Investigation of Usavi_25 and Usavi_26

AF thoughts: Usavi_25 NKF. Usavi_26 could possibly be TAC? See below.

Usavi 25:

MATIRARARIEIDEAALERESGEHLRAFHRSLTRRIANQSRVAVPVRTGNLGRTIGELPQVYTPFRVRGGVEATADYAAPVHEGSRPHAI RARNAQYLHFWWHGREMFRKSVWHPGTRARPFMRNSAQRVVTNDPRVRMT

Usavi 26:

MATFNADGKTVTDEQLSPPAEYLPDLTDSEREADYQAEQARIEAAIADAAAADDDDELPEVEVLTPPAGLATTDEQPGTEVATFERFNV EEVESWTYDKLEFKGDLLGIRLPTKAALAGFSLASSKYVSLGVKNDLTGLFIARHLSPESYGRVFSRLMDPDDVDYDVDAVGELFNAIV TAAVESDDE

This started with an investigation of Usavi_25, which has many <u>pham members that appear to have been mis-called as TAC</u>.

Forum: https://seaphages.org/forums/topic/4518/

"As of Dec 2022 I see many calls for a (single) tail assembly chaperone gene (Usavi_25) in B1 phages, so I assume sufficient supporting evidence has been found to make this call - but I'm not sure what it is?

HHpred results for Usavi_25: https://toolkit.tuebingen.mpg.de/jobs/Usavi_25 (Seem to indicate some kind of structural gene, maybe capsid-related, but not a TAC)"

Debbie's reply:

"I agree. While this has to be a structural protein, i agree it likely not a TAC. Interestingly, **Usavi_26 [downstream gene]** does have a Pfam hit of 94%, e-value of .95 (which is not great) to TAC. (That is not typical at all.)

This article might shed some light.

A conserved spiral structure for highly diverged phage tail assembly chaperones. **Pell** LG, Cumby N, Clark TE, Tuite A, Battaile KP, Edwards AM, Chirgadze NY, Davidson AR, Maxwell KL. J. Mol. Biol. 425, 2436-49, (2013). <u>View article PMID</u>: 23542344

Let me know what you find, debbie

PS, change what databases you are checking to see what else you can find."

Usavi 26 HHpred search on Pfam, SCOPE, CDD, PDB:

https://toolkit.tuebingen.mpg.de/jobs/Usavi 26



The Pell paper mentions another TAC that matches the top hit (**PF10963**) in the first HHpred search.

"By contrast, HHpred identified a number of putative TACs as possessing similarity to several different Pfam families with varying probabilities. For example, the Pfam group containing the phage T1 TAC (PF08748) could be connected to families containing the TACs from PBSX (PF08890) and HP1 (PF10963) as well as the TAC from phage PY54 and the 3FGX structure. "

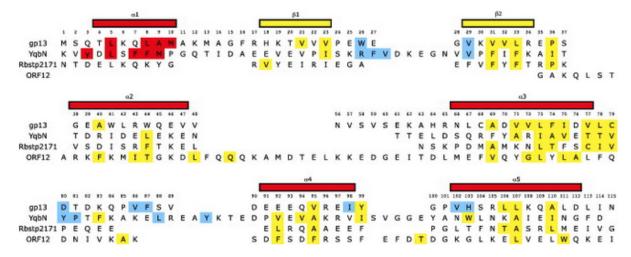
The InterPro record for the TAC from HP1 (PF10963) is here: https://www.ebi.ac.uk/interpro/protein/reviewed/P51730/

```
Template alignment | CDD
1. PF10963.11; Phage_TAC_10; Phage tail assembly chaperone
    Probability: 94.07%, E-value: 1.1, Score: 34.89, Aligned cols: 71, Identities: 10%, Similarity: 0.037, Template Neff: 6.3
                       0 ss pred
    Q Q_Usavi_26 99 KLEFKGDLLGIRLPTKAALAGFSLASSKYVSLGVKNDLTGLFIARHLSPESYGRVFSRLMDPDDVDVDVDVDAVGELFNAI 177 (187)
                  99 ~lev~G~~l~aR~P~p~Al~aLs~A~s~~~q~~~~lFv~~HLspeSy~rll~RMmDPDD~~f~~d~lgeL~~aI 177 (187)
    Q Consensus
                       .|++.|..+++ |+..+..-+.-.+++. .+.....||.+++.++++.|..-|-.|. .....+.|++..
                       ~l~v~~~~F~=pt~~yn~~in~~~~==-nky~~a~nfL~~~V~~e~k~~L~~~l~~PG---~~~i~~l~ee~
    T Consensus
                  2
                                                                                             72 (82)
    T PF10963.11
                       TLNVAGVDFTFI-RDNAAYDQMLNDIESN----NKVTPIKDYLLAIVAREQKEALLEIIHVPT---LAAQLAAKVNEVF
                       EEEECCEEEEe-cCHHHHHHHHHHccC----CCCHHHHHHHHHH
    T ss_pred
```

The region of Usavi_26 that aligns is from 99-177 of Usavi, and it aligns with 2-72 of PF10963. The Pell paper below, describes both helices and sheets throughout the entire gene (with the caveat that the Pell paper didn't actually determine the structure of PF10963).

Is it possible that we are seeing alignment with b2, a2, a3, a4, and a5 (see below)? The gaps between these motifs vary in size from Usavi_26 to the given sequences from the paper, but they also point out that these sequences [of TACs] are highly diverse.

From the Pell paper:



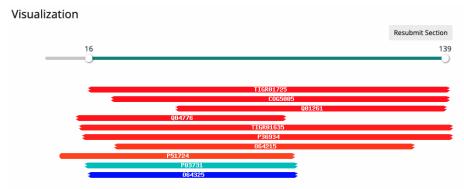
Download: Download high-res image (425KB)

Download: Download full-size image

Fig. 5. Structural alignment of the TACs. The <u>amino acid sequences</u> of gp13, YqbN, Rbstp2171, and ORF12 are aligned based on their structural overlays. The residue positions that are greater than 95% buried in the monomers are shown in yellow. The residues that are buried in the interprotomer interfaces in gp13 and YqbN are highlighted according to their positions in the helix (yellow) or the groove that the helix packs against in the adjacent <u>protomer</u> (red).

Usavi_26 HHpred search on UniPROT, COG, TIGRAM, SMART:

https://toolkit.tuebingen.mpg.de/jobs/Usavi 26b



The top TIGR hit is for HK97_gp10 and it is very strong. (Usavi_25 has a very similar and strong match to the same entry!)

However, the paper that Debbie linked states that the TAC in HK97 is gp13.