

1. INTRODUCCIÓN

«Nada hay más monárquico que una molécula de ADN...La vida misma es el resultado del gobierno absoluto desoxirribonucleico.»

Salvador Dalí (1904-1989)

1. INTRODUCCIÓN

1.1. Las proteínas HERC

1.1.1. Definición

De acuerdo con el Comité de Nomenclatura Genética de la Organización del Genoma Humano (*Human Genome Organization (HUGO) Gene Nomenclature Committee (HGNC)*), son definidas como proteínas HERC todas aquellas proteínas en cuya estructura coexistan un dominio de homología al extremo carboxilo de la proteína E6AP (es decir un dominio HECT) y uno o más dominios homólogos a la proteína RCC1 (es decir dominios RLD). De la fusión de los términos HECT y RCC1 surge la palabra HERC.

1.1.2. El dominio RLD

Se conocen como dominios RLD (*RCC1-like domain*) aquellos dominios proteicos que presentan un elevado grado de similitud con la proteína RCC1 (*Regulator of Chromosome Condensation-1*). La secuencia de dicho dominio se caracteriza por presentar siete repeticiones de 51-68 residuos aminoacídicos cada una, dando lugar pues a dominios de entre aproximadamente 350 y 450 aa. La estructura del dominio ha sido resuelta por cristalografía de rayos X en el caso de RCC1 (Renault et al. (1998) y (2001) y Figura 1), lo cual ha permitido comprobar que las siete repeticiones descubiertas originalmente por análisis secuencial corresponden en realidad a cada uno de los siete lados de un dominio con simetría heptagonal formado por cadenas beta, que algunos han asemejado a una hélice de siete aspas (*seven-bladed β -propeller*). Hasta la fecha se han descrito en la literatura unas dieciocho proteínas con dominios RLD (aunque existen muchas más en las bases de datos (Bateman et al. (2002))). De ellas, RCC1 fue la primera en identificarse y es también aquélla cuya función (además de su estructura) se

conoce en mayor detalle. Concretamente, RCC1 es capaz de unirse a la cromatina a través de su interacción con las histonas H2A y H2B (Nemergut et al. (2001)), lo cual

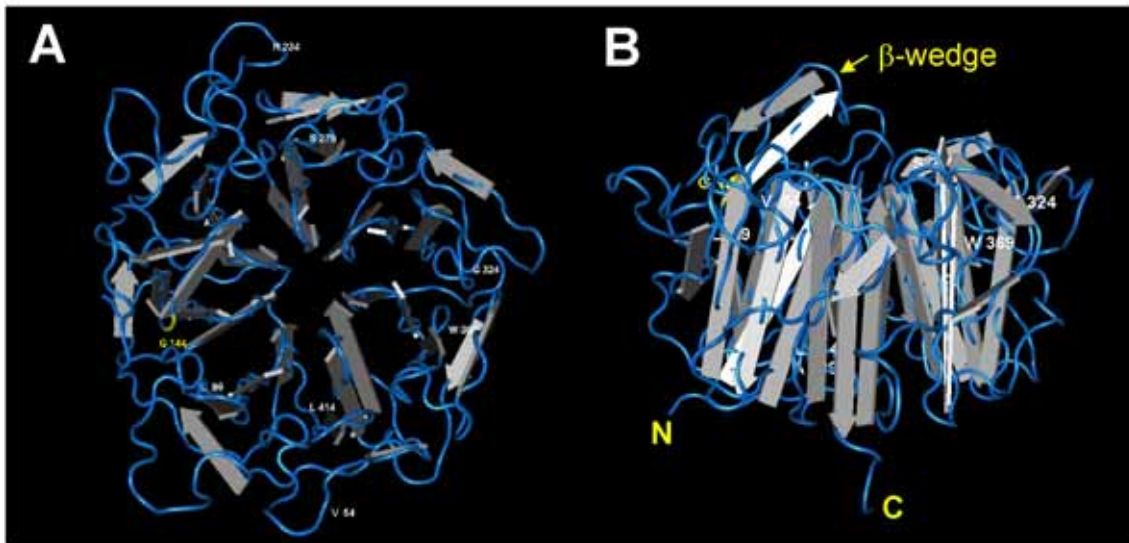


Figura 1. Estructura del dominio RLD de RCC1. (A) Vista de RCC1 por la zona de unión a Ran. Desde esta perspectiva es posible apreciar la simetría heptagonal del β -propeller o propulsor- β (B) Vista lateral de RCC1. La zona de unión a la cromatina se halla abajo en la figura, donde se encuentran también los extremos amino y carboxilo terminal de la proteína, mientras que la unión a Ran se produce por la zona situada arriba en la figura. Para la actividad GEF de RCC1 sobre Ran es esencial la cuña- β (β -wedge), formada por dos cadenas β antiparalelas que protruyen por encima del propulsor- β y contribuyen a desplazar el GDP del centro activo de Ran.

además estimula su capacidad de actuar como un factor intercambiador de nucleótidos de guanina (GEF: *guanine nucleotide exchange factor*) sobre Ran, una proteína G monomérica de la superfamilia de Ras (Figura 2). De las dos caras del propulsor- β (β -propeller) en que se pliega RCC1, una de ellas, la que posee los extremos N y C-terminal de la proteína, es la que interacciona con las susodichas histonas, mientras que la otra es la encargada de interactuar con Ran y catalizar el intercambio de GDP por GTP en el centro activo de ésta (Nemergut et al. (2001) y Renault et al. (2001)). Mediante esta actividad, RCC1 regula varios procesos importantes para las células. En primer lugar, el transporte vectorial entre el núcleo y el citoplasma de la célula depende de la existencia de un gradiente de RanGTP entre ambos compartimentos. Dicho gradiente se genera como consecuencia de la distinta localización subcelular de las proteínas activadora (RCC1, nuclear) e inhibidora (RanGAP1, citoplásmica) de Ran, así como de la permeabilidad de la envoltura nuclear a esta GTPasa. La forma activa de

Ran, es decir RanGTP, es capaz de regular la afinidad de importinas y exportinas (las proteínas encargadas del transporte a través de los poros nucleares) por aquéllas

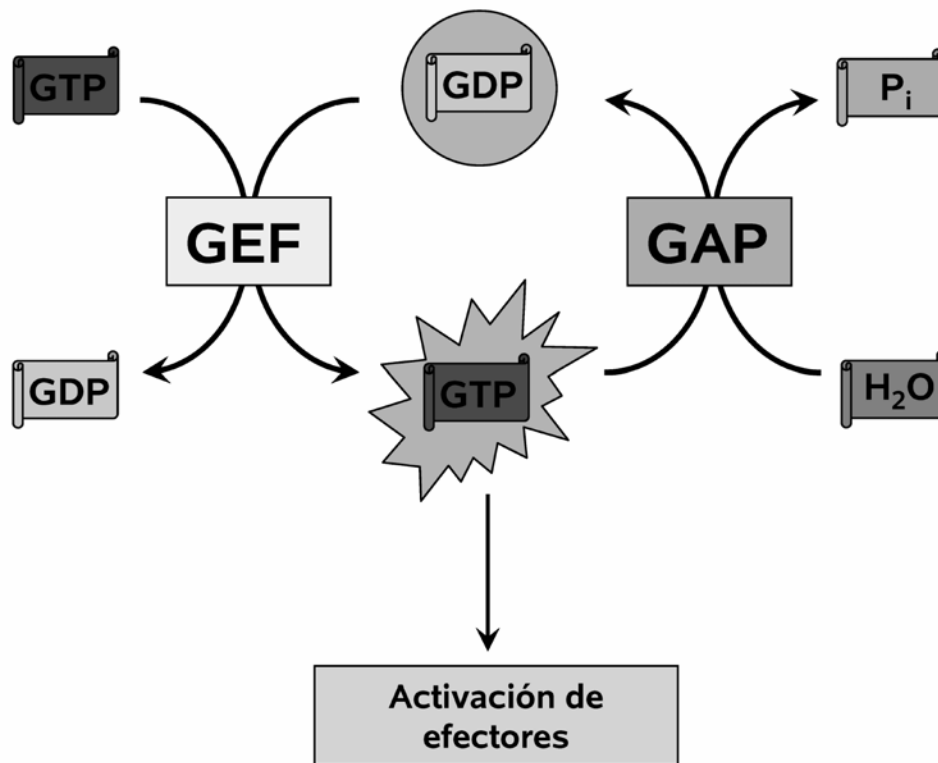


Figura 2. Regulación de la actividad de las proteínas G monoméricas de la superfamilia de Ras. Estas proteínas actúan en las células a modo de interruptor, oscilando entre una forma inactiva (o apagada) en la que se encuentran unidas a GDP y una forma activa (o encendida) en la que se hallan unidas a GTP y son capaces de unirse y regular la actividad de sus efectores. Sin embargo, tanto el proceso de activación como el de inactivación son extremadamente lentos a menos que estén catalizados enzimáticamente por dos tipos de enzimas llamados GEFs y GAPs. Así, los factores intercambiadores de nucleótidos (GEFs: *guanine nucleotide exchange factors*) aceleran la activación de la proteína G al estimular la disociación de GDP de su centro activo y la subsiguiente unión a éste de una molécula de GTP. Por su parte, las proteínas estimuladoras de la actividad GTPasa (GAPs: *GTPase activating proteins*), como su nombre indica, aumentan la velocidad a la que el GTP es hidrolizado a GDP y fosfato en el centro activo de la proteína G, con lo cual participan en la inactivación de la misma.

proteínas a las que deben transportar (el “*cargo*”, que aquí traduciré como la mercancía). Así, dentro del núcleo RanGTP hace que las exportinas se unan a su mercancía y que las importinas se deshagan de ella. Por contra, en el citoplasma ocurre a la inversa: la ausencia de RanGTP hace que las exportinas “descarguen” su mercancía y que las importinas se unan a ella. Además del transporte nucleocitoplásmico, RCC1, a través de Ran, también regula la formación del huso mitótico. El papel de RCC1 en este proceso es análogo al ya descrito para el transporte nucleocitoplásmico: en este caso el gradiente de RanGTP se forma entre los cromosomas metafásicos a los que se halla

asociado RCC1 y los polos del huso, con el que interacciona RanGAP1. Sin embargo, contrariamente a lo que sucede en células interfásicas, durante la mitosis RCC1 debe ser

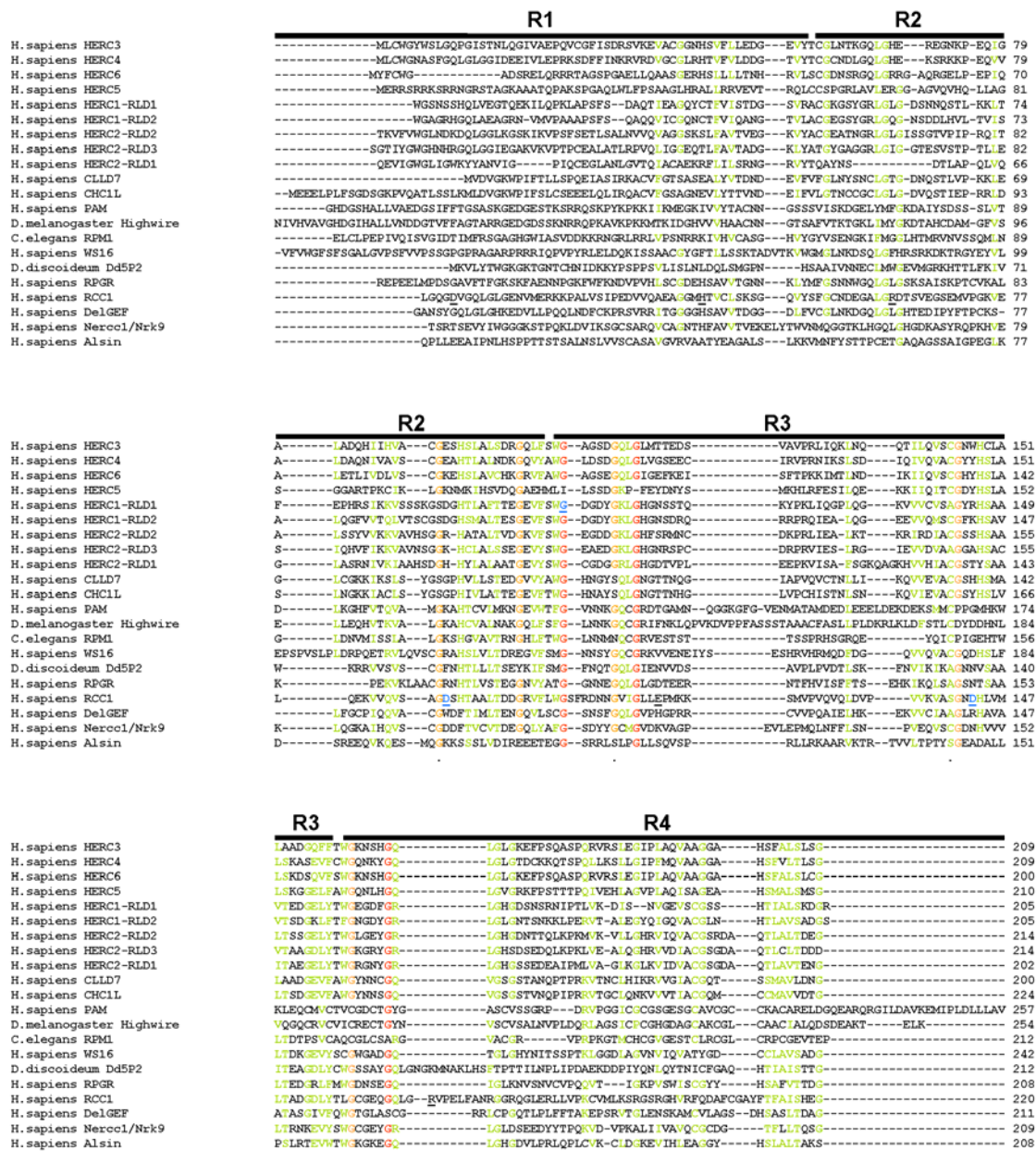


Figura 3 (esta página y la siguiente). Alineamiento múltiple de los RLDs de varias de las proteínas descritas en el texto. Los residuos conservados en la mayor parte de RLDs se han indicado, en orden creciente de conservación, en color verde, naranja o rojo. Además, el residuo de glicina de HERC1 mutado a glutámico en los ratones tambaleante ha sido subrayado y marcado en azul. También en azul y subrayados se encuentran los tres residuos de RCC1 cuya mutación a alanina abolece por completo la actividad GEF de esta proteína. Otros residuos cuya mutación disminuye, aunque no anula, la actividad GEF de RCC1 han sido simplemente subrayados (Azuma et al. (1999)).

fosforilado por Cdk1 en sus serinas 2 y 11 para de este modo evitar que las importinas- α y β se unan a la señal de localización nuclear (NLS) de RCC1 e impidan así la unión de éste a los cromosomas metafásicos y la formación del gradiente de RanGTP (Li & Zheng (2004)). En cuanto al papel de RanGTP en este proceso, se ha postulado que

consistiría en su capacidad de estabilizar la unión entre los microtúbulos del huso y los cromosomas. Por último, la activación de Ran por RCC1 también estimula el

	R4	R5	
H.sapiens HERC3	AVFGKGM	NNAGQLGLS	DEKD---RESCHVKLLRTRKVVY 246
H.sapiens HERC4	AIFGKGR	NKFGQLGLN	DEND---RYVENLIRKSLRSQKIVY 246
H.sapiens HERC6	TSEFGKS	NSAGQLALS	GRNVFQSNKPLSGALKNLGVVY 240
H.sapiens HERC5	NIYSGKK	NECGQLGLG	HTES---KDDPSLIEGLDNQKVEF 242
H.sapiens HERC1-RLD1	TWVSEGG	GDNGQLGHG	DTNR---VYKPKVLEALQKQKIRK 247
H.sapiens HERC1-RLD2	MWVAFCD	GDYGLGLG	NSTA---KSSPQKDVLCGIGIRK 242
H.sapiens HERC2-RLD2	LVEFKGD	GDYGLGLRG	GSEG---CNIPQNIEALNGQGVQ 251
H.sapiens HERC2-RLD3	TWVSKGD	GDYGLGLRG	GSDG---CKVFMKIDSLTGLGVK 251
H.sapiens HERC2-RLD1	QVWFKGD	GDYGLGLRG	GSDG---CKTKLIEKLQDLVVK 239
H.sapiens CLLD7	EYVWQY	NNGQLGLG	NNGN---QLTQVRAALHSVQVQ 237
H.sapiens CHC11	EYVWQY	NNGQLGLG	NSGN---QPTQCRVAALGQIRVQR 261
H.sapiens PAM	PVPGVNIIEHQLRQEEKRQVIRRHLEEGRPVLFAGPIEMHREQALARLRSHPAQLKHKRDKHDKSSGEREKDASKITTYPGSFRDCELRAVQ 357		
D.melanogaster Highwire	SDVQQRQQRSKTLMRRK	ERKSELETG	AAGGGAATPTDLD
C.elegans RRM1	AQPGR	AQHLVFSST	AAPQRSTLHPSRVLSQGHVVKVSS 252
H.sapiens WS16	GLEFGKN	SEYLLASVT	DSTQ---VNVFRCIHFSGVGRVQ 280
D.discoideum Dd5P2	AIFSGNS	NQYQQLGLGHCN	AKLVFVPEIHLDNSFERRAGHINSLGEIVTTDETAKTIRHIVVD 274
H.sapiens RPGR	ELVVEGE	PENGLGLP	NQLLGNHRTFQVSEIP-EKV 246
H.sapiens RCC1	HVYGFGL	SNYHGLTPT	GTES---CFIDQNLTSFKNSKTSW 257
H.sapiens DelGEF	EYVWNS	NKHQLANEA	APLVEFQK EAHCFQNEKV 247
H.sapiens Neroccl/Nrk9	KVLACGL	NEFNQLGNQMSG	IINHEAYHEVYVTSFELAKQLSEYK 256
H.sapiens Alsin	QVYSGNS	NTEGQLGHSDFP	TTVFLRAKISSENGVWS 244

	R5	R6	R7
H.sapiens HERC3	SQEEHTAVLTKSGVFTFGAGSQG	LGHDSMDEVNRRVLE	LMGSEVTQIACGRQHTLAEVPSGLIYAFGGGARQGL 327
H.sapiens HERC4	CGEDHTAALTKSGVFTFGAGGYGQ	LGHNSTSHEINRRKFE	LMGSLVTEIACGRQHTSAFVPSGRIVSEGLGGRQGL 327
H.sapiens HERC6	CGDHTAVLTKSGVFTFGDNRSGQ	LGYSPTEPKRGPQLVE	RIDGLVSQLDCGSYHTLAVYHTTQGVVSGHG 315
H.sapiens HERC5	ACGSHSALITQDELFTFGAGKHGQ	LGHSTQNELRCLVAE	LVGVRVQIACGRVHTLAVYVSDLGKVESGSGRQGL 328
H.sapiens HERC1-RLD1	CAGSQSIALITSTQVYWGCGACLG	CSSEAT-ALRKLIEE	LAATRIVDVSIGDSSICLALSHDN-EVYAWGNMSSQ 321
H.sapiens HERC1-RLD2	AGTQFSVALTKDQVYTFGQDRLIG	IEEERARNHNRQQFV	LAVIIEVAVGAEITLALASNG-DVYAWGNMSSQ 322
H.sapiens HERC2-RLD2	EQAQFSIALTKSIVYTWKGDYFR	LGHSDVHVRRQVVEG	LRGKIVHVAAGLACIADVDSG-QVYAWGNMSSQ 331
H.sapiens HERC2-RLD3	EQSQFSIALTKSIVYTWKGDYFR	LGHSDDHVRRQVVEG	LQGRKIVIAIATGSLHVCCTEDG-EVYAWGNMSSQ 331
H.sapiens HERC2-RLD1	RCQSFSIALTKDQVYTWKGDYFR	LGHTEEHVRYKLEGG	LQGRKIVIAIATGSLHVCCTEDG-EVYAWGNMSSQ 319
H.sapiens CLLD7	VGYAHTLALITDEQLYANGANTYQ	LGTGRNRLLSAHMV	EKERVIEIAACHSHTSAAKTQGGHVIWAG-QCRSQ 316
H.sapiens CHC11	AGYAHTLALITDEQLYANGANTYQ	LGTGRNRLLSAHMV	EKERVIEIAACHSHTSAAKTQGGHVIWAG-QCRSQ 340
H.sapiens PAM	SGPHHSVIMENQVYTFGQDRIG	LGHQVNSRGCITLVA	LRGP-STQTAGSNHTAVLMDG-QVYTFGSGKGL 436
D.melanogaster Highwire	AGLHHTVTLAEVYTFGSNQYQ	LGSIDLQVSVRQV	VEGA-ISQVAGSNHTVTLTSGM-SVYTFGNYKQGL 399
C.elegans RRM1	SCNFTVILASDRRVTFGSNCHGQ	LGVDTLSKNTQQVIL	PSDTVIVQVAGSNHTLBRANDG-SVYTFGAFKQGL 332
H.sapiens WS16	AGGTGAVNGEHVYTWGILRGG	PNIVESAVPEMIPITL	GLTEFNIEIQVSRRCGLSHFAALNKG-ELVWGNKIRVCI 366
D.discoideum Dd5P2	ACGNFHSALITLEDLFLMGSNQYQ	LQDSDLDQSVKTKG	SLEPVKINRAQGNHTLITLTKEQ-DVYTFGSGNYQGL 355
H.sapiens RPGR	AGGHTVLTENAVYTFGLQYFQ	LGLTFLFETSEPKVIE	IRDQTISYISCGENHTALITDID-LMVTGQGRHK 326
H.sapiens RCC1	FGGQHTVQDSEKAYSLGRAEYGR	LGLEGAEEKSIPTLIS	RLPAVSSVACGASVYAVTRDG-RVYAWGNMSSQ 338
H.sapiens DelGEF	IMSGHTHLVAQTEKQVWTRDYGGLGRKLETYEGMKLEKQDSFLPCSRPNSMPSHPHCLTGATEVSCGSENLAIIGGV-CYSHWGNHMC 344		
H.sapiens Neroccl/Nrk9	IRTIAPKRHTAAIDERRLTFGCNKCSQ	LGVNYKRLGINLIGG	PLGGQVIRVSCGDEFTIATDDN-HIFAWGNMSSQ 340
H.sapiens Alsin	AAGRDYSLFVDTEDQPLGYSSGRD	PTEGDNLPENHSGSKTFLVLLSCKLGYISRVTAQRDYSILAVDRKNIMGVIAISHELATT 331	

	R7
H.sapiens HERC3	GTGHTC-----NVKCSIPVKGYYAAHSGQLSARADRFRYHIVKQIFSGDQTFVLCRK 380
H.sapiens HERC4	GTGSTS-----NRKSFTRKGNVYYPNGQCLFDIDSEYFVCKRIFSGDQSFVSHYSS 380
H.sapiens HERC6	PSDTS-----KPTHFEALTENFDIS---CLISAEDFVDVQVHIFAGTYANFVTHQ 364
H.sapiens HERC5	QNGTR-----DQIMLPKVSSEEE-----LKLSEHTSEKELI 362
H.sapiens HERC1-RLD1	QQNSTG-----PITKPKVSGLDGIAIQSASTHLSAWTA 359
H.sapiens HERC1-RLD2	GLHTN-----HVRETLVTLGQGRVQRITSAIRCHSAWTA 359
H.sapiens HERC2-RLD2	GNSTT-----VNRKTLVQGLEGKIRVACSSHSVAWTTVDVAT 373
H.sapiens HERC2-RLD3	GDSTN-----AIQRRLVAALQGRKRVNACSAHTLAWSTSKPAS 373
H.sapiens HERC2-RLD1	FDTLRV-----TKPEAALPGLDTHIVGIACPAQSEAFWSCSSEMSIGLRVFP 368
H.sapiens CLLD7	-----SVILHLHLSCTDDVFACFATPAVWRLLS 347
H.sapiens CHC11	-----SVILHLHLSCTDDVFACFATPAVWRLLS 371
H.sapiens PAM	GRPILDVPYNAKPAFNIKSGYGRKATWIGASDQTFRLIDEALINSHVLTASEIFASKHIG 501
D.melanogaster Highwire	GR-----LP---SDYGLKPPQD-----DDSPVG 420
C.elegans RRM1	ARPAGEKAGNNAIPEKVSFGGPFNAFAGWIGADSSSIHSHALLSS 381
H.sapiens WS16	GIRLED-----QYFWRVMPGEP 386
D.discoideum Dd5P2	CHSIGGRN-----RLYFTQIQKPKVMASDIFAAVDCSAYFVLDSDSPV 399
H.sapiens RPGR	GLLEN-----FTNHFIPLTCSNFRFIVKLVACGCHMVFVAAPHRGV 370
H.sapiens RCC1	GTQDED-----ANSPVEMGRQLENRWLVSSGQHTVLLVVD 378
H.sapiens DelGEF	GDTEAN-----WAKPVQALLSSGSLLVGCGAGHSIALCQL 382
H.sapiens Neroccl/Nrk9	AMPTFERPHGSDICTSWRPIFGLSHPVLDLSCRWHTILIVERVLNS 388
H.sapiens Alsin	ERRFYS-----RLSDTKSQTLRPLLSLENLSTT 360

ensamblaje postmitótico de la envoltura nuclear mediante un mecanismo que parece basarse en el reclutamiento por parte de RanGTP de la importina-β, la cual a su vez reclutaría a las nucleoporinas en lo que constituirían los primeros pasos de la génesis de una nueva envoltura nuclear (para una revisión reciente del complejo funcionamiento de los procesos regulados por RCC1 y Ran referirse a Quimby & Dasso (2003)).

La capacidad de RCC1 para actuar como GEF de Ran plantea la cuestión de hasta qué punto los dominios RLD existentes en otras proteínas pueden operar de forma similar al RLD de RCC1. Como ya se ha mencionado más arriba, sólo dieciocho (incluyendo RCC1) de las varias docenas de proteínas poseedoras de RLDs existentes en las bases de datos han sido estudiadas hasta la fecha, y de éstas sólo en algunos casos se han hallado funciones para los dominios RLD (Figuras 3 y 4). A continuación se resumen los conocimientos disponibles acerca de estas diecisiete proteínas y sus RLDs: seis de estas diecisiete son los miembros de la familia HERC humana (ver sección 1.4). De éstos, sólo en el caso de HERC1 se han descrito posibles funciones para sus dos dominios RLD. Así, mientras que el dominio RLD2 de HERC1 posee afinidad por ARF1, una proteína G monomérica de la superfamilia de Ras involucrada en tráfico intracelular, y por la cadena pesada de la clatrina, otra proteína clave en el transporte vesicular (Rosa et al. (1996) y Rosa & Barbacid (1997)), el dominio RLD1 de HERC1 fue descrito inicialmente como poseedor de actividad GEF sobre ARF1 y otras GTPasas implicadas en tráfico como son Rab3a y Rab5 (Rosa et al. (1996). Sin embargo, análisis más detallados han mostrado recientemente que éste no es exactamente el caso, sino que el RLD1 de HERC1, más que como un GEF para dichas GTPasas, actúa sobre ellas como un factor disociador de nucleótidos de guanina (GRF: *guanine nucleotide release factor*), para lo cual es además estrictamente necesario que RLD1 esté unido al fosfolípido fosfatidilinositol-4,5-bisfosfato (Garcia-Gonzalo et al. (2005)). De entre las once proteínas restantes con RLDs, otro caso interesante es el de Claret, una proteína de *Drosophila* involucrada en la biogénesis de los gránulos pigmentarios del ojo de este díptero. Aunque se ha postulado que Claret pueda ser un GEF para Lightoid (una GTPasa de la familia Rab de *Drosophila*), lo cierto es que lo único que ha sido demostrado es que el dominio RLD de Claret se une a Lightoid (y también a sus

homólogos humanos Rab32 y Rab38) y lo hace con mayor afinidad por la forma libre de nucleótido de esta GTPasa (Ma et al. (2004)). Otra “aspirante” a GEF es la proteína PRAF-1 (*PH, RLD and FYVE-1*) de *Arabidopsis thaliana*. En esta ocasión, ha sido demostrado que el dominio RLD de PRAF-1 es capaz de disociar GDP del centro activo

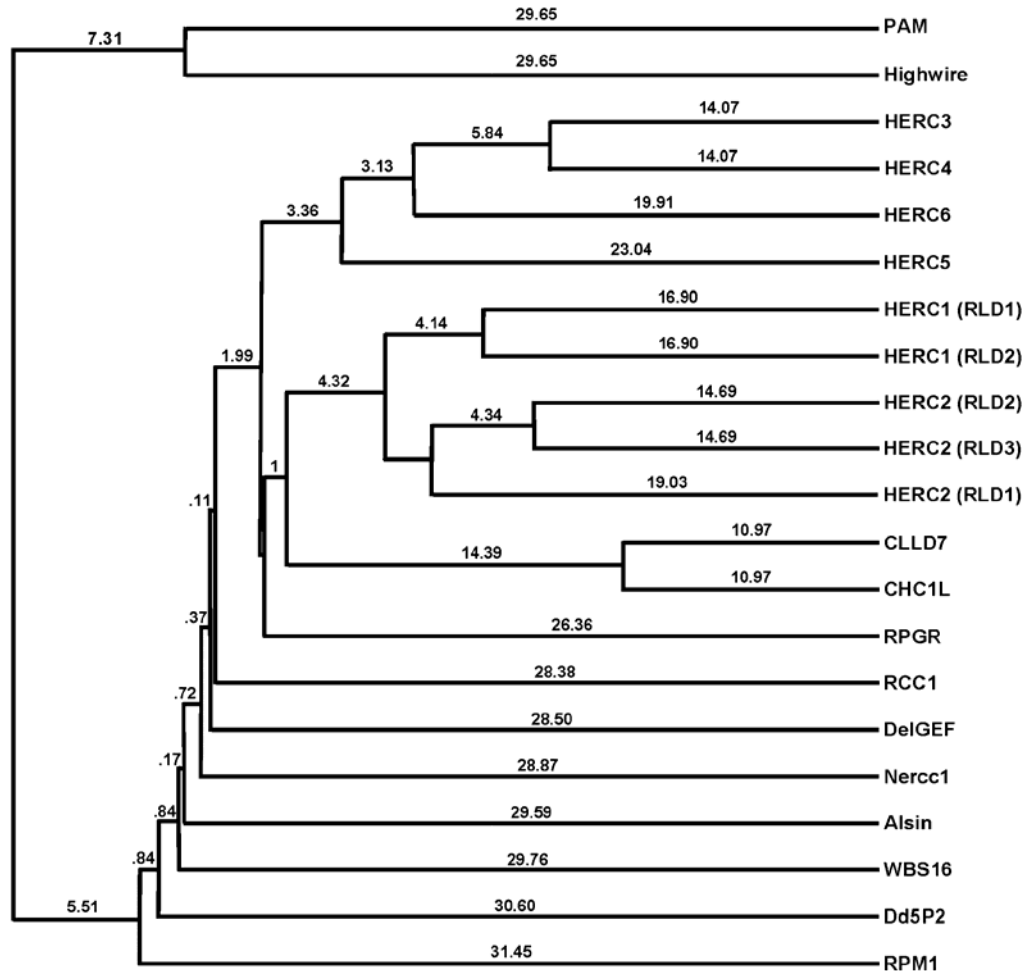


Figura 4. Árbol filogenético basado en el alineamiento múltiple mostrado en la Figura 3. Se indican las distancias evolutivas relativas. Como puede verse, los RLDs de PAM y su homólogo en *Drosophila*, Highwire, son los que más se diferencian de todos los demás. Por contra, los RLDs más parecidos entre sí son los de las proteínas CLLD7 y CHC1L, los cuales, a su vez, son los más cercanamente emparentados con los RLDs de las HERCs grandes. De hecho, la similitud entre los RLDs de RPGR, CHC1L, CLLD7, HERC1 y HERC2 es incluso mayor a la similitud existente entre los RLDs de las HERCs grandes y los de las pequeñas, lo cual subraya el hecho de que HERCs grandes y pequeñas, pese a ser consideradas dentro de una misma familia por motivos estructurales, constituyen en realidad dos familias distintas genéticamente hablando.

de Rab8 (y en menor medida Rab11) pero no de otros miembros de las familias Rab, ARF o Ran (Jensen et al. (2001)). Aunque parece probable que PRAF-1 sea en efecto un GEF para Rab8, sería un error considerarlo como tal a menos que sea probado que PRAF-1 estimula el *intercambio* de nucleótidos sobre Rab8 y no sólo su *disociación* (la omisión de esta distinción ya llevó, como hemos visto, a la errónea conclusión de que

HERC1 era un GEF para ARF1 y Rabs). Un caso curioso lo constituye la proteína Alsln, el producto del gen ALS2 implicado en la etiopatogénesis de la esclerosis lateral amiotrófica: aunque ha sido probado que Alsln actúa como GEF de Rab5 y de Rac1, dichas actividades no requieren de la presencia de su dominio RLD, del cual se desconoce la función, sino de otros dominios situados en otros lugares de la misma proteína (Topp et al. (2004)). Es particularmente interesante el caso de la proteína Nercc1, también llamada Nrck9 (Roig et al. (2002)). Se trata de una quinasa homóloga a la proteína codificada por el gen *nimA* del hongo *Aspergillus nidulans* (*NIMA-like kinase*) cuya disrupción da lugar a alteraciones en la formación del huso mitótico y en el agarre de los cromosomas a dicho huso que se asemejan considerablemente al fenotipo ocasionado por la delección de RCC1 o Ran. Interesantemente, Nercc1 tiene afinidad por la forma inactiva de Ran, a la que se une tanto a través de su dominio quinasa como de su dominio RLD. Nercc1 forma homodímeros que durante la interfase se hallan exclusivamente en el citoplasma, donde permanecen inactivados por la presencia de un factor todavía por determinar pero que se ha postulado que podría ser el propio RanGDP. Según este modelo, durante la mitosis la desaparición de la envoltura nuclear permitiría a Nercc1 acercarse a los cromosomas donde, dados los muy bajos niveles de RanGDP en las cercanías de éstos (por el gradiente mencionado anteriormente), Nercc1 se activaría y podría ejercer sus efectos, probablemente mediante la fosforilación de alguna proteína/s reguladora de la adhesión entre los cromosomas y los microtúbulos del huso. En cualquier caso, parece claro que Nercc1 no posee actividad GEF sobre Ran, pues varios de los residuos necesarios para ello en RCC1 no se hallan conservados en el RLD de Nercc1 (Azuma et al. (1999)). Todas las proteínas con RLDs mencionadas hasta ahora (RCC1, HERC1, Claret, PRAF-1, Alsln y Nercc1) tienen en común su capacidad para asociarse con proteínas G monoméricas (no necesariamente para

activarlas y en el caso de Alsin ni siquiera, por lo que se sabe, a través del RLD). No obstante, esta característica tampoco parece ser demasiado generalizable a las proteínas con RLDs, como veremos en seguida. Otras proteínas para cuyos RLDs se han descrito funciones incluyen PAM, RPGR y DelGEF. Por lo que respecta a PAM (*Protein associated with Myc*), se trata de una proteína gigante que juega un papel importante en el sistema nervioso en procesos tales como la sinaptogénesis o la nocicepción (Ehnert et al. (2004)). A nivel molecular, las funciones de PAM parecen explicarse en base a por lo menos tres acciones distintas: (1) inhibición del enzima adenilato ciclasa (Pierre et al. (2004)) (2) interacción con el factor de transcripción Myc (Guo et al. (1998)) y (3) unión al complejo TSC1/TSC2 (Murthy et al. (2004)). De estas tres actividades, sólo la primera es llevada a cabo por uno de los dos RLDs de la proteína. Así, el dominio RLD2 de PAM se une al dominio C2 de la adenilato ciclasa tipo V (AC-V) e inhibe la actividad de la misma (Scholich et al. (2001)). Es más, en un estudio reciente, Gao & Patel (2005) han sido capaces de separar unión e inhibición al lograr un mutante del RLD2 de PAM que se une pero no inhibe a AC-V. En dicho mutante, dos residuos de histidina consecutivos del RLD2 (His-912 y His-913) han sido sustituidos por sendas alaninas. Después de RCC1 y HERC1, la siguiente proteína con RLDs en ser identificada fue RPGR (*Retinitis pigmentosa GTPase regulator*). Esta proteína, cuya disfunción ocasiona la variante ligada al sexo de la retinitis pigmentosa, RP3 (Meindl et al. (1996)), posee un dominio RLD en su mitad amino terminal, mientras que el resto de la misma no contiene ningún otro dominio conservado. De hecho, todas las mutaciones causantes de RP3 identificadas en RPGR se encuentran en su RLD y todas ellas afectan a la capacidad de RPGR de interactuar con dos proteínas: RPGRIP y la subunidad δ de la fosfodiesterasa de cGMP. Mientras que la primera es un regulador clave de la morfogénesis de los discos fotosensibles en los fotorreceptores retinianos y parece ser el

principal efector de RPGR (Zhao et al. (2003)), la segunda podría estar regulando la unión/disociación del grupo isoprenilo de RPGR a la membrana (Linari et al. (1999) y Yan et al. (1998)). En cuanto a DelGEF (*Deafness locus GEF*), tratase de una proteína soluble de expresión ubicua que fue inicialmente identificada como el producto de un gen potencialmente implicado en ciertos casos de sordera congénita (Uhlmann et al. (1999)), aunque poco después se comprobó que los genes responsables de dichas alteraciones eran otros. La proteína DelGEF consiste casi exclusivamente en su dominio RLD a través del cual se une a otras dos proteínas: el homólogo humano de Sec5 (de forma dependiente de cationes Mg^{2+} y estimulada por GTP y dCTP) y DelGIP1, una proteína muy conservada de sólo 9 kDa. El papel de DelGEF parece que podría estar relacionado con la exocitosis, dado que (1) Sec5 es un componente del exocisto, un complejo molecular implicado en la exocitosis de vesículas y (2) reducciones en los niveles de tanto DelGEF como DelGIP1 provocan un aumento en la tasa de secreción de proteoglicanos al medio extracelular (Sjolinder et al. (2002) y (2004)). Existen aún algunas otras proteínas descritas con RLDs, aunque las funciones de éstos se ignoran por completo. Estas incluyen CLLD7 y CHC1L/RLG, dos proteínas muy homólogas cuyos genes se hallan muy cercanos en el cromosoma 13q14 y que son candidatos a genes supresores de tumores en leucemia linfocítica crónica de células B y cáncer de próstata, respectivamente (Mabuchi et al. (2001) y Latil et al. (2003)), otra proteína, WBS16, cuyo gen se halla en la región crítica del síndrome de Williams-Beuren (Merla et al. (2002)) y una de las cuatro inositol-5-fosfatasas funcionalmente redundantes que se encuentran en *Dictyostelium discoideum*.

En resumen, los dominios RLD parecen extremadamente versátiles y difíciles de encasillar en una función conservada, aunque en general parece que tienen facilidad para interactuar con otras proteínas (y a veces incluso lípidos), especialmente

proteínas G monoméricas a las que, en casos más bien excepcionales, pueden incluso activar. Por lo tanto, la conservación en el caso de los RLDs es más estructural que funcional, mientras que, como veremos ahora, en el caso de los dominios HECT es de ambos tipos.

1.1.3. El dominio HECT

Los dominios HECT (*Homologous to E6AP COOH Terminus*) son aquéllos que presentan homología con la región carboxilo terminal de la proteína E6AP. Esta proteína se identificó originalmente por su capacidad de asociarse a la oncoproteína vírica E6 de los subtipos oncogénicos del virus del papiloma humano. Más adelante se averiguó que los tumores inducidos por estos virus son debidos en buena parte al hecho que E6 interactúa dentro de la célula con la proteína supresora de tumores p53 al mismo tiempo que lo hace con E6AP. La formación del complejo ternario p53-E6-E6AP permite que p53 sea ubiquitinada por E6AP, lo cual induce la degradación de aquélla en el proteasoma y facilita el desarrollo tumoral (Scheffner et al. (1993)). Estudios posteriores han demostrado que la actividad ubiquitina ligasa de E6AP no se restringe a contextos patológicos como el mencionado, sino que la ubiquitinación de proteínas es su función habitual en el interior de las células, con la sola diferencia que en condiciones normales son otras proteínas y no p53 las que son ubiquitinadas por E6AP. Es más, no sólo E6AP, sino, por lo que se sabe hasta el momento, todas las proteínas poseedoras de un dominio HECT son capaces de actuar como ubiquitina ligasas y de este modo regular el funcionamiento de aquellas proteínas a las que ubiquitan, es decir sus sustratos. El proceso mediante el cual se efectúa la ubiquitinación de proteínas es complejo y requiere de la acción de como mínimo tres tipos de enzimas distintos: el enzima activador de la ubiquitina, también llamado E1, los enzimas conjugadores de

ubiquitina o E2 y finalmente las ubiquitina ligasas o E3 (para una explicación más detallada del sistema ubiquitina-proteasoma referirse a Pickart (2004)). Las proteínas que poseen dominios HECT constituyen pues un subtipo de E3/ubiquitina ligasas,

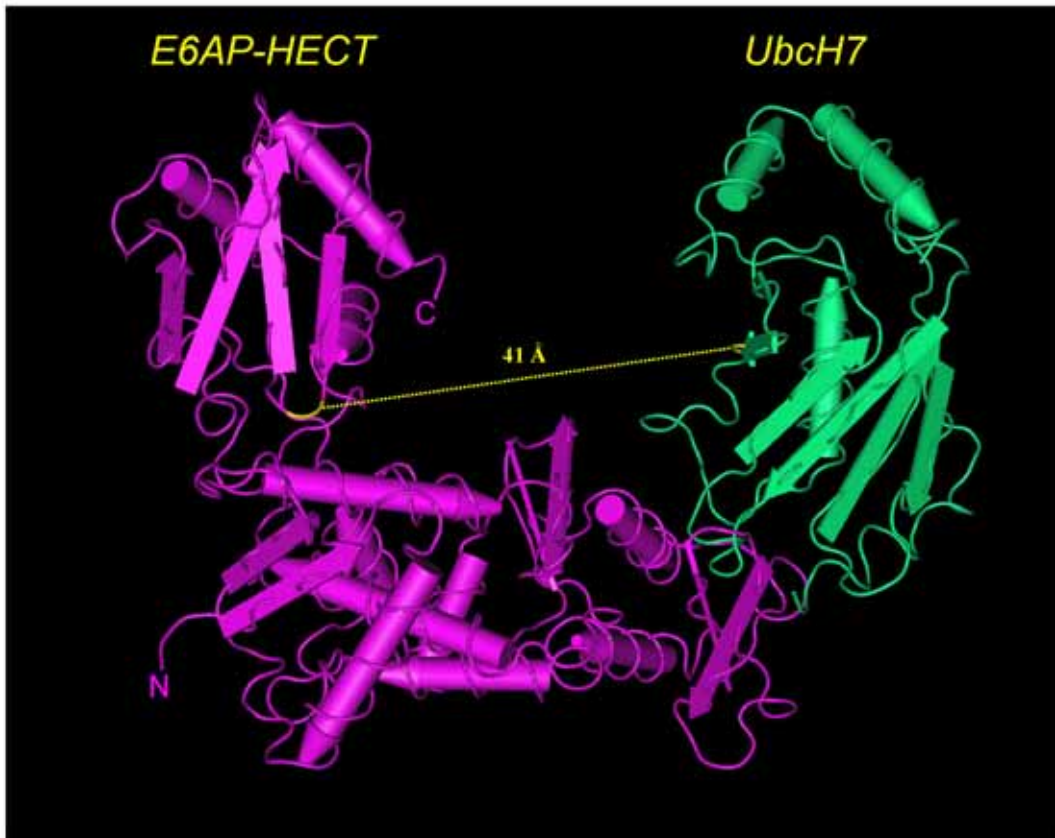


Figura 5. Estructura del dominio HECT de E6AP unido al enzima de tipo E2 UbcH7. Obsérvese como el dominio HECT (en color rosáceo) está formado por dos subdominios, denominados lóbulo N (abajo) y lóbulo C (arriba a la izquierda). El lóbulo N posee la región de unión al enzima E2 (en verde), mientras que el centro activo se halla en la zona de contacto entre ambos lóbulos. Las cisteínas conservadas a las que se une la ubiquitina tanto en el E2 (UbcH7) como en el E3 (E6AP) han sido marcadas en amarillo. Dichas cisteínas se encuentran cara a cara (no hay obstáculos entre ambas) a una distancia de 41Å. Esta distancia debe reducirse considerablemente para que se pueda producir la transferencia de ubiquitina entre una y otra.

siendo el otro principal subtipo el de las E3 con dedos de zinc de tipo RING (*RING finger E3s*). Contrariamente al resto de E3 conocidas, las E3 de tipo HECT se caracterizan por unirse covalentemente a la ubiquitina antes de transferirla definitivamente a sus sustratos. Dicha unión tiene lugar entre una cisteína conservada del centro activo del dominio HECT y el extremo carboxilo terminal de la ubiquitina. La estructura tridimensional del dominio HECT de E6AP unido al E2 UbcH7 ha sido resuelta por cristalografía de rayos X (ver Figuras 5-7). El dominio HECT se pliega

formando dos lóbulos, el lóbulo N y el lóbulo C, en cuya región de contacto se encuentra la cisteína conservada, situada en una *loop* del lóbulo C. Es el lóbulo N el que interacciona con UbcH7 por el extremo opuesto al que contacta con el lóbulo C

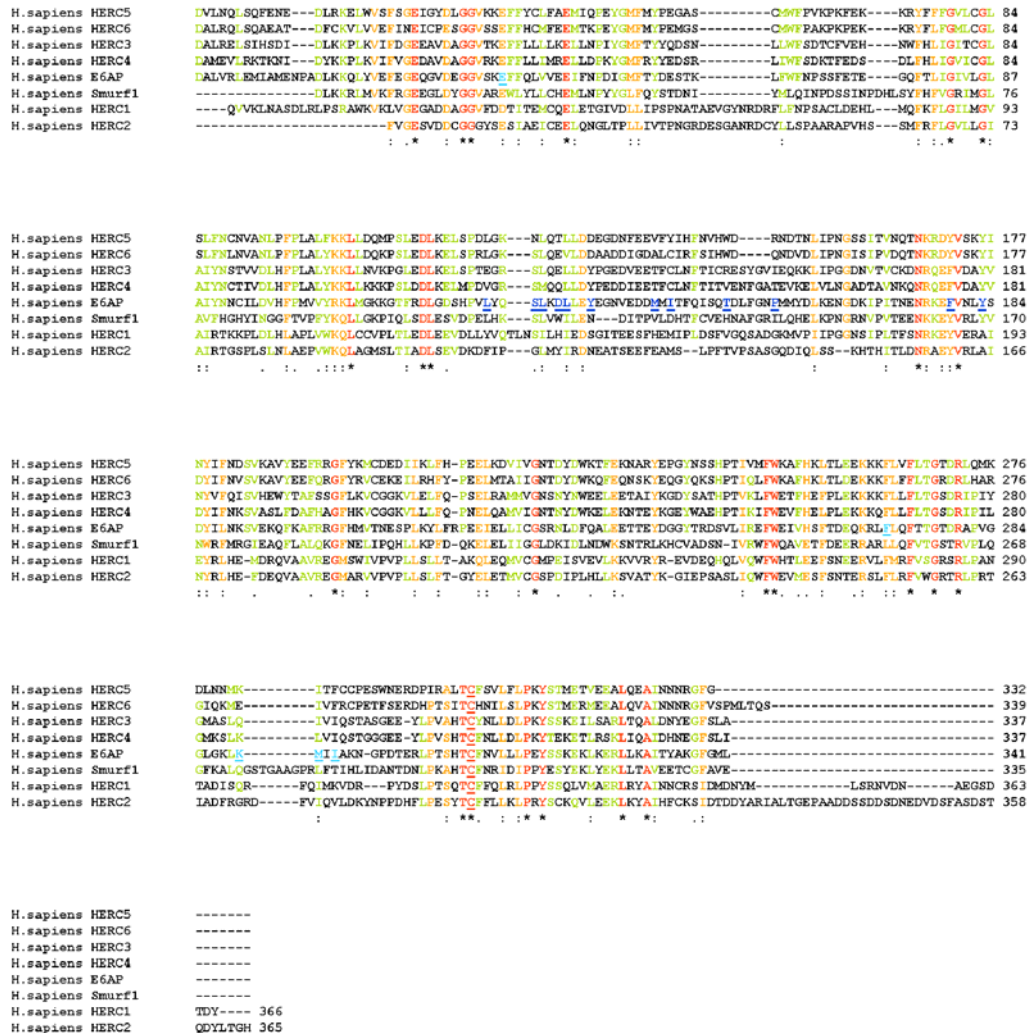


Figura 6. Alineamiento múltiple de los dominios HECT de las seis proteínas HERC humanas y de otras dos proteínas humanas con dominios HECT, E6AP y Smurf1. El código de colores es el siguiente: verde significa elevado grado de conservación, naranja muy elevado y rojo conservación total. La cisteína que forma el enlace tioéster con la ubiquitina ha sido subrayada en las ocho secuencias. Asimismo, sobre la secuencia de E6AP se han subrayado aquellos residuos que se han encontrado mutados en el síndrome de Angelman (azul claro) así como los implicados en la unión al enzima conjugador de ubiquitina o E2 (azul marino).

mediante una bisagra formada por sólo tres aminoácidos (Huang et al. (1999)). Dicha estructura, sin embargo, presenta un pequeño problema: desde la cisteína del centro activo de UbcH7 a la cisteína del centro activo del HECT hay una distancia de 41 Å, demasiado grande para que la cisteína del HECT pueda llevar a cabo el ataque

importancia de las proteínas HECT y de todas las E3/ubiquitina ligasas en general, no estriba en su capacidad de transferir ubiquitina a sus sustratos (de hecho en las E3 de

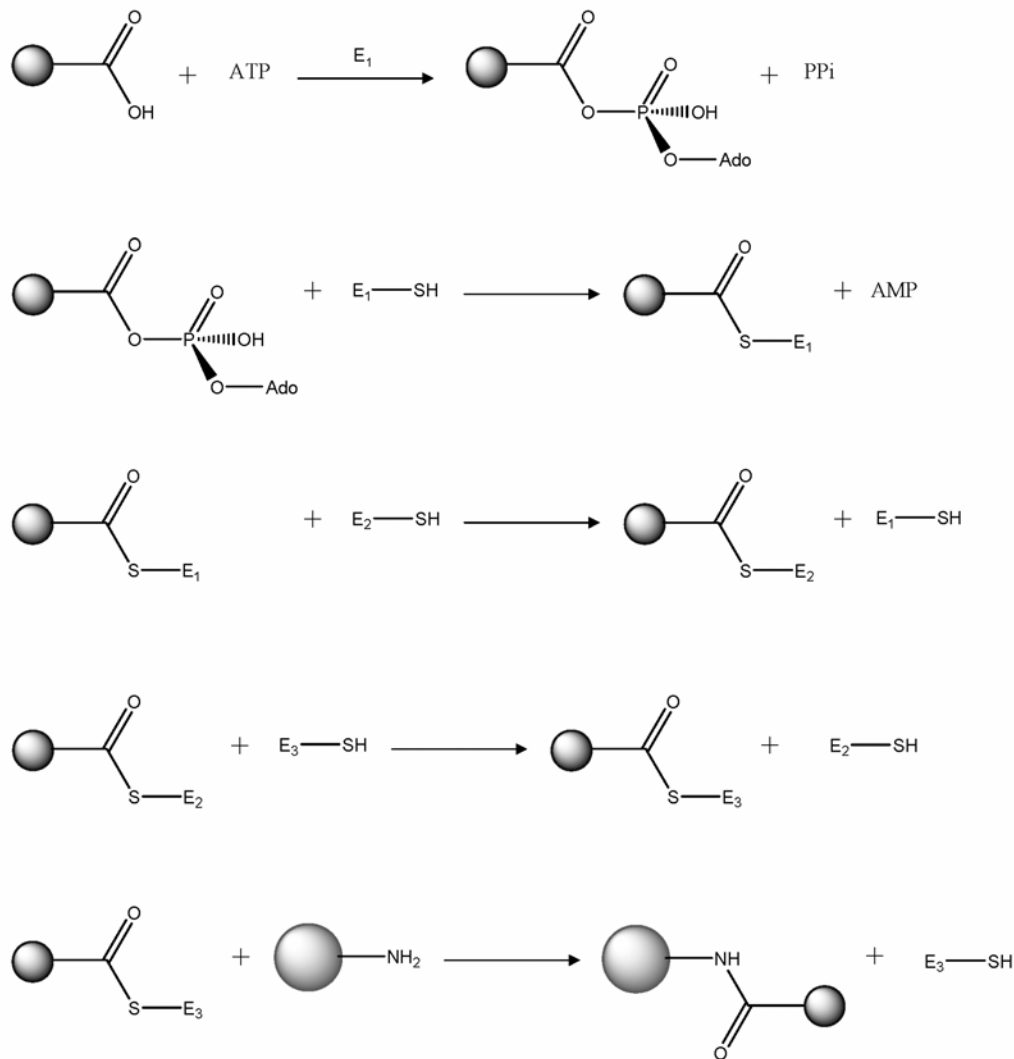


Figura 8. Esquema general del mecanismo de ubiquitinación de proteínas. El proceso por el cual una molécula de ubiquitina se conjuga a un sustrato incluye cinco reacciones secuenciales que involucran a tres enzimas además de al sustrato. Estas reacciones son: (1) activación del extremo carboxilo terminal de la ubiquitina mediante el uso de ATP para formar ubiquitina-adenilato y pirofosfato, todo ello catalizado por el enzima activador de la ubiquitina o E1; (2) conjugación del extremo C-terminal de la ubiquitina activada a una cisteína del centro activo del propio E1, liberándose adenilato como subproducto de la reacción; (3) transferencia de la ubiquitina desde la cisteína del E1 a la cisteína del centro activo del enzima conjugador de ubiquitina o E2, en una reacción de transtioesterificación que resulta en la regeneración del E1 libre; (4) transferencia de la ubiquitina desde la cisteína del E2 a la cisteína del centro activo de la ubiquitina ligasa o E3, en una reacción análoga a la anterior; (5) transferencia de la ubiquitina desde la cisteína del E3, que se regenera, a una lisina del sustrato, que queda pues ubiquitinado. Nótese que el mecanismo mostrado aquí es el utilizado por las E3 de tipo HECT, ya que todas las demás E3 no se unen covalentemente a la ubiquitina, la cual es transferida al sustrato directamente desde el E2. En estos casos el papel de la E3 es exclusivamente el de formar un puente entre el E2 y el sustrato.

tipo RING esta tarea es llevada a cabo directamente por los enzimas conjugadores de ubiquitina/E2), sino en el hecho que las E3 aportan la necesaria especificidad a la vía de ubiquitinación: en efecto, mientras que las células poseen solamente un enzima

activador de la ubiquitina (E1) y un número reducido de E2, existe un gran número de E3/ubiquitina ligasas, cada una de las cuales interacciona específicamente con uno o

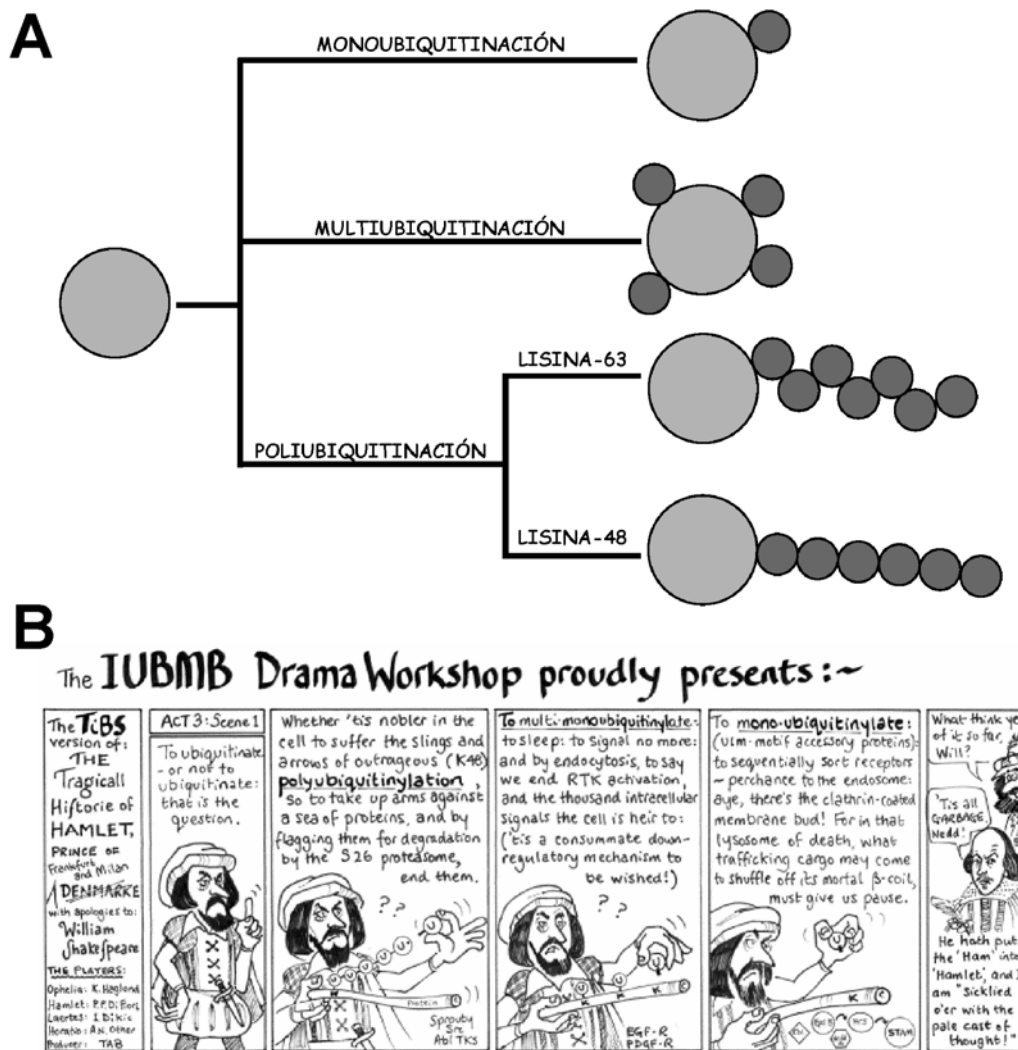


Figura 9. Tipos de ubiquitinación y su relevancia funcional. (A) Se distinguen en la actualidad tres grandes tipos de ubiquitinación, llamados mono-, multi- y poliubiquitinación (Haglund et al. (2003a)). La monoubiquitinación consiste en la conjugación de una sola molécula de ubiquitina a una lisina del sustrato y juega un papel importante en procesos tales como la regulación del tráfico vesicular, la gemación virica, la expresión génica o la reparación del DNA (Hicke (2001) y Sun y Chen (2004)). Por otra parte, la multiubiquitinación tiene lugar cuando varias lisinas del sustrato son modificadas cada una con una sola ubiquitina y se ha visto que tiene importancia en la internalización y tráfico posterior de receptores de membrana con actividad tirosina quinasa (Haglund et al. (2003b)). Finalmente, la poliubiquitinación se produce cuando una cadena de múltiples ubiquitinas es conjugada a una lisina del sustrato. Existen dos tipos mayoritarios de poliubiquitinación según los enlaces entre ubiquitinas sucesivas de la cadena se formen usando la lisina-48 (K48) o la lisina-63 (K63) de la ubiquitina proximal (la más cercana al sustrato: la ubiquitina distal se une por su extremo C-terminal). Mientras que las cadenas de tipo K63 son poco frecuentes y parecen ser importantes en la regulación de la actividad de ciertos enzimas y la reparación del DNA (Sun y Chen (2004)), las cadenas de tipo K48 son muy comunes y actúan como señal de reconocimiento por el proteasoma (siempre y cuando posean más de cuatro ubiquitinas) (Hicke (2001)). En levadura se han descrito incluso cadenas de tipo K29, aunque se desconoce su función (Hicke (2001)). (B) Tira cómica que ilustra lo comentado en el apartado anterior (sacada de Haglund et al. (2003a)).

unos pocos sustratos y con uno o unos pocos E2. Por lo tanto, según cuál sea la combinación de enzimas de tipo E2 y E3 que se hallan expresados y activos en una célula en un momento dado, las proteínas de esa célula que serán ubiquitinadas serán

unas u otras. En particular, sólo se ubiquitarán aquellas proteínas para las que la célula posea por lo menos un E3 activo que se una específicamente a ellas y un E2 activo que se una al susodicho E3. Tradicionalmente se había considerado que la ubiquitinación conllevaba la degradación de las proteínas ubiquitinadas en el proteasoma. No obstante, en los últimos años se ha demostrado que hay otros muchos procesos importantes para la célula que son regulados mediante ubiquitinación, tales como la expresión génica, la endocitosis mediada por receptor, la actividad de ciertos enzimas, etcétera (Figura 9) (para una revisión referirse a Pickart (2004)). En resumen, pues, las proteínas con dominios HECT permiten el correcto y fino funcionamiento de las células al especificar cuáles son las proteínas que deben ubiquitinarse en cada momento.

1.1.4. Las proteínas HERC en el ser humano

En *Homo sapiens* existen seis proteínas HERC, codificadas por seis genes ubicados en distintos lugares del genoma (Figuras 10 y 11). Las seis proteínas HERC de humanos pueden clasificarse en dos grupos en función de su tamaño. Así, los miembros

<i>Miembro de la familia</i>	<i>Localización cromosómica</i>	<i>Tamaño del gen (kb)</i>	<i>Número de exones</i>	<i>ID humano ratón (%)</i>	<i>Tamaño proteína (kDa)</i>
HERC1	15q22	225	78	96	532
HERC2	15q13	211	93	95	528
HERC3	4q21	116	26	92	117
HERC4	10q21	153	29	91	118
HERC5	4q22	49	23	n.a.	117
HERC6	4q22	64	25	66	115

Figura 10. Los genes HERC. En la tabla pueden observarse la localización cromosómica, tamaño y número de exones de los seis genes HERC humanos. Asimismo, en las dos columnas de la derecha se muestran la conservación (identidad humano-ratón) y el tamaño de las proteínas codificadas por dichos genes. (n.a.: no aplicable, puesto que no existe HERC5 en el ratón).

HERC1 y HERC2 se caracterizan por ser proteínas gigantes (cada una de ellas posee cerca de 5000 aa y pesa más de 500 kDa) y por contener en sus secuencias más de un dominio RLD además del dominio HECT y otros dominios adicionales. En cambio, los demás miembros de la familia, HERC3, HERC4, HERC5 y HERC6, presentan una

estructura mucho más compacta, que incluye un dominio RLD en el extremo amino terminal, una zona intermedia no conservada y un dominio HECT en el extremo carboxilo. Estas proteínas contienen algo más de 1000 aa y pesan cerca de 120 kDa. A continuación se resumen los conocimientos disponibles para cada una de estas seis proteínas.

1.1.4.1. La proteína HERC1

HERC1 puede ser considerado el miembro fundador de la familia, en tanto que fue el primero en ser descubierto. Hasta ahora ha sido también el más estudiado, aunque, como veremos, ello todavía no se refleja en un conocimiento detallado de su función. La proteína HERC1 (originalmente denominada p619 y luego también p532) fue identificada en el curso de una búsqueda de secuencias humanas oncogénicas en ensayos de tumorigenicidad en ratones atímicos (*nude mouse assays*) (para esta técnica ver Fasano et al. (1984)). Resumidamente, fibroblastos murinos NIH-3T3 fueron transfectados con DNA humano procedente de un adenocarcinoma mamario para luego ser inyectados en ratones carentes de inmunidad celular. Como consecuencia de la presencia de oncogenes humanos en las células inyectadas, éstas formaban tumores en los ratones indefensos. El DNA de estos tumores era a continuación aislado y en él se buscaban las secuencias humanas responsables de la transformación celular. Como resultado de este proceso, se logró finalmente hallar una secuencia de 150 kb con la capacidad de inducir tumores en los ensayos susodichos (curiosamente, las células transfectadas con este fragmento de DNA causaban tumores al ser inyectadas en ratones pero en cambio no generaban focos tumorales al ser cultivadas *in vitro*). Dicho fragmento, denominado OncH, poseía DNA procedente de como mínimo tres loci distintos, dos de los cuales eran de origen humano. El paso siguiente fue tratar de hallar

dentro de *OncH* secuencias que codificaran por alguna proteína. Ello se llevó a cabo mediante la técnica de “caza” de exones (*exon trapping*) y dio como resultado un exón de 382 bp procedente de uno de los dos loci humanos presentes en *OncH* (Rosa et al. (1996)). Análisis de la secuencia de este exón mostró una elevada homología con el gen *RCC1*, por entonces ya conocido como un importante regulador del ciclo celular. Partiendo, pues, de este fragmento de DNA codificante se procedió al aislamiento del cDNA completo a partir de una genoteca de cDNAs de cerebro fetal humano. Tras siete rondas consecutivas de cribaje (*screening*), se pudo reconstituir la secuencia completa del cDNA de este gen, al que se denominó entonces p619 y al que me referiré en adelante como *HERC1*. El cDNA de *HERC1* posee un marco abierto de lectura (ORF: *open reading frame*) de 14,586 bp que codifica por una proteína gigante de 4861 aa, 532

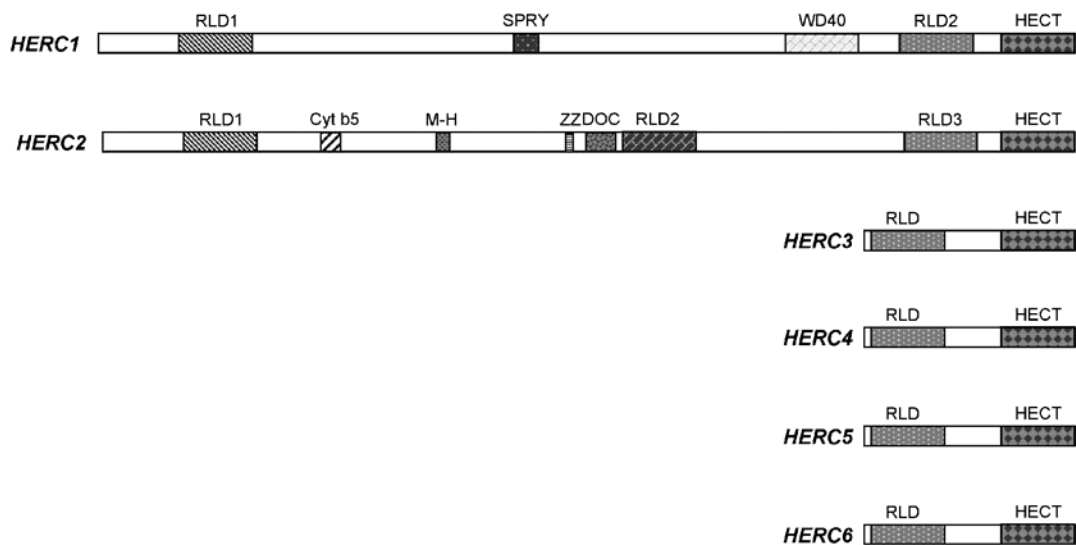


Figura 11. Las proteínas HERC. Esquema de las estructuras de las seis proteínas HERC de humanos. Las longitudes en la figura guardan relación con el número real de residuos aminoacídicos de las proteínas y dominios en cuestión. Los conocimientos disponibles sobre estas proteínas y los dominios que poseen se comentan en el texto.

kDa y un punto isoelectrónico teórico de 5.62. La comparación de la secuencia aminoacídica de *HERC1* con las de otras proteínas presentes en las bases de datos pone de manifiesto la presencia en *HERC1* de varias regiones conservadas evolutivamente. En primer lugar, *HERC1* posee dos dominios RLD altamente similares entre ellos (48%

idénticos) y que como se ha dicho están emparentados con la proteína RCC1, con la que comparten un 30% de identidad (ver sección 1.1.2 y Figuras 3-4). Además de los dominios RLD, HERC1 posee un dominio HECT bien conservado en su extremo C-terminal (ver sección 1.1.3 y figuras 6-7), un dominio WD40 (aa 3424-3775) y un dominio SPRY (aa 2072-2190). En cuanto al dominio WD40, se trata de un dominio presente en numerosas proteínas, entre ellas las subunidades β de las proteínas G heterotriméricas, cuya estructura es similar a la mencionada para los dominios RLD (*seven-bladed β -propeller*) (Sondek et al. (1996)) y cuya función parece estar en la mayoría de casos relacionada con su capacidad para interactuar con otras proteínas (Chen et al. (2004)). Por lo que respecta al dominio SPRY (*splA and RyR*), poco se sabe de su estructura (se ha postulado que podría plegarse de forma similar a las inmunoglobulinas (Seto et al. (1999)) y función, aunque ésta podría tener que ver también con interacciones con otras proteínas o incluso con RNAs (Ponting et al. (1997)). Otras regiones potencialmente interesantes en la secuencia de HERC1 son las siguientes: tres regiones ricas en prolina que podrían unir dominios SH3, regiones ricas en aminoácidos polares y ácidos y una secuencia similar a una cremallera de leucinas. El análisis de la expresión del mRNA de HERC1 muestra una distribución ubicua de este mensajero con niveles un poco mayores en cerebro y testículos. Asimismo, el mensajero de HERC1 se halla en todas las líneas celulares investigadas hasta la fecha, con la interesante peculiaridad de que los niveles son significativamente mayores en líneas transformadas con respecto a líneas que no lo están (Rosa et al. (1996)). Por su parte, los niveles de la proteína HERC1 parecen reflejar bastante fielmente los de su mRNA, habiendo sido aquella hallada en todas las líneas celulares y tejidos estudiados hasta el momento (FG y JLR, observaciones no publicadas). Por lo que concierne a su posible actividad oncogénica, el cDNA de HERC1 no induce la transformación de las

células NIH-3T3, ni siquiera en ensayos con ratones atímicos como sí hacía *OncH*. Esta observación puede tener por lo menos tres explicaciones: (1) de los dos loci humanos identificados en *OncH*, sólo uno contenía un oncogén y éste no era *HERC1* (el otro locus no se llegó a caracterizar); (2) *OncH* poseía una versión mutada de *HERC1* con actividad oncogénica, mientras que la forma salvaje finalmente clonada no actuaría como tal; (3) dado que se observó que los niveles de la proteína *HERC1* en las células NIH-3T3 transfectadas apenas superaban a los de las células no transfectadas, podría ser que la acción tumorigénica de *HERC1* requiriese de una mayor sobreexpresión de ésta (Rosa et al. (1996)). A nivel de localización subcelular, *HERC1* se encuentra exclusivamente en el citoplasma, donde aparece en estructuras vesiculares así como asociada al aparato de Golgi, de donde es dislocada por tratamiento con el metabolito fúngico brefeldina A. En cuanto a la función de *HERC1*, los primeros estudios en realizarse estuvieron encaminados a averiguar si los dominios RLD pueden actuar, como lo hacen en el caso de *RCC1*, como GEFs para proteínas G monoméricas. Dado que la localización subcelular de *HERC1* apuntaba a un posible papel en el tráfico intracelular de membranas, las GTPasas de las familias ARF y Rab constituyeron los primeros candidatos a posibles sustratos de *HERC1*. Efectivamente, se pudo comprobar que el dominio RLD1 purificado de *HERC1* es capaz de estimular la disociación de nucleótidos de guanina del centro activo de ARF1, Rab3a y Rab5, mientras que este mismo dominio no afecta a la tasa de disociación de nucleótido en otras GTPasas, como son Ran y R-Ras2/TC21. El dominio RLD2, por su parte, no posee actividad sobre ninguna de estas proteínas pero en cambio es capaz de interactuar con la forma miristoilada de ARF1 (pero no con Rab3a y Rab5), lo cual no ocurre con RLD1. Estos resultados, pues, apuntaban a un modelo según el cual *HERC1* sería un GEF para ARF1 (como veremos más adelante, esto no es estrictamente cierto), al que se asociaría por el

dominio RLD2 y activaría por el dominio RLD1 (algo similar podría ocurrir con Rab3a y Rab5). Estos datos eran además perfectamente compatibles con la localización subcelular de HERC1, puesto que ARF1 es un factor esencial para el correcto funcionamiento del aparato de Golgi, mientras que Rab3a y Rab5 regulan procesos exo- y endocíticos, respectivamente. La convicción de que HERC1 debe jugar un papel en el tráfico intracelular se vio reforzada poco más tarde por la observación de que HERC1, a través del dominio RLD2, puede unirse a la cadena pesada de la clatrina (CHC: *clathrin heavy chain*), un componente de la cubierta de muchas de las vesículas que se forman tanto en la red trans-Golgi (TGN: *trans-Golgi network*) como en la membrana plasmática (Rosa & Barbacid (1997)). Dicha interacción tiene lugar sólo cuando HERC1 se halla en la fracción citosólica, mientras que la fracción de HERC1 que se halla asociada a membranas intracelulares no interacciona con CHC. De acuerdo con esto, HERC1 no parece ser un componente constitutivo de las vesículas de clatrina (CCVs: *clathrin-coated vesicles*). A partir de estos datos, parece que lo más probable es que HERC1 actúe como regulador ya sea de la deposición o del desensamblaje de las cubiertas de clatrina sobre las vesículas. Esta última posibilidad es especialmente atractiva dado el hallazgo de que la chaperona Hsp70, previamente implicada en el proceso de desensamblaje (*uncoating*) de las CCVs, también se halla unida al complejo CHC-HERC1, del cual se disocia en presencia de ATP. Una posibilidad alternativa, o complementaria, a la anterior sería que HERC1 regulara la unión entre las cadenas ligera (CLC) y pesada (CHC) de la clatrina, lo cual a su vez podría afectar a la dinámica de formación o desensamblaje de las cubiertas de clatrina. Ello es compatible con la observación de que HERC1 se une a CHC por una zona que comprende los aminoácidos del 1315 al 1557 de ésta (Rosa & Barbacid (1997)), dentro de los cuales también está el lugar de unión a CLC. Para averiguar si es éste el caso sería necesario

ver si CLC también se halla asociada al complejo ternario HERC1-CHC-Hsp70. Por último, hay también estudios hechos que aportan datos sobre la funcionalidad del dominio HECT de HERC1 (Schwarz et al. (1998)). Si bien el HECT de HERC1 no es capaz de suplir al HECT de E6AP en la ubiquitinación de p53 mediada por la oncoproteína papilomavírica E6, el HECT de HERC1 sí parece ser plenamente funcional en tanto que es capaz de unirse a ubiquitina mediante un enlace sensible a ditiotreitól (con toda probabilidad un enlace tioéster entre la cisteína-4811 de HERC1 y el extremo COOH de la ubiquitina). La formación de dicho enlace requiere además de un enzima E2 específico: el HECT de HERC1 sólo acepta ubiquitina procedente de UbcH5, pero no de UbcH1, UbcH6 o UbcH7. Sin embargo, la pregunta más interesante, es decir a qué otra proteína/s transfiere HERC1 ubiquitina una vez la ha recibido de UbcH5, es todavía un misterio.

1.1.4.2. La proteína HERC2

Aunque HERC1 fuese el primer miembro de la familia en ser identificado, la historia de HERC2 se remonta bastante más tiempo en el pasado. Así, HERC2 fue hallado como consecuencia del análisis de la complementación entre distintos alelos mutantes del locus *p* situado en el cromosoma 7C del ratón. El locus *p* (*pink-eyed dilute*) fue uno de los primeros en ser estudiados dados sus efectos en la pigmentación tanto de los ojos como del pelaje de estos roedores. Los mutantes de este locus, obtenidos inicialmente por irradiación de las espermatogonia con rayos X, se caracterizaban por distintos grados de hipopigmentación, pero eran al margen de esto perfectamente viables y fértiles. Sin embargo, ya en 1960 fueron descritos mutantes de este locus que, además de la hipopigmentación característica, presentaban un cuadro bastante más grave de alteraciones que incluía un crecimiento reducido (*runtiness*),

movimientos descoordinados (*jerky gait*), fertilidad muy disminuida en las hembras, deficiencias en el comportamiento de las madres para con sus crías, espermatogénesis anómala y total esterilidad en los machos y viabilidad reducida (Hollander et al. (1960)). Más adelante este síndrome sería bautizado como *rjs* (*runty, jerky, sterile*) o también *jdf2* (*juvenile development and fertility-2*). Inicialmente, tan elevado grado de pleiotropía parecía incompatible con la implicación de un solo gen. Así, refiriéndose a la multiplicidad de efectos observada en estos mutantes, Hollander et al. (1960) escriben: “...*these seem impossible to fit into the one-gene-one-enzyme hypothesis of gene action, and similar theories, namely that all the effects can be traced to a single derangement of a primordial physiological process. It is at present impossible to conceive of such a common source of the changes here observed...*”. A pesar del escepticismo de Hollander et al., el origen común de las alteraciones mencionadas acabaría siendo descubierto, aunque todavía tendrían que transcurrir 38 años. No obstante, no tendría que pasar tanto tiempo para que los científicos se convencieran del origen monogénico de la pleiotropía observada en el síndrome *rjs*. Así, un análisis exhaustivo de la complementación entre distintos alelos del locus p mostró claramente que aquellos alelos que daban lugar al fenotipo *rjs* eran incapaces de complementarse entre si, mientras que sí eran complementados por aquellos alelos del locus p que generaban exclusivamente alteraciones pigmentarias (Lyon et al. (1992)). Estos datos, juntamente con la incapacidad de segregar los diversos efectos observados en los ratones *rjs* en distintos mutantes, ya indicaba que probablemente el fenotipo *rjs* era debido a un solo gen, distinto a su vez del gen causante de la hipopigmentación. Prueba irrefutable de esto último se alcanzaría tres años después cuando Rinchik et al. (1995) fueron capaces de crear, mediante un protocolo de mutagénesis química con N-etil-N-nitrosourea (ENU), tres mutantes puntuales (con un único cambio en la secuencia

nucleotídica) que en homocigosis presentaban la totalidad de los síntomas del síndrome *rjs*. El siguiente paso era el clonaje del gen responsable de estos efectos y fue logrado tres años más tarde por Lehman et al. (1998) e independientemente por Ji et al. (1999). Mientras que en el primero de estos trabajos se demostraba que tres de los mutantes con el fenotipo *rjs* tenían deleciones en el gen HERC2, una de las cuales tenía su principio y fin dentro de este gen, todavía quedaba abierta la posibilidad, aunque remota, de que el fenotipo no fuese causado por el gen HERC2 sino por otro gen ubicado dentro de un intrón de aquél que también hubiese sido delecionado. Dicha posibilidad quedó totalmente descartada cuando Ji et al. (1999) demostraron que las tres mutaciones puntuales causantes del síndrome *rjs* creadas previamente por Rinchik et al. (1995) consistían en mutaciones del gen HERC2 que eliminaban lugares consenso de *splicing* dentro del pre-mRNA de HERC2 y por consiguiente ocasionaban la pérdida de exones (*exon skipping*) en el mRNA maduro y la formación de proteínas truncadas o con deleciones internas. Así quedaba, pues, demostrado que alteraciones en la proteína HERC2 eran la causa del síndrome *rjs/jdf2* en ratón. Ji et al. (1999) también clonaron el ortólogo humano de HERC2 y mostraron además que duplicaciones y translocaciones recientes (durante la evolución de los primates e incluso del hombre) de porciones de este gen forman las repeticiones transcritas de bajo número de copias (*transcribed low-copy repeats*), también denominadas duplicones, que se encuentran en los lugares de corte de las deleciones (*deletion breakpoint hotspots*) que dan lugar a un 70% de los casos de síndrome de Prader-Willi en humanos. Así, el síndrome de Prader-Willi (PWS) aparece cuando durante la profase I de la meiosis que tiene lugar durante la espermatogénesis del padre se produce un error en la recombinación entre los cromosomas 15 homólogos (error ocasionado por la existencia de duplicones de HERC2 en las regiones 15q11 y 15q13), de tal modo que uno de los cromosomas

formados posee una delección de unas 4 Mb entre 15q11 y 15q13. El nuevo individuo será pues heterocigoto para dicha delección, pero padecerá PWS porque los genes situados dentro de un fragmento de 2 Mb en la región 15q11-15q13 de su cromosoma 15 materno habrán sido silenciados durante la oogénesis o las primeras fases de desarrollo del embrión en un fenómeno conocido como *imprinting* genómico. Aunque el propio gen HERC2 ha sido hallado alterado en algunos casos de PWS y los síntomas de éste recuerdan en cierta medida a los de los ratones *rjs*, el hecho de que el gen HERC2 no sea silenciado en el cromosoma materno y por tanto los individuos con PWS sean hemicigotos para este gen, juntamente con la herencia recesiva del fenotipo *rjs* en ratón, hacen pensar que no es HERC2 sino otros genes que sí son silenciados en el cromosoma 15 materno los que dan lugar a PWS (para una revisión sobre este complejo asunto ver Nicholls & Knepper (2001)). En cualquier caso, los duplicones de HERC2, pese a ser a veces transcritos y quizá traducidos, no parecen tener relevancia funcional alguna, puesto que además de constituir solamente trozos del gen HERC2, están sujetos a elevadas tasas de mutación (las mismas que cualquier intrón) y sus ORFs poseen codones stop prematuros en todas las pautas de lectura. De hecho, estos duplicones son considerados como pseudogenes, siendo el único gen auténtico el gen HERC2 que ha dado lugar a ellos y que es el único que se encuentra en el ratón.

Por lo que respecta a la proteína HERC2, se trata de una proteína gigante de 4834 aa, 528 kDa y que posee varias regiones conservadas en su secuencia, entre las que cabe destacar el dominio HECT, tres dominios RLD, un dominio DOC, un dominio M-H, una región homóloga al citocromo b5 y un dedo de zinc de tipo ZZ. El dominio DOC fue descrito inicialmente por Grossberger et al. (1999) como un dominio de unos 200 aa con homología a la subunidad APC10/Doc1 del complejo promotor de la anafase (APC: *anaphase-promoting complex*), también llamado ciclosoma. Este complejo

multiproteico es un regulador clave de la mitosis gracias a su actividad E3 ubiquitina ligasa sobre proteínas que controlan la progresión a través de la fase M como son la securina o la ciclina B. Recientemente se ha descrito que APC10/Doc1 se pliega formando un β -*sandwich* (Wendt et al. (2001) y Au et al. (2002)) y que dicha estructura contribuye a aumentar considerablemente la afinidad de APC por sus sustratos y por consiguiente a acelerar la velocidad de ubiquitinación de éstos (Carroll & Morgan (2002) y Passmore et al. (2003)). Dominios DOC han sido hallados en varias otras proteínas, todas ellas poseedoras de ya sea dominios HECT o dominios *Cullin*, también implicados en procesos de ubiquitinación (Grossberger et al. (1999)). Dado que el dominio DOC de HERC2, así como el HECT, están muy conservados en la evolución, es probable que el dominio DOC contribuya de forma importante a la unión de HERC2 a sus sustratos, los cuales serían subsiguientemente ubiquitinados por el dominio HECT. Otro dominio que podría estar implicado en la actividad de HERC2 como ubiquitina ligasa es el dominio M-H (Mib/Herc2). Aunque se desconoce la función de dicho dominio, su elevado grado de conservación así como su presencia en Mind Bomb, una ubiquitina ligasa del pez cebra (Itoh et al. (2003)), apuntan a su posible importancia funcional. También muy conservados están el hipotético dedo de zinc de tipo ZZ, cuya posible función en HERC2 se desconoce, y el dominio homólogo al citocromo b5. Este último carece de los dos residuos de histidina esenciales para la coordinación del grupo hemo en el citocromo b5. Aún y así, la cavidad hidrofóbica ocupada por el hemo en el citocromo b5 podría dar cabida en HERC2 a algún otro ligando, como parece ocurrir en el caso de los receptores de membrana de la progesterona (MAPRs: *membrane-associated progesterone receptors*) (Mifsud & Bateman (2002)). Aunque todavía no existen estudios hechos con la proteína HERC2, el análisis de su mRNA así como de los distintos mutantes generados en el ratón permite sacar algunas conclusiones

interesantes. Así, el mensajero de HERC2 es expresado de forma aparentemente ubicua, aunque con niveles significativamente mayores en cerebro y testículos (Lehman et al. (1998)). Mientras que la mayoría de mutantes estudiados expresan proteínas con truncaciones que abarcan toda o buena parte de la misma, existen dos mutaciones que revisten un especial interés: en la primera, la proteína codificada es normal excepto por la ausencia de los residuos 4428-4748 (321 aa), lo cual supone una delección de buena parte del dominio HECT y con toda seguridad la pérdida de función del mismo. En el segundo caso, la delección es de sólo 53 aa (3716-3768) y se encuentra en una región de la proteína (entre los dominios RLD2 y RLD3) que no presenta homología con ninguna otra proteína conocida. Ambas mutaciones dan lugar al mismo fenotipo, el síndrome *rjs*, cuando se encuentran en homocigosis. Por lo tanto, parece obvio que este fenotipo se debe a la pérdida por parte de HERC2 de la capacidad de ubiquitinar a uno o más de sus sustratos. En este sentido, es probable que los aminoácidos 3716-3768 sean necesarios para la unión de este/estos sustratos a HERC2. Los procesos en los que la actividad ubiquitina ligasa de HERC2 es requerida *in vivo* pueden ser vagamente delineados en base a los síntomas presentados por los ratones *rjs*. Así, los déficits en el crecimiento y en la maduración testicular y ovárica podrían ser explicados por una reducción en la secreción de hormonas hipotálamicas tales como la somatotropina (GH) y las gonadotropinas (FSH y LH), cuya secreción está a su vez controlada por la acción de neuronas hipotálamicas, algunos de cuyos axones neurosecretorios se hallan degenerados en los ratones *rjs* (Johnson & Hunt (1975)). Otras neuronas podrían asimismo dar cuenta de los defectos observados en la coordinación de los movimientos y el comportamiento maternal. Por último, ha sido demostrado que los defectos en la espermatogénesis no pueden ser explicados en términos de disfunciones del sistema hipotálamo-hipófiso-testicular, sino que son intrínsecos a la propia línea germinal

(Handel et al. (1987)). De hecho, los espermatozoides aberrantes de los ratones *rjs* parecen sufrir problemas principalmente con la formación de sus acrosomas. Si se tiene en cuenta que dichos orgánulos son algo semejante a un gran “gránulo” secretor derivado del aparato de Golgi, resulta atractivo especular con la posibilidad de que todos los efectos derivados de la pérdida de función de HERC2 se deban a un defecto concreto en las vías de secreción celular, el cual se pondría más de manifiesto en aquellos órganos, cerebro y testículos, en que HERC2 parece jugar un papel más importante. Sin embargo, muchos más estudios fisiológicos con los ratones *rjs* y moleculares con la proteína HERC2 van a ser necesarios para poder establecer con firmeza la función de esta proteína.

1.1.4.3. La proteína HERC3

El cDNA de la proteína HERC3 humana, por entonces designada aún como D25215, fue identificado por primera vez en una búsqueda aleatoria de cDNAs mayores de 2 kb (Nomura et al. (1994)). No fue hasta 7 años más tarde que apareció en la literatura la identificación y caracterización preliminar del producto de este gen ubicado en nuestro cromosoma 4 (Cruz et al. (1999 y 2001)). Se trata de una proteína de 117 kDa y expresión más o menos ubicua (aunque la mayor expresión, al menos en ratón, parece hallarse en el cerebro, y más concretamente en la corteza piriforme, el hipocampo y la amígdala (Davies et al. (2004)) que se localiza en el citoplasma de las células, donde se halla asociada al menos parcialmente a estructuras vesiculares en las que colocaliza con marcadores de las vías de tráfico intracelular como β -COP, Rab5 y ARF, pero no con marcadores lisosomales (LIMP-II) o del aparato de Golgi (GMPT1). Aunque no puede descartarse que el dominio RLD de HERC3 posea actividad disociadora de nucleótido para proteínas G monoméricas como es el caso de RCC1 y

HERC1, dicha actividad no se halló para ninguna de 9 de las susodichas GTPasas, a saber Ran, Rab3a, Rab5, Rab8, ARF1, ARF6, Ras, Rac y RhoA. Por otro lado, el dominio HECT de HERC3 sí parece ser plenamente funcional, puesto que es capaz de conjugar ubiquitina in vitro en presencia del E2 UbcH7 y sólo en ausencia del reductor ditiotreitól, lo cual sugiere que la unión HECT-ubiquitina tiene lugar a través de una cisteína, muy probablemente la cisteína-1018 del centro activo (Schwarz et al. (1998)). De hecho, los resultados de Cruz et al. (2001) indican que HERC3 es capaz de unir ubiquitina incluso cuando la cisteína-1018 ha sido mutada a alanina, aunque con menor afinidad. Dichos resultados pueden en principio ser explicados ya sea por una unión no covalente de la ubiquitina al centro activo del HECT o bien por la existencia en HERC3 de un lugar de unión a ubiquitina independiente de su centro activo. En cualquier caso, HERC3, además de unir ubiquitina por su centro activo como corresponde a una ubiquitina ligasa de tipo HECT, también la une, independientemente, a través de uno o más de sus residuos de lisina. En efecto, HERC3 no sólo es un enzima ubiquitinador de otras proteínas aún por identificar, sino que él mismo también es el sustrato de otro enzima ubiquitinador o incluso de sí mismo. Esta última posibilidad es sin embargo poco probable, pues HERC3-Cys1018Ala es ubiquitinado con la misma intensidad y su vida media en células es la misma que la de la forma salvaje. Sea como fuere, lo que sí está claro es que HERC3 es poliubiquitinado y ello conlleva su degradación en el proteasoma, con una vida media aproximada de 8 horas. En resumen, HERC3 es una ubiquitina ligasa que a su vez es regulada por ubiquitinación y que probablemente interviene en el control del tráfico intracelular.

1.1.4.4. La proteína HERC4

El cDNA de HERC4 fue originalmente descrito, como el de HERC3, como resultado de una búsqueda de cDNAs que codificaran proteínas grandes (Nagase et al. (2000)). Sin embargo, la secuencia descrita inicialmente resultó ser incompleta y la publicación del cDNA entero no tuvo lugar hasta muy recientemente (Hochrainer et al. (2005)). Los pocos conocimientos disponibles hasta la fecha sobre la proteína HERC4 y el gen que la codifica han sido publicados en este mismo trabajo. En cuanto a la expresión del mRNA de HERC4, éste parece estar distribuido de forma bastante ubicua, aunque existen diferencias notables en los niveles del mismo en distintos órganos, siendo el cerebro y los testículos los que albergan las mayores cantidades (en esto último se asemeja HERC4 a otras proteínas de la familia). Por otra parte, análogamente a las demás proteínas HERC estudiadas, HERC4 se localiza en el citoplasma de la célula, presentando por inmunocitoquímica un marcaje punteado indicativo de asociación a estructuras vesiculares. Sin embargo, el aspecto más sorprendente de los resultados de Hochrainer et al. con respecto a HERC4 (y a HERC6, como veremos) es sin duda el referente al procesamiento del transcrito primario de este gen. Así, el gen HERC4 posee 29 exones potenciales, de los cuales normalmente sólo 25 o 26 pasan a formar parte del mRNA maduro (4.45 kb), dando lugar a sendas proteínas de 1049 y 1057 aa que constituyen las formas mayoritarias de la proteína HERC4 dentro de las células. No obstante, aunque poco a menudo, el procesamiento del pre-mRNA toma otros derroteros. En estos casos, pueden suceder cosas tales como la exclusión de los 5 primeros exones usados habitualmente y sustitución de éstos por otro exón que contiene un codón alternativo de inicio de la traducción (dando así lugar a una proteína que carece del trozo inicial del dominio RLD), la exclusión de los exones 24 y 25 (ambos en la zona que codifica el dominio HECT, el cual, pues, queda delecionado), la inclusión tras el tercer u onceavo exón de un exón no usado normalmente y que posee un codón

STOP (generando proteínas truncadas, en el primer caso de sólo 110 aa y sin ni siquiera el dominio RLD) o la exclusión de los exones 5 y 9, causando un desplazamiento de la pauta de lectura y por tanto de nuevo la formación de una proteína truncada (de 118 aa). La relevancia funcional, si la hay, de tal maremagnum de procesamientos (*splicings*) alternativos es totalmente desconocida, del mismo modo en que lo es, por el momento, la función de HERC4.

1.1.4.5. La proteína HERC5

HERC5, inicialmente descrita con el nombre de Ceb1 (*Cyclin E-binding protein-1*) (Mitsui et al. (1999)) y posteriormente también denominada HERC4 (Cruz et al. (2001) y Garcia-Gonzalo et al. (2003 y 2004)), es una proteína de 117 kDa cuyo mRNA se halla altamente expresado en testículos (100X) y cerebro fetal (30X) y en mucho menor grado (~1X) en otros órganos como son los ovarios, el páncreas, el corazón, la placenta y el músculo esquelético. La proteína HERC5 fue identificada inicialmente al ser “pescado” su cDNA en una genoteca de linfocitos B en el transcurso de un ensayo de doble híbrido en levadura en el que se usaba como cebo a la proteína ciclina E (Mitsui et al. (1999)). Más recientemente, HERC5 reapareció en la literatura científica al ser su cDNA clonado durante una búsqueda de genes estimulados por citoquinas pro-inflamatorias (LPS y TNF- α) en células endoteliales microvasculares de piel humana (Kroismayr et al. (2004)). Ambos trabajos aportan contribuciones muy valiosas para la comprensión de la función de esta proteína, las cuales se resumen a continuación. Al margen de la distribución tisular ya citada, la localización subcelular de HERC5 es análoga a la observada para las demás proteínas HERC pequeñas, es decir citoplásmica, con un componente soluble o citosólico y otro asociado a estructuras vesiculares, visibles en el microscopio como puntos discretos en la zona perinuclear (Mitsui et al.

(1999)) (cabe decir aquí que Kroismayr et al. afirman en su artículo, erróneamente a mi parecer, que “*HERC5 does not localize to cytosolic vesicles*” (Kroismayr et al. (2004), pág.4755). Ello se debe probablemente a una mala interpretación de las palabras de Mitsui et al. (1999), pág.119, en las que estos hablan de “*exclusive cytoplasmic staining*”; en cualquier caso, las imágenes en este último trabajo no dejan lugar a dudas.). Contrariamente a los demás genes HERC estudiados, los cuales parecen expresarse constitutivamente en todos los tejidos, la expresión del gen HERC5 es objeto de una regulación bastante fina. En efecto, si bien la expresión de HERC5 es muy baja en la mayoría de los tejidos, los niveles del mRNA y de la proteína HERC5 aumentan considerablemente cuando las células son sometidas a tratamientos, tales como la expresión de oncoproteínas víricas, que inactivan a las proteínas supresoras de tumores p53 y Rb (Mitsui et al. (1999)). Por otro lado, los niveles del mRNA de HERC5 también aumentan en respuesta a estímulos pro-inflamatorios (Kroismayr et al. (2004)). En particular, las citoquinas LPS, TNF- α e IL-1 β inducen un claro aumento de la cantidad del transcrito de HERC5 a las 8 horas tras la estimulación con estas sustancias de endotelios de piel humana (*Human skin microvascular endothelial cells, HSMECs*). Dicho aumento requiere de la previa activación del factor de transcripción NF κ B, así como de la síntesis de nuevas proteínas, por lo que probablemente la transcripción de HERC5 (o la estabilidad de su mensajero) dependa de algún factor cuya expresión, a su vez, esté regulada por NF κ B. Tal interpretación estaría además en concordancia con la observación de que la zona proximal del promotor de HERC5 no contiene ningún lugar consenso de unión a NF κ B y sí en cambio a otros factores de transcripción relacionados con la inflamación. Sin embargo, tras el tratamiento de HSMECs con LPS, los niveles de la proteína HERC5 no reflejan los cambios que se producen en los niveles de su mRNA. Ello es debido a un efecto independiente del LPS

sobre la vida media de la proteína. Este efecto es mucho más veloz que el que tiene lugar con el mRNA y conlleva una rápida degradación de la proteína, cuya vida media pasa de ser de 9 horas en ausencia de LPS a 2 horas en su presencia. Así, los niveles de la proteína HERC5 caen rápidamente en respuesta a LPS para recuperarse luego, tras aproximadamente 12 horas, gracias al aumento en los niveles del mRNA. Aunque no ha sido probado, la degradación de HERC5 inducida por LPS se debe con toda probabilidad a la poliubiquitinación y degradación de ésta en el proteasoma, lo cual hace de HERC5 la segunda proteína HERC, después de HERC3, de la que se describe que es un sustrato de ubiquitinación, además de una ubiquitina ligasa activa. Por último en cuanto a la regulación de HERC5, parece razonable suponer que su elevada y bastante selectiva expresión en testículos sea debida, aunque caben otras posibilidades, a la existencia en el promotor de su gen de algún elemento de respuesta a algún factor de transcripción específico de este órgano. Todos estos datos aportan pistas acerca de la posible función o funciones de HERC5 dentro de las células. En primer lugar, por lo que respecta a los testículos, es probable que el papel de HERC5 en éstos tenga que ver con la destrucción masiva de proteínas que tiene lugar en el transcurso de la espermatogénesis, y que se lleva a cabo por vía del sistema ubiquitina-proteasoma. Parece plausible que HERC5, en virtud de su actividad E3/ubiquitina ligasa, participe en este proceso, puesto que las E3 asociadas al mismo todavía están por descubrir. Algo semejante podría ocurrir con la inflamación. Hemos visto que HERC5 desaparece durante la fase temprana de la respuesta inflamatoria, lo cual permitiría que se acumulase/n su/s sustrato/s temporalmente, hasta que los niveles de HERC5 se restablecieran al cabo de unas 12 horas, contribuyendo así a la resolución de la inflamación: esto encajaría además con la demostrada importancia de la degradación proteica durante la resolución de la inflamación (Majetschak et al. (2000)). Por último,

algunos de los datos presentados apuntan a una posible intervención de HERC5 en el ciclo celular. Concretamente, la regulación de HERC5 por parte de p53 y Rb así como su unión a la ciclina E (la cual se une a la mitad C-terminal de HERC5) invita a pensar que HERC5 podría jugar un papel en la entrada a la fase S del ciclo. No obstante, hasta la fecha no ha sido demostrada ninguna oscilación en los niveles de HERC5 a lo largo del ciclo celular. Asimismo, resulta difícil hallar en este contexto una explicación satisfactoria para el hecho de que la interacción de HERC5 con ciclina E no sea específica para ésta: las ciclinas A, B1 y D1 también parecen tener afinidad por HERC5, la cual contiene dos secuencias consenso de unión a ciclinas. En resumen, aunque la información disponible ya permite percibir vagamente en qué consisten las tareas que HERC5 lleva a cabo dentro de las células, todavía hace falta mucho trabajo para poder disponer de explicaciones detalladas a este respecto.

1.1.4.6. La proteína HERC6

El gen HERC6 ha sido identificado y caracterizado de forma preliminar, juntamente con el gen HERC4, en un trabajo de este mismo año (Hochrainer et al. (2005)). Del mismo modo que hemos visto que ocurría con HERC4, quizá el aspecto más llamativo del estudio realizado sobre HERC6 lo constituye la complejidad existente en el procesamiento del pre-mRNA. En el caso de HERC6, nos enfrentamos a un gen formado por 25 exones, de los cuales normalmente sólo 23, todos menos el 9 y el 17, acaban integrando el mensajero maduro (3.89 kb), el cual entonces codifica para una proteína, la forma mayoritaria de HERC6 en la célula, de 1022 aa de longitud. Con menor frecuencia, no obstante, el transcrito primario de HERC6 es procesado de forma distinta para dar lugar a por lo menos tres mRNAs minoritarios, en los que faltan los exones 9, 16 y 17 (proteína de 986 aa con delección interna en una zona no conservada),

9, 10 y 14 (el empalme entre los exones 8 y 11 provoca entonces el desplazamiento de la pauta de lectura y la formación de una proteína de 364 aa que sólo contiene el dominio RLD) o 7, 9, 10, 14 y 17 (proteína de 322 aa). De nuevo, se desconoce aquí el posible papel de tales formas minoritarias de HERC6, aunque cabe especular con la posibilidad de que las formas que poseen el dominio RLD pero no el HECT puedan intervenir en la regulación de la función de los presuntos sustratos de ubiquitinación de HERC6, al unirse a ellos sin ser capaces de ubiquitinarlos. Finalmente, el análisis de la distribución tisular y localización subcelular de HERC6 muestran un patrón muy similar al de HERC4: expresión ubicua con mayores niveles en cerebro y testículos así como marcaje citoplásmico punteado.

1.1.5. Evolución de los genes *HERC*

El gran desarrollo experimentado por la ciencia genómica a lo largo de los últimos años está permitiendo conocer en mayor profundidad de qué manera los genomas de los seres vivos han ido evolucionando a lo largo de la historia de la vida en la Tierra. La información disponible en este campo ha sido aplicada recientemente al estudio de la evolución de los genes de la familia HERC (Hochrainer et al. (2004)). Dicho estudio ha permitido extraer algunas conclusiones interesantes. En primer lugar, al elaborar un árbol filogenético basado en las secuencias disponibles de genes HERC de distintas especies (Figura 12), salta a la vista la existencia de las dos familias mencionadas más arriba para las correspondientes proteínas: las HERC grandes, HERC1 y HERC2, por un lado, y las HERC pequeñas, HERC3, HERC4, HERC5 y HERC6, por otro. Interesantemente, el gen HERC4 de *Caenorhabditis elegans* surge directamente del tronco del árbol, sugiriendo que HERC4 representa el gen HERC ancestral, a partir del cual derivan todos los demás. El hecho de que organismos como

Saccharomyces cerevisiae o *Arabidopsis thaliana* no presenten ningún gen HERC (es decir, ningún gen que contenga a la vez dominios RLD y HECT) pero sí posean en cambio genes RLD y genes HECT, sugiere que el propio HERC4 habría aparecido

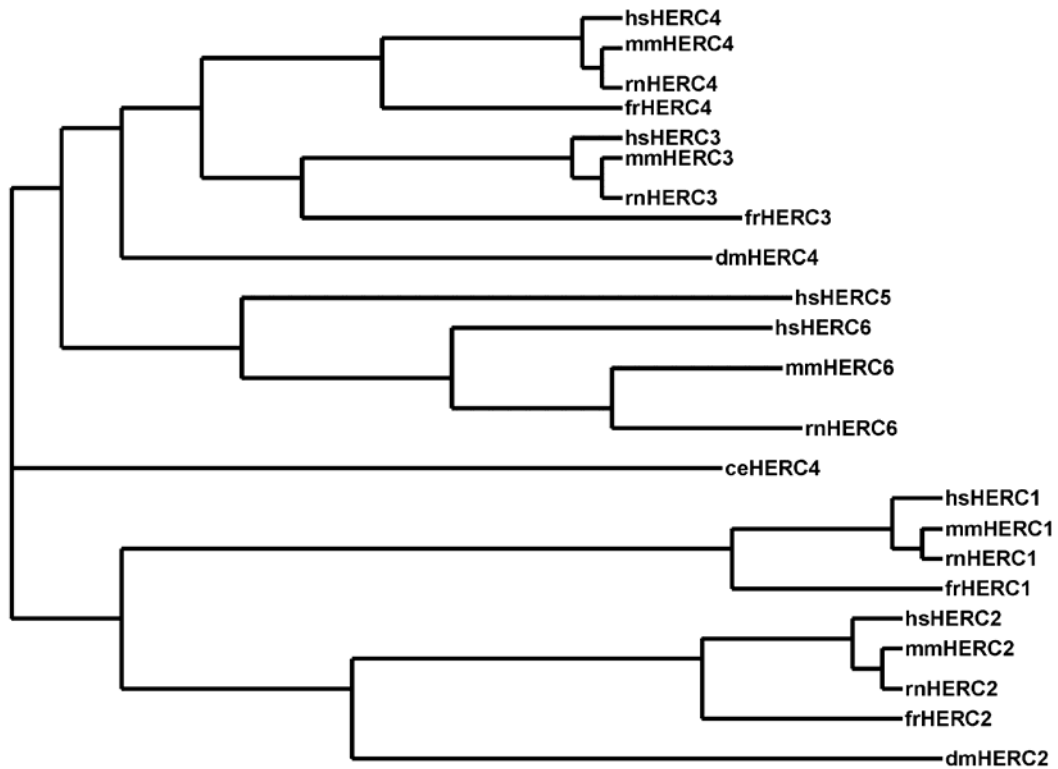


Figura 12. Árbol filogenético de las proteínas HERC (procedente de Hochrainer et al. (2005)). En el árbol se observa de nuevo la separación clara entre HERCs grandes (abajo) y pequeñas (arriba). Sin embargo, la proteína HERC4 de *C.elegans* (centro) surge directamente del tronco del árbol y por lo tanto se asemeja por igual a HERCs grandes y pequeñas. Aunque esto último ha llevado a sugerir que esta proteína debe asemejarse mucho a un hipotético ancestro común a todos los miembros de la familia, argumentos de carácter evolutivo hacen pensar que esto no es exactamente cierto (ver texto y figura 13). (hs: *Homo sapiens*; mm: *Mus musculus*; rn: *Rattus norvegicus*; fr: *Fugu rubripes*; dm: *Drosophila melanogaster*; ce: *Caenorhabditis elegans*).

durante la evolución temprana de los animales invertebrados mediante un evento de fusión génica. Sin embargo, la interpretación de Hochrainer et al. del árbol filogenético de los genes HERC pasa por alto un punto importante: a pesar de que *C. elegans* no posee ningún gen HERC grande, esto no significa que sus antepasados no lo poseyesen y, de hecho, un simple análisis de la filogenia animal permite llegar a esta conclusión (Figura 13). En efecto, los genes HERC grandes pueden hallarse en la actualidad en especies que van desde *Drosophila melanogaster*, un artrópodo, hasta el ser humano, un cordado. Ello significa que los antepasados comunes entre la mosca y el hombre ya

debieron tener genes HERC grandes. Cabe ahora recordar que los antepasados comunes de artrópodos (protóstomos) y cordados (deuteróstomos) vivieron muchos años antes que los últimos ancestros comunes entre artrópodos y nematodos (ambos protóstomos).

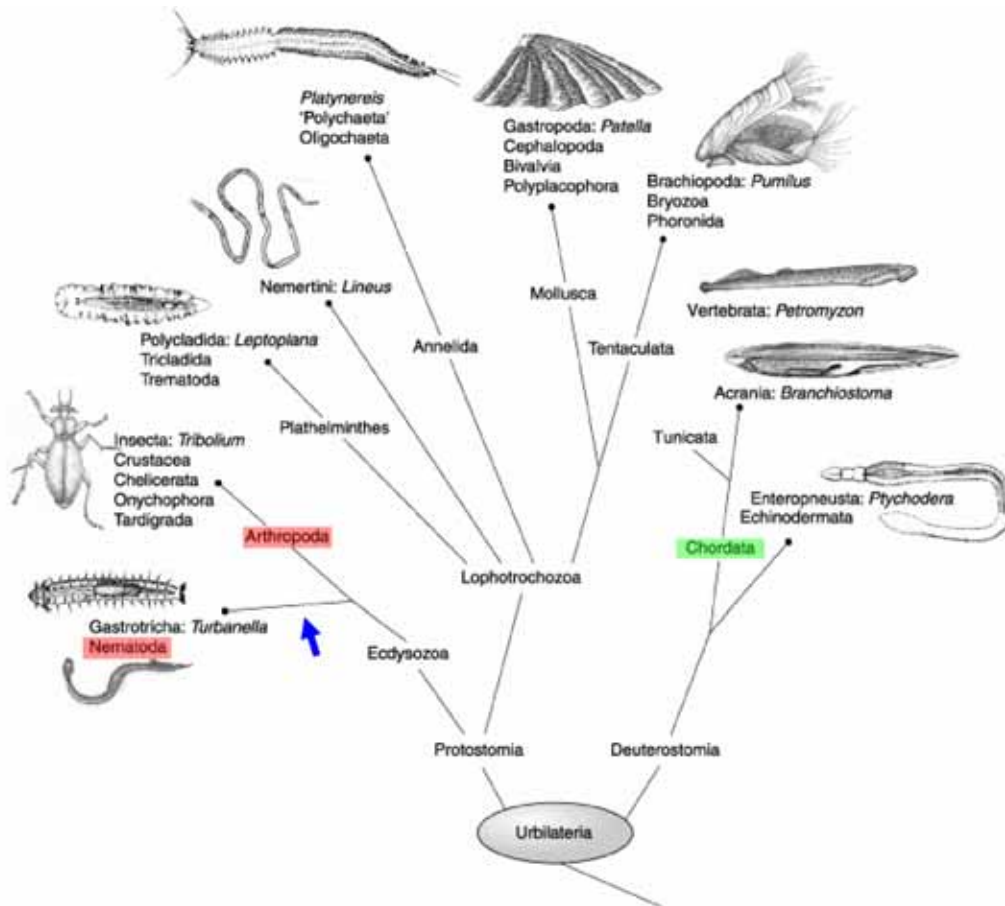


Figura 13. Orígenes de los genes HERC (imagen adaptada de Tessmar-Raible y Arendt (2003)). El hecho de que se hayan hallado HERCs grandes en los genomas de varias especies de artrópodos (hasta el momento se sabe de la existencia de ortólogos de HERC2 en tres insectos dípteros (*Drosophila melanogaster*, *Drosophila pseudoobscura* y *Anopheles gambiae*) y uno himenóptero (*Apis mellifera*)) así como en los genomas de cordados indica que el último ancestro común de ambos, es decir Urbilateria, ya debió poseer genes HERC grandes (lo mismo vale para los pequeños). En vista de esto, el hecho de que los genomas de los nematodos *Caenorhabditis elegans* y *Caenorhabditis briggsae* no posean ningún gen HERC grande (sólo poseen un gen HERC pequeño) parece indicar que HERC2 se perdió en algún momento posterior a la divergencia artrópodos-nematodos y anterior al último antepasado común entre *C.elegans* y *C.briggsae* (flecha).

Por tanto cabría esperar que *C. elegans* hubiese heredado de sus antepasados una copia de un gen HERC grande, tal y como la hemos heredado las moscas y nosotros mismos. El hecho, pues, de que *C. elegans* no contenga tal gen quiere decir que éste fue perdido secundariamente en algún punto de la rama evolutiva que condujo a *C. elegans*, y que esto ocurrió posteriormente a la divergencia evolutiva entre nematodos y artrópodos. Esta nueva interpretación corrige, pues, la formulada anteriormente y establece como

genes ancestrales de la familia no sólo a HERC4, sino también a HERC2 (*Drosophila* posee estos dos genes solamente). Queda todavía por averiguar cual de ambos genes fue el primero en aparecer y si ambos se formaron independientemente o bien uno procede del otro. En cualquier caso, no cabe duda de que los demás genes de la familia sí proceden de HERC4 y HERC2. En particular, HERC1 parece haber aparecido, a partir de HERC2, muy pronto en la evolución de los cordados, puesto que se encuentra en todos los miembros analizados del filo, incluyendo tetrápodos, peces y hasta un hemicordado, la ascidia *Ciona intestinalis*, pero no se ha hallado hasta la fecha en ningún invertebrado. El mecanismo por el cual HERC2 dio lugar a HERC1 probablemente esté relacionado con la duplicación genómica que tuvo lugar temprano en la evolución de los cordados. Dicho evento también dio lugar, con toda probabilidad, al gen HERC3, el cual se generó por duplicación del gen HERC4. La duplicación genómica explicaría por qué HERC3 y HERC4, del mismo modo que HERC1 y HERC2, se hallen en distintos cromosomas (HERC1 y HERC2 se hallan en humanos en el mismo cromosoma, el 15, pero en regiones sinténicas distintas). Debe remarcarse además que los genes HERC1 y HERC2 no son completamente homólogos en toda su secuencia, como puede observarse por la estructura en dominios de las proteínas: ello probablemente sea consecuencia de las diferencias acumuladas entre ambos genes posteriormente a la formación de HERC1. Por último, los genes HERC5 y HERC6 se formaron por duplicación génica (por contraposición a la duplicación del genoma entero mencionada antes) a partir de HERC3 durante la evolución de los vertebrados y todos ellos (HERC3, HERC5 y HERC6) se encuentran muy cerca en el mismo cromosoma (el 4 humano). HERC6 apareció antes y se encuentra en mamíferos pero no en peces, mientras que HERC5 constituye una adquisición evolutiva reciente, puesto que apareció durante la evolución de los mamíferos y se le encuentra en primates pero no en

roedores. A medida que se hagan públicos más genomas de más especies pertenecientes a los distintos taxones del reino animal se hará posible indagar más en los detalles de la historia evolutiva de la familia HERC y de este modo comprobar la validez de las hipótesis aquí sugeridas.