SupeRNAlign Tutorial

Installing SupeRNAlign

Download .zip file with the latest version of SupeRNAlign from http://genesilico.pl. Follow the installation instructions in `README.md`.

Install the superposition tool of your choice; currently R3D Align, SARA, ARTS and LaJolla are supported.

> **Note:** When installing R3D Align, please use the older version from our website. The latest version is incompatible with our program.

Set the correct paths in `config_file.cfg`.

> **Note:** SupeRNAlign comes bundled with two auxiliary programs: Clarnet and Smatch. They improve the quality of alignments produced by SupeRNAlign, but - in rare cases - you may want to turn them off. Both programs may be disabled in the config file.

Obtaining input data

Download two tRNA structures: 3KFU and 4WJ4 from RCSB Protein Data Bank. (Click *Download Files / PDB Format*.)

3KFU (chain K) will serve as the reference structure, and 4WJ4 (chain B) will be superimposed on it.

Superimposing input structures

Enter the directory where the input .pdb files are located. Run SupeRNAlign:

```
supernalign.py -c 3kfu.pdb 0 K 4wj4.pdb 0 B -o tRNA_superposition
```

(Assuming SupeRNAlign directory is in your system path; if not, you should precede `supernalign.py` with the path, e.g. `~/supernalign/supernalign.py` - if SupeRNAlign is located in the `supernalign` subdirectory of your home directory.)

- `-c` tells SupeRNAlign to clean and renumber the input structures, and remove all modifications. This is to make sure that SupeRNAlign won't encounter any 'illegal' residues; if you're sure that your structures don't need cleaning, you may omit this option.
- `3kfu.pdb` is the name of the first input file.
- `0` is the model number of the first structure; usually 0, unless the model is explicitly set in the input file.

- `K` is the chain ID of the first structure (see the summary page in PDB).
- `4kj4.pdb` is the second input file.
- `0` is the model number of the second structure (see above).
- `B` is the chain ID of the second structure.
- `-o tRNA_superposition` sets the names of the output files to `tRNA_superposition.pdb` (superimposed structures) and `tRNA_superposition.aln` (sequence alignment). If you omit this option, SupeRNAlign will save output to `supernalign_output`.

SupeRNAlign will inform you on the progress:

```
2D structure processing...
Preliminary superposition...
Processing...
```

Retrieving output data

After a while (usually a few minutes), SupeRNAlign will save output data and print the sequence alignment on the screen.

`tRNA_superposition.pdb` will hold the superimposed structures. The first (reference) structure - 3KFU - will be stored as chain A, and the second structure - 4WJ4 - as chain B (see Figure 1). You can view the structures using molecular visualization software, e.g. PyMol.

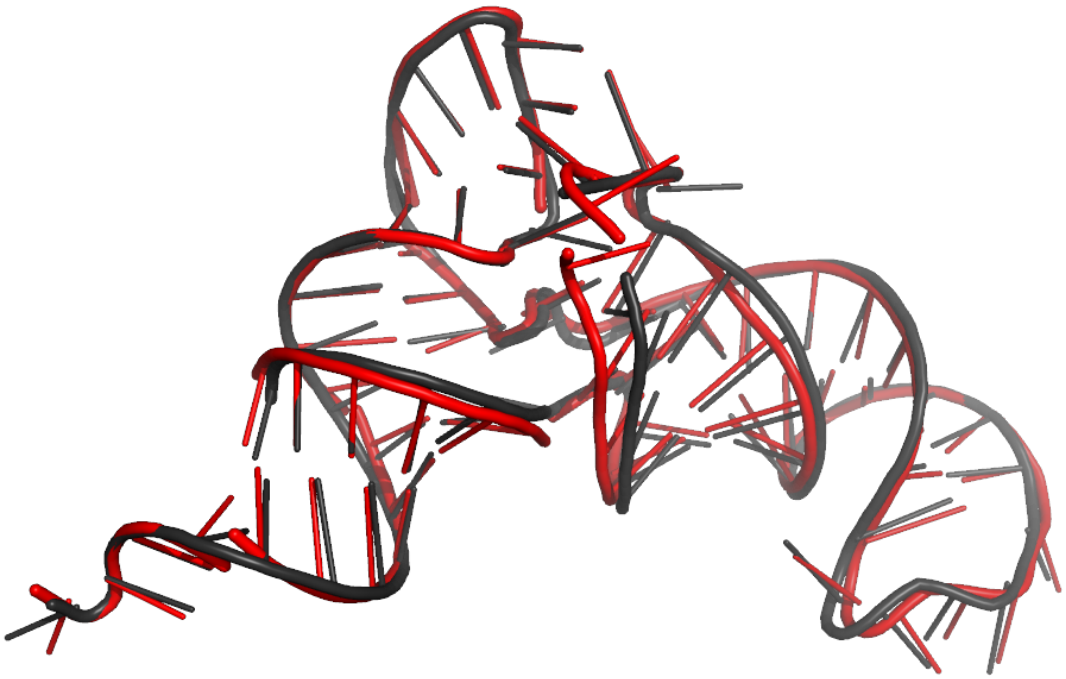

Figure 1. Structures after superposition. Dark grey - reference structure (3KFU); red - aligned structure (4WJ4).

Sequence alignment will be saved in `tRNA_superposition.aln`, in FASTA format. `Chain A` denotes

the reference structure (3KFU), and `Chain B` is the aligned structure - 4WJ4 (Figure 2).

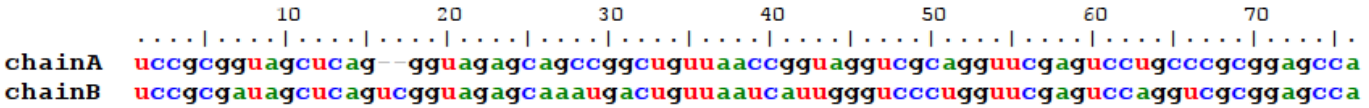

Figure 2. Sequence alignment returned by SupeRNAlign.

Pre-aligned structures

If your input structures are already superimposed, use the `-p` option and SupeRNAlign will skip the preliminary superposition step.

Enter ModeRNA server's website: http://iimcb.genesilico.pl/modernaserver/.

- Click *Submit / Analyse structure*.
- Select `3kfu.pdb` - the first downloaded input file.
- Enter `K` as chain ID.
- Select *clean structure*, *remove modifications*, *renumber chain*.
- Click *Analyse*.
- Wait for the results, download the file and rename it to `3kfu_cleaned.pdb`.

Repeat the above steps for `4wj4.pdb`.

Once you have your input files prepared, enter Rclick website: http://mspc.bii.a-star.edu.sg/minhn/rclick.html.

Upload `3kfu_cleaned.pdb` and `4wj4_cleaned.pdb`, and click *Run Rclick*. When the results are ready, click *Superimposed structures*; unpack the downloaded file:

```
tar xvf 3kfu_cleaned-4wj4_cleaned.2.tar.gz
```

From the working directory type:

```
supernalign.py -p 3kfu_cleaned-4wj4_cleaned.2.pdb 0 K
4wj4_cleaned-3kfu_cleaned.2.pdb 0 B -o tRNA_Rclick
```

Results will be saved to `tRNA_Rclick.pdb` and `tRNA_Rclick.aln`.

SupeRNAlign-Coffee

To further improve the alignments, SupeRNAlign has an interface to T-Coffee. You need to install T-Coffee as described in the documentation.

From your working directory type:

```
make_coffee.py tRNA_superposition.pdb 0 A tRNA_superposition.pdb 0 B > library.lib
```

T-Coffee library file `library.lib` will be created. Type:

```
t_coffee -lib library.lib -mode rcoffee -output fasta -outfile sa-coffee.fasta
```

Sequence alignment will be written to `sa-coffee.fasta`.

Paweł Piątkowski, 21.07.2016