

Structure Determination using Experimental Data

Marion Sauer

marion.f.sauer@vanderbilt.edu

Rosetta Workshop

Introduction

- Determination of X-ray crystal structures at moderate to high resolutions has recently accelerated
- Vast improvements in the resolutions achievable by single-particle cryo-electron microscopy due to improved image data analysis
- Currently available X-ray crystallographic tools perform relatively poorly with density-map resolutions worse than 3Å

Different Methods

- Relax into Density Map
 - Iterative local rebuilding
 - Rosetta.Phenix Refinement
- } Model Refinement Methods
-
- RosettaCM
 - *De novo* model building
(not available yet)
- } Model Building Methods

EM Density Maps

Scenario0

```
<ROSETTASCRIPTS>
  <SCOREFXNS>
    <dens weights=talaris2014_cart>
      <Reweight scoretype=elec_dens_fast weight=20.0/>
    </dens>
  </SCOREFXNS>

  <MOVERS>
    <SetupForDensityScoring name=setupdens/>
    <LoadDensityMap name=loaddens mapfile="1lissA_6A.mrc"/>
    <FastRelax name=relaxcart scorefxn=dens repeats=1 cartesian=1/>
  </MOVERS>

  <PROTOCOLS>
    <Add mover=setupdens/>
    <Add mover=loaddens/>
    <Add mover=relaxcart/>
  </PROTOCOLS>

  <OUTPUT scorefxn=dens/>
</ROSETTASCRIPTS>
```

```
elec_dens_fast=20.0
elec_dens_window=2.0
```

EM Density Maps

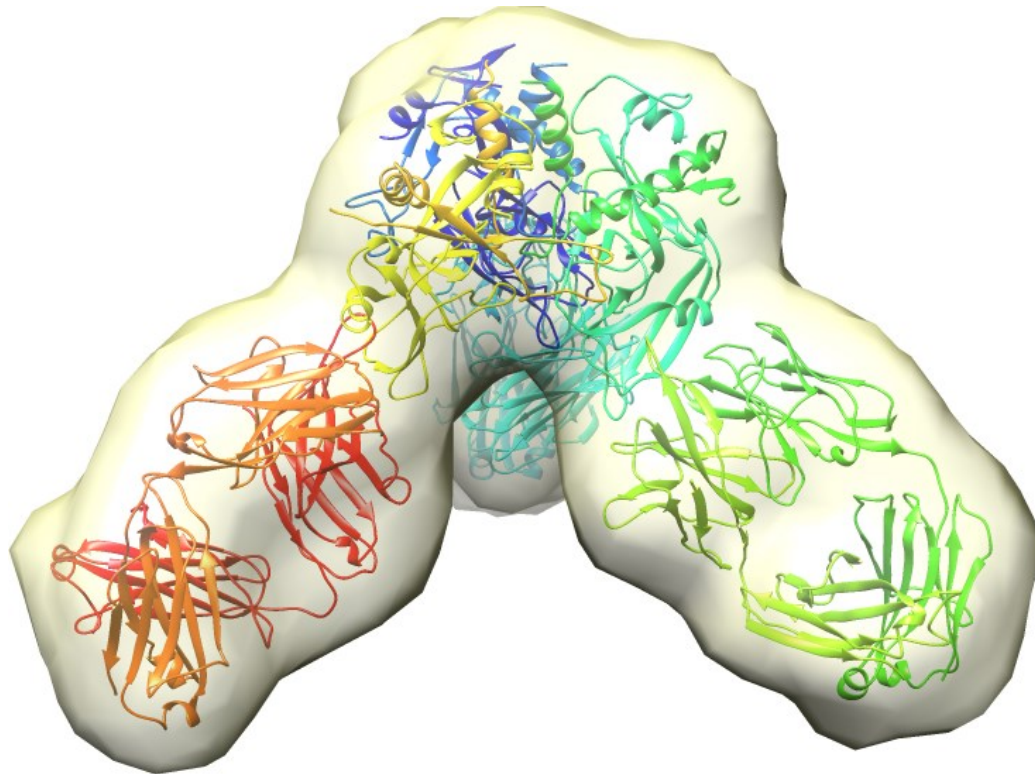
Scenario0

```
~/rosetta_workshop/rosetta/main/source/bin/rosetta_scripts.linuxgccrelease \  
-database ~/rosetta_workshop/rosetta/main/database/ \  
-in::file::s 1lsrA.pdb \  
-parser::protocol ex_B1_run_RS_relax_density.xml \  
-ignore_unrecognized_res \  
  
-edensity::mapreso 5.0 \  
-edensity::cryoem_scatterers \  
-crystal_refine \  
-out::suffix _relax \  
-default_max_cycles 200
```

X-ray scattering factors is the default option.

EM Density Maps

Scenario0



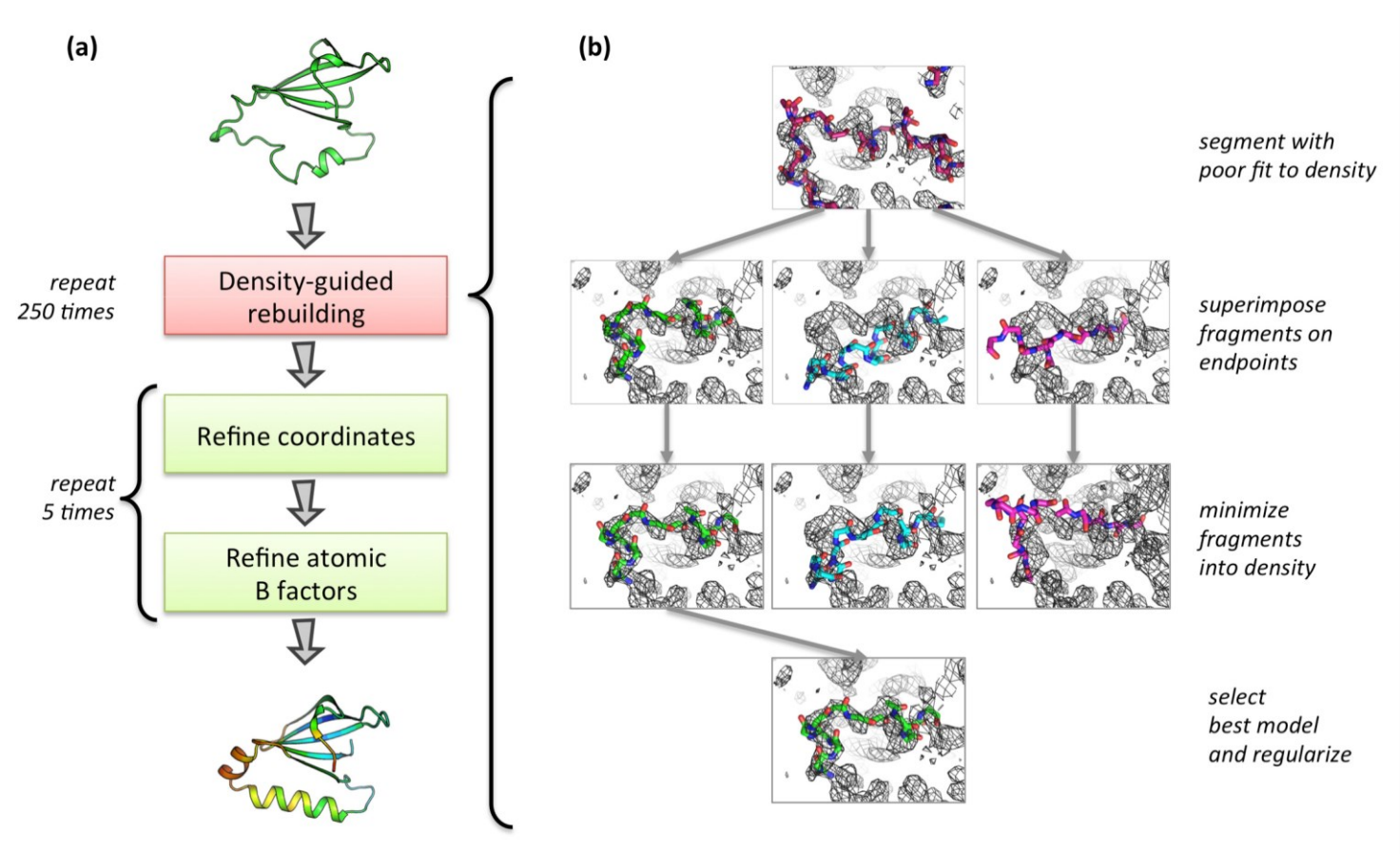
<http://www.cgl.ucsf.edu/chimera/data/tutorials/emfit09/emfit.html>

Iterative Local Building

scenario1

- Monte Carlo sampling with local density-guided optimization
- Rosetta all-atom refinement
- Real-space B-factor fitting

This method integrates approaches from crystallographic refinement, *ab initio* structure prediction, and segment rebuilding and refinement in comparative modeling.



Iterative Building

scenario1

```
<ROSETTASCRIPTS>
  <SCOREFXNS>
    <cen weights="score4_smooth_cart">
      <Reweight scoretype=elec_dens_fast weight=20/>
    </cen>

    <dens_soft weights="soft_rep">
      <Reweight scoretype=cart_bonded weight=0.5/>
      <Reweight scoretype=pro_close weight=0.0/>
      <Reweight scoretype=fa_sol weight=0.0/> # membrane protein
      <Reweight scoretype=elec_dens_fast weight=25/>
    </dens_soft>

    <dens weights=talaris2014_cart>
      <Reweight scoretype=elec_dens_fast weight=25/>
      <Reweight scoretype=fa_sol weight=0.0/> # membrane
        protein
      <Set scale_sc_dens_byres="R:0.76,K:0.76,E:0.76,
        D:0.76,M:0.76,C:0.81,Q:0.81,H:0.81,N:0.81,
        T:0.81,S:0.81,Y:0.88,W:0.88,A:0.88,F:0.88,
        P:0.88,I:0.88,L:0.88,V:0.88"/>
    </dens>
  </SCOREFXNS>
```

<MOVERS>

<SetupForDensityScoring name=setupdens/>

<LoadDensityMap name=loaddens mapfile="./trpv1_half1.mrc"/>

**<MinMover name=cenmin scorefxn=cen type=lbfgs_armijo_nonmonotone
max_iter=200 tolerance=0.00001 bb=1 chi=1 jump=ALL/>**

**<CartesianSampler name=cen5_50 automode_scorecut=-0.5 scorefxn=cen
mcscscorefxn=cen fascorefxn=dens_soft strategy="auto"
fragbias="density" rms=2.0 ncycles=200 fullatom=0 bbmove=1
nminsteps=25 temp=4 fraglens=3 nfrags=25/>**

<FastRelax name=relaxcart scorefxn=**dens** repeats=1 cartesian=1/>

</MOVERS>

<PROTOCOLS>

<Add mover=setupdens/>

<Add mover=loaddens/>

<Add mover=cenmin/>

<Add mover=relaxcart/>

<Add mover=**cen5_50**/>

<Add mover=relaxcart/>

<Add mover=relaxcart/>

</PROTOCOLS>

<OUTPUT scorefxn=dens/>

</ROSETTASCRIPTS>

Iterative Local Building

scenario1

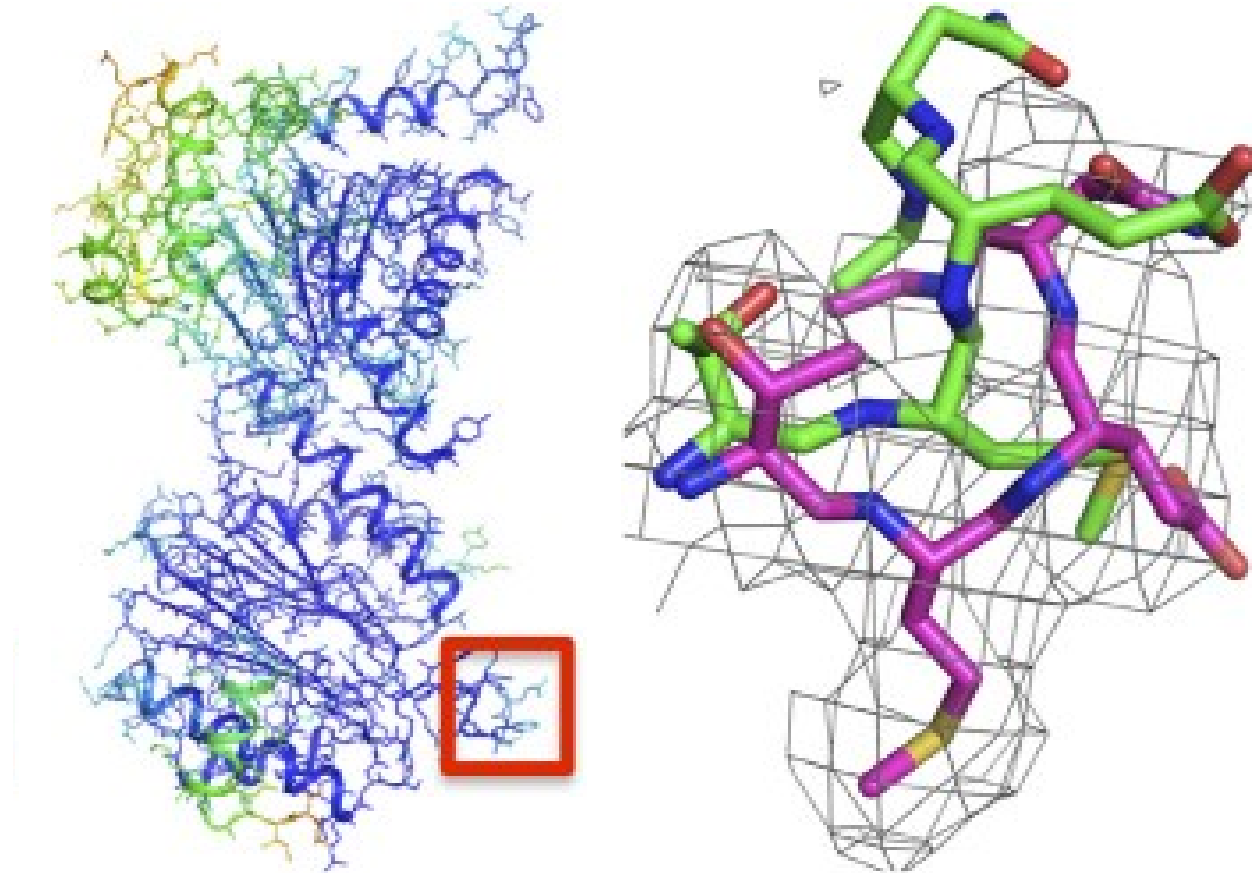
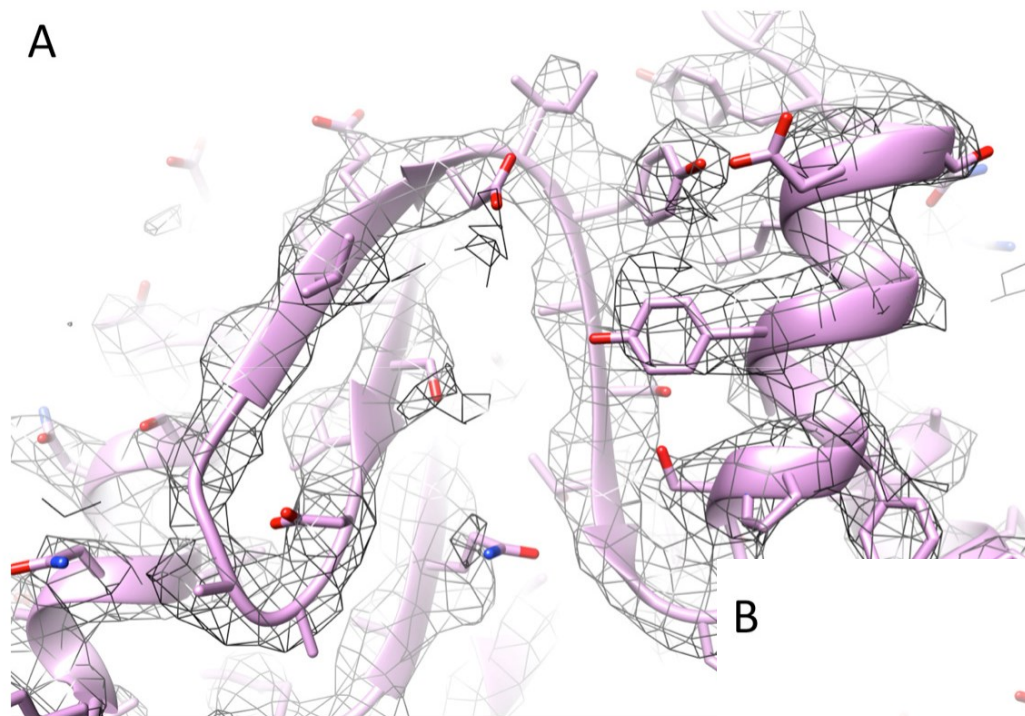


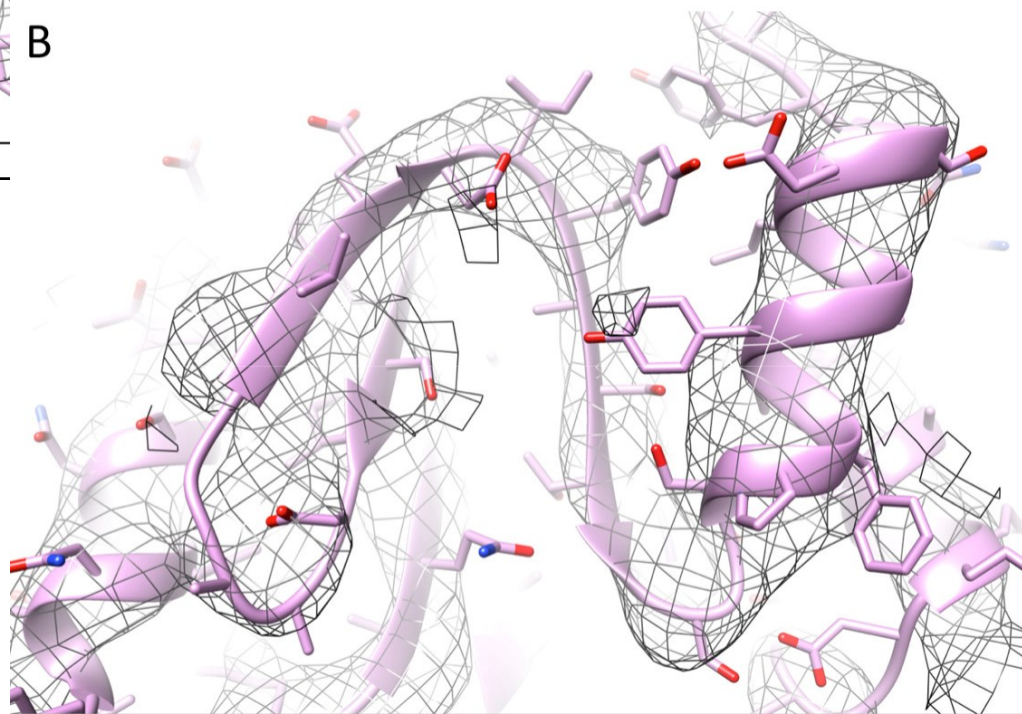
Figure: Refinement of 20S proteasome crystal structure into high-resolution cryo-EM map

DiMaio et al., 2015 (12), Nature Methods, 361-365.

A



B



RosettaCM with Density Maps

scenario2

Combine pieces from multiple templates guided by density

- Get the templates using BLAST like algorithms
- Align and partially thread the templates sequences onto target
- Prepare the symmetry file
- Combine pieces from different templates using Hybridize
- Full atom refinement and relax

RosettaCM with Density Maps

Scenario2

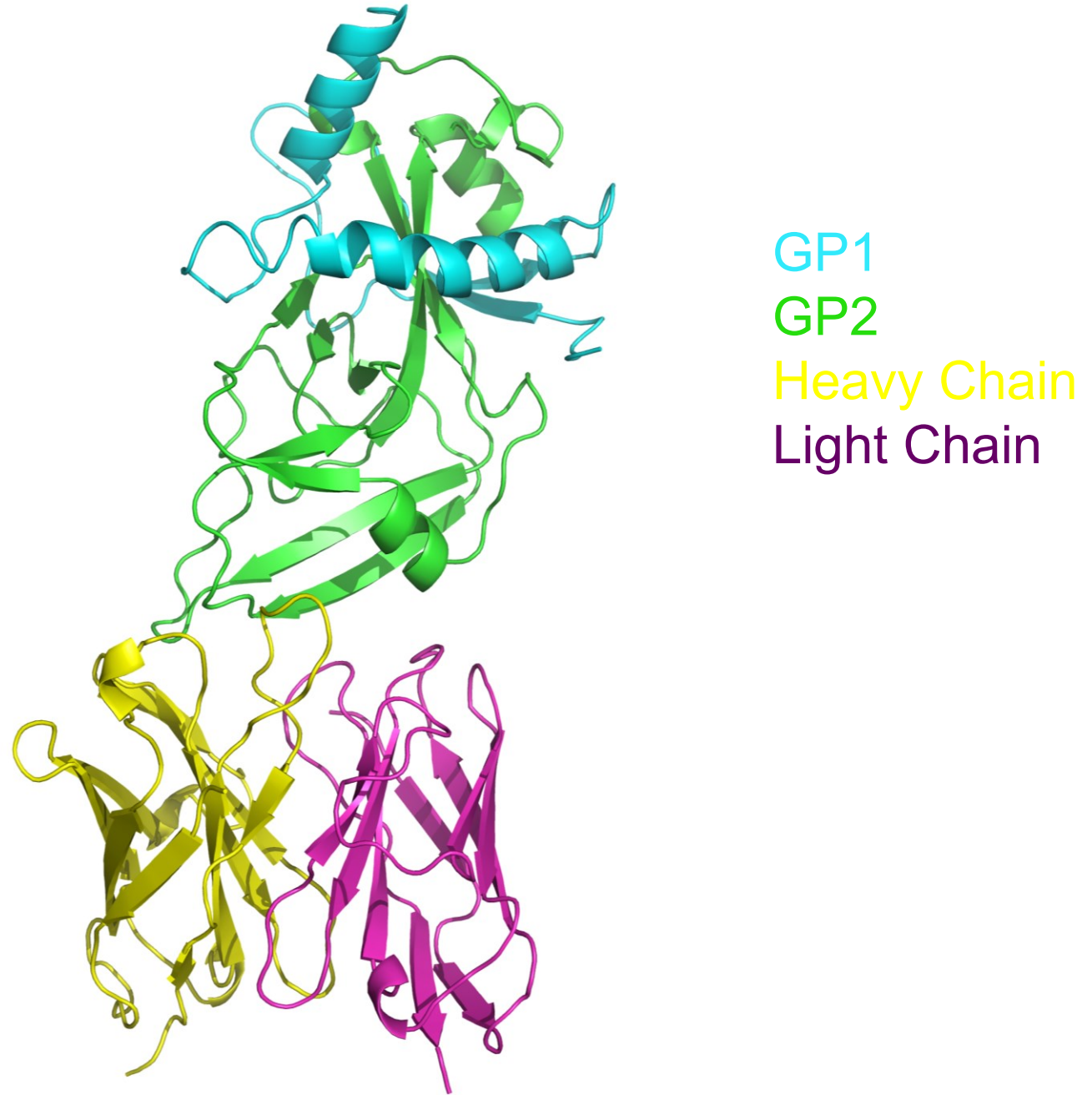
```
<ROSETTASCRIPTS>
  <SCOREFXNS>
    <stage1 weights="score3" symmetric=1>
      <Reweight scoretype=atom_pair_constraint weight=0.25/>
    </stage1>
    <stage2 weights="score4_smooth_cart" symmetric=1>
      <Reweight scoretype=atom_pair_constraint weight=0.25/>
    </stage2>
    <fullatom weights="talaris2014_cart" symmetric=1>
      <Reweight scoretype=atom_pair_constraint weight=0.25/>
    </fullatom>
  </SCOREFXNS>
  <MOVERS>
    <Hybridize name=hybridize stage1_scorefxn=stage1
      stage2_scorefxn=stage2 fa_scorefxn=fullatom batch=1>
      <Template pdb="1liruA_thread.pdb" weight=1.0
        cst_file="auto" symm_file="D7_edit.symm"/>
    </Hybridize>
  </MOVERS>
  <PROTOCOLS>
    <Add mover=hybridize/>
  </PROTOCOLS>
</ROSETTASCRIPTS>
```

RosettaCM with Density Maps

Scenario2

```
~/rosetta_workshop/rosetta/main/source/bin/rosetta_scripts.default.linuxgccrelease \  
-database ~/rosetta_workshop/rosetta/main/database \  
-in:file:fasta t20s.fasta \  
-parser:protocol ex_B1_rosettaCM_singletarget.xml \  
-nstruct 50 \  
  
-edensity::mapfile t20S_41A_half1.mrc \  
-edensity::mapreso 5.0 \  
-edensity::cryoem_scatterers \  
  
-relax:minimize_bond_angles \  
-relax:min_type lbfgs_armijo_nonmonotone \  
-relax:jump_move true \  
-relax:default_repeats 2 \  
-default_max_cycles 200
```

Marburg Virus GP in complex with MR78 antibody



Refinement using Phenix.Rosetta

- Method combines the Rosetta sampling methodology and energy function with X-ray refinement in Phenix
- Rosetta force-field constrains refinement to physically plausible conformations; density maps restrain Rosetta side-chain and backbone sampling in real-space

Refinement using Phenix.Rosetta

```
<PROTOCOLS>
```

```
  <Add mover=setup_opts/>  
  <Add mover=setup_symm/>
```

```
  <Add mover=tag_i/>  
  <Add mover=stage1cycle/>  
  <Add mover=stage1cycle/>  
  <Add mover=stage1cycle/>  
  <Add mover=stage2cycle/>  
  <Add mover=stage2cycle/>  
  <Add mover=stage2cycle/>  
  <Add mover=stage2cycle/>  
  <Add mover=stage2cycle/>  
  <Add mover=stage3cycle/>  
  <Add mover=stage3cycle/>  
  <Add mover=tag_f/>
```

```
  <Add mover=extract_asu/>  
</PROTOCOLS>
```

```
<ParsedProtocol name=stage1cycle>  
  <Add mover=recompute_dens/>  
  <Add mover=screpack_soft/>  
  <Add mover=set_cryst_wt_tors/>  
  <Add mover=min_tors_xtal/>  
  <Add mover=fit_bs/>  
  <Add mover=tag_c/>
```

```
</ParsedProtocol>
```

```
<ParsedProtocol name=stage2cycle>  
  <Add mover=recompute_dens/>  
  <Add mover=screpack/>  
  <Add mover=min_dens/>  
  <Add mover=set_cryst_wt_tors/>  
  <Add mover=min_tors_xtal/>  
  <Add mover=fit_bs/>  
  <Add mover=tag_c/>
```

```
</ParsedProtocol>
```

```
<ParsedProtocol name=stage3cycle>  
  <Add mover=recompute_dens/>  
  <Add mover=screpack/>  
  <Add mover=set_cryst_wt_cart/>  
  <Add mover=min_cart_xtal/>  
  <Add mover=fit_bs/>  
  <Add mover=tag_c/>
```

```
</ParsedProtocol>
```

```
</MOVERS>
```

Refinement using Phenix.Rosetta

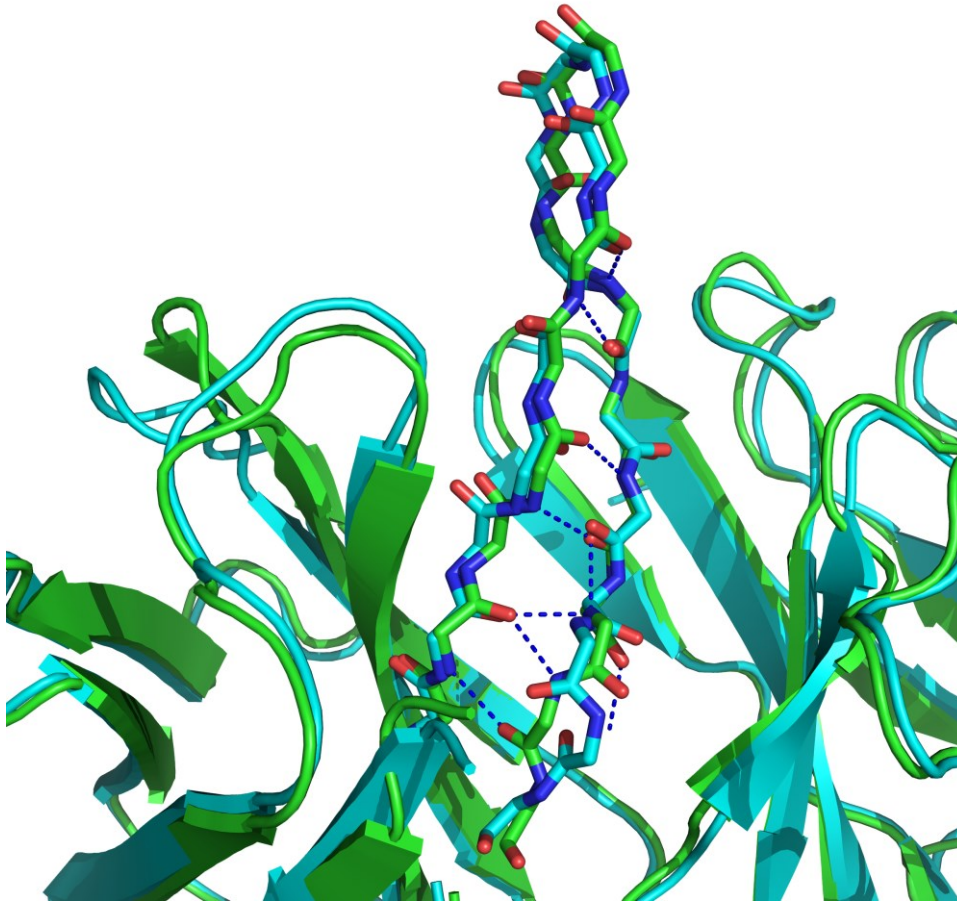


Figure: Phenix.Rosetta refined structure of Marv+MR78 complex

De novo Structure Prediction

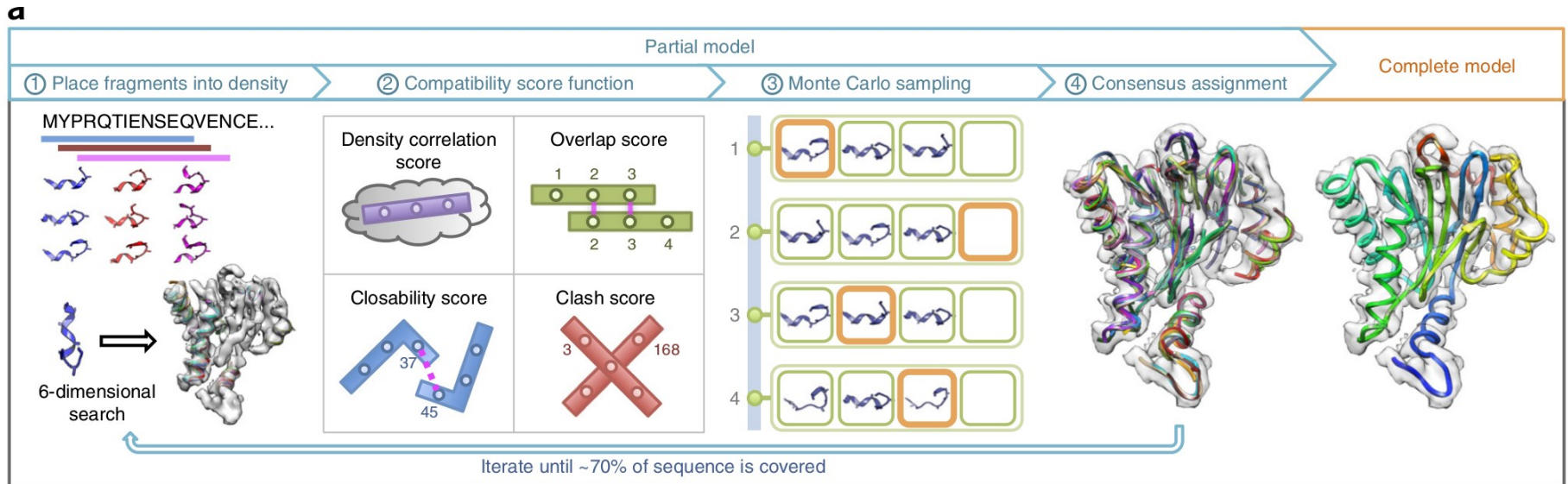


Figure: Various steps in *de novo* model building using density maps.

Wang et al, 2015, Nature Methods, 335-338.

De novo Structure Prediction

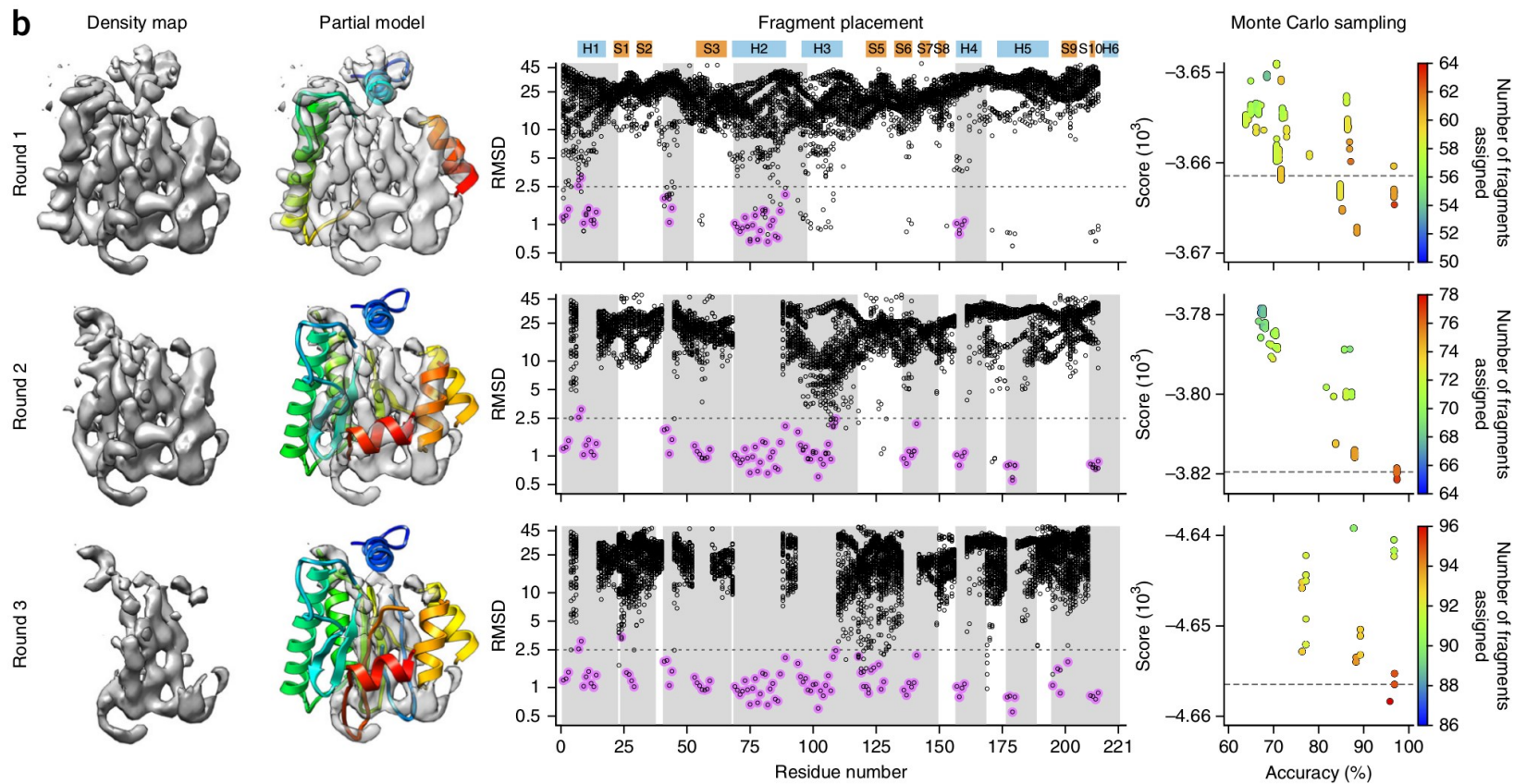


Figure: Partial model building steps

De novo Structure Prediction

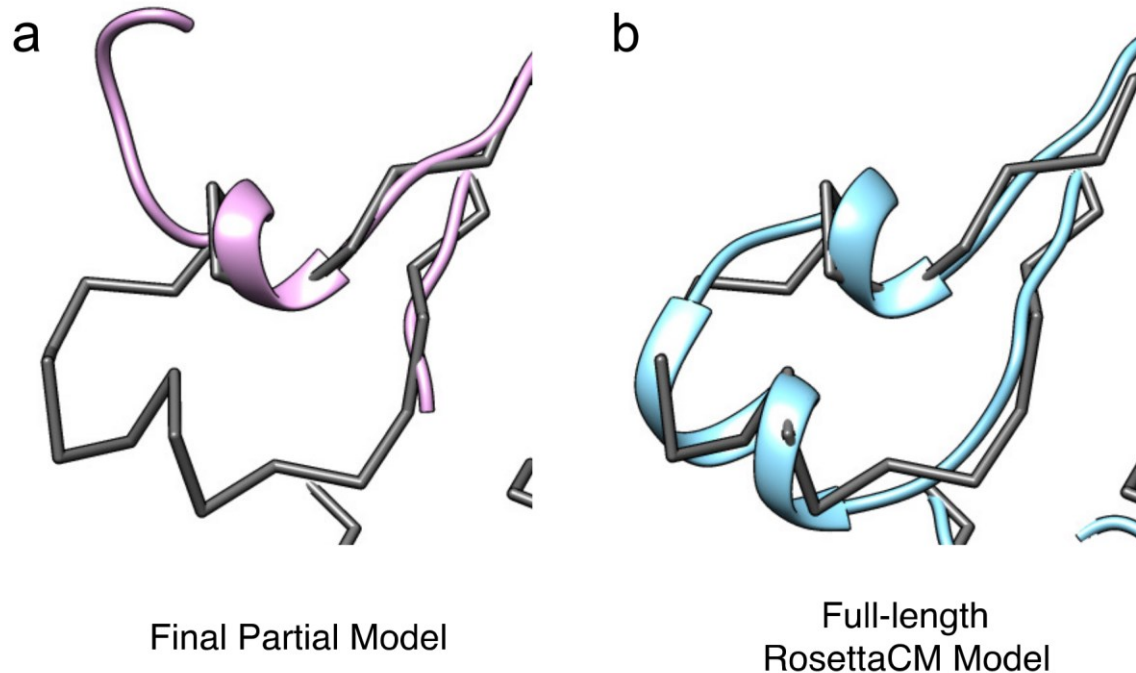


Figure: Partial model completed by RosettaCM.