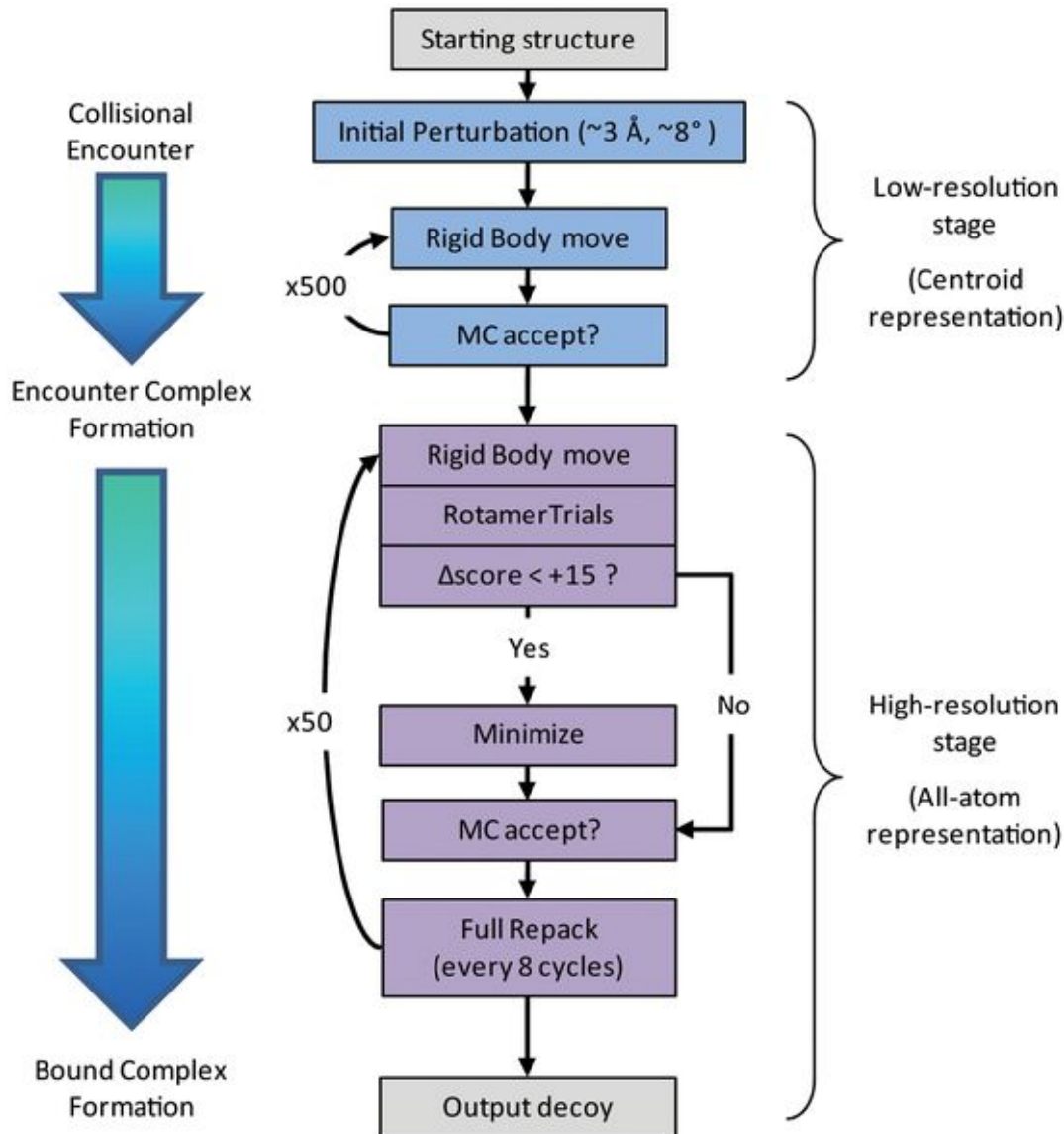# Protein-Protein Docking

Nina Bozhanova

Rosetta Workshop
November 7, 2018

Overview - Protein Docking

- Mimics natural process of protein interaction
- Centroid-based coarse grain stage
- Fine conformational search with side chain flexibility

- Hints:
  - Incorporate experimental data to decrease conformational search space

Chaudhury *et al*, PlosONE, 2011

General Docking Workflow

Prepare input files
- –Input structures
- –RosettaScripts XML File
- –Options File

Perform docking calculations

Analyze results
- –Plot Score vs. RMSD
- –Examine interface character

Example System: Influenza Hemagglutinin bound by mAb CR6261

- Crystallized both in complex with H1 and H5 hemagglutinin (HA)

- Cross-docking experiment – antibody from one PDB (3GBN) and HA from another (3GBM)

- Influenza H5 hemagglutinin
  - 2 chains: HA1, HA2

- CR6261
  - Human derived monoclonal antibody
  - 2 chains: Heavy and Light

Developing your docking XML file

Please open the docking_full.xml file

gedit
~/rosetta_workshop/tutorials/protein-protein_docking/input_files
/docking_full.xml

Where should you start?
```
<PROTOCOLS>
        <Add mover="dock_low"/>
        <Add mover="srsc" />
        <Add mover="dock_high" />
        <Add mover="minimize_interface" />
</PROTOCOLS>
```

Docking movers

```
<Docking name=(string) score_low=(string) score_high=(string)
fullatom=(bool) local_refine=(bool) jumps=(Integer vector)
optimize_fold_tree=(bool) conserve_foldtree=(bool) design=(bool)
ignore_default_docking_task=(bool) task_operations=(comma-separated list)/>
```

Terms	Definition
score_low	scorefxn for centroid-level docking
score_high	scorefxn for full atom docking
fullatom	if true, do full atom docking
local_refine	if true, skip centroid.
jumps	where should we carry out rigid body motions (not used here)
optimize_fold_tree	should DockingProtocol make the fold tree for this pose
conserve_fold_tree	should DockingProtocol reset the fold tree to the input one after it is done
design	interface design for all chains downstream of the rb_jump
ignore_default_docking_task	allows you to ignore the default docking task and only use the ones defined in your task_operations section

Docking movers

```
<PROTOCOLS>
            <Add mover="dock_low"/>
            <Add mover="srsc" />
            <Add mover="dock_high" />
            <Add mover="minimize_interface" />
</PROTOCOLS>
```

```
<Docking name="dock_low" score_low="score_docking_low" score_high="REF2015"
fullatom="0" local_refine="0" ignore_default_docking_task="0" design="0"
task_operations="ifcl,prfrp" jumps="1"/>
```

```
<Docking name="dock_high" score_low="score_docking_low" score_high="REF2015"
fullatom="1" local_refine="1" design="0" task_operations="ifcl,prfrp"
jumps="1"/>
```

```
<SaveAndRetrieveSidechains name="srsc" allsc="0" /> Speeds the move from
centroid to full atom mode
```

Important options

Option Terms	Definition
docking:partners AB_HL	sets rigid body docking partners – chains must be in this order in the PDB
docking:randomize1	randomize the first docking partner – should be passed to run global docking (not used in this tutorial)
docking:dock_pert 3 8	set coarse perturbation parameters (angstroms and degrees)
docking:dock_mcm_trans_magnitude 0.1	refinement translational perturbation
docking:dock_mcm_rot_magnitude 5.0	refinement rotational perturbation

TaskOperations

```
<TASKOPERATIONS>
    <InitializeFromCommandline name="ifcl"/>
    <RestrictToRepacking name="rtr" />

    <RestrictToInterfaceVector name="rtiv" chain1_num="1,2"
chain2_num="3,4" CB_dist_cutoff="10.0" nearby_atom_cutoff="5.5"
vector_angle_cutoff="75" vector_dist_cutoff="9.0" />

    <PreventResiduesFromRepacking name="prfrp"
residues="11,41,345" />

</TASKOPERATIONS>
```

Minimization movers

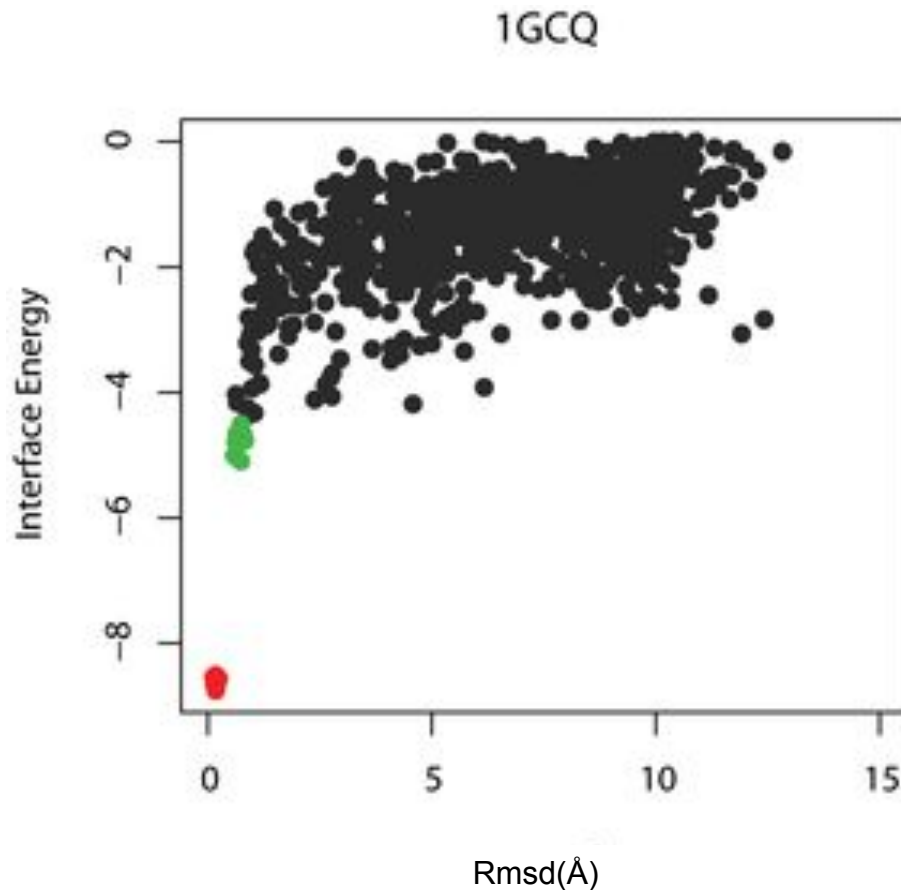Use a single round of FastRelax to refine the backbone of both docking partners

<FastRelax name="minimize_interface" scorefxn="REF2015" repeats="1" task_operations="ifcl,rtr,rtiv,prfrp" />

•Restrict to repacking – default behavior is design (very important)

•Interface from command line – allows the mover to read options from the command line

Analysis

- Score vs. RMSD plots to identify a docking funnel
    - In experimental studies, the 'native structure' is a good scoring structure, one that fulfills experimental restraints, or a cluster center
- It may be necessary to cluster and dock again, using the top cluster centers as starting templates. This will help to drive the models into an energy well. The 'correct' model will continue down into an energy well while incorrect models will stall.
- Look at the structures
    - Unsatisfied polar residues
    - Binding density
    - Average Degree of Residue Burial

Fleishman, S. J.; Baker, D., Role of the biomolecular energy gap in protein design, structure, and evolution. *Cell* **2012,** *149* (2), 262-73
Fleishman, S. J.; Whitehead, T. A.; Strauch, E. M.; Corn, J. E.; … ; Baker, D., Community-wide assessment of protein-interface modeling suggests improvements to design methodology. *Journal of molecular biology* **2011,** *414* (2), 289-302.

Score vs. RMSD plots – identifying a docking funnel

1GCQ

- Key features:
  - Native complex is kept at low RMSD, low scoring
  - Good models are sampled (< 2 A RMSD)
  - Score function distinguishes between low and high RMSD models

Gray – Full docking protocol
Green – Top models
Red – Refinement of Native complex

Today's Tutorial

```
cd
~/rosetta_workshop/tutorials/protein-p
rotein_docking/

firefox
protein-protein_docking_tutorial.html
```

Reference List

- **Docking**
    - Gray, J. J.; Moughon, S.; Wang, C.; Schueler-Furman, O.; Kuhlman, B.; Rohl, C. A.; Baker, D., Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *Journal of molecular biology* **2003,** *331* (1), 281-99.
    - Daily, M. D.; Masica, D.; Sivasubramanian, A.; Somarouthu, S.; Gray, J. J., CAPRI rounds 3-5 reveal promising successes and future challenges for RosettaDock. *Proteins* **2005,** *60* (2), 181-6.
    - Chaudhury, S.; Sircar, A.; Sivasubramanian, A.; Berrondo, M.; Gray, J. J., Incorporating biochemical information and backbone flexibility in RosettaDock for CAPRI rounds 6-12. *Proteins* **2007,** *69* (4), 793-800.
    - Sircar, A.; Chaudhury, S.; Kilambi, K. P.; Berrondo, M.; Gray, J. J., A generalized approach to sampling backbone conformations with RosettaDock for CAPRI rounds 13-19. *Proteins* **2010,** *78* (15), 3115-23.
    - Sivasubramanian, A.; Sircar, A.; Chaudhury, S.; Gray, J. J., Toward high-resolution homology modeling of antibody Fv regions and application to antibody-antigen docking. *Proteins* **2009,** *74* (2), 497-514.
    - Sircar, A.; Gray, J. J., SnugDock: paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models. *PLoS computational biology* **2010,** *6* (1), e1000644

Reference list continued

- **Analysis**
  – Fleishman, S. J.; Baker, D., Role of the biomolecular energy gap in protein design, structure, and evolution. *Cell* **2012,** *149* (2), 262-73.
  – Fleishman, S. J.; Khare, S. D.; Koga, N.; Baker, D., Restricted sidechain plasticity in the structures of native proteins and complexes. *Protein science : a publication of the Protein Society* **2011,** *20* (4), 753-7.
  – Fleishman, S. J.; Whitehead, T. A.; Strauch, E. M.; Corn, J. E.; … ; Baker, D., Community-wide assessment of protein-interface modeling suggests improvements to design methodology. *Journal of molecular biology* **2011,** *414* (2), 289-302.