

# **MGV Database: Extracting Connector Protein**

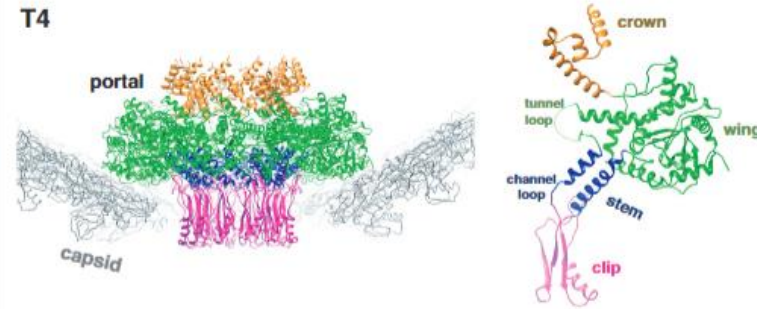
# Importance of Bacteriophages

- Affect our microbiome
- The microbiome has been shown to influence everything from physical health (ie obesity) to mental disorders (ie anxiety, depression)

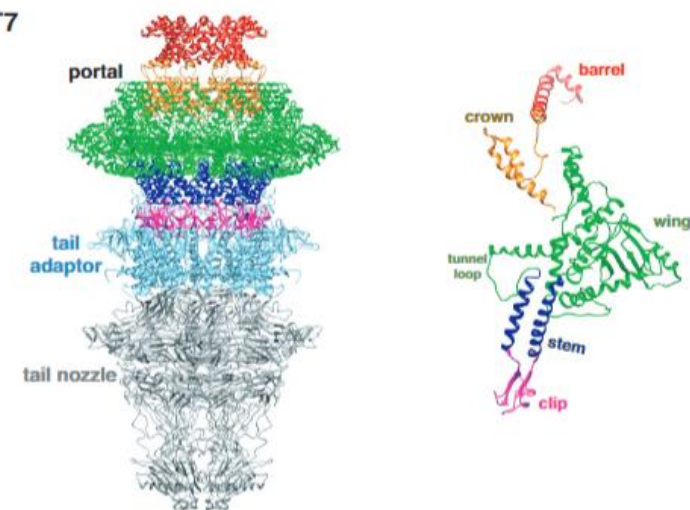
# Connector Protein

- Helps package DNA into the capsid
- Assembly of the motor

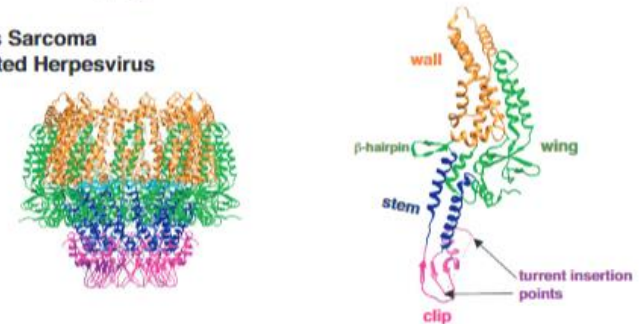
T4



T7



Kaposi's Sarcoma  
Associated Herpesvirus



# Goal

1. Find connector protein sequence
2. Simulate sequences

Improve understanding of the general connector protein structure

Eventual goal: We want to simulate the structure of an entire virus

# Viral Protein Homology

- Connector protein sequences aren't similar because the phages evolved the protein independently
- Common methods of annotating proteins:
  - Structural: modelling and comparing
  - Sequence: comparing sequence to database
- Homology = sequence related because of common ancestor
- Similarity = degree of likeness between sequences

BUT the genomic organization (order of genes) and structure is similar

# Phages have adapted the same protein fold to fulfill multiple functions in virion assembly

Check for updates

Lia Cardarelli<sup>a</sup>, Lisa G. Pell<sup>a,b</sup>, Philipp Neudecker<sup>a,c,d</sup>, Nawaz Pirani<sup>a,b</sup>, Amanda Liu<sup>c</sup>, Lindsay A. Baker<sup>a,b</sup>, John L. Rubinstein<sup>a,b</sup>, Karen L. Maxwell<sup>c</sup>, and Alan R. Davidson<sup>a,c,1</sup>

Departments of <sup>a</sup>Biochemistry and <sup>c</sup>Molecular Genetics, University of Toronto, Toronto, ON, Canada M5S 1A8; <sup>b</sup>Molecular Structure and Function Program, The Hospital for Sick Children Research Institute, Toronto, ON, Canada M5G 1X8; and <sup>d</sup>Department of Chemistry, University of Toronto, Toronto, ON, Canada M5S 3H6

Edited\* by Michael G. Rossmann, Purdue University, West Lafayette, IN, and approved June 21, 2010 (received for review April 28, 2010)

Evolutionary relationships may exist among very diverse groups of proteins even though they perform different functions and display little sequence similarity. The tailed bacteriophages present a uniquely amenable system for identifying such groups because of their huge diversity yet conserved genome structures. In this work, we used structural, functional, and genomic context comparisons to conclude that the head–tail connector protein and tail tube protein of bacteriophage  $\lambda$  diverged from a common ancestral protein. Further comparisons of tertiary and quaternary structures indicate that the baseplate hub and tail terminator proteins of bacteriophage may also be part of this same family. We propose that all of these proteins evolved from a single ancestral tail tube protein fold, and that gene duplication followed by differentiation

the connector and passes down the tail into the cell. The portion of the connector that is inserted into the head is composed of a dodecameric ring of the product of gene *B* (gpB), also known as the portal protein. The bottom surface of the connector (Fig. 1A), which interacts with the tail, is composed of gpFII (5). Another protein, gpW, is required for the stabilization of the DNA within the head and for the addition of gpFII (6, 7), suggesting that it may be positioned in the connector between gpB and gpFII. *Bacillus subtilis* phage SPP1 gp16, a protein with the same structure, function, and genomic position as gpFII (2) (Fig. 1A and C), has been shown by cryoelectron microscopy (cryoEM) to form a 12-membered ring within the connector (8, 9). Although the number of molecules of gpFII in assembled phage particles has

# Data

- The paper analyzed bacteriophages found in the human gut microbiome
- 189,680 viral genomes
- Found many NEW species
- Clustered viral proteins using MMseqs2 and annotated using HMMER (pHMM that searches database for sequence homologs) against HMM databases (KEGG, TIGRFAM, Pfam, VOGDB)

## Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome

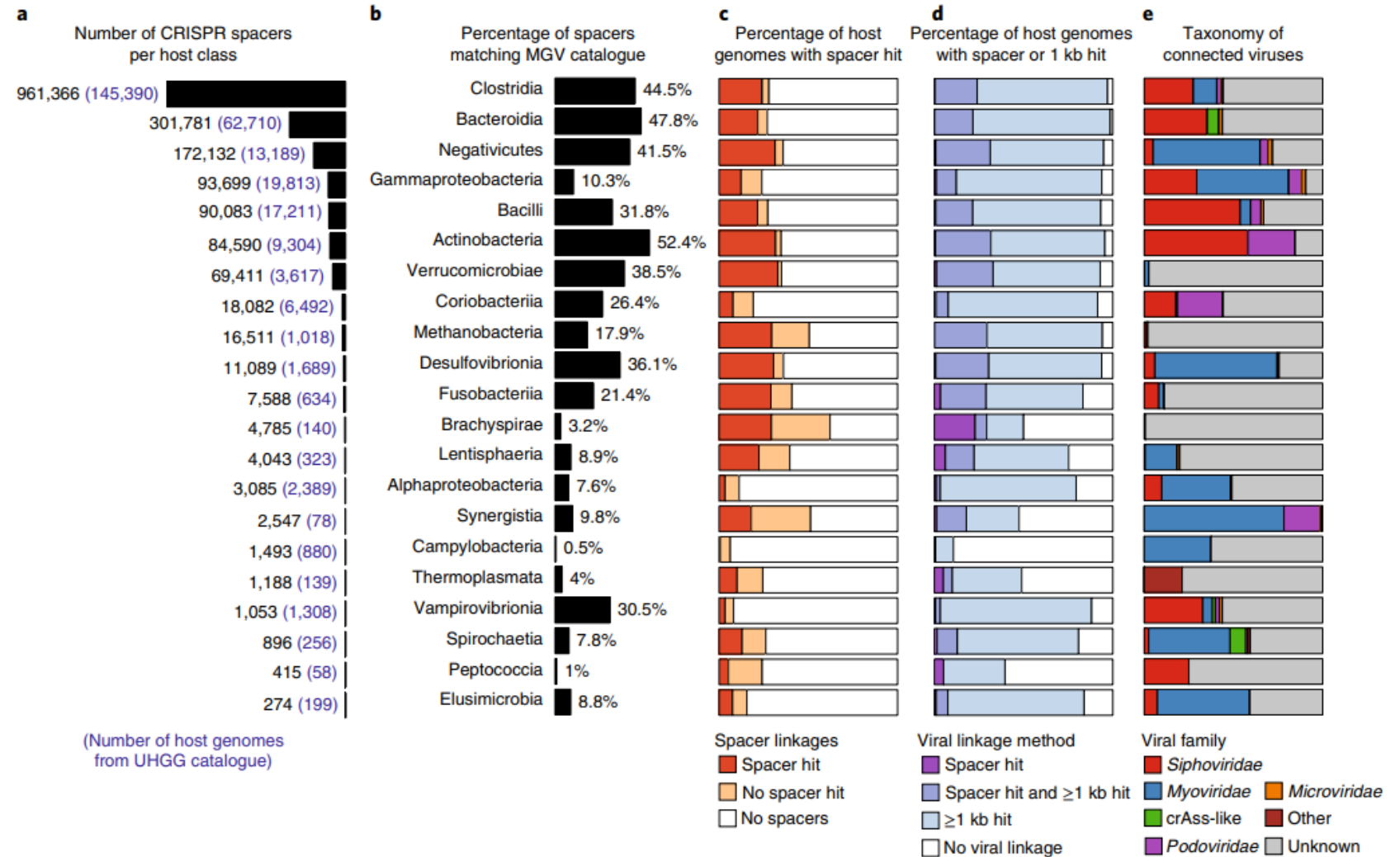
Stephen Nayfach<sup>1,2</sup>  , David Páez-Espino<sup>1,2</sup> , Lee Call<sup>1,2</sup>, Soo Jen Low<sup>3</sup>, Hila Sberro<sup>4,5</sup>, Natalia N. Ivanova<sup>1,2</sup> , Amy D. Proal<sup>6</sup>, Michael A. Fischbach<sup>1,7,8,9,10</sup> , Ami S. Bhatt<sup>4,5</sup> , Philip Hugenholtz<sup>3</sup>  and Nikos C. Kyrpides<sup>1,2</sup>  

Bacteriophages have important roles in the ecology of the human gut microbiome but are under-represented in reference databases. To address this problem, we assembled the Metagenomic Gut Virus catalogue that comprises 189,680 viral genomes from 11,810 publicly available human stool metagenomes. Over 75% of genomes represent double-stranded DNA phages that infect members of the Bacteroidia and Clostridia classes. Based on sequence clustering we identified 54,118 candidate viral species, 92% of which were not found in existing databases. The Metagenomic Gut Virus catalogue improves detection of viruses in stool metagenomes and accounts for nearly 40% of CRISPR spacers found in human gut Bacteria and Archaea. We also produced a catalogue of 459,375 viral protein clusters to explore the functional potential of the gut virome. This revealed tens of thousands of diversity-generating retroelements, which use error-prone reverse transcription to mutate target genes and may be involved in the molecular arms race between phages and their bacterial hosts.



# Family

- “Distribution of known viral families that are associated with each host class. Each host class is infected by a distinct repertoire of viral families”
- Interested in what families the proteins we find come from





# Metadata

## mgv\_contig\_info.tsv

- Metadata for the 189,680 viral genomes. Fields include:
- `votu_id`: indicate the species-level viral OTU the genome belongs to
- `checkv_quality`: medium quality (50-90% complete), high quality (>90% complete), complete (closed genome)
- `prophage`: whether or not the contig was flanked by DNA from the host (these regions were removed)
- `temperate_score`: BACPHLIP output indicating the probability the virus lives a temperate lifestyle
- `virulent_score`: BACPHLIP output indicating the probability the virus lives a virulent lifestyle
- `completeness`: CheckV estimated completeness
- `gc`: GC content
- `stop_codon_readthrough`: indicates whether the virus is predicted to read through a particular stop codon
- `baltimore`: baltimore classification
- `ictv_order`, `ictv_family`, `ictv_genus`: annotations based on the ICTV taxonomy

contig_id	votu_id	length	checkv_quality	prophage	temperate_score	virulent_score	completeness	gc	stop_codon_readthrough	baltimore	ictv_order	ictv_family	
MGV-GENOME-0364295	OTU-61123	97376	Complete	No	0.0375	0.9625	98.26	31.6166	TAG	dsDNA	Caudovirales	crAss-phage	NULL
MGV-GENOME-0364296	OTU-61123	97376	Complete	No	0.0375	0.9625	98.26	31.6146	TAG	dsDNA	Caudovirales	crAss-phage	NULL
MGV-GENOME-0364303	OTU-05782	97388	Complete	No	0.0357402	0.96426	98.28	27.9706	NULL	dsDNA	Caudovirales	crAss-phage	NULL
MGV-GENOME-0364311	OTU-01114	97394	Complete	No	0.0375	0.9625	98.38	31.4485	TAG	dsDNA	Caudovirales	crAss-phage	NULL
MGV-GENOME-0364312	OTU-23935	97395	Complete	No	0.0138753	0.986125	99.25	33.5777	TAG	dsDNA	Caudovirales	crAss-phage	NULL

# Data

## Proteins

### mgv\_proteins

- `protein_id`
- `sequence (aa)`

(# lines in file: 23,674,396)

(# of sequences: 11,837,198)

### mgv\_pc\_info

- `pc_id`
- `size`
- `avg_gene_length`
- `min_gene_length`
- `max_gene_length`
- `rep_id`
- `gene_ids (protein_id)`

(# clusters: 459,375)

### mgv\_pc\_func

- `pc_id` (protein cluster)
- `gene family (annotation)`
- `description`
- `fraction_pc_with_annotation`

(# annotated: 95,164)

# Data

- Baltimore Classification:
- 7 classes based on **nucleic acid** (DNA / RNA), **strandness** (double / single), **sense**, **method of replication**

## mgv\_sample\_info

- contig\_id
- assembly\_source
- assembly\_name
- study\_accession (# unique: 179,323)
- sample\_accession (# unique: 188,684)
- run\_accessions
- continent
- country\_code
- sex
- age
- health
- disease

## mgv\_contig\_info.tsv.gz

- Metadata for the 189,680 viral genomes. Fields include:
- contig\_id: indicate the species-level viral OTU the genome belongs to
- checkv\_quality: medium quality (50-90% complete), high quality (>90% complete), complete (closed genome)
- prophage: whether or not the contig was flanked by DNA from the host (these regions were removed)
- temperate\_score: BACPHLIP output indicating the probability the virus lives a temperate lifestyle
- virulent\_score: BACPHLIP output indicating the probability the virus lives a virulent lifestyle
- completeness: CheckV estimated completeness
- gc: GC content
- stop\_codon\_readthrough: indicates whether the virus is predicted to read through a particular stop codon
- baltimore: baltimore classification
- ictv\_order, ictv\_family, ictv\_genus: annotations based on the ICTV taxonomy

## mgv\_votu\_representatives

- contig\_id
- vOTU

(# vOTUs: 54,118)

## mgv\_contigs

- contig\_id
- sequence (DNA)

(# lines: 189,681)

(# sequences: 189,680)

## mgv\_host\_assignments.tsv.gz

- contig
- host: (# unique: 246)
- host\_phylum: (# unique: 102000)
- host\_class: (# unique: 102197)
- host\_order: (# unique: 127,548)
- host\_family: (# unique: 145047)
- host\_genus: (# unique: 141839)
- host\_species: (# unique: 112148)

(# lines: 170,093)

# Annotations in Dataset

- **Paper:**
  - Clustered data using MMseqs2
  - Annotated 20% using HMMER:
    - HMMER: detects homology by comparing a profile-HMM (a Hidden Markov model constructed explicitly for a particular search) to either a single sequence or a database of sequences.
- **Pfam:**
  - collection of protein families (MSA and HMMs)
- **Results**
  - 411 portal proteins
  - 146 connector

annotation	vpc_id	protein_id	protein_seq
⌵	⌵	⌵	⌵
Phage gp6-like head-tail connector protein	VPC-8627	MGV-GENOME-0282701_34	MSLDDEKILEKIKFSCRIDDDI
Phage gp6-like head-tail connector protein	VPC-16699	MGV-GENOME-0270537_10	MLSMADFEDTVLINVKEDLA
Phage gp6-like head-tail connector protein	VPC-135993	MGV-GENOME-0232097_34	MSIKNLMGTVTDDDLQLTKT
Phage gp6-like head-tail connector protein	VPC-545	MGV-GENOME-0260596_65	MEYTTLEQVKIRLKQFHIDTV
Phage gp6-like head-tail connector protein	VPC-456140	MGV-GENOME-0209946_11	MSGEAAAFKPPNRTERTKER

# Workflow

- 1. Find connector protein domain** (in literature / NCBI)
- 2. BLAST**
- 3. Filter alignment results**
  - want very low e-values and high bit scores
- 4. Check if the proteins are annotated / clustered** (mgv\_pc\_info.tsv.gz)
- 5. Model some proteins in list / cluster**
  - to make sure the results are correct

# 1. Connector Domain

- **Query:** Phage connector domain (from NCBI, bacteriophage phi29) →
- **Database:** MGv database
- **BLAST results:** 540 sequences
- Filtered by e-value:  $2e-60$  as a threshold (was just the highest “significant” e-value)
- Found protein cluster (in database) that matched most of the BLAST outputs (5464 sequences)

Sequence Alignment

include consensus sequence ?

Reformat

Format: 

Hypertext

Row Display: 

All 4 rows

Color Bits: 

2.0 bit

Type Selection: 

top listed sequences

1IJG\_I

8

TYRS----INEIQRQK----RNR--WFIHYLN

YLSLAYQLFEWENLP

PTINPSFLEKSIHQFGYVGFYKDPVISYIACN

77

Bacillus virus phi29

Q37891

7

SYKS----INDIQRMR----GNR--WYHYHYQ

YLCSLAYQLFEWERLP

SPVDP

SYLEKSIHQFGYVGFYKDPRIQYIACQ

76

Bacillus virus B103

Q37995

2

SYKNykrhLGKIELNKetveRNRlaFFEFY

FNYFYNI

VVNYFTWEGLPNDIDELFIEKKLIENGHVAFFHDDTFGYIAQG

81

Streptococcus phage Cp-1

Q9FZW5

7

SYKT----IGEIQRRR----GNL--WFRTYQ

RYLFLSLAYQMFEWQGLPKTVDP

IFLEKQLHQRGFVAFYKDEMYGLGVQ

76

Bacillus virus GA1

1IJG\_I

78

GALSGQRDVYNQATVFR---AASPVYQ

KEFKLYN----

YR---DMKEEDMG-----

VVIYNNDMAFPTTPTLELFAAEL

141

Bacillus virus phi29

Q37891

77

GALSGTVDHYNLPDRFH---ASSVGYQ

NTFKLYN----

YS---DMKEKNMG-----

VAIYNNDLKCSLTLPALMFAQDL

140

Bacillus virus B103

Q37995

82

GTRGERLNHYDQPLTYQpvnASSMNYFKQ

MEIAYtendFRvieELHKDNPDKikr

pcIVIPNNNFYEPYIGYLELFCCKL

161

Streptococcus phage Cp-1

Q9FZW5

77

GTLSGQINLYNQPNFYT---ASAPTYQ

KSFPLYW----

YDmgeDLNEKGQG-----

IVIYNNLERMPTLDILNLYAMNL

143

Bacillus virus GA1

1IJG\_I

142

AELKEIISVNQNAQKTPVLIRANDNNQ

LSLKQVYNQYEGNAPVIFAHEALD-----

SDSIEVFKTDAPYVVDKLN

211

Bacillus virus phi29

Q37891

141

AELKEIIAVNQNAQKTPVLIAANDNNQ

LSLKIYNQYEGNAPVIFVHESLD-----

LDNLKVFKTDAPYVVDKLN

210

Bacillus virus B103

Q37995

162

ADIELTIQLNRNAQITPYFIFADNTN

VLSMKNIFNKIANFEPVVLN

KQKdqgqdsfkqlSDYIQVFR

TAPFLDLKLH

241

Streptococcus phage Cp-1

Q9FZW5

144

AELKETIYVNQNAQKTPVIKAGDNDL

FSMKQVYNKYEGNEPVIFAGKKFN-----

TDDIEVLKTDAPYVADKLT

213

Bacillus virus GA1

1IJG\_I

212

AQKNAVWNEMMTFLGIKANLEKKERM

VTDEVSSNDEQIESSGTVFLKSREE

ACEKINELYGLNVKVKFRYDIV

285

Bacillus virus phi29

Q37891

211

AQKNAVWNEVMTYLGIKANLEKKERM

VTSEVDSNDEQIESSGNIYLKARQE

ACNKISELYGLNLKVKFRYDIV

284

Bacillus virus B103

Q37995

242

DEKL

RVMNQLLTFIGINNPSDKKERLVSE

ISNNGVISANIEVGWKSRRKFVEL

INKCYGLEISVKPAETIQ

315

Streptococcus phage Cp-1

Q9FZW5

214

MLFKDQWNEAMTFLGLSNANTDKKER

LIQSEVESNNDQIQGSANIYLAPRQE

ACRLINEYYGLNVSVKLRKELV

287

Bacillus virus GA1

## 2&3. BLAST results

```
MGV-GENOME-0119675_3 # 4070 # 4969 # -1 # ID=1077_3;partial=00;st... 201 2e-60
MGV-GENOME-0089456_15 # 11494 # 12366 # 1 # ID=606_15;partial=00;... 200 2e-60
MGV-GENOME-0098546_17 # 11992 # 12870 # 1 # ID=1255_17;partial=00... 200 2e-60
MGV-GENOME-0102802_18 # 12383 # 13201 # 1 # ID=206_18;partial=00;... 199 2e-60
MGV-GENOME-0102795_18 # 12383 # 13201 # 1 # ID=204_18;partial=00;... 199 2e-60
MGV-GENOME-0102791_18 # 12383 # 13201 # 1 # ID=203_18;partial=00;... 199 2e-60
```

- **E-value**
  - number of expected hits of similar quality (score) that could be found just by chance
- **Bit score**
  - Required size of a sequence database in which the current match could be found just by chance



## 2&3. BLAST

Sequences producing significant alignments:	Score (Bits)	E Value
MGV-GENOME-0212193_24 # 19174 # 20226 # 1 # ID=267_24;partial=00;...	140	2e-36
MGV-GENOME-0159433_29 # 20711 # 21538 # 1 # ID=491_29;partial=00;...	138	2e-36
MGV-GENOME-0210500_19 # 15471 # 16532 # 1 # ID=861_19;partial=00;...	139	3e-36
MGV-GENOME-0117354_6 # 5749 # 6810 # 1 # ID=496_6;partial=00;star...	139	4e-36
MGV-GENOME-0222640_14 # 9874 # 10935 # -1 # ID=1055_14;partial=00...	136	4e-35
MGV-GENOME-0209211_21 # 16647 # 17711 # 1 # ID=1419_21;partial=00...	135	1e-34
MGV-GENOME-0131812_19 # 12507 # 12902 # -1 # ID=367_19;partial=00...	76.6	9e-15
MGV-GENOME-0191353_6 # 4015 # 4557 # -1 # ID=2682_6;partial=00;st...	75.1	7e-14
MGV-GENOME-0191353_5 # 3534 # 3938 # -1 # ID=2682_5;partial=00;st...	63.9	3e-10
MGV-GENOME-0105632_1 # 3 # 485 # -1 # ID=725_1;partial=10;start_t...	58.9	3e-08
MGV-GENOME-0214625_28 # 24444 # 25526 # 1 # ID=514_28;partial=00;...	55.5	3e-06
MGV-GENOME-4395318_4 # 2697 # 3779 # 1 # ID=1794_4;partial=00;sta...	53.5	1e-05
MGV-GENOME-0215696_35 # 28616 # 29707 # -1 # ID=1451_35;partial=0...	51.6	5e-05

- BLAST results from using the entire connector protein as the query.
- After checking if the results correspond to any protein clusters, I found VPC\_8016, which has exactly the same number of proteins in the cluster as there are in the filtered BLAST results (269 proteins)

## 4. Check for annotation & family

### Annotation

```
VPC-34 pfam|PF05352.12 Phage Connector (GP10) 1
```

### Families

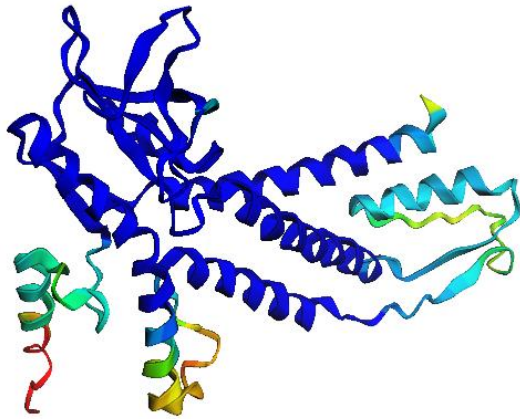
```
167 NULL
  1 Papillomaviridae
5296 Podoviridae
```

## 5. Modeling

- Modeled some of the proteins found in the protein cluster and in BLAST results (VPC-34)

- Result:

- AlphaFold:



- Swiss-MODEL:

VPC-34: MGV-GENOME-0100472\_17 Created: today at 17:54

Summary

Templates 21

Models 2

Project Data

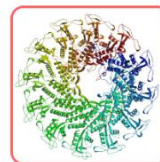
Model Results

Order by: GMQE



Automodel is running - more models are still to be built for this project.

Modelling job 02 is RUNNING.



Model 01

Structure Assessment

Oligo-State  
Homo-12-mer  
(matching prediction)

GMQE  
0.59

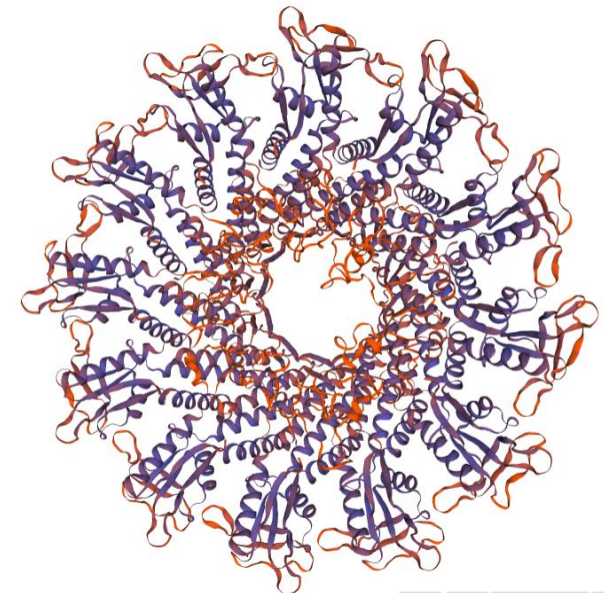
QMEANDisCo Global:  
0.62 ± 0.05

QMEANDisCo Local  
QMEAN Z-Scores

Template

7pv2.1.A Head-tail connector (Portal protein)  
GA1 bacteriophage portal protein

Seq Identity  
37.64%



pLDDT: ■ Very low (<50) ■ Low (60) ■ OK (70) ■ Confident (80) ■ Very high (>90)

# 5. Modelling

- Modeled protein found in VPC-34 but not in BLAST results

MGV-GENOME-0113413\_9 Created: today at 19:53

Summary

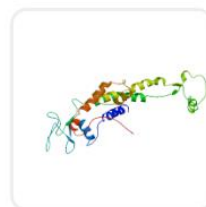
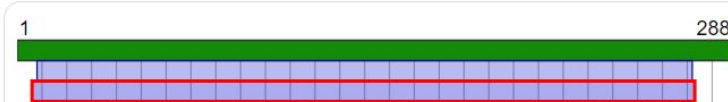
Templates 16

Models 2

Project Data

Model Results

Order by: GMQE



Model 02

Structure Assessment



Oligo-State  
Monomer

GMQE  
0.42

QMEANDisCo Global:  
0.42 ± 0.05

QMEANDisCo Local

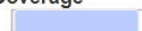
QMEAN Z-Scores

Template

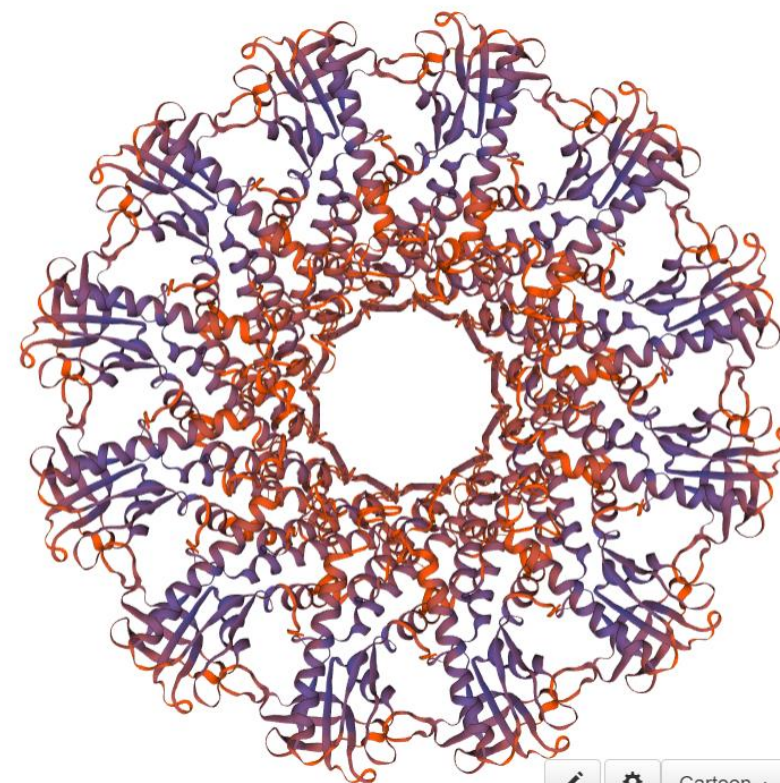
6q3g.40.A Portal protein  
Structure of native bacteriophage P68

Seq Identity  
17.39%

Coverage



Model-Template Alignment



Cartoon



## MGV-GENOME-0113413\_9 Created: today at 19:53

[Summary](#)[Templates](#) 16[Models](#) 2[Project Data](#) ▼

## VPC-34: MGV-GENOME-0100472\_17 Created: today at 17:54

[Summary](#)[Templates](#) 21[Models](#) 2[Project Data](#) ▼

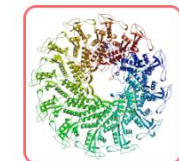
### Model Results ⓘ

Order by: GMQE ▼



Automodel is running - more models are still to be built for this project. ⚙️

Modelling job 02 is RUNNING. ⚙️



Model 01 ▼  
Structure  
Assessment

Oligo-State  
Homo-12-mer  
(matching prediction)

GMQE  
0.59

QMEANDisCo Global:  
0.62 ± 0.05

QMEANDisCo Local  
QMEAN Z-Scores

Template

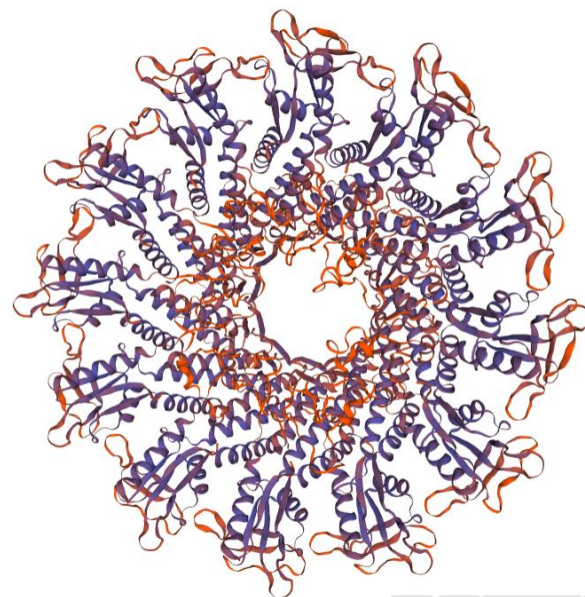
7pv2.1.A Head-tail connector (Portal protein)

Seq Identity

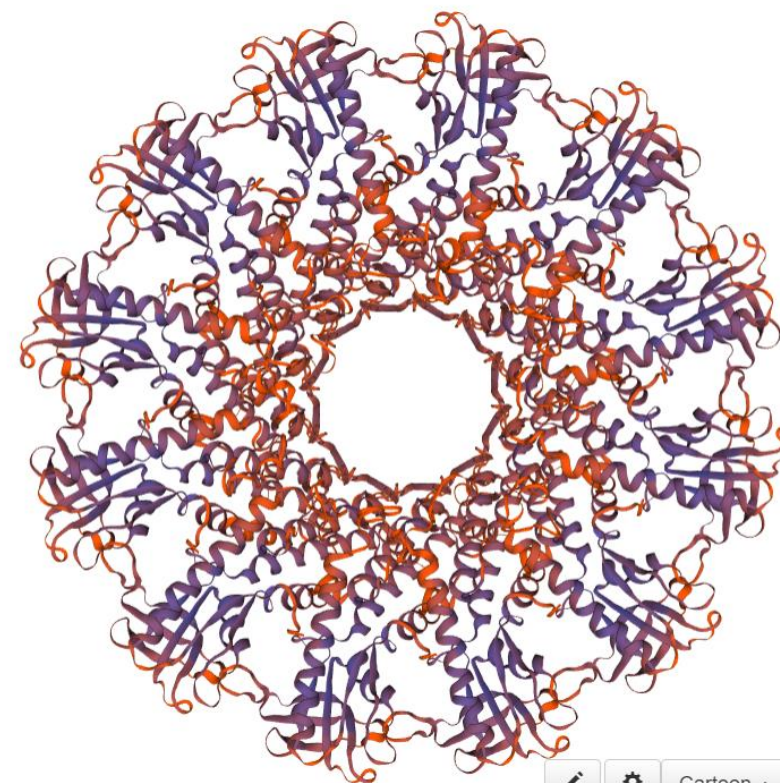
GA1 bacteriophage portal protein

37.64%

Model-template Alignment



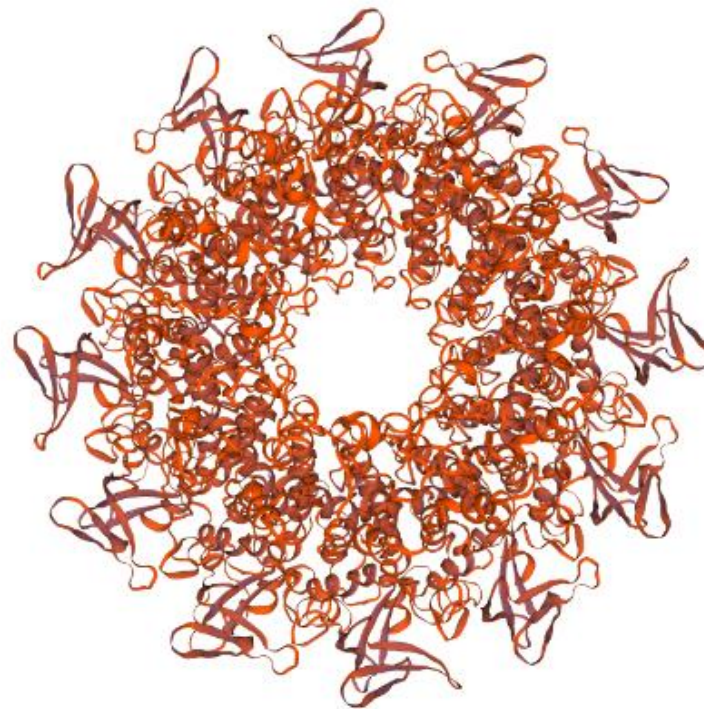
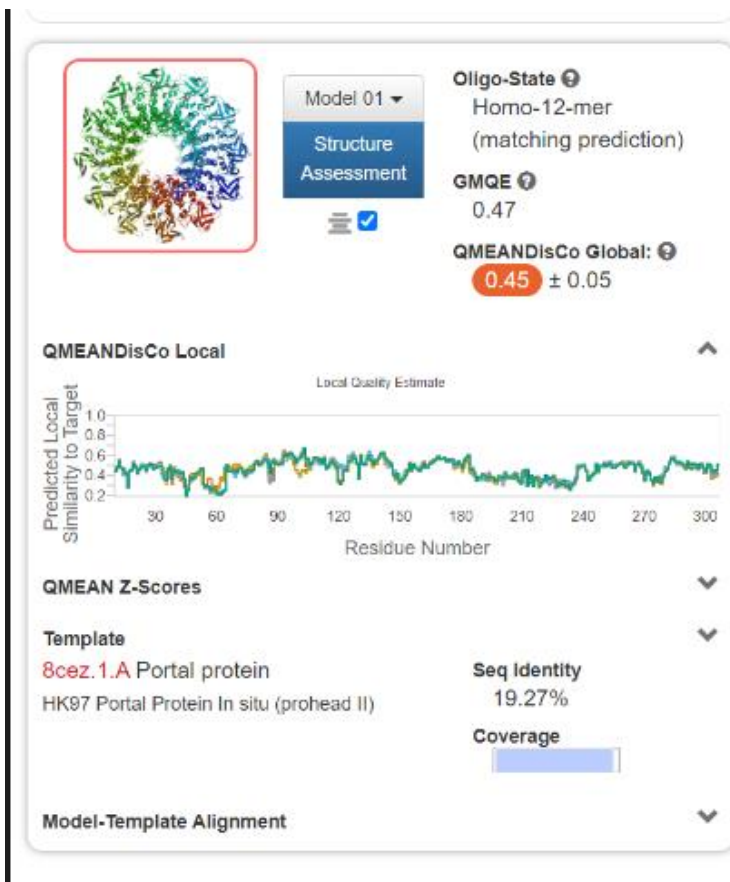
🔍 ⚙️ Cartoon ▲ 📷 ▲



🔍 ⚙️ Cartoon ▲ 📷 ▲

# Query Sequence

>YP\_010091615.1 portal protein [uncultured Caudovirales phage]





# BLAST: Query Sequence

```
(base) claireh@claire-virtualbox:/media/sf_shared_folder/Bacteriophage/MGV$ cat blast_protein/filtered_query.txt
MGV-GENOME-0192513_23 # 17497 # 18534 # 1 # ID=3086_23;partial=00... 172 8e-49
MGV-GENOME-0212292_17 # 12116 # 13165 # 1 # ID=320_17;partial=00;... 170 5e-48
MGV-GENOME-0172516_15 # 10958 # 12007 # -1 # ID=1136_15;partial=0... 170 5e-48
MGV-GENOME-0210072_17 # 11748 # 12797 # 1 # ID=55_17;partial=00;s... 170 5e-48
MGV-GENOME-0226859_40 # 31378 # 32427 # 1 # ID=795_40;partial=00;... 170 6e-48
MGV-GENOME-0187969_20 # 13343 # 14392 # -1 # ID=372_20;partial=00... 169 1e-47
MGV-GENOME-0183924_16 # 11035 # 12084 # -1 # ID=1677_16;partial=0... 169 2e-47
MGV-GENOME-0165826_10 # 10580 # 11620 # 1 # ID=2611_10;partial=0... 160 2e-47
```

```
(base) claireh@claire-virtualbox:/media/sf_shared_folder/Bacteriophage/MGV$ cat blast_protein/filtered_connector_dom
MGV-GENOME-0104393_9 # 6790 # 7770 # 1 # ID=513_9;partial=00;star... 254 6e-81
MGV-GENOME-0122635_24 # 15385 # 16266 # 1 # ID=1448_24;partial=00... 236 1e-74
MGV-GENOME-4313378_4 # 3458 # 4342 # -1 # ID=783_4;partial=00;sta... 233 3e-73
MGV-GENOME-0094600_4 # 3582 # 4466 # -1 # ID=892_4;partial=00;sta... 233 3e-73
MGV-GENOME-0094502_14 # 10883 # 11767 # 1 # ID=877_14;partial=00;... 233 3e-73
MGV-GENOME-0095706_5 # 3584 # 4468 # -1 # ID=1053_5;partial=00;st... 231 2e-72
MGV-GENOME-0099638_13 # 9158 # 10123 # 1 # ID=1418_13;partial=00;... 231 2e-72
```

- BLAST results using query sequence aren't quite as good as using connector domain but they're still good



## Query Seq. (cont)

### Annotation

MGV-GENOME-0192513_23	VPC-8016	Cytidine and deoxycytidylate deaminase zinc-binding region
MGV-GENOME-0212292_17	VPC-8016	Cytidine and deoxycytidylate deaminase zinc-binding region
SMGV-GENOME-0172516_15	VPC-8016	Cytidine and deoxycytidylate deaminase zinc-binding region
MGV-GENOME-0210072_17	VPC-8016	Cytidine and deoxycytidylate deaminase zinc-binding region
MGV-GENOME-0226859_40	VPC-8016	Cytidine and deoxycytidylate deaminase zinc-binding region
MGV-GENOME-0187969_20	VPC-8016	Cytidine and deoxycytidylate deaminase zinc-binding region
MGV-GENOME-0183924_16	VPC-8016	Cytidine and deoxycytidylate deaminase zinc-binding region
MGV-GENOME-0165836_19	VPC-8016	Cytidine and deoxycytidylate deaminase zinc-binding region
MGV-GENOME-0142829_6	VPC-8016	Cytidine and deoxycytidylate deaminase zinc-binding region

# Query Seq. (cont)

MGV-GENOME-0192513\_23, labelled as Cytidine and deoxycytidylate deaminase zinc-binding region

Created: Nov. 21, 2023 at 03:29

Summary

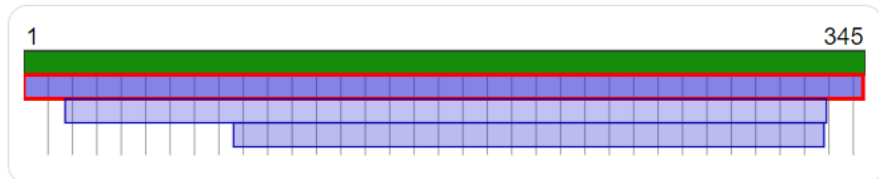
Templates 20

Models 3

Project Data ▾

## Model Results ?

Order by: GMQE ▾



Model 01 ▾

Structure  
Assessment



Oligo-State ?

Monomer

GMQE ?

0.80

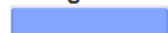
### Template

**A0A3Z9LJW2.1.A** Phage portal protein  
AlphaFold DB model of A0A3Z9LJW2\_SALER (gene:  
A0A3Z9LJW2\_SALER, organism: Salmonella enterica  
(Salmonella choleraesuis))

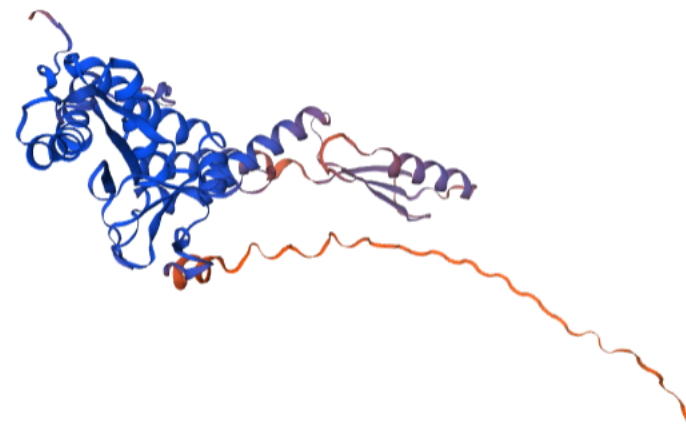
Seq Identity

79.94%

Coverage



Model-Template Alignment ▾



Cartoon ▲