

# **Análisis *in silico* de proteínas involucradas en la formación del biofilm en bacterias enteropatógenicas**

Viloria-Álvarez J.C.<sup>1</sup>, Barrera-Chávez I.D.<sup>1</sup>, García-Andrade A.<sup>1</sup>, García-Infante S.I.<sup>1</sup>, López-Briones S.<sup>1</sup>, Hernández-Luna M.A.<sup>1</sup>

1. Departamento de Medicina y Nutrición, División Ciencias de la Salud. Universidad de Guanajuato, Campus León

## **Resumen**

El biofilm es un agregado de bacterias que se encuentran sumergidas en una matriz polimérica de polisacáridos, proteínas, lípidos y ADN extracelular que las protege de la respuesta inmune del huésped y los antibióticos. El biofilm ha tomado relevancia en el ámbito hospitalario y la industria de alimentos, y en los últimos años también se ha asociado al desarrollo de neoplasias del tracto gastrointestinal. Por lo que el estudio de las proteínas implicadas en la formación del biofilm resulta importante para el desarrollo de estrategias para la prevención, diagnóstico y tratamiento de infecciones bacterianas. En este trabajo hicimos un análisis *in silico* de las proteínas del biofilm de bacterias enteropatógenicas asociadas al cáncer gastrointestinal como *F. nucleatum*, *B. fragilis*, *E. coli* y *Salmonella enterica*, con la finalidad de buscar regiones conservadas entre este tipo de proteínas que en algún momento permitan predecir la interacción que podría existir entre ellas. Para el análisis bioinformático se realizó el alineamiento de las secuencias de aminoácidos y el modelamiento de la estructura terciaria de las proteínas. Nuestros resultados muestran regiones conservadas entre las proteínas del biofilm de las bacterias enteropatógenicas estudiadas, así mismo, se realizó una propuesta de la posible interacción entre las proteínas EnvC, BT3147 y pefC de *Fusobacterium nucleatum*, *Bacteroides fragilis* y *Salmonella enterica* respectivamente. Este tipo de estudios podría resultar en el desarrollo de métodos para la detección, inhibición y eliminación del biofilm.

**Palabras clave:** Biofilm; bacterias enteropatógenicas; análisis bioinformático.

## **Abstract**

Biofilm is an aggregate of bacteria that are immersed in a polymeric matrix of polysaccharides, proteins, lipids and extracellular DNA that protects them from the host immune response and antibiotics. Biofilm has become relevant at hospital environmental and food industry, and in recent years it has also been associated with the development of neoplasms of the gastrointestinal tract. Therefore, the study of the proteins involved in biofilm formation is important for the development of strategies for prevention, diagnosis and treatment of bacterial infections. In this work, we performed an *in-silico* analysis of the biofilm proteins of enteropathogenic bacteria associated with gastrointestinal cancer such as *F. nucleatum*, *B. fragilis*, *E. coli* and *Salmonella enterica*, with the aim of searching for conserved domains between these proteins that at some point would allow predicting the interaction that could exist between them. For the bioinformatic analysis, the alignment of amino acid sequences and the modeling of the tertiary structure of the proteins were performed. Our results show conserved domains among the biofilm proteins of the enteropathogenic bacteria studied, likewise, a proposal of the possible interaction between EnvC, BT3147 and pefC proteins of *Fusobacterium nucleatum*, *Bacteroides fragilis* and *Salmonella enterica*, respectively, was made. This kind of studies could result in the development of new methods for the detection, inhibition and elimination of biofilm.

**Keywords:** Biofilm; protein; bioinformatic analysis.

## **Introducción**

El biofilm se ha definido como “un agregado de bacterias” donde están sumergidas por una matriz de sustancias extracelulares poliméricas (SEP) compuesta principalmente por polisacáridos, proteínas, lípidos y ADN extracelular (eADN) (Flemming et al., 2016). La formación del biofilm depende, la mayoría de las veces de estímulos percibidos como negativos del ambiente, como, el pH la temperatura o la disponibilidad de nutrientes. Las bacterias se adhieren a estructuras bióticas o abióticas, produciendo el SEP, permitiéndoles protegerse de los cambios en el ambiente, incluyendo la respuesta inmune del huésped y los antibióticos (di Domenico et al., 2017).

Los efectos del biofilm bacteriano han tomado relevancia en el ambiente hospitalario y la industria de alimentos principalmente, y han creado la urgencia de identificar y caracterizar esta estructura, con el objetivo de

desarrollar nuevos métodos para la prevención, diagnóstico y el tratamiento de aquellas infecciones ocasionadas por bacterias que tienen la maquinaria para formar el biofilm (Flemming et al., 2016).

Adicionalmente, algunas bacterias formadoras de biofilm también se han estudiado por su posible papel en el desarrollo y progresión de neoplasias del tracto gastrointestinal, en específico, hablamos de neoplasias de colon, recto y vesícula biliar. Se ha descrito el rol de bacterias como *Fusobacterium nucleatum*, *Escherichia coli*, *Bacteroides fragilis* y *Salmonella enterica* como posibles factores de riesgo para el desarrollo de neoplasias gastrointestinales, a través de diferentes mecanismos, como, la disbiosis intestinal, la inflamación crónica o provocando cambios en la respuesta inmune del huésped infectado. (Hernández et al., 2019)

*Fusobacterium nucleatum* (*F. nucleatum*) es una bacteria anaerobia Gram negativa adherente e invasiva que reside principalmente en la cavidad oral y generalmente se asocia con enfermedad periodontal, ya que su presencia es fundamental en la formación de la placa dental. Aunque el principal sitio de residencia de *F. nucleatum* es la cavidad oral, esta bacteria se ha encontrado en otros sitios del organismo como el tubo digestivo, así como en biopsias de lesiones primarias y heces de pacientes con cáncer de colon, por lo que se ha relacionado con el desarrollo y progresión de esta neoplasia. *F. nucleatum* también se ha relacionado con complicaciones obstétricas como el parto prematuro, la sepsis neonatal y la muerte prematura. (Hernández et al., 2019). Se han identificado factores que promueven la interacción y coagregación de *F. nucleatum* con otras bacterias, la mayoría de esos factores son adhesinas, de las cuales algunas están directamente implicadas con la formación del biofilm (Wu, C et al., 2018).

*Bacteroides fragilis* (*B. fragilis*) es una bacteria anaerobia obligada que constituye aproximadamente del 1 al 2% de los microorganismos normales en el intestino de las personas sanas (Shin et al., 2020). Esta bacteria desempeña un rol importante en el desarrollo del sistema inmune asociado a mucosas y en la homeostasis intestinal. Sin embargo, es un patógeno oportunista que está relacionado con el desarrollo de la enfermedad intestinal inflamatoria (IBD). Se han descrito dos cepas de *B. fragilis*: *B. fragilis* no toxigénica (NTBF) y *B. fragilis* enterotoxigénica (ETBF). La cepa enterotoxigénica produce una toxina llamada fragilisina, que está asociada con la enfermedad diarreica aguda, la enfermedad de Crohn y el cáncer colorrectal (Kim et al., 2019). Las cepas de *B. fragilis* forman biofilms en la membrana mucosa del colon. La detección del biofilm puede predecir un riesgo incrementado de desarrollo de cáncer y de IBD (Chew et al., 2020).

*Escherichia coli* (*E. coli*) es una bacteria Gram negativa, anaerobia facultativa, en forma de bastón y no esporulante. Ampliamente estudiada debido a su alta presencia en nuestro organismo, también debido a la facilidad con la que se puede cultivar en el laboratorio. La mayoría de las especies de *E. coli* se encuentran en el intestino de organismos de sangre caliente (Sharma et al., 2016). *E. coli* es la bacteria anaerobia facultativa comensal más abundante del microbiota de nuestro tracto gastrointestinal y junto con otros microorganismos es esencial para el funcionamiento correcto del proceso digestivo. entonces podemos considerarla como una bacteria “buena” ya que la mayoría de las especies no son patógenas y ayudan a prevenir la colonización de bacterias que si lo son. Sin embargo, existen especies de *E. coli* que son los causantes de diversas complicaciones, como es el caso de infecciones asociadas a dispositivos médicos, como catéteres uretrales e intravasculares, articulaciones y derivaciones e injertos protésicos. Las cepas asociadas a las infecciones gastrointestinales son denominadas *E. coli* diarreogénicas o intestinales, mientras que las asociadas a infecciones en otros aparatos y sistemas, como, por ejemplo, el tracto urinario, sanguíneo y nervioso, son denominadas *E. coli* patogénicas extraintestinales (ExPEC). Debido a que este tipo de cepas forma biofilm sus infecciones resultan difíciles de tratar farmacológicamente.

*Salmonella enterica* (*S. enterica*) es una bacteria, bacilo Gram negativo, flagelado aeróbico. Causante de la fiebre enterica o fiebre tifoidea, caracterizada por una fiebre alta acompañada de síntomas como malestar general, pérdida del apetito y malestar general. La infección por *Salmonella enterica* suele darse después del consumo de alimentos o agua contaminada y puede producir una infección sistémica después de atravesar la superficie mucosa del intestino. Alrededor de 4% de las personas infectadas desarrollan una infección crónica asintomática, situada en los litos de la vesícula biliar a través del biofilm que se forma en su superficie, esto lleva a una inflamación crónica asociada con el carcinoma de vesícula biliar (di Domenico et al., 2017). Sus serotipos *Typhi*, *Typhimurium* y *enteritidis* son los organismos estudiados en la búsqueda de las proteínas que forman parte de la estructura del biofilm en diferentes materiales y bajo diversas condiciones, como los litos de colesterol de la vesícula biliar. Estudios previos han mostrado que las mutaciones en genes involucrados en la motilidad flagelar, el metabolismo del ácido colánico, las fimbrias, los lipopolisacáridos y la celulosa afectan la formación del biofilm (Solano et al., 2002).

En el presente trabajo realizamos un análisis bioinformático para comparar las proteínas que se han asociado a la formación del biofilm de *Fusobacterium nucleatum*, *Escherichia coli*, *Bacteroides fragilis enterotoxigénica* y *Salmonella enterica* con el objetivo de saber si las proteínas de estas bacterias podrían en algún momento interactuar entre ellas y formar en conjunto el biofilm, el cual podría participar en el desarrollo progresión del cáncer de colon.

## Material y Métodos

### Búsqueda de secuencias de proteínas

Se realizó la búsqueda de publicaciones en revistas indexadas en JCR utilizando como palabras clave el nombre de cada bacteria, por ejemplo, "*Fusobacterium nucleatum*" y los términos "biofilm" y "protein" en PubMed. Se seleccionaron las publicaciones de acuerdo con los siguientes parámetros: 1) Proteínas relacionadas directamente con la formación del biofilm, 2) Que los genes que codifican para las proteínas asociadas al biofilm estén registradas GenBank y 3) Que las proteínas tengan descrita su secuencia de aminoácidos mediante el código FASTA.

Las proteínas seleccionadas fueron las siguientes A) FtsX, EnvC, RadD, Fap2, CmpA de *Fusobacterium nucleatum* B) ompA, BT3148, BT3147, bmeB3 de *Bacteroides fragilis*, C) csgD, hha, bcsA operon, pgaC, fimB de *Escherichia coli* y D) fimH, wcaM, WCAA, WZA, pefC, ipfE, BapA, WCAL, wzxE, UDP-f, yhiO de *Salmonella enterica*.

### Alineamiento de proteínas

El alineamiento de las proteínas se hizo con la herramienta bioinformática Multiple Sequence Alignment de Clustal Omega, utilizando los códigos FASTA de aminoácidos de cada una de las proteínas seleccionadas para las cuatro bacterias. El alineamiento se hizo entre las proteínas del biofilm producidas por la misma especie bacteriana y posteriormente se hizo el alineamiento entre las proteínas con características similares de las cuatro especies bacterianas.

### Modelamiento de la estructura de las proteínas

Se realizó el modelo tridimensional de la estructura de cada una de las proteínas utilizando la base de datos de RCSB PDB (Protein Data Bank), seleccionando los modelos de aquellas proteínas que estaban previamente reportadas. Para las proteínas que la estructura tridimensional no se ha reportado se utilizó la herramienta SWISS-MODEL utilizando el código FASTA de las proteínas.

## Resultados y Discusión

### Análisis de publicaciones

Después de realizar la búsqueda de publicaciones en PubMed con las palabras clave, donde el tópico de la primera búsqueda fueron las proteínas implicadas en la formación de biofilm de las bacterias *Fusobacterium nucleatum*, *Escherichia coli*, *Salmonella enterica* y *Bacteroides fragilis*, se analizaron un total de 24 Artículos, en los cuales se identificaron las proteínas que se muestran en la Tabla 1.

**Tabla 1.** Códigos de GenBank proteínas encontradas implicadas en el biofilm de *Fusobacterium nucleatum*, *Bacteroides fragilis*, *Escherichia coli* y *Salmonella enterica*.

	Proteína	GenBank
<i>Fusobacterium nucleatum</i>	FtsX	WP_005903510.1
	EnvC	EFG94960.1
	RadD (Fn1526)	AL93652.1
	Fap2 (FN1449)	AAL95642.1
	CmpA (Fn1554)	AAL93680.1
<i>Bacteroides fragilis</i>	OmpA	AY028459.1
	BT3148	AAO78254.1
	BT3147	AAO78253.1
	AcrB (bmeB3)	AUI48642.1
	csgD	CAD6011295.1

<i>Escherichia coli</i>	hha	CAD6015104.1
	bcsA operon	BAE77761.1
	pgaC	CAD6017156.1
	fimB	QJ11330.1
<i>Salmonella enterica</i>	fimH	NP_459542.1
	WcaM	WP_001111841.1
	WCAA	WP_000205695.1
	WZA	WP_000978077.1
	pefC	WP_000007893.1
	lpfABCDE	AAA73970.1
	BapA,	ACS91539.1
	WCAL	NP_461045.1
	wzxE	NP_462816.1
	UDP-f	CAC86199.1
yhiO	NP_462897.1	

## Alineamiento y modelamiento de proteínas

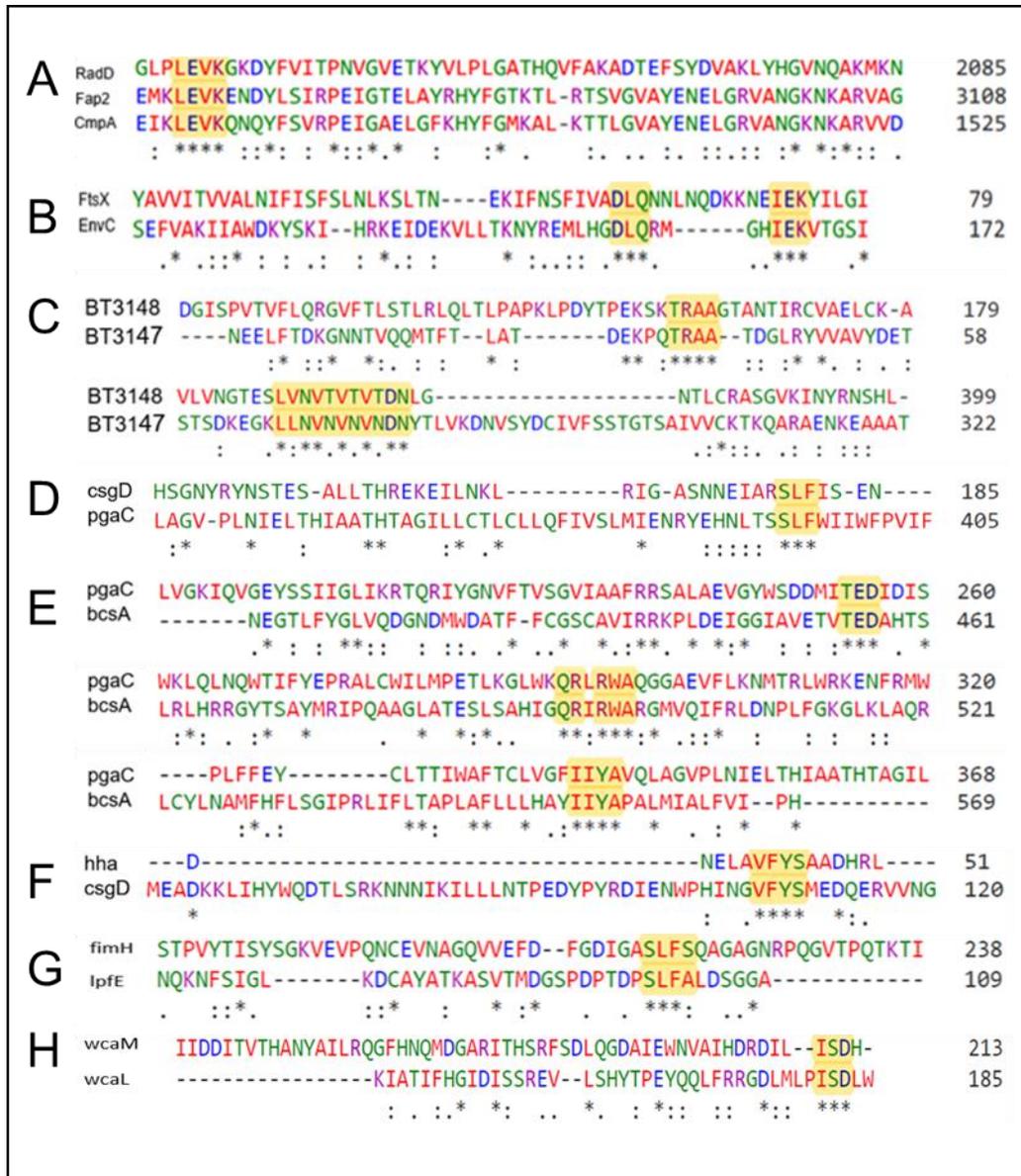
Una vez identificados las proteínas del biofilm de las cuatro bacterias enteropatógenas, se realizó el alineamiento de las secuencias utilizando los códigos FASTA. Primero se hizo una comparación entre las secuencias de las proteínas del biofilm de cada bacteria basados en el tipo de proteína (proteínas de superficie, proteínas transportadoras, proteínas reguladoras, etc), la alineación se realizó a través de la plataforma en línea Clustal Omega. Los resultados observados se describen a continuación.

Para el alineamiento de las proteínas de *Fusobacterium nucleatum*, los resultados más significativos fueron, el alineamiento de las proteínas de membrana externa RadD, Fap2 y CmpA (figura 1A), que mostró 7 regiones conservadas, además de múltiples sitios donde había concordancia y/o similitud entre los aminoácidos de las 3 proteínas. El alineamiento de la proteína transportadora ABC "FtsX" y la proteína peptidasa M23 "EnvC" de *F. nucleatum* (figura 1B), mostró 3 regiones conservadas, además de múltiples sitios donde había concordancia y/o similitud entre los aminoácidos de las 2 proteínas.

Para el alineamiento de las proteínas de *Bacteroides fragilis*, los resultados más significativos se encontraron en el alineamiento de las proteínas hipotéticas BT3148 y BT3147, que (figura 1C) mostró 4 regiones conservadas, además de múltiples sitios donde había concordancia y/o similitud entre los aminoácidos de las 2 proteínas.

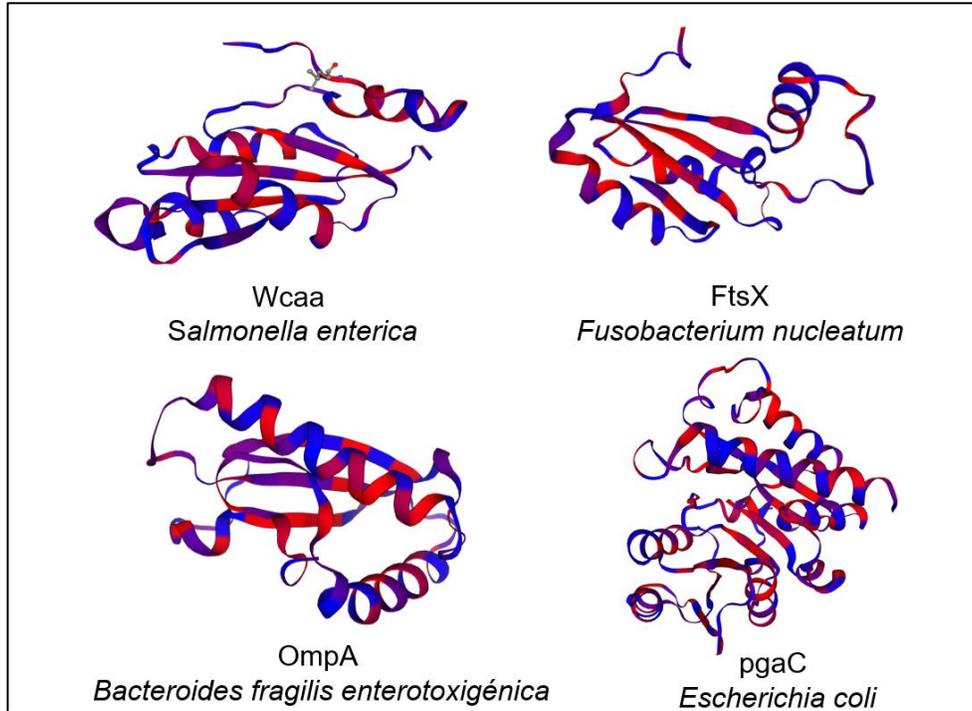
Para el alineamiento de las proteínas de *Escherichia coli*, los resultados más significativos fueron, el alineamiento de las proteínas csgD y pgaC (figura 1D), ambas proteínas de tipo sintasa, mostraron 3 regiones conservadas. El alineamiento de las proteínas bcsA operón y pgaC de *E. coli* (figura 1E) mostró 4 regiones conservadas, y el alineamiento de las proteínas hha y csgD de *E. coli*, ambos reguladores transcripcionales, mostró 3 regiones conservadas (figura 1F).

Para el alineamiento de las proteínas de la bacteria de *Salmonella enterica*, los resultados más significativos fueron, el alineamiento de FimH y de lpfE (figura 1G), que mostró 3 regiones conservadas con tres aminoácidos cada una, además de algunas zonas con similitudes entre los aminoácidos. El alineamiento de las proteínas relacionadas con la síntesis del ácido colánico wcaA, wcaL y wcaM resultó en cuatro diferentes tipos de combinaciones, pero solo se muestra el alineamiento de wcaL y wcaM (figura 1H) que mostró más de 3 regiones conservadas. El resto solo mostró zonas similares entre ellos.

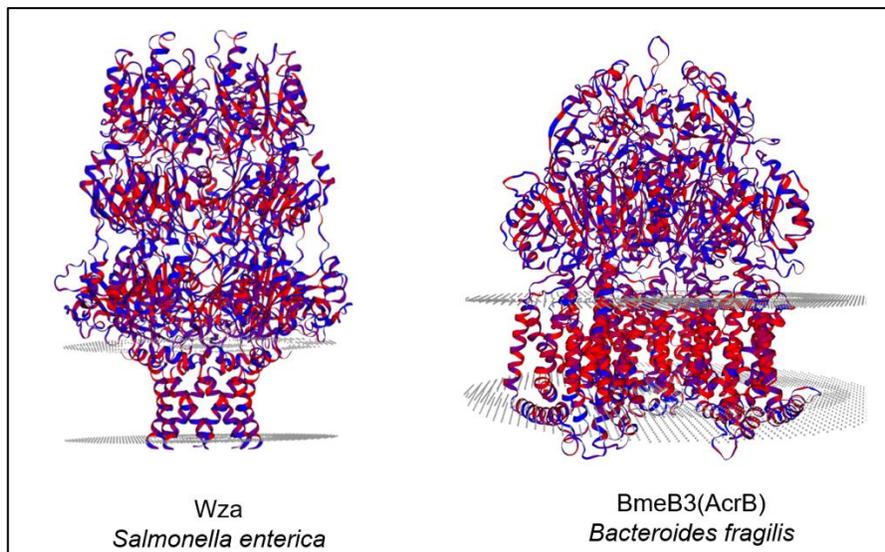


**Figura 1.** Alineamiento de las proteínas del biofilm de las bacterias enteropatógenas *Fusobacterium nucleatum*, *Escherichia coli*, *Bacteroides fragilis* y *Salmonella enterica*. A) RadD, Fap2 y CmpA de *F. nucleatum*, B) FtsX y EnvC de *F. nucleatum*, C) BT3148 y BT3147 de *B. Fragilis*, D) CsgD y PgaC de *E. coli*, E) bcsA operon y pgaC de *E. coli*, F) hha y csgD de *E. coli*, G) FimH y lpfE de *Salmonella enterica* y H) wcaL y wcaM de *Salmonella enterica*.

Posteriormente, se realizó el análisis de las estructuras secundarias de las proteínas del biofilm para determinar posibles similitudes entre ellas. Para ello se buscó en la base de datos Protein Data Bank (PDB) las estructuras de las proteínas del biofilm de las cuatro bacterias enteropatógenas, sin embargo, solo se encontraron 3 proteínas pertenecientes a *E. coli*. Por este motivo, se realizó un modelaje a través de la plataforma en línea Swiss-Model, esto con el fin de crear las posibles estructuras de las proteínas no reportadas en PDB. Nuestros resultados mostraron que las proteínas Wcaa, FtsX, OmpA y pgaC presentan similitud en sus estructuras secundarias al estar conformadas por  $\alpha$ -hélices y hojas plegadas (figura 2), así como Wza y BmeB3, ambos transportadores de membrana (figura 3).



**Figura 2.** Estructuras secundarias de las proteínas del biofilm de *S. enterica*, *F. nucleatum*, *B. fragilis* y *E. coli*.



**Figura 3.** Transportadores transmembrana de *S. enterica* y *B. fragilis*.

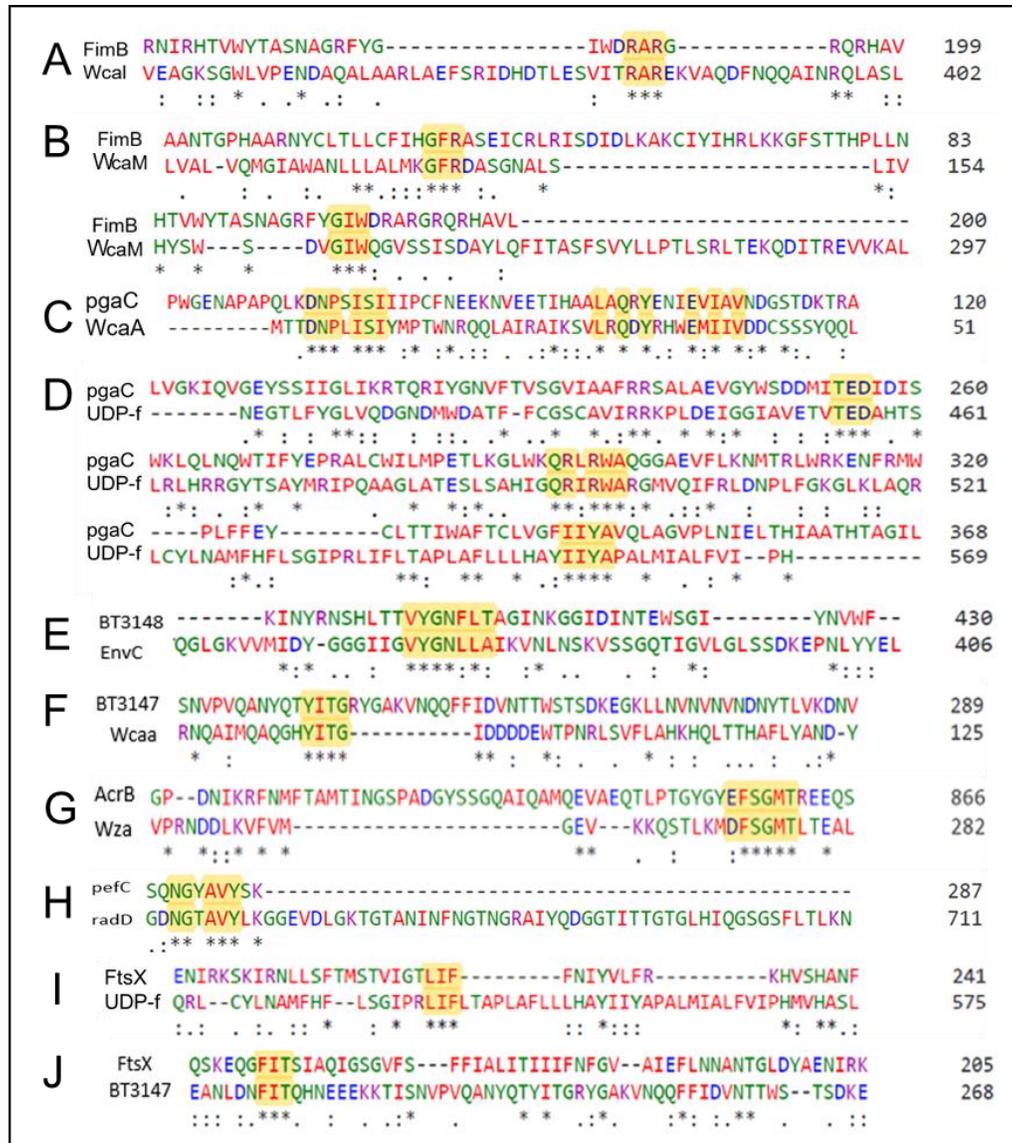
Una vez realizada la comparación entre las estructuras secundarias de las proteínas del biofilm de las cuatro bacterias enteropatógenas, se procedió a la realización del alineamiento de secuencias entre las proteínas cuya estructura secundaria fuera parecida, esto con la finalidad de identificar regiones conservadas entre ellas. Los resultados de esta práctica se muestran a continuación.

De manera interesante, las proteínas bcsA de *E. coli* y UDP-f de *Salmonella enterica* mostraron un alto grado de conservación (figura 4), lo que nos dice que ambas bacterias, al ser de la misma familia, tienen proteínas en común.

60	bscA	MSILTRHLLIPPVNRILGRVDRYRHGASAFSATLGCFWMLAWIFIPLEHPRMQRIRA	60	bscA	DGHDMDATFFCGSCAVIRRRKPLDEIGGIAVETVEDAHTSLRLHRRGYTSAYMRRIPQAA	480
60	UDP-f	MSALSRLILIPPVSRILSERVQGYRRHGASPFSAALGCWTLAKIVFPLEHPRMQRIRD	60	UDP-f	DGHDMDATFFCGSCAVIRRRKPLDEIGGIAVETVEDAHTSLRLHRRGYTSAYMRRIPQAA	480
120	bscA	EHKMLYPHINASRPRPLDPVRYLIQTWLLIGASRKETPKPRRRAFSGLQNIIRGRVYHQIM	120	bscA	GLATESLSAHIGQIRIRWARGNVQIFRLDNPLTGKGLKFAQRLCYVNIAMFHLGSGIPRLIF	540
120	UDP-f	GHKALYPHINAARPRPLDPARYLIQTLMLVMSSTKERHEPRWRFSARLKDVRGRVYHQIM	120	UDP-f	GLATESLSAHIGQIRIRWARGNVQIFRLDNPLFGKGLKLAQRLCYLNIAMFHLGSGIPRLIF	540
180	bscA	NELPERVSHKTHLDEKKEKGLHSAGARRLILGIIVFSLILALICVTQPFNPLAQFIFL	180	bscA	LTAPLAFLLHAYIYAPALMIALFVLPHPHISLTHSKIQGKYRHSFWSIEYTVLAMY	600
180	UDP-f	DTLPERVRQKTTLHLEKEKGLHSAGARRLILGIIVFSLILALICVTQPFNPLSQFIFL	180	UDP-f	LTAPLAFLLHAYIYAPALMIALFVLPHPHISLTHSKIQGKYRHSFWSIEYTVLAMY	600
240	bscA	MLLWGVALLVRRMPGRFSALMLIVLSLTVSCRYIMWRYTSTLINDDPVSLVCGILLFAE	240	bscA	IAPPTLVALINPHKGFNVTAKGGLVEEYVDWISRPIYFLVLLNLVGAAGVIRYFYG	660
240	UDP-f	LLLWGVALLVRRMPGRFSALMLIVLSLTVSCRYIMWRYTSTLINDDPVSLVCGILLFAE	240	UDP-f	IAPPTLVALINPHKGFNVTAKGGLVEEYVDWISRPIYFLVLLNLVGAAGVIRYFYG	660
300	bscA	TYANIVLVLYGFQVWVPLNRQPVLPKDMSLWSPVDFVPTYIEDLNWKNITIASLGID	300	bscA	PPTENLTVVSMWVFNILVLLGGAVAVSVEKQVRRHRVETMPAARAREDEGLFSCT	720
300	UDP-f	TYANIVLVLYGFQVWVPLNRQPVLPKEMSQWPTVDIFVPTYIEDLNWKNITIASLGID	300	UDP-f	PENETLTVIVSLVWVFNILVLLGGAVAVSVEKQVRRHRVETMPAARAREDEGLFSCT	720
360	bscA	WPKDKLNLIWIDDGGRREEFRQAVQVGVKYIARTTHEHAKAGININNAKHKAGGEFVAFID	360	bscA	VQDFSDGGLGKINGQAQILEGQVNLKRGQQEYVFPQVAVMGNVGLKMLPLTTQ	780
360	UDP-f	WPKDKLNLIWIDDGGRREEFRHFAHVGHYIARTTHEHAKAGININNAKHKAGGEFVAIFD	360	UDP-f	VHDFSDGGLGKINGQAQILEGQVNLKRGQQEYVFPQVAVMGNVGLKMLPLTTK	780
420	bscA	CDHVPTRSFLQMTMGWFLKEKQLAWWQTPHHFFSPDPFERNLGRFRKTPNEGTLFYGLVQ	420	bscA	QHIDFVQCTFARADTWALWQDSYPEDKPLESLLDILKLGFRGYRHLAEFAPSSVKGFVRV	840
420	UDP-f	CDHVPTRSFLQMTMGWFLKEKQLAWWQTPHHFFSPDPFERNLGRFRKTPNEGTLFYGLVQ	420	UDP-f	QHIDFVQCTFARADTWALWQDSYPEDKPLESLLDILKLGFRGYRHLAEFAPSSVKVIFRS	840
	bscA			bscA	LTSLSWVVSFIPRRPERSETAQPSSDQALAAQ--	872
	UDP-f			UDP-f	LTALIAIVVSFIPRRPERQAATQPSSDRVMAQAQQ	874

Figura 4. Alineamiento entre las proteínas bscA y UDP-f de *Escherichia coli* y *S. enterica* respectivamente.

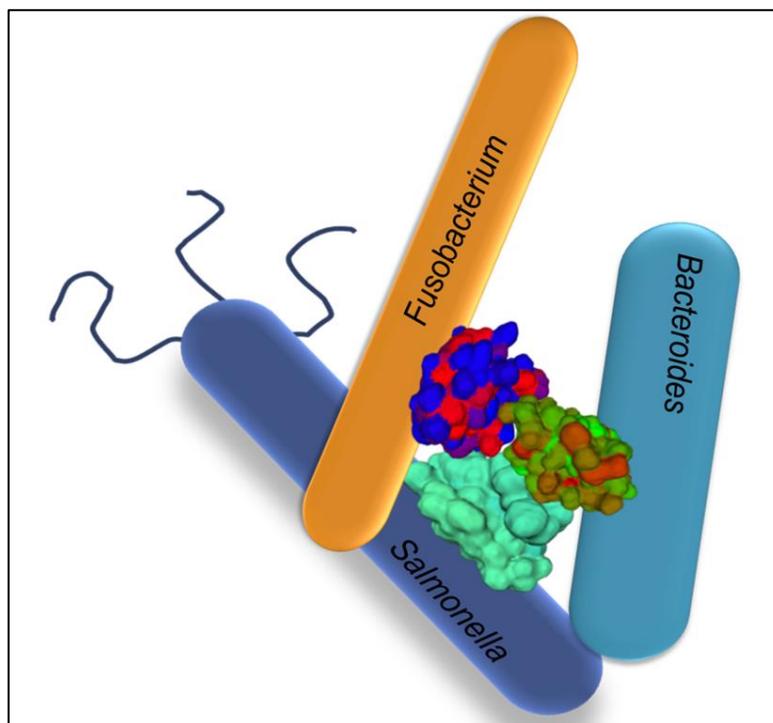
Por otra parte, las proteínas del biofilm de las bacterias *F. nucleatum*, *E. coli*, *B. fragilis* y *S. enterica* que mostraron regiones conservadas fueron las siguientes: FimB de *E. coli* y WcaI de *S. enterica*, mostrando 3 regiones conservadas (figura 5A), FimB de *E. coli* y wcaM de *S. enterica* (figura 5B), mostraron un par de 3 regiones conservadas consecutivas y un par de 3 regiones conservadas separadas por un aminoácido entre cada uno. El alineamiento de secuencia de aminoácidos entre pgaC de *E. coli* y WcaA de *S. enterica* (figura 5C), mostrando un par de regiones de aminoácidos conservadas de manera consecutiva separadas únicamente por un aminoácido, y un par de 3 regiones conservadas separadas por un aminoácido entre cada uno. El alineamiento entre pgaC de *E. coli* y UDP-f de *S. enterica* (figura 5D), mostro en la primera columna de la figura tres regiones conservadas, en la segunda columna una dupla aminoácidos conservada separada por un aminoácido ambos de la misma especie de una tripleta conservada y por último en la tercera columna un grupo de 4 aminoácidos consecutivos conservados. El alineamiento entre BT3148 de *B. fragilis* y EnvC de *F. nucleatum* (figura 5E), mostraron 4 regiones conservadas, de no ser por la separación por un aminoácido del mismo tipo en ambas secuencias, se tendrían 6 regiones. El alineamiento entre BT3147 de *B. fragilis* y Wcaa de *S. enterica*, (figura 5F) mostró un grupo de 4 aminoácidos conservado en ambas secuencias. Entre AcrB de *B. fragilis* y Wza de *S. enterica* (figura 5G), se encontraron 5 regiones conservadas. Entre radD de *F. nucleatum* y pefC de *S. enterica* (figura 5H), se encontraron 6 regiones conservadas. El alineamiento entre Ftsx de *F. nucleatum* y UDP-f de *S. enterica* (figura 5I) mostrando 3 regiones conservadas en ambas secuencias y, por último, el alineamiento entre FtsX de *F. nucleatum* y UDP-f de *S. enterica* (figura 5J) mostraron 3 regiones conservadas.



**Figura 5.** Alineamiento entre las proteínas del biofilm de *F. nucleatum*, *E. coli*, *B. fragilis* y *S. enterica*. A) FimB de *E. coli* y WcaI de *S. enterica*, B) FimB de *E. coli* y WcaM de *S. enterica*, C) FimB de *E. coli* y WcaA de *S. enterica*, D) pgaC de *E. coli* y UDP-f de *S. enterica*, E) BT3148 de *B. fragilis* y EnvC de *F. nucleatum*, F) BT3147 de *B. fragilis* y WcaA de *S. enterica*, G) AcrB de *B. fragilis* y Wza de *S. enterica* H) radD de *F. nucleatum* y pefC de *S. enterica*, I) FtsX de *F. nucleatum* y UDP-f de *S. enterica* y J) FtsX de *F. nucleatum* y BT3147 *B. fragilis*.

## Acoplamiento entre las proteínas del biofilm de bacterias enteropatógenicas

Por último, se hizo la representación de una posible interacción entre *F. nucleatum*, *B. fragilis* y *S. enterica* a través de las proteínas que forman el biofilm de estas bacterias, utilizando proteínas que tuvieran estructuras similares y que tuvieran sitios donde se pudieran unir otras moléculas. El acoplamiento se realizó de forma manual, utilizando las estructuras de superficie de las proteínas EnvC, BT3147 y pefC de *F. nucleatum*, *B. fragilis* y *S. enterica* respectivamente (figura 6).



**Figura 6.** Modelo hipotético de la interacción entre las proteínas EnvC de *F. nucleatum*, BT3147 de *B. fragilis* y pefC de *S. enterica*.

## Conclusiones

En este trabajo se realizó la búsqueda de las secuencias de aminoácidos, alineamiento y análisis de las proteínas FtsX, EnvC, RadD, Fap2, CmpA de *F. nucleatum*; ompA, BT3148, BT3147, bmeB3, *B. fragilis*; csgD, hha, bcsA operon, pgaC, fimB de *E. coli*; fimH, wcaM, WCAA, WZA, pefC, ipfE, BapA, WCAL, wzxE, UDP-f, yhiO de *S. enterica*, mismas que participan en la formación del biofilm. Así mismo, se realizó el modelamiento de las estructuras secundarias y terciarias de dichas proteínas, para predecir la interacción que puede haber entre ellas.

Nuestros resultados mostraron la presencia de varias regiones conservadas entre las proteínas que conforman el biofilm de las cuatro bacterias enteropatógenas analizadas en este trabajo, lo que podría permitir la interacción entre ellas como se muestra en nuestro modelo hipotético.

Por último, este tipo de trabajos muestran la importancia de realizar análisis bioinformático con la finalidad de buscar interacciones entre proteínas y así diseñar moléculas que interfieran con la formación del biofilm, para que en el futuro se puedan diseñar nuevos fármacos para su inhibición y/o eliminación, lo que representaría una estrategia prometedora para la prevención y tratamiento de distintas enfermedades asociadas a esta estructura, incluidas las neoplasias del tracto gastrointestinal, como el cáncer colorrectal.

## Referencias

- Chew, S., Tan, L., Law, J., Pusparajah, P., Goh, B., Ab Mutalib, N., & Lee, L. (2020). Targeting Gut Microbial Biofilms—A Key to Hinder Colon Carcinogenesis? *Cancers*, 12(8), 2272. <https://doi.org/10.3390/cancers12082272>
- Copenhagen-Glazer, S., Sol, A., Abed, J., Naor, R., Zhang, X., Han, Y., & Bachrach, G. (2015). Fap2 of *Fusobacterium nucleatum* Is a Galactose-Inhibitable Adhesin Involved in Coaggregation, Cell Adhesion, and Preterm Birth. *Infection And Immunity*, 83(3), 1104-1113. <https://doi.org/10.1128/iai.02838-14>

- di Domenico, E. G., Cavallo, I., Pontone, M., Toma, L., & Ensoli, F. (2017). Biofilm Producing Salmonella Typhi: Chronic Colonization and Development of Gallbladder Cancer. *International Journal of Molecular Sciences*, 18(9), 1887. <https://doi.org/10.3390/ijms18091887>
- Flemming, H. C., Wingender, J., Szewzyk, U., Steinberg, P., Rice, S. A., & Kjelleberg, S. (2016). Biofilms: an emergent form of bacterial life. *Nature Reviews Microbiology*, 14(9), 563–575. <https://doi.org/10.1038/nrmicro.2016.94>
- Hernández-Luna, M., López-Briones, S., & Luria-Pérez, R. (2019). The Four Horsemen in Colon Cancer. *Journal Of Oncology*, 2019, 1–12. <https://doi.org/10.1155/2019/5636272>
- Jang, H., & Eom, Y. (2019). Antibiofilm and antibacterial activities of repurposing aurano-fin against *Bacteroides fragilis*. *Archives Of Microbiology*, 202(3), 473–482. <https://doi.org/10.1007/s00203-019-01764-3>
- Kim, H., Rhee, K., & Eom, Y. (2019). Anti-biofilm and antimicrobial effects of zerumbone against *Bacteroides fragilis*. *Anaerobe*, 57, 99–106. <https://doi.org/10.1016/j.anaerobe.2019.04.001>
- Ledeboer, N. A., & Jones, B. D. (2005). Exopolysaccharide Sugars Contribute to Biofilm Formation by *Salmonella enterica* Serovar Typhimurium on HEp-2 Cells and Chicken Intestinal Epithelium. *Journal of Bacteriology*, 187(9), 3214–3226. <https://doi.org/10.1128/jb.187.9.3214-3226.2005>
- Sharma, G., Sharma, S., Sharma, P., Chandola, D., Dang, S., Gupta, S., & Gabrani, R. (2016). *Escherichia coli* biofilm: development and therapeutic strategies. *Journal of applied microbiology*, 121(2), 309–319. <https://doi.org/10.1111/jam.13078>
- Shin, D., Rhee, K., & Eom, Y. (2020). Effect of Probiotic *Clostridium butyricum* NCTC 7423 Supernatant on Biofilm Formation and Gene Expression of *Bacteroides fragilis*. *Journal Of Microbiology and Biotechnology*, 30(3), 368–377. <https://doi.org/10.4014/jmb.2001.01027>
- Solano, C., García, B., Valle, J., Berasain, C., Ghigo, J. M., Gamazo, C., & Lasa, I. (2002). Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. *Molecular Microbiology*, 43(3), 793–808. <https://doi.org/10.1046/j.1365-2958.2002.02802.x>
- Wu, C., Al Mamun, A., Luong, T., Hu, B., Gu, J., & Lee, J. et al. (2018). Forward Genetic Dissection of Biofilm Development by *Fusobacterium nucleatum*: Novel Functions of Cell Division Proteins FtsX and EnvC. *Mbio*, 9(2). <https://doi.org/10.1128/mbio.00360-18>