

**Hirsfeld Institute of Immunology and Experimental Therapy
Polish Academy of Sciences**



Doctoral dissertation

**Study of biological functions of *Yersinia enterocolitica* bacteriophage tail
proteins**

Badanie funkcji biologicznych białek ogonka bakteriofagów
Yersinia enterocolitica

Karolina Filik

Dissertation Supervisor: dr hab. Ewa Brzozowska

The doctoral thesis was conducted in the Laboratory of Medical Microbiology of the
Department of Immunology of Infectious Diseases

Research was funded by National Science Center of Poland,
Sonata Bis 7 UMO-2017/26/E/NZ1/00249

Wrocław, 2022

1. Acknowledgments

I would like to thank Ewa Brzozowska, my dissertation advisor. Thank you for the trust you have placed in me, for your patience, all valuable advice and constant motivation for scientific development.

I would like to thank also Bożena Szermer-Olearnik for all advice, motivation and kindness.

I would like to thank all the students with whom I worked during my doctorate: Natalia, Dominika, Kay, Janek, Filip and Sabina – I am truly honored that I could work with you all during this 4 years.

I would like to thank my fiancé Irwin for always believing in me, always trying to help me, for discovering the scientific world together and especially for patience.

*I want to thank my family, the biggest thanks to my mum without whom I would not be here.
Mum this is for you.*

Table of content

1. Acknowledgments
2. Streszczenie
3. Abstract
4. The list of publication included in the dissertation
5. Author's declarations
6. Publications included in the dissertation
7. Achievements and scientific contribution resulting from the research carried out as part of this dissertation
8. Autoreferat

2. Streszczenie

Bakteriofagi są obiektem badań głównie w kontekście użycia ich w terapii przeciwko zakażeniom bakteryjnym. Pomimo korzyści jakie posiada terapia fagowa, istnieją także ograniczenia używania cząstek fagowych takie jak np. horyzontalny transfer genów (HTG). Najpoważniejszym skutkiem HTG może być nabywanie czynników wirulencji przez bakterie, stąd terapia fagowa pozostaje terapią wysokiego ryzyka. W związku z powyższym wiele badań naukowych koncentruje się na charakterystyce poszczególnych elementów wirionów fagowych, a mianowicie białkach, które mogą mieć funkcje enzymatyczne lub mogą być wykorzystane w diagnostyce patogenów. Początkowo bardzo intensywnie badano enzymy fagowe takie jak depolimerazy czy endolizyny powodujące lizę bakterii od wewnątrz komórki. Jednak obecnie uwaga naukowców została skierowana na białka ogonka. Białka ogonka bakteriofagów, do których zaliczamy białka tubularne oraz włókienkowe oprócz funkcji strukturalnej czyli budulcowej mogą pełnić funkcje enzymatyczne względem zewnętrznych struktur na powierzchni bakterii. Białka włókienkowe pełnią funkcję rozpoznającą i wiążącą receptor na komórkach bakteryjnych, czyli są białkami adhezyjnymi. Wśród białek ogonkowych występuje duża różnorodność, wynikająca z procesów adaptacyjnych, które w dużej mierze są wynikiem koewolucji fagów oraz ich gospodarzy. Z tego też względu na podstawie analizy porównawczej sekwencji nukleotydowych nie jest możliwe określenie przewidywanej funkcji tych białek. Dopiero badania eksperymentalne dają możliwość poznania ich właściwości.

Do badań będących tematem niniejszej rozprawy doktorskiej wykorzystano fagi ogonkowe *Yersinia enterocolitica* ϕ 80-18 oraz ϕ YeO3-12 należące do rodziny *Podoviridae*. Gospodarzem dla tych fagów jest *Yersinia enterocolitica*, patogen jelitowy, który powoduje chorobę u ludzi i niektórych zwierząt. Zakażenie tym patogenem wiąże się z wystąpieniem głównego objawu, którym jest wodnista biegunka, wynikająca ze stanu zapalnego toczącego się w jelitach. Jest to choroba odzwierzęca, którą możemy się głównie zarazić poprzez spożycie wieprzowiny. Według ECDC (Europejskie Centrum Zapobiegania i Kontroli Chorób) w 2019 zanotowano 7048 przypadków jersiniozy w 29 europejskich krajach. Ze względu na charakter choroby liczba ta wydaje się niedoszacowana i faktyczny odsetek chorujących prawdopodobnie jest większy. Szersza diagnostyka tego schorzenia jest przeprowadzana dopiero w momencie występowania większego ogniska zakażeń, bądź kiedy infekcja prowadzi do hospitalizacji. Diagnostyka *Yersinia enterocolitica* jest oparta głównie o metody hodowlane, które w dużej mierze są czasochłonne, oraz dodatkowo wymagają potwierdzenia w metodach molekularnych czy z wykorzystaniem spektrometrii mas. Gatunek *Yersinia enterocolitica* stanowi heterogenna grupa szczepów, które są wyróżniane na podstawie biotypowania oraz serotypowania. Wyróżniamy odpowiednio 6 biotypów oraz 57 O serotypów, co dodatkowo komplikuje identyfikację w obrębie gatunku. Znanym w literaturze podejściem, jest możliwość użycia białek ogonkowych do diagnostyki patogenów. Białka te często posiadają w swojej sekwencji regiony, które specyficznie rozpoznają dany patogen, może się to przyczyniać do szybszego stawiania diagnozy już w momencie zaistnienia infekcji.

Celem rozprawy doktorskiej była charakterystyka białek ogonka fagów pod kątem właściwości enzymatycznych oraz adhezyjnych i ocena możliwości praktycznego ich wykorzystania zarówno w zwalczaniu jak i detekcji *Yersinia enterocolitica*.

Badania prowadzące do osiągnięcia powyższego celu zostały ujęte w opublikowanym cyklu czterech prac oryginalnych. Są one poszerzone także o badania właściwości biologicznych bakteriofaga ϕ 80-18. W pierwszej pracy badano funkcje oraz mechanizm działania białka TTPAgp11 (ang. Tail Tubular Protein A) pochodzącego z faga ϕ YeO3-12. Jest to białko z grupy białek TTPA, do której należy gp31 z faga *Klebsiella pneumoniae* KP32 wykazującego aktywność α -1,4-glukozydazy. W celu wykazania, czy białko TTPAgp11 również posiada tę samą aktywność enzymatyczną co białko TTPAgp31, wyprodukowano białko rekombinowane TTPAgp11, następnie użyto skrobi (komercyjny substrat Red-Starch) i maltozy jako substratów w testach trawienia z wykorzystaniem metody kolorymetrycznej i chromatografii gazowej. Wykazano, że białko to posiada aktywność hydrolityczną względem badanych substratów i możemy je sklasyfikować do enzymów z rodziny α -1,4-glukozydaz. Dzięki temu potwierdzono hipotezę o dwufunkcyjności białek ogonkowych (ang. dual-function protein) również w przypadku fagów *Yersinia*.

Druga praca cyklu dotyczyła ogólnej charakterystyki bakteriofaga ϕ 80-18. Została ona zainicjowana trudnościami z namnożeniem faga, a co za tym idzie, brakiem możliwości ekstrakcji matrycy DNA niezbędnej do syntezy genów białek ogonkowych. Okazało się także, że brakuje danych literaturowych na temat jego biologii i ogólnej charakterystyki. Bakteriofag został wyizolowany w latach 90, natomiast dopiero w 2020 roku został opisany właśnie przy okazji realizacji niniejszej pracy doktorskiej. Dla faga ϕ 80-18 wyznaczono krzywą wzrostu (ang. one-step growth curve), określono jego stabilność w szerokim zakresie pH oraz temperatury, przeanalizowano przynależność filogenetyczną, ale także przeprowadzono analizę genomu i proteomu oraz wyznaczono zakres gospodarzy. Analiza filogenetyczna potwierdziła przynależność faga do rodziny *Podoviridae* oraz podrodziny *Autographivirinae*. Gospodarzem dla tego faga jest patogenny szczep *Y. enterocolitica* O:8 biotyp 1B, ale również serotypy: O:4, O:4,32, O:20 i O:21. Otrzymane wyniki wskazują na możliwość wykorzystania ϕ 80-18 w biokontroli amerykańskiego biotypu 1B *Y. enterocolitica*.

Trzecia przedstawiona praca dotyczyła charakterystyki białek TTPBgp12 (ang. Tail Tubular Protein B) oraz TFPgp17 (ang. Tail Fiber Protein), które pochodzą z faga ϕ YeO3-12. Białka te przynależą do odpowiednich grup białek opisanych wcześniej dla bakteriofagów *Klebsiella pneumoniae*. Z tego też względu, białko TTPBgp12 oprócz funkcji strukturalnej, było rozpatrywane pod kątem posiadania funkcji enzymatycznej. Nie udało się wyznaczyć specyficzności substratowej dla białka TTPBgp12. W toku badań stwierdzono natomiast, że białko to hamuje wzrost bakterii i spowalnia rozwój biofilmu bakteryjnego *Y. enterocolitica*. W pracy przedstawiono obszerną analizę porównawczą *in silico* dla obu białek. Ciekawy wynik uzyskano z analizy dla białka TFPgp17, dla którego wykazano podobieństwo do białka RsaA obecnego w warstwie S u bakterii Gram-ujemnych i Gram-dodatnich. Podobieństwo strukturalne obu białek fagowego i bakteryjnego może być wynikiem ewolucyjnego przystosowania się wirusa do efektywnego infekowania swojego gospodarza. W pracy przedstawiono także wyniki testów dotyczących stabilności białek ogonka w obecności

wybranych cukrów. Najistotniejszym okazał się wpływ stabilizujący N-acetylogalaktozaminy na białko TFPgp17. GalNac jest głównym źródłem węgla dla bakterii *Yersinia* w przewodzie pokarmowym człowieka oraz świń. Nie dziwi więc fakt, że ten aminocukier może stabilizować całe cząstki fagowe tam obecne. Wykonane badania mogą częściowo potwierdzać tę hipotezę.

Ostatnia przedstawiona praca dotyczyła białka TFPgp17 i jego wykorzystania jako specyficznej adhezyny wykrywającej patogenny serotyp O:3 *Yersinia enterocolitica*. Białko TFPgp17 wyprodukowano w kompleksie z MBP (ang. maltose binding protein) z metką histydynową, co było konieczne do umożliwienia detekcji, nie zakłócało właściwości adhezyjnych białka fagowego i dodatkowo pozwalało na znaczne skrócenie procedury oczyszczania. Ponadto obecność metki daje możliwość immobilizacji białka (za pomocą metki) do nośnika np. czujnika światłowodowego w ściśle ukierunkowany, nieprzypadkowy sposób i wykorzystania do opracowania czujników np. optycznych. W przeprowadzonych badaniach wykonano test ELISA, w którym analizowano nie tylko szczepy *Y. enterocolitica* o różnych serotypach, ale także inne szczepy m.in. *K. pneumoniae*, *S. aureus*, *P. aeruginosa*. Test miał na celu ocenę specyficzności białka TFPgp17. Wykazano, że było ono wysoce specyficzne i umożliwiało detekcję tylko *Y. enterocolitica* o serotypie O:3. W celu potwierdzenia uzyskanych wyników wykonano analizę w mikroskopii TEM, gdzie immobilizowano bakterie do siatki, a następnie inkubowano z TFPgp17. Detekcja utworzonego połączenia między białkiem, a bakterią była możliwa właśnie dzięki obecności zastosowanej metki, która była wykrywana przeciwciałami anty-metka zmodyfikowanymi złotem koloidalnym. Obrazowanie TEM pokazało przyczepianie się kompleksu TFPgp17-MBP/his tag na powierzchni komórki bakteryjnej *Y. enterocolitica* O:3, co było potwierdzeniem wyników otrzymanych w teście ELISA.

Podsumowując, cel pracy doktorskiej został osiągnięty. W toku badań udało się uzyskać i scharakteryzować 3 białka ogonka bakteriofagów. Dla dwóch z nich, obok funkcji strukturalnej, udało się wykazać właściwości biologiczne i tak dla białka TTPAgp11 określono funkcję enzymatyczną, α -1,4-glukozydazy, dla białka TTPBgp12 wykazano znaczący wpływ hamujący wzrost bakterii patogennych i tworzenia się biofilmu bakteryjnego. Wskazanie antyseptycznego potencjału białka TTPBgp12 może być wykorzystane do rozwoju preparatów biologicznie czynnych w przyszłości. W przypadku białka włókienkowego TFPgp17 potwierdzono jego adhezyjną funkcję. Określono jego specyficzność względem bakterii. Ponadto opracowano wydajną i szybką metodę uzyskiwania tego białka w formie zmodyfikowanej do szerokiego zastosowania w rozwoju biosensorów do szybkiej detekcji skażeń *Yersinia enterocolitica* serotypu O:3.

3. Abstract

Bacteriophages are the subject of research mainly in the context of their use in therapy against bacterial infections. Despite the benefits of phage therapy, there are also limitations in using phage particles, such as horizontal gene transfer (HTG). The most serious effect of HTG may be the acquisition of virulence factors by bacteria, which why phage therapy still remains a high-risk therapy. Therefore, many scientific studies have focused on the characterizing individual elements of phage virions. These studies have mainly focused on proteins that may have enzymatic functions or may be used for diagnostic purposes. Initially, phage enzymes such as depolymerases or endolysins, which lyse bacteria from inside the cell, were studied intensively. However, scientists' attention has now been focused on tail proteins. The proteins of the bacteriophage tail, which include tubular and filamentary proteins, in addition to their structural function, can perform enzymatic functions on bacterial surface structures. Fiber proteins perform the function of recognizing and binding the receptor on bacterial cells and they are called adhesin. There is a great diversity among the tail proteins which have many adaptive processes. These proteins are largely the result of the co-evolution of phages and their hosts. Therefore, comparative analysis of the nucleotide sequences alone can not determine the predicted function of these proteins. Only experimental research can be used to learn about their properties.

The *Yersinia enterocolitica* ϕ 80-18 and ϕ YeO3-12 tail phages belonging to the *Podoviridae* family were used for the research subject this doctoral dissertation. The host for these phages is *Yersinia enterocolitica*, an enteropathogen that causes disease in humans and some animals. The main symptom of infection is watery diarrhea which is a result of inflammation in the intestines. Yersiniosis is a zoonotic disease, infection which is spread mainly by contaminated pork. According to ECDC (European Center for Disease Prevention and Control) report, in 2019 there were 7,058 cases of yersiniosis in 29 European countries. Due to the nature of the disease, this number seems to be an underestimation and the actual percentage of patients with the disease is probably higher. Widespread diagnostic protocols of this disease are carried out only when there is a larger outbreak of infection, or when the reported infections lead to hospitalization. Diagnosis of *Yersinia enterocolitica* is based mainly on culture methods, which are largely time-consuming, and additionally require confirmation in molecular methods or with mass spectrometry. *Yersinia enterocolitica* is a heterogeneous group of strains that are distinguished on the basis of biotyping and serotyping. We distinguish 6 biotypes and 57 O serotypes, respectively, this complicates the identification within the species. An approach known in the literature is the possibility of using tail proteins for diagnosing pathogens. These proteins often have regions in their sequences that specifically recognize a pathogen. This may contribute to faster diagnosis already at the time of infection.

The aim of this doctoral dissertation was to characterize phage tail proteins in terms of enzymatic and adhesive properties and to evaluate the possibility of their practical use both in the control and detection of *Yersinia enterocolitica*.

Research leading to the achievement of the above goal was included in the published cycle of four original papers. They also include research on the biological properties of the ϕ 80-18 bacteriophage. In the first paper, the function and mechanism of the action of the TTPAgp11 (Tail Tubular Protein A) protein derived from phage ϕ YeO3-12 was investigated. Gp11 is a protein from the group of TTPA proteins, which includes also gp31 from *Klebsiella pneumoniae* KP32 phage, which shows α -1,4-glucosidase activity. In order to demonstrate whether the TTPAgp11 protein also has the same enzymatic activity as the TTPAgp31 protein, the recombinant TTPAgp11 protein was produced, then starch (a commercial Red-Starch substrate) and maltose were used as substrates in the colorimetric and gas chromatographic enzymatic assays. It has been shown that this protein has a hydrolytic activity towards the tested substrates and we can classify it into enzymes from the α -1,4-glucosidase family. Due to these results, the hypothesis of dual-function proteins was confirmed also in the case of *Yersinia* phages.

The second paper was concerned with the general characteristics of the ϕ 80-18 bacteriophage. In the beginning there were difficulties with phage multiplication, and thus we were unable to extract the DNA template necessary for the synthesis of tail protein genes. It turned out that there is a lack of literature data on its biology and general characteristics. The bacteriophage was isolated in the 1990s, but it only was described in 2020 during one of the studies presented in this doctoral dissertation. For ϕ 80-18 phage, a one-step growth curve was determined, its stability in a wide range of pH and temperature was determined, phylogenetic affiliation was analyzed, but also genome and proteome analysis was performed and the hosts range was determined. Phylogenetic analysis confirmed that the phage belonged to the *Podoviridae* family and the *Autographivirinae* subfamilies. The host for this phage is the pathogenic strain of *Y. enterocolitica* O: 8 biotype 1B, but also the serotypes: O: 4, O: 4.32, O: 20 and O: 21. The obtained results indicate the possibility of using ϕ 80-18 in the biocontrol of the American biotype 1B *Y. enterocolitica*.

The third paper presented was concerned with the characteristics of the TTPBg12 (Tail Tubular Protein B) and TFPgp17 (Tail Fiber Protein) proteins, which are derived from the phage ϕ YeO3-12. These proteins belong to the appropriate groups of proteins described earlier for *Klebsiella pneumoniae* bacteriophages. Therefore, the TTPBg12 protein, in addition to its structural function, has been considered to have some enzymatic function. The substrate specificity for the TTPBg12 protein could not be determined. In the course of the research, it was found that this protein inhibits the growth of bacteria and slows down the development of *Y. enterocolitica* bacterial biofilm. The paper presents an extensive comparative in silico analysis for both proteins. An interesting result was obtained from the analysis for the TFPgp17 protein, which was similar to the RsaA protein present in the S layer in gram-negative and gram-positive bacteria. The structural similarity of both phage and bacterial proteins may be the result of the evolutionary adaptation of the virus to effectively infect its host. The paper also presents the results of tests on the stability of the tail proteins in the presence of selected sugars. The stabilizing effect of N-acetylgalactosamine on the TFPgp17 protein turned out to be the most important. GalNac is the main source of carbon for *Yersinia* bacteria in the digestive tract

of humans and pigs. It is therefore, not surprising that this amino sugar can stabilize the entire phage particles present there. These tests partially confirm this hypothesis.

The latest paper was about the TFPgp17 protein and its use as a specific adhesin to detect the pathogenic *Yersinia enterocolitica* serotype O:3. The TFPgp17 protein was produced in a complex with MBP (maltose binding protein) with a histidine tag, which was necessary to enable detection. It did not disturb the adhesion properties of the phage protein and additionally allowed for a significant shortening of the purification procedure. In addition, the presence of a tag makes it possible to immobilize the protein (with a tag) to a carrier in a strictly targeted way and use it to develop sensors, e.g. optical sensors. In the conducted research, an ELISA test was performed, which analyzed not only *Y. enterocolitica* strains of various serotypes, but also other strains, e.g. *K. pneumoniae*, *S. aureus*, *P. aeruginosa*. The test was designed to assess the specificity of the TFPgp17 protein. It was shown to be highly specific and able to detect only *Y. enterocolitica*, serotype O:3. To confirm the obtained results, the TEM microscopy analysis was performed, where the bacteria were immobilized to a mesh and then incubated with TFPgp17. We detected the creation of a connection between the protein and the bacteria. This was possible thanks to the presence of the applied tag, which was detected with anti-tag antibodies modified with colloidal gold. TEM imaging showed the attachment of the TFPgp17-MBP / his tag complex on the surface of the *Y. enterocolitica* O:3 bacterial cell, which confirmed the results obtained in the ELISA test.

In conclusion, the goal of this dissertation has been achieved. In the course of the research, it was possible to obtain and characterize 3 proteins of the bacteriophage tail. For two of them, apart from the structural function, it was possible to demonstrate their biological properties. For the TTPAgp11 protein the enzymatic function of α -1,4-glucosidase was determined. For the TTPBgp12 protein it was shown a significant inhibitory effect on the growth of pathogenic bacteria and the formation of bacterial biofilm. The indication of the antiseptic potential of the TTPBgp12 protein can be used for the development of biologically active preparations in the future. In the case of TFPgp17 protein, its adhesin function was confirmed. Its specificity for the bacteria was determined. In addition, an efficient and rapid method of obtaining this protein in a modified form was developed for wide use in the development of biosensors for the rapid detection of *Yersinia enterocolitica* O:3 contamination.

4. The list of publication included in the dissertation

- I. Pyra, A., Urbańska, N., **Filik, K.**, Tyrlik, K., & Brzozowska, E. (2020). Biochemical features of the novel Tail Tubular Protein A of *Yersinia* phage phiYeO3-12. *Scientific reports*, 10(1), 1-11.
- II. **Filik, K.***, Szermer-Olearnik, B.*, Wernecki, M., Happonen, L. J., Pajunen, M. I., Nawaz, A., Qasim M. S., Jun J. W., Mattinen, L., Skurnik, M.*, Brzozowska, E. (2020). The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*. *Frontiers in Microbiology*, 11, 1356. *autor korespondencyjny
- III. Pyra, A., **Filik, K.**, Szermer-Olearnik, B., Czarny, A., & Brzozowska, E. (2020). New Insights on the Feature and Function of Tail Tubular Protein B and Tail Fiber Protein of the Lytic Bacteriophage ϕ YeO3-12 Specific for *Yersinia enterocolitica* Serotype O: 3. *Molecules*, 25(19), 4392.
- IV. **Filik, K.**, Szermer-Olearnik, B., Niedziółka-Jönson, J., Roźniecka, E., Ciekot, J., Pyra, A., Matyjaszczyk, I., Skurnik, M., Brzozowska, E. (2022). ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3. *AMB Expr* 12, 1.

5. Author's declarations

Anna Pyra
 Faculty of Chemistry
 University of Wrocław
 14 F. Joliot-Curie St.
 58-383 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Anna Pyra, Natalia Urbańska, Karolina Filik, Katherine Tyrlik, Ewa Brzozowska, Biochemical features of the novel Tail Tubular Protein A of Yersinia phage ϕ YeO3-12, *Scientific reports*, 2020, 10(1), 1-11.

IF₂₀₂₀=4,38; MEiN=140

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Anna Pyra	50%	Conceptualization; methodology: bioinformatics analysis; CD measurements ESI-MS analysis; Gas chromatography analysis; raw data analysis; figures and tables preparation; manuscript preparation; writing – editing
Natalia Urbańska	10%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, hydrolytic analysis towards red starch, protein folding and aggregation tests
Karolina Filik	30%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, protein folding and aggregation tests; supervision of the students; results analysis; writing – review
Katherine Tyrlik	5%	Methodology: Protein production, SDS-PAGE; writing – English corrections
Ewa Brzozowska	5%	Conceptualization; supervision; results analysis; writing – review and editing; funding acquisition

Anna Pyra

12.01.2022

Natalia Urbańska
Hirszfeld Institute of Immunology and Experimental Therapy
Polish Academy of Science
Weigl 12
53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Anna Pyra, Natalia Urbańska, **Karolina Filik**, Katherine Tyrlik, Ewa Brzozowska, Biochemical features of the novel Tail Tubular Protein A of Yersinia phage ϕ YeO3-12, *Scientific reports*, 2020, 10(1), 1-11.

IF₂₀₂₀=4,38; MEiN=140

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Anna Pyra	50%	Conceptualization; methodology: bioinformatics analysis; CD measurements ESI-MS analysis; Gas chromatography analysis; raw data analysis; figures and tables preparation; manuscript preparation; writing – editing
Natalia Urbańska	10%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, hydrolytic analysis towards red starch, protein folding and aggregation tests
Karolina Filik	30%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, protein folding and aggregation tests; supervision of the students; results analysis; writing – review
Katherine Tyrlik	5%	Methodology: Protein production, SDS-PAGE; writing – English corrections
Ewa Brzozowska	5%	Conceptualization; supervision; results analysis; writing – review and editing; funding acquisition



Karolina Filik
 Hirszfeld Institute of Immunology and Experimental Therapy
 Polish Academy of Science
 Weigl 12
 53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Anna Pyra, Natalia Urbańska, **Karolina Filik**, Katherine Tyrlik, Ewa Brzozowska, Biochemical features of the novel Tail Tubular Protein A of Yersinia phage phiYeO3-12, *Scientific reports*, 2020, 10(1), 1-11.

IF₂₀₂₀=4,38; MEiN=140

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Anna Pyra	50%	Conceptualization; methodology: bioinformatics analysis; CD measurements ESI-MS analysis; Gas chromatography analysis; raw data analysis; figures and tables preparation; manuscript preparation; writing – editing
Natalia Urbańska	10%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, hydrolytic analysis towards red starch, protein folding and aggregation tests
Karolina Filik	30%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, protein folding and aggregation tests; supervision of the students; results analysis; writing – review
Katherine Tyrlik	5%	Methodology: Protein production, SDS-PAGE; writing – English corrections
Ewa Brzozowska	5%	Conceptualization; supervision; results analysis; writing – review and editing; funding acquisition

Karolina Filik

12.01.2022

Katherine Tyrlik
Hirsfeld Institute of Immunology and Experimental Therapy
Polish Academy of Science
Weigl 12
53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Anna Pyra, Natalia Urbańska, **Karolina Filik**, Katherine Tyrlik, Ewa Brzozowska, Biochemical features of the novel Tail Tubular Protein Λ of Yersinia phage ϕ YeO3-12, *Scientific reports*, 2020, 10(1), 1-11.

IF₂₀₂₀=4,38; MEiN=140

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Anna Pyra	50%	Conceptualization; methodology: bioinformatics analysis; CD measurements ESI-MS analysis; Gas chromatography analysis; raw data analysis; figures and tables preparation; manuscript preparation; writing – editing
Natalia Urbańska	10%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, hydrolytic analysis towards red starch, protein folding and aggregation tests
Karolina Filik	30%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, protein folding and aggregation tests; supervision of the students; results analysis; writing – review
Katherine Tyrlik	5%	Methodology: Protein production, SDS-PAGE; writing – English corrections
Ewa Brzozowska	5%	Conceptualization; supervision; results analysis; writing – review and editing; funding acquisition

Katherine Tyrlik

Ewa Brzozowska
 Hirszfeld Institute of Immunology and Experimental Therapy
 Polish Academy of Science
 Weigl 12
 53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Anna Pyra, Natalia Urbańska, **Karolina Filik**, Katherine Tyrlik, Ewa Brzozowska, Biochemical features of the novel Tail Tubular Protein A of Yersinia phage phiYeO3-12, *Scientific reports*, 2020, 10(1), 1-11.

IF₂₀₂₀=4,38; MEiN=140

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Anna Pyra	50%	Conceptualization; methodology: bioinformatics analysis; CD measurements ESI-MS analysis; Gas chromatography analysis; raw data analysis; figures and tables preparation; manuscript preparation; writing – editing
Natalia Urbańska	10%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, hydrolytic analysis towards red starch, protein folding and aggregation tests
Karolina Filik	30%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, protein folding and aggregation tests; supervision of the students; results analysis; writing – review
Katherine Tyrlik	5%	Methodology: Protein production, SDS-PAGE; writing – English corrections
Ewa Brzozowska	5%	Conceptualization; supervision; results analysis; writing – review and editing; funding acquisition

Brzozowska

Karolina Filik

Hirszfeld Institute of Immunology and Experimental Therapy
Polish Academy of Science
Weigl 12, 53-114 Wrocław, Poland

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska, The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*, *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5,640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology: stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of physical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assembly of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acquisition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acquisition

Karolina Filik

Bożena Szermer-Olearnik
 Hirszfild Institute of Immunology and Experimental Therapy
 Polish Academy of Science
 Weigl 12
 53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska, The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*, *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5,640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology: stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of physical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assembly of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acquisition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acquisition

Bożena Szermer-Olearnik

Maciej Wernecki
 Department of Microbiology
 University of Wrocław
 Przybyszewskiego 63
 51-148 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska, The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*, *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5,640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology: stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of physical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assembly of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acquisition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acquisition

Maciej Wernecki



12.01.2022

Lotta J. Happonen
Infection Medicine
Department of Clinical Sciences Lund
Lund University
Lund, Sweden

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska, The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*, *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5,640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology: stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of phisical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assembly of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acquisition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acquisition

Maria I. Pajunen
 Department of Bacteriology and Immunology
 University of Helsinki
 Helsinki, Finland

Declaration

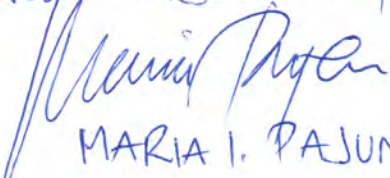
I hereby declare that my contribution to the following manuscript:

Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska, The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*, *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5,640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology: stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of phisical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assembly of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acquisition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acquisition

Approved, 14 Jan 2022, Helsinki

 MARIA I. PAJUNEN



Ayesha Nawaz
Department of Bacteriology and Immunology
University of Helsinki
Helsinki, Finland

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska. The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*. *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5.640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology; stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of physical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assembly of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acquisition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acquisition

Muhammad Suleman Qasim
 Faculty of Biological and Environmental Science
 University of Helsinki
 Helsinki, Finland

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska, The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*, *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5,640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology: stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of phisical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assembly of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acquisition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acquisition

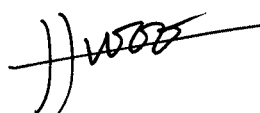
13.1.2022

X 

Muhammad Suleman Qasim

Signed by: Suleman Qasim

Jin Woo Jun
 Department of Aquaculture
 The Korea National College of Agriculture and Fisheries
 Jeonju, South Korea

Jin Woo Jun


Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska, The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*, *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5,640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology: stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of phisical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assemblby of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acqusition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acqusition

Laura Mattinen
 Department of Bacteriology and Immunology
 University of Helsinki
 Helsinki, Finland

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska, The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*, *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5,640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology: stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of physical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assembly of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acquisition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acquisition



Mikael Skurnik
 Department of Bacteriology and Immunology
 University of Helsinki
 Helsinki, Finland

Declaration

I hereby declare that my contribution to the following manuscript:

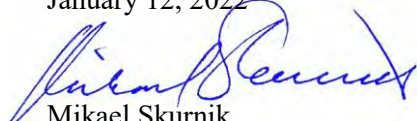
Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska, The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*, *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5,640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology: stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of phisical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assembly of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acquisition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acquisition

January 12, 2022



Mikael Skurnik

Ewa Brzozowska
 Hirszfeld Institute of Immunology and Experimental Therapy
 Polish Academy of Science
 Weigl 12
 53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik*, Bożena Szermer-Olearnik*, Maciej Wernecki, Lotta J. Happonen, Maria I. Pajunen, Ayesha Nawaz, Muhammad Suleman Qasim, Jin Woo Jun, Laura Mattinen, Mikael Skurnik*, Ewa Brzozowska, The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*, *Frontiers in Microbiology*, 2020, 11, 1356. *correspondence

Impact Factor: 5,640 (2020); MEiN: 100

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	35%	Conceptualization; methodology: stability experiments, one-step growth curve, phylogenetic analysis; raw data analysis; figures preparation; statistical analysis; preparation of the manuscript; correspondence
Bożena Szermer-Olearnik	20%	Conceptualization; Methodology: TEM photography, performing stability experiments and one-step growth curve; writing – review and editing; correspondence
Maciej Wernecki	5%	Phylogenetic analysis
Lotta J. Happonen	5%	Initial annotation
Maria I. Pajunen	5%	Methodology: Purification and preparation phages particles for proteomic analysis
Ayesha Nawaz	5%	Initial annotation
Muhammad Suleman Qasim	5%	Determination of physical end of the genome
Jin Woo Jun	5%	Host range analysis
Laura Mattinen	5%	Methodology: Preparation of the phage genome for Illumina sequencing and carried out of the initial de novo assembly of genome
Mikael Skurnik	5%	Conceptualization; Final annotation of the genome; host range analysis; raw data analysis; supervision; writing – review and editing; funding acquisition; correspondence
Ewa Brzozowska	5%	Supervision; writing – review and editing; funding acquisition

Brzozowska

Anna Pyra
 Faculty of Chemistry
 University of Wrocław
 14 F. Joliot-Curie St.
 58-383 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Anna Pyra, **Karolina Filik**, Bożena Szermer-Olearnik, Anna Czarny, Ewa Brzozowska, New Insights on the Feature and Function of Tail Tubular Protein B and Tail Fiber Protein of the Lytic Bacteriophage ϕ YeO3-12 Specific for *Yersinia enterocolitica* Serotype O:3, *Molecules*, 2020, 25(19):4392.

Impact Factor: 4.412 ; MEiN: 100.

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Anna Pyra	45%	Conceptualization; methodology; bioinformatics analysis; CD measurements; raw data analysis; figures preparation; manuscript preparation; writing – editing
Karolina Filik	35%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, protein folding and aggregation tests (nanoDSF); raw data analysis; figures preparation; writing – review
Bożena Szermer-Olearnik	10%	Methodology: biofilm inhibition test; Light-microscopy photography
Anna Czarny	5%	Methodology: biofilm inhibition test
Ewa Brzozowska	5%	Conceptualization: supervision; results analysis; writing – review and editing; funding acquisition

Anna Pyra

Karolina Filik

Hirszfeld Institute of Immunology and Experimental Therapy

Polish Academy of Science

Weigl 12

53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Anna Pyra, **Karolina Filik**, Bożena Szermer-Olearnik, Anna Czarny, Ewa Brzozowska, New Insights on the Feature and Function of Tail Tubular Protein B and Tail Fiber Protein of the Lytic Bacteriophage ϕ YeO3-12 Specific for *Yersinia enterocolitica* Serotype O:3, *Molecules*, 2020, 25(19):4392.

Impact Factor: 4,412 ; MEiN: 100.

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Anna Pyra	45%	Conceptualization; methodology: bioinformatics analysis; CD measurements; raw data analysis; figures preparation; manuscript preparation; writing – editing
Karolina Filik	35%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, protein folding and aggregation tests (nanoDSF); raw data analysis; figures preparation; writing – review
Bożena Szermer-Olearnik	10%	Methodology: biofilm inhibition test; Light-microscopy photography
Anna Czarny	5%	Methodology: biofilm inhibition test
Ewa Brzozowska	5%	Conceptualization; supervision; results analysis; writing – review and editing; funding acquisition

Karolina Filik

Bożena Szermer-Olearnik
 Hirszfeld Institute of Immunology and Experimental Therapy
 Polish Academy of Science
 Weigl 12
 53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Anna Pyra, **Karolina Filik**, Bożena Szermer-Olearnik, Anna Czarny, Ewa Brzozowska, New Insights on the Feature and Function of Tail Tubular Protein B and Tail Fiber Protein of the Lytic Bacteriophage ϕ YeO3-12 Specific for Yersinia enterocolitica Serotype O:3, Molecules, 2020, 25(19):4392.

Impact Factor: 4,412 ; MEiN: 100.

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Anna Pyra	45%	Conceptualization; methodology: bioinformatics analysis; CD measurements; raw data analysis; figures preparation; manuscript preparation; writing – editing
Karolina Filik	35%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, protein folding and aggregation tests (nanoDSF); raw data analysis; figures preparation; writing – review
Bożena Szermer-Olearnik	10%	Methodology: biofilm inhibition test; Light-microscopy photography
Anna Czarny	5%	Methodology: biofilm inhibition test
Ewa Brzozowska	5%	Conceptualization; supervision; results analysis; writing – review and editing; funding acquisition

Bożena Szermer - Olearnik

Anna Czarny
 Hirszfeld Institute of Immunology and Experimental Therapy
 Polish Academy of Science
 Weigl 12
 53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Anna Pyra, **Karolina Filik**, Bożena Szermer-Olearnik, Anna Czarny, Ewa Bzozowska, New Insights on the Feature and Function of Tail Tubular Protein B and Tail Fiber Protein of the Lytic Bacteriophage ϕ YeO3-12 Specific for *Yersinia enterocolitica* Serotype O:3, *Molecules*, 2020, 25(19):4392.

Impact Factor: 4,412 ; MEiN: 100.

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Anna Pyra	45%	Conceptualization; methodology: bioinformatics analysis; CD measurements; raw data analysis; figures preparation; manuscript preparation; writing – editing
Karolina Filik	35%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, protein folding and aggregation tests (nanoDSF); raw data analysis; figures preparation; writing – review
Bożena Szermer-Olearnik	10%	Methodology: biofilm inhibition test; Light-microscopy photography
Anna Czarny	5%	Methodology: biofilm inhibition test
Ewa Brzozowska	5%	Conceptualization; supervision; results analysis; writing – review and editing; funding acquisition

Anna Czarny

Ewa Brzozowska
 Hirszfeld Institute of Immunology and Experimental Therapy
 Polish Academy of Science
 Weigl 12
 53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Anna Pyra, **Karolina Filik**, Bożena Szermer-Olearnik, Anna Czarny, Ewa Brzozowska, New Insights on the Feature and Function of Tail Tubular Protein B and Tail Fiber Protein of the Lytic Bacteriophage ϕ YeO3-12 Specific for *Yersinia enterocolitica* Serotype O:3, *Molecules*, 2020, 25(19):4392.

Impact Factor: 4,412 ; MEiN: 100.

is correctly characterized in the table below:

Contributor	Contribution [%]	Description of main tasks
Anna Pyra	45%	Conceptualization; methodology: bioinformatics analysis; CD measurements; raw data analysis; figures preparation; manuscript preparation; writing – editing
Karolina Filik	35%	Methodology: Protein production and purification via IMAC and SEC chromatography, SDS-PAGE, protein folding and aggregation tests (nanoDSF); raw data analysis; figures preparation; writing – review
Bożena Szermer-Olearnik	10%	Methodology: biofilm inhibition test; Light-microscopy photography
Anna Czarny	5%	Methodology: biofilm inhibition test
Ewa Brzozowska	5%	Conceptualization; supervision; results analysis; writing – review and editing; funding acquisition

Brzozowska

Karolina Filik
 Hirszfeld Institute of Immunology and Experimental Therapy
 Polish Academy of Science
 Weigl 12
 53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik, Bożena Szermer-Olearnik, Joanna Niedziółka-Jönson, Ewa Roźniecka, Jarosław Ciekot, Anna Pyra, Irwin Matyjaszczyk, Mikael Skurnik, Ewa Brzozowska, ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3, AMB Express, 2022, 12, 1.

Impact Factor : 3,298 , MEiN: 70.

is correctly characterized in the table below.

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	66%	Conceptualization, methodology: protein production and purification, selection of bacterial strains for ELISA, optimalization of ELISA test conditions, selection of conditions for the experiment with TEM; raw data analysis; preparation of figures and tables; statistical analysis; preparation of the manuscript and editing (KF wrote the first version of the manuscript)
Bożena Szermer-Olearnik	10%	Development of a preparation method for electron microscopy visualization; TEM photography
Joanna Niedziółka-Jönson	3%	Conceptualization – UV-Vis measurments; writing – review
Ewa Roźniecka	3%	UV-Vis measurments, nanoparticles modification
Jarosław Ciekot	3%	SEC chromatography and UPLC-MS analysis
Anna Pyra	2%	Review of manuscript
Irwin Matyjaszczyk	3%	Statistical analysis
Mikael Skurnik	5%	Review of manuscript
Ewa Brzozowska	5%	Conceptualization, funding acquisition, supervision; writing – review and editing

Karolina Filik

Bożena Szermer-Olearnik
 Hirszfeld Institute of Immunology and Experimental Therapy
 Polish Academy of Science
 Weigl 12
 53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik, Bożena Szermer-Olearnik, Joanna Niedziółka-Jönson, Ewa Roźniecka, Jarosław Ciekot, Anna Pyra, Irwin Matyjaszczyk, Mikael Skurnik, Ewa Brzozowska, ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3, AMB Express, 2022, 12, 1.

Impact Factor : 3,298 , MEiN: 70.

is correctly characterized in the table below.

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	66%	Conceptualization, methodology: protein production and purification, selection of bacterial strains for ELISA, optimalization of ELISA test conditions, selection of conditions for the experiment with TEM; raw data analysis; preparation of figures and tables; statistical analysis; preparation of the manuscript and editing (KF wrote the first version of the manuscript)
Bożena Szermer-Olearnik	10%	Development of a preparation method for electron microscopy visualization; TEM photography
Joanna Niedziółka-Jönson	3%	Conceptualization – UV-Vis measurments; writing – review
Ewa Roźniecka	3%	UV-Vis measurments, nanoparticles modification
Jarosław Ciekot	3%	SEC chromatography and UPLC-MS analysis
Anna Pyra	2%	Review of manuscript
Irwin Matyjaszczyk	3%	Statistical analysis
Mikael Skurnik	5%	Review of manuscript
Ewa Brzozowska	5%	Conceptualization, funding acquisition, supervision; writing – review and editing

Bożena Szermer-Olearnik

Joanna Niedziółka-Jönson
 Institute of Physical Chemistry
 Polish Academy of Science
 Kasprzaka 44/52
 01-224 Warsaw

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik, Bożena Szermer-Olearnik, Joanna Niedziółka-Jönson, Ewa Roźniecka, Jarosław Ciekot, Anna Pyra, Irwin Matyjaszczyk, Mikael Skurnik, Ewa Brzozowska, ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3, AMB Express, 2022, 12, 1.

Impact Factor : 3,298 , MEiN: 70.

is correctly characterized in the table below.

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	66%	Conceptualization, methodology: protein production and purification, selection of bacterial strains for ELISA, optimalization of ELISA test conditions, selection of conditions for the experiment with TEM; raw data analysis; preparation of figures and tables; statistical analysis; preparation of the manuscript and editing (KF wrote the first version of the manuscript)
Bożena Szermer-Olearnik	10%	Development of a preparation method for electron microscopy visualization; TEM photography
Joanna Niedziółka-Jönson	3%	Conceptualization – UV-Vis measurements; writing – review
Ewa Roźniecka	3%	UV-Vis measurements, nanoparticles modification
Jarosław Ciekot	3%	SEC chromatography and UPLC-MS analysis
Anna Pyra	2%	Review of manuscript
Irwin Matyjaszczyk	3%	Statistical analysis
Mikael Skurnik	5%	Review of manuscript
Ewa Brzozowska	5%	Conceptualization, funding acquisition, supervision; writing – review and editing



Ewa Roźniecka
 Institute of Physical Chemistry
 Polish Academy of Science
 Kasprzaka 44/52
 01-224 Warsaw

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik, Bożena Szermer-Olearnik, Joanna Niedziółka-Jönson, Ewa Roźniecka, Jarosław Ciekot, Anna Pyra, Irwin Matyjaszczyk, Mikael Skurnik, Ewa Brzozowska, ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3, AMB Express, 2022, 12, 1.

Impact Factor : 3,298 , MEiN: 70.

is correctly characterized in the table below.

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	66%	Conceptualization, methodology: protein production and purification, selection of bacterial strains for ELISA, optimalization of ELISA test conditions, selection of conditions for the experiment with TEM; raw data analysis; preparation of figures and tables; statistical analysis; preparation of the manuscript and editing (KF wrote the first version of the manuscript)
Bożena Szermer-Olearnik	10%	Development of a preparation method for electron microscopy visualization; TEM photography
Joanna Niedziółka-Jönson	3%	Conceptualization – UV-Vis measurements; writing – review
Ewa Roźniecka	3%	UV-Vis measurements, nanoparticles modification
Jarosław Ciekot	3%	SEC chromatography and UPLC-MS analysis
Anna Pyra	2%	Review of manuscript
Irwin Matyjaszczyk	3%	Statistical analysis
Mikael Skurnik	5%	Review of manuscript
Ewa Brzozowska	5%	Conceptualization, funding acquisition, supervision; writing – review and editing

Roźniecka

12.01.2022

Jarosław Ciekot
Hirsfeld Institute of Immunology and Experimental Therapy
Polish Academy of Science
Weigl 12
53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik, Bożena Szermer-Olearnik, Joanna Niedziółka-Jönson, Ewa Roźniecka, Jarosław Ciekot, Anna Pyra, Irwin Matyjaszczyk, Mikael Skurnik, Ewa Brzozowska, ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3, AMB Express, 2022, 12, 1.

Impact Factor : 3,298 , MEiN: 70.

is correctly characterized in the table below.

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	66%	Conceptualization, methodology: protein production and purification, selection of bacterial strains for ELISA, optimalization of ELISA test conditions, selection of conditions for the experiment with TEM; raw data analysis; preparation of figures and tables; statistical analysis; preparation of the manuscript and editing (KF wrote the first version of the manuscript)
Bożena Szermer-Olearnik	10%	Development of a preparation method for electron microscopy visualization; TEM photography
Joanna Niedziółka-Jönson	3%	Conceptualization – UV-Vis measurements; writing – review
Ewa Roźniecka	3%	UV-Vis measurements, nanoparticles modification
Jarosław Ciekot	3%	SEC chromatography and UPLC-MS analysis
Anna Pyra	2%	Review of manuscript
Irwin Matyjaszczyk	3%	Statistical analysis
Mikael Skurnik	5%	Review of manuscript
Ewa Brzozowska	5%	Conceptualization, funding acquisition, supervision; writing – review and editing

J. Ciekot

Anna Pyra
 Faculty of Chemistry
 University of Wrocław
 14 F Joliot-Curie St.
 58-383 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik, Bożena Szermer-Olearnik, Joanna Niedziółka-Jönson, Ewa Roźniecka, Jarosław Ciekot, Anna Pyra, Irwin Matyjaszczyk, Mikael Skurnik, Ewa Brzozowska, ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3, AMB Express, 2022, 12, 1.

Impact Factor : 3.298 , MEiN: 70.

is correctly characterized in the table below.

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	66%	Conceptualization, methodology: protein production and purification, selection of bacterial strains for ELISA, optimalization of ELISA test conditions, selection of conditions for the experiment with TEM; raw data analysis; preparation of figures and tables; statistical analysis; preparation of the manuscript and editing (KF wrote the first version of the manuscript)
Bożena Szermer-Olearnik	10%	Development of a preparation method for electron microscopy visualization; TEM photography
Joanna Niedziółka-Jönson	3%	Conceptualization – UV-Vis measurments; writing – review
Ewa Roźniecka	3%	UV-Vis measurments, nanoparticles modification
Jarosław Ciekot	3%	SEC chromatography and UPLC-MS analysis
Anna Pyra	2%	Review of manuscript
Irwin Matyjaszczyk	3%	Statistical analysis
Mikael Skurnik	5%	Review of manuscript
Ewa Brzozowska	5%	Conceptualization, funding acquisition, supervision; writing – review and editing

Anna Pyra

Irwin Matyjaszczyk
 Department of Mycology and Genetics
 University of Wrocław
 Przybyszewskiego 63
 51-148 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik, Bożena Szermer-Olearnik, Joanna Niedziółka-Jönson, Ewa Roźniecka, Jarosław Ciekot, Anna Pyra, Irwin Matyjaszczyk, Mikael Skurnik, Ewa Brzozowska, ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3, AMB Express, 2022, 12, 1.

Impact Factor : 3,298 , MEiN: 70.

is correctly characterized in the table below.

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	66%	Conceptualization, methodology: protein production and purification, selection of bacterial strains for ELISA, optimization of ELISA test conditions, selection of conditions for the experiment with TEM; raw data analysis; preparation of figures and tables; statistical analysis; preparation of the manuscript and editing (KF wrote the first version of the manuscript)
Bożena Szermer-Olearnik	10%	Development of a preparation method for electron microscopy visualization; TEM photography
Joanna Niedziółka-Jönson	3%	Conceptualization – UV-Vis measurements; writing – review
Ewa Roźniecka	3%	UV-Vis measurements, nanoparticles modification
Jarosław Ciekot	3%	SEC chromatography and UPLC-MS analysis
Anna Pyra	2%	Review of manuscript
Irwin Matyjaszczyk	3%	Statistical analysis
Mikael Skurnik	5%	Review of manuscript
Ewa Brzozowska	5%	Conceptualization, funding acquisition, supervision; writing – review and editing



Mikael Skurnik
 Department of Bacteriology and Immunology
 University of Helsinki
 Helsinki, Finland

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik, Bożena Szermer-Olearnik, Joanna Niedziółka-Jönson, Ewa Roźniecka, Jarosław Ciekot, Anna Pyra, Irwin Matyjaszczyk, Mikael Skurnik, Ewa Brzozowska, φYeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3, AMB Express, 2022, 12, 1.

Impact Factor : 3,298 , MEiN: 70.

is correctly characterized in the table below.

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	66%	Conceptualization, methodology: protein production and purification, selection of bacterial strains for ELISA, optimization of ELISA test conditions, selection of conditions for the experiment with TEM; raw data analysis; preparation of figures and tables; statistical analysis; preparation of the manuscript and editing (KF wrote the first version of the manuscript)
Bożena Szermer-Olearnik	10%	Development of a preparation method for electron microscopy visualization; TEM photography
Joanna Niedziółka-Jönson	3%	Conceptualization – UV-Vis measurements; writing – review
Ewa Roźniecka	3%	UV-Vis measurements, nanoparticles modification
Jarosław Ciekot	3%	SEC chromatography and UPLC-MS analysis
Anna Pyra	2%	Review of manuscript
Irwin Matyjaszczyk	3%	Statistical analysis
Mikael Skurnik	5%	Review of manuscript
Ewa Brzozowska	5%	Conceptualization, funding acquisition, supervision; writing – review and editing

January 12, 2022



Mikael Skurnik

Ewa Brzozowska
 Hirsfeld Institute of Immunology and Experimental Therapy
 Polish Academy of Science
 Weigl 12
 53-114 Wrocław

Declaration

I hereby declare that my contribution to the following manuscript:

Karolina Filik, Bożena Szermer-Olearnik, Joanna Niedziółka-Jönson, Ewa Roźniecka, Jarosław Ciekot, Anna Pyra, Irwin Matyjaszczyk, Mikael Skurnik, Ewa Brzozowska, φYeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3, AMB Express, 2022, 12, 1.

Impact Factor : 3,298 , MEiN: 70.

is correctly characterized in the table below.

Contributor	Contribution [%]	Description of main tasks
Karolina Filik	66%	Conceptualization, methodology: protein production and purification, selection of bacterial strains for ELISA, optimalization of ELISA test conditions, selection of conditions for the experiment with TEM; raw data analysis; preparation of figures and tables; statistical analysis; preparation of the manuscript and editing (KF wrote the first version of the manuscript)
Bożena Szermer-Olearnik	10%	Development of a preparation method for electron microscopy visualization; TEM photography
Joanna Niedziółka-Jönson	3%	Conceptualization – UV-Vis measurements; writing – review
Ewa Roźniecka	3%	UV-Vis measurements, nanoparticles modification
Jarosław Ciekot	3%	SEC chromatography and UPLC-MS analysis
Anna Pyra	2%	Review of manuscript
Irwin Matyjaszczyk	3%	Statistical analysis
Mikael Skurnik	5%	Review of manuscript
Ewa Brzozowska	5%	Conceptualization, funding acquisition, supervision; writing – review and editing

Brzozowska

6. Publications included in the dissertation

OPEN

Biochemical features of the novel Tail Tubular Protein A of *Yersinia* phage phiYeO3-12

Anna Pyra^{1*}, Natalia Urbańska^{2,3}, Karolina Filik³, Katherine Tyrlik³ & Ewa Brzozowska^{3*}

Tail Tubular Protein A (TTPA) was long thought to be strictly a structural protein of environmental bacteriophages. However, our recent work has suggested that some TTPAs have additional functional features and thus are dual-function proteins. This study introduces a new TTPA family member, TTPAgp11, which belongs to *Yersinia* phage phiYeO3-12. We cloned the gene, expressed it and then purified the phage protein. The protein, including its hydrolytic activity, was characterized. Our enzymatic activity tests showed that TTPAgp11 displayed hydrolytic activity towards Red-starch, suggesting that this enzyme could be classified as part as the $\alpha - 1, 4$ -glucosidase family. Protein folding and aggregation tests indicated that TTPAgp11 is a single-domain protein whose aggregation can be induced by maltose or N-acetylglucosamine. The spatial structure of TTPAgp11 seemed to resemble that of the first reported dual-function TTPA, TTPAgp31, which was isolated from *Klebsiella pneumoniae* phage 32.

Antibiotic-resistant and biofilm-forming bacteria are widespread problems in many aspects of human life, including medicine and food production. A biofilm is a large aggregated structure that is formed when bacteria stick to one another on a solid surface¹. This adhesion is governed by extracellularly secreted polysaccharides called exopolysaccharides (EPS)². EPS can exist as a shell that is ionically or covalently connected to the cell (capsular EPS) or as a mucus (slime EPS). A promising strategy for eradicating drug-resistant biofilm-forming bacteria is the use of lytic phages that contain enzymes capable of depolymerizing (and thereby destroying) a biofilm^{2,3}. Although phages are well known to have antibacterial efficiency⁴⁻⁸, they are not commonly used in therapeutic settings. Phages have a high risk of undergoing mutagenesis, resulting in the alteration of their biological nature. Also, our inability to monitor their presence in the body makes phage therapy implication challenging. Obligatory lytic phages are virulent and quickly kill their bacterial hosts by lysing them. A one-time, high dose of bacterial endotoxin can be risky for humans. On the other hand, lysogenic bacteriophages can transfer some virulence factors and thus may be associated with pathogenicity, in a process called “lysogenic conversion”; this is very common in *Escherichia coli*, *Streptococcus pyogenes*, *Salmonella enterica*, and *S. aureus*. Prophages, for example, can encode exotoxins, such as those causing the major pathogenicity of *E. coli* EHEC, by inter-phage interaction (verocytotoxins or Shiga-toxins) or by *Vibrio cholerae* (A-B-type exotoxin mediated by prophage CTX). In addition, a rather broad spectrum of other proteins play a significant role in bacterial virulence⁹.

Macromolecules, such as proteins, offer more control opportunities, beginning with their production process and ending with the relevant therapeutic effects. Creating a treatment plan using a bacteriophage “cocktail” is not trivial, since several parameters of phage biology have to be considered. First, the pharmaceutical phage production process must be strictly defined and accepted by national authorities including the European Medicines Agency (EMA) in Europe. Then a suitable phage treatment must be individualized for each patient. This means modifying the application frequency, duration of therapy, dosage, and pharmaceutical form for each clinical trial or individual therapy¹⁰.

Tail tubular protein A (TTPA) has been described as a structural protein¹¹. TTPA is also called the gatekeeper protein, since it forms the attachment for the tail spikes and is also thought to mediate the initiation of infection through sensing the deflection of the side fibers upon cell wall binding. TTPA, during infection of bacteria, interacts with the side fibers and constitutes the conformational switch by which the side fibers engage causing

¹University of Wrocław, Faculty of Chemistry, 14F. Joliot-Curie St, Wrocław, 50383, Poland. ²University of Wrocław, Faculty of Biological Sciences, Institute of Experimental Biology, 6 Kanonia St, Wrocław, 50328, Poland. ³Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, 12R. Weigl St, Wrocław, 53114, Poland. *email: anna.pyra@chem.uni.wroc.pl; ewa.brzozowska@hirszfeld.pl

the release of the capsid contents. TTPA remains in contact with tubular tail protein B (TTPB) which forms a nozzle-like structure mounted below TTPA and is thought to extend the tube through which the DNA travels. In contrast to TTPA, TTPB seems more like an adapter for mounting additional functions than an essential component of the virion¹².

Most phage depolymerases are encoded in phage structural proteins such as with tail fibers, baseplates and neck structures¹³. Thus, these proteins were long thought to function only as structural proteins. However, as previously reported by our group, the TTPAs from *Klebsiella* phages KP32 and KP34 (TTPAgp31 and TTPAgp44, respectively) have biological activity: TTPAgp31 acts as α -1, 4- glucosidase (EC 3.2.1.20), while TTPAgp44 acts as a α -1, 1- trehalase (EC 3.2.1.28)^{14,15}. Due to the structural and hydrolytic features of both TTPAs, they were designated as dual-function proteins. The crystal structure of TTPAgp31 (PDB code: 5 mu4) was also described, and it appeared that the protein contained 3D structural elements that distinguished it from all others found in the Protein Databank (PDB)¹⁴. It was hypothesized that the lytic properties of these TTPAs arised from the presence of an additional structural element that is not found in the other TTPAs reported to date.

TTPA has also been found among the structural proteins of *Yersinia enterocolitica* bacteriophage phi-YeO3-12¹⁶. *Yersinia enterocolitica* is a Gram-negative bacterium that causes food-borne acute or chronic gastrointestinal diseases¹⁷. The identified phage infects pathogenic strains of *Yersinia enterocolitica* with serotypes O:3 by recognizing bacterial lipopolysaccharides¹⁸.

Here, is the first report regarding the biochemical properties of the TTPA, *Yersinia* TTPAgp11. This dual-function TTPA acts as a hydrolytic enzyme and is structurally similar to TTPAgp31, which was previously described by our group^{14,15}.

Results and Discussion

The enzymatic activities of the dual-function TTPAs are only just beginning to be studied. As previously reported, the TTPAs encoded by genes 31 and 44 of the *Klebsiella* phages (KP32 and KP34) can hydrolyze saccharide substrates (e.g., maltose, trehalose and starch) as well as bacterial EPS^{14,15}. TTPAgp31 resembles maltase in its substrate specificity but lacks a catalytic domain homologous to that of maltase, whereas TTPAgp44 shows trehalase-like activity (UniProt code: D1L2Y9). The question was whether any other TTPAs from bacteriophages could potentially act against biofilm-forming bacteria. After searching the UniProt database¹⁹ some interesting candidates appeared. From these, the uncharacterized TTPA (UniProt code: Q9T106) encoded by gene 11 of *Yersinia* phage phiYeO3-12 was selected (GenBank code: AJ251805). BLAST amino acid sequence analysis²⁰ showed that TTPAgp11 is similar to the TTPAs of phages that infect other types of bacteria with the highest similarity to bacteriophages infecting *Citrobacter* and *Enterobacter* (99–100%) (Supplementary Data). Amino acid sequence similarity was also found between TTPAgp11 and tail fiber protein B from phages of *Yersinia*, *E. coli* and *Stenotrophomonas* (76–81%). The amino acid sequence diversity found in the TTPAs is likely a reflection of the adaptations that phages acquire in response to environmental changes and/or bacterial cell variability.

Previous studies have shown that some tail tubular proteins have hydrolytic activity^{14,15}. For this study, our goal was to determine how similar TTPAgp11 is to other proteins in terms of primary structure, since primary structures affect the biological function of proteins. The Clustal Omega tool²¹ was used to perform multiple protein sequence alignments (Fig. 1). For this analysis, two proteins previously reported as enzymes, TTPAgp31 and TTPAgp44^{14,15} were selected. In addition, TTPAgp45, originating from *Yersinia* phage phi80–18²², and T7_TTPgp11, which was classified as a tail tubular protein of *Enterobacteria* phage T7 were also chosen. Both of the structures and functions of these proteins have been well documented¹¹. Our sequence alignment analysis generated interesting results: TTPAgp11 was found to be most similar to T7_TTPgp11 of phage T7 (80% identity) and TTPAgp31 of phage KP32 (58% identity), whereas it showed far less similarity to TTPAgp44 of phage KP34 (21%) and TTPAgp45 of *Yersinia* phage phi80–18 (20%). It is particularly interesting that the last protein listed, TTPAgp45, showed the lowest similarity with TTPAgp11, although both proteins originated from *Yersinia* phages. If it is assumed that the function of a protein is determined by its primary structure, then it is expected that the biochemical features of TTPAgp11 should be more similar to those of T7_TTPgp11 (from T7 bacteriophage) than those of TTPAgp31. Notably, TTPAgp31 was 64% identical to T7_TTPgp11, but only the former has been reported to exhibit hydrolytic activity^{14,15}.

Our Phyre server analysis²³ showed that 92% and 93% of the amino acids of TTPAgp11 took on the same predicted structures as those of TTPAgp31 and T7_TTPgp11, respectively. The confidence level was 100% for each of these matches. Although the confidence level does not specifically speak to the accuracy of the model, this high level of confidence showed that TTPAgp11 adopted the overall folding and protein core of our models. Any difference in the 3D structure was likely due to deviation of the surface loops relative to those placed in the templates (PDB code: 5 mu4)¹⁴.

The I-Tasser server^{24,25} was used to further investigate the 3D structure of TTPAgp11. I-Tasser generated five models based on the comparison to the protein structures deposited in PDB (Fig. 2). Interestingly, the server used the structure of TTPAgp31 (PDB code: 5 mu4)¹⁴, not T7_TTPgp11 (PDB code: 3j4b)¹¹, as template, suggesting that the spatial structure was more similar to that of TTPAgp31 versus T7_TTPgp11. In this prediction, the α -helical and β -stranded elements matched well between TTPAgp11 and TTPAgp31, while some differences were seen in the loop regions. For example, the predicted loop containing amino acid residues number 88 to 96 of TTPAgp11 was a bit longer than the corresponding loop of TTPAgp31 (residues 86 to 91). The structural variations between the two models were largely related to the N- and C-termini of the polypeptide chains, but for different reasons. The N-terminal region of TTPAgp31 was nine amino acids longer than that of TTPAgp11, whereas the C-terminus of TTPAgp31 is shorter as in the structure determined by X-ray diffraction data the C-terminus could not be modeled as there was no electron density in this region.

Further analysis was performed using the HHPred tool²⁶ and it was predicted that TTPAgp11 contained a peptidoglycan hydrolase domain (probability, 68%). Based on our analyses of the primary structure, it was

TTPAgp11	MRSYEMNIETAEELSAVNDILASIGEPVSTLEGDANADVANARRVLNKNRQIQSRGWT	60
T7_TTPgp11	MRSYDMNVETAELSAVNDILASIGEPVSTLEGDANADAANARRILNKNRQIQSRGWT	60
TTPAgp31	MNMQDAYFGSAAELDAVNEMLAATIGESPVTTLDEGSDVADVARRILNRINRQIQSKGWA	60
TTPAgp44	-----MRELDAINLTLEALGESRVMDINT-SNPSAGLARSALARNRRLSTGFW	49
TTPAgp45	-----MITELNVVNSCLATIGEMPLVELQD-EHPMVAAARQNFEEARVAEVAVQWW	50
TTPAgp11	FNIEEGVTLPLDAFSGMIPFSSDYLSVMATSGQTQYINRGGYLYDRSAKTRDFPSGVQVN	120
T7_TTPgp11	FNIEEGITLLPDVYSNLIIVYSDDYLSLMSTSGQSIYVNRGGYVYDRTSQSDRFDGSGITVN	120
TTPAgp31	FNINESATLTPDVSTGLIPFRPAYLSILG----GQYVNRGGWVYDKSTGTDTFSGPITVT	116
TTPAgp44	FNVVE-REVTP-TADGFIKVPWNQLAVYDAGSDSKYGVDRGNLYDLMEQNQYFDSVVKLK	107
TTPAgp45	FNTDR-VTLQPANGEGFIYVPRDAVAVTPL-DRQDLGMRGRRLYNHQESTYVI GAPVRCV	108
TTPAgp11	LIRLREFDEMPECFRNYIVTKASRQFNNRFFGAPEVDGVLQEEQEAWRACFEYE-LDYG	179
T7_TTPgp11	IIRLRDYDEMPECFRYIIVTKASRQFNNRFFGAPEVEGVLQEEDEARRLCMEYE-MDYG	179
TTPAgp31	LITLQDYDEMPECFRQWIIVTKASRQFNSRFFGAEDVENSQAQEEEMEARMACNEYE-MDFG	175
TTPAgp44	IVLDLDFEDLPEHAAMWVANYTTAQVYLNLDLGGDSNYANYAQ-EAERYKSMVLRHEHLRNQ	166
TTPAgp45	VVRDIPFDDLPAQAQLLVKHACVLIQFQMNDADEAKSAKLEALYAKAYRT-LNAEHIRQV	167
TTPAgp11	NYNMLDGDFTSGLLNR-----	196
T7_TTPgp11	GYNMLDGDFTSGLLNR-----	196
TTPAgp31	QYNMLDGDAYVQGLIGR-----	192
TTPAgp44	RFST-----SKTRFARRIRRRAREMV-----	186
TTPAgp45	ALN-----GLLMPQIAQARMAVGGNRRIGNRIPVR	197

Figure 1. Sequence alignment of TTPAgp11, TTPAgp31 from the tail of bacteriophage KP32 (PDB code: 5 mu4), T7_TTPgp11 from bacteriophage T7 (PDB code: 3j4b), TTPAgp44 from bacteriophage KP34 (UniProt code: D1L2Y9) and TTPAgp45 from *Yersinia* phage phi80–18 (UniProt code: I7K3G0). Identities between TTPAgp11 and all other selected proteins are shown in red; those between TTPAgp11, T7_TTPgp11 and TTPAgp31 are shown in blue; those between TTPAgp11 and T7_TTPgp11 are shown in green, and those between TTPAgp11 and TTPAgp31 are gray-shadowed. For clarity, the identities between TTPAgp11 and TTPgp44 (21%) or TTPAgp45 (20%) are not shown. The catalytic motif proposed herein is shown in yellow.

hypothesized that TTPAgp11 could contain a domain in the α -helical region, at amino acid residues 130 to 177, which corresponded to the hydrolase domain of TTPAgp31¹⁴.

The catalytic motif of most glycosyl hydrolases contained two carboxylates in a D-X-E motif²⁷. This short sequence in two regions of TTPAgp31 was previously identified, and it was proposed that one of them is likely to play a crucial role in the enzymatic reaction catalyzed by this protein¹⁴. Here, two similar and corresponding motifs were found in TTPAgp11 (Fig. 1), but note that they are inverted to the E-X-D sequence (E126-X-D128, E155-X-D157). This suggested, with a high probability, that this motif could be responsible for binding as well as for catalysis of sugar hydrolysis. In our previous studies, in case of TTPAgp31 of the KP32 bacteriophage, it was proposed that D131 and/or D133 were the most likely catalytic amino acid residues. *In silico* analysis presented by Świętnicki & Brzozowska²⁸ confirmed that D133 participates in a stable substrate binding. D133 is positioned near a scissile bond, potentially making it a catalytic residue. Our initial *in silico* studies (data not shown) of the putative catalytic mechanism of TTPAgp11 showed that a good catalytic candidate is the E126 residue. In the previously mentioned studies²⁸, the authors also confirmed the conformational fit of the saccharide substrate in the catalytic pocket of the protein. Our analysis showed that the catalytic pocket in TTPAgp11 is structurally similar to the catalytic pocket in TTPAgp31. Therefore, it is thought that there is a great probability that, in TTPAgp31, and TTPAgp11, maltose spatially fits into the substrate binding pocket.

It was observed that TTPAgp11, just like TTPAgp31, possessed an additional antiparallel β -sheet carrying a lectin-like domain (residues 56 to 103), which was responsible for sugar binding. The amino acid sequences of these domains in TTPAgp11 and in TTPAgp31 showed 62% identity (Fig. 1).

The gene for TTPAgp11 was cloned into the pMCSG9 vector²⁹. Afterwards, the overexpressed protein was purified and analyzed with SDS-PAGE (Fig. 3), electrospray ionization mass spectrometry (ESI-MS), as well as circular dichroism spectroscopy (CD). The obtained ESI-MS spectrum indicated that the apparent molecular weight of this protein was 22.69 kDa (Fig. 4). The results of the CD spectrum analysis (Fig. 5) suggested that TTPAgp11 adopts a mixed α/β secondary structure with a large proportion of helical structure.

To test the hydrolytic activity of TTPAgp11, disaccharides substrates such as α -lactose, β -lactose, trehalose, melibiose, cellobiose, maltose and saccharose were first used. Then a chromogenic substrate, Red-starch, which was previously found to be hydrolyzed by TTPAgp31, was also used¹⁴. The hydrolytic activity tests for disaccharides were performed according to the same procedure described in¹⁵. Our observation was that only maltose was hydrolyzed, however, our results were not unambiguous. To verify the hydrolytic activity of TTPAgp11 towards maltose, gas chromatography of two samples after hydrolysis of this disaccharide by TTPAgp11 was performed. The retention times of standards of α -D-glucose and β -D-glucose were assigned as 8.599 min and 9.161 min, respectively. The GC analysis of the tested samples revealed the presence of α - and β -glucose in both samples (Fig. 6) which were absent in negative control. On the base of the sugars peak sizes it was estimated that around 15–18% of maltose was hydrolyzed. This result undoubtedly indicated hydrolytic activity of TTPAgp11 towards maltose.

In order to confirm the activity, another substrate was used. The relevant enzymatic activity was examined both in solution and on filter paper. Our results confirmed that TTPAgp11 hydrolyzed the glycosidic bonds α – 1,

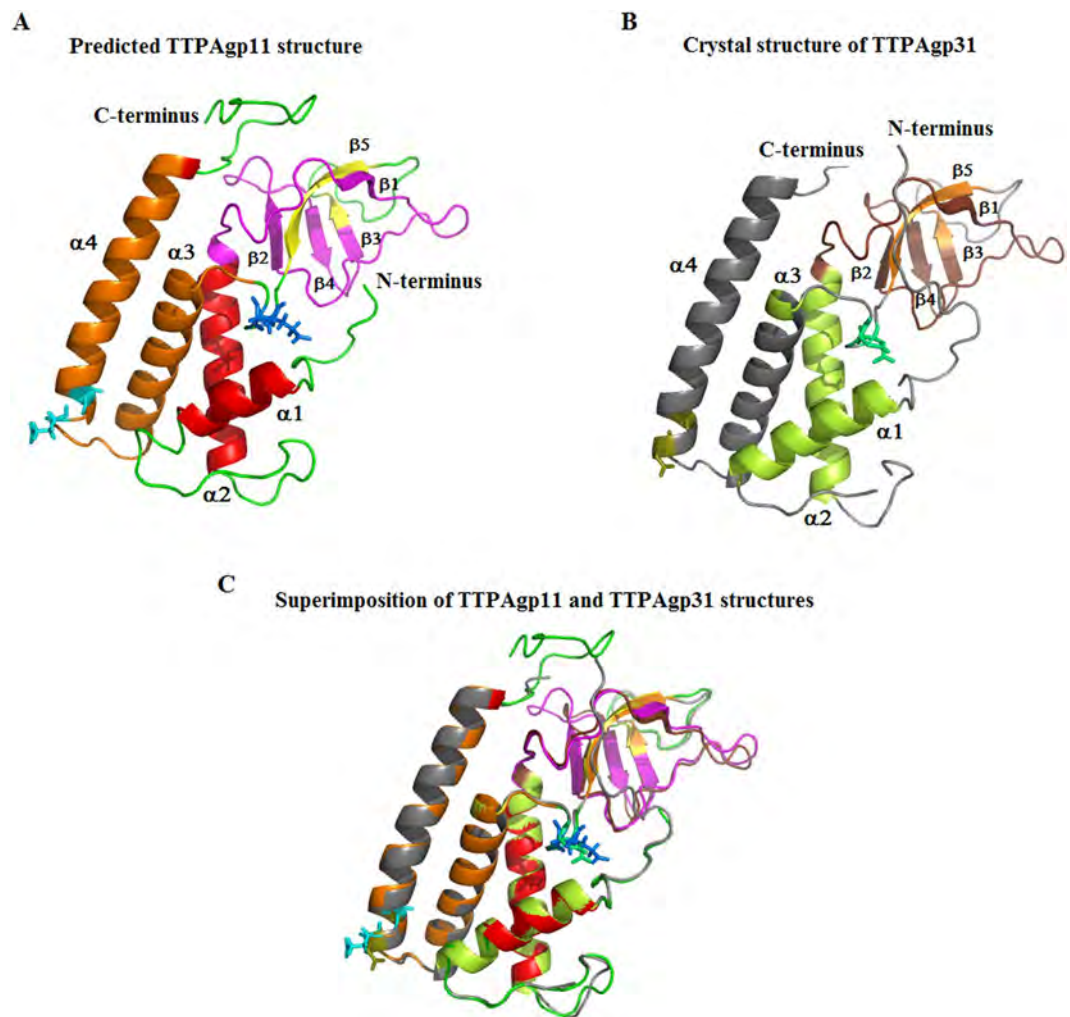


Figure 2. The predicted 3D structural model of TTPAgp11 generated using the I-Tasser server^{24,25}. (A) The presented structure was the best calculated model, and was obtained using the crystal structure of TTPAgp31 (B) deposited in PDB (PDB code: 5 mu4). The residues of two potential catalytic triads are shown as blue and cyan sticks in TTPAgp11 (E126-X-D128 or E155-X-D157) along with green and deepolive sticks (D122-X-D124 or D151-X-E153) in TTPAgp31. A predicted hydrolase domain is shown in orange in TTPAgp11 and in grey in TTPAgp31. A potential lectin-like domain is shown in magenta and brown in TTPAgp11 and TTPAgp31 respectively. (C) Predicted structure of TTPAgp11 superimposed onto crystal structure of the TTPAgp31 monomer.

4 in Red-starch to release glucose molecules from the non-reducing end of the substrate (Fig. 7). These findings indicated that, similar to TTPAgp31^{14,15}, TTPAgp11 exhibited α -glucosidase-like activity and can be considered to act as an α -1, 4-glucosidase (EC 3.2.1.20). This further suggested that TTPAgp11 could serve as an antibiofilm factor that acts against the EPS of pathogenic bacteria. In general, this class of enzymes is very diverse in their substrate specificity, optimal reaction temperature and transglucosylation activities³⁰.

To quantify the stability of TTPAgp11, thermal unfolding and protein aggregation studies were performed using NanoDSF and backreflection technology. The stabilization test was performed in the presence of different monosaccharides (that we considered as potential binders). Moreover, the assessment of protein stability was key to determining the importance for protein's basic features characterization. Therefore our goal was to characterize the thermal and colloidal stability of TTPAgp11 along with establishing the optimal conditions for large-scale production and long-term storage. The obtained results (Fig. 8, Table 1) showed that TTPAgp11 exhibited one unfolding event at ~ 53 °C. This suggested that TTPAgp11 is a single-domain protein that can unfold at this temperature. It was further observed that different sugars have different effects on protein stability. The addition, some of them even caused aggregation of TTPAgp11. Aggregation was induced by maltose and N-acetylglucosamine (GlcNAc) but not melibiose, galactose, glucose, lactose or N-acetylgalactosamine (GalNAc). The presence of GlcNAc, but not the other sugars, was associated with an inversion of the unfolding profile. This suggested that this sugar may have a dramatic influence on the conformation of TTPAgp11. These sugar moieties were chosen because they are components of *Y. enterocolitica* lipopolysaccharide. The outer core of the lipopolysaccharide contains two GalNAc moieties, glucose and galactose¹⁸. Maltose and melibiose were also tested as potential substrates for TTPAgp11.

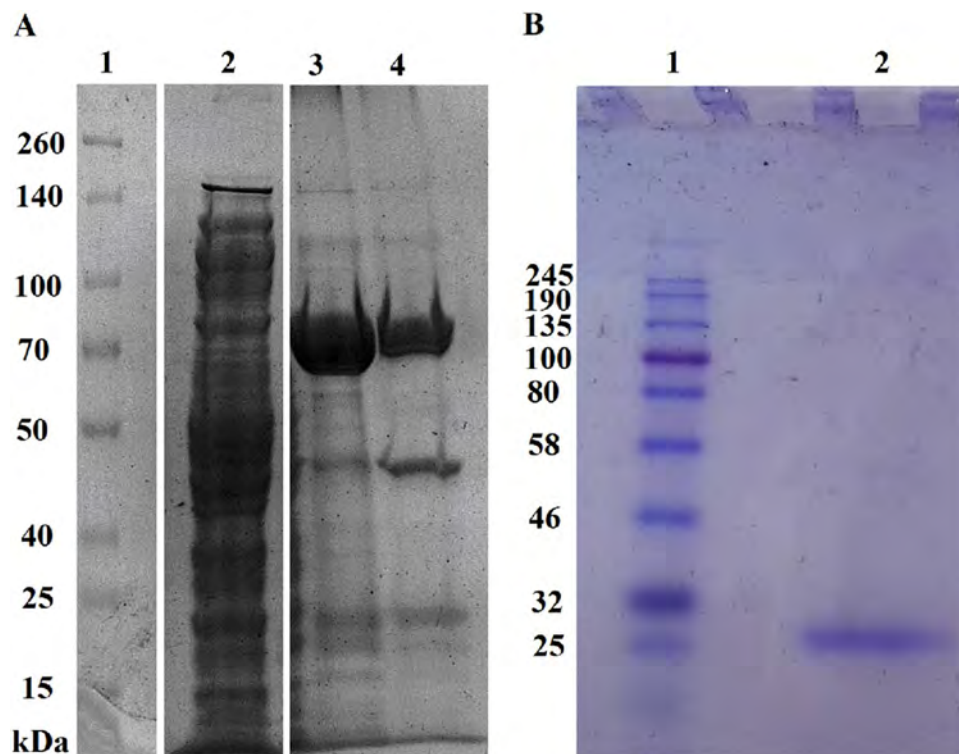


Figure 3. Analysis of TTPAgp11 from *Yersinia* phage phiYeO3-12 using 10% (A) and 12.5% (B) SDS-PAGE. (A) The lanes are as follows: (1) Spectra Multicolor Broad Range Protein Ladder (Bio-Rad); (2) crude extract; (3) pooled fractions after the first round of Ni²⁺ -affinity chromatography; (4) protein solution after TEV protease cleavage. (B) The lanes are as follows: (1) Color Prestained Protein Standard, Broad Range (Bio-Rad); (2) pure target protein - TTPAgp11. The original pictures of the gels are attached in a Supplementary Data File.

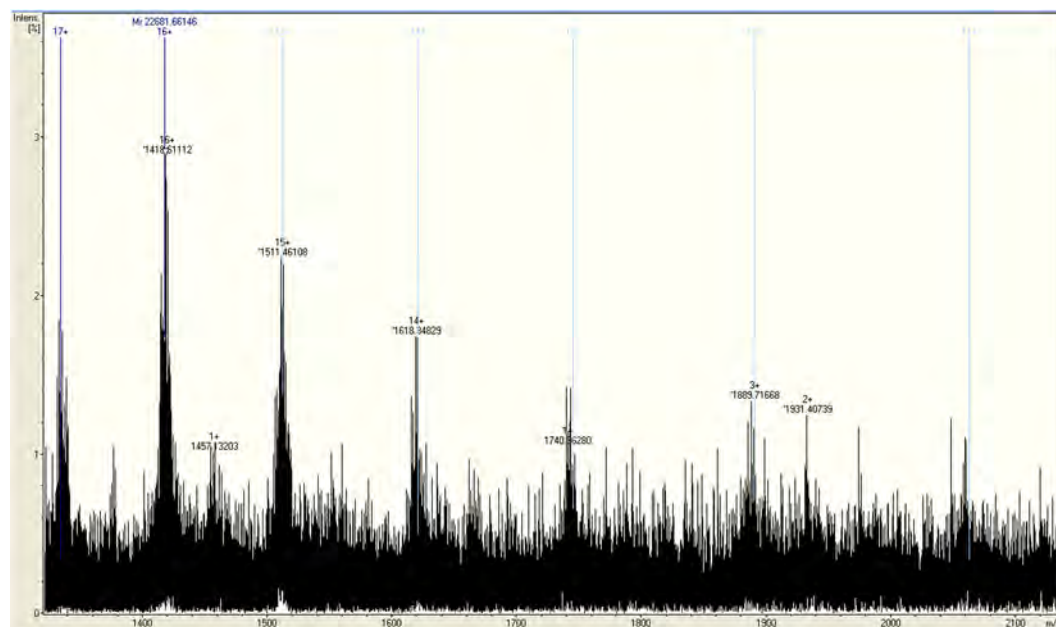


Figure 4. The ESI-MS spectrum of TTPAgp11.

Glycoside hydrolases catalyze the cleavage of the glycosidic bonds using two commonly found mechanisms of action with either net retention or inversion of anomeric configuration. Hydrolysis, with net retention of anomeric configuration, is achieved via a two steps double displacement mechanism involving a covalent glycosyl-enzyme intermediate. The reaction occurs with acid/base and nucleophilic assistance provided by two

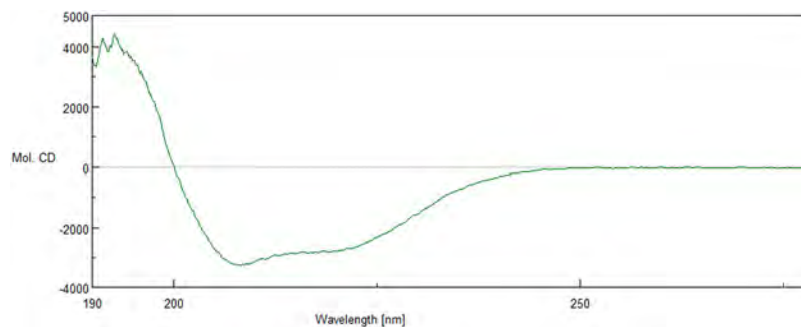


Figure 5. The CD spectrum of TTPAgp11 in units of molar ellipticity [$\text{deg cm}^2 \text{dmol}^{-1}$].

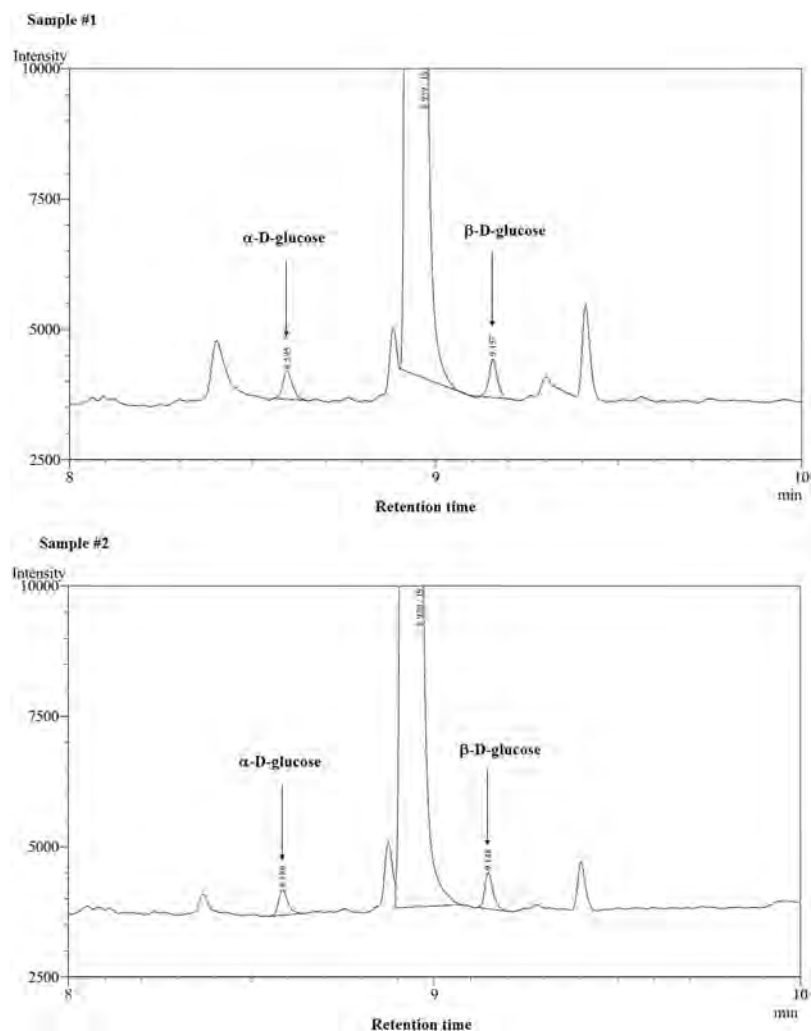


Figure 6. The GC analysis of hydrolytic activity of TTPAgp11 towards maltose. The experiment was performed twice, so two chromatograms are presented. The products of enzymatic reaction are released glucose molecules indicated by the arrows (peaks of α -D-glucose and β -D-glucose). The retention times of the standards of α -D-glucose and β -D-glucose were assigned as 8.599 min and 9.161 min, respectively.

amino acid side chains, typically glutamate or aspartate. The second type of hydrolysis - the inversion of anomeric configuration - is achieved via a one step, single-displacement mechanism. The reaction typically occurs with general acid and general base assistance from two amino acid chains usually glutamic or aspartic acids. The distance between both catalytic residues (general acid and general base) is very well conserved, being 5.5 Å and 10.5 Å in retaining and inverting glycoside hydrolases respectively³¹. In the presented studies, two catalytic pairs have been proposed in TTPAgp11. The first pair is E126 - X - D128 and the second one is E155 - X - D157. In the

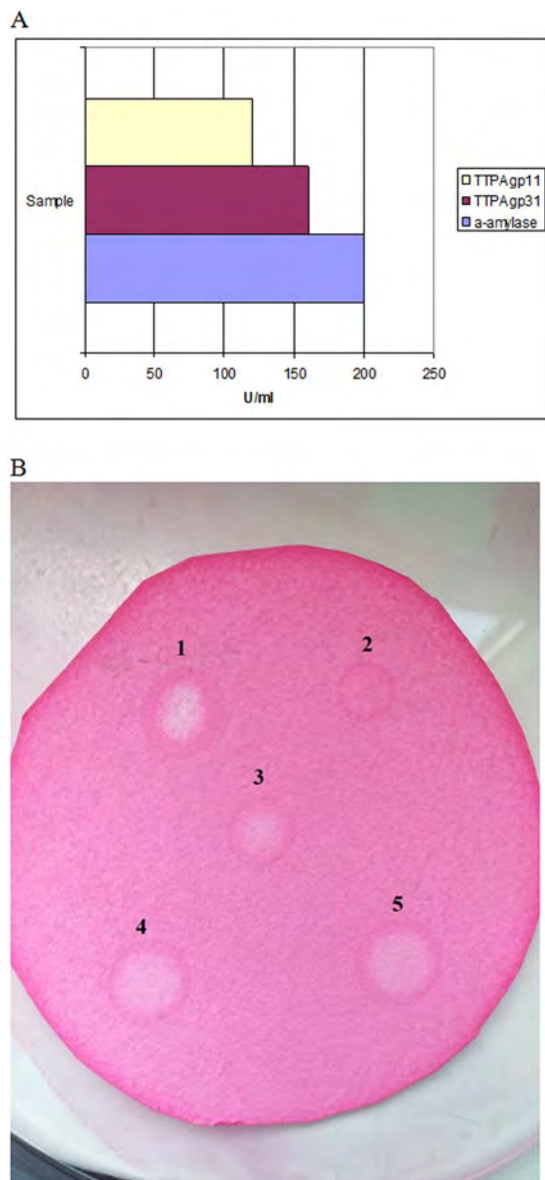


Figure 7. Hydrolytic activity of TTPAgp11 towards Red-starch. *Bacillus subtilis* α-amylase was used as a positive control. **(A)** The α-amylase activity against Red-starch was calculated as described in the provided protocol (Megazyme). All results are presented as averages of results from three independent replicates in three parallel trials. **(B)** Assays were also performed using Red-starch-saturated filter paper treated with: (1) *B. subtilis* α-amylase, as a positive control; (2) phosphate buffer pH 6.8, as a negative control; (3) TTPAgp31 as a positive control; and (4, 5) TTPAgp11.

first pair, the distance between acid/base residues is about 6 Å which indicates the retention configuration, while the distance between acid/base in the second pair is over 10 Å indicating the inversion configuration. However, the distances between the putative catalytic residues have been estimated using the predicted 3D structure of the protein (The I-Tasser server). Therefore, the more likely catalytic mechanism of action is retention due to the known proximity of acid/base residues. Moreover, the same configuration was also proposed by Świetnicki & Brzozowska²⁸ for the homologous protein possessing the same enzymatic activity. Their conclusion was supported by the results of MD simulations in that work.

Further structural and functional studies are needed to validate the enzymatic function of one or both of these motifs.

Conclusion

TTPAs were long thought to function solely as structural proteins of bacteriophage tails. However recent studies have shown that they can display hydrolytic activity towards saccharides, and thus act as dual-function proteins. Although some studies have examined the biochemical features of TTPAs, these macromolecules are still not well understood, especially in the context of their enzymatic activity and 3D structure. Here, the goal was to identify

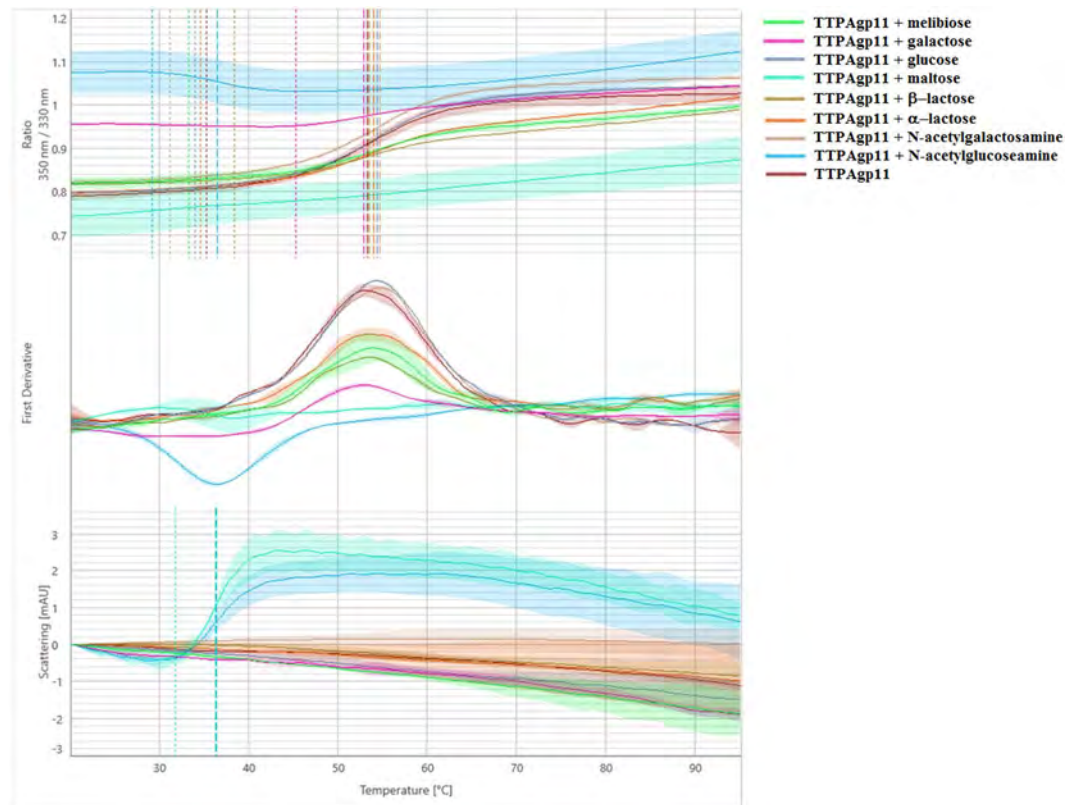


Figure 8. Melting analysis of TTPAgp11 with various sugars. The fluorescence ratio (350 nm/330 nm) is shown in the top panel, the first derivative is shown in the middle panel and scattering is shown in the bottom panel. Thermal unfolding and aggregation onsets (indicated as the vertical dotted lines) as well as unfolding and aggregation transitions (indicated as the vertical dashed lines) are indicated by vertical lines in the graph.

Sample	Tm		
	Fluorescence ratio 350 nm/330 nm [°C]	First derivative [°C]	Scattering [°C]
TTPAgp11/melibiose	53.93	32.22	NA
TTPAgp11/galactose	52.85	45.25	NA
TTPAgp11/glucose	54.31	33.97	NA
TTPAgp11/maltose	NA	NA	36.26
TTPAgp11/ ^β -lactose	53.44	38.34	NA
TTPAgp11/ ^α -lactose	53.93	34.55	NA
TTPAgp11/N-acetylgalactosamine	54.68	31.19	NA
TTPAgp11/N-acetylglucosamine	36.41	29.11	36.38
TTPAgp11	53.27	35.23	NA

Table 1. TTPAgp11 thermal unfolding and aggregation transitions midpoints (Tm [°C]) obtained from a melting scan of TTPAgp11 mixed with sugars, showing the fluorescence ratio (350 nm/330 nm), the first derivative onset and scattering. Each sample contained 0.25 mg/ml of TTPAgp11 and 0.15 mg/ml of the sugar.

a previously uncharacterized TTPA of a new origin, as well as assess its biological function and structure. This research has practical relevance because TTPAs could potentially serve as antimicrobial factors for the treatment of antibiotic-resistant bacterial infection.

This study identified a novel TTPA from *Yersinia* phage phiYeO3-12, and designated it as TTPAgp11. This protein was purified for the first time via a ligation-independent cloning procedure in which gene overexpression was performed in competent *E. coli* cells and the protein was purified by nickel-affinity chromatography. Our hydrolytic activity tests showed that TTPAgp11 can hydrolyze maltose and Red-starch, indicating that it functions as an α -glucosidase. The protein was further classified as part of the α -1, 4-glucosidase family (EC 3.2.1.20), which also contains the previously reported TTPAgp31. These findings indicated that TTPAgp11 could act against the EPS of biofilm-forming pathogenic bacteria. Our thermal unfolding measurements suggested that TTPAgp11 is

a single-domain protein that may unfold around 53 °C and it appeared to aggregate in the presence of maltose or N-acetylglucosamine. Our primary structure analysis revealed high levels of amino acid sequence identity between TTPAgp11 and TTPAgp31 from phage KP32 (58%) and T7_TTPgp11 from phage T7 (80%). Potential catalytic motifs consisting of E-X-D residues were identified and it was proposed that one or both could play a role in the enzymatic reaction. The 3D structure analysis predicted that the spatial structure of TTPAgp11 would be more similar to that of TTPAgp31 than that of T7_TTPgp11. It was speculated that this may reflect the presence of additional structural elements in the crystal structure of TTPAgp31 compared to that of T7_TTPgp11. These elements are likely to be responsible for the hydrolytic activity of the dual-function TTPAs. Clearly, more structural studies are needed to fully clarify the structure and enzymatic reaction center(s) of TTPAgp11.

Materials and Methods

Gene cloning and protein overexpression, purification and analysis. The genome of *Yersinia* phage phiYeO3-12 can be found in GenBank under accession number AJ251805. The gp11, the gene that encodes the TTPA studied, was amplified by polymerase chain reaction (PCR) using the following primers: TTPAgp11_FW, TACTTCCAATCCAATGCCATGCGCTCTTATGAGATGAAC and TTPAgp11_RV, TTATCCACTTCCAATGTTATTAGCGGTTAAGTAGACCAGAGG. The PCR reaction was performed using genomic DNA as a template (100 ng). The DNA was isolated from phage lysate using viral DNA extraction kit (Biocompare). The reaction comprised of a 3-min denaturation at 94 °C followed by 35 cycles of 30 s at 94 °C, 45 s at 55–62 °C, and 1 min at 72 °C, and a final extension for 10 min 72 °C. The PCR product was cloned into the pMCSG9 vector using a T4 DNA polymerase, according to the previously described procedure for ligation-independent cloning²⁹. The construct was transformed into *E. coli* DH5 α cells using the heat-shock method, and confirmed by sequencing. The plasmid was transformed into competent *E. coli* BL21 Star(DE3) cells, which were grown to an OD₆₀₀ of 0.9 and then induced overnight with 0.25 mM IPTG at 19 °C in Luria-Bertani (LB) medium with ampicillin. The cells were harvested by centrifugation and sonicated 10 times (30-sec pulses separated by 15-sec breaks) in 50 mM Tris/HCl buffer, pH 8.0, containing 300 mM NaCl, 5% glycerol and 5 mM β -mercaptoethanol (buffer A). The cell disruption by sonication was performed on ice (30 s/30 s cycles, 40% amplitude) using a UP200S ultrasonic disintegrator (Dr. Hielscher GmbH). The cell debris was pelleted and the supernatant was loaded onto a nickel-immobilized affinity column equilibrated with buffer A. Unbound proteins were washed with this buffer and TTPAgp11 was eluted with buffer A containing 250 mM imidazole. The tail protein A fractions were pooled together, diluted to a final concentration of 125 mM imidazole and then mixed with TEV protease at a 1:100 ratio with the addition of 0.5 mM EDTA. This mixture was incubated for 2 hr at 30 °C and then overnight at 4 °C to allow cleavage of the MBP-TTPAgp11 fusion protein. The digested proteins were then precipitated with ammonium sulfate (95% saturation) and pelleted to remove the EDTA and imidazole. The proteins were dissolved in buffer A and applied onto a second nickel column. The TTPAgp11-containing His₆-tag-MBP flow through fractions were collected, precipitated by ammonium sulfate (95% saturation) and analyzed by 10% and 12.5% SDS-PAGE³², electrospray ionization mass spectrometry (ESI-MS) and circular dichroism (CD) spectroscopy. Both, for ESI-MS and CD analysis, the protein was dissolved in water with a final concentration of 24 μ M. ESI-MS, and measurement was carried out using Bruker micrOTOF-Q mass spectrometer (Bruker Daltonics). The solution of methanol with a 0.1% acetic acid was added to the protein sample in the ratio of 1:1 (v/v) and the mixture was mechanically applied to “electrospray” at 120 ml/min and a 4.5 kV needle voltage. After measurement the molecular mass of the protein was determined using Bruker software³³. CD measurements were performed four times in a 1 mm cuvette for a one TTPAgp11 sample in the wavelength range of 350 to 190 nm using J-1000 Circular Dichroism spectrophotometer (Jasco Inc.). The results of the CD measurement were given both in units of absorbance and ellipticity. To interpret the obtained spectra, data with the units of ellipticity (millidegrees, mdeg) were used. These results were exported to a spreadsheet and then the values of millidegrees were extracted for 51 points in the wavelength range of 190–240 nm (every 1 nm) and these data were interpreted using K2D3 method that allows to estimate the secondary structure of the protein³⁴. The protein concentration was determined by the BCA method³⁵.

Measurement of hydrolytic activity towards a chromogenic substrate and maltose. As a chromogenic substrate for TTPAgp11, 2% Red-starch (Megazyme) was used and prepared in 5 mM KCl. TTPAgp11 (100 μ l of 47 μ M) was mixed with the substrate (25 μ l), water was added to a final volume of 1.5 ml and the sample was incubated at 37 °C for 20 hr. The reaction was stopped using 96% ethanol (2.5 ml). As positive controls, 50 μ l of 50 μ M α -amylase from *Bacillus subtilis* (Sigma-Aldrich) was used and TTPAgp31 prepared as described for TTPAgp11. After the reactions, the solution's absorbance at 510 nm were measured. The activity calculations were performed based on a protocol for assay of α -amylase using Red-starch from Megazyme.

The assay was also performed on grade 1 Whatman filter paper, as previously described by Martin *et al.*³⁶. Briefly, the filter paper was saturated with 0.6% Red-starch in 50 mM phosphate buffer, pH 6.8, dried, and then spotted with 15 μ l of TTPAgp11 at 47 μ M, 15 μ l of TTPAgp31 at 47 μ M (positive control), 10 μ l of α -amylase from *Bacillus subtilis* (positive control) and 10 μ l of 50 mM phosphate buffer (negative control). The filter paper was incubated at 37 °C for 20 min. Positive results were characterized by a marked color change from dark pink to white.

The hydrolytic activity of TTPAgp11 has been tested towards the disaccharide substrate such as maltose. TTPAgp11 (20 μ l of 47 μ M) was mixed with the maltose solution (100 μ l, 1 mg/ml) prepared in PBS buffer. This buffer was added to a final volume of 0.2 ml and the sample was incubated at room temperature for 24 hr. The test was performed in duplicate. The negative control did not contain the protein. Then the sample was lyophilized and chromatographed in GC-2010 Plus system (Shimadzu, Japan). The sample was derivatized as follows: the dry mass was dissolved in 500 μ l of anhydrous pyridine and reacted with 500 μ l MTBSTFA (N-tert-Butyldimethylsilyl-N-methyltrifluoroacetamide) as well as 50 μ l TMCS (Trimethylchlorosilane). Then

the sample was vortexed for 30 seconds, heated at 60 °C for 2 hr and the sample was loaded in split 1:30 on the column Zebron, ZB-5 (L = 30 m, ID = 25 mm, FT = 0.25 µm, Phenomenex, USA). The starting temperature of chromatographic analysis was 100 °C and the temperature increased 15 °C per minute until the temperature reached 320 °C. The column was standardized with sugars, such as α- and β-glucose what allows to indicate the presence such sugars in the analyzed sample.

Protein folding and aggregation tests using nanoDSF technology (NanoTemper). The thermal unfolding experiments were carried out using nanoDSF technology (NanoTemper). The Prometheus instrument was used to monitor unfolding-related tryptophan and tyrosine fluorescence at the emission wavelengths of 330 nm and 350 nm, respectively. The thermal unfolding transition midpoint, T_m [°C], which is the inflection point of the unfolding curve and the point at which half of the protein population is unfolded, was determined automatically from the derivative of the curve using the PR.ThermControl software³⁷. This method circumvents the need to subjectively determine the baseline levels and allows for the determination of a single or multiple unfolding transition midpoints. The PR.ThermControl software automatically calculates the onset of unfolding using the transition midpoint and the slope of the unfolding signal.

The protein aggregation tests were carried out using backreflection technology (NanoTemper) and the Prometheus instrument. It emits near-UV light at a wavelength that is scattered by aggregated proteins, so that that only non-scattered light reaches the detector. The reduction in backreflected light is taken as a direct measure of aggregation, and is plotted as mAU (Attenuation Units) against temperature.

Protein unfolding and aggregation were detected simultaneously. Briefly, 10 µl of TTPAgp11 (0.25 mg/ml, 23.5 µM) was placed in the capillary either alone or mixed with one of the following sugars (0.15 mg/ml): glucose, galactose, maltose, *N*-acetylglucosamine, *N*-acetylgalactosamine, melibiose, α-lactose and β-lactose. The samples were subjected to a temperature ramp of 2 °C/min from 20 °C to 95 °C and fluorescence was constantly monitored. Data were analyzed with the PR.ThermControl and PR.StabilityAnalysis software packages³⁷. The profile of the 350/330 nm fluorescence ratio was used to calculate T_m [°C]. The first derivative profile showed the onset temperatures of unfolding and aggregation as well as the occurrence of conformational changes in the protein sample. The scattering curve was used to determine the onset temperature of protein aggregation.

Received: 21 May 2019; Accepted: 20 February 2020;

Published online: 06 March 2020

References

- Donlan, R. M. Preventing biofilms of clinically relevant organisms using bacteriophage. *Trends in Microbiology*. **17**, 66–72 (2009).
- Nwodo, U. U., Green, E. & Okoh, A. I. Bacterial Exopolysaccharides: Functionality and Prospects. *Int. J. Mol. Sci.* **13**, 14002–14015 (2012).
- Kassa, T. & Chhibber, S. Thermal treatment of the bacteriophage lysate of *Klebsiella pneumoniae* B5055 as a step for the purification of capsular depolymerase enzyme. *J. Virol. Methods*. **179**, 135–141 (2012).
- Sulakvelidze, A., Alavidze, Z. & Morris, J. G. Bacteriophage Therapy. *Antimicrob. Agents Chemother.* **45**, 649–659 (2001).
- Kucharewicz-Krukowska, A. & Slopek, S. Immunogenetic effect of bacteriophage in patients subjected to phage therapy. *Arch. Immunol. Et Therapie Experimental*. **35**, 553–561 (1987).
- Cislo, M., Dabrowski, M., Weber-Dabrowska, B., Woyton, A., Bacteriophage treatment of suppurative skin infections. *Arch. Immunol. Et Therapie Experimental*. **35**, 175–183 (1987).
- Slopek, S., Weber-Dabrowska, B., Dabrowski, M. & Kucharewicz-Krukowska, A. Results of bacteriophage treatment of suppurative bacterial infections in the years 1981–1986. *Arch. Immunol. Et Therapie Experimental*. **35**, 569–583 (1987).
- Weber-Dabrowska, B., Mulczyk, M. & Gorski, A. Bacteriophage Therapy of bacterial infections: an update of our institute's experience. *Arch. Immunol. Et Therapie Experimental*. **48**, 547–551 (2000).
- Navarro, F. & Muniesa, M. Phages in the Human Body. *Front. Microbiol.* **8**, 566–572 (2017).
- Rohde, C., Wittmann, J. & Kutter, E. Bacteriophages: A Therapy Concept against Multi-Drug-Resistant Bacteria. *Surg. Infect (Larchmt)* **19**, 737–744 (2018).
- Cuervo, A. *et al.* Structural characterization of the bacteriophage t7 tail machinery. *J. Biol. Chem.* **288**, 26290–26299 (2013).
- Hardies, S. C. *et al.* Identification of structural and morphogenesis genes of Pseudoalteromonas phage phiRIO-1 and placement within the evolutionary history of Podoviridae. *Virology* **489**, 116–127 (2016).
- Pires, D. P., Oliveira, H., Melo, L. D. R., Sillankorova, S. & Azeredo, J. Bacteriophage-encoded depolymerases: their diversity and biotechnological application. *Appl. Microbiol Biotechnol.* **100**, 2141–2151 (2016).
- Pyra, A. *et al.* Tail tubular protein A: a dual – function tail protein of *Klebsiella pneumoniae* bacteriophage KP32. *Scient. Rep.* **7**, 2223 (2017).
- Brzozowska, E. *et al.* Hydrolytic activity determination of Tail Tubular Protein A of *Klebsiella pneumoniae* bacteriophages towards saccharide substrates. *Scient. Rep.* **7**, 18048 (2017).
- Pajunen, M. I., Kiljunen, S. J., Söderholm, M. E. & Skurnik, M. Complete Genomic Sequence of the Lytic Bacteriophage phiYeO3-12 of *Yersinia enterocolitica* Serotype O:3. *J. Bacteriol.* **183**, 1928–1937 (2001).
- Cover, T. L. & Aber, R. C. *Yersinia enterocolitica*. *N. Engl. J. Med.* **321**, 16–24 (1989).
- Pinta, E. *et al.* Characterization of the six glycosyltransferase involved in the biosynthesis of *Yersinia enterocolitica* serotype O:3 lipopolysaccharide outer core. *J. Biol. Chem.* **283**, 28333–28342 (2010).
- UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **46**, 2699 (2018).
- Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
- Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
- Zhang, L. & Skurnik, M. Isolation of an R¹M+ mutant of *Yersinia enterocolitica* serotype O:8 and its application in construction of rough mutants utilizing mini-Tn5 derivatives and lipopolysaccharide-specific phage. *J. Bacteriol.* **176**, 1756–1760 (1994).
- Kelly, L. A. & Sternberg, M. J. E. Subject Categories: Bioinformatics In silico modeling structural analysis Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols* **4**, 363–371 (2009).
- Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5**, 725–738 (2010).
- Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40 (2008).

26. Soding, J., Biegerd, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, 243–248 (2005).
27. Moak, M. & Molineux, I. J. Peptidoglycan hydrolytic activities associated with bacteriophage virions. *Mol. Microbiol.* **51**, 1169–1183 (2004).
28. Swietnicki, W. & Brzozowska, E. In silico analysis of bacteriophage tail tubular proteins suggests a putative sugar binding site and a catalytic mechanism. *J. Mol. Graph. Model* **92**, 8 (2019).
29. Eschenfeldt, W. H., Stols, L., Sanville Millard, C., Joachimiak, A. & Donnelly, M. I. A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. *Methods Mol. Biol.* **498**, 105–115 (2009).
30. Nashiru, O., Koh, S., Lee, S. Y. & Lee, D. S. Novel α -Glucosidase from Extreme Thermophile *Thermus caldophilus* GK24. *Biochem. Mol. Biol.* **34**, 347–354 (2001).
31. Gloster, T. M., Torkenburg, J. P., Potts, J. R., Henrissat, B. & Davis, G. J. Divergence of catalytic mechanism within a glycosidase family provides insight into evolution of carbohydrate metabolism by human gut flora. *Chem. Biol.* **15**, 1058–1067 (2008).
32. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
33. Bruker Compass DataAnalysis ver. 4.0 (2011).
34. Louis-Juene, C., Andrade-Navarro, M. A. & Perez-Iratxeta, C. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins* **80**, 374–381 (2012).
35. Smith, P. K. *et al.* Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150**, 76–85 (1985).
36. Martin, N. C., Clayson, N. J. & Scrimger, D. G. The sensitivity and specificity of Red-Starch paper for the detection of saliva. *Scient. Tech.* **46**, 97–105 (2006).
37. NanoTemper, PR. ThermControl & PR. StabilityAnalysis software (ver. 2018).

Acknowledgements

We would like to thank Prof. Michael Skurnik for providing the phage, Prof. Piotr Chmielewski for providing access to the CD spectrophotometer, Dr. Mariusz Dziadas for providing access to GC and result analysis and the NanoTemper Company for allowing us to perform protein folding and aggregation measurements. This research was supported by the National Science Center (Poland), grant no.2017/26/E/NZ1/00249.

Author contributions

Conceived the experiments and analyzed the results: A.P. and E.B. Conducted the experiments: N.U., K.F. and K.T. Prepared article for a publication: A.P. and E.B. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-61145-5>.

Correspondence and requests for materials should be addressed to A.P. or E.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

Biochemical features of the novel Tail Tubular Protein A of *Yersinia* phage phiYeO3-12

Anna Pyra^{1*}, Natalia Urbańska^{2,3}, Karolina Filik³, Katherine Tyrlik³ & Ewa Brzozowska^{3**}

¹ University of Wrocław, Faculty of Chemistry, 14 F. Joliot-Curie St, Wrocław, 50383, Poland

² University of Wrocław, Faculty of Biological Sciences, Institute of Experimental Biology, 6 Kanonia St, Wrocław, 50328, Poland

³ Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, 12 R. Weigl St, Wrocław, 53114, Poland

* anna.pyra@chem.uni.wroc.pl

** ewa.brzozowska@hirszfeld.pl

BLAST amino acid sequence analysis [20] showing similarity of TTPAgp11 to other TTPs.

Tail Protein (TP)	Bacteria	% of amino acid sequence identity
Tail Tubular Protein A (TTPA)	<i>Yersinia</i>	72-100
	<i>Citrobacter</i>	99-100
	<i>E. coli</i>	68-100
	<i>Enterobacter</i>	99
	<i>Enterobacteria</i>	74-99
	<i>Salmonella</i>	76-99
	<i>Serratia</i> and <i>Leclercia</i>	98
	<i>Pectobacterium</i> , <i>Erwinia</i> and <i>Kluyvera</i>	72-75
	<i>Dickeya</i>	68-70
	<i>Pseudomonas</i> , <i>Pseudobacterium</i> and <i>Morganella</i>	62-64
Tail Fiber Protein B (TFPB)	<i>Yersinia</i> , <i>E. coli</i> and <i>Stenotrophomonas</i>	76-81

[20] Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402 (1997).

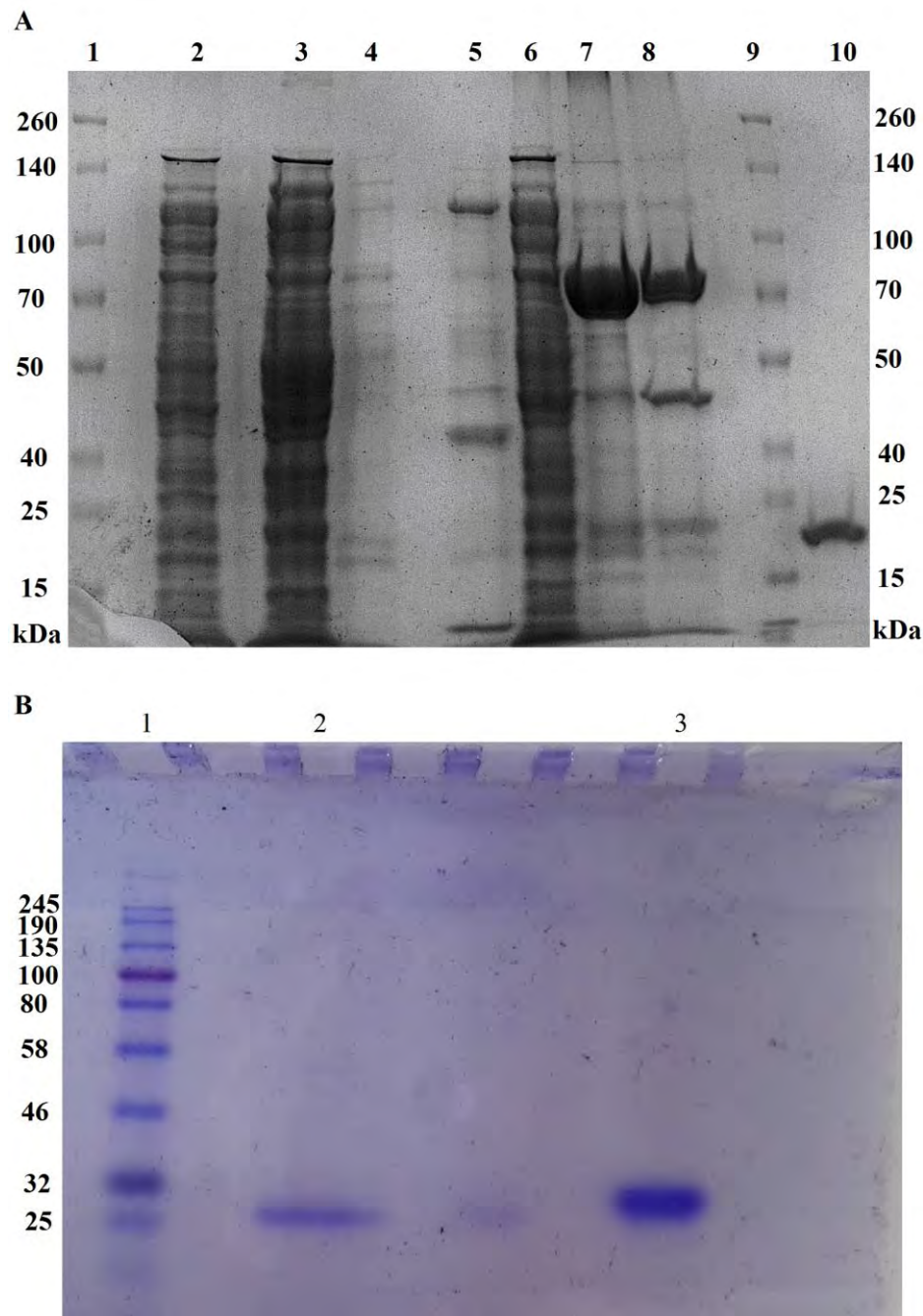


Figure 3. Analysis of TTPAgp11 from *Yersinia* phage phiYeO3-12 using 10% (A) and 12.5% (B) SDS-PAGE. (A) The lanes are as follows: (1) Spectra Multicolor Broad Range Protein Ladder (Bio-Rad); (2) the pellet and (3, 4) the supernatant after sonication and centrifugation of *E. coli* BL21 Star(DE3) cells (lane 4 - 15 times less sample volume); (5) the fraction eluted with 20 mM imidazole and (6) the flow through fraction after the first round of nickel-immobilized affinity column; (7) MBP-TTPAgp11 fusion protein after purification by the first round of Ni²⁺-affinity chromatography; (8) MBP-TTPAgp11 fusion protein after purification by the first round of Ni²⁺-affinity chromatography and TEV protease cleavage; (9) Spectra Multicolor Broad Range Protein Ladder (Bio-Rad); (10) TEV protease. (B) The lanes are as follows: (1) Color Prestained Protein Standard, Broad Range (Bio-Rad); (2) TTPAgp11 after purification by the second round of Ni²⁺-affinity chromatography; (3) TEV protease.



The Podovirus ϕ 80-18 Targets the Pathogenic American Biotype 1B Strains of *Yersinia enterocolitica*

OPEN ACCESS

Edited by:

William Michael McShan,
The University of Oklahoma Health
Sciences Center, United States

Reviewed by:

Konstantin Anatolievich
Miroshnikov,
Institute of Bioorganic Chemistry
(RAS), Russia
Inmaculada Garcia-Heredia,
University of Alicante, Spain

*Correspondence:

Karolina Filik
karolina.filik@hirsfeld.pl
Bożena Szermer-Olearnik
bozena.szermer-olearnik@hirsfeld.pl
Mikael Skurnik
mikael.skurnik@helsinki.fi

† Present address:

Lotta J. Happonen,
Department of Clinical Sciences Lund,
Infection Medicine, Lund University,
Lund, Sweden

Specialty section:

This article was submitted to
Virology,
a section of the journal
Frontiers in Microbiology

Received: 14 April 2020

Accepted: 27 May 2020

Published: 19 June 2020

Citation:

Filik K, Szermer-Olearnik B,
Wernecki M, Happonen LJ,
Pajunen MI, Nawaz A, Qasim MS,
Jun JW, Mattinen L, Skurnik M and
Brzozowska E (2020) The Podovirus
 ϕ 80-18 Targets the Pathogenic
American Biotype 1B Strains
of *Yersinia enterocolitica*.
Front. Microbiol. 11:1356.
doi: 10.3389/fmicb.2020.01356

Karolina Filik^{1*}, Bożena Szermer-Olearnik^{1*}, Maciej Wernecki², Lotta J. Happonen[†],
Maria I. Pajunen⁴, Ayesha Nawaz⁴, Muhammad Suleman Qasim⁵, Jin Woo Jun⁶,
Laura Mattinen⁴, Mikael Skurnik^{4,7*} and Ewa Brzozowska¹

¹ Hirsfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Wrocław, Poland, ² Department of Microbiology, Institute of Genetics and Microbiology, Faculty of Biological Sciences, University of Wrocław, Wrocław, Poland, ³ Department of Biosciences, Institute of Biotechnology, University of Helsinki, Helsinki, Finland, ⁴ Research Programme Unit Immunobiology, Department of Bacteriology and Immunology, Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland, ⁵ Molecular and Integrative Biosciences Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland, ⁶ Department of Aquaculture, The Korea National College of Agriculture and Fisheries, Jeonju, South Korea, ⁷ Division of Clinical Microbiology, Helsinki University Hospital, HUSLAB, Helsinki, Finland

We report here the complete genome sequence and characterization of *Yersinia* bacteriophage vB_YenP_ ϕ 80-18. ϕ 80-18 was isolated in 1991 using a *Y. enterocolitica* serotype O:8 strain 8081 as a host from a sewage sample in Turku, Finland, and based on its morphological and genomic features is classified as a podovirus. The genome is 42 kb in size and has 325 bp direct terminal repeats characteristic for podoviruses. The genome contains 57 predicted genes, all encoded in the forward strand, of which 29 showed no similarity to any known genes. Phage particle proteome analysis identified altogether 24 phage particle-associated proteins (PPAPs) including those identified as structural proteins such as major capsid, scaffolding and tail component proteins. In addition, also the DNA helicase, DNA ligase, DNA polymerase, 5'-exonuclease, and the lytic glycosylase proteins were identified as PPAPs, suggesting that they might be injected together with the phage genome into the host cell to facilitate the take-over of the host metabolism. The phage-encoded RNA-polymerase and DNA-primase were not among the PPAPs. Promoter search predicted the presence of four phage and eleven host RNA polymerase –specific promoters in the genome, suggesting that early transcription of the phage is host RNA-polymerase dependent and that the phage RNA polymerase takes over later. The phage tolerates pH values between 2 and 12, and is stable at 50°C but is inactivated at 60°C. It grows slowly with a 50 min latent period and has apparently a low burst size. Electron microscopy revealed that the phage has a head diameter of about 60 nm, and a short tail of 20 nm. Whole-genome phylogenetic analysis confirmed that ϕ 80-18 belongs to the *Autographivirinae* subfamily of the *Podoviridae* family, that it is 93.2% identical to *Yersinia* phage fHe-Yen3-01. Host range analysis showed that ϕ 80-18 can infect in addition to *Y. enterocolitica* serotype O:8 strains

also strains of serotypes O:4, O:4,32, O:20 and O:21, the latter ones representing similar to *Y. enterocolitica* serotype O:8, the American pathogenic biotype 1B strains. In conclusion, the phage φ80-18 is a promising candidate for the biocontrol of the American biotype 1B *Y. enterocolitica*.

Keywords: bacteriophage, *Yersinia enterocolitica*, phage biocontrol, phylogenetic, podovirus, proteome, genome

INTRODUCTION

Yersinia enterocolitica is a gram-negative bacterium that belongs to the *Enterobacteriaceae* family. It is a human enteropathogen (Thomson et al., 2006). *Y. enterocolitica* strains are classified into six biogroups based on phenotypic characteristics, and to 57-O serogroups based mainly on difference in the lipopolysaccharide (LPS) O-antigen structures (Fàbrega and Vila, 2012). Yersiniosis is a zoonotic foodborne infection of animals and humans caused by pathogenic strains of *Y. enterocolitica* that mainly belong to bioserotypes 1B/O:8, 2/O:5,27, 2/O:9, 3/O:3, and 4/O:3. The strains of bioserotype 4/O:3 cause the majority of the infections in Europe, Japan, Canada and the United States (Bottone, 1999; Fredriksson-Ahomaa et al., 2006). In Europe and China, the most prevalent are the *Y. enterocolitica* serogroups O:3 and O:9, whereas in United States the predominant serogroup is O:8 (Sabina et al., 2011). In recent years, *Y. enterocolitica* infections have also spread between continents through human travel and transportation of pigs, and this has resulted in the higher occurrence of *Y. enterocolitica* O:8 infections in Europe (Rastawicki et al., 2009) and also in Japan (Ichinohe et al., 1991). The main reservoir of pathogenic *Y. enterocolitica* are pigs, and infections are caused especially by consumption of raw or undercooked pork, but dogs have also been implicated as a potentially significant source in rural communities. In addition, direct and indirect contact with feces from contaminated livestock can also lead to infection (Sabina et al., 2011; Wang et al., 2011). In humans, the infection is usually localized to the gastrointestinal track and the bacteria may also cause mesenteric lymphadenitis. The most common symptoms of the infection are acute enteritis, fever, vomiting, inflammatory and watery diarrhea (Fàbrega and Vila, 2012).

The pathogenic *Y. enterocolitica* strains are characterized by the presence of virulence factors encoded by the genes located either in the chromosome or in the 70 kb virulence plasmid, pYV. The most important virulence factors are the LPS, the adhesins/invasins (Inv, YadA, Ail), the flagella, the type 3 secretion system (T3SS) and the enterotoxin Yst. These virulence properties help *Y. enterocolitica* bacteria to survive and colonize the human host and cause the symptomatic infections (Fàbrega and Vila, 2012).

The invention of antibiotics has certainly saved millions of lives, but currently the rapid acquisition of antibiotic resistance by bacteria has become a major epidemiological problem. According to World Health Organization, antibiotic resistance is one of the biggest threats to global health and food security. Therefore, we have to take into use alternative approaches to combat the drug-resistant bacteria.

Bacteriophages are the most abundant organisms on Earth. The total number of phages has been estimated to be around 10^{31} particles (Hendrix, 2002; Leon-Velarde et al., 2019). The therapeutic potential of phages was recognized in the early twentieth century; specifically in the 1930s and 1940s (Fischetti et al., 2006). Lytic phages have been used as therapeutic and prophylactic agent in controlling bacterial infections (Jun et al., 2018). Phage therapy is becoming an interesting being an alternative to antibiotic therapy. Since Alexander Fleming's discovery of antibiotics, the overuse of antibiotics has imposed selective pressures on microorganisms. This has caused microorganisms to develop resistance mechanisms such as enzymatic mechanisms of drug modification, enhanced efflux pump expression, mutated drug target, etc. (Aleksun and Levy, 2007). During the last 10 years, phage research has become very popular especially scientist are now focused on the genome and evolution of bacterial viruses as well as horizontal gene transfer (HTG) which is the main cause of diversity (Leon-Velarde et al., 2019).

Bacteriophages characterized by exceptional specificity and selectivity, can only infect and reproduce inside the host bacteria (Ventola, 2015; Zhao et al., 2019). This specificity makes them an excellent tool to fight the pathogenic bacteria and it also provides a number of possibilities for diagnostic applications. Therefore, learning about the biology of bacterial viruses is an important research topic. Bacteriophages have long been utilized as tools in bacterial genetics and systematics. Indeed, the first suspicions that the genus *Yersinia* belongs to the *Enterobacteriaceae* was made on the basis of common sensitivities to phages (Brubaker, 1972). Phages have also been used in epidemiological characterization and other studies on *Y. enterocolitica* strains (Nicolle et al., 1967; Baker and Farmer, 1982).

Several bacteriophages infecting *Y. enterocolitica* have been isolated and characterized in the Skurnik laboratory (Skurnik, 1999). By using different host strains for enrichment, phages with different specificities were obtained and several of them were shown to use different parts of the *Y. enterocolitica* LPS as receptor (Skurnik, 1999). Detailed characterizations of several bacteriophages have been published including the T3-related φYeO3-12 (Pajunen et al., 2000, 2001; Kiljunen et al., 2005b), the giant myovirus φR1-37 (Kiljunen et al., 2005a; Skurnik et al., 2012; Leskinen et al., 2016), and the T4-like myovirus φR1-RT (Leon-Velarde et al., 2016). Genetic and structural data showed that the surface receptors of phages φR1-37 and φYeO3-12 are the outer core (OC) hexasaccharide and the O-antigen of the *Y. enterocolitica* O:3 LPS, respectively (Al-Hendy et al., 1991; Skurnik, 1999; Pajunen et al., 2000; Pinta et al., 2010; Skurnik et al., 2012), and that phage φR1-RT uses both OmpF and LPS inner core as receptors (Leon-Velarde et al., 2016).

In this paper, we describe the characterization of the *Y. enterocolitica* serotype O:8 specific phage φ80-18 that was isolated in 1991 and used as a tool in genetic selections (Zhang and Skurnik, 1994). We have earlier shown that purified O:8 LPS inhibits the phage and that the phage can infect an *E. coli* strain expressing the *Y. enterocolitica* serotype O:8 O-antigen, confirming that the O:8 O-antigen is the host receptor of φ80-18 (Zhang and Skurnik, 1994; Zhang et al., 1997). However, a detailed characterization of the phage has been missing and is presented here.

MATERIALS AND METHODS

Bacteriophage, Bacteria and Culture Media

The bacterial strains used in this work are described in **Supplementary Table S1**. Isolation of the phage φ80-18 has been described earlier (Zhang and Skurnik, 1994). Both the phage φ80-18 and its host strain *Y. enterocolitica* serotype O:8 strain 8081-c have been deposited to Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH – Leibniz – Institut DSMZ under catalog numbers DSMZ 23253 and DSMZ 23249, respectively. Bacteria and bacteriophages were grown in lysogeny broth (LB, Bertani, 2004) at room temperature (22–25°C RT) unless otherwise indicated.

Bacteriophage Propagation

Yersinia enterocolitica strain 8081-c was grown in LB for 16 h, and 0.1 ml of the culture added to 5 ml of LB. The bacteria were grown aerated at 28°C to exponential phase ($OD_{600} = 0.3–0.5$), 0.1 ml of a crude phage lysate (4.5×10^7 PFU/ml) was added, and the culture was then incubated overnight at 28°C with shaking. The obtained phage lysate was filter-sterilized using a 0.22 μm Millipore membrane. In addition, a bacteriophage propagation experiment at 4, 28, and 37°C was performed, according to this scheme.

Determination of Host Ranges and Efficiency of Plating

To evaluate the host range of the phage, the infectivity of the membrane-filtered phage lysate (10^8 PFU/ml) was tested on the bacterial strains listed in **Supplementary Table S1**, using either the drop-test or plaque formation assay on soft-agar embedded bacteria. The formation of lysis zone or individual plaques was determined after 24 h of incubation. For the efficiency of plating (EOP), the PFU measurements were determined using the double-layer agar method. The EOP was calculated as the ratio between the PFU of the test strain to that of the original host strain *Y. enterocolitica* serotype O:8 strain 8081 (Stor ID 1258, **Supplementary Table S1**). The EOP assays were performed in triplicate.

Genome Sequencing, Assembly and Annotation

Phage DNA was obtained from high-titer phage preparations as described earlier (Sambrook et al., 2001). Phage DNA was sequenced using the Illumina GAIx (Genome Analyzer) technology at the FIMM Sequencing unit (Helsinki, Finland). The sequence assembly was done with the NextGene¹ and Staden software packages (Staden et al., 2003). The Artemis genome-browsing and annotation tool (Rutherford et al., 2000) was used for genome annotation. The physical ends of the phage genome and the terminal repeats (approx. 200 bp) of φ80-18 could not be identified from the *de novo*-assembled genomic sequence. To carry out this we used the approach described in details previously (Salem and Skurnik, 2018). Briefly, a 500 bp PCR-amplified fragment of the *fliC* gene of *Y. enterocolitica* O:3 was ligated with phosphorylated phage genomic DNA. The ligation mix was then used as a template for PCR using a primer pair of which one primer was *fliC*-specific and the other primer one of the phage-specific primers predicted to be close to the physical ends of the phage genomes as described (Salem and Skurnik, 2018). The resulting PCR products were purified and sequenced using a *fliC*-specific nested primed located ca. 200 bp upstream of the ligation junction. The PSI-BLAST (Altschul et al., 1997) and HHPred (Söding et al., 2005) programs were used to identify homologous proteins. Genome identity analysis between different viruses was carried out using StretcherN at EBI (Li et al., 2015). The PHIRE search tool was used to identify phage-encoded RNA polymerase promoters (Lavigne et al., 2004). The sigma-70 specific bacterial promoters and rho-independent terminators were searched using the search tools BPROM and FindTerm, respectively (Solovyev and Salamov, 2011). The annotated genome sequence of phage φ80-18 has been deposited into the nucleotide sequence databases under the accession numbers HE956710 and NC_019911.2.

Proteomics

Phage particle proteomes were analyzed by liquid chromatography coupled with mass spectrometry (LC-MS/MS) at the Proteomics Unit, Institute of Biotechnology, University of Helsinki. The phage with a titer $>10^{10}$ pfu/mL was used for the analysis. Prior to digestion of proteins to peptides with trypsin, the proteins in the samples were reduced with tris (2-carboxyethyl) phosphine (TCEP) and alkylated with iodoacetamide. Tryptic peptide digests were purified by C18 reversed-phase chromatography columns (Varjosalo et al., 2013) and the mass spectrometry (MS) analysis was performed on an Orbitrap Elite Electron-Transfer Dissociation (ETD) mass spectrometer (Thermo Scientific, Waltham, MA, United States), using Xcalibur version 2.2, coupled to a Thermo Scientific nLC1000 nanoflow High Pressure Liquid Chromatography (HPLC) system. Peak extraction and subsequent protein identification were achieved using Proteome Discoverer 1.4 software (Thermo Scientific). Calibrated peak files were searched against all amino acid sequences of all six open reading frames

¹<http://www.softgenetics.com>

of φ80-18 by a SEQUEST search engine. Error tolerances on the precursor and fragment ions were ± 15 ppm and ± 0.8 Da, respectively. Hits with at least two identified tryptic peptides were regarded as true hits.

Electron Microscopy

The purified bacteriophage was applied to the surface of formvar carbon-coated copper grids and negatively stained with 2% uranyl acetate for 1 min. The excess of uranyl acetate was then removed from the grids using filter paper and the grids were allowed to air dry for 20 min (Ackermann, 2009). Preparations were visualized using a JEOL JEM-1200 EX 80 kV TEM. The dimensions of the bacteriophages were determined using RADIUS EM Imaging Software.

Thermal and pH Stability Tests

To determine the thermal stability of phage φ80-18, phage samples (4.5×10^7 PFU/ml) were incubated at 4, 25, 40, 50, 60, and 80°C for 2 h. Phage survival was determined from samples collected after 20, 40, 60, 80, 100, and 120 min incubation using the double-layer agar method (Chen et al., 2016; Zhao et al., 2019).

To determine the pH stability of phage φ80-18, 200 μl samples of the phage (4.5×10^7 PFU/ml) were incubated under various pH conditions (2, 3, 5, 6, 7, 8, 10, and 12) for 2 h at 28°C. Bacteriophage preparations were mixed with different pH solutions in the volume ratio 1:1. Phage titers in the tubes were determined using the double-layer agar plate method.

One-Step Growth Curve

Yersinia enterocolitica 8081-c bacteria, grown in 5 ml of LB to an OD_{600} of 0.5, were centrifuged at $12000 \times g$ for 15 min at 4°C, and resuspended in 5 ml of fresh LB medium. The bacteria were then infected with phage φ80-18 at a MOI of 0.01, and the phages were allowed to adsorb to the bacteria for 5 min at 28°C. To remove the unadsorbed phages the suspension was centrifuged at $14000 \times g$ for 1 min, the bacterial pellet washed twice with fresh LB, and finally resuspended to 5 ml of LB, followed by incubation at 28°C. 100 μl from the sample were withdrawn from the tube every 10 min and the phage titers assayed using the double-layer agar method. The experiment was repeated three times (Zhao et al., 2019).

Phylogenetics Analysis

The phylogeny of phage φ80-18 was determined using both the whole genome nucleotide and the RNA polymerase (RNAP) amino acid sequences for the analysis. The genomic sequences of representative *Autographivirinae* (taxid:542835) phages most closely related to φ80-18 (NC_019911.2) were identified using the BLASTN search. The genome-based phylogenetic tree was constructed using the VICTOR web service (Meier-Kolthoff and Göker, 2017) based on the Genome-BLAST Distance Phylogeny (GBDP) method (Meier-Kolthoff et al., 2013) and FastME software. This included 100 pseudo-bootstrap replicates and SPR post-processing (Lefort et al., 2015). The amino acid sequences of the most closely

related RNAP proteins were identified using the BLASTP search. The sequences were aligned with MAFFT v7.429 under the L-INS-i strategy (Katoh and Standley, 2013). The best-fit model for tree reconstruction (LG+F+I+G4, chosen according to BIC) was calculated with ModelFinder (Kalyaanamoorthy et al., 2017). The RNAP phylogenetic tree was inferred by maximum likelihood method with IQ-TREE v1.6.11, performing ultrafast bootstrap with 1000 replicates for calculating branch support (Hoang et al., 2018). The phylogenetic trees were visualized with FigTree (Rambaut, 2006) and tanglegram was constructed with Dendroscope (Huson et al., 2007).

RESULTS

Genome Analysis of φ80-18

Phage φ80-18 has a linear double-stranded DNA genome of 42,406 bp with the GC content of 47.64% that is close to that of *Y. enterocolitica* strain 8081 (47%) (Thomson et al., 2006). Altogether 57 genes were predicted from the sequence, all in the forward strand (Figure 1). No tRNA coding genes were found. The physical ends of the genome contain 325 bp direct repeats (Figure 1). While the function of altogether 29 predicted gene products showed no similarity to any known genes in the databases and remained therefore unassigned, similarity searches by BLASTP (Altschul et al., 1997) and HHPred (Söding et al., 2005) assigned a putative function to 17 gene products. The remaining 11 predicted gene products were identified as phage particle-associated proteins (PPAPs) in the phage particle proteome analysis (Figure 1 and Supplementary Table S2). Altogether 25 PPAPs were detected by LC-MS/MS analysis including those identified as structural proteins such as major capsid (Gp44), phage collar protein (Gp42), scaffolding protein (Gp43) and identified tail component proteins (Gp45, Gp46, and Gp50) as well as the DNA packaging proteins A and B (Gp52 and Gp53). Also the peptidoglycan penetrating lytic murein transglycosylase protein (Gp49) was identified as a PPAP. The catalytic domain of the 1259 residue Gp49 occupies 150 N-terminal residues, thus it is likely that the remaining protein functions as a tape measure protein to determine the length of the tail tube that is extended upon adsorption of the phage particle on host bacteria (Hu et al., 2013). In addition, also the DNA helicase (Gp20), DNA ligase (Gp25), DNA polymerase (Gp28), and 5'-exonuclease (Gp31) proteins were identified as PPAPs, suggesting that they might be injected together with the phage genome into the host cell to facilitate the take-over of the host metabolism. In contrast, the phage-encoded RNA-polymerase and DNA-primase were not PPAPs. For the 56 predicted genes, the initiation codon was ATG and only for the *g37* gene encoding the DNA-directed RNA polymerase, it was GTG.

Promoters

Using the PHIRE search tool, four 25 nt long phage promoters, designated P1 – P4 with a consensus sequence of -TGAT(T/a)(c/g)TCTACCCATATAG(c/t)AA(C/t)(A/t),

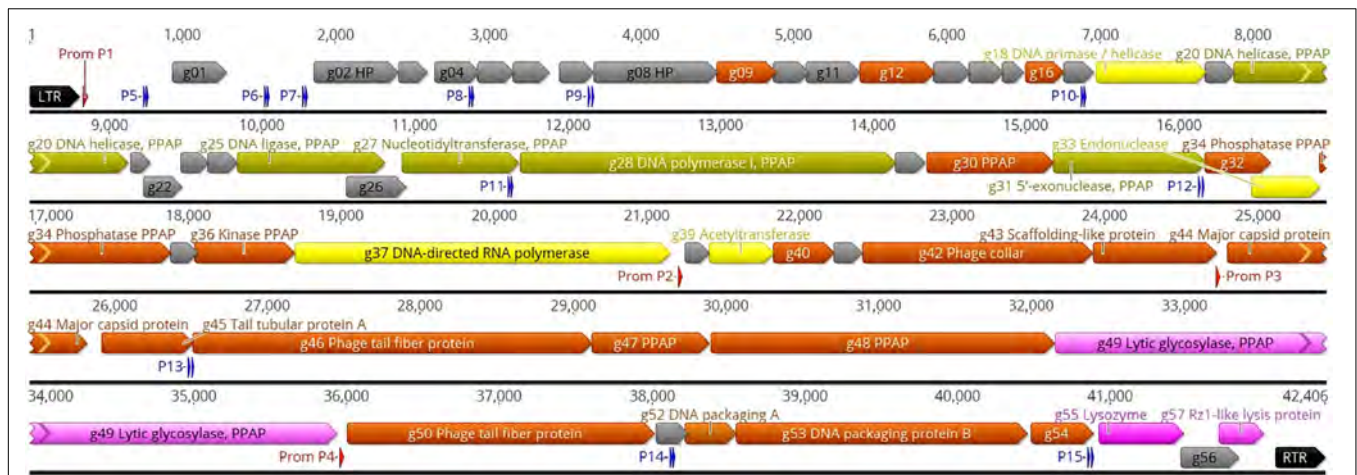


FIGURE 1 | Genomic map of phage φ80-18. The nucleotide sequence of the phage is represented by the black horizontal line, above which are indicated the left and right terminal repeats (LTR and RTR, respectively) as black arrows, the phage RNAP promoters as red arrows, and the sigma-70 host RNAP promoters as blue double arrows representing the -35 and -10 boxes. The promoters are numbered and detailed information of them is given in **Supplementary Table S3**. All the predicted genes are indicated by different-colored arrows and the gene names and predicted functions are indicated either inside or outside the arrows. The genes encoding hypothetical proteins (HP) are gray. The genes encoding phage particle-associated proteins (PPAP) are dirty green (for genes predicted to encode enzymes) and brown (for genes encoding predicted structural proteins). The genes predicted to encode phage particle-associated lytic glycosylase, and the lysozyme, are pink, and the genes predicted to encode DNA primase, endonuclease and RNA polymerase are yellow. The map was produced using the Geneious 10.2.6 (www.geneious.com).

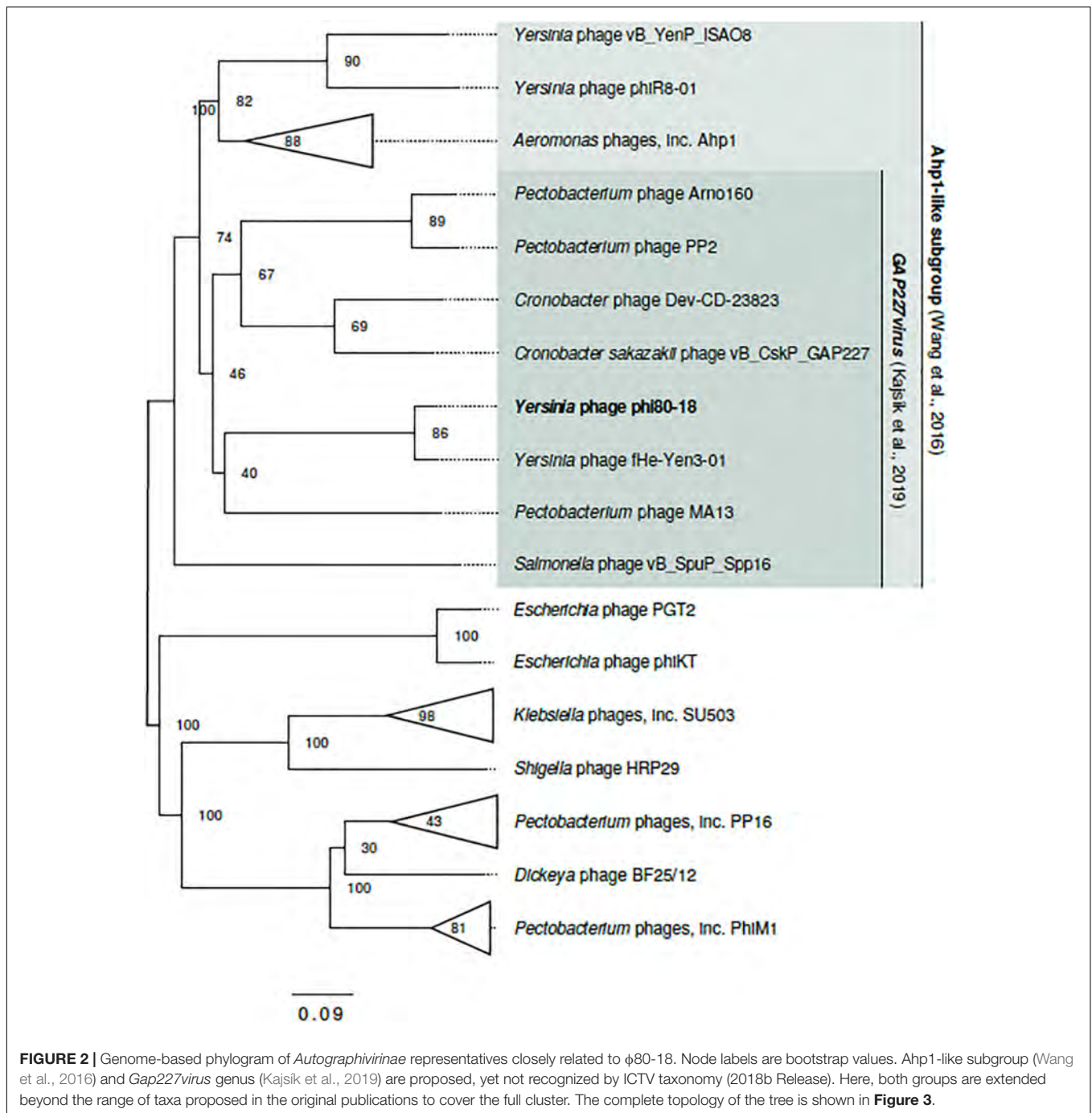
TABLE 1 | Overview of 15 phages most closely related to phage φ80-18.

Description	Genome size (Kb)	Query coverage	E-Value	Identity-%	Accession no.
Yersinia phage f80-18 complete genome	42.41	100,00%	0.0	100.00	NC_019911.2
Yersinia phage fHe-Yen3-01, complete genome	42.77	97,00%	0.0	98.31	KY318515.1
Pectobacterium phage MA13, partial genome	42.46	47,00%	0.0	75.54	MN509793.1
Cronobacter sakazakii phage vB_CskP_GAP227, complete genome	41.8	42,00%	0.0	73.48	KC107834.1
Cronobacter phage Dev-CD-23823 complete sequence	41.62	44,00%	0.0	71.70	LN878149.1
Pectobacterium phage PP2, complete genome	41.84	41,00%	0.0	72.05	KX756572.1
Yersinia phage vB_YenP_ISA08, complete genome	41.45	33,00%	0.0	71.80	KT184661.1
Pectobacterium phage Arno160, complete genome	41.38	41,00%	0.0	71.11	MK053931.1
Yersinia phage phiR8-01 complete genome	42.09	32,00%	0.0	71.97	HE956707.2
Aeromonas phage 25AhydR2PP, complete genome	42.70	31,00%	0.0	70.46	MH179473.2
Aeromonas phage ZPAH7B, complete genome	30.79	28,00%	0.0	69.93	MK330684.1
Aeromonas phage ZPAH7, complete genome	30.79	28,00%	0.0	69.93	MH992513.1
Aeromonas phage phiAS7, complete genome	41.57	33,00%	0.0	70.00	JN651747.1
Salmonella phage vB_SpuP_Spp16, complete genome	41.83	32,00%	0.0	69.17	MG878892.2
Aeromonas phage LAh5, complete genome	41.99	19,00%	6,00E-170	68.19	MK838111.1

typical for the *Autographiviridae*, were identified upstream the *g01*, *g38*, *g44*, and *g50* genes (**Figure 1** and **Supplementary Table S3**). These promoters likely regulate the expression of the phage genes during different phases of the infection cycle. In addition, using the BPROM search tool for bacterial sigma-70 type promoters we identified 11 bacterial promoter candidates, designated P5 – P15 (**Figure 1** and **Supplementary Table S3**). While the functionality of these promoters awaits experimental evidence, the highest scores were predicted to P5 located leftmost in the genome and very likely the first one to start transcription upon the injection of the phage genome into the bacterial cell. We

did not detect the phage encoded RNA polymerase in the phage particle so it has to be synthesized *de novo* before the phage promoters can be utilized, therefore, the presence of eleven sigma-70 type promoters scattered around the phage genome will allow transcription of the necessary phage genes, including *g37* encoding the phage RNA polymerase. Only one rho-independent terminator was detected by the FindTerm program, located inside the *g37* gene encoding the phage RNA polymerase.

While in general, the genomes of many podoviruses can be divided into three regions comprising early, middle and late genes for virus-host interactions, DNA metabolism and

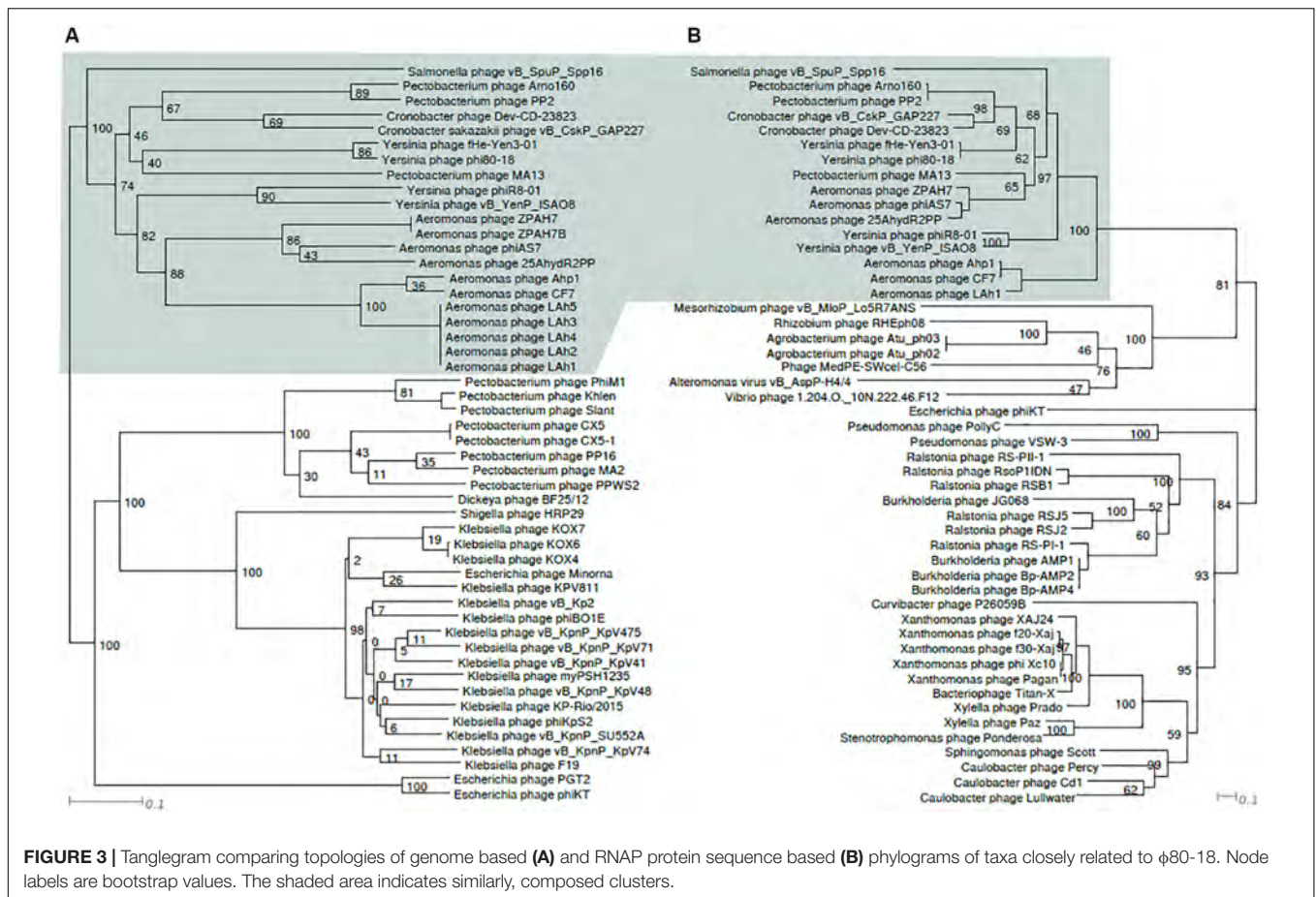


virion structure and assembly, respectively (Wang et al., 2016), this classification could not be directly applied to phage φ80-18 genome. While the phage RNAP promoters P2, P3, and P4 all seem to direct the transcription of the late genes, only phage RNAP promoter P1 remains for the first half of the genome. Therefore, the sigma-70 type promoters that are scattered around the genome might be involved in the transcription of the early and middle genes. To the latter ones based on functional predictions would belong the genes *g18–g37* (Figure 1). Then the predicted early genes are *g01–g17*, among

which are located also genes *g09*, *g12*, and *g16*, that encode PPAPs of unknown function.

Phylogenetic Analysis of φ80-18

The phage φ80-18 has been assigned to the *Podoviridae* family and the *Autographivirinae* subfamily like the model bacteriophage T7 or T3. BLASTN analysis revealed the highest sequence identity of 98 with 97% coverage (total identity of 93.2%, as determined by the EMBOSS stretcher alignment tool) to another *Yersinia* phage fHe-Yen3-01 that



we recently isolated in Finland (Jun et al., 2018), followed by *Pectobacterium* phage MA13 (75,5%) and *Cronobacter sakazakii* phage vB_CskP_GAP227 (73,5%) (Table 1). Whole-genome phylogenetic tree (Figure 2) places f80-18 in well-defined clade, which was defined as *Gap227virus* (Kajsik et al., 2019) or broader as Ahp1-like subgroup (Wang et al., 2016). This significant phylogenetic association is supported by a tree inferred using single marker, RNA polymerase (RNAP) (Figure 3). Genome alignment of selected phages from this clade (Figure 4) revealed that most of the similarities (local sequence identity ≥ 60%) come from the predicted genes coding for DNA helicase (*g20*), DNA polymerase (*g28*), phosphoesterase (*g34*), RNA polymerase (*g37*), phage collar (*g42*), major capsid protein (*g44*), lytic glycosylase (*g49*) and DNA packaging protein (*g53*). The major genomic diversity regions are located to the early gene and the tail fiber protein encoding gene (*g50*) that score the lowest local identity results (<60%). Notably, the genome of the nearly identical phage fHe-Yen3-01 differs from f80-18 mainly by the absence of the *g03* gene. On the other hand, the fHe-Yen3-01 possesses the gene *g29* that is not related to any φ80-18 genes. The only other major difference between the two phages resides in the N-terminal parts of their respective tail fiber proteins that are only 56% identical, explaining the distinct differences in the host ranges between these phages (Jun et al., 2018).

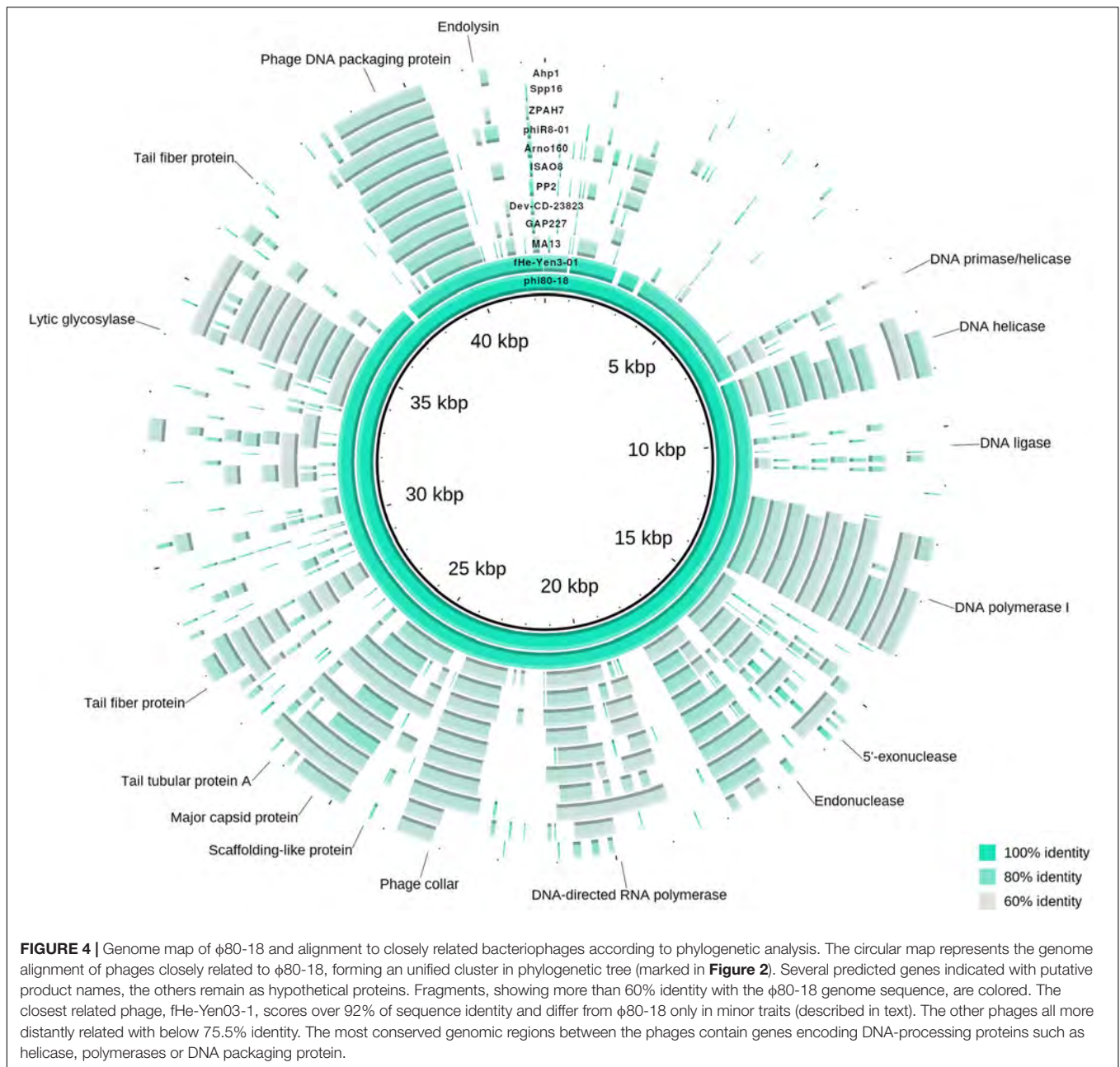
Characterization of Bacteriophage φ80-18 Growth and Stability

To characterize the biological properties of φ80-18 its one-step growth curve was determined for the host strain *Y. enterocolitica* 8081-c (Figure 5). The phage seems to grow rather slowly in this host showing an apparent 50 min latent period, and low burst size of 8-10 PFU per infected bacterium. Comparing bacteriophage φ80-18 propagation at different temperatures similar effectivity was achieved at temperatures 4°C (4,0 × 10⁸ PFU/ml) and 28°C (8,7 × 10⁷ PFU/ml) and much lower efficiency at 37°C (7,3 × 10⁴ PFU/ml).

In the thermostability test the phage was stable for 2 h between +4 and 50°C, and was slowly inactivated at 60°C the titer dropping one log every 20 min, however, at 80°C it was completely inactivated already after 20 min incubation (Figure 6). The phage tolerated well pH values between 2 and 12 and had apparently optimal pH of 7-8 (Figure 7).

Morphology of Bacteriophage φ80-18

Genome sequence and phylogenetic analysis showed that φ80-18 belongs to *Podoviridae* family of bacteriophages. Transmission Electron Microscopy confirmed that this phage has icosahedral capsid and short non-contractile tail with tail fibers. The dimensions of the phage are 59.0 ± 2.28 nm (n = 14) for capsid

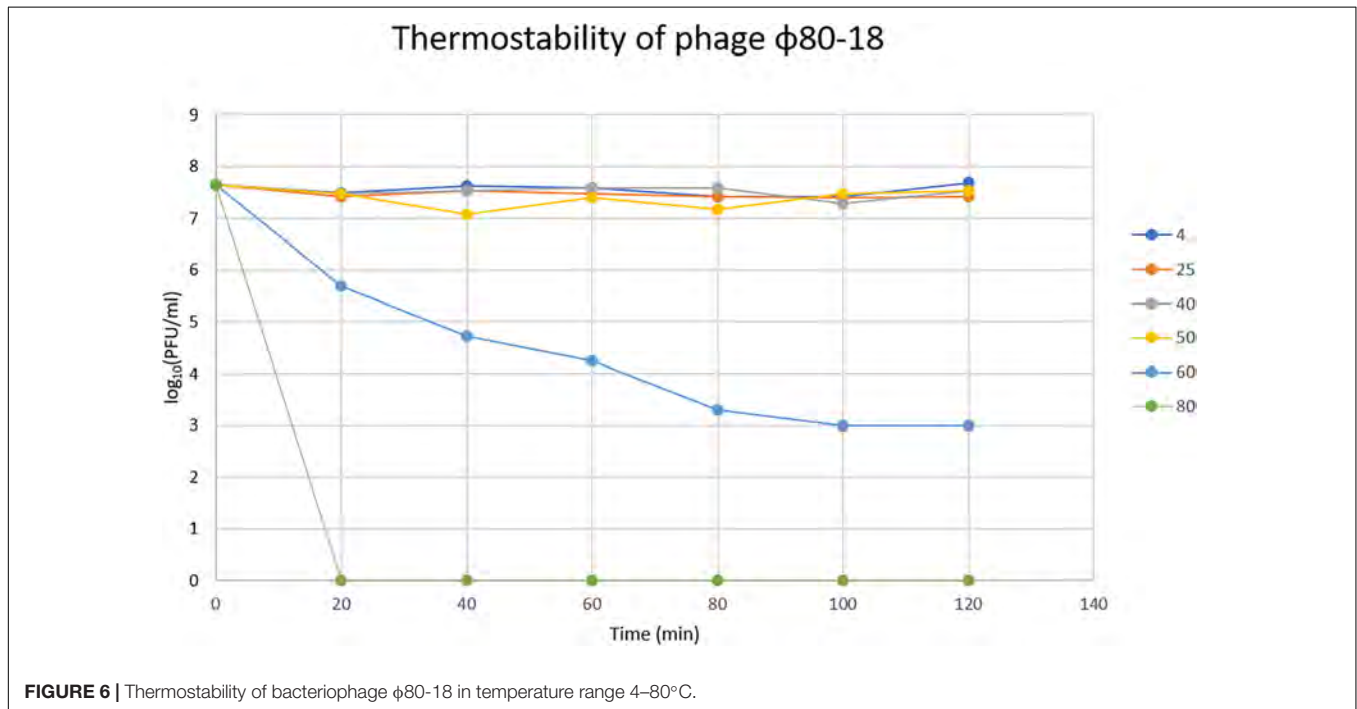
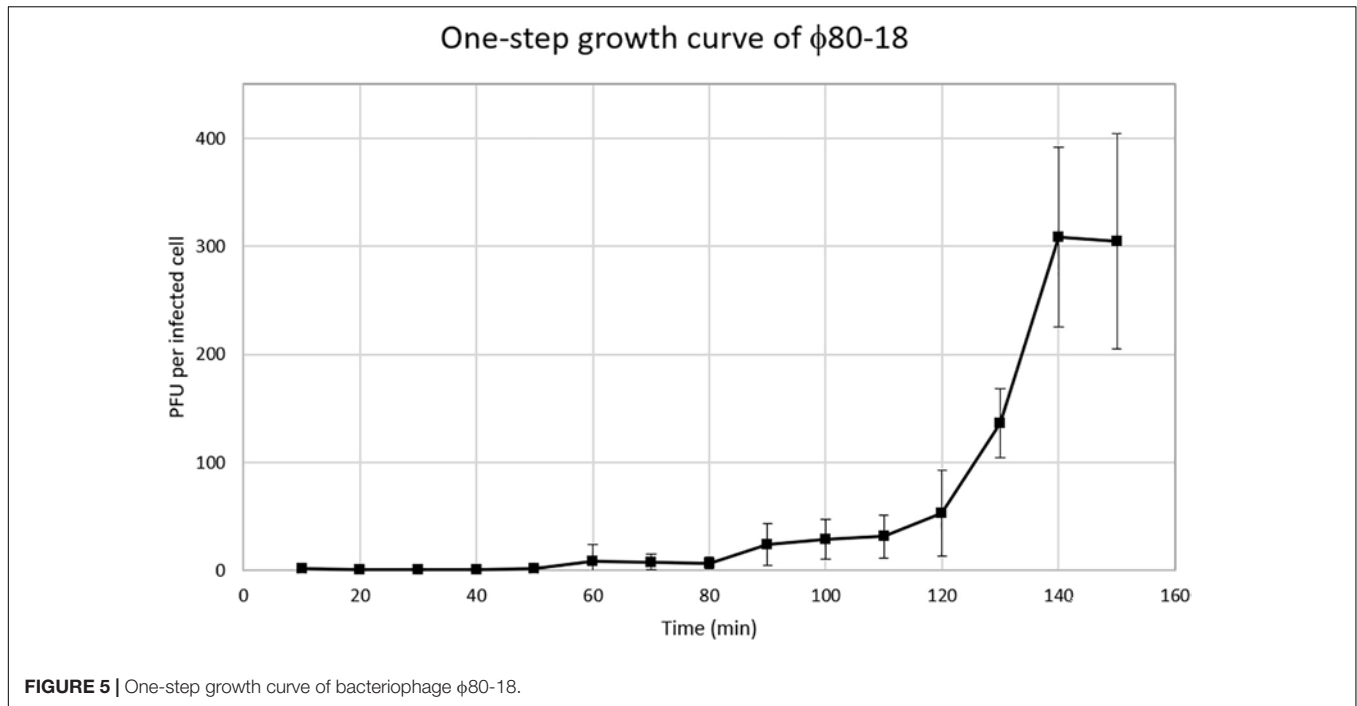


vertex to vertex, 59.0 ± 2.9 nm ($n = 14$) for capsid face to face and 18.3 ± 1.44 nm long tail (**Figure 8**).

Host Range

The host range of phage φ80-18 was tested using 115 *Yersinia* strains representing *Y. aleksiciae*, *Y. bercovieri*, *Y. enterocolitica*, *Y. frederiksenii*, *Y. intermedia*, *Y. kristensenii*, *Y. mollaretii*, *Y. nurmii*, *Y. pekkannenii*, *Y. ruckeri*, and *Y. pseudotuberculosis* (**Supplementary Table S1**). Bacteriophage φ80-18 was able to infect 16 strains. The host range analysis showed that φ80-18 can infect in addition to *Y. enterocolitica* serotype O:8 strains also strains of serotypes O:4, O:4,32, O:20 and O:21, the latter ones representing similar to serotype O:8 the American pathogenic

Y. enterocolitica serotypes. In addition, also strains of the non-pathogenic serotype O:7,8 and two of the bioserotype 1A/O:5 strains were infected by φ80-18 (**Supplementary Table S1**). The LPS O-antigen composed of pentasaccharide repeat units was shown to function as a receptor for phage φ80-18 (Zhang et al., 1997), and a very similar structure is present in serotype O:7,8 O-antigen, however, the O-antigen repeat unit structure of O:4,32 shares only the reducing-end sugar, *N*-acetylgalactosamine, with the O:8 structure (Skurnik and Zhang, 1996). The structures of serotype O:20 and O:21 O-repeat units are not known, however, it is possible that they also contain an O-unit with a reducing-end *N*-acetylgalactosamine. If so, the phage receptor structure could be composed of the junction between the



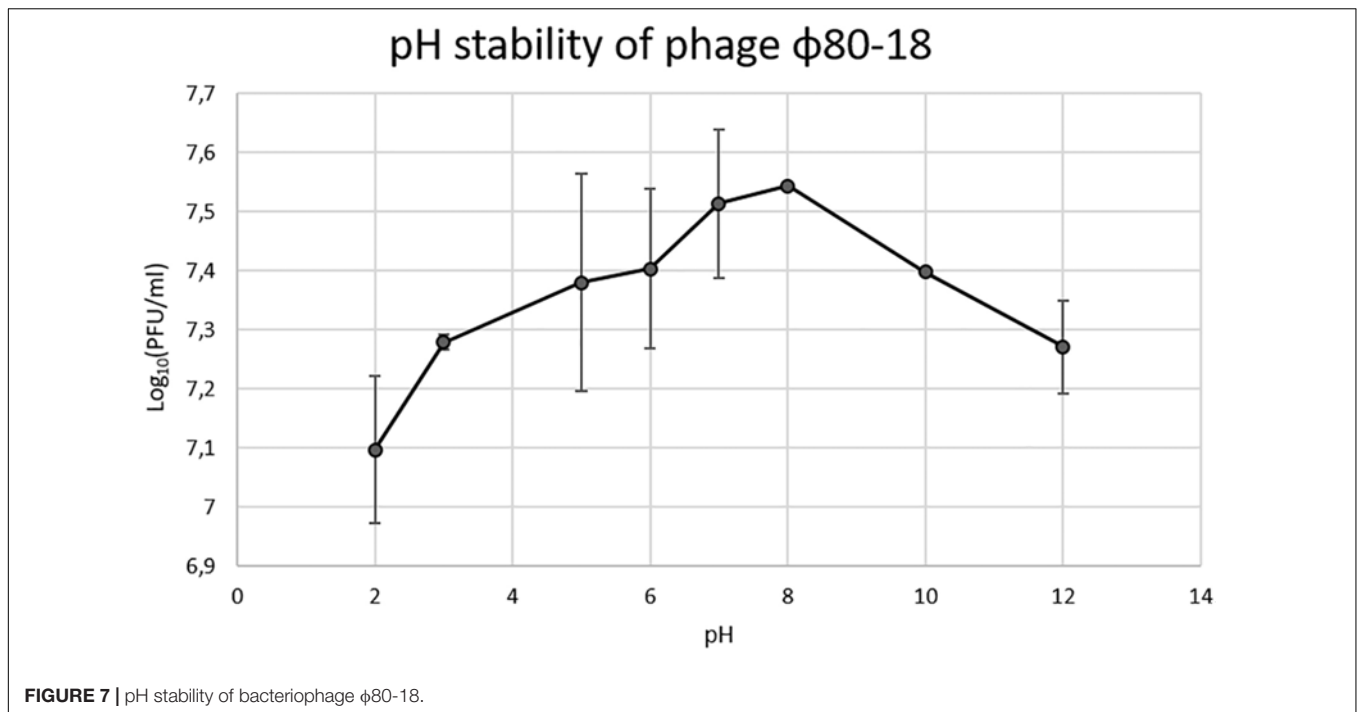
LPS core and the reducing-end *N*-acetylgalactosamine of the O-antigen.

DISCUSSION

Foodborne illnesses are still common despite of the use of many antibacterial methods during food production such as

pasteurization, high pressure processing (HPP), irradiation or chemical disinfectants. *Y. enterocolitica* is a food-borne zoonotic pathogen which is able to grow at 4°C, making it dangerous when contaminated product is stored at low temperatures. The most common source of this pathogen is raw pork (Leon-Velarde et al., 2019).

We show here that bacteriophage φ80-18 is stable and active in a wide range of pH (from 2 to 12) (Figure 7) and temperature



(from 4 to 50°C) (**Figure 6**). These properties of φ80-18 make it a potential candidate for further research on the elimination of *Y. enterocolitica* serotype O:8 and possible other American serotypes, for example, during the processing of food products. And these properties suggest that the phage would be easy to maintain and store for longer periods. Furthermore, the tolerance to pH 2 further indicates that the phage particles might survive the exposure to gastric juices after oral administration of the phage. The tail fiber of phage φ80-18 is also a good candidate to be used for specific detection of the American pathogenic *Y. enterocolitica* serotype bacteria.

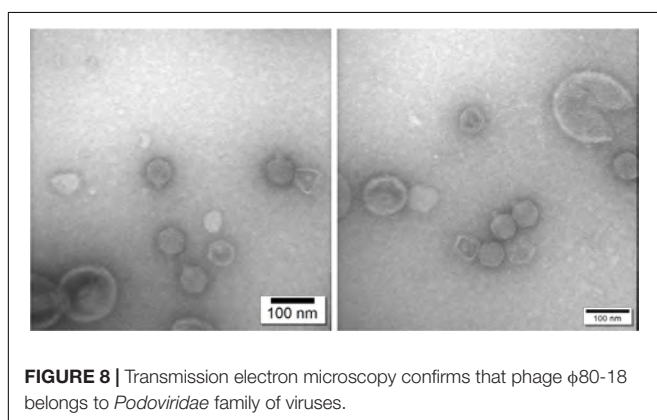
Phylogenetic trees constructed using the whole genome sequence (**Figure 2**) or RNAP protein sequence (**Figure 3**) confirmed that φ80-18 belongs to *Autographivirinae* subfamily in *Podoviridae* and shows the highest similarity to *Yersinia* phage fHe-Yen3-01 (**Table 1**), followed by *Pectobacterium*

phage MA13, *Cronobacter sakazakii* phage vB_CskP_GAP227, *Cronobacter* phage Dev-CD-23823, *Pectobacterium* phage PP2 and *Pectobacterium* phage Arno160. Including other closely related phages such as *Yersinia* phage PhiR8-01, *Aeromonas* phage ZPAH7, *Salmonella* phage vB_SpuP_Spp16 and *Aeromonas* phage Ahp1, it forms a cluster that is stable in both phylogenetic analyses (**Figure 2**).

We recently demonstrated that *Yersinia* bacteriophage fHe-Yen3-01 can be used to treat of kitchen utensils (wooden and plastic cutting boards, knives) and artificial hands contaminated by *Y. enterocolitica* (Jun et al., 2018). After treatment with the phages, CFU counts remained constant for the first 2 h of the experiment. However, after 2 h, there were no detectable bacteria. The results of this experiments proves the potential of using *Yersinia* phages in the food industry. Phage fHe-Yen3-01 is closely related to phage φ80-18 (92% of nucleotide sequence identity) indicating that φ80-18 is also a good candidate for this type of research.

Additionally, the research on the PY100 phage was interesting because of its lytic properties and activity in controlling *Yersinia* in meat. PY100 significantly reduced the number of bacteria at 4°C in pork (the best results were obtained at a MOI 10⁴, when the number of bacteria decreased by up to 5 log₁₀ units) (Orquera et al., 2012). This is also an encouraging argument for the possibility of using *Yersinia* phages in the food industry.

The food poisoning is still one of the major causes of hospitalization or even patients death around the world (Moye et al., 2018). Pasteurization and HPP, are methods used for inactivating microbes in liquids, dairy products and pre-cooked meals. However, these methods cannot be used with fresh meats due to their influence on the color as well as the nutritional



content (Wolbang et al., 2008; Bajovic et al., 2012; Moye et al., 2018).

Irradiation is effective in reducing pathogenic bacteria in food, but it can also affect the food's organoleptic properties. Chemical sanitizers, such as chloride, reduce bacteria from fruits and vegetables surface, but also these chemicals affect the environment (Beuchat and Ryu, 1997; Sohaib et al., 2016; Moye et al., 2018). More consumers now do not tolerate chemical additives in foods for example, because of allergies. However, all the microbes (pathogenic bacteria, normal flora or probiotic bacteria) are killed by all these methods.

A completely another approach is to use lytic bacteriophages for specific foodborne bacteria in foods thereby circumventing any adverse influence on normal, most of the time beneficial microflora. Currently, phage biocontrol is the most environmentally friendly method which can be used to eradicate pathogens from food products (Moye et al., 2018). At present, several phages have been approved by the FDA for use in the food industry.

Phages are used in the food industry to combat pathogens such as *E. coli* 0157:H7, *Listeria monocytogenes*, *Salmonella* spp., *Shigella* spp. (Moye et al., 2018). These are the first steps toward using lytic bacteriophages as safe and natural antibacterial agents. Bacteriophage biocontrol can be used both pre-harvest (e.g., live animals) and post-harvest (e.g., applied to food surface, packing materials) to remove pathogens.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and

accession number(s) can be found here: <https://www.ncbi.nlm.nih.gov/genbank/>, HE956710.

AUTHOR CONTRIBUTIONS

BS-O performed the stability experiments and took TEM photography of the phage. KF took part in conducting experiments. LH and AN performed the initial annotation. MS performed the final annotation of the genome and analyzed the data. MP purified and prepared the phage particles for proteomic analysis. MQ determined the physical ends of the genome. JJ and MS carried out the host range analyses. LM prepared the phage genome for Illumina sequencing and carried out the initial genome *de novo* assembly of the genome. EB and MS supervised the studies. KF and MW performed the phylogenetic and genome analysis. All authors read and approved the final manuscript.

FUNDING

This study was supported by the National Science Centre, Poland (Grant Numbers UMO-2017/26/E/NZ1/00249). This work was supported by Academy of Finland (project 114075) to MS.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.01356/full#supplementary-material>

REFERENCES

- Ackermann, H. W. (2009). "Basic phage electron microscopy," in *Bacteriophages. Methods in Molecular Biology*TM, Vol. 501, eds M. R. Clokie, and A. M. Kropinski, (Totowa NJ: Humana Press), doi: 10.1007/978-1-60327-164-6_12
- Alekshun, M. N., and Levy, S. B. (2007). Molecular mechanisms of antibacterial multidrug resistance. *Cell* 128, 1037–1050. doi: 10.1016/j.cell.2007.03.004
- Al-Hendy, A., Toivanen, P., and Skurnik, M. (1991). The effect of growth temperature on the biosynthesis of *Yersinia enterocolitica* 0:3 lipopolysaccharide: temperature regulates the transcription of the *rfb* but not of the *rfa* region. *Microb. Pathogenesis* 10, 81–86. doi: 10.1016/0882-4010(91)90068-L
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Bajovic, B., Bolumar, T., and Heinz, V. (2012). Quality considerations with high pressure processing of fresh and value added meat products. *Meat Sci.* 92, 280–289. doi: 10.1016/j.meatsci.2012.04.024
- Baker, P. M., and Farmer, J. J. (1982). New bacteriophage typing system for *Yersinia enterocolitica*, *Yersinia kristensenii*, *Yersinia frederiksenii*, and *Yersinia intermedia*: correlation with serotyping, biotyping, and antibiotic susceptibility. *J. Clin. Microbiol.* 15, 491–502. doi: 10.1128/jcm.15.3.491-502.1982
- Bertani, G. (2004). Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *J. Bacteriol.* 186, 595–600. doi: 10.1128/jb.186.3.595-600.2004
- Beuchat, L. R., and Ryu, J. H. (1997). Produce handling and processing practices. *Emerg. Infect. Dis.* 3:459. doi: 10.3201/eid0304.970407
- Bottone, E. J. (1999). *Yersinia enterocolitica*: overview and epidemiologic correlates. *Microbes Infect.* 1, 323–333. doi: 10.1016/S1286-4579(99)80028-8
- Brubaker, R. R. (1972). *The Genus Yersinia: Biochemistry and Genetics of Virulence With 3 Figures: Current Topics in Microbiology and Immunology*. Berlin: Springer, 111–158.
- Chen, M., Xu, J., Yao, H., Lu, C., and Zhang, W. (2016). Isolation, genome sequencing and functional analysis of two T7-like coliphages of avian pathogenic *Escherichia coli*. *Gene* 582, 47–58. doi: 10.1016/j.gene.2016.01.049
- Fàbrega, A., and Vila, J. (2012). *Yersinia enterocolitica*: pathogenesis, virulence and antimicrobial resistance. *Enfermed. Infecc. Microbiol. Clin.* 30, 24–32. doi: 10.1016/j.eimc.2011.07.017
- Fischetti, V. A., Nelson, D., and Schuch, R. (2006). Reinventing phage therapy: are the parts greater than the sum? *Nat. Biotechnol.* 24, 1508–1511. doi: 10.1038/nbt1206-1508
- Fredriksson-Ahoma, M., Stolle, A., and Korkeala, H. (2006). Molecular epidemiology of *Yersinia enterocolitica* infections. *FEMS Immunol. Med. Microbiol.* 47, 315–329. doi: 10.1111/j.1574-695X.2006.00095.x
- Hendrix, R. W. (2002). Bacteriophages: evolution of the majority. *Theoret. Pop. Biol.* 61, 471–480. doi: 10.1006/tpbi.2002.1590
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522. doi: 10.1093/molbev/msx281
- Hu, B., Margolin, W., Molineux, I. J., and Liu, J. (2013). The bacteriophage t7 virion undergoes extensive structural remodeling during infection. *Science* 339, 576–579. doi: 10.1126/science.1231887

- Huson, D. H., Richter, D. C., Rausch, C., DeZulian, T., Franz, M., and Rupp, R. (2007). Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8:460. doi: 10.1186/1471-2105-8-460
- Ichinohe, H., Yoshioka, M., Fukushima, H., Kaneko, S., and Maruyama, T. (1991). First isolation of *Yersinia enterocolitica* serotype O:8 in Japan. *J. Clin. Microbiol.* 29, 846–847. doi: 10.1128/jcm.29.4.846-847.1991
- Jun, J. W., Park, S. C., Wicklund, A., and Skurnik, M. (2018). Bacteriophages reduce *Yersinia enterocolitica* contamination of food and kitchenware. *Int. J. Food Microbiol.* 271, 33–47. doi: 10.1016/j.ijfoodmicro.2018.02.007
- Kajsik, M., Bugala, J., Kadlíčková, V., Szemes, T., Turdó, J., and Drahovská, H. (2019). Characterization of Dev-CD-23823 and Dev-CT57, new Autographivirinae bacteriophages infecting *Cronobacter* spp. *Arch. Virol.* 164, 1383–1391. doi: 10.1007/s00705-019-04202-3
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285
- Katoh, K., and Standley, D. M. (2013). MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kiljunen, S., Hakala, K., Pinta, E., Huttunen, S., Pluta, P., Gador, A., et al. (2005a). Yersiniophage φR1-37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by deoxyuridine. *Microbiology* 151, 4093–4102. doi: 10.1099/mic.0.28265-0
- Kiljunen, S., Vilen, H., Pajunen, M., Savilahti, H., and Skurnik, M. (2005b). Nonessential genes of phage φYeO3-12 include genes involved in adaptation to growth on *Yersinia enterocolitica* serotype O:3. *J. Bacteriol.* 187, 1405–1414. doi: 10.1128/JB.187.4.1405-1414.2005
- Lavigne, R., Sun, W. D., and Volckaert, G. (2004). PHIRE, a deterministic approach to reveal regulatory elements in bacteriophage genomes. *Bioinformatics* 20, 629–635. doi: 10.1093/bioinformatics/btg456
- Lefort, V., Desper, R., and Gascuel, O. (2015). FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program: table 1. *Mol. Biol. Evol.* 32, 2798–2800. doi: 10.1093/molbev/msv150
- Leon-Velarde, C. G., Happonen, L., Pajunen, M., Leskinen, K., Kropinski, A. M., Mattinen, L., et al. (2016). *Yersinia enterocolitica*-specific infection by bacteriophages TG1 and φR1-RT is dependent on temperature-regulated expression of the phage host receptor OmpF. *Appl. Environ. Microbiol.* 82, 5340–5353. doi: 10.1128/AEM.01594-16
- Leon-Velarde, C. G., Jun, J. W., and Skurnik, M. (2019). *Yersinia* phages and food safety. *Viruses* 11:1105. doi: 10.3390/v11121105
- Leskinen, K., Blasdel, B. G., Lavigne, R., and Skurnik, M. (2016). RNA-sequencing reveals the progression of phage-host interactions between φR1-37 and *Yersinia enterocolitica*. *Viruses* 8:111. doi: 10.3390/v8040111
- Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., et al. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* 43, W580–W584. doi: 10.1093/nar/gkv279
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H.-P., and Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. doi: 10.1186/1471-2105-14-60
- Meier-Kolthoff, J. P., and Göker, M. (2017). VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics* 33, 3396–3404. doi: 10.1093/bioinformatics/btx440
- Moye, Z. D., Woolston, J., and Sulakvelidze, A. (2018). Bacteriophage applications for food production and processing. *Viruses* 10:205. doi: 10.3390/v10040205
- Nicolle, P., Mollaret, H., Hamon, Y., Vieu, J. F., Brault, J., and Brault, G. (1967). Lysogenic, bacteriocinogenic and phage-typing study of species *Yersinia enterocolitica*. *Ann. Inst. Pasteur* 112:86.
- Orquera, S., Gözl, G., Hertwig, S., Hammer, J., Sparborth, D., Joldic, A., et al. (2012). Control of campylobacter spp. and yersinia enterocolitica by virulent bacteriophages. *J. Mol. Genet. Med.* 6:273. doi: 10.4172/1747-0862.1000049
- Pajunen, M., Kiljunen, S., and Skurnik, M. (2000). Bacteriophage φYeO3-12, specific for yersinia enterocolitica serotype O:3, is related to coliphages T3 and T7. *J. Bacteriol.* 182, 5114–5120. doi: 10.1128/JB.182.18.5114-5120.2000
- Pajunen, M. I., Kiljunen, S. J., Söderholm, M. L., and Skurnik, M. (2001). Complete genomic sequence of the lytic bacteriophage φYeO3-12 of *Yersinia enterocolitica* serotype O:3. *J. Bacteriol.* 183, 1928–1937. doi: 10.1128/JB.183.6.1928-1937.2001
- Pinta, E., Duda, K. A., Hanuszkiewicz, A., Salminen, T. A., Bengoechea, J. A., Hyttiäinen, H., et al. (2010). Characterization of the six glycosyltransferases involved in the biosynthesis of *Yersinia enterocolitica* serotype O:3 lipopolysaccharide outer core. *J. Biol. Chem.* 285, 28333–28342. doi: 10.1074/jbc.M110.111336
- Rambaut, A. (2006). *FigTree 1.4.3 - a Graphical Viewer of Phylogenetic Trees and a Program for Producing Publication-Ready Figures*. Available online at: <http://tree.bio.ed.ac.uk/software/figtree/> (accessed June 9, 2020).
- Rastawicki, W., Szych, J., Gierczyński, R., and Rokosz, N. (2009). A dramatic increase of *Yersinia enterocolitica* serogroup O:8 infections in Poland. *Eur. J. Clin. Microbiol. Infect. Dis.* 28, 535–537. doi: 10.1007/s10096-008-0647-7
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., et al. (2000). Artemis: sequence visualization and annotation. *Bioinformatics* 16, 944–945. doi: 10.1093/bioinformatics/16.10.944
- Sabina, Y., Rahman, A., Ray, R. C., and Montet, D. (2011). *Yersinia enterocolitica*: mode of transmission, molecular insights of virulence, and pathogenesis of infection. *J. Pathog.* 2011:429069. doi: 10.4061/2011/429069
- Salem, M., and Skurnik, M. (2018). Genomic characterization of sixteen *Yersinia enterocolitica*-infecting podoviruses of pig origin. *Viruses* 10:174. doi: 10.3390/v10040174
- Sambrook, J., Fritsch, E., and Maniatis, T. J. (2001). *Molecular Cloning: a Laboratory Manual*, 3rd Edn. New York, NY: Cold Spring Harbor Laboratory Press, 9.14–19.19.
- Skurnik, M. (1999). *Molecular Genetics of Yersinia Lipopolysaccharide. Genetics of Bacterial Polysaccharides*. Boca Raton, FL: CRC Press, 23–51.
- Skurnik, M., Hyttiäinen, H. J., Happonen, L. J., Kiljunen, S., Datta, N., Mattinen, L., et al. (2012). Characterization of the genome, proteome, and structure of yersiniophage φR1-37. *J. Virol.* 86, 12625–12642. doi: 10.1128/JVI.01783-12
- Skurnik, M., and Zhang, L. (1996). Molecular genetics and biochemistry of *Yersinia lipopolysaccharide*. *APMIS* 104, 849–872. doi: 10.1111/j.1699-0463.1996.tb04951.x
- Söding, J., Biegert, A., and Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33(Suppl. 2), W244–W248. doi: 10.1093/nar/gki408
- Sohaib, M., Anjum, F. M., Arshad, M. S., and Rahman, U. U. (2016). Postharvest intervention technologies for safety enhancement of meat and meat based products; a critical review. *J. Food Sci. Technol.* 53, 19–30. doi: 10.1007/s13197-015-1985-y
- Solovyev, V., and Salamov, A. (2011). “Automatic annotation of microbial genomes and metagenomic sequences,” in *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies*, ed. R. W. Li, (Hauptpage, NY: Nova Science Publishers), 61–78.
- Staden, R., Judge, D. P., and Bonfield, J. K. (2003). *Analyzing Sequences Using the Staden Package and EMBOSS: Introduction to Bioinformatics*. Totowa, NJ: Humana Press, 393–410.
- Thomson, N. R., Howard, S., Wren, B. W., Holden, M. T., Crossman, L., Challis, G. L., et al. (2006). The complete genome sequence and comparative genome analysis of the high pathogenicity *Yersinia enterocolitica* strain 8081. *PLoS Genet.* 2:e206. doi: 10.1371/journal.pgen.0020206
- Varjosalo, M., Keskitalo, S., Van Drogen, A., Nurkkala, H., Vichalkovski, A., Aebersold, R., et al. (2013). The protein interaction landscape of the human CMG kinase group. *Cell Rep.* 3, 1306–1320. doi: 10.1016/j.celrep.2013.03.027
- Ventola, C. L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *Pharm. Therap.* 40:277.
- Wang, J.-B., Lin, N.-T., Tseng, Y.-H., and Weng, S.-F. (2016). Genomic characterization of the novel aeromonas hydrophila phage Ahp1 suggests the derivation of a new subgroup from phiKMV-like family. *PLoS One* 11:e0162060. doi: 10.1371/journal.pone.0162060
- Wang, X., Li, Y., Jing, H., Ren, Y., Zhou, Z., Wang, S., et al. (2011). Complete genome sequence of a *Yersinia enterocolitica* “Old World”(3/O:9) strain and comparison with the “New World”(1B/O:8) strain. *J. Clin. Microbiol.* 49, 1251–1259. doi: 10.1128/JCM.01921-10
- Wolbang, C. M., Fitos, J. L., and Treeby, M. T. (2008). The effect of high pressure processing on nutritional value and quality attributes of *Cucumis melo* L. *Innov. Food Sci. Emerg. Technol.* 9, 196–200. doi: 10.1016/j.ifset.2007.08.001
- Zhang, L., Radziejewska-Lebrecht, J., Krajewska-Pietrasik, D., Toivanen, P., and Skurnik, M. (1997). Molecular and chemical characterization of the lipopolysaccharide O-antigen and its role in the virulence of *Yersinia*

- enterocolitica serotype O: 8. *Mol. Microbiol.* 23, 63–76. doi: 10.1046/j.1365-2958.1997.1871558.x
- Zhang, L., and Skurnik, M. (1994). Isolation of an R-M+ mutant of *Yersinia enterocolitica* serotype O: 8 and its application in construction of rough mutants utilizing mini-Tn5 derivatives and lipopolysaccharide-specific phage. *J. Bacteriol.* 176, 1756–1760. doi: 10.1128/jb.176.6.1756-1760.1994
- Zhao, F., Sun, H., Zhou, X., Liu, G., Li, M., Wang, C., et al. (2019). Characterization and genome analysis of a novel bacteriophage vB_SpuP_Spp16 that infects *Salmonella enterica* serovar pullorum. *Virus Genes* 55, 532–540. doi: 10.1007/s11262-019-01664-0
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Filik, Szermer-Olearnik, Wernecki, Happonen, Pajunen, Nawaz, Qasim, Jun, Mattinen, Skurnik and Brzozowska. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

SUPPLEMENTARY TABLES

The podovirus ϕ 80-18 targets the pathogenic American biotype 1B strains of *Yersinia enterocolitica*

Karolina, Filik^{1#}, Bożena Szermer-Olearnik^{1#}, Maciej Wernecki², Lotta J. Happonen^{3,7}, Maria I. Pajunen⁴, Ayesha Nawaz⁴, Suleman Mohammed Qasim⁴, Jin Woo Jun⁵, Laura Mattinen⁴, Mikael Skurnik^{4,6#} and Ewa Brzozowska¹

¹ *Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Wrocław, Poland*

² *Department of Microbiology, Institute of Genetics and Microbiology, University of Wrocław, Wrocław, Poland*

³ *Institute of Biotechnology and Department of Biosciences, University of Helsinki, Helsinki, Finland.*

⁴ *Department of Bacteriology and Immunology, Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland*

⁵ *Department of Aquaculture, Korea National College of Agriculture and Fisheries, Jeonju 54874, Republic of Korea*

⁶ *Division of Clinical Microbiology, Helsinki University Hospital, HUSLAB, Helsinki, Finland*

⁷ *Lund University, Department of Clinical Sciences Lund, Infection Medicine, Lund, Sweden.*

Coresponding authors:

E-mails:

mikael.skurnik@helsinki.fi;

karolina.filik@hirszfeld.pl;

bozena.szermer-olearnik@hirszfeld.pl

Table S1 List of *Yersinia* strains used in phage ϕ 80-18 host range experiments. All the strains are from the Skurnik lab strain collection. Efficiency of plating (EOP) with strain 8081 set as 1.0.

Bacterial Strain	Origin / Characteristics	Serotype ^a	Stor ID	Phage sensitivity (EOP)	Reference
<i>Yersinia aleksiciae</i>					
404/81	Human, Finland	O:16	284	-	[1]
317/82	Human, Finland	O:16	390	-	
<i>Yersinia bercovieri</i>					
3016/84	Human, Finland	O:58,16	781	-	[2]
127/84	Human, Finland	NT	789	-	[2]
3984/84	Human, Finland	O:58,16	790	-	[2]
<i>Yersinia enterocolitica</i>					
10/84	Human, Finland	NT	760	-	[3]
1539	Human, Finland	NT	286	-	[3]
8533/84	Human, Finland	NT	841	-	[3]
659/83	Human, Finland	K1, NT	741	-	[3]
4367/83	Human, Finland	K1, NT	533	-	[3]
gk132	Chinchilla	O:1	908	-	[4]
JDE029	Human	O:1	1006	-	[5]
gk2943	Goat, Norway	O:2	942	-	[3]
gk1142	Hare, Norway	O:2	943	-	[6]
20373/79	Human, Finland	O:3	2	-	[3]
5854	Human, Finland	O:3	3	-	[3]
9568/79	Human, Finland	O:3	4	-	[3]
6471/76	Human, Finland	O:3	1254	-	[7]
(YeO3)					
6471/76-c	Human, Finland	O:3	1255	-	[7]
(YeO3-c)					
YeO3-c-OC	LPS mutant, derivative of YeO3-c	O:3	2477		[8]
YeO3-c-OCR	LPS mutant, derivative of YeO3-c	O:3 (rough)	2684		[8]
JDE766	Human, North America	O:1,2,3	1002	-	[5]
gc3973-76	Stool, USA	O:4	1029	+ (0.23±0.05)	[9]
JDE701	Human, North America	O:4,32	952	+++ (3.73±0.69)	[10]
4209 CR+	Unknown, Canada	O:5	968	+	[11]
14693/84	Human, Finland	O:5	876	-	[3]
298/85a2	Human, Finland	O:5	908	+	[3]
20109/83	Human, Finland	O:5	730	-	[3]
14779/83	Human, Finland	O:5	664	-	[4]
17223/83	Human, Finland	O:5	693	-	[3]
18710/83	Human, Finland	O:5	709	-	[3]
gk7500	Coypu, Unknown	O:5,27	944	-	[6]
JDE657	Human, North America	O:5,27	1004	-	[5]
JDE654	Human, North America	O:5,27	1007	-	[5]
gc815-73	Unknown, USA	O:5,27	1026	-	[9]
590/80	Human, Finland	O:6	96	-	[7]

Bacterial Strain	Origin / Characteristics	Serotype^a	Stor ID	Phage sensitivity (EOP)	Reference
266/84	Human, Finland	O:6	838	-	[3]
189/80	Human, Finland	O:6,30	92	-	[12]
6737/80	Human, Finland	O:6,30	200	-	[12]
3604/80	Human, Finland	O:6,30	201	-	[12]
438/80	Human, Finland	O:6,31	217	-	[7]
1309/80	Human, Finland	O:6,31	16	-	[7]
605	Human, Finland	O:7,8	98	+	[7]
22848/79	Human, Finland	O:7,8	28	+	[7]
17869/83	Human, Finland	O:7,8	698	+++ (3.37±0.67)	[3]
p310	Unknown	O:8	252	+	[3]
CDCA2635	Milk, USA	O:8	277	+	[13]
TAMU-75	Human, USA	O:8	278	+	[13]
WA	Human, USA	O:8	322		[13]
JDE661	Human, North America	O:8	999	+	[5]
8081 (YeO8)	Human septicemia, USA	O:8	1258	++ (1.00)	[14]
8081-c	pYV-cured derivative of 8081	O:8	1259	+++ (3.03±0.12)	[14]
8081-R2	Rough derivative of 8081	rough O:8	1978		[15]
8081-c-R2	Rough derivative of 8081-c	rough O:8	1978		[15]
YeO8-c:: ΔwbcEGB	LPS mutant of 8081-c	Semi-rough O::8	2427		[16]
YeO8::Δwzz GB	LPS mutant of 8081		2429		[16]
YeO8:: Δwbc EGB	LPS mutant of 8081	Semi-rough O::8	2446		[16]
277/74	Human, Finland	O:9	1	-	[3]
4945/74	Human, Finland	O:9	21	-	[3]
767/73	Human, Finland	O:9	54	-	[3]
467/73	Human, Finland	O:9	55	-	[17]
3672/74	Human, Finland	O:9	97	-	[3]
13752/73	Human, Finland	O:9	137	-	[3]
Ruokola/71-c	Human, Finland	O:9	1257	-	[7]
3102/80	Human, Finland	O:10	102	-	[12]
3788/80	Human, Finland	O:10	210	-	[12]
2640/84	Human, Finland	O:10	777	-	[3]
10927/84	Human, Finland	O:10	851	-	[3]
gc1209-79	Blood, USA	O:13	1025	-	[9]
ST5081	Unknown, USA	O:13a,13b	1047	-	[18]
421/84	Human, Finland	O:13,7	862	-	[3]
2446/84	Human, Finland	O:13,7	775	-	[3]
gc9312-78	Stool, USA	O:13,18	1023	-	[9]
15712/83	Human, Finland	O:14	679	-	[1]
gc874-77	Stool, USA	O:20	1021	+	[9]
gc1223-75	Stool, USA	O:20	1028	+++ (6.82±0.85)	[9]
E736	North America	O:21	953	-	[10]
WI-81-50	North America	O:21	956	+ (0.32±0.10)	[19]
80-EA-63	North America	O:21	957	-	[19]

Bacterial Strain	Origin / Characteristics	Serotype^a	Stor ID	Phage sensitivity (EOP)	Reference
431/84	Human, Finland	O:25	878	-	[1]
63/84	Human, Finland	O:26,44	768	-	[20]
18425/83	Human, Finland	O:25,26,44	704	-	[20]
5186/84	Human, Finland	O:28,50	813	-	[20]
gc2139-72	Unknown, USA	O:34	1027	-	[9]
248/84	Human, Finland	O:35,52	824	-	[20]
7104/83	Human, Finland	O:35,36	568	-	[20]
264/85	Human, Finland	O:41,43	904	-	[1]
626/83	Human, Finland	O:41(27),42	728	-	[1]
9613/83	Human, Finland	O:41(27), K1	590	-	[3]
19942/83	Human, Finland	O:41(27),42, K1	729	-	[20]
1346/84	Human, Finland	O:41(27),43	761	-	[20]
647/83	Human, Finland	O:41(27),43	740	-	[20]
3229	Human, Finland	O:50	209	-	[4]
<i>Yersinia frederiksenii</i>					
38/83	Human, Finland	O:48	502	-	[1]
3400/83	Human, Finland	O:16	532	-	[1]
3317/84	Human, Finland	O:35	785	-	[1]
498/85	Human, Finland	NT	910	-	[21]
28/85	Human, Finland	K1, NT	915	-	[21]
IP23047	Institut Pasteur, France	O:3	2287	-	[22]
<i>Yersinia intermedia</i>					
9/85	Human, Finland	O:16,21	914	-	[1]
821/84	Human, Finland	O52,54	757	-	[4]
<i>Yersinia kristensenii</i>					
4336/83	Human, Finland	UT	535	-	[3]
19602/83	Human, Finland	NT	720	-	[3]
119/84	Human, Finland	O:12,25	783	-	[4]
IP22828	Institut Pasteur, France	O:3	2288	-	[22]
<i>Yersinia mollaretii</i>					
92/84	Human, Finland	O:59(20,36,7)	778	-	[2]
IP22404	Institut Pasteur, France	O:3	2289	-	[22]
<i>Yersinia nurmii</i>					
DSM 22296	Meat, Finland	UT	5607	-	[23]
<i>Yersinia pekkanenii</i>					
A125KOH2	Lettuce, Finland	UT	3059	-	[24]
<i>Yersinia pseudotuberculosis</i>					
2812/79	Human, Finland	O:1b	192	-	[7]
677/82	Human, Finland	O:1b	407	-	[3]
324/80	Human, Finland	O:3	214	-	[4]
1261/79	Human, Finland	O:3	245	-	[7]
<i>Yersinia ruckeri</i>					
OMBL3	Fish isolate, Finland	UT	884	-	[4]
OMBL4	Fish isolate, Finland	UT	885	-	[21]

^a NT, non-typeable; UT, untyped.

REFERENCES

1. Skurnik, M.; Toivonen, S. Identification of distinct lipopolysaccharide patterns among *Yersinia enterocolitica* and *Y. enterocolitica*-like bacteria. *Biochemistry (Mosc)* **2011**, *76*, 823-831.
2. Wauters, G.; Janssens, M.; Steigerwalt, A.D.; Brenner, D.J. *Yersinia mollaretii* sp. nov. and *Yersinia bercovieri* sp. nov., formerly called *Yersinia enterocolitica* biogroups 3A and 3B. *Int J Syst Bacteriol* **1988**, *38*, 424-429.
3. Leon-Velarde, C.G.; Happonen, L.; Pajunen, M.; Leskinen, K.; Kropinski, A.M.; Mattinen, L.; Rajtor, M.; Zur, J.; Smith, D.; Chen, S., et al. *Yersinia enterocolitica*-specific infection by bacteriophages TG1 and ϕ R1-RT is dependent on temperature-regulated expression of the phage host receptor OmpF. *Appl Environ Microbiol* **2016**, *82*, 5340-5353.
4. Skurnik, M.; Toivanen, P. Intervening sequences (IVSs) in the 23S ribosomal RNA genes of pathogenic *Yersinia enterocolitica* strains. The IVSs in *Y. enterocolitica* and *Salmonella typhimurium* have common origin. *Mol. Microbiol.* **1991**, *5*, 585-593.
5. Schiemann, D.A.; Devenish, J.A. Relationship of HeLa cell infectivity to biochemical, serological, and virulence characteristics of *Yersinia enterocolitica*. *Infect. Immun.* **1982**, *35*, 497-506.
6. Kapperud, G.; Skarpeid, H.-J.; Solberg, R.; Bergan, T. Outer membrane proteins and plasmids in different *Yersinia enterocolitica* serogroups isolated from man and animals. *Acta path microbiol immunol scand Sect B* **1985**, *93*, 27-35.
7. Skurnik, M. Lack of correlation between the presence of plasmids and fimbriae in *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. *J. Appl. Bact.* **1984**, *56*, 355-363.
8. Biedzka-Sarek, M.; Venho, R.; Skurnik, M. Role of YadA, Ail, and lipopolysaccharide in serum resistance of *Yersinia enterocolitica* serotype O:3. *Infect. Immun.* **2005**, *73*, 2232-2244.
9. Kay, B.A.; Wachsmuth, K.; Gemski, P.; Feeley, J.C.; Quan, T.J.; Brenner, D.J. Virulence and phenotypic characterization of *Yersinia enterocolitica* isolated from humans in the united states. *J. Clin. Microbiol.* **1983**, *17*, 128-138.
10. Perry, R.D.; Brubaker, R.R. Vwa⁺ phenotype of *Yersinia enterocolitica*. *Infect. Immun.* **1983**, *40*, 166-171.
11. Prpic, J.K.; Robins-Browne, R.M.; Davey, R.B. Differentiation between virulent and avirulent *Yersinia enterocolitica* isolates by using congo red agar. *J. Clin. Microbiol.* **1983**, *18*, 486-490.
12. Skurnik, M.; Nurmi, T.; Granfors, K.; Koskela, M.; Tiilikainen, A.S. Plasmid associated antibody production against *Yersinia enterocolitica* in man. *Scand. J. Inf. Dis.* **1983**, *15*, 173-177.
13. Gemski, P.; Lazere, J.R.; Casey, T. Plasmid associated with pathogenicity and calcium dependency of *Yersinia enterocolitica*. *Infect Immun* **1980**, *27*, 682-685.

14. Portnoy, D.A.; Falkow, S. Virulence-associated plasmids from *Yersinia enterocolitica* and *Yersinia pestis*. *J. Bacteriol.* **1981**, *148*, 877-883.
15. Zhang, L.; Radziejewska-Lebrecht, J.; Krajewska-Pietrasik, D.; Toivanen, P.; Skurnik, M. Molecular and chemical characterization of the lipopolysaccharide O-antigen and its role in the virulence of *Yersinia enterocolitica* serotype O:8. *Mol. Microbiol.* **1997**, *23*, 63-76.
16. Bengoechea, J.A.; Najdenski, H.; Skurnik, M. Lipopolysaccharide O antigen status of *Yersinia enterocolitica* O:8 is essential for virulence and absence of O antigen affects the expression of other *Yersinia* virulence factors. *Mol. Microbiol.* **2004**, *52*, 451-469.
17. Kiljunen, S.; Hakala, K.; Pinta, E.; Huttunen, S.; Pluta, P.; Gador, A.; Lönnberg, H.; Skurnik, M. Yersiniophage ϕ R1-37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by deoxyuridine. *Microbiology* **2005**, *151*, 4093-4102.
18. Toma, S.; Wauters, G.; McClure, H.M.; Morris, G.K.; Weissfeld, A.S. O:13a,13b, a new pathogenic serotype of *Yersinia enterocolitica*. *J Clin Microbiol* **1984**, *20*, 843-845.
19. Schiemann, D.A. Antigenic identity of *Yersinia enterocolitica* serotypes O:Tacoma and O21. *J Clin Microbiol* **1984**, *20*, 831-832.
20. Skurnik, M. Studies on the virulence plasmids of *Yersinia* species. PhD, University of Oulu, Oulu, 1985.
21. Reuter, S.; Connor, T.R.; Barquist, L.; Walker, D.; Feltwell, T.; Harris, S.R.; Fookes, M.; Hall, M.E.; Petty, N.K.; Fuchs, T.M., *et al.* Parallel independent evolution of pathogenicity within the genus *Yersinia*. *Proc Natl Acad Sci U S A* **2014**, *111*, 6768-6773.
22. Pajunen, M.; Kiljunen, S.; Skurnik, M. Bacteriophage ϕ YeO3-12, specific for *Yersinia enterocolitica* serotype O:3, is related to coliphages T3 and T7. *J. Bacteriol.* **2000**, *182*, 5114-5120.
23. Murros-Kontiainen, A.E.; Fredriksson-Ahomaa, M.; Korkeala, H.; Johansson, P.; Rahkila, R.; Björkroth, J. *Yersinia nurmii* sp. nov. *Int. J. Sys. Evol. Microbiol.* **2011**.
24. Murros-Kontiainen, A.E.; Johansson, P.; Niskanen, T.; Fredriksson-Ahomaa, M.; Korkeala, H.; Björkroth, J. *Yersinia pekkanenii* sp. nov. *Int. J. Sys. Evol. Microbiol.* **2011**.

Table S2. Predicted genes and gene products of bacteriophage ϕ 80-18 (acc. no. HE956710).

Gene		Gene product			BLAST and Hhpred similarity searches				Predicted function
Name	Location	Size (aa) ²	Molecular mass (kDa) ²	pI ³	PSI-BLAST ⁴ (id-%)	Organism	Sequence ID	Hhpred domain search ⁵ Probability % / e-value	
g01	933-1283	116	13.4	10.2	MA13_gp04 (72%)	Pectobacterium phage MA13	QGF20948.1	No hits	hypothetical protein
g02	1857-2414	185	20.4	8.5	PP2_002 (66 %)	Pectobacterium phage PP2	AOT25368.1	No hits	hypothetical protein
g03	2411-2602	63	7.1	4.3	No hits			cd13329; Rho guanine nucleotide exchange factor Pleckstrin homology domain 67.57 / 5.1	hypothetical protein
g04	2648-2941	97	10.8	5.6	Gp1.05	Escherichia phage IMM-002	ATI16975.1	No hits	hypothetical protein
g05	2931-3164	77	9.2	7.6	73 aa protein (36%)	Pantoea phage vB_PagP-SK1	QFR42365.1	No hits	hypothetical protein
g06	3164-3397	77	8.7	9.6	BN110_019 (55%)	Yersinia phage phiR8-01	CC188389.2	No hits	hypothetical protein
g07	3464-3691	75	8.6	9.2	Arno160_gp08	Pectobacterium phage Arno160	AZF88070.1	No hits	hypothetical protein
g08	3688-4497	269	29.8	9.0	Only hit fHeYen301_7 (73%)	Yersinia phage fHe-Yen3-01	APU00340.1	cd03819; GT4 family glycosyltransferase 72.81 / 5.1	hypothetical protein
g09	4494-4886	130	14.5	5.4	MA13_gp08 (44%)	Pectobacterium phage MA13	QGF20952.1	cd16380; Bacillus subtilis YitT domain 90.48 / 0.95 cd00953; archeal 2-keto-3-deoxygluconate aldolase 74.99 / 10	PPAP
g10	4867-5097	76	8.5	9.4	BN110_029 (53%)	Yersinia phage phiR8-01	CC188400.2	No hits	hypothetical protein
g11	5081-5434	117	12.8	9.9	EJ02DRAFT_439484 (33%)	Clathrospora elynae Cronobacter phage	KAF1935050.1	No hits	hypothetical protein
g12	5434-5916	160	17.3	5.4	GAP227_11 (38%)	vB_CskP_GAP227	YP_007348330.1	No hits	PPAP
g13	5916-6149	77	8.7	5.6	HRP29_gp12 (30%)	Shigella phage HRP29	QBP32912.1	No hits	hypothetical protein
g14	6146-6370	74	8.4	6.0	HRP29_gp12 (33%)	Shigella phage HRP29	QBP32912.1	No hits	hypothetical protein
g15	6367-6501	44	4.8	8.7	Only hit fHeYen301_14 (73%)	Yersinia phage fHe-Yen3-01	APU00347.1	No hits	hypothetical protein
g16	6515-6766	83	9.4	9.4	BN110_026 (53%)	Yersinia phage phiR8-01	CC188396.2	No hits	PPAP
g17	6768-6962	64	7.8	10.1	Only hit fHeYen301_16 (98%)	Yersinia phage fHe-Yen3-01	APU00349.1	No hits	hypothetical protein

g18	6978-7694	238	27.2	8.9	DNA primase (54%)	Pectobacterium phage Arno160	AZF88076.1	cd03364; DnaG primase nucleotidyltransferase/hydrolase domain 92.98 / 0.19	DNA primase/helicase; PPAP
g19	7691-7885	64	7.7	10.1	Only hit fHeYen301_18 (38%)	Yersinia phage fHe-Yen3-01	APU00351.1	cd09804; mRNA decapping enzyme 72.52 / 4.2	hypothetical enzyme
g20	7879-9123	414	46.2	5.9	DNA helicase (76%)	Cronobacter phage vB_CskP_GAP227	YP_007348335.1	cd01122; homohexameric 5'-3' helicase 99.85 / 2.3e-19	DNA helicase; PPAP
g21	9140-9268	42	4.5	4.7	Only hit fHeYen301_20 (95%)	Yersinia phage fHe-Yen3-01	APU00353.1	No hits	hypothetical protein
g22	9225-9476	83	9.2	5.3	MP2_gp20 (35%)	Morganella phage vB_MmoP_MP2	YP_009291546.1	cd08769; peptidase 34.3 / 10	hypothetical peptidase
g23	9469-9651	60	6.6	4.6	hypothetical protein (25%)	Salmonella phage vB_SpuP_Spp16	AVI05053.1	cd15485; leucine zipper domain 72.18 / 0.93	hypothetical protein
g24	9644-9841	65	7.7	5.6	E5A41_05345 (44%)	Salmonella enterica	KAA6658853.1	No hits	hypothetical protein
g25	9841-10800	319	36.4	6.8	DNA ligase (52%)	Pectobacterium phage MA13	QGF20957.1	cd07902; Adenylation DNA ligase 99.88 / 2.4e-21	DNA ligase; PPAP
g26	10556-10939	127	14.6	6.4	No hits			cd18525; BACK domain 46.0 / 11	hypothetical protein
g27	10917-11681	254	29	4.6	nucleotidyl transferase (20%)	Erwinia phage vB_EamP-S2	AUV57215.1	cd05398; nucleotidyltransferase domain 99.8 / 2.8e-19	Putative nucleotidyltransferase; PPAP
g28	11694-14147	817	92.6	5.6	DNA polymerase (82%)	Pectobacterium phage MA13	QGF20958.1	cd06139; 3'-5' exonuclease 99.07 / 1e-9	DNA polymerase; PPAP
g29	14147-14338	63	6.7	4.6	Sodium channel protein type 5 subunit alpha (28%)	Tetraena socialis	PNH09896.1	cd08962; Glu-tRNA amidotransferase 40.23 / 22	hypothetical amidotransferase
g30	14355-15179	274	29.7	4.8	MA13_Gp15 (71%)	Pectobacterium phage MA13	QGF20959.1	No hits	PPAP
g31	15179-16177	332	37.9	6.7	DNA exonuclease (73 %)	Pectobacterium phage PP2	AOT25387.1	cd00008; 5'-3' exonuclease, T5-5' nuclease 99.86 / 7e-22	5'-exonuclease; PPAP
g32	16174-16599	141	15.6	4.6	MA13_Gp17 (54%)	Pectobacterium phage MA13	QGF20961.1	No hits	PPAP
g33	16475-16921	148	16.5	9.9	putative DNA endonuclease (67%)	Aeromonas phage 25Ahydr2PP	AWH15420.1	cd12870; MqsA antitoxin of MqsR 58.9 / 4.1	Phage endonuclease; weak similarity to antitoxin
g34	16921-17886	321	35.6	7.8	BN110_037 (82%)	Yersinia phage phiR8-01	CC188408.2	cd07424; metallophosphatase domain of type I serine/threonine and tyrosine phosphatases 99.46 / 3.5e-14	Metallophosphatase; PPAP
g35	17887-18063	58	6.8	6.1	putative ATP-dependent RNA helicase (31%)	Trypanosoma rangeli	XP_029241962.1	cd16127; ubiquitin-like domain 37.02 / 10	hypothetical protein

g36	18044-18691	215	24.5	5.4	Tonnikala_41 (35%)	Escherichia phage tonnikala	QHR71312.1	cd02022; DPCK 98.49 / 4.2e-8 cd00227; CPT 98.31 / 2.7e-6 cd01672; TMPK 97.72 / 1.1e-5 cd02023; UMPK 97.47 / 3.4e-5	Kinase phosphorylating CoA, chloramphenicol, TMP or uridine; PPAP
g37	18702-21155	817	92.7	7.8	Phage RNA polymerase (65%)	Cronobacter phage Dev-CD-23823	YP_009223403.1	cd08642; DNA polymerase 92.66 / 0.5	DNA-directed RNA polymerase
g38	21249-21410	53	5.8	8.3	PP2_029 (73 %)	Pectobacterium phage PP2	AOT253395.1	cd14765; hemoglobin 67.12 / 5.5	hypothetical protein
g39	21407-21832	141	16.0	8.8	Hypothetical protein (65%)	Cronobacter phage Dev-CD-23823	YP_009223405.1	cd04301; N-acyltransferase superfamily 98.66 / 2.1e-7 cd02169; citrate lyase ligase acetylates CoA 98.3 / 3.1e-6	N-acyltransferase or acetyltransferase
g40	21832-22227	131	14.1	9.3	Hypothetical protein CF7_11 (51%)	Aeromonas phage CF7	ASZ71957.1	No hits	PPAP
g41	22227-22406	59	6.4	4.6	Anti-sigma regulatory factor (33%)	Fortiea contorta	WP_017652703.1	No hits	hypothetical protein
g42	22415-23920	501	56.2	5.0	Putative head portal protein (71%)	Aeromonas phage phiAS7	YP_007007810.1	No hits	Phage collar; PPAP
g43	23923-24729	268	28.0	4.6	Scaffolding protein (55%)	Pectobacterium phage MA13	QGF20968.1	No hits	Scaffolding-like protein; PPAP
g44	24804-25820	338	36.8	5.3	Major capsid protein (83%)	Pectobacterium phage MA13	QGF20969.1	No hits	Major capsid protein; PPAP
g45	25917-26510	197	22.1	6.4	Putative tail tubular protein A (62%)	Aeromonas phage phiAS7	YP_007007807.1	No hits	Tail tubular protein A; PPAP
g46	26514-29123	869	97.6	5.8	Putative tail tubular protein B (60%)	Aeromonas phage phiAS7	YP_007007806.1	No hits	Phage tail fiber protein; PPAP
g47	29125-29892	255	27.3	6.3	hypothetical protein (47%)	Aeromonas phage 25AhydR2PP	AWH15406.1	No hits	PPAP
g48	29903-32152	749	82.0	5.6	Hypothetical protein (53%)	Cronobacter phage Dev-CD-23823	YP_009223413.1	No hits	PPAP
g49	32156-35935	1259	137.5	5.7	Lytic glycolase (62%)	Cronobacter phage vB_CskP_GAP227	YP_007348360.1	cd13403; lytic murein transglycolase 98.63 / 9.2e-8	Lytic transglycosylase; PPAP
g50	36005-38014	669	70.8	5.0	Putative tail fiber protein (23%)	Aeromonas phage phiAS7	YP_007007802.1	No hits	Phage tail fiber protein; PPAP
g51	38029-38232	67	7.3	9.2	Holin (52 %)	Pectobacterium phage PP2	AOT25407.1	cd14263; diacylglycerol kinase 32.86 / 35	hypothetical kinase
g52	38213-38551	112	12.5	5.2	GAP227_42 (69%) Putative DNA maturase A or terminase small subunit	Cronobacter phage vB_CskP_GAP	YP_007348363.1	cd03460; Catechol 1,2 dioxygenase 53.15 / 20	DNA packaging protein A, likely also PPAP

g53	38551-40455	634	71.5	5.7	DNA maturase B (75 %) or terminase large subunit	Pectobacterium phage PP2	AOT25409.1	cd17921; DEXH-box helicase domain 98.57 / 1.9e-7 cd13962; Holliday junction resolvase RuvC 94.97 / 0.53	DNA packaging protein B; PPAP
g54	40483-40890	135	14.8	5.0	hypothetical protein (29%)	Salmonella phage vB_SpuP_Spp16	AVI05064.1	No hits	PPAP
g55	40926-41480	184	20.5	9.1	Phage lysin (65%)	Yersinia phage phiR8-01	CC188419.2	cd16900; endolysin R21 99.72 / 4.1e-16	Lysozyme
g56	41459-41839	126	13.8	5.6	Only hit fHeYen301_55 (95%)	Yersinia phage fHe-Yen3-01	APU00388.1	No hits	hypothetical protein
g57	41709-42002	97	10.8	6.6	Rz1-like lysis protein (47%)	Pectobacterium phage Arno160	AZF88111.1	No hits	Rz1-like lysis protein

¹ <http://www.endmemo.com/bio/gc.php>

² <https://www.uniprot.org/>

³ https://web.expasy.org/compute_pi/

⁴ The fHe-Yen3-01 hits were excluded from this column. In general, all ϕ 80-18 gene products are >97% identical to corresponding ones in fHe-Yen3-01.




⁵ NCBI conserved domains database v_3.17

Table S3. The predicted phage and host RNA polymerase promoters in the ϕ 80-18 genome. The predicted -35 and -10 boxes of the host RNA polymerase promoters, and the BPROM scores of the boxes, are highlighted in yellow and green, respectively.

Promoter	Location (next gene)	Sequence		
Phage RNA polymerase promoters				
Phage P1	351-375 (<i>g01</i>)	CTGATTGTCTACCCATATAGTAACA		
Phage P2	21210-21234 (<i>g38</i>)	CTGATACTCTACCCATATAGCAACT		
Phage P3	24726-24750 (<i>g44</i>)	TTGATTGTCTACCCATATAGCAATA		
Phage P4	35959-35983 (<i>g50</i>)	TTGATTCTCTACCCATATAGTAACA		
	Consensus	-TGATt-TCTACCCATATAG-AAca		
Host RNA polymerase promoters				
		-35 box	-10 box	Scores
Host P5	741-769 (<i>g01</i>)	TTGACAGC	TTCAGAGTAACAAGTTAGTAT	57 / 66
Host P6	1533-1561 (<i>g02</i>)	TGGCTGG	TTGAAACCCTCGGCTATATAAA	33 / 10
Host P7	1785-1811 (<i>g02</i>)	TAGTCA	GTTGGTTAATGTAAAGTATCAT	57 / 16
Host P8	2873-2898 (<i>g05</i>)	TTTATC	TACAAGGATGCGGATATATT	59 / 17
Host P9	3654-3687 (<i>g08</i>)	TGGAAT	CATACTGCCGTCGATTTGGAGTTAAATT	61 / 18
Host P10	6883-6910 (<i>g18</i>)	TAGAAG	TCAAGCATAATGATGGTAAGGT	47 / 17
Host P11	11610-11638 (<i>g28</i>)	TTGAAA	GTACATGATGCACGTCCTAATAT	53 / 60
Host P12	16130-16162 (<i>g32</i>)	TTGAAA	CAACAAGCTAAACTTAAC TGGTTTATT	36 / 60
Host P13	26485-26513 (<i>g46</i>)	TTGGCA	ATCGTATTCCAGTGAGGTAATAT	68 / 38
Host P14	38116-38147 (<i>g52</i>)	CTGCA	AGACTGGGTGTATATCTTAACTATTA	55 / 20
Host P15	40853-40886 (<i>g55</i>)	CTGAAA	CAGAGCCAGCCGTAAGAACGGTATCGT	48 / 25

Article

New Insights on the Feature and Function of Tail Tubular Protein B and Tail Fiber Protein of the Lytic Bacteriophage ϕ YeO3-12 Specific for *Yersinia enterocolitica* Serotype O:3

Anna Pyra ^{1,*} , Karolina Filik ² , Bożena Szermer-Olearnik ², Anna Czarny ² and Ewa Brzozowska ^{2,*} 

¹ Faculty of Chemistry, University of Wrocław, 14 F. Joliot-Curie St, 50383 Wrocław, Poland

² Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, 12 R. Weigl St, 53114 Wrocław, Poland; karolina.filik@hirszfeld.pl (K.F.); bozena.szermer-olearnik@hirszfeld.pl (B.S.-O.); anna.czarny@hirszfeld.pl (A.C.)

* Correspondence: anna.pyra@chem.uni.wroc.pl (A.P.); ewa.brzozowska@hirszfeld.pl (E.B.); Tel.: +48-71-375-7240 (A.P.); +48-71-337-1172 (E.B.)

Received: 29 August 2020; Accepted: 23 September 2020; Published: 24 September 2020



Abstract: For the first time, we are introducing TTPBgp12 and TFPgp17 as new members of the tail tubular proteins B (TTPB) and tail fiber proteins (TFP) family, respectively. These proteins originate from *Yersinia enterocolitica* phage ϕ YeO3-12. It was originally thought that these were structural proteins. However, our results show that they also inhibit bacterial growth and biofilm formation. According to the bioinformatic analysis, TTPBgp12 is functionally and structurally similar to the TTP of *Enterobacteria* phage T7 and adopts a β -structure. TFPgp17 contains an intramolecular chaperone domain at its C-terminal end. The N-terminus of TFPgp17 is similar to other representatives of the TFP family. Interestingly, the predicted 3D structure of TFPgp17 is similar to other bacterial S-layer proteins. Based on the thermal unfolding experiment, TTPBgp12 seems to be a two-domain protein that aggregates in the presence of sugars such as maltose and N-acetylglucosamine (GlcNAc). These sugars cause two unfolding events to transition into one global event. TFPgp17 is a one-domain protein. Maltose and GlcNAc decrease the aggregation temperature of TFPgp17, while the presence of N-acetylgalactosamine (GalNAc) increases the temperature of its aggregation. The thermal unfolding analysis of the concentration gradient of TTPBgp12 and TFPgp17 indicates that with decreasing concentrations, both proteins increase in stability. However, a decrease in the protein concentration also causes an increase in its aggregation, for both TTPBgp12 and TFPgp17.

Keywords: tail tubular proteins; tail fiber proteins; biofilm

1. Introduction

Phage therapy is a promising alternative to antibiotics that are becoming increasingly less effective. Many phages that are effective against pathogenic strains have already been isolated. More than 95% of all phages belong to the *Caudovirales* of order, which includes the *Myoviridae*, *Podoviridae* and *Siphoviridae* families [1]. The bacteriophage ϕ O3-12 belongs to the *Podoviridae* family and it is classified as part of the T7 group [2]. It was reported that most of the predicted gene products of phage ϕ YeO3-12 are over 70% identical to those of the T3 genome. Therefore, the ϕ YeO3-12 phage genome is extensively similar to the *Enterobacteria* T3 and T7 genomes [2].

Yersinia enterocolitica is a Gram-negative bacterium from the *Enterobacteriaceae* family. It is a widespread bacterium that can be found in soil and water environments. The reservoir of *Yersinia*

enterocolitica serotypes is mainly found in pigs. People can become infected by eating uncooked pork contaminated with bacteria or through contact with the feces of livestock [3–6]. It is one of three known species of *Yersinia*, that is both pathogenic to humans and animals. *Y. enterocolitica* usually causes infections specific to the gastrointestinal track. Some common symptoms of infection are watery diarrhea, fever, vomiting, etc. [7,8].

The bacterial cell wall of *Yersinia*, as with other members of the whole Gram-negative bacteria group, contains an outer membrane in which the major component is lipopolysaccharide (LPS). LPS consists of three main parts: the lipid A, the core oligosaccharide that can often be divided into an inner and outer core, and an O-specific polysaccharide [9]. The presence of the complete LPS is required for the full virulence of *Yersinia enterocolitica* [10]. The outer core (OC) of the LPS is a particularly important virulence factor in *Yersinia enterocolitica* serotype O:3 (YeO3) [11].

The OC saccharide of YeO3 LPS contains 2-acetamido-2,6-dideoxy-D-xylo-hex-4-ulopyranose (Sugp), two glucose (Glc), one galactose (Gal) and two *N*-acetylgalactosamine (GalNAc) residues [9]. It was reported that the external part of the LPS also functions as a bacteriophage receptor [9,12]. Therefore, phages, that display specificity for a particular bacterial strain, could be used as potential diagnostic or therapeutic tools.

Phages attach to their host bacteria with the end of their tails. During infection, DNA is ejected from the capsid through a tail complex to the bacterial host. This stage is well characterized for the T7 phage [13]. The tail machine is composed of the following proteins: the portal (gp8), the adaptor (gp11), the nozzle (gp12, T7_TTPgp12) and the fiber (gp17, T7_TFPgp17). The portal protein, which exists in two conformations, acts as a valve on the portal pore. Thus, it is responsible for the regulation of the passage of DNA passage into the capsid. The adaptor protein interacts with the bottom of the portal protein, allowing DNA to slip along the tail up to the nozzle protein. The nozzle protein, T7_TTPgp12, contains four domains: the platform (interacting with the adaptor protein), the fiber dock interaction domain, the central β -propeller domain and the most distal nozzle tip domain. The nozzle, T7_TTPgp12, is the main protein responsible for closing and securing DNA inside the tail of the mature virus. The tail fiber protein, T7_TFPgp17, interacts with the bacterial LPS [14]. It has been reported that the C-terminal receptor-binding domain of T7_TTPgp17 is responsible for host-range determination, most likely by binding to a specific LPS region that varies between bacterial strains [15]. Additionally, it is well known that phages have evolved to be able to infect different hosts by adapting the receptor domain of TFP [16]. The adaptation is propelled by gene transfer, which accidentally leads to the acquisition new features. One example is the intramolecular chaperone (classified as S74 peptidase family) which contains a conserved C-terminal domain of an endosialidase found in the ϕ K1F phage. The gene encoding the chaperone has become a common feature of other TFPs and it can be removed by a self-cleaving mechanism [17,18]. The first 175 N-terminal amino acid residues of this endosialidase show homology to the N-terminal regions of T7_TFPgp17, TFP of T3 phage (T3_TFPgp17) and TFP of ϕ YeO3-12 phage [17]. The C-terminals of the TFPs differ since the C-terminals form the distal part of TFPs, which bind to the host receptors [19].

In our previous paper, we reported the biochemical features of TTPAgp11 of *Yersinia* phage ϕ YeO3-12, which, in addition to its structural function, also acts as a sugar hydrolyzing enzyme [20]. In this paper, we are presenting our results of our bioinformatic analysis, biochemical characterization and antibacterial tests of two phage proteins. These proteins are the tail tubular protein B (TTPBgp12) and tail fiber protein (TFPgp17) originating from the *Yersinia* phage ϕ YeO3-12. First, the nano differential scanning fluorimetry (nanoDSF) was used to assess the folding and stability of both proteins. Moreover, the stability in presence of different saccharides molecules was measured.

2. Materials and Methods

2.1. Gene Cloning and Protein Overexpression, Purification and Analysis

The complete genome of bacteriophage ϕ YeO3-12 was found in GenBank under accession number AJ251805. The genes that encode the tail tubular protein B (TTPB) and the tail fiber protein (TFP) (the gp12 and gp17), were amplified by polymerase chain reaction (PCR) using the following primers: TTPBgp12_FW, TACTTCCAATCCAATGCCATGCGCTCTTATGAGATGAAC and TTPBgp12_RV, TTATCCACTTCCAATGTTATTAGCGGTTAAGTAGACCAGAGG, TFPgp17_FW, TACTTCCAATCCAATGCCATGGCTACAACATAAAGACCG and TFPgp17_RV, TTATCCACTTCCAATGTTACTAAGTCTTGTCCTTCTCCAAC. The genomic DNA was isolated from the phage lysate using a viral DNA extraction kit (Biocompare) and served as a template (100 ng) in the PCR reaction. A ligation-independent cloning method was used to clone the PCR products into the pMCSG9 vector using a T4 DNA polymerase [21]. Gene overexpression, protein isolation and purification were performed according to a slightly modified procedure described by Pyra et al. [20]. *E. coli* BL21(DE3)pLysS competent cells were used to overexpress both genes. The competent cells were added to Luria-Bertani (LB) medium with ampicillin, as well as chloramphenicol, in working concentrations of 100 μ g/mL and 25 μ g/mL, respectively. Both proteins were purified using two rounds of nickel-affinity chromatography. Additionally, gel filtration using a Protein Arc 6–600 16/60 HR Sec column equilibrated with 50 mM Tris/HCl buffer, pH 8.0, containing 300 mM NaCl, 5% glycerol and 1 mM dithiothreitol, was used for purification. After the second nickel column, the flow-through fraction containing TTPBgp12 or TFPgp17 was precipitated using ammonium sulfate (95% saturation) overnight at 4 °C. It was then pelleted by centrifugation, dissolved in the buffer mentioned above and then applied to a gel-filtration column. The TTPBgp12 or TFPgp17 fractions were collected and precipitated by ammonium sulfate (95% saturation) and their purity was analyzed by 12% SDS-PAGE [22].

The secondary structure of both proteins was analyzed by circular dichroism (CD) spectroscopy. For the CD analysis, the proteins were dissolved in water with a final concentration of \sim 7 μ M. CD measurements were performed four times using a 1 mm cuvette. The measurements were taken in the wavelength range of 350 to 190 nm using a J-1000 Circular Dichroism spectrophotometer (Jasco Inc., Easton, MD, USA). These data were interpreted using the K2D3 method, which provided an estimation of the secondary structure of the proteins [23]. The proteins concentration was determined by the BCA method [24].

2.2. Protein Folding and Aggregation Tests

The thermal unfolding and protein aggregation experiments were carried out using a Prometheus instrument with nanoDSF and backreflection technology (NanoTemper, South San Francisco, CA, USA), using methods that were previously described [20].

NanoDSF technology allows for the automatic determination of thermal unfolding transition midpoints, T_m (°C). This technology also calculates the onset of unfolding using the transition midpoint and the slope of the unfolding signal. This method monitors unfolding-related tryptophan and tyrosine fluorescence at the emission wavelengths of 330 nm and 350 nm, respectively.

The Prometheus device emits near-UV light at a wavelength that is scattered by aggregated proteins, which causes only non-scattered light to reach the detector. A direct measure of protein aggregation is the result of a reduction in backreflected light.

Both protein unfolding and aggregation tests were performed simultaneously. Then, 10 μ L of TTPBgp12 (0.25 mg/mL, 2.78 μ M) or TFPgp17 (0.25 mg/mL, 3.6 μ M) was placed in the capillary alone or mixed with one of the following sugars (0.15 mg/mL): melibiose, galactose, glucose, maltose, β -lactose, α -lactose, GalNAc and GlcNAc.

The samples were heated to 95 °C, and the fluorescence was monitored throughout the experiment. Data were analyzed with the PR.ThermControl and PR.StabilityAnalysis software packages (Publisher, South San Francisco, CA, USA) [25]. As for the results, melting scans of the studied proteins showed

the following: the fluorescence ratio (350/330 nm) profile used to calculate T_m ($^{\circ}\text{C}$), the first derivative profile showing the initial temperatures of unfolding, aggregation and the occurrence of conformational changes in the protein sample, as well as the scattering profile plotted as attenuation units (mAU) against temperature ($^{\circ}\text{C}$) indicating the onset temperature of protein aggregation.

2.3. Measuring the Effect on Biofilm Formation

One hundred microliters of inoculum, containing 2.2×10^7 CFU/mL of *Yersinia enterocolitica* O:3 was added to each well on 96-round well plates. Then, 10 μL of 0.5 mM of TTPBgp12 or TFPgp17 was added to the test samples. Ten microliters of lysis buffer (300 mM NaCl, 20 mM Tris-HCl pH 8.0, 5% glycerol, 5 mM BME), in which proteins were suspended, was added to the bacterial inoculum and considered as a negative control. As a positive control, 10 μL of 10 mg/mL [26] gentamicin and 10 μL of bacteriophage $\phi\text{Ye O3-12}$ (4.5×10^7 PFU/mL) were used. All samples were triplicated. The 96-round well plates were incubated at 28 $^{\circ}\text{C}$ for 24 h. After incubation, the resulting biofilm was suspended in PBS and plated on agar plates. Agar plates were incubated at 28 $^{\circ}\text{C}$ for 24 h, after this time the number of colonies on the plates was counted. The bacteria layer was also placed on aLab-Tek[®] Chamber SlideTM System (Thermo Fisher Scientific Inc., Rochester, NY, USA), and incubated at 28 $^{\circ}\text{C}$ for 24 h. Then the slides were rinsed with water, dried, and stained with 1% crystal violet. After 30 min, the slides were rinsed repeatedly with water and then observed under a light Olympus microscope (100 \times magnification).

3. Results and Discussion

3.1. Bioinformatics Analysis

While studying the proteins of bacteriophage tails, it was discovered that some structural tail tubular A proteins (TTPAs), of the environmental *Klebsiella* (KP32 and KP34) and *Yersinia* ($\phi\text{YeO3-12}$) phages, can have dual-functionality [20,27,28]. These proteins were only previously considered to be structural phage tail proteins. Our results have shown that they exhibit a second function—a specific saccharide hydrolytic activity on biofilms formed by bacteria. This specificity can be used as an antibacterial agent with a wide spectrum of applications. So far, we have only examined TTPAs. The main question was whether other structural tail phage proteins, like tail tubular proteins B (TTPBs) or tail fiber proteins (TFPs), could also play an antibacterial role. The uncharacterized TTPB (UniProt code: Q9T105) and TFP (UniProt code: Q9T0Z9) encoded by gene 12 and 17, respectively, of *Yersinia* phage $\phi\text{YeO3-12}$ (GenBank code: AJ251805) were chosen for this study. Firstly, the primary structure of both selected proteins was analyzed using BLAST [29], Clustal Omega [30] and HHPred [31].

BLAST amino acid sequence analysis [29] of TTPBgp12 showed that it is similar to the TTPBs of phages that infect other types of bacteria with the highest similarity to bacteriophages infecting *Yersinia*, *Escherichia*, *Salmonella*, *Citrobacter*, *Serratia*, *Klebsiella*, *Leclercia* and *Enterobacteria* (97–99%) (Supplementary data). Notably, amino acid sequence similarity was also found between TTPBgp12 and TFP from phages of *Enterobacter* (99–100%). The observed amino acid sequence diversity is likely a reflection of the phage's adaptations in response to environmental changes and/or bacterial cell variability. Although a lot of proteins with high amino acid sequence similarity to TTPBgp12 were found, only one of them appeared to have been well characterized [13]. This protein was a tail tubular protein of *Enterobacteria* phage T7, T7_TTPgp12, which is a nozzle tail protein. This is the main protein responsible for closing and securing the DNA inside the tail, as well as phage absorption to the host bacteria outer membrane [13,32].

The Clustal Omega tool was used [24] to perform the amino acid sequence alignment of both proteins (Supplementary data) which allowed us to determine that TTPBgp12 is 69% identical to T7_TTPgp12.

The HHPred tool [31] confirmed a high similarity of both proteins predicting that TTPBgp12 is largely structurally homologous to T7_TTPgp12 with 100% probability. This means that the overall fold of TTPBgp12 should be the same as in T7_TTPgp12.

The Phyre server [33] analysis showed that TTPBgp12 adopts probably a mainly β -structure (56%) with only six very short α -helical fragments (4%) and some disordered regions (18%). Domain analysis led to the finding of twenty short regions of polypeptide chain of TTPBgp12, which corresponds to some known domains of a different protein families. Although the confidence level does not specifically reflect the accuracy of the model, the level of confidence in the TTPBgp12 domain prediction was very low, about 20% or lower. Therefore, the conclusion was that there was no specific domain found in the structure of TTPBgp12.

The I-Tasser [34] server was used for further bioinformatic analysis of the 3D structure of TTPBgp12. The I-Tasser server modeled the structure by using the structures deposited in Protein Data bank (PDB) as a template. As a result, five three-dimensional models were generated. The structure of T7_TTPgp12, determined by cryo-electron microscopy (PDB code: 6r21), presented us with the best template for TTPBgp12 modelling (Z-score > 1). The best obtained model of the TTPBgp12 structure suggests it adopts mainly a β -structure (Figure 1). In this prediction, all secondary structure elements matched very well between TTPBgp12 and T7_TTPgp12. This is not surprising, given the almost 70% similarity of the amino acid sequence of both proteins. Interestingly, the only difference was seen in the one loop region which lacked two residues (R744–R745) in the T7_TTPgp12 structure. However, it modelled well in the corresponding loop of the predicted TTPBgp12 structure.

To build the 3D model for TTPBgp12, the Swiss-Model server [35] was used. Seven protein templates were found. Six of them were considered to be less suitable for modeling than the first one with the PDB code: 6r21. Thus, the only one model for TTPBgp12 was predicted. The obtained results were in accordance with TTPBgp12 structure predicted by the I-Tasser server. This was not surprising since both servers used the same protein template for model building.

In conclusion, the amino acid sequence and structure prediction analysis showed a high level of primary and tertiary structure similarity between TTPBgp12 and T7_TTPgp12. This suggests that the proteins have similar biochemical features and functions.

Based on the information deposited in the UniProt database, it is known that TFPgp17 contains a S74 peptidase domain at its C-terminus end (539–645 amino acid residues). This C-terminus end might be an intramolecular chaperone [17] removed after auto-proteolysis [18].

The BLAST analysis [29] showed that TFPgp17 had the highest amino acid similarity (99–100%) to TFps of other phages infecting *Enterobacter*, *Yersinia*, *Serratia*, *Citrobacter*, *Shigella* and *Escherichia*. There were also a lot of phage TFP amino acid sequences with a lower similarity found in the range of 40 to 90% (Supplementary data). No similarity was found between TFPgp17 and the TTPAs or the TTPBs. As for TTPBgp12, the amino acid sequence diversity among the TFps is also reflective of the phage adaptations to environmental changes. Unfortunately, the most similar proteins to TFPgp17, in terms of their amino acid sequence, have not been studied. Therefore, we decided to align the amino acid sequence of TFPgp17 with amino acid sequences of only well biochemically or structurally characterized proteins. The following proteins were chosen for multiple protein sequence alignment analysis by BLAST [23] and Clustal Omega tool [24]: TFP of *Enterobacteria* phage T7 (T7_TFP), TFP of *Enterobacteria* phage T3 (T3_TFP) of tail spike protein of *Enterobacteria* phage K1F (K1F_TSP), L-shaped TFP of *Escherichia* phage T5 (T5_TFP), side TFP of prophage K12 (K12_TFP) and TFP of *Enterobacteria* phage lambda (L_TFP). The BLAST results revealed that most of the similarities in the amino acid sequences of the selected proteins were observed in the N-terminal regions of TFPgp17 and T7_TFP, as well as T3_TFP and K1F_TSP (Supplementary data). T5_TFP, K12_TFP and L_TFP only had very short fragments in their amino acid sequences that were similar to the polypeptide chain of TFPgp17, mainly present in the center and C-end regions (Supplementary data). To clarify, the Clustal Omega alignments were performed for two sets of amino acid sequences. The first input included the protein sequences with a higher BLAST alignment (TFPgp17 and T7_TFP, T3_TFP and K1F_TSP). The second

set contained the protein sequences from the BLAST analysis with a smaller degree of similarity (TFPgp17 and T5_TFP, K12_TFP and L_TFP). All these analyses found that there is not a high similarity between the primary structures of TFPgp17 and the selected proteins, with similarity in the range of only 25% to 36% (Supplementary data).

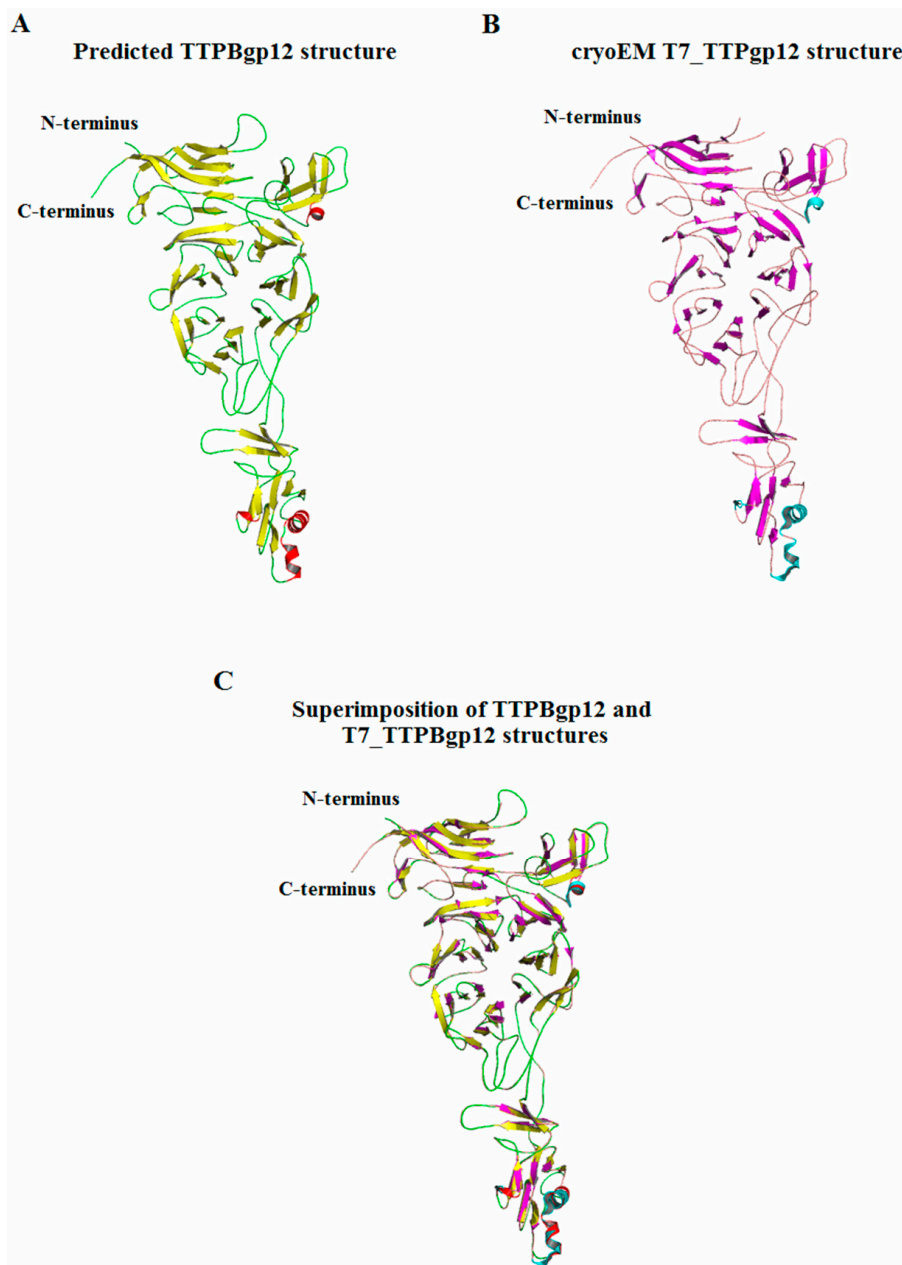


Figure 1. The predicted 3D structural model of tail tubular proteins B (TTPB)gp12 generated using the I-Tasser server [34]. (A) The presented structure was the best calculated model and was obtained using the cryoEM structure of T7_TTPgp12, (B) deposited in PDB (PDB code: 6r21). (C) The predicted structure of TTPBgp12 was superimposed onto cryoEM structure of the T7_TTPgp12 monomer.

Interestingly, the N-terminal of all the compared proteins was much more similar than the C-terminal region, even up to ≈ 78 . This was previously reported by Pajunen et al. [2]. This is in agreement with other findings that state that in many phages, the gene regions coding the C-terminus of TFPs evolve faster than other phage genes, as a result of intense host range selection [36]. This is due to the function of TFP that is responsible for binding to host cell receptors. Our analysis using the HHPred tool [31] also demonstrates a great diversity in the amino acid sequences of TFPs. The obtained results

suggest that the polypeptide chain of TFPgp17 contains some structurally homologous fragments, ~100 to 200 amino acid residues long, to six phage proteins with a probability of 99% (Supplementary data). These proteins include: endo-*N*-acetylneuraminidase (*Enterobacteria* phage K1F), receptor recognition protein (*Salmonella* phage vB_SenMS16), L-shaped tail fiber protein (*Enterobacteria* phage T5), neck appendage protein (*Bacillus* phage GA-1), phiAB6 tail spike (unidentified phage) and tail spike protein (*Acinetobacter* phage vB_AbaP_AS12). The homology of the rest of the hits from the HHPred analysis of TFPgp17 was predicted with a probability of ~60% and lower. Most of the listed proteins were structural tail phage proteins. The neck appendage protein of *Bacillus* phage GA-1 is classified as a chaperone, same as endo-*N*-acetylneuraminidase of *Enterobacteria* phage K1F, which additionally acts as hydrolase. That allows us to speculate that TFPgp17 is mainly a structural tail fiber protein. This protein could play the role of a chaperone protein, a receptor binding protein responsible for the bacterial cell recognition or could even exhibit hydrolytic activity.

The Phyre server analysis [33] found that the best protein templates for TFPgp17 domain prediction (with confidence 99%) are in accordance with the results performed by the HHPred tool. These templates include the following: the neck appendage protein (*Bacillus* phage GA-1), the endo-*N*-acetylneuraminidase (*Enterobacteria* phage K1F) and the L-shaped tail fiber protein (*Enterobacteria* phage T5). All of the previously listed proteins represent chaperone proteins. The obtained results suggest that TFPgp17 could contain a chaperone domain on its C-terminus of its polypeptide chain. It is worth noting that the results obtained by both bioinformatic tools indicated that the acquired domains had a rather low percentage of identity, within the range of 13–24% for HHPred and 18–33% for Phyre analysis. The Phyre server analysis also discovered that TFPgp17 could contain an adhesion domain on its N-terminus and a confidence range of 90–96%. The secondary structure and disorder prediction performed by the Phyre server showed that TFPgp17 probably contains more α -helical elements (36%) than β -strands (25%). In addition, the disordered regions are on a similar level as the ordered helices (36%).

The I-Tasser server [34] was used to further analyze the spatial structure of TFPgp17. It generated five models of TFPgp17 based on the comparison to the protein structures deposited in the PDB. The best template proteins for modelling appeared to be the X-ray structures of endo-*N*-acetylneuraminidase of *Enterobacteria* phage K1F, (K1F_Nase, PDB code: 3gw6) [18] and the L-shaped tail fiber protein of *Enterobacteria* phage T5 (T5_TFP, PDB code: 4uw8) [36]. These two templates served as models for the prediction of the C-terminal of TFPgp17. The best obtained model of TFPgp17 structure suggests it mainly adopts a β -structure (Figure 2). In this prediction, only the C-terminal β -structure elements of TTPBgp12 and T5_TFP matched up well.

Sizeable differences were observed between the rest of the modeled TFPgp17 molecule and the templates proteins. This is not surprising, since the amino acid similarity of these regions, in TFPgp17 and T5_TFP, as well as K1F_Nase, is 15% and 13%, respectively. Therefore, we could assess that the overall structure of TFPgp17 predicted by the I-Tasser server was not correct. Moreover, this server has listed ten structural analogs in PDB to the predicted TFPgp17. The first required structure was a solid fit to the TFPgp17 model (Figure 2). That was the bacterial membrane protein RsaA with PDB code: 5n8p, which represents S-layer proteins [38]. This group of proteins includes a diverse class of molecules found in the surface layer (S-layer) of Gram-negative, Gram-positive bacteria and most archaea. The S-layer consists of repeating molecules of S-layer proteins that protect bacterial cells from the external environment, as well as play a role in pathogenicity. Based on this analysis, the following questions arise: Is it possible that the overall fold of TFPgp17 is actually similar to that of RsaA? Or, is it possible that, given the similar predicted structure of TFPgp17 to RsaA, could the phage protein exhibit similar features as have been seen with RsaA? Thus, could it build into the outer membrane of host bacterial cells, which would help the phage to infect bacteria.

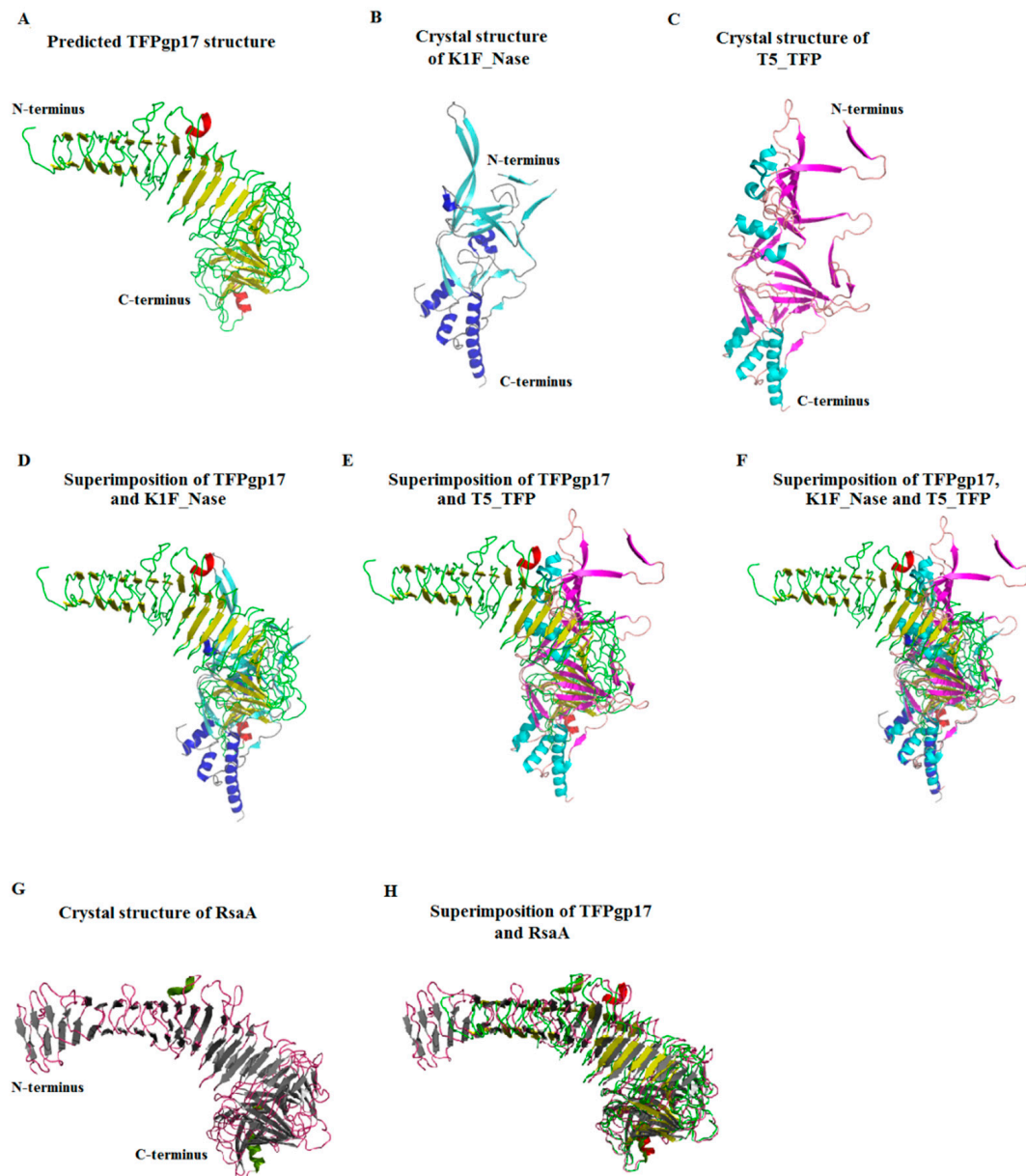


Figure 2. The predicted 3D structural model of TFPgp17 generated using the I-Tasser server [34]. (A) The presented structure was the best calculated model, and was obtained using the crystal structure of (B) endo-*N*-acetylneuraminidase of *Enterobacteria* phage ϕ K1F, (K1F_Nase, PDB code: 3gw6) [18] and (C) L-shaped tail fiber protein of *Enterobacteria* phage T5 (T5_TFP, PDB code: 4uw8) [37]. The predicted structure of TFPgp17 superimposed onto the crystal structure of K1F_Nase (D), T5_TFP (E) and both template proteins (F). (G) The crystal structure of RsaA (PDB code: 5n8p) [38], the structural analog of the predicted TFPgp17. (H) The predicted structure of TFPgp17 superimposed onto RsaA.

Another result from the Swiss-Model server presented with some interesting new information. In this study, 33 templates from PDB for TFPgp17 amino acid sequences were filtered. The templates with the highest quality, predicted from features of the target–template alignment, were selected for model building. The Swiss-Model server generated three partial models of the 3D TFPgp17 structure (Figure 3).

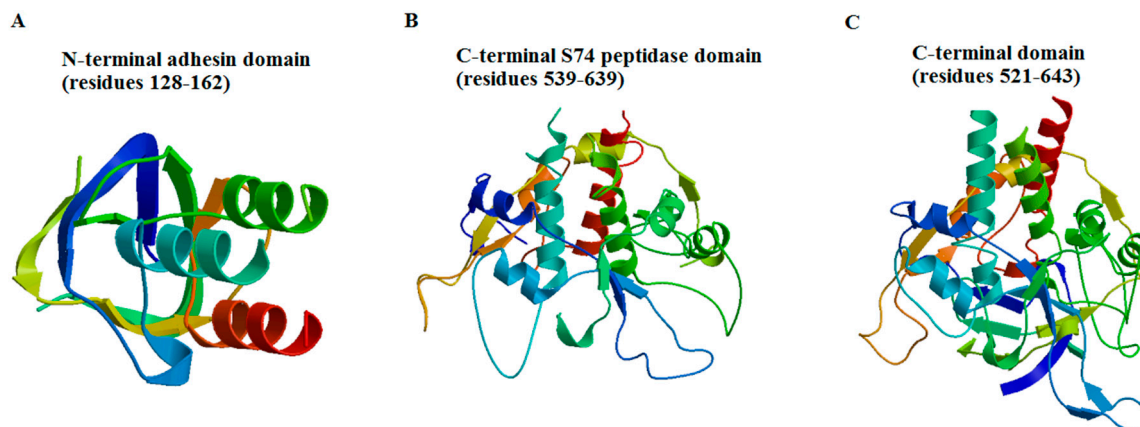


Figure 3. The predicted 3D partial structural models of tail fiber protein (TFP)gp17 generated using the Swiss-Model server [35]. (A) The model #1—N-terminal region consisting of the amino acid residues 128–162 build on the template of adhesin A (PDB code: 3d9x) [39]. (B) The model #2—the C-terminal S74 peptidase domain consisting of the amino acid residues 539–639 build on the template of the neck appendage protein (intramolecular chaperone) of *Bacillus* phage GA-1 (PDB code: 3gud) [18]. (C) The model #3—C-terminal region consisting of the amino acid residues 521–643 build on the template of the L-shaped tail fiber protein of *Enterobacteria* phage T5 (PDB code: 4uw8) [37].

Model no. 1 was built on the template for adhesin A (PDB code: 3d9x) [39] and contains a fragment of the amino acid, containing residues 128–162, which adopts a mixed α/β secondary structure (37% sequence similarity, 5% coverage). The template for model no. 2 was a neck appendage protein (intramolecular chaperone) of *Bacillus* phage GA-1 (PDB code: 3gud) [18]. The predicted model includes the C-terminal S74 peptidase domain built of 539–640 amino acid residues (34% sequence similarity, 14% coverage). Additionally, the last 3D model of TFPgp17 was predicted for its C-terminal amino acid sequence (residues 521–643) using an L-shaped tail fiber protein from *Enterobacteria* phage T5 (T5_TFP, PDB code: 4uw8) [24] as a template (27% sequence similarity, 16% coverage). Models nos. 2 and 3 consist of a mixed α/β secondary structure, with an abundance of a helical structures.

In summary, all the bioinformatic tools used for the TFPgp17 amino acid sequence analysis allowed us to state that TFPgp17 undoubtedly belongs to the phage tail fiber protein family having an intramolecular chaperone domain at its C-terminal end. The N-terminal end of TFPgp17 showed a much higher similarity with another well-characterized representative of this protein family than its C-terminal regions. The great variability of the C-terminal regions is a common feature of TFPs due to being responsible for binding to the host receptors. The partially predicted TFPgp17 3D structure showed that this protein could possess an adhesion domain at the N-terminal region and an autoproteolytic S74 peptidase domain at the C-terminus of its polypeptide chain. The 3D structure of the whole TFPgp17 molecule was also predicted. It demonstrated that this protein could adopt an overall fold very similar to RsaA, which is reported as being a bacterial S-layer protein providing mechanical stability of the cell and protecting bacteria from the outside conditions. For a better characterization of biochemical features of TTPBgp12 and TFPgp17, experimental studies were carried out.

3.2. Gene Cloning and Protein Overexpression, Purification and Analysis

The genes for TTPBgp12 and TFPgp17 were cloned into a pMCSG9 vector and both overexpressed proteins were purified in two rounds of nickel-affinity chromatography. This was followed by gel filtration and analyzed with 12% SDS-PAGE (Figures 4 and 5).

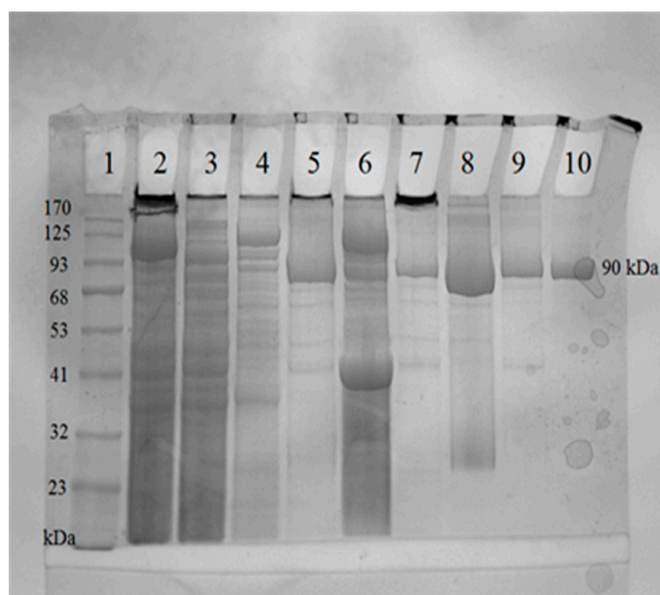


Figure 4. Analysis of TTPBgp12 from *Yersinia* phage ϕ YeO3-12 using 12% SDS-PAGE. The lanes are as follows: (1) prestained protein ladder—mid-range molecular weight (10–180 kDa) (Nippon Genetics); (2) crude extract; (3) the flow through fraction and (4) pooled fractions eluted with 250 mM imidazole after the first round of nickel-immobilized affinity column; (5) TTPBgp12 fractions after Tobacco Etch Virus (TEV) protease cleavage; (6) pooled fractions eluted with 250 mM imidazole after the second round of Ni^{2+} -affinity chromatography; (7) denatured and undenatured (8) TTPBgp12 fractions after ammonium sulfate precipitation; (9) denatured and undenatured (10) TTPBgp12 solution after gel-filtration chromatography.

The SDS-PAGE showed that the purity of both proteins was over 90% and indicated that the molecular weight of TTPBgp12 was ~90 kDa. This is in accordance with the theoretical value. It is interesting that the theoretical molecular mass of TFPgp17 is 69.44 kDa. However, based on the SDS-PAGE, the molecular weight was determined to be ~58 kDa. This was not surprising, since TFPgp17 is a protein containing a S74 peptidase domain, which is responsible for protein autolysis. Hence, the molecular weight of the MBP-TFPgp17 fusion protein was ~110 kDa. However, after TEV protease cleaved off the fusion protein and the second round of Ni^{2+} -affinity chromatography, the molecular weight of TFPgp17 was indicated to be ~58 kDa with SDS-PAGE. During this step of purification, a small protein, ~11 kDa, was also observed in the gel. This corresponded to the molecular mass of the C-terminal S74 peptidase domain (539–645 amino acid residues).

3.3. Protein Folding and Aggregation Tests

The CD spectrum analysis (Figure 6) showed that both purified proteins were folded. They adopted a mixed α/β secondary structure with a large proportion of beta structure in TTPBgp12 and with similar content of helical and beta structures in TFPgp17.

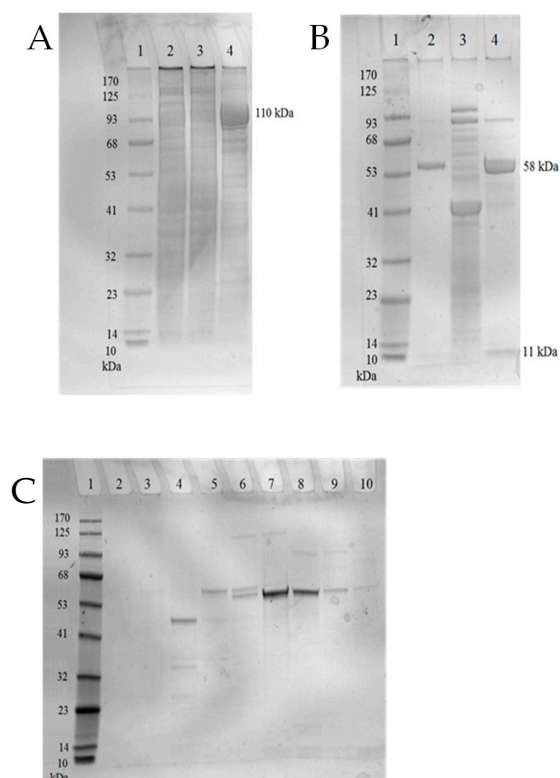
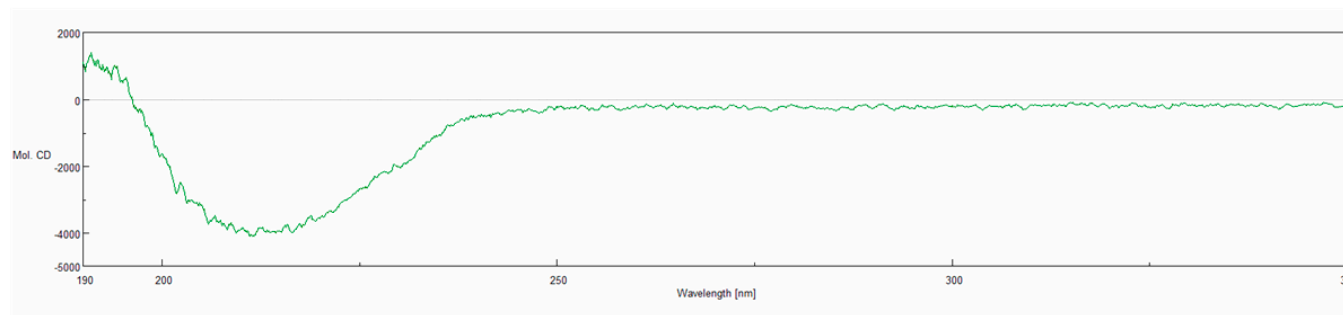


Figure 5. Analysis of TFPgp17 from *Yersinia* phage ϕ YeO3-12 using 12% SDS-PAGE. **(A)** The first round of nickel-immobilized affinity column. The lanes are as follows: (1) prestained protein ladder—mid-range molecular weight (10–170 kDa) (Nippon Genetics); (2) crude extract; (3) the flow through fraction; (4) pooled fractions eluted with 250 mM imidazole. **(B)** The second round of nickel-immobilized affinity column. The lanes are as follows: (1) prestained protein ladder—mid-range molecular weight (10–170 kDa) (Abcam); (2) TFPgp17 fractions after TEV protease cleavage; (3) pooled fractions eluted with 250 mM imidazole; (4) TFPgp17 fractions after ammonium sulfate precipitation. **(C)** Fractions obtained after gel-filtration chromatography: (1) mid-range molecular weight (10–170 kDa) (ABCAM), (2) collected fraction 1, (3) collected fraction 2, (4) collected fraction 3, (5) collected fraction 4, (6) collected fraction 5, (7) collected fraction 6, (8) collected fraction 7, (9) collected fraction 8, (10) collected fraction 9.

NanoDSF and backreflection technology was used to measure the stability, thermal unfolding and aggregation of TTPBgp12 and TFPgp17. Estimating the protein's stability is very important to researchers studying these macromolecules. This quantification reveals information about their thermal and colloidal features. Consequently, it is possible to determine the optimal conditions for large-scale production and long-term storage of proteins. We have also performed stabilization experiments for proteins in the presence of different mono- and disaccharides that are potential binders. In the experiment, the most popular saccharide moieties, which build bacterial polysaccharides components, were used. Some of them are components of the outer core of the lipopolysaccharide (LPS) of *Y. enterocolitica* such as GalNAc, glucose, galactose [40]. Increasing the stability of the proteins in the presence of saccharide molecules could be the premise for complex formation with the protein. However, the nanoDSF experiment is a preliminary screen test and cannot support whether the proteins can interact with sugars. Our results present differences in the stability and aggregation of TTPBgp12 and TFPgp17 (Figures 7 and 8, and Table 1).

A)



B)

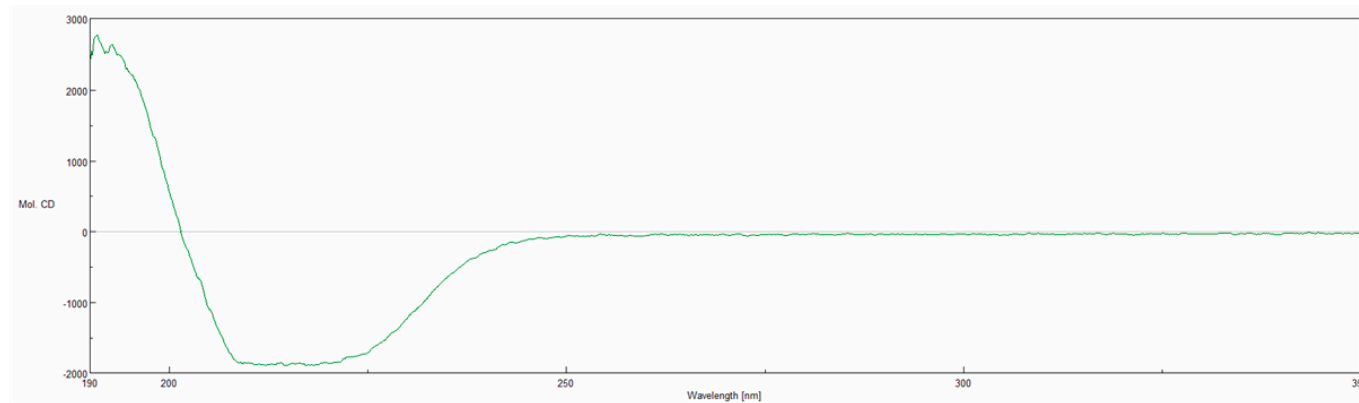


Figure 6. The circular dichroism (CD) spectrum of TTPBgp12 (A) and TFPgp17 (B) in units of molar ellipticity (deg cm² dmol⁻¹).

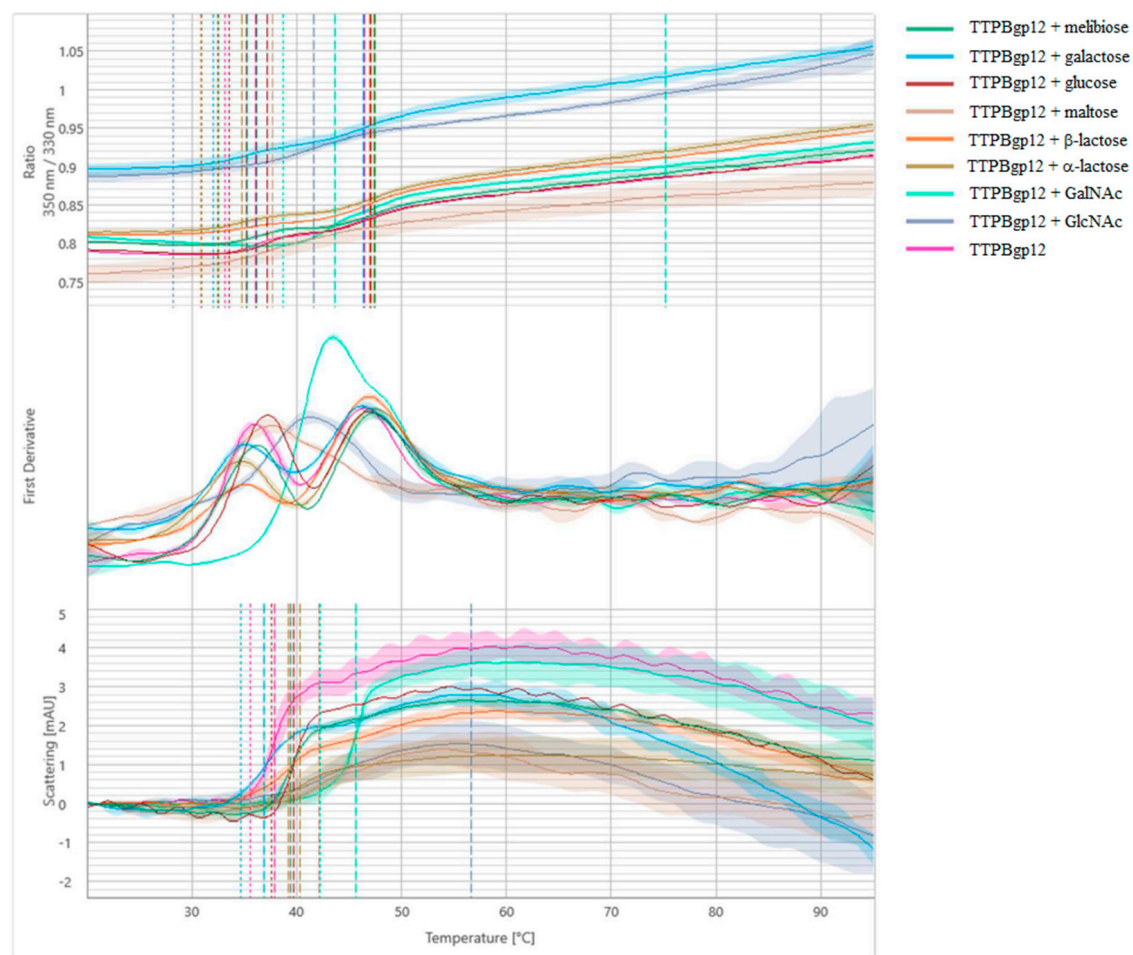


Figure 7. Melting analysis of TTPBgp12 with various sugars. The fluorescence ratio (350 nm/330 nm) is shown in the top panel, the first derivative is shown in the middle panel and scattering is shown in the bottom panel. Thermal unfolding and aggregation onsets (indicated as the vertical dotted lines), as well as unfolding and aggregation transitions (indicated as the vertical dashed lines), are indicated by vertical lines in the graph.

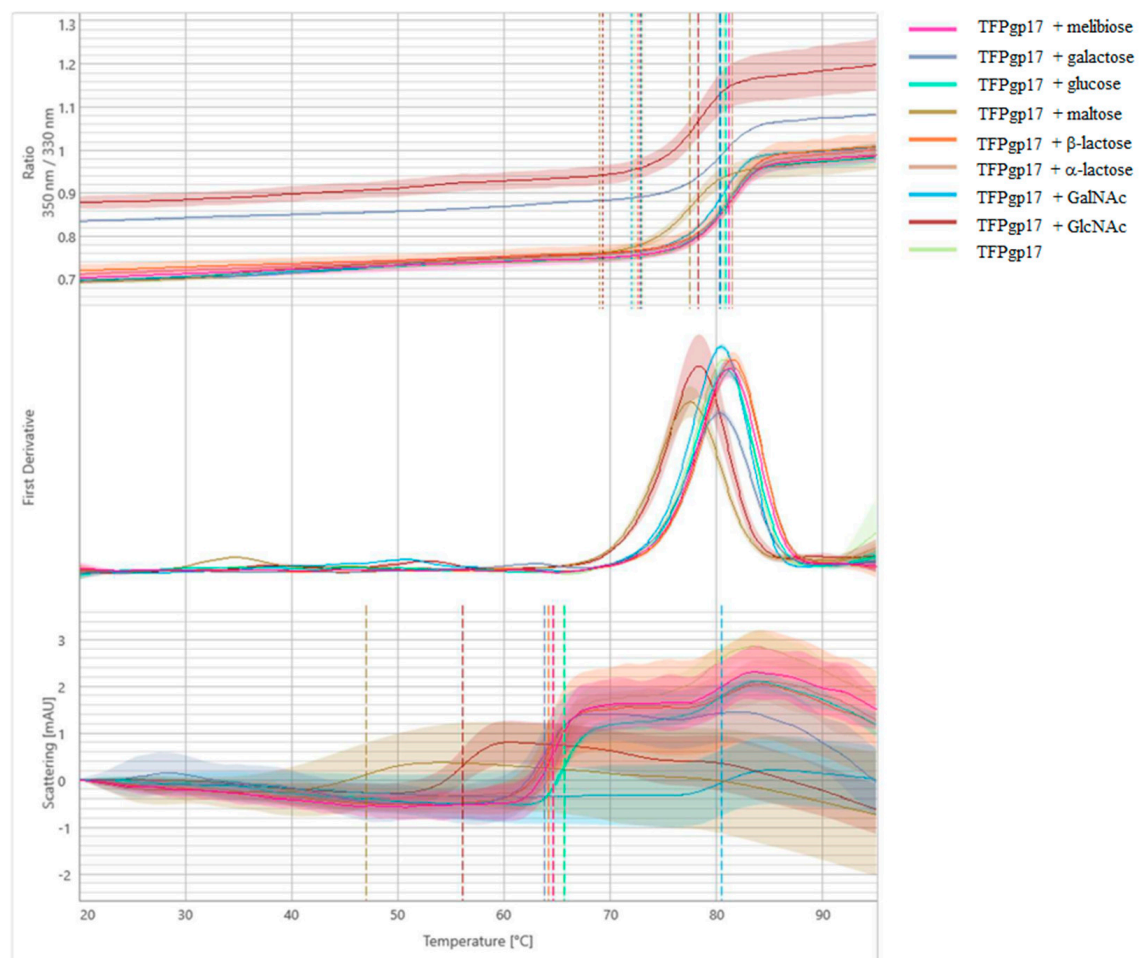


Figure 8. Melting analysis of TFPgp17 with various sugars. The fluorescence ratio (350 nm/330 nm) is shown in the top panel, the first derivative is shown in the middle panel and scattering is shown in the bottom panel. Thermal unfolding and aggregation onsets (indicated as the vertical dotted lines), as well as unfolding and aggregation transitions (indicated as the vertical dashed lines), are indicated by vertical lines in the graph.

Table 1. TTPBgp12 and TFPgp17 thermal unfolding and aggregation transitions midpoints (T_m (°C)) obtained from a melting scan of both proteins mixed with sugars, showing the fluorescence ratio 350 nm/330 nm, the first derivative onset and scattering.

Sample	T_m (°C)			
	Fluorescence Ratio 350 nm/330 nm		First Derivative	Scattering
	#1	#2		
TTPBgp12/melibiose	36.20	47.45	32.51	39.43
TTPBgp12/galactose	35.30	46.40	32.02	36.93
TTPBgp12/glucose	37.21	46.98	NA	33.76
TTPBgp12/maltose	37.64	NA	30.85	42.11
TTPBgp12/ β -lactose	35.13	46.96	32.46	39.19
TTPBgp12/ α -lactose	34.74	47.42	30.99	40.31
TTPBgp12/GalNAc	43.65	75.19	38.73	45.65
TTPBgp12/GlcNAc	41.66	NA	28.23	56.61
TTPBgp12	36.04	46.46	33.13	37.89
TFPgp17/melibiose	81.16	NA	72.83	64.64
TFPgp17/galactose	80.30	NA	72.88	63.78
TFPgp17/glucose	80.86	NA	72.89	65.71
TFPgp17/maltose	77.49	NA	69.03	46.99
TFPgp17/ β -lactose	81.45	NA	72.57	64.19
TFPgp17/ α -lactose	81.51	NA	73.01	64.66
TFPgp17/GalNAc	80.41	NA	71.96	80.45
TFPgp17/GlcNAc	78.29	NA	69.28	56.06
TFPgp17	80.70	NA	72.15	65.55

Each sample contained 0.25 mg/mL of the tail protein and 0.15 mg/mL of the sugar.

TTPBgp12 was shown to have two unfolding events (at ~36 and 46 °C), suggesting that the protein is a dimer. We have observed that sugars overall were shown to have a different effect on the general stability of TTPBgp12, which aggregated in the presence of all the tested sugars. Interestingly, GlcNAc and GalNAc significantly increased the temperature at which the aggregation was detected, to ~56 and ~45 °C, by GlcNAc and GalNAc, respectively. It is worth noting that maltose and GlcNAc shift two unfolding events to one global unfolding. Interestingly, in the crystal structure of T7_TTPgp12 (the structural homolog of TTPBgp12), four domains were determined [13]. If we assume that TTPBgp12 adopts the spatial structure very similar to T7_TTPgp12, then it is possible that some domains unfold at ~36 and other domains unfold at ~46 °C. We think that the nozzle tip domain is the most distal and thus a separate domain. Therefore, it could unfold independently of the other domains.

TFPgp17 was shown to have a higher stability than TTPBgp12 with the one unfolding event at ~80 °C. Interestingly, the aggregation of TFPgp17 precedes the main unfolding event (at ~65 °C). It was observed that sugars can play various roles in the overall stability of TFPgp17. Among the sugars tested, maltose and GlcNAc were shown to have the highest effect on TFPgp17 aggregation. The sugars caused a large decrease in the observed temperature, to ~47 and ~56 °C, by maltose and GlcNAc, respectively. The presence of GalNAc increased the temperature of aggregation to ~80 °C, leading to the protein unfolding and aggregation events at the same time. NanoDSF and backreflection technology allowed us to determine the effect of the protein concentration on its stability (Figures 9 and 10 and Table 2). The analysis of thermal unfolding profiles of the concentration gradient of TTPBgp12 and TFPgp17 clearly showed that with decreasing concentrations, both proteins increased in stability. As previously mentioned, the unfolding of individual domains of TTPBgp12 can even be detected at a concentration of ~0.25 mg/mL and higher. It is worth noting, that at the same time, a decrease in protein concentration can cause an increase in aggregation, for both TTPBgp12 and TFPgp17. These inverted patterns suggest that the optimization of the amount of protein in the buffer solution might help increase purification yield of the protein. It is more prone to aggregate when diluted, which is

the same condition presented during the flow through the chromatography columns. Additionally, for long-term storage it might be more optimal to lower the aggregation under unfolded conditions and establish refolding conditions to keep higher concentrations of functional folded protein in the experimental assays.

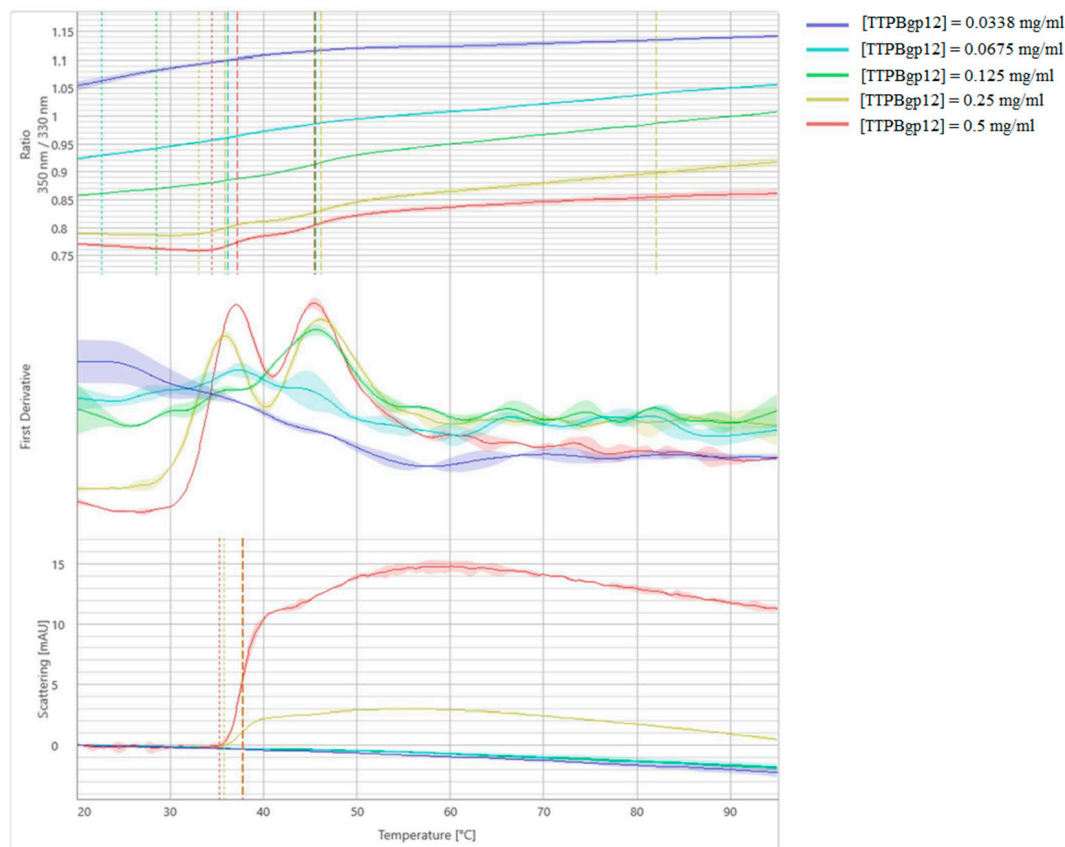


Figure 9. Melting analysis of the concentration gradient of TTPBgp12. The fluorescence ratio (350 nm/330 nm) is shown in the top panel, the first derivative is shown in the middle panel and scattering is shown in the bottom panel. Thermal unfolding and aggregation onsets (indicated as the vertical dotted lines), as well as unfolding and aggregation transitions (indicated as the vertical dashed lines), are indicated by vertical lines in the graph.

Table 2. TTPBgp12 and TFPgp17 thermal unfolding and aggregation transitions midpoints (T_m (°C)) obtained from a melting scan of concentration gradient of both proteins, showing the fluorescence ratio (350 nm/330 nm), the first derivative onset and scattering.

Sample		T_m (°C)			
		Fluorescence Ratio 350 nm/330 nm		First Derivative	Scattering
Protein	Concentration (mg/mL)	#1	#2		
TTPBgp12	0.0338	NA	NA	NA	NA
	0.0675	NA	NA	NA	NA
	0.125	45.58	NA	28.50	NA
	0.25	35.87	46.13	33.03	37.68
	0.5	37.17	45.46	34.46	37.79
TFPgp17	0.0338	71.30	NA	59.45	NA
	0.0675	76.80	NA	66.78	NA
	0.125	79.70	NA	72.49	NA
	0.25	80.54	NA	72.18	65.34
	0.5	80.67	NA	72.24	64.50

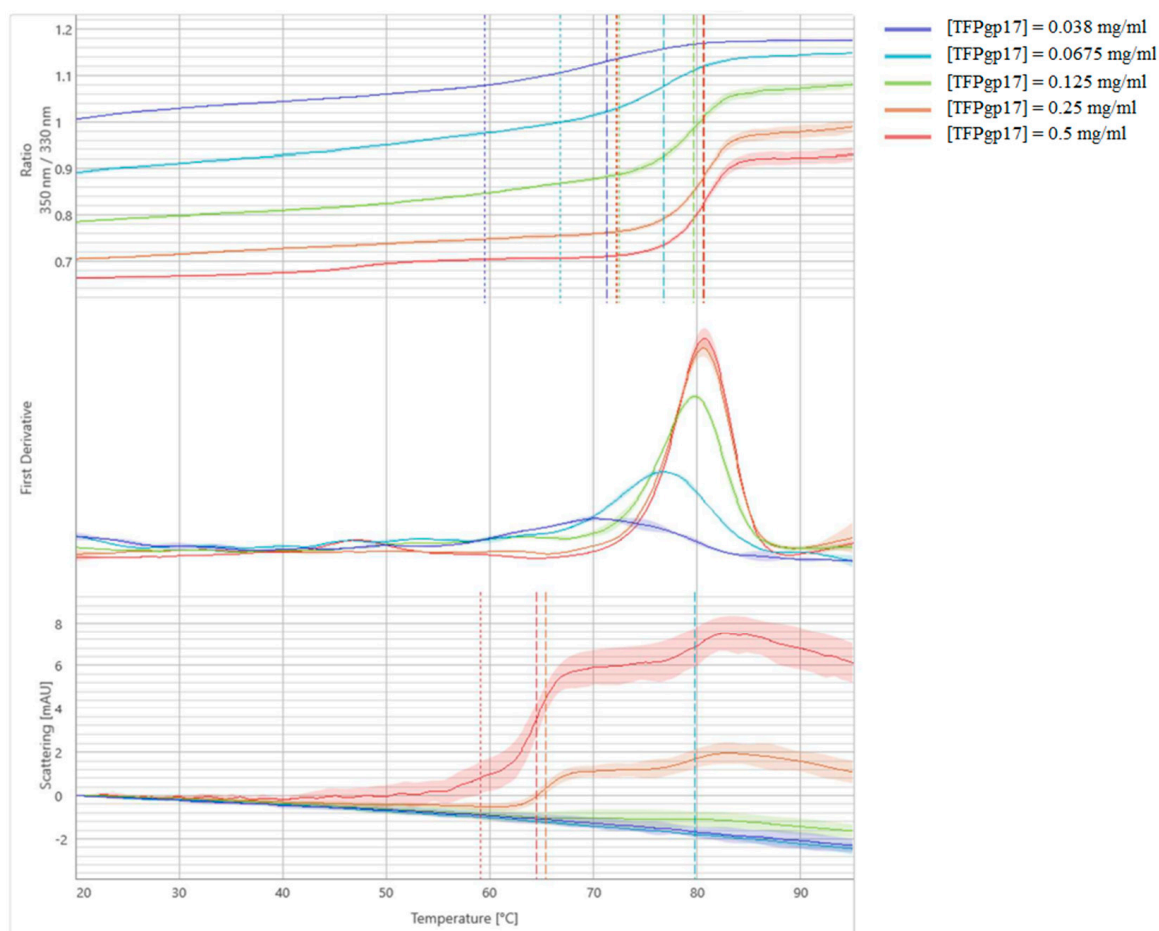


Figure 10. Melting analysis of the concentration gradient of TFPgp17. The fluorescence ratio (350 nm/330 nm) is shown in the top panel, the first derivative is shown in the middle panel and scattering is shown in the bottom panel. Thermal unfolding and aggregation onsets (indicated as the vertical dotted lines), as well as unfolding and aggregation transitions (indicated as the vertical dashed lines), are indicated by vertical lines in the graph.

3.4. Influence of TTPBgp12 and TFPgp17 on Biofilm Formation

As previously mentioned, tail phage proteins display a structural function, but they can play additional biological roles as well [20,27,28]. The purified TTPBgp12 and TFPgp17 were used in experiments regarding their effect on the bacterial *Yersinia enterocolitica* O:3 cells growth and biofilm formation. During bacteria incubation, TTPBgp12 or TFPgp17 was shown to have an inhibiting effect on bacterial culture growth. The effect was visible both in the bacterial suspension and during biofilm formation compared to the control samples (Figure 11).

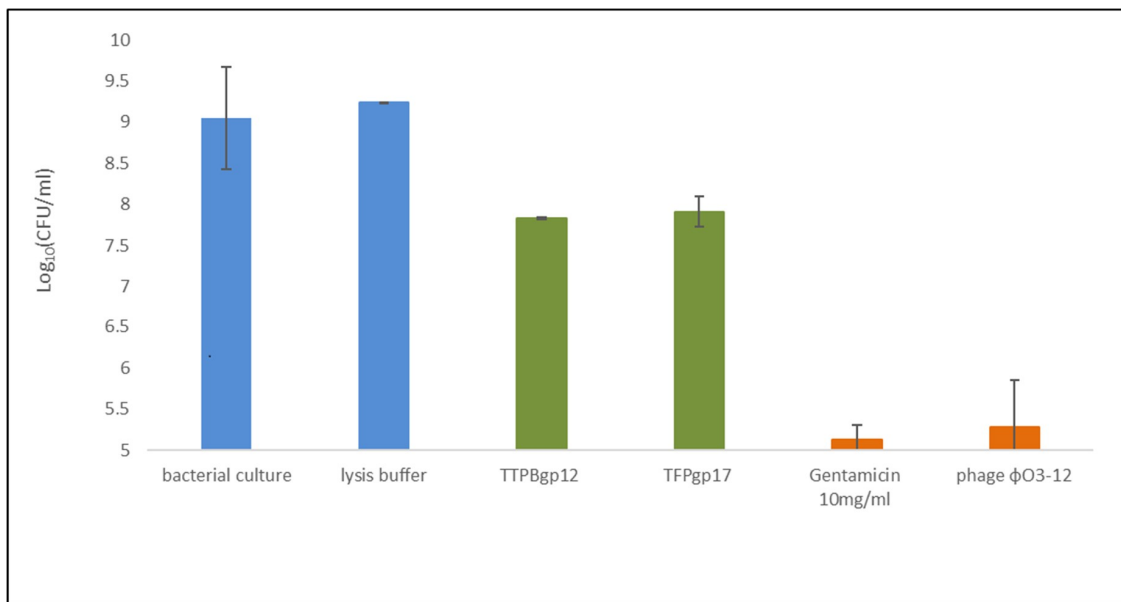


Figure 11. The effect of TTPBgp12 and TFPgp17 (in green) on *Yersinia enterocolitica* O:3 growth and the biofilm formation. The negative control (in blue) was lysis buffer (buffer in which proteins were suspended during the test) and the positive control (in orange) were gentamicin (10 mg/mL) and phage ϕ YeO3-12. The untreated bacterial culture was shown in yellow. All results were presented as averages of results from three independent replicates in three parallel trials. Error bars represent the means, standard deviations. CFU, colony-forming unit.

Moreover, the visualization of biofilm layers stained with crystal violet confirmed this observation (Figure 12).

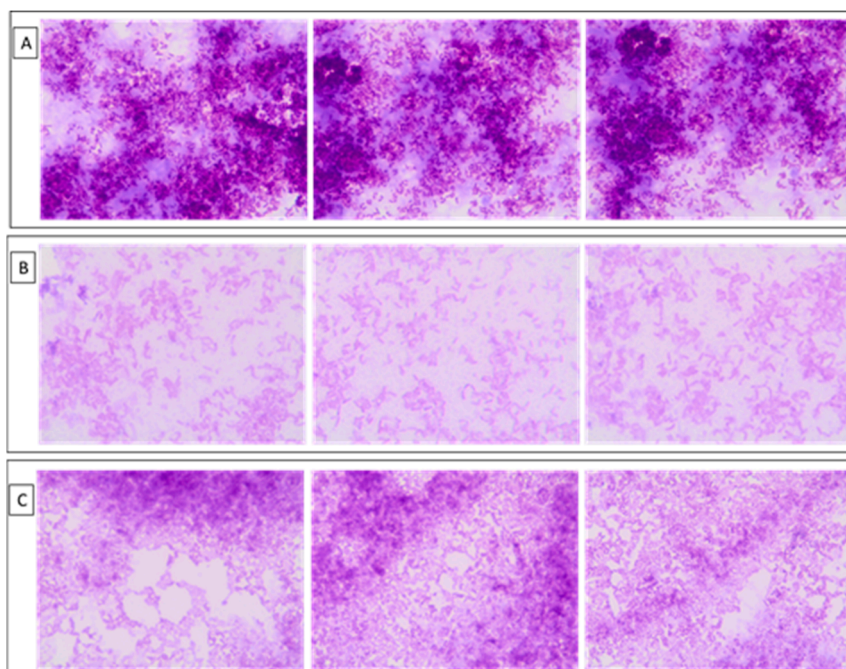


Figure 12. The effect of TTPBgp12 and TFPgp17 on the biofilm formation observed in a light microscope Olympus BX43 (magnitude 100 \times). The biofilm formed by *Yersinia enterocolitica* O:3 was stained with 1% crystal violet. (A) The control sample (*Yersinia enterocolitica* O:3 alone). (B) The bacteria incubated with TTPBgp12. (C) The bacteria incubated with TFPgp17.

The obtained results showed that both TTPBgp12 and TFPgp17 were able to inhibit the growth of bacteria and could influence biofilm formation. This observation could open a wide field for further research towards explaining the mechanism of this phenomenon. In natural conditions, both TTPBgp12 and TFPgp17 were the part of the tail phage machinery involved in bacterial cells infection. The role of TTPBgp12 is still undetermined. However, it was shown that this protein can bind to bacterial cells after specific recognition by TFP. TTPBgp12 can limit bacterial communication with the external environment, decreasing the flow between the inside and the outside of the bacteria. In contrast, TFPgp17 was observed to be responsible for recognition and binding to the host receptors. It is most likely that the binding is reversible until the TTPAgp11 and TTPBgp12 come into action. Moreover, the bioinformatics analysis allows us to speculate that TFPgp17 itself could even be incorporated into the bacterial cell wall, resulting in the formation of pores. This could increase the cell membrane permeability of bacteria, efflux of a substances, and consequently even bacterial cell death. For now, these are our assumptions. Further research will allow us to find out the answers to our questions.

4. Conclusions

The goal of this work was to characterize the biochemical, functional structure, as well as the antibacterial activity of new members of the tail tubular protein B (TTPB) and tail fiber protein (TFP) families. We have chosen TTPBgp12 and TFPgp17, previously classified as tail phage proteins. The UniProt database was lacking on the information of the transcript, and homology levels of these proteins. Bioinformatic tools were used to analyze the similarity of the selected proteins to other members of their families. TTPBgp12 showed a high similarity of primary (69%) and spatial structure to the tail tubular protein of *Enterobacteria* phage T7.

TFPgp17 did not display a high similarity to other known TFPS. However, its N-terminal amino acid sequence is much more similar to corresponding fragments of other TFPS than its C-terminal regions. The C-terminus of TFPgp17 contains an intramolecular chaperone domain, which is common among the TFP family. The overall fold of the predicted 3D structure of TFPgp17 is very similar to S-layer proteins responsible for the stability of the bacteria cells.

For the first time, the genes of both proteins were overexpressed in *E. coli* cells and the proteins were purified by nickel-affinity and gel-filtration chromatography. The thermal unfolding experiments showed that TTPBgp12 is a two-domain protein in contrast to the single-domain protein, TFPgp17. TFPgp17 is a much more stable protein than TTPBgp12 which is indicated by the unfolding temperature of ~80 °C for TFPgp17, while the unfolding temperatures for the domains in TTPBgp12 are ~36 and 46 °C.

The presence of sugars affects the stability of both proteins causing their aggregation. It has also been observed that the stability of TTPBgp12 and TFPgp17 increases when of their concentrations are decreased. However, decreasing the protein concentration causes an increase in its aggregation, both for TTPBgp12 and TFPgp17. The incubation of TTPBgp12 or TFPgp17 with *Yersinia enterocolitica* O:3 inhibited bacterial cell growth, as well as impacted the biofilm formation. These results revealed the antibacterial activity of both tested phage proteins. The mechanism of this activity is not yet known. We hypothesize that TTPBgp12, which acts as an adhesive, may limit the uptake of nutrients in bacteria culture. TFPgp17, similar to other S-layer proteins, can form pores in the bacterial membrane, resulting in an increase in permeabilization and transport interruption. More studies are needed to further understand the structure and function of TTPBgp12 and TFPgp17. Our results suggest that both TTPBgp12 and TFPgp17 could potentially be used as a biofilm inhibiting factor.

Supplementary Materials: The following are available online: the BLAST amino acid sequence analysis, the HHpred analysis.

Author Contributions: A.P.: conceptualization, methodology, results analysis, writing and editing; E.B.: conceptualization, methodology, results analysis, writing and editing, revision, supervision and funding acquisition; K.F.: methodology, results analysis, experiments conduction, revision; B.S.-O.: methodology, experiment conduction, results analysis, revision; A.C.: experiment conduction, results analysis. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Center (Poland), grant number 2017/26/E/NZ1/00249.

Acknowledgments: We would like to thank Mikael Skurnik for providing the phage, Piotr Chmielewski for providing access to the CD spectrophotometer, Katherine Tyrlik for language correction, and the NanoTemper Company for allowing us to perform protein folding and aggregation measurements.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ackermann, H.W.; Prangishvili, D. Prokaryote viruses studied by electron microscopy. *Arch. Virol.* **2012**, *157*, 1843–1849. [[CrossRef](#)] [[PubMed](#)]
2. Pajunen, M.I.; Kiljunen, S.J.; Söderholm, M.E.; Skurnik, M. Complete genomic sequence of the lytic bacteriophage phiYeO3-12 of *Yersinia enterocolitica* Serotype O:3. *J. Bacteriol.* **2001**, *183*, 1928–1937. [[CrossRef](#)] [[PubMed](#)]
3. Bottone, E.J. *Yersinia enterocolitica*: The charisma continues. *Clin. Microbiol. Rev.* **1997**, *10*, 257–276. [[CrossRef](#)]
4. Bottone, E.J. *Yersinia enterocolitica*: Overview and epidemiologic correlates. *Clin. Microbes. Infect.* **1999**, *1*, 323–333. [[CrossRef](#)]
5. Sabina, Y.; Rahman, A.; Ray, R.C.; Montet, D. *Yersinia enterocolitica*: Mode of transmission, molecular insights of virulence, and pathogenesis of infection. *J. Pathog.* **2011**. [[CrossRef](#)]
6. Wang, X.; Li, Y.; Jing, H.; Ren, Y.; Zhou, Z.; Wang, S.; Wang, L. Complete genome sequence of a *Yersinia enterocolitica* “Old World”(3/O: 9) strain and comparison with the “New World”(1B/O: 8) strain. *J. Clin. Microbiol.* **2011**, *49*, 1251–1259. [[CrossRef](#)]
7. Cover, T.L.; Aber, R.C. *Yersinia enterocolitica*. *N. Engl. J. Med.* **1987**, *321*, 16–24. [[CrossRef](#)]
8. Fàbrega, A.; Vila, J. *Yersinia enterocolitica*: Pathogenesis, virulence and antimicrobial resistance. *Enfermedades Infecciosas Microbiología Clínica* **2012**, *30*, 24–32. [[CrossRef](#)] [[PubMed](#)]
9. Pinta, E.; Duda, K.A.; Hanuszkiewicz, A.; Kaczyński, Z.; Lindner, B.; Miller, W.L.; Hyytiäinen, H.; Vogel, C.; Borowski, S.; Kasperkiewicz, K.; et al. Identification and role of a 6-deoxy-4-keto-hexosamine in the lipopolysaccharide outer core of *Yersinia enterocolitica* serotype O:3. *Chem. A Eur. J.* **2009**, *15*, 9747–9754. [[CrossRef](#)]
10. al-Hendy, A.; Toivanen, P.; Skurnik, M. Lipopolysaccharide O side chain of *Yersinia enterocolitica* O:3 is an essential virulence factor in an orally infected murine model. *Infect. Immun.* **1992**, *60*, 870–875. [[CrossRef](#)] [[PubMed](#)]
11. Skurnik, M.; Venho, R.; Bengoechea, J.A.; Moriyón, I. The lipopolysaccharide outer core of *Yersinia enterocolitica* serotype O:3 is required for virulence and plays a role in outer membrane integrity. *Mol. Microbiol.* **1999**, *31*, 1443–1462. [[CrossRef](#)] [[PubMed](#)]
12. Erridge, C.; Bennett-Guerrero, E.; Poxton, I.R. Structure and function of lipopolysaccharides. *Microbes. Infect.* **2002**, *4*, 837–851. [[CrossRef](#)]
13. Cuervo, A.; Fabrega-Ferrer, M.; Machon, C.; Conesa, J.J.; Fernandez, F.J.; Perez-Lique, R.; Perez-Ruiz, M.; Pous, J.; Vega, M.C.; Carrascosa, J.L.; et al. Structures of T7 bacteriophage portal and tail suggest a viral DNA retention and ejection mechanism. *Nat. Commun.* **2019**, *10*, 3746–3756. [[CrossRef](#)]
14. Molineux, I.J. No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. *Mol. Microbiol.* **2001**, *40*, 1–8. [[CrossRef](#)] [[PubMed](#)]
15. Garcia-Doval, C.; van Raaij, M.J. Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. *Prac. Natl. Acad. Sci. USA* **2012**, *109*, 9390–9395. [[CrossRef](#)] [[PubMed](#)]
16. Veessler, D.; Cambillau, C. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiol. Mol. Biol. Rev.* **2011**, *75*, 423–433. [[CrossRef](#)]
17. Muhlenhoff, M.; Stummeyer, K.; Grove, M.; Sauerborn, M.; Gerardy-Schahn, R. Proteolytic processing and oligomerization of bacteriophage-derived endosialidases. *J. Biol. Chem.* **2003**, *278*, 12634–12644. [[CrossRef](#)]

18. Schulz, E.C.; Dickmanns, A.; Urlaub, H.; Schmitt, A.; Muhlenhoff, M.; Stummeyer, K.; Schwarzer, D.; Gerardy-Schahn, R.; Ficner, R. Crystal structure of an intramolecular chaperone mediating triple-beta-helix folding. *Nat. Struct. Mol. Biol.* **2010**, *17*, 210–215. [[CrossRef](#)]
19. Steven, A.C.; Trus, B.L.; Maizel, J.V.; Unser, M.; Parry, D.A.D.; Wall, J.S.; Hainfeld, J.F.; Studier, F.W. Molecular substructure of a viral receptor-recognition protein. The gp17 tail-fiber of bacteriophage T7. *J. Mol. Biol.* **1988**, *200*, 351–365. [[CrossRef](#)]
20. Pyra, A.; Urbańska, N.; Filik, K.; Tyrlik, K.; Brzozowska, E. Biochemical features of the novel Tail Tubular Protein A of Yersinia phage phiYeO3-12. *Sci. Rep.* **2020**, *10*, 4196–4206. [[CrossRef](#)]
21. Eschenfeldt, W.H.; Stols, L.; Sanville Millard, C.; Joachimiak, A.; Donnelly, M.I. A family of LIC vectors for high-throughput cloning and purification of proteins. *Methods Mol. Biol.* **2009**, *498*, 105–115. [[PubMed](#)]
22. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680–685. [[PubMed](#)]
23. Louis-Juene, C.; Andrade-Navarro, M.A.; Perez-Iratxeta, C. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins* **2012**, *80*, 374–381.
24. Smith, P.K.; Krohn, R.I.; Hermanson, G.T.; Mallia, A.K.; Gartner, F.H.; Provenzano, M.D.; Fujimoto, E.K.; Goeke, N.M.; Olson, B.J.; Klenk, D.C. Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **1985**, *150*, 76–85. [[PubMed](#)]
25. Hernandez, E.; Girardet, M.; Ramisse, F.; Vidal, D.; Cavallo, J.D. Antibiotic susceptibilities of 94 isolates of *Yersinia pestis* to 24 antimicrobial agents. *J. Antimicrob. Chemother.* **2003**, *52*, 1029–1031. [[PubMed](#)]
26. NanoTemper, Version 2018; PR.ThermControl & PR.StabilityAnalysis Software: San Francisco, CA, USA, 2018.
27. Pyra, A.; Brzozowska, E.; Pawlik, K.; Gamian, A.; Dauter, M.; Dauter, Z. Tail tubular protein A: A dual—Function tail protein of Klebsiella pneumoniae bacteriophage KP32. *Scient. Rep.* **2017**, *7*, 2223. [[CrossRef](#)]
28. Brzozowska, E.; Pyra, A.; Pawlik, K.; Janik, M.; Górska, S.; Urbańska, N.; Drulis-Kawa, Z.; Gamian, A. Hydrolytic activity determination of Tail Tubular Protein A of Klebsiella pneumoniae bacteriophages towards saccharide substrates. *Sci. Rep.* **2017**, *7*. [[CrossRef](#)]
29. Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.
30. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.
31. Soding, J.; Biegerd, A.; Lupas, A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **2005**, *33*, 243–248.
32. Kemp, P.; Garcia, L.R.; Molineux, L.J. Changes in bacteriophage T7 virion structure at the initiation of infection. *Virology* **2005**, *340*, 307–317. [[CrossRef](#)] [[PubMed](#)]
33. Kelley, L.; Mezulis, S.; Yates, C.; Wass, M.; Sternberg, N. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [[CrossRef](#)] [[PubMed](#)]
34. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinform.* **2008**, *9*, 40. [[CrossRef](#)] [[PubMed](#)]
35. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic. Acids Res.* **2018**, *46*, W296–W303. [[CrossRef](#)] [[PubMed](#)]
36. Haggård-Ljungquist, E.; Halling, C.; Calendar, R. DNA sequences of the tail fiber genes of bacteriophage P2: Evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. *J. Bacteriol.* **1992**, *174*, 1462–1477. [[CrossRef](#)] [[PubMed](#)]
37. Garcia-Doval, C.; Castón, J.R.; Luque, D.; Granell, M.; Otero, J.M.; Llamas-Saiz, A.L.; Renouard, M.; Boulanger, P.; van Raaij, M.J. Structure of the receptor-binding carboxy-terminal domain of the bacteriophage T5 L-shaped tail fibre with and without its intra-molecular chaperone. *Viruses* **2015**, *7*, 6424–6440. [[CrossRef](#)]
38. Bharat, T.A.M.; Kureisai-Ciziene, D.; Hardy, G.G.; Yu, E.W.; Devant, J.M.; Hagen, W.J.H.; Brun, Y.V.; Briggs, J.A.G.; Lowe, J. Structure of the hexagonal surface layer on *Caulobacter crescentus* cells. *Nat. Microbiol.* **2017**, *2*, 17059. [[CrossRef](#)]

39. Szczesny, P.; Linke, D.; Ursinus, A.; Bar, K.; Schwarz, H.; Riess, T.M.; Kempf, V.A.J.; Lupas, A.N.; Martin, J.; Zeth, K. Structure of the head of the bartonella adhesin BadA. *PLoS Pathog.* **2008**, *4*, e1000119. [[CrossRef](#)]
40. Białas, N.; Kasperkiewicz, K.; Radziejewska-Lebrecht, J.; Skurnik, M. Bacterial cell surface structures in *Yersinia enterocolitica*. *Arch. Immunol. Ther. Exp.* **2012**, *60*, 199–209. [[CrossRef](#)]

Sample Availability: Samples of the compounds *TTFBgp12*, *TFPgp17* are available from the authors.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

New insights on the feature and function of Tail Tubular Protein B and Tail Fiber Protein of the Lytic Bacteriophage ϕ YeO3-12 specific for *Yersinia enterocolitica* Serotype O:3

Anna Pyra^{1*}, Karolina Filik², Bożena Szermer-Olearnik², Anna Czarny², Ewa Brzozowska^{2**}

¹ University of Wrocław, Faculty of Chemistry, 14 F. Joliot-Curie St, Wrocław, 50383, Poland

² Hirsfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, 12 R. Weigl St, Wrocław, 53114, Poland

Correspondence to:

* anna.pyra@chem.uni.wroc.pl

** ewa.brzozowska@hirsfeld.pl

BLAST amino acid sequence analysis [28] showing similarity of TTPBgp12 to other TPs.

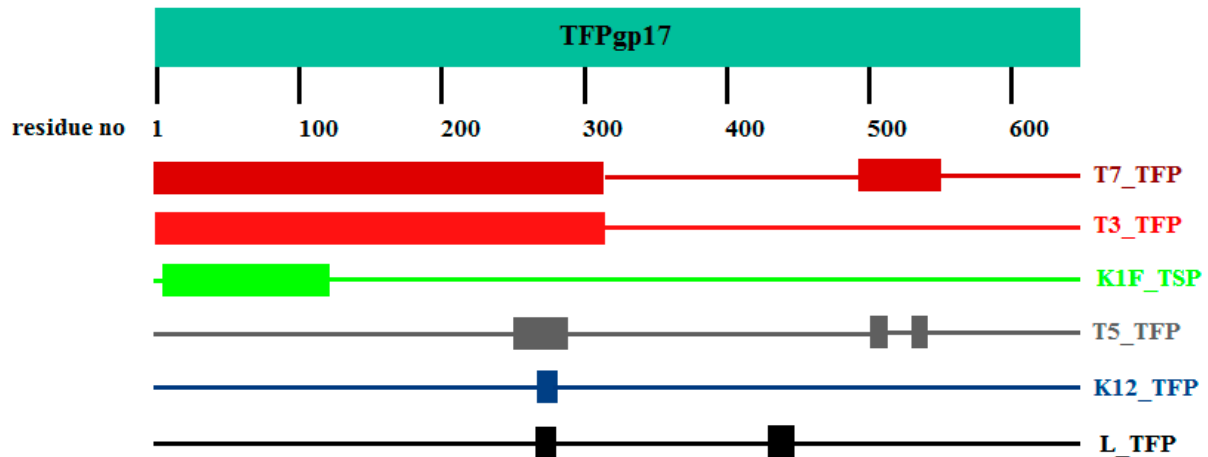
Phage Tail Protein (TP)	Phage of	% of amino acid sequence identity of TP
Tail Tubular Protein B (TTPB)	<i>Yersinia, Escherichia, Salmonella, Citrobacter, Serratia, Klebsiella, Leclercia, Enterobacteria</i>	97-99
	<i>Salmonella, Serratia, Escherichia, Enterobacteria, Erwinia, Yersinia, Pectobacterium</i>	65-70
	<i>Pectobacterium, Erwinia, Pantoea, Yersinia, Escherichia, Enterobacteria, Salmonella, Klebsiella</i>	62-65
Tail Fiber Protein (TFP)	<i>Enterobacter</i>	99-100
	<i>Escherichia, Stenotrophomonas, Klebsiella</i>	62-67

BLAST amino acid sequence analysis [29] showing similarity of TFPgp17 to other phage TFPs.

Phage of	% of amino acid sequence identity of TFP to TFPgp17
<i>Enterobacter, Yersinia, Serratia, Citrobacter, Shigella, and Escherichia</i>	99-100
<i>Dickeya</i>	92
<i>Yersinia</i>	81-84
<i>Pseudomonas and Klebsiella</i>	76 -78
<i>Dickeya</i>	68
<i>Klebsiella, Escherichia, Stenotrophomonas and Pseudomonas</i>	51-55
<i>Leclercia, Escherichia, Pectobacterium, Erwinia, Enterobacteria, Yersinia, Xylella, Klebsiella, Serratia, Kluyvera and Pseudomonas</i>	47-48
<i>Pseudomonas, Serratia, Citrobacter and Salmonella</i>	43-46
<i>Pseudomonas</i>	24

The BLAST amino acid sequence alignment scheme [28] showing similarity of TFPgp17 and TFP of *Enterobacteria* phage T7 (T7_TFP), TFP of *Enterobacteria* phage T3 (T3_TFP) of tail spike protein of *Enterobacteria* phage K1F (K1F_TSP), L-shaped TFP of *Escherichia* phage T5 (T5_TFP), side TFP of prophage K12 (K12_TFP) and TFP of *Enterobacteria* phage lambda (L_TFP).

Distribution of the top Blast Hits on 6 sequences



Sequence alignment of TTPBgp12 and T7_TTPgp12 from bacteriophage T7 (PDB code: 6r21) [13] using The Clustal Omega tool [24]. Identities between TTPBgp12 and T7_TTPgp12 are shown in red.

TTPBgp12	MALISQSIKLNKGGISQQPDILRFAEQGSVQINGNSSESEGLQKRPPMIHLKTLGPAQYV	60
T7_TTPgp12	MALISQSIKLNKGGISQQPDILRYPDQGSRQVNGWSSETEGLQKRPPLVFLNLTLDNGAL	60
TTPBgp12	GAQPYVHLINRDEFQYFVFTGEDIKVFDLDGKEYQVR--GDRSYVRATANPREDLRMIT	118
T7_TTPgp12	GQAPYIHLINRDEHEQYYAVFTGSGIRVFDLSGNEKQVRYPNGSNYIKTANPRNDRMVT	120
TTPBgp12	VADYTFVTNRKVVVQSNLQSVNLFQFQDQDALINVRGGQYGRRLSIEFNGAERAQVQLP	178
T7_TTPgp12	VADYTFIVNRNVVAQKNTKSVNLFNPNFNQDGLINVRGGQYGRELIVHINGKDVAKYKIP	180
TTPBgp12	DGSQPAHVNEVDGQAIAEKLAQLRNINLGNPNNDQDPNKWRFNVGPGFIHILAFNNDNVW	238
T7_TTPgp12	DGSQPEHVNTDAQWLAEELAKQMRNLT-----SDWTVNVQGGFIHVVTAPSGQQID	231
TTPBgp12	GLQTKDGYADQLINPVTHYTQSFQKLPINAPDGYIVKIVGDTSKTADQYVVRFDLNRKVV	298
T7_TTPgp12	SFTTKDGYADQLINPVTHYAQSFQKLPINAPDGYIVKIVGDTASKSADQYVRYVDAERKVV	291
TTPBgp12	VETIGWNTTRTHLYYHTMPWALVRASDGNFDFKWLWEGARTVGGDDTNPYPSFTGQTINDI	358
T7_TTPgp12	TETLGNTEDEQVLWETMPHALVRAADGNFDFKWLWESPKSCGDVDTNPWFPSFVGSINDV	351
TTPBgp12	FFFRNRLGFLSGENIILSRTSKYFNFPASVSNYSDDDPIDVAVSHNRVSTLKYAVPFSE	418
T7_TTPgp12	FFFRNRLGFLSGENIILSRTAKYFNFPASIANLSDDDPIDVAVSTNRRIALKYAVPFSE	411
TTPBgp12	ELLWSDQAQFVLTAAGILSSRSVELNLTQFDVQDRARPHGVGRNVYFASPRASFTSIN	478
T7_TTPgp12	ELLIWSDEAQFVLTAAGILSSRSVELNLTQFDVQDRARPHGIGRNVYFASPRASFTSIH	471
TTPBgp12	RYYAVQDVSSVKNAEDMTAHVPNYIPNGVFSISGTTAENFAAILTSGAFNRVYIYKFLYI	538
T7_TTPgp12	RYYAVQDVSSVKNAEDITSHVPNYIPNGVFSICGSGTENFCSVLSHGDPISKIFMYKFLYL	531
TTPBgp12	DEEIRQQSWSHWDFGDNVTVFAAQVINSTMVLMGNEHAVMGRHLHFTKNSIDIPGEPYR	598
T7_TTPgp12	NEELRQQSWSHWDFGENVQVLAQQSISSDMYVILRNEFNFTFLARISFTKNAIDLQGEFYR	591
TTPBgp12	LYIDAKRKYTIPAGTYNDDTYQTSISLATIYGMNFTKGRVSVVFPDGKIEVDQFINGWS	658
T7_TTPgp12	AFMDMKIRYTIPIPSGTYNDDTFTSIHIPTIYGANFGRGKITVLEPDGKITVFEQFAGWN	651
TTPBgp12	SDPVLRLDGNQEGQVVIYGFNIPFTYTFPSKFLIKKTAEDGSTATEDIGRLQLRRRAWVNYE	718
T7_TTPgp12	SDPVLRLSGNLEGRMVVIYGFNINEVYEFPSKFLIKQTADDGSTATEDIGRLQLRRRAWVNYE	711
TTPBgp12	DSGAFTIRVNNLSREFIYTMAGARLGSNLRVGRSNIGTGQYRFPVVGNAQTNLVTEISD	778
T7_TTPgp12	NSGTFDIYVENQSSNWKYTMAGARLGSNLRVGRSNIGTGQYRFPVVGNAKFNFTVYILSD	771
TTPBgp12	ASTPLNIIGCGWEGNYLRRSSGI	801
T7_TTPgp12	ETPLNIIGCGWEGNYLRRSSGI	794

The results of HHPred analysis of TFPgp17 [31].

No	Homologous template protein	Source	PDB code	% of probability	Amino acid residues of template protein	Amino acid residues of TFPgp17	The length of polypeptide chain
1	Endo-N-acetylneuraminidase (chaperone)	<i>Enterobacteria</i> phage K1F	3GW6	99.4	10-275	419-643	208
2	Receptor recognition protein	<i>Salmonella</i> phage vB_SenMS16	6F45	99.3	78-193	524-645	113
3	L-shaped tail fiber protein	<i>Enterobacteria</i> phage T5	4UW8	99.2	232-426	412-645	156
4	Neck appendage protein (chaperone)	<i>Bacillus</i> phage GA-1	3GUD	99.1	1-123	539-641	93
5	phiAB6 tailspike	unidentified phage	5JSD	99.1	21-179	2-188	156
6	Tail spike protein	<i>Acinetobacter</i> phage vB_AbaP_AS12	6EU4	98.9	569-716	500-638	125

Multiple sequence alignment of TFPgp17 and TFP from phage T7 (T7_TFP, UniProt code: P03748), TFP from phage T3 (T3_TFP, UniProt code: P10308), tail spike protein from phage K1F (K1F_TSP, UniProt code: Q04830). Identities between TFPgp17 and all other selected proteins are shown in dark blue; those between TFPgp17, T7_TFP and T3_TFP are shown in red; those between TFPgp17, T7_TFP and K1F_TSP are shown in grey; those between TFPgp17 and T7_TFP are shown in turquoise; those between TFPgp17 and T3_TFP are shown in cyan and those between TFPgp17 and K1F_TSP are shown in orange.

TFPgp17	MATIKIVMTYFLDGGSTDFNIPFEYLARKFVVRTLIGVDR---- KELILNQYRFATKT	56
T7_TFP	MANVINTVLTYLQDGSNRQFNIPFEYLARKFVVRTLIGVDR---- KVLTINTDYRFATRT	56
T3_TFP	MANVINTVLTYLQDGSNRQFNIPFEYLARKFVVRTLIGVDR---- KVLSINADYRFATRT	56
K1F_TSP	MS-CI---- T-QFFSGNQQYRIEDYLARTFVVVTLVNSSNPTLNRVLEVGROYRFINPT	54
TFPgp17	TISTTRALGFADGYTLEIRRFSDATDRLVDFDTGGSILRAYDLNISQVQTLRVAREARDL	116
T7_TFP	TISLTKANGFADGYTLEIRRVFTSTIDRLVDFDTGGSILRAYDLNVAQIQTMRVAREARDL	116
T3_TFP	TISLTKANGFADGYTLEIRRVFTSTIDRLVDFDTGGSILRAYDLNVAQIQTMRVAREARDL	116
K1F_TSP	MIEDEL---- VDQSGFDIVRHKQF-CEDLVVDPRNGSVLTAEDLTTAELQATMIAEEGRDQ	110
TFPgp17	TADTIGVNDGHLD----- ARGRRIVNLANAVDRAVFPQQL-----KIDNQNS	161
T7_TFP	TADTIGVNDGHLD----- ARGRRIVNLANAVDRAVFPQQL-----KIDNQNS	161
T3_TFP	TADTIGVNDGHLD----- ARGRRIVNLANAVDRAVFPQQL-----KIDNQNS	161
K1F_TSP	TVDLAHEYADAAGSAGNAKDSDEARRIAESIRAA---- GLIGYITFRSFEKGYNVT	165
TFPgp17	LNSANNKQEQADPATARANDANNANASASASASSACSAELAKMHWATSD-TVVVDIEEG	220
T7_TFP	WQARNEALQFNEAETFRQAEGFKNE----- STNATNTKQRDET-KGYRDEAK--	211
T3_TFP	WQARNEALQFNEAETFRQAEGFKNE----- SGTNATNTKQRDEA-NGSRDEAE--	211
K1F_TSP	WSE---- VLLMEEDGDYYRMDGLPKQVFAQSTPETSOGIG-LGAWVSVGDAAALRQQISN-	220
TFPgp17	RTYALHSMELYNETKDADRAAVSETHAK-- ASECGAANSAAAKVSETNAKA-----	271
T7_TFP	-RPFNTAGQYATSAAGNSAS----- AAH--QSEVNAENSATASANSALAEQ-----	254
T3_TFP	-QFNTAGQYATSAAGNSAS----- TAT--QSEVNAENSATADANSALAEQ-----	254
K1F_TSP	---- PEGAILYFELHEAR----NLDHNDARGHGAKDGVYDDTALTSALADTFVQGKING	273
TFPgp17	-SEERAITKASKLQDSNDFAAATSEVTVGNDVQSGQAVSSPGNITG-- GGLVSTGAASIQK	328
T7_TFP	-QADRAREADKLEWYNGLAGAIDKVDGTVYNYKGNIEANGRLYMTNDFDC-GQYQOFF	312
T3_TFP	-HADRLLEADKLGIFNGLAGRIDKGDGTVYNYKGNIEANGRLYMTNDFDC-GQYQOFF	312
K1F_TSP	NGKTYKVTSLFDISRFINTRFVYERIPQGPLYVYSEEFVQGLFKITDTFY-----	324
TFPgp17	GALVGEDLIVCRDITAKQDMYSQR----- NIAVAGTYAQGGIEQTLATNIYNK	377
T7_TFP	CGVTHRYSVSWGCDENGCMWVQPRRSTALGCGNIQLVYNGQIITQCCMHPGQLK----	367
T3_TFP	CGSACRYSVSWGIEKAWLVNFRFRERTALVDNIQLVYNGHIAQGGDMHPGFLK----	367
K1F_TSP	----- YNARPDQKA-----FVYENI-----	340
TFPgp17	LYRLHINNPQKVGQRQGLHIGNESGSGESNFTINRAGSGGFVFRVNAENSVEVTRVY	437
T7_TFP	----- LQNGHVLQLESAEDHAHYILSKDGNRNNWHIGRSDN-----NN	406
T3_TFP	----- LQNGHLYLESAEDHAQYILSKDGNRNNWHIGRSDN-----NN	406
K1F_TSP	-LAFINGGD-- RHGVSRLHVSIVKSDDGGQVTSVF-----EWTLDLHPDY--FTVNY	387
TFPgp17	DITGGQDIY----- ANILQRSARIE-----	459
T7_TFP	DCTFHSYVH----- GTLTLKQDYAVV-----	428
T3_TFP	DCTFHSYVY----- GTLTLKQDYAVV-----	428
K1F_TSP	HCMSHGVCRNRLFAMIESTRLANALTNCALNDRPHMSLHLTGQITKAAHQYATINVF	447
TFPgp17	GNNNIVCQ----- NLYAGGSDMFEQGNLTGGINA----	490
T7_TFP	NKH----- FHYGQAVVATDQNIQGT-----	449
T3_TFP	NKR----- FHYGQAVVATDQNIQGT-----	449
K1F_TSP	DHGLFVCDFVNFNSAVTVGSGMTVATVIDKDNFTVLTFNQQTSDLNAGKQWHDGTSF	507
TFPgp17	---QWGNLMSGLNNSLFAKFFGGQLTAR-GGYLLEKRDQI--- AVGTRMFSQR	541
T7_TFP	---NCGKGLDAYLRDSFVAKSKANTQVMSGS-ACGGVSVT/SQDL--- RFRNIWIKCAN	501
T3_TFP	---NCRKGLDAYLRDTPYKKTMAHTQVMAAA-SDSVMGGGQ--- TDTLHRTCS	497
K1F_TSP	HKSPKRTDLGLIPS----- VTEVHSFATIDNNGFAMVYHQGVAFREVGLFYFDAAF	560
TFPgp17	R----- IKEDIKVVRSACDMLNIIRSYIPVSYKYKQDASVTDNQRNTNIEGKSRAGFIT	596
T7_TFP	N---SNNFFRTGPDG----- T-----YFTASDGGNLRPQIMS-----NGLGFFN	537
T3_TFP	A--TYGIRPFTTIGTISELVLV----- STSFQPRGLKQFIMS-----NGRVFQN	541
K1F_TSP	NSFSNYVRRQFSETEFQASEPCIKYIDGL----- LITROT---RCDLGGSSLHR	609
TFPgp17	Q-ELIRIIFRAVDVMSDGGQSPFNQIICGLMLLVNLDARIQLEKDKT-----	645
T7_TFP	I-ADSRVFNALISVENE-----	553
T3_TFP	I-ADRAATPTAIAVEDV-----	557
K1F_TSP	SRLIQGTKSLR-FPHDVRHTTLFAVUCDLIMFGSERAENKWEAGAFDDRYKASTPR-	667

Multiple sequence alignment of TFPgp17 and L-shaped TFP from phage T5 (T5_TFP, UniProt code: P13390), side TFP from prophage K12 (K12_TFP, UniProt code: P76072) and TFP from phage lambda (L_TFP, UniProt code: P03764). Identities between TFPgp17 and all other selected proteins are shown in dark blue; those between TFPgp17, T5_TFP and K12_TFP are shown in turquoise; those between TFPgp17, K12_TFP and L_TFP are shown in red; those between TFPgp17, T5_TFP and L_TFP are shown in grey; those between TFPgp17 and T5_TFP are shown in pink; those between TFPgp17 and K12_TFP are shown in orange and those between TFPgp17 and L_TFP are shown in purple


TFPgp17	MATTIKTVMYFLDGSSTDFNIPFEYLAKF--VRVTLIGVDRK-EL-ILNQCYRFAFK	55
T5_TFP	-----MGI--TK	5
K12_TFP	SAVKISGVVKDGT-GKFPVQ-NCTIQLKAKRNSTVVVNTLASENPDLAGRYSMDVEYGY	58
L_TFP	SAVKISGVVKDGT-GKFPVQ-NCTIQLKAKRNSTVVVNTVQSENPDLAGRYSMDVEYGY	58
TFPgp17	TTISTTRALGFADGYT--LIEIRRFPSATDRLVDFT---DGSILRAYDINIQQVQLHV	109
T5_TFP	-IIL-----QQKVTMDQNSITAKYFKITVVLKNSIS---	36
K12_TFP	SVILLVEGFFPSHAGTITVYEDSQPOTLNDFLCAMTEDDARPEALRFEIMVEEVARNAS	118
L_TFP	SVILLQVDFGFFPSHAGTITVYEDSQPOTLNDFLCAMTEDDARPEALRFEIMVEEVARNAS	118
TFPgp17	AEEARDLTADTIGVNDGNLDAIGRRIVNVAQAQDVGDAI--NLGQIQIRIND-----	159
T5_TFP	-----SITA---ADVTSAIESKA---SGFAAKQS-----	60
K12_TFP	AVA--QNTAAAKSASDASTSARE-AATHAADADARAASSTAGQAASQAASASASAGT	175
L_TFP	VVA--QSTADAKKSAGASASAAQ-VAALVTDATDSARAASSTAGQAASQAASASAGT	175
TFPgp17	-----SALNSANRA--KQEAADRA	175
T5_TFP	---EINAKQSELNAKDSENE-----ARISATSSQQSATQSSASATASANAKAA	106
K12_TFP	ASTKATEAKSAAAESSKSAATSAAGAATSETNASASLQSAATSASTA--TTKASEA	232
L_TFP	ASAKATEAKSAAAESSKSAATSAAGAATSETNAAASQQSAATSASTA--ATKASEA	232
TFPgp17	TARANDANONASASAS--SASSAGSAELAKRMTSDTVVESDLESSRTYALHSHLYRN	232
T5_TFP	KTSETNANQSKRAANTSETNAAASASAS-----S-----FAT	139
K12_TFP	ATSARDAAASKEAASSETNASSSASAA-----S-----SAT	265
L_TFP	ATSARDAAASKEAASSETNASSSAGRAA-----S-----SAT	265
TFPgp17	ETKDSADRAAVSETNAKASEG-----	253
T5_TFP	AAENSARAARTSETNAGNSAQAADASKTAAANSATAAK-----TSETNAK-----	184
K12_TFP	AAENSARAARTSETNARSSETAAGQASASAAAGSKTAAASASASASTSAGQASASATAAGK	325
L_TFP	AAENSARAARTSETNARSSETAARERSASAAADAKTAAA-----	303
TFPgp17	-----GAANSAAAKAVSETNAKASEERAITEASKLQGD---	286
T5_TFP	-----KSETAARTSETNAKASEERAITEASKLQGD---	217
K12_TFP	SAESAASSASTATTKAGEATEQASAAARSAARTSETNAKASETSAESKTAASASSASS	385
L_TFP	-----	303
TFPgp17	-----NDFAAAIEVTCNDV--DQKGAVSSPGNITGGGLV-STGAASIQKCALV	332
T5_TFP	VT-----QYDW--FVGTNNSVYVIAKLTDF-GAVS-CHLTL-----MITNCGNY	259
K12_TFP	AASSASASASAKDETRQAIAAKSATTASTKATEAAGSAAAQSKSTAESAAATRAETA	445
L_TFP	-----GSASTASTKATEAAGSAVASQSKSAEAAALPAONS	340
TFPgp17	GEDLIVGRDITAKQMYSQRNIAVAVTYAQGGIEQTLATINYNKLYRLHINSNQRVGG	392
T5_TFP	GSS--YGNL---GF--VEISARLNDARQVTS---ENIT-NFLSVRRLSF----	299
K12_TFP	AKR--AEDI-----ASAVALEAS-TTK---KGIV-QLSSATNSTSETLAAT	485
L_TFP	AKR--AEDI-----ASAVALEAD-TTR---KGIV-QLSSATNSTSETLAAT	380
TFPgp17	RQGLHIGH-NESSG-----SGESNFTNKG---AGSGGFVFRVYNAENSV	432
T5_TFP	---NLANDNQLRYGLV--EGDGYFEVWCYQRAIKETRVAVLAQTGR--TELYIFEGFV	351
K12_TFP	PKAVKSAVDNAENRIQDQNGADIPDKGCFLNQ--INAVSKTDFADKRIMRYVRYNAPAGA	544
L_TFP	PKAVKVVVDENRKAFLDSPALT-----	403
TFPgp17	ETRV---DITGGVIY--ANHLQVRSGLVIEG---INNIVGQNLVAC-----MGST	476
T5_TFP	-SQDTQ-----PSGFIESLAARIYDQVHKPT--KADLGLNMLVGFGLGGNGLYS	401
K12_TFP	TS-KYYPVVVRSAGSVSELSARVIITATATAGDFMNCENGFVMPGNW-----	595
L_TFP	-----GTPAPTALROT-----	415
TFPgp17	MFEGRGNLGGIWAQGNLMSGLANNLSLFAKFP-GVQLFTA-----FGGYILEGRVD	528
T5_TFP	SVQSNVDLINKLKAQQVWRAARESGARVDIMHSGSGFYSHCGDTHAAINQVINTGIV-	460
K12_TFP	-----TDRGRYAYGMFWYQNNKRAIHSIM-----MSNKGDDL--SVFTVDCAA-	638
L_TFP	-----NNT-----QI--AN-----TA-	424
TFPgp17	GTAVGFRWFQSDRRLKEDIKVVRSADDCNLTIR---SYIFVSYKYDASYTDNRGRN	583
T5_TFP	-----KVLATTURNL-----ASDIVYANTLY--GTANKPS--KSD	491
K12_TFP	-----F-----PVFAPIEDGLSISAPGADLVVNDTTYFGATNPATECIAAD	680
L_TFP	-----F-----VLAALADVIDA-SFDALNTINELAAALGN-----	453
TFPgp17	TIEGKRLRAQFI-----TQ-----DLIRLMPAVDVMSDGMQ-SFD	618
T5_TFP	VGLGNVYNDQAVYKAGDVMISGDLIRKETPSIRLKSQGNALMFPQNDGGERGVINSPF	551
K12_TFP	VLDLFFGRGFYESHSLIVNDNLSCPKLFAIDELVARGGN---QIRMIGGETGALWQND	736
L_TFP	-----DPDFATMTNALAGK-----QFK-----N-	472
TFPgp17	PKQII-----GLMLLV-----KNLARIQELEKQK-----	645
T5_TFP	NNGSLGEHIRAKTSDGTSTGDFIVRHDRIEAKDAMISYKISSRATRFSDNTNTAATN	611
K12_TFP	GAKTYLL-----TNQGVYGNNTIR-----PFALNATGLVIGTKLS---	776
L_TFP	-----A-TL-----T-----	476

ORIGINAL ARTICLE

Open Access



ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3

Karolina Filik¹, Bożena Szermer-Olearnik¹, Joanna Niedziółka-Jönson², Ewa Roźniecka², Jarosław Ciekot¹, Anna Pyra³, Irwin Matyjaszczyk⁴, Mikael Skurnik^{5,6} and Ewa Brzozowska^{1*} 

Abstract

Yersiniosis is an infectious zoonotic disease caused by two enteropathogenic species of Gram-negative genus *Yersinia*: *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. Pigs and other wild and domestic animals are reservoirs for these bacteria. Infection is usually spread to humans by ingestion of contaminated food. Yersiniosis is considered a rare disease, but recent studies indicate that it is overlooked in the diagnostic process therefore the infections with this bacterium are not often identified. Reliable diagnosis of Yersiniosis by culturing is difficult due to the slow growth of the bacteria easily overgrown by other more rapidly growing microbes unless selective growth media is used. Phage adhesins recognizing bacteria in a specific manner can be an excellent diagnostic tool, especially in the diagnosis of pathogens difficult for culturing. In this study, it was shown that Gp17, the tail fiber protein (TFP) of phage ϕ YeO3-12, specifically recognizes only the pathogenic *Yersinia enterocolitica* serotype O:3 (YeO:3) bacteria. The ELISA test used in this work confirmed the specific interaction of this protein with YeO:3 and demonstrated a promising tool for developing the pathogen recognition method based on phage adhesins.

Keywords: Phage, Yersiniosis, Tail fiber protein, Phage adhesins, *Yersinia enterocolitica*, Diagnostic, ELISA

Key points

1. TFP of ϕ YeO3-12 phage was shown to be an excellent tool for YeO:3 detection.
2. It is beneficial to leave MBP in the complex with TFP.
3. The specific interaction of H/MTFP-Gp17 and the pathogenic bacteria was shown in ELISA and TEM.

Introduction

Yersiniosis is an infection in human caused by *Yersinia enterocolitica* (Ye) bacteria. Most often the infection is caused by eating raw or undercooked pork but also dairy products contaminated by the bacteria. It was shown that

Ye is the third most common enteric pathogen responsible for food poisonings with dairy products. In the United States, Ye causes almost 117,000 illnesses, 640 hospitalizations, and 35 deaths every year. In Europe, 7048 confirmed cases of yersiniosis (caused by *Y. enterocolitica* and *Y. pseudotuberculosis*) were reported by the European Food and Waterborne Diseases and Zoonoses (FWD) Network for 2019 and was the fourth most commonly reported foodborne zoonotic disease in the European Union. Children are infected more often than adults, and the infection is more common in cooler climates. Yersiniosis is typically an enteric disease and signs may include diarrhea, weight loss, severe abdominal pain, dehydration, and bloody feces (Watkins and Fredman, 2020; Triantafyllidi et al., 2020). In all countries, the infection rate is most likely much higher since only serious cases are registered, and for many reasons the infections may be overlooked (Wielkoszynski et al., 2018). The bacteria of genus *Yersinia* are Gram-negative coccobacilli

*Correspondence: ewa.brzozowska@hirszfild.pl

¹ Hirszfild Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, 12 R. Weigl St, 53114 Wrocław, Poland
Full list of author information is available at the end of the article

that belong to the *Enterobacteriales* order, *Yersiniaceae* family. Biochemically and serologically they have been categorized into three species (*Y. enterocolitica*, *Y. pestis*, *Y. pseudotuberculosis*), which are responsible for infections in humans. *Y. enterocolitica* isolates, based on their biochemical properties have been divided into six biotypes, and on their antigenic properties, into 70 serotypes (Shoib et al., 2019). According to pathogenicity and geographical distribution, it is categorized into six distinct groups corresponding to the biotypes defined by their biochemical properties: 1A, 1B, 2, 3, 4, 5. Within these biotypes of Ye are different serotypes and some of them belong to each biota as follows: 1A (O:5; O:6, 30; O:7, 8; O:18; O:46), 1B (O:8; O:4; O:13a, 13b; O:18; O:20; O:21), 2 (O:9; O:5, 27), 3 (O:1, 2, 3; O:5, 27), 4 (O:3) and 5 (O:2,3). Typically, serotypes O:3, O:8, O:9, and O:5, 27 have been associated with virulence and cause most of the infections. In addition, there is a geographical distribution; serotypes O:4, O:8, O:13a/b, O:18, O:20, and O:21 are prevalent in USA, while serotypes O:3 and O:9 dominate in Europe and Japan (Wielkoszynski et al., 2018; Simonova et al., 2007). The serotypes are sometimes described as bio/serotypes such as 4/O:3. 4/O:3 is the most frequently isolated *Y. enterocolitica* in Europe that cause asymptomatic infections in pigs thereby contaminating the pork meat and causing human infections (Batzilla et al., 2011). YeO:3 appears to be an example of a zoonotic pathogen perfectly adapted to infect humans. It can be seen in the genomic variations of YeO:3 that streamline the physiology and metabolism of the bacteria (Schmühl et al., 2019). It is commonly known that the composition and modification of mucin is a critical defense mechanism in the prevention of pathogenic bacteria in the intestine. The amount of *N*-acetyl-*D*-galactosamine (GalNAc) is nearly twice that of any of the other sugars present in the mucin of the small intestines of pigs. In contrast, *N*-acetyl-*D*-glucosamine (GlcNAc) is the major amino sugar in human mucin. One of the adaptive features of the YeO:3 bacteria in contrast to the non-pathogenic Ye bacteria, is the ability to uptake GlcNAc and GalNAc as a source of carbon (Schmühl et al., 2019).

Ye remains a challenge for researchers and food handlers due to its ability to grow at refrigeration temperature, low concentrations in samples, morphological similarities with other bacteria and lack of rapid, cost-effective, and accurate detection methods. The recommended method of Ye isolation is carried out by using ISO 10273-2003 (Morka et al., 2018). In this protocol, body fluids (from the peritoneum, wounds, or abscesses) or stools are taken to analysis, and bacteria are inoculated on enrichment peptone sorbitol bile (PSB) broth,

irgasan-ticarcillin-potassium chlorate (ITC) broth, and cefsulodin-irgasan-novobiocin (CIN). To isolate presumptive strains of Ye, the ITC broth and CIN agar media are recommended. Colonies of *Yersinia* sp. on CIN agar have the bull's eye morphology with a red center and colorless translucent rims (Morka et al., 2018). After bacterial colony isolation, the Ye biotypes are identified via a biochemical characterization using commercial systems such as API 20E or 50CH (bioMérieux), PCR or MALDI TOF MS (Morka et al., 2018; Laporte et al., 2015). Furthermore, the determination of serotypes for enteropathogenic *Yersinia* species can be achieved using serotype-specific antisera. However, this technique is available only in specialized laboratories (Laporte et al., 2015).

The O-specific polysaccharide (also known as O-antigen) is used for the serological characterization of Ye strains. O-antigen is the outermost structure of lipopolysaccharide (LPS) that is essential for the efficient colonization and invasion of the pathogenic strains (Fàbrega and Vila, 2012; Al-Hendy et al., 1992; Kenyon et al., 2016). The O-antigen is linked via the core oligosaccharide to the lipid A of LPS. The O-antigen is a virulence factor for many bacteria (Al-Hendy et al., 1992). LPS of Ye serotype O:3 has a unique structure in which the outer core (OC) forms a branch. The lipid A moiety is abridged via 3-deoxy-*D*-manno-2-octulopyranosonic acid (Kdo) to the inner core heptose residues onto which either the OC or the O-antigen is linked (Skurnik et al., 1999; Pinta et al., 2010). OC is important for the resistance of YeO:3 to cationic antimicrobial peptides, and it functions as the receptor of bacteriophage ϕ R1-37 (Pinta et al., 2010; Leskinen et al., 2016; Leon-Velarde et al., 2019). The O-antigen of YeO:3, a homopolymer of 6-deoxy-*L*-altrose moieties, on the other hand, functions as the receptor of phage ϕ YeO3-12 (Leon-Velarde et al., 2019).

Our previous results suggested that tail fiber protein TFP-Gp17 belonging to ϕ YeO3-12 phage tail fiber contained a chaperone domain on the C-terminus of its polypeptide chain and the adhesion domain on its N-terminus end. The 3D structure of this protein was predicted using the Swiss-Model server suggesting to adopt the tertiary structure as adhesin A (PDB code: 3d9x) (Pyra et al., 2020). That might suggest the adhesive feature for this protein. In this paper, we characterize the ability of bacteria-specific binding by the TFP-Gp17. The TFP-Gp17 binds to the O-antigen of YeO:3 and can be applied to identify strains of this pathogenic serotype. As no other serotype was recognized by TFP-Gp17, the specific detection of YeO:3 bacteria with the help of TFP-Gp17 becomes fast and effective.

Materials and methods

Gene cloning, protein overexpression, purification and analysis

The annotated nucleotide sequence of the phage ϕ YeO3-12 genome is available at GenBank under the accession number AJ251805 (Pajunen et al., 2000). The phage genomic DNA was isolated from the phage lysate using a viral DNA extraction kit (Biocompare) and used as a template (20 ng) in the PCR reaction. The TFP-Gp17 encoding gene, g17, was amplified as a 1937 bp fragment by PCR using primers: TFPgp17_FW 5'-TACTTCCAA TCCAATGCCATGGCTACAACATTAAGACCG and TFPgp17_RV 5'-TTATCCACTTCCAATGTTACTAAG TCTTGTCCTTCTCCAAC. The ligation-independent cloning method was used to clone the PCR fragment into the pMCSG9 vector using a T4 DNA polymerase (Eschenfeldt et al., 2009). This way the protein coding sequence would be fused to hexa-His tag—MBP (Maltose Binding Protein). MBP is one of the most popular fusion components for recombinant proteins produced in a bacterial expressing system. MBP facilitates the proper folding and solubility of the target proteins, increasing the effectivity of proteins production (Lebendiker and Danieli, 2010). The construct was transformed into *E. coli* DH5 α cells using the heat-shock method, and confirmed by sequencing. The obtained plasmid pMCSG9-6HMBP-ypQ9T0Z9 was transformed into competent *E. coli* BL21(DE3)pLysS.

The *E. coli* BL21(DE3)pLysS/pMCSG9-6HMBP-ypQ9T0Z9 bacteria were inoculated into Luria–Bertani (LB) medium supplemented with ampicillin and chloramphenicol, at 100 and 25 μ g/ml, respectively. The bacteria were grown at 37 °C with shaking 120 rpm to an OD600 of 0.7 and the gene expression was induced by addition of isopropyl β -D-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM, and the bacteria were then incubated overnight at 18 °C. The bacteria were harvested by centrifugation (5000g, 5 min) and sonicated 10 times (30-s pulses separated by 15-s breaks) in buffer containing 20 mM Tris/HCl buffer, pH 8.0, 300 mM NaCl, 5% glycerol and 5 mM β -mercaptoethanol (buffer A). The cell disruption by sonication was performed on ice using a UP200S ultrasonic disintegrator (Dr. Hielscher GmbH). The cell debris was pelleted and the supernatant was treated with viscolase (AA&Biotechnology) to reduce the viscosity of the bacterial lysate and filtered through a 0.45 μ m filter. Then supernatant was loaded onto Super Nickel NTA Affinity Resin (Protein Ark) equilibrated with buffer A. TFP-Gp17 was purified using two rounds of nickel-affinity chromatography. Unbound proteins were washed with buffer A and TFP-Gp17 was eluted with buffer A containing 250 mM imidazole. Before the second round, the eluted protein fraction

was precipitated with 0,65 g/ml ammonium sulfate and harvested by centrifugation (20,000g, 4 °C, 45 min) and resuspended in buffer A. To remove the MBP and the his-tag the protein solution was digested by TEV protease overnight at 4 °C. After digestion, the tag-free protein appeared in the unbound fraction of proteins (flow through) during nickel-affinity chromatography. The flow through fraction was again precipitated by addition of 0,65 g/ml ammonium sulfate overnight in 4 °C and harvested by centrifugation (20,000g, 4 °C, 45 min). The Knauer system of chromatography was used for protein purification by affinity chromatography. The system was controlled using the Purity Chrome software.

Size Exclusion Chromatography (SEC) was performed using the Dionex Ultimate 3000 System (Dionex Corporation, USA) equipped with LPG-3400SD pump, WPS-3000T(B) FC Analytical autosampler, TCC-3000SD column compartment and a DAD-3000 diode array detector. Protein separation was performed using Superdex[®]200 10/300 GL Code No 17-5175-01 Id No 0710071 (GE Healthcare). Nominal separation range for this column is 10–600 kDa (globular proteins). A PBS buffer (pH 7.2) was used as the mobile phase. The flow rate was 0.9 ml/minute and the eluent was monitored at 220 nm at room temperature. The inject sample volume was 200 μ l. Control of the system, and data acquisition and treatment were performed using Chromeleon software (Dionex).

A Dionex 3000 RS-HPLC equipped with a DGP-3600 pump, a WPS-3000 TLS TRS autosampler, a TCC-3000 RS column compartment (Dionex Corporation, USA) and a Bruker micrOTOF-QII mass spectrometry as a detector (Bruker Daltonics, Germany) were used to determine the molecular weight (MW) of TFP-Gp17. The chromatography column was a 100 \times 1 (i.d)-millimeter Thermo Scientific BioBasic-8 with 5-micron particles (Part No. 72205-101030, Serial No. 10158875). The injected sample volume was 2 μ l. The flow rate was 0.1 ml/minute and the eluent was monitored using mass spectrometry. The mobile phase: solvent A—0.1% formic acid in water and solvent B—0.1% formic acid in acetonitrile. The ramp: 0 min—5%B, 1 min—5% B, 16 min—95% B, 17 min—95% B, 17.1 min—5% B. The mass spectrometer was calibrated at the beginning of each run with 10 mM sodium formate and the following settings in positive ESI mode were used. Scan range: 300–3000 m/z, End plate offset: – 500 V, Capillary: – 4000 V, Nebulizer gas (N2): 1 bar, Dry gas (N2): 8 L/min, Dry Temperature: 180 °C.

The protein sample was analyzed via 12% SDA-PAGE (Laemmli, 1970) and the concentration was determined by the nanodrop at 280 nm absorbance (Denovix) and by the BCA method (Smith et al., 1985).

Bacterial cell-based sandwich ELISA

For the bacterial cell-based sandwich ELISA the protocol by Paton et al. (2001) with some modification was used. The bacterial strains used for the experiment are shown in Table 1. Firstly, the bacterial culture was incubated overnight (O/N). Next day, bacterial cultures were refreshed and incubated at 37 °C and 28 °C with shaking 120 rpm until the OD600 reached 0.5. When the OD600 reached the appropriate value, the cultures were washed with PBS-T (PBS supplemented with 0,1% Tween 20) by centrifugation (2000×g, 5 min) to get rid of the medium. The bacterial pellet was resuspended in fresh PBS and diluted tenfold. The bacterial suspension was added to Maxisorp 96-well microplates for overnight at 4 °C. After O/N incubation, plate was centrifuged (600×g, 4 °C, 20 min). The plate was fixed with 0.1% glutaraldehyde for 30 min in room temperatures. Then solution was removed by pipetting. The plate was then treated with a solution of 0.1% BSA in PBS supplemented with 0.1 M glycine for 2 h at RT. The plates were blocked with 4% (w/v) BSA (BSA Blocker, Thermo Scientific) for 2 h at RT. After washing with PBS-T, the plate was incubated with protein H/MTFP-Gp17 (8.5 µg/ml diluted in PBS) for 2 h at RT. After washing, HRP-conjugated IgG anti-HisTag monoclonal antibody (Biorad), diluted 1:200 in blocking buffer was added to the wells, and the plate was incubated with antibody for 60 min at 37 °C. After washing, the plate was incubated with TMB substrate (Thermo

Scientific) for 5–15 min at RT. Reaction was stopped with 0.18 M sulphuric acid. Color development was measured at 450 nm on a microplate reader (Biotek).

Immunogold labelling—visualization of the interaction of the phage tail fiber protein with Ye:O3 surface using transmission electron microscopy (TEM)

YeO:3 was grown for 24 h at 28°C, after this time the bacteria were multiplied by transferring to fresh LB medium and incubated until the culture reached OD600 of 0.4–0.5. The culture was centrifuged (3000×g, 10 min), the pellet was washed with PBS and centrifuged again. The preparation was applied to a 200-mesh nickel grid with carbon film for 15 min. Next the excess liquid was sucked away, the H/MTFP-Gp17 protein at concentration of 1 mg/ml was applied to the grid for 30 min. Then the monoclonal anti-Maltose Binding Protein (MBP) antibody (produced in mouse, SIGMA) at a concentration of 1 mg/ml was applied to the grid for 30 min and in the next stage the anti-mouse IgG1 conjugated with HRP (produced in rabbit, SIGMA) at a concentration of 1 mg/ml was applied to the grid for 30 min. The grid was rinsed with PBS buffer supplemented with BSA and the Protein A-Gold with 20 nm colloidal gold was added for 30 min. After this time, the grid was washed with PBS and Milli Q water and contrasted with 2% uranyl acetate. Samples were visualized using a JEOL JEM-1200 EX 80kV TEM.

Table 1 Strains used to perform ELISA assay

No	Bacterial strain	Serotype	Source
1	<i>Yersinia enterocolitica</i> 6471/76-c	O:3	Skurnik, 1984
2	<i>Yersinia enterocolitica</i> YeO3-R1 (spontaneous rough derivative of 6471/76-c)		Al-Hendy et al., 1992
3	<i>Yersinia enterocolitica</i> DSMZ 23,249	O:8	DSMZ-German Collection of Microorganisms and Cell Cultures GmbH
4	<i>Yersinia enterocolitica</i> PCM 1879	4/O:3	Polish Collection of Microorganisms PAS (PCM)
5	<i>Yersinia enterocolitica</i> PCM 1880	5/O:3	PCM
6	<i>Yersinia enterocolitica</i> PCM 1881	O:3	PCM
7	<i>Yersinia enterocolitica</i> PCM 2072	O:1	PCM
8	<i>Yersinia enterocolitica</i> PCM 2080	O:8	PCM
9	<i>Yersinia enterocolitica</i> PCM 2081	O:9	PCM
10	<i>Yersinia enterocolitica</i> PCM 1883	1A/O:5A	PCM
11	<i>Yersinia enterocolitica</i> PCM 1884	2/O:8	PCM
12	<i>Yersinia enterocolitica</i> PCM 1886	O:7,8	PCM
13	<i>Yersinia enterocolitica</i> PCM 2090	O:27	PCM
14	<i>Escherichia coli</i> PCM 2337	–	PCM
15	<i>Escherichia coli</i> B PCM 1630	–	PCM
16	<i>Pseudomonas aeruginosa</i> PCM 499	–	PCM
17	<i>Enterobacter aerogenes</i> PCM 532	–	PCM

Results

Gene cloning, protein overexpression, purification and analysis

The gene *g17* was cloned to the expression vector pMCSG9 in order to obtain both a tagged and untagged variant of the adhesin TFP-Gp17. The tagged variant carried the MBP-His tag to be used as a target for the anti-His and anti-MBP antibodies.

The overexpressed TFP-Gp17 carrying the MBP-His tag (H/MTFP-Gp17) was purified using two rounds of nickel-affinity chromatography followed by SEC as described in the Materials and Methods section. The recovered protein was analyzed in SDS-PAGE (Fig. 1).

Based on these analyses, the purity of both TFP-Gp17 and H/MTFP-Gp17 was over 95%. After digestion of TFP-Gp17 with TEV protease two polypeptides were identified in mass analysis with molecular masses after

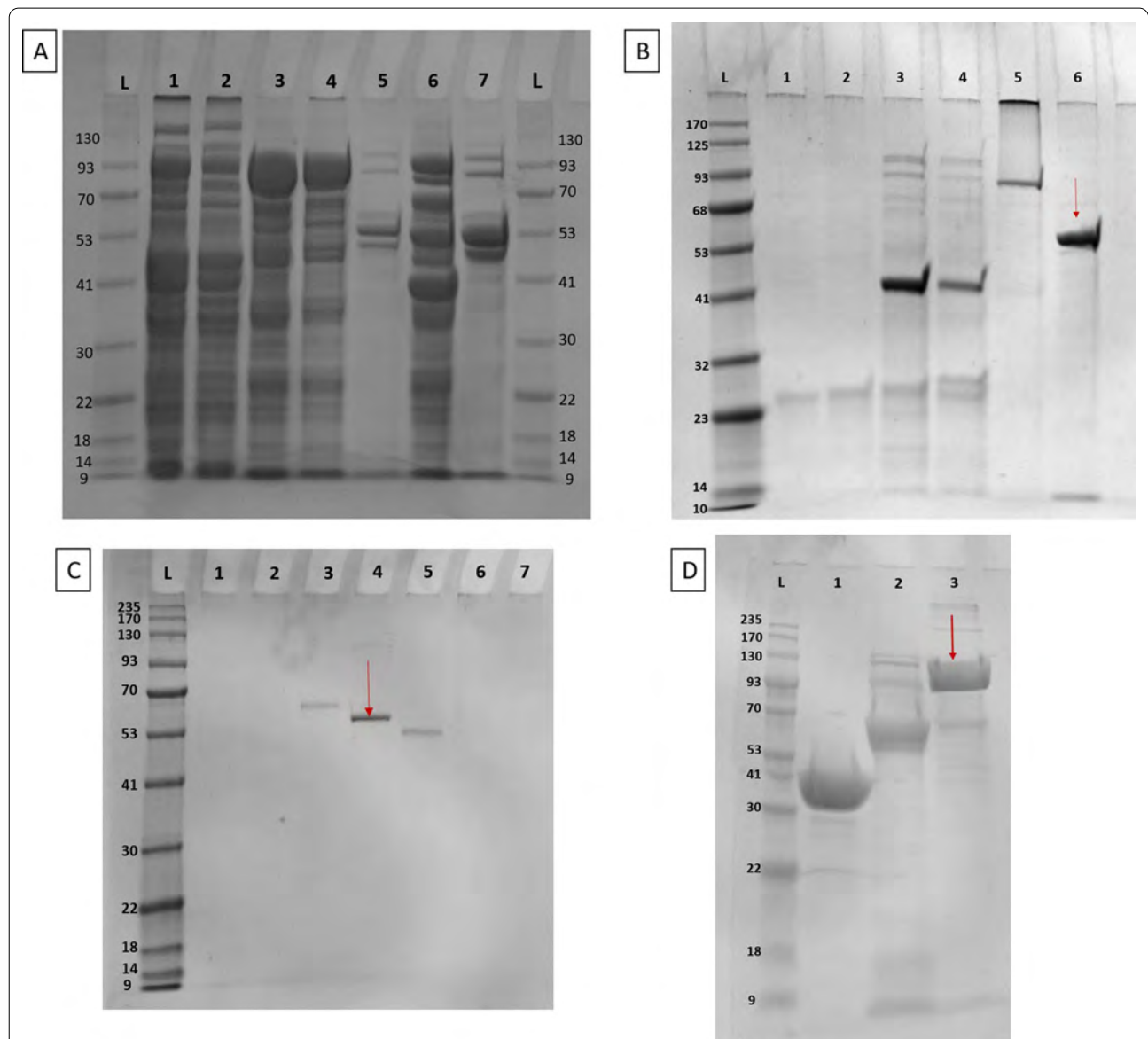


Fig. 1 SDS-PAGE of TFP-Gp17. **A** Samples from all purification steps by nickel-affinity chromatography. Lanes: 1, supernatant from bacterial lysate; 2, flow through 1 (FT1) containing the unbound proteins; 3, elution fraction 1 (elution with buffer A supplemented with 250 mM imidazole); 4, elution fraction after desalting with ammonium sulfate; 5,7, FT2, containing the unbound proteins after cleavage with the TEV protease; 6, elution fraction 2. **B** TFP-Gp17 before SEC chromatography is shown in the lane 6. **C** Protein fractions after SEC chromatography; TFP-Gp17 in 4 lane is indicated by the red arrow. **D** The H/MTFP-Gp17 protein after SEC chromatography is indicated in lane by the red arrow

deconvolution of 57.423 and 12.242 kDa. Their combined mass of 69.665 kDa corresponds well with the predicted mass of 69.4 kDa of TFP-Gp17 before autoprolysis. The mass spectrometry indicated that the molecular mass of H/MTFP-Gp17 was 100.846 kDa, and that also another protein with a molecular mass of 12.242 kDa was present. Their combined mass of 113.088 kDa corresponds to the predicted mass of H/MTFP-Gp17 before autoprolysis (Additional file 1: Figures S1 and S2). The autoprolysis indicates that the C-terminal fragment of TFP-Gp17 is cleaved, supporting our previous observation that TFP-Gp17 contains an S74 peptidase domain, which is responsible for protein autolysis (Pyra et al., 2020). Importantly, the results also demonstrated that there are no sterical obstacles for the autoprolysis in the H/MTFP-Gp17 construct where the His-MBP is fused to the N-terminal end of TFP-Gp17.

Bacterial cell-based sandwich ELISA

We used an ELISA-based method for the specific detection of the pathogenic YeO:3 bacteria employing H/MTFP-Gp17 as the recognition agent. Microtiter plate wells were coated with whole bacteria and the bound H/MTFP-Gp17 was detected by the HRP-conjugated anti-His antibodies. The specificity of the assay was tested against pathogenic serotype O:3, O:8, O:9, and O:5 strains responsible for Yersiniosis in Europe, and against several other non-Ye pathogens. We also included the rough YeO:3 mutant strain YeO3-R1 devoid missing the O-antigen. As negative controls we used samples that contained the YeO:3 bacteria without H/MTFP-Gp17 or with His-MBP (Fig. 2), while H/MTFP-Gp17 alone served as positive control for the protein interaction with anti-HisTag antibody. The experiment was repeated three times and each time the signal from O:3 strains was the highest.

Among the four YeO:3 strains, the biotype 5 strain had the highest signal with the absorbance at 450 nm of over 3.5. The absorbances of the other O:3 strains were

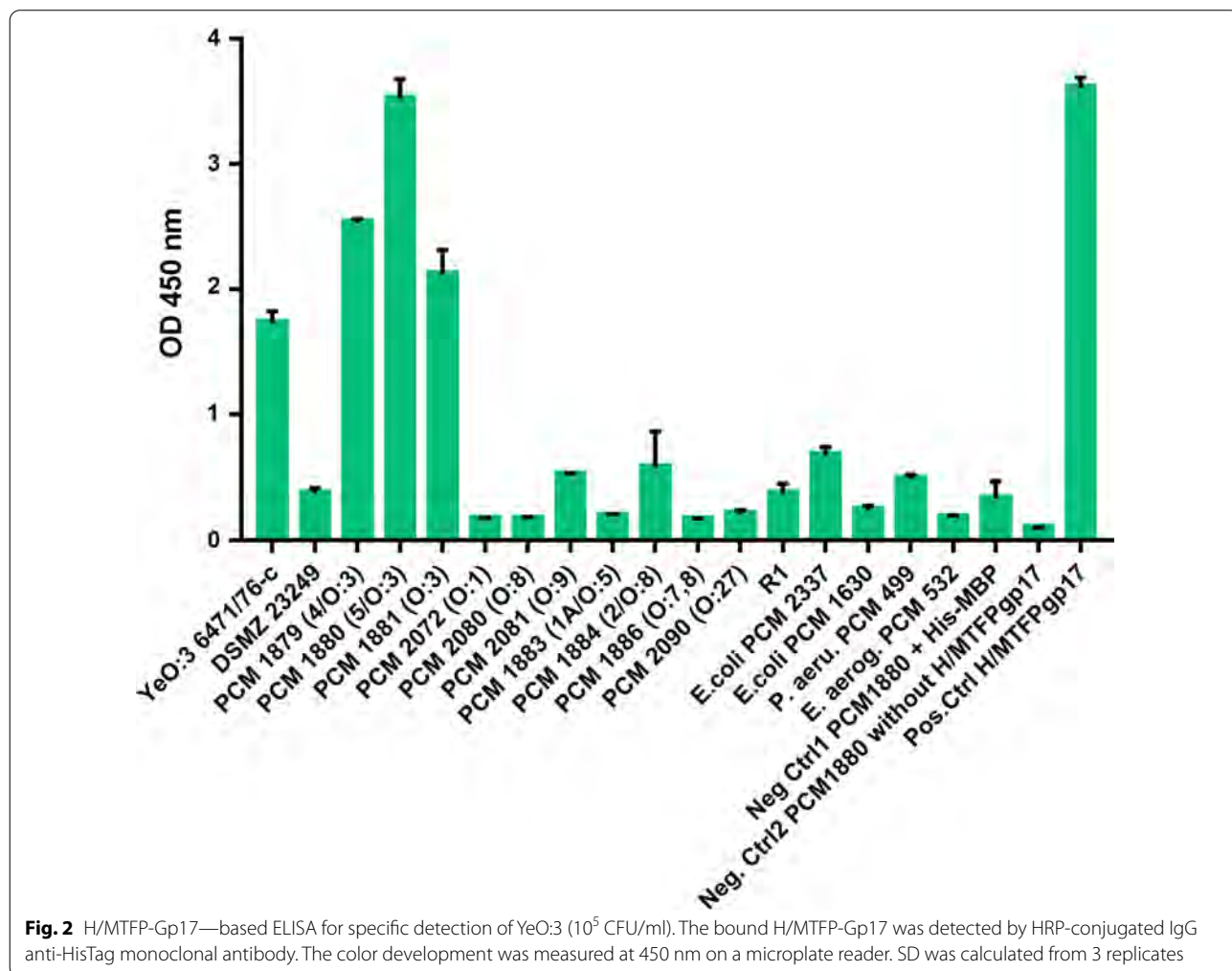


Fig. 2 H/MTFP-Gp17—based ELISA for specific detection of YeO:3 (10⁵ CFU/ml). The bound H/MTFP-Gp17 was detected by HRP-conjugated IgG anti-HisTag monoclonal antibody. The color development was measured at 450 nm on a microplate reader. SD was calculated from 3 replicates

between 1.5 and 2.5. The O-antigen negative mutant YeO3-R1 was not recognized by the phage adhesion (The absorbance was below 0.5). Among the non-Ye:O3 strain, *E coli* PCM2337 gave the highest signal, ~ 0.7, at the same test conditions. In the test conditions, a signal higher than 1.0 at 450 nm may be a safe cut-off above which the detected strain certainly belongs to the YeO:3 serotype. These results indicated that TFP-Gp17 is responsible for the recognition of the bacteria and that the O-antigen of YeO:3 bacteria is the receptor. The results also clearly show that MBP does not bind non-specifically to bacteria.

The sensitivity of the ELISA was assessed by coating the wells with final bacterial concentrations of 10⁶, 10⁵ and 10⁴ CFU/ml of the four YeO:3 strains (Fig. 3).

Taking into account the safe cut-off value (estimated as 1.0), the limit of bacteria detection in this ELISA conditions may be determined as not less than 10⁵ CFU.

Visualization of the phage adhesin interaction with YeO:3 using TEM

For the visualization of the interaction between phage protein and bacterial cell, we designed a sandwich-type preparation method based on immunogold labelling according to the scheme represented below (Fig. 4).

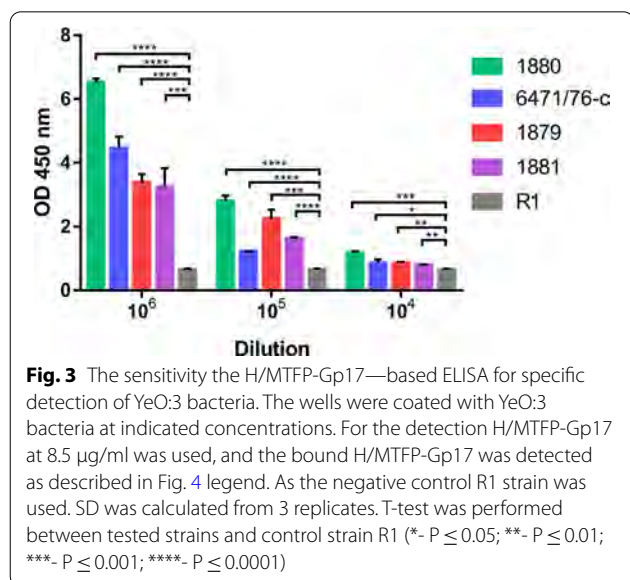
As a negative control, we used YeO3-R1, the rough derivative of the YeO:3 wild type strain 6471/76-c. The immunogold labelling for both types of bacteria was performed identically. In Fig. 5B gold nanoparticles can be observed on the YeO:3 surface, which indicates a reaction with the antibody conjugated with protein on the bacteria surface. We didn't observe such effect in the case of R1 mutant (Fig. 5A). This is another evidence showing the

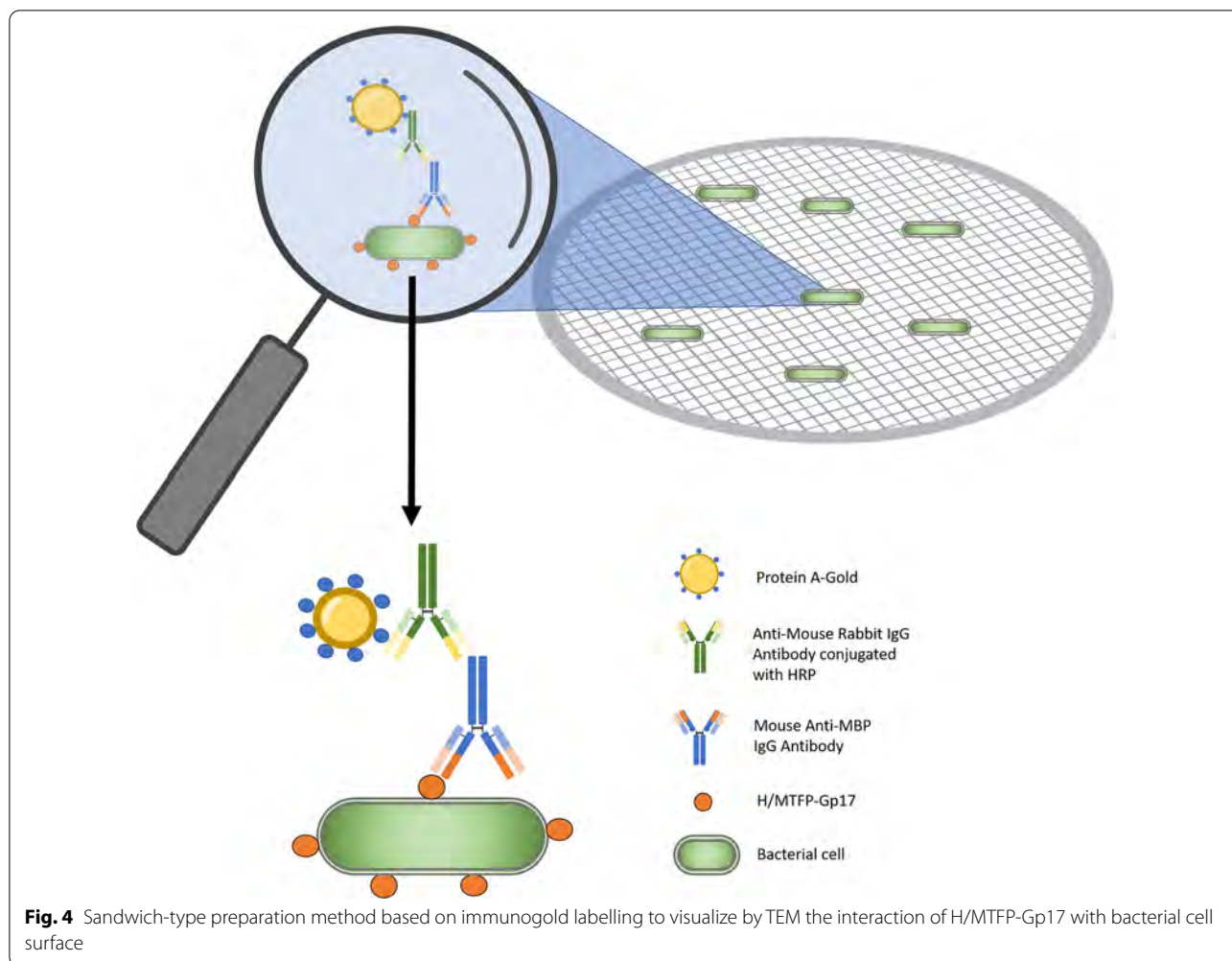
specific and selective interaction of the protein with bacteria YeO:3 surface (Additional file 1: Figure S3).

Discussion

Yersiniosis is somewhat underestimated threat to human health. In some countries, *Yersinia* infections have overtaken *Shigella* and *Salmonella* species as the most common cause of bacterial gastroenteritis. Diagnosis towards Ye infections is infrequently performed routinely in clinical laboratories because of Ye specific growth characteristics, which make it difficult to isolate and culture (Aziz and Yalamanchili, 2021). Moreover, current isolation procedures are time-consuming and expensive, thus leading to underestimates of the incidence of enteric yersiniosis, inappropriate prescriptions of antibiotic treatments, and unnecessary appendectomies (Laporte et al., 2015; Weagant and Feng, 2017). Since the Ye serotype O:3 is the most common pathogenic serotype encountered in Europe, we have demonstrated for the first time the ability of the phage ϕ YeO3-12 receptor-binding-protein to interact with the serotype O:3 bacteria in the highly sensitive ELISA. The TFP-Gp17 is most specific towards both pathogenic biotype 5 and 4 of Ye with O:3 serotype (Leon-Velarde et al., 2014). The differences between these two biotypes are determined by presence of some genes responsible for biological activity (Morka et al., 2018). The ability of some tail fiber phage adhesins for bacteria detection was demonstrated previously by Leon-Velarde et al. (2019). The example was RBP Gp17 derived from the Podovirus phage vB_YenP_AP5. The protein was identified as a ligand with specificity also for the O-antigen of serotype O:3 strains. These two adhesins, TFP-Gp17 of ϕ YeO3-12 and RBPgp17 of phage vB_YenP_AP5 are 89% identical with highest similarity at the N-terminal parts of these proteins (Leon-Velarde et al., 2014). Another example was a distal long tail fiber protein, RBP Gp37, derived from the Myovirus phage vB_YenM_TG1, however, it was identified as a ligand for the outer membrane protein OmpF of serotype O:3, O:5,27 and O:9 (Leon-Velarde, 2017). In the case of TFP-Gp17, the specificity is strictly restricted to the serotype O:3.

Our results correspond to those reported by Laporte et al. (2015) where EIA (Enzymatic Immunoassay) was presented for fast Ye detection. Although, Laporte used the monoclonal antibodies in that assay the detection limits are the same also for TFP-Gp17. The advantage of using the phage protein is due to the fact, the procedure of phage adhesin production and purification is faster and easier than monoclonal antibody generation. TFP-Gp17 is produced in complex with MBP. The His-MBP is not cleaved out from the complex since its presence is necessary for the phage adhesin detection after bacteria binding.





To visualize and confirm the interaction of the phage adhesin with whole bacteria we used immunogold labelling method based on the workflow from ELISA with slight modifications. As a negative control, YeO3-R1 mutant was used (Kaur et al., 2002). This experiment confirmed the previous results, gold nanoparticles were abundantly present on the YeO:3 surface in contrast to the YeO3-R1 mutant, which indicates the high specificity of the complex.

In our previous report, we indicated that two amino sugars GalNAc and GlcNAc stabilize the phage adhesin having impact on increase of its thermal stability (Pyra et al., 2020). Initially we thought that this could be due to interaction with the aminosugar moieties presented in the outer core (OC) of YeO:3 (Al-Hendy et al., 1992). For that reason, we took the YeO3-R1 mutant with exposed OC missing the O-antigen to check the interaction. Neither H/MTFP-Gp17 nor TFP-Gp17 bound to YeO3-R1 bacteria in opposition to the wild type YeO:3 bacteria. We decided to assess whether the interaction occurs

with those two aminosugars using modified AuNP and UV-Vis spectroscopy (Additional file 1: Figure S4). As the LSPE band is sensitive to the nanoparticles properties, UV-Vis spectroscopy was used to monitor the LSPR band of AuNPgalSH30 and AuNPglukSH30 before and after incubation with the phage adhesin. According to the obtained results, the interaction between the phage adhesin and these two aminosugars took place. We assume that the interaction stabilizes not only the TFP-Gp17 but also the whole phage in the gastrointestinal tract of both humans and pigs since GalNAc as well as GlcNAc are released by enzymes coming from the commensal bacteria (Sicard et al., 2017). As it was previously mentioned, these aminosugars are also a carbon source for Ye strains (Schmühl et al., 2019). Moreover, Ye contain mucin-degrading enzyme(s) increasing the permeability of the mucus gel layer and allowing the bacteria to move more easily through the mucin (Sicard et al., 2017).

For the first time, the TFP-Gp17 of phage ϕ YeO3-12 was demonstrated as a highly specific adhesin towards

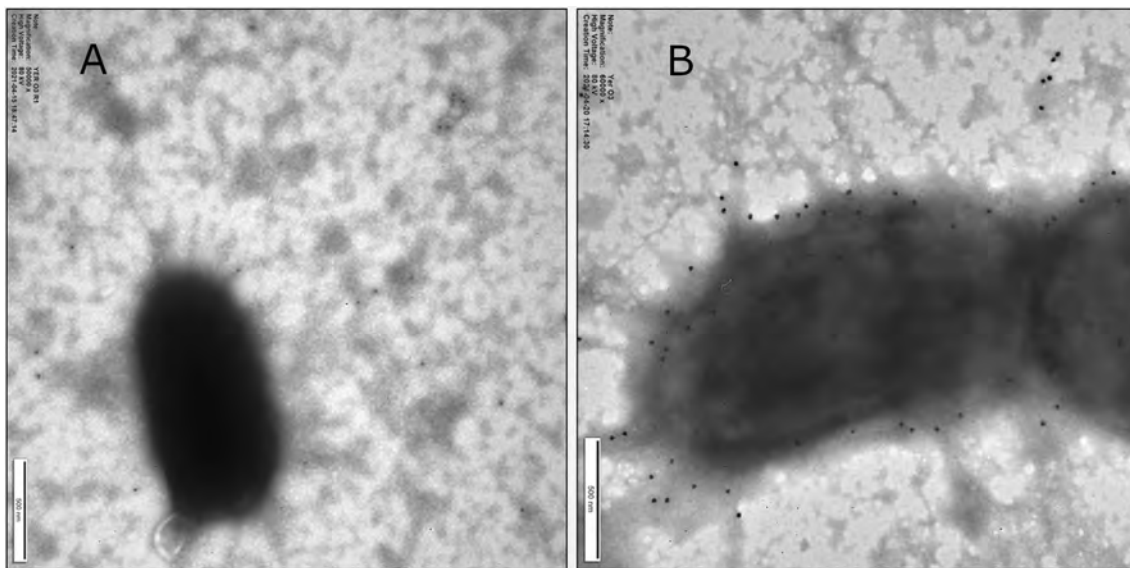


Fig. 5 Specific interaction of H/MTFP-Gp17 with YeO:3 bacteria visualized by TEM using immunogold labelling. **A** R1 mutant, the negative reaction with Protein A-Gold, **B** YeO:3 wild type strain 6471/76-c, positive reaction with Protein A-Gold

the pathogenic Ye serotype O:3. It was shown that TFP-Gp17 carrying the 6His-MBP tag could be used as a sensing molecule to detect the YeO:3 strains. The phage adhesin recognizes the serotype O:3 O-antigen of the bacterial LPS. The YeO3-R1 mutant strain lacking the O-antigen was not recognized by the phage protein. Both the specificity and sensitivity of the ELISA test were determined. We also showed that the phage receptor binding adhesin protein interacts with the two major mucus components GalNAc and GlcNAc which suggests that the amino sugars stabilize the whole phage particle in the intestinal tracts of mammals.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13568-021-01341-2>.

Additional file 1: Additional file contains partial implementation, Figures S1–S4.

Acknowledgements

Not applicable

Authors' contributions

KF, IM—Experimental part—protein production and purification, ELISA, Writing, Figures. BSO—Experimental part—TEM photography. ER—Experimental part—UV-Vis measurement, nanoparticles modification. JNJ—Conceptualization, Writing-review. JC—Experimental part- SEC, MW analysis. AP—Review. MS—Writing—review. EB—Conceptualization, Funding Acquisition, Supervision, Writing. All authors read and approved the final manuscript.

Funding

This study was supported by National Science Center, Poland (Grant Numbers UMO-2017/26/E/NZ1/00249).

Availability of data and materials

All data generated or analysed during this study are included in this published article (and its additional files).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Hirsfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, 12 R. Weigl St, 53114 Wrocław, Poland. ²Institute of Physical Chemistry, Polish Academy of Sciences, Kasprzaka 44, 5201-224 Warsaw, Poland. ³Faculty of Chemistry, University of Wrocław, 14 F. Joliot-Curie St, 50383 Wrocław, Poland. ⁴Department of Mycology and Genetics, Institute of Genetics and Microbiology, University of Wrocław, 51-148 Wrocław, Poland. ⁵Department of Bacteriology and Immunology, Faculty of Medicine, Human Microbiome Research Program, University of Helsinki, Helsinki, Finland. ⁶Division of Clinical Microbiology, Helsinki University Hospital, HUSLAB, Helsinki, Finland.

Received: 20 October 2021 Accepted: 19 December 2021

Published online: 06 January 2022

References

- Al-Hendy A, Toivanen P, Skurnik M (1992) Lipopolysaccharide O side chain of *Yersinia enterocolitica* O: 3 is an essential virulence factor in an orally infected murine model. *Infect Immun* 60(3):870–875
- Aziz M, Yalamanchili VS (2021) *Yersinia enterocolitica*. StatPearls Publishing, Treasure Island

- Batzilla J, Antonenka U, Höper D, Heeseemann J, Rakin A (2011) *Yersinia enterocolitica* palearctica serobiotyp O: 3/4-a successful group of emerging zoonotic pathogens. *BMC Genomics* 12(1):1–11
- Eschenfeldt WH, Lucy S, Millard CS, Joachimiak A, Mark ID (2009) A family of LIC vectors for high-throughput cloning and purification of proteins In high throughput protein expression and purification. *Methods Mol Biol* 498:105–115
- Fàbrega A, Vila J (2012) *Yersinia enterocolitica*: pathogenesis, virulence and antimicrobial resistance. *Enferm Infecc Microbiol Clin* 30(1):24–32
- Kaur R, Dikshit KL, Rajee M (2002) Optimization of immunogold labeling TEM: an ELISA-based method for evaluation of blocking agents for quantitative detection of antigen. *J Histochem Cytochem* 50(6):863–873
- Kenyon JJ, Duda KA, De Felice A, Cunneen MM, Molinaro A, Laitinen J, De Castro C (2016) Serotype O: 8 isolates in the *Yersinia pseudotuberculosis* complex have different O-antigen gene clusters and produce various forms of rough LPS. *Innate Immun* 22(3):205–217
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227(5259):680–685
- Laporte J, Savin C, Lamourette P, Devilliers K, Volland H, Carniel E, Simon S (2015) Fast and sensitive detection of enteropathogenic *Yersinia* by immunoassays. *J Clin Microbiol* 53(1):146–159
- Lebendiker M, Danieli T (2010) Purification of proteins fused to maltose-binding protein. *Methods Mol Biol* 681:281–293
- Leon-Velarde CG, Kropinski AM, Chen S, Abbasifar A, Griffiths MW, Odumeru JA (2014) Complete genome sequence of bacteriophage vB_YenP_AP5 which infects *Yersinia enterocolitica* of serotype O: 3. *Virology J* 11(1):1–14
- Leon-Velarde CG, Jun JW, Skurnik M (2019) *Yersinia* phages and food safety. *Viruses* 11(12):1105
- Leon-Velarde CG (2017) The application of bacteriophage host recognition binding proteins for the isolation of *Yersinia enterocolitica* in Foods, Doctoral dissertation. <https://atrium.lib.uoguelph.ca/xmlui/handle/10214/10451>
- Leskinen K, Blasdel BG, Lavigne R, Skurnik M (2016) RNA-sequencing reveals the progression of phage-host interactions between ϕ R1-37 and *Yersinia enterocolitica*. *Viruses* 8(4):111
- Morka K, Bystron J, Bania J, Korzeniowska-Kowal A, Korzekwa K, Guz-Regner K, Bugla-Ploskońska G (2018) Identification of *Yersinia enterocolitica* isolates from humans, pigs and wild boars by MALDI TOF MS. *BMC Microbiol* 18(1):1–10
- Pajunen M, Kiljunen S, Skurnik M (2000) Bacteriophage ϕ YeO3-12, specific for *Yersinia enterocolitica* serotype O: 3, is related to coliphages T3 and T7. *J Bacteriol* 182(18):5114–5120
- Paton JC, Rogers TJ, Morona R, Paton AW (2001) Oral administration of formaldehyde-killed recombinant bacteria expressing a mimic of the Shiga toxin receptor protects mice from fatal challenge with Shiga-toxicogenic *Escherichia coli*. *Infect Immun* 69(3):1389–1393
- Pinta E, Duda KA, Hanuszkiewicz A, Salminen TA, Bengoechea JA, Hyytiäinen H, Skurnik M (2010) Characterization of the six glycosyltransferases involved in the biosynthesis of *Yersinia enterocolitica* serotype O: 3 lipopolysaccharide outer core. *J Biol Chem* 285(36):28333–28342
- Pyra A, Filik K, Szermer-Olearnik B, Czarny A, Brzozowska E (2020) New insights on the feature and function of tail tubular protein b and tail fiber protein of the lytic bacteriophage ϕ YeO3-12 specific for *Yersinia enterocolitica* serotype O: 3. *Molecules* 25(19):4392
- Schmühl C, Beckstette M, Heroven AK, Bunk B, Spröer K, McNally A, Overmann J, Derscha P (2019) Comparative transcriptomic profiling of *Yersinia enterocolitica* O:3 and O:8 reveals major expression differences of fitness and virulence-relevant genes indicating ecological separation. *M Syst* 4(2):e00239-18
- Shoib M, Shehzad A, Raza H, Niazi S, Khan IM, Akhtar W, Wang Z (2019) A comprehensive review on the prevalence, pathogenesis and detection of *Yersinia enterocolitica*. *RSC Adv* 9(70):41010–41021
- Sicard JF, Le Bihan G, Voegeléer P, Jacques M, Harel J (2017) Interactions of intestinal bacteria with components of the intestinal mucus. *Front Cell Infect Microbiol* 7:387
- Simonova J, Vazlerova M, Steinhauserova I (2007) Detection of pathogenic *Yersinia enterocolitica* serotype O: 3 by biochemical, serological, and PCR methods. *Czech J Food Sci* 25(4):214
- Skurnik M (1984) Lack of correlation between the presence of plasmids and fimbriae in *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. *J. Appl. Bacteriol.* 56:355–363
- Skurnik M, Venho R, Bengoechea JA, Moriyón I (1999) The lipopolysaccharide outer core of *Yersinia enterocolitica* serotype O: 3 is required for virulence and plays a role in outer membrane integrity. *Mol Microb* 31(5):1443–1462
- Smith PE, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano M, Klenk DC (1985) Measurement of protein using bicinchoninic acid. *Anal Biochem* 150(1):76–85
- Triantafyllidis JK, Thomaidis T, Papalois A (2020) Terminal Ileitis due to *Yersinia* infection: an underdiagnosed situation. *Biomed Res Int.* <https://doi.org/10.1155/2020/1240626>
- Watkins LKF, Fredman CR (2020) *Yersiniosis*, Yellow Book, Traveler's Health, CDC, Oxford University Press, Ch. 4
- Weagant DS, Feng P (2017) Nutrition. *BAM Chapter 8.Science & Food.*
- Wielkoszynski T, Moghaddam A, Bäckman A, Broden J, Piotrowski R, Mond-Paszek R, Wilczynska M (2018) Novel diagnostic ELISA test for discrimination between infections with *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. *Eur J Clin Microbiol Infect Dis* 37(12):2301–2306

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

Additional file 1.

AMB Express

ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3

Karolina Filik¹, Bożena Szermer-Olearnik¹, Joanna Niedziółka-Jönson², Ewa Rożniecka², Jarosław Ciekot¹, Anna Pyra³, Irwin Matyjaszczyk⁴, Mikael Skurnik^{5,6} and Ewa Brzozowska^{1*}

1. Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, 12 R. Weigl St, 53114, Wrocław, Poland

2. Institute of Physical Chemistry, Polish Academy of Sciences, Kasprzaka 44/52 01-224 Warsaw, Poland

3. University of Wrocław, Faculty of Chemistry, 14 F. Joliot-Curie St, Wrocław, 50383, Poland

4. Department of Mycology and Genetics, Institute of Genetics and Microbiology, University of Wrocław, 51-148, Poland

5. Department of Bacteriology and Immunology, Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

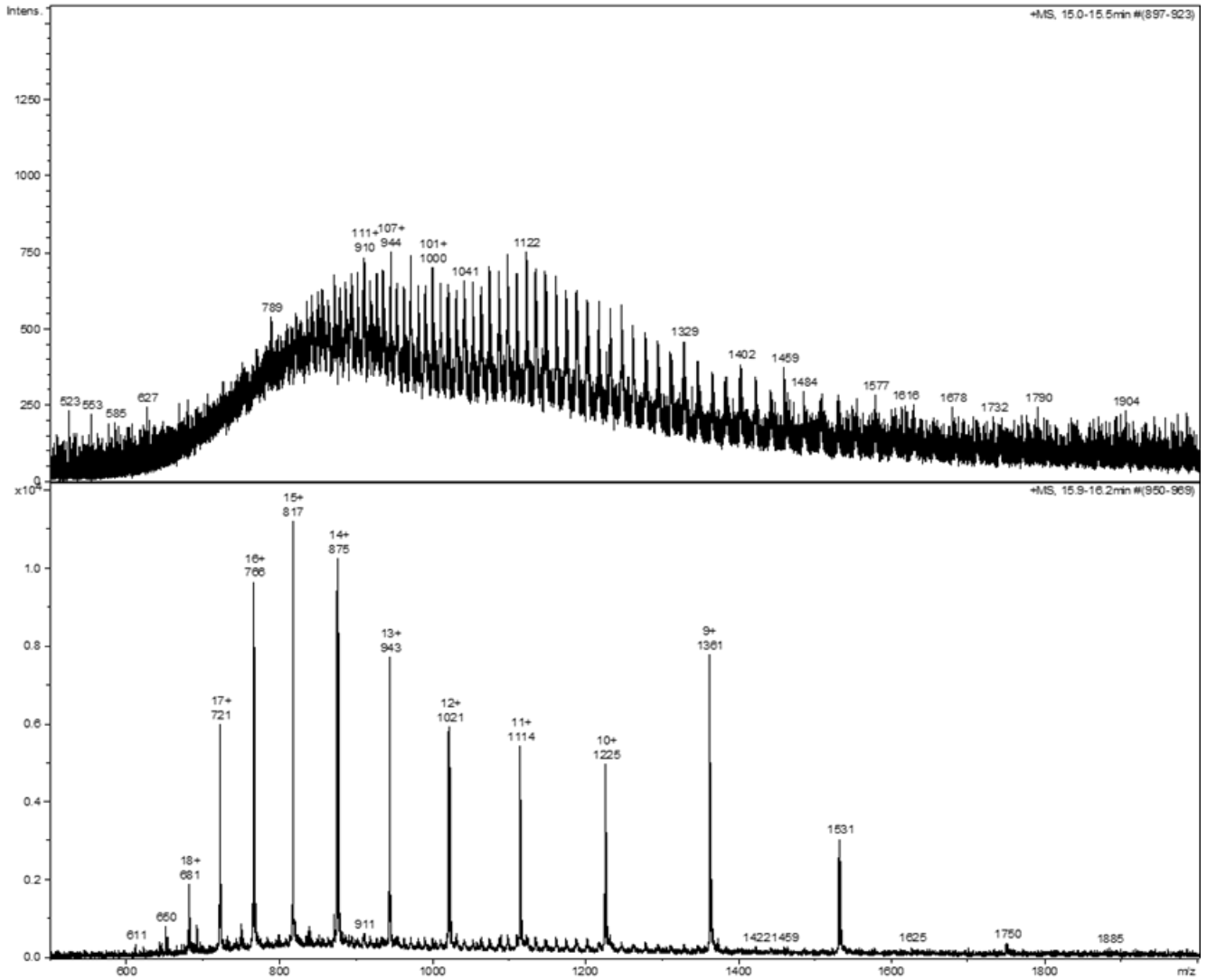
6. Division of Clinical Microbiology, Helsinki University Hospital, HUSLAB, Helsinki, Finland

*Correspondence: ewa.brzozowska@hirszfeld.pl;

Tel.: +48 71 370 99 01, Fax: +48 71 337 13 82

Results:

Mass spectra



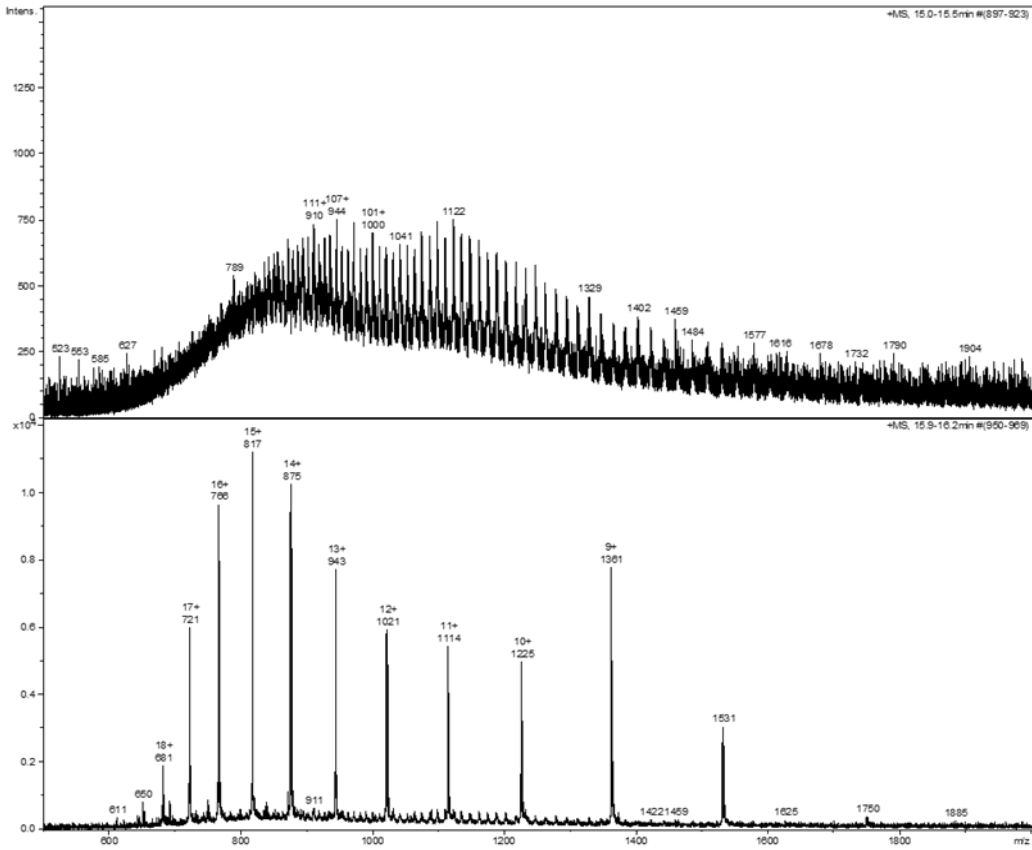
Deconvolution 1:

Component	Mass	Molecule	Abundance	Abund.[%]	StdDeviation
A	100846	[M + H] ⁺	19079	100	0.781884
B	100900	[M + H] ⁺	11159	58	0.710151
C	102119	[M + H] ⁺	8590	45	0.644540

Deconvolution 2:

Component	Mass	Molecule	Abundance	Abund.[%]	StdDeviation
A	12242	[M + H] ⁺	75070	100	0.234035

Figure 1S. Mass spectra of H/MTFP-Gp17 complex



Deconvolution 1:

Component	Mass	Molecule	Abundance	Abund.[%]	StdDeviation
A	100846	[M + H] ⁺	19079	100	0.781884
B	100900	[M + H] ⁺	11159	58	0.710151
C	102119	[M + H] ⁺	8590	45	0.644540

Deconvolution 2:

Component	Mass	Molecule	Abundance	Abund.[%]	StdDeviation
A	12242	[M + H] ⁺	75070	100	0.234035

Figure 2S. Mass spectra of TFP-Gp17 after digestion with TEV protease

Visualization of the phage adhesin interaction with YeO:3 using TEM

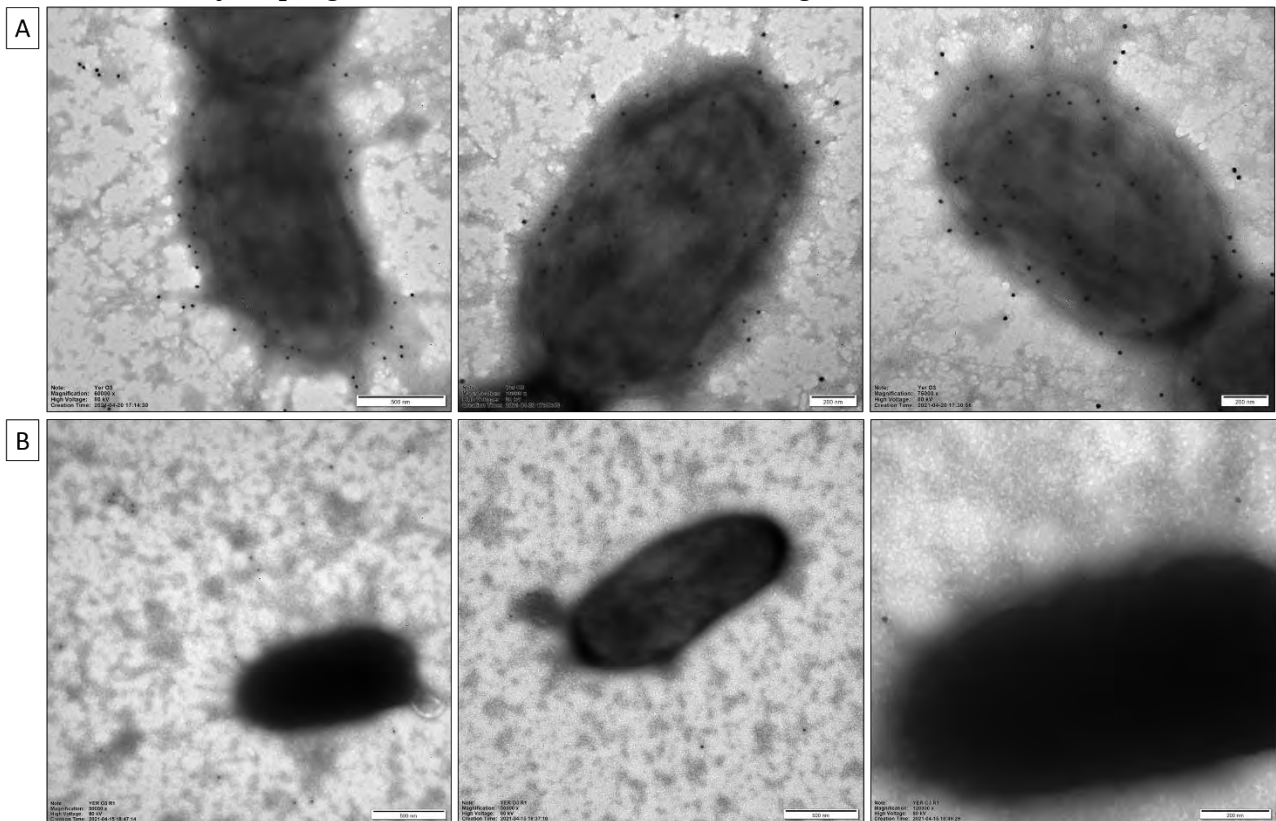


Figure 3S. Specific interaction of H/MTFP-Gp17 with YeO:3 bacteria visualized by TEM using immunogold labelling. (A) YeO:3 wild type strain 6471/76-c, positive reaction with Protein A-Gold, (B) R1 mutant, the negative reaction with Protein A-Gold.

Material and Methods for Assessment of binding between TFP-Gp17 and gold nanoparticles coated by thiol derivatives 218 of GalNAc and GlcNAc 219

Synthesis of gold nanoparticles (AuNP) stabilized with citrates was performed as 220 described before [23] 150 ml of a 0.5 mM HAuCl₄ 3×H₂O solution was heated up to 95° C 221 on a heating plate with intense stirring (900 rpm). Then, 15 ml of 38.8 mM trisodium citrate 222 dihydrate was added. The solution was further heated for 18 min and cooled. Next, 5 ml 223 of AuNP stock suspension and 5 ml of ethanol were added to 2 vials. Then, 30 μl of 2.12 224 μM ethanolic solution of 11-(α-D-galactopyranosyl) undecane-1-thiol (gal-SH) 225 (ProChimina) and 11-(α-D-glucopyranosyl) undecane-1-thiol (gluc-SH) (ProChimina) 226 were added, respectively. The modified nanoparticles were marked as AuNPgalSH30 and 227 AuNPglucSH30. UV-Vis spectroscopy was used to characterize synthesized AuNP and 228 their surface modification, as well as tracking the interaction between the immobilized 229 aminosugars and the adhesin TFP-Gp17, by following the changes of the localized surface 230 plasmon resonance (LSPR) band with a maximum wavelength (λ_{max}) at around 530 nm. 231 All measurements were performed in a BRAND semi-micro cuvette made from PMMA 232 with an optical path of 10 mm. First, the suspensions of AuNP as synthesized and 233 modified with aminosugars were measured with addition of PBS and BSA as reference 234

probes. Finally, TFP-Gp17 in PBS was added to a final concentration of 10 $\mu\text{g/ml}$ and after 235 2 h, the spectra were measured once again.

Results for Assessment of the binding between TFP-Gp17 and gold nanoparticles modified with thiol derivatives of GalNAc and GlcNAc 340

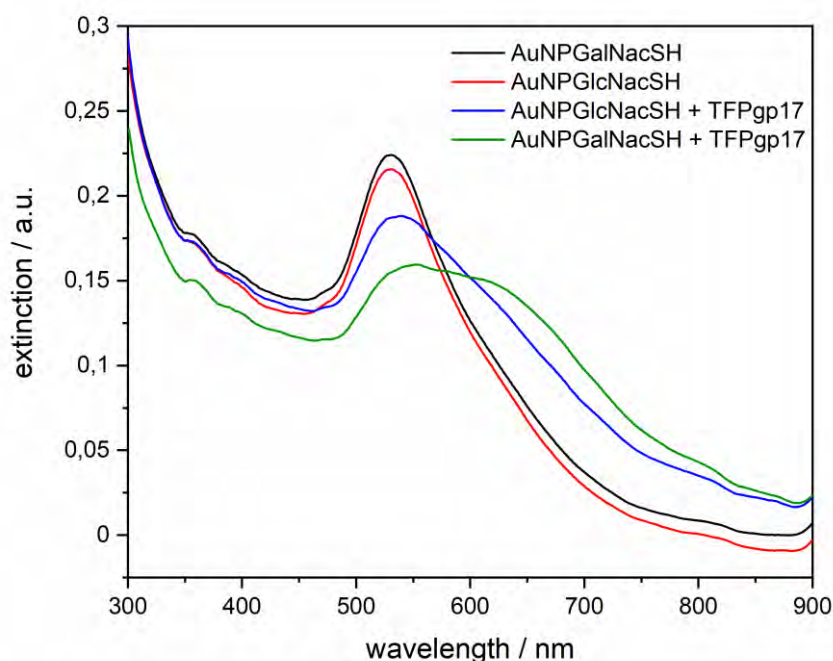


Figure 4S Extinction spectra for AuNPGalNacSH (black) and AuNPGlcNacSH (red) suspensions before and after (green and blue) 2h incubations with 10 $\mu\text{g/ml}$ TFP-Gp17.

In our previous report, we indicated that TFP-Gp17 is stabilized by the aminosugars GalNAc and GlcNAc [1] suggesting an interaction of the adhesin with these sugars. To assess the interaction between the aminosugars and the TFP-Gp17 adhesin we used colorimetric detection [2]. This is possible because the wavelength of the localized surface plasmon resonance of the AuNPs is strongly dependent on the dielectric constant of the local environment. Therefore, we synthesized gold nanoparticles with covalently immobilized GalNAc and GlcNAc marked as AuNPGalNacSH and AuNPGlcNacSH, respectively. In the first control experiments PBS and BSA were added to the AuNPGalNacSH and AuNPGlcNacSH suspensions to see if a change of the ionic strength, in the case of PBS, or sugar protein interaction, in the case of BSA, affected the position of the LSPR. No changes in the spectra were observed (data not shown). Then the spectra of AuNPGalNacSH and AuNPGlcNacSH nanoparticles were measured before (black and red color in Figure 1) and after incubation with the TFP-Gp17 (blue and green color) under the same experimental conditions. The results are presented in Figure 4.

The LSPR bands at 531 nm of AuNPGlcNacSH and AuNPGalNacSH after incubation with the adhesin were shifted to 537 nm and 553 nm respectively, and the extinction decreased significantly. This directly indicates an interaction between the phage protein and the GalNac and GlcNac coated particles. The shift for GalNac was more pronounced than that for GlcNac. This shows that refractive index in the closest vicinity of AuNPGalNacSH changed significantly. Moreover, a second band around 610 nm is observed due to changes of the interparticle distance and aggregates of AuNPGalNacSH formation caused by the complex formation between the aminosugar and adhesin [2,3]. Only a small shoulder is seen in this position in the case of AuNPGlcNacSH.

Bibliography:

[1] Matyjewicz, J., Lesniewski, A., & Niedziolka-Jonsson, J. (2014). Click chemistry modification of glassy carbon electrode with gold nanoparticles for electroactive ion discrimination. *Electrochemistry communications*, 48, 73-76.

doi: 10.1016/j.elecom.2014.08.020

[2] Kannan, P., Los, M., Los, J. M., & Niedziolka-Jonsson, J. (2014). T7 bacteriophage induced changes of gold nanoparticle morphology: biopolymer capped gold nanoparticles as versatile probes for sensitive plasmonic biosensors. *Analyst*, 139(14), 3563-3571.

doi: 10.1039/C3AN02272B

[3] Rechberger, W., Hohenau, A., Leitner, A., Krenn, J. R., Lamprecht, B., & Aussenegg, F. R. (2003). Optical properties of two interacting gold nanoparticles. *Optics communications*, 220(1-3), 137-141. doi: 10.1016/S0030-4018(03)01357-9

7. Achievements and scientific contribution resulting from the research carried out as part of this dissertation:

- I. The bacteriophage ϕ 80-18 was characterized. Its belonging to the *Podoviridae* family, the subfamily *Autographivirinae* was determined. It has been shown to be highly stable both in a wide range of pH 1-12 and temperatures (it is stable after incubation at 50°C). Phage grows slowly, has a long latent period (50 minutes) and a low "burst size" (8-10 PFU per infected bacterial cell). It seems to be an appropriate tool for the biocontrol of contamination with pathogenic strain 1B *Y. enterocolitica* O:8. Due to the precise determination of the biological characteristics of the phage, it can be indicated that it has a protective potential, e.g. of food contaminated with pathogenic strain *Yersinia enterocolitica* O:8 (Filik et al., 2020).
- II. It has been shown that the TTPAgp11 protein, apart from its structural function, also has an enzymatic function - it is an α -1,4-glucosidase. Thus, another phage protein belonging to the TTPA family with a dual function was discovered. The amino acid triad, the E-X-D motif present in the protein, which may be responsible for its enzymatic activity, has also been indicated (Pyra et al., 2020).
- III. The potential of TTPBgp12 protein to inhibit bacterial growth was demonstrated and formation of a bacterial biofilm of *Yersinia enterocolitica* O: 3 strains. This property can be used to develop antiseptic substances for general use (Pyra et al., 2020).
- IV. 4. It has been confirmed that the TFPgp17 protein is a protein with adhesive properties - it enables the diagnosis of the pathogenic *Y. enterocolitica* O: 3 serotype with high specificity. An efficient, fast and cheap method of producing this protein in fusion with the MBP protein and the His tag has been developed, which gives great opportunities to use this protein to develop new types of bio-sensors while maintaining the specificity and binding capacity of bacteria (Filik et al., 2022).

8. Autoreferat

Mgr Karolina Filik

Laboratorium Mikrobiologii Lekarskiej; Zakład Immunologii Chorób Zakaźnych

Instytut Immunologii i Terapii Doświadczalnej im. L. Hirszfelda PAN

Edukacja:

10.2018 – obecnie - Studium doktoranckie, Instytut Immunologii i Terapii Doświadczalnej im. Ludwika Hirszfelda PAN

02.2018 – 07.2018 - Erasmus+, Université catholique de Louvain, Life Sciences Institute, Molecular Physiology (FYMO) unit, Research Project pod kierunkiem prof. Francois'a Chaumont: Study the interactions between ZmPIP2;5 aquaporin and proteins involved in trafficking

2016 – 2018 - Uniwersytet Wrocławski, Wydział Nauk Biologicznych, Studia magisterskie, Mikrobiologia

2013 - 2016 – Uniwersytet Wrocławski, Wydział Nauk Biologicznych, Studia licencjackie, Biologia specjalność mikrobiologia

Publikacje:

1. Pyra, A., Urbańska, N., **Filik, K.**, Tyrlik, K., & Brzozowska, E. (2020). Biochemical features of the novel Tail Tubular Protein A of Yersinia phage phiYeO3-12. *Scientific reports*, 10(1), 1-11.
IF₂₀₂₀=4,38; MEiN=140
2. Fox, A. R., Scochera, F., Laloux, T., **Filik, K.**, Degand, H., Morsomme, P., ... & Chaumont, F. (2020). Plasma membrane aquaporins interact with the endoplasmic reticulum resident VAP27 proteins at ER–PM contact sites and endocytic structures. *New Phytologist*, 228(3), 973-988.
IF₂₀₂₀=10,151; MEiN=140
3. **Filik, K.***, Szermer-Olearnik, B.*, Wernecki, M., Happonen, L. J., Pajunen, M. I., Nawaz, A., Qasim M. S., Jun J. W., Mattinen, L., Skurnik, M.*, Brzozowska, E. (2020). The Podovirus φ80-18 Targets the Pathogenic American Biotype 1B Strains of Yersinia enterocolitica. *Frontiers in Microbiology*, 11, 1356. *autor korespondencyjny
IF₂₀₂₀=5,64; MEiN=100
4. Pyra, A., **Filik, K.**, Szermer-Olearnik, B., Czarny, A., & Brzozowska, E. (2020). New Insights on the Feature and Function of Tail Tubular Protein B and Tail Fiber Protein of the Lytic Bacteriophage φYeO3-12 Specific for Yersinia enterocolitica Serotype O: 3. *Molecules*, 25(19), 4392.
IF₂₀₂₀=4,412; MEiN=100

5. Cal, M., Matyjaszczyk, I., **Filik, K.**, Ogórek, R., Ko, Y., & Ułaszewski, S. (2021). Mitochondrial Function Are Disturbed in the Presence of the Anticancer Drug, 3-Bromopyruvate. *International Journal of Molecular Sciences*, 22(12), 6640. IF₂₀₂₀=5,924; MEiN=140
6. **Filik, K.**, Szermer-Olearnik, B., Niedziółka-Jönson, J., Roźniecka, E., Ciekot, J., Pyra, A., Matyjaszczyk, I., Skurnik, M., Brzozowska, E. ϕ YeO3-12 phage tail fiber Gp17 as a promising high specific tool for recognition of *Yersinia enterocolitica* pathogenic serotype O:3. *AMB Expr* 12, 1 (2022). IF₂₀₂₀=3,298; MEiN=70

Sumaryczny IF= 33,805

Sumaryczna wartość punktów MEiN=690

Uczestnictwo w projektach naukowych:

1. Wykonawca w projekcie NCN Sonata 9 „Badanie genotoksyczności przeciwnowotworowego związku 3-bromopirogronianu u drożdży *Saccharomyces cerevisiae*”; zadanie nr 5 – Analiza stabilności genomu w obecności 3-bromopirogronianu.
(kierownik projektu: dr Magdalena Cal, Zakład Mykologii i Genetyki, Uniwersytet Wrocławski; 07.2017 - 01.2018)
2. Realizacja projektu Sonata Bis 7 jako doktorant-stypendysta; grant pt: "Badanie funkcji i struktury białek ogonka bakteriofagów na przykładzie bakteriofagów *Yersinia Enterocolitica*". (kierownik projektu: dr hab. Ewa Brzozowska, IITD PAN; 10.2018 – 05.2022)
3. Wykonawca w projekcie Tango II: "Wykonanie badań mających na celu optymalizację warunków testu do diagnostyki zakażeń/nosicielstwa *S. agalactiae*". (kierownik projektu: prof. dr hab. Monika Brzychczy-Włoch; Collegium Medicum, Uniwersytet Jagielloński; 10.2019 – 12.2019)
4. Wykonawca w projekcie pt. „Badania nad uzyskaniem nowatorskiej szczepionki przeciwko wirusowi SARS-CoV-2 odpowiedzialnemu za chorobę COVID-19”. (kierownik projektu: prof. dr hab. Andrzej Gamian, IITD PAN; 01.2021 – 03.2022)

Uczestnictwo w konferencjach:

1. Udział w symposium “12th NanoTemper Technologies Symposium & Workshop – East” 22-23 listopada 2018, Kraków.
2. Prezentacja posteru na konferencji EUREKA, styczeń 2019; tytuł posteru: "Białka ogonka bakteriofagów – związki o potencjale antybakteryjnym"; autorzy: **Filik K.**, Ogorzelski F., Brzozowska E.
3. Prezentacja posteru na konferencji „Multi-omika – biologia systemów w badaniach medycznych” 28 listopada 2019 r. UMed Wrocław; tytuł posteru” Badanie funkcji

białek bakteriofagowych ze szczególnym uwzględnieniem zdolności do biodegradacji biofilmu”;

autorzy: **Filik K.**, Szermer-Olearnik B., Brykała J., Brzozowska E.

4. Prezentacja posteru na konferencji krystalograficznej we Wrocławiu- Joint Polish-German Crystallographic Meeting 2020 24 - 27 February 2020 • Wrocław/Poland; tytuł posteru: “Dual-function tail tubular proteins of bacteriophages”; autorzy: Pyra A., **Filik K.**, Brzozowska E.
5. Współautorstwo posteru prezentowanego na konferencji w Lipsku: 6th Joint Conference of the DGHM & VAAM – 72nd Annual Meeting of the DGHM & Annual Meeting of the VAAM Leipzig, Mar 8, 2020 bis Mar 11, 2020; Tytuł posteru: “Bacteriophages and their secrets: discovering unknown, lytic properties of tail tubular proteins from *Y. enterocolitica* and *K. pneumoniae*-infecting viruses”; autorzy: Urbanska N., Brzozowska E. , Pyra A. , **Filik K.**