



UNIVERSIDAD NACIONAL **UNAM**
AUTÓNOMA DE MÉXICO **POSGRADO** 

FACULTAD DE QUÍMICA

PROGRAMA DE MAESTRÍA Y DOCTORADO
EN CIENCIAS BIOQUÍMICAS

DISEÑO DE PROTEÍNAS DE FUSION PARA
UNION A CARBOHIDRATO CON FINES
ANALÍTICOS Y BIOTECNOLOGICOS

T E S I S A

QUE PARA OBTENER EL GRADO DE :

MAESTRO EN CIENCIAS (BIOQUÍMICAS)

P R E S E N T A :

DANIEL ALEJANDRO GUILLÉN SANTOS



Tutor: DRA. ROMINA RODRÍGUEZ SANOJA

MÉXICO, D. F.

Julio 2010



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

DISEÑO DE PROTEÍNAS DE FUSION PARA UNION A CARBOHIDRATO CON FINES ANALÍTICOS Y BIOTECNOLOGICOS

RECONOCIMIENTOS

Esta tesina de maestría se realizó bajo la dirección de la Dra. Romina Rodríguez Sanoja en el Departamento de Biología Molecular y Biotecnología del Instituto de Investigaciones Biomédicas de la Universidad Nacional Autónoma de México.

Durante la elaboración de la tesina goce de una beca otorgada por CONACYT.

Esta tesina fue defendida en examen presentado el día

El Jurado de Examen Doctoral estuvo constituido por:

Presidente	Dr. Agustín Lopez-Munguía Canales	Instituto de Biotecnología, UNAM
Vocal	Dr. Ismael Bustos Jaimes	Facultad de Medicina, UNAM
Vocal	Dra. Ma. Imelda López Villaseñor	Instituto Investigaciones Biomédicas, UNAM
Vocal	Dr. Enrique García Hernandez	Instituto de Química, UNAM
Secretario	Dra. Amelia Farrés González-Saravia	Facultad de Química, UNAM
Suplente	Dr. Roberto Arreguín Espinosa	Instituto de Química, UNAM
Suplente	Dr. Rogelio Rodríguez Sotres	Facultad de Química, UNAM

Diseño de proteínas de fusión para unión a carbohidratos con fines analíticos o biotecnológicos

Introducción

Definición y Propiedades de los **Módulos de Unión a Carbohidrato (MUC)**

Clasificación de los MUC

- Clasificación en base a su secuencia.
- Clasificación en base a su estructura.
- Clasificación en base a la topología del sitio de unión.

Interacción Carbohidrato-MUC

Especificidad de los MUC

Elaboración de Proteínas de Fusión con Afinidad a Carbohidrato

Diseño de un Péptido Antimicrobiano (PAM) Fusionado a un Dominio de Unión a Almidón

Aplicaciones de los MUC

Conclusión

Introducción

Los carbohidratos son moléculas que juegan un papel muy importante en diversos procesos biológicos. Se les puede encontrar siendo parte de elementos estructurales y protectores, unidos a otras moléculas como lípidos o proteínas formando glicoconjugados, involucrados en procesos de reconocimiento, señalización y adhesión celular y además son la fuente de energía por excelencia de los seres vivos.

Para que los carbohidratos puedan intervenir en los procesos mencionados, es necesario que exista un reconocimiento de sus estructuras. Para ello, animales, plantas y microorganismos sintetizan una gran variedad de proteínas con capacidad de modificar o reconocer a los carbohidratos. Entre dichas proteínas se encuentran lectinas, anticuerpos, transportadores de azúcares, glucósido-hidrolasas y glucosil-transferasas, por mencionar algunos ejemplos.

Los numerosos estudios existentes acerca de las proteínas con capacidad de unirse a carbohidrato y los avances en biología estructural, han permitido reconocer ciertas similitudes entre las distintas clases de proteínas que reconocen carbohidratos. Algunas de estas similitudes se observan a nivel estructural y en la naturaleza de los residuos que interaccionan con los carbohidratos. La racionalización de esta información nos ayuda a comprender como es que se llevan a cabo, en los seres vivos, las diferentes funciones en las que los carbohidratos se encuentran involucrados. Por otra parte, es posible emplear dichos conocimientos en el desarrollo de nuevas tecnologías con aplicación tanto en ciencia como en industria.

Lectinas y en menor medida en anticuerpos, son de las proteínas con afinidad a carbohidrato más estudiadas, sin embargo, cada vez se ha puesto mayor atención a las glucósido-hidrolasas y glucosil-transferasas. La estructura de muchas de estas proteínas ha sido dilucidada y se ha observado la presencia de una región con la propiedad de discriminar entre los diferentes arreglos que pueden tener los carbohidratos así como mediar la unión de la proteína a esta clase de sustratos. Dicha región se pliega de manera independiente y no presenta actividad catalítica por lo que ha sido denominada como dominio o módulo de unión a carbohidrato (MUC).

Las características que permiten el reconocimiento específico o la promiscuidad en el reconocimiento han generado un importante número de estudios, en especial por la

posibilidad de aplicaciones que estos dominios tienen. La definición de estas características, así como la posibilidad de utilizar estos dominios en proteínas de fusión con fines analíticos o biotecnológicos es analizada en el presente documento.

Definición y Propiedades de los MUC

Actualmente, a los MUC se les define como una secuencia de aminoácidos localizada en una enzima que presenta actividad sobre carbohidrato, dicha región se pliega de manera autónoma y es capaz de unirse a uno o más carbohidratos (1, 2). Algunas excepciones de MUCs independientes son los Modulos de Unión a Celulosa (CBM por sus siglas en inglés) que actúan como andamiaje del celulosoma y algunos otros casos aislados. Los MUC se encuentran extensamente distribuidos en la naturaleza y reconocen azúcares, enlaces y estructuras tan diversos como los que se encuentran presentes en celulosa, quitina, almidón, glucógeno, pululano, inulina, β -glucanos, fucosa, lactosa, galactosa, arabinofuranosa, ácido poligalacturónico, *N*-acetilglucosamina, lipopolisacáridos, el antígeno A/B del grupo sanguíneo, entre otros.

En enzimas como las glucósido hidrolasas y las glucosil-transferasas, los MUC se pueden localizar, en el extremo N- o C-terminal de la enzima, entre dos dominios catalíticos, o bien, una misma enzima puede poseer más de un MUC y éstos ubicarse en tándem o estar separados por otros dominios, también puede darse el caso de que un MUC constituya toda la proteína, aunque esto es menos frecuente (Fig 1).

En el caso de múltiples MUC presentes en una sola enzima, estos pueden compartir la especificidad por un sustrato (3), presentar diferente afinidad por un mismo sustrato (4, 5) o bien pueden tener especificidades diferentes (6, 7). La presencia de múltiples MUC puede conferir un incremento en la afinidad por sustrato en la enzima que los posee (3,

8), permitir el reconocimiento de diferentes regiones en un mismo sustrato o aumentar el rango de polisacáridos que pueden ser reconocidos (9, 10).

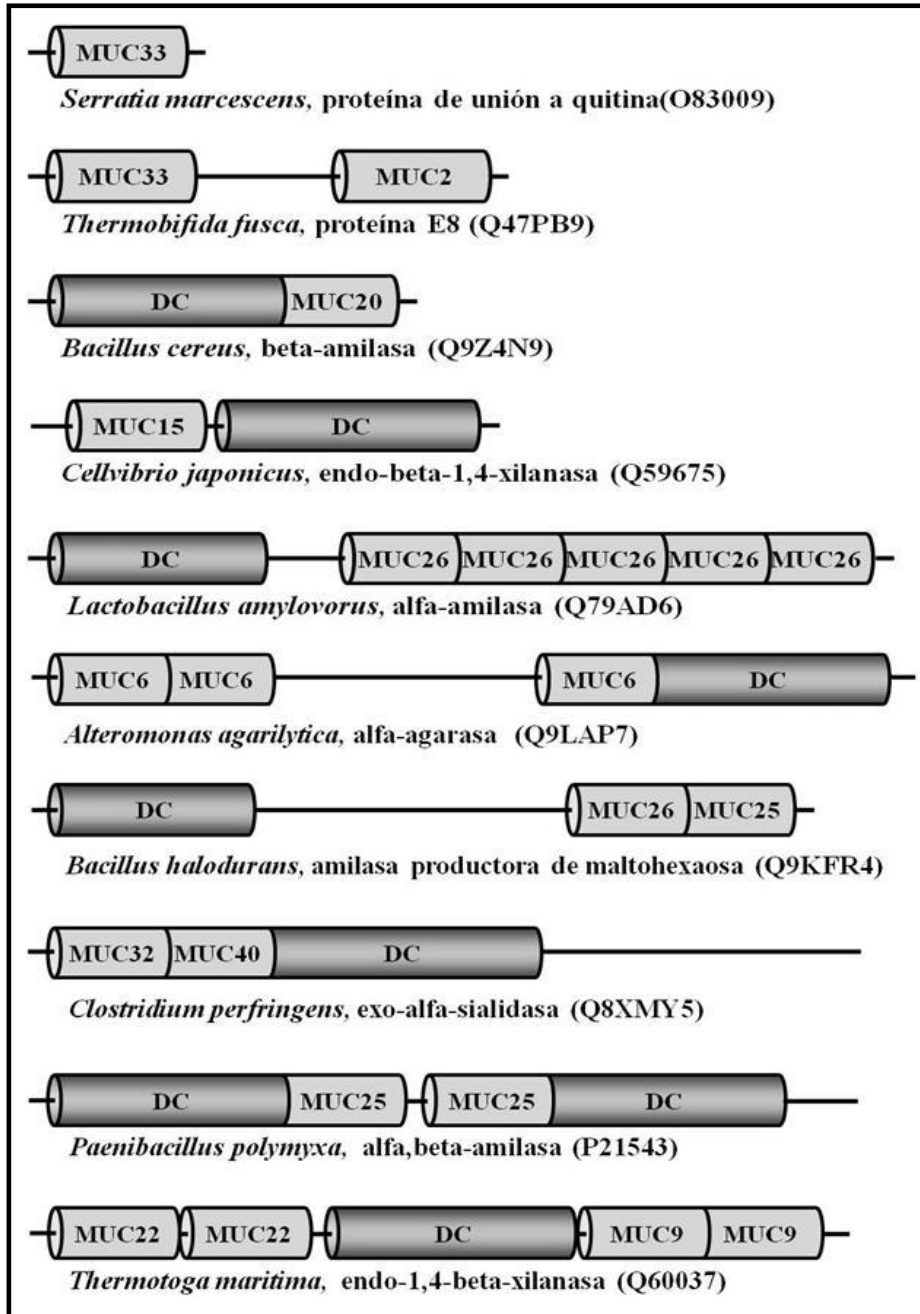


Figura 1: Distribución de MUC en diferentes proteínas. Pueden presentarse en diferente número y localización en la cadena polipeptídica. DC simboliza la posición del dominio catalítico en varias glucósido-hidrolasas. La clave en paréntesis corresponde a la base de datos UniProt.

Tomando en cuenta lo antes mencionado, la principal función de los MUC es anclar a la enzima al sustrato insoluble (celulosa, almidón, quitina, etc.) y por tanto incrementar el tiempo de contacto del sitio catalítico con el sustrato (11), lo que conduce a un aumento en la actividad enzimática. La relevancia de los MUC como dominios auxiliares presentes en las enzimas, se denota por la reducción y hasta la pérdida total de la actividad enzimática contra sustratos insolubles cuando el MUC es removido de la enzima a la cual pertenece (12). Sin embargo, es de notar que la actividad frente a sustratos solubles no se ve afectada la mayoría de las veces (6, 13) y en algunas ocasiones se ve mejorada con respecto a la original (14).

Además, se ha reportado que algunos MUC tienen la capacidad de desorganizar la superficie del sustrato al cual se unen, como consecuencia, las cadenas de polisacáridos adquieren un arreglo menos cristalino, facilitándose el ataque enzimático sobre el polisacárido insoluble. Esta es una función que había sido controversial en años pasados, pero los resultados obtenidos en diferentes estudios dirigidos para demostrar esta propiedad en algunos MUC son convincentes (1). Una prueba de la capacidad de los MUC para desorganizar la estructura del sustrato y facilitar la hidrólisis es dada por la proteína de unión a quitina CBP21 de *Serratia marcescens*, que promueve la degradación de quitina cristalina por quitinasas a través de interacciones polares que modifican el arreglo del sustrato (15).

Clasificación de los MUC

- **Clasificación en base a su secuencia**

Originalmente los MUC fueron clasificados como dominios de fijación a celulosa, basados en el hecho del descubrimiento inicial de varios módulos que fijaban celulosa, sin embargo, conforme el conocimiento y el número de MUC se iba incrementando,

surgió la necesidad de tener un sistema de clasificación más generalizado. En la actualidad, los MUC se encuentran clasificados en función de su secuencia de aminoácidos en 59 familias (10) (CAZY -carbohydrate active enzyme database, <http://www.cazy.org/>). Debido a que el plegamiento de los MUC se encuentra mejor conservado que las secuencias, algunas de las familias se pueden agrupar en clanes o superfamilias dependiendo de su estructura.

En dicha clasificación, existen familias de MUC que presentan afinidad por un único ligando o sus derivados, como la familia 17 que presenta afinidad sólo por celulosa amorfa o las familias 12 y 18 en las cuales solo se ha reportado que presentan afinidad por quitina. También hay casos de familias de MUC cuyos miembros presentan afinidad por distintos carbohidratos, tal es el caso de la familia 37 que incluye miembros con afinidad por celulosa, xilano, quitina y algunos sustratos más heterogéneos como tallo de plátano o pared celular de alfalfa.

La clasificación mencionada, ha resultado muy útil para la identificación de secuencias que podrían corresponder a MUC aún no identificados experimentalmente, en algunos casos para predecir el ligando al cual se une un dominio, revelar posibles relaciones evolutivas, inferir el plegamiento y ayudar en la identificación de residuos de aminoácidos involucrados en la unión a sustrato (16) ya que generalmente se encuentran conservados (Fig 2).

Las familias de MUC que tienen especificidad por diversos ligandos, conservan a los aminoácidos que interactúan directamente con el sustrato pero presentan cambios en aminoácidos funcionales secundarios, lo que proporciona diversidad en la especificidad (17).

P96992	AGGSWTAVPGVGMR-AACAGWKKDVDLGAATSLTAAFNNGNGVWDNNKGAD	230
Q9L036	AGGSWTAVPGVGMR-AACAGWKKDVDLGAATSLTAAFNNGNGVWDNNKGAD	230
C4EHJ9	TGGTWTVPVPGVAMDETACADWARKTVDLGTATGLKAAFNNGSGTWDNNNGAD	583
C7R010	DGGSWTSVPGDSMT-AACNGWVTRTIAS-AGKNVTAAFNNGSGTWDNNSGKD	575
Q5WHK9	NQEAWTTLPGVPLNSSEYEGYAKITIEETEEGSQLRAAFNNGRGQWDNNQGRD	942
Q9KFR4	NQGAWTTLPGVPLTKSEYEGYVKVTIEAEEGSQLRAAFNNGSGQWDNNQGRD	944
Q08T95	AGGTWTTTPPGIAMPDAEVPGHAKYTVNLATATQLEAVFNNGSGTWDSDNNGNN	1062
P21543	AGGSWTAAAPGVKMQDAEISGYAKITVDIGSASQLEAAFNDGNNNWDSDNNTKN	537

P62695	HYGQCGGIGYSGPTVCASGTTTCQVLNPFYYSQCL	513
P62694	HYGQCGGIGYSGPTVCASGTTTCQVLNPFYYSQCL	513
Q9P8P3	HYGQCGGTGWTPTRCASGYTCQVLNPFYYSQCL	505
Q7LIJ0	QWGQCGGIGYSGSTTCASPYTCHVLNPFYYSQCY	510
C5MRS3	QWGQCGGQGYTGPTACVSGTTCKAQNPFYYSQCL	477
A1DD18	LYGQCGSGWTGPTACASGATCKVLNSYYSQCL	216
Q5ZNB1	HWGQCGGIGWSGPTICVSPYTCQVLNPFYYSQCL	407
O42638	HWGQCGGIGYTGPTVCASGFTCHVLNPFYYSQCY	32

Fig. 2: Alineamiento de MUC de la familia 25 con afinidad a almidón (arriba) y la familia 1 con afinidad a celulosa (abajo). Residuos implicados en la unión se denotan en cajas grises. Las claves corresponden a UniProt.

- **Clasificación en base a su estructura**

De las 59 familias reportadas, se conoce la estructura de al menos un miembro de 41 de ellas. A pesar de las numerosas familias en que los MUC se han clasificado y de la diversidad en estructura primaria, los arreglos que adquieren los MUC son relativamente pocos, siendo el plegamiento dominante el de β -sandwich (2, 10). Este arreglo consta de dos capas de láminas- β , cada una formada por tres a seis láminas- β antiparalelas. Además del arreglo en β -sandwich, los MUC presentan otros seis diferentes arreglos, lo que ha llevado a su clasificación o agrupación con base en su plegamiento (10) (Figura 3 y Tabla 1).

Es de mencionar que, aunque los MUC pueden ser agrupados en familias con base en los plegamientos conservados que se presentan, dicha clasificación no sirve para predecir la posible especificidad, ya que MUC que poseen plegamientos similares pueden tener diferencias en los elementos funcionales como lo son los aminoácidos involucrados con la unión a sustrato.

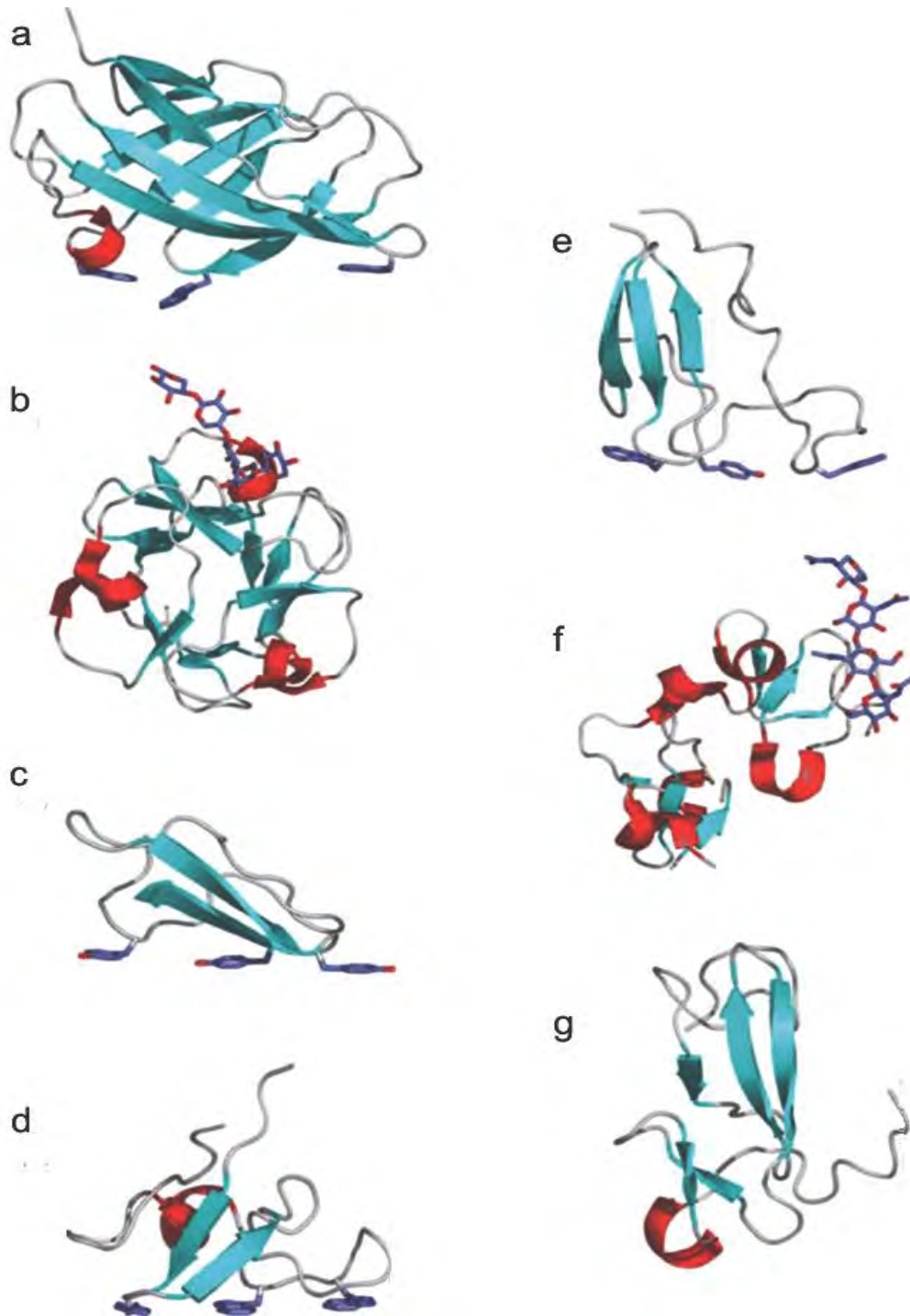


Figura 3: Imagen de los diversos plegamientos de los MUC. **a.** MUC con plegamiento β -sandwich (1EXG); **b.** MUC con plegamiento β -trébol en complejo con xilopentaosa (1MC9); **c.** MUC con plegamiento de nudo de cisteínas (1CBH); **d.** MUC con plegamiento de unión a oligo- nucleótido/sacárido (1E830); **e.** MUC con plegamiento único (clave PDB 1AIW); **f.** MUC con plegamiento de heveina unido a quitotriosa (1EN2); **g.** MUC con plegamiento único similar a heveina (1DQC). Las claves en paréntesis corresponden al Protein Data Bank. Modificada de (10).

Familias por plegamiento	Plegamiento	Familias de MUC
1	β -sandwich	2, 3, 4, 6, 9, 11, 15, 16, 17, 20, 21, 22, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 40, 41, 42, 44, 47, 48, 51
2	β -trébol	13, 42
3	Nudo de Cisteínas	1
4	Único	5, 12
5	Plegamiento de unión a oligo- nucleótido/sacárido	10
6	Plegamiento de heveina	18
7	Único (Similar a heveina)	14

Tabla 1: Clasificación de los MUC en 7 diferentes familias en base a su plegamiento.

- **Clasificación con base en su topología**

Tomando en cuenta las similitudes estructurales y funcionales de los MUC, se ha propuesto otra clasificación basándose en la topología del sitio de unión a sustrato, que como resultado los agrupa en tres distintos tipos.

Los MUC tipo A se caracterizan por sus residuos de aminoácidos que adoptan una arquitectura plana, es decir que forman una superficie que es complementaria a la superficie plana del sustrato al cual se unen; por ejemplo, la celulosa y la quitina son

sustratos reconocidos por MUC tipo A, en cambio, este tipo de MUC muestran poca o nula afinidad por polisacáridos solubles (18).

En contraste con lo anterior, la topología del sitio de unión de los MUC tipo B es descrita como una “hendidura” que varía en profundidad y en la cual se acomodan cadenas sencillas de polisacáridos. Los aminoácidos involucrados en la unión a carbohidrato adoptan una arquitectura tipo sándwich o torcida que permite cierta plasticidad en el reconocimiento tanto de la composición como del tipo de enlace presente en el esqueleto de una cadena de polisacárido así como de los diferentes carbohidratos que se encuentran unidos a la cadena principal, así encontramos en este grupo proteínas que se unen a xilano, celulosa, laminarina, liquenano, etc (4).

El último grupo corresponde a los MUC tipo C o tipo-lectina que carecen de hendiduras extendidas como los MUC tipo B y en consecuencia únicamente se unen óptimamente a mono-, di- o trisacáridos. Entre los ligandos reconocidos por los MUC tipo C se encuentran los extremos reductores de polisacáridos como celulosa o xilano y glicanos presentes en la superficie de células eucariotas (10, 19) (Fig 4).

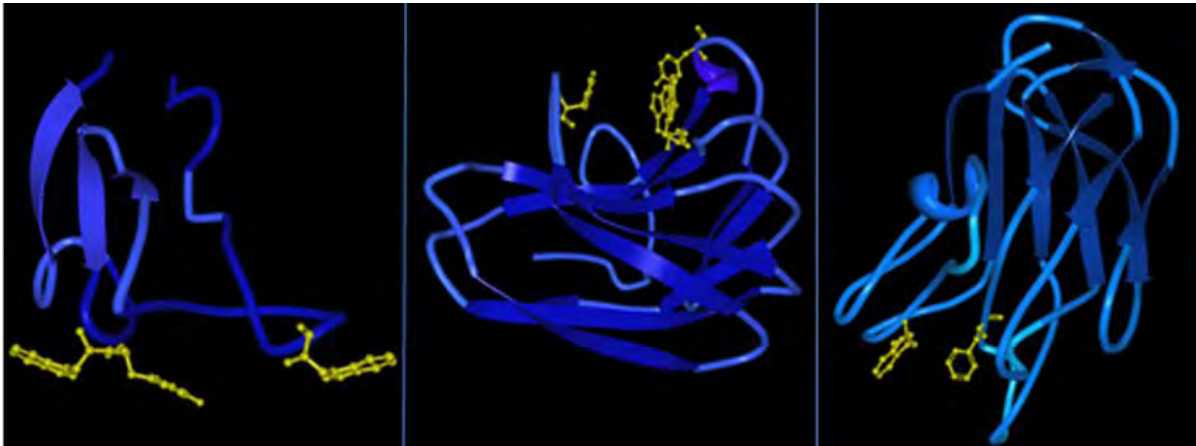


Figura 4: Diferentes topologías de los MUC. De izquierda a derecha: MUC Tipo A de la endoglucanasa Z de *Erwinia chrysantemi*, clave PDB 1AIW; MUC Tipo B de la laminarina de *Thermotoga marítima*, clave PDB 1GUI; MUC Tipo C de la *N*-acetil-hexosaminidasa de *Clostridium perfringens*, clave PDB 2J1E. Los residuos en varillas y esferas corresponden a los residuos implicados en la unión a carbohidrato.

Especificidad de los MUC

Los MUC pueden ser muy específicos en el reconocimiento de un ligando o pueden ser promiscuos y reconocer a varios azúcares. La razón para esta plasticidad o especificidad recae en la topología del sitio de unión así como en la orientación de los residuos de aminoácidos involucrados en la interacción con el carbohidrato (10).

Como ejemplo de lo anterior, en un estudio de afinidad por pared celular de plantas, se emplearon cuatro MUC tipo A que tienen afinidad por celulosa cristalina y tres MUC tipo B con afinidad por celulosa amorfa. Los resultados obtenidos demostraron diferencias en el reconocimiento de pared celular primaria y secundaria entre los MUC tipo A y los MUC tipo B, incluso se observó que el reconocimiento entre los cuatro MUC tipo A era dependiente del tipo de tejido y del origen botánico. Los MUC tipo B también presentaron diferente reconocimiento de la pared celular vegetal dependiendo de su origen, dando muestra de una alta especificidad en el reconocimiento de su ligando (9).

Contrario al ejemplo anterior donde diferentes tipos de MUC reconocen específicamente distintas estructuras de la pared celular de plantas, un estudio realizado con un MUC tipo B de la familia 29 proveniente de *Pyromices equi*, demostró su afinidad por varios ligandos entre los que se incluyen xilano, galactomanano, glucomanano e hidroxietilcelulosa (20). La plasticidad en reconocimiento que presenta le permite unirse a distintas estructuras de la planta, lo cual puede ser aprovechado para llevar a cabo una degradación más eficiente de la pared celular, ya que esta proteína forma parte de un complejo enzimático que incluye la presencia de celulasas y hemicelulasas (21, 22).

Resulta particularmente interesante que diferentes MUC que parecen tener especificidad *in vitro* por un solo polisacárido purificado, son capaces de discriminar en el tejido los distintos arreglos que pueden presentar los carbohidratos, así pues, la diversidad en MUC es un reflejo de la diversidad en carbohidratos y la heterogeneidad estructural que pueden adoptar (9). Por otra parte, la presencia de MUC con cierta promiscuidad por el ligando puede conducir a un reconocimiento más amplio de las diferentes estructuras que presentan los carbohidratos, por ejemplo los que componen a la pared celular de plantas, confiriendo cierta ventaja a las enzimas o proteínas a las cuales se encuentran anexados o asociados dichos MUC.

Interacción Carbohidrato-MUC

Con la finalidad de determinar a los residuos de aminoácidos involucrados en la unión a carbohidrato, diversos estudios estructurales y de mutagénesis sitio-dirigida se han llevado a cabo (23-25). En general, los residuos que se encuentran conservados en las diversas familias de MUC y que interactúan directamente con el sustrato son residuos aromáticos, entre ellos triptófano y tirosina son los que se reportan con mayor frecuencia. Una observación común es que la mutación de estos residuos lleva asociada consigo la reducción o pérdida de la función de unión de estas proteínas.

En la unión MUC-carbohidrato, las cadenas laterales de los residuos aromáticos crean interacciones de tipo hidrofóbico con las unidades de carbohidratos, mientras que algunos de los residuos polares que se encuentran cercanos forman puentes de hidrógeno, ya sea directamente con los grupos hidroxilo presentes en el carbohidrato o bien mediados por moléculas de agua, lo que permite estabilizar la interacción (Fig 5).

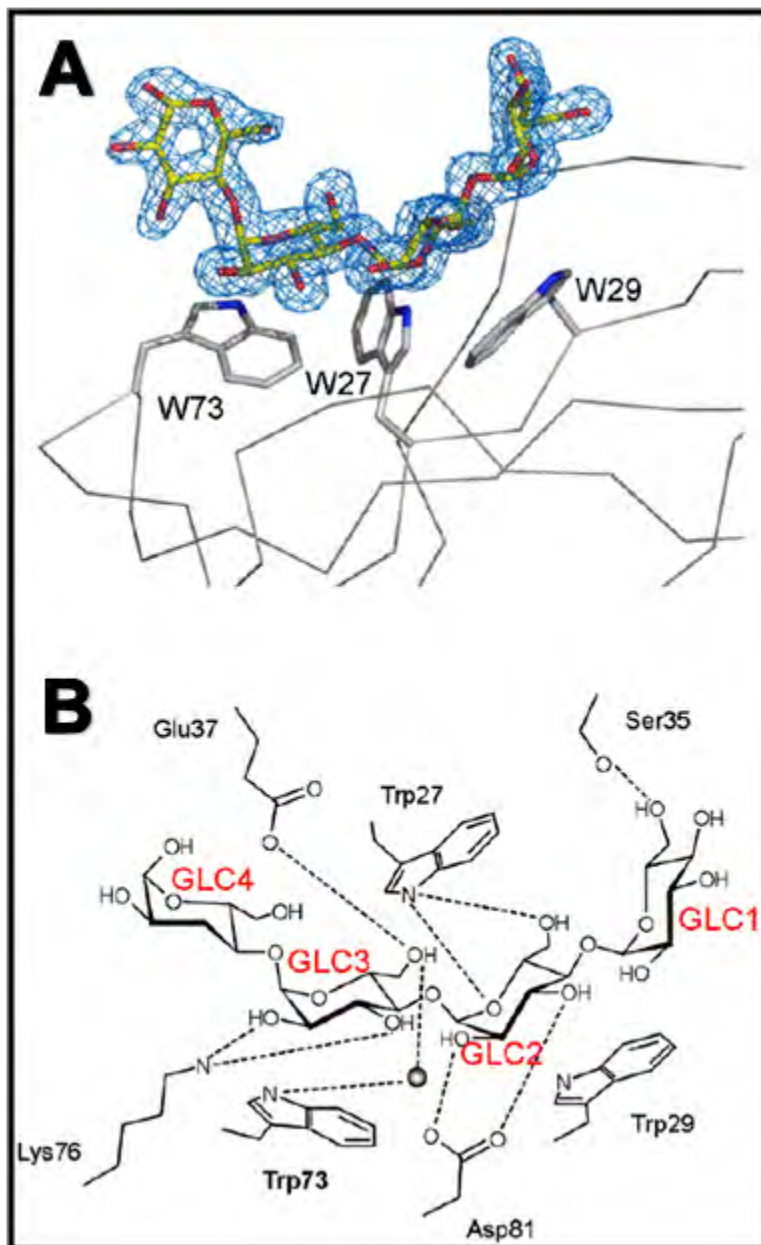


Figura 5: Esquema de un MUC de la familia 41 en presencia de maltotetraosa. (A) Los residuos aromáticos del sitio de unión se muestran en varillas y el mapa de densidad electrónica se muestra alrededor del ligando. (B) Arquitectura del sitio de unión con los posibles puentes de hidrógeno que se forman entre el sitio de unión y el ligando. Tomada de (26).

Dependiendo del tipo de MUC, los puentes de hidrógeno cobran mayor o menor importancia. En los MUC tipo A por ejemplo, el mutar residuos polares que forman puentes de hidrógeno directamente con polisacáridos, tiene poco impacto en la afinidad de la proteína por su ligando. Esto sugiere que las interacciones hidrofóbicas desempeñan el papel más importante, mientras que los puentes de hidrógeno no tienen tanta relevancia en los MUC tipo A. Por otro lado, mutación de residuos involucrados en la formación de puentes de hidrógeno en los MUC tipo B y C, provoca la disminución en afinidad e incluso la total pérdida de la capacidad de unión, por tanto en los MUC tipo B y C, los puentes de hidrógeno como las interacciones hidrofóbicas tienen un rol significativo en la interacción con el carbohidrato (10).

Como ya se había mencionado, de los análisis estructurales realizados, se deduce que los residuos aromáticos involucrados en la unión a sustrato, se encuentran orientados de tal manera que reflejan la estructura del carbohidrato al cual se unen. Por otra parte, existen MUC de diferentes familias que presentan un plegamiento muy parecido y una orientación similar del sitio de unión a carbohidrato, mientras que los residuos que interactúan directamente con el ligando no son compartidos o bien, las diferencias se presentan en residuos funcionales secundarios. Dichas diferencias modifican la topología del sitio de unión y dictan la especificidad de los MUC pertenecientes a diferentes familias o incluso que pertenecen a una misma familia (27, 28).

La comprensión de características de los MUC como su especificidad o la forma en que interactúan con su ligando, ha servido como plataforma para poder llevar a cabo diversos estudios, dentro de los cuales, se incluye el empleo de los MUC en proteínas de fusión para conferir capacidad de unirse a carbohidrato a una proteína de interés.

Elaboración de Proteínas de Fusión con Afinidad a Carbohidrato

Entre las razones existentes para preferir diseñar proteínas que presenten afinidad a carbohidratos, es que éstos resultan ser matrices que se encuentran de manera abundante, por ejemplo la celulosa y la quitina son los polímeros más abundantes que existen en la naturaleza y otros polisacáridos, como almidón y lignina, también se encuentran ampliamente distribuidos. Además de su abundancia, son biodegradables y biocompatibles, de hecho la celulosa, el almidón y la quitina tienen diversas aplicaciones farmacéuticas y en otras industrias como la alimentaria y de cosméticos; resultan ser agentes químicamente inertes a los que se les puede dar diferentes presentaciones como filtros, membranas, polvos, fibras, geles, etc. Todo esto en conjunto, hace de los carbohidratos matrices simples y económicas, comparados con otras matrices que son más complejas y requieren de modificaciones químicas o de acoplar a sus ligandos de manera covalente, lo que aumenta su costo y limita su uso en aplicaciones a gran escala.

En cuanto a los MUC, éstos se han ido estableciendo como una alternativa para el diseño de proteínas de fusión con afinidad a carbohidrato, debido a propiedades ya discutidas, como el plegamiento autónomo que les permite ser funcionales en proteínas de fusión y su especificidad de unión.

Por otra parte, la especificidad y afinidad de los MUC por su ligando puede ser manipulada (29-32), lo que se puede aprovechar para adaptar la unión a carbohidrato a conveniencia de las condiciones en las que se quiere producir la proteína o de los usos que se le quiere dar (Fig. 6). Se ha reportado también que algunos MUC, pueden ser de utilidad para mejorar la estabilidad y solubilidad de las proteínas en las cuales se encuentran incorporados (33, 34).

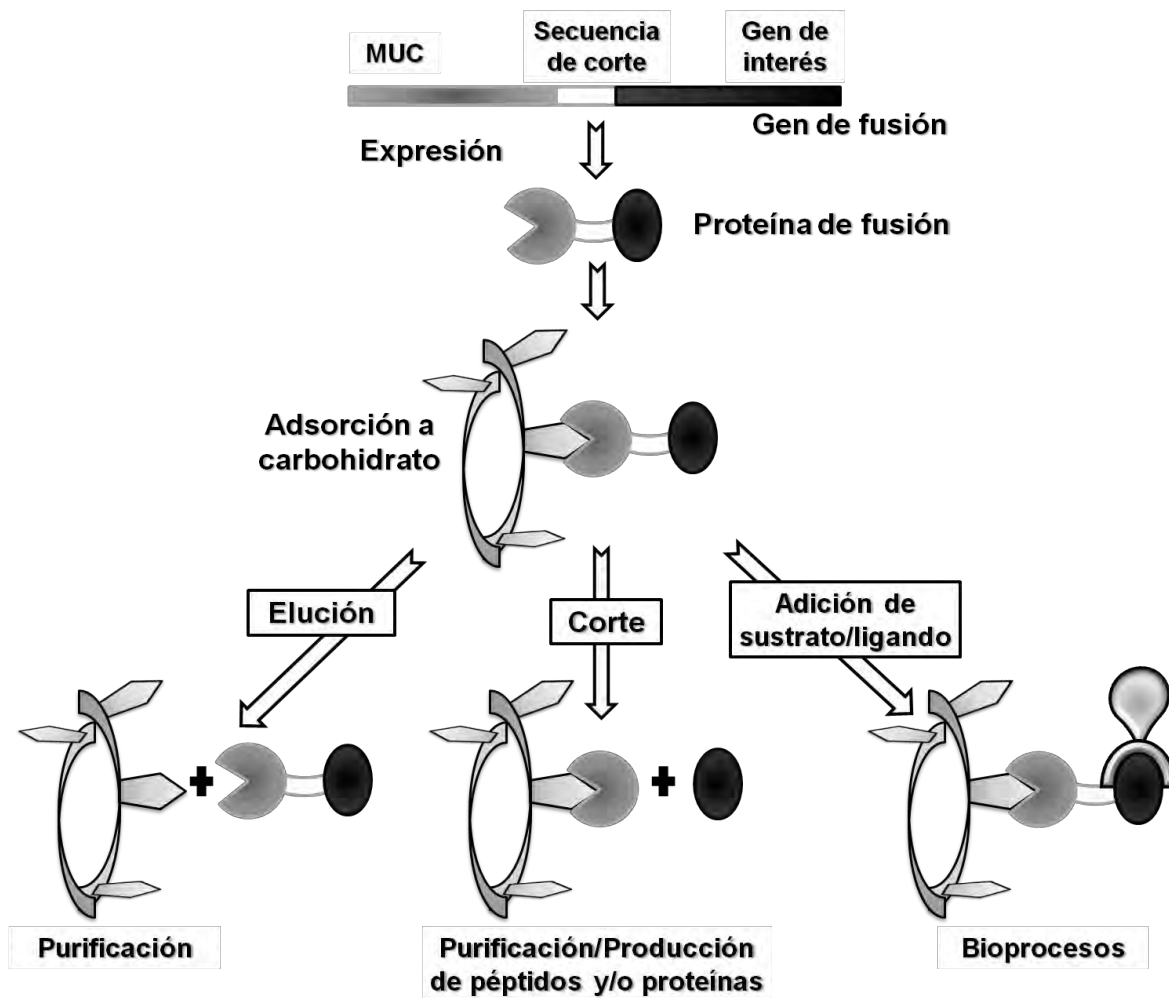


Figura 6: Generación de una proteína de fusión incluyendo la presencia de un MUC. Dependiendo del interés, la proteína de fusión puede recuperarse de la matriz donde se encuentra adsorbida, inmovilizarse para llevar a cabo bioprocesos o recuperarse sólo la proteína de interés al separarla del MUC.

La elección de un determinado tipo de MUC al momento de diseñar una proteína de fusión con afinidad a carbohidrato, dependerá de factores como la naturaleza de la proteína a la cual se desea fusionar, la aplicación que se le va a dar a la proteína recombinante, el carbohidrato por el cual se desea que tenga especificidad y la fuerza con que se desea que se una la proteína. A continuación se presentan algunos ejemplos de aplicaciones desarrolladas y publicadas de los MUC en proteínas de fusión con afinidad a carbohidrato.

Aplicaciones de los MUC

MUC como Etiquetas de Afinidad

Como proteínas de fusión con afinidad por ligando, los MUC compiten con otras proteínas o etiquetas que confieren afinidad por distintas matrices a las proteínas a las cuales se encuentran fusionados, entre los sistemas de etiquetas de afinidad actualmente en el mercado se encuentran la etiqueta de histidinas, el péptido de unión a calmodulina y la glutatión S-transferasa, entre otros (Tabla 2).

Etiqueta de afinidad	Tamaño (kDa)	Matriz/ligando	Elución
Poli-His	0.84	Ion metálico de transición inmovilizado	Imidazol 50 - 250mM
Proteína A	31	IgG	pH 2.8 - 3.0
Glutatión S-transferasa	26	Glutatión	5 - 10mM de glutatión reducido
Proteína de unión a maltosa	40	Amilosa entrecruzada	10mM de maltosa
Biotina	0.24	Avidina monomérica	20mM de biotina
Péptido de unión a calmodulina	2.96	Calmodulina	EGTA
Dominio de unión a celulosa	4-20	Celulosa	Urea 4M o etilenglicol
Dominio de unión a quitina	5.59	Quitina	30 - 50 mM ditiotreitól o β -mercaptoetanol

Tabla 2: Etiquetas de afinidad comúnmente usadas y sus propiedades.

Como etiquetas de afinidad “ideales”, los MUC y otros tipos de etiquetas, deben cumplir con ciertas características como: 1) adsorción y elución de una matriz en un solo paso, 2) no afectar la estructura terciaria ni la actividad biológica de la proteína a la cual se está fusionando, 3) ser fácil de remover de la proteína quimérica y 4) ser aplicable a diferentes proteínas (35). A la fecha existen diversos estudios utilizando MUC como etiquetas de afinidad a diferentes polisacáridos (36-38). Entre los que se usan con más frecuencia, se encuentran los dominios de unión a celulosa (DUC), algunos de los cuales se unen irreversiblemente a la celulosa, entendiendo por irreversibilidad el hecho de que la concentración de MUC adsorbido no se reduce a pesar de remover o diluir la fase soluble. Éstos MUC catalogados como de unión irreversible, son empleados para inmovilización de proteínas de fusión o de células (39), mientras que los que se unen reversiblemente, son más utilizados para purificación (35).

Ejemplos de MUC que reconocen otros polisacáridos incluyen al dominio de unión a quitina (DUQ), el cual se ha empleado directamente fusionado a la proteína de interés (40, 41), o bien, separado de está por regiones denominadas inteínas, las cuales son secuencias de aminoácidos a las que se les puede inducir auto-proteólisis para separar la proteína de interés (35, 42). Por su parte, los dominios de unión a almidón también han encontrado aplicación como etiquetas de afinidad con éxito (43), aunque no de manera comercial.

Aunque la elección de la etiqueta de afinidad a fusionar en una proteína depende de las características de dicha proteína y del uso al cual está destinada, los MUC tienen gran potencial como etiquetas de afinidad en proteínas recombinantes o si se emplean MUC con diferente especificidad, mejorar la purificación de la proteína recombinante mediante un segundo paso de cromatografía de afinidad; estudios sobre estas posibilidades proveerán más información en un futuro cercano.

Además de la purificación de proteínas recombinantes, la inmovilización de proteínas de fusión es también uno de los usos más frecuentes que se les da a los MUC. Para inmovilizar se usan MUC que se unen de manera irreversible a su sustrato, al contrario de los que se emplean como etiquetas de purificación que se unen de manera reversible.

Las enzimas inmovilizadas en los carbohidratos mantienen su actividad, incluso en algunos casos se reporta un aumento en la actividad enzimática después de ser inmovilizada (1, 44, 45) y pueden ser usadas de manera continua en reactores de lecho fluidizado y/o en procesos cíclicos (46, 47).

Inmovilización celular

Los MUC no sólo pueden inmovilizar enzimas, también pueden ser empleados para la inmovilización de células a través de la presentación en membrana de éstos dominios. Por ejemplo, se ha inmovilizado *E. coli*, que expresa en su superficie un DUC proveniente de *Cellulomonas fimi*, en diferentes soportes de celulosa, las monocapas de células son capaces de unirse en un amplio rango de pH, siendo el grado de inmovilización dependiente de la cantidad de DUC expuesto en la superficie (48). Resultados similares se han obtenido utilizando quitina como soporte y el DUQ de la quitinasa A1 de *Bacillus circulans* WL-12 (49).

Una estrategia diferente que se siguió para la inmovilización de células de mamífero en celulosa bacteriana, fue la fusión del DUC de *C. fimi* al péptido de unión celular RGD; de esta manera, la proteína de fusión tiene capacidad de adherirse a la celulosa por un extremo y unirse a la célula por el otro, anclando las células a la matriz de celulosa (50).

Biorremediación

Recientemente, se modificó genéticamente a *E. coli* de tal manera que expresara de manera simultánea un DUC y una hidrolasa de compuestos órganofosforados. La bacteria es capaz de fijarse a una matriz de celulosa y de hidrolizar compuestos organofosforados que son muy tóxicos para el sistema nervioso. La alta capacidad de degradación de este sistema lo hace atractivo para su aplicación en descontaminación, así como en destoxificación y para que en un futuro se emplee en filtros contra esta clase de compuestos, presentes por ejemplo, en el gas nervioso y algunos plaguicidas (51).

En la búsqueda de sistemas eficientes para la depuración de agentes contaminantes, se han utilizado enzimas recombinantes fusionadas a distintos MUC, algunos de los ejemplos incluyen enzimas inmovilizadas para remover metales pesados como cadmio con gran eficacia (39) o bien, enzimas como la atrazina clorohidrolasa acopladas a resinas de celulosa que pueden ser usadas para destoxificar aguas de desecho industrial (45)

MUC como herramientas analíticas y de diagnóstico

El uso de un reportero como la proteína verde fluorescente (GFP) fusionada a distintos MUC, ha permitido la localización *in vivo* y el análisis de los polisacáridos presentes en distintos organismos. Usando un MUC con afinidad a manano fusionado a la GFP, se visualizó la regeneración de la pared celular de un alga roja (52). Empleando también la fusión DUA-GFP se ha logrado hacer el seguimiento de almidón de granos de maíz durante su extracción para su uso como materia prima (53).

Pensando en la rápida detección de microorganismos medicamente relevantes, también se han empleado a los MUC. La fusión de dos DUC provenientes de *Trichoderma reesei*

permite la detección específica de quistes de *Acanthamoeba* que contienen celulosa, a diferencia de muchos otros que contienen quitina, lo que permite pensar en la aplicación de este sistema como una herramienta diagnóstica para la detección de protozoarios que presentan estructuras con celulosa (54).

Para detectar a *Staphylococcus aureus*, se diseñó una proteína pentamérica consistente en cinco DUC y cinco anticuerpos de un solo dominio variable. El pentámero conserva la propiedad de unirse a celulosa y a las células de *S. aureus*. El diseño de esta proteína permitió la captura y detección de *S. aureus* que se hizo pasar a través de filtros de celulosa incubados previamente con el pentámero (55).

Para el análisis por ELISA de anticuerpos IgG, se diseñó una fusión entre la proteína A de *S. aureus* y un DUC de *Clostridium cellulovorans*. La proteína de fusión se adsorbe en microplatos con celulosa, lo que permite la posterior captura de IgG por medio de la proteína A. Esta misma fusión se utilizó para llevar a cabo la detección de IgG en western-blot, donde la proteína quimérica se empleó en lugar del anticuerpo secundario, revelando con partículas de celulosa teñidas (56).

Otras aplicaciones en biomedicina

El uso de los MUC se va volviendo cada vez más amplio, por ejemplo, la fusión de una citocina a un DUC, permite la inmovilización de la proteína quimérica en una matriz de celulosa y potencia el efecto en la estimulación y proliferación de líneas celulares dependientes de la citocina (57). En inmunología, antígenos fusionados a MUC que reconocen y se anclan a la pared celular de bacterias lácticas son prometedores para la administración oral de vacunas vivas (58) además, el uso de polisacáridos como vehículos para la administración de antígenos adsorbidos a ellos vía MUC es otro tema que se está estudiando (59).

Como se ha descrito hasta el momento, el uso de los MUC en proteínas de fusión puede ser muy variado, no obstante, resulta difícil generalizar una sola estrategia para la elaboración de proteínas de fusión que contengan MUC, sin embargo, en el siguiente ejemplo se describe el diseño de una proteína fusionada a un MUC tomando en cuenta los detalles a cuidar durante la confección de la proteína recombinante con afinidad a carbohidrato

Diseño de un Péptido Antimicrobiano Fusionado a un Dominio de Unión a Almidón

El objetivo del ejemplo a desarrollar es diseñar una proteína de fusión entre un Péptido Antimicrobiano (PAM) de interés y un MUC con afinidad a almidón. Dicha proteína tendría la propiedad de adsorberse al almidón y presentar actividad antimicrobiana, generándose novedosos bioproductos con propiedades antimicrobianas, además, el almidón puede ser usado con fines farmacéuticos, por su inocuidad.

Por otra parte, usualmente los PAM son péptidos pequeños, lo que dificulta su expresión de manera individual. Los bajos niveles de expresión junto con toxicidad para la célula, proteólisis y pobre solubilidad, son problemas que se pueden superar mediante la fusión a un MUC (60).

Estrategia general:

- 1) *El Dominio de unión a almidón (DUA) a utilizar.* Si se desea conferirle afinidad por almidón a una proteína de fusión, la búsqueda de un DUA se debe llevar a cabo en las enzimas adecuadas. Entre las glucósido hidrolasas, las familias de las α -amilasas, glucoamilasas y β -amilasas son las que cuentan con más representantes y cerca del 10% de estas enzimas poseen un DUA (61). Uno de los DUA más estudiados es el de la glucoamilasa de *Rhizopus oryzae*, el cual tiene comprobada

función de adsorción al almidón insoluble y a algunos oligosacáridos (62, 63), por lo que resulta un candidato viable para fusionarse a una proteína.

- 2) *Características de la construcción.* El siguiente aspecto a decidir, es la posición donde se desea fusionar el DUA con respecto al PAM, en la glucoamilasa de *R. oryzae*, el DUA se encuentra ubicado en el extremo N-terminal. Sin embargo, estudios colocándolo en la posición N- o C-terminal de una proteína como la verde fluorescente, han demostrado su eficiencia para adsorber la proteína de fusión al almidón, independientemente de su localización (43). Para efectos del ejemplo que se desarrolla, el gen del DUA de *R. oryzae* se clona en el extremo 5' del gen del PAM, es decir, se coloca en el extremo N-terminal de la proteína de fusión, simulando la posición que tiene en la enzima original.

En el diseño de la proteína de fusión, se puede incorporar una secuencia de corte específica para enterocinasa, dicha secuencia se colocaría entre el DUA y el PAM y serviría para poder separarlos en caso de ser necesario. Por otra parte, la enterocinasa en una proteasa intestinal, la acción de la enterocinasa permitiría una liberación del PAM *in vivo* en el tracto intestinal, de tal manera que se pueda dirigir la acción del PAM contra microorganismos patógenos causantes de enfermedades intestinales. Una segunda función de la secuencia de corte para enterocinasa se relaciona con estudios de inmuno-detección, ELISAs, Western-blot, etc., ya que comercialmente existen anticuerpos dirigidos contra esta secuencia y que permitirían llevar a cabo dichos ensayos.

En sus enzimas, los DUA pueden estar localizados de manera contigua al dominio catalítico o separados por medio de una secuencia entre 10 y 20 aminoácidos ó linker O-glicosilado (61). En estudios con proteínas de fusión, se ha reportado que la separación por medio de esta secuencia confiere flexibilidad

y puede favorecer el plegamiento correcto y la acción independiente de las proteínas o dominios fusionados (64, 65). La glucoamilasa de *R. oryzae* presenta un linker que separa al DUA del dominio catalítico. Aprovechando que de manera natural esta enzima posee el DUA y un linker, se propondría clonar la región correspondiente al DUA incluyendo también la región que corresponde al linker.

Resumiendo, la proteína de fusión con afinidad a carbohidrato incluiría el DUA en el extremo N-terminal seguido de un linker, la secuencia de corte para enterocinasa y finalmente el péptido de interés en el extremo C-terminal.

- 3) *Sistema de expresión.* Proteínas de fusión que incluyen MUC han sido expresadas en *E. coli* y en eucariontes como *Pichia pastoris* (43, 66) con buenos resultados. Sin embargo, específicamente para fusiones conteniendo el DUA de *R. oryzae*, la expresión y purificación a partir de *E. coli* ha resultado más eficiente que en *P. pastoris* (43), por lo que en este ejemplo, *E. coli* sería el sistema de expresión elegido.
- 4) *Purificación.* La adición del DUA a la proteína de fusión podría ejercer función de etiqueta de purificación, evitando la incorporación de alguna otra etiqueta de afinidad, además de que la matriz de purificación resultaría menos compleja como ya se discutió anteriormente. Ya que se tiene la proteína recombinante pura, se analizaría la capacidad de adsorción de esta proteína al almidón insoluble, permitiendo de esta manera, cuantificar la cantidad máxima de proteína que se adsorbe por miligramo de almidón, además se evaluaría la actividad antimicrobiana de la proteína de fusión adsorbida al almidón y así poder determinar su efectividad como agente antimicrobiano.

5) *Optimización de la función.* Por último, habría que analizar la posibilidad de introducir mejoras en el diseño de la proteína en caso de que fuera necesario, por ejemplo, la afinidad de la proteína se podría mejorar agregando más de un MUC, ya que se ha reportado que la presencia de múltiples MUC pueden mejorar la afinidad hacia un ligando (3, 8). Por otra parte, los MUC fusionados al antígeno de interés pueden tener diferente especificidad, por ejemplo a celulosa y a almidón, esto pensando en aprovechar las nuevas tecnologías en polímeros biodegradables basados en estos dos polisacáridos que pueden usarse para el transporte de moléculas de interés (67), en este caso la proteína de fusión.

Conclusión

En el diseño de proteínas de fusión con afinidad a carbohidrato, los MUC resultan ser una alternativa interesante y empleada con frecuencia para generar esta clase de proteínas. La comprensión de las propiedades de los MUC y la forma en que éstos reconocen e interactúan con su ligando, ha sido de gran importancia para generar diversos estudios que tienen la finalidad de explotar las cualidades de los MUC en distintos campos de investigación, sobre todo en el de la biotecnología como componentes moleculares en proteínas de fusión.

Dada la variedad de proteínas a las que se pueden fusionar los MUC, la amplia variedad de carbohidratos existentes en la naturaleza y las ventajas que su utilización significa, las aplicaciones de los MUC en proteínas de fusión son muy diversas y generalmente, con buenos resultados.

Finalmente, aunque en este trabajo se discute el empleo de los MUC en proteínas de fusión, es de mencionar que el uso de MUC como proteínas independientes también se encuentra bajo investigación. Por ejemplo, el tratamiento de materiales que contienen

celulosa como papel o algodón con DUC puede alterar propiedades mecánicas y/o de afinidad por colorantes de dichos materiales (1), también como agentes antifúngicos (68) y en productos de higiene oral donde la adición de MUC puede ayudar a remover y prevenir la formación de la placa dentobacteriana (69), Estas características hace de los MUC una herramienta de la que se puede obtener bastante provecho en futuras investigaciones.

BIBLIOGRAFIA

- 1 Shoseyov, O., Shani, Z. & Levy, I. (2006) Carbohydrate binding modules: Biochemical properties and novel applications *Microbiol. Mol. Biol. Rev.*, 70(2), pp. 283-295.
- 2 Hashimoto, H. (2006) Recent structural studies of carbohydrate-binding modules *Cell. Mol. Life Sci.*, 63(24), pp. 2954-2967.
- 3 Guillen, D., Santiago, M., Linares, L., Perez, R., Morlon, J., Ruiz, B., Sanchez, S. & Rodriguez-Sanoja, R. (2007) Alpha-amylase starch binding domains: cooperative effects of binding to starch granules of multiple tandemly arranged domains *Appl. Environ. Microbiol.*, 73(12), pp. 3833-3837.
- 4 Najmudin, S., Guerreiro, C.I.P.D., Carvalho, A.L., Prates, J.A.M., Correia, M.A.S., Alves, V.D., Ferreira, L.M.A., Romao, M.J., Gilbert, H.J., Bolam, D.N. & Fontes, C.M.G.A. (2006) Xyloglucan is recognized by carbohydrate-binding modules that interact with beta-glucan chains *J. Biol. Chem.*, 281(13), pp. 8815-8828.
- 5 Bae, B., Ohene-Adjei, S., Kocherginskaya, S., Mackie, R.I., Spies, M.A., Cann, I.K.O. & Nair, S.K. (2008) Molecular basis for the selectivity and specificity of ligand recognition by the family 16 carbohydrate-binding modules from *Thermoanaerobacterium polysaccharolyticum* ManA *J. Biol. Chem.*, 283(18), pp. 12415-12425.
- 6 Waeonukul, R., Pason, P., Kyu, K.L., Sakka, K., Kosugi, A., Mori, Y. & Ratanakhanokchai, K. (2009) Cloning, sequencing, and expression of the gene encoding a multidomain endo-beta-1,4-xylanase from *Paenibacillus curdlanolyticus* B-6, and characterization of the recombinant enzyme *J. Microbiol. Biotechnol.*, 19(3), pp. 277-285.
- 7 Zhao, G., Ali, E., Araki, R., Sakka, M., Kimura, T. & Sakka, K. (2005) Function of the family-9 and family-22 carbohydrate-binding modules in a modular β -1,3-1,4-glucanase/xylanase derived from *Clostridium stercorarium* Xyn10B *Biosci. Biotechnol. Biochem.*, 69(8), pp. 1562-1567.
- 8 Boraston, A.B., McLean, B.W., Chen, G., Li, A., Warren, R.A.J. & Kilburn, D.G. (2002) Co-operative binding of triplicate carbohydrate-binding modules from a thermophilic xylanase *Mol. Microbiol.*, 43, pp. 187-194.
- 9 Blake, A.W., McCartney, L., Flint, J.E., Bolam, D.N., Boraston, A.B., Gilbert, H.J. & Knox, J.P. (2006) Understanding the biological rationale for the diversity of cellulose-directed carbohydrate-binding modules in prokaryotic enzymes *J. Biol. Chem.*, 281(39), pp. 29321-29329.
- 10 Boraston, A.B., Bolam, D.N., Gilbert, H.J. & Davies, G.J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition *Biochem. J.*, 382(3), pp. 769-781.
- 11 Bolam, D.N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M.P., Rixon, J.E., Boraston, A., Hazlewood, G.P. & Gilbert, H.J. (1998) *Pseudomonas* cellulose-binding domains mediate their effects by increasing enzyme substrate proximity *Biochem. J.*, 331(3), pp. 775-781.
- 12 Rodriguez Sanoja, R., Morlon-Guyot, J., Jore, J., Pintado, J., Juge, N. & Guyot, J.P. (2000) Comparative Characterization of Complete and Truncated Forms of *Lactobacillus amylovorus* alpha - Amylase and Role of the C-Terminal Direct Repeats in Raw-Starch Binding *Appl. Environ. Microbiol.*, 66(8), pp. 3350-3356.

- 13 Jörg, K. & Wolfgang, L. (2006) Comparative characterization of deletion derivatives of the modular xylanase XynA of *Thermotoga maritima* *Extremophiles*, 10(5), pp. 373-381.
- 14 Wang, Y., Yuan, H., Wang, J. & Yu, Z. (2009) Truncation of the cellulose binding domain improved thermal stability of endo-beta-1,4-glucanase from *Bacillus subtilis* JA18 *Bioresour. Technol.*, 100(1), pp. 345-349.
- 15 Vaaje-Kolstad, G., Houston, D.R., Riemen, A.H.K., Eijsink, V.G.H. & van Aalten, D.M.F. (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21 *J. Biol. Chem.*, 280(12), pp. 11313-11319.
- 16 Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics *Nucl. Acids Res.*, 37(suppl_1), pp. 233-238.
- 17 Ficko-Blean, E. & Boraston, A.B. (2009) N-acetylglucosamine recognition by a family 32 carbohydrate-binding module from *Clostridium perfringens* NagH *J. Mol. Biol.*, 390(2), pp. 208-220.
- 18 Nagy, T., Simpson, P., Williamson, M.P., Hazlewood, G.P., Gilbert, H.J. & Orosz, L. (1998) All three surface tryptophans in Type IIa cellulose binding domains play a pivotal role in binding both soluble and insoluble ligands *FEBS Letters*, 429(3), pp. 312.
- 19 Abbott, D.W., Eirin-Lopez, J.M. & Boraston, A.B. (2008) Insight into Ligand Diversity and Novel Biological Roles for Family 32 Carbohydrate-Binding Modules *Mol Biol Evol*, 25(1), pp. 155-167.
- 20 Flint, J., Bolam, D.N., Nurizzo, D., Taylor, E.J., Williamson, M.P., Walters, C., Davies, G.J. & Gilbert, H.J. (2005) Probing the mechanism of ligand recognition in family 29 carbohydrate-binding modules *J. Biol. Chem.*, 280(25), pp. 23718-23726.
- 21 Obembe, O., Jacobsen, E., Timmers, J., Gilbert, H., Blake, A., Knox, J., Visser, R. & Vincken, J.-P. (2007) Promiscuous, non-catalytic, tandem carbohydrate-binding modules modulate the cell-wall structure and development of transgenic tobacco (*Nicotiana tabacum*) plants *J. Plant Res.*, 120, pp. 605-617.
- 22 Simon, J.C., Bolam, D.N., Nurizzo, D., Szabó, L., McKie, V.A., Gilbert, H.J. & Gideon, J.D. (2002) Promiscuity in ligand-binding: The three-dimensional structure of a *Piromyces* carbohydrate-binding module, CBM29-2, in complex with cello- and mannohexaose *Proc. Natl. Acad. Sci. U. S. A.*, 99(22), pp. 14077-14082.
- 23 Sorimachi, K., Gal-Coëffet, M.-F.L., Williamson, G., Archer, D.B. & Williamson, M.P. (1997) Solution structure of the granular starch binding domain of *Aspergillus niger* glucoamylase bound to β -cyclodextrin *Structure*, 5(5), pp. 647-661.
- 24 Pantoom, S., Songsiritthigul, C. & Suginta, W. (2008) The effects of the surface-exposed residues on the binding and hydrolytic activities of *Vibrio carchariae* chitinase A *BMC Biochemistry*, 9(1), pp. 2.
- 25 Rodríguez-Sanoja, R., Oviedo, N., Escalante, L., Ruiz, B. & Sánchez, S. (2009) A single residue mutation abolishes attachment of the CBM26 starch-binding domain from *Lactobacillus amylovorus* α -amylase *J. Ind. Microbio. Biotechnol.*, 36(3), pp. 341-346.

- 26 van Bueren, A.L. & Boraston, A.B. (2007) The structural basis of alpha-glucan recognition by a family 41 carbohydrate-binding module from *Thermotoga maritima* *J. Mol. Biol.*, 365(3), pp. 555.
- 27 Boraston, A.B., Nurizzo, D., Notenboom, V., Ducros, V., Rose, D.R., Kilburn, D.G. & Davies, G.J. (2002) Differential oligosaccharide recognition by evolutionarily-related beta-1,4 and beta-1,3 glucan-binding modules *J. Mol. Biol.*, 319(5), pp. 1143-1156.
- 28 Boraston, A.B., Notenboom, V., Warren, R.A.J., Kilburn, D.G., Rose, D.R. & Davies, G. (2003) Structure and ligand binding of carbohydrate-binding module CsCBM6-3 reveals similarities with fucose-specific lectins and "galactose-binding" domains *Journal of Molecular Biology*, 327(3), pp. 659.
- 29 Linder, M., Nevanen, T. & Teeri, T.T. (1999) Design of a pH-dependent cellulose-binding domain *FEBS Letters*, 447, pp. 13.
- 30 Takashima, S., Ohno, M., Hidaka, M., Nakamura, A., Masaki, H. & Uozumi, T. (2007) Correlation between cellulose binding and activity of cellulose-binding domain mutants of *Humicola grisea* cellobiohydrolase 1 *FEBS Lett.*, 581(30), pp. 5891-5896.
- 31 Simpson, P.J., Xie, H., Bolam, D.N., Gilbert, H.J. & Williamson, M.P. (2000) The structural basis for the ligand specificity of family 2 carbohydrate-binding modules *J. Biol. Chem.*, 275(52), pp. 41137-41142.
- 32 Cicortas gunnarsson, L., Montanier, C., Tunnicliffe, R.B., Williamson, M.P., Gilbert, H.J., Nordberg karlsson, E. & Ohlin, M. (2007) Novel xylan-binding properties of an engineered family 4 carbohydrate-binding module *Biochem. J.*, 406(2), pp. 209-214.
- 33 Ahn, J.O., Choi, E.S., Lee, H.W., Hwang, S.H., Kim, C.S., Jang, H.W., Haam, S.J. & Jung, J.K. (2004) Enhanced secretion of *Bacillus stearothermophilus* L1 lipase in *Saccharomyces cerevisiae* by translational fusion to cellulose-binding domain *Appl. Environ. Microbiol.*, 64(6), pp. 833.
- 34 Koichiro, M., Akihiko, K. & Roy, H.D. (2003) Solubilization of cellulosomal cellulases by fusion with cellulose-binding domain of noncellulosomal cellulase engD from *Clostridium cellulovorans* *Proteins Struct. Funct. Genet.*, 50(4), pp. 620-628.
- 35 Terpe, K. (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems *Appl. Microbiol. Biotechnol.*, 60(5), pp. 523.
- 36 Craig, S.J., Foong, F.C. & Nordon, R. (2006) Engineered proteins containing the cohesin and dockerin domains from *Clostridium thermocellum* provides a reversible, high affinity interaction for biotechnology applications *J. Biotechnol.*, 121(2), pp. 165.
- 37 Boraston, A.B., McLean, B.W., Guarna, M.M., Amandaron-Akow, E. & Kilburn, D.G. (2001) A Family 2a Carbohydrate-Binding Module Suitable as an Affinity Tag for Proteins Produced in *Pichia pastoris* *Protein Expr. Purif.*, 21, pp. 417.
- 38 Kavoosi, M., Meijer, J., Kwan, E., Creagh, A.L., Kilburn, D.G. & Haynes, C.A. (2004) Inexpensive one-step purification of polypeptides expressed in *Escherichia coli* as fusions with the family 9 carbohydrate-binding module of xylanase 10A from *T. maritima* *J. Chromatogr. B*, 807(1), pp. 87.
- 39 Xu, Z., Bae, W., Mulchandani, A., Mehra, R.K. & Chen, W. (2002) Heavy Metal Removal by Novel CBD-EC20 Sorbents Immobilized on Cellulose *Biomacromolecules*, 3(3), pp. 462.

- 40 Bernard, M.P., Cao, D., Myers, R.V. & Moyle, W.R. (2004) Tight attachment of chitin-binding-domain-tagged proteins to surfaces coated with acetylated chitosan *Anal. Biochem.*, 327(2), pp. 278.
- 41 Chern, J.-T. & Chao, Y.-P. (2005) Chitin-binding domain based immobilization of d-hydantoinase J. *Biotechnol.*, 117(3), pp. 267.
- 42 Sharma, S.S., Chong, S. & Harcum, S.W. (2006) Intein-mediated protein purification of fusion proteins expressed under high-cell density conditions in *E. coli* J. *Biotechnol.*, 125(1), pp. 48.
- 43 Lin, S.-C., Lin, I.P., Chou, W.-I., Hsieh, C.-A., Liu, S.-H., Huang, R.-Y., Sheu, C.-C. & Chang, M.D.-T. (2009) CBM21 starch-binding domain: A new purification tag for recombinant protein engineering *Protein Expr. Purif.*, 65(2), pp. 261.
- 44 Hwang, S., Ahn, J., Lee, S., Lee, T.G., Haam, S., Lee, K., Ahn, I.-S. & Jung, J.-K. (2004) Evaluation of cellulose-binding domain fused to a lipase for the lipase immobilization *Biotechnol. Letters*, 26(7), pp. 603.
- 45 Kauffmann, C., Shoseyov, O., Shpigel, E., Bayer, E.A., Lamed, R., Shoham, Y. & Mandelbaum, R.T. (2000) Novel methodology for enzymatic removal of atrazine from water by CBD-fusion protein immobilized on cellulose *Environ. Sci. Technol.*, 34(7), pp. 1292.
- 46 Etai, S., Arie, G., Gilat, E., Amos, A., Adi, E., Mara, D. & Oded, S. (1999) Immobilization of recombinant heparinase I fused to cellulose-binding domain *Biotechnol. and Bioeng.*, 65(1), pp. 17-23.
- 47 Volkov, I.Y., Lunina, N.A. & Velikodvorskaya, G.A. (2004) Prospects for the practical application of substrate-binding modules of glycosyl hydrolases *Appl. Biochem. and Microbiol.*, 40(5), pp. 427.
- 48 Aijun, A.W., Ashok, M. & Wilfred, C. (2001) Whole-cell immobilization using cell surface-exposed cellulose-binding domain *Biotechnol. Prog.*, 17(3), pp. 407-411.
- 49 Wang, J.-Y. & Chao, Y.-P. (2006) Immobilization of cells with surface-displayed chitin-binding domain *Appl. Environ. Microbiol.*, 72(1), pp. 927-931.
- 50 Andrade, F.K., Moreira, S.M.G., Domingues, L. & Gama, F.M.P. (2009) Improving the affinity of fibroblasts for bacterial cellulose using carbohydrate-binding modules fused to RGD *Journal of Biomedical Materials Research Part A*.
- 51 Wang, A.A., Mulchandani, A. & Chen, W. (2002) Specific adhesion to cellulose and hydrolysis of organophosphate nerve agents by a genetically engineered *Escherichia coli* strain with a surface-expressed cellulose-binding domain and organophosphorus hydrolase *Appl. Environ. Microbiol.*, 68(4), pp. 1684-1689.
- 52 Umemoto, Y. & Araki, T. (2009) Cell wall regeneration in *Bangia atropurpurea* (Rhodophyta) protoplasts observed using a mannan-specific carbohydrate-binding module *Mar. Biotechnol.*
- 53 Stephanie, E.P., Bryon, S.D., Kyle, E.B., Qi, X., Shi-You, D., Todd, B.V., Charles, A.A. & Michael, E.H. (2007) Microscopic analysis of corn fiber using starch- and cellulose-specific molecular probes *Biotechnol. and Bioeng.*, 98(1), pp. 123-131.
- 54 Linder, M., Winiecka-Krusnell, J. & Linder, E. (2002) Use of recombinant cellulose-binding domains of *Trichoderma reesei* cellulase as a selective immunocytochemical marker for cellulose in protozoa *Appl. Environ. Microbiol.*, 68(5), pp. 2503-2508.

- 55 Hussack, G., Luo, Y., Veldhuis, L., Hall, J.C., Tanha, J. & MacKenzie, R. (2009) Multivalent anchoring and oriented display of single-domain antibodies on cellulose *Sensors*, 9(7), pp. 5351-5367.
- 56 Shpigel, E., Goldlust, A., Eshel, A., Ber, I.K., Efroni, G., Singer, Y., Levy, I., Dekel, M. & Shoseyov, O. (2000) Expression, purification and applications of staphylococcal protein A fused to cellulose-binding domain *Biotechnol. Appl. Biochem.*, 31 (Pt 3), pp. 197-203.
- 57 Doheny, J.G., Jervis, E.J., Guarna, M.M., Humphries, R.K., Warren, R.A. & Kilburn, D.G. (1999) Cellulose as an inert matrix for presenting cytokines to target cells: production and properties of a stem cell factor-cellulose-binding domain fusion protein *Biochem. J.*, 339(2), pp. 429-434.
- 58 Raha, A.R., Varma, N.R.S., Yusoff, K., Ross, E. & Foo, H.L. (2005) Cell surface display system for *Lactococcus lactis*: a novel development for oral vaccine *Appl. Microbiol. Biotechnol.*, 68(1), pp. 75.
- 59 Maurice, S., Dekel, M., Shoseyov, O. & Gertler, A. (2003) Cellulose beads bound to cellulose binding domain-fused recombinant proteins; an adjuvant system for parenteral vaccination of fish *Vaccine*, 21(23), pp. 3200.
- 60 Guerreiro, C.I.P.D., Fontes, C.M.G.A., Gama, M. & Domingues, L. (2008) *Escherichia coli* expression and purification of four antimicrobial peptides fused to a family 3 carbohydrate-binding module (CBM) from *Clostridium thermocellum* *Protein Expr Purif*, 59(1), pp. 161.
- 61 Rodríguez-Sanoja, R., Oviedo, N. & Sanchez, S. (2005) Microbial starch binding domains *Curr. Opin. Microbiol.*, 8(3), pp. 260-267.
- 62 Chou, W.-I., Pai, T.-W., Liu, S.-H., Hsiung, B.-K. & Chang, M.D.T. (2006) The family 21 carbohydrate-binding module of glucoamylase from *Rhizopus oryzae* consists of two sites playing distinct roles in ligand binding *Biochem J*, 396(3), pp. 469-477.
- 63 Tung, J.-Y., Chang, M.D.-T., Chou, W.-I., Liu, Y.-Y., Yeh, Y.-H., Chang, F.-Y., Lin, S.-C., Qiu, Z.-L. & Sun, Y.-J. (2008) Crystal structures of the starch-binding domain from *Rhizopus oryzae* glucoamylase reveal a polysaccharide-binding path *Biochem J*, 416(1), pp. 27-36.
- 64 George, R.A. & Heringa, J. (2002) An analysis of protein domain linkers: their classification and role in protein folding *Protein Eng.*, 15(11), pp. 871-879.
- 65 Mojgan, K., Creagh, A.L., Douglas, G.K. & Charles, A.H. (2007) Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in *Escherichia coli* *Biotechnol. and Bioeng.*, 98(3), pp. 599-610.
- 66 Boraston, A.B., McLean, B.W., Guarna, M.M., Amandaron-Akow, E. & Kilburn, D.G. (2001) A family 2a carbohydrate-binding module suitable as an affinity tag for proteins produced in *Pichia pastoris* *Protein Expr Purif*, 21(3), pp. 417.
- 67 Levy, I., Paldi, T. & Shoseyov, O. (2004) Engineering a bifunctional starch-cellulose cross-bridge protein *Biomaterials*, 25(10), pp. 1841.
- 68 Van den Bergh, K.B., Rougé, P., Proost, P., Coosemans, J., Krouglova, T., Engelborghs, Y., Peumans, W. & Damme, E.M. (2004) Synergistic antifungal activity of two chitin-binding proteins from spindle tree (*Euonymus europaeus* L.) *Planta*, 219(2), pp. 221.

69 Fuglsang, C.C. & Tsuchiya, R. July 2001 Cellulose binding domains (CBDs) for oral care products.
U.S. patent 6,264,925