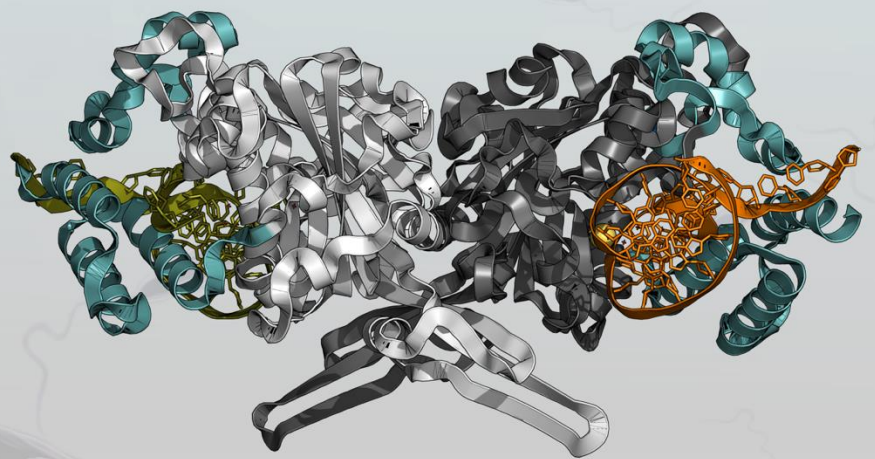


Caracterización funcional y estructural de la Endonucleasa G de *Leishmania infantum*. Exploración de su potencial como diana para el desarrollo de agentes terapéuticos con actividad leishmanicida

Cristina Oliva Correia





Universidad de Alcalá

Doctorado en Señalización Celular

CARACTERIZACIÓN FUNCIONAL Y ESTRUCTURAL DE LA ENDONUCLEASA G
DE *LEISHMANIA INFANTUM*. EXPLORACIÓN DE SU POTENCIAL COMO
DIANA PARA EL DESARROLLO DE AGENTES TERAPÉUTICOS CON
ACTIVIDAD LEISHMANICIDA.

Tesis Doctoral Presentada por

CRISTINA OLIVA CORREIA

Directores

DR. ANTONIO JIMÉNEZ RUIZ

&

DR. FEDERICO GAGO BADENAS

Alcalá de Henares, 2020

¡Gracias!

No podía empezar esta tesis con otra palabra que no fuese gracias. Ahora que parece que ya solo queda el último pasito para completar esta etapa, no puedo evitar echar la vista atrás y darme cuenta de que me llevo un saco lleno de buenos recuerdos gracias a todos los que me habéis acompañado a lo largo del camino.

Gracias Antonio por hacer que me enamorara de la Biología Molecular con tus clases y por invitarme a tu laboratorio cuando ni si quiera había terminado la carrera. Gracias por ser tan buen maestro, por guiarme en cada etapa de este proceso haciéndome aprender de mis errores sin dejar de destacar mis éxitos pero, sobre todo, gracias por no perder la esperanza en esta tesis incluso cuando yo ya lo había hecho. Y a Federico, gracias por aceptarme en tu grupo sin ni si quiera conocerme, por tu paciencia infinita conmigo y por todas las ideas y trabajo que has aportado para que este proyecto haya quedado tan bonito.

Gracias a todos los miembros del Departamento de Biología de Sistemas por formar parte en mayor o menor medida de este proyecto. En especial, gracias a Begoña, a Santiago y a Chilo por aceptarme como okupa en vuestros laboratorios durante los cafés que a veces se alargaban un poco más de la cuenta, por las cañas por los bares de Alcalá y por tener siempre palabras de apoyo y ánimo cuando las he necesitado. Gracias también a todos los miembros de las unidades de apoyo a la investigación sin las cuales este trabajo habría sido imposible. Gracias a las chicas de la Unidad de Cultivos y de la Unidad de Biología Molecular por prestarme su ayuda siempre que la he necesitado y a Juan y Guillermo por ayudarme en mis primeras incursiones con la radiactividad y por esperarme siempre con una sonrisa cuando aparecía por la instalación a dos minutos del cierre.

Gracias infinitas a todos los que hicisteis que el laboratorio se convirtiera en mi segundo hogar. No creo que nunca me vaya a sentir tan a gusto trabajando en ningún sitio como lo he hecho con vosotros. Gracias a Eva, por enseñarme a dar mis primeros pasos en el laboratorio y dejarme en herencia este proyecto tan bonito. Gracias también a David, Carlos, Kilian, Pedro, Héctor y a todos los pollitos por haber sido tan buenos compañeros y estar siempre ahí para echarme una mano, un pie, un buffer o lo que hiciera falta. Gracias a Miguel por ser el mejor compañero de poyata del mundo. A veces en la vida hay golpes de suerte y uno de

los míos fue que te concedieran una beca para hacer el doctorado justo a mi lado. Gracias por ayudarme (casi) siempre sin rechistar, por las risas en los momentos buenos y los abrazos en los que no lo fueron tanto, por las canciones con la guitarra (las de ukelele no te las agradezco tanto), por mi preciosa portada y por traer a nuestro grupito a Edwin y a Lina. A Ana, gracias por contagiar tu alegría allá por donde vas, por esas carcajadas que estoy segura se escuchaban hasta en la biblioteca, por descubrirme Asturias y por cuidarme tanto. A los cuatro, gracias por formar parte de esta pequeña familia de doctorandos que parece que al fin va a coronar, por las vacaciones, los cumpleaños, las navidades, las borracheras y, en definitiva, por haber sido una parte tan importante de esta etapa.

Gracias también a Eva y a Chusi por los desayunos, las escapadas y los capítulos de tu cara me suena (o lo que hubiera), sin vosotras esto no habría sido igual. Gracias también a todos los compañeros del resto de laboratorios y que de alguna forma habéis formado parte de esta tesis. Han sido tantos años que si intento mencionaros a todos seguro que me dejo a alguien, pero no quería dejar de agradeceros que hayáis hecho de esta experiencia algo tan especial.

Quiero agradecer también a mis amigos de TJ de toda la vida el haberme apoyado siempre, aun cuando no tenían claro qué era lo que estaba haciendo (¡que no, que no quería ser bióloga!). Todo esto empezó con vosotros cuando aún estudiábamos juntos en la biblioteca y, aunque nuestras vidas han cambiado mucho desde entonces, siempre encontramos ese ratito para reunirnos y que parezca que no ha pasado el tiempo.

Gracias Alberto por aguantarme y quererme a pesar de mis defectos y mis maldades, por ser mi mayor apoyo durante todos los años que llevamos juntos y por creer más en mí de lo que lo hago yo misma. Te quiero amod.

Por último, quiero dar las gracias a los que más debo. Gracias papá y mamá por darme la oportunidad de estudiar aun cuando las condiciones no eran las más favorables, por apoyarme siempre sin cuestionar mis decisiones, por cuidarme y por quererme tanto. Sin vuestro esfuerzo y trabajo nada de esto habría sido posible, esta tesis es para vosotros. Gracias a mi hermano Luis por ser además el amigo con el que sé que siempre puedo contar y a mi hermano Álex por dejarme verle crecer y convertirse en una personita. Enano, espero que este trabajo te ayude a ver que con esfuerzo puedes llegar a alcanzar todas las metas que te propongas, no te rindas sin al menos intentarlo.

A mis padres, Luis y Álex.

A Alberto.

Tabla de contenidos

Sinopsis	1
Abreviaturas y acrónimos	3
Introducción	5
1.1 Leishmaniasis y particularidades de <i>Leishmania spp.</i>	5
1.1.1 Definición, manifestaciones clínicas y epidemiología de la enfermedad	5
1.1.2 Generalidades de <i>Leishmania spp.</i> , morfología y ciclo de vida	9
1.1.3 Particularidades bioquímicas, celulares y moleculares de <i>Leishmania</i>	14
1.1.3.1 Organización genómica y regulación de la expresión génica	14
1.1.3.2 Mitocondria y kinetoplasto	15
1.1.3.3 Glicosomas y acidocalcisomas	16
1.1.3.4 Metabolismo	17
1.1.4 Tratamiento farmacológico de la leishmaniasis	20
1.2 Mecanismos de muerte celular	24
1.2.1 Apoptosis en organismos pluricelulares	32
1.2.2 Muerte celular regulada y apoptosis en <i>Leishmania spp.</i>	38
1.3 Endonucleasa G	43
1.3.1 Estructura, localización y caracterización enzimática	45
1.3.2 Funciones descritas	48
1.3.3 La EndoG de <i>Leishmania infantum</i> (LiEndoG)	54
Hipótesis	59
Objetivo general	59
Resultados	61

4.1	Los derivados de la 5'-trilitimidina como una nueva clase de compuestos antileishmania: la Endonucleasa G de <i>Leishmania infantum</i> como posible diana terapéutica	62
4.1.1	Hipótesis.....	62
4.1.2	Objetivos	62
4.1.3	Resultados principales	62
4.1.4	Resumen	63
4.1.5	Artículo 1.....	69
4.2	La EndoG de <i>Leishmania infantum</i> es una endo/exonucleasa esencial para la supervivencia del parásito	83
4.2.1	Hipótesis.....	83
4.2.2	Objetivos	83
4.2.3	Resultados principales	83
4.2.4	Resumen	84
4.2.5	Artículo 2.....	89
4.3	Asignación de los dominios presentes en la EndoG de <i>Leishmania infantum</i> en base a su estructura: caracterización de una región que regula la actividad de esta enzima en función del pH y una extensión C-terminal que dicta en gran medida su preferencia por los distintos sustratos de ADN	99
4.3.1	Hipótesis.....	99
4.3.2	Objetivos	99
4.3.3	Resultados principales	99
4.3.4	Resumen	100
4.3.4.1	Características del centro catalítico de LiEndoG	101
4.3.4.2	LiEndoG existe como homodímero	103
4.3.4.3	LiEndoG posee un dominio de autoinhibición que regula su actividad en función del pH.....	103

4.3.4.4	La extensión en el extremo carboxilo terminal de <i>LiEndoG</i> determina su preferencia por los sustratos de cadena sencilla	105
4.3.5	Artículo 3.....	111
	Conclusiones	135
	Bibliografía	139
	Anexo	157

Sinopsis

La familia EndoG está compuesta por un grupo de nucleasas no específicas con dominio $\beta\beta\alpha$ de unión a metal presentes en todos los organismos cuyo genoma ha sido secuenciado, incluyendo los tripanosomátidos. Sin embargo, a pesar de la ubicuidad de estas proteínas, y de que su existencia se conoce hace más de cincuenta años, todavía no están claras cuáles son sus funciones en el interior de las células. Una de las que parece estar mejor caracterizada es la de migrar de la mitocondria al núcleo para degradar el ADN genómico durante la apoptosis, función que ha sido confirmada para la EndoG de *Leishmania infantum* (*LiEndoG*) por nuestro grupo de investigación. No obstante, también se ha demostrado que algunos miembros de esta familia, como por ejemplo los de *S. cerevisiae* (Nuc1p) o *C. elegans* (CPS-6), son necesarios para el correcto desarrollo de estos organismos, apuntando hacia un posible papel dual provida/promuerte para este grupo de enzimas. En este sentido, el trabajo llevado a cabo en esta tesis nos ha permitido comprobar que *LiEndoG*, además de participar en el proceso apoptótico de los parásitos, es esencial para su desarrollo y supervivencia, dado que la reducción de su expresión produce un retraso significativo de su crecimiento y hace que su infectividad sobre los macrófagos sea mucho menor.

Adicionalmente, nuestro trabajo nos ha permitido identificar varios derivados de la 5'-trilitimidina con una potente actividad leishmanicida frente a promastigotes y amastigotes axénicos de *Leishmania infantum*, dentro de los cuales destaca la 3'-O-(isoleucilisoleucil)-5'-O-(3,3,3,-trifenilpropanoil)timidina (**compuesto 28**) por mostrar una actividad particularmente potente sobre los amastigotes intracelulares. El tratamiento de los parásitos con este compuesto produce alteraciones significativas en su potencial de membrana mitocondrial, causa un aumento en la concentración de anión superóxido y cursa con una disminución significativa de la degradación del ADN durante la muerte celular, sugiriendo que podría tener como diana en *L. infantum* una proteína mitocondrial involucrada en la degradación del ADN durante la muerte celular, ambos requisitos presentes en *LiEndoG*. Dado que los nuevos compuestos generados que tienen actividad leishmanicida son además capaces de inhibir la actividad nucleasa de *LiEndoG in vitro*, y que el efecto citotóxico producido sobre los parásitos por el compuesto **Lei49 (compuesto 37)** puede ser contrarrestado cuando se sobreexpresa esta enzima, nuestros resultados apuntan a que

LiEndoG es la diana sobre la que están actuando los derivados de la 5'-trilitimidina. Por tanto, el conjunto de los resultados obtenidos en este trabajo sugiere que la inhibición de *LiEndoG* podría estar produciendo un efecto citotóxico sobre los parásitos, apoyando aún más el papel provida de esta enzima y demostrando el potencial que podría tener como diana terapéutica para el desarrollo de nuevos fármacos con actividad leishmanicida.

Finalmente, en el laboratorio hemos desarrollado un nuevo modelo molecular para una *LiEndoG* dimérica e identificado varios dominios que confieren a esta nucleasa propiedades prácticamente únicas. El primero de estos dominios está constituido por una inserción rica en residuos ácidos que, según nuestro modelo, se localiza de forma muy próxima al centro activo de la enzima y en la misma región espacial ocupada por el inhibidor de la EndoG de *Drosophila melanogaster* en el cristal *DmEndoG/DmEndoGI*. Dado que resultados previos obtenidos en nuestro laboratorio demostraban que *LiEndoG* está inhibida a pHs básicos (8 o superior), pensamos que este dominio podría estar actuando como un sensor de pH y estar produciendo la autoinhibición de la enzima. De hecho, en este trabajo mostramos que la variante de *LiEndoG* en la que se ha eliminado esta región (residuos 145-180) es mucho más activa que su homóloga silvestre a pHs mayores de siete, por lo que esta región constituye un dominio de autoinhibición que regula la actividad de la enzima en función del pH. Adicionalmente, hemos sido capaces de demostrar que la extensión del extremo carboxilo terminal de *LiEndoG* conforma un dominio que dicta la preferencia que tiene esta enzima por digerir ADN de cadena sencilla. Finalmente, hemos descubierto que: i) *LiEndoG* existe como un homodímero donde cada uno de los dos sitios activos se localizan en lados opuestos de la interfaz de dimerización; ii) el reemplazo de la Ser211 del motivo SRGH del centro catalítico por el aspartato canónico de la mayoría de las nucleasas de esta familia da lugar a una enzima catalíticamente deficiente; y iii) la actividad de la variante S211D puede ser rescatada cuando se reemplaza de forma simultánea la Ala247 por una arginina. En conjunto, nuestros resultados apoyan la idea de que *LiEndoG* es una proteína multidominio cuya actividad nucleasa puede ser modulada o incluso abolida mediante la introducción de cambios en su arquitectura como consecuencia de las condiciones en el ambiente intracelular y la interacción con otras proteínas.

Abreviaturas y acrónimos

5hmC	5-hidroximetilcitosina
ADN	Ácido desoxirribonucleico
ADNcs	ADN cadena sencilla
ADNdc	ADN de doble cadena
ADNg	ADN genómico
ADNk	ADN del kinetoplasto
ADNmt	ADN mitocondrial
ADNn	ADN nuclear
AIF	Factor inductor de apoptosis (<i>Apoptosis inducing factor</i>)
Apaf-1	Proteasa apoptótica activadora del factor 1
ARN	Ácido ribonucleico
ARNi	ARN de interferencia
ARNm	ARN mensajero
ARNr	ARN ribosomal
BER	Reparación por escisión de base (<i>Base excision repair</i>)
BERmit	Reparación por escisión de base mitocondrial
CAD	ADNasa activada por caspasas (<i>caspase-activated DNase</i>)
CASP	Caspasa
CPS-6	Proteína de <i>C. elegans</i> homóloga a EndoG
cs	Cadena sencilla
CSR	Recombinación de conmutación de clase (<i>class switch recombination</i>)
CTD	Dominio carboxilo terminal
DAMPs	Patrones moleculares asociados a daño (<i>Damage-associated molecular patterns</i>)
DAPI	4',6-diamidino-2-fenilindol
dc	Doble cadena
DD	Dominio de muerte (<i>Death Domain</i>)
DISC	Complejo inducido por señales de muerte (<i>Death-inducing signaling complex</i>)
DmEndoG	Endonucleasa G de <i>Drosophila melanogaster</i>
DmEndoGI	Inhibidor de la EndoG de <i>Drosophila melanogaster</i>
EDO	Enfermedad de Declaración Obligatoria
eGFP	Proteína fluorescente verde mejorada
EndoG	Endonucleasa G
EndoGI	Inhibidor de EndoG de <i>Drosophila melanogaster</i>
FADD	Proteína asociada con dominio de muerte
FASL	Ligando del fragmento estimulador de la apoptosis
FASR	Receptor del fragmento estimulador de la apoptosis
FEN-1	Flap endonucleasa 1
gp63	Glicoproteína 63
GSP	Gel secretado por promastigotes
hEndoG	Endonucleasa G humana
hEXOg	Exonucleasa G humana

Hsp	Proteínas de choque térmico
HSV-1	Virus del herpes simple tipo 1
IAP	Proteínas inhibidoras de apoptosis (<i>inhibitor of apoptosis protein</i>)
iCAD	Inhibidor de la ADNasa activada por caspasas
LC	Leishmaniasis cutánea
LCD	Leishmaniasis cutánea difusa
LdEndoG	EndoG de <i>Leishmania donovani</i>
LdFEN1	Flap endonucleasa 1 de <i>Leishmania donovani</i>
LDPK	Leishmaniasis dérmica post kala-azar
LiEndoG	Endonucleasa G de <i>Leishmania infantum</i>
LiFEN1	Flap endonucleasa 1 de <i>Leishmania infantum</i>
LMC	Leishmaniasis mucocutánea
LPG	Lipofosfoglicano
LV	Leishmaniasis visceral
MCA	Muerte celular accidental
MCP	Muerte celular programada
MCR	Muerte celular regulada
MEM	Membrana externa mitocondrial
NCCD	Comité de Nomenclatura de Muerte Celular (<i>Nomenclature Committee on Cell Death</i>)
NETs	Trampas celulares de neutrófilos (<i>Neutrophil Extracellular Traps</i>)
NTD	Enfermedades Tropicales Desatendidas
Nuc1p	Nucleasa 1 de <i>S. cerevisiae</i>
NucA	Nucleasa A de <i>Anabaena spp.</i>
NuiA	Inhibidor de la Nucleasa A de <i>Anabaena spp.</i>
Omi/HtrA2	Proteína A2 con requerimiento de alta temperatura
OMS	Organización Mundial de la Salud
PARP1	Poli (ADP-ribosa) polimerasa 1
PDB	<i>Protein Data Bank</i>
PMEM	Permeabilización de la membrana externa mitocondrial
Poly	ADN polimerasa γ
PS	Fosfatidilserina
rLiEndoG	EndoG recombinante de <i>L. infantum</i>
ROS	Especies reactivas de oxígeno
RRP	Receptor de reconocimiento de patógenos
SIDA	Síndrome de inmunodeficiencia adquirida
SMAC/DIABLO	Segundo activador de caspasas derivado de la mitocondria/Proteína de unión directa a IAP con bajo punto isoeléctrico
Spp.	Especies
TGF-β	Factor transformante del crecimiento β
TNF	Factor de necrosis tumoral
TUNEL	Marcaje de extremos rotos del ADN por la transferasa terminal con dUTP
UTR	Regiones no transcritas (<i>untranslated regions</i>)
VIH	Virus de la inmunodeficiencia humana
$\Delta\psi_m$	Potencial de membrana mitocondrial

Introducción

1.1 Leishmaniasis y particularidades de *Leishmania spp.*

1.1.1 Definición, manifestaciones clínicas y epidemiología de la enfermedad

El término leishmaniasis agrupa un conjunto de enfermedades causadas por parásitos protozoos del género *Leishmania*, los cuales se transmiten a través de la picadura de insectos dípteros del género *Lutzomyia* y *Phlebotomus* en el Nuevo y Viejo Mundo, respectivamente. Existen cerca de 20 especies de *Leishmania* patógenas para los seres humanos que pueden ser transmitidas por más de 90 especies diferentes de insectos vectores (1). Estas especies patógenas están presentes en zonas tropicales y subtropicales de América, Europa, Asia, África e incluso se han detectado casos en Australia (2). La transmisión de la enfermedad presenta dos escenarios fundamentales: a) un ciclo zoonótico, en el que animales silvestres y domésticos actúan como reservorio del parásito, incluyéndose en este grupo perros, gatos, roedores, monos o marsupiales, y b) un ciclo antroponótico, presente principalmente en la India y Sudán, en el que el hombre es la única fuente de infección para el insecto vector (3).

La leishmaniasis presenta un amplio espectro de manifestaciones clínicas, yendo desde manifestaciones cutáneas que cicatrizan espontáneamente hasta infecciones viscerales de gravedad que pueden llegar a causar la muerte de los pacientes. La forma en la que se va a manifestar la enfermedad en cada caso estará determinada por la interacción de todos los agentes involucrados en el desarrollo de la misma, es decir, de las características del parásito causante de la infección, del vector que haya transmitido la enfermedad y de la reacción inmune producida en el organismo hospedador como respuesta a la infección (4). De forma general, se reconocen tres tipos de manifestaciones clínicas en la leishmaniasis (1,4–6)

- Leishmaniasis cutánea (LC) es la forma más común y menos grave de la enfermedad. Inicia con el desarrollo de una pápula en el sitio de la picadura del insecto vector que generalmente crece de tamaño y termina ulcerándose. Cabe destacar que, si bien este tipo de manifestación no suele conllevar gravedad, las lesiones suelen producirse en zonas de la piel expuestas como son la cara, el cuello y las extremidades, por lo que pueden conducir a la estigmatización social y a efectos psicológicos sobre los pacientes infectados. Existen también variaciones de la enfermedad como la

leishmaniasis cutánea difusa (LCD), con lesiones diseminadas en cara, pies y manos que abarcan grandes extensiones de la piel.

- Leishmaniasis mucocutánea (LMC) ocurre meses e incluso años después del inicio de la LC como resultado de la diseminación de los parásitos y de una respuesta inmunitaria hiperactiva frente a estos. Esta manifestación de la enfermedad se caracteriza por la destrucción extensiva de las mucosas de la nariz, boca y región rinofaríngea, pudiendo llegar a producir la perforación del tabique nasal y la inflamación de la faringe y laringe hasta incluso llegar a ocasionar su obstrucción. Las lesiones mucocutáneas, además de producir una desfiguración notable, incrementan el riesgo de infecciones secundarias y pueden llegar a imposibilitar la ingesta de alimentos, lo que lleva a un aumento significativo de la mortalidad.
- Leishmaniasis visceral (LV), también conocida como kala-azar, es la forma más grave de la enfermedad y se caracteriza por producir fiebre, pérdida de peso, hepatoesplenomegalia, linfadenopatía, pancitopenia e hipergammaglobulinemia. La LV generalmente cursa de modo crónico y puede ser mortal sin tratamiento. Adicionalmente, tras la recuperación, algunos pacientes pueden desarrollar una forma crónica de LC denominada “leishmaniasis dérmica post kala-azar” (LDPK), que habitualmente requiere tratamientos largos y costosos (7).

El tropismo del parásito por determinados órganos o tejidos depende tanto de la especie causante de la infección como del estado inmune del mamífero hospedador. Por ejemplo, se ha visto que la coinfección de *Leishmania* y el virus de la inmunodeficiencia humana (VIH) incrementa en un factor de entre 100 y 2000 veces el riesgo de desarrollar LV en zonas endémicas para VIH. Así mismo, la infección por VIH está relacionada con una menor respuesta al tratamiento e incrementa la posibilidad de recidivas. Por su parte, la propia LV promueve la progresión clínica del VIH, produciendo que se desarrollen las condiciones que definen el síndrome de inmunodeficiencia adquirida (SIDA) (8). De forma semejante, en los últimos años se ha visto aumentado el número de individuos inmunodeprimidos por causas diferentes al VIH que han desarrollado LV. La gran mayoría de estos individuos son pacientes trasplantados o pacientes reumatológicos, hematológicos u oncológicos que han sido sometidos a tratamientos inmunosupresores. Estos tratamientos obstaculizan el control

inmunológico de la LV y promueven la reactivación de una infección latente o la imposibilidad de controlar una nueva infección (9).

La leishmaniasis presenta una distribución global con importantes focos de infección en Centro y Suramérica, el sur de Europa, África, Oriente Medio y la India, representando un problema de salud pública principalmente en países en vía de desarrollo. Según datos de la Organización Mundial de la Salud (OMS), en 2018 esta enfermedad era endémica en regiones tropicales y subtropicales de casi 100 países en todo el mundo (**Figura 1**) (10) y se estima que cada año se producen entre 0.7 y 1 millón de nuevos casos, con una mortalidad asociada de entre 26 mil y 65 mil personas anualmente (11). Estas cifras destacan la importancia de la enfermedad como problema de salud pública no solo en países en vías de desarrollo, sino también en países desarrollados. Además, se trata de una enfermedad dinámica cuya incidencia se ve afectada por cambios medioambientales, desplazamientos masivos de población desencadenados por conflictos geopolíticos, así como por el calentamiento global que afecta a la ecología y distribución del insecto vector (12), por lo que tanto su distribución como su incidencia pueden verse afectados durante los próximos años. Adicionalmente, hay que tener en cuenta que los datos oficiales subestiman la incidencia real de la enfermedad debido a que en muchos países la leishmaniasis no es una enfermedad de notificación obligatoria y a que se trata de una enfermedad infradiagnosticada, especialmente en zonas rurales con difícil acceso a servicios médicos (1).

Afortunadamente, la incidencia global de la LV ha disminuido sustancialmente en la pasada década, pasando de 200-400 mil nuevos casos en 2012 a 50-90 mil en 2017. De todos los casos reportados a la OMS, más del 95% se concentran en Bangladesh, Brasil, China, Etiopía, India, Kenia, Nepal, Somalia, Sudán y Sudán del Sur. Por su parte, para la LC y LMC, la OMS estima una incidencia anual de entre 0.6 y 1 millón de casos al año, los cuales se concentran en Afganistán, Argelia, Brasil, Colombia, Irán, Iraq y Siria (4,11).

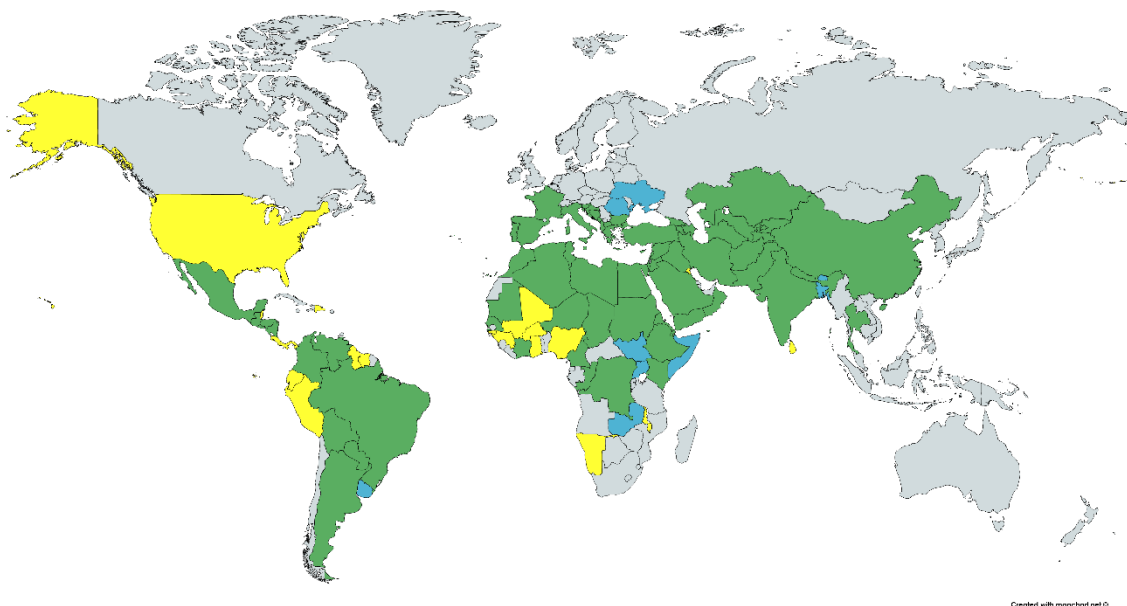


Figura 1. Representación de los países endémicos para leishmaniasis cutánea (amarillo), visceral (azul) o cutánea y visceral (verde). Datos de la OMS basados en los casos reportados en 2018. Mapa creado con mapchart.net.

En España, tanto la LC como la LV son zoonosis endémicas, siendo la principal especie parásita causante de ambas formas de la enfermedad *L. infantum*. Su principal reservorio es el perro, aunque también pueden serlo gatos, roedores y otras especies silvestres como lince, lobos, y liebres. Los flebótomos *Phlebotomus perniciosus* y *Phlebotomus ariasi*, responsables de la transmisión de *L. infantum* en nuestro país, se encuentran distribuidos por todo el territorio peninsular, por lo que el riesgo de infección está presente en cualquier área en la que haya circulado el parásito (13,14). Desde el año 1997 la leishmaniasis es una Enfermedad de Declaración Obligatoria (EDO) sujeta a vigilancia epidemiológica en España. Los primeros reportes obtenidos a partir de esa fecha mostraban una incidencia de unos 15-25 casos anuales, sin embargo, entre 2009 y 2013, se reportaron más de 400 casos, la mayoría de ellos repartidos entre las localidades de Fuenlabrada y Leganés, suponiendo el brote de leishmaniasis más grande del que se tenga conocimiento en España y en Europa (14,15). En la actualidad, este brote se encuentra controlado y, según el último informe anual de resultados de Vigilancia Epidemiológica de las enfermedades infecciosas, en 2016 se comunicaron 197 casos de leishmaniasis en España, siendo 4 de estos casos importados de África. Madrid fue la Comunidad Autónoma con mayor número de casos notificados (66),

seguida de Baleares (36) y Cataluña (35), estando solo exentas de notificación seis Comunidades Autónomas (16).

1.1.2 Generalidades de *Leishmania spp.*, morfología y ciclo de vida

El género *Leishmania* pertenece a la familia de los parásitos protozoos *Trypanosomatidae*, ubicada dentro del orden *Kinetoplastida*. Dicho orden debe su nombre a la presencia de una sola mitocondria que se extiende a lo largo del parásito y cuyo ADN crea una estructura única situada próxima al cuerpo basal del flagelo denominada kinetoplasto (17). La familia *Trypanosomatidae* incluye, además de *Leishmania*, otros organismos parásitos como *Trypanosoma cruzi*, causante de la enfermedad de Chagas, o *Trypanosoma brucei*, causante de la enfermedad del sueño.

Dentro del género *Leishmania* se diferencian los subgéneros *Leishmania (Leishmania)* y *Leishmania (Vianna)*, clasificando las distintas especies de parásitos en función de si proliferan en la región suprapilórica o perpilórica del tracto digestivo del flebótomo, respectivamente (2,18).

Los parásitos del género *Leishmania* presentan un ciclo de vida dimórfico, alternando entre la forma promastigote en el insecto vector, y la forma amastigote cuando está en el hospedador mamífero. La diferenciación de *Leishmania* a cada una de estas formas es resultado de la adaptación de los parásitos a las diferentes condiciones ambientales con las que se encuentra al alojarse en sus distintos hospedadores (invertebrado y mamífero).

Morfológicamente, los promastigotes presentan un cuerpo celular alargado de un tamaño comprendido entre 5-15 μm de longitud y un flagelo anterior largo y móvil que se extiende fuera del bolsillo flagelar. Este flagelo proporciona a los parásitos la fuerza propulsora necesaria para su desplazamiento a través del tracto digestivo del insecto (**Figura 2A**). En cuanto a los amastigotes, presentan un cuerpo celular de forma ovalada o redondeada con un tamaño de unos 2-4 μm de longitud y un flagelo corto e inmóvil (**Figura 2B**). Si bien es cierto que existen diferencias morfológicas significativas entre los estados promastigote y amastigote, ambos conservan varios elementos de la arquitectura celular básica como son el núcleo, orgánulos de copia única como el aparato de Golgi y la mitocondria que se extiende hasta la región anterior del núcleo; y el flagelo, que se extiende desde el cuerpo basal del

parásito. En la base del flagelo se encuentra además el bolsillo flagelar, una invaginación de la membrana celular que constituye el único sitio de endocitosis y exocitosis en estos parásitos (19,20) (**Figura 2A y 2B**).

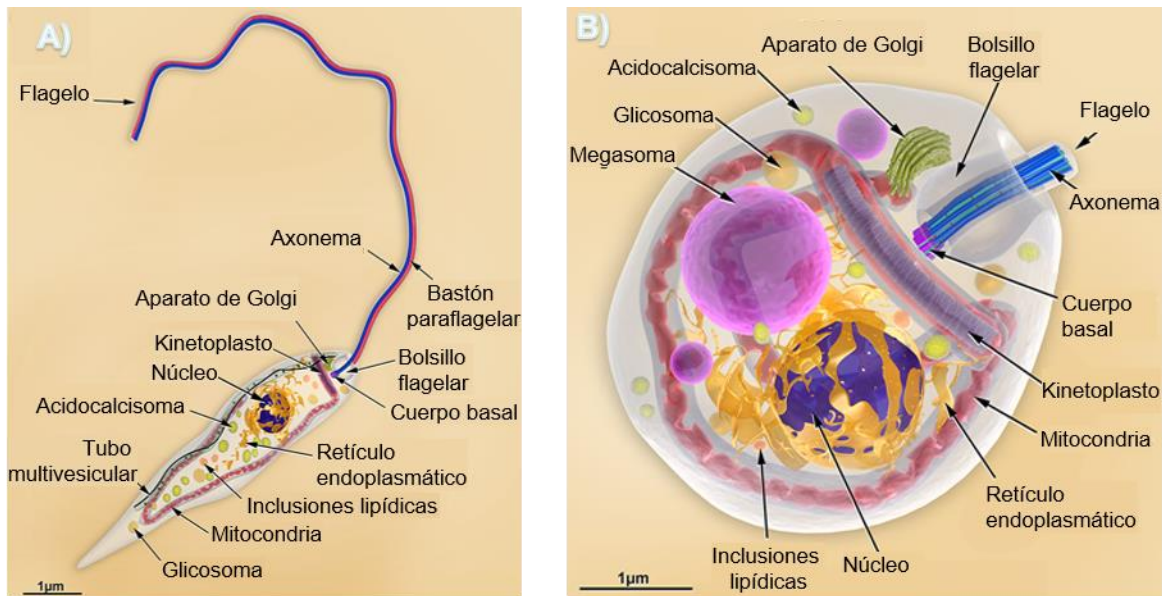


Figura 2. Morfología de las formas A) promastigote y B) amastigote de *Leishmania* spp. Imágenes adaptadas de (19).

El ciclo de transmisión de la enfermedad comienza cuando un flebótomo hembra se alimenta de la sangre de un mamífero infectado, ingiriendo con ello macrófagos parasitados con la forma amastigote de *Leishmania* (**Figura 3.1**). Estos amastigotes van a ser liberados del macrófago rápidamente (**Figura 3.2 y 3.3**) y se van a diferenciar a promastigotes procíclicos (**Figura 3.4**), estando esta diferenciación propiciada por cambios en las condiciones del entorno como la disminución de la temperatura y el aumento del pH. Esta forma del parásito es poco infecciosa para el mamífero, tiene poca motilidad y es altamente replicativa. Los promastigotes procíclicos van a replicarse en el interior de la matriz peritrófica (**Figura 3.5**), estructura formada por una malla de quitina y proteínas secretadas por las células epiteliales del intestino que encierra la sangre que está siendo digerida en la parte posterior del intestino del insecto vector.

Pasados unos pocos días, los promastigotes procíclicos disminuyen su capacidad replicativa y se diferencian a una forma elongada, con un flagelo alargado y mucha motilidad denominada

promastigote nectomonado. Los promastigotes nectomonados van a escapar de la matriz peritrófica y van a comenzar a desplazarse hacia la parte anterior del intestino del insecto. Estos promastigotes tienen además la capacidad de insertarse en las microvellosidades del epitelio intestinal a través de su flagelo, evitando con ello ser eliminados por los movimientos peristálticos del intestino. Se ha visto que este proceso de adhesión está facilitado por remodelaciones estructurales del glucoconjugado de superficie lipofosfoglicano (LPG) y su interacción con proteínas del epitelio intestinal.

Una vez que estos protozoos alcanzan el intestino medio, se van a diferenciar a promastigotes leptomonados, los cuales vuelven a ser altamente replicativos y van a producir una infección masiva del intestino medio anterior (**Figura 3.6**). Los promastigotes leptomonados son además los responsables de la producción del “gel secretado por promastigotes” (GSP). Este gel juega un papel fundamental en la transmisión al hospedador mamífero, puesto que produce que el insecto vector tenga que regurgitarlo durante la ingesta de sangre, inoculando los parásitos embebidos en él al hospedador.

Finalmente, algunos de los leptomonados se diferencian a su forma final conocida como promastigote metacíclico (**Figura 3.7**). Los promastigotes metacíclicos tienen un tamaño de 15-20 μm x 1.5-3.5 μm y un flagelo de 15-25 μm , son altamente infectivos en mamíferos y se distribuyen tanto en la parte anterior del intestino del insecto como en su aparato picador. El proceso completo de transformación de los promastigotes procíclicos a promastigotes metacíclicos se conoce con el nombre de metaciclologénesis. En ella, además de los cambios morfológicos anteriormente descritos, tienen lugar una serie de transformaciones bioquímicas y moleculares como son la remodelación de los glucolípidos de membrana, siendo las más estudiadas las remodelaciones del LPG y la expresión de proteínas de membrana específicas del estado promastigote, como la metaloproteasa gp63, también conocida como leishmanolisina. Cuando el flebótomo se alimenta de nuevo de sangre (**Figura 3.8**), deposita la forma metacíclica del parásito en el interior de la piel (**Figura 3.9**), dándole continuidad al ciclo de transmisión de la infección (21–25).

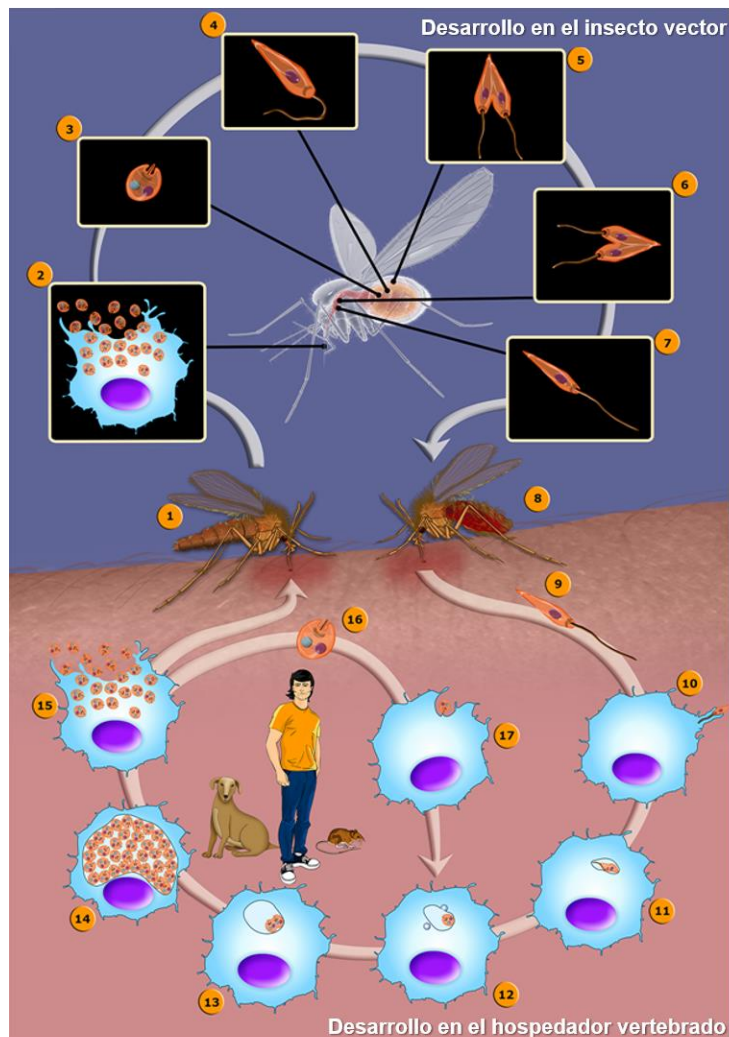


Figura 3. Ciclo de vida de *Leishmania* spp. El ciclo de vida comienza cuando un flebótomo hembra se alimenta de un mamífero infectado (1). Los amastigotes contenidos en los macrófagos son liberados en el insecto (2). Los amastigotes (3) se diferencian a promastigotes procíclicos (4), forma altamente replicativa que se multiplica en la matriz peritrófica (5). Los parásitos migran a la parte anterior del intestino donde reinician su replicación (6), colonizando el tracto digestivo del insecto. Finalmente, los parásitos se diferencian a promastigotes metacíclicos (7), forma infectiva del parásito. El flebótomo hembra se alimenta de un nuevo mamífero hospedador (8) e inocula los promastigotes metacíclicos vía regurgitación (9). Estos promastigotes son fagocitados por células fagocíticas del organismo hospedador (10) donde se diferencian a amastigotes (11) dentro la vacuola parasitófora (12) y se multiplican (13). La replicación intracelular de los amastigotes (14) culmina con la ruptura de los macrófagos (15) y la consecuente liberación de los amastigotes contenidos en los mismos (16), infectando así nuevas células (17), lo que lleva a la propagación y establecimiento de la enfermedad. En la parte central de la imagen se representan algunos de los reservorios mamíferos más importantes en el ciclo de vida del parásito. Imagen adaptada de (19).

En el interior del hospedador mamífero, *Leishmania* es un patógeno intracelular obligado a vivir en el interior de las células fagocíticas (**Figura 3.10**), por lo que ha desarrollado diferentes estrategias encaminadas a evadir la respuesta del sistema inmune del hospedador. Las moléculas de superficie de *Leishmania* LPG y gp63 que hemos mencionado anteriormente juegan un papel importante en este proceso, ya que ambas actúan bloqueando la lisis

mediada por el complemento (26,27). Más concretamente, la metaloproteasa gp63 va a procesar la proteína del complemento C3b, dando lugar a la forma inactiva iC3b que no es capaz de llevar a cabo la lisis, pero si conserva la capacidad de opsonizar a los parásitos, marcándolos como diana para las células fagocíticas y favoreciendo su internalización en sus células hospedadoras. Por su parte, la mayor longitud del LPG en los promastigotes metacíclicos previene la unión de las subunidades C5b-C9 (28,29).

Adicionalmente, se ha demostrado que la presencia de promastigotes apoptóticos en el inóculo es necesaria para que la infección concluya con éxito (30). Bien es sabido que la exposición de fosfatidilserina (PS) en la cara externa de las células apoptóticas promueve su internalización en los fagocitos para su limpieza (31). Además, la fagocitosis mediada por el reconocimiento de PS estimula la producción del factor TGF- β por parte de los macrófagos, citoquina que actúa inhibiendo la respuesta inflamatoria (32). De esta forma, la inoculación de parásitos apoptóticos que exponen PS en su membrana hace que se active este mecanismo de fagocitosis, evitando la respuesta inflamatoria y favoreciendo la supervivencia de otros promastigotes no apoptóticos fagocitados por la misma célula (30).

Los neutrófilos son los primeros en fagocitar a los promastigotes metacíclicos y actúan como células hospedadoras intermediarias en las cuales los parásitos permanecen sin replicarse (33). De hecho, se ha propuesto que *Leishmania* utiliza a los neutrófilos como un “caballo de Troya” para entrar silenciosamente en los macrófagos, ya que estos reconocen la exposición de PS por parte de los neutrófilos que han entrado en apoptosis tras internalizar al parásito y los fagocitan (34). Una vez dentro del macrófago, los promastigotes se van a ubicar en el interior la vacuola parasitófora (**Figura 3.11**), donde el aumento de temperatura a 32-37°C (en función de si se trata de una infección cutánea o visceral) y la acidificación del medio (pH 4.5-5.5) va a producir que se diferencien a amastigotes intracelulares (**Figura 3.12**)(35). Ya en el interior del macrófago, los amastigotes comenzarán a replicarse (**Figura 3.13 y 3.14**) y, tras varios días desde el inicio de la infección, causarán la lisis de la célula hospedadora y la consiguiente liberación de los amastigotes al exterior celular (**Figuras 3.15 y 3.16**). Estos amastigotes serán incorporados por nuevas células fagocíticas, extendiéndose así la infección (**Figura 3.17**).

Finalmente, cuando un mosquito se alimenta de la sangre de un hospedador mamífero infectado, se lleva consigo los amastigotes alojados en los macrófagos de la piel, cerrándose

así el ciclo de vida de este parásito. En función del lugar donde el parásito se establezca y multiplique, se desarrollarán las distintas formas clínicas de la enfermedad. Cuando ocurre en la dermis se desarrollará LC; si es en el hígado y el bazo LV y si migran a la mucosa orofaríngea se desarrollará LMC.

1.1.3 Particularidades bioquímicas, celulares y moleculares de *Leishmania*

1.1.3.1 Organización genómica y regulación de la expresión génica

A pesar de que, como hemos visto, existe gran variabilidad en cuanto a patogenicidad y tropismo celular entre las diferentes especies del género *Leishmania*, todas ellas presentan una estructura genómica muy similar con un alto grado de conservación en contenido y orden de los genes (sintenia). Los genomas de las especies del Viejo Mundo (*L. major*, *L. donovani*, *L. infantum*) presentan 36 cromosomas, mientras que las especies del Nuevo Mundo presentan 34 (*L. mexicana*) o 35 cromosomas (*L. braziliensis*) (36). Durante muchos años se creyó que el genoma de estos protozoos era estrictamente diploide, pero se ha demostrado que *Leishmania* presenta lo que se ha denominado aneuploidía mosaico, una característica particular de estos parásitos por la cual cada célula de una población puede presentar distinto número de copias de uno o más cromosomas. Se ha propuesto que la presencia de esta plasticidad genética en *Leishmania* le dota de una poderosa estrategia para adaptarse a las condiciones cambiantes del entorno y para el desarrollo de resistencia a los medicamentos (37).

El análisis de los genomas de diferentes especies de *Leishmania* ha permitido identificar cerca 8300 genes que codifican proteínas y unos 900 genes que codifican ARNs. De los genes que codifican proteínas, solo el 35% tienen asignada una función biológica en base a caracterizaciones experimentales (3.7%) o a la similitud de su secuencia con proteínas de función conocida en otros organismos (31.5%). Además, estos genomas se caracterizan por presentar una elevada densidad génica y un elevado contenido en G+C tanto en regiones codificantes (62.5%) como no codificantes (57.3%) (38).

En *Leishmania*, los genes codificantes se organizan en largas secuencias policistrónicas que pueden alcanzar los cientos de genes y cuya función habitualmente no está relacionada entre sí. Estas regiones policistrónicas pueden ser divergentes, cuando el sentido de la transcripción

es hacia los telómeros; o convergentes, cuando el sentido de la transcripción es desde los telómeros. La transcripción se inicia a partir de una región *switch* desde donde la ARN polimerasa II inicia su actividad de forma bidireccional (38).

El ARN policistrónico generado tras la transcripción va a sufrir un proceso de *trans-splicing* y poliadenilación que culminará con la generación de los correspondientes ARN mensajeros (ARNm) maduros preparados para su traducción. Estos ARNm maduros contienen una secuencia de 39 nucleótidos denominada secuencia del mini-exón o *spliced leader* en su extremo 5' y una cola de poliadeninas en su extremo 3'. Cabe señalar que la secuencia del mini-exón se sitúa a solo 30-100 nucleótidos corriente arriba del codón iniciador ATG, por lo las regiones no transcritas o UTRs (del inglés, *untranslated regions*) son relativamente cortas en los tripanosomátidos (39).

Dada la organización policistrónica de los genes codificantes de proteínas, y la ausencia de promotores específicos para la ARN polimerasa II, la expresión génica en *Leishmania* está regulada principalmente por el número de copias génicas, el procesamiento postranscripcional del ARN policistrónico, la tasa de degradación de los ARNm y la eficiencia de su traducción (39).

1.1.3.2 Mitocondria y kinetoplasto

A diferencia de los cientos o miles de mitocondrias que podemos encontrar en la mayoría de las células eucariotas, los tripanosomátidos contienen una única mitocondria, haciendo que su correcto funcionamiento sea vital para la supervivencia del parásito. Esto ha propiciado que en los últimos años haya aumentado el interés en este orgánulo como posible diana terapéutica. La mitocondria de *Leishmania* se extiende a lo largo de todo el cuerpo del parásito (**Figura 4**) y, dependiendo de las condiciones ambientales y los recursos nutricionales disponibles, puede llegar a ocupar hasta el 12% del volumen celular (40).

Como ya mencionamos anteriormente en esta introducción, los organismos del orden *Kinetoplastida* deben su nombre a la presencia en su mitocondria de una región especializada que contiene el ADN mitocondrial (ADNmt) denominada kinetoplasto. En el parásito, el kinetoplasto se localiza próximo al cuerpo basal del flagelo (**Figura 4**), siendo su posición

relativa respecto al núcleo un factor importante para determinar la fase del ciclo celular en la que se encuentra.

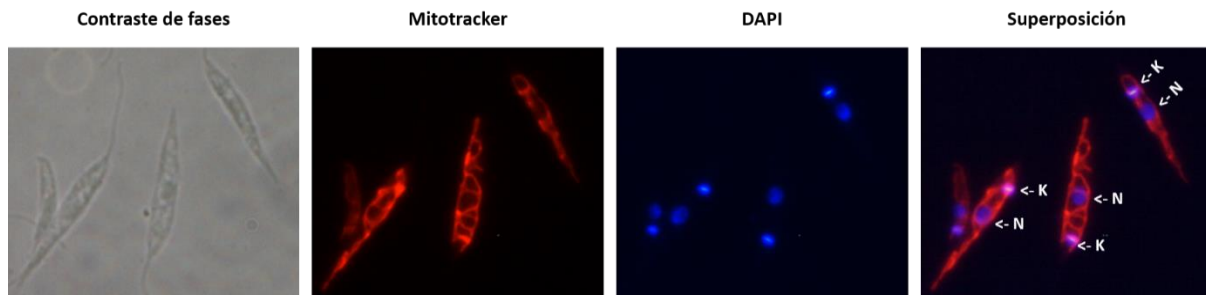


Figura 4. Imagen de microscopía de promastigotes de *L. infantum* sometidos a tinciones de su mitocondria (Mitotracker Red) y ácidos nucleicos (DAPI). El Mitotracker marca de forma específica la mitocondria, mediante su utilización puede visualizarse una única mitocondria que se extiende a lo largo del cuerpo celular de los promastigotes. El DAPI (4',6-diamidino-2-fenilindol) se une a regiones enriquecidas en adenina y timina marcando el ADN, permitiendo visualizar tanto el núcleo (N) como el kinetoplasto (K) de los parásitos. Por último, la **superposición** de las señales de fluorescencia del Mitotracker y el DAPI permite visualizar la posición relativa del núcleo, el kinetoplasto y la mitocondria en el cuerpo celular del promastigote. (Imágenes originales obtenidas en nuestro laboratorio).

El ADN del kinetoplasto (ADNk), representa aproximadamente el 30% del ADN celular total y está constituido por miles de moléculas circulares entrelazadas formando una red única. Existen dos tipos de estructuras de ADN circular en el kinetoplasto: los minicírculos, que típicamente tienen un tamaño de 0.5 a 10 kb (dependiendo de la especie) y de los que hay entre 2.000 y 10.000 unidades; y los maxicírculos, cuyo tamaño oscila entre las 20 y 40 kb y de los cuales solo hay unas 20-25 copias idénticas. Estos maxicírculos son análogos al ADNmt de eucariotas superiores y codifican ARNs ribosomales (ARNr) y algunas proteínas de la cadena respiratoria mitocondrial. Por su parte, los minicírculos son heterogéneos en su secuencia de nucleótidos y codifican ARNs guía que participan en el proceso de edición de ARN transcrito de los maxicírculos para dar lugar a ARN funcionales (17).

1.1.3.3 Glicosomas y acidocalcisomas

Los organismos pertenecientes al orden *Kinetoplastida* poseen unos orgánulos globulares con una única membrana lipídica denominados glicosomas (41). Los glicosomas son peroxisomas modificados que contienen la mayor parte de las enzimas involucradas en las rutas glucolítica y gluconeogénica, así como enzimas implicadas en otros procesos celulares como la ruta de

las pentosas fosfato, la beta-oxidación de ácidos grasos, la protección frente al estrés oxidativo, la síntesis de pirimidinas y la recuperación de purinas. La cantidad de glicosomas presentes en una misma célula y su contenido enzimático varía tanto según la especie de tripanosomátido como, dentro de una misma especie, según la disponibilidad de nutrientes y de las distintas necesidades energéticas durante su ciclo de vida (42).

Otros orgánulos característicos de los tripanosomátidos son los acidocalcisomas. Estos orgánulos, similares a los lisosomas, se caracterizan por tener un pH ácido y altas concentraciones de fosfatos y cationes de sodio, potasio, magnesio, zinc, calcio y hierro. Entre las funciones descritas para los acidocalcisomas están el almacenamiento de compuestos de alta energía, calcio y otros cationes, así como la regulación osmótica y del pH intracelular. Adicionalmente, al tratarse de la principal fuente de almacenamiento de Ca^{2+} intracelular, está implicado en los procesos de señalización por calcio, fundamentales para la adaptación a los cambios ambientales y durante el proceso de invasión de la célula hospedadora (43,44).

1.1.3.4 Metabolismo

Como hemos comentado anteriormente en esta introducción, *Leishmania* presenta un ciclo de vida dimórfico en el que alterna entre el estado amastigote intracelular en el mamífero hospedador y el estado promastigote en el interior del insecto vector. Esta situación hace que los parásitos se enfrenten a situaciones cambiantes de disponibilidad de recursos energéticos, pH, concentración de O_2 y CO_2 y ambiente redox, por lo que deben poseer distintas vías metabólicas que les permitan adaptarse y sobrevivir a las mismas. A continuación, se describen brevemente algunas de las particularidades del metabolismo de *Leishmania spp.* (45).

- Metabolismo de los carbohidratos

Leishmania expresa en todos sus estadios las enzimas necesarias para la glucólisis y la vía de las pentosas fosfato, las cuales utiliza para la generación de energía y equivalentes de reducción, y la generación de precursores para la síntesis de ADN y ARN, respectivamente. Como hemos mencionado, muchas de estas enzimas están parcial o totalmente secuestradas en el interior del glicosoma, siendo el principal orgánulo implicado en el metabolismo de carbohidratos. Es en el interior de este orgánulo donde se inicia la glucólisis llegando hasta

la formación de 3-fosfo-glicerato, el cual es liberado al citoplasma para ser transformado a piruvato con la consiguiente formación de ATP a partir de ADP (45).

- Ciclo de Krebs y respiración mitocondrial

La mitocondria de *Leishmania* contiene todos los componentes necesarios cadena respiratoria mitocondrial funcional, estando sus proteínas codificadas tanto en genes nucleares como mitocondriales.

La respiración mitocondrial parece ser importante para todos los estadios de *Leishmania*. Cuando la biodisponibilidad de glucosa es alta, el piruvato producto de la glucólisis es transportado a la mitocondria y oxidado a CO₂, con la consiguiente obtención de energía. La dependencia del ciclo de Krebs en los amastigotes es aún mayor, dado que en los estadios intracelulares, a pesar de que la biodisponibilidad de glucosa es baja, hay concentraciones relativamente altas de aminoácidos y lípidos que pueden ser catabolizados a Acetil-CoA para iniciar el ciclo de Krebs (45).

- Almacenamiento de energía

Leishmania utiliza un único oligosacárido de reserva denominado manán formado por cadenas lineales de β-1,2-manosa. El manán se sintetiza en el citoplasma y se acumula en los promastigotes. También se ha comprobado que el manán se acumula en amastigotes intracelulares, indicando que las fuentes de carbono no son limitantes en el fagosoma del macrófago (45).

En cuanto al almacenamiento de energía a largo plazo, al igual que en otros protistas y metazoos, *Leishmania* emplea lípidos de reserva. Los ácidos grasos, que pueden ser sintetizados *de novo* o recogidos del medio extracelular, son incorporados en moléculas de triacilglicerol que serán posteriormente almacenadas en cuerpos lipídicos. En ausencia de nutrientes, estos ácidos grasos son liberados y catabolizados por β-oxidación en los glicosomas y la mitocondria (45).

- Biosíntesis de lípidos

El genoma de *Leishmania* codifica la maquinaria de la ácido graso sintetasa tipo II. Sin embargo, esta vía parece no ser la principal responsable de la síntesis de ácidos grasos *de novo* en tripanosomátidos. En su lugar, existe una familia de β-cetoacil-CoA sintetasas o

elongasas que parecen ser las principales encargadas de la biosíntesis de ácidos grasos. Adicionalmente, estos parásitos son capaces de sintetizar esfingolípidos *de novo* por una vía iniciada por la serinpalmitoiltransferasa. Curiosamente, esta enzima solo se expresa en los promastigotes, por lo que los amastigotes tienen que obtener los esfingolípidos del hospedador (45).

- Metabolismo de aminoácidos

Los promastigotes y amastigotes de *Leishmania* pueden utilizar los aminoácidos como su fuente principal o única de carbono en ausencia de carbohidratos. De hecho, existen estudios que sugieren que el catabolismo de aminoácidos es importante para la proliferación de los amastigotes en el fagolisosoma del macrófago (45).

- Recuperación y síntesis de nucleótidos

Los parásitos del género *Leishmania* son auxótrofos para purinas, puesto que no poseen las enzimas necesarias para su síntesis *de novo*. Los promastigotes expresan nucleotidasas y nucleasas que median la degradación de nucleótidos y ácidos nucleicos extracelulares, además de transportadores de membrana que permiten la posterior incorporación de nucleósidos de purina y pirimidina al interior celular. Adicionalmente, se cree que los amastigotes expresan transportadores de membrana adicionales o alternativos que permiten también la incorporación de nucleobases de purina y uracilo (45).

Estos parásitos, además de incorporar nucleósidos de pirimidina del medio extracelular, tienen la maquinaria enzimática necesaria para su síntesis *de novo*. La síntesis de pirimidina se lleva a cabo a partir de glutamato, bicarbonato y aspartato y transcurre entre el citoplasma y el glicosoma, donde tienen lugar los dos últimos pasos de la ruta biosintética (45).

- Vitaminas y cofactores

Leishmania es auxótrofa para varias vitaminas y cofactores por lo que tienen que ser obtenidas del medio. Entre estos compuestos encontramos: ácido fólico, biopterina, ácido pantotico, piridoxina, biotina, tiamina, riboflavina, ácido nicotínico y cianocobalamina (45).

- Sistema redox

Los tripanosomátidos sintetizan glutatión, glutationil-espermidina, tripanotión y ovotiol, cuatro tioles de bajo peso molecular que conjuntamente forman un sistema redox único de tripanosomátidos. De todos ellos, el tripanotión es el más abundante dentro de la célula, por lo que debe permanecer en un estado reducido, presentando una concentración aproximadamente constante durante todas las fases de desarrollo del parásito. Además del papel fundamental que tiene el tripanotión en el mantenimiento del equilibrio redox de estos parásitos, también se ha descrito que actúa como donante de equivalentes de reducción para algunas enzimas (45).

1.1.4 Tratamiento farmacológico de la leishmaniasis

Antes de adentrarnos en las estrategias terapéuticas que existen en la actualidad para el tratamiento de la leishmaniasis, es importante reseñar que durante los últimos años se han ido implementado diversas estrategias para tratar de impedir el ciclo de transmisión de la enfermedad y controlar así su propagación. Estas estrategias se han centrado principalmente en el control del insecto vector mediante la utilización de insecticidas y en la vigilancia epidemiológica de los animales reservorios de la enfermedad. No obstante, la falta de continuidad de estas medidas solo ha sido capaz de producir resultados a corto plazo que, si bien resultaron prometedores, no han sido suficientes para controlar la enfermedad (46). Adicionalmente, durante los últimos años se han hecho grandes esfuerzos de investigación para intentar desarrollar vacunas eficaces para humanos mediante el empleo de parásitos modificados genéticamente o muertos, extractos antigénicos y antígenos recombinantes. Lamentablemente, hasta la fecha, estas vacunas solo han mostrado una eficacia parcial en ensayos clínicos, por lo que de momento no existe una vacuna eficaz y segura que ofrezca una alta inmunogenicidad y capacidad protectora y pueda ser empleada para la profilaxis de la leishmaniasis en humanos (47). Por su parte, desde 2004 se han aprobado cuatro vacunas para combatir la leishmaniasis canina, dos en Brasil, Leishmune® y Leish-Tec®; y otras dos en Europa, CaniLeish® y LetiFend®. A pesar de que los resultados de los ensayos clínicos iniciales eran alentadores, después varios años de comercialización existen bastantes dudas sobre la eficacia y efectividad de estas vacunas. De hecho, la licencia de comercialización de Leishmune® fue retirada en 2014 por el Ministerio de Agricultura de Brasil tras la falta de

efectividad demostrada en ensayos clínicos fase III. Adicionalmente, los últimos estudios realizados con las tres vacunas cuyo uso continúa autorizado ponen en entredicho su utilidad para el control de la leishmaniasis canina, sugiriendo incluso que la vacunación puede llevar a una infravaloración del número de cánidos infectados porque: i) reducen la apariencia o gravedad de los síntomas, haciendo que individuos infectados pasen desapercibidos; y ii) pueden interferir en el diagnóstico serológico de los perros infectados. De esta forma, el uso de estas vacunas podría de hecho actuar en detrimento del control global de la leishmaniasis canina, por lo que es necesario continuar investigando en este campo (48,49).

Una vez establecida la enfermedad, el control de la leishmaniasis recae en un diagnóstico y tratamiento adecuado de los pacientes infectados. Actualmente, el tratamiento de la leishmaniasis comprende cinco clases de fármacos: antimoniales pentavalentes, anfotericina B, pentamidina, paramomicina y miltefosina. A pesar de que se dispone de varias alternativas terapéuticas, ninguna de ellas es considerada ideal ya que en su mayoría producen una alta toxicidad, tienen un precio elevado y precisan tratamientos largos que, en muchos casos, son de administración parenteral, lo que dificulta la adherencia a estas terapias. Además, con mucha frecuencia se detecta falta de respuesta terapéutica a los tratamientos debido principalmente al desarrollo de parásitos resistentes a los mismos, lo que hace aún más complicado elegir la terapia más adecuada para tratar la enfermedad (46). A continuación, describiremos brevemente los tratamientos que existen actualmente para la leishmaniasis:

- Antimoniales pentavalentes

El estibogluconato sódico (Pentostam®) y el antimoniato de meglumina (Glucantime®) son los antimoniales pentavalentes recomendados como tratamiento de primera línea por la OMS para la leishmaniasis. A pesar de que se empezaron a utilizar hace más de 60 años, su mecanismo de acción exacto es aún desconocido, aunque probablemente dependa de la reducción del antimonio pentavalente (Sb^V) a antimonio trivalente (Sb^{III}). Se ha demostrado que estos compuestos alteran el metabolismo energético a través de la inhibición de la glucólisis y la β -oxidación de los ácidos grasos, además de producir una acumulación de las formas oxidadas de tripanotión y glutatión, alterando con ello el equilibrio redox de los parásitos. La acción de los antimoniales pentavalentes produce un proceso de muerte celular similar a apoptosis en estos organismos, involucrando la fragmentación de su ADN y la exposición de PS en la cara extracelular de su membrana (50,51). En cuanto a sus desventajas,

cabe destacar que los regímenes terapéuticos son largos y de administración parenteral, lo que lleva a poca adherencia al tratamiento por parte de los pacientes. Así mismo, presenta efectos secundarios como mialgias, pancreatitis y arritmias cardíacas que en ocasiones hacen que se tenga que detener su utilización. Por último, cabe señalar que durante los últimos años han aparecido casos de parásitos resistentes a los antimoniales pentavalentes, lo cual ha limitado notablemente su uso (52).

- Anfotericina B

La anfotericina B, comercializada como Fungizone, es un compuesto antifúngico indicado para como segunda línea de tratamiento de la LV y LMC que no responde a antimoniales pentavalentes. Su actividad leishmanicida se achaca a su interacción con el ergosterol, esteroide principal de la membrana lipídica de *Leishmania*, provocando la formación de poros que aumentan la permeabilidad a iones y conducen a la lisis celular. Los efectos adversos de la anfotericina B incluyen la nefrotoxicidad severa, la hipocalcemia y la fiebre (53).

Existe también una formulación liposomal de la anfotericina B comercializada bajo el nombre de Ambisome. Esta alternativa presenta menos efectos adversos y mayor eficacia, pero su elevado coste hace que su uso se vea limitado en países en vías de desarrollo (54).

- Pentamidina

La pentamidina es una diamina aromática sintética desarrollada como fármaco hipoglucemiante, aunque su utilidad como antiparasitario fue rápidamente descubierta. Hace más de 50 años que se emplea como segunda línea de tratamiento para la LC, LCD y casos de LV que no responden a los antimoniales pentavalentes. Este fármaco actúa inhibiendo el metabolismo de poliaminas, altera el potencial de membrana mitocondrial y se une al ADN dificultando su replicación. Entre sus efectos adversos se encuentran hipotensión, hipoglucemia, diabetes, miocarditis y nefrotoxicidad, por lo que su utilización no está muy extendida y se ha ido abandonando progresivamente como segunda línea de tratamiento (55).

- Paramomicina

La paramomicina es un aminoglucósido de amplio espectro usado para el tratamiento de la LV en formulación parenteral y de la LC tanto en administración tópica como parenteral. Este

compuesto actúa inhibiendo la síntesis de proteínas mediante su unión a la subunidad 30S ribosomal, lo que lleva a la acumulación de un complejo de iniciación anormal en el codón de inicio del ARNm. Sus principales efectos adversos son la nefrotoxicidad severa, hepatotoxicidad y ototoxicidad (56).

- Miltefosina

La miltefosina (hexadecilfosfocolina) es un alquilfosfolípido sintético que se desarrolló inicialmente como un agente antitumoral para el tratamiento del cáncer de mama y los linfomas cutáneos, aunque su actividad leishmanicida se conoce desde finales de los 80. Respecto a su mecanismo de acción, la miltefosina produce alteraciones en la biosíntesis de fosfatidilcolina y altera el potencial de membrana mitocondrial a través de la inhibición de la citocromo c oxidasa, induciendo una muerte celular tipo apoptosis en los parásitos (57).

A pesar de que la administración oral de la miltefosina mejora la adherencia al tratamiento de los pacientes, este fármaco presenta efectos adversos gastrointestinales, principalmente vómitos y diarrea, y puede producir hepatotoxicidad y nefrotoxicidad. Además, es un compuesto teratogénico, lo cual restringe su uso en mujeres embarazadas y en edad fértil. Por último, la miltefosina tiene una vida media larga, lo que potencia el desarrollo de parásitos resistentes a este tratamiento (58).

Debido a la creciente incidencia de falta de respuesta terapéutica a los tratamientos en monoterapia y a la emergencia de parásitos resistentes a los tratamientos, la OMS ha recomendado el uso de terapias combinadas que empleen compuestos con actividad sinérgica o aditiva que actúen sobre diferentes dianas. El objetivo perseguido con estas terapias combinadas es conseguir reducir las dosis, la duración de los tratamientos, los efectos adversos y prevenir la aparición de resistencias (59).

Además de las distintas opciones quimioterapéuticas descritas anteriormente, se están empleando también terapias físicas para el tratamiento de la LC en aquellos casos en los que la localización, tamaño y número de lesiones cutáneas lo permite y/o el tratamiento sistémico está contraindicado. Entre estas terapias podemos encontrar la termoterapia y la crioterapia, basada en la destrucción de los parásitos por la aplicación directa sobre las lesiones de calor o frío, respectivamente; el láser de CO₂, que produce la termólisis específica del tejido infectado; o la terapia fotodinámica, que a través del uso de agentes fotosensibilizadores

consigue generar especies reactivas de oxígeno en las células tras la irradiación de las lesiones con luz visible de baja intensidad, promoviendo la muerte celular. Estas terapias pueden producir algunos efectos secundarios menores como la hiper e hipopigmentación, la aparición de infecciones secundarias, eritema, cicatrización hipertrófica, ardor, prurito, enrojecimiento, edema y dolor (60).

1.2 Mecanismos de muerte celular

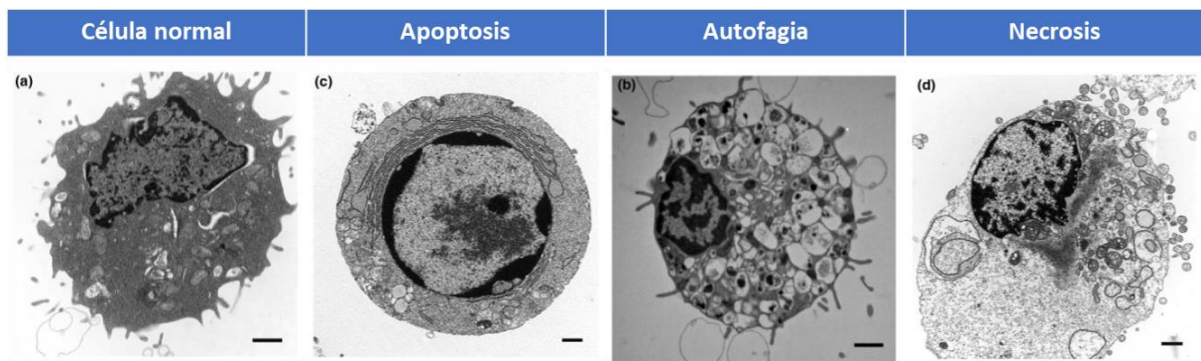
Durante las últimas décadas, se han realizado innumerables trabajos de investigación encaminados a conocer mejor los mecanismos celulares que controlan los procesos de muerte celular regulada (MCR) que se encargan de la eliminación selectiva de células superfluas, irreversiblemente dañadas y/o potencialmente dañinas (61). Curiosamente, estos procesos de MCR no son exclusivos de las formas de vida multicelulares, donde juegan un papel fundamental para ayudar al organismo a mantener su homeostasis en entornos fisiológicos y patológicos, sino que también se han visto variantes simplificadas de la misma en organismos eucariotas unicelulares como levaduras y protozoos, tema que abordaremos más adelante en esta tesis (62). En marcado contraste con la muerte celular accidental (MCA), que supone la desaparición instantánea y catastrófica de las células expuestas a graves daños físicos (como altas presiones, temperaturas o fuerzas osmóticas), químicos (como variaciones extremas de pH) o mecánicos (como fuerzas de corte), la MCR emplea una maquinaria molecular dedicada, regulada y específica, lo que implica que puede ser modulada (es decir, retrasada o acelerada) mediante intervenciones farmacológicas o genéticas (61).

Aunque, como veremos a continuación, existen distintos mecanismos de MCR, de forma general se pueden diferenciar dos escenarios diametralmente opuestos. Por un lado, existen mecanismos de MCR que forman parte de procesos puramente fisiológicos como pueden ser el desarrollo o la renovación de los tejidos, es decir, no están relacionados con perturbaciones en la homeostasis celular y, por tanto, no ocurren en un contexto de falta de adaptación a una situación de estrés. Estos procesos de MCR puramente fisiológicos son los que se suelen conocer como procesos de muerte celular programada (MCP). Por otro lado, la MCR puede iniciarse como respuesta a perturbaciones en el ambiente intracelular o extracelular que son demasiado intensas o prolongadas como para que la célula pueda recuperarse y restaurar su

homeostasis, por lo que terminan produciendo la muerte celular. No obstante, es importante destacar que la MCR impulsada por estrés también constituye una estrategia para la preservación del equilibrio biológico ya que, además de producir la eliminación de células inútiles o potencialmente moribundas, da la capacidad a estas células de exponer o liberar moléculas que alertan al organismo sobre una amenaza potencial.

La muerte celular exhibe además manifestaciones morfológicas características. Dichos morfotipos, junto con los mecanismos por los cuales se eliminan las células muertas y sus fragmentos, han sido empleados históricamente para clasificar los distintos tipos de muerte celular en tres clases (**Figura 5**) (63):

- Muerte celular tipo I o apoptosis, la cual cursa con redondeamiento de la célula, encogimiento citoplasmático, condensación de la cromatina (picnosis), retracción de los pseudópodos, fragmentación nuclear (cariorexis), pequeña o inexistente modificación de los orgánulos citoplasmáticos, formación de cuerpos apoptóticos y mantenimiento de la membrana plasmática hasta los últimos estadios del proceso. Estos cuerpos apoptóticos son absorbidos por células fagocíticas vecinas y degradados en el interior de los lisosomas.
- Muerte celular tipo II o autofagia, que ocurre en ausencia de condensación de la cromatina y se manifiesta con una extensa vacuolización citoplasmática. Este tipo de muerte celular culmina de forma similar a la apoptosis con la captación fagocítica y posterior degradación lisosomal.
- Muerte celular tipo III o necrosis, que no muestra ninguna de las características distintivas de la muerte celular tipo I o II y se manifiesta por aumento del tamaño citoplasmático y de los orgánulos celulares, ruptura de la membrana plasmática y una moderada condensación de la cromatina. Este tipo de muerte celular culmina con la eliminación de los restos celulares sin intervención fagocítica y lisosomal aparente.



Características morfológicas	Apoptosis (muerte celular tipo I)	Autofagia (muerte celular tipo II)	Necrosis (muerte celular tipo III)
Tamaño citoplasmático	Encogido	-	Aumentado
Condensación de la cromatina	++	+	-
Vacuolización (autofagosomas)	-	++	+
Externalización de fosfatidilserina	++	+	-

Figura 5. Características de la muerte celular apoptótica, autofágica y necrótica. En la figura se han incluido la microscopía electrónica de una célula normal, apoptótica, autofágica y necrótica, así como un breve resumen de sus características morfológicas principales. Imagen adaptada de (64).

A pesar de que es importante conocer esta clasificación morfológica, ya que todavía es ampliamente utilizada, se sabe que está muy limitada ya que existen mecanismos bioquímicos muy dispares que pueden conducir a muertes celulares morfológicamente similares. Por este motivo, los editores de la revista *Cell Death and Differentiation* (revista del grupo *Nature* especializada en muerte celular) decidieron crear en el año 2005 el Comité de la Nomenclatura de Muerte Celular (NCCD, de sus siglas en inglés *Nomenclature Committee on Cell Death*) con el objetivo de: 1) abordar los problemas relacionados con el uso de una nomenclatura de muerte celular basada únicamente en características morfológicas; 2) unificar y definir con precisión las principales modalidades de muerte celular sobre una base genética, bioquímica, farmacológica y funcional (en lugar de morfológica) y mantener estas definiciones actualizadas; 3) distinguir los aspectos causales esenciales del proceso de muerte; y 4) diferenciar entre “células muertas”, definidas como aquellas que presentan una permeabilización irreversible de su membrana plasmática o fragmentación celular completa, de células que se están muriendo como “proceso” (61,65).

Puesto que este campo continúa creciendo y las vías de señalización que regulan la MCR todavía se están caracterizando, el NCCD ha ido actualizando progresivamente sus guías,

siendo la última la publicada en 2018. En esta última actualización, el NCCD ha diferenciado un total de 12 mecanismos diferentes de MCR que describiremos brevemente a continuación (61), si bien más adelante en esta introducción se describirán con más detalle los dos procesos apoptóticos por ser de especial relevancia para esta tesis.

a) Apoptosis intrínseca

La apoptosis intrínseca es un tipo de MCR iniciada por perturbaciones en el ambiente celular incluyendo, entre otros, la desaparición de factores de crecimiento, daño en el ADN, estrés en el retículo endoplásmico, acumulación de especies reactivas de oxígeno (ROS), fallos en la replicación, alteraciones en los microtúbulos o defectos mitóticos. Las células apoptóticas mantienen su integridad plasmática y actividad metabólica (hasta cierto punto) durante el proceso de muerte celular, lo cual permite su rápida limpieza por los macrófagos u otras células fagocíticas mediante un proceso conocido como esferocitosis.

El punto de no retorno de la apoptosis intrínseca es la permeabilización de la membrana externa mitocondrial (PMEM), la cual está controlada por factores pro-apoptóticos y anti-apoptóticos de la familia de proteínas reguladoras de la apoptosis Bcl-2, y es finalmente ejecutada por proteínas de la familia de las caspasas, principalmente CASP3 (61).

b) Apoptosis extrínseca

La apoptosis extrínseca es una variante de la MCR iniciada por perturbaciones en el microambiente extracelular que son detectadas por receptores de membrana específicos, de los cuales se diferencian dos tipos; 1) receptores de muerte, cuya activación depende de la unión de sus ligandos específicos; y 2) receptores de dependencia, cuya activación tiene lugar cuando los niveles de sus ligandos específicos disminuyen por debajo de una cantidad límite. La unión de ligandos específicos a estos receptores promueve la activación de proteínas tipo caspasa, jugando un papel especialmente relevante en la propagación de la señal la caspasa CASP8 y en la ejecución de la apoptosis CASP3 (61).

c) Necrosis dirigida por la permeabilidad mitocondrial transitoria

Este tipo de muerte celular se inicia por perturbaciones específicas del microambiente intracelular como estrés oxidativo grave y sobrecarga de Ca^{2+} citosólico. Estos estímulos producen una permeabilidad mitocondrial transitoria, término empleado para definir la

pérdida abrupta de la impermeabilidad de la membrana interna mitocondrial a pequeños solutos, lo cual resulta en una rápida disipación del potencial de membrana mitocondrial ($\Delta\psi_m$), la pérdida del balance osmótico entre la membrana interna y externa de este orgánulo y, finalmente, la muerte celular. Este tipo de MCR normalmente se manifiesta con un fenotipo necrótico.

A nivel bioquímico, se ha propuesto que este tipo de muerte celular depende de la apertura de un poro no específico denominado “poro de permeabilidad mitocondrial transitoria”, complejo supramolecular formado entre las membranas mitocondriales interna y externa. No obstante, la composición, regulación y mecanismo de acción de este complejo está todavía bajo investigación (61).

d) Necroptosis

La necroptosis, también conocida como “necrosis programada”, es un tipo de MCR que se inicia por perturbaciones del microambiente intra o extracelular que son detectadas por receptores de muerte específicos o receptores de reconocimiento de patógenos (RRPs) y que se manifiesta generalmente con un fenotipo necrótico (61).

e) Ferroptosis

La ferroptosis es una forma de MCR dependiente de hierro que se inicia por perturbaciones en el balance oxidativo intracelular que conducen a la formación de especies reactivas de oxígeno (ROS) y la formación de lípidos peroxidados. La ferroptosis produce un morfotipo necrótico con alteraciones mitocondriales significativas que incluyen su encogimiento, la reducción o desaparición de las crestas mitocondriales y la ruptura de la membrana externa mitocondrial (61).

f) Piroptosis

La piroptosis se desencadena por perturbaciones de la homeostasis extra o intracelular relacionadas con la inmunidad innata (por ejemplo, por la invasión de patógenos) y se manifiesta con características morfológicas específicas que incluyen una forma particular de condensación de la cromatina (diferente a la observada en apoptosis), así como la inflamación celular que culmina con la permeabilización de la membrana plasmática. El término piroptosis fue inicialmente acuñado para definir un tipo particular de MCR similar a la apoptosis pero

que dependía de la caspasa CASP1, vinculada a la aparición de pirexia (aumento de temperatura). Inicialmente, se pensaba que la piroptosis solo participaba en la desaparición de monocitos o macrófagos sometidos a la activación de CASP1, sin embargo, hallazgos recientes indican que también puede ser impulsada por otras caspasas como CASP3, puede ocurrir en células no monocíticas, puede tener un papel importante en la inmunidad innata contra patógenos intracelulares y puede estar involucrada en afecciones patológicas como el shock séptico letal (61).

g) Partanatos

El partanatos es una forma de MCR producida por la hiperactivación de un componente específico de la maquinaria de respuesta de daño en el ADN llamada poli (ADP-ribosa) polimerasa 1 (PARP1). Curiosamente, este tipo de MCR parece no ocurrir únicamente como respuesta a un daño grave o prolongado en el ADN, sino que también puede producirse como respuesta al estrés oxidativo, hipoxia, hipoglucemia o señales inflamatorias.

La hiperactivación de PARP1 causa daños citotóxicos mediante dos vías 1) la depleción de NAD^+ y ATP, causando un desequilibrio en la homeostasis redox y bioenergética; y 2) la acumulación de polímeros de poli-ADP-ribosa y proteínas poli-ADP-ribosiladas en la mitocondria, que conducen a la disipación del $\Delta\psi_m$ y la permeabilización de la membrana externa mitocondrial.

Uno de los procesos clave en el partanatos es la unión de los polímeros de poli-ADP-ribosa al factor inductor de apoptosis mitocondrial AIF (de sus siglas en inglés para *Apoptosis Inducing Factor*), induciendo su liberación al citosol y posterior traslocación al núcleo, donde mediará en la fragmentación del ADN y la condensación de la cromatina (61).

h) Muerte celular entótica

La muerte celular entótica es una forma de “canibalismo celular” que ocurre tanto en tejidos sanos como malignos y que implica la absorción de células viables por células no fagocíticas. A menudo, aunque no siempre, las células internalizadas (conocidas como células entóticas) desaparecen tras ser internalizadas. La entosis se desencadena principalmente por el desprendimiento de células epiteliales de su matriz extracelular y la consecuente pérdida de señalización por integrinas, aunque también se han descrito mecanismos alternativos.

El modelo actual para este tipo de MCR propone que la internalización de las células entóxicas se produce por invasión celular en lugar de por fagocitosis. En consecuencia, la captación de estas células es un proceso independiente de integrinas que está promovido por la formación de uniones entre la célula envolvente y la célula entóxica a través de cadherinas y cateninas (61).

i) Muerte celular netóxica

El término muerte celular netóxica describe un tipo de MCR inicialmente caracterizado en neutrófilos y que estaba vinculada a la extrusión de una malla de fibras denominadas trampas celulares de neutrófilos o NETs (*Neutrophil Extracellular Traps*). Estas NETs, que son producidas como respuesta a activadores microbianos o por la estimulación de receptores específicos, constituyen una plataforma extracelular estable para atrapar y degradar los microbios. A pesar de su nombre, se ha visto que otras células como mastocitos, eosinófilos o basófilos son capaces de producir estructuras similares a los NETs. Por tanto, la muerte celular netóxica se ha definido como una modalidad de MCR restringida a células originadas a partir de células hematopoyéticas y asociada con la extrusión de trampas celulares de neutrófilos. Se ha propuesto que la muerte celular netóxica transcurre a través de la activación de la NADPH oxidasa y la consiguiente generación de ROS en el interior de la célula (61).

j) Muerte celular dependiente de lisosomas

La muerte celular dependiente de lisosomas es una subrutina de MCR iniciada por perturbaciones en la homeostasis intracelular y demarcada por la permeabilización de las membranas lisosomales. Este tipo de MCR se ha visto asociada con el desarrollo de varias afecciones fisiopatológicas que incluyen la inflamación, la remodelación de tejidos (por ejemplo, la involución de la glándula mamaria después de la lactancia), el envejecimiento, la neurodegeneración, ciertos trastornos cardiovasculares y la respuesta a patógenos intracelulares.

A nivel bioquímico, la muerte celular dependiente de lisosomas se desencadena tras la permeabilización de la membrana lisosomal y la liberación de su contenido (incluyendo enzimas de la familia de las catepsinas) al citosol (61). No obstante, el mecanismo por el cual se produce esta permeabilización es todavía desconocido.

k) Muerte celular por autofagia

La autofagia es un proceso fisiológico por el cual componentes celulares, como macromoléculas u organelas completas, son secuestrados en el interior de una vacuola de doble membrana, los autofagosomas, para su posterior degradación en los lisosomas. La fusión de los autofagosomas con los lisosomas da lugar a la formación de los autolisosomas, en los cuales el contenido del autofagosoma va a ser degradado por hidrolasas lisosomales para la recuperación y reutilización de sus componentes, es decir, es una especie de mecanismo de “reciclaje celular” (**Figura 6**). Por tanto, el proceso de autofagia puede ocurrir sin necesidad de que tenga lugar la muerte de la célula (66). Es importante distinguir entre la autofagia “como proceso” y la muerte celular por autofagia, siendo esta definida por el NCCD como una subrutina de MCR cuyo mecanismo depende de la maquinaria autofágica o componentes de esta. Para evitar confusiones, este término debe evitarse en ausencia de evidencia experimental que relacione la muerte de la célula con componentes de la maquinaria autofágica (idealmente, con más de uno) y cuando las manipulaciones farmacológicas o genéticas de la maquinaria autofágica impacten sobre una subrutina de MCR diferente a esta (61).

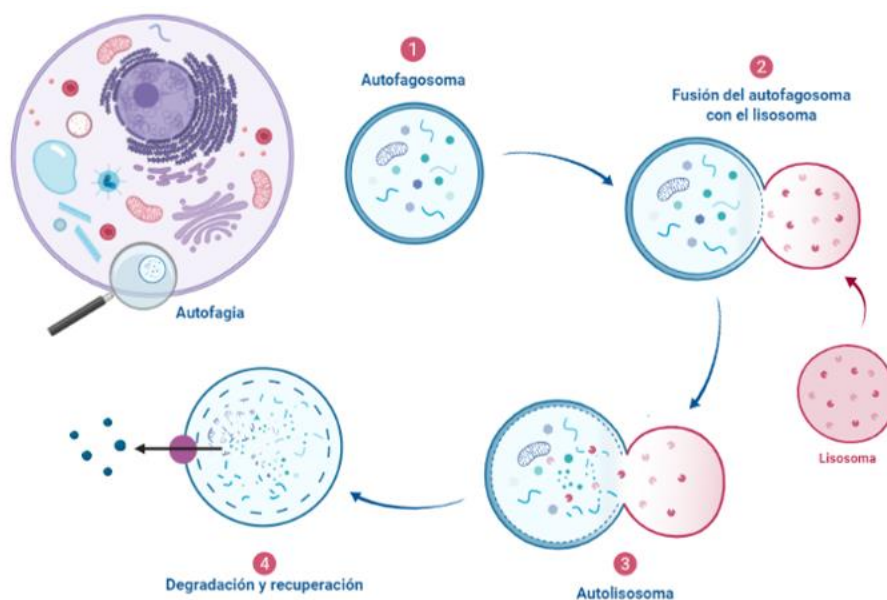


Figura 6. Representación esquemática del proceso de autofagia. Imagen creada en www.biorender.com

l) Muerte celular inmunogénica

La muerte celular inmunogénica es un tipo de MCR con una función peculiar, ya que es capaz de activar la respuesta inmune adaptativa específica frente a antígenos endógenos (celulares) o exógenos (virales) expresados por las células que se están muriendo. Este tipo de muerte celular puede iniciarse por un conjunto relativamente limitado de estímulos, incluyendo las infecciones virales, algunos quimioterapéuticos, formas específicas de radioterapia y por la terapia fotodinámica basada en hipericina. Estos estímulos producen la liberación de una serie “patrones moleculares asociados a daño” o DAMPs (*Damage-Associated Molecular Patterns*), cuyo reconocimiento por los RRP de las células del sistema inmune advierte al organismo de una situación de peligro, resultando así en la activación de una respuesta inmune generalmente asociada con el establecimiento de la memoria inmunológica. Por todo ello, el NCCD ha definido la muerte celular inmunogénica como un tipo de MCR que es suficiente para activar una respuesta inmune adaptativa en hospedadores inmunocompetentes.

Finalmente, es importante destacar que, a pesar de que todas estas subrutinas presentan características genéticas, bioquímicas, farmacológicas y funcionales específicas que las hacen diferentes a unas de otras, existe cierto grado de similitud e interconexión entre ellas.

1.2.1 Apoptosis en organismos pluricelulares

El término “apoptosis” fue acuñado por Kerr, Wyllie y Currie en 1972, siendo además los primeros en describir algunos de los cambios morfológicos característicos de este tipo de muerte celular (67). No obstante, no fue hasta 1999 cuando Robert Horvitz y su equipo demostraron la existencia de genes y proteínas específicas que controlaban la muerte celular usando como modelo el nemátodo *Caenorhabditis elegans* (68,69), trabajos que le llevaron a ganar el premio Nobel de fisiología junto a Sydney Brenner y John E. Sulston en 2002. En el desarrollo del nemátodo *C. elegans* adulto hermafrodita se generan 1090 células somáticas, de las cuales 131 sufren apoptosis de forma invariable, es decir, se trata de un fenómeno controlado. Para estudiar cómo tenía lugar este proceso, Horvitz y su equipo generaron distintos mutantes genéticos con los que consiguieron demostrar que el desarrollo de este nemátodo dependía de la presencia y expresión de tan solo cuatro genes: *ced-3*, *ced-4*, *ced-9* y *egl-1* (68,69). CED-3 es una proteína tipo caspasa, es decir, una cisteín-proteasa con

especificidad por residuos de aspartato que es sintetizada como un zimógeno que necesita ser procesado para poder actuar como ejecutor de la muerte celular. CED-4 actúa como proteína adaptadora uniéndose a CED-3 y promoviendo su auto-procesamiento y activación. Por su parte, CED-9 actúa como represor de la muerte celular, ya que al unirse a CED-4 impide la activación de CED-3. Por último, EGL-1 es un antagonista de CED-9, al cual se une para evitar su efecto represor y permitir la activación de CED-3 (**Figura 7**) (68,69).

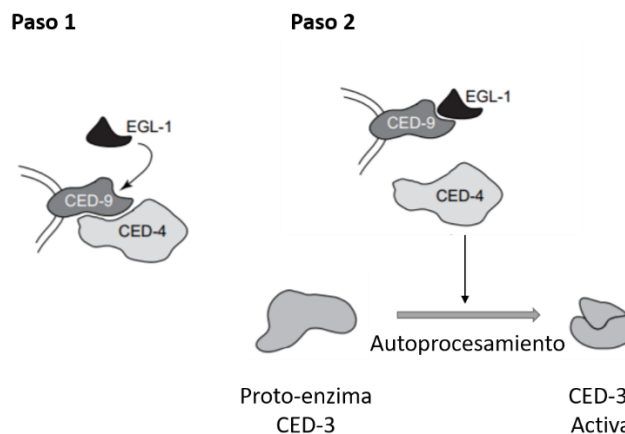


Figura 7. Modelo de activación de la muerte celular en *C. elegans*. Paso 1: La activación de EGL-1 hace que se una a la proteína CED-9 y produce así la liberación de CED-4 de su represor. Paso 2: CED-4 es capaz de promover la proteólisis de la proto-enzima CED-3, haciendo que se genere una CED-3 activa que actuará como caspasa ejecutora del proceso de muerte celular. Figura adaptada de (73).

Tras la publicación de estos hallazgos, se realizaron numerosos estudios que llevaron a la identificación de proteínas homólogas a las descubiertas en *C. elegans* en todos los organismos vertebrados, así como en la mosca de la fruta *D. melanogaster*, que rápidamente se convirtió en el segundo modelo favorito para estudiar la apoptosis (70). Adicionalmente, los estudios realizados en células humanas y de ratón mostraron que las vías de señalización implicadas en la regulación de la apoptosis de estos organismos son mucho más vastas y complejas que las implicadas en la apoptosis de *C. elegans*. En estos organismos existen alrededor de veinte homólogos del represor CED-9 y su antagonista EGL-1, los cuales se engloban dentro de la familia Bcl-2. Dentro de esta familia se diferencia entre proteínas represoras de muerte, (Bcl-2, Bcl_{X_L}, Bcl_W, MCL1 y BFL1/A1), proteínas pro-apoptóticas (Bax y Bak) y proteínas con un solo dominio BH3 o *BH3-only* (Bad, Bim, Noxa y Puma), que actúan

activando efectores pro-apoptóticos y/o neutralizando los miembros anti-apoptóticos de la familia Bcl-2. En ambos casos, ya sea de forma directa o indirecta, las proteínas *BH3-only* tienen un papel pro-apoptótico.

De igual forma, existen catorce homólogos de la proteína ejecutora CED-3, los cuales se engloban dentro de la familia de las caspasas. En esta familia podemos diferenciar entre caspasas iniciadoras (caspasas 2, 8, 9 y 10), caspasas ejecutoras (caspasas 3, 6, 7 y 14) y caspasas activadoras de citoquinas (caspasas 1, 4, 5, 11, 12 y 13).

Por último, se ha identificado al menos un homólogo de CED-4 conocido como proteasa apoptótica activadora del factor 1 o Apaf-1 (71,72).

Como vimos anteriormente en esta introducción, el NDDC diferencia dos mecanismos o vías de señalización apoptóticos diferentes: la vía intrínseca, cuando la apoptosis por perturbaciones en el ambiente intracelular; o la vía extrínseca, donde la apoptosis se inicia tras la detección de señales de muerte extracelulares (61). A continuación, describiremos ambas vías en detalle. Como veremos, a pesar de que son diferentes, las dos vías están ligadas y las moléculas de una u otra pueden influenciarse mutuamente (**Figura 8**).

- Vía intrínseca

La vía intrínseca, también conocida como vía mitocondrial, se inicia por una serie de estímulos que producen señales intracelulares que terminan repercutiendo en la mitocondria. Estos estímulos pueden ser de naturaleza negativa, como la ausencia de factores de crecimiento, hormonas o citoquinas; o positiva, como la exposición a radiación, hipoxia, toxinas, radicales libres, hipertermia o infecciones virales. Adicionalmente, el daño en el ADN puede inducir la apoptosis por esta vía, proceso que suele estar mediado por la proteína p53 (72).

Cuando ciertos orgánulos o compartimentos celulares experimentan perturbaciones en su homeostasis como consecuencia de los estímulos anteriormente descritos, se produce la activación de los miembros *BH3-only* de la familia Bcl-2 (**Figura 8** – derecha). Una vez activadas, estas proteínas *BH3-only* tienen la capacidad de interactuar físicamente con las proteínas pro-apoptóticas Bax y Bak presentes en la membrana externa mitocondrial (MEM), causando su oligomerización. Esta oligomerización produce la formación de un poro toroidal lipídico y va a desencadenar la permeabilización de la membrana externa mitocondrial

(PMEM), paso crítico en la apoptosis intrínseca. Cabe señalar que, en ausencia de estímulos apoptóticos, Bax y Bak se encuentran inhibidos por los miembros anti-apoptóticos de la familia Bcl-2. Esta inhibición puede llevarse a cabo por una unión directa entre ambos, que previene su oligomerización y la formación del poro mitocondrial; o de forma indirecta, mediante el secuestro de los activadores del grupo *BH3-only* por parte de los miembros anti-apoptóticos. Por lo tanto, podemos decir que la PMEM, paso crítico para la apoptosis intrínseca, está regulada por el equilibrio entre los distintos miembros de la familia Bcl-2 (61).

La PMEM produce la pérdida del potencial de membrana mitocondrial ($\Delta\Psi_m$) y la liberación de dos grupos de proteínas pro-apoptóticas desde el espacio intermembrana al citosol. El primer grupo está constituido por el citocromo C, la proteína SMAC/DIABLO y la serín-proteasa HtrA2/Omi, las cuales van a causar la activación de la vía mitocondrial dependiente de caspasas (72). Tras liberarse al citosol, el citocromo C se va a unir a Apaf-1 y a la pro-caspasa 9 formando un complejo conocido como apoptosoma. La formación del apoptosoma va a promover la activación autocatalítica de la pro-caspasa 9, dando lugar a la caspasa 9 (CASP9). Esta CASP9 activa, a su vez, va a catalizar la activación proteolítica de CASP3 y CASP7, caspasas ejecutoras responsables del desmantelamiento celular durante la apoptosis tanto por su vía intrínseca como extrínseca, convergiendo ambas vías en este punto. La actividad catalítica de las caspasas ejecutoras es responsable de muchas de las características morfológicas y bioquímicas asociadas a la apoptosis, incluyendo la fragmentación del ADN, la exposición de PS y la formación de cuerpos apoptóticos. Por ejemplo, CASP3 promueve la fragmentación del ADN catalizando la inactivación proteolítica de la subunidad alfa del factor de fragmentación del ADN (iCAD), liberando así la actividad catalítica de la ADNasa activada por caspasas CAD (caspase-activated DNase). Por su parte, SMAC/DIABLO y HtrA2/Omi también promueven la activación de las caspasas mediante un mecanismo indirecto, ya que son capaces de secuestrar y/o degradar las proteínas inhibidoras de apoptosis o IAPs (*inhibitor of apoptosis*), que cuando se encuentran libres actúan inhibiendo a las caspasas.

Por último, tras la PMEM también se libera un segundo grupo de proteínas pro-apoptóticas incluyendo la Endonucleasa G (EndoG) y AIF, las cuales actúan en estadios más tardíos de la apoptosis. Tras su liberación, estas proteínas mitocondriales se translocan al núcleo donde van a participar en la fragmentación ADN y la condensación de la cromatina junto con CAD. Cabe señalar que CAD interviene exclusivamente en procesos dependientes de caspasas,

mientras que EndoG y AIF pueden participar también en procesos de MCR independientes de caspasas (72,74). De la segunda de estas nucleasas hablaremos en detalle más adelante por ser de especial relevancia para esta tesis.

- Vía extrínseca

La apoptosis extrínseca está dirigida por dos tipos de receptores de membrana: 1) receptores de muerte, cuya activación depende de la unión de su ligando específico; y 2) receptores dependientes, cuya activación tiene lugar cuando los niveles de sus ligandos específicos caen por debajo de un valor límite (61).

Los receptores de muerte más estudiados son los miembros de la superfamilia de receptores del factor de necrosis tumoral (TNF), los cuales presentan un dominio intracelular de aproximadamente 80 aminoácidos conocido como dominio de muerte o DD, de sus siglas en inglés para "*death domain*". Dentro de esta familia los receptores más estudiados son FAS, TRAILR1 y TRAILR2, cuya unión a sus ligandos específicos FASL y TRAIL, respectivamente, produce la estabilización de los homotrímeros de receptor presentes en la membrana celular. Esta estabilización induce un cambio conformacional en su dominio citoplasmático que permite la asociación de la proteína adaptadora FADD (**Figura 8** – izquierda). La asociación de FADD a través de los DD citosólicos va a promover el reclutamiento de varias proteínas, incluyendo la procaspasa-8 (o la 10, en determinados casos), dando lugar a la formación de un complejo multiproteico conocido como DISC (*Death-Inducing Signaling Complex*). La formación del complejo DISC va a permitir el procesamiento autocatalítico de las pro-caspasas 8 o 10, según sea el caso, dando lugar a la liberación de la correspondiente caspasa activa CASP8 o CASP10 (61,72,75).

En determinados tipos celulares, incluyendo los linfocitos (los cuales se han englobado dentro de las "células tipo I"), CASP8 cataliza directamente la maduración proteolítica de CASP3, desencadenando así la fase ejecutora de la apoptosis intrínseca sin intervención de la mitocondria. En otras células como hepatocitos, células pancreáticas tipo β y la mayoría de las células cancerosas ("células tipo II"), en las que la activación de CASP3 y CASP7 está restringida por la proteína tipo IAP XIAP, CASP8 va a desencadenar la apoptosis mediante la proteólisis de la proteína Bid. Esta proteólisis produce una forma truncada de Bid, tBid, que se transloca hasta la MEM donde va a actuar promoviendo la formación del poro de

membrana mitocondrial mediado por Bax y Bak. La PMEM y la liberación de factores mitocondriales va a producir la activación de la cascada de caspasas iniciada por CASP9 tal y como se describió para la vía intrínseca de la apoptosis (61,72,75).

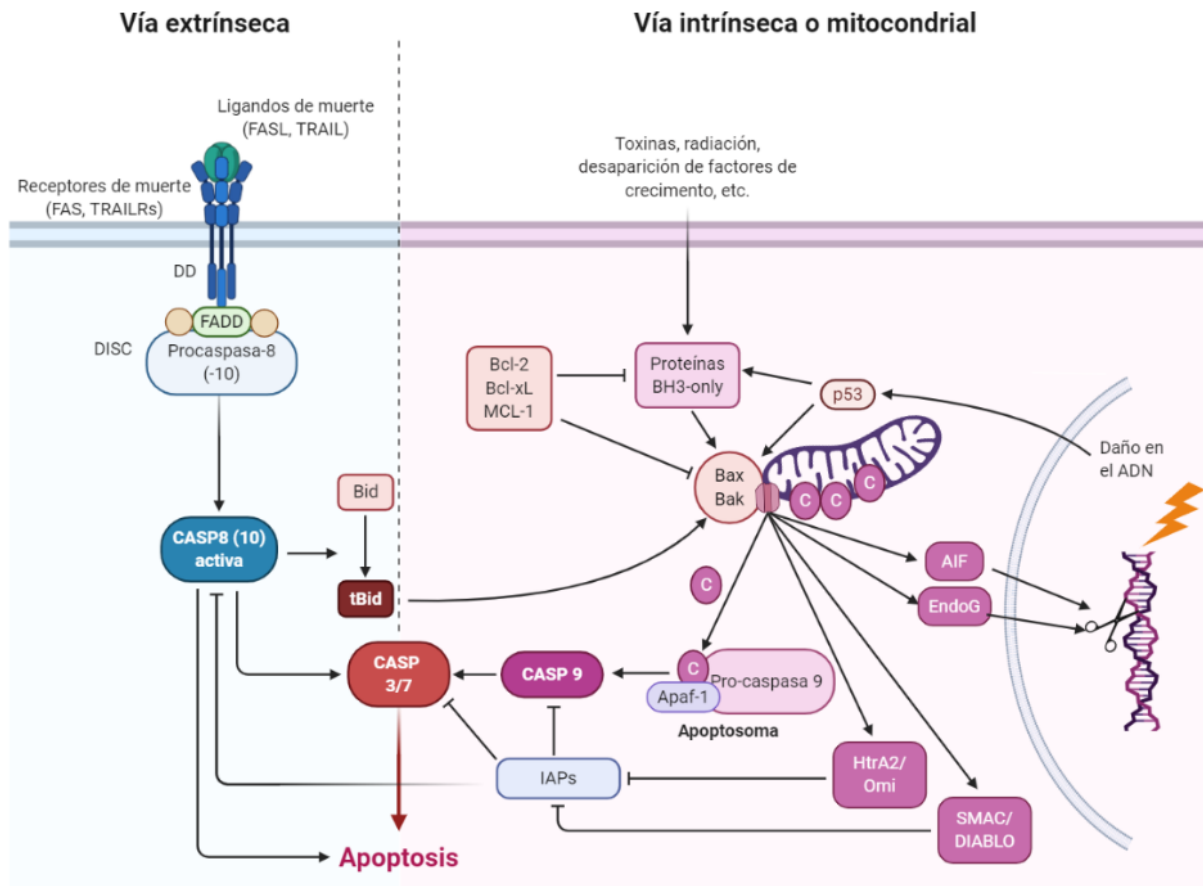


Figura 8. Principales componentes de apoptosis extrínseca e intrínseca, así como las interacciones y convergencia entre ellas. Imagen creada en www.biorender.com.

Por su parte, existen más de veinte tipos de receptores dependientes que pueden activar distintas cascadas de señalización cuando la disponibilidad de su ligando cae por debajo de un nivel limitante, implicando generalmente la activación de caspasas. No obstante, los mecanismos exactos por los que estos ligandos promueven la muerte celular todavía no están bien elucidados (61).

1.2.2 Muerte celular regulada y apoptosis en *Leishmania spp.*

Tal y como hemos descrito en las secciones previas de esta introducción, la existencia de procesos de MCR en organismos multicelulares es fundamental para la eliminación selectiva de células superfluas, irreversiblemente dañadas y/o potencialmente dañinas para el resto del organismo y juega un papel clave en numerosos procesos fisiológicos como el desarrollo, la homeostasis de los tejidos, la inflamación y la inmunidad (74). En este contexto, la “autodestrucción” de una célula individual puede ser entendida como una forma de cooperación en la que se perjudica a un individuo, la célula, para obtener un beneficio a un nivel superior, el del organismo. Desde este punto de vista, podría parecer que la existencia de mecanismos de autodestrucción en organismos unicelulares carecería de sentido, ya que la desaparición de la célula conllevaría también la muerte del propio organismo. Sin embargo, se han descrito procesos de MCR similares a los de metazoos en numerosos organismos eucariotas unicelulares, incluyendo los del género *Leishmania*. Además, los resultados obtenidos sugieren que al menos algunos de los elementos involucrados en estos procesos están evolutivamente relacionados con los presentes en la MCR de los organismos multicelulares (76).

Uno de los factores que apoyan la existencia de procesos de MCR en organismos unicelulares es que, al igual que ocurre en metazoos, estos procesos podrían desempeñar un papel fisiológico en el desarrollo de estos organismos *in vivo*. En concreto, en *Leishmania* se ha descrito que la MCR podría regular el número de individuos de la población en el interior del intestino del insecto vector como respuesta a los recursos limitados que se encuentran en este entorno (77,78). Por ejemplo, se ha descrito que los promastigotes que no se han diferenciado a la forma infectiva dentro del insecto vector pueden sufrir procesos de MCR, dejando de consumir nutrientes y beneficiando con ello a la población de promastigotes metacíclicos infectivos (79). En este sentido, la MCR parece constituir un mecanismo altruista para la selección de los parásitos más válidos para la transmisión de la enfermedad. De forma similar, parece que la MCR podría controlar la densidad parasitaria dentro del hospedador mamífero, evitando así un hiperparasitismo que podría producir su muerte y, por tanto, impedir la transmisión de la enfermedad (80,81). Además, como ya comentamos anteriormente en esta introducción, se ha demostrado que la presencia de parásitos apoptóticos en el inóculo es importante para que la infección concluya con éxito. La

fagocitosis de los parásitos apoptóticos induce la secreción de moléculas antiinflamatorias, como IL-10 o TGF- β , inhibe la secreción de agentes proinflamatorios, como el TNF- α , y disminuye la presentación de antígenos por parte de los macrófagos. De esta forma, se favorece la supervivencia de otros promastigotes no apoptóticos fagocitados por la misma célula. De hecho, se ha demostrado que si se eliminan los parásitos apoptóticos del inóculo los parásitos fagocitados no son capaces de sobrevivir en el interior de la célula fagocítica *in vitro* y pierden su habilidad para inducir la enfermedad *in vivo* (30,82).

Hasta la fecha, se han descrito diversos estímulos y tratamientos con actividad leishmanicida que pueden producir que la muerte celular muestre un fenotipo similar a apoptosis en *Leishmania*. Dentro de los estímulos podemos encontrar algunos que imitan las condiciones que se encuentran los parásitos dentro de sus hospedadores, como el choque térmico o las especies reactivas de oxígeno (NO y H₂O₂) (83–86). Por su parte, entre los tratamientos, podemos encontrar compuestos de naturaleza muy diversa, siendo algunos de los ejemplos más representativos la estaurosporina, cúrcuma, anfotericina B, antimicina A, miltefosina, edelfosina, novobiocina, pentamidina o camptotecina (87–89). Estos estímulos producen una muerte celular en la que se observa redondeamiento y encogimiento celular, cambios en la membrana plasmática, alteraciones en la permeabilidad mitocondrial, condensación de la cromatina y fragmentación del ADN nuclear (88). Es decir, muestran muchas de las características que definen el proceso apoptótico en eucariotas superiores. No obstante, como hemos visto anteriormente, en la actualidad se trata de definir las diferentes subrutinas de MCR en función de los agentes genéticos y bioquímicos implicados, en lugar de basarse únicamente en las características morfológicas expresadas. Lamentablemente, el conocimiento actual de las proteínas implicadas en la muerte celular en *Leishmania* es muy limitado, por lo que la utilización del término apoptosis en estos parásitos debe entenderse en un contexto principalmente morfológico. De hecho, algunas de las principales proteínas implicadas en la apoptosis en eucariotas superiores, como son las caspasas, los receptores de muerte, o los principales miembros de la familia Bcl-2, no han podido ser identificadas en *Leishmania*, ni si quiera mediante la utilización de técnicas bioinformáticas y de secuenciación (87,90). Curiosamente, si se ha descrito que la expresión de la proteína Bcl-2 humana tiene un efecto protector en *L. infantum* tras ser sometido al tratamiento con edelfosina (91), así como que la respuesta celular de estos parásitos a varios estímulos de muerte puede ser

modulada por la presencia de inhibidores de caspasas como zVAD-fmk (84,85,92). A continuación, describiremos algunas de las moléculas cuya participación en la muerte celular de *Leishmania* si ha sido demostrada, implicando la existencia de mecanismos que regulan la muerte celular en estos parásitos.

- Metacaspasas

Las metacaspasas son cisteín proteasas que han sido identificadas en hongos, plantas y protozoos. De hecho, los análisis genómicos realizados han permitido identificar secuencias codificantes para metacaspasas en todas las especies de *Leishmania*. Estas metacaspasas poseen un dominio tipo caspasa o “*caspasa-like*” homólogo al de la subunidad p20 de las caspasas, con la diada catalítica histidina/cisteína; y un dominio C-terminal rico en prolina similar al de la subunidad p10 de las caspasas (93,94). Sin embargo, mientras que las caspasas cortan de forma específica en residuos de aspartato, las metacaspasas muestran especificidad por residuos de arginina o lisina (93,95).

Se ha propuesto que la metacaspasa de *L. major*, LmjMCA, podría desempeñar un papel importante durante la muerte celular en estos parásitos. La sobreexpresión de LmjMCA induce la apoptosis de *L. major* tras ser sometida a choque térmico o estrés oxidativo y tras el tratamiento con miltefosina o cúrcuma (94), mientras que su eliminación inhibe la muerte celular inducida por estos dos tratamientos (96). Adicionalmente, se ha demostrado que LmjMCA podría estar implicada en el proceso de autofagia en estos parásitos, dotándole así de un rol provida no descrito en las caspasas (96).

- Acuaporina Li-BH3AQP

La acuaporina Li-BH3AQP fue identificada en *L. infantum* por nuestro grupo de investigación. Esta proteína contiene un dominio similar a los dominios BH3 que se encuentra en todos los miembros de la familia Bcl-2. Ensayos de co-inmunoprecipitación han demostrado que Li-BH3AQP puede unirse a la proteína anti-apoptótica Bcl-X_L humana. Adicionalmente, su expresión reduce la viabilidad celular en células de mamífero, pudiendo contrarrestarse este efecto mediante su coexpresión con Bcl-X_L. En *L. infantum*, la sobreexpresión de Li-BH3AQP reduce la viabilidad celular e induce la fragmentación del ADN tras el tratamiento con estaurosporina y antimicina A. Esta actividad pro-apoptótica desaparece tras la mutación de dos residuos altamente conservados del dominio BH3, demostrando que este dominio es

clave para poder desempeñar esta función. Sorprendentemente, la sobreexpresión de esta acuaporina también hizo a los parásitos más resistentes al estrés hipotónico o la privación de nutrientes, mostrando así un posible papel proveído para esta proteína (97).

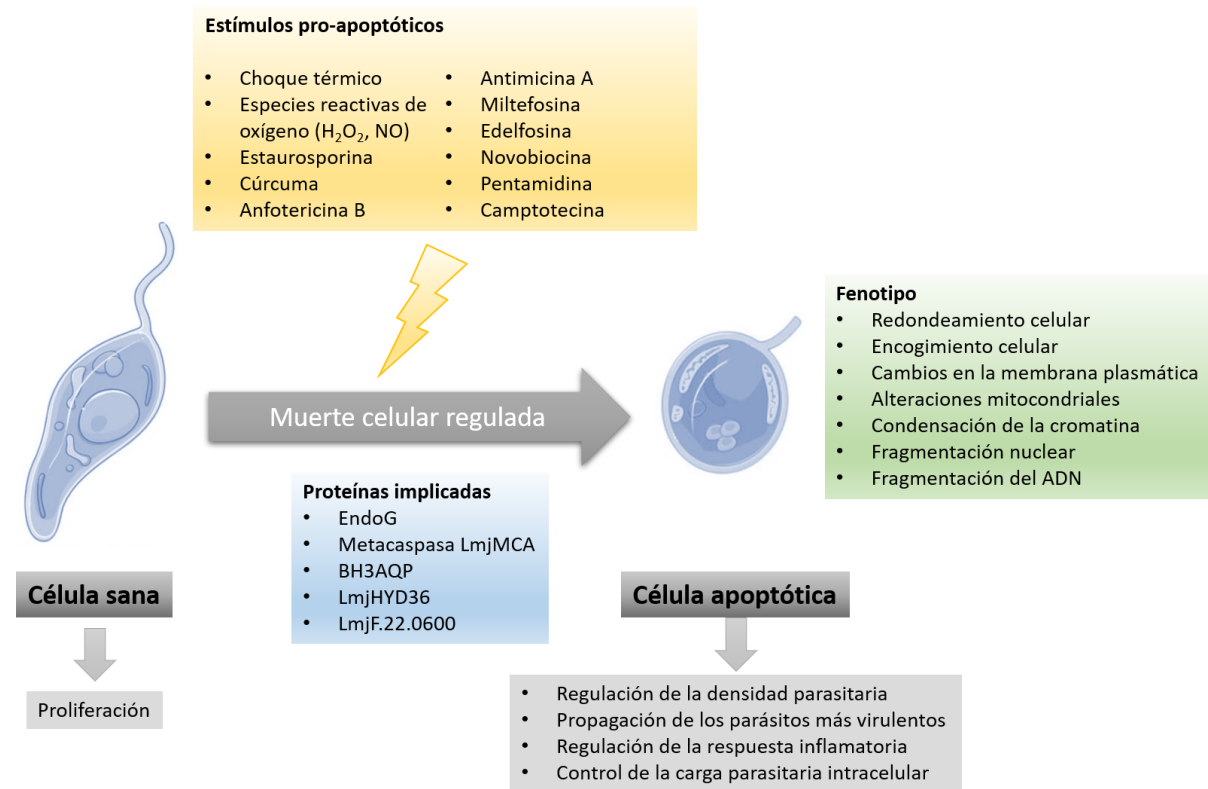


Figura 9. Representación esquemática de los aspectos más importantes implicados en la muerte celular regulada en *Leishmania*. Figura adaptada de (102).

- LmjF.36.6540 y Lmj.22.0666

Recientemente se han identificado dos nuevas proteínas en *L. major* que podrían actuar como efectores de la muerte celular: LmjF.36.6540 y Lmj.22.0666 (98,99).

La proteína LmjF.36.6540 (o LmjHYD36, nombre propuesto por los autores que la han descrito) contiene la triada catalítica cisteína/histidina/aspartato típica de las α/β -hidrolasas y está altamente conservada en *Leishmania spp.* Durante el pasado año, Basmacıyan y su equipo decidieron estudiar la posible implicación de LmjHYD36 en la muerte celular en *L. major*. Sus resultados mostraron que al tratar promastigotes de *L. major* con cúrcuma y miltefosina aumentaba la expresión de la proteína, mientras que el tratamiento con

pentamidina no tenía efecto sobre su expresión. Adicionalmente, vieron que la sobreexpresión del gen *LmjF.36.6540* aumentaba la muerte inducida por cúrcuma y pentamidina, mientras que no tuvo ningún efecto sobre la muerte inducida por miltefosina. En conjunto, estos resultados demostraron que esta proteína está involucrada en la muerte celular inducida por cúrcuma, miltefosina y pentamidina al menos en *L. major* (98). Adicionalmente, este mismo grupo identificó una segunda proteína, *LmjF.22.0666*, que se sobreexpresa cuando los parásitos se diferencian a sus distintos estadios durante su ciclo de vida. *LmjF.22.0666* es una potencial acetiltransferasa conservada en las distintas especies de *Leishmania*. Tras llevar a cabo experimentos similares a los descritos para el caso anterior, vieron que esta proteína se sobreexpresaba tras someter a los parásitos a diferentes estímulos de muerte, así como que el aumento en la expresión del gen aumentaba la muerte celular tras el tratamiento con cúrcuma y miltefosina, sugiriendo que esta proteína podría actuar también como efector en la muerte celular en estos parásitos (99).

- Endonucleasa G

Como comentamos cuando describimos la vía intrínseca de la apoptosis, EndoG está implicada en la degradación del ADN nuclear independiente de la intervención de las caspasas en mamíferos. Se han encontrado miembros de la familia EndoG en todos los organismos cuyo genoma ha sido secuenciado, incluyendo en *Leishmania*, donde se han estudiado y caracterizado especialmente dos especies: *L. donovani* y *L. infantum*. En ambos casos, se ha demostrado que esta enzima posee características similares a las observadas en la EndoG de eucariotas superiores: posee actividad nucleasa *in vitro*, se localiza en la mitocondria en condiciones fisiológicas y migra al núcleo en respuesta a estímulos apoptóticos como H₂O₂ o edelfosina, donde va a participar en la degradación el ADN nuclear, lo que convierte a EndoG en uno de los efectores de la apoptosis en estos parásitos (100,101). La caracterización de la EndoG de *Leishmania infantum* (*LiEndoG*) es el objetivo central de esta tesis, por lo que en el siguiente apartado de esta introducción nos centraremos en describir algunas de las EndoGs más representativas de esta familia de proteínas y en resumir lo que sabemos hasta ahora sobre *LiEndoG*.

En conclusión, el reciente descubrimiento de todos estos potenciales efectores de muerte celular en *Leishmania* parece indicar que la muerte en estos parásitos si está regulada y no es un hecho accidental, como se ha sugerido por varios autores (90,103). No obstante, es

necesario continuar investigando para poder describir con exactitud las vías moleculares que desencadenan estos procesos de MCR. En la **Figura 9** se resumen las implicaciones fisiológicas descritas para la MCR en *Leishmania*, los aspectos fenotípicos apoptóticos más característicos en estos parásitos, algunos de sus estímulos pro-apoptóticos más representativos y las proteínas descritas como efectores en el proceso apoptótico hasta la fecha.

1.3 Endonucleasa G

Los primeros resultados que muestran la existencia de la endonucleasa mitocondrial que ahora conocemos como Endonucleasa G (EndoG) datan de mediados de la década de los 60, si bien en ese momento no le fue otorgado este nombre. De forma casi simultánea, Curtis *et al.* demostraron la presencia de una proteína con actividad nucleasa en los extractos mitocondriales de hígados de rata (104); mientras que, por su parte, Linn y Lehman mostraron que también existía una proteína con actividad similar en los extractos mitocondriales de *Neurospora crassa* (105). Los resultados expuestos por estos autores mostraron que ambas nucleasas tenían propiedades similares, siendo su actividad dependiente de Mg^{2+} o Mn^{2+} y produciendo como resultado de esta la liberación de oligonucleótidos con un grupo fosfato libre en su extremo 5'. Adicionalmente, ambas enzimas presentaban mayor actividad a pHs ligeramente ácidos y eran capaces de hidrolizar extensivamente tanto ARN como ADN de cadena sencilla (cs), mientras que hidrolizaban con menos avidez el ADN de doble cadena (ADNdc) (104–106). No obstante, estudios posteriores mostraron que la nucleasa de *Neurospora* era de hecho una endo-exonucleasa, ya que poseía actividad 5'-exonucleasa además de la actividad endonucleasa ya descrita (107,108). Esta actividad exonucleasa era capaz de crear las regiones de cadena sencilla dentro del ADNdc que son requeridas para que se pueda producir la recombinación homóloga, por lo que se propuso que esta enzima podría estar involucrada en este proceso (109). Adicionalmente, en 1988, Zassenhaus y su equipo consiguieron purificar una endo-exonucleasa mitocondrial de la levadura *Saccharomyces cerevisiae*, denominada Nuc1p, con características muy similares a la de *Neurospora*. De hecho, el anticuerpo policlonal contra la nucleasa de *Neurospora* era capaz de reconocer e inhibir la a Nuc1p y, de igual forma, el anticuerpo contra Nuc1p reconocía la nucleasa de *Neurospora*. Esta reactividad cruzada, junto con la similitud de sus propiedades enzimáticas,

sugería que la actividad endo-exonucleasa mitocondrial podría estar conservada al menos en los eucariotas unicelulares (110,111).

Por otra parte, diversos grupos de investigación describieron la presencia de actividad nucleasa en lisados mitocondriales de distintas células de mamífero, si bien no fueron capaces de purificar la proteína causante de esta actividad. No fue hasta 1986 que Tomkinson y Linn consiguieron purificar por primera vez una nucleasa de las mitocondrias de la línea celular de ratón MCP-11 (112). Esta enzima fue descrita como un homodímero en su forma nativa y, de forma similar a lo observado para las proteínas anteriores, su actividad dependía de Mg^{2+} y degradaba preferentemente sustratos de cadena sencilla, si bien también era capaz de degradar ADNdc. Adicionalmente, esta proteína también se veía afectada por el anticuerpo preparado contra Nuc1p, sugiriendo que podría tratarse del homólogo en mamíferos de las enzimas mitocondriales descritas en eucariotas unicelulares (112). Un año después, se consiguió purificar una nucleasa de extractos mitocondriales de células bovinas con propiedades similares a las nucleasas que hemos descrito hasta ahora; se trataba de un homodímero que requería Mg^{2+} o Mn^{2+} para su actividad y era capaz de degradar tanto ARN como ADN de cadena sencilla o doble. Sin embargo, esta proteína solo presentaba actividad endonucleasa, a diferencia de la actividad endo-exonucleasa descrita en las enzimas de *Neurospora* o *Saccharomyces* (113).

Curiosamente, el término Endonucleasa G (EndoG) no fue acuñado hasta 1987 para una nucleasa que purificaron Ruiz-Carrillo y Renaud en 1987 de eritrocitos de gallina (114). En las condiciones experimentales propuestas por estos autores, esta proteína presentaba una clara preferencia por digerir sitios $(dG)_n \cdot (dC)_n$, con mayor especificidad por la hebra G, motivo por el cual le otorgaron el nombre de Endonucleasa G (EndoG) (114). Posteriormente, Gerschenson *et al.* demostrarían que EndoG y las distintas nucleasas mitocondriales de mamífero descritas hasta la fecha eran de hecho proteínas homólogas, ya que presentaban un alto grado de identidad en sus secuencias y las mismas propiedades bioquímicas (115). A pesar de que los estudios posteriores han mostrado que estas enzimas pueden digerir el ADN de forma aleatoria sin tener apenas especificidad por secuencias concretas, el nombre de Endonucleasa G ha persistido hasta la actualidad, por lo que todas ellas se agrupan dentro de esta familia. De hecho, a pesar de las diferencias entre la actividad de las nucleasas mitocondriales de *Saccharomyces* y *Neurospora* respecto a las de mamífero, el clonaje y secuenciación de los

genes codificantes para estas proteínas mostró que existe una elevada homología de secuencia entre ellas, corroborando que las mitocondrias de todos los organismos eucariotas contienen una EndoG evolutivamente conservada (116).

En la actualidad, EndoG ha sido identificada en todos los genomas completos que existen de eucariotas, incluyendo *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Xenopus laevis*, *Drosophila Melanogaster*, *Saccharomyces cerevisiae*, *Caenorhabditis elegans* y *Neurospora crassa*, entre otros (116). Curiosamente, en el genoma de eucariotas superiores se ha encontrado un gen parálogo al de EndoG denominado *endonuclease G-like-1*. Estudios recientes han mostrado que este gen contiene la secuencia codificante de la nucleasa mitocondrial EXOG que, a diferencia de su paróloga EndoG, sí presenta actividad endo y exonucleasa. Estos resultados sugieren que estas dos enzimas podrían complementarse para desempeñar las mismas funciones que se han propuesto para las endo-exonucleasas de otros organismos (117). Adicionalmente, también se ha descrito que algunas bacterias como *Serratia marcescens* y *Anabaena sp.* contienen genes que codifican para nucleasas con un 42% y 28% de identidad de secuencia respecto a la EndoG bovina, respectivamente, lo cual apoya la teoría endosimbionte del origen de la mitocondria. Estas nucleasas bacterianas son secretadas y utilizadas para la obtención de nucleótidos del medio extracelular (116).

Dada la especial relevancia de esta familia de proteínas para el desarrollo de esta tesis, en los siguientes apartados profundizaremos un poco más en su estructura, actividad enzimática, mecanismo de acción y funciones descritas hasta ahora.

1.3.1 Estructura, localización y caracterización enzimática

EndoG pertenece a la superfamilia de nucleasas no específicas con dominio $\beta\beta\alpha$ de unión a metal, las cuales se han encontrado en numerosos organismos tanto eucariotas como procariotas. Dentro de este último grupo de organismos, cabe destacar la nucleasa extracelular de *Serratia marcescens* por ser la que más se ha estudiado y mejor se conoce. De hecho, tanto el modelo estructural como el mecanismo de acción propuesto para la nucleasa de *Serratia* (118,119) fueron empleados como modelos para elaborar el primer modelo molecular y funcional de la EndoG bovina (120). El dominio $\beta\beta\alpha$ de unión a metal, como su nombre indica, está compuesto por una horquilla beta seguida de una hélice alfa que se estructuran conformando el sitio de unión para el ion metálico, imprescindible para la

actividad catalítica de estas enzimas. De hecho, se ha demostrado que la actividad de estas nucleasas es dependiente de cationes divalentes como Mg^{2+} , Mn^{2+} o Co^{2+} (incluso Ca^{2+} , en el caso de EXOG), puede ser inhibida por la presencia de cationes monovalentes como el K^+ o el Na^+ y presenta mayor actividad a pHs ligeramente ácidos (120,121). En cuanto al estado de oligomerización de la proteína, los primeros estudios con EndoG recombinante apuntaban a que se trataba de una proteína dimérica, aunque estudios posteriores matizaron que podría existir un equilibrio monómero-dímero (120). Cuando la proteína se encuentra en estado dimérico, los centros activos de cada uno de los monómeros se sitúan de forma distante y apuntando en sentidos opuestos, lo que sugiere que cada uno de ellos une e hidroliza sustratos de manera independiente (122,123).

Otra de las características de esta familia de nucleasas es que la mayoría de ellas presentan un motivo DRGH en su dominio $\beta\beta\alpha$ que contiene varios residuos importantes para la catálisis, entre los cuales destaca la histidina. Este residuo, altamente conservado, actúa como una base general activando una de las moléculas de agua que rodean al catión metálico del centro activo. Una vez activada, la molécula de agua va a llevar a cabo un ataque nucleófilo sobre el átomo de fósforo del ADN que producirá la hidrólisis de su enlace fosfodiéster, siendo este el mecanismo por el cual EndoG lleva a cabo su actividad nucleasa. La sustitución de esta His por un residuo de Ala produce una enzima incapaz de degradar el ADN, haciendo patente su papel clave en el proceso catalítico. Por su parte, el residuo de Asp parece estar implicado en la conformación del centro activo y su sustitución por un residuo de Ala produce una enzima con aproximadamente la mitad de actividad que su homóloga no mutada. Por último, el residuo de Arg participa en la unión del ADN y su correcto posicionamiento respecto al sitio catalítico, produciendo su sustitución por Ala una enzima sin actividad catalítica (120,122). Curiosamente, el aspártico de este motivo DRGH está reemplazado por una serina tanto en EXOG como en las EndoGs de tripanosomátidos, característica que exploraremos más adelante en esta tesis. Adicionalmente, existen otros residuos en la estructura de EndoG cuya presencia es fundamental para su actividad. Entre ellos, cabe destacar la Asn del centro catalítico, necesaria para la coordinación del catión metálico, así como varios residuos de Arg implicados en la unión del sustrato (120,122).

Al igual que ocurre con muchas otras proteínas mitocondriales, EndoG está codificada por un gen nuclear que contiene una región codificante para un péptido señal en su extremo amino

terminal que produce su direccionamiento a la mitocondria (116,124). La compartimentalización mitocondrial ha mostrado ser un mecanismo para el silenciamiento de la actividad nucleasa de EndoG, ya que su expresión ectópica extra-mitocondrial puede inducir la muerte celular (120). Además, como ya hemos comentado previamente, la localización de EndoG cambia durante el proceso apoptótico, translocándose de la mitocondria al núcleo para degradar el ADN genómico (125). No obstante, existen otros mecanismos para el silenciamiento de estas enzimas. Por ejemplo, las nucleasas de *Serratia marcescens* y *Vibrio vulnificus* tienen que ser activadas mediante la formación de puentes disulfuro entre residuos de cisteína, permaneciendo inactivas en el citoplasma en ausencia de los mismos (126,127). Otro ejemplo lo constituye NucA, nucleasa de *Anabaena*, la cual se encuentra inhibida en el citoplasma celular mediante a su unión a un polipéptido específico denominado NuiA, previniendo con ello su toxicidad (116). También se ha reportado un mecanismo de silenciamiento similar en la EndoG de *Drosophila melanogaster* (*DmEndoG*). A pesar de que en este insecto *DmEndoG* está silenciada por su compartimentalización en el espacio intermembrana mitocondrial, se ha descrito que existe una segunda proteína localizada en el núcleo, *DmEndoGI*, que podría actuar como un “seguro de vida” en caso de que parte de la EndoG llegara al núcleo por un fallo en el sistema de transporte a la mitocondria. Esta proteína inhibidora está compuesta por dos dominios altamente conservados unidos por una cadena polipeptídica flexible. La estructura cristalina del complejo *DmEndoG/DmEndoGI* muestra como cada uno de estos dominios está unido a los centros activos de cada subunidad del homodímero *DmEndoG*, bloqueando así los sitios de unión del sustrato e impidiendo la actividad enzimática (123).

Por último, estas enzimas se caracterizan por ser nucleasas no específicas, es decir, son capaces de digerir ácidos nucleicos de diversa naturaleza. EndoG es activa frente a ARN, ADN e híbridos ARN/ADN de cadena doble o cadena sencilla, mostrando preferencia por este último. De hecho, cuando EndoG tiene como sustrato ADN_{dc}, genera primero cortes de cadena sencilla (113,115,128). No obstante, cabe destacar que existen diferencias notables en la actividad catalítica de distintos miembros de este grupo. Por ejemplo, la proteína homóloga a EndoG en *C. elegans* (CPS-6) y la EndoG humana (hEndoG) solo tienen actividad endonucleasa, mientras que Nuc1p, proteína homóloga a EndoG en *S. cerevisiae*, y la EXOG humana (hEXOG), tienen tanto actividad 5'-exonucleasa como endonucleasa (121). En

conjunto, las diferencias encontradas en las actividades endo/exonucleasa de estas enzimas apuntan a que podrían tener diferentes mecanismos de acción o desempeñar distintos papeles dentro de la célula. En el siguiente apartado, describiremos algunas de las funciones más importantes que se han descrito para estas nucleasas.

1.3.2 Funciones descritas

A pesar de que EndoG ha sido extensivamente caracterizada desde su descubrimiento, todavía no están realmente claras cuáles son sus funciones dentro de la célula, si bien se ha propuesto su implicación en numerosos procesos fisiológicos. Por un lado, como ya se ha mencionado en esta introducción, se ha demostrado que EndoG es liberada de la mitocondria tras la permeabilización de su membrana externa y migra al núcleo para degradar el ADN nuclear (ADNn) durante el proceso de apoptosis (125,129–131). No obstante, al menos en algunos organismos, EndoG parece no estar sola en esta tarea. Por ejemplo, se ha visto que hEndoG es capaz de interactuar con otras proteínas mitocondriales como el factor inductor de apoptosis (AIF), la Hsp70 o la flap-endonucleasa-1 (FEN-1), produciéndose como resultado de estas interacciones un aumento de la actividad nucleasa de la enzima (132). Además, AIF contiene una señal de localización nuclear en su secuencia, por lo que se ha sugerido que su asociación con EndoG podría ser la responsable de la translocación al núcleo durante la apoptosis (133). De forma similar, se ha demostrado que CPS-6 interactúa con la proteína WAH-1, homóloga a AIF en *C. elegans*, y que esta interacción aumenta la capacidad de EndoG de degradar el ADN durante la muerte celular (134). De hecho, se han identificado otras cuatro nucleasas en este nemátodo que parecen estar involucradas en el progreso normal de la apoptosis: CRN-1 (proteína homóloga a FEN-1), CRN-4, CRN-5 y CYP-13. Estas nucleasas son capaces de interactuar entre ellas y con CPS-6, sugiriendo que podrían estar formando un complejo multiproteico denominado “degradosoma” junto a WAH-1 y llevando a cabo de forma conjunta la fragmentación y degradación del ADN *in vivo* (135,136). Un poco más particular es el caso de *S. cerevisiae*, donde se ha demostrado que Nuc1p puede inducir la apoptosis de forma independiente a AIF. En su lugar, Nuc1p interactúa con la carioferina KAP123, que facilita su translocación al núcleo, así como con la histona 2B fosforilada del ADN nuclear, siendo esta unión necesaria para la acción nucleolítica de Nuc1p durante la muerte

celular (131). Estas diferencias apuntan a que las proteínas de la familia EndoG podrían estar implicadas en la ejecución de la muerte celular por distintas vías.

Por otra parte, existen indicios de que EndoG podría estar involucrada en el metabolismo del ADN mitocondrial (ADNmt) y genómico (ADNg) mediante su participación en procesos de recombinación, reparación, replicación y degradación de estos. Por ejemplo, como ya hemos comentado anteriormente, el hecho de que las proteínas homólogas a EndoG de *Neurospora crassa* y *Saccharomyces cerevisiae* presenten actividad 5'-exonucleasa, además de endonucleasa, sugiere que podrían jugar un papel en la recombinación del ADNmt. Más concretamente, se propuso que esta actividad 5'-exonucleasa podría actuar ampliando cortes de cadena sencilla o "gaps" existentes en el ADNdc para producir los tramos de ADNcs dentro de su estructura necesarios para que tengan lugar los eventos de recombinación (109,137). Este tipo de actividad ha sido confirmada *in vitro* para Nuc1p, hEXOg y CPS-6, por lo que *a priori* también podrían estar implicados en la recombinación del ADNmt (121). No obstante, la primera confirmación de la participación en este proceso *in vivo* se obtuvo en *S. cerevisiae*, donde la inactivación de la proteína Nuc1p produjo una disminución en la frecuencia de eventos de recombinación y de conversión génica en el ADNmt (137). Por su parte, hEndoG es el único de los miembros estudiados de esta familia que no tiene actividad frente a este tipo de sustratos, por lo que su participación en recombinación a través de este mecanismo parece improbable, lo cual no implica que no pueda estar haciéndolo mediante otros mecanismos. Por ejemplo, Huang *et al.* demostraron que hEndoG era la única enzima aislada de células HeLa capaz de producir cortes de doble cadena en la secuencia de recombinación del virus del herpes simple tipo 1 (HSV-1), secuencia responsable de la inversión que se produce durante la replicación del ADN de estos virus (138). Además, existen varias publicaciones que muestran como la EndoG de mamíferos podría estar implicada en la recombinación del ADNn. En este punto, es importante señalar que, pese a que EndoG es una proteína principalmente mitocondrial, existen numerosos estudios que muestran que parte de la proteína puede encontrarse en el núcleo (128,138–141), por lo que no es descabellado que pueda desempeñar un papel fisiológico dentro de este orgánulo. En este sentido, se ha mostrado que la hEndoG puede generar los cortes de doble cadena en las regiones *switch* del ADN necesarias para que se inicie la recombinación de conmutación de clase o CSR (*Class Switch Recombination*), siendo este proceso necesario para que tenga lugar la conmutación

de clase de inmunoglobulinas en los linfocitos B (142). Otra publicación del año 2014 muestra que EndoG puede inducir la recombinación en la región *MLLbcr* del gen *MLL* tras la inducción de estrés replicativo. Estos autores sugieren que EndoG podría estar cortando en la región *MLLbcr* para reparar vía recombinación las horquillas de replicación que están estancadas como resultado de este estrés replicativo. Sin embargo, estos cortes permiten a su vez que tengan lugar reorganizaciones en el gen *MLL*, las cuales se asocian con la gran parte de las leucemias de mal pronóstico (141). Por último, un ensayo llevado por Robertson *et al.* mostró que EndoG digiere preferentemente sustratos que contienen residuos de 5-hidroximetilcitosina (5hmC) en regiones ricas en guanina (G) y citosina (C), produciendo cortes de doble cadena como resultado de esta actividad (143). Estas regiones ricas en GC son considerados puntos calientes o *hotspots* de recombinación. Así mismo, se ha visto que los residuos de 5hmC tienden a acumularse en regiones en las que existe daño en el ADN y en *hotspots* de recombinación (144). Por último, un estudio publicado este mismo año ha demostrado que, además de digerir preferentemente residuos 5hmC, EndoG presenta una afinidad hasta 50 veces mayor sobre las uniones de Holliday (estructuras del ADN en forma de cruz que se forman durante la recombinación) que sobre el ADNdc (145). En su conjunto, estos resultados sugieren que EndoG está implicada en la reparación del ADNn mediante mecanismos de recombinación.

Otro de los papeles que se ha otorgado a EndoG es el de participar en la reparación del ADNmt. Actualmente, se sabe que las especies reactivas de oxígeno (ROS) producidas como consecuencia de la fosforilación oxidativa mitocondrial son la principal causa de daño en el ADNmt. En este sentido, se ha demostrado que existe una correlación directa entre los niveles de expresión de EndoG y el consumo de oxígeno en los distintos tejidos de rata, sugiriendo que los niveles de EndoG varían con el ratio de daño oxidativo en el ADNmt (146). Adicionalmente, se ha demostrado que la EndoG bovina presenta mayor afinidad por el ADN dañado como consecuencia del tratamiento con agentes como el ácido L-ascórbico o la peplomicina, que producen cortes de cadena sencilla por la acción de los radicales activos de oxígeno sobre este; así como sobre el ADN tratado con cisplatino, que causa el entrecruzamiento intracatenario entre guaninas adyacentes. En conjunto, estos resultados sugieren que EndoG podría participar en la reorganización y reparación de las distorsiones locales que causan estos agentes en el ADN y, por lo tanto, estar involucrada en el

mantenimiento del ADNmt (140). Por otra parte, se ha propuesto que hEXOG podría participar en el mecanismo reparación del ADN mitocondrial por escisión de bases (BERmit), principal mecanismo de reparación del daño oxidativo en el ADN de este orgánulo (147). En la mitocondria, la ADN polimerasa γ (Poly) se encarga tanto de la replicación como de la reparación del ADNmt, es decir, no existe una ADN polimerasa específica para cada uno de estos procesos como sí sucede en el núcleo. Una de las características de la Poly es que es muy poco eficiente reparando cortes puntuales de cadena sencilla en el ADN. Sin embargo, su actividad aumenta notablemente cuando el número de nucleótidos a incluir en el proceso de polimerización es mayor de dos. En este contexto, se ha propuesto que hEXOG podría colaborar en el proceso de BERmit elongando estos cortes puntuales gracias a su actividad 5'-exonucleasa y producido así un sustrato óptimo para que la Poly pueda sintetizar la hebra complementaria de forma eficiente (**Figura 10**) (148). En este sentido, se ha demostrado que la disminución en los niveles de hEXOG produce una acumulación de cortes de cadena sencilla en el ADNmt, no viéndose este efecto cuando se disminuyen los niveles de FEN-1 o DNA2, enzimas mitocondriales que desempeñan un papel similar en el proceso de BER nuclear. Adicionalmente, se ha visto que la disminución de los niveles de hEXOG produce una acumulación de ROS intracelular, probablemente a causa del mal funcionamiento mitocondrial como consecuencia de la acumulación de daños en su ADN, que finalmente conducen a la apoptosis celular. Este hallazgo muestra además que la acumulación de daños en el ADNmt puede iniciar la apoptosis intrínseca en células de mamífero (149). Finalmente, cabe destacar que hEXOG contiene una extensión en su extremo carboxilo terminal, ausente en la mayoría de los miembros de la familia EndoG, que se estructura formando un dominio en forma de ala sobre el centro catalítico y va a ser el responsable de la actividad 5'-exonucleasa de esta enzima (148). Curiosamente, la EndoG de tripanosomátidos también contienen una extensión similar en su extremo C-terminal, característica que será explorada más adelante en los resultados de esta tesis.

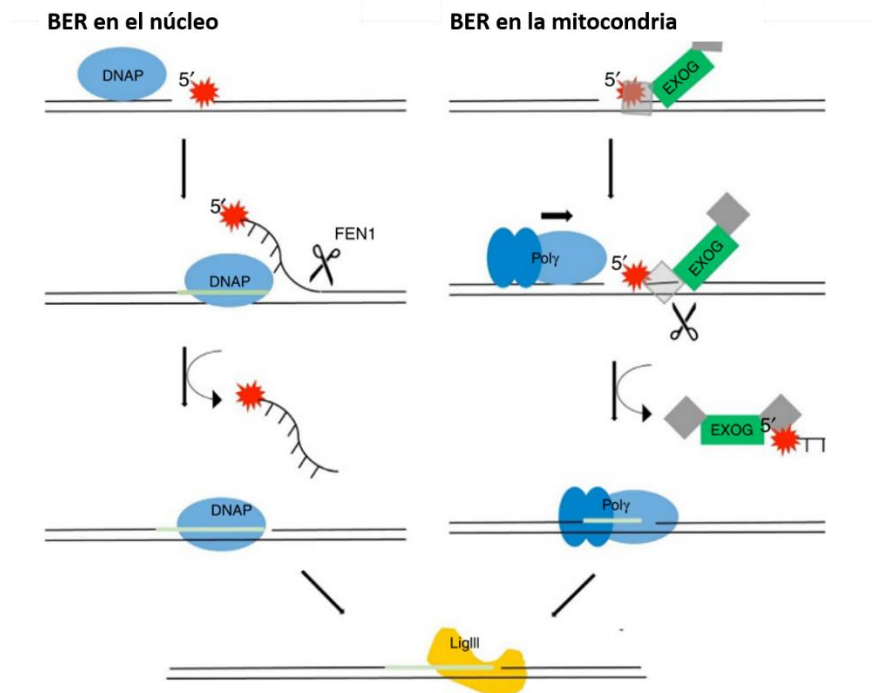


Figura 10. Representación esquemática del mecanismo de reparación por escisión de bases en el núcleo (izquierda) y en la mitocondria (derecha). Figura adaptada de (148).

Por último, como hemos mencionado anteriormente, también se ha propuesto que EndoG podría participar en la replicación del ADNmt. Un ensayo llevado a cabo por Côte y Carrillo mostró que, al menos *in vitro*, EndoG es capaz de generar los oligonucleótidos de ARN necesarios para que la Poly comience la replicación del ADNmt gracias a su actividad ARNasa H (actividad que permite degradar el ARN en los dímeros ADN:ARN) (124). Sin embargo, esta función ha sido discutida, ya que a pesar de que EndoG tiene actividad ARNasa H, no presenta especificidad por los dímeros ADN:ARN, siendo esta cualidad típica de este tipo de enzimas (116). Por otra parte, una publicación mucho más reciente ha demostrado que el silenciamiento de la hEndoG por ARNi reduce significativamente el inicio de la transcripción del ADNmt, apuntando nuevamente a la implicación de la enzima en este proceso (150). Adicionalmente, en esta misma publicación, los autores muestran que esta enzima también podría estar involucrada en la degradación del ADNmt dañado, mostrando el conjunto de sus resultados que hEndoG puede regular el equilibrio entre la degradación y la síntesis del ADNmt y, por tanto, participar en el mantenimiento de su homeostasis (150). Finalmente, un estudio del pasado año indica que hEXOg también podría formar parte de la maquinaria de replicación del ADNmt. Según estos autores, hEXOg actuaría eliminando los cebadores de

ARN del extremo 5' de la cadena de ADN que se está sintetizando, permitiendo con ello que la ADN ligasa III pueda unir los extremos de la molécula circular de ADNmt una vez terminada la replicación. No obstante, la capacidad de llevar a cabo este proceso solo ha sido demostrado *in vitro*, por lo que sería necesario realizar estudios adicionales que demuestren su papel *in vivo* (151).

En conjunto, la gran variedad de funciones en las que se ha involucrado a EndoG señala que esta enzima tiene un papel dual en el desarrollo celular, siendo importante tanto para la regulación de la supervivencia como de la muerte de las células. Este papel dual ha sido bien caracterizado en la levadura *S. cerevisiae*, donde la delección de Nuc1p produce un incremento de la muerte celular necrótica cuando la fosforilación oxidativa está reprimida (en condiciones de fermentación), mientras que disminuye la muerte celular apoptótica cuando la respiración mitocondrial está incrementada (131). Otro ejemplo de este doble papel lo encontramos en *C. elegans*, donde la reducción de los niveles de CPS-6 mediante ARNi lleva a un retraso en el crecimiento celular durante el desarrollo del nemátodo; mientras que, por otra parte, también afecta a la degradación del ADN durante la apoptosis (129). Sin embargo, el papel provida de EndoG en mamíferos ha sido mucho más discutido. Por un lado, un estudio *in vitro* mostró que el silenciamiento de EndoG por ARNi producía un defecto en la proliferación celular y un cambio en la distribución del ciclo celular que se asemejaba a una parada en la fase G₂, estando este efecto normalmente asociado a la acumulación de daños no reparados en el ADNn (128). Sin embargo, el papel de EndoG *in vivo* es mucho más controvertido, ya que se han obtenido resultados contradictorios en ratones *knockout* en los que se ha eliminado completamente el gen que codifica esta proteína. Para entender mejor estos resultados, es importante señalar que en ratones el gen de EndoG está solapado con otro gen de función desconocida. Los grupos de investigación que publicaron que los ratones EndoG^{-/-} no mostraban ningún rasgo fenotípico atípico se limitaron a eliminar únicamente el gen que codifica para EndoG. Por el contrario, en el trabajo en el que se describe que EndoG es necesaria para la embriogénesis, la apoptosis y la fragmentación del ADN nuclear, ambos genes habían sido eliminados (152–154). Adicionalmente, cabe destacar que todos estos resultados son previos al descubrimiento de EXOG. Esta proteína surge como resultado de la duplicación de un gen ancestral que generó las subfamilias parálogas EndoG y EXOG en eucariotas superiores, de forma que se ha propuesto que estas dos enzimas podrían tener

funciones complementarias y análogas (117). Teniendo esto en cuenta, parece plausible que al eliminar el gen de EndoG parte de sus funciones fueran compensadas por EXOG, por lo que sería interesante estudiar el efecto de la delección conjunta de ambos genes sobre el desarrollo y la apoptosis en eucariotas superiores.

1.3.3 La EndoG de *Leishmania infantum* (LiEndoG)

Como ya se ha discutido anteriormente en esta introducción, la idea de que existan mecanismos de muerte celular regulada en organismos unicelulares ha sido materia de controversia y discusión durante los últimos años, si bien parece que esta posibilidad está cada vez más aceptada. Aun así, los mecanismos que regulan este proceso y las proteínas involucradas en el mismo son todavía muy poco conocidas. En este sentido, la finalización de varios proyectos de secuenciación genética en tripanosomátidos ha permitido que se realicen numerosas búsquedas informáticas encaminadas a identificar proteínas similares a las que regulan la muerte celular en eucariotas superiores. Siguiendo estas estrategias, nuestro grupo de investigación consiguió identificar una putativa Endonucleasa G en *Leishmania infantum* (LiEndoG) cuya homología de secuencia con las EndoGs de otros organismos la hizo postularse como una clara candidata en la búsqueda de efectores del proceso apoptótico en este parásito (101). Por este motivo, la caracterización de LiEndoG ha sido una de las principales líneas de investigación seguidas por nuestro grupo durante los últimos años, siendo también el objetivo principal perseguido en esta tesis. Por ello, describiremos a continuación las características de LiEndoG que ya conocíamos antes de comenzar con este trabajo y que sirvieron como base sobre la cual continuar nuestra investigación y, por tanto, para obtener los resultados que se mostrarán en el siguiente capítulo de esta tesis.

Una vez conocida la secuencia de LiEndoG, nuestro grupo se planteó construir un modelo tridimensional de la proteína mediante el empleo de herramientas bioinformáticas y modelos estructurales proteicos depositados en la base de datos *Protein Data Bank* (www.pdb.org). Los primeros análisis mostraron que tanto la endonucleasa secretada de *Serratia marcescens* como la NucA de *Anabaena sp.* tenían una alta identidad de secuencia con LiEndoG y un patrón de plegamiento similar, gracias a lo cual se pudo construir un primer modelo molecular con un elevado nivel de confianza ($p\text{-valor} < 0.0001$). Estas dos proteínas son miembros de la superfamilia de nucleasas no específicas con dominio $\beta\beta\alpha$ de unión a metal, indicando que

*Li*EndoG debía formar parte también de esta familia. Un análisis más detallado de los residuos implicados en el proceso catalítico sugirió además que el mecanismo de hidrólisis del enlace fosfodiéster llevado a cabo por *Li*EndoG debía ser similar al descrito para las nucleasas de *Serratia* y *Anabaena*. Otra de las similitudes estructurales que encontramos entre *Li*EndoG y la nucleasa extracelular de *Serratia* es que, al igual que esta, parece que su actividad catalítica depende de la formación de enlaces disulfuro en su estructura, ya que *Li*EndoG pierde su actividad catalítica en presencia de β -mercaptoetanol. De hecho, un análisis de la estructura de *Li*EndoG llevado a cabo con el servidor DISULFIND predijo la formación de dos enlaces disulfuro entre las posiciones C100-C116 y C434-C439 con un alto nivel de confianza. No obstante, si bien es cierto que *Li*EndoG presenta en general un alto grado de homología estructural con estas proteínas, también tiene características propias que la hacen diferente. Una de estas diferencias es que la EndoG de *Leishmania infantum*, así como en otras especies de *Leishmania* y *Trypanosoma*, presenta en su sitio catalítico un motivo SRGH en lugar del motivo DRGH típico de estas enzimas, aunque las implicaciones de esta diferencia nos eran desconocidas al inicio de esta tesis.

Nuestro siguiente paso fue purificar una *Li*EndoG recombinante (*rLi*EndoG) para poder caracterizar su actividad nucleasa *in vitro*. Los resultados obtenidos por nuestro grupo mostraron que era posible expresar y purificar la proteína *rLi*EndoG de *E. coli*. No obstante, es importante destacar que, tras la inducción de la expresión, la proteína se agregaba rápidamente formando cuerpos de inclusión, siendo probablemente este el motivo por el que no tuvo un efecto tóxico para las bacterias. Así mismo, el hecho de que la proteína precipitase hizo que no fuese posible obtenerla directamente en su forma nativa de los lisados celulares. Por ello, hubo que someterla a un proceso de desnaturalización y posterior renaturalización para tratar de obtener una proteína recombinante correctamente plegada tras la purificación.

De forma similar a lo observado para las EndoGs de otros organismos, nuestros resultados mostraron que *rLi*EndoG requería la presencia de cationes divalentes como Mg^{2+} , Mn^{2+} o Co^{2+} para su actividad. Por el contrario, la presencia de concentraciones moderadas de Na^+ y K^+ tuvieron un fuerte efecto inhibitorio sobre la actividad de esta nucleasa. Esta capacidad inhibitoria fue aún más marcada en el caso del potasio, siendo capaz de abolir completamente la actividad de *rLi*EndoG a una concentración de 150mM. Según Sen *et al.* (92), en los promastigotes la concentración intracelular de K^+ está habitualmente en torno a 140mM,

mientras que disminuye a 60mM poco tiempo después de ser tratados con camptotecina, compuesto que se utiliza para inducir la muerte celular en *Leishmania*. Nuestros resultados mostraron que a una concentración de K^+ de 75mM la actividad de rLiEndoG solo estaba ligeramente inhibida, por lo que la disminución de la concentración de este catión durante la muerte celular descrita por Sen *et al.* podría suponer un mecanismo para la activación de la enzima durante la apoptosis.

Como hemos comentado anteriormente, se ha mostrado que la EndoG de eucariotas superiores se dirige a la mitocondria gracias a la presencia de un péptido señal en su extremo N-terminal. De forma similar, el análisis de la secuencia de aminoácidos de LiEndoG empleando el programa MITOPROT reveló que nuestra proteína también contenía una secuencia de direccionamiento mitocondrial en este extremo. Posteriormente, la localización mitocondrial de LiEndoG fue confirmada mediante técnicas inmunohistoquímicas que empleaban anticuerpos anti-rLiEndoG.

Para estudiar si, tal y como se había descrito para otras EndoGs, LiEndoG era capaz de migrar desde la mitocondria al núcleo como respuesta a estímulos apoptóticos, se emplearon promastigotes transfectados con la proteína quimérica LiEndoG-eGFP. Bajo condiciones de crecimiento normales, LiEndoG-eGFP estaba claramente localizada en la mitocondria de los promastigotes. Sin embargo, tras el tratamiento con edelfosina, la señal de fluorescencia producida por la proteína quimérica se localizaba principalmente en el núcleo. Estos resultados fueron a su vez confirmados por Western blot comparando la presencia de LiEndoG en extractos nucleares, citosólicos y mitocondriales de promastigotes control y tratados con edelfosina. En su conjunto, estos resultados mostraron que LiEndoG es una proteína mitocondrial capaz de migrar al núcleo como respuesta a estímulos apoptóticos.

Por último, quisimos saber si LiEndoG podría estar participando en la degradación del ADN durante la muerte celular. Para testar esta hipótesis, se transfectaron promastigotes de *L. infantum* con un fragmento de ADN que contenía la secuencia codificante de LiEndoG, de forma que estos parásitos sobreexpresaran constitutivamente la proteína. El posible efecto de la sobreexpresión de LiEndoG sobre la fragmentación ADN de los parásitos fue cuantificado empleando el método TUNEL, técnica citométrica basada en cuantificar la cantidad de dUTPs marcados que han sido incorporados a los extremos 3'OH libres presentes en el ADN gracias a la acción catalítica de la deoxinucleotidil transferasa terminal o TdT. Los resultados de estos

experimentos mostraron que la sobreexpresión de *LiEndoG* en los promastigotes no producía ninguna alteración significativa en su patrón de TUNEL. Sin embargo, tras el tratamiento con edelfosina, se observó un aumento significativo de la señal de TUNEL en los parásitos que sobreexpresaban *LiEndoG* en comparación con los controles, indicando que bajo estas condiciones la sobreexpresión de *LiEndoG* producía un aumento en la degradación del ADN.

En su conjunto, los resultados obtenidos por nuestro grupo soportan la conclusión de que *Leishmania infantum* expresa una nucleasa mitocondrial, homóloga a la EndoG de otros eucariotas, que migra al núcleo como respuesta a estímulos apoptóticos para participar en la ejecución del proceso de muerte celular en estos parásitos (101).

Hipótesis

Además de su demostrado papel durante el proceso de muerte similar a apoptosis, la Endonucleasa G de *L. infantum* es necesaria para el correcto crecimiento de los parásitos y, por ello, puede ser considerada como una diana terapéutica para el desarrollo de fármacos frente a la leishmaniasis.

Objetivo general

Caracterizar el papel provida de *LiEndoG* y la función de sus dominios estructurales específicos.

Resultados

La OMS define las enfermedades tropicales desatendidas o NTD (*Neglected Tropical Diseases*) como un conjunto diverso de enfermedades con alta prevalencia en zonas tropicales y subtropicales de todo el planeta que afectan cada año a más de mil millones de personas (una sexta parte de la población mundial), provocan más de un millón de muertes y suponen un coste asociado de billones de dólares anualmente (157). Estas enfermedades, que afectan en mayor medida a las zonas más pobres del planeta, pueden llegar a suponer hasta el 11% de la carga mundial de enfermedad, por lo que representan uno de los mayores retos a los que se enfrenta la ciencia y la salud pública. Sin embargo, el grifo de producción de nuevos fármacos para su tratamiento está prácticamente cerrado, recibiendo menos de un 1% de la financiación total para la investigación y desarrollo de nuevos medicamentos (158). Dentro de estas NTD algunas de las que tienen unos índices de mortalidad más elevados son las causadas por agentes de la familia *Trypanosomatidae* como son la enfermedad de Chagas (*Trypanosoma cruzi*), la enfermedad del sueño (*Trypanosoma brucei*) y la leishmaniasis (*Leishmania spp.*). Como mencionamos en la introducción de esta tesis, se estima que sólo la leishmaniasis afecta cada año a cerca de 1 millón de personas y puede llegar a producir más de 50 mil muertes (11). Esta enfermedad ha ido ganando importancia durante los últimos años en Europa, sobre todo en la cuenca del Mediterráneo, donde la leishmaniasis ya es considerada una enfermedad endémica. En estos países, la influencia de diversos factores como la globalización, el cambio climático o las migraciones masivas asociadas a conflictos armados desde zonas endémicas ha facilitado que se puedan reunir las condiciones necesarias para que se dé la enfermedad (la presencia del vector, un hospedador susceptible y un reservorio), exponiendo cada vez a un mayor número de individuos a contraerla (12). A esto hay que sumarle que los facultativos de estas zonas al no estar familiarizados con esta enfermedad puedan en muchas ocasiones no realizar un correcto diagnóstico diferencial, con lo que no se administran los tratamientos adecuados a los pacientes y se facilita la libre circulación del parásito (159). Por último, esta situación se ve aún más agravada debido a que los fármacos de los que se dispone actualmente para el tratamiento de esta enfermedad son limitados, anticuados, tóxicos y, en su mayoría, tienen varias décadas de antigüedad, lo que ha permitido la aparición de parásitos resistentes a estos tratamientos. Por todo ello, es necesario y urgente identificar tanto nuevos agentes terapéuticos con potencial para tratar

estas enfermedades de forma efectiva y segura, como nuevas dianas moleculares que puedan mejorar el diseño y descubrimiento de nuevos fármacos (160), siendo ambas necesidades el *leitmotiv* del trabajo desarrollado en esta tesis. El resultado de este trabajo de investigación ha permitido la publicación de los tres artículos científicos que resumimos y presentamos a continuación.

4.1 Los derivados de la 5'-tritiltimidina como una nueva clase de compuestos antileishmania: la Endonucleasa G de *Leishmania infantum* como posible diana terapéutica

4.1.1 Hipótesis

Dado que los parásitos del género *Leishmania* son capaces de obtener nucleósidos de pirimidina del medio extracelular, la utilización de compuestos derivados de este tipo moléculas puede ser una aproximación válida para el desarrollo de nuevos agentes con actividad leishmanicida.

4.1.2 Objetivos

- Evaluar la actividad leishmanicida de una serie de compuestos derivados de la 5'-tritiltimidina.
- Caracterizar el proceso de muerte celular inducido por estos compuestos.
- Identificar la posible diana terapéutica sobre la que estos compuestos están actuando dentro de las células.

4.1.3 Resultados principales

En el trabajo que exponemos a continuación se testaron dos series de derivados de timidina con sustituyentes 5'-trifenilmetilo (5'-tritilo) frente a promastigotes y amastigotes axénicos de *Leishmania infantum*. Varios de estos compuestos mostraron una potente actividad leishmanicida, con valores de IC₅₀ en el rango micromolar. Dentro de ellos, la 3'-O-(isoleucilisoleucil)-5'-O-(3,3,3,-trifenilpropanoil)timidina, denominado **compuesto 28** en este artículo, mostró una actividad leishmanicida particularmente potente sobre los amastigotes

intracelulares. Los ensayos llevados a cabo para caracterizar la muerte celular inducida por este compuesto sobre los parásitos mostraron que es capaz de producir alteraciones significativas en el potencial de membrana mitocondrial, aumentar la concentración de anión superóxido y causar una disminución significativa de la degradación del ADN durante el proceso de muerte celular. En su conjunto, nuestros resultados sugieren que esta familia de compuestos podría tener como diana en *L. infantum* una proteína mitocondrial involucrada en la degradación del ADN durante la muerte celular, ambos requisitos presentes en la nucleasa *LiEndoG* previamente caracterizada por nuestro grupo. Finalmente, nuestros resultados muestran que los miembros de esta familia de compuestos con actividad leishmanicida son capaces de inhibir la actividad nucleasa de *LiEndoG in vitro*, lo que apoya la hipótesis de que esta proteína podría ser la diana sobre la que actúan.

4.1.4 Resumen

Una aproximación racional para buscar nuevos fármacos frente a las enfermedades parasitarias consiste en explorar las diferencias bioquímicas que existen entre el parásito y el hospedador humano. Una de estas diferencias la encontramos en el metabolismo de los nucleósidos, dado que los parásitos del género *Leishmania* carecen de las herramientas necesarias para la síntesis *de novo* de purinas, necesitando obtenerlos del medio extracelular. Por ello, la búsqueda de inhibidores del catabolismo de purinas y/o de sus transportadores a través de la membrana ha sido una de las estrategias que se han empleado durante los últimos años para el desarrollo de compuestos con actividad leishmanicida. Sin embargo, las investigaciones llevadas a cabo en este sentido han mostrado que estas rutas son altamente redundantes, pudiendo muchos de los pasos llevarse a cabo por varias enzimas o a través de sub-rutas accesorias, por lo que su inhibición se vuelve complicada (161). Por otra parte, si bien *Leishmania* es capaz de sintetizar nucleósidos pirimidinas *de novo*, se ha mostrado que estos parásitos también emplean rutas de rescate del medio extracelular para su obtención. En este caso, existen menos ejemplos en la literatura en la que se haya analizado el efecto de inhibidores o antimetabolitos basados en análogos de pirimidina, aunque si se han descrito varios compuestos con una prometedora actividad leishmanicida (162–167). Basándonos en estos resultados, y aprovechando que nuestro laboratorio tiene puesto a punto un procedimiento para la evaluación de la actividad leishmanicida de compuestos que nos

proporcionan grupos de síntesis química con los que mantenemos una estrecha colaboración, decidimos probar una colección de derivados de nucleósidos de pirimidina para evaluar su efecto frente a estos parásitos. Más concretamente, esta serie de compuestos estaba formada por un grupo derivados de timidina a los que se habían agregado distintos sustituyentes en el carbono 5' y/o 3' de la desoxirribosa, así como pequeñas modificaciones en la propia base pirimidínica. Los resultados de estos ensayos dieron lugar a la publicación del artículo 1 (*5'-Trityl-Substituted Thymidine Derivatives as a Novel Class of Antileishmanial Agents: Leishmania infantum EndoG as a Potential Target*; (156)), el cual resumimos a continuación y puede consultarse en el apartado 4.1.5 de esta tesis.

Para testar la actividad leishmanicida de los derivados de timidina se determinó en primer lugar la dosis de cada uno capaz de inhibir a la mitad la proliferación de promastigotes y amastigotes de *L. infantum*, es decir, su IC₅₀. Estos primeros ensayos permitieron comprobar que una alta proporción de estas moléculas tenía una buena actividad frente a los promastigotes de *L. infantum*, aunque mostraban valores de IC₅₀ ligeramente superiores cuando se trataba de amastigotes (artículo 1, tabla 1). Los compuestos que mostraron mayor actividad leishmanicida (IC₅₀ < 25 µM) se testaron posteriormente en células hospedadoras (Jurkat) para evaluar así su citotoxicidad. Los valores de IC₅₀ obtenidos para cada compuesto y línea celular, así como la estructura molecular de los mismos, se han adaptado del artículo mencionado anteriormente y recogido en la tabla del **Anexo** incluido al final de esta tesis (156).

Nuestros resultados muestran que los distintos grupos funcionales añadidos a la desoxirribosa en sus posiciones 3' o 5' tienen una clara influencia sobre el efecto que ejercen estos compuestos sobre las células. En este sentido, se puede comprobar que estas moléculas necesitan de la presencia de un grupo 3'-O-aminoacil, o un derivado de este, para ejercer su actividad leishmanicida, ya que la presencia de un grupo hidroxilo libre en esta posición (**compuestos 3 y 4**) da lugar a moléculas incapaces de inhibir la proliferación de los parásitos a las concentraciones estudiadas. Adicionalmente, si nos fijamos en la serie de moléculas con un grupo tritil (trifenilmetil) éster en su posición 5' y un aminoácido en la posición 3' (**compuestos 1 y 6-9**), comprobamos como la presencia de una cadena lateral hidrofóbica parece ser determinante para que el compuesto tenga una buena actividad leishmanicida. Dos buenos ejemplos de este fenómeno son los **compuestos 6 y 7**, con una L-isoleucina o una

L-ciclohexilglicina en la posición 3', respectivamente. Adicionalmente, se ensayaron moléculas sustituidas en su posición 3' con dipéptidos (**compuestos 27, 28 y 29**) o un espaciador succinato-glicol entre el aminoácido y el nucleósido (**compuestos 32 y 33**). Este tipo de sustituyentes consiguieron en la mayoría de los casos disminuir la citotoxicidad estas moléculas sin apenas afectar a su actividad leishmanicida, por lo que *a priori* se postulaban como mejores candidatos para eliminar los parásitos en el interior de la célula hospedadora. Por último, se evaluó si la introducción de pequeñas variaciones en la propia base pirimidínica (manteniendo la L-isoleucina en la posición 3') podían afectar a la actividad, pero tanto el derivado con uracilo (**compuesto 20**) como con 3-N-metilamina (**compuesto 21**) presentaron valores de IC₅₀ superiores a los obtenidos para la timina, por lo que se mantuvo esta base nitrogenada para el resto de los compuestos.

En cuanto a los sustituyentes en 5', se evaluó el efecto que podía producir realizar modificaciones sobre el grupo tritilo o sobre el tipo de enlace que le unía al carbono 5' de la ribosa. Nuestros resultados muestran como el valor de IC₅₀ aumenta significativamente cuando el trifenilmetilester es reemplazado por un grupo difenilmetilester (**compuesto 24**), mientras que apenas se ve alterado cuando los tres fenilos se sustituyen por 4-clorofenilos (**compuestos 25 y 26**), demostrando la importancia de este grupo voluminoso para la actividad de estas moléculas. Adicionalmente, se ensayaron compuestos en los que el grupo tritilo se unía a la ribosa mediante un enlace amida o éter en lugar de un enlace éster. En el primer caso los **compuestos 15 y 16** fueron inactivos, por lo que podemos afirmar que el enlace tipo amida actúa en detrimento de la actividad de este tipo de moléculas. En el segundo caso, los **compuestos 2, 34, 37 y 38** mostraron buena actividad frente a promastigotes y amastigotes, así como baja citotoxicidad, por lo que la unión del grupo tritilo a través de un enlace éter también produce compuestos con buena actividad leishmanicida.

El análisis químico realizado hasta este momento nos sirvió para definir la influencia que tienen los distintos grupos funcionales sobre la capacidad de estas moléculas para actuar en los promastigotes y amastigotes axénicos. Sin embargo, era necesario comprobar que este efecto se trasladaba a los parásitos intracelulares, ya que este es el estado en el que se encuentran cuando están infectando al organismo hospedador. Por ello, seleccionamos **los compuestos 2, 6, 28, 32, 33, 37 y 38** por tener una actividad leishmanicida potente y una citotoxicidad moderada frente a células Jurkat y se exploró su actividad antileishmania sobre

los amastigotes intracelulares contenidos en células THP-1 infectadas. La mayoría de estos compuestos consiguieron disminuir el número de células infectadas entre un 20-40%, pero sólo el **compuesto 28**, con una disminución de casi el 90% de macrófagos infectados, mostró una actividad leishmanicida semejante a la obtenida para el tratamiento con edelfosina que se incluyó como control (artículo 1, figura 2).

Teniendo en cuenta los resultados anteriores, pensamos que podía ser interesante explorar el proceso de muerte que producía el **compuesto 28** en *Leishmania*. Nuestros resultados muestran que esta molécula es capaz de inducir cambios en la actividad de la única mitocondria presente en estos parásitos, produciendo una repentina hiperpolarización de su membrana mitocondrial (artículo 1, figura 3a) acompañada de un aumento de la producción de anión superóxido ($O_2^{\cdot-}$) en este orgánulo (artículo 1, figura 3b), pudiendo estos eventos ser los desencadenantes de la muerte celular. Así mismo, evaluamos por citometría de flujo el contenido de ADN de los parásitos tratados con este compuesto, observándose una disminución significativa en el número de parásitos con el ADN degradado (hipoploides) respecto a aquellos tratados con la droga control (estaurosporina), a pesar de que en ambos tratamientos el porcentaje de parásitos no viables era aproximadamente del 50% (artículo 1, figura 4).

La disminución en el porcentaje de parásitos hipoploides tras el tratamiento con el **compuesto 28**, unida a su efecto sobre el $\Delta\psi_m$ y la producción de $O_2^{\cdot-}$, parecían indicar que la diana sobre la que estaba actuando esta molécula podría ser una proteína mitocondrial que mediase en la degradación del ADN durante el proceso de muerte celular, coincidiendo esta localización y papel con el descrito por nuestro grupo para *LiEndoG* (101). Es importante no olvidar que en eucariotas superiores la principal nucleasa mitocondrial encargada de degradar el ADN durante la muerte celular es CAD, no obstante, no se han encontrado proteínas homólogas a esta nucleasa en *Leishmania infantum*, siendo al menos hasta el momento *LiEndoG* la principal candidata a ejercer esta función en estos parásitos.

Por tanto, nuestro siguiente paso fue comprobar si estas moléculas realmente podían estar actuando sobre *LiEndoG*. Para ello, llevamos a cabo una serie de ensayos fluorimétricos que nos permitían evaluar el efecto que tenían estos compuestos sobre la actividad nucleasa de esta enzima (artículo 1 figura 5). Estos experimentos no solo nos permitieron demostrar que los derivados de timidina son capaces de inhibir la actividad nucleasa de *LiEndoG in vitro*, sino

que además mostraron que existe una relación directa entre la inhibición enzimática que producían y su capacidad de inhibir la proliferación de los promastigotes y amastigotes de *L. infantum*. Es decir, los compuestos que presentaban valores de IC₅₀ más bajos y, por tanto, tenían mayor capacidad leishmanicida, eran a su vez los que más inhibían la actividad de *LiEndoG*, y viceversa (artículo 1, figura 5b). Así mismo, pudimos comprobar como el efecto del **compuesto 28** sobre la actividad de *LiEndoG* era dependiente de su dosis, por lo que inequívocamente estaba actuando sobre esta proteína (artículo 1, figura 5a). Este mismo resultado se obtuvo para el **compuesto 37**, denominado **Lei49** en el artículo 2 de esta tesis (artículo 2, figura 5) (168).

El conjunto de los resultados mostrados hasta ahora sugiere que *LiEndoG* es probablemente la diana sobre la cual están actuando estos derivados de timidina en los parásitos. Nuestro grupo ya había demostrado que *LiEndoG* es un miembro de la familia de las nucleasas no específicas con un dominio $\beta\beta\alpha$ de unión a metal que participa en el proceso apoptótico en *L. infantum* translocándose al núcleo y contribuyendo a la degradación del ADN genómico (101). Sin embargo, como se comentó ampliamente en la introducción de esta tesis, se han descrito otros papeles para las proteínas de esta familia además del de la ejecución de la muerte celular. En este sentido, estas nucleasas han mostrado ser necesarias para el crecimiento normal de varios organismos, apuntando a que podrían desempeñar también un papel provida que probablemente está relacionado con su implicación en la homeostasis del ADN mitocondrial y genómico (129,131,169). Otro ejemplo de este posible papel provida lo encontramos en el artículo publicado por McDermott-Roe *et al.*, en el cual demuestran que los ratones carentes de EndoG (*EndoG*^{-/-}) tienen la respiración mitocondrial deteriorada y muestran un aumento en la producción de especies reactivas de oxígeno, así como una disminución en la ratio de ADN mitocondrial frente al genómico (170). Nuestros experimentos nos han permitido comprobar que, al menos el **compuesto 28**, produce un efecto similar al observado por McDermott-Roe sobre la mitocondria de estos parásitos, causando una hiperpolarización de la membrana mitocondrial y un aumento en la concentración de anión superóxido. Teniendo en cuenta todo lo anteriormente expuesto, el conjunto de nuestros resultados sugiere que el bloqueo de *LiEndoG* y, por tanto, de sus posibles funciones provida para el mantenimiento de la homeostasis celular, podría ser la responsable de la actividad

leishmanicida de esta familia de compuestos, convirtiéndola en una potencial diana terapéutica.

DOI: 10.1002/cmdc.201300129

5'-Trityl-Substituted Thymidine Derivatives as a Novel Class of Antileishmanial Agents: *Leishmania infantum* EndoG as a Potential Target

Elena Casanova,^[a] David Moreno,^[b] Alba Gigante,^[a] Eva Rico,^[b] Carlos Mario Genes,^[b] Cristina Oliva,^[b] María-José Camarasa,^[a] Federico Gago,^[c] Antonio Jiménez-Ruiz,^[b] and María-Jesús Pérez-Pérez^{*[a]}

Two series of 5'-triphenylmethyl (trityl)-substituted thymidine derivatives were synthesized and tested against *Leishmania infantum* axenic promastigotes and amastigotes. Several of these compounds show significant antileishmanial activity, with IC₅₀ values in the low micromolar range. Among these, 3'-O-(isoleucylisoleucyl)-5'-O-(3,3,3-triphenylpropanoyl)thymidine displays particularly good activity against intracellular amastigotes.

Assays performed to characterize the nature of parasite cell death in the presence of the tritylthymidines indicated significant alterations in mitochondrial transmembrane potential, an increase in superoxide concentrations, and also significant decreases in DNA degradation during the cell death process. Results point to the mitochondrial nuclease LiEndoG as a target for the action of this family of compounds.

Introduction

Leishmaniasis, one of the most prevalent neglected diseases, is a major health problem in tropical and subtropical countries, with more than 12 million people affected and about 300 million persons at risk of infection. The manifestation of this disease can be cutaneous, diffuse cutaneous, mucocutaneous, or visceral.^[1] It is transmitted to humans by the bite of an infected insect vector, the female *Phlebotominae* sand fly. The *Leishmania* parasite is present in two different forms: the promastigote form in the vector, and the amastigote form in the mammalian host.

Unfortunately, the arsenal of available drugs to combat the various forms of leishmaniasis is very limited. Pentavalent antimonials, pentamidine, amphotericin B, and miltefosine are currently the most effective compounds against this family of diseases.^[2,3] However, treatment with most of these is usually accompanied by significant side effects. These include severe acute pancreatitis and cardiac arrhythmia for the antimonials; nephrotoxicity or even anaphylaxis for amphotericin B; renal, pancreatic, and hepatic toxicities for pentamidine; and an extremely long half-life for miltefosine that favors the develop-

ment of resistance, which is a serious problem common to many of these drugs.^[4] Therefore, there is an urgent need for the development of new antileishmanial agents based on molecular scaffolds distinct from those explored thus far. It can be speculated that by exploring different molecular frameworks, we may obtain bioactive compounds with a mechanism of action different from those of drugs in current use.

Focusing on nucleoside metabolism, *Leishmania* species lack the metabolic pathway for de novo synthesis of purine nucleosides. Therefore, salvage of purines is an obligatory pathway and has been the focus of considerable interest.^[5] In contrast, the situation is quite different for pyrimidines, the metabolism of which has generally been considered less susceptible to therapeutics than the pathway of purine salvage.^[6] Indeed, to the best of our knowledge, there are very few examples of pyrimidine nucleosides such as 5-substituted pyrimidine nucleosides^[7–10] with reported antileishmanial activity in vitro. Nonetheless, it is reasonable to expect that improved knowledge about the enzymes involved in pyrimidine biosynthesis may afford new targets. In this respect, a thymidine kinase in *L. major* was recently identified and shown to be involved in flagellum formation, promastigote shape and growth, and virulence.^[11] Another interesting example comes from the search for inhibitors of the parasitic dUTPase, in which 5'-O-trityl-3'-O-TBDMSi-2'-deoxyuridine was identified as an inhibitor of *L. donovani* replication, with an IC₅₀ value of 17 μM. However, further careful research revealed that this activity could not be ascribed to the inhibition of the related *L. major* dUTPase, as the compound was virtually inactive against the isolated enzyme, indirectly suggesting the involvement of a yet unidentified molecular target.^[12]

[a] Dr. E. Casanova, A. Gigante, Prof. M.-J. Camarasa, Prof. M.-J. Pérez-Pérez
Instituto de Química Médica (IQM-CSIC)
Juan de la Cierva 3, 28006, Madrid (Spain)
E-mail: mjperez@iqm.csic.es

[b] Dr. D. Moreno, Dr. E. Rico, C. M. Genes, C. Oliva, Prof. A. Jiménez-Ruiz
Departamento de Bioquímica y Biología Molecular
Unidad Asociada de I+D+i al CSIC
Campus Universitario, Universidad de Alcalá
Alcalá de Henares, 28871 Madrid (Spain)

[c] Prof. F. Gago
Departamento de Farmacología
Unidad Asociada de I+D+i al CSIC, Campus Universitario
Universidad de Alcalá, Alcalá de Henares (Spain)

Based on these findings, we decided to evaluate our own collection of 5'-trimethylphenyl (trityl)-substituted pyrimidine nucleosides and analogues against *L. infantum* promastigotes. We found that 5'-tritylthymidine derivatives with a valine ester at the 3'-position (compounds **1** and **2**, Figure 1) were able to efficiently inhibit promastigote replication at a concentration of 12.5 μM . Interestingly, the corresponding 3'-hydroxy analogues **3** and **4** (Figure 1) were virtually inactive.

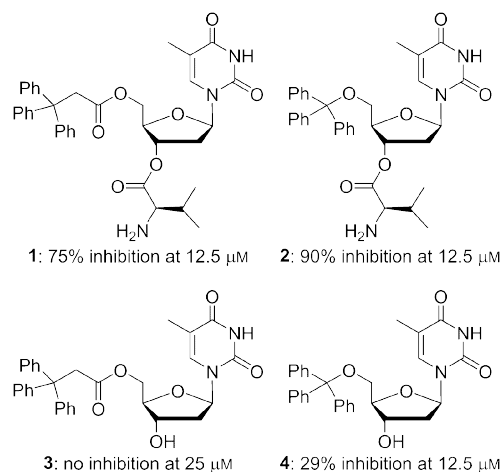


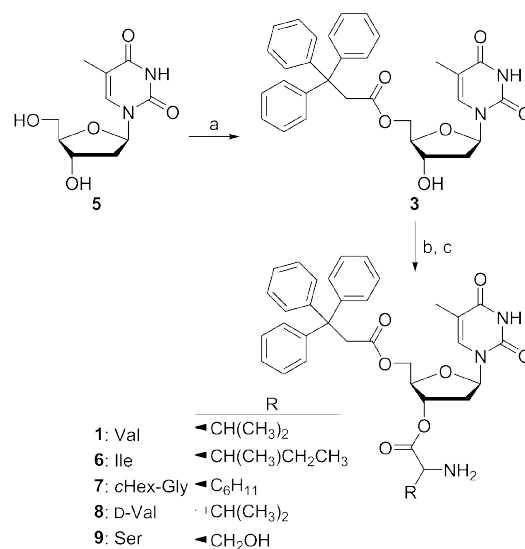
Figure 1. 5'-Trityl nucleosides initially tested against *L. infantum* promastigotes.

The work described herein includes the synthesis and biological evaluation of various analogues of our prototype compounds **1** and **2** by varying the nature of the amino acid at the 3'-position, by modifying the nexus at the 5'-position connecting the nucleoside and the trityl substituent, by introducing subtle changes at the base moiety, and by elongating the substituent at the 3'-position either by preparing dipeptidyl derivatives or by including a spacer between the 3'-hydroxy group and the amino acid residue. The observed changes in mitochondrial activity, together with the surprisingly low percentage of hypoploid cells in the population of treated parasites, led us to the identification of *L. infantum* endonuclease G (LiEndoG), a mitochondrial nuclease, as a target for these molecules.

Results and Discussion

Chemistry

The synthesis of prototype compound **1** and 3'-O-aminoacyl analogues **6–9** is illustrated in Scheme 1. Reaction of thymidine **5** with 3,3,3-triphenylpropionic acid under Mitsunobu conditions (Ph_3P and DBAB) effectively afforded the 5'-substituted compound **3** in 88% yield. The coupling of **3** with various amino acids protected as their corresponding Boc derivatives was performed in the presence of PyBOP and triethylamine. The amino acid derivatives incorporated were Boc-Val-OH, Boc-Ile-OH, Boc-cHexGly-OH, Boc-D-Val-OH, and Boc-Ser(OSi-



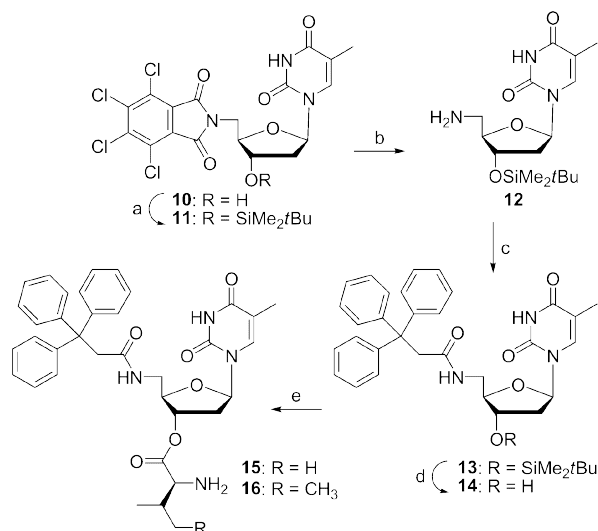
Scheme 1. Reagents and conditions: a) triphenylpropionic acid (2.0 equiv), PPh_3 (2.0 equiv), DBAD (2.0 equiv), DMF, 0 $^\circ\text{C}$, 1.5 h, 88%; b) 1. Boc-Xaa-OH (2.0 equiv), PyBOP (1.2 equiv), Et_3N (2.5 equiv), CH_2Cl_2 , 0 $^\circ\text{C}$, overnight; 2. TFA (23 equiv), CH_2Cl_2 , RT, 3–5 h (**1**, 63%; **6**, 63%; **7**, 52%; **8**, 62%; **9**, 52%).

Me_2tBu)-OH. After coupling, treatment with trifluoroacetic acid (TFA) removed the Boc protecting group together with the Si- Me_2tBu group in the side chain of the Ser derivative to afford the NH_2 -free compounds **1** and **6–9**. Based on the results obtained and synthetic accessibility, the amino acids of choice for the next series of compounds were L-Val and L-Ile.

Our next aim was to modify the linker at the 5'-position connecting the nucleoside and the trityl substituent by replacement of the ester group at the 5'-position by the corresponding amide. This required the synthesis of 5'-amino-5'-deoxythymidine as the key intermediate that had been described by Telzlaff et al.^[13] Thus, treatment of thymidine with tetrachlorophthalimide under Mitsunobu conditions, as described, afforded tetrachlorophthalimido compound **10** (Scheme 2). However, treatment of **10** with hydrazine hydrate to obtain 5'-amino-5'-deoxythymidine gave, in our hands, a compound difficult to isolate and handle. We therefore decided to introduce a protecting group at the 3'-position of **10** prior to its transformation into the amino derivative.

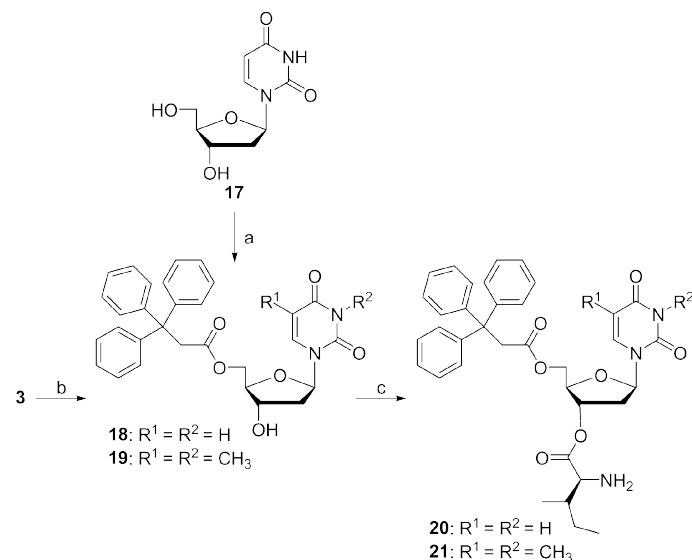
Thus, compound **10** was allowed to react with $\text{tBuMe}_2\text{SiCl}$ in the presence of imidazole to provide the 3'-protected nucleoside **11** (Scheme 2). Reaction of **11** with hydrazine hydrate led to the 5'-amino derivative **12** in 96% yield. This amino nucleoside was subjected to a BOP-mediated coupling with 3,3,3-triphenylpropionic acid to afford amide **13** in excellent yield. Removal of the 3'-O-silyl protecting group (compound **14**) followed by coupling of the amino acids at the 3'-position and Boc deprotection afforded the desired nucleosides **15** and **16**, with Val and Ile residues at the 3'-position, respectively (Scheme 2).

Our next goal was to evaluate how subtle modifications at the pyrimidine base of the nucleoside affects their activity

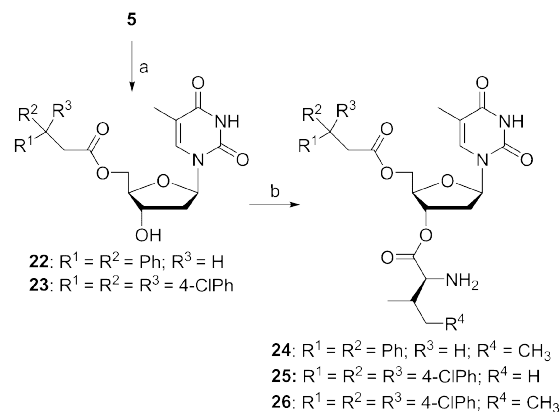


Scheme 2. Reagents and conditions: a) TBDMSCl (1.5 equiv), imidazole (3.0 equiv), DMF, RT, overnight, 77%; b) H₂NNH₂·H₂O (3.0 equiv), EtOH, reflux, 1 h, 96%; c) triphenylpropionic acid (1.5 equiv), BOP (1.5 equiv), Et₃N (1.5 equiv), CH₂Cl₂, RT, overnight 95%; d) 1 N HCl, THF, RT, overnight, 93%; e) 1. Boc-Val-OH or Boc-Ile-OH (2.0 equiv), PyBOP (1.2 equiv), Et₃N (2.5 equiv), CH₂Cl₂, 0 °C, overnight; 2. TFA (23 equiv), CH₂Cl₂, RT, 3–5 h (15, 60%; 16, 70%).

against *in vitro* parasite growth. Reaction of 2'-deoxyuridine (17, Scheme 3) with 3,3,3-triphenylpropionic acid under Mitsunobu conditions afforded the 5'-substituted nucleoside 18 in 70% yield. Alternatively, reaction of the 5'-substituted thymidine derivative 3 with methyl iodide in acetone in the presence of potassium carbonate under microwave (MW) irradiation



Scheme 3. Reagents and conditions: a) triphenylpropionic acid (2.0 equiv), PPh₃ (2.0 equiv), DBAD (2.0 equiv), DMF, 0 °C, 1.5 h, 70%; b) MeI (4.0 equiv), K₂CO₃ (0.5 equiv), acetone, 100 °C, 1 h, MW, 85%; c) 1. Boc-Ile-OH (2.0 equiv), PyBOP (1.2 equiv), Et₃N (2.5 equiv), CH₂Cl₂, 0 °C, overnight; 2. TFA (23 equiv), CH₂Cl₂, RT, 3–5 h (20, 63%; 21, 80%).



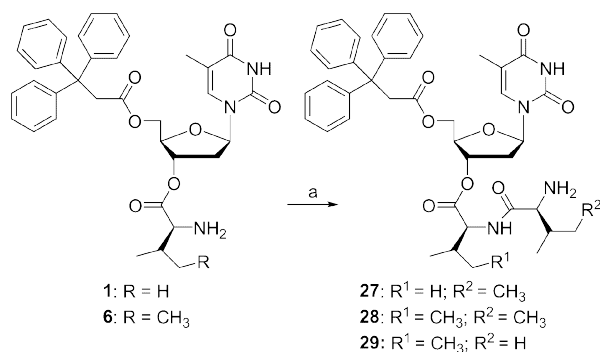
Scheme 4. Reagents and conditions: a) diphenylpropionic acid or 3,3,3-tris(4-chlorophenyl)propionic acid (2.0 equiv), PPh₃ (2.0 equiv), DBAD (2.0 equiv), DMF, 0 °C, 1.5 h (22, 51%; 23, 61%); b) 1. Boc-Ile-OH or Boc-Val-OH (2.0 equiv), PyBOP (1.2 equiv), Et₃N (2.5 equiv), CH₂Cl₂, 0 °C, overnight; 2. TFA (23 equiv), CH₂Cl₂, RT, 3–5 h (24, 58%; 25, 55%; 26, 66%).

at 100 °C for 1 h led to the 3-*N*-methyl analogue 19 in 85% yield. Both nucleosides 18 and 19 were subjected to PyBOP-mediated coupling with Boc-Ile-OH followed by treatment with TFA to yield the respective Ile derivatives 20 and 21.

We also modified the ester group at the 5'-position of the nucleoside by replacing the triphenylpropanoyl group with structurally related analogues. For this purpose, we selected two commercial acids: 3,3-diphenylpropionic acid and 3,3,3-tris(4-chlorophenyl)propionic acid. Thus, reaction of thymidine 5 with the above-mentioned acids under Mitsunobu conditions afforded the 5'-substituted thymidine derivatives 22 and 23 in 51 and 61% yields, respectively (Scheme 4). Reaction of 22 with Boc-Ile-OH followed by treatment with TFA yielded 24 (58% yield). Similarly, reaction of 23 with Boc-Ile-OH and Boc-Val-OH, followed by Boc removal, led to the target compounds 25 and 26 in 55 and 66% respective yields.

We next focused our interest on the 3'-position of the nucleoside. Starting from the 3'-*O*-aminoacyl derivatives 1 and 6, we decided to incorporate an additional amino acid to prepare dipeptidyl derivatives. Thus, reaction of 1 with Boc-Ile-OH and treatment with TFA afforded the dipeptidyl nucleoside 27 (Scheme 5). Similarly, reaction of the isoleucyl derivative 6 with Boc-Ile-OH or Boc-Val-OH, followed by Boc removal, yielded the Ile-Ile compound 28 and the Val-Ile analogue 29.

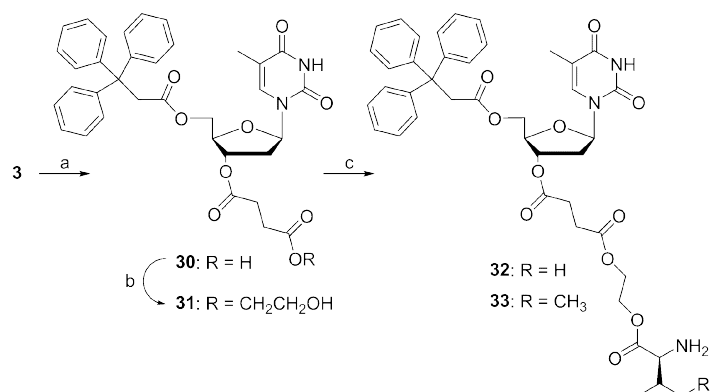
A second series of modifications at the 3'-position of the prototype compound 1 consisted of introducing a spacer between the 3'-hydroxy group of the nucleoside and the amino acid derivative. The proposed spacer incorporated a succinate and an ethylene glycol, as recently employed in the coupling of amino acids to the antiviral cidofovir.^[14] Thus, reaction of 3 with succinic anhydride in pyridine in the presence of DMAP afforded the hemisuccinate 30 in



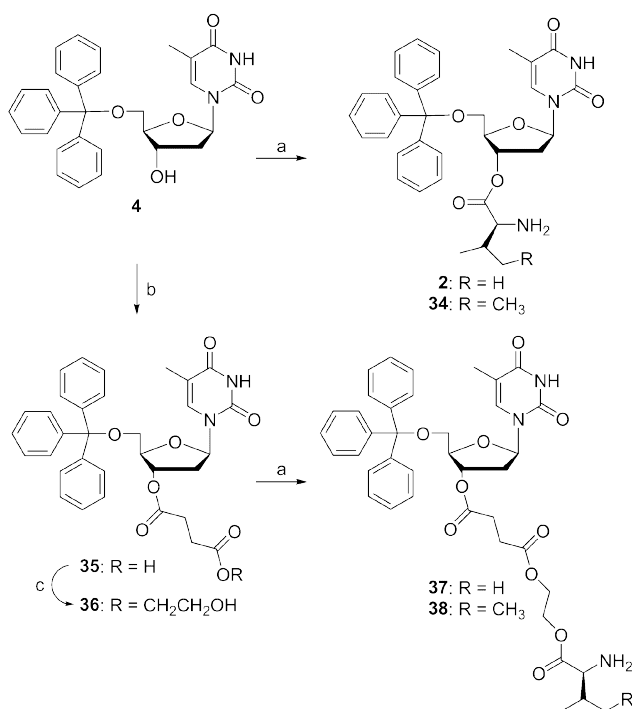
Scheme 5. Reagents and conditions: a) 1. Boc-Ile-OH or Boc-Val-OH (2.0 equiv), PyBOP (1.2 equiv), Et₃N (2.5 equiv), CH₂Cl₂, 0 °C, overnight; 2. TFA (23 equiv), CH₂Cl₂, RT, 3–5 h (27, 54%; 28, 50%; 29, 82%).

82% yield (Scheme 6). Further reaction of **30** with ethylene glycol using PyBOP as the condensation agent afforded alcohol **31** (53% yield). This alcohol was treated with Boc-Val-OH or Boc-Ile-OH in the presence of PyBOP, followed by removal of the Boc protecting group. In this way, conjugates **32** and **33** were obtained (36 and 35% respective yields for the two steps).

The trityl ether **2** had also been identified as an interesting hit in our initial screen of *L. infantum* promastigote replication inhibitors. Based on the significant structural similarity between ester **1** and ether **2**, in the trityl ether series we performed only those modifications that led to the most successful compounds in the ester series. Condensation of 5'-O-tritylthymidine^[15] (**4**, Scheme 7) with Fmoc-Val-OH or Fmoc-Ile-OH was performed in the presence of PyBOP, followed by removal of the Fmoc protecting group with piperidine. In this way the 3'-O-Val **2** and 3'-O-Ile derivatives **34** were obtained (74 and 65% yield, respectively). On the other hand, reaction of 5'-O-tritylthymidine (**4**) with succinic anhydride in dichloromethane in the presence of triethylamine and DMAP afforded



Scheme 6. Reagents and conditions: a) succinic anhydride (2.0 equiv), DMAP (1.0 equiv), pyridine, RT, overnight (82%); b) ethylene glycol (10.0 equiv), PyBOP (1.3 equiv), Et₃N (1.3 equiv), CH₂Cl₂, RT, overnight (53%); c) 1. Boc-Val-OH or Boc-Ile-OH (2.0 equiv), PyBOP (1.2 equiv), Et₃N (2.5 equiv), CH₂Cl₂, 0 °C, overnight; 2. TFA (23 equiv), CH₂Cl₂, RT, 3–5 h (32, 36%; 33, 35%).



Scheme 7. Reagents and conditions: a) 1. Fmoc-Val-OH or Fmoc-Ile-OH (2.0 equiv), PyBOP (1.2 equiv), Et₃N (2.5 equiv), CH₂Cl₂, 0 °C, overnight; 2. piperidine, CH₂Cl₂, RT, 1 h (2, 74%; 34, 65%); b) succinic anhydride (2.0 equiv), DMAP (1.0 equiv), pyridine, RT, 7 h (37, 96%; 38, 71%).

the hemisuccinate **35** in 82% yield. BOP-mediated condensation with ethylene glycol gave alcohol **36** (92% yield), which was coupled to Fmoc-Val-OH or Fmoc-Ile-OH, followed by piperidine treatment to afford conjugates **37** and **38** (96 and 71% respective yields).

Biological evaluation

The synthesized compounds were initially screened against the promastigote form of *L. infantum*. Those compounds showing significant activity in this assay (IC₅₀ < 25 μM) were also tested against amastigotes and, to measure their cytotoxicity against human cell lines, proliferating Jurkat cells. The IC₅₀ values obtained are listed in Table 1. Miltefosine was included as a reference compound.

In the trityl ester series and concerning the evaluation of the 3'-O-aminoacyl derivatives **1** and **6–9** against promastigotes, it is clear that the presence of an amino acid is required for effective inhibition. Moreover, the nature of the amino acid side chain has a clear impact on the antileishmanial effect. Whereas the L-Ser derivative **9** was virtually inactive (IC₅₀ > 25 μM), moderate activity (5 μM < IC₅₀ < 10 μM) was observed for L- and D-Val derivatives **1** and **8**, and good activity (IC₅₀ < 5 μM) was shown when L-Ile and L-cHexGly were included at the 3' position (com-

Table 1. In vitro activity of the synthesized trityl esters, amides, and ethers against the promastigote form of *L. infantum*, the amastigote form of *L. infantum*, and Jurkat cells.^[a]

Compd	IC ₅₀ [μM]		
	Promastigote	Amastigote	Jurkat
1	7.67 ± 0.21	> 25	16.9 ± 0.5
2	2.88 ± 0.13	10.63 ± 0.44	17.1 ± 2.2
3	> 25	ND	ND
4	> 25	ND	ND
6	3.4 ± 0.6	7.33 ± 0.04	9.87 ± 1.8
7	3.5 ± 0.6	6.33 ± 0.01	13.9 ± 0.7
8	6.51 ± 0.03	22.60 ± 1.52	15.3 ± 0.4
9	> 25	ND	ND
15	> 25	ND	ND
16	> 25	ND	ND
20	7.31 ± 3.22	12.76 ± 0.25	9.3 ± 0.3
21	5.87 ± 0.75	9.68 ± 0.36	10.6 ± 0.9
24	17.29 ± 1.64	> 25	22.4 ± 1
25	2.59 ± 0.24	4.49 ± 0.52	6.83 ± 0.1
26	6.56 ± 0.05	7.07 ± 0.02	9.9 ± 0.3
27	6.67 ± 0.23	22.33 ± 0.31	16.2 ± 1.1
28	1.31 ± 0.09	11.75 ± 0.74	15.9 ± 0.6
29	5.99 ± 0.33	13.42 ± 0.14	8.28 ± 1.0
32	6.38 ± 1.17	14.52 ± 1.40	24.9 ± 2.3
33	3.78 ± 0.14	8.65 ± 0.51	20.9 ± 1.2
34	2.22 ± 0.09	6.11 ± 0.01	7.0 ± 0.9
37	4.33 ± 0.05	6.44 ± 0.01	17.2 ± 0.9
38	4.51 ± 0.65	7.37 ± 0.02	15.6 ± 1.5
miltefosine	46.45 ± 3.03	3.96 ± 0.53	48

[a] Values are expressed as the mean ± SEM of three independent experiments; ND: not determined.

pounds **6** and **7**, respectively). Therefore, it can be concluded that a hydrophobic side chain of the amino acid is the preferred substituent for this position. Compounds **6** and **7** maintained good activity when assayed against amastigotes.

Interestingly, the antiparasitic effect was significantly decreased by changing the linker between the trityl substituent and the 5'-position of the nucleoside from an ester to an amide (compare esters **1** and **6** with the corresponding amides **15** and **16**). On the other hand, subtle changes in the pyrimidine base (while maintaining L-Ile at the 3'-position) decreased activity against promastigotes, as evidenced by the fact that both the uracil (**20**) and 3-N-methylthymine (**21**) derivatives showed IC₅₀ values > 5 μM. The IC₅₀ value increased significantly if the triphenylmethyl ester was replaced by the corresponding diphenylmethyl derivative (compound **24**). On the other hand, the 5'-tris(4-chlorophenyl)propanoyl compounds with either Val or Ile at the 3'-position (**25** and **26**, respectively) maintained IC₅₀ values close to 5 μM in promastigotes, showing a strikingly similar activity in amastigotes. IC₅₀ values against Jurkat cells are twofold higher than against the parasites. Moving back to the triphenylpropanoyl compounds, coupling of dipeptides at the 3'-position (**27–29**), or incorporation of a succinate-glycol spacer between the nucleoside and the amino acid (**32** and **33**) caused minimal variations in IC₅₀ values relative to the single amino acid derivatives **1** and **6**, except that toxicity toward Jurkat cells was notably decreased for most of them. In general, the compounds are less active against amastigotes than against promastigotes, yet a signifi-

cant number of them show IC₅₀ values in the low micromolar range (< 10 μM) against the amastigote form of the parasite. Among the most active compounds are Ile derivative **6**, cHexGly derivative **7**, 5'-tris(4-chlorophenyl)propanoyl compounds with Val or Ile at the 3'-position (**25** and **26**, respectively), those with dipeptides coupled at the 3'-position (compound **28**), and the succinate-glycol-Ile derivative **33**.

Concerning the trityl ether derivatives **2**, **4**, **34**, **37**, and **38**, all of the 3'-substituted compounds were active against *L. infantum* promastigotes, with IC₅₀ values < 5 μM. Except for compound **2** (IC₅₀ ≈ 10 μM), all of them also showed significant inhibitory activity against amastigotes, with IC₅₀ values in the 5 μM range. Also in this case, IC₅₀ values against Jurkat cells were three- to fourfold higher than against the parasites.

Leishmanicidal activity was also assayed in amastigote-infected phorbol myristate acetate (PMA)-differentiated THP-1 cells by flow cytometry.^[16] Seven compounds were selected on the basis of their high activity against axenic parasites and moderate toxicity in Jurkat cells. Accordingly, compounds **2**, **6**, **28**, **32**, **33**, **37**, and **38** were assayed over the course of 48 h in differentiated THP-1 cells infected with *L. infantum* amastigotes expressing green-fluorescent protein (eGFP); infected cells were identified by the green fluorescence emitted by their intracellular amastigotes (Figure 2).

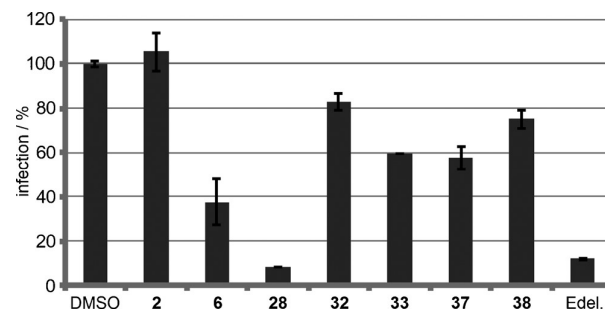


Figure 2. Inhibition of intracellular infection rates: PMA-differentiated THP-1 cells were infected with amastigotes expressing green-fluorescent protein (eGFP). Percentages of infection were evaluated by counting the number of green-fluorescent PMA-differentiated THP-1 cells by flow cytometry. All compounds were assayed at 15 μM. Edelfosine (Edel.) was assayed at 3 μM. Error bars represent standard deviation; results are representative of three independent experiments.

Most of the compounds assayed were able to decrease the number of infected cells by approximately 20–40%, but only compound **28** showed a leishmanicidal activity against intracellular parasites similar to that observed for the control drug edelfosine (87% decrease in the number of infected macrophages), with an estimated IC₅₀ value of 8.0 ± 0.15 μM. According to these results, compound **28** was selected for the study of death-related processes in the parasites. Changes in the mitochondrial transmembrane potential (ΔΨ_m) of logarithmically growing *L. infantum* promastigotes were measured with the potentiometric probe TMRM.^[17] The results presented in Figure 3a indicate a rapid but transient increase in the mean ΔΨ_m when the parasites were incubated with **28** at 10 μM, as

shown by the significant increase in the TMRM-derived fluorescence observed during the first 20 min of treatment. The mean values of fluorescence are restored to that of the control parasites after 80 min exposure to the compound. The mitochondrial proton gradient uncoupler carbonyl cyanide *m*-chlorophenyl hydrazone (CCCP, 100 μM) was added as a positive control of mitochondrial depolarization.

Changes in mitochondrial transmembrane potential are frequently associated with variations in superoxide ($\text{O}_2^{\cdot-}$) production that can be detected by flow cytometry with specific probes such as MitoSOXTM. The increase in MitoSOX-derived fluorescence over the first 20 min of treatment with compound **28**, similar to that observed during the same period in the anti-

mycin-treated cells, is indicative of an accumulation of $\text{O}_2^{\cdot-}$ inside the mitochondrion (Figure 3b). Accordingly, compound **28** induces changes in the activity of the single mitochondrion of the parasites, causing its sudden hyperpolarization and a concomitant increase in superoxide production that eventually leads to cell death.

Figure 4a shows bi-parametric plots of forward scatter (Fs) and propidium iodide (PI) staining of control, staurosporine-, and **28**-treated parasites. The lower right quadrant (C4) represents healthy parasites displaying no alterations in shape or plasma membrane integrity, which constitute 96% of the population of the untreated control parasites. This number is decreased to 54 and 56% as a consequence of 24 h treatment with staurosporine (17 μM) or **28** (10 μM), respectively. Events in any of the other three quadrants are representative of parasites that are either dead or in the process of dying. DNA content in the treated parasites was analyzed by flow cytometry. Surprisingly, although staurosporine and compound **28** at the indicated concentrations cause a similar deleterious effect in the population (Figure 4a), the process of DNA degradation associated with cell death is much lower in **28**-treated parasites (19% of hypoploid cells) than in staurosporine-treated parasites (39% of hypoploid cells) (Figure 4b). The decreased percentage of hypoploid parasites after treatment with **28**, together with its effect on $\Delta\Psi_m$ and $\text{O}_2^{\cdot-}$ production, led us to postulate that the target for this compound could be a mitochondrial protein that mediates nuclear DNA degradation during the cell-death process. The recently characterized LiEndoG protein fulfills both requisites.^[18]

Accordingly, the nuclease activity of this enzyme in the presence of increasing concentrations of **28** was evaluated in a fluorimetric assay. Purified recombinant LiEndoG was incubated with a double-stranded DNA probe (16 nucleotides long) that contains 6-carboxyfluorescein (FAM) attached to both 3' ends and tetramethylrhodamine (TAMRA) bound to both 5' ends. In the undigested probe, FAM-derived fluorescence is quenched by the proximity of TAMRA. Digestion of any of the two oligonucleotides causes separation of the fluorophore from the quencher, so that a fluorescent signal can then be emitted. Fluorescence generated as a consequence of the nuclease activity of LiEndoG clearly diminishes in the presence of increasing concentrations of compound **28** (Figure 5a). This inhibitory effect on LiEndoG was also assayed for compounds **3**, **4**, **6**, and **25**. Reactions in the presence of compounds **3** and **4** at 25 μM were very slightly inhibited relative to the control reaction, which indicates that the substituent at the 3' position is important for LiEndoG inhibitory activity. Compound **6** showed moderate inhibitory activity, whereas compounds **25** and **28** displayed high capacity for LiEndoG inhibition. Thus, the ability of the compounds to inhibit LiEndoG activity strongly correlates with their leishmanicidal capacity, and the presence of substituents at the 3'-position is essential for both activities.

Based on the ability of compound **28** to inhibit LiEndoG activity and also to decrease intracellular infection rates without causing severe cell damage, we decided to gauge the evolution of this compound in PMA-differentiated THP-1 cells. Thus a sample of **28** was incubated with in cell-free extract of these

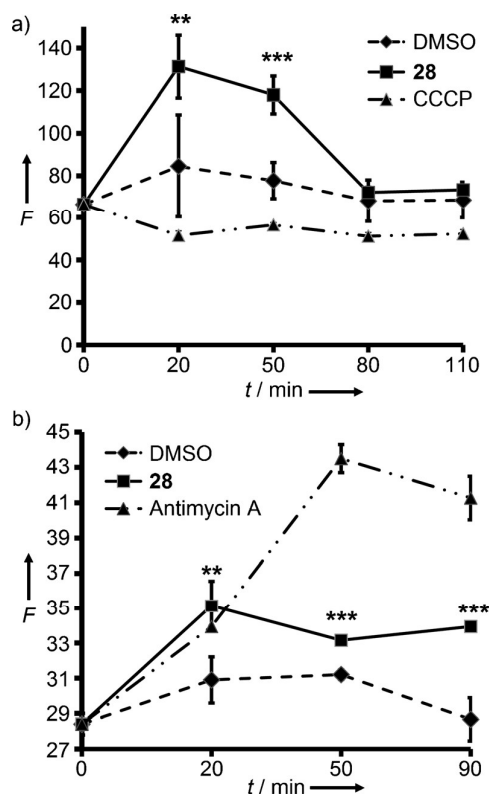


Figure 3. a) Analysis of mitochondrial transmembrane potential ($\Delta\Psi_m$): TMRM-derived fluorescence in *L. infantum* promastigotes treated with compound **28** (10 μM) or with the uncoupler CCCP (100 μM). Y-axis values represent the mean fluorescence (F) of the entire population of living cells. Parasites treated with an equivalent concentration of DMSO used as solvent are included as a control (DMSO). Fluorescence was analyzed by flow cytometry ($n=3$). Samples were collected and analyzed at 20, 50, 80, and 110 min of treatment. Error bars represent standard deviation; results are representative of three independent experiments. b) Analysis of intra-mitochondrial superoxide concentration: MitoSOX-derived fluorescence in *L. infantum* promastigotes treated with compound **28** (10 μM) or with antimycin A (20 μM). Y-axis values represent the mean fluorescence (F) of the entire population of living cells. Parasites treated with an equivalent concentration of DMSO used as solvent are included as a control (DMSO). Fluorescence was analyzed by flow cytometry ($n=3$). Samples were collected and analyzed at 20, 50, and 90 min of treatment. Error bars represent standard deviation; results are representative of three independent experiments. The mean value of the treatments differ from that of the DMSO control with p values <0.05 (**) or 0.005 (***).

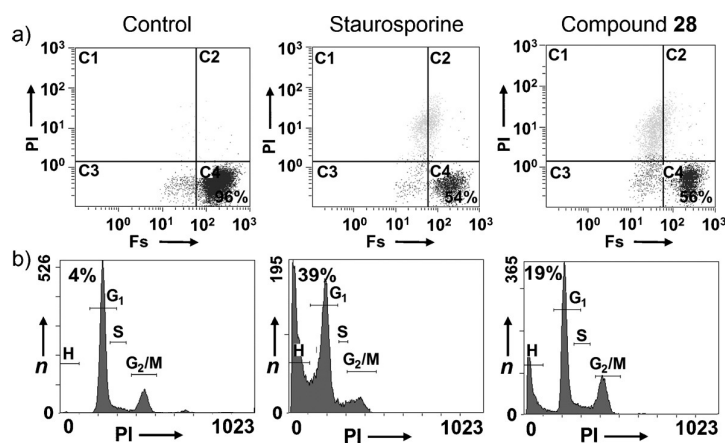


Figure 4. DNA degradation in treated parasites: a) Bi-parametric plots of forward scatter (Fs) and propidium iodide (PI) staining of DMSO (Control), 17 μM staurosporine- and 10 μM compound-28-treated parasites. Healthy parasites without alterations in shape or plasma membrane integrity are located in the lower right quadrant (C4). b) DNA content in the parasite populations shown in panel a). Cell-cycle distributions of the parasites along the G₁, S, or G₂/M phases are indicated. Percentages of parasites with decreased DNA content (hypoploid parasites) are indicated (H); n = number of events; PI = propidium iodide-derived fluorescence.

cells at 37 °C, and the stability of the compound was monitored by HPLC at various time points. Compound **28** evolved toward the Ile derivative **6** and further to the free 3'-hydroxy compound **3**. In contrast, compound **28** is perfectly stable in PBS after 24 h incubation at 37 °C, and no hydrolysis products were detected.

Therefore, it can be hypothesized that the presence of the dipeptide at the 3'-position in compound **28** favors entrance into the differentiated THP-1 cells, where it exerts its leishmanicidal effect until it is enzymatically cleaved toward the Ile derivative **6** and eventually to **3**. It must be pointed out that compound **6** also has significant leishmanicidal activity and is able to cause moderate inhibition of LiEndoG, which may extend the initial activity of compound **28** once it has been modified by endogenous esterases.

The results reported above suggest that LiEndoG is a likely target for the compounds described herein. EndoGs are members of a family of DNA/RNA nonspecific $\beta\beta\alpha$ metal nucleases that have been demonstrated to take part in the apoptotic process, whereby they translocate to the nucleus and contribute to the degradation of genomic DNA into oligonucleosomal fragments.^[18] Apart from their involvement in the death process, these nucleases have been shown to be necessary for the proper growth of several unicellular organisms, indicating a dual (pro-life and pro-death) function, most likely due to their postulated role in mitochondrial DNA recombination.^[19–22] A recent study reveals that *EndoG*^{-/-} mice have impaired mitochondrial respiration and increased production of reactive oxygen species, also associated with a decrease in the ratios of mitochondrial to genomic DNA.^[23] Accordingly, blockage of these pro-life activities of LiEndoG may be responsible for the antileishmanial activity of this family of compounds.

Conclusions

Two series of 5'-trityl-substituted thymidine derivatives were synthesized and assayed against *L. infantum* promastigotes and amastigotes. A significant number of these compounds showed pronounced inhibitory activity against *L. infantum* promastigotes, and had slightly higher IC₅₀ values against amastigotes. Because of its good activity against intracellular amastigotes, compound **28** was selected for analysis of its cellular effects. Profound changes in mitochondrial $\Delta\Psi_m$ and O₂⁻ production were observed as early as 20 min after compound addition. Because of the low percentage of hypoploid parasites

observed after **28**-induced cell death, the nuclease LiEndoG was assayed as a putative target for this family of compounds. The activity of this enzyme in the presence of selected compounds was significantly inhibited. The varying activity of the derivatives against intracellular amastigotes strongly correlates with their ability to inhibit this enzyme. Taken together, LiEndoG can be considered a macromolecular target responsible for the toxic activity of these compounds on *L. infantum*.

Experimental Section

Chemistry

General procedures: Melting points were obtained on a Reichert-Jung Kofler apparatus and are uncorrected. Microanalysis was obtained with a Heraeus CHN-O-RAPID instrument. The elemental compositions of the compounds agreed to within $\pm 0.4\%$ of the calculated values. Electrospray mass spectra were measured on a quadrupole mass spectrometer equipped with an electrospray source (Hewlett-Packard, LC-MS HP 1100). ¹H and ¹³C NMR spectra were recorded on a Varian INNOVA 300 operating at 299 (¹H) and 75 MHz (¹³C), and a Varian INNOVA-400 operating at 399 (¹H) and 99 MHz (¹³C).

Analytical TLC was performed on silica gel 60 F₂₅₄ (Merck) pre-coated plates (0.2 mm). Spots were detected under UV light (254 nm) and/or by charring with phosphomolybdic acid. Separations on silica gel were performed by preparative centrifugal circular thin-layer chromatography (CCTLC) on a Chromatotron (Kieselgel 60 PF₂₅₄ gipshaltig, Merck), with a layer thickness of 1 and 2 mm and flow rate of 4 or 8 mL min⁻¹, respectively. Flash chromatography was performed using a force flow Horizon HPLC system (Biotage) with Flash 25 or 40 silica gel cartridges.

All experiments involving water-sensitive compounds were conducted under scrupulously dry conditions. Et₃N and CH₃CN were dried by holding at reflux over CaH. THF was dried by reflux over sodium/benzophenone. Anhydrous *N,N'*-dimethylformamide (DMF) was purchased from Aldrich. Microwave reactions were performed

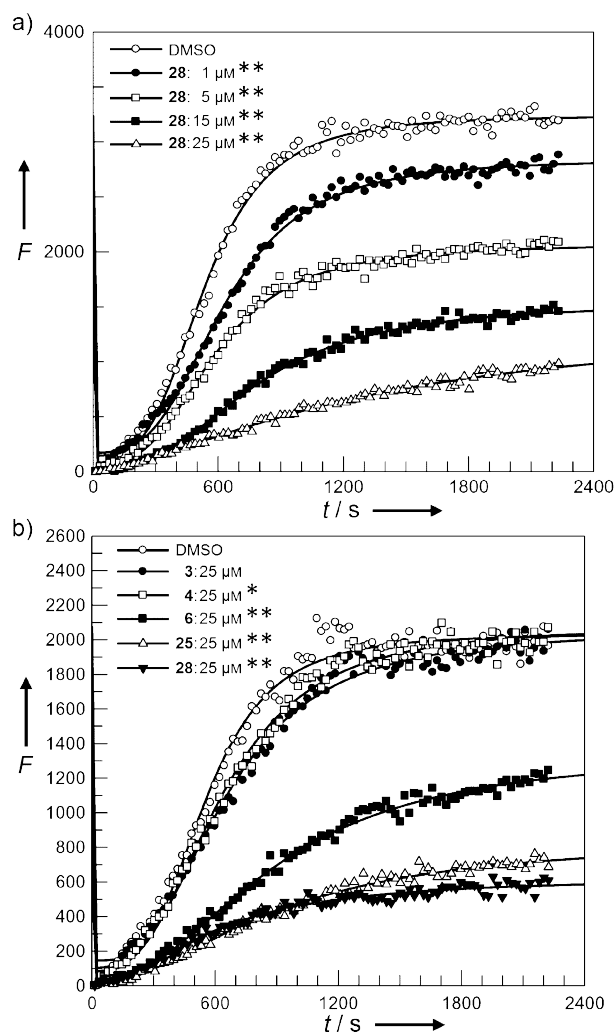


Figure 5. Inhibition of LiEndoG activity was assayed by the cleavage of a dual-labeled double-stranded probe. a) LiEndoG-mediated DNA degradation is followed by an increase in fluorescence (F) observed over 40 min of probe digestion in the presence of increasing amounts of compound **28**. b) LiEndoG-mediated DNA degradation in the presence of compounds **3**, **4**, **6**, **25**, and **28** at $25 \mu\text{M}$. The slopes of the linear section of the curves differ from that of the DMSO control with p values $< 10^{-6}$ (**) or 10^{-3} (*).

with a Biotage Initiator 2.0 single-mode cavity instrument from Biotage (Uppsala, Sweden). Experiments were carried out in sealed microwave process vials using the standard absorbance level (400 W maximum power). The temperature was measured with an IR sensor on the outside of the reaction vessel.

5'-O-(3,3,3-Triphenylpropanoyl)thymidine (3): A solution of dry di-*tert*-butylazodicarboxylate (DBAD; 460 mg, 2.0 mmol) in DMF (2.5 mL) was added dropwise to a solution containing thymidine (**5**; 242 mg, 1.0 mmol), triphenylpropionic acid (604 mg, 2.0 mmol), and Ph_3P (525 mg, 2.0 mmol) in dry DMF (2.5 mL) at 0°C . The mixture was stirred at 0°C under argon for 1.5 h. MeOH (2.5 mL) was added, and volatiles were removed. The residue was treated with 4 N HCl in dioxane (5 mL), and stirred for an additional hour. After removal of the solvent in vacuo, the residue was dissolved in CH_2Cl_2 (50 mL) and washed with 4 N HCl (30 mL) and a solution of NaHCO_3 (30 mL). The organic phase was dried on anhydrous

MgSO_4 , filtered, and evaporated. The residue was purified by flash chromatography ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 100:0–96:4) to yield 463 mg (88%) of **3** as a white solid; mp: $94\text{--}96^\circ\text{C}$ ($\text{CH}_2\text{Cl}_2/\text{MeOH}$); MS (ES, +) m/z 549 [$M+\text{Na}$] $^+$; $^1\text{H NMR}$ (CDCl_3): $\delta = 1.85$ (d, $J = 1.0$ Hz, 3H, 5- CH_3), 1.95 (m, 1H, H-2'), 2.27 (m, 1H, H-2''), 3.67 (d, $J = 15.6$ Hz, 1H, CH_2CO), 3.74 (m, 1H, H-4'), 3.82 (m, 1H, H-3'), 3.91 (d, $J = 15.6$ Hz, 1H, CH_2CO), 3.94 (dd, $J = 12.1, 4.1$ Hz, 1H, H-5'), 4.17 (dd, $J = 12.1, 4.3$ Hz, 1H, H-5''), 6.15 (pseudo-t, $J = 6.4$ Hz, 1H, H-1'), 7.15 (d, $J = 1.1$ Hz, 1H, H-6), 7.20–7.29 (m, 15H, CPh_3), 8.31 ppm (brs, 1H, 3-NH); Anal. calcd for $\text{C}_{31}\text{H}_{30}\text{N}_2\text{O}_6$: C 70.71, H 5.74, N 5.32, found: C 70.54, H 5.66, N 5.19.

General procedure for the synthesis of 3'-amino acid ester derivatives: To a solution of the nucleoside derivative (1.0 mmol) in CH_2Cl_2 (5 mL) at 0°C , the corresponding Boc-protected amino acid (2.0 mmol), PyBOP (624 mg, 1.2 mmol), and Et_3N (345 μL , 2.5 mmol) were added, and the solution was adjusted to pH 11 by the addition of DMAP. The reaction mixture was stirred overnight, diluted with CH_2Cl_2 (50 mL), and washed with a 10% aqueous solution of citric acid (3×20 mL), H_2O (3×20 mL), a 5% aqueous solution of NaHCO_3 (3×20 mL), and brine (3×20 mL). The organic phase was dried on anhydrous MgSO_4 , filtered, and evaporated. The residue was purified as specified. The fractions containing the desired compound were evaporated. A solution of the Boc-protected compound (1.0 mmol) in CH_2Cl_2 (10 mL) was treated with trifluoroacetic acid (TFA; 1.7 mL, 23 mmol). The reaction mixture was stirred at room temperature for 3–5 h, and volatiles were removed. The residue was purified by CCTLC in the Chromatotron, as specified in each case.

5'-O-(3,3,3-Triphenylpropanoyl)-3'-O-valylthymidine (1): Reaction of **3** (100 mg, 0.19 mmol) with Boc-Val-OH (82 mg, 0.38 mmol), PyBOP (118 mg, 0.23 mmol), and Et_3N (66 μL , 0.47 mmol) afforded a residue that was purified in the Chromatotron (hexane/EtOAc, 1:1) to yield 103 mg of the Boc-protected ester [MS (ES, +) m/z 726 [$M+1$] $^+$]. Treatment of this compound with TFA, as described in the general procedure, afforded a residue that was purified in the Chromatotron ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 10:1) to yield 75 mg (63%, two steps) of **1** as a white solid; mp: $85\text{--}86^\circ\text{C}$ ($\text{CH}_2\text{Cl}_2/\text{MeOH}$); MS (ES, +) m/z 626 [$M+1$] $^+$; $^1\text{H NMR}$ ($[\text{D}_6]\text{DMSO}$): $\delta = 0.81$ (d, $J = 6.8$ Hz, 3H, $\text{CH}_3\text{-}\gamma$), 0.86 (d, $J = 6.8$ Hz, 3H, $\text{CH}_3\text{-}\gamma'$), 1.72 (s, 3H, 5- CH_3), 1.85 (m, 1H, CH- β), 2.17 (m, 1H, H-2'), 2.34 (m, 1H, H-2''), 3.14 (d, $J = 5.2$ Hz, 1H, CH- α), 3.84 (s, 2H, CH_2CO), 3.87–4.00 (m, 3H, H-4', H-5'), 4.97 (m, 1H, H-3'), 6.12 (m, 1H, H-1'), 7.15–7.27 (m, 15H, CPh_3), 7.44 (s, 1H, H-6), 11.37 ppm (brs, 1H, NH); Anal. calcd for $\text{C}_{36}\text{H}_{39}\text{N}_3\text{O}_7$: C 69.10, H 6.28, N 6.72, found: C 68.97, H 6.31, N 6.80.

3'-O-Isoleucyl-5'-O-(3,3,3-triphenylpropanoyl)thymidine (6): Following the general procedure, reaction of **3** (100 mg, 0.19 mmol) with Boc-Ile-OH (87 mg, 0.38 mmol), PyBOP (118 mg, 0.23 mmol), and Et_3N (66 μL , 0.47 mmol) afforded a residue that was purified in the Chromatotron ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 10:1) to yield the Boc-protected ester [MS (ES, +) m/z : 762 [$M+\text{Na}$] $^+$]. Treatment of this compound with TFA afforded a residue that was purified in the Chromatotron ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 10:1) to yield 76 mg (63%, two steps) of **6** as a white solid; mp: $78\text{--}79^\circ\text{C}$ ($\text{CH}_2\text{Cl}_2/\text{MeOH}$); MS (ES, +) m/z 640 [$M+1$] $^+$; $^1\text{H NMR}$ ($[\text{D}_6]\text{DMSO}$): $\delta = 0.79\text{--}0.84$ (m, 6H, $\text{CH}_3\text{-}\gamma$, $\text{CH}_3\text{-}\delta$), 1.12 (m, 1H, CH'- γ), 1.40 (m, 1H, CH''- γ), 1.58 (m, 1H, CH- β), 1.72 (s, 3H, 5- CH_3), 2.18 (m, 1H, H-2'), 2.34 (m, 1H, H-2''), 3.18 (d, $J = 5.6$ Hz, 1H, CH- α), 3.83 (s, 2H, CH_2CO), 3.88–4.00 (m, 3H, H-4', H-5'), 4.96 (m, 1H, H-3'), 6.12 (dd, $J = 6.6, 6.3$ Hz, 1H, H-1'), 7.14–7.30 (m, 15H, CPh_3), 7.44 (s, 1H, H-6), 11.36 ppm (brs, 1H, NH); Anal. calcd for $\text{C}_{37}\text{H}_{41}\text{N}_3\text{O}_7$: C 69.47, H 6.46, N 6.57, found: C 69.19, H 6.44, N 6.51.

3'-O-Cyclohexylglycyl-5'-O-(3,3,3-triphenylpropanoyl)thymidine (7): Following the general procedure, reaction of **3** (90 mg, 0.17 mmol) with Boc-cHexGly-OH (88 mg, 0.34 mmol), PyBOP (107 mg, 0.20 mmol) and Et₃N (59 μ L, 0.43 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 20:1) to yield the corresponding ester [MS (ES, +) *m/z* 788 [M+Na]⁺]. Treatment of this compound with TFA and purification in the Chromatotron (CH₂Cl₂/MeOH, 10:1) afforded 59 mg (52%, two steps) of **7** as a white solid; mp: 84–86 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 666 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 0.97–1.16 (m, 5H, cHex), 1.48–1.70 (m, 6H, cHex), 1.72 (d, *J* = 1.0 Hz, 3H, 5-CH₃), 2.18 (m, 1H, H-2'), 2.32 (m, 1H, H-2''), 3.19 (d, *J* = 5.6 Hz, 1H, CH- α), 3.83 (s, 2H, CH₂CO), 3.88 (m, 1H, H-4'), 3.95 (m, 2H, H-5'), 4.96 (m, 1H, H-3'), 6.12 (dd, *J* = 7.9, 6.4 Hz, 1H, H-1'), 7.14–7.27 (m, 15H, CPh₃), 7.44 (d, *J* = 1.0 Hz, 1H, H-6), 11.33 ppm (s, 1H, 3-NH); Anal. calcd for C₃₉H₄₃N₃O₇: C 70.36, H 6.76, N 4.59, found: C 70.12, H 6.91, N 4.62.

5'-O-(3,3,3-Triphenylpropanoyl)-3'-O-D-valylthymidine (8): Following the general procedure, reaction of **3** (122 mg, 0.23 mmol) with Boc-D-Val-OH (100 mg, 0.46 mmol), PyBOP (140 mg, 0.27 mmol), and Et₃N (80 μ L, 0.56 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 30:1) to yield the desired ester [MS (ES, +) *m/z* 748 [M+Na]⁺]. Treatment of the Boc-protected ester with TFA afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield 89 mg (62%, two steps) of **8** as a white solid; mp: 66–68 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 626 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 0.84 (d, *J* = 6.8 Hz, 3H, CH₃- γ), 0.88 (d, *J* = 6.8 Hz, 3H, CH₃- γ), 1.72 (d, *J* = 1.2 Hz, 3H, 5-CH₃), 1.85 (m, 1H, CH- β), 2.14 (m, 1H, H-2'), 2.32 (m, 1H, H-2''), 3.16 (d, *J* = 5.4 Hz, 1H, CH- α), 3.81 (d, *J* = 15.9 Hz, 1H, CH₂CO'), 3.88 (d, *J* = 15.9 Hz, 1H, CH₂CO''), 3.93–4.00 (m, 3H, H-4', H-5'), 4.93 (m, 1H, H-3'), 6.12 (dd, *J* = 8.5, 6.0 Hz, 1H, H-1'), 7.13–7.27 (m, 15H, CPh₃), 7.43 (d, *J* = 1.2 Hz, 1H, H-6), 11.37 ppm (s, 1H, 3-NH); Anal. calcd for C₃₆H₃₉N₃O₇: C 68.88, H 6.38, N 6.69, found: C 69.05, H 6.66, N 6.71.

3'-O-Seryl-5'-O-(3,3,3-triphenylpropanoyl)thymidine (9): Following the general procedure, reaction of **3** (82 mg, 0.16 mmol) with Boc-Ser(TBDMS)-OH^[24] (100 mg, 0.31 mmol), PyBOP (97 mg, 0.19 mmol), and Et₃N (55 μ L, 0.40 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 30:1) to yield the protected ester [MS (ES, +) *m/z* 850 [M+Na]⁺]. Treatment of this compound with TFA afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield 51 mg (52%, two steps) of **9** as a white solid; mp: 85–87 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 614 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 1.72 (s, 3H, 5-CH₃), 2.18 (m, 1H, H-2'), 2.32 (m, 1H, H-2''), 3.39 (t, *J* = 4.6 Hz, 1H, CH- α), 3.56 (m, 2H, CH₂- β), 3.85 (s, 2H, CH₂CO), 3.94–4.00 (m, 3H, H-4', H-5'), 4.87 (brs, 1H, OH), 4.95 (m, 1H, H-3'), 6.13 (dd, *J* = 8.2, 6.2 Hz, 1H, H-1'), 7.15–7.27 (m, 15H, CPh₃), 7.43 ppm (s, 1H, H-6); Anal. calcd for C₃₄H₃₅N₃O₈·0.5H₂O: C 65.58, H 5.83, N 6.75, found: C 65.36, H 6.16, N 6.78.

5'-(1,3-Dihydro-1,3-dioxo-4,5,6,7-tetrachloroisindol-2-yl)-3'-O-(tert-butylidimethylsilyl)-5'-deoxythymidine (11): To a solution containing *tert*-butylidimethylsilyl chloride (107 mg, 0.71 mmol) and imidazole (96 mg, 1.41 mmol) in dry DMF (2 mL), compound **10**^[14] (240 mg, 0.47 mmol) was added. The reaction mixture was stirred at room temperature overnight, and volatiles were removed. The residue was purified by CCTLC in the Chromatotron (EtOAc/hexane, 1:2→1:1) to afford 225 mg (77%) of **11** as a white solid; mp: 110–112 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 622 [M+1]⁺ with 4Cl isotopic pattern; ¹H NMR ([D₆]DMSO): δ = 0.90 (s, 9H, (CH₃)₃), 2.03 (s, 3H, 5-CH₃), 2.09 (m, 1H, H-2'), 2.31 (ddd, *J* = 13.5, 5.5, 2.4 Hz, 1H, H-2''), 3.86 (d, *J* = 7.0 Hz, 2H, H-5'), 4.24–4.30 (m, 2H, H-

3', H-4'), 6.19 (dd, *J* = 8.0, 5.5 Hz, 1H, H-1'), 7.39 (d, *J* = 1.0 Hz, 1H, H-6), 8.14 ppm (s, 1H, 3-NH).

5'-Amino-3'-O-(tert-butylidimethylsilyl)-5'-deoxythymidine (12): To a suspension of **11** (195 mg, 0.31 mmol) in EtOH (5 mL), H₂NNH₂·H₂O (49 μ L, 1.00 mmol) was added. The reaction mixture was stirred at reflux for 1 h. The white solid precipitated was filtered off. The filtrate was evaporated, and the residue was purified by CCTLC in the Chromatotron (CH₂Cl₂/MeOH, 5:1) to afford 107 mg (96%) of **12** as a colorless oil. The analytical and spectroscopic data are identical to those previously described.^[25]

3'-O-(tert-Butylidimethylsilyl)-5'-(3,3,3-triphenylpropanamide)-5'-deoxythymidine (13): To a solution of **12** (109 mg, 0.31 mmol) in CH₂Cl₂ (1.5 mL), 3,3,3-triphenylpropionic acid (139 mg, 0.46 mmol), BOP (204 mg, 0.46 mmol), and Et₃N (64 μ L, 0.46 mmol) were added. The reaction mixture was stirred at room temperature overnight, and volatiles were removed. The residue was purified by CCTLC in the Chromatotron (EtOAc/hexane, 1:2→1:1) to yield 186 mg (95%) of **13** as a white solid; mp: 94–96 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 641 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 0.84 (s, 9H, (CH₃)₃), 1.72 (s, 3H, 5-CH₃), 1.91 (m, 1H, H-2'), 2.02 (m, 1H, H-2''), 3.03 (m, 2H, H-5'), 3.49 (m, 1H, H-4'), 3.60 (s, 2H, CH₂), 4.14 (m, 1H, H-3'), 6.19 (pseudo-t, *J* = 7.1 Hz, 1H, H-1'), 7.11–7.23 (m, 15H, CPh₃), 7.41 (s, 1H, H-6), 7.33 (t, *J* = 5.6 Hz, 1H, NH), 11.30 ppm (s, 1H, 3-NH).

5'-(3,3,3-Triphenylpropanamide)-5'-deoxythymidine (14): A suspension of **13** (140 mg, 0.22 mmol) in THF (2 mL) was treated with 1 N HCl (0.8 mL). The reaction mixture was stirred at room temperature overnight. The reaction was diluted with EtOAc (25 mL) and washed with a solution of NaHCO₃ (2 \times 10 mL) and H₂O (2 \times 10 mL). The organic phase was dried on anhydrous MgSO₄, filtered, and evaporated. The residue was purified by CCTLC in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to afford 106 mg (93%) of **14** as a white solid; mp: 116–118 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 526 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 1.72 (s, 3H, 5-CH₃), 1.97 (m, 2H, H-2'), 3.04 (m, 2H, H-5'), 3.49 (m, 1H, H-4'), 3.60 (s, 2H, CH₂), 3.93 (m, 1H, H-3'), 5.16 (d, *J* = 4.0 Hz, 1H, OH), 6.19 (pseudo-t, *J* = 6.7 Hz, 1H, H-1'), 7.11–7.24 (m, 15H, CPh₃), 7.39 (s, 1H, H-6), 7.78 (t, *J* = 5.7 Hz, 1H, NH), 11.28 ppm (s, 1H, 3-NH); Anal. calcd for C₃₁H₃₁N₃O₅·2H₂O: C 66.30, H 6.28, N 7.48, found: C 65.94, H 6.54, N 7.58.

5'-(3,3,3-Triphenylpropanamide)-3'-O-valyl-5'-deoxythymidine (15): Following the general procedure for the synthesis of 3'-amino acid ester derivatives, reaction of **14** (57 mg, 0.11 mmol) with Boc-Val-OH (47 mg, 0.22 mmol), PyBOP (67 mg, 0.13 mmol), and Et₃N (37 μ L, 0.27 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield the coupling derivative [MS (ES, +) *m/z* 725 [M+1]⁺], that was treated with TFA for Boc deprotection. The final residue was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield 42 mg (60%, two steps) of **15** as a white solid; mp: 124–125 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 625 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 0.93 (d, *J* = 6.8 Hz, 3H, CH₃- γ), 0.94 (d, *J* = 6.8 Hz, 3H, CH₃- γ), 1.73 (s, 3H, 5-CH₃), 2.05–2.25 (m, 3H, H-2', CH- β), 3.09–3.17 (m, 3H, H-5', CH- α), 3.61 (s, 2H, CH₂CO), 3.77 (m, 1H, H-4'), 5.08 (m, 1H, H-3'), 6.10 (dd, *J* = 8.2, 6.1 Hz, 1H, H-1'), 7.14–7.24 (m, 15H, CPh₃), 7.46 (s, 1H, H-6), 7.87 (t, *J* = 4.8 Hz, 1H, NH), 11.35 ppm (s, 1H, 3-NH); Anal. calcd for C₃₆H₄₀N₄O₆: C 69.21, H 6.45, N 8.97, found: C 69.44, H 6.69, N 9.01.

3'-O-Isoleucyl-5'-(3,3,3-triphenylpropanamide)-5'-deoxythymidine (16): Following the general procedure for the synthesis of 3'-amino acid ester derivatives, reaction of **14** (44 mg, 0.08 mmol) with Boc-Ile-OH (39 mg, 0.17 mmol), PyBOP (52 mg, 0.10 mmol), and Et₃N (29 μ L, 0.21 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to provide the coupling

product [MS (ES, +) m/z 739 $[M+1]^+$]. This was treated with TFA for Boc deprotection. The final residue was purified in the Chromatotron ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 10:1) to yield 36 mg (70%, two steps) of **16** as a white solid; mp: 83–85 °C ($\text{CH}_2\text{Cl}_2/\text{MeOH}$); MS (ES, +) m/z 639 $[M+1]^+$; $^1\text{H NMR}$ ($[\text{D}_6]\text{DMSO}$): δ = 0.80–0.84 (m, 6H, $\text{CH}_3\text{-}\gamma$, $\text{CH}_3\text{-}\delta$), 1.15 (m, 1H, $\text{CH}'\text{-}\gamma$), 1.40 (m, 1H, $\text{CH}'\text{-}\gamma$), 1.58 (m, 1H, $\text{CH-}\beta$), 1.72 (s, 3H, 5- CH_3), 2.06–2.26 (m, 2H, H-2'), 3.11 (m, 2H, H-5'), 3.17 (d, J = 5.5 Hz, 1H, $\text{CH-}\alpha$), 3.60 (s, 2H, CH_2CO), 3.69 (m, 1H, H-4'), 4.98 (m, 1H, H-3'), 6.06 (dd, J = 8.3, 6.1 Hz, 1H, H-1'), 7.11–7.23 (m, 15H, CPh₃), 7.47 (s, 1H, H-6), 7.82 (t, J = 5.4 Hz, 1H, NH), 11.34 ppm (s, 1H, 3-NH); Anal. calcd for $\text{C}_{37}\text{H}_{42}\text{N}_4\text{O}_6$: C 69.57, H 6.63, N 8.77, found: C 69.66, H 6.54, N 8.57.

2'-Deoxy-5'-O-(3,3,3-triphenylpropanoyl)uridine (18): To a solution containing 2'-deoxyuridine (**17**; 100 mg, 0.44 mmol), 3,3,3-triphenylpropionic acid (302 mg, 0.87 mmol) and Ph_3P (262 mg, 0.87 mmol) in dry DMF (1.1 mL), a solution of DBAD (230 mg, 0.87 mmol) in DMF (1.1 mL) was added dropwise. The mixture was stirred at 0 °C under argon for 1.5 h. MeOH (2.5 mL) was added, and volatiles were removed. The work-up was performed as described for **3**. The residue that was purified by HPFC ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 100:0→92:8) to yield 158 mg (70%) of **18** as an amorphous white solid. MS (ES, +) m/z 513 $[M+1]^+$; $^1\text{H NMR}$ ($[\text{D}_6]\text{DMSO}$): δ = 2.04 (m, 2H, H-2'), 3.69 (m, 1H, H-4'), 3.80 (brs, 2H, CH_2CO), 3.87 (m, 1H, H-3'), 3.94 (m, 2H, H-5'), 5.30 (d, J = 4.2 Hz, 1H, OH), 5.57 (d, J = 8.0 Hz, 1H, H-5'), 6.07 (pseudo-t, J = 6.7 Hz, 1H, H-1'), 7.15–7.28 (m, 15H, CPh₃), 7.51 (d, J = 8.1 Hz, 1H, H-6), 11.31 ppm (brs, 1H, 3-NH).

N³-Methyl-5'-O-(3,3,3-triphenylpropanoyl)thymidine (19): To a solution of **3** (125 mg, 0.24 mmol) in acetone (2 mL) in 5 mL Pyrex microwave process vial, K_2CO_3 (16 mg, 0.12 mmol) and MeI (60 μL , 0.95 mmol) were added. The reaction vessel was sealed, stirred, and subsequently irradiated for 60 min at 100 °C in a single-mode microwave reactor. The mixture was purified by flash chromatography ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 100:0→96:4) to afford 152 mg (85%) of **19** as a white solid; mp: 56–58 °C ($\text{CH}_2\text{Cl}_2/\text{MeOH}$); MS (ES, +) m/z 541 $[M+1]^+$; $^1\text{H NMR}$ ($[\text{D}_6]\text{DMSO}$): δ = 1.75 (d, J = 1.0 Hz, 3H, 5- CH_3), 2.05 (m, 1H, H-2') 3.15 (s, 3H, NCH₃), 3.41 (m, 1H, H-2'), 3.72 (m, 1H, H-4'), 3.82 (s, 2H, CH_2CO), 3.90–4.01 (m, 3H, H-5', H-3'), 5.31 (d, J = 4.2 Hz, 1H, OH), 6.16 (pseudo-t, J = 6.8 Hz, 1H, H-1'), 7.14–7.27 (m, 15H, CPh₃), 7.43 ppm (d, J = 1.0 Hz, 1H, H-6).

2'-Deoxy-3'-O-isoleucyl-5'-O-(3,3,3-triphenylpropanoyl)uridine (20): Following the general procedure for the synthesis of 3'-amino acid ester derivatives, reaction of **18** (67 mg, 0.13 mmol) with Boc-Ile-OH (60 mg, 0.26 mmol), PyBOP (81 mg, 0.16 mmol), and Et_3N (46 μL , 0.33 mmol) afforded a residue that was purified in the Chromatotron ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 30:1) to yield 72 mg of the coupling product [MS (ES, +) m/z 748 $[M+\text{Na}]^+$]. This compound was treated with TFA for removal of the Boc protecting group. After removal of volatiles, the residue was purified in the Chromatotron ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 10:1) to yield 51 mg (63%, two steps) of **20** as a white solid; mp: 81–83 °C ($\text{CH}_2\text{Cl}_2/\text{MeOH}$); MS (ES, +) m/z 626 $[M+1]^+$; $^1\text{H NMR}$ ($[\text{D}_6]\text{DMSO}$): δ = 0.79–0.84 (m, 6H, $\text{CH}_3\text{-}\gamma$, $\text{CH}_3\text{-}\delta$), 1.12 (m, 1H, $\text{CH}'\text{-}\gamma$), 1.38 (m, 1H, $\text{CH}'\text{-}\gamma$), 1.58 (m, 1H, $\text{CH-}\beta$), 2.17–2.36 (m, 2H, H-2'), 3.18 (d, J = 5.6 Hz, $\text{CH-}\alpha$), 3.82 (s, 2H, CH_2CO), 3.88–4.98 (m, 3H, H-4, H-5'), 4.94 (m, 1H, H-3'), 5.62 (d, J = 8.0 Hz, 1H, H-5), 6.07 (m, 1H, H-1'), 7.14–7.27 (m, 15H, CPh₃), 7.56 (d, J = 8.1 Hz, 1H, H-6), 11.35 ppm (brs, 1H, 3-NH); Anal. calcd for $\text{C}_{36}\text{H}_{39}\text{N}_5\text{O}_7$: C 69.10, H 6.28, N 6.72, found: C 68.94, H 6.06, N 6.68.

3'-O-Isoleucyl-N³-methyl-5'-O-(3,3,3-triphenylpropanoyl)thymidine (21): Following the general procedure for the synthesis of 3'-amino acid ester derivatives, reaction of **19** (148 mg, 0.27 mmol)

with Boc-Ile-OH (127 mg, 0.55 mmol), PyBOP (171 mg, 0.33 mmol), and Et_3N (95 μL , 0.68 mmol) afforded a residue that was purified by flash chromatography ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 100:96:4) to yield 174 mg of the coupling product [MS (ES, +) m/z 776 $[M+\text{Na}]^+$]. This was deprotected by treatment with TFA. After removal of volatiles, the residue was purified by flash chromatography ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 100:0→92:8) to yield 141 mg (80%, two steps) of **21** as a white solid; mp: 67–69 °C ($\text{CH}_2\text{Cl}_2/\text{MeOH}$); MS (ES, +) m/z 654 $[M+1]^+$; $^1\text{H NMR}$ ($[\text{D}_6]\text{DMSO}$): δ = 0.86–0.91 (m, 6H, $\text{CH}_3\text{-}\gamma$, $\text{CH}_3\text{-}\delta$), 1.27 (m, 1H, $\text{CH}'\text{-}\gamma$), 1.43 (m, 1H, $\text{CH}'\text{-}\gamma$), 1.79 (s, 3H, 5- CH_3), 1.87 (m, 1H, $\text{CH-}\beta$), 2.27 (m, 1H, H-2'), 2.39 (m, 1H, H-2'), 3.16 (m, 4H, N- CH_3 , $\text{CH-}\alpha$), 3.81 (d, J = 16.3 Hz, 1H, $\text{CH}_2\text{CO}'$), 3.87 (d, J = 16.3 Hz, 1H, $\text{CH}_2\text{CO}''$), 3.99–4.03 (m, 3H, H-4', H-5'), 5.07 (m, 1H, H-3'), 6.21 (pseudo-t, J = 7.1 Hz, 1H, H-1'), 7.14–7.27 (m, 15H, CPh₃), 7.50 (s, 1H, H-6), 8.34 ppm (brs, 2H, NH₂); Anal. calcd for $\text{C}_{38}\text{H}_{43}\text{N}_5\text{O}_7$: C 69.81, H 6.63, N 6.43, found: C 69.72, H 6.78, N 6.52.

5'-O-(3,3-Diphenylpropanoyl)thymidine (22): As described for the synthesis of **3**, thymidine (**5**; 100 mg, 0.41 mmol) reacted with 3,3-diphenylpropionic acid (186 mg, 0.82 mmol) in the presence of Ph_3P (215 mg, 0.82 mmol) and DBAD (189 mg, 0.82 mmol). The residue was purified by flash chromatography ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 100:0→96:4) to yield 93 mg (51%) of **22** as an amorphous white solid. MS (ES, +) m/z 451 $[M+1]^+$; $^1\text{H NMR}$ ($[\text{D}_6]\text{DMSO}$): δ = 1.74 (s, 3H, 5- CH_3), 1.99–2.18 (m, 2H, H-2'), 3.17 (m, 2H, CH_2CO), 3.80 (m, 1H, H-4'), 4.08–4.15 (m, 3H, H-3', H-5', H-5''), 4.43 (t, J = 8.0 Hz, 1H, CHPh_2), 5.37 (d, J = 4.1 Hz, 1H, OH), 6.15 (pseudo-t, J = 6.8 Hz, 1H, H-1'), 7.14–7.32 (m, 10H, CPh₂), 7.40 (s, 1H, H-6), 11.33 ppm (brs, 1H, 3-NH).

5'-O-[3,3,3-Tris(4-chlorophenyl)propanoyl]thymidine (23): As described for the synthesis of **3**, thymidine (**5**; 100 mg, 0.41 mmol) reacted with 3,3,3-tris(4-chlorophenyl)propionic acid (332 mg, 0.82 mmol) in the presence of Ph_3P (215 mg, 0.82 mmol) and DBAD (189 mg, 0.82 mmol). The final residue was purified by flash chromatography ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 100:0→96:4) to yield 156 mg (61%) of **23** as an amorphous white solid. MS (ES, +) m/z 652 $[M+\text{Na}]^+$ with 3 Cl isotopic pattern; $^1\text{H NMR}$ ($[\text{D}_6]\text{DMSO}$): δ = 1.70 (s, 3H, 5- CH_3), 2.01–2.15 (m, 2H, H-2'), 3.69 (m, 1H, H-4'), 3.80 (d, J = 17.0 Hz, 1H, $\text{CH}_2\text{CO}'$), 3.87 (d, J = 17.0 Hz, 1H, $\text{CH}_2\text{CO}''$), 3.92–4.04 (m, 3H, H-3', H-5'), 5.33 (d, J = 4.2 Hz, 1H, OH), 6.12 (pseudo-t, J = 6.4 Hz, 1H, H-1'), 7.16 (d, J = 8.5 Hz, 6H, CPh₃), 7.32 (d, J = 8.5 Hz, 6H, CPh₃), 7.37 (s, 1H, H-6), 11.32 ppm (s, 1H, 3-NH).

3'-O-Isoleucyl-5'-O-(3,3-diphenylpropanoyl)thymidine (24): Following the general procedure for the synthesis of 3'-amino acid ester derivatives, reaction of **22** (80 mg, 0.18 mmol) with Boc-Ile-OH (82 mg, 0.36 mmol), PyBOP (111 mg, 0.21 mmol), and Et_3N (62 μL , 0.44 mmol) afforded a residue that was purified in the Chromatotron ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 10:1) to yield the coupling product [MS (ES, +) m/z 687 $[M+\text{Na}]^+$]. Treatment with TFA followed by removal of volatiles afforded a residue that was purified in the Chromatotron ($\text{CH}_2\text{Cl}_2/\text{MeOH}$, 10:1) to yield 59 mg (58%, two steps) of **24** as a white solid; mp: 58–60 °C ($\text{CH}_2\text{Cl}_2/\text{MeOH}$); MS (ES, +) m/z 564 $[M+1]^+$; $^1\text{H NMR}$ ($[\text{D}_6]\text{DMSO}$): δ = 0.80–0.84 (m, 6H, $\text{CH}_3\text{-}\gamma$, $\text{CH}_3\text{-}\delta$), 1.12 (m, 1H, $\text{CH}'\text{-}\gamma$), 1.39 (s, m, 1H, $\text{CH}'\text{-}\gamma$), 1.58 (m, 1H, $\text{CH-}\beta$), 1.75 (d, J = 1.1 Hz, 3H, 5- CH_3), 2.22 (m, 1H, H-2'), 2.40 (m, 1H, H-2'), 3.18 (m, 3H, CH_2CO , $\text{CH-}\alpha$), 4.01 (m, 1H, H-4'), 4.15 (m, 2H, H-5'), 4.43 (t, J = 8.0 Hz, 1H, CHPh_2), 5.09 (m, 1H, H-3'), 6.15 (dd, J = 8.2, 6.2 Hz, 1H, H-1'), 7.10–7.17 (m, 2H, CHPh_2), 7.21–7.32 (m, 8H, CHPh_2), 7.47 (d, J = 1.1 Hz, 1H, H-6), 11.38 ppm (s, 1H, 3-NH); Anal. calcd for $\text{C}_{31}\text{H}_{37}\text{N}_5\text{O}_7\cdot 2\text{H}_2\text{O}$: C 62.21, H 7.21, N 7.30, found: C 62.09, H 6.89, N 7.01.

3'-O-Valyl-5'-O-[3,3,3-tris(4-chlorophenyl)propanoyl]thymidine (25): Following the general procedure the synthesis of 3'-amino acid ester derivatives, reaction of **23** (93 mg, 0.15 mmol) with Boc-Val-OH (64 mg, 0.30 mmol), PyBOP (92 mg, 0.18 mmol), and Et₃N (51 μ L, 0.37 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield the corresponding ester [MS (ES, +) *m/z* 852 [M+Na]⁺ with a 3Cl isotopic pattern]. Treatment with TFA and removal of volatiles afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield 60 mg (55%, two steps) of **25** as a white solid; mp: 105–107 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 728 [M+1]⁺ with a 3Cl isotopic pattern; ¹H NMR ([D₆]DMSO): δ = 0.81 (d, *J* = 6.8 Hz, 3H, CH₃- γ), 0.86 (d, *J* = 6.8 Hz, 3H, CH₃- γ), 1.72 (d, *J* = 0.9 Hz, 3H, 5-CH₃), 1.83 (m, 1H, CH- β), 2.21 (m, 1H, H-2'), 2.40 (m, 1H, H-2''), 3.13 (d, *J* = 5.4 Hz, 1H, CH- α), 3.79–4.09 (m, 5H, H-4', H-5', CH₂CO), 5.02 (m, 1H, H-3'), 6.13 (dd, *J* = 7.6, 6.2 Hz, 1H, H-1'), 7.17 (d, *J* = 8.7 Hz, 6H, CPh₃), 7.32 (d, *J* = 8.7 Hz, 6H, CPh₃), 7.45 (d, *J* = 0.9 Hz, 1H, H-6), 11.36 ppm (brs, 1H, NH); Anal. calcd for C₃₆H₃₆Cl₃N₃O₇: C 59.31, H 4.98, N 5.76, Cl 14.59, found: C 59.18, H 4.91, N 5.99, Cl 14.26.

3'-O-Isoleucyl-5'-O-[3,3,3-tris(4-chlorophenyl)propanoyl]thymidine (26): Following the general procedure the synthesis of 3'-amino acid ester derivatives, reaction of **23** (115 mg, 0.18 mmol) with Boc-Ile-OH (84 mg, 0.36 mmol), PyBOP (112 mg, 0.22 mmol), and Et₃N (62 μ L, 0.45 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield 106 mg of the coupling product [MS (ES, +) *m/z* 866 [M+Na]⁺ with a 3Cl isotopic pattern]. After treatment with TFA and removal of volatiles, the residue was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield 88 mg (66%, two steps) of **26** as a white solid; mp: 122–125 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 744 [M+1]⁺ with a 3Cl isotopic pattern; ¹H NMR ([D₆]DMSO): δ = 0.84–0.89 (m, 6H, CH₃- γ , CH₃- δ), 1.20 (m, 1H, CH'- γ), 1.43 (m, 1H, CH'- γ), 1.73 (s, 3H, 5-CH₃), 1.80 (m, 1H, CH- β), 2.24 (m, 1H, H-2'), 2.41 (m, 1H, H-2''), 3.80–3.85 (m, 2H, CH₂CO', CH- α), 3.90 (d, *J* = 16.4 Hz, 1H, CH₂CO''), 3.95–4.10 (m, 3H, H-4, H-5'), 5.10 (m, 1H, H-3'), 6.18 (dd, *J* = 7.8, 6.6 Hz, 1H, H-1'), 7.17 (d, *J* = 8.7 Hz, 6H, CPh₃), 7.32 (d, *J* = 8.7 Hz, 6H, CPh₃), 7.45 (s, 1H, H-6), 11.39 ppm (s, 1H, 3-NH); Anal. calcd for C₃₇H₃₈Cl₃N₄O₆: C 59.81, H 5.15, Cl 14.31, N 5.65, found: C 59.72, H 5.09, Cl 14.23, N 5.82.

3'-O-(Isoleucylvalyl)-5'-O-(3,3,3-triphenylpropanoyl)thymidine (27): Following a similar procedure to that described for the synthesis of 3'-amino acid ester derivatives, reaction of **1** (48 mg, 0.08 mmol) with Boc-Ile-OH (35 mg, 0.15 mmol), PyBOP (47 mg, 0.09 mmol), and Et₃N (27 μ L, 0.19 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 30:1) to yield the coupling product [MS (ES, +) *m/z* 839 [M+1]⁺]. Treatment with TFA and removal of volatiles afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield 32 mg (54%, two steps) of **27** as a white solid; mp: 93–95 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 740 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 0.78–0.95 (m, 12H, CH₃- γ -Ile, CH₃- γ -Val, CH₃- δ -Ile), 1.08 (m, 1H, CH'- γ -Ile), 1.43 (m, 1H, CH'- γ -Ile), 1.57 (m, 1H, CH- β -Ile), 1.72 (s, 3H, 5-CH₃), 2.03 (m, 1H, CH- β -Val), 2.17 (m, 1H, H-2'), 2.34 (m, 1H, H-2''), 3.06 (d, *J* = 5.7 Hz, 1H, CH- α -Ile), 3.78–4.00 (m, 5H, CH₂CO, H-4', H-5'), 4.16 (pseudo-t, *J* = 6.6 Hz, 1H, CH- α -Val), 4.97 (m, 1H, H-3'), 6.12 (dd, *J* = 8.0, 6.4 Hz, 1H, H-1'), 7.14–7.27 (m, 15H, CPh₃), 7.43 (d, *J* = 0.9 Hz, 1H, H-6), 7.95 (brs, 2H, NH₂), 8.12 (d, *J* = 7.4 Hz, 1H, NH-Val), 11.25 ppm (s, 1H, 3-NH); Anal. calcd for C₄₂H₅₀N₄O₈·1.5H₂O: C 65.86, H 6.97, N 7.32, found: C 65.93, H 6.86, N 7.39.

3'-O-(Isoleucylisoleucyl)-5'-O-(3,3,3-triphenylpropanoyl)thymidine (28): As described for **27**, reaction of **6** (90 mg, 0.14 mmol) with Boc-Ile-OH (65 mg, 0.28 mmol), PyBOP (87 mg, 0.17 mmol),

and Et₃N (49 μ L, 0.35 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 30:1) to yield the coupling product [MS (ES, +) *m/z* 876 [M+Na]⁺]. Treatment with TFA and purification of the residue in the Chromatotron (CH₂Cl₂/MeOH, 10:1) afforded 58 mg (55%, two steps) of **28** as a white solid; mp: 111–113 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 754 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 0.80–0.91 (m, 12H, CH₃- γ -Ile, CH₃- δ -Ile), 1.11 (m, 2H, CH'- γ -Ile), 1.23 (m, 2H, CH'- γ -Ile), 1.72 (s, 3H, 5-CH₃), 1.69–1.88 (m, 2H, CH- β -Ile), 2.14 (m, 1H, H-2'), 2.23 (m, 1H, H-2''), 3.56 (d, *J* = 5.8 Hz, 1H, CH- α -Ile), 3.78–4.02 (m, 5H, CH₂CO, H-4', H-5'), 4.25 (pseudo-t, *J* = 6.5 Hz, 1H, CH- α -Ile), 4.96 (m, 1H, H-3'), 6.13 (dd, *J* = 8.1, 6.2 Hz, 1H, H-1'), 7.13–7.27 (m, 15H, CPh₃), 7.43 (s, 1H, H-6), 8.15 (d, *J* = 6.3 Hz, 1H, NH-Ile), 11.37 ppm (s, 1H, 3-NH); Anal. calcd for C₄₃H₅₂N₄O₈: C 68.60, H 6.90, N 7.44, found: C 68.81, H 7.05, N 7.31.

3'-O-(Valyisoleucyl)-5'-O-(3,3,3-triphenylpropanoyl)thymidine (29): As described for **27**, reaction of **6** (155 mg, 0.25 mmol) with Boc-Val-OH (108 mg, 0.50 mmol), PyBOP (155 mg, 0.30 mmol), and Et₃N (86 μ L, 0.62 mmol) afforded a residue that was purified by HPFC (CH₂Cl₂/MeOH, 100:0→96:4) to yield the coupling product [MS (ES, +) *m/z* 836 [M+1]⁺]. Treatment with TFA and removal of volatiles afforded a residue that was purified by HPFC (CH₂Cl₂/MeOH, 100:0→90:10) to yield 151 mg (82%, two steps) of **29** as a white solid; mp: 132–133 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 739 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 0.82–0.96 (m, 12H, CH₃- γ -Ile, CH₃- γ -Val, CH₃- δ -Ile), 1.15 (m, 1H, CH'- γ -Ile), 1.45 (m, 1H, CH'- γ -Ile), 1.72 (s, 3H, 5-CH₃), 1.83 (m, 1H, CH- β -Ile), 2.03 (m, 1H, CH- β -Val), 2.17 (m, 1H, H-2'), 2.34 (m, 1H, H-2''), 3.65 (d, *J* = 5.6 Hz, 1H, CH- α -Val), 3.78–3.90 (m, 3H, CH₂CO, H-4'), 3.94–4.00 (m, 2H, H-5'), 4.25 (pseudo-t, *J* = 6.4 Hz, 1H, CH- α -Ile), 4.96 (m, 1H, H-3'), 6.13 (dd, *J* = 8.4, 6.1 Hz, 1H, H-1'), 7.13–7.27 (m, 15H, CPh₃), 7.43 (d, *J* = 1.1 Hz, 1H, H-6), 7.95 (brs, 2H, NH₂), 8.65 (d, *J* = 6.8 Hz, 1H, NH-Ile), 11.38 ppm (s, 1H, 3-NH); Anal. calcd for C₄₂H₅₀N₄O₈·2H₂O: C 68.27, H 6.82, N 7.58, found: C 68.10, H 6.98, N 7.39.

4-[(2R,3S,5R)-5-(Thymin-1-yl)-2-(3,3,3-triphenylpropanoyloxymethyl)tetrahydrofuran-3-yloxy]-4-oxobutanoic acid (30): To a solution of **3** (132 mg, 0.25 mmol) in anhydrous pyridine (4 mL), DMAP (31 mg, 0.25 mmol), and succinic anhydride (50 mg, 0.50 mmol) were added. The reaction mixture was stirred at room temperature overnight, diluted with CH₂Cl₂ (50 mL), and washed with a 10% aqueous solution of KH₂PO₄ (2×30 mL), and H₂O (30 mL). The organic phase was dried on anhydrous MgSO₄, filtered, and evaporated. The residue was purified by flash chromatography (CH₂Cl₂/MeOH, 100:0→90:10) to afford 129 mg (82%) of **30** as an amorphous solid. MS (ES, +) *m/z* 627 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 1.71 (s, 3H, 5-CH₃), 2.15 (m, 1H, H-2'), 2.30 (m, 1H, H-2''), 2.59–2.72 (m, 4H, CH₂CH₂), 3.78–4.05 (m, 5H, H-4', H-5', CH₂CO), 4.92 (m, 1H, H-3'), 6.11 (dd, *J* = 8.1, 6.2 Hz, 1H, H-1'), 7.17–7.27 (m, 15H, CPh₃), 7.43 (s, 1H, H-6), 11.38 ppm (s, 1H, 3-NH).

2-Hydroxyethyl-[(2R,3S,5R)-5-(thymin-1-yl)-2-(3,3,3-triphenylpropanoyloxymethyl)tetrahydrofuran-3-yl] succinate (31): To a solution of **30** (85 mg, 0.13 mmol) in CH₂Cl₂ (1.0 mL), ethylene glycol (76 μ L, 1.35 mmol), PyBOP (88 mg, 0.17 mmol), and Et₃N (23 μ L, 0.17 mmol) were added. The reaction mixture was stirred at room temperature overnight, diluted with CH₂Cl₂ (25 mL), and washed with a 5% aqueous solution of citric acid (10 mL), NaHCO₃ solution (10 mL), and brine. The organic phase was dried on anhydrous MgSO₄, filtered, and evaporated. The residue was purified by flash chromatography (CH₂Cl₂/MeOH, 100:0→95:5) to afford 48 mg (53%) of **31** [MS (ES, +) *m/z* 671 [M+Na]⁺]. This compound was used as such for the subsequent steps.

2-[(S)-2-Amino-3-methylbutanoyloxy]ethyl-[(2R,3S,5R)-5-(thymine-1-yl)-2-(3,3,3-triphenylpropanoyloxymethyl)tetrahydrofuran-3-yl] succinate (32): Following a similar procedure to that described for the synthesis of 3'-amino acid ester derivatives, reaction of **31** (80 mg, 0.12 mmol) with Boc-Val-OH (52 mg, 0.24 mmol), PyBOP (74 mg, 0.14 mmol), and Et₃N (42 μ L, 0.30 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield the coupling product [MS (ES, +) *m/z* 893 [M+Na]⁺]. Treatment with TFA and purification of the residue in the Chromatotron (CH₂Cl₂/MeOH, 10:1) yielded 33 mg (36%, two steps) of **32** as a white solid; mp: 63–65 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 771 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 0.86 (pseudo-t, *J* = 6.5 Hz, 6H, CH₃- γ), 1.72 (s, 3H, 5-CH₃), 1.90 (m, 1H, CH- β), 2.15 (m, 1H, H-2'), 2.30 (m, 1H, H-2''), 2.58 (m, 4H, OCC₂H₂CO), 3.20 (d, *J* = 5.0 Hz, 1H, CH- α), 3.85 (m, 2H, CH₂CO), 3.90–3.99 (m, 3H, H-4', H-5'), 4.20–4.33 (m, 4H, OCH₂CH₂O), 4.93 (m, 1H, H-3'), 6.10 (dd, *J* = 8.4, 6.1 Hz, 1H, H-1'), 7.14–7.27 (m, 15H, CPh₃), 7.43 (s, 1H, H-6), 11.36 ppm (s, 1H, 3-NH); Anal. calcd for C₄₂H₄₇N₃O₁₁: C 65.53, H 6.15, N 5.46, found: C 65.80, H 6.34, N 5.22.

2-[(S)-2-Amino-3-methylpentanoyloxy]ethyl-[(2R,3S,5R)-5-(thymine-1-yl)-2-(3,3,3-triphenylpropanoyloxymethyl)tetrahydrofuran-3-yl] succinate (33): As described for the synthesis of **32**, reaction of **31** (67 mg, 0.10 mmol) with Boc-Ile-OH (46 mg, 0.20 mmol), PyBOP (62 mg, 0.12 mmol) and Et₃N (34 μ L, 0.25 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/EtOAc, 6:4) to yield 57 mg of the coupling product [MS (ES, +) *m/z* 884 [M+1]⁺]. Treatment with TFA and purification of the residue in the Chromatotron (CH₂Cl₂/MeOH, 10:1) yielded 28 mg (35%, two steps) of **33** as a white solid; mp: 58–59 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 784 [M+1]⁺; ¹H NMR ([D₆]DMSO): δ = 0.79–0.83 (m, 6H, CH₃- γ , CH₃- δ), 1.14 (m, 1H, CH'- γ), 1.40 (m, 1H, CH''- γ), 1.58 (m, 1H, CH- β), 1.71 (s, 3H, 5-CH₃), 2.15 (m, 1H, H-2'), 2.30 (m, 1H, H-2''), 2.58 (m, 4H, OCC₂H₂CO), 3.24 (d, *J* = 5.4 Hz, 1H, CH- α), 3.78–4.00 (m, 5H, H-4', H-5', CH₂CO), 4.20–4.32 (m, 4H, OCH₂CH₂O), 4.92 (m, 1H, H-3'), 6.10 (dd, *J* = 8.2, 6.3 Hz, 1H, H-1'), 7.14–7.27 (m, 15H, CPh₃), 7.43 (s, 1H, H-6), 11.36 ppm (s, 1H, 3-NH); Anal. calcd for C₄₅H₄₉N₃O₁₁: C 65.89, H 6.30, N 5.36, found: 65.88, H 6.45, N 5.29.

5'-O-Trityl-3'-O-valylthymidine (2): Following the general procedure the synthesis of 3'-amino acid ester derivatives, reaction of 5'-O-tritylthymidine^[6] (**4**; 150 mg, 0.31 mmol), with Fmoc-Val-OH (210 mg, 0.62 mmol), PyBOP (193 mg, 0.37 mmol), and Et₃N (107 μ L, 0.77 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂) to yield the coupling product [MS (ES, +) *m/z* 829 [M+Na]⁺]. Treatment with piperidine (85 μ L) and removal of volatiles afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield 133 mg (74%, two steps) of **2** as a white solid; mp: 88–90 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 607 [M+Na]⁺; ¹H NMR ([D₆]DMSO): δ = 0.74 (d, *J* = 6.5 Hz, 3H, CH₃- γ), 0.82 (d, *J* = 6.5 Hz, 3H, CH₃- γ), 1.46 (s, 3H, 5-CH₃), 1.78 (m, 1H, CH- β), 2.30 (m, 2H, H-2'), 3.10 (d, *J* = 5.1 Hz, CH- α), 3.25 (m, 2H, H-5'), 4.02 (m, 1H, H-4'), 5.32 (m, 1H, H-3'), 6.21 (pseudo-t, *J* = 6.8 Hz, 1H, H-1'), 7.27–7.39 (m, 15H, CPh₃), 7.52 (s, 1H, H-6), 11.38 ppm (brs, 1H, 3-NH); Anal. calcd for C₃₄H₃₇N₃O₆: C 69.96, H 6.39, N 7.20, found: C 69.85, H 6.21, N 7.05.

3'-O-Isoleucyl-5'-O-tritylthymidine (34): Following the general procedure the synthesis of 3'-amino acid ester derivatives, reaction of 5'-O-tritylthymidine^[6] (**4**; 100 mg, 0.21 mmol), with Fmoc-Ile-OH (146 mg, 0.41 mmol), PyBOP (129 mg, 0.25 mmol), and Et₃N (72 μ L, 0.52 mmol) afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 30:1) to yield the coupling product [MS (ES, +) *m/z* 842 [M+Na]⁺]. Treatment with piperidine (72 μ L) followed by purification of the residue in the Chromatotron (CH₂Cl₂/MeOH,

10:1) yielded 81 mg (65%, two steps) of **34** as a white solid; mp: 84–86 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 620 [M+Na]⁺; ¹H NMR ([D₆]DMSO): δ = 0.76–0.87 (m, 6H, CH₃- γ , CH₃- δ), 1.03 (m, 1H, CH'- γ), 1.23 (m, 1H, CH''- γ), 1.47 (s, 3H, 5-CH₃), 1.53 (m, 1H, CH- β), 2.30–2.50 (m, 2H, H-2'), 3.15 (d, *J* = 5.5 Hz, 1H, CH- α), 3.26 (m, 2H, H-5'), 4.03 (m, 1H, H-4'), 5.30 (m, 1H, H-3'), 6.21 (pseudo-t, *J* = 7.0 Hz, 1H, H-1'), 7.27–7.38 (m, 15H, CPh₃), 7.52 (s, 1H, H-6), 11.37 ppm (brs, 1H, 3-NH); Anal. calcd for C₃₅H₃₉N₃O₆: C 70.33, H 6.58, N 7.03, found: C 70.12, H 6.57, N 6.98.

4-[(2R,3S,5R)-5-(Thymine-1-yl)-2-(trityloxymethyl)tetrahydrofuran-3-yl]-4-oxobutanoic acid (35): To a suspension of 5'-O-tritylthymidine^[6] (**4**; 150 mg, 0.31 mmol) in CH₂Cl₂ (7 mL), Et₃N (86 μ L, 0.62 mmol), DMAP (37 mg, 0.31 mmol), and succinic anhydride (62 mg, 0.62 mmol) were added. The reaction mixture was stirred at room temperature under argon for 7 h, diluted with CH₂Cl₂ (30 mL), and washed with a 10% aqueous solution of KH₂PO₄ (2 \times 15 mL), and H₂O (2 \times 15 mL). The organic phase was dried on anhydrous MgSO₄, filtered, and evaporated. The residue was purified in a SPE silica cartridge (CH₂Cl₂/MeOH, 300:1 \rightarrow 200:1) to afford 149 mg (82%) of **35** as a white solid; mp: 97–99 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 584 [M]⁺; ¹H NMR (CDCl₃): δ = 1.38 (s, 3H, 5-CH₃), 2.41 (m, 1H, H-2'), 2.51 (m, 1H, H-2''), 2.59–2.72 (m, 4H, OCH₂CH₂O), 3.42, 3.48 (dd, *J* = 10.6, 2.5 Hz, 2H, H-5'), 4.17 (m, 1H, H-4'), 5.46 (m, 1H, H-3'), 6.38 (dd, *J* = 9.0, 5.4 Hz, 1H, H-1'), 7.24–7.41 (m, 15H, CPh₃), 7.61 (s, 1H, H-6), 9.72 ppm (s, 1H, 3-NH).

2-Hydroxyethyl-[(2R,3S,5R)-5-(thymine-1-yl)-2-(trityloxymethyl)tetrahydrofuran-3-yl] succinate (36): To a solution of **35** (130 mg, 0.22 mmol) in CH₂Cl₂ (1 mL) was added ethylene glycol (50 μ L, 0.89 mmol), BOP (119 mg, 0.27 mmol), and Et₃N (37 μ L, 0.27 mmol). The reaction mixture was stirred at room temperature for 3 h, and diluted with EtOAc (25 mL), washed with a 5% aqueous solution of citric acid (10 mL), a NaHCO₃ solution (10 mL), and brine. The organic phase was dried on anhydrous MgSO₄, filtered, and evaporated. The residue was purified by CCTLC in the Chromatotron (CH₂Cl₂/MeOH, 20:1) to afford 128 mg (92%) of **36** as a white solid; mp: 73–74 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 651 [M+Na]⁺; ¹H NMR (CDCl₃): δ = 1.36 (d, *J* = 0.9 Hz, 3H, 5-CH₃), 2.47 (m, 2H, H-2'), 2.67 (m, 4H, COCH₂CH₂CO), 3.47 (d, *J* = 2.4 Hz, 2H, H-5'), 3.82 (m, 2H, CH₂OH), 4.14 (m, 1H, H-4'), 4.24 (m, 2H, OCH₂), 5.49 (m, 1H, H-3'), 6.40 (dd, *J* = 8.8, 5.7 Hz, 1H, H-1'), 7.24–7.40 (m, 15H, CPh₃), 7.60 (d, *J* = 1.1 Hz, 1H, H-6), 8.24 ppm (s, 1H, 3-NH).

2-[(S)-2-Amino-3-methylbutanoyloxy]ethyl-[(2R,3S,5R)-5-(thymine-1-yl)-2-(trityloximethyl)tetrahydrofuran-3-yl] succinate (37): Following a similar procedure to that described for the synthesis of 3'-amino acid ester derivatives, compound **36** (85 mg, 0.14 mmol) reacted with Fmoc-Val-OH (130 mg, 0.38 mmol), PyBOP (187 mg, 0.30 mmol), and Et₃N (37 μ L, 0.30 mmol). The crude was dissolved in DMF (1 mL) and treated with piperidine (72 μ L). Removal of volatiles afforded a residue that was purified in the Chromatotron (CH₂Cl₂/MeOH, 10:1) to yield 60 mg (96%) of **37** as a white solid; mp: 62–64 °C (CH₂Cl₂/MeOH); MS (ES, +) *m/z* 728 [M+1]⁺; ¹H NMR (CDCl₃): δ = 0.90 (d, *J* = 6.8 Hz, 3H, CH₃- γ), 0.97 (d, *J* = 6.8 Hz, 3H, CH₃- γ), 1.36 (s, 3H, 5-CH₃), 2.01 (m, 1H, CH- β), 2.45 (m, 2H, H-2'), 2.67 (m, 4H, COCH₂CH₂CO), 3.33 (d, *J* = 4.9 Hz, 1H, CH- α), 3.47 (d, *J* = 2.2 Hz, 2H, H-5'), 4.13 (m, 1H, H-4'), 4.32 (m, 4H, CH₂O), 5.50 (m, 1H, H-3'), 6.40 (dd, *J* = 8.2, 6.2 Hz, 1H, H-1'), 7.25–7.43 (m, 15H, CPh₃), 7.58 ppm (s, 1H, H-6); Anal. calcd for C₄₀H₄₅N₃O₁₀ · H₂O: C 64.42, H 6.35, N 5.63, found: C 64.18, H 6.03, N 5.70.

2-[(S)-2-Amino-3-methylpentanoyloxy]ethyl-[(2R,3S,5R)-5-(thymine-1-yl)-2-(trityloximethyl)tetrahydrofuran-3-yl] succinate (38): Following a reaction sequence similar to that described for

37, alcohol **36** (110 mg, 0.17 mmol) reacted with Fmoc-Ile-OH (93 mg, 0.26 mmol), PyBOP (109 mg, 0.21 mmol), and Et₃N (30 μ L, 0.21 mmol), and was deprotected with piperidine (85 μ L). The residue was purified by HPFC (CH₂Cl₂/MeOH, 100:0→96:4) to yield 91 mg (71%) of **38**; mp: 57–59 °C; MS (ES, +) *m/z* 742 [*M*+1]⁺; ¹H NMR (CDCl₃): δ =0.77–0.85 (m, 6H, CH₃- γ , CH₃- δ), 1.08 (m, 1H, CH'- γ), 1.36 (m, 1H, CH''- γ), 1.43 (s, 3H, 5-CH₃), 1.55 (m, 1H, CH- β), 2.31–2.50 (m, 2H, H-2'), 2.57 (m, 4H, COCH₂CH₂CO), 3.14–3.19 (m, 2H, H-5', CH- α), 3.30 (m, 1H, H-5''), 4.05 (m, 1H, H-4'), 4.16–4.29 (m, 4H, CH₂O), 5.30 (m, 1H, H-3'), 6.40 (dd, *J*=8.3, 6.2 Hz, 1H, H-1'), 7.27–7.39 (m, 15H, CPh₃), 7.50 (s, 1H, H-6), 11.36 ppm (brs, 1H, NH); Anal. calcd for C₄₁H₄₇N₃O₁₀: C 66.38, H 6.39, N 5.66, found: C 66.09, H 6.60, N 5.91.

Biology

Cells and culture conditions: *L. infantum* promastigotes (MCAN/ES/89/IPZ229/1/89) were grown in RPMI-1640 medium (Sigma–Aldrich, St. Louis, MO, USA) supplemented with 10% heat-inactivated fetal calf serum (FCS), antibiotics, and 25 mM HEPES (pH 7.2) at 26 °C. *L. infantum* axenic amastigotes were grown in M199 medium (Invitrogen, Leiden, The Netherlands) supplemented with 10% heat-inactivated FCS, 1 g L⁻¹ β -alanine, 100 mg L⁻¹ L-asparagine, 200 mg L⁻¹ sucrose, 50 mg L⁻¹ sodium pyruvate, 320 mg L⁻¹ malic acid, 40 mg L⁻¹ fumaric acid, 70 mg L⁻¹ succinic acid, 200 mg L⁻¹ α -ketoglutaric acid, 300 mg L⁻¹ citric acid, 1.1 g L⁻¹ sodium bicarbonate, 5 g L⁻¹ MES, 0.4 mg L⁻¹ hemin, and 10 mg L⁻¹ gentamicin, pH 5.4 at 37 °C. Jurkat cells were grown in RPMI-1640 medium (Sigma–Aldrich) supplemented with 10% heat-inactivated FCS, antibiotics, and 10 mM HEPES, pH 7.2 at 37 °C and 5% CO₂. THP-1 cells were grown in RPMI-1640 medium (Gibco, Leiden, The Netherlands) supplemented with 10% heat-inactivated FCS, antibiotics, 1 mM HEPES, 2 mM glutamine, and 1 mM sodium pyruvate, pH 7.2 at 37 °C and 5% CO₂.

Cell proliferation assays: Drug treatment of promastigotes and amastigotes was performed during logarithmic growth phase at a concentration of 2 \times 10⁶ parasites mL⁻¹ at 26 °C or 1 \times 10⁶ parasites mL⁻¹ at 37 °C for 24 h, respectively. Drug treatment of Jurkat and THP-1 cells was performed during the logarithmic growth phase at a concentration of 4 \times 10⁵ cells mL⁻¹ at 37 °C and 5% CO₂ for 24 h. The percentage of living cells was evaluated by flow cytometry by the propidium iodide (PI) exclusion method.^[16] Briefly, treated parasites were stained for 10 min with 5 μ g mL⁻¹ PI. The number of PI-positive parasites was determined in a Beckman–Coulter FC500 flow cytometer.

Leishmania infection assays: THP-1 cells were seeded at 120 000 cells mL⁻¹ in 24-multiwell plates (Nunc, Roskilde, Denmark) and differentiated to macrophages for 24 h in 1 mL RPMI-1640 medium containing 10 ng mL⁻¹ phorbol 12-myristate 13-acetate (PMA, Sigma–Aldrich). The culture medium was removed, and *L. infantum* amastigotes (1.2 \times 10⁶ in 1 mL THP-1 medium) were added to each well; 3 h later, all medium with non-infecting amastigotes was removed, washed three times with 1 \times phosphate-buffered saline (PBS) and replaced with new THP-1 medium and the corresponding treatment. After 48 h of treatment, medium was removed; THP-1 cells were washed three times with 1 \times PBS and detached with TrypLE Express (Invitrogen, Leiden, The Netherlands) according to the manufacturer's indications. Infection was measured by flow cytometry.

Cytotoxicity assays in human cells: Drug treatment of Jurkat cells was performed during logarithmic growth phase at a concentration of 4 \times 10⁵ cells mL⁻¹ at 37 °C and 5% CO₂ for 24 h. The percentage

of living cells was evaluated by flow cytometry by the propidium iodide (PI) exclusion method.

Hydrolysis studies in cell extracts: Cell extracts from PMA-differentiated THP-1 were prepared according to the procedure described by Saboulard et al.^[26] Hydrolysis studies of **28** were performed as follows: Aliquots of 200 μ L of compound **28** (100 μ M) were mixed with 10 μ L cell extract, and the samples were incubated at 37 °C. At various time points (0, 0.5, 1, 2, 4, and 24 h), MeOH (100 μ L) was added to inactivate the enzymes. These samples were centrifuged, and the supernatants were injected on HPLC. Identification of the hydrolysis products was based on the retention times of the reference compounds and co-elution experiments under identical analytical conditions.

HPLC analysis of hydrolysis products: Analysis of the hydrolysis products after cell extract incubation was performed by HPLC injecting 30 μ L of the supernatant. HPLC spectra were recorded on an Agilent 1120 Compact LC using a diode array detector (λ 230–400 nm), and an analytical ACE5 C₁₈-300 column (150 \times 4.6 mm, 300 \AA). Solvents used were CH₃CN (0.05% TFA) for A, and H₂O (0.05% TFA) for B, and the flow rate was 1 mL min⁻¹. The gradient used was 30→60% A over 20 min. The retention times of the analytes were: 16.8 min for compound **28**, 14.2 min for compound **6**, and 13.5 min for compound **1**.

Mitochondrial transmembrane potential: Parasites were incubated in RPMI medium with 10% FCS at 1.5 \times 10⁶ mL with 50 nM tetramethylrhodamine methyl ester perchlorate (TMRM, Sigma Cat. #T5428) probe for 15 min at 26 °C and then analyzed by flow cytometry. A positive control of mitochondrial depolarization was obtained in parallel by incubating logarithmic promastigotes at 1.5 \times 10⁶ mL with the protonophore (H⁺ ionophore) and uncoupler of oxidative phosphorylation in mitochondria, carbonyl cyanide 3-chlorophenylhydrazone (CCCP) during the indicated times and then stained with TMRM as described above.

Mitochondrial superoxide detection: Parasites were preloaded for 10 min with 2 μ M MitoSOX, and then fluorescence emission was measured at the indicated times in a Beckman Coulter FC500 flow cytometer.

DNA content analysis by flow cytometry: Parasites (1.5 \times 10⁶) were centrifuged at 1000 *g* for 5 min, the pellet was resuspended in 100 μ L PBS, cold (–20 °C) 70% EtOH (600 μ L) was added, and then parasites were incubated on ice for 30 min. After incubation, the parasites were washed with 1 mL PBS/50 mM EDTA, pelleted at 1000 *g*, and then resuspended in 400 μ L PBS/50 mM EDTA/50 μ g mL⁻¹ RNase and incubated for 30 min at 37 °C. Propidium iodide (PI) was then added to a final concentration of 5 μ g mL⁻¹, and the parasites were immediately analyzed for PI fluorescence in a Beckman Coulter FC500 flow cytometer.

LiEndoG activity: The nuclease activity of LiEndoG in the presence of selected compounds was monitored by measuring the increase in fluorescence derived from the digestion of a double-stranded probe. 6-Carboxyfluorescein (FAM)-derived fluorescence is quenched by the proximity of tetramethylrhodamine (TAMRA) in the undigested probe constructed by hybridization of the oligonucleotides FAM-5'-CTG TCG CTA CCT GTG G-3'-TAMRA and FAM-5'-CCA CAG GTA GCG ACA G-3'-TAMRA. Digestion of any of the two oligonucleotides causes separation of the fluorophore and quencher, and then a fluorescent signal can be emitted; 30 pmol of the double-stranded probe were digested with LiEndoG 2.5 ng μ L⁻¹. After digestion, both fluorophores separate, and FAM-derived fluorescence can be detected. Reactions were monitored in a Victor

1420 Multilabel Counter (Wallac) at an excitation wavelength of 492 nm and an emission wavelength of 517 nm.

Acknowledgements

E.C. thanks the Comunidad de Madrid and the Fondo Social Europeo (FSE) for a predoctoral fellowship. A.G. has a JAE predoctoral fellowship financed by the CSIC and FSE. We thank Ms. Maria Nares and Mr. Kilian Gutierrez for excellent technical assistance. Financial support to Ms. M. Nares by the Spanish MEC and FSE is gratefully acknowledged. This work was supported by a grant from the Spanish CICYT (SAF2009-13914-C02-01 and SAF2009-13914-C02-02). The authors also acknowledge financial support from Viceconsejería de Ciencia y Tecnología, Junta de Comunidades de Castilla-La Mancha (grant POII10-0180-7897).

Keywords: amino acids · endonucleases · leishmania · thymidine · trityl groups

- [1] P. Desjeux, *Comp. Immunol. Microbiol. Infect. Dis.* **2004**, *27*, 305–318.
- [2] F. Chappuis, S. Sundar, A. Hailu, H. Ghalib, S. Rijal, R. W. Peeling, J. Alvar, M. Boelaert, *Nat. Rev. Microbiol.* **2007**, *5*, 873–882.
- [3] T. S. Tiuman, A. O. Santos, T. Ueda-Nakamura, B. P. Dias, C. V. Nakamura, *Int. J. Infect. Dis.* **2011**, *15*, e525–e532.
- [4] S. L. Croft, S. Sundar, A. H. Fairlamb, *Clin. Microbiol. Rev.* **2006**, *19*, 111–126.
- [5] M. Berg, P. Van der Veken, A. Goeminne, A. Haemers, K. Augustyns, *Curr. Med. Chem.* **2010**, *17*, 2456–2481.
- [6] "Purine and Pyrimidine Metabolism in *Leishmania* in Drug Targets in Kinetoplastid Parasites" N. S. Carter, P. Yates, C. S. Arendt, J. M. Boitz, B. Ullman in *Book Series: Advances in Experimental Medicine and Biology*, Vol. 625, Springer, Berlin, **2008**, pp. 141–154.
- [7] C. Peyron, R. Benhida, C. Bories, P. M. Loiseau, *Bioorg. Chem.* **2005**, *33*, 439–447.
- [8] P. F. Torrence, X. S. Fan, X. Y. Zhang, P. M. Loiseau, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5047–5051.
- [9] X. S. Fan, X. Y. Zhang, C. Bories, P. M. Loiseau, P. F. Torrence, *Bioorg. Chem.* **2007**, *35*, 121–136.
- [10] X. S. Fan, D. Feng, Y. Y. Qu, X. Y. Zhang, J. J. Wang, P. M. Loiseau, G. Andrei, R. Snoeck, E. De Clercq, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 809–813.
- [11] M. Thiel, S. Harder, M. Wiese, M. Kroemer, I. Bruchhaus, *Mol. Biochem. Parasitol.* **2008**, *158*, 152–162.
- [12] C. Nguyen, G. Kasinathan, I. Leal-Cortijo, A. Musso-Buendia, M. Kaiser, R. Brun, L. M. Ruiz-Perez, N. G. Johansson, D. Gonzalez-Pacanoska, I. H. Gilbert, *J. Med. Chem.* **2005**, *48*, 5942–5954.
- [13] C. N. Tetzlaff, I. Schwoppe, C. F. Blecinski, J. A. Steinberg, C. Richert, *Tetrahedron Lett.* **1998**, *39*, 4215–4218.
- [14] U. Eriksson, J. M. Hilfinger, J. S. Kim, S. Mitchell, P. Kijek, K. Z. Borysko, J. M. Breitenbach, J. C. Drach, B. A. Kashemirov, C. E. McKenna, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 583–586.
- [15] "5'-O-Tritylthymidine, Selective Tritylation of a Pyrimidine Nucleoside": R. H. Munson, Jr. in *Synthetic Procedures in Nucleic Acid Chemistry* (Eds.: W. W. Zorbach, R. S. Tipson) Wiley, New York, **1968**; pp. 321–322.
- [16] D. Plano, Y. Baquedano, D. Moreno-Mateos, M. Font, A. Jiménez-Ruiz, J. A. Palop, C. Sanmartin, *Eur. J. Med. Chem.* **2011**, *46*, 3315–3323.
- [17] J. F. Alzate, A. A. Barrientos, V. M. González, A. Jiménez-Ruiz, *Apoptosis* **2006**, *11*, 161–171.
- [18] E. Rico, J. F. Alzate, A. A. Arias, D. Moreno, J. Clos, F. Gago, I. Moreno, M. Dominguez, A. Jiménez-Ruiz, *Mol. Biochem. Parasitol.* **2009**, *163*, 28–38.
- [19] E. Rico, J. F. Alzate, C. Oliva, C. M. Genes, F. Gago, A. Jiménez-Ruiz, unpublished results.
- [20] J. Parrish, L. Li, K. Klotz, D. Ledwich, X. Wang, D. Xue, *Nature* **2001**, *412*, 90–94.
- [21] S. Büttner, D. Carmona-Gutierrez, I. Vitale, M. Castedo, D. Ruli, T. Eisenberg, G. Kroemer, F. Madeo, *Cell Cycle* **2007**, *6*, 1072–1076.
- [22] S. Büttner, T. Eisenberg, D. Carmona-Gutierrez, D. Ruli, H. Knauer, C. Ruckenstuhl, C. Sigrist, S. Wissing, M. Kollroser, K. U. Fröhlich, S. Sigrist, F. Madeo, *Mol. Cell* **2007**, *25*, 233–246.
- [23] C. McDermott-Roe, J. Ye, R. Ahmed, X.-M. Sun, A. Serafin, J. Ware, L. Botto, P. Muckett, X. Canas, J. Zhang, G. C. Rowe, R. Buchan, H. Lu, A. Braithwaite, M. Mancini, D. Hauton, R. Marti, E. Garcia-Arumi, N. Hubner, H. Jacob, T. Serikawa, V. Zidek, F. Papousek, F. Kolar, M. Cardona, M. Ruiz-Meana, D. Garcia-Dorado, J. X. Comella, L. E. Felkin, P. J. R. Barton, Z. Arany, M. Pravenec, E. Petretto, D. Sanchis, S. A. Cook, *Nature* **2011**, *478*, 114–118.
- [24] D. Yoo, J. S. Oh, D.-W. Lee, Y. G. Kim, *J. Org. Chem.* **2003**, *68*, 2979–2982.
- [25] P. Luo, J. C. Leitzel, Z.-Y. J. Zhan, D. G. Lynn, *J. Am. Chem. Soc.* **1998**, *120*, 3019–3031.
- [26] D. Saboulard, L. Naesens, D. Cahard, A. Salgado, R. Pathirana, S. Velazquez, C. McGuigan, E. De Clercq, J. Balzarini, *Mol. Pharmacol.* **1999**, *56*, 693–704.

Received: March 21, 2013

Published online on ■■■■■, 0000

4.2 La EndoG de *Leishmania infantum* es una endo/exonucleasa esencial para la supervivencia del parásito

4.2.1 Hipótesis

La EndoG de *Leishmania infantum*, al igual que sus homólogas en otros organismos, podría desempeñar un papel provida y ser necesaria para el correcto desarrollo de los parásitos.

4.2.2 Objetivos

- Caracterizar si el tipo de actividad nucleasa de *LiEndoG* es similar al de aquellas homólogas para las que se ha demostrado un papel provida.
- Evaluar el efecto que produce sobre los parásitos la eliminación del gen que codifica para esta enzima.
- Confirmar el papel provida de *LiEndoG*.

4.2.3 Resultados principales

Como sabemos, EndoG pertenece a la superfamilia de nucleasas no específicas con dominio $\beta\beta\alpha$ de unión a metal, las cuales se han encontrado en numerosos organismos, incluyendo los tripanosomátidos. En estos parásitos, se ha demostrado que esta nucleasa participa en la ejecución del programa apoptótico migrando de la mitocondria al núcleo para degradar el ADN genómico, paso característico de este tipo de muerte celular (155). En el trabajo que exponemos a continuación demostramos que la Endonucleasa G de *Leishmania infantum* (*LiEndoG*) es una endo-exonucleasa que tiene actividad preferencial 5'-exonucleasa sobre el ADN bicatenario lineal. Este tipo de actividad, que también ha sido descrito en otros miembros de esta familia como Nuc1p o EXOG, se ha asociado con la capacidad de estas enzimas de participar en la recombinación y reparación del ADN mitocondrial, indicando que *LiEndoG* podría estar también implicada en este tipo de procesos. Adicionalmente, a parte del papel de *LiEndoG* durante la muerte celular apoptótica, en este trabajo mostramos que esta enzima parece ser necesaria para el desarrollo normal de estos parásitos, dado que los parásitos *hemi-knockouts* para EndoG presentan una tasa de crecimiento significativamente menor que los parásitos control, así como una infectividad sobre células THP-1 diferenciadas

mucho menor. Finalmente, el papel provida de *LiEndoG* se puede corroborar por la mayor supervivencia observada en parásitos que sobreexpresan esta proteína cuando son tratados con el inhibidor de *LiEndoG* **Lei49** (denominado compuesto 37 en el artículo 1 de esta memoria (156)) . En conjunto, nuestros resultados demuestran que *LiEndoG* juega un papel dual en estos parásitos, siendo fundamental tanto para la supervivencia como para la muerte en *Leishmania*.

4.2.4 Resumen

Hasta el inicio de esta tesis, nuestro grupo había logrado caracterizar *LiEndoG* como una nucleasa involucrada en la degradación del ADN durante el proceso de muerte celular de *L. infantum* (101). Sin embargo, nada sabíamos sobre su posible papel provida, si bien suponíamos que podría estar involucrada en alguno de los múltiples procesos fisiológicos que se han propuesto para sus homólogas en otros organismos. En este sentido, el hecho de que el grupo de derivados de timidina descritos en el artículo 1 pareciera ejercer su actividad leishmanicida mediante la inhibición de *LiEndoG* nos motivó a seguir investigando la posible implicación de esta proteína en la supervivencia de estos organismos. Los resultados de esta investigación dieron como fruto la publicación del artículo 2 (*Leishmania infantum* EndoG is an endo/exo-nuclease essential for parasite survival; (168)), recogida en el apartado 4.2.5 de esta tesis y los cuales resumimos a continuación.

En primer lugar, decidimos caracterizar qué tipo de actividad nucleasa poseía *LiEndoG* porque, como se comentó en la introducción de esta tesis, gran parte de los papeles fisiológicos que se han asignado a los miembros de la familia EndoG dependen del tipo de actividad que presentan. En este sentido, el análisis del patrón de digestión generado por *LiEndoG* muestra como su actividad inicial sobre plásmidos superenrollados da lugar a la aparición de moléculas circulares abiertas (artículo 2, figura 1). Este resultado sugiere que *LiEndoG* no es capaz de generar cortes lo suficientemente próximos en ambas hebras como para que se produzca la linealización del plásmido (como hacen por ejemplo las enzimas de restricción), si no que corta de forma aleatoria cada una de las hebras del ADN. Por ello, la linealización del plásmido solo se llega a observar cuando se han generado suficientes cortes independientes en cada una de las hebras como para que no se pueda mantener el apareamiento de las bases. Por tanto, podemos concluir que *LiEndoG* genera

preferentemente cortes puntuales de cadena sencilla sobre el ADN de doble cadena (ADNdc) superenrollado. Este resultado está en consonancia con el tipo de actividad que se ha descrito para la EndoG de *Drosophila melanogaster*. La estructura cristalina de esta enzima dimérica muestra como cada uno de los dos centros activos se disponen de forma distante y orientados en sentidos opuestos, haciéndoles incapaces de procesar oligonucleótidos de doble cadena de forma simultánea sin la existencia de una gran curvatura del ADN (123). Esta misma conclusión se puede extraer de nuestro nuevo modelo molecular para la *Li*EndoG dimérica, ya que la disposición de los dos centros activos del homodímero es similar a la que acabamos de describir para *Dm*EndoG (artículo 3, figura 3) (171).

Como sabemos, también existen diferencias respecto al tipo de actividad endo o exonucleasa que tienen los distintos miembros de la familia EndoG. Por ejemplo, mientras que las EndoGs de metazoos son exclusivamente endonucleasas, las nucleasas ortólogas a *Li*EndoG en *S. cerevisiae* (Nuc1p) y *Neurospora* o las EXOG de metazoos poseen actividad 5'-exonucleasa además de su característica actividad endonucleasa (116,117). Por ello, nuestro siguiente paso fue determinar en cuál de estos dos grupos podíamos clasificar nuestra proteína. Nuestros resultados muestran que *Li*EndoG tiene una clara preferencia por generar cortes cercanos al extremo 5' cuando digiere una molécula lineal de ADNdc (artículo 2, figura 2), demostrando que esta enzima podría ser una endo-exonucleasa. No obstante, es importante mencionar que las limitaciones de la técnica empleada para este experimento no nos permitieron determinar el tamaño exacto de los oligonucleótidos generados como producto de esta digestión.

En conjunto, los resultados anteriores nos permitieron demostrar que *Li*EndoG es una endo-exonucleasa capaz de generar cortes de cadena sencilla sobre el ADNdc superenrollado circular, mientras que presenta preferentemente una actividad tipo 5'-exonucleasa sobre ADNdc lineal. Este tipo de actividad coincide con el descrito para Nuc1p y EXOG y se ha asociado con su capacidad para participar en la recombinación y reparación del ADN mitocondrial, respectivamente (137,148). Por ello, es plausible pensar que *Li*EndoG pueda estar ejerciendo funciones similares en *Leishmania infantum*.

En segundo lugar, quisimos evaluar si la eliminación de *Li*EndoG podría tener algún efecto sobre los parásitos. Para ello, se reemplazó uno de los alelos codificantes para *Li*EndoG por un gen de resistencia a antibiótico, obteniéndose con ello promastigotes *hemi-knockouts* para

EndoG (*LiEndoG*^{-/+}). Nuestros resultados muestran que estos parásitos, en los que se había conseguido reducir significativamente la expresión de *LiEndoG*, presentan un crecimiento significativamente más lento que el de los controles en condiciones de cultivo normales, confirmando que *LiEndoG* es esencial para el desarrollo normal de los promastigotes de *L. infantum* (artículo 2, figura 3). Este hallazgo está además en consonancia con los reportados para otros organismos. Por ejemplo, se ha descrito que la reducción de los niveles de expresión de CPS-6 mediante ARNi en *C. elegans* produce un retraso del crecimiento celular durante el desarrollo de este nemátodo (129). De forma similar, se ha demostrado que la delección de Nuc1p en *S. cerevisiae* produce una disminución en su tasa de crecimiento, efecto que se atribuyó a la implicación de esta proteína en la recombinación del ADNmt (131).

Adicionalmente, hemos demostrado que los parásitos deficientes en *LiEndoG* son mucho menos infectivos para los macrófagos, ya que tanto el número de células infectadas como la carga parasitaria en el interior de estas es significativamente menor cuando la infección es llevada a cabo con los parásitos *LiEndoG*^{-/+} (artículo 2, figuras 4b y c). Por tanto, podemos concluir que los parásitos con un nivel de expresión de *LiEndoG* disminuido tienen mermada su capacidad para infectar macrófagos, así como la de proliferar y/o sobrevivir una vez se encuentran infectando los mismos, demostrando que esta proteína también juega un papel fundamental en los amastigotes intracelulares.

Finalmente, dada la demostrada localización mitocondrial de *LiEndoG* (101), decidimos evaluar si la reducción en los niveles de expresión de esta proteína produce algún efecto sobre el funcionamiento de este orgánulo. En este sentido, comprobamos que la deficiencia de *LiEndoG* genera un aumento en el potencial de membrana mitocondrial de los parásitos (artículo 2, figura 4d). Este resultado puede parecer *a priori* contradictorio, ya que cabría esperar que la deficiencia de *LiEndoG* produjese deterioro de la actividad mitocondrial. Sin embargo, se había demostrado previamente que los daños en el ADN mitocondrial en *L. donovani* producían un aumento de su $\Delta\psi_m$ (92), por lo que es posible hipotetizar que la hiperpolarización mitocondrial observada sea debida a la acumulación de este tipo de daños como consecuencia de la deficiencia de *LiEndoG*.

Por último, para terminar de comprobar el posible papel dual de *LiEndoG* en la supervivencia y muerte de los parásitos, utilizamos cultivos de promastigotes que sobreexpresan de forma constitutiva esta nucleasa. Nuestro grupo ya había demostrado previamente que la

sobreexpresión de *LiEndoG* no tiene ningún efecto sobre el crecimiento de los parásitos en condiciones normales, pero si produce un aumento del porcentaje de parásitos con el ADN degradado (TUNEL positivos) cuando estos se tratan con el estímulo apoptótico edelfosina (101). Por ello, decidimos comprobar si los inhibidores de *LiEndoG* que hemos descrito en el apartado anterior eran capaces de revertir la “sensibilización” a la muerte por estímulos apoptóticos que estaba provocando la sobreexpresión de esta nucleasa. Para este fin decidimos seleccionar el **compuesto 37**, denominado **Lei49** en el artículo que estamos discutiendo (168), porque tenía un potente efecto inhibitorio sobre la actividad de *LiEndoG* *in vitro*, una buena actividad leishmanicida frente a promastigotes y amastigotes de *L. infantum* y una citotoxicidad baja. Los resultados de nuestros experimentos muestran que los parásitos que sobreexpresan *LiEndoG* están protegidos frente a la muerte producida por concentraciones moderadas del **Lei49** pero, como cabía esperar, este efecto desaparece a medida que se aumentamos la concentración de la molécula (artículo 2, figura 6b). De forma contraria, y tal y como cabía esperar por nuestros resultados previos, estos mismos parásitos eran más susceptibles a la muerte producida por edelfosina que los parásitos control, probablemente como consecuencia de tener su maquinaria molecular pro-apoptótica aumentada (artículo 2, figura 6c). En su conjunto, los resultados presentados en el artículo 2 demuestran que *LiEndoG* es, al menos, una de las dianas sobre las que actúa el **Lei49** en el interior de los parásitos, siendo esta afirmación probablemente extensible al resto de derivados de timidina con actividad leishmanicida.

En su conjunto, nuestros resultados indican que *LiEndoG* es importante tanto para la supervivencia como para la ejecución de la muerte en *L. infantum*, tal y como se ha descrito también para sus ortólogos en *S. cerevisiae* o *C. elegans* (129,131). No obstante, sería necesario determinar cuáles son las funciones fisiológicas en las que está implicada *LiEndoG* y que le otorgan este papel provida. En este sentido, se ha sugerido que las EndoGs de *S. cerevisiae* y *Neurospora* podrían estar iniciando el proceso de recombinación homóloga mediante la generación de regiones de cadena sencilla en los cromosomas gracias a su actividad 5'-exonucleasa (109,137). Así mismo, se ha sugerido que este mismo tipo de actividad podría permitir a EXOG ampliar los cortes de cadena sencilla producidos en el ADNmt como consecuencia del daño oxidativo para facilitar su reparación (147,148). Nuestros resultados muestran que *LiEndoG*, al igual que las proteínas mencionadas

anteriormente, tiene actividad tipo 5'-exonucleasa, por lo que no sería descabellado que también pudiese estar involucrada en la recombinación y reparación del ADNmt. No obstante, sería necesario caracterizar de forma más detallada el tipo de actividad nucleasa de *LiEndoG*, tarea que se abordó en el tercer artículo presentado en esta memoria. Por último, podemos afirmar que el hecho de que *LiEndoG* esté involucrada en la supervivencia de estos parásitos confirma que esta proteína es una potencial diana terapéutica hacia la cual se podrían desarrollar nuevos compuestos con actividad leishmanicida.

Leishmania infantum EndoG Is an Endo/Exo-Nuclease Essential for Parasite Survival

Eva Rico^{1,2}, Cristina Oliva^{1,3}, Kilian Jesús Gutierrez¹, Juan Fernando Alzate², Carlos Mario Genes¹, David Moreno¹, Elena Casanova³, Alba Gigante³, María-Jesús Pérez-Pérez³, María-José Camarasa³, Joachim Clos⁴, Federico Gago⁵, Antonio Jiménez-Ruiz^{1*}

1 Departamento de Biología de Sistemas-Unidad Asociada al Consejo Superior de Investigaciones Científicas (CSIC), Universidad de Alcalá, Alcalá de Henares, Madrid, Spain, **2** Departamento de Microbiología y Parasitología, Facultad de Medicina, Universidad de Antioquia, Medellín, Colombia, **3** Instituto de Química Médica - Consejo Superior de Investigaciones Científicas (IQM-CSIC), Madrid, Spain, **4** Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany, **5** Departamento de Ciencias Biomédicas - Unidad Asociada al Consejo Superior de Investigaciones Científicas (CSIC), Alcalá de Henares, Madrid, Spain

Abstract

EndoG, a member of the DNA/RNA non-specific $\beta\beta\alpha$ -metal family of nucleases, has been demonstrated to be present in many organisms, including Trypanosomatids. This nuclease participates in the apoptotic program in these parasites by migrating from the mitochondrion to the nucleus, where it takes part in the degradation of genomic DNA that characterizes this process. We now demonstrate that *Leishmania infantum* EndoG (LiEndoG) is an endo-exonuclease that has a preferential 5' exonuclease activity on linear DNA. Regardless of its role during apoptotic cell death, this enzyme seems to be necessary during normal development of the parasites as indicated by the reduced growth rates observed in LiEndoG hemi-knockouts and their poor infectivity in differentiated THP-1 cells. The pro-life role of this protein is also corroborated by the higher survival rates of parasites that over-express this protein after treatment with the LiEndoG inhibitor Lei49. Taken together, our results demonstrate that this enzyme plays essential roles in both survival and death of *Leishmania* parasites.

Citation: Rico E, Oliva C, Gutierrez KJ, Alzate JF, Genes CM, et al. (2014) *Leishmania infantum* EndoG Is an Endo/Exo-Nuclease Essential for Parasite Survival. PLOS ONE 9(2): e89526. doi:10.1371/journal.pone.0089526

Editor: Ger van Zandbergen, Federal Institute for Vaccines and Biomedicines, Germany

Received: July 26, 2013; **Accepted:** January 21, 2014; **Published:** February 26, 2014

Copyright: © 2014 Rico et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Sources of funding: Spanish MEC/MICINN Project SAF2009-13914-C02; Comunidad de Madrid Project BIPEDD-2-CM ref S-2010/BMD-2457; Junta de Comunidades de Castilla la Mancha Project POI10-0180-7897. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: antonio.jimenez@uah.es

☛ These authors contributed equally to this work.

Introduction

Members of the Endonuclease G (EndoG) family have been found in all organisms whose genomes have been fully sequenced. One of the most important roles described for these enzymes is their participation in the apoptotic process whereby EndoGs translocate to the nucleus and contribute to the degradation of genomic DNA into oligonucleosomal fragments [1]. All of these proteins belong to the superfamily of non-specific $\beta\beta\alpha$ -metallonucleases [2] and their nuclease activity depends on the presence of divalent cations such as Mg^{2+} , Mn^{2+} or Co^{2+} whereas it is inhibited by moderate concentrations of monovalent cations such as K^+ or Na^+ [2,3,4]. EndoGs are encoded by a nuclear gene and imported into the mitochondrion as a consequence of the presence of an N-terminal signal peptide [5]. Mitochondrial localization is essential because its ectopic expression in the cytosol induces cell death [2].

Despite the similarities among EndoG family members, their nuclease activities are not identical: yeast and *Neurospora* enzymes contain an intrinsic 5' exonuclease activity that is absent in their mammalian counterparts [6]. In fact, mammalian mitochondria also contain EXOG, an endo-exonuclease that is considered to have originated from the duplication of an ancestral nuclease gene

that generated the paralogous EndoG- and EXOG-protein subfamilies in higher eukaryotes. This way, the full endo/exonuclease activities found in mitochondria of lower eukaryotes were maintained [7].

Mammalian members of this family of nucleases attack the nucleotide sequence in duplex DNA in a highly nonrandom fashion and show a particularly strong, but not exclusive, preference for nicking at positions adjacent to guanines [8]. This striking preference to attack guanine tracts is the hallmark of the endonuclease activity of the mammalian members of this family and the reason why these enzymes were named Endonucleases G. In contrast, the yeast endo-exonuclease has no preference to nick specific sequences in double-stranded DNA (dsDNA) [6].

Apart from their involvement in the death process, a pro-life role for these nucleases has also been suggested in several organisms. Thus, reduction in the levels of the *C. elegans* EndoG ortholog *cps-6* by RNA interference (RNAi) delays cellular growth during development of this worm [1]. Similarly, deletion of the yeast EndoG ortholog *Nuc1p* sensitized cells to cell death using standard culture conditions but, on the contrary, diminished apoptotic death when mitochondrial respiration was increased. These findings point to a dual (pro-life and pro-death) role [9].

In this work we show that *Leishmania infantum* EndoG (LiEndoG) not only displays endonuclease activity but, similarly to its counterparts from yeast and *Neurospora*, also functions as a 5'-exonuclease. In an attempt to evaluate its possible pro-life role, we used a gene replacement strategy to diminish endogenous LiEndoG levels. This decrease reduced the growth rate of the parasites and limited their capacity to infect and survive inside differentiated THP-1 cells. Moreover, results from assays on control and LiEndoG-transfected parasites using a recently described LiEndoG inhibitor support the vital function of this protein.

Materials and Methods

Cells and culture conditions

L. infantum promastigotes (M/CAN/ES/96/BCN150 MON-1), kindly provided by Dr. Alonso (CBMSO- Universidad Autónoma Madrid, Spain), were grown in RPMI-1640 medium (Gibco, Paisley, UK) supplemented with 10% heat-inactivated fetal calf serum (FCS), antibiotics, and 25 mM HEPES at 26°C. Parasite death was induced by addition of edelfosine (Calbiochem) or **Lei49** (2-[(S)-2-Amino-3-methylbutanoyl-oxy]ethyl [(2R,3S,5R)-5-(thymine-1-yl)-2-(trityloximethyl)tetrahydrofuran-3-yl] succinate) to the culture medium [10]. Growth curves were plotted as the product of cell concentration versus time. The graph shows the log₂ values of the parasite concentration. Parasites were kept at the log phase by daily dilution to 1.1×10^6 – 1.5×10^6 parasites/mL and successive dilutions were considered to calculate the number of parasites.

Sequence cloning and construct design

DNA fragments containing the *LiEndoG* 5'UTR were obtained after amplification of genomic DNA with 5'GGGGAATTCT-GAACATCACAGTGTGGAG3' and 5'CTCATGGGTACCT-GATGAATCGTTTTTCGCGCA3' primers (underlined sequences highlight the EcoRI and the KpnI restriction sites respectively). DNA fragments containing the LiEndoG 3'UTR were obtained after amplification of genomic DNA with 5'GGGCTGCAGCT-TGTTCGCGGTTTTGG3' and 5'GGGGATCCTTGGGGT-TGCCCTCGTTGC3' primers (underlined sequences highlight the PstI and the BamHI restriction sites respectively). Gene replacement was achieved by transfection of the parasites with DNA constructions containing the *LiEndoG* 5' UTR and 3'UTR ends (cloned using the EcoRI/KpnI and BamHI/PstI restriction sites, respectively) within the pUC19 vector containing either the Puromycin *N*-acetyl-transferase (PAC) (pUC195'UTRLiEndoG:*pur*-o:3'UTRLiEndoG) or the neomycin resistance gene (*neo*) (pUC195'UTRLiEndoG:*neo*:3'UTRLiEndoG). Both *pac* and *neo* genes had been previously inserted into the pUC19 vector using the KpnI and BamHI restriction sites.

EndoG expression and purification

E. coli BL21(DE3) pLys bacterial cells were transformed with pRSET-LiEndoG. Recombinant LiEndoG (rLiEndoG) expression was then induced during 30 minutes with 1 mM IPTG at 37°C. Because most of the expressed protein was found in the non-soluble fraction, *E. coli* cells were lysed in 6 M GuHCl and the cleared lysate was loaded onto a Ni-NTA resin under denaturing conditions according to the manufacturer's instructions (Qiagen). On-column refolding of the protein was achieved by progressive reduction of the urea concentration from 6 M to 1 M. The protein was used for monoclonal antibody production according to standard protocols [11].

Nuclease activity assay

1 µg of plasmid DNA (pRSETa) was digested in a final volume of 20 µL with increasing amounts of rLiEndoG (20–150 ng) for 1 h at 37°C. Digested DNA samples were denatured using 0.1 volumes of NaOH 1 M, heated 2 minutes at 55°C, neutralized with 0.2 volumes of TrisHCl 1 M, pH 4.0, chilled in ice and finally diluted in Milli-Q water in order to avoid the effect of the salts on the migration of the samples during the subsequent electrophoresis. The DNA samples were finally analyzed on an agarose gel at 1.2% w/v, stained with ethidium bromide and visualized under UV light.

Nuclease activity was also tested against a 500-bp PCR product amplified with a forward primer fluorescently labeled at its 5' end with 6-carboxyfluorescein (FAM) (Applied Biosystems). The PCR fragment with a fluorescent 5' end was purified using a spin-column (Qiagen) in order to remove the primers used for the amplification. The PCR fragments were digested in a final volume of 20 µL with increasing amounts of rLiEndoG or DNase I (Boehringer Mannheim) at 37°C for 1 h. 5 µL of the digested samples were heat-denatured and diluted 1:10 and analyzed by capillary electrophoresis using the 3130 Genetic Analyzer (Applied Biosystems). The results obtained were examined using the Peak Scanner v1.0 (Applied Biosystems) software. The remaining 15 µL from each digestion were analyzed by electrophoresis on an agarose gel at 1.2% w/v, stained with ethidium bromide and visualized under UV light.

Promastigote transfection

The parasites were harvested in logarithmic growth phase and transfected by electroporation as previously described [12]. Stably transfected strains were selected in RPMI/20%FCS with 20 µg/mL of puromycin for hemi-KO parasites and with 20 µg/mL of puromycin and 30 µg/mL of G418 for double-KO parasites.

The pUC195'UTRLiEndoG:*pur*:3'UTRLiEndoG construct was linearized using the restriction enzymes HindIII and EcoRI. The pUC195'UTRLiEndoG:*neo*:3'UTRLiEndoG construct was linearized using the restriction enzyme EcoRI. The digestion products were gel purified (Illustra GFX gel purification kit, General Electric) prior to transfection. Individual clones were isolated in agar plates [13]. Cultured cells were diluted 1:5 when the density exceeded 5×10^6 cells/mL and cell densities were determined by triplicates using cell counting grids (Marienfeld, with a depth of 0.1 mm).

Protein Extraction and Western Blotting

40×10^6 parasites were lysed in 100 µL of Laemmli lysis buffer and boiled for 10 minutes. 20 µL of each sample were used for SDS-PAGE analysis in 10% acrylamide gels. For Western blot analysis, proteins were transferred from the gels to a PROTRAN nitrocellulose membrane (Whatman) in transfer buffer (25 mM Tris-HCl, 192 mM glycine, 20% methanol, 0.02%SDS, pH 8.3). The membranes were first incubated for 2 h in a blocking solution consisting of TBS-T buffer (150 mM NaCl, 10 mM Tris, 0.1% Tween, pH 8) supplemented with 5% of BSA (bovine serum albumin) (Sigma) and then with anti-LiEndoG monoclonal antibody for 16 hours at 4°C with shaking. An anti-mouse monoclonal antibody HRP-conjugated (Sigma) diluted 1:5000 in the blocking solution was used as the secondary antibody.

Polyclonal anti-Hsp70 antibody diluted 1:2000 was used as a loading control. In this case, a polyclonal anti-rabbit antibody conjugated with HRP (Santa Cruz) diluted 1:5000 was used as the secondary antibody. Antibodies were recognized with ECL reagent (Thermo Scientific).

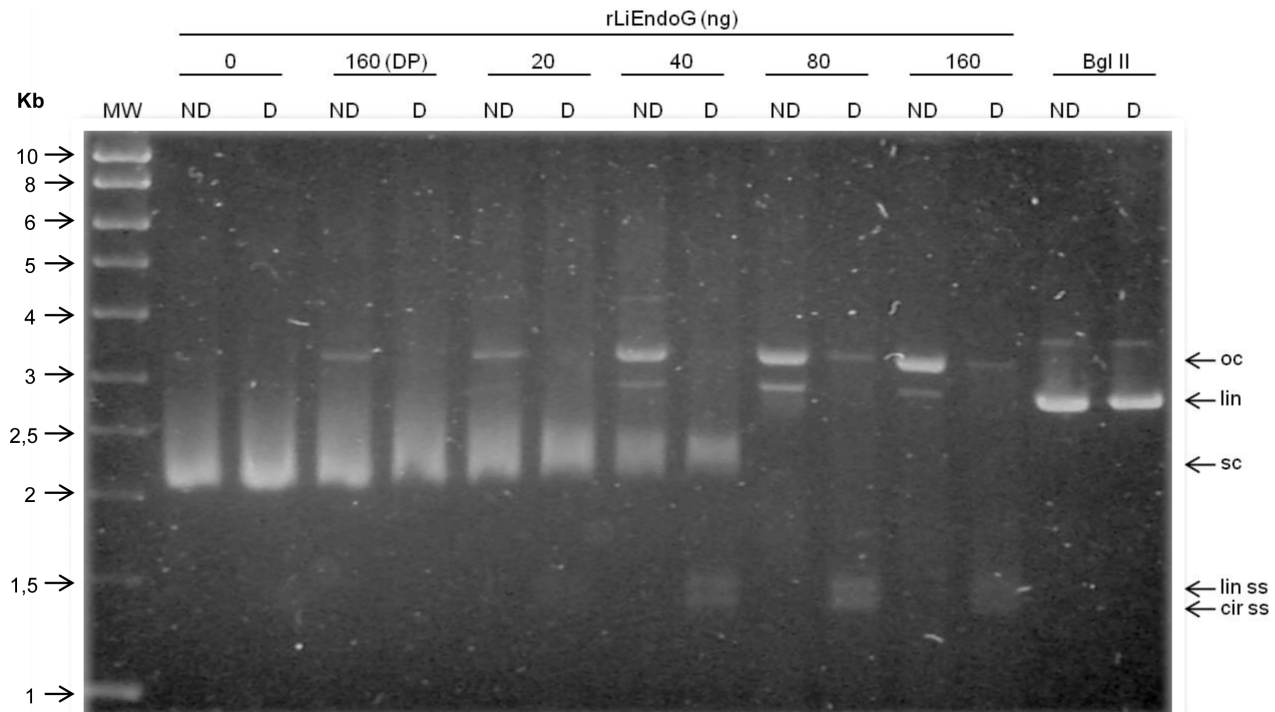


Figure 1. Analysis of rLiEndoG activity on double-stranded DNA. 1 μ g of supercoiled plasmid DNA (sc) was digested with increasing amounts of rLiEndoG (0, 20, 40, 80 and 160 ng) and with 160 ng of denatured rLiEndoG protein (160DP) as a negative control. The restriction enzyme Bgl II (Bgl II) was used as a control to generate linear DNA (lin). Each sample of digested DNA was split into two aliquots and one of them was heat-denatured (D) while the other one remained non-denatured (ND). All samples were electrophoresed in a 1% agarose gel and visualized under UV light. The different forms of plasmid DNA are indicated: open circular (oc); linear double-stranded (lin); supercoiled (sc); linear single-stranded (lin ss); circular single-stranded (cir ss).

doi:10.1371/journal.pone.0089526.g001

THP-1 infections

THP-1 cells were seeded at 120,000 cells mL^{-1} in 24-multiwell plates (Nunc, Roskilde, Denmark) and differentiated to macrophages for 24 h in 1 mL RPMI-1640 medium containing 10 ng mL^{-1} phorbol 12-myristate 13-acetate (PMA, Sigma-Aldrich) at 37°C. Then the culture medium was removed and *L. infantum* promastigotes in a ratio 20:1 to THP-1 cells were added in RPMI-1640 medium. After 24 hours of incubation all medium with non-infecting promastigotes was removed. Cells were washed three times with 1 \times phosphate-buffered saline (PBS) and detached with TrypLE Express (Invitrogen, Leiden, The Netherlands) according to the manufacturer's indications. Infection rates were measured by flow cytometry as the percentage of GFP positive cells. The number of amastigotes inside the cells was indirectly determined based on the intensity of the green fluorescence of the infected cells.

Mitochondrial transmembrane potential determination

Logarithmically growing promastigotes (1×10^6 cells) were incubated with Mitotracker Red CMXRos (Life Technologies) to a final concentration of 100 nM and maintained under normal growing conditions for 45 minutes. As a mitochondrial membrane depolarization control the protonophore carbonyl cyanide m-chloro phenyl hydrazone (CCCP) was used to a final concentration of 100 μ M under the same conditions. Subsequently, samples were analyzed using a FC500 MPL flow cytometer (Beckman-Coulter).

LiEndoG activity

The nuclease activity of LiEndoG in the presence of **Lei49** was monitored by measuring the increase in fluorescence derived from the digestion of a piece of dsDNA. 6-Carboxyfluorescein (FAM)-derived fluorescence is quenched by the proximity of tetramethylrhodamine (TAMRA) in the undigested probe constructed by hybridization of the oligonucleotides FAM-5'-CTG TCG CTA CCT GTG G-3'-TAMRA and FAM-5'-CCA CAG GTA GCG ACA G-3'-TAMRA. Digestion of any of the two oligonucleotides causes separation of the fluorophore from the quencher giving rise to emission of a fluorescent signal. 30 pmol of the double-stranded probe were digested with 2.5 ng mL^{-1} LiEndoG. Reactions were monitored in a Victor 1420 Multilabel Counter (Wallac) at an excitation and emission wavelengths of 492 and 517 nm, respectively.

Cell viability assay

Drug treatment of promastigotes was performed during logarithmic growth phase at a concentration of 2×10^6 parasites/mL at 26°C. The percentage of living cells was evaluated by flow cytometry by the propidium iodide (PI) exclusion method [14]. Briefly, treated parasites were stained for 10 min with 5 μ g/mL PI. The number of PI-negative parasites was determined in a Beckman Coulter FC500 flow cytometer.

Confocal microscopy

All experiments were done starting with 10×10^6 logarithmically growing promastigotes and visualized under a LEICA TCS-SP5 confocal microscope. Mito-Tracker Red CMXRos (Life Technol-

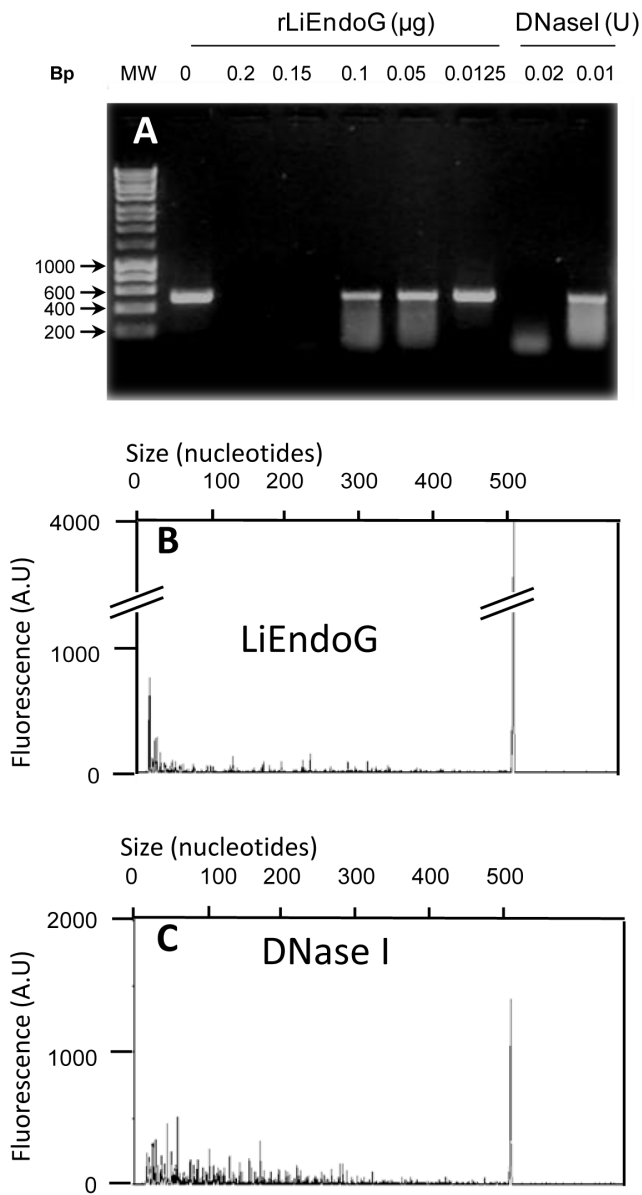


Figure 2. Exonuclease activity of rLiEndoG. 500 ng of a PCR fragment 5'-labeled with FAM were digested with rLiEndoG or DNase I. The amount of enzyme used for digestion is indicated. Digestion products were processed by both agarose gel (1%) and capillary electrophoresis. **A)** Agarose gel of the DNA fragments generated after 1 h of digestion with rLiEndoG or DNase I. **B)** Capillary electrophoresis of the DNA probe digested for 1 h with 0.1 µg of rLiEndoG. **C)** Capillary electrophoresis results obtained for the DNA probe digested for 1 hour with 0.01 units of DNase I. Digested DNA was heat-denatured prior to capillary electrophoresis. Fluorescence intensities (arbitrary units) of the single-stranded DNA fragments generated after digestion and denaturation are shown on the y axis. Sizes of the ssDNA fragments (in nucleotides) are shown on the x axis. Fragment sizes were analyzed with the Peak Scanner (Applied Biosystems) software. Accurate sizes can only be predicted for fragments longer than 50 nucleotides. doi:10.1371/journal.pone.0089526.g002

ogies) was used for mitochondrial localization at a final concentration of 200 nM according to the manufacturer's instructions. Briefly, cell cultures were centrifuged to obtain a cell pellet and gently suspended in warm (26°C) staining solution containing 200 nM Mito-Tracker Red CMXRos. After 45 minutes at 26°C

the cells were re-pelleted by centrifugation and suspended in fresh warm medium for 30 minutes. Finally, the cells were re-pelleted again and washed in PBS. Cells were then fixed in 2% paraformaldehyde in PBS for 10 minutes at room temperature and washed with PBS. 1.5×10^6 cells were spread on poly-lysine pre-coated slides and permeabilized with 0.1% Triton X-100 (TX-100) in PBS solution for 10 minutes. Preparations were then incubated with: i) blocking solution (2% BSA, 0.1% TX-100 in PBS) for 60 minutes at room temperature, ii) primary monoclonal antibody anti-LiEndoG (diluted 1:2 in blocking solution) for 16 h at 4°C, and iii) antibody Alexa Fluor 488 anti-mouse (diluted 1:1000 in blocking solution) (Life Technologies). Each incubation was followed by 3 washes with $1 \times$ PBS.

Finally, the samples were covered with ProLong Gold Antifade Reagent with DAPI liquid mountant (Life Technologies) and visualized under the microscope.

Statistics

All the analyses were carried out with a minimum of three independent experiments. The statistical significance of the differences between treatments was evaluated by a two-tailed unpaired t-test.

Results

The endonuclease activity of LiEndoG generates single-stranded cuts in the DNA double helix

To determine whether, as described for other endonucleases [6], LiEndoG is able to produce nicks in dsDNA, we proceeded to digest supercoiled (sc) plasmid DNA with increasing amounts of rLiEndoG expressed in *E. coli*. Each digested sample was divided into two aliquots, and one of them was denatured to follow the migration of the distinct forms of released single-stranded DNA. The restriction enzyme Bgl II was used as a control for the simultaneous digestion of both strands to generate a linear (lin) DNA molecule (Figure 1; Bgl II). rLiEndoG is able to digest supercoiled closed circular DNA, thus demonstrating that this enzyme displays endonuclease activity. The enzyme primarily generates open circular (oc) intermediates while only a very faint band of linear (lin) DNA can be observed. Denaturation of the samples (D lanes) allowed the detection of the fully-sized linear and circular single-stranded (ss) DNA molecules that made up the open circular forms generated after digestion. Because ethidium bromide has much higher affinity for dsDNA, these ss molecules are weakly stained compared to the oc DNA. Accordingly, rLiEndoG displays endonucleolytic activity but digests only one of the two strands of the double helix in every digestion step. However, this enzymatic activity may render linear DNA as a minor digestion product when two independent cuts in the complementary strands are so close together that base pairing can no longer maintain the circular structure (Figure 1; lin). The progressive decrease in the total ethidium-derived fluorescence reveals a process of DNA degradation that can be explained by the release of small oligonucleotides from the double helix after accumulation of single strand breaks. Thus, the enzyme seems to cut stochastically any of the two strands, as has been described for *Drosophila melanogaster* EndoG [15].

rLiEndoG shows exonuclease activity

While EndoGs from higher eukaryotes display only endonuclease activity [6], Nuc1p from *S. cerevisiae* behaves both as an endonuclease and as an exonuclease [16]. Once established that LiEndoG is an endonuclease, we addressed the possibility that it might also display exonuclease activity. To this end, genomic

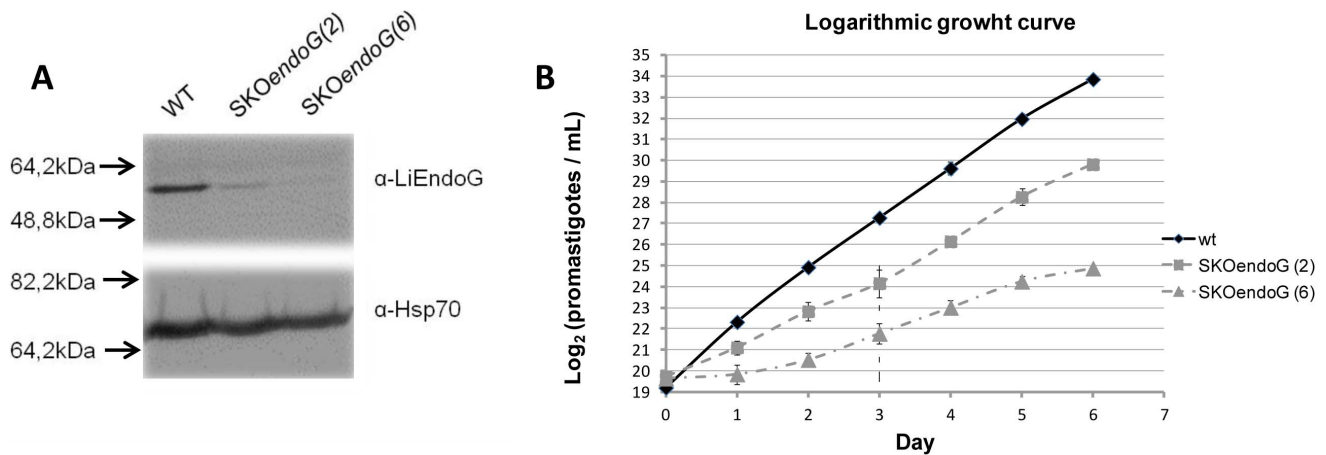


Figure 3. LiEndoG single allele replacement reduces *L. infantum* promastigotes growth rate. **A)** Immuno-blot using the monoclonal antibody α -rLiEndoG against cellular extracts from wild-type (WT) promastigotes and from two clones of *L. infantum* promastigotes with a single copy of the *LiEndoG* gene: SKOendoG(2) and SKOendoG(6). A Polyclonal α -Hsp70 (1:5000) antibody was used as a loading control. **B)** Logarithmic growth curve for WT, SKOendoG(2) and SKOendoG(6) promastigotes. The curve is expressed as the \log_2 parasite number per mL. Error bars represent standard deviations; results are representative of three independent experiments. doi:10.1371/journal.pone.0089526.g003

DNA was amplified by PCR using a FAM-labeled forward primer so as to generate a PCR product with a fluorescent 5' end. The resulting PCR product was digested with increasing concentrations of rLiEndoG. DNase I was used as a control for endonuclease activity [17,18]. The products resulting from the digestion were processed by both agarose gel electrophoresis (Figure 2A) and capillary electrophoresis (Figure 2B and 2C). Whereas gel electrophoresis allows the detection of all the dsDNA fragments generated after digestion, only ss FAM-labelled fragments (fragments containing the labelled 5' end) are detected in capillary electrophoresis.

Digestion with rLiEndoG mostly generated fluorescent low-molecular-weight ss- DNA fragments, being the major peak the one closest to the y axis (Figure 2B). This accumulation of small DNA fragments is not caused by extensive digestion because the conditions of the reaction were adjusted to allow the detection of a major peak of undigested DNA (500 bases). By contrast, the peak pattern observed after digestion with DNase I is much more dispersed along the x axis and the smallest fragments are not the most abundant (Figure 2C). Taken these results together, the nicking activity of rLiEndoG appears to be preferentially directed against the 5' end of DNA.

LiEndoG has a pro-life role

EndoGs from other lower eukaryotes have been reported to play a dual role in both cell death and survival [9]. A gene replacement strategy was followed to determine whether LiEndoG is also involved in pro-life roles in *Leishmania*. All attempts to delete both *LiEndoG* alleles were unsuccessful but parasites remained viable after single allele replacement (*LiEndoG*^{-/+}). The results shown in Figure 3 indicate that LiEndoG expression is highly reduced in the *LiEndoG*^{-/+} promastigotes (Figure 3A). Growth of these parasites was analyzed by counting the number of cells for 6 consecutive days. Parasites were kept in the logarithmic growth phase by daily dilutions (Figure 3B). Since the y axis represents the \log_2 of the number of parasites an increase of one unit denotes one round of division of the entire population. Results for two different randomly selected *LiEndoG*^{-/+} clones (SKOendoG(2) and SKOendoG(6)) are shown. *LiEndoG*^{-/+} promastigotes grow at a much lower rate than do their wild-type counterparts under optimal

growth conditions. In fact, a difference of 9 \log_2 units between clone 6 and the wild-type parasites is apparent after 6 days of culture. Variations in the growth rates of clones 2 and 6 seem to correlate with differences observed in the residual levels of expression of LiEndoG. Thus, the disadvantage of the parasites harboring reduced levels of LiEndoG is consistent with a pro-life role for this protein.

LiEndoG-deficient parasites were also tested for their ability to grow inside infected macrophages. Clones 2, 6 and wild-type parasites were transfected with a plasmid containing the sequence coding for eGFP. After incubation with nourseothricin and selection of two random clones (SKO-GFP-2.1 and SKO-GFP-6.1), the fluorescent parasites were incubated with differentiated THP-1 cells and both the infection rates and the mean fluorescent values of the infected cells were evaluated by flow cytometry. The number of infected macrophages is significantly reduced when incubating with SKO-GFP-2.1 and SKO-GFP-6.1 clones (Figure 4B). Moreover, the parasite load inside the infected THP-1 cells is also strongly decreased as indicated by the much lower mean fluorescence intensities of the infected cells (Figure 4C). Accordingly, LiEndoG-deficient parasites also have growth defects during macrophage infection, confirming the pro-survival function of this protein in amastigotes.

Because of the mitochondrial location of this protein, we also evaluated any possible disturbances in the morphology of this organelle in the *LiEndoG*^{-/+} clones but no significant differences could be observed. However, the LiEndoG-deficient parasites show increased Mitotracker-derived fluorescence intensities when compared with WT ones (Figure 4D), which demonstrates an increased mitochondrial transmembrane potential ($\Delta\psi_m$) in the *LiEndoG*^{-/+} clones. This constitutes an unexpected result as LiEndoG deficiency is expected to impair the mitochondrial function. However, it has already been demonstrated that DNA damages in the mitochondrial DNA of *Leishmania* parasites cause an increase in their $\Delta\psi_m$ [19].

We have recently identified several molecules able to reduce the activity of LiEndoG *in vitro* [10]. Among them, **Lei49** behaves as one of the best inhibitors (Figure 5) and also as one of the compounds with better cytotoxic activity against *L. infantum* promastigotes ($IC_{50} = 4.33 \pm 0.05 \mu M$) and amastigotes

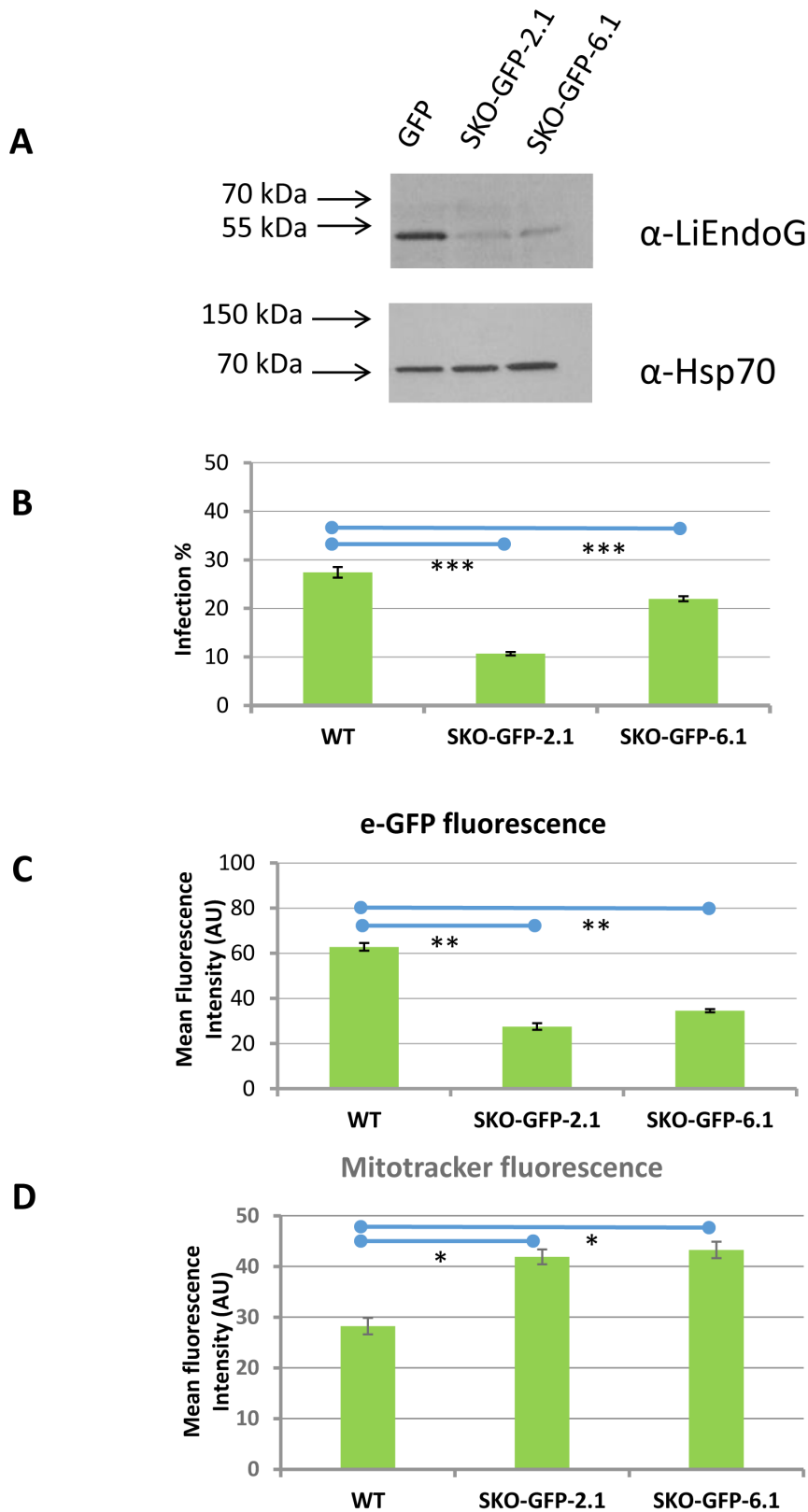


Figure 4. LiEndoG single allele replacement impairs macrophage infection, reduces intracellular survival and increases mitochondrial transmembrane potential. A) Immuno-blot using the monoclonal antibody α -rLiEndoG against cellular extracts from GFP-expressing parasites. GFP: wild-type promastigotes. SKO-GFP-2.1 and SKO-GFP-6.1: Two independent GFP-expressing *LiEndoG*^{-/-} promastigote clones. A Polyclonal α -Hsp70 (1:5000) antibody was used as a loading control. **B)** and **C).** Intracellular infections. PMA-differentiated THP-1 cells were infected with promastigotes expressing the green-fluorescent protein (eGFP). **B).** Percentages of infection were evaluated by counting the number of green-fluorescent PMA-differentiated THP-1 cells by flow cytometry. **C)** Parasite load was determined based on the mean fluorescence intensity

(arbitrary units; AU) of the infected THP-1 cells. **D**). Analysis of mitochondrial membrane potential based on the relative fluorescence due to Mitotracker Red CMXRos. Error bars symbolize standard deviations; results are representative of three independent experiments. The mean value of the clones differ from that of WT with p values $<10^{-3}$ (*), $<10^{-4}$ (**), $<10^{-5}$ (***).
doi:10.1371/journal.pone.0089526.g004

($IC_{50} = 6.44 \pm 0.01 \mu\text{M}$). To demonstrate that LiEndoG activity is essential for parasite survival, **Lei49**-induced death rates were analyzed in control parasites transfected with an empty vector and in parasites overexpressing LiEndoG (Figure 6A). Parasite viability upon treatment with increasing concentrations of **Lei49** or edelfosine was measured based on the percentage of propidium iodide (PI) negative cells. The results presented in Figure 6B demonstrate that parasites overexpressing this enzyme are significantly protected against cell death caused by moderate concentrations of **Lei49** but, as expected, increasing concentrations of the compound abolish the protective effect of overexpression. By contrast, as previously described [3], the same parasites are more susceptible to edelfosine-induced cell death (Figure 6C), probably as a consequence of increased activity of the proapoptotic molecular machinery. As shown in Figure 6D the overexpressed protein, similarly to the endogenous one, is correctly targeted to the mitochondrion. These results support the relevance of LiEndoG for parasite survival in addition to its better known pro-death role.

Discussion

We show that LiEndoG is an endo-exonuclease that is able to generate single-strand breaks in circular supercoiled DNA whereas it has a preferential 5' exonuclease activity over linear DNA. Previous findings from us and from other groups have already demonstrated that this nuclease activity is relevant during the

apoptotic cell death process of the parasites [3,20,21]. The results presented herein demonstrate that this enzyme is also essential for survival. Firstly, we show that parasites which, as a consequence of single allele replacement, contain reduced levels of LiEndoG protein grow at much lower rates than do their wild-type counterparts. Secondly, we demonstrate that these LiEndoG-deficient parasites are much less infectious for macrophages. Thirdly, we prove that the cytotoxic activity of the LiEndoG inhibitor **Lei49** is significantly reduced when the protein is overexpressed in the parasites. Our results agree with those reported by Buttner et al. who showed that deletion of Nu1p reduces yeast growth rates, an effect attributed to its possible role in mtDNA recombination [9]. Indeed, orthologs of EndoG in *S. cerevisiae* and in *Neurospora* initiate homologous recombination by generating single-stranded regions in the chromosomes by means of their 5' exonuclease activity [22,23]. The 5' exonuclease activity described for LiEndoG in this article could be considered as an indication of its possible involvement in recombination and repair of mitochondrial DNA, even though more work will be necessary to prove this point. Taken together, our results strongly indicate that LiEndoG does indeed play a dual pro-life and pro-death role in *L. infantum* similar to that described for the EndoG orthologs in yeast and *C. elegans*. Moreover, the involvement of LiEndoG in the survival of these parasites makes this protein a potential target against which new leishmanicidal drugs could be developed.

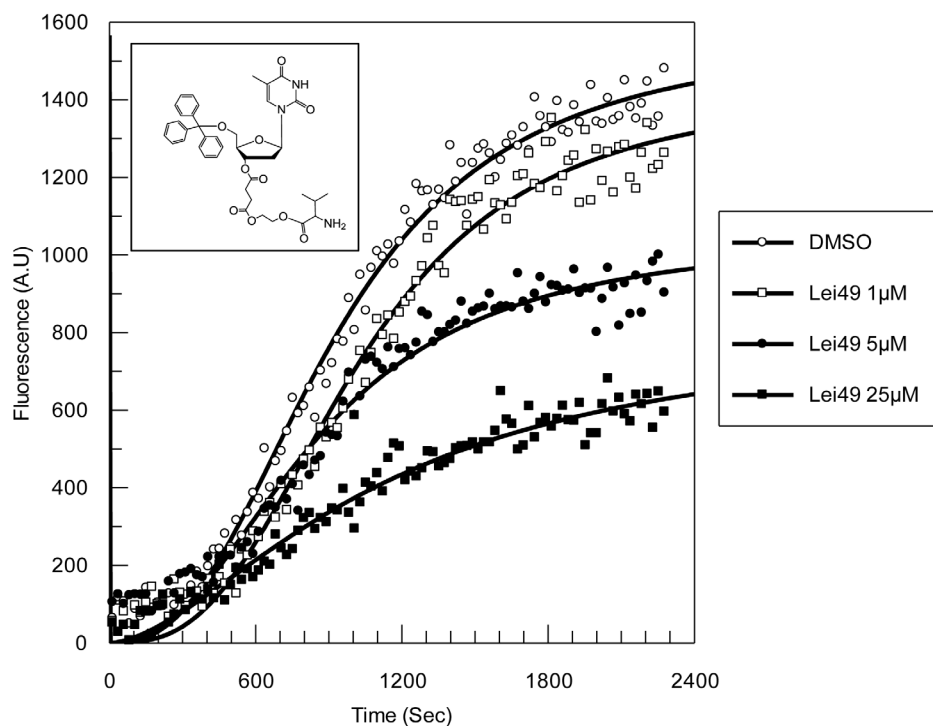


Figure 5. Inhibition of the nuclease activity of LiEndoG by Lei49. LiEndoG activity was assayed by the cleavage of a dually labeled double-stranded probe. DNA degradation is followed by the increase in fluorescence observed over 40 min of probe digestion in the presence of increasing amounts of **Lei49**.
doi:10.1371/journal.pone.0089526.g005

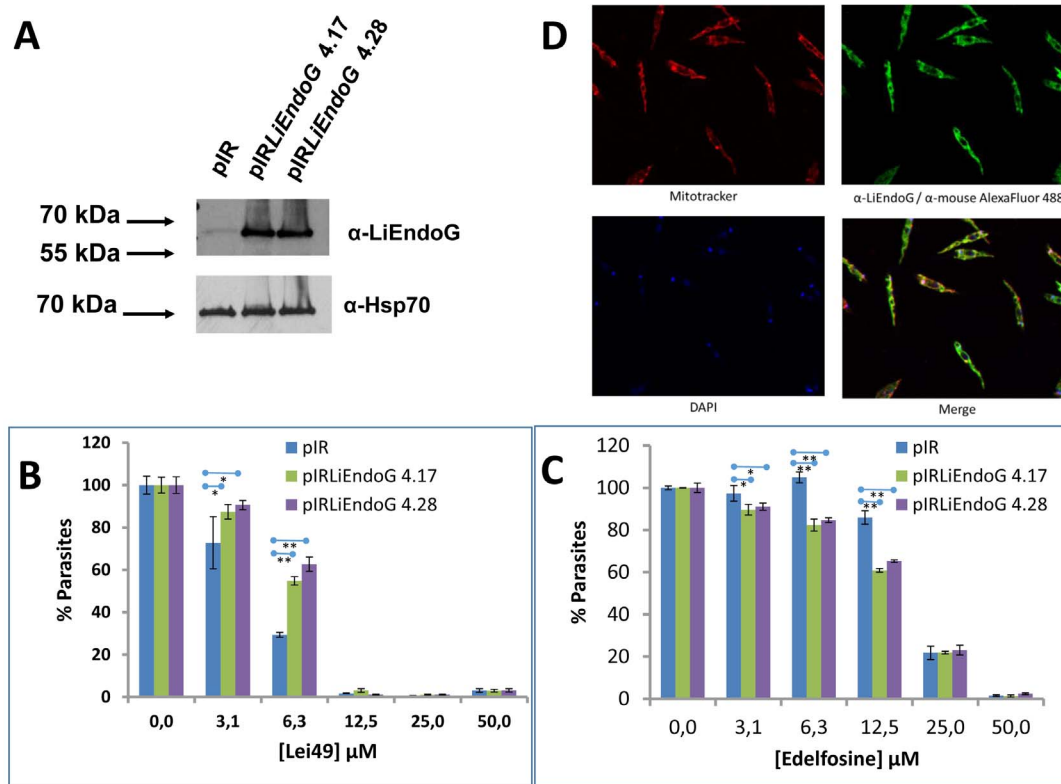


Figure 6. Drug-induced cell death in parasites overexpressing LiEndoG. **A)** Immuno-blot using the monoclonal antibody α -rLiEndoG against cellular extracts from parasites transfected with an empty vector (pIR) and from two independent clones overexpressing LiEndoG (pIRLiEndoG 4.17; pIRLiEndoG 4.28). A polyclonal α -Hsp70 (1:5000) antibody was used as a loading control. **B)** Percentages of PI negative (alive) cells in the three cell lines treated with increasing concentrations of **Lei49**. **C)** Percentages of PI negative (alive) cells in the three cell lines treated with increasing concentrations of **Edelfosine**. Error bars symbolize standard deviations; results are representative of three independent experiments. **D)** Mitochondrial localization of LiEndoG in the overexpressing clone pIRLiEndoG 4.28. Similar results (not shown) were obtained for clone pIRLiEndoG 4.17. The mean values of the clones differ from that transfected with an empty vector with p values $<10^{-1}$ (*) or $<10^{-2}$ (**). doi:10.1371/journal.pone.0089526.g006

Author Contributions

Conceived and designed the experiments: ER JFA JC FG AJR DM MJPP MJC. Performed the experiments: ER JFA CO CMG DM EC KJG.

Analyzed the data: ER JFA CO CMG FG AJR DM JC. Contributed reagents/materials/analysis tools: ER JFA CO EC MJPP MJC JC AG. Wrote the paper: ER FG AJR.

References

- Parrish J, Li L, Klotz K, Ledwich D, Wang X, et al. (2001) Mitochondrial endonuclease G is important for apoptosis in *C. elegans*. *Nature* 412: 90–94.
- Schafer P, Scholz SR, Gimadudinow O, Cymerman IA, Bujnicki JM, et al. (2004) Structural and functional characterization of mitochondrial EndoG, a sugar non-specific nuclease which plays an important role during apoptosis. *J Mol Biol* 338: 217–228.
- Rico E, Alzate JF, Arias AA, Moreno D, Clos J, et al. (2009) *Leishmania infantum* expresses a mitochondrial nuclease homologous to EndoG that migrates to the nucleus in response to an apoptotic stimulus. *Mol Biochem Parasitol* 163: 28–38.
- Ruiz-Carrillo A, Renaud J (1987) Endonuclease G: a (dG)n X (dC)n-specific DNase from higher eukaryotes. *Embo J* 6: 401–407.
- Cote J, Ruiz-Carrillo A (1993) Primers for mitochondrial DNA replication generated by endonuclease G. *Science* 261: 765–769.
- Low RL (2003) Mitochondrial Endonuclease G function in apoptosis and mtDNA metabolism: a historical perspective. *Mitochondrion* 2: 225–236.
- Cymerman IA, Chung I, Beckmann BM, Bujnicki JM, Meiss G (2008) EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Res* 36: 1369–1379.
- Cummings OW, King TC, Holden JA, Low RL (1987) Purification and characterization of the potent endonuclease in extracts of bovine heart mitochondria. *J Biol Chem* 262: 2005–2015.
- Buttner S, Eisenberg T, Carmona-Gutierrez D, Ruli D, Knauer H, et al. (2007) Endonuclease G regulates budding yeast life and death. *Mol Cell* 25: 233–246.
- Casanova E, Moreno D, Gigante A, Rico E, Genes CM, et al. (2013) 5'-Trityl-Substituted Thymidine Derivatives as a Novel Class of Antileishmanial Agents: *Leishmania infantum* EndoG as a Potential Target. *ChemMedChem* 8: 1161–1174.
- Yokoyama WM, Christensen M, Santos GD, Miller D (2006) Production of monoclonal antibodies. *Curr Protoc Immunol Chapter 2: Unit 2.5*.
- Alzate JF, Barrientos AA, Gonzalez VM, Jimenez-Ruiz A (2006) Heat-induced programmed cell death in *Leishmania infantum* is reverted by Bcl-X(L) expression. *Apoptosis* 11: 161–171.
- Quijada L, Soto M, Alonso C, Requena JM (2003) High-efficiency plating method for *Leishmania infantum*. *Mol Biochem Parasitol* 130: 139–141.
- Plano D, Baquedano Y, Moreno-Mateos D, Font M, Jimenez-Ruiz A, et al. (2011) Selenocyanates and diselenides: a new class of potent antileishmanial agents. *Eur J Med Chem* 46: 3315–3323.
- Loll B, Gebhardt M, Wahle E, Meinhardt A (2009) Crystal structure of the EndoG/EndoGI complex: mechanism of EndoG inhibition. *Nucleic Acids Res* 37: 7312–7320.
- Dake E, Hofmann TJ, McIntire S, Hudson A, Zassenhaus HP (1988) Purification and properties of the major nuclease from mitochondria of *Saccharomyces cerevisiae*. *J Biol Chem* 263: 7691–7702.
- Bernardi G, Gaillard C, Bernardi G (1975) The specificity of five DNAases as studied by the analysis of 5'-terminal doublets. *Eur J Biochem* 52: 451–457.
- Bernardi G, Ehrlich SD, Thierry JP (1973) The specificity of deoxyribonucleases and their use in nucleotide sequence studies. *Nat New Biol* 246: 36–40.
- Sen N, Das BB, Ganguly A, Mukherjee T, Tripathi G, et al. (2004) Camptothecin induced mitochondrial dysfunction leading to programmed cell death in unicellular hemoflagellate *Leishmania donovani*. *Cell Death and Differentiation* 11: 924–936.

20. Bosedasgupta S, Das BB, Sengupta S, Ganguly A, Roy A, et al. (2008) The caspase-independent algorithm of programmed cell death in *Leishmania* induced by baicalein: the role of LdEndoG, LdFEN-1 and LdTatD as a DNA 'degradesome'. *Cell Death Differ*.
21. Gannavaram S, Vedvyas C, Debrabant A (2008) Conservation of the proapoptotic nuclease activity of endonuclease G in unicellular trypanosomatid parasites. *J Cell Sci* 121: 99–109.
22. Zassenhaus HP, Denniger G (1994) Analysis of the role of the NUC1 endo/exonuclease in yeast mitochondrial DNA recombination. *Curr Genet* 25: 142–149.
23. Fraser MJ, Hatahet Z, Huang XT (1989) The actions of *Neurospora* endo-exonuclease on double strand DNAs. *J Biol Chem* 264: 13093–13101.

4.3 Asignación de los dominios presentes en la EndoG de *Leishmania infantum* en base a su estructura: caracterización de una región que regula la actividad de esta enzima en función del pH y una extensión C-terminal que dicta en gran medida su preferencia por los distintos sustratos de ADN

4.3.1 Hipótesis

La secuencia polipeptídica de *LiEndoG* presenta varias inserciones que son prácticamente exclusivas de tripanosomátidos y que podrían estar dotándola de propiedades enzimáticas únicas.

4.3.2 Objetivos

- Mejorar el modelo molecular de *LiEndoG* e incluir una previsión de estructura tridimensional para estas inserciones.
- Asignar las posibles funciones de estas regiones en función de su estructura, secuencia y localización.
- Comprobar experimentalmente el papel que tienen estas regiones sobre la actividad nucleasa de *LiEndoG*.

4.3.3 Resultados principales

Una vez demostrado el papel dual provida/promuerte de *LiEndoG* en los parásitos, decidimos llevar a cabo la caracterización molecular de esta enzima para conocer mejor su funcionamiento. En este sentido, nuestros resultados muestran que *LiEndoG* presenta una inserción rica en residuos ácidos que, según nuestro nuevo modelo molecular, se localiza de forma muy próxima al centro activo de la enzima y en la misma región espacial ocupada por el inhibidor de la EndoG de *Drosophila melanogaster* en el cristal *DmEndoG/DmEndoGI*. Dado que resultados previos obtenidos en nuestro laboratorio demostraban que *LiEndoG* está inhibida a pHs básicos (8 o superior), la presencia esta inserción de carácter ácido, así como su localización, nos llevó a postular que podría estar actuando como un sensor de pH y produciendo la autoinhibición de la enzima. De hecho, en este trabajo mostramos que la variante de *LiEndoG* en la que se han eliminado esta región (residuos 145-180) es mucho más

activa que su homóloga silvestre a pHs mayores de siete. Adicionalmente, hemos descubierto que: i) *LiEndoG* existe como un homodímero y cada uno de los dos sitios activos se localizan en lados opuestos de la interfaz de dimerización; ii) el reemplazo de la Ser211 del motivo SRGH del centro catalítico por el aspartato canónico de la mayoría de las nucleasas de esta familia da lugar a una enzima catalíticamente deficiente; iii) la actividad de la variante S211D puede ser rescatada cuando se reemplaza de forma simultánea la Ala247 por una arginina; y iv) la presencia de una extensión en el extremo carboxilo terminal de *LiEndoG* es la responsable de la preferencia de esta enzima por digerir ADN de cadena sencilla y las uniones entre ADN de doble cadena y cadena sencilla. En conjunto, nuestros resultados apoyan la idea de que *LiEndoG* es una proteína multidominio cuya actividad nucleasa puede ser modulada o incluso abolida mediante cambios en su arquitectura generados en respuesta a modificaciones del ambiente intracelular y a la interacción con otras proteínas.

4.3.4 Resumen

Llegados a este punto, podemos afirmar que nuestro grupo de investigación ha sido capaz de demostrar que la proteína *LiEndoG* posee características funcionales y estructurales muy similares a las mostradas por las EndoGs de otros eucariotas, incluyendo la localización mitocondrial y la capacidad de translocarse al núcleo para degradar el ADN como respuesta a un estímulo apoptótico (101). Adicionalmente, hemos demostrado que *LiEndoG* juega también un papel provida en *Leishmania infantum*, ya que la disminución de su expresión produce un retraso en el crecimiento de estos parásitos y su inhibición mediante el tratamiento con derivados de timidina tiene un efecto citotóxico sobre los mismos, efecto que además puede ser contrarrestado mediante la sobreexpresión de esta proteína (156,168). Teniendo en consideración este papel dual y el potencial que ello le otorga como posible diana terapéutica para el desarrollo de nuevos agentes antileishmania, pensamos que podría ser interesante llevar a cabo una caracterización detallada de la enzima para conocerla en mayor profundidad.

Para ello, en nuestro laboratorio se comenzó a mejorar el modelo molecular que habíamos desarrollado previamente para *LiEndoG* (101). Afortunadamente, en los últimos años se habían resuelto y depositado en *Protein Data Bank* (PDB) las estructuras cristalográficas de varios miembros de esta familia de nucleasas, entre las que se encontraban la EndoG de

Drosophila melanogaster (*DmEndoG*) en complejo con su inhibidor (*DmEndoGI*), varias variantes de la EndoG de *Caenorhabditis elegans* (CPS-6 o *CeEndoG*) y la endo-exonucleasa mitocondrial humana EXOG, proteínas con las cuales *LiEndoG* guarda un alto grado de homología y cuyas estructuras pudimos emplear para mejorar nuestro modelo. Un primer análisis del alineamiento múltiple de las secuencias de todas estas enzimas nos permitió detectar que *LiEndoG* presenta cuatro inserciones que no están presentes en sus homólogos (artículo 3, figura 1), por lo que nos pareció que podía ser interesante caracterizarlas y comprobar si otorgaban alguna propiedad característica a nuestra enzima. Además, tal y como ya se ha comentado previamente en esta tesis, una de las peculiaridades más llamativas de las EndoGs de tripanosomátidos es que presentan un motivo SRGH en el sitio catalítico - en lugar del motivo DRGH canónico de este tipo de nucleasas - por lo que pensamos que también podía ser interesante investigar más sobre las implicaciones que podía tener esta diferencia en nuestra enzima.

Finalmente, el conjunto de las técnicas de modelado empleadas nos permitió desarrollar un nuevo modelo tridimensional de *LiEndoG* en su forma dimérica que incluía estas cuatro inserciones (artículo 3, figura 3) y que proporcionaba una explicación estructural a algunas de las peculiaridades enzimáticas de *LiEndoG* que describiremos a continuación.

4.3.4.1 Características del centro catalítico de *LiEndoG*

Como se comentó en la introducción de esta tesis, nuestro trabajo de investigación previo mostraba que el centro catalítico de *LiEndoG* guarda una alta similitud con el de las nucleasas extracelulares de *S. marcescens* y *Anabaena sp.*, siendo ambas proteínas miembros de la superfamilia de las nucleasas inespecíficas con un dominio $\beta\beta\alpha$ de unión a metal. Basándonos en la estructura del centro catalítico de estas enzimas, fuimos capaces de generar un modelo del sitio activo de *LiEndoG* en el que se muestra que el Mg^{2+} catalítico está coordinado directamente con el oxígeno carboxamídico de la cadena lateral de la Asn246. Adicionalmente, pudimos comprobar como una de las moléculas de agua de la esfera de coordinación del catión divalente está unida mediante un enlace de hidrógeno al N δ del anillo imidazólico de la His214, residuo que a su vez actúa como una base general activando la molécula de agua que llevará a cabo el ataque nucleofílico sobre el átomo de fósforo del sustrato que se esté hidrolizando (101). Esta histidina catalítica está altamente conservada y

suele formar parte del motivo DRGH característico de esta familia de nucleasas. Sin embargo, tanto en tripanosomátidos como en EXOG, el aspártico de este motivo está reemplazado por un residuo de serina, por lo que estas enzimas presentan un motivo SRGH en lugar del motivo DRGH canónico en su centro activo. Por ello, decidimos explorar cuál era la relevancia de la Ser211 en *LiEndoG*, para lo que expresamos dos variantes en las que este residuo se había reemplazado por una Ala (S211A) o el aspártico canónico (S211D) y evaluamos su actividad nucleasa.

Nuestros resultados muestran que el reemplazo de la Ser211 por un aspártico (*LiEndoG*^{S211D}) produce una enzima con una actividad catalítica drásticamente reducida (artículo 3, figura 4). Este resultado fue doblemente chocante porque, por un lado, se había mostrado que el mismo tipo de reemplazo en EXOG daba lugar a una enzima con tan solo una ligera inhibición de su actividad respecto a la proteína silvestre (117) y, por otro lado, el reemplazo de esta misma Ser por una Ala en *LiEndoG* (*LiEndoG*^{S211A}) solo afectaba levemente a su actividad (artículo 3, figura 4). Por ello, pensamos que el efecto deletéreo observado para la mutación S211D debía ser consecuencia de las posibles interacciones y/o distorsiones que podía estar generando el hecho de introducir un residuo con carga negativa en esta posición de la proteína.

Tras revisar las estructuras de *DmEndoG* y CPS-6 comprobamos que en estas enzimas el carboxilato del aspártico del motivo DRGH está “anclado” mediante la formación de un puente salino con el grupo guanidinio de una Arg colindante. Dado que en *LiEndoG* la posición equivalente a este residuo está ocupada por la Ala247, pensamos que su reemplazo por una Arg podría reestablecer la pérdida de actividad observada en la variante S211D. En efecto, nuestros resultados muestran como la introducción de esta segunda mutación para generar una versión de la proteína con dos sustituciones (*LiEndoG*^{S211A/A247R}) es capaz de rescatar la actividad de la enzima hasta alcanzar aproximadamente un 75% de la observada en la proteína silvestre (artículo 3, figura 4). En conjunto, el análisis de nuestro modelo y los resultados obtenidos sugieren que la introducción de un Asp en la posición 211 de *LiEndoG* produce que el carboxilato de su cadena lateral desorganice el sitio activo y cause una disminución drástica de la actividad de la enzima, efecto que puede ser contrarrestado si se introduce una Arg en la posición 247 capaz de “anclar” este carboxilato mediante la formación de un puente salino.

Para completar nuestros experimentos, reemplazamos también His y la Arg del motivo SRGH por alaninas y evaluamos el efecto que producía cada uno de estos cambios sobre la actividad enzimática. Tal y como cabía esperar, la mutante *LiEndoG*^{H214A} fue incapaz de degradar el ADN, dado el papel catalítico crucial que juega este residuo como base general que activa la molécula de agua necesaria para el ataque nucleofílico. De forma similar, el reemplazo de la Arg212 por una Ala generó una variante prácticamente inactiva, lo cual no es sorprendente teniendo en cuenta que este residuo está implicado en la unión y posicionamiento del ADN sobre el sitio catalítico (artículo 3, figura 4).

4.3.4.2 *LiEndoG* existe como homodímero

Si bien suponíamos que *LiEndoG*, al igual que sus homólogas, podía existir como un homodímero, no teníamos ninguna evidencia de que efectivamente así fuera. En este sentido, la superposición de las estructuras tridimensionales de *DmEndoG*, CPS-6 y EXOG nos permitió identificar varios residuos involucrados en la estabilización de la interfaz de dimerización de estas enzimas que estaban altamente conservados. Muchos de estos residuos también estaban presentes en *LiEndoG*, lo cual aumentó nuestra confianza sobre la posibilidad de que nuestra proteína fuese dimérica. Finalmente, la realización de ensayos de equilibrio de sedimentación y ultracentrifugación analítica nos permitió obtener una confirmación experimental de que *LiEndoG* existe principalmente en su forma homodimérica, si bien también se puede observar la existencia de proteína monomérica y tetramérica de forma muy minoritaria (artículo 3, figura 2). Teniendo en cuenta estos resultados, nuestro modelo de *LiEndoG* refleja una proteína dimérica en la que cada uno de los centros activos se disponen de forma distante y orientados en sentidos opuestos respecto a la interfaz de dimerización, al igual que ocurre en las proteínas empleadas como modelo estructural (artículo 3, figura 3).

4.3.4.3 *LiEndoG* posee un dominio de autoinhibición que regula su actividad en función del pH

Tal y como se ha comentado a lo largo de esta tesis, nuestros resultados muestran que *LiEndoG* juega un papel dual en los parásitos y está involucrada tanto en la supervivencia como en la muerte de estos (101,156,168). Por tanto, parece evidente que deben existir mecanismos moleculares y/o celulares capaces de regular la actividad de esta enzima y de

asegurar el correcto funcionamiento de este equilibrio vida/muerte. Además, no podemos olvidar que *LiEndoG* es una nucleasa no específica, por lo que estos mecanismos también pueden servir para proteger al material genético frente a un ataque nucleolítico que podría ser autodestructivo.

En este sentido, se ha descrito que algunos organismos utilizan sistemas como la compartimentalización celular, la formación de puentes disulfuro, la regulación por pH o la presencia de proteínas inhibidoras específicas para modular la actividad de este tipo de nucleasas (123,126,127,136,172). Uno de los ejemplos más caracterizados de esta regulación, comentado también en la introducción de esta tesis, lo conforman la *EndoG* de *Drosophila melanogaster* (*DmEndoG*) y su inhibidor específico *DmEndoGI*. En este organismo, la formación del complejo *DmEndoG:DmEndoGI* permite que la proteína inhibidora encaje un bucle o *loop* rico en residuos ácidos en el surco de unión del sustrato de *DmEndoG* mediante la formación de numerosos puentes salinos, bloqueando así la unión de los oligonucleótidos e inhibiendo la actividad nucleasa (123). Si bien es cierto que hasta la fecha no ha sido descrito ningún inhibidor de este tipo en *Leishmania*, nuestro modelo nos permitió comprobar que la primera de las cuatro inserciones detectadas en la secuencia de *LiEndoG* (residuos 145-180) se localizaba cerca del sitio activo y aproximadamente en la misma región espacial ocupada por *DmEndoGI* en el complejo *DmEndoG:DmEndoGI*. La presencia de numerosos residuos ácidos de aspártico y glutámico en esta inserción, sumado a que los resultados previos de nuestro laboratorio mostraban que *LiEndoG* está inhibida a pHs iguales o superiores a 8, nos llevó a pensar que este tramo podría estar actuando como un dominio de autorregulación y ser el responsable de la dependencia de pH mostrada por esta enzima. Para testar esta hipótesis, los residuos de esta inserción exclusiva de tripanosomátidos fueron reemplazados por la secuencia AKNDA que tiene *DmEndoG* en esta región y la actividad de la variante sin el putativo dominio de autoinhibición (*LiEndoGΔi*) fue comparada con la de la proteína silvestre (*LiEndoG*) a pHs comprendidos entre 6.5 (pH óptimo para la actividad de *LiEndoG*) y 8.

Nuestros resultados muestran que *LiEndoGΔi* permanece catalíticamente activa en todo el intervalo de pHs estudiado mientras que, por el contrario, la actividad de la proteína silvestre disminuye progresivamente según se incrementa el pH hasta 8.0, (artículo 3, figura 5A). Es importante señalar que tanto la actividad de la proteína silvestre como la de la mutante sin el putativo dominio de autoinhibición es máxima a pH 6.5. En consecuencia, la delección de

este dominio no modifica el valor de pH óptimo para la actividad catalítica, sino que amplía el intervalo de valores de pH en los cuales esta enzima permanece activa. Por tanto, podemos concluir que los residuos comprendidos entre las posiciones 145-180 constituyen un dominio de autoinhibición en *LiEndoG* capaz de regular su actividad en función del pH.

Existen varios resultados que apuntan a la presencia de un pH próximo a 8 en la matriz mitocondrial y de aproximadamente una unidad menos en el citosol (173,174). Según esto, *LiEndoG* estaría casi completamente inhibida en el interior de la mitocondria y podría activarse una vez es liberada al citosol durante la muerte celular, pudiendo esta activación estar además potenciada por el descenso del pH citosólico observado durante el proceso apoptótico (175,176). Sin embargo, dado que tanto el pH citosólico como mitocondrial pueden verse fuertemente reducidos durante la activación de las células con agonistas de movilizadores de Ca^{2+} (174), la actividad de *LiEndoG* dentro de la mitocondria podría estar estimulada bajo diferentes condiciones celulares. De hecho, cualquier agente que cause una disminución de la actividad mitocondrial puede producir una reducción en el transporte de protones y provocar una reducción del valor del pH en el interior de la matriz mitocondrial que, eventualmente, podría llevar a la activación indirecta de *LiEndoG*.

Finalmente, quisimos evaluar si la presencia de residuos ácidos en este dominio de autoinhibición era realmente relevante para la modulación de la actividad de la enzima. Para ello, reemplazamos los tres residuos de aspártico localizados en la región central del dominio (160DDD162) por tres alaninas para generar la variante *LiEndoG*^{D160-162A} y evaluamos su actividad nucleasa a distintos pHs (artículo 3, figura 5b). Nuestros resultados demuestran que estos tres residuos son esenciales para la autoinhibición de *LiEndoG*, dado que la variante *LiEndoG*^{D160-162A} muestra básicamente el mismo comportamiento que *LiEndoG* Δ i sobre el rango de pHs estudiado. El hecho de que el reemplazo de estos aspárticos por alaninas de lugar a una proteína que es todavía activa a pH 8.0 refuerza la relevancia que tienen estos tres residuos ácidos en el mecanismo de autoinhibición dependiente de pH en *LiEndoG*.

4.3.4.4 La extensión en el extremo carboxilo terminal de *LiEndoG* determina su preferencia por los sustratos de cadena sencilla

Como hemos comentado anteriormente, tras realizar el alineamiento múltiple de secuencias de *LiEndoG* con sus homólogos en otros organismos, fuimos capaces de detectar la presencia

de cuatro inserciones en nuestra proteína que no tenían una función previa asignada (artículo 3, figura 1). La última de estas inserciones la componen los residuos 440-510 y supone una extensión del extremo carboxilo terminal en *LiEndoG*. Curiosamente, cuando fuimos a modelar esta región en base a la predicción de su estructura secundaria, vimos que el único modelo estructural similar disponible en PDB era el dominio carboxilo terminal (CTD, *carboxyl terminal domain*) de la EXOG humana, incluso a pesar de que estas dos regiones tienen una identidad de secuencia relativamente baja. Partiendo de este modelo, y empleando varias técnicas de refinamiento estructural, en el laboratorio se generó un modelo tridimensional para el dominio CTD de *LiEndoG* compuesto por la agrupación de tres hélices alfa, la primera de las cuales queda localizada de forma muy próxima al surco de unión del sustrato. Esta localización, sumada al hecho de que el CTD en EXOG es responsable tanto de su afinidad por el ADN de cadena sencilla como de su actividad 5'-exonucleasa (117), nos hizo pensar que este dominio podría estar jugando un papel similar en nuestra enzima. Para testar esta hipótesis, generamos una proteína truncada carente de sus 70 últimos aminoácidos (*LiEndoG*^{CTD}) y comparamos su actividad respecto a la proteína silvestre (*LiEndoG*) frente a distintos sustratos.

En primer lugar, digerimos un sustrato de ADN bicatenario con extremos romos marcado en 5' con P³² y separamos los productos de estas reacciones mediante una electroforesis desnaturante en gel de acrilamida (artículo 3, figura 6B). Tal y como esperábamos, los resultados de este experimento muestran que tanto *LiEndoG* como *LiEndoG*^{CTD} son endonucleasas sin especificidad de secuencia, ya que a sus concentraciones más bajas ambas enzimas generan una colección de fragmentos de ADN de tamaños más o menos aleatorios que van desapareciendo progresivamente a medida que aumentamos la concentración de proteína. No obstante, también es posible apreciar algunas diferencias en la forma que tienen las dos enzimas de digerir este sustrato. Una de ellas es el tamaño de los fragmentos de ADN más pequeños que son capaces de generar, siendo de 5, 6, 8 y 10 nucleótidos para *LiEndoG* y de 6 y 10 para *LiEndoG*^{CTD}. No obstante, la diferencia más notable la encontramos en que *LiEndoG* genera rápidamente fragmentos pequeños de ADN (8 nucleótidos) incluso cuando todavía queda una alta proporción de sustrato sin digerir (artículo 3, figura 6B-línea 3), lo cual parece indicar que parte de su actividad está dirigida sobre el extremo 5' del sustrato. Esta preferencia por el extremo 5' fue confirmada tras marcar el sustrato en su extremo 3' y

comprobar que *LiEndoG* generaba mayoritariamente un fragmento de gran tamaño a las concentraciones más bajas de enzima, si bien la longitud exacta de este oligonucleótido no se pudo determinar con los marcadores empleados (resultado no mostrado). De forma contraria, con *LiEndoG*^{CTD} tenemos que esperar a que el sustrato inicial esté digerido casi por completo para poder observar altas concentraciones de oligonucleótidos pequeños, indicando que estos deben surgir como resultado de la digestión progresiva de los productos de reacción iniciales.

Nuestro siguiente paso fue testar la actividad de estas enzimas sobre ADN de doble cadena (dc) y cadena sencilla (cs), para lo cual seleccionamos un sustrato con una región central de 48 pares de bases flanqueada en sus extremos por oligonucleótidos de cadena sencilla de tamaño variable (artículo 3, Figura 6C). Nuestros resultados muestran que, al menos en este tipo de sustratos, *LiEndoG* tiene una afinidad extraordinaria por el ADNcs, no pudiéndose observar la aparición de fragmentos correspondientes a la digestión sobre la región de ADNdc. Sin embargo, esta afinidad se pierde cuando eliminamos el CTD, ya que la variante *LiEndoG*^{CTD} da lugar a un patrón de digestión muy similar al observado con el sustrato de extremos romos anterior. Adicionalmente, pudimos comprobar que *LiEndoG* tiene una clara preferencia por digerir este sustrato en las posiciones 7 y 11, ya que estos son los fragmentos mayoritarios que observamos a la concentración más baja de enzima (de entre las que son capaces de digerir la sonda). Dado que la posición 11 marca exactamente el primer nucleótido de la región de ADNdc, este resultado indica que *LiEndoG* es capaz de eliminar los oligonucleótidos de cadena sencilla u *overhangs* que sobresalen en 5' del ADNdc, siendo esta una actividad que típicamente llevan a cabo las flap endonucleasas. Posteriormente, este fragmento de 11 nucleótidos es degradado rápidamente hasta generar hexa y pentanucleótidos, siendo aparentemente este el tamaño crítico de sustrato de cadena sencilla que es capaz de digerir *LiEndoG*. Este resultado está en consonancia con lo predicho por nuestro modelo, ya que los oligonucleótidos de menor tamaño pierden algunas de las interacciones más relevantes con la enzima, dificultando que se unan a esta con suficiente afinidad.

Por último, para comprobar de forma inequívoca si el CTD es el responsable de la actividad preferente de *LiEndoG* sobre el ADNcs, utilizamos como sustrato un ADN de solo 14 nucleótidos incapaz de generar estructuras bicatenarias (artículo 3, figura 6D). Nuestros

resultados muestran que *LiEndoG* Δ^{CTD} tiene tan baja afinidad por este sustrato que es prácticamente incapaz de procesarlo, pudiéndose solo observar cierta digestión a la mayor concentración de enzima empleada. Por el contrario, *LiEndoG* digiere este ADNcs rápidamente, confirmando así que el dominio CTD es determinante para la afinidad de esta enzima por los sustratos de cadena sencilla.

En conjunto, nuestros resultados muestran de forma inequívoca que el dominio carboxilo terminal de *LiEndoG* es responsable de la preferencia que tiene esta enzima por digerir ADN de cadena sencilla, ya sea de forma aislada o formando parte de un híbrido de ADNcs/dc. Este resultado es similar al reportado para EXOG, donde la eliminación del dominio CTD o el reemplazo de la Gly277 (próxima a este dominio) por una Val da lugar a variantes con mucha menor actividad sobre el ADNcs, indicando la relevancia que tiene tanto la presencia como el posicionamiento de este dominio sobre su actividad catalítica (117). Adicionalmente, en este mismo trabajo se muestra como la actividad 5'-exonucleasa de EXOG, que da lugar a la aparición de mono y dinucleótidos como producto de digestión, desaparece cuando la hidrólisis es llevada a cabo con las dos variantes mencionadas anteriormente (117). Cabe destacar que *LiEndoG*, aun teniendo un CTD muy similar al de EXOG, no es capaz de generar estos mono y dinucleótidos, por lo que no podemos considerarla como una exonucleasa "clásica". No obstante, es importante recordar *LiEndoG* fue capaz de generar rápidamente oligonucleótidos de pequeño tamaño (8 nucleótidos) al digerir el sustrato de doble cadena, por lo que sí parece tener cierta polaridad 5'. Esta preferencia por digerir rápidamente los extremos 5' para generar oligonucleótidos de pequeño tamaño se vio corroborada por experimentos en los que se marcó el extremo 3' de una de las dos hebras complementarias y se comprobó la aparición de un fragmento de gran tamaño resultante de un acortamiento inicial del extremo 5'. En cualquier caso, parece claro que tanto en EXOG como en *LiEndoG* la preferencia por los sustratos de cadena sencilla está dictada por la presencia de esta extensión carboxilo terminal. De hecho, tras la eliminación de este dominio, *LiEndoG* Δ^{CTD} se comporta básicamente como la EndoG humana o CPS-6, endonucleasas sin preferencia por el ADNcs y que carecen de forma natural de la extensión carboxilo terminal (121). Este comportamiento puede entenderse si tenemos en cuenta que las tres hélices del dominio CTD, cuando se mantienen en su posición con la ayuda de la formación de un puente disulfuro entre la Cys 434 y Cys439, están organizadas de forma que se posicionan próximas al centro

activo, dificultando la unión del ADNdc. Sin embargo, el paso del ADNcs sigue siendo posible, de ahí la gran afinidad que muestra *LiEndoG* tanto por este sustrato como por las uniones ADNcs/ADNdc. Por el contrario, cuando el CTD no está frente al sitio activo, ya sea porque se ha desplazado como consecuencia de un movimiento tipo bisagra del cuerpo principal de la proteína o porque ha sido eliminado, *LiEndoG* es capaz de unir sustratos de doble cadena y comportarse como una endonucleasa clásica.

Por último, el hecho de que este dominio CTD no esté presente ni en *DmEndoG* ni en CPS-6 (enzimas carentes de actividad exonucleasa), sumado a que se cree que la actividad 5'-exonucleasa de EXOG es esencial para la reparación y recombinación del ADNmt (148), nos hace pensar que *LiEndoG* puede actuar como una enzima exonucleolítica y/o tener la capacidad de cortar moléculas de ADN ramificadas con fragmentos de cadena sencilla como los que surgen en los procesos de replicación, recombinación y reparación del ADN (177). Adicionalmente, esta capacidad de *LiEndoG* de eliminar los 5'-*overhangs* del ADNdc nos hizo darnos cuenta de que nuestra enzima también guarda cierta similitud con la proteína CRN-1, proteína homóloga a la flap endonucleasa 1 (FEN-1) en *C. elegans*. CRN-1 es una enzima nuclear con actividad 5'-exonucleasa, flap-endonucleasa y 5'-*gap* endonucleasa a la que habitualmente se ha asociado con procesos de reparación y replicación del ADN. Adicionalmente, se ha mostrado que CRN-1 interacciona y coopera con CPS-6 durante el proceso de muerte celular para llevar a cabo de forma conjunta la fragmentación del ADN. De hecho, se ha propuesto que la cooperación de estas enzimas es crítica para pasar del estado de replicación/reparación al de degradación del ADN durante la apoptosis (136). De forma similar, se ha demostrado que *LdEndoG* y *LdFEN1* coinmunoprecipitan en los extractos nucleares de *Leishmania donovani* cuando estos parásitos son tratados con un inhibidor de la topoisomerasa IB (100). Si bien es cierto que aún no se sabe si *LiEndoG* y *LiFEN1* también cooperan en *Leishmania infantum*, nuestros resultados indican que *LiEndoG* por sí sola podría tener una actividad equivalente a la de CPS-6 y CRN-1 actuando juntas.

En conclusión, nuestro trabajo muestra que *LiEndoG* tiene un tamaño mucho mayor al de otras EndoGs (101) debido a la presencia de inserciones en su secuencia que no existen fuera de la familia de los tripanosomátidos y que le dotan de propiedades prácticamente únicas. En este trabajo, hemos identificado y caracterizado el papel de dos de estas inserciones, siendo la primera de ellas un dominio de autoinhibición que modula la actividad de la enzima en

función del pH y la última una extensión del dominio carboxilo terminal que le confiere selectividad por el ADN de cadena sencilla, tanto por si solo como cuando está formando parte de una unión ADNcs/ADNdc. Por consiguiente, nuestros resultados muestran que *LiEndoG* está dotada de una especificidad estructural que le permite cortar preferentemente en la base de un saliente de ADNcs de una forma similar como lo hacen las flap endonucleasas (178). Adicionalmente, el modelo generado en nuestro laboratorio sugiere que las dos inserciones restantes podrían estar implicadas en la unión del sustrato y la formación de interacciones con otras proteínas, si bien estas propuestas no han sido contrastadas experimentalmente todavía. No obstante, este estudio allana el camino para la caracterización de estos y otros dominios y nos motiva a buscar las posibles proteínas con las que puede estar interaccionando *LiEndoG*.

Structure-based domain assignment in *Leishmania infantum* EndoG: characterization of a pH-dependent regulatory switch and a C-terminal extension that largely dictates DNA substrate preferences

Cristina Oliva¹, Pedro A. Sánchez-Murcia², Eva Rico¹, Ana Bravo², Margarita Menéndez³, Federico Gago^{2,*} and Antonio Jiménez-Ruiz^{1,*}

¹Departamento de Biología de Sistemas, Universidad de Alcalá, E-28805 Alcalá de Henares, Madrid, Spain,

²Departamento de Ciencias Biomédicas y "Unidad Asociada IQM-CSIC", Universidad de Alcalá, E-28805 Alcalá de Henares, Madrid, Spain and ³Instituto de Química Física Rocasolano, Consejo Superior de Investigaciones Científicas (CSIC), E-28006 Madrid, Spain

Received April 07, 2017; Revised July 06, 2017; Editorial Decision July 07, 2017; Accepted July 11, 2017

ABSTRACT

Mitochondrial endonuclease G from *Leishmania infantum* (LiEndoG) participates in the degradation of double-stranded DNA (dsDNA) during parasite cell death and is catalytically inactive at a pH of 8.0 or above. The presence, in the primary sequence, of an acidic amino acid-rich insertion exclusive to trypanosomatids and its spatial position in a homology-built model of LiEndoG led us to postulate that this peptide stretch might act as a pH sensor for self-inhibition. We found that a LiEndoG variant lacking residues 145–180 is indeed far more active than its wild-type counterpart at pH values >7.0. In addition, we discovered that (i) LiEndoG exists as a homodimer, (ii) replacement of Ser211 in the active-site SRGH motif with the canonical aspartate from the DRGH motif of other nucleases leads to a catalytically deficient enzyme, (iii) the activity of the S211D variant can be restored upon the concomitant replacement of Ala247 with Arg and (iv) a C-terminal extension is responsible for the observed preferential cleavage of single-stranded DNA (ssDNA) and ssDNA–dsDNA junctions. Taken together, our results support the view that LiEndoG is a multidomain molecular machine whose nuclease activity can be subtly modulated or even abrogated through architectural changes brought about by environmental conditions and interaction with other binding partners.

INTRODUCTION

Sugar nonspecific endonucleases belong to a family of enzymes that cleave single- and double-stranded DNA (ssDNA and dsDNA, respectively) and/or RNA molecules with little or no sequence specificity. Among them, endonuclease G (EndoG) is highly conserved in eukaryotes (1) and is targeted to the mitochondrial intermembrane space from where it is released in response to a death stimulus. Upon subsequent translocation to the nucleus, EndoG is thought to participate in the process of DNA degradation that is associated with programmed cell death (2).

Organisms have sophisticated mechanisms for genome protection against self-destructive nuclease activity. Although EndoG is usually silenced by compartmentalization into mitochondria, additional protective strategies have evolved in nature, such as activation via cross-linking of cysteine residues, as in the nucleases secreted by *Serratia marcescens* (3) and *Vibrio vulnificus* (4), or active site blockade, as exemplified by the specific protein inhibitors co-crystallized with nucleases from *Anabaena spiroides* (5) and *Drosophila melanogaster* (6). These mechanisms avoid irreversible damage caused by stray EndoG.

In the last decade, the 3D structures of several eukaryotic endo/exo-nucleases from different organisms have been solved and deposited in the Protein Data Bank (PDB) (7). From the structural point of view, the common core of these metallonucleases is made up of two antiparallel β -sheets and one α -helix that contains an essential and strictly conserved asparagine residue whose carboxamide oxygen is used to coordinate a divalent metal ion. This cation is usually magnesium although manganese or, in the case of human EXOG (8), calcium is also catalytically competent.

*To whom correspondence should be addressed. Tel: +34 918 855 109; Fax: +34 918 854 585; Email: antonio.jimenez@uah.es
Correspondence may also be addressed to Federico Gago. Tel: +34 918 854 514; Fax: +34 918 854 591; Email: federico.gago@uah.es

We previously showed that the protozoan parasite *Leishmania infantum* expresses a protein, *LiEndoG*, with structural and functional characteristics that strongly resemble those of endonucleases G from higher eukaryotes, including mitochondrial localization and migration to the nucleus in response to a death stimulus (9). More recently, we demonstrated that *LiEndoG* also has a pro-life role in the parasite, as indicated by the reduced growth rates observed in *LiEndoG* hemi-knockouts and the cytotoxicity of the *LiEndoG* inhibitor *Lei49* on promastigotes and amastigotes, which can be counteracted by *LiEndoG* overexpression (10).

Herein, guided and supported by an atomic three-dimensional model of the homodimeric enzyme, we seek a rationale for the occurrence of the SRGH motif in *EndoGs* from trypanosomatids in place of the canonical DRGH and pursue the assignment of specific functions to several discrete domains that appear to be unique to these enzymes. Two of these assigned functions are validated in *LiEndoG* through the experimental generation and biochemical characterization of several protein variants whose enzymatic activities on a variety of substrates are compared to that of the wild-type enzyme.

MATERIALS AND METHODS

Molecular modeling of *LiEndoG*

The dimers observed in the crystal structures of *Drosophila melanogaster* *EndoG* (*DmEndoG*, PDB id. 3ISM) (6), the inactive H148A (11), H148A/F122A (12) and H148A/Q130K (13) variants of *Caenorhabditis elegans* *EndoG* (*CeEndoG*, a.k.a. CED-3 protease suppressor-6 [CPS-6], PDB ids. 3S5B, 5GKC and 4QN0, respectively), and human mitochondrial endo-exonuclease G (EXOG, PDB id. 4A1N) (11) are considered to be valid models for eukaryotic *EndoGs* because of the high conservation, from yeast to human, of many residues that are crucial for both activity and dimerization (6,14). Comparison and inspection of the incomplete homology models provided by the threading methods implemented in the Phyre2 (15,16) and Swiss-Model (17) web servers allowed us to construct a consensus ‘gapped’ model comprising residues 99–144, 181–264, 301–315 and 364–439 of *LiEndoG* (UniProtKB entry B0ZTD3) that was elaborated further with the aid of additional iterative and incremental modeling, as described below. Although human EXOG has been biochemically characterized in detail (18), the three-dimensional coordinates deposited in the PDB for this protein remain unpublished and suffer from some uncertainties. For this reason, we refined them against the accompanying electron density map using Coot (19) and Phenix (20,21). The resulting model (Supplementary Figure S1) was then used as a template by the one-to-one threading methods implemented in the Phyre (15) and SwissModel (17) servers to generate a consensus 3-helix bundle for the C-terminal domain (CTD) of *LiEndoG*. Structure visualization, molecular editing, and figure preparation were performed with PyMOL (22).

No suitable templates were found for building a homology model of inserts 1–3 (Figure 1) but the algorithms (23,24) implemented in Robetta (25) were able to produce a

reasonable structure for *insert 1* (residues 145–180), which comprises the (Asp/Glu)-rich region exclusive to nucleases from trypanosomatids, and we tentatively modeled the polyglutamine tract in *insert 2* as an α -helix (26,27). The flap in *insert 3*, one β -strand of which is proposed to form an intermolecular β -sheet in the dimer, was constructed by following an incremental approach that took into account the putatively equivalent positioning at the tip of the β -hairpin of ³²¹RPDG³²⁴ in *LiEndoG* and ²¹⁴EGDG²¹⁷, ²²¹EDDG²²⁴ and ²⁰⁵RGDG²⁰⁸ in *CeEndoG*, *DmEndoG* and EXOG, respectively. The Random Coordinate Descent (RCD) loop closure algorithm (28) was used profusely for the *ab initio* modeling of the remaining gaps.

The two putative disulfide bonds between residues 100–116 and 434–439, earlier predicted by the DISULFIND server (29) on the basis of the primary sequence (9), and now found to be feasible in 3D given the spatial closeness and suitable side-chain orientation of the two cysteines making up each pair (Figure 1), were incorporated into the structural model of each monomer. The catalytic Mg²⁺ ion and the water molecules completing its octahedral solvation shell were introduced in agreement with the geometries observed in both *DmEndoG* (6) and EXOG (11), which are coincident with that described previously for our simplified model of the active site of *LiEndoG* (9) based on *Vibrio vulnificus* nuclease (*Vvn*) (30). Finally, all the domains were integrated into a *LiEndoG* dimer that was visually and computationally inspected. Minimal manual adjustments alleviated some steric clashes involving some side-chain rotamers and protein backbone atoms. The resulting protein structure was then refined in the AMBER force field (31) by carrying out, essentially as described (30), a series of C α -restrained energy minimizations followed by molecular dynamics (MD) simulations at 300 K in a box of TIP3P water molecules and counterions using the *pmemd* module of the AMBER 14.0 suite of programs (31).

The complexes of wild-type *Vvn* and its inactive H80A variant with a dsDNA octamer (32) and a dsDNA 16-mer (33), respectively, as well as the complex between the H148A/F122A variant of *CeEndoG* and a 5'-TTTTTGT-3' ssDNA (12), provided the structural information necessary for building the initial *LiEndoG*:DNA complexes prior to refinement using the same molecular mechanics methods that were employed on the protein alone. In this respect, further crystallographic refinement of the *CeEndoG*(H148A/F122A):ssDNA complex (PDB code: 5GKP) against the accompanying electron density map using Coot (19) and Phenix (20,21) allowed us to unambiguously trace most of the intact octanucleotide chains bound to both subunits and place all the bases in the right register and correct positions (Supplementary Figure S2 and Table S1). This step was instrumental in deriving a reliable molecular model of *LiEndoG*:ssDNA in which the residues that interact with the phosphate backbone can be unequivocally identified (Table 1). As a template for our ssDNA–dsDNA junction we used the single-crystal structure of a DNA Holliday junction model (34) (PDB code: 4GS2) and changed base composition by means of the *ad hoc* utility implemented in Coot (19).

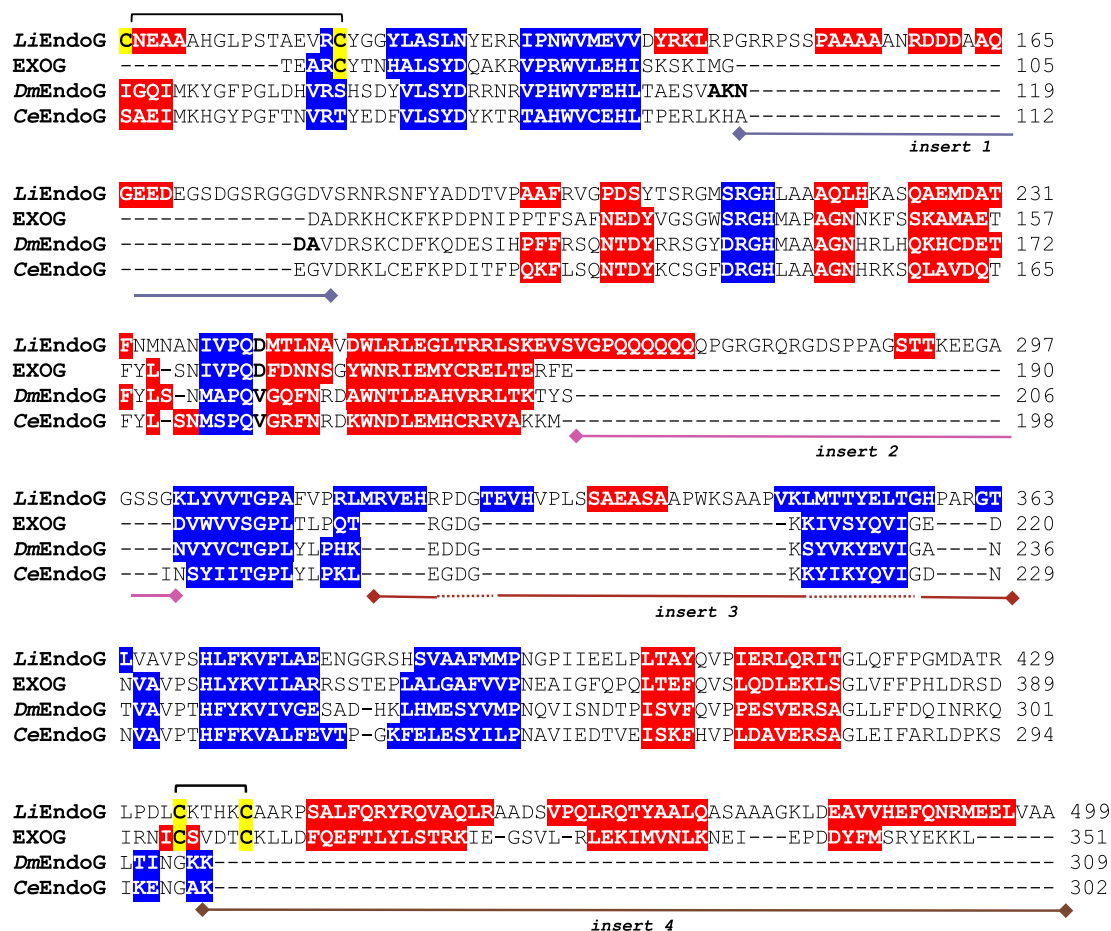


Figure 1. Structure-based multiple sequence alignment, as provided by the updated Dali server (70), of our *LiEndoG* model against the three most suitable templates found in the Protein Data Bank. α -helices and β -strands are boxed in a red or blue background, respectively. Note the presence of the four insertions ($\blacktriangleleft\blacktriangleright$) in *LiEndoG* for which the only counterpart in the other three enzymes is the EXOG CTD, which shows some homology to *insert 4*. The two rotated square brackets connecting the cysteines boxed in a yellow background represent the two putative disulfide bonds. The AKNDA sequence from *DmEndoG* highlighted in bold was used to replace the putative self-inhibitory domain (*insert 1*) in the *LiEndoG* Δ i variant.

Table 1. Protein atoms involved in phosphate recognition of the oligonucleotides to be cleaved by *CeEndoG*, *EXOG* and *LiEndoG*^a

Nuclease	DNA phosphodiester bond																	
	-5		-4		-3		-2		-1		0		+1		+2		+3	
	OP1	OP2	OP1	OP2	OP1	OP2	OP1	OP2	OP1	OP2	OP1 ^b	OP2	OP1	OP2	OP1	OP2	OP1	OP2
<i>CeEndoG</i>	—	—	—	—	—	—	(NH1 R192) ^c	NH2 R156	NH G153 NH2 R192	NH1 R117	Mg ²⁺	NH2 R117	NE R146	NH2 R146	NH C141	NZ K140	—	NZ K140
<i>EXOG</i>	[L302] ^d	[E184]	[L309]	[L309]	[E317]	[L302]	NH1 R183	NE R314 NK148	NH G145 NH2 R183	NH1 R109	Mg ²⁺	NH2 R109	NE R138	NH2 R138	NH G133	[V132]	—	[V132]
<i>LiEndoG</i>	(NE R442)	(NH1 R259)	(NH1 R449)	(NH2 R442)	(NH2 R457)	(NH2 R442)	NH2 R258	OH Y450	NH Q219	NH1 R183	Mg ²⁺	NH2 R183	NE R212	NH2 R212	NH S207	(OG1 T206)	—	(OG S207)
									NH2 R258									

^aThe reference frame for numbering the phosphodiester bonds is upstream (–) and downstream (+) of site 0 (i.e. the cleavage site, which is defined as the bond linking the Mg²⁺-bound O3' of one nucleotide to the P atom of the next nucleotide). OP1 and OP2 are the phosphate oxygens involved in recognition.

^bThe Mg²⁺ ion is included for completeness and as an aid for establishing the frame of reference.

^cResidue names in brackets are thought to be involved in phosphate recognition but these particular contacts are not observed in the reference crystal structure.

^dResidue names in square brackets are included for completeness and comparative purposes but they cannot establish a favorable interaction with a phosphate group.

Expression of wild-type *LiEndoG* and engineered variants

Escherichia coli Rosetta (DE3) pLysS bacterial cells were transformed with pET28a-endog. Expression of *LiEndoG* was then induced in LB medium at 0.6 OD_{600nm} with 1 mM IPTG at 37°C during 3 h. Because most of the expressed protein was found in the non-soluble fraction, *E. coli* cells were lysed in a pH 8.0 buffer containing 8 M urea, 100 mM NaH₂PO₄, 10 mM imidazole, and 10 mM Tris. The cleared lysate was loaded onto a HisTrap Ni-NTA resin column (GE Healthcare) in which the protein was refolded by progressive removal of urea prior to elution with 500 mM imidazole.

The DNA coding for the *LiEndoG* variant lacking amino acids 440–510 (*LiEndoG*Δ^{CTD}) was obtained by PCR amplification of the *LiEndoG* gene sequence from its start position up to the nucleotides encoding amino acid 439. This PCR-amplified gene was cloned into the BamHI/XhoI site of the expression vector pET-28a.

For *LiEndoG*Δ_i construction, the sequence corresponding to the putative self-inhibitory domain encoding amino acids 145–180 was replaced by nucleotides encoding amino acids AKNDA to make this region resemble the positionally equivalent region of *DmEndoG* (see Figure 1). This was achieved by two separate PCR amplifications followed by a fusion PCR via overlap sequences. The fusion construct was also cloned into the BamHI/XhoI site of the expression vector pET-28a.

The S211A, S211D, H214A, R212A, S211D/A247R and D(160–162)A variants were generated using the QuikChange XL-Site-Directed mutagenesis kit (Agilent) and oligos that were designed following the manufacturer's instructions. The presence of the desired mutations was checked by DNA sequencing.

Fluorescence-based nuclease activity assay

The enzymatic activity of *LiEndoG* was monitored by measuring the increase in fluorescence obtained upon digestion of a dsDNA probe constructed by hybridization of the oligonucleotides FAM-5'-CTGTCGCTACCTGTGG-3' and 5'-CCACAGGTAGCGACAG-3'-TAMRA. The intrinsic fluorescence of 6-carboxyfluorescein (FAM) is quenched by the proximity of tetramethylrhodamine (TAMRA) in the undigested probe. Digestion of any of the strands causes separation of the fluorophore from the quencher which is apparent by emission of a fluorescent signal. The dsDNA probe (30 pmol) was digested with the different nucleases (1 ng/μl) in 10 mM NaAc, 10 mM MES, 20 mM Tris-HCl and 10 mM MgCl₂. The pH value was adjusted to 6.5 unless otherwise indicated. The reactions were monitored either in a Victor 1420 Multilabel Counter (Wallac) or in an EnSpire Microplate Reader, at excitation and emission wavelengths of 492 nm and 517 or 495 and 520 nm, respectively.

Nuclease activity assay on supercoiled circular DNA

pRSETa plasmid DNA (0.3 μg) was digested in a final volume of 20 μl pH 6.5 buffer containing 22.5 mM Tris, 50 mM NaCl, MES 10 mM, NaAc 10 mM, 2.5 mM MgCl₂ and 5% glycerol during 1 h at 16°C using the amounts of enzyme

indicated in the figure. The digested DNA was resolved on 1.0% (w/v) agarose gels, stained with ethidium bromide and visualized under UV light.

Activity assays using radioactively labeled oligonucleotide substrates

The ssDNA oligonucleotides shown in Table 2 were purchased from Invitrogen and radioactively labeled at the 5'-end with [γ -³²P]-ATP using polynucleotide kinase (Thermo Scientific). Unincorporated radionucleotides were removed using the QIAquick Nucleotide Removal kit (Qiagen). To generate the labeled dsDNA substrates, each corresponding radioactively labeled ssDNA was incubated in a 1:1.25 molar ratio with its complementary strand in 20 mM Tris-HCl (pH 7.5) and 50 mM NaCl. Annealing was achieved by heating the solution at 95°C for 5 min and then cooling down slowly (-1°C per minute) to 25°C.

Enzymatic cleavage assays were done at 37°C in a pH 6.5 buffer containing 22.5 mM Tris, 50 mM NaCl, MES 10 mM, NaAc 10 mM, 2.5 mM MgCl₂ and 5% glycerol in a final volume of 10 μl. Cleavage products were resolved by 7 M urea 20% polyacrylamide gel electrophoresis (PAGE). DNase I and exonuclease T7 (Thermo Scientific) were used as positive controls of endonuclease and exonuclease activities, respectively.

Analytical ultracentrifugation

Sedimentation velocity experiments were run at 45 000 rpm using cells with double sector Epon-charcoal centrepieces. Differential sedimentation coefficients were calculated by least-squares boundary modeling of the experimental data with the program SEDFIT. The experimental values were corrected to the viscosity and density of water at 20°C ($s_{20,w}$) (35). Sedimentation equilibrium experiments were carried out at different rotor speeds, from 6500 to 11 000 rpm, as described previously (36), and weight-average molecular weights ($M_{w,app}$) were calculated using the HETEROANALYSIS program, version 1.1.2 (<http://www.biotech.uconn.edu/auf/>). All measurements were performed in 50 mM MES, 500 mM NaCl, 0.4 mM KCl buffer (pH 6.5) at 20°C using protein concentrations ranging from 0.07 to 0.77 mg/ml in an Optima XL-A analytical ultracentrifuge (Beckman Coulter) with an AN50-Ti rotor. The partial specific volume of *LiEndoG*Δ^{CTD} and the buffer's density and viscosity were estimated from the amino acid sequence and the buffer composition, respectively, using the SEDNTERP program (37).

RESULTS

An updated *LiEndoG* molecular model

Our initial homology model for *LiEndoG* (9) afforded enough detail about the active site environment but provided a reliable structure only for the catalytic core of a single monomer. To follow up on this work, we elaborated on our model further and then focused on four conspicuous regions for which no functional assignment has been reported yet: 145–180 (*insert 1*), 265–300 (*insert 2*), 316–363 (*insert 3*) and 440–510 (*insert 4*). *Insert 1*, which appears to be unique

Table 2. Oligonucleotides used for radioactive labeling on the 5'-end

dsDNA probe ^a
forward 5'- GCCGCTAACAGAGACGCGGAGGCCGCGCAGGGTGAAGAAGAC -3'
reverse 5'- GTCTTCTTCACCCCTGCGGGCCCTCCGCGTCTCTGTTAGCGGC -3'
ssDNA probe
5'-CGCCAGATCTTCCC-3'
dsDNA with ssDNA overhangs
forward 5'-
TTTTTTTTTTTTTTTTTTTTTTGGTTAATACGACTCACTATAGGAGAACCCTTAAAGGTTTAACTTTAAAGACCCCCCCCCCCCCCCCCCCCCCCC-3'
reverse 5'-TTTTTTTTTTGGGTCTTAAAGTTAAACCTTAAAGGTTCTCTATAGTGAGTCGTATTAA-3'

^aNucleotides involved in base-pair formation are highlighted in bold.

to trypanosomatids (Figure 1), contains numerous aspartic and glutamic acid residues that are clustered into two distinct regions. *Insert 2* comprises a Q,G-rich region that is identified as part of a compositionally biased (38) and intrinsically disordered domain (39). *Insert 3* represents an expansion of a β -hairpin region at the dimerization interface that is much shorter in the reference templates. *Insert 4* is a C-terminal extension that, of all available templates, is only present, albeit with low sequence identity, in the CTD of human EXOG. Our updated model is now a homodimer that incorporates all insertions and provides a structural rationale to the functions experimentally assigned to inserts 1 and 4.

Active site signature. The active site of *LiEndoG* consists of a $\beta\beta\alpha$ -metal finger that contains an SRGH signature (9) in place of the canonical DRGH ProSite motif (PDOC00821) characteristic of most other members of this family of nucleases (4). In the apo state, the catalytic Mg^{2+} ion is coordinated directly by the carboxamide oxygen of the Asn246 side chain and five water molecules, four of which are hydrogen bonded to the carboxylate oxygens of Glu254, the imidazole N δ of His214 and the side-chain carboxamide of Gln241 (9).

Dimerization domain. The 3D structural superposition of *DmEndoG*, *CeEndoG* and EXOG demonstrates that the interface residues involved in dimer stabilization are highly conserved in these three enzymes. Since a multiple sequence alignment shows that many of these residues are also present in *LiEndoG* (Figure 1) we considered *LiEndoG* to be, in all likelihood, a homodimer as well. This view is strongly supported by results from analytical ultracentrifugation proving that the biologically active form of a *LiEndoG* lacking only the CTD (*LiEndoG* Δ^{CTD}) is indeed a homodimer (Figure 2). Given this premise, it is noteworthy that the two active sites are displayed, as is true for the template dimeric enzymes as well, on opposite sides of the dimerization interface (Figure 3; *LiEndoG*.apo.pdb in Supplementary material).

Disulfide bridges. Our updated *LiEndoG* molecular model also gave us additional confidence on the existence—for which there is no precedent in any of the other nucleases used as templates—of the two disulfide bridges postulated earlier (9), one between Cys100 and Cys116 and another linking Cys434 and Cys439 (Figure 1). With respect to the latter, which is just upstream of the CTD, the positionally equivalent thiol groups of Cys294

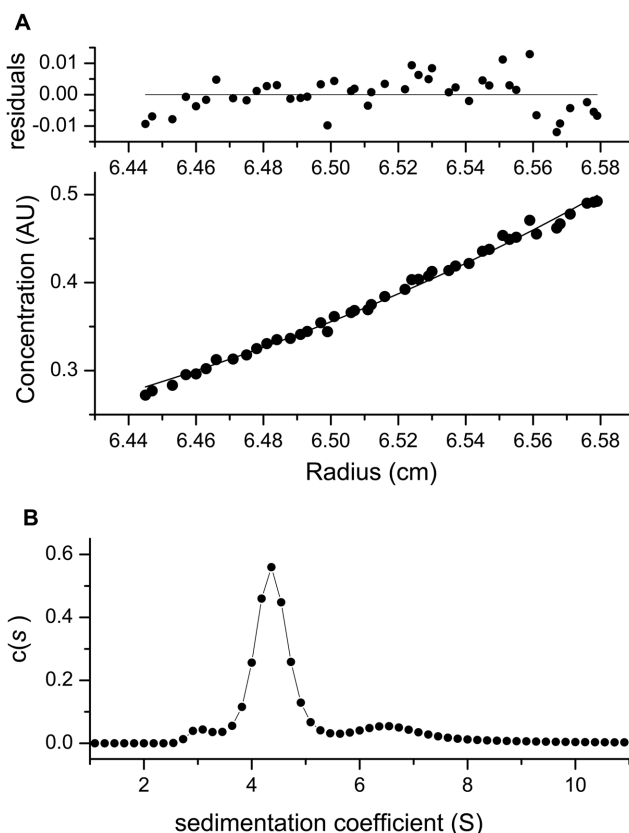


Figure 2. *LiEndoG* quaternary structure. (A) Sedimentation equilibrium profile of *LiEndoG* (0.77 mg/ml) in 50 mM MES, 500 mM NaCl, 0.4 mM KCl buffer (pH 6.5) at 8,000 rpm and 20°C. The lower part shows the experimental data (circles) and the best fit (solid line) to a single dimeric species of 91 ± 2 kDa. The upper part shows residuals of the theoretical fit. (B) Distribution of experimental sedimentation coefficients of the same sample in sedimentation velocity experiments (45 000 rpm, 20°C). Data shown are representative of at least four independent experiments.

and Cys299 in PDB entry 4A1N appear close enough for a disulfide bond to form under oxidative conditions but this EXOG protein was expressed in the reducing environment of the *E. coli* cytoplasm (40,41) and no details about the experimental procedures are publicly available (11). Remarkably, the amino acid sequence of yeast endonuclease G Nuc1p (*ScNuc1p*) (42,43) also shows two cysteines in equivalent positions (Supplementary Figure S3), which suggests that the existence of a disulfide bond between these two residues is highly plausible and may play a functional role. In any case, it must be noted that all of

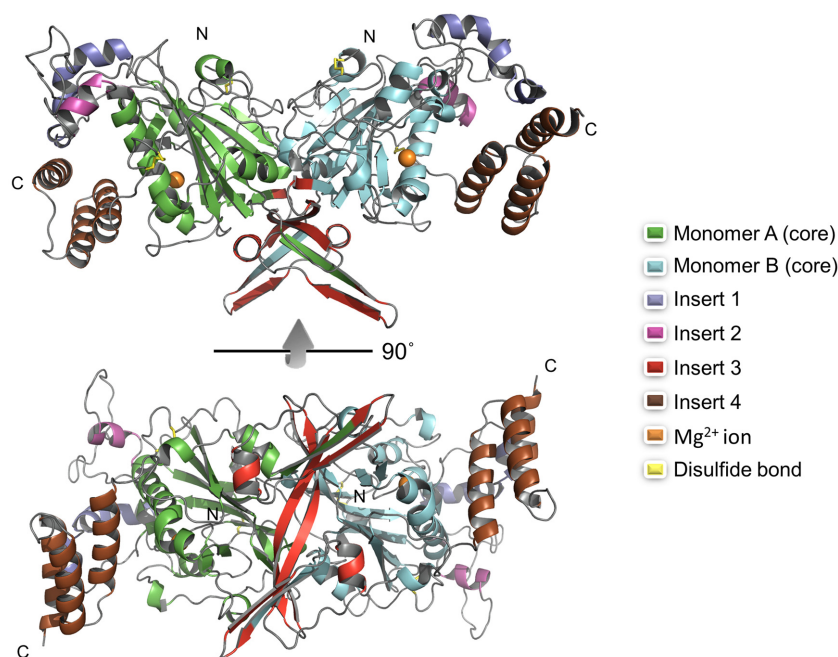


Figure 3. Molecular model of homodimeric *LiEndoG*. Two side views differing in a 90° rotation, as indicated, are shown. The central cores of the two monomers are colored in green and cyan. The Mg^{2+} ions bound in each active site, close to the slightly distorted α -helix, are displayed as orange spheres. Inserts 1, 2, 3 and 4 in both monomers are colored in blue, purple, red and brown, respectively. The putative disulfide bonds are displayed as yellow sticks. Note the extended interfacial β -strands giving rise to a long antiparallel β -sheet in the dyad axis.

these postulated cystines are intrasubunit and we do not find any structural clue pointing to the possible existence of intermonomeric crosslinks in *LiEndoG*.

Insert 1 (residues 145–180). *DmEndoG*'s activity is known to be negatively regulated by an inhibitory protein (*DmEndoGI*) that binds to it with subpicomolar affinity and consists of two highly homologous domains linked by a polypeptide chain. Similarly, the activity of the cyanobacterial, nonspecific and extracellular nuclease NucA is kept in check by the inhibitor NuiA. The structural resolution of *DmEndoG*:*DmEndoGI* (6) and NucA:NuiA (5) complexes by X-ray crystallography (PDB entries: 3ISM and 2O3B, respectively) has provided valuable insights into the binding determinants and mechanisms of inhibition, which are totally unrelated. Thus, *DmEndoGI* docks a loop region rich in acidic residues into the oligonucleotide-binding groove of *DmEndoG* to form a plethora of salt bridges with Arg side-chains (6) whereas NuiA makes use of its C-terminal threonine to directly coordinate to the catalytic Mg^{2+} cation (5,44). To the best of our knowledge, no such inhibitors have been described in *Leishmania* but the presence of this insertion in the *LiEndoG* sequence (Figure 1) made us suspect that this region could adopt a definite secondary structure (Figure 3) which, under certain conditions, might block the active site and prevent binding of the negatively charged polynucleotides. The modeling of this ~30-aa segment as a candidate domain for self-regulation of the nuclease activity was far from trivial because no suitable templates could be found in the whole PDB. Nonetheless, the algorithm implemented in the Robetta web server (25) filled in the ¹⁴⁵Pro–Asp¹⁸⁰ gap and provided us with a suitable geometry for this partially disordered region, which likely in-

cludes a short α -helix encompassing the polyalanine expansion (45).

Insert 2 (residues 265–300) represents a sequence stretch which extends from the C-terminus of the α -helix that contains the catalytic Asn246 up to the N-terminus of a β -strand that makes up the long central intermonomeric β -sheet. We tentatively modeled the polyglutamine tract as an α -helix (26,27), filled in the gap (28), and relaxed the conformation of the resulting loop by means of a restrained MD simulation. This region is spatially close to *insert 1* (Figure 3) and contains a positively charged ²⁷⁷GRGRQR²⁸² segment that we think provides affinity for polyphosphate binding.

Insert 3 (residues 316–363) is found within a region that is structurally conserved in all homologous enzymes used as templates (Figure 1). Thus, EXOG, *DmEndoG* and *CeEndoG* show a common 'flap' or 'wing' that protrudes from the globular domain of each monomer and is located at one end of an antiparallel β -sheet, the longer β -strand of which (residues 347–358) contributes to the dimerization interface by forming an intermolecular β -sheet (Figure 3). Significantly, in *LiEndoG* the Thr352 residues from each subunit come close together midway on one side of this β -sheet whereas the ends appear to be stabilized by a salt bridge between Lys348 in one monomer and Glu354 in the other. This electrostatic interaction seems to be highly relevant for dimer stabilization because it is also conserved in homodimeric *DmEndoG* (6) whereas the positionally equivalent residues in human EXOG (18) and *CeEndoG* are Lys209:Gln215 and Lys218:Gln224, respectively, also allowing for the formation of hydrogen bonds. On the other side of this β -sheet, the Tyr353 residues from both subunits

also meet, and their phenol rings are involved in strong hydrogen bonding interactions with ion-pairing side-chains protruding from the ¹²⁷ERR¹²⁹ loop, which also appears to be key for dimerization.

On the basis of the multiple sequence alignment shown in Figure 1, we predict that residues 314–328 in *LiEndoG* make up an extended β -hairpin. A sequence similarity search using FASTA (46) on the UniProt Knowledgebase (47) revealed that the amino acid stretch immediately downstream is unique to *Leishmania*. To model this region, we reasoned that the ³²¹RPDG³²⁴ stretch must be located at the tip of the hairpin because both Asp323 and Gly324 are structurally conserved in *DmEndoG*, *CeEndoG* and human EXOG as the central residues in the distinct flaps (or ‘wings’) that, together in the dimer, resemble a saddle. We believe that this turn is potentially involved in crucial protein-protein interactions because a similar loop with an identical ¹⁸RPDGT²² sequence at the tip is also present in the cytosolic protein PYM (PDB entry: 1RK8) (48), in which it defines the effector-binding domain to the core of the exon junction protein Mago-Y14 that participates in postslicing mRNA metabolism (49).

Interestingly, placing ³²¹RPDG³²⁴ in *LiEndoG* at the turn of the β -hairpin suggested to us that ³¹⁶MRVEH³²⁰ and ³²⁵TEVHV³²⁹ could also be part of the antiparallel β -strand. By carrying out this modeling work, we managed to reduce the length of sequence without an assigned structure. Next, we realized that ³³³SAEASAAPWKSA³⁴⁴ in *insert 3* is highly similar to the ¹¹⁸⁴SKTASASPWKSA¹¹⁹⁵ sequence present in an amphiphilic α -helical region of the cytoskeletal protein α II spectrin that has been shown by X-ray crystallography (PDB entry: 2FOT) to fit snugly within the hydrophobic binding groove of Ca²⁺-bound calmodulin (50). We could then hypothesize that amino acids 314–344 in each *LiEndoG* monomer adopt a $\beta\beta\alpha$ structure that allows extension of the common flap in the eukaryotic family of EndoG nucleases while maintaining a similar fold in this region (Figure 3). We found indirect support to this assumption in the crystal structure of TM1367 protein from *Thermotoga maritima* (51), which includes a sequence identical to ³¹⁶MRVE³¹⁹ of *LiEndoG* at an N-terminal $\beta_1\beta_2\alpha_1$ motif. The fact that this stretch is flanked by two prolines is suggestive of the need for a forced turn in the secondary structure so as to fit in this insertion while keeping the overall shape of the rest of the protein. Although these findings are surely not enough to guarantee the goodness of our model, we are confident about the feasibility of these structural features, which pave the way for further research in the absence of an X-ray crystal structure.

Insert 4. Multiple sequence alignments and homology modeling of residues 440–510 of *LiEndoG* consistently found a similar CTD in human EXOG (18) as the only structural template available in the PDB, on the basis of its predicted secondary structure, despite its relatively low sequence identity (11). Unfortunately, due to poor electron density, a number of side chains were missing from the model deposited in this database and reasonable doubts arose about the plausibility of the interhelical loop regions, particularly involving helix 2. Likewise, we could not be sure about the exact composition and relative orientation of the second and third helices. Refinement of PDB en-

try 4A1N using Coot (19) and Phenix (20,21) substantially improved the figures of merit ($R_{\text{work}} = 0.191$ and $R_{\text{free}} = 0.250$) and provided a Ramachandran map free of outliers (Supplementary Figure S1) and a MolProbity (52) overall score of 2.49. The optimized sequence alignment obtained when this refined structure—with the right register of amino acids—was used as input for the one-to-one threading method implemented in Phyre (16) gave us extra confidence that the CTD of *LiEndoG* consists of a three-helix bundle that places helix 1 spatially close to the DNA-binding groove.

Substrate recognition and binding. Re-refinement of PDB entry 5GKP allowed us to trace the position of nucleotides 3–8 of 5'-d(TTTTTTGT)-3' in chain C and 1–7 in chain D bound to monomers A and B of the H148A/F122A variant of *CeEndoG*, respectively (Supplementary Figure S2). The Mg²⁺ located within the $\beta\beta\alpha$ -metal motif is coordinated by the phosphate oxygens O3' and OP1 of T5 and T6, respectively, in subunit A, and O3' and OP1 of T2 and T3 in subunit B. Thus, we considered the phosphodiester bond between these O3' atoms and the phosphorus of the following nucleotide as site 0 for cleavage (Table 1). Protein residues involved in fixating the phosphates upstream and downstream of this site were also identified. A best-fit root-mean-square C α superposition of EXOG and *LiEndoG* onto *CeEndoG* then allowed us to locate the equivalent residues in these two enzymes and find a rationale for the differences in substrate selectivity, as will be discussed below.

Mutations leading to replacement of Ser211 with Ala or Asp in the active site strongly affect *LiEndoG* DNA cleavage efficiency

The SRGH active site motif in *LiEndoG* is rather unique among nonspecific $\beta\beta\alpha$ -metal finger nucleases (53), which usually contain the ‘canonical’ DRGH signature. To assess the importance of Ser211 in the DNA cleavage ability of *LiEndoG*, we expressed the Ser211Ala and Ser211Asp variants and assayed their nuclease activities. Whereas the Ser211Ala variant showed only a moderate reduction in activity, replacement of this Ser with the ‘canonical’ Asp resulted in a dramatic loss of catalytic efficiency (Figure 4). This observation is in consonance with results reported for the activity on ssDNA of human EXOG upon replacement of the equivalent Ser137 with Asp (18).

Inspection of our template structures revealed that the active site Asp in the $\beta\beta\alpha$ -domains of *DmEndoG* and *CeEndoG* establishes a strong electrostatic interaction with an arginine located at the N-terminus of an α -helix. Because in *LiEndoG* an Ala occupies this position, we thought that a second Ala \rightarrow Arg mutation could restore the lost activity in the S211D variant. Indeed, when we expressed and tested the S211D/A247R double mutant, we found that 75% of the activity of the wild-type enzyme had been recovered (Figure 4). This finding supports the view that in *LiEndoG*^{S211D/A247R} (as happens with ‘true’ DRGH-containing $\beta\beta\alpha$ -metal finger nucleases) the guanidinium of Arg247 is used to anchor one of the side-chain carboxylate oxygens of Asp211 in place so that the other oxygen

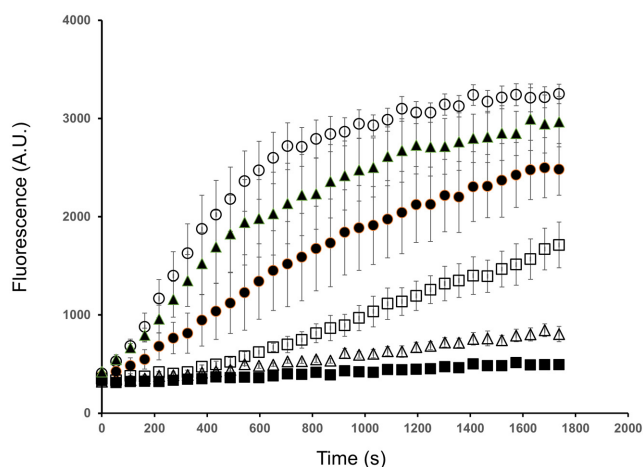


Figure 4. Time course of the digestion of the 5'-FAM/3'-TAMRA-labeled 16-bp dsDNA probe by wild-type *LiEndoG* (○) and by the protein variants generated by site-directed mutagenesis of the coding sequence: *LiEndoG*^{S211A} (●), *LiEndoG*^{S211D} (□), *LiEndoG*^{H214A} (■), *LiEndoG*^{R212A} (△), and *LiEndoG*^{S211D/A247R} (▲) over 30 min of incubation. Error bars represent the standard error obtained from four independent experiments ($n = 4$).

hydrogen bonds to the carboxamide nitrogen of the active site Asn246. In the wild-type *LiEndoG* this carboxamide nitrogen is hydrogen bonded to the hydroxyl oxygen of Ser211, which points away from the hydrophobic side chain of Met243.

For comparative purposes and completeness, we also replaced His214 with Ala. As expected, this H214A variant was unable to degrade DNA, given the crucial role of this residue as the general base that activates a water molecule for the nucleophilic in-line attack on the phosphorous atom during the nuclease reaction (30). Similarly, replacement of Arg212 with Ala generates a variant that is virtually inactive (Figure 4).

***LiEndoG* is endowed with a pH-sensitive self-inhibitory domain**

Insert 1 is located in close proximity to the active site and approximately in the same spatial region that is occupied by *DmEndoGI* in the *DmEndoG:DmEndoGI* complex (6) (Figure 3). The presence of numerous aspartic and glutamic acid residues in the sequence of this *EndoG* inhibitor, as well as previous results from our laboratory demonstrating that *LiEndoG* is inhibited at pH 8.0, led us to think that this stretch could be acting as a self-inhibitory domain responsible for the pH-dependence of this enzyme. To test this hypothesis, the Asp/Glu-rich stretch comprising residues 145–180, which is exclusive to trypanosomatids (Supplementary Figure S4), was replaced by the sequence AKNDA that occupies the equivalent position in *DmEndoG* (Figure 1). The activity of this *LiEndoG* Δ i variant was compared to that of the wild-type enzyme over a range of pH values. The results obtained (Figure 5A) demonstrate that deletion of this region abolishes inhibition at pH values of 7.5–8. Indeed, whereas the wild-type enzyme is extremely sensitive to 0.5–1.5 pH increments, *LiEndoG* Δ i preserves its catalytic proficiency over the whole range of pH values studied.

Once the pH-dependent modulation of this self-inhibitory domain was demonstrated, we focused on analysing the relevance of some of the negatively charged residues present in this sequence stretch. We replaced the three contiguous aspartic residues located in the central region of the domain (¹⁶⁰DDD¹⁶²) with three alanines to generate *LiEndoG*^{D(160–162)A} and the nuclease activity of this variant was studied at different pH values. The results shown in Figure 5B demonstrate that these three aspartates are essential for *LiEndoG*'s self-inhibition as the *LiEndoG*^{D(160–162)A} variant essentially recapitulates the behavior of *LiEndoG* Δ i over the range of pHs tested.

Deletion of the C-terminal domain affects the DNA cleavage pattern

Alignment of *EndoG* sequences from trypanosomatids with those from endo/exonucleases of other organisms such as *D. melanogaster*, *C. elegans*, *S. cerevisiae* and *H. sapiens* (Supplementary Figure S5) reveals the presence of an extended CTD with low sequence identity but predicted high structural similarity to the CTD of human EXOG. To define the role of this CTD, we generated a truncated *LiEndoG* protein lacking the last 70 residues (*LiEndoG* Δ ^{CTD}), which was expressed, refolded, and purified following the same procedures that were used to obtain wild-type *LiEndoG*.

The equilibrium sedimentation profiles of *LiEndoG* Δ ^{CTD} could be fitted to a single ideal species model (Figure 2A) with an apparent molecular mass ($M_{w,app}$) of 91 ± 2 kDa (mean of six independent experiments), which is 1.98-fold the value measured by mass spectrometry. Sedimentation velocity experiments carried out in parallel showed the presence of three different species (Figure 2B), the major one being compatible with a dimer ($s_{20,w} = 5.0 \pm 0.07$ S; $M_w \sim 81$ kDa) and the other two with a monomer ($s_{20,w} = 3.2 \pm 0.1$ S; $M_w \sim 44$ kDa) and a tetramer ($s_{20,w} = 7.6 \pm 0.2$ S; $M_w \sim 159$ kDa; mean of four independent experiments). Their percentages were consistent with the $M_{w,app}$ calculated from equilibrium experiments and an upper limit of 14 nM was estimated for the dimer dissociation constant from the percentages of species (around 81%, 4% and 15%, respectively) measured at the lowest *LiEndoG* Δ ^{CTD} concentration used.

When the activities of wild-type *LiEndoG* and *LiEndoG* Δ ^{CTD} on plasmid DNA were compared (Figure 6A), similar nicking of the supercoiled substrate was observed although the concentration of *LiEndoG* Δ ^{CTD} had to be increased fourfold to reach an equivalent efficiency. Accordingly, in all subsequent comparative experiments, protein concentrations were adjusted so as to maintain similar enzymatic activities.

The nuclease activity of both enzymes was then assayed on a 5'-³²P-labeled 42-bp-long dsDNA helix and the generated fragments were subjected to denaturing PAGE. Digestion of the same probe with DNase I was included as a control. Both *LiEndoG* and *LiEndoG* Δ ^{CTD} produced DNA fragments of many different sizes (Figure 6B). As expected from digestion of a 5'-labeled probe by an endonuclease with no sequence specificity, fragment size decreased on increasing the amount of enzyme. Low enzyme concentrations gave rise to an almost random distribution of DNA

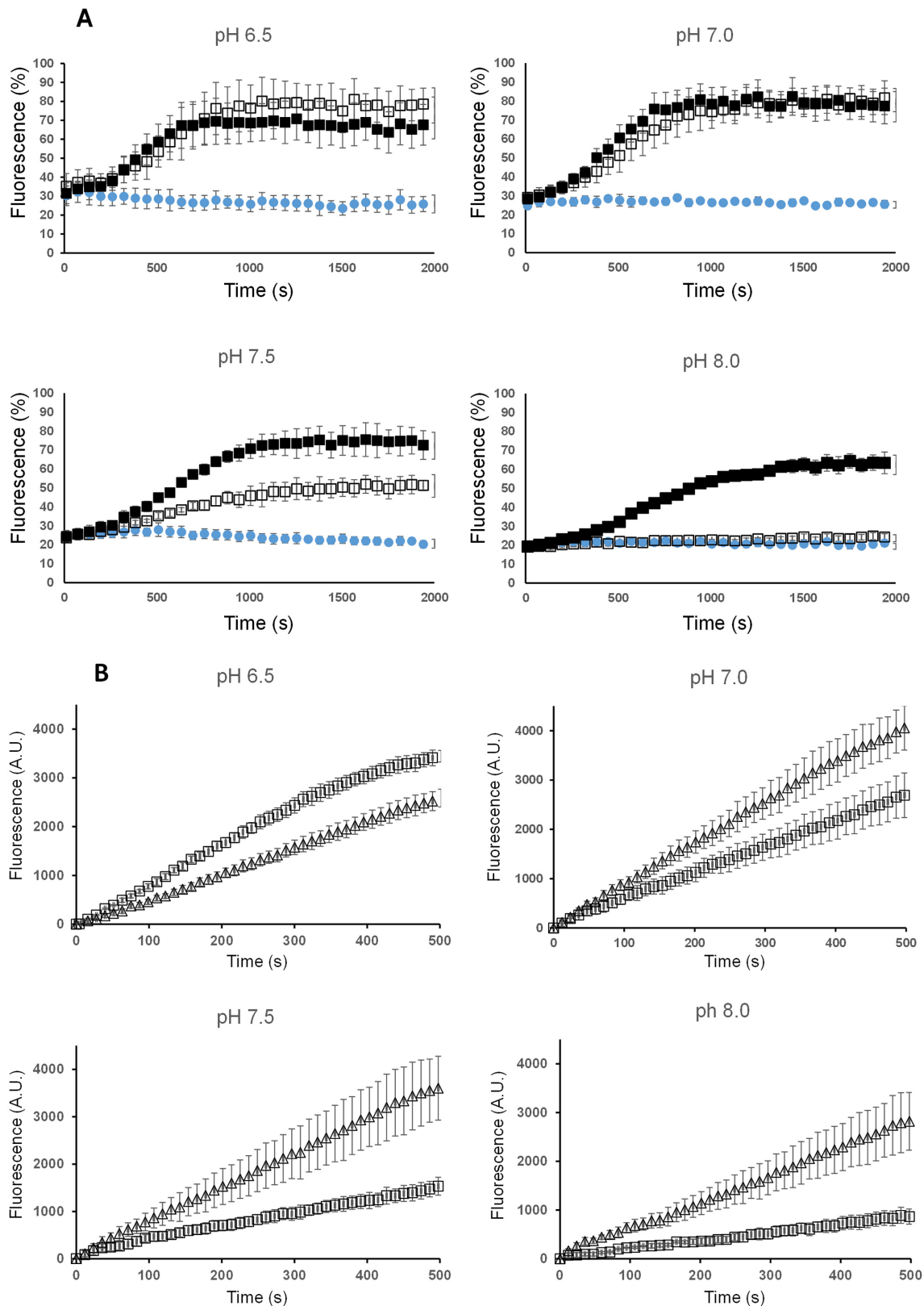


Figure 5. (A) Comparison of the endonucleolytic activities on a dually 5'-FAM/3'-TAMRA-labeled dsDNA probe of wild-type *LiEndoG* (□) and *LiEndoG*Δi (■) at different pH values over >30 min of incubation. (●) Fluorescence of the probe in the absence of enzyme. All data are normalized against the fluorescence values of the 5'-FAM-labeled ssDNA probe. (B) Activity of wild-type *LiEndoG* (□), and *LiEndoG*^{D(160-162)A} (Δ) on the same probe at different pH values over the first 10 min of incubation. Error bars represent the standard error obtained from three independent experiments ($n = 3$).

fragments that progressively disappeared as the concentration of enzyme increased. The smallest DNA fragments detected were 5-, 6-, 8- and 10-mers for *LiEndoG* and 6- and 10-mers for *LiEndoG* Δ ^{CTD}. A remarkable difference was that *LiEndoG* quickly generated small-sized fragments. Note that approximately half of the radioactive signal is associated to an octanucleotide, *lane 3*, even when there is still a large proportion of undigested probe. In contrast, upon incubation with *LiEndoG* Δ ^{CTD} high concentrations of short oligonucleotides are visible only after all the original DNA probe has been completely digested.

To test the activity of both enzymes on ssDNA and dsDNA, we selected a probe with a 48-nt-long central dsDNA stem flanked on each side by a differently sized ssDNA overhang (Figure 6C). Strikingly different behaviors were found: the wild type showed an extraordinary affinity for the ssDNA region, as revealed by the complete absence of DNA fragments longer than 12 nucleotides (Figure 6C), whereas this particular preference for ssDNA was not observed for *LiEndoG* Δ ^{CTD}, which gave rise to a degradation pattern quite similar to that observed upon cleavage of a dsDNA substrate possessing blunt-end termini (Figure 6B). In addition, only *LiEndoG* showed a definite preference for cleaving the probe at positions 7 and 11, as fragments of these two sizes were the first to be observed upon increasing concentrations of the enzyme (*lane 5*). Since position 11 marks precisely the first nucleotide of the dsDNA stem, it seems feasible that *LiEndoG* can eliminate 5' overhangs appended to a dsDNA, an activity typically displayed by flap endonucleases. Once generated, the 11-nt-long fragment is quickly cleaved and shorter ones appear but 5–6-mers are not degraded further.

Finally, to reinforce the idea that the CTD is responsible for the preferential activity of *LiEndoG* on ssDNA, we selected as substrate a short ssDNA fragment unable to form any dsDNA structure. The results shown in Figure 6D reveal that *LiEndoG* Δ ^{CTD} displays such a low affinity for ssDNA that only a negligible degradation pattern is observed at the highest enzyme concentration assayed (*lane 1*, *LiEndoG* Δ ^{CTD}). In stark contrast, this probe is readily cleaved by wild-type *LiEndoG* (starting in *lane 4*).

Analysis of the final products of DNA digestion by *LiEndoG* consistently shows that the smallest ssDNA substrate that is cleaved needs to be longer than a pentamer or a hexamer. We take this as an indication, in agreement with our molecular models (Figure 7, Supplementary Figure S6, *LiEndoG*.ssDNA.pdb and *LiEndoG*.ssDNA+dsDNA.pdb in Supplementary material), that shorter oligonucleotides do not bind with sufficient affinity because they lose some relevant interactions with the enzyme (Table 1).

DISCUSSION

We previously reported (9) that the core of the catalytic domain of *LiEndoG* is likely to display an overall similarity to other structurally well-characterized extracellular nucleases such as those from *Serratia marcescens* (54) and *Anabaena* sp. (55). These two proteins belong to the $\beta\beta\alpha$ -metal superfamily of nucleases which share a structurally conserved active site scaffold and utilize a divalent metal

ion to catalyze the hydrolysis of the phosphodiester bond (56). We proposed that, in *LiEndoG*, the hydrated Mg^{2+} ion is coordinated directly through the side-chain carboxamide oxygen of Asn246 and also that one of the Mg^{2+} -bound water molecules is hydrogen-bonded to the imidazole N8 of His214, which behaves as a general base to activate a water molecule for the nucleophilic in-line attack on the phosphorous atom of DNA or RNA substrates (9). This catalytic His214 is part of the 'SRGH' motif characteristic of a second family of sugar non-specific nucleases in higher eukaryotes, represented by EXOG, which also have an additional CTD consisting of ~ 70 amino acids (18). Human EXOG and EndoG were shown to have complementary and partially overlapping enzymatic activities. When Ser137 in the catalytic site of EXOG was replaced by the aspartic acid residue present in the canonical 'DRGH' motif the enzymatic activity of the resulting S137D variant was lower than that of the wild type (18). In our hands, the DNA cleavage ability of the S211D variant of *LiEndoG* was strongly reduced. This was intriguing, especially in view of the additional finding that this nuclease activity was affected only slightly in the S211A variant (Figure 4). Therefore, we thought that the deleterious effect of Asp211 was mostly due to the negative charge in the side chain of this amino acid. Re-examination of the macromolecular structures revealed that in *DmEndoG* and *CeEndoG* the carboxylate of the equivalent Asp is involved in a salt bridge with the guanidinium of Arg181, the spatial counterpart of which is an alanine (Ala247) in *LiEndoG*. When an arginine was introduced at this position, the nuclease activity of the S211D/A247R variant arising from the double mutation was $\sim 75\%$ of that of the wild-type enzyme. Therefore, we propose that the carboxylate of Asp211 disorganizes the active site and faces the nucleotide phosphates unless it is firmly anchored in place by the guanidinium of Arg247.

As we demonstrated earlier, *LiEndoG* plays a dual role in both cell death and survival (9,10). Due to the importance of controlling these sugar nonspecific nucleases, different molecular mechanisms have been found in nature to finely modulate their activity. Among them, cellular compartmentalization, the existence of specific inhibiting proteins, and regulation by pH are good examples (5,6,44). For *LiEndoG* we have demonstrated abolition of nuclease activity at high pH values ($pH \geq 8.0$). In this respect, different approaches indicate pH values close to 8 in the mitochondrial matrix and a reduction of about 1 pH unit in the cytosol (57,58). Accordingly, *LiEndoG* would be almost completely inhibited inside the mitochondrion but could become activated upon its release to the cytosol during cell death. This activation could be further potentiated by the decrease in cytosolic pH values observed during apoptosis-like processes (2,59). However, since both cytosolic and mitochondrial pHs can be strongly reduced during activation of cells with Ca^{2+} -mobilizing agonists (58), the activity of *LiEndoG* inside the mitochondrion might be stimulated in different cellular conditions. In fact, any decrease in mitochondrial activity would decrease proton transport rates and cause a reduction in the pH within the matrix that, eventually, would indirectly activate *LiEndoG*.

Our updated molecular model of *LiEndoG* indicates that, in all probability, the domain corresponding to

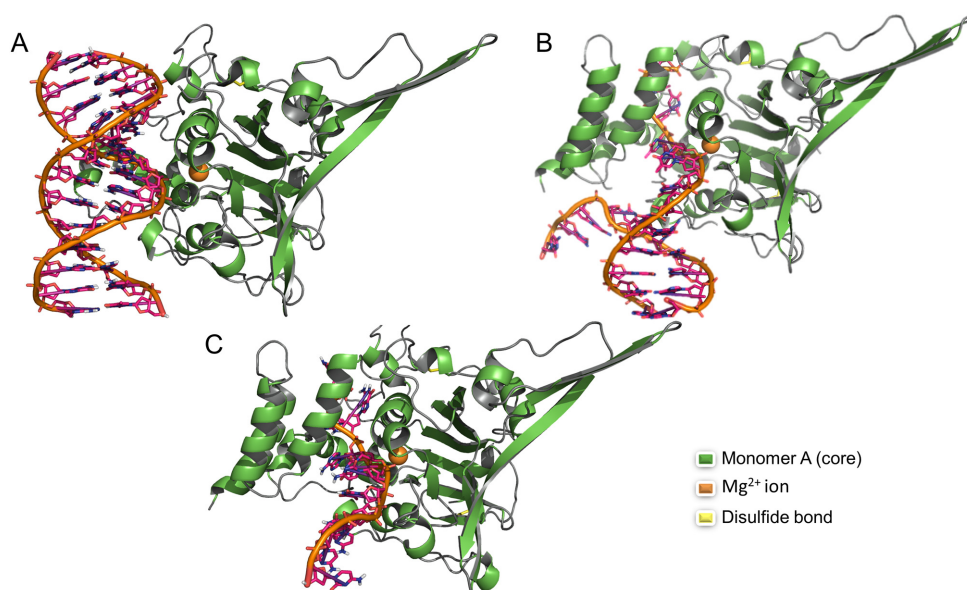


Figure 7. Molecular models of (A) *LiEndoG* Δ CTD in complex with a dsDNA molecule, (B) wild-type *LiEndoG* in complex with a ssDNA–dsDNA junction and (C) wild-type *LiEndoG* in complex with a ssDNA oligonucleotide. Only one protein monomer is shown for clarity and a rainbow coloring scheme has been chosen to facilitate tracing of the backbone from the N-terminus (blue) to the C-terminus (red). The corresponding DNA sequences are displayed in Table 2. The dotted arrows in (B) and (C) point to residues F447, Y450 and R451 in helix 1 of the CTD (shown as sticks).

residues 145–180 is located in the same spatial region that is occupied by *DmEndoG*I in its complex with *DmEndoG* (6). To demonstrate the self-inhibitory function of this unique amino acid stretch, we first expressed a variant enzyme in which this domain was deleted (*LiEndoG* Δ i) and then measured the nuclease activity of this engineered protein in the presence of a 5'-FAM/3'-TAMRA-labeled dsDNA oligonucleotide at different pH values ranging from 6.5 to 8.0. As shown in Figure 5A, the nuclease activity of *LiEndoG* Δ i is not reduced at basic pH values (7.0–8.0). It must be pointed out that the activity of both the wild-type enzyme and this variant is maximum at pH 6.5. Accordingly, deletion of this domain does not shift the optimum pH value but broadens the range of pH values at which this enzyme is active. Furthermore, the putative role of the three central Asp residues (¹⁶⁰DDD¹⁶²) in the mechanism of self-inhibition was assessed by site-directed mutagenesis. The finding that replacement of these three aspartates with three alanines renders a protein (*LiEndoG*^{D(160-162)A}) that is still active at pH 8.0 reinforces the relevance of these acidic residues and points to an effective pH-dependent self-regulatory mechanism in *LiEndoG*.

Our experimental results indicate that the C-terminal extension present in *LiEndoG*, which –to the best of our knowledge—has a known structural counterpart only in human mitochondrial EXOG, is responsible for its high affinity toward ssDNA. Remarkably, deletion of this domain in EXOG or replacement of Gly277, closely upstream of the CTD, with valine results in variants that are much less active or as inactive on ssDNA as the H140A enzyme, respectively (18). This indicates the relevance of both the presence and the positioning of this domain in EXOG for its enzymatic activity. Furthermore, when working on ssDNA or dsDNA substrates, EXOG produces di- and mononucleotides or mostly dinucleotides, respectively, but

this 5'–3' exonuclease activity is abolished in the G277V and Δ CTD variants (18). Since no mono- or dinucleotides are detected upon digestion with *LiEndoG*, despite the presence of a CTD and significant sequence similarity with EXOG, *LiEndoG* cannot be considered a *bona fide* exonuclease. In any case, it seems clear that the substrate preferences are in both cases dictated by the presence of this C-terminal extension.

The modeling work presented here leads us to think that the CTD of *LiEndoG* folds into a three-helix bundle similar to that found in human EXOG and different from that reported for DnaT, a primosomal protein that is required for the stalled replication fork restart in *E. coli* (60). The unpublished crystallographic work on EXOG (11), improved upon by our own refinement (Supplementary Figure S1 and Table S1), sheds light on the preferential cleavage by these enzymes because it shows in atomic detail the spatial relation between the CTD and the nucleotide-binding groove in the vicinity of the active site. When the EXOG structure is best-fit superimposed onto that of the *CeEndoG*(H148A/Q130K) in its complex with ssDNA, insight is gained about the contacts established between these nucleases and this type of substrate. This knowledge can then be safely extrapolated to *LiEndoG* to complement what is already known about its interaction with dsDNA (9). Thus, we learnt that the three helices in the CTD bundle are arranged in such a way that, when positioned close to the active site, they prevent binding of a dsDNA to the enzyme's active site (Figure 7). Nonetheless, passage of a ssDNA is still possible, hence the affinity of *LiEndoG* for ssDNA–dsDNA junctions. At the junction, the displaced unpaired DNA strand is bent at a single phosphodiester bond by the loop connecting helices 1 and 2 of the CTD in a manner reminiscent of that described for helices α 2 and α 3 of human Exo1 (61). The hydroxyl group of the strictly con-

served Ser461, at the tip of this loop, and the guanidinium of Arg492 in helix 3, can theoretically hydrogen bond to a phosphate in the displaced DNA strand, so as to stabilize the unpaired state (Figure 7B), with the aid of the side chains of Asn184 and Asn187.

In addition, helix 1 of the CTD projects the side chains of the strictly conserved Phe447, Tyr450 and Arg451 toward the ssDNA substrate. We propose that the side chain of Tyr450, with the structural support of Phe447, plays the roles of (i) preventing the stacking of the base in the second nucleotide 5'-upstream of that undergoing the cleavage—in a manner reminiscent of the *hydrophobic wedge* observed in FEN-1 family members (61), (ii) establishing a hydrogen bond with a phosphate oxygen in the nucleotidic substrate and (iii) introducing a turn in the ssDNA to be cleaved such that upstream nucleotides are directed toward the polyglutamine domain in *insert 2*, thus accounting for the much higher affinity for this type of substrate of the wild-type enzyme relative to that of its Δ^{CTD} counterpart (Figure 6). Although our model strongly suggests that *insert 2* could behave as a ssDNA-binding domain, we cannot assign to it a definite structure with confidence because it is predicted as an intrinsically disordered region and lacks homology with the limited set of proteins with known structure and similar function (62).

Since this CTD is present neither in *DmEndoG* nor in *CeEndoG*, both lacking exonuclease activity, and also because the 5'-3' exonuclease activity is thought to be essential in human EXOG for DNA recombination and repair (18), we suspected that *LiEndoG* might be competent as an exonucleolytic enzyme or have the capacity to trim bifurcated or branched DNA molecules containing free 5'-single-stranded ends (i.e. 5'-flaps) such as those emerging from replication, recombination and repair processes (63). The CTD of *LiEndoG* directs substrates containing ssDNA toward the active site of the enzyme. Upon its removal, *LiEndoG* Δ^{CTD} essentially behaves as do human *EndoG* and *CeEndoG* (no preference for ssDNA) (8), which naturally lack this C-terminal extension and are regarded as endonucleases (Figure 1). This behavior can be rationalized because it is the C-terminal three-helix bundle, presumably when held in place with the aid of the disulfide linkage between Cys434 and Cys439, that confers affinity for ssDNA, be it alone or as part of a dsDNA-ssDNA junction. In the absence of the CTD in front of the active site, either because it is displaced by a hinge motion from the main body of the enzyme or because it has been deleted, *LiEndoG* can bind dsDNA in the active site and behaves as a typical endonuclease. In connection with this reasoning, the reported high amounts of catalytically incompetent molecules and the opposed differences in activity on a hairpin substrate of recombinant *CeEndoG* (~3-fold) and EXOG (~8-fold) enzymes expressed in *E. coli* when they are purified either from the soluble fraction or from inclusion bodies followed by refolding (8) could be related to the correct formation of these disulfide bonds.

Interestingly, the preferential ability of *LiEndoG* to produce 5–7-mers (Figure 6B–D) rather than the mono- and dinucleotides characteristically produced by EXOG, despite the presence of a structurally similar CTD in both enzymes, can be mostly ascribed to a correlated mutation (64) result-

ing in the reciprocal positional exchange of positively and negatively charged amino acids in the Arg457:Glu488 and Glu317:Arg346 pairs that stabilize the interaction between helices 1 and 3 through a salt bridge in the respective enzymes (Supplementary Figure S5A). Arg457 is involved in substrate phosphate recognition three nucleotides upstream of the phosphodiester bond that is cleaved in the active site of *LiEndoG* (Table 1) but this favourable electrostatic interaction cannot take place in EXOG through the carboxylate of the positionally equivalent Glu317. Instead, a reciprocal Ala \leftrightarrow Arg exchange at positions 454 and 314 makes it possible for Arg314 in EXOG (PDB numbering), but not for Ala454 in *LiEndoG*, to presumably recognize a terminal 5'-phosphate. In contrast, the positions occupied by Arg442, Arg449 and Arg259 in *LiEndoG*, which can recognize a terminal 5'-phosphate 4–5 nucleotides upstream of the cleavage site (Table 1), are occupied in EXOG by Leu302, Leu309 and Glu184, which are incapable of such recognition. This structural rationale accounts well for the observed differences in nuclease activities exerted by these enzymes (8).

Because the nuclease activity in the S211D variant of *LiEndoG* was restituted up to 75% of that of the wild-type enzyme upon the concomitant introduction of a A247R replacement, and because EXOG also has (i) an ¹³⁷SRGH¹⁴⁰ signature, (ii) a hydrophobic residue (Phe168) next to Ser137, and (iii) a similar CTD, we decided to check the amino acid composition of EXOG at the position equivalent to that of Ala247, which turns out to be a serine (Ser172). Interestingly, the distance (~10.0–14.0 Å) between the C β atoms of Ser172 and Arg314 in the X-ray crystal structure of EXOG (or Ala247 and Arg454 in our *LiEndoG* model) is less than or almost equal to twice the length of an Arg side chain. This fact virtually rules out, on electrostatic and steric grounds, the possibility of having two Arg residues simultaneously in these positions and strongly suggests that these sequence differences, which spatially cluster around the 'wedge' Tyr350, go hand in hand (Supplementary Figure S5). Thus, our observations provide a rationale for the reported substrate differences among members of this superfamily of nucleases, which appear to be dictated by the intimate relation in these enzymes between the nuclease domain and the CTD facing the active site cleft.

Remarkably, *ScNuc1p* (42,43) can produce not only endonucleolytic products but also dinucleotides due to preferential cleavage close to the 5'-end (8). We note that the CTD of this enzyme, although similarly linked through a (presumably) redox-dependent disulfide bridge to the globular protein, is much shorter than the CTD of both EXOG and *LiEndoG* and unlikely to adopt an α -helical structure. The three-dimensional structure of *ScNuc1p* has not been reported yet but we believe that its CTD can fold similarly to the synaptic CTD (residues 220–276) of Maedi-Visna virus integrase, which has been recently solved both in isolation as a dimer by X-ray crystallography at 1.78 Å resolution (UniProtKB entry P35956, PDB code 5LLJ) and, at much lower resolution by cryo-electron microscopy (PDB code 5M0R), as part of the supramolecular assembly that mediates lentiviral DNA integration (65). In this complex, the β_1 - β_2 loops of two CTDs get inserted into expanded DNA major grooves and contribute to target DNA bending. The

Arg-rich cluster present in the CTD of ScNuc1p could conceivably play the role, if properly positioned, of recognizing the terminal 5'-phosphate located two nucleotides upstream of the cleavage site.

The finding that EndoG from trypanosomatids removes 5' overhangs from dsDNA in a highly specific and efficient manner suggests some similarities with CRN-1, a *C. elegans* ortholog of flap endonuclease FEN-1 that localizes to nuclei and cooperates with CeEndoG to promote step-wise DNA fragmentation utilizing the endonuclease activity of CeEndoG and both the 5'-3' exonuclease and gap-dependent endonuclease activities of CRN-1 (66). Cooperation between both enzymes has been proposed to play a critical role in switching the state of cells from DNA replication/repair to DNA degradation during apoptosis. The existence and properties of a DNA flap endonuclease from *Trypanosoma cruzi* (TcFEN1) have just been reported (67) and orthologs have been detected in *L. major* (67) and *L. donovani* (68). TcFEN1 localizes to the nucleus in transfected epimastigotes and its overexpression increases both parasite proliferation and survival against exogenous oxygen peroxide. LdEndoG and LdFEN1 have been shown to co-immunoprecipitate in nuclear extracts following treatment of the parasites with a topoisomerase IB inhibitor (68). Whether or not LiEndoG and LiFEN1 also cooperate in processing flap fragments during DNA replication and repair is not known yet but LiEndoG on its own might well have an activity equivalent to that of CeEndoG and CRN-1 acting together.

In conclusion, the much larger size of LiEndoG compared to that of other EndoGs (9) is due to sequence insertions that are not found outside the trypanosomatid family and provide this enzyme with rather unique properties. In the present work, we have identified and characterized the role of two of these insertions, namely a pH-regulated self-inhibitory domain and an extended CTD that confers substrate selectivity for ssDNA, both as a free molecule and as part of a ssDNA-dsDNA junction. Consequently, LiEndoG emerges as an endonuclease with the added structure-specific ability to cut preferentially at the base of a ssDNA overhang, in a manner reminiscent of that carried out by two-metal-ion flap endonucleases (69). In addition, the elongated flaps (or 'wings') that make up a saddle-shaped part of the dimerization domain are suggestive of the participation of this enzyme in multiprotein assemblies. All in all, our study paves the way for the characterization of other domains and prompts the investigation of LiEndoG's binding partners.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

C.O. and E.R. gratefully acknowledge being recipients of FPI predoctoral fellowships from the Spanish Government. We are grateful to Kilian Gutiérrez (UAH), for his assistance in protein purification, and Javier Rodríguez-Villanueva, for help in some early stages of the molecular modeling work.

FUNDING

Spanish MEC/MICINN [SAF2012-39760-C02-02 and SAF2015-64629-C2-2-R to F.G. and A.J.R.]; Comunidad Autónoma de Madrid [BIPEDD-2-CM Project, ref. S-2010/BMD-2457 to M.M., F.G. and A.J.R.]; Junta de Comunidades de Castilla la Mancha [POII10-0180-7897 to A.J.R.]. Funding for open access charge: MICINN Project [SAF2015-64629-C2-2-R].

Conflict of interest statement. None declared.

REFERENCES

- Low, R.L. (2003) Mitochondrial Endonuclease G function in apoptosis and mtDNA metabolism: a historical perspective. *Mitochondrion*, **2**, 225–236.
- Ramsdale, M. (2012) Programmed cell death in the cellular differentiation of microbial eukaryotes. *Curr. Opin. Microbiol.*, **15**, 646–652.
- Ball, T.K., Suh, Y. and Benedik, M.J. (1992) Disulfide bonds are required for *Serratia marcescens* nuclease activity. *Nucleic Acids Res.*, **20**, 4971–4974.
- Wu, S.-I., Lo, S.-K., Shao, C.-P., Tsai, H.-W. and Hor, L.-I. (2001) Cloning and characterization of a periplasmic nuclease of *Vibrio vulnificus* and its role in preventing uptake of foreign DNA. *Appl. Environ. Microbiol.*, **67**, 82–88.
- Ghosh, M., Meiss, G., Pingoud, A.M., London, R.E. and Pedersen, L.C. (2007) The nuclease A-inhibitor complex is characterized by a novel metal ion bridge. *J. Biol. Chem.*, **282**, 5682–5690.
- Loll, B., Gebhardt, M., Wahle, E. and Meinhart, A. (2009) Crystal structure of the EndoG/EndoGI complex: mechanism of EndoG inhibition. *Nucleic Acids Res.*, **37**, 7312–7320.
- Rose, P.W., Prlc, A., Bi, C., Bluhm, W.F., Christie, C.H., Dutta, S., Green, R.K., Goodsell, D.S., Westbrook, J.D., Woo, J. *et al.* (2015) The RCSB Protein Data Bank: views of structural biology for basic and applied research and education. *Nucleic Acids Res.*, **43**, D345–D356.
- Kieper, J., Lauber, C., Gimadutdinov, O., Urbanska, A., Cymerman, I., Ghosh, M., Szczesny, B. and Meiss, G. (2010) Production and characterization of recombinant protein preparations of endonuclease G-homologs from yeast, *C. elegans* and humans. *Protein Expression Purif.*, **73**, 99–106.
- Rico, E., Alzate, J.F., Arias, A.A., Moreno, D., Clos, J., Gago, F., Moreno, I., Domínguez, M. and Jiménez-Ruiz, A. (2009) *Leishmania infantum* expresses a mitochondrial nuclease homologous to EndoG that migrates to the nucleus in response to an apoptotic stimulus. *Mol. Biochem. Parasitol.*, **163**, 28–38.
- Rico, E., Oliva, C., Gutierrez, K.J., Alzate, J.F., Genes, C.M., Moreno, D., Casanova, E., Gigante, A., Pérez-Pérez, M.-J., Camarasa, M.-J. *et al.* (2014) *Leishmania infantum* EndoG is an endo/exo-nuclease essential for parasite survival. *PLoS ONE*, **9**, e89526.
- Welin, M., Moche, M., Arrowsmith, C.H., Berglund, H., Bountra, C., Collins, R., Edwards, A.M., Flodin, S., Graslund, S., Hammarstrom, M. *et al.* (2012) Human mitochondrial endo-exonuclease. *Protein Data Bank*, doi:10.2210/pdb4a1n/pdb.
- Lin, J.L., Wu, C.C., Yang, W.Z. and Yuan, H.S. (2016) Crystal structure of endonuclease G in complex with DNA reveals how it nonspecifically degrades DNA as a homodimer. *Nucleic Acids Res.*, **44**, 10480–10490.
- Lin, J.L., Nakagawa, A., Skeen-Gaar, R., Yang, W.Z., Zhao, P., Zhang, Z., Ge, X., Mitani, S., Xue, D. and Yuan, H.S. (2016) Oxidative stress impairs cell death by repressing the nuclease activity of mitochondrial endonuclease G. *Cell Rep*, **16**, 279–287.
- Lin, J.L., Nakagawa, A., Lin, C.L., Hsiao, Y.-Y., Yang, W.-Z., Wang, Y.-T., Doudeva, L.G., Skeen-Gaar, R.R., Xue, D. and Yuan, H.S. (2012) Structural insights into apoptotic DNA degradation by CED-3 Protease Suppressor-6 (CPS-6) from *Caenorhabditis elegans*. *J. Biol. Chem.*, **287**, 7110–7120.
- Kelley, L.A. and Sternberg, M.J.E. (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nat. Protoc.*, **4**, 363–371.

16. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.
17. Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L. *et al.* (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.*, **42**, W252–W258.
18. Cymerman, I.A., Chung, I., Beckmann, B.M., Bujnicki, J.M. and Meiss, G. (2008) EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Res.*, **36**, 1369–1379.
19. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
20. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Echols, N., Headd, J.J., Hung, L.W., Jain, S., Kapral, G.J., Grosse Kunstleve, R.W. *et al.* (2011) The Phenix software for automated determination of macromolecular structures. *Methods*, **55**, 94–106.
21. 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. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
22. DeLano, W.L. (2015) *The PyMOL Molecular Graphics System, version 1.8.2.0*. Schrödinger, LLC.
23. Stein, A. and Kortemme, T. (2013) Improvements to robotics-inspired conformational sampling in Rosetta. *PLoS One*, **8**, e63090.
24. Mandell, D.J., Coutsias, E.A. and Kortemme, T. (2009) Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat. Methods*, **6**, 551–552.
25. Kim, D.E., Chivian, D. and Baker, D. (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.*, **32**, W526–W531.
26. Miller, J., Rutenber, E. and Muchowski, P.J. (2009) Polyglutamine dances the conformational cha-cha-cha. *Structure*, **17**, 1151–1153.
27. Robertson, A.L., Horne, J., Ellisdon, A.M., Thomas, B., Scanlon, M.J. and Bottomley, S.P. (2008) The structural impact of a polyglutamine tract is location-dependent. *Biophys. J.*, **95**, 5922–5930.
28. López-Blanco, J.R., Canosa-Valls, A.J., Li, Y. and Chacón, P. (2016) RCD+: Fast loop modeling server. *Nucleic Acids Res.*, **44**, W395–W400.
29. Ceroni, A., Passerini, A., Vullo, A. and Frascioni, P. (2006) DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. *Nucleic Acids Res.*, **34**, W177–W181.
30. Bueren-Calabuig, J.A., Coderch, C., Rico, E., Jiménez-Ruiz, A. and Gago, F. (2011) Mechanistic insight into the catalytic activity of $\beta\beta\alpha$ -metallonucleases from computer simulations: *Vibrio vulnificus* periplasmic nuclease as a test case. *ChemBioChem*, **12**, 2615–2622.
31. Case, D.A.T.A.D., Cheatham, T.E. III, Simmerling, C.L., Wang, J., Duke, R.E., Luo, R., Walker, R.C., Zhang, W., Merz, K.M., Roberts, B. *et al.* (2012) University of California.
32. Li, C.L., Hor, L.I., Chang, Z.F., Tsai, L.C., Yang, W.Z. and Yuan, H.S. (2003) DNA binding and cleavage by the periplasmic nuclease Vvn: a novel structure with a known active site. *EMBO J.*, **22**, 4014–4025.
33. Wang, Y.T., Yang, W.J., Li, C.L., Doudeva, L.G. and Yuan, H.S. (2007) Structural basis for sequence-dependent DNA cleavage by nonspecific endonucleases. *Nucleic Acids Res.*, **35**, 584–594.
34. Carter, M., Voth, A.R., Scholfield, M.R., Rummel, B., Sowers, L.C. and Ho, P.S. (2013) Enthalpy-entropy compensation in biomolecular halogen bonds measured in DNA junctions. *Biochemistry*, **52**, 4891–4903.
35. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.*, **78**, 1606–1619.
36. Monterroso, B., Saiz, J.L., Garcia, P., Garcia, J.L. and Menendez, M. (2008) Insights into the structure-function relationships of pneumococcal cell wall lysozymes, LytC and Cpl-1. *J. Biol. Chem.*, **283**, 28618–28628.
37. Laue, T.M., Shah, B.D., Ridgeway, T.M. and Pelletier, S.L. (1992) In: Harding, S.E. and Horton, J.C. (eds). *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Royal Soc. Chem., Cambridge, pp. 90–125.
38. Wootton, J.C. and Federhen, S. (1996) Analysis of compositionally biased regions in sequence databases. *Methods Enzymol.*, **266**, 554–571.
39. Ward, J.J., McGuffin, L.J., Bryson, K., Buxton, B.F. and Jones, D.T. (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics*, **20**, 2138–2139.
40. de Marco, A. (2009) Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microb. Cell Fact.*, **8**, 26.
41. Stewart, E.J., Aslund, F. and Beckwith, J. (1998) Disulfide bond formation in the *Escherichia coli* cytoplasm: an *in vivo* role reversal for the thioredoxins. *EMBO J.*, **17**, 5543–5550.
42. Buttner, S., Eisenberg, T., Carmona-Gutierrez, D., Ruli, D., Knauer, H., Ruckenstein, C., Sigrist, C., Wissing, S., Kollros, M., Frohlich, K.U. *et al.* (2007) Endonuclease G regulates budding yeast life and death. *Mol. Cell*, **25**, 233–246.
43. Burhans, W.C. and Weinberger, M. (2007) Yeast endonuclease G: complex matters of death, and of life. *Mol. Cell*, **25**, 323–325.
44. Meiss, G., Gimadudinow, O., Haberland, B. and Pingoud, A. (2000) Mechanism of DNA cleavage by the DNA/RNA-nonspecific *Anabaena* sp. PCC 7120 endonuclease NucA and its inhibition by NuiA. *J. Mol. Biol.*, **297**, 521–534.
45. Polling, S., Ormsby, A.R., Wood, R.J., Lee, K., Shoubridge, C., Hughes, J.N., Thomas, P.Q., Griffin, M.D., Hill, A.F., Bowden, Q. *et al.* (2015) Polyalanine expansions drive a shift into α -helical clusters without amyloid-fibril formation. *Nat. Struct. Mol. Biol.*, **22**, 1008–1015.
46. McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P. and Lopez, R. (2013) Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.*, **41**, W597–W600.
47. UniProt, C. (2015) UniProt: a hub for protein information. *Nucleic Acids Res.*, **43**, D204–D212.
48. Bono, F., Ebert, J., Unterholzner, L., Güttler, T., Izaurralde, E. and Conti, E. (2004) Molecular insights into the interaction of PYM with the Mago-Y14 core of the exon junction complex. *EMBO Rep.*, **5**, 304–310.
49. Le Hir, H., Sauliere, J. and Wang, Z. (2016) The exon junction complex as a node of post-transcriptional networks. *Nat. Rev. Mol. Cell Biol.*, **17**, 41–54.
50. Simonovic, M., Zhang, Z., Cianci, C.D., Steitz, T.A. and Morrow, J.S. (2006) Structure of the calmodulin α II-spectrin complex provides insight into the regulation of cell plasticity. *J. Biol. Chem.*, **281**, 34333–34340.
51. Jin, K.K., Krishna, S.S., Schwarzenbacher, R., McMullan, D., Abdubek, P., Agarwalla, S., Ambing, E., Axelrod, H., Canaves, J.M., Chiu, H.J. *et al.* (2006) Crystal structure of TM1367 from *Thermotoga maritima* at 1.90 Å resolution reveals an atypical member of the cyclophilin (peptidylprolyl isomerase) fold. *Proteins*, **63**, 1112–1118.
52. Chen, V.B., Arendall, W.B. 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.
53. Hsia, K.-C., Li, C.-L. and Yuan, H.S. (2005) Structural and functional insight into sugar-nonspecific nucleases in host defense. *Curr. Opin. Struct. Biol.*, **15**, 126–134.
54. Lunin, V.Y., Levnikov, V.M., Shlyapnikov, S.V., Blagova, E.V., Lunin, V.V., Wilson, K.S. and Mikhailov, A.M. (1997) Three-dimensional structure of *Serratia marcescens* nuclease at 1.7 Å resolution and mechanism of its action. *FEBS Lett.*, **412**, 217–222.
55. Ghosh, M., Meiss, G., Pingoud, A., London, R.E. and Pedersen, L.C. (2005) Structural insights into the mechanism of nuclease A, a $\beta\beta\alpha$ metal nuclease from *Anabaena*. *J. Biol. Chem.*, **280**, 27990–27997.
56. Dupureur, C.M. (2008) Roles of metal ions in nucleases. *Curr. Opin. Chem. Biol.*, **12**, 250–255.
57. Llopis, J., McCaffery, J.M., Miyawaki, A., Farquhar, M.G. and Tsien, R.Y. (1998) Measurement of cytosolic, mitochondrial, and Golgi pH in single living cells with green fluorescent proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 6803–6808.
58. Santo-Domingo, J. and Demarex, N. (2012) The renaissance of mitochondrial pH. *J. Gen. Physiol.*, **139**, 415–423.
59. Gannavaram, S. and Debrabant, A. (2012) Programmed cell death in *Leishmania*: biochemical evidence and role in parasite infectivity. *Front. Cell Infect. Microbiol.*, **2**, 95.

60. Liu,Z., Chen,P., Wang,X., Cai,G., Niu,L., Teng,M. and Li,X. (2014) Crystal structure of DnaT84-153-dT10 ssDNA complex reveals a novel single-stranded DNA binding mode. *Nucleic Acids Res.*, **42**, 9470–9483.
61. Orans,J., McSweeney,E.A., Iyer,R.R., Hast,M.A., Hellinga,H.W., Modrich,P. and Beese,L.S. (2011) Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell*, **145**, 212–223.
62. Dickey,T.H., Altschuler,S.E. and Wuttke,D.S. (2013) Single-stranded DNA-binding proteins: multiple domains for multiple functions. *Structure*, **21**, 1074–1084.
63. Ceska,T.A. and Sayers,J.R. (1998) Structure-specific DNA cleavage by 5' nucleases. *Trends Biochem. Sci.*, **23**, 331–336.
64. Pazos,F., Helmer-Citterich,M., Ausiello,G. and Valencia,A. (1997) Correlated mutations contain information about protein-protein interaction. *J. Mol. Biol.*, **271**, 511–523.
65. Ballandras-Colas,A., Maskell,D.P., Serrao,E., Locke,J., Swuec,P., Jonsson,S.R., Kotecha,A., Cook,N.J., Pye,V.E., Taylor,I.A. *et al.* (2017) A supramolecular assembly mediates lentiviral DNA integration. *Science*, **355**, 93–95.
66. Parrish,J.Z., Yang,C., Shen,B. and Xue,D. (2003) CRN-1, a *Caenorhabditis elegans* FEN-1 homologue, cooperates with CPS-6/EndoG to promote apoptotic DNA degradation. *EMBO J.*, **22**, 3451–3460.
67. Ponce,I., Aldunate,C., Valenzuela,L., Sepulveda,S., Garrido,G., Kemmerling,U., Cabrera,G. and Galanti,N. (2016) A flap endonuclease (TcFEN1) is involved in *Trypanosoma cruzi* cell proliferation, DNA repair, and parasite survival. *J. Cell. Biochem.*, **118**, 1722–1732.
68. BoseDasgupta,S., Das,B.B., Sengupta,S., Ganguly,A., Roy,A., Dey,S., Tripathi,G., Dinda,B. and Majumder,H.K. (2008) The caspase-independent algorithm of programmed cell death in *Leishmania* induced by baicalein: the role of LdEndoG, LdFEN-1 and LdTatD as a DNA 'degradesome'. *Cell Death Differ.*, **15**, 1629–1640.
69. Balakrishnan,L. and Bambara,R.A. (2013) Flap endonuclease 1. *Annu. Rev. Biochem.*, **82**, 119–138.
70. Holm,L. and Laakso,L.M. (2016) Dali server update. *Nucleic Acids Res.*, **44**, W351–W355.

SUPPLEMENTAL DATA

Structure-based domain assignment in *Leishmania infantum* EndoG: characterization of a pH-dependent regulatory switch and a C-terminal extension that largely dictates DNA substrate preferences

Cristina Oliva, Pedro A. Sánchez-Murcia, Eva Rico, Ana Bravo, Margarita Menéndez, Federico Gago, and Antonio Jiménez-Ruiz

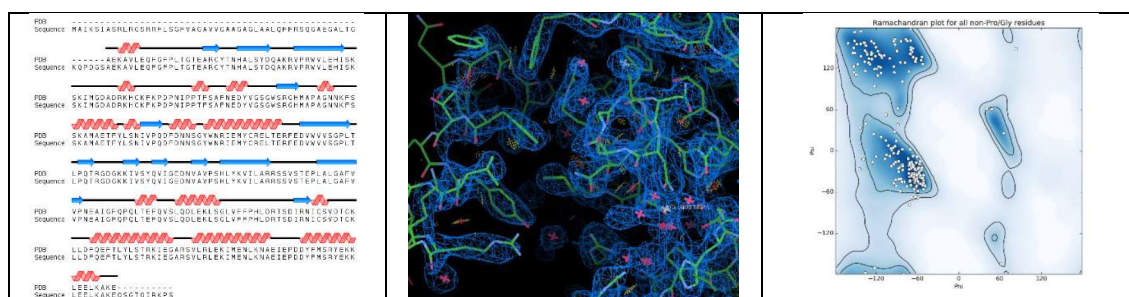


Figure S1. Sequence coverage (left), detail of the structural model fitted to the electron density map contoured at 1.0 σ in the vicinity of the active site and helix 1 of the CTD (middle), and Ramachandran plot (right) of PDB entry 4A1N (EXO) after refinement using Coot (1) and Phenix (2,3).

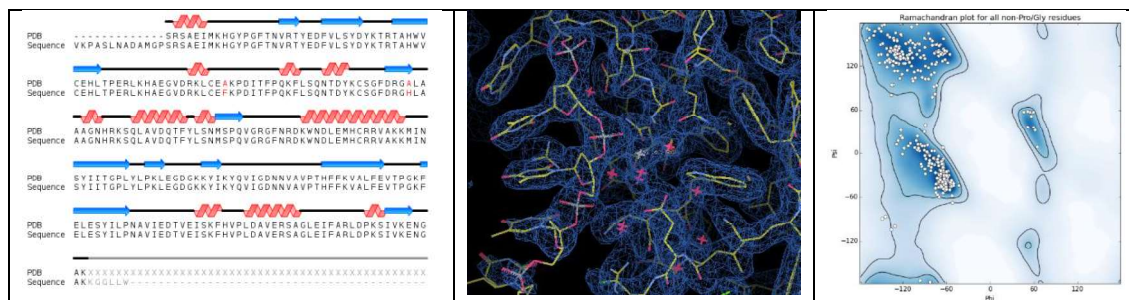


Figure S2. Sequence coverage (left), detail of the structural model fitted to the electron density map contoured at 1.0 σ in the vicinity of the bound oligonucleotide and Mg^{2+} ion (white cross) in chain A (middle), and Ramachandran plot (right) of PDB entry 5GKP (CeEndoG(H148A/F122A) bound to 5'-TTTTTGT-3' ssDNA) after refinement using Coot (1) and Phenix (2,3).


```

LiEndoG      ERLQRITGLQFFPGMDATRLP-DLCKTHKCAARPSALFQRYRQVAQLRAADSVPLRQTY 469
LiEndoG      ERLQRITGLQFFPGMDATRLP-DLCKTHKCDARPSALFQRYRQVAQLRAADSVPLRQTY 466
LmEndoG      EQLQRITGLQFFPGMDVTRLP-DLCKAHKCDARPSALFQRYRQVAQLRAADSVPLQQAY 461
LdEndoG      ERLQRITGLQFFPGMDATRLP-DLCKTHKCDARPSALFQRYRQVAQLRAADSVPLRQTY 466
LxEndoG      ERLQRITGLQFFPGMDATRLP-DLCKTHKCDARPSALFQRYRQVAQLRAADSVPLRKTY 468
LbEndoG      EQLQRITGLQFFPGMDATRLP-DLCKTHKCDARPPALFQRYRQVAQLRAAKSVPLRQTY 461
LpEndoG      EQLQRITGLQFFPGMDATRLP-DLCKTHKCDARPPALFQRYRQVAQLRAAKSVPLRQTY 461
HsEXOG       QDLEKLSGLVFFPHLDRTSDIRNICSVDTCKLLDFQEFTLYLSTRKIEGARSVLRLEKIM 329
ScNuc1p      DALERSTGLELLQKVPPSKKK-ALCKEVNCKIIVVR-DFS----NAAIKQSKDVK----LL 322
:  :::  :**  ::  :          *      *

```

Figure S3. Multiple sequence alignment showing the positions of the conserved (*) Cys residues (highlighted in yellow) immediately upstream of the CTD that can be involved in intramolecular disulfide bond formation. Sequences were retrieved from UniProtKB (4) using the following entries, from top to bottom: *Leishmania infantum* EndoG (B0ZTD3 and A4HUH3), *Leishmania major* EndoG (Q4QHF4), *Leishmania donovani* EndoG (E9BAB1), *Leishmania mexicana* EndoG (E9AN71), *Leishmania braziliensis* EndoG (A4H653), *Leishmania panamensis* EndoG (A0A088S3R7), human mitochondrial EXOG (Q9Y2C4), and *Saccharomyces cerevisiae* Nuc1p (C7GUE6).

(A)

```
LiEndoG      LASLNYERRIPNWVMEVVDYRKLHPGRRPSSPAAAAANRDDDAAQGEDEGSDGSRGGGD 180
LiEndoG      LASLNYERRIPNWVMEVVDYRKLHSGRRPSSPAAAAANRDDDAAQAADDEGSDGSRGGGD 180
LmEndoG      LASLNYERRIPNWVMEVVDYRKLHPGRRPSSPAAAAASRDDDAQDEDEGSDGSRGGGD 175
LcEndoG      LASLNYERRIPNWVMEVVDYRKLHSGRRPSSPAAAAANRDDDAAQAADDEGSDGSRGGGD 180
LxEndoG      LASLNYERRIPNWVMEVVDYRKLHSGRRPSSPAAAAANRDDDAAHGKEEGCNDTRGGDD 180
LbEndoG      LASLNYERRIPNWVMEVVDYRKLRSERRPSSPAVAATS--GDDAAPGEEVGCDSRGGDE 177
LpEndoG      LASLNYERRIPNWVMEVVDYRKLRSERRPSSPAVATTS--GDDAAPGEEVGCDSRGGDE 177
*****: *****.*:: * * .: *.: :*** :
```

(B)

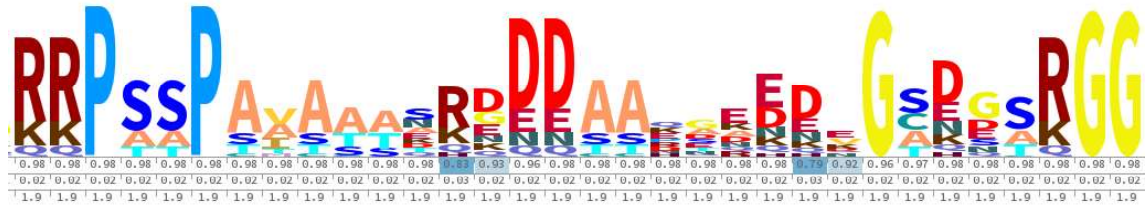


Figure S4. (A) CLUSTAL O 1.2.4 multiple sequence alignment (5) of the region encompassing the putative self-inhibitory domains of EndoG from *Leishmania* spp.: *Leishmania infantum* EndoG (B0ZTD3 and A4HUH3), *Leishmania major* EndoG (Q4QHF4), *Leishmania donovani* EndoG (E9BAB1), *Leishmania mexicana* EndoG (E9AN71), *Leishmania braziliensis* EndoG (A4H653), and *Leishmania panamensis* EndoG (A0A088S3R7). (B) WebLogo (6) produced by the SANSparallel server (7) for *Leishmania* EndoG sequences. Note the three contiguous Asp residues in the middle region that were replaced with Ala in our *LiEndoG*^{D(160-162)A} variant.

```

(A)
LiEndoG GLQFFPGMDAT-RLPDI C KTHKCAARPSALFQRYRQVAQLRAADSVLPQLRQTYAALQASAAGKLDQAVVHVEFQFMEEELVAAVGLIDQAERR--- 510
LiEndoG GLQFFPGMDAT-RLPDI C KTHKCDARPSALFQRYRQVAQLRAADSVLPQLRQTYAALQASAAGKLDQAVVHVEFQFMEEELVAAVGLIDQAERR--- 507
LmEndoG GLQFFPGMDVT-RLPDI C KAHKCDARPSALFQRYRQVAQLRAADSVLPQLQAYAAALQASAAGKLDQAVVHVEFQFMEEELVAAVGLIDQAERR--- 502
LdEndoG GLQFFPGMDAT-RLPDI C KTHKCDARPSALFQRYRQVAQLRAADSVLPQLRQTYAALQASAAGKLDQAVVHVEFQFMEEELVAAVGLIDQAERR--- 507
LxEndoG GLQFFPGMDAT-RLPDI C KTHKCDARPSALFQRYRQVAQLRAADSVLPQLRQTYAALQASAAGKLDQAVVHVEFQFMEEELVAASVGRIDHAEQR--- 509
LbEndoG GLQFFPGMDAT-RLPDI C KTHQCDAARPPALFQRYRQVAQLRAAKSVPLRQTYATLQASAAGKLDQAVVHVEFQFMEEELLAGAAGSGGREWRRD- 504
LpEndoG GLQFFPGMDAT-RLPDI C KTHQCDAARPPALFQRYRQVAQLRAAKSVPLRQTYATLQASAAGKLDQAVVHVEFQFMEEELLAGAAGSGGREWRRD- 504
HsEXOG GLVFFPHLDRTSDIRNI C SVDTC KLLLDQFETL YLSTKIKI EGARSVLRLEKIMVNLKNA--EIEPDDYFMSRYEKLLEELKAKEQSGTQ--IRKPSA 329
** *** : * * : : ** * * * : : * * : : * * : : * * : : * * : : * * : : * * : : * * : : * * : : * * : : * * : : * * : : * * : : * * : :

```

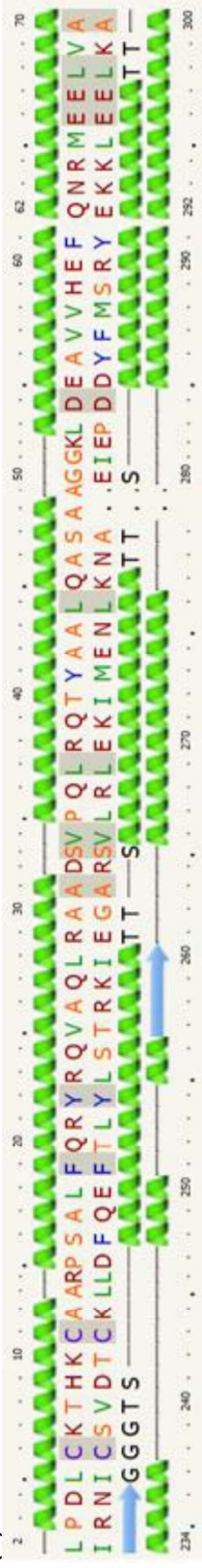
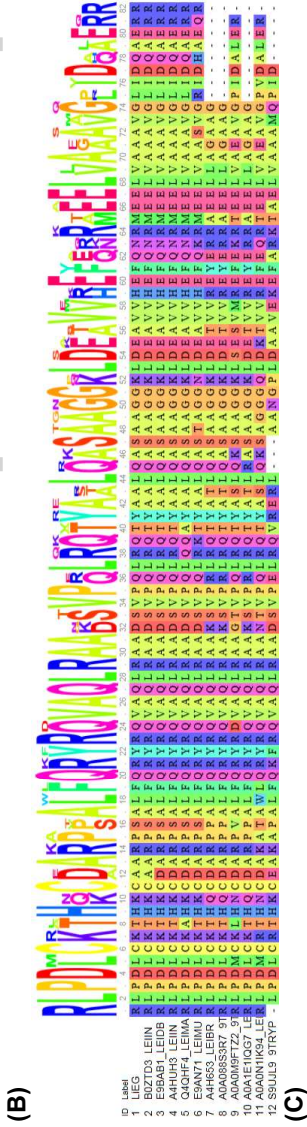


Figure S5. (A) CLUSTAL O 1.2.4 multiple sequence alignment (5) of the C-terminal domains (CTD) of EndoG from *Leishmania* spp. and human EXOG (same UniprotKB entry numbers as in Figure S3). Highlighted in yellow are Cys434 and Cys439 of *L*EndoG (positionally equivalent to Cys294 and Cys299 [PDB numbering] in EXOG), which are spatially close enough to be involved in a presumably redox-dependent disulfide bridge. The likely existence of this disulfide bond in EXOG is partially supported by the available X-ray crystal structure, which shows both thiol groups in close proximity although with no clear electron density between the sulfur atoms likely due to the fact that the protein was expressed in the cytoplasm of *E. coli* cells, where the thioredoxin/thioredoxin reductase and glutathione/glutaredoxin pathways provide a reducing environment (8). Highlighted in cyan are the aromatic residues responsible for base unstacking and backbone twisting of the substrate ssDNA 5'-upstream of the cleavage site. The three columns highlighted in grey correspond to residues that we propose are mostly responsible for the differences in exonuclease activity between *L*EndoG and EXOG, respectively; note that the Glu→Arg and Arg→Glu changes at these positions occur in a correlated fashion conceivably to maintain the helix1-helix3 electrostatically driven attraction in a manner similar to what is known for protein-protein interfaces (9). (B) SANSparallel (7) multiple sequence alignment of the CTD of EndoGs from trypanosomatids showing that the residues highlighted in (A) are strictly conserved. (C) Structure-based sequence alignment of *L*EndoG (top) and EXOG (bottom) CTDs. The respective secondary structures predicted by Phyre2 (10) are shown above and below, whereas the cartoon underneath the sequence corresponds to the secondary structure observed in our refined (1,3) crystal structure of EXOG (PDB entry 4A1N) (11).

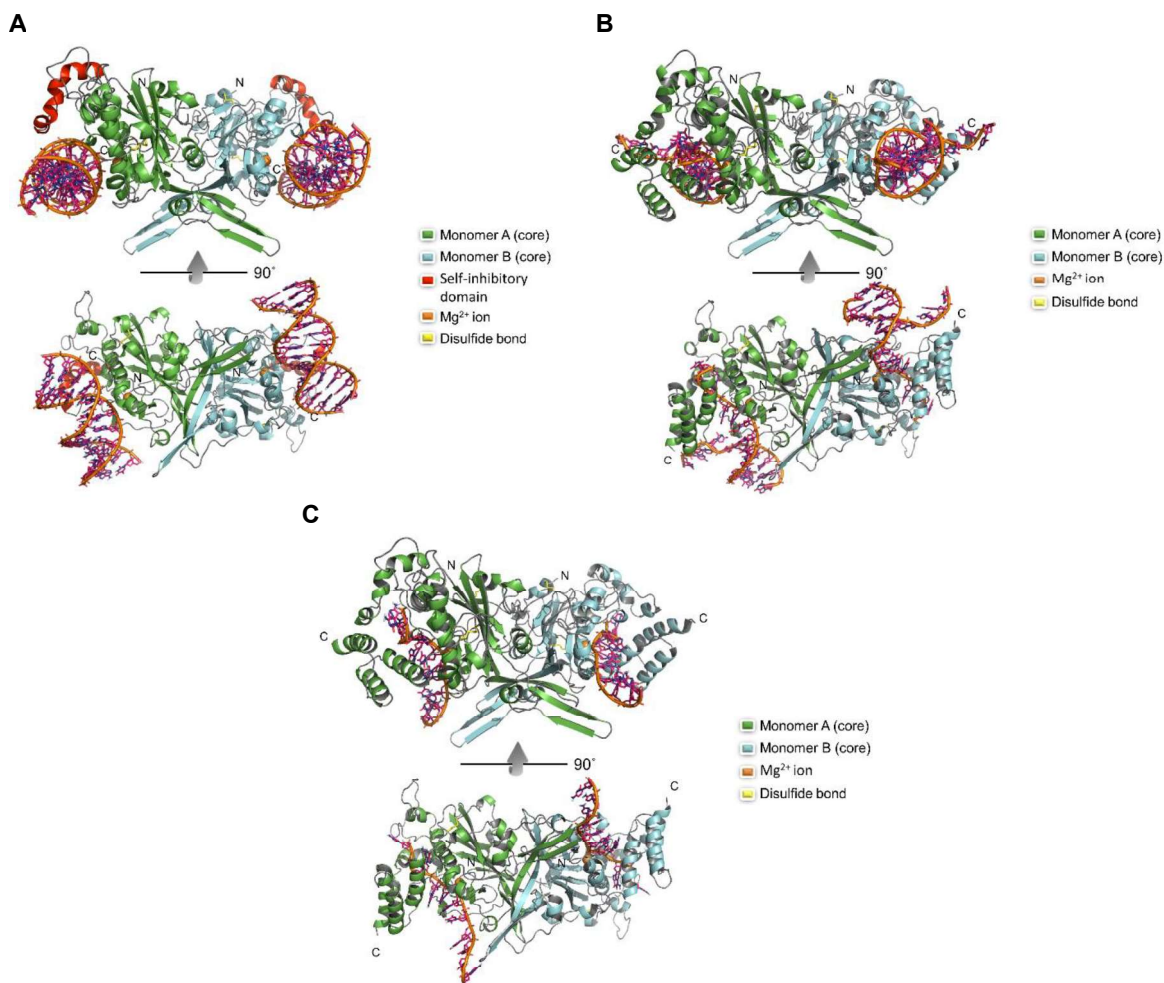


Figure S6. Molecular models of (A) homodimeric *LiEndoG* Δ ^{CTD} in complex with one dsDNA molecule bound to each monomer, (B) homodimeric wild-type *LiEndoG* in complex with one ssDNA-dsDNA junction bound to each monomer, and (C) homodimeric wild-type *LiEndoG* in complex with one ssDNA oligonucleotide bound to each monomer. The pH-dependent self-inhibitory domain in (A) has been coloured in red to highlight its spatial relation to the DNA-binding groove. The corresponding DNA sequences are displayed in Table 2 of the main text. Separate files containing the coordinates of these refined models are provided in PDB format as supplementary files *LiEndoG_apo.pdb*, *LiEndoG_ssDNA+dsDNA.pdb*, and *LiEndoG_ssDNA.pdb*.

Table S1. Refinement statistics^a of the PDB entries used as templates.

Metric	PDB entry								
	4A1N				5GKP				
	PDB ^b	PDB_REDO ^c	UAH ^d	PDB ^b	PDB_REDO ^c	UAH ^d	PDB ^b	PDB_REDO ^c	UAH ^d
R	0.1725	0.1815	0.1614	0.1660	0.1685	0.1576	0.1737	0.1844	0.1589
R-free	0.2227	0.2254	0.2215	0.2100	0.1990	0.2005	0.2493	0.2289	0.2247
Bond length RMSD	0.013	0.031	0.007	0.007	0.016	0.006	0.014	0.0108	0.007
Bond angle RMSD	1.74	1.97	0.82	1.05	1.53	0.793	1.41	1.307	0.837
Ramachandran outliers (%)	2.1	1.8	0	0	0	0	0	0	0
Ramachandran favored (%)	94.4	93.3	96.0	98.1	97.9	98.1	96.4	98.3	96.8
Rotamer outliers (%)	5.2	3.8	3.7	0.5	0.2	1.0	5.9	2.7	4.3
Clashscore	3.53	5.17	4.77	5.12	2.47	1.68	3.6	1.16	3.21
Overall score	2.07	2.15	1.81	1.27	1.06	0.92	1.98	1.15	1.79

^a Computed using MolProbity (12). The overall score represents the experimental resolution expected for a model of this quality; ideally, the score should be lower than the actual resolution.

^b Calculated from the coordinates deposited in the Protein Data Bank.

^c Calculated from the fully optimised coordinates deposited in the PDB_REDO database (13).

^d After our own refinement using Coot (1) and Phenix (2,3).

References

1. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*, **66**, 486-501.
2. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Echols, N., Headd, J.J., Hung, L.W., Jain, S., Kapral, G.J., Grosse Kunstleve, R.W. *et al.* (2011) The Phenix software for automated determination of macromolecular structures. *Methods*, **55**, 94-106.
3. 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. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, **66**, 213-221.
4. UniProt, C. (2015) UniProt: a hub for protein information. *Nucleic Acids Res.*, **43**, D204-212.
5. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **7**, 539.
6. Schneider, T.D. and Stephens, R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.*, **18**, 6097-6100.
7. Somervuo, P. and Holm, L. (2015) SANSparallel: interactive homology search against Uniprot. *Nucleic Acids Res.*, **43**, W24-29.
8. de Marco, A. (2009) Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microb Cell Fact*, **8**, 26.
9. Pazos, F., Helmer-Citterich, M., Ausiello, G. and Valencia, A. (1997) Correlated mutations contain information about protein-protein interaction. *J. Mol. Biol.*, **271**, 511-523.
10. Kelley, L.A. and Sternberg, M.J.E. (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nat. Protoc.*, **4**, 363-371.
11. Welin, M., Moche, M., Arrowsmith, C.H., Berglund, H., Bountra, C., Collins, R., Edwards, A.M., Flodin, S., Graslund, S., Hammarstrom, M. *et al.* (2012) Human mitochondrial endo-exonuclease. The Protein Data Bank, entry i.d. 4A1N (<http://www.rcsb.org/pdb/explore.do?structureId=4a1n>).
12. Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, **66**, 12-21.
13. Touw, W.G., Baakman, C., Black, J., te Beek, T.A.H., Krieger, E., Joosten, R.P. and Vriend, G. (2014) A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.*, **43**, D364-D368.

Conclusiones

El conjunto de resultados obtenidos en este trabajo, unido a los que habíamos obtenido previamente en nuestro laboratorio (101), muestran que *LiEndoG* es importante tanto para la supervivencia como para la ejecución de la muerte celular en *L. infantum*, tal y como se ha descrito también para sus ortólogos en otros organismos (129,131). Este novedoso papel provida de *LiEndoG* quedó demostrado al comprobar que la reducción de su expresión en *L. infantum* causaba un retraso en el crecimiento de los parásitos y disminuía su capacidad de infectar macrófagos. Así mismo, el hecho de que la inhibición de *LiEndoG* por acción de los derivados de la 5'-trilitimidina descritos en esta memoria produjera un efecto citotóxico sobre los promastigotes y amastigotes de *L. infantum*, y que este efecto pudiera ser contrarrestado por la sobreexpresión de *LiEndoG*, confirmaron la importancia que tiene esta proteína para la supervivencia de los parásitos y su potencial como diana terapéutica para el desarrollo de nuevos compuestos con actividad leishmanicida.

Nuestro trabajo nos ha permitido, además, construir un modelo tridimensional actualizado de *LiEndoG* en su forma dimérica e identificar varios dominios que confieren a esta enzima características prácticamente únicas. Una de estas características es la presencia de un dominio de autoinhibición rico en residuos ácidos de aspártico y glutámico que es capaz de inhibir la actividad enzimática de *LiEndoG* a pHs básicos, apuntando a un posible mecanismo de regulación por el cual esta enzima estaría prácticamente inactiva en condiciones normales en la mitocondria y se activaría tras ser liberada al citosol durante el proceso de muerte celular (173,174). Así mismo, hemos sido capaces de demostrar que el dominio carboxilo terminal de *LiEndoG* es responsable de la clara preferencia que muestra esta enzima por digerir sustratos de cadena sencilla y sobre la unión entre ADN de doble cadena y cadena sencilla, siendo este tipo de actividad similar al mostrado para otras nucleasas implicadas en la reparación y recombinación del ADN (116,117,136). Adicionalmente, hemos mostrado que el reemplazo de la serina del motivo SRGH - presente en las nucleasas de tripanosomátidos y en EXOG - por el aspártico canónico del motivo DRGH de esta familia de nucleasas da lugar a una enzima con una actividad nucleasa fuertemente mermada, efecto que puede ser contrarrestado en gran medida mediante el reemplazo concomitante de la arginina 247 por

una alanina. Finalmente, tanto nuestro modelo como los resultados experimentales obtenidos indican que *LiEndoG* es una enzima principalmente dimérica en la cual cada uno de los dos centros activos quedan localizados en lados opuestos de la interfaz de dimerización, sugiriendo que cada uno de ellos une e hidroliza sustratos de manera independiente.

En resumen, las conclusiones principales de los trabajos presentados en esta memoria se enumeran en el siguiente listado:

1. Los derivados de la 5' timidina con sustituyentes tipo tritilo en el carbono 5', y aminoácidos hidrofóbicos - o derivados de estos - en el carbono 3', constituyen un grupo novedoso de agentes con actividad leishmanicida capaces de producir un efecto citotóxico sobre promastigotes y amastigotes de *L. infantum*.
2. Dentro de estos derivados de timidina, la 3'-O-(isoleucilisoleucil)-5'-O-(3,3,3,-trifenilpropanoil)timidina, denominada **compuesto 28** en esta memoria, ha mostrado ser el que tiene mayor potencial para actuar sobre los amastigotes intracelulares, reduciendo hasta en un 90% el porcentaje de macrófagos infectados tras el tratamiento con este compuesto a una concentración de 15 μ M.
3. El tratamiento de los parásitos con el **compuesto 28** produce un aumento de su potencial de membrana mitocondrial y de la producción de anión superóxido, eventos que pueden ser los desencadenantes de la muerte celular producida por esta molécula. Adicionalmente, los parásitos que mueren por acción de este compuesto muestran una menor degradación del ADN que aquellos que lo hacen por efecto del tratamiento control.
4. Varios de estos compuestos han mostrado ser capaces de inhibir la actividad nucleasa de *LiEndoG in vitro*, existiendo una relación entre su capacidad inhibitoria y el efecto citotóxico que causan sobre los parásitos e indicando que esta proteína podría ser la diana sobre la que están actuando en los parásitos.

5. La sobreexpresión de *LiEndoG* hace que los parásitos sean más resistentes a la muerte inducida por el compuesto **Lei49 (compuesto 37)**.
6. La disminución de los niveles de expresión de *LiEndoG* en los parásitos produce una reducción de su tasa de crecimiento y hace que sean mucho menos infectivos para los macrófagos.
7. Mediante la utilización de múltiples técnicas de modelado molecular y basándonos en las similitudes que existen entre *LiEndoG* y las nucleasas de *C. elegans*, *D. melanogaster* o EXOG, en el laboratorio hemos construido un nuevo modelo tridimensional de *LiEndoG* en su forma dimérica.
8. La serina 211 del motivo SRGH de *LiEndoG* no parece ser esencial para la actividad catalítica, dado que su reemplazo por una alanina da lugar a una variante que muestra una actividad nucleasa tan solo ligeramente reducida.
9. Sin embargo, el reemplazo de esta Ser211 por el aspártico canónico del motivo DRGH de otras EndoGs da lugar a una enzima cuya actividad está drásticamente reducida. No obstante, la actividad de la variante S211D puede ser recuperada tras el reemplazo concomitante de la Ala247 por una Arg.
10. La Arg y la His del motivo SRGH son indispensables para la actividad nucleasa de *LiEndoG* dado su papel en la unión del sustrato y la catálisis, respectivamente.
11. *LiEndoG* es una proteína principalmente homodimérica (si bien también se ha demostrado la existencia de la forma monomérica y tetramérica) donde cada uno de los dos centros activos quedan localizados en lados opuestos de la interfaz de dimerización.

12. Los residuos comprendidos entre las posiciones 145-180 de *LiEndoG* constituyen un dominio de autoinhibición que se localiza de forma próxima al sitio activo y es capaz de inhibir la actividad de la enzima a pHs iguales o superiores a 8. Este mecanismo de regulación podría ser importante para mantener la actividad nucleasa de esta enzima bajo control en condiciones de crecimiento normales en los parásitos.

13. *LiEndoG* produce cortes de cadena sencilla sobre el ADN bicatenario superenrollado y tiene preferencia por digerir en las proximidades del extremo 5' sobre el ADN de doble cadena lineal.

14. La extensión de la región carboxilo terminal de *LiEndoG*, que abarca desde el residuo 440 hasta el 510, conforma un dominio que dicta la preferencia que tiene esta enzima por digerir sustratos de cadena sencilla ya sea de forma aislada o en la base de la unión entre un ADN de cadena sencilla y un ADN de doble cadena.

Bibliografía

1. Pace D. Leishmaniasis. *J Infect.* 2014;69(S1):S10–8.
2. Akhoundi M, Kuhls K, Cannet A, Votýpka J, Marty P, Delaunay P, et al. A Historical Overview of the Classification, Evolution, and Dispersion of Leishmania Parasites and Sandflies. *PLoS Negl Trop Dis.* 2016;10(3):e0004349.
3. Ashford RW. The leishmaniasis as emerging and reemerging zoonoses. *Int J Parasitol.* 2000;30(12–13):1269–81.
4. Burza S, Croft SL, Boelaert M. Leishmaniasis. *Lancet.* 2018;392(10151):951–70.
5. Murray HW, Berman JD, Davies CR, Saravia NG. Advances in leishmaniasis. *Lancet.* 2005;366(9496):1561–77.
6. Herwaldt BL. Leishmaniasis. *Lancet.* 1999;354(9185):1191–9.
7. Desjeux P. Leishmaniasis: current situation and new perspectives. *Comp Immunol Microbiol Infect Dis.* 2004;27(5):305–18.
8. Alvar J, Aparicio P, Aseffa A, Den Boer M, Cañavate C, Dedet JP, et al. The relationship between leishmaniasis and AIDS: The second 10 years. *Clin Microbiol Rev.* 2008;21(2):334–59.
9. Van Griensven J, Carrillo E, López-Vélez R, Lynen L, Moreno J. Leishmaniasis in immunosuppressed individuals. *Clin Microbiol Infect.* 2014;20(4):286–99.
10. World Health Organization - Global Health Observatory data repository - Leishmaniasis [Internet]. [cited 2020 Feb 7]. Available from: <http://apps.who.int/gho/data/node.main.NTDLEISH?lang=en>
11. World Health Organization - Leishmaniasis Fact Sheet [Internet]. [cited 2020 Feb 7]. Available from: www.who.int/news-room/fact-sheets/detail/leishmaniasis
12. Dujardin JC, Campino L, Cañavate C, Dedet JP, Gradoni L, Soteriadou K, et al. Spread of vector-borne diseases and neglect of leishmaniasis, Europe. *Emerg Infect Dis.* 2008;14(7):1013–8.

13. Gálvez R, Descalzo MA, Guerrero I, Miró G, Molina R. Mapping the Current Distribution and Predicted Spread of the Leishmaniosis Sand Fly Vector in the Madrid Region (Spain) Based on Environmental Variables and Expected Climate Change. *Vector-Borne Zoonotic Dis.* 2011;11(7):799–806.
14. Gomez-Barroso D, Herrador Z, San Martín J V, Gherasim A, Aguado M, Romero-Maté A, et al. Spatial distribution and cluster analysis of a leishmaniasis outbreak in the south-western Madrid region, Spain, September 2009 to April 2013. *Eurosurveillance.* 2015;20(7):1–10.
15. Chicharro C, Llanes-Acevedo P, García E, Nieto J, Moreno J, Cruz I. Molecular typing of leishmania infantum isolates from a leishmaniasis outbreak in madrid, Spain, 2009 to 2012. *Eurosurveillance.* 2013;18(30):1–13.
16. Red Nacional de Vigilancia Epidemiológica. Resultados de la Vigilancia Epidemiológica de las enfermedades transmisibles. Informe anual. Año 2016.
17. Cavalcanti DP, De Souza W. The kinetoplast of trypanosomatids: From early studies of electron microscopy to recent advances in atomic force microscopy. *Scanning.* 2018;2018:9603051.
18. Dostálová A, Volf P. Leishmania development in sand flies: Parasite-vector interactions overview. *Parasites and Vectors.* 2012;5(1):1–12.
19. Teixeira DE, Benchimol M, Rodrigues JCF, Crepaldi PH, Pimenta PFP, de Souza W. The Cell Biology of Leishmania: How to Teach Using Animations. *PLoS Pathog.* 2013;9(10):e1003594.
20. Sunter J, Gull K. Shape, form, function and Leishmania pathogenicity: from textbook descriptions to biological understanding. *Open Biol.* 2017;7(9):170165.
21. Sacks DL. Metacyclogenesis in Leishmania promastigotes. *Exp Parasitol.* 1989;69(1):100–3.
22. Bates P, Rogers M. New Insights into the Developmental Biology and Transmission Mechanisms of Leishmania. *Curr Mol Med.* 2005;4(6):601–9.
23. Bates PA. Transmission of Leishmania metacyclic promastigotes by phlebotomine sand

- flies. *Int J Parasitol.* 2007;37(10):1097–106.
24. Sacks DL, Saraiva EM, Rowton E, Turco SJ, Pimenta PF. The role of the lipophosphoglycan of *Leishmania* in vector competence. *Parasitology.* 1994;108(S1):S55–62.
 25. Voth BR, Kelly BL, Joshi PB, Ivens AC, McMaster WR. Differentially expressed *Leishmania major* gp63 genes encode cell surface leishmanolysin with distinct signals for glycosylphosphatidylinositol attachment. *Mol Biochem Parasitol.* 1998;93(1):31–41.
 26. Brittingham A, Morrison CJ, McMaster WR, McGwire BS, Chang K-P, Mosser DM. Role of the *Leishmania* surface protease gp63 in complement fixation, cell adhesion, and resistance to complement-mediated lysis. *Parasitol Today.* 1995;11(12):445–6.
 27. Cunningham AC. Parasitic adaptive mechanisms in infection by *Leishmania*. *Exp Mol Pathol.* 2002;72(2):132–41.
 28. Wilson ME, Hardin KK, Donelson JE. Expression of the major surface glycoprotein of *Leishmania donovani* chagasi in virulent and attenuated promastigotes. *J Immunol.* 1989;143(2):678–84.
 29. Kane MM, Mosser DM. *Leishmania* parasites and their ploys to disrupt macrophage activation. *Curr Opin Hematol.* 2000;7(1):26–31.
 30. van Zandbergen G, Bollinger A, Wenzel A, Kamhawi S, Voll R, Klinger M, et al. *Leishmania* disease development depends on the presence of apoptotic promastigotes in the virulent inoculum. *Proc Natl Acad Sci U S A.* 2006;103(37):13837–42.
 31. Fadok VA, Voelker DR, Campbell PA, Cohen JJ, Bratton DL, Henson PM. Exposure of phosphatidylserine on the surface of apoptotic lymphocytes triggers specific recognition and removal by macrophages. *J Immunol.* 1992;148(7):2207–16.
 32. McDonald PP, Fadok VA, Bratton D, Henson PM. Transcriptional and translational regulation of inflammatory mediator production by endogenous TGF-beta in macrophages that have ingested apoptotic cells. *J Immunol.* 1999;163(11):6164–72.
 33. Laufs H, Müller K, Fleischer J, Reiling N, Jahnke N, Jensenius JC, et al. Intracellular

- survival of *Leishmania major* in neutrophil granulocytes after uptake in the absence of heat-labile serum factors. *Infect Immun.* 2002;70(2):826–35.
34. van Zandbergen G, Klinger M, Mueller A, Dannenberg S, Gebert A, Solbach W, et al. Cutting Edge: Neutrophil Granulocyte Serves as a Vector for *Leishmania* Entry into Macrophages. *J Immunol.* 2004;173(11):6521–5.
 35. Antoine JC, Prina E, Lang T, Courret N. The biogenesis and properties of the parasitophorous vacuoles that harbour *Leishmania* in murine macrophages. *Trends Microbiol.* 1998;6(10):392–401.
 36. Cantacessi C, Dantas-Torres F, Nolan MJ, Otranto D. The past, present, and future of *Leishmania* genomics and transcriptomics. *Trends Parasitol.* 2015;31(3):100–8.
 37. Sterkers Y, Lachaud L, Bourgeois N, Crobu L, Bastien P, Pagès M. Novel insights into genome plasticity in Eukaryotes: mosaic aneuploidy in *Leishmania*. *Mol Microbiol.* 2012;86(1):15–23.
 38. Myler PJ. Genome structure and content. In: *Leishmania after the genome*. Caister Academic Press; 2008. p. 15–23.
 39. De Gaudenzi JG, Noé G, Campo VA, Frasch AC, Cassola A. Gene expression regulation in trypanosomatids. *Essays Biochem.* 2011;51(1):31–46.
 40. Fidalgo LM, Gille L. Mitochondria and trypanosomatids: Targets and drugs. *Pharm Res.* 2011;28(11):2758–70.
 41. Opperdoes FR, Michels PAM. The glycosomes of the Kinetoplastida. *Biochimie.* 1993;75(3–4):231–4.
 42. Haanstra JR, González-Marcano EB, Gualdrón-López M, Michels PAM. Biogenesis, maintenance and dynamics of glycosomes in trypanosomatid parasites. *Biochim Biophys Acta - Mol Cell Res.* 2016;1863(5):1038–48.
 43. Miranda K, Docampo R, Grillo O, Franzen A, Attias M, Vercesi A, et al. Dynamics of polymorphism of acidocalcisomes in *Leishmania* parasites. *Histochem Cell Biol.* 2004;121(5):407–18.
 44. Docampo R, Huang G. Calcium signaling in trypanosomatid parasites. *Cell Calcium.*

- 2015;57(3):194–202.
45. McConville M, De Souza D, Saunders E, Pyke J, Naderer T, Ellis M, et al. Analysis of the *Leishmania* metabolome. In: *Leishmania after the genome*. Caister Academic Press; 2008. p. 75–106.
 46. Davies CR, Kaye P, Croft SL, Sundar S. Leishmaniasis: New approaches to disease control. *Br Med J*. 2003;326(7385):377–82.
 47. Ghorbani M, Farhodi R. Leishmaniasis in humans: Drug or vaccine therapy? *Drug Des Devel Ther*. 2018;12:25–40.
 48. Velez R, Gállego M. Commercially approved vaccines for canine leishmaniasis: a review of available data on their safety and efficacy. *Trop Med Int Heal*. 2020;tmi.13382.
 49. Velez R, Domenech E, Rodríguez-Cortés A, Barrios D, Tebar S, Fernández-Arévalo A, et al. Evaluation of canine leishmaniasis vaccine CaniLeish® under field conditions in native dog populations from an endemic Mediterranean area—A randomized controlled trial. *Acta Trop*. 2020;205:105387.
 50. Wyllie S, Cunningham ML, Fairlamb AH. Dual action of antimonial drugs on thiol redox metabolism in the human pathogen *Leishmania donovani*. *J Biol Chem*. 2004;279(38):39925–32.
 51. Ouellette M, Drummel-Smith J, Papadopoulou B. Leishmaniasis: Drugs in the clinic, resistance and new developments. *Drug Resist Updat*. 2004;7(4–5):257–66.
 52. Croft SL, Barrett MP, Urbina JA. Chemotherapy of trypanosomiasis and leishmaniasis. *Trends Parasitol*. 2005;21(11):508–12.
 53. Ramos H, Valdivieso E, Gamargo M, Dagger F, Cohen BE. Amphotericin B kills unicellular leishmanias by forming aqueous pores permeable to small cations and anions. *J Membr Biol*. 1996;152(1):65–75.
 54. De Menezes JPB, Guedes CES, De Oliveira Almeida Petersen AL, Fraga DBM, Veras PST, Bezerra de Menezes J, et al. Advances in Development of New Treatment for Leishmaniasis. *Biomed Res Int*. 2015;2015:15–8.
 55. Mishra J, Saxena A, Singh S. Chemotherapy of Leishmaniasis: Past, Present and Future.

- Front Med Chem. 2013;97–130.
56. Sundar S, Chakravarty J. Paromomycin in the treatment of leishmaniasis. *Expert Opin Investig Drugs*. 2008;17(5):787–94.
 57. Verma NK, Dey CS. Possible mechanism of miltefosine-mediated death of *Leishmania donovani*. *Antimicrob Agents Chemother*. 2004;48(8):3010–5.
 58. Sundar S, Chakravarty J. Leishmaniasis: An update of current pharmacotherapy. *Expert Opin Pharmacother*. 2013;14(1):53–63.
 59. Sundar S, Chakravarty J. An update on pharmacotherapy for leishmaniasis. *Expert Opin Pharmacother*. 2015;16(2):237–52.
 60. Nassif PW, De Mello TFP, Navasconi TR, Mota CA, Demarchi IG, Aristides SMA, et al. Safety and efficacy of current alternatives in the topical treatment of cutaneous leishmaniasis: A systematic review. *Parasitology*. 2017;144(8):995–1004.
 61. Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P, et al. Molecular mechanisms of cell death: Recommendations of the Nomenclature Committee on Cell Death 2018. *Cell Death Differ*. 2018;25(3):486–541.
 62. Green DR, Fitzgerald P. Just So Stories about the Evolution of Apoptosis. *Curr Biol*. 2016;26(13):R620–7.
 63. Galluzzi L, Maiuri MC, Vitale I, Zischka H, Castedo M, Zitvogel L, et al. Cell death modalities: Classification and pathophysiological implications. *Cell Death Differ*. 2007;14(7):1237–43.
 64. Edinger AL, Thompson CB. Death by design: Apoptosis, necrosis and autophagy. *Curr Opin Cell Biol*. 2004;16(6):663–9.
 65. Kroemer G, El-Deiry WS, Golstein P, Peter ME, Vaux D, Vandenabeele P, et al. Classification of cell death: recommendations of the Nomenclature Committee on Cell Death. *Cell Death Differ*. 2005;12 Suppl 2:1463–7.
 66. Madeo F, Carmona-Gutierrez D, Ring J, Büttner S, Eisenberg T, Kroemer G. Caspase-dependent and caspase-independent cell death pathways in yeast. *Biochem Biophys Res Commun*. 2009;382(2):227–31.

67. Kerr JFR, Wyllie AH, Currie AR. Apoptosis: A basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer*. 1972;26(4):239–57.
68. Horvitz HR. Genetic control of programmed cell death in the nematode *Caenorhabditis elegans*. *Cancer Res*. 1999;59(7 SUPPL.):1–13.
69. Gumienny TL, Lambie E, Hartwig E, Horvitz HR, Hengartner MO. Genetic control of programmed cell death in the *Caenorhabditis elegans* hermaphrodite germline. *Development*. 1999;126(5):1011–22.
70. Zmasek CM, Godzik A. Evolution of the animal apoptosis network. *Cold Spring Harb Perspect Biol*. 2013;5(3).
71. Ameisen JC. On the origin, evolution, and nature of programmed cell death: A timeline of four billion years. *Cell Death Differ*. 2002;9(4):367–93.
72. Kiraz Y, Adan A, Kartal Yandim M, Baran Y. Major apoptotic mechanisms and genes involved in apoptosis. *Tumor Biol*. 2016;37(7):8471–86.
73. Metzstein MM, Stanfield GM, Horvitz HR. Genetics of programmed cell death in *C. elegans*: Past, present and future. *Trends Genet*. 1998;14(10):410–6.
74. Carmona-Gutierrez D, Bauer MA, Zimmermann A, Aguilera A, Austriaco N, Ayscough K, et al. Guidelines and recommendations on yeast cell death nomenclature. *Microb Cell*. 2018;5(1):4–31.
75. Galluzzi L, Vitale I, Abrams JM, Alnemri ES, Baehrecke EH, Blagosklonny M V., et al. Molecular definitions of cell death subroutines: Recommendations of the Nomenclature Committee on Cell Death 2012. *Cell Death Differ*. 2012;19(1):107–20.
76. Nedelcu AM, Driscoll WW, Durand PM, Herron MD, Rashidi A. On the paradigm of altruistic suicide in the unicellular world. *Evolution (N Y)*. 2011;65(1):3–20.
77. Welburn SC, Barcinski MA, Williams GT. Programmed cell death in trypanosomatids. *Parasitol Today*. 1997;13(1):22–6.
78. Lee N, Bertholet S, Debrabant A, Muller J, Duncan R, Nakhasi HL. Programmed cell death in the unicellular protozoan parasite *Leishmania*. *Cell Death Differ*. 2002;9(1):53–64.

79. Debrabant A, Nakhasi H. Programmed cell death in trypanosomatids: is it an altruistic mechanism for survival of the fittest? *Kinetoplastid Biol Dis.* 2003;2(1):7.
80. Lüder CG, Campos-Salinas J, Gonzalez-Rey E, Van Zandbergen G. Impact of protozoan cell death on parasite-host interactions and pathogenesis. *Parasites and Vectors.* 2010;3(1):116.
81. Bruchhaus I, Roeder T, Rennenberg A, Heussler VT. Protozoan parasites: programmed cell death as a mechanism of parasitism. *Trends Parasitol.* 2007;23(8):376–83.
82. Wanderley JLM, Pinto da Silva LH, Deolindo P, Soong L, Borges VM, Prates DB, et al. Cooperation between apoptotic and viable metacyclics enhances the pathogenesis of Leishmaniasis. *PLoS One.* 2009;4(5):e5733.
83. Moreira ME, Del Portillo HA, Milder R V, Balanco JM, Barcinski MA. Heat shock induction of apoptosis in promastigotes of the unicellular organism *Leishmania (Leishmania) amazonensis*. *J Cell Physiol.* 1996;167(2):305–13.
84. Das M, Mukherjee SB, Shaha C. Hydrogen peroxide induces apoptosis-like death in *Leishmania donovani* promastigotes. *J Cell Sci.* 2001;114(13):2461–9.
85. Holzmüller P, Sereno D, Cavaleyra M, Mangot I, Daulouede S, Vincendeau P, et al. Nitric oxide-mediated proteasome-dependent oligonucleosomal DNA fragmentation in *Leishmania amazonensis* amastigotes. *Infect Immun.* 2002;70(7):3727–35.
86. Alzate JF, Arias AA, Moreno-Mateos D, Alvarez-Barrientos A, Jiménez-Ruiz A. Mitochondrial superoxide mediates heat-induced apoptotic-like death in *Leishmania infantum*. *Mol Biochem Parasitol.* 2007;152(2):192–202.
87. Smirlis D, Duszenko M, Ruiz AJ, Scoulica E, Bastien P, Fasel N, et al. Targeting essential pathways in trypanosomatids gives insights into protozoan mechanisms of cell death. *Parasit Vectors.* 2010;3(1):107.
88. Jiménez-Ruiz A, Alzate JF, MacLeod ET, Lüder CGK, Fasel N, Hurd H. Apoptotic markers in protozoan parasites. *Parasites and Vectors.* 2010;3(1):104.
89. Das R, Roy A, Dutta N, Majumder HK. Reactive oxygen species and imbalance of calcium homeostasis contributes to curcumin induced programmed cell death in *Leishmania*

- donovani. Apoptosis. 2008;13(7):867–82.
90. Menna-Barreto RFS. Cell death pathways in pathogenic trypanosomatids: lessons of (over)kill. *Cell Death Dis.* 2019;10(2):1–11.
 91. Alzate JF, Arias A, Mollinedo F, Rico E, de la Iglesia-Vicente J, Jiménez-Ruiz A. Edelfosine induces an apoptotic process in *Leishmania infantum* that is regulated by the ectopic expression of Bcl-XL and Hrk. *Antimicrob Agents Chemother.* 2008;52(10):3779–82.
 92. Sen N, Das BB, Ganguly A, Mukherjee T, Tripathi G, Bandyopadhyay S, et al. Camptothecin induced mitochondrial dysfunction leading to programmed cell death in unicellular hemoflagellate *Leishmania donovani*. *Cell Death Differ.* 2004;11(April):924–36.
 93. Vercammen D, Declercq W, Vandenabeele P, Van Breusegem F. Are metacaspases caspases? *J Cell Biol.* 2007;179(3):375–80.
 94. Zalila H, González IJ, El-Fadili AK, Delgado MB, Desponds C, Schaff C, et al. Processing of metacaspase into a cytoplasmic catalytic domain mediating cell death in *Leishmania major*. *Mol Microbiol.* 2011;79(1):222–39.
 95. González IJ, Desponds C, Schaff C, Mottram JC, Fasel N. *Leishmania major* metacaspase can replace yeast metacaspase in programmed cell death and has arginine-specific cysteine peptidase activity. *Int J Parasitol.* 2007;37(2):161–72.
 96. Casanova M, Gonzalez IJ, Sprissler C, Zalila H, Dacher M, Basmaciyan L, et al. Implication of different domains of the *Leishmania major* metacaspase in cell death and autophagy. *Cell Death Dis.* 2015;6(10):e1933.
 97. Genes CM, de Lucio H, González VM, Sánchez-Murcia PA, Rico E, Gago F, et al. A functional BH3 domain in an aquaporin from *Leishmania infantum*. *Cell death Discov.* 2016;2(1):16043.
 98. Basmaciyan L, Jacquet P, Azas N, Casanova M. A novel hydrolase with a pro-death activity from the protozoan parasite *Leishmania major*. *Cell Death Discov.* 2019;5(1):99.
 99. Basmaciyan L, Azas N, Casanova M. A potential acetyltransferase involved in

- Leishmania major metacaspase-dependent cell death. *Parasit Vectors*. 2019;12(1):266.
100. BoseDasgupta S, Das BB, Sengupta S, Ganguly A, Roy A, Dey S, et al. The caspase-independent algorithm of programmed cell death in Leishmania induced by baicalein: the role of LdEndoG, LdFEN-1 and LdTatD as a DNA “degradesome”. *Cell Death Differ*. 2008;15(10):1629–40.
 101. Rico E, Alzate JF, Arias AA, Moreno D, Clos J, Gago F, et al. Leishmania infantum expresses a mitochondrial nuclease homologous to EndoG that migrates to the nucleus in response to an apoptotic stimulus. *Mol Biochem Parasitol*. 2009;163(1):28–38.
 102. Basmacıyan L, Casanova M. Cell death in Leishmania. *Parasite*. 2019;26:71.
 103. Proto WR, Coombs GH, Mottram JC. Cell death in parasitic protozoa: Regulated or incidental? *Nat Rev Microbiol*. 2013;11(1):58–66.
 104. Curtis PJ, Burdon MG, Smellie RM. The purification from rat liver of a nuclease hydrolysing ribonucleic acid and deoxyribonucleic acid. *Biochem J*. 1966;98(3):813–7.
 105. Linn S, Lehman IR. An endonuclease from mitochondria of Neurospora crassa. *J Biol Chem*. 1966;241(11):2694–9.
 106. Curtis PJ, Smellie RM. Properties of a purified rat-liver nuclease. *Biochem J*. 1966;98(3):818–25.
 107. Kwong S, Fraser MJ. Neurospora endoexonuclease and its inactive (precursor?) form. *Can J Biochem*. 1978;56(6):370–7.
 108. Fraser MJ, Tjeerde R, Matsumoto K. A second form of the single-strand specific endonuclease of Neurospora crassa which is associated with a double-strand exonuclease. *Can J Biochem*. 1976;54(11):971–80.
 109. Fraser MJ, Hatahet Z, Huang XT. The actions of Neurospora endo-exonuclease on double strand DNAs. *J Biol Chem*. 1989;264(22):13093–101.
 110. Dake E, Hofmann TJ, McIntire S, Hudson A, Zassenhaus HP. Purification and properties of the major nuclease from mitochondria of Saccharomyces cerevisiae. *J Biol Chem*. 1988;263(16):7691–702.

111. Vincent RD, Hofmann TJ, Zassenhaus HP. Sequence and expression of NUC1, the gene encoding the mitochondrial nuclease in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 1988;16(8):3297–312.
112. Tomkinson AE, Linn S. Purification and properties of a single strand-specific endonuclease from mouse cell mitochondria. *Nucleic Acids Res.* 1986;14(24):9579–93.
113. Cummings OW, King TC, Holden JA, Low RL. Purification and characterization of the potent endonuclease in extracts of bovine heart mitochondria. *J Biol Chem.* 1987;262(5):2005–15.
114. Ruiz-Carrillo A, Renaud J. Endonuclease G: a (dG)_n X (dC)_n-specific DNase from higher eukaryotes. *EMBO J.* 1987;6(2):401–7.
115. Gerschenson M, Houmiel KL, Low RL. Endonuclease G from mammalian nuclei is identical to the major endonuclease of mitochondria. Vol. 23, *Nucleic Acids Research.* 1995.
116. Low RL. Mitochondrial Endonuclease G function in apoptosis and mtDNA metabolism: a historical perspective. *Mitochondrion.* 2003;2(4):225–36.
117. Cymerman I a, Chung I, Beckmann BM, Bujnicki JM, Meiss G. EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Res.* 2008;36(4):1369–79.
118. Miller MD, Tanner J, Alpaugh M, Benedik MJ, Krause KL. 2.1 Å structure of serratia endonuclease suggests a mechanism for binding to doublestranded dna. *Nat Struct Biol.* 1994;1(7):461–8.
119. Friedhoff P, Kolmes B, Gimadutdinow O, Wende W, Krause KL, Pingoud A. Analysis of the mechanism of the *Serratia* nuclease using site-directed mutagenesis. *Nucleic Acids Res.* 1996;24(14):2632–9.
120. Schäfer P, Scholz SR, Gimadutdinow O, Cymerman IA, Bujnicki JM, Ruiz-Carrillo A, et al. Structural and functional characterization of mitochondrial EndoG, a sugar non-specific nuclease which plays an important role during apoptosis. *J Mol Biol.* 2004;338(2):217–28.
121. Kieper J, Lauber C, Gimadutdinow O, Urbańska A, Cymerman I, Ghosh M, et al.

- Production and characterization of recombinant protein preparations of Endonuclease G-homologs from yeast, *C. elegans* and humans. *Protein Expr Purif.* 2010;73(1):99–106.
122. Lin JLJL, Wu C-C, Yang W-Z, Yuan HS. Crystal structure of endonuclease G in complex with DNA reveals how it nonspecifically degrades DNA as a homodimer. *Nucleic Acids Res.* 2016;44(21):10480–90.
 123. Loll B, Gebhardt M, Wahle E, Meinhart A. Crystal structure of the EndoG/EndoGI complex: Mechanism of EndoG inhibition. *Nucleic Acids Res.* 2009;37(21):7312–20.
 124. Côté J, Ruiz-Carrillo A. Primers for mitochondrial DNA replication generated by endonuclease G. *Science (80-).* 1993;261(5122):765–9.
 125. Li LY, Luo X, Wang X. Endonuclease G is an apoptotic DNase when released from mitochondria. *Nature.* 2001;412(6842):95–9.
 126. Ball TK, Suh Y, Benedik MJ. Disulfide bonds are required for *Serratia marcescens* nuclease activity. *Nucleic Acids Res.* 1992;20(19):4971–4.
 127. Wu SI, Lo SK, Shao CP, Tsai HW, Hor LI. Cloning and characterization of a periplasmic nuclease of *Vibrio vulnificus* and its role in preventing uptake of foreign DNA. *Appl Environ Microbiol.* 2001;67(1):82–8.
 128. Huang K-J, Ku C-C, Lehman IR. Endonuclease G: a role for the enzyme in recombination and cellular proliferation. *Proc Natl Acad Sci U S A.* 2006;103(24):8995–9000.
 129. Parrish J, Li L, Klotz K, Ledwich D, Wang X, Xue D. Mitochondrial endonuclease G is important for apoptosis in *C. elegans*. *Nature.* 2001;412(6842):90–4.
 130. Van Loo G, Schotte P, Van Gorp M, Demol H, Hoorelbeke B, Gevaert K, et al. Endonuclease G: A mitochondrial protein released in apoptosis and involved in caspase-independent DNA degradation. *Cell Death Differ.* 2001;8(12):1136–42.
 131. Büttner S, Eisenberg T, Carmona-Gutierrez D, Ruli D, Knauer H, Ruckenstuhl C, et al. Endonuclease G Regulates Budding Yeast Life and Death. *Mol Cell.* 2007;25(2):233–46.
 132. Kalinowska M, Garncarz W, Pietrowska M, Garrard WT, Widlak P. Regulation of the human apoptotic DNase/RNase endonuclease G: involvement of Hsp70 and ATP. *Apoptosis.* 2005;10(4):821–30.

133. Susin SA, Lorenzo HK, Zamzami N, Marzo I, Snow BE, Brothers GM, et al. Molecular characterization of mitochondrial apoptosis-inducing factor. *Nature*. 1999;397(6718):441–6.
134. Wang X, Yang C, Chai J, Shi Y, Xue D. Mechanisms of AIF-mediated apoptotic DNA degradation in *Caenorhabditis elegans*. *Science* (80-). 2002;298(5598):1587–92.
135. Parrish JZ, Xue D. Functional genomic analysis of apoptotic DNA degradation in *C. elegans*. *Mol Cell*. 2003;11(4):987–96.
136. Parrish JZ, Yang C, Shen B, Xue D. CRN-1, a *Caenorhabditis elegans* FEN-1 homologue, cooperates with CPS-6/EndoG to promote apoptotic DNA degradation. *EMBO J*. 2003;22(13):3451–60.
137. Zassenhaus HP, Denniger G. Analysis of the role of the NUC1 endo/exonuclease in yeast mitochondrial DNA recombination. *Curr Genet*. 1994;25(2):142–9.
138. Huang K-J, Zemelman B V, Lehman IR. Endonuclease G, a candidate human enzyme for the initiation of genomic inversion in herpes simplex type 1 virus. *J Biol Chem*. 2002;277(23):21071–9.
139. Cote J, Ruizcarrillo A, Côté J, Ruiz-Carrillo A. Primers for Mitochondrial DNA Replication Generated by Endonuclease-G. *Science* (80-). 1993;261(5122):765–9.
140. Ikeda S, Ozaki K. Action of mitochondrial endonuclease G on DNA damaged by L-ascorbic acid, peplomycin, and cis-diamminedichloroplatinum (II). *Biochem Biophys Res Commun*. 1997;235(2):291–4.
141. Gole B, Baumann C, Mian E, Ireno CI, Wiesmüller L. Endonuclease G initiates DNA rearrangements at the MLL breakpoint cluster upon replication stress. *Oncogene*. 2014;34(February):1–11.
142. Zan H, Zhang J, Al-Qahtani A, Pone EJ, White CA, Lee D, et al. Endonuclease G plays a role in immunoglobulin class switch DNA recombination by introducing double-strand breaks in switch regions. *Mol Immunol*. 2011;48(4):610–22.
143. Robertson AB, Robertson J, Fusser M, Klungland A. Endonuclease G preferentially cleaves 5-hydroxymethylcytosine-modified DNA creating a substrate for

- recombination. *Nucleic Acids Res.* 2014;42(21):1–14.
144. Kafer GR, Li X, Horii T, Suetake I, Tajima S, Hatada I, et al. 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes Genome Stability. *Cell Rep.* 2016;14(6):1283–92.
 145. Vander Zanden CM, Czarny RS, Ho EN, Robertson AB, Ho PS. Structural adaptation of vertebrate endonuclease G for 5-hydroxymethylcytosine recognition and function. *Nucleic Acids Res.* 2020;
 146. Houmiel KL, Gerschenson M, Low RL. Mitochondrial endonuclease activity in the rat varies markedly among tissues in relation to the rate of tissue metabolism. *Biochim Biophys Acta (BBA)/Protein Struct Mol.* 1991;1079(2):197–202.
 147. Kazak L, Reyes A, Holt IJ. Minimizing the damage: Repair pathways keep mitochondrial DNA intact. *Nat Rev Mol Cell Biol.* 2012;13(10):659–71.
 148. Szymanski MR, Yu W, Gmyrek AM, White MA, Molineux IJ, Lee JC, et al. A domain in human EXOG converts apoptotic endonuclease to DNA-repair exonuclease. *Nat Commun.* 2017;8:14959.
 149. Tann AW, Boldogh I, Meiss G, Qian W, Van Houten B, Mitra S, et al. Apoptosis induced by persistent single-strand breaks in mitochondrial genome: critical role of EXOG (5'-EXO/endonuclease) in their repair. *J Biol Chem.* 2011;286(37):31975–83.
 150. Wiehe RS, Gole B, Chatre L, Walther P, Calzia E, Ricchetti M, et al. Endonuclease G promotes mitochondrial genome cleavage and replication. *Oncotarget.* 2018;9(26):18309–26.
 151. Wu C-C, Lin JIJ, Yang-Yen H-F, Yuan HS. A unique exonuclease ExoG cleaves between RNA and DNA in mitochondrial DNA replication. *Nucleic Acids Res.* 2019;47(10):5405–19.
 152. David KK, Sasaki M, Yu S-W, Dawson TM, Dawson VL. EndoG is dispensable in embryogenesis and apoptosis. *Cell Death Differ.* 2006;13(7):1147–55.
 153. Irvine RA, Adachi N, Shibata DK, Cassell GD, Yu K, Karanjawala ZE, et al. Generation and Characterization of Endonuclease G Null Mice. *Mol Cell Biol.* 2005;25(1):294–302.

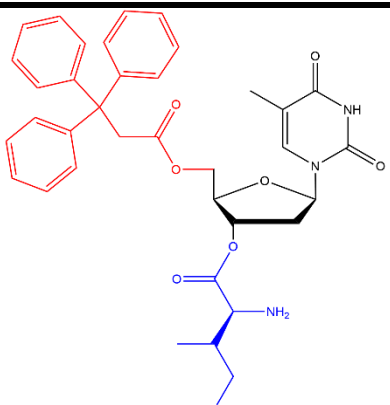
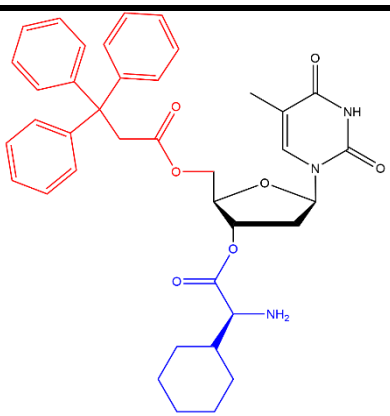
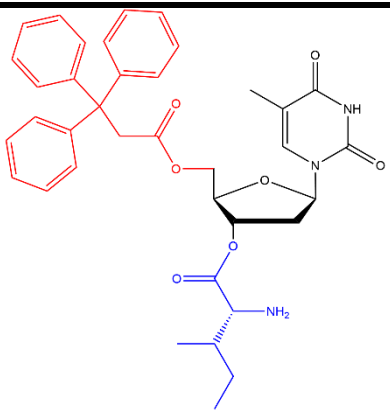
154. Zhang J, Dong M, Li L, Fan Y, Pathre P, Dong J, et al. Endonuclease G is required for early embryogenesis and normal apoptosis in mice. *Proc Natl Acad Sci U S A*. 2003;100(26):15782–7.
155. Kroemer G, Galluzzi L, Vandenabeele P, Abrams J, Alnemri ES, Baehrecke EH, et al. Classification of cell death: Recommendations of the Nomenclature Committee on Cell Death 2009. *Cell Death Differ*. 2009;16(1):3–11.
156. Casanova E, Moreno D, Gigante A, Rico E, Genes CM, Oliva C, et al. 5'-trityl-substituted thymidine derivatives as a novel class of antileishmanial agents: *Leishmania infantum* EndoG as a potential target. *ChemMedChem*. 2013;8(7):1161–74.
157. World Health Organization - Neglected Tropical Diseases [Internet]. Available from: https://www.who.int/neglected_diseases/diseases/en/
158. Røttingen JA, Regmi S, Eide M, Young AJ, Viergever RF, Ardal C, et al. Mapping of available health research and development data: what's there, what's missing, and what role is there for a global observatory? *Lancet*. 2013;382(9900):1286–307.
159. Antoniou M, Gramiccia M, Molina R, Dvorak V, Volf P. The role of indigenous phlebotomine sandflies and mammals in the spreading of leishmaniasis agents in the mediterranean region. *Eurosurveillance*. 2013;18(30).
160. Gilbert IH. Drug discovery for neglected diseases: Molecular target-based and phenotypic approaches. *J Med Chem*. 2013;56(20):7719–26.
161. Arora K, Rai AK. Dependence of *Leishmania* parasite on host derived ATP: an overview of extracellular nucleotide metabolism in parasite. *J Parasit Dis*. 2019;43(1):1.
162. Sunduru N, Agarwal A, Katiyar SB, Nishi, Goyal N, Gupta S, et al. Synthesis of 2,4,6-trisubstituted pyrimidine and triazine heterocycles as antileishmanial agents. *Bioorganic Med Chem*. 2006;14(23):7706–15.
163. Guo S, Wang J, Zhang X, Cojean S, Loiseau PM, Fan X. Synthesis of 5-isoxazol-3-yl-pyrimidine nucleosides as potential antileishmanial agents. *Bioorganic Med Chem Lett*. 2015;25(13):2617–20.
164. Coimbra ES, Nora de Souza MV, Terror MS, Pinheiro AC, da Trindade Granato J.

- Synthesis, biological activity, and mechanism of action of new 2-pyrimidinyl hydrazone and N-acylhydrazone derivatives, a potent and new classes of antileishmanial agents. *Eur J Med Chem.* 2019;184:111742.
165. Atta KFM, Ibrahim TM, Farahat OOM, Al-Shargabi TQ, Marei MG, Bekhit AA, et al. Synthesis, modeling and biological evaluation of hybrids from pyrazolo[1,5c]pyrimidine as antileishmanial agents. *Future Med Chem.* 2017;9(16):1913–29.
166. Fan X, Feng D, Qu Y, Zhang X, Wang J, Loiseau PM, et al. Practical and efficient synthesis of pyrano[3,2-c]pyridone, pyrano[4,3-b]pyran and their hybrids with nucleoside as potential antiviral and antileishmanial agents. *Bioorganic Med Chem Lett.* 2010;20(3):809–13.
167. Torrence PF, Fan X, Zhang X, Loiseau PM. Structurally diverse 5-substituted pyrimidine nucleosides as inhibitors of *Leishmania donovani* promastigotes in vitro. *Bioorganic Med Chem Lett.* 2006;16(19):5047–51.
168. Rico E, Oliva C, Gutierrez KJ, Alzate JF, Genes CM, Moreno D, et al. *Leishmania infantum* EndoG is an endo/exo-nuclease essential for parasite survival. *PLoS One.* 2014;9(2):e89526.
169. Büttner S, Carmona-Gutierrez D, Vitale I, Castedo M, Ruli D, Eisenberg T, et al. Depletion of endonuclease G selectively kills polyploid cells. *Cell Cycle.* 2007;6(9):1072–6.
170. McDermott-Roe C, Ye J, Ahmed R, Sun X-M, Serafín A, Ware J, et al. Endonuclease G is a novel determinant of cardiac hypertrophy and mitochondrial function. *Nature.* 2011;478(7367):114–8.
171. Oliva C, Sánchez-Murcia PA, Rico E, Bravo A, Menéndez M, Gago F, et al. Structure-based domain assignment in *Leishmania infantum* EndoG: Characterization of a pH-dependent regulatory switch and a C-Terminal extension that largely dictates DNA substrate preferences. *Nucleic Acids Res.* 2017;45(15):9030–45.
172. Ghosh M, Meiss G, Pingoud AM, London RE, Pedersen LC. The nuclease a-inhibitor complex is characterized by a novel metal ion bridge. *J Biol Chem.* 2007;282(8):5682–90.

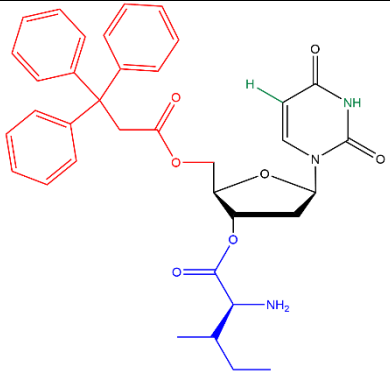
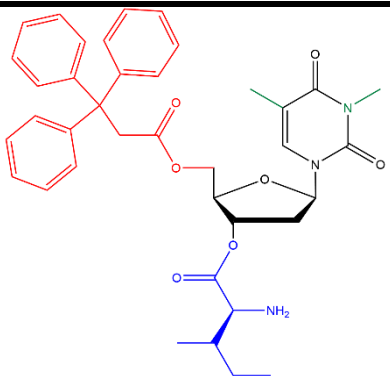
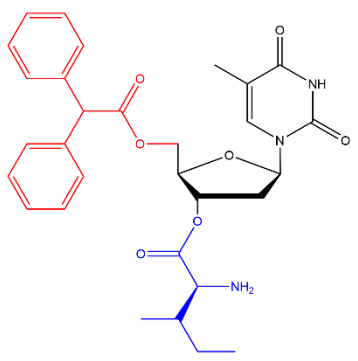
173. Llopis J, McCaffery JM, Miyawaki A, Farquhar MG, Tsien RY. Measurement of cytosolic, mitochondrial, and Golgi pH in single living cells with green fluorescent proteins. *Proc Natl Acad Sci U S A*. 1998;95(12):6803–8.
174. Santo-Domingo J, Demaurex N. The renaissance of mitochondrial pH. *J Gen Physiol*. 2012;139(6):415–23.
175. Ramsdale M. Programmed cell death in the cellular differentiation of microbial eukaryotes. *Curr Opin Microbiol*. 2012;15(6):646–52.
176. Gannavaram S, Debrabant A. Programmed cell death in *Leishmania*: biochemical evidence and role in parasite infectivity. *Front Cell Infect Microbiol*. 2012;2:95.
177. Ceska TA, Sayers JR. Structure-specific DNA cleavage by 5' nucleases. *Trends Biochem Sci*. 1998;23(9):331–6.
178. Balakrishnan L, Bambara RA. Flap Endonuclease 1. *Annu Rev Biochem*. 2013;82(1):119–38.

Anexo

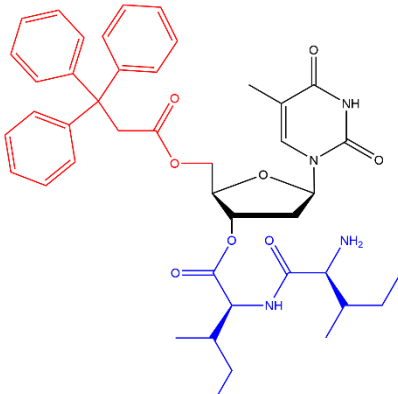
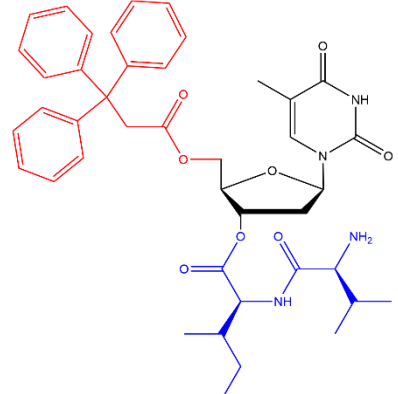
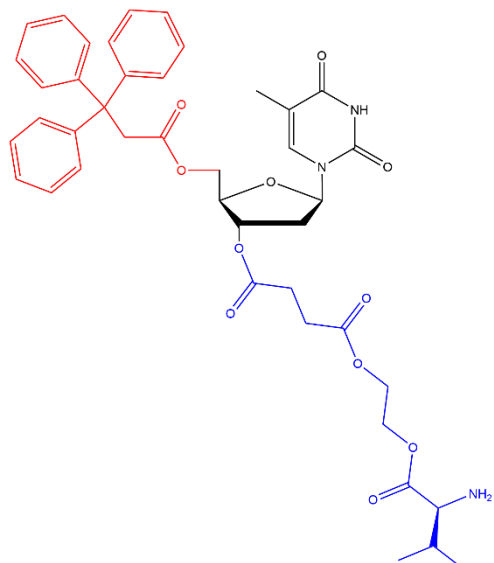
Compuesto	(μM)	Estructura molecular	
1	IC ₅₀ promastigotes	7.67 ± 0.21	
	IC ₅₀ amastigotes	>25	
	IC ₅₀ Jurkat	16.9 ± 0.5	
2	IC ₅₀ promastigotes	2.88 ± 0.13	
	IC ₅₀ amastigotes	10.63 ± 0.44	
	IC ₅₀ Jurkat	17.1 ± 2.2	
3	IC ₅₀ promastigotes	>25	
	IC ₅₀ amastigotes	ND	
	IC ₅₀ Jurkat	ND	
	Inhibición <i>Li</i> EndoG	Nula	
4	IC ₅₀ promastigotes	>25	
	IC ₅₀ amastigotes	ND	
	IC ₅₀ Jurkat	ND	
	Inhibición <i>Li</i> EndoG	Nula	

6	IC ₅₀ promastigotes	3.4 ± 0.6	
	IC ₅₀ amastigotes	7.33 ± 0.04	
	IC ₅₀ Jurkat	9.87 ± 1.8	
	Inhibición <i>rLi</i> EndoG	Media	
7	IC ₅₀ promastigotes	3.5 ± 0.6	
	IC ₅₀ amastigotes	6.33 ± 0.01	
	IC ₅₀ Jurkat	13.9 ± 0.7	
8	IC ₅₀ promastigotes	6.51 ± 0.03	
	IC ₅₀ amastigotes	22.60 ± 1.52	
	IC ₅₀ Jurkat	15.3 ± 0.4	

9	IC ₅₀ promastigotes	>25	
	IC ₅₀ amastigotes	ND	
	IC ₅₀ Jurkat	ND	
15	IC ₅₀ promastigotes	>25	
	IC ₅₀ amastigotes	ND	
	IC ₅₀ Jurkat	ND	
16	IC ₅₀ promastigotes	>25	
	IC ₅₀ amastigotes	ND	
	IC ₅₀ Jurkat	ND	

20	IC ₅₀ promastigotes	7.31 ± 3.22	
	IC ₅₀ amastigotes	12.76 ± 0.25	
	IC ₅₀ Jurkat	9.3 ± 0.3	
21	IC ₅₀ promastigotes	5.87 ± 0.75	
	IC ₅₀ amastigotes	9.68 ± 0.36	
	IC ₅₀ Jurkat	10.6 ± 0.9	
24	IC ₅₀ promastigotes	17.29 ± 1.64	
	IC ₅₀ amastigotes	>25	
	IC ₅₀ Jurkat	22.4 ± 1.0	

25	IC ₅₀ promastigotes	2.59 ± 0.24	
	IC ₅₀ amastigotes	4.49 ± 0.52	
	IC ₅₀ Jurkat	6.83 ± 0.1	
	Inhibición <i>rLi</i> EndoG	Elevada	
26	IC ₅₀ promastigotes	6.56 ± 0.05	
	IC ₅₀ amastigotes	7.07 ± 0.02	
	IC ₅₀ Jurkat	9.9 ± 0.3	
27	IC ₅₀ promastigotes	6.67 ± 0.23	
	IC ₅₀ amastigotes	22.33 ± 0.31	
	IC ₅₀ Jurkat	16.2 ± 1.1	

28	IC ₅₀ promastigotes	1.31 ± 0.09	
	IC ₅₀ amastigotes	11.75 ± 0.74	
	IC ₅₀ Jurkat	15.9 ± 0.6	
	Inhibición <i>Li</i> EndoG	Elevada	
29	IC ₅₀ promastigotes	5.99 ± 0.33	
	IC ₅₀ amastigotes	13.42 ± 0.14	
	IC ₅₀ Jurkat	8.28 ± 1.0	
32	IC ₅₀ promastigotes	6.38 ± 1.17	
	IC ₅₀ amastigotes	14.52 ± 1.40	
	IC ₅₀ Jurkat	24.9 ± 2.3	

33	IC ₅₀ promastigotes	3.78 ± 0.14	
	IC ₅₀ amastigotes	8.65 ± 0.51	
	IC ₅₀ Jurkat	20.9 ± 1.2	
34	IC ₅₀ promastigotes	2.22 ± 0.09	
	IC ₅₀ amastigotes	6.11 ± 0.01	
	IC ₅₀ Jurkat	7.0 ± 0.9	
37 o Lei49	IC ₅₀ promastigotes	4.33 ± 0.05	
	IC ₅₀ amastigotes	6.44 ± 0.01	
	IC ₅₀ Jurkat	17.2 ± 0.9	
	Inhibición <i>Li</i> EndoG	Elevada	

	IC ₅₀ promastigotes	4.51 ± 0.65	
38	IC ₅₀ amastigotes	7.37 ± 0.02	
	IC ₅₀ Jurkat	15.6 ± 1.5	
	IC ₅₀ promastigotes	46.45 ± 3.03	
Miltefosina	IC ₅₀ amastigotes	3.96 ± 0.53	-
	IC ₅₀ Jurkat	48	