Rosetta Basics:
IO and Navigating Rosetta

Shannon Smith
shannon.t.smith.1@vanderbilt.edu
Graduate Student
Meiler Lab, Vanderbilt University

Goals for this Talk

1. General Rosetta Concepts:
   - How do I run basic Rosetta applications?
   - Input/Output: file types, options, etc.

2. Learn where things are in Rosetta
   - Your working directory is independent of these Rosetta directories
     - (AKA your data is stored outside of Rosetta)
   - Highly encourage everybody to follow along from your command line
   - Full path = path to file from home directory
     - (~ = /usr/people/molgraph/)
   - Relative path = path to file from current working directory

- Please stop me for questions!

How do we get Rosetta?

- https://www.rosettacommons.org/software/license-and-download
- Weekly Releases: (e.g. “2016.37”)
  - Latest version of the code, released roughly every week
  - Every revision passes scientific performance tests
- Numbered Releases (e.g. “3.9”)
  - A weekly release that’s relabeled, released roughly every 6 months
- All tutorials use version 2017.36.59679 (May 2018)

- Links to documentation, forum and demos:
  - https://www.rosettacommons.org/docs/latest/Home
  - https://www.rosettacommons.org/demos/latest/Home

cd ~/rosetta_workshop/tutorials/short_talks/RosettaIO/

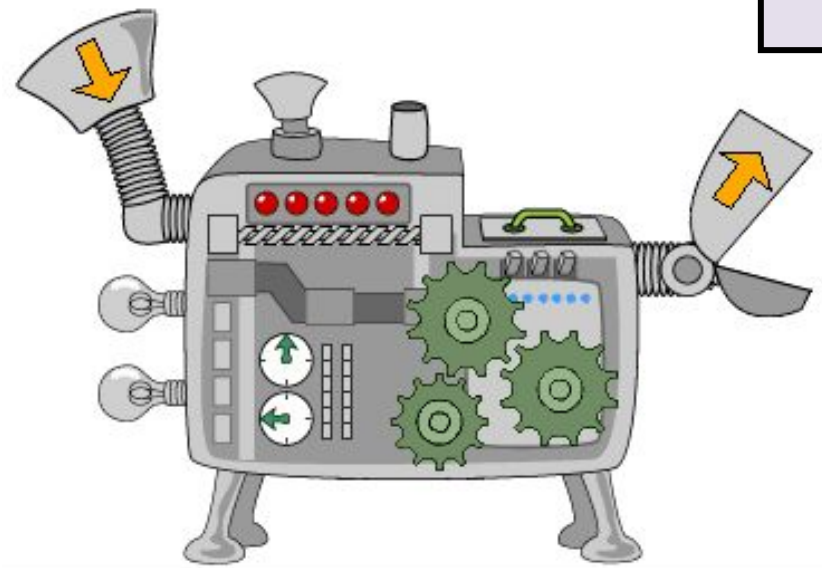
- Where you will find all the files associated with this talk
- input_files/
- output_files/
- scripts/

- Notice these files are located outside of Rosetta (aka you do NOT want to store your input/output files in directories within ~/rosetta_workshop/rosetta/*) !!

Tip: Open each file in your favorite text editor (gedit, vi, emacs, etc.) as we introduce them throughout the talk

Basic Pipeline for Running a Rosetta Application

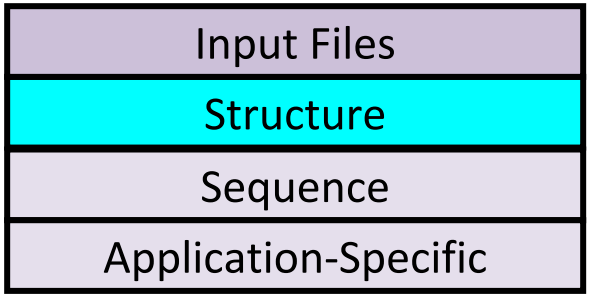
Input Files
Structure
Sequence
Application-Specific



Rosetta Application

Output Files
Structures
Scores
Log/Tracer

Structure Files: PDB



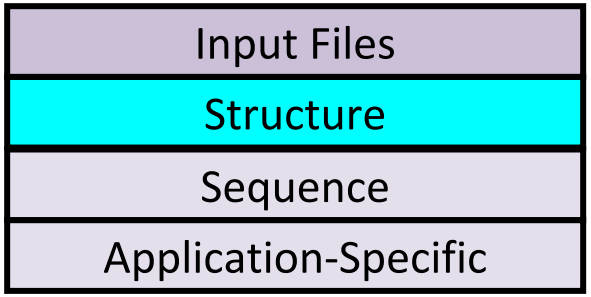
- PDBs are broken down by atom lines
- Each atom is characterized by the following (left to right across each row)
  - Atom, atom number, atom name, residue name, chain identifier, residue sequence number, x, y, z orthogonal coordinates, occupancy, temperature factor, element symbol, atom charge.
- Note: CIF not supported

Snippet from 1qys.pdb

```
419 ATOM     57  N   ILE A  10       3.073  -2.269  16.587  1.00 40.88           N
420 ATOM     58  CA  ILE A  10       3.976  -3.337  16.187  1.00 40.88           C
421 ATOM     59  C   ILE A  10       3.531  -4.697  16.730  1.00 40.88           C
422 ATOM     60  O   ILE A  10       3.465  -4.900  17.944  1.00 70.73           O
423 ATOM     61  CB  ILE A  10       5.399  -3.068  16.702  1.00 61.80           C
424 ATOM     62  CG1 ILE A  10       5.870  -1.693  16.253  1.00 61.80           C
425 ATOM     63  CG2 ILE A  10       6.343  -4.127  16.180  1.00 61.80           C
426 ATOM     64  CD1 ILE A  10       7.241  -1.341  16.788  1.00 61.80           C
427 ATOM     65  N   ASP A  11       3.225  -5.621  15.824  1.00 68.81           N
428 ATOM     66  CA  ASP A  11       2.802  -6.960  16.220  1.00 68.81           C
429 ATOM     67  C   ASP A  11       3.892  -7.970  15.878  1.00 68.81           C
430 ATOM     68  O   ASP A  11       4.569  -7.850  14.850  1.00 68.81           O
```

Structure Files: Silent File

- Rosetta-specific structure file-type
- Useful when storing many structures
  - Easy conversion between PDB <--> Silent files



Snippet from 1qys.out

```
SEQUENCE: DIQVQVNIDDNGKNFDYTYTVTTESELQKVLNELMDYIKKQGAKRVRISITARTKKEAEKFAAILKVFAELGYNDINVTFDGDTVTVEGQLwwwwwwwX
SCORE:     score     fa_atr     fa_rep     fa_sol    fa_intra_rep    fa_elec     pro_close    hbond_sr_bb    hbond_lr_bb    hbond_bb_sc    hbond_sc    dslf_fa13    coordinate_constraint       rama       omega
REMARK BINARY SILENTFILE
SCORE:  -145.658    -416.906     48.038    235.048           1.023    -47.764        0.000        -25.252        -27.431         -4.739     -10.754        0.000                   19.154     -4.561      4.169
RES_NUM A:3-94 A:107-113 z:1 1qys_0001
FOLD_TREE  EDGE 100 48 8  EDGE 48 1 -1  EDGE 48 92 -1  EDGE 1 93 1  EDGE 1 94 2  EDGE 1 95 3  EDGE 1 96 4  EDGE 1 97 5  EDGE 1 98 6  EDGE 1 99 7  1qys_0001
RT 0.0482407 0.667083 0.74342 0.668251 0.531625 -0.520399 -0.74237 0.521896 -0.420133 -23.3245 8.31606 -8.68819   1qys_0001
RT -0.396168 0.11999 0.910304 0.116778 0.989958 -0.0796675 -0.910722 0.0747416 -0.406202 -26.5704 -2.74183 -1.85648   1qys_0001
RT 0.628098 -0.724121 -0.284853 0.188735 -0.213372 0.958568 -0.754898 -0.655836 0.0026483 -9.05427 1.73199 -13.4752   1qys_0001
RT 0.920183 0.0507995 0.388179 0.347555 -0.562384 -0.750286 0.180192 0.825314 -0.535152 -4.59601 4.10013 7.78226   1qys_0001
RT 0.216437 0.133709 -0.967097 0.963388 -0.189798 0.189366 -0.158233 -0.972676 -0.169893 -19.6994 17.4952 0.560284   1qys_0001
RT 0.69639 -0.500263 0.514565 -0.0851439 0.654345 0.751388 -0.712595 -0.567071 0.413085 -14.5353 -4.17559 -6.15321   1qys_0001
RT 0.794518 -0.141481 0.590529 0.448149 0.792838 -0.413003 -0.409762 0.592783 0.693328 -3.98858 7.73926 -6.37077   1qys_0001
RT -0.0824693 0.014772 -0.996484 0.91817 -0.387663 -0.0817348 -0.387507 -0.921683 0.018407 2.46349 -1.61798 -0.14082   1qys_0001
ANNOTATED_SEQUENCE: D[ASP:NtermProteinFull]IQVQVNIDDNGKNFDYTYTVTTESELQKVLNELMDYIKKQGAKRVRISITARTKKEAEKFAAILKVFAELGYNDINVTFDGDTVTVEGQL[LEU:CtermProteinFull]wwwwwwwX 1qys_0001
CHAIN_ENDINGS  92 1qys_0001
LlotiAbKNVG0oHjYQjDjOAvKDSGEY8nYQVWlIAv/EIGkkmOYQncfWAzsaCGUcR8XQevMJAnLIQGUqHXZQneRKAfeSaGEo6xZQ1NkJAnD0iGEtvfZQUkyLATXbZGknSYaQXHJkArjxbGE6d0YQsVRkALWtWGECUEYQMLanAbQQPG0mVtYQ8zjEAXPjYG0OBcYQPTUUAnpSKG0vc
L9Qbn/uM9FGkW5LYQHZvN/+Dt5FEJ4mXQwGTH/OHTlF0vDOYQEqQZ+sjRjFkpqqYQip9K/0M9+FEudBXQoQ4G/8QtJGkWvLWQRd8l/sNWsF0NFMWQuEtv+qdIJGktOCVQLyRF/GlzKGk4QbYQDy9s/G3U2F0LryWQt66r/gSABGU0a1XQwG4p+c0aQGkZc0WQhHCz/0LcLGkxk
LeVtv/6CQWFUIsDYQ/oUz/aEMCFUbwaYQOqpb/iKkvEUq/CYQfKMs/yTGpEUqQBXQjuQRAvCB9EEm2iYQ718UA/tooEU1S8YQB+PmAPxYkEEEaEZQ3HOtAXWTvEkR7yYQobioAb5dUEExBfZQxRpBAncBZFkbKVXQWOrj/G2+DFU+t5YQ0JtYAPXMKFE6L0YQHekZAHAb7E0wW
LngMD+kNFnEEl6VYQWtTZ/Q2hVEEyEDYQ1l+K/QWeAEUs3aYQqxfg/Uti8DkcRAZQGIYGAdA6aE04PDYQ9nMTAtAyIEkqjeXQ/XPKAhJyvE0g6WXQC+62+0jHtE0AayYQzw2E/02hTE0I5DXQmvwLAFnhcEEPRkYQSJFiAhY2MEUbWfXQ+FpQA9cR0D0TgBYQwwAOAVvXGEUqq
L/9Tt9gfFiDk/oEYQ2ZUN+SZu3Ck4QWYQu2xK/M7rWCkaf8XQwA4G/c9ZTCUwduWQyuR1/SdnsC0bFTYQ32dAAznVBCU55lYQgAaYA7AuyBUdvlYQKLkiArGVWCUpRuYQ8FseA3v4jAkrAcYQITQO+GFeoDkSTMXQB6Om9oUC4CUgM4YQoLBEATc8EDkkmjYQl8D//aVysC0/N
L66Av/86Z9BkRJWYQy52FARRk9A06FEYQ3DEAAVcwJ/zxqVYQ44QBAV9os+ziH8YQkVSdAxk9SBU1APYQGVImAhcBQAEHc4XQunbhA1r3VCUe58XQZBYu/Yx+CCEJL2YQERTEAp49+AERrCXQp0TfAFOORBkmpxYQ7DYuAl3IgAUYQEYQ4NLkA1yui+jTZJYQsPOlAhimOAkbF
LuqVz/gc2n8DFfzXQemCo/Ebev874HIYQPdKDAZbLh/Ll7rXQ0zEEANTLAAMHyeWQNnMQ+6Cq+97js4XQwxRl/eVPN1ri8QYQdF+7/G3cx8z/21XQp05m/CklD2rdy6YQzTVx/w2ng9DQj0WQuMRr/kve58Lk3qYQbQUn+yzkw9r9YzWQ9zK3+ew0AAcIGGYQwOH7/uUFL8z+O
LqlPPAR1fYA8yuSYQTb5eARylZBMdcHYQ3bMYABISZCcg7WYQ4sIXAp7bhCsrd9YQ0LfqAhVBDB8dKaYQkuUuApLQZ/L64IYQ6WjyAtWfEC8t8PYQ26Y4Ati1X+r3IeYQBNVNA9fXIA80yxYQZ5TgAdciaB8UhJXQGo2pAxeZ8A81x8YQv90vA9vxj/LLTNXQhlDoANqkz9rMR
LH4pTANwz1CcHYxXQIZ1MAlHIhD8AhDYQF32eANfVAEM4a0XQNSxkABZ59D8sByWQs2hx/8LmtDMBvUXQ1ZPr+cdQSDM7WrXQvJAx+U7t8C8UwVYQH8pH/yULTD8rR8WQv77UARIbtCcTKzWQfoIJA1gNiDMPmlYQfjx3/oKgqDMbXQWQYiyp/wThHE8HPhXQ 1qys_0001
LfPEhAthbPEMXcXYQZzwpAl2neEsPSSYQXQdoAFPQyEscsuYQ3mquAN4f1EsJiMZQ+Ge0ARxyTEslDcYQ4uD+ANbshEcsnQYQmrf8A5pJyEs8o6XQ8PbDBVPubEczSbYQz2mYAtPgQEsodxYQ0J2pARfFkEsGTiXQgwB1A9j0EE8uSKYQ4830Alb4PEsZB+YQ 1qys_0001
L9wXgAlYN/E8BVlYQYEGeARaHUF8Ht5YQolbcAlP1TFcYaqZQH8KmA1oSXFMXoAaQAmKoANKsiFccVsYQbKToAxO5mF8q75XQm8DgAdaUsFMq4TXQJ3fxA9zfkFMvGTXQ2a7VA9qY6E8bVPYQgpJPAZ3KbFsRcuYQ0UxvAZpebFMdj1YQa1fnAJo6xFM1I9YQMFEyAtEFnFM4E
LT3GJA9DAQFcSs6ZQXqvFABsiOFMjroaQpILOARVe6E86H+aQfIiOAxuE5EspclbQeYB5/kDEOFMg0mZQr/Vp/EfcPFMAmvaQAcTNAh5IcFslZ4aQ 1qys_0001
Lh7AVA9I8qE8SekaQALAeAF5MXEsHd1aQgP5WAtgQEEs+WaaQrsRgAdhA/D8Dg5ZQnyNrAp5aYEsAF3aQLUlvABBPpE8DuVbQajv7AFcpoEsvZXbQz9DABtpj5EM7u1bQOpAGBFqR5Esad3bQVW9TAJC3sE86cEaQo1GYAN7+UEM5/VbQNwTuAZV9bEc1OXaQEEcuAhI8IE8ur
LYe0FAtVMzDcXKoaQ6oF7/452NDs3pRaQsoUEA9UtkC8VSpaQql9BAZOGhCsQVQbQGoMr+kofQD8lMNaQfqn19qVv2DMlqzZQSc8I/M6dDEc2iYZQPGsp/CxvCEsVz7ZQdKy8/QvO9DMZDDbQQvTEAFSbND86dxZQJtG89WRGSDcd4saQNN0B9C6lzCsNr9ZQEOz0/SdcPEMUy
LDwENAV2DGCsEcRaQPiYUA9FC8AsfGkaQznJHAhcQo/Le8QaQ3bBDApZsm/7onqZQMXGmAt4UtA8KhbaQdHKtAxwr0BMECraQJ5lvANRpcCcxmQaQm3XxApPZzBMhEUbQEbF2Alff9CMS2eaQEyY33AdVdaCMOYibQ4dO6AxrZ8C8EuHbQpwrOANusKCsZnxZQjEISAZUO/Asir
LZGRAAdIUi9bJJsaQRq9k/U9Bi5z9JfaQJMG2/UAuI/DhwwaQGyMx/Mi/v/jtgWbQUNXj+ysnOyLtvuaQpCxn/SXNJ+T4fgaQJyTS/2CYGAU8WSaQkS2GAr5qd9DF5jaQM23EAZ2U09rmCLbQMSlk/0vQJ6TGT8ZQuzkI/Kq459rMtiaQ9/ok+SZ4M27PmRbQ 1qys_0001
LEBQFAZfzWAURMVaQKMZOAdvxpB0HXhaQXWS//A8NZC0JVTaQIRJy/0vabCU3cuZQU/ViA1PP1BE77NaQ9USqAp3hxAkTobaQJvLrAZOLM/zTgEaQMArwAR+3+AUfT/aQCBcyAhQ/g8zuARaQ9N73AduCBAUzzLbQEV04AhN+X9jfz0aQZXBABFlHz3rrQBbQY4rGAl6kCAESo
LSv/0/4jo1CUeMxaQRn4U/c+laDkc4naQ7Vpz/sxQBEUL6yaQi6yBApudDEUjrWbQts3/+G9mZDkL9AbQxOpn/2w+zC0Up1aQ87pq/GshAEEa62aQgVAAAZxVxCUnRPbQh4wC/AzgZDUUNGaQie/k+idvYDEnKjbQN0WR/epvkCUQjfaQkzPEA/H0/DkB7IbQZLxR/is2OE0V0
LtJE4/shfPEk4hTaQ7VLIAtIXjEkWtYaQwjw0/Aku1E0+mIaQjDJU/s+DzE0TNsZQwIvcAR+3jEUol9ZQUmBnAF1eTEEpxGaQH8imAhwx9DEyV1ZQb1GwAZTHZEETVeaQLg2uABQ/gD0WE7ZQgBd4AJz0KEk/IkaQXty3AJ6NtDkiSSaQckGABREORDky6WaQnOBq/4uWMEkvG
L0gS8/QQgIF03saaQ09zl/w3rbF0R1NaQGAdEAlaAsFEy4IaQ5BfQAlA8vFUN4maQGZLA+cZBjFE/juaQ0KeV/GnTSFkR80aQOyeF/WhC3F0I0eaQCINJAlUbJFEUMyaQA9kG/Q/HYF01IxZQ6bFH/44dmFkeBNbQ0twg/WDgLFEuDaaQZUfq/ul/7F0kW2aQRXHE+Yv6BGkFS
LswZGAdeE1FkuxiZQfJPWAxH2CGE/7bZQHroMAZ2qMGUP6GZQTYJ2/w6SMGUHt0YQvpSkApnI9F0A1/YQLjtpANPLqFkQVVZQ8hAgANN52FEzAUYQgu55/oPjwFkzKKZQY7ucAxeNFGEs66ZQkaUqA5jzEGkOP8YQHSLWAhXVkFkzBBZQM99sAd2JvFkMH0ZQt14jAZsbdFUCP
LchdYAtjdVGEFHKZQCPCSA9QyfGUI34YQiIUjAFqllGEQstYQvGWlAxYPuG0fP+YQLVkEA5gWmGkBsYZQldTQAx5wnG0as/ZQ8bTe/YxrgGkATiZQwXUjAlD1UGk9aZZQalgHAx0aeGkQldYQc/KBAVkMuGEZELZQ22eSAB2QvGk5vEaQzeJy+4OclGE904ZQy7P0+UgrfGEGw
Lfr8pAJWzgG0weSYQyA8zA9Z+lGUxuFYQgwC3AZ+hjGkEXvWQsJ35A5/uqGEiw+VQ/uJ9AR0oiGkxlJYQvlk+A1KRXGUunhYQVtv6ABs5lGEteRZQ+0IoAhMiZG02WGYQTUCzAN4luGEdwJYQ2CRCBBlfmGUXQZYQUeB6ANe5TGUMZ2YQiArABd1ejGkfHmZQCJw5A9ajuG0gJ
LUCi2AF+TZGUBKYWQzv85ANOpVGkH0DVQkREDBJpjVGEcGKVQab6FBBEUUGERkKUQEYC2AV39cG0W17TQ74DqARYKfGEUC9TQv4djAVuRVG0ZLuTQPc7mAZ9eOGk3t+SQkM6VABIhUGE0WSUQjFrzAdc7TGk1lCXQqgc2AlmyNG0LE4UQh0K6A5bokGUZr/TQ8iA4AVJaZGE+H
LQA+EBFU7WGU33YWQRNpKBhdmXGkChsWQv+jLBlF8OG0rpvXQrCNPBl5YIGkNTnXQo+LMBpSkiGUI2NXQZrpRBVDAjG0pghXQJNSCBdFhXG0w8IXQmV5MB5cdWG0faxVQlSPLBZXmoG0w1dWQbP2JBhZZkG0hZDYQIy6RB9d6mGU/GLYQ 1qys_0001
```

PDB files

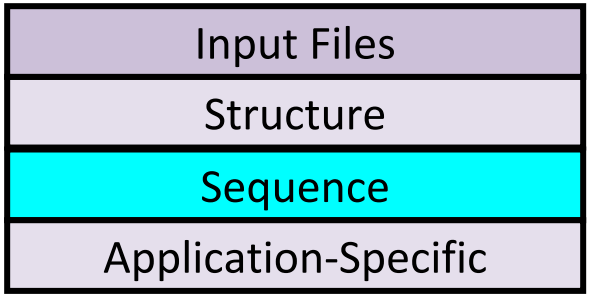
- International standard
- Useful for a small number of structures
- One line per atom
- Readable by PyMol, MOE, Chimera, etc
- Must be processed with clean_pdb.py prior to use in Rosetta

- Full info: www.wwpdb.org/documentation/file-format

Silent files

- Specific to Rosetta
- Useful for archiving many structures
- One line per residue
- Binary files: more compact, but not human-readable
- Silent --> PDB using $ROSETTA/main/source/bin/extract_pdbs.linuxgccrelease

- Full info: https://www.rosettacommons.org/docs/latest/rosetta_basics/file_types/silent-file



Sequence Files:

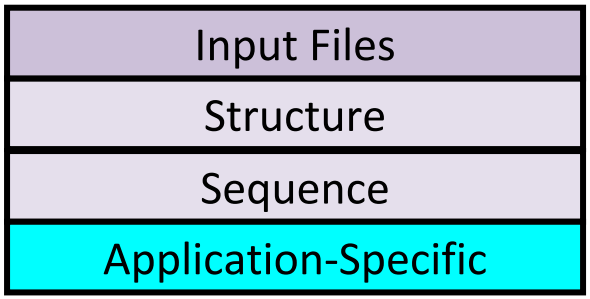
- fasta: string of 1-letter amino acid names
- clustal: common for multiple sequence alignments
- grishin: RosettaCM-specific

1qys.fasta

```
>1QYS:A|PDBID|CHAIN|SEQUENCE
MGDIQVQVNIDDNGKNFDYTYTVTTESELQKVLNELMDYIKKQGAKRVRISITARTKKEA
EKFAAILIKVFAELGYNDIN
VTFDGDTVTVEGQLEGGSLEHHHHHH
```

Application-Specific Files:

- More details later during tutorials



- Res Files: Indicates which residue positions should be designed
- Params File: Custom parameters for small molecules or non-canonical amino acids
- Span File: Defines which residues are in the membrane
- Loops File: Identifies the loop residues for loop closure
- Constraint File: Experimentally derived restraints
- Fragment File: Short protein segments used for comparative modeling and de novo folding

Now we have our input files...what do we do now?

$ROSETTA/main/source/bin/<your favorite Rosetta application> -option1 <argument1> -option2 -option3 <argument3a> <argument3b>...

Note: Some options can take one or multiple arguments, while some don't take any.

Now we have our input files...what do we do now?

$ROSETTA/main/source/bin/<your favorite Rosetta application> -option1 <argument1> -option2 -option3 <argument3a> <argument3b>...

Note: Some options can take one or multiple arguments, while some don't take any.

```
>> $ROSETTA/main/source/bin/AbinitioRelax.linuxgccrelease
-in:file:fasta ./input_files/1qys_A.fasta -in:file:native
./input_files/1qys.pdb -in:file:frag3
./input_files/aa1elwA03_05.200_v1_3 -in:file:frag9
./input_files/aa1elwA09_05.200_v1_3 -abinitio:relax -relax:fast
-abinitio::increase_cycles 10 -abinitio::rg_reweight 0.5
-abinitio::rsd_wt_helix 0.5 -abinitio::rsd_wt_loop 0.5
-use_filters true -psipred_ss2 ./input_files/1elwA.psipred_ss2
-kill_hairpins -out:file:silent ./output_files/1qys_10.out
-out:file:scorefile ./output_files/1qys_silent_scores.sc -nstruct
10
```

Above command-line taken directly from: Barth, P., Schonbrun, J., Baker, D. *Towards high-resolution prediction and design of transmembrane helical protein structures* (2007).

Now we have our input files...what do we do now?

$ROSETTA/main/source/bin/<your favorite Rosetta application> -option1 <argument1> -option2 -option3 <argument3a> <argument3b>...

Note: Some options can take one or multiple arguments, while some don't take any.

```
>> $ROSETTA/main/source/bin/AbinitioRelax.linuxgccrelease
-in:file:fasta ./input_files/1qys_A.fasta -in:file:native
./input_files/1qys.pdb -in:file:frag3
./input_files/aa1elwA03_05.200_v1_3 -in:file:frag9
./input_files/aa1elwA09_05.200_v1_3 -abinitio:relax -relax:fast
-abinitio::increase_cycles 10 -abinitio::rg_reweight 0.5
-abinitio::rsd_wt_helix 0.5 -abinitio::rsd_wt_loop 0.5
-use_filters true -psipred_ss2 ./input_files/1elwA.psipred_ss2
-kill_hairpins -out:file:silent ./output_files/1qys_10.out
-out:file:scorefile ./output_files/1qys_silent_scores.sc -nstruct
10
```

Above command-line taken directly from: Barth, P., Schonbrun, J., Baker, D. *Towards high-resolution prediction and design of transmembrane helical protein structures* (2007).

$ROSETTA/main/source/bin/AbinitioRelax.linuxgccrelease @options.txt

- Contains all the same options as the command line, just in a more user-friendly fashion
- Easier to read, edit, log, etc.

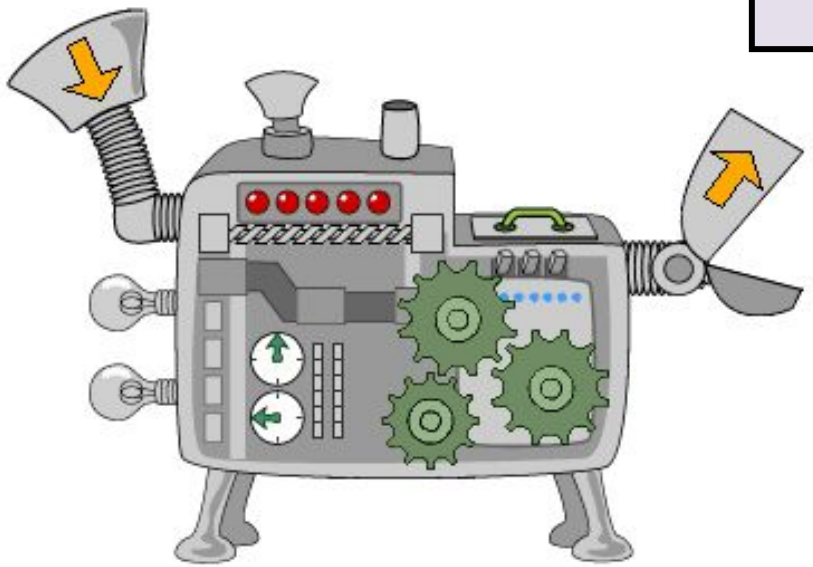
options.txt

```
## This is a comment

## input files
-in:file
        -fasta ./input_files/1qys_A.fasta ## specifies location of fasta sequence file
        -native ./input_files/1qys.pdb ## specifies location of native PDB; "native" files are often used to compare final output structures
        -frag9 ./input_files/aa1elwA03_05.200_v1_9 ## not in this directory, but are 9-mer residue fragments generated on Robetta server
        -frag3 ./input_files/aa1elwA03_05.200_v1_3 ## not in this directory, but are 3-mer residue fragments generated on Robetta server
        -psipred_ss2 ./input_files/1elwA.psipred_ss2 ## not in this directory, but secondary structure prediction outputs from Robetta server

## application-specific flags
-abinitio:relax
-relax:fast
-abinitio::increase_cycles 10
-abinitio::rg_reweight 0.5
-abinitio::rsd_wt_helix 0.5
-abinitio::rsd_wt_loop 0.5
-use_filters true
-kill_hairpins

## output options
-out:file:silent ./output_files/1qys_10.out
-out:file:scorefile ./output_files/1qys_silent_scores.sc
-nstruct 10
```

Basic Pipeline for Running a Rosetta Application

Input Files
Structure
Sequence
Application-Specific

Output Files
Structures
Scores
Log/Tracer



```
$ROSETTA/main/source/bin/AbinitioRelax.linuxgccrelease @options.txt
```

Output Files

Scores

Structures

Log/Tracer

Score Files:

- tab-delimited file containing all score terms for each output

./output_files/1qys_silent_scores.sc

1	SCORE:	score	fa_atr	fa_rep	fa_sol	fa_intra_rep	fa_elec		ret	time	description
2	SCORE:	-145.658	-416.906	48.038	235.048	1.023	-47.764	…	-12.643	121.000	1qys_0001
3	SCORE:	-145.660	-420.838	48.025	238.756	1.060	-46.720	….	-12.643	119.000	1qys_0002
4	SCORE:	-147.668	-414.585	46.523	235.308	1.033	-48.171	…	-12.643	115.000	1qys_0003
5	SCORE:	-147.324	-413.085	46.068	232.161	1.025	-46.810	…	-12.643	116.000	1qys_0004
6	SCORE:	-148.283	-423.099	47.967	241.560	1.046	-48.530	…	-12.643	110.000	1qys_0005
7	SCORE:	-147.020	-418.127	47.198	237.669	1.018	-48.364	…	-12.643	121.000	1qys_0006
8	SCORE:	-148.368	-417.523	46.620	235.293	1.049	-47.662	…	-12.643	111.000	1qys_0007
9	SCORE:	-147.644	-414.716	45.941	232.969	1.038	-46.471	…	-12.643	119.000	1qys_0008
10	SCORE:	-147.763	-416.130	46.386	235.877	1.023	-47.914	…	-12.643	118.000	1qys_0009
11	SCORE:	-146.560	-421.948	49.483	239.471	1.031	-49.106	…	-12.643	115.000	1qys_0010

Structure Files:

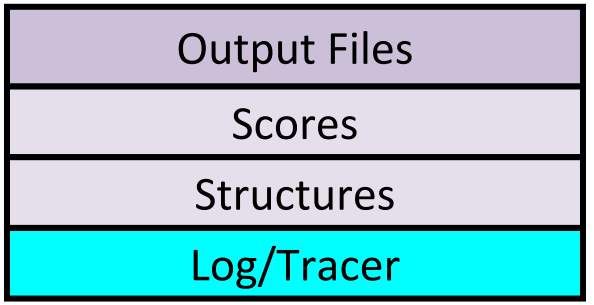
- PDB or silent file -- same as input files, but note that output PDBs contain per residue score terms after ATOM lines.

./output_files/1qys_0001.pdb

```
#BEGIN_POSE_ENERGIES_TABLE 1qys_0001.pdb
label fa_atr fa_rep fa_sol fa_intra_rep fa_elec pro_close hbond_sr_bb hbond_lr_bb hbond_bb_sc hbond_sc dslf_fa13 rama omega fa_dun p_aa_pp yhh_planarity ref linear_chainbreak overlap_chainbreak total
weights 1 0.55 0.9375 0.005 0.875 1.25 1.17 1.17 1.17 1.1 1.25 0.25 0.625 0.7 0.4 0.625 1 1.33333 1 NA
pose -423.638 49.1168 241.309 1.0398 -46.1456 0 -25.4907 -26.9979 -3.93408 -11.2344 0 -4.90455 4.21053 109.662 -13.6029 0.23024 -12.6427 0 0 -163.023
ASP:NtermProteinFull_1 -2.66563 0.27015 2.41637 0.02503 -0.26885 0 0 0 0 -0.32408 0 0 0.0102 1.89419 0 0 -1.63002 0 0 -0.27263
ILE_2 -5.61825 0.23689 2.80238 0.02592 -0.12286 0 0 0 0 0 0 -0.26155 0.00661 0.81411 -0.34843 0 1.0806 0 0 -1.38458
GLN_3 -4.84167 0.48106 4.18521 0.00631 -1.21224 0 0 0 0 -0.48562 0 0.06079 0.00207 2.96363 0.1074 0 -1.51717 0 0 -0.25024
VAL_4 -4.61075 0.23581 1.49989 0.01618 -0.54587 0 0 0 0 0 0 -0.24278 0.0034 0.03458 -0.42478 0 0.97964 0 0 -3.05468
GLN_5 -4.16136 0.16768 3.36407 0.00633 -1.06934 0 0 0 0 -0.60901 0 0.04214 0.02118 2.77942 0.12364 0 -1.51717 0 0 -0.85242
VAL_6 -5.59614 0.88153 1.82246 0.014 -0.69849 0 0 0 0 0 0 -0.2164 0.00792 0.02014 -0.46296 0 0.97964 0 0 -3.24828
ASN_7 -4.85911 0.20247 3.95038 0.00414 -1.45872 0 0 0 0 -0.60901 0 -0.0151 0.00501 3.51664 0.05469 0 -1.19118 0 0 -0.39979
ILE_8 -5.46549 0.36001 1.8693 0.02274 -0.68642 0 0 0 0 0 0 -0.22814 0.00246 0.20363 -0.48548 0 1.0806 0 0 -3.3268
ASP_9 -3.90212 0.1626 3.5745 0.00384 -1.65188 0 0 0 0 -0.46353 0 0.12839 0.04956 2.18701 0.36117 0 -1.63002 0 0 -1.18048
ASP_10 -3.84216 0.45653 3.22502 0.01376 -0.64368 0 0 0 0 -0.41813 0 0.55824 0.52659 1.7871 -0.93152 0 -1.63002 0 0 -0.89825
ASN_11 -1.70761 0.57979 1.45451 0.00575 0.08728 0 0 0 0 0 0 0.70672 0.03429 1.84768 -0.73549 0 -1.19118 0 0 1.08173
GLY_12 -0.80224 0.12458 0.5163 1e-05 0.04398 0 0 0 0 0 0 -0.35846 0.02351 0 -0.8684 0 0.17333 0 0 -1.1474
LYS_13 -3.51045 0.33392 1.99225 0.00495 -0.05999 0 0 0 0 0 0 0.39242 0.00549 1.4125 -0.12885 0 -0.35857 0 0 0.08368
ASN_14 -2.79919 0.07672 2.07492 0.00412 -0.26575 0 0 0 0 0 0 0.01568 0.00684 2.02524 0.20226 0 -1.19118 0 0 0.14967
PHE_15 -5.49746 0.25492 2.43742 0.03039 -0.62276 0 0 0 0 0 0 -0.19658 0.06517 2.03196 -0.15724 0 0.61937 0 0 -1.03481
ASP_16 -2.92446 0.08288 2.94318 0.02006 -0.24577 0 0 0 0 0 0 0.05442 0.00203 1.72485 -0.29358 0 -1.63002 0 0 -0.26641
TYR_17 -5.88808 0.44904 2.60429 0.03033 -0.15087 0 0 0 0 0 0 -0.09114 0.03337 2.10883 -0.04977 0.00033 0.1625 0 0 -0.79117
THR_18 -1.9813 0.05965 1.85734 0.00877 -0.21462 0 0 0 0 0 0 0.02411 0.047 0.04309 -0.12051 0 0.20134 0 0 -0.07513
TYR_19 -6.68966 0.80179 3.35153 0.0269 -1.08788 0 0 0 0 -0.41091 0 -0.20557 0.01282 1.67279 -0.30295 0.00237 0.1625 0 0 -2.66626
THR_20 -2.62243 0.31036 2.76908 0.00722 -0.10022 0 0 0 0 -0.16154 0 0.12349 0.04786 0.09622 -0.08035 0 0.20134 0 0 0.59103
VAL_21 -3.80377 0.75681 1.67726 0.01702 -0.122 0 0 0 0 0 0 -0.19285 0.00444 0.02274 -0.31811 0 0.97964 0 0 -0.97881
```

Log Files/Tracer:

- Printed version of what runs on the terminal
- Keeps track of what's occurring during a simulation-- VERY important for keeping records
- These files get large:
  - Silence certain tracers:
    - -mute core.chemical.ResidueTypeSet
  - Change verbosity level (Error/Warning/Info/Debug/Trace)
    - -out:levels all:Warning core.init:Info

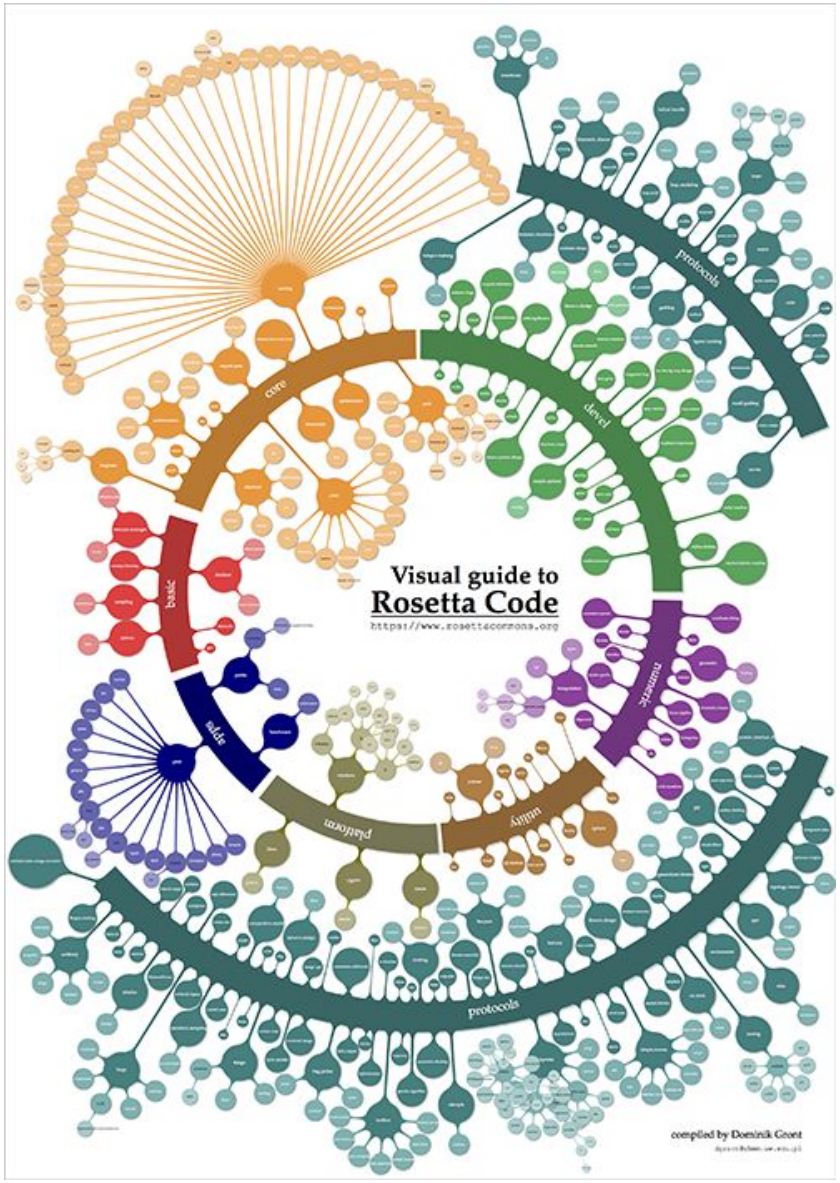


$ROSETTA/main/source/bin/AbinitioRelax.linuxgccrelease @options.txt >> run.log

```
core.init: Rosetta version exported  from http://www.rosettacommons.org
core.init: command: /dors/meilerlab/apps/rosetta/rosetta_2016.08.58479/main/source/bin/rosetta_scripts.default.linuxgccrelease @docking.options -parser:protocol docking
core.init: 'RNG device' seed mode, using '/dev/urandom', seed=1059677151 seed_offset=0 real_seed=1059677151
core.init.random: RandomGenerator:init: Normal mode, seed=1059677151 RG_type=mt19937
core.init: Resolved executable path: /dors/meilerlab/apps/rosetta/rosetta_2016.08.58479/main/source/build/src/release/linux/2.6/64/x86/gcc/5.2/default/rosetta_scripts.d
core.init: Looking for database based on location of executable: /dors/meilerlab/apps/rosetta/rosetta_2016.08.58479/main/database/
protocols.jd2.PDBJobInputter: Instantiate PDBJobInputter
protocols.jd2.PDBJobInputter: PDBJobInputter::fill_jobs
protocols.jd2.PDBJobInputter: pushed 3gbm_HA_3gbn_Ab.pdb nstruct indices 1 - 50
protocols.evaluation.ChiWellRmsdEvaluatorCreator: Evaluation Creator active ...
protocols.jd2.JobDistributor: Parser is present.  Input mover will be overwritten with whatever the parser creates.
protocols.jd2.PDBJobInputter: PDBJobInputter::pose_from_job
protocols.jd2.PDBJobInputter: filling pose from PDB 3gbm_HA_3gbn_Ab.pdb
core.chemical.ResidueTypeSet: Finished initializing fa_standard residue type set.  Created 384 residue types
core.chemical.ResidueTypeSet: Total time to initialize 0.43 seconds.
core.conformation.Conformation: Found disulfide between residues 7 461
...
protocols.rosetta_scripts.ParsedProtocol.REPORT: ============End report for ==================
protocols.rosetta_scripts.ParsedProtocol.REPORT: ============Begin report for ==================
protocols.rosetta_scripts.ParsedProtocol.REPORT: ============End report for ==================
protocols.jd2.JobDistributor: 3gbm_HA_3gbn_Ab_full_0050 reported success in 381 seconds
protocols.jd2.JobDistributor: no more batches to process...
protocols.jd2.JobDistributor: 50 jobs considered, 50 jobs attempted in 16297 seconds
```

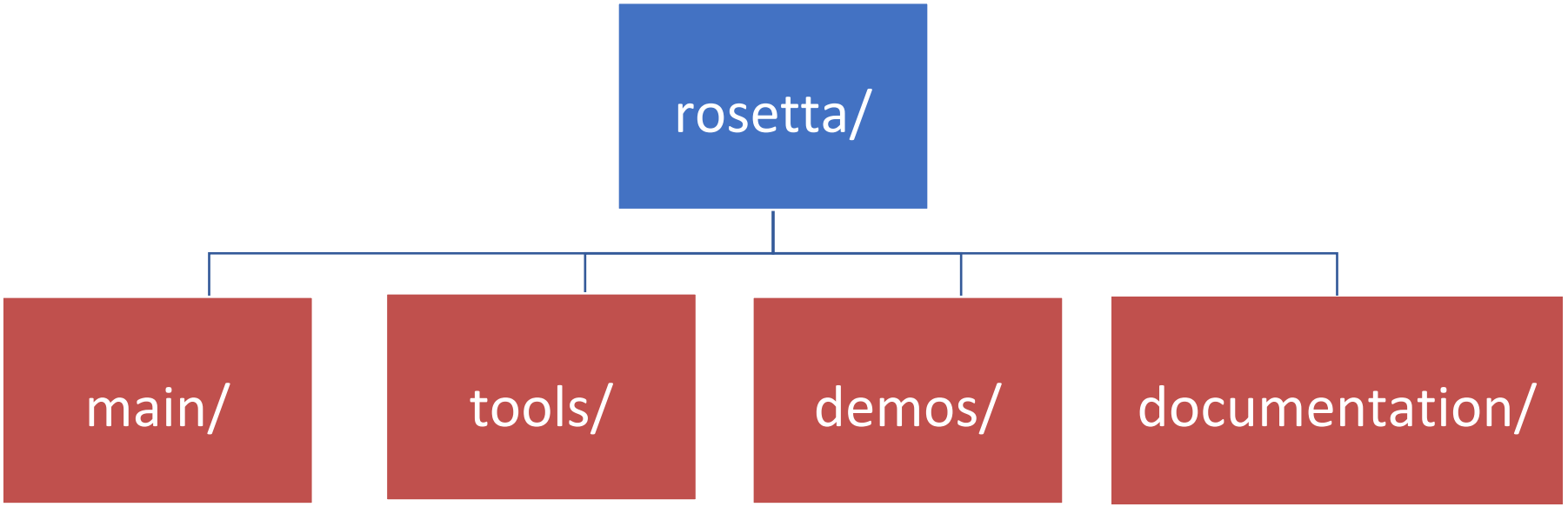
Any Questions?

With great power comes great responsibility



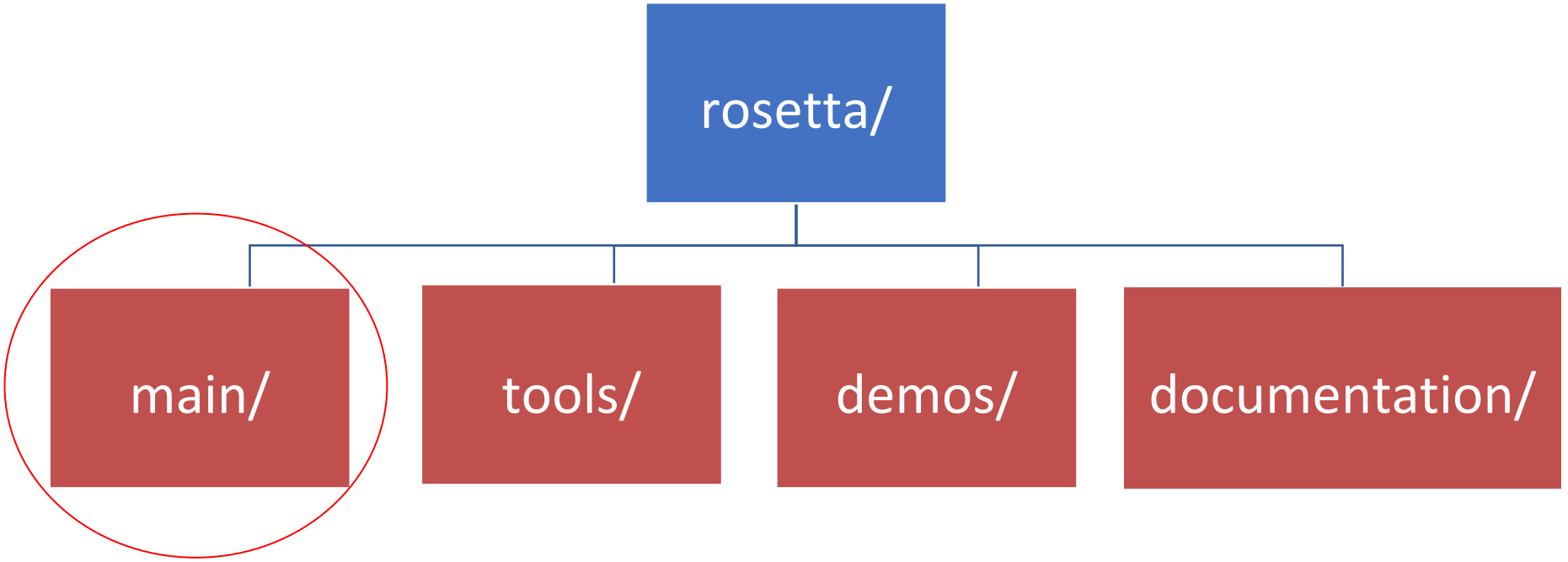
cd ~/rosetta_workshop/rosetta/

Basic Rosetta Structure

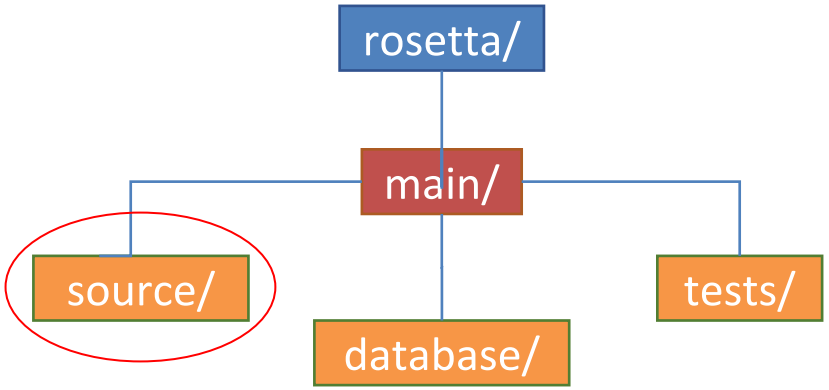


cd ~/rosetta_workshop/rosetta/

Basic Rosetta Structure

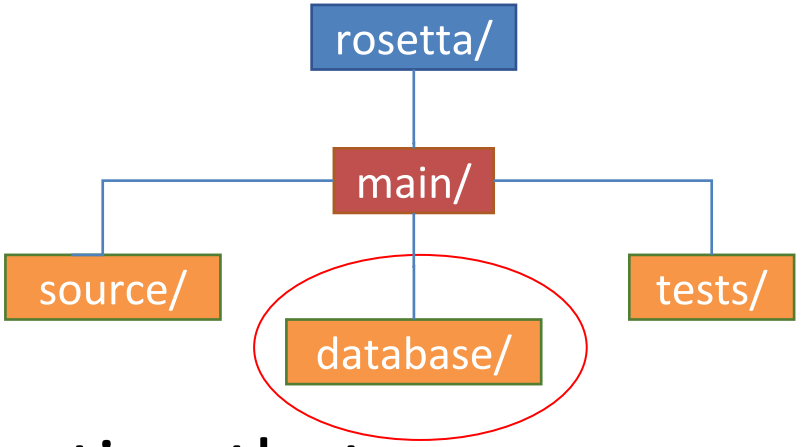


cd ~/rosetta_workshop/rosetta/main/source/



- Rosetta/main/source/bin/
  - Most applications you will run are calling programs within the bin directory
    - rosetta_scripts.default.linuxgccrelease
    - score_jd2.default.linuxgccrelease
    - relax.default.linuxgccrelease
- Rosetta/main/source/scons.py
  - Used for compiling
- Rosetta/main/source/src/
  - This is where all of the code lives

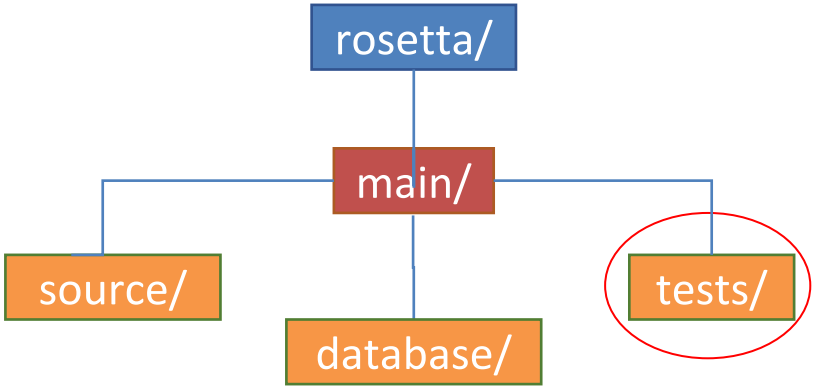
cd ~/rosetta_workshop/rosetta/main/database/
(relative path from source/ == cd ../database/)



- Contains pre-defined information that Rosetta needs for runs

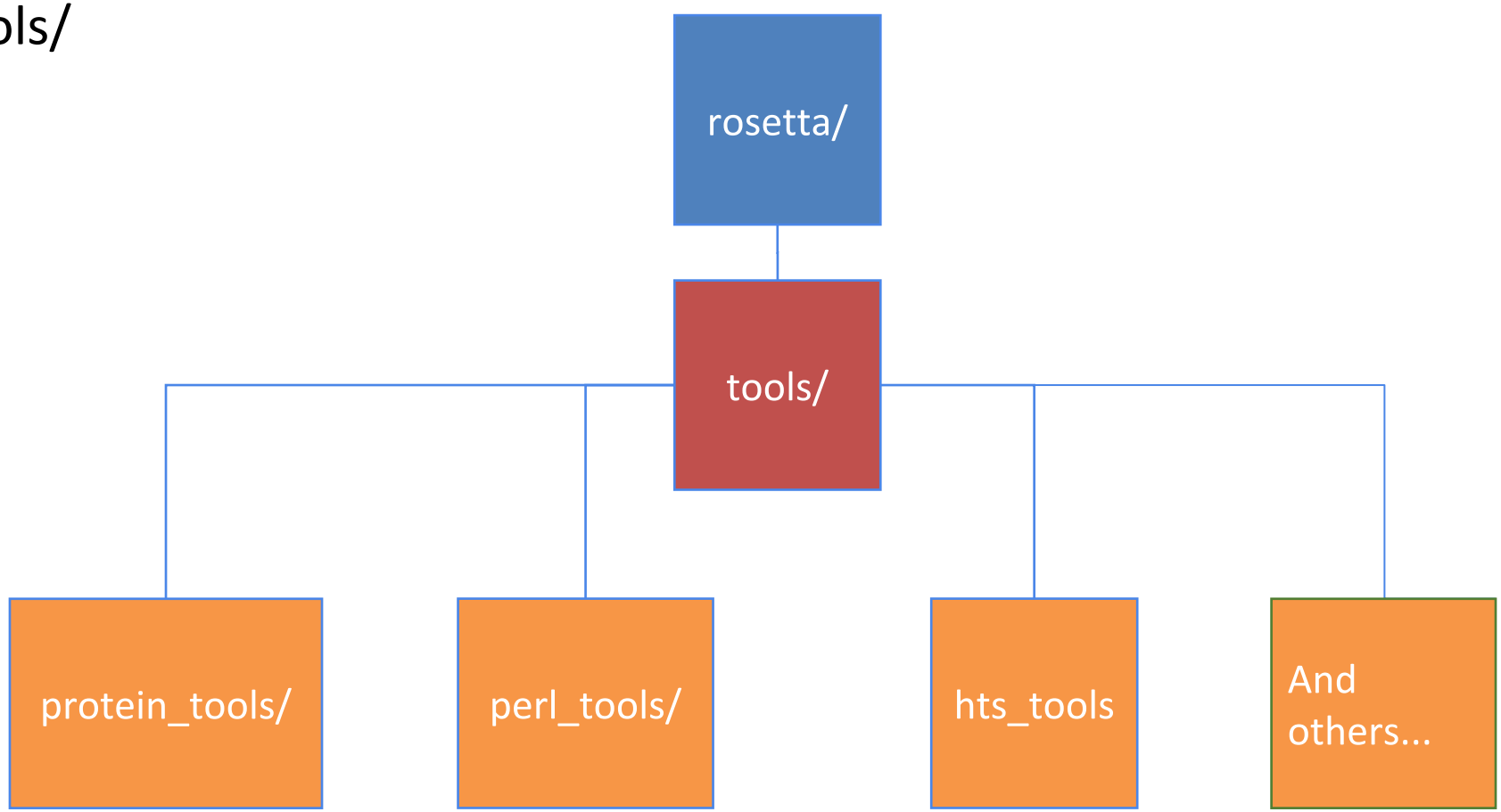
- Rosetta/main/database/chemical/
  - Residue information--params files
  - Atom sets
- Rosetta/main/database/scoring/
  - Default weights files
  - Rotamer libraries

cd ~/rosetta_workshop/rosetta/main/tests/

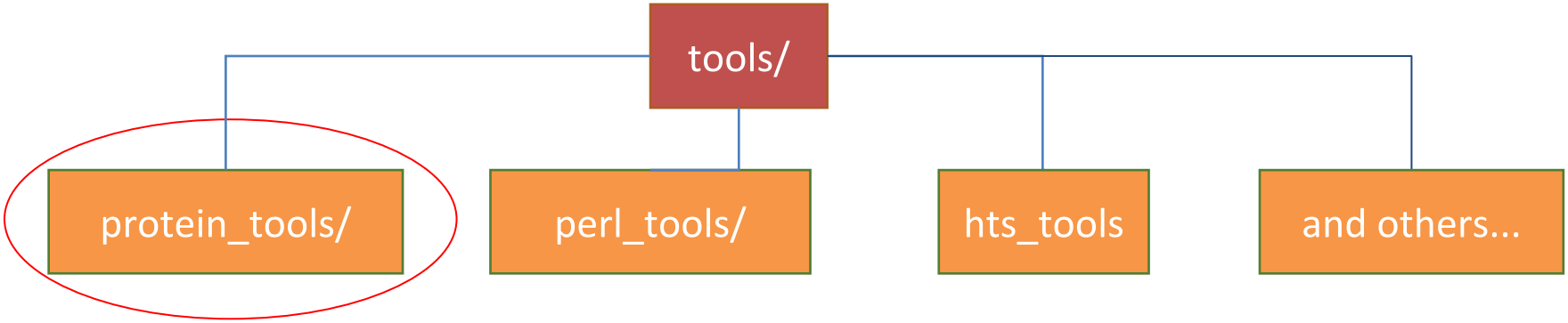


- Rosetta/main/tests/
  - This is for developer use--don’t worry about this (fun to read though!)

cd
~/rosetta_workshop/rosetta
/tools/

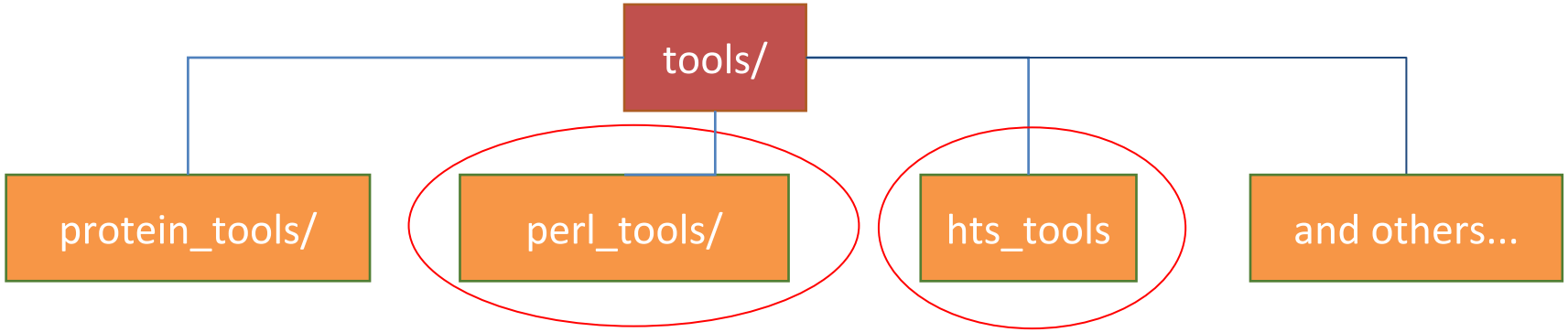


- These scripts are incredibly help for smaller, more basic tasks
- Used mainly to setup or analyze runs



Rosetta/tools/protein_tools/scripts/

- clean_pdb.py
  - Makes a PDB “Rosetta-proof” and used at the beginning of almost any protocol
- pdb_renumber.py
- score_vs_rmsd.py
  - Setup for score vs. RMSD plots
- top_n_percent.py
  - Extracts tags (protein names) for top given percent of models based on score term

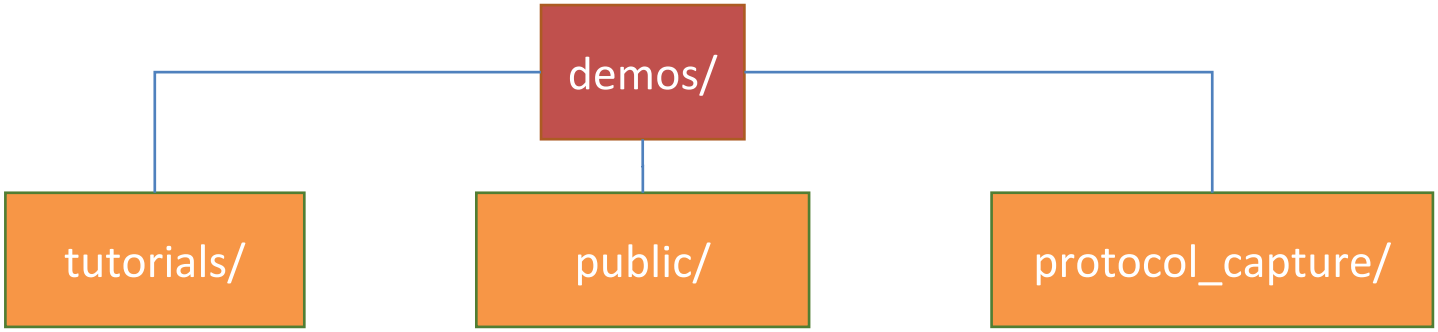


- Rosetta/tools/perl_tools/
  - addchain.pl
  - getFastaFromCoords.pl
  - getCAcoords.pl

- Rosetta/tools/hts_tools/

Scripts used mainly to setup high throughput screen runs (> 1000 compounds)

  - make_params.py
  - get_descriptor_data.py
  - setup_screening_project.py

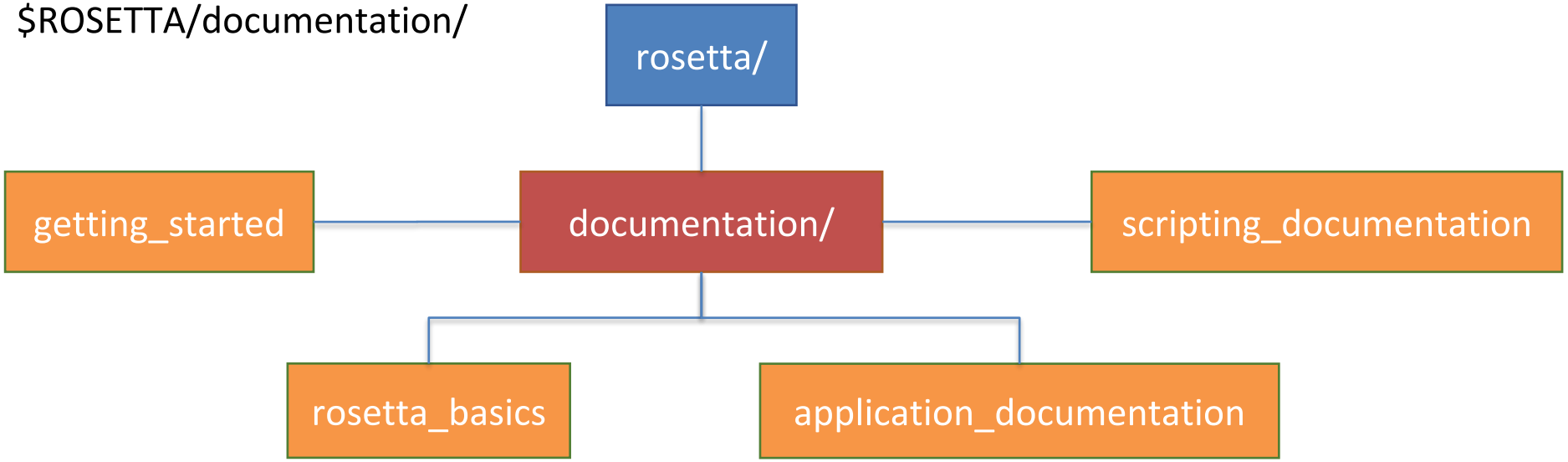


Input files/options are used in protocols--good place to start when you're doing learning a new method

- Rosetta/demos/tutorials/
  - these tutorials
- Rosetta/demos/public/
  - used to go through sample procedures
- Rosetta/demos/protocol_capture/
  - protocols associated with a publication

DISCLAIMER: May be out of date, always check Wiki/RosettaCommons/forum for latest information!

$ROSETTA/documentation/



(Everything here is also on the RosettaCommons site!)

- Understanding general Rosetta concepts
  - Where to find FAQs (How long does this run take?)
  - Options list, file types
  - Protocols you can use
- General structural biology FAQs
  - How do I do X?

Other Rosetta Resources:

- https://www.rosettacommons.org/docs/latest/Home
  - Forum
  - Documentation
  - User guides
  - Everything Rosetta-related you could ever dream of!

GO HAVE FUN!