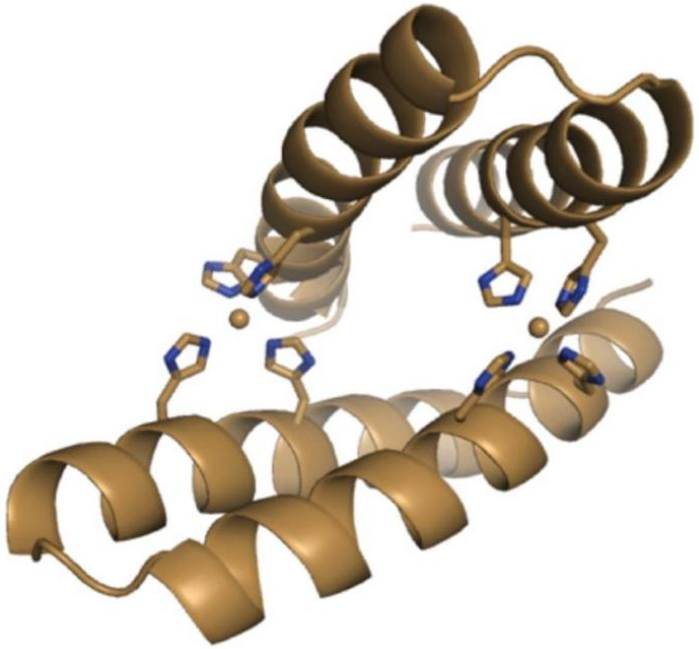
General Design & Thermostabilization

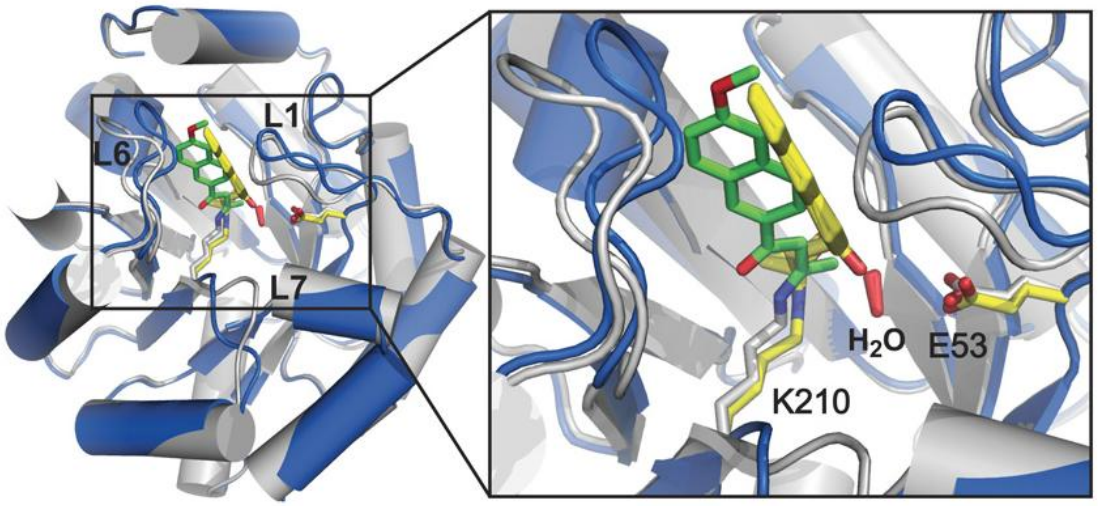
Diego del Alamo, Meiler Lab

8 May 2018

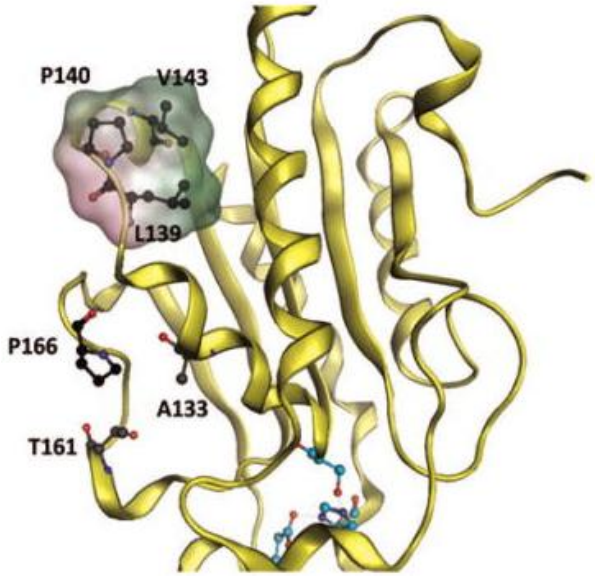
Rosetta Workshop



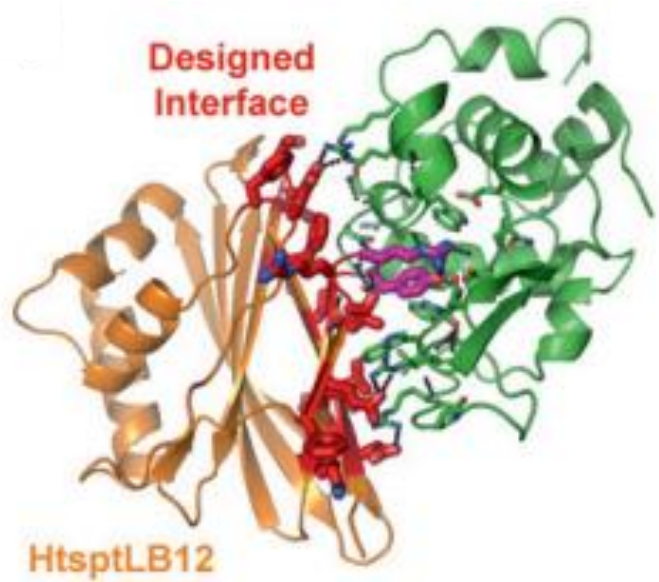
Dimers with metal ion-binding interfaces



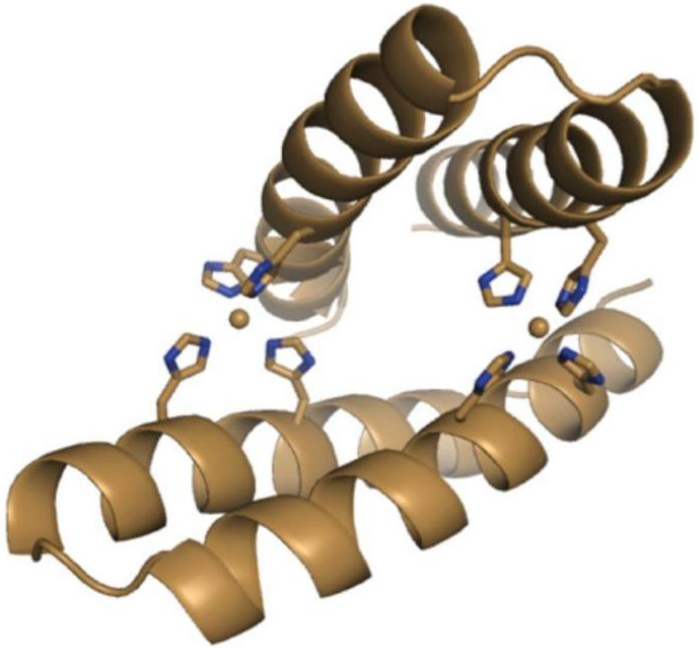
Novel enzymes



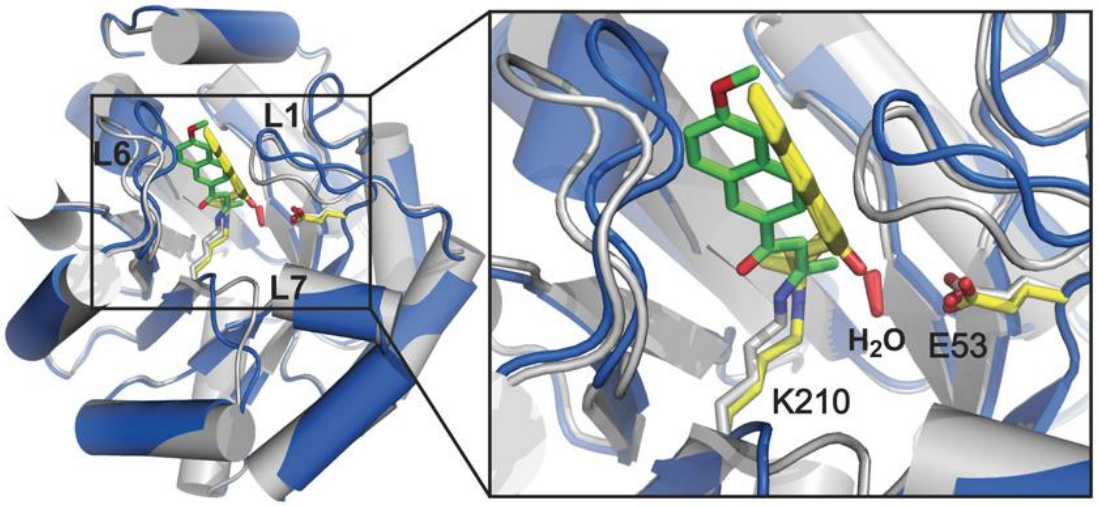
Thermostabilized proteins



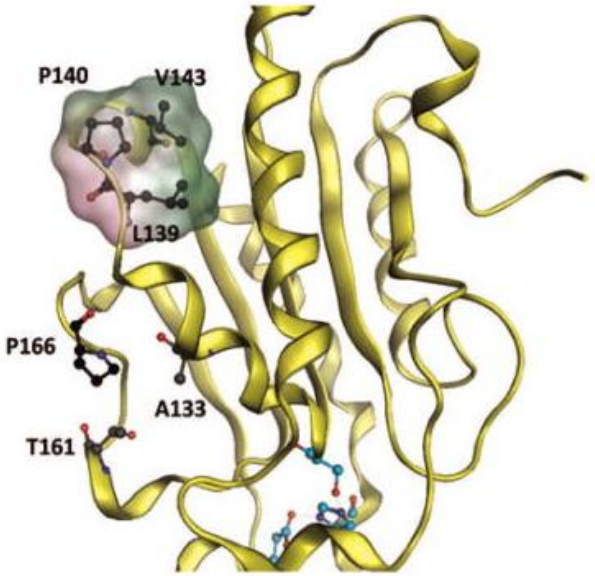
Enzyme-inhibiting proteins



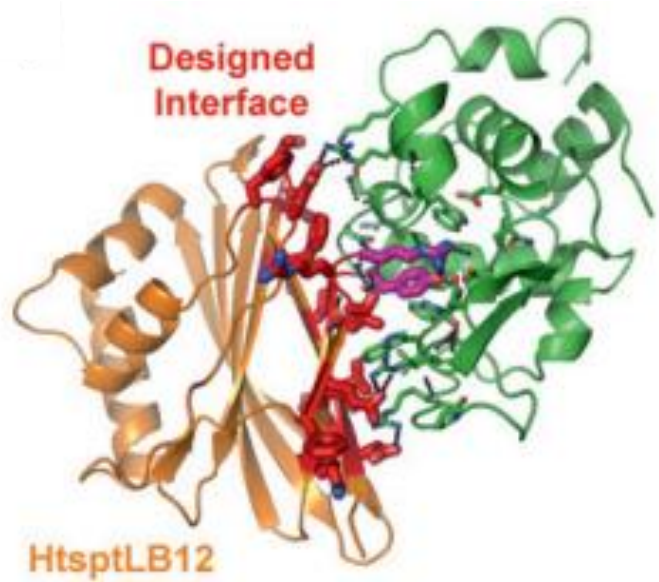
Dimers with metal ion-binding interfaces



Novel enzymes



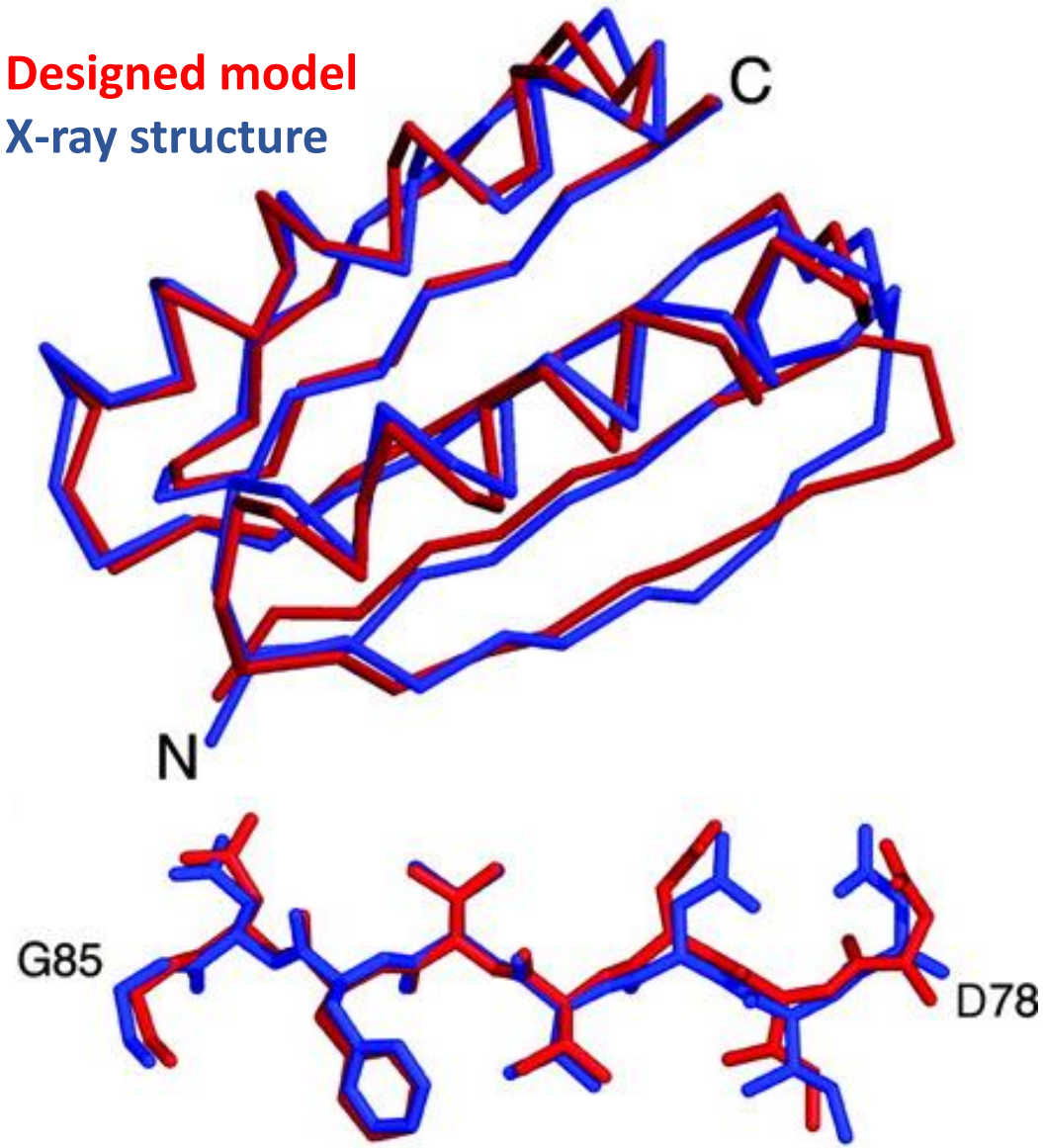
Thermostabilized proteins



Enzyme-inhibiting proteins

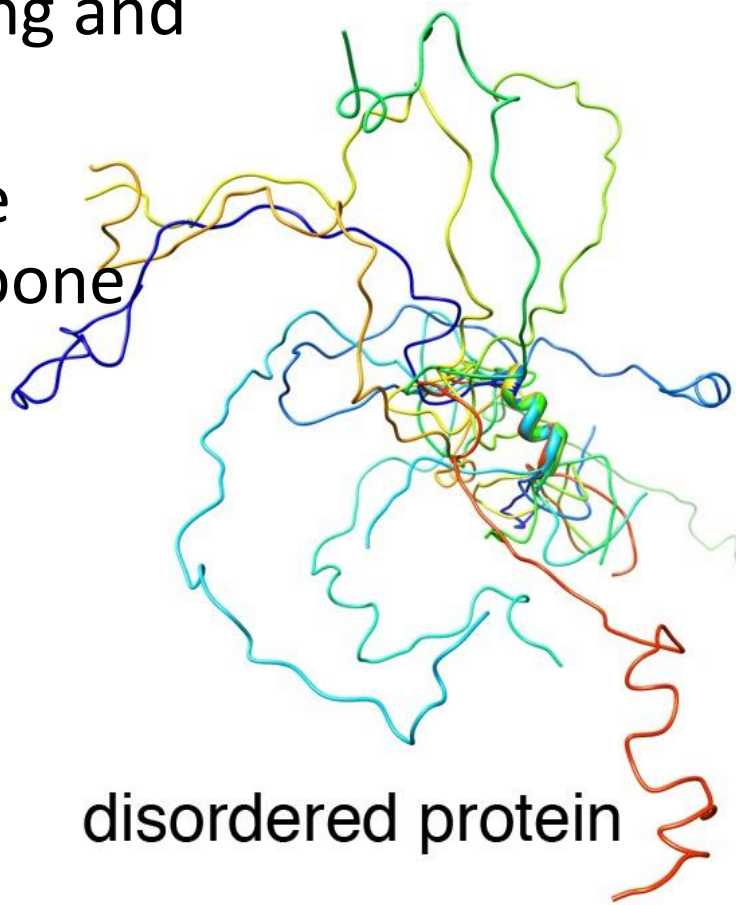
Protein Design: What is this?

- Protein design: given the structure, what is the sequence?
- (Contrasts with protein structure prediction: given the sequence, what is the structure?)
- Both problems use the same knowledge potentials, the same energy terms, etc
- Highly customizable

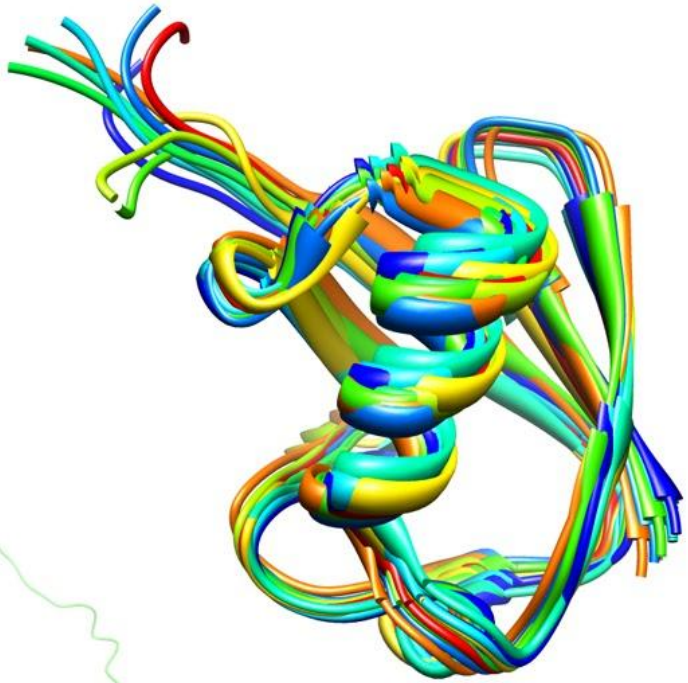


Thermostabilization

- Proteins undergo Brownian motion, conformational changes, folding and unfolding, etc.
- Thermostabilization: sequence changes to favor a given backbone configuration
  - (this can also be achieved by destabilizing alternate states)



disordered protein



well-folded protein

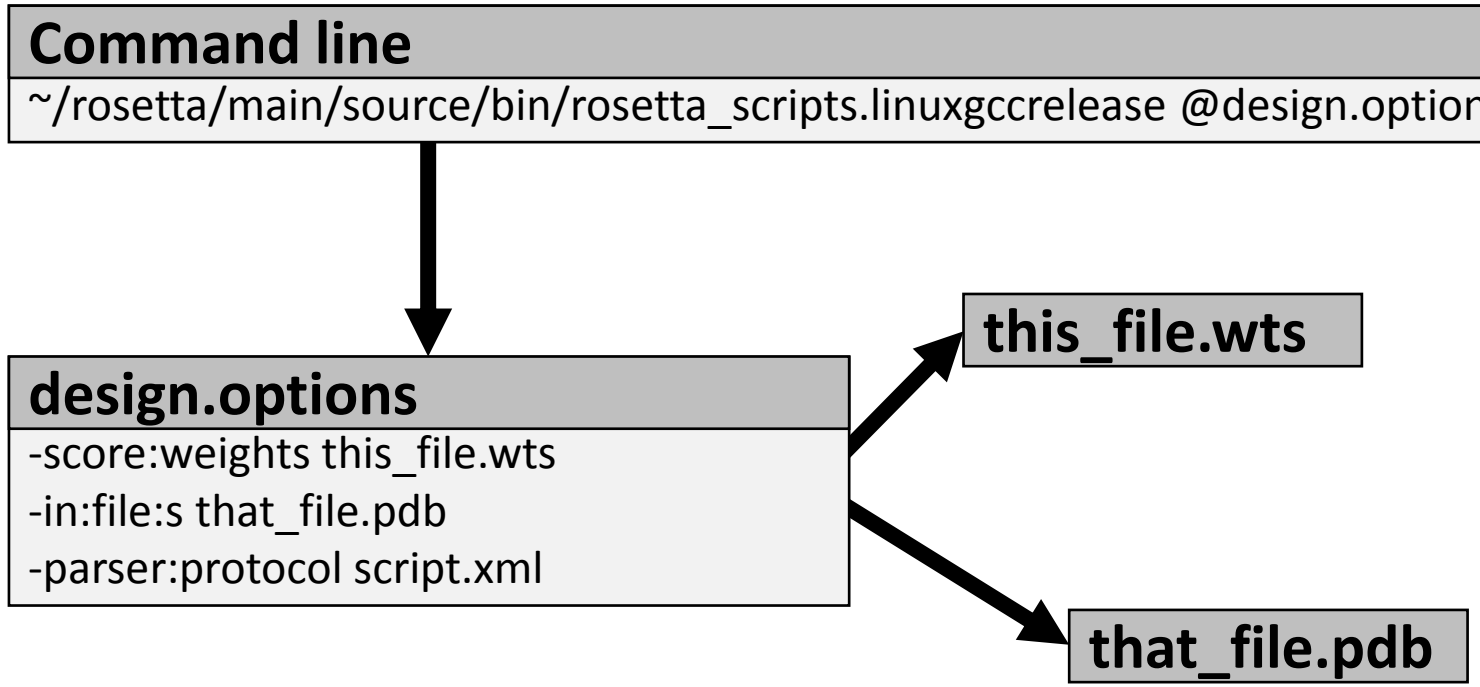
Command line
~/rosetta/main/source/bin/rosetta_scripts.linuxgccrelease @design.options

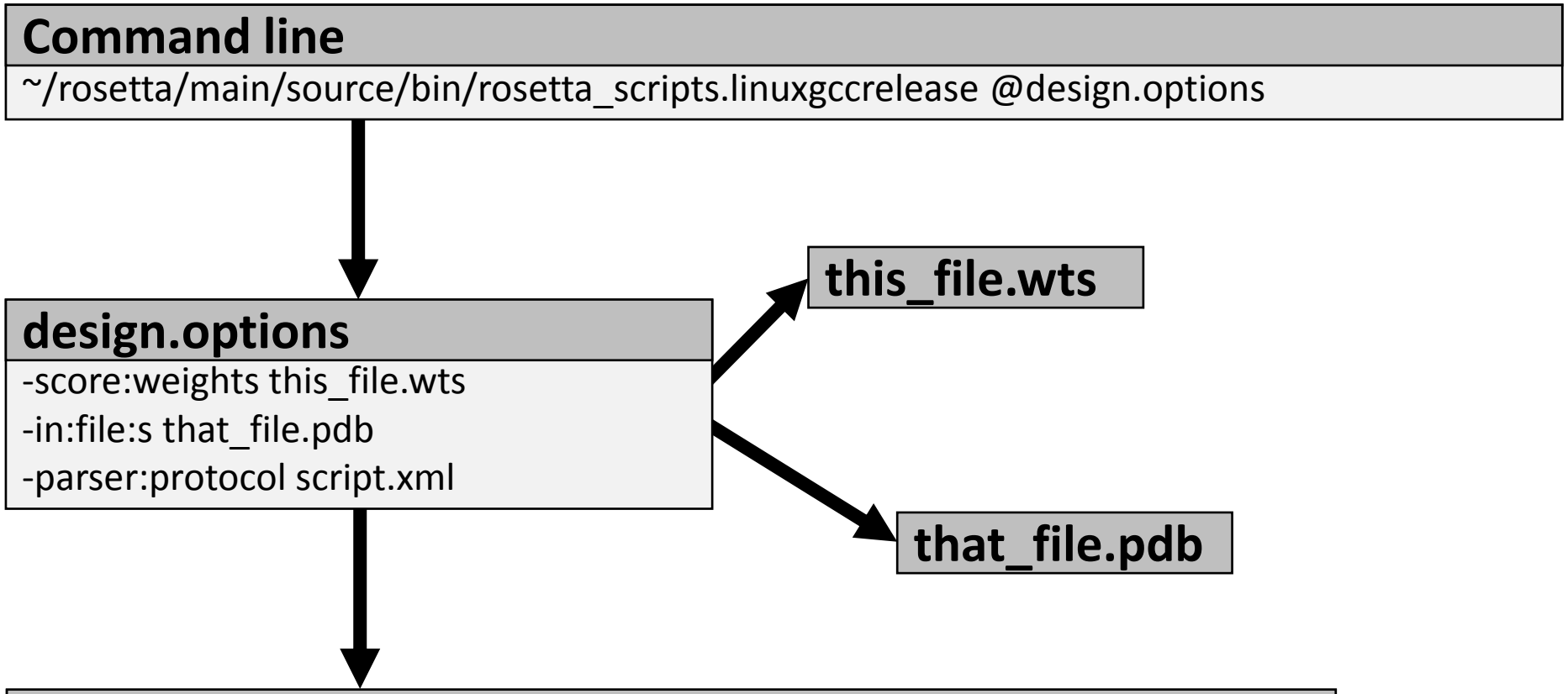
Command line

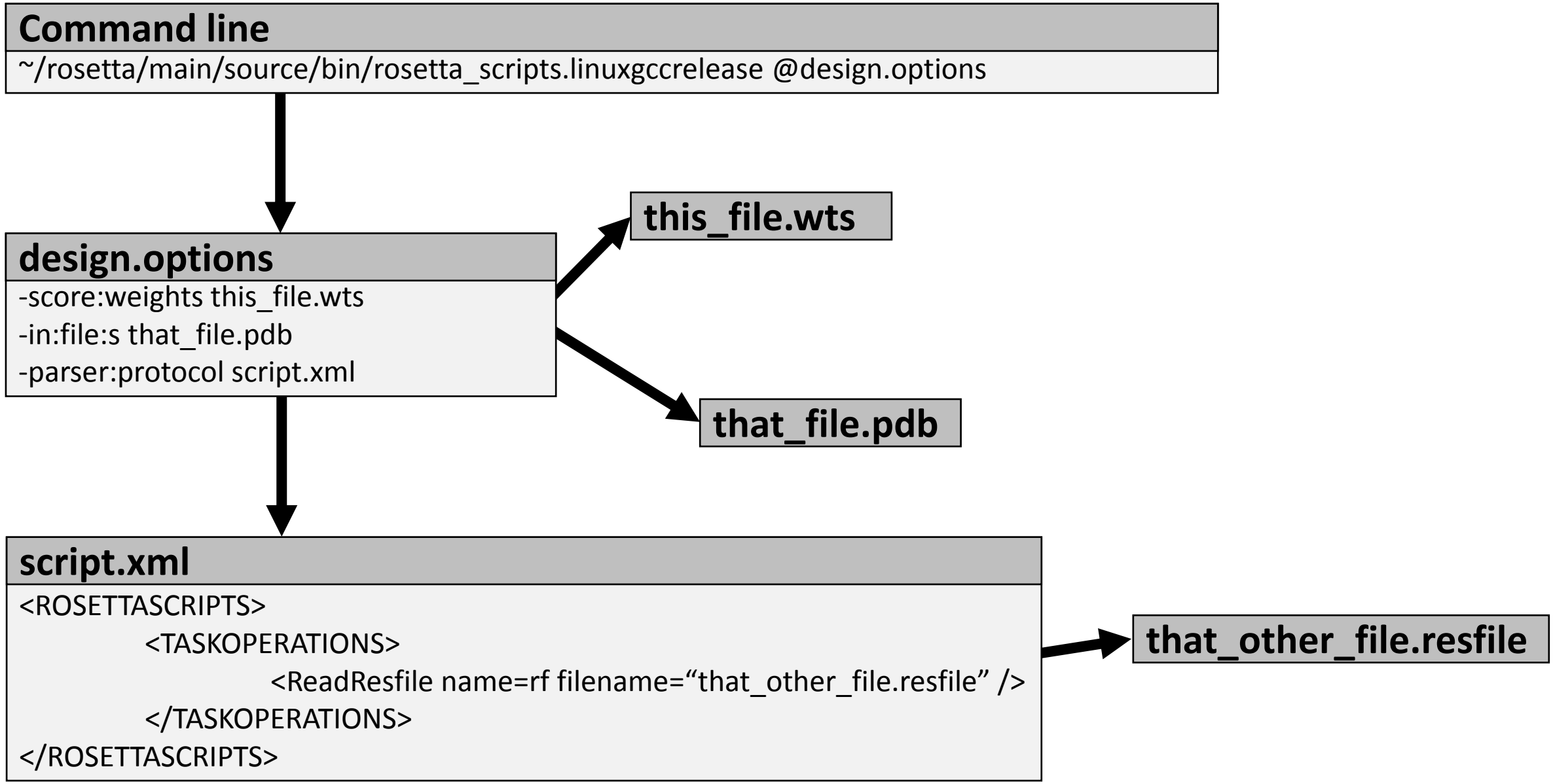
~/rosetta/main/source/bin/rosetta_scripts.linuxgccrelease @design.options

design.options

-score:weights this_file.wts
-in:file:s that_file.pdb
-parser:protocol script.xml


Command line
~/rosetta/main/source/bin/rosetta_scripts.linuxgccrelease @design.options
design.options
-score:weights this_file.wts
-in:file:s that_file.pdb
-parser:protocol script.xml
this_file.wts
that_file.pdb


Command line
~/rosetta/main/source/bin/rosetta_scripts.linuxgccrelease @design.options
design.options
-score:weights this_file.wts
-in:file:s that_file.pdb
-parser:protocol script.xml
this_file.wts
that_file.pdb
script.xml
<ROSETTASCRIPTS>
<TASKOPERATIONS>
<ReadResfile name=rf filename="that_other_file.resfile" />
</TASKOPERATIONS>
</ROSETTASCRIPTS>


Command line
~/rosetta/main/source/bin/rosetta_scripts.linuxgccrelease @design.options
design.options
-score:weights this_file.wts
-in:file:s that_file.pdb
-parser:protocol script.xml
this_file.wts
that_file.pdb
script.xml
<ROSETTASCRIPTS>
<TASKOPERATIONS>
<ReadResfile name=rf filename="that_other_file.resfile" />
</TASKOPERATIONS>
</ROSETTASCRIPTS>
that_other_file.resfile

Command line

~/rosetta/main/source/bin/rosetta_scripts.linuxgccrelease @design.options

design.options

-score:weights this_file.wts
-in:file:s that_file.pdb
-parser:protocol script.xml

this_file.wts

that_file.pdb

File types used in this tutorial:
-Fasta files
-Span files
-Weights/score files
-Resfiles

script.xml

```
<ROSETTASCRIPTS>
        <TASKOPERATIONS>
                <ReadResfile name=rf filename="that_other_file.resfile" />
        </TASKOPERATIONS>
</ROSETTASCRIPTS>
```

that_other_file.resfile

Resfiles: the protein design decider

- Very short, describe parameters for design
- Case-insensitive, whitespace-delimited, commented with ‘#’
- Contains a “header” section that applies to all residues and a “body” section for residue-specific design decisions
- Used to introduce noncanonical/unnatural amino acid
- Layout:
  - Header: sets parameters for default residue design
  - “START”
  - Body: sets parameters for individual residues or ranges of residues

Resfiles: the protein design decider

- Very short, describe parameters for design
- Case-insensitive, whitespace-delimited, commented with ‘#’
- Contains a “header” section that applies to all residues and a “body” section for residue-specific design decisions
- Used to introduce noncanonical/unnatural amino acid
- Layout:
  - Header: sets parameters for default residue design
  - “START”
  - Body: sets parameters for individual residues or ranges of residues

full_redesign.resfile

ALLAA *# all amino acids are permitted*
START

Resfiles: the protein design decider

- Very short, describe parameters for design
- Case-insensitive, whitespace-delimited, commented with ‘#’
- Contains a “header” section that applies to all residues and a “body” section for residue-specific design decisions
- Used to introduce noncanonical/unnatural amino acid
- Layout:
  - Header: sets parameters for default residue design
  - “START”
  - Body: sets parameters for individual residues or ranges of residues

full_redesign.resfile

ALLAA *# all amino acids are permitted*
START

no_effect.resfile

NATRO *# keep existing sidechain in place*
START

Resfiles: the protein design decider

- Very short, describe parameters for design
- Case-insensitive, whitespace-delimited, commented with '#'
- Contains a "header" section that applies to all residues and a "body" section for residue-specific design decisions
- Used to introduce noncanonical/unnatural amino acid
- Layout:
  - Header: sets parameters for default residue design
  - "START"
  - Body: sets parameters for individual residues or ranges of residues

full_redesign.resfile

ALLAA *# all amino acids are permitted*
START

no_effect.resfile

NATRO *# keep existing sidechain in place*
START

partial_redesign.resfile

NATAA *# keep existing amino acid, permit repacking*
START
10 – 20 A ALLAA *# redesign 10 to 20, chain A*
21 – 30 A POLAR *# redesign, polar residues only*
31 – 40 A APOLAR *# redesign, apolar residues only*
41 A PIKAA AILV *# redesign, Ala, Ile, Leu, or Val only*

Resfile syntax

- NATRO: Keep rotamers in place
- NATAA: Keep residue identity unchanged, allow repacking
- ALLAA: Allow full redesign
- ALLAAwc: Same as ALLAA
- ALLAAxc: Full redesign except cysteine
- PIKAA: Any residue from among the following (one-letter codes)
- NOTAA: Any residue except from among the following
- POLAR: Polar residues only
- APOLAR: Apolar residues only

Today’s design protocol

introduce_mutations_a.options

```
-parser:protocol introduce_mutations_a.xml
-score:weights mpframework_smooth_fa_2012.wts
-mp:setup:spanfiles LeuT.span
-mp:scoring:hbond true
```

introduce_mutations_a.xml

```
<ROSETTASCRIPTS>
        <SCOREFXNS>
                <ScoreFunction name="memb_hires" weights="mpframework_smooth_fa_2012.wts" />
        </SCOREFXNS>
        <TASKOPERATIONS>
                <ReadResfile name="resfile" filename="mutations_a.resfile" />
                <InitializeFromCommandline name="ifcl"/>
                <IncludeCurrent name="ic" />
                <RestrictToRepacking name="rtr" />
        </TASKOPERATIONS>
        <MOVERS>
                <AddMembraneMover name="add_memb" />
                <MembranePositionFromTopologyMover name="init_pose" />
                <PackRotamersMover name="pack" scorefxn="memb_hires" task_operations="resfile,ifcl" />
                <FastRelax name="relax" scorefxn="memb_hires" repeats="2" task_operations="ifcl,ic,rtr" />
        </MOVERS>
        <PROTOCOLS>
                <Add mover="add_memb" />
                <Add mover="init_pose" />
                <Add mover="pack" />
                <Add mover="relax" />
        </PROTOCOLS>
</ROSETTASCRIPTS>
```

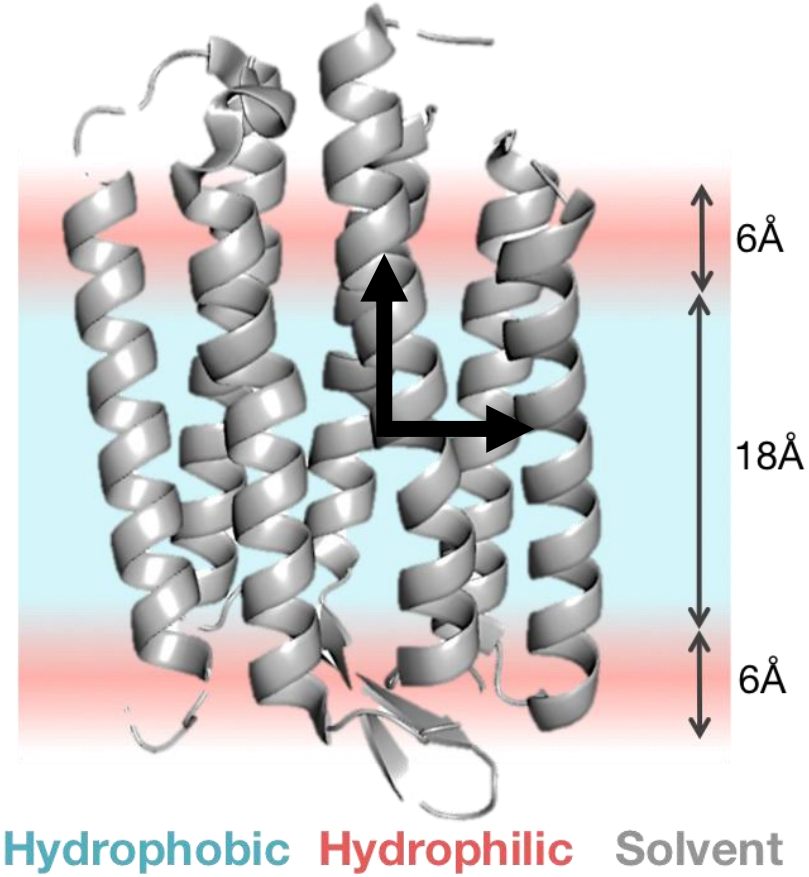
Today’s design protocol

introduce_mutations_a.options

```
-parser:protocol introduce_mutations_a.xml
-score:weights mpframework_smooth_fa_2012.wts
-mp:setup:spanfiles LeuT.span
-mp:scoring:hbond true
```

introduce_mutations_a.xml

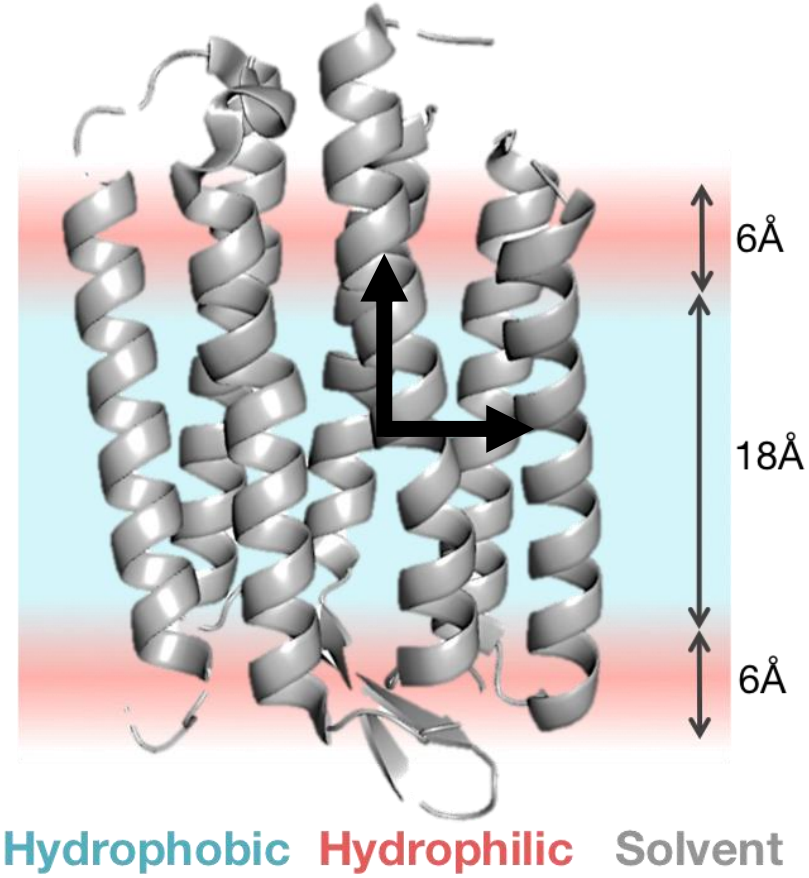
```
<ROSETTASCRIPTS>
        <SCOREFXNS>
                <ScoreFunction name="memb_hires" weights="mpframework_smooth_fa_2012.wts" />
        </SCOREFXNS>
        <TASKOPERATIONS>
                <ReadResfile name="resfile" filename="mutations_a.resfile" />
                <InitializeFromCommandline name="ifcl"/>
                <IncludeCurrent name="ic" />
                <RestrictToRepacking name="rtr" />
        </TASKOPERATIONS>
        <MOVERS>
                <AddMembraneMover name="add_memb" />
                <MembranePositionFromTopologyMover name="init_pose" />
                <PackRotamersMover name="pack" scorefxn="memb_hires" task_operations="resfile,ifcl" />
                <FastRelax name="relax" scorefxn="memb_hires" repeats="2" task_operations="ifcl,ic,rtr" />
        </MOVERS>
        <PROTOCOLS>
                <Add mover="add_memb" />
                <Add mover="init_pose" />
                <Add mover="pack" />
                <Add mover="relax" />
        </PROTOCOLS>
</ROSETTASCRIPTS>
```



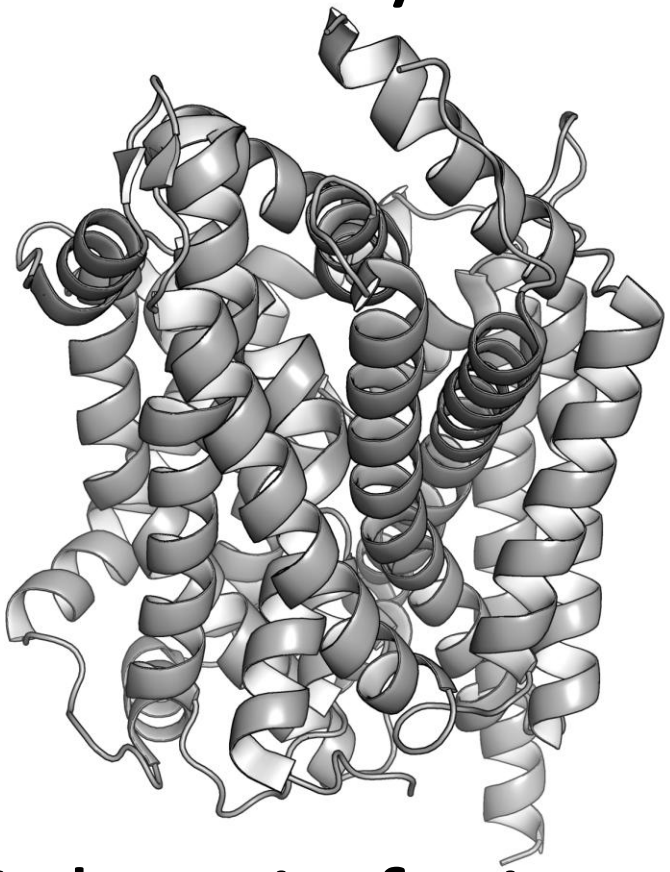
Today’s design protocol

model.pdb

```
...
ATOM   8345  HE  ARG A 519      12.606  -6.930  32.167  1.00  0.00           H
ATOM   8346 1HH1 ARG A 519      10.637  -8.525  29.654  1.00  0.00           H
ATOM   8347 2HH1 ARG A 519       9.277  -8.706  30.747  1.00  0.00           H
ATOM   8348 1HH2 ARG A 519      10.814  -7.215  33.503  1.00  0.00           H
ATOM   8349 2HH2 ARG A 519       9.367  -7.984  32.893  1.00  0.00           H
TER
HETATM 8351 THKN MEM B 520      15.000   0.000   0.000  1.00  0.00           X
HETATM 8352 CNTR MEM B 520       0.000   0.000   0.000  1.00  0.00           X
HETATM 8353 NORM MEM B 520       0.000   0.000   1.000  1.00  0.00           X
TER
CONECT 8351 8352
CONECT 8352 8351 8353
CONECT 8353 8352
# All scores below are weighted scores, not raw scores.
#BEGIN_POSE_ENERGIES_TABLE 2A65_full_0001.pdb
label fa_atr fa_rep fa_intra_rep pro_close fa_pair hbond_sr_bb hbond_lr_bb hbond_
dslf_ss_dih dslf_ca_dih rama omega fa_dun p_aa_pp ref fa_mpenv fa_mpenv_smooth li
weights 0.8 0.44 0.004 1 0.49 1.17 1.17 2.34 2.2 0.35 0.5 2 5 5 0.2 0.5 0.56 0.32
...
```

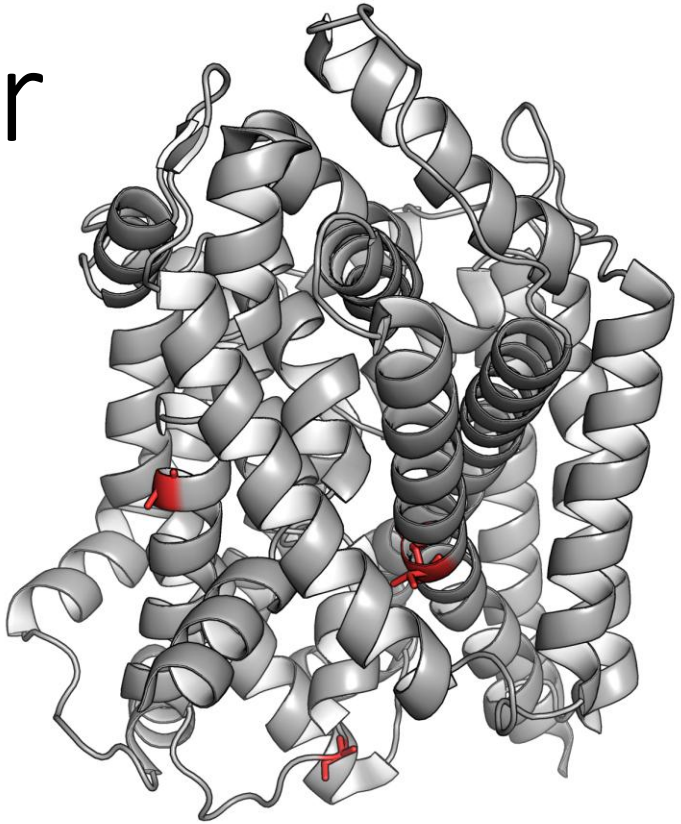


Today's test case: bacterial leucine transporter



Periplasmic-facing

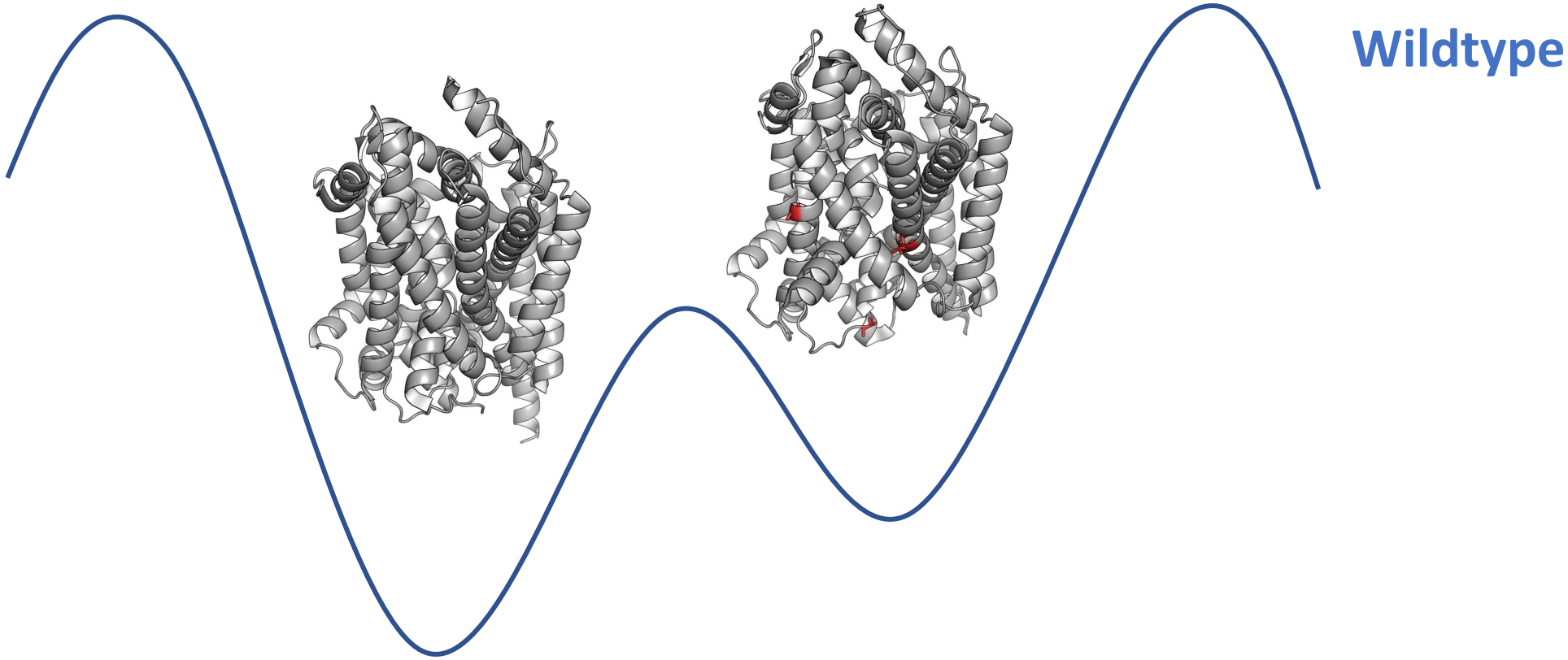
- Crystallized at 1.65 Å resolution
- No mutations added



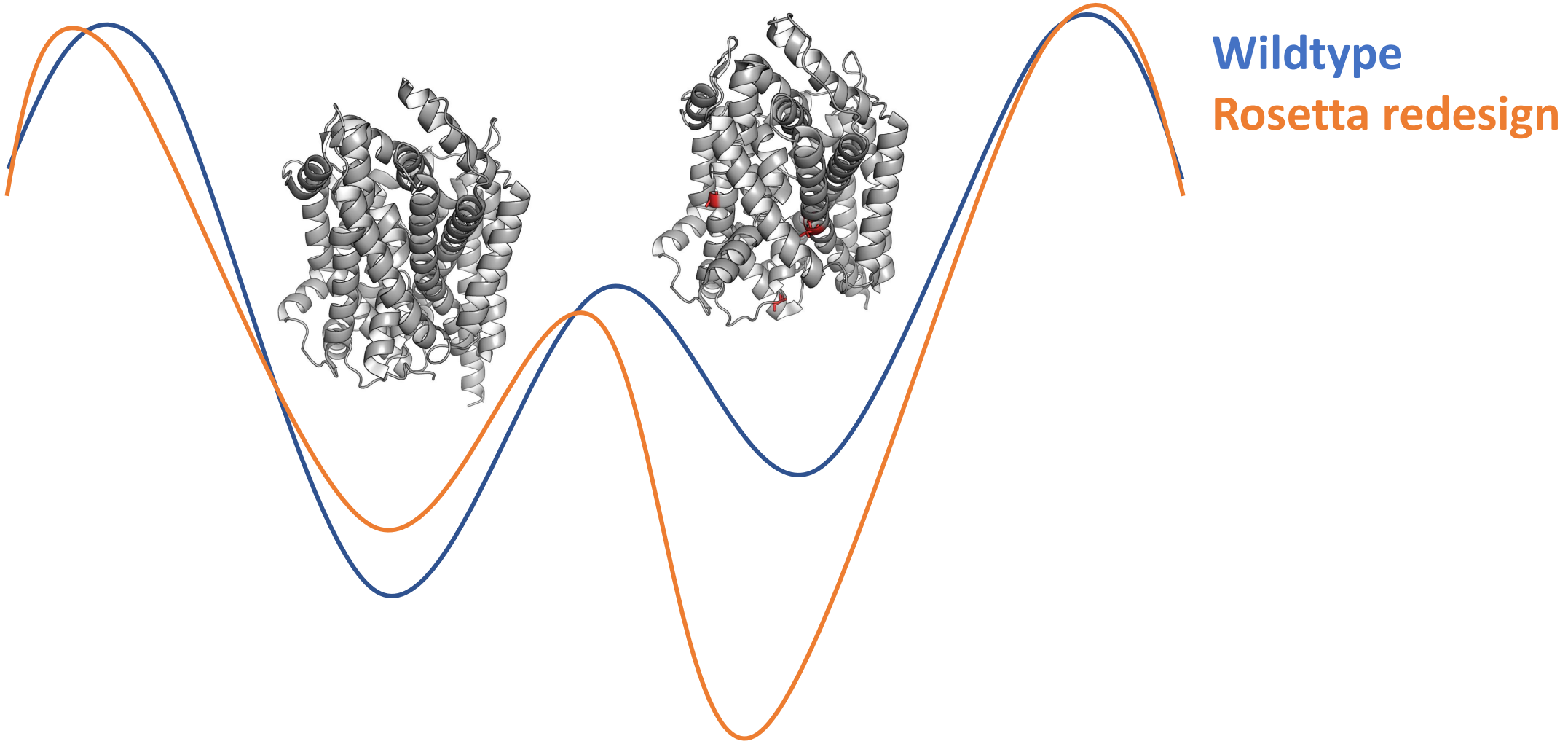
Cytoplasmic-facing

- Crystallized at 3.3 Å resolution
- Four mutations added

Today’s test case: bacterial leucine transporter



Today’s test case: bacterial leucine transporter



Today's test case: stabilize cytoplasmic-facing conformation

- First: Identify the lower-energy conformer
- Second: Introduce the four thermostabilizing mutations used to crystallize the periplasmic-facing conformation
- Third: Identify new mutations at those sites
- Fourth: Redesign the whole protein
- Fifth: Redesign the whole protein for an aqueous environment (outside the membrane)