Multi-Template Comparative Modeling with RosettaCM

Greg Sliwoski

gregory.r.sliwoski@vanderbilt.edu

Rosetta Workshop 2014

Target: bovine rhodopsin

- PDB ID: 1U19
- Class A G-protein coupled receptor
- No high identity templates

- 7 transmembrane helices
- 3 extracellular loops, 3 intracellular loops
- Highly conserved GPCR residues

Low identity templates

template	PDB ID	score
β2 adrenergic receptor	2RH1	16
β1 adrenergic receptor	2VT4	18
A2A adenosine receptor	3EML	20
CXCR4 chemokine receptor	3ODU	16
Dopamine D3 receptor	3PBL	26
Histamine H1 receptor	3RZE	18
M2 muscarinic acetylcholine receptor	3UON	21
Sphingosine 1-phosphate receptor 1	3V2W	19
M3 muscarinic acetylcholine receptor	4DAJ	22
κ-opioid receptor	4DJH	18
μ-opioid receptor	4DKL	21
N/OFQ opioid receptor	4EA3	21
δ-opioid receptor	4EJ4	19
5-HT1B receptor	4IAR	20
5-HT2B receptor	4IB4	20

Class A GPCR's

Highly conserved GPCR residues

Similar fold profiles

Low sequence identity (especially loops)

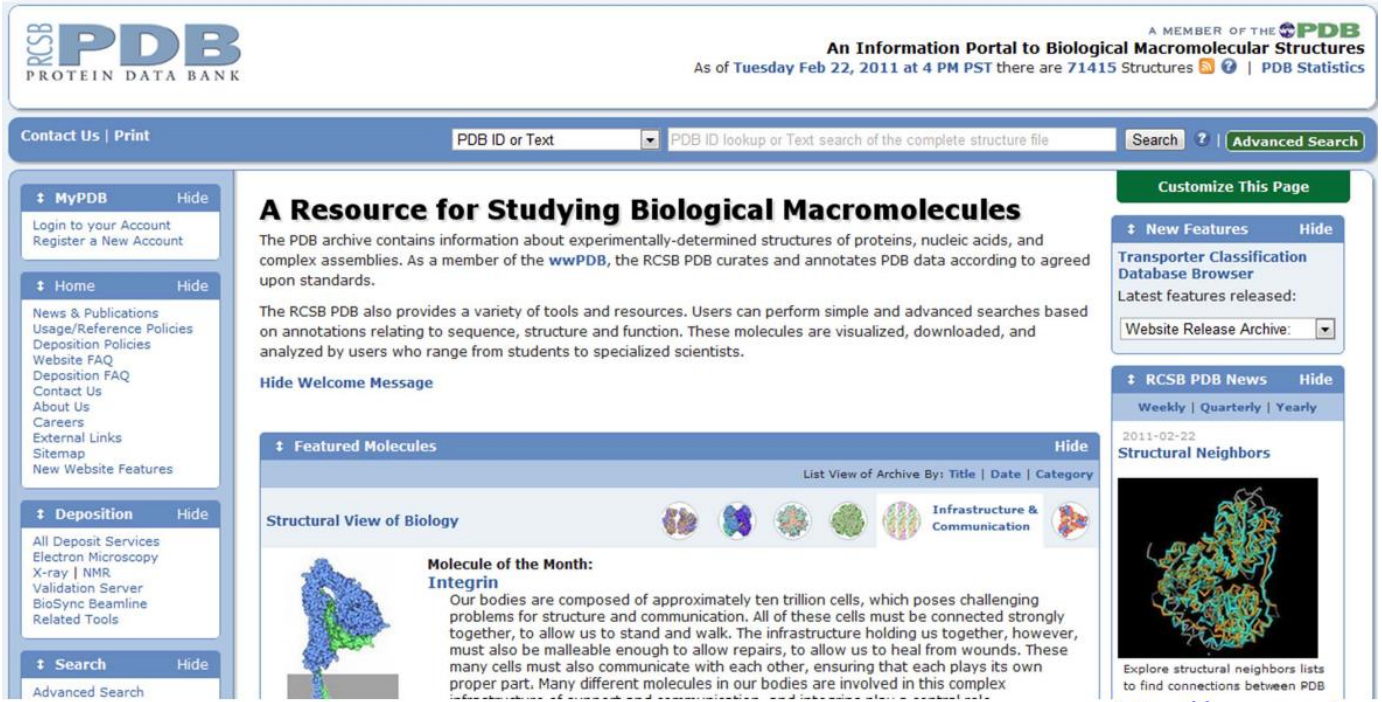
Template PDBs

Human β2-adrenergic receptor (PDB: 2rh1)

Human A2A adenosine receptor (PDB: 3eml)

Human CXC chemokine 4 receptor (PDB: 3odu)

Find these files at /day06/rosetta_cm/1_setup/



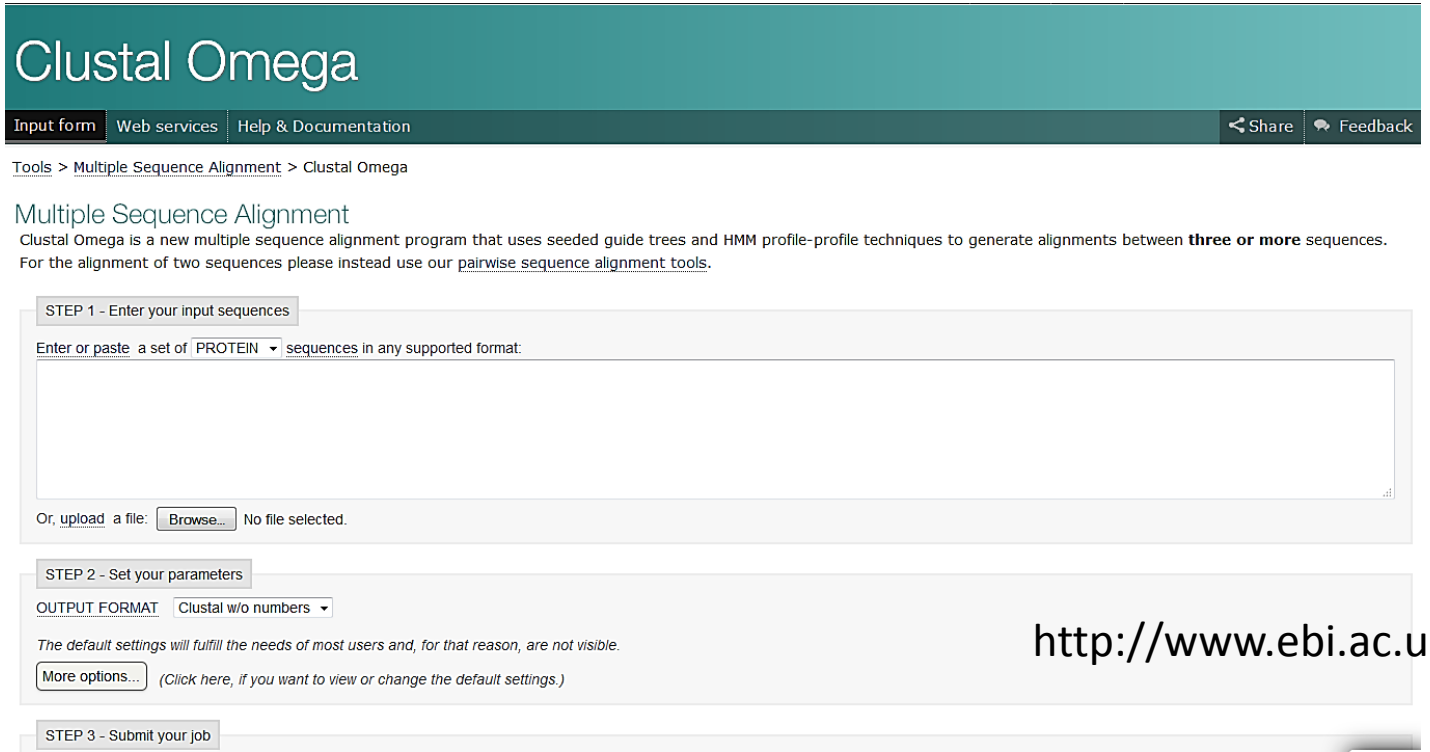
http://www.rcsb.o

Multiple Sequence Alignment

Find this file at /day06/rosetta_cm/2_threading/1u19_2rh1_3eml_3odu.aln

```
CLUSTAL O(1.2.1) multiple sequence alignment


1u19          -----PWQFSM--LAAYMFLLIMLGFPINFLTLYVTVQHKKLRTPLNYILLNLAVADLFM
3ODU_A        ANFNKIFL------PTIYSIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLF
2RH1_A        ---DEVWVVGMGIVMS---LIVLAIVFGNVLVITAIAKFERLQTVTNYFITSLACADLVM
3EML_A        -------IMGSSVYITVELAIAVLAILGNVLVCWAVWLNSNLQNVTNYFVVSLAAADIAV
                             :    : :  .  * *.  .    ..*:.  :     *: **: .
```



http://www.ebi.ac.uk/Tools/msa/clustalo/

Adjusting multiple sequence alignment

Experimental expectations:

- Highly conserved residues
- Secondary structure elements

Raw ClustalO alignment

```
1u19    - - - - - P W Q F S M - - L A A Y M F L L I M L G F P I N F L T L Y V T V Q H K K
3ODU_A  A N F N K I F L - - - - - - P T I Y S I I F L T G I V G N G L V I L V M G Y Q K K
2RH1_A  - - - D E V W V V G M G I V M S - - - L I V L A I V F G N V L V I T A I A K F E R
3EML_A  - - - - - - - I M G S S V Y I T V E L A I A V L A I L G N V L V C W A V W L N S N
```

Adjusted alignment

```
1u19    - - - - - - - P W Q F S M L A A Y M F L L I M L G F P I N F L T L Y V T V Q H
3ODU_A  A N F - - - - - - N K I F L P T I Y S I I F L T G I V G N G L V I L V M G Y Q
2RH1_A  - - - D - - - E V W V V G M G I V M S L I V L A I V F G N V L V I T A I A K F
3EML_A  - - - - - - - I M G S S V Y I T V E L A I A V L A I L G N V L V C W A V W L N
```

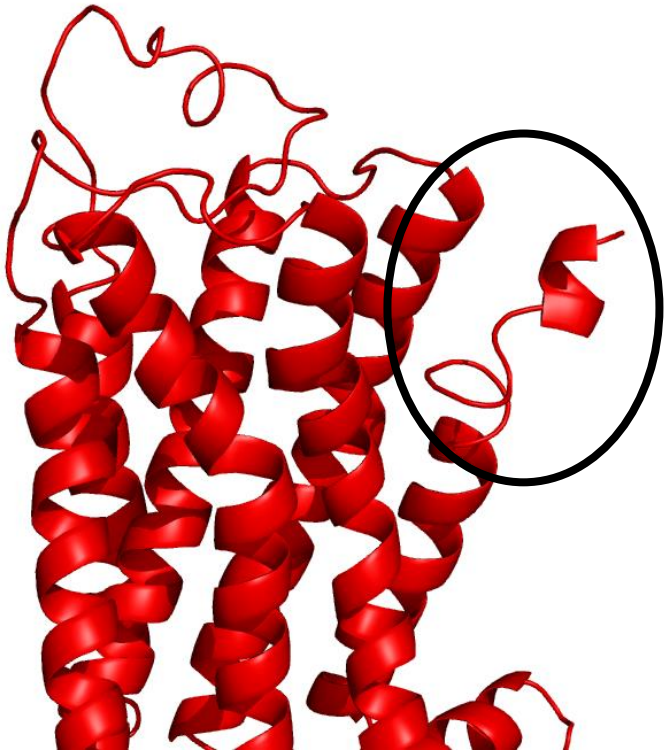
helix regions

highly conserved residues

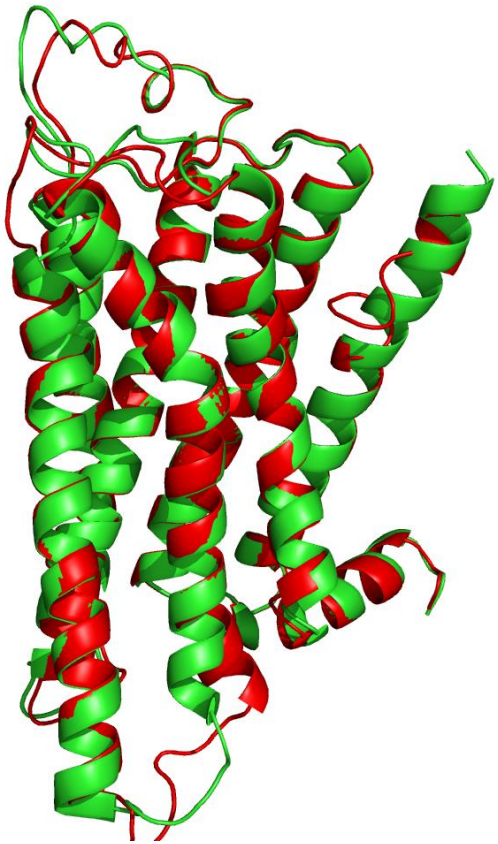
Alignment issues to be resolved

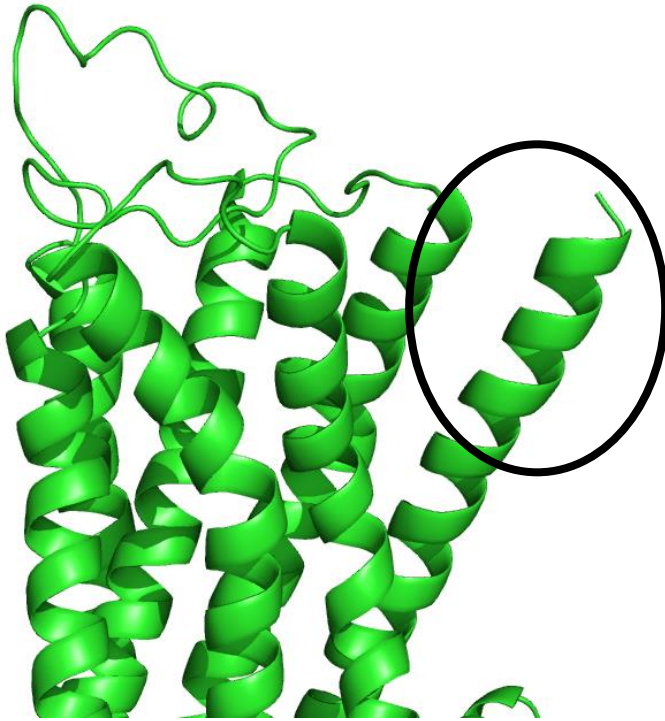
predicted membrane spanning region from OCTOPUS

Removing helix gaps



Raw alignment

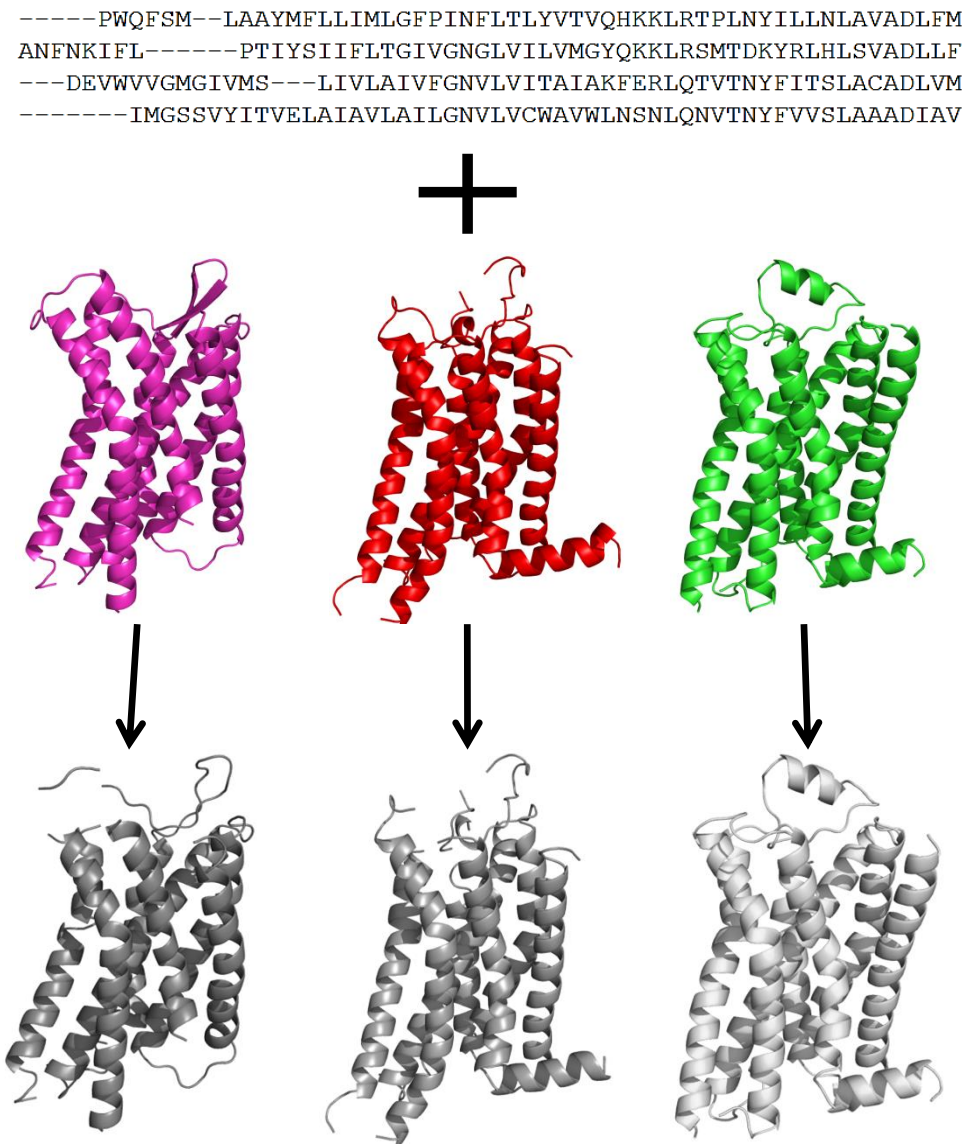




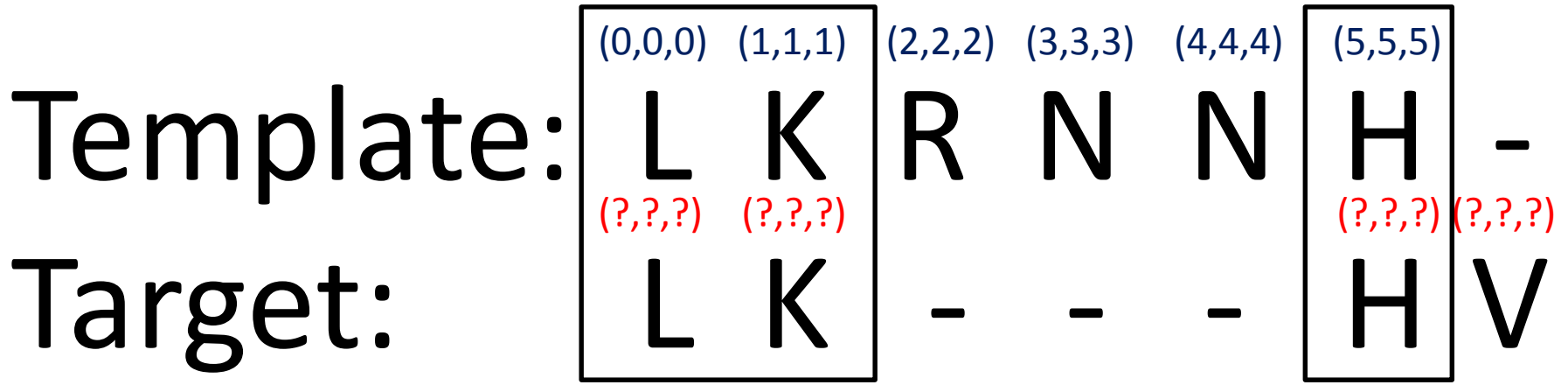
Adjusted alignment

Comparative Modeling Protocol

- **Step 1**: Align target sequence with template sequences
- **Step 2**: Partial-thread the target sequence onto template structures
- **Step 3**: Combine pieces from different templates using RosettaCM Hybridize
- **Step 4**: Full-atom refinement and relax



Threading



Thread Coordinates

Grishin Format

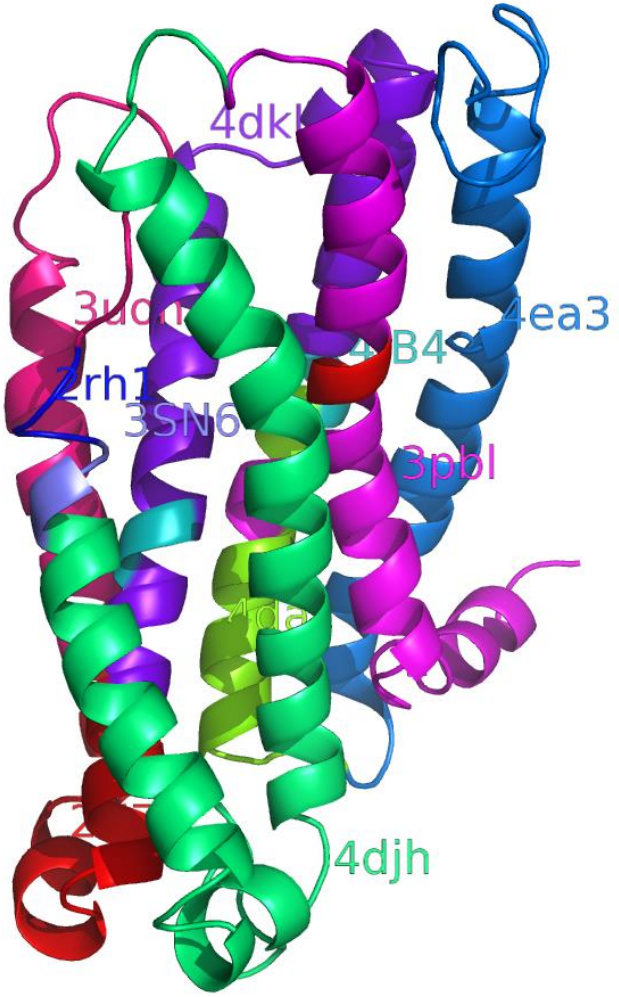
- ClustalO:
  - all sequences in one file
  - Sequences broken up over several lines
- Grishin:
  - one file per alignment pair
  - sequences continuous over one line each
  - Contains header information

Find converted alignment files at /day06/rosetta_cm/2_threading/

(1u19_2rh1.grishin, 1u19_3eml.grishin, 1u19_3odu.grishin)

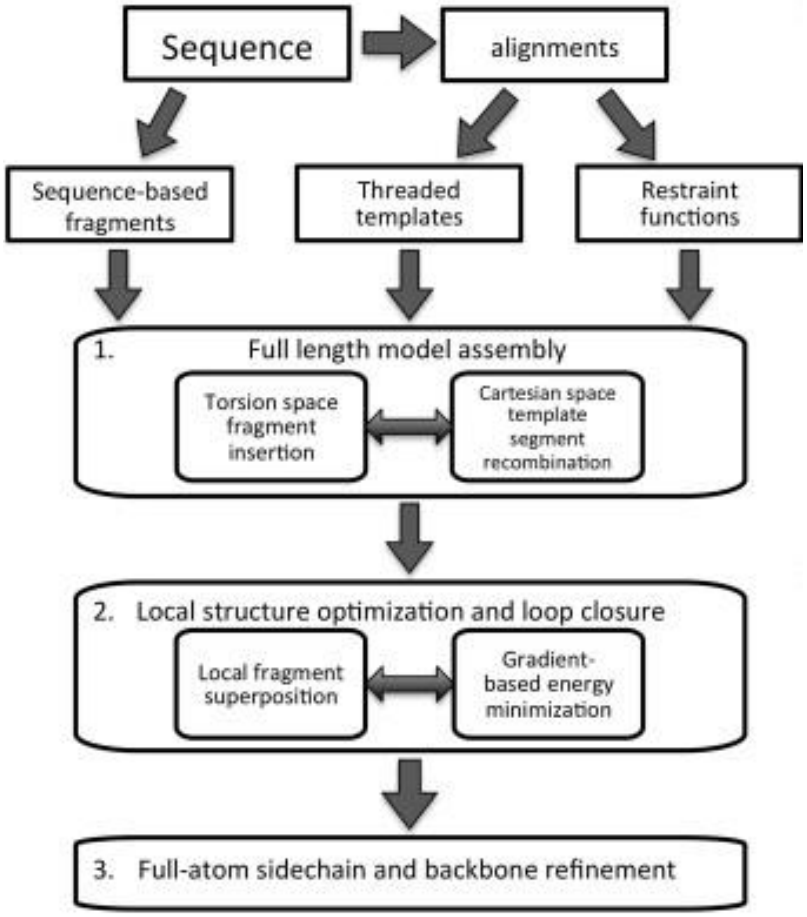
Comparative Modeling Protocol

- **Step 1**: Align target sequence with template sequences
- **Step 2**: Partial-thread the target sequence onto template structures
- **Step 3**: Combine pieces from different templates using RosettaCM Hybridize
- **Step 4**: Full-atom refinement and relax



RosettaCM: Three Stages

1. Generate initial models from template alignments
2. **Explore deviations from templates and close loops with 2-step MC:**
   - Randomly select de novo or template-based fragment and substitute into current conformation
   - Cartesian space full-backbone minimization
3. Full atom refinement and final relax



Song, et al. 2013

Hybridize input

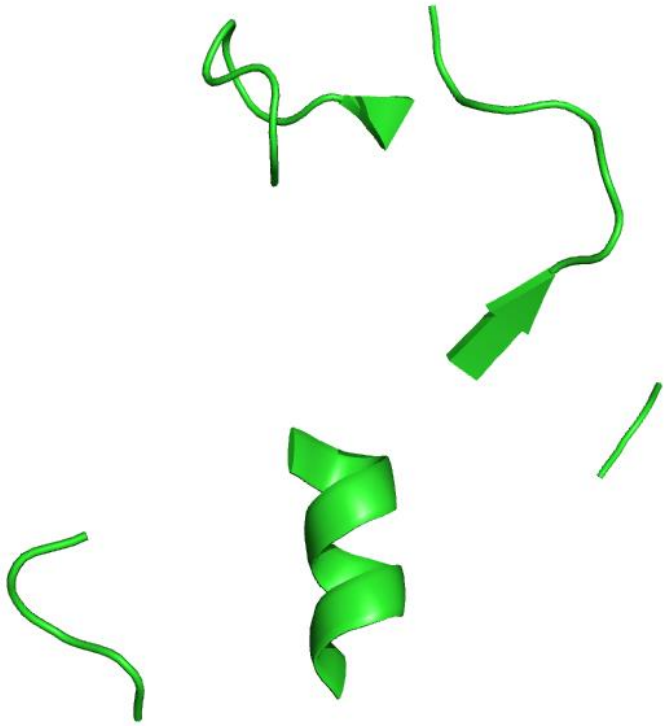
- Partial-threads
- Fragment files (3mer and 9mer)
- Membrane spanning regions (span file)
- Mover definition and options
- Weight patches

Fragment Files

Find these files at

/day06/rosetta_cm/3_hybridize/1u19_3.frags c

/day06/rosetta_cm/3_hybridize/1u19_9.frags



www.bakerlab.org

ROBETTA BETA

Full-chain Protein Structure Prediction Server

Structure Prediction Fragment Libraries Alanine Scanning DNA Interface Scan

[Queue] [Submit] [Queue] [Submit] [Queue] [Submit] [Queue] [Submit]

[Register / Update] [Docs / FAQs] [Login]

Submit a job to the Fragment Server

*Please submit one job at a time

Required

Registered Username: or Registered Email Address:

Target Name:

Paste Fasta

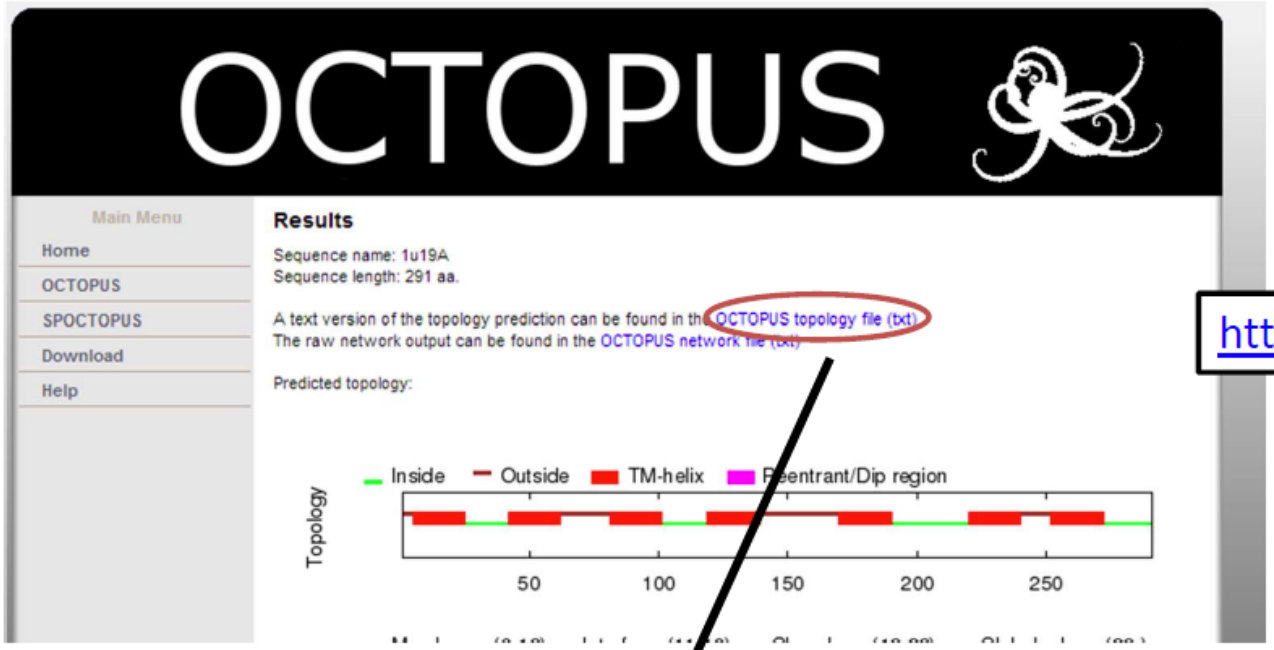
or Upload Fasta: Choose File No file chosen

Submit

http://robetta.bakerlab.org/fragmentsubmit.jsp

Membrane spanning regions

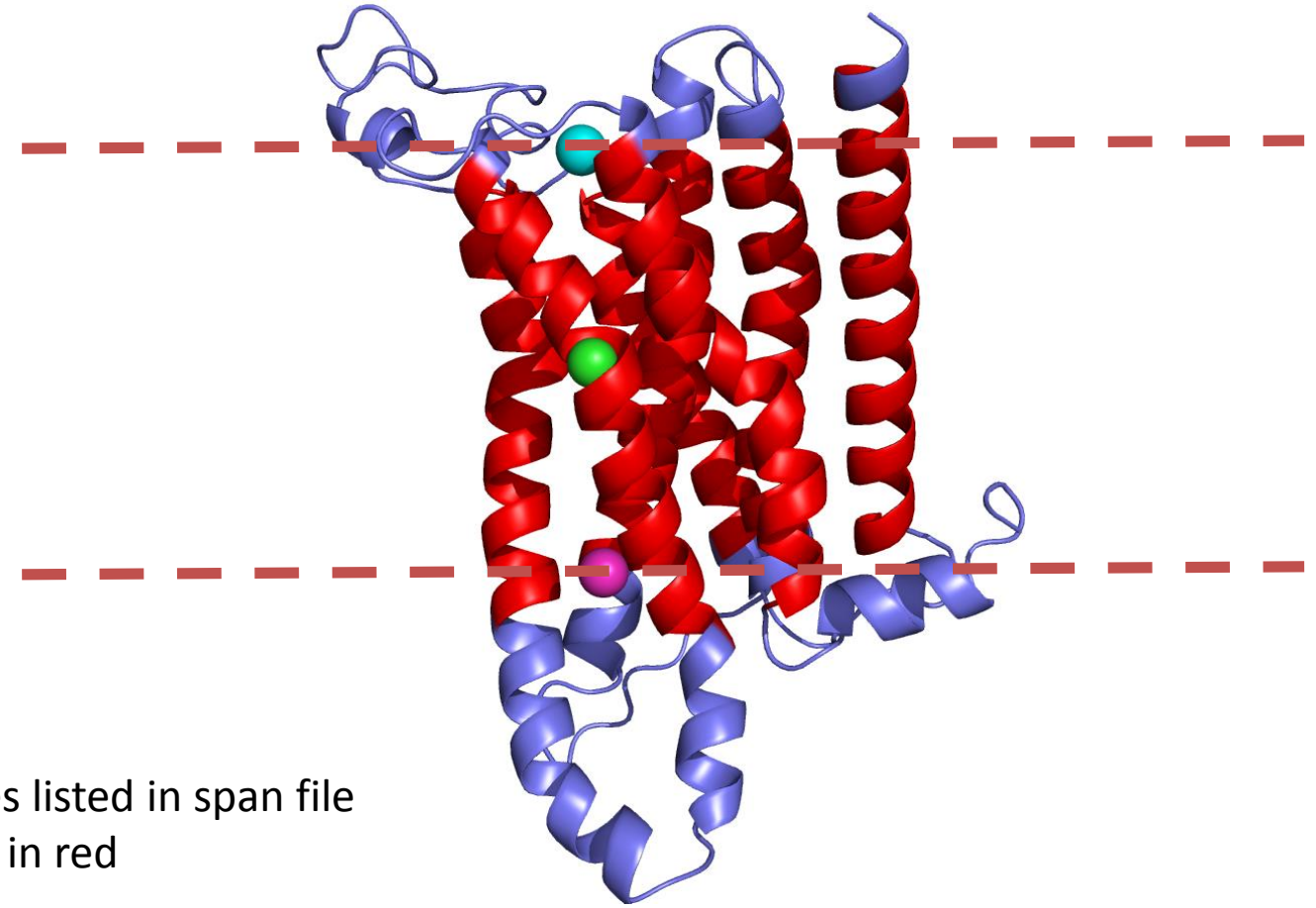
Find this file at /day06/rosetta_cm/3_hybridize/1u19.span



http://octopus.cbr.su.se/

octopus2span.pl 1u19A.octopus

Rosetta Membrane



Residues listed in span file
colored in red

Hybridize Mover

/day06/rosetta_cm/3_hybridize/rosetta_cm.xml

```
<SCOREFXNS>
        <stage1 weights="stage1_membrane.wts" symmetric=0>
                <Reweight scoretype=atom_pair_constraint weight=1/>
        </stage1>
        <stage2 weights="stage2_membrane.wts" symmetric=0>
                <Reweight scoretype=atom_pair_constraint weight=0.5/>
        </stage2>
        <fullatom weights="stage3_rlx_membrane.wts" symmetric=0>
                <Reweight scoretype=atom_pair_constraint weight=0.5/>
        </fullatom>
</SCOREFXNS>
```

*Find all **.wts** files in /day06/rosetta_cm/3_hybridize/*

```
<MOVERS>
        <Hybridize name=hybridize stage1_scorefxn=stage1 stage2_scorefxn=stage2
          fa_scorefxn=fullatom batch=1 stage1_increase_cycles=1.0
          stage2_increase_cycles=1.0 linmin_only=1>
                <Fragments 3mers="1u19_3.frags" 9mers="1u19_9.frags"/>
                <Template pdb="1u19_on_2rh1.pdb" cst_file="AUTO" weight=  1.000 />
                <Template pdb="1u19_on_3eml.pdb" cst_file="AUTO" weight=  1.000 />
                <Template pdb="1u19_on_3odu.pdb" cst_file="AUTO" weight=  1.000 />
        </Hybridize>
</MOVERS>
```

RosettaCM Options

/day06/rosetta_cm/3_hybridize/rosetta_cm.options

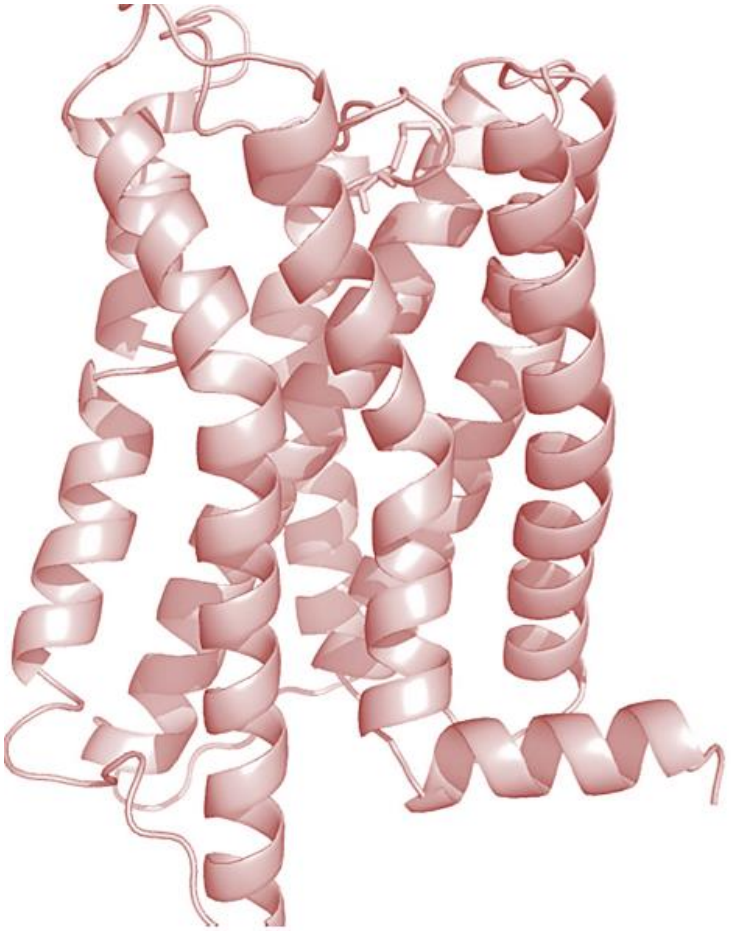
```
# i/o
-in:file:fasta 1u19.fasta
-parser:protocol rosetta_cm.xml

# relax options
-relax:minimize_bond_angles
-relax:minimize_bond_lengths
-relax:jump_move true
-default_max_cycles 200
-relax:min_type lbfgs_armijo_nonmonotone
-score:weights stage3.wts
-use_bicubic_interpolation
-hybridize:stage1_probability 1.0

#Initialize membrane
-in:file:spanfile 1u19.span
-membrane:no_interpolate_Mpair
-membrane:Menv_penalties
-rg_reweight .1
```

Comparative Modeling Protocol

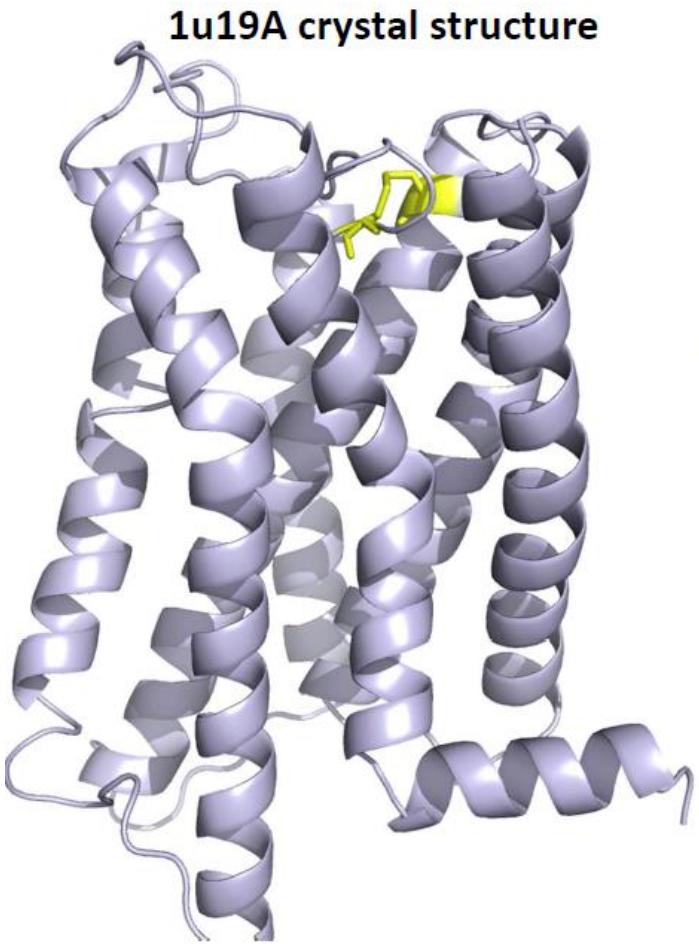
- **Step 1**: Align target sequence with template sequences
- **Step 2**: Partial-thread the target sequence template structures
- **Step 3**: Combine pieces from different templates using RosettaCM Hybridize
- **Step 4**: Full-atom refinement and relax

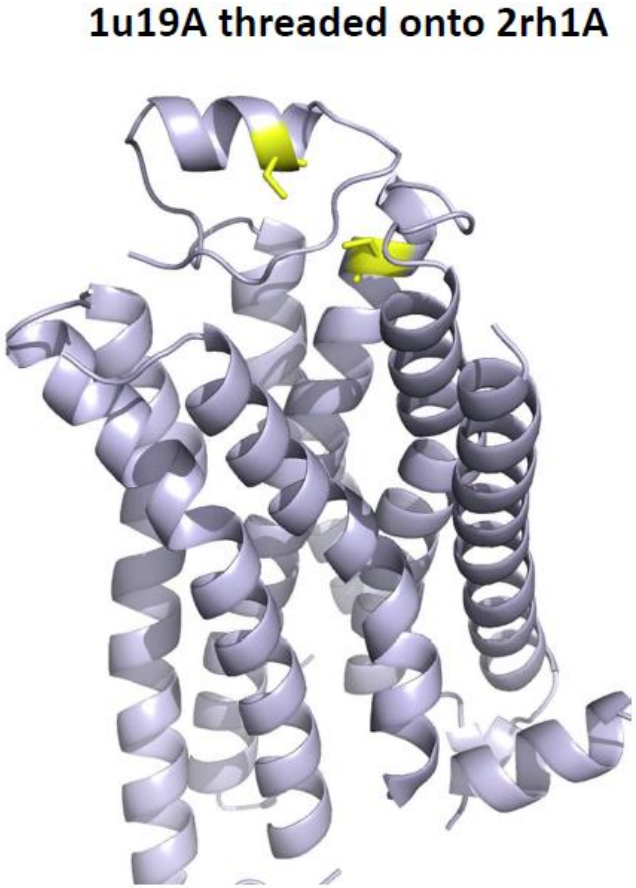


Disulfide constraints

Find this file at /day06/rosetta_cm/4_relax/1u19.disulfide

77 154





Relax options

/rosetta_cm/4_relax/relax.options

 # i/o
-in:fix_disulf 1u19.disulfide
-out:file:fullatom
-out:pdb

 # Dualspace settings
-relax:dualspace
-relax:minimize_bond_angles
-set_weights cart_bonded .5 pro_close 0
-default_max_cycles 200

#Initialize membrane
-in:file:spanfile 1u19.span
-membrane:no_interpolate_Mpair
-membrane:Menv_penalties
-score:weights membrane_highres_Menv_smooth.wts

Tutorial

Comparative modeling of rhodopsin with three class A GPCR's

Five stages:

1. Setup
2. Threading
3. RosettaCM hybridize
4. Dualspace relax
5. Final model selection

References

- **Rosetta 3.5 User Guide**

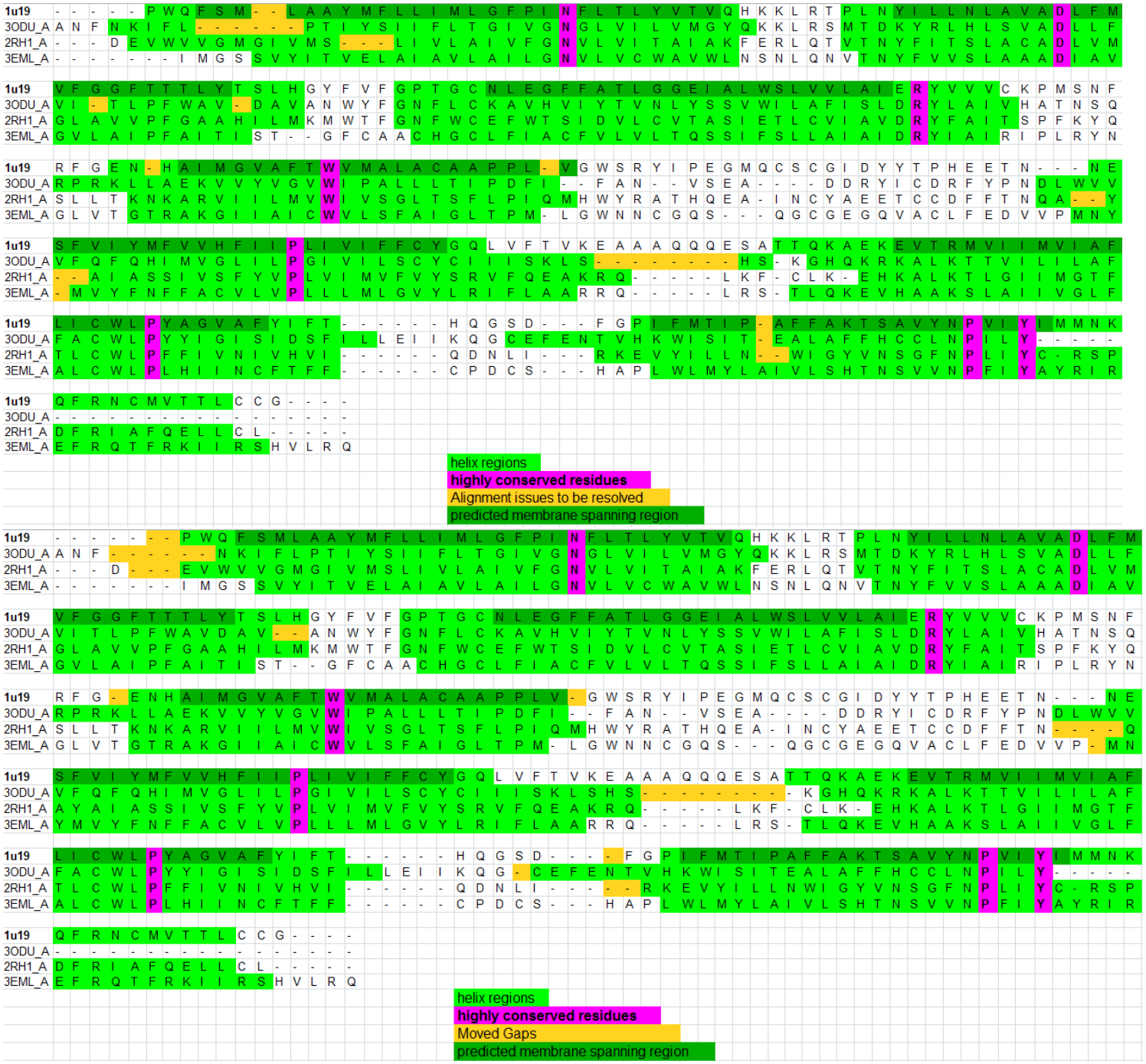
https://www.rosettacommons.org/manuals/archive/rosetta3.5_user_guide/index.html

- **Modeling Membrane Proteins**

https://www.rosettacommons.org/manuals/archive/rosetta3.5_user_guide/db/d38/membrane_abinitio.html

- **RosettaCM**

Yifan Song, et al. (2013). High-Resolution Comparative Modeling with RosettaCM. Structure, *21*(10), 1735-1742.



Initial alignment

Adjusted alignment