

VILNIUS UNIVERSITY

EUGENIJUS ŠIMOLIŪNAS

**CONSTRUCTION OF SELF-ASSEMBLING
NANOSTRUCTURES BASED ON STRUCTURAL
PROTEINS OF BACTERIOPHAGES**

Summary of doctoral dissertation,
Physical sciences, Biochemistry (04 P)

Vilnius, 2017

This dissertation was prepared during the period of 2012–2016 at the Institute of Biochemistry, Vilnius University.

Supervisor – dr. **Rolandas Meškys** (Vilnius University, physical sciences, biochemistry – 04 P).

The dissertation is defended at the Council of Biochemistry Sciences of Vilnius University:

Chairman – prof. dr. **Nomeda Kuisienė** (Vilnius University, biomedical sciences, biology – 01 B).

Members:

prof. habil. dr. **Rimantas Daugelavičius** (Vytautas Magnus University, biomedical sciences, biology – 01 B);

assoc. prof. dr. **Elena Bakienė** (Vilnius University, physical sciences, biochemistry – 04 P);

dr. **Milda Plečkaitytė** (Vilnius University, physical sciences, biochemistry – 04 P);

dr. **Milda Stuknytė** (University of Milan, Italy, biomedical sciences, biology – 01 B).

The dissertation will be defended at the public session of the Council of Biochemistry Sciences at 3:00 pm on 21th of June, 2017, in the R402 auditorium at the Life Sciences Center, Vilnius University, (Sauletekio av. 7, LT-10257, Vilnius, Lithuania).

The summary of doctoral dissertation was sent on 19th of May, 2017.

This dissertation is available at the Library of Vilnius University and at the website of Vilnius University: <http://www.vu.lt/naujienos/ivykiu-kalendorius>

VILNIAUS UNIVERSITETAS

EUGENIJUS ŠIMOLIŪNAS

**BAKTERIOFAGŲ STRUKTŪRINIŲ BALTYMŲ
PANAUDOJIMAS SAVITVARKIŲ NANOSTRUKTŪRŲ
SURINKIMUI**

Daktaro disertacijos santrauka,
Fiziniai mokslai, Biochemija (04 P)

Vilnius, 2017

Disertacija rengta 2012–2016 metais Vilniaus universiteto Biochemijos institute.

Mokslinis vadovas – dr. **Rolandas Meškys** (Vilniaus universitetas, fiziniai mokslai, biochemija – 04 P).

Disertacija ginama Vilniaus universiteto Biochemijos mokslo krypties taryboje:

Pirmininkė – prof. dr. **Nomeda Kuisienė** (Vilniaus universitetas, biomedicinos mokslai, biologija – 01 B).

Nariai:

prof. habil. dr. **Rimantas Daugelavičius** (Vytauto Didžiojo universitetas, biomedicinos mokslai, biologija – 01 B);

doc. dr. **Elena Bakienė** (Vilniaus universitetas, fiziniai mokslai, biochemija – 04 P);

dr. **Milda Plečkaitytė** (Vilniaus universitetas, fiziniai mokslai, biochemija – 04 P);

dr. **Milda Stuknytė** (Milano universitetas, Italija, biomedicinos mokslai, biologija – 01 B).

Disertacija bus ginama viešame Biochemijos mokslo krypties tarybos posėdyje 2017 m. birželio mėn. 21 d. 15 val. Vilniaus universiteto Gyvybės mokslų centre, R402 auditorijoje (Saulėtekio al. 7, LT-10257, Vilnius, Lietuva).

Disertacijos santrauka išsiųsta 2017 m. gegužės mėn. 19 d.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje bei svetainėje adresu:

<http://www.vu.lt/naujienos/ivykiu-kalendorius>

CONTENT

LIST OF ABBREVIATIONS	6
INTRODUCTION.....	7
1. MATERIALS AND METHODS	10
2. RESULTS AND DISCUSSION	23
2.1 Characterization of a new bacteriophages	23
2.1.1 <i>Klebsiella</i> bacteriophage vB_KleM-RaK2	23
2.1.2 <i>Arthrobacter</i> bacteriophage vB_ArS-ArV2.....	25
2.1.3 <i>Arthrobacter</i> bacteriophage vB_ArtM-ArV1	26
2.1.4 <i>Escherichia coli</i> bacteriophage vB_EcoS_NBD2	28
2.1.4 <i>Escherichia coli</i> bacteriophage vB_EcoM-FV3	29
2.2 Investigation of structural proteins of new bacteriophages.....	29
2.2.1 Investigation of structural proteins of phage RaK2.....	30
2.2.2 Investigation of structural proteins of phages ArV1 and ArV2.....	31
2.2.3 Investigation of structural proteins of phage NBD2.....	31
2.2.4 Investigation of structural proteins of phage FV3	32
2.3 Investigation of nanostructures formed by gp053 of phage FV3	33
2.3.1 Bioinformatics analysis of gp053	33
2.3.2 Expression, purification and stability studies of gp053	34
2.3.3 Investigation of gp053 deletion mutants.....	37
2.3.4 Investigation of gp053 mutants with inserted cloning sites.....	40
2.3.5 Investigation of gp053-based hybrid proteins	43
2.3.6 Hybrids of gp053 polysheaths and gold nanoparticles	45
2.3.7 Investigation of oligomers of gp053 mutant.....	46
CONCLUSIONS	48
LIST OF PUBLICATIONS	49
FINANCIAL SUPPORT.....	53
CURRICULUM VITAE	54
ACKNOWLEDGEMENTS	55
REZIUMÉ	56
REFERENCES.....	58

LIST OF ABBREVIATIONS

a.a. – amino acid
ATCC – American Type Culture Collection
CPMV – cowpea mosaic virus
CPV – canine parvovirus
Cryo-EM – cryo-electron microscopy
DSMZ – Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
e.o.p. – efficiency of plating
EDTA – N,N,N',N'-ethylenediaminetetraacetic acid
FHV – flock house virus
gDNA – genomic DNA
HBV – hepatitis B virus
IPTG – isopropyl β -D-1-thiogalactopyranoside
LB – Luria-Bertani broth
LC-MS/MS – liquid chromatography–mass spectrometry
MCIC – MacConkey-inositol-carbenicillin agar
MOI – multiplicity of infection
NCBI – National Center for Biotechnology Information
NCP – noncontracted polysheaths
OD – optical density
ORF – open reading frame
PAA – polyacrylamide
PB – phosphate buffer
PCR – polymerase chain reaction
PDB – Protein Data Bank
PFU – plaque forming unit
PLTS – phage tail-like protein translocation structures
PVX – potato virus X
SCAI – Simmons citrate agar inositol
SDS-PAGE – sodium dodecyl sulfate polyacrylamide gel electrophoresis
T6SS – type VI secretion system)
TCP – tail completion protein
TEM – transmission electron microscopy
TMP – tape measure protein
TMV – tobacco mosaic virus
TRIS – 2-amino-2-(hydroxymethyl)propane-1,3-diol
TrP – tail terminator protein
TTC – tail tip complex
TTP – tail-tube protein
VNPs – virus-based nanoparticles

INTRODUCTION

Over recent decades, in the field of science and technology, particular attention is paid to construction of various nanostructures, based on a self-assembling biomolecules. Research in this area is in special interest due to the enormous potential of application of these structures in industry, medicine and other fields (Busseron et al., 2013; Lee et al., 2016). For this reason, a variety of biomolecules are used as possible candidates for self-assembly. Self-assembling nanostructures are constructed of nucleotides (Pinheiro et al., 2011; Zhang et al., 2014), peptides (Mandal et al., 2014; Sharma et al., 2015), proteins (Kostiainen et al., 2013; Glover et al., 2016) and viruses, including bacteriophages (Lee et al., 2012; Molek and Bratkovič, 2015; Pires, 2016).

In nanobiotechnology bacteriophages (and other viruses) or their structural proteins are attractive tools not only because of unique physical, chemical and thermodynamic properties of their proteins, but also because of the ability of these proteins to be easily modified both chemically and by genetic engineering (Pokorski and Steinmetz, 2011). In addition, the particle sizes, varying from tens to several thousand nanometers, is also particularly suitable for the formation of nanostructures (Douglas and Young, 2006). Finally, the efficient production of nanostructures and expression of proteins in cells, as well as convenient purification are cheap and environmentally friendly processes (Li and Wang, 2014).

To date, the studies on polymerization and nanostructure formation by the structural proteins from bacteriophages have been carried out with a limited number of specimens. Most of the investigations of structural proteins of bacteriophage into practical applications such as creation of biosensors, fabrication of energy storage devices, synthesis of nanobiopolymers or tissue engineering have been made using bacteriophage M13 and its relatives (Moon et al., 2015; Pires et al., 2016). Self-assembling nanostructures based on the structural proteins of bacteriophages T4 (Daube et al., 2007; Yokoi et al., 2010), P22 (Bhardwaj et al., 2008) or phi29 (Guo, 2005) have been also constructed. On the other hand, a tremendous diversity of the structural proteins from the less studied phages have been not yet examined neither in any detail nor as the targets for genetic manipulations.

This work was carried out in order to expand current knowledge and explore a biotechnological potential of the self-assembling nanostructures of structural proteins from newly identified bacteriophages. In addition, a construction of the mutants or hybrid proteins, based on the self-assembling structural proteins from phages, as well as the analysis of nanostructures formed by these proteins have been carried out.

The aim of this study was: to construct self-assembling hybrid nanostructures using the structural proteins of bacteriophages.

The following tasks have been formulated to achieve this aim:

1. To identify and characterize new bacteriophages and their structural proteins.
2. To determine which of the identified structural proteins are suitable for the construction of self-assembling nanostructures *in vivo* or *in vitro*.
3. To construct mutants of the structural proteins of bacteriophages and investigate the structures they form.
4. To construct the hybrid proteins based on the structural proteins of bacteriophages and investigate the structures formed by these proteins.

Scientific novelty and practical value

In this study five new bacteriophages have been identified and characterized. One of these phages, *Klebsiella* sp. infecting phage RaK2, with a genome size of 345,809 bp, was the second largest bacteria infecting virus and the largest *Klebsiella* spp. phage with completely sequenced and in NCBI database deposited genome at a time. Also, for the first time in NCBI database was published genome of *Arthrobacter* sp. infecting bacteriophage (phage ArV2). Furthermore, another object of this study –*Arthrobacter* phage ArV1 was the first published *Arthrobacter* spp. infecting myovirus with completely sequenced and in NCBI database deposited genome at a time. During this study, morphology, physiology, bioinformatics analysis of bacteriophages and many other studies were made. The results of these studies are essential for deepening the theoretical knowledge herewith rising awareness of practical application abilities of these the most abundant entities on the Earth.

Although investigation of nanostructures formed by recombinant tail sheath proteins from bacteriophages have been started a long time ago (most of these works were carried out with *E. coli* bacteriophage T4), these structures were examined in order to understand the polymerization properties of a protein and not for application purposes in nanobiotechnology. To our knowledge, no literature data about construction of self-assembling structures made of the tail sheath-based hybrid proteins of bacteriophages were published to date. For this reason, an investigation of nanostructures based on the tail sheath protein (gp053) from *E. coli* phage FV3 as well as construction of gp053 mutants and hybrid proteins are particularly important in finding new, alternative biomolecules for synthesis of self-assembling nanostructures.

For the first time, this study demonstrated, that the recombinant tail tube protein (gp39) from bacteriophage NBD2 polymerized into ordered, flexible, and particularly long (up to a 3 μm) tubular structures *in vivo* in the absence of other viral proteins. These results are important not only for the construction of new, bacteriophage structural protein-based self-assembling nanostructures, which could be applied in the future in various fields of nanobiotechnology, but also provide new knowledge about the bacteriophage tail tube protein properties.

Thesis statements:

1. Recombinant gp041 from *Klebsiella* phage RaK2, gp39 from Enterobacteria phage NBD2 and gp053 from *Escherichia* phage FV3 *in vivo*, in the absence of other phage proteins, self-assemble into long, regular, tubular structures.
2. Polysheaths formed by recombinant gp053 from FV3 are very stable and especially resistant structures to various environmental factors.
3. C-terminal amino acid deletions have less impact on the polymerization properties of gp053 from FV3 than truncations of the N-terminus of this protein.
4. FV3 gp053 mutants and hybrid proteins, based on gp053, self-assemble into nanotubular structures with properties depending on the number of removed or inserted amino acids as well as on their position in the protein sequence.

1. MATERIALS AND METHODS

Bacteriophages

Bacteriophages used in this study were: *Klebsiella* sp. infecting phage vB_KleM-RaK2 (RaK2), *Escherichia coli* phages vB_EcoM-FV3 (FV3) and vB_EcoS_NBD2 (NBD2), *Arthrobacter* spp. phages vB_ArS-ArV2 (ArV2) and vB_ArtM-ArV1 (ArV1). All these bacteriophages were isolated from samples collected in different places in Lithuania.

Bacterial strains

Bacterial strains used in this study are listed in Table 1.

Plasmids and oligonucleotides

Plasmids used in this study were: pJET1.2, pET16b, pET21a, pET21b, pET28a. Sequences of all oligonucleotides and primers used in this study are listed in Table 2 and Table 3.

Chemicals and enzymes

All chemicals used in this study were analytically pure. T4 Polynucleotide kinase, T4 DNA ligase, FastDigest restriction enzymes, Rnase A, Dnase I, proteinase K, Pfu and DreamTaq DNA polymerases and Phusion DNA polymerases were obtained from Thermo Fisher Scientific. All these products were used according to manufacturer's protocols.

Table 1. Bacterial strains used in this study.

Bacterial strain	Purpose of use	Reference
<i>Acinetobacter baumannii</i> #16	RaK2 host range determination	Prof. E. Sužiedėlienė
<i>Acinetobacter baumannii</i> #46	ArV1, ArV2, RaK2 host range determination	Prof. E. Sužiedėlienė
<i>Acinetobacter</i> gen. sp. 13#23	ArV1, ArV2, RaK2 host range determination	Prof. E. Sužiedėlienė
<i>Acinetobacter</i> gen. sp. 3#9	RaK2 host range determination	Prof. E. Sužiedėlienė
<i>Arthrobacter alkaliphilus</i> DSM 23368	ArV1 host range determination	DSMZ
<i>Arthrobacter aurescens</i> DSM 20116	ArV1 host range determination	DSMZ
<i>Arthrobacter chlorophenolicus</i> DSM 12829	ArV1, ArV2 host range determination	DSMZ
<i>Arthrobacter citreus</i> DSM 20133	ArV1 host range determination	DSMZ
<i>Arthrobacter crystallopoetes</i> DSM 20117	ArV1, ArV2, RaK2 host range determination	DSMZ
<i>Arthrobacter defluvii</i> DSM 18782	ArV1, ArV2 host range determination	DSMZ
<i>Arthrobacter gandavensis</i> DSM 15046	ArV1 host range determination	DSMZ
<i>Arthrobacter globiformis</i> DSM 20124	ArV1, ArV2 host range determination	DSMZ
<i>Arthrobacter globiformis</i> NRRL B-2979	RaK2 host range determination	Casaite et al, 2011
<i>Arthrobacter histidinovorans</i> DSM 20115	ArV1, ArV2 host range determination	DSMZ
<i>Arthrobacter ilicis</i> DSM 20138	ArV1, ArV2 host range determination	DSMZ
<i>Arthrobacter koreensis</i> DSM 16760	ArV1 host range determination	DSMZ
<i>Arthrobacter luteolus</i> DSM 13067	ArV1 host range determination	DSMZ
<i>Arthrobacter methylophilus</i> DSM 14008	ArV1, ArV2 host range determination	DSMZ
<i>Arthrobacter nicotinovorans</i> DSM 420	ArV1, ArV2 host range determination	DSMZ
<i>Arthrobacter nitroguajacolicus</i> DSM 15232	ArV1, ArV2 host range determination	DSMZ
<i>Arthrobacter oxydans</i> DSM 20119	ArV1, ArV2 host range determination	DSMZ
<i>Arthrobacter</i> sp. 68b	ArV1, ArV2 isolation, propagation, adsorption experiments, host range determination	Stanislauskienė et al. 2011
<i>Arthrobacter</i> sp. 68m	ArV1, ArV2 host range determination	laboratory collection
<i>Arthrobacter</i> sp. 83b	ArV1, ArV2 host range determination	laboratory collection
<i>Arthrobacter</i> sp. 85	ArV1, ArV2 host range determination	laboratory collection
<i>Arthrobacter</i> sp. 94	ArV1, ArV2 host range determination	laboratory collection

See next page for Table 1 extension

Extension of the Table 1

<i>Arthrobacter</i> sp. 96	ArV1, ArV2 host range determination	laboratory collection
<i>Arthrobacter</i> sp. 25DMP1	ArV1, ArV2 host range determination	Kutanovas et al. 2013
<i>Arthrobacter</i> sp. 25DOT1	ArV1, ArV2 host range determination	Kutanovas et al. 2013
<i>Arthrobacter</i> sp. BL-3	RaK2 host range determination	Stanislauskiene et al. 2012
<i>Arthrobacter</i> sp. IN13	ArV1, ArV2 host range determination	Gasparavičiūtė et al. 2006
<i>Arthrobacter</i> sp. PY11	ArV1, ArV2 host range determination	Semėnaitė et al. 2003
<i>Arthrobacter</i> sp. PY21	ArV1, ArV2, RaK2 host range determination	Stanislauskiene et al. 2012
<i>Arthrobacter</i> sp. PY22	ArV1, ArV2, RaK2 host range determination	Stanislauskiene et al. 2012
<i>Arthrobacter</i> sp. PRH1	ArV1, ArV2 host range determination	Stanislauskiene et al. 2012
<i>Arthrobacter</i> sp. VM22	ArV1, ArV2 host range determination	Semėnaitė et al. 2003
<i>Arthrobacter</i> sp. VP3	ArV1, ArV2 host range determination	Semėnaitė et al. 2003
<i>Arthrobacter karbamidofaciens</i> DSM 20126	ArV1, ArV2 host range determination	DSMZ
<i>Buttiauxella</i> sp. S1-1	RaK2 host range determination	Šimoliūnas et al., 2013
<i>Citrobacter freundii</i>	ArV1, ArV2, RaK2 host range determination	Prof. E. Sužiedėlienė
<i>Enterobacter cloacae</i>	ArV1, ArV2, RaK2 host range determination	Prof. E. Sužiedėlienė
<i>Enterobacter</i> sp. VT1-1	RaK2 host range determination	Šimoliūnas et al., 2013
<i>Erwinia carotovora</i> 8982	ArV1, ArV2, RaK2 host range determination	Prof. E. Sužiedėlienė
<i>Erwinia carotovora</i> 961-63	RaK2 host range determination	Prof. E. Sužiedėlienė
<i>Escherichia coli</i> B ^E	ArV1, ArV2, RaK2 host range determination	Dr. L. W. Black
<i>Escherichia coli</i> BL21 (DE3)	ArV1, ArV2, RaK2 host range determination, plasmid recipient	Avidis
<i>Escherichia coli</i> B834 (DE3)	RaK2 host range determination	Dr. L. W. Black
<i>Escherichia coli</i> CR63	RaK2 host range determination	Dr. K. N. Kreuzer
<i>Escherichia coli</i> DH5 α	ArV1, ArV2, RaK2 host range determination	Pharmacia
<i>Escherichia coli</i> DH10 β	ArV1, ArV2, RaK2 host range determination, plasmid recipient	Invitrogen
<i>Escherichia coli</i> GM2163	RaK2 host range determination	Fermentas
<i>Escherichia coli</i> JM109	RaK2 host range determination	Fermentas
<i>Escherichia coli</i> MH1	RaK2 host range determination	Dr. K. N. Kreuzer
<i>Escherichia coli</i> Nova Blue (DE3)	RaK2 host range determination	Avidis

See next page for Table 1 extension

Extension of the Table 1

<i>Escherichia coli</i> Rosetta (DE3)	Plasmid recipient	Novagen
<i>Escherichia coli</i> XL1 Blue	RaK2 host range determination	Stratagene
<i>Klebsiella oxytoca</i> ATCC 8724	RaK2 host range determination	ATCC
<i>Klebsiella pneumoniae</i> ATCC BAA-1705	RaK2 host range determination	ATCC
<i>Klebsiella pneumoniae</i> 279	ArV1, ArV2, RaK2 host range determination	Prof. E. Sužiedēlienē
<i>Klebsiella</i> sp. KV-1	RaK2 host range determination	Šimoliūnas et al., 2013
<i>Klebsiella</i> sp. KV-3	ArV1, ArV2, RaK2 host range determination, RaK2 isolation, propagation and other experiments	Šimoliūnas et al., 2013
<i>Kocuria palustris</i> DSM 11925	ArV1 host range determination	DSMZ
<i>Kribella catacumbae</i> DSM 19601	ArV1 host range determination	DSMZ
<i>Nesterenkonia aethiopica</i> DSM 17733	ArV1 host range determination	DSMZ
<i>Pseudomonas aeruginosa</i> PAO1	ArV1, ArV2, RaK2 host range determination	Prof. E. Sužiedēlienē
<i>Pseudomonas brenneri</i> D14	RaK2 host range determination	Prof. E. Sužiedēlienē
<i>Pseudomonas</i> sp. PV1-1	RaK2 host range determination	Šimoliūnas et al., 2013
<i>Pseudomonas</i> sp. RA1-1	RaK2 host range determination	Šimoliūnas et al., 2013
<i>Pseudomonas</i> sp. RA1-3	RaK2 host range determination	Šimoliūnas et al., 2013
<i>Pseudomonas</i> sp. RA1-11	RaK2 host range determination	Šimoliūnas et al., 2013
<i>Rhodococcus erythropolis</i> SQ1	RaK2 host range determination	Quan et al., 1993
<i>Rhodococcus</i> sp. PY11	RaK2 host range determination	Semėnaitė et al, 2000
<i>Rothia aeria</i> DSM 14556	ArV1 host range determination	DSMZ
<i>Salmonella enterica</i> ser. Typhimurium 292	ArV1, ArV2 host range determination	Prof. E. Sužiedēlienē
<i>Solitalea canadensis</i> DSM 3403	ArV1 host range determination	DSMZ
<i>Yaniella soli</i> DSM 22211	ArV1 host range determination	DSMZ

Table 2. Oligonucleotides used in this study for cloning structural genes from different bacteriophages.

Gene (protein function)	Primers used for gene amplification and sequence of oligonucleotides (5'→3')	Plasmid
RaK2_g041 (tail sheath protein)	Rak2_orf041_Nco_F ggaggattttccatggcagatttaatc Rak2_041_Xho_R_n ccccttttgatactcgagagtatttc	pET21d
RaK2_g042 (tail tube protein)	RaK2_42NdeF ggagaaattatcatatggcg RaK2_42XhoR gctccgacgtctcgagcatcg	pET21a
RaK2_g043 (hypothetical protein)	RaK2_43NdeF cagaggttaacatattggctgatatg RaK2_43XhoR caagtacagtctcgagaacatttg	pET21a
RaK2_g106 (capsid decoration protein)	RaK2106NdeIF ttgactaaggagattttcatatgagtc RaK2106BamHIR cggtattagtagtattaggatccatag	pET16b, pET21a
RaK2_g107 (capsid decoration protein)	RaK2107NdeIF attacatattggctactaataactactaataac RaK2107BamHIR tccagacattctctcgaggatttaaac	pET16b, pET21a
ArV1_g15 (tail sheath protein)	ArV1_put_tail_sheatF gagacaggagaaaacatattggctattg ArV1_put_tail_sheatR gttgtctttgtctcgagtcagttattc	pET16b, pET21a
ArV2_g03 (portal protein)	ArV2_portal_NdeIF ggtgggtgcacccatattgtaatttc ArV2_portal_BamHIR gttgcaactgtacgggatccacgtcg	pET16b
ArV2_g05 (capsid maturation protein)	ArV2_scaffold_NdeIF gacggcggtcatattgtctggtag ArV2_scaffold_BamHI gcgggtcaacaggatccttagtgcc	pET16b
ArV2_g06 (major capsid protein)	ArV2_majcap_NdeIF ctttaggagtttcatattggcgcag ArV2_majcap_BamHIR ggtcacctcgcggatccccaccag	pET16b
ArV2_g11 (major tail protein)	ArV2_maj-tail_NdeIF ggaaaccgaaaacatattgcccttg ArV2_maj-tail_BamHIR gttgagagcgggagatccaccac	pET16b
ArV2_g15 (tape measure protein.)	ArV2_tape-m_NdeIF gccagctctgccgggtctcatattggccg ArV2_tape-m_XhoIR tgggtccggcggagctcgagggatgg	pET16b
ArV2_g16 (tail protein)	ArV2_tail(1)_NdeIF catttagggaggccatattggtgggtatc ArV2_tail(1)_XhoIR caataaactcccaccctcgagaaacca	pET16b
ArV2_g17 (tail protein)	ArV2_tail(2)_NdeIF gtcactaacaagatcatattggcgttc ArV2_tail(2)_XhoIR ctcttctgtactgctcgagccattag	pET16b
ArV2_g18 (tail protein)	ArV2_min-tail_NdeIF catggacaggatggcccgcattatggc ArV2_min-tail_BamHIR gccagccattgcagcggatccac	pET16b
ArV2_g20 (tail fiber protein)	ArV2_adhes_NdeIF ccggaggacaaagcatattgagtgaatc ArV2_adhes_BamHIR cggcggcgggaactggatccatcatg	pET16b
ArV2_g24 (tail fiber protein)	ArV2_tail(3)_NdeIF gaggaggtcggttccatattgggttac ArV2_tail(3)_BamHIR gcggctttgggatcctgctcgcgttc	pET16b
NBD2_g39 (tail tube protein)	NBD2-mtail-NdeIF caaaggagtttcatattgtctcttc NBD2_mtail_BamHIR ctcttgggatccagtcgc	pET16b, pET21a
	NBD2-mtail-NdeIF caaaggagtttcatattgtctcttc NBD2_mtail_XhoR gcaatgcccccttctggctcggagaacttc	pET21a
FV3_g052 (tail tube protein)	FV52_Nde_F1 gactgaataaggaagacatattgaaactc FV52_BamHI_R1 cgaaggtgaggatcctattgctaaagtg	pET21a, pET16b
	FV52_Nde_F1 gactgaataaggaagacatattgaaactc FV52_XhoI_R1 gcctattgctaaagtgtgctcgaggataatacc	pET21a
FV3_g53 (tail sheath protein)	FV53_NdeI_F1 ggaacatattgcatatttagataaagtg FV-53_BamHI_R1 ttgaggatccatctagctctccttattc	pET21a, pET16b
FV3_g53 (tail sheath protein)	FV53_NdeI_F1 ggaacatattgcatatttagataaagtg FV-53_XhoI_R1 catctagctctcctcgagttcagtcattg	pET21a

Table 3. Oligonucleotides used in this study for construction of gp053 mutants and hybrid proteins of FV3.

Primer	Sequence (5'→3')
FV53_NdeI_F1	ggaaacatatgccatatttagataaagtg
FV53_BamHI_R1	ttgaggatccatctagtcttcttattc
FV53_XhoI_R1	catctagtcttctcgcagttcagtcattg
FV53_Nde10F	taaagtggtacatatgacagttaac
FV53_Bam10R	cagcattggtggatccttaaccaac
FV53_Xho10R	cagcattggtggactcgcagaccaac
FV53_Nde20F	cccagcctattcatatggtagggttc
FV53_Bam20RN	gtccagaaggatccttaccggatcag
FV53_Xho20R	gaaggacgttgatctcgcagacgaac
FV53_Nde25F	ctgtagggttccatatgccgctg
FV53_Bam25R	gttgatccggatccgaacttagtgc
FV53_Xho25R	gatccggatcgcctcgcaggtgcag
FV53_Nde30F	gaaactccgctgcatatggctatcc
FV53_Bam30R	cgaagtgccggtccttgttttacac
FV53_Bam50R	ctgattggatccgatctatgcacg
FV53_Bam75R	ctgtcaggatccctcaattcttaag
FV53_Bam100R	cggagatcaggatccgactactcatg
FV53_Bam150R	caacaccgtggatcctcagtagaag
FV53_Bam200R	cacctatgataggatcctacggaaaatc
PilnasMaF	gatatacatatggctagcacaggtaccccatatttagataaagtgg
PilnasMaR	gtggtgctcgcagttacgtacgtgtactagttcagtcattgcaacagc
PilnasMaRH	gtggtgctcgcagcgtacgtgtactagttcagtcattgcaacagc
10MaF	gatatacatatggctagcacaggtaccgtaacctcggaaaccagc
10MaR	gtggtgctcgcagttacgtacgtgtactagtaccaactttgtccagaagg
10MaRH	gtggtgctcgcagcgtacgtgtactagtaccaactttgtccagaagg
20MaF	gatatacatatggctagcacaggtaccgtaggtttcgaaactccgc
20MaRH	gtggtgctcgcagcgtacgtgtactagtatacgaacgaagtgcagg
SmaBamR	gagttgcacctgttccaccagggatccccgggatcagagaaagtattgtag
SmaBamF	ctacaatacttctctgataccgggatccctggtggaacaggtgcaactc
SBv2BamR	cactaggatccccgggattgtactaccttttgaata
SBv2SmaF	acaatccccgggatcctagtgcctatctctggtggaac
SBv3BamR	agtatggatccccgggttttgaatagtcattacgaca
SBv3SmaF	caaaaccgggatccatacttctctgatagtgcctat
SBv4BamR	accttggatccccgggcattacgacatcagatcagct
SBv4SmaF	taatgccccgggatccaaggtagctacaatacttctct
SBv5BamR	tgcacggatccccgggagagataggcactatcagagaa
SBv5SmaF	tcttccccgggatccgtgcaactcagattgcagagtc
SBv6BamR	tgactggatccccgggagttgcacctgttccaccagag
SBv6SmaF	caactccccgggatccagtcattgaagactaaaattga
SBv7BamR	ttcaaggatccccgggcattgactctgcaatctgagtt
SBv7SmaF	caatgccccgggatcctgaagcagatgatagcttgaa
Kilpa239R	ttgtgggatccccgggtttggtggcgatgttactatct
Kilpa239F	caaaccgggatccacaatatgatagcgtcggcatg
Kilpa2392R	ttgtgggatccccgggtttggtggcgatgttactatct

See next page for Table 3 extension

Extension of the Table 3

Kilpa2392F	caaaccgggatccacaatatgatagcgtcggcatg
FV53_Sma75R	caggatctcgagctactggatcccgggattgatcgggtgttcac
FV53_Gly_Xho	ggagttaaattctcggaggcttctcctcctcctcagtcattgccaac
100C_Gly_Xho	gaacagacttgataacctcgaggcttctcctcctcctcctcatgcttaaac
AgE6F	ctagcgaagaagaagaagaagaag
AgE6R	gtaccttcttcttcttcttctcg
Ti1F	ctagccgtaaaaaacgtacaaaaacccgaccacaaaactgggtgggtgggg
Ti1R	gtacccaaccaccaccagtttggtggtcgggttttggtacgtttttacgg
SmaAgD6F	gacgacgacgacgacgacgg
SmaAgD6R	gatcccgtcgtcgtcgtcgtcgtc
SmaAgE6F	gaagaagaagaagaagaagg
SmaAgE6R	gatcccttcttcttcttcttctc
SmaAgE6F	gaagaagaagaagg
SmaAgE6R	gatcccttcttcttcttctc
Pfl23Strep	tggtgctcgagttattttcgaactgcccgggtggctccacgtacgttcagtcattgcaacagc

Bacteria cultivation

Bacteria were cultivated in Luria-Bertani (LB) broth at 28–37°C until the cell suspension (A_{600nm}) reached 0.4–1. Bacterial culture was used for bacteriophage isolation, propagation, titration, host-range determination, efficiency of plating (e.o.p.), adsorption and single step experiments as well as the preparation of competent cells and gene expression experiments. Bacterial cultures were also grown on LB agar broth at 18–48°C from 12 h up to 7 days (*Kribella catacumbae* DSM 19601 strain was cultivated on R2A agar medium). MCIC and SCAI selective media were used for isolation of *Klebsiella* spp. related bacterial strains.

Isolation of bacteriophages

Isolation of bacteriophages from soil samples

Soil samples (1–10 g) were collected in sterile test tubes or bottles and mixed with 10 ml of LB medium. After 15–60 min of incubation by shaking at 30–37°C the mixture was centrifuged at 3000 g for 15 min. The supernatant fluid was sequentially filtered through sterile 0.45 and 0.2 mm membrane filters and assayed for plaque forming units by the soft agar overlay method. Additional concentration of filtered samples was performed by centrifugation for 1 h at 16,000 g, 4°C. The supernatant was decanted and the pellet resuspended in 1 ml of PB buffer and stored at 4°C for at least 2 hours or overnight.

Isolation of bacteriophages from water samples

Water samples (10–100 ml) were collected in sterile test tubes or bottles and centrifuged for 15 min at 3000 g, 4°C. The supernatant was sequentially filtered through sterile 0.45 and 0.2 mm membrane filters and used for titration or additional concentration procedures as described previously.

Propagation of bacteriophages

Propagation of bacteriophages using LB medium

E. coli strain MH1 was used for bacteriophage FV3 propagation, while *Klebsiella* sp. KV3 was used for bacteriophage RaK2 propagation. Bacteria were cultivated in LB broth at 37°C (*E. coli* strain MH1) or 30°C (*Klebsiella* sp. KV3) until the cell suspension ($A_{600\text{nm}}$) reached 0.4–0.5. Then, the bacteria culture was infected by a single lysis zone content (fresh phage lysis zone was suspended in 10 µl of LB media) and continued to grow until full or partial bacterial lysis occurred. Lysed bacteria culture was centrifuged for 1 h at 16,000 g, 8°C. The supernatant was decanted and the pellet was resuspended in 1 ml of PB buffer. Additionally, 1/10 part of CHCl_3 and 2 µl of DNaseI (1 u/µl) were added to the suspension. The mixture was incubated for 1 h at 37°C, then suspended and placed in a new tube. A disrupted bacterial cell pellets were removed by centrifugation at 3000 g for 10–15 min. The supernatant with phages was collected and stored at 4°C.

Propagation of bacteriophages using double agar method

Propagation of phages ArV1 and ArV2 was performed using a standard double agar method described by Adams et al. (Adams et al., 1959) with few modifications using *Arthrobacter* strain 68b as a host. For propagation of bacteriophages, fresh (within 24 hours) titrating plates with infectious centers occupying about ½ of an area were used. Phage particles were subsequently collected by adding 3 ml of LB broth to the surface of each plate. The top agar was scraped off and the suspension recovered. After 15–30 min of incubation at 4°C with mild stirring, the mixture was centrifuged at 6000 g for 15 min. The phage-containing supernatant was decanted and filtered through sterile 0.45 and 0.2 mm membrane filters. Further purification was performed using a CsCl step gradient.

Titering of bacteriophages

Bacteriophages ArV1 and ArV2 were titrated on *Arthrobacter* sp. strain 68b, phage RaK2 was titrated on *Klebsiella* sp. KV3, phage FV3 was titrated on *E. coli* strains. Titering of bacteriophages was performed using a standard double agar method with a few modifications. Briefly, 0.1 ml of diluted phage suspension was mixed with 0.5 ml of indicator cells, then the mixture was added to 2.5 ml of 0.5% (w/v) soft agar and poured over the 1.2% LB agar plate as a uniform layer. The plates were incubated for 12–48 h at 28–37°C before the enumeration of plaques.

Host-range determination of bacteriophages

Bacteriophage host-range determination was performed by titration or spot-tests using bacteria strains described in Table 1. At the beginning of spot-tests, 500 µl of an exponential growth phase bacterial culture (A_{600nm} 0.4–1) was mixed with 3 ml of soft agar (0.5%), the mixture was then quickly spread on a Petri dish with the LB agar medium (1.2%). Plates were dried for 30 min and spot-tests were carried out by dropping 10 or 20 µl of diluted phage stocks on them. The plates were then incubated for 12–48 h at 28–37°C.

Bacteriophage efficiency of plating, adsorption and single step experiments

Determination of the efficiency of plating (e.o.p.) was performed as described by Kaliniene et al. (Kaliniene et al., 2010). Plates were incubated in temperature ranging from 10 to 47°C, incubation time varied from 12 h up to 7 days. Optimal phage development temperature was determined on the basis of the number of plaque forming units (PFU), as well as their size and morphology. The adsorption tests were carried out as described by Kropinski (Kropinski, 2009) with a few modifications. Bacterial culture at the exponential growth phase was mixed with bacteriophage suspension to a final MOI of 0.1 (bacteria and bacteriophage ratio in a mixture is 10:1) and incubated at 30–37°C for 15–20 min. Sampling was executed in a certain time intervals. The bacteria and complexes of bacteria and bacteriophages, attached on their surface, were removed from suspension by addition of chloroform (1/10 volume of the solution) and centrifugation at 5000 g for 5 min. The number of unadsorbed phage particles in the supernatant was determined by titration using a standard double agar method. The one-step growth (or

single step) experiments were carried out as described by Carlson and Miller (Carlson and Miller, 1994).

Purification of bacteriophages using a CsCl step gradient

Phage purification using a CsCl step gradient was performed as described by Sambrook et al. (Sambrook et al., 2001) with a few modifications. Concentrated suspension of phage was deposited on the top of CsCl step gradient (densities: 1.1 g/ml, 0.9 g/ml, 0.7 g/ml, 0.5 g/ml) and centrifuged in a Spinco SW39 rotor for 2–3 h at 24,000 rpm, 4°C. The resulting phage band with the highest opalescence was collected with a syringe and dialyzed against three changes of phage dialysis buffers at 4°C. Finally, phage was dialyzed against SM buffer (100 mM NaCl, 8 mM MgSO₄, 50mM Tris-HCl, pH 7.5) and stored at 4°C.

Sample preparation for TEM analysis

Phage particles purified by CsCl density gradient were diluted to approximately 10¹¹ PFU/ml with distilled water, 5 µl of the sample were directly applied on the carbon-coated nitrocellulose (produced by prof. J. Staniulis), copper or nickel (Agar Scientific) grids. After 2–5 min of incubation, excess liquid was drained with filter paper before staining with two successive drops of 2% uranyl acetate (pH 4.5), then dried and examined in MorgagniTM 268(D) or Tecnai G2 F20 X-TWIN (FEI, Oregon, USA) transmission electron microscopes. Samples of purified or in soluble fraction presented recombinant proteins were prepared for TEM analysis accordingly.

Isolation of phage DNA

Aliquots of phage suspension were subjected to phenol/chloroform extraction and ethanol precipitation as described by Kricker and Carlson (Kricker and Carlson, 1994) with a few modifications. 100–200 µl of phage suspension (10¹¹–10¹² PFU/ml) was thoroughly mixed with an equal volume of TE buffer (pH 8.0) saturated with phenol. The mixture was centrifuged for 5 min at 6 000 g. After centrifugation, the upper fraction was gently collected into a new tube and an equal volume of phenol and chloroform (1:1) was added. The solution was thoroughly mixed and centrifugated for 5 min at 6000 g and again upper fraction was collected into a new tube. The procedure was repeated 3 times. Final collected phage suspension was poured over 600 µl of cold 96%

ethanol and 10 μ l of 3 M ammonium acetate. Formed DNA "thread" was quickly gathered by a sterile glass rod, which was then left to dry, washed with 2–3 drops of 70% ethanol, and dried again. Dry glass rod was covered with 50 μ l of water without nucleases for DNA to elute. Isolated phage DNA was subsequently used in restriction analysis, for PCR or was subjected to genome sequencing. Phage gDNA has also been isolated using Quick-gDNA Miniprep Kit (Zymo Research) according to the manufacturer's recommendations.

Isolation of plasmid DNA

Plasmid DNA was isolated using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) in accordance with the manufacturer's recommendations. This way isolated plasmid DNA was used in restriction analysis, sequencing and transformation reactions.

Polymerase chain reaction (PCR) and agarose gel electrophoresis

DNA fragments were multiplied by polymerase chain reaction (PCR). Phage genomic DNA, purified phage virions or suspension of bacterial clones were used as a matrix for reaction. The PCR was carried out using specific synthetic oligonucleotide primers and the polymerase chain reaction kits (Thermo Fisher Scientific). 30 reaction cycles were performed according to standard schemes proposed by the manufacturer's, due to the primer melting temperatures and the length of amplifying fragments. The amplified products were detected by electrophoresis and visualized using UV. The PCR products were purified using GeneJET Gel Extraction and DNA Cleanup Micro Kit according to the manufacturer's recommendations.

Construction of plasmid vectors

To construct plasmid vectors, containing targeted genes, purified PCR products (as well as vectors) were digested using the appropriate restriction endonucleases (Thermo Fisher Scientific). Reaction buffers, amount of reagent, temperature, and restriction time were chosen according to the manufacturer's recommendations. Reaction products were analyzed in the agarose gel electrophoresis, further purification of fragments was performed by using GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Fisher Scientific). Ligation of plasmid vectors and DNA fragments was performed using T4 ligase (Thermo Fisher Scientific) following the manufacturer's

ligation protocol. The ligation reactions were also carried out using the Rapid DNA Ligation Kit (Thermo Fisher Scientific) in accordance with the manufacturer's recommendations.

Bacterial transformation

The chemical calcium-dependent transformation of competent cells was carried out as described by Mandel and Higa (Mandel and Higa, 1970). Electroporation of competent *E. coli* cells was performed using a standard electroporation protocol described by Sambrook and Rusell (Sambrook and Rusell, 2001). The positive transformants were selected by screening on selective LB agar medium with ampicillin (50 µg/ml).

Bioinformatic analysis

Analysis of sequenced DNA fragments or genome sequence was performed using Fasta-Protein, Fasta-Nucleotide, Fasta-Genome, BLASTP, Transeq (http://www.ebi.ac.uk/Tools/st/emboss_transeq), Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo>), Sequence Editor (<http://www.fr33.net/seqedit.php>) and Geneious v5.5.6. (<http://www.geneious.com/>). tRNAscan-SE 1.21 was used to search for tRNAs (<http://lowelab.ucsc.edu/tRNAscan-SE/>). Artemis was used for comparative genomic analysis (Carver et al., 2008). Phylogenetic analysis was performed using MEGA 5.0 (Tamura et al, 2011). Search for the gp053 fold was done using the HHpred (<https://toolkit.tuebingen.mpg.de/hhpred>) (Söding et al, 2005). Three-dimensional (3D) structure of gp053 was predicted using I-TASSER server (Zhang, 2008).

Production of recombinant proteins

E. coli BL21-DE3 cells with recombinant plasmid were grown in LB medium containing ampicillin (50 mg/ml) at 37°C until A_{600} reached 0.5. Protein expression was induced with IPTG to a final concentration of 0.1 mM. The cell growth temperature then was lowered to 30°C and the incubation was carried out for 3 hours (in the case of proteins with a low solubility or for the production of a higher amount of biomass, after induction, the cell culture was grown 16–18 hours at 20°C). Cells were harvested by centrifugation for 10 min at 4000 g, resuspended in TE (20 mM Tris–HCl (pH 7.8), 1

mM EDTA) or His-Wash (50 mM sodium phosphate buffer (pH 7.7), 300 mM NaCl, 50 mM imidazole, 0.03% Triton X-100) buffers and disrupted by sonication. Crude extracts were centrifuged at 4°C for 15 min at 21,000 g to remove the debris. Supernatant and pellets were directly analyzed by SDS-PAGE as well as by TEM. SDS-PAGE was carried out in 4.5% concentrating and 14% running gel using the Laemmli method (Laemmli, 1970). After electrophoresis gels were stained with Coomassie Brilliant Blue dye (Thermo Fisher Scientific) and washed in 7% acetic acid solution.

Purification of recombinant proteins

Protein purification using metal-chelating sorbent

Recombinant his-tagged proteins were purified using Zymo Research His-Spin Protein Miniprep kit according to manufacturer's recommendations.

Protein purification by precipitation with ammonium sulfate

The polysheaths formed by recombinant gp053 were precipitated from the supernatant by addition of ammonium sulfate to a final concentration of 10%. After incubation for 10 minutes on ice and centrifugation for 15 min at 9000 g, 4°C, the supernatant was removed, pellet was suspended in TE or HEPES buffer (1/10 of the initial volume) and stored at 4°C. Shorter, less ordered polysheaths were purified by addition of ammonium sulfate to a final concentration of 15–20% and/or by repeated centrifugation for 15 min at 9000 g, 4°C. Concentration of purified proteins was determined using method described by Lowry (Lowry et al., 1951).

2. RESULTS AND DISCUSSION

The aim of this study was to construct self-assembling hybrid nanostructures based on the structural proteins of newly isolated bacteriophages. Thus, a detailed characterization of new bacteriophages was carried out. Morphology, physiology as well as genome bioinformatics analysis and proteomic approaches were used to characterize new bacteriophages and also to identify their structural proteins, some of which were chosen for a more detailed analysis.

2.1 Characterization of a new bacteriophages

2.1.1 *Klebsiella* bacteriophage vB_KleM-RaK2

Klebsiella sp. infecting bacteriophage vB_KleM-RaK2 (abbreviated name RaK2) – a giant singleton virus of the family *Myoviridae*. The unique feature of this bacteriophage – particularly large genome (345,809 bp). In 2012, after publishing the genome of bacteriophage in NCBI database, it was the second largest bacteria infecting virus and the largest *Klebsiella* phage published until that day. Furthermore, phage RaK2, based on the genome size, is the sixth largest phage to date (Yuan and Gao, 2017).

Another interesting feature of this bacteriophage – its morphology. TEM revealed that phage RaK2 is characterized by an isometric head of 123 nm in diameter and a contractile tail (128 nm in length and 21.5 nm in width in extended state and 42 nm in width in contracted state). The phage has a clearly defined neck (12.3 nm), baseplate (35 nm), collar (18.5 nm) and six baseplate-associated ramified long tail fibers (Fig. 2.1). These fibers are generously studded with spike-like structures, making it impossible to accurately determine the actual length of these structures. Similar structures have been observed only in phages from "viunalikeviruses" group (Adriaenssens et al., 2012). However, based on bioinformatics and phylogenetic analyzes, a close connection between tail fibers from RaK2 and these phages was not found.

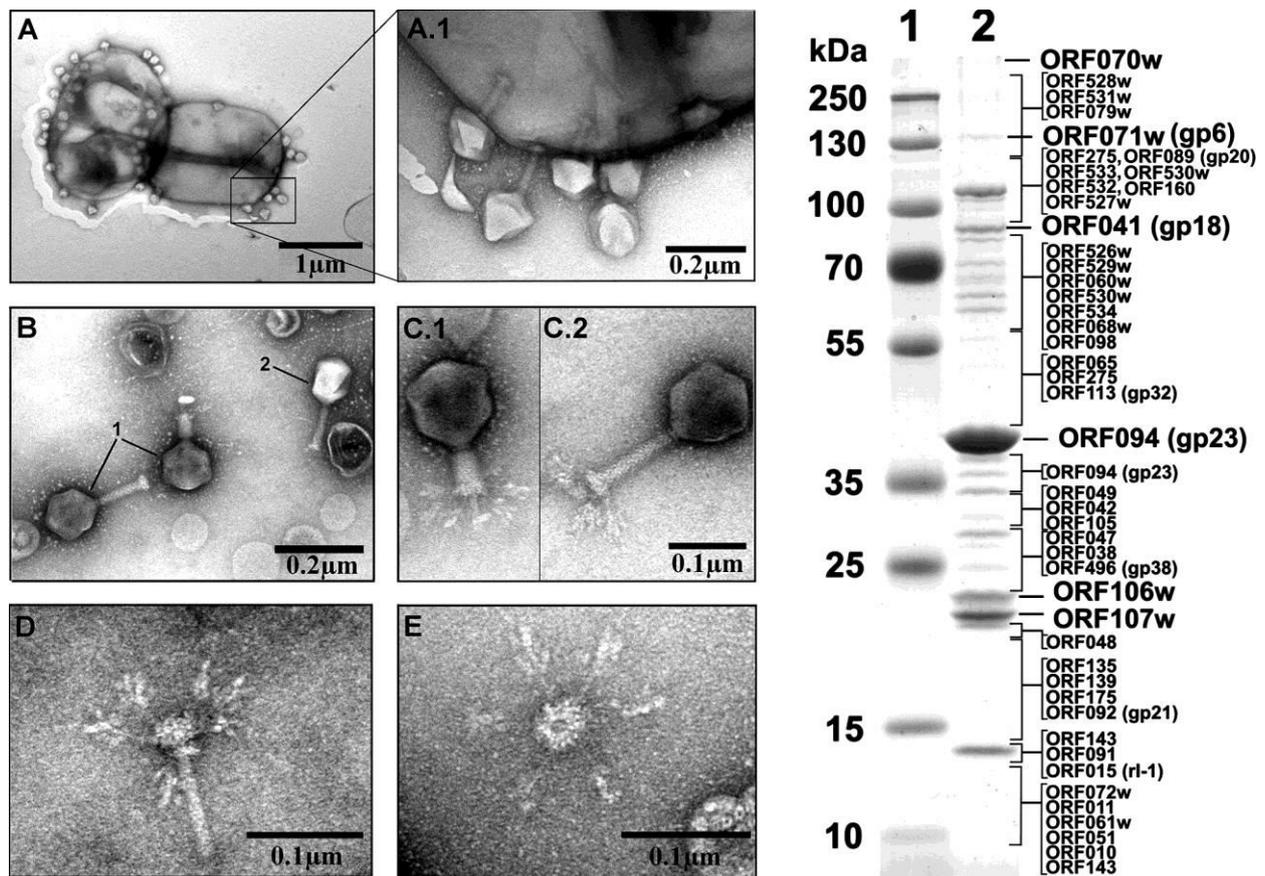


Fig. 2.1. TEM (left) and SDS-PAGE (right) analysis of *Klebsiella* phage RaK2 virion. (A and A.1) Phage RaK2 particles adsorbed to the surface of *Klebsiella* sp. KV-3 cells. (B) Purified phage RaK2 particles (1) and one particle of phage T4 (2). (C) RaK2 particle with contracted (C.1) and extended (C.2) tail. (D) Inner tail tube with baseplate and baseplate-associated ramified tail fiber structures. (E) Baseplate with six long tail fibers. Lanes: 1 – molecular mass marker Page Ruler™ prestained Protein Ladder Plus (Thermo Fisher Scientific), 2 – phage RaK2. Relative migrations of MW marker proteins are indicated on the left. Proteins identified by MS/MS are indicated on the right.

The host range determination test revealed that with the exception of *Klebsiella* sp. veterinary isolate KV-3, all of 40 bacterial strains tested in this study (Table 1.1) were resistant to this phage. The e.o.p. test unveil that the phage has an optimum temperature for plating around 30°C. Meanwhile the adsorption tests showed that a high percentage (93%) of the RaK2 particles adsorb to *Klebsiella* sp. veterinary isolate KV-3 cells after 10 min of incubation. The single step experiments revealed that latent period of RaK2 is 60 min, the eclipse period is 40 min and the phage generates an average burst of 140 virions per infected cell.

The genome sequence analysis revealed that the circularly permuted linear genome of RaK2 is 345,809 bp long with an overall G+C content of 31.7%. RaK2 has a total of 534 probable protein-encoding genes, 5 genes for tRNA and 2 pseudo-tRNA

genes. However, based on the similarity to biologically defined proteins, only 79 of RaK2 ORFs were given a functional annotation. Reversed-phase nano-liquid chromatography directly coupled with LC-MS/MS analysis of the structural Rak2 proteins separated by SDS PAGE led to the experimental identification of 54 virion proteins, including 19 that were predicted by bioinformatics approaches.

2.1.2 *Arthrobacter* bacteriophage vB_ArS-ArV2

Arthrobacter bacteriophage vB_ArS-ArV2 (abbreviated name ArV2) – first *Arthrobacter* sp. infecting bacteriophage with completely sequenced and published genome. The host range determination test revealed that with the exception of *Arthrobacter* sp. 68b, all of 40 bacterial strains tested in this study (Table 1.1) were resistant to this phage. The e.o.p. test revealed that the phage has an optimum temperature for plating around 30°C. Attempts to obtain a one-step growth curve of ArV2 were unsuccessful because of the slow adsorption kinetics: only about 50% of the PFU adsorb in 5 min, and after 15 min as many as 25% of the original PFU remain unattached.

TEM analysis revealed that phage ArV2 belongs to the family *Siphoviridae* and is characterized by an isometric head (diameter 62.86 nm) and an apparently non-contractile, flexible tail (194.46 nm in length and 11.86 nm in width) (Fig. 2.2). A baseplate was observed, although its diameter was not clearly distinguishable. Tail fibers are not obviously visible but upon closer inspection, six baseplate-associated short tail fibers (7.56 nm in length) can be seen in the vicinity of the tail tips (Fig. 2.2 D, E).

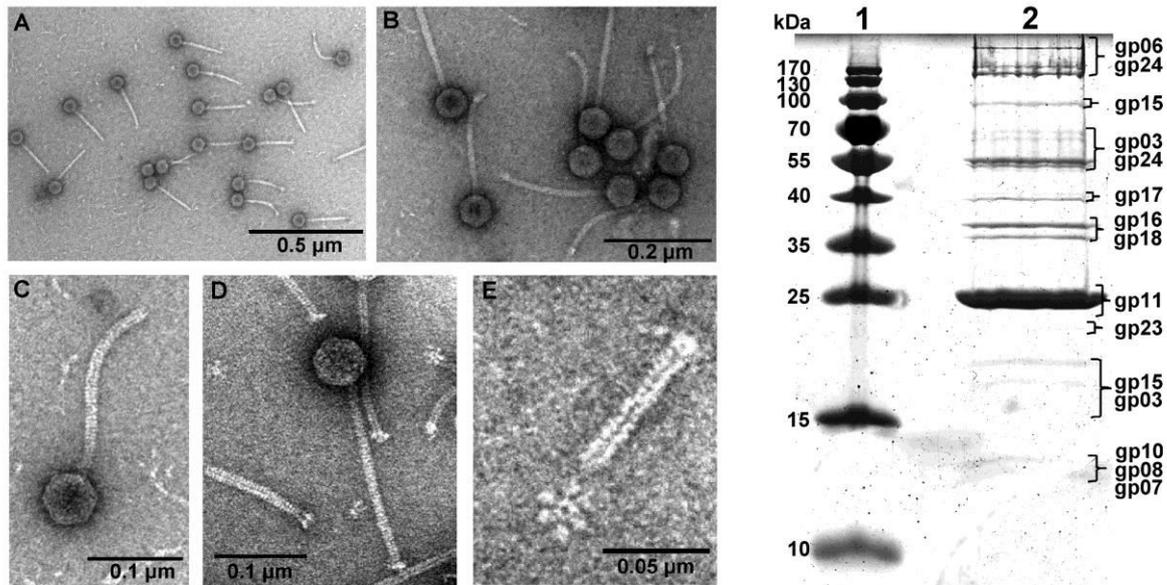


Fig. 2.2. TEM (left) and SDS-PAGE (right) analysis of *Arthrobacter* phage ArV2 virions. (A–D) ArV2 virions. (E) ArV2 tail tube and baseplate with six baseplate-associated short tail fibers. Lanes: 1 – molecular mass marker Page Ruler™ prestained Protein Ladder Plus (Thermo Fisher), 2 – phage ArV2. Relative migrations of MW marker proteins are indicated on the left. Proteins identified by LC-MS/MS are indicated on the right.

Phage ArV2 has a linear, double stranded DNA genome consisting of 37,372 bp with a G+C content of 62.73%. The genome sequence analysis revealed that ArV2 has a total of 68 probable protein-encoding genes and no genes for tRNA. The genome analysis revealed that 41% of the ArV2 genes (28 out of 68 ORFs) encode unique proteins that have no reliable identity to database entries. LC-MS/MS analysis of the structural ArV2 proteins separated by SDS PAGE and filter-aided protein sample preparation (FASP) led to the experimental identification of 14 virion proteins, including 9 that were predicted by bioinformatics approaches.

2.1.3 *Arthrobacter* bacteriophage vB_ArtM-ArV1

Arthrobacter bacteriophage vB_ArtM-ArV1 (abbreviated name ArV1) – first *Arthrobacter* spp. infecting phage from the family *Myoviridae* with completely sequenced and published genome. Interestingly, TEM analysis demonstrated that ArV1 is a member of the family *Myoviridae* but phylogenetic and comparative sequence analyzes, however, revealed that ArV1 has more genes in common with phages from the family *Siphoviridae* than it does with any myovirus characterized to date. Based on the

results of this study, it was demonstrated that ArV1 forms a discrete clade that seems to occupy a position somewhat intermediate between myo- and siphoviruses.

Host range determination test revealed that out of 51 bacterial strains tested (Table 1.1) phage ArV1 was capable of infecting 6 phylogenetically related *Arthrobacter* species. The e.o.p. test revealed that the phage has an optimum temperature for plating around 28°C. Bacteriophage ArV1 was failed to reproduce after inoculation into liquid bacterial culture; therefore, the one-step growth experiment was not performed.

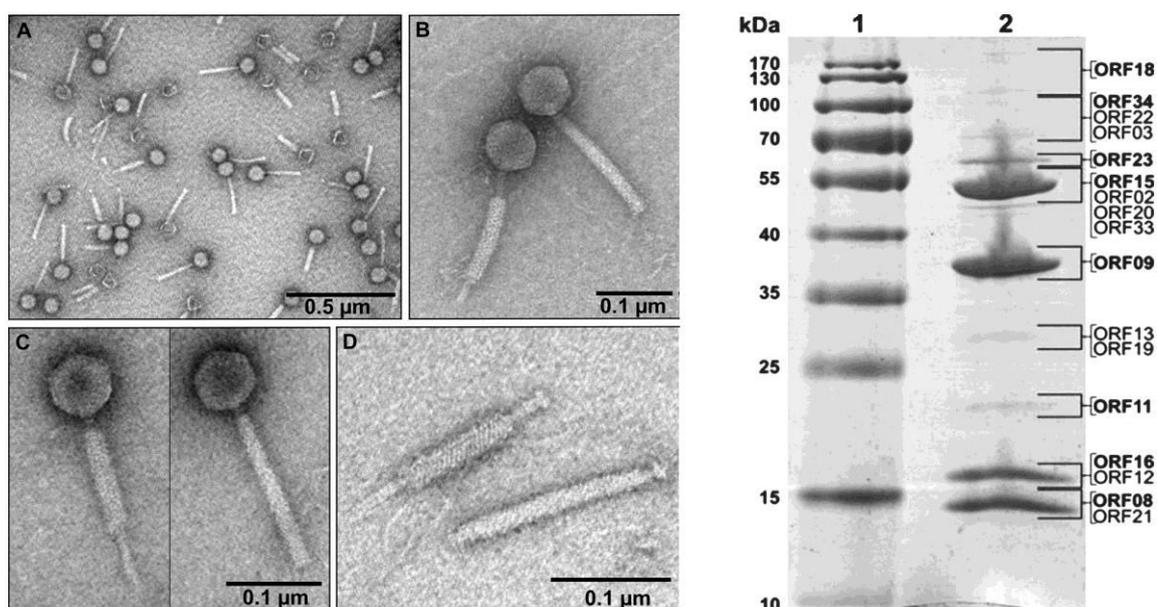


Fig. 2.3. TEM (left) and SDS-PAGE (right) analysis of *Arthrobacter* phage ArV1 virions. (A–C) ArV2 virions. (D) ArV1 tail with contracted (top) and extended (bottom) tail sheath. Lanes: 1 – molecular mass marker Page Ruler™ prestained Protein Ladder Plus (Thermo Fisher Scientific), 2 – phage ArV2. Relative migrations of MW marker proteins are indicated on the left. Proteins identified by LC-MS/MS are indicated on the right.

TEM analysis showed that ArV1 has an isometric head (~74 nm in diameter) and a contractile, nonflexible tail (~192 nm). The genome of ArV1 is a linear, circularly permuted, double-stranded DNA molecule (71,200 bp) with a G+C content of 61.6%. The genome sequence analysis revealed that ArV1 has a total of 101 probable protein-encoding genes and no genes for tRNA. More than 50% of ArV1 genes encode unique proteins that either have no reliable identity to database entries or have homologues only in *Arthrobacter* phages, both sipho- and myoviruses. Using bioinformatics approaches, 13 ArV1 structural genes were identified, including those coding for head, tail, tail fiber,

and baseplate proteins. A further 6 ArV1 ORFs were annotated as putative structural proteins based on the results of proteomic analysis.

2.1.4 *Escherichia coli* bacteriophage vB_EcoS_NBD2

Bacteriophage vB_EcoS_NBD2 (abbreviated name NBD2) – a member of the family *Siphoviridae*. TEM analysis showed that NBD2 has an isometric head (~65 nm in diameter) and noncontractile, flexible tail (~170 nm in length and ~12 nm in width) (Fig. 2.4).

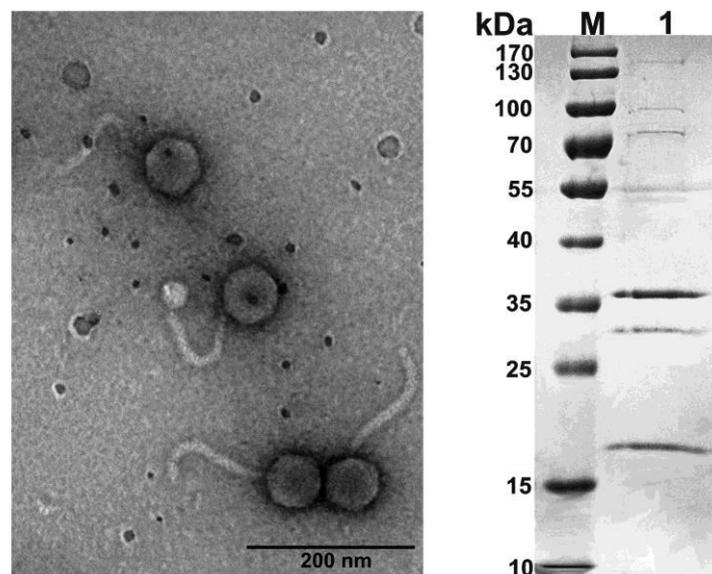


Fig. 2.4. TEM (left) and SDS-PAGE (right) analysis of *E. coli* phage NBD2 virions. Lanes: 1 – molecular mass marker Page Ruler™ prestained Protein Ladder Plus (Thermo Fisher Scientific), 2 – phage NBD2.

Although neither base plate nor tail fiber proteins were seen on electron micrographs of phage NBD2, based on the genome bioinformatics analysis (work carried out by dr. L. Kalinienė), genes encoding these structures (respectively ORF44, ORF45 and ORF47, ORF57) were identified. In total, 16 structural proteins encoding genes, including the already mentioned above, were predicted in genome of NBD2 based on the bioinformatics approaches. All of these structural proteins have been approved and following LC-MS/MS analysis. In addition, two more NBD2 structural proteins (gp26 and gp49), which have no reliable homology with annotated structural proteins, were identified during proteomic analysis.

2.1.4 *Escherichia coli* bacteriophage vB_EcoM-FV3

Bacteriophage vB_EcoM_FV3 is a member of "rV5-like viruses". The host-range determination analysis revealed that this phage infects *E. coli* K-12-derived laboratory strains and replicates at high temperature (up to 47°C). Electron micrographs showed that FV3 is morphologically similar to the members of the family *Myoviridae* and has an isometric head of 85 nm in apical diameter and a necked contractile tail of ~120 nm in length and ~18 nm in width in extended state and ~27 nm in width in contracted state. FV3 has six slightly kinked 60 nm long tail fibers connected to the baseplate (Fig. 2.5).

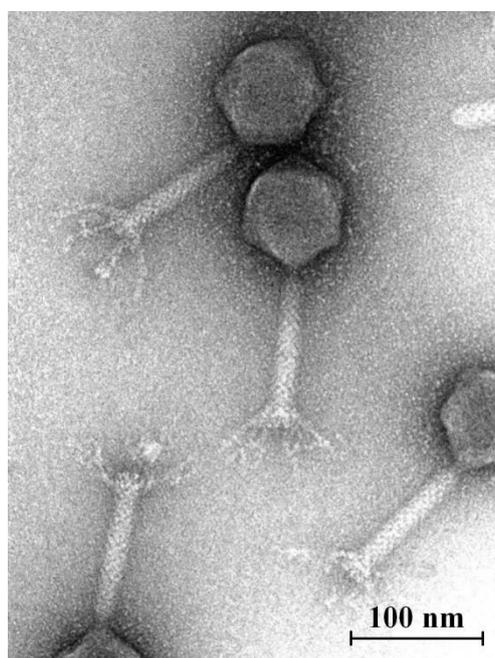


Fig. 2.5. Electron micrographs of *E. coli* phage FV3.

Genome of FV3 (136,947 bp) contains 218 open reading frames and encodes 5 genes for tRNA, genome G+C content is 43.7%. Bioinformatics analysis revealed that FV3 genome organization closely resembles that of phage rV5. In total, 212 out of 218 FV3 ORFs are homologous to the corresponding rV5 ORFs and have an average level of amino acid identity of 93.1%. Based on the bioinformatics approaches, 19 structural protein encoding genes were identified in genome of FV3.

2.2 Investigation of structural proteins of new bacteriophages

To determine which of the identified new structural proteins of phages are suitable for the construction of self-assembling nanostructures *in vivo* or *in vitro*, detailed

analysis of selected structural proteins was performed: a plasmid vectors containing structural proteins encoding genes were designed (see Table 2), gene expression in *E. coli* was carried out, finally, samples of soluble recombinant proteins were analyzed by TEM.

2.2.1 Investigation of structural proteins of phage RaK2

Five structural proteins of phage RaK2 were selected for more detailed analysis: tail sheath protein (gp041), tail tube protein (gp042) and three hypothetical proteins gp043, gp106 and gp107. The last two proteins were found to be one the most abundant proteins of RaK2 virion (Fig. 2.1) and were thought to be associated with a tail of RaK2. However, during this study it was demonstrated that these proteins are components of RaK2 head. SDS-PAGE analysis demonstrated that recombinant gp106 and gp107 are soluble proteins, but, according to the results of TEM, did not form the ordered self-assembling structures. Nanostructures were not detected in the case of recombinant gp042 and gp043, too. Moreover, SDS-PAGE analysis demonstrated that these proteins were insoluble under investigated conditions.

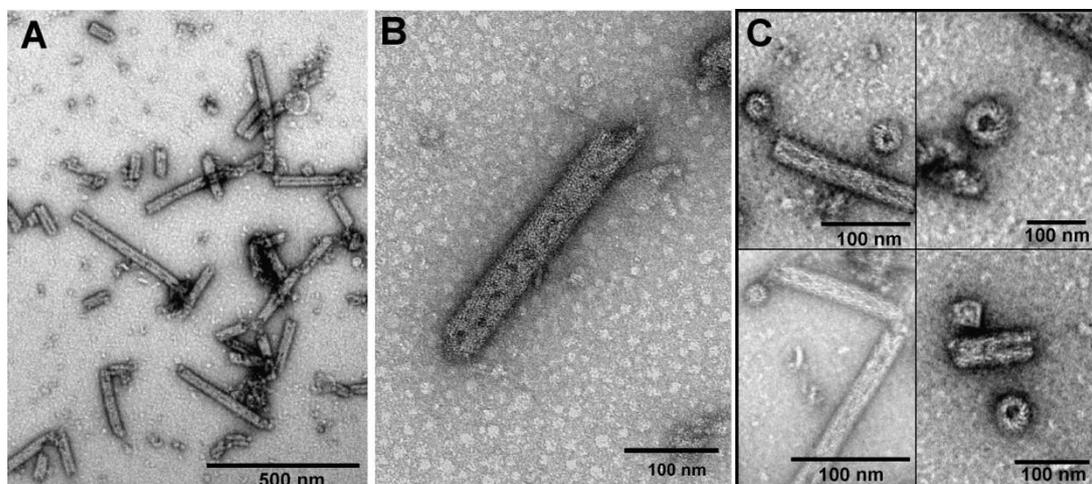


Fig. 2.6. Electron micrographs of nanostructures formed by recombinant gp041 of phage RaK2. (A, B) Regular tubular nanostructures formed by recombinant gp041. **(C)** Individual rings formed by recombinant gp041.

On the other hand, it was demonstrated that recombinant protein encoded by gene 041, which was cloned in plasmid vector pET21d and expressed in *E. coli* BL21-DE3 cells (work was carried out by dr. S. Povilonienė), was soluble and *in vivo*, in the absence of any other phage proteins, self-assembles into regular tubular structures. TEM

analysis revealed that the length of these structures varied from 10 nm to 750 nm, meanwhile the diameter of these structures (~41 nm) corresponds to the width of a contracted tail sheath (~42 nm) of the phage RaK2 (Fig. 2.6).

2.2.2 Investigation of structural proteins of phages ArV1 and ArV2

Ten structural proteins of phage ArV2 were selected for more detailed analysis: gp03 (portal protein), gp05 (capsid maturation protein), gp06 (major capsid protein), gp11 (major tail protein), gp15 (tape measure protein), gp16, gp17, gp18 (tail proteins), gp20 and gp24 (tail fiber proteins). Further, tail sheath protein (gp15) of phage ArV1 also was chosen for investigation.

Based on the SDS-PAGE analysis, most of studied proteins were insoluble in investigated conditions. Only gp05, gp18 (expressed in *E. coli* BL21-DE3) and gp24 (expressed in *E. coli* Rosetta) from ArV2 as well as gp15 (expressed in *E. coli* Rosetta) from ArV1 were soluble. Unfortunately, TEM analysis revealed that none of these proteins were capable to polymerase into regular structures.

2.2.3 Investigation of structural proteins of phage NBD2

Tail tube protein (or major tail protein) was the only one structural protein of siphovirus NBD2 selected for more detailed studies. It was demonstrated that this protein, encoded by gene 39, which was cloned in plasmid vector pET21a and expressed in *E. coli* BL21-DE3 cells, was soluble and *in vivo*, in the absence of any other phage proteins, self-assembles into regular tubular structures. Based on TEM analysis, tubular structures, formed by recombinant gp39, were flexible and extremely long (up to 2–3 μm in length). The diameter of these structures (~12 nm) corresponds to the width of a tail of phage NBD2 (Fig. 2.7).

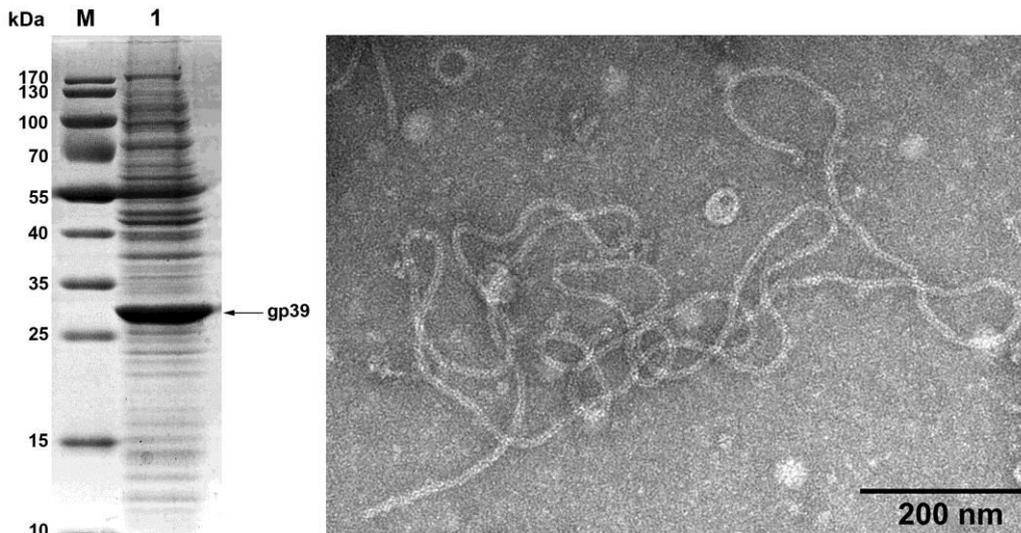


Fig. 2.7. SDS-PAGE (left) and TEM (right) analysis of recombinant gp39, expressed in *E. coli* BL21-DE3 cells. Protein expression was induced by IPTG (0.1 mM final concentration), the cells were incubated for 16 h at 20°C. Lanes: 1 – molecular mass marker Page Ruler™ prestained Protein Ladder Plus (Thermo Fisher Scientific), 2 – soluble protein fraction.

According to the literature, the tail tip complex (TTC) is necessary for initiation of the tail tube protein polymerization (Katsura and Tsugita, 1977; Katsura, 1983; Davidson, 2012). Only recently, it was demonstrated that the recombinant tail tube protein (gp17.1) from siphovirus SPP1 polymerized into tubular structures in the absence of other phage proteins (Langlois et al., 2015). A self-assembly process of gp17.1 occurred *in vitro* and required a prolonged incubation (up to 10 days). Furthermore, according to the published data, the length of nanostructures, formed by gp17.1 was significantly shorter than in the case of nanostructures formed by recombinant gp39 from NBD2. For this reason, the gp39 studies are of particular interest not only because of the possibility to construct the hybrid nanostructures, but, in general, for elucidation of the polymerization process of a bacteriophage tail tube assembly, which is not fully understood.

2.2.4 Investigation of structural proteins of phage FV3

Two structural proteins of phage FV3 were selected for more detailed analysis: tail tube protein (gp052) and tail sheath protein (gp053). Plasmid vectors containing genes encoding these structural proteins were constructed and gene expression in *E. coli* BL21-DE3 cells was analyzed (dr. L. Truncaitè, unpublished). Based on the SDS-PAGE analysis, the recombinant gp052 and gp053 were soluble, but only one of these proteins

(gp053) self-assembled *in vivo* into the regular tubular structures in the absence of any other phage proteins.

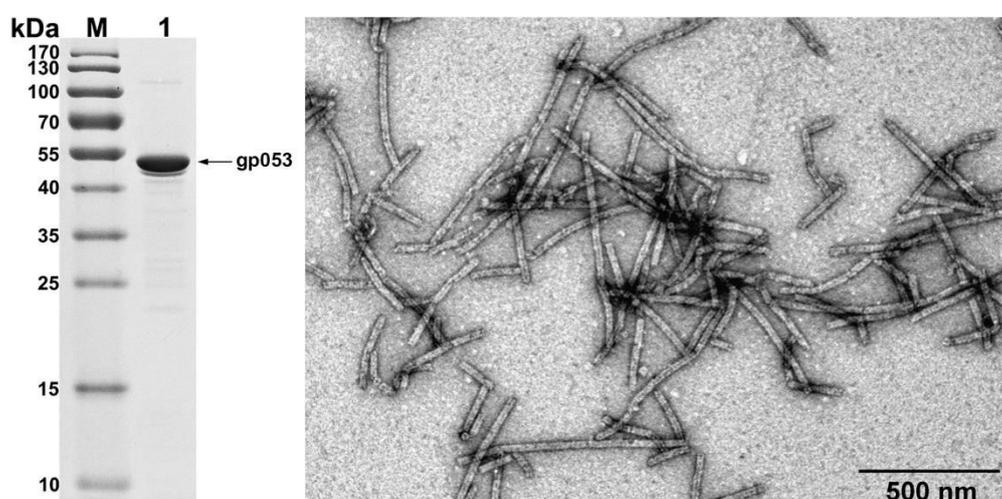


Fig. 2.8. SDS-PAGE (left) and TEM (right) analysis of recombinant gp053, expressed in *E. coli* BL21-DE3 cells. Protein expression was induced by IPTG (0.1 mM final concentration), the cells were incubated for 3 h at 30°C. Lanes: 1 – molecular mass marker Page Ruler™ prestained Protein Ladder Plus (Thermo Fisher Scientific), 2 – purified gp053.

TEM analysis revealed that the diameter of nanostructures (27.13 ± 2.69 nm) formed by recombinant gp053 corresponded to the width of a contracted tail (26.79 ± 1.78 nm) of the phage FV3. The length of nanotubes varied from separate rings (inner surface of 11.51 ± 0.78 nm, outer surface of 27.83 ± 2.68 nm) to nanotubes almost 1000 nm long (Fig. 2.8).

2.3 Investigation of nanostructures formed by gp053 of phage FV3

2.3.1 Bioinformatics analysis of gp053

According to bioinformatics analysis, gp053 has the highest identity (98–99%) to tail sheath proteins of six *Escherichia* spp. infecting phages from the group of "rV5-like viruses". Despite this, it is still an open question of molecular architecture of gp053 of FV3. To our knowledge, none of crystal structures of the tail sheath proteins from aforementioned bacteriophages have been identified to date. On the other hand, it was observed that, despite low similarity in amino acid level, the tail sheath proteins of different *Myoviridae* phages appear to have similar helical parameters and function in a

similar manner (Aksyuk et al. 2011; Leiman and Shneider, 2012; Fokine and Rossman, 2014).

Search for the gp053 fold was done using the HHpred server. It was shown that only C-terminal fragment of gp053 (residues 258 to 442) was predicted to adopt the fold of the C-terminal fragment (residues 470 to 643) of the tail sheath protein gp18 of phage T4 (PDB ref 3J2M) with a probability of 97.2 (E-value= 0.00034). On the other hand, HHpred analysis revealed that predicted fold of FV3 gp053 was more similar to three other contractile phage-like structures. It was shown that gp053 residues 1 to 442 were predicted to adopt the fold of the R-type pyocin from *Pseudomonas aeruginosa* (PDB ref 3J9Q) with a probability of 98.6 (E-value=9.9e-07). Meanwhile, residues 18 to 442 were aligned to the tail sheath protein encoded by gene *lin1278* from prophage infecting *Listeria innocua* (PDB ref 3LML) with a probability of 98.5 (E-value=1.3e-05). Similarly, residues 16 to 442 were aligned to the tail sheath protein encoded by gene *dasy3957* from prophage infecting *Desulfitobacterium hafniense* (PDB ref 3HXL) with a probability of 98.5 (E-value=6.3E-06).

The results of HHpred analysis are not unexpected due to two reasons. Firstly, a limited number of crystal structures of the tail sheath proteins from different phages have been solved. Secondly, it has been shown that the contractile structures from phages have a very similar structure to the contractile molecular machines found in many prokaryotes: R-type pyocins, the Type VI secretion system (T6SS) and phage tail-like protein translocation structures (PLTS) (Leiman et al. 2009; Sarris et al., 2014; Kube, 2015). Therefore, it is unsurprising that a predicted fold of gp053 from FV3 is more similar to R-type pyocin from *P. aeruginosa* than to the tail sheath protein from phage T4.

2.3.2 Expression, purification and stability studies of gp053

Expression of gp053 was carried out in *E. coli* strain BL21-DE3, protein expression was induced with 0.1 mM IPTG, cell culture was incubated at 30°C for 3 hours or overnight. Purification of gp053 was performed using several methods. Firstly, N- and C-terminus his-tagged gp053 were constructed. Proteins were purified using His-Spin Protein Miniprep kit (Zymo Research) according to manufacturer's

recommendations. Based on the results of SDS-PAGE and TEM analysis, gp053 purification by affinity chromatography was ineffective (Fig. 2.9).

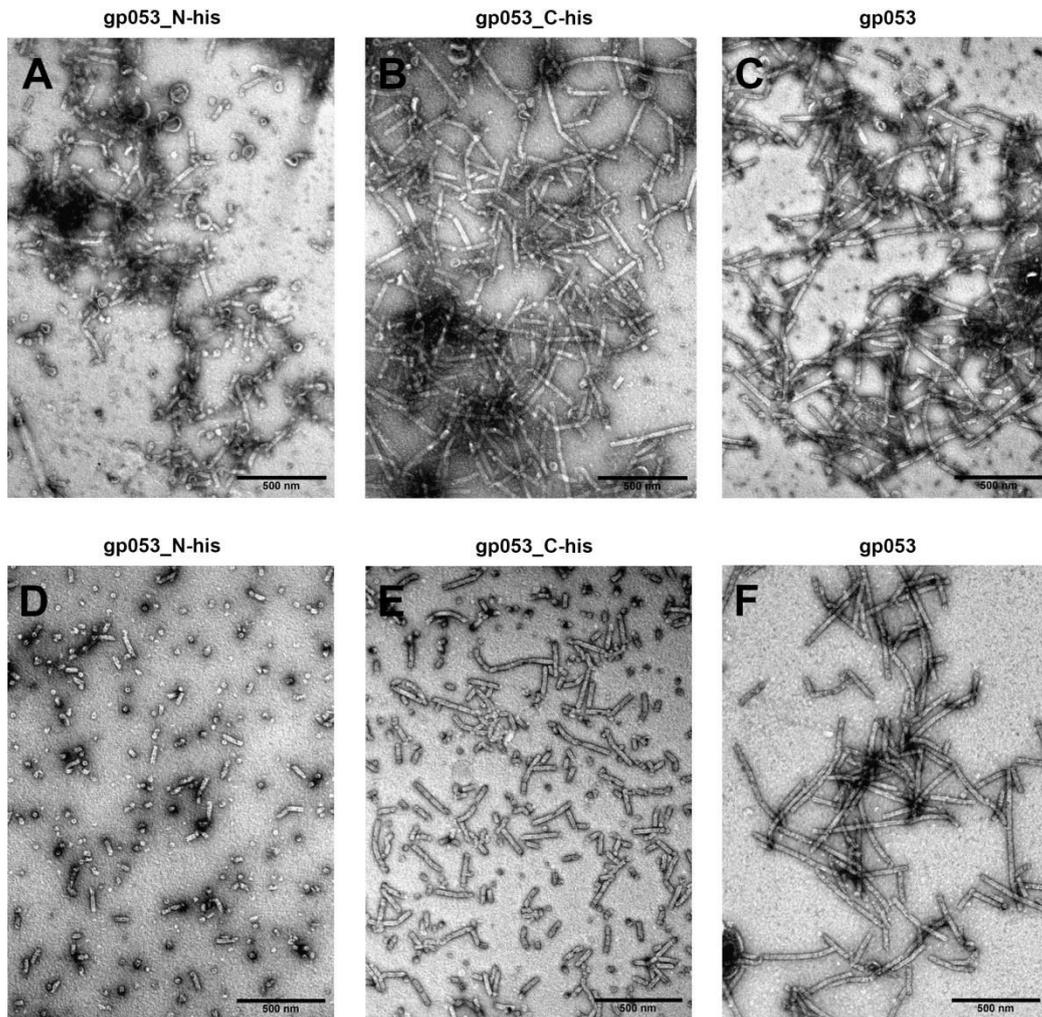


Fig. 2.9. TEM analysis of recombinant gp053 before (above) and after (below) purification. (D) gp053_N-his and (E) gp053_C-his purified using His-Spin Protein Miniprep kit (Zymo Research). (F) gp053 purified using ammonium sulfate (10% final concentration).

This could be due to several reasons. Firstly, after cell sonication, most of the gp053 is found in a state of comparatively long polysheaths, which could be a limiting factor for usual protein purification procedures using affinity chromatography. Secondly, based on the conserved model of the tail sheath structure of phage T4, N- and C-terminus of the tail sheath protein (gp18) are positioned further towards the interior of the tail sheath (Aksyuk et al., 2009; Aksyuk et al., 2011; Leiman and Shneider, 2012). Therefore, considering the general conservation of structural proteins in bacteriophages and given results of our study, it seems that the gp053 most likely have similar helical parameters as gp18 from T4 and its N- and C-termini are positioned in the internal part of the sheath.

Alternatively, gp053 polysheaths were purified using ammonium sulfate. It was demonstrated that precipitation of gp053 nanostructures from the supernatant by an addition of ammonium sulfate (final concentration 10%) was very effective, as the purification could be carried out in just half an hour, the method was cheap (only ammonium sulfate was used) and productive (yield of 1–5 mg/ml). The SDS-PAGE analysis of the purified gp053 is shown in Fig. 2.8, the results of TEM analysis are presented in Figs. 2.8 and 2.9 F.

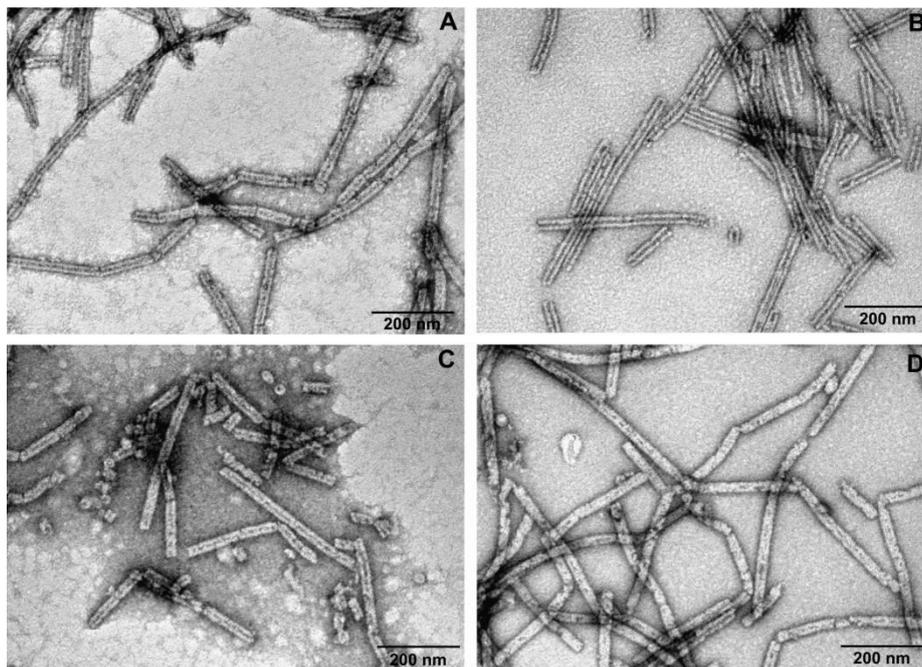


Fig. 2.10. TEM analysis of recombinant gp053 polysheaths stability. Samples incubated in: (A) TE buffer at 4°C for 12 months; (B) 16 h with trypsin (0.02 mg/ml); (C) 16 h with 8 M urea; (D) 30 min. in boiling water.

According to the published data, the polysheaths formed from the tail sheath protein of phage T4 were found to be very resistant to various physical and chemical factors (Arisaka et al., 1981; Arisaka et al., 1990). This extraordinary stability of polysheaths as well as contracted sheaths is thought to be associated to the very low-energy configuration of these structures, which is formed during the process of tail sheath contraction (Aksyuk et al., 2009; Leiman and Shneider, 2012). This study demonstrated that polysheaths formed by gp053 were also very stable structures. Nanotubes remain unchanged after a prolonged storage (more than one year) in TE buffer at 4°C, 16 hours of incubation in 8M urea, 16 hours of incubation with trypsin (final concentration of 0.02 mg/ml) and even boiling for 30 min. (Fig. 2.10).

2.3.3 Investigation of gp053 deletion mutants

In this study it was constructed a set of the gp053 mutants truncated at the N- and C-terminus, in order to understand the polymerization properties of gp053 and to obtain truncated recombinant proteins that could form a regular stable tubular structures. These recombinant proteins corresponded to full-length gp053 from 98.0% (gp053_N Δ 9) to 43.7% (gp053_C Δ 200) (Fig. 2.11).

TEM study of protein samples revealed that a deletion of 9, 20 and 25 amino acids from the N-terminal region did not disturb a protein assembly into the polysheaths. Meanwhile gp053_N Δ 29 mutant protein (deletion of 29 amino acids from the N-terminal region) was found in a soluble fraction but folded to unordered protein ribbons without observable formation of the typical polysheaths. The polysheaths were not also detected in the case of gp053_N Δ 25_C-strep. Moreover, the recombinant protein was insoluble. Thus, it was demonstrated that polymerization properties of gp053 were determined not only by the number of truncated amino acids from N-terminus, but by amino acids located on the C-terminus as well. Mutant proteins harbouring deletions of 11, 31, 51, 76 and even 100 amino acids from the C-terminus formed the ordered tubular polysheaths and only deletion of 152 C-terminal amino acids (gp053_C Δ 152 mutant) resulted in a formation of soluble, but not self-assembled proteins. These results were in accord with a previous observation that deletions of residues from the C-terminus of the tail sheath protein (gp18) from phage T4 reduced its polymerization ability much less than elimination of residues from the N-terminus (Kuznetsova et al., 1998; Poglazov et al., 1999).

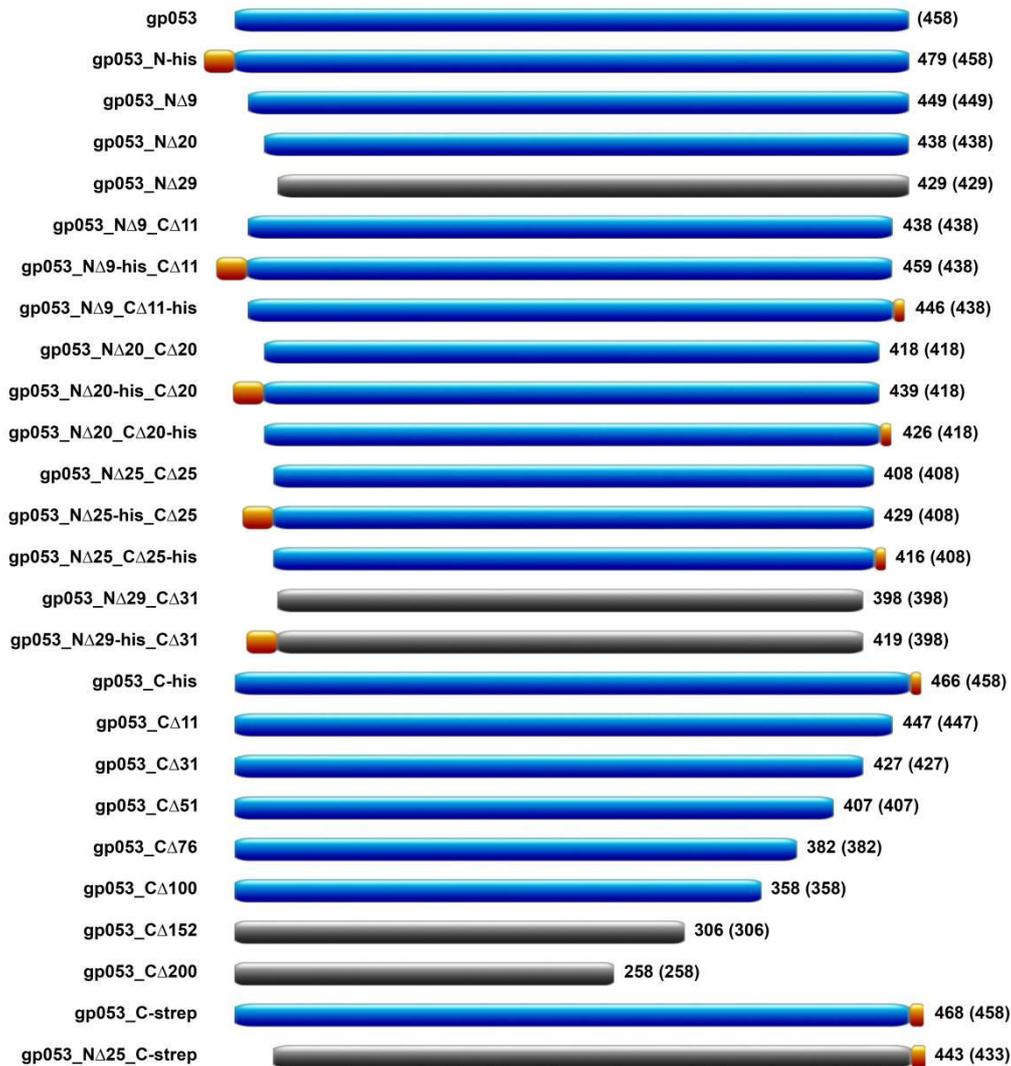


Fig. 2.11. Analysis of polymerization properties of gp053 deletion mutants. The mutants that can polymerize (including the full-length protein) are shown in blue, those that do not polymerize are shown in black, his-tags and strep-tags are shown in orange. All mutants entitled according to N-terminal or C-terminal mutations and distribution of his-tag or strep-tag (shown on the left). The amount of amino acids found in mutants are shown on the right, in brackets – the number of unchanged amino acids from native gp053.

According to the literature, a number of T4 gp18 mutants assemble into thinner filaments called "noncontracted polysheaths" (NCP) (Poglazov et al, 1999). Meanwhile TEM analysis of polysheaths, formed by the majority of gp053 deletion mutants, revealed that the structures itself visually had no significant differences in comparison to polysheaths formed by recombinant gp053 (Fig. 2.12). Contrarily, gp053_CΔ100 maintained the ability to assemble into even longer (~2000 nm) and slightly larger in diameter (~32.47 nm) tubular structures than those of a gp053 (Fig.2.12 I).

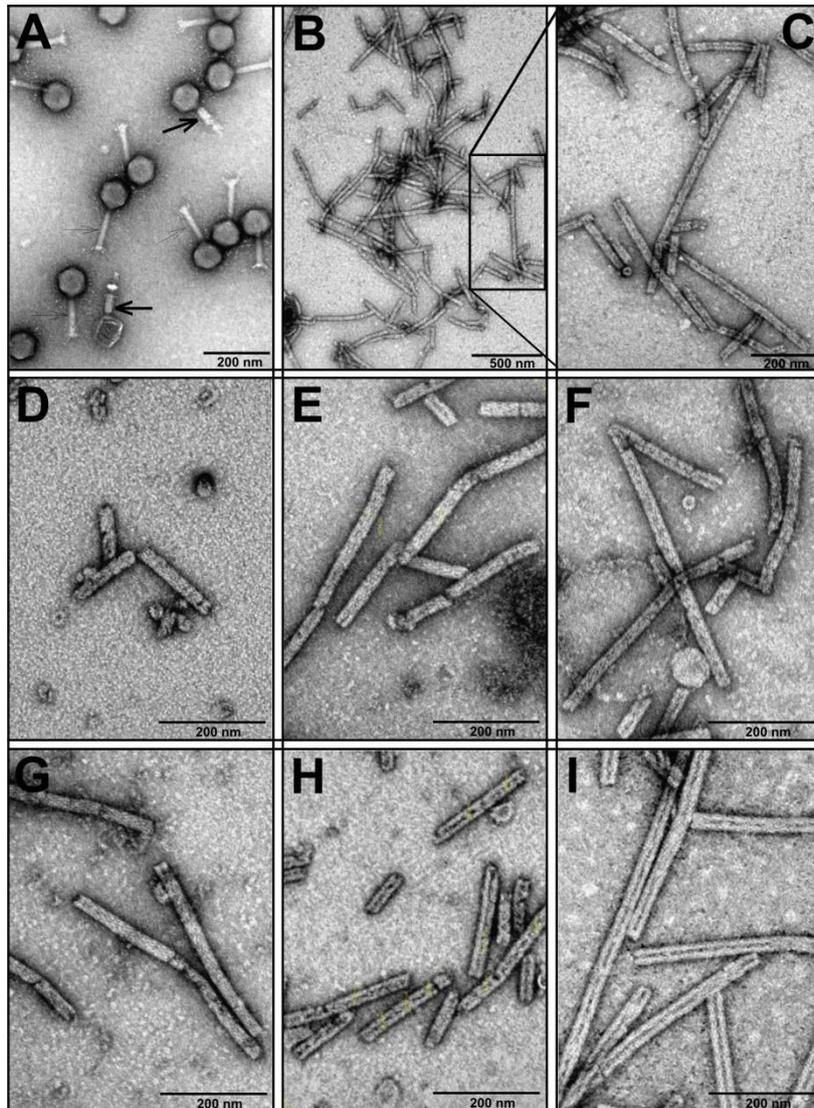


Fig. 2.12. TEM analysis of phage FV3 particles and polysheaths formed by gp053 and its mutants. (A) phage FV3 virions (arrows show the contracted tail sheaths); (B–I) nanostructures formed by recombinant gp053 and its mutants: (B,C) gp053; (D) gp053_N-his; (E) gp053_C-his; (F) gp053_NΔ9-his_CΔ11; (G) gp053_NΔ25-his_CΔ25; (H) gp053_NΔ20; (I) gp053_CΔ100.

In addition, it was observed that some of the tubular structures as formed, for example, by gp053_CΔ75 (Fig. 2.16 C), gp053_NΔ20 and gp053_CΔ100 (Fig. 2.12 H and I, respectively) as well as by full-length gp053 after long storage in TE buffer or incubation with trypsin (Fig. 2.10 A and B, respectively) were characterized by a better defined internal channel than those of a recombinant gp053 (Fig. 2.12 C). Similarly, polysheaths with more visible internal channel in the case of gp029 from *Pseudomonas aeruginosa* phage phiKZ were also observed. Probably, these structures are caused by less compact packing of protein subunits in the polysheaths that facilitates stain penetration into the internal channel (Kurochkina et al, 2009).

2.3.4 Investigation of gp053 mutants with inserted cloning sites

In this study it was constructed a set of the gp053 mutants with cloning sites inserted at locations encoding N- and C-terminus that were situated in the inner part of the polysheath as well as in the middle of gene changing the outer surface of the polysheath (Fig. 2.13).

It was determined that additional five amino acids in the N- and C-terminus of gp053 (mutant m_N_C) and gp053_C-his (mutant m_N_C-his) resulted in a formation of the shorter (up to 500 nm) nanostructures (Fig. 2.14 A, B). On the other hand, an addition of the same five amino acids in the N- and C-terminus of the deletion mutants gp053_N Δ 10_C Δ 11 (mutant m_N_C) and gp053_C-his (mutant m_N_C-his) led to formation of nanostructures up to 800 nm and 600 nm, respectively (Fig. 2.14 C and D). Even longer polysheaths (up to 1000 nm, Fig. 2.14 E) were observed after insertion of aforementioned amino acid in the N- and C-terminus of gp053_N Δ 20_C Δ 21-his (mutant m_N Δ 20_C Δ 21-his). The same tendency (more amino acids are removed – the more stable structures are formed) was observed in mutants with the longer insertions. Regular but short (up to 250 nm only) nanostructures were observed after addition of 34 amino acids in the C-terminus of gp053 (mutant m_C) (Fig. 2.14 K), meanwhile an insertion of the same amino acids in the C-terminus of gp053_C Δ 81 (m_C_C Δ 81) led to formation of nanotubes up to 500 nm long (Fig. 2.14 J). On the other hand, it was found that insertion of 13 amino acids in gp053_C Δ 100 (mutant m_C Δ 100) induced a formation of shorter nanostructures, up to 300 nm long only (Fig. 2.14 J), meanwhile the length of the structures formed by gp053_C Δ 100, as it was mentioned before, reached 2000 nm.

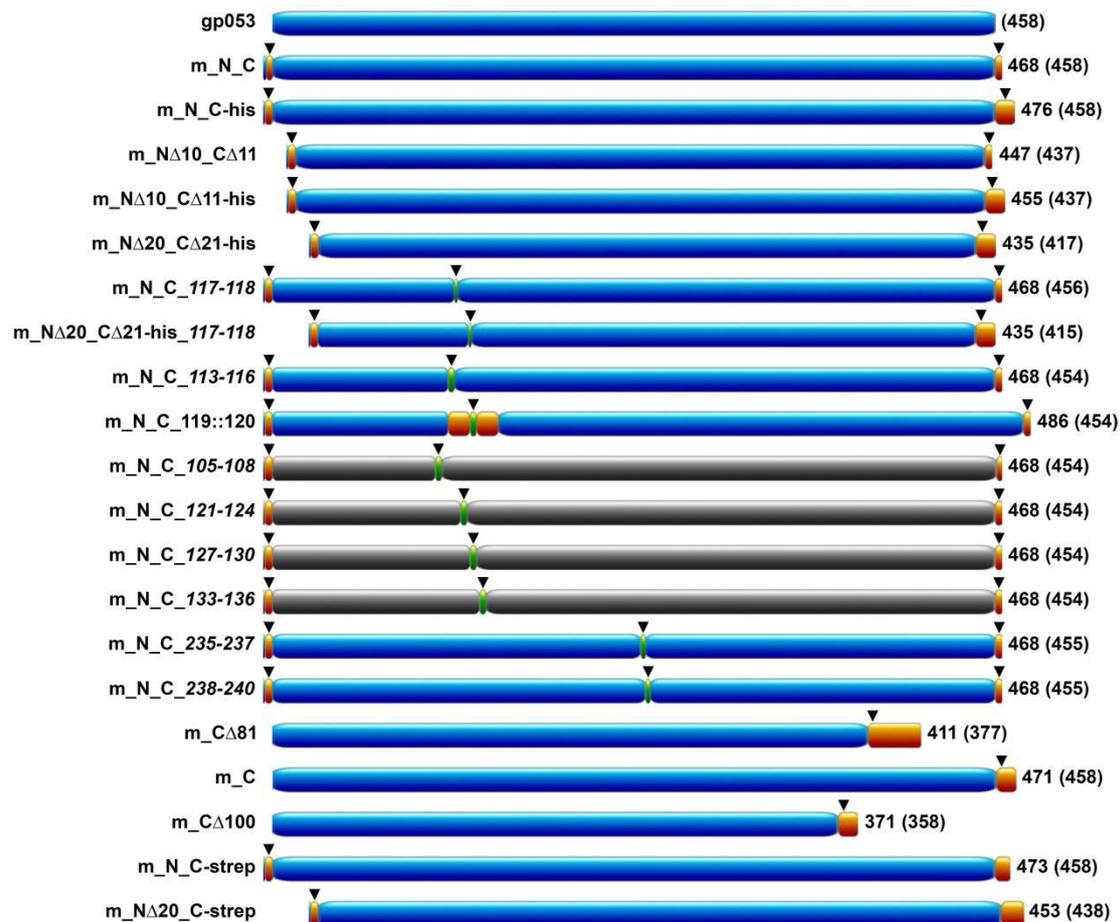


Fig. 2.13. Analysis of polymerization properties of gp053 mutants with inserted cloning sites. The mutants that can polymerize (including the full-length protein) are shown in blue, those that do not polymerize are shown in black, inserted amino acids with targeted cloning sites are shown in orange (black inverted triangles represents cloning site position). Symbols in entitled mutants (shown on the left) mean: gp053 mutant; N – cloning site inserted in the N-terminus; C – cloning site inserted in the C-terminus; Δ – amino acid deletion; $::$ – position of inserted amino acids (in the sequence of native protein); numbers in italics represents the position of amino acids of native protein, which were substituted with unnatural amino acids containing cloning sites. The amount of amino acids found in mutants are shown on the right, in brackets – the number of unchanged amino acids from native gp053.

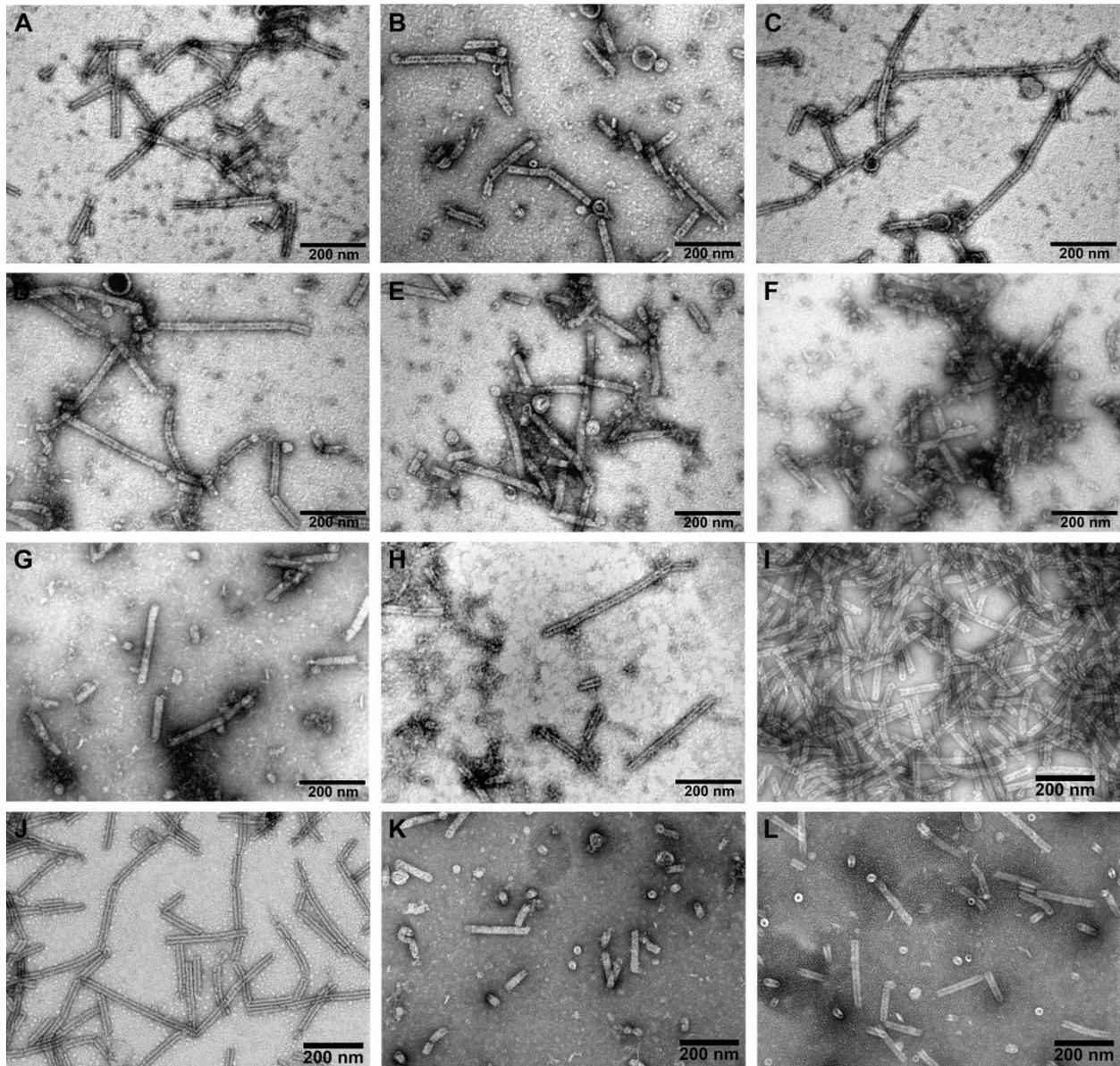


Fig.2.14. TEM analysis of polysheaths formed by gp053 mutants with inserted targeted cloning sites. (A) m_N_C; (B) m_N_C-his; (C) m_N Δ 10_C Δ 11; (D) m_N Δ 10_C Δ 11-his; (E) m_N Δ 20_C Δ 21-his; (F) m_N_C_117-118; (G) m_N Δ 20_C Δ 21-his_117-118; (H) m_N_C_113-116; (I) m_N_C_119::120; (J) m_C Δ 81; (K) m_C; (L) m_C Δ 100.

An even greater adverse impact on gp053 polymerization was observed after amino acids insertions or substitutions in the middle of the protein. Four recombinant mutated gp053 proteins (m_N_C_105-108, m_N_C_121-124, m_N_C_127-130 and m_N_C_133-136) with replaced four amino acids (positions indicated in the names of mutants) were insoluble (Fig. 2.13). Mutant proteins m_N_C_117-118, m_N Δ 20_C Δ 21-his_117-118 and m_N_C_113-116 formed regular, but short nanostructures up to 200 nm, 250 nm and 400 nm, respectively (Fig. 2.14 F, G, H). Polysheaths up to 400 nm were observed in the case of the mutant protein m_N_C_119::120 with inserted

additional 18 amino acids in position between 119 and 120 amino acids. Aforementioned mutants with the inserted cloning sites that formed the most stable polysheaths were selected for construction of the hybrid proteins.

2.3.5 Investigation of gp053-based hybrid proteins

In this study, gp053-based hybrid proteins were constructed by inserting Ag+ or/and TiO₂ binding peptides into cloning sites of gp053 mutants described previously (Fig. 2.15). Similarly to the results of studies of gp053 mutants with inserted cloning sites, it was demonstrated that mutants with inserted peptides in the N-terminus (m_NTi_C, m_NAgE6_C) formed less regular, shorter (50–150 nm) structures than mutants with inserted the same peptides in the C-terminus (m_N_CTi, m_N_CAgE6). On the other hand, insertions of peptides in truncated gp053 mutants (m_NΔ10AgE61_CΔ11, m_NΔ10_CΔ11_AgE6 and m_NΔ20AgE6_CΔ21Ti-his) resulted in formation of long (600–1000 nm), regular nanostructures (Fig. 2.16 B).



Fig. 2.15. Analysis of polymerization properties of gp053-based hybrid proteins. The mutants that can polymerize (including the full-length protein) are shown in blue, those that do not polymerize are shown in black, inserted amino acids with targeted cloning sites are shown in orange. Peptides inserted in cloning sites are shown in grey. Symbols in entitled mutants (shown on the left) mean: gp053 mutant; N – cloning site inserted in the N-terminus; C – cloning site inserted in the C-terminus; Δ – amino acid deletion; :: – position of inserted amino acids (in the sequence of native protein); numbers in italics represents the position of amino acids of native protein, which were substituted with unnatural amino acids containing cloning sites. The amount of amino acids found in mutants are shown on the right, in brackets – the number of unchanged amino acids from native gp053.

Hybrids m_CΔ81_AgD6, m_CΔ81_AgE6 and m_CΔ81_AgE4 formed rather short (up to 250–300 nm) but regular polysheaths (Fig.2.16 D, E, F). Insertion of peptides in cloning site of m_N_C_119::120 had especially negative impact on gp053 polymerization: structures formed by hybrids m_N_C_119::120_AgD6 and m_N_C_119::120_AgE6 were least regular, in most cases the length of these structures varied from 50 to 150 nm. In the case of m_N_C_119::120_AgE4 more regular, up to 200 nm structures were observed (Fig. 2.16 H, I).

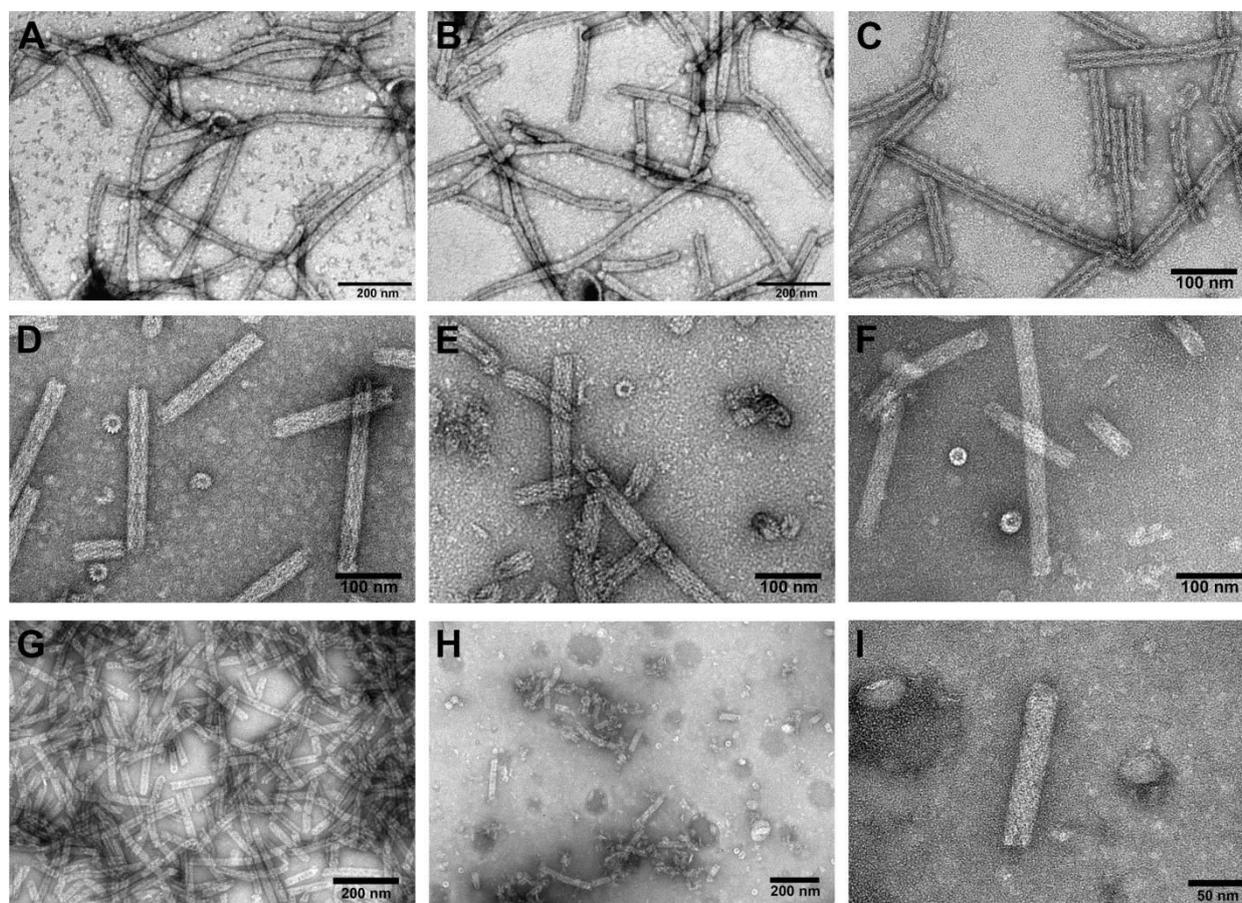


Fig. 2.16. TEM analysis of polysheaths formed by gp053 mutants with inserted targeted cloning sites and gp053-based hybrid proteins. (A) m_NΔ20_CΔ21-his; (B) m_NΔ20_AgE6_CΔ21Ti-his; (C) m_CΔ81; (D) m_CΔ81_AgD6; (E) m_CΔ81_AgE6; (F) m_CΔ81_AgE4; (G) m_N_C_119::120; (H-I) m_N_C_119::120_AgE6.

Based on the literature data, peptides D6, E6, or E4 exhibited in a variety of biological scaffolds (for example, on the surface of bacteriophage M13 or yeast cells) cause photoreduction of Ag^+ ions in the presence of ambient light. As a result, the solution color changes from colorless to red, morphologically noticeable changes on surfaces with exhibited peptides occurs (Nam et al., 2008). In this study, it was

demonstrated that color changes in samples of gp053-based hybrids (m_C Δ 81_AgD6, m_C Δ 81_AgE6, m_C Δ 81_AgE4 and especially in m_N Δ 20_AgE6_C Δ 21Ti-his and m_N_C_119::120_AgE4) were more extensive than in samples of gp053 mutants with cloning sites (m_C Δ 81, m_N Δ 20_C Δ 21-his and m_N_C_119::120). Unfortunately, TEM analysis revealed no morphological changes of the polysheaths before and after incubation with AgNO₃ (final AgNO₃ concentration 1 mM, reaction carried out in HEPES buffer (10 mM, pH 7,5), reaction time – 24 hours).

2.3.6 Hybrids of gp053 polysheaths and gold nanoparticles

The hybrid polysheaths of gp053 covered with gold nanoparticles were constructed (Fig. 2.17). gp053 was modified with biotin and incubated with neutravidin-conjugated gold nanoparticles (work was carried in collaboration with dr. S. Povilonienė). TEM analysis revealed that the ends of these polysheaths were covered with gold nanoparticles (Fig. 2.17).

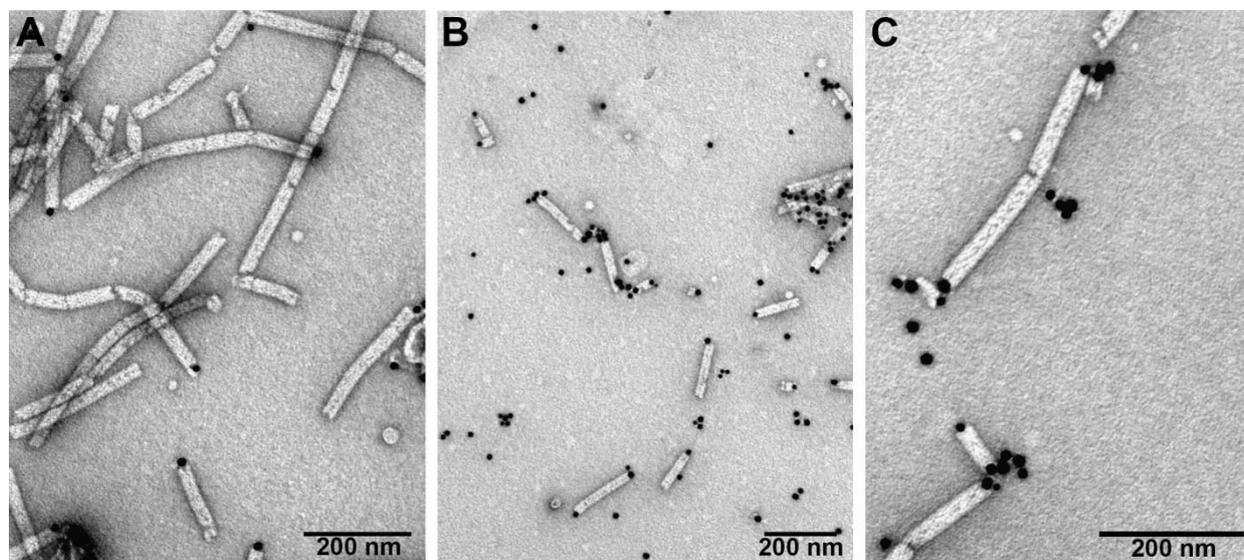


Fig. 2.17. TEM analysis of recombinant gp053 modified with biotin and incubated with neutravidin conjugated gold nanoparticles.

As can be seen in Fig. 2.17, the neutravidin-conjugated gold nanoparticles in most of the cases were observed attached to the ends or "break points" of the polysheaths, single rings or short gp053 aggregates, but not to the rest of the outer surface of the regular long polytubes. Thus, these studies not only allowed construction of the self-assembling hybrid structures, but also confirmed that four cysteine residues (positions of

160, 198, 208 and 320) are located in the inner part of the sheath and can be accessible at the termini of the polysheaths only.

2.3.7 Investigation of oligomers of gp053 mutant

Based on the conserved model of the tail sheath of phage T4, these structures are formed only of the tail sheath protein subunits, which interact with neighbors and form hexameric rings (Kostyuchenko et al., 2005; Aksyuk et al., 2009; Leiman and Shneider, 2012). Thus, construction of tail sheath-based hybrid proteins allows production of only homogenic tail sheaths. In order to determine whether it is possible to obtain polysheaths with controlled disposition of various peptides in specific locations on its surface, oligomers of gp053 were constructed and analysis of polymerization potential of these proteins was carried out. m_N_C mutant protein with cloning sites in the N- and C-terminus (Fig. 2.13) was selected for these studies.

Plasmid vectors with m_N_C monomeric, dimeric, trimeric, tetrameric, pentameric and hexameric proteins coding genes were constructed, gene expression in *E. coli* BL21–DE3 strain was carried out. It was found that all oligomers were soluble. TEM analysis revealed that only in the case of m_N_C monomer, long (up to 400 nm) ordered tubular structures were visible in samples analyzed immediately after cell sonication (Fig 2.18 A). Dimeric m_N_C formed sparse, short tubular structures, the rest of the oligomers formed indefinite structures (Fig 2.18 B-F).

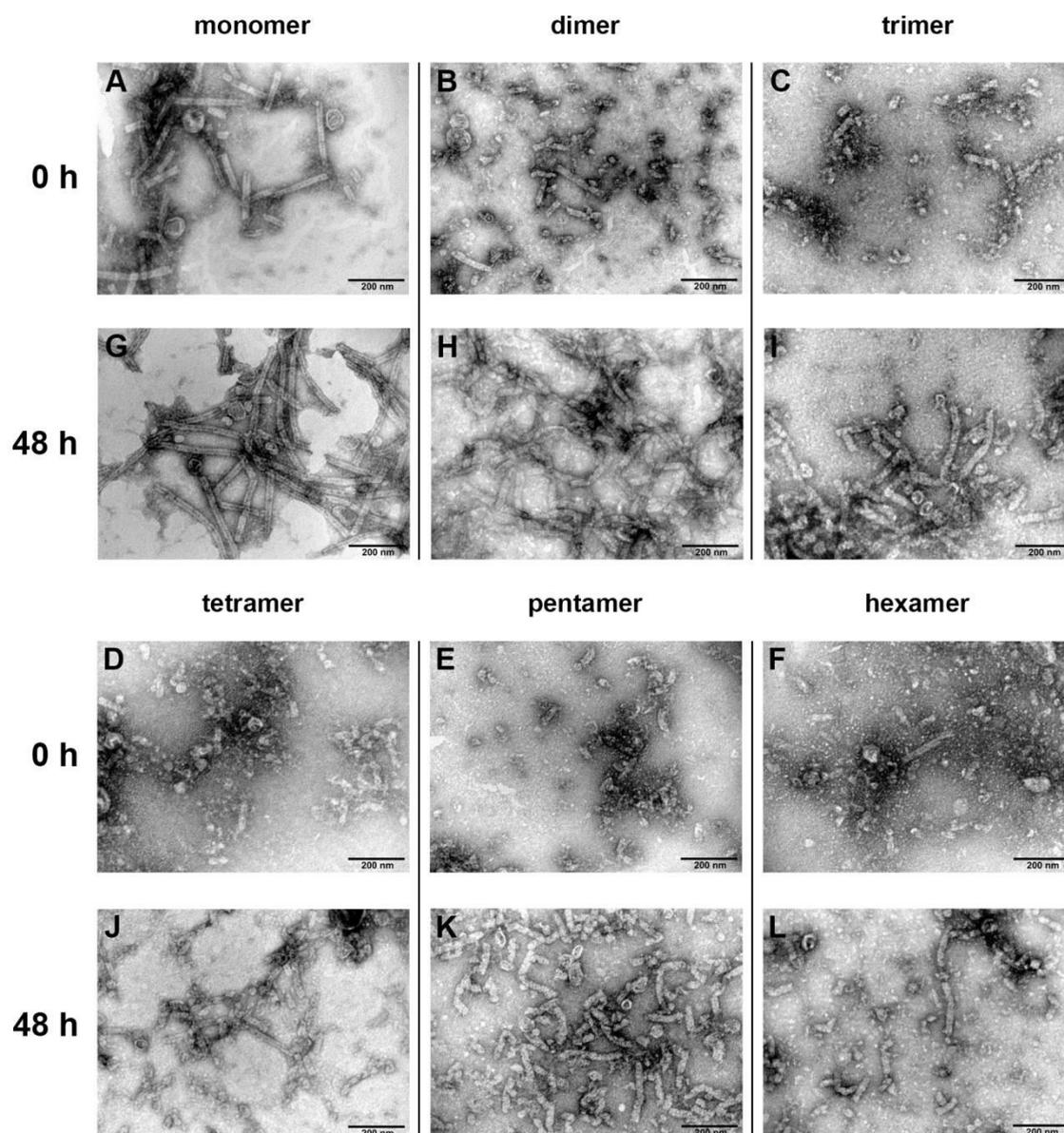


Fig. 2.18. TEM analysis of structures formed by m_NC oligomers. Titles above represent a number of m_NC duplications. Hours indicated in the left represent incubation time. Duplicates were incubated in TE buffer with periodical shaking at 22°C.

On the other hand, it was demonstrated that self-assembling process of these duplicates proceeds *in vitro*. TEM analysis revealed that m_NC oligomers form tubular, although less regular structures, after incubation for 48 hours in TE buffer with periodical shaking at 22°C (Fig 2.18 H-L). These results are promising in construction of nanostructures with various materials exhibited in specific locations on their surfaces as well as in more detailed investigation of polymerization processes of bacteriophage tail sheath assembly.

CONCLUSIONS

1. Of the 19 tested structural bacteriophage virion proteins (RaK2 gp041–gp043, gp106, gp107, ArV1 gp15, ArV2 gp05, gp06, gp11, gp15–gp18, gp20, gp24, NBD2 gp39, FV3 gp052, gp053), 3 recombinant proteins (RaK2 gp041, NBD2 gp39 and FV3 gp053) *in vivo*, in the absence of other phage proteins, self-assemble into ordered tubular structures.
2. Polysheaths formed by recombinant gp053 of FV3 are extremely stable and resistant to various environmental factors structures. The diameter of these nanostructures (~27 nm) corresponds to the width of a contracted tail of the phage FV3, the length varied from 10 to 1000 nm.
3. C-terminal amino acid deletions have less impact on the polymerization properties of gp053 of FV3 than the elimination of the amino acids from the N-terminus.
4. gp053 mutants and gp053-based hybrid proteins of FV3 *in vivo* self-assembles into nanotubular structures, whose morphology depends on the number of removed or inserted amino acids as well as on their position in the protein sequence.

LIST OF PUBLICATIONS

The thesis is based on the following original publications:

1. Kaliniene L*, **Šimoliūnas E***, Truncaitė L, Zajančkauskaitė A, Nainys J, Kaupinis A, Valius M, Meškys R. (2017) Molecular analysis of *Arthrobacter* myovirus vB_ArtM-ArV1: we blame it on the tail. *J Virol.* 91: e00023-17.
2. **Šimoliūnas E**, Kaliniene L, Stasilo M., Truncaitė L, Zajančkauskaitė A, Staniulis J, Nainys J, Kaupinis A, Valius M, Meškys R. (2014) Isolation and characterization of vB_ArS-ArV2 – first *Arthrobacter* sp. infecting bacteriophage with completely sequenced genome. *PLoS ONE* 9: e111230.
3. **Šimoliūnas E**, Kaliniene L, Truncaitė L, Zajančkauskaitė A, Staniulis J, Kaupinis A, Ger M, Valius M, Meškys R. (2013) *Klebsiella* phage vB_KleM-RaK2 – a giant singleton virus of the family *Myoviridae*. *PLoS ONE* 8: e60717.
4. Truncaite L, **Šimoliūnas E**, Zajančkauskaite A, Kaliniene L, Mankevičiūtė R, Staniulis J, Klausas V, Meškys R. (2012) Bacteriophage vB_EcoM-FV3: a new member of "rV5-like viruses". *Arch Virol.* 157: 2431–2435.
5. **Simoliūnas E**, Kaliniene L, Truncaite L, Klausas V, Zajanckauskaite A, Meskys R. (2012) Genome of *Klebsiella* sp.-infecting bacteriophage vB_KleM-RaK2. *J Virol.* 86: 5406.

* – authors contributed equally to publication

Results not discussed in PhD thesis were published in these original publications:

1. Alijošius L, **Šimoliūnas E**, Kaliniene L, Meškys R, Truncaitė L. (2017) Complete genome sequence of *Escherichia coli* phage vB_EcoM_Alf5. *Genome Announc.* (*accepted for publication*).
2. **Šimoliūnas E**, Vilkaitytė M, Kaliniene L, Zajančkauskaitė A, Kaupinis A, Staniulis J, Valius M, Meškys R, Truncaitė L. (2015) Incomplete LPS core-specific *Felix01likevirus* vB_EcoM_VpaE1. *Viruses.* 7: 6163-6181.
3. Gelzinis A., Verikas A., Vaiciukynas E., Backauskienė M., Sulčius S., **Simoliūnas E.**, Staniulis J., Paskauskas R. (2015) Automatic detection and morphological delineation of bacteriophages in electron microscopy images. *Comput Biol Med.* 64:101-116.
4. Kaliniene L, Zajančkauskaitė A, **Šimoliūnas E**, Truncaitė L, Meškys R. (2015) Low-temperature bacterial viruses VR – a small but diverse group of *E. coli* phages. *Arch Virol.* 160:1367-1370.
5. Šulčius S, **Šimoliūnas E**, Staniulis J, Koreivienė J, Baltrušis P, Meškys R, Paškauskas R. (2014) Characterization of a lytic cyanophage that infects the bloom-

forming cyanobacterium *Aphanizomenon flos-aquae*. FEMS Microbiology Ecology 91:1-7.

CONFERENCE PRESENTATIONS

Results discussed in PhD thesis were presented in these conferences:

1. **Šimoliūnas E**, Povilonienė S, Truncaitė L, Rutkienė R, Časaitė V, Kalinienė L, Pajeda S, Goda K, Meškys R. The application of structural proteins from phages for construction of self-assembling hybrid nanostructures. National Conference of Young Scientist "Biofuture: Perspectives of Nature and Life Sciences", 12 07 2016, Lithuanian Academy of Sciences, Vilnius, Lithuania (*Oral presentation*).
2. **Šimoliūnas E**, Časaitė V, Kalinienė L, Rutkienė R, Truncaitė L, Vilkaitytė M, Goda K, Meškys R. Investigation of tail-forming proteins of *Escherichia* phage vB_EcoS_NBD2. The 3rd Congress of Baltic Microbiologists (CBM2016), 10 18–21 2016, Vilnius, Lithuania (*Oral presentation*).
3. Vilkaitytė M, Kalinienė L, Truncaitė L, Zajančauskaitė A, **Šimoliūnas E**, Goda K, Meškys R. Characterization of low-temperature *Escherichia* phage vB_EcoS_NBD2. The 3rd Congress of Baltic Microbiologists (CBM2016), 10 18–21 2016, Vilnius, Lithuania (*Poster*).
4. Kaliniene L, Zajanckauskaite A, **Simoliunas E**, Vilkaityte M, Urbonavicius J, Meskys R, Truncaite L. Host-Specificity Determinants of Bacteriophage vB_EcoM_FV3. The 41st FEBS Congress "Molecular and System Biology for a Better Life", 09 03–08 2016, Ephesus/Kuşadası, Turkey (*Poster*).
5. Kaliniene L, Truncaite L, Zajanckauskaite A, **Simoliunas E**, Serviene E, Meskys R. Low-temperature enterobacteria phage vB_EcoS_NBD2 isolated from agricultural soil. The 41st FEBS Congress "Molecular and System Biology for a Better Life", 09 03–08 2016, Ephesus/Kuşadası, Turkey (*Poster*).
6. **Šimoliūnas E**, Povilonienė S, Rutkienė R, Kalinienė L, Truncaitė L, Urbelienė N, Nainys J, Meškys R. Investigation of self-assembling tail proteins of phages. XIV International Conference of the Lithuanian Biochemical Society, 06 28–30, 2016, Druskininkai, Lithuania (*Poster*).
7. **Šimoliūnas E**, Povilonienė S, Kalinienė L, Meškys R. Investigation of structural proteins from phages. IX National PhD Students Conference "Science for Health", 04 13 2016, Lithuanian University of Health Sciences, Kaunas, Lithuania (*Oral presentation*).
8. **Šimoliūnas E**, Povilonienė S, Kalinienė L, Truncaitė L, Zajančauskaitė A, Meškys R. Investigation and application of structural proteins from phages. 2015 National Conference of Young Scientist "Biofuture: Perspectives of Nature and Life Sciences", 12 10 2015, Lithuanian Academy of Sciences, Vilnius, Lithuania (*Oral presentation*).

9. **Simoliunas E**, Kaliniene L, Stasilo M, Truncaite L, Zajanckauskaite A, Meškys R. *Arthrobacter* phage vB_ArtM-ArV1: a solitary myovirus among the phages from family *Siphoviridae*. Conference "Phages 2015", 09 01–02 2015, St Hilda's College, Oxford, UK (*Poster*).
10. **Simoliunas E**, Poviloniene S, Kaliniene L, Truncaite L, Zajanckauskaite A, Meškys R. Investigation of tail-forming proteins of *Klebsiella* phage vB_KleM-RaK2. XXIV Biennial Conference "Phage/Virus Assembly", 06 07–12 2015, Les Diablerets, Switzerland (*Poster*).
11. **Šimoliūnas E**, Stasilo M, Nainys J, Meškys R. Bacteriophage vB_ArtM-ArV1 – first *Arthrobacter* sp. infecting myovirus with completely sequenced genome. VIII National PhD Students Conference "Science for Health", 04 10 2015, Lithuanian University of Health Sciences, Kaunas, Lithuania (*Oral presentation*).
12. **Simoliunas E**, Kaliniene L, Stasilo M, Truncaite L, Zajanckauskaite A, Meškys R. Genomic characterization of the first bacteriophage infecting *Arthobacter* sp. Conference "Phages 2014: Bacteriophage in Medicine, Food and Biotechnology", 09 16–18 2014, St Hilda's College, Oxford, UK (*Poster*).
13. **Šimoliūnas E**, Stasilo M, Kaliniene L, Truncaitė L, Zajančkauskaitė A, Meškys R. Bacteriophage vB_ArS-ArV2 – first *Arthrobacter* sp. infecting phage with completely sequenced genome. VII National PhD Students Conference "Science for Health", 04 09 2014, Lithuanian University of Health Sciences, Kaunas, Lithuania (*Oral presentation*).
14. **Šimoliūnas E**, Stasilo M, Kaliniene L, Truncaitė L, Zajančkauskaitė A, Meškys R. Characterization of vB_ArS-ArV2 – first *Arthobacter* sp. infecting bacteriophage with completely sequenced genome. Conference of Natural and Life sciences "The COINS'14", 03 3-8 2014, Faculty of Natural Sciences, Vilnius university, Vilnius, Lithuania (*Poster*).
15. Kaliniene L, **Šimoliūnas E**, Stasilo M, Truncaitė L, Zajančkauskaitė A, Meškys R. Characterization of newly isolated *Arthrobacter* bacteriophages ArV1 and ArV2. 20th Biennial Evergreen International Phage Meeting, 08 4-9 2013, The Evergreen State College, Olympia, WA 98505, USA (*Poster*).
16. **Šimoliūnas E**, Kaliniene L, Truncaitė L, Zajančkauskaitė A, Staniulis J, Meškys R. *Klebsiella* sp. infecting phage RaK2 – unique member of the family *Myoviridae*. VI National PhD Students Conference "Science for Health", 04 05 2013, Lithuanian University of Health Sciences, Kaunas, Lithuania (*Oral presentation, winner of section*).
17. Kaliniene L, **Simoliunas E**, Truncaite L, Zajanckauskaite A, Meškys R. Unique *Klebsiella* phage vB_KleM-RaK2 representing a novel genus within a family of *Myoviridae* of tailed bacteriophages. Conference "EuroPhages 2012: Bacteriophage in Medicine, Food and Biotechnology", 09 24–26 2012, St Hilda's College, Oxford, UK (*Poster*).
18. **Šimoliūnas E**, Kalinienė L, Zajančkauskaitė A, Truncaitė L. Investigation of DNA

modification profiles in phages. XII Lithuanian Biochemical Society conference "50 years of Biochemistry Studies in Lithuania", 06 28–30 2012 Tolieja, Lithuania (*Thesis*).

FINANCIAL SUPPORT

National projects funded by Research Council of Lithuania (RCL) contract No.: MIP-002/2014, SVE-04/2012, MIP-10240.

Research Council of Lithuania (RCL) doctoral scholarships (2013–2015).

Research Council of Lithuania (RCL) financial support for doctoral visits (2014–2015).

CURRICULUM VITAE

Name Eugenijus Šimoliūnas

Date of birth 1985 08 27

Work address Department of Molecular Microbiology and Biotechnology
Institute of Biochemistry, Life Sciences Centre, Vilnius University
Saulėtekio av. 7, LT-10257, Vilnius, Lithuania

Phone +37065070467

E-mail eugenijus.simoliunas@bchi.vu.lt

Education

2012–2016 PhD studies in Biochemistry
(Vilnius University, Institute of Biochemistry, Department of Molecular Microbiology and Biotechnology)

2009–2011 MSc in Microbiology
(Vilnius University, Faculty of Natural Sciences)

2004–2009 BSc in Biology (qualification of molecular microbiology)
(Vilnius University, Faculty of Natural Sciences)

Work experience

2010–present Specialist
(Department of Molecular Microbiology and Biotechnology, Institute of Biochemistry, Vilnius University, (from 2016 – Life Sciences Centre))

ACKNOWLEDGEMENTS

I am grateful to my supervisor dr. Rolandas Meškys for the opportunity to grow, learn and work in this area, as well as for advices, discussions and assistance in the preparation of scientific publications and my PhD thesis.

My special thanks to dr. Lidija Truncaitė for initial cloning of gp053 of FV3, as well as for scientific discussions, support, encouragement, faith and constant support in everyday work as well as in the preparation of scientific publications.

I would like to thank my student Miroslav Stasilo for the cloning and expression studies of the structural genes of ArV2, my student Simonas Pajeda and dr. Vida Časaitė for the cloning and expression studies of the gp039 of NBD2, dr. Simona Povilonienė for the cloning and expression studies of gp041 of RaK2, dr. Laura Kalinienė for the help preparing publications, dr. Algirdas Kaupinis for proteomic analysis of the phages virions and gp053 of FV3, prof. Juozas Staniulis, Juozas Nainys, Karolis Goda and Martynas Skapas for assistance with TEM.

Many thanks to Monika Vilkaitytė, dr. Aurelija Zajančkauskaitė, dr. Rasa Rutkienė, Virginija Dzekevičienė, Algimantas Krutkis, Lina Juškienė, Nijolė Uždavinienė and to all former and current members of the Department of Molecular Microbiology and Biotechnology for their help, support and guidance in daily work.

REZIUMĖ

Įvairių nanostruktūrų, paremtų savaime susirenkančiomis biomolekulėmis, konstravimas – ypač aktuali ir perspektyvi tyrimų sritis. Ieškant naujų, potencialių baltymų, kurie galėtų būti sėkmingai panaudoti šioje srityje, šiame darbe atlikti naujų bakteriojų ir jų struktūrinių baltymų tyrimai.

Charakterizuoti penki bakteriofagai: *Klebsiella* sp. infekuojantis bakteriofagas RaK2, *Arthrobacter* sp. fagai ArV1 ir ArV2, bei *Escherichia coli* bakteriofagai FV3 ir NBD2. Identifikuoti 122 minėtųjų bakteriofagų struktūriniai baltymai, 19 iš jų pasirinkti detalesniems tyrimams. Nustatyta, kad fagų RaK2 gp041 ir FV3 gp053, formuojantys bakteriojų uodegėlės apvaskalą, bei bakteriofago NBD2 uodegėlės vamzdelio baltymas gp39, *in vivo*, nesant kitų faginių baltymų, formuoja taisyklingas savitvarkes vamzdelines struktūras.

Optimizuotas bakteriofago FV3 gp053 formuojamų nanostruktūrų gryninimas, atlikti nanovamzdelių atsparumo įvairiems aplinkos veiksniams tyrimai. Nustatyta, kad rekombinantinio gp053 formuojami nanovamzdeliai – įvairiems neigiamiems aplinkos veiksniams itin atsparios struktūros. Tiriant galinių aminorūgščių svarbą baltymo polimerizacinėms savybėms, sukonstruoti 25 gp053 deleciniai mutantai, atlikta jų formuojamų nanostruktūrų analizė. Nustatyta, kad C-galinių aminorūgščių delecijos turi mažiau neigiamos įtakos FV3 gp053 susirinkimui į nanostruktūras nei aminorūgščių pašalinimas nuo baltymo N-galo.

Tiriant papildomai įterptų arba pakeistų svetimomis aminorūgščių įtaką gp053 polimerizacinėms savybėms, sukonstruota 20 gp053 mutantų su tikslinėmis klonavimo vietomis baltymo N- ir C-galuose, taip pat baltymo vidinėje dalyje. Nustatyta, kad papildomų aminorūgščių įterpimas į baltymo vidines sritis baltymo polimerizaciją riboja žymiai stipriau, nei svetimų aminorūgščių įterpimas į baltymo N- arba C-galus.

Taip pat sukonstruota 16 hibridinių baltymų, sudarytų iš gp053 arba gp053 mutantų ir į baltymo N- ir/arba C-galus arba vidinėje baltymo dalyje esančias tikslines klonavimo vietas įterptų sidabro jonus ir/arba titano oksidą surišančių peptidų. Nustatyta, kad hibridinių baltymų formuojamų nanostruktūrų morfologija priklauso nuo pašalintų ar įterptų aminorūgščių skaičiaus, taip pat jų pozicijos baltymo sekoje. Gautos hibridinės nanostruktūros su neutravidinu konjuguotomis aukso nanodalelėmis, kurios atrankiai

tvirtinasi prie biotinu modifikuoto gp053 formuojamų nanovamzdelių galų. Siekiant nustatyti, ar įmanoma gauti hibridinius, gp053 pagrindu sukurtus vamzdelius, kuriuose būtų galima kontroliuoti tam tikrų peptidų ar fermentų išsidėstymą vamzdelio struktūroje, atlikti mutantinio gp053 oligomerų tyrimai. Nustatyta, kad mutantinio gp053 oligomerai *in vivo* susirenka į trumpas, netvarkingas vamzdelines struktūras.

Apibendrinant rezultatus, galima teigti, kad šiame darbe atlikti bakteriofagų charakterizavimo tyrimai yra ypač svarbūs gilinant teorines žinias apie šių bakterijų virusų įvairovę, jų tarpusavio ryšius bei jų pačių ar jų produktų praktinio taikymo galimybes ateityje. Bakteriofagų struktūrinių baltymų, jų mutantų ar šių baltymų pagrindu sukurtų hibridinių baltymų tyrimai neabejotinai suteikia daug naujos informacijos ieškant naujų biomolekulių, tinkamų savaime susirenkančių nanostruktūrų konstravimui.

REFERENCES

1. Adams M (1959) Bacteriophages. Interscience publishers, Inc., New York.
2. Adriaenssens EM, Ackermann H-W, Anany H, Blasdel B, Connerton IF, Goulding D, Griffiths MW, Hooton SP, Kutter EM, Kropinski AM, Lee J-H, Maes M, Pickard D, Ryu S, Sepehrizadeh Z, Shahrabak SS, Toribio AL, Lavigne R (2012) A suggested new bacteriophage genus: “Viunalikevirus.” Arch Virol 157:2035–2046.
3. Aksyuk AA, Kurochkina LP, Fokine A, Forouhar F, Mesyanzhinov VV, Tong L, Rossmann MG (2011) Structural conservation of the *Myoviridae* phage tail sheath protein fold. Structure 19:1885–1894.
4. Aksyuk AA, Leiman PG, Kurochkina LP, Shneider MM, Kostyuchenko VA, Mesyanzhinov VV, Rossmann MG (2009) The tail sheath structure of bacteriophage T4: a molecular machine for infecting bacteria. EMBO J 28:821–829.
5. Arisaka F, Engel J, Klump H (1981) Contraction and dissociation of the bacteriophage T4 tail sheath induced by heat and urea. Prog Clin Biol Res 64:365–379.
6. Arisaka F, Takeda S, Funane K, Nishijima N, Ishii S (1990) Structural studies of the contractile tail sheath protein of bacteriophage T4. 2. Structural analyses of the tail sheath protein, gp18, by limited proteolysis, immunoblotting, and immunoelectron microscopy. Biochemistry 29:5057–5062.
7. Bhardwaj A, Walker-Kopp N, Wilkens S, Cingolani G (2008) Foldon-guided self-assembly of ultra-stable protein fibers. Protein Sci 17:1475–1485.
8. Busseron E, Ruff Y, Moulin E, Giuseppone N (2013) Supramolecular self-assemblies as functional nanomaterials. Nanoscale 5:7098–7140.
9. Carlson K, Miller E (1994) Experiments in T4 genetics. In: Karam J (ed) Molecular biology of bacteriophage T4. ASM Press, Washington, DC.
10. Carver T, Berriman M, Tivey A, Patel C, Böhme U, Barrell BG, Parkhill J, Rajandream M-A (2008) Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. Bioinformatics 24:2672–2676.
11. Casaite V, Poviloniene S, Meskiene R, Rutkiene R, Meskys R (2011) Studies of dimethylglycine oxidase isoenzymes in *Arthrobacter globiformis* cells. Curr Microbiol 62:1267–1273.
12. Daube SS, Arad T, Bar-Ziv R (2007) Cell-free co-synthesis of protein nanoassemblies: tubes, rings, and doughnuts. Nano Lett 7:638–641.
13. Davidson AR, Cardarelli L, Pell LG, Radford DR, Maxwell KL (2012) Long noncontractile tail machines of bacteriophages. Adv Exp Med Biol 726:115–142.
14. Douglas T, Young M (2006) Viruses: making friends with old foes. Science 312:873–875.
15. Fokine A, Rossmann MG (2014) Molecular architecture of tailed double-stranded DNA phages. Bacteriophage 4:e28281.
16. Glover DJ, Giger L, Kim SS, Naik RR, Clark DS (2016) Geometrical assembly of ultrastable protein templates for nanomaterials. Nat Commun 7:11771.
17. Guo P (2005) Bacterial virus phi29 DNA-packaging motor and its potential applications in gene therapy and nanotechnology. Methods Mol Biol 300:285–324.
18. Yokoi N, Inaba H, Terauchi M, Stieg AZ, Sanghamitra NJM, Koshiyama T, Yutani K, Kanamaru S, Arisaka F, Hikage T, Suzuki A, Yamane T, Gimzewski JK, Watanabe Y, Kitagawa S, Ueno T (2010) Construction of robust bio-nanotubes using the controlled self-assembly of component proteins of bacteriophage T4. Small 6:1873–1879.
19. Yuan Y, Gao M (2017) Jumbo Bacteriophages: An Overview. Front Microbiol. 2017 Mar 14;8:403.
20. Kaliniene L, Klausas V, Truncaite L (2010) Low-temperature T4-like coliphages vB_EcoM-VR5, vB_EcoM-VR7 and vB_EcoM-VR20. Arch Virol 155:871–880.
21. Katsura I (1981) Structure and function of the major tail protein of bacteriophage lambda. Mutants having small major tail protein molecules in their virion. J Mol Biol 146:493–512.
22. Katsura I, Tsugita A (1977) Purification and characterization of the major protein and the terminator protein of the bacteriophage λ tail. Virology 76:129–145.
23. Kostianen MA, Hiekkataipale P, Laiho A, Lemieux V, Seitsonen J, Ruokolainen J, Ceci P (2013) Electrostatic assembly of binary nanoparticle superlattices using protein cages. Nat Nano

- 8:52–56.
24. Kostyuchenko VA, Chipman PR, Leiman PG, Arisaka F, Mesyanzhinov VV, Rossmann MG (2005) The tail structure of bacteriophage T4 and its mechanism of contraction. *Nat Struct Mol Biol* 12:810–813.
 25. Krickler M, Carlson K (1994) Isolation of T4 phage DNA. In: Karam J (ed) *Molecular biology of bacteriophage T4*. ASM Press, Washington, DC, pp 455–456.
 26. Kropinski A (2009) Measurement of the rate of attachment of bacteriophage to cells. *Methods Mol Biol* 501:151–155.
 27. Kube S, Wendler P (2015) Structural comparison of contractile nanomachines. *AIMS Biophysics* 2:88–115.
 28. Kurochkina LP, Aksyuk AA, Sachkova MY, Sykilinda NN, Mesyanzhinov VV (2009) Characterization of tail sheath protein of giant bacteriophage ϕ KZ *Pseudomonas aeruginosa*. *Virology* 395:312–317.
 29. Kutanovas S, Rutkienė R, Urbelis G, Tauraitė D, Stankevičiūtė J, Meškys R (2013) Bioconversion of methylpyrazines and pyridines using novel pyrazines-degrading microorganisms. *Chemija* 24:67–73.
 30. Kuznetsova TA, Efimov AV, Aijrich LG, Kireeva IY, Marusich EI, Cappuccinelli P, Fiori P, Rappelli P, Kurochkina LP, Poglazov BF, Mesyanzhinov VV (1998) Properties of recombinant bacteriophage T4 tail sheath protein and its deletion fragments. *Biochemistry (Mosc)* 63:702–709.
 31. Laemli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
 32. Langlois C, Ramboarina S, Cukkemane A, Auzat I, Chagot B, Gilquin B, Ignatiou A, Petitpas I, Kasotakis E, Paternostre M, White HE, Orlova EV, Baldus M, Tavares P, Zinn-Justin S (2015) Bacteriophage SPP1 tail tube protein self-assembles into beta-structure-rich tubes. *J Biol Chem* 290:3836–3849.
 33. Lee EJ, Lee NK, Kim I-S (2016) Bioengineered protein-based nanocage for drug delivery. *Adv Drug Deliv Rev* 106:157–171.
 34. Lee S-Y, Lim J-S, Harris MT (2012) Synthesis and application of virus-based hybrid nanomaterials. *Biotechnol Bioeng* 109:16–30.
 35. Leiman PG, Shneider MM (2012) Contractile tail machines of bacteriophages. *Adv Exp Med Biol* 726:93–114.
 36. Li F, Wang Q (2014) Fabrication of nanoarchitectures templated by virus-based nanoparticles: strategies and applications. *Small* 10:230–245.
 37. Lowry O, Rosebrough N, Farr A, Randall R (1951) Protein measurement with the folin phenol reagent. *J Biol Chem* 193:265–275.
 38. Mandal D, Nasrolahi Shirazi A, Parang K (2014) Self-assembly of peptides to nanostructures. *Org Biomol Chem* 12:3544–3561.
 39. Molek P, Bratkovic T (2015) Bacteriophages as scaffolds for bipartite display: designing swiss army knives on a nanoscale. *Bioconj Chem* 26:367–378.
 40. Moon J-S, Kim W-G, Kim C, Park G-T, Heo J, Yoo SY, Oh J-W (2015) M13 bacteriophage-based self-assembly structures and their functional capabilities. *Mini-Reviews in Organic Chemistry* 12:271–281.
 41. Nam KT, Lee YJ, Krauland EM, Kottmann ST, Belcher AM (2008) Peptide-mediated reduction of silver ions on engineered biological scaffolds. *ACS Nano* 2:1480–1486.
 42. Pinheiro AV, Han D, Shih WM, Yan H (2011) Challenges and opportunities for structural DNA nanotechnology. *Nat Nano* 6:763–772.
 43. Pires DP, Cleto S, Sillankorva S, Azeredo J, Lu TK (2016) Genetically engineered phages: a review of advances over the last decade. *Microbiol Mol Biol Rev* 80:523–543.
 44. Poglazov BF, Efimov AV, Marco S, Carrascosa J, Kuznetsova TA, Aijrich LG, Kurochkina LP, Mesyanzhinov VV (1999) Polymerization of bacteriophage T4 tail sheath protein mutants truncated at the C-termini. *J Struct Biol* 127:224–230.
 45. Pokorski JK, Steinmetz NF (2011) The art of engineering viral nanoparticles. *Mol Pharm* 8:29–43.
 46. Quan S, Dabbs ER (1993) Nocardioform arsenic resistance plasmid characterization and improved *Rhodococcus* cloning vectors. *Plasmid* 29:74–79.

47. Sambrook J, Russell D (eds) (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, NY, USA, pp 430–433.
48. Sarris PF, Ladoukakis ED, Panopoulos NJ, Scoulica EV (2014) A phage tail-derived element with wide distribution among both prokaryotic domains: a comparative genomic and phylogenetic study. *Genome Biol Evol* 6:1739–1747.
49. Semėnaitė R, Duran R, Marcinkevičienė L, Bachmatova I, Pacenkaitė J (2000) Degradation of pyridine and 2-hydroxypyridine by environmental *Rhodococcus* and *Arthrobacter* strains. *Biologija* 2:329–331.
50. Semenaitė R, Gasparavičiūtė R, Duran R, Precigou S, Marcinkevičiūtė L, Bachmatova I, Meškys R (2003) Genetic diversity of 2-hydroxypyridine-degrading soil bacteria. *Biologija* 2:27–29.
51. Sharma PP, Rathi B, Rodrigues J, Gorobets NY (2015) Self-assembled peptide nanoarchitectures: applications and future aspects. *Curr Top Med Chem* 15:1268–1289.
52. Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction *Nucleic Acids Res.* 33(Web Server issue): W244–W248.
53. Stanislauskienė R, Gasparavičiūtė R, Vaitekunas J, Meskiene R, Rutkiene R, Casaite V, Meskys R (2012) Construction of *Escherichia coli*-*Arthrobacter*-*Rhodococcus* shuttle vectors based on a cryptic plasmid from *Arthrobacter rhombi* and investigation of their application for functional screening. *FEMS Microbiol Lett* 327:78–86.
54. Stanislauskienė R, Rudenkov M, Karvelis L, Gasparavičiūtė R, Meškienė R, Časaitė V, Meškys R (2011) Analysis of phthalate degradation operon from *Arthrobacter* sp. 68b. *Biologija* 57:45–54.
55. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739.
56. Zhang F, Nangreave J, Liu Y, Yan H (2014) Structural DNA nanotechnology: state of the art and future perspective. *J Am Chem Soc* 136:11198–11211.
57. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9:40.