More about the use of the 'Re_Fit' mode

Tutorial by:
F. Wang,[1] J.-P. Becker, [1] P. Cieplak,[2] and F.-Y. Dupradeau[1]

[1] Université de Picardie - Jules Verne, Amiens
[2] Sanford-Burnham Institute for Medical Research, La Jolla, CA

November 2014

The 'Re_Fit' mode allows fast and efficient update of a force field from a previous PyRED job after modification of different keywords and/or providing new force field parameters.

The quantum mechanics (QM) Gaussian, GAMESS or Firefly program is selected during the input submission procedure using the web interface: one can use the public account or a private account. See http://q4md-forcefieldtools.org/REDServer-Development/faq.php#3

Input molecules of PyRED are provided in the PDB file format:
See http://q4md-forcefieldtools.org/REDServer-Development/Documentation/readme.txt

RUN A FIRST PyRED JOB

1- Define specific keywords in the Project.config and System.config files

See http://q4md-forcefieldtools.org/REDServer-Development/Documentation/Project.config
The user has to define the total charge and spin multiplicity of the molecules involved in force field generation if they are different from zero and one, respectively. She/he can also define bond(s), angle(s) and/or dihedral(s), which can be constrained during geometry optimization. Charge constraints can be requested during the charge fitting step to lead to molecular fragments as well, etc...

See http://q4md-forcefieldtools.org/REDServer-Development/Documentation/System.config
Among others a user has to define the charge model and force field parameter set to be used.

2- Create an archive file to upload the different PyRED inputs

```
tar -zcvf Archive.tgz Mol_red*.pdb Project.config System.config
```

3- Download the first PyRED job

See http://q4md-forcefieldtools.org/REDServer-Development/faq.php#7 to download a job.

PREPARE AND RUN THE SECOND PyRED JOB

After downloading and studying the data generated for the first PyRED job, a user can apply the Re_Fit mode (*i. e.* define `Re_Fit = On` in the *System.config* file), to produce an updated force field for the input molecules based on the QM computations performed in the first PyRED job. To be able to

execute the second PyRED job the data generated in that first PyRED job (*i. e.* the whole `Data-R.E.D.Server` directory) has to be provided in the uploaded archive with the PDB input file(s) and the updated *Project.config*, *System.config* and/or *frcmod.user* files.

1- Modify the keywords in the Project.config file

The user can request the definition of:

- modified rules for chemical equivalencing:

see the `MOLECULE1-CHEMEQSM`, `MOLECULE1-CHEMEQIA`, `MOLECULE1-CHEMEQMM` keywords.

- new/updated charge constraints applied during charge fitting:

see the `MOLECULE1-INTRA-MCC1`, `MOLECULE1-INTRA-MCC2`, `MOLECULE-INTER-MCC1`, `MOLECULE-INTER-MCC2`, `MOLECULE-INTER-MEQA`, `MOLECULE-INTER-MEQB` keywords.

- new names and types for the atoms:

see the `MOLECULE1-ATMNAME` and `MOLECULE1-ATMTYPE` keywords.

- new molecular topology:

see the `MOLECULE1-CALCONECT` keyword with atom connectivities provided in the PDB input file.

- extra-points and/or united carbon atoms:

see http://q4md-forcefieldtools.org/Tutorial/Tutorial-4-demo3.pdf
and http://q4md-forcefieldtools.org/Tutorial/Tutorial-4-demo4.pdf

- etc...

On the contrary the user cannot request the modification of the keywords related to QM calculations (geometry optimization as well as molecular electrostatic or frequency computation), *i. e.* for instance one cannot modify:

- the elements, the total charge and the spin multiplicity of the input molecules,
- the atom order of the input molecules,
- the RBRA versus QMRA procedure (or *vice versa*).

2- Modify the keywords in the System.config file

The user must set the `Re_Fit` variable to `On` in the *System.config* file.

The user can request the definition of:

- a different accuracy for the charge values: see the `COR_CHR` keyword.
- a new format for the force field libraries: see the `OUTPUT_FORMAT` keyword.
- energy decomposition using various scaling factors for 1-4 interactions:

see the `ENERGY_CALC`, `SCALE_FACTOR_EEL`, `SCALE_FACTOR_VDW` keywords.

- rounding off of force field constants: see the `FFPARM_ROUND` keyword, etc...

The user might request the definition of:

- a new charge model if the QM calculations used for the charge models in the first and second jobs are identical. For instance, after a `CHR_TYP = RESP-A1` job a user might perform a `ESP-A1` or a `RESP-O1` job, but cannot run a `RESP-C1` or `RESP-X1` job.
  See http://q4md-forcefieldtools.org/REDServer-Development/popup/popkeyword.php

- a new force field set if the QM calculations used for the force field generated in the first and second jobs are identical. For instance, the `AMBERFF03` and `GLYCAM04` force fields request specific QM calculations, while equivalent QM theory levels are used for the `AMBERFF94`, `AMBERFF96`, `AMBERFF98`, `AMBERFF99`, `AMBERFF99SB`, `AMBERFF99SBBSC` and `AMBERFF10` force fields. See the `FFPARM` keyword and references at http://q4md-forcefieldtools.org/REDServer-Development/popup/popforcefield.php.

On the contrary the user cannot request the modification of the theory levels used in QM calculations or perform frequency computation (see the `METHOD_OPTCALC`, `BASSET_OPTCALC`, `METHOD_MEPCALC`, `BASSET_MEPCALC` and `SURFMK_MEPCALC` keywords related to the Gaussian program).

3- Define specific force field parameters in the frcmod.user file

A user can construct a *frcmod.user* file to provide missing and/or mandatory force field parameters in agreement with the force field generated in the first job, and/or in agreement with the modifications requested in the second job.

4- Create an archive file to upload the different PyRED inputs

```
tar -zcvf Archive.tgz Mol_red*.pdb *.config frcmod.user Data-R.E.D.Server/*
```

Using the bzip2 compressor leads to a smaller archive file to be uploaded:

```
tar -jcvf Archive.tar.bz2 Mol_red*.pdb *.config frcmod.user Data-R.E.D.Server/*
```

5- Download and study the second PyRED job

After download of the second PyRED job, the data of the second PyRED job are available in the `Data-R.E.D.Server` directory, while the data of the first PyRED job have been moved in the `Data-R.E.D.Server1` directory.

6- Using the force field generated by PyRED within the AmberTools

The force field generated by PyRED can be found in the following sub-directories:

- The *force field libraries* are available in the `Mol_m1` or in the `Mol_MM` sub-directory (the mol3 file format is the default; see http://q4md-forcefieldtools.org/Tutorial/leap-mol3.php) if the number of PDB input file(s) equals one, or is strictly greater than one, respectively.

- The *force field parameter files* are available in the *Data-Default-Proj* sub-directory.
  Three files are generated:
  - the *frcmod.known* file lists the known force field parameters determined by PyRED based on the atom types assigned to the input molecules,
  - the *frcmod.corespondence* file lists the force field parameters, which were determined by analogy to the existing ones,

- the *frcmod.unknown* file lists the force field parameters, which remains unknown.
  The 'Re_Fit' mode can be used to provide a *frcmod.user* input file in agreement with the atom types assigned by PyRED, or after update of these atom types (see the `MOLECULE1-ATMTYPE` keyword):
  See http://q4md-forcefieldtools.org/Tutorial/Tutorial-4-demo1.pdf

– The *leaprc.q4mdfft* script available in the *Data-Default-Proj* sub-directory allows the direct use of the force field generated by PyRED within the LEaP program (see the Amber tools at http://ambermd.org/#AmberTools):

```
# Extract the PXXXX.tar.bz2 R.E.D. Server Dev./PyRED job after download
tar -jxvf PXXXX.tar.bz2
cd PXXXX/Data-R.E.D.Server/Data-Default-Proj
# Edit to update the LEaP script and extend its use
gedit leaprc.q4mdfft &
# Run the LEaP program:
# see http://q4md-forcefieldtools.org/REDServer-Development/images/PyRED.gif
xleap -f leaprc.q4mdfft
```

Remarks

– This strategy of force field update can be repeated as many times as ones wishes by using the Re_Fit mode.
– A Re_Fit job is generally fast (takes a few minutes) compared to a job involving geometry optimization and molecular electrostatic potential computation.