

**DEPARTAMENT DE BIOQUÍMICA I  
BIOLOGIA MOLECULAR**

**TESI DOCTORAL**

**DETERMINACIÓ DE L'ESTRUCTURA TRIDIMENSIONAL DE LA  
GLICOGEN SINTASA DE *PYROCOCCUS ABYSSI***

**Cristina Horcajada Garro**

**Barcelona, 2005**

# ÍNDEX

<b>PRESENTACIÓ I OBJECTIUS .....</b>	<b>1</b>
<b>INTRODUCCIÓ .....</b>	<b>5</b>
1. Metabolisme del glicogen .....	7
1.1. Glicogen .....	7
1.2. Síntesi i degradació del glicogen .....	8
2. Glicogen sintases .....	10
2.1. Glicogen sintases de mamífers i fongs .....	11
2.2. Glicogen sintases de bacteris i midó sintases de plantes .....	13
2.3. Glicogen sintases d'arqueons .....	13
3. Glicosiltransferases .....	14
3.1. Mecanisme catalític de les glicosiltransferases .....	16
<b>CAPÍTOL 1: Clonació, expressió, purificació i cristal·lització de la glicogen sintasa de <i>Pyrococcus abyssi</i> .....</b>	<b>19</b>
<b>Introducció .....</b>	<b>21</b>
<b>Resultats .....</b>	<b>22</b>
1.1. Clonació de la PaGS .....	22
1.2. Expressió de la PaGS .....	22
1.3. Purificació dels cossos d'inclusió .....	23
1.4. Caracterització de la PaGS en solució .....	25
1.4.1. Dicroisme circular .....	25
1.4.2. Activitat enzimàtica .....	26
1.4.3. Gels nadius i gel filtració .....	27
1.4.4. Ultracentrifugació analítica .....	29
1.4.5. Microscòpia electrònica .....	30
1.5. Cristal·lització de la PaGS .....	31
1.5.1. Condicions inicials .....	31
1.5.2. Microsembra .....	32
1.5.3. Macrosembra .....	33
1.5.4. Cocrystal·lització .....	33
1.5.5. Criopreservació .....	34
1.6. Aproximacions experimentals al problema de la fase .....	34
1.6.1. Selenometionines .....	34
1.6.2. Àtoms pesats .....	34
<b>Discussió .....</b>	<b>36</b>

---

<b>CAPÍTOL 2: Clonació, expressió, purificació i cristal·lització del domini C-terminal de la glicogen sintasa de <i>Pyrococcus abyssi</i></b> .....	<b>39</b>
<b>Introducció</b> .....	<b>41</b>
<b>Resultats</b> .....	<b>42</b>
2.1. Clonació del domini C-terminal de la PaGS .....	42
2.2. Expressió del domini C-terminal de la PaGS .....	43
2.3. Purificació del domini C-terminal de la PaGS .....	43
2.4. Cristal·lització del domini C-terminal de la PaGS .....	45
2.4.1. Condicions inicials i optimització .....	45
2.4.2. Cocrystal·litzacions i incubacions amb substrats .....	46
2.4.3. Criocrystal·litzacions .....	46
<b>Discussió</b> .....	<b>47</b>
<b>CAPÍTOL 3: Resolució de les estructures de la glicogen sintasa de <i>Pyrococcus abyssi</i> i del seu domini C-terminal</b> .....	<b>49</b>
<b>Introducció</b> .....	<b>51</b>
<b>Resultats</b> .....	<b>51</b>
3.1. Resolució del domini C-terminal de la PaGS .....	51
3.1.1. Recollida dels espectres de difracció i processament .....	51
3.1.2. Resolució de l'estructura del domini C-terminal .....	53
3.2. Resolució de la PaGS .....	56
3.2.1. Recollida dels espectres de difracció i processament .....	56
3.2.2. Resolució de l'estructura de la PaGS .....	60
3.2.2.1. Dispersió anòmala per selenometionines .....	60
3.2.2.2. Reemplaçament isomorf per àtoms pesats .....	62
3.2.2.3. Reemplaçament molecular. Funcions bloquejades .....	63
3.2.2.4. Reemplaçament molecular a partir del domini C-terminal ...	65
<b>Discussió</b> .....	<b>70</b>
<b>CAPÍTOL 4: Descripció estructural de la glicogen sintasa de <i>Pyrococcus abyssi</i></b> .....	<b>73</b>
<b>Introducció</b> .....	<b>75</b>
<b>Resultats</b> .....	<b>76</b>
4.1. Descripció general .....	76
4.1.1. Estructura del domini C-terminal de la PaGS .....	76
4.1.2. Empaquetament cristal·lí del domini C-terminal .....	78
4.1.3. Estructura de la PaGS .....	80
4.1.4. Empaquetament cristal·lí de la PaGS .....	83
4.1.5. Comparació del domini C-terminal aïllat i formant part de PaGS ....	84
4.2. Comparació de la PaGS amb altres glicosiltransferases .....	85
4.2.1. Comparació de la PaGS amb la AtGS .....	85
4.2.2. Comparació de la PaGS amb la resta de glicosiltransferases .....	87

4.3. Unió de substrat .....	89
4.3.1. Unió d'una molècula d'acetat .....	89
4.3.2. Comparació del lloc d'unió de nucleòtid de PaGS i AtGS .....	91
4.3.3. Comparació del lloc d'unió de nucleòtid de PaGS i OtsA .....	92
4.3.4. Model d'unió de UDP .....	95
4.4. Centre catalític .....	96
4.4.1. Modelat de la PaGS activa .....	96
4.4.2. Residus catalítics .....	97
4.4.2.1. Àcid glutàmic 339 .....	100
4.4.2.2. Histidina 151 .....	100
4.4.2.3. Asparagina 217 .....	102
4.4.2.4. Arginina 257 .....	102
4.4.2.5. Glicina 17 .....	102
4.5. Unió a glicogen .....	103
4.5.1. Unió a les glucoses +1 i +2 .....	103
4.5.2. Unió a les glucoses +3 i +4 .....	105
4.6. Oligomerització .....	112
4.6.1. Trimerització de la PaGS .....	112
4.6.2. Comparació entre els estats d'oligomerització de PaGS i AtGS .....	114
4.6.3. Predicció de l'estat d'oligomerització de les GS eucariotes .....	115
4.7. Unió a dioxans .....	116
4.8. Modelat dels llocs de fosforilació de l'extrem C-terminal en GS eucariotes	118
4.9. Model mecànic .....	119
<b>Discussió .....</b>	<b>121</b>
<b>DISCUSSIÓ GENERAL .....</b>	<b>125</b>
<b>CONCLUSIONS .....</b>	<b>131</b>
<b>MATERIALS I MÈTODES .....</b>	<b>135</b>
A. Biologia molecular .....	137
A.1. Tècniques estàndard .....	137
A.1.1. PCR .....	137
A.1.2. Clonatge .....	137
A.1.3. Transformació de cèl·lules competents .....	137
A.1.4. Obtenció de DNA plasmídic .....	138
A.1.5. Seqüenciació de DNA .....	138
A.1.6. Electroforesi de DNA .....	138
A.2. Obtenció dels vectors d'expressió .....	139
B. Tècniques d'expressió recombinant .....	139
C. Tècniques de purificació .....	140
C.1. Cromatografia d'afinitat .....	140
C.2. Replegament de la proteïna en una columna d'afinitat .....	141
C.3. Digestió amb trombina .....	142
C.4. Cromatografia de filtració en gel .....	142

D. Tècniques d'electroforesi .....	142
D.1. Electroforesi en gel desnaturalitzant .....	142
D.2. Electroforesi en gel nadiu .....	143
D.3. Tinció amb blau de Coomassie .....	143
D.4. Seqüenciació de proteïnes .....	144
E. Determinació de l'activitat enzimàtica i concentració de proteïna .....	144
E.1. Determinació de l'activitat enzimàtica .....	144
E.2. Mètode Bradford .....	144
F. Tècniques de caracterització en solució .....	144
F.1. Dicroisme circular .....	144
F.2. Ultracentrifugació analítica .....	145
G. Microscòpia electrònica .....	145
H. Tècniques de cristallització .....	145
H.1. Gota penjant .....	146
H.2. Microsembra i macrosembra .....	146
H.3. Criopreservació .....	146
I. Recollida dels espectres de difracció .....	147
I.1. Generador d'ànode rotatori .....	147
I.2. Radiació de sincrotró .....	147
I.3. Processament dels espectres de difracció .....	148
J. Mètodes de resolució de l'estructura .....	148
J.1. SAD, MAD, SIR, MIR .....	148
J.2. Reemplaçament molecular .....	149
J.2.1. Models .....	149
J.2.2. Modes normals de vibració .....	150
J.2.3. Simetria no cristal·logràfica .....	150
J.3. Afinament per cos rígid .....	150
K. Extensió de les fases .....	150
K.1. Construcció de màscares .....	151
K.2. Promitjat per simetria no cristal·logràfica .....	151
L. Afinament de l'estructura .....	151
M. Representacions de l'estructura i superposicions .....	151
<b>BIBLIOGRAFIA .....</b>	<b>153</b>
<b>ANNEX .....</b>	<b>167</b>

## ABREVIATURES

ADPG	adenildifosfat-glucosa
AMPc	adenilmonofosfat cíclic
AtGS	glicogen sintasa d' <i>Agrobacterium tumefaciens</i>
BE	enzim ramificador del glicogen
bp	parells de bases nucleotídiques
BSA	albúmina de sèrum boví
CK	caseïna quinasa
Da	dalton
DME	enzim desramificant del glicogen
GFP	proteïna verda fluorescent
G(M)S	glicogen midó sintasa
GN	glicogenina
Glc 1-P	glucosa 1-fosfat
Glc 6-P	glucosa 6-fosfat
GP	glicogen fosforilasa
GS	glicogen sintasa
GSD	malaltia d'emmagatzemament de glicogen
GSK-3	glicogen sintasa quinasa 3
Gsy2p	glicogen sintasa de llevat 2
GT	glicosiltransferasa
HK	hexoquinasa
kb	kiloparells de bases nucleotídiues
kDa	kilodaltons
LGS	glicogen sintasa de fetge
MAD	difracció anòmala múltiple
MalP	maltodextrin fosforilasa
MGS	glicogen sintasa muscular
MIR	reemplaçament isomorf múltiple

MR	reemplaçament molecular
MS	midó sintasa
NIDDM	diabetis mellitus no depenent d'insulina
PDB	protein data bank
PP	proteïna fosfatasa
PKA	proteïna quinasa dependent d'AMPc
SAD	difracció anòmala simple
SIR	reemplaçament isomorf simple
UDPG	uridildifosfat-glucosa

## PRESENTACIÓ i OBJECTIUS

Estudis previs del nostre grup d'investigació, basats en l'alineament de seqüències i la predicció d'estructures, han mostrat que les glicogen sintases (GS) d'animals i fongs de la família 3 de les glicosiltransferases (GT), i les GS de bacteris i midó sintases de plantes (GM), membres de la família 5 de les GT, tenen en comú un conjunt de característiques estructurals que fan suposar que tots aquests enzims operen a través d'un mecanisme catalític molt similar (Cid et al. 2002). D'altra banda, per a les GS de arqueons, classificades en la família GT-5, s'han predit propietats estructurals i funcionals presents als membres d'una, de l'altra, o d'ambdues famílies. La seva mida reduïda, és el membre més petit d'aquest superfamília, i el seu caràcter termofílic fan de la GS de l'arqueon *Pyrococcus abyssi* un bon model per abordar estudis estructurals i mecànics en aquest grup de glicosiltransferases. En el moment d'iniciar aquesta tesi no s'havia resolt l'estructura de cap membre de les famílies GT-3 o GT-5.

Així doncs, l'objectiu central d'aquesta tesi fou la **caracterització bioquímica i la determinació de l'estructura tridimensional de la glicogen sintasa de *P. abyssi* (PaGS) mitjançant cristal·lografia de raigs X**. Aquest objectiu es pots desglossar en els apartats següents:

- 1. Determinació de les condicions òptimes per l'obtenció de PaGS amb un grau de puresa adequat per a realitzar estudis bioquímics i estructurals.**

Era necessari dissenyar un protocol d'expressió heteròloga i purificació necessaris per a obtenir PaGS recombinant amb una quantitat i una qualitat adients per als posteriors estudis bioquímics i estructurals.

- 2. Caracterització bioquímica de la PaGS en solució.**

Atès que no s'havien descrit amb anterioritat les característiques de la GS d'aquest termòfil calia mostrar, en primer lloc, que posseïa efectivament activitat GS, així com determinar la seva especificitat per substrat i l'estat d'oligomerització en solució.

- 3. Resolució de l'estructura tridimensional de la PaGS per cristal·lografia de raigs X.**

Per a la consecució d'aquest objectiu s'havien d'obtenir cristalls de l'enzim que difractessin a alta resolució i, a partir de les dades de difracció, resoldre la seva estructura per tècniques de reemplaçament isomorf, difracció anòmala o reemplaçament molecular.



#### **4. Anàlisi de l'estructura de la PaGS.**

Un cop resolta l'estructura de l'enzim, calia analitzar les bases moleculars que determinen les seves característiques estructurals i la seva especificitat per substrat. Finalment, interessava aportar dades sobre el possible mecanisme catalític, fins ara desconegut, a través del qual operen aquest grup d'enzims i sobre les propietats estructurals i de regulació de GS eucariotes més complexes.

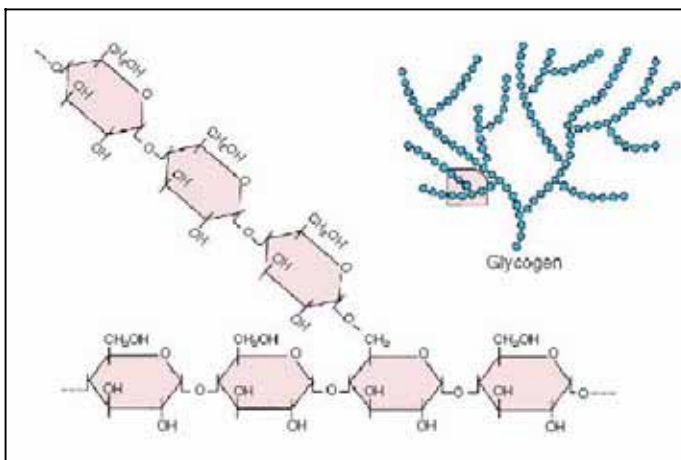
# INTRODUCCIÓ

## 1. METABOLISME DEL GLICOGEN:

### 1.1. Glicogen:

Gairebé tots els organismes vius acumulen polímers de glucosa com a forma d'emmagatzematge d'energia (Buleon *et al.* 1998; Roach *et al.* 1998). Els animals i els bacteris sintetitzen glicogen, mentre que el reservori de glucosa de les plantes és el midó.

El glicogen és un polímer d'unitats de D-glucosa unides per enllaços  $\alpha$ -1,4-glicosídics amb punts de ramificació formats per unions  $\alpha$ -1,6-glicosídiques (Fig.1). La cadena de glicogen creix fins que no pot incorporar més residus de glucosa a causa de l'impediment estèric (Melendez *et al.* 1998), fet que produeix el que es coneix com a partícula  $\beta$ . Una partícula  $\beta$  té un diàmetre d'uns 25 nm i un pes molecular aproximat de  $10^7$  Da. Aquest és el tipus d'organització que es troba en el glicogen muscular. Al fetge, en canvi, s'ha observat la presència de partícules amb un diàmetre quatre o cinc cops més gran que el de la partícula  $\beta$ . Aquestes acumulacions, anomenades partícules  $\alpha$ , estan formades per l'agregació en forma de roseta de partícules  $\beta$ , mitjançant un ordenament desconegut (Drochmans 1962).



**Fig.1. Estructura del glicogen.** Esquema d'una molècula de glicogen. Els residus de glucosa s'uneixen mitjançant enllaços glicosídics  $\alpha$ -1,4 i  $\alpha$ -1,6 en els punts de ramificació.

La formació del polímer permet l'acumulació de grans quantitats de glucosa sense alterar l'osmolaritat intracel·lular, mentre que l'estructura ramificada de la molècula permet la seva ràpida degradació. A més, l'elevat grau de ramificació del glicogen fa que aquest sigui soluble, a diferència del midó, i permet la seva distribució pel citoplasma.

En els vertebrats, la majoria de cèl·lules contenen glicogen però, en quantitats absolutes, el fetge i el múscul esquelètic constitueixen els majors dipòsits d'aquest polisacàrid. El paper biològic del glicogen en aquests dos teixits és diferent. El glicogen hepàtic s'utilitza per a mantenir l'homeòstasi de la glucosa en sang (glucèmia), mentre

que el del múscul esquelètic és emprat per al consum propi en la contracció d'aquest teixit.

Els llevats (Sillje *et al.* 1999) i els bacteris (Preiss i Romeo 1989) sintetitzen glicogen passada l'etapa de creixement exponencial o bé en resposta a la manca d'algun nutrient. En ambdós casos, les mutacions que inactiven els enzims implicats en la síntesi del polisacàrid continuen essent viables tot i créixer més lentament. Malgrat això, la distribució pràcticament universal del metabolisme del glicogen en éssers vius ha portat a alguns autors a suggerir que l'absència de la maquinària necessària per a sintetitzar aquest polímer de glucosa en microorganismes és un indicador d'un comportament parasitari o simbiòtic (Henrissat *et al.* 2002).

S'ha descrit l'acumulació de glicogen en arqueons en resposta a la manca d'algun nutrient, com el nitrogen (König *et al.* 1982; Maitra *et al.* 2001; Gruyer *et al.* 2002). Com succeeix en bacteris, l'acumulació de polisacàrid és màxima al final de la fase del creixement exponencial.

Moltes proteïnes s'uneixen a les partícules de glicogen. El 95% del contingut proteic està format per la glicogen sintasa (GS), la glicogen fosforilasa (GP), l'enzim desramificant i algunes proteïna-quinases (Cohen 1978). Aquesta associació afavoreix les interaccions entre les diferents proteïnes, com les de les subunitats de les proteïna-fosfatases (PP) associades a glicogen amb la GS i la GP (Newgard *et al.* 2000). S'ha proposat que el glicogen i les proteïnes associades podrien formar un orgànul anomenat glicosoma (Rybicka 1996). Molt recentment, Prats i col·laboradors han descrit la associació del glicogen a un nou compartiment intracel·lular implicat en el metabolisme d'aquest polisacàrid en el múscul esquelètic *in vivo*. Es tracta d'unes estructures esfèriques presents en el múscul estimulat i formades a partir de la reorganització del citoesquelet d'actina, a les quals també transloquen la GS i la GP a l'inici de la resíntesi de glicogen (Prats *et al.* 2005).

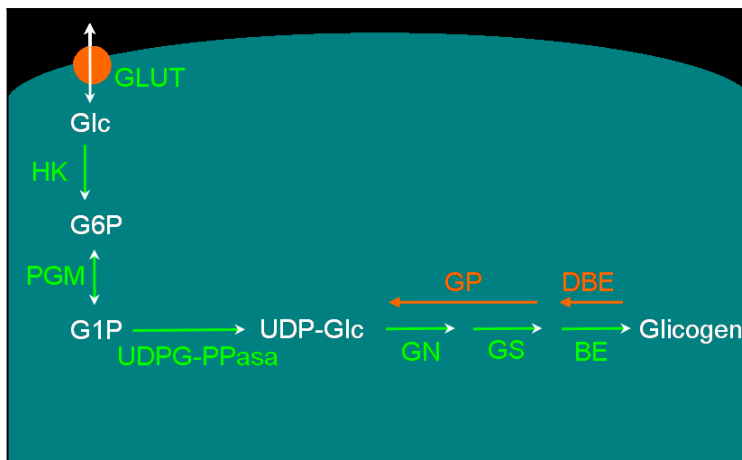
## 1.2. Síntesi i degradació del glicogen:

Un cop la glucosa entra dins la cèl·lula és convertida a glucosa 6-fosfat (Glc 6-P) pels enzims hexoquinasa o glucoquinasa. La ruta biosintètica del glicogen s'inicia amb la formació de la glucosa 1-fosfat (Glc 1-P) a partir de la Glc 6-P per acció de la fosfoglucomutasa, la Glc 1-P és posteriorment transformada en UDP-glucosa (UDPG) en vertebrats o ADP-glucosa en bacteris mitjançant les corresponents ADP- o UDP-pirofosforilases. La GS utilitza aquest glucòsid de nucleòtid per a allargar un polisacàrid preexistent, mentre que l'enzim ramificant, una transglicosidasa, introdueix els punts de ramificació. En el cas dels animals i els fongs, l'etapa limitant en la síntesi de glicogen és la catalitzada per la GS, mentre que per als bacteris i les plantes l'etapa limitant és la catalitzada per l'ADP-glucosa pirofosforilasa.

En la biogènesi del glicogen eucariota, les fases d'iniciació i creixement del polisacàrid estan catalitzades per diferents enzims. La iniciació és catalitzada per una proteïna encebadora, la glicogenina (GN) capaç d'autoglicosilar-se en un residu tirosina i

allargar la cadena fins a uns set residus de glucosa. Aquest oligosacàrid unit a la glicogenina és elongat per la GS i ramificat per l'enzim ramificant (Cao *et al.* 1993). S'ha demostrat recentment que, al contrari del que succeeix en cèl·lules eucariotes, la GS del bacteri *A.tumefaciens* és capaç de catalitzar ambdós passos d'iniciació per autoglicosilació i elongació, i per tant sintetitzar glicogen *de novo* (Ugalde *et al.* 2003).

La degradació del glicogen es produeix per l'acció de la GP i de l'enzim desramificant. La primera degrada les cadenes lineals i produeix Glc 1-P fins que arriba a una distància de tres o quatre unitats de glucosa d'un punt de ramificació. Aquesta branca de curta longitud unida a la macromolècula per un enllaç  $\alpha$ -1,6 és transferida a una altra cadena mitjançant la formació d'un enllaç  $\alpha$ -1,4.



**Fig. 2. Metabolisme del glicogen.** Esquema de la síntesi (verd) i degradació del glicogen (taronja) en cèl·lules de mamífer. La glucosa entra a la cèl·lula mitjançant el transportador GLUT i es transforma a Glc 6-P per l'hexoquinasa (HK) que al seu torn és convertida a Glc 1-P per la fosfoglucomutasa (PGM). La UDPG-pirofosforilasa (UDPG-PPasa) transforma Glc 1-P en UDP-Glc i aquesta última s'utilitza per a iniciar amb la glicogenina (GN), elongar amb la glicogen sintasa (GS). La cadena de glicogen és ramificada per l'enzim ramificant (BE) una cadena de glicogen. El pas invers, la degradació, està catalitzat principalment pels enzims glicogen fosforilasa (GP) i enzim desramificant (DBE).

En els darrers anys s'ha fet evident que el metabolisme del glicogen en animals superiors està compartimentalitzat en diferents espais subcel·lulars i que existeixen moviments dels enzims implicats que depenen de l'estat metabòlic i poden constituir un punt de regulació addicional d'aquest metabolisme. Ja es coneixia des dels anys 60 que la glucosa activa la GS (Katz *et al.* 1979) al fetge i als hepatòcits aïllats de rata. Més recentment ha estat descrit per Fernández-Novell i col·laboradors (Fernandez-Novell *et al.* 1992; Fernandez-Novell *et al.* 1994) que la glucosa provoca canvis en la distribució subcel·lular de l'enzim. En resposta a aquest monosacàrid, l'enzim transloca de la fracció soluble a la fracció que sedimenta a 10000xg en l'homogeneïtzat d'hepatòcits. També ha estat descrit que el tractament amb glucosa d'hepatòcits en suspensió fa canviar la distribució uniforme en el citoplasma de la GS cap a una acumulació a la perifèria cel·lular, on comença la síntesi de glicogen (Fernandez-Novell *et al.* 1997). Resultats similars han estat obtinguts en hepatòcits en cultiu (Garcia-Rocha *et al.* 2001).

D'una altra banda, estudis realitzats per Ferrer i col·laboradors (Ferrer *et al.* 1997), (Cid *et al.* 2000) (Cid 2005) també han demostrat que la GS muscular sofreix un canvi en la seva distribució intracel·lular en resposta a la glucosa. En aquest cas, l'enzim s'acumula al nucli en absència de glucosa i passa a tenir una disposició citoplasmàtica en presència d'aquest monosacàrid. Com hem exposat en l'apartat anterior, molt recentment s'ha demostrat que existeix una redistribució cel·lular de la GS i GP en múscul esquelètic *in vivo* a l'inici de la resíntesi de glicogen (Prats *et al.* 2005).

S'ha demostrat que defectes en el metabolisme del glicogen estan associats a certes patologies humanes. Així, l'absència o disminució de l'activitat GS en fetge dona lloc a una malaltia associada a l'emmagatzematge de glicogen anomenada GSD 0. Aquesta malaltia apareix en nadons o infants de curta edat i provoca hipoglicèmia, hiperetonèmia i baixes concentracions de lactat en sang durant el dejú. La ingesta disminueix aquests símptomes però provoca hiperglicèmia i hiperlacticèmia. El fenotip és degut a diferents mutacions en el gen hepàtic de la GS (GSY2) que provoquen una pèrdua de funció (Orho *et al.* 1998).

En la patologia de la diabetis mellitus no dependent d'insulina (NIDDM), s'observa un menor emmagatzematge de glicogen. A més, en pacients amb una resistència a la insulina precoç, la primera disfunció relacionada amb l'acció de la insulina afecta el transport i l'emmagatzematge de glucosa en forma de glicogen en el múscul (Warram *et al.* 1990), fet que esdevé un punt crític en la patofisiologia de la malaltia.

D'altra banda, la sobreexpressió d'una forma activada de la GS muscular (Manchester *et al.* 1996; Azpiazu *et al.* 2000) en ratolins transgènics es correlaciona amb una sobreacumulació de glicogen. Aquestes observacions posen de manifest la importància d'aquest enzim en el metabolisme de la glucosa en la cèl·lula.

Com s'ha demostrat, la GS és un enzim clau en l'acumulació de glicogen i en la homeòstasi de la glucosa en la cèl·lula. Una aproximació estructural al millor coneixement de la seva regulació i mecanisme catalític ens pot ajudar a entendre el seu funcionament.

## 2. GLICOGEN SINTASES:

Les GS eucariotes, bacterianes i d'arqueons pertanyen a la superfamília de les glicosiltransferases (GT) i catalitzen essencialment la mateixa reacció: l'addició de residus de glucosa mitjançant enllaços  $\alpha$ -1,4 a una cadena creixent de polisacàrid. De tota manera, a causa de que no presenten una identitat de seqüència significativa, utilitzen diferents donadors de residus glicosil (UDPG o ADPG) i altres diferències, s'han classificat en famílies diferents dins del conjunt de les GT i històricament s'han estudiat per separat.

Així doncs, les  $\alpha$ -1,4-glucà sintases han estat classificades fonamentalment en funció de la similitud de seqüència en GS, integrants de la família 3 de les GT, i MS i midó sintases (MS), integrants de la família 5 de les GT (Campbell *et al.* 1997; Campbell *et al.*

1998). Els enzims d'animals i fongs pertanyen al primer grup, mentre que els de bacteris, arqueons i plantes es classifiquen en el segon grup. Totes operen amb retenció de configuració del carboni anomèric del residu de glucosa transferit.

## 2.1. Glicogen sintases de mamífers i fongs:

Com a característiques comunes d'aquests enzims podem destacar que usen exclusivament UDP-glucosa com a donador de sucre i que la seva activitat està altament regulada. D'una banda, aquests enzims són inactivats per fosforilació en residus serina i treonina localitzats en els extrems N- i/o C-terminals (Roach 1991) i de l'altra, són activats al·lostèricament per Glc 6-P, de tal forma que aquest metabòlit pot arribar a contrarestar la inactivació produïda per fosforilació (Villar-Palasi i Guinovart 1997).

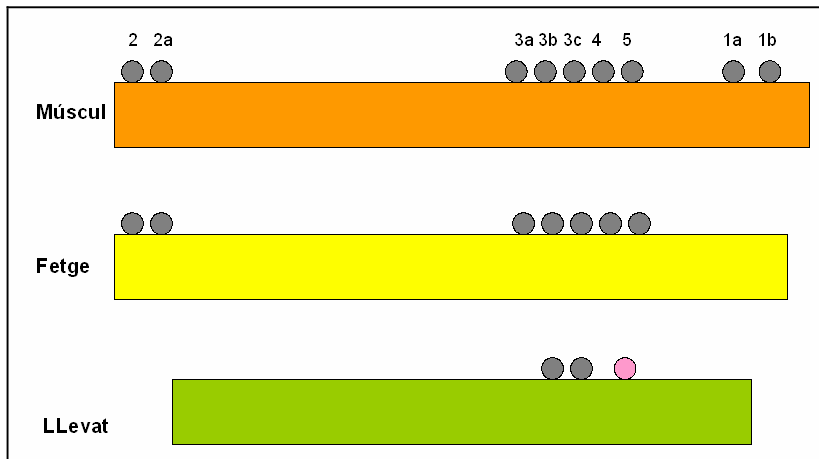
El llevat *Saccharomyces cerevisiae* conté dues isoformes de la GS, Gsy1p i Gsy2p, de les quals la Gsy2p sembla ser determinant en l'acumulació de glicogen (Farkas *et al.* 1991). Ambdós isoenzims només es fosforilen en la regió C-terminal. Concretament en Gsy2p s'han identificat els residus Ser-650, Ser-654 i Thr-667 com punts de regulació per fosforilació (Hardy i Roach 1993).

En mamífers han estat descrites dues isoformes de la GS, l'hepàtica i la muscular, que provenen de dos gens diferents. La seva existència ha estat demostrada per aproximacions bioquímiques (Camici *et al.* 1984; Hardy i Roach 1993), immunològiques (Kaslow *et al.* 1985) i per clonació dels seus respectius cDNA (Browner *et al.* 1989; Bai *et al.* 1990; Nuttall *et al.* 1994). La isoforma de múscul s'expressa en la majoria de teixits: múscul esquelètic, teixit adipós, cor, cervell, ronyó, mentre que la isoforma hepàtica sembla ser exclusiva de fetge. L'homologia entre la GS humana de fetge i la de múscul és d'un 70%, aproximadament. Ambdós isoenzims presenten la màxima divergència en l'extrem carboxi-terminal, que és 34 aminoàcids més llarg en l'isoenzim muscular (Nuttall *et al.* 1994).

Les GS de mamífers presenten una regulació complexa per fosforilació en múltiples residus, combinat amb l'acció al·lostèrica de la Glc 6-P. Mentre la fosforilació associada amb la inactivació de l'enzim redueix la  $V_{\max}$  en la isoforma hepàtica, en la muscular redueix l'afinitat envers el seu substrat, UDP-glucosa. Fins ara, s'ha descrit la fosforilació de nou residus de serina (1a, 1b, 2, 2a, 3a, 3b,3c, 4 i 5) en la GS muscular de conill que corresponen a les serines 697, 710, 7, 10, 640, 644, 648, 652 i 656, respectivament (Skurat i Roach 1995), mentre que la GS hepàtica només presenta homologia de seqüència en set d'aquests residus, ja que no posseeix la regió C-terminal on es troben els llocs de fosforilació 1a i 1b (Bai *et al.* 1990).

Existeixen moltes quinases implicades en la inactivació de la GS. Així, s'ha descrit que el lloc 2 és fosforilat per l'acció de la quinasa dependent d'AMPc (PKA) (Cohen i Hardie 1991). Posteriorment es podria fosforilar el lloc 2a per caseïna quinasa-1, ja que reconeix el motiu fosforilat prèviament Ser(P)-X-X-Ser (Skurat *et al.* 1994). Els altres residus importants en la regulació de la GS, a part de les posicions 2 i 2a, són els llocs 3a i 3b, que serien fosforilats per la glicogen sintasa quinasa-3 (GSK-3). Prèviament,

s'hauria de fosforilar la posició 5 per acció de la caseïna quinasa-2 perquè la GSK-3 pugui fosforilar seqüencialment els llocs 4, 3c, 3b i 3a (Fiol *et al.* 1990). De la seva banda, la fosforilació dels llocs 1a i 1b no afecta aparentment l'activitat de l'enzim i es desconeix la seva funció.



**Fig.3. Llocs de fosforilació de les glicogen sintases.** Esquema dels llocs de fosforilació de la GS de múscul de conill ( 2, 2a, 3a, 3b, 3c, 4,5, 1a, 1b ) i comparació amb els llocs equivalents a fetge de rata i llevat. La GS de fetge no posseeix els llocs 1a i 1b. La GS de llevat només conserva les fosforilacions similars a 3a i 3b i en una treonina (rosa).

Tant el control per fosforilació com l'al·lostèric responen a estímuls metabòlics de la cèl·lula. La GS hepàtica respon a increments en la concentració de glucosa sanguínea i a senyals hormonals, com la insulina i el glucagó, i la isoforma muscular a la contracció i a senyals hormonals, concretament a la insulina i a l'adrenalina.

L'enzim de llevat Gsy2p presenta aproximadament un 50% d'identitat de seqüència amb les GS de mamífer. Com ja hem introduït, la Gsy2p és inactivada per fosforilació en tres residus, dos dels quals (Ser-650 i Ser-654) són similars als residus de mamífer Ser-640 i Ser-644. De tota manera, la conservació de seqüència en aquesta zona és molt baixa entre ambdós enzims i el mecanisme plantejat de fosforilació ordenada de CK-2 i GSK-3 per a mamífers no sembla ser aplicable a llevat. Una altra diferència és que la proteòlisi parcial de la GS de mamífer desfosforilada i activa dona lloc a un enzim amb un comportament equivalent a la forma fosforilada (Takeda i Larner 1975). En llevat, en canvi, la proteòlisi parcial de la forma fosforilada (Huang i Cabib 1974) o l'eliminació per mutagènesi dirigida de l'extrem C-terminal a partir del residu 643 (Hardy i Roach 1993) activen la Gsy2p. La presència d'un major nombre de centres de fosforilació incloent-t'hi l'extrem N-terminal fa pensar que la regulació de l'enzim de mamífer serà més complexa que en l'enzim de llevat Gsy2p.

A l'inici d'aquesta tesi, les bases estructurals per a l'activació de la Glc 6-P continuaven sent desconegudes. Pederson i col·laboradors (Pederson *et al.* 2000) van construir dos mutants completament insensibles a Glc 6-P: el mutant R586A/R588A/R591A, que manté l'activitat i, al igual que l'enzim salvatge, és inactivat per fosforilació, i el mutant R579A/R580A/R582A, l'activitat del qual és, a més a més, insensible a fosforilació. A l'analitzar els dos mutants equivalents en l'enzim de mamífer,

que també són insensibles a la presència de Glc 6-P, es va arribar sorprenentment a resultats oposats als obtinguts per a llevat, l'activitat del mutant R585A/R587A/R590A no s'alterava per fosforilació mentre que el mutant R578A/R579A/R581A sí que era inactivat per fosforilació (Hanashiro i Roach 2002). Així, tot i que es demostra que aquesta zona de la GS rica en arginines és clau per a la seva regulació al·lostèrica per Glc-6P i per fosforilació tan en llevat com en mamífer, existeixen subtils diferències entre les funcions de cada residu en cada enzim.

## 2.2. Glicogen sintases de bacteris i midó sintases de plantes:

Aquests enzims presenten diferències significatives respecte a la família de les GT-3: utilitzen ADP-glucosa com a donador de residus glicosil, la seva activitat no està regulada per fosforilació, no presenten activació al·lostèrica per Glc-6P i no són l'etapa limitant en la síntesi del corresponent polisacàrid. De fet, com ja s'ha indicat abans, l'enzim limitant és en aquest cas l'ADP-glucosa pirofosforilasa, el qual regula la disponibilitat d'ADP-glucosa a la cèl·lula.

Aquests enzims tenen una mida menor (50 kDa aproximadament) i la similitud de seqüència amb els seus homòlegs de mamífer és marginal, tot i que la zona que correspon al centre catalític està molt conservada.

Molt recentment s'ha descrit l'estructura de la GS d' *Agrobacterium tumefaciens*, AtGS (Buschiazzo *et al.* 2004), el primer enzim resolt de la família de les GS. L'anàlisi d'aquest enzim també ha permès suggerir que, a diferència de les GS eucariotes, les GS bacterianes catalitzen la iniciació de la síntesi de glicogen, a més de la seva elongació (Ugalde *et al.* 2003).

## 2.3. Glicogen sintases d'arqueons:

S'ha descrit l'existència de GS en diferents tipus d'arqueons com ara *Sulfolobus acidocaldarius* (Cardona *et al.* 2001), *Thermococcus hydrothermalis* (Gruyer *et al.* 2002) i *Methanococcoides methylutens* (Maitra *et al.* 2001), entre d'altres.

Tot i que la seqüència de les GS d'arqueons és més semblant a les GS bacterianes (20-40%), fet que ha conduït a la seva classificació en la mateixa família de glicosiltransferases (GT-5), s'ha descobert recentment que les GS de *Sulfolobus acidocaldarius* (Cardona *et al.* 2001), *Thermococcus hydrothermalis* (Gruyer *et al.* 2002) i *Pyrococcus furiosus* (Zea *et al.* 2003) poden d'utilitzar de forma indiferent ADPG i UDPG com a donadors de glucosil.

De la relació filogenètica entre els enzims de la família GT-3 i GT-5 es desprèn que les GS d'arqueons són el grup d'aquesta segona família més proper a les GS eucariotes i per tant representen un nexa d'unió entre les dues (Cid *et al.* 2002). Les GS i les MS haurien evolucionat a partir d'un ancestre comú capaç d'utilitzar tant UDPG com ADPG com a fonts de grups glicosil, i mentre les GS eucariotes, les GS bacterianes i les MS de



plantes haurien desenvolupat especificitats exclusives per un o altre donador, les GS d'arqueons mantindrien aquesta dualitat.

La comparació entre membres de les famílies 3 i 5 mitjançant estudis de predicció d'estructura ha permès determinar que comparteixen un plegament similar i molt probablement un mecanisme catalític comú tot i la baixa similitud de seqüència. Atès que totes les  $\alpha$ -1,4-glucà sintases comparteixen un nucli estructural comú i la GS de *Pyrococcus abyssi* (PaGS) és el membre més petit d'aquesta superfamília, aquest enzim representa la unitat catalítica mínima de les GS (Cid *et al.* 2002).

Tot i l'existència de nombrosos estudis enfocats a l'anàlisi de la regulació de les GS, en particular de les de mamífer, el mecanisme catalític d'aquesta família d'enzims continua essent desconegut. Aquesta manca de coneixement és deguda principalment a la impossibilitat de produir quantitats suficients de proteïna funcional i pura per realitzar estudis estructurals (Ferrer i col·laboradors, resultats no publicats).

La reduïda mida de la PaGS, que continua mantenint la informació essencial necessària per a la seva funció catalítica, la fa una candidata excepcional per poder abordar nous estudis estructurals que condueixin a la caracterització mecanística d'aquesta superfamília d'enzims. S'afegeix el fet de que el seu caràcter termoestable ha de facilitar, *a priori*, els experiments d'expressió heteròloga i posterior purificació.

### 3. GLICOSILTRANSFERASES:

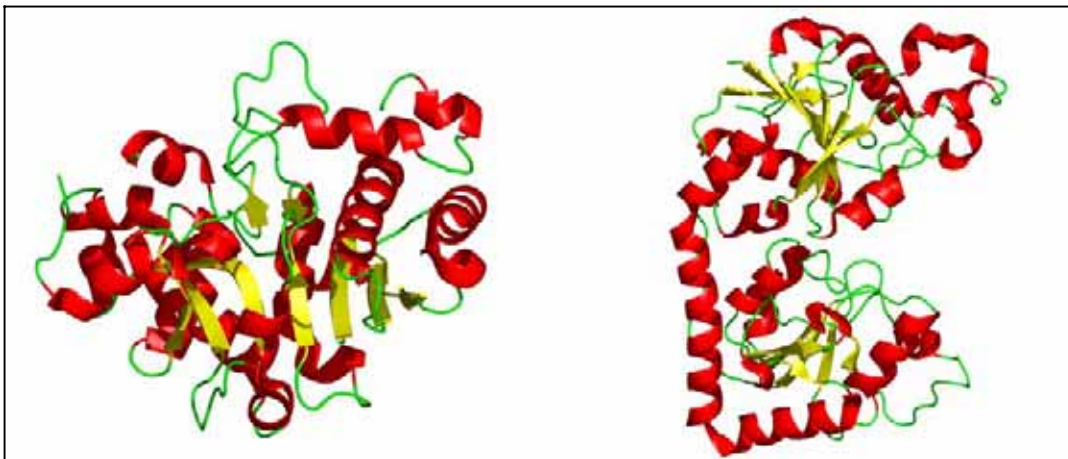
Atesa la dificultat per a obtenir suficient quantitat de l'enzim per a encetar estudis estructurals, Cid i col·laboradors van utilitzar les similituds de seqüència amb altres membres de la superfamília de les glicosiltransferases que actuen amb retenció de configuració per identificar residus claus i així desxifrar noves claus del mecanisme desconegut de les GS (Cid *et al.* 2000). D'aquesta manera s'adonaren que un motiu format per dos glutàmics separats per set aminoàcids (E-X<sub>7</sub>-E) present a la família GT-4 també estava conservat en les GS. Aquest estudi relacionava molt estretament les GS amb altres membres de la superfamília de les GT.

Les glicosiltransferases (GT) catalitzen la transferència d'un grup glicosídic des d'un donador activat a un acceptor específic. Aquest últim pot ser un oligosacàrid, una proteïna, un lípid o una cadena de DNA. Com a molècules donadores poden actuar polisacàrids, fosfats de sucres o, més freqüentment, glucòsids de nucleòtids. Aquestes reaccions són en termes quantitius, les transformacions biològiques més importants a la biosfera, ja que són les responsables de sintetitzar les macromolècules que constitueixen la major proporció de biomassa (Sinnot 1990). D'altra banda, la biosíntesi d'oligosacàrids i de glicoconjugats és fonamental, ja que aquestes molècules presenten una gran diversitat i un ampli espectre de funcions que engloba papers estructurals, de reserva energètica i fins i tot reconeixement molecular i senyalització (Dwek 1996). Tot i la seva rellevància, les GT han estat fins molt recentment una àrea poc coneguda en termes mecanístics i estructurals.

En base a les similituds de seqüència, més de 7200 seqüències de GT s'han classificat en 78 famílies (Maig, 2005) (Campbell *et al.* 1997; Coutinho 1999; Coutinho *et al.* 2003) (Coutinho, P.M. & Henrissat, B. (1999) Carbohydrate-Active Enzymes server at URL: <http://afmb.cnrs-mrs.fr/CAZY/>). De la mateixa forma que les glicosidases, les GT també s'han classificat atenent a les estereoquímiques relatives del seu substrat i del producte de la reacció que catalitzen. Així, els enzims operen amb retenció de configuració quan l'estereoquímica del carboni anomèric del sucre transferit és la mateixa en el donador de glicosil que en el producte de la reacció, i amb inversió de configuració en el cas contrari.

La informació estructural de les GT és en la majoria dels casos molt recent. Una característica clau de les glicosiltransferases és que totes les proteïnes resoltes s'engloben en només dos plegaments fonamentals anomenats GT-A i GT-B. A més, l'anàlisi de seqüències suggereix que aquesta situació continuarà sent certa per a les GT de la resta de famílies, l'estructura de les quals encara no ha estat resolta (Wrabl i Grishin 2001). S'ha suggerit que l'existència de només dos tipus de plegaments es pot deure a les limitacions imposades per l'existència d'un lloc d'unió a nucleòtid o per un possible origen evolutiu comú a partir d'unes poques seqüències precursors (Coutinho *et al.* 2003).

El plegament tipus GT-A comprèn un únic domini amb plegament de Rossmann i ha estat identificat tant en enzims que actuen amb retenció de configuració com amb inversió. Els enzims amb plegament tipus GT-B tenen dos dominis clarament diferenciats amb plegament de Rossmann i també inclouen enzims que operen amb retenció i amb inversió de la configuració anomèrica.



**Fig.4. Plegaments GT-A i GT-B.** Esquema dels dos tipus de plegaments generals de les glicosiltransferases. A l'esquerra es representa el plegament tipus GT-A de la  $\alpha$ -1,3-galactosiltransferasa. A la dreta, el plegament tipus GT-B de la  $\beta$ -glicosiltransferasa del bacteriòfag T4.

S'ha resolt l'estructura d'onze glicosiltransferases que actuen amb inversió de configuració: la  $\beta$ -glicosiltransferasa del bacteriòfag T4 (Vrieling *et al.* 1994), la transferasa de glucòsid de difosfonucleòtid SpA de *Bacillus subtilis* (Charnock i Davies 1999), la  $\beta$ -4-galactosiltransferasa bovina (Gastinel *et al.* 1999), l'enzim MurG d' *E.coli*

(Ha *et al.* 2000), la glucuroniltransferasa I (Pedersen *et al.* 2000), la transferasa de N-acetilglucosamina de conill (Unligil *et al.* 2000), la glicosiltransferasa GtfB (Mulichak *et al.* 2001), la glicosiltransferasa GtfA (Mulichak *et al.* 2003), la uridina fosforilasa (Burling *et al.* 2003), la sialiltransferasa CstII (Chiu *et al.* 2004) i la glicosiltransferasa GtfD (Mulichak *et al.* 2004), representatives de les vuit famílies GT-1, -2, -7, -9, -13, -28, -42, -43, -63.

Disposem de deu exemples d'enzims que actuen amb retenció de configuració. Tres d'aquestes GT es classifiquen dins de la família de les GT-6, són l' $\alpha$ -1,3-galactosiltransferasa bovina, que intervé en la síntesi de l'epítot  $\alpha$ -Gal (Gastinel *et al.* 2001), una N-acetilgalactosaminiltransferasa i una galactosiltransferasa humanes, responsables de l'última etapa de síntesi dels antígens A i B del grup sanguini respectivament (Patenaude *et al.* 2002). A la família GT-8 pertanyen l' $\alpha$ -1,4-galactosiltransferasa del bacteri *Neisseria meningitidis*, que participa en la síntesi de lipooligosacàrids (Persson *et al.* 2001) i la glicogenina de múscul de conill, implicada en la síntesi de novo del glicogen (Gibbons *et al.* 2002). L'enzim Kre2, una  $\alpha$ -1,2-manosiltransferasa implicada en la síntesi de glicoproteïnes de la paret cel·lular (Gibbons *et al.* 2002), la ppGalNTasa-T1, un enzim que inicia la síntesi d'O-glicans mitjançant la transferència de N-acetilgalactosamina a residus de serina o treonina (Fritz *et al.* 2004) i la ExtI2, una  $\alpha$ -1,4-N-acetilhexosaminiltransferasa, que participa en la síntesi d'heparan sulfat (Pedersen *et al.* 2003) són els representants de les famílies GT-15, GT-27 i GT-64, respectivament. Tot i pertànyer a famílies diferents i no presentar una identitat de seqüència global significativa, totes aquestes GT són molt similars a nivell de plegament global, GT-A, així com en el domini d'unió a la molècula donadora de glicosil. Una altra característica en comú és la presència d'un metall divalent coordinat als carboxilats de dos residus aspàrtic del motiu DXD en el centre catalític, imprescindible per a la catàlisi (Unligil *et al.* 2000).

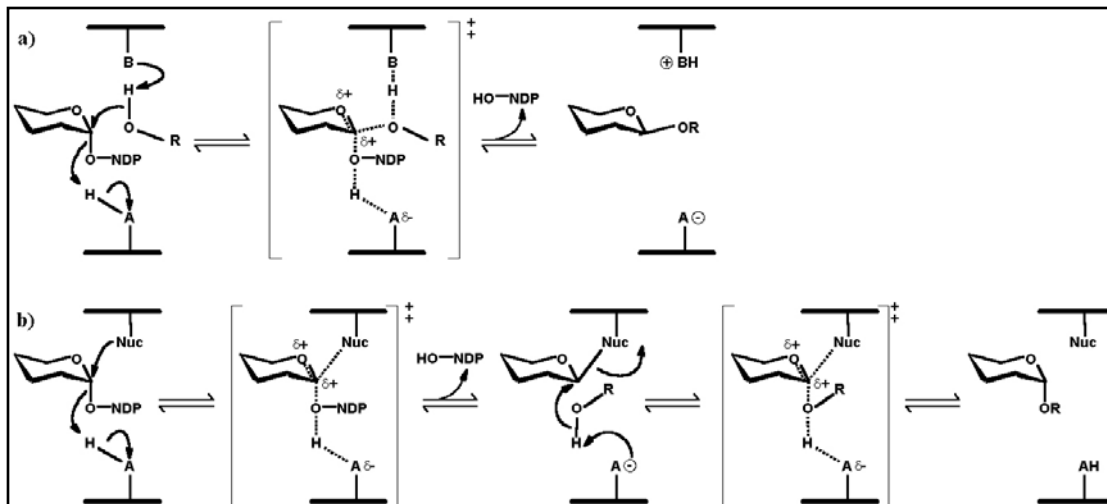
La resta de GT resoltes que retenen el sucre transferit presenten un plegament global comú, GT-B. Són la glicogen fosforilasa (Fletterick *et al.* 1976) i la maltodextrin fosforilasa (Watson *et al.* 1997) de la família GT-35, les quals utilitzen Glc 1-P com a donador, la trealosa-fosfat sintasa OtsA (Gibson *et al.* 2002) que pertany a la família GT-20, i la GS d'*Agrobacterium tumefaciens* (Buschiazzo *et al.* 2004), de la família GT-5, resolta recentment.

### 3.1. Mecanisme catalític de les glicosiltransferases (GT):

Tot i que les GS s'han pogut relacionar amb altres membres de la família de les GT, no s'ha pogut desxifrar el seu mecanisme catalític. Ja que aquest mecanisme no s'ha pogut determinar amb certesa ni tan sols per a aquells enzims dels que es coneix la seva estructura.

A la dècada passada, el nostre coneixement de les estructures i mecanismes de les glicosidases, encarregades d'hidrolitzar glicòsids, ha avançat substancialment. Per simple analogia, s'ha hipotetitzat que les GT operen mitjançant un mecanisme similar a les glicosidases, tant si inverteixen com si retenen la configuració anomèrica del sucre transferit. En el primer cas, la reacció té lloc mitjançant un mecanisme de desplaçament

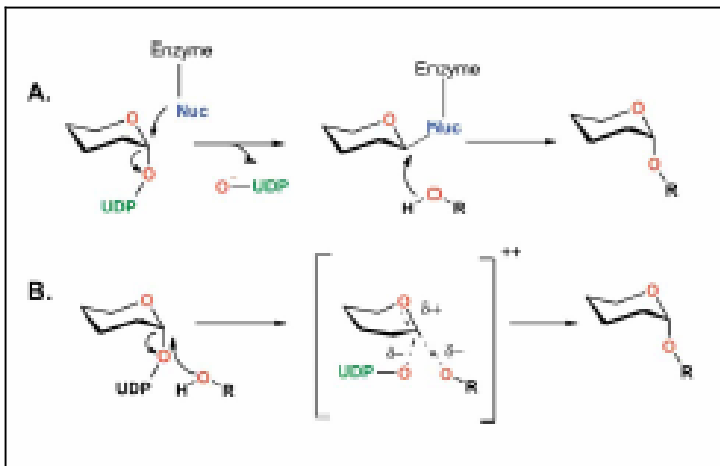
directe, en el que l'estat de transició és tipus ió oxocarbeni (Vocadlo *et al.* 2001). Aquest procés requereix la presència d'una base que assisteixi en la desprotonació de l'hidroxil reactiu de la molècula acceptora, i un residu que proporcioni alhora la catàlisi àcida per a l'heteròlisi de l'enllaç exocíclic C1"-O1". En alguns casos un ió metàl·lic unit a l'enzim pot desenvolupar aquesta segona funció.



**Fig.5. Mecanismes proposats per a les GT.** Esquema dels mecanismes proposats per a les GT que operen amb inversió (a) o retenció (b) de configuració del carboni anomèric.

En el cas de les GT que operen amb retenció de configuració, s'ha postulat que la catàlisi tindria lloc a través d'un mecanisme de doble desplaçament, amb dos estats de transició caracteritzats per dos ions oxocarbeni i que implicaria la formació i posterior destrucció d'un intermedi glicosil-enzim. Segons aquest mecanisme, un residu nucleofílic de l'enzim atacaria al carboni anomèric del sucre donador per formar el complex glicosil-enzim. La catàlisi àcida pot assistir en el trencament de l'enllaç C1"-O1", mentre que la catàlisi bàsica seria imprescindible per desprotonar l'hidroxil reactiu de la molècula acceptora.

Tot i que les estructures disponibles ratifiquen el mecanisme de desplaçament directe proposat per les GT que operen amb inversió de configuració, les dades experimentals de les quals es disposa per a les GT que retenen la configuració no permeten confirmar que el mecanisme de doble desplaçament sigui el que utilitzen aquests enzims (Persson *et al.* 2001). Per això alguns autors han proposat un mecanisme alternatiu conegut amb el nom de retorn intern "S<sub>N</sub>i" (Sinnot 1990; Persson *et al.* 2001). Aquest mecanisme seria similar al S<sub>N</sub>2 amb la diferència que l'atac del grup nucleofílic seria per la mateixa cara que el grup sortint de manera concertada, però necessàriament asíncrona (Gibson *et al.* 2002). Aquest mecanisme tampoc disposa de moment de suport experimental.



**Fig.6. Comparació entre els mecanismes possibles per a les GT que actuen amb retenció de configuració.** Esquema dels mecanismes proposats per a les GT que actuen amb retenció de configuració del carboni anomèric: model per homologia amb les glicosiltransferases de doble desplaçament (A) i model de retorn intern (B).

La determinació de l'estructura tridimensional de la GS de *Pyrococcus abyssi* podria obrir noves perspectives en l'estudi de les GS i, donades les característiques abans esmentades, podria constituir un bon model per aprofundir en el coneixement del seu mecanisme d'acció.

En aquesta memòria es mostra la caracterització i la determinació de l'estructura tridimensional de la GS de l'arqueon *P.abysyi*. Aquesta estructura serà utilitzada com a model per a ampliar el coneixement de les GS i el seu mecanisme d'acció, així com l'aplicació d'aquestes observacions a altres famílies de les GT, en especial les GS eucariotes.

# CAPÍTOL 1: CLONACIÓ, EXPRESSIÓ, PURIFICACIÓ I CRISTAL·LITZACIÓ DE LA GLICOGEN SINTASA DE *PYROCOCCUS ABYSSI*.

## INTRODUCCIÓ

Les  $\alpha$ -1,4-glucà sintases han estat classificades per similitud de seqüència en dues famílies: glicogen sintases (GS) de mamífers i fongs, que pertanyen a la família GT-3, i les GS de bacteris i arqueons i midó sintases (MS) de plantes, que pertanyen a la família GT-5 (Coutinho *et al.* 2003).

La comparació de seqüències i la predicció d'estructures ha permès determinar que els membres d'aquestes dues famílies tenen un plegament molt similar i, per tant, un mecanisme catalític previsiblement molt semblant, tot i la baixa identitat de seqüència. De tots els membres d'aquesta superfamília, la hipotètica glicogen sintasa de *P.abyssii* (PaGS) és el membre més petit descrit, el qual representa la unitat catalítica mínima de les famílies GT-3 i GT-5 (Cid *et al.* 2002).

Tot i que els complexos mecanismes de regulació d'aquests enzims han estat estudiats àmpliament, les bases moleculars d'aquests continuen sent un misteri. D'altra banda, tampoc es coneix el mecanisme catalític segons el qual actuen aquests enzims. Aquesta manca de coneixement es deu en bona part a la dificultat que suposa obtenir aquests enzims en quantitat i qualitat suficient per a poder abordar estudis estructurals. Resultats obtinguts anteriorment al nostre laboratori havien demostrat la impossibilitat d'expressar en soques recombinants d'*E. coli* quantitats suficients de les GS de mamífer de forma soluble (Ferrer i col·laboradors, resultats no publicats). A les dificultats tècniques inicials s'afegia el risc que el patró de fosforilació múltiple dels enzims eucariotes nadius tingués un paper important en el seu correcte plegament, ja que aquest estudis es van realitzar emprant soques bacterianes que no disposen dels mecanismes de fosforilació de proteïnes.

Atès que la PaGS té una mida menor que els enzims de mamífer i, a més a més, no es fosforila, constitueix un candidat idoni per encetar l'anàlisi d'aquest grup d'enzims. La major simplicitat d'aquest enzim sumat al seu caràcter termòfil, podria facilitar *a priori* els passos d'expressió i purificació.

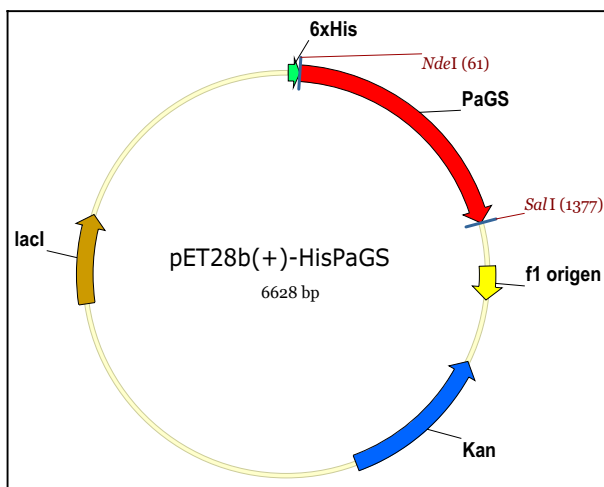
Vam iniciar doncs la clonació, expressió i purificació de la PaGS tal i com es descriu en aquest capítol. A continuació vam realitzar la caracterització de l'enzim en solució i vam determinar les condicions de cristal·lització òptimes per obtenir cristalls que difractessin a resolució atòmica.

## RESULTATS

### 1.1. Clonació de la PaGS:

Es va clonar el marc obert de lectura PAB2292 que codificava per una hipotètica GS a partir del genoma de *P.abyssi*, cedit pel Dr. Guerrero. Aquest es va amplificar per PCR, usant uns oligonucleòtids específics i l'enzim *Pfu*. Com a producte de l'amplificació, vam obtenir una sola banda de la mida esperada (1326 bp). Es procedí a la seva purificació mitjançant una reïna de bescanvi iònic i es clonà dins del vector pGFPCR. Com a senyal d'inserció el vector usa la GFP, de manera que s'analitza per restricció la presència d'insert en aquelles colònies que no presentaven fluorescència. Es realitzà una preparació a gran escala del DNA plasmídic d'una de les colònies positives i es comprovà per seqüenciació l'absència d'errors en l'insert.

El cDNA es va subclonar en el vector d'expressió pET28a(+), el qual afegia una cua de sis histidines a l'extrem amino terminal de la proteïna expressada i una diana per a la digestió amb trombina, i va donar lloc al plasmidi pET28PaGS. Es van utilitzar els enzims de restricció *NdeI* i *SalI* per a tallar específicament el vector i l'insert, creant extrems cohesius que van poder ser lligats amb l'enzim *T<sub>4</sub>* DNA lligasa. La construcció va ser verificada per seqüenciació.

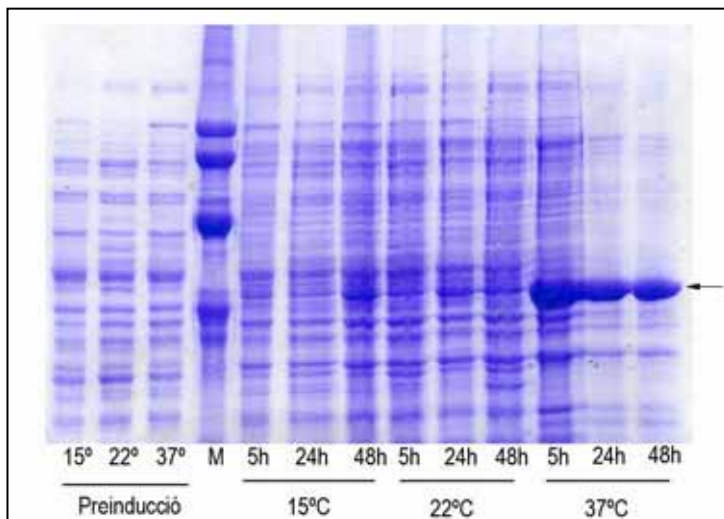


**Fig.1.1. Vector d'expressió pET28HisPaGS.** Dibuix esquemàtic del vector d'expressió pET28b(+) al qual s'ha insertat el cDNA que codifica per la PaGS.

### 1.2. Expressió de la PaGS:

Es va utilitzar el sistema d'expressió basat en la polimerasa T7 per a obtenir proteïna recombinant a partir de cultius bacterians (Studier i Moffatt 1986). Es va transformar una soca de *E. coli* BL21-CodonPlus(DE3)-RIL amb la construcció pET28PaGS i es comprovà la incorporació del plasmidi mitjançant amplificació per PCR, usant oligonucleòtids específics pel vector.

Es van realitzar proves d'expressió a 37, 22 i 15 °C, i es va observar que la proteïna recombinant s'acumulava majoritàriament a la fracció insoluble en forma d'agregats. Diversos assajos realitzats variant les temperatures i els temps d'inducció i expressió, van conduir en tots els casos a l'obtenció majoritària de la proteïna en forma insoluble. Així doncs, ens vam plantejar la purificació de la HisPaGS a partir dels cossos d'inclusió obtinguts després de créixer les cèl·lules a 37° durant 24 hores.



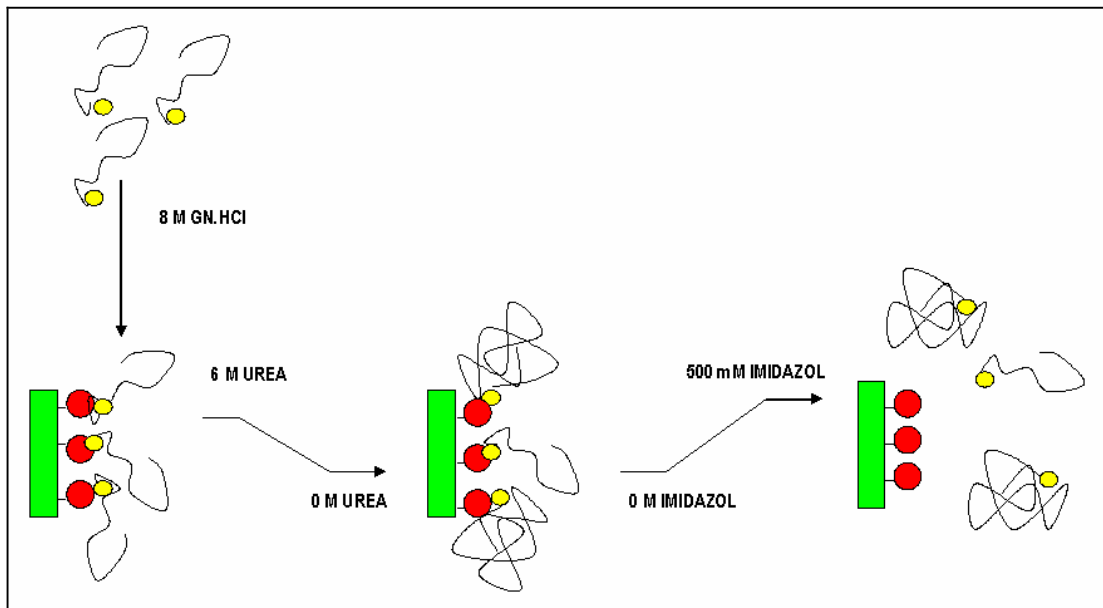
**Fig.1.2. Expressió de la HisPaGS.** Els extractes dels bacteris que havien sobreexpressat la proteïna HisPaGS a 15, 22 i 37°C a diferents temps van ser homogeneïtzats i centrifugats a 10000 g segons el descrit a l'apartat de Materials i Mètodes. Els sediments (10 µg de proteïna total) van ser sotmesos a electroforesi en un gel del 10% d'acrilamida i tenyits amb blau de Coomassie. La banda corresponent a la proteïna sobreexpressada apareix assenyalada amb una fletxa.

### 1.3. Purificació a partir de cossos d'inclusió:

Transcorregut el temps de creixement indicat, les cèl·lules es van centrifugar i trencar per sonicació. L'homogenat obtingut es va sotmetre de nou a centrifugació, es va descartar el sobrenedant i la fracció insoluble es va rentar amb una solució que contenia tritó al 2%, per tal d'eliminar proteïnes de membrana i altres contaminants. Immediatament, els agregats van ser dissolts en una solució que contenia hidròclorur de guanidini com a agent desnaturalitzant. Es va comprovar que temps llargs d'incubació (48 hores) milloraven considerablement el rendiment de la proteïna soluble.

La proteïna desnaturalitzada i solubilitzada es va carregar en una columna d'afinitat de níquel i es va sotmetre a un gradient decreixent d'urea. De la mateixa forma, es comprovà que el temps era un factor determinant en el procés de renaturalització, i es va establir un temps òptim de replegament de 12 hores. En l'etapa següent, la proteïna recombinant va ser eluïda mitjançant un gradient creixent d'imidazole. Les fraccions eluïdes es van analitzar amb gels desnaturalitzants de poliàcrilamida i es comprovà que contenien una proteïna de la massa molecular esperada (51 kDa) amb una puresa aproximada del 90%.

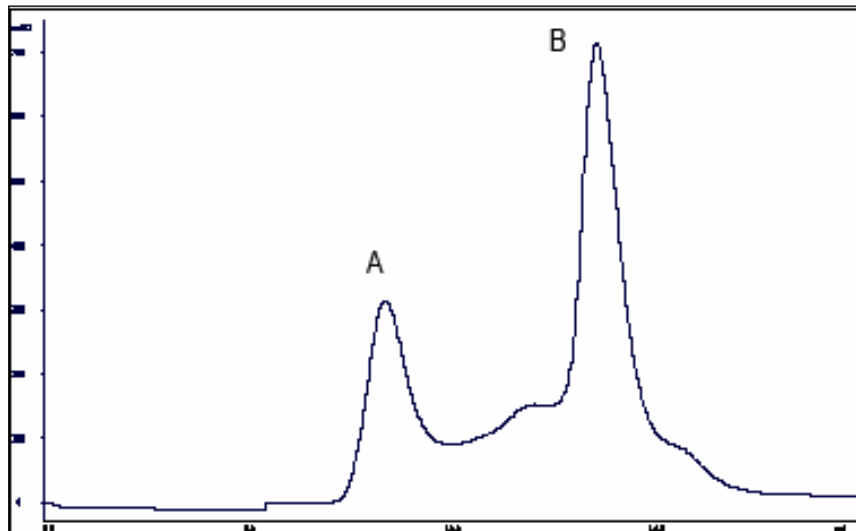




**Fig.1.3. Replegament de la HisPaGS en una columna d'afinitat per histidines.** La proteïna es mostra en línies negres, la cua d'histidines està representada per cercles grocs i la reïna amb el níquel unit per un rectangle verd amb boles vermelles.

Les fraccions positives es van concentrar i purificar per gel filtració en tampó d'elevada força iònica. En tots els casos apareixia en el volum d'exclusió un pic que quan s'analitzava en gel desnaturant de poliacrilamida presentava una massa molecular d'uns 51 kDa. Atesos els resultats experimentals, vam concloure que es tractava de proteïna recombinant que no s'havia replegat correctament, tot i que no es podia descartar que es tractés d'un estat oligomèric de massa molecular superior al límit d'exclusió de la columna. En segon lloc apareixia un pic de massa molecular aparent de 150 kDa que donava una única banda de 51 kDa en ser analitzat en gel desnaturant de poliacrilamida. Aquestes observacions suggerien que la PaGS era un homotrimer.

No vam aconseguir digerir la proteïna amb trombina per eliminar la cua d'histidines. Tot i que vam repetir l'experiment a diferents temperatures i temps d'incubació, no s'assoliren en cap cas uns rendiments raonables de digestió. La impossibilitat de tallar amb trombina es podia atribuir a la presència de traces dels agents caotròpics en el medi, els quals podien afectar a l'estabilitat de l'enzim proteolític o, més probablement, al fet que les dianes per aquesta proteasa estiguessin amagades en la proteïna replegada. Vam decidir continuar endavant tot mantenint la cua d'histidines en l'extrem N-terminal de la proteïna recombinant.



**Fig.1.4. Cromatograma de la gel filtració.** Part de la proteïna HisPaGS era eluïda amb el volum d'exclusió de la columna (A) i una part majoritària ho feia a un volum corresponent a una massa molecular estimada de 150 kDa.

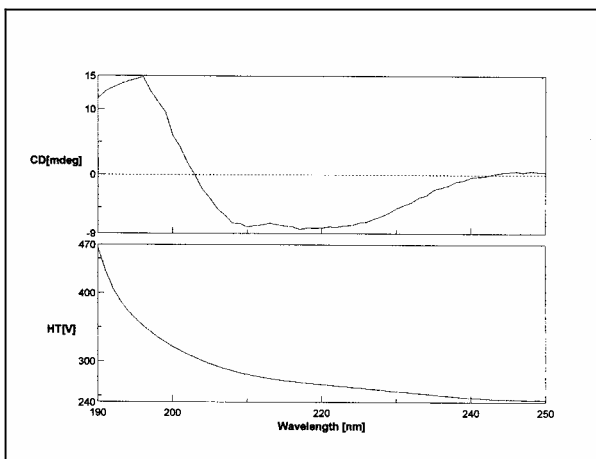
La seqüenciació pel mètode de degradació d'Edman dels primers nou aminoàcids de la mostra confirmava que es tractava de la proteïna d'interès.

Els rendiments del procés de purificació eren molt baixos ja que de mitjana obteníem de 0.3 a 0.6 mg de proteïna purificada i replegada per litre de cultiu bacterià.

## 1.4. Caracterització de la PaGS en solució:

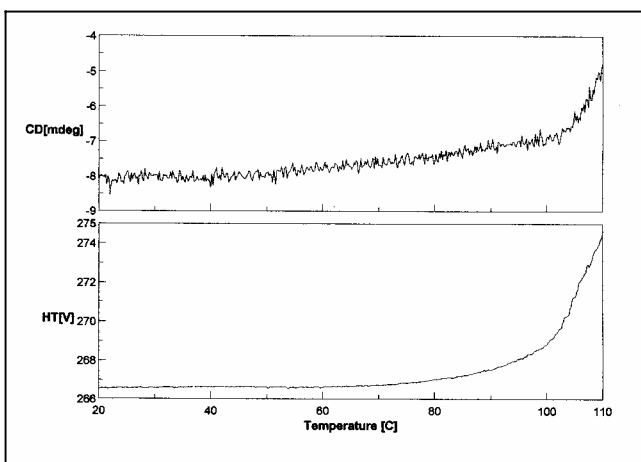
### 1.4.1. Dicroisme circular:

El procés de replegament necessari per purificar la proteïna recombinant a partir de cossos d'inclusió havia alterat dràsticament les condicions físiques de la mostra. Era necessari comprovar la seva integritat estructural i ens vam plantejar el dicroisme circular a l'ultraviolat proper. La mesura de l'elipticitat en l'interval de 190 a 250 nm va permetre determinar la presència d'elements d'estructura secundària a la proteïna. La deconvolució d'aquest espectre va donar uns valors estimats d'un contingut de 20% d'hèlix  $\alpha$  i 42% de full  $\beta$ , aproximadament.



**Fig.1.5. Espectre de dicroisme circular variant la longitud d'ona.** Espectre de dicroisme circular (mdeg) en l'interval de longitud d'ona entre 190 i 250 nm (superior) d'una solució de PaGS (0.5 mg/mL), així com la mesura del voltatge (V) en l'interval analitzat.

Per a estudiar l'estabilitat tèrmica de la proteïna recombinant, vam mesurar l'elipticitat a una longitud fixa de 220nm i vam anar incrementant la temperatura de 20 a 110°. L'estructura secundària es mantenia en augmentar la temperatura fins arribar a 105 °. A partir d'aquest punt, l'elipticitat mesurada experimentava canvis importants. Tot i així, no vam aconseguir determinar amb precisió la temperatura de transició a causa del límit de temperatura de l'aparell.

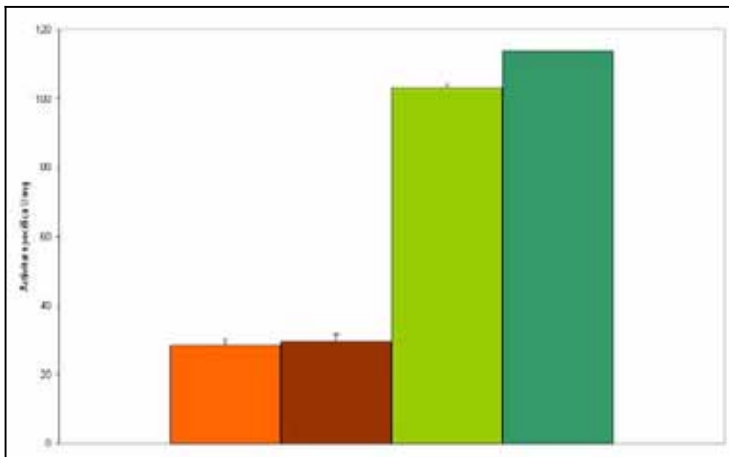


**Fig.1.6. Espectre de dicroisme circular variant la temperatura.** Variació l'elipticitat a 220 nm amb la temperatura (superior) d'una solució de PaGS (0.5 mg/mL), així com la variació del voltatge durant l'experiment.

#### 1.4.2. Activitat enzimàtica:

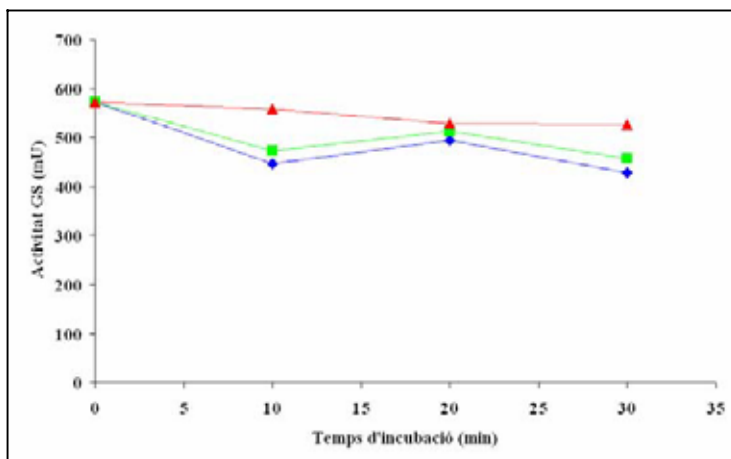
L'activitat glicogen sintasa de la proteïna replegada es va assajar a diferents temperatures, es va comprovar que l'activitat màxima s'assolia a 80°C. Aquesta activitat es va analitzar en mostres procedents de les dues últimes etapes de purificació i es va mostrar que catalitzava la transferència d'unitats glucosil a una cadena creixent de glicogen, tan des d'ADPG com d'UDPG (Fig.1.7). Treballant a concentracions

saturants de substrat, les activitats eren comparables utilitzant els dos glucosildifosfonucleòtids. L'activitat específica després del primer pas de purificació per replegament en columna d'afinitat era d'unes 30 U/mg, mentre que el pas de gel filtració augmentava aquest valor fins a 100 U/mg. Aquests darrer valor és comparable a l'obtingut per Zea i col laboradors en l'estudi de les propietats cinètiques de la GS de *P. furiosus* (Zea et al. 2003).



**Fig.1.7. Especificitat per substrat d'extractes purificats de HisPaGS.** Activitat GS mesurada a partir de la incorporació de UDP-glucosa (taronja i verd clar) o ADP-glucosa (marró i verd fosc). Mesures realitzades en mostres després de purificar per columna d'afinitat (taronja i marró) i gel filtració (verds).

Es comprovà que aquests enzims eren resistents a temperatures elevades, ja que tractaments perllongats a 80°, 90° i 100° només feien disminuir l'activitat en un 10%.

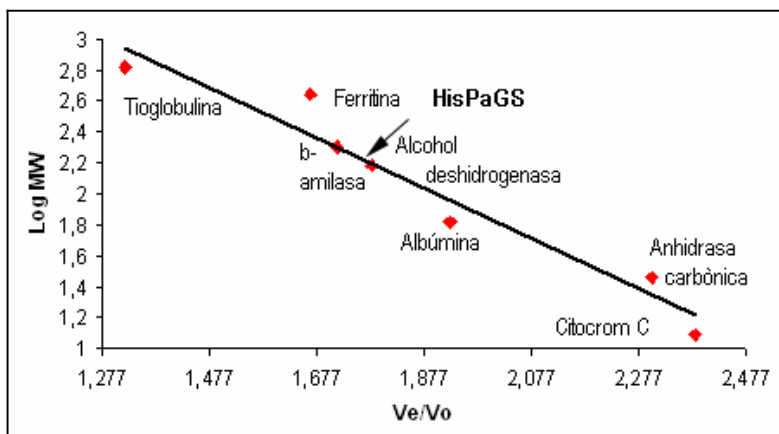


**Fig.1.8. Incubacions a diferents temperatures de la HisPaGS purificada.** Activitat de la proteïna purificada després d'incubar a 80°C (blau), 90°C (verd) i 100°C (vermell) durant diferents intervals de temps. L'activitat es va determinar a 80°C utilitzant la UDP-glucosa com a substrat.

### 1.4.3. Gels nadius i gel filtració:

Una primera aproximació a la determinació de l'estat d'oligomerització de la proteïna nadiua s'havia aconseguit a l'última etapa de purificació per cromatografia d'exclusió molecular. Es procedí a la calibració de la columna de gel filtració mitjançant

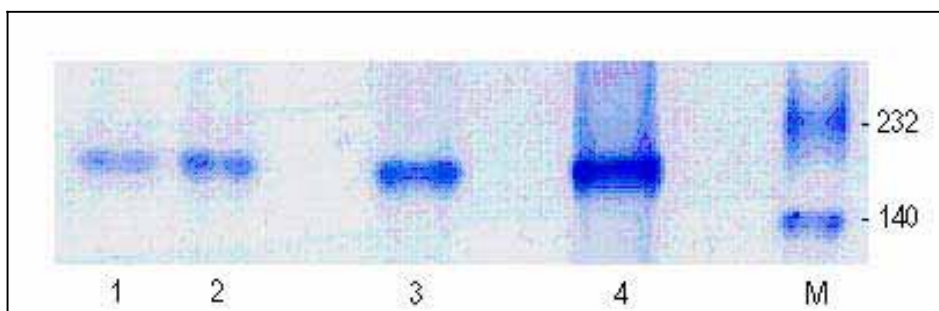
marcadors de massa molecular específics i blau de dextrà com a marcador del volum d'exclusió de la mateixa. Es representà el logaritme de les masses moleculars enfront el quocient entre el volum al qual s'el·luïa cada marcador i el volum d'exclusió de la columna. Un cop obtinguda la corba de calibració es determinà que la mostra problema apareixia com un únic pic d'elució corresponent a una massa molecular propera als 150 kDa.



**Fig.1.9. Calibració de la columna de GF.** Corba de calibració obtinguda utilitzant el paquet de marcadors de pes molecular, tal i com s'exposa a l'apartat de Materials i Mètodes.

Aquest experiment es repetí després de 24, 48 i 72 hores de mantenir la mostra a temperatura ambient i s'obtingué el mateix resultat. Aquests resultats indicaven que aparentment la PaGS és un homotrimer i que no té tendència a agregar-se en les condicions assajades.

Paral·lelament s'analitzà la proteïna purificada en gels nadius de poliàcrilamida. En totes les concentracions assajades, obteníem una única banda corresponent a una massa molecular aproximada de 150 kDa .

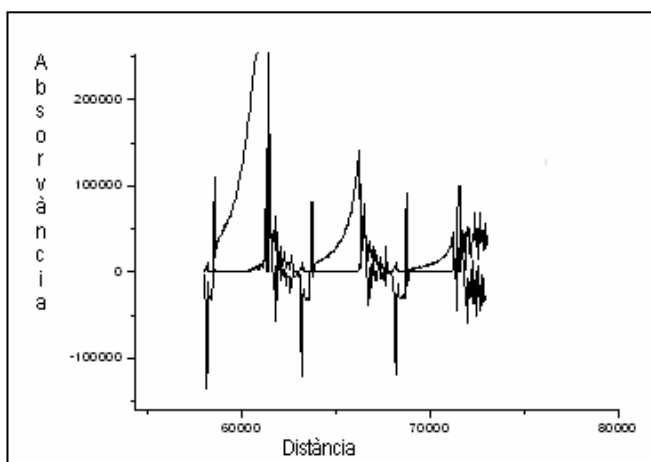


**Fig.1.10. Gel nadiu.** Es carregaren en un gel nadiu diferents quantitats de PaGS purificada: 3, 5, 10, 20 µg de proteïna en els carrils 1, 2, 3 i 4 respectivament. Els gels van ser sotmesos a electroforesi i tenyits amb blau de Coomassie, tal i com s'indica a l'apartat de Materials i Mètodes.

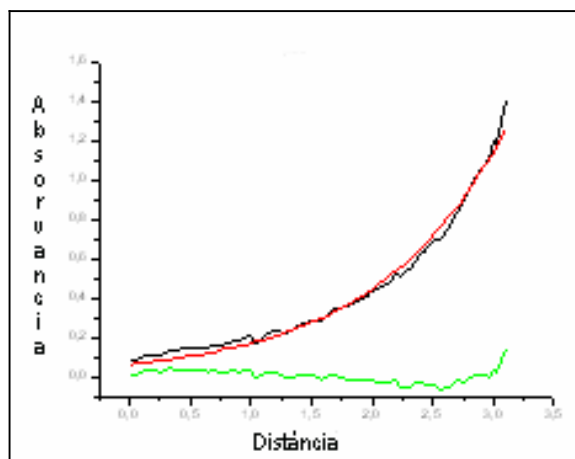
#### 1.4.4. Ultracentrifugació analítica:

Per a poder determinar amb major precisió la massa molecular de la PaGS vam utilitzar el mètode de l'equilibri de sedimentació amb ultracentrífuga analítica a tres concentracions de mostra purificada (0.1, 0,5 i 1 mg/ml). Un cop vam obtenir les distribucions d'equilibri a cadascuna de les concentracions, vam ajustar les corresponents corbes a un model exponencial i vam aplicar els valors calculats a l'equació teòrica.

Els tres casos s'ajustaven molt bé a una equació de primer ordre, indicatiu d'una única espècie en solució. Es va arribar al mateix valor experimental de 158700 Da, massa molt propera al valor teòric de 153000 Da del trímer.



**Fig.1.11. Ultracentrifugació analítica.** Representació de l'absorbància en funció de la distància radial de tres mostres de proteïna un cop assolit l'equilibri de sedimentació. Les concentracions corresponen d'esquerra a dreta a 1, 0.5 i 0.1 mg/ml respectivament.

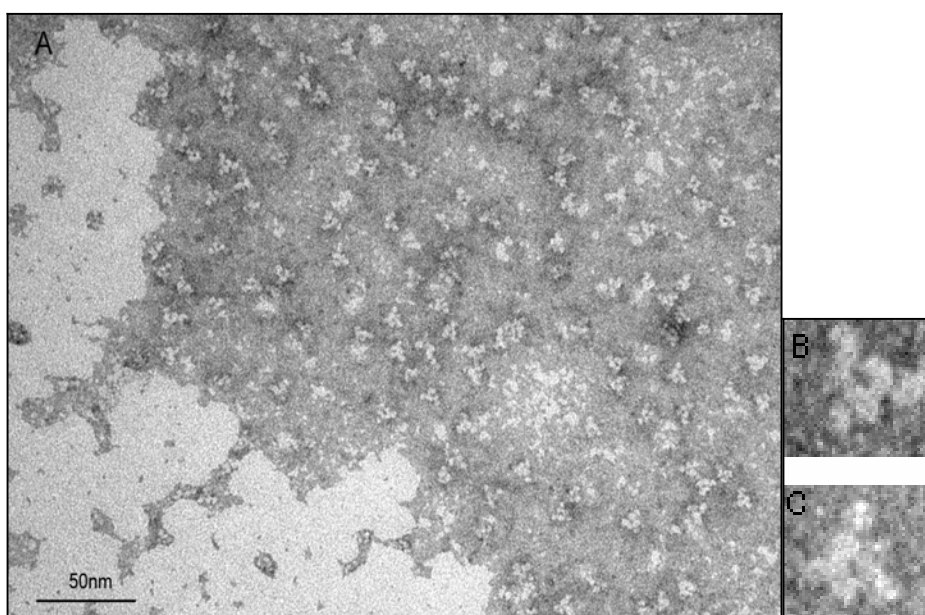


**Fig.1.12. Corba d'equilibri de sedimentació.** Corba experimental obtinguda per la PaGS a una concentració de 0.5 mg/ml (negre). Corba teòrica ajustada a la corba experimental (vermell) i error entre la corba teòrica i la corba experimental (verd).

### 1.4.5. Microscòpia electrònica

La tècnica de microscòpia electrònica per tinció negativa fou emprada per determinar l'estructura de la proteïna a baixa resolució. L'observació al microscopi a 300,000 augments de les mostres dipositades sobre una reixeta carregada i tenyides amb sal de tungstè va proporcionar una primera indicació de la forma molecular de l'enzim.

S'observa a la figura 1.13 la forma triangular plana amb uns 130 Å de costat que és compatible amb un estat d'oligomerització trimèric de la PaGS. Es tracta d'una estructura molt estesa en que cadascun dels monòmers està format per dues subunitats amb un solc molt profund entre elles.



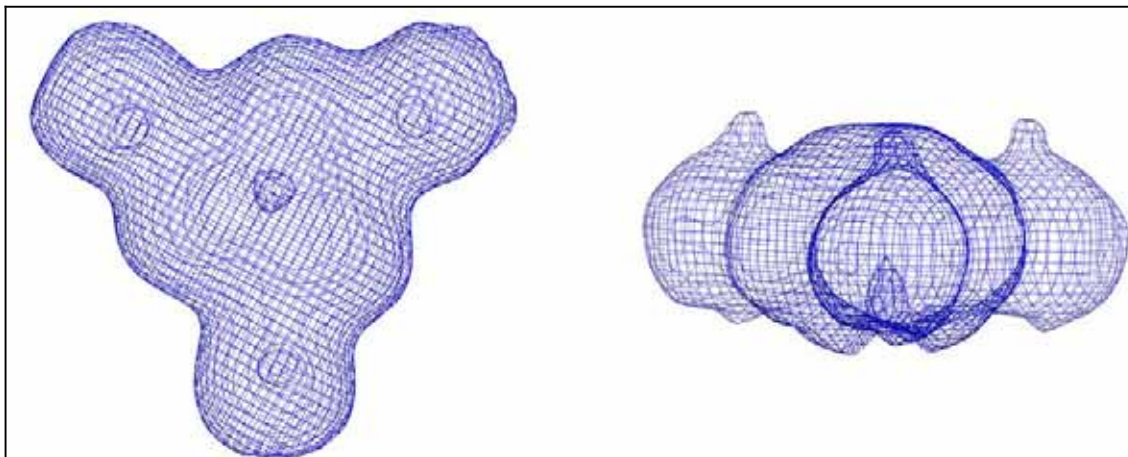
**Fig.1.13. Imatges de microscòpia electrònica de la PaGS.** Imatges a 300.000 augments de la tinció negativa de les mostres de PaGS, segons el que s'amplia a Materials i Mètodes (A). Ampliacions de camps d'aquesta mateixa imatge en que es mostra la forma triangular de la molècula (B, C).

A partir d'aquests primers experiments, la Dra. Irina Gutsche i el Dr. Jorge Navaza es van involucrar en el projecte, i van reproduir i optimitzar la tinció negativa per a poder fer una reconstrucció tridimensional de la proteïna. En general, el límit de resolució per a una reconstrucció tridimensional de criomicroscòpia és d'uns 300 kDa, mentre que per a la tinció negativa és tres vegades menor. Fins ara, només quatre partícules de 100 a 150 kDa s'han pogut reconstruir a partir de tinció negativa, la nostra inclosa.

Per a la reconstrucció, 2670 imatges van ser escollides a mà i s'en van rebutjar 712 en base als criteris de selecció. La mà de cada imatge es va escollir després de provar a cara o creu (*flipping*) i en tots els casos donava una solució inequívoca. No s'establí cap restricció de simetria ja que la simetria ternària apareixia per si sola. En conjunt, la reconstrucció tenia una resolució de 18 Å en la direcció xy, mentre que en la direcció z

només arribava als 25 Å. Aquesta anisotropia es devia a l'adsorció preferencial de la molècula en l'orientació xy.

Els tres eixos imaginaris que relacionarien cadascun dels dos dominis de cada monòmer serien gairebé perpendiculars a l'eix trimèric molecular, tot formant un angle d'uns 110°.



**Fig.1.14. Màscara de microscòpia electrònica.** Dibuix de la reconstrucció tridimensional de la PaGS a partir de les imatges de microscòpia electrònica. Projectcions en la direcció xy (esquerra) i z (dreta).

## 1.5. Cristal·lització de la PaGS:

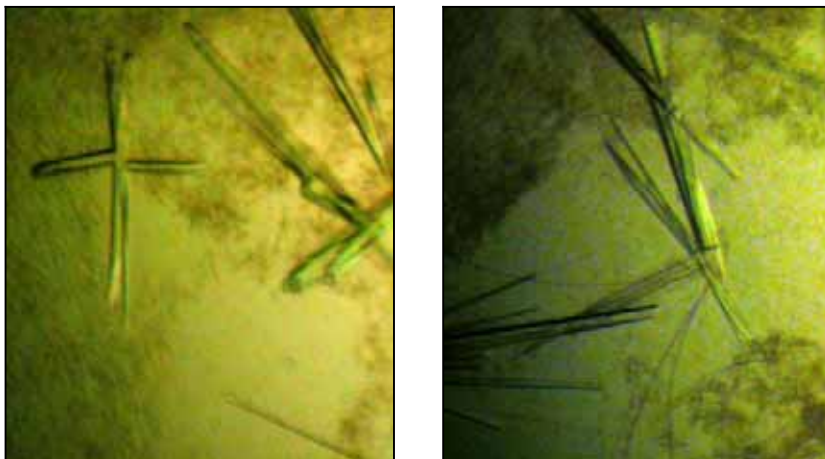
### 1.5.1. Condicions inicials:

Les mostres de proteïna purificades i concentrades fins a 6 mg/mL es van utilitzar per a determinar sistemàticament les possibles condicions de cristal·lització mitjançant la tècnica de la gota penjant, basada en la difusió de vapor. Com a base per a l'anàlisi, es van utilitzar els equips de Hampton basats en el rastreig de condicions positives per a d'altres experiments de cristal·lització anteriors. Els primers cristalls van aparèixer després de quatre dies a 20 °C en una condició del criatge que contenia citrat de sodi pH 5.6, 20% PEG 4000 i 20% isopropanol. Es comprovà que els cristalls eren efectivament de proteïna dissolent-los en el tampó de purificació i carregant-los en un gel desnaturalitzant de proteïna. La tinció del gel amb plata mostrava una única banda d'una massa molecular de 51 kDa. Els cristalls eren en realitat macles en forma de microagulles ramificades que precisaven més passos d'optimització abans de poder utilitzar-los per a estudis de difracció.

Els primers passos d'optimització es centraren en fer petites variacions de les condicions inicials de pH i concentració de precipitants. En aquest punt ens enfrontàrem a un problema afegit, ja que l'elevada concentració d'isopropanol feia gairebé impossible la manipulació de les agulles. L'optimització de les condicions ens va permetre reduir la concentració de l'alcohol fins al 9% però, tot i així, no resultava suficient ja que es continuaven produint turbulències a causa de la ràpida evaporació



de l'agent precipitant. Vam iniciar doncs, proves enfocades a substituir completament aquest alcohol i vam arribar a unes noves condicions de cristallització amb 18% de dioxà, mantenint 20% de PEG 4000 i tampó citrat. La naturalesa dels cristalls continuava sent de macla, però les agulles eren més llargues i menys ramificades. D'altra banda, gràcies a la menor volatilitat del dioxà, la manipulació d'aquests cristalls deixà de ser una limitació.



**Fig.1.15. Cristalls de PaGS.** Fotografies dels cristalls de PaGS en forma de macla obtinguts mitjançant la tècnica de la gota penjant, usant 20% PEG4000, 18% de dioxà i citrat de sodi pH 5.6, tal i com s'exposa a l'apartat de Materials i Mètodes.

### 1.5.2. Microseembra:

Tot i que havíem optimitzat les condicions, l'elevada nucleació continuava impossibilitant obtenir monocristalls de grandària suficient. Les tècniques de sembra van permetre treballar en condicions de menor sobresaturació, en què la nucleació no era espontània però sí el creixement a partir de cristalls polvoritzats, anomenats *llavors*.

Els millors cristalls es van obtenir gràcies a mantenir concentracions altes de dioxà però disminuït la concentració de PEG 4000 fins al 15 %. A partir de cada nucli o *llavor* creixien diverses agulles, però aquestes ja eren monocristalls sense ramificacions.



**Fig.1.16. Microseembra de cristalls de PaGS.** Fotografia de diferents agulles crescudes a partir d'una única llavor. Les condicions de microseembra implicaven una concentració menor de PEG 4000, tal i com s'indica a Materials i Mètodes.

### 1.5.3. Macroseembra:

Tot i l'acceptable qualitat dels cristalls anteriors, ens vam plantejar un últim pas d'optimització a fi d'augmentar la mida de les agulles. Utilitzant el mateix principi d'equilibri metaestable de l'apartat anterior, es varen emprar monocristalls com a noves *llavors*. Les condicions del licor de cristallització també es mantingueren idèntiques a l'apartat anterior. A cada gota s'hi afegia un nou cristall i després d'uns quatre dies la mida era suficient per poder encetar els estudis de difracció.



**Fig.1.17. Macroseembra dels cristalls de PaGS.** Imatge d'un monocristall crescut segons la tècnica de macroseembra exposada a Materials i Mètodes.

### 1.5.4. Cocrystal·lització:

Vam intentar obtenir cristalls de la PaGS unida als seus substrats. Com es tractava d'un altre sistema, vam tornar a fer un escombrat local de les condicions de

crystal·lització al voltant de les condicions que havien donat cristalls per la proteïna aïllada. Els substrats assajats independentment (ADPG, UDPG, maltotetraosa, maltopentosa i maltoheptosa) es van incubar amb la proteïna i es van intentar cristal·litzar. Només es van obtenir cristalls com a resultat d'incubar la proteïna amb UDPG, tot i que en aquest punt desconeixíem si la molècula s'havia unit a l'enzim. Aquests resultats ens feien pensar que a la majoria dels casos, la proteïna patiria un canvi conformacional importat en unir-se al substrat i això faria que canviessin dràsticament les condicions de cristal·lització. Aquestes observacions coincidien amb els resultats obtinguts anteriorment per a la proteïna homòloga d'*Agrobacterium tumefaciens* (Buschiazzo *et al.* 2004). En aquest cas, els cristalls de proteïna es trencaven després d'incubar-los amb els substrats.

### **1.5.5. Criopreservació:**

Un cop els cristalls havien assolit la mida esperada, es congelaven en nitrogen líquid i s'emmagatzemaven fins que eren sotmesos a difracció per raigs X. Les condicions òptimes de criopreservació consistien a l'afegir glicerol fins a una concentració final del 15%. El glicerol facilita la vitrificació de l'aigua i així s'evita la formació de cristalls de gel que destrueixen l'ordre intern del cristall.

## **1.6. Aproximacions experimentals al problema de la fase:**

Dues de les aproximacions per a poder resoldre el problema de la fase requerien cristalls que continguessin àtoms dispersors anòmals (selenometionines) o bé àtoms altament dispersors (àtoms pesats).

### **1.6.1. Selenometionines:**

Vam recórrer a l'obtenció de proteïna recombinant en la que els residus de metionina fossin substituïts per selenometionines, ja que el seleni és un dispersor anòmal adequat. Cadascun dels monòmers de la proteïna contenia set metionines a substituir, valor suficient, en principi, per poder obtenir un bon senyal anòmal.

Vam expressar la PaGS en la soca de *E. coli* BL21-CodonPlus(DE3)-RIL en medi mínim suplementat amb els aminoàcids essencials (excepte Gly, Ala, Pro, Asn, Cys i Met) i selenometionina (Van Duyne *et al.* 1993). El nivell d'expressió era inferior al que obteníem amb la proteïna nadiua i es continuava acumulant en cossos d'inclusió. La proteïna amb selenometionines es va purificar seguint el mateix protocol que s'havia utilitzat per la proteïna nadiua. De la mateixa forma, les condicions i passos de cristal·lització van ser idèntics als emprats anteriorment.

### **1.6.2. Àtoms pesats:**

Com a segona aproximació, vam intentar introduir àtoms pesats en els cristalls de proteïna nadiua. Vam submergir els cristalls en diferents solucions d'àtoms pesats a

diferents concentracions (Taula 1.1) per comprovar la seva possible derivatització. Els cristalls van ser submergits durant 24 hores, ja que controls amb temps d'incubació més curts indicaven que la majoria de cristalls suportaven el tractament. Els cristalls que van sobreviure es van assajar per analitzar la seva capacitat de difracció de raigs X. En primer lloc es provaren els cristalls que havien estat en contacte amb les solucions més concentrades ja que tenien més probabilitat d'haver incorporat el metall. D'aquells que van difractar es va intentar recollir un espectre complet per a comprovar si produïen algun canvi en la intensitat dels raigs X difractats.

Àtom pesat	Compost	Concentració	Temps	Recollida	Línia del sincrotró
Pt (IV)	K <sub>2</sub> PtCl <sub>4</sub>	5 mM	24 hores	-	ID13
		10mM	24 hores	-	
		20mM	24 hores	3.5 Å	
Pt (VI)	K <sub>2</sub> PtCl <sub>6</sub>	5 mM	24 hores	-	ID13
		10mM	24 hores	-	
		20mM	24 hores	4 Å	
Hg (II)	HgCl <sub>2</sub>	10 mM	24 hores	-	ID14-EH2
		20 mM	24 hores	3.0 Å	
Ta (V)	Ta <sub>6</sub> Br <sub>14</sub>	5 mM	24 hores	-	ID13 ID13
		10mM	24 hores	3.0 Å	
		20mM	24 hores	No difracta	
W	K <sub>6</sub> HSiW <sub>9</sub> V <sub>3</sub> O <sub>40</sub>	5 mM	24 hores	-	ID14-EH2 ID14-EH2
		10mM	24 hores	3.5 Å	
		20mM	24 hores	No difracta	

**Taula 1.1. Recollida de l'espectre de difracció dels cristalls incubats amb àtoms pesats.** Els cristalls nadius van ser incubats amb cada solució d'àtom pesat durant 24 hores a diferents concentracions (5, 10, 20 mM) i a continuació es van difractar a les línies del sincrotró ID13 i ID14-EH2, tal i com s'exposa a Materials i Mètodes.

Com es mostra a la Taula 1.1. vam poder recollir un espectre de cada cristall incubat amb diferents sals de metalls pesats. Tot i així, en aquest punt no podíem saber si els cristalls havien incorporat el metall i per tant s'havien derivatitzat. Era necessari el processament de les dades i el seu anàlisi per a determinar si el metall havia adoptat alguna posició definida dins del cristall, tal i com s'amplia a l'apartat 3.2.2.2.

## DISCUSSIÓ

La clonació, expressió i purificació del marc obert de lectura PAB2292 del genoma de *P. abyssi* ha permès determinar que aquest gen codifica en efecte per a un enzim amb activitat glicogen sintasa.

Atès que aquesta proteïna s'expressa en *E. coli* exclusivament en forma insoluble, hem desenvolupat un mètode de purificació que implica el correcte replegament de la proteïna a partir dels cossos d'inclusió desnaturalitzats. L'èxit d'aquesta tècnica obre la porta a nous projectes encaminats a purificar altres GS que també formen cossos d'inclusió quan s'expressen, com és el cas de les GS de mamífer.

De tota manera, és sorprenent que una proteïna d'una mida relativament petita i d'origen hipertermòfil tingui tanta tendència a formar cossos d'inclusió. L'anàlisi en profunditat de l'estructura que es durà a terme en els següents capítols ens ajudarà a entendre millor les bases moleculars d'aquest problema.

Un cop ben plegada, però, la PaGS és extremadament estable, com ho demostra el seguiment de la desnaturalització per temperatura dut a terme en l'experiment de dicromisme circular. Experiments paral·lels de calorimetria duts a terme per un altre membre del nostre grup i dirigits pel Dr. E. Freire van conduir a resultats molt similars (Gambetti, E. i col·laboradors, resultats no publicats). Mitjançant un experiment de calorimetria diferencial d'escombrat (DSC, *Differential Scanning Calorimetry*) van poder determinar que la temperatura de fusió  $T_m$  de la proteïna és de 105.7°C, un valor molt elevat només assolit per proteïnes d'organismes hipertermòfils.

Altres autors havien descrit que les glicogen sintases dels arqueons *T. hydrothermalis* (Gruyer *et al.* 2002) i *Sulfolobus acidocaldaius* (König *et al.* 1982) podien utilitzar UDP-glucosa i ADP-glucosa com a donadors de glicosil. Vam poder determinar que el mateix succeïa amb la PaGS, tot i que al treballar en condicions saturants no vam poder determinar si hi havia diferències d'afinitat entre ambdós substrats. Durant el decurs d'aquest projecte, Zea i col·laboradors van publicar un estudi cinètic de la GS de *P. furiosus* on mostraven que ambdós substrats tenien la mateixa constant de Michaelis ( $K_m = 4.1$  mM), tot i que la velocitat màxima era una mica superior utilitzant ADP-glucosa que d'UDP-glucosa (Zea *et al.* 2003). Aquest comportament és previsiblement molt similar al de PaGS, ja que ambdós enzims presenten un 97% d'identitat de seqüència.

El fet que aquests organismes puguin utilitzar qualsevol dels dos nucleòsids de glucosa com a donadors de residus glicosil, els situa en un terme mig entre les GS de la família GT-3, que només utilitzen UDP-glucosa, i les de la família GT-5, que utilitzen exclusivament ADP-glucosa com a donador de residus glicosil.

Diferents experiments ens han permès determinar que la PaGS és un trímer en solució. A més, la microscòpia electrònica per tinció negativa ens mostra una disposició

triangular plana de la molècula. L'obtenció d'una reconstrucció a tan alta resolució d'una proteïna de "només" 150 kDa per microscòpia electrònica permet plantejar la seva validesa per a ser usada com a model de reemplaçament molecular. Totes aquestes consideracions estructurals seran analitzades amb més detall en els capítols 3 i 4.

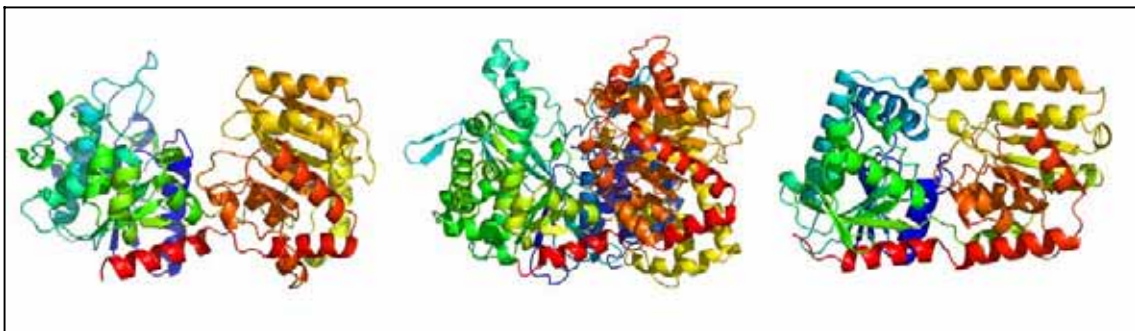
L'obtenció i optimització dels cristalls de la PaGS va permetre iniciar els estudis de determinació de l'estructura tridimensional de l'enzim per difracció de raigs X.

## CAPÍTOL 2: EXPRESSIÓ, PURIFICACIÓ I CRISTAL·LITZACIÓ DEL DOMINI C-TERMINAL DE LA PAGS

### INTRODUCCIÓ

Per alineament de seqüència i per predicció d'estructura, s'ha inferit que les GS de les famílies GT-3 i GT-5 posseeixen un plegament global tipus GT-B (Coutinho *et al.* 2003). A aquesta superfamília també pertanyen la glicogen fosforilasa (GP) (Fletterick *et al.* 1976) i la maltodextrin fosforilasa (MalP) (Watson *et al.* 1997) de la família GT-35, les quals utilitzen glucosa 1-fosfat com a donador de glucosil i la trealosa-fosfat sintasa OtsA (Gibson *et al.* 2002), que es classifica dins la família GT-20. Molt recentment s'ha resolt la GS d'*Agrobacterium tumefaciens* (Buschiazzo *et al.* 2004) de la família GT-5 i s'ha comprovat que efectivament presenta aquest tipus de plegament (Fig. 2.1).

Els enzims amb plegament tipus GT-B tenen dos dominis amb plegament de Rossmann separats per un solc molt profund en la interfase que inclou el centre catalític. En tots els casos coneguts l'última hèlix del domini C-terminal experimenta un gir fent que els últims residus de la proteïna creuin al domini N-terminal en forma d'hèlix  $\alpha$  (Wrabl i Grishin 2001).



**Fig. 2.1. Plegament tipus GT-B.** Dibuix de la AtGS (esquerre), MalP (centre) i OtsA (dreta), tres proteïnes amb plegament tipus GT-B que actuen amb retenció de la configuració del sucre transferit. La gradació de colors del blau al vermell indica la direcció N- a C-terminal.

Per tant, els dominis N- i C-terminals són molt compactes, però estan units entre ells només per dues cadenes d'aminoàcids, la cadena que va del domini N- al C-terminal i la cadena que experimenta un gir retornant del domini C- al N-terminal. Així doncs, es poden plantejar estudis de l'enzim truncat en què cada domini es pugui expressar i analitzar per separat.

L'anàlisi de les estructures resoltes del grup de les GT-B mostra que la gran majoria dels residus implicats en la unió al donador de residu glicosil es troben en el domini C-terminal de la proteïna, per tant l'expressió independent d'aquest domini pot aportar informació sobre aquesta unió.

Atès que aquest domini C-terminal és compacte i la seva mida és aproximadament la meitat que la proteïna sencera, existeixen probabilitats *a priori* que aquest domini s'expressi de forma soluble.

L'alineament de seqüència de les proteïnes del grup de les GT-B mostra que la similitud de seqüència és més gran en aquest domini C-terminal que en la proteïna sencera. Aquest fet podria facilitar la resolució del domini C-terminal de la PaGS per reemplaçament molecular utilitzant com a model inicial algun dels dominis C-terminal de les estructures resoltes. Un cop resolta l'estructura d'aquest domini C-terminal aïllat, es podria utilitzar com a model inicial per abordar la resolució de l'estructura de l'enzim sencer per reemplaçament molecular, ja que aquest seria gairebé idèntic al domini C-terminal de l'estructura de la PaGS sencera.

Per aquests motius i davant la impossibilitat de resoldre directament l'estructura de la proteïna sencera, ens vam plantejar la clonació, expressió, purificació i cristallització del domini C-terminal aïllat de la PaGS.

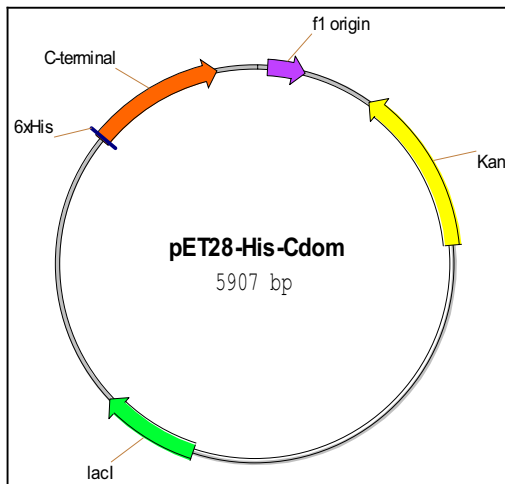
## RESULTATS

### 2.1. Clonació del domini C-terminal de la PaGS:

Per a determinar els límits del domini C-terminal, es va alinear la seqüència de la GS de *P.abysyi* amb l'alineament estructural de les diferents GT-B d'estructura coneguda que operen, com la PaGS, amb retenció de configuració (AtGS, OtsA, fosforilasa). En aquestes tres estructures les meitats N- i C-terminal formen dos dominis perfectament diferenciats, però presenten una hèlix característica a l'extrem C-terminal que retorna al domini N per formar-ne part. Per tant, aquesta hèlix no va ser considerada en la clonació del domini C-terminal de la PaGS. A partir d'aquesta anàlisi vam hipotetitzar que aquest domini C-terminal contenia des del residu 218 fins al 413.

Vam amplificar el DNA que codificava pel domini C-terminal a partir del vector pGFPCRPaGS, utilitzant oligonucleòtids específics i l'enzim *Pfu* polimerasa. Vam obtenir una única banda de la mida esperada (548 bp), la qual va ser purificada mitjançant una reïna d'intercanvi iònic i clonada en el vector pET28 emprant els enzims de restricció *NdeI* i *Sall* (pET28HisCterminal). Aquesta construcció codifica per a una proteïna a la que s'ha afegit una cua de 6 histidines i una seqüència específica de tall per trombina a l'extrem amino del domini C-terminal (Fig. 2.2). La nova construcció va ser verificada per seqüenciació.

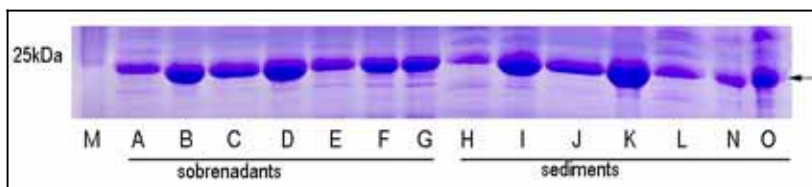




**Fig.2.2. Vector d'expressió pET28HisCterminal.** Dibuix esquemàtic del vector d'expressió pET28b(+) al qual s'ha inserat el cDNA que codifica pel domini C-terminal la PaGS.

## 2.2. Expressió del domini C-terminal de la PaGS:

Es va transformar la soca d' *E.coli* BL21-CodonPlus(DE3)-RIL amb la construcció pET28HisCterminal i es va comprovar la incorporació del plasmidi per amplificació amb oligonucleòtids específics. Es van realitzar proves d'expressió a 22 i 37 °C, i es va comprovar que en els dos casos la fracció soluble contenia aproximadament el 50% de la proteïna recombinant i que a 37°C el nivell d'expressió total era aproximadament un 20 % superior a l'obtingut a 22 °C (Fig. 2.3) .

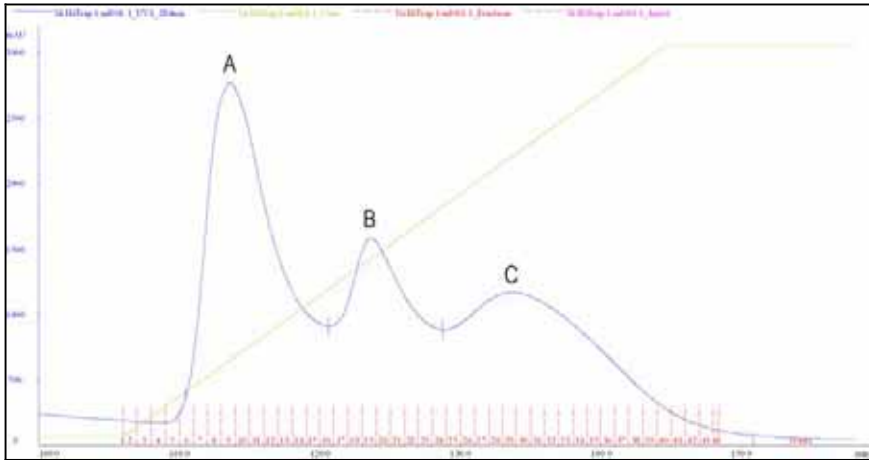


**Fig.2.3. Expressió del domini C-terminal de la PaGS.** Els extractes de bacteris en què s'havia sobreexpressat el domini C-terminal de la PaGS a 22°C (C,E,G,J,L,N) i 37°C (A,B,D,F,H,I,K,O) a diferents temperatures: preinducció (A i H), 3h (B,C,I,J), 7h (D,E,K,L) i 24 h (F,G,N,O), van ser homogeneïtzats i centrifugats a 10000 g segons el que es descriu a l'apartat de Materials i Mètodes. Els sobrenadants (A-G) i els sediments (H-O) (10 µg de proteïna total) van ser sotmesos a electroforesi en un gel del 10% d'acrilamida i tenyits amb blau de Coomassie.

## 2.3. Purificació del domini C-terminal de la PaGS:

Les cèl·lules obtingudes a partir d'un cultiu transformat amb pET28HisCterminal crescudes durant 24 hores a 37°C foren lisades per sonicació. La fracció soluble es carregà a una columna d'afinitat de níquel i es va eluir amb un gradient creixent d'imidazole. El cromatograma obtingut va indicar que existien tres poblacions de proteïnes amb diferent afinitat per la columna de níquel (Fig. 2.4). L'anàlisi en gel desnaturalitzant d'acrilamida va permetre determinar que el primer pic corresponia a

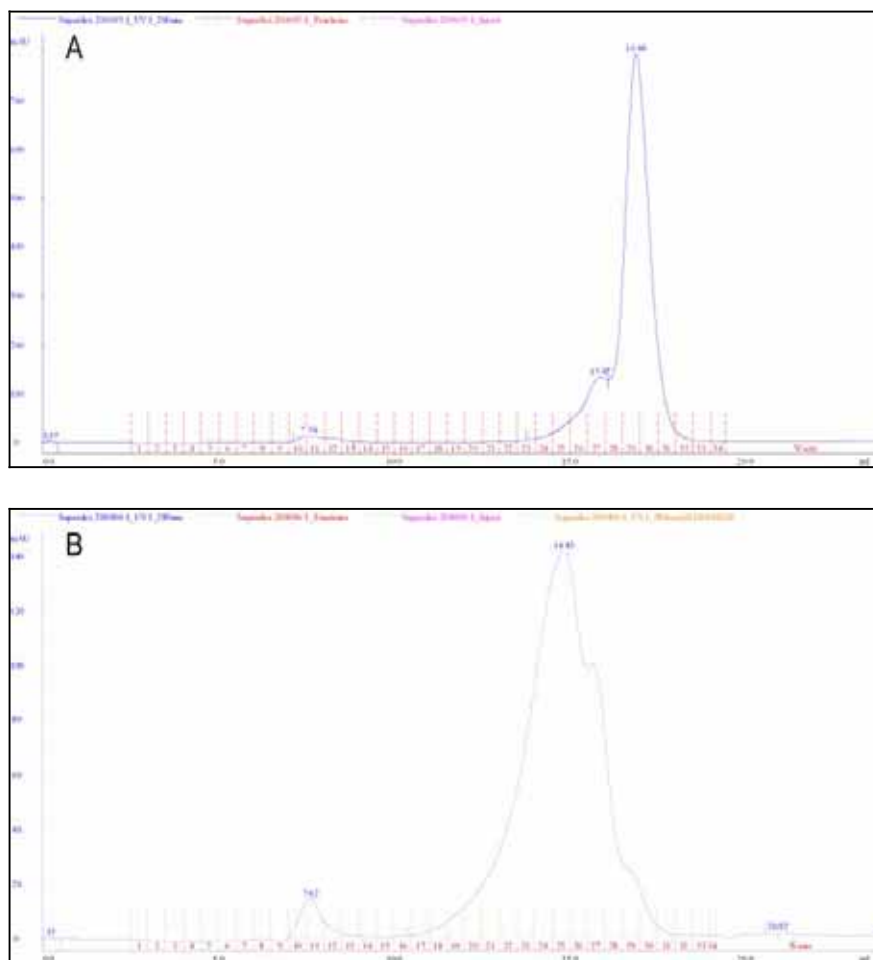
proteïnes contaminants, mentre que el segon i el tercer pic corresponien a una proteïna amb la massa molecular esperada, d'uns 25 kDa. Vam suposar que en els dos casos es tractava del domini C-terminal de la PaGS.



**Fig.2.4. Perfil d'elució del domini C-terminal de la PaGS d'una columna d'afinitat.** Les proteïnes unides a la columna d'afinitat elueixen en tres pics (A,B i C) de menys a més afinitat en resposta a un gradient d'imidazole, tal i com es descriu a l'apartat de Materials i Mètodes.

Davant la diferent naturalesa de les dues poblacions vam decidir continuar purificant-les per separat. Cadascuna de les fraccions va ser incubada amb trombina per eliminar la cua d'histidines i es van sotmetre novament a cromatografia d'afinitat per níquel. Aquesta vegada però ens interessà recollir la fracció no retinguda ja que ens interessava la proteïna tallada, que havia perdut l'afinitat per níquel. Cadascuna de les fraccions així purificades es carregà a continuació en una columna de gel filtració. Mentre que el producte del segon pic inicial donava lloc a un pic majoritari de massa molecular aparent d'uns 25 kDa, corresponent a un monòmer (Fig. 2.5 A), el tercer pic donava una barreja de dos pics de menor volum d'elució i per tant compatibles amb estats d'oligomerització superiors (Fig. 2.5 B). Atesa la seva homogeneïtat, s'escollí la fracció purificada que provenia del segon pic per continuar els estudis de cristal·lització.

En aquest procés obteníem un rendiment aproximat de 2-3 mg de proteïna purificada per cada litre de cultiu bacterià. Es concentraren les fraccions que contenien la proteïna fins a 10 mg/mL i es procedí a la seva cristal·lització.



**Fig.2.5. Cromatogrames de les filtracions per gel.** Cromatogrames de les columnes d'exclusió molecular de les fraccions eluïdes en el segon (A) i el tercer (B) pic de la columna d'afinitat segons el que es descriu a Materials i Mètodes. En el cromatograma A s'observa l'elució d'una proteïna majoritària a un volum de 16.88 mL, mentre que la gel filtració de la segona mostra dona un patró complex de pics en gel filtració.

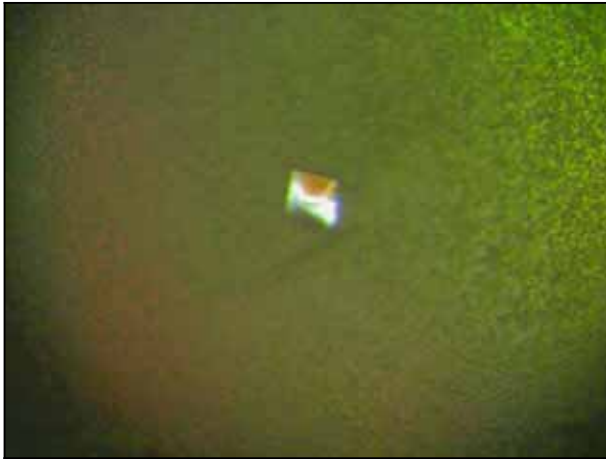
## 2.4. Cristal·lització del domini C-terminal:

### 2.4.1. Condicions inicials i optimització:

De la mateixa forma que amb la proteïna sencera, es van analitzar sistemàticament les possibles condicions de cristal·lització usant el mètode de difusió de vapor en gota penjant. L'escombrat de condicions es va fer utilitzant els equips d'Emerald Biostructures i els primers cristalls van aparèixer després de cinc dies a 4 °C. El licor de cristal·lització contenia 30% de PEG 8000, 0.2 M de sulfat de liti i acetat de sodi 0.1M a pH 4.5 com a tampó. Aquests primers cristalls tenien forma cúbica i una mida aproximada de 0,04x0,04x0,04 mm. Tot i que eren molt petits es van poder utilitzar per comprovar que eren cristalls de proteïna obtenint un parell d'imatges de difracció.

De tota manera, el volum dels cristalls era insuficient per obtenir dades completes i, a més, les primeres imatges mostraven que no eren monocristalls. Es va iniciar doncs

l'optimització del procés de cristallització mitjançant un rastreig sistemàtic de les condicions properes a les obtingudes inicialment. Vam aconseguir una millora notable en la mida i la qualitat dels cristalls establint les condicions òptimes a 19% de PEG8000 i 0.05 M de sulfat de liti en el tampó d'acetat de sodi 0.1M a pH 4.5. La mida dels cristalls així obtinguts fou 0.1x0.1x0.1 mm (Fig. 2.6).



**Fig.2.6. Cristall del domini C-terminal de la PaGS.** Cristalls obtinguts usant el mètode de difusió de vapor en gota penjant. El licor de cristallització contenia 19% PEG 8K, 0.05 M sulfat de liti i acetat de sodi pH 4.5, segons el que s'ha descrit a Materials i Mètodes.

#### **2.4.2. Cocrystal·litzacions i incubació amb substrats:**

Per homologia de seqüència del domini C-terminal de la PaGS amb els dominis equivalents d'altres proteïnes relacionades estructuralment (AtGS, OtsA i GP) es podia concloure que la majoria dels residus implicats en la unió al nucleòtid del substrat donador de glucosil estaven en aquest domini. Per aquest motiu, es va plantejar el creixement de nous cristalls en presència d'aquests substrats i productes (ADP, UDP, UDP-glucosa i glucosa). Es van seguir dues aproximacions: en primer lloc es va incubar el substrat amb la proteïna abans de cristal·litzar-la, en previsió que es pogués formar el complex i aquest cristal·litzés. En segon lloc, es van incubar cristalls de proteïna nadiua amb els diferents substrats i productes en previsió que aquests difonguessin dins dels cristalls i s'unissin específicament a la proteïna.

#### **2.4.3. Criocristal·lització:**

De la mateixa manera que per la proteïna sencera, es van determinar les condicions òptimes de criopreservació en nitrogen líquid. Com a crioprotector es va afegir glicerol, i va augmentar la seva concentració en dues etapes fins al 10 i 20%, respectivament.

## DISCUSSIÓ

La clonació, l'expressió i la purificació de l'hipotètic domini C-terminal de la PaGS han mostrat que, en efecte, aquest és capaç de plegar-se correctament tot i estar aïllat de l'altre meitat de la molècula.

La seva mida més reduïda i la seva naturalesa compacta han permès expressar-lo parcialment de forma soluble, a diferència de la proteïna sencera que només es podia obtenir en forma de cossos d'inclusió.

Tot i la presència de diferents poblacions del domini C-terminal amb propietats i estats d'oligomerització diferents, s'ha pogut purificar una fracció que conté majoritàriament el domini C-terminal en forma monomèrica i s'han obtingut cristalls de qualitat suficient per poder obtenir un espectre de difracció a resolució atòmica, tal i com s'amplia al capítol 3.

## **CAPÍTOL 3: RESOLUCIÓ DE LES ESTRUCTURES DE LA GLICOGEN SINTASA DE *PYROCOCCUS ABYSSI* I DEL SEU DOMINI C-TERMINAL.**

### **INTRODUCCIÓ**

Per a resoldre l'estructura de la PaGS i del seu domini C-terminal aïllat, havíem d'obtenir dades de difracció dels cristalls (capítol 1 i 2) i resoldre el problema de les fases.

Com a primera aproximació per a resoldre el problema de les fases, vam intentar utilitzar tècniques tradicionals de reemplaçament isomorf (MIR) i de difracció anòmala (MAD) dels cristalls derivats amb àtoms pesats o selenometionines, respectivament.

Un mètode alternatiu i aconsellable quan es disposa d'un bon model és el reemplaçament molecular (MR). En aquest cas però, no disposàvem d'un bon model ja que tot i que en el decurs d'aquesta tesi s'havia resolt l'estructura d'una GS de la família GT-5, la GS d'*Agrobacterium tumefaciens* (AtGS) (Buschiazzi *et al.* 2004), la identitat de seqüència global amb la PaGS és tan sols del 21%. A més, l'estructura l'AtGS està formada per dos dominis units per un solc molt profund que confereix molta flexibilitat al sistema i dificulta la resolució per MR.

Ateses aquestes dificultats, vam plantejar una aproximació inusual al problema que consistia en resoldre en primer lloc per MR un dels dos dominis de la PaGS, el domini C-terminal aïllat. En aquest cas la identitat de seqüència amb AtGS era una mica més gran (22%) i a més eliminàvem els problemes de mobilitat entre dominis. A partir de l'estructura resolta del domini C-terminal de PaGS, vam procedir a resoldre l'estructura de la PaGS sencera, tal i com s'exposa en aquest capítol.

### **RESULTATS**

#### **3.1. Resolució del domini C-terminal de la PaGS:**

##### **3.1.1. Recollida dels espectres de difracció i processament de les dades del cristall del domini C-terminal:**

Vam obtenir un primer espectre de difracció a partir dels cristalls cúbics descrits al capítol 2, utilitzant el generador d'ànode rotatori. L'òptica microfocus del sistema, caracteritzada per un feix de raigs X d'elevat flux i de petit diàmetre, va permetre obtenir un espectre a resolució atòmica tot i la reduïda mida del cristall, 0.05x0.05x0.05 mm (Fig. 3.1). L'espectre es va poder indexar i integrar amb el programa Mosflm i es va escalar amb Scala.



**Fig. 3.1. Imatge d'un cristall del domini C-terminal.** Cristall muntat en un *loop* de 0.05 mm, llest per a ser difractat.

Tot i que la resolució era gairebé atòmica i l'estadística correcta, vam decidir recollir un espectre al sincrotró per a millorar la resolució de l'espectre i la relació senyal-soroll, atesa la major brillantor, el menor diàmetre, el major paral·lelisme i la naturalesa altament monocromàtica del seu feix. Vam poder recollir un espectre complet d'un únic cristall a la línia BM16 del sincrotró i vam poder indexar i integrar les taques amb el programa Mosflm i posteriorment escalar amb el programa Scala.

	<i>Domini C-terminal</i>	<i>Domini C-terminal</i>
<b>Raigs X</b>	Ànode rotatori	Sincrotró
<b>Grup espacial</b>	P2 <sub>1</sub>	P2 <sub>1</sub>
<b>Paràmetres de cel·la (Å, °)</b>	a=39.8, b=52.4, c=42.6, β=105.7	a=39.6, b=50.8, c=44.2, β=106.8
<b>Resolució (Å)</b>	25-2.5 (2.6-2.5)	50-1.8 (1.9-1.8)
<b>Reflexions úniques</b>	4772	15701
<b>Completesa (%)</b>	99.8 (98.0)	97.6 (85.9)
<b>Mosaicitat (°)</b>	0.5	0.7
<b>R<sub>mescla</sub> (%)</b>	9.6 (24.2)	4.3 (24.8)
<b>Multiplicitat</b>	2.0 (2.0)	2.4 (2.0)
<b>&lt;I/σ(I)&gt;</b>	39.2 (64.0)	15.1 (3.5)
<b>B factor (Pendent de Wilson, Å<sup>2</sup>)</b>	29.6	25.4

**Taula 3.1. Estadística dels espectres de difracció dels cristalls del domini C-terminal.** Estadística de les dades de difracció recollides a la font d'ànode rotatori (esquerra) i al sincrotró (dreta).

$$R_{\text{mescla}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \{I(hkl)\}|}{\sum_{hkl} \sum_i I_i(hkl)}$$

Com es mostra a la taula 3.1, la proteïna empaquetava en el grup espacial monoclínic P2<sub>1</sub>. Aquest és un dels empaquetaments més comuns de les proteïnes, amb una freqüència del 13.6% (Teplyakov 1990), i es caracteritza per tenir un únic eix binari helicoidal a través de l'eix b de la cel·la unitària. Atesa la simetria del grup espacial, calgué recollir 180° per a poder disposar de dades de difracció completes. El residual de

mescla ( $R_{\text{mescla}}$ ) de l'espectre complet era d'un 4.3% i el factor de temperatura (B) calculat a partir del pendent de Wilson era  $25.4 \text{ \AA}^2$ .

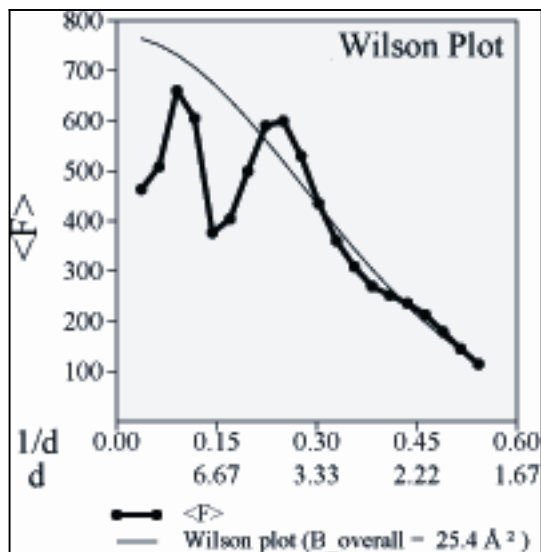


Fig. 3.2. Gràfic de Wilson. Factor de temperatura (B) calculat a partir del pendent de la corba .

A la imatge de difracció, les taques estaven molt separades, atès que la distància a l'espai recíproc és inversament proporcional a la mida de la cel·la unitària. La integració de les imatges va permetre determinar els paràmetres de cel·la (taula 3.1), que proporcionaren un volum de  $86150 \text{ \AA}^3$ . Assumint que la massa molecular del monòmer és d'uns 22 KDa, el càlcul del coeficient de Matthews (Matthews 1968) va donar una solució única a 1.9, corresponent a una molècula per unitat asimètrica i un contingut en solvent del 36%.

Es tractava, doncs, d'uns cristalls molt compactes amb un coeficient d'empaquetament molt proper al màxim permès per a esferes perfectes (Andersson i Hovmoller 2000). Quedava palesa la clara tendència del grup espacial  $P2_1$  a assolir uns elevats graus d'empaquetament. Davant dels molts graus de llibertat del grup, estarien afavorits energèticament aquells empaquetaments que maximitzessin el nombre de contactes entre les molècules i la densitat del cristall i minimitzessin, per tant, el volum mort.

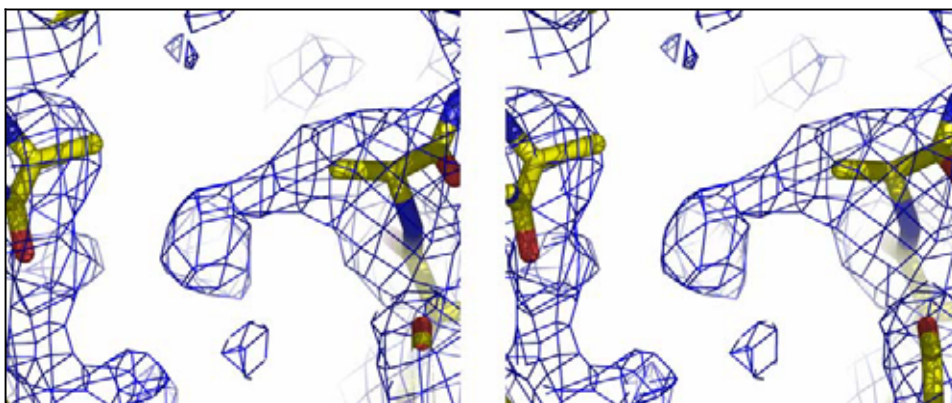
### 3.1.2. Resolució de l'estructura del domini C-terminal:

L'estructura del domini C-terminal de la glicogen sintasa a de *P.abyssi* es va resoldre per reemplaçament molecular (MR), a partir de les dades experimentals obtingudes utilitzant l'ànode rotatori i usant com a model el domini C-terminal de la proteïna homòloga d' *A. tumefaciens*. Amdòs dominis presenten un 22% d'identitat de seqüència. El programa de MR Molrep (Vaguin i Teplyakov 1997) va donar una primera solució inequívoca amb un coeficient de correlació del 43% i un  $R_{\text{factor}}$  del 50% a  $4.5 \text{ \AA}$  de resolució, usant el model del domini C-terminal d' *A. tumefaciens* convertit a polialanines en que s'havia eliminat l'hèlix  $\alpha 15$  (Fig. 4.4. Capítol 4). L'empaquetament d'aquesta primera solució era correcta i gairebé no s'observava cap contacte cristal·logràfic.



Va ser necessari utilitzar l'afinament del cos rígid amb el programa Refmac5 (CCP4) al voltant de les millors posicions i orientacions obtingudes anteriorment. Es van afinar cadascuna de les cinc dimensions (tres rotacions i dos translacions) per a cada cos rígid. En l'afinament es van utilitzar les dades de l'espectre de difracció obtingut a la línia BM16 del sincrotró.

Davant del fet que en aquest punt disposàvem d'un model a baixa resolució, vam estendre les fases de forma iterativa mitjançant el programa DM (CCP4). Vam construir una màscara al voltant del model i vam aplicar els criteris de distribució de la densitat electrònica (*histogram matching*), l'aplanament de solvent (*solvent flattening*) i la connectivitat entre zones contínues de més alta densitat (*skeletonization*). D'aquesta forma vam obtenir els primers mapes interpretables. Tot i que el model que havíem utilitzat era el de polialanines, es podien distingir les cadenes laterals dels aminoàcids, indicant que havíem arribat a una solució real i no a una sobreestimació del model (Fig. 3.3).



**Fig. 3.3. Diferències entre el mapa de densitat electrònica i el model.** Dibuix estereogràfic que mostra el traçat de densitat electrònica calculada per DM corresponent a cadenes laterals tot i que el model només està compostat per una cadena de polialanines.

El traçat de l'estructura es va fer manualment, usant el programa gràfic O (Jones *et al.* 1991). Novament de forma iterativa, es van anar incloent etapes d'afinat de les posicions dels àtoms així com de la seva estereoquímica, usant Refmac5 (CCP4). Els graus de restricció es van anar alleugerint a mida que avançava l'afinament. Finalment es van afegir les molècules de solvent de forma automàtica amb el programa ARP/wARP (CCP4) i van ser comprovades manualment.

Com s'observa a la taula 3.2, el model final contenia 191 residus, que corresponien a tots els residus del domini C-terminal, des de la Gly 218 fins a la Ser 413, més una extensió al seu N-terminal de tres residus (Gly-Ser-His) procedents del vector. L'elevada qualitat de la densitat electrònica i l'empaquetament tan compacte del cristall van permetre determinar la doble configuració de disset residus (Cys 221, Glu 226, Ser 242, Asp 247, Met 253, Ile 255, Arg 257, Arg 260, Lys 263, Ile 288, Ser 302, Met 316, Val 322, Val 332, Leu 351, Glu 392, Met 410), així com la localització de 198 molècules d'aigua (Fig. 3.4). El mapa de densitat electrònica també mostrava la presència de dos ions acetat i quatre ions sulfat procedents del tampó de cristallització.

Afinament	Domini C-terminal
Marge de resolució (Å)	40 – 1.8
Reflexions (%)	14581 (833)
R factor (%)	18.3 (23.5)
B mitjà (Å <sup>2</sup> )	15.64
<b>Nombre d'àtoms</b>	1866
Àtoms proteics	1669
Molècules d'acetat	2
Molècules de sulfat	4
Molècules d'aigua	198
<b>RMSD dels valors ideals</b>	
Rmsd enllaç (Å)	0.008
Rmsd angle (°)	0.98
<b>Diagrama de Ramachandran</b>	
Més afavorits (%)	91.8
Permesos(%)	8.2
Permesos generosament(%)	0.0
Prohibits (%)	0.0

Taula 3.2. Afinament de l'estructura del domini C-terminal. Afinament utilitzant el programa Refmac5.

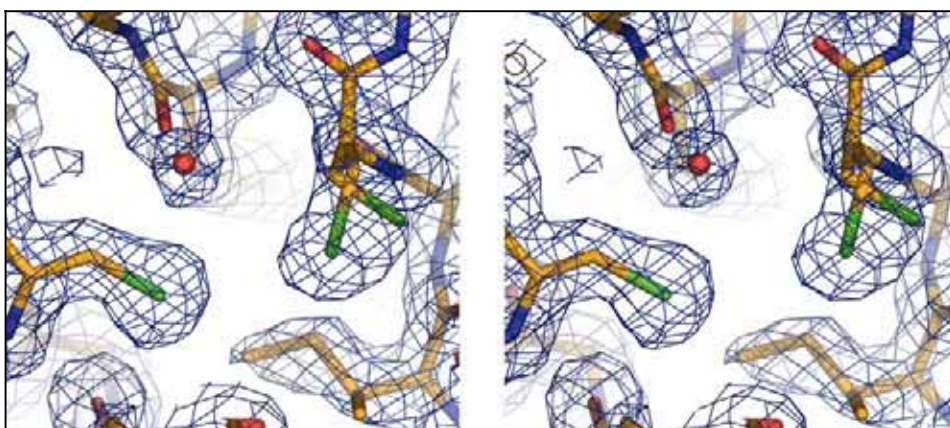
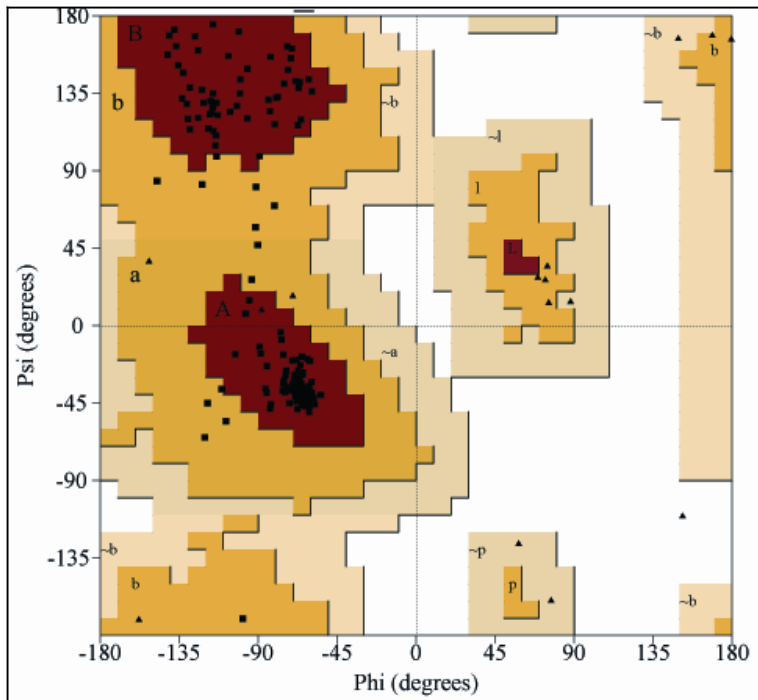


Fig. 3.4. Doble configuració de la Cys221. Dibuix estereogràfic de la doble configuració de la Cys221 en el cristall del domini C-terminal.

La qualitat estereoquímica del model, comprovada mitjançant el programa Procheck (Laskowski *et al.* 1993), quedava palesa pel fet que tots els angles conformacionals de la cadena principal estaven en zones permeses del Diagrama de Ramachandran (Fig. 3.5).



**Fig. 3.5. Diagrama de Ramachandran.** Totes les conformacions dels angles de la cadena principal estan dins de les zones permeses.

## 3.2. Resolució de la PaGS:

### 3.2.1. Recollida dels espectres de difracció i processament de les dades del cristall de PaGS:

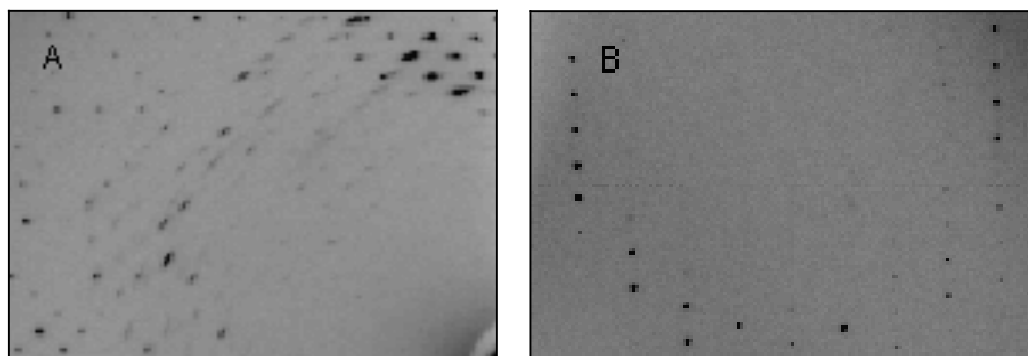
Atesa la naturalesa dels cristalls de la proteïna nadiua, era imprescindible usar una font de radiació molt potent, per la qual cosa van ser difractats a la línia ID14-EH2 del sincrotró. Es van recollir dos espectres de dos cristalls únics, amb i sense UDPG, respectivament.

Com mostra la taula 3.3, el cristall que havia crescut en presència d'UDPG difractava a més resolució. Vam concloure que el substrat afegit al licor de cristallització millorava la qualitat dels cristalls.

	PaGS	UDPG-PaGS
<b>Línia de sincrotró</b>	ID14-EH2	ID14-EH2
<b>Grup espacial</b>	C2	C2
<b>Cel·la (Å, °)</b>	a=202.8, b=73.5, c=148.3, β=131.2	a=202.9, b=73.9, c=148.2, β=131.2
<b>Resolució (Å)</b>	30-3.5 (3.6-3.5)	30-2.8 (2.9-2.8)
<b>Reflexions úniques</b>	20885	41544
<b>Completesa (%)</b>	99.7 (99.3)	99.9 (99.8)
<b>Mosaicitat (°)</b>	0.45	0.6
<b>R<sub>mescla</sub> (%)</b>	14.7 (32.3)	6.6 (45)
<b>Multiplicitat</b>	2.5 (2.5)	3.7 (3.7)
<b>&lt;I/σ(I)&gt;</b>	6.7 (2.4)	14.1 (1.8)
<b>B factor (Wilson, Å<sup>2</sup>)</b>	58.0	61.0

**Taula 3.3. Estadística de l'espectre de la PaGS.** Estadística de les dades de difracció del cristall nadiu de PaGS i el crescut en presència de UDPG.  $R_{mescla} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \{I(hkl)\}|}{\sum_{hkl} \sum_i I_i(hkl)}$ .

La forma d'agulla dels cristalls feia que aquests fossin altament anisotròpics, i que presentessin unes intensitats de difracció molt diferents en funció del volum de cristall que travessava el raig en cada difracció. Així doncs, si el feix travessava l'agulla longitudinalment la difracció era molt més intensa que si només ho feia transversalment. Imatges de difracció d'altres cristalls obtingudes anteriorment eren gairebé improcessables ja que les taques no eren escalables. Un altre problema greu va ser l'alta mosaicitat d'aquests cristalls en la direcció longitudinal referida, causat per un desordre en la xarxa cristal·lina, que feia que les taques s'estenguessin fins a 4 o 5° (Fig. 3.6). Aquest desordre era tan intens en alguns cristalls que es podia observar macroscòpicament un cert recargolament de les agulles.



**Fig. 3.6. Anisotropia dels cristalls de PaGS.** Dues imatges de difracció del mateix cristall: direcció on la mosaicitat és molt elevada (A) i on aquesta mosaicitat és aparentment no observable (B).

Per tal d'evitar aquest fenomen, vam muntar les agulles de forma que el seu eix més llarg quedés a uns 20° de l'eix de rotació. D'aquesta manera, el feix de raigs X

perpendicular a l'eix de rotació travessava en tota la recollida un volum similar de cristall (evitant l'anisotropia) i a més ho feia en la zona de menor volum (evitant la major mosaicitat). Atès que era molt probable que algun dels eixos cristal·logràfics coincidís amb l'eix de l'agulla, es va evitar que aquest últim coincidís al seu torn amb l'eix de rotació i això fes disminuir la completesa de l'espectre.

Ambdós espectres es van processar amb el paquet Denzo/Scalepack i es van obtenir dades a 3.5 i 2.8 Å amb un 97.9 i un 99.9% de completesa per al cristall nadiu i per al que s'havia obtingut en presència d'UDPG, respectivament. Els dos cristalls eren isomorfs i pertanyien al grup espacial C2, amb els paràmetres de cel·la que s'indiquen a la taula 3.3. El volum de la cel·la era de 1674800 Å<sup>3</sup> i, assumint que la massa molecular de cada monòmer és d'uns 51 kDa, el càlcul dels coeficients de Matthews resultà en el que es mostra en la taula 3.4.

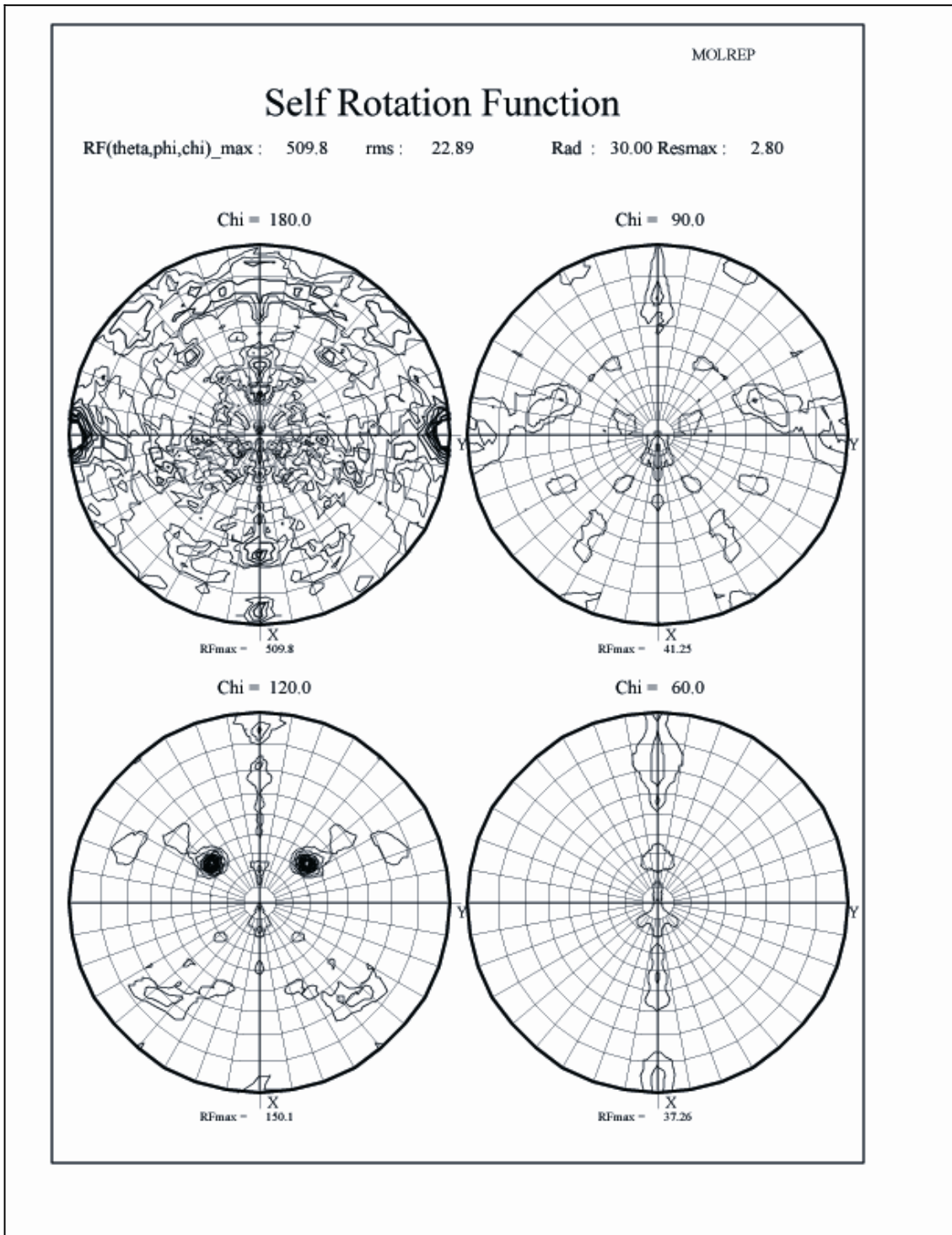
Nmol/asim	Coef. Matthews	%solvent
1	8.2	84.9
2	4.1	69.8
<b>3</b>	<b>2.7</b>	<b>54.7</b>
4	2.1	39.6
5	1.6	24.5
6	1.4	9.4

**Taula 3.4. Càlcul del coeficient de Matthews.** Atesos els paràmetres de cel·la, el seu volum i la massa molecular del monòmer, el càlcul del coeficient de Matthews permet estimar el nombre de monòmers en la unitat asimètrica i el % de solvent.

Tot i que la presència de tres molècules a la unitat asimètrica era el més probable, amb un 55% de solvent, no es podia descartar que es tractés d'un dímer o un tetràmer cristal·logràfic, ja que ambdós tenien contingut de solvent acceptable per a formar un cristall de proteïna (70 i 40%, respectivament).

La prova definitiva que confirmà que la PaGS cristal·litzava en forma de trímer va ser el càlcul de la funció d'autorotació (*self-rotation*) amb el programa Molrep (Vaguin i Teplyakov 1997). Amb la funció d'autorotació podíem determinar la correlació de les reflexions de Patterson amb elles mateixes a una rotació diferent. S'observen a la figura 3.7, a més de l'eix binari cristal·logràfic a  $\chi=180^\circ$ , dos pics molt intensos relacionats per la simetria del grup espacial a  $\chi=120^\circ$  i amb un eix de rotació a  $63^\circ$  respecte l'eix cristal·logràfic C2, paral·lel a b, i a  $35^\circ$  de l'eix z. Vam concloure que la unitat asimètrica estava composta per un homotrímer de PaGS.

La presència d'aquest eix no-cristal·logràfic podria facilitar la resolució de l'estructura per reemplaçament molecular ja que d'una banda permetria aplicar un filtre a les solucions, essent només possibles aquelles que respectessin la simetria trimèrica. D'una altra banda, també afavoriria l'extensió de les fases inicials (qualsevol que fos el seu origen) per terme mitjà no cristal·logràfic i en l'afinat. Tots aquests aspectes seran tractats a l'apartat 3.2.2.



**Fig.3.7. Diagrama de l'autorotació del cristall de PaGS.** En el diagrama es mostren els pics de la funció d'autorotació calculada per Molrep en els angles de rotació de 180°, 90°, 120°, 60°. Els dos pics que apareixen a 180° són producte de la simetria cristal·logràfica del grup C2 mentre que els de 120° corresponen a una simetria no cristal·logràfica.

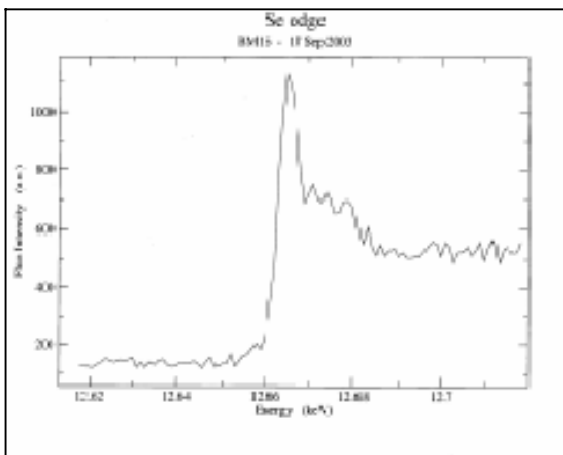
### 3.2.2. Resolució de l'estructura de la PaGS:

Per a resoldre l'estructura de la PaGS es va abordar el problema de la fase mitjançant les tres aproximacions més comuns: senyal anòmal per selenometionines (simple SAD i múltiple MAD), reemplaçament isomorf per derivatització amb àtoms pesats i, per últim, reemplaçament molecular.

#### 3.2.2.1 Dispersió anòmala per selenometionines:

Els cristalls de proteïna recombinant, en què els àtoms de sofre de les metionines havien estat substituïts per selenis, obtinguts tal com s'indica en l'apartat 1.6.1, van ser difractats per llum de sincrotró. Aquest mètode presentava dos avantatges clars: en primer lloc, els selenis s'havien incorporat homogèniament a les proteïnes, ja que en formaven part constitutivament. En segon lloc, el seleni com a àtom pesat presenta el seu màxim d'absorció a una longitud d'ona molt adequada per als cristalls de proteïna i molt comú en els sincrotrons, donant lloc al senyal anòmal.

Es confirmà la incorporació de l'àtom pesat als cristalls duent a terme un experiment de fluorescència per raigs X (Fig. 3.8), en què s'observa el màxim d'absorció a 12,658 KeV.



**Fig. 3.8. Fluorescència del Seleni en els cristalls de SeMet-PaGS.** Gràfic on es representa la intensitat de flux en funció de l'energia del feix incident.

En primer lloc, es realitzà un experiment de dispersió anòmala simple, recollint un espectre de difracció a la línia ID13 del sincrotró a 3.5 Å de resolució. Aquesta es caracteritza per la seva òptica *microfocus*, donant lloc a un feix molt col·limat d'alta energia. Aquest feix és molt adequat per a analitzar cristalls petits ja que d'una banda maximitza la difracció, però d'altra banda, augmenta molt la divergència del raig i també la mosaicitat aparent del cristall. La intensitat de la radiació incident va fer que augmentés exponencialment el dany en el cristall, fet que va obligar a anar desplaçant el cristall a mida que avançava la recollida. Tot i aquestes precaucions, la difracció del cristall anava decaient.

Vam indexar independentment cada bloc de recollida amb el programa Denzo, ja que la translació al llarg de l'agulla no assegurava que les condicions en cada punt fossin idèntiques. A continuació vam escalar conjuntament tot l'espectre amb el programa Scalepack, permetent que refinés els diferents paràmetres d'escalat per separat (Taula 3.5).

No vam poder detectar senyal anòmal en aquests cristalls.

	SAD	MAD		
		$\lambda$ pic	$\lambda$ inflexió	$\lambda$ remota
<b>Línia de sincrotró</b>	ID13	ID29	ID29	ID29
<b>Longitud d'ona (<math>\lambda</math>)</b>	0.9789	0.97976	0.979916	0.96862
<b>Grup espacial</b>	C2	C2	C2	C2
<b>Paràmetres de cel·la (<math>\text{Å}</math>, °)</b>	a=202.2,b=73.3, c=149.3, $\beta$ =131.2	a=203.4,b=73.3, c=148.4, $\beta$ =131.3	a=204.2,b=73.4, c=148.7, $\beta$ =131.4	a=203.9,b=73.2, c=148.5, $\beta$ =131.3
<b>Resolució (<math>\text{Å}</math>)</b>	30-3.4 (3.5-3.4)	25-3.0 (3.1-3.2)	25-3.3 (3.2-3.3)	25-3.6 (3.5-3.6)
<b>Reflexions úniques</b>	22978	32671	24860	17130
<b>Completesa (%)</b>	88.0 (90.5)	99.0 (98.9)	98.8 (98.4)	87.6 (85.5)
<b>Mosaicitat (°)</b>	0.8	0.6	0.75	1.1
<b><math>R_{\text{mescla}}</math> (%)</b>	13.5 (49.2)	11.1 (48.2)	13.6 (52.0)	18.1 (61.7)
<b>Multiplicitat</b>	1.9 (2.0)	3.5 (3.5)	3.4 (3.4)	2 (2)
<b><math>\langle I/\sigma(I) \rangle</math></b>	7.9 (1.8)	12.0 (3.1)	10.1 (2.9)	5.5 (1.5)
<b>Bfactor (Wilson,<math>\text{Å}^2</math>)</b>	60.0	45.5	52.3	60.6

**Taula 3.5. Estadística de la recollida de SAD i MAD de la SeMet-PaGS.** Dades estadístiques d'un experiment de SAD a la línia ID13 i un MAD a la línia ID29, recollits a l'ESRF.  $R_{\text{mescla}} = \sum_{hkl} \sum_i |I_i(hkl) - \{I(hkl)\}| / \sum_{hkl} \sum_i I_i(hkl)$

A la línia ID29 del sincrotró, que permetia modular la longitud d'ona del feix, vam recollir un espectre complet de dispersió anòmala múltiple (MAD). Vam utilitzar l'espectre de fluorescència per a determinar quines eren les longituds d'ona en les que l'absorció era màxima (pic), en el punt d'inflexió i en un punt allunyat després d'aquest màxim (remota). I vam recollir a cadascuna d'aquestes longituds  $180^\circ$ , en el mateix ordre (pic, inflexió i remota). Com s'observa a la taula 3.V, la qualitat de les dades anava disminuint a mesura que avançava la recollida a causa del dany per radiació que patia el cristall en una de les línies més intenses del sincrotró.

Després de processar les dades amb el paquet Denzo/Scalepack es procedí a analitzar el senyal anòmal amb el programa Xprep. Només teníem senyal anòmal a baixa resolució, inferior a  $5 \text{ Å}$ .



Vam intentar trobar les posicions del seleni mitjançant el programa Solve i SnB, però no ho vam aconseguir. Tot i que la proteïna havia incorporat el seleni, en cap dels dos experiments (SAD i MAD) no vam poder aconseguir senyal anòmal a elevada resolució a causa dels problemes intrínsecs que presentaven els cristalls. La mosaicitat, l'anisotropia i el dany per radiació van provocar un soroll experimental que en tots els casos fou major que les diferències anòmales.

### 3.2.2.2. Reemplaçament isomorf per àtoms pesats:

Com a segona aproximació a la resolució del problema de la fase, es va utilitzar el mètode basat en la capacitat de difracció més intensa dels àtoms pesats. Com s'ha indicat al capítol 1, els cristalls nadius es van submergir en una solució que contenia l'àtom pesat, mètode que minimitzava el risc de que el cristall deixés de ser isomorf. Tot i gairebé assegurar l'isomorfisme no es podia garantir una incorporació de l'àtom metàl·lic a la proteïna de forma ordenada i amb una elevada ocupació, amdos factors limitants per a poder resoldre l'estructura. Com s'observa a la taula 3.6, es van poder processar les dades completes de tres possibles cristalls derivatitzats (Hg, Ta i W), ja que no vam poder escalar els espectres de Pt a causa de la poca qualitat de les dades.

Àtom pesat	Hg (II)	Ta (V)	W
<b>Compost</b>	HgCl <sub>2</sub>	Ta <sub>6</sub> Br <sub>14</sub>	K <sub>6</sub> HSiW <sub>9</sub> V <sub>3</sub> O <sub>40</sub>
<b>Línia de sincrotró</b>	ID14-EH2	ID14-EH2	ID14-EH2
<b>Grup espacial</b>	C2	C2	C2
<b>Paràmetres de cel·la (Å, °)</b>	203.9, 73.7, 149.3, 131.4	203.1, 73.4, 148.3, 131.2	203.5, 74.2, 150.0, 131.4
<b>Resolució (Å)</b>	30-3.7	30-3.5 (3.62-3.5)	30-3.5
<b>Reflexions úniques</b>	23346	21927	22287
<b>Completesa (%)</b>	98.1 (85.0)	98.6 (87.1)	98.4 (89.0)
<b>Mosaicitat (°)</b>	0.55	0.55	0.3
<b>R<sub>mescla</sub> (%)</b>	7.4 (20.0)	7.3 (19.8)	7.1 (19.4)
<b>Multiplicitat</b>	3.1 (3.0)	3.2 (3.8)	3.3 (3.8)
<b>&lt;I/σ(I)&gt;</b>	13.3 (3.0)	17.9 (5.1)	19.0 (4.9)
<b>B factor (Wilson, Å<sup>2</sup>)</b>	60.0	59.2	58.0

**Taula 3.6. Àtoms pesats.** Estadística de les dades de difracció recollits pels cristalls incubats amb clorur de mercuri i amb els clústers de tàntal i tungstè.  $R_{mescla} = \sum_{hkl} \sum_i |I_i(hkl) - \{I(hkl)\}| / \sum_{hkl} \sum_i I_i(hkl)$ .

Amb cadascun dels cristalls vam intentar resoldre l'estructura utilitzant tècniques de reemplaçament isomorf únic (SIR) i vam intentar combinar-ho amb el possible senyal anòmal dels cristalls (SIRAS) a diferents resolucions. Vam utilitzar el cristall incubat amb

UDPG com a cristall nadiu, i vam usar els programes SOLVE i BnP. Vam treballar a unes resolucions de 5-6 Å en el cas dels clústers de Ta i W, i a 3.7 Å pel Hg. No vam obtenir senyal anòmal per sobre de 8-10 Å de resolució. Els experiments de reemplaçament isomorf múltiple (MIR o MIRAS) també van donar resultats negatius.

En cap dels casos analitzats vam aconseguir resoldre el problema de la fase mitjançant aquesta aproximació.

### 3.2.2.3. Reemplaçament molecular. Funció de rotació i de translació bloquejades (*locked rotation function*):

Al mateix temps que intentàvem resoldre la nostra estructura, el grup del Dr. Alzari va resoldre l'estructura d'una proteïna homòloga, la glicogen sintasa d' *Agrobacterium tumefaciens*, la seqüència aminoacídica de la qual és només 21% idèntica a la de PaGS. D'altra banda, la GS consta de dos dominis separats per un solc molt profund i això ens feia pensar que la posició relativa entre els dos dominis podia ser sensiblement diferent en ambdues proteïnes. Com a dades experimentals vam usar l'espectre del cristall de PaGS amb UDPG, ja que tenia més resolució.

Vam fer una primera aproximació clàssica al problema utilitzant les funcions de rotació i translació ordinàries dels programes de MR Molrep i Amore per a posicionar els tres monòmers a la unitat asimètrica. Tot i que vam fer nombrosos intents considerant diferents resolucions (6.0-3.0 Å), diferents models (amb i sense cadenes laterals) i dividint la cerca per dominis N- i C-terminal del model, no vam arribar a una solució final amb una correlació elevada i en cap de les solucions proposades es mantenia la simetria no-cristal·logràfica entre els monòmers.

En aquest punt, ens vam plantejar incloure la informació de la simetria no-cristal·logràfica del cristall per a millorar la relació senyal-soroll, permetent l'aplicació del valor mitjà entre les subunitats. A partir de la funció d'autorotació anteriorment exposada, podíem determinar la posició relativa del conjunt no-cristal·logràfic respecte al cristall. Vam utilitzar el programa GLRF (*general locked rotation function*) per calcular la funció de rotació bloquejada (*locked cross-rotation function*) per un dels monòmers (els altres dos quedaven definits per simetria no-cristal·logràfica) i una posterior funció de translació bloquejada (*locked translation function*), que determinava la posició d'aquest monòmer respecte al centre del conjunt no-cristal·logràfic. A partir d'aquest punt teníem una primera solució del conjunt centrat a l'origen de coordenades. Vam utilitzar la translació ordinària del programa Molrep per a poder posicionar el conjunt en el cristall i vam analitzar les possibles solucions i l'empaquetament a què donaven lloc.

Tot i el potencial de la tècnica, no vam ser capaços de discriminar cap solució del conjunt de possibles, ja que l'espectre de correlacions era gairebé continu i totes donaven lloc a greus contactes en els suposats empaquetaments. D'altra banda, no era possible analitzar moltes condicions alhora, ja que aquest programa necessitava molt temps de càlcul.

Necessitàvem un programa que fos capaç d'automatitzar la cerca de solucions en el mínim temps de càlcul possible i que alhora permetés una major discriminació. El programa AMoRe permetia l'automatització del mètode explorant ràpidament múltiples configuracions al cristall. Utilitzava uns algorismes de càlcul molt potents que permetien determinar les funcions de rotació i translació ràpidament, així com el refinament de les posicions per mínims quadrats. Mentre que GLRF només permetia calcular les translacions per a l'orientació de major correlació, AMoRe permetia analitzar totes les translacions per a totes les orientacions calculades. Aquesta cerca massiva era molt interessant ja que era molt probable que la correcta orientació no fos la que tingués la millor correlació i el millor  $R_{\text{factor}}$ . A l'inici d'aquesta etapa el programa AMoRe disposava de les funcions de rotació i translació tradicionals així com de la funció de rotació bloquejada (*locked rotation function*). El mateix Dr. Jorge Navaza es va oferir per a implementar la funció de translació bloquejada (*locked translation function*) i a col·laborar en la resolució del nostre problema.

Com a models inicials vam usar el monòmer de glicogen sintasa d'*A. tumefaciens* i onze modificacions d'aquest obtingudes a partir del càlcul dels Modes Normals de Vibració (NM) de menor energia (NM 7) considerats independentment (models 1-11) o tots superposats (model 12) tal i com s'exposa a materials i mètodes (apartat 10.2.2).

Aquesta aproximació ens permetia identificar possibles diferències conformacionals entre model i solució. Es van usar les orientacions corresponents als cent primers pics de la funció de rotació ordinària. Per a cada model i per a cada orientació, només s'escollí la solució de translació bloquejada de major coeficient de correlació (CC) aplicant la restricció de la simetria trimèrica no-cristal·logràfica. Després de refinar per cos rígid, els models trimèrics van ser sotmesos a una translació ordinària i les cent millors solucions sotmeses a refinament per cos rígid. Els resultats no van ser concloents i no vam poder discriminar la solució correcta.

Un cop resolta l'estructura de la proteïna tal i com s'exposa l'apartat següent, el Dr. Jorge Navaza i el seu equip van reavaluar aquest mètode usant l'estructura resolta com a referència. Van concloure que la solució correcta apareixia en primera posició per al model monomèric d'*A.tumefaciens* (model 0) i per als models modificats corresponents als NM (6, 7, 9, 10 i 12) quan es considerava P1 com a grup espacial en el primer refinament per cos rígid, en comptes del veritable C2. Com es mostra a la taula 3.7, els models 9 i 10 presentaven els valors més alts de correlació i un contrast raonable amb la segona millor solució, permetent la identificació de la solució correcta. Per últim, les solucions considerades correctes van ser sotmeses a un refinament per cos rígid per dominis (dominis N i C-terminals).

Model No.:	ref.	0	1	2	3	4	5	6	7	8	9	10	11	12
<b>Superposition of the monomeric models onto the reference model</b>														
rmsd (Å) after:														
- density superposition	0.0	2.53	3.61	3.34	3.08	2.84	2.60	2.40	2.21	2.07	1.96	1.90	1.90	—
- atom superposition	0.0	2.33	3.59	3.32	3.05	2.79	2.55	2.33	2.13	1.97	1.84	1.76	1.75	—
<b>Crystallographic rigid body fit around the reference model MR positions</b>														
At 7.0 Å res.:														
- CC (%)	69.0	28.8	21.9	21.0	20.5	21.9	24.3	26.7	30.9	32.2	33.2	32.3	32.1	28.0
- monomer A rmsd (Å)	0.00	2.91	3.94	4.32	3.28	2.91	3.23	3.03	2.65	2.40	2.24	2.28	2.26	—
- monomer B "	0.00	4.02	5.26	3.56	3.66	3.06	2.75	2.64	3.12	2.99	2.48	2.16	2.30	—
- monomer C "	0.00	2.59	4.60	3.80	3.88	3.91	2.65	2.60	2.59	2.67	2.90	2.36	2.65	—
At 5.5 Å res.:														
- CC (%)	70.2	32.4	29.0	29.4	28.6	30.2	32.0	33.3	34.8	35.1	36.6	36.9	36.6	34.7
- monomer A rmsd (Å)	0.00	2.69	4.33	4.02	3.24	3.69	2.95	2.76	2.37	2.24	2.18	2.46	2.43	—
- monomer B "	0.00	2.48	3.94	3.84	3.36	2.98	2.71	2.45	2.29	2.19	2.27	2.31	2.26	—
- monomer C "	0.00	2.78	4.32	3.93	3.46	3.07	2.91	2.71	2.61	2.53	2.83	2.45	2.29	—

**Taula 3.7. Comparacions dels models usats pel reemplaçament molecular.** Comparacions entre l'estructura de la PaGS resolta i el model de la AtGS i les 12 modificacions d'aquest calculades per NM. Els àtoms considerats per al càlcul del rmsd són els àtoms de la cadena principal del conjunt de residus equivalents (distància màxima entre C $\alpha$ -C $\alpha$  de 3.5 Å) que resulten de l'alineament estructural independent dels dominis N- i C-terminals.

### 3.2.2.4. Reemplaçament molecular a partir de l'estructura del domini C-terminal:

Donada la impossibilitat de resoldre l'estructura per cap dels mètodes tradicionals exposats, ens vam plantejar una aproximació completament diferent. Vam construir un model híbrid a partir del domini C-terminal de la glicogen sintasa de *P.abysyi* resolt a l'apartat 3.1 i el domini N-terminal de la glicogen sintasa d'*A.tumefaciens* (apartat J.2.1, Materials i Mètodes). Com a dades experimentals inicials vam usar l'espectre de difracció a 2.8 Å del cristall de proteïna amb UDPG, ja que era el de major resolució.

Disposàvem de l'estructura del domini C-terminal aïllat de la glicogen sintasa de *P.abysyi* a 1.7 Å de resolució. Tot i les diferències degudes als diferents entorns, aquest model havia de ser gairebé idèntic al domini C-terminal en l'estructura sencera. En primer lloc vam utilitzar les funcions ordinàries de rotació i translació a 3.0 Å del programa Molrep per a posicionar el domini C-terminal en cadascuna de les tres subunitats del trímer, considerant-ne les cadenes laterals. Tot i la similitud del model, el programa només trobava dos dels tres dominis relacionats per l'eix ternari no cristal·logràfic (Taula 3.8).

Sol	RF		theta	phi	chi	alpha	beta	gamma	Rf	Rf/sigma
Sol	RF	1	59.25	108.26	100.32	49.76	82.58	13.23	142.8	7.21
Sol	RF	2	153.75	-143.73	169.55	42.09	52.26	149.55	115.7	5.84
Sol	RF	3	161.11	-176.55	162.55	12.67	37.33	185.76	96.62	4.88
Sol	RF	4	152.24	-22.32	179.86	157.61	55.51	22.24	96.59	4.88
Sol	RF	5	157.93	-158.77	159.90	32.06	43.43	169.60	94.06	4.75
<b>Domini C-terminal 1</b>										
Sol	RF	TF	theta	phi	chi	tx	ty	tz	Rfac	Corr
Sol	2	1	153.75	-143.73	169.55	0.194	0.000	0.306	0.583	0.119
Sol	1	1	59.25	108.26	100.32	0.310	0.000	0.435	0.581	0.118
Sol	3	7	161.11	-176.55	162.55	0.563	0.000	0.445	0.587	0.091
<b>Domini C-terminal 2 (fixat domini C-terminal 1)</b>										
Sol	RF	TF	theta	phi	chi	tx	ty	tz	Rfac	Corr
Sol	1	1	59.25	108.26	100.32	0.810	0.288	0.935	0.566	0.164
Sol	4	1	152.24	-22.32	179.86	0.981	0.040	0.218	0.585	0.109
Sol	9	1	70.66	98.96	88.97	0.019	0.097	0.718	0.584	0.108

**Taula 3.8. Solucions de les funcions de rotació i translació per Molrep a 3Å.** Obtingudes utilitzant com a model inicial el domini C-terminal resolt independentment.

Vam afinar aquestes dues posicions per afinament del cos rígid utilitzant el programa Refmac5. El model que contenia els dos dominis C-terminals afinats es va incloure en un nou cicle de Molrep com a model fixat i es va iniciar la cerca a 3.5 Å dels corresponents dos dominis N-terminals, utilitzant com a model el domini N-terminal de l' *A.tumefaciens*. Aquest segon model presentava un 21% d'homologia de seqüència amb el domini homòleg de *P.abysssi*. Per a millorar el model, es mantingueren les cadenes laterals de les zones idèntiques a PaGS, s'eliminaren les zones que no eren presents a PaGS i la resta fou convertit a cadena de polialanines, tal i com s'indica a Materials i Mètodes (apartat J.2.1). Vam obtenir dues solucions per a les dues posicions dels dominis N. Aquest resultat era molt encoratjador ja que de totes les posicions possibles a la cel·la, els dos dominis N-terminals presentaven una connectivitat perfecta amb els corresponents dominis C-terminals. Els dos monòmers híbrids estaven relacionats per 114°. Vam afinar aquest nou model per afinament del cos rígid, considerant els dominis N i C-terminals com a cossos independents.

Atès que ens faltava trobar el tercer component del trímer, ens vam plantejar tornar a utilitzar Molrep i vam usar com a model inicial a 3-5 Å de resolució el nou monòmer híbrid que havíem construït. El programa va donar tres solucions inequívokes que van permetre traçar un primer model trimèric. L'empaquetament d'aquesta primera solució era correcte i gairebé no s'observava cap contacte cristal·logràfic (Taula 3.9)

Sol_RF	theta	phi	chi	alpha	beta	gamma	Rf	Rf/sigma
Sol_RF 1	0.00	-45.00	1.13	190.57	0.00	170.57	21.47	10.32
Sol_RF 2	143.60	-50.97	115.54	167.10	60.26	89.04	13.58	6.53
Sol_RF 3	35.23	131.30	122.35	97.33	60.71	14.73	11.79	5.67
Sol_RF 4	146.33	-28.80	127.21	182.01	59.56	59.61	8.685	4.18
Sol_RF 5	145.56	-34.32	114.13	183.83	56.68	72.47	8.421	4.05

Monòmer 1									
Sol_RF	TF	theta	phi	chi	tx	ty	tz	Rfac	Corr
Sol__1__1	1	0.00	-45.00	1.13	0.265	0.000	0.094	0.572	0.136
Sol__2__1	2	143.60	-50.97	115.54	0.164	0.000	0.311	0.577	0.110

Monòmer 2 (fixat Monòmer 1)									
Sol_RF	TF	theta	phi	chi	tx	ty	tz	Rfac	Corr
Sol__2__1	1	143.60	-50.97	115.54	0.664	0.075	0.311	0.543	0.215
Sol__3__1	2	35.23	131.30	122.35	0.552	0.353	0.439	0.555	0.182

Monòmer 3 (fixats Monòmer 1 i 2)									
Sol_RF	TF	theta	phi	chi	tx	ty	tz	Rfac	Corr
Sol__3__1	1	35.23	131.30	122.35	0.551	0.351	0.439	0.520	0.282
Sol__5__3	2	145.56	-34.32	114.13	0.862	0.102	0.445	0.550	0.198

**Taula 3.9. Solucions de les funcions de rotació i translació per Molrep a 3Å.** Obtinguts utilitzant com a model inicial la molècula híbrida construïda a partir del domini C-terminal de la PaGS i el domini N-terminal de l'AtGS amb polialanina posicionats per Molrep.

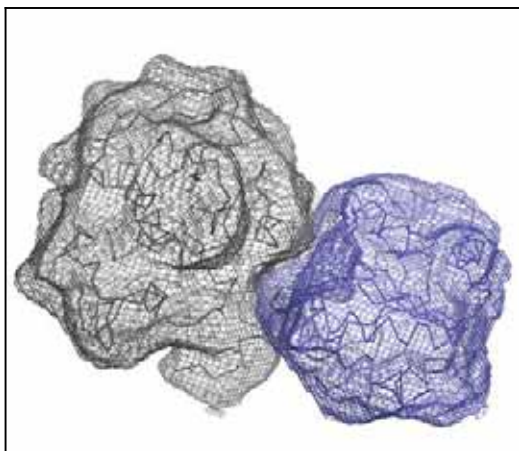
Va ser necessari utilitzar l'afinament del cos rígid amb el programa Refmac5 (CCP4), al voltant de les millors posicions i orientacions obtingudes anteriorment. Es van afinar cadascuna de les sis dimensions (tres rotacions i tres translacions) per a cada cos rígid (donat que es tractava del grup espacial C2, en el càlcul del primer cos, només s'afinaren cinc dimensions).

Per a poder estudiar les diferències dels angles entre els dominis N i C-terminals, vam plantejar una aproximació amb sis cossos rígids (tres dominis N-terminals i tres dominis C-terminals). Després de 20 cicles, el valor de  $R_{\text{factor}}$  havia baixat a 48.3, mentre que el valor de  $R_{\text{free}}$  era de 48.2.

A continuació, vam estendre les fases de forma iterativa mitjançant el programa DM (CCP4). Per a modular l'efecte del solvent sobre les amplituds, es van construir màscares que separessin les zones de proteïna de les de solvent desordenat. Com es mostra a la figura 3.9, vam calcular dues màscares utilitzant el programa NCSMask al voltant d'un domini N-terminal i d'un domini C-terminal. La màscara del domini C-terminal la vam fer més ajustada, perfilant a una distància de 2 Å, ja que aquesta part del model es preveia gairebé idèntica i disposava de les cadenes laterals. En canvi, la màscara pel domini N-terminal era més ample, perfilant a una distància de 3 Å, i deixava espai suficient per a les cadenes laterals que no havien estat contemplades en el model amb tot alanines.

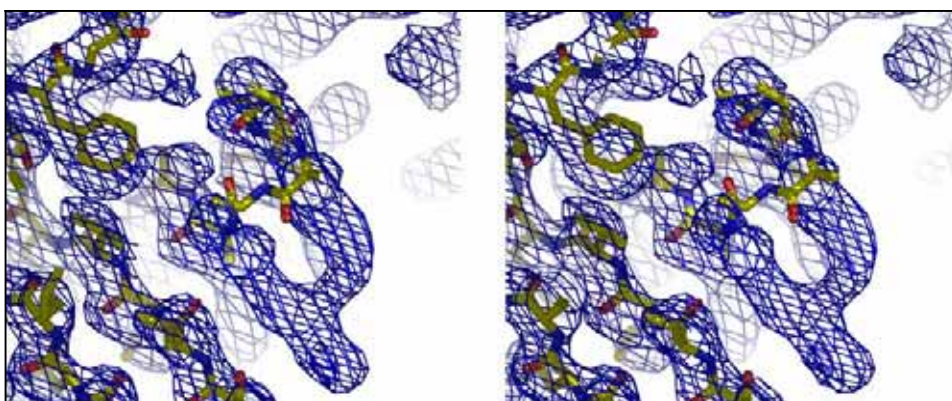
Es van aplicar les transformacions que relacionaven els diferents elements amb simetria no-cristal·logràfica en la unitat asimètrica, és a dir les que relacionaven els tres

N-terminals i els tres C-terminals. D'aquesta manera, durant l'afinat es podien sobreposar i promitjar aquestes densitats.



**Fig. 3.9. Màscares dels dominis N- i C-terminal de la PaGS.** Màscares del domini N-terminal (gris) i del domini C-terminal (blau) calculades amb el programa NCSMask, tal i com es descriu a Materials i Mètodes.

Aquest terme promitjat de la simetria no-cristal·logràfica va resultar molt efectiu i es combinà amb els criteris de distribució de la densitat electrònica (*histogram matching*), l'aplanament de solvent (*solvent flattening*) i la connectivitat entre zones contínues de més alta densitat (*skeletonization*). D'aquesta forma, vam obtenir els primers mapes interpretables. Les densitats electròniques així calculades no seguien exactament el model i en les zones que només havíem utilitzat polialanines, es podien distingir les cadenes laterals dels aminoàcids, indicant que havíem arribat a una solució real i no a una sobreestimació del model (Fig. 3.10).



**Fig. 3.10. Diferències entre el mapa de densitat electrònica i el model.** Dibuix estereogràfic del mapa de densitat electrònica calculat per DM sobre el model utilitzat. S'observa clarament que el mapa conté informació addicional.

El traçat de l'estructura es feu manualment, usant el programa gràfic O (Jones *et al.* 1991). Novament de forma iterativa, es van anar incloent etapes d'afinat de les posicions dels àtoms així com de la seva estereoquímica, usant Refmac5 (CCP4). Durant l'afinament es van utilitzar les dades completes de difracció, excepte un 5% escollit a

l'atzar que s'utilitzà per a calcular el  $R_{free}$ . Es va utilitzar la simetria no-cristal·logràfica com a restricció en l'afinament i es va anar alleugerint a mida que avançava el procés. Cap al final de l'afinament es va aplicar l'afinament anisotròpic per TLS, considerant la presència de dos cossos rígids per molècula (N- i C-terminal).

Afinament	PaGS
Marge de resolució (Å)	20 – 2.8
Reflexions (%)	38891 (2822)
R factor (%)	20.0 (26.6)
B mitjà (Å <sup>2</sup> )	49.0
<b>Nombre d'àtoms</b>	10569
Àtoms de proteïna	10449
Molècules de glucosa	2
Molècules d'UDP	1
Molècules de dioxà	9
Molècules de glicerol	7
<b>RMSD dels valors ideals</b>	
Rmsd enllaç (Å)	0.012
Rmsd angle (°)	1.4
<b>Diagrama de Ramachandran</b>	
Més afavorits (%)	88.7
Permesos (%)	10.8
Permesos generosament(%)	0.3
Prohibits (%)	0.2

**Taula 3.10. Afinament de l'estructura de la PaGS.**

Com s'observa a la taula 3.10, el model final contenia 440 residus per monòmer que corresponien als 437 aminoàcids del monòmer, des de la Met 1 fins a la Leu 437, més una extensió en el seu N-terminal de tres residus (Gly-Ser-His) procedents del vector.

El mapa de densitat electrònica també mostrava la presència de 9 molècules de dioxà i 7 molècules de glicerol per unitat asimètrica, procedents del tampó de cristallització. A més es van poder posicionar dues molècules de glucosa i una d'UDP del substrat UDP-glucosa afegit al tampó de cristallització. La qualitat estereoquímica del model va ser analitzada pel programa Procheck (Laskowski *et al.* 1993), mostrant que només un 0.2% dels aminoàcids presentaven conformacions de la cadena principal situades en zones prohibides del diagrama de Ramachandran (Fig. 3.11).



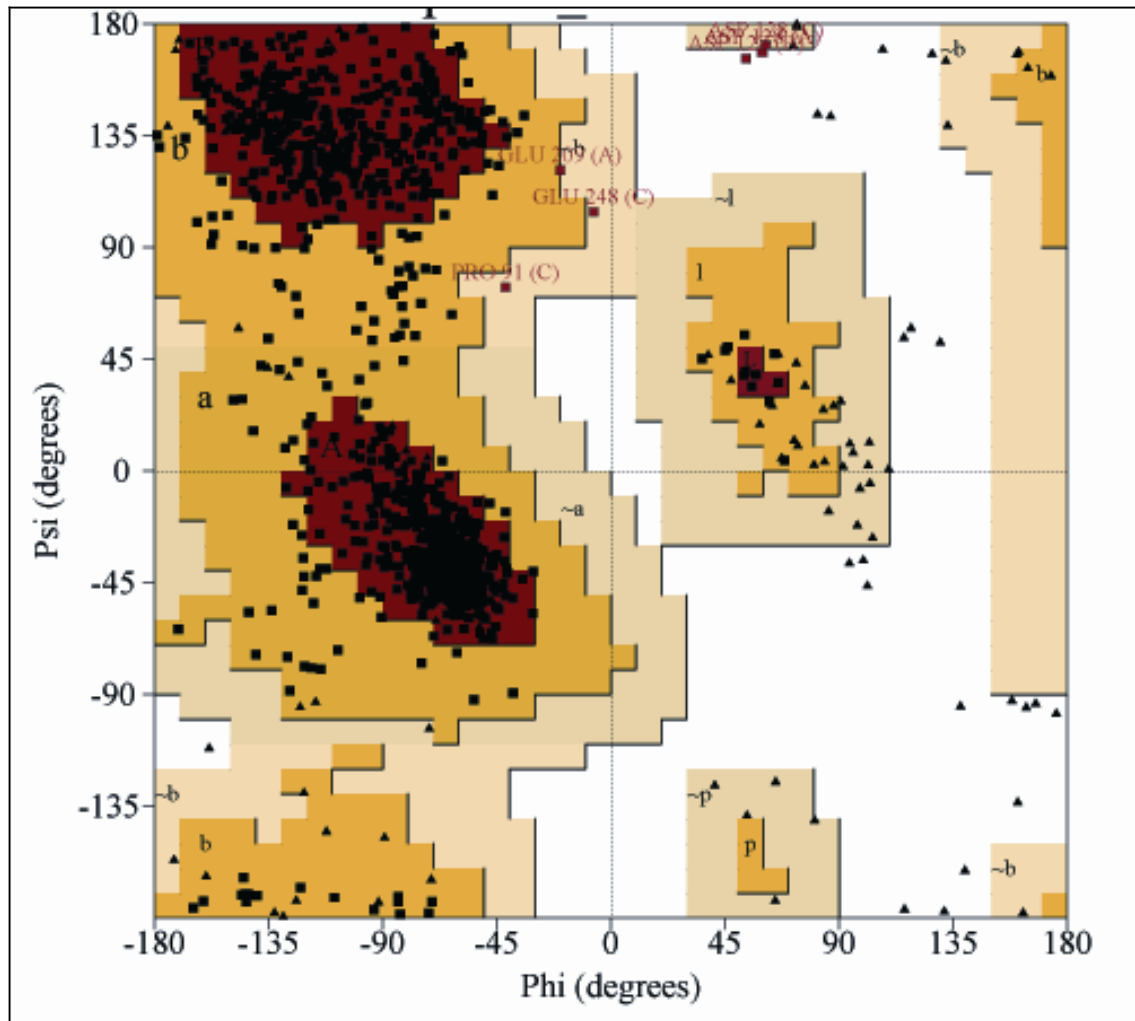


Fig. 3.11. Diagrama de Ramachandran de l'estructura de la PaGS.

## DISCUSSIÓ

Els cristalls del domini C-terminal de la PaGS empaqueten de forma molt compacte en el grup espacial  $P2_1$  amb un monòmer per unitat asimètrica. Aquests cristalls difracten a molt alta resolució.

Hem resolt l'estructura del domini C-terminal aïllat de la PaGS mitjançant el mètode de reemplaçament molecular, tot i no disposar d'un model amb elevada similitud de seqüència. L'existència d'un únic monòmer a la unitat asimètrica, la qualitat de l'espectre i la similitud de l'estructura tridimensional amb el domini C-terminal de l'AtGS, utilitzat com a model tot i la baixa identitat de seqüència, han permès trobar la solució correcta.

L'elevada resolució de la difracció ha permès determinar les configuracions dobles d'alguns residus i la majoria d'aigües del sistema, ja que les primeres esferes de

solvatació d'un monòmer se solapen amb les de la molècula veïna, donat l'elevat grau de compactació del cristall.

La resolució de la PaGS ha resultat més complexa, en bona part degut a la naturalesa anisotròpica dels cristalls. La difracció no homogènia en les diferents orientacions del cristall ha impossibilitat la resolució per mètodes de difracció anòmala, ja que el senyal anòmal quedava emmascarat sota el soroll de les dades experimentals. Tampoc ha estat possible la resolució per reemplaçament isomorf.

Tot i així, el fet de que la PaGS empaqueti en el grup espacial C2 amb tres monòmers per unitat asimètrica, relacionats per simetria no-cristal·logràfica, ha permès aplicar les restriccions d'aquesta simetria a la resolució per reemplaçament, a l'extensió de les fases i a l'afinat.

Els intents per a aplicar les restriccions de la simetria no-cristal·logràfica a la resolució per reemplaçament molecular ens van portar a utilitzar les funcions de rotació i translació bloquejades. Donada la poca similitud de la seqüència del model emprat, l'AtGS, no vam poder trobar la solució *a priori*. Tot i així, la capacitat de discriminació aplicant aquests criteris de simetria no-cristal·logràfica ha quedat palesa posteriorment per estudis realitzats en el laboratori del Dr. Navaza. La utilització de la reconstrucció del trímer a partir de les imatges de microscòpia electrònica (capítol 1) com a model inicial per al programa de reemplaçament molecular AMoRe, l'extensió teòrica de les fases a alta resolució i l'aplicació de les restriccions no-cristal·logràfiques els ha permès determinar la solució de reemplaçament correcta (Dr. Navaza *et al*, resultats no publicats).

Per a determinar l'estructura de la PaGS va caldre però obtenir un model més similar, per aquest motiu s'utilitzà l'aproximació poc comú de resoldre en primer lloc el domini C-terminal i a continuació resoldre la proteïna sencera per MR, utilitzant el model híbrid format pel domini N-terminal de l'AtGS i el domini C-terminal de la pròpia PaGS.

D'aquesta manera, vam resoldre les estructures de la PaGS i del seu domini C-terminal a 2.8 Å i 1.7 Å, l'anàlisi de les quals es mostra en el següent capítol.

## CAPÍTOL 4: DESCRIPCIÓ ESTRUCTURAL DE LA GLICOGEN SINTASA DE *PYROCOCCUS ABYSSI*.

### INTRODUCCIÓ

La GS d'*A. tumefaciens* (AtGS) ha estat la primera estructura tridimensional resolta d'un membre de les famílies GT-3 i GT-5 (Buschiazzo *et al.* 2004). Com la resta de GS, pertany al grup de proteïnes amb plegament tipus GT-B que actuen amb retenció de la configuració del sucre transferit. Dins d'aquest grup han estat resoltes les estructures de la GP, la MalP i l'OtsA (Fletterick *et al.* 1976; Watson *et al.* 1997; Gibson *et al.* 2002). La resolució de l'estructura de la PaGS amplia el nombre d'estructures d'aquest grup resoltes i obre noves vies en l'estudi dels seus mecanismes catalítics i reguladors.

Mentre s'accepta que les glicosidases que retenen la configuració del centre anomèric operen a través d'un mecanisme de doble desplaçament, el qual implica la formació i el trencament d'un intermedi covalent amb l'enzim (Vocadlo *et al.* 2001) o amb el substrat (Zechel i Withers 2000), no existeix evidència experimental d'un mecanisme similar a les GT. A més, els residus del centre actiu de les estructures de GT resoltes no estan col·locats adequadament per a poder dur a terme els rols necessaris d'àcid/base i de nucleòfil en un mecanisme d'aquest tipus.

Aquestes observacions han portat a proposar un mecanisme alternatiu tipus  $S_Ni$ , en el que l'atac del grup nucleofílic seria per la mateixa cara que el grup sortint de manera concertada, però necessàriament asíncrona (Gibson *et al.* 2002). Aquest mecanisme tampoc disposa encara d'evidències experimentals.

Tot i així s'ha obtingut recentment un resultat inesperat per a la lipopolisacàril- $\alpha$ -1,4-galactosiltransferasa C de *Neisseria meningitidis* (LgtC), una GT amb plegament tipus GT-A, que actua amb retenció de la configuració anomèrica. La incubació del seu substrat donador, UDP-galactosa (UDP-Gal) amb l'enzim mutat en Gln189 per Glu dona lloc a un intermedi en el que el residu galactosil es troba covalentment unit al residu adjacent, Asp180, la cadena lateral del qual es troba a 8,9 Å del substrat donador en l'estructura cristal·logràfica disponible (Lairson *et al.* 2004). Per tant, és possible que les orientacions relatives i les distàncies dels residus catalítics en aquests enzims experimentin canvis importants en resposta a la unió dels seus substrats.

Mentre la LgtC (Ly *et al.* 2002) i altres membres de la superfamília GT-A actuen segons un mecanisme cinètic ordenat tipus bi-bi en el que s'ha d'unir en primer lloc el donador de glicosil, les GS eucariotes (McVerry i Kim 1974; Plesner *et al.* 1974) i la GP (Davies 1997) presenten un mecanisme aleatori d'equilibri ràpid. En aquest cas, els dos substrats han d'estar units a l'enzim sense un ordre preestablert per a que es produeixi la reacció. Sembla, doncs, que en el cas de les GS, i possiblement en altres GT-B que operen amb retenció de configuració, és la unió dels dos substrats la causa de la

reorientació dels residus en el centre actiu que fa que aquests adoptin les posicions adequades per a la catàlisi.

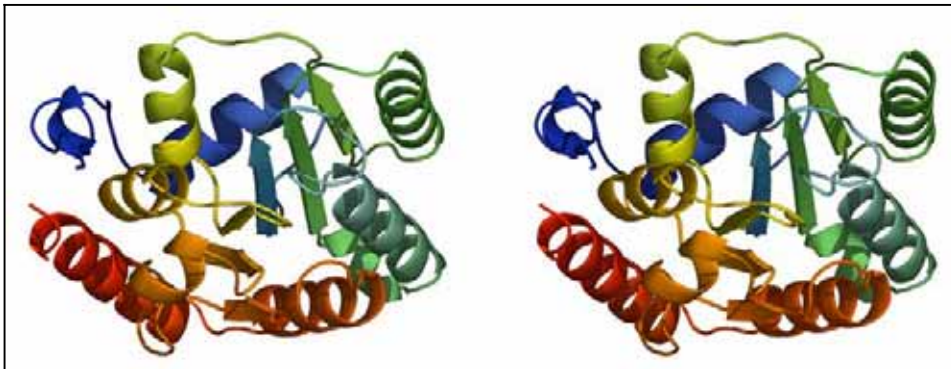
A més de les incògnites que desperta el seu mecanisme, hi ha molts altres aspectes de les GS encara desconeguts, com ara la selectivitat dels diferents membres per diferents donadors de sucre, els seus estats d'oligomerització, la unió a glicogen o la seva regulació. En aquest capítol s'analitza l'estructura de la PaGS i s'aprofundeix en l'estudi d'aquests aspectes.

## RESULTATS

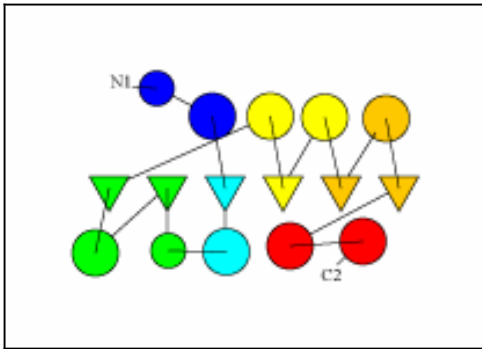
### 4.1. Descripció general:

#### 4.1.1. Estructura del domini C-terminal de la PaGS:

El domini C-terminal de la PaGS es va determinar a 1.8 Å de resolució (Fig. 4.1). Aquest domini s'estructura en forma de plegament de Rossmann (*Rossmann fold*). El cor del domini C està compost per sis fulles  $\beta$  paral·leles envoltades per nou hèlix  $\alpha$  (Fig. 4.1).



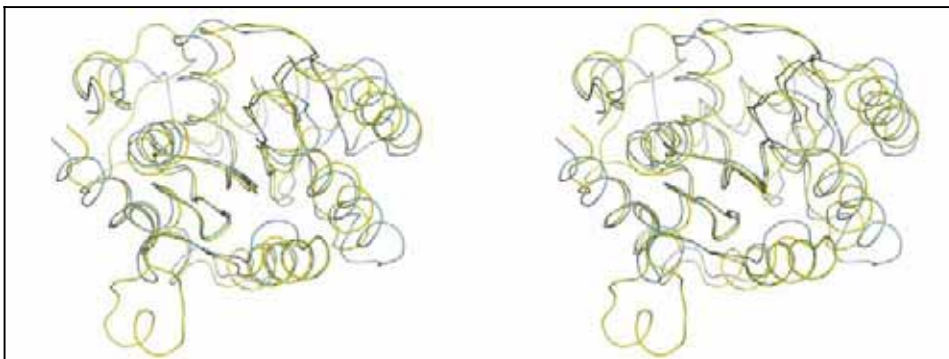
**Fig. 4.1. Estructura del domini C-terminal de la glicogen sintasa de *P. abyssi*.** Dibuix estereogràfic que mostra el plegament general i l'estructura secundària del domini C-terminal de la PaGS. La gradació de colors del blau al vermell indica la direcció de N- a C-terminal.



**Fig. 4.2. Topologia del domini C-terminal de la glicogen sintasa de *P. abyssi*.** Esquema de la disposició de l'estructura secundària del domini, realitzada amb el programa TOPS. Els triangles indiquen els fulls  $\beta$  i els cercles les hèlixs  $\alpha$ . La gradació de colors del blau al vermell indica la direcció de N- a C-terminal.

Aquest tipus de plegament és molt similar al del domini C-terminal de la majoria de glicosiltransferases resoltes fins ara que s'engloben en la superfamília GT-B. Totes elles pertanyen a la classe de les proteïnes alfa/beta (Classificació estructural de les proteïnes, SCOP. <http://scop.mrc-lmb.cam.ac.uk/scop>).

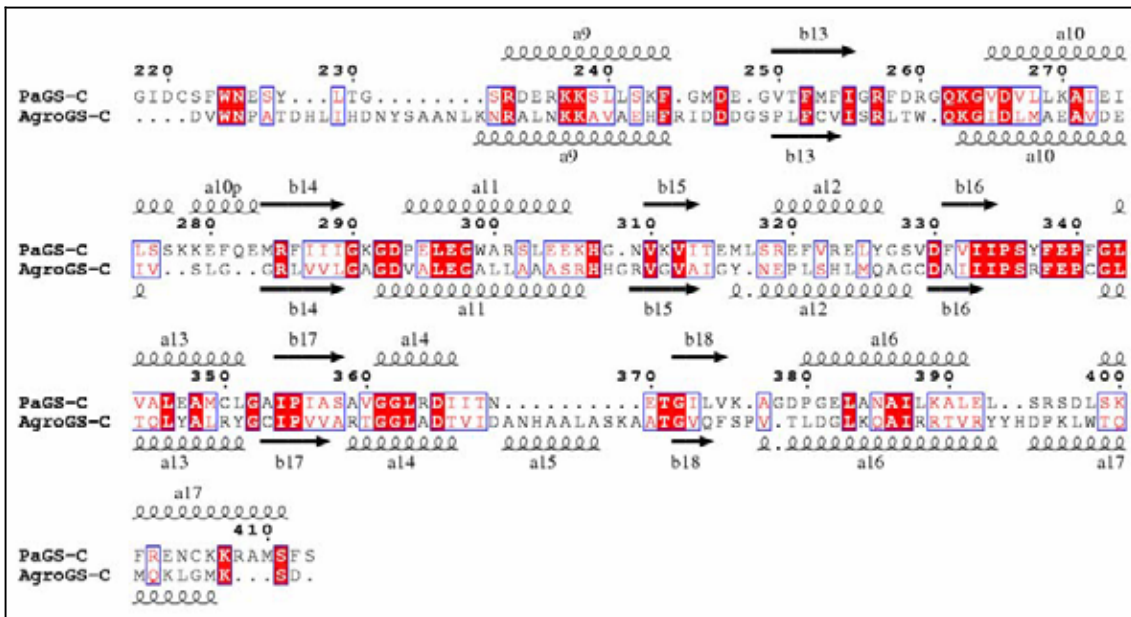
Tot i la baixa homologia de seqüència (22%), en superposar l'estructura del domini C-terminal de PaGS amb el seu homòleg d'*A. tumefaciens* es comprova que ambdós tenen un plegament molt similar amb una desviació estàndard de 1.3 Å per 183 residus (86%) estructuralment equivalents (Figura 4.3).



**Fig. 4.3. Superposició dels dominis C-terminal de PaGS i AtGS.** Dibuix estereogràfic de la superposició de les estructures tridimensionals de la PaGS (blau) i l' AtGS (groc), mitjançant el programa SHP. Es mostra la vista en que l'extrem C-terminal del domini està més proper a l'observador.

Com es mostra a la figura 4.3, els fulls  $\beta$  centrals tenen una topologia idèntica en els dos dominis C-terminal de les glicogen sintases de *P.abyssi* i d'*A. tumefaciens*, mentre que s'aprecien algunes diferències respecte a l'emboïllat format per les hèlixs  $\alpha$ .

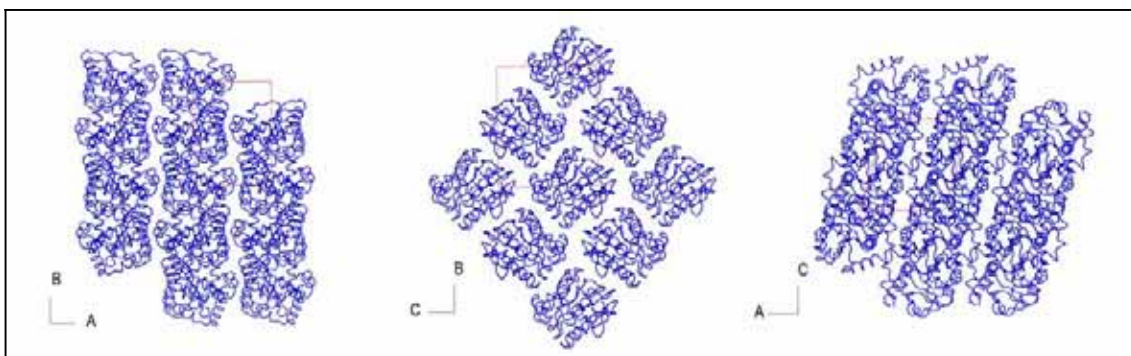
En l'alineament estructural entre ambdós dominis (Fig. 4.4) s'observa l'elevada equivalència estructural, exceptuant l'hèlix  $\alpha_{10p}$  del domini C-terminal de PaGS que no és present en el seu homòleg AtGS. A més a més, el domini C-terminal de l'AtGS conté una inserció que correspon a l'hèlix  $\alpha_{15}$  (entre la hèlix  $\alpha_{14}$  i el full  $\beta_{18}$ ) responsable de l'oligomerització d'AtGS i que serà tractada amb més detall a l'apartat 4.6.2.



**Fig. 4.4. Alineament estructural del domini C-terminal de la PaGS amb el de la AtGS.** Les regions estructuralment equivalents apareixen enquadrades. Els residus conservats es mostren de color blanc sobre fons vermell, mentre que els residus similars es mostren de color vermell sobre fons blanc. Els elements d'estructura secundària (hèlixes  $\alpha$  i fulls  $\beta$ , marcats amb les lletres a i b, respectivament) de la PaGS i de l'AtGS es mostren a la part superior i inferior dels alineaments, respectivament.

#### 4.1.2. Empaquetament cristal·lí del domini C-terminal de la PaGS:

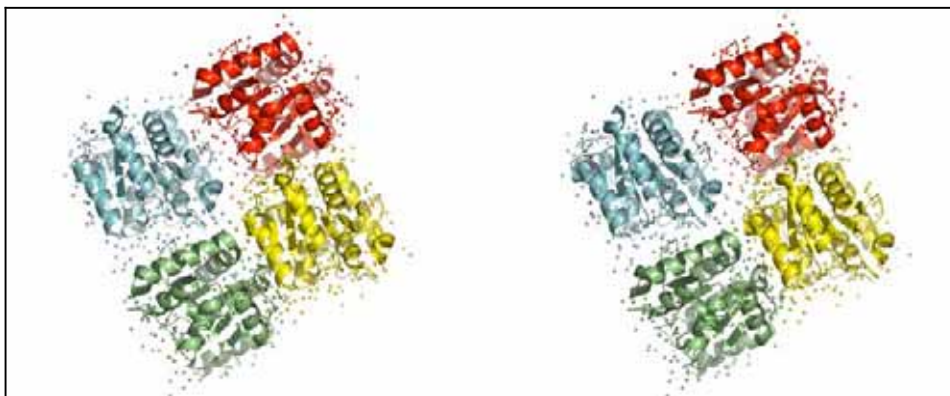
Com es mostra a la figura 4.5, el cristall format pel domini C-terminal de la PaGS és molt dens en totes les direccions. En el tipus d'empaquetament  $P2_1$ , les molècules es troben disposades de forma helicoidal al voltant d'un eix imaginari paral·lel a l'eix cristal·logràfic B.



**Fig. 4.5. Empaquetament cristal·lí del domini C-terminal de la PaGS.** Vistes al llarg de l'eix C, A i B respectivament del cristall del domini C-terminal. En vermell apareix dibuixada la cel·la unitària.

L'elevat grau d'empaquetament d'aquest cristall provoca que els monòmers s'ordenin molt propers uns als altres. Aquest fet i la difracció a alta resolució van permetre determinar gairebé totes les aigües del conjunt. Les esferes de solvatació d'un monòmer es sobreposen a les del veí, per la qual cosa no queden molècules de solvent

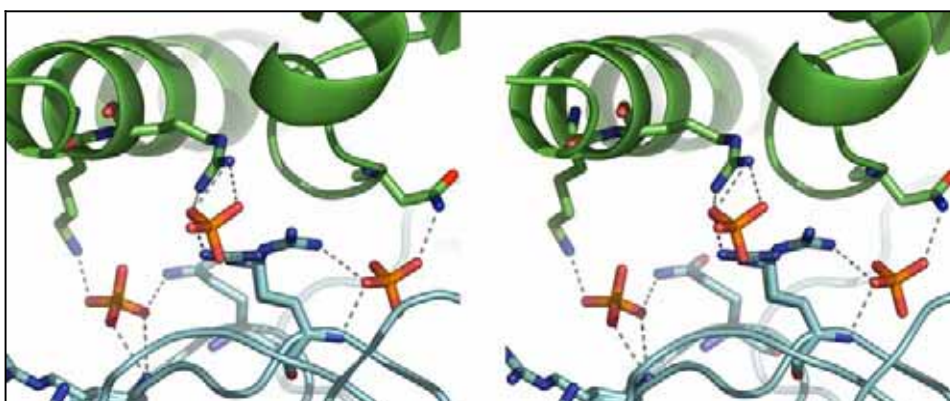
desordenades. D'aquesta forma es van localitzar gairebé 200 molècules d'aigua per molècula (fig. 4.6).



**Fig. 4.6. Esferes de solvatació en el cristall del domini C-terminal de la PaGS.** Representació estereogràfica de la superposició d'esferes de solvatació de quatre monòmers diferents. Les molècules d'aigua es representen mitjançant boles del color del monòmer al qual s'assignen.

A més de les molècules de solvent, va ser possible localitzar molècules del tampó de cristallització que comparteixen interaccions amb dos monòmers veïns. Com es mostra a la figura 4.7, tres sulfats es coordinen a la interfase, establint interaccions polars entre els seus oxígens i els nitrògens de la proteïna. En aquests tres casos, els grups sulfat només utilitzen tres dels quatre oxígens disponibles.

Segons la figura 4.7, un oxigen de l'ió sulfat esquerre interacciona amb el nitrogen de la cadena lateral de la Lys407 d'un monòmer, i amb els nitrògens de la cadena lateral de Gln262 i la cadena principal de l'Arg260 de l'altre monòmer.



**Fig. 4.7. Sulfats compartits per dos monòmers en el cristall.** Dibuix estereogràfic que mostra la localització de tres sulfats (taronja) compartits entre dos monòmers de color blau (B) i verd (V). En línies discontinües s'indiquen les interaccions polars entre els sulfats i els residus proteics. El sulfat de l'esquerra interacciona amb la Lys407 (V), amb el N de la cadena principal de l'Arg260 (B) i amb la Gln262 (B). El sulfat central interacciona amb l'Arg408 (V) i amb l'Arg257 en una de les seves dues configuracions (B). El sulfat de la dreta interacciona amb l'Asn369 i amb l'Arg257 en l'altre configuració (B).

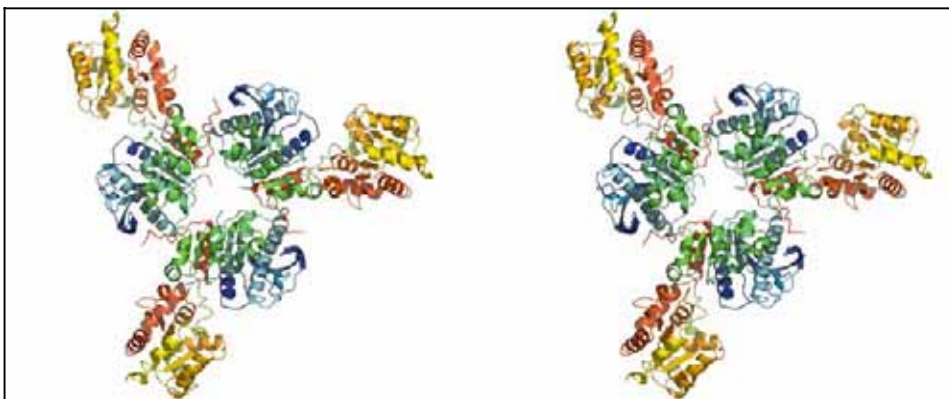
Dos dels oxígens del sulfat central interaccionen amb els dos nitrògens de l'Arg408 del monòmer superior (verd) mentre que un altre dels oxígens ho fa amb un nitrogen de l'Arg257 de l'altre monòmer (blau), quan aquest residu pren la configuració més propera al sulfat central.

En la configuració alternativa, l'Arg257 (blau) estableix un pont polar amb un dels oxígens del sulfat dret, que al mateix temps interacciona amb el nitrogen de la cadena principal del mateix residu. Per últim, el tercer oxigen interacciona amb el nitrogen de la cadena lateral de l'Asn369 de l'altre monòmer (verd). La possible rellevància biològica d'aquestes interaccions es tractarà amb detall als apartats 4.3.1, 4.4.2 i 4.8.

L'obtenció de l'espectre de difracció a alta resolució ha permès determinar la mobilitat d'un cert nombre de residus, que presenten dobles configuracions per a les seves cadenes laterals (apartat 3.1.2). Com mostrarem més endavant, la majoria d'aquests residus estan implicats en la interacció amb substrat. Un cas interessant és la doble configuració de la Cys221 que no forma enllaç disulfur amb la Cys350 pròxima (Fig. 3.4, capítol 3), contràriament al que succeeix a la proteïna sencera, en la qual aquests dos residus estan enllaçats. La no formació d'aquest pont a l'estructura es podria deure al diferent entorn en què es troba el domini C-terminal aïllat respecte a l'integrat en la proteïna sencera.

#### 4.1.3. Estructura de la glicogen sintasa de *P.abysyi*:

L'estructura de la PaGS trimèrica es determinà a 2.8 Å de resolució. Cadascuna de les subunitats del trímer es plega en dos dominis tipus plegament de Rossmann (*Rossmann fold*). Aquests dos dominis estan separats per un solc molt profund on es situa el centre catalític de l'enzim (Fig. 4.8)

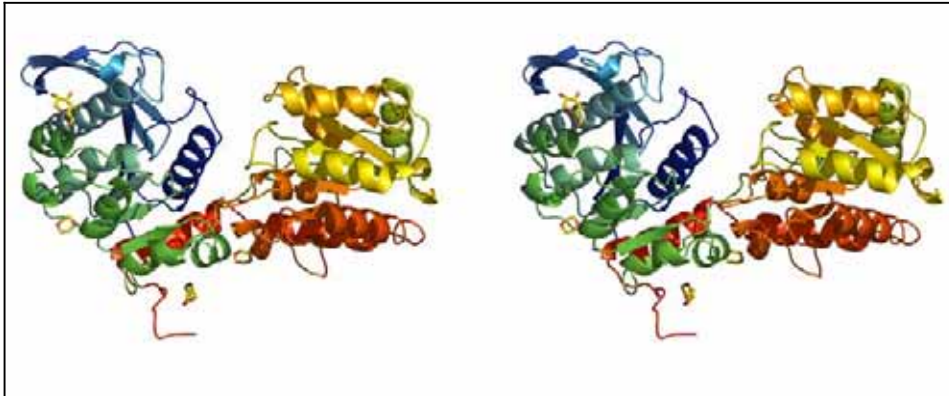


**Fig 4.8. Estructura de la glicogen sintasa de *P.abysyi*.** Dibuix estereogràfic que mostra el plegament general i l'estructura secundària del glicogen sintasa de *P.abysyi* GS. La gradació de colors del blau al vermell indica la direcció de N- a C-terminal per a cada monòmer.

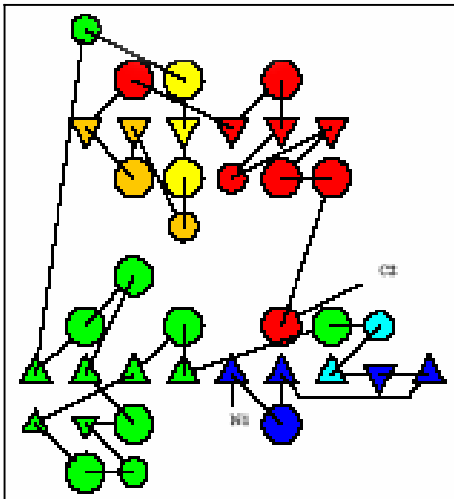
En cada monòmer de PaGS, el domini N-terminal engloba els residus des de la Met1 fins a l'Asn217. El domini C-terminal ha estat definit entre els residus Gly218 i Ser413. La última hèlix del domini C-terminal (hèlix  $\alpha$  17) experimenta un gir i torna cap al domini N-



terminal. A partir de la posició 413 i fins al residu Thr426 forma una hèlix  $\alpha$  que creua el domini N-terminal (hèlix  $\alpha$  18) i en què s'engloba (Fig. 4.9).



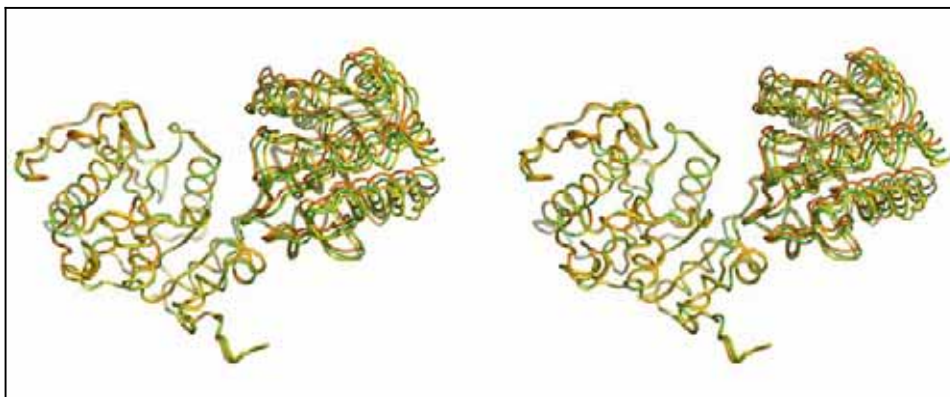
**Fig. 4.9. Estructura d'una subunitat de la glicogen sintasa de P.abysyi.** Dibuix estereogràfic que mostra el plegament general i l'estructura secundària de la subunitat A de la glicogen sintasa de P.abysyi GS. La gradació de colors del blau al vermell indica la direcció de N- a C-terminal.



**Fig. 4.10. Topologia d'una subunitat de la glicogen sintasa de P. abysyi.** Esquema de la disposició de l'estructura secundària d'una subunitat, realitzada amb el programa TOPS. Els triangles indiquen els fulls  $\beta$  i els cercles les hèlixes  $\alpha$ . La gradació de colors del blau al vermell indica la direcció de N- a C-terminal.

En la PaGS, l'hèlix  $\alpha$ 18 s'estén en forma de cua protuberant formada per residus hidrofòbics (Gly427-Leu437) i interactua amb una cavitat hidrofòbica del domini N-terminal de la subunitat veïna (Leu53-Arg55, Leu109-Leu120). D'aquesta manera, i com s'observa a la figura 4.8, els tres dominis N-terminal (tons blaus i verds) estan units molt estretament mentre que els dominis C-terminals (tons vermells i grocs) no participen en les interaccions entre subunitats. Segons aquest model trimèric, els moviments entre els dos dominis no estarien restringits i els dominis C-terminals tindrien llibertat per a apropar-se i allunyar-se, i donarien lloc a un tancament i a una obertura del sot catalític, respectivament. De fet, els dos dominis estan units per només dues cadenes d'aminoàcids que corresponen als residus Asn217-Gly218-Ile219 en la direcció del domini N- al C-terminal, i els residus Phe412-Ser413-Trp414, entre les hèlix  $\alpha$ 17 i  $\alpha$ 18. Aquest fet donaria flexibilitat al sistema.

Aquesta llibertat de moviment es reflecteix en el fet que les tres subunitats del trímer presenten diferents graus d'obertura. Com es mostra a la figura 4.11, la superposició dels dominis N-terminal de les tres subunitats col·loca els dominis C-terminal en llocs lleugerament diferents. Aquesta superposició és representativa de la situació exposada anteriorment, en la que els dominis C-terminal es podrien moure respecte als dominis N-terminal fixos.



**Fig. 4.11. Superposició segons els dominis N-terminal de les subunitats del trímer de la GS de P.abysyi.** Dibuix estereogràfic que mostra la superposició dels dominis N-terminal de les subunitats A (taronja), B (groc) i C (verd) del trímer de la PaGS amb el programa SHP, com es descriu a Materials i Mètodes.

Mentre els dominis N-terminal es relacionen per l'eix ternari molecular amb molta precisió, els dominis C-terminal presenten desviacions a aquesta simetria degudes a la variabilitat en el grau d'obertura molecular. La taula 4.1 posa de manifest aquestes desviacions en el càlcul de l'angle chi, el qual relaciona els dominis N- i C-terminal entre les tres subunitats.

Subunitats	Domini N-terminal		Domini C-terminal	
	RMSD	angle chi	RMSD	angle chi
A – B	0.26	118.2	0.42	113.7
A – C	0.27	120.1	0.45	121.9
B – C	0.25	120.7	0.50	124.4

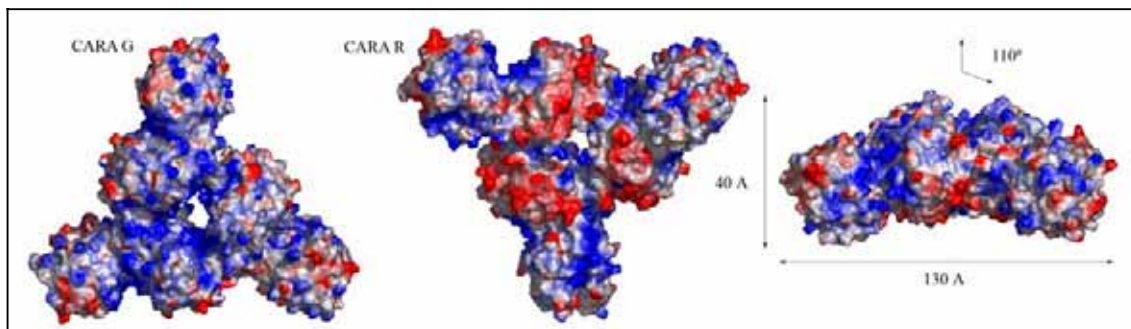
**Taula 4.1. Superposicions dels dominis N i C-terminal entre subunitats.** Valors de desviació estàndard (RMSD) i angles (chi) que relacionen els dominis N i C-terminal de les subunitats A, B i C. Aquests valors s'han obtingut amb el programa SHP, com es descriu a Materials i Mètodes.

Analizant aquests valors, s'observa que els dominis N- i C-terminal de la subunitat A es relacionen amb els de la subunitat C amb una major precisió de simetria ternària que la subunitat B amb les dues anteriors. Aquest comportament ha estat justificat per la naturalesa dels contactes cristal·logràfics, que serà tractada en l'apartat 4.1.4.

Com s'observa a la figura 4.9, la disposició de les subunitats en el trímer de PaGS, amb l'eix ternari pràcticament perpendicular als eixos longitudinals de cada monòmer (aproximadament  $110^\circ$ ), fa que la molècula tingui una forma triangular molt evident, amb uns 130 Å de costat i un gruix d'uns 40 Å. La forma del trímer permet definir dues cares, que anomenarem G (per glicogen) i R (per reguladora) pels motius que s'exposen en els apartats 4.5 i 4.9, i seguint la descripció feta pel cas similar de la GP (Rath *et al.* 2000).

Prenent com a referència la vista de la cara G, el solc catalític que separa cada subunitat N- i C-terminal es situa de costat i s'enfoca cap a l'exterior de la molècula. Segons el model plantejat de moviment, els dominis C-terminals es mouran en el sentit contrari a les agulles del rellotge per a tancar i en sentit horari per a obrir el solc catalític.

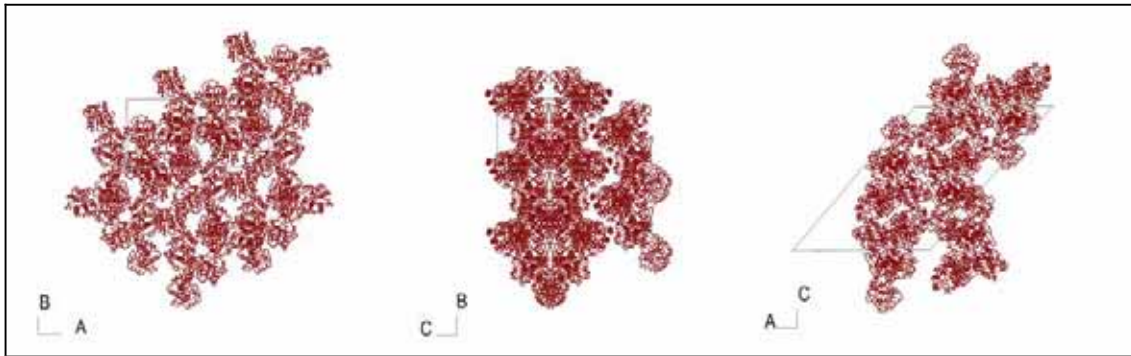
La distribució de càrregues entre les cares G i R és molt diferent. Els residus exposats a la cara G tenen un caràcter majoritàriament bàsic, mentre que en la cara R residus àcids s'acumulen en els dominis N-terminals (centrals). El solc catalític té una naturalesa principalment bàsica.



**Fig. 4.12. Representació de la superfície molecular de la PaGS.** Superfícies calculades amb el programa Pymol amb vistes de la cares G, R i visió lateral on s'indiquen les dimensions del trímer i l'angle que forma l'eix ternari amb el pla molecular. El color blau indica les zones positives, el vermell les negatives i el gris les neutres.

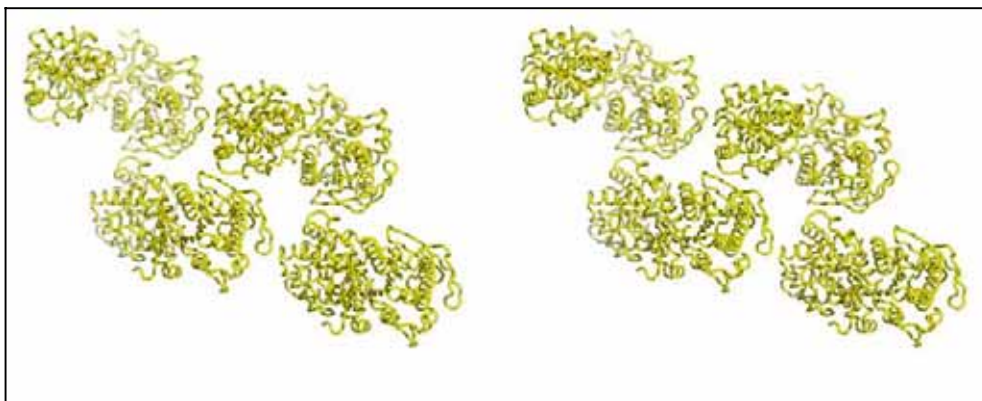
#### 4.1.4. Empaquetament cristal·lí de la PaGS:

L'empaquetament cristal·lí de la PaGS s'engloba en el grup espacial C2. Amb un trímer per unitat asimètrica, la xarxa cristal·lina és especialment densa en la direcció de l'eix cristal·logràfic B.



**Fig. 4.13. Empaquetament cristal·lí de la glicogen sintasa de *P.abysssi*.** Vistes al llarg dels eixos C, A i B respectivament. En gris es representa la cel·la unitària.

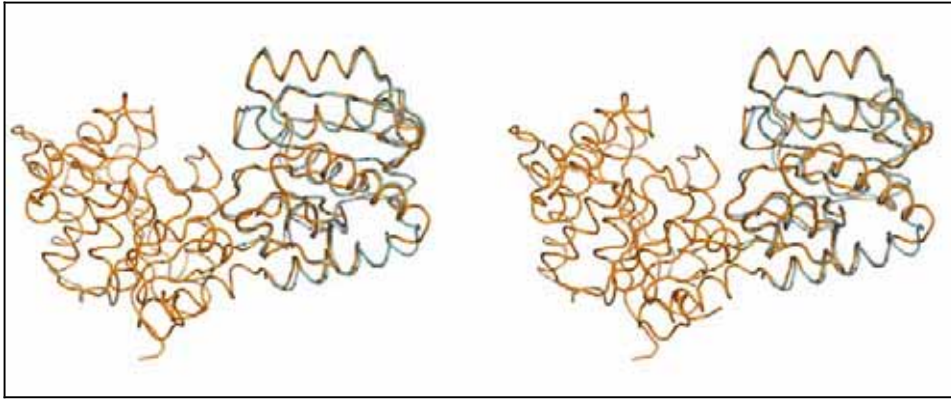
A l'hora d'analitzar el tipus d'empaquetament, s'observa que les subunitats B són les que més contactes cristal·logràfics pateixen (Fig. 4.14). Tal com hem exposat a l'apartat anterior, aquesta és la subunitat que més s'allunya d'una disposició ideal respecte a la simetria ternària amb les subunitats A i C. Els forts contactes cristal·logràfics explicarien aquest comportament, ja que estarien forçant el moviment de les subunitats B per a poder adaptar-se a un espai molt reduït, limitat per les subunitats B veïnes i per les subunitats A i C del propi trímer.



**Fig. 4.14. Contactes cristal·lins de les subunitats B.** Dibuix estereogràfic dels contactes cristal·lins entre la subunitat B d'un trímer i les subunitats B simètriques.

#### 4.1.5. Comparació del domini C-terminal aïllat i dins de la proteïna sencera:

El plegament del domini C-terminal de la PaGS aïllat o unit al domini N-terminal és gairebé idèntic, amb tres excepcions. La primera correspon a una lleugera distorsió en els llaços entre les hèlixs  $\alpha_{16}$  i  $\alpha_{17}$ . La segona es produeix al llaç entre el full  $\beta_{15}$  i l'hèlix  $\alpha_{12}$  i probablement és causada per la presència d'una molècula d'acetat al lloc d'unió a nucleòtid en el domini C-terminal aïllat (apartat 4.3.1). Aquesta observació es correspon amb la descrita per a l'estructura de l'AtGS, en què també s'observa un canvi lleuger en aquest llaç en resposta a la unió d'ADP. Per últim, es forma un enllaç disulfur entre la Cys350 i la Cys221 a l'estructura sencera, que no s'observa a l'estructura del domini C-terminal aïllat (apartat 4.1.2).



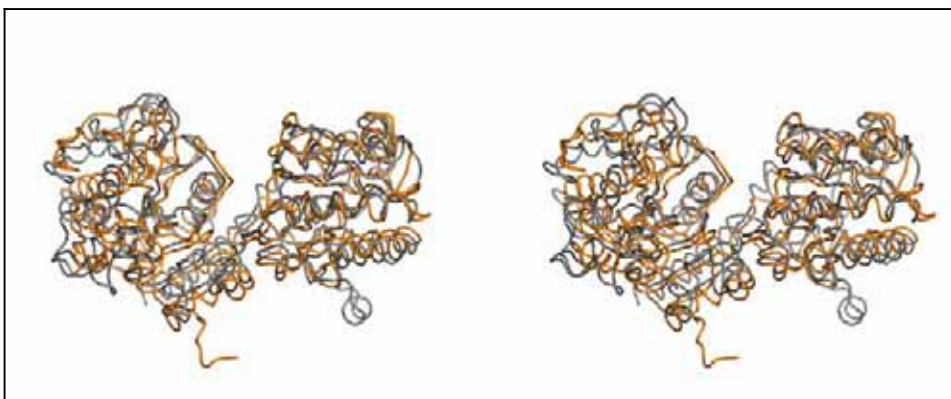
**Fig. 4.15. Comparació de les estructures del domini C-terminal de la PaGS aïllat i formant part de la proteïna sencera.** Dibuix estereogràfic de la superposició de l'estructura del domini C-terminal aïllat (blau) sobre el domini C-terminal de l'estructura de la proteïna sencera (taronja).

## 4.2. Comparació de la PaGS amb altres glicosiltransferases:

### 4.2.1. Comparació de la glicogen sintasa de *P. abyssi* amb la d'*A. tumefaciens*:

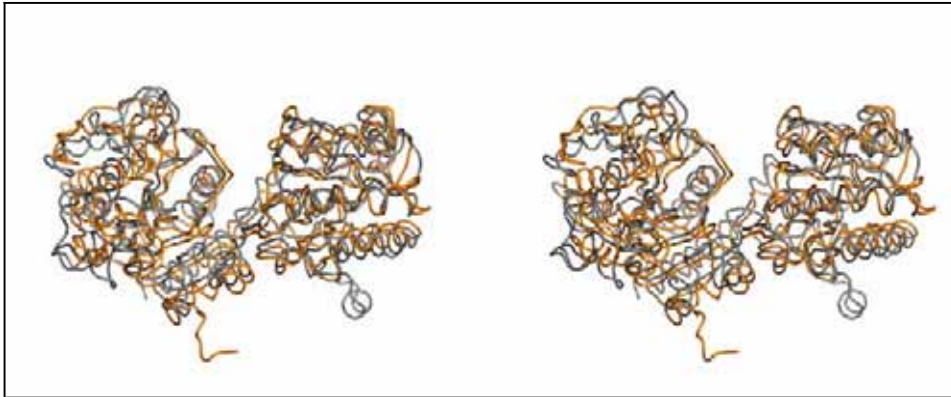
El plegament de la glicogen sintasa de *P.abyssi* és molt similar al descrit per a *A. tumefaciens* (Buschiazzi *et al.* 2004), tot i la baixa identitat de seqüència (21%). El domini N-terminal de la PaGS (residus 1-217 i 414-437) presenta una desviació quadràtica mitja (rmsd) de 2.4 Å per 219 residus equivalents a l'estructura d'AtGS (79% de residus equivalents). Pel domini C-terminal de la PaGS (residus 218-413), la desviació quadràtica mitja (rmsd) és de 1.3 Å per 183 residus estructuralment equivalents a AtGS (86% de residus equivalents).

La superposició d'ambdues estructures a partir de la superposició dels dominis C-terminal permet apreciar la similitud entre plegaments (Fig. 4.16). De totes formes, s'observa que els dominis N-terminal en aquesta figura no es superposen exactament, pel fet que l'obertura del sot catalític dels dos enzims és lleugerament diferent.



**Fig. 4.16. Superposició de la AtGS sobre la PaGS.** Dibuix estereogràfic de la superposició, utilitzant el programa SHP, de la PaGS (taronja) i la AtGS (gris). La superposició dels dominis C-terminal s'ha utilitzat per a superposar les dues molècules.

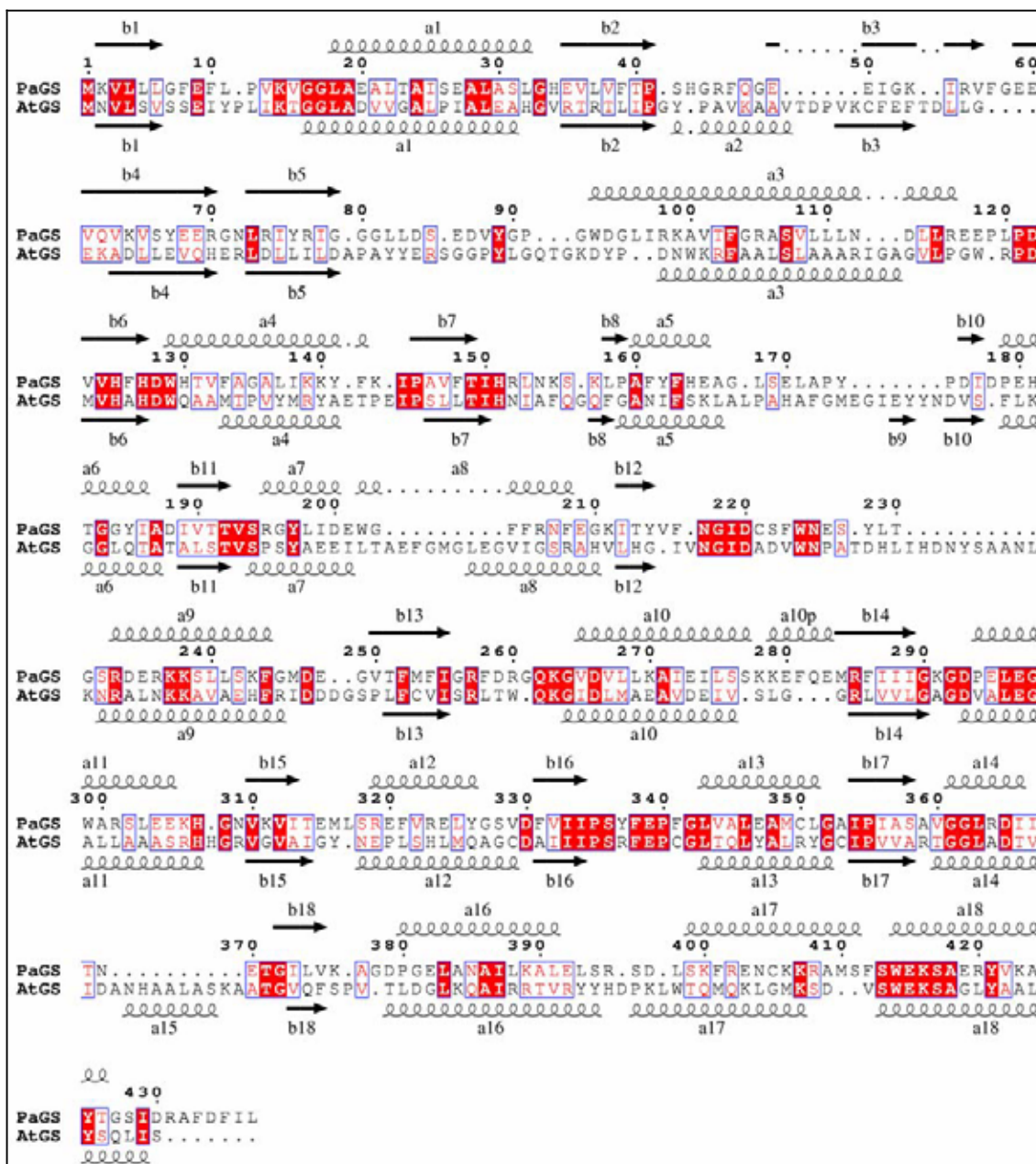
Per a poder comparar els dos enzims, era necessari eliminar la diferència causada pel diferent grau d'obertura. D'aquesta manera, vam construir una molècula quimera a partir de la superposició per separat de cadascun dels dominis N- i C-terminal de l'AtGS als seus homòlegs en PaGS. En unir els dominis N- i C-terminals superposats, vam obtenir una molècula d'AtGS quimera amb la mateixa obertura que PaGS (Fig. 4.17).



**Fig. 4.17. Molècula quimera.** Superposició de la PaGS (taronja) i la molècula quimera d'AtGS (gris) amb la mateixa obertura que la PaGS, obtinguda a partir de la superposició dels dominis N- (esquerra) i C-terminal (dreta) per separat de l'AtGS sobre els corresponents de la PaGS.

Mentre que la similitud estructural dels dominis C-terminals és molt elevada, els dominis N-terminals presenten més diferències, especialment en la zona més allunyada del solc catalític. En aquest solc les dues estructures són gairebé idèntiques en ambdós dominis, ja que s'hi troben les zones més conservades de l'enzim. En qualsevol cas, les dues diferències més visibles entre ambdues estructures són: d'una banda la protuberància de la PaGS (taronja) en el seu extrem C-terminal (situat al domini N-terminal) i que no és present a AtGS. De l'altra, l'AtGS (gris) té una protuberància diferent en forma d'hèlix  $\alpha$  en el seu domini C-terminal i que no és present en PaGS. Aquestes diferències són les principals responsables dels estats d'oligomerització dels dos enzims i seran tractades amb detall a l'apartat 4.6.

L'alineament estructural de les GS de *P.abysyi* i *A. tumefaciens* (Fig 4.18) mostra l'elevat grau de similitud. En AtGS hi ha tres insercions d'elements d'estructura secundària respecte a PaGS que corresponen a les hèlix  $\alpha_2$  i  $\alpha_{15}$  i al full  $\beta_9$ . La PaGS només presenta la inserció d'una hèlix que s'ha anomenat  $\alpha_{10p}$ .



**Fig. 4.18.** Alineament estructural de les glicogen sintases de *P.abyssi* i *A. tumefaciens*. Les regions estructuralment equivalents apareixen enquadrades. Els residus conservats es mostren de color blanc sobre fons vermell, mentre que els residus similars es mostren de color vermell sobre fons blanc. Els elements d'estructura secundària (hèlixs  $\alpha$  i fulls  $\beta$ , representats per a i b respectivament) de la PaGS i de l'AtGS es mostren a la part superior i inferior dels alineaments, respectivament. La numeració de les estructures secundàries de l'AtGS s'ha utilitzat com a referència. La numeració de la seqüència d'aminoàcids correspon a la PaGS.

#### 4.2.2. Comparació de la PaGS amb la resta de glicosiltransferases:

Per a trobar altres homòlegs estructurals de la PaGS, vam utilitzar el servidor DALI (Holm i Sander 1997). D'aquesta forma, les nostres coordenades es comparen amb totes les proteïnes amb estructura 3D coneguda que estan dipositades a la base de dades PDB. Aquesta eina és capaç de trobar semblances no detectables per comparació simple de seqüències.

A la taula 4.2 s'enumeren els homòlegs estructurals. Com era previsible, l'estructura més similar a la PaGS és l'AtGS. Després d'aquesta, les estructures més similars són la trehalosa-6-P sintasa, OtsA (Gibson *et al.* 2002), i la glicogen fosforilasa, GP (Fletterick *et al.* 1976), amb uns rmsd de 3.6 i 3.3 Å respectivament. Tots aquests enzims pertanyen a la superfamília de les GT-B i actuen amb retenció de configuració del carboni anomèric del sucre transferit.

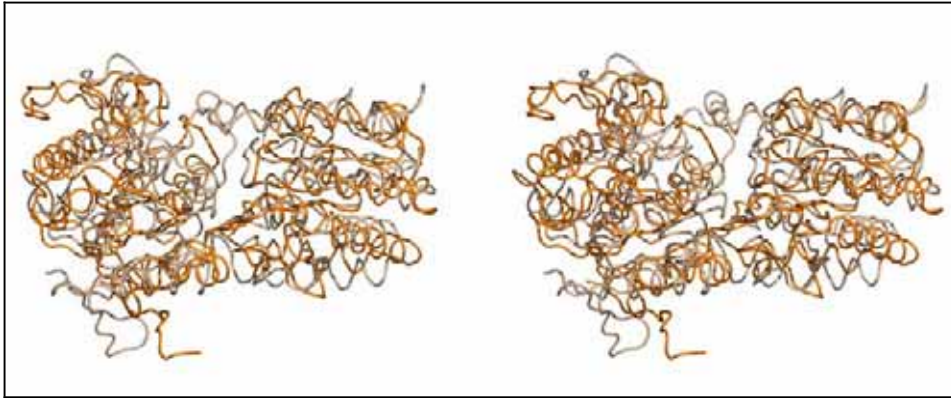
La següent estructura relacionada correspon a una epimerasa (Campbell *et al.* 2000) i a continuació apareixen tres glicosiltransferases de la superfamília GT-B, però que actuen amb inversió de configuració: T4βGT (Vrielink *et al.* 1994), MurG (Ha *et al.* 2000) i Gtfb (Mulichak *et al.* 2001).

Proteïna	PDB	Z	RMSD	LALI	LSEQ2	%IDE
AtGS	1rzu	35.8	2.9	415	477	25
Otsa	1gz5	23.5	3.6	341	452	18
GP	1a8i	20.9	3.3	418	813	14
Epi	1f6d	20.7	4.9	326	376	15
T4βGT	1c3j	14.2	4.0	269	333	13
Murg	1f0k	10.1	4.7	232	351	15
Gtfb	1iir	9.3	4.7	248	382	13

**Taula 4.2. Homòlegs estructurals de la PaGS.** Homòlegs estructurals de la PaGS calculats amb el programa DALI: glicogen sintasa d'A. tumefaciens (AtGS), trehalosa-6-P sintasa (OtsA), glicogen fosforilasa (GP), UDP-N-acetilglucosamina 2-epimerasa (Epi), T4 Fag β-glicosiltransferasa (T4βGT), UDP-N-Acetilglucosamina transferasa (MurG), UDP-glicosiltransferasa (Gtfb). Les estructures estan ordenades pel valor de Z (desviació estàndard de les similituds estructurals). També es tabulen els valors de la desviació quadràtica mitja de la posició dels àtoms Cα superposats (RMSD), nombre total de residus equivalents (LALI), nombre total de residus de la proteïna (LSEQ2) i percentatge d'identitat de seqüència entre les posicions equivalents.

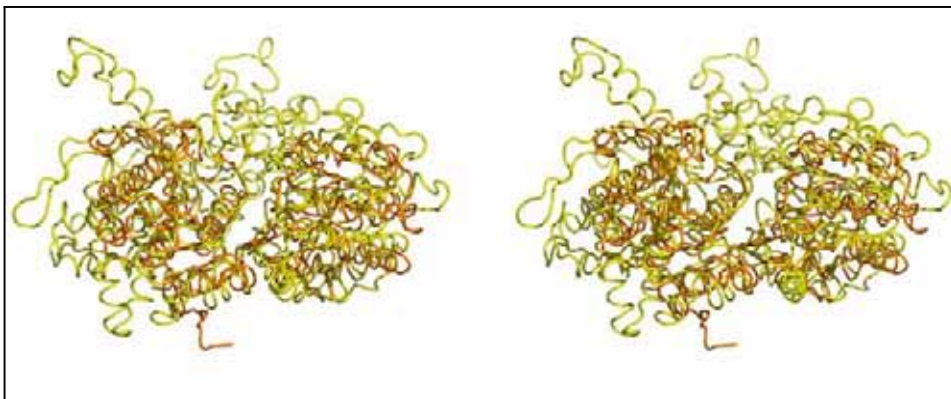
La superposició d'una subunitat de l'OtsA sobre una subunitat de la PaGS permet observar l'elevat grau de similitud en el plegament (Fig. 4.19). Les majors diferències s'observen a l'entrada del solc catalític, que en l'OtsA està protegit per dos llaços en la part superior. També són destacables les diferències en la terminació de l'última hèlix (hèlix a18 en PaGS) i que en PaGS és responsable de la seva oligomerització.





**Fig. 4.19. Superposició d'OtsA sobre PaGS.** Dibuix estereogràfic de la superposició realitzada amb el programa SHP d'un monòmer de la OtsA (rosa) sobre un monòmer de la PaGS (taronja). Els dominis N-terminals queden a l'esquerra i els C-terminals a la dreta.

Respecte a la superposició de la GP sobre la PaGS, els fulls  $\beta$  i les hèlixs  $\alpha$  centrals estan molt conservats entre ambdues estructures (Fig. 4.20). De tota manera, la GP té moltes insercions, majoritàriament situades en la superfície de la molècula. Els residus del sot catalític presenten una elevada homologia de seqüència i una idèntica disposició espacial. Aquest tema serà tractat amb més profunditat a l'apartat 4.4 i 4.5.



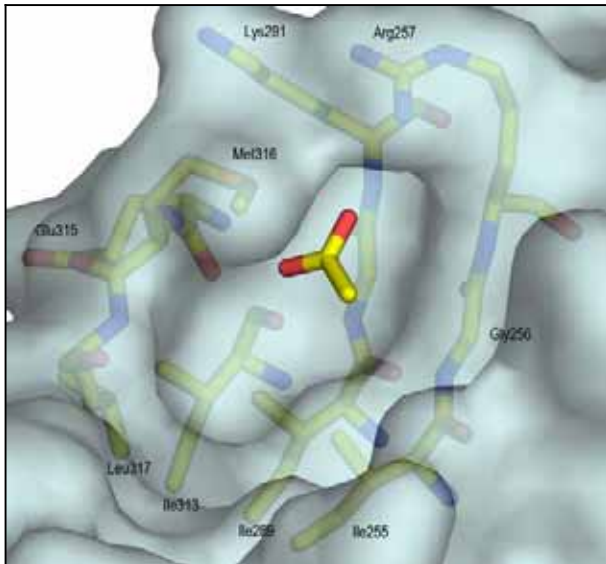
**Fig. 4.20. Superposició de GP sobre PaGS.** Dibuix estereogràfic de la superposició realitzada amb el programa SHP d'un monòmer de la GP (groc) sobre un monòmer de la PaGS (taronja). Els dominis N-terminals queden a l'esquerra i els C-terminals a la dreta.

### 4.3. Unió de substrat:

#### 4.3.1. Unió d'una molècula d'acetat:

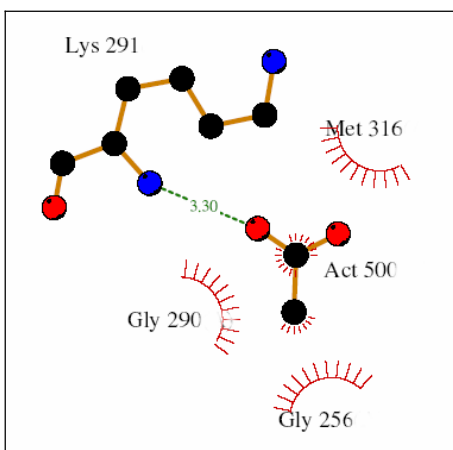
En resoldre el cristall del domini C-terminal de la PaGS, vam localitzar una molècula d'acetat procedent del tampó de cristallització en el lloc d'unió a nucleòtid. La molècula d'acetat semblava estar mimetitzant algunes de les interaccions que podrien establir els substrats naturals de l'enzim, UDPG i ADPG.

Com s'observa a la figura 4.21, la superfície molecular del domini C-terminal forma una cavitat per a poder acomodar la base nitrogenada del substrat donador de glucosil (uridina o adenina). La cavitat ve definida pel full  $\beta$ 13 (residus 255-257) i els llaços entre  $\beta$ 14 i  $\alpha$ 1 (residu 291) i entre  $\beta$ 15 i  $\alpha$ 12 (residus 315-316). Els residus implicats són Lys291, Arg257, Gly256, Ile255, Ile299, Ile 313, Leu317, Glu315 i Met316.



**Fig. 4.21.** Lloc d'unió d'una molècula d'acetat al domini C-terminal de la PaGS. Dibuix de superfície del lloc d'unió a nucleòtid en el domini C-terminal de la PaGS. Sota la superfície es detallen els residus implicats en la unió: Lys291, Arg257, Gly256, Ile255, Ile299, Ile 313, Leu317, Glu315, Met316.

Les interaccions de la molècula d'acetat amb els residus del domini C-terminal de la PaGS es detallen a la figura 4.22.



**Fig. 4.22.** Interaccions entre la molècula d'acetat i el domini C-terminal de la PaGS. Representació esquemàtica, realitzada amb el programa Ligplot, de les interaccions entre el domini C-terminal de la PaGS i una molècula d'acetat. Els punts d'acetat són representats per punts vermells i negres. Els residus Lys 291, Met 316, Gly 290, Act 500 i Gly 256 són representats amb símbols vermells. Els punts d'acetat són representats per punts vermells i negres. Els residus Lys 291, Met 316, Gly 290, Act 500 i Gly 256 són representats amb símbols vermells. Els punts d'acetat són representats per punts vermells i negres. Els residus Lys 291, Met 316, Gly 290, Act 500 i Gly 256 són representats amb símbols vermells.

La cadena lateral de la Met316 estableix interaccions hidrofòbiques en forma de *stacking* sobre la molècula d'acetat. També són hidrofòbiques les interaccions amb les cadenes principals de la Gly256 i la Gly290. El grup amida de la cadena principal de la Lys291 forma un pont d'hidrogen amb un dels oxígens de l'acetat.

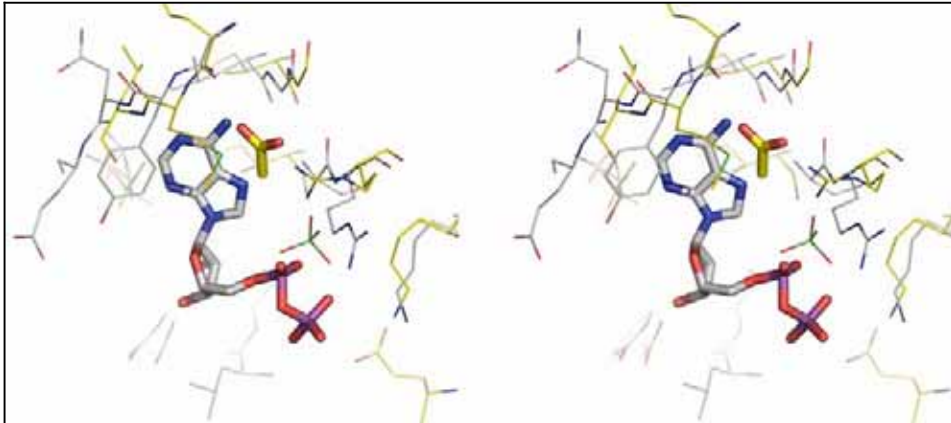
#### 4.3.2. Comparació del lloc d'unió de nucleòtid de la PaGS i l'AtGS:

Disposàvem de l'estructura del complex de l'AtGS amb ADP (Buschiazzo *et al.* 2004) i del domini C-terminal de la PaGS amb acetat unit. Atès que la majoria d'interaccions amb el substrat donador de glucosil es produeixen en aquest domini i que la unió d'acetat només s'observa en el cristall del domini C-terminal de la PaGS, ens vam limitar a l'anàlisi en aquest domini. Així doncs, vam superposar aquests dominis C-terminals amb el programa SHP (Fig. 4.23).

En l'AtGS, la cavitat pel ADP es defineix pels fulls  $\beta$ 13 (residus 297-299) i  $\beta$ 14 (326-328) i pel llaç entre  $\beta$ 15 i  $\alpha$ 12 (352-356) (Buschiazzo *et al.* 2004). Com es mostra a la figura 4.23, les interaccions impliquen les cadenes principals dels residus Ile297, Gly353 i Asn355, i les cadenes laterals dels residus Ser298, Arg299, Tyr354, Ser359, Glu376 i Thr381. L'especificitat per l'adenina sembla definida per dos únics ponts d'hidrogen febles establerts entre els àtoms N1 i N6 de l'heterocicle amb els grups amida i carbonil de l'esquelet proteic, corresponents a Gly353 i Asn355, respectivament. El residu Tyr354 també juga un paper important en l'especificitat, ja que estableix una interacció hidrofòbica forta tipus *stacking* amb l'heterocicle i adopta diferents conformacions en presència i absència de substrat (Buschiazzo *et al.* 2004).

Com ja s'ha comentat a l'apartat anterior, en el cas de PaGS aquesta interacció hidrofòbica la realitza la Met316, que també s'ha observat que pot adoptar conformacions alternatives. Tot i la mobilitat d'ambdós residus hidrofòbics, la cadena lateral de la metionina posseeix una major flexibilitat, la qual cosa li permetria adaptar-se als heterocicles d'adenina i d'uridina.

Els grups fosfats de l'ADP interaccionen amb la cadena lateral de l'Arg299 i amb la Lys304 de l'AtGS, que al seu torn forma un pont salí amb el Glu376 (glutàmic essencial que serà tractat en l'apartat 4.4.2). En el cristall del domini C-terminal de la PaGS, un sulfat del tampó de cristal·lització ocupa una posició equivalent al fosfat unit a la ribosa en el complex AtGS-ADP. Aquesta observació experimental i l'estricta conservació d'aquests residus (Arg257, Lys263, Glu339) fa pensar que aquesta interacció serà pràcticament idèntica en el cas de la PaGS.



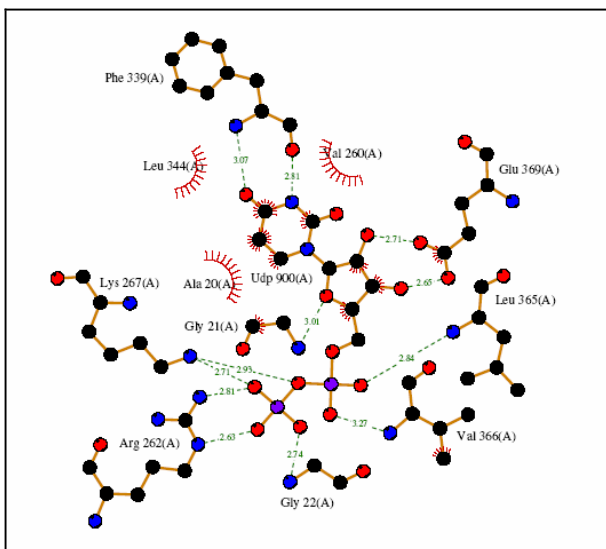
**Fig. 4.23. Comparació dels llocs d'unió a nucleòtid de l'AtGS i la PaGS.** Dibuix estereogràfic de la superposició realitzada amb el programa SHP entre els dominis C-terminal de les GS de *A.tumefaciens* i *P.abyssii*. Detall dels llocs d'unió a nucleòtid de l'AtGS (gris) unida a ADP (línies més gruixudes) i de la PaGS (groc) unida a acetat (línies més gruixudes).

#### 4.3.3. Comparació del lloc d'unió a nucleòtid de la PaGS i l'OtsA:

Atès que la PaGS pot utilitzar tant ADPG com UDPG com a substrats, volíem comparar el lloc d'unió a nucleòtid amb algun enzim relacionat que utilitzés UDPG. Vam escollir l'OtsA, ja que a més d'utilitzar aquest nucleòtid, és el segon enzim d'estructura coneguda més semblant a la PaGS (Taula 4.2). De la mateixa manera que per l'AtGS, vam superposar els dos dominis C-terminal de l'OtsA i la PaGS amb el programa SHP.

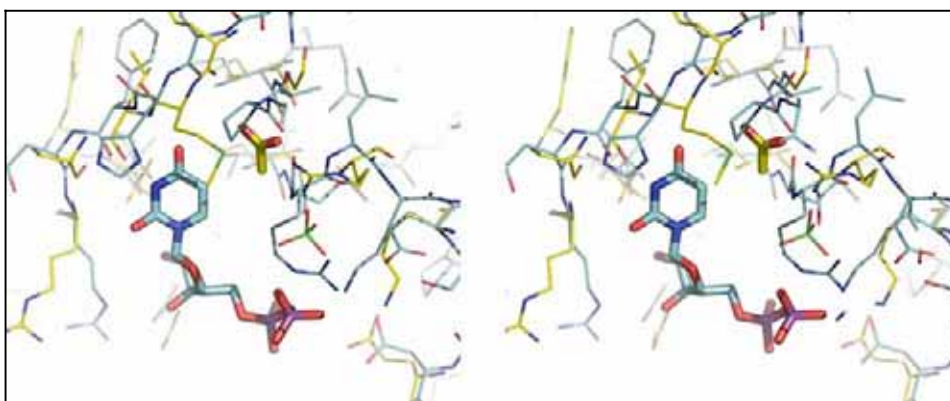
Com a PaGS, l'OtsA presenta una especificitat per nucleòtid relaxada, podent utilitzar tant UDPG com GDPG (Gibson *et al.* 2004). La resolució de l'estructura d'OtsA formant complex amb UDP, mostra que l'especificitat per l'anell d'uracil es produeix a través de les interaccions dels àtoms O4 i N3 amb els grups amida i carbonil del residu Phe339 (Fig.4.24). L'O2 de l'UDP no estableix cap interacció significativa, fet que explica la possible interacció amb GDP a través de N1 i O6.

Com s'observa a la figura 4.24 els grups hidroxil O2 i O3 de la ribosa interaccionen amb la cadena lateral del Glu 369, estructuralment equivalent al residu Glu347 en PaGS. La conservació d'aquest residu i la seva implicació en la catàlisi serà tractat a l'apartat 4.4.2. Els oxígens O1 i O2 del fosfat unit a l'anell de ribosa interaccionen amb els grups amida de la cadena principal del residu Leu365 i Val366. El segon fosfat interacciona amb les cadenes laterals de l'Arg262 i la Lys267 de forma equivalent a l'ió sulfat unit al domini C-terminal d'AtGS.



**Fig. 4.24. Interaccions entre els residus de l'OtsA i UDP.** Representació esquemàtica, realitzada amb el programa Ligplot, de les interaccions entre els residus de l'OtsA i una molècula d'UDP. Els punts d'hidrogen i les interaccions iòniques es mostren amb línia discontinua indicant les distàncies en Å. Les interaccions hidrofòbiques es marquen amb símbols vermells.

A la figura 4.25 es mostra la superposició dels dos dominis C-terminal d'OtsA i PaGS. S'observa que la His338 estableix interaccions hidrofòbiques tipus *stacking* sobre l'anell d'uracil en una posició estructuralment equivalent a la Met316 de PaGS. De tota manera, existeix un cert desplaçament cap a l'esquerra de la cavitat d'OtsA respecte a PaGS atès que el lloc d'unió de l'acetat a PaGS està ocupat per el residu Pro297 a OtsA.

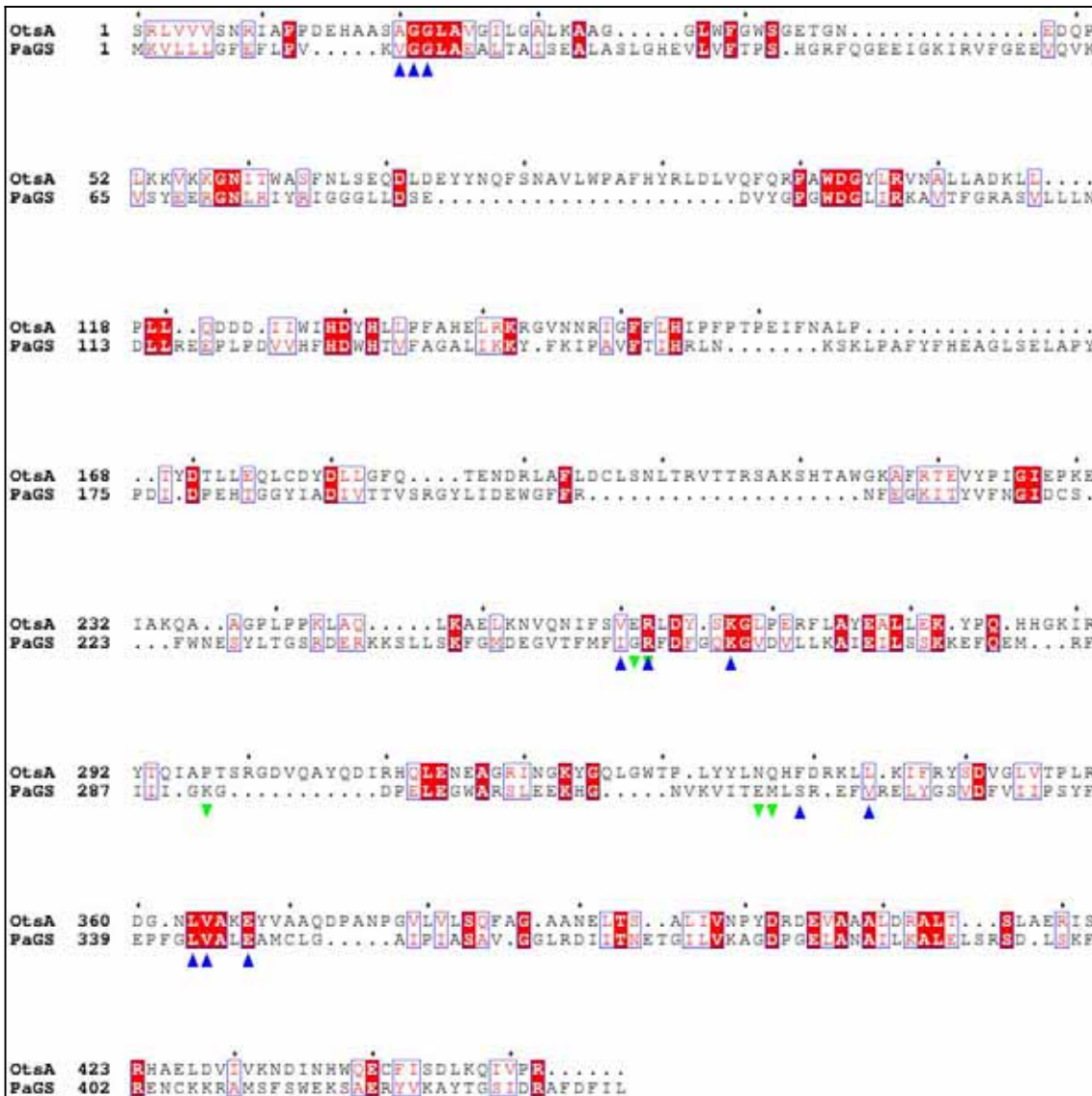


**Fig. 4.25. Comparació dels llocs d'unió a nucleòtid de l'OtsA i la PaGS.** Dibuix estereogràfic de la superposició realitzada amb el programa SHP, dels llocs d'unió a nucleòtid de l'OtsA (blau) unida a UDP (línies més gruixudes) i de la PaGS (groc) unida a acetat (línies més gruixudes).

Així doncs, tot i que els tipus d'interaccions de l'OtsA amb l'UDP seran previsiblement molt similars als establerts per PaGS, els residus que proporcionen l'especificitat per l'anell d'uracil no són els mateixos. Així, la Phe339 (que forma dos punts d'hidrogen amb l'heterocicle a través de la seva cadena principal) no està conservada en PaGS, com tampoc ho està la His338 (que forma interaccions tipus *apilament* amb l'anell). Aquesta especificitat ha estat traslladada als residus Glu315 i Met316 (Fig. 4.26).

Pel que fa referència a la interacció amb l'anell de ribosa, el Glu369 a OtsA té un residu estructuralment equivalent a PaGS (Glu347). Evidències experimentals reafirmen la importància d'aquest residu en les GS, ja que la mutació puntual del residu equivalent Glu518 a Ala en la GS muscular humana fa disminuir dràsticament l'activitat de l'enzim (Cid *et al.* 2000).

De la mateixa manera, la interacció amb els fosfats també sembla estar conservada en PaGS i resulta homòloga a AtGS. Així, els residus Leu365, Val366, Arg262 i Lys 267 a OtsA equivalen estructuralment als residus Leu343, Val344, Arg257 i Lys263 a PaGS.



**Fig. 4.26. Alineament estructural de l'Otsa i la PaGS.** Les regions estructuralment equivalents apareixen enquadrades. Els residus conservats es mostren de color blanc sobre fons vermell, mentre que els residus equivalents es mostren de color vermell sobre fons blanc. Cada deu residu en l'Otsa està marcat amb un punt a la part superior de l'alineament. El símbol ▲ marca els residus que interaccionen amb UDP a Otsa i el símbol ▼ marca els residus de PaGS que interaccionarien amb l'uridil de l'UDP en el nostre model.

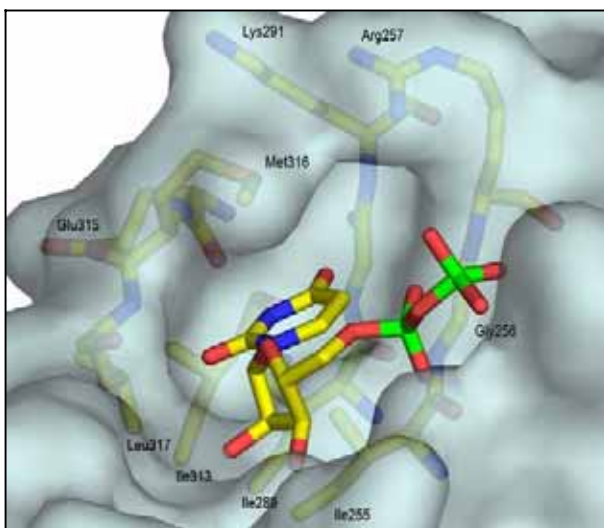
Tot i no disposar de dades experimentals del complex PaGS-UDP, l'alineament estructural amb OtsA ens permet predir que el motiu conservat Gly21-Gly22-Leu23, implicat en la unió del substrat donador, també està present en PaGS. En l'estructura d'OtsA la Gly21 forma un pont d'hidrogen amb l'oxigen endocíclic de la ribosa i la Gly22 interactua a través de la seva cadena principal amb el fosfat distal (Fig. 4.24). Donat l'elevat grau de conservació d'aquesta zona, podem inferir que aquest residus jugaran papers equivalents en PaGS.

#### 4.3.4. Model de la unió d'UDP:

Les dades experimentals disponibles de la PaGS i la comparació amb les d'AtGS i OtsA ens van permetre modelar una molècula d'UDP a la cavitat d'unio a nucleòtid del domini C-terminal de la PaGS.

Tant l'ADP a AtGS com l'UDP a OtsA estableixen ponts d'hidrogen entre l'heterocicle i els grups amida i carbonil de la cadena principal de la proteïna. La PaGS permet establir el mateix patró d'interacció d'enllaços d'hidrogen tant per a l'UDP com per a l'ADP. En el cas de l'UDP, es podrien formar tres ponts d'hidrogen. Els àtoms O2, N3 i O4 interaccionarien amb els grups amida i carbonil de les cadenes principals dels residus Leu317, Glu315 i Lys291 respectivament (Fig. 4.27). En el cas d'ADP només es podrien establir dos ponts d'hidrogen entre els nitrògens N1 i N6 amb el grup amida i el carbonil dels residus Glu315 i Lys291 respectivament.

La flexibilitat de la cadena lateral de la Met316 facilitaria les interaccions hidrofòbiques tipus *stacking* amb les dues bases. Respecte a les interaccions de la resta del nucleòtid difosfat aquestes serien molt similars a les establertes a OtsA, com ja s'ha comentat anteriorment.



**Fig. 4.27. Model d'unio d'UDP a PaGS.** Dibuix de superfície molecular del lloc d'unio a nucleòtid en el domini C-terminal de la PaGS i modelat de l'UDP. Sota la superfície es detallen els residus que estarien implicats en la unió: Lys291, Arg257, Gly256, Ile255, Ile299, Ile 313, Leu317, Glu315, Met316.

## 4.4. Centre catalític:

### 4.4.1. Modelat de la PaGS activa

Com hem introduït en l'apartat 4.2.2, les estructures resoltes d'altres proteïnes de la superfamília GT-B suggereixen l'existència d'una conformació oberta, que correspondria a un estat inactiu de la proteïna i que precisaria el seu tancament per a poder apropar els residus essencials i generar així un centre catalític actiu.

En el cas de OtsA, l'estructura global està més "oberta" quan només està unit el donador de glucosil (UDPG) que quan la proteïna està unida a la part difosfonucleòtid del donador (UDP) i al seu acceptor (G6P). En concret aquesta diferència és de 1-1.5 Å (Gibson *et al.* 2004).

Totes les estructures d'OtsA resoltes fins ara contenen el grup donador (UDPG) o una part d'aquest (UDP), i per tant, aquests models no permeten analitzar els possibles canvis en grau d'obertura o tancament de l'estructura deguts a l'absència o presència del donador.

Per a l'AtGS, la unió d'ADP manté l'estructura oberta de la proteïna i només es produeixen canvis locals en els subdominis (Buschiazzi *et al.* 2004).

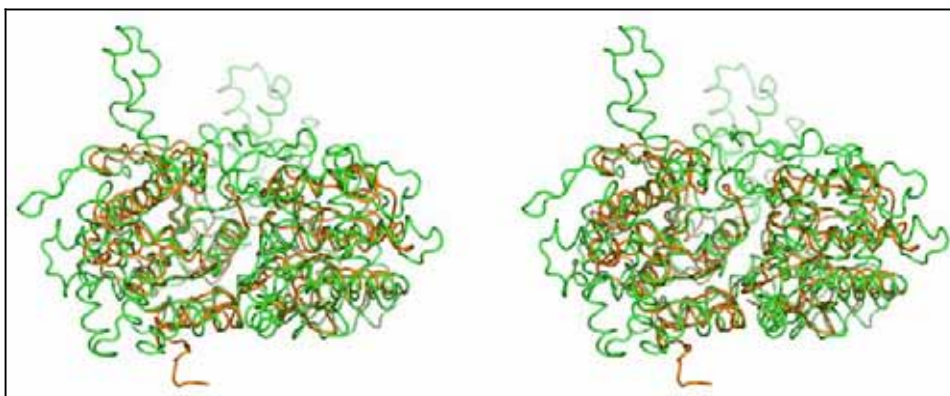
En el cas de MurG, la diferència entre l'enzim lliure i unit al seu grup donador (UDP-GlcNAc) és d'uns 2 Å. Aquest canvi conformacional es tradueix en un moviment de cos rígid del domini C-terminal, rotant uns 10° respecte a la seva posició d'enzim lliure inicial (Hu *et al.* 2003).

Per a la  $\beta$ -glicosiltransferasa del bacteriòfag T4 (BGT) (Vrieling *et al.* 1994), el canvi conformacional induït per la unió a UDPG consisteix en una rotació del domini C-terminal de 5° cap a la conformació tancada.

La PaGS lliure presenta una estructura oberta. Per a poder estudiar la implicació dels residus catalítics, ens vam plantejar modelar la PaGS sobre alguna proteïna de plegament tipus GT-B resolt en presència dels seus substrats i, per tant, amb una configuració més tancada. Vam escollir la maltodextrínfosforilasa (MalP) d'*E.coli*. *In vivo*, aquest enzim catalitza la fosforòlisi de cadenes lineals de maltodextrines per a donar lloc a glucosa 1-fosfat (Glc 1-P), encara que, *in vitro* i depenent de les concentracions de reactius, també catalitza la reacció inversa d'elongació d'una cadena d'oligosacàrid amb un residu de glucosa procedent de Glc 1-P amb l'alliberament de fosfat inorgànic. Aquesta última reacció és idèntica a la que catalitzen les glicogen sintases, amb l'única diferència que el grup alliberat en aquest darrer cas és un difosfonucleòtid. De totes les MalP resoltes, vam escollir l'obtinguda a partir del complex ternari format per l'enzim, una molècula de Glc 1-P i una cadena de maltopentaosa. En aquestes condicions, es produeix en la matriu cristal·lina la reacció inversa de la MalP, és a dir, l'elongació de la maltopentaosa amb un nou residu de glucosa i la formació d'una molècula de fosfat inorgànic (Geremia *et al.* 2002).

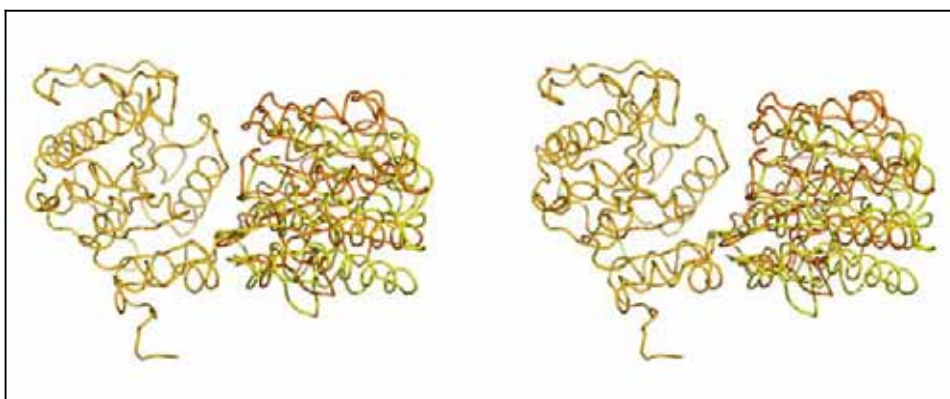


D'aquesta manera, vam construir una molècula de PaGS quimèrica amb la mateixa obertura que MalP, superposant per separat els dominis N- i C-terminal de la PaGS sobre els equivalents en MalP.



**Fig. 4.28. Molècula quimera.** Superposició de la MalP (verd) i la molècula quimera de PaGS amb la obertura de la MalP (taronja), obtinguda a partir de la superposició de cada domini N (esquerra) i C-terminal (dreta) de la PaGS sobre la MalP.

El model presentava una rotació dels dominis de PaGS d'aproximadament  $19.5^\circ$  entre les estructures oberta i tancada (Fig.4.28). El domini N-terminal de la PaGS (residus 1-217 i 414-437), presenta una desviació quadràtica mitja (rmsd) de  $1.83 \text{ \AA}$  per 207 residus equivalents (86%) a l'estructura de la MalP. Pel domini C-terminal de la PaGS (residus 218-413) la desviació quadràtica mitja (rmsd) és de  $1.68 \text{ \AA}$  per 179 residus estructuralment equivalents a MalP (91%).

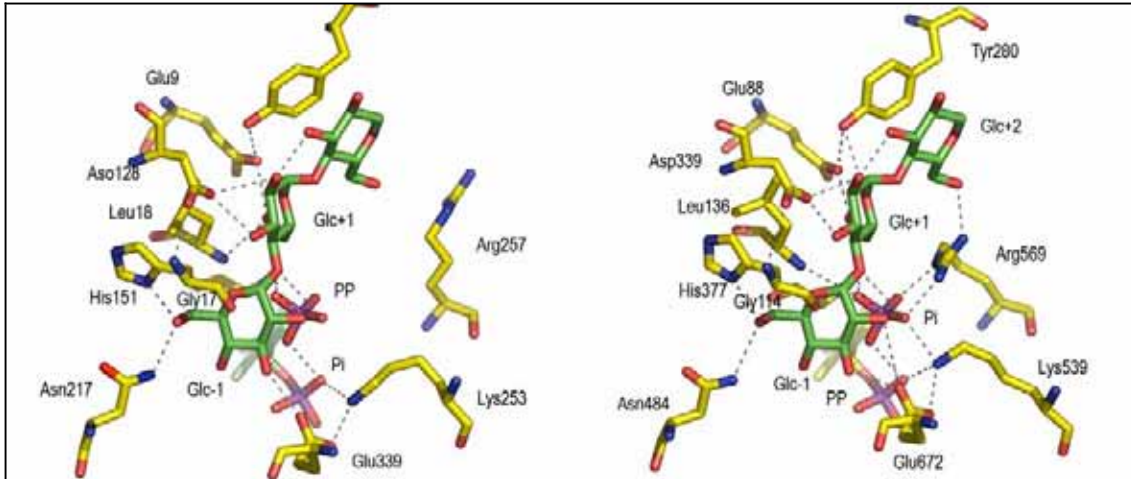


**Fig. 4.29. Superposició de la molècula quimera de PaGS (tancada) sobre la PaGS (oberta).** Superposició de la PaGS (groc) i la molècula quimera de PaGS amb l'obertura de la MalP (taronja), obtinguda a partir de la superposició dels dominis (esquerra).

#### 4.4.2. Residus catalítics:

En comparar en detall els centres catalítics dels dos enzims sorprèn l'elevada similitud entre ells, tant en la conservació dels residus com en la seva disposició i interaccions (Fig. 4.30). Aquest fet fa pensar que actuaran sota un mecanisme similar, com ja ha estat suggerit (Geremia *et al.* 2002; Buschiazzo *et al.* 2004). Les úniques

diferències destacables són les configuracions de l'Arg257 i la Lys253 a PaGS respecte a les de l'Arg569 i la Lys539 degut a l'absència de fosfat en la PaGS (apartat 4.3.2). A més de les interaccions amb les cadenes laterals, existeixen un gran nombre d'interaccions entre el sucre i els grups carbonils de les cadenes principals dels aminoàcids His151, Leu18 i Gly19, que són responsables la manca de plasticitat del sucre donador dins del solc catalític. Aquestes similituds s'estenen, encara que en menor mesura, a altres membres de la família GT-B, com és el cas de l'OtsA.

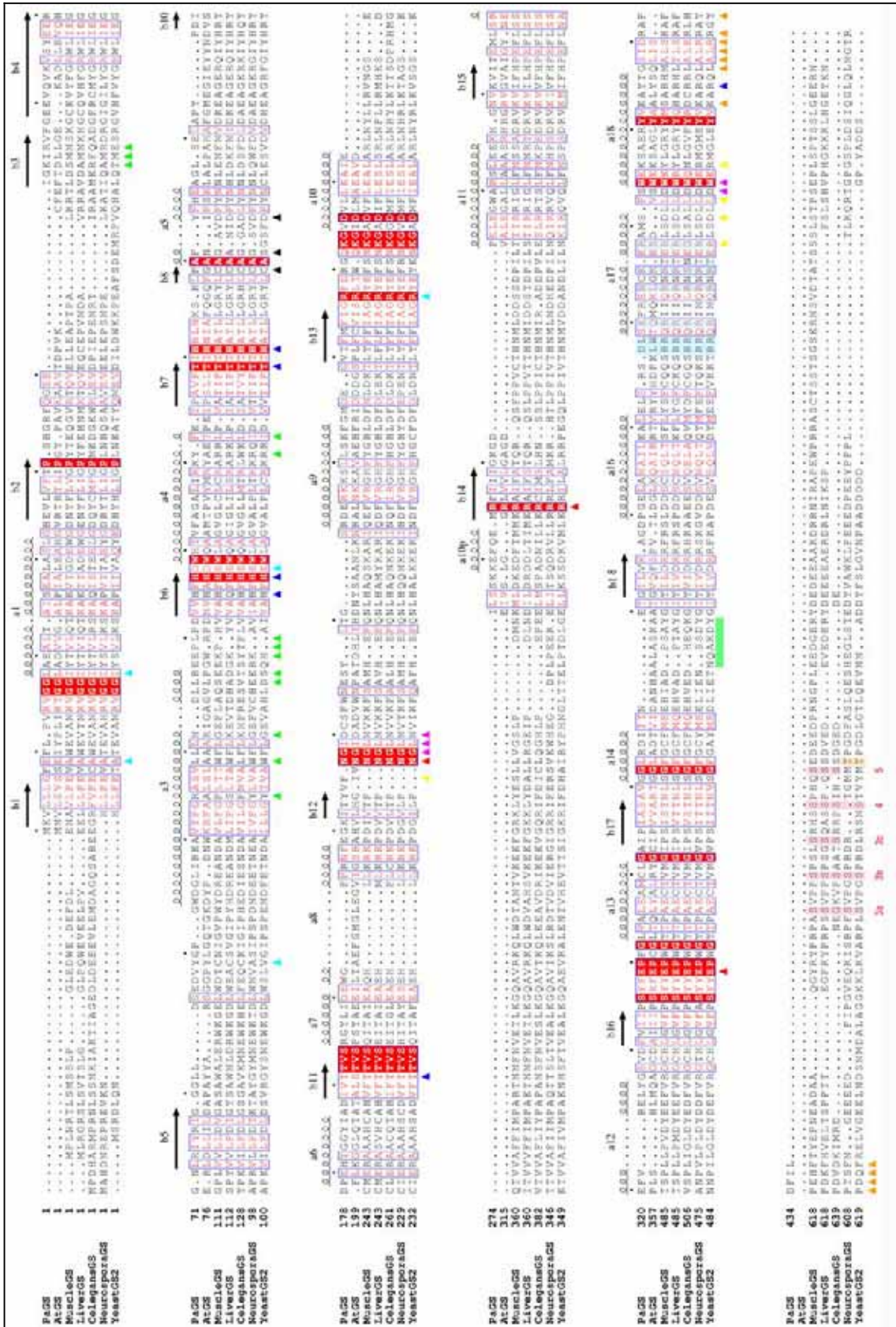


**Fig. 4.30. Centres catalítics de PaGS i MalP.** Representació comparativa dels centres catalítics de la PaGS tancada (esquerra) i de la MalP. Ambdues representacions contenen un fosfat, piridoxal fosfat i les tres primeres glucoses de la maltopentaosa originals del cristall de MalP. Per coherència amb la bibliografia publicada, la numeració de la MalP en realitat correspon a la numeració de la fosforilasa de múscul de conill.

El grup piridoxal de la MalP i el seu grup fosfat (Geremia, Campagnolo et al. 2002) ocupen una posició equivalent a la ribosa del nucleòtid difosfat i al fosfat proximal a les GS (Geremia *et al.* 2002; Buschiazzo *et al.* 2004). De la mateixa manera, el fosfat inorgànic de la MalP és equivalent al fosfat distal del nucleòtid difosfat en les sintases.

**Fig.4.31. Alineament de les glicogen sintases eucariotes al perfil estructural de la PaGS i l'AtGS.**

Les regions estructuralment equivalents apareixen enquadrades. Els residus conservats es mostren de color blanc sobre fons vermell, mentre que els residus equivalents es mostren de color vermell sobre fons blanc. Cada deu residus en PaGS està marcat amb un punt a la part superior de l'alineament. El símbol ▲ marca els residus del centre catalític, el símbol ▲ marca els residus implicats en la transferència de càrrega (*charge relay*), el símbol ▲ marca els residus que interaccionen amb la Glc +1 i la Glc +2. El símbol ▲ marca els residus que interaccionen amb la glucosa a la superfície de la molècula. El símbol ▲ marca els residus que formen la cua hidrofòbica implicada en la trimerització i el símbol ▲ representa els residus que formen la cavitat hidrofòbica de trimerització. El símbol ▲ marca els residus que interaccionen amb la molècula de dioxà i el símbol ▲ marca els residus que uneixen els dominis N- i C-terminals. La zona responsable de la dimerització de l'AtGS es marca amb ■. Les arginines sensibles a Glc 6-P en GS eucariotes s'enquadren sobre fons blau, les serines fosforilables, sobre fons rosa i la treonina fosforilable en llevat, sobre fons taronja. A la part superior es mostra l'estructura secundària de la PaGS com a referència (els fulls α i β es representen per a i b respectivament).



A continuació s'analitza amb detall la implicació dels aminoàcids que formen part del centre catalític i als quals s'ha assignat un paper funcional:

#### 4.4.2.1. Àcid glutàmic 339:

En els membres d'estructura coneguda de les GT-B que operen amb retenció de configuració es troba sempre un aminoàcid àcidic (Glu o Asp) en la posició equivalent al Glu339 de PaGS. Aquest correspon al Glu672 de MalP (utilitzem el sistema de numeració per a la GP de múscul de conill) (Geremia *et al.* 2002) o al residu Asp361 en OtsA (Gibson *et al.* 2004). Aquest residu estableix un pont d'hidrogen amb l'O3 del sucre del donador de glucosil. La mutació a Ala o Gln del residu equivalent en la GS d'*E.coli*, Glu377, disminueix la  $V_{max}$  de l'enzim més de 10.000 vegades, sense afectar la seva afinitat aparent per ADPG o glicogen (Yep *et al.* 2004). De forma similar, la mutació a Ala del residu equivalent en la GS de múscul humà (HmGS), Glu510, condueix a una pèrdua de l'activitat enzimàtica (Cid *et al.* 2000). Aquestes observacions subratllen la importància d'aquest residu en la catàlisi, encara que el seu rol exacte en el mecanisme catalític és pel moment desconegut. Les estructures resoltes amb substrats en el centre catalític mostren que, a menys que tingui lloc una reordenació conformacional important, la cadena lateral d'aquest àcid glutàmic (o aspàrtic en el cas d'OtsA) es troba massa lluny del carboni anomèric del sucre en el donador de glucosil com per a poder efectuar un atac nucleofílic sobre el mateix.

#### 4.4.2.2. Histidina 151:

La His151 en PaGS ocupa la mateixa posició que la His377 en MalP i que la His154 en Otsa. En MalP, es considera que la His377 és el primer residu de la regió mòbil que s'ha anomenat loop 380 (residus 377-384). Aquest llaç es plega sobre el centre catalític en la conformació tancada i contribueix al reconeixement i al correcte posicionament de l'oligosacàrid en els complexos MalP-oligosacàrid estudiats. A més, estableix un pont d'hidrogen entre l'àtom ND1 de la seva cadena lateral i el grup hidroxil de C6 del sucre donador.

En l'OtsA es descriu el mateix tipus d'interacció entre la cadena lateral de la His154 i l'O6 del sucre donador (UDPG). La interacció de la cadena lateral de la histidina amb l'hidroxil de C6 s'ha relacionat amb una estabilització de l'estat de transició (Gibson *et al.* 2004).

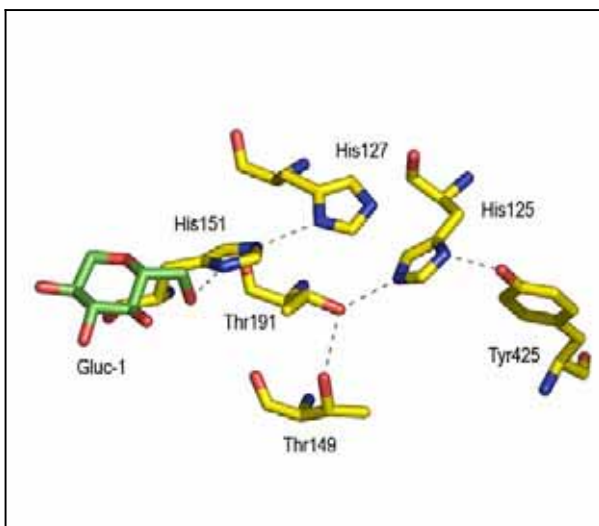
S'ha suggerit que l'oxigen del grup carbonil de la cadena principal d'aquest residu podria estar implicat en la catàlisi, actuant com a nucleòfil (Withers *et al.* 1981; Mitchell *et al.* 1996). Aquest carbonil de la cadena principal ocupa una posició equivalent a l'amida de la cadena lateral de la Gln189 en la  $\alpha$ -galactosiltransferasa LgtC de *Neisseria meningitidis* (Persson *et al.* 2001). De tota manera, el mutant Gln189Ala de LgtC manté una activitat relativament elevada i suggereix que aquest no és el veritable nucleòfil (Persson *et al.* 2001).

D'altra banda, aquest residu està conservat en tots els membres de les famílies GT-3 i GT-5 i el mutant His163Ala d'AtGS és com completament inactiu (Buschiazzi *et al.*

2004). En el mutant equivalent de la GS d'*E.coli*, l'afinitat aparent per ADPG decreix un ordre de magnitud i la  $V_{\max}$  decreix 710 vegades (Yep *et al.* 2004). Totes aquestes observacions suggereixen que la cadena lateral d'aquesta histidina té un paper important, a part del merament estructural.

L'anàlisi de l'entorn de la His151 en PaGS ens ha permès determinar un grup de residus, majoritàriament histidines, que recorda un ordenament del tipus transmissió de càrregues (*charge relay system*), el qual es troba al centre catalític de les serinproteases i que permet la fàcil transferència de protons des del substrat cap a l'exterior i a l'inrevés.

Com es mostra a la figura 4.32, el N $\Sigma$  de la His151 forma un pont d'hidrogen amb el ND de la His127. El N $\Sigma$  de la His127 forma un pont d'hidrogen amb el ND de la His125 i el N $\Sigma$  d'aquesta histidina està a una distància raonable de la Thr149 per a formar un pont d'hidrogen amb l'O6 i aquest, al seu torn, formaria un pont amb l'O6 de la Thr191. El ND de l'His125 podria estar alhora interaccionant amb l'O de la Tyr425. Aquest entorn és certament suggerent, ja que permetria fer entrar i sortir protons del centre catalític amb una relativa facilitat.



**Fig. 4.32. Entorn de la His151.** Dibuix de la interacció de la His151 de PaGS amb la glucosa-1 modelada a partir de la superposició a la MalP. Estudi de les interaccions en l'hipotètic mecanisme de transmissió de càrrega que implicaria als residus His127, His125, Thr149, Thr191, Tyr425.

Aquests residus estan molt conservats en la resta de GS. En AtGS corresponen a His163, His137, His135, Ser211, Thr161 i Tyr472. La importància d'aquestes interaccions es reflecteix en el fet que la substitució d'His167 per un residu alanina en la GS de llevat (Gsy2p), estructuralment equivalent a His127 en PaGS, rendeix un enzim que només manté l'1% de l'activitat de l'enzim salvatge (Pederson *et al.* 2000).

Mentre a OtsA es conserven la His154 i la His125, estructuralment equivalents a la His153 i a la His127 en PaGS, a les fosforilases només es troba la His377, que és isoestructural a la His154 de PaGS. Tot i les semblances entre els centres catalítics de les fosforilases i les GS, els mecanismes catalítics entre les dues famílies d'enzims podrien no ser totalment equivalents. De fet, el mecanisme proposat per a la reacció

reversa de les fosforilases, segons el qual la formació de l'enllaç glicosídic  $\alpha$ -1,4 s'iniciaria per la transferència d'un protó des del grup fosfat del cofactor piridoxal fosfat (PLP) al fosfat del substrat donador, Glc 1-P (Watson *et al.* 1999; Geremia *et al.* 2002) no és aplicable a les GS. En aquests enzims, la transferència de protó no és possible, ja que el fosfat distal al sucre de l'ADPG o UDPG, que és estructuralment equivalent al fosfat de PLP en GP, estarà probablement desprotonat. El  $pK_a$  de la primera i única ionització d'aquest fosfat distal està al voltant de 2, mentre que la segona ionització del fosfat de PLP és propera a la neutralitat.

#### **4.4.2.3. Asparagina 217:**

A més de l'His151, l'Asn217 també interacciona amb l'O6 de la glucosa -1 (Fig. 4.30). Aquesta interacció és equivalent a la que estableix l'Asn484 a les fosforilases. A OtsA no existeix aquesta equivalència amb la mateixa topologia. Tot i això, la Gln185 fa un paper equivalent des d'una posició oposada de l'estructura.

#### **4.4.2.4. Arginina 257:**

Com hem descrit anteriorment, la configuració de l'Arg257 canvia en funció de la presència o absència de fosfat (Figura 4.30). En PaGS, la cadena lateral d'aquest residu està més allunyada del centre actiu, mentre que quan està present un ió fosfat del substrat (fosfat inorgànic en el cas de les fosforilases o el fosfat distal d'UDPG/ADPG en el cas de les GS) s'hi apropa. Un comportament idèntic presenta l'Arg299 d'AtGS. Com s'observa a la figura 4.30, a més d'interaccionar amb el fosfat distal, l'Arg pot formar un pont d'hidrogen amb el C6 de la glucosa en posició +2 (Figura 4.30), tal i com s'ampliarà a l'apartat 4.5.1.

El residu equivalent en la glicogen fosforilasa, l'Arg569, s'ha suggerit que és un element important de la transició al·lostèrica entre els estats T (inactiu)-R (actiu) que contribuiria a crear un lloc d'unió d'elevada afinitat pel fosfat (Barford i Johnson 1989).

D'altra banda, les mutacions Arg569Gln i Lys574Arg a MalP produeixen una disminució dràstica de l'eficiència catalítica, sense gairebé alterar la unió a substrats (Schinzel i Drueckes 1991). Com a resultat d'aquestes observacions, s'ha postulat que aquests residus estarien implicats en la xarxa de càrregues que mantindria el balanç de protonació en els fosfats i estabilitzaria l'estat de transició.

#### **4.4.2.5. Glicina 17:**

Com s'ha introduït al capítol 4.3.3., la Gly17 és equivalent a la Gly22 en OtsA, la qual interacciona amb el fosfat distal de UDP a través de la seva cadena principal. Com s'observa a la figura 4.30, la Gly17 també és equivalent a la Gly114 en la fosforilasa. En tots els casos corresponen a la segona glicina del motiu Lys-X-Gly-Gly-Leu conservat. En tots els casos es demostra una clara interacció entre les dues glicines i l'UDP, que es relaciona amb el tancament de l'enzim al apropar-se el motiu del domini N-terminal cap al lloc d'unió del substrat al domini C-terminal (Gibson *et al.* 2002). La importància d'aquesta

regió en les GS es posa de manifest en la substitució de Gly18 per Ala (del motiu Lys-X-Gly-Gly) a la GS d'*E.coli*, estructuralment equivalent a la Gly17 de PaGS, que dona lloc a un mutant amb una constant catalítica (kcat) 3 vegades menor, mentre que el mutant Gly17Ala té una kcat tres ordres de magnitud inferior a la de l'enzim salvatge (Furukawa *et al.* 1993).

## 4.5. Unió a glicogen:

### 4.5.1. Unió de les glucoses +1 i +2:

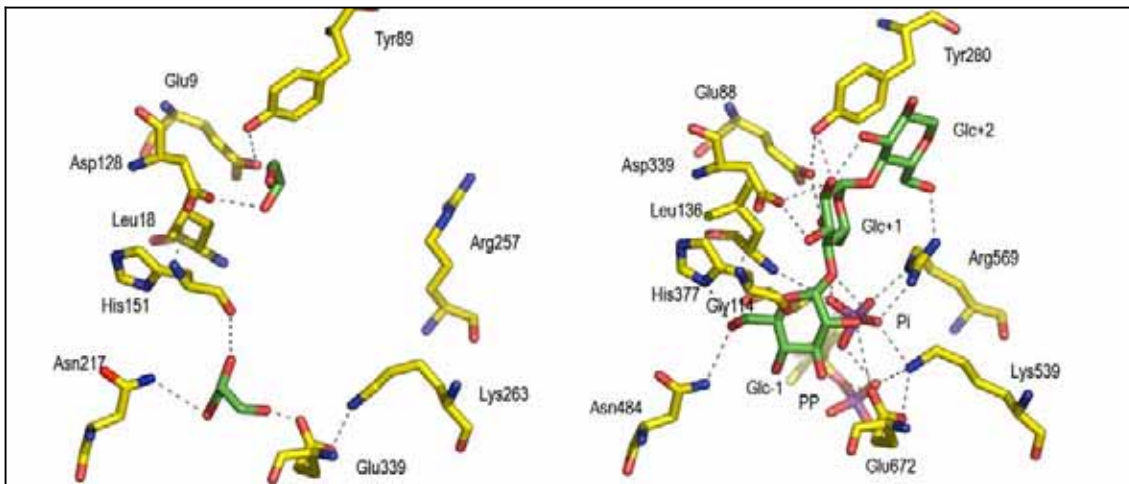
Els residus que interaccionen amb les glucoses en posició +1 i +2 de la cadena de polisacàrid estan conservats entre les GP i les GS, i en particular entre MalP i la PaGS, com s'observa a la figura 4.30.

Interaccionant amb la glucosa en posició +1, en les fosforilases es troben el Glu88 i la Leu136, que corresponen al Glu9 i a la Leu18 de PaGS. El glutàmic interacciona a través de la seva cadena lateral amb l'O6 mentre que la leucina ho fa a través del grup amino de la seva cadena principal amb l'O3.

L'Asp128 interacciona amb els oxígens O2 i O3 de la glucosa en posició +1 i, de forma equivalent a MalP (Asp339) i AtGS (Asp138), presenta uns valors pels angles  $\Phi, \Psi$  desfavorables, ocupant una zona prohibida en el diagrama de Ramachandran (Fig. 3.11, capítol 3). La importància d'aquest residu en la catàlisi es posa de manifest amb la mutació a Asn en la midó sintasa de blat SSIIb, enzim que pertany a la mateixa família de GTs que les GS bacterianes i d'arqueons (GT-5), i que dona lloc a un enzim inactiu (Nichols *et al.* 2000).

La glucosa en posició +2 forma un pont d'hidrogen entre el seu O6 i la cadena lateral de l'Arg569 en MalP. En PaGS aquest residu, l'Arg257, té una configuració diferent al no estar units els substrats. A més, com hem descrit a l'apartat 4.4.2.4, la configuració d'aquest residu canvia per a contribuir a crear un lloc d'unió amb elevada afinitat per fosfat.

La Tyr280 a les fosforilases estableix una interacció hidrofòbica tipus *stacking* amb la glucosa +2, a més de formar un pont d'hidrogen amb l'O2 de la glucosa en posició +1. Aquest residu és equivalent a Tyr89 en PaGS i Tyr96 en AtGS.



**Fig. 4.33. Molècules unides als centres catalítics de PaGS i MalP.** Representació comparativa de les molècules unides als centres catalítics de la PaGS tancada (esquerra) i de la MalP (dreta). Dos molècules de glicerol estan unides a PaGS mentre que MalP conté un fosfat, piridoxal fosfat i les tres primeres glucoses de la maltopentaosa original. Per coherència amb la bibliografia, la numeració de la MalP, en realitat, correspon a la numeració de la fosforilasa de múscul de conill.

Tot i que no disposem del complex de l'enzim amb el substrat acceptor, hem pogut validar algunes de les interaccions proposades en aquest model per homologia amb la MalP. Com es mostra a la figura 4.33, la presència de dos molècules de glicerol, que s'ha utilitzat com a agent crioprotector, en les posicions -1 i +1 mimetitzen algunes d'aquestes interaccions. Així, el glicerol a la posició -1 forma un pont d'hidrogen amb l'Asn257 (ocupant el lloc de l'O6 de la glucosa -1), un pont d'hidrogen amb el Glu339 (enlloc de l'O3 de la glucosa -1) i per últim, interaccionant amb el carbonil de la cadena principal de la His151 implicat en la catàlisi, segons s'ha descrit a l'apartat anterior. De la seva banda, el glicerol que ocupa la posició +1 interacciona amb l'Asp128, imitant la interacció que s'establiria entre aquest residu i l'O3 de la glucosa en posició +1.

De fet, la unió de molècules de glicerol també ha estat descrita per altres enzims relacionats. Per exemple, en el cas de la MalP, s'han resolt estructures en què molècules d'aquest alcohol ocupen exactament les mateixes posicions que en PaGS i només s'han pogut eliminar del centre actiu substituint el tampó de criopreservació per un que no contingui glicerol (Watson *et al.* 1999).

Totes aquestes observacions indiquen que la posició dels dos primers residus de glucosa (+1 i +2) del substrat acceptor es disposen de forma idèntica en el centre actiu de la MalP i la PaGS.

Les conformacions dels oligosacàrids resolts presenten una particularitat en l'enllaç glicosídic entre les glucoses en posicions -1 i +1. Els angles de torsió exhibeixen una geometria no estàndard que és imprescindible perquè el sucre es pugui unir a la posició -1. Existeix una rotació d'uns  $-60^\circ$  en l'angle de torsió entre O4' i C4' ( $\Psi$ ), i una rotació de  $-30^\circ$  al voltant de l'angle de torsió C1-O4' ( $\Phi$ ), que trenca el pont d'hidrogen intramolecular entre els hidrògens O2' i O3' (Watson *et al.* 1999; Geremia *et al.* 2002). Atesa l'elevada

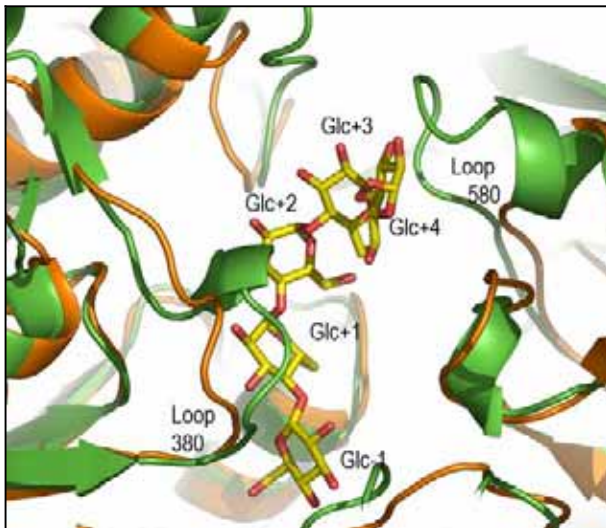


similitud entre els entorns proteics de les fosforilases i les sintases, preveiem que aquestes alteracions en la conformació glucídica també tindran lloc en la PaGS.

#### 4.5.1. Unió de les glucoses +3 i +4:

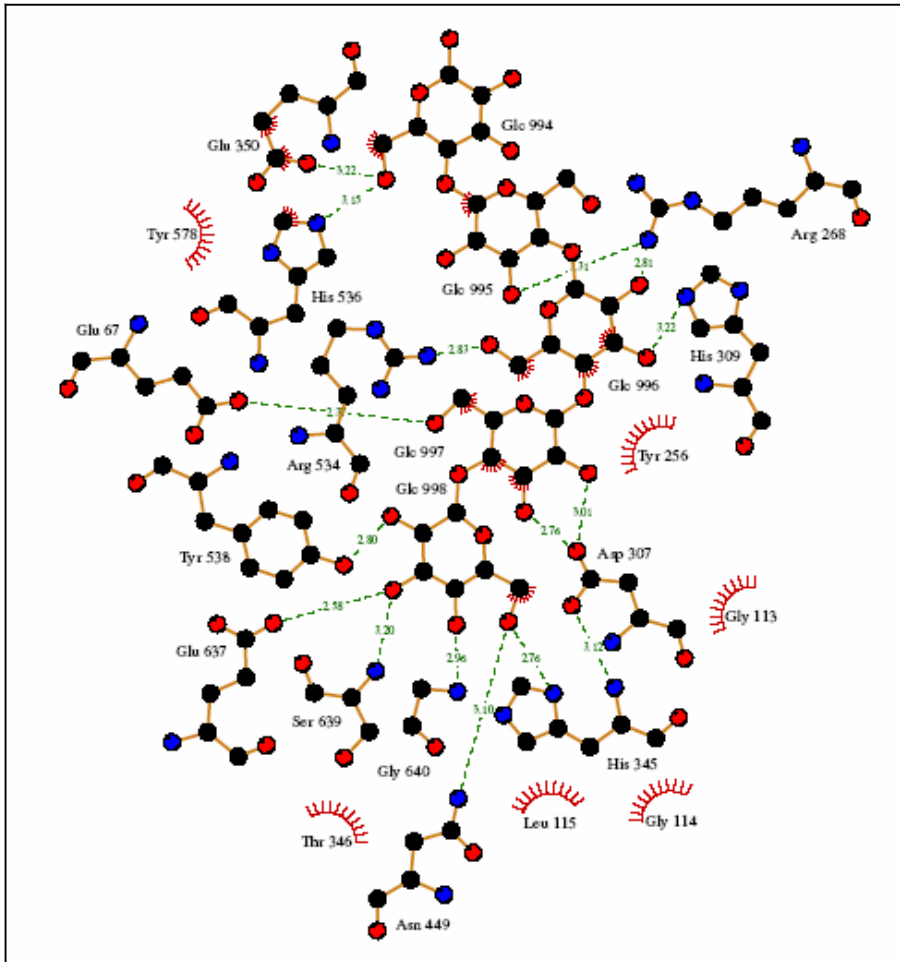
Tot i que el model de la glicogen sintasa de PaGS tancada mostra que els residus implicats en les interaccions amb les glucoses en posició +1 i +2 estan conservats en MalP, el contrari succeeix en els residus implicats en la unió a les glucoses +3 i +4. De fet, les diferències entre ambdós enzims no es limiten únicament a la conservació dels aminoàcids implicats, sinó que també impliquen una estructura secundària diferent.

Com s'observa a la figura 4.34, el loop580 del domini C-terminal (s'utilitza aquesta nomenclatura per homologia amb les fosforilases) és molt més llarg a MalP que a PaGS. De fet, en el model de la PaGS "tancada" aquesta regió ni tan sols s'apropa a la cadena d'oligosacàrid. En canvi, la MalP pot interaccionar abastament amb els dos últims residus de glucosa a través dels aminoàcids que formen aquest llaç.



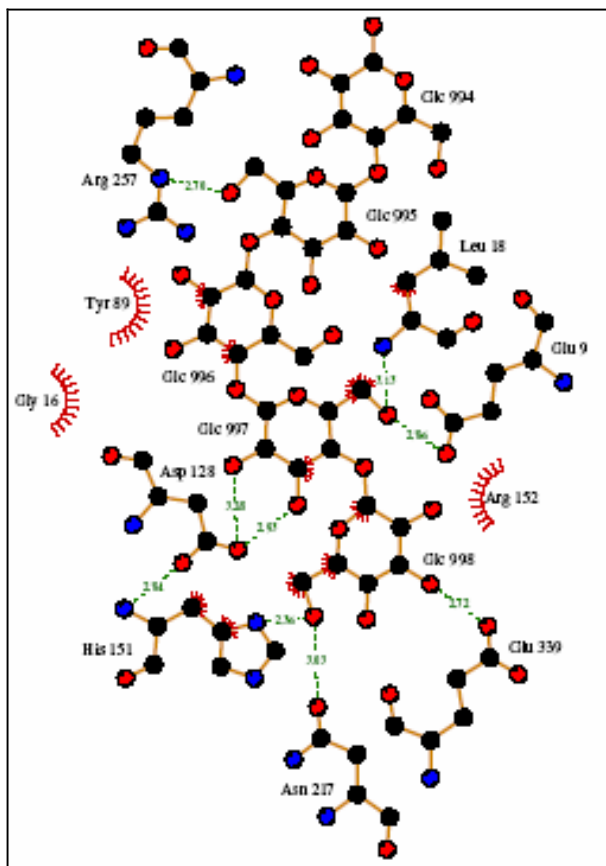
**Fig. 4.34. Túnel d'unió a oligosacàrid de PaGS i MalP.** Dibuix de la superposició del model tancat de la PaGS (taronja) a la MalP (verd) en el túnel d'unió a maltopentaosa de la MalP. Es destaquen els elements d'estructura secundària: loop380(esquerra, domini N-terminal) i loop580(dreta, domini C-terminal).

La MalP interacciona amb els residus +3 i +4, principalment a través d'una interacció hidrofòbica tipus *stacking* amb la Tyr630, segons la numeració GP (en la representació esquemàtica de la figura 4.35 apareix com a Tyr578 ja que conserva la numeració de MalP original) i per ponts d'hidrogen dels grups hidroxil del sucre amb la proteïna (Fig. 4.35).



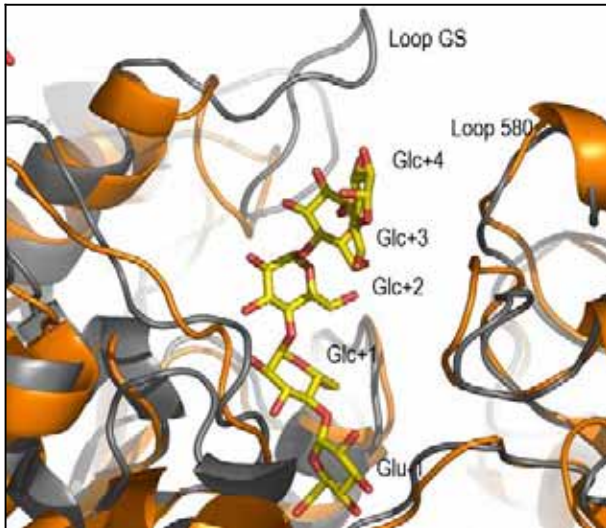
**Fig.4.35. Interaccions entre la MalP i el seu substrat maltopentaosa.** Esquema de les interaccions hidrofòbiques i ponts d'hidrogen entre la MalP i la maltopentaosa. La nomenclatura del sucre es correspon als llocs d'unió segons: 998(-1), 997(+1), 996(+2), 995(+3), 994(+4). Els ponts d'hidrogen i les interaccions iòniques es mostren amb línia discontinua indicant les distàncies en Å. Les interaccions hidrofòbiques es marquen amb símbols vermells.

El túnel catalític a PaGS estaria molt més obert i accessible que a MalP, ja que com s'observa a la figura 4.34, no disposaria delsllaços superiors (especialment el loop580 en el domini C-terminal) que encaixarien la cadena d'oligosacàrid. Aquesta observació estaria en consonància amb els resultats obtinguts per calorimetria que demostren la baixa afinitat de la PaGS per oligosacàrids lineals, com la maltopentaosa i la maltoheptaosa (Gambetti et al, resultats no publicats). Com es mostra a la figura 4.36, un model hipotètic d'unió d'oligosacàrid a PaGS per homologia a MalP suposaria la gairebé absència d'interaccions entre la proteïna i les glucoses +4 i +5.



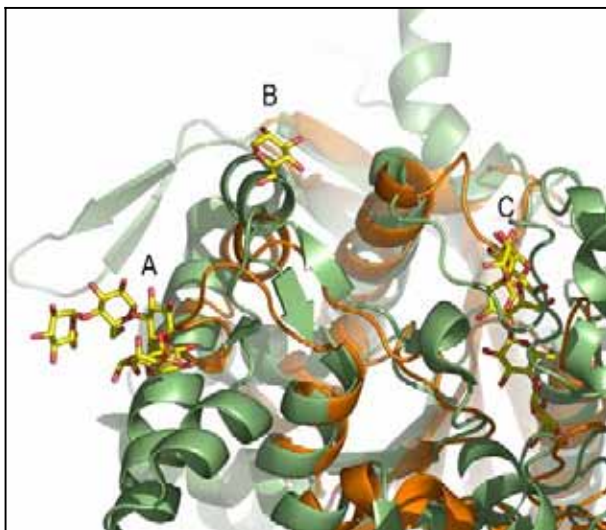
**Fig.4.36. Model de la PaGS tancada amb maltopentaosa.** Esquema del model teòric de les interaccions hidrofòbiques i ponts d'hidrogen que s'establirien entre la PaGS i la maltopentaosa per homologia amb el complex MalP-maltopentaosa (Fig.4.34). La nomenclatura del sucre es correspon als llocs d'unió segons: 998(-1), 997(+1), 996(+2), 995(+3), 994(+4). Els ponts d'hidrogen i les interaccions iòniques es mostren amb línia discontinua indicant les distàncies en Å. Les interaccions hidrofòbiques es marquen amb símbols vermells.

Una alternativa a l'absència d'aquest llaç en totes les GS analitzades podria ser el llaç que hem anomenat loopGS. Com s'observa en la superposició de models tancats de la PaGS i de la AtGS en la figura 4.37, només la glicogen sintasa d'AtGS té un llaç en el domini N-terminal molt similar al loop580 de MalP, però en una posició oposada. Tot i que no es disposa de l'estructura de la AtGS coordinada a un oligosacàrid, la posició d'aquest llaç ens permet hipotetitzar que podria tenir una funció similar d'interacció amb les glucoses +3 i +4 de l'acceptor. En qualsevol cas, aquest element estructural protegeix l'entrada del túnel catalític de l'enzim.



**Fig. 4.37. Model del túnel d'unió a oligosacàrid de PaGS i AtGS.** Dibuix de la superposició del model tancat de la PaGS (taronja) i de la AtGS (gris) al túnel d'unió a maltopentaosa de la MalP. La maltopentaosa ha estat modelada a partir del complex amb la MalP. Es destaquen els elements d'estructura secundària: loopGS (esquerra, domini N-terminal) i loop580 (dreta, domini C-terminal).

La MalP, al contrari que la GP, no mostra afinitat per glicogen. La interacció de GP amb el glicogen es defineix pels anomenats llocs d'emmagatzematge de glicogen (major i menor) (Johnson *et al.* 1988; Goldsmith *et al.* 1989; Johnson *et al.* 1990). La comparació de seqüències mostra que MalP només conserva dos (Asn407 i Lys437) dels cinc residus que a GP (Tyr404, Asn407, Gln408, Glu433, Lys437) formen els contactes essencials amb les glucoses de l'oligosacàrid en el lloc d'emmagatzematge. S'ha postulat que aquest lloc participa en el correcte posicionament de la cadena de glicogen en el centre catalític a les GP eucariotes.

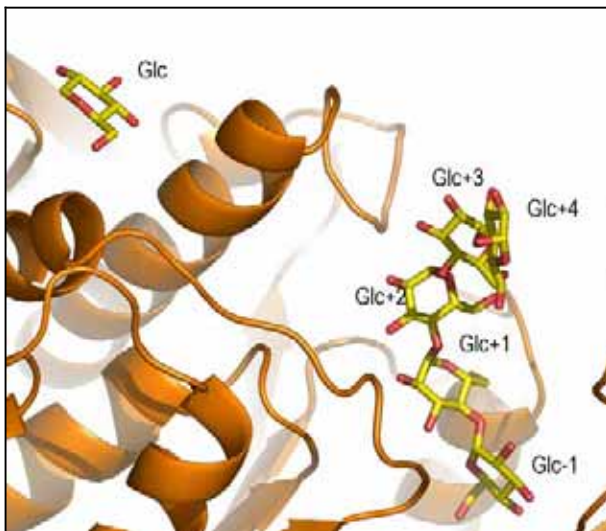


**Fig. 4.38. Superposició de la GP i la PaGS mostrant el lloc d'emmagatzematge del glicogen.** Dibuix de la superposició de la GP(verd) unida a maltopentaosa en el lloc d'emmagatzematge del glicogen (A) sobre el model tancat de la PaGS (taronja) amb una molècula de maltopentaosa modelada en el centre actiu a partir de l'estructura del complex del polisacàrid amb la MalP(C) i una glucosa localitzada en la superfície de l'estructura (B).

L'absència d'un lloc d'unió a glicogen en la MalP i, per tant, la impossibilitat d'un correcte posicionament d'un oligosacàrid ramificat podria explicar la preferència de MalP per oligosacàrids lineals (Fig 4.38) (Watson *et al.* 1999).

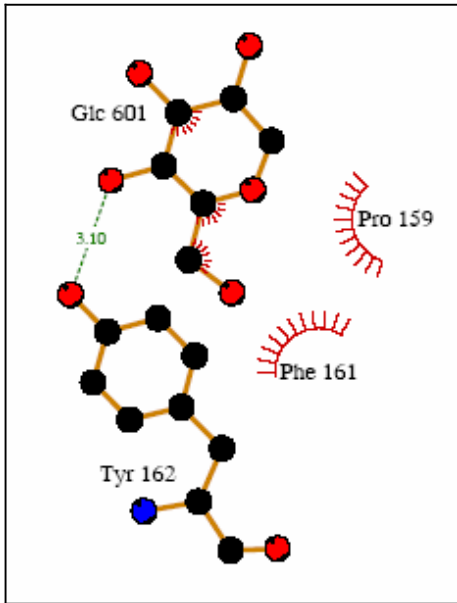
L'alineament de seqüències mostra que les GS no conserven els llocs d'unió a glicogen presents a les GP. De tota manera, existeixen múltiples evidències que les GS (almenys les eucariotes) també estan unides a glicogen en diferents condicions fisiològiques (Ferrer *et al.* 1997; Garcia-Rocha *et al.* 2001; Prats *et al.* 2005). Encara que els residus implicats en la interacció amb la molècula de glicogen seran diferents.

Al resoldre l'estructura de la PaGS, vam localitzar una molècula de glucosa en dues de les subunitats (A i C) del trímer, probablement procedent de l'UDP-glucosa que havíem afegit al tampó de cristallització. Aquesta glucosa no estava dins del centre catalític sinó que es trobava en la superfície de la proteïna, tot i que a poca distància del solc catalític (Fig. 4.39).



**Fig. 4.39. Localització d'una glucosa en les proximitats del centre catalític de la PaGS.** Model tancat de la PaGS (taronja) amb una molècula de maltopentosa modelada en el centre actiu a partir de l'estructura del complex de la MalP amb aquest polisacàrid i una glucosa localitzada en la superfície de l'estructura de la GS.

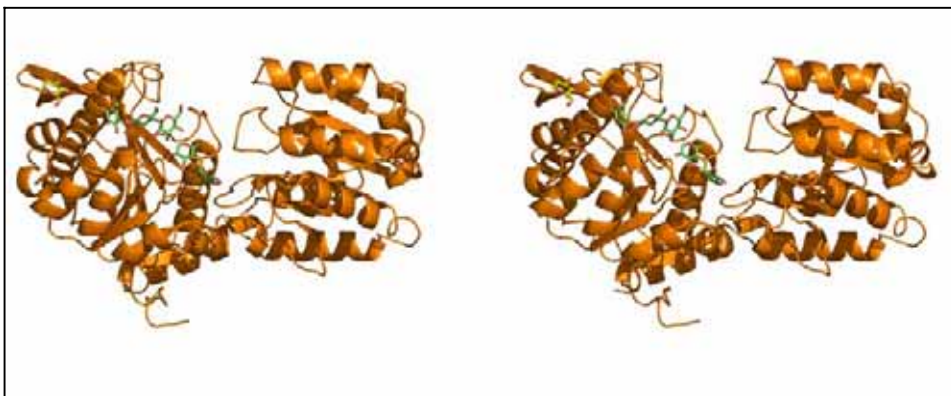
Les interaccions entre el residu de glucosa i la PaGS es detallen en la figura 4.39. La Tyr162 estableix un pont d'hidrogen amb l'O5 de la glucosa i els residus Pro159 i Phe161 estableixen interaccions hidrofòbiques de tipus *stacking*.



**Fig. 4.39. Interaccions entre el residu de glucosa i la PaGS.** Dibuix esquemàtic de les interaccions tipus ponts d'hidrogen, les interaccions iòniques (línia discontinua indicant les distàncies en Å) i les interaccions hidrofòbiques (símbols vermells).

La distància entre la glucosa i el centre catalític (glucosa en posició -1) és d'uns 30Å i aquesta és la mida aproximada d'una cadena de cinc molècules de glucosa. Si analitzem l'entorn de la PaGS al voltant de la glucosa i del centre catalític ens adonem que existeix un canal que comunica ambdues zones.

Considerant les diferències entre la PaGS i la MalP en la zona d'interacció de les dos últimes glucoses i la presència d'aquest canal, es podria hipotetitzar que aquest fos, en realitat, el lloc d'unió de la cadena emergent de glicogen (Fig.4.40). Tot i així, no afavorim aquesta hipòtesi, ja que el modelat d'aquesta cadena d'oligosacàrid en el nou canal implicaria un canvi molt important de les interaccions que estableix la glucosa +2 amb l'enzim en relació a la MalP, i això és poc probable pel descrit anteriorment.

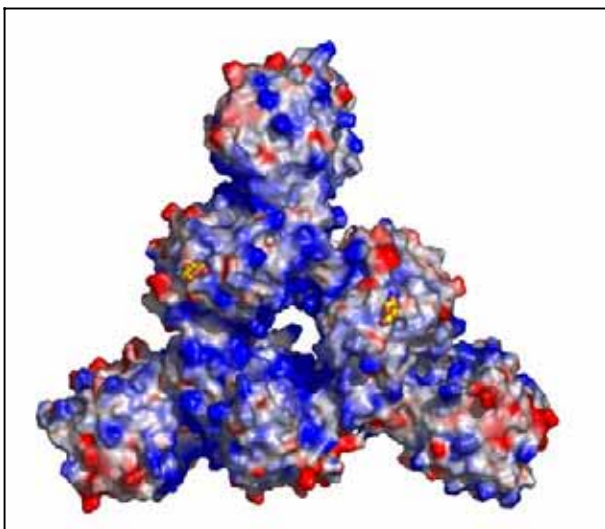


**Fig.4.40. Canal hipotètic d'unió a glicogen.** Dibuix estereogràfic d'una molècula de maltopentaosa modelada sobre l'estructura de PaGS que definiria la unió d'aquest sucre al hipotètic canal.

En concret, la glucosa en posició +2 interaccionaria amb la Tyr89 i la His130, la posicionada en +3 interaccionaria amb Trp129 i Leu29. La glucosa en posició +4 podria interaccionar amb el Trp93. Aquest residu podria resultar clau, ja que un petit canvi en la seva configuració permetria deixar entrar la molècula en el canal i protegir-la amb un nou canvi a la posició inicial. Amb aquest residu també interaccionarien la Leu158 i la Pro159. Com ja hem descrit anteriorment, la glucosa que localitzem en el cristall de PaGS ocuparia la posició +5 i podria interaccionar amb la Tyr162, Phe161 i Pro159 (Fig 4.41). D'altra banda, la baixa conservació dels residus que conformen el canal en la resta de GS faria pensar que aquestes unirien el seu substrat acceptor de mode diferent a la PaGS o, més probablement, que aquest no és el lloc d'unió de glicogen real en aquesta família d'enzims.

Una hipòtesi alternativa a l'anterior, a propòsit del significat biològic de la molècula de glucosa en la superfície de la PaGS, relacionaria aquesta amb un hipotètic lloc d'emmagatzematge del glicogen en les GS. Com s'observa a la figura 4.38, la glucosa ocuparia un lloc intermedi entre el centre catalític i el lloc d'emmagatzematge de glicogen en GP. Per tant, molt probablement senyala una zona d'elevada afinitat per sucre que ajudaria a posicionar la molècula de glicogen en el solc catalític.

Aquesta anàlisi ens permet definir una cara G en l'estructura de la PaGS trimèrica, que seria la que s'uniria a les partícules de glicogen *in vivo*, augmentant la concentració efectiva del polisacàrid en els extrems ramificats (Fig. 4.41). Com en el cas de les fosforilases, el domini encarregat d'unir glicogen seria el domini N-terminal. Segons aquesta hipòtesi, els tres dominis N-terminal quedarien fixats per les interaccions d'oligomerització i per la unió a glicogen. El domini C-terminal, principal responsable de la unió del substrat donador de glucosil, podria apropar-se o allunyar-se del domini N-terminal, tancant o obrint el solc catalític, respectivament.

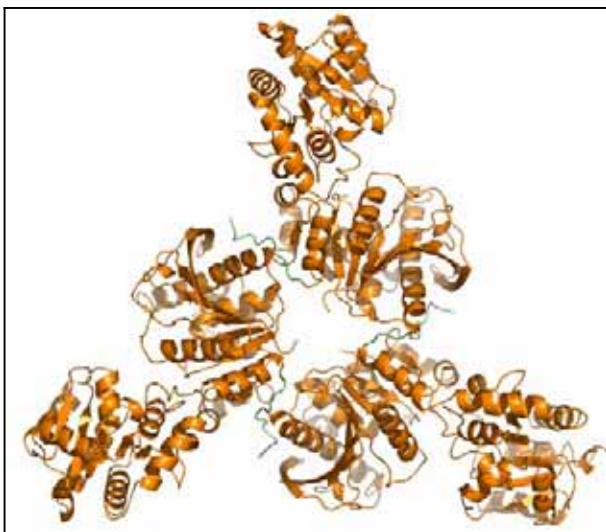


**Fig.4.41. Dues molècules de glucosa unides a la superfície de la PaGS.** Dibuix de la superfície de la cara G de la PaGS a la que estan unides dues molècules de glucosa (groc). El color blau indica les zones positives, el vermell les negatives i el gris les neutres.

## 4.6. Oligomerització:

### 4.6.1. Trimerització de la PaGS:

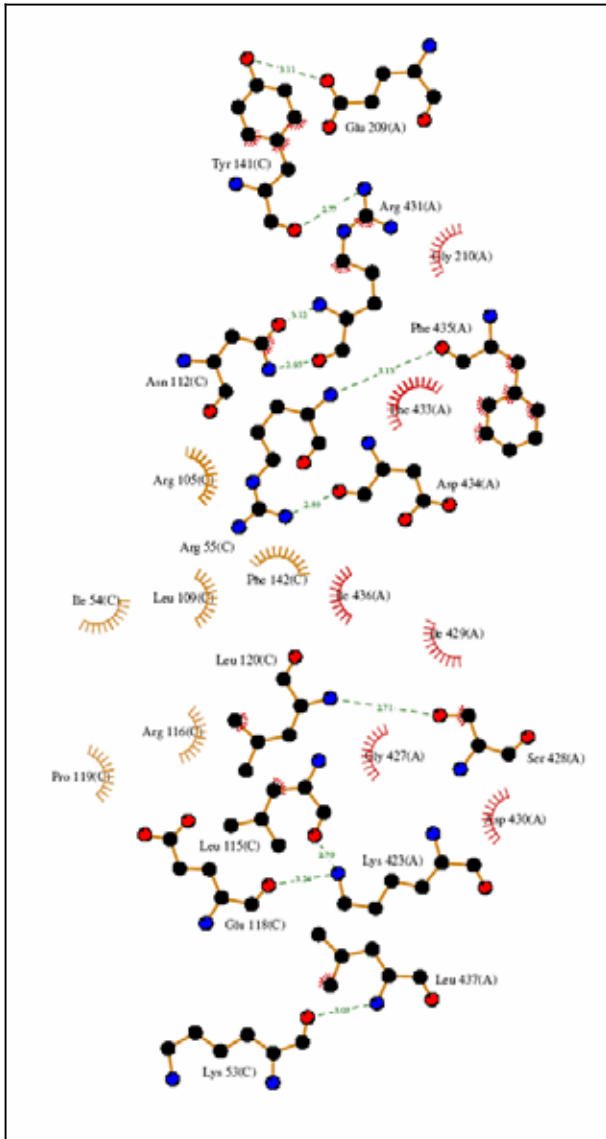
La GS de *P.abysyi*, a diferència de la majoria de GT de les que es coneix l'estructura i que oligomeritzen en forma de dímers o tetràmers, s'organitza en forma de trímers. Per a oligomeritzar, la PaGS utilitza els seus deu últims residus de naturalesa hidrofòbica (Gly427-Leu437) per a projectar-los cap a una cavitat hidrofòbica del domini N-terminal de la subunitat veïna formada pel full  $\beta$  a3 i l'hèlix  $\alpha$ 3, essencialment amb els residus Lys53-Arg55 i Leu109-Leu120 (Fig. 4.42).



**Fig.4.42. trímer de PaGS.** Dibuix de la PaGS trimèrica (taronja) en què s'han pintat de color verd les cues de 10 residus que participen en l'oligomerització.

A la figura 4.43, es mostren les interaccions que s'estableixen en la superfície d'oligomerització entre la cua hidrofòbica de la subunitat A i la cavitat hidrofòbica de la subunitat C. Aquestes interaccions impliquen fonamentalment les cadenes laterals dels residus Phe435, Phe433, Ile436, Ile429, Asp430, i les cadenes principals de la Gly210 i la Gly427 de la subunitat A, d'una banda, i les cadenes laterals dels residus Arg105, Phe142, Leu109, Ile54, Arg116, Pro119, de l'altra. Els deu ponts d'hidrogen que es detecten entre subunitats impliquen majoritàriament els àtoms de les cadenes principals i algunes cadenes laterals. L'O de la cadena lateral de la Tyr141(C) forma un pont d'hidrogen amb la cadena lateral del Glu209(A), mentre que l'O amídic interacciona amb el N de la cadena lateral de l'Arg431(A), que al seu torn interacciona amb l'O i el N de la cadena lateral de l'Asn112(C) a través dels N i O de l'enllaç peptídic. L'Arg55(C) interacciona amb l'O de la cadena principal de la Phe435 a través del seu N amídic i amb el carbonil de la cadena principal de l'Asp434 a través del N de la seva cadena lateral. El mateix succeeix entre el carbonil de la Ser428(A) i el N amida de la Leu120(C). La Lys423(A) interacciona amb els O de les cadenes principals de la Leu i el Glu. La Leu437(A) forma un pont d'hidrogen entre el seu N amídic i l'O de la cadena principal de la Lys53.

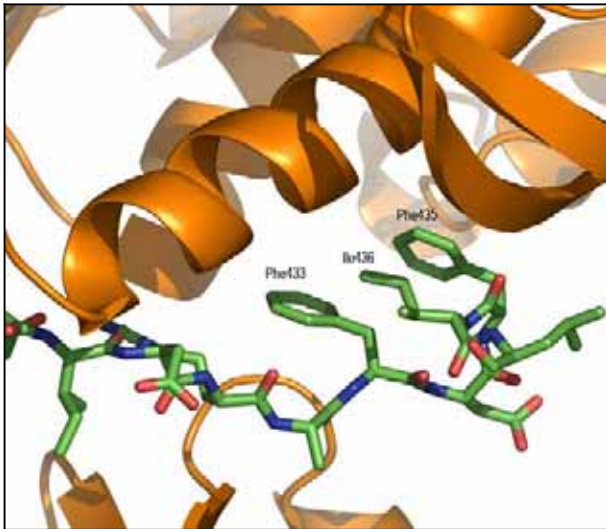




**Fig. 4.43. Interaccions entre els residus implicats en l'oligomerització.** Esquema de les interaccions hidrofòbiques i ponts d'hidrogen entre les subunitats A i C del trimer de PaGS. Els ponts d'hidrogen i les interaccions iòniques es mostren amb línia discontinua i les distàncies s'indiquen en Å. Les interaccions hidrofòbiques es marquen amb símbols vermells.

Malgrat l'existència de les interaccions de tipus polar esmentades, són les de caràcter hidrofòbic les que dominen la formació del trimer. Aquest tipus d'interacció és molt freqüent en proteïnes oligomèriques d'organismes termòfils, ja que la força de l'enllaç augmenta amb la temperatura. El motiu FDFIL (Phe433, Asp434, Phe435, Ile436, Leu437) en la cua hidrofòbica és essencial per a la formació d'aquest trimer. Com es mostra a la figura 4.44, la disposició dels dos anells aromàtics de les dues fenilalanines en forma de forca fa que s'ajustin perfectament a la cavitat hidrofòbica complementària de la subunitat veïna, evitant d'aquesta manera la seva exposició al medi aquós.

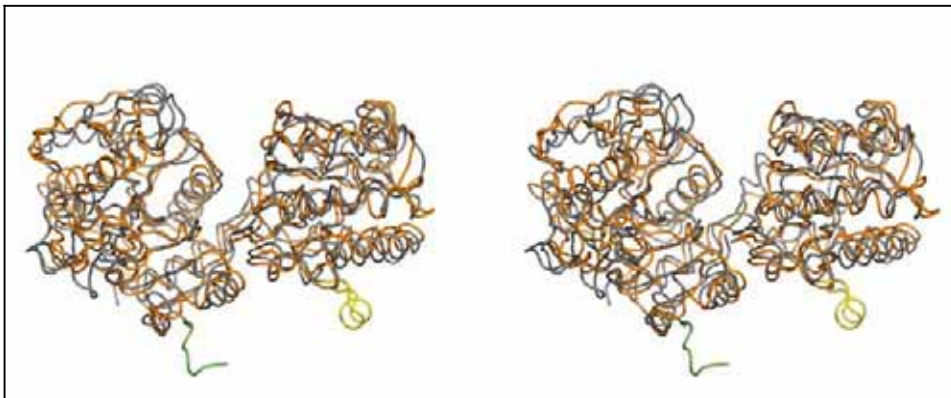
Tot i que els residus implicats en l'oligomerització són pocs en relació al nombre total d'aminoàcids de l'estructura, cada superfície de contacte és de 1140 Å<sup>2</sup>, i per tant, la superfície total de contacte en el trimer és de 3420 Å<sup>2</sup>.



**Fig.4.44. Interacció del motiu FDFIL en la superfície d'oligomerització.** Dibuix de la disposició del motiu FDFIL (en verd) en interaccionar amb la cavitat hidrofòbica de la molècula veïna (taronja)

#### 4.6.2. Comparació entre els estats d'oligomerització de la PaGS i l'AtGS:

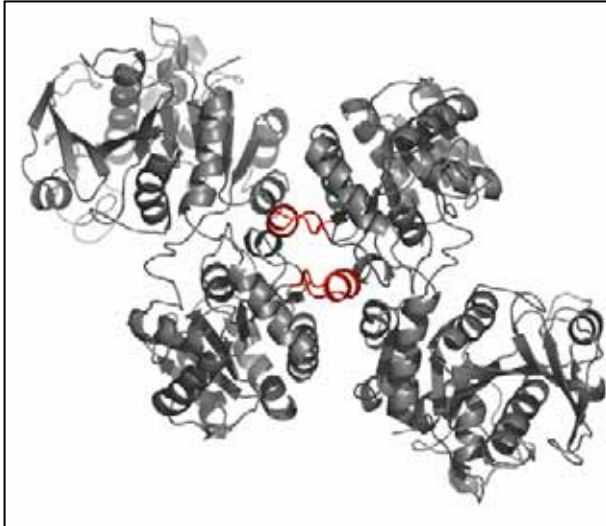
Com s'ha introduït a l'apartat 4.1.5., la superposició d'un monòmer de PaGS a un d'AtGS mostra les poques diferències estructurals entre elles. Precisament són aquestes diferències puntuals les principals responsables dels diferents estat d'oligomerització, dímers en el cas de AtGS i trímers en el de PaGS (Fig. 4.45).



**Fig.4.45. Comparació dels motius d'oligomerització.** Dibuix estereogràfic de la superposició d'un monòmer de PaGS (taronja) amb el motiu de trimerització en verd sobre un monòmer d'AtGS (gris) amb l'hèlix de dimerització de color groc.

L'alineament estructural entre ambdues proteïnes (Fig.4.18) mostra que l'AtGS conté una hèlix extra ( $\alpha 15$ ) en el domini C-terminal entre l'hèlix  $\alpha 14$  i el full  $\beta 18$ , formada pels residus 407-416, que és responsable de la seva dimerització. En aquest cas, i a diferència del que succeeix amb la PaGS, les interaccions entre subunitats només impliquen els dominis C-terminals dels dos monòmers, mentre que els dominis N-terminals queden lliures per a permetre els moviments d'obertura i tancament del solc catalític imprescindible per a la catàlisi (Fig. 4.46).

D'una altra banda, l'AtGS no conté la cua hidrofòbica de la PaGS (Gly427-Leu437), la qual manté units els dominis N-terminals en el trímer de PaGS i també permet la catàlisi a partir del moviment relatiu dels dominis C-terminals lliures.



**Fig.4.47. Dímer de AtGS.** Dibuix del dímer de l'AtGS (gris) en què les hèlix responsables d'aquesta dimerització apareixen de color vermell.

L'alineament de les seqüències de les sintases bacterianes mostra que totes elles contenen la inserció responsable de la dimerització en AtGS i no contenen la cua hidrofòbica present a PaGS. Aquesta observació suggereix que molt probablement l'organització molecular de les sintases bacterianes és dimèrica. Tanmateix, les GS d'arqueons no contenen la inserció responsable de la dimerització, mentre que posseeixen la cua hidrofòbica, suggerint que molt probablement formaran trímers.

#### 4.6.3. Predicció de l'estat d'oligomerització de les GS eucariotes:

Per a predir l'estat d'oligomerització de les GS eucariotes de la família GT-3, de les quals no es coneix l'estructura tridimensional, es van alinear amb les seqüències d'AtGS i PaGS. Tot i que la majoria dels sintases eucariotes mantenen part de l'hèlix de dimerització de l'AtGS ( $\alpha 15$ ), la seqüència d'aminoàcids no està conservada. D'altra banda, les sintases eucariotes contenen una regió equivalent a la cua C-terminal de la PaGS responsable de la trimerització en aquest enzim. Encara que la identitat de seqüència no és molt elevada entre eucariotes i arqueons en aquesta zona, la majoria dels canvis són conservatius. Aquestes característiques estructurals podrien indicar que les sintases eucariotes són trimèriques.

La GS de *Neurospora crassa* ha estat descrita com a trímer (Takahara i Matsuda 1978). Es va utilitzar el programa InterPreTS (Aloy i Russell 2002) per a predir si el tipus d'oligomerització d'aquesta sintasa seria equivalent al de PaGS. En efecte, la predicció indica que la glicogen sintasa de *N.crassa* conté els motius d'oligomerització equivalents a la PaGS i que pot formar aquest tipus de trímer amb més afinitat que la pròpia PaGS.

L'estat d'oligomerització de les GS eucariotes ha estat un tema molt controvertit. A causa de no disposar de l'estructura tridimensional de cap membre d'aquest grup, els seus estats d'oligomerització han estat inferits a partir d'experiments de gel filtració o ultracentrifugació, a partir del coeficient de sedimentació o de l'equilibri de sedimentació. Així, les sintases de mamífer han estat descrites com dímers, trímers i tetràmers per diferents autors. D'altra banda, la coneguda tendència de les GS de mamífer a formar agregats i la dependència d'aquest procés amb l'estat de fosforilació de la proteïna (McVerry i Kim 1972) compliquen aquest tipus d'anàlisi i poden ser la causa d'aquestes discrepàncies.

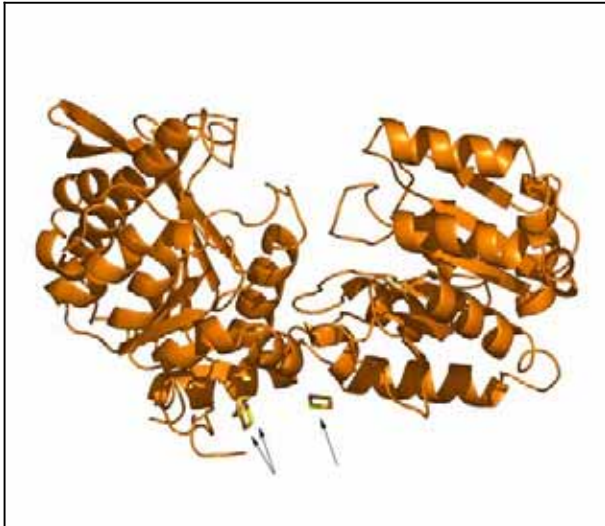
Dos estudis conclouen, a partir d'estudis d'equilibri de sedimentació, que la GS de múscul de conill i la de fetge de rata són trimèriques. Aquesta tècnica, també utilitzada per a determinar l'estat d'oligomerització de la PaGS és la més acurada, ja que és independent de la forma de la proteïna.

Prats i col·laboradors han donat evidències que la GS muscular podria existir com a trímer en determinades condicions (Prats *et al.* 2005). L'anàlisi de fibres musculars de conill deplecionades de glicogen, per microscòpia electrònica de transmissió, mostra la GS acumulada en unes estructures esfèriques que presenten un patró repetitiu pseudocristal·lí de formes triangulars. La mida del costat del triangle és d'aproximadament 180 Å, lleugerament superior als 130 Å calculats per a l'estructura de la PaGS, i que podria estar reflectint la major massa molecular de l'enzim muscular.

#### **4.7. Unió a dioxans:**

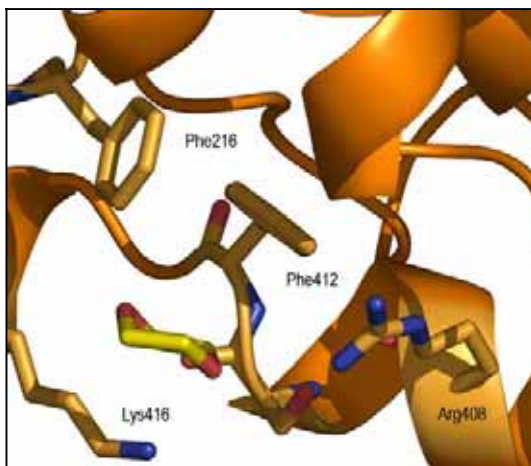
Diverses molècules de dioxà, que s'han utilitzades com a agent precipitant en el tampó de cristal·lització, s'han localitzat en l'estructura de la PaGS. La molècula de dioxà té una mida i una forma molt similars a la molècula de glucosa, aquest fet permet hipotetitzar que les zones on es localitzen els dioxans podrien esdevenir llocs amb afinitat per monosacàrids similars a la glucosa.

Es van localitzar dues molècules de dioxà en el colze que forma la cua C-terminal per a oligomeritzar amb la molècula veïna (Fig.4.48). La localització d'aquests residus, tot i ser suggerent ja que coincideix amb la zona d'oligomerització, no ens va permetre assignar-ne un significat fisiològic. De tota manera, s'ha de mencionar que és precisament en la zona de contacte entre subunitats on es troba el punt de regulació al·lostèric per AMP en la GP (Barford *et al.* 1991). Les interaccions que estableixen aquestes dues molècules amb la proteïna no semblen gaire específiques, ja que són fonamentalment hidrofòbiques, tipus apilament (*stacking*). Un dioxà interacciona amb les cadenes laterals dels residus Ile199 i Phe202, i l'altre amb els residus Tyr214 i Arg431.



**Fig. 4.48. Localització dels dioxans en l'estructura de la PaGS.** Dibuix d'un monòmer de PaGS on apareixen assenyalats amb fletxes les localitzacions dels tres dioxans localitzats en l'estructura.

La tercera molècula de dioxà es localitza en la part posterior de la zona d'unió entre dominis, unit a la cavitat definida en el gir entre les hèlixs  $\alpha 17$  i  $\alpha 18$ . Com es mostra a la figura 4.49, la molècula de dioxà estableix interaccions hidrofòbiques d'apilament (*stacking*) amb les cadenes laterals dels residus Phe216, Ser411, Phe412, Lys416, a més a més d'un pont d'hidrogen entre un dels oxígens i el nitrogen del grup guanidini de l'Arg408. La comparació de seqüències amb la resta de membres de la superfamília de les GS indica que aquests residus estan conservats, o bé que els canvis són conservatius. Cal destacar que l' Arg408 està conservada en totes les GS eucariotes, excepte en *C. elegans* (substituint per Gly), mentre que en bacteris l'equivalent estructural és una serina (Figura 4.31). El paper d'aquest residu serà tractat en el següent apartat.



**Fig. 4.49. Cavitat del dioxà en la part posterior de la unió entre dominis.** Dibuix de la interacció d'una molècula de dioxà amb els residus que formen la cavitat en la part posterior de la unió entre dominis: Phe216, Phe412, Arg408 i Lys416.

L'existència d'aquesta cavitat que pot acomodar una molècula tipus glucosa en un lloc que pot afectar els moviments del solc catalític resulta molt suggerent com a possible

punt de regulació d'aquests enzims. A partir d'aquestes observacions, es pot dissenyar un model mecànic simple on el control de la regulació estaria dirigit a l'obertura i tancament del solc catalític per la unió de molècules tipus glucosa a la part del darrera d'aquest.

Aquest principi de regulació de proteïnes per la part posterior de la frontissa formada entre els dos dominis de la proteïna ha estat descrit anteriorment en la literatura. Recentment, s'ha descrit un lloc de regulació alostèric per a la glucoquinasa (GK), l'enzim que catalitza la transformació de glucosa en Glc 6-P al fetge. Aquest enzim també presenta dos dominis clarament definits separats per un profund solc, en què es troba el centre catalític. La unió de molècules petites a la part posterior del seu solc catalític modula les propietats cinètiques de l'enzim (Grimsby *et al.* 2003; Kamata *et al.* 2004). En la glucoquinasa, el compost que s'uneix al lloc alostèric actua en forma de cuny, interacciona amb residus hidrofòbics de la part posterior del solc i desplaça l'equilibri entre les formes oberta (inactiva) i tancada (activa) cap a aquesta última.

En el cas de la PaGS, els residus que formen l'estreta frontissa, que coincideix amb el lloc d'unió de la molècula de dioxà, està formada per residus amb les seves configuracions esteses, molt conservats en les famílies GT-3 i GT-5 (Fig. 4.31). Aquests corresponen a Asn217, Gly218 i Ile219 en la direcció del N- al C-terminal i Phe412, Ser413 i Trp414 entre les hèlixs terminals  $\alpha 17$  i  $\alpha 18$ . La Gly218 s'apila sobre el Trp414 i pot rrelliscar sobre el grup aromàtic quan canvia l'apertura del solc catalític. D'altra banda, el tancament de l'enzim sembla augmentar l'exposició al solvent dels residus hidrofòbics Phe216 i Phe412. Aquesta exposició desfavorable faria desplaçar l'equilibri cap a la forma oberta (inactiva) de l'enzim en absència de substrats. La interacció d'un activador alostèric amb aquests residus hidrofòbics desplaçaria l'equilibri cap a una forma tancada i, per tant, activa de l'enzim.

Com hem introduït anteriorment, les GS eucariotes (família GT-3) estan regulades per fosforilació i activades alostèricament per Glc 6-P. En canvi, els paràmetres cinètics descrits per a la GS de *Pyrococcus furiosus* mostren que aquesta no es veu afectada per aquest sucre fosfat (Zea *et al.* 2003). Per tant, el possible lloc d'unió a una molècula tipus glucosa en l'estructura de la PaGS podria no tenir cap paper regulador o, per contra, unir un altre tipus d'activador alostèric no descrit i diferent a la Glc 6-P. En el cas de les sintases més complexes, com les eucariotes, aquest lloc hauria evolucionat fins a convertir-se en un lloc d'unió específic per a l'activador alostèric, Glc-6P.

#### **4.8. Modelat dels llocs de fosforilació de l'extrem C-terminal en GS eucariotes. Possible interacció amb la zona sensible a Glc-6P rica en arginines:**

Com hem exposat al llarg d'aquesta memòria, les GS eucariotes (GT-3) estan altament regulades per la fosforilació en residus serina i treonina, així com per l'activador alostèric Glc 6-P. L'acció d'aquest metabòlit pot arribar a contrarestar la inactivació

produïda per la fosforilació. Els llocs de fosforilació en mamífers es troben en els extrems N- i C-terminals mentre que en llevats només es fosforila la cua C-terminal.

L'alineament de seqüència entre la PaGS i les GS eucariotes mostra que les últimes tenen una cua C-terminal molt més llarga, la qual posseeix els llocs de fosforilació. La resolució de l'estructura de la PaGS ha permès determinar que existeix un gir en els últims aminoàcids implicat en l'oligomerització de l'enzim. El modelat de les estructures eucariotes sobre la PaGS indica que aquest gir apropa la cua C-terminal cap a la zona que en eucariotes conté les arginines implicades en la regulació al·lostèrica per part de la Glc 6-P i la inactivació per fosforilació. D'aquesta manera, es pot postular una interacció iònica directa entre els residus fosforilats de la cua C-terminal, carregats negativament, i les cadenes laterals positives de les arginines. L'afinitat d'aquests residus bàsics pels grups fosfat es posa de manifest en l'estructura del domini C-terminal de la PaGS en que els residus Lys407 i Arg408 interaccionen amb els ions sulfats del medi, equivalents als ions fosfat (apartat 4.1.2).

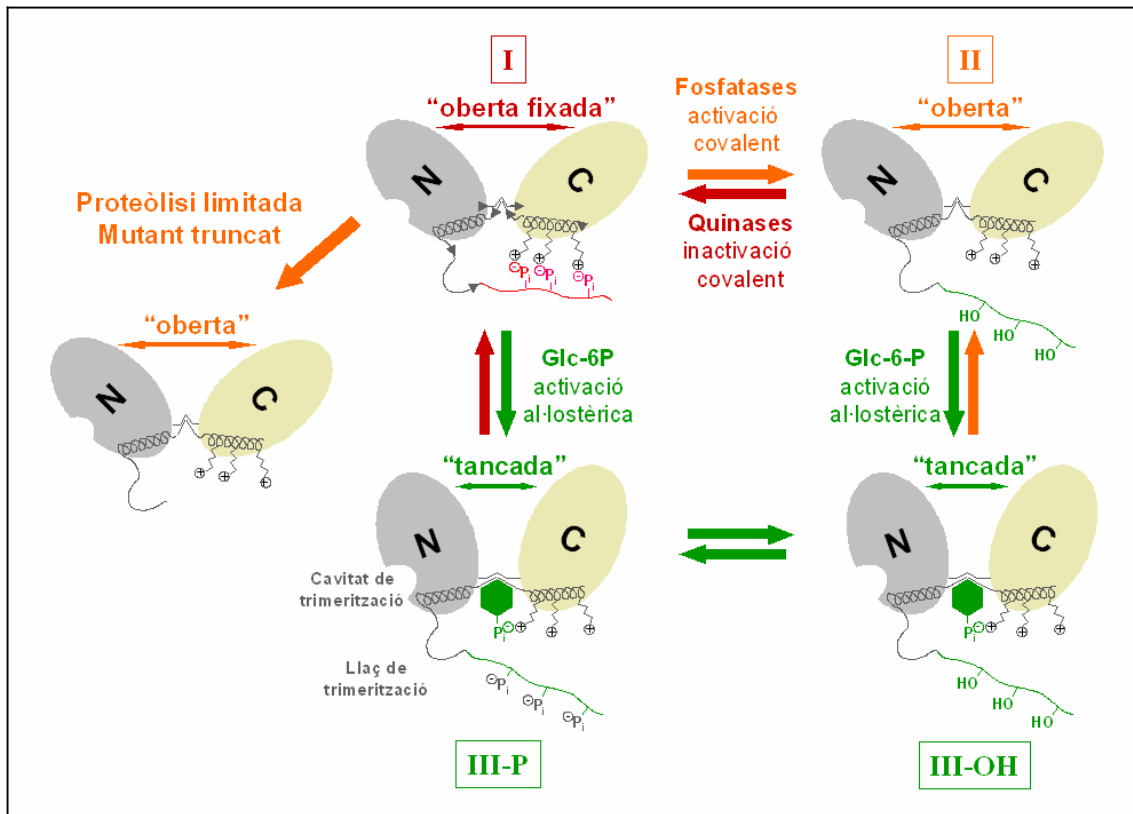
Si analitzem amb detall les seqüències de les cues C-terminals eucariotes, ens adonem que l'aminoàcid equivalent a l'últim residu de la PaGS i el primer lloc de fosforilació estan separats per 19 aminoàcids en totes les GS eucariotes excepte en les GS de llevat, en què aquest nombre ascendeix a 28 aminoàcids. L'elevat contingut en prolines en aquesta zona de les GS de mamífers fa pensar que estarà majoritàriament desplegada. A més, la comparació d'aquesta seqüència amb la base de dades del PDB la relaciona amb dominis interns estesos o bé coincideix amb dominis exposats no ordenats. La predicció de l'estructura secundària calculada amb el programa Jpred també assigna una disposició *random coil* a aquest polipèptid. La longitud equivalent a una cadena d'aquestes característiques és d'uns 20-25 Å, exactament la distància necessària per a fer arribar els llocs de fosforilació a la zona de les arginines.

La informació de què disposem a partir de l'estructura de la PaGS no ens permet però determinar com seran les característiques estructurals de la cua fosforilable N-terminal de les sintases de mamífer.

#### 4.9. Model mecànic:

La possible oligomerització trimèrica d'aquests enzims, juntament amb la possible localització de la cavitat d'unió per a la Glc 6-P en la part del darrera del solc catalític, permeten suggerir un model simple de regulació per a les GS eucariotes basat en el control mecànic de la transició entre l'obertura i el tancament del solc catalític.

Per a hipotetitzar un mecanisme que justifiqui les observacions experimentals, ens hem centrat en la GS de llevat. Es tracta d'una GS eucariota més senzilla, ja que només disposa d'una cua C-terminal, a diferència de les GS de mamífer, les quals presenten dues cues regulables (Fig. 4.50).



**Fig. 4.50. Mecanisme mecànic per les GS de llevat.** Esquema del mecanisme mecànic proposat per a les GS de llevat que es representen pels dos lòbuls N (gris) i C (beix). La cua C-terminal és de color verd quan els residus estan desfosforilats i vermella quan estan fosforilats.

La forma fosforilada de l'enzim es mantindria en una configuració "oberta fixada" mitjançant la interacció electrostàtica dels residus fosforilats amb els residus d'arginina situats en la última hèlix del domini C-terminal i que correspondrien a l'hèlix  $\alpha 17$  en PaGS (estat I). Una desfosforilació de l'enzim per l'acció de les fosfatases, faria desaparèixer aquesta interacció electrostàtica i l'enzim quedaria lliure per a obrir-se i tancar-se lliurement, donant lloc a un enzim actiu (estat II).

La proteòlisi limitada de l'enzim fosforilat així com l'expressió de la seva forma truncada donen lloc a un enzim activat (Hardy i Roach 1993), equivalent a la seva configuració oberta i lliure (estat II). Aquesta evidència experimental està en consonància amb el model mecànic.

L'activació al·lostèrica per Glc 6-P és capaç de contrarestar la inactivació per fosforilació (Villar-Palasi i Guinovart 1997). Segons el model proposat, en la GS fosforilada inactiva, la Glc 6-P s'uniria a la cavitat posterior del solc catalític i competiria per la interacció amb l'arginina més propera, corresponent a l'Arg 408 en PaGS. En llevat, l'arginina equivalent és l'Arg 591, la qual pertany, juntament amb les Arg 586 i Arg 588, al grup de residus que en ser substituïts per alanina donen lloc a un mutant insensible a l'efecte activador de la Glc 6-P (Hardy i Roach 1993). Aquest mutant però continua inactivant-se per fosforilació. Aquest fet es podria explicar tenint en compte que la cadena de residus entre l'últim residu homòleg a la PaGS fins al primer residu



fosforilable és de 28 en comptes de 19 aminoàcids (mamífers), per això els residus fosforilats de llevat poden continuar interaccionant amb els residus d'arginina més allunyats del lloc d'unió a Glc 6-P. L'efecte de cuny que exerciria la Glc 6-P podria fer tancar el solc catalític, de manera que es trencarien les interaccions entre els residus fosforilats i les arginines, i l'enzim quedaria alliberat per l'activació al·lostèrica (estat III<sub>P</sub>).

En una GS de llevat activa i desfosforilada, la unió de Glc 6-P en forma de cuny a la part posterior del solc catalític, provocaria un desplaçament de l'equilibri des d'una forma oberta lliure cap a una forma més tancada i, per tant, més activa (estat III<sub>OH</sub>).

D'aquesta manera es podria definir en les GS una cara reguladora, cara R, en la que es trobaria la cavitat d'unió a Glc 6-P i la zona d'arginines, i amb les que interaccionaria la cua fosforilable dels enzims eucariotes.

## DISCUSSIÓ

Tal i com s'havia predit per alineaments de seqüència i predicció d'estructura (Cid *et al.* 2002), la PaGS es plega en dos dominis  $\alpha\beta$  tipus plegament de Rossmann, comú a totes les GT del grup GT-B (Coutinho *et al.* 2003). De les estructures resoltes, les més semblants a la PaGS són les dels membres de la superfamília GT-B que, de la mateixa manera que la PaGS, actuen amb retenció de configuració del centre anomèric. De totes, l'AtGS, que pertany a la mateixa família GT-5, és la més similar a la PaGS (Buschiazzo *et al.* 2004).

A més de les similituds en el plegament global, els residus implicats en catàlisi i en la unió als substrats estan molt conservats, fet que permet la comparació funcional entre ells i fer prediccions per a altres membres de les famílies GT-3 i GT-5.

L'especificitat pel glucòsid de difosfonucleòtid en les diferents GS ve determinada per la interacció d'uns pocs residus amb l'anell heterocíclic del nucleòtid. Així, la cadena lateral de la Tyr354 d'AtGS estableix interaccions hidrofòbiques tipus apilament exclusivament amb l'adenina de l'ADPG, el qual forma dos ponts d'hidrogen amb dos grups de l'esquelet proteic. La cavitat d'unió a nucleòtid en la PaGS permet la formació de ponts d'hidrogen amb un anell heterocíclic d'adenina o d'uridina (ADPG o UDPG respectivament), simplement amb lleugers desplaçaments entre les posicions d'unió de les bases. A més, la flexibilitat de la cadena lateral de la Met316, encarregada de les interaccions hidrofòbiques amb l'anell facilita els apilaments en cadascuna de les posicions alternatives. En les GS eucariotes, que utilitzen exclusivament UDPG com a donador de glucosil, el residu equivalent a la Met316 és una Phe en tots els casos. Aquest residu podrà interaccionar amb l'anell de la uridina, el qual establirà previsiblement ponts d'hidrogen específics amb l'esquelet de la proteïna equivalents a PaGS.

En el centre catalític, la conservació dels residus és molt elevada entre els diferents membres del grup GT-B que actuen amb retenció de configuració del centre anomèric (GS, GP, OtsA). Tots aquests residus, que corresponen a Glu339, His151, Asn217, Arg257 i Gly17 en PaGS, estan conservats en la resta d'enzims. A més, hem identificat la presència d'una complexa xarxa de ponts d'hidrogen entre el residu His151, His127, His125, Thr149, Thr191 i Tyr425. Aquest grup d'aminoàcids podria estar involucrat en un possible mecanisme de transmissió de càrregues (*charge relay system*), que constituïria un canal de protons entre el centre catalític i l'exterior. Aquests residus estan conservats en la resta de GS de les famílies GT-3 i GT-5 però no en GP. Aquest i altres indicis suggereixen que les GS i les GP podrien actuar a través de mecanismes catalítics no estrictament equivalents.

Quant a la unió del substrat acceptor, la comparació de les GS amb les GP permet predir les interaccions que s'establiran amb les glucoses en posició +1 i +2, però a partir de la posició +3 es perd l'homologia estructural i augmenten les diferències, fins i tot entre les pròpies GS.

D'altra banda, la localització d'una glucosa a la superfície de la PaGS ha permès hipotetitzar sobre l'existència d'un segon lloc d'unió a glicogen, el qual permetria a l'enzim mantenir-se unit al seu substrat, augmentant la seva concentració efectiva, de la mateixa forma que s'ha descrit per a les GP eucariotes (Johnson *et al.* 1988; Goldsmith *et al.* 1989; Johnson *et al.* 1990). Aquest fet estaria en consonància amb els resultats experimentals que indiquen que les GS eucariotes es mantenen unides al glicogen en l'interior de la cel·lula (Ferrer *et al.* 1997; Garcia-Rocha *et al.* 2001; Prats *et al.* 2005).

L'estructura cristal·logràfica de la PaGS, a diferència de la resta de les estructures resoltes del grup GT-B que formen dímers o tetràmers, és trimèrica. L'altre GS resolta, AtGS, forma un dímer cristal·logràfic. Aquesta diferència d'oligomerització dins d'una mateixa família GT-5 fa difícil predir l'estat d'oligomerització de les GS eucariotes. No existeix cap estructura resolta de la família GT-3, i els resultats obtinguts per a la determinació de l'estat d'oligomerització de les GS en solució han estat en molts casos contradictoris. L'alineament estructural de les estructures de les dues GS resoltes i l'aplicació d'aquest perfil a la resta de seqüències de GS eucariotes ens ha permès hipotetitzar que aquestes últimes podrien, en efecte, formar trímers.

L'oligomerització de la PaGS només implica residus dels dominis N-terminals de diferents subunitats deixant lliure el moviment dels dominis C-terminals. La implicació d'un sol domini en l'oligomerització és comú entre les GT-B d'aquest grup i es relaciona amb la llibertat de moviment necessària per obrir i tancar el solc catalític.

Aquest tipus d'oligomerització de la PaGS dona lloc a una estructura triangular plana en la qual es poden definir dues cares: la cara G que seria la que s'uniria a glicogen i la cara contrària, que contindria la maquinària de regulació de l'enzim i que s'ha anomenat cara R.

La localització d'una cavitat capaç d'unir molècules tipus glucosa en la part posterior del solc catalític ens ha permès suggerir que aquest seria un bon punt de regulació de l'activitat enzimàtica a través del control de l'obertura i tancament d'aquest solc, com succeeix en altres enzims (Grimsby *et al.* 2003; Kamata *et al.* 2004). Els residus que formen aquesta cavitat així com els que es troben en la unió entre els dos dominis estan conservats en les GS eucariotes i els pocs canvis que existeixen són conservatius. Per tant, aquest podria ser el lloc d'unió de Glc 6-P en GS eucariotes, responsable de la seva activació al·lostèrica.

A més d'activar-se al·lostèricament per Glc 6-P, els enzims eucariotes s'inactiven per fosforilació en residus serina i/o treonina situats en les cues N- i C-terminal, que no estan presents en les GS de bacteris ni en les d'arqueons. El modelat de la cua C-terminal de les GS eucariotes, a partir de l'extensió de l'estructura de la PaGS, ha permès hipotetitzar que aquests residus podrien interaccionar directament amb una zona d'arginines implicada en la regulació per Glc 6-P i per fosforilació. Aquesta regió bàsica es troba molt propera a la hipotètica cavitat d'unió a l'activador al·lostèric, fins al punt que una de les arginines que la integren formaria part d'aquesta cavitat. En absència de Glc 6-P, els llocs de fosforilació de la cua C-terminal interaccionarien en forma de cremallera amb les arginines, bloquejant l'enzim en una configuració oberta inactiva. En presència de l'activador al·lostèric, el fosfat de la Glc 6-P competiria amb el primer lloc de fosforilació per la unió a aquesta primera arginina, eliminant així la inactivació provocada per la fosforilació. A més a més, la unió de Glc-6-P afavoriria el tancament de la cavitat catalítica, conduint a l'activació de l'enzim.

En aquest model mecànic, les zones de regulació per l'activador al·lostèric, Glc 6-P, i per fosforilació controlarien l'enzim per la part posterior del solc catalític de cada subunitat, en el que s'ha anomenat cara R. Aquesta zona seria accessible per a les quinases i les fosfatases encarregades de fosforilar i desfosforilar la cua de les GS eucariotes ja que la interacció amb les branques de glicogen es produiria per la cara oposada, la cara G.

## DISCUSSIÓ GENERAL

Les GS catalitzen un pas clau en la síntesi de glicogen, polisacàrid present en la gran majoria d'eucariotes, bacteris i arqueons. De la mateixa forma, en plantes, les MS catalitzen l'elongació de midó. Tot i catalitzar la mateixa reacció, les GS d'eucariotes han desenvolupat uns mecanismes de regulació molt complexos en relació a la resta de membres de les famílies GT-3 i GT-5. En comparar les seqüències d'aquestes GS, es comprova que l'aparició de la regulació per fosforilació en les GS eucariotes és producte de l'addició de fragments de seqüència en els seus extrems, els quals són absents en les GS bacterianes, d'arqueons i en les MS de plantes. El diferent grau de sofisticació en els mecanismes de regulació es pot deure al fet que els organismes eucariotes depenen en gran mesura de la metabolització de la glucosa externa, i fan imprescindible un control estricte del seu emmagatzematge.

Per a conèixer les bases moleculars de la catàlisi i la regulació d'aquests enzims, es feia imprescindible disposar de la seva estructura tridimensional. Els projectes encetats per a determinar les estructures de GS eucariotes, paradigmes de la regulació, no havien donat fruits a causa de la insolubilitat de les proteïnes obtingudes de forma recombinant en sistemes bacterians. El projecte de resoldre l'estructura d'una GS més senzilla, la GS de l'arqueon *P. abyssi*, facilitaria *a priori* les etapes inicials d'obtenció de proteïna.

Les proves d'expressió heteròloga de la PaGS ens enfrontaren al mateix problema d'insolubilitat. Aquesta repetida manca de solubilitat pot ser deguda a causes diverses, però també pot indicar l'existència d'un patró comú en el comportament d'aquestes GS. En aquest segon cas, hauria d'existir alguna característica en les GS eucariotes, diferent a les cues de fosforilació, que causés la insolubilitat, ja que les GS bacterianes, que no disposen d'aquestes cues, s'expressen de forma soluble.

Comparant les estructures de la PaGS i de l'AtGS, existeixen poques diferències estructurals significatives. Les més destacables són les responsables de l'estat d'oligomerització en cadascun d'aquests enzims. Així, la PaGS posseeix una cua hidrofòbica, inexistent en AtGS, que es projecta cap a una cavitat hidrofòbica complementària de la subunitat veïna, evitant així el contacte d'ambdues regions amb el solvent. Durant l'expressió heteròloga de l'enzim, aquesta cua exposada al solvent podria ser responsable de la inestabilitat i de la seva consegüent insolubilitat. Els enzims eucariotes haurien conservat aquests residus hidrofòbics tot i que haurien afegit la cua de fosforilació a continuació. Un experiment relativament senzill que aportaria informació valuosa al respecte consistiria en l'expressió del mutant per deleció de l'enzim d'arqueon en el que s'hagués eliminat aquesta cua hidrofòbica. Aquest enzim hauria perdut la capacitat per oligomeritzar i potser també la tendència a formar cossos d'inclusió.

Ens vam proposar la renaturalització dels cossos d'inclusió formats per la PaGS i vam aconseguir-ho, posant a punt un mètode de replegament en columna d'afinitat. Aquest mètode podria ser també aplicable a les GS eucariotes.

La PaGS forma un trímer molt estable en solució, ja que en cap de les condicions assajades hem observat dissociació del complex. Adopta una forma triangular plana en que l'oligomerització només implica els dominis N-terminal. Els dominis C-terminal no participen en les interaccions entre subunitats i, per tant, l'organització trimèrica de la PaGS no restringeix els moviments a través de l'estreta frontissa (*hinge*) que connecta els dos dominis. L'existència de frontisses mòbils en aquest complex sistema trimèric ha estat una de les causes que han dificultat la seva cristal·lització, així com la seva resolució per reemplaçament molecular.

Les característiques de les GS d'arqueons les situen en un terme mig entre les GS bacterianes i les GS de mamífers i fongs (Cid *et al.* 2002). Atesa la naturalesa aparentment primitiva d'aquests organismes, les GS d'arqueons podrien conservar les propietats d'una GS ancestral a partir de la qual haurien evolucionat la resta de membres. D'aquesta forma, s'explicaria la seva promiscuïtat per utilitzar UDPG i ADPG com a donadors de residus glicosil, oposada a l'especificitat divergent que han adquirit les GS de bacteris (ADPG) i eucariotes (UDPG). Per aquestes raons i altres similituds estructurals, les GS eucariotes s'assemblen més a les d'arqueons que a les GS bacterianes. D'altra banda aquesta observació no és inesperada, atesa la coneguda major proximitat filogenètica dels enzims eucariotes amb els enzims d'arqueons, en comparació amb els bacterians (Woese *et al.* 1990).

La conservació de la seqüència en les zones de la PaGS implicades en la trimerització en eucariotes, i altres evidències experimentals fan pensar que les GS superiors hagin pogut mantenir aquesta estructura quaternària. En eucariotes, la interacció dels residus hidrofòbics de l'extrem C-terminal a la cavitat de la subunitat veïna esdevindrien un punt d'anclatge perfecte per a poder projectar la cua C-terminal fosforilable cap a la zona de regulació de l'enzim.

El model mecànic que hem proposat és el primer model amb base estructural que explicaria el mecanisme bàsic pel qual les GS eucariotes són regulades. La regulació de l'enzim es basaria en controlar els moviments d'obertura i tancament del solc catalític i bloquejaria el seu moviment dels elements de seqüència que connecten ambdós dominis per la part oposada al centre actiu. Si visualitzem una frontissa d'una porta, resulta senzill imaginar que si hi col·loquem un cuny per la part posterior bloquejarem completament el seu moviment d'obertura. Postulem que l'acció realitzada per l'activador al·lostèric Glc 6-P podria anar en aquesta direcció. La cua fosforilada de l'enzim interaccionant amb la zona rica en arginines actuaria de forma comparable a la d'una corda enganxada al pom de la porta i l'ancoraria formant un determinat angle per la part del darrera de la frontissa. D'aquesta forma s'impedeix el moviment de tancament.

Per tant, la regulació de l'enzim es localitzaria en la superfície posterior de l'enzim i oposada a la superfície del solc catalític. La topologia seria alhora molt important ja que permetria regular l'enzim mitjançant la desfosforilació i fosforilació de l'enzim, per l'acció de les fosfatases i quinases respectivament, i de forma independent a la unió d'aquest a un dels seus substrats naturals, el glicogen. Amb aquest tipus de trímer pla, l'existència

d'una cara G (pel glicogen) i una cara R (per la regulació) evitaria conflictes estèrics entre el glicogen i el conjunt d'enzims i proteïnes que interactuen amb les GS eucariotes.

Resulta essencial contrastar les nostres hipòtesis amb l'estructura d'una GS eucariota. L'optimització de mètodes de purificació alternatius ens pot permetre obtenir proteïna per a encetar aquests estudis estructurals. Si escollim l'expressió heteròloga en bacteris, el model de les GS eucariotes a partir de la PaGS permet determinar que les cues d'aquestes GS estaran possiblement desordenades i mòbils, donat que estaran desfosforilades, i per tant, seran poques les possibilitats de poder cristal·litzar l'enzim en aquestes condicions. Una alternativa podria ser la fosforilació *in vitro* de la cua C-terminal, de tota manera aquesta podria donar lloc a l'aparició de diferents poblacions amb diferents graus de fosforilació. Una segona alternativa implicaria la construcció d'enzims truncats en què s'haguessin eliminat aquestes cues i/o el seu punt d'anclatge. En definitiva, el modelat d'aquestes GS eucariotes sobre la PaGS ens permetrà determinar "per on" tallar.

A més, el mecanisme catalític a través del qual actuen aquests enzims continua essent desconegut. La PaGS és un excelent candidat per abordar determinats tipus d'estudis estructurals. Atesa la seva naturalesa termoestable, la seva activitat a temperatura ambient és molt menor que per a la resta d'enzims, i per tant és possible que es puguin preparar complexos de l'enzim a baixa temperatura en que estiguin presents ambdós substrats sense que s'arribi a produir la reacció. Aquests resultats serien molt valuosos, ja que ens podrien donar informació sobre el moment previ a la catàlisi. La posició relativa dels residus en aquesta fase podria aclarir quina és la implicació real d'aquests, ja que fins al moment, la informació estructural de la que es disposa no permet definir el paper específic dels residus del centre actiu en la catàlisi.

La determinació de l'estructura de la PaGS ha permès resoldre i suggerir alguns aspectes del funcionament de les GS i, al mateix temps, ha plantejat noves incògnites, la resolució de les quals esdevindrà tot un repte pels propers anys.

## CONCLUSIONS

El marc obert de lectura PAB2292 del genoma de *P.abysyi* codifica per a un enzim amb activitat glicogen sintasa. Aquest enzim és termoestable i presenta una activitat màxima a 80°C i és capaç d'utilitzar indistintament ADPG i UDPG com a donadors de glucosil.

La PaGS forma de cossos d'inclusió en totes les condicions d'expressió en *E.coli* assajades. Amb el mètode de renaturalització en columna hem aconseguit que la PaGS es replegui correctament i ens ha permès obtenir la quantitat de proteïna pura necessària per encetar estudis de cristal·lització i de caracterització en solució.

El domini C-terminal de la PaGS s'expressa de forma soluble i la seva purificació ens ha permès disposar de quantitat suficient de proteïna per a realitzar estudis de cristal·lització. Aquest domini és monomèric tant en solució com en el cristall.

La resolució de l'estructura del domini C-terminal a 1.7 Å ha permès utilitzar-lo com a model inicial per a resoldre l'estructura de la PaGS a 2.8 Å per reemplaçament molecular.

La PaGS és un trímer tant en solució com en el cristall, amb una disposició triangular plana. Cadascuna de les seves subunitats està formada per dos dominis  $\alpha\beta$  tipus plegament de Rossmann separats per un solc molt profund on es troba el centre catalític de l'enzim. L'estructura global és molt similar a la de la resta d'estructures resoltes del grup GT-B que actuen amb retenció de la configuració anomèrica.

La trimerització de la PaGS implica exclusivament els dominis N-terminals de l'enzim, quedant els dominis C-terminals lliures per a poder realitzar la transició d'obertura i tancament necessària per a la catàlisi.

El domini C-terminal conté la cavitat d'unió a nucleòtid, mentre que la majoria d'interaccions amb l'oligosacàrid acceptor es produeixen a través de residus del domini N-terminal. L'elevada similitud del centre actiu entre les GS i les GP suggereix que ambdues operen a través d'un mecanisme catalític similar, si be no idèntic.

L'estructura de la PaGS i la seva comparació amb l'estructura de l'AtGS ofereixen les bases per a entendre l'especificitat de les GS eucariotes per UDPG, com a substrat donador de glucosil, i la promiscuïtat de les GS d'arqueons per unir tant UDPG com ADPG.

La localització d'una molècula de glucosa en la superfície de la PaGS i la comparació amb les GP eucariotes permeten suggerir un possible lloc d'emmagatzematge del glicogen en les GS.

La presència d'una molècula de dioxà en la part posterior del solc catalític suggereix que la unió de molècules similars a la glucosa en aquest lloc podria controlar l'activitat de l'enzim, mitjançant la regulació de l'obertura i el tancament d'aquest solc.

La comparació de seqüències i l'alineament estructural suggereixen que les GS eucariotes probablement també presenten una organització molecular trimèrica. El modelat simple d'aquestes GS a partir de l'estructura de la PaGS ha permès dissenyar un model dels mecanismes de regulació alostèrica i per fosforilació de l'activitat de les GS de fongs, compatible amb les evidències experimentals.



## MATERIALS I MÈTODES

### A. BIOLOGIA MOLECULAR:

#### A.1. Tècniques estàndards:

##### A.1.1. PCR:

La reacció en cadena de la polimerasa, PCR (Saiki *et al.* 1988) s'utilitza per a amplificar fragments de DNA a partir de DNA genòmic, així com a partir de plasmidis. La PCR utilitza la Pfu DNA polimerasa (Stratagene), un enzim d'elevada fidelitat de còpia, capaç de sintetitzar DNA a les temperatures necessàries per a deshibridar les cadenes de DNA i hibridar els oligonucleòtids a les cadenes complementàries de DNA, atesa la seva naturalesa termoestable. L'aparell emprat fou Applied Biosystems GeneAmp PCR System 2400.

El tampó adequat per a la reacció conté magnesi, dNTPs i dos oligonucleòtids, directe i invers, complementaris als límits del fragment de DNA a amplificar. Els oligonucleòtids han de ser el més específics possible a la regió del DNA on s'han d'hibridar, amb una llargada mitjana d'uns 20 nucleòtids i un contingut mínim del 50% en bases G i C, evitant la formació d'estructures secundàries.

##### A.1.2. Clonatge:

Procés a partir del qual s'aconsegueix la lligació del DNA, el qual implica reaccions de digestió i obtenció de DNA amb extrems roms o cohesius segons allò descrit per Sambrook i col·laboradors (Sambrook 2001).

L'enzim T<sub>4</sub> DNA lligasa (Rapid Ligation Kit, Roche) s'emprà per a lligar inserts de DNA i plasmidis durant 2 o 3 hores a temperatura ambient. Les quantitats totals de DNA oscil·laven entre 10-500 ng de DNA amb un excés de les molècules d'insert respecte a les de plasmidi (relació 3 :1).

Per a clonar els fragments de DNA que provenien de la PCR, es va utilitzar el paquet comercial Sure Clone Rapid Ligation Kit (Pharmacia). El cDNA clonat es lligà al vector pGFPCR (Cormack 1997) digerit amb SmaI.

##### A.1.3. Transformació de cèl·lules competents:

Tècnica a partir de la qual s'introdueix DNA exògen, plasmidi que conté el DNA d'interès, en cèl·lules procaríotes. Per a aconseguir la internalització del plasmidi, s'altera la permeabilitat de la membrana bacteriana mitjançant xoc tèrmic (Sambrook 2001).

Es va afegir el DNA plasmídic (relació 10:1) a 50 µl de cèl·lules competents pretractades amb clorur de calci. Després de 10 minuts en gel, s'incubà la barreja 30 s a 42°C i 5 minuts a 4°C. S'afegí 1 mL de medi LB (1% triptona, 0.5% extracte de llevat i 1% NaCl) i s'incubà durant 1 hora a 37°C, en agitació constant. A continuació, es van sembrar 200 i 800 µL de la suspensió cel·lular sobre plaques de LB-agar (1.5% agar), que contenien l'antibiòtic, al qual era resistent cada tipus de plasmidi, i s'incubaren a 37°C durant tota la nit.

S'utilitzaren cèl·lules competents d'E.coli DH5α per als passos inicials de biologia molecular i la soca d' E.coli BL21-Codon Plus (DE3)-RIL (Stratagene), transformada amb els plasmidis d'expressió per a produir la proteïna recombinant. Els plasmidis utilitzats, pET28b(+) (Novagen) i pGFPCR, conferien resistència a kanamicina i a ampicilina, respectivament. La concentració final de l'antibiòtic era de 100 µg/mL. A més, la soca BL21-CodonPlus (DE3)-RIL contenia plasmidis que conferien resistència a cloranfenicol, utilitzat a una concentració final de 34 µg/ml.

#### **A.1.4. Obtenció de DNA plasmídic:**

Per a aïllar el DNA plasmídic dels cultius cel·lulars, es va utilitzar la lisi inicial de la paret cel·lular mitjançant tractament alcalí (lisi alcalina), seguida per l'absorció del DNA a una reïna de sílice en presència d'un tampó d'alta sal i la posterior elució amb un tampó de baixa sal (Tris-HCl 10mM, EDTA 0.1mM, pH8.0).

Es va obtenir DNA plasmídic utilitzant dos protocols, en funció del volum de cultiu bacterià inicial: miniprep en el cas de cultius de 2 mL o maxiprep per a cultius de 200mL (Qiagen Plasmid Maxiprep).

#### **A.1.5. Seqüenciació del DNA:**

Per a comprovar l'absència d'errors, es van seqüenciar totes les construccions de DNA mitjançant el paquet de seqüenciació ABI-Prism DNA (Applied Biosystems). Es dugué a terme una reacció de PCR i es van usar dNTPs marcats amb diferents grups fluoròfors. A continuació, les cadenes de DNA foren analitzades per electroforesi capil·lar i fluorimetria amb el seqüenciador ABI Prism 3700 (Applied Biosystems) al servei de seqüenciació dels Serveis Científic-Tècnics (SCT) de la Universitat de Barcelona al Parc Científic de Barcelona.

#### **A.1.6. Electroforesi de DNA:**

Per a separar molècules de DNA de diferents mides s'utilitzà l'electroforesi en gel d'agarosa segons allò descrit per Sambrook i col·laboradors (Sambrook 2001).

El gel s'obtingué dissolent 1.5% d'agarosa en tampó TAE (TrisHCl 40 mM, EDTA 2mM), es va escalfar, s'afegiren 0.5 µg/mL de bromur d'etidi i es deixà gelificar dins de la cubeta d'electroforesi. Es carregaren les mostres, a les quals s'havia afegit el tampó de càrrega a una concentració final de 5% de glicerol i 0.05% de blau de bromofenol i es va

sotmetre a una diferència de potencial de 70 V en el tampó d'electroforesi TAE. El DNA és visible a l'UV atès que les molècules de bromur d'etidi presents en el gel d'agarosa s'intercalen entre les cadenes de DNA i donen lloc a una fluorescència rosada.

## A.2. Obtenció dels vectors d'expressió:

S'amplificà per reacció de PCR el marc obert de lectura que codificava per a la hipotètica PaGS (PA2292) a partir del genoma de *Pyrococcus abyssi*, utilitzant l'oligonucleòtid PYA-UP (TTCCATATGAAGGTGTTATTGCTTGG), que conté la diana per l'enzim de restricció *NdeI*, i l'oligonucleòtid PYA-DN (GTCTGACTCATAGGATGAAGTCAAACGC), que conté la diana per l'enzim de restricció *SalI*. El producte resultant fou clonat en el vector pGFPCR, que havia estat digerit prèviament amb l'enzim *SmaI* (pGFPCRPaGS). S'obtingué el DNA plasmídic dels clons positius i es digerí amb els enzims *NdeI* i *SalI*. L'insert fou clonat en el vector pET28b(+) tallat amb els mateixos enzims de restricció, i va donar lloc a pET28HisPaGS. S'obtingué DNA plasmídic de les col·lònies positives i es seqüencià completament la zona codificant per a la PaGS per a descartar possibles errors. Per últim, es transformà la soca BL21-CodonPlus (DE3)-RIL amb el plasmídi pET28HisPaGS.

Per a obtenir el plasmídi que contingués el DNA que codifica per a l'hipotètic domini C-terminal de la PaGS, s'amplificà aquest cDNA per PCR utilitzant pGFPCRPaGS com a motlle. Com encebadors s'utilitzaren els oligonucleòtids HIS-CTAL-UP (CGCGCGGCA GCCATAATGGGATTGATTGTAGC i HIS-CTAL-DN (CCGCAAGCTTGTCGATC AGGAAAATGACATTGC), els quals contenen les dianes *NdeI* i *SalI*, respectivament. Vam obtenir DNA plasmídic de les col·lònies positives, el vam digerir amb els enzims i vam clonar l'insert en el vector pET28b(+), tallat amb els mateixos enzims (pET28HisCterminal). El DNA plasmídic obtingut de les col·lònies positives es va seqüenciar completament per a descartar possibles errors. Per últim, es transformà la soca BL21-CodonPlus (DE3)-RIL amb el plasmídi pET28HisCterminal.

## B. TÈCNiques D'EXPRESSIÓ RECOMBINANT:

S'utilitzà el sistema de la T7 polimerasa per a expressar les proteïnes heteròlogues en cultius bacterians (Studier i Moffatt 1986). La soca escollida, BL21-CodonPlus (DE3)-RIL, conté el gen de la T7 RNA polimerasa (lisogen  $\lambda$ DE3) per a transcriure el gen d'estudi, el qual es troba sota el control del promotor T7, induïble per IPTG. A més, aquesta soca conté còpies extres dels gens, que codifiquen pels tRNAs que habitualment limiten la traducció de les proteïnes heteròlogues (argU, ileY i LeuW).

Per a cadascuna de les proves d'expressió a diferent temperatura, es van inocular amb una col·lònia positiva 2 mL de medi LB suplementat amb kanamicina 100  $\mu$ g/mL i cloranfenicol 34  $\mu$ g/mL, es van créixer 12h a 37°C, amb agitació constant. La suspensió cel·lular es va utilitzar per a inocular 200 mL de LB amb kanamicina i cloranfenicol i es va créixer a 37°C, fins que la densitat òptica mesurada a 600 nm va arribar a 0.6. Cada

cultiu es va atemperar a la temperatura d'assaig i es va induir l'expressió afegint IPTG 0.4 mM.

A partir d'aquest punt, els cultius bacterians van créixer a la temperatura d'estudi: 15, 22 o 37°C. Es van anar prenent mostres en diferents temps per a analitzar els nivells d'expressió mitjançant electroforesi amb gels desnaturalitzants de poliàcrilamida.

Un cop establertes les condicions òptimes d'expressió, 24 hores a 37°C, tant per a la proteïna sencera com per al domini C-terminal aïllat, es procedí a l'escalat de la tècnica. Es preparava un primer preinòcul de 2mL que, un cop crescut s'afegia a 200mL de medi LB fresc amb antibiòtics. Quan aquesta segona suspensió cel·lular havia crescut a saturació, s'utilitzava per a inocular 2 L de medi LB. Quan el nou cultiu assolía una densitat òptica de 0.6 a 600nm, era induït amb IPTG 0.4 mM i es deixava créixer durant 24 hores a 37°C.

Per a produir proteïna amb selenometionines, ens basàrem en el protocol establert per Van Duyne i col·laboradors (Van Duyne *et al.* 1993). Els preinòculs van ser equivalents als de les proteïnes nadiues. Un cop aquests havien crescut a saturació, s'afegien a 2L de medi mínim suplementat amb 0.2% glucosa, 2 µM MgSO<sub>4</sub>, 0.1 µM CaCl<sub>2</sub> i una barreja d'aminoàcids essencials a 40 mg/l que no contenia Gly, Ala, Pro, Asn, Cys ni Met. Quan el cultiu havia assolit una densitat òptica de 0.6 a 600 nm, s'afegien 50 mg de selenometionina per litre de cultiu. Al mateix temps s'inhibia la síntesi de metionina afegint 100 mg de Phe, Thr i Lys i 50 mg de Leu, Ile, Val i Pro per litre de cultiu. Després de 15 minuts, s'induïa l'expressió afegint IPTG fins a una concentració final de 0.4mM i es deixava incubant a 37°C durant 24 hores.

Per a obtenir una quantitat suficient de proteïna i poder desenvolupar els passos següents, cada experiment implicava el creixement de 14 L de cultiu per a la HisPaGS i 4 L per al domini C-terminal.

## **C. TÈCNiques DE PURIFICACIÓ:**

Els extractes bacterians foren homogenitzats a 4°C amb el tampó A (500 mM clorur de sodi, 20 mM fosfat de sodi, pH 7.4), trencant les membranes bacterianes per sonicació. A continuació, es centrifugà a 10000 g i es separà el sobrenedant del sediment. Per a purificar les proteïnes a estudiar, utilitzàrem diferents tècniques cromatogràfiques.

### **C.1. Cromatografia d'afinitat:**

Aquesta tècnica permet la separació de proteïnes en base a la interacció reversible entre un lligand específic unit covalentment a una matriu i una proteïna. En el nostre estudi, s'utilitzà la cua de 6 histidines en l'extrem N-terminal per a interaccionar amb l'ió

Ni<sup>2+</sup> unit a la reïna. Aquesta tècnica fou escollida com a primera etapa en la purificació del domini C-terminal de la PaGS.

Es va utilitzar el sistema de cromatografia líquida AKTA-FPLC (Amersham Biosciences) i es carregà una columna d'afinitat HiTrap (Amersham Biosciences) de 1 mL amb una solució de sulfat de níquel (II). A continuació s'equilibrà la columna amb el tampó A i s'hi féu passar la fracció soluble de l'homogenat, descartant la fracció no unida. La proteïna unida a la columna s'eluí mitjançant un gradient d'imidazole fins a 0.5 M, utilitzant el tampó B (500 mM imidazole, 500 mM clorur de sodi, 20mM fosfat de sodi, pH 7.4). Les fraccions eluïdes foren analitzades per electroforesi en gel desnaturalitzant de poliacrilamida.

## **C.2. Replegament de la proteïna en una columna d'afinitat a partir de cossos d'inclusió:**

Es desenvolupà una tècnica per a replegar les proteïnes agregades en la fracció insoluble basada en la cromatografia per afinitat. Donat que les histidines es poden unir a l'ió Ni<sup>2+</sup> en presència d'elevades concentracions d'agents caotròpics, es poden eliminar els contaminants i replegar la proteïna en la columna en un sol pas. A més, la disposició de les proteïnes separades unes de les altres en diferents punts de la reïna fa disminuir la concentració efectiva de la mostra i afavoreix el replegament. Aquesta tècnica va ser emprada per a purificar els cossos d'inclusió que formava la PaGS.

Vam recollir la fracció insoluble agregada i vam rentar els cossos d'inclusió amb la solució I (500 mM NaCl, 2% (p/v) Tritó X100, 20 mM Tris-HCl pH 8.0). A continuació els vam dissoldre en el tampó G (6 M clorur de guanidini, 500 mM NaCl, 5 mM imidazole, 5 mM 2-mercaptoetanol, 20 mM Tris-HCl pH 8.0) i s'incubà a temperatura ambient durant 24 hores.

La proteïna desplegada solubilitzada es va carregar en una columna d'afinitat de 5 mL HiTrap (Amersham Biosciences), prèviament carregada amb Ni<sup>2+</sup> i preequilibrada amb el tampó G, mitjançant el sistema de cromatografia líquida AKTA-FPLC (Amersham Biosciences). Es va descartar la fracció no unida a la columna i s'intercanvià el tampó G pel tampó U (6 M urea, 500 mM NaCl, 5 mM imidazole, 5 mM 2-mercaptoetanol, 20 mM Tris-HCl pH 8.0) de forma que l'agent caotròpic clorur de guanidini va ser reemplaçat per urea. El replegament de la proteïna unida a la columna es produí gràcies al gradient linial decreixent d'urea 6-0 M durant 12 hores a 0.1mL/min.

La proteïna replegada es va eluir amb un gradient creixent d'imidazole de 5 mM a 0.5 M, usant el tampó C (0.5 M NaCl, 0.5 M imidazole, 5 mM 2-mercaptoetanol, 20 mM Tris-HCl pH 8.0). Les fraccions eluïdes foren analitzades per electroforesi en gel desnaturalitzant de poliacrilamida.

### **C.3. Digestió amb trombina:**

Per a tallar la cua d'histidines, vam utilitzar l'enzim trombina, ja que la proteïna de fusió disposava d'una diana específica entre la cua N-terminal i la proteïna d'interès. Vam utilitzar el paquet Thrombin CleanCleave Kit (Sigma-Aldrich). En aquest sistema, la trombina està unida covalentment a una reïna d'agarosa. Després de 12 hores de digestió a temperatura ambient, el 85% de la proteïna havia estat tallada i la reïna podia ser eliminada per centrifugació.

Per a separar la proteïna digerida de la resta, s'afegí un nou pas de cromatografia d'afinitat. Aquesta vegada vam recollir la fracció no unida, ja que era la que havia perdut la cua d'histidines. Aquesta etapa de digestió s'emprà per a tallar la cua d'histidines del domini C-terminal de la PaGS.

### **C.4. Cromatografia de filtració en gel:**

Aquest tipus de cromatografia es basa en la separació de les molècules en funció de la seva mida. Aquesta tècnica va ser escollida com a últim pas tant en la purificació de la PaGS com en la del seu domini C-terminal.

Les fraccions eluïdes positives es van concentrar amb les unitats de filtració per centrifugació Centriprep (Millipore), es van carregar en una columna de gel filtració Superdex 200 HR10/30 (Amersham Biosciences) i la proteïna va ser eluïda en un tampó que contenia 500 mM NaCl, 1 mM DTT en 20 mM Tris-HCl pH 7.4. S'analitzaren les fraccions eluïdes per electroforesi en gel desnaturalitzat i aquelles que contenien la proteïna purificada van ser concentrades amb unitats de filtració Ultrafree (Millipore).

Per a calibrar la columna, s'utilitzà el paquet de marcadors de pes molecular específic per a cromatografia de filtració en gel (Sigma-Aldrich).

## **D. TÈCNiques D'ELECTROFORESI:**

### **D.1. Electroforesi en gel desnaturalitzant (SDS/PAGE):**

L'electroforesi discontinua en gel de poliacrilamida amb SDS (SDS/PAGE) és un dels mètodes més emprats per a analitzar barreges de proteïnes en funció dels seus pesos moleculars. La tècnica fou descrita per Laemmli (Laemmli 1970). Aquesta tècnica es basa en utilitzar dos tipus de gels, amb diferent concentració d'acrilamida i pH: el gel concentrador (stacking), a la part superior i amb porus grans, i el gel separador (running), a la part inferior. Aquestes diferències fan que les mostres s'apil·lin en estretes bandes abans de produir-se la separació durant la migració de les proteïnes en el gel separador.

L'electroforesi es desenvolupa en gels de poliàcrilamida i SDS, de grandària 10 x 8 cm i de 0.75 mm de gruix emprant l'aparell d'electroforesi *miniprotein* (Bio-Rad). Els gels utilitzats van ser de 8-10 % d'acrilamida (p/v).

Referit a la preparació de les mostres, els cultius bacterians s'homogenitzaren segons el protocol exposat a l'apartat 2 i es prengueren mostres del sobrenedant i el sediment. També es prengueren mostres de les fraccions eluïdes en cada columna cromatogràfica.

Les diferents mostres es van diluir en tampó de càrrega, la composició final del qual era Tris/HCl 50 mM a pH 6.8, DTT 10 mM, 2% (p/v) SDS, 2-mercaptoetanol 4% (v/v), glicerol 10%, blau de bromofenol 0.1% (p/v) (tinció). S'incubà durant 3-10 minuts (depenent del volum de mostra) a 100°C, per desnaturalitzar les proteïnes. Es van utilitzar els marcadors d'elevada massa molecular (Sigma), formats per una barreja de proteïnes amb massa de 205, 116, 97.4, 66, 45 i 29 kDa.

El tampó d'electroforesi fou Tris/HCl 25 mM, glicina 192 mM, SDS al 0.1%, pH 8.3. L'electroforesi es desenvolupa durant 45 minuts aproximadament amb un voltatge fix de 200V i una intensitat variable d'uns 80-30 mA.

## **D.2. Electroforesi en gel nadiu:**

En l'electroforesi en condicions nadiues, a diferència de l'anterior, les proteïnes conserven la seva estructura tridimensional. Així, les proteïnes migren en funció del seu tamany, càrrega i forma. L'electroforesi es desenvolupa en gels de 5-7.5 % d'acrilamida (p/v).

L'aparell utilitzat i les preparació de les mostres va ser equivalent a l'apartat anterior, tot i que el tampó de càrrega en aquest cas estava compost per Tris/HCl 50 mM a pH 6.8, glicerol 10%, blau de bromofenol 0.1% (p/v) (tinció). Es van utilitzar els marcadors nadius de massa molecular SeeBlue Pre-stained Standard (Invitrogen), formats per la barreja de proteïnes amb pesos: 250, 98, 64, 50, 36, 30, 16 kDa.

El tampó d'electroforesi contenia Tris/HCl 20 mM, glicina 150 mM pH 8.3. L'electroforesi es desenvolupa durant 2 hores aproximadament a un voltatge fix de 90-100V i a 4°C.

## **D.3. Tinció amb blau de Coomassie:**

La tècnica emprada per a detectar les proteïnes separades pels gels de poliàcrilamida va ser la tinció amb blau de Coomassie. Aquest mètode permet detectar fins a uns 0.5 µg de proteïna i és quantitatiu fins a uns 15 µg.

Un cop s'han fixat les proteïnes amb una solució al 10%(v/v) d'acid acètic i 20%(v/v) d'etanol, es tenyeix el gel amb una solució de blau de Coomassie al 0.2%(p/v) i 10%(v/v) d'acid acètic. Passats 10 minuts, es destenyeix amb una solució 10%(v/v) d'acid acètic.

#### **D.4. Seqüenciació de proteïnes:**

Per a comprovar que la proteïna expressada i purificada era correcta, vam seqüenciar el domini N-terminal mitjançant la degradació química d'Edman. La seqüenciació es va dur a terme a la Unitat de Proteòmica dels Serveis Científic-Tècnics (SCT) de la Universitat de Barcelona, al Parc Científic de Barcelona. L'aparell utilitzat fou el seqüenciador Procise 492cLC (Applied Biosystems).

### **E. DETERMINACIÓ DE L'ACTIVITAT ENZIMÀTICA I CONCENTRACIÓ DE PROTEÏNA:**

#### **E.1. Determinació de l'activitat enzimàtica:**

L'activitat enzimàtica fou determinada mesurant la incorporació d'UDP-[<sup>14</sup>C]-glucosa o d'ADP-[<sup>14</sup>C]-glucosa a glicogen, seguint les pautes generals de la tècnica descrita per Thomas i col·laboradors (Thomas *et al.* 1968). Aquesta radioactivitat incorporada a una molècula de glicogen creixent es mesurava en un comptador de cintil·lació líquida (LKB 1217 RACKBETA, Wallac, USA). Els resultats de l'activitat total d'aquest enzim s'expressaren en U/mg de proteïna.

#### **E.2. Determinació de la concentració de proteïna pel mètode Bradford:**

Es quantificà la proteïna total mitjançant el mètode de Bradford, basat en el canvi colorimètric del blau de Coomassie en funció de la concentració de proteïna (Bradford 1976). En una solució àcida, el blau de Coomassie canvia el màxim d'absorbància de 465 a 595 nm, quan interacciona amb proteïnes.

La corba patró es realitzà a partir d'una solució mare de BSA (albúmina de sèrum boví, Pierce) a 2 mg/ml. Els punts de la corba van correspondre a 2.5, 5, 10, 15, 20 µg de BSA. Els resultats s'expressaren en µg/µL.

### **F. TÈCNiques DE CARACTERITZACIÓ EN SOLUCIÓ:**

#### **F.1. Dicroisme circular:**

El dicroisme circular és un bon mètode per a analitzar l'estructura secundària de les proteïnes en solució, ja que les proteïnes són òpticament actives per la disposició quiral dels seus enllaços peptídics en l'espai.

Les mesures de dicroisme circular es van realitzar en l'espectropolarímetre J-810 (Jasco, Japó) del Centro Nacional de Investigaciones Oncológicas (CNIO, Madrid). Aquest aparell està equipat amb un sistema de control de temperatura Peltier (Neslab,



USA). Es va recollir l'espectre en l'UV llunyà entre 190 i 250 nm de la proteïna a concentracions de 0.1 i 0.5 mg/mL en una cel·la de 2 mm.

La desnaturalització tèrmica fins a 110°C es va seguir a 220nm, amb una velocitat d'escalfament d'1°C per minut.

## **F.2. Ultracentrifugació analítica. Mètode de l'equilibri de sedimentació:**

La tècnica d'equilibri de sedimentació és molt adequada per a determinar l'estructura quaternària de les proteïnes, determinant la massa molecular promig de les espècies presents en solució. Aquest mètode es basa en la mesura de la distribució de la concentració a la cel·la quan s'ha arribat a les condicions d'equilibri, és a dir, quan la sedimentació és compensada per la difusió (Svedberg 1940).

L'experiment d'equilibri de sedimentació es va dur a terme amb una ultracentrífuga Beckman Optima XL-A (Beckman) equipada amb un sistema d'absorbència òptica. Les dilucions de la solució de proteïna a 1, 0.5 i 0.01 mg/mL foren sotmeses a 10000 g i es recolliren espectres de les diferents absorbàncies en funció les distribucions radials en l'equilibri.

Les dades es van analitzar utilitzant el programa Origin v 6.1.

## **G. MICROSCÒPIA ELECTRÒNICA:**

S'escollí la microscòpia per tinció negativa, on la mostra es deshidrata i queda immersa en una capa de metall, la qual la protegeix de la radiació electrònica i del buit al qual està sotmesa la columna del microscopi. Atès que els àtoms del metall pesat dispersen més els electrons que els àtoms lleugers de la mostra, la imatge generada prové fonamentalment del contrast entre les zones que contenen l'àtom pesat i les que no en tenen perquè es troba la mostra biològica. Atesa la mida de la sal del metall pesat, aquesta tècnica té una resolució màxima de 15-20 Å.

Una gota de mostra de 20 µl s'adsorbí sobre una reixeta de coure coberta per una capa de carbó i activada per llum UV. Es va rentar amb dues gotes d'aigua desionitzada i es tenyí amb una gota d'acetat d'uranil al 2%(p/v). Un cop deshidratades, es van observar al microscopi electrònic de transmissió Jeol JEM 1010 (Jeol). Es van prendre imatges a 300,000x augments.

## **H. TÈCNiques DE CRISTAL·LITZACIÓ:**

Les cristal·litzacions es van dur a terme al laboratori del Dr. I. Fita, a l'Institut de Biologia Molecular de Barcelona (IBMB) del CSIC.

## H.1. Gota penjant:

S'escollí el mètode de difusió de vapor per gota penjant per a assolir la sobresaturació (Gómez-Moreno 2003).

La proteïna purificada i concentrada a 6-10  $\mu\text{g}/\mu\text{L}$  fou sotmesa a un extens rastreig de condicions inicials de cristal·lització. Es barrejà 1  $\mu\text{L}$  de proteïna amb 1  $\mu\text{L}$  de la cada solució precipitant i s'equilibrà amb 1mL de solució precipitant continguda en la cambra de cristal·lització. Com a cambres de cristal·lització, s'utilitzàren plaques de 24 pouets (Linbro) cobertes per cobreobjectes de 22x22mm tractats amb diclorometilsilà (Merck) i segellades amb greix (Dow Corning).

Com a base del rastreig, vam utilitzar entre d'altres els equips de Hampton (Crystal Screen, Crystal Screen 2, Index, Natrix, PEG-ion Screen) i de Emerald Biostructures (Wizard) a 4 i 20°C.

Els cristalls de la proteïna sencera aparegueren en la condició 39 del Crystal Screen i els del domini C-terminal en la condició 17 de Wizard I.

Per a optimitzar la qualitat dels cristalls, es féu un segon cribatge al voltant de les condicions inicials.

## H.2. Microseembra i macroseembra:

Per a millorar la qualitat i el tamany dels cristalls, s'utilitzaren tècniques de sembra segons el descrit per McRee D.E. (McRee 1993). Per a aconseguir un millor creixement dels cristalls, s'ha de reduir el nivell de sobresaturació, així com la velocitat d'aquest creixement.

En l'experiment de microseembra, es pulveritzaren cristalls amb l'ajut de boles de cristall (Sigma-Aldrich). Es prepararen succesives dilucions d'aquestes llavors i s'afegiren 0.2  $\mu\text{L}$  a diferents gotes en què s'havia variat la concentració dels agents precipitants.

En l'experiment de macroseembra, s'escolliren cristalls com a llavors i, amb l'ajut de capil·lars, es rentàren i s'afegiren a gotes fresques que contenien condicions menys saturants dels agents precipitants.

## H.3. Criopreservació:

La radiació de raigs X és molt energètica i s'ha de protegir el cristall perquè el dany per radiació no el destrueixi. Una alternativa és mantenir el cristall a 100 K durant la recollida.

S'assajaren diferents condicions de criopreservació per a garantir la integritat del cristall en ser congelat a 100 K en nitrogen líquid. A fi i efecte de minimitzar les

alteracions fisico-químiques del cristall, vam afegir l'agent criopreservant en dues etapes, i així facilitar-ne l'adaptació.

Els cristalls es recolliren en loops de 0.05 a 0.4 mm (Hampton Research) i es congelaren immediatament en nitrògen líquid.

## I. RECOLLIDA DELS ESPECTRES DE DIFRACCIÓ:

### I.1. Generador d'ànode rotatori:

El generador d'ànode rotatori està format per una cambra de buit que conté un filament que emet electrons (càtode), els quals són accelerats mitjançant una diferència de potencial fins que xoquen amb l'ànode de Cu, provocant l'emissió de raigs X a la longitud d'ona precisa per a l'experiment de difracció. Per evitar que l'ànode es fongui a causa del bombardeig d'electrons, se'l manté rotant i refrigerat (Drenth 1994).

S'utilitzà el generador Rigaku Micromax 007 (Rigaku) de la Plataforma Automatitzada de Cristal·lografia del Parc Científic de Barcelona (PAC-PCB). Aquest generador té un càtode de 0.3x3 mm i una potència de 4.5 kW, amb òptica microfocus ConBlueR2 i un detector Image Plate Mar345 (Mar). El sistema Oxford Cryosystem mantingué els cristalls a 100 K durant la recollida.

### I.2. Radiació de sincrotró:

La radiació de sincrotró és la que es desprèn en canviar la trajectòria de les partícules carregades (electrons i positrons) que circulen properes a la velocitat de la llum (Helliwell 1992). Aquestes partícules viatgen dins d'un anell que està al buit. A les zones corbes de l'anell se situen els imans corbadors (*bending magnets*) que com el seu propi nom indica, corben la trajectòria de les partícules tot generant raigs X i donen lloc a les línies BM del sincrotró. En les zones rectes de l'anell, s'intercalen els onduldadors (*insertion device magnet*), les quals produeixen una radiació de major intensitat que els imans corbadors i donen lloc a les línies ID. La longitud d'ona a estudi és filtrada pels monocromadors i els col·limadors afinen el raig (Drenth 1994).

Tots els cristalls es van difractar al sincrotró ESRF (Grenoble), amb un anell de 844 metres de perímetre i una energia de 6 GeV.

Els cristalls del domini C-terminal de la PaGS van ser difractats a la línia BM16 i es va utilitzar un detector MarCCD165. En aquesta línia també vam recollir la fluorescència del seleni en els cristalls de selenometionina, emprant el detector Rontec.

Els cristalls de PaGS nadius i els cocrystal·litzats amb UDPG es van recollir a la línia ID14-EH2, que conté un monocromador Diamond (111) Ge (220) i un detector ADSC Q4. També es van recollir en aquesta línia els cristalls incubats amb les sals de Hg i W.

L'experiment de SAD es realitzà a la línia ID13, la qual es caracteritza per una òptica microfocus i un doble monocromador de Si (111). També es van recollir cristalls incubats amb les sals de Pt i Ta.

L'experiment de MAD es dugué a terme a la línia ID29 de longitud d'ona modulable, amb un monocromador de Si (111) i un detector ADSC Q4.

### **I.3. Processament dels espectres de difracció:**

L'indexat de les primeres imatges permet determinar el grup espacial, l'orientació del cristall i la cel·la del cristall. Un cop es disposa d'aquesta informació, es pot determinar el nombre mínim de graus que s'han de recollir per a disposar d'un espectre complet. A partir de l'indexat, es poden predir els patrons de les imatges següents i així, durant la integració, es quantifica l'energia en cadascuna de les reflexions. En l'escalat, s'ajunten les reflexions repetides i s'intenten reconstruir les que estan disperses en diverses imatges a causa de la mosaïcitat.

L'indexat i la integració de les dades dels cristalls del domini C-terminal es dugué a terme amb el programa Mosfilm (Leslie 1991) i l'escalat amb el programa Scala, del paquet CCP4 (Project 1994). Pels cristalls de PaGS s'utilitzà el paquet HKL (Denzo i Scalepack) (Otwinowski i Minor 1997).

## **J. MÈTODES DE RESOLUCIÓ DE L'ESTRUCTURA:**

### **J.1. SAD, MAD, SIR, MIR:**

El reemplaçament molecular (apartat J2) acostuma a ser el mètode més directe per a determinar l'estructura d'una proteïna. Però si no disposem d'un model de partida, el problema de la fase pot ser resolt per reemplaçament isomorf i/o reemplaçament anòmal. En aquest cas, reduïm l'estructura a un nombre petit d'àtoms i intentem afrontar-lo mitjançant el mètode de Patterson o per mètodes directes (Drenth 1994). Els àtoms pesats tenen una major contribució i presenten una difracció anisotròpica a la longitud d'ona de raigs X corresponent al seu màxim d'absorció.

Els mètode SIR (*Single Isomorphous Replacement*) i MIR (*Multiple Isomorphous Replacement*) utilitzen la major contribució a la difracció per part dels àtoms pesats units a la proteïna. Requereix la difracció del cristall nadiu i, com a mínim, d'un derivat que contingui un àtom pesat, i que entre ells siguin isomorfs. Per a intentar resoldre l'estructura per àtoms pesats, vam utilitzar els programes SOLVE (Project 1994) i BnP (Weeks i Miller 1999).

Els mètodes SAD (*Single Anomalous Replacement*) i MAD (*Multiple Anomalous Signal*) eludeixen la necessitat de trobar dos cristalls isomorfs degudament derivats, ja que només s'utilitza un cristall derivat. Aquest es recull a la longitud d'ona del seu màxim

d'absorció, perquè en aquestes condicions es produeix una anisotropia en la difracció (senyal anòmal). En el mètode SAD recollim un espectre al màxim d'absorció. En el mètode MAD es recullen tres espectres del mateix cristall a tres longituds d'ona (pic, inflexió, remota). Per a intentar resoldre l'estructura per difracció anòmala a partir de cristalls de selenometionina, vam utilitzar els programes SOLVE (Project 1994), BnP (Weeks i Miller 1999) i SHELXD (Schneider i Sheldrick 2002).

## **J.2. Reemplaçament molecular:**

El mètode del reemplaçament molecular es basa en una optimització global per a localitzar en la millor posició un model d'una proteïna, de manera que en aquesta posició les intensitats calculades (per aquest model) seran el més properes possible a les observades experimentalment, per a la proteïna de la qual s'en desconeix l'estructura (Drenth 1994).

La cerca sistemàtica de possibles posicions del model dins de la cel·la unitària condueix a un problema de sis dimensions (tres rotacions i tres translacions). Aquest es redueix a un problema de tres dimensions, buscant primer les rotacions i en segon lloc les translacions.

Molrep (Vaguin i Teplyakov 1997) i AMoRe (Navaza 1994) han estat els programes de reemplaçament emprats en aquest treball.

### **J.2.1. Models:**

En general, s'accepta que per a resoldre amb èxit una estructura per reemplaçament molecular, la identitat de seqüència entre la proteïna model i la problema ha de ser igual o superior al 40%. En el nostre cas, aquesta identitat era d'un 21%, utilitzant la GS d'*Agrobacterium tumefaciens* (AtGS) com a model.

Per a resoldre el domini C-terminal de la PaGS, vam utilitzar com a model inicial per Molrep el domini C-terminal de l'AtGS (residus 271-456) convertit en una cadena de polialanines i en què s'havia eliminat l'hèlix  $\alpha$ 15 (residus 406-417).

Per a resoldre la PaGS, es dissenyà un model híbrid per Molrep format per l'estructura sencera del domini C-terminal de la PaGS i pel domini N-terminal de l'AtGS modificat. Aquestes modificacions es basen en conservar només les cadenes laterals de les zones amb seqüències idèntiques entre AtGS i PaGS (residus 1-21), eliminar les zones que clarament no són presents a PaGS (residus 50-63, 224-232, 255-267) i convertir a polialanines la resta de la cadena.

Per a resoldre la PaGS segons el mètode de les funcions de rotació i translació bloquejades, s'utilitzà el model monomèric de l'AtGS sense modificacions.

### **J.2.2. Modes normals de vibració (NM):**

S'ha demostrat que en el 50% de les proteïnes que tenen dos conformacions diferents, el moviment que les relaciona pot ser descrit mitjançant un o dos modes normals de vibració de baixa freqüència (Krebs *et al.* 2002). Aquest fet és molt beneficiós en la resolució d'una estructura, ja que es poden usar els models perturbats pels modes normals de vibració com a models inicials en el reemplaçament molecular.

Vam calcular els modes normals de vibració utilitzant la plataforma disponible a internet *eINémo* (Suhre i Sanejouand 2004). De tots els modes calculats, vam escollir els de menor energia.

### **J.2.3. Simetria no cristal·logràfica (NCS):**

La simetria no cristal·logràfica que existeix en la cel·la asimètrica causa restriccions significatives en la determinació estructural per reemplaçament molecular (McRee 1993).

Atesa l'existència de còpies múltiples d'una molècula relacionades per simetria rotacional, les solucions de la funció de rotació ordinària (*cross rotation function*) es poden restringir a les que compleixen la relació especificada per la funció d'autorotació (*locked rotation function*). Així, si s'inclou la informació de simetria no-cristal·logràfica, es pot millorar la relació senyal-soroll. De la mateixa manera, la funció de translació també es pot restringir a la relació de la funció d'autorotació (*locked translation function*).

La simetria no-cristal·logràfica per a resoldre l'estructura s'aplicà amb els programes de reemplaçament molecular Gllr (Tong i Rossmann 1997) i AMoRe (Navaza 1994).

### **J.3. Afinament per cos rígid:**

Un cop trobada la primera solució que permetia posicionar el model inicial en el cristall, es procedí a l'afinament per cos rígid amb Refmac5 (Project 1994) de les tres rotacions i les tres translacions que determinaven la seva posició.

S'afinaren els dominis per separat, en un sistema de sis cossos rígids: tres dominis N-terminals (residus 1-217, 414-437) i tres dominis C-terminals (residus 218-413).

## **K. EXTENSIÓ DE LES FASES:**

Per a calcular les fases a una major resolució, s'utilitzà el programa de modificació de densitat DM (Project 1994). Aquest programa aplica restriccions en l'espai real basades en característiques conegudes del mapa de densitat electrònica (aplanament de solvent, distribució de la densitat electrònica, esqueletonització i mitjana per simetria no-cristal·logràfica).

### K.1. Construcció de màscares:

Es construïren màscares per a definir una porció del cristall on aplicarem una operació de modificació de densitat electrònica, sigui per diferenciar l'estructura cristal·logràfica del contingut de solvent o bé per promitjar les zones de densitat electrònica relacionades en la unitat asimètrica per simetria no-cristal·logràfica. Vam dissenyar dues màscares al voltant del domini N- i C-terminal d'una molècula model amb el programa NCSMask (Project 1994). La màscara del domini N-terminal s'ajustà a una distància de 3Å i a 2Å del domini C-terminal, a la màxima resolució. Per a evitar solapaments entre les màscares en la zona de contacte entre dominis, s'utilitzà el programa Mapman (Jones i Thirup 1986). Aquest permetia ajustar la màscara del domini N-terminal mitjançant l'eliminació de les zones que interpenetraven en la màscara del domini C-terminal.

### K.2. Promitjat per simetria no-cristal·logràfica:

S'utilitzaren els operadors de rotació i translació que relacionen els dominis N- i C-terminals de les tres molècules per simetria no cristal·logràfica en la unitat asimètrica per promitjar les densitats.

### L. AFINAMENT DE L'ESTRUCTURA:

Es calcularen els primers mapes de diferència ( $2F_o - F_c$ ),  $\Phi_c$  i ( $F_o - F_c$ ),  $\Phi_c$  amb el programa FFT (Project 1994). Aquests mapes s'obtenen fent la síntesi de Fourier entre les amplituds observades ( $F_o$ ) i les calculades ( $F_c$ ) i les fases calculades ( $\Phi_c$ ). Els mapes ( $2F_o - F_c$ ),  $\Phi_c$  són més "reals", ja que el pes negatiu de  $F_c$  fa disminuir el biaix del mapa (*bias*). Els mapes ( $F_o - F_c$ ),  $\Phi_c$  ajuden a resoldre les zones conflictives dels mapes.

L'afinament es dugué a terme alternant cicles d'afinament de les posicions atòmiques i mantenint l'estereoquímica amb el programa Refmac5 amb cicles manuals de traçat usant el programa O (Jones *et al.* 1991).

La qualitat estereoquímica del model fou analitzada amb el programa Procheck (Laskowski *et al.* 1993).

### M. REPRESENTACIONS DE L'ESTRUCTURA I SUPERPOSICIONS (SHP):

Tots els dibuixos que representen les estructures, densitats electròniques i superfícies de potencial s'han realitzat amb el sistema de visualització molecular PyMOL (Delano 2002).

Per a superposar les diferents estructures, vam utilitzar el programa SHP (Stuart *et al.* 1979). Per a construir les proteïnes quimèriques amb l'angle d'obertura del model motlle, vam superposar cadascun dels dominis per separat i, a continuació, els vam unir en un únic arxiu pdb.



## BIBLIOGRAFIA

- Aloy, P. and R. B. Russell (2002). "Interrogating protein interaction networks through structural biology." *Proc Natl Acad Sci U S A* 99(9): 5896-901.
- Andersson, K. M. and S. Hovmöller (2000). "The protein content in crystals and packing coefficients in different space groups." *Acta Crystallogr D Biol Crystallogr* 56 ( Pt 7): 789-90.
- Azpiazu, I., J. Manchester, A. V. Skurat, P. J. Roach and J. C. Lawrence, Jr. (2000). "Control of glycogen synthesis is shared between glucose transport and glycogen synthase in skeletal muscle fibers." *Am J Physiol Endocrinol Metab* 278(2): E234-43.
- Bai, G., Z. J. Zhang, R. Werner, F. Q. Nuttall, A. W. Tan and E. Y. Lee (1990). "The primary structure of rat liver glycogen synthase deduced by cDNA cloning. Absence of phosphorylation sites 1a and 1b." *J Biol Chem* 265(14): 7843-8.
- Barford, D. and L. N. Johnson (1989). "The allosteric transition of glycogen phosphorylase." *Nature* 340(6235): 609-16.
- Barford, D., S. H. Hu and L. N. Johnson (1991). "Structural mechanism for glycogen phosphorylase control by phosphorylation and AMP." *J Mol Biol* 218(1): 233-60.
- Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." *Anal Biochem* 72: 248-54.
- Browner, M. F., K. Nakano, A. G. Bang and R. J. Fletterick (1989). "Human muscle glycogen synthase cDNA sequence: a negatively charged protein with an asymmetric charge distribution." *Proc Natl Acad Sci U S A* 86(5): 1443-7.
- Buleon, A., P. Colonna, V. Planchot and S. Ball (1998). "Starch granules: structure and biosynthesis." *Int J Biol Macromol* 23(2): 85-112.
- Burling, F. T., R. Kniewel, J. A. Buglino, T. Chadha, A. Beckwith and C. D. Lima (2003). "Structure of Escherichia coli uridine phosphorylase at 2.0 Å." *Acta Crystallogr D Biol Crystallogr* 59(Pt 1): 73-6.
- Buschiazzo, A., J. E. Ugalde, M. E. Guerin, W. Shepard, R. A. Ugalde and P. M. Alzari (2004). "Crystal structure of glycogen synthase: homologous enzymes catalyze glycogen synthesis and degradation." *Embo J* 23(16): 3196-205.
- Camici, M., A. A. DePaoli-Roach and P. J. Roach (1984). "Rabbit liver glycogen synthase. Purification and comparison of the properties of glucose-6-P-dependent and glucose-6-P-independent forms of the enzyme." *J Biol Chem* 259(6): 3429-34.

Campbell, J. A., G. J. Davies, V. Bulone and B. Henrissat (1997). "A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities." *Biochem J* 326 ( Pt 3): 929-39.

Campbell, J. A., G. J. Davies, V. V. Bulone and B. Henrissat (1998). "A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities." *Biochem J* 329 (Pt 3): 719.

Campbell, R. E., S. C. Mosimann, M. E. Tanner and N. C. Strynadka (2000). "The structure of UDP-N-acetylglucosamine 2-epimerase reveals homology to phosphoglycosyl transferases." *Biochemistry* 39(49): 14993-5001.

Cao, Y., A. M. Mahrenholz, A. A. DePaoli-Roach and P. J. Roach (1993). "Characterization of rabbit skeletal muscle glycogenin. Tyrosine 194 is essential for function." *J Biol Chem* 268(20): 14687-93.

Cardona, S., F. Remonsellez, N. Guiliani and C. A. Jerez (2001). "The glycogen-bound polyphosphate kinase from *Sulfolobus acidocaldarius* is actually a glycogen synthase." *Appl Environ Microbiol* 67(10): 4773-80.

Charnock, S. J. and G. J. Davies (1999). "Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms." *Biochemistry* 38(20): 6380-5.

Chiu, C. P., A. G. Watts, L. L. Lairson, M. Gilbert, D. Lim, W. W. Wakarchuk, S. G. Withers and N. C. Strynadka (2004). "Structural analysis of the sialyltransferase CstII from *Campylobacter jejuni* in complex with a substrate analog." *Nat Struct Mol Biol* 11(2): 163-70.

Cid, E., R. R. Gomis, R. A. Geremia, J. J. Guinovart and J. C. Ferrer (2000). "Identification of two essential glutamic acid residues in glycogen synthase." *J Biol Chem* 275(43): 33614-21.

Cid, E., R. A. Geremia, J. J. Guinovart and J. C. Ferrer (2002). "Glycogen synthase: towards a minimum catalytic unit?" *FEBS Lett* 528(1-3): 5-11.

Cid, E., Cifuentes, D., Baqué, S., Ferrer, J.C., Guinovart, J.J. (2005). "Determinants of the nucleocytoplasmic shuttling of muscle glycogen synthase." *FEBS Journal* in press.

Cohen, P. (1978). *Curr. Top. Cell. Regul.* 14: 117-196.

Cohen, P. and D. G. Hardie (1991). "The actions of cyclic AMP on biosynthetic processes are mediated indirectly by cyclic AMP-dependent protein kinase." *Biochim Biophys Acta* 1094(3): 292-9.

Cormack, R. S., Somssich, I. E. (1997). *Tech. Tips Online*.

- Coutinho, P. M., E. Deleury, G. J. Davies and B. Henrissat (2003). "An evolving hierarchical family classification for glycosyltransferases." *J Mol Biol* 328(2): 307-17.
- Davies, G., Withers, S.G., Sinnott, M.L. (1997). Glycogen Transfer. *Comprehensive Biological Catalysis*. I. Academic Press.
- Delano, W. (2002). "The PyMOL molecular graphic system."
- Drenth, J. (1994). *Principles of Protein X-ray Crystallography*.
- Drochmans, P. (1962). "[Morphology of glycogen. Electron microscopic study of the negative stains of particulate glycogen.]" *J Ultrastruct Res* 6: 141-63.
- Dwek, R. A. (1996). "Glycobiology: Toward Understanding the Function of Sugars." *Chem Rev* 96(2): 683-720.
- Farkas, I., T. A. Hardy, M. G. Goebel and P. J. Roach (1991). "Two glycogen synthase isoforms in *Saccharomyces cerevisiae* are coded by distinct genes that are differentially controlled." *J Biol Chem* 266(24): 15602-7.
- Fernandez-Novell, J. M., J. Arino, S. Vilaro, D. Bellido and J. J. Guinovart (1992). "Role of glucose 6-phosphate in the translocation of glycogen synthase in rat hepatocytes." *Biochem J* 288 ( Pt 2): 497-501.
- Fernandez-Novell, J. M., J. Arino and J. J. Guinovart (1994). "Effects of glucose on the activation and translocation of glycogen synthase in diabetic rat hepatocytes." *Eur J Biochem* 226(2): 665-71.
- Fernandez-Novell, J. M., D. Bellido, S. Vilaro and J. J. Guinovart (1997). "Glucose induces the translocation of glycogen synthase to the cell cortex in rat hepatocytes." *Biochem J* 321 ( Pt 1): 227-31.
- Ferrer, J. C., S. Baque and J. J. Guinovart (1997). "Muscle glycogen synthase translocates from the cell nucleus to the cytosol in response to glucose." *FEBS Lett* 415(3): 249-52.
- Fiol, C. J., A. Wang, R. W. Roeske and P. J. Roach (1990). "Ordered multisite protein phosphorylation. Analysis of glycogen synthase kinase 3 action using model peptide substrates." *J Biol Chem* 265(11): 6061-5.
- Fletterick, R. J., J. Sygusch, N. Murray and N. B. Madsen (1976). "Low-resolution structure of the glycogen phosphorylase alpha monomer and comparison with phosphorylase beta." *J Mol Biol* 103(1): 1-13.

Fritz, T. A., J. H. Hurley, L. B. Trinh, J. Shiloach and L. A. Tabak (2004). "The beginnings of mucin biosynthesis: the crystal structure of UDP-GalNAc:polypeptide alpha-N-acetylgalactosaminyltransferase-T1." *Proc Natl Acad Sci U S A* 101(43): 15307-12.

Furukawa, K., M. Tagaya, K. Tanizawa and T. Fukui (1993). "Role of the conserved Lys-X-Gly-Gly sequence at the ADP-glucose-binding site in Escherichia coli glycogen synthase." *J Biol Chem* 268(32): 23837-42.

Garcia-Rocha, M., A. Roca, N. De La Iglesia, O. Baba, J. M. Fernandez-Novell, J. C. Ferrer and J. J. Guinovart (2001). "Intracellular distribution of glycogen synthase and glycogen in primary cultured rat hepatocytes." *Biochem J* 357(Pt 1): 17-24.

Gastinel, L. N., C. Cambillau and Y. Bourne (1999). "Crystal structures of the bovine beta4galactosyltransferase catalytic domain and its complex with uridine diphosphogalactose." *Embo J* 18(13): 3546-57.

Gastinel, L. N., C. Bignon, A. K. Misra, O. Hindsgaul, J. H. Shaper and D. H. Joziase (2001). "Bovine alpha1,3-galactosyltransferase catalytic domain structure and its relationship with ABO histo-blood group and glycosphingolipid glycosyltransferases." *Embo J* 20(4): 638-49.

Geremia, S., M. Campagnolo, R. Schinzel and L. N. Johnson (2002). "Enzymatic catalysis in crystals of Escherichia coli maltodextrin phosphorylase." *J Mol Biol* 322(2): 413-23.

Gibbons, B. J., P. J. Roach and T. D. Hurley (2002). "Crystal structure of the autocatalytic initiator of glycogen biosynthesis, glycogenin." *J Mol Biol* 319(2): 463-77.

Gibson, R. P., J. P. Turkenburg, S. J. Charnock, R. Lloyd and G. J. Davies (2002). "Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA." *Chem Biol* 9(12): 1337-46.

Gibson, R. P., C. A. Tarling, S. Roberts, S. G. Withers and G. J. Davies (2004). "The donor subsite of trehalose-6-phosphate synthase: binary complexes with UDP-glucose and UDP-2-deoxy-2-fluoro-glucose at 2 Å resolution." *J Biol Chem* 279(3): 1950-5.

Goldsmith, E. J., S. R. Sprang, R. Hamlin, N. H. Xuong and R. J. Fletterick (1989). "Domain separation in the activation of glycogen phosphorylase a." *Science* 245(4917): 528-32.

Gomis, R. R., E. Cid, M. Garcia-Rocha, J. C. Ferrer and J. J. Guinovart (2002). "Liver glycogen synthase but not the muscle isoform differentiates between glucose 6-phosphate produced by glucokinase or hexokinase." *J Biol Chem* 277(26): 23246-52.

Gómez-Moreno, C., Sancho, J. (2003). Estructura de proteínas. Ariel Ciencia.

- Grimsby, J., R. Sarabu, W. L. Corbett, N. E. Haynes, F. T. Bizzarro, J. W. Coffey, K. R. Guertin, D. W. Hilliard, R. F. Kester, P. E. Mahaney, L. Marcus, L. Qi, C. L. Spence, J. Teng, M. A. Magnuson, C. A. Chu, M. T. Dvorozniak, F. M. Matschinsky and J. F. Grippo (2003). "Allosteric activators of glucokinase: potential role in diabetes therapy." *Science* 301(5631): 370-3.
- Gruyer, S., E. Legin, C. Bliard, S. Ball and F. Duchiron (2002). "The endopolysaccharide metabolism of the hyperthermophilic archeon *Thermococcus hydrothermalis*: polymer structure and biosynthesis." *Curr Microbiol* 44(3): 206-11.
- Ha, S., D. Walker, Y. Shi and S. Walker (2000). "The 1.9 Å crystal structure of *Escherichia coli* MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis." *Protein Sci* 9(6): 1045-52.
- Hanashiro, I. and P. J. Roach (2002). "Mutations of muscle glycogen synthase that disable activation by glucose 6-phosphate." *Arch Biochem Biophys* 397(2): 286-92.
- Hardy, T. A. and P. J. Roach (1993). "Control of yeast glycogen synthase-2 by COOH-terminal phosphorylation." *J Biol Chem* 268(32): 23799-805.
- Helliwell, J. R. (1992). *Macromolecular Crystallography with Synchrotron Radiation*.
- Henrissat, B., E. Deleury and P. M. Coutinho (2002). "Glycogen metabolism loss: a common marker of parasitic behaviour in bacteria?" *Trends Genet* 18(9): 437-40.
- Holm, L. and C. Sander (1997). "Dali/FSSP classification of three-dimensional protein folds." *Nucleic Acids Res* 25(1): 231-4.
- Horcajada, C., E. Cid, J. J. Guinovart, N. Verdaguer and J. C. Ferrer (2003). "Crystallization and preliminary X-ray analysis of the glycogen synthase from *Pyrococcus abyssi*." *Acta Crystallogr D Biol Crystallogr* 59(Pt 12): 2322-4.
- Hu, Y., L. Chen, S. Ha, B. Gross, B. Falcone, D. Walker, M. Mokhtarzadeh and S. Walker (2003). "Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases." *Proc Natl Acad Sci U S A* 100(3): 845-9.
- Huang, K. P. and E. Cabib (1974). "Yeast glycogen synthetase in the glucose 6-phosphate-dependent form. II. The effect of proteolysis." *J Biol Chem* 249(12): 3858-61.
- Johnson, L. N., J. Cheetham, P. J. McLaughlin, K. R. Acharya, D. Barford and D. C. Phillips (1988). "Protein-oligosaccharide interactions: lysozyme, phosphorylase, amylases." *Curr Top Microbiol Immunol* 139: 81-134.

Johnson, L. N., K. R. Acharya, M. D. Jordan and P. J. McLaughlin (1990). "Refined crystal structure of the phosphorylase-heptulose 2-phosphate-oligosaccharide-AMP complex." *J Mol Biol* 211(3): 645-61.

Jones, T. A. and S. Thirup (1986). "Using known substructures in protein model building and crystallography." *Embo J* 5(4): 819-22.

Jones, T. A., J. Y. Zou, S. W. Cowtan and M. Kjeldgaard (1991). "Improved methods for building protein models in electron density maps and the location of errors in these maps." *Acta Crystallographica Section A* 47: 110-119.

Kamata, K., M. Mitsuya, T. Nishimura, J. Eiki and Y. Nagata (2004). "Structural basis for allosteric regulation of the monomeric allosteric enzyme human glucokinase." *Structure (Camb)* 12(3): 429-38.

Kaslow, H. R., D. D. Lesikar, D. Antwi and A. W. Tan (1985). "L-type glycogen synthase. Tissue distribution and electrophoretic mobility." *J Biol Chem* 260(18): 9953-6.

Katz, J., S. Golden and P. A. Wals (1979). "Glycogen synthesis by rat hepatocytes." *Biochem J* 180(2): 389-402.

König, G., R. Skorko, W. Zillig and W. D. Reiter (1982). "Glycogen in thermoacidophilic archaebacteria of the genera *Sulfolobus*, *Thermoproteus*, *Desulfurococcus* and *Thermococcus*." *Arch. Microbiol.* 132: 297-303.

Krebs, W. G., V. Alexandrov, C. A. Wilson, N. Echols, H. Yu and M. Gerstein (2002). "Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic." *Proteins* 48(4): 682-95.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." *Nature* 227(5259): 680-5.

Lairson, L. L., C. P. Chiu, H. D. Ly, S. He, W. W. Wakarchuk, N. C. Strynadka and S. G. Withers (2004). "Intermediate trapping on a mutant retaining alpha-galactosyltransferase identifies an unexpected aspartate residue." *J Biol Chem* 279(27): 28339-44.

Laskowski, R. A., D. S. Moss and J. M. Thornton (1993). "Main-chain bond lengths and bond angles in protein structures." *J Mol Biol* 231(4): 1049-67.

Leslie, A. (1991). *Crystallographic Computing V*. Oxford University Press.

Ly, H. D., B. Loughheed, W. W. Wakarchuk and S. G. Withers (2002). "Mechanistic studies of a retaining alpha-galactosyltransferase from *Neisseria meningitidis*." *Biochemistry* 41(16): 5075-85.

- Maitra, P. K., S. B. Bhosale, D. C. Kshirsagar, T. Y. Yeole and A. N. Shanbhag (2001). "Metabolite and enzyme profiles of glycogen metabolism in *Methanococcoides methylutens*." *FEMS Microbiol Lett* 198(1): 23-9.
- Manchester, J., A. V. Skurat, P. Roach, S. D. Hauschka and J. C. Lawrence, Jr. (1996). "Increased glycogen accumulation in transgenic mice overexpressing glycogen synthase in skeletal muscle." *Proc Natl Acad Sci U S A* 93(20): 10707-11.
- Matthews, B. W. (1968). "Solvent content of protein crystals." *J Mol Biol* 33(2): 491-7.
- McRee, D. E. (1993). *Practical Protein Crystallography*. Academic Press, Inc.
- McVerry, P. and K. H. Kim (1972). "Dephosphorylation and aggregation of rat liver glycogen synthetase." *Biochem Biophys Res Commun* 48(6): 1636-40.
- McVerry, P. H. and K. H. Kim (1974). "Purification and kinetic mechanism of rat liver glycogen synthase." *Biochemistry* 13(17): 3505-11.
- Melendez, R., E. Melendez-Hevia, F. Mas, J. Mach and M. Cascante (1998). "Physical constraints in the synthesis of glycogen that influence its structural homogeneity: a two-dimensional approach." *Biophys J* 75(1): 106-14.
- Mitchell, E. P., S. G. Withers, P. Ermert, A. T. Vasella, E. F. Garman, N. G. Oikonomakos and L. N. Johnson (1996). "Ternary complex crystal structures of glycogen phosphorylase with the transition state analogue nojirimycin tetrazole and phosphate in the T and R states." *Biochemistry* 35(23): 7341-55.
- Mulichak, A. M., H. C. Losey, C. T. Walsh and R. M. Garavito (2001). "Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics." *Structure (Camb)* 9(7): 547-57.
- Mulichak, A. M., H. C. Losey, W. Lu, Z. Wawrzak, C. T. Walsh and R. M. Garavito (2003). "Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway." *Proc Natl Acad Sci U S A* 100(16): 9238-43.
- Mulichak, A. M., W. Lu, H. C. Losey, C. T. Walsh and R. M. Garavito (2004). "Crystal structure of vancosaminyltransferase GtfD from the vancomycin biosynthetic pathway: interactions with acceptor and nucleotide ligands." *Biochemistry* 43(18): 5170-80.
- Navaza, J. (1994). "AMoRe: an automated package for molecular replacement." *Acta Crystallographica Section A* 50: 157-163.
- Navaza, J. (2001). "Implementation of molecular replacement in AMoRe." *Acta Crystallogr D Biol Crystallogr* 57(Pt 10): 1367-72.

Newgard, C. B., M. J. Brady, R. M. O'Doherty and A. R. Saltiel (2000). "Organizing glucose disposal: emerging roles of the glycogen targeting subunits of protein phosphatase-1." *Diabetes* 49(12): 1967-77.

Nichols, D. J., P. L. Keeling, M. Spalding and H. Guan (2000). "Involvement of conserved aspartate and glutamate residues in the catalysis and substrate binding of maize starch synthase." *Biochemistry* 39(26): 7820-5.

Nuttall, F. Q., M. C. Gannon, G. Bai and E. Y. Lee (1994). "Primary structure of human liver glycogen synthase deduced by cDNA cloning." *Arch Biochem Biophys* 311(2): 443-9.

Orho, M., N. U. Bosshard, N. R. Buist, R. Gitzelmann, A. Aynsley-Green, P. Blumel, M. C. Gannon, F. Q. Nuttall and L. C. Groop (1998). "Mutations in the liver glycogen synthase gene in children with hypoglycemia due to glycogen storage disease type 0." *J Clin Invest* 102(3): 507-15.

Otwinowski, Z. and W. Minor (1997). *Methods in Enzymology* 276: 307-326.

Patenaude, S. I., N. O. Seto, S. N. Borisova, A. Szpacenko, S. L. Marcus, M. M. Palcic and S. V. Evans (2002). "The structural basis for specificity in human ABO(H) blood group biosynthesis." *Nat Struct Biol* 9(9): 685-90.

Pedersen, L. C., K. Tsuchida, H. Kitagawa, K. Sugahara, T. A. Darden and M. Negishi (2000). "Heparan/chondroitin sulfate biosynthesis. Structure and mechanism of human glucuronyltransferase I." *J Biol Chem* 275(44): 34580-5.

Pedersen, L. C., J. Dong, F. Taniguchi, H. Kitagawa, J. M. Krahn, L. G. Pedersen, K. Sugahara and M. Negishi (2003). "Crystal structure of an alpha 1,4-N-acetylhexosaminyltransferase (EXTL2), a member of the exostosin gene family involved in heparan sulfate biosynthesis." *J Biol Chem* 278(16): 14420-8.

Pederson, B. A., C. Cheng, W. A. Wilson and P. J. Roach (2000). "Regulation of glycogen synthase. Identification of residues involved in regulation by the allosteric ligand glucose-6-P and by phosphorylation." *J Biol Chem* 275(36): 27753-61.

Persson, K., H. D. Ly, M. Dieckelmann, W. W. Wakarchuk, S. G. Withers and N. C. Strynadka (2001). "Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs." *Nat Struct Biol* 8(2): 166-75.

Plesner, L., I. W. Plesner and V. Esmann (1974). "Kinetic mechanism of glycogen synthase D from human polymorphonuclear leukocytes." *J Biol Chem* 249(4): 1119-25.



- Prats, C., J. A. Cadefau, R. Cusso, K. Qvortrup, J. N. Nielsen, J. F. Wojtaszewki, D. G. Hardie, G. Stewart, B. F. Hansen and T. Ploug (2005). "Phosphorylation-dependent translocation of glycogen synthase to a novel structure during glycogen re-synthesis." *J Biol Chem*.
- Preiss, J. and T. Romeo (1989). "Physiology, biochemistry and genetics of bacterial glycogen synthesis." *Adv Microb Physiol* 30: 183-238.
- Project, C. C. (1994). "The CCP4 suite: programs for protein crystallography." *Acta Crystallogr D Biol Crystallogr* 50(Pt 5): 760-3.
- Rath, V. L., M. Ammirati, P. K. LeMotte, K. F. Fennell, M. N. Mansour, D. E. Danley, T. R. Hynes, G. K. Schulte, D. J. Wasilko and J. Pandit (2000). "Activation of human liver glycogen phosphorylase by alteration of the secondary structure and packing of the catalytic core." *Mol Cell* 6(1): 139-48.
- Roach, P. J. (1991). "Multisite and hierarchal protein phosphorylation." *J Biol Chem* 266(22): 14139-42.
- Roach, P. J., C. Cheng, D. Huang, A. Lin, J. Mu, A. V. Skurat, W. Wilson and L. Zhai (1998). "Novel aspects of the regulation of glycogen storage." *J Basic Clin Physiol Pharmacol* 9(2-4): 139-51.
- Rybicka, K. K. (1996). "Glycosomes--the organelles of glycogen metabolism." *Tissue Cell* 28(3): 253-65.
- Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis and H. A. Erlich (1988). "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase." *Science* 239(4839): 487-91.
- Sambrook, J., Russell, D.W. (2001). *Molecular cloning*, Cold Spring Harbor Laboratory Press.
- Schinzel, R. and P. Drueckes (1991). "The phosphate recognition site of Escherichia coli maltodextrin phosphorylase." *FEBS Lett* 286(1-2): 125-8.
- Schneider, T. R. and G. M. Sheldrick (2002). "Substructure solution with SHELXD." *Acta Crystallogr D Biol Crystallogr* 58(Pt 10 Pt 2): 1772-9.
- Sillje, H. H., J. W. Paalman, E. G. ter Schure, S. Q. Olsthoorn, A. J. Verkleij, J. Boonstra and C. T. Verrips (1999). "Function of trehalose and glycogen in cell cycle progression and cell viability in *Saccharomyces cerevisiae*." *J Bacteriol* 181(2): 396-400.
- Sinnot, M. L. (1990). "Catalytic mechanism of enzymatic glycosyl transfer." *Chem. Rev.* 90: 1171-1202.

Skurat, A. V., Y. Wang and P. J. Roach (1994). "Rabbit skeletal muscle glycogen synthase expressed in COS cells. Identification of regulatory phosphorylation sites." *J Biol Chem* 269(41): 25534-42.

Skurat, A. V. and P. J. Roach (1995). "Phosphorylation of sites 3a and 3b (Ser640 and Ser644) in the control of rabbit muscle glycogen synthase." *J Biol Chem* 270(21): 12491-7.

Stuart, D. I., Levine M., Murlhead, H. and D. K. Stammers (1979). "Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 Å." *J Mol Biol* 134: 109-142.

Studier, F. W. and B. A. Moffatt (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." *J Mol Biol* 189(1): 113-30.

Suhre, K. and Y. H. Sanejouand (2004). "EINemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement." *Nucleic Acids Res* 32(Web Server issue): W610-4.

Svedberg, T., Pederson, K.O. (1940). "The Ultracentrifuge." Oxford and Charendon Press: 5.

Takahara, H. and K. Matsuda (1978). "Biosynthesis of glycogen in *Neurospora crassa*. Purification and properties of the UDPglucose:glycogen 4-alpha-glucosyltransferase." *Biochim Biophys Acta* 522(2): 363-74.

Takeda, Y. and J. Larner (1975). "Structural studies on rabbit muscle glycogen synthase. II. Limited proteolysis." *J Biol Chem* 250(23): 8951-6.

Teplyakov, A., Vainstein, B. (1990). *Sov. Phys. Crystallogr.* 35(3): 414-418.

Thomas, J. A., K. K. Schlender and J. Larner (1968). "A rapid filter paper assay for UDPglucose-glycogen glucosyltransferase, including an improved biosynthesis of UDP-14C-glucose." *Anal Biochem* 25(1): 486-99.

Tong, L. and M. G. Rossmann (1997). "Rotation function calculations with GLRF program." *Methods Enzymol* 276: 594-611.

Ugalde, J. E., A. J. Parodi and R. A. Ugalde (2003). "De novo synthesis of bacterial glycogen: *Agrobacterium tumefaciens* glycogen synthase is involved in glucan initiation and elongation." *Proc Natl Acad Sci U S A* 100(19): 10659-63.

Unligil, U. M., S. Zhou, S. Yuwaraj, M. Sarkar, H. Schachter and J. M. Rini (2000). "X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily." *Embo J* 19(20): 5269-80.

- Vaguin, A. and A. Teplyakov (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* 30: 1022-1025.
- Van Duyne, G. D., R. F. Standaert, P. A. Karplus, S. L. Schreiber and J. Clardy (1993). "Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin." *J Mol Biol* 229(1): 105-24.
- Villar-Palasi, C. and J. J. Guinovart (1997). "The role of glucose 6-phosphate in the control of glycogen synthase." *Faseb J* 11(7): 544-58.
- Vocadlo, D. J., G. J. Davies, R. Laine and S. G. Withers (2001). "Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate." *Nature* 412(6849): 835-8.
- Vrielink, A., W. Ruger, H. P. Driessen and P. S. Freemont (1994). "Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose." *Embo J* 13(15): 3413-22.
- Warram, J. H., B. C. Martin, A. S. Krolewski, J. S. Soeldner and C. R. Kahn (1990). "Slow glucose removal rate and hyperinsulinemia precede the development of type II diabetes in the offspring of diabetic parents." *Ann Intern Med* 113(12): 909-15.
- Watson, K. A., R. Schinzel, D. Palm and L. N. Johnson (1997). "The crystal structure of Escherichia coli maltodextrin phosphorylase provides an explanation for the activity without control in this basic archetype of a phosphorylase." *Embo J* 16(1): 1-14.
- Watson, K. A., C. McCleverty, S. Geremia, S. Cottaz, H. Driguez and L. N. Johnson (1999). "Phosphorylase recognition and phosphorolysis of its oligosaccharide substrate: answers to a long outstanding question." *Embo J* 18(17): 4619-32.
- Weeks, C. M. and R. Miller (1999). "Optimizing Shake-and-Bake for proteins." *Acta Crystallogr D Biol Crystallogr* 55 ( Pt 2): 492-500.
- Withers, S. G., N. B. Madsen, B. D. Sykes, M. Takagi, S. Shimomura and T. Fukui (1981). "Evidence for direct phosphate-phosphate interaction between pyridoxal phosphate and substrate in the glycogen phosphorylase catalytic mechanism." *J Biol Chem* 256(21): 10759-62.
- Woese, C. R., O. Kandler and M. L. Wheelis (1990). "Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya." *Proc Natl Acad Sci U S A* 87(12): 4576-9.
- Wrabl, J. O. and N. V. Grishin (2001). "Homology between O-linked GlcNAc transferases and proteins of the glycogen phosphorylase superfamily." *J Mol Biol* 314(3): 365-74.

Yep, A., M. A. Ballicora and J. Preiss (2004). "The active site of the Escherichia coli glycogen synthase is similar to the active site of retaining GT-B glycosyltransferases." *Biochem Biophys Res Commun* 316(3): 960-6.

Zea, C. J., S. W. MacDonell and N. L. Pohl (2003). "Discovery of the archaeal chemical link between glycogen (starch) synthase families using a new mass spectrometry assay." *J Am Chem Soc* 125(45): 13666-7.

Zechel, D. L. and S. G. Withers (2000). "Glycosidase mechanisms: anatomy of a finely tuned catalyst." *Acc Chem Res* 33(1): 11-8.