

DEPARTAMENT DE BIOQUÍMICA I BIOLOGIA MOLECULAR  
**UNIVERSITAT DE BARCELONA**

**EVOLUCIÓ DELS MECANISMES DE CONTROL  
DEL METABOLISME DEL GLICOGEN**

**DANIEL CIFUENTES BUIRA**

Barcelona, maig de 2006

Programa de Doctorat de Biotecnologia  
Bienni 2000-2002

## **C. RESULTATS**

... era per dir-te que la recerca de les lleis explicatives, en els fets naturals, procedeix d'una manera tortuosa. Davant d'alguns fets inexplicables has de fer per manera d'imaginar moltes lleis generals, tot i que encara no vegis el lligam que tenen amb els fets de que t'ocupes.

Quart dia, vespres.

Umberto Eco, "*El nom de la rosa*"

**CAPÍTOL 1**  
**EVOLUCIÓ DE LA GLICOGEN SINTASA**



## CAPÍTOL 1: EVOLUCIÓ DE LA GLICOGEN SINTASA

### 1.1 INTRODUCCIÓ

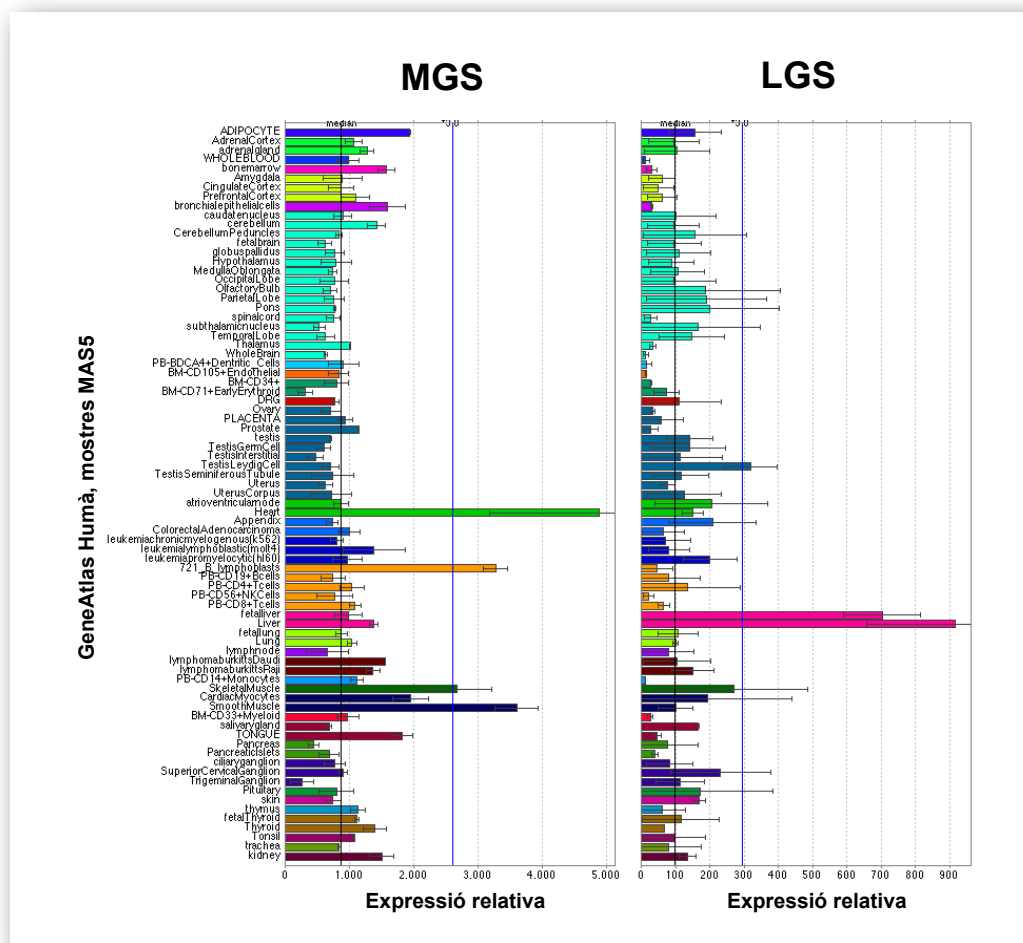
La glicogen sintasa és un enzim que catalitza la formació d'enllaços  $\alpha(1-4)$  entre el residu glucosil de l'UDP-glucosa i l'extrem reductor una cadena de glicogen preexistent. En humans trobem dos isoenzims de glicogen sintasa, l'isoforma específica de fetge (GYS2, LGS) i l'isoforma muscular (GYS1, MGS) (taula 1.1). Aquestes dues isoformes s'expressen a partir de dos gens diferents i mantenen un 67% d'homologia a nivell de la seqüència d'aminoàcids, amb les majors diferències concentrades a l'extrem C-terminal de la proteïna i coincidint amb l'últim exó.

Fins al moment no s'ha descrit cap isoforma per ús d'exons alternatius en cap dels dos gens, la qual cosa és molt significativa tenint en compte la seva longitud i que ambdós gens s'estructuren en 16 exons.

**Taula 1.1** Isoformes de glicogen sintasa humanes.

	GS Hepàtica	GS Muscular
Gen	<i>GYS2</i>	<i>GYS1</i>
Cromosoma	<i>12p12.2</i>	<i>19q13.3</i>
Nº d'exons	<i>16</i>	<i>16</i>
Nº d'aminoàcids	<i>702</i>	<i>737</i>
Teixits	<i>Fetge</i>	<i>Múscul esquelètic, cardíac i llis, cervell, ronyó, adipós</i>
Associacions clíniques	<i>GSD0</i>	<i>Possible Diabetis Tipus II, síndrome d'ovari policístic, hipertensió.</i>

El que destaca de les dues isoformes de glicogen sintasa però, és el seu patró d'expressió. Mentre que fins ara només s'ha detectat la presència de l'LGS en fetge, l'MGS és un enzim ubic que s'expressa virtualment en qualsevol teixit excepte el fetge, sobretot en el múscul esquelètic i cardíac, d'on pren el nom d'isoforma muscular (figura 1.1).

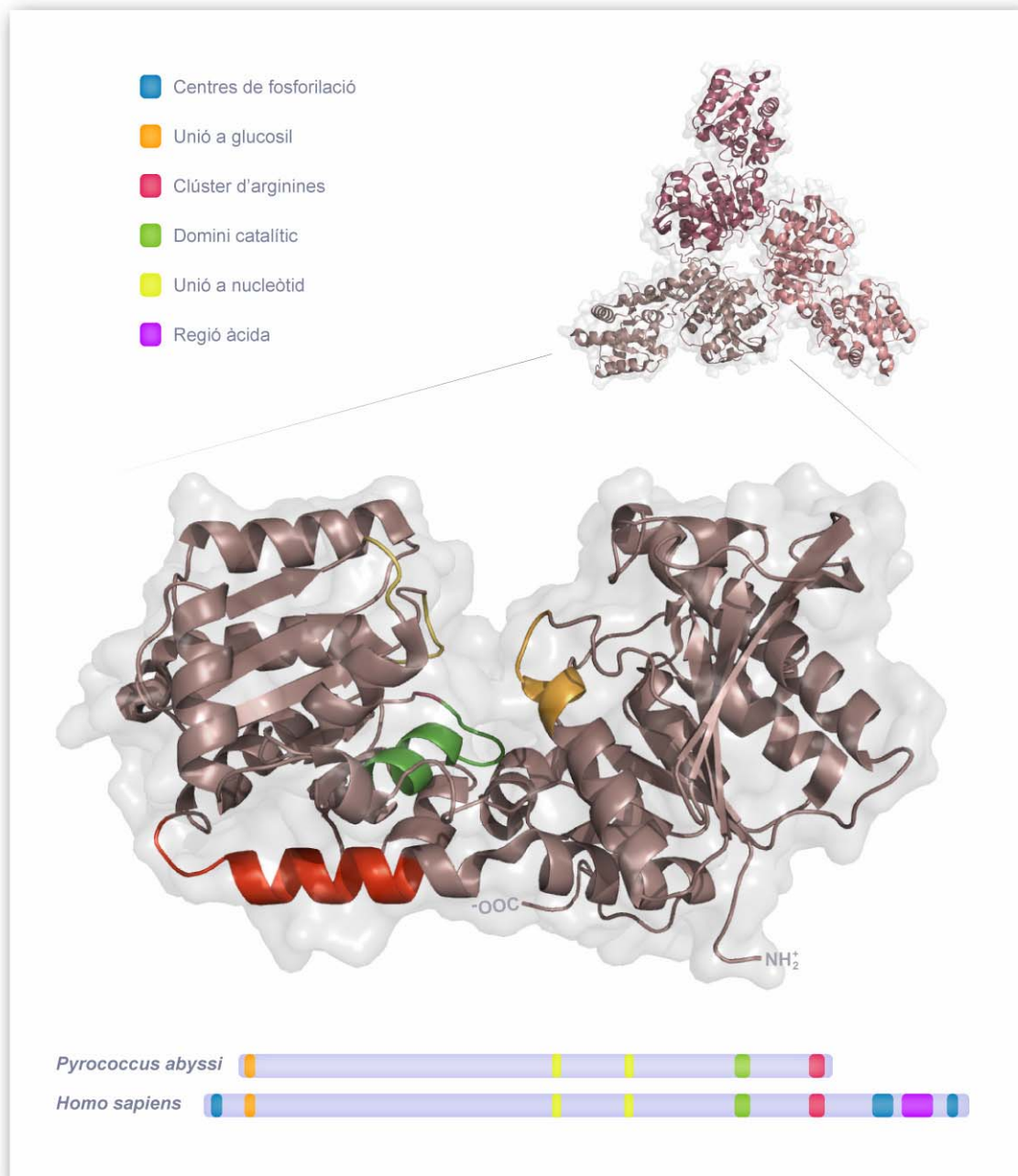


**Figura 1.1** Patró d'expressió de l'mRNA de l'MGS i l'LGS en teixits humans. L'mRNA de 79 teixits humans es va hibridar en un xip GNF1H (Su et al., 2004) d'alta densitat d'oligonucleòtids. La figura ens mostra els nivells relatius d'expressió de l'MGS i l'LGS. La línia negra indica la mitjana d'expressió del gen en qüestió en tots els teixits, i la línia blava indica el llindar situat tres cops per sobre la mitjana. Cal notar la diferència d'escala entre les gràfiques de MGS i LGS, així l'LGS només té expressió significativa en fetge. La detecció de MGS en fetge no es pot descartar que sigui un artefacte a l'hora de purificar mRNA, donat que el fetge és un teixit complex que no només està format per hepatòcits sinó també per altres tipus cel·lulars com les cèl·lules estrellades i teixit epitelial.

Les GS humanes pertanyen a la família de les glicosiltransferases GT-3 (Coutinho et al., 2003) on també trobem les GS d'animals i fongs. S'estructuren en dos dominis amb un plegament Rossman del tipus  $\beta\alpha\beta$ . Es caracteritzen per utilitzar UDP-glucosa com a donador de residus glucosil i la seva activitat es regula per fosforilació reversible i per efectors al·lostèrics, principalment G6P.

Les glicosiltransferases de la família GT-3 estan relacionades amb les GS de bacteris i les midó sintases de plantes, englobades en la família GT-5. Aquestes utilitzen exclusivament ADP-glucosa com a font de residus glucosil i la seva activitat no està regulada.

Recentment s'ha resolt l'estructura de la GS de *Pyrococcus abyssi* (Horcajada et al., 2003; Horcajada et al., 2006), que representa l'enzim més petit de la superfamília de les glicogen (midó) sintases (GT-3 i GT-5) i comparteix característiques estructurals amb els membres d'ambdues famílies (figura 1.2) (Cid et al., 2002).



**Figura 1.2 Estructura de la GS de *Pyrococcus abyssi*.** Representació esquemàtica de l'estructura terciària i quaternària de la GS de *Pyrococcus abyssi* on es mostra l'estat d'oligomerització de la proteïna i es destaquen els dominis implicats en la funció de la proteïna. També s'indica la disposició relativa d'aquests dominis al llarg de la seqüència primària en comparació amb la GS muscular humana.



## 1.2 RESULTATS

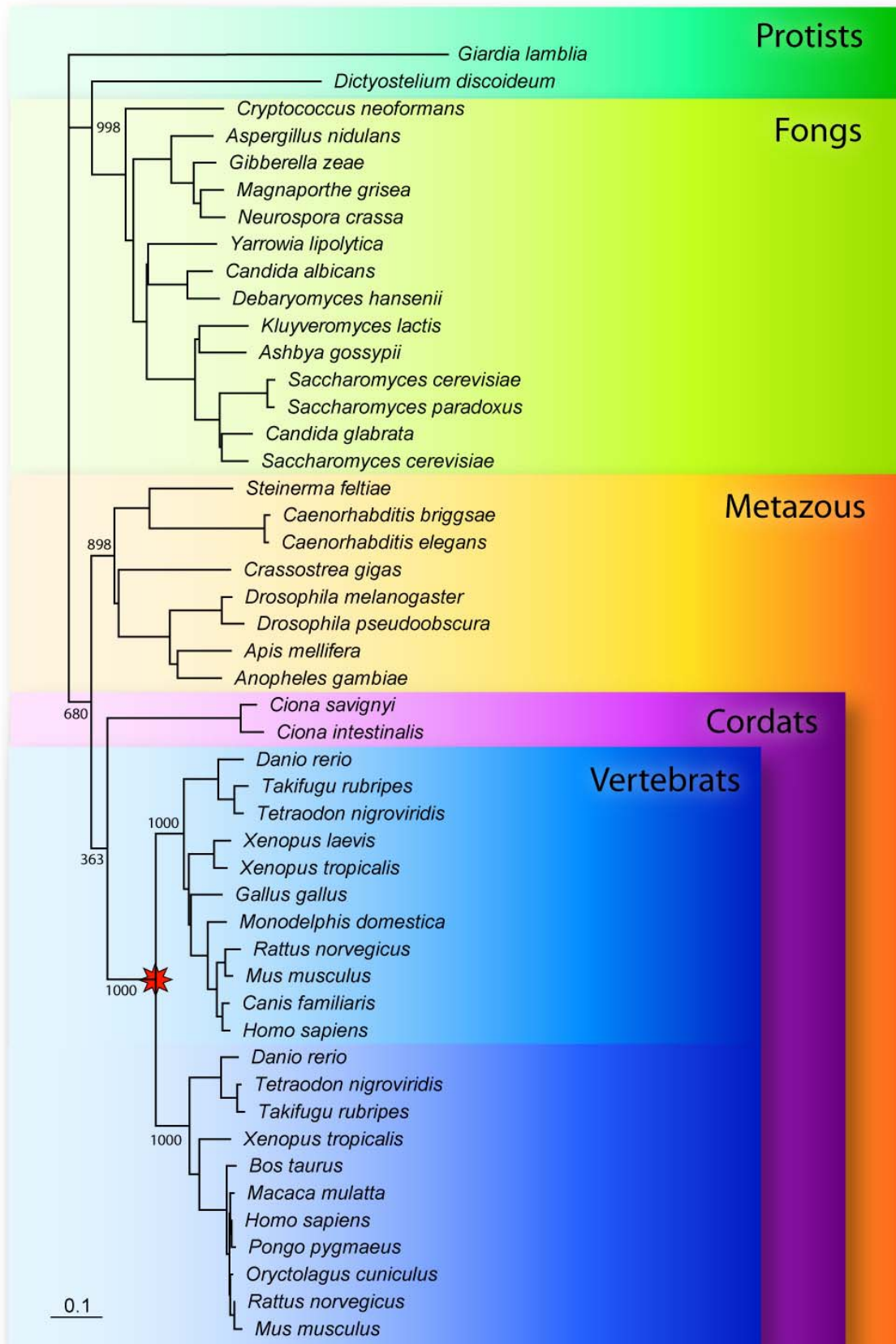
### 1.2.1 Filogènia de la glicogen sintasa.

A fi de comprendre l'evolució de les glicogen sintases eucariotes i els seus mecanismes de regulació, vàrem dur a terme anàlisis de seqüències i estudis filogenètics. En el moment de començar aquest treball només 10 seqüències de GS eucariotes (Cid et al., 2002) estaven disponibles en les bases de dades (GenBank). Aquest nombre era insuficient per a poder realitzar estudis evolutius en profunditat perquè hi havia branques del regne eucariota que no tenien cap representant. Per això calia buscar i identificar les seqüències de GS d'organismes representatius que ens permetessin obtenir la visió de conjunt de l'evolució de la GS en els eucariotes.

Mitjançant cerques amb el programa BLAST (Altschul et al., 1997) en els projectes genoma dels diferents organismes (Hillier et al., 2004; Jaillon et al., 2004; Lander et al., 2001; Lindblad-Toh et al., 2005; Waterston et al., 2002) i amb la posterior edició vam obtenir una base de dades amb més de cinquanta seqüències completes de GS, representatives de totes les branques del regne eucariota (veure Apèndix).

Els gens candidats a codificar per a la glicogen sintasa que torbàvem seguint aquest procediment eren confirmats mitjançant alineaments de seqüència amb les GS humanes i de llevats que estaven disponibles en el GenBank. La presència de regions d'homologia amb les seqüències consens del centre catalític o del clúster d'arginines en les seqüències candidates eren proves necessàries per a validar-les com a GS.

Així vam trobar seqüències que codificaven per a la GS en cada genoma eucariota disponible, amb l'excepció de les plantes (que acumulen midó enlloc de glicogen) i la dels organismes paràsits obligats (Henrissat et al., 2002). Però només en els vertebrats érem capaços d'identificar dos possibles gens de GS per genoma. Tota la resta d'eucariotes, des de els Diplomonads fins als Urocordats posseeixen una sola còpia del gen. La còpia extra que trobem en *Saccharomyces cerevisiae* no està present en el seu ancestre *Ashbya gossypii* i s'origina molt probablement en un fenomen independent de duplicació de tot el genoma exclusiu d'aquesta família (Dietrich et al., 2004).

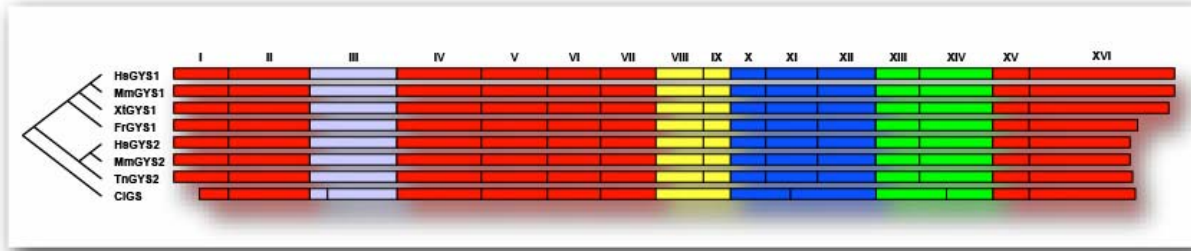


**Figura 1.3 Filogènia de les glicogen sintases eucariotes.** Arbre *Neighbor-joining* de les glicogen sintases eucariotes calculat a partir d'una matriu de distàncies. La base de l'arbre l'hem situat en la *Giardia lamblia*, l'eucariota més primitiu del qual es disposa la seqüència. El valors de *bootstrap* s'indiquen en els nodes de les branques. L'estrella vermella indica el punt de duplicació gènica.

A partir dels alineaments de les seqüències vam calcular una matriu de distàncies que ens va permetre inferir les relacions filogenètiques de les GS en un arbre de *Neighbor-joining* (figura 1.3). L'arbre consens ens mostra una evolució lineal de la GS fins arribar als vertebrats, on aquest clade es divideix en dues branques independents, que ens separen les isoformes de fetge i de múscul. Això ens indica que el gen de la glicogen sintasa es va duplicar en el moment de l'aparició dels vertebrats, donant origen als gens paràlegs GYS1 i GYS2, corresponents a les dues isoformes de glicogen sintasa.

Cal remarcar que aquesta duplicació gènica no la observem en els ascidis *Ciona intestinalis* i *Ciona savignyi*, que pertanyen al grup dels urocordats i es consideren el grup més proper als vertebrats (Canestro et al., 2003; Delsuc et al., 2006). Molts autors sostenen que l'aparició dels vertebrats va anar acompanyada d'una sèrie de duplicacions genòmiques completes (Abi-Rached et al., 2002; Amores et al., 1998; Jaillon et al., 2004) que van originar moltes parelles de gens duplicats, malgrat que amb el temps l'evolució només va mantenir una de les còpies de la majoria de parelles. Com que aquests fenòmens duplicatius es van produir després de l'aparició dels Urocordats i en vertebrats trobem dos gens de glicogen sintasa, GYS1 i GYS2 podrien representar una de les parelles supervivents d'aquelles duplicacions genòmiques en les que l'evolució no n'ha eliminat cap membre.

La comparació de l'estructura gènica dels dos gens de GS de vertebrats respecte a la GS de l'urocordat *Ciona intestinalis* (CiGS) ens va proporcionar una prova addicional de l'origen de les dues isoformes de GS específiques de teixit. En efecte, l'organització en introns i exons del gen de la CiGS respecte a un conjunt representatiu de les isoformes de fetge i múscul de vertebrats indica que la CiGS comparteix una estructura gènica similar amb les GS de vertebrats (figura 1.4) i que tots els empalmaments entre exons es localitzen en la mateixa posició tant en les isoformes hepàtiques com en les musculars. Això suggereix que les isoformes de glicogen sintasa de fetge i múscul provenen de la duplicació d'un mateix ancestre comú.



**Figura 1.4 Estructura exònica de les GS de cordats.** Representació esquemàtica dels gens de les GS de cordats, on els exons es representen per blocs de tamany proporcional a la seva longitud. L'arbre filogenètic ens indica les relacions entre les GS representades. Els colors destaquen els exons que difereixen entre ciona i els vertebrats.

Tot i que les GS de les ciones es poden considerar properes a l'ancestre comú de les glicogen sintases musculars i hepàtiques de vertebrats, hi ha més percentatge d'identitat entre les isoformes de múscul i fetge (67% de promig) que entre el conjunt de les GS de vertebrat comparades amb les GS de les ciones (54%), tal com ens mostra la taula 1.2.

	GS <i>Ciona intestinalis</i>	GS <i>Ciona savignyi</i>	GYS2 <i>Tetraodon nigrovidiris</i>	GYS2 <i>Danio rerio</i>	GYS2 <i>Xenopus tropicalis</i>	GYS2 <i>Rattus norvegicus</i>	GYS2 <i>Mus musculus</i>	GYS2 <i>Canis familiaris</i>	GYS1 <i>Homo sapiens</i>	GYS1 <i>Tetraodon nigrovidiris</i>	GYS1 <i>Danio rerio</i>	GYS1 <i>Xenopus tropicalis</i>	GYS1 <i>Rattus norvegicus</i>	GYS1 <i>Mus musculus</i>	GYS1 <i>Canis familiaris</i>	GYS1 <i>Homo sapiens</i>
-	92	53	54	55	54	55	56	55	54	52	52	52	52	52	52	52
-	54	55	56	55	55	56	55	55	56	53	53	53	53	53	53	53
-		86	79	77	78	79	79	68	69	67	66	65	65	65	65	65
-			79	78	79	80	80	70	71	68	67	66	66	66	66	66
-				81	82	84	83	70	72	70	68	68	66	68	66	68
-					96	92	93	67	69	67	66	65	64	65	64	65
-						94	94	68	70	68	66	66	64	66	64	66
-							96	69	71	68	68	66	65	67	65	67
-								69	71	68	68	66	65	67	65	67
-									89	77	76	76	74	77	74	77
-										77	78	76	75	77	75	77
-											81	81	78	81	78	81
-												98	95	97	95	97
-													94	96	94	96
-														95	95	95
-															95	95

**Identitat global:**  
■ GYS1= 83 %  
■ GYS2= 84 %  
■ GYS1 vs. GYS2=67 %

**Taula 1.2 Percentatge d'identitat entre les GS de cordats.** A partir de l'alineament de les GS de cordats calculem els percentatges d'identitat entre proteïnes. En negre es destaquen les comparacions entre les GS d'ascidis i les de vertebrats, en verd les comparacions entre GS hepàtiques i musculars. En vermell i blau identifiquem les comparacions entre les isoformes del mateix tipus.

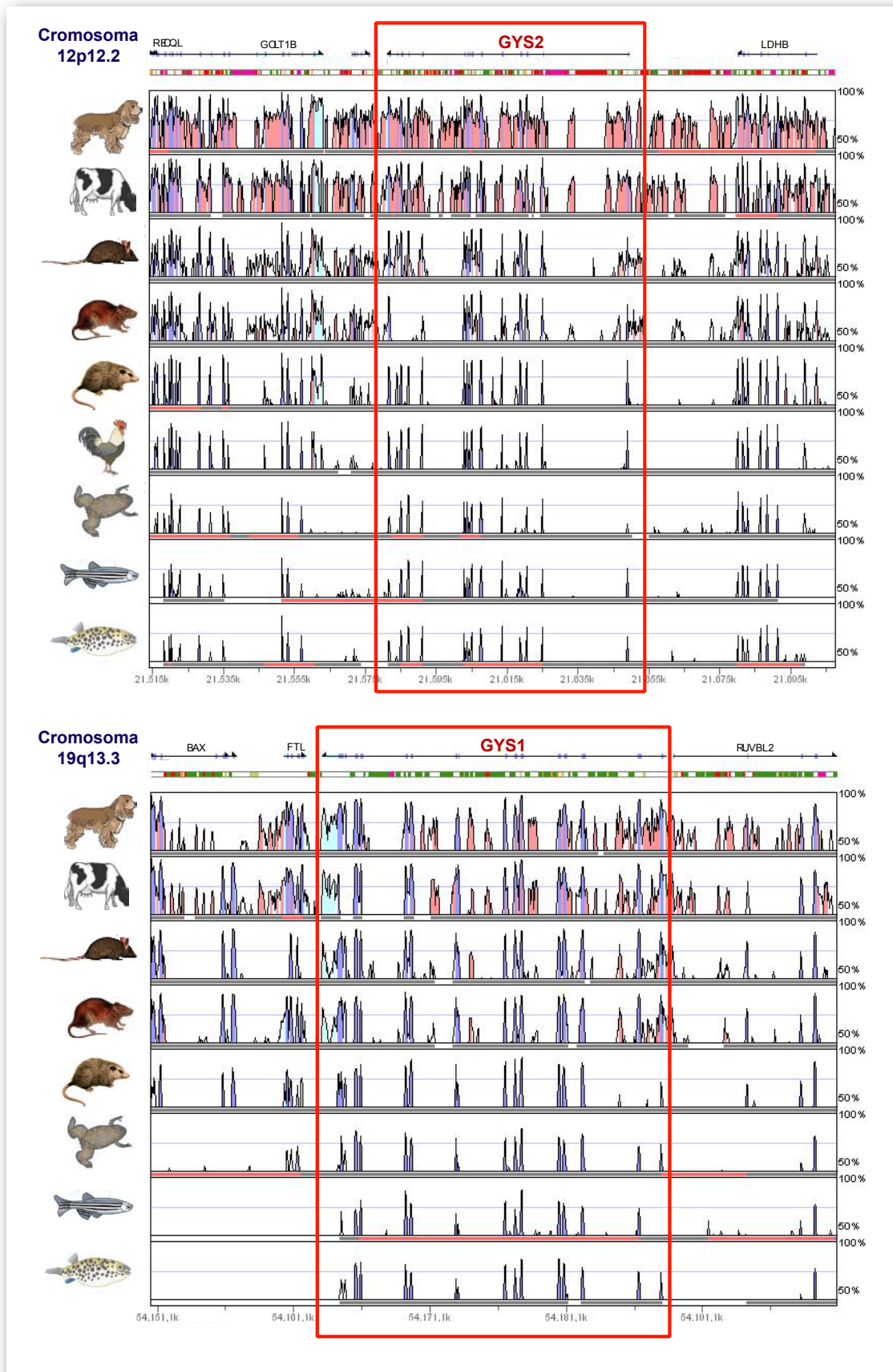
A continuació vam estudiar el grau de sintènia dels locus genètics de l'LGS i MGS en els vertebrats, és a dir, la conservació de la seqüència en les regions dels cromosomes on es troben les isoformes de GS. Això també implica estudiar la posició relativa dels gens disposats al seu entorn en aquests trams de seqüència. Per a determinar la sintènia vam utilitzar els alineaments dels genomes d'organismes representatius, prenent com a referència el genoma humà.

La representació esquemàtica dels alineaments genòmics generats per l'interfaç *Vista* (Couronne et al., 2003) (figura 1.5) ens mostra les regions cromosòmiques corresponents a l'LGS i l'MGS amb els seus introns i exons, així com els gens flanquejants. Els pics ens indiquen la posició dels màxims d'identitat entre les seqüències del genoma humà i de l'organisme indicat.

A partir d'aquests estudis podem deduir que el grau de sintènia en vertebrats és molt elevat tant per al locus de l'LGS com el de l'MGS perquè des de peixos fins a humans es conserven els trams de seqüència codificants (els exons), així com també es manté la posició relativa dels gens de LGS o MGS respecte als gens del seu entorn. Per exemple: en vertebrats l'LGS sempre la trobem al costat de la lactat deshidrogenasa B (isoforma H) i l'MGS es situa molt a prop de la recombinasa RUVBL2.

També cal destacar la informació addicional que podem extreure d'aquest tipus d'alineament, en referència als trams de seqüència no codificants conservats (SNC). Davant l'allau d'informació genòmica actual, molts autors es sorprenen de l'alt grau de conservació que presenten algunes regions no codificants. Per justificar la pressió selectiva que mantindria aquestes regions amb poc canvis postulen que es podria tractar d'elements amb algun paper en la regulació transcripcional (Nobrega et al., 2004) i això explicaria el fet de que es trobessin conservades malgrat no codificar per cap proteïna ni estar en regions promotores.

En els gens de l'LGS i l'MGS podem trobar regions SNC, marcades en rosa. En aquests trams es podria trobar la clau per entendre fenòmens com l'especificitat de teixit de les isoformes de GS o l'augment de MGS en múscul després d'un període d'entrenament.



**Figura 1.5 Sintènia dels locus genètics de les glicogen sintasa hepàtiques i musculars.** La figura mostra els alineaments de les regions cromosòmiques corresponents a les LGS i MGS humanes respecte a les mateixes regions en altres genomes de vertebrats. Els punts de màxima identitat i conservació entre seqüències s'indiquen per pics, en blau els que corresponen a exons, en rosa els que es troben en regions no codificants. Els organismes representats són: gos, vaca, ratolí, rata, opossum, gallina, xenopus, peix zebra i peix globus.

Aquestes regions guanyen més pes com a seqüències candidates a contenir elements reguladors si tenim present que el promotor de l'MGS humana (i de la resta de mamífers estudiats) tan sols té uns 750 parells de bases de longitud perquè es troba solapat amb el del gen adjacent RUVBL2. Una regió genòmica tan petita probablement és un espai insuficient per contenir tots els elements d'unió dels factors de transcripció necessaris per controlar una expressió tan complexa.

### **1.2.2 Dominis de les glicogen sintases eucariotes**

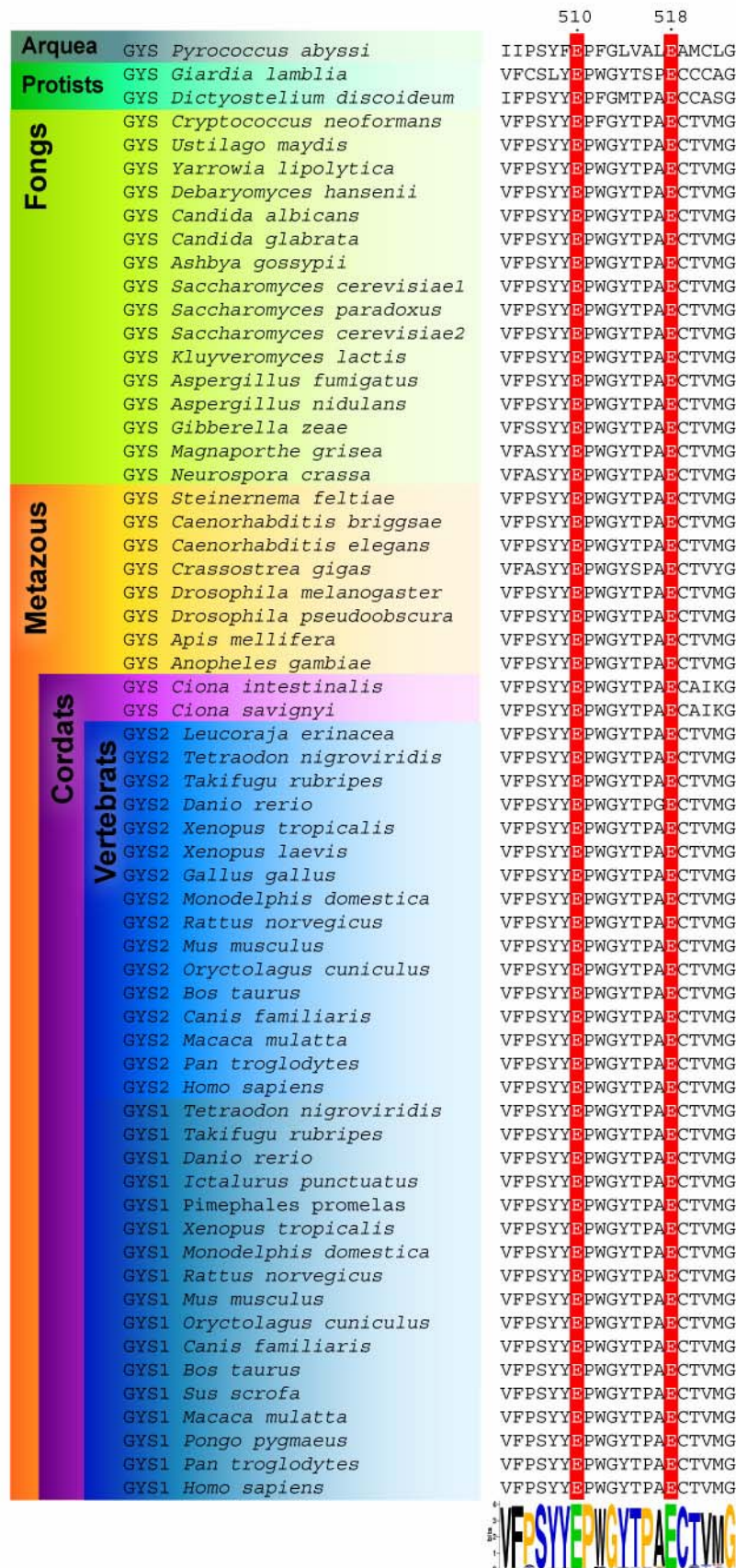
Els alineaments de les seqüències de GS recopilades es converteixen en una eina molt valuosa per a estudiar l'evolució dels diferents dominis catalítics i reguladors de la GS.

#### **1.2.2.1 Aminoàcids importants per a la catàlisi.**

Els residus importants per a la catàlisi en la glicogen sintasa van ser identificats per Cid i col·laboradors (Cid et al., 2002; Cid et al., 2000) mitjançant la comparació de seqüències de bacteris, arquees i eucariotes. Van descriure una regió molt conservada amb la seqüència consens EX<sub>7</sub>E, que també trobaven en altres famílies de glicosil transferases. Aquests glutàmics es situaven en les posicions 510 i 518 de l'MGS humana.

La prova que validava aquests residus com a components essencials del centre catalític era que quan es substituïen per alanina, el mutant E518A mantenia un 6% de l'activitat respecte l'enzim salvatge i el mutant E510A perdia totalment l'activitat malgrat que s'expressava i es plegava correctament.

Tot i que el mecanisme catalític de la glicogen sintasa no es coneix en detall, aquests experiments feien suposar que el residu E510 actua com a nucleòfil i l'E518 com catalitzador general àcid/base.



**Figura 1.6 Alineament de la regió del centre catalític de les glicogen sintases.** Per a emfatitzar el grau de conservació del domini s’ha inclòs una GS d’arquea en l’alineament de les GS eucariotes. Els glutàmics essencials per a la catàlisi estan ressaltats. La numeració de la seqüència correspon a les posicions que els residus marcats ocupen en l’MGS humana. L’alçada dels aminoàcids en la seqüència consens és equivalent a la seva freqüència.



Vam calcular l'alineament dels residus importants per a la catàlisi de les GS eucariotes a partir de les seqüències completes però en aquest cas també vàrem incloure espècies per a les quals només es tenien fragments parcials de la seqüència però que cobrien el domini que volíem estudiar.

El nostre alineament de la regió catalítica (figura 1.6) confirma que els glutàmics equivalents als E510 i E518 estan absolutament conservats en totes les GS eucariotes identificades, a més de retenir un alt grau de conservació en el seu entorn.

### 1.2.2.2 Clúster d'arginines

Aquesta regió ha estat descrita en GS de llevats i humans com a responsable de la transducció del senyal d'activació al·lostèrica que desencadena la G6P (Hanashiro & Roach, 2002; Pederson et al., 2000). Juntament amb la regulació covalent per fosforilació, el clúster d'arginines contribueix a la regulació de l'activitat de la GS. El motiu del clúster consisteix en sis arginines agrupades en un tram molt curt de seqüència segons el consens **RRxRxxxRxRxxR**.

L'alineament d'aquesta regió (figura 1.7) ens mostra l'importància d'aquestes arginines donat que es troben conservades en tots els eucariotes, amb la notable excepció dels protistes *Giardia lamblia* i *Dyctiostelium discoideum*. Ni les GS d'arquees, representades en l'alineament per la GS de *Pyrococcus abyssi*, ni les GS de protistes presenten la seqüència estricta d'arginines consens malgrat que mantenen una certa homologia en aquest tram amb les GS d'eucariotes superiors.

Podem entendre el possible origen d'aquest domini si tenim en compte que aquesta regió no sorgeix *de novo* per incorporació d'un nou tram de seqüència a les GS més evolucionades, sinó que prové d'una regió que ja trobàvem present en GS procariotes i que paulatinament ha anat incorporant més arginines fins assolir la configuració definitiva en els llevats i que es mantindrà fins als mamífers.

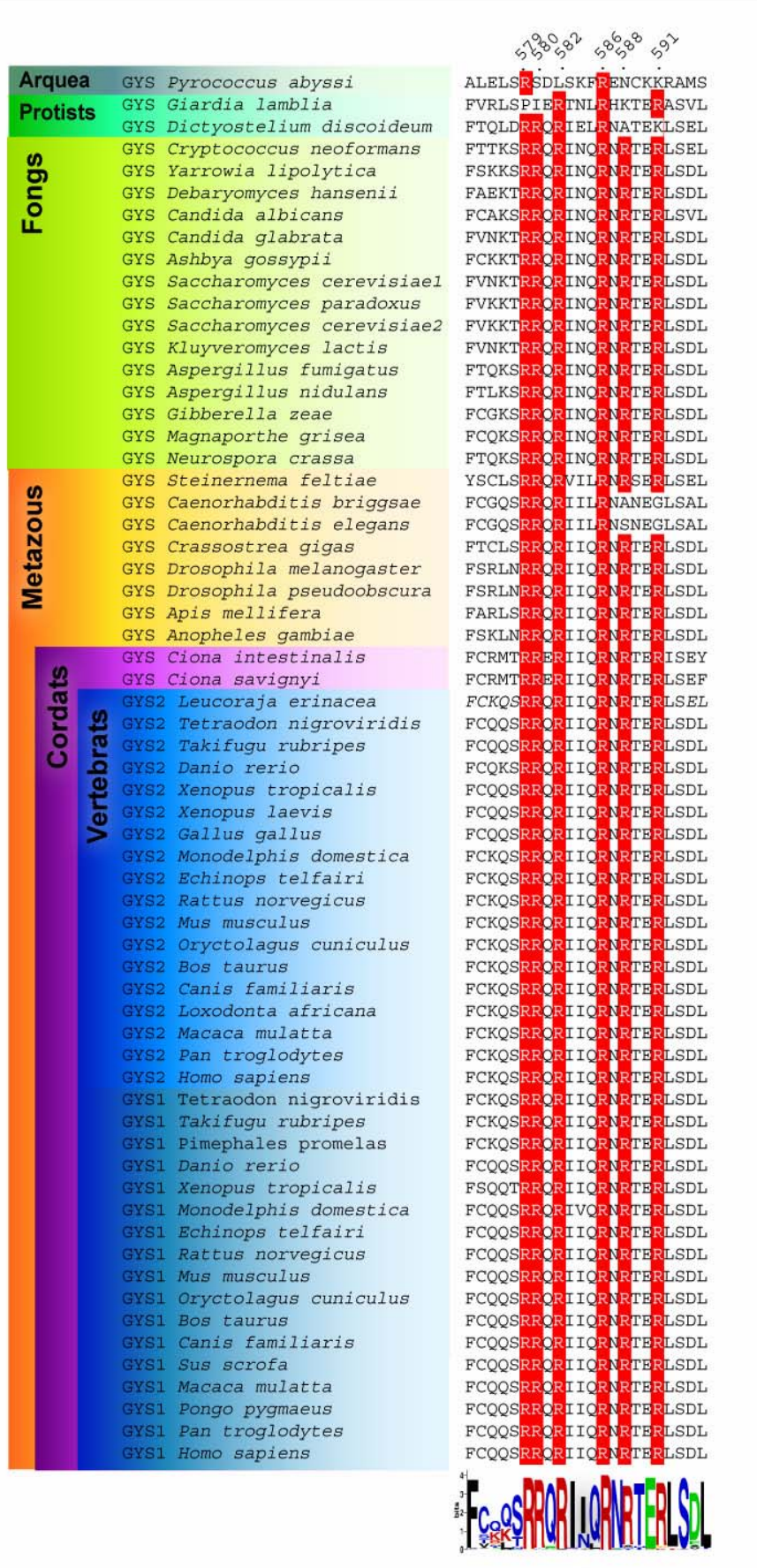


Figura 1.7 Alineament del clúster d'arginines. Es ressalten les sis arginines conservades que configuren la seqüència consens. La numeració correspon a les posicions que els residus marcats ocupen en l'MGS humana.

Donat que les GS de *Caenorhabditis* no tenen la seqüència canònica, podem suposar que hi ha una particularitat en la regulació al·lostèrica de la GS, específica del clade dels nemàtodes.

A fi de comprovar el paper regulador del clúster d'arginines (RR), vam determinar en cèl·lules COS-1 (línia fibroblasts immortals de *Macaca mulatta* fàcils de transfectar) les activitats específiques de la GS de *Giardia lamblia* (G/GS) i dels mutants de l'MGS humana (HsMGS) als quals se'ls havia substituït part o la totalitat de les arginines del clúster per alanines.

Vam treballar amb tres mutants de HsMGS. Al mutant R1 li havíem mutat les tres primeres arginines del clúster a alanina (R579A, R580A, R582A), al mutant R2 li mutàrem les tres darreres (R586A R588A i R591). El tercer mutant, R1R2 tenia totes sis arginines mutades a alanina.

La funcionalitat del domini RR, i per tant la magnitud de l'activació al·lostèrica per G6P, l'expressa la relació d'activitats I/T on I és l'activitat glicogen sintasa de la mostra en absència de G6P (Independent de l'activador) i T és l'activitat de la mostra amb l'activador (activitat Total).

Les activitats i les relacions I/T de la taula 1.3 ens mostren com la G6P té menys efecte sobre les activitats de les GS que no tenen el clúster canònic d'arginines que sobre les que mantenen la integritat del domini. Això és vàlid tant per les GS a les quals les hi hem alterat la seqüència consens per mutagènesi dirigida com per a la G/GS que de forma natural no presenta totes les arginines.

**Taula 1.3**

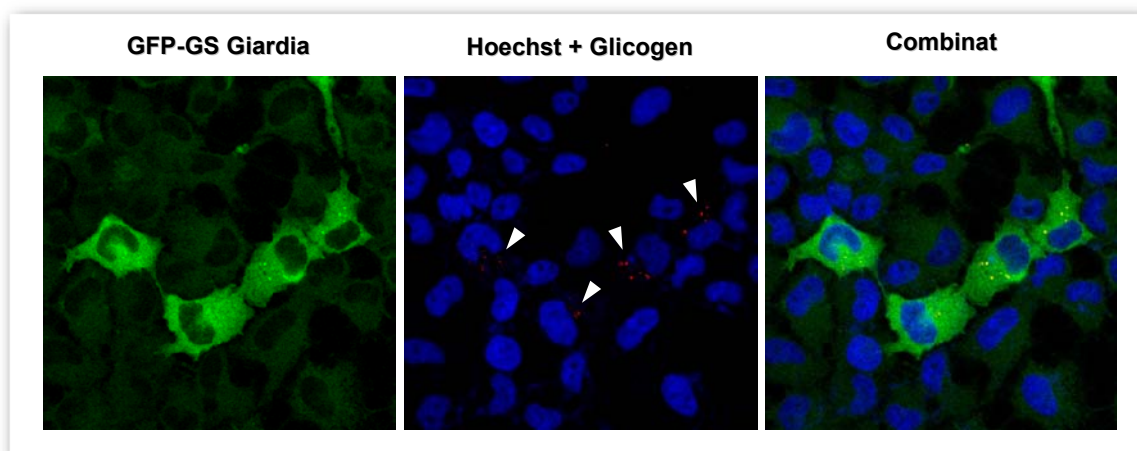
Construcció	Activitat GS (mU/mg proteïna)		Ratio d'activitat I/T (-/+ G6P)
	- G6P	+ G6P	
GFP-HsMGS	20,7	51,3	0,4
GFP-HsMGS R1	22,8	34,8	0,6
GFP-HsMGS R2	60,0	52,9	1,1
GFP-HsMGS R1R2	32,4	36,1	0,9
GFP-G/GS	55,7	73,7	0,8

Podria ser que l'estat de fosforilació variés entre els mutants estudiats i l'MGS salvatge i per això observariem un canvi en la relació I/T. Però com tots els assajos d'activitat s'ha fet a partir de cèl·lules en les mateixes condicions, qualsevol canvi en el patró de fosforilació també es podria atribuir a un efecte col·lateral induït per la mutació del clúster d'arginines.

La semblança entre el valor de les relacions I/T del mutant R1R2 i la G/GS ens reforça la idea que la G/GS no està regulada al·lostèricament per G6P, igual que les GS procariotes i les midó sintases de plantes i ens indiquen que la funció del clúster d'arginines és una innovació de les GS d'eucariotes més evolucionades.

Si la G/GS no es regula al·lostèricament, un dels possibles punts de control de la seva activitat catalítica pot ser la disponibilitat de substrats, com en el cas de les glicogen (midó) sintases de bacteris i plantes (Ball & Morell, 2003). Per avaluar aquesta hipòtesi estudiem la síntesi de glicogen de la G/GS en les cèl·lules FTO2B. En aquestes cèl·lules la GS hepàtica de rata no pot sintetitzar glicogen malgrat la presència constant del substrat UDP-glucosa, a menys que es sobreexpressi la GK i sintetitzi G6P per activar-la.

Vam transfectar cèl·lules FTO2B amb la construcció de la G/GS fusionada a GFP. Quan incubàvem amb glucosa, només les cèl·lules que expressaven la G/GS eren capaces d'iniciar la síntesi de glicogen (figura 1.8)



**Figura 1.8 G/GS-GFP en FTO2B.** La localització de la G/GS es segueix per la fluorescència de la GFP. La síntesi de glicogen es detecta mitjançant la immunodetecció amb un anticòs específic. Les puntes de fletxa indiquen llocs d'acumulació de glicogen.

Els resultats de síntesi de glicogen, combinats amb les dades dels alineaments de seqüència i la poca variació de l'activitat en presència o absència de G6P indiquen que la GS de *Giardia lamblia* no es regula al·lostèricament i que el punt de control es deu situar en una etapa anterior de la via, com per exemple la síntesi d'UDP-glucosa.

Aquestes característiques defineixen a la GS de *Giardia lamblia* com la baula entre les GS de procariotes i les GS d'eucariotes superiors.

### 1.2.2.3 Centres de fosforilació.

La glicogen sintasa va ser la primera proteïna descrita que s'inactivava per fosforilació (Friedman & Larner, 1962; Roach & Larner, 1976; Roach et al., 1976). La inactivació de les glicogen sintases per fosforilació seqüencial en serines ha estat una adquisició evolutiva de les GS eucariotes. S'han descrit fins a nou llocs de fosforilació en l'MGS de conill, malgrat que hi ha evidències que en poden ser molts més (Cid, 2002). Aquests centres de fosforilació estan concentrats en quatre blocs dels extrems N- i C-terminal de la proteïna, precisament en segments de proteïna que no estaven presents en les GS procariotes o en les GS eucariotes més primitives com la *G/GS*. Malgrat els esforços per establir homologies amb altres proteïnes fosforilades, no hem pogut aclarir l'origen d'aquests nous dominis.

Vàrem alinear les seqüències de GS eucariotes per la zona C-terminal on es concentren la majoria del centres de fosforilació (figura 1.9). L'alineament ens indica que aquests centres no han aparegut tots de cop, sinó que s'han incorporat a la seqüència de forma paulatina. L'antiguitat i grau de conservació d'alguns centres posa en relleu la jerarquia en l'ordre de fosforilació que han esmentat altres autors (Roach, 1990; Roach et al., 1991). Els centres **3a** i **3b** ja estan presents des de *Dictyostelium* i es mantenen en totes les GS eucariotes, mentre que els llocs de fosforilació **3c**, **4** i **5** només apareixen a partir dels metazous.

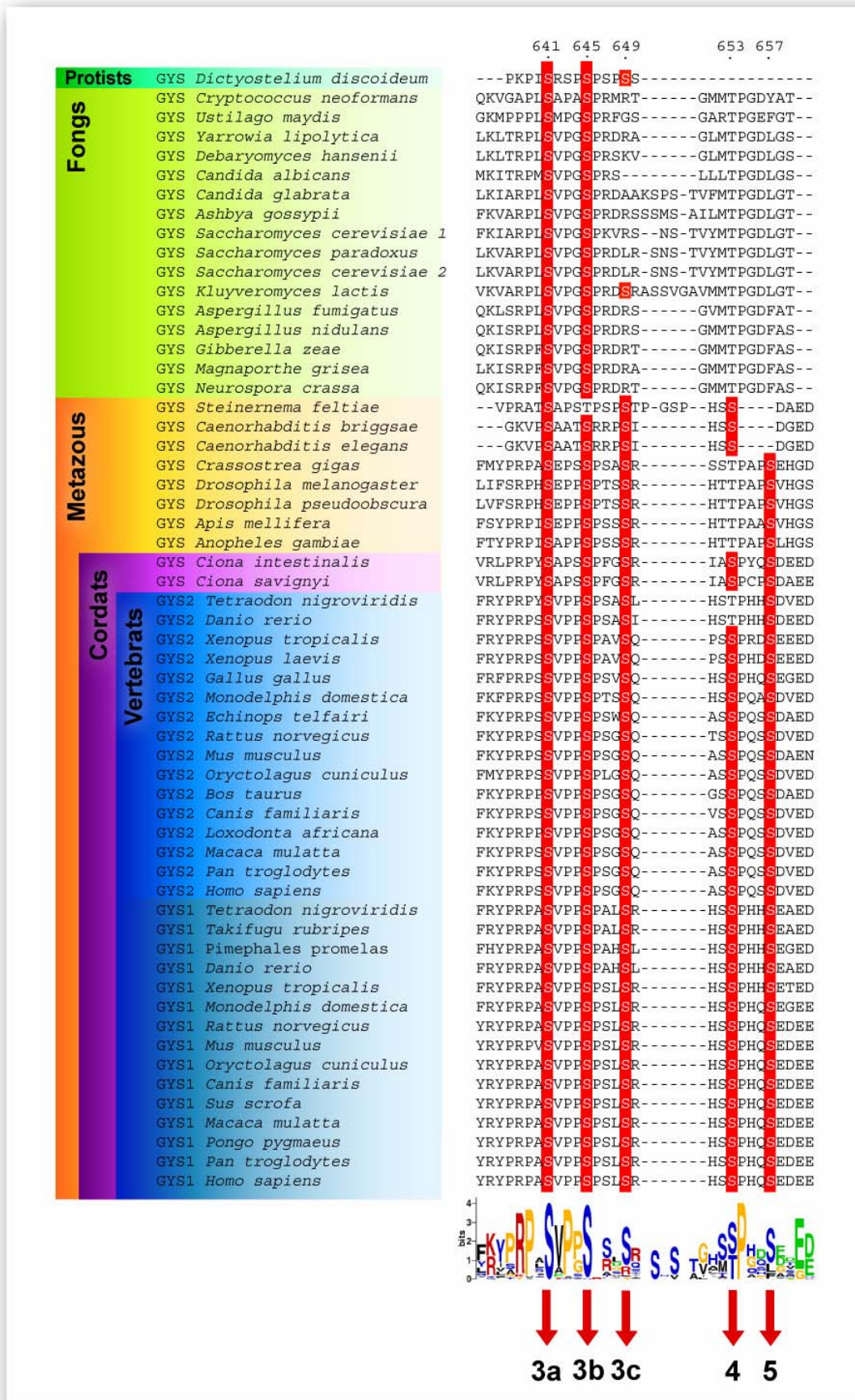


Figura 1.9 Alineament dels centres de fosforilació 3abc, 4 i 5. Es ressalten les serines conservades que constitueixen els centres de fosforilació indicats.

La GS de Giardia no està present en aquest alineament perquè la proteïna s'acaba abans d'arribar a aquesta regió C-terminal. Això ens indica que aquesta regió rica en centres de fosforilació s'ha incorporat com un domini nou al llarg de l'evolució del eucariotes, elongant la proteïna, i no pas per modificacions sobre una regió preexistent, com en el cas del clúster d'arginines.

Un altre cop els nemàtodes es perfilen con un grup particular que trenca l'evolució continua dels centres de fosforilació. Malgrat que l'alineament indica que aquesta regió en general està molt conservada, les seqüències corresponents als nemàtodes (*Caenorhabditis elegans*, *Caenorhabditis briggsae* i *Steinernema feltiae*) no mantenen una homologia clara i variaven la seva posició exacta segons l'alineament.

Per establir l'alineament definitiu que es mostra en la figura hem seguit la tendència que hem observat en la resta de seqüències, segons la qual la serina corresponent al centre de fosforilació 3a és el més conservat. Un cop fixada aquesta posició, s'observa que els nemàtodes perden el centre de fosforilació **3c**, mantenen els centres **3a** i **3b** i incorporen *de novo* els centres **4** i **5**.

Els centres de fosforilació de l'extrem N-terminal es designen **2** i **2a**, i corresponen a les posicions de les serines 8 i 11 de l'MGS humana. Quan alineem les seqüències de les GS eucariotes de que disposem per aquesta regió (figura 1.10) observem que aquest tram de seqüència ha estat una innovació dels metazous, ja que els fongs i els protistes no tenen l'extrem N-terminal tan llarg.

No obstant, els nemàtodes, mol·luscs, artròpodes i ascidis encara mostren vacil·lacions en la seqüència i no és fins als vertebrats que les serines S8 i S11 s'incorporen definitivament a la seqüència tant de la GS hepàtica com de la muscular.

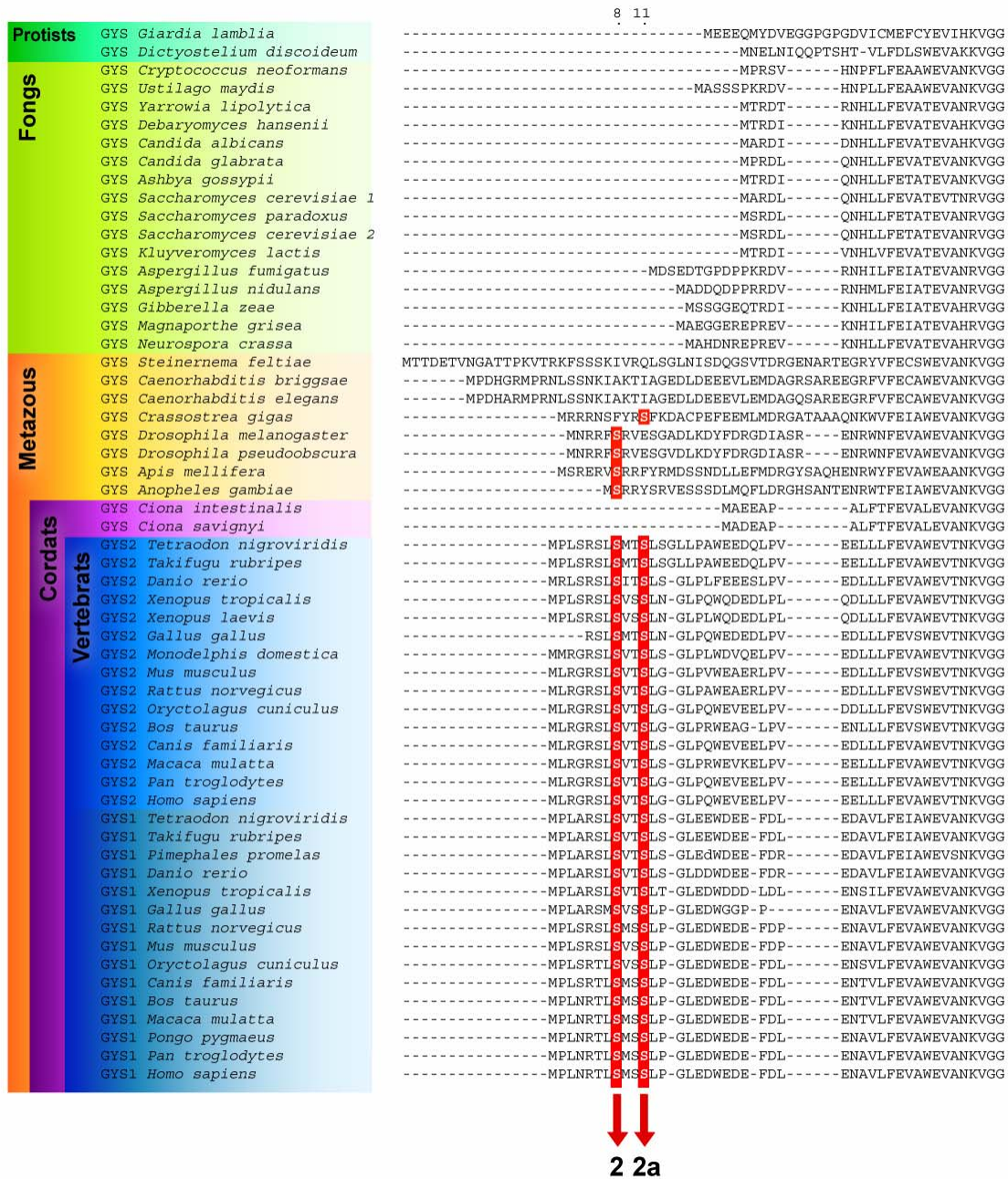


Figura 1.10 Alineament dels centres de fosforilació 1a i 1b. En blanc sobre vermell es destaquen les serines conservades que constitueixen els centres de fosforilació indicats. S’observa clarament com les LGS de protists i fongs són més curtes que les de metazous per aquest extrem N-terminal.

Els centres de fosforilació que comentarem a continuació són els centres 1a i 1b. Malgrat aquesta nomenclatura, són els darrers centres en evolucionar perquè només els trobem en l’extrem C-terminal de les GS musculars de mamífer, en una regió exclusiva d’aquestes isoformes. Aquesta



“cua” és la responsable de que les MGS de mamífer siguin una trentena d’aminoàcids més llargues que les seves homòlogues hepàtiques.

Aquesta zona de la proteïna és difícil d’alinejar perquè segons les prediccions d’estructura secundària és una regió desestructurada i a més és on hi ha menys homologia entre les isoformes musculars i les hepàtiques. Tanmateix podem observar clarament com les cues C-terminal de les GS hepàtiques són pràcticament idèntiques des de peixos fins a primats (figura 1.11), en canvi veiem com aquesta cua es va expandint en les isoformes musculars fins a arribar a la seva configuració definitiva en els mamífers.

En aquesta regió amb tanta variabilitat és on trobem els centres de fosforilació **1a** i **1b**, en les posicions de les serines S698 i S710 de l’MGS humana. Seguint estrictament l’alineament, aquestes serines fosforilables apareixerien només a partir dels mamífers placentaris, malgrat que en l’amfibi *Xenopus tropicalis* (granota amb urpes africana) i en el mamífer marsupial *Monodelphis domestica* (opossum) trobem serines molt properes a les posicions descrites.

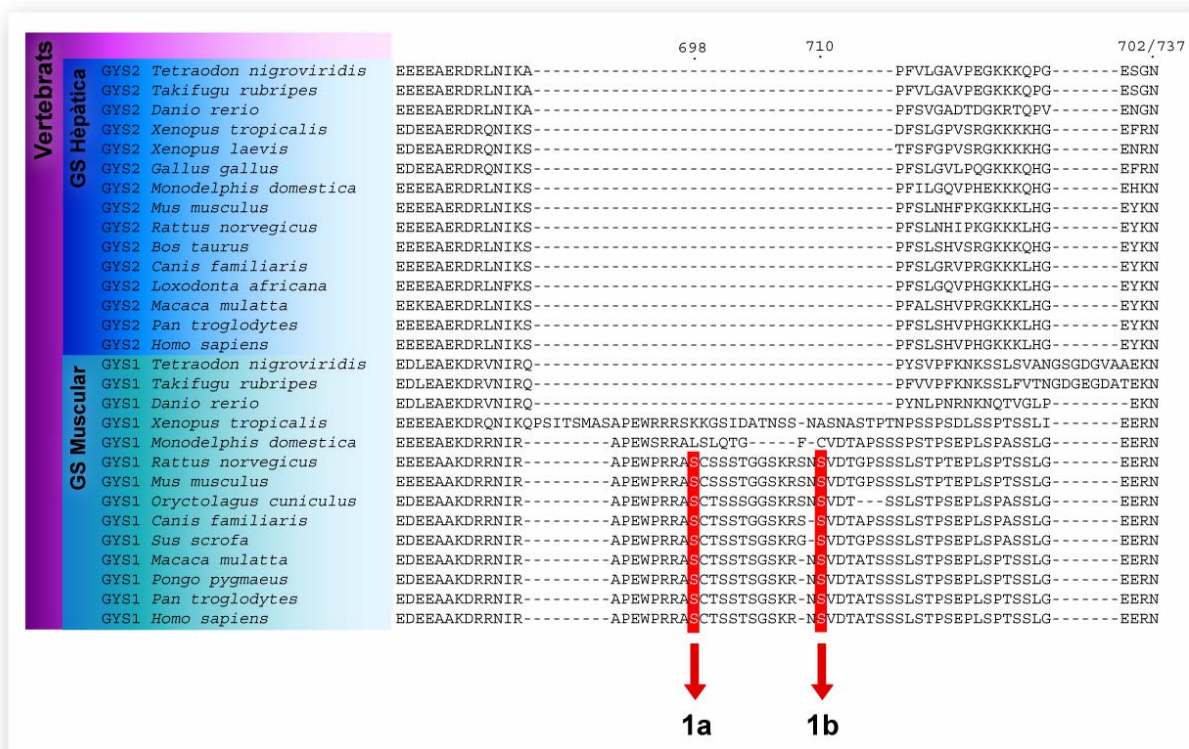
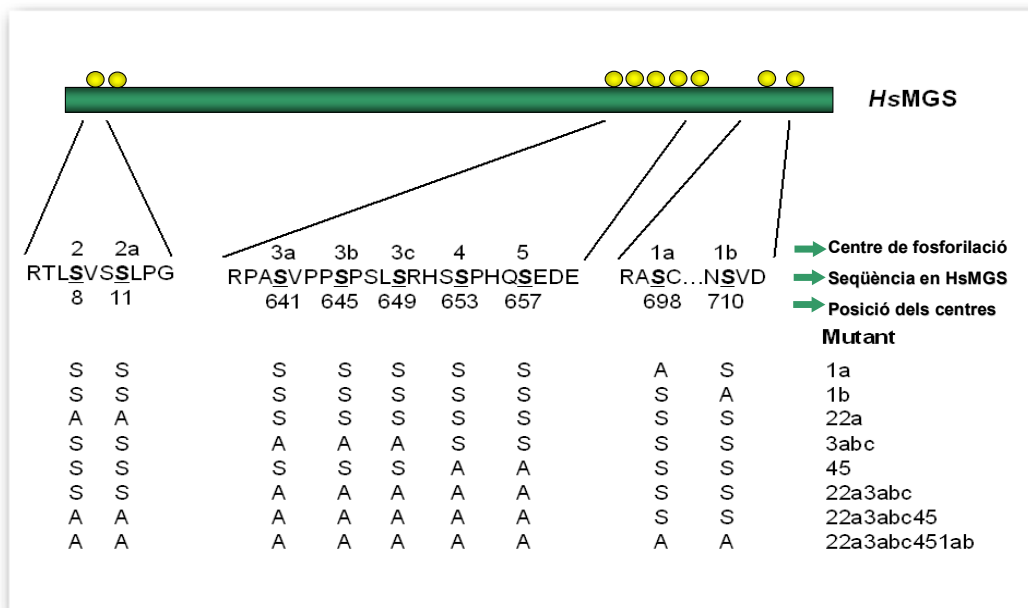


Figura 1.11 Alineament dels centres de fosforilació 1a i 1b. Es mostra l’alineament de la zona C-terminal de les GS de vertebrats. En vermell es destaquen les serines conservades que constitueixen els centres de fosforilació indicats.

Per a comprovar l'efecte dels centres de fosforilació sobre l'activitat de la glicogen sintasa vam recórrer un cop més als mutants puntuals. Com en el cas dels mutants del clúster d'arginines, vàrem substituir en l'MGS humana fusionada a GFP les serines dels centres de fosforilació per alanines, en les combinacions descrites en l'esquema de la figura 1.10. La nomenclatura per designar els mutants indica els centres que s'han mutat en cada construcció.

Les activitats dels mutants i de la proteïna salvatge es van determinar en cèl·lules COS-1 transfectades amb cadascuna de les construccions indicades en la figura 1.12. Les activitats mesurades es resumeixen en la taula 1.4. Les construccions de HsMGS que tenen mutats els centres de fosforilació **3a**, **3b** i **3c** mostren una tendència cap a valors de relacions I/T més alts.



**Figura 1.12 Situació i nomenclatura dels mutants dels centres de fosforilació.** Les serines dels centres de fosforilació s'han substituït per alanines en les combinacions indicades. La columna de l'esquerra mostra el nom dels corresponents mutants.

Aquests resultats indiquen que els centres de fosforilació amb més efecte sobre l'activitat de l'MGS humana són els centres **3a**, **3b** i **3c** i concorden amb els obtinguts per a l'MGS de conill (Skurat & Roach, 1995; Skurat et al., 1994).

**Taula 1.4**

Construcció	Activitat GS (mU/mg proteïna)		Ratio d'activitat I/T (-/+ G6P)
	- G6P	+ G6P	
GFP- <i>HsMGS</i>	20,7	51,3	0,40
1a	47,3	104,4	0,45
1b	21,2	54,4	0,39
22a	31,3	67,7	0,46
3abc	34,4	60,3	0,57
45	39,4	96,0	0,41
22a3abc45	36,5	74,6	0,49
22a3abc451ab	30,7	53,9	0,57

Les dades d'activitat ens confirmen els alineaments de seqüència que indicaven que els centres de fosforilació **3a** i **3b** eren els més antics. A partir d'aquests centres de fosforilació bàsics, la glicogen sintasa hauria anat adquirint d'altres centres al llarg de l'evolució per modular i regular més finament l'activitat.

### 1.2.3 Impacte dels centres de fosforilació i del clúster d'arginines sobre la translocació nucleocitoplasmàtica de l'*HsMGS*.

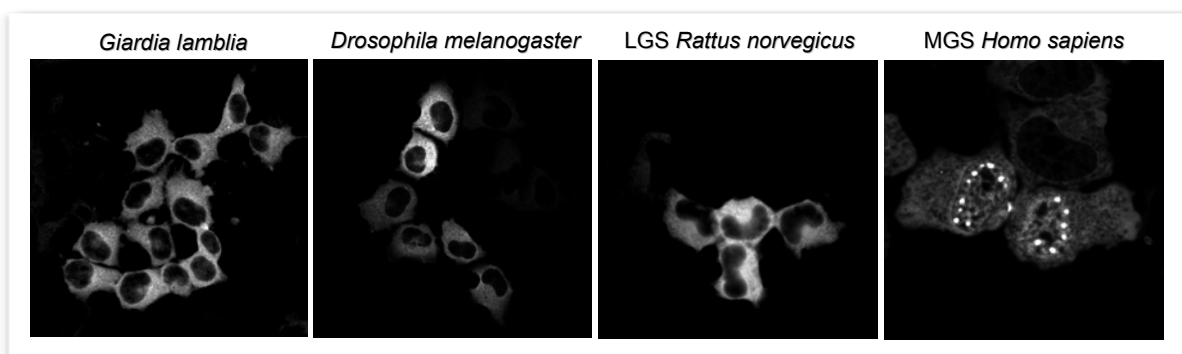
Estudis del nostre laboratori han demostrat *in vivo* que l'LGS de rata (*RnLGS*) i l'MGS humana (*HsMGS*) transloquen a diferents localitzacions subcel·lulars en funció de la presència o absència de glucosa i glicogen.

La *RnLGS* es distribueix uniformement pel citosol en absència de glucosa, i transloca a la perifèria cel·lular en presència de glucosa, on té lloc la síntesi de glicogen. En canvi, l'*HsMGS* transloca des d'uns compartiments subnuclears cap al citosol en resposta a glucosa.

Davant l'origen duplicatiu de les isoformes de GS, sorgeix la qüestió de quin és el patró de translocació de les glicogen sintases ancestrals, anteriors al fenomen de duplicació gènica.

Per això vam clonar les GS de *Giardia lamblia* (*GIGS*) i *Drosophila melanogaster* (*DmGS*) i vam estudiar si translocaven al nucli de cèl·lules FTO2B per comparar-la amb les *RnLGS* i *HsMGS*. Altre cop aquesta línia cel·lular es presenta com el millor sistema d'estudi degut a que l'*HsMGS* pot translocar al nucli de qualsevol tipus cel·lular, en canvi l'*LGS*, que mai s'acumula al nucli d'hepatòcits, quan la transfectem en cèl·lules que expressen MGS, entra al nucli no pels seus propis mitjans sinó perquè probablement oligomeritza amb l'MGS endògena d'aquestes cèl·lules. Així doncs, per evitar artefactes vam transfectar les construccions de GS ancestrals fusionades a GFP en aquest hepatoma que no expressa MGS.

La figura 1.13 ens mostra que la distribució de la *GIGS* i de la *DmGS* en cèl·lules deplecionades de glucosa és estrictament citosòlica, com la descrita per la *RnLGS*, indicant així un cop més que les GS hepàtiques conserven característiques ancestrals i que la translocació vers el nucli de l'*HsMGS* és una innovació evolutiva de les GS musculars.



**Figura 1.13 Translocació de les glicogen sintases eucariotes en resposta a la depleció de glucosa.** Es mostra la localització subcel·lular de GS eucariotes representatives fusionades a GFP transfectades en cèl·lules FTO2B i incubades en absència de glucosa. Només la isoforma muscular humana entra i s'agrega dins el nucli.

Continua essent necessari però, explicar perquè les dues isoformes de GS transloquen a llocs diferents i trobar els dominis de la proteïna que controlen la seva localització subcel·lular.

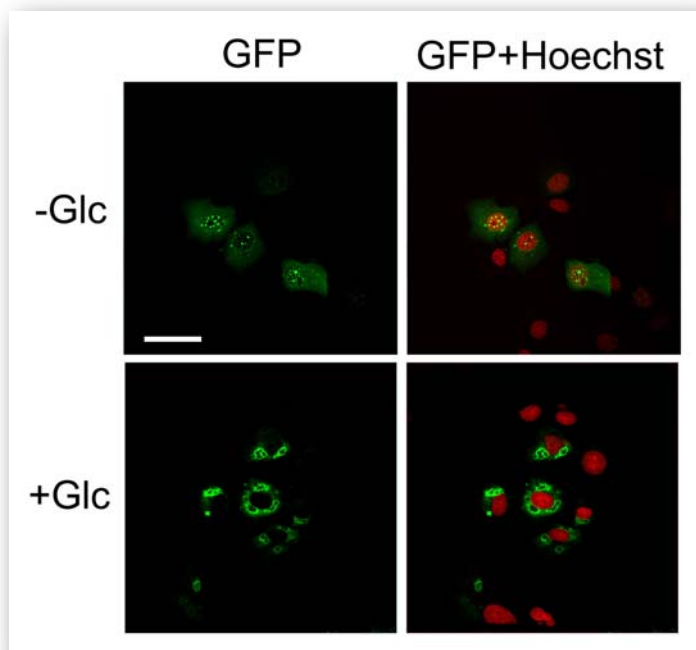
L'estudi evolutiu dels dominis més rellevants de la glicogen sintasa que hem realitzat ens pot servir com a punt de partida per determinar els dominis de la proteïna involucrats en la seva translocació.

Ens vam centrar en l'estudi de la translocació nuclear de l'*HsMGS* perquè experimentalment és la més clara de definir i podia ser útil en investigacions futures sobre el paper en el nucli de la glicogen sintasa. Les cèl·lules on vam dur a terme aquests experiments van ser de la línia COS-1, perquè la GS endògena no interfereix en la translocació de l'*HsMGS*.

En primer lloc vam estudiar la localització subcel·lular dels mutants de fosforilació que havíem caracteritzat en l'apartat anterior. Les construccions de GFP-*HsMGS* amb els centres de fosforilació **1a** i **1b** mutats eren fermes candidates a mostrar alteracions en la translocació cap al nucli perquè els centres de fosforilació **1a** i **1b** no estan presents en les isoformes hepàtiques de glicogen sintasa, i la *RnLGS* no transloca al nucli.

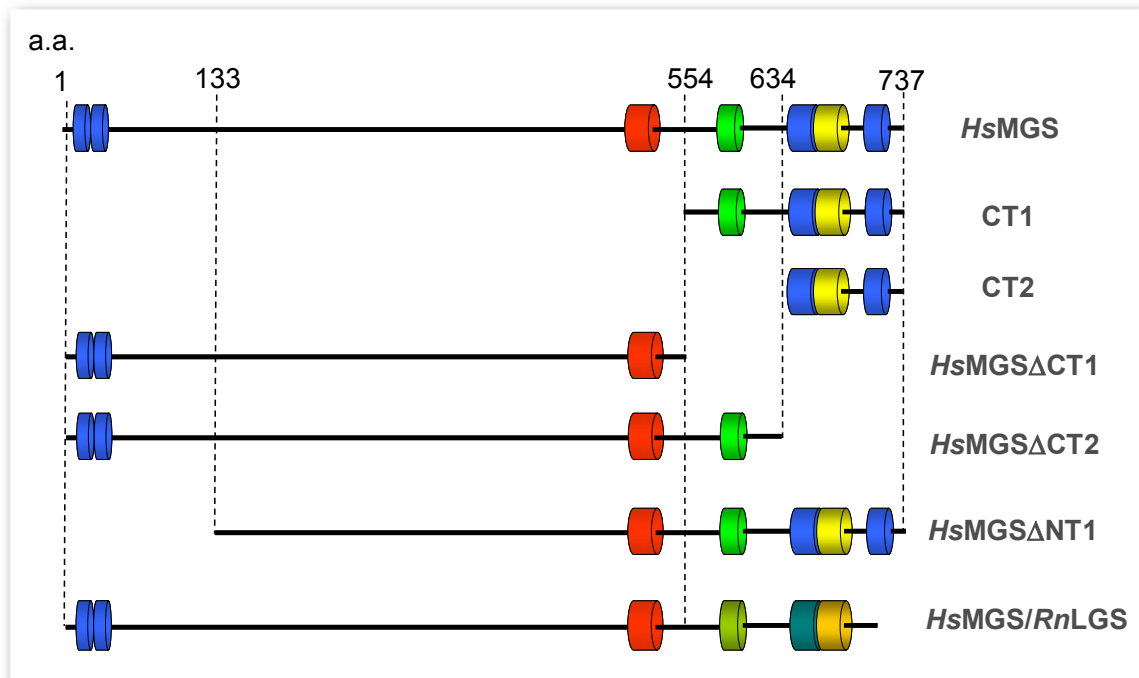
Però quan vam deplecionar de glucosa les cèl·lules transfectades amb cadascun dels mutants de fosforilació indicats en la figura 1.12, ni el mutants **1a** i **1b** ni cap de les altres construccions va mostrar alteracions en la translocació i la seva localització subcel·lular seguia el mateix patró que l'*HsMGS* salvatge. La distribució representativa de l'*HsMGS* salvatge en funció de la glucosa es mostra en la figura 1.14.

**Figura 1.14** Localització en cèl·lules COS-1 de l'*HsMGS* salvatge fusionada a GFP, en presència o absència de 25 mM glucosa. En el canal verd es mostra la localització de la *HsMGS* i en vermell els nuclis. Aquesta distribució subcel·lular també és representativa de la localització dels mutants de fosforilació. La barra blanca equival a 50 µm.



Així doncs, cap dels centres de fosforilació analitzats sembla regular la translocació de la proteïna.

A continuació vam abordar la recerca dels dominis implicats en la translocació de l'*HsMGS* de manera més sistemàtica, seguint l'aproximació clàssica de dissenyar construccions truncades. Tots els mutants de delecció de l'*HsMGS* es van fusionar a GFP. La nomenclatura dels mutants i la seva estructura en dominis es detalla en la figura 1.15 i el seu comportament en resposta a la presència o absència de glucosa queda reflexat en les fotografies de microscòpia confocal de la figura 1.16.



**Figura 1.15 Dominis de l'*HsMGS* i els respectius mutants de delecció.** Els blocs en blau representen la situació aproximada dels centres de fosforilació, en verd s'indica el clúster d'arginines, en vermell el centre catalític i en groc la zona rica en glutàmics i aspàrtics.

De tots els mutants de delecció, només el *HsMGS* $\Delta$ *CT1* tenia sèriament afectada la translocació al nucli. Aquest mutant era el que li mancava la zona corresponent al domini ric d'arginines.

El més sorprenent però, és que precisament aquesta zona està perfectament conservada entre les isoformes hepàtica i muscular, tal com hem determinat en els alineaments de seqüència.

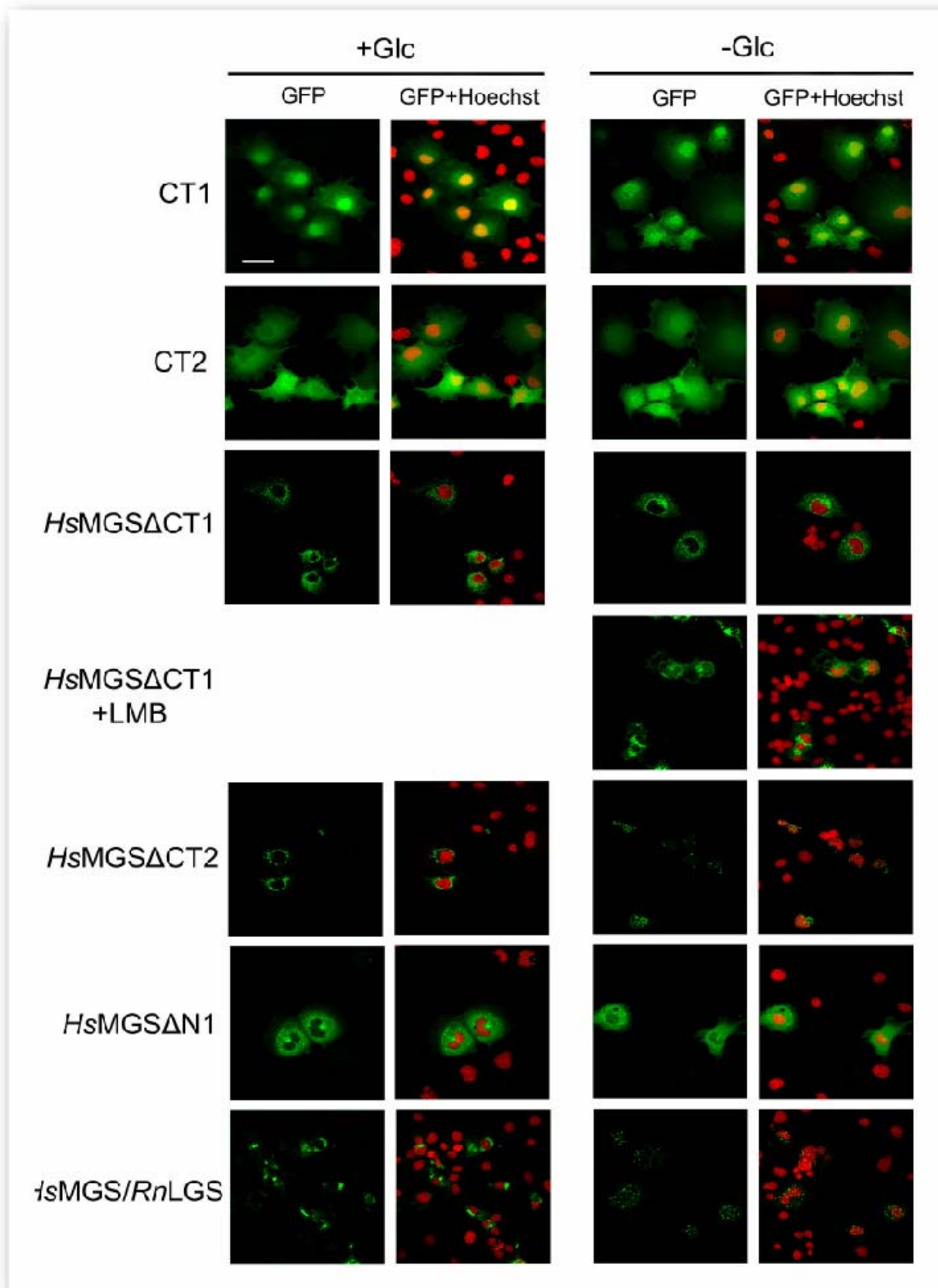
Les cues CT1 i CT2 es localitzaven al nucli independentment de les concentracions de glucosa, perquè no superaven el tamany d'exclusió del porus nuclear i difonien lliurement al nucli.

Donat que la translocació al nucli és una característica exclusiva de la GS muscular, era esperable que els mutants amb la translocació més afectada fossin aquells amb delecions en les zones de màxima divergència de seqüència, com els de les cues N- i C-terminal, però en canvi no observem cap efecte.

És més, si substituïem el tram de seqüència de l'*HsMGS* des del l'inici clúster d'arginines fins l'extrem C-terminal pel tram corresponent de la *RnLGS*, la quimera resultant presenta una translocació nucleocitoplàsmica normal. Per una banda això era previsible, doncs els clúster d'arginines de l'*HsMGS* i la *RnLGS* són idèntics, però sorprenia pel fet que la part més C-terminal de la *RnLGS* no afecta la translocació malgrat que és tan diferent de la cua de l'*HsMGS* i és on es localitzen la regió àcida i la majoria de centres de fosforilació,.

Davant la rellevància que semblava tenir la zona rica en arginines en la translocació, els següents experiments es van encaminar cap a l'estudi més fi i detallat d'aquest clúster.

Enlloc de treballar amb mutants de delecio que perden grans porcions de proteïna, vam transfectar els mutants R1 i R2, que hem descrit anteriorment, als quals els hi havíem substituït les tres primeres (mutant R1) o tres últimes (mutant R2) arginines de la seqüència consens d'aquest domini per alanines.



**Figura 1.16 Localització subcel·lular dels mutants de deleció de l'*HsMGS*.** Les diferents construccions de deleció fusionades a GFP van ser transfectades en cèl·lules COS-1 i incubades en presència o absència de glucosa 25 mM. Les construccions d'*MGS* es mostren en el canal verd i en vermell es marquen els nuclis. La barra blanca equival a 50  $\mu$ m.



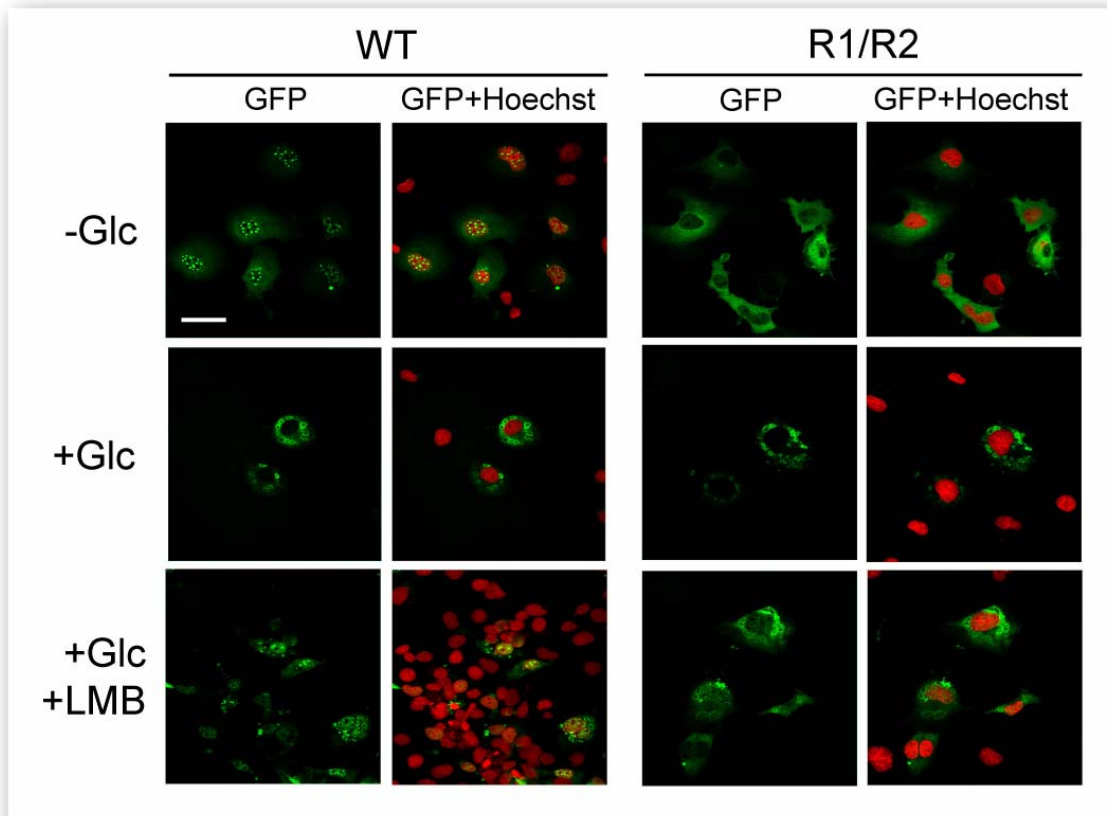
Primer va semblar que aquests mutants eren capaços de translocar al nucli amb normalitat, però un estudi més en profunditat ens va revelar que la seva cinètica de translocació cap al citosol en resposta a glucosa era més ràpida que la de la proteïna salvatge.

Això ens va conduir a mutar les sis arginines del clúster a alanines (mutant R1R2) per veure si es potenciava aquest efecte de sortida del nucli. Efectivament, el mutant R1R2 no s'acumulava al nucli en absència de glucosa, sinó que quedava distribuït per tot el citosol. En resposta a glucosa però, era capaç de sintetitzar glicogen com la proteïna salvatge (figura 1.17).

Per verificar que aquest doble mutant no entrava en el nucli vam fer un experiment de comprovació amb cèl·lules transfectades amb el mutant R1R2 i tractades amb leptomicina B (LMB) (figura 1.17). Aquesta droga interfereix l'export de proteïnes del nucli mediat per CRM1 (Ossareh-Nazari et al., 1997) i en experiments previs havia demostrat que també bloquejava la sortida de l'HsMGS salvatge del nucli. (Cid et al., 2005)

Vam induir la sortida del nucli amb glucosa al mateix temps que bloquejàvem l'export nuclear amb LMB en cèl·lules que prèviament havien estat deplecionades de glucosa. Sorprenentment, les cèl·lules així tractades mostraven acúmuls d'HsMGS/R1R2 en el nucli, la qual cosa indicava que el mutant R1R2 es troba en un cicle dinàmic de translocació nucleocitoplàsmica, que només es fa evident quan bloquegem la sortida del nucli.

Aquests resultats suggereixen que l'MGS està constantment entrant i sortint del nucli, i el fet que es localitzi en el citoplasma o en el nucli depèn de l'afinitat relativa per cadascun dels compartiments. El factor de retenció en el citoplasma és el glicogen, però cal esbrinar què és el que reté l'MGS en el nucli.



**Figura 1.17 Localització subcel·lular dels mutants del clúster d'arginines.** L'*HsMGS* salvatge i el mutant del clúster d'arginines, ambdós fusionats a GFP, van ser transfectats en cèl·lules COS-1 i incubats en presència o absència de glucosa 25 mM.

A la llum d'aquestes observacions podem afirmar que el clúster d'arginines participa en la retenció nuclear de l'*HsMGS* però aparentment no està implicat en el mecanisme de translocació ni organitza el patró d'agregació nuclear. La informació per a la translocació de l'*HsMGS* probablement està continguda en la estructura de la proteïna i no en dominis concrets de la seqüència primària.

### 1.3 DISCUSSIÓ

El coneixement sobre el metabolisme del glicogen s'ha incrementat molt des de que en 1959 el Professor Luís Leloir descobrí la glicogen sintasa (L. F. Leloir & Cardini, 1957). A partir de llavors s'han assolit noves fites que han permès comprendre la funció d'aquest enzim i la seva regulació.

Amb els treballs pioners realitzats en animals i extractes de teixit es va descobrir que la glicogen sintasa és un enzim regulat per efectors al·lostèrics (Leloir et al., 1959) i per fosforilació covalent (Friedman & Larner, 1962)

Després de l'adveniment de les tècniques de biologia molecular es va poder determinar que existeixen dos gens (Browner et al., 1989; Nuttall et al., 1994) que codifiquen per a la glicogen sintasa i amb el coneixement de la seqüència primària es va assignar les funcions als diferents dominis (Cid et al., 2000; Pederson et al., 2000; Skurat et al., 1994). Aquests esforços han culminat amb recent resolució de l'estructura de dos membre de la família de les glicogen sintases (Buschiazzo et al., 2004; Horcajada et al., 2006).

Actualment gràcies a l'avenç en les tècniques de seqüenciació el nombre de genomes disponibles ha augmentat fins a cobrir raonablement tot el ventall evolutiu. Així doncs, ara es donen les condicions per abordar amb solvència l'estudi de la glicogen sintasa i els seus mecanismes de regulació a la llum de l'evolució.

El primers resultats obtinguts seguint aquesta nova aproximació suggereixen que les dues isoformes de glicogen s'originen per duplicació gènica d'un ancestre comú dels vertebrats. La consistència d'aquesta conclusió rau en el fet que arribem al mateix resultat tant a partir dels estudis filogenètics com de la comparació de l'estructura gènica, i l'alt grau de sintènia ens indica la conservació d'aquest escenari que es va configurar a l'origen dels vertebrats.

Tot i que hem detectat l'existència de dos gens de GS en totes les espècies de vertebrats amb genomes completament seqüenciats que hem analitzat, en la resta de genomes eucariotes només hem trobat una còpia del gen de la GS. L'estudi de la presència i funcionalitat de dues isoformes de GS en el vertebrats serà l'objectiu dels següents capítols.

L'únic organisme no vertebrat amb dues còpies funcionals de GS (GSY1 i GSY2) és el fong *Saccharomyces cerevisiae* (Farkas et al., 1990), originades en un fenomen de duplicació gènica independent del que ocorre en vertebrats. Està documentat que aquesta espècie de fong ha patit duplicacions en bloc de part del genoma, originant multitud de gens duplicats (Scannell et al., 2006). La majoria d'aquests gens han incorporat mutacions fins a perdre totalment l'activitat o han evolucionant fins a adquirir noves funcions (Lynch & Conery, 2000). En el cas de les glicogen sintases de llevat, la funció metabòlica de sintetitzar glicogen la duu a terme la GSY2, mentre que la GYS1 aparentment no té un paper destacat en l'acumulació de reserves (Farkas et al., 1991).

Durant la recerca de seqüències putatives de glicogen sintasa hem identificat alguns pseudogens que no es mostren en els alineaments, sobretot en els rosegadors. Els hem identificat per la presència de truncacions prematures, abundància de mutacions en la pauta oberta de lectura i pèrdua d'introns. En cas de dubte també hem analitzat els gens de l'entorn per comprovar que es mantenia la sintènia.

Els alineaments de seqüència ens afegixen una dimensió més a la informació que podem extreure d'una sola seqüència perquè mostren el camí que ha seguit la proteïna per arribar a la configuració actual i fan destacar les regions conservades.

Així trobem dominis conservats on no només es mantenen els aminoàcids implicats en la funció catalítica sinó també el seu l'entorn. Aquest és el cas del clúster d'arginines o dels residus de la regió catalítica, que mantenen l'homologia gairebé en tots els eucariotes. La conservació de l'entorn dels dominis de fosforilació és menys estricta i denota una incorporació successiva de dianes de fosforilació a la seqüència. Una explicació plausible per aquesta menor conservació fora dels residus fosforilables podria ser que els dominis de fosforilació estan situats en zones de la proteïna sense estructura secundària, on les restriccions evolutives per mantenir l'estructura són més lleus.

La regulació al·lostèrica per G6P i la fosforilació són mecanismes de regulació característics les GS eucariotes. Gràcies als alineaments però, observem que aquestes innovacions no eren presents en l'ancestre comú de totes les GS eucariotes sinó que es van configurar paulatinament sobre els elements de seqüència originals. Així doncs, la GS de *Giardia lamblia*, tot i pertànyer a un organisme eucariota ancestral encara no posseeix el clúster d'arginines consens, ni les regions de fosforilació.

Les dades filogenètiques i la comparació de seqüències suggereixen que la regulació al·lostèrica i la fosforilació covalent de la glicogen sintasa evolucionen simultàniament. El clúster d'arginines responsable de l'activació al·lostèrica apareix al mateix temps que ho fan els primers centres de fosforilació (3a i 3b). En els protistes encara es veuen vacil·lacions de la seqüència, però a partir dels primers fongs ja es fixa un consens. Aquestes observacions reforcen les dades que indiquen un cert grau de complementarietat entre la regulació al·lostèrica i covalent perquè l'acció de la G6P sobre la GS afavoreix l'acció de les fosfatases (Villar-Palasi & Guinovart, 1997). Els models d'estructura de les GS humanes indiquen que és factible una interacció física entre la regió rica en arginines i els llocs de fosforilació 2a i 3a. Aquesta interacció probablement estaria regulada per l'estat de fosforilació.

L'evolució en paral·lel possiblement també es posa de manifest a la inversa i explicaria el cas de les GS de *Caenorhabditis*. Les GS del clade dels nemàtodes no tenen ni el clúster d'arginines canònic (la qual cosa és molt significativa degut a l'elevada conservació d'aquest domini), ni els centres de fosforilació consens. Seran necessaris més estudis per poder entendre les alteracions que suposen aquests canvis en el metabolisme del glicogen d'aquests organismes.

La translocació de la glicogen sintasa s'ha estudiat en profunditat des de que va ser descoberta en el nostre laboratori (Fernandez-Novell et al., 1992b; Ferrer et al., 1997). Les diferències en el patró de translocació de les isoformes muscular i humana feien impossible predir quin era el comportament de les GS més primitives. Gràcies a la clonació de GS ancestrals com la de *Giardia lamblia* o *Drosophila melanogaster* i a l'estudi de la seva localització

subcel·lular podem concloure que les GS dels organismes precordats es comporten com les isoformes hepàtiques de vertebrats.

No obstant degut a la variació de seqüència entre les isoformes musculars de vertebrats no podem afirmar que totes les MGS transloquin vers al nucli fins que no es cloni i s'analitzi la translocació de glicogen sintases representatives d'altres grups de vertebrats, com els peixos i amfibis.

Un altre aspecte important en l'estudi de la translocació és la identificació de la regió responsable d'aquest comportament. Els resultats obtinguts amb mutants puntuals i de deleció de la *HsMGS* fusionada a GFP ens permeten concloure que la translocació en resposta a glucosa no està regulada per fosforilació i que el clúster d'arginines és important per a la retenció nuclear de la *HsMGS* però que no afecta ni l'entrada al nucli de la proteïna ni el patró d'acumulació nuclear. Probablement la informació per la translocació de la *HsMGS* està codificada al llarg de tota l'estructura, enlloc de en dominis discrets i aïllats.

El bloqueig de l'export nuclear mediat per leptomicina B indica que la sortida de la *HsMGS* està controlada per la proteïna CRM1. En el cas invers, d'entrada al nucli, la manca d'una seqüència consens de localització nuclear (NLS) suggereix que l'MGS ha d'interaccionar amb una segona proteïna amb senyal NLS propi per poder translocar cap al nucli.

El coneixement dels mecanismes de translocació que generen aquests resultats és una eina molt valuosa que permetrà emprendre futurs estudis a fi d'esbrinar la funció de la translocació nuclear de la glicogen sintasa muscular.