README document for the project files provided as supporting information to the manuscript "Ligand-protein interactions in lysozyme investigated through a dual-resolution model"

February 2020

Overview

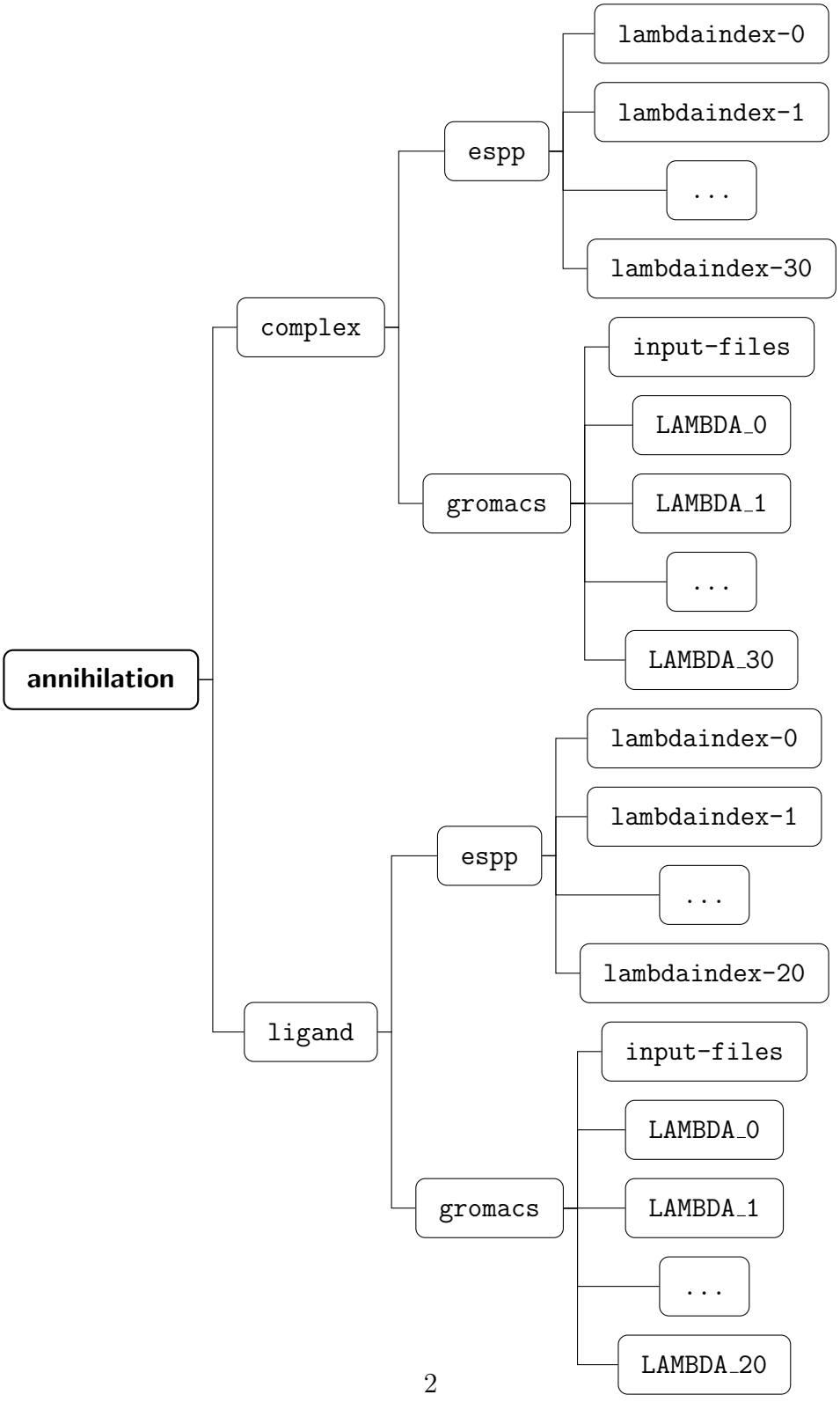
The main folder is organised in three subfolders (see the tree diagrams in each section):

1. `annihilation/`
2. `decoupling/`
3. `density/`

The figure deltaG_binding_ann_dec_comparison.png shows the results of binding free energy calculations comparing the values obtained both for annihilation and decoupling.

In each section of this document, it is possible to notice a tree diagram representing the sequence of directories present for easier global reading.

1 Annihilation



The `annihilation/` folder contains all results concerning the calculation of binding free energy in case of annihilation and it is divided in two parts:

- `complex/`
- `ligand/`

The figure deltaG_binding_annih_gromacs_espp.png displays the results for Binding FE, comparing the values obtained in GROMACS and ESPResSo++.

1.1 complex

In `complex/` are reported the results of Ligand-Protein FE both in ESPResSo++ and GROMACS. All simulations are fully-atomistic.

The protein-ligand complex free energy (ΔG_{compl}) calculation uses 11 λ values per ΔG_{restr_on} 5 evenly spaced λ values per ΔG_{LJ} (with separation 0.20) and 15 λ values per $\Delta G_{coul,c}$, with 1 ns of simulation per λ.

It is possible to find the trajectory files in the subdirectories `lambdaindex-0/` and `lambdaindex-30/` in case of ESPResSo++, or `LAMBDA_0/` and `LAMBDA_30/` in case of GROMACS simulation.

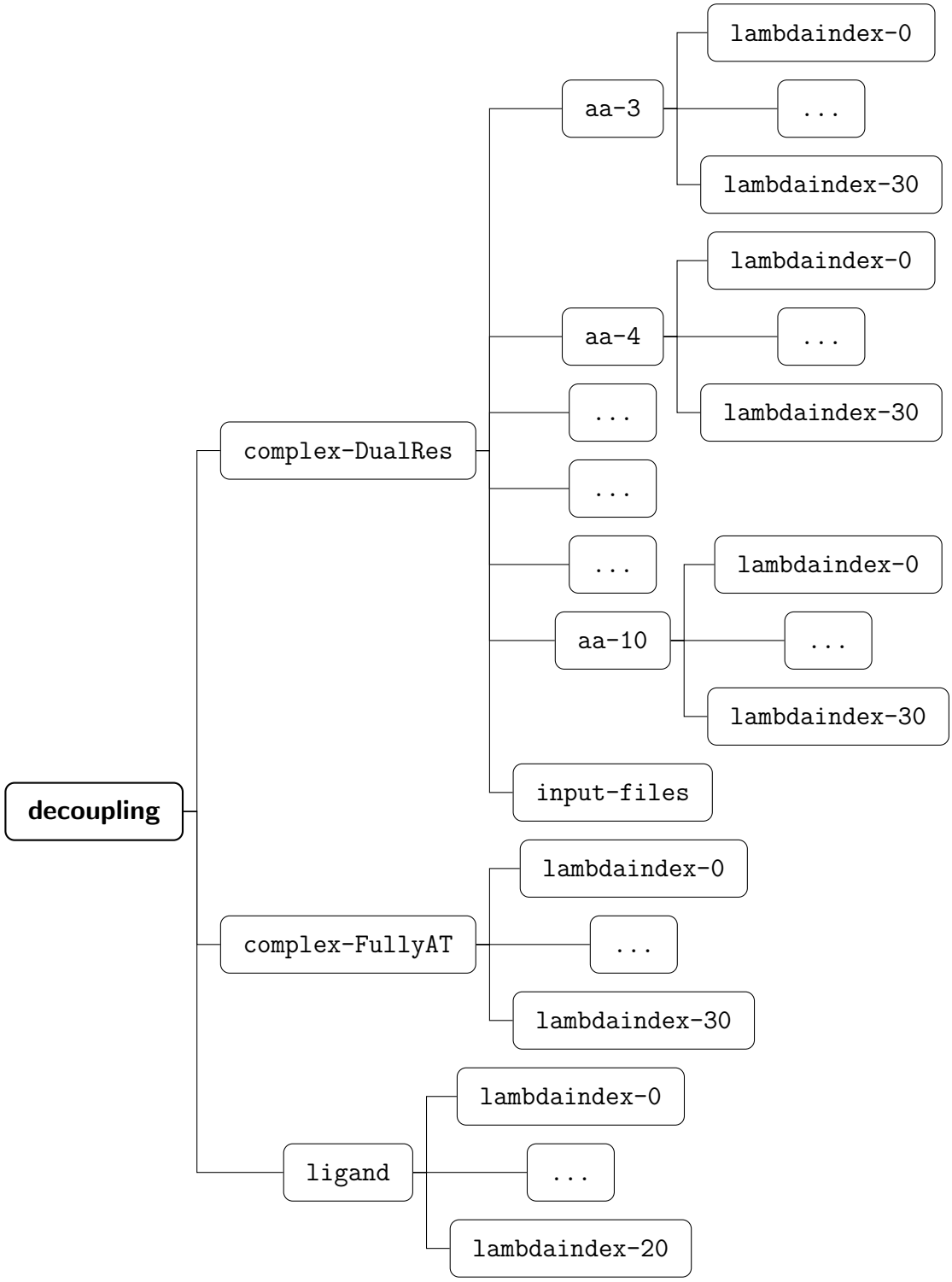
1.2 Ligand

In `ligand/` are reported the results of ligand solvation free energy both in ESPResSo++ and GROMACS. All simulations are fully-atomistic.

The ligand solvation free energy (ΔG_{lig}) calculation uses 5 evenly spaced λ-values per $\Delta G_{coul,\ell}$ (with separation 0.20) and 16 λ-values per $\Delta G_{LJ,\ell}$ with 1 ns of simulation of each λ-value.

It is possible to find the trajectory file in the subdirectories `lambdaindex-0/` and `lambdaindex-20/` in case of ESPResSo++, or `LAMBDA_0/` and `LAMBDA_20/` in case of GROMACS simulation.

2 Decoupling



This folder contains all results concerning the calculation of binding free energy in case of decoupling and it is divided in three parts:

- `complex-DualRes/`
- `complex-FullyAT/`
- `ligand/`

2.1 complex-DualRes

In `complex-DualRes/` are reported the results of Ligand-Protein FE only in ESPResSo++ (GROMACS cannot do decoupling). The system is simulated in Dual-Resolution. In particular, according with the number of the protein active site residues modelled atomistically (from 3 to 10), the ligand-protein FE (ΔG_{compl}) has been computed (`aa-3/`, `aa-4/`, ... , `aa-10/`).

For each case, the protein-ligand complex free energy calculation uses 11 λ values per ΔG_{restr_on} 5 evenly spaced λ values per ΔG_{LJ} (with separation 0.20) and 15 λ values per $\Delta G_{coul,c}$, with 4 ns of simulation.

It is possible to find the trajectory files in the sub-directories `lambdaindex-0/` and `lambdaindex-30/`.

2.2 complex-fullyAT

In `complex-fullyAT/` are reported the results of Ligand-Protein FE only in ESPResSo++. The system simulated is fully-atomistic. The protein-ligand complex free energy (ΔG_{compl}) calculation uses 11 λ values per ΔG_{restr_on} 5 evenly spaced λ values per ΔG_{LJ} (with separation 0.20) and 15 λ values per $\Delta G_{coul,c}$, with 600 ps of simulation per λ.

It is possible to find the trajectory file in the sub-directories `lambdaindex-0/` and `lambdaindex-30/`.

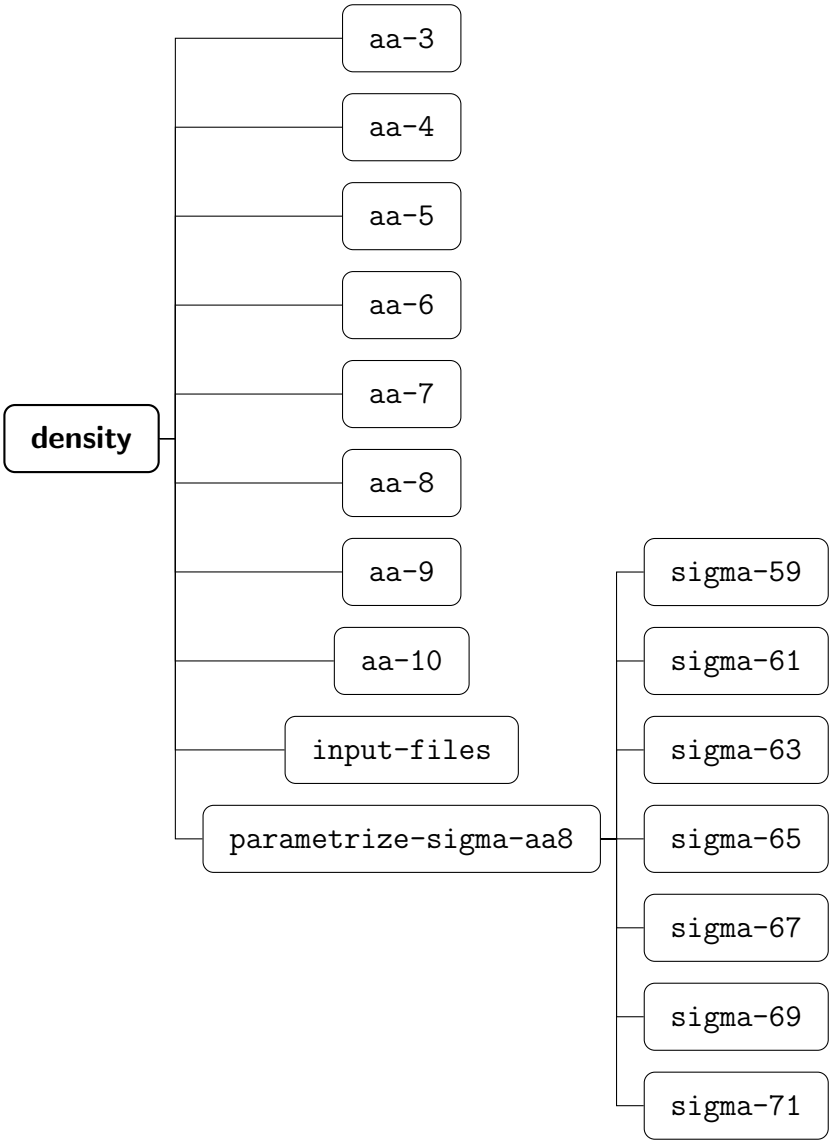
2.3 Ligand

In `ligand/` are reported the results of ligand solvation free energy only in ESPResSo++. All simulations are fully-atomistic.

The ligand solvation free energy (ΔG_{lig}) calculation uses 5 evenly spaced λ-values per $\Delta G_{coul,\ell}$ (with separation 0.20) and 16 λ-values per $\Delta G_{LJ,\ell}$ with 600 ps of simulation of each λ-value.

It is possible to find the trajectory file in the sub-directories `lambdaindex-0/` and `lambdaindex-20/`

3 Density



When the Elastic Network Model (ENM) is employed in multi-resolution simulations, an excluded volume interaction between ENM nodes and solvent molecules is required, in order to prevent from penetrating the protein and solvating the atomistic binding site from the interior. Thus, a WCA interaction is applied between C_α nodes and all the solvent molecules.

In its formulation, WCA needs two parameters: ϵ and σ. The former has

a value of 0.34 kJ $\cdot$ mol^{-1}, arbitrarily chosen as the value for carbon in the atomistic forcefield, whilst $\sigma_i = R_{g,i} \cdot c$, where $R_{g,i}$ is the radius of gyration of a given residue i out of the twenty possible amino acids and c is the same for all amino acids. The latter is not known a priori, because its value has to be tuned to give the correct bulk water density for a protein-water system (i.e. the water density far from the protein) from fully atomistic simulation: the results of these calculations are reported in `density/` folder.

First, in order to find the proper value of c we started with the 8 atomistic residues protein model (the reference folder is `parametrize-sigma-aa8/`) launching different dual resolution simulations of 1 ns, varying its value (`sigma-59/`, `sigma-61/`, ... , `sigma-71/`).

After finding the correct c such that the density of water in the atomistic and dual-res system are comparable (that is $c = 0.658$), we checked that such value is correct running, this time, 1 ns simulations with different numbers of atomistic residues keeping c fixed (`aa-3/`, `aa-4/`, ... , `aa-10/`).