Structure and function of phage AR9 RNA polymerase

Petr Leiman

University of Texas Medical Branch

Galveston, Texas

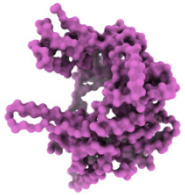
Transcription

DNA-dependent RNA polymerases (RNAPs)

‘right-hand-shaped’ RNAPs

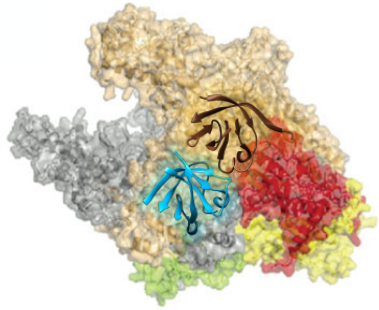
‘two-barrel’ RNAPs

Mitochondria
Chloroplasts
Bacteriophages



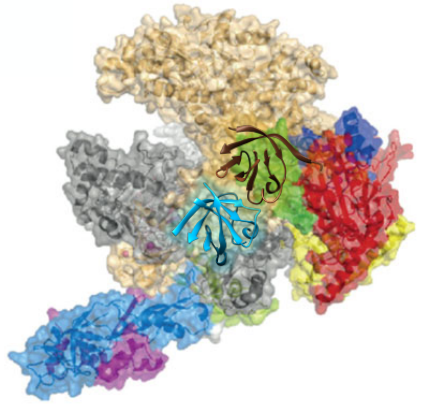
T7 phage
RNAP
PDB 4RNP

Bacteria



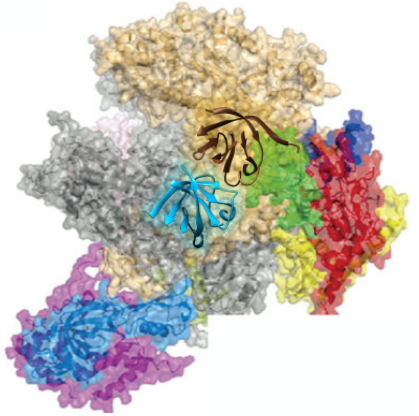
Thermus aquaticus
RNAP
PDB 1I6V

Archaea



Sulfolobus solfataricus
RNAP
PDB 2PMZ

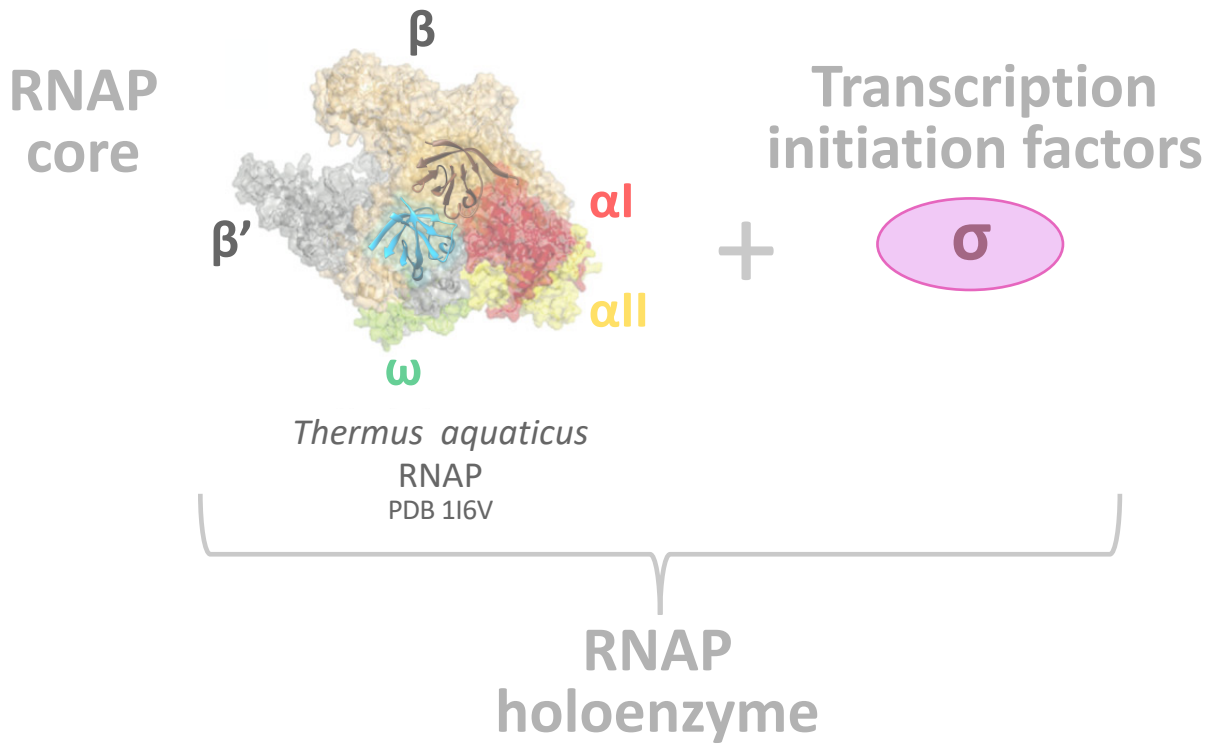
Eukarya



Saccharomyces cerevisiae
RNAP II
PDB 1NT9

Werner F., Trends Microbiology, 2008

A canonical RNA polymerase



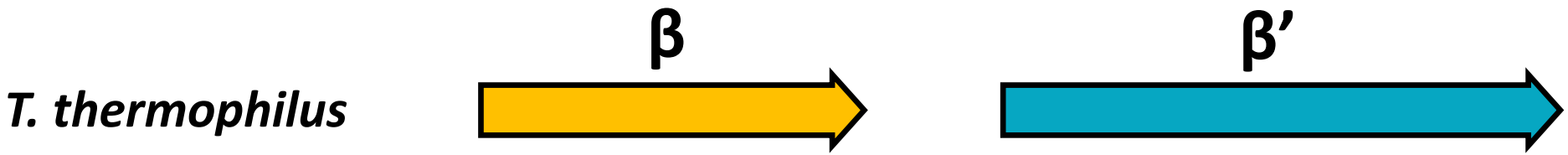
Some bacteriophages encode homologs of β/β' subunits but lack homologs of other subunits (α, ω, σ):

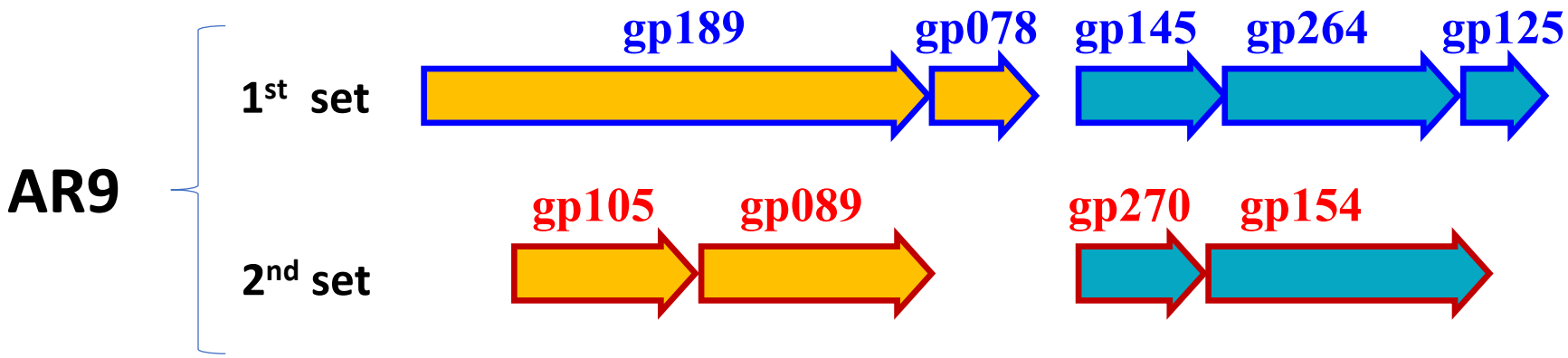
- **new mechanisms of assembly**
- **novel promoter recognition strategies**

Non-canonical RNAPs

Bacteriophage AR9

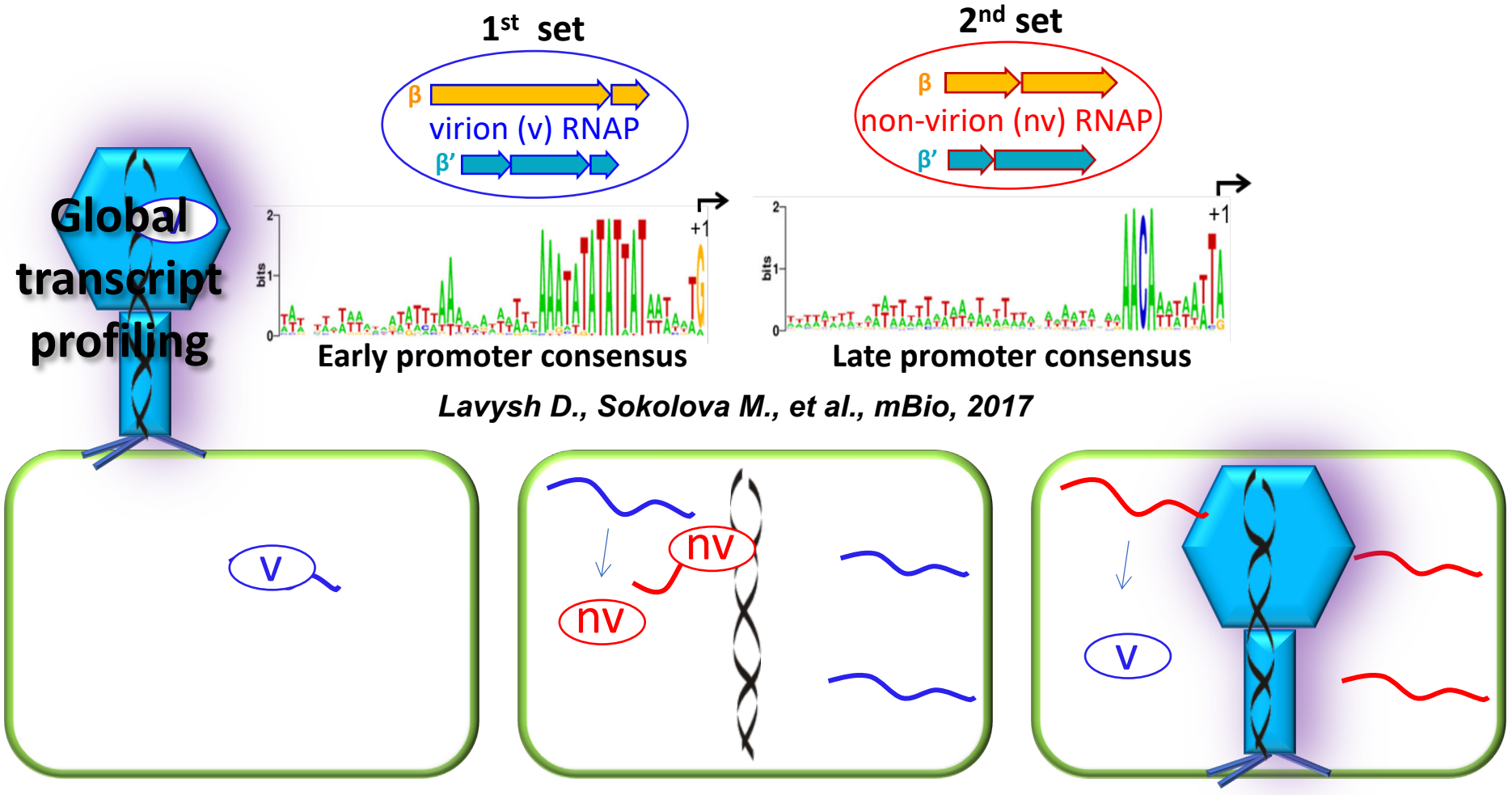
- **infects:** ***Bacillus subtilis***
- **genome size:** **251 kbs**
- **genome composition:** **A, U, G, C (T → U)**





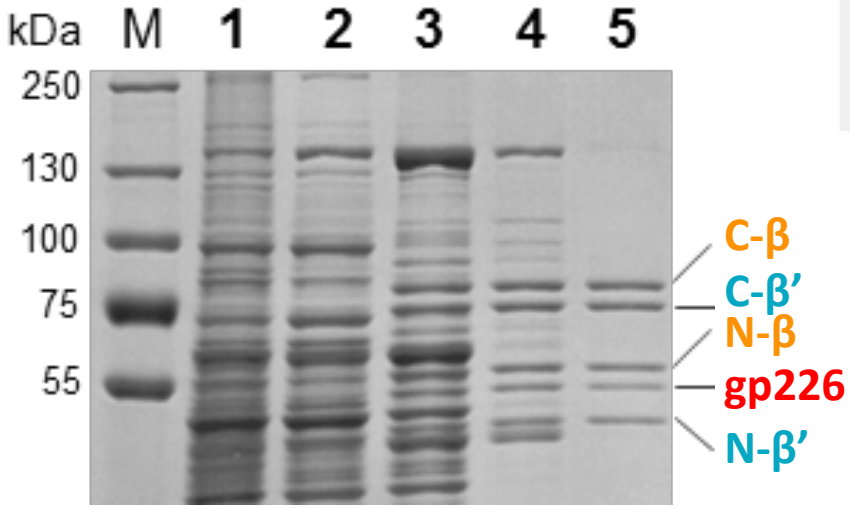
Lavysh D., Sokolova M., et al., Virology, 2016

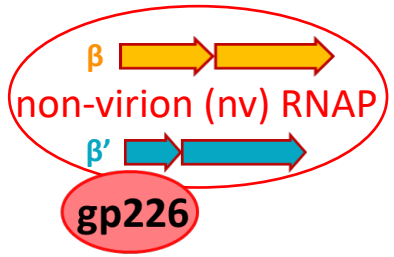
Transcription strategy of the AR9 phage



Purification of the AR9 nvRNAP from *B. subtilis* cells infected with AR9

1 Cleared cell lysate
↓
2 Polymin P fractionation
↓
3 Heparin-Sepharose
↓
4 Superdex 200
↓
5 MonoQ





Descriptions | Graphic Summary | Alignments | Taxonomy

Sequences producing significant alignments Download Manage Columns Show 100

select all *4 sequences selected* GenPept Graphics Distance tree of results Multiple alignment

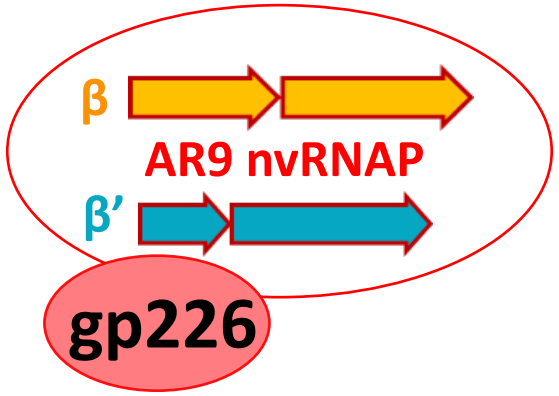
Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
DNA-directed RNA polymerase subunit [Bacillus phage AR9]	915	915	100%	0.0	100.00%	YP_009283130.1
hypothetical protein DA469_21760 [Bacillus subtilis]	897	897	98%	0.0	99.56%	PTU25816.1
hypothetical protein Bp8pS_281 [Bacillus phage vB_BpuM-BpSp]	495	495	95%	1e-169	58.39%	ALN97960.1
g245 [Yersinia phage phiR1-37]	89.0	89.0	46%	6e-15	28.96%	YP_004934479.1

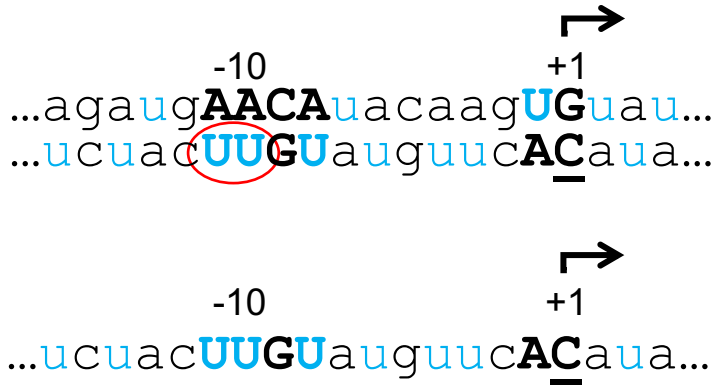
Predicted subunits

+

Protein with unknown function

Properties of AR9 nvRNAP

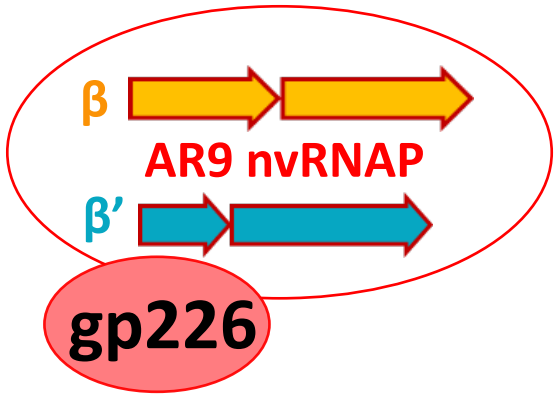




AR9 nvRNAP

- **consists of the homologs of β/β' subunits of bacterial RNAP and a protein with unknown function**
- **transcribes from AR9 late promoters *in vitro***
- **requires the presence of two uracils at the -11, -10 positions**
- **recognizes the promoter on the template DNA strand**
- **promoter-specifically transcribes ssDNA**

The protein with previously unknown function (gp226) is a promoter specificity subunit.



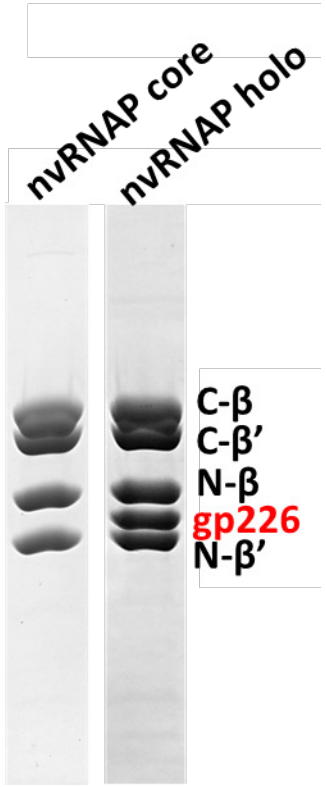
Structure of the enzyme?

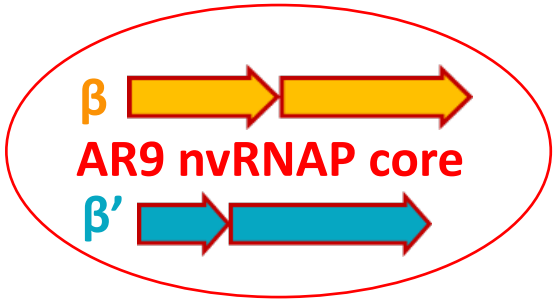
4S particle: ~2300 residues

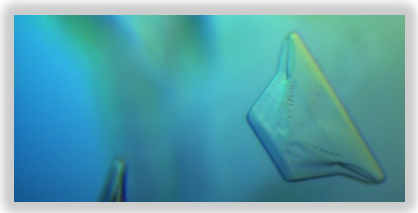
5S particle: ~2800 residues

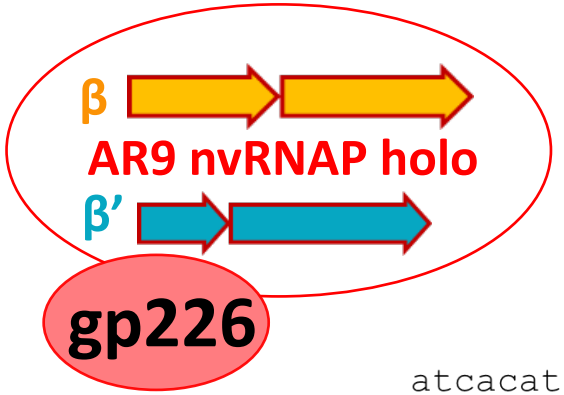
Crystallization of AR9 nvRNAP

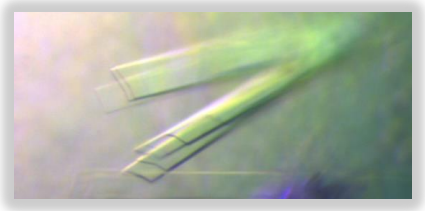
Recombinant:



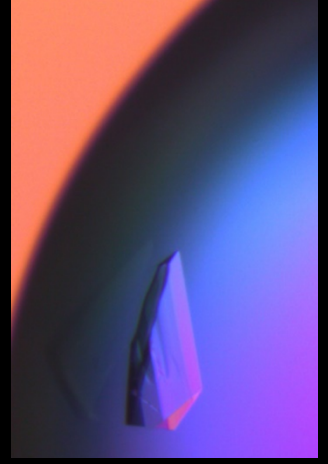


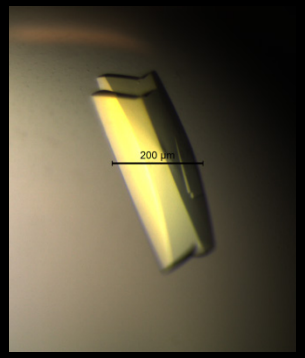


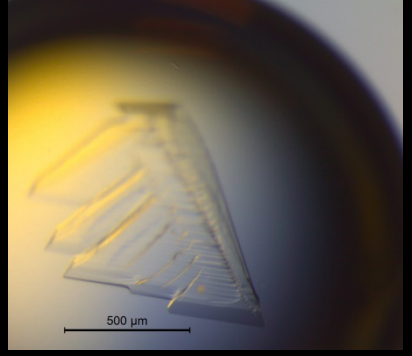


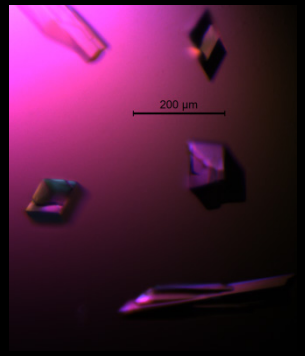


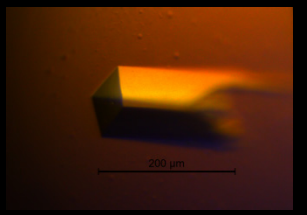
atcacatattggag
gttauUUGUuatataATatagtgtataacctc

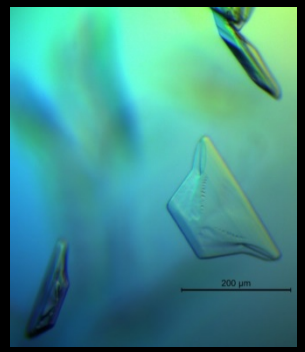


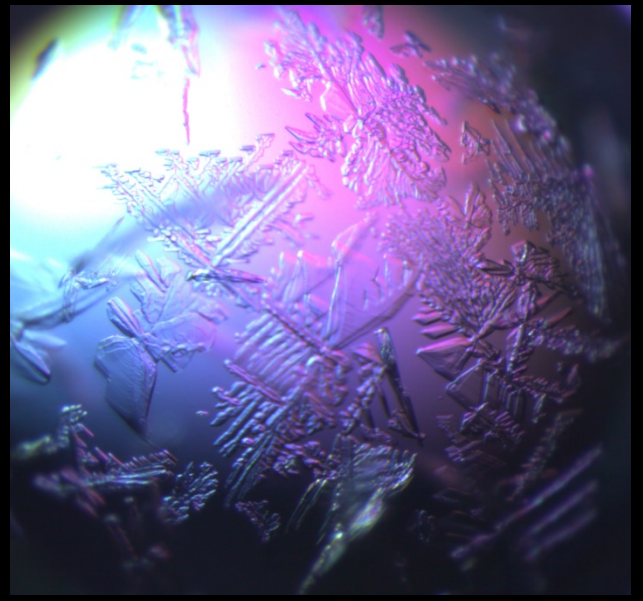

200 µm

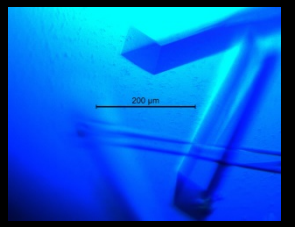

500 µm

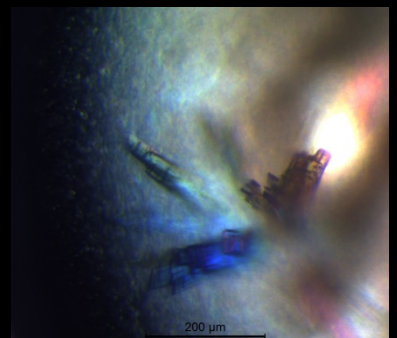

200 µm

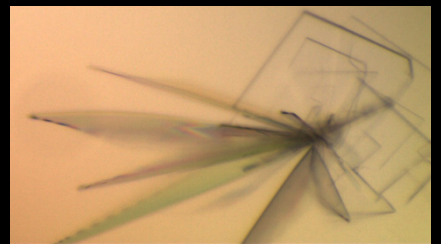

200 µm

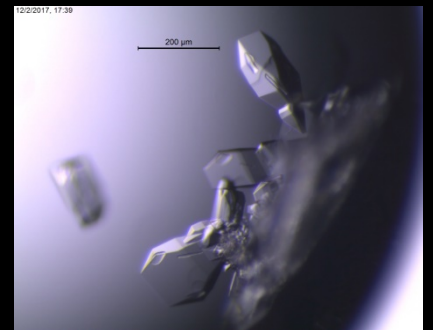

200 µm




200 µm


200 µm



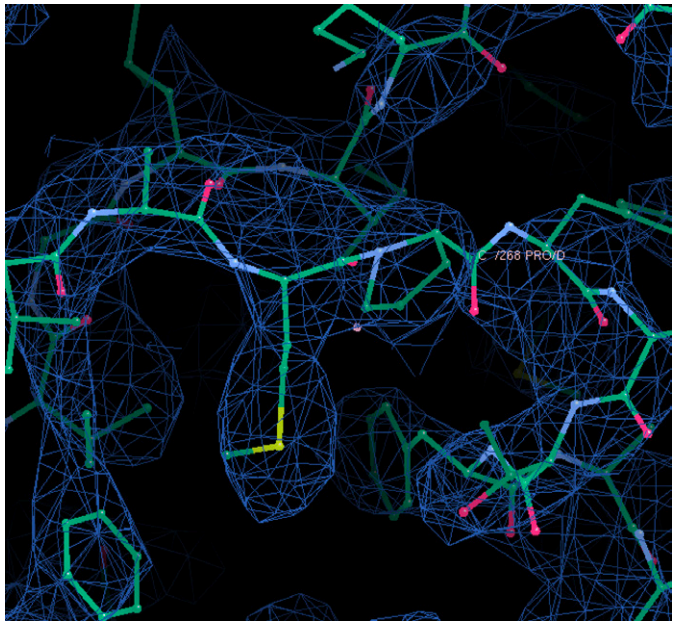

12/2/2017, 17:39
200 µm

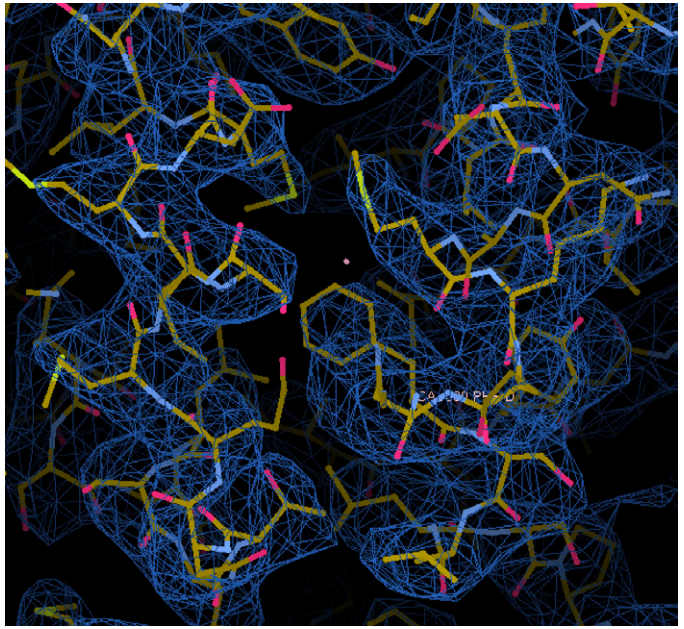
Path to atomic structure

X-ray crystallography

Total number of crystals shot: 300+ (of which most were heavy atoms soaks).
Total number of datasets collected and analyzed: 50+
Total number of days and/or nights spent collecting X-ray data: ~14.
Total time spent on analyzing diffraction data and phasing: months.

Best datasets: Several natives (3.2 Å), Ta derivative (6 Å), Hg derivative (very large unit cell, 3.8 Å).
"Interpretable" density obtained by SAD+MR+NCS+multicrystal averaging (3.8 Å) of several 4S datasets. Then switched to cryoEM.





CryoEM

Total data sets collected (different Krios-K2/K3 Summit):

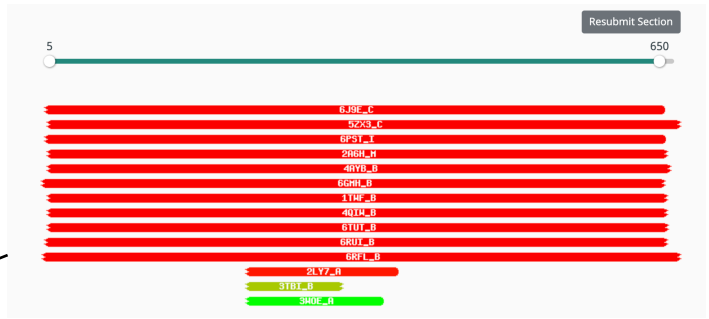
- 4S on thin carbon (no protein in the holes at all otherwise), terrible preferred orientation;
- 5S on thin carbon with 30-degree tilt, some preferred orientation;
- 5S + DNA – no tilt and no support film, little preferred orientation.

CryoSPARC gave interpretable map at 3.8 Å resolution. Relion map was not easily interpretable.

Helpful bioinformatics - HHpred

gp105

gp89



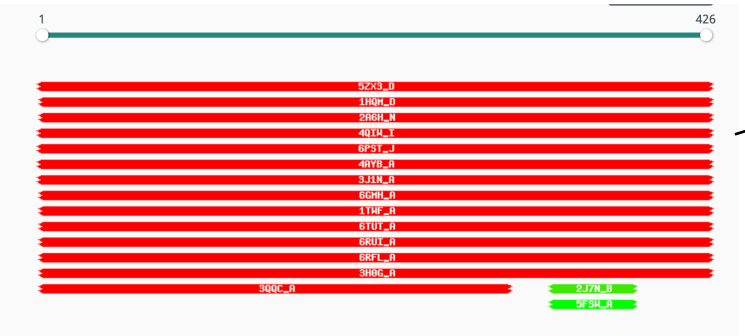
β

AR9 nvRNAP holo

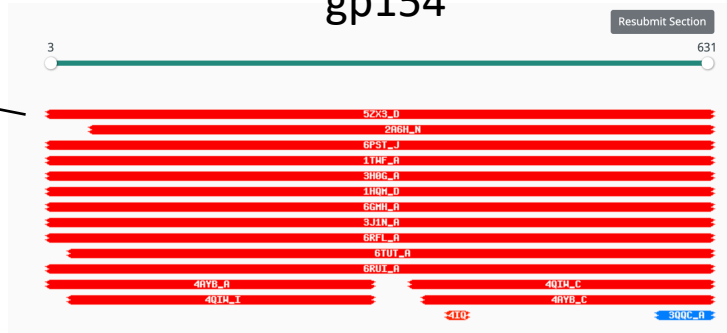
β'

gp226

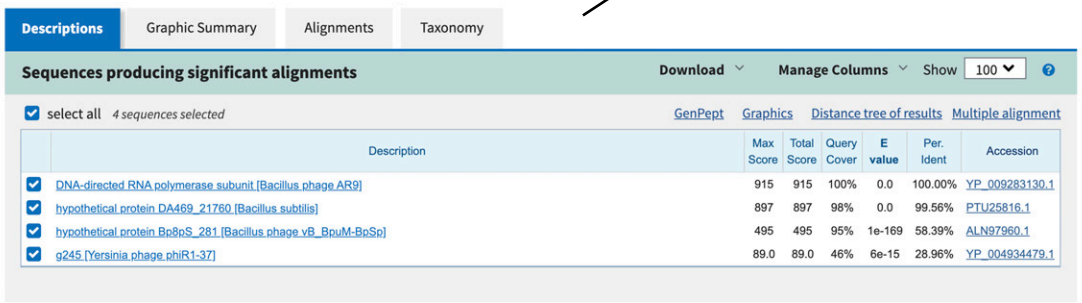
gp270



gp154



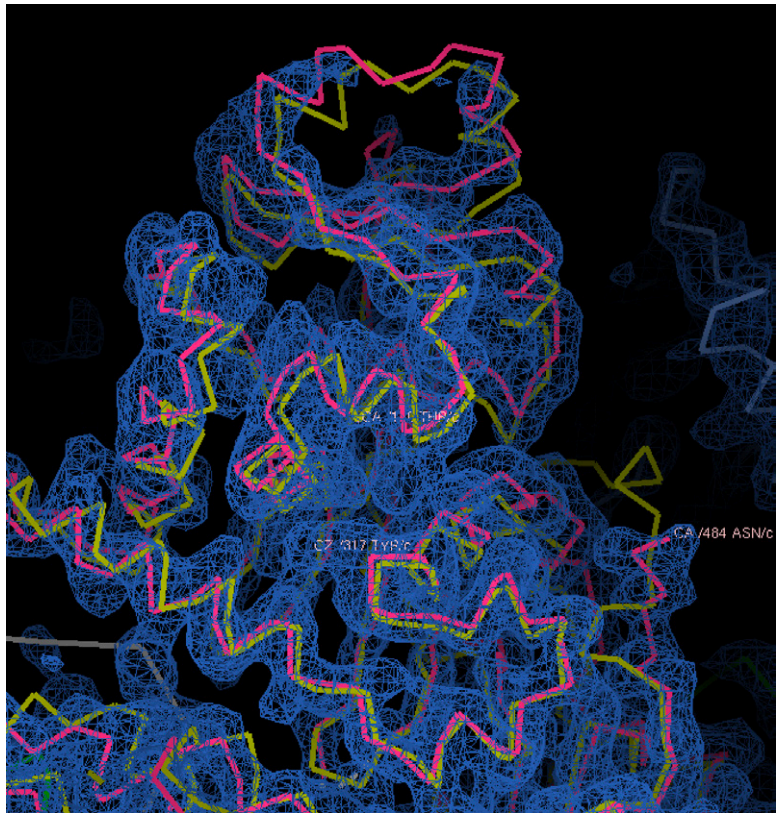
???



For highly similar targets (gp89, gp270 and gp154): Modeller models after much trimming (30% left) and adjustments were used in the MR-SAD-NCS averaging procedure for X-ray data.

Astonishing quality of group 427 models

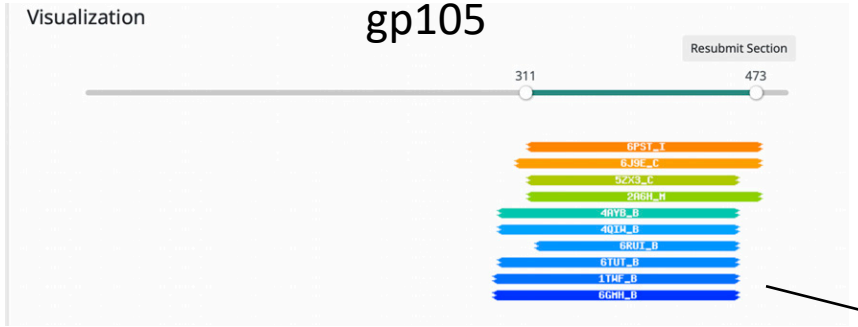
gp105



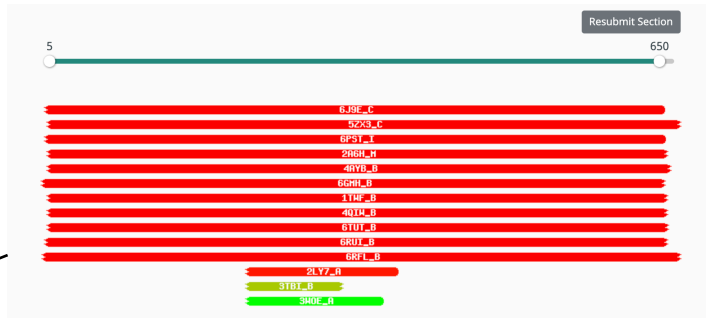
Group 427 model (yellow) is fitted into the cryoEM map.

Helpful bioinformatics

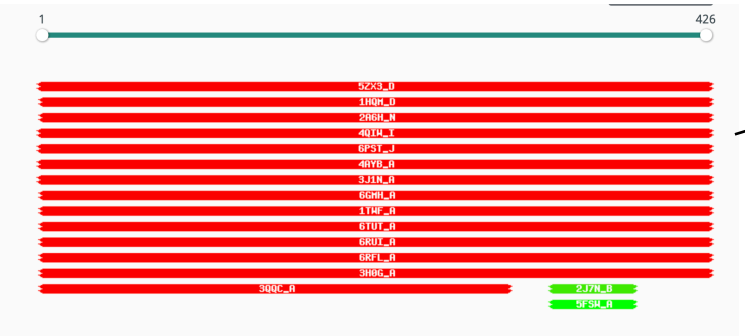
gp105



gp89



gp270



β

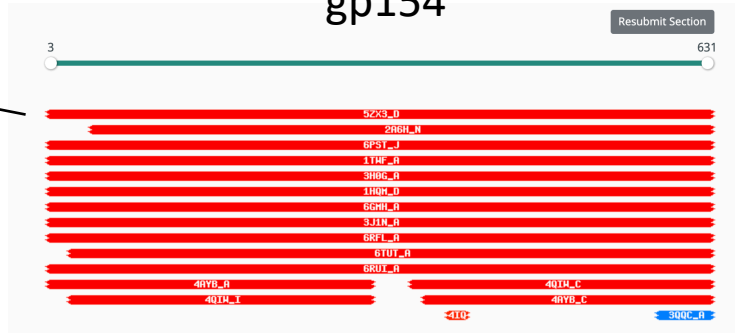
AR9 nvRNAP holo

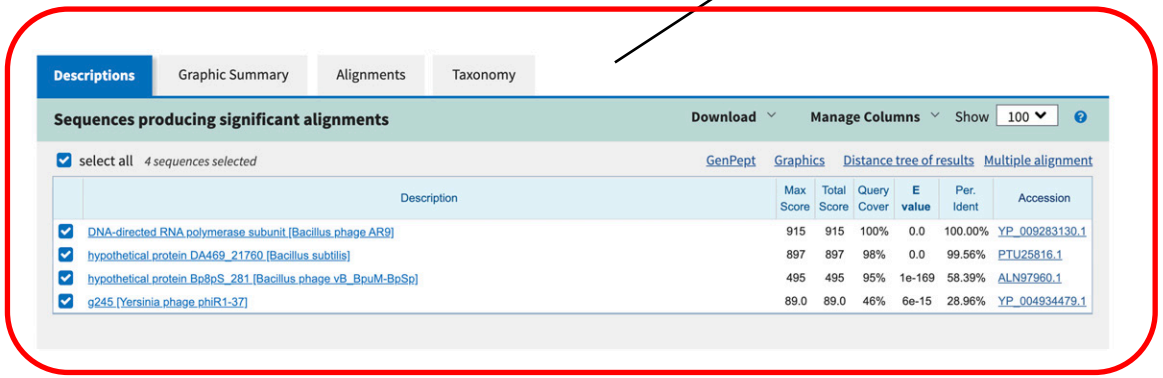
β'

gp226

???

gp154

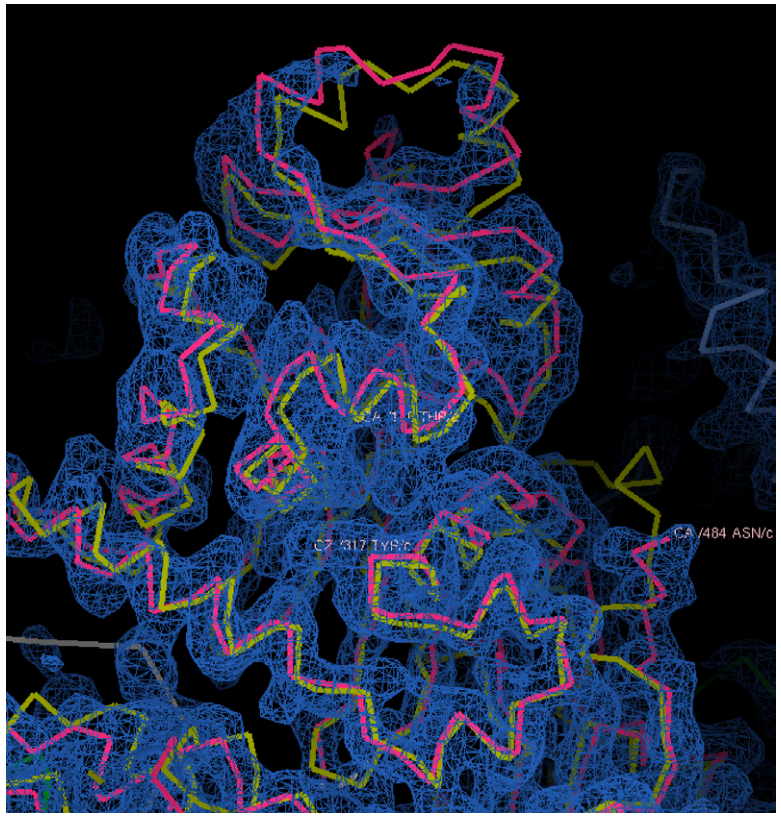




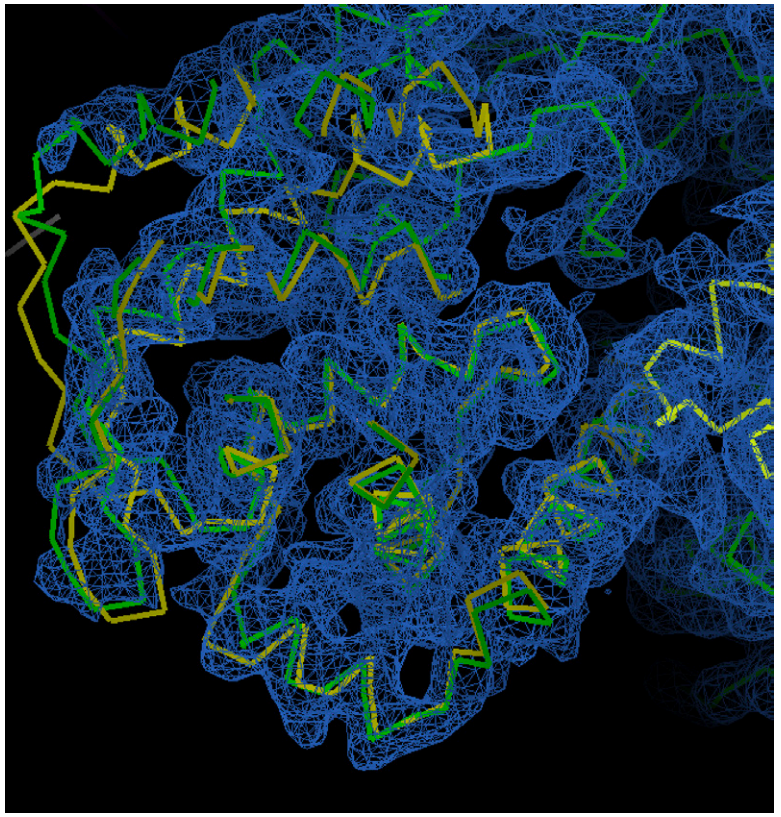
For highly similar targets (gp89, gp270 and gp154): Modeller models after much trimming (30% left) and adjustments were used in the MR-SAD-NCS averaging procedure for X-ray data.

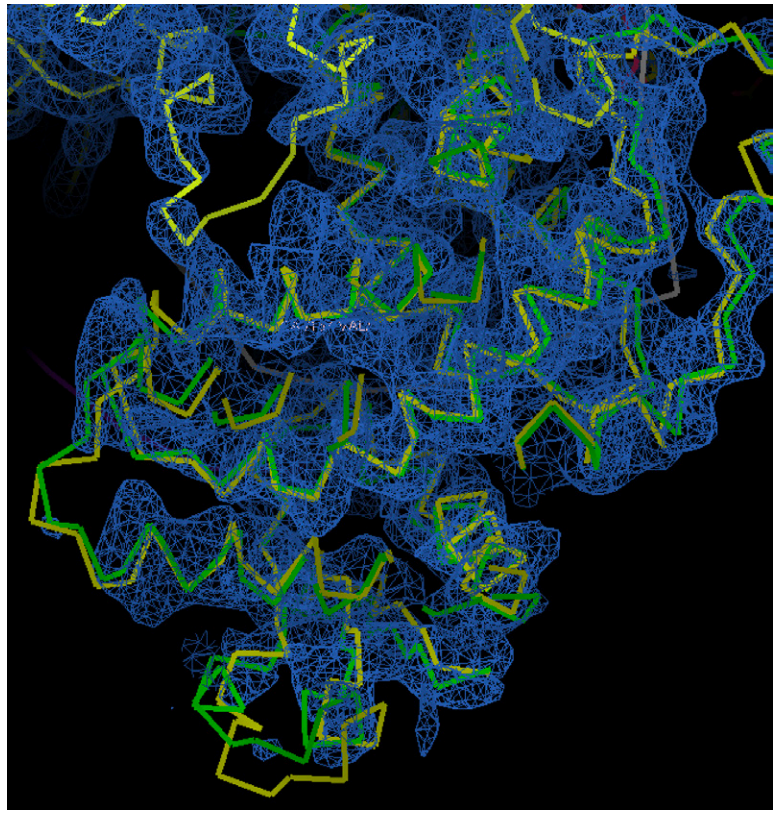
Astonishing quality of group 427 models

gp105



gp226

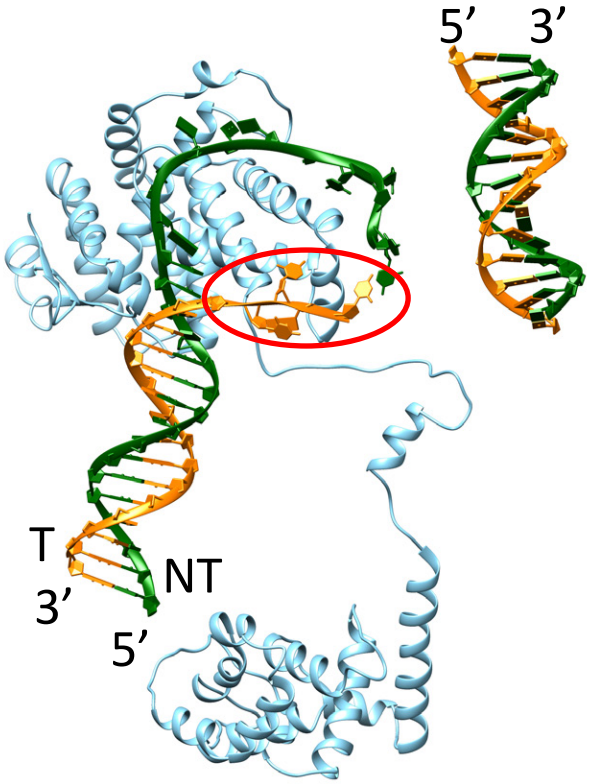




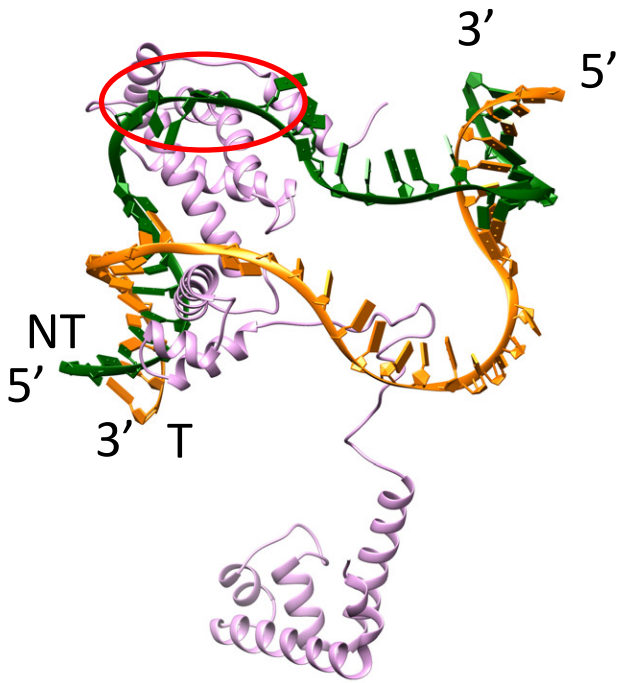
Group 427 model (yellow) is fitted into the cryoEM map.

Promoter recognition by gp226

gp226

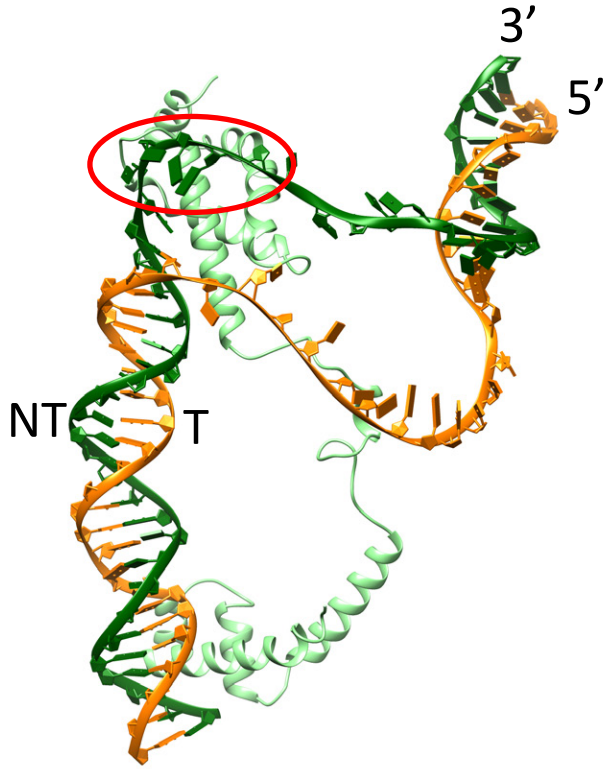


E. coli Sigma S



PDB: 5ipm

E. coli Sigma E



PDB: 6jbq

Acknowledgements

Konstantin Severinov,
Skoltech, Rutgers

Arina Drobysheva,
Skoltech

Maria Sokolova,
Skoltech, UTMB

Alec Fraser,
UTMB

Sergei Borukhov,
Rowan University

Group 427 (Alphafold2)!!!

U.S. Department of Energy's Argonne National Laboratory, Chicago, USA

Berkeley Center for Structural Biology, Berkeley, USA