES pH 7.4 y con concentraciones de 25, 50, 75, 100, 125, 150, 175, 200, 250, 500 mM de NaCl.

Estudio del pH

Los complejos de mTERF con los oligonucleótidos de 12, 15, 21, 27, 29 pb fueron dializados contra diferentes soluciones tamponadas de 200 mM NaCl, 10 mM BME: acetato sódico pH 4.5, citrato sódico pH 5.0, citrato sódico pH 5.5, MES pH 6.0, fosfato pH 6.5, Hepes pH 7.0 y Hepes pH7.4

3.2 Obtención del complejo para la cristalización

Previamente a la formación del complejo con DNA, se debe hacer un annealing de los oligonucleótidos (Biomers), que requiere mezclarlos en 20 mM KCl, calentarlos 20min a 80° C y dejarlos enfriar muy lentamente para que se forme la doble cadena.

Oligonucleótidos mayores de 23bp

Los primeros oligonucleótidos que se probaron en la cristalización del complejo fueron de 27pb, 28pb, 29pb, 30pb y 31pb con extremos romos o cohesivos, dependiendo del caso (Tabla 2.8).

La proteína a 0.5mg ml⁻¹ en tampón (50mM HEPES pH7.4, 500mM NaCl, 10mM BME) y el DNA (0.5mM) se mezclaron en relación molar 1:1 y la muestra se dializó contra el tampón (50mM HEPES pH7.4, 200mM NaCl, 10mM BME).

Una vez dializado, el complejo se concentró y se pasó por una columna de exclusión molecular (Superdex200). Se chequearon las fracciones de elución mediante electroforesis en geles de poliacrilamida-SDS con tinción plata o por retardo en un gel nativo, y aquellas que contenían el complejo se concentraron hasta 7mg ml⁻¹ para la posterior cristalización.

Oligonucleótidos menores de 23bp

En este caso, a diferencia de los oligonucleótidos mayores de 23pb, el complejo precipitaba al intentar dializarlo a concentraciones menores de sal. Así pues, la muestra utilizada para la cristalización se obtenía mezclando directamente la proteína (8mg ml⁻¹) en tampón (50mM citrato de sodio pH6.0, 700mM NaCl, 10mM BME) y el DNA (0.5mM) en una relación molar 1:2.

	10pb	11pb	12pb
cohesivo	5' CGGGCTCTGC 3' 3' GGCCCGAGAC 5'	5' CGGGCTCTGCC 3' 3' GGCCCGAGACG 5'	5' CGGGCTCTGCCc 3' 3' GGCCCGAGACGG 5'
romo	5' CGGGCTCTGC 3' 3' GGCCCGAGACG 5'	5' CGGGCTCTGCC 3' 3' GGCCCGAGACG 5'	5' CCGGGCTCTGCC 3' 3' GGCCCGAGACGG 5'
	13pb	14pb	15pb
cohesivo	5' CCGGGCTCTGCCA 3' 3' TGGCCCGAGACGG 5'	5' CCGGGCTCTGCCAT 3' 3' AGGCCCGAGACGGT 5'	5' TACCGGGCTCTGCCA 3' 3' TGGCCCGAGACGGTA 5'
romo	5' CCGGGCTCTGCCA 3' 3' GGCCCGAGACGGT 5'	5' ACCGGGCTCTGCCA 3' 3' TGGCCCGAGACGGT 5'	5' ACCGGGCTCTGCCAT 3' 3' TGGCCCGAGACGGTA 5'
	16pb		17pb
cohesivo	5' TACCGGGCTCTGCCAT 3' 3' TGGCCCGAGACGGTAA 5'		5' TACCGGGCTCTGCCATC 3' 3' TGGCCCGAGACGGTAGA 5'
	25pb		27pb
cohesivo	5' GATTACCGGGCTCTGCCATCTTAAC 3' 3' GCTAATGGCCCGAGACGGTAGAATT 5'		5' CGATTACCGGGCTCTGCCATCTTAACA 3' 3' CGCTAATGGCCCGAGACGGTAGAATTG 5'
	31pb		29pb
cohesivo	5' TTGCGATTACCGGGCTCTGCCATCTTAACAA 3' 3' ACGCTAATGGCCCGAGACGGTAGAATTGTTA 5'		5' TGCATTACCGGGCTCTGCCATCTTAACA 3' 3' CGCTAATGGCCCGAGACGGTAGAATTGTA 5'
	19pb		21pb
romo	5' TTACCGGGCTCTGCCATCT 3' 3' AATGGCCCGAGACGGTAGA 5'		5' ATTACCGGGCTCTGCCATCTT 3' 3' TAATGGCCCGAGACGGTAGAA 5'
	28pb		
romo	5' GCGATTACCGGGCTCTGCCATCTTAACA 3' 3' CGCTAATGGCCCGAGACGGTAGAATTGT 5'		

Tabla 3.8_ Oligonucleótidos utilizados en los ensayos cristalográficos

3.3 Análisis y caracterización del complejo

Cromatografía de exclusión molecular

Como se ha indicado anteriormente, la cromatografía de exclusión molecular proporciona también un método para determinar el peso molecular y por tanto, en este caso también para identificar la estequiometría del complejo según el V_e de éste.

Técnicas electroforéticas

Geles nativos para analizar el retardo de los complejos

Las muestras para cargar en el gel se preparan mezclando 5-10 μ l del complejo previamente dializado o de las fracciones eluidas de la columna de purificación con DNA *dye loading buffer* 6x.

El gel nativo de dimensiones 8cm x 10cm x 1.5mm se prepara mezclando 30ml de H₂O, 7.33ml solución A (29:1 acrilamida, bisacilamida), 2ml

TAE 10x, 100 μ l de PSA 10% y 80 μ l de TEMED y se debe pre-correr a 4° C durante unos 20min en tampón TAE 0.5. Una vez cargadas las muestras, se corre a 4° C y 200V durante aproximadamente 25 minutos (hasta que la banda azul haya corrido $\frac{3}{4}$ partes del gel), se tiñe con SyBr green (Invitrogen) y se visualiza con luz UV. Se puede observar el DNA libre y el que está formando complejo con la proteína que aparece como una banda retardada. Se puede teñir con comassie para verificar que la banda retardada contiene proteína.

Esta técnica se puede utilizar para determinar la estequiometría aproximada del complejo manteniendo constante la cantidad de DNA y aumentando la proporción de proteína evaluando cuál es la concentración de proteína necesaria para retardar todo el DNA.

3.4 Ensayos de especificidad

En estos ensayos, los geles utilizados son nativos de composición igual que los descritos anteriormente y de dimensiones 18cm x 20cm x 1mm de grosor.

Para los ensayos de especificidad de los diferentes constructos, las muestras se preparan mezclando 3.2 μ l de tampón (125mMHEPES pH7.4, 500mMNaCl, 25mM MgCl₂), 1 μ l de poly dI-dC, 1 μ l de DNA_{TER} 28pb (relación 1:1 con la proteína) marcado con un fluoróforo FAM 494nm-520nm (AlexaFluor 488) y H₂O hasta un volumen final de 15 μ l. Añadimos 1 μ l de proteína (1mg ml⁻¹ en tampón D y 10% glicerol), e incubamos en hielo durante 20 minutos. Después se añaden 3.5 μ l de tampón de carga 6x (preparado: 60% de glicerol y una punta de espátula de azul de bromofenol).

Se debe hacer un primer ensayo para determinar la cantidad óptima de poly dI-dC y para confirmar qué constructos unen DNA. Se probaron cantidades crecientes de 100, 250, 500, 1000ng y se determinó que la cantidad óptima para los siguientes ensayos era 500ng.

El gel se pre-corre a 200V durante 20-30min a 4° C en TAE 0.5x. Una vez cargadas las muestras se corrian en las mismas condiciones durante 1.5-2h. Se visualizaban directamente las bandas de DNA con luz UV.

Para los ensayos de competición manteníamos 500ng de poly dI-dC y añadíamos cantidades crecientes del oligo competidor no marcado (1:0, 1:0.5, 1:1, 1:5), manteniendo un volumen final de 20µl.

4_Cristalización

La obtención de cristales cuya difracción por rayos X dé una resolución adecuada, es el paso limitante de la cristalografía de proteínas, y sin éstos es imposible obtener datos estructurales a alta resolución. Para tratar de obtener cristales se realizan ensayos cuyo objetivo es provocar una sobresaturación de la solución de proteína para que se dé una precipitación ordenada en forma de cristal. Como las variables que dan lugar a la cristalización son numerosas (pH, concentración de precipitante, temperatura, concentración de proteína...) es extremadamente difícil poder predecir las condiciones apropiadas para la nucleación o el crecimiento de cristales ordenados (Rupp, 2004). Por esa razón, se determinan mediante un proceso de criba de prueba y error en el cual se exploran numerosas combinaciones de las variables mencionadas anteriormente (*screenings*).

La cristalización consta de dos fases, la nucleación y el crecimiento del cristal. En la nucleación las moléculas de proteína dispersas en el solvente empiezan a formar pequeños clusters, que empiezan a ser estables -no se redisuelven- a partir de un tamaño crítico. Éstos forman los núcleos de cristalización. En esta fase, la disposición de una manera ordenada de las moléculas de proteína determina la estructura de la red cristalina. La aparición de los cristales se debe al crecimiento de estos núcleos.

Las condiciones óptimas para la nucleación se dan en niveles de sobresaturación elevados, mientras que el crecimiento del cristal que debe ser lento y ordenado, se favorece a niveles de saturación menores.

Diagrama de fases y métodos de cristalización

El proceso de cristalización se puede ilustrar gráficamente mediante un diagrama de fases. Estos diagramas constituyen la base para controlar el sistema de cristalización de una proteína.

En la cristalografía de proteínas existen cuatro métodos estándares esquematizados en la (Fig 3.9): microbatch, difusión de vapor, diálisis y difusión líquido líquido. En este trabajo se utilizó básicamente la difusión de vapor -gota sentada- por ser el método más habitual y reproducible.

Difusión de vapor

En esta técnica se establece un equilibrio de vapor entre una gota que contiene la proteína y la solución de cristalización y un reservorio líquido que contiene solamente la solución de cristalización. Se produce un intercambio de H_2O entre ambas soluciones y, en el caso de ciertas condiciones también de especies volátiles, hasta que la presión de vapor entre la gota y el reservorio se iguala. La concentración del agente precipitante es mayor en el reservorio, ya que en la gota se mezcla con la solución acuosa de proteína, así que se produce un desplazamiento de agua de la gota al reservorio, disminuyendo el volumen de la gota y provocando por tanto la concentración de sus componentes. Se produce entonces la sobresaturación de la proteína que precipita de forma más o menos amorfa y, si tenemos suerte, da lugar a la formación de cristales. Otro factor a tener en cuenta es que la técnica de difusión de vapor es sensible a las variaciones de temperatura. Las placas, una vez dispensadas, se guardan en cámaras de temperatura constante a $20^{\circ} C$ o $4^{\circ} C$.

Para esta técnica, existen dos estrategias, gota colgante o gota sentada (Fig 3.9 b y c). En el sistema de gota colgante la gota con la mezcla de proteína y solución de cristalización se deposita sobre un cubreobjetos, que puede ser de cristal previamente tratado con diclorometilsilano para impedir deformaciones y desplazamiento de la gota o de plástico (Hampton), y se sella con silicona encima del reservorio. En el sistema de gota sentada, la mezcla se deposita en un pequeño reservorio y se cierra mediante cinta

óptica adhesiva transparente. El volumen de la gota y del reservorio depende de las placas utilizadas.

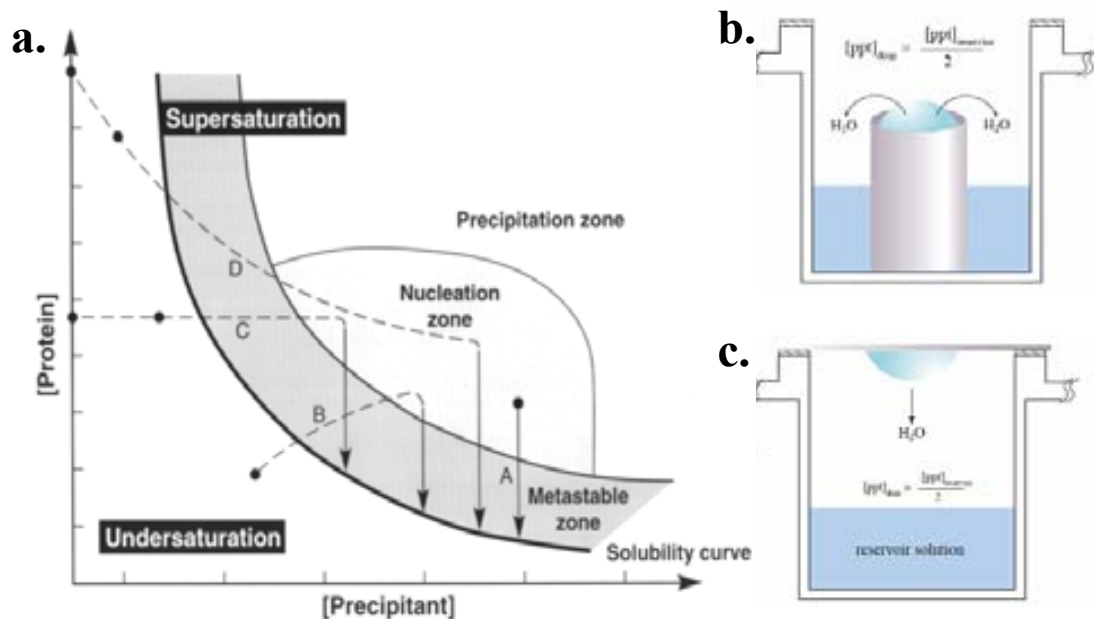


Fig 3.9_ (a) Esquema de un diagrama de fases de cristalización de una proteína en función de dos de los parámetros más comúnmente variados, la concentración de proteína y precipitante. El círculo negro representa la condición inicial y la línea representa la evolución en el diagrama de fases de cada uno de los cuatro métodos mayoritarios de cristalización; (A) cristalización en batch, (B) difusión de vapor, (C) diálisis, (D) difusión líquido-líquido. En el caso de la diálisis y la difusión líquido-líquido se muestran dos puntos iniciales alternativos, ya que la solución no saturada de proteína puede contener solo proteína o proteína con bajas concentraciones de agente precipitante. Dentro del método de cristalización mediante la difusión de vapor, existen dos técnicas (b) Esquema de la técnica de gota sentada (la utilizada mayoritariamente en esta tesis). (c) Esquema de la técnica de gota colgante.

Determinación de las condiciones de cristalización

Los screenings iniciales para los constructos mTERF-NHis, mTERF-CHis y mTERF- Δ N y para el complejo mTERF- Δ N con DNAs mayores de 25pb se realizaron a mano según el método de difusión de vapor en placas de 24 pocillos mediante la técnica de gota sentada (Hampton) mezclando 1 μ l proteína+ 1 μ l de solución precipitante en la gota de cristalización y dispensando 500 μ l de solución precipitante en el reservorio.

Los screenings iniciales para los complejos de mTERF- Δ N con DNAs menores de 25pb y mTERF-NHis o mTERF-CHis con diferentes DNAs se realizaron en la Plataforma automatizada de cristalografía (PAC) del PCB. El método es el mismo que el anteriormente descrito, pero en este caso se utilizaron placas de 96x2 pocillos. Las soluciones del reservorio se prepararon y dispensaron con el robot

Freedom EVO (Tecan) y, el dispensado de la proteína y el reservorio en la gota de cristalización (200nl finales) fue realizado mediante el robot Cartesian nanodrop (Genomic Solutions).

Las técnicas robóticas de cristalización, con respecto a las manuales, permiten utilizar una cantidad de proteína mucho menor en el *screening* inicial de las condiciones de cristalización. Además, al tratarse de gotas más pequeñas la relación superficie/volumen es mayor y el proceso de cristalización es más rápido. El principal inconveniente es el escalado de nanogota a microgota, ya que no siempre son reproducibles los resultados a mayor escala.

La PAC dispone también de dos *crystal farm* (incubadores automatizados) a 20° y 4° C, donde se guardan las placas dispensadas, selladas e identificadas por un código de barras. Las placas se mantienen a la temperatura deseada y se someten a un análisis fotográfico automático según intervalos de tiempo previamente establecidos por el usuario.

4.1 Optimización de los cristales

Una vez se obtienen cristales en una o varias condiciones se deben escalar, si la gota ha sido dispensada por el robot y optimizar en placas de 24 pocillos.

La optimización consiste en probar condiciones de pH, sal y concentración de precipitante alrededor de la condición en la que se han obtenido los primeros cristales para obtener cristales mejores y más grandes. En este caso, las soluciones del reservorio se prepararon y dispensaron con el robot Freedom EVO (Tecan), utilizándose el programa informático que incorpora el robot para diseñar las placas de optimización. Para una condición de cristalización concreta se puede diseñar una placa de cristalización en base al correspondiente diagrama de fases y dispensarlo con el Tecan y el Cartesian en una placa de 96 pocillos.

Métodos previos a la cristalización

En un experimento de cristalización ideal, una vez se forman los núcleos, la concentración de proteína en la solución disminuye, de manera que el sistema alcanza la zona metaestable (Fig 3.9) donde se produce el crecimiento del cristal.

A la práctica muy a menudo se da la formación de numerosos núcleos de cristalización dando lugar a muchos cristales pequeños y de baja calidad.

En este caso se aplicaron técnicas de *seeding* y la utilización de aceites como estrategias para la optimización de los cristales.

Seeding

Una manera de evitar los problemas de nucleación masiva es añadiendo núcleos preformados a una nueva gota de cristalización en condiciones metaestables (Bergfors, 2003). De esta manera se limita el número de núcleos y se favorece el crecimiento de éstos. Se utilizaron tres técnicas:

Streackseeding

Esta técnica consiste en friccionar los cristales con un pelo limpio de bigote de gato (o similar), que después se desliza por la gota previamente equilibrada.

Microseeding

Esta técnica se puede utilizar para que el crecimiento de los cristales sea más efectivo o incluso para obtener cristales en otras condiciones donde la gota puede estar en la zona metaestable y no observamos la formación de cristales.

Para hacer la solución stock de microseeding, se trituran los cristales con bolitas de vidrio (*seed beads*). Para controlar la cantidad de núcleos que se transfieren a la nueva gota se realizan diferentes diluciones de la solución stock. Es importante la calidad de los cristales utilizados en la solución stock, ya que los defectos de los cristales se acumulan a medida que crecen.

Se añade un volumen muy pequeño (por ejemplo, 0.25 en gotas de 2 μ l) de la mezcla de cristales triturados o de las correspondientes diluciones, a las gotas de cristalización previamente equilibradas o a un *screening* general. En éste último caso, el objetivo es la

obtención de cristales en otras condiciones, en aquellas gotas que estén en la zona metaestable. Esta técnica se puede realizar tanto a mano como con el Cartesian (Walter *et al.*, 2008).

Macroseeding

Esta técnica sirve básicamente para el crecimiento de los cristales. En este caso, se pesca el cristal, se lava cuidadosamente y se deposita en una gota previamente equilibrada. En el caso de mTERF-ΔN con DNA de 29bp con esta técnica los cristales podían llegar a crecer 1/3 de su volumen inicial. Hay que tener mucho cuidado en la manipulación de los cristales para no crear imperfecciones en la superficie de éstos, ya que sirven como núcleo de crecimiento para otros cristales.

Utilización de aceites

La utilización de aceites permite variar la cinética en un experimento de difusión de vapor. La velocidad de difusión depende del tipo y del grosor de la capa de aceite. Se utilizan mezclas de parafina (muy poco permeable) y silicona (muy permeable) para controlar la difusión final. El grosor de la capa suele ser de entre 1 y 3mm. En este caso se utilizó el aceite comercial Al's Oil mezcla 50:50 (v/v) de parafina y silicona, aunque no se obtuvieron mejoras. El aceite se puede dispensar en el reservorio o encima de la gota (parecido a la técnica de *microbatch* en aceite).

Métodos postcristalización

Existen varias técnicas para intentar mejorar la difracción de los cristales una vez formados (Heras & Martin, 2005).

Annealing

La posterior manipulación y congelación de los cristales pueden provocar desórdenes en la red cristalográfica y por tanto un aumento de la mosaicidad y una menor resolución. Este fenómeno es especialmente importante en cristales grandes o con alto contenido de solvente. El annealing puede ayudar en estos casos (Harp *et al.*, 1998; Harp *et al.*,

1999). Esta técnica se puede llevar a cabo de dos maneras; una vez el cristal está montado, se bloquea la corriente de gas criogénica temporalmente (1-2s) o bien se sumerge de nuevo el cristal en el crioprotector (aprox. 3min) y se vuelve a congelar. Es un método fácil y rápido en el que se homogeneiza la concentración de agente crioprotector de manera que cuando el cristal se vuelve a congelar la contracción del solvente y de la red cristalina del cristal es muy parecida.

Deshidratación

El agua juega un papel muy importante en el mantenimiento de la estructura y en la actividad de las proteínas en solución y en el cristal. La reducción del contenido de solvente en un cristal puede favorecer un mayor empaquetamiento de éste aumentando así la resolución del patrón de difracción. Normalmente el contenido de solvente es del 50%, así que la deshidratación suele dar resultado en cristales con un contenido de solvente superior.

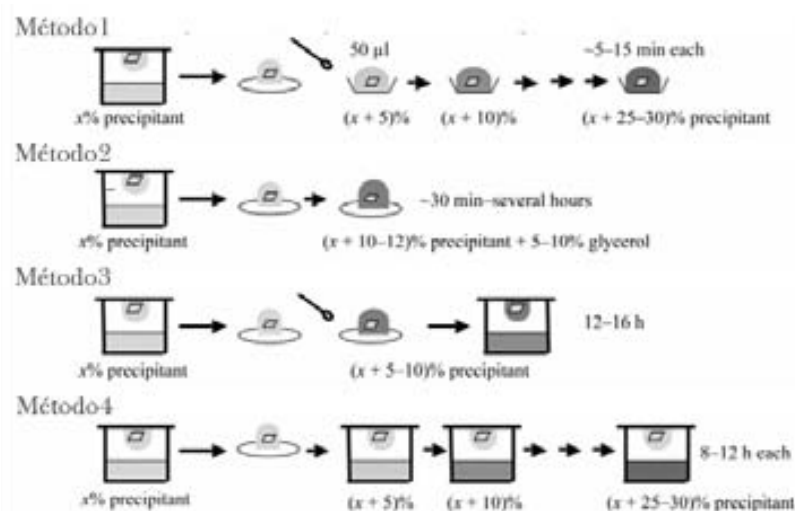


Fig 3.10_ **Método1:** Los cristales se transfieren secuencialmente a gotas de $50\mu\text{l}$ que contengan cantidades crecientes de precipitante. Dependiendo de la estabilidad del cristal la concentración de agente deshidratante se puede incrementar del 5% al 30% y en pasos de 0.5 a 5%. El tiempo de soaking puede variar entre 5-15min a días, en tiempos elevados tenemos que tener en cuenta la deshidratación de la gota y hay que mantenerla equilibrada contra un reservorio. **Método2:** Se añade lentamente la solución deshidratante a la gota (incremento 10-12% precipitante + 5-10% de agente crioprotector) y se deja deshidratar al aire durante unos 30min. **Método3:** El cristal se transfiere a una gota con la solución deshidratante que se equilibra contra un reservorio con la misma solución (incremento 5-10% precipitante) y se incuba durante 12-16h. **Método4:** Una vez formado el cristal, la gota de cristalización se equilibra contra reservorios que contienen concentraciones crecientes de precipitante (aumentos del 5%). El tiempo de incubación para cada paso es de 8-12h.

Existen varios métodos químicos de deshidratación, el esquema de los cuales se muestra en la (Fig 3.10). De estos métodos se utilizó básicamente el 4, y aunque no se obtuvo un aumento en la resolución en la difracción, aumentó la reproducibilidad de la difracción en cristales diferentes.

La plataforma de cristalografía del PCB dispone también del *Proteros Free Mounting System* (Rigaku), un aparato que permite controlar la humedad relativa del cristal de una manera más reproducible y eficaz. Este sistema permite realizar protocolos de deshidratación-rehidratación recogiendo para cada % de humedad el patrón de difracción, pudiendo determinar cuál es la humedad que dé un patrón de difracción de mayor resolución.

Crosslinking

Se utiliza en cristales frágiles para hacerlos más resistente durante los cambios de solución, en la congelación o en el manejo. Se usó glutaraldehído como agente entrelazante que actúa principalmente sobre los residuos de lisina.

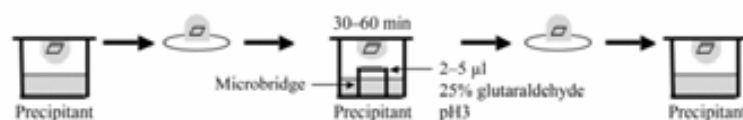


Fig 3.11_ Esquema del protocolo para el crosslinking. Se utiliza un micropuente donde se dispensa el glutaraldehído para poder incubar la gota colgante con el cristal durante 30-60min.

4.2 Análisis de la presencia de DNA en los cristales

Se utilizaron varias técnicas para la comprobación de la presencia de DNA en los cristales.

Electroforesis en geles de poliacrilamida-SDS

Se pescan varios cristales, se lavan bien con la solución del reservorio y se cargan en un gel de poliacrilamida-SDS. La tinción puede ser con plata o con Sybr safe, ambas permiten la visualización del DNA.

Cristalización con DNA marcado con un cromóforo

En la cristalización del complejo mTERF-ΔN con DNA 29pb, se compró una de las hebras del DNA con el cromóforo Cy5 (max. absorción 649nm, max. emisión 670nm) (Georgescu *et al.*, 2008). Se realizó el *annealing* del DNA y se siguió el mismo procedimiento que el descrito en el apartado de obtención del complejo para la cristalización. Se cristalizó en las mismas condiciones que el complejo “nativo”. Si los cristales contienen DNA se deben ver de color más intenso con respecto al de la solución –en este caso azul- ya que la concentración de DNA en el cristal es mayor que en la solución.

Detección fluorescente de ácidos nucleicos en cristales

Es un ensayo específico y altamente sensible para la detección de ácidos nucleicos en cristales de complejos macromoleculares. Se basa en la incubación de los cristales con un fluoróforo que une específicamente ácidos nucleicos y la posterior observación de éstos mediante un microscopio de fluorescencia (Kettenberger & Cramer, 2006).

Para asegurar la completa eliminación de los ácidos nucleicos que puedan estar asociados no específicamente, los cristales son lavados y seguidamente incubados en la solución de cristalización a 20° C o/n.

Como control se utilizaron cristales de lisozima crecidos en presencia de DNA y se siguió el mismo protocolo.

Se añadió Sybr green a concentración final 1x a ambas soluciones y se incubó 1 hora. La unión de este fluoróforo a ácidos nucleicos incrementa considerablemente su fluorescencia mientras que las proteínas no muestran esencialmente ningún efecto (Cosa *et al.*, 2001). La fluorescencia máxima se alcanza aproximadamente en una hora, e incubaciones prolongadas hasta 12h no varían significativamente la señal.

Las imágenes fueron tomadas en el servicio de microscopía del PCB. Las condiciones de medida se realizaron con una lente objetivo UPLAPO 20x, PMT Voltaje Ch1= 301 y apertura confocal de 800.0 μm.

4.3 Derivatización de los cristales con átomos pesados

La obtención de cristales derivados con átomos pesados es imprescindible para la utilización de los métodos de faseado experimental que se explican con detalle más adelante.

Para la derivatización existen dos tipos de técnicas.

Tratamientos precristalización

Se pueden obtener cristales derivados mediante la producción de proteína con SeMet (explicado anteriormente), la cocrystalización y en el caso de la cristalización de complejos, se pueden usar DNAs derivados con bromuro o yoduro (Biomers).

La cocrystalización consiste en crecer los cristales en presencia del compuesto que contiene el átomo pesado, que se puede añadir a la solución proteica o en la gota de cristalización. Durante el crecimiento del cristal el compuesto es incorporado en la red cristalina.

Se pueden utilizar DNAs derivados con haluros, para que estéricamente sea lo más equivalente posible, se sustituye el grupo 5-metilo de la timina por un bromo o un yodo (Fig 3.12). Los derivados de yodo, aunque la señal sea menor, pueden ser medidos en un difractor de ánodo rotatorio (PAC).

El DNA halogenado en dos y tres posiciones se compró en biomers, y se probó con los cristales de mTERF-ΔN-29pb y mTERF-ΔN-15pb. En éste último caso se obtuvieron unos cristales grandes y muy bonitos pero que no difractaban. Y los cristales de mTERF-ΔN-29pb con Br no tenían señal anómala.



Fig 3.12_ La halogenación se produce en el carbono 5 del uracilo (sustituyendo el H por un yodo o bromo) resultando en 5-Iodo/Bromo-2'-deoxy-uridina.

Tratamientos postcristalización

El soaking consiste en, una vez obtenidos los cristales, añadir en la gota la solución el compuesto que contiene el átomo pesado o los transferirlos a una solución de condiciones similares a las de cristalización que contiene el átomo pesado.

En este experimento es importante la concentración de átomo pesado (10-0.1mM) y el tiempo de *soaking* (30min-varios días), el pH, la concentración de sal y la temperatura, que deben ser determinados para cada compuesto. Después del *soaking* los cristales se lavan 1min en la solución de cristalización sin el átomo pesado para eliminar el exceso de éste (*back-soaking*). Existe otro protocolo *soaking* donde los cristales se sumergen tiempos muy cortos (10min-2h) en soluciones del átomo pesado próximas a la saturación (Sun *et al.*, 2002).

Existen muchos factores que determinan la incorporación del átomo pesado en el cristal, así pues, aunque hay ciertas condiciones que favorecen la incorporación de cierto tipo de átomo pesado (Garman & Murray, 2003; Agniswamy *et al.*, 2008), a la práctica sólo lo podemos determinar mediante ensayo.

Se realizaron *soackings* a diferentes tiempos y concentraciones con varios compuestos de Hg, Au y Pt, y con algún compuesto de Sm, Os e Yb para mTERF-ΔN con DNA 29pb, y derivatizaciones mediante cortos *crio-soackings* con haluros (Dauter *et al.*, 2000), aunque ninguno dio señal anómala.

Con los cristales de mTERF-NHis- DNA 15pb se realizaron *soackings* con compuestos de Hg, Au, Pb y Pt. En este caso se obtuvo señal anómala con el derivado de Au y el de benzoato de Hg.

4.4 Crioprotección y montaje de los cristales

Crioprotección

La interacción entre los rayos X y la materia sólida es un fenómeno complejo. A parte de la dispersión elástica que causa la difracción, los fotones incidentes pueden transferir parte de su energía e ionizar los átomos al interactuar con la materia.

Así pues, los cristales de proteína expuestos a un haz de rayos X sufren daño por radiación que provoca que los datos de difracción pierdan progresivamente resolución y calidad. En el daño por radiación se pueden distinguir el daño primario, debido a la fotoionización de los átomos del cristal y el daño secundario debido a la difusión de estos radicales libres a lo largo del cristal. Este último se reduce considerablemente a temperaturas criogénicas ya que la velocidad de difusión de los radicales disminuye considerablemente.

Las desventajas de la criocristalización son que las condiciones de crioprotección se deben establecer empíricamente y que la mosaicidad del cristal suele aumentar cuando se congela.

Durante la congelación se debe evitar la formación de gel cristalino en el interior de la muestra, porque distorsiona el orden interno del cristal e interfiere en el patrón de cristalización de la proteína. Para este fin se añaden agentes crioprotectores que en la congelación dan lugar a un estado vítreo del solvente.

Como crioprotectores se pueden utilizar:

- los que penetran en el solvente del cristal: como glicerol, MPD, alcoholes, PEGs de bajo peso molecular...
- los que envuelven el cristal: aceites como paratone, parafina y aceite mineral de alta densidad.

La solución crioprotectora normalmente se formula sustituyendo un % de agua de la solución de cristalización por un crioprotector. Para determinar el % se debe difractar la solución crioprotectora congelada, siendo óptima aquella que no presenta anillos de gel en el patrón de difracción. Normalmente la presencia del

crystal hace que debamos aumentar un 5% la cantidad del crioprotector determinada inicialmente.

Como las soluciones de cristalización donde se han obtenido los cristales no eran en ninguno de los casos criogénicas, se siguieron dos métodos de crioprotección.

- Transferencia del cristal a soluciones con concentraciones crecientes de agente criogénico, mayoritariamente glicerol (5, 10, 20% y en algunos casos 30%). El tiempo de incubación en cada solución era de 1min aprox.
- Sustitución gradual en la gota de la solución de cristalización por la solución crioprotectora. Este tratamiento es más suave y consiste en quitar la mayor parte de la solución de la gota e ir añadiendo la solución con concentraciones crecientes de crioprotector.

Montaje de los cristales

Una vez criogenizado, el cristal se pesca con un loop (Fig 3.13) de medida adecuada -lo más parecida al cristal para minimizar el volumen de solución y reducir su dispersión de rayos X- y se congela sumergiéndolo rápidamente en nitrógeno líquido.



Fig 3.13_ (a) Elementos para pescar y guardar el cristal en nitrógeno líquido. En los experimentos de difracción se utilizaron loops de 20 y 10 micras de grosor. (b) Podemos observar el loop montado sobre diferentes bases (cap). De izquierda a derecha: CrystalCap (de rosca), CrystalCap HT (magnético) -usados actualmente en todas las líneas del sincrotrón- y CrystalCap con base de cobre. Este último se utilizó en los experimentos de difracción de Göttingen ya que la difracción duró varios días y esta base ayuda a reducir la formación de hielo durante la recogida. (c) Cristal montado en el loop. Imagen tomada en el sincrotrón de un cristal de complejo mTERF- Δ N con DNA 29pb.

5_ Difracción de rayos X y resolución de la estructura

5.1 Difracción de rayos X

El uso de la radiación electromagnética para visualizar objetos requiere que la radiación tenga una longitud de onda comparable a la distancia más pequeña que se quiera resolver (Fig 3.14). Por tanto la radiación que debe ser utilizada para la resolución de estructuras atómicas debe tener una longitud de onda del orden de la distancia entre dos átomos unidos covalentemente, que en el caso del enlace C-C es de 1.2-1.54Å.

La difracción de una sola molécula da una señal demasiado ténue para ser medida, así que se utiliza un conjunto tridimensional ordenado de moléculas, el cristal, para amplificar la señal. La calidad del orden interno del cristal determina la resolución de la difracción.

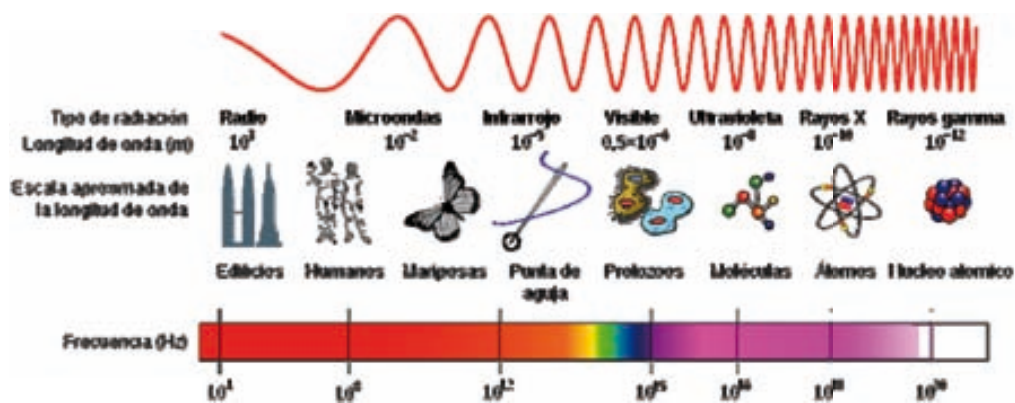


Fig 3.14_ Se denomina espectro electromagnético a la distribución energética del conjunto de las ondas electromagnéticas. El espectro electromagnético se extiende desde la radiación de menor longitud de onda, como los rayos gamma y los rayos X, pasando por la luz ultravioleta, la luz visible y los rayos infrarrojos, hasta las ondas electromagnéticas de mayor longitud de onda, como son las ondas de radio.

Cuando un haz de rayos X incide sobre un cristal, éste interacciona con los electrones de los átomos que componen el cristal, haciéndolos vibrar acopladamente a las variaciones periódicas de su campo eléctrico. Los electrones oscilantes se convierten así en focos de nueva radiación X que se emite en todas direcciones. Este fenómeno se conoce como dispersión, y normalmente es elástica

-la longitud de onda de la radiación emitida es la misma que la de la radiación que impacta el cristal-

El cristal se comporta como una red de difracción, en el cual los rayos X dispersados por los electrones pueden interferir entre sí, generando efectos constructivos o destructivos según la geometría del cristal (Fig 3.15). Éste es el fenómeno de la difracción.

La interferencia es constructiva cuando la diferencia de fase entre la radiación emitida por diferentes átomos es proporcional a 2π . Esta condición se expresa en la ley de Bragg:

$$n\lambda = 2d\sin(\theta)$$

donde,

- n es un número entero,
- λ es la longitud de onda de los rayos X,
- d es la distancia entre los planos de la red cristalina y,
- θ es el ángulo entre los rayos incidentes y los planos de dispersión.

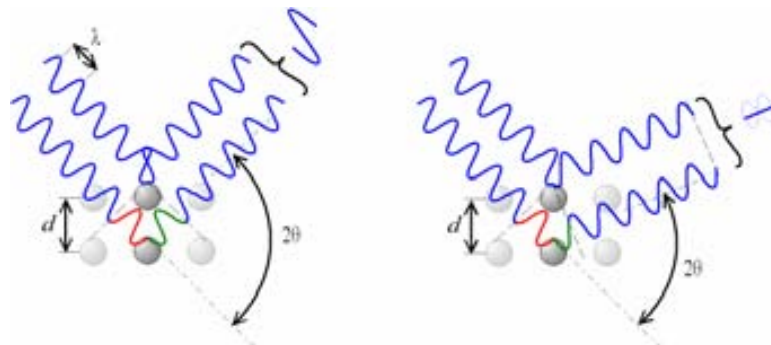


Fig 3.15_ De acuerdo al ángulo de desviación (2θ), el cambio de fase de las ondas produce interferencia constructiva (figura izquierda) o destructiva (figura derecha).

La imagen resultante recogida por el detector consiste en una serie de puntos discretos llamados reflexiones, producto del efecto constructivo de la dispersión de los rayos X de los electrones en la red cristalina de átomos. Cada reflexión contiene información de todos los átomos de la estructura.

Para que se produzcan todos los haces difractados posibles, el cristal debe ser girado de tal modo que todos los planos virtuales del cristal se coloquen en la posición adecuada para cumplir la ley de Bragg. Este fenómeno se puede explicar usando la red recíproca y la esfera de Ewald -simplificación del modelo de Bragg- (Fig 3.16).

Los rayos X son difractados por los electrones, por tanto, tras la resolución de la estructura (véase más adelante), el resultado es un mapa tridimensional de la distribución de los electrones en la estructura proteica.

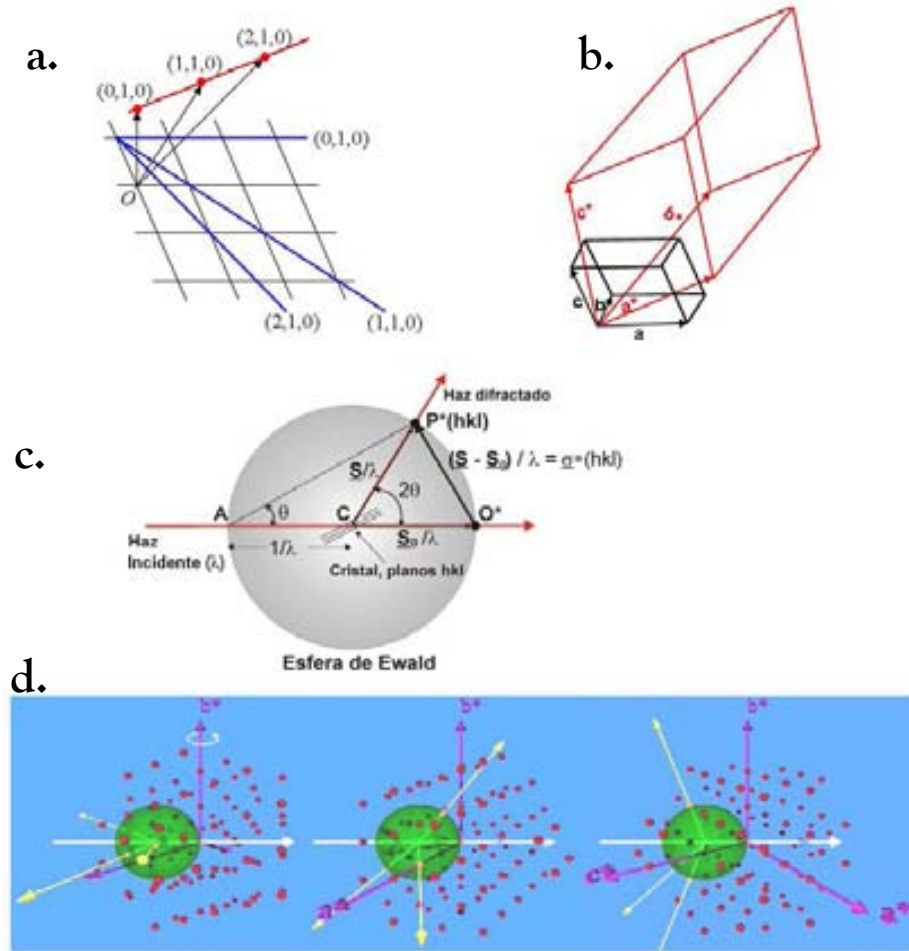


Fig 3.16_ (a) Construcción de la red recíproca a partir de las familias de planos de Braggs (planos equivalentes que se encuentran a una distancia $2\pi/d_{hkl}$ del origen O y orientados en la dirección normal al plano del mismo índice). Por claridad del dibujo el tercer eje de la red directa sería perpendicular al dibujo. Cada una de estas familias está representada por un solo plano (azul) perpendicular al dibujo. Los puntos recíprocos (rojo), están situados en la dirección perpendicular a su familia de planos y a una distancia del origen inversamente proporcional al espaciado d de cada familia. Estos puntos se identifican con los mismos índices de Miller (h, k, l) que los planos. Por ejemplo, el punto recíproco de índices (2,1,0) está situado sobre el vector perpendicular al plano (2,1,0) y su distancia al origen O es inversamente proporcional al espaciado de dicha familia de planos. (b) Sobre esta red recíproca se puede definir también una celda unidad (celda recíproca en rojo) cuyas traslaciones periódicas vienen determinadas por tres ejes recíprocos que forman entre sí unos ángulos recíprocos. Los ejes y ángulos de la celda unidad recíproca se denominan con las mismas letras que la celda unidad directa, añadiéndoles un asterisco. Estos ejes recíprocos (a^* , b^* , c^*) corresponden a los vectores de índices (1,0,0), (0,1,0) y (0,0,1), respectivamente, de forma que cualquier vector recíproco se puede expresar como una combinación lineal de estos tres vectores recíprocos. (c) Interpretación de la difracción según el modelo de Ewald. Cuando un punto recíproco P (hkl) se sitúa sobre la superficie de la esfera de radio $1/\lambda$ se genera un haz difractado que emergiendo del centro de la esfera pasa por P(hkl). (d) Los rayos incidentes de longitud de onda λ (línea blanca), llevan asociados una esfera imaginaria (verde), la esfera de Ewald. La red recíproca (puntos rojos), se mueve con el cristal y cada vez que un punto recíproco coincide con la superficie de la esfera de Ewald (verde) se produce un haz difractado (amarillo) que se visualiza en forma de un punto negro (reflexión) en la imagen de difracción.

Recogida de datos

Los datos de difracción cristalográfica de este trabajo fueron tomados en el sincrotrón European Synchrotron Radiation Facility (ESRF) de Grenoble (Francia), y con generadores de ánodo rotatorio en el laboratorio de George Sheldrick, Göttingen (Alemania) y en la PAC del PCB (Barcelona). Para la generación de rayos X en el sincrotrón, los electrones se aceleran hasta energías elevadas, y posteriormente se varía su trayectoria mediante imanes. En cada cambio de dirección, los electrones emiten tangencialmente al giro parte de su energía en forma de rayos X. El haz de rayos X de un sincrotrón es, además de tener una elevada intensidad, monocromático y colimado (todas las ondas son paralelas y corren en la misma dirección).

Para su recogida, los cristales previamente montados en el loop se colocan en la cabeza goniométrica que permite el posicionamiento adecuado y el giro del cristal con respecto al haz de rayos X (Fig 3.17). Los cristales se mantienen congelados durante toda la recogida mediante un flujo de nitrógeno gas a 100° K suministrado por un criostato.

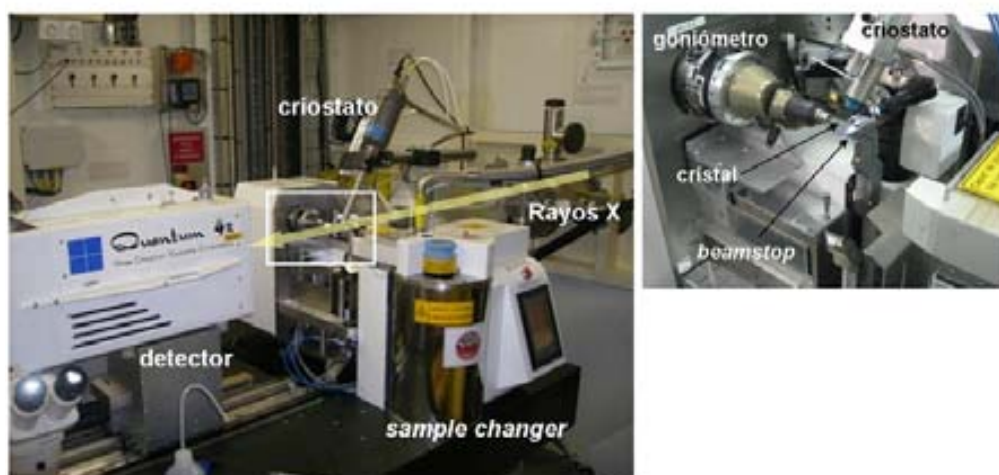


Fig 3.17_ Disposición de aparatos de la zona de experimentación en un experimento de difracción en el sincrotrón. El *sample changer* es el robot que monta y desmonta automáticamente los cristales en la cabeza del goniómetro. El *beamstop* evita que el haz incidente de rayos X impacte sobre el detector ya que su intensidad es tan fuerte que no se podría apreciar la difracción.

En la recogida primero se debe determinar la simetría del cristal, los parámetros de la celda unidad, la orientación del cristal, el límite de resolución y el intervalo máximo de rotación por foto (en el que no se superpongan reflexiones diferentes).

Estos parámetros fueron determinados con MOSFLM (Leslie, 1992) que nos permite también establecer una estrategia de recogida que optimiza la resolución y la completitud de los datos.

Se debe tener en cuenta también la dosis de radiación a la que se somete el cristal. Se debe evitar que las reflexiones de intensidad elevada saturen la capacidad de medición del detector o que sean tan bajas que no se distingan del ruido de fondo. En el caso que estas no sean compatibles se pueden realizar dos recogidas, una a alta resolución y otra a baja. Además, en la mayoría de las líneas del sincrotrón debemos tener en cuenta el daño causado por la radiación. Debemos ajustar pues la transmisión del haz y el tiempo por foto ya que interesa que el cristal pueda aguantar una recogida completa.

Mini Kappa

Es un goniómetro multi-ejes de manera que nos permite orientar el cristal en las tres dimensiones y no únicamente a lo largo de un solo eje como ocurre con los goniómetros usados normalmente. Permite realizar recogidas con diferentes orientaciones, lo que nos asegura una buena completitud y redundancia ya que para obtener una buena estimación de la intensidad de una reflexión es necesario medirla varias veces.

Además, elimina el efecto de distorsión por el daño por radiación en la recogida de reflexiones entre las cuales existe una diferencia anómala que suele ser tenue y puede perderse a causa del “envejecimiento” del cristal durante la recogida, ya que se puede escoger la orientación adecuada para que sean recogidas a la vez las dos reflexiones que contienen la señal anómala.

En la Id14.4 se ha desarrollado un software STAC (STrategy for Aligned Crystals, S. Brockhauser) que guía al usuario en el proceso de alineamiento del cristal y la generación de una estrategia de recogida en dos pasos para tener una buena completitud.

La utilización del goniómetro mini-kappa ha sido muy importante en este estudio ya que los datos de difracción de los cristales en un solo eje no eran óptimos.



Fig 3.18_ Imagen de un goniómetro mini-Kappa donde se muestra la rotación de los diferentes ángulos. 2θ varía en la base del goniómetro, el ángulo ω rota perpendicular al plano del goniómetro; el ángulo del κ rota sobre un eje en 50° del eje del ω ; y, el ángulo ϕ sobre el eje de rotación del cristal durante la recogida. Cuando el ángulo del κ es cero, se alinean ω y ϕ .

Procesamiento de los datos

El conjunto de imágenes obtenidas durante la recogida de datos de difracción de rayos X se debe procesar para obtener un fichero con todas las reflexiones medidas. Los parámetros que se evalúan son *la posición de cada rayo difractado*, que depende de la red cristalina y *la intensidad de cada reflexión*, que depende de la distribución de los átomos en la celda.

Indexación, integración y escalado

El proceso de indexación consiste en determinar el grupo espacial, la orientación del cristal y la celda unidad a partir de las imágenes de difracción. Los programas de indexación analizan la disposición de las reflexiones en el difractograma -normalmente con dos imágenes 0° y 90° es suficiente- y calculan la orientación del cristal. Teniendo en cuenta la longitud de onda de la radiación y la distancia entre el detector y el cristal se calcula la medida de la celda unidad y el grupo espacial cristalográfico. Una vez determinados estos datos, se pueden utilizar para predecir las reflexiones en el resto de imágenes, lo que permite, para cada imagen recogida, una medida de la intensidad de cada una de las reflexiones predichas. Se asigna a cada reflexión un índice de Miller (hkl) (Fig 3.16).

La integración consiste en medir la intensidad de cada reflexión. Cada reflexión se encuadra en una caja tridimensional (en el plano de la imagen y a lo largo de la oscilación en φ) para determinar el ruido de fondo y la distribución de la intensidad se ajusta a una distribución gaussiana. Se genera entonces un fichero que contiene el índice, la intensidad integrada y la desviación estándar para cada reflexión.

El escalado promedia todas las reflexiones que se repiten y agrupa las que están dispersas en varias imágenes debido a la mosaicidad. A causa de las pequeñas variaciones en el tiempo de exposición, la intensidad del haz o el daño por radiación a lo largo de la recogida, se debe aplicar un factor de escala entre reflexiones recogidas en diferentes imágenes.

Cuando se da un fenómeno de absorción, las reflexiones equivalentes generadas por los elementos de simetría en el cristal tienen intensidades diferentes que las reflexiones equivalentes generadas por un centro de inversión. Es decir, el fenómeno de absorción provoca que la pareja de reflexiones h,k,l y \bar{h},\bar{k},\bar{l} llamada par de Friedel, tenga un valor de intensidad diferente y que no se cumpla la ley de Friedel. Así pues, en el caso de un experimento de difracción anómala, durante el escalado estos pares se deben procesar por separado, ya que en este caso no se cumple la ley de Friedel y las reflexiones opuestas no tienen la misma amplitud.

Los datos recogidos en este trabajo han sido indexados e integrados con los programas MOSFLM (Leslie, 1992) y XDS (Kabsch, 1993) y escalados con SCALA (Evans, 2006) y XSCALE (Kabsch, 1993).

Al final del procesamiento se obtienen unos valores -ordenados por capas de resolución e intensidad- que nos dan información sobre la calidad de los datos (Glusker, 1994).

- La *completitud* indica el porcentaje de las reflexiones únicas observadas con respecto a las posibles.
- La relación $\langle I \rangle / \langle \sigma_I \rangle$ nos indica la intensidad promedio de los datos respecto al promedio de error de los mismos.

- El R_{merge} y R_{sym} son residuales que miden la similitud entre reflexiones idénticas recogidas varias veces o relacionadas por simetría. Valores bajos indican que las reflexiones están bien recogidas y procesadas.

$$R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I(hkl)};$$

donde, $I_i(hkl)$ es la medida i de la reflexión hkl .

Esta fórmula da valores mayores de R-factor cuando la redundancia es mayor por estar comparándose un mayor número de reflexiones (Diederichs & Karplus, 1997). Así pues los conjuntos de datos de baja redundancia parecerían mejores que los de mayor redundancia, lo que no nos da una óptima indicación de la calidad de los datos. R_{meas} , en cambio, nos da un valor residual independiente de la redundancia.

$$R_{meas} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_i^n |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I(hkl)}$$

Para atribuir un límite de resolución se tomó como criterio que en el último intervalo de resolución la completitud fuera superior al 55%, las reflexiones debían tener una intensidad de al menos 1.2 veces el error y su R_{merge} no debía exceder el 55-60%.

Estimación del número de moléculas por unidad asimétrica y del porcentaje de solvente de los cristales.

La unidad asimétrica es la menor unidad estructural a partir de la cual aplicando los elementos de simetría del correspondiente grupo espacial se puede generar la celda unidad del cristal. La celda unidad es aquella que duplicada y trasladada en el espacio puede generar la totalidad del cristal.

La determinación del número de macromoléculas por unidad asimétrica y del contenido de solvente de los cristales aporta una información valiosa para la mejora de las fases (se explica más adelante).

Para ello se utilizó el coeficiente de Matthews V_M , que relaciona el volumen de la celda unidad con el peso molecular de la macromolécula

(Matthews, 1968; Kantardjieff & Rupp, 2003). Existe una web <http://www.ruppweb.org/Mattprob/> que hace una estimación del número de moléculas por unidad asimétrica y el contenido de solvente

$$V_M = \frac{V_C}{M_w \cdot Z \cdot X};$$

donde, V_C es el volumen de la celda

M_w es el peso molecular de la macromolécula

Z es el número de unidades asimétricas de la celda unidad

X es el número de moléculas por unidad asimétrica.

V_M está relacionado con el volumen de solvente en un cristal, que para un complejo proteína-DNA:

$$V_{solvente} = 1 - \frac{1.66 \cdot (0.74N_{prot} + 0.5N_{DNA})}{V_M};$$

donde, 0.74 y 0.5 son los volúmenes específicos parciales (cm^3/g) de las proteínas y del DNA respectivamente.

N_{prot} y N_{DNA} son las proporciones de proteína:DNA en el complejo en tanto por 1.

Corrección anisotrópica

La difracción anisotrópica se da cuando existe una dependencia direccional en la calidad de la difracción. Comúnmente se observa en la cristalografía de proteínas, y se atribuye a la vibración anisotrópica, que puede ser debida a que las interacciones de empaquetamiento del cristal son más uniformes en una dirección con respecto a la otra. Frecuentemente, la dependencia direccional coincide con las direcciones recíprocas de la celda unidad, a^* , b^* y c^* .

En los casos que presentan anisotropía severa, el refinado se puede encallar en valores de R-factors altos y los mapas de densidad electrónica pueden aparecer sin rasgos distintivos impidiendo el trazado de la proteína.

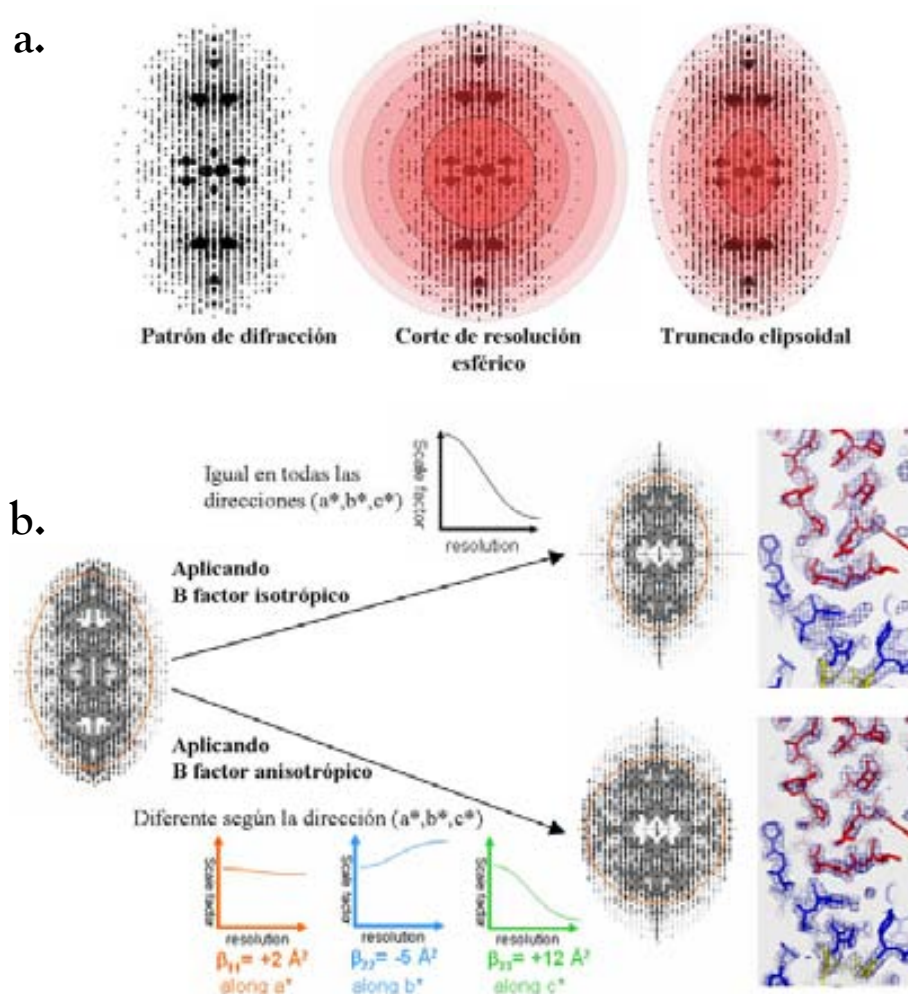


Fig 3.19_ (a) La difracción anisotrópica da lugar a una forma elipsoidal, ya que la resolución de la difracción depende de la dirección. El truncado elipsoidal permite eliminar de los datos las reflexiones más débiles de la dirección con difracción más pobre sin perder las reflexiones de alta resolución de la dirección con mejor difracción. (b) (arriba) Idealmente el B-factor varía según la resolución de igual forma en las tres direcciones. Aplicando un B-factor isotrópico arrastramos el problema de la anisotropía y los mapas resultantes son difícilmente interpretables. (abajo) El factor de escalado anisotrópico varía con la resolución, como el B-factor, pero también con la dirección. Tiene tres componentes principales β_{11} , β_{22} y β_{33} que actúan como B-factors a lo largo de las direcciones a^* , b^* y c^* respectivamente. Así, un conjunto de datos anisotrópico puede ser convertido a isotrópico mediante la aplicación de un factor de escala apropiado que incrementa el la dirección de difracción pobre o disminuye en la dirección buena de la difracción, dando lugar a mapas interpretables.

En el procesamiento de los datos, si se quieren incluir todas las reflexiones en la dirección “buena” de la difracción, se incluyen también las reflexiones de la dirección de difracción pobre, ya que el corte de resolución en los programas es esférico, mientras que en un caso de anisotropía, la intensidad del patrón de difracción es elipsoidal. El truncado elipsoidal consiste en imponer un corte de resolución elipsoidal que se ajuste a los datos recogidos, de manera que eliminamos de los datos las reflexiones más débiles (Fig 3.19a).

En los mapas electrónicos, la falta de detalles se puede paliar mediante el uso de algoritmos que aplican factores de escala (equivalentes a B-factors) dependientes de la resolución a las tres direcciones principales del conjunto de datos, de manera que las magnitudes de los factores de estructura (descritos más adelante) decrecen proporcionalmente con la resolución en las tres dimensiones. Así pues, se aumenta la magnitud de las reflexiones en la dirección pobre de la difracción y se disminuye la magnitud de las reflexiones en la dirección buena de la difracción. El efecto colateral de este algoritmo es que, en la dirección buena de la difracción, las reflexiones a alta resolución contribuyen muy poco a la densidad electrónica. Para solucionar este problema se aplica al conjunto de datos un B-factor isotrópico negativo, para restaurar la magnitud de las reflexiones de alta resolución. Como resultado, los mapas de densidad electrónica mejoran (Strong *et al.*, 2006) (Fig 3.19b).

Para evaluar la anisotropía y tener una noción del límite de resolución para los tres ejes principales del elipsoide, se usó el servidor <http://www.doe-mbi.ucla.edu/~sawaya/anisyscale/>. La corrección de los datos se llevó a cabo con el programa XPREP (Sheldrick, 2008). En este caso únicamente se aplicó el escalado anisotrópico para no perder la información, que aunque débil, contenía el eje de menor difracción. En la zona, aproximadamente a 3.4Å, donde se observaba el *spot* alargado característico del *stacking* de un B-DNA se llevó a cabo un escalado local.

5.2 Resolución de la estructura

El problema de la fase

El objetivo de un estudio estructural es la determinación de la densidad electrónica, $\rho(x,y,z)$, para cada punto de la celda unidad cristalina. Para ello partimos de la información del espacio recíproco (difracción). La transformación entre la difracción (espacio recíproco) y la densidad electrónica (espacio real), implica la transformada de Fourier:

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l |F(hkl)| \cdot \cos 2\pi(hx + ky + lz - \phi(hkl))$$

donde,

$F(hkl)$: factores de estructura definidos en cada punto del espacio recíproco (hkl) y que contienen la resultante de la dispersión de los rayos X emitidos por los electrones de la celda unidad en esa dirección.

Los factores de estructura son ondas, y se pueden describir por:

- módulos $|F(hkl)|$: representan las amplitudes de los haces difractados y su valor se obtiene directamente de las intensidades que se recogen en el detector. La relación entre el módulo de los factores de estructura y la intensidad de los puntos del espectro de difracción:

$$I(hkl) = K \cdot A \cdot L \cdot p \cdot |F(hkl)|^2, \quad \text{donde } K \text{ es un factor de escala, } A \text{ es un factor de absorción, } L \text{ es el factor de Lorentz y } p \text{ es el factor de polarización}$$

- fases $\phi(hkl)$: esta información no se puede medir durante el experimento de difracción, y por tanto a partir de éste no se puede determinar directamente la densidad electrónica ρ para un punto (x,y,z) del espacio.

La posición e intensidad de cada punto (hkl) se determina a partir del patrón de difracción, pero para obtener un mapa de densidad electrónica también deben ser determinadas las fases de las ondas que formaron cada punto. Este problema se puede resolver perturbando la estructura y el patrón de difracción (**reemplazamiento isomórfico** y **dispersión anómala**) o usando estructuras similares conocidas (**reemplazamiento molecular**) (Taylor, 2003).

Reemplazo isomórfico

Este método fue introducido por Max F. Perutz y John Kendrew al resolver por primera vez la estructura cristalográfica de la hemoglobina. Consiste en introducir átomos pesados –grandes dispersores de los rayos X–, en la celda unidad cristalina sin que ello distorsione la estructura cristalina de la proteína nativa. Los cristales derivados deben ser, por tanto, isomorfos con respecto a los nativos. En la práctica este reemplazo isomórfico se lleva a cabo difundiendo complejos de metales pesados a través de los canales que poseen los cristales de proteína, y en

donde normalmente existen cadenas laterales de aminoácidos con capacidad de coordinar a los átomos metálicos (técnica de *soaking*). Puesto que los átomos metálicos tienen un mayor poder de dispersión, pueden llegar a modificar sensiblemente la intensidad del patrón de difracción del derivado. La diferencia de intensidad entre las reflexiones del cristal nativo y derivado se puede usar para calcular un mapa de Patterson de diferencias -mapa de vectores de situación relativa entre los átomos pesados-, a partir del cual se puede determinar la posición de los átomos pesados en la celda unidad. La función de Patterson se obtiene a partir de la función de densidad electrónica $\rho(xyz)$, omitiendo la información de las fases y reemplazando los módulos de los factores de estructura por sus cuadrados. Esta función definida en un nuevo espacio de coordenadas u, v, w , se puede calcular directamente a partir de los datos obtenidos en el experimento de difracción y sus máximos contienen información de las distancias interatómicas en el cristal.

$$P(uvw) = \sum_u \sum_v \sum_w |F(hkl)|^2 \cdot \cos 2\pi(hu + kv + lw)$$

En el caso del reemplazamiento isomórfico simple (SIR), la contribución de los átomos pesados en las amplitudes y fases se puede ilustrar con el diagrama de Argand.

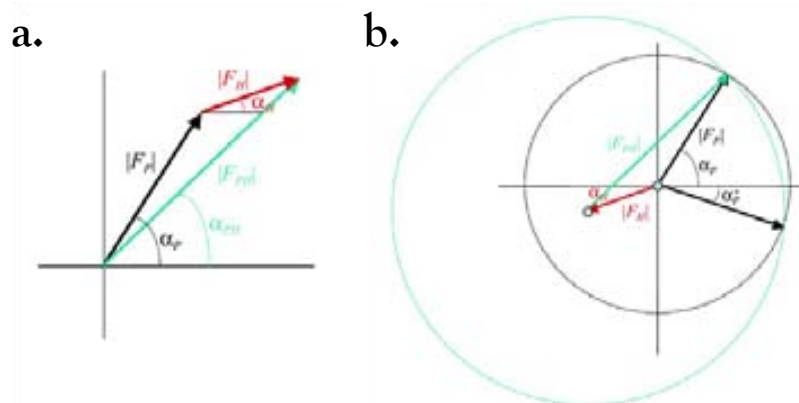


Fig 3.20_ (a) $|F_P|$ y $|F_{PH}|$ se obtienen experimentalmente a partir de los datos de difracción de un cristal nativo y un derivado respectivamente. $|F_P|$ es la amplitud de la reflexión del cristal nativo y $|F_{PH}|$ es la amplitud de la reflexión del derivado. La diferencia isomórfica entre ambos $|F_H| \approx |F_{PH}| - |F_P|$ puede ser utilizada para estimar la amplitud del factor de estructura de los átomos pesados y determinar su posición usando Patterson o métodos directos. Una vez localizados, se pueden definir los parámetros de los átomos pesados (posiciones xyz, ocupación y factores térmicos Debye-Waller) y usarlos para calcular de manera más precisa $|F_H|$ y así como la fase correspondiente α_H . (b) Las fases de la proteína nativa α_P , se pueden estimar usando la regla del coseno, $\alpha_P = \alpha_H + \cos^{-1} [(F_{PH}^2 - F_P^2 - F_H^2)/2F_P F_H]$, dando lugar a dos posibles soluciones distribuidas simétricamente α_P y α_P' . Esta ambigüedad se ilustra con la construcción de Harker,

donde las dos posibles soluciones son las intersecciones de los círculos, centrados en los puntos azules y de radio $|F_P|$ y $|F_{PH}|$.

En realidad, existen errores asociados con las medidas de los factores de estructura y las posiciones y con la ocupación de los átomos pesados. Asumiendo que todos los errores residen en $F_{PH(calc)}$ y que éstos siguen una distribución gaussiana, la probabilidad de que una fase tenga un valor cierto es:

$$P(\alpha_P) \propto \exp\left(\frac{-\varepsilon^2}{2E^2}\right), \quad \text{donde } E = \left\langle \left[F_{PH(obs)} - F_{PH(calc)} \right]^2 \right\rangle$$

Muchos programas de faseado calculan las probabilidades de 0 a 360° en intervalos de 10°, dando lugar a una distribución de probabilidad de la fase cuya forma puede representarse por cuatro coeficientes llamados coeficientes de Hendrickson-Lattman.

El uso de más de un derivado en el reemplazo isomórfico múltiple (MIR) puede romper esa ambigüedad. La probabilidad de la fase se obtiene multiplicando las probabilidades individuales para una misma reflexión, pero esta vez los dos o más derivados con átomos pesados resultan en una distribución unimodal con mayor contraste y con una figura de mérito mayor (Fig 3.21).

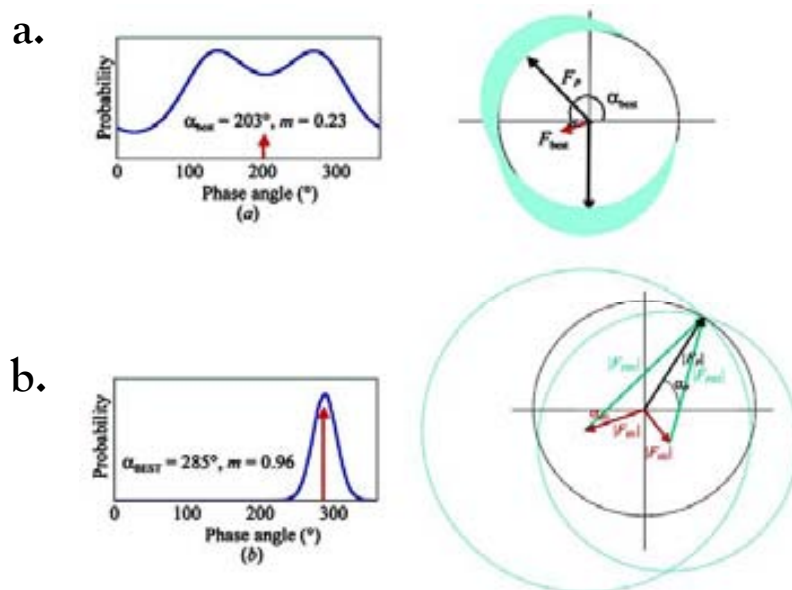


Fig 3.21_ (a) Para un experimento SIR. (dcha.) Representación de la distribución de probabilidad de fase para una reflexión y (izq.) diagrama de Harker (se representa la probabilidad en azul alrededor del círculo). (b) Para un experimento MIR. (dcha.) Probabilidad de fase de 3 derivados y (izq.) diagrama de Harper para 2 derivados H1 y H2 donde se puede determinar la fase correcta α_P común a las dos representaciones.

Dispersión anómala

Los cambios que se provocan en la intensidad de difracción al introducir átomos pesados en los cristales de proteína, se pueden considerar como *modificaciones químicas* de la difracción. De un modo análogo, se pueden provocar cambios en las intensidades de difracción *modificando las propiedades físicas* de los átomos. Si la radiación incidente tiene una frecuencia próxima a la frecuencia natural de oscilación de los electrones de un determinado átomo se produce un fenómeno de absorción y con ello una **dispersión anómala** que modifica el factor de dispersión atómico (Fig 3.22b). Este fenómeno provoca el no cumplimiento de la ley de Friedel (Fig 3.22c), de manera que se produce un cambio de intensidad entre las parejas de reflexiones que, en condiciones normales, deberían tener la misma amplitud e idénticas fases pero con signos opuestos.

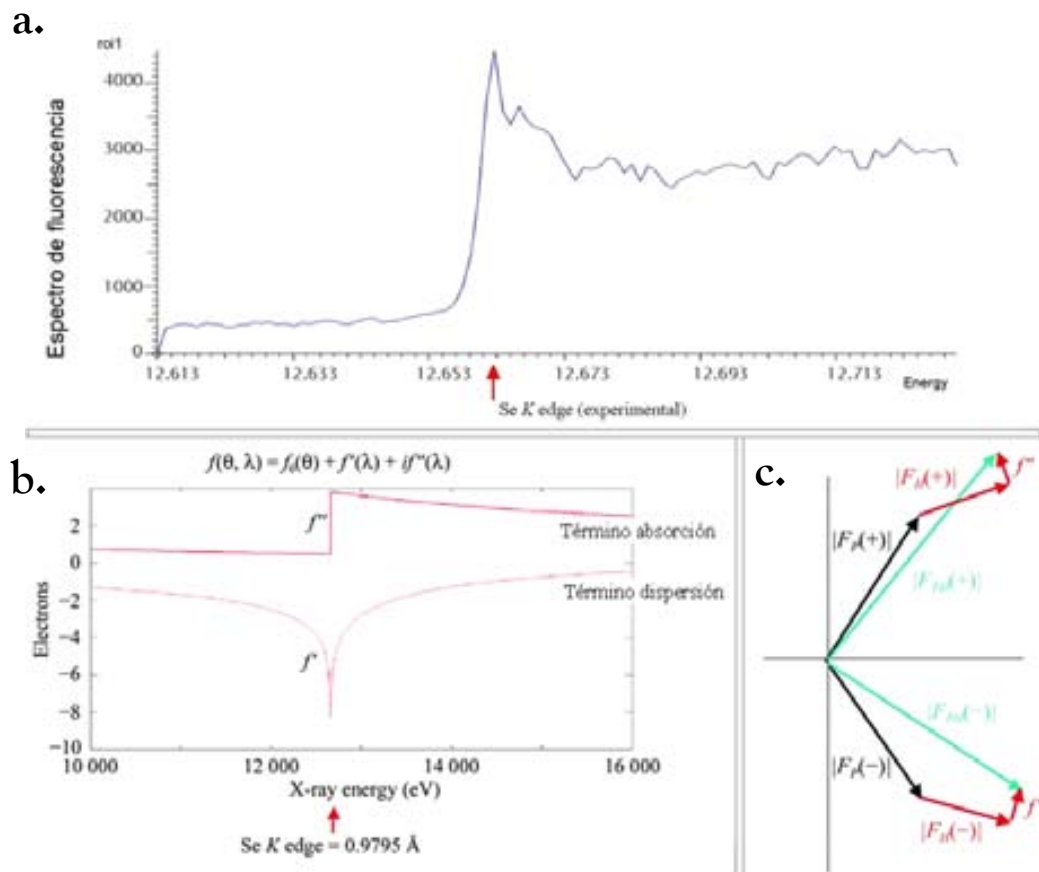


Fig 3.22_ (a) Espectro de absorción de la radiación de rayos X de un cristal de SeMet de mTERF- Δ N DNA29bp medido en la Id29 del sincrotrón. Se observa un salto en la absorción llamado “borde de absorción” (*absortion edge*) debido al efecto de la frecuencia incidente sobre los electrones del átomo que varía su factor de dispersión. A partir de la curva se determinan el pico de absorción y el punto de inflexión. (b) La modificación de f según la frecuencia incidente se puede representar con la ecuación indicada en la figura $f(\theta, \lambda) = f_0(\theta) + f'(\lambda) + i f''(\lambda)$. Siendo f' y f'' los ‘componentes ligados al átomo’ relacionados por la ecuación de kramers-kronig. El gráfico muestra la variación de las componentes real (f') e imaginaria (f'') del factor de dispersión atómico del selenio en función de la energía de los rayos X incidentes. En la web

http://skuld.bmsc.washington.edu/scatter/AS_periodic.html se pueden consultar las *absortion edge* de todos los elementos de la tabla periódica. (c) No cumplimiento de la ley de Friedel $F_{hkl} \neq F_{\bar{h}\bar{k}\bar{l}}$, cuando está presente un dispersor anómalo. La componente f'' distorsiona $|F_{PH}(+)|$ y $|F_{PH}(-)|$ de manera distinta.

En estos casos, la recogida de datos en el sincrotrón se realiza en líneas donde la λ es modulable, escogiendo la energía donde se maximice la señal anómala.

Esta dispersión anómala que da lugar a diferencias entre reflexiones que de otra manera serían equivalentes, puede servir, de la misma manera que la diferencia isomórfica en el Patterson, para localizar los átomos que la producen. Las fases para los factores de estructura nativos se pueden deducir de manera similar al caso SIR o MIR.

La dispersión anómala se puede utilizar también para romper la ambigüedad de fase de un experimentos SIR dando lugar a un **SIRAS** (reemplazo isomórfico único con dispersión anómala). El uso de reemplazo isomórfico múltiple con dispersión anómala se denomina **MIRAS**.

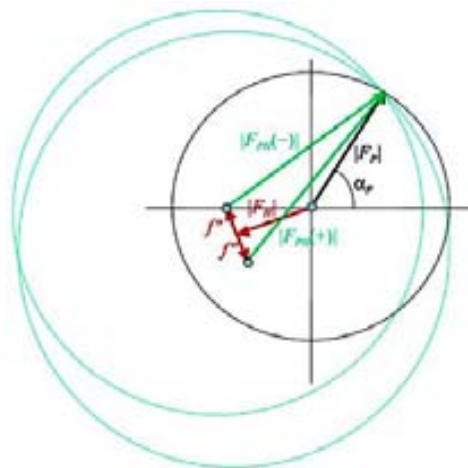


Fig 3.23_ Diagrama de Harker para un experimento SIRAS. $|F_P|$ y $|F_{PH}(+)|$ y $|F_{PH}(-)|$ se obtienen experimentalmente a partir de los datos de difracción de un cristal nativo y un derivado medido en la longitud de onda donde presenta una dispersión anómala. El incumplimiento de la ley de Friedel, permite distinguir la ambigüedad resultante de un experimento SIR (dos posibles soluciones para las fases de la proteína nativa), siendo la solución la intersección de los tres círculos, uno centrado en los puntos azules indicados en la figura y de radio $|F_P|$, $|F_{PH}(+)|$ y $|F_{PH}(-)|$.

La técnica de SIRAS y MIRAS fue utilizada con los derivados de Hg y Au de los cristales del complejo mTERF-NHis con el DNA de 15pb.

MAD

El método MAD, desarrollado por Hendrickson y Kahn, implica la medida de los datos de difracción de un cristal de proteína que contenga un dispersor anómalo fuerte, usando radiaciones de distinta energía.

Normalmente se utiliza como dispersor anómalo la proteína recombinante expresada en presencia de SeMet, asegurándose de esta manera la presencia del átomo pesado en el cristal. En caso de que la proteína con SeMet no cristalice, una alternativa es introducir átomos pesados en la red cristalina o bien explotar la señal anómala de azufre presente de forma natural en las proteínas. El inconveniente de la incorporación de átomos pesados, es la alteración del cristal incrementando la mosaicidad o disminuyendo las intensidades de las reflexiones. El inconveniente del S es la baja energía de la λ utilizada ($\lambda = 5.0155\text{\AA}$)

Hay que tener en cuenta que los cambios en las amplitudes del factor de estructura de la dispersión anómala son generalmente pequeños y se requiere una medida precisa de las intensidades. Se recomienda, por tanto recoger, sobre la misma imagen de difracción, las reflexiones que muestran las diferencias.

Actualmente el espectro de absorción del átomo pesado se puede determinar experimentalmente con un escáner de fluorescencia en el sincrotrón, ya que el entorno químico del dispersor anómalo puede afectar a la longitud de onda de absorción teórica (Fig 3.22a). Se necesita, por tanto, una óptica excelente para una elevada precisión de la longitud de onda de medida.

Se recogen varios conjuntos de datos a diferentes longitudes de onda, usualmente 3, para maximizar los efectos de absorción y dispersión. Normalmente, las λ escogidas corresponden al pico de absorción f'' , (λ_1), en el punto de inflexión de la curva de absorción (λ_2), donde el término dispersivo (la derivada de la curva f'') tiene un mínimo y a una longitud de onda remota (λ_3 y/o λ_4) (Fig 3.24a). Combinando estos conjuntos de datos de difracción y analizando las diferencias entre ellos, es posible calcular la distribución de amplitudes y fases que generan los dispersores anómalos. El uso de las fases generadas por estos dispersores anómalos,

como una primera aproximación a las fases globales, permite calcular la densidad electrónica para la proteína.

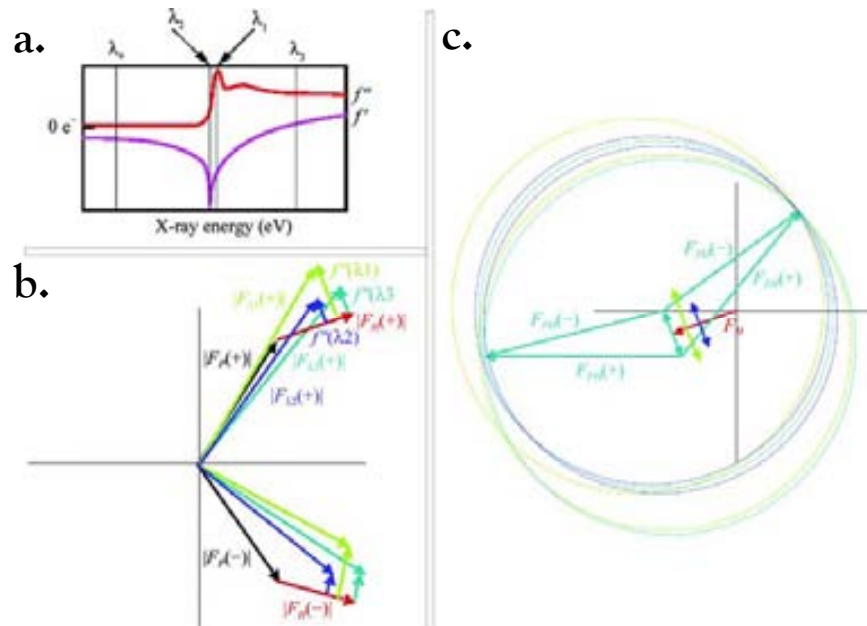


Fig 3.24_ Faseado con MAD. (a) Curva de absorción de un dispersor anómalo con las respectivas longitudes de onda de medida. (b) Diagrama de fase. Como $|F_p|$ no se ha medido, se escoge una de las λ 3 o 4 como 'nativo'. (c) Construcción de Harper donde la fase correcta queda inequívocamente determinada.

Las ventajas que presenta este método con respecto al MIR son:

- al recoger la difracción con un único cristal desaparecen los problemas de falta de isomorfismo -debidos a cambios en la celda unidad, reorientaciones de la proteína, cambios conformacionales y cambios en la sal y los iones solvatados- tan habituales en la técnica MIR.
- Mientras que el factor de dispersión atómico en ausencia de dispersión anómala f_0 decrece drásticamente con el ángulo de dispersión, la componente anómala del factor de dispersión atómico ($f' + if''$) aumenta con la resolución del espectro -a ángulos de dispersión elevados-. Así, las estimaciones de fases mediante este método son, en general, mejores a alta resolución. En un experimento MIR, la falta de isomorfismo es más acusado en las reflexiones de alta resolución, por lo que en la mayoría de casos la información de fases mediante esta técnica no se puede obtener a partir de reflexiones de resolución mayor a 3.5 \AA

SAD

Este método, consiste en recoger a una única longitud de onda, normalmente la del pico de absorción, y usar los protocolos de *density-modification* (DM), para romper la ambigüedad de las fases y dar lugar a mapas interpretables (Dodson, 2003).

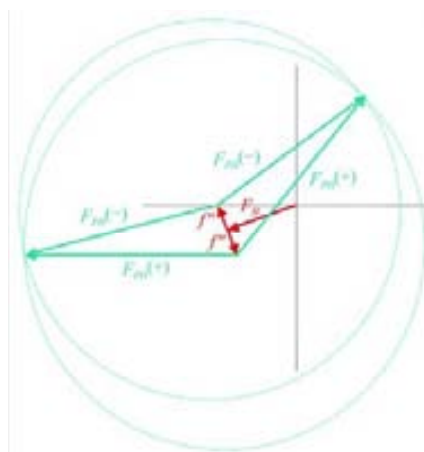


Fig 3.25_ Construcción de Harker para SAD. $\Delta F^{\pm} = |F_{PH}^{+}| - |F_{PH}^{-}|$ se usa para encontrar la subestructura de los dispersores anómalos, y determinar así la fase. Para romper la ambigüedad mostrada en la figura, se deben aplicar los métodos de la mejora de fases.

Las técnicas de SAD y MAD fueron utilizadas para obtener fases de los cristales de mTERF- Δ N con los DNAs de 29 y 15pb.

Se utilizaron programas como el XPREP-SHELXD-SHELXE (Sheldrick, 2008) o SHARP (Bricogne *et al.*, 2003) para resolver la estructura.

Reemplazo molecular.

Este método consiste en asignar las fases de la estructura modelo como fases iniciales para la estructura que queremos determinar. Así pues, la estructura conocida se rota y traslada al nuevo sistema cristalino hasta obtener un buen encaje con los datos experimentales. Como regla general, se necesita una identidad de secuencia superior al 25% entre el modelo y la proteína de interés y una desviación cuadrática media (r.m.s.) menor a 2Å entre los carbonos α del modelo y de la estructura nueva final, aunque existen excepciones.

Si se obtiene una solución con sentido, en la que las moléculas simétricas se distribuyan de una forma razonable dentro del empaquetamiento (no se superpongan), se pueden calcular a partir de ella las fases que nos permiten, con las amplitudes experimentales, calcular un mapa de densidad electrónica. La limitación de éste método es la dependencia de las fases con el modelo inicial que puede provocar *model bias*.

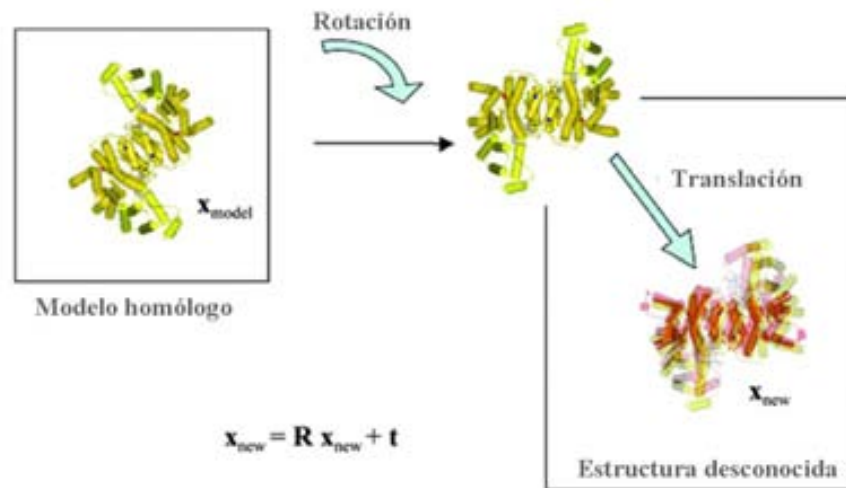


Fig 3.26_ Proceso del reemplazo molecular. Consiste en la rotación y traslación del modelo homólogo (amarillo) para obtener un buen encaje con la estructura desconocida (rojo).

Este método fue utilizado en el faseado de los cristales de complejo mTERF-NHis y mTERF-CHis con el DNA de 15pb, tomando como modelo mTERF- Δ N con y sin DNA mediante el programa MolRep (Vagin, 1997).

Optimización de las fases

Raramente unas fases experimentales son suficientes para dar un mapa de densidad electrónica completamente interpretable. Las fases experimentales son normalmente sólo el punto de partida de la mejora de fases basada en conocimientos previos de ciertas características que un buen mapa de densidad electrónica debería tener. Estos métodos están integrados en programas como DM, DMMULTI (Cowtan, 1994), RESOLVE (Terwilliger, 2000) y CNS (Abrahams & Leslie, 1996).

La modificación de la densidad es un proceso cíclico que comporta modificar el mapa experimental según varios criterios (ver más adelante), calcular la transformada de Fourier para generar nuevas fases, recombinar estas fases con las

experimentales y calcular un nuevo mapa el cual será posteriormente modificado siguiendo el proceso cíclico hasta que se consigue la convergencia (Fig 3.27). Si se dispone de datos nativos a mayor resolución, se pueden extender las fases.

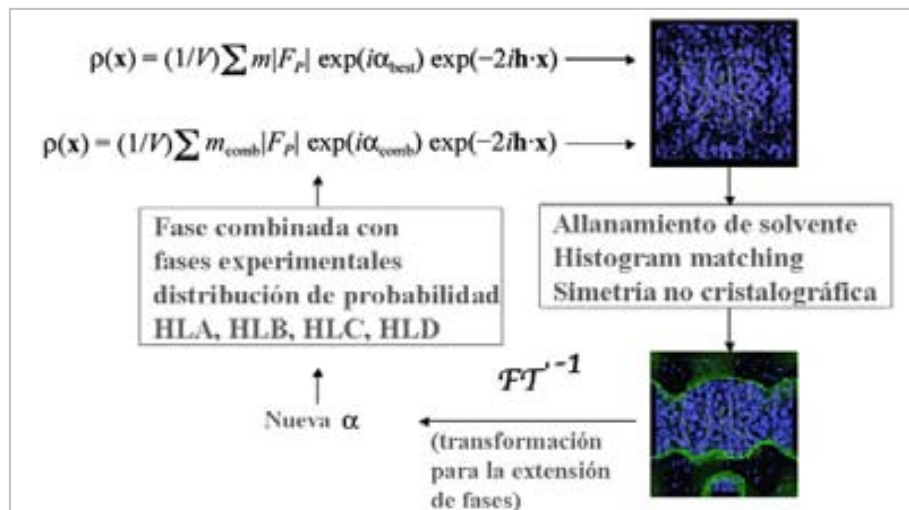


Fig 3.27_ Mejora de las fases mediante DM.

Aplanamiento de solvente

En un cristal típico de proteína, el 50% del volumen está ocupado por solvente. Este solvente desordenado, debería no tener densidad electrónica. Se emplean métodos automáticos para definir la frontera entre proteína y solvente desarrollados por Wang (1985) y trasladados al espacio recíproco por Leslie (1988). Esta técnica elimina la densidad electrónica negativa y da un valor de densidad electrónica $0.33\text{e}\text{\AA}^{-3}$ a las regiones de solvente, en contraste con las zonas de proteína de valor $0.43\text{e}\text{\AA}^{-3}$.

Simetría no cristalográfica y *cross-crystal averaging*

Frecuentemente las proteínas cristalizan con más de una copia en la unidad asimétrica de la celda. En otros casos, las proteínas cristalizan en diferentes formas cristalinas. Cuando la misma proteína aparece en diferentes lugares en un mapa de densidad electrónica, o en mapas de diferentes cristales, la densidad debe ser muy parecida, y si existe diferencias son debidas a errores en las fases. Este método realiza un promedio de la densidad, de manera que se cancelan algunos errores aleatorios y se incrementa por tanto la fidelidad de las fases.

Histogram matching

Las proteínas están formadas por los mismos tipos de átomos y las mismas distancias entre ellos. En consecuencia, en los mapas de densidad electrónica se pueden apreciar el mismo tipo de valores de densidad para diferentes proteínas. Si un mapa en la región de la proteína no tiene la distribución de densidades mayores y menores esperada, es probable que sea por el error en las fases. Alterando la distribución de los valores de densidad mediante un algoritmo, se mejoran las fases correspondientes.

Resolución cristalográfica de la estructura de una proteína *ab initio*

Los métodos *ab initio* están muy restringidos a proteínas de unos 1000 átomos que difractan a resolución atómica o proteínas que en su estructura contienen átomos pesados. De estas dos barreras, la resolución es la más difícil de superar, y hemos de tener en cuenta que la atomicidad es una potente restricción en el espacio real y recíproco. A bajas resoluciones pues, esta restricción debe ser substituida por el conocimiento de que las estructuras macromoleculares están compuestas por fragmentos de geometría conocida (hélices α , hojas β y pares de bases) que son una buena primera aproximación. Es muy efectivo utilizar este hecho para ayudar en el proceso de faseado una vez obtenido el mapa experimental preliminar, así, programas como RESOLVE (Terwilliger, 2003), ARP/wARP (Perrakis *et al.*, 1999) o SHELXE (Sheldrick, 2002) lo han implementado a los algoritmos de autotrazado. Las pruebas han demostrado que el emplazamiento correcto de fragmentos perfectos representando el 13% de la estructura, pueden ser suficientes para el faseado a través de la modificación de la densidad con ACORN (Jia-Xing *et al.*, 2005). Esta aproximación consiste en un cuadro de varias soluciones que combina la localización de pequeños fragmentos modelo con la modificación de la densidad y el autotrazado en los mapas resultantes. El refinado de *B-factor* del mejor trazado mejora la interpretabilidad de los mapas (Rodríguez *et al.*, 2009).

Para el faseado de los cristales del complejo mTERF- Δ N/29bp-DNA se usaron derivados de SeMet. La solución resultante resultó ser insuficiente para la

resolución de la estructura, dando lugar a mapas de baja calidad debido a la severa anisotropía de los datos aún con la corrección anisotrópica de XPREP. Para resolver este problema se utilizó un proceso iterativo con una versión del programa SHELXE (<http://shelx.uni-ac.gwdg.de/SHELX/>), alternando el autotrazado automático de las hélices de la cadena principal con la modificación de la densidad, mejorando así los mapas para permitir la construcción manual del modelo. Las soluciones se pueden mejorar refinando el *B-value* de los átomos de la cadena principal trazados por SHELXE con el programa REFMAC (Murshudov *et al.*, 1997), dando así lugar a una densidad electrónica más fácilmente interpretable, a partir del cual se prosiguió con el trazado manual de la estructura de la proteína.

5.3 Construcción del modelo atómico y refinado

Solventado el problema de la fase, se puede calcular la función de densidad electrónica, la representación gráfica de la cual da lugar al *mapa de densidad electrónica*. Si la distribución de fases estimadas es correcta para el conjunto de módulos de los factores de estructura experimentales, el mapa mostrará zonas con densidad electrónica positiva sobre las cuales deberá acoplarse el esqueleto polipeptídico. Esta labor se realiza mediante programas gráficos y equipos informáticos que permiten, de modo interactivo, trazar no sólo la cadena de carbonos alfa sino las cadenas laterales, adecuándose a la señal de densidad electrónica que muestra el mapa.

Para el trazado de la molécula, en el caso del reemplazo molecular es más sencillo porque ya se dispone de un conjunto de coordenadas atómicas como punto de partida. En el caso del reemplazo isomórfico o dispersión anómala, sólo se dispone de un mapa y se debe construir el modelo entero.

Para construir el modelo atómico, normalmente se coloca un esqueleto proteico en el interior del mapa de densidad electrónica y, a medida que la resolución lo permite, se va insertando la secuencia aminoacídica siempre y cuando la densidad permita asignar el tipo de cadena lateral. El grado de detalle visible depende de la

resolución y la calidad de las fases. Cuando se dispone de las posiciones de los átomos de Se en el caso de los cristales con SeMet, se pueden localizar las metioninas, facilitando así la asignación de la secuencia. En el caso de mTERF, tanto la proteína recombinante con SeMet como los mutantes puntuales por Met realizados para aumentar la señal anómala, fueron de gran ayuda en la asignación de la secuencia.

Para el trazado manual se utilizaron los programas TURBO-FRODO (Roussel, 1989) y Coot (Emsley *et al.*; Emsley & Cowtan, 2004).

Encaje automático y refinado

El modelo estructural viene definido por las coordenadas atómicas (x , y , z) y un factor de temperatura (B) para cada átomo, que expresa la movilidad térmica del átomo alrededor de su posición de equilibrio y que va relacionado con el error de localización del átomo en una posición. En el refinamiento estructural, se ajustan los parámetros estructurales (x , y , z , B) para optimizar el acuerdo entre las observaciones y la predicción. Existen varios métodos para conseguir esta optimización.

Se puede considerar como otra forma de *modificación de la densidad*. Las estructuras proteicas están formadas por átomos. Si la densidad puede ser interpretada en términos de un modelo atómico y los átomos se colocan más o menos en el lugar correcto, la distribución de densidad estará más cercana a la realidad y por tanto en el espacio recíproco las fases correspondientes serán mejores. Para reducir las posibilidades de ajuste de los parámetros estructurales se imponen constricciones o restricciones en la estereoquímica del modelo, lo que es relativamente sencillo ya que se conocen muchos parámetros que determinan cómo los átomos están colocados unos respecto a otros en una proteína -ángulos y distancias de enlace, conectividad química definida por la secuencia aminoacídica, etc.-.

A partir de los primeros mapas de densidad electrónica interpretables, se aplica iterativamente la construcción del modelo en la densidad electrónica seguido de un refinado automático para alcanzar una mayor concordancia del modelo con los datos de difracción observados. Las nuevas fases, calculadas a partir del modelo y refinadas automáticamente, se emplean para calcular un nuevo mapa.

Opcionalmente otras técnicas de modificación de densidad, como el allanamiento de solvente y el *averaging* (promediado de mapas correspondientes a moléculas de la unidad asimétrica), se pueden aplicar antes de la construcción del modelo en cada nuevo ciclo.

Los programas utilizados en el proceso de refinado de las estructuras de mTERF fueron PHENIX (Adams *et al.*) y REFMAC, incorporando TLS (Winn *et al.*, 2001; Winn *et al.*, 2003). El TLS permite reducir el R_{free} y el R_{factor}, ya que produce un modelo estructural que representa mejor los datos experimentales. Se realiza definiendo grupos de refinado. Existe un servidor *TLS Motion Determination (TLSMD)* que, partiendo de la secuencia de aminoácidos, recomienda varias posibilidades de cómo segmentar la secuencia (para una misma secuencia existen diferentes grupos de TLS) (Painter & Merritt, 2006).

La construcción del modelo se realizó utilizando el programa Coot, en base a mapas de densidad electrónica y mapas de Fourier calculados con coeficientes: $2F_{\text{obs}}-F_{\text{calc}}$ y $F_{\text{obs}}-F_{\text{calc}}$, éstos últimos para comprobar las zonas negativas que corresponden a partes del modelo no observadas en los datos experimentales, o positivas, debidas a datos experimentales que no están explicados en el modelo.

5.4 Validación del modelo estructural

Parámetros estadísticos

Los indicadores utilizados para medir el buen ajuste del modelo a los datos experimentales son el R_{factor} y el R_{free} (Brunger, 1992; Morris *et al.*, 1992).

R_{factor}

Es un indicador que estima el desacuerdo entre los F_{obs} y los F_{calc}, de manera que indica lo bien que el modelo predice las amplitudes medidas de las reflexiones.

$$R_{\text{factor}} = \frac{\sum_{hkl} \left| |F_{\text{obs}}| - k|F_{\text{calc}}| \right|}{\sum_{hkl} |F_{\text{obs}}|} \cdot 100$$

Este algoritmo de refinamiento puede conllevar a la disminución del factor R sin una mejor calidad del modelo. En general, aunque depende de la resolución, con valores por debajo del 20% son indicativos de la buena calidad del modelo

R_{free}

Mide de forma objetiva el grado en que el modelo atómico construido predice las amplitudes observadas al ser calculado, empleando el mismo algoritmo que el R_{factor}, a partir de un pequeño conjunto de reflexiones tomadas aleatoriamente y no utilizadas en el refinamiento (Brunger, 1992). En general, cuando el modelo trazado da lugar a valores menores de 30% se considera que explica de forma aceptable la estructura cristalográfica 'real', aunque ésta debe cumplir los criterios estereoquímicos (ver más adelante).

Parámetros estereoquímicos

La consistencia del modelo ha de ser contrastada haciendo uso de criterios químicos establecidos, ausencia de choques entre cadenas, distancias y ángulos de enlace adecuados y parámetros térmicos razonables. Además de estos parámetros es importante contrastar que las conformaciones que adoptan los enlaces N-C α y C α -C del enlace peptídico no dan lugar a aproximaciones interatómicas energéticamente improbables.

Las torsiones del enlace peptídico (ϕ , phi y ψ , psi) tienen, en principio, libertad de giro. Sin embargo, cada uno de estos ángulos está energéticamente limitado a ciertos intervalos definidos en el diagrama de *Ramachandran*, que muestra las regiones permitidas de estos ángulos para cada uno de los tipos de estructura secundaria (Ramachandran *et al.*, 1963) (Fig 3.28). Este diagrama fue calculado justo antes de que fuera resuelta la primera estructura atómica de una proteína. Después de cuatro décadas de investigación en este área, existen aun discrepancias con respecto a la forma exacta de las 'regiones permitidas'. El diagrama de Ramachandran se puede generar de dos maneras, una basado en cálculos teóricos y la otra en observaciones experimentales. El diagrama de Ramachandran original, fue construido mediante el análisis estérico de modelos esféricos de los

aminoácidos. Durante estos años de investigación, se han introducido cálculos más sofisticados, pero ninguno refleja satisfactoriamente las observaciones experimentales de los 20 tipos de aminoácido (Ho *et al.*, 2003). Igualmente, el resultado no es totalmente satisfactorio si se usan los datos experimentales para la construcción de un diagrama de Ramachandran. Varios grupos han intentado construir el diagrama usando la información existente en el banco de datos de proteínas (Protein Data Bank PDB) (Walther & Cohen, 1999; Chakrabarti & Pal, 2001). El diagrama experimental más ampliamente utilizado es el del programa PROCHECK (Laskowski *et al.*, 1993). Pero existen también otros diagramas más recientes y diferentes del anterior (Hovmoller *et al.*, 2002; Lovell *et al.*, 2003).

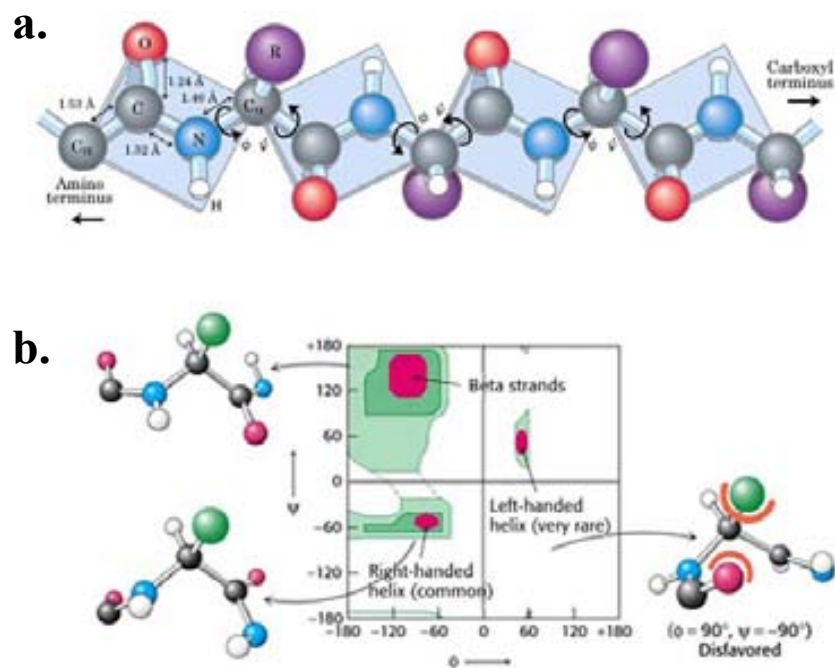


Fig 3.28_ (a) Ángulos de torsión del enlace peptídico. El ángulo ϕ se define como el ángulo de torsión alrededor del enlace $N-C\alpha$, y ψ alrededor del enlace $C\alpha-C$. (b) Diagrama de Ramachandran que muestra las conformaciones más favorables del enlace peptídico en términos de los dos ángulos de torsión.

6_Dispersión a bajo ángulo

La dispersión de rayos X a bajo ángulo (*Small-angle X ray scattering SAXS*) y la dispersión de neutrones a bajo ángulo (*Small-angle neutron scattering SANS*) son las dos técnicas complementarias conocidas como dispersión a bajo ángulo (*small-angle scattering SAS*).

En aplicaciones biológicas esta técnica se utiliza en la determinación de la estructura (promedio de tamaños y formas) de las partículas en una muestra. Se puede también obtener información sobre el ratio superficie-volumen.

Es un método preciso, mayoritariamente no destructivo, aunque las moléculas biológicas son susceptibles al daño por radiación (especialmente SAXS), y que nos permite estudiar el sistema en solución.

Los experimentos de SAS consisten en la exposición de una muestra en solución a los rayos X o neutrones. La dispersión de radiación a “ángulos bajos” –muy cerca del haz primario- de dicha muestra es registrada por un detector. Por esta razón el haz de rayos X o neutrones debe ser altamente colimado. Normalmente los experimentos se llevan a cabo en fuentes de radiación sincrotronal, porque las muestras biológicas están diluidas y tienen una dispersión débil.

La curva de SAS –intensidad versus ángulo de dispersión- se usa para crear un modelo de la proteína a baja resolución (Fig 3.29). Se pueden usar estos datos para encajar en el modelo estructuras de rayos X o RMN.

En un experimento de dispersión, una solución de macromoléculas es expuesta a los rayos X ($\lambda \sim 0.15\text{nm}$) o a un haz de neutrones ($\lambda \sim 0.5\text{nm}$). Si las partículas están distribuidas aleatoriamente y sus orientaciones y posiciones no están correlacionadas, las intensidades de dispersión se suman (no hay interferencias). Este fenómeno da lugar a una distribución isotrópica continua de la intensidad de la muestra, la cual, para partículas monodispersas que no interactúan, es proporcional a la dispersión de una sola partícula promediando todas las orientaciones.

$$I(s) = \langle I(s) \rangle_{\Omega}$$

La intensidad dispersada $I(s)$ es registrada en función del *momentum transfer*, s (Fig 3.29):

$$s = \frac{4\pi}{\lambda} \cdot \sin \theta; \quad 2\theta \text{ es el ángulo entre el haz incidente y la radiación dispersada}$$

La dispersión neta de la partícula es proporcional al cuadrado del contraste -la diferencia de dispersión entre la partícula y el solvente-.

$$I(s) = \left\langle \left| A_a(s) - \rho_s A_s(s) + \delta\rho_b A_b(s) \right|^2 \right\rangle_{\Omega}$$

Donde $A_a(s)$, $A_s(s)$ y $A_b(s)$ son respectivamente las amplitudes de dispersión de la partícula al vacío, del volumen excluido y de la capa de hidratación. ρ_s y ρ_b son las densidades electrónicas del solvente y de la capa de hidratación respectivamente.

En SANS, se puede variar el contraste usando diferentes mezclas de H_2O/D_2O para obtener información adicional.

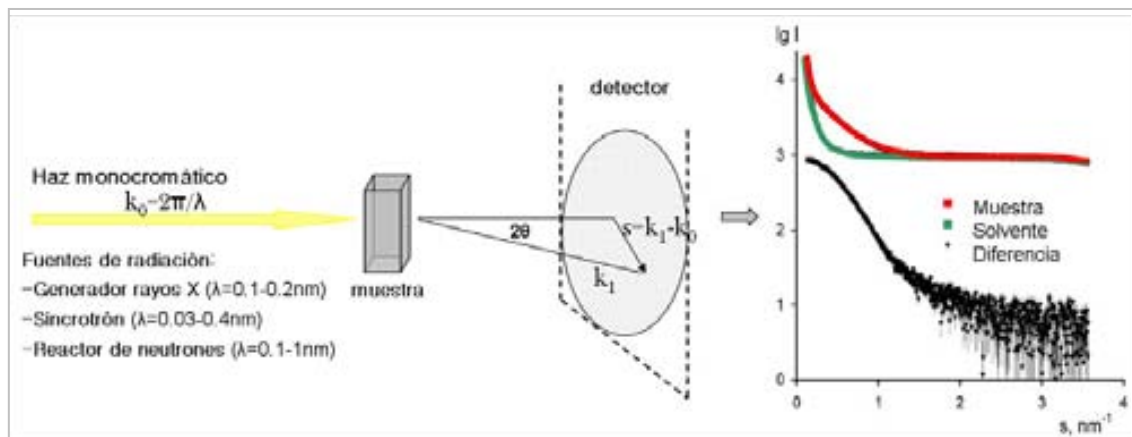


Fig 3.29_ A la izq., representación esquemática de un experimento SAS. A la dcha., patrón típico de dispersión de rayos X de una muestra de BSA. Se debe medir la muestra, el solvente, y la curva resultante es la diferencia entre ambas. A ángulos bajos las curvas decaen rápidamente en función de s , esencialmente según la forma de la partícula. A una resolución media la diferencia es menos pronunciada y a alta resolución las curvas son muy similares. SAS contiene, por tanto, información sobre los grandes rasgos estructurales -forma, estructura cuaternaria y terciaria- pero no es apropiado para el análisis de estructuras atómicas.

Dependiendo del objetivo del experimento la estrategia de medición será diferente. En un experimento de SAS, se suelen medir varias soluciones de la macromolécula de interés a diferentes concentraciones. Extrapolando las curvas de dispersión medidas a diferentes concentraciones a concentración cero, se puede obtener la curva de dispersión de "dilución infinita", donde los efectos de la concentración no deben afectar a la curva de dispersión. Debemos tener en cuenta que a mayores concentraciones la señal es mayor

pero aumentan también los problemas derivados de la interacción de las partículas en solución.

El análisis de los datos empieza con el estudio del principio de la curva, si sigue la aproximación de Guinier la muestra no está agregada (Fig 3.32). La forma de la partícula se puede determinar por varios métodos.

El *small angle scattering* de partículas, puede ser utilizado para determinar la forma de la partícula o su distribución de tamaño. Un patrón de SAS puede ser encajado con intensidades calculadas a partir de modelos de diferentes formas cuando la distribución de tamaño es conocida. Si conocemos la forma, se puede encajar a la intensidad una distribución de tamaño. Una parte importante de la determinación de la forma es la función de distribución $p(r)$, que se calcula a partir de la intensidad utilizando la transformada de Fourier (coordenadas esféricas) (Svergun & Koch, 2003):

$$p(r) = \frac{r^2}{2\pi} \int_0^\infty s^2 I(s) \frac{\sin sr}{sr} ds$$

La forma de la función de distribución está relacionada con la forma de la partícula, si la función es muy simétrica la partícula también.

Información que nos proporciona esta técnica:

- El radio de giro R_g (diferente del radio promedio de la partícula en solución), parámetro característico de la forma y el tamaño, puede ser estimado mediante la aproximación de Guinier.
- La determinación de la intensidad de dispersión a $2\theta=0$, $I(0)$ permite la estimación del peso molecular.
- La transformada de Fourier de la curva de dispersión resulta en la función de distribución que nos permite calcular de una manera más precisa R_g , $I(0)$ y la dimensión máxima lineal de la proteína, D_{max} .
- La función de distribución se puede usar como input para los algoritmos de determinación de estructuras *ab initio*, la cual produce modelos tridimensionales llamados “envolturas” que describen el tamaño y la forma.
- La función de distribución se puede combinar con otras técnicas de alta resolución, como NMR o cristalografía de rayos X para crear una imagen completa y precisa de la macromolécula en solución.

SANS

En un experimento de SANS, la muestra se expone a un haz de neutrones que interactúan con los núcleos de los átomos presentes en la muestra. Esto supone una ventaja sobre las técnicas de rayos X que interactúan muy débilmente con el hidrógeno, el elemento más abundante.

Los neutrones pueden ser considerados como ondas planas con una longitud de onda (0.5-10Å) mucho mayor que el tamaño de los núcleos dispersores. La onda incidente:

$$\psi(z) = \exp(ik_0z); \text{ donde } k_0 \text{ es el vector de onda del neutrón } k_0 = 2\pi / \lambda .$$

La onda de dispersión es una onda esférica:

$$\psi(r) = -\frac{b}{r} \exp(ik \cdot r); \text{ donde } b \text{ es la longitud de dispersión del átomo (Tabla 3.30).}$$

Normalmente b es un número real, lo que significa que no hay absorción o ésta es negligible (Jacrot, 1976).

H	D	C	N	O	P	S
-0.3742	0.6671	0.6651	0.940	0.5804	0.517	0.2847

Tabla 2.30_ Longitudes de dispersión de varios elementos (en 10⁻¹²cm)

La técnica de *variación de contraste* se basa en la diferente dispersión del hidrógeno y el deuterio. En los sistemas biológicos el hidrógeno puede ser intercambiado por deuterio, que normalmente tiene un efecto mínimo en la muestra pero un gran efecto en la dispersión de ésta.

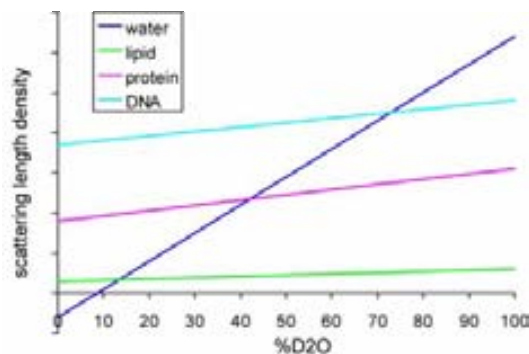


Fig 3.31_ Muestra la longitud de dispersión del agua y de varias macromoléculas biológicas en función de la concentración de deuterio. Las muestras biológicas se disuelven normalmente en agua, así que sus hidrógenos se pueden intercambiar con los deuterios del solvente. La dispersión de una molécula depende del ratio de hidrógeno:deuterio ya que su dispersión depende de todos sus componentes. Existen ciertos ratios de H₂O:D₂O donde la dispersión de la molécula es equivalente a la del solvente, de manera que su dispersión queda anulada cuando la dispersión del

buffer se sustrae a la de la muestra. Por ejemplo a una concentración de 40-45% de D₂O, la dispersión de una proteína es indistinguible de la del buffer.

6.1 Preparación y medida de las muestras de mTERF y el complejo con DNA

Para los experimentos de SAXS y SANS se añadió 10% de glicerol en todos los pasos de purificación. La cromatografía final en exclusión molecular -último paso de la purificación de la proteína- se realizó con el tampón E (50mM citrato sódico pH6.0, 1M NaCl, 10% glicerol, 10mM BME).

El complejo proteína-DNA se obtuvo mezclando mTERF-NHis o - Δ N (0.25 mg/mL) con el DNA de 28bp (oligo Tabla 3.8) a un ratio 1:1 y dializándolo en membranas de diálisis (Spectra/Por) en 7 pasos, hasta una concentración final de sal de 25 mM. La muestra resultante se inyectó en una columna de gelfiltración (Superdex 200) equilibrada previamente con el tampón F (50 mM citrato sódico pH 6.0, 25 mM NaCl, 50 mM MgCl₂, 10 mM BME y 10% glicerol).

En ambas muestras, proteína y complejo, con las fracciones eluidas de la columna se preparó una batería de muestras a distintas concentraciones. Todas las muestras se filtraron (0.1 nm tamaño poro, Millipore) justo antes de ser medidas.

Para los experimentos de SAS, es muy importante que la muestra sea monodispersa. Se utilizó un aparato de *dynamic light scattering* (DLS) para optimizar y comprobar la monodispersidad de las muestras. Esta técnica mide la luz (láser) dispersada por las macromoléculas disueltas o las partículas en suspensión de una muestra en periodos de tiempo muy cortos. Dado que el movimiento de las moléculas es función de su tamaño, el resultado es una función de correlación que permite calcular el coeficiente de difusión y la distribución del radio hidrodinámico de las moléculas. La polidispersidad es la desviación estándar relativa de la muestra. Según este parámetro se pueden distinguir: muestra monodispersa si la polidispersidad es menor del 20%, muestra medio dispersa si la polidispersidad es entre 20-30% y muestra polidispersa si la polidispersidad es mayor del 30%.

Los datos de SAXS fueron tomados en el sincrotrón de Grenoble (ID14.3) y en el sincrotrón de Hamburgo (DORIS X33) -la intensidad del haz es menor y por

tanto es menor también el daño por radiación-. Todas las medidas se tomaron a 10° C. Para cada medición se deben tomar 2 blancos, antes y después de la medida de la muestra, el promedio de los cuales se debe sustraer a la curva de la muestra.

En Grenoble, debido al daño por radiación que se observa sobretodo en las muestras de proteína sola, la estrategia de recogida consistió en medir de cada muestra dos recogidas de 10 repeticiones. La primera de 3seg por imagen para obtener la primera parte de la curva -es en la que se observa este daño- y la segunda de 30seg por imagen para obtener la segunda parte.

En Hamburgo, las mediciones de muestra o tampón fueron de 30seg y se realizaron por cuadruplicado.

SANS

En este caso, quisimos utilizar la variación de contraste en el estudio de los complejos mTERF-NHis y Δ N con el DNA de 28bp. Así pues se realizaron antes de la medida para cada concentración de complejo tres diálisis o/n con diferentes proporciones de agua deuterada. Medimos cuatro muestras por concentración, una sin agua deuterada, otra al 40% -donde la proteína es invisible-, otra al 80% -donde el DNA es invisible- y otra al 100% de D₂O. Las concentraciones medidas fueron de 2, 6, 8 y 14mg·ml⁻¹ para el complejo de mTERF-NHis y DNA_{TER}(28pb) y 3, 5.5 y 8mg·ml⁻¹ para el complejo de mTERF- Δ N y DNA_{TER}(28pb).

La medición se realizó en el ILL en Grenoble.

6.2 Procesamiento

A las curvas medidas de cada muestra, resultantes de la media de los replicados, se les debe sustraer la media de los blancos medidos antes y después. En el caso que la muestra presente daño por radiación -como las muestras de proteína medidas en el sincrotrón de Grenoble-, éste afecta al primer tramo de la curva. Por tanto, la curva experimental se obtiene combinando la primera parte del promedio de las

curvas medidas a 3s (s entre 0 y ~0.5) y el resto de las curvas medidas a 30s. Las muestras del complejo, en cambio, no mostraron daño por radiación.

La curva resultante debe se escaló según la concentración de la muestra. Estas operaciones fueron realizadas mediante PRIMUS (Konarev *et al.*, 2003) usando protocolos estándar.

Los valores de $I(0)$ y el radio de giro R_g , fueron calculados mediante la aproximación de Guinier (Fig 3.32).

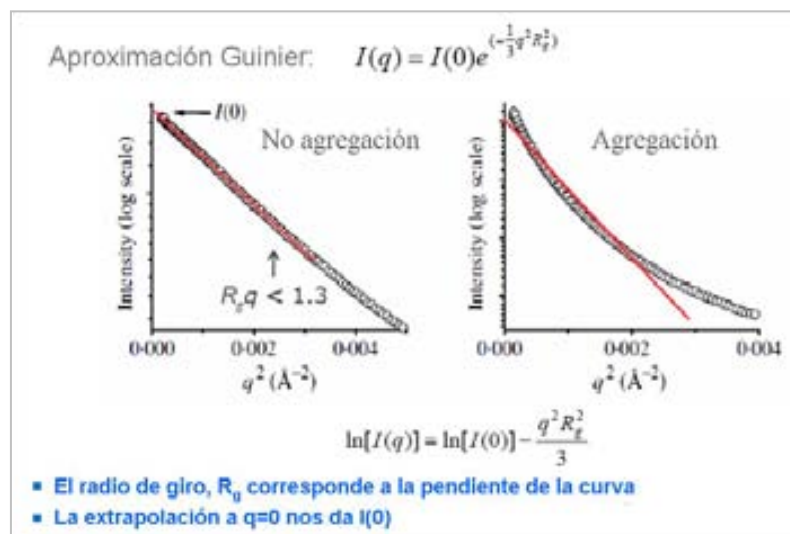
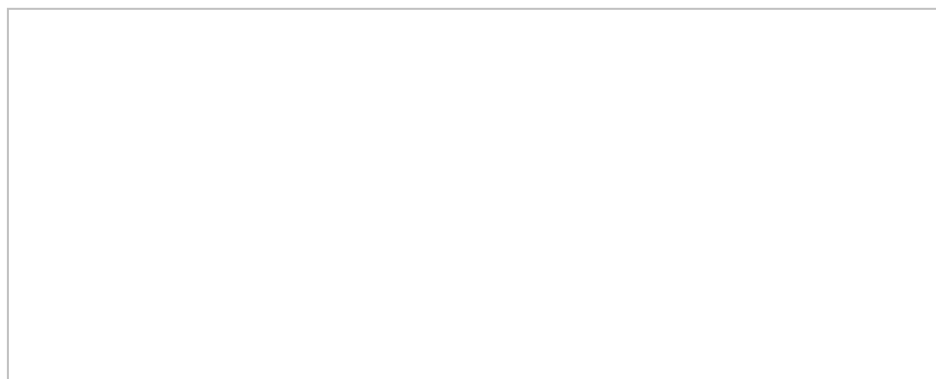


Fig 3.32_ Según la aproximación de Guinier, para valores bajos de q ($< 1.3/R_g$), se encuentra que la representación $\log(I(q))$ vs. q^2 muestra comportamientos lineales, de manera que la pendiente corresponde al radio de giro R_g y la extrapolación a $q=0$ nos da $I(0)$. (Izq.) Si la curva experimental no muestra un comportamiento lineal a valores bajos de q puede indicar que hay agregación en la muestra.

La dimensión máxima de la partícula D_{\max} , y la función de distribución $\rho(r)$, se calcularon a partir de los patrones de dispersión mediante el programa GNOM (Svergun, 1992). El peso molecular de las muestras se estimó a partir de la dispersión de éstas en referencia a la dispersión de una solución de referencia medida en las mismas condiciones de BSA (3.63 mg ml^{-1}).



- **Calibración con una proteína estándar**

$$MW_{\text{sample}} = I(0)_{\text{sample}} \times \frac{MW_{\text{standard}}}{I(0)_{\text{standard}}}$$

- **Medida precisa de la concentración de proteína**
- **Se asume que el volumen específico parcial es el mismo**
- **El error es ~10%**

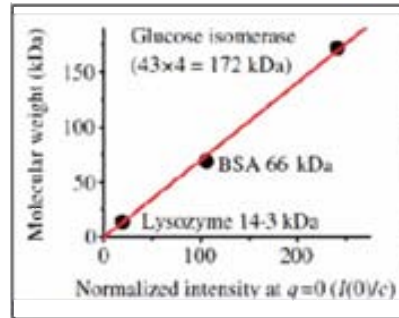


Fig 3.33_Requerimientos para la estimación del peso molecular de una muestra. La intensidad se debe normalizar con respecto a la concentración $I(0)/\text{Conc}$.

En este caso concreto nos interesaba estudiar el complejo en solución con el oligonucleótido DNA_{TER} completo (28pb) y ver si éste pasaba a través de la proteína, ya que el modelo cristalográfico de una molécula de mTERF unida a dos oligonucleótidos idénticos, no acababa de explicar la interacción descrita de la proteína con el DNA mitocondrial continuo. Se generaron modelos en base a las estructuras cristalográficas de mTERF-ΔN-DNA12pb, donde los dos DNAs simétricos unidos a la proteína están en el mismo eje, y de mTERF-NHis-DNA15pb, donde los dos DNAs simétricos se encuentran relacionados por un ángulo de ~36°. Dichos modelos se generaron deslizando una molécula DNA de 28pb, curvado en el caso de mTERF-NHis y recto en el caso de mTERF-ΔN, a lo largo del eje que forman los dos DNAs que une cada molécula de mTERF en el cristal (Fig 3.34). El modelo de B-DNA curvado se realizó con el programa 3D-DART (Van Dijk & Bonvin, 2009) y la superposición y los modelos se generaron con Coot. El ajuste entre los modelos y las curvas experimentales se evaluó con el programa CRY SOL (Svergun, 1995). En este caso se generaron como *input* 50 y 30 pdb's modelo, para los complejos de mTERF-NHis y mTERF-ΔN respectivamente. El programa calcula el patrón de dispersión del promedio esférico de la macromolécula teniendo en cuenta también la capa de hidratación. Dada una curva experimental de SAXS, CRY SOL puede encajar la curva teórica de dispersión correspondiente a los modelos generados, promediando el volumen de solvente desplazado, contrastando la capa de hidratación y comparando el fondo, para minimizar la discrepancia entre curvas.

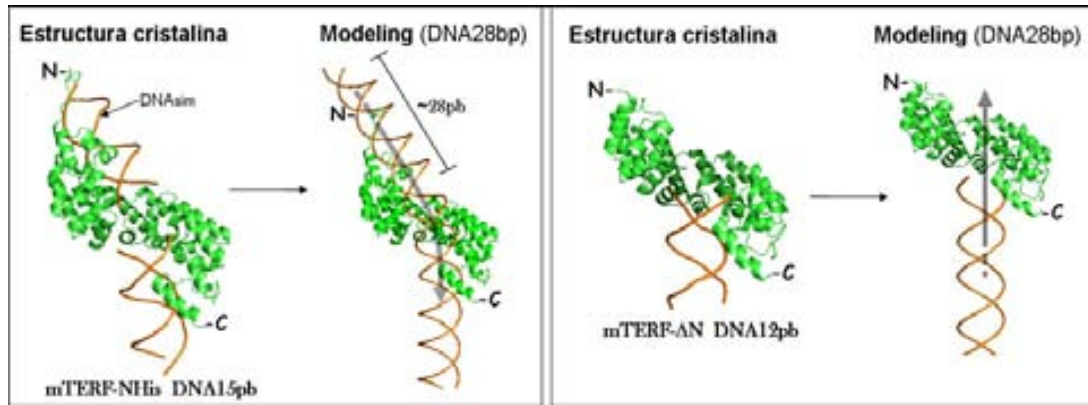


Fig 3.34_ Generación de los modelos para SAXS mediante el deslizamiento de una molécula DNA de 28pb a lo largo del eje que forman los dos DNAs que une cada molécula de mTERF en el cristal. (Izq.) Modelos para mTERF-NHis. La flecha indica cómo se desliza el DNA de 28pb en cada modelo. (Dcha.) Modelos para mTERF-ΔN. En este caso, en los modelos, el DNA de 28pb se desliza en el eje vertical.

Bibliografia

- Abrahams, J.P. & Leslie, A.G. (1996) Methods used in the structure determination of bovine mitochondrial f1 atpase. *Acta Crystallogr D Biol Crystallogr*, **52**, 30-42
- Abrahams, J.P., Leslie, A.G., Lutter, R. & Walker, J.E. (1994) Structure at 2.8 a resolution of f1-atpase from bovine heart mitochondria. *Nature*, **370**, 621-8
- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C. & Zwart, P.H. Phenix: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, **66**, 213-21
- Agniswamy, J., Joyce, M.G., Hammer, C.H. & Sun, P.D. (2008) Towards a rational approach for heavy-atom derivative screening in protein crystallography. *Acta Crystallogr D Biol Crystallogr*, **64**, 354-67
- Albring, M., Griffith, J. & Attardi, G. (1977) Association of a protein structure of probable membrane derivation with hela cell mitochondrial DNA near its origin of replication. *Proc Natl Acad Sci U S A*, **74**, 1348-52
- Aloni, Y. & Attardi, G. (1971) Symmetrical in vivo transcription of mitochondrial DNA in hela cells. *Proc Natl Acad Sci U S A*, **68**, 1757-61
- Amalric, F., Merkel, C., Gelfand, R. & Attardi, G. (1978) Fractionation of mitochondrial rna from hela cells by high-resolution electrophoresis under strongly denaturing conditions. *J Mol Biol*, **118**, 1-25
- Andersen, G.R., Thirup, S., Spemulli, L.L. & Nyborg, J. (2000) High resolution crystal structure of bovine mitochondrial ef-tu in complex with gdp. *J Mol Biol*, **297**, 421-36
- Anderson, S., Bankier, A.T., Barrell, B.G., De Bruijn, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J., Staden, R. & Young, I.G. (1981) Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457-65
- Andersson, U. & Scarpulla, R.C. (2001) Pgc-1-related coactivator, a novel, serum-inducible coactivator of nuclear respiratory factor 1-dependent transcription in mammalian cells. *Mol Cell Biol*, **21**, 3738-49
- Andrade, M.A., Petosa, C., O'donoghue, S.I., Muller, C.W. & Bork, P. (2001) Comparison of arm and heat protein repeats. *J Mol Biol*, **309**, 1-18
- Andreu, A.L. & Gonzalo-Sanz, R. (2004) [mitochondrial disorders: A classification for the 21st century]. *Neurologia*, **19**, 15-22
- Arnoult, D. (2007) Mitochondrial fragmentation in apoptosis. *Trends Cell Biol*, **17**, 6-12
- Asin-Cayuela, J., Helm, M. & Attardi, G. (2004) A monomer-to-trimer transition of the human mitochondrial transcription termination factor (mterf) is associated with a loss of in vitro activity. *J Biol Chem*, **279**, 15670-7
- Asin-Cayuela, J., Schwend, T., Farge, G. & Gustafsson, C.M. (2005) The human mitochondrial transcription termination factor (mterf) is fully active in vitro in the non-phosphorylated form. *J Biol Chem*, **280**, 25499-505
- Asin-Cayuela, J.S., Y Gustafsson, Cm (2008) Initial characterization of mterf4, a paralogue of mterf1 (mterf). In: *EUROMIT 7*, p. pg 20, Stockholm, Sweden
- Attardi, G. (1985) Animal mitochondrial DNA: An extreme example of genetic economy. *Int Rev Cytol*, **93**, 93-145
- Baker, N.A., Sept, D., Joseph, S., Holst, M.J. & Mccammon, J.A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci U S A*, **98**, 10037-41
- Banerjee, S., Chalissery, J., Bandey, I. & Sen, R. (2006) Rho-dependent transcription termination: More questions than answers. *J Microbiol*, **44**, 11-22
- Barrell, B.G., Bankier, A.T. & Drouin, J. (1979) A different genetic code in human mitochondria. *Nature*, **282**, 189-94

- Bartsch, I., Schoneberg, C. & Grummt, I. (1988) Purification and characterization of ttf1, a factor that mediates termination of mouse ribosomal DNA transcription. *Mol Cell Biol*, **8**, 3891-7
- Bellance, N., Lestienne, P. & Rossignol, R. (2009) Mitochondria: From bioenergetics to the metabolic regulation of carcinogenesis. *Front Biosci*, **14**, 4015-34
- Berdanier, C.D. (2005) *Mitochondria in health and disease*. Taylor & Francis/CRC Press, Boca Raton, FL.
- Bereiter-Hahn, J. & Voth, M. (1994) Dynamics of mitochondria in living cells: Shape changes, dislocations, fusion, and fission of mitochondria. *Microsc Res Tech*, **27**, 198-219
- Berg, O.G. & Kurland, C.G. (2000) Why mitochondrial genes are most often found in nuclei. *Mol Biol Evol*, **17**, 951-61
- Bergfors, T. (2003) Seeds to crystals. *J Struct Biol*, **142**, 66-76
- Berthier, F., Renaud, M., Alziari, S. & Durand, R. (1986) Rna mapping on drosophila mitochondrial DNA: Precursors and template strands. *Nucleic Acids Res*, **14**, 4519-33
- Bogenhagen, D., Gillum, A.M., Martens, P.A. & Clayton, D.A. (1979) Replication of mouse l-cell mitochondrial DNA. *Cold Spring Harb Symp Quant Biol*, **43 Pt 1**, 253-62
- Bogenhagen, D.F., Rousseau, D. & Burke, S. (2008) The layered structure of human mitochondrial DNA nucleoids. *J Biol Chem*, **283**, 3665-75
- Bogenhagen, D.F., Wang, Y., Shen, E.L. & Kobayashi, R. (2003) Protein components of mitochondrial DNA nucleoids in higher eukaryotes. *Mol Cell Proteomics*, **2**, 1205-16
- Boore, J.L. (1999) Animal mitochondrial genomes. *Nucleic Acids Res*, **27**, 1767-80
- Bowmaker, M., Yang, M.Y., Yasukawa, T., Reyes, A., Jacobs, H.T., Huberman, J.A. & Holt, I.J. (2003) Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone. *J Biol Chem*, **278**, 50961-9
- Boyer, P.D. (1997) The atp synthase--a splendid molecular machine. *Annu Rev Biochem*, **66**, 717-49
- Brewer, B.J. & Fangman, W.L. (1991) Mapping replication origins in yeast chromosomes. *Bioessays*, **13**, 317-22
- Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003) Generation, representation and flow of phase information in structure determination: Recent developments in and around sharp 2.0. *Acta Crystallogr D Biol Crystallogr*, **59**, 2023-30
- Brown, T.A. (2006) "synthesis and processing of rna." *genomes* 3rd edn.
- Brunger, A.T. (1992) Free r value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472-5
- Camasamudram, V., Fang, J.K. & Avadhani, N.G. (2003) Transcription termination at the mouse mitochondrial h-strand promoter distal site requires an a/t rich sequence motif and sequence specific DNA binding proteins. *Eur J Biochem*, **270**, 1128-40
- Cantatore, P. & Attardi, G. (1980) Mapping of nascent light and heavy strand transcripts on the physical map of hela cell mitochondrial DNA. *Nucleic Acids Res*, **8**, 2605-25
- Cantatore, P., Flagella, Z., Fracasso, F., Lezza, A.M., Gadaleta, M.N. & De Montalvo, A. (1987) Synthesis and turnover rates of four rat liver mitochondrial rna species. *FEBS Lett*, **213**, 144-8
- Cantatore, P., Roberti, M., Loguercio Polosa, P., Mustich, A. & Gadaleta, M.N. (1990) Mapping and characterization of paracentrotus lividus mitochondrial transcripts: Multiple and overlapping transcription units. *Curr Genet*, **17**, 235-45
- Cantatore, P., Roberti, M., Rainaldi, G., Gadaleta, M.N. & Saccone, C. (1989) The complete nucleotide sequence, gene organization, and genetic code of the mitochondrial genome of paracentrotus lividus. *J Biol Chem*, **264**, 10965-75
- Capaldi, R.A. & Aggeler, R. (2002) Mechanism of the f(1)f(0)-type atp synthase, a biological rotary motor. *Trends Biochem Sci*, **27**, 154-60
- Carmona, M.C., Iglesias, R., Obregon, M.J., Darlington, G.J., Villarroya, F. & Giralto, M. (2002) Mitochondrial biogenesis and thyroid status maturation in brown fat require ccaat/enhancer-binding protein alpha. *J Biol Chem*, **277**, 21489-98
- Casas, F., Pesseme, L., Grandemange, S., Seyer, P., Gueguen, N., Baris, O., Lepourry, L., Cabello, G. & Wrutniak-Cabello, C. (2008) Overexpression of the mitochondrial t3 receptor p43 induces a shift in skeletal muscle fiber types. *PLoS One*, **3**, e2501
- Casas, F., Rochard, P., Rodier, A., Cassar-Malek, I., Marchal-Victorion, S., Wiesner, R.J., Cabello, G. & Wrutniak, C. (1999) A variant form of the nuclear triiodothyronine receptor c-erbaalpha1 plays a direct role in regulation of mitochondrial rna synthesis. *Mol Cell Biol*, **19**, 7913-24
- Cingolani, G., Petosa, C., Weis, K. & Muller, C.W. (1999) Structure of importin-beta bound to the ibb domain of importin-alpha. *Nature*, **399**, 221-9
- Clayton, D.A. (1982) Replication of animal mitochondrial DNA. *Cell*, **28**, 693-705
- Clayton, D.A. (2003) Mitochondrial DNA replication: What we know. *IUBMB Life*, **55**, 213-7

- Coates, J.C. (2003) Armadillo repeat proteins: Beyond the animal kingdom. *Trends Cell Biol*, **13**, 463-71
- Collins, T.J. & Bootman, M.D. (2003) Mitochondria are morphologically heterogeneous within cells. *J Exp Biol*, **206**, 1993-2000
- Conti, E. & Kuriyan, J. (2000) Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure*, **8**, 329-38
- Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. (1998) Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell*, **94**, 193-204
- Cosa, G., Focsaneanu, K.S., Mclean, J.R., Mcnamee, J.P. & Scaiano, J.C. (2001) Photophysical properties of fluorescent DNA-dyes bound to single- and double-stranded DNA in aqueous buffered solution. *Photochem Photobiol*, **73**, 585-99
- Cowtan, K. (1994) Dmmulti: A multi-crystal density modification package. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, 34-38
- Cuezva, J.M., Ostronoff, L.K., Ricart, J., Lopez De Heredia, M., Di Liegro, C.M. & Izquierdo, J.M. (1997) Mitochondrial biogenesis in the liver during development and oncogenesis. *J Bioenerg Biomembr*, **29**, 365-77
- Chakrabarti, P. & Pal, D. (2001) The interrelationships of side-chain and main-chain conformations in proteins. *Prog Biophys Mol Biol*, **76**, 1-102
- Chang, D.D. & Clayton, D.A. (1984) Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. *Cell*, **36**, 635-43
- Chang, D.D. & Clayton, D.A. (1985) Priming of human mitochondrial DNA replication occurs at the light-strand promoter. *Proc Natl Acad Sci U S A*, **82**, 351-5
- Chang, D.D. & Clayton, D.A. (1987) A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. *EMBO J*, **6**, 409-17
- Chang, D.D. & Clayton, D.A. (1989) Mouse mrnaase mrp rna is encoded by a nuclear gene and contains a decamer sequence complementary to a conserved region of mitochondrial rna substrate. *Cell*, **56**, 131-9
- Chang, D.D., Hauswirth, W.W. & Clayton, D.A. (1985) Replication priming and transcription initiate from precisely the same site in mouse mitochondrial DNA. *EMBO J*, **4**, 1559-67
- Chen, H., Chomyn, A. & Chan, D.C. (2005a) Disruption of fusion results in mitochondrial heterogeneity and dysfunction. *J Biol Chem*, **280**, 26185-92
- Chen, H., Mccaffery, J.M. & Chan, D.C. (2007) Mitochondrial fusion protects against neurodegeneration in the cerebellum. *Cell*, **130**, 548-62
- Chen, X.J. & Butow, R.A. (2005) The organization and inheritance of the mitochondrial genome. *Nat Rev Genet*, **6**, 815-25
- Chen, Y., Zhou, G., Yu, M., He, Y., Tang, W., Lai, J., He, J., Liu, W. & Tan, D. (2005b) Cloning and functional analysis of human mterfl encoding a novel mitochondrial transcription termination factor-like protein. *Biochem Biophys Res Commun*, **337**, 1112-8
- Chipuk, J.E. & Green, D.R. (2005) Do inducers of apoptosis trigger caspase-independent cell death? *Nat Rev Mol Cell Biol*, **6**, 268-75
- Chomyn, A., Enriquez, J.A., Micol, V., Fernandez-Silva, P. & Attardi, G. (2000) The mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episode syndrome-associated human mitochondrial trnaleu(uur) mutation causes aminoacylation deficiency and concomitant reduced association of mrna with ribosomes. *J Biol Chem*, **275**, 19198-209
- Chomyn, A., Martinuzzi, A., Yoneda, M., Daga, A., Hurko, O., Johns, D., Lai, S.T., Nonaka, I., Angelini, C. & Attardi, G. (1992) Melas mutation in mtdna binding site for transcription termination factor causes defects in protein synthesis and in respiration but no change in levels of upstream and downstream mature transcripts. *Proc Natl Acad Sci U S A*, **89**, 4221-5
- Chook, Y.M. & Blobel, G. (1999) Structure of the nuclear transport complex karyopherin-beta2-ran x gppnhp. *Nature*, **399**, 230-7
- Christianson, T.W. & Clayton, D.A. (1986) In vitro transcription of human mitochondrial DNA: Accurate termination requires a region of DNA sequence that can function bidirectionally. *Proc Natl Acad Sci U S A*, **83**, 6277-81
- Christianson, T.W. & Clayton, D.A. (1988) A tridecamer DNA sequence supports human mitochondrial rna 3'-end formation in vitro. *Mol Cell Biol*, **8**, 4502-9
- Daga, A., Micol, V., Hess, D., Aebersold, R. & Attardi, G. (1993) Molecular characterization of the transcription termination factor from human mitochondria. *J Biol Chem*, **268**, 8123-30
- Dairaghi, D.J., Shadel, G.S. & Clayton, D.A. (1995a) Addition of a 29 residue carboxyl-terminal tail converts a simple hmg box-containing protein into a transcriptional activator. *J Mol Biol*, **249**, 11-28

- Dairaghi, D.J., Shadel, G.S. & Clayton, D.A. (1995b) Human mitochondrial transcription factor a and promoter spacing integrity are required for transcription initiation. *Biochim Biophys Acta*, **1271**, 127-34
- Dauter, Z., Dauter, M. & Rajashankar, K.R. (2000) Novel approach to phasing proteins: Derivatization by short cryo-soaking with halides. *Acta Crystallogr D Biol Crystallogr*, **56**, 232-7
- Davis, K.A. & Hatefi, Y. (1972) Resolution and reconstitution of complex ii (succinate-ubiquinone reductase) by salts. *Arch Biochem Biophys*, **149**, 505-12
- Denslow, N.D., Michaels, G.S., Montoya, J., Attardi, G. & O'Brien, T.W. (1989) Mechanism of mrna binding to bovine mitochondrial ribosomes. *J Biol Chem*, **264**, 8328-38
- Detmer, S.A. & Chan, D.C. (2007a) Complementation between mouse mfn1 and mfn2 protects mitochondrial fusion defects caused by cmt2a disease mutations. *J Cell Biol*, **176**, 405-14
- Detmer, S.A. & Chan, D.C. (2007b) Functions and dysfunctions of mitochondrial dynamics. *Nat Rev Mol Cell Biol*, **8**, 870-9
- Diederichs, K. & Karpplus, P.A. (1997) Improved r-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol*, **4**, 269-75
- Doda, J.N., Wright, C.T. & Clayton, D.A. (1981) Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. *Proc Natl Acad Sci U S A*, **78**, 6116-20
- Dodson, E. (2003) Is it jolly sad? *Acta Crystallogr D Biol Crystallogr*, **59**, 1958-65
- Doersen, C.J., Guerrier-Takada, C., Altman, S. & Attardi, G. (1985) Characterization of an rnase p activity from hela cell mitochondria. Comparison with the cytosol rnase p activity. *J Biol Chem*, **260**, 5942-9
- Dubin, D.T., Montoya, J., Timko, K.D. & Attardi, G. (1982) Sequence analysis and precise mapping of the 3' ends of hela cell mitochondrial ribosomal rnas. *J Mol Biol*, **157**, 1-19
- Edwards, T.A., Pyle, S.E., Wharton, R.P. & Aggarwal, A.K. (2001) Structure of pumilio reveals similarity between rna and peptide binding motifs. *Cell*, **105**, 281-9
- Emsley, P. & Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-32
- Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of coot. *Acta Crystallogr D Biol Crystallogr*, **66**, 486-501
- Enriquez, J.A., Fernandez-Silva, P., Garrido-Perez, N., Lopez-Perez, M.J., Perez-Martos, A. & Montoya, J. (1999) Direct regulation of mitochondrial rna synthesis by thyroid hormone. *Mol Cell Biol*, **19**, 657-70
- Enriquez, J.A., Fernandez-Silva, P., Perez-Martos, A., Lopez-Perez, M.J. & Montoya, J. (1996) The synthesis of mrna in isolated mitochondria can be maintained for several hours and is inhibited by high levels of atp. *Eur J Biochem*, **237**, 601-10
- Ericsson, U.B., Hallberg, B.M., Detitta, G.T., Dekker, N. & Nordlund, P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem*, **357**, 289-98
- Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr*, **62**, 72-82
- Falkenberg, M., Gaspari, M., Rantanen, A., Trifunovic, A., Larsson, N.G. & Gustafsson, C.M. (2002) Mitochondrial transcription factors b1 and b2 activate transcription of human mtdna. *Nat Genet*, **31**, 289-94
- Fan, L., Sanschagrin, P.C., Kaguni, L.S. & Kuhn, L.A. (1999) The accessory subunit of mtdna polymerase shares structural homology with aminoacyl-trna synthetases: Implications for a dual role as a primer recognition factor and processivity clamp. *Proc Natl Acad Sci U S A*, **96**, 9527-32
- Feagin, J.E. (2000) Mitochondrial genome diversity in parasites. *Int J Parasitol*, **30**, 371-90
- Fernandez-Silva, P., Enriquez, J.A. & Montoya, J. (2003) Replication and transcription of mammalian mitochondrial DNA. *Exp Physiol*, **88**, 41-56
- Fernandez-Silva, P., Martinez-Azorin, F., Micol, V. & Attardi, G. (1997) The human mitochondrial transcription termination factor (mterf) is a multizipper protein but binds to DNA as a monomer, with evidence pointing to intramolecular leucine zipper interactions. *EMBO J*, **16**, 1066-79
- Fernandez-Silva, P., Polosa, P.L., Roberti, M., Di Ponzio, B., Gadaleta, M.N., Montoya, J. & Cantatore, P. (2001) Sea urchin mtdbp is a two-faced transcription termination factor with a biased polarity depending on the rna polymerase. *Nucleic Acids Res*, **29**, 4736-43
- Fish, J., Raule, N. & Attardi, G. (2004) Discovery of a major d-loop replication origin reveals two modes of human mtdna synthesis. *Science*, **306**, 2098-101

- Fisher, R.P. & Clayton, D.A. (1985) A transcription factor required for promoter recognition by human mitochondrial rna polymerase. Accurate initiation at the heavy- and light-strand promoters dissected and reconstituted in vitro. *J Biol Chem*, **260**, 11330-8
- Fisher, R.P. & Clayton, D.A. (1988) Purification and characterization of human mitochondrial transcription factor 1. *Mol Cell Biol*, **8**, 3496-509
- Fisher, R.P., Lisowsky, T., Parisi, M.A. & Clayton, D.A. (1992) DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. *J Biol Chem*, **267**, 3358-67
- Fisher, R.P., Topper, J.N. & Clayton, D.A. (1987) Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. *Cell*, **50**, 247-58
- Frey, T.G., Renken, C.W. & Perkins, G.A. (2002) Insight into mitochondrial structure and function from electron tomography. *Biochim Biophys Acta*, **1555**, 196-203
- Fuste, J.M., Wanrooij, S., Jemt, E., Granycome, C.E., Cluett, T.J., Shi, Y., Atanassova, N., Holt, I.J., Gustafsson, C.M. & Falkenberg, M. Mitochondrial rna polymerase is needed for activation of the origin of light-strand DNA replication. *Mol Cell*, **37**, 67-78
- Gaines, G., Rossi, C. & Attardi, G. (1987) Markedly different atp requirements for rna synthesis and mtdna light strand transcription versus mrna synthesis in isolated human mitochondria. *J Biol Chem*, **262**, 1907-15
- Gardner, L.P., Mookhtiar, K.A. & Coleman, J.E. (1997) Initiation, elongation, and processivity of carboxyl-terminal mutants of t7 rna polymerase. *Biochemistry*, **36**, 2908-18
- Garman, E. & Murray, J.W. (2003) Heavy-atom derivatization. *Acta Crystallogr D Biol Crystallogr*, **59**, 1903-13
- Garrido, N., Griparic, L., Jokitalo, E., Wartiovaara, J., Van Der Blik, A.M. & Spelbrink, J.N. (2003) Composition and dynamics of human mitochondrial nucleoids. *Mol Biol Cell*, **14**, 1583-96
- Gaspari, M., Larsson, N.G. & Gustafsson, C.M. (2004) The transcription machinery in mammalian mitochondria. *Biochim Biophys Acta*, **1659**, 148-52
- Gasteiger E., H.C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; (2005) *Protein identification and analysis tools on the expasy server*.
- Gelfand, R. & Attardi, G. (1981) Synthesis and turnover of mitochondrial ribonucleic acid in hela cells: The mature ribosomal and messenger ribonucleic acid species are metabolically unstable. *Mol Cell Biol*, **1**, 497-511
- Georgescu, R.E., Kim, S.S., Yurieva, O., Kuriyan, J., Kong, X.P. & O'donnell, M. (2008) Structure of a sliding clamp on DNA. *Cell*, **132**, 43-54
- Gerber, J.K., Gogel, E., Berger, C., Wallisch, M., Muller, F., Grummt, I. & Grummt, F. (1997) Termination of mammalian rdna replication: Polar arrest of replication fork movement by transcription termination factor ttf-i. *Cell*, **90**, 559-67
- Ghivizzani, S.C., Madsen, C.S., Nelen, M.R., Ammini, C.V. & Hauswirth, W.W. (1994) In organello footprint analysis of human mitochondrial DNA: Human mitochondrial transcription factor a interactions at the origin of replication. *Mol Cell Biol*, **14**, 7717-30
- Glusker, J.P. (1994) X-ray crystallography of proteins. *Methods Biochem Anal*, **37**, 1-72
- Goldbach, R.W. (1977) The structure of tetrahymena pyriformis mitochondrial DNA. The alkali stability of the DNA. *FEBS Lett*, **80**, 157-60
- Gray, M.W., Lang, B.F., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., Brossard, N., Delage, E., Littlejohn, T.G., Plante, I., Rioux, P., Saint-Louis, D., Zhu, Y. & Burger, G. (1998) Genome structure and gene content in protist mitochondrial dnas. *Nucleic Acids Res*, **26**, 865-78
- Green, D.R. & Reed, J.C. (1998) Mitochondria and apoptosis. *Science*, **281**, 1309-12
- Grigorieff, N. (1998) Three-dimensional structure of bovine nadh:Ubiquinone oxidoreductase (complex i) at 2.2 Å in ice. *J Mol Biol*, **277**, 1033-46
- Grohmann, K., Amairic, F., Crews, S. & Attardi, G. (1978) Failure to detect "Cap" Structures in mitochondrial DNA-coded poly(a)-containing rna from hela cells. *Nucleic Acids Res*, **5**, 637-51
- Groves, M.R., Hanlon, N., Turowski, P., Hemmings, B.A. & Barford, D. (1999) The structure of the protein phosphatase 2a pr65/a subunit reveals the conformation of its 15 tandemly repeated heat motifs. *Cell*, **96**, 99-110
- Grummt, I., Kuhn, A., Bartsch, I. & Rosenbauer, H. (1986a) A transcription terminator located upstream of the mouse rdna initiation site affects rrna synthesis. *Cell*, **47**, 901-11
- Grummt, I., Rosenbauer, H., Niedermeyer, I., Maier, U. & Ohrlein, A. (1986b) A repeated 18 bp sequence motif in the mouse rdna spacer mediates binding of a nuclear factor and transcription termination. *Cell*, **45**, 837-46

- Hackenbrock, C.R. (1968) Ultrastructural bases for metabolically linked mechanical activity in mitochondria. II. Electron transport-linked ultrastructural transformations in mitochondria. *J Cell Biol*, **37**, 345-69
- Hackenbrock, C.R., Rehn, T.G., Weinbach, E.C. & Lemasters, J.J. (1971) Oxidative phosphorylation and ultrastructural transformation in mitochondria in the intact ascites tumor cell. *J Cell Biol*, **51**, 123-37
- Harp, J.M., Hanson, B.L., Timm, D.E. & Bunick, G.J. (1999) Macromolecular crystal annealing: Evaluation of techniques and variables. *Acta Crystallogr D Biol Crystallogr*, **55**, 1329-34
- Harp, J.M., Timm, D.E. & Bunick, G.J. (1998) Macromolecular crystal annealing: Overcoming increased mosaicity associated with cryocrystallography. *Acta Crystallogr D Biol Crystallogr*, **54**, 622-8
- Helm, M., Brule, H., Friede, D., Giege, R., Putz, D. & Florentz, C. (2000) Search for characteristic structural features of mammalian mitochondrial trnas. *RNA*, **6**, 1356-79
- Henderson, S. & Sollner-Webb, B. (1986) A transcriptional terminator is a novel element of the promoter of the mouse ribosomal rna gene. *Cell*, **47**, 891-900
- Hendrickson, W.A. & Ogata, C. (1997) Phase determination from multiwavelength anomalous scattering diffraction measurements. *Methods Enzymol.*, **276**, 494-522
- Heras, B. & Martin, J.L. (2005) Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Crystallogr D Biol Crystallogr*, **61**, 1173-80
- Hess, J.F., Parisi, M.A., Bennett, J.L. & Clayton, D.A. (1991) Impairment of mitochondrial transcription termination by a point mutation associated with the melas subgroup of mitochondrial encephalomyopathies. *Nature*, **351**, 236-9
- Hirano, M. & Dimauro, S. (2001) Ant1, twinkle, polg, and tp: New genes open our eyes to ophthalmoplegia. *Neurology*, **57**, 2163-5
- Hirsch, M. & Penman, S. (1974) The messenger-like properties of the poly(a)plus rna in mammalian mitochondria. *Cell*, **3**, 335-9
- Hixson, J.E. & Clayton, D.A. (1985) Initiation of transcription from each of the two human mitochondrial promoters requires unique nucleotides at the transcriptional start sites. *Proc Natl Acad Sci U S A*, **82**, 2660-4
- Ho, B.K., Thomas, A. & Brasseur, R. (2003) Revisiting the ramachandran plot: Hard-sphere repulsion, electrostatics, and h-bonding in the alpha-helix. *Protein Sci*, **12**, 2508-22
- Holt, I.J., Lorimer, H.E. & Jacobs, H.T. (2000) Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell*, **100**, 515-24
- Holzmann, J., Frank, P., Löffler, E., Bennett, K.L., Gerner, C. & Rossmann, W. (2008) Rnase p without rna: Identification and functional reconstitution of the human mitochondrial trna processing enzyme. *Cell*, **135**, 462-74
- Hood, D.A. (2001) Invited review: Contractile activity-induced mitochondrial biogenesis in skeletal muscle. *J Appl Physiol*, **90**, 1137-57
- Hoppins, S., Lackner, L. & Nunnari, J. (2007) The machines that divide and fuse mitochondria. *Annu Rev Biochem*, **76**, 751-80
- Hovmöller, S., Zhou, T. & Ohlson, T. (2002) Conformations of amino acids in proteins. *Acta Crystallogr D Biol Crystallogr*, **58**, 768-76
- Huber, A.H., Nelson, W.J. & Weis, W.I. (1997) Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell*, **90**, 871-82
- Hurd, D.D. & Saxton, W.M. (1996) Kinesin mutations cause motor neuron disease phenotypes by disrupting fast axonal transport in drosophila. *Genetics*, **144**, 1075-85
- Hyvarinen, A.K., Kumanto, M.K., Marjavaara, S.K. & Jacobs, H.T. Effects on mitochondrial transcription of manipulating mterf protein levels in cultured human hek293 cells. *BMC Mol Biol*, **11**, 72
- Hyvarinen, A.K., Pohjoismaki, J.L., Holt, I.J. & Jacobs, H.T. Overexpression of mterfd1 or mterfd3 impairs the completion of mitochondrial DNA replication. *Mol Biol Rep*
- Hyvarinen, A.K., Pohjoismaki, J.L., Reyes, A., Wanrooij, S., Yasukawa, T., Karhunen, P.J., Spelbrink, J.N., Holt, I.J. & Jacobs, H.T. (2007) The mitochondrial transcription termination factor mterf modulates replication pausing in human mitochondrial DNA. *Nucleic Acids Res*, **35**, 6458-74
- Iborra, F.J., Kimura, H. & Cook, P.R. (2004) The functional organization of mitochondrial genomes in human cells. *BMC Biol*, **2**, 9
- Ishihara, N., Jofuku, A., Eura, Y. & Mihara, K. (2003) Regulation of mitochondrial morphology by membrane potential, and drp1-dependent division and fzo1-dependent fusion reaction in mammalian cells. *Biochem Biophys Res Commun*, **301**, 891-8
- Iverson, T.M., Luna-Chavez, C., Cecchini, G. & Rees, D.C. (1999) Structure of the escherichia coli fumarate reductase respiratory complex. *Science*, **284**, 1961-6

- Jaattela, M. (2004) Multiple cell death pathways as regulators of tumour initiation and progression. *Oncogene*, **23**, 2746-56
- Jacrot, B. (1976) The study of biological structures by neutron scattering from solution *Rep. Prog. Phys.*, **39**, 911
- Jameson, D., Gibson, A.P., Hudelot, C. & Higgs, P.G. (2003) OGRE: A relational database for comparative analysis of mitochondrial genomes. *Nucleic Acids Res.*, **31**, 202-6
- Jeppesen, M.G., Navratil, T., Spemulli, L.L. & Nyborg, J. (2005) Crystal structure of the bovine mitochondrial elongation factor tu.Ts complex. *J Biol Chem*, **280**, 5071-81
- Jia-Xing, Y., Woolfson, M.M., Wilson, K.S. & Dodson, E.J. (2005) A modified acorn to solve protein structures at resolutions of 1.7 Å or better. *Acta Crystallogr D Biol Crystallogr*, **61**, 1465-75
- Jia, L., Dienhart, M., Schramm, M., Mccauley, M., Hell, K. & Stuart, R.A. (2003) Yeast oxal1 interacts with mitochondrial ribosomes: The importance of the c-terminal region of oxal1. *EMBO J*, **22**, 6438-47
- Johnson, A.A. & Johnson, K.A. (2001) Exonuclease proofreading by human mitochondrial DNA polymerase. *J Biol Chem*, **276**, 38097-107
- Jones, D.T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol*, **292**, 195-202
- Kabsch, W.J. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J.Appl.Cryst*, 795-800
- Kaltschmidt, B., Kaltschmidt, C., Hofmann, T.G., Hehner, S.P., Droge, W. & Schmitz, M.L. (2000) The pro- or anti-apoptotic function of nf-kappab is determined by the nature of the apoptotic stimulus. *Eur J Biochem*, **267**, 3828-35
- Kanehisa Laboratories (2008) Oxidative phosphorylation In: *KEGG: Kyoto Encyclopedia of Genes and Genomes (Online)*. <<http://www.genome.jp/kegg/>>
- Kantardjiev, K.A. & Rupp, B. (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci*, **12**, 1865-71
- Karbowski, M., Arnoult, D., Chen, H., Chan, D.C., Smith, C.L. & Youle, R.J. (2004) Quantitation of mitochondrial dynamics by photolabeling of individual organelles shows that mitochondrial fusion is blocked during the bax activation phase of apoptosis. *J Cell Biol*, **164**, 493-9
- Karlberg, O., Canback, B., Kurland, C.G. & Andersson, S.G. (2000) The dual origin of the yeast mitochondrial proteome. *Yeast*, **17**, 170-87
- Karlin, S. & Brocchieri, L. (2000) Heat shock protein 60 sequence comparisons: Duplications, lateral transfer, and mitochondrial evolution. *Proc Natl Acad Sci U S A*, **97**, 11348-53
- Karrasch, S. & Walker, J.E. (1999) Novel features in the structure of bovine atp synthase. *J Mol Biol*, **290**, 379-84
- Kasamatsu, H., Robberson, D.L. & Vinograd, J. (1971) A novel closed-circular mitochondrial DNA with properties of a replicating intermediate. *Proc Natl Acad Sci U S A*, **68**, 2252-7
- Kasamatsu, H. & Vinograd, J. (1974) Replication of circular DNA in eukaryotic cells. *Annu Rev Biochem*, **43**, 695-719
- Kelly, D.P. & Scarpulla, R.C. (2004) Transcriptional regulatory circuits controlling mitochondrial biogenesis and function. *Genes Dev*, **18**, 357-68
- Kenney, W.C., Walker, W.H. & Singer, T.P. (1972) Studies on succinate dehydrogenase. Xx. Amino acid sequence around the flavin site. *J Biol Chem*, **247**, 4510-3
- Kettenberger, H. & Cramer, P. (2006) Fluorescence detection of nucleic acids and proteins in multi-component crystals. *Acta Crystallogr D Biol Crystallogr*, **62**, 146-50
- Koc, E.C. & Spemulli, L.L. (2002) Identification of mammalian mitochondrial translational initiation factor 3 and examination of its role in initiation complex formation with natural mRNAs. *J Biol Chem*, **277**, 35541-9
- Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. & Svergun, D.I. (2003) *Primus*: A windows pc-based system for small-angle scattering data analysis. *J. Appl. Cryst.*, **36**, 1277-1282
- Koshihara, T., Detmer, S.A., Kaiser, J.T., Chen, H., Mccaffery, J.M. & Chan, D.C. (2004) Structural basis of mitochondrial tethering by mitofusin complexes. *Science*, **305**, 858-62
- Kravchenko, J.E., Rogozin, I.B., Koonin, E.V. & Chumakov, P.M. (2005) Transcription of mammalian messenger RNAs by a nuclear RNA polymerase of mitochondrial origin. *Nature*, **436**, 735-9
- Krissinel, E. & Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr*, **60**, 2256-68
- Krissinel, E. & Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, **372**, 774-97
- Kroemer, G., Galluzzi, L. & Brenner, C. (2007) Mitochondrial membrane permeabilization in cell death. *Physiol Rev*, **87**, 99-163

- Kruse, B., Narasimhan, N. & Attardi, G. (1989) Termination of transcription in human mitochondria: Identification and purification of a DNA binding protein factor that promotes termination. *Cell*, **58**, 391-7
- Kunkel, T.A. (1985) The mutational specificity of DNA polymerases-alpha and -gamma during in vitro DNA synthesis. *J Biol Chem*, **260**, 12866-74
- Kunkel, T.A. & Mosbaugh, D.W. (1989) Exonucleolytic proofreading by a mammalian DNA polymerase. *Biochemistry*, **28**, 988-95
- Kurland, C.G. & Andersson, S.G. (2000) Origin and evolution of the mitochondrial proteome. *Microbiol Mol Biol Rev*, **64**, 786-820
- Lagerkvist, U. (1978) "Two out of three": An alternative method for codon reading. *Proc Natl Acad Sci U S A*, **75**, 1759-62
- Lancaster, C.R., Kroger, A., Auer, M. & Michel, H. (1999) Structure of fumarate reductase from *Wolfinella succinogenes* at 2.2 Å resolution. *Nature*, **402**, 377-85
- Lang, B.F., Burger, G., O'Kelly, C.J., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M. & Gray, M.W. (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature*, **387**, 493-7
- Lang, B.F., Seif, E., Gray, M.W., O'Kelly, C.J. & Burger, G. (1999) A comparative genomics approach to the evolution of eukaryotes and their mitochondria. *J Eukaryot Microbiol*, **46**, 320-6
- Langst, G., Becker, P.B. & Grummt, I. (1998) Ttf-i determines the chromatin architecture of the active rDNA promoter. *EMBO J*, **17**, 3135-45
- Langst, G., Blank, T.A., Becker, P.B. & Grummt, I. (1997) Rna polymerase i transcription on nucleosomal templates: The transcription termination factor ttf-i induces chromatin remodeling and relieves transcriptional repression. *EMBO J*, **16**, 760-8
- Laskowski, R.A., Moss, D.S. & Thornton, J.M. (1993) Main-chain bond lengths and bond angles in protein structures. *J Mol Biol*, **231**, 1049-67
- Lee, D.Y. & Clayton, D.A. (1996) Properties of a primer rna-DNA hybrid at the mouse mitochondrial DNA leading-strand origin of replication. *J Biol Chem*, **271**, 24262-9
- Lee, D.Y. & Clayton, D.A. (1997) Rnase mitochondrial rna processing correctly cleaves a novel r loop at the mitochondrial DNA leading-strand origin of replication. *Genes Dev*, **11**, 582-92
- Lee, J., Sharma, S., Kim, J., Ferrante, R.J. & Ryu, H. (2008) Mitochondrial nuclear receptors and transcription factors: Who's minding the cell? *J Neurosci Res*, **86**, 961-71
- Lee, Y.S., Kennedy, W.D. & Yin, Y.W. (2009) Structural insight into processive human mitochondrial DNA synthesis and disease-related polymerase mutations. *Cell*, **139**, 312-24
- Legros, F., Lombes, A., Frachon, P. & Rojo, M. (2002) Mitochondrial fusion in human cells is efficient, requires the inner membrane potential, and is mediated by mitofusins. *Mol Biol Cell*, **13**, 4343-54
- Legros, F., Malka, F., Frachon, P., Lombes, A. & Rojo, M. (2004) Organization and dynamics of human mitochondrial DNA. *J Cell Sci*, **117**, 2653-62
- Leist, M. & Jaattela, M. (2001) Four deaths and a funeral: From caspases to alternative mechanisms. *Nat Rev Mol Cell Biol*, **2**, 589-98
- Leslie, A.G.W. (1992) Recent changes to the mosflm package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, **No. 26**
- Letunic, I., Doerks, T. & Bork, P. (2009) Smart 6: Recent updates and new developments. *Nucleic Acids Res*, **37**, D229-32
- Levinger, L., Morl, M. & Florentz, C. (2004) Mitochondrial trna 3' end metabolism and human disease. *Nucleic Acids Res*, **32**, 5430-41
- Liao, H.X. & Spremulli, L.L. (1990a) Effects of length and mrna secondary structure on the interaction of bovine mitochondrial ribosomes with messenger rna. *J Biol Chem*, **265**, 11761-5
- Liao, H.X. & Spremulli, L.L. (1990b) Identification and initial characterization of translational initiation factor 2 from bovine mitochondria. *J Biol Chem*, **265**, 13618-22
- Liao, H.X. & Spremulli, L.L. (1991) Initiation of protein synthesis in animal mitochondria. Purification and characterization of translational initiation factor 2. *J Biol Chem*, **266**, 20714-9
- Linder, T., Park, C.B., Asin-Cayuela, J., Pellegrini, M., Larsson, N.G., Falkenberg, M., Samuelsson, T. & Gustafsson, C.M. (2005) A family of putative transcription termination factors shared amongst metazoans and plants. *Curr Genet*, **48**, 265-9
- Litonin, D., Sologub, M., Shi, Y., Savkina, M., Anikin, M., Falkenberg, M., Gustafsson, C.M. & Temiakov, D. (2010) Human mitochondrial transcription revisited: Only tfam and tfb2m are required for transcription of the mitochondrial genes in vitro. *J Biol Chem*, **285**, 18129-33

- Lodeiro, M.F., Uchida, A.U., Arnold, J.J., Reynolds, S.L., Moustafa, I.M. & Cameron, C.E. Identification of multiple rate-limiting steps during the human mitochondrial transcription cycle in vitro. *J Biol Chem*, **285**, 16387-402
- Loguercio Polosa, P., Roberti, M., Musicco, C., Gadaleta, M.N., Quagliariello, E. & Cantatore, P. (1999) Cloning and characterisation of mtdbp, a DNA-binding protein which binds two distinct regions of sea urchin mitochondrial DNA. *Nucleic Acids Res*, **27**, 1890-9
- Lonsdale, D.M., Hodge, T.P. & Fauron, C.M. (1984) The physical map and organisation of the mitochondrial genome from the fertile cytoplasm of maize. *Nucleic Acids Res*, **12**, 9249-61
- Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, De Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S. & Richardson, D.C. (2003) Structure validation by calpha geometry: Phi,psi and cbeta deviation. *Proteins*, **50**, 437-50
- Lowell, B.B. & Spiegelman, B.M. (2000) Towards a molecular understanding of adaptive thermogenesis. *Nature*, **404**, 652-60
- Lutter, R., Saraste, M., Van Walraven, H.S., Runswick, M.J., Finel, M., Deatherage, J.F. & Walker, J.E. (1993) F1f0-atp synthase from bovine heart mitochondria: Development of the purification of a monodisperse oligomycin-sensitive atpase. *Biochem J*, **295 (Pt 3)**, 799-806
- Maiti, R., Van Domselaar, G.H. & Wishart, D.S. (2005) Moviemaker: A web server for rapid rendering of protein motions and interactions. *Nucleic Acids Res*, **33**, W358-62
- Manam, S. & Van Tuyle, G.C. (1987) Separation and characterization of 5'- and 3'-trna processing nucleases from rat liver mitochondria. *J Biol Chem*, **262**, 10272-9
- Mannella, C.A., Marko, M. & Buttle, K. (1997) Reconsidering mitochondrial structure: New views of an old organelle. *Trends Biochem Sci*, **22**, 37-8
- Mannella, C.A., Marko, M., Penczek, P., Barnard, D. & Frank, J. (1994) The internal compartmentation of rat-liver mitochondria: Tomographic study using the high-voltage transmission electron microscope. *Microsc Res Tech*, **27**, 278-83
- Mannella, C.A., Pfeiffer, D.R., Bradshaw, P.C., Moraru, Ii, Slepchenko, B., Loew, L.M., Hsieh, C.E., Buttle, K. & Marko, M. (2001) Topology of the mitochondrial inner membrane: Dynamics and bioenergetic implications. *IUBMB Life*, **52**, 93-100
- Marcotte, E.M. (2000) Computational genetics: Finding protein function by nonhomology methods. *Curr Opin Struct Biol*, **10**, 359-65
- Marcotte, E.M., Xenarios, I., Van Der Bliek, A.M. & Eisenberg, D. (2000) Localizing proteins in the cell from their phylogenetic profiles. *Proc Natl Acad Sci U S A*, **97**, 12115-20
- Margulis, L. (1981) *Symbiosis in cell evolution.*, New York.
- Martin, M., Cho, J., Cesare, A.J., Griffith, J.D. & Attardi, G. (2005) Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rna synthesis. *Cell*, **123**, 1227-40
- Martin, W. & Herrmann, R.G. (1998) Gene transfer from organelles to the nucleus: How much, what happens, and why? *Plant Physiol*, **118**, 9-17
- Martinez-Diez, M., Santamaria, G., Ortega, A.D. & Cuezva, J.M. (2006) Biogenesis and dynamics of mitochondria during the cell cycle: Significance of 3'utrs. *PLoS One*, **1**, e107
- Marzuki, S. & Hibbs, A.R. (1986) Are all mitochondrial translation products synthesized on membrane-bound ribosomes? *Biochim Biophys Acta*, **866**, 120-4
- Masters, B.S., Stohl, L.L. & Clayton, D.A. (1987) Yeast mitochondrial rna polymerase is homologous to those encoded by bacteriophages t3 and t7. *Cell*, **51**, 89-99
- Mattenberger, Y., James, D.I. & Martinou, J.C. (2003) Fusion of mitochondria in mammalian cells is dependent on the mitochondrial inner membrane potential and independent of microtubules or actin. *FEBS Lett*, **538**, 53-9
- Matthews, B.W. (1968) Solvent content of protein crystals. *J Mol Biol*, **33**, 491-7
- Mcculloch, V., Seidel-Rogol, B.L. & Shadel, G.S. (2002) A human mitochondrial transcription factor is related to rna adenine methyltransferases and binds s-adenosylmethionine. *Mol Cell Biol*, **22**, 1116-25
- Mestay, B. & Reeder, R.H. (1986) A termination site for xenopus rna polymerase i also acts as an element of an adjacent promoter. *Cell*, **47**, 913-20
- Meeusen, S., Mccaffery, J.M. & Nunnari, J. (2004) Mitochondrial fusion intermediates revealed in vitro. *Science*, **305**, 1747-52
- Merli, A., Capaldi, R.A., Ackrell, B.A. & Kearney, E.B. (1979) Arrangement of complex ii (succinate-ubiquinone reductase) in the mitochondrial inner membrane. *Biochemistry*, **18**, 1393-400
- Micol, V., Fernandez-Silva, P. & Attardi, G. (1996) Isolation and assay of mitochondrial transcription termination factor from human cells. *Methods Enzymol*, **264**, 158-73

- Micol, V., Fernandez-Silva, P. & Attardi, G. (1997) Functional analysis of in vivo and in organello footprinting of hela cell mitochondrial DNA in relationship to atp and ethidium bromide effects on transcription. *J Biol Chem*, **272**, 18896-904
- Michikawa, Y., Mazzucchelli, F., Bresolin, N., Scarlato, G. & Attardi, G. (1999) Aging-dependent large accumulation of point mutations in the human mtdna control region for replication. *Science*, **286**, 774-9
- Minn, A.J., Velez, P., Schendel, S.L., Liang, H., Muchmore, S.W., Fesik, S.W., Fill, M. & Thompson, C.B. (1997) Bcl-x(l) forms an ion channel in synthetic lipid membranes. *Nature*, **385**, 353-7
- Miquel, J. (1998) An update on the oxygen stress-mitochondrial mutation theory of aging: Genetic and evolutionary implications. *Exp Gerontol*, **33**, 113-26
- Mitchell, P. (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature*, **191**, 144-8
- Miyakawa, I., Aoi, H., Sando, N. & Kuroiwa, T. (1984) Fluorescence microscopic studies of mitochondrial nucleoids during meiosis and sporulation in the yeast, *saccharomyces cerevisiae*. *J Cell Sci*, **66**, 21-38
- Miyakawa, I., Sando, N., Kawano, S., Nakamura, S. & Kuroiwa, T. (1987) Isolation of morphologically intact mitochondrial nucleoids from the yeast, *saccharomyces cerevisiae*. *J Cell Sci*, **88 (Pt 4)**, 431-9
- Montoya, J. (2005) Biogénesis y patología mitocondrial. *Rev. Real Academia de Ciencias. Zaragoza.*, **60**, 7-28
- Montoya, J., Christianson, T., Levens, D., Rabinowitz, M. & Attardi, G. (1982) Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. *Proc Natl Acad Sci U S A*, **79**, 7195-9
- Montoya, J., Gaines, G.L. & Attardi, G. (1983) The pattern of transcription of the human mitochondrial rna genes reveals two overlapping transcription units. *Cell*, **34**, 151-9
- Montoya, J., Lopez-Perez, M.J. & Ruiz-Pesini, E. (2006) Mitochondrial DNA transcription and diseases: Past, present and future. *Biochim Biophys Acta*, **1757**, 1179-89
- Montoya, J., Ojala, D. & Attardi, G. (1981) Distinctive features of the 5'-terminal sequences of the human mitochondrial mtnas. *Nature*, **290**, 465-70
- Morris, A.L., Macarthur, M.W., Hutchinson, E.G. & Thornton, J.M. (1992) Stereochemical quality of protein structure coordinates. *Proteins*, **12**, 345-64
- Morris, R.L. & Hollenbeck, P.J. (1993) The regulation of bidirectional mitochondrial transport is coordinated with axonal outgrowth. *J Cell Sci*, **104 (Pt 3)**, 917-27
- Morris, R.L. & Hollenbeck, P.J. (1995) Axonal transport of mitochondria along microtubules and f-actin in living vertebrate neurons. *J Cell Biol*, **131**, 1315-26
- Murphy, W.I., Attardi, B., Tu, C. & Attardi, G. (1975) Evidence for complete symmetrical transcription in vivo of mitochondrial DNA in hela cells. *J Mol Biol*, **99**, 809-14
- Murshudov, G.N., Vagin, A.A. & Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, **53**, 240-55
- Mutvei, A., Kuzela, S. & Nelson, B.D. (1989) Control of mitochondrial transcription by thyroid hormone. *Eur J Biochem*, **180**, 235-40
- Nagaike, T., Suzuki, T., Tomari, Y., Takemoto-Hori, C., Negayama, F., Watanabe, K. & Ueda, T. (2001) Identification and characterization of mammalian mitochondrial trna nucleotidyltransferases. *J Biol Chem*, **276**, 40041-9
- Nakada, K., Inoue, K., Ono, T., Isobe, K., Ogura, A., Goto, Y.I., Nonaka, I. & Hayashi, J.I. (2001) Inter-mitochondrial complementation: Mitochondria-specific system preventing mice from expression of disease phenotypes by mutant mtdna. *Nat Med*, **7**, 934-40
- Nakada, K., Sato, A. & Hayashi, J. (2009) Mitochondrial functional complementation in mitochondrial DNA-based diseases. *Int J Biochem Cell Biol*, **41**, 1907-13
- Narasimhan, N. & Attardi, G. (1987) Specific requirement for atp at an early step of in vitro transcription of human mitochondrial DNA. *Proc Natl Acad Sci U S A*, **84**, 4078-82
- Nass, M.M. (1980) Analysis of the two heavy and light strand origins and the direction of replication of mitochondrial DNA within a detailed physical map of this genome in transformed hamster cells. *J Mol Biol*, **140**, 231-56
- Newmeyer, D.D. & Ferguson-Miller, S. (2003) Mitochondria: Releasing power for life and unleashing the machineries of death. *Cell*, **112**, 481-90
- Nunnari, J., Marshall, W.F., Straight, A., Murray, A., Sedat, J.W. & Walter, P. (1997) Mitochondrial transmission during mating in *saccharomyces cerevisiae* is determined by mitochondrial fusion and fission and the intramitochondrial segregation of mitochondrial DNA. *Mol Biol Cell*, **8**, 1233-42

- O'Brien, T.W. & Denslow, N.D. (1996) 13bovine mitochondrial ribosomes. *Methods Enzymol*, **264**, 237-48
- O'Brien, T.W. & Kalf, G.F. (1967) Ribosomes from rat liver mitochondria. II. Partial characterization. *J Biol Chem*, **242**, 2180-5
- Ojala, D., Merkel, C., Gelfand, R. & Attardi, G. (1980) The trna genes punctuate the reading of genetic information in human mitochondrial DNA. *Cell*, **22**, 393-403
- Ojala, D., Montoya, J. & Attardi, G. (1981) Trna punctuation model of rna processing in human mitochondria. *Nature*, **290**, 470-4
- Okamoto, K. & Shaw, J.M. (2005) Mitochondrial morphology and dynamics in yeast and multicellular eukaryotes. *Annu Rev Genet*, **39**, 503-36
- Painter, J. & Merritt, E.A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing tls motion. *Acta Crystallogr D Biol Crystallogr*, **62**, 439-50
- Palmer, J.D. (1997) Genome evolution. The mitochondrion that time forgot. *Nature*, **387**, 454-5
- Parisi, M.A. & Clayton, D.A. (1991) Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. *Science*, **252**, 965-9
- Park, C.B., Asin-Cayuela, J., Camara, Y., Shi, Y., Pellegrini, M., Gaspari, M., Wibom, R., Hultenby, K., Erdjument-Bromage, H., Tempst, P., Falkenberg, M., Gustafsson, C.M. & Larsson, N.G. (2007) Mterf3 is a negative regulator of mammalian mtdna transcription. *Cell*, **130**, 273-85
- Peifer, M., Berg, S. & Reynolds, A.B. (1994) A repeating amino acid motif shared by proteins with diverse cellular roles. *Cell*, **76**, 789-91
- Pellegrini, M., Asin-Cayuela, J., Erdjument-Bromage, H., Tempst, P., Larsson, N.G. & Gustafsson, C.M. (2009) Mterf2 is a nucleoid component in mammalian mitochondria. *Biochim Biophys Acta*, **1787**, 296-302
- Perkins, G., Renken, C., Martone, M.E., Young, S.J., Ellisman, M. & Frey, T. (1997) Electron tomography of neuronal mitochondria: Three-dimensional structure and organization of cristae and membrane contacts. *J Struct Biol*, **119**, 260-72
- Perkins, G.A. & Frey, T.G. (2000) Recent structural insight into mitochondria gained by microscopy. *Micron*, **31**, 97-111
- Perrakis, A., Morris, R. & Lamzin, V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol*, **6**, 458-63
- Petit, P.X., Goubern, M., Diolez, P., Susin, S.A., Zamzami, N. & Kroemer, G. (1998) Disruption of the outer mitochondrial membrane as a result of large amplitude swelling: The impact of irreversible permeability transition. *FEBS Lett*, **426**, 111-6
- Pfanner, N., Rassow, J., Van Der Klei, I.J. & Neupert, W. (1992) A dynamic model of the mitochondrial protein import machinery. *Cell*, **68**, 999-1002
- Pfanner, N., Rassow, J., Wienhues, U., Hergersberg, C., Sollner, T., Becker, K. & Neupert, W. (1990) Contact sites between inner and outer membranes: Structure and role in protein translocation into the mitochondria. *Biochim Biophys Acta*, **1018**, 239-42
- Pham, X.H., Farge, G., Shi, Y., Gaspari, M., Gustafsson, C.M. & Falkenberg, M. (2006) Conserved sequence box ii directs transcription termination and primer formation in mitochondria. *J Biol Chem*, **281**, 24647-52
- Pohjoismaki, J.L., Holmes, J.B., Wood, S.R., Yang, M.Y., Yasukawa, T., Reyes, A., Bailey, L.J., Cluett, T.J., Goffart, S., Willcox, S., Rigby, R.E., Jackson, A.P., Spelbrink, J.N., Griffith, J.D., Crouch, R.J., Jacobs, H.T. & Holt, I.J. Mammalian mitochondrial DNA replication intermediates are essentially duplex but contain extensive tracts of rna/dna hybrid. *J Mol Biol*, **397**, 1144-55
- Polosa, P.L., Deceglie, S., Falkenberg, M., Roberti, M., Di Ponzio, B., Gadaleta, M.N. & Cantatore, P. (2007) Cloning of the sea urchin mitochondrial rna polymerase and reconstitution of the transcription termination system. *Nucleic Acids Res*, **35**, 2413-27
- Polosa, P.L., Deceglie, S., Roberti, M., Gadaleta, M.N. & Cantatore, P. (2005) Contrahelicase activity of the mitochondrial transcription termination factor mtdbp. *Nucleic Acids Res*, **33**, 3812-20
- Prieto-Martin, A., Montoya, J. & Martinez-Azorin, F. (2004) New DNA-binding activity of rat mitochondrial transcription termination factor (mterf). *J Biochem*, **136**, 825-30
- Psarra, A.M. & Sekeris, C.E. (2008a) Nuclear receptors and other nuclear transcription factors in mitochondria: Regulatory molecules in a new environment. *Biochim Biophys Acta*, **1783**, 1-11
- Psarra, A.M. & Sekeris, C.E. (2008b) Steroid and thyroid hormone receptors in mitochondria. *IUBMB Life*, **60**, 210-23
- Puranam, R.S. & Attardi, G. (2001) The rnaase p associated with hela cell mitochondria contains an essential rna component identical in sequence to that of the nuclear rnaase p. *Mol Cell Biol*, **21**, 548-61

- Putter, V. & Grummt, F. (2002) Transcription termination factor ttf-i exhibits contrahelicase activity during DNA replication. *EMBO Rep*, **3**, 147-52
- Ramachandran, G.N., Ramakrishnan, C. & Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol*, **7**, 95-9
- Rassow, J., Hartl, F.U., Guiard, B., Pfanner, N. & Neupert, W. (1990) Polypeptides traverse the mitochondrial envelope in an extended state. *FEBS Lett*, **275**, 190-4
- Ravagnan, L., Roumier, T. & Kroemer, G. (2002) Mitochondria, the killer organelles and their weapons. *J Cell Physiol*, **192**, 131-7
- Raven, P.H. (1970) A multiple origin for plastids and mitochondria. *Science*, **169**, 641-6
- Reyes, A., Yang, M.Y., Bowmaker, M. & Holt, I.J. (2005) Bidirectional replication initiates at sites throughout the mitochondrial genome of birds. *J Biol Chem*, **280**, 3242-50
- Richardson, J.P. (1993) Transcription termination. *Crit Rev Biochem Mol Biol*, **28**, 1-30
- Robberson, D.L., Kasamatsu, H. & Vinograd, J. (1972) Replication of mitochondrial DNA. Circular replicative intermediates in mouse l cells. *Proc Natl Acad Sci U S A*, **69**, 737-41
- Roberti, M., Bruni, F., Loguercio Polosa, P., Manzari, C., Gadaleta, M.N. & Cantatore, P. (2006a) Mterf3, the most conserved member of the mterf-family, is a modular factor involved in mitochondrial protein synthesis. *Biochim Biophys Acta*, **1757**, 1199-206
- Roberti, M., Bruni, F., Polosa, P.L., Gadaleta, M.N. & Cantatore, P. (2006b) The drosophila termination factor dmttf regulates in vivo mitochondrial transcription. *Nucleic Acids Res*, **34**, 2109-16
- Roberti, M., Fernandez-Silva, P., Polosa, P.L., Fernandez-Vizarrá, E., Bruni, F., Deceglie, S., Montoya, J., Gadaleta, M.N. & Cantatore, P. (2005) In vitro transcription termination activity of the drosophila mitochondrial DNA-binding protein dmttf. *Biochem Biophys Res Commun*, **331**, 357-62
- Roberti, M., Mustich, A., Gadaleta, M.N. & Cantatore, P. (1991) Identification of two homologous mitochondrial DNA sequences, which bind strongly and specifically to a mitochondrial protein of paracentrotus lividus. *Nucleic Acids Res*, **19**, 6249-54
- Roberti, M., Polosa, P.L., Bruni, F., Manzari, C., Deceglie, S., Gadaleta, M.N. & Cantatore, P. (2009) The mterf family proteins: Mitochondrial transcription regulators and beyond. *Biochim Biophys Acta*, **1787**, 303-11
- Roberti, M., Polosa, P.L., Bruni, F., Musicco, C., Gadaleta, M.N. & Cantatore, P. (2003) Dmttf, a novel mitochondrial transcription termination factor that recognises two sequences of drosophila melanogaster mitochondrial DNA. *Nucleic Acids Res*, **31**, 1597-604
- Rodeheffer, M.S. & Shadel, G.S. (2003) Multiple interactions involving the amino-terminal domain of yeast mtrna polymerase determine the efficiency of mitochondrial protein synthesis. *J Biol Chem*, **278**, 18695-701
- Rodriguez, D.D., Grosse, C., Himmel, S., Gonzalez, C., De Ilarduya, I.M., Becker, S., Sheldrick, G.M. & Uson, I. (2009) Crystallographic ab initio protein structure solution below atomic resolution. *Nat Methods*, **6**, 651-3
- Rojo, G., Chamorro, M., Salas, M.L., Vinuela, E., Cuezva, J.M. & Salas, J. (1998) Migration of mitochondria to viral assembly sites in african swine fever virus-infected cells. *J Virol*, **72**, 7583-8
- Rosignol, R., Faustin, B., Rocher, C., Malgat, M., Mazat, J.P. & Letellier, T. (2003) Mitochondrial threshold effects. *Biochem J*, **370**, 751-62
- Rossmann, W., Tullo, A., Potuschak, T., Karwan, R. & Sbisá, E. (1995) Human mitochondrial trna processing. *J Biol Chem*, **270**, 12885-91
- Roussel, A.A.C., C. (1989) Turbo-frodo. In silicon graphics geometry partner directory. Silicon graphics, mountain view, ca. 77-78
- Rupp, B.A.W., J.-W. (2004) Predictive models for protein crystallization. *Methods*, **34**, 391-408
- Sanchirico, M.E., Fox, T.D. & Mason, T.L. (1998) Accumulation of mitochondrially synthesized saccharomyces cerevisiae cox2p and cox3p depends on targeting information in untranslated portions of their mRNAs. *EMBO J*, **17**, 5796-804
- Satoh, M. & Kuroiwa, T. (1991) Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Exp Cell Res*, **196**, 137-40
- Sauvanet, C., Duvezin-Caubet, S., Di Rago, J.P. & Rojo, M. (2009) Energetic requirements and bioenergetic modulation of mitochondrial morphology and dynamics. *Semin Cell Dev Biol*
- Sazanov, L.A. & Hinchliffe, P. (2006) Structure of the hydrophilic domain of respiratory complex i from thermus thermophilus. *Science*, **311**, 1430-6
- Scalettar, B.A., Abney, J.R. & Hackenbrock, C.R. (1991) Dynamics, structure, and function are coupled in the mitochondrial matrix. *Proc Natl Acad Sci U S A*, **88**, 8057-61

- Scarpulla, R.C. (1997) Nuclear control of respiratory chain expression in mammalian cells. *J Bioenerg Biomembr*, **29**, 109-19
- Scarpulla, R.C. (2008) Transcriptional paradigms in mammalian mitochondrial biogenesis and function. *Physiol Rev*, **88**, 611-38
- Schendel, S.L., Xie, Z., Montal, M.O., Matsuyama, S., Montal, M. & Reed, J.C. (1997) Channel formation by antiapoptotic protein bcl-2. *Proc Natl Acad Sci U S A*, **94**, 5113-8
- Schiffer, S., Rosch, S. & Marchfelder, A. (2002) Assigning a function to a conserved group of proteins: The trna 3'-processing enzymes. *EMBO J*, **21**, 2769-77
- Schillace, R., Preiss, T., Lightowers, R.N. & Capaldi, R.A. (1994) Developmental regulation of tissue-specific isoforms of subunit via of beef cytochrome c oxidase. *Biochim Biophys Acta*, **1188**, 391-7
- Schlerf, A., Droste, M., Winter, M. & Kadenbach, B. (1988) Characterization of two different genes (cdna) for cytochrome c oxidase subunit via from heart and liver of the rat. *EMBO J*, **7**, 2387-91
- Schultz, J., Milpetz, F., Bork, P. & Ponting, C.P. (1998) Smart, a simple modular architecture research tool: Identification of signaling domains. *Proc Natl Acad Sci U S A*, **95**, 5857-64
- Schwartz, R.M. & Dayhoff, M.O. (1978) Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science*, **199**, 395-403
- Shadel, G.S. & Clayton, D.A. (1997) Mitochondrial DNA maintenance in vertebrates. *Annu Rev Biochem*, **66**, 409-35
- Shang, J. & Clayton, D.A. (1994) Human mitochondrial transcription termination exhibits rna polymerase independence and biased bipolarity in vitro. *J Biol Chem*, **269**, 29112-20
- Sheldrick, G.M. (2002) Macromolecular phasing with shelxe. *Z. Kristallogr.*, **217**, 644-650
- Sheldrick, G.M. (2008) A short history of shelx. *Acta Crystallogr A*, **64**, 112-22
- Shoubridge, E.A. (2001) Nuclear genetic defects of oxidative phosphorylation. *Hum Mol Genet*, **10**, 2277-84
- Shuey, D.J. & Attardi, G. (1985) Characterization of an rna polymerase activity from hela cell mitochondria, which initiates transcription at the heavy strand rna promoter and the light strand promoter in human mitochondrial DNA. *J Biol Chem*, **260**, 1952-8
- Shutt, T.E., Lodeiro, M.F., Cotney, J., Cameron, C.E. & Shadel, G.S. Core human mitochondrial transcription apparatus is a regulated two-component system in vitro. *Proc Natl Acad Sci U S A*, **107**, 12133-8
- Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J. & Li, Q. (2009) Practical protocols for production of very high yields of recombinant proteins using escherichia coli. *Protein Sci*, **18**, 936-48
- Slomovic, S., Laufer, D., Geiger, D. & Schuster, G. (2005) Polyadenylation and degradation of human mitochondrial rna: The prokaryotic past leaves its mark. *Mol Cell Biol*, **25**, 6427-35
- Smid, A., Finsterer, M. & Grummt, I. (1992) Limited proteolysis unmasks specific DNA-binding of the murine rna polymerase i-specific transcription termination factor ttfi. *J Mol Biol*, **227**, 635-47
- Solano, A., Playan, A., Lopez-Perez, M.J. & Montoya, J. (2001) [genetic diseases of the mitochondrial DNA in humans]. *Salud Publica Mex*, **43**, 151-61
- Sondheimer, N., Fang, J.K., Polyak, E., Falk, M. & Avadhani, N.G. The leucine-rich pentatricopeptide-repeat containing protein regulates mitochondrial transcription. *Biochemistry*
- Song, Z., Ghochani, M., Mccaffery, J.M., Frey, T.G. & Chan, D.C. (2009) Mitofusins and opa1 mediate sequential steps in mitochondrial membrane fusion. *Mol Biol Cell*, **20**, 3525-32
- Spahr, H., Samuelsson, T., Hallberg, B.M. & Gustafsson, C.M. Structure of mitochondrial transcription termination factor 3 reveals a novel nucleic acid-binding domain. *Biochem Biophys Res Commun*, **397**, 386-90
- Spelbrink, J.N. Functional organization of mammalian mitochondrial DNA in nucleoids: History, recent developments, and future challenges. *IUBMB Life*, **62**, 19-32
- Spithill, T.W., Trembath, M.K., Lukins, H.B. & Linnane, A.W. (1978) Mutations of the mitochondrial DNA of saccharomyces cerevisiae which affect the interaction between mitochondrial ribosomes and the inner mitochondrial membrane. *Mol Gen Genet*, **164**, 155-62
- Staub, J.M. & Castora, F.J. (1993) Mammalian mitochondrial DNA sequences can function as in vivo bacterial transcription terminators. *Biochem Biophys Res Commun*, **192**, 616-26
- Strong, M., Sawaya, M.R., Wang, S., Phillips, M., Cascio, D. & Eisenberg, D. (2006) Toward the structural genomics of complexes: Crystal structure of a pe/ppe protein complex from mycobacterium tuberculosis. *Proc Natl Acad Sci U S A*, **103**, 8060-5
- Studier, F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*, **41**, 207-34

- Suen, D.F., Norris, K.L. & Youle, R.J. (2008) Mitochondrial dynamics and apoptosis. *Genes Dev*, **22**, 1577-90
- Suhre, K. & Sanejouand, Y.H. (2004) Elnemo: A normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res*, **32**, W610-4
- Sun, P.D., Radaev, S. & Kattah, M. (2002) Generating isomorphous heavy-atom derivatives by a quick-soak method. Part i: Test cases. *Acta Crystallogr D Biol Crystallogr*, **58**, 1092-8
- Svergun, D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.*, **25**, 495-503
- Svergun, D.I., C. Barberato, and M. H. J. Koch (1995) Crysol - a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.*, **28**, 768 - 773
- Svergun, D.I. & Koch, M.H.J. (2003) Small-angle scattering studies of biological macromolecules in solution *Rep. Prog. Phys.*, **66**
- Szyrach, G., Ott, M., Bonnefoy, N., Neupert, W. & Herrmann, J.M. (2003) Ribosome binding to the oxal complex facilitates co-translational protein insertion in mitochondria. *EMBO J*, **22**, 6448-57
- Taylor, G. (2003) The phase problem. *Acta Crystallogr D Biol Crystallogr*, **59**, 1881-90
- Terwilliger, T.C. (2000) Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr*, **56**, 965-72
- Terwilliger, T.C. (2003) Improving macromolecular atomic models at moderate resolution by automated iterative model building, statistical density modification and refinement. *Acta Crystallogr D Biol Crystallogr*, **59**, 1174-82
- Tewari, R., Bailes, E., Bunting, K.A. & Coates, J.C. Armadillo-repeat protein functions: Questions for little creatures. *Trends Cell Biol*, **20**, 470-81
- Tiranti, V., Savoia, A., Forti, F., D'apolito, M.F., Centra, M., Rocchi, M. & Zeviani, M. (1997) Identification of the gene encoding the human mitochondrial rna polymerase (h-mtrpol) by cyberscreening of the expressed sequence tags database. *Hum Mol Genet*, **6**, 615-25
- Tobias, J.W., Shrader, T.E., Rocard, G. & Varshavsky, A. (1991) The n-end rule in bacteria. *Science*, **254**, 1374-7
- Topper, J.N. & Clayton, D.A. (1989) Identification of transcriptional regulatory elements in human mitochondrial DNA by linker substitution analysis. *Mol Cell Biol*, **9**, 1200-11
- Topper, J.N. & Clayton, D.A. (1990) Characterization of human mrp/th rna and its nuclear gene: Full length mrp/th rna is an active endoribonuclease when assembled as an rnp. *Nucleic Acids Res*, **18**, 793-9
- Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R. & Yoshikawa, S. (1996) The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å. *Science*, **272**, 1136-44
- Vagin, A.A.T., A. (1997) *Molrep*: An automated program for molecular replacement. *J. Appl. Cryst.*, **1022-1025**
- Vaidya, A.B., Akella, R. & Suplick, K. (1989) Sequences similar to genes for two mitochondrial proteins and portions of ribosomal rna in tandemly arrayed 6-kilobase-pair DNA of a malarial parasite. *Mol Biochem Parasitol*, **35**, 97-107
- Van Dijk, M. & Bonvin, A.M. (2009) 3d-dart: A DNA structure modelling server. *Nucleic Acids Res*, **37**, W235-9
- Vander Heiden, M.G., Chandel, N.S., Li, X.X., Schumacker, P.T., Colombini, M. & Thompson, C.B. (2000) Outer mitochondrial membrane permeability can regulate coupled respiration and cell survival. *Proc Natl Acad Sci U S A*, **97**, 4666-71
- Vieira, H.L., Haouzi, D., El Hamel, C., Jacotot, E., Belzacq, A.S., Brenner, C. & Kroemer, G. (2000) Permeabilization of the mitochondrial inner membrane during apoptosis: Impact of the adenine nucleotide translocator. *Cell Death Differ*, **7**, 1146-54
- Viscomi, C., Spinazzola, A., Maggioni, M., Fernandez-Vizarra, E., Massa, V., Pagano, C., Vettor, R., Mora, M. & Zeviani, M. (2009) Early-onset liver mtdna depletion and late-onset proteinuric nephropathy in mpv17 knockout mice. *Hum Mol Genet*, **18**, 12-26
- Vogel, F., Bornhovd, C., Neupert, W. & Reichert, A.S. (2006) Dynamic subcompartmentalization of the mitochondrial inner membrane. *J Cell Biol*, **175**, 237-47
- Walter, T.S., Mancini, E.J., Kadlec, J., Graham, S.C., Assenberg, R., Ren, J., Sainsbury, S., Owens, R.J., Stuart, D.I., Grimes, J.M. & Harlos, K. (2008) Semi-automated microseeding of nanolitre crystallization experiments. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, **64**, 14-8
- Walther, D. & Cohen, F.E. (1999) Conformational attractors on the ramachandran map. *Acta Crystallogr D Biol Crystallogr*, **55**, 506-17

- Wang, X., Mclachlan, J., Zamore, P.D. & Hall, T.M. (2002) Modular recognition of rna by a human pumilio-homology domain. *Cell*, **110**, 501-12
- Wang, Y. & Bogenhagen, D.F. (2006) Human mitochondrial DNA nucleoids are linked to protein folding machinery and metabolic enzymes at the mitochondrial inner membrane. *J Biol Chem*, **281**, 25791-802
- Wenz, T., Luca, C., Torraco, A. & Moraes, C.T. (2009) Mterf2 regulates oxidative phosphorylation by modulating mtdna transcription. *Cell Metab*, **9**, 499-511
- Wilson, K.S. & Von Hippel, P.H. (1995) Transcription termination at intrinsic terminators: The role of the rna hairpin. *Proc Natl Acad Sci U S A*, **92**, 8793-7
- Winn, M.D., Isupov, M.N. & Murshudov, G.N. (2001) Use of tls parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr*, **57**, 122-33
- Winn, M.D., Murshudov, G.N. & Papiz, M.Z. (2003) Macromolecular tls refinement in refmac at moderate resolutions. *Methods Enzymol*, **374**, 300-21
- Wolstenholme, D.R. (1992) Animal mitochondrial DNA: Structure and evolution. *Int Rev Cytol*, **141**, 173-216
- Wrutniak-Cabello, C., Casas, F. & Cabello, G. (2001) Thyroid hormone action in mitochondria. *J Mol Endocrinol*, **26**, 67-77
- Wrutniak, C., Cassar-Malek, I., Marchal, S., Rascle, A., Heusser, S., Keller, J.M., Flechon, J., Dauca, M., Samarut, J., Ghysdael, J. & Et Al. (1995) A 43-kda protein related to c-erb a alpha 1 is located in the mitochondrial matrix of rat liver. *J Biol Chem*, **270**, 16347-54
- Wu, Z., Puigserver, P., Andersson, U., Zhang, C., Adelmant, G., Mootha, V., Troy, A., Cinti, S., Lowell, B., Scarpulla, R.C. & Spiegelman, B.M. (1999) Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator pgc-1. *Cell*, **98**, 115-24
- Wurm, C.A. & Jakobs, S. (2006) Differential protein distributions define two sub-compartments of the mitochondrial inner membrane in yeast. *FEBS Lett*, **580**, 5628-34
- Xia, D., Yu, C.A., Kim, H., Xia, J.Z., Kachurin, A.M., Zhang, L., Yu, L. & Deisenhofer, J. (1997) Crystal structure of the cytochrome bcl complex from bovine heart mitochondria. *Science*, **277**, 60-6
- Xu, B. & Clayton, D.A. (1992) Assignment of a yeast protein necessary for mitochondrial transcription initiation. *Nucleic Acids Res*, **20**, 1053-9
- Xu, B. & Clayton, D.A. (1996) Rna-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: An implication for rna-DNA hybrids serving as primers. *EMBO J*, **15**, 3135-43
- Yakubovskaya, E., Mejia, E., Byrnes, J., Hambardjieva, E. & Garcia-Diaz, M. Helix unwinding and base flipping enable human mterf1 to terminate mitochondrial transcription. *Cell*, **141**, 982-93
- Yang, J.C. & Cortopassi, G.A. (1998) Induction of the mitochondrial permeability transition causes release of the apoptogenic factor cytochrome c. *Free Radic Biol Med*, **24**, 624-31
- Yang, M.Y., Bowmaker, M., Reyes, A., Vergani, L., Angeli, P., Gringeri, E., Jacobs, H.T. & Holt, I.J. (2002) Biased incorporation of ribonucleotides on the mitochondrial l-strand accounts for apparent strand-asymmetric DNA replication. *Cell*, **111**, 495-505
- Yankovskaya, V., Horsefield, R., Tornroth, S., Luna-Chavez, C., Miyoshi, H., Leger, C., Byrne, B., Cecchini, G. & Iwata, S. (2003) Architecture of succinate dehydrogenase and reactive oxygen species generation. *Science*, **299**, 700-4
- Yasukawa, T., Reyes, A., Cluett, T.J., Yang, M.Y., Bowmaker, M., Jacobs, H.T. & Holt, I.J. (2006) Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *EMBO J*, **25**, 5358-71
- Yasukawa, T., Yang, M.Y., Jacobs, H.T. & Holt, I.J. (2005) A bidirectional origin of replication maps to the major noncoding region of human mitochondrial DNA. *Mol Cell*, **18**, 651-62
- Yoon, Y.S., Yoon, D.S., Lim, I.K., Yoon, S.H., Chung, H.Y., Rojo, M., Malka, F., Jou, M.J., Martinou, J.C. & Yoon, G. (2006) Formation of elongated giant mitochondria in dfo-induced cellular senescence: Involvement of enhanced fusion process through modulation of fis1. *J Cell Physiol*, **209**, 468-80
- Yoshida, M., Muneyuki, E. & Hisabori, T. (2001) Atp synthase--a marvellous rotary engine of the cell. *Nat Rev Mol Cell Biol*, **2**, 669-77
- Zimmermann, K.C., Bonzon, C. & Green, D.R. (2001) The machinery of programmed cell death. *Pharmacol Ther*, **92**, 57-70

Siglas y acrónimos

aa	Aminoácido
ATP	Adenosina 5' trifosfato
ATPase	Adenosina trifosfatasa
BME	β -mercaptoetanol
DLS	Dynamic light scattering
DNA	Ácido desoxirribonucleico
DNA _n	DNA nuclear
DNA _{mt}	DNA mitocondrial
DNA _{TER}	DNA _{mt} de 28pb en el extremo 3' del 16S rRNA que queda protegido por mTERF en el ensayo de <i>footprint</i> con DNAasa (Kruse <i>et al.</i> , 1989)
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Etilenediaminotetraacético
EMSA	Electrophoretic mobility shift assay. Ensayo electroforético de retardo
ESRF	European Synchrotron Radiation Facility
GF	Cromatografía de exclusión molecular
IPTG	isopropil- β -D-1-tiogalactopiranosido
ITC	Isothermal Titration Calorimetry
kb	kilobases
kDa	kiloDaltons
LB	Medio crecimiento bacteriano Luria-Bertani broth
MAD	Multiple-wavelength Anomalous Dispersion
MIR	Multiple Isomorphous Replacement
MPD	2-metil-2,4-pentenediol
MR	Molecular Replacement. Reemplazo molecular
MS	Mass spectrometry. Espectrometría de masas
OD	Optical density. Densidad óptica
PAC	Plataforma automatizada de cristalografía (PCB)
PAGE	Polyacrylamide gel electrophoresis. Electroforesis en gel de poliacrilamida
pb	pares de bases
PCB	Parque científico de Barcelona

PCR	Polymerase chain reaction
PDB	Protein Data Bank. Banco de datos de proteínas
PEG	Polietilenglicol
pI	Punto isoeléctrico
pM	Peso molecular
PMSF	Phenylmethylsulfonyl fluoride
rmsd	Root Mean Square Deviation. Desviación media cuadrática
RNA	Ácido ribonucleico
mRNA	RNA mensajero
rRNA	RNA ribosómico
SAD	Single-wavelength Anomalous Dispersion
SAS	Small angle scattering. Dispersión a bajo ángulo
SAXS	Small-angle X ray scattering. Dispersión de rayos X a bajo ángulo
SANS	Small-angle neutron scattering. Dispersión de neutrones a bajo ángulo
SB	Medio crecimiento bacteriano super broth
SDS	Dodecil sulfato sódico
SeMet	Selenometionina
T _m	Temperature of melting. Temperatura de fusión
TAE	Tris-acetato EDTA
Tris	Tris(hidroximetil)aminometano

Código de una y tres letras para los aminoácidos

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

Agradecimientos

Vull agrair especialment a la Maria Solà i al Miquel Coll, directora i codirector d'aquesta tesi, per acollir-me al seu grup i donar-me l'oportunitat d'endinsar-me en el món de la investigació.

A la Isabel Usón, perquè sense la seva valuosa ajuda cristal·logràfica, segurament aquest projecte no hagués estat possible.

Al Pau Bernadó, per la seva proximitat. Per estar sempre disposat a explicar, tornar a explicar i discutir sobre SAXS. Segur que ben aviat, en algun lloc sabran valorar la feina ben feta durant tots aquests anys.

Al Jordi Bernués pel seu interès. Al Ramon Crehuet per la seva disposició. Pel temps que els he robat i pels dubtes resolt. Al Quico Canals per la caracterització proteica per espectrometria de masses.

I a la Jenny Colom per la seva infinita paciència. Per la valuosa ajuda en la part de purificació.

A tota la gent de Cri1-7, gràcies per tota l'ajuda científica durant aquests anys, en la poiata i en la part informàtica. Pels sincrotrons i SAUs que hem viscut junts.

En el viatge a Nepal, un gurú ens instruí sobre els mandales. Això és el recordo d'aquella curiosa trobada: un mandala pot representar la vida d'un mateix i a mida que el vas dibuixant, vas agraint en cada etapa a les persones i les situacions que han estat importants. De moment queda pendent el dibuix, però no vull deixar passar l'oportunitat per agrair-vos a tots haver estat presents en aquest camí que m'ha portat fins aquí, fins avui.

A mis padres, por empeñar en mí hasta el último beso. A mis hermanas. Os quiero con locura, estando cerca y en la distancia. Y a mi recién adquirido cuñado Santi por lo feliz que haces a mi hermana.

A mis abuelos y a toda la familia porque sé que siempre podré contar con ellos.

A tota la família del Pep per acollir-me com si fos la meva.

A les meves nenes, les que sempre han estat i que estic orgullosa de conservar, tot i la pila d'anys que han passat. Per les complicitats i els moments màgics. Per les vetllades a casa vostra. Pel viatges, entre d'altres, a Galícia, Argentina, Lisboa, Marroc i Nepal. Espero que repetim aviat. Em sento molt afortunada de tenir-vos.

A la Martona, per ser tant afectuosa i alegre. Per estar sempre tan aprop. Per compartir les primeres passes en la nostra independència. A l'Asun, per ser tant dolça, tant atenta. Per ser tan especial. Per créixer juntes, descalces, ballant. A la Cris, per preocupar-se sempre perquè tot vagi bé. Pels bitllets, al final no va sortir bé l'intent de quedar-nos a Lisboa un dies més. Per posar cos i ànima en tot el que fas. A la Betty, per ser tant divertida. Per la teva fortalesa. Per posar seny, altre cop, quan quasi ens quedem a Boulmane per una vaga d'autobusos. Per tots els cotxes que cal substituir per bicis. A l'Eli, per la seva vitalitat i el munt de plans que sempre té. Per gaudir tant de la teva feina. Perquè després de l'hivern, sempre reneix la primavera. A la Judith, per l'amistat recuperada després dels anys a Canàries. Pels teus pastissos boníssims.

I als nois que hem anat incorporant a les nostres vides. I als que vindran. A l'Oriol, per assessorar-nos en totes aquestes qüestions a nivell econòmic que escapen a la nostra intel·ligència limitada, al François per su inagotable energía cuando se trata de subir una montaña a 5000metros, al David i al Jose.

Als antics veïns de Gràcia, al jeringuilla Roger, per tenir sempre obertes les portes del teu pis i a l'Oriole, per la teva capacitat única de contagiar d'energia a la gent que t'envolta. Per ser la festa personificada. Pels teus flams 'caserus'. Per descobrir-nos Montcortés. I als nous veïns, a Juantxo, por todos los momentos que ahora recordamos con nostalgia. Por los antiguos amigos y las sidras. Y a Sebas, por su dulzura. Por el tiempo que vivimos juntos. Es curioso cómo los cursos de doctorado me dieron la oportunidad de formarme y conocer a gente maravillosa. A Mertxe, a Pau, a Gus, a Frá, a Ornella, a Héctor, a Pablo, a Wifi, a Silvia.

Y a mis pequeñas brujis. Por ser tan especiales.

Por la pócima mágica que añadisteis a mi vida con vuestra amistad.

A Paloma, por su sonrisa pícaro y su dulce voz. Por los cuentos de Cortázar. No se muy bien si fue porque le crecieron alas o se fue en escoba, pero nos espera en París, cumpliendo su sueño de música.

A Babi, a menina que mais quero, por tantas lunas llenas y otras tantas puestas de sol... Por la calma de tu mirada, de tus abrazos. A Oriol, por cuidarte. Espero tenerte siempre cerquita.

I per aquests meravellosos viatges, sí, aquests en els que son necessàries mil tardes de birres per poder alleugerar la ressaca que provoquen...

Per la mezquita azul i el bazar de les espècies. Pel raki, la neu i la incursió a la piscina. A tots els bioquímics, pels mil records condensats en els dos últims anys de carrera, l'últim intent d'aferrar-me a la vida universitària. Perquè en breu hauré de recórrer mig món per poder veure-us a tots. Al Xavi Muñoz, que li ha donat temps a marxar i a tornar, i nosaltres tot just acabem la tesi... A la Cris, que ha abandonat el nostre estimat barri i ja està a NY. Espero que allà es contagiïn de la teva alegria tant com ho hem fet nosaltres. A l'Esterilla, per ser tan dolça i riallera. Per defensar allò que creus. Al Gerard, per fer tan especials certs moments i records. Pels poemes i les cançons. Al Martí, per deixar un doctorat a temps i submergir-se sense por en la incertesa de la vida laboral. Per les ganes d'enredar-se en mil coses. Al Rubisco, per ser tant atent i estar sempre disposat a tot. Per la teva passió pel regne vegetal. A la Gemma, per ser sempre la nostra 'mami' i gaudir del repte de ser-ho ara, de veritat. A l'Albert, que és l'únic que es guanya bé la vida. Pel teatre i la música. Perquè es capaç de contagiar-te sempre de bon humor. Al David, l'inventor del cubating, perquè la ciència és, de vegades, entremaliada amb els que més s'ho mereixen. A Isaac, por su empeño y fascinación por todo lo que hace. Al Xavi Castells, per la seva espontaneïtat i les inigualables gimcanes. I a la Mireia, la bioquímica adoptada, per la seva sensibilitat. Perquè et devem una visita a Sevilla.

I per Mèxic, un congrés de cristal·lografia i dues setmanes de runes espectaculars, peripècies varies, cremades solars, lidocaïna, peixos i banys nocturns. Al Mià, per ser una de les retrobades més especials que la vida em tenia guardada. Pels moments jugats a ser peixos, per tenir sempre les paraules exactes. I a la Nunu. Segur que la petita Aina sap que, amb els pares que té, serà la nena més afortunada de Gràcia... A l'Arnau, per ser el meu company fidel de "la tierra de nunca jamás", per seguir-se emocionat amb cada cim

escalat (després d'uns deu mil diria jo...). Perquè sé que sempre puc comptar amb tu. Tenim un viatge pendent! Al Jordi, perquè vaig anar descobrint mica en mica el teu encant quasi addictiu. Per les festes de disfresses i les inacabables nits de confessions. Us trobem a faltar tant, a tu i al Bladi! A Tiago, por ser um dos garotos mais carinhosos que conheço. Por tu vitalidad inagotable. A David por su espontaneidad, por los lanzamientos a grandes distancias, por todas las anécdotas que recuerdo con tanto cariño. A Tomislav, por tu independencia. Por toda la discreción que te bebiste el día de tu cumpleaños en la Barceloneta. Por llevarnos a conocer Croacia y preocuparte porque todo saliera a pedir de boca. Y a Annie, por ser la mama del viaje y preocuparte por todos nosotros, incluidas las desaparecidas latas de atún.

Vam aconseguir recuperar-nos a base de Jarres. Mil gràcies Paco per tenir sempre les portes obertes.

Al petit grup dels MITO. Els que encara hi son i els que han marxat. A l'Anna perquè guardes, encara no se on, un "frasquito de esencia de alegría". Per tantes complicitats, per les tardes i els sincrotrons inacabables fent-nos mútua companyia. Pels balls. Per un viatge inoblidable a Nepal i pel que ens espera. Perquè haver-te conegut em fa sentir molt afortunada. A Jas, por esas deliciosas noches en tu casa de pintura y arroz persa. Por obligarnos a ir al Oktober fest. Porque fue genial compartir poyata contigo. A Pablo por tu desparpajo y sentido del humor. Porque invades de risas y palabras. A Cristiña porque fue deliciosa la manera que tuviste de hacerte un hueco entre nosotros sin saber muy bien donde te metías. A Alvarito, porque con tu espontaneidad conseguiste, en unos pocos meses, ganarte el cariño de todos. Seguro que te irá muy bien allá donde estés. A Cuppi, porque unos ciclos se cierran y otros se abren. Te deseo mucha suerte.

Als companys recents de laboratori, a la Cris, per ser la rubia més intel·ligent del labo, sobretot ara que marxo. No deixis que et prenguin el pèl. A la Laia pels correbars i les festes de Gràcia. Et desitjo molta sort en el doctorat. A la Sol que ha decidit creuat el mar per poder oferir-nos casa a Buenos Aires. T'enyorem petita...

Als ex-companys del pis de dalt CR1. Als que ja hi eren quan vaig arribar, al Dani i la Raquelilla, per tenir paciència i ensenyar-me en l'estada com a estudiant de pràctiques. A l'Albert, l'Àlex, la Mònica, la Marta i la Carme, per ajudar-me sempre que ho he

necessitat. I a l'Esther F., per estar disposada en qualsevol moment a solucionar-nos qualsevol problema logístic del laboratori. Perquè tens un cor enorme.

A todos los que han pasado por el laboratorio y ya se han ido o están a punto de hacerlo, a Thil, Seb, Carlo... A María la sevillana, por encontrar finalmente tu sitio. Por las risas mientras estuviste aquí. A Silvia por estar siempre dispuesta a echar un cable.

A Leo, porque es tan fácil encariñarse contigo... Por tu espontaneidad y tu locura. Por tus bailes bailongos. Y a Miguelote por su simpatía.

A l'Esther P., per ser tant independent. Per la teva senzillesa. Per aguantar situacions extremes. Que sàpigues que ens tens al teu costat pel que necessitis.

A Mailys, por saber lo que quieres. Por ser tan decidida. Por el viaje tan emocionante que te espera y del que quizá, con un poco de suerte, pueda disfrutar un poquito.

A Lionel, porque silenciosamente como el ladrón que llega de noche, se ha ido haciendo un hueco en nuestras vidas dándonos cuenta especialmente en estos últimos tiempos en que te echamos tanto de menos.

Y a los que aun están, Fabio, Juliana, Zuzanna, Radoslaw y Diana. A la Nayibe per la seva simpatia. A Roeland, no sé si sabe cuán importante fueron para mí las clases aceleradas de cristalografía. Mil gracias. A Robert por las inolvidables fiestas en tu casa. Por los disfraces y la purpurina.

Y a Judith por su alegría y vivacidad, porque es como un divertido ratoncito que nunca para. Por ser tan cariñosa y emotiva. Por apostar por un futuro con Josep.

A la Vane, per la seva energia boja que sempre m'enreda. Per les complicitats. Pels concerts. Perquè 'no estarás sola'. Perquè mica en mica la vida ens mostra els camins amb els que sempre vam somiar.

Als ex-veïns CR2,

A la Rosa i la Maria, la parella més cotilla. A la Rosa per la valuosa ajuda durant aquests anys i a la Maria per la seva rialla sincera. Al Xavi, per les seves reflexions. Per anar fins al final amb les teves conviccions. Per la coope. A la MariQueri, per la teva alegria i els teus mails provocadors. A Luca por su simpatía y su sonrisa. I al Garreta, per trobar, mica en mica, el teu raconet.

I al Pep, per fer-me riure. Perquè m'encanta aixecar-me cada matí entre els teus braços.

Perquè és amb tu amb qui vull compartir el fascinant viatge de la VIDA.