

# COMPUTATIONAL ANALYSIS OF UNKNOWN DNA SEQUENCES USING INTEGRATED BIOINFORMATICS TOOLS



## RAW DNA SEQUENCE

ATGGACATCGATTTCAAGAAGTTGGCTCCCTGGAAGTGGTTCAAGAACGAGCAGCAAGAGCAGCA  
GACCGTCTCTTCCCTGCCGGTGACGCGCAATGACCTGCCAGCGGCGAGCGGCCAGTCAGCCCG  
ATCCTGCAACTGCATCGGGAAATCAGCCGGCTGTTTCGATGACGCATTCCGGGGCTTCGGTTTTCCG  
GCGTTGAACATGCGCCGAGTGCCATCCGATTGGTCGGGCATGCTGAAGCCGGCCCTGGACATCCA  
GGAAACCGACAAGCAGTACAAGATTGCCCTGGAAGTGCCCGGTGTCGAGGAGAAGGACATCCAG  
ATCACTCTCGACAACGACGTGCTGATGGTCGTCGAGGAGCGCCAGGAACAGGAGAAGAAGG  
AAGGTGGCTTCCACCGTGTGGAGCGCTCTACGGCAGCTTTCAGCGTGCCTTGAACCTGCCTGAC  
GACGCCAACCGAGATTTCGATCAAGCATCGTTCAAGAACGGGGTGTCTACGGTACGATCGACAA  
GCGCGAGGTCAGCGCGCCGAAGCAGGGACGCTCGATCCCGATCAACGGCTGA

- **GENE:** hsp20 (Heat Shock Protein 20) / ibpB
- **ORGANISM:** Klebsiella pneumoniae
- **LOCATION:** Plasmid (Mobile genetic element)
- **PROTEIN:** Small Heat Shock Protein (sHSP)
- **Bp/Amino Acids:** 570 bp/189 aa
- **ACCESSION:** CP146304.1 (First hit from BLAST)
- **Total GC Content**=59% with no tandem repeats



## DNA>PROTEIN TRANSLATION FILE

5'3' Frame 1
MDIDFKKLAPWNWFKNEQQEQTVSSLPVQRNDLPAASGVPSPILQLHREIDRLFDDAIFRGFPALNMPQWSDWSGNLKPALDIQ ETDKQYKIALEVPGEVEKDIIQTLNDVLMVGRKEKQEQKEGGFHRVERSYGSFQRALNLPDDANQDSIKASFKNGLVTVTIDKR EVSAPKQGRSIPING
5'3' Frame 2
WTSTSRSLWPGTGSRTSSKSSRPSLPCRCASMTCCRRAGQSARSNCIGKSTGCSMTHSGSVFR - TCRSGHPTGRAC - SRPWTSR KPTSTSLRPMKCPVSRRTSRSLSTTTC - WCVARSNRRRRKVVASTVWSAPTAASFVP - TCLTTPTRISKHSRTGCSRSTSA RSARRSDARSSTA
5'3' Frame 3
GHRFQEVGSLLELVQEARAADRLLFPAGAAQ - PASGERASQPDPTASGNRPVAVR - RIPGLRFSGVEHAVAIRLVGHAEAGPGHPG NRQAQVQCPGSGARCGEHPDHSRRRRADGAWREAPGTGEEGRWLPCCGALLRQLSACLEPA - RRQPGFDQSIQVQERGAHGHDRQAR GQRAEAGTLDPDQRL

## DNA ANALYSIS GENE LEVEL

### NCBI BLASTn

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download

Select columns

Show100

☐ select all

3 sequences selected

GenBank

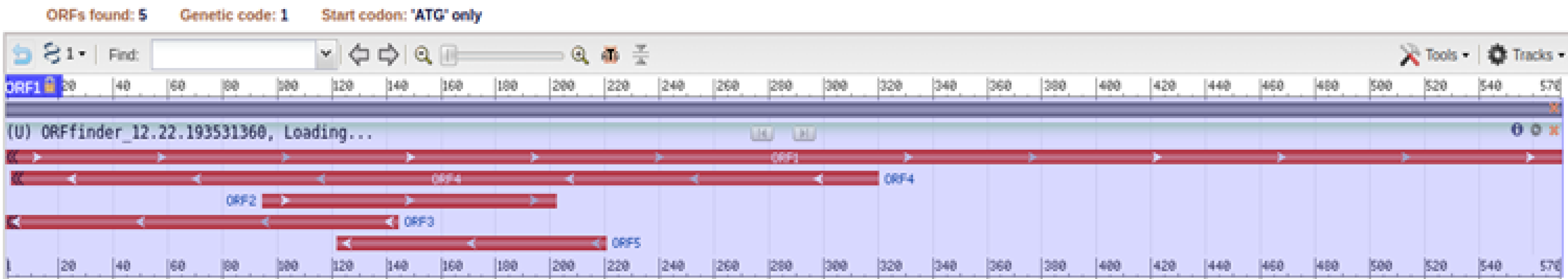
Graphics

Distance tree of results

MSA Viewer

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	<a href="#">Klebsiella pneumoniae strain AMC116165 plasmid unnamed1 . complete sequence</a>	<a href="#">Klebsiella pneu...</a>	1053	1053	100%	0.0	100.00%	172716	<a href="#">CP185177.1</a>
<input checked="" type="checkbox"/>	<a href="#">Klebsiella pneumoniae subsp. pneumoniae strain L2304797 plasmid p1-L2304797 . complete sequence</a>	<a href="#">Klebsiella pneu...</a>	1053	1053	100%	0.0	100.00%	207532	<a href="#">CP165951.1</a>
<input checked="" type="checkbox"/>	<a href="#">Klebsiella pneumoniae subsp. pneumoniae strain M7233093 plasmid p1-M7233093 . complete sequence</a>	<a href="#">Klebsiella pneu...</a>	1053	1053	100%	0.0	100.00%	370711	<a href="#">CP165904.1</a>

## ORF FINDER



## PROTEIN ANALYSIS (STRUCTURE AND PROPERTIES)



## DOMAIN ARCHITECTURE and FUNCTION PREDICTION

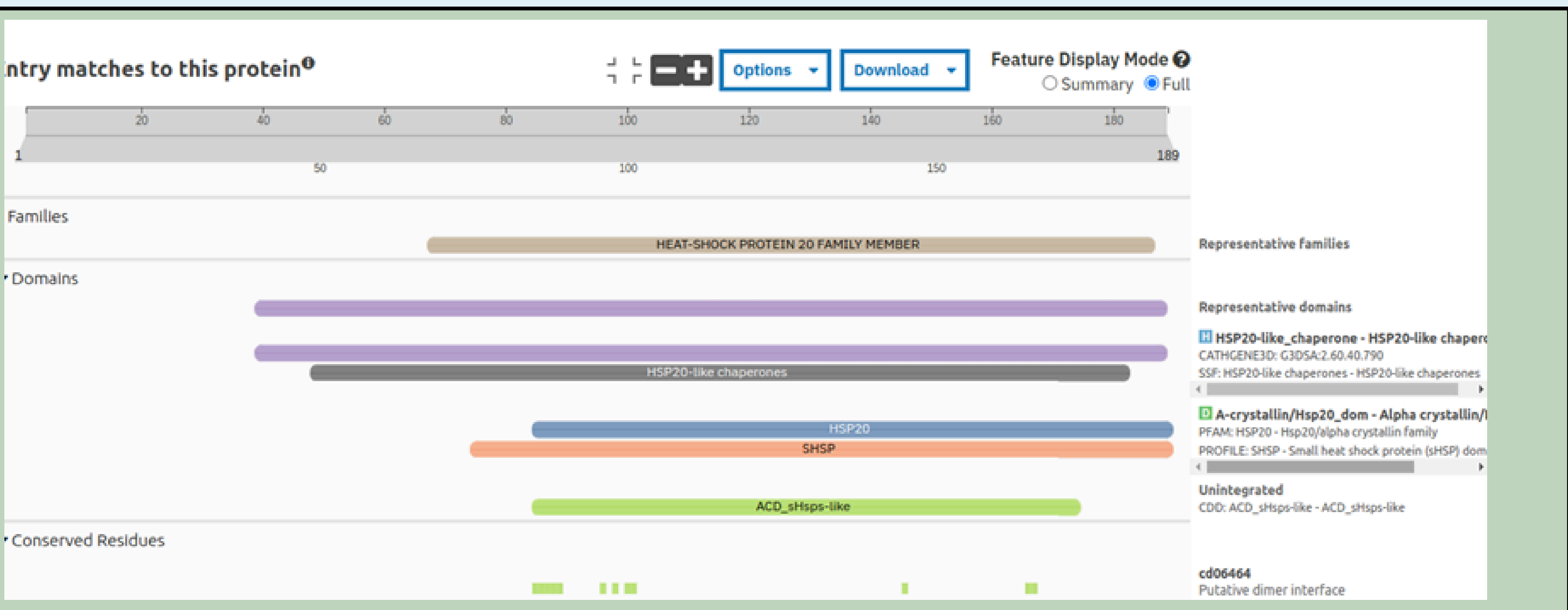
### FUNCTION PREDICTION

#### Chaperone Activity:

Prevents protein aggregation and misfolding during stress.

#### Environmental Fitness:

Protects against heat, oxidative stress, and pH fluctuations.



## PHYLOGENETIC ANALYSIS

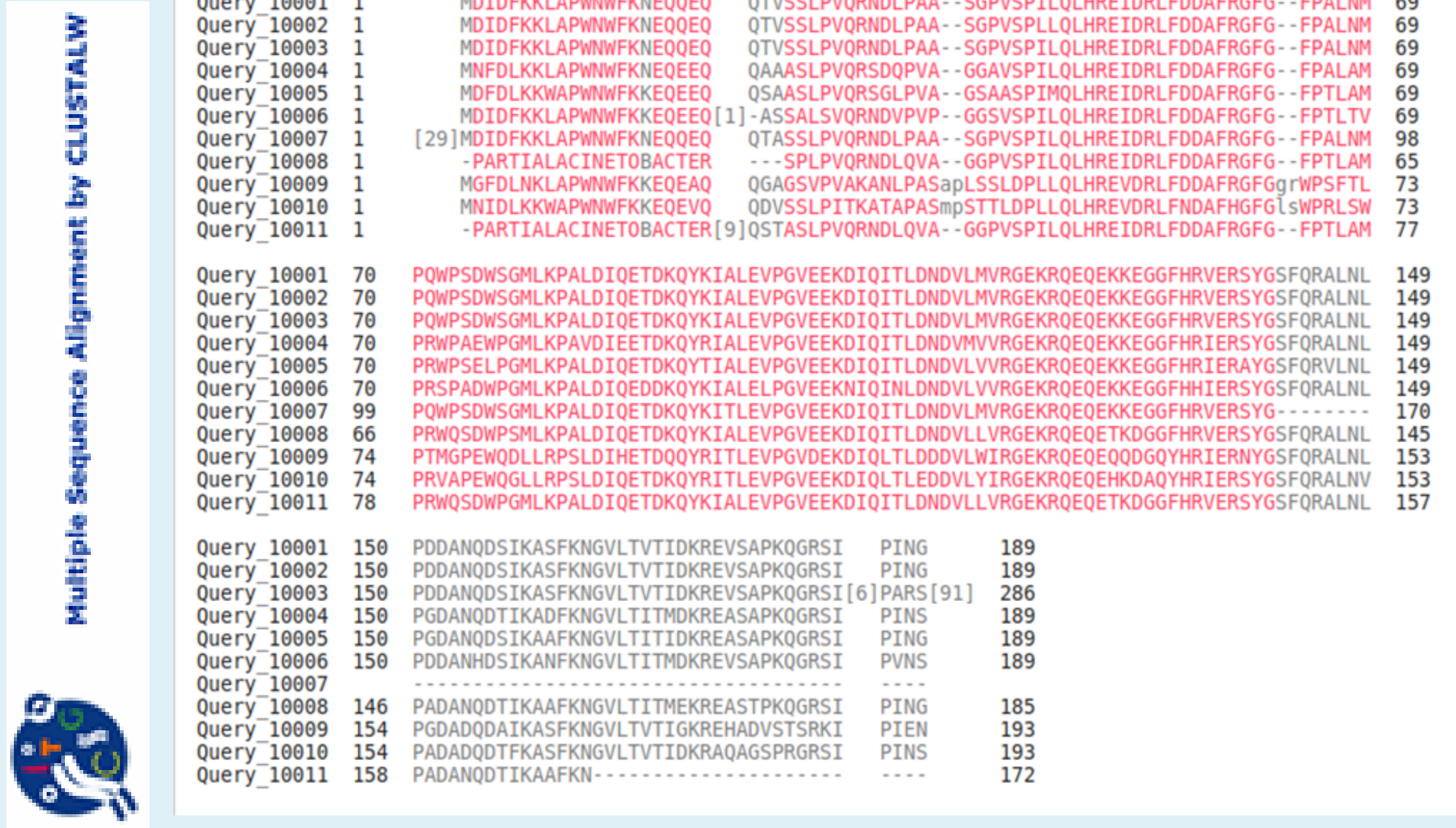
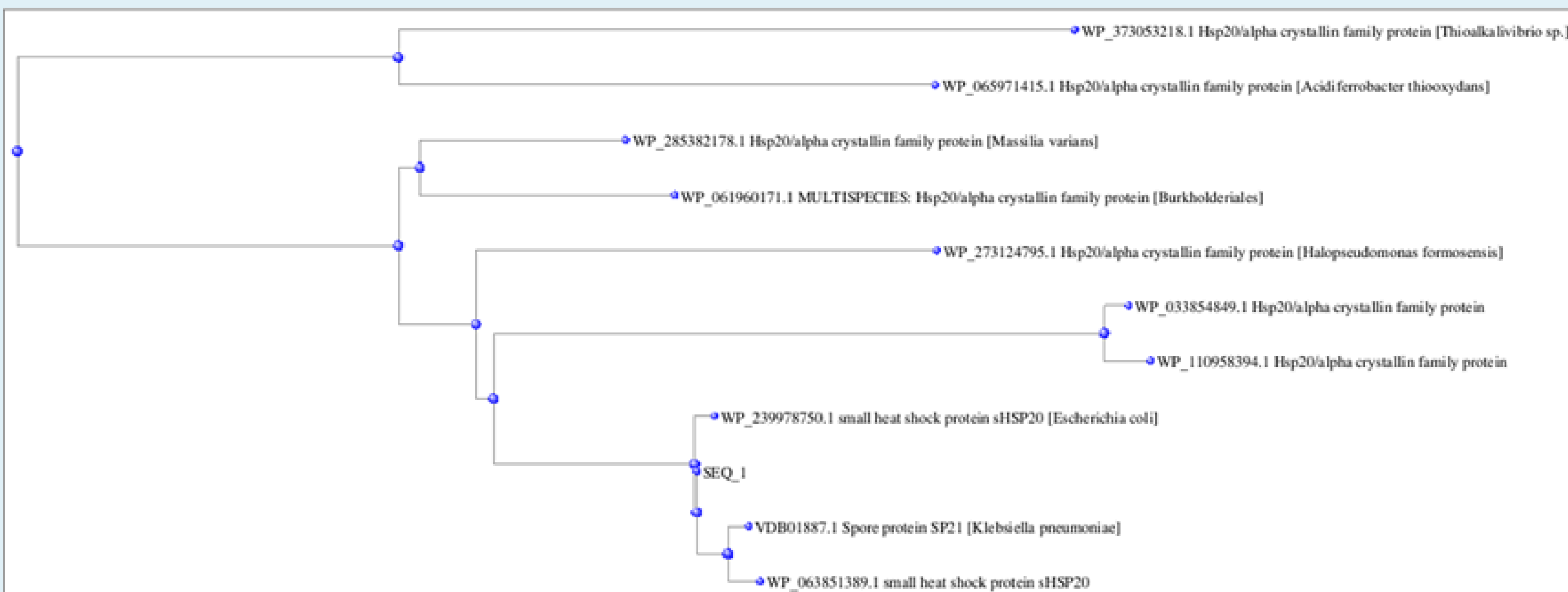
COBALT

Constraint-based Multiple Alignment Tool

### Multiple Sequence Alignment

### PHYLOGENETIC TREE

The HSP20 gene is a highly conserved, stable component of the Klebsiella pneumoniae plasmid genome.



## SUBCELLULAR LOCALIZATION PREDICTION

### TargetP - 2.0

Protein type	Other	Signal peptide	Mitochondrial transfer peptide
Likelihood	0.9999	0	0

### DeepLoc - 2.0

Protein sequence  
Predicted localizations: Cytoplasm, Lysosome/Vacuole  
Predicted signals:

Localization	Cytoplasm	Nucleus	Extracellular	Cell membrane	Mitochondrion	Plastid	Endoplasmic reticulum	Lysosome/Vacuole	Golgi apparatus	Peroxisome
Probability	0.5280	0.4833	0.0956	0.0661	0.2767	0.0335	0.1775	0.7258	0.2205	0.2465

**TargetP 2.0:** 0.9999 "Other" score confirms no signal peptide; protein is not secreted.

**DeepLoc 2.0:** High probability of cytoplasmic localization based on amino acid features.

**Justification:** As an HSP20 chaperone, it functions in the cytosol to stabilize soluble proteins.

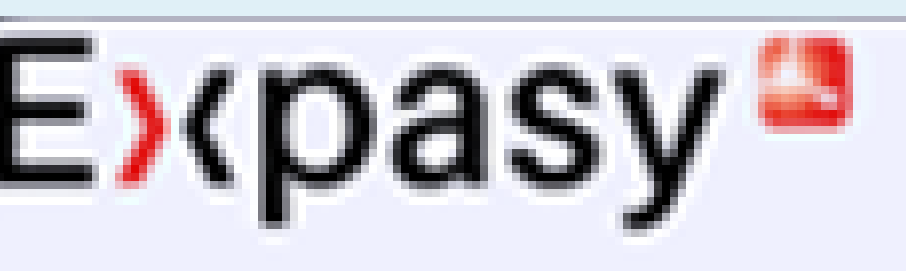
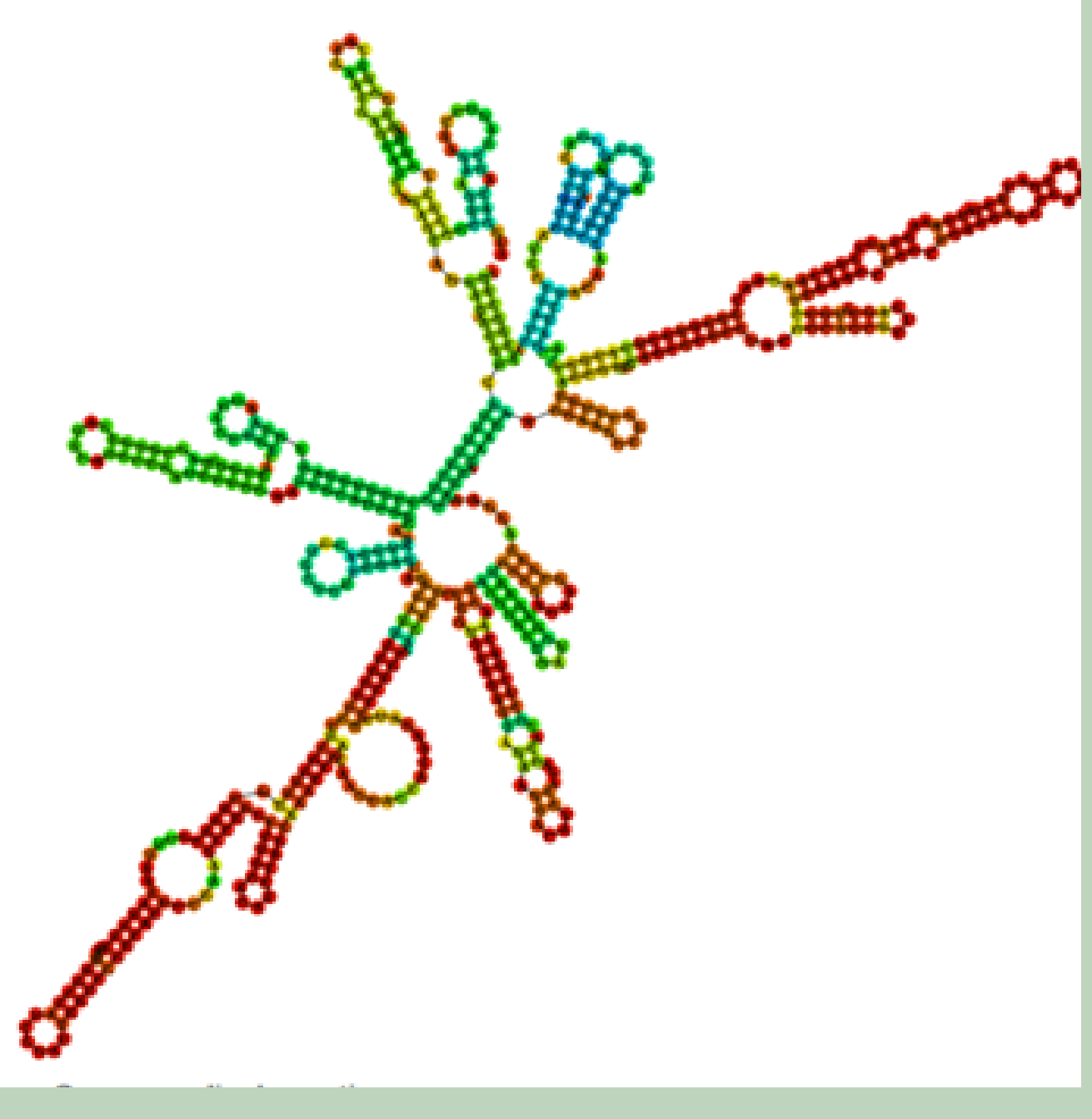
## RNA ANALYSIS

### ViennaRNA Web Services

Institute for Theoretical Chemistry

#### RNA FOLD (mRNA Secondary structure)

The mRNA secondary structure was predicted using the **Minimum Free Energy (MFE)** model. The structure contains multiple stem-loops and junctions, indicating a highly folded RNA. A very low free energy ( **$\Delta G = -217.30$  kcal/mol**) reflects strong thermodynamic stability. This stability protects the mRNA from degradation and can influence translation by controlling ribosome access.



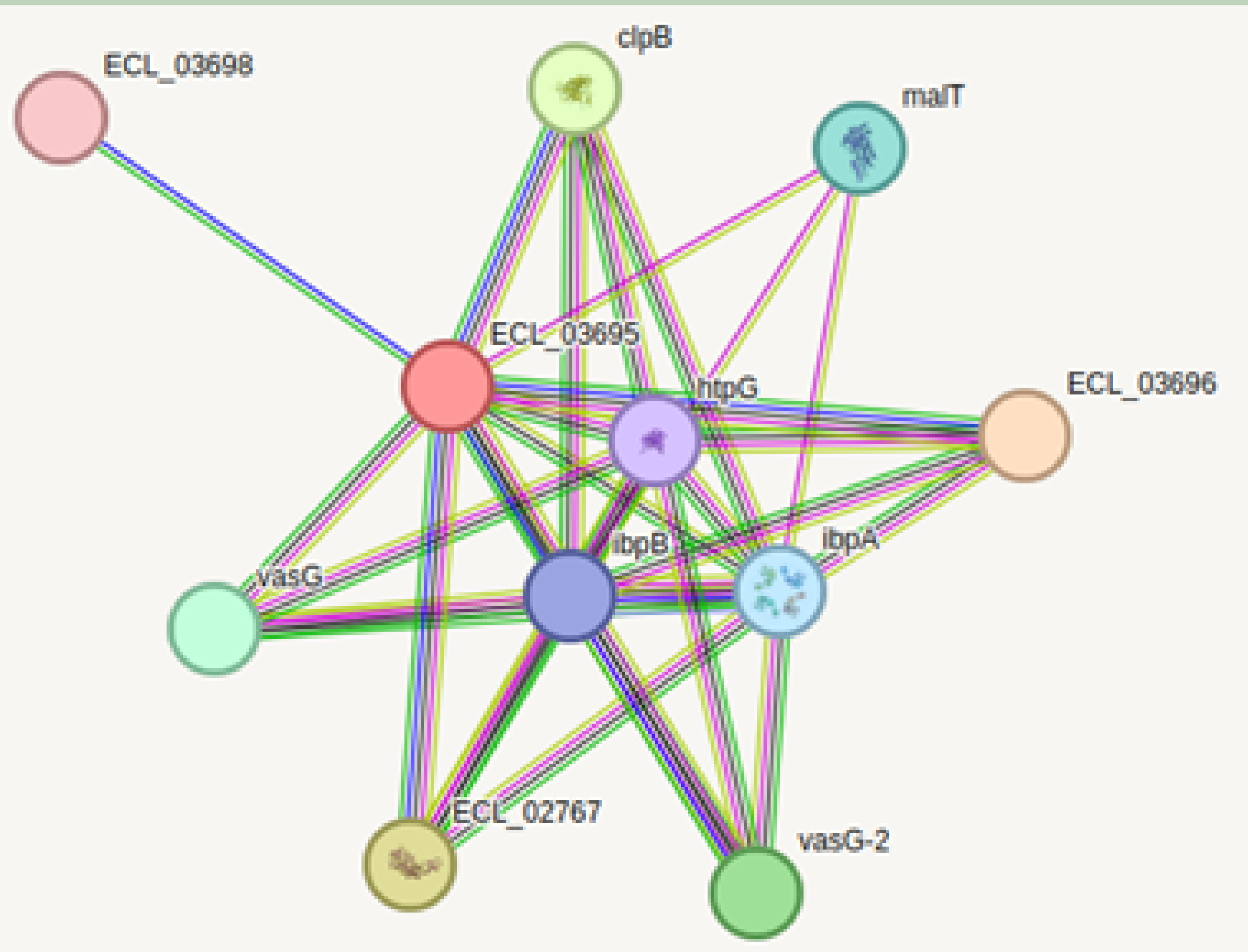
### PROT PARAM

### TABLE OF PROTIEN PROPERTIES

Number of amino acids	189
Theoretical pI	5.32
Molecular weight	21506.25
Instability index	48.70
Aliphatic index	76.35
Grand average of hydropathicity (GRAVY)	0.717

## 4

## PROTIEN –PROTIEN INTERACTION ANALYSIS



### BIOLOGICAL SIGNIFICANCE

**Network Role:** Operates within a highly interconnected multichaperone system (including ibpA, ibpB, htpG, and clpB).

**Synergy:** Sequesters misfolded proteins alongside ibpA/B for downstream refolding by ATP-dependent chaperones like clpB.

**Significance:** Provides a robust defense against aggregation, directly enhancing bacterial pathogenicity and antibiotic survival