Reacción Combinada de la Amilosacarasa y la Ciclodextrina Glucosiltransferasa para la producción de Ciclodextrinas a Partir de Sacarosa y Lixiviados de Mucílago de Cacao

Leidy Rocío Pico Martínez
Bióloga

Trabajo de Grado para Optar al Título de Magister en Biología

Director del proyecto:

Jorge Hernández Torres
Biólogo, Ph.D
Escuela de Biología

Codirector del proyecto:

Hermínsul de Jesús Cano Calle
Químico, Ph.D
Escuela de Química

Universidad Industrial de Santander
Facultad de Ciencias
Escuela de Biología
Maestría en Biología
Bucaramanga
2021

Agradecimientos

Un agradecimiento especial al proyecto “Desarrollo de nuevos procesos y productos para la valorización de mucílagos y granos de cacao en el Departamento de Santander – BPIN 2016000100046”, financiado con recursos del Fondo de Ciencia Tecnología e Innovación FCTel, el Sistema General de Regalías -SGR, y a la Universidad Industrial de Santander.

A mi tutor, el profesor Jorge Hernández, por brindarme la oportunidad de pertenecer al grupo de investigación CINBIN por más de 8 años; por su apoyo y acompañamiento incondicional durante la elaboración de este proyecto y vida académica. Gracias por seguir apoyándome en mi camino como investigadora, por la confianza y por siempre inspirarme a superar mis límites.

Al profesor Hermínsul Cano Calle, por confiar en mis capacidades y brindarme la oportunidad de pertenecer a este proyecto de investigación.

Al profesor Cristian Blanco y a la profesora Marianny Yajaira Combariza, por la confianza impuesta en mi al brindarme la oportunidad de ejecutar este proyecto y de pertenecer al grupo de investigación GIFTEX.

Un agradecimiento especial a los grupos de investigación UIS y a sus integrantes, por la colaboración con el uso de equipos y servicios brindados, los cuales fueron esenciales para el desarrollo de este proyecto: Laboratorio de Actividades Biológicas; SC3 UIS – grupo de Supercomputación y Cálculo Científico; Centro de Investigaciones en Enfermedades Infecciosas – CINTROP; Grupo de Investigación en Fisicoquímica Teórica y Experimental – GIFTEX y Laboratorio de Espectrometría de Masas.

A mi familia, por su apoyo constante y amor. A mi madre por su entrega incondicional, esfuerzo, dedicación y por siempre creer en mí. A mi padre y mi hermana por su valiosa compañía, amor y apoyo.

A John Alexander Soracá, mi compañero de vida, por alentarme y ayudarme a iniciar esta meta que hoy veo materializada. Gracias por tu amor, por construir sueños conmigo y por estar siempre aquí siendo mi soporte incondicional. Gracias por animarme a seguir adelante ante las adversidades y obstáculos. Fuiste fundamental durante este proceso de aprendizaje.

A Elizabeth Quintero, mi colega y amiga por más de 10 años, por escucharme y compartir conmigo momentos únicos. Gracias por tu apoyo incondicional en mi vida y por tu gran espíritu colaborativo y siempre estar dispuesta a ayudarme. Especialmente, gracias por tu apoyo con el préstamo de los equipos de laboratorio, los cuales fueron indispensables para la ejecución de este proyecto.

A Juan David Bayona, mi colega y amigo, por escucharme siempre, porque aún en la distancia las risas, la amistad y el apoyo incondicional se sintieron más cerca que nunca. Gracias por tus consejos y asesoría constante durante el desarrollo de este proyecto. Gracias por ser mi editor de tesis y de vida.

A Nathalie Ñañez y Jhonatan Carreño, por su valiosa amistad. Por siempre recibirme con tanto amor y cariño. Cada palabra y momento compartido fueron valiosos y me ayudaron a sobrellevar este proceso.

A mis amigos y compañeros de Maestría, Paulin Polanco, David Rueda, Nubia Castro, Carlos Pedraza y Sergio Bolívar, por las reuniones con café, las charlas y momentos compartidos. Gracias a ustedes este recorrido se hizo más divertido y ameno.

A Schrödinger y Leia, por brindarme paz y tranquilidad en los momentos de crisis. Por seleccionarme para hacer parte de sus vidas gatunas… por amarme cada uno a su manera.

Tabla de contenido

Introducción 13

1. Objetivos 18

1.2 Objetivo General 18

1.3 Objetivos Específicos 18

2. Materiales Y Métodos 19

2.1 Diseño *in silico* de las enzimas recombinantes AS, CGTase y fusión (AS·CGTase). 19

2.1.1 Diseño de las estructuras primarias de las proteínas AS, CGTase y AS·CGTase. 19

2.1.2 Modelado 3D de la proteína de fusión AS·CGTase. 19

2.1.3 Evaluación de la estabilidad estructural de los modelos mediante dinámica molecular 19

2.1.4 Diseño de los vectores recombinantes pET–GST-AS-His, pET–DacD-cgt-His y pET–DacD-AS-cgt-His. 20

2.2 Sobrexpresión en *Escherichia coli* de los polipéptidos GST·AS·His, DacD·CGTasa·His y DacD·AS·CGTasa·His. 21

2.2.1 Transformación de *E. coli*. 21

2.2.2 Control de fidelidad de los constructos. 22

2.2.3 Expresión heteróloga de las proteínas recombinantes. 22

2.2.4 Solubilización de Proteínas 23

2.3 Ensayos de Actividad 23

2.3.1 Ensayo de actividad AS: determinación de azúcares reductores 23

2.3.2 Ensayo de actividad CGTasa: Producción de CDs a partir de almidón 24

2.3.3 Determinación del pH óptimo 25

2.3.4 Producción de CDs por reacción combinada de GST·AS·His y DacD·CGTasa·His 25

2.3.5 Producción de CDs por DacD·AS·CGTasa·His a partir de sacarosa 25

2.3.6 Producción de CDs a partir de lixiviados de mucílago de cacao 26

2.3.7 Determinación de la concentración de proteínas 26

2.3.8 Cálculo de las actividades enzimáticas 26

2.3.9 Electroforesis Desnaturalizante de Proteínas (SDS-PAGE) 27

2.4 Análisis estadístico 27

2.5 Detección de CDs por espectrometría de masas (MALDI-TOF) 28

3. Resultados 28

3.1 Diseño *in silico* de los polipéptidos GST·AS, DacD·CGTasa y DacD·AS·CGTasa. 28

3.1.1 Selección de estructuras de las proteínas AS y CGTasa 28

3.1.2 Selección de *linkers* para la fusión de ORFs 30

3.1.3 Modelamiento de estructuras 3D y elección de la estructura primaria de la fusión 30

3.1.4 Evaluación *in silico* de la estabilidad de las fusiones mediante Dinámica Molecular 41

3.1.5 Diseño de los vectores de expresión recombinantes. 45

3.1.6 Verificación de la calidad de los constructos. 48

3.2 Sobreexpresión de las enzimas recombinantes en *E. coli* 50

3.2.1 Expresión en *E. coli* del vector pET-GST-AS-His 51

3.2.2 Expresión en *E. coli* de pET-DacD-cgt-His 54

3.2.3 Sobreexpresión de pET-DacD-AS-cgt-His y actividades AS y CGTasa 58

3.2.4 Visualización de las proteínas recombinantes en geles SDS-PAGE 63

3.3 Producción de CDs en una sola reacción a partir de sacarosa como sustrato 64

3.4 Producción de CDs en una sola reacción a partir de lixiviados de mucílago de cacao como sustrato. 64

3.5 Detección de CDs por espectrometría de masas (MS-MALDI-TOF) 68

4. Discusión 73

5. Conclusiones 83

6. Recomendaciones 84

Referencias Bibliográficas 85

Apéndices 100

Lista de Tablas

Tabla 1. Información general de los modelos obtenidos por el portal web SynLinker. Longitud AA = Longitud de aminoácidos, EM = Energía de minimización **..... 31**

Tabla 2. Resumen de los valores obtenidos en los gráficos de Ramachandran. **..... 32**

Tabla 3. Valores de RMSD por los alineamientos estructurales de los modelos con respecto a los moldes AS y CGTasa**..... 41**

Tabla 4. Valores de actividad específica y producción de azúcar reductor (mM) mediante la reacción de hidrólisis de 0.1 M sacarosa, de la AS y AS·CGTasa. **..... 61**

Tabla 5. Actividad específica de la CGTasa (fracción FS) y de la fusión AS·CGTasa (fracción FS) en la producción de α-CDs y β-CDs mediante la reacción de ciclización con 1.5% almidón**.**
..... 63

Lista de Figuras

Figura 1. Dominios estructurales de la AS (PDB 1g5a) y la CGTasa (PDB 4jcl)..... **29**

Figura 2. Modelos AS·[GGGS]n·CGTasa SynLinker..... **34**

Figura 3. Modelos 3D fusión CGTasa·[GGGS]n·AS SynLinker..... **37**

Figura 4. Sitios activos AS y CGTasa en los modelos 3D **38**

Figura 5. Alineamiento estructural de los modelos AS-(GGGGS)n-CGTasa con las proteínas molde..... **39**

Figura 6. Alineamiento estructural de los modelos CGTasa-(GGGGS)n-AS con las proteínas molde..... **40**

Figura 7. Trayectoria de estabilidad de los modelos 3D y las proteínas Molde. **42**

Figura 8. Radio de giro de los modelos y las proteínas molde. **44**

Figura 9. Esquema de construcción de los vectores. a) pET-GST-AS-His, b) pET-DacD-cgt-His, c) pET-DacD-AS-cgt-His **46**

Figura 10. Análisis de restricción de los constructos..... **49**

Figura 11. Curva de crecimiento de E. coli transformada con pET-GST-AS-His vs actividad especifica (U/mg) de cada tiempo de inducción. **52**

Figura 12. Actividad específica (AE) de hidrolisis de sacarosa de extractos crudos FIC y FM de E. coli BL21 transformada con pET–GST-AS-His y pET-21a. **53**

Figura 13. Influencia del pH sobre la actividad específica de hidrolisis de 0,1 M de sacarosa de AS, usando extractos de FIC de E. coli transformada con pET–GST-AS-His..... **54**

Figura 14. Curva de crecimiento de E. coli transformada con pET-DacD-cgt-His vs actividad específica de ciclización de almidón (AE)..... **56**

Figura 15. Actividad específica (AE) de extractos crudos FS (rosado) y FM (blanco) de E. coli BL21 transformada con pET-AS-cgt-His y pET-21a. **57**

Figura 16. Influencia del pH sobre la actividad específica de ciclización de almidón de la CGTasa cgt-His. **58**

Figura 17. Curva de crecimiento de E. coli transformada con pET–DacD-AS-cgt-His vs actividad específica de hidrólisis de sacarosa. **59**

Figura 18. Actividad específica de hidrólisis de sacarosa de las enzimas recombinantes AS·CGTasa y AS. **60**

Figura 19. Actividad específica de ciclización para producir α-CDs y β-CDs (FS) de la CGTasa, la fusión AS·CGTasa **62**

Figura 20. Gel SDS-PAGE de las fracciones FIC de E. coli recombinante, inducida con IPTG. **64**

Figura 21. Producción de α-CDs y β-CDs, por reacción combinada de AS + CGTasa y de la Fusión AS·CGTasa con sacarosa. **66**

Figura 22. Producción de α-CDs y β-CDs por reacción combinada de AS + CGTasa y de la fusión AS·CGTasa con lixiviados de mucílago de cacao como sustrato **67**

Figura 23. Espectro de masas (MALDI-TOF-MS) de estándares comerciales de CDs α, β y γ. **69**

Figura 24. Espectro de masas (MALDI-TOF-MS) del producto de la reacción de la CGTasa con almidón como sustrato. **70**

Figura 25. Espectro de masas (MALDI-TOF-MS) del producto de la reacción de la AS + CGTasa con sacarosa como sustrato. **71**

Figura 26. Espectro de masas (MALDI-TOF-MS) del producto de la reacción de la fusión AS·CGTasa con lixiviados de mucílago de cacao como sustrato. **72**

Lista de Apéndices

Apéndice 1. Validación de los modelos SynLinker y las estructuras de las proteínas nativas AS (pdb 1g5a) y CGTasa (1cgt), mediante las herramientas PROSA WEB y Rampage Ramachandran..... **100**

Apéndice 2. Sitio activo de AS y CGTasa **109**

Apéndice 3. Secuencias GST·AS, CGTasa y fusión AS·CGTasa. **111**

Apéndice 4. Representación 3D de los sitios activos de AS·CGTasa (modelo ASL1CGT) con sacarosa y α-(1,4)-glucanos. **117**

Resumen

Título: Reacción Combinada de la Amilosacarasa y la Ciclodextrina Glucosiltransferasa para la Producción de Ciclodextrinas a Partir de Sacarosa y Lixiviados del Mucílago de Cacao[*]
Autor: Leidy Rocío Pico Martínez[**]
Palabras Clave: Ciclodextrinas, amilosacarasa, ciclodextrina glucosiltransferasa, proteína fusión, mucílago de cacao.

Descripción: Las ciclodextrinas (CDs) son oligosacáridos cíclicos compuestos de diferentes números de subunidades de glucopiranosa enlazadas en α-(1,4). Los CDs se producen comercialmente a partir de almidón, mediante la acción de la ciclodextrina glucosiltransferasa (CGTasa), que generalmente cataliza la ciclación del almidón y los α-(1,4)-glucanos producidos mediante reacciones de transglicosilación intramolecular. Un trabajo reciente desarrolló un proceso enzimático novedoso, utilizando sacarosa como materia prima en lugar de almidón, para reducir el tiempo y consumo de energía en la producción de CDs a nivel industrial. En este proceso, la amilosacarasa (AS) cataliza la síntesis de polímeros de α-(1,4)-glucanos a partir de moléculas de sacarosa, abriendo así la posibilidad de explorar otros sustratos ricos en azúcar disponibles comercialmente. Este trabajo tuvo como objetivo evaluar la acción combinada de AS de y CGTasa de *Bacillus macerans*, para producir CDs en una sola reacción a partir de mucílago de cacao y sacarosa como sustratos. Se utilizaron dos estrategias metodológicas: 1) probar la acción combinada de la AS y la CGTasa agregadas simultáneamente; 2) Evaluar la proteína AS-CGTasa, que surge como producto de la fusión en tándem de los genes AS y CGTasa. El primer paso implicó el uso de herramientas bioinformáticas para el diseño y análisis de la fusión AS-CGTasa. Como resultado, los modelos 3D de AS·CGTasa adquirieron una conformación con dos subunidades AS y CGTasa unidas por un motivo GGGS y se analizaron mediante simulaciones de dinámica molecular (GROMACS) para seleccionar el modelo más estable y evaluarlo experimentalmente. Los análisis in vitro mostraron que AS·CGTasa se comporta como una multienzima que es capaz de sintetizar CDs (sacarosa: α-CDs = 325 nmol, β-CDs = 297 nmol y Mucílago: α-CDs = 287 nmol, β-CDs = 259 nmol) en un solo paso a través de las reacciones acopladas de AS y CGTasa. En general, los resultados experimentales demostraron la capacidad de la reacción combinada de AS + CGTasa y la fusión AS·CGTasa para producir CDs en una reacción de un solo paso, a partir de sustratos no convencionales como el mucílago de cacao y la sacarosa. A partir de los resultados concluimos que el mucílago de cacao es un sustrato prometedor para la síntesis de CDs gracias a su aporte de monosacáridos y oligosacáridos. Además, se discute el papel de los monosacáridos como intermediarios en la reacción de polimerización de α-(1,4)-glucanos. Estos resultados pueden ser aplicados a futuro para optimizar los procesos de bioconversión de CDs a partir de mucílago de cacao u otros sustratos ricos en azúcar y que no requieran procesamiento industrial convencional.

---

[*] Trabajo de Grado

**Facultad de Ciencias. Escuela de Biología
Director: Jorge Hernández Torres, Biólogo, Ph.D.
Director: Hermínsul de Jesús Cano Calle, Químico, Ph.D.

Abstract

Title: Combined Reaction of Amylosucrase and Cyclodextrin Glucanotransferase to Produce Cyclodextrins from Sucrose and Cocoa Mucílage.*

Author: Leidy Rocío Pico Martínez**

Key Words: Cyclodextrins, Amylosucrase, Cyclodextrin glucanotransferase, one-pot synthesis, fusion protein, cocoa mucilage.

Description: Cyclodextrins (CDs) are cyclic oligosaccharides composed of different numbers of α-(1,4)-linked glucopyranose subunits. CDs are commercially produced from starch, through the action of cyclodextrin glucanotransferase (CGTase), which generally catalyzes the cyclization of starch and related α-(1,4)-glucans via intramolecular transglycosylation reactions. A recent work developed a novel enzymatic process, utilizing sucrose as raw material instead of starch, to reduce time and energy consumption of CDs production at the industrial level. In this process, a glucosyltransferase named amylosucrase (AS) catalyses the synthesis of α-(1,4)-glucan polymers from sucrose molecules, opening the possibility of exploring other commercially available sugar-rich substrates. The aim of this project was to test the combined action of AS from *Neisseria polysaccharea* and CGTase from *Bacillus macerans*, to produce CDs in a one-pot reaction from cocoa mucilage and sucrose as substrates. Two methodological strategies were used: 1) test the combined actions of AS and CGTase added simultaneously. 2) To evaluate the AS·CGTase fusion protein, which arises as a product of the tandem fusion of the AS and CGTase genes. The first step involved the use of bioinformatics tools for the design and analysis of the AS·CGTase fusion. As a result, 3D models of AS.CGTase acquired a conformation with two AS and CGTase subunits linked by a GGGS motif and were analyzed based on molecular dynamics simulations (GROMACS) to select the most stable ones from laboratory testing. In vitro analyzes showed that AS.CGTase behaves like a multienzyme that is capable of synthesizing CDs (sucrose: α-CDs = 325 nmol, β-CDs = 297 nmol and mucilage: α-CDs = 287 nmol, β-CDs = 259 nmol) in a single step through the coupled reactions of AS and CGTase. In general, experimental results demonstrated the ability of the dual enzyme reaction of AS+CGTase, and the AS.CGTase fusion to produce CDs in a one-pot reaction, from non-conventional substrates such as cocoa mucilage and sucrose. Cocoa mucilage is a promising substrate for the synthesis of CDs thanks to its contribution of monosaccharides and oligosaccharides. The role of monosaccharides as an intermediate in the polymerization reaction of α-(1,4)-glucans are discussed. These results can be applied in the future to optimize CD bioconversion processes from cocoa mucilage or other substrates rich in sugar and that do not require conventional industrial processing.

---

* Trabajo de Grado

** Facultad de Ciencias. Escuela de Biología
Director: Jorge Hernández Torres, Biólogo, Ph.D.
Director: Hermínsul de Jesús Cano Calle, Químico, Ph.D.

Introducción

Las ciclodextrinas (CDs) son oligosacáridos circulares compuestos de diferentes unidades de D-glucopiranosa, unidas por enlaces (α-1,4 glucano) (Koh et al., 2016; Van Der Veen et al., 2000). Se caracterizan por presentar una forma cónica con una cavidad interior hidrofóbica y la superficie exterior hidrófila (Tiwari et al., 2010). Esta configuración les confiere la capacidad de actuar como contenedores, atrapando moléculas en su cavidad interior (Gawande et al., 1999). En consecuencia, las CDs han despertado gran interés, ya que llevan a cabo acciones como la eliminación de sustancias no deseadas de una mezcla, ayuda en la liberación controlada de fármacos, además retiene el sabor y aroma de los compuestos alimenticios (Mizrahi y Wezel, 1983; Hedges, 1998; Gawande et al., 1999). Al estar presente en los alimentos, las CDs mejoran la estabilidad de los componentes, prolongado la vida útil de los productos. También son utilizadas para mejorar el valor nutracéutico de los alimentos, al encapsular compuestos como fitoesteroles vitamina E, D y K, carotenoides y otros antioxidantes (Stewart et al., 1999). Dentro de los productos alimenticios cotidianos, las CDs han sido exitosas en la producción de jugos con vitaminas, chicles aromatizados y, en la industria del cacao, en la elaboración barras de chocolate con mejores sabores, aromas y contenido nutricional (McKee & Karwic 2006). Las CDs son producidas comercialmente a partir de almidón, como resultado de reacciones de transglicosidación, comúnmente conocidas como ciclización. Esta reacción es catalizada por la enzima ciclodextrina glucosiltransferasa, i.e., CGTasa (Schmid, 1989; Crabb & Mitchinson, 1997). En la naturaleza, las CGTasas son producidas por diferentes géneros del reino bacteriano, tales como *Bacillus, Klebsiella, Pseudomonas, Brevibacterium* y *Micrococcus* (Gawande et al., 1999). *Bacillus* es el principal género, con varias especies productoras y enzimas mayormente caracterizadas (Gawande et al., 1999).

Recientemente, la demanda de CGTasa ha sido considerable debido a su rol fundamental en la producción de CDs (Han et al., 2014). Los estudios para la producción de CGTasa a nivel industrial se han enfocado en la optimización de la expresión recombinante en *E. coli* y otros hospederos (Jeang et al., 2005; Bonilha et al., 2006; Jemli et al., 2008; Kriaa et al., 2012; Han et al., 2014; Sonnendecker et al., 2017). Por otro lado, se han adelantado esfuerzos para mejorar la estabilidad frente a solventes orgánicos y altas temperaturas, mediante la fusión con otras proteínas y etiquetas de péptidos (Kaulpiboon et al., 2010). No obstante, el almidón utilizado como materia prima para la producción industrial de CDs debe pasar por procesos previos con alto consumo de tiempo y temperatura (>90 ºC), ya que el almidón en su estado natural no puede ser utilizado por la enzima (Koh et al., 2016). Además, los estudios han demostrado que el proceso descarta alrededor del 80% de material residual, con almidón como sustrato (Biwer et al., 2002; Urban et al., 2012). Los trabajos recientes han sido direccionados hacia mejorar la producción de CDs con bajo consumo de energía, y con metodologías más limpias y naturales. Un reporte reciente exploró la producción de CDs utilizando sacarosa como sustrato, en lugar de almidón (Koh et al., 2016). La estrategia consistió en la acción conjunta de dos enzimas adquiridas comercialmente: la amilosacarasa de *Neisseria polysaccharea* (*Np*AS) para producir α-(1,4)-glucanos lineales (almidón modificado) y la CGTasa de *Bacillus macerans* para producir α-, β- y γ-CDs a partir de los glucopolímeros, mediante la reacción de ciclización. En dicha reacción, se logró obtener cerca de 21% de producción de CDs (Koh et al., 2016). Los resultados de Koh et al. (2016) resultan alentadores, ya que pueden ser aplicados a futuro en la producción de CDs en masa que pueden ser incluidas en la industria alimenticia y farmacéutica. No obstante, es muy probable que, con la aplicación de la ingeniería de proteínas, esta tecnología sea susceptible de mejoramiento.

La amilosacarasa (AS) es una α-glucosiltransferasa bacteriana (EC 2.4.1.4) que, con sacarosa como sustrato, cataliza la biosíntesis de polímeros similares a la amilosa (α-1,4 glucano), con liberación concomitante de fructosa (Poctoki et al., 2000; Pizzut-Serin et al., 2005). Esta enzima versátil lleva a cabo tres reacciones diferentes: 1) Hidrólisis de sacarosa para liberar moléculas de glucosa y fructosa. 2) Polimerización, en donde la glucosa inicialmente liberada se utiliza como aceptor para producir maltosa. Las transferencias sucesivas de unidades glucosídicas dan como resultado la producción de un polímero similar a la amilosa. 3) Síntesis de isómeros de sacarosa (turanosa, isomaltosa, trehalulosa) (Cambon et al., 2014; Montalk et al., 2000; Skov et al., 2013). A diferencia de otras enzimas de su familia, la AS no requiere de azúcares activados como donadores (Zhang et al., 2017a). Gracias a esta propiedad única, las amilosacarasas han sido aplicables en la producción industrial de glucopolímeros hiperramificados y otros glucanos conjugados (Tian et al., 2018). Actualmente, los estudios sobre esta enzima se han centrado en la producción de AS recombinantes, porque no ha sido rentable extraerlas de las bacterias de origen. Se ha logrado la obtención de AS recombinantes con mejores rendimientos, gracias a la optimización de los sistemas de expresión bacterianos y a la construcción de quimeras con proteínas de fusión (ej. glutatión S-transferasa y diacilglicerol aciltransferasa) (Montalk et al., 1999; Emond et al., 2008).

El grupo de Investigación en Fisicoquímica Teórica y Experimental de la Escuela de Química llevó a cabo un macroproyecto enfocado hacia el aprovechamiento del fruto del cacao y sus subproductos. Colombia es el cuarto productor de cacao en Latinoamérica después de Brasil, Ecuador y República Dominicana. Para 2012, se alcanzó una producción alrededor de 50.000 ton (Ortiz-Rodríguez et al., 2016). Dentro de este sistema de producción, se genera un gran volumen

de residuos que no son utilizados, y que, generalmente, quedan acumulados en los cultivos, favoreciendo el crecimiento de microorganismos y plagas. Cada ton de semilla representa cerca de 10 ton de cáscara (peso fresco) y de 800 kg de semillas frescas derivan aproximadamente 40 L de pulpa o mucílago (Ortiz y Álvarez, 2015). En la actualidad, debido a sus características y contenido de azúcares, el lixiviado de mucílago de cacao se emplea como materia prima para la preparación de productos alimenticios como refrescos, vinos y mermeladas (Anvoh et al., 2009; Oddoye et al., 2013). Sin embargo, el lixiviado de mucílago puede aprovecharse en otros frentes, pues producido en exceso, como ocurre actualmente, termina siendo descartado como desecho. Una alternativa por explorar sería el uso combinado de una AS y una CGTasa para producir CDs. Producir ciclodextrinas a partir de lixiviados de mucílago de cacao añadiría un valor agregado a los productos alimenticios derivados del cacao, como la elaboración de jugos, barras de chocolate, entre otros. Producir CDs por este método, implica una estrategia de producción ecológica, limpio y libre de químicos ya que, además de mitigar la contaminación y los problemas que surgen a partir de la acumulación residual del mucílago, es una forma segura y natural para producir estas sustancias químicas incorporadas en los alimentos.

En este contexto, el presente trabajo de grado tuvo como objetivo evaluar la acción combinada de la AS de *N. polysaccharea* y la CGTasa de *B. macerans* para producir CDs en una sola reacción, a partir de sustratos no convencionales como la sacarosa y los lixiviados de mucílago de cacao. Para este fin, se siguió dos enfoques metodológicos: 1) evaluar la acción combinada de la AS y la CGTasa recombinantes, agregadas simultáneamente en la reacción. 2) mediante la acción de la proteína AS·CGTasa recombinante, construida a partir de la fusión de los genes AS y CGTasa. Los resultados experimentales demuestran la capacidad de la reacción acoplada de la AS

y la CGTasa con lixiviados de mucílago de cacao, para producir CDs en un solo paso. Además, se reporta por primera vez una proteína fusión AS·CGTasa bifuncional, que conserva la función de cada uno de sus dominios AS y CGTasa, con los sustratos convencionales (sacarosa y almidón). Asimismo, la AS·CGTasa tiene la capacidad de sintetizar CDs a partir de sacarosa y lixiviados de mucilago de cacao, mediante la sinergia de las reacciones de polimerización de la AS y ciclización de la CGTasa. Estos resultados sirven como punto de partida para optimizar los procesos de bioconversión de CDs a partir de lixiviados de mucílago de cacao u otros sustratos ricos en azúcar y que no requieran de los procesamientos industriales convencionales. La síntesis de CDs a partir de un sustrato natural como el lixiviado de mucílago de cacao, aporte un valor agregado, ya que su producción implica un proceso más natural que podría ser incorporado en la industria alimenticia, incluida en la línea de producción de alimentos derivados del cacao.

1. Objetivos

1.2 Objetivo General

Evaluar la acción combinada de la amilosacarasa (AS) y la ciclodextrina glucosiltransferasa (CGTasa), para la producción de ciclodextrinas a partir de la sacarosa y lixiviados del mucílago de cacao.

1.3 Objetivos Específicos

- Construir los vectores de expresión recombinantes con los genes que codifican para la AS de *Neisseria polysaccharea*, la CGTasa de *Bacillus macerans* y una fusión AS·CGTasa .
- Sobreexpresar en *Escherichia coli* las proteínas recombinantes.
- Comprobar las actividades enzimáticas con extractos crudos de las enzimas recombinantes utilizando sacarosa y mucílago de cacao como sustrato.
- Constatar la producción de CDs mediante espectrometría de masas (MALDI-TOF).

2. Materiales y métodos

2.1 Diseño *in silico* de las enzimas recombinantes AS, CGTasa y fusión (AS·CGTasa).

2.1.1 Diseño de las estructuras primarias de las proteínas AS, CGTasa y *AS·CGTasa*.

Las secuencias aminoacídicas de la AS de *Neisseria polysaccharea* (Skov et al. 2001) y la CGTasa de *Bacillus macerans* (Klein & Schulz 1991) fueron seleccionadas de la base de datos RCSB Protein D2.ata Bank (códigos PDB 1g5a y 4jcl, respectivamente).

2.1.2 Modelado 3D de la proteína de fusión AS·CGTasa.

La combinación de modelos 3D se llevó a cabo en el portal SynLinker (Yu et al. 2015). Para el diseño de la proteína de fusión, las estructuras 3D de la AS y la CGTasa fueron concatenadas utilizando los motivos GGGGS(n) y GGGS(n) como región linker (Argos 1990; Cortes, Wedd, & Xiao 2015; Yu et al. 2015). Seleccionados los linkers de interés, se procedió a fusionar los archivos PDB, utilizando las estructuras cristalinas de referencia, i.e., 1g5a y 4jcl. La conservación geométrica de las estructuras se determinó mediante superposición 3D de moldes y modelos, utilizando el programa PyMOL 1.7.0 (Schrodinger, LLC. 2010). La validación de los modelos se realizó utilizando los servidores PROSA, SAVES y Rampage Ramachandran. Los diagramas de Ramachandran (Ramachandran, Ramakrishnan, & Sasisekharan 1963) se construyeron con el servidor web PROCHECK (Laskowski et al. 1993).

2.1.3 Evaluación de la estabilidad estructural de los modelos mediante dinámica molecular

Las simulaciones de dinámica molecular se llevaron a cabo utilizando el programa GROMACS 5.1 (Abraham et al 2019). Se utilizó el campo de fuerza OPLS-AA (Jorgensen,

Maxwell, & Tirado-Rives 1996), incluido en GROMACS 5. El sistema fue solvatado dentro de una caja cúbica, empleando el modelo de agua SPC-216 (*Simple Point Charge*) (Berendsen et al. 1981). Las cargas de cada complejo fueron neutralizadas con 100 mM NaCl. Se llevó a cabo una minimización de energía para lograr un valor de energía potencial máxima de 1000 kJ/mol. El sistema fue equilibrado bajo condiciones isotérmicas e isobáricas a temperatura constante de 300 K (NVT) y presión de 1 bar (NPT). Finalmente, la trayectoria fue obtenida durante 10 ns de simulación, utilizando la opción *mdrun* de GROMACS. La estabilidad de los modelos fue evaluada mediante la comparación de la desviación cuadrática media (RMSD) y el radio de giro (Rg) (Abraham et al. 2015). Se hicieron pruebas de estabilidad y flexibilidad *in silico,* variando el orden N–terminal y C–terminal de las proteínas y el número de repeticiones de los motivos GGGS y GGGGS. Se contrastaron las variantes AS $_{\text{N–terminal}}$–CGTasa $_{\text{C–terminal}}$ y CGTasa $_{\text{N–terminal}}$–AS $_{\text{C–terminal}}$ (ver resultados).

2.1.4 Diseño de los vectores recombinantes pET–GST-AS-His, pET–DacD-cgt-His y pET–DacD-AS-cgt-His.

Para la construcción del plásmido pET–GST-**AS**-His recombinante, el *ORF* inició con 687 pb del gen de la enzima glutatión S-transferasa (GST, Genbank: U78874.1), entre los sitios *Nde* I y *Bam* HI del vector pET–21a, seguidos de 1908 pb codificantes para la AS de *N. polysaccharea* (Genbank: AJ011781.1). Trabajos anteriores han demostrado que la fusión *in frame* GST-AS favorece la expresión recombinante de la AS (Joucla et al. 2004; Montalk et al. 1999; Montalk et al. 2000). Al final del *ORF*, se añadieron los nucleótidos codificantes para una etiqueta de polihistidina (His·Tag), para futura purificación en columna de afinidad.

Igualmente, el vector pET–DacD-**cgt**-His inició con la adición de 63 pb del péptido señal DacD, para mejorar la expresión recombinante (Sonnendecker et al. 2017). Seguidamente, se clonó el ORF del gen α–*cgt* de *B. macerans* (Genbank: M12777.1) *in frame*, entre los sitios de restricción *Nde* I y *Bam* HI del vector pET–21a.

Finalmente, el ORF de la fusión AS–CGTasa más estable, según las simulaciones bioinformáticas, se construyó añadiendo en orden consecutivo 63 pb del péptido señal DacD, las secuencias nucleotídicas de la AS y la CGTasa (conectadas por el *linker* GGGGS) y, finalmente, una etiqueta de polihistidina (His·Tag). El constructo en su conjunto, clonado en el mismo vector, se denominó pET–DacD-**AS-cgt**-His. La síntesis química de los genes, previa optimización con el uso de codones de *E. coli*, fue contratada con la empresa General Biosystems.

2.2 Sobrexpresión en *Escherichia coli* de los polipéptidos GST·AS·His, DacD·CGTasa·His y DacD·AS·CGTasa·His.

2.2.1 Transformación de E. coli.

E. coli BL21(DE3) fue transformada con los plásmidos pET–GST-AS-His, pET–DacD-cgt-His, pET–DacD-AS-cgt-His y pET–21d-NR (control negativo). Los precultivos fueron incubados durante la noche, en 5 ml de medio Luria-Bertani (LB) a 37 °C y agitación a 180 rpm. Un ml de cultivo se repicó en 20 ml de medio LB hasta alcanzar una OD600nm = 1.0. Se procedió a preparar células competentes para la transformación por tratamiento con $CaCl_2$, según Sambrook & Russel (2001). Posteriormente, fueron transformadas con los plásmidos anteriormente citados, siguiendo las instrucciones de los mismos autores.

2.2.2 Control de fidelidad de los constructos.

De las cepas transformadas, se seleccionaron colonias únicas que fueron utilizadas para minipreparaciones de ADN de plásmido en pequeña cantidad (Sambrook & Russel 2001). Los plásmidos extraídos fueron sometidos a análisis de restricción con las enzimas *Nde* I y *Bam* HI, en las condiciones descritas por el proveedor (Fermentas). Los productos de la digestión fueron visualizados en geles 0,8% agarosa y coloreados con *GelGreen Nucleic Acid Gel Stain*, 10.000X (Biotium).

2.2.3 Expresión heteróloga de las proteínas recombinantes.

Las cepas seleccionadas de *E. coli* BL21(DE3) recombinante fueron cultivadas en 5 ml de medio LB con 100 µg/ml ampicilina, a 37 °C y agitación a 180 rpm, durante toda la noche. Los 5 ml de precultivo se adicionaron a 250 ml de medio LB + 100 µg/ml ampicilina e incubados en las mismas condiciones, hasta alcanzar una OD600nm = 0.6. Las células fueron sedimentadas mediante centrifugación a 5000 *g* por 5 min a 4 °C y resuspendidas en el mismo volumen de medio LB fresco [pET–GST-**AS**-His (Albenne, Potocki De Montalk, et al. 2002) y pET–DacD-AS-cgt-His] y *Terrific Broth* (TB) [pET–DacD-**cgt**-His (Li et al. 2010)]. La expresión fue inducida mediante la adición de 1 mM [pET–GST-AS-His] (Emond et al. 2007; Joucla et al. 2004) y 0.5 Mm IPTG [pET–DacD-**cgt**-His y pET–DacD-**AS-cgt**-His] (Sonnendecker et al. 2017) e incubación a 25 °C con agitación, hasta alcanzar una OD600nm = 1.0. Alícuotas de 15 ml de cultivo fueron recuperadas a 0, 2, 4, 6, 12 y 24 h, después de inducción con IPTG. De cada alícuota se efectuó la separación de las fracciones de medio (FM) y citoplasma insoluble (FIC) siguiendo el manual de Novagen para el sistema pET (10th Ed.).

2.2.4 Solubilización de Proteínas

Para solubilizar las proteínas recombinantes se modificaron los protocolos reportados en el manual "*Purifying Challenging Proteins*" (GE Healthcare 2007) y (Kim et al. 2000), así: Centrifugar 15 ml de cultivo (post-inducción) a 10.000 *g* a 4 °C durante 10 min y descartar el sobrenadante. Luego, resuspender el *pellet* en 7 ml de buffer 20 mM Tris-HCl pH 8.0. Después, sonicar las células en hielo, amplitud 65%, pulso 5 x 10 s, durante 5 min. Centrifugar la suspensión a 20.000 *g* a 4 °C durante 30 min. Descartar sobrenadante. Posteriormente, resuspender el pellet en ½ vol de solución fría de: 50 mM Tris-HCl pH 8.0, 20 mM EDTA, 100 mM NaCl, 2% Tritón X-100. Sonicar nuevamente bajo las mismas condiciones mencionadas. Finalmente, la suspensión se centrifuga a 20.000 *g* a 4 °C durante 30 min y se colecta el sobrenadante con las proteínas solubilizadas, llamada Fracción Soluble (FS).

2.3 Ensayos de Actividad

Para todos los ensayos de actividad, se utilizó como muestra control los extractos proteicos del clon pET-21d no recombinante (pET-21d-NR).

2.3.1 Ensayo de actividad AS: determinación de azúcares reductores

La actividad de hidrólisis de GST·**AS**·His y DacD·**AS·CGTasa**·His fue determinada a 40 ºC, midiendo la concentración de azúcares reductores liberados mediante el método colorimétrico con ácido 3.5-dinitrosalicílico o DNS (Sumner & Howell 1934), en un espectrofotómetro Shimadzu UV-VIS 1800. La mezcla de reacción de 2 ml incluía: 1 ml 1X Buffer fosfato pH 7.0, 500 µl de extracto proteico (FIC) y 0.1 M sacarosa como sustrato control. La reacción fue incubada por 10 min a 40 °C y se detuvo agregando 1 ml de DNS, seguido de calentamiento a 95 °C por 10

min. Para cuantificar la concentración de azúcares reductores, primero se llevó a cabo una curva estándar utilizando fructosa. Una unidad (1 U) de AS fue definida como la cantidad de enzima que libera 1 µmol de fructosa por min.

2.3.2 Ensayo de actividad CGTasa: Producción de CDs a partir de almidón

Para evaluar la producción de CDs por la actividad de ciclización de la CGTasa, se utilizó el método colorimétrico de fenolftaleína propuesto por Makela et al. (2000). La fenolftaleína actúa como compuesto acomplejante de la β-CD presente en la solución. Por lo tanto, en presencia de β-CDs, la solución se tornará de color rosado claro. Para evaluar la concentración de β-CDs, primero fue necesario realizar una curva de calibración estándar de 0 a 2 mg/ml de estándares comerciales de β-CDs, con una solución de 0.4 mM fenolftaleína y 1 M NaCO3. La concentración de β-CD fue determinada midiendo el ΔA 550 nm (Makela et al. 2002), en un espectrofotómetro Shimadzu UV-VIS 1800, a temperatura ambiente.

Igualmente se evaluó la producción de α-CDs para las mejores condiciones de reacción, utilizando el método colorimétrico de naranja de metilo, propuesto por Lejeune (1989) (Lejeune, Sakaguchi, & Imanaka 1989). El método se basa en la disminución de la absorbancia cuando una solución de naranja de metilo forma complejo con las α-CDs producidas en la reacción (Cramer, Saenger, & Spatz 1967). Se realizó una curva de calibración estándar de 0 a 1 mg/ml de estándares comerciales de α-CDs, con una solución de 0.035 mM naranja de metilo y 0.3 N HCl, por 20 min a 16 °C. La concentración de α-CDs fue determinada midiendo el ΔA 546 nm (Lejeune, Sakaguchi, & Imanaka 1989) en un espectrofotómetro Shimadzu UV-VIS 1800, a temperatura ambiente.

La mezcla de reacción de 2.5 ml incluía: 1 ml de 50 mM Tris-HCl pH 8.0; 1 ml de extractos crudos de DacD·**CGTasa**·His y DacD·**AS·CGTasa**·His (FM y FS); como sustrato control, una solución

de 1.5% almidón de yuca, diluido en el buffer de reacción y autoclavado durante 5 min. La preparación fue incubada durante 30 minutos a 45 °C. La reacción se detuvo con calentamiento a 95 °C y enfriamiento durante 10 min. Posteriormente, a 700 µl de reacción fueron agregados 300 µl de fenolftaleína; se dejó reposar durante 10 min y se procedió a medir la OD550nm.

2.3.3 Determinación del pH óptimo

Para determinar el pH óptimo en las reacciones, se prepararon mezclas siguiendo las especificaciones de los numerales 2.3.1 y 2.3.2. Los intervalos de pH evaluados fueron los siguientes: buffer 50 mM Citrato pH 3.0 y pH 4.0, buffer Acetato de sodio pH 5.5, buffer 1X Fosfato pH 7.0, buffer 50 mM Tris-HCl pH 8.0 y buffer 50 mM Glicina-NaOH pH 10 (Sambrook & Russel 2001).

2.3.4 Producción de CDs por reacción combinada de GST·AS·His y DacD·CGTasa·His

Para sintetizar CDs en una sola reacción (3 ml finales), extractos crudos de 0.588 mg/ml GST·**AS**·His (0,5 ml) y 1.23 mg/ml de DacD·**CGTasa**·His (1 ml) fueron incubados con 0.1 M sacarosa en 1 ml de 50 mM Tris-HCl pH 7.5, durante 24 h a 45 °C. La reacción fue detenida por calentamiento durante 10 min y enfriamiento. Posteriormente, la mezcla de reacción fue centrifugada a 9000 *g* por 10 min a temperatura ambiente. La producción de CDs se corroboró, adicionando 600 µl de mezcla de reacción a 200 µl de fenolftaleína, al igual que en el ítem 2.3.2.

2.3.5 Producción de CDs por DacD·AS·CGTasa·His a partir de sacarosa

Se preparó una mezcla de reacción de 2.5 ml de vol final que contenía: 1 ml de extractos crudos de (1.096 mg/ml) DacD·**AS·CGTasa**·His, 1 ml de 50 Mm Tris-HCl pH 7.5 y 0.1 M

sacarosa. Se siguieron las condiciones de incubación y cuantificación de CDs mencionadas anteriormente.

2.3.6 Producción de CDs a partir de lixiviados de mucílago de cacao

Se siguió la metodología de los ítems 2.3.4 y 2.3.5. Para neutralizar el pH ácido del mucílago, se preparó una solución 2 : 1, de lixiviado de mucílago de cacao : 0,5 M Tris-HCl pH 8.0. Posteriormente, se prepararon reacciones de 3 ml que contenían: 0.588 mg/ml crudos de GST·**AS**·His (0.5 ml) y 1.23 mg/ml crudos de DacD·**CGTasa**·His (1 ml), 1 ml de 50 mM Tris-HCl pH 7.5 y 1 ml de la solución de mucílago de cacao, por 24 h a 45 °C. Para la reacción de la fusión DacD·**AS·CGTasa**·His, la mezcla de reacción fue la misma que para el ítem 2.3.5, reemplazando la sacarosa por 1 ml de solución de mucílago de cacao. La cuantificación de CDs se realizó con el método de fenolftaleína y naranja de metilo siguiendo las condiciones anteriormente relacionadas.

2.3.7 Determinación de la concentración de proteínas

La concentración de proteína total en los extractos crudos de FM, FIC y FS (en mg/ml) se determinó mediante el microensayo de Bradford (Ernst and Zor 2010), propuesto en el manual de Ausubel (1992). Como proteína de referencia se usó Albúmina de Suero Bovino (BSA).

2.3.8 Cálculo de las actividades enzimáticas

Para la actividad AS con el método de DNS, una unidad enzimática (1 U) fue definida como la cantidad de enzima que cataliza la liberación de 1 µmol de fructosa por min de reacción. En el caso la actividad CGTasa con los métodos de fenolftaleína y naranja de metilo, una unidad

enzimática fue definida como la cantidad de enzima que es capaz de producir 1 µmol de CDs por min de reacción. La actividad específica (AE) se calculó con base en el número de unidades enzimáticas por mg de proteína en extractos crudos. Para las pruebas de actividad de la proteína fusión AS·CGTasa y la acción combinada AS + CGTasa con mucílago de cacao y sacarosa, se calculó el número de μmol^{-1} min^{-1} partiendo de la concentración final de producto y el tiempo de reacción.

2.3.9 Electroforesis Desnaturalizante de Proteínas (SDS-PAGE)

Los extractos proteicos de las fracciones FM fueron separados por electroforesis desnaturalizante en geles 9% SDS-PAGE, de acuerdo con el protocolo sugerido por Laemmli (1970). Se empleó el *Prestained Protein Ladder* (Thermo Scientific, USA) como marcador de peso molecular.

2.4 Análisis estadístico

En todos los ensayos de actividad enzimática y producción de α-CDs y β-CDs producidas, se utilizó un diseño experimental completamente aleatorio con tres réplicas. De cada uno de los resultados, se calcularon los valores medios de actividad específica con sus correspondientes desviaciones estándar. Los datos fueron analizados mediante la prueba de Kruskal-Wallis, utilizando el programa SPSS y se consideró un valor de significancia $p < 0.05$. Los gráficos de barras y curvas de crecimiento fueron realizados con el programa GraphPad Prism 6.0.

2.5 Detección de CDs por espectrometría de masas (MALDI-TOF)

Para constatar la producción de las CDs (α, β, y γ), el producto de las reacciones anteriores fue analizado por la técnica analítica de espectrometría de masas MALDI-TOF-MS (Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry). El producto de cada reacción fue centrifugado a 9000g 25 ºC durante 10 minutos. Luego, 1 ul de sobrenadante, con el contenido de CDs, fue mezclado con un 1.0 µl de ácido 2.5-dihidroxibenzoico (DHB) (10 mg/ml) utilizado como matriz. Se utilizaron estándares comerciales de α, β, y γ CDs (USP Reference Standards), como control y punto comparación de los espectros de masa molecular.

3. Resultados

3.1 Diseño *in silico* de los polipéptidos GST·AS, DacD·CGTasa y DacD·AS·CGTasa.

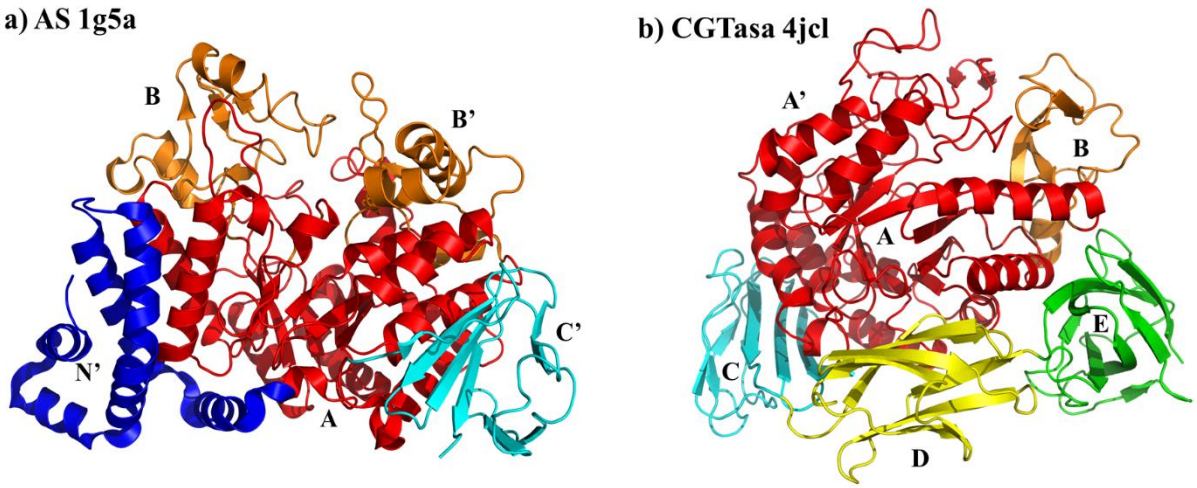
3.1.1 Selección de estructuras de las proteínas AS y CGTasa

El punto de partida para la construcción de una fusión AS·CGTasa viable fue la búsqueda de estructuras 3D disponibles en la PDB. Seleccionamos la AS de *N. polysaccharea* que es la única que, además de contar con estudios bioquímicos, moleculares de mutantes y estructurales (Skov, Mirza, Henriksen, De Montalk, et al. 2001), también tiene una estructura libre de sustratos (PDB 1g5a). La isoforma de *N. polysaccharea* tiene un peso molecular de 72.3 kDa y constituye de monómero de 628 residuos aminoacídicos. La AS pertenece a la familia de la glucósido hidrolasa (GH) 13, superfamilia α-amilasa (Davies & Henrissat 1995). La estructura 3D consta de cinco dominios denominados N, A, B, B’ y C (Figura 1a), de los cuales B’ y N son exclusivos de la AS (Pizzut-Serin et al. 2005).

Una situación similar ocurrió con la búsqueda de estructuras de la CGTasa. Solo la estructura 3D de la isoforma de *B. macerans* (PDB 4jcl) está libre de sustratos y cuenta con bastantes estudios bioquímicos y moleculares (Wu et al., 2014). Al igual que la AMS, la CGTasa hace parte de la familia GH 13, superfamilia α-amilasa (Qi & Zimmermann 2005). La proteína tiene un peso molecular de 76.9 kDa (Figura 1b) y consta de 714 aminoácidos distribuidos en cinco dominios A (A'), B, C, D y E (Chang, Irwin, & Nikolov 1998). En conclusión, las estructuras 1g5a y 4jcl fueron seleccionadas para la construcción de la proteína de fusión, por lo cual se procedió a realizarles las pruebas bioinformáticas sustanciales.

Figura 1.

Dominios estructurales de la AS (PDB 1g5a) y la CGTasa (PDB 4jcl).



Nota: Estructura PDB de las proteínas moldes para la construcción de la proteína fusión y las proteínas individuales.

3.1.2 Selección de linkers para la fusión de ORFs

La selección de la región *linker* entre proteínas de fusión es fundamental para reducir el impedimento estérico y conservar la estructura y función individual de cada proteína (Chen, Zaro, & Shen 2013; Yu et al. 2015). Las bases de datos ofrecen una amplia gama de posibilidades; no obstante, los *linkers* elegidos, i.e., [GGGGS]n y [GGGS]n, son de los motivos más utilizados por su flexibilidad y disponibilidad para construir modelos 3D (Horikawa et al. 2010; Van Rosmalen, Krom, & Merkx 2017). Se ha establecido que ambos motivos permiten a las proteínas fusionadas desarrollar sus actividades individuales (Argos 1990; Chen, Zaro, & Shen 2013). Por todo lo anterior, se incluyeron los motivos [GGGGS]n y [GGGS]n dentro de las simulaciones de dinámica molecular, para elegir la combinación más adecuada. La idea era obtener la mejor compactación sin comprometer la flexibilidad y al mismo tiempo preservar la estabilidad en solventes acuosos, a través de la formación de enlaces de hidrógeno con el agua (Argos 1990; Chen, Zaro, & Shen 2013; George & Heringa 2002). La longitud del número de repeticiones en tándem de los motivos [GGGS]n y [GGGGS]n (1–4) se ajustaron mediante ensayo y error, apoyados en los resultados de dinámica molecular, como se describe a continuación.

3.1.3 Modelamiento de estructuras 3D y elección de la estructura primaria de la fusión

Un segundo punto crítico en el diseño de la fusión fue la elección del orden consecutivo de las proteínas individuales, entre AS·CGTasa y CGTasa·AS. Está establecido que ciertos dominios o regiones de una de las proteínas puede afectar la funcionalidad de la otra (Chen, Zaro, & Shen 2013). Por consiguiente, fue necesario probar *in silico* las dos combinaciones posibles, i.e., AS·CGTasa y CGTasa·AS, enlazadas por los *linkers* seleccionados. Como resultado se evaluaron

ocho posibles variantes de proteína fusión: **AS**·$[GGGGS]_{1–4}$·**CGTasa** y **CGTasa** ·$[GGGGS]_{1–4}$·**AS** y dos variantes con motivos [GGGS].

Se procedió a evaluar la flexibilidad y estabilidad de modelos tridimensionales de cada variante de fusión, con el portal web SynLinker (Liu, Chin, & Lee 2015). El primer paso fue conseguir modelos 3D de los *linkers* y afortunadamente $[GGGGS]_{1–4}$ y [GGGS] estaban disponibles en la base de datos. El paquete *'fusedomains*' de SynLinker (Liu, Chin, & Lee 2015) inició con la estructura atómica resuelta para cada polipéptido y los linkers anteriormente seleccionados. En 5 pasos fundamentales se construyeron los modelos: 1) Traducción de coordenadas; 2) comprobación de choques estéricos; 3) rotación de coordenadas; 4) fusión de dominios; y 5) minimización de energía. De 10 fusiones testeadas, SynLinker modeló con éxito seis estructuras, tres para AS·CGTasa y tres para CGTasa·AS. En la Tabla 1 se muestran las variantes viables y se excluyeron las que no pasaron el número umbral de rotación exigido y reportaron errores.

Tabla 1.

Información general de los modelos obtenidos por el portal web SynLinker.

Modelo	***Abreviatura***	***Linker***	***AA***	***EM***
AS·$[GGGS]_1$·**CGTasa**	$AS_{L1}CGT$	$^{629}GGGS^{632}$	1319	-249.87
AS·$[GGGGS]_3$·**CGTasa**	$AS_{L3}CGT$	$^{629}GGGGSGGGGSGGGS^{643}$	1330	-240.93
AS·$[GGGGS]_4$·**CGTasa**	$AS_{L4}CGT$	$^{629}GGGGSGGGGSGGGGSGGGGS^{648}$	1335	-252.41
CGTasa·$[GGGGS]_1$·**AS**	$CGT_{L1}AS$	$^{685}GGGGS^{689}$	1317	-251.22
CGTasa·$[GGGGS]_3$·**AS**	$CGT_{L3}AS$	$^{685}GGGGSGGGGSGGGS^{698}$	1327	-252.56
CGTasa·$[GGGGS]_4$·**AS**	$CGT_{L4}AS$	$^{685}GGGGSGGGGSGGGGSGGGGS^{704}$	1332	-249.77

Nota: Longitud AA = Longitud de aminoácidos, EM = Energía de minimización

Con la herramienta Prosa web (Wiederstein & Sippl 2007), se evaluó la calidad estructural de los modelos de la Tabla 1. Como puede apreciarse en el Apéndice 1.1, los valores de Z-score se encuentran dentro de los rangos típicos encontrados en proteínas nativas resueltas experimentalmente por cristalografía de rayos X. Por lo tanto, se concluyó que las estructuras 3D se encuentran dentro de un rango de puntuación de calidad aceptable, para continuar con los demás análisis bioinformáticos. Seguidamente, se representaron diagramas de Ramachandran de los modelos, para evaluar la calidad estereoquímica (detalles en el Apendice 1.2). Como puede apreciarse en la Tabla 2, los modelos AS_{L1}CGT y AS_{L3}CGT presentaron >90% de aminoácidos en regiones favorables y solo 0,1% en regiones no permitidas, revelando una muy buena calidad de los modelos. Los demás modelos presentaron una calidad estereoquímica aceptable, ya que 86–87% de los residuos se observan en las regiones favorables, ~12% en las regiones permitidas y solo 0.2–0.3% en las regiones no permitidas. Se comprobó la validez de las estructuras 3D y se consideró que los modelos cumplían con las condiciones para nuevos análisis.

Tabla 2.

Resumen de los valores obtenidos en los gráficos de Ramachandran para los aminoácidos favorables y no favorables de los modelos 3D.

Parámetros evaluados	**$AS_{L1}CGT$**	**$AS_{L3}CGT$**	**$AS_{L4}CGT$**	**$CGT_{L1}AS$**	**$CGT_{L3}AS$**	**$CGT_{L4}AS$**
Regiones favorables	1032 [90.4%]	1032 [90.3%]	1001 [87.5%]	991 [86.9%]	991 [86.9%]	1009 [86.9%]
Regiones adicionales permitidas	104 [9.1%]	103 [9.0%]	133 [11.6%]	138 [12.1%]	143 [12.5%]	139 [12%]
Regiones generosamente repetidas	4 [0.4]	7 [0.6%]	7 [0.6%]	10 [0.9%]	7 [0.6%]	10 [0.9%]
Regiones no permitidas	1 [0.1%]	1 [0.1%]	3 [0.3%]	2 [0.2%]	2 [0.2%]	3 [0.3%]

Una visión global de la representación 3D (Figuras 2 y 3), revela que los modelos de fusión de la Tabla 1 conformaron dos dominios grandes (subunidades) y bien definidos, con una región *linker* (color rojo) cada vez más larga y flexible. En el orden de proteínas AS-L_n–CGTasa, los *linkers* no adoptaron estructuras secundarias canónicas, pero sí el mismo número de residuos, incluyendo los dos últimos residuos de la AS y los 16 los primeros aminoácidos de la CGTasa.

El modelo $AS_{L1}CGT$ (Figura 2a) presentó un *linker* con una longitud de 22 residuos (627 a 648), formando dos bucles antiparalelos. Para el modelo $AS_{L3}CGT$ (Figura 2b) el *linker* forma un bucle con los motivos GGGGS (cian), seguido de una región irregular (rojo), con una longitud de 33 aminoácidos, i.e., 627 a 659. Finalmente, el modelo $AS_{L4}CGT$ (Figura 2c) dio como resultado un largo *linker* (amarillo) de 38 residuos (627 a 664), con un patrón lineal. Además de la longitud y conformación de las regiones *linker*, otro aspecto que diferenció a los modelos fue la posición relativa en el espacio de las regiones N–terminal y C–terminal. Como se aprecia en la Figura 2a, en ASL_1CGT y ASL_4CGT la región N–terminal de AS se ubica en la zona superior – central, mientras que el extremo C–terminal de la CGTasa se posiciona en el lado opuesto. En

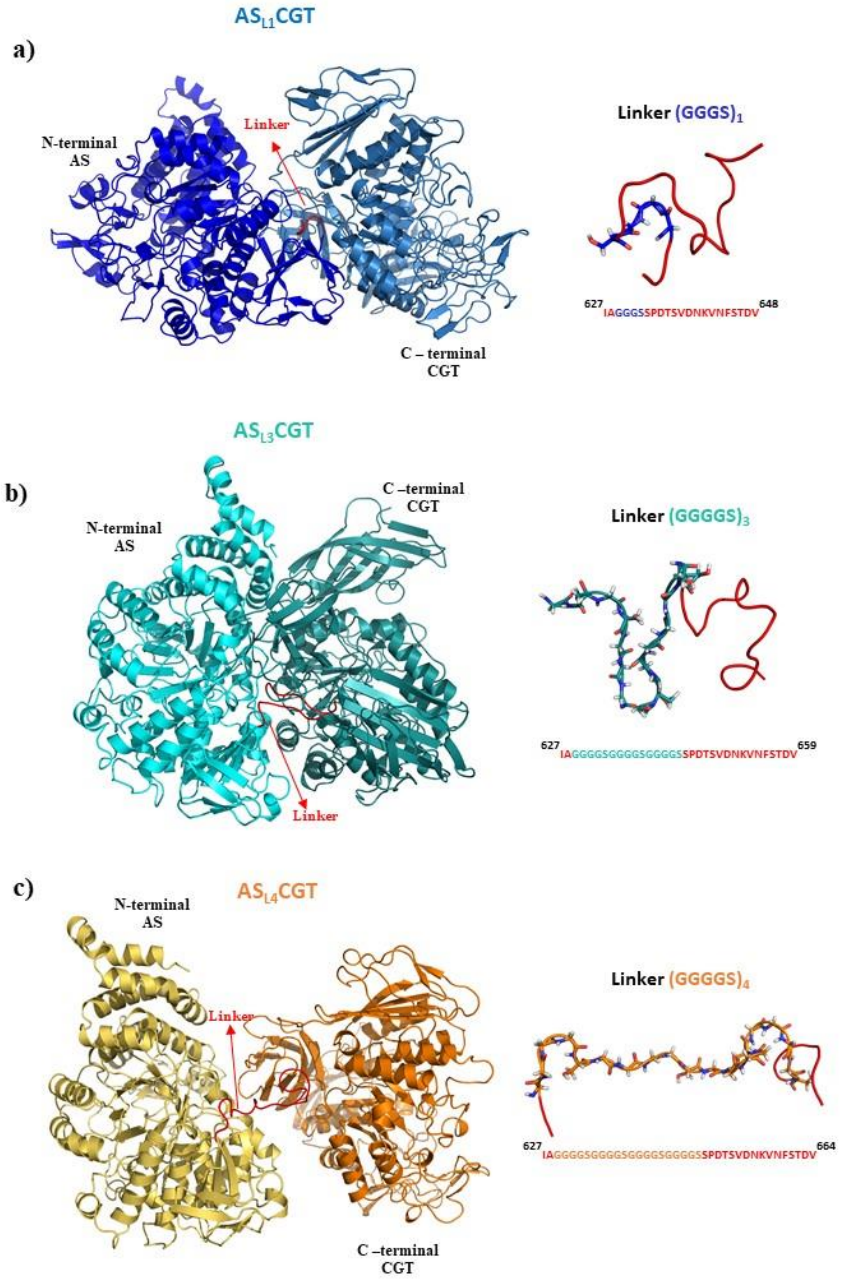
cambio, en el modelo $AS{L_3}CGT$ los extremos N-terminal y C-terminal de la AS se encuentran en el mismo plano. Esta disposición de los extremos refleja una mayor libertad de movimiento entre subunidades y más adelante se explica el impacto estructural y sobre todo funcional de este comportamiento.

Por otra parte, los modelos del orden **CGTasa**·$[GGGGS]_n$·**AS** presentaron regiones *linker* con una longitud menor (Figura 3), iniciando con los dos últimos aminoácidos (WQ) de la CGTasa y terminando con los dos primeros residuos (PN) de la AS. El modelo $CGT_{L1}AS$ exhibió un *linker* de 10 residuos, i.e., 683-692 (Figura 3a). El *linker* del modelo $CGT_{L3}AMS$, de 20 aminoácidos (683 a 702) formó un bucle extendido (Figura 3b). Finalmente, el modelo $CGT_{L4}AMS$, conformado por 25 residuos aminoacídicos, i.e., 683 a 707, $AS_{L3}CGT$, presentó una porción alargada, con un final en zigzag.

En cuanto a la posición relativa de las regiones N– y C–terminal, la flexibilidad fue similar a la fusión AS·CGTasa. En efecto, en los modelos $CGT_{L1}AS$ y $CGT_{L3}AS$, las regiones N– y C–terminal se ubicaron en el mismo plano, mientras que en $CGT_{L4}AS$ se disponen de manera opuesta (Figura 3).

Figura 2.

Modelos AS·[GGGS]n·CGTasa SynLinker.



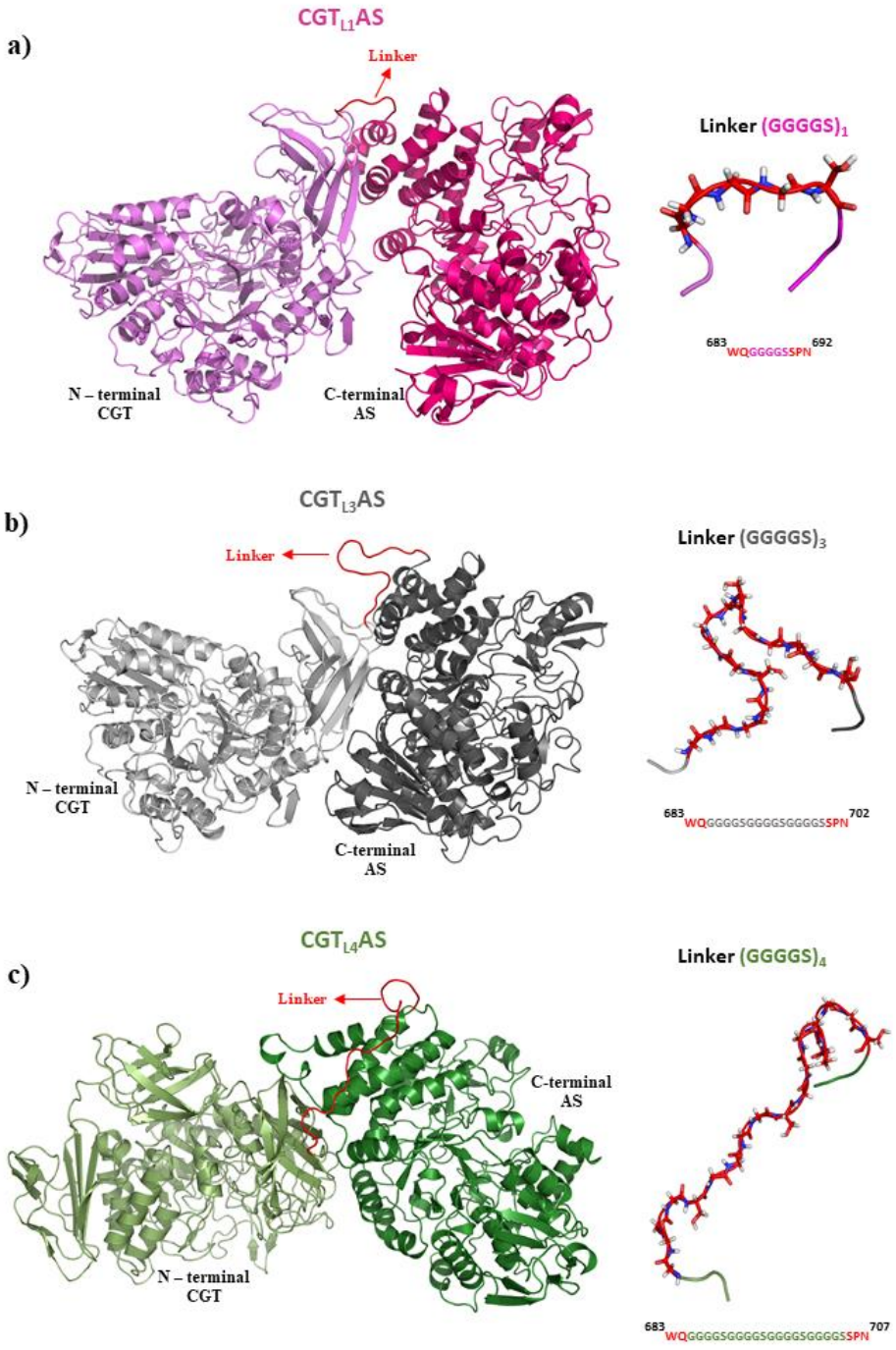
Nota: Estructura 3D de los modelos de la fusión AS·[GGGS]n·CGTasa obtenidos con SynLinker.a) AS$_{L1}$CGT, b) AS$_{L3}$CGT, c) AS$_{L4}$CGT.

Era válido suponer que la fusión de las proteínas AS y CGTasa tuviera repercusiones, principalmente en la accesibilidad de los sitios de unión al sustrato. Para revelar la posición relativa de los centros catalíticos, se realizaron representaciones de superficie, resaltando los residuos claves para la actividad enzimática (Figura 4, Apéndice 2). En la Figura 4 se puede observar que para el modelo $AS_{L1}CGT$, los sitios activos AS y CGTasa se situaron libres de obstáculos y ampliamente distantes. Contrariamente, los modelos $AS_{L3}CGT$ y $AS_{L4}CGT$ reflejan una disposición en la que el sitio activo de la AS queda embebido entre las dos subunidades, con mayor proporción en $AS_{L4}CGT$ (Figura 4c). Lo anterior, podría configurar una desventaja, ya que algunas regiones de la CGTasa podrían obstaculizar la debida interacción entre el bolsillo de unión y la sacarosa. Por el contrario, el sitio activo CGTasa, en ambos modelos se localizó hacia los costados externos de la proteína de fusión.

De otro lado, en los modelos $CGT_{L1}AS$ y $CGT_{L3}AS$, los sitios activos se ubicaron en el mismo plano, pero distantes, sin obstrucción visible en los bolsillos de unión (Figura 4 d y e). Al igual que con $AS_{L1}CGT$, el modelo $CGT_{L4}AS$ presentó los sitios activos CGTasa y AS en caras opuestas (Figura 4f). Una de las funciones asociadas al dominio E de la CGTasa (Figura 1) es la de servir como sitio de unión para el sustrato usual, i.e., almidón (Penninga et al. 1996)

Figura 3.

Modelos 3D fusión CGTasa·[GGGS]n·AS SynLinker.

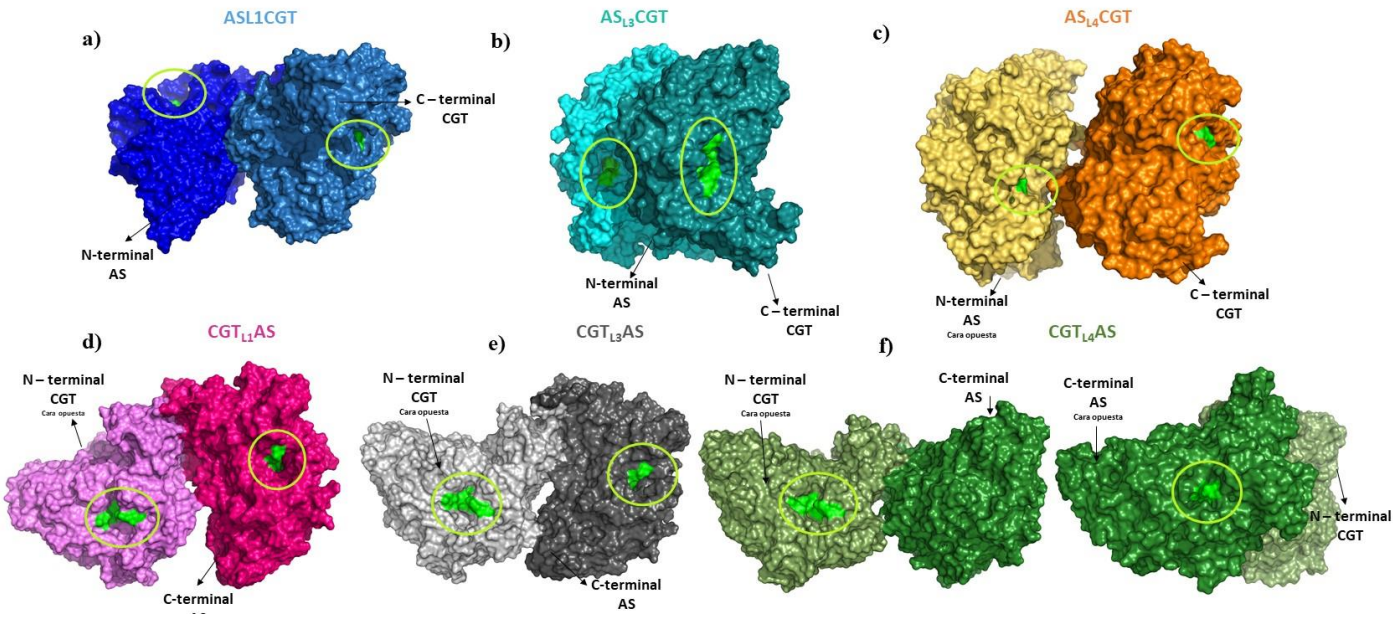


Nota: Estructura 3D de los modelos fusión CGTasa·[GGGS]n·AS obtenidos con el portal Web SynLinker. a) CGT$_{L1}$AS, b) CGT$_{L3}$AS, c) CGT$_{L4}$AS.

Con base en lo anterior, es altamente probable que la función del dominio E, en el ordenamiento **CGTasa**·[GGGGS]$_n$·**AS** pueda verse afectada, ya que la proximidad con el *linker* y la región N-terminal de la AS podría obstaculizar el acceso al almidón.

Figura 4.

Sitios activos AS y CGTasa en los modelos 3D

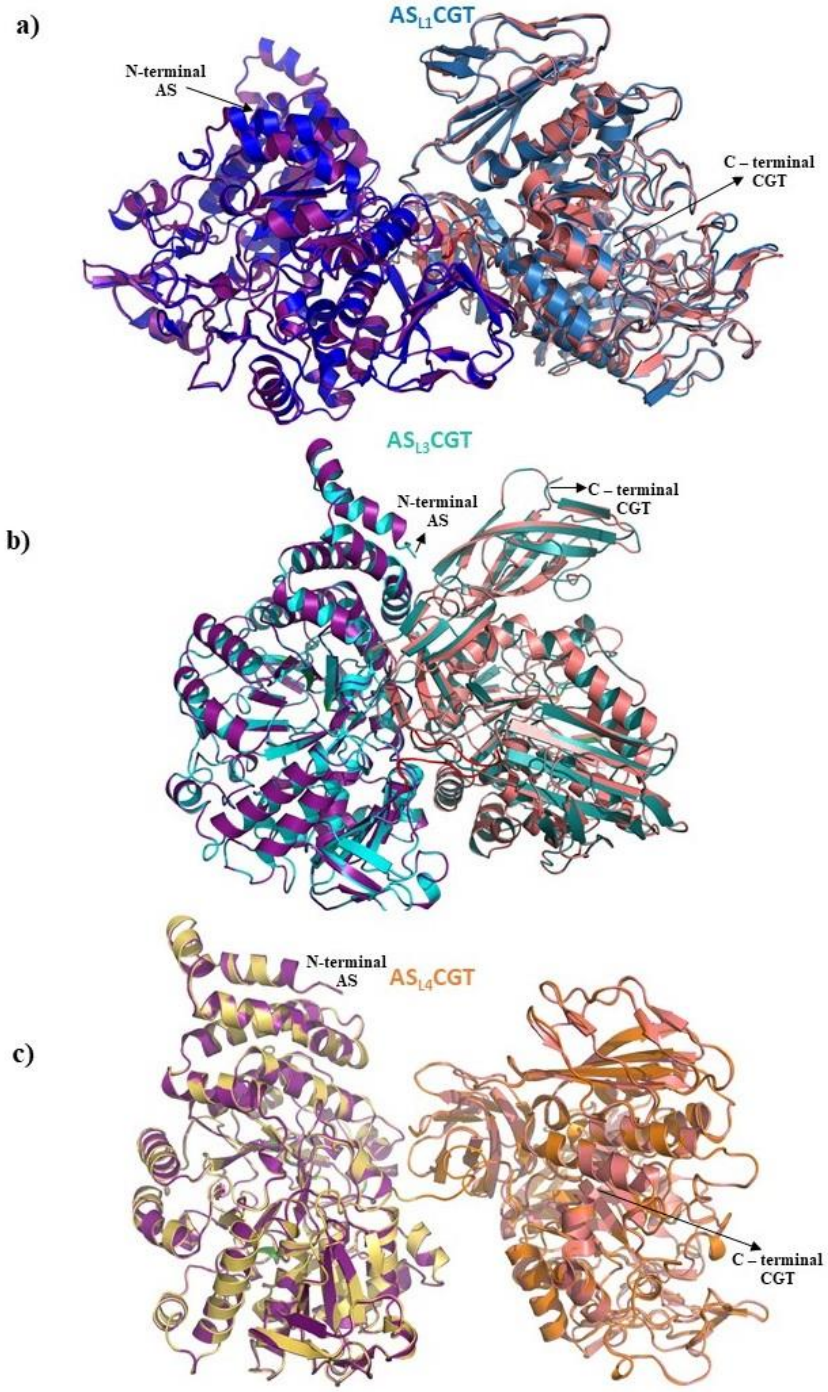


Nota: Representación en superficie de los modelos AS·CGTasa y CGTasa·AS. Ubicación del sitio activo de cada dominio AS y CGTasa (verde) en visualización tipo Surface.

Para constatar la fidelidad de las estructuras con respecto a los modelos originales (1g5a y 4jcl), procedimos a realizar alineamientos 3D y calcular el RMSD entre pares estructuras (Figuras 5 y 6). Los bajos valores de RMSD validan la correcta conformación de las subunidades (Tabla 3).

Figura 5.

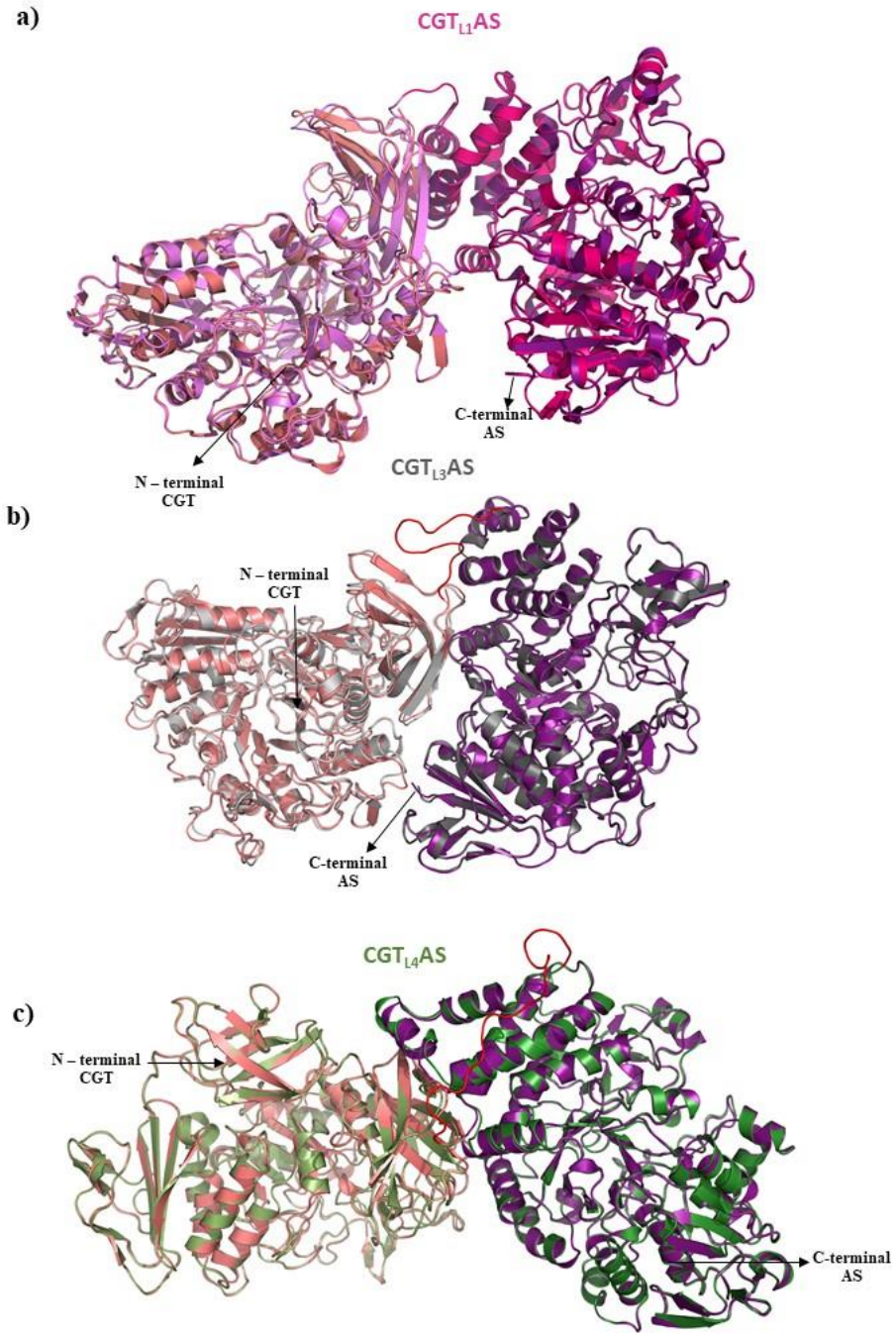
Alineamiento estructural de los modelos AS-(GGGGS)n-CGTasa con las proteínas molde.



Nota: Superposición de estructuras 3D de los modelos AS-(GGGGS)n-CGTasa con las proteínas AS (morado) y CGTasa (rosado). a) AS$_{L1}$CGT, b) AS$_{L3}$CGT y c) AS$_{L4}$CGT

Figura 6.

Alineamiento estructural de los modelos CGTasa-(GGGGS)n-AS con las proteínas molde.



Nota: Superposición de estructuras 3D de los modelos CGTasa-(GGGGS)n-AS con las proteínas AS (morado) y CGTasa (rosado). a) CGT$_{L1}$AS, b) CGT$_{L3}$AS y c) CGT$_{L4}$AS.

Tabla 3.

Valores de RMSD (Raíz de la Desviación Cuadrática Media) generados a partir de los alineamientos estructurales de los modelos con respecto a los moldes AS y CGTasa.

Modelo	RMSD Å [AS]	RMSD Å [CGTasa]
$AS_{L1}CGT$	0.271	0.318
$AS_{L3}CGT$	0.210	0.117
$AS_{L4}CGT$	0.299	0.287
$CGT_{L1}AS$	0.322	0.331
$CGT_{L3}AS$	0.313	0.328
$CGT_{L4}AS$	0.316	0.339

Tomados en conjunto los resultados del modelamiento se concluyó que fue posible fusionar las dos enzimas in silico, en cualquiera de las combinaciones AS·CGTasa o CGTasa·AS. Los valores de RMSD indican que no se afectó la estructura de cada subunidad; además, los sitios activos no se ven estructuralmente interferidos. En principio, al menos teóricamente, los dominios AS y CGTasa podrían ser funcionales. Se procedió entonces a realizar pruebas de estabilidad.

3.1.4 Evaluación in silico de la estabilidad de las fusiones mediante Dinámica Molecular

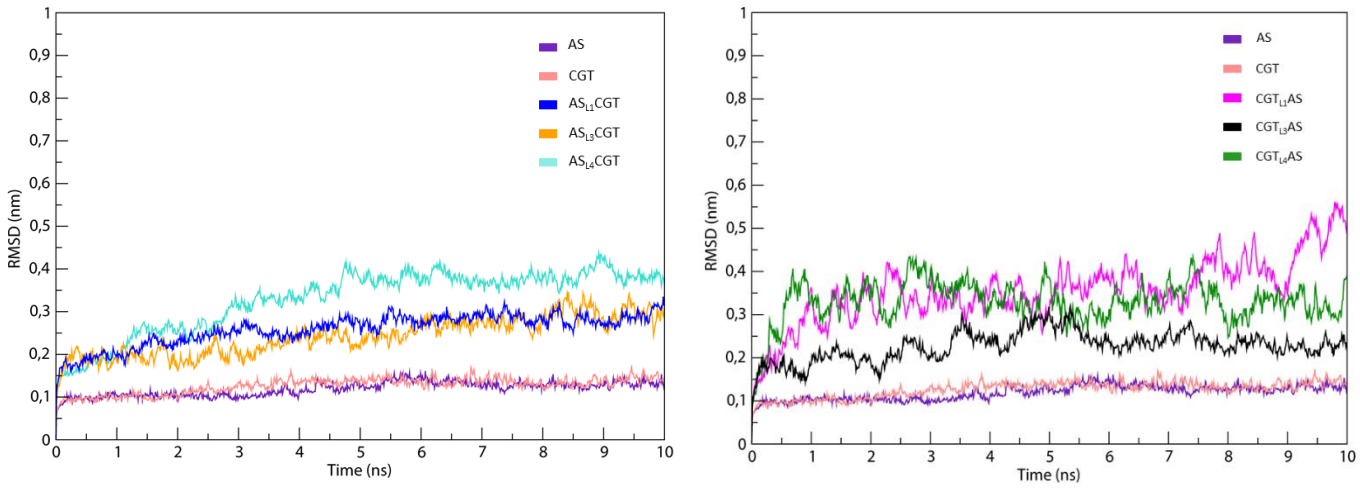
Con el fin de predecir la estabilidad estructural de los modelos viables, se llevaron a cabo simulaciones de dinámica molecular (Pikkemaat et al. 2002). La calidad de cada estructura fue evaluada mediante la trayectoria de RMSD y el Radio de Giro (RD) en 10 ns de simulación. En la Figura 7a se presentan las trayectorias de los modelos **AS**·[GGGGS]n·**CGTasa**, con respecto a las estructuras de partida (AS y CGTasa). Los controles (proteínas AS y CGTasa) presentaron un valor de RMSD de 0.16 ±0.016 nm y 0.17 ±0.018 nm, respectivamente. AS presentó una ligera fluctuación entre los 2.5 a 4.5 ns, mientras que CGTasa no presentó variaciones. El valor RMSD del modelo $AS_{L1}CGT$ indica que mantiene su equilibrio durante la trayectoria con un valor de 0.25

±0.036 nm, presentando pequeñas fluctuaciones aprox. a los 5 y 8 ns. $AS_{L3}CGT$ presentó un comportamiento similar al anterior, pero con mayores picos de fluctuación; el valor máximo de RMSD fue 0.345 ±0.04 nm. $AS_{L4}CGT$ fue el modelo que presentó mayor variación en los valores de RMSD. En la parte inicial, se observa un comportamiento ascendente entre 0–4.5 ns, con valores de 0.15 a 0.437 ±0.07 nm. Puede concluirse que, bajo los parámetros de la simulación, los modelos originales AS·CGTasa mantienen una trayectoria con pequeñas fluctuaciones, pero finalmente bastante estables. Los modelos con linkers cortos presentan una fluctuación un poco mayor que los controles, sin que el RMSD sufra una variación significativa. En cambio, el comportamiento conformacional otorgado por la región linker al modelo $AS_{L4}CGT$ refleja un aparente exceso de movilidad de las subunidades, proclive a la interferencia estérica, lo que podría resultar innecesario para el propósito de este trabajo.

Figura 7.

Trayectoria de estabilidad de los modelos 3D y las proteínas Molde.

a. b.



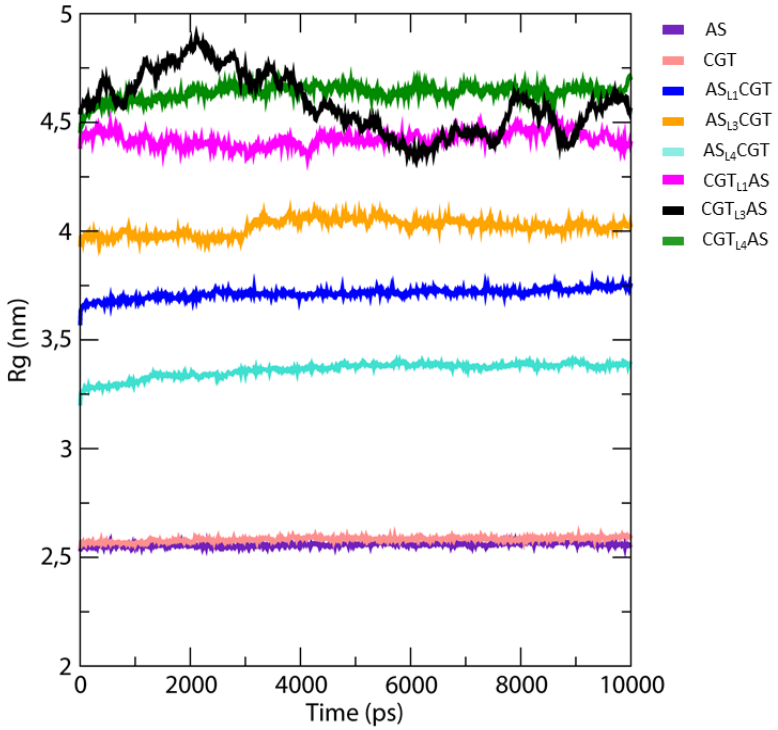
Nota: Estabilidad estructural de los modelos calculada mediante dinámica molecular. a) Trayectoria $AS_{L1}CGT$, $AS_{L3}CGT$, $AS_{L4}CGT$. b) $CGT_{L1}AS$, $CGT_{L4}AS$, $CGT_{L3}AS$.

Por el contrario, los modelos **CGTasa**·$[GGGGS]_n$·**AS** reflejaron una menor estabilidad estructural (Figura 7b); en efecto, los valores de RMSD fueron superiores y las trayectorias con picos irregulares durante los 10 ns de simulación. $CGT_{L1}AS$ y $CGT_{L4}AS$, presentaron la mayor variación con RMSD de 0.527 ±0.06 y 0.433 ±0.048, respectivamente. En las trayectorias se evidencian fluctuaciones que no terminan de estabilizarse. $CGT_{L3}AS$ presentó fluctuaciones entre 0–8 ns, con un valor máximo de RMSD = 0.319 ±0.034, al final de los cuales la tendencia aparente es a alcanzar la estabilidad. Curiosamente, los modelos **CGTasa**·$[GGGGS]_n$·**AS** presentaron *linkers* de menor tamaño, en comparación con el orden AS·CGTasa y, sin embargo, las fluctuaciones a nivel estructural fueron mayores. En conclusión, las simulaciones de dinámica molecular indican mayor estabilidad estructural para el ordenamiento AS·CGTasa y específicamente los modelos ASL_1CGT y $AS_{L3}CGT$.

El radio de giro Rg es un parámetro que relaciona la estabilidad de una proteína con respecto a su deformación y compactación en el tiempo (Lobanov, Bogatyreva, & Galzitskaya 2008). A medida que incrementa el valor del Rg la proteína es menos compacta. En la Figura 8, se comparan los valores de Rg de los modelos, en relación con los controles. AS y CGT presentaron valores de Rg = ±2.5 nm, que se mantuvieron estables en el tiempo. Aunque con valores superiores a los controles, los modelos $AS_{L4}CGT$ (3,25 nm) y $AS_{L1}CGT$ (3,7 nm), presentaron un menor nivel de compactación. Por el contrario, los modelos del ordenamiento **CGTasa**·$[GGGGS]_n$·**AS** exhibieron mayores grados de deformación. $CGT_{L3}AS$, que era un buen candidato en las pruebas de estabilidad de estructural, fue el de mayor variación en el radio de giro, presentando una mayor variación en la compactación.

Figura 8.

Radio de giro de los modelos y las proteínas molde.



Nota: Estabilidad estructural, Radio de giro de los modelos con respecto a las estructuras de AS y CGTasa. a) $AS_{L1}CGT$, $AS_{L3}CGT$, $AS_{L4}CGT$. b) $CGT_{L1}AS$, $CGT_{L4}AS$, $CGT_{L3}AS$.

Tomados en conjunto los resultados de dinámica molecular y radio de giro, los modelos AS·CGTasa presentaron mejor estabilidad y deformación limitada. Aunque, las simulaciones de dinámica molecular indican que los modelos $AS_{L1}CGT$ y $AS_{L3}CGT$ presentaron mejores patrones de estabilidad competitivos, el radio de giro de $AS_{L3}CGT$ no le es favorable. Por lo tanto, concluimos que el modelo $AS_{L1}CGT$ es el mejor para ser probado *in vitro*.

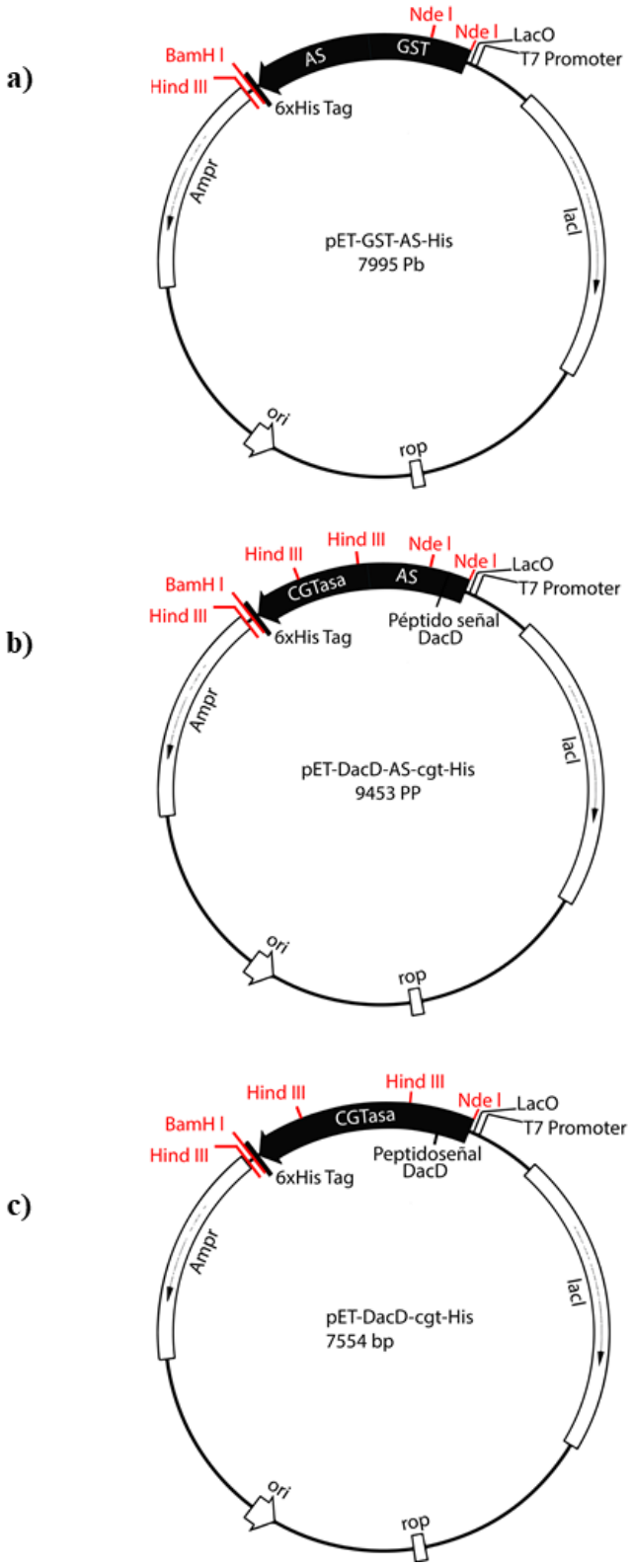
3.1.5 Diseño de los vectores de expresión recombinantes.

Las primeras acciones en el laboratorio estuvieron encaminadas al diseño de los ORFs que codificaran para cada una de las proteínas recombinantes. Se construyeron plásmidos recombinantes con genes sintéticos para expresar en *E. coli* los polipéptidos GST·**AS**·His, DacD·**CGTasa**·His y DacD·**AS·CGTasa**·His.

Los trabajos de expresión heteróloga de AS reportados hasta la fecha han recurrido a la fusión en 5’ de la AS con la Glutation S-Transferasa (GST), para aumentar la cantidad de AS recombinante (Skov et al. 2001). Además, una etiqueta de polihistidina (His·Tag) en 3’ ha sido útil a la hora de purificar la enzima con columna de afinidad (Emond et al. 2007; van der Veen et al. 2004). Atendiendo la experiencia previa de los autores, decidimos emular la fusión GST·**AS**·His en el diseño del plásmido recombinante. Para la construcción del ORF GST-AS (ver Apéndice 3.1), en 5’ del gen de la AS (GenBank: AJ011781.1) se incluyeron 687 pb que codifican para la GST, inspirados la secuencia homóloga del vector pGEX-6P (Genbank: AAB37352, locus U78874.1). Entre los sitios Nde I y Bam HI del vector pET-21a se clonó el ORF codificante para la fusión GST·**AS**·His, con los detalles reportados por Montalk et al., (1999) (Montalk et al. 1999), en cuanto a la eliminación del péptido señal de los primeros 10 residuos de la AS (Montalk et al. 1999).

Figura 9.

Esquema de construcción de los vectores. a) pET-GST-AS-His, b) pET-DacD-cgt-His, c) pET-DacD-AS-cgt-His



Además, se mutaron las posiciones 11 y 12 (I →N y L→S, respectivamente) (Montalk et al. 1999), que en nuestra secuencia adoptarían las posiciones N232 y S233 (Apéndice 3.1). En la Figura 9a se presenta el esquema general del plásmido recombinante pET-GST-**AS**-His (7995 pb).

Para la construcción de la secuencia codificante para la CGTasa, nos basamos en la secuencia con número de acceso GenBank M12777.1 (Apéndice 3.2). Trabajos recientes han demostrado que adicionar el péptido señal DacD a la CGTasa incrementa la secreción extracelular de la proteína recombinante en *E. coli*, en relación con el péptido señal original (Sonnendecker et al. 2017; K. Zhang et al. 2017b; Zheng, Li, & Wu 2019). En los sitios de restricción Nde I y Bam HI del vector pET-21a, se clonaron en orden consecutivo 67 pb que codifican para el péptido DacD, seguidos del gen cgt y la secuencia His·Tag. Fue necesario remover 27 residuos aminoacídicos que conforman el péptido señal original de la CGTasa de *B. macerans* y reemplazarlos por la región codificante para DacD (ver detalles en el Apéndice 3.3). Como resultado de la construcción, obtuvimos el vector recombinante denominado pET-DacD-**cgt**-His (Figura 9b).

Finalmente, el diseño de la fusión AS·CGTasa se basó en la secuencia aminoacídica del modelo **AS$_{L1}$CGT** (Apéndice 3.4). En este nuevo vector no se incluyó la GST, porque equivaldría a una triple fusión. No obstante, para estimular la producción de proteína recombinante hacia la fracción extracelular, incluimos el péptido señal DacD en la región N-terminal. En los sitios de restricción Nde I y Bam HI del plásmido pET-21a, se clonaron en orden consecutivo las secuencias codificantes para el péptido DacD, seguida por la fusión AS·CGTasa (separadas por el linker GGGS) y la secuencia His·Tag en el extremo 3'.

Para unir las secuencias AS y CGTasa, se eliminó el codón stop de la AS y se dio continuidad en fase con el gen cgt, mediante los nucleótidos que codifican para la región linker (Apéndice 3.4). El vector pET-DacD-**AS-cgt**-His (Figura 9c) contiene los genes AS y CGTasa fusionados con un solo marco de lectura abierto de 4055 pb que codifica para la proteína Fusión **ASL$_1$CGT**.

3.1.6 Verificación de la calidad de los constructos.

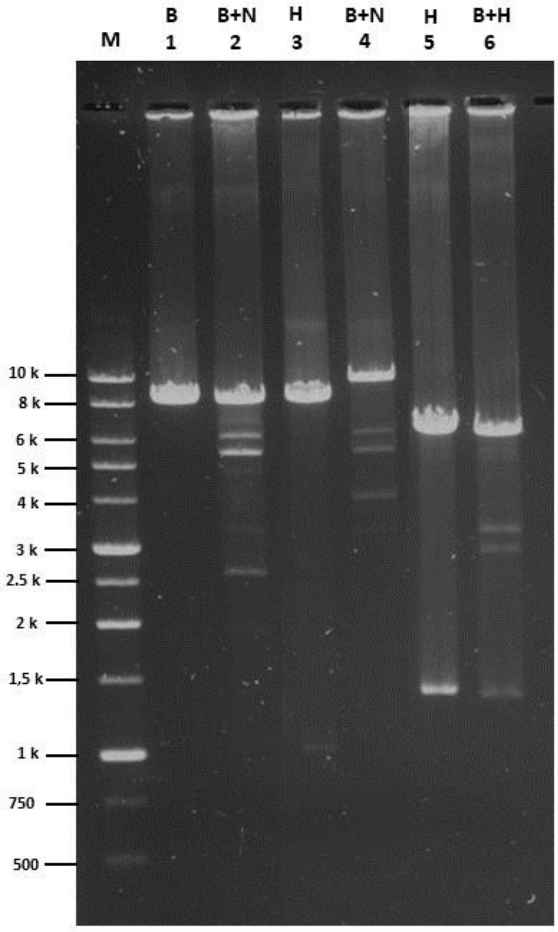
Para evaluar la expresión y funcionalidad de las proteínas, *E. coli* BL21 (DE3) fue transformada con cada uno de los vectores construidos, incluyendo pET-21a (5443 bp) no recombinante (NR), utilizado como control para los ensayos de actividad. Para comprobar la fidelidad de los constructos, se realizó un análisis de restricción con las enzimas de restricción Nde I, Bam HI y Hind III (Figura 11). Los carriles 1 y 2 corresponden al vector pET-GST-**AS**-His (tamaño de inserto 2597 pb). En el carril 1 se evidencia una banda de ~8 kb correspondiente al plásmido linearizado con la enzima Bam HI. En el carril 2 se observa la digestión del plásmido con las dos enzimas Bam HI y Nde I. La banda de ~5.4 kb corresponde al vector pET-21a y más abajo se aprecian los fragmentos de 2597 (GST-AS) y 1907 (AS). Varias bandas de menor brillo sugieren que la digestión del carril 2 es parcial. La más visible es una de 6 kb que pertenecería al vector pET-21a con el fragmento GST (i.e., 5443 + 687 pb).

En los carriles 3 y 4, se encuentran los patrones de digestión del plásmido pET–DacD-**AS-cgt**-His (tamaño de inserto 4055 pb), cortado con las enzimas Hind III y Bam HI + Nde I, respectivamente. El vector tiene tres sitios de restricción Hind III, dos de ellos dentro del gen CGTasa y uno en el polylinker (Figura 10). En el carril 3, se observa una banda >8 kb,

correspondiente a pET-21a con la mayor parte del DNA del inserto. Se logra apreciar una banda esperada de ~1 kb, formada por los 1014 pb que hay entre los dos sitios Hind III del gen CGTasa; en cambio, por razones de rango dinámico, no es visible la banda de 373 pb del fragmento CGTasa-polylinker. Del carril 4 se deduce que las digestiones con Bam HI + Nde I del plásmido pET–DacD-**AS-cgt**-His resultaron parciales. La banda cercana a los 10 kb hace referencia al plásmido linearizado con un solo corte. La banda de ~5.5 kb, corresponde al vector pET-21a y la de ~4 kb es el inserto de 4050 pb (fusión AS·CGTasa). Una banda adicional de ~ 6 kb podría ser la forma superenrollada del plásmido sin digerir.

Figura 10.

Análisis de restricción de los constructos.



Nota: 1. pET-GST-AS-His + *Bam* HI (B). 2. Corte con *Bam* HI (B) y *Nde* I (N). 3. pET-DacD-**AS-cgt**-His + *Hind* III, 4. Corte con *Bam* HI (B) y *Nde* I (N). 5. pET-DacD-cgt-His digestión con *Hind* III (H). 6. corte con *Bam* HI (B) y *Hind* III (H). M: DM3100 ExcelBandTM 1 kb DNA Ladder.

Finalmente, en los carriles 5 y 6 se observa el patrón de digestión del vector pET-DacD-**cgt**-His (tamaño de inserto 2157 pb). La banda que supera los 6 kb en ambos carriles corresponde al vector desprovisto de la porción que delimitan los sitios Hind III interno de la CGTasa y del polylinker (Figura 10). Asimismo, se observan los fragmentos Hind III de 1387 pb y Hind III + Bam HI 1362 bp, respectivamente.

También se emplearon las enzimas de restricción disponibles en el laboratorio, i.e., Bam HI y Hind III. Aunque las digestiones resultaron parciales y las enzimas empleadas fueron únicamente las que estaban disponibles en el laboratorio, es razonable concluir que los plásmidos son portadores de los insertos correctos. Sin embargo, para no dar lugar a la incertidumbre, los tres plásmidos fueron enviados a secuenciar a Macrogen y los resultados coincidieron con nuestros diseños.

3.2 Sobreexpresión de las enzimas recombinantes en *E. coli*

Para esta sección, con el fin de facilitar la lectura, nombraremos las proteínas recombinantes solo como AS, CGTasa y AS·CGTasa. En esta primera parte, evaluamos la expresión y funcionalidad de las enzimas utilizando los sustratos convencionales sacarosa y almidón, para cada una de las reacciones individuales (hidrólisis y ciclización). Lo anterior, con el fin de determinar

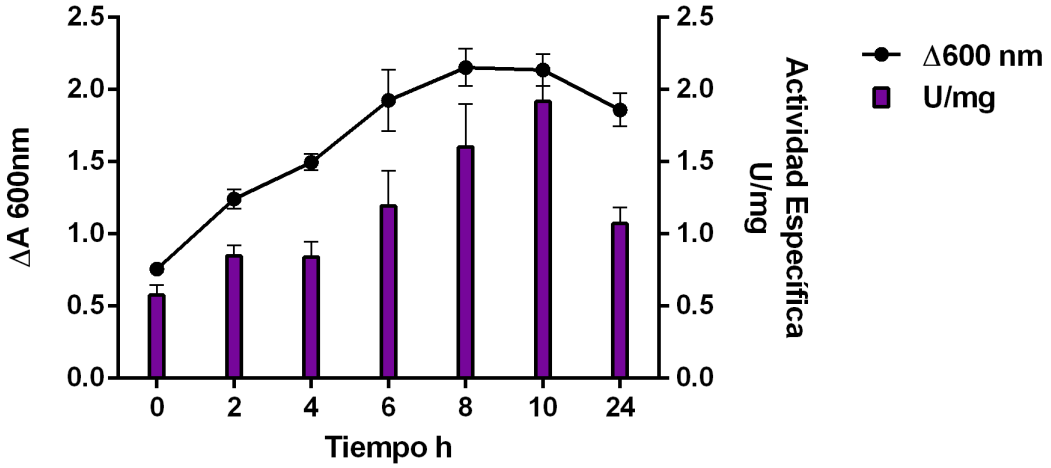
condiciones reacción funcionales para, finalmente, probarlas con los lixiviados de mucílago de cacao.

3.2.1 Expresión en E. coli del vector pET-GST-AS-His

En conformidad con la literatura, la proteína de fusión AS se caracteriza por ser de tipo intracelular, cuando se expresa en *E. coli* (Albenne, Skov, et al. 2002; Montalk et al. 1999). Por lo tanto, para corroborar la expresión y funcionalidad de AS, se monitoreó la actividad de hidrólisis de sacarosa de la fracción intracelular (FIC). La cepa de *E. coli* transformada con pET-GST-**AS**-His fue cultivada en medio LB e inducida con 1 mM IPTG (Emond et al. 2007; Joucla et al. 2004)**.** En primer lugar, se evaluó la actividad *in vitro* de AS en función de la expresión, después de añadir el inductor. Se tomaron fracciones del extracto FIC a 0, 2, 4, 6, 8, 10 y 24 h. Se adoptaron condiciones de reacción basados en la literatura. Para evaluar la actividad hidrolítica de sacarosa, se empleó el ensayo de DNS para detección de azúcares reductores (i.e., fructosa producida). En la Figura 11, se observa la curva de crecimiento (ΔA 600 nm) de la cepa de *E. coli BL21* y, en paralelo, la actividad hidrolítica de sacarosa expresada como actividad específica (AE), en la fracción intracelular (FIC). Como puede apreciarse, la curva de crecimiento del clon pET-GST-AS-His crece exponencialmente, alcanzando un pico máximo a las 8 h (ΔA 600 nm 2.15 ± 0.074) y una disminución al llegar a las 24 h (ΔA 600 nm 1.8 ± 0.067). Los valores de actividad específica (AE) evidencian que la producción de la AS aumenta exponencialmente hasta alcanzar un máximo a 10 h de inducción (AE = 1.912 ± 0.25 U/mg). El comportamiento de la curva de crecimiento y los valores de actividad evidencian una relación directa entre el crecimiento bacteriano y la producción enzimática de AS. Para futuros experimentos, se definió un tiempo de inducción de 10 h antes de proceder con las extracciones.

Figura 11.

Curva de crecimiento de E. coli transformada con pET-GST-AS-His vs actividad especifica de cada tiempo de inducción.

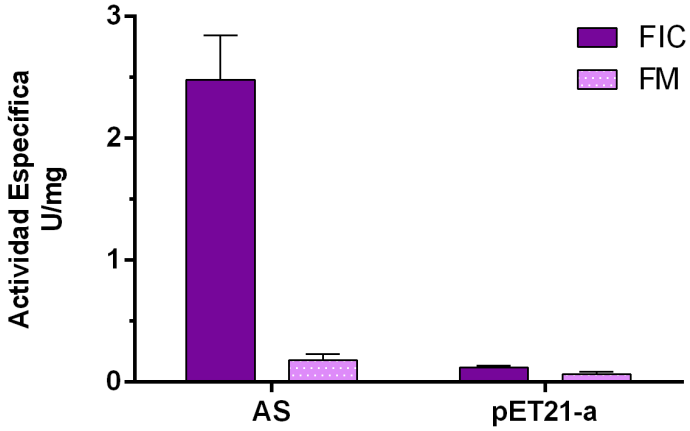


Nota: Eje Y izquierdo: crecimiento bacteriano medido con ΔA 600 nm (•). Eje Y derecho: actividad específica (AE) en U/mg de AS, medida de la AE de hidrólisis de sacarosa. Eje X: tiempos de toma de muestras post-inducción con IPTG (barras moradas). Condiciones de reacción: Extractos crudos de *E. coli* inducida con 1 mM IPTG. 1X Buffer fosfato pH 7.0 y 0.1 M sacarosa, a 40 °C durante 10 min de reacción (Ryu et al. 2010).

Para verificar si había presencia de AS en el medio de cultivo, el mismo ensayo se aplicó a la fracción extracelular (FM) (Figura 12). Como control negativo, se llevaron a cabo los mismos ensayos de actividad AS con las fracciones FIC y FM de *E. coli* transformada con pET-21a. En la Figura 12, se puede observar que la actividad de hidrólisis de AS para la fracción FIC (AE = 2.48 ± 0.36 U/mg) fue significativamente mayor a la obtenida con la fracción FM (p=0.025). En esta última, el valor de AE para AS fue de 0.18 ± 0.048 U/mg, es decir, mínima. Como estaba previsto, las muestras de pET-21a no presentaron valores significativos de actividad AS. Los resultados coinciden con la literatura, en que la AS se produce mayormente de manera intracelular en *E. coli*.

Figura 12.

Actividad específica (AE) de hidrolisis de sacarosa de extractos crudos FIC y FM de E. coli BL21 transformada con pET–GST-AS-His y pET-21a.



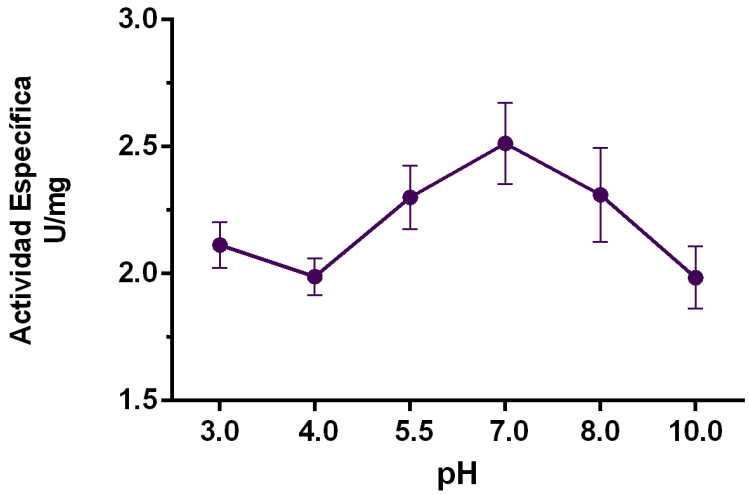
Nota: Condiciones: Buffer fosfato pH 7.0 y 0.1 M sacarosa, 40 ºC durante 10 min de reacción.

Trabajos previos sugieren que el pH óptimo para la hidrólisis de sacarosa por la AS es de 7.0 (Koh et al. 2016; G. P. De Montalk et al. 1999; Pizzut-Serin et al. 2005; Ryu et al. 2010). Para llevar a cabo la reacción combinada de la AS y la CGTasa en la producción de CDs, era necesario definir un pH que favoreciera las dos reacciones simultáneamente. Por consiguiente, evaluamos el efecto del pH sobre la actividad AS en las condiciones de nuestro laboratorio. En la Figura 13, se presenta la actividad específica de AS en función del pH (rango 3 a 10) y se observa que la actividad máxima se alcanza a un pH de 7.0 (2.51 ± 0.160 U/mg), tal como lo reportan los autores anteriormente citados. El segundo pico de AE se presenta a pH 5.5 y 8.0 con valores de AE relativamente altos (2.30 ± 0.125 y 2.31 ± 0.185 U/mg), respectivamente (Figura 13). Estas cifras de AE resultan importantes en el marco de este proyecto, ya que indican que la AS podría funcionar de manera favorable en un rango de pH más amplio (i.e., 5.5 a 8.0), en caso de ser necesario un

ajuste de las condiciones de reacción para la reacción combinada. En general, podemos concluir que la actividad de AS no se ve significativamente afectada por el cambio de pH (p = 0,089), aunque 7.0 puede ser considerado el pH óptimo para futuras reacciones.

Figura 13.

Influencia del pH sobre la actividad específica de hidrolisis de sacarosa de AS



Nota: Condiciones de reacción, extractos crudos de FIC de *E. coli* transformada con pET–GST-AS-His con 0,1 M de sacarosa de AS.

3.2.2 Expresión en E. coli de pET-DacD-cgt-His

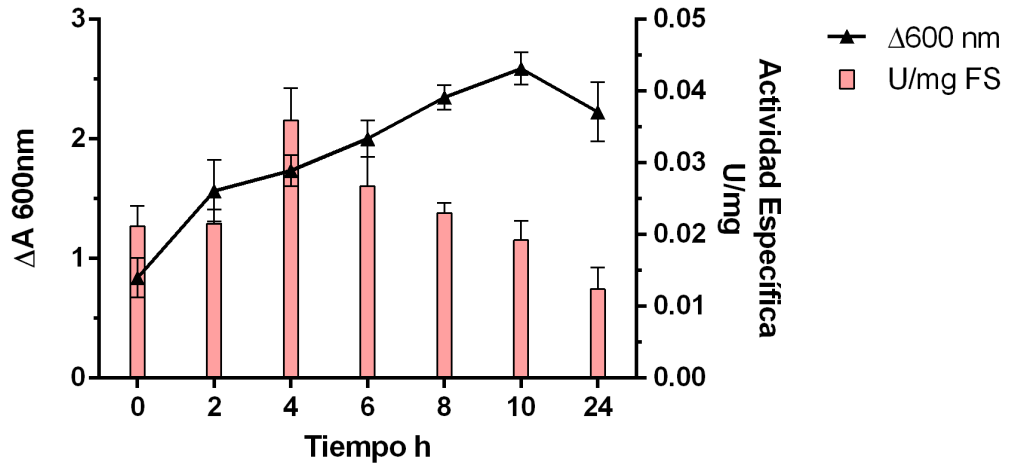
El vector pET–DacD-**cgt**-His incluyó el péptido señal DacD para la exportación de CGTasa a la fracción extracelular. No obstante, nuestras primeras pruebas de expresión no dieron resultados de actividad aceptables en esa fracción, posiblemente debido a la formación de cuerpos de inclusión, un fenómeno reportado con anterioridad para esta enzima (Jeang 1999; Li et al. 2010). Por esta razón, fue necesario reajustar las condiciones de expresión y el protocolo de extracción

de proteínas, incluyendo una metodología para solubilizar las proteínas mediante 2% Triton-100x y diversos pasos de sonicación y centrifugación (ver materiales y métodos). De este método de extracción surge la fracción soluble a la que denominamos FS. Para la expresión de la enzima CGTasa, se tuvo en cuenta adicionar 0.5 mM IPTG como publicado previamente (Sonnendecker et al. 2017). Inicialmente, evaluamos la producción de CDs mediante la actividad de ciclización de la CGTasa, utilizando extractos FS y almidón como sustrato.

Como con el vector anterior, determinamos la expresión de la proteína recombinante en función del tiempo, después de añadir el inductor. Los tiempos recolectados de FS post-inducción fueron 0, 2, 4, 6, 8, 10 y 24 h. Las condiciones de reacción adoptadas para la actividad de ciclización fueron : Tris-HCl pH 8.0, 1.5% almidón como sustrato control, a 45 °C durante 10 min de reacción (Koh et al. 2016; Li et al. 2009). En esta parte inicial, nos enfocamos en evaluar la producción de β-CDs, ya que éstas presentan un mayor interés comercial (Mourtzinos & Goula 2019; Sebaaly et al. 2018) y para este proyecto. La producción de β-CDs se evaluó mediante el método colorimétrico de fenolftaleína (Makela et al. 2002). En la Figura 14, puede apreciarse el paralelo entre la curva de crecimiento (ΔA 600 nm) de la cepa de *E. coli* transformada con pET–DacD-**cgt**-His y la actividad de ciclización de la fracción FS, expresada en U/mg. La curva de crecimiento de la cepa evidencia un patrón exponencial con un valor máximo de ΔA 600 nm = 2.58 ± 0.078 nm a las 10 h, el cual disminuye al alcanzar las 24 h. No obstante, la producción de enzima CGTasa en la fracción FS presentó un máximo de producción a 4 h de inducción (0.037 ± 0.04 U/mg), cuando ΔA 600 nm estuvo por encima de 2. Por consiguiente, adoptamos este tiempo para proceder a los extractos de FS, ya que como lo muestra la Figura 14, a partir de 4 h la enzima pierde su actividad.

Figura 14.

Curva de crecimiento de E. coli transformada con pET-DacD-cgt-His vs actividad específica de ciclización de la CGTasa con almidón.



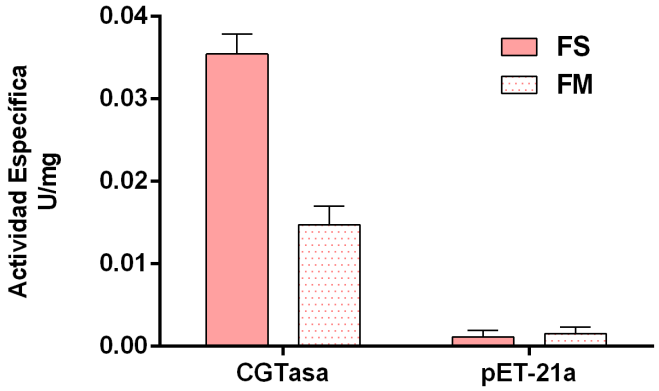
Nota: Eje Y izquierdo: crecimiento bacteriano medido con ΔA 600 nm (▲). Eje Y derecho: actividad específica (AE) en U/mg de CGTasa (actividad de ciclización). Eje X: tiempos de toma de muestras post-inducción con IPTG (barras rosadas). Condiciones de reacción: Extractos crudos de *E. coli* inducida con 0,5 mM IPTG. Sustrato 1.5% almidón, pH 8.0, 30 min de reacción.

La caída de la AE después de 4 h podría reflejar exportación de la enzima al espacio extracelular, degradación o muerte celular. Por lo tanto, se llevaron a cabo pruebas de actividad de extractos de medio de cultivo (FM), a 10 h de inducción. Como en el caso anterior, los extractos FS y FM de pET-21a no recombinante fueron utilizados como control negativo. En la Figura 15, se puede apreciar un comparativo de la AE de ciclización entre la FS y la FM, de la CGTasa aquí producida. El valor de la AE de la FS superó significativamente a la FM ($p = 0,025$). Sin embargo, la fracción FM presentó una actividad de ciclización ostensible con un AE = 0.015 ±0.004 U/mg, equivalente a ≈42% de la actividad encontrada en la fracción FS (AE = 0.035 ± 0.004 U/mg). Los extractos del control negativo pET-21a no presentaron valores de AE, confirmando que la

actividad de ciclización detectada corresponde netamente a la CGTasa recombinante. Por lo tanto, podemos concluir que la CGTasa se produce y es funcional en las fracciones FS y FM, y además logra producir CDs utilizando como almidón sustrato.

Figura 15.

Actividad específica de extractos crudos FS y FM de la CGTasa.

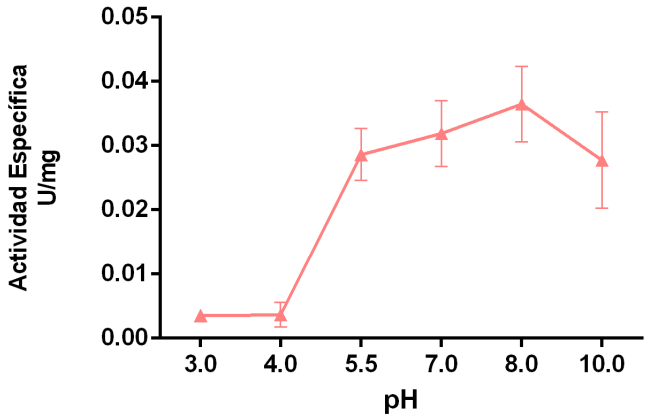


Nota: Condiciones de reacción: 1.5% Almidón como sustrato, pH 8.0, temperatura 45 ºC durante 30 min de reacción.

Finalmente, al igual que con la AS, evaluamos el efecto del pH sobre la actividad de ciclización de la CGTasa, con el fin de encontrar un valor de pH que favoreciera la reacción combinada AS + CGTasa. Como se puede apreciar en la Figura 16, la CGTasa presenta actividad de ciclización a partir de pH 5.5, alcanzando un valor máximo de AE a pH 8.0 (0.036 ± 0.006 U/mg). El valor de AE CGTasa a pH 7.0 (0.032 ± 0.005) equivale a 89% de la AE de la AS a pH 8.0, indicando que este rango de pH podría funcionar al momento de llevar a cabo la reacción combinada.

Figura 16.

Influencia del pH sobre la actividad específica de ciclización de almidón de la CGTasa.



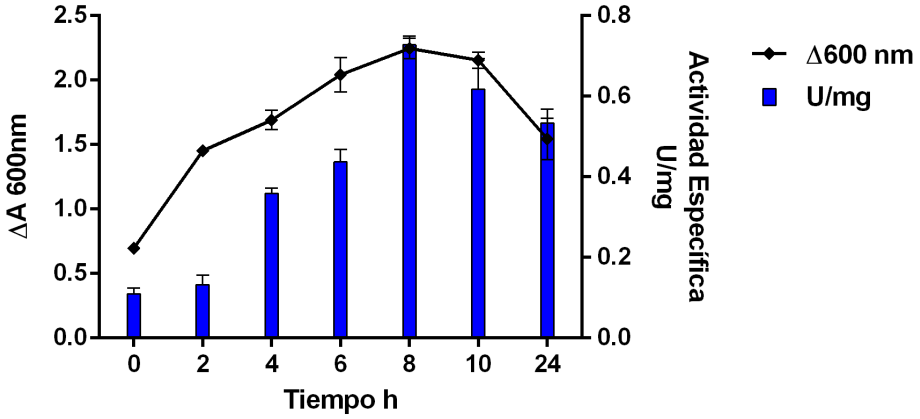
Nota: Extractos de FS de *E. coli* transformada con pET–DacD-cgt-His con 1.5% de Almidón

3.2.3 Sobreexpresión de pET-DacD-AS-cgt-His y actividades AS y CGTasa en extractos crudos

En primer lugar, evaluamos la expresión de la fusión AS·CGTasa, en función del tiempo de inducción con IPTG. Decidimos detectar solo la actividad AS, bajo el supuesto de que reflejaría la presencia simultánea de las dos enzimas AS y CGTasa, ya que están fusionadas. Por consiguiente, se realizaron pruebas de la función individual AS mediante hidrólisis de sacarosa, siguiendo las mismas condiciones de reacción previamente utilizadas para la AS. La curva de crecimiento de la Figura 17 evidencia un comportamiento exponencial alcanzando un valor máximo de ΔA 600 nm = 2.247 ± 0,078 a 8 h y luego empieza a decrecer. Por otro lado, durante los primeros tiempos de inducción (0 y 2 h) la actividad de la enzima es baja, pero a partir de 4 h aumenta exponencialmente hasta alcanzar un máximo a 8 h de inducción 0.728 ± 0.022 U/mg. Por lo tanto, para análisis posteriores, partimos de extractos de FS de 8 h de inducción.

Figura 17.

Curva de crecimiento de E. coli transformada con pET–DacD-AS-cgt-His vs actividad específica de hidrólisis de sacarosa.



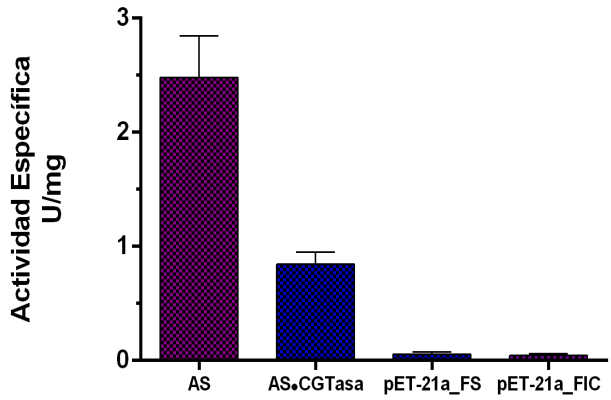
Nota: Eje Y izquierdo: crecimiento bacteriano ΔA 600 nm (♦). Eje Y derecho: actividad específica (U/mg) de AS (hidrólisis de sacarosa). Eje X: tiempos de toma de muestras post-inducción con IPTG (barras azules). Condiciones de reacción: Extractos crudos de *E. coli* inducida con 0.5 mM IPTG. 1X Buffer fosfato pH 7.0 y 0.1 M sacarosa, a 40 °C, durante 10 min de reacción.

Posteriormente, se procedió a evaluar las actividades individuales de AS y CGTasa, empleando extractos de FS tomados a 8 h post-inducción. La actividad de hidrólisis de la subunidad AS ya había sido demostrada, así que seguimos las mismas condiciones de reacción, que eran compatibles con la CGTasa. En la Figura 18 se aprecia un paralelo en la actividad hidrolasa de sacarosa, entre extractos crudos de *E. coli* con las enzimas recombinantes AS (fracción FS) y fusión AS·CGTasa (fracción FIC). La AS·CGTasa presentó el 34% de la actividad específica de hidrólisis de sacarosa de la AS ($p=0.05$). En vista de que el contraste de actividades enzimáticas se realizó entre extractos crudos, la comparación resulta improcedente. De hecho, los valores netos

de producción de azúcares reductores no son coherentes con ese 34%, ya que la fusión AS·CGTasa produce aprox. 5 mM de fructosa, mientras que la AS sola presentó un valor de 7 mM, i.e., el 34% ascendería a 71,4% (Tabla 4). En la discusión se tratará este punto; sin embargo, para el propósito final del proyecto, es válido concluir que las dos enzimas se expresan y son activas, lo suficiente para los ensayos futuros.

Figura 18.

Actividad específica de hidrólisis de sacarosa de las enzimas recombinantes la AS y la fusión.



Nota: AS·CGTasa (fracción FIC, azul) y AS (fracción FS, morado). Condiciones de reacción: extractos crudos de *E. coli* transformada con pET–GST-**AS**-His, pET–DacD-AS-cgt-His y pET-21a NR (control); 0.1 M sacarosa, pH 7.0, 40 ºC, 10 min reacción.

Tabla 4.

Valores de actividad específica y producción de azúcar reductor (mM) mediante la reacción de hidrólisis de 0.1 M sacarosa, de la AS y AS·CGTasa.

Sacarosa	Fracción	Proteína (mg/ml)	µmol/min	AE (U/mg)	*Azúcar reductor**
AS	FIC	0.580 ± 0.07	1.436 ± 0.1	2.48 ± 0.36	7 ± 0.01
AS·CGTasa	FS	1.096 ± 0.2	0.911 ± 0.2	0.84 ± 0.1	5 ± 0.00

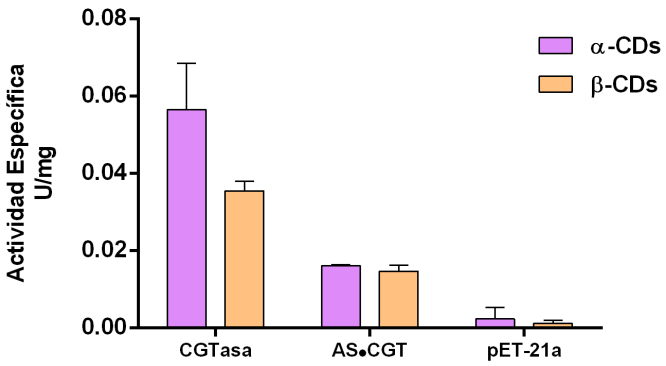
Nota: * Azúcar reductor en mM

Por otra parte, para determinar la actividad de ciclización de la subunidad CGTasa de la fusión, se evaluó la producción de α-CDs y β-CDs, utilizando almidón como sustrato control (Figura 19, Tabla 5). En este caso, también son evidentes los menores valores de la AE de la fusión (AE α-CDs = 0.016 – 631.2 nmol, AE β-CDs = 0.015- 613.6 nmol) en relación con la CGTasa (AE α-CDs = 0.056 – 1206.3 nmol, AE β-CDs = 0.035 – 751.1 nmol), pero como hemos dicho, esta comparación podría ser engañosa. En cambio, podemos, concluir que sí se detecta una actividad CGTasa, lo que significa que la fusión está cumpliendo su propósito bifuncional.

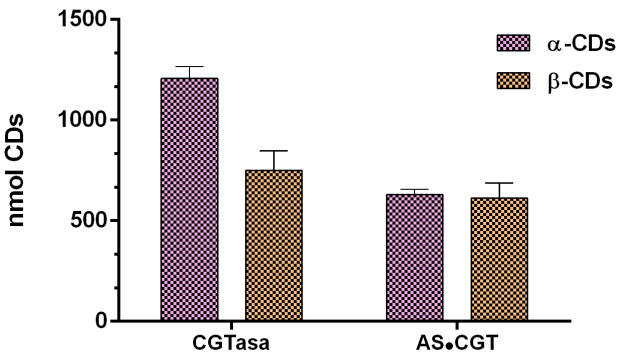
Figura 19.

Actividad específica de ciclización para producir α-CDs y β-CDs (FS) de la CGTasa, la fusión AS·CGTasa

a)



b)



Nota: pET-21a NR (control negativo) con 1.5% de almidón como sustrato. Condiciones de reacción: 1.5% almidón, pH 8.0, 45 ºC, 30 min. α-CDs (naranja) y β-CDs (rosado). a) Actividad expresada en actividad especifica (AE); b) producción de CDs expresada en nmol totales.

En cuanto a la producción de CDs, en la Tabla 5 podemos apreciar que, para la CGTasa, la producción de α-CDs (1.235 mg/ml) fue mayor en comparación de las β-CDs (0.918 mg/ml). Este comportamiento era de esperarse, ya que la CGTasa de *B. macerans* produce en su mayoría α-CDs (Urban et al. 2012). No obstante, para la AS·CGTasa, la proporción en la producción de α-CDs (0.642 mg/ml) y β-CDs (0.673 mg/ml) no representó diferencias significativas (p = 0,564). A partir de estos resultados podemos concluir que la proteína AS·CGTasa es funcional a nivel de actividades individuales (hidrólisis y ciclización), con sustratos convencionales.

Tabla 5.

Actividad específica de la CGTasa (fracción FS) y de la fusión AS·CGTasa (fracción FS) en la producción de α-CDs y β-CDs mediante la reacción de ciclización con 1.5% almidón.

	Proteína (mg/ml)	***µmol/min***		***AE (U/mg)***		***CDs (mg/ml)***	
		A	**β**	**α**	**β**	**α**	**β**
CGTasa	0.610 ± 0.09	0,034 ±0,001	0.022 ± 0.017	0,056 ±0,007	0.036 ± 0.003	1.235 ± 0,06	0.918 ± 0.08
AS·CGTasa	1.096 ± 0,2	0,018 ±0,008	0.016 ± 0.008	0,016 ±0,005	0.015 ± 0.002	0.642 ± 0,03	0.673 ± 0.04

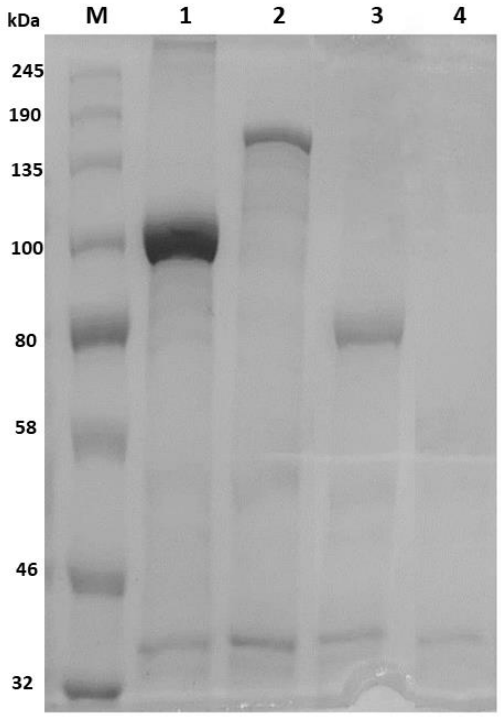
3.2.4 Visualización de las proteínas recombinantes en geles SDS-PAGE

Antes de proceder a la producción combinada de CDs en una sola reacción, un control natural sería visualizar las proteínas recombinantes en geles SDS-PAGE. La AS presentó mejor actividad en la fracción FIC y las otras dos proteínas también, aunque más baja. Finalmente, optamos por hacer la separación electroforética con dicha fracción, pues todas las proteínas evidenciaban actividad en la FIC y el perfil de bandas se haría más comparable (Figura 20). En el carril 1 (AS), se destaca una fuerte banda cerca de 100 kDa que coincide con el peso molecular de la proteína (96 kDa) y que refleja altos niveles de expresión. En el carril 2 es evidente una banda de menor intensidad entre 135 y 190 kDa, la cual corresponde al peso de la fusión AS·CGTasa (150 kDa). Finalmente, en el carril 3 se puede apreciar una banda de ~ 80 kDa, que pertenecería a la proteína CGTasa, ya que el peso molecular concuerda plenamente. Desafortunadamente, no contamos con anticuerpos para revelar las proteínas mediante *western blot;* sin embargo, los pesos moleculares esperados coinciden perfectamente con lo esperado. Es de precisar que las bandas en

cuestión no figuran en los extractos de pET21-NR, luego provienen de la expresión de los vectores (control, carril 4).

Figura 20.

Gel SDS-PAGE de las fracciones FIC de E. coli recombinante, inducida con IPTG.



Nota: Carril 1: AS; Carril 2: fusión AS·CGTasa; Carril 3: CGTasa; Carril 4: pET-21a (control negativo). M: Referencia de peso molecular 11 a 245 kDa (*Prestained Protein Standard, Broad Range, BioLabs*).

3.3 Producción de CDs en una sola reacción a partir de sacarosa como sustrato

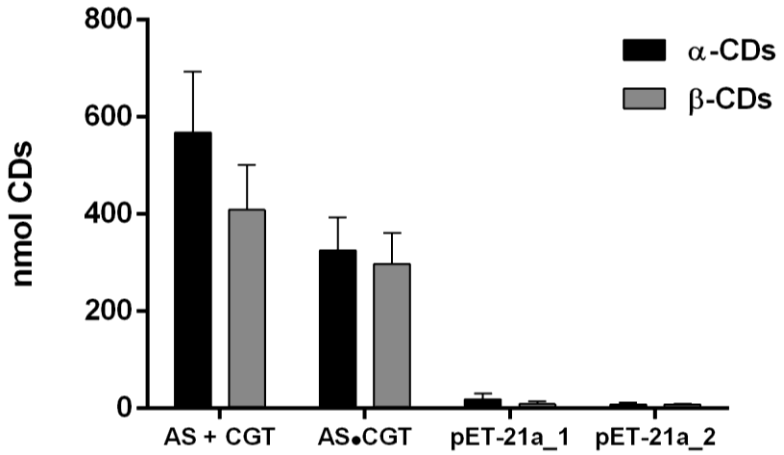
Comprobada la correcta expresión y funcionalidad de cada enzima, procedimos a evaluar la producción de α- y β-CDs con sacarosa como sustrato control, en una reacción mixta de AS +

CGTasa y, en otra, con la proteína fusión AS·CGTasa. Para la reacción combinada, se emplearon extractos crudos conteniendo que mostraron mejor actividad: FIC para la AS (10 h de inducción, 0.1 mM IPTG) y FS para la CGTasa (4 h de inducción, 0.5 mM IPTG). Por otro lado, para los ensayos con la fusión AS·CGTasa, se utilizaron extractos FS de 8 h de inducción, 0.5 mM IPTG. Las pruebas fueron realizadas a pH 7.5, por ser un valor intermedio entre los óptimos aquí determinados (Figuras 14 y 17). En la Figura 21 se muestran los resultados de la síntesis de CDs en una sola reacción en nmoles totales. Cuando se emplearon las enzimas libres (AS + CGTasa), la reacción favoreció la producción de α-CDs (0.55 mg/ml, i.e., 568.4 nmol) y un poco menos de β-CDs (0.46 mg/ml, i.e., 408.8 μmol^{-1} min^{-1}), sin diferencias significativas (p=0.76). La fusión AS·CGTasa, que logró sintetizar CDs en una sola reacción, presentó un comportamiento similar: produjo un poco más de α-CDs = 325.35 nmol, que de β-CDs = 297.35 nmol, sin diferencias significativas (p=0.982). Por tanto, los resultados exhiben las mismas tendencias de actividades individuales AS y CGTasa, con almidón como sustrato control (Tabla 5).

Los resultados entre las dos reacciones AS + CGTasa y AS·CGTasa sí presentaron diferencias significativas (p = 0.0028), lo cual es reflejo, entre otros factores, del número de unidades enzimáticas funcionales que se producen en cada sistema de expresión. En efecto, las enzimas individuales superaron a la fusión en ~ 22-32% de producción de CDs. Sin embargo, también hay que tener en cuenta que se trata de una reacción no optimizada en sus condiciones fisicoquímicas, fuera de los alcances de este proyecto. Se puede entonces concluir que parte del objetivo principal se halla cumplido, ya que logramos construir una fusión AS·CGTasa que sintetiza CDs en una sola reacción, si bien el sustrato continúa siendo el control (sacarosa).

Figura 21.

Producción de α-CDs y β-CDs, por reacción combinada de AS + CGTasa y de la Fusión AS·CGTasa con sacarosa.



Nota: Fracciones: AS (FIC) + CGTasa (FS), Fusión AS·CGTasa (FS). Síntesis de CDs expresada en nmoles totales de producción. Condiciones: 0.1 M sacarosa, pH 7.5, 45 ºC. Se emplearon extractos FIC (pET-21a_1) y FS (pET-21a_2) del plásmido no recombinante, como control negativo.

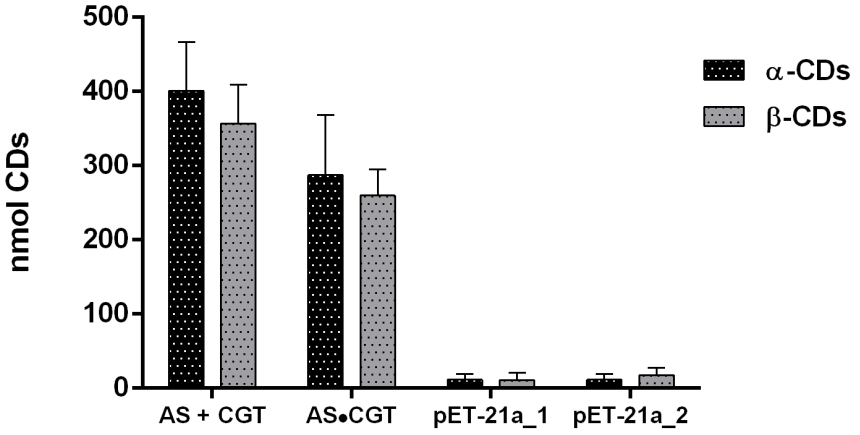
3.4 Producción de CDs en una sola reacción a partir de lixiviados de mucílago de cacao como sustrato.

Para evaluar la producción de CDs a partir de lixiviados de mucílago de cacao, objetivo principal de este trabajo, se realizaron los mismos ensayos de actividad en una sola reacción. Se usaron las mismas condiciones que para la actividad con sacarosa, pero en su lugar se adicionó 1 ml de lixiviados de mucílago de cacao, clon CCN-51. El lote de mucílago presentó esta concentración de azúcares totales (g/L): 86.95 g/L glucosa (0.48 M), 90.86 g/L fructosa (0.5 M) y 0.27 g/L sacarosa (0.0078 M). En la Figura 22, podemos apreciar la primera evidencia de producción de CDs en una sola reacción, a partir de lixiviados de mucílago de cacao. Las

concentraciones de CD obtenidas en la reacción combinada AS + CGTasa fueron de α-CDs = 400.9 nmol y β-CDs = 356.8 nmol, sin diferencias significativas entre las proporciones (p=0.30). Por otra parte, con la fusión AS·CGTasa fueron de α-CDs = 287.3 nmol y β-CDs = 259.9 nmol, sin diferencias significativas entre α-CDs y β-CDs (p=0.43). La combinación AS + CGTasa superó a la fusión AS·CGTasa en ≈23%; no obstante, este incremento no representa diferencias significativas (p = 0.11).

Figura 22.

Producción de α-CDs y β-CDs por reacción combinada de AS + CGTasa y de la fusión AS·CGTasa con lixiviados de mucílago de cacao como sustrato.



Nota: Fracciones: AS (FIC) + CGTasa (FS) y de la fusión AS·CGTasa (FS). síntesis expresada en nmoles totales de producción. Condiciones: 1 ml de mucílago, pH 7.5, 45 ºC. Se emplearon extractos FIC y FS de pET-21a no recombinante como control negativo.

Los resultados evidencian que la reacción combinada AS + CGTasa fue más eficiente en la producción de CDs con los dos sustratos (sacarosa y mucílago), en comparación con la proteína

fusión. Las reacciones con lixiviados de mucílago de cacao presentaron una leve disminución en la producción de CDs en relación con la sacarosa, lo cual puede atribuirse a una baja concentración de sacarosa presente en el mucílago. No obstante, el mucílago de cacao contiene una amplia variedad de componentes orgánicos e inorgánicos potencialmente inhibitorios, que deben estudiarse por separado (Ortiz & Álvarez 2015).

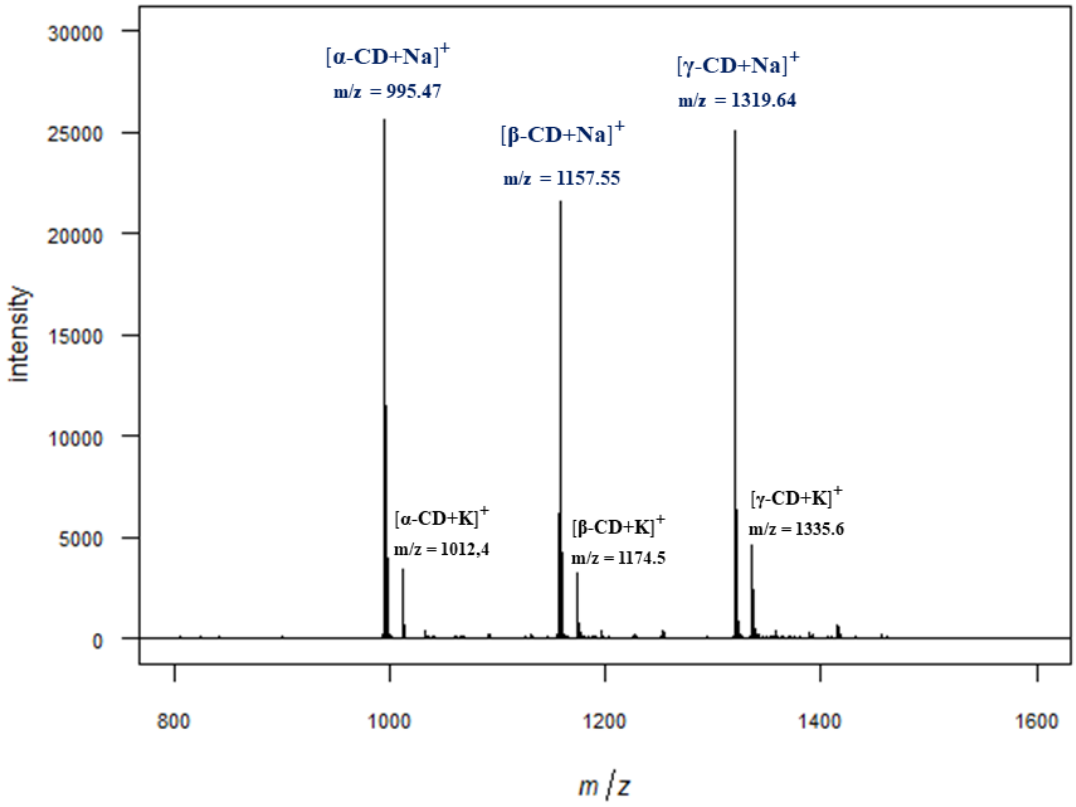
A partir de los resultados obtenidos podemos concluir que sí es posible la producción de CDs por acción conjunta de la AS y la CGTasa, utilizando mucílago de cacao como sustrato. De esta conclusión deriva que el diseño teórico de la proteína de fusión DacD·**AS·cgt**·His fue acertado y se expresa en cantidades suficientes en *E. coli* para realizar pruebas con pequeñas cantidades de extractos crudos.

3.5 Detección de CDs por espectrometría de masas (MS-MALDI-TOF)

Hasta el momento, la detección de CDs en las reacciones se ha efectuado con los métodos de fenolftaleína y naranja de metilo. Para obtener una evidencia más precisa de la presencia de CDs, recurrimos al análisis de los productos enzimáticos mediante MS-MALDI-TOF. Se analizaron los productos de las reacciones catalizadas por las enzimas recombinantes de este trabajo: CGTasa con almidón, AS + CGTasa con sacarosa y AS·CGTasa con mucílago de cacao. Una mezcla de CDs comerciales fueron analizados con anterioridad, para comparar las masas moleculares de las CDs: α = 972.84 Da, β = 1135 Da y γ = 1297 Da. En la Figura 23, se puede apreciar el espectro de masas y la abundancia relativa de la mezcla de CDs comerciales α, β y γ, con tres picos mayores que corresponden a los aductos con Na^+ (23 Da) y tres picos de menor intensidad con el aducto con K^+ (39 Da).

Figura 23.

Espectro de masas (MALDI-TOF-MS) de estándares comerciales de CDs α, β y γ.

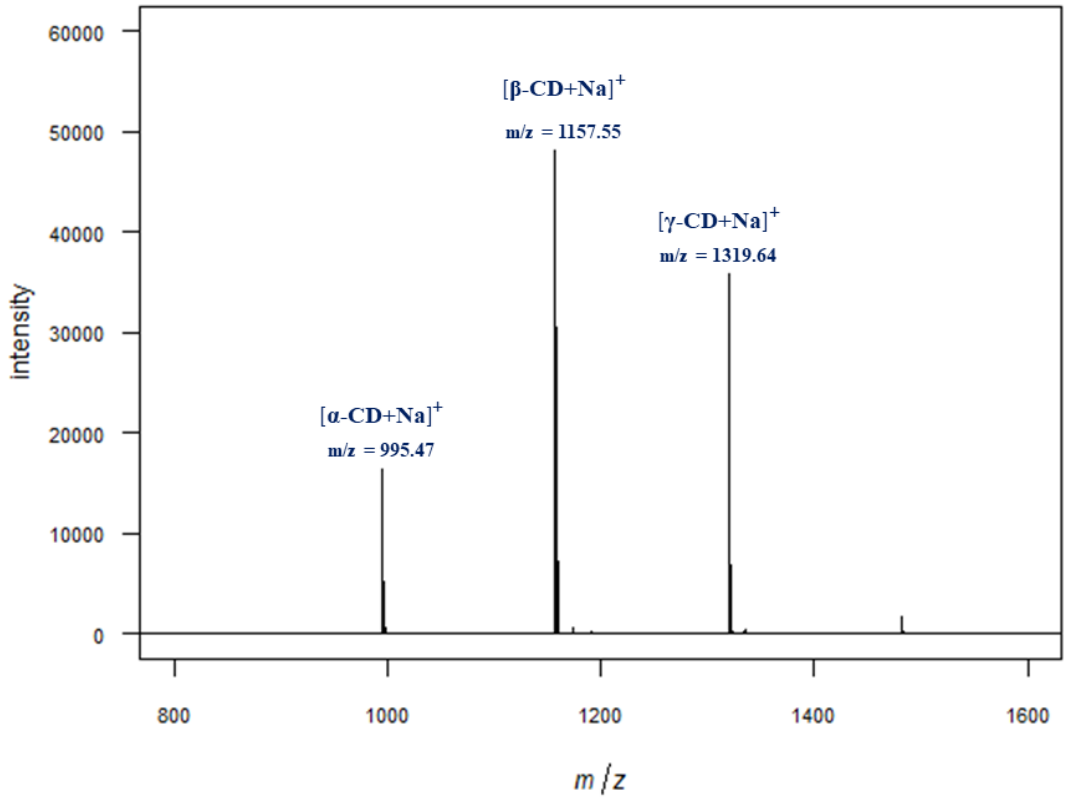


Nota: En el eje x se aprecia el peso en relación masa/carga (m/z) y en el eje Y se observa la intensidad de la señal de los iones. El valor m/z de cada pico es equivalente al peso molecular de cada CD junto con la masa de los iones sodio (Na+, 23 Da) y potasio (K+, 39 Da).

En el espectro del producto generado por la reacción de la CGTasa con almidón (Figura 24), se pueden apreciar tres señales: α-CD m/z = 995.47, β-CD m/z = 1157.55 y γ-CD m/z = 1319, todas con el aducto con Na^+.

Figura 24.

Espectro de masas (MALDI-TOF-MS) del producto de la reacción de la CGTasa con almidón como sustrato.

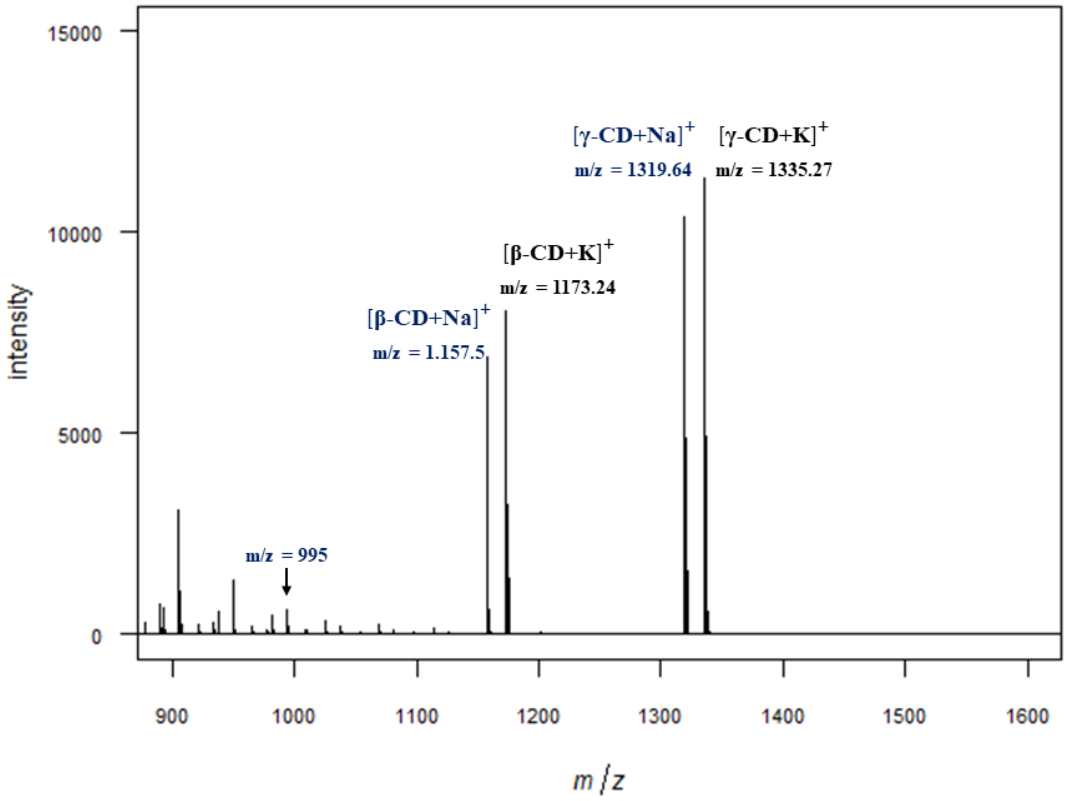


Nota: El m/z de cada pico es equivalente al peso molecular de cada CD junto con la masa del ion sodio (Na^+), i.e., 23 Da.

Por otro lado, la reacción combinada AS + CGTasa con sacarosa como sustrato, arrojó un espectro un tanto diferente, con cuatro señales de mayor intensidad correspondientes a β-CD (m/z= 1157.5, m/z= 1174.24) y γ-CD (m/z= 1319.64, m/z= 1335.27), ambas con aductos con Na^+ y K^+ (Figura 25). Una señal de baja intensidad con m/z= 995 posiblemente haga referencia a la α-CD con el aducto Na^+.

Figura 25.

Espectro de masas (MALDI-TOF-MS) del producto de la reacción de la AS + CGTasa con sacarosa como sustrato.



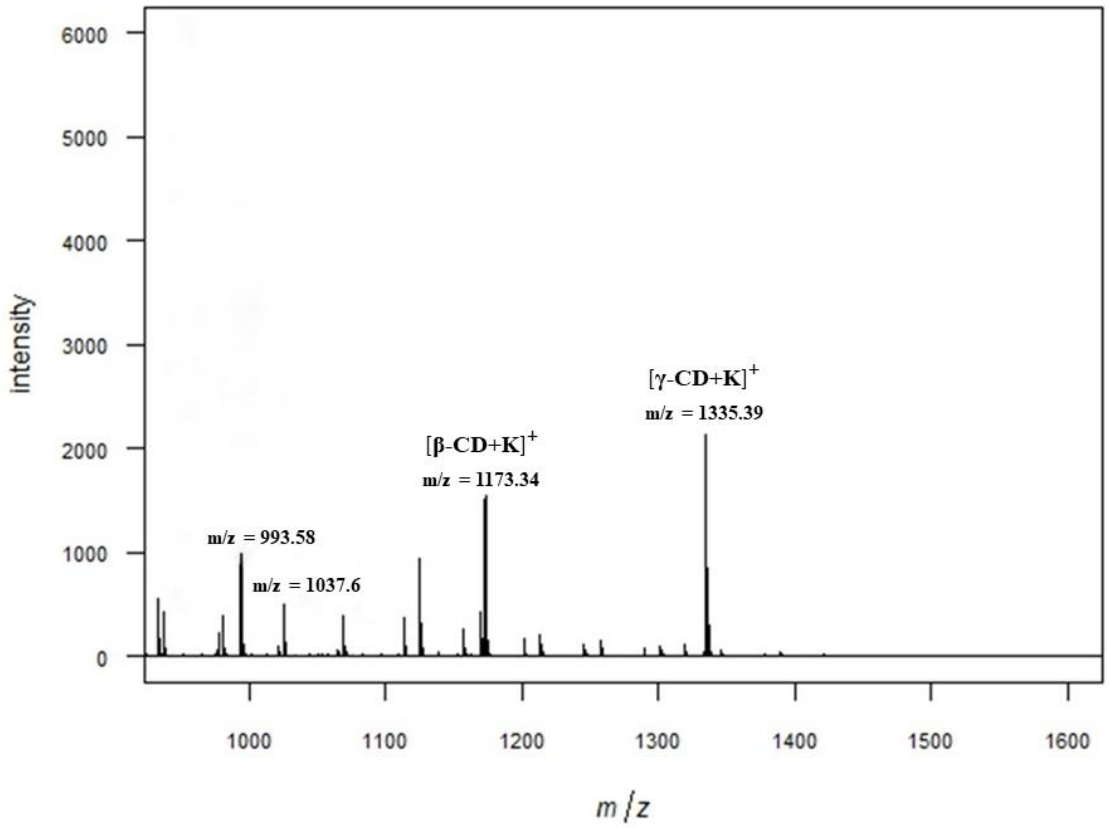
Nota: El m/z de cada pico es equivalente al peso molecular de cada CD junto con la masa de los iones sodio (Na^+), i.e., 23 Da y potasio (K^+), i.e., 39 Da.

Finalmente, la producción de CDs en la reacción con la fusión AS·CGTasa y lixiviados de mucílago de cacao se evidencia en el espectro de la Figura 26. Al igual que la reacción combinada AS + CGTasa, se observan dos señales de mayor intensidad, pero con valores relativos menores, que corresponden a la β-CD (m/z= 1173.34) y γ-CD (m/z= 1335.39), en un aducto con Na^+. Interesantemente, no hubo presencia de señal para la α-CD. A partir de los resultados obtenidos

por espectrometría de masas, confirmamos la presencia irrefutable de CDs y, por tanto, todos los resultados previos quedan íntegramente validados.

Figura 26.

Espectro de masas (MALDI-TOF-MS) del producto de la reacción de la fusión AS·CGTasa con lixiviados de mucílago de cacao como sustrato.



Nota: El m/z de cada pico es equivalente al peso molecular de cada CDs con la masa con el ion potasio (K^+), i.e., 39 Da.

4. Discusión

En el presente trabajo de grado se logró construir una proteína de fusión entre la AS de *N. polysaccharea* y la CGTasa de *B. macerans.* Se demostró la funcionalidad de las dos subunidades en la producción de CDs en una sola reacción, utilizando sacarosa y mucílago de cacao como sustrato. Los resultados obtenidos son de gran relevancia en las áreas de biotecnología y la ingeniería de proteínas (Kaulpiboon et al. 2010).

Las predicciones bioinformáticas fueron fundamentales para el diseño; en efecto, permitieron anticipar las probabilidades de éxito en la construcción de la fusión. Sin embargo, únicamente las pruebas *in vitro* podrían aclarar preguntas pendientes. ¿Cada subunidad adoptaría la conformación 3D suficiente para hacerlas funcionales? ¿La región *linker* aportaría la distancia conveniente? ¿Qué tan estable sería la fusión AS·CGTasa? ¿Qué impacto tendría la fusión en los sitios activos? ¿Surgiría una conformación cuaternaria no prevista? Algunas de estas preguntas hallaron respuesta en este trabajo, con la funcionalidad demostrada de cada subunidad. No obstante, otras siguen sin respuesta, ya que están por fuera del alcance de este proyecto. Sin duda, La purificación de la fusión y la determinación de su estructura tridimensional aclarará estos y otros interrogantes. Por el momento, los resultados de este trabajo demuestran que es posible fusionar las dos proteínas y que pueden ejercer su función catalítica de manera simultánea. Además, la capacidad de la fusión de actuar sobre varios sustratos (almidón, sacarosa y mucílago de cacao) permite imaginar nuevos e interesantes experimentos orientados al uso biotecnológico de la fusión.

El diseño y construcción de una proteína fusión representa un gran desafío, pues el éxito funcional de las subunidades dependerá de diversos factores: la orientación espacial de cada subunidad, la interferencia mutua entre subunidades, el acceso al sustrato de los sitios activos y las posibles vías intermoleculares de reacción, entre otros (Rizk, Antranikian & Elleuche, 2012; Yu et al., 2015). En general, a la hora de construir proteínas de fusión, se le ha dado preferencia a la evaluación directa *in vitro* antes que *in silico* (Yu et al., 2015). En la actualidad, las herramientas bioinformáticas actuales permiten anticipar con precisión los problemas de diseño. En este sentido, las predicciones computacionales que pudimos realizar nos brindaron un buen acercamiento al comportamiento estructural de la fusión AS·CGTasa.

Una de las principales preocupaciones era la correcta selección del *linker* entre subunidades. La composición aminoacídica, longitud y conformación estructural del *linker* (Figura 2a) estuvo en consonancia con lo establecido por Argos (1990) y George & Heringa (2002) para proteínas de fusión. Los *linkers* inspirados en las proteínas encontradas en la naturaleza adoptan estructuras secundarias (α-hélices, hojas β o bucles –no estructurados–) y logran ejercer su función (Chen, Zaro & W.-C. Shen, 2013). En nuestro caso, basados en los testimonios de muchos autores y las predicciones *in silico*, decidimos adoptar una región espaciadora presente en la mayoría de linkers naturales, con la idea subyacente de favorecer una suficiente flexibilidad (Argos, 1990). En este sentido, la contribución de pequeños aminoácidos no polares como Gly y polares como Ser, Thr y Asn (Figura 2) lograron la flexibilidad esperada del *linker* en soluciones acuosas (Chen, Zaro & W. C. Shen, 2013). Este comportamiento se vio reflejado en los análisis de dinámica molecular, por la estabilidad de las trayectorias de los modelos evaluados. A título de comparación, estudios de dinámica molecular para la AS (Champion et al. 2012) y CGTasa (Chen et al. 2015),

presentan patrones de trayectoria similares a los aquí obtenidos, con valores de RMSD dentro del rango estable.

Estudios recientes reportan que, en algunos casos, la cercanía de los centros activos en las fusiones mejora la eficiencia de la reacción global. El producto de la primera reacción es activa o pasivamente transportado al segundo sitio activo por canales intramoleculares (Huang, Holden, & Raushel 2001; Iturrate et al. 2010; Lim et al. 2016; Raushel, Thoden, & Holden 2003). Aunque no existe una pauta de la distancia ideal entre sitios activos, si la hay entre la longitud de los canales y la distancia de cada centro catalítico. La longitud promedio de los canales estaría entre 15 y 30 Å, con un máximo razonable de 127 Å (i.e., catalasa de *Penicillium janthinellum*) (Pravda et al. 2014). La fusión AS·CGTasa (modelo $AS_{L1}CGT$) presentó una distancia aproximada de 59 Å entre las regiones internas de los sitios catalíticos (ver Apéndice 4a). De ser posible la formación de canales en la fusión AS·CGTasa en la producción de CDs, la longitud promedio de los sitios activos estaría dentro esperado (Pravda et al. 2014). Otro de los posibles mecanismos involucrados en la cascada de reacciones en proteínas de fusión es por difusión libre. El producto de la primera reacción alcanza la saturación y es liberado por la parte externa de la proteína, hacia el segundo sitio activo (Pettersson 2001). Sería válido especular que la libre difusión de producto–sustrato dentro de la AS·CGTasa podría direccionarse desde la parte superficial del extremo inferior de la de AS (C-terminal) hacia la CGTasa (Figura 2 y Apéndice 4b).

Para evaluar la acción combinada de la AS y la CGTasa en la producción de CDs *in vitro*, iniciamos con ensayos individuales de cada enzima. La hidrólisis de sacarosa por la fracción FIC, acorde con una localización intracelular, fue similar a lo reportado para la AS de *N. polysaccharea*

(Albenne et al. 2002; Büttcher et al. 1997) El pH óptimo de la AS en la reacción de hidrólisis de sacarosa debe estar cercano a la neutralidad (Montalk, et al. 2000; Albenne et al. 2002). No obstante, no hay un reporte de la influencia del pH sobre la reacción con sacarosa. Nuestros resultados indican que la actividad hidrolítica de sacarosa puede funcionar a pH ácido y alcalino, aunque el mayor pico de actividad es en efecto, cercano a pH 7.0 (Figura 13).

En la construcción de la CGTasa se adicionó el péptido señal DacD (Figura 9c) y para la sobreexpresión se variaron las condiciones i.e., 0.5 mM de IPTG, medio TB y la temperatura de incubación post inducción se disminuyó a 25 ºC, con la expectativa de favorecer la secreción extracelular de la proteína recombinante en *E. coli* (Sonnendecker et al. 2017). Sin embargo, la actividad de ciclización en la FM representó solo el 42% de la actividad total. La mayor parte de la actividad fue obtenida con la fracción FS, después de una etapa de solubilización de la proteína. El resultado sugiere la formación de cuerpos de inclusión, los cuales bloquean los canales de translocación e impiden la secreción de la proteína (Zheng, Li & Wu, 2019). De hecho, este comportamiento ya fue reportado en la sobreexpresión de la CGTasa en *E. coli* (Kim et al. 2000; Kweon, Kim, and Seo 2004). Por otro lado, Zheng et al. (2019) demostraron que a pesar de que DacD es el péptido señal más efectivo para la producción de CGTasa extracelular, la mayor parte de la proteína sigue siendo retenida en el citoplasma insoluble. La influencia del pH en los parámetros cinéticos de la CGTasa, con almidón como sustrato, mostró un amplio rango de pH (5.5 a 10), con actividad óptima a pH 8.0 (Figura 16). Por tanto, nuestros resultados son consistentes con estudios previos de expresión recombinante en *E. coli*, de la CGTasa de *B. macerans* (Li et al., 2010; Yang, Shan, & Wang 2017).

El diseño del vector pET–DacD-**AS-cgt**-His también incluyó el péptido señal DacD, con las mismas expectativas de secreción extracelular. No obstante, bajo las condiciones evaluadas en este trabajo, las actividades individuales AS y CGTasa de la fusión fueron mayoritarias en los extractos FS, i.e., fracción intracelular. Por analogía con la CGTasa, es válido atribuir este comportamiento a la formación de cuerpos de inclusión. Así, la solubilidad de la fusión será un asunto que deberá estudiarse más en detalle para mejorar la expresión y las pruebas de actividad. Interesantemente, las actividades individuales de hidrólisis (AS) y ciclización (CGTasa) de la AS·CGTasa con sacarosa (Figura 18) y almidón (Figura 19) demostraron que cada una mantuvo su función individual. Este resultado es importante porque se podrán realizar reacciones específicas para cada subunidad, a partir de un solo extracto de la fusión, según las nuevas preguntas de investigación que surjan en el futuro.

Las actividades específicas tanto en la hidrólisis de sacarosa como de ciclización con almidón disminuyeron notablemente en la fusión (Figuras 18 y 19), en comparación con los valores reportados para la AS y la CGTasa (Figuras 12 y 14). No obstante, si se observa la producción neta de azúcares reductores (Tabla 4) y concentración de α y β–CDs (Tabla 5), los valores a comparar no son tan distantes. Esto se debe a que los cálculos se basaron en la concentración de proteína total de los extractos, la cual varió entre el extracto de FS (AS·CGT y CGTasa) y FIC (AS). Además, se debe tener en cuenta que los extractos crudos de la proteína de interés están contaminados con otros polipéptidos propios de *E. coli*. Así, los valores de proteína total se incrementan y podrían estar afectando negativamente los valores de actividad específica. La evaluación futura de la actividad específica de la proteína purificada podría aclarar este punto.

Una de las precauciones más destacadas al diseñar una proteína fusión es la similaridad funcional, perfiles de temperatura y pH cercanos, de las subunidades (Rizk, Antranikian, & Elleuche 2012; Yu et al. 2015). Dos elementos fueron claves para el éxito de la reacción dual de la AS + CGTasa y en fusión AS·CGTasa para la producción de CDs: la compatibilidad de las condiciones fisicoquímicas inferida en la caracterización de las actividades individuales (AS y la CGTasa) y el hecho de que ambas pertenecen a la familia de glucósido hidrolasas. Las condiciones de reacción para la síntesis de CDs en una sola reacción (pH 7.5, 40 ºC) fueron acertadas, si bien podrían estudiarse con más detalle. Lo importante, en el marco de este trabajo, era demostrar que era posible producir las CDs en una sola reacción, pero sin duda, un nuevo trabajo debería consagrarse a la optimización de la reacción. En cuanto al tiempo de reacción en la producción de CDs, es de reconocer que ya existía al menos un reporte previo exitoso, i.e., 24 h (Koh 2016). Por esta razón, adoptamos esta condición como punto de partida.

Concentraciones bajas de sacarosa favorecen la producción CDs, siendo 0.1 M la concentración con mayor rendimiento en la reacción combinada de AS + CGTasa (Koh et al. 2016). Inspirados en este trabajo, lo más conveniente era mantener la concentración de 0.1 M sacarosa para evaluar nuestras reacciones de AS + CGTasa y AS·CGTasa. Las pruebas de actividad con sacarosa demostraron que AS·CGTasa era capaz de sintetizar CDs proporciones similares de α y β–CDs en una sola reacción (Figura 21). Aunque la reacción combinada AS + CGTasa presentó una mayor producción de CDs (α = 568.4 nmol, β = 408.8 nmol) en comparación con AS·CGTasa (α = 325 nmol, β = 297 nmol), sería apresurado atribuir esta diferencia a fallas estructurales o funcionales en la fusión. En efecto, las condiciones de reacción para la fusión no han sido optimizadas, asunto que deberá ser resuelto en futuros trabajos.

Los resultados de las pruebas de actividad de la Figura 22 dan cumplimiento al objetivo general de este trabajo. Específicamente, quedó suficientemente demostrado que el tratamiento dual con AS + CGTasa y la fusión AS·CGTasa son capaces de producir CDs utilizando sustratos no convencionales como los lixiviados de mucílago de cacao. Al comparar las reacciones con sacarosa y mucílago de cacao, la sacarosa resultó, por ahora, un mejor sustrato (Figura 21 y 22). Esta diferencia podría ser consecuencia de los bajos niveles de sacarosa en el mucílago (7.8 mM inicial, 0.025 mM final), lo que dificulta un poco compararlos con 0.1 M sacarosa. Lo interesante es que, a pesar de que el porcentaje de sacarosa en el mucílago es mínimo, la síntesis de CDs no se vio tan gravemente afectada. En efecto, la producción de CDs con sacarosa superó solo en ~10–20% la producción con mucílago de cacao. Estos resultados nos hacen sospechar que la presencia de glucosa libre (480 mM) podría favorecer la producción de CDs y a la vez arroja nuevas expectativas una vez se logre la optimización de las condiciones fisicoquímicas más favorables. La AS hidroliza sacarosa para liberar glucosa y fructosa (Montalk et al. 2000). Posteriormente, por reacciones de transglicosilación –o polimerización, para algunos autores– (Tian et al. 2018; Van Der Veen et al. 2006), la AS utiliza la glucosa libre para catalizar la síntesis de maltooligosacáridos como los α-(1,4)-glucanos lineales (Jung et al. 2020; Tian et al. 2018). A su vez, estos productos son el sustrato de la CGTasa para la síntesis de CDs (Kim et al. 2011; Koh et al. 2016). Sobre la base de esta ruta metabólica, resulta viable concebir que los valores superiores de glucosa libre del mucílago jueguen un rol activo en la etapa de polimerización. Así, las unidades libres de glucopiranosa también servirían de insumo para la síntesis de α-(1,4)-glucanos, aumentando la producción de CDs. Esta hipótesis se basa en la observación previa de formación del ‘complejo D-glucosa’ evidenciado en estudios estructurales de AS de *N. polysaccharea* (Mirza et al. 2001).

Durante la formación del complejo, una glucosa se une al 'subsitio-1' del sitio activo (Glu328, Asp393, His187, His392 y Arg284) y otra glucosa se integra a la 'interfaz de empaquetamiento', formando así una conformación favorable para la polimerización de α-(1,4)-glucanos (Mirza et al. 2001; Skov et al. 2001). Por otra parte, otros estudios demuestran la versatilidad de la AS para admitir otros donadores y aceptores de las diferentes reacciones que cataliza (Jung et al. 2020; Park et al. 2016; Rha et al. 2020). Por ejemplo, la adición de fructosa a la reacción de AS + sacarosa facilitó la producción de turanosa por isomerización, gracias a la fructosa extra (Park et al. 2016). Esto refuerza la idea de la posible incorporación de glucosa libre en la ruta enzimática.

Las proporciones relativas de α y β–CDs variaron entre tratamientos. En la reacción de la CGTasa con almidón (Figura 19), la producción de α-CDs fue mayor que de las β-CDs, lo cual es congruente con la literatura (Ding et al. 2010; Jeang 1999). Del mismo modo, en las pruebas de actividad de la reacción combinada AS + CGTasa con sacarosa y mucílago, la producción de α-CDs superó ligeramente a la β–CDs, al igual que lo reportado por Koh (2016). Por último, los ensayos de actividad de la fusión AS·CGTasa con almidón, sacarosa y mucílago indican que las α y β–CDs se sintetizaron casi en igual proporción. Este fenómeno es normal y se ha demostrado con anterioridad que las condiciones de reacción enzimática (tiempo de incubación, temperatura y concentración de sustrato) (Rendleman 1999), así como el tipo de sustrato o fuente de almidón influye directamente en las proporciones de α, β y γ CDs sintetizadas (Urban et al. 2012).

Un resultado totalmente imprevisto pero muy interesante resultó ser que, aunque los análisis de espectrometría de masas (MALDI-TOF) revelaron la presencia de las CDs, la señal de α-CD fue ínfima en el producto AS + CGTasa y ausente en la fusión. No sobra recordar que la

presencia de α–CDs se revela por un método colorimétrico con naranja de metilo (Lejeune et al. 1989) y las ß–CDs por otro método colorimétrico con fenolftaleína (Makela et al. 2000). Ambas técnicas colorimétricas son de amplio uso y están reconocidas por los principales autores que han trabajado en la producción enzimática de CDs (Ismail et al. 2011; J. Makela et al. 2002; Jiang et al. 2018; Lejeune, Sakaguchi, & Imanaka 1989; Pendergast & Connors 1984; Qi & Zimmermann 2005). También hay que aclarar que en este trabajo no estaba prevista la detección colorimétrica de γ–CDs. Lo intrigante es que el pico de α–CDs sí aparece en la reacción de la CGTasa con almidón, pero está ausente en las reacciones AS + CGTasa con sacarosa y AS·CGTasa con lixiviados de mucílago de cacao, como sustratos. En otras palabras, ¿cómo explicar que las α–CDs son detectadas por colorimetría, pero no por MALDI-TOF? Lo primero que deberíamos mencionar es que en nuestros laboratorios ya se habían obtenido resultados similares (Mercado Villamizar & Pinzón Joya 2020), sin que se hubiera ofrecido explicación alguna. También podemos confiar en la alta precisión de los resultados de MALDI-TOF, por lo cual no se debe tratar de un asunto metodológico. En cambio, debería considerarse la posibilidad de una degradación selectiva de las α-CDs por parte de la CGTasa. En efecto, la enzima es capaz de ciclizar la glucosa en CDs, pero también de catalizar la reacción inversa, es decir hidrolizar las CDs (Terada et al., 2001; Koh et al. 2016). Koh et al. reportaron en su trabajo lo siguiente: "*However, the amount of produced CDs became lower after 48 h due to CGTase-catalyzed hydrolysis of α-(1,4)-linkages, which represents a minor effect of this enzyme*". Interesantemente, las muestras que se procesaron con el MALDi-TOF se conservaron más de una semana por razones de un operario y un turno disponibles. De esta manera, habría habido tiempo suficiente para la acción hidrolítica de la CGTasa. Más interesante aún, resulta la selectividad en la degradación de las α-CDs, lo que constituye una nueva pregunta de investigación de gran interés para el futuro.

Tomados en conjunto los resultados aquí mostrados, es válido concluir que se logró construir una proteína de fusión AS·CGTasa bifuncional. También, que la fusión es capaz de catalizar la síntesis de CDs en una solo reacción, lo cual abre inmensas perspectivas para minimizar los pasos y los costos de la producción industrial de CDs. Finalmente, el lixiviado de mucílago de cacao resultó ser un sustrato promisorio para producir CDs y evitar al mismo el impacto negativo ambiental que tiene botarlo como desecho. No obstante, nuestros resultados son apenas un punto de partida para estudios futuros, pero los primeros pasos se han dado aquí. En este sentido, la síntesis de CDs a partir de un sustrato natural como el lixiviado de mucílago de cacao, podría aportar un valor agregado importante en la línea de producción. Este método implica un proceso más natural, que podría ser incorporado en la industria alimenticia, incluida en la línea de producción de alimentos derivados del cacao. Eso sí sin descartar la posibilidad de experimentar con otros sustratos naturales ricos en azúcares.

5. Conclusiones

- Los análisis bioinformáticos demostraron que, a nivel predictivo, la fusión AS·CGTasa adquirió una conformación estructural estable, formada por dos dominios grandes bien definidos y separados por una región linker.
- La fusión de la AS y la CGTasa, no afectó drásticamente la funcionalidad de los centros activos de cada proteína. Los resultados obtenidos en las pruebas de actividad indican que la AS·CGTasa es una proteína bifuncional, ya que los dominios AS y CGTasa cumplen cada uno con sus actividades individuales, con sustratos convencionales.
- La AS·CGTasa se comporta como una multienzima que es es capaz de sintetizar CDs en un solo paso mediante las reacciones acopladas de la AS y la CGTasa, a partir de sustratos ricos en azucares como los lixiviados de mucílago de cacao y sacarosa.
- Se comprobó la capacidad de la reacción combinada de los extractos crudos de AS (FIC) + CGTasa (FS) para la síntesis de CDs a partir de mucílago de cacao y sacarosa bajo las condiciones de reacción: Buffer Tris-HCl pH 7.5, temperatura 40ºC, 24 h de incubación. La producción de CDs de la reacción combinada fue mayor en relación con la fusión AS·CGTasa.
- El mucílago de cacao puede ser utilizado como sustrato para la obtención de CDs, a partir de la reacción combinada de de AS + CGTasa y la proteína fusión AS·CGTasa, gracias a su aporte de oligosacáridos y monosacáridos. Los resultados obtenidos en este trabajo son prometedores, ya que demostraron la obtención de CDs a partir de este sustrato alternativo con las enzimas recombinantes aún en condiciones no optimizadas.

6. Recomendaciones

1) Purificar las proteínas AS (pET·GST·AS·His), CGTasa (pET·DacD·cgt·His), y AS·CGTasa (pET·DacD·AS·cgt·His).

2) Determinar la estructura 3D de la proteína AS·CGTasa (pET·DacD·AS·cgt·His)

3) Caracterizar las condiciones expresión (Medio de cultivo, concentración de IPTG y TºC y tiempo de de incubación) de pET·DacD·AS·cgt·His.

4) Determinar las condiciones óptimas de producción de CDs con la fusión AS·CGTasa, para los sustratos mucílago de cacao y sacarosa.

5) Analizar el producto de las reacciones con mucílago de cacao y sacarosa mediante HPLC.

6) Hacer un análisis completo de todas las reacciones evaluadas en el trabajo mediante MALDI-TOF.

Referencias Bibliográficas

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindah, E. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001

Albenne, C., Montalk, G., Monsan, P., Skov, L., Mirza, O., Gajhede, M., & Remaud-Simeon, M. (2002). Site-directed mutagenesis of key amino acids in the active site of amylosucrase from *Neisseria polysaccharea*. *Biologia - Section Cellular and Molecular Biology*, *57*(SUPPL. 11), 119–128.

Albenne, C., Skov, L. K., Mirza, O., Gajhede, M., Potocki-Véronèse, G., Monsan, P., & Remaud-Simeon, M. (2002). Maltooligosaccharide disproportionation reaction: An intrinsic property of amylosucrase from *Neisseria polysaccharea*. *FEBS Letters*, *527*(1–3), 67–70. https://doi.org/10.1016/S0014-5793(02)03168-X

Anvoh, K. Y. B., Zoro, B., & Gnakri, D. (2009). Production and characterization of juice from mucilage of cocoa beans. *Pakistan Journal of Nutrition*, *8*(2), 129–133.

Argos, P. (1990). An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *Journal of Molecular Biology*, *211*(4), 943–958. https://doi.org/10.1016/0022-2836(90)90085-Z

Ausubel, Frederick M. 1992. Short Protocols in Molecular Biology : A Compendium of Methods from Current Protocols in Molecular Biology. Greene Pub. Associates.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., & Hermans, J. (1981). Interaction Models for Water in Relation to Protein Hydration (pp. 331–342). Springer Netherlands. https://doi.org/10.1007/978-94-015-7658-1_21

Biwer, A., Antranikian, G., & Heinzle, E. (2002). Enzymatic production of cyclodextrins. *Applied Microbiology and Biotechnology*, *59*(6), 609–617. https://doi.org/10.1007/s00253-002-1057-x

Bonilha, P. R. M., Menocci, V., Goulart, A. J., Polizeli, M. D. L. T. D. M., & Monti, R. (2006). Cyclodextrin glycosyltransferase from Bacillus licheniformis: Optimization of production and its properties. *Brazilian Journal of Microbiology*, *37*(3), 317–323. https://doi.org/10.1590/S1517-83822006000300022

Büttcher, V., Welsh, T., Willmitzer, L., Kossmann, J., & Kossmann, J. (1997). Cloning and characterization of the gene for amylosucrase from Neisseria polysaccharea : production of a linear alpha-1 , 4-glucan . Cloning and Characterization of the Gene for Amylosucrase from Neisseria polysaccharea : Production of a, *179*(10), 3324–3330. https://doi.org/10.1128/jb.179.10.3324-3330.1997

Cambon, E., Barbe, S., Pizzut-serin, S., Remaud-simeon, M., & Andr, I. (2014). Essential Role of Amino Acid Position 226 in Oligosaccharide Elongation by Amylosucrase From Neisseria polysaccharea. *Biotechnology and Bioengineering*, *111*(9), 1719–1728. https://doi.org/10.1002/bit.25236

Champion, Elise et al. 2012. "Applying Pairwise Combinations of Amino Acid Mutations for Sorting out Highly Efficient Glucosylation Tools for Chemo-Enzymatic Synthesis of Bacterial Oligosaccharides." *Journal of the American Chemical Society* 134(45): 18677–88.

Chang, H. yin, Irwin, P. M., & Nikolov, Z. L. (1998). Effects of mutations in the starch-binding domain of *Bacillus macerans* cyclodextrin glycosyltransferase. *Journal of Biotechnology*, *65*(2–3), 191–202. https://doi.org/10.1016/S0168-1656(98)00115-1

Chen, Fangjin et al. 2015. "Molecular Dynamic Analysis of Mutant Y195I α-Cyclodextrin Glycosyltransferase with Switched Product Specificity from α-Cyclodextrin to γ-

Cyclodextrin." *Journal of Molecular Modeling* 21(8).

Chen, X., Zaro, J. L., & Shen, W.-C. (2013). Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews*, *65*(10), 1357–1369. https://doi.org/10.1016/j.addr.2012.09.039

Cortes, L., Wedd, A. G., & Xiao, Z. (2015). The functional roles of the three copper sites associated with the methionine-rich insert in the multicopper oxidase CueO from *E. coli*. *Metallomics*, *7*(7), 776–785. https://doi.org/10.1039/c5mt00001g

Crabb, W. D., & Mitchinson, C. (1997). Enzymes involved in the processing of starch to sugars. *Trends in Biotechnology*, *15*(9), 349–352. https://doi.org/10.1016/S0167-7799(97)01082-2

Cramer, F., Saenger, W., & Spatz, H. C. H. (1967). Inclusion Compounds. XIX.1a The Formation of Inclusion Compounds of a-Cyclodextrin in Aqueous Solutions. Thermodynamics and Kinetics. *Journal of the American Chemical Society*, *89*(1), 14–20. https://doi.org/10.1021/ja00977a003

Daudé, D., Topham, C. M., Remaud-Siméon, M., & André, I. (2013). Probing impact of active site residue mutations on stability and activity of *Neisseria polysaccharea* amylosucrase. *Protein Science*, *22*(12), 1754–1765. https://doi.org/10.1002/pro.2375

Davies, G., & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*, *3*(9), 853–859. https://doi.org/10.1016/S0969-2126(01)00220-9

Montalk, G. P., Remaud-Simeon, M., Willemot, R. M., Planchot, V., & Monsan, P. (1999). Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. *Journal of Bacteriology*, *181*(2), 375–381. https://doi.org/10.1128/jb.181.2.375-381.1999

Ding, R., Li, Z., Chen, S., Wu, D., Wu, J., & Chen, J. (2010). Enhanced secretion of recombinant

α-cyclodextrin glucosyltransferase from *E. coli* by medium additives. *Process Biochemistry*, *45*(6), 880–886. https://doi.org/10.1016/j.procbio.2010.02.009

Emond, S., Mondeil, S., Jaziri, K., André, I., Monsan, P., Remaud-Siméon, M., & Potocki-Véronèse, G. (2008). Cloning, purification and characterization of a thermostable amylosucrase from Deinococcus geothermalis. *FEMS Microbiology Letters*, *285*(1), 25–32. https://doi.org/10.1111/j.1574-6968.2008.01204.x

Emond, S., Potocki-Véronèse, G., Mondon, P., Bouayadi, K., Kharrat, H., Monsan, P., & Remaud-Simeon, M. (2007). Optimized and automated protocols for high-throughput screening of amylosucrase libraries. *Journal of Biomolecular Screening*, *12*(5), 715–723. https://doi.org/10.1177/1087057107301978

Ernst, Orna, & Tsaffrir Zor. 2010. "Linearization of the Bradford Protein Assay." *Journal of visualized experiments : JoVE* (38). http://www.ncbi.nlm.nih.gov/pubmed/20386536 (December 14, 2018).

Gawande, B. N., Goel, A., Patkar, A. Y., & Nene, S. N. (1999). Purification and properties of a novel raw starch degrading cyclomaltodextrin glucanotransferase from Bacillus firmus. *Applied Microbiology and Biotechnology*, *51*(4), 504–509. https://doi.org/10.1007/s002530051424

GE Healthcare. (2007). Purifying Challenging Proteins. *General Electric Company*, *28-9095–31*, 1–107. Retrieved from http://www.mcdb.ucla.edu/Research/Jacobsen/LabWebSite/PDFOthers/GESeminar.pdf

George, R. A., & Heringa, J. (2002). An analysis of protein domain linkers: Their classification and role in protein folding. *Protein Engineering*, *15*(11), 871–879. https://doi.org/10.1093/protein/15.11.871

Han, R., Li, J., Shin, H. dong, Chen, R. R., Du, G., Liu, L., & Chen, J. (2014). Recent advances in

discovery, heterologous expression, and molecular engineering of cyclodextrin glycosyltransferase for versatile applications. *Biotechnology Advances*, *32*(2), 415–428. https://doi.org/10.1016/j.biotechadv.2013.12.004

Hedges, A. R. (1998). Industrial applications of cyclodextrins. *Chemical Reviews*, *98*(5), 2035–2044. https://doi.org/10.1021/cr970014w

Horikawa, K., Yamada, Y., Matsuda, T., Kobayashi, K., Hashimoto, M., Matsu-Ura, T., … Nagai, T. (2010). Spontaneous network activity visualized by ultrasensitive Ca 2+ indicators, yellow Cameleon-Nano. *Nature Methods*, *7*(9), 729–732. https://doi.org/10.1038/nmeth.1488

Huang, Xinyi, Hazel M Holden, & Frank M Raushel. 2001. "C Hanneling of S Ubstrates And." *New York*.

Ismail, N. F., Hamdan, S., Mahadi, N. M., Murad, A. M. A., Rabu, A., Bakar, F. D. A., … Illias, R. M. (2011). A mutant l-asparaginase II signal peptide improves the secretion of recombinant cyclodextrin glucanotransferase and the viability of Escherichia coli. *Biotechnology Letters*, *33*(5), 999–1005. https://doi.org/10.1007/s10529-011-0517-8

Iturrate, Laura et al. 2010. "Preparation and Characterization of a Bifunctional Aldolase/Kinase Enzyme: A More Efficient Biocatalyst for C-C Bond Formation." *Chemistry - A European Journal* 16(13): 4018–30.

Jeang, C L. (1999). Characterization of the *Bacillus macerans* Cyclodextrin Glucanotransferase Overexpressed in Escherichia coli, *23*(2), 62–68.

Jeang, Chii Ling, Lin, D. G., & Hsieh, S. H. (2005). Characterization of cyclodextrin glycosyltransferase of the same gene expressed from *Bacillus macerans*, Bacillus subtilis, and Escherichia coli. *Journal of Agricultural and Food Chemistry*, *53*(16), 6301–6304. https://doi.org/10.1021/jf0503356

Jemli, S., Ben Messaoud, E., Ben Mabrouk, S., & Bejar, S. (2008). The Cyclodextrin Glycosyltransferase of *Paenibacillus pabuli* US132 Strain: Molecular Characterization and Overproduction of the Recombinant Enzyme. *Journal of Biomedicine and Biotechnology*, *2008*, 1–9. https://doi.org/10.1155/2008/692573

Jiang, Y., Liu, B., Xu, J., Pan, K., Hou, H., Hu, J., & Yang, J. (2018). *Cross-linked chitosan/β-cyclodextrin composite for selective removal of methyl orange: Adsorption performance and mechanism*. *Carbohydrate Polymers* (Vol. 182). Elsevier Ltd. https://doi.org/10.1016/j.carbpol.2017.10.097

Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society*, *118*(45), 11225–11236. https://doi.org/10.1021/ja9621760

Joucla, G., Veen, B. A. Van Der, Potocki-ve, G., Monsan, P., & Ã, M. R. (2004). Combinatorial engineering to enhance amylosucrase performance : construction , selection , and screening of variant libraries for increased activity. *Federation of European Biochemical Societies*, *560*, 91–97. https://doi.org/10.1016/S0014-5793(04)00077-8

Jung, Y. S., Kim, Y. J., Kim, A. T., Jang, D., Kim, M. S., Seo, D. H., … Kim, D. O. (2020). Enrichment of polyglucosylated isoflavones from soybean isoflavone aglycones using optimized amylosucrase transglycosylation. *Molecules*, *25*(1). https://doi.org/10.3390/molecules25010181

Kaulpiboon, J., Prasong, W., Rimphanitchayakit, V., Murakami, S., Aoki, K., & Pongsawasdi, P. (2010). Expression and characterization of a fusion protein-containing cyclodextrin glycosyltransferase from Paenibacillus sp. A11. *Journal of Basic Microbiology*, *50*(5), 427–435. https://doi.org/10.1002/jobm.200900039

Kelly, R. M., Dijkhuizen, L., & Leemhuis, H. (2009). The evolution of cyclodextrin

glucanotransferase product specificity. *Applied Microbiology and Biotechnology*, *84*(1), 119–133. https://doi.org/10.1007/s00253-009-1988-6

Kim, C. I., Kim, M. D., Park, Y. C., Han, N. S., & Seo, J. H. (2000). Refolding of *Bacillus macerans* cyclodextrin glucanotransferase expressed as inclusion bodies in recombinant Escherichia coli. *Journal of Microbiology and Biotechnology*.

Kim, J. H., Wang, R., Lee, W. H., Park, C. S., Lee, S., & Yoo, S. H. (2011). One-pot synthesis of cycloamyloses from sucrose by dual enzyme treatment: Combined reaction of amylosucrase and 4-α-glucanotransferase. *Journal of Agricultural and Food Chemistry*, *59*(9), 5044–5051. https://doi.org/10.1021/jf2002238

Klein, C., & Schulz, G. E. (1991). Structure of cyclodextrin glycosyltransferase refined at 2.0 Å resolution. *Journal of Molecular Biology*, *217*(4), 737–750. https://doi.org/10.1016/0022-2836(91)90530-J

Koh, D. W., Park, M. O., Choi, S. W., Lee, B. H., & Yoo, S. H. (2016). Efficient biocatalytic production of cyclodextrins by combined action of amylosucrase and cyclodextrin glucanotransferase. *Journal of Agricultural and Food Chemistry*, *64*(21), 4371–4375. https://doi.org/10.1021/acs.jafc.6b01080

Kriaa, M., Zouari Ayadi, D., Jemli, S., Sahnoun, M., Bejar, S., & Kammoun, R. (2012). Improvement of cyclodextrin glycosyltransferase (CGTase) production by recombinant Escherichia coli pAD26 immobilized on the cotton. *Biologia (Poland)*, *67*(6), 1049–1055. https://doi.org/10.2478/s11756-012-0124-8

Kweon, Dae Hyuk, Sung Gun Kim, & Jin Ho Seo. 2004. "Purification and Refolding of Cyclodextrin Glycosyltransferase Expressed from Recombinant Escherichia Coli." *Journal of Inclusion Phenomena* 50(1–2): 37–41.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a

program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(2), 283–291. https://doi.org/10.1107/s0021889892009944

Lejeune, A., Sakaguchi, K., & Imanaka, T. (1989). A spectrophotometric assay for the cyclization activity of cyclomaltohexaose (α-cyclodextrin) glucanotransferase. *Analytical Biochemistry*, *181*(1), 6–11. https://doi.org/10.1016/0003-2697(89)90385-0

Li, Z. F., Li, B., Liu, Z. G., Wang, M., Gu, Z. B., Du, G. C.,Chen, J. (2009). Calcium leads to further increase in glycine-enhanced extracellular secretion of recombinant α-cyclodextrin glycosyltransferase in Escherichia coli. *Journal of Agricultural and Food Chemistry*, *57*(14), 6231–6237. https://doi.org/10.1021/jf901239k

Li, Z., Li, B., Gu, Z., Du, G., Wu, J., & Chen, J. (2010). Extracellular expression and biochemical characterization of α-cyclodextrin glycosyltransferase from Paeni*Bacillus macerans*. *Carbohydrate Research*, *345*(7), 886–892. https://doi.org/10.1016/j.carres.2010.02.002

Lim, S. I., Yang, B., Jung, Y., Cha, J., Cho, J., Choi, E. S., I. (2016). Controlled Orientation of Active Sites in a Nanostructured Multienzyme Complex. *Scientific Reports*, *6*(December), 1–9. https://doi.org/10.1038/srep39587

Liu, C., Chin, J. X., & Lee, D. Y. (2015). SynLinker: An integrated system for designing linkers and synthetic fusion proteins. *Bioinformatics*, *31*(22), 3700–3702. https://doi.org/10.1093/bioinformatics/btv447

Lobanov, M. Y., Bogatyreva, N. S., & Galzitskaya, O. V. (2008). Radius of gyration as an indicator of protein structure compactness. *Molecular Biology*, *42*(4), 623–628. https://doi.org/10.1134/S0026893308040195

Makela, M., K. Korpela, T., Puisto, J., & V. Laakso, S. (2002). Nonchromatographic cyclodextrin assays: evaluation of sensitivity, specificity, and conversion mixture applications. *Journal of Agricultural and Food Chemistry*, *36*(1), 83–88. https://doi.org/10.1021/jf00079a021

McKee, D., & Karwic, A. (2006). Product and method for oral administration of nutraceuticals (75). *Naval Architect, US 7,632,5*. https://doi.org/10.1037/t24245-000

Mercado Villamizar, M. F., & Pinzón Joya, J. R. (2020). *Producción de Ciclodextrinas a partir de Almidón de Yuca y Mucílago de Cacao*. Universidad Industrial de Santander.

Mirza, O., Skov, L. K., Remaud-Simeon, M., Montalk, G. P., Albenne, C., Monsan, P., & Gajhede, M. (2001). Crystal structures of amylosucrase from *Neisseria polysaccharea* in complex with D-glucose and the active site mutant Glu328Gln in complex with the natural substrate sucrose. *Biochemistry*, *40*(30), 9032–9039. https://doi.org/10.1021/bi010706l

Mizrahi, A., & Wezel, A. L. van. (1983). *Production, characterization, and application of cyclodextrins*. *Adv. Biotechnol. ProcessesAdvances in Biotechnological Processes* (Vol. 6). Alan R. Liss. Retrieved from https://www.tib.eu/en/search/id/ceaba%3ACEAB1970094865/Production-characterization-and-application-of/

Montalk, G. P. D. E., Willemot, R. M., Planchot, V., & Monsan, P. (1999). Sequence Analysis of the Gene Encoding Amylosucrase from *Neisseria polysaccharea* and Characterization of the Recombinant Enzyme, *181*(2), 375–381.

Mourtzinos, I., & Goula, A. (2019). *Polyphenols in Agricultural Byproducts and Food Waste*. *Polyphenols in Plants* (2nd ed.). Elsevier Inc. https://doi.org/10.1016/b978-0-12-813768-0.00002-5

Oddoye, E. O. K., Agyente-Badu, C. K., & Gyedu-Akoto, E. (2013). Cocoa and Its By-Products: Identification and Utilization. In *Chocolate in Health and Nutrition* (pp. 23–37). Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-61779-803-0_3

Ortiz-Rodríguez, O. O., Villamizar-Gallardo, R. A., Naranjo-Merino, C. A., García-Caceres, R. G., & Castañeda-galvís, M. T. (2016). Carbon footprint of the colombian cocoa production. *Engenharia Agrícola*, *36*(2), 260–270. https://doi.org/10.1590/1809-4430-Eng.Agric.v36n2p260-270/2016

Ortiz, K., & Álvarez, R. (2015).Efecto del vertimiento de subproductos del beneficio de cacao *(Theobroma cacao L.)* sobre algunas ropiedades químicas y biológicas en los suelos de una finca cacaotera,Municipio de Yaguará (Huila, Colombia). *Boletín Científico. Centro de Museos. Museo de Historia Natural*, *19*(1), 65–84. https://doi.org/10.17151/bccm.2015.19.1.5

Park, M. O., Lee, B. H., Lim, E., Lim, J. Y., Kim, Y., Park, C. S., … Yoo, S. H. (2016). Enzymatic Process for High-Yield Turanose Production and Its Potential Property as an Adipogenesis Regulator. *Journal of Agricultural and Food Chemistry*, *64*(23), 4758–4764. https://doi.org/10.1021/acs.jafc.5b05849

Pendergast, D. D., & Connors, K. A. (1984). Improved competitive indicator methods for the study of α-cyclodextrin complexes. *Journal of Pharmaceutical Sciences*, *73*(12), 1779–1783. https://doi.org/10.1002/jps.2600731230

Penninga, D., van der Veen, B. A., Knegtel, R. M. A., van Hijum, S. A. F. T., Rozeboom, H. J., Kalk, K. H., … Dijkhuizen, L. (1996). The raw starch binding somain of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251. *Journal of Biological Chemistry*, *271*(51), 32777–32784. https://doi.org/10.1074/jbc.271.51.32777

Pettersson, Henrik. 2001. “Kinetics of the Coupled Reaction Catalysed by a Fusion Protein of β-Galactosidase.” 1549: 6–11.

Pikkemaat, M. G., Linssen, A. B. M., Berendsen, H. J. C., & Janssen, D. B. (2002). Molecular dynamics simulations as a tool for improving protein stability. *Protein Engineering, Design and Selection*, *15*(3), 185–192. https://doi.org/10.1093/protein/15.3.185

Pizzut-Serin, S., Potocki-Véronèse, G., Van Der Veen, B. A., Albenne, C., Monsan, P., & Remaud-Simeon, M. (2005). Characterisation of a novel amylosucrase from Deinococcus radiodurans. *FEBS Letters*, *579*(6), 1405–1410. https://doi.org/10.1016/j.febslet.2004.12.097

Pravda, Lukáš et al. 2014. *Anatomy of Enzyme Channels*. http://www.biomedcentral.com/1471-2105/15/379 (March 25, 2021).

Montalk, G., Remaud-Simeon, M., Willemot, R. M., Sarçabal, P., Planchot, V., & Monsan, P. (2000). Amylosucrase from *Neisseria polysaccharea*: Novel catalytic properties. *FEBS Letters*, *471*(2–3), 219–223. https://doi.org/10.1016/S0014-5793(00)01406-X

Qi, Q., & Zimmermann, W. (2005). Cyclodextrin glucanotransferase: From gene to applications. *Applied Microbiology and Biotechnology*, *66*(5), 475–485. https://doi.org/10.1007/s00253-004-1781-5

Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*. J Mol Biol. https://doi.org/10.1016/S0022-2836(63)80023-6

Raushel, Frank M., James B. Thoden, & Hazel M. Holden. 2003. "Enzymes with Molecular Tunnels." *Accounts of Chemical Research* 36(7): 539–48.

Rendleman, J. A. (1999). The Production of Cyclodextrins Using CGTase from *Bacillus macerans*, *10*, 89–101. https://doi.org/10.1007/978-1-59259-261-6_8

Rha, C. S., Kim, H. G., Baek, N. I., Kim, D. O., & Park, C. S. (2020). Using Amylosucrase for the Controlled Synthesis of Novel Isoquercitrin Glycosides with Different Glycosidic Linkages. *Journal of Agricultural and Food Chemistry*, *68*(47), 13798–13805. https://doi.org/10.1021/acs.jafc.0c05625

Rizk, M., Antranikian, G., & Elleuche, S. (2012). End-to-end gene fusions and their impact on the production of multifunctional biomass degrading enzymes. *Biochemical and Biophysical Research Communications*, *428*(1), 1–5. https://doi.org/10.1016/j.bbrc.2012.09.142

Ryu, J., Lee, B., Seo, D., Baik, M., Park, C., Wang, R., & Yoo, S. (2010). Production and characterization of digestion-resistant starch by the reaction of *Neisseria polysaccharea* amylosucrase, 221–228. https://doi.org/10.1002/star.200900182

Sambrook, J., & Russel, D. (2001). Molecular Cloning: A Laboratory Manual. Third Ed.

Schmid, G. (1989). Cyclodextrin glycosyltransferase production: yield enhancement by overexpression of cloned genes. *Trends in Biotechnology*, *7*(9), 244–248. https://doi.org/10.1016/0167-7799(89)90015-2

Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 2.0

Sebaaly, C., Charcosset, C., Fourmentin, S., & Greige-Gerges, H. (2018). Potential Applications of Cyclodextrin Inclusion Complexes, Liposomes, and Drug-in-Cyclodextrin-in-Liposome in Food Industry and Packaging. In *Role of Materials Science in Food Bioengineering* (pp. 187–234). Elsevier. https://doi.org/10.1016/B978-0-12-811448-3.00006-1

Skov, L. K., Mirza, O., Henriksen, A., Montalk, G. P., Remaud-Simeon, M., Sarçabal, P., … Gajhede, M. (2001). Amylosucrase, a Glucan-synthesizing Enzyme from the α-Amylase Family. *Journal of Biological Chemistry*, *276*(27), 25273–25278. https://doi.org/10.1074/jbc.M010998200

Skov, L. K., Pizzut-Serin, S., Remaud-Simeon, M., Ernst, H. A., Gajhede, M., & Mirza, O. (2013). The structure of amylosucrase from *Deinococcus radiodurans* has an unusual open active-site topology. *Acta Crystallographica Section F Structural Biology and Crystallization Communications*, *69*(9), 973–978. https://doi.org/10.1107/S1744309113021714

Sonnendecker, C., Wei, R., Kurze, E., Wang, J., Oeser, T., & Zimmermann, W. (2017). Efficient extracellular recombinant production and purification of a Bacillus cyclodextrin glucanotransferase in Escherichia coli. *Microbial Cell Factories*, *16*(1), 1–6. https://doi.org/10.1186/s12934-017-0701-1

Stewart, D. J., Milanova, R. K., Zawistowski, J., & Wallis, S. H. (1999). Compositions comprising phytosterol and(or) phytostanol having enhanced solubility and dispersability.

Sumner, J. B., & Howell, S. F. (1934). A method for determination of saccharase activity. *Medical, Sumner Biochemistry, Stacey Cornell, October F Howell*, (5), 51–55.

Tian, Y., Xu, W., Zhang, W., Zhang, T., Guang, C., & Mu, W. (2018). Amylosucrase as a transglucosylation tool: From molecular features to bioengineering applications. *Biotechnology Advances*, *36*(5), 1540–1552. https://doi.org/10.1016/j.biotechadv.2018.06.010

Tiwari, G., Tiwari, R., & Rai, A. K. (2010). Cyclodextrins in delivery systems: Applications. *Journal of Pharmacy & Bioallied Sciences*, *2*(2), 72–79. https://doi.org/10.4103/0975-7406.67003

Uitdehaag, J. C. M., Kalk, K. H., Van Der Veen, B. A., Dijkhuizen, L., & Dijkstra, B. W. (1999a). The cyclization mechanism of cyclodextrin glycosyltransferase (CGTase) as revealed by a γ-cyclodextrin-CGTase complex at 1.8-Å resolution. *Journal of Biological Chemistry*, *274*(49), 34868–34876. https://doi.org/10.1074/jbc.274.49.34868

Uitdehaag, J. C., Mosi, R., Kalk, K. H., van der Veen, B. A., Dijkhuizen, L., Withers, S. G., & Dijkstra, B. W. (1999b). X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the {\textalpha{}}-amylase family. *Nat Struct Biol*, *6*(5), 432–436.

Urban, M., Beran, M., Adámek, L., Drahoŕd, J., Molík, P., & Matušová, K. (2012). Cyclodextrin

production from amaranth starch by cyclodextrin glycosyltransferase produced by paeni*Bacillus macerans* CCM 2012. *Czech Journal of Food Sciences*, *30*(1), 15–20. https://doi.org/papers3://publication/uuid/3EFBB50C-27CC-44EF-993C-AD64E28AAFD5

van der Veen, B. A., Potocki-Véronèse, G., Albenne, C., Joucla, G., Monsan, P., & Remaud-Simeon, M. (2004). Combinatorial engineering to enhance amylosucrase performance: construction, selection, and screening of variant libraries for increased activity. *FEBS Letters*, *560*(1–3), 91–97. https://doi.org/10.1016/S0014-5793(04)00077-8

Van Der Veen, B. A., Skov, L. K., Potocki-Véronèse, G., Gajhede, M., Monsan, P., & Remaud-Simeon, M. (2006). Increased amylosucrase activity and specificity, and identification of regions important for activity, specificity and stability through molecular evolution. *FEBS Journal*, *273*(4), 673–681. https://doi.org/10.1111/j.1742-4658.2005.05076.x

Van Der Veen, B. A., Van Alebeek, G. J. W. M., Uitdehaag, J. C. M., Dijkstra, B. W., & Dijkhuizen, L. (2000). The three transglycosylation reactions catalyzed by cyclodextrin glycosyltransferase from Bacillus circulans (strain 251) proceed via different kinetic mechanisms. *European Journal of Biochemistry*, *267*(3), 658–665. https://doi.org/10.1046/j.1432-1327.2000.01031.x

Van Rosmalen, M., Krom, M., & Merkx, M. (2017). Tuning the Flexibility of Glycine-Serine Linkers to Allow Rational Design of Multidomain Proteins. *Biochemistry*, *56*(50), 6565–6574. https://doi.org/10.1021/acs.biochem.7b00902

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*(SUPPL.2), 407–410. https://doi.org/10.1093/nar/gkm290

Yang, Yi Nan, Wen Xin Shan, & Pi Wu Wang. 2017. “Upscale Production of a Recombinant Cyclodextrin Glycosyltransferase from Paenibacillus Macerans in Escherichia Coli.” *3 Biotech* 7(3): 1–10.

Yu, K., Liu, C., Kim, B. G., & Lee, D. Y. (2015). Synthetic fusion protein design and applications. *Biotechnology Advances*, *33*(1), 155–164. https://doi.org/10.1016/j.biotechadv.2014.11.005

Zhang, H., Zhou, X., He, J., Wang, T., Luo, X., Wang, L., & Wang, R. (2017a). Impact of amylosucrase modification on the structural and physicochemical properties of native and acid-thinned waxy corn starch. *Food Chemistry*, *220*, 413–419. https://doi.org/10.1016/j.foodchem.2016.10.030

Zhang, K., Su, L., Duan, X., Liu, L., & Wu, J. (2017b). High-level extracellular protein production in Bacillus subtilis using an optimized dual-promoter expression system. *Microbial Cell Factories*, *16*(1), 1–15. https://doi.org/10.1186/s12934-017-0649-1

Zheng, J., Li, X., & Wu, H. (2019). High-level extracellular secretion and characterization of the thermophilic β-cyclodextrin glucanotranferase from Paenibacillus campinasensis in Escherichia coli. *3 Biotech*, *9*(10), 1–11. https://doi.org/10.1007/s13205-019-1909-z

Apéndices

Apéndice 1. Validación de los modelos SynLinker y las estructuras de las proteínas nativas AS (pdb 1g5a) y CGTasa (1cgt), mediante las herramientas PROSA WEB y Rampage Ramachandran.

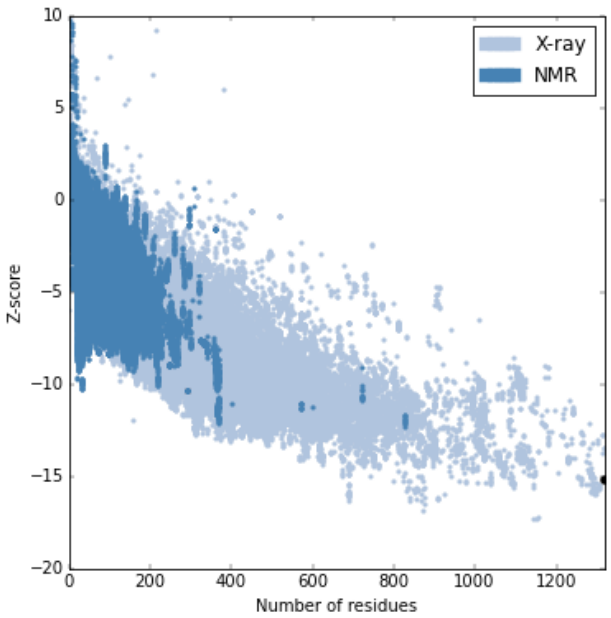
1.1 PROSA WEB

El Z-score indica la calidad general del modelo, valores de Z-score fuera del rango característico de las proteínas nativas indican estructuras erróneas. Con el fin de facilitar la interpretación de la puntuación de Z-score del modelo especifico, su valor particular se muestra en un gráfico que contiene valores Z-scorte de todas las cadenas de proteínas determinadas experimentalmente.

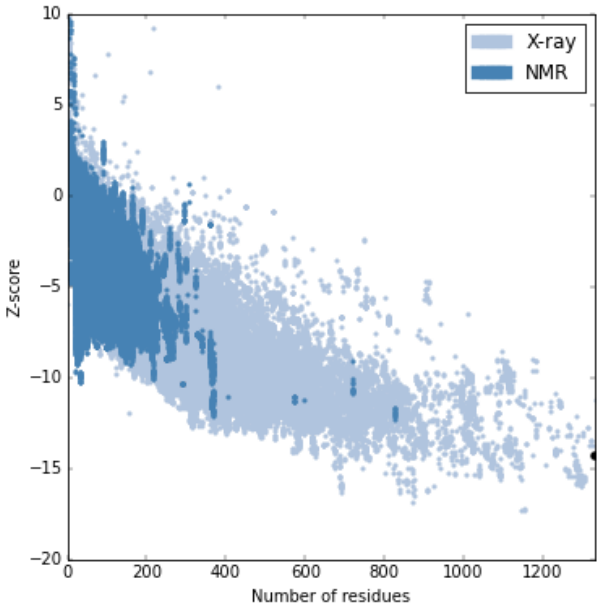
Modelos $AS_{N\text{-}terminal}$ -$CGT_{C\text{-}terminal}$

Como se observa en los gráficos de energía de Z-score, existe un número pequeño de proteínas de tamaño similar a los modelos de proteína fusión y con estructura 3D experimentalmente resulta. No obstante, los valores de Z-score para los modelos 3D de la variante $AMS_{N\text{-}terminal}$ -$CGT_{C\text{-}terminal}$ se encuentran dentro del rango de puntuación típico para las proteínas determinadas experimentalmente y disponibles en la pdb.

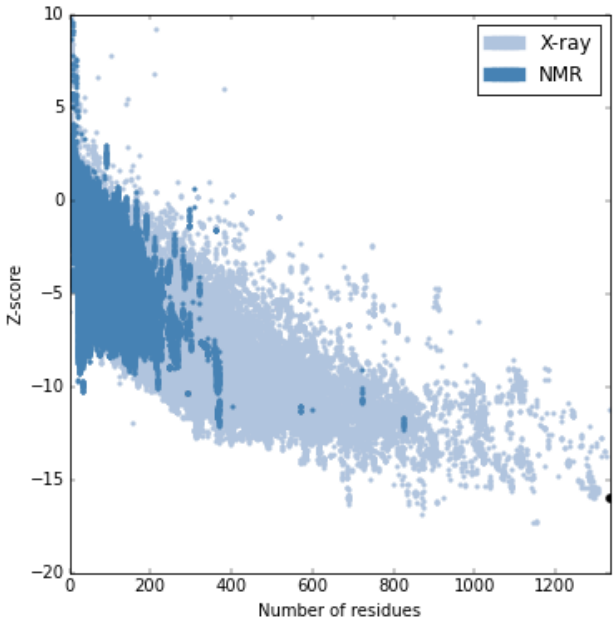
AS-$_{(GGGGS)1}$-CGTasa
Z-Score = -15.7



AS-$_{(GGGGS)3}$-CGTasa
Z-Score = -14.26



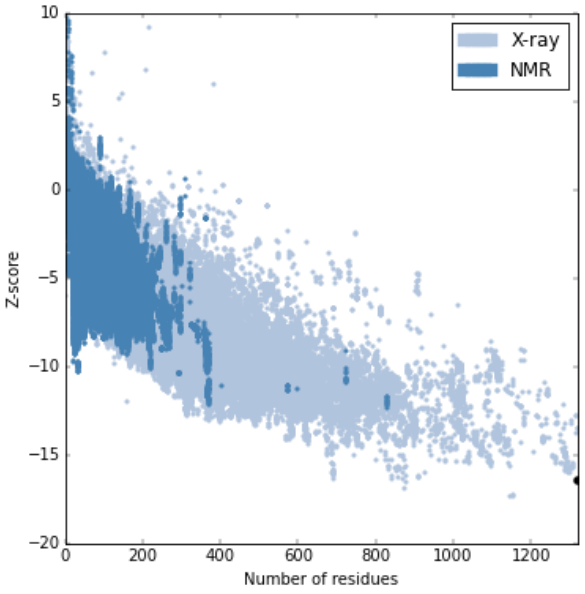
AS-$_{(GGGGS)4}$-CGTasa
Z-Score = -15.9



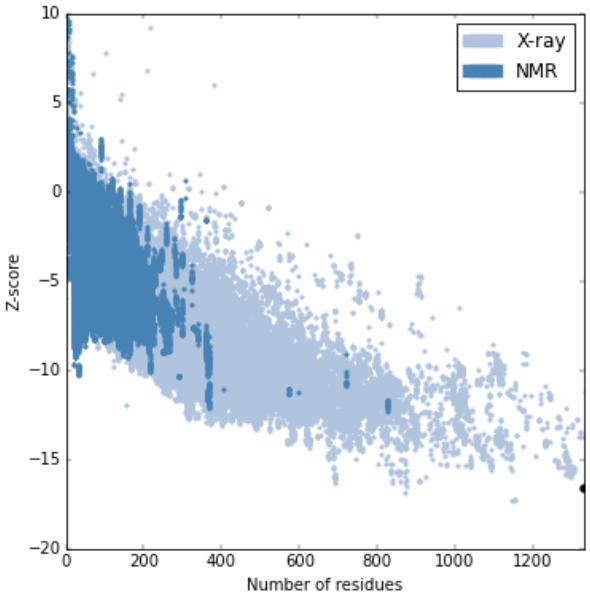
Modelos $CGT_{N\text{-terminal}}$-$AS_{C\text{-terminal}}$

El Z-score de los modelos $CGT_{N\text{-terminal}}$-$AMS_{C\text{-terminal}}$ se encuentra dentro de los valores teóricamente aceptables, pero presenta una ligera desviación con respecto a proteínas de tamaño similar resueltas experimentalmente. La desviación de los valores posiblemente esté relacionada con el valor de Z-Score observado en la proteína individual AMS (1g5a) que, a pesar de ser una estructura resuelta experimentalmente, presenta un valor de puntuación atípico.

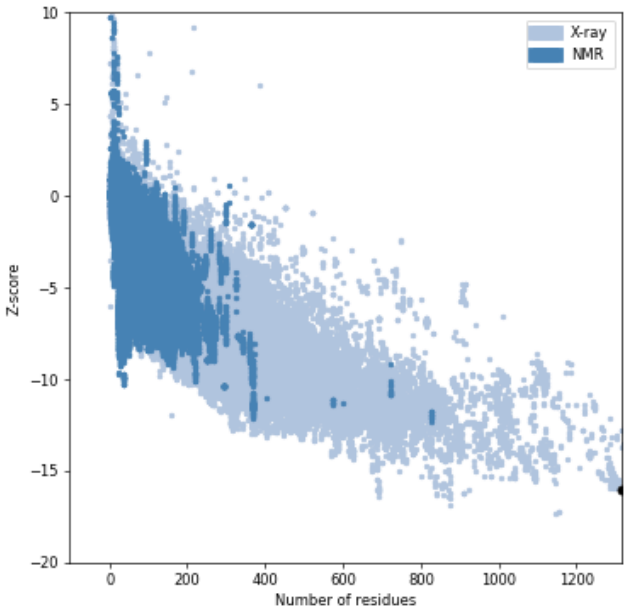
CGTasa-$_{(GGGGS)1}$-AS
Z-Score = -16.4



CGTasa-$_{(GGGGS)3}$-AS
Z-Score = -16.56

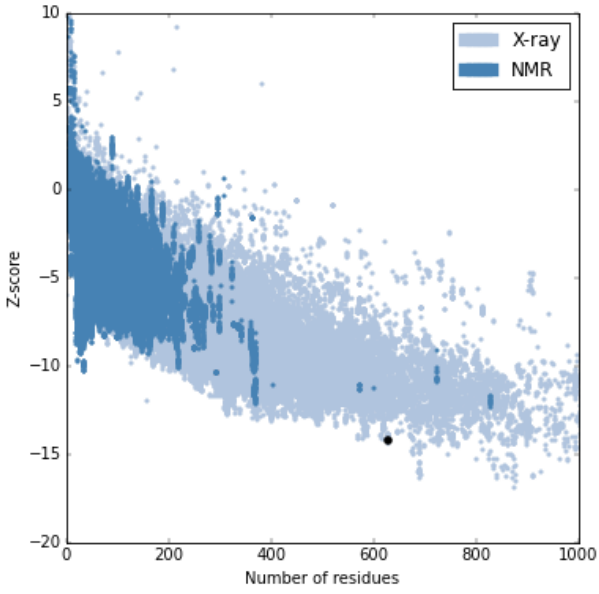


CGTasa-$_{(GGGGS)4}$-AS
Z-Score = -16.01

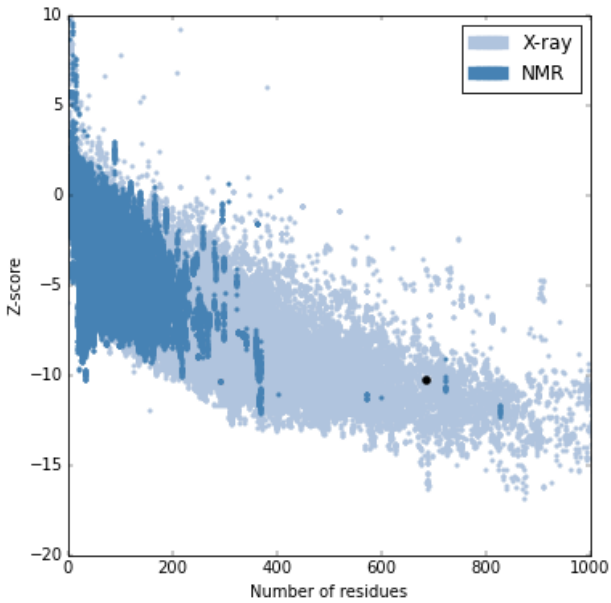


Estructuras pdb AMS y CGTasa

AS (1g5a)
Z-Score = -14.17

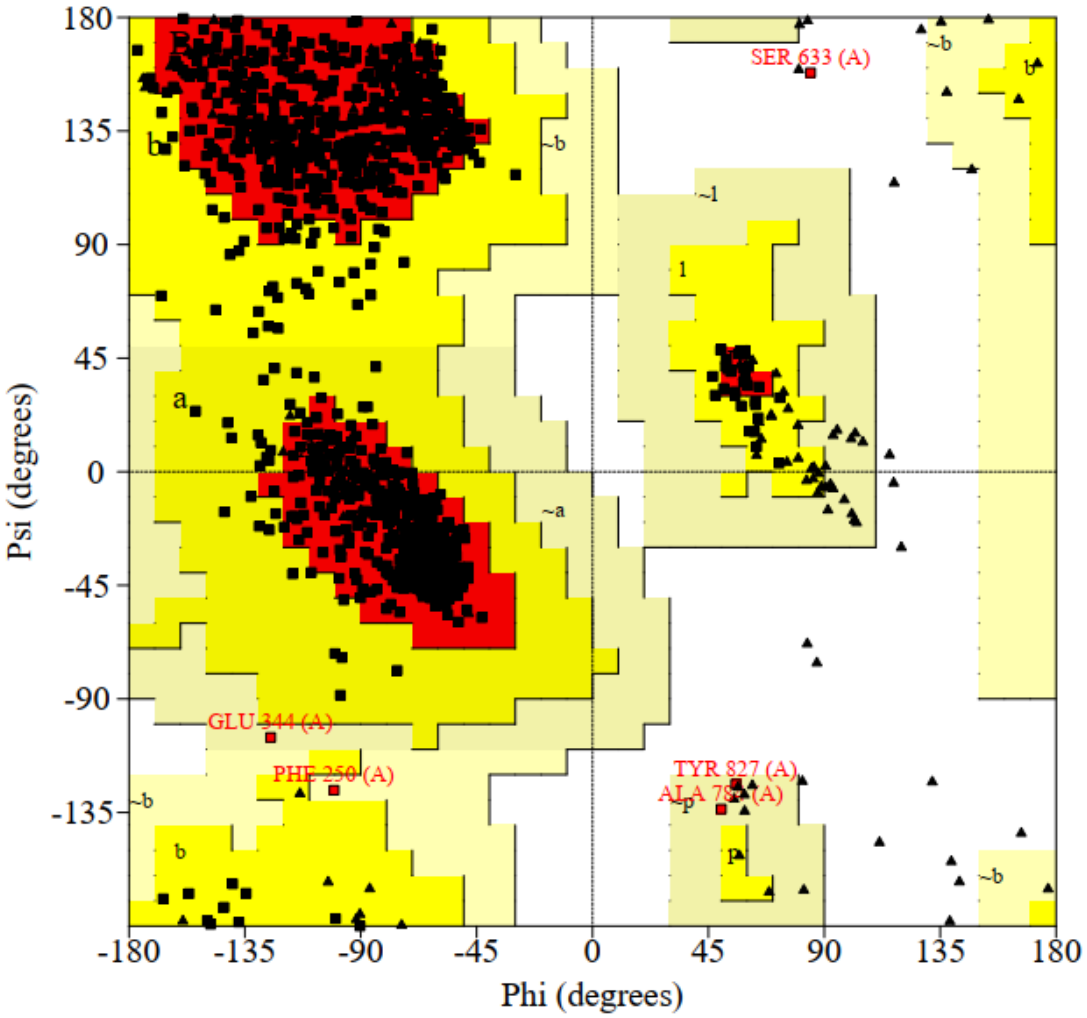


CGTasa (1cgt)
Z-Score = -10.25



1.2 Ramachandran plot

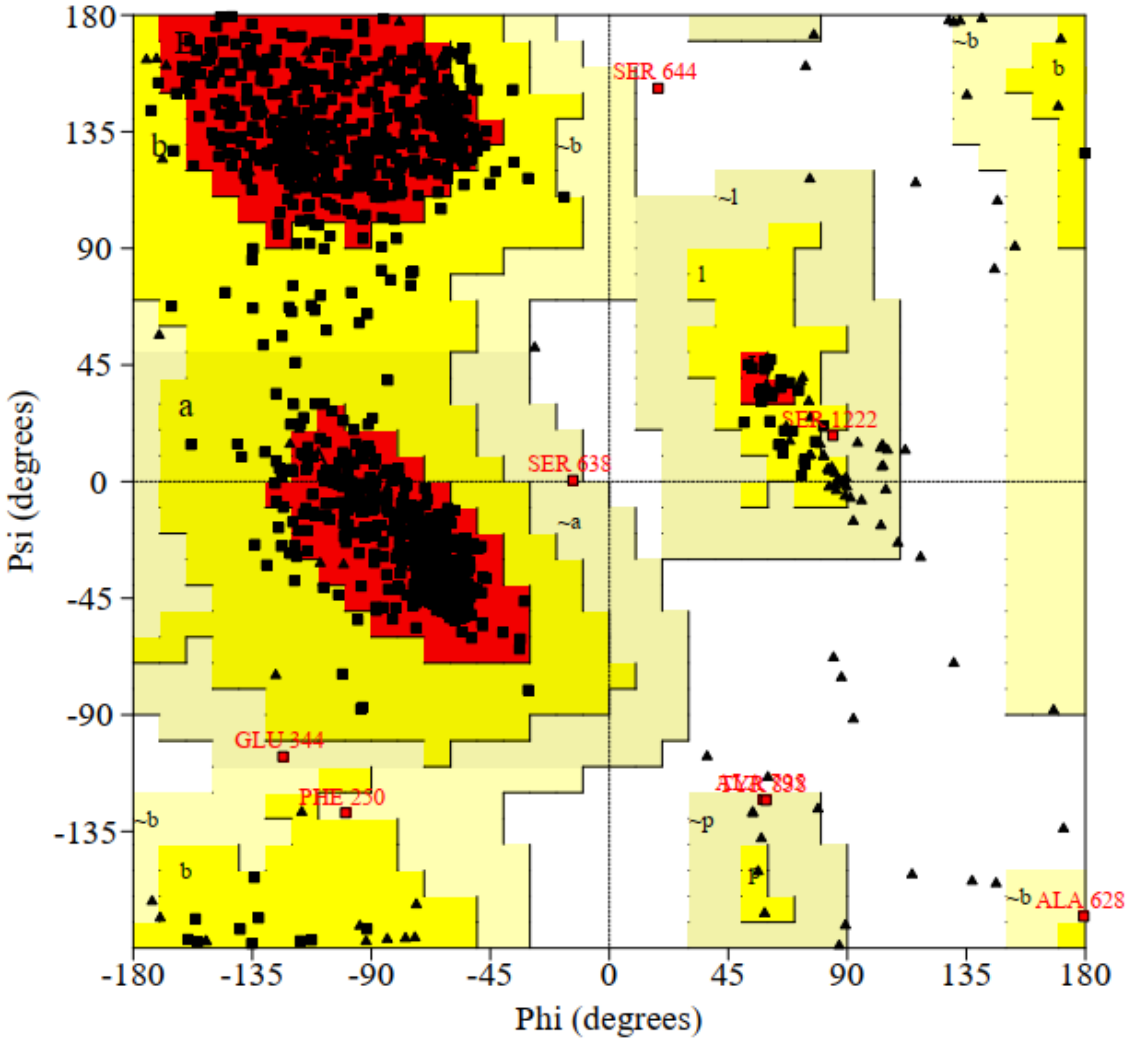
AS$_{L1}$CGT



Plot statistics

Residues in most favoured regions [A,B,L]	1032	90.4%
Residues in additional allowed regions [a,b,l,p]	104	9.1%
Residues in generously allowed regions [~a,~b,~l,~p]	4	0.4%
Residues in disallowed regions	1	0.1%
Number of non-glycine and non-proline residues	1141	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	122	
Number of proline residues	54	
Total number of residues	1319	

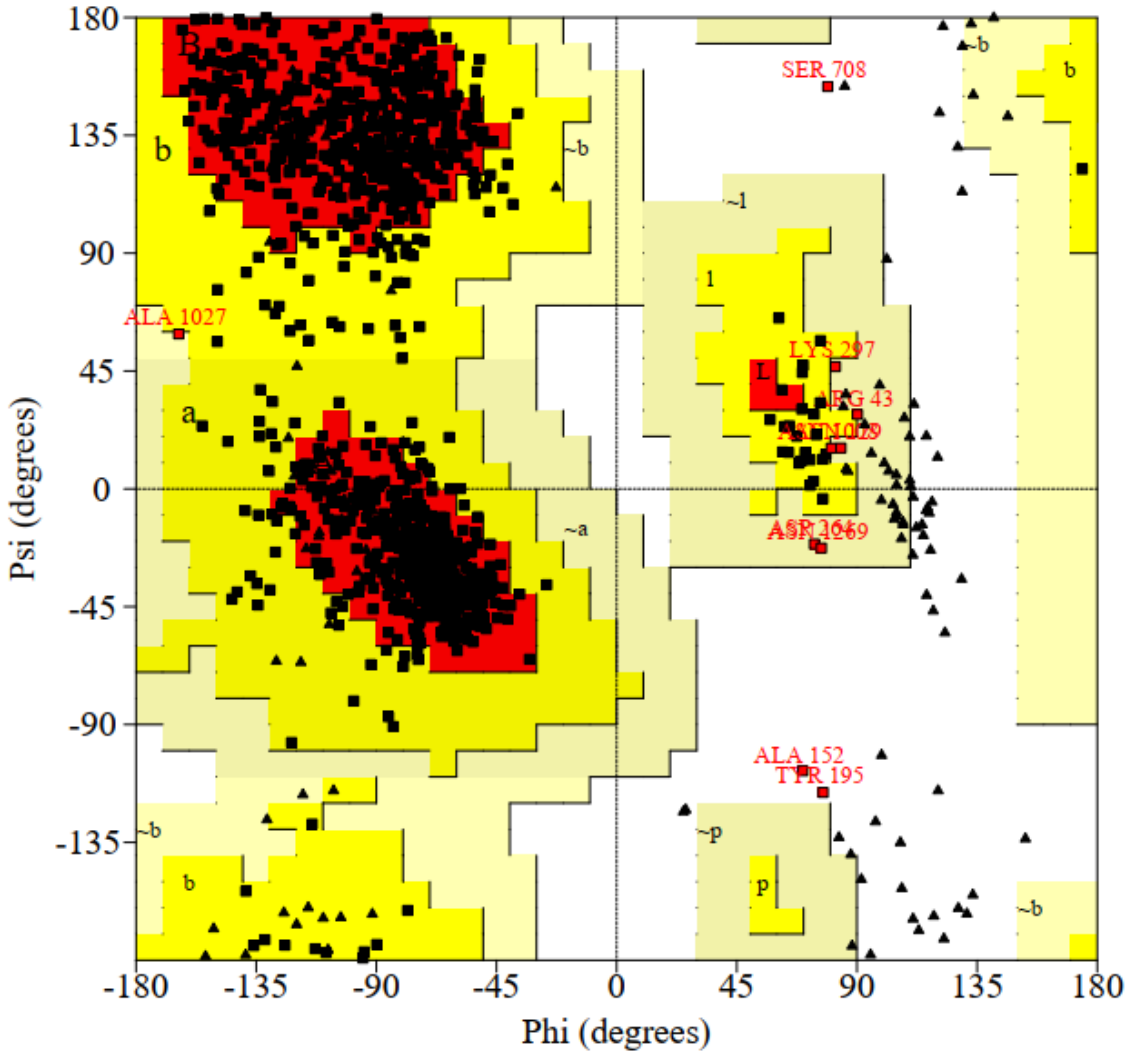
$AS_{L3}CGT$



Plot statistics

Residues in most favoured regions [A,B,L]	1032	90.3%
Residues in additional allowed regions [a,b,l,p]	103	9.0%
Residues in generously allowed regions [~a,~b,~l,~p]	7	0.6%
Residues in disallowed regions	1	0.1%
Number of non-glycine and non-proline residues	1143	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	131	
Number of proline residues	54	
Total number of residues	1330	

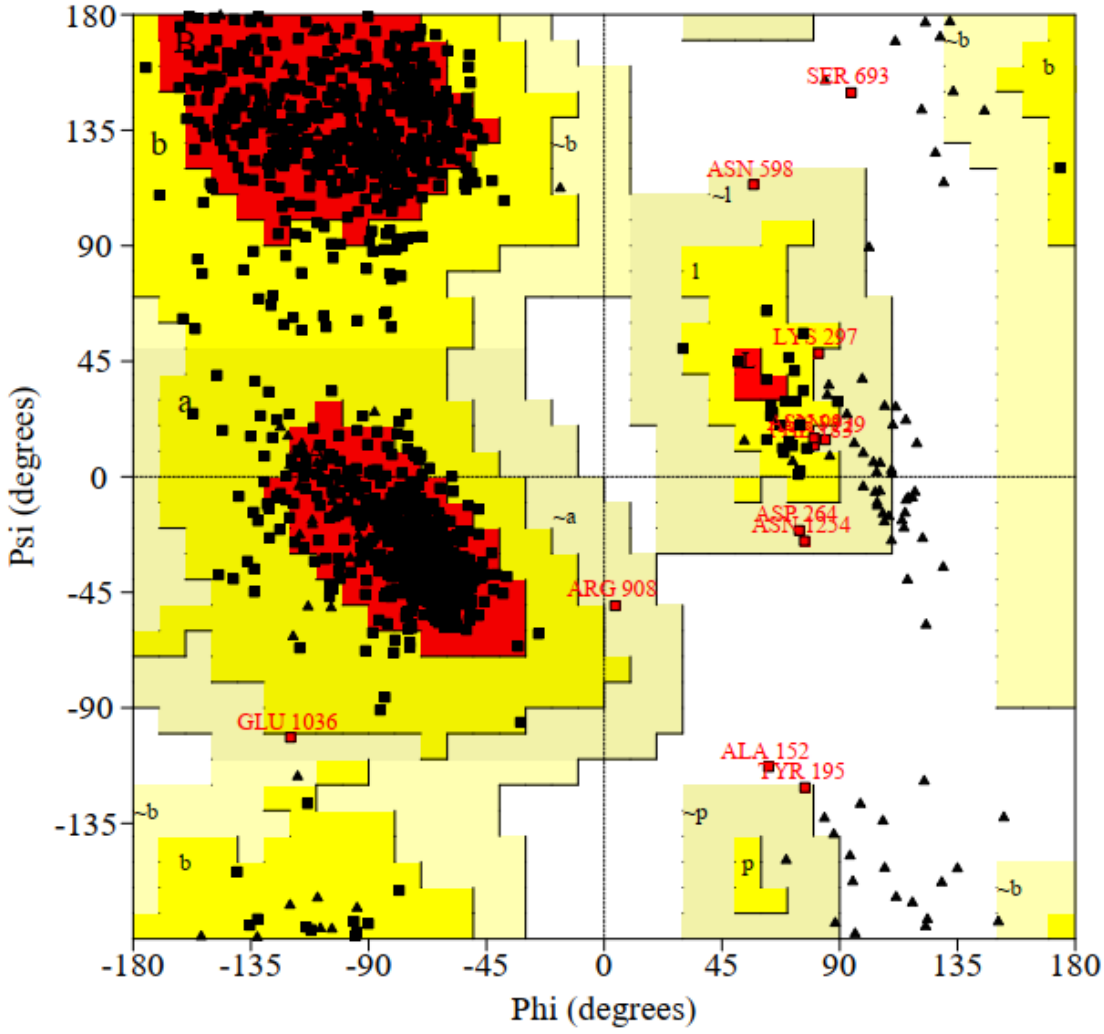
AS$_{L4}$CGT



Plot statistics

Residues in most favoured regions [A,B,L]	1001	87.5%
Residues in additional allowed regions [a,b,l,p]	133	11.6%
Residues in generously allowed regions [~a,~b,~l,~p]	7	0.6%
Residues in disallowed regions	3	0.3%
Number of non-glycine and non-proline residues	1144	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	135	
Number of proline residues	54	
Total number of residues	1335	

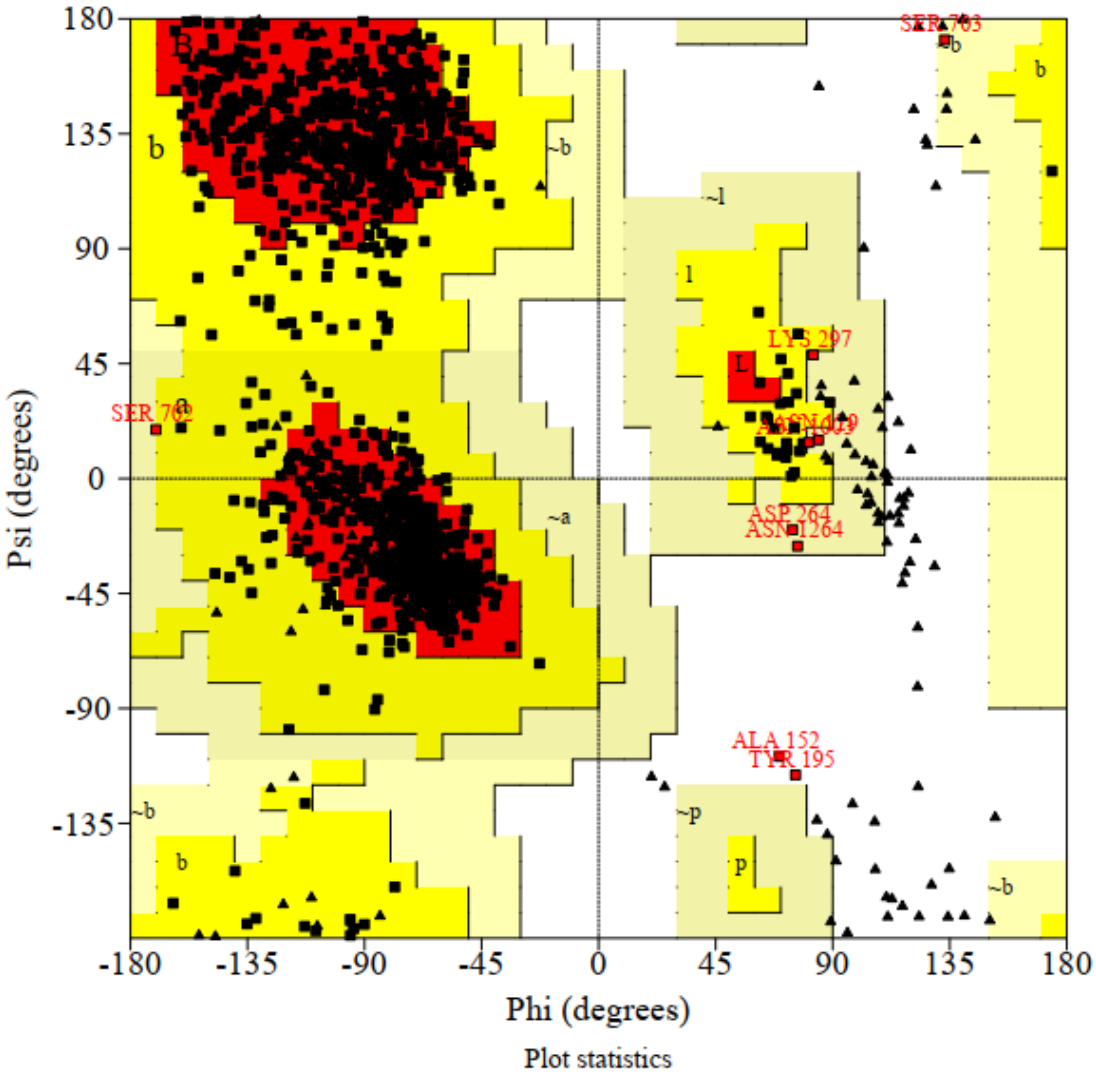
$GT_{L1}AS$



Plot statistics

Residues in most favoured regions [A,B,L]	991	86.9%
Residues in additional allowed regions [a,b,l,p]	138	12.1%
Residues in generously allowed regions [~a,~b,~l,~p]	10	0.9%
Residues in disallowed regions	2	0.2%
Number of non-glycine and non-proline residues	1141	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	123	
Number of proline residues	54	
Total number of residues	1320	

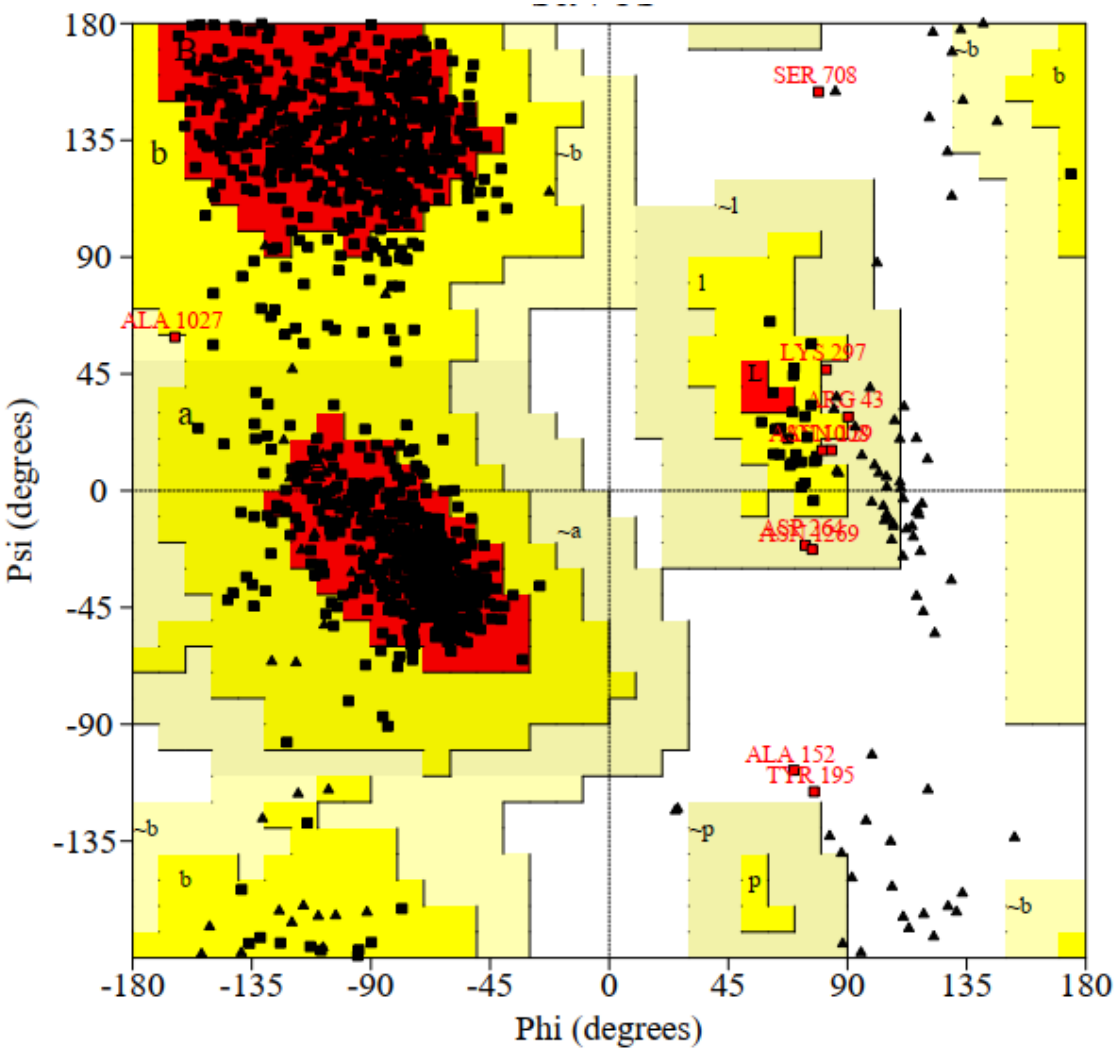
$CGT_{L3}AS$



Plot statistics

Residues in most favoured regions [A,B,L]	991	86.7%
Residues in additional allowed regions [a,b,l,p]	143	12.5%
Residues in generously allowed regions [~a,~b,~l,~p]	7	0.6%
Residues in disallowed regions	2	0.2%
Number of non-glycine and non-proline residues	1143	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	131	
Number of proline residues	54	
Total number of residues	1330	

$CGT_{L4}AS$



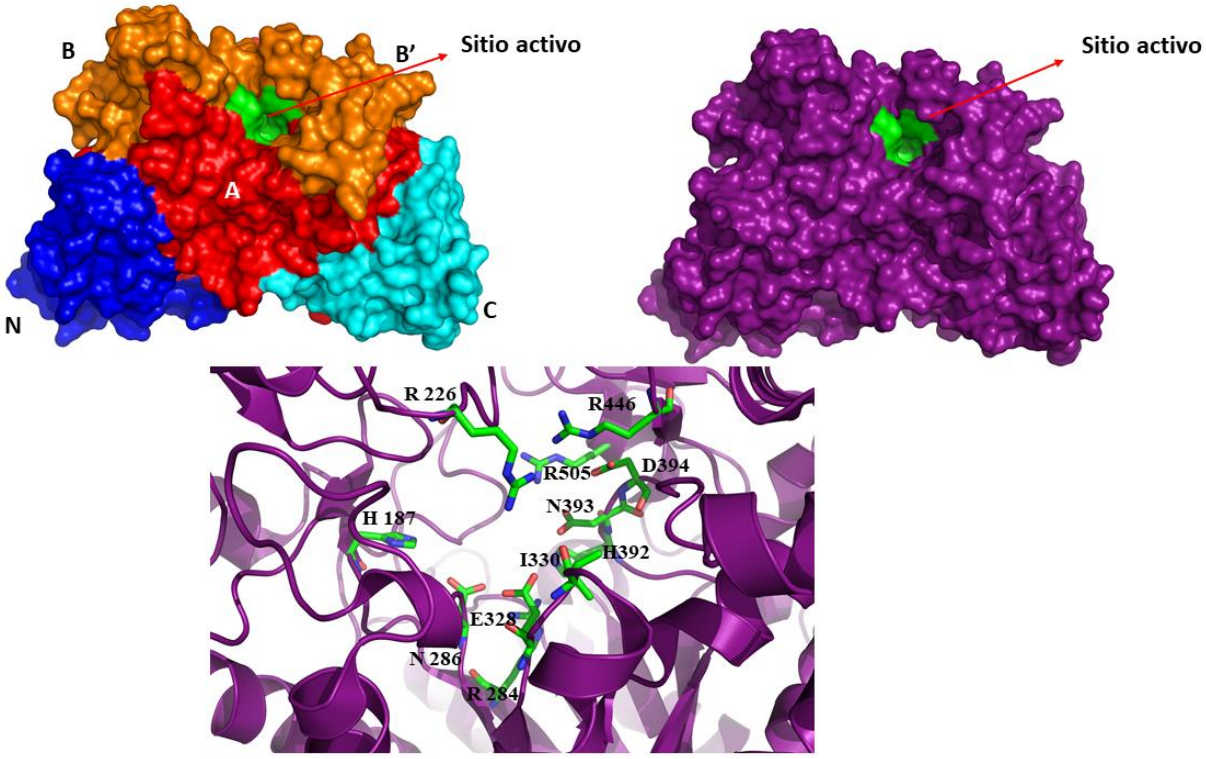
Plot statistics

Residues in most favoured regions [A,B,L]	1001	87.5%
Residues in additional allowed regions [a,b,l,p]	133	11.6%
Residues in generously allowed regions [~a,~b,~l,~p]	7	0.6%
Residues in disallowed regions	3	0.3%
Number of non-glycine and non-proline residues	1144	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	135	
Number of proline residues	54	
Total number of residues	1335	

Apéndice 2. Sitio activo de AS y CGTasa

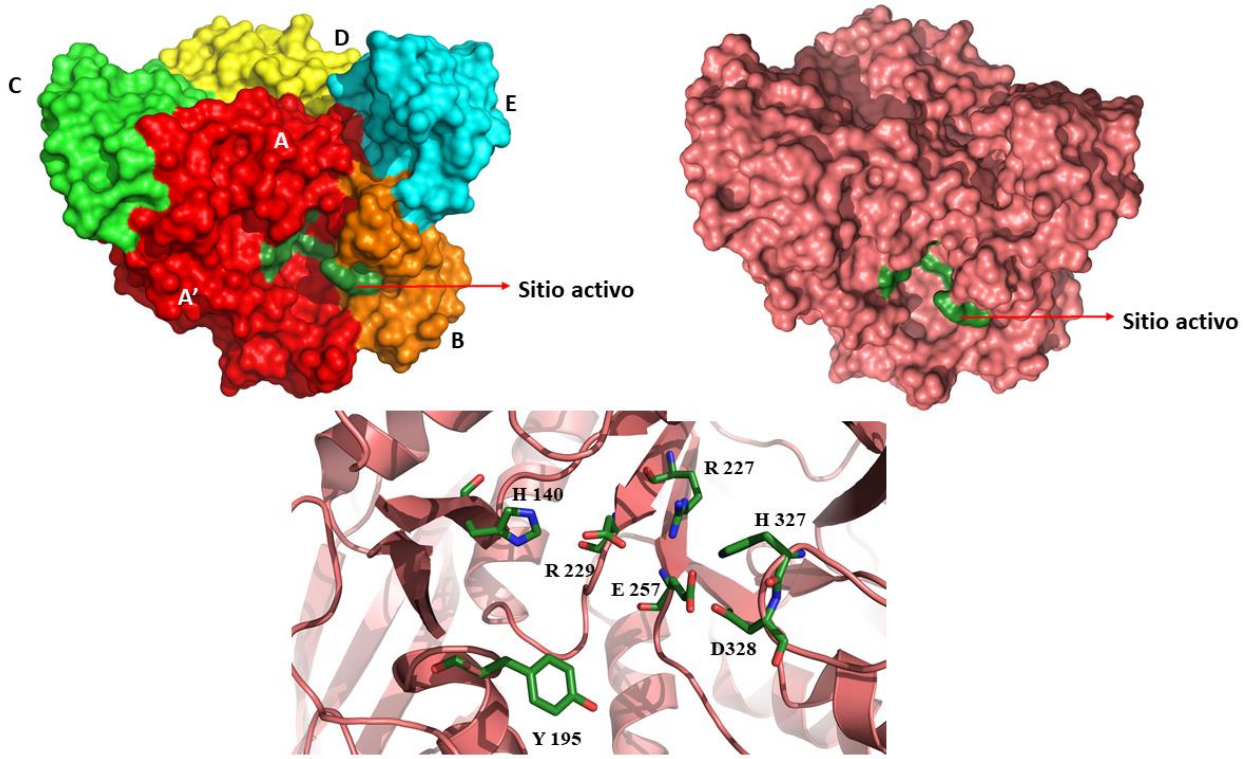
2.1 Amilosacarasa

En la amilosacarasa el bolsillo de unión (Binding pocket) se encuentra en la superficie cerca a los dominios B' y B, conformado por los aminoácidos H392, D144, R226, I330, R509 y H187 (subsitios -1 y +1) (Daudé et al. 2013; Skov et al. 2001a), mientras que los residuos catalíticos son E328, D393 y D285.



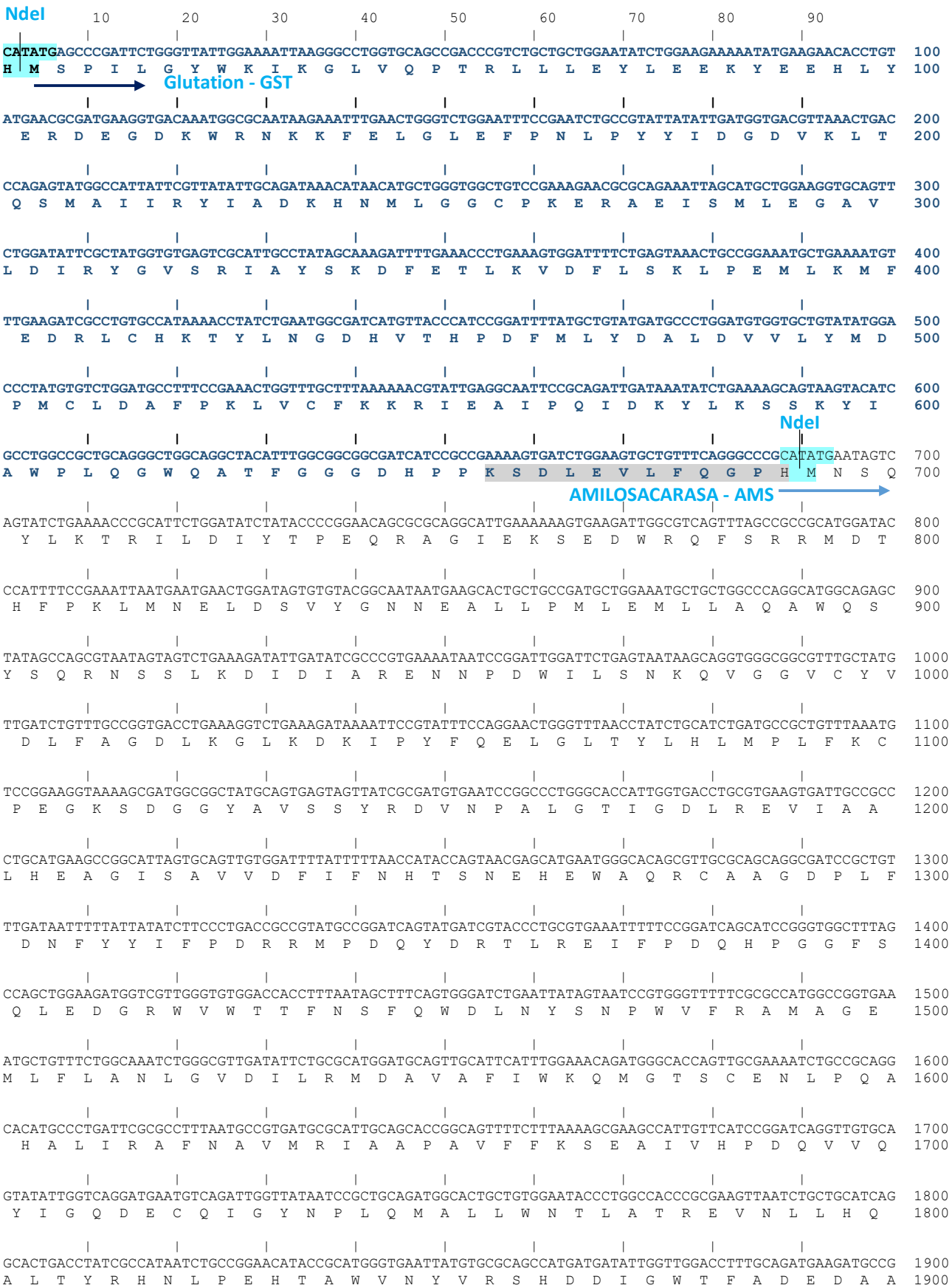
2.2 CGTasa

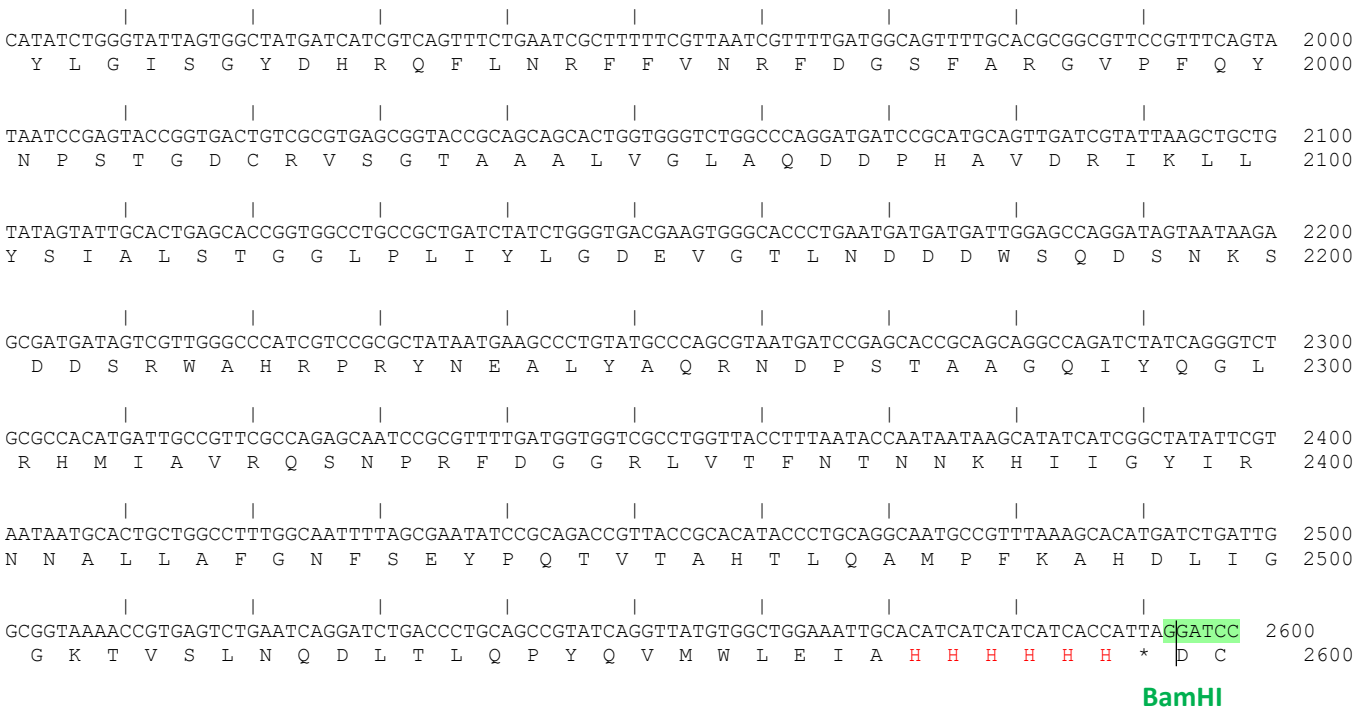
En la CGTasa, el sitio activo está ubicado en el Dominio A, conformado por los aminoácidos E257, D229, D328, H140, H327 y R227 (Uitdehaag et al. 1999a; Uitdehaag et al. 1999b). Otro aminoácido importante es Y195, que cumple un papel importante en la determinación de la especificidad del tipo de CDs que se producen (Kelly, Dijkhuizen, & Leemhuis 2009). Por otro lado, el dominio E funciona como un dominio de unión del almidón (Penninga et al. 1996).



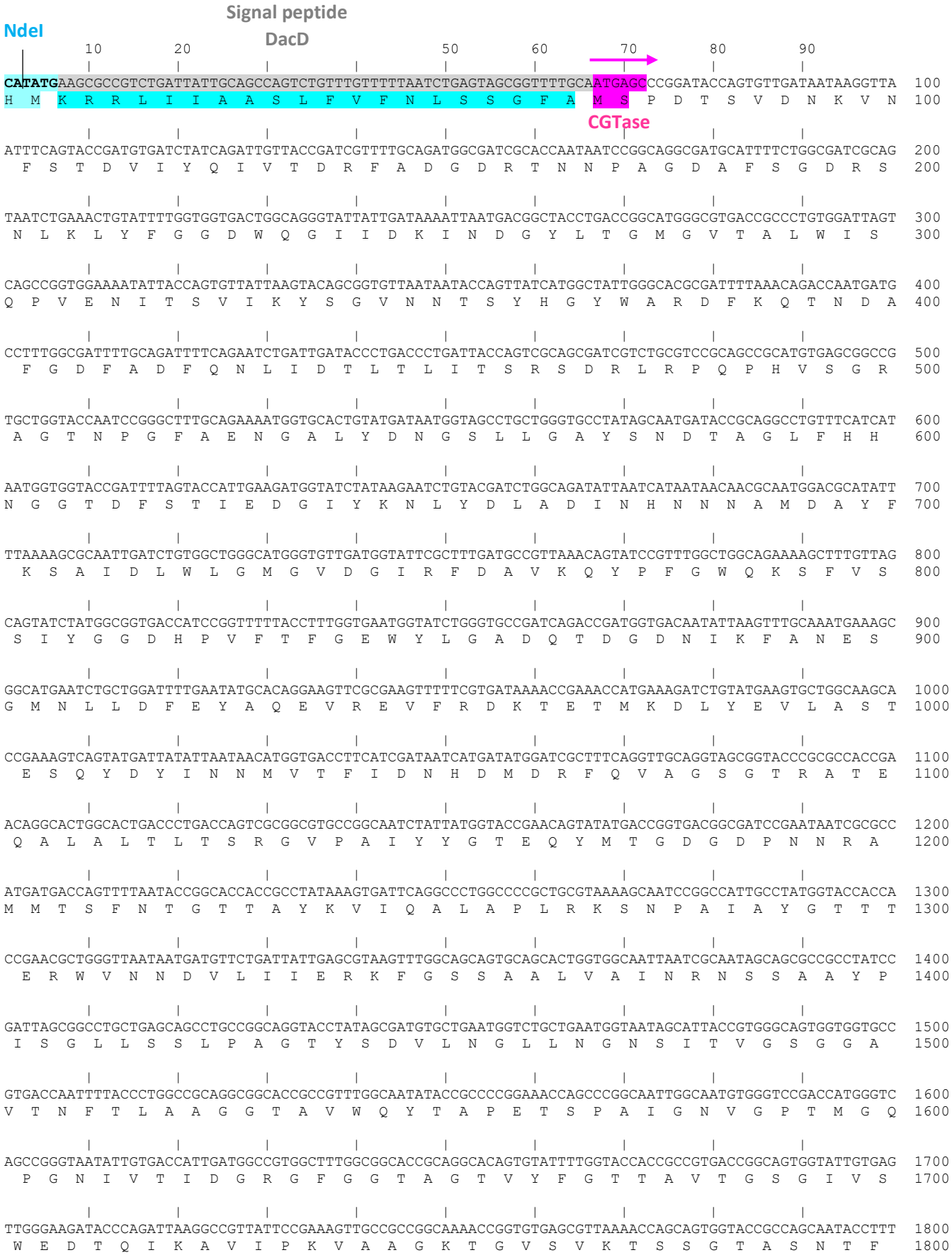
Apéndice 3. Secuencias GST·AS, CGTasa y fusión AS·CGTasa.

3.1 Secuencia GST·AS + Sitios de Restricción




CATATCTGGGTATTAGTGGCTATGATCATCGTCAGTTTCTGAATCGCTTTTTCGTTAATCGTTTTGATGGCAGTTTTGCACGCGGCGTTCCGTTTCAGTA 2000
Y L G I S G Y D H R Q F L N R F F V N R F D G S F A R G V P F Q Y 2000
TAATCCGAGTACCGGTGACTGTCGCGTGAGCGGTACCGCAGCAGCACTGGTGGGTCTGGCCCAGGATGATCCGCATGCAGTTGATCGTATTAAGCTGCTG 2100
N P S T G D C R V S G T A A A L V G L A Q D D P H A V D R I K L L 2100
TATAGTATTGCACTGAGCACCGGTGGCCTGCCGCTGATCTATCTGGGTGACGAAGTGGGCACCCTGAATGATGATGATTGGAGCCAGGATAGTAATAAGA 2200
Y S I A L S T G G L P L I Y L G D E V G T L N D D D W S Q D S N K S 2200
GCGATGATAGTCGTTGGGCCCATCGTCCGCGCTATAATGAAGCCCTGTATGCCCAGCGTAATGATCCGAGCACCGCAGCAGGCCAGATCTATCAGGGTCT 2300
D D S R W A H R P R Y N E A L Y A Q R N D P S T A A G Q I Y Q G L 2300
GCGCCACATGATTGCCGTTCGCCAGAGCAATCCGCGTTTTGATGGTGGTCGCCTGGTTACCTTTAATACCAATAATAAGCATATCATCGGCTATATTCGT 2400
R H M I A V R Q S N P R F D G G R L V T F N T N N K H I I G Y I R 2400
AATAATGCACTGCTGGCCTTTGGCAATTTTAGCGAATATCCGCAGACCGTTACCGCACATACCCTGCAGGCAATGCCGTTTAAAGCACATGATCTGATTG 2500
N N A L L A F G N F S E Y P Q T V T A H T L Q A M P F K A H D L I G 2500
GCGGTAAAACCGTGAGTCTGAATCAGGATCTGACCCTGCAGCCGTATCAGGTTATGTGGCTGGAAATTGCACATCATCATCATCACCATTAGGATCC 2600
G K T V S L N Q D L T L Q P Y Q V M W L E I A H H H H H H * D C 2600
BamHI

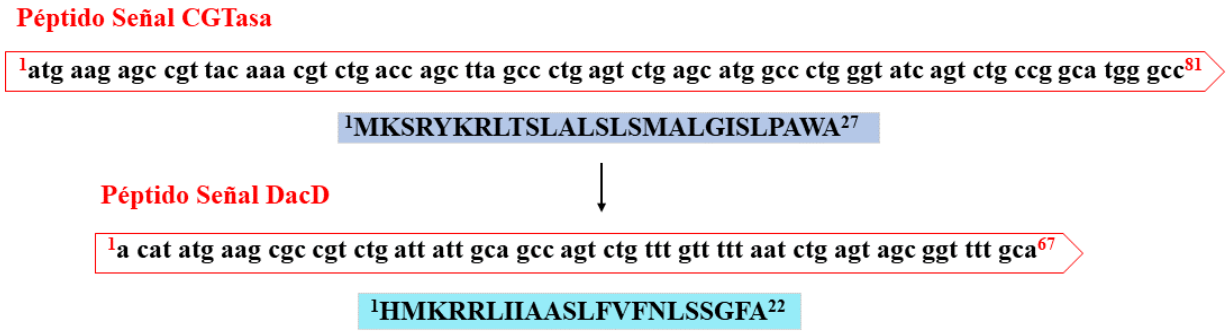
3.2 Secuencia CGTasa + Sitios de Restricción



```
         |         |         |         |         |         |         |         |         |
AAAAGTTTTAATGTGCTGACCGGCGATCAGGTGACCGTTCGTTTTCTGGTTAATCAGGCAAATACCAATTATGGTACCAATGTGTATCTGGTGGGTAATG  1900
K  S  F  N  V  L  T  G  D  Q  V  T  V  R  F  L  V  N  Q  A  N  T  N  Y  G  T  N  V  Y  L  V  G  N  A  1900

         |         |         |         |         |         |         |         |         |
CAGCAGAACTGGGTACCTGGGACCCTAATAAGGCCATTGGTCCGATGTATAATCAGGTTATTGCAAAATATCCGAGCTGGTATTATGATGTGAGTGTTCC  2000
  A  E  L  G  T  W  D  P  N  K  A  I  G  P  M  Y  N  Q  V  I  A  K  Y  P  S  W  Y  Y  D  V  S  V  P   2000

         |         |         |         |         |         |         |         |         |
GGCAGGTACAAAACTGGATTTTAAATTCATTAAGAAGGGCGGTGGTACCGTTACCTGGGAAGGTGGTGGTAATCATACCTATACCACCCCGGCAAGTGGC  2100
 A  G  T  K  L  D  F  K  F  I  K  K  G  G  G  T  V  T  W  E  G  G  G  N  H  T  Y  T  T  P  A  S  G   2100

         |         |         |         |         |         |         |         |         |
GTGGGCACCGTTACCGTTGATTGGCAGAATCATCATCATCATCACCATTAGGATCC                                               2200
V  G  T  V  T  V  D  W  Q  N  H  H  H  H  H  H  *  D  C                                               2200
```

BamHI

3.3 Esquema péptido señal CGTasa : Reemplazo del péptido señal CGTasa por DacD.



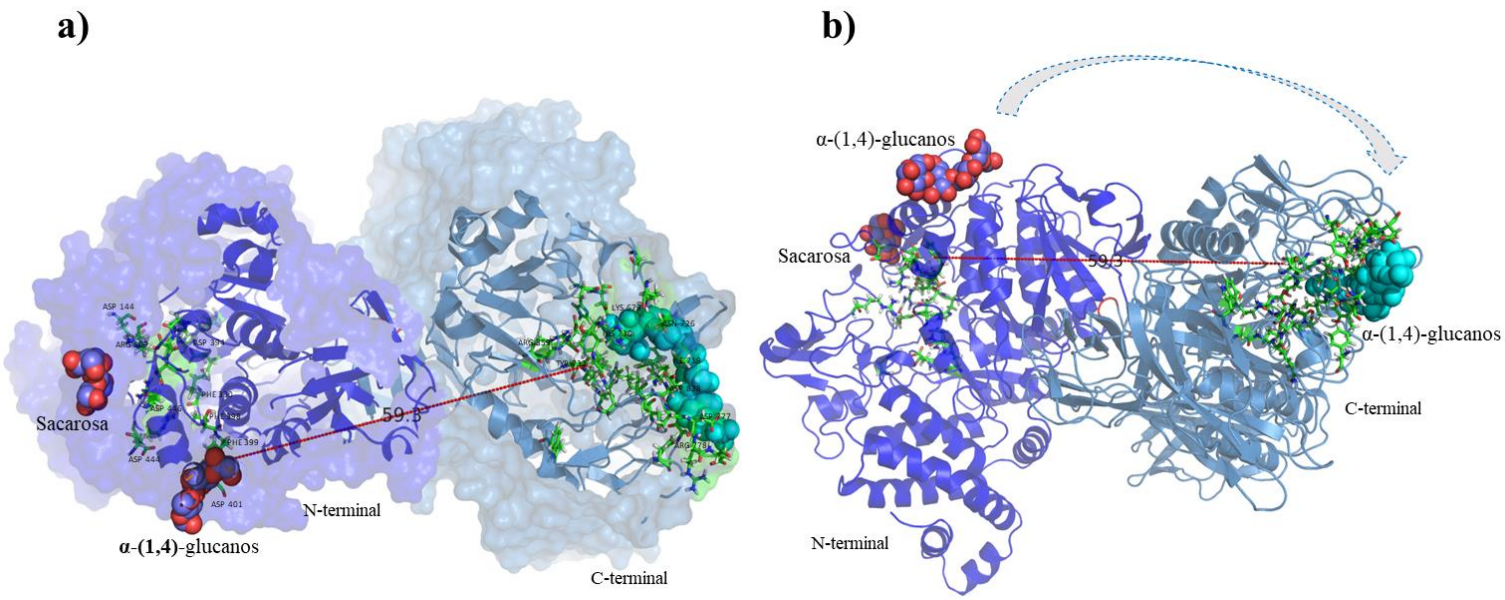
3.4 Secuencia Fusión AS·CGTasa + Sitios de Restricción

```
NdeI           Signal peptide DacD                                   NdeI
                    10        20        30        40        50        60        70        80        90
CATATGAAGCGCCGTCTGATTATTGCAGCAAGTCTGTTTGTTTTTAACCTGAGCAGCGGTTTTGCCGCCCATATGAATAGCCAGTATCTGAAAACCCGTA  100
H  M  K  R  R  L  I  I  A  A  S  L  F  V  F  N  L  S  S  G  F  A  A  H  M  N  S  Q  Y  L  K  T  R  I   100
                                                                         → Amilosacarasa

TTCTGGATATCTATACCCCGGAACAGCGCGCAGGCATTGAAAAAAGCGAAGATTGGCGCCAGTTTAGTCGTCGTATGGATACCCATTTTCCGAAACTGAT  200
 L  D  I  Y  T  P  E  Q  R  A  G  I  E  K  S  E  D  W  R  Q  F  S  R  R  M  D  T  H  F  P  K  L  M   200

GAATGAACTGGATAGTGTTTATGGTAATAACGAAGCCCTGCTGCCGATGCTGGAAATGCTGCTGGCACAGGCATGGCAGAGTTATAGTCAGCGCAATAGC  300
 N  E  L  D  S  V  Y  G  N  N  E  A  L  L  P  M  L  E  M  L  L  A  Q  A  W  Q  S  Y  S  Q  R  N  S    300

AGTCTGAAAGATATTGATATTGCCCGTGAAAATAACCCGGATTGGATTCTGAGCAATAAGCAGGTTGGTGGTGTGTGTTATGTGGATCTGTTTGCCGGTG  400
S  L  K  D  I  D  I  A  R  E  N  N  P  D  W  I  L  S  N  K  Q  V  G  G  V  C  Y  V  D  L  F  A  G  D  400

ACCTGAAAGGCCTGAAAGATAAAATTCCGTATTTTCAGGAACTGGGTCTGACCTATCTGCATCTGATGCCGCTGTTTAAATGTCCGGAAGGTAAAAGTGA  500
 L  K  G  L  K  D  K  I  P  Y  F  Q  E  L  G  L  T  Y  L  H  L  M  P  L  F  K  C  P  E  G  K  S  D   500

TGGTGGTTATGCAGTGAGCAGCTATCGTGATGTTAATCCGGCCCTGGGTACCATTGGCGATCTGCGCGAAGTGATTGCAGCACTGCATGAAGCCGGTATT  600
 G  G  Y  A  V  S  S  Y  R  D  V  N  P  A  L  G  T  I  G  D  L  R  E  V  I  A  A  L  H  E  A  G  I    600

AGTGCCGTTGTTGATTTTATTTTTAACCATACCAGCAACGAACATGAATGGGCCCAGCGTTGTGCAGCAGGTGACCCGCTGTTTGATAATTTTTATTATA  700
S  A  V  V  D  F  I  F  N  H  T  S  N  E  H  E  W  A  Q  R  C  A  A  G  D  P  L  F  D  N  F  Y  Y  I  700

TCTTCCCTGACCGCCGTATGCCGGATCAGTATGATCGCACCCTGCGTGAAATTTTTCCGGATCAGCATCCGGGTGGCTTTAGTCAGCTGGAAGATGGTCG  800
 F  P  D  R  R  M  P  D  Q  Y  D  R  T  L  R  E  I  F  P  D  Q  H  P  G  G  F  S  Q  L  E  D  G  R   800

CTGGGTTTGGACCACCTTTAATAGCTTTCAGTGGGATCTGAATTATAGTAATCCGTGGGTTTTTCGTGCCATGGCCGGTGAAATGCTGTTTCTGGCCAAT  900
 W  V  W  T  T  F  N  S  F  Q  W  D  L  N  Y  S  N  P  W  V  F  R  A  M  A  G  E  M  L  F  L  A  N    900

CTGGGCGTGGATATTCTGCGCATGGATGCCGTGGCATTCATTTGGAAACAGATGGGTACCAGCTGTGAAAATCTGCCGCAGGCCCATGCCCTGATTCGTG  1000
L  G  V  D  I  L  R  M  D  A  V  A  F  I  W  K  Q  M  G  T  S  C  E  N  L  P  Q  A  H  A  L  I  R  A  1000

CCTTTAATGCCGTTATGCGTATTGCAGCCCCGGCAGTGTTTTTCAAAAGCGAAGCAATTGTGCATCCGGATCAGGTGGTGCAGTATATTGGCCAGGATGA  1100
 F  N  A  V  M  R  I  A  A  P  A  V  F  F  K  S  E  A  I  V  H  P  D  Q  V  V  Q  Y  I  G  Q  D  E   1100

ATGCCAGATTGGTTATAATCCGCTGCAGATGGCACTGCTGTGGAATACCCTGGCCACCCGCGAAGTGAATCTGCTGCATCAGGCACTGACCTATCGTCAT  1200
 C  Q  I  G  Y  N  P  L  Q  M  A  L  L  W  N  T  L  A  T  R  E  V  N  L  L  H  Q  A  L  T  Y  R  H    1200

AATCTGCCGGAACATACCGCATGGGTGAATTATGTGCGCAGCCATGATGATATTGGCTGGACCTTTGCCGATGAAGATGCAGCATATCTGGGCATTAGCG  1300
N  L  P  E  H  T  A  W  V  N  Y  V  R  S  H  D  D  I  G  W  T  F  A  D  E  D  A  A  Y  L  G  I  S  G  1300

GTTATGATCATCGTCAGTTTCTGAATCGCTTTTTCGTGAATCGTTTTGATGGTAGTTTTGCACGCGGTGTGCCGTTTCAGTATAATCCGAGTACCGGTGA  1400
 Y  D  H  R  Q  F  L  N  R  F  F  V  N  R  F  D  G  S  F  A  R  G  V  P  F  Q  Y  N  P  S  T  G  D   1400

CTGTCGCGTTAGCGGTACCGCCGCCGCCCTGGTTGGCTTAGCCCAGGATGATCCGCATGCAGTGGATCGTATTAAGCTGCTGTATAGCATTGCCCTGAGT  1500
 C  R  V  S  G  T  A  A  A  L  V  G  L  A  Q  D  D  P  H  A  V  D  R  I  K  L  L  Y  S  I  A  L  S    1500

ACCGGTGGCCTGCCGCTGATCTATCTGGGTGACGAAGTGGGCACCCTGAATGATGATGATTGGAGTCAGGATAGCAATAAGAGTGATGATAGCCGCTGGG  1600
T  G  G  L  P  L  I  Y  L  G  D  E  V  G  T  L  N  D  D  D  W  S  Q  D  S  N  K  S  D  D  S  R  W  A  1600

CACATCGTCCGCGTTATAATGAAGCACTGTATGCCCAGCGTAATGATCCGAGCACCGCAGCCGGCCAGATCTATCAGGGCCTGCGTCACATGATTGCCGT  1700
 H  R  P  R  Y  N  E  A  L  Y  A  Q  R  N  D  P  S  T  A  A  G  Q  I  Y  Q  G  L  R  H  M  I  A  V   1700

GCGTCAGAGCAATCCGCGTTTTGATGGCGGCCGCCTGGTGACCTTTAATACCAATAATAAGCATATCATCGGCTATATTCGCAATAATGCACTGCTGGCC  1800
 R  Q  S  N  P  R  F  D  G  G  R  L  V  T  F  N  T  N  N  K  H  I  I  G  Y  I  R  N  N  A  L  L  A    1800
```

```
TTTGGTAATTTTAGCGAATATCCGCAGACCGTTACCGCCCATACCCTGCAGGCAATGCCGTTTAAAGCACATGATCTGATTGGTGGTAAAACCGTGAGTC  1900
F  G  N  F  S  E  Y  P  Q  T  V  T  A  H  T  L  Q  A  M  P  F  K  A  H  D  L  I  G  G  K  T  V  S  L  1900

TGAATCAGGATCTGACCCTGCAGCCGTATCAGGTTATGTGGCTGGAAATTGCAGGTGGTGGTAGCATGAGCCCGGATACCAGCGTTGATAATAAGGTGAA  2000
  N  Q  D  L  T  L  Q  P  Y  Q  V  M  W  L  E  I  A  G  G  G  S  M  S  P  D  T  S  V  D  N  K  V  N  2000

TTTTAGTACCGATGTGATCTATCAGATTGTGACCGATCGCTTTGCCGATGGCGATCGTACCAATAATCCGGCAGGTGACGCCTTTAGCGGTGACCGTAGC  2100
 F  S  T  D  V  I  Y  Q  I  V  T  D  R  F  A  D  G  D  R  T  N  N  P  A  G  D  A  F  S  G  D  R  S   2100

AATCTGAAACTGTATTTTGGTGGTGACTGGCAGGGCATTATTGATAAAATTAATGATGGTTACCTGACCGGCATGGGCGTGACCGCACTGTGGATTAGCC  2200
N  L  K  L  Y  F  G  G  D  W  Q  G  I  I  D  K  I  N  D  G  Y  L  T  G  M  G  V  T  A  L  W  I  S  Q  2200

AGCCGGTTGAAAATATTACCAGCGTGATTAAGTATAGCGGTGTTAATAATACCAGTTATCATGGTTATTGGGCCCGTGATTTTAAACAGACCAATGATGC  2300
  P  V  E  N  I  T  S  V  I  K  Y  S  G  V  N  N  T  S  Y  H  G  Y  W  A  R  D  F  K  Q  T  N  D  A   2300

CTTTGGCGATTTTGCAGATTTTCAGAATCTGATTGATACCCTGACCCTGATTACCAGTCGTAGTGATCGTCTGCGCCCGCAGCCGCATGTTAGTGGTCGC  2400
 F  G  D  F  A  D  F  Q  N  L  I  D  T  L  T  L  I  T  S  R  S  D  R  L  R  P  Q  P  H  V  S  G  R   2400

GCAGGTACCAATCCGGGTTTTGCCGAAAATGGCGCCCTGTATGATAATGGCAGCCTGCTGGGCGCCTATAGCAATGATACCGCCGGTCTGTTTCATCATA  2500
A  G  T  N  P  G  F  A  E  N  G  A  L  Y  D  N  G  S  L  L  G  A  Y  S  N  D  T  A  G  L  F  H  H  N  2500

ATGGCGGTACCGATTTTAGTACCATTGAAGATGGCATCTATAAAAATCTGTACGATCTGGCCGATATTAATCATAATAATAACGCAATGGACGCATATTT  2600
  G  G  T  D  F  S  T  I  E  D  G  I  Y  K  N  L  Y  D  L  A  D  I  N  H  N  N  N  A  M  D  A  Y  F   2600

TAAGAGCGCAATTGATCTGTGGCTGGGTATGGGCGTTGATGGTATTCGTTTTGATGCCGTTAAACAGTATCCGTTTGGCTGGCAGAAAAGCTTTGTTAGT  2700
 K  S  A  I  D  L  W  L  G  M  G  V  D  G  I  R  F  D  A  V  K  Q  Y  P  F  G  W  Q  K  S  F  V  S   2700

AGCATCTATGGCGGCGATCATCCGGTGTTTACCTTTGGTGAATGGTATCTGGGCGCCGATCAGACCGATGGTGACAATATTAAGTTTGCAAATGAAAGCG  2800
S  I  Y  G  G  D  H  P  V  F  T  F  G  E  W  Y  L  G  A  D  Q  T  D  G  D  N  I  K  F  A  N  E  S  G  2800

GTATGAATCTGCTGGATTTTGAATATGCCCAGGAAGTTCGCGAAGTTTTTCGCGATAAAACCGAAACCATGAAAGATCTGTATGAAGTGCTGGCCAGCAC  2900
  M  N  L  L  D  F  E  Y  A  Q  E  V  R  E  V  F  R  D  K  T  E  T  M  K  D  L  Y  E  V  L  A  S  T   2900

CGAAAGTCAGTATGATTATATTAATAACATGGTGACCTTCATCGATAATCATGATATGGATCGTTTTCAGGTTGCCGGCAGCGGCACCCGTGCCACAGAA  3000
 E  S  Q  Y  D  Y  I  N  N  M  V  T  F  I  D  N  H  D  M  D  R  F  Q  V  A  G  S  G  T  R  A  T  E   3000

CAGGCACTGGCCCTGACCCTGACCAGTCGTGGTGTGCCGGCCATCTATTATGGTACCGAACAGTATATGACCGGTGACGGCGATCCGAATAATCGTGCCA  3100
Q  A  L  A  L  T  L  T  S  R  G  V  P  A  I  Y  Y  G  T  E  Q  Y  M  T  G  D  G  D  P  N  N  R  A  M  3100

TGATGACCAGTTTTAATACCGGTACCACCGCATATAAAGTGATTCAGGCACTGGCGCCGCTGCGCAAAAGTAATCCGGCCATTGCCTATGGTACCACCAC  3200
  M  T  S  F  N  T  G  T  T  A  Y  K  V  I  Q  A  L  A  P  L  R  K  S  N  P  A  I  A  Y  G  T  T  T   3200

CGAACGTTGGGTTAATAATGATGTGCTGATTATTGAACGTAAGTTTGGCAGCAGTGCCGCACTGGTGGCCATTAATCGCAATAGCAGCGCCGCCTATCCG  3300
 E  R  W  V  N  N  D  V  L  I  I  E  R  K  F  G  S  S  A  A  L  V  A  I  N  R  N  S  S  A  A  Y  P   3300

ATTAGTGGTCTGCTGAGTAGTCTGCCGGCAGGTACCTATAGTGATGTGCTGAATGGTCTGCTGAATGGTAATAGCATTACCGTGGGCAGCGGTGGTGCAG  3400
I  S  G  L  L  S  S  L  P  A  G  T  Y  S  D  V  L  N  G  L  L  N  G  N  S  I  T  V  G  S  G  G  A  V  3400

TGACCAATTTTACCCTGGCCGCAGGTGGCACCGCAGTGTGGCAGTATACCGCCCCGGAAACCAGCCCGGCAATTGGTAATGTTGGTCCGACCATGGGTCA  3500
  T  N  F  T  L  A  A  G  G  T  A  V  W  Q  Y  T  A  P  E  T  S  P  A  I  G  N  V  G  P  T  M  G  Q   3500

GCCGGGCAATATTGTTACCATTGATGGTCGTGGCTTTGGTGGCACCGCCGGTACCGTTTATTTTGGTACCACCGCCGTTACCGGCAGTGGCATTGTGAGC  3600
 P  G  N  I  V  T  I  D  G  R  G  F  G  G  T  A  G  T  V  Y  F  G  T  T  A  V  T  G  S  G  I  V  S   3600

TGGGAAGATACCCAGATTAAGGCAGTTATTCCGAAAGTGGCAGCCGGCAAAACCGGTGTGAGTGTTAAAACCAGCAGTGGTACCGCAAGTAATACCTTTA  3700
W  E  D  T  Q  I  K  A  V  I  P  K  V  A  A  G  K  T  G  V  S  V  K  T  S  S  G  T  A  S  N  T  F  K  3700
```

```
        |         |         |         |         |         |         |         |         |
AAAGCTTTAATGTGCTGACCGGTGACCAGGTGACCGTTCGCTTTCTGGTTAATCAGGCAAATACCAATTATGGTACCAATGTTTATCTGGTTGGCAATGC  3800
  S   F   N   V   L   T   G   D   Q   V   T   V   R   F   L   V   N   Q   A   N   T   N   Y   G   T   N   V   Y   L   V   G   N   A   3800

        |         |         |         |         |         |         |         |         |
CGCCGAACTGGGTACCTGGGACCCTAATAAGGCAATTGGTCCGATGTATAATCAGGTTATTGCAAAATATCCGAGCTGGTATTATGATGTTAGTGTGCCG  3900
 A   E   L   G   T   W   D   P   N   K   A   I   G   P   M   Y   N   Q   V   I   A   K   Y   P   S   W   Y   Y   D   V   S   V   P    3900

        |         |         |         |         |         |         |         |         |
GCCGGCACCAAACTGGATTTTAAATTCATTAAGAAGGGCGGCGGCACCGTGACCTGGGAAGGTGGTGGTAATCATACCTATACCACCCCGGCAAGTGGTG  4000
A   G   T   K   L   D   F   K   F   I   K   K   G   G   G   T   V   T   W   E   G   G   G   N   H   T   Y   T   T   P   A   S   G   V  4000
|         |         |         |         |         |         |         |         |
TTGGCACCGTGACAGTTGATTGGCAGAATCATCATCATCATCACCATTAGGATCC                                              4100
  G   T   V   T   V   D   W   Q   N   H   H   H   H   H   H   *   D   C                              4100
```

Apéndice 4. Representación 3D de los sitios activos de AS·CGTasa (modelo $AS_{L1}CGT$) con sacarosa y α-(1,4)-glucanos.



a) Distancia entre los sitios de unión de los α-(1,4)-glucanos en la AS y la CGTasa (antes de la conversión a CDs) y posible ruta para la formación de los canales intramoleculares.

b) Posible ruta de libre difusión del producto-sustrato α-(1,4)-glucano de la AS hacía el sitio de unión de la CGTasa.