Modifying Lennard-Jones Parameters in the Amberff14SB force field

In this tutorial you will learn to:

- Modify the self-interaction Lennard-Jones parameters of oxygens of carboxylate groups in proteins.
- Modify the Lennard-Jones parameters defining the interaction between sodium ions and the oxygens of carboxylate groups in proteins.

Hosein Geraili Daronkola
Ana Vila Verde

Max Planck Institute of Colloids and Interfaces
Theory & Bio-systems department
March 2019

Introduction

- The Amber formulation of the 6-12 Lennard-Jones (LJ) potential, V_{ij}, between 2 atoms *i* and *j* is:

$$V_{i,j} = \varepsilon_{i,j}\left(\left(\frac{R_{\min,i,j}}{r_{i,j}}\right)^{12} - 2\left(\frac{R_{\min,i,j}}{r_{i,j}}\right)^{6}\right) \qquad \text{(eq. 1)}$$

- Here $R_{min,i,j}$ is the center-to-center distance between *i* and *j* at which the potential is at the minimum $\varepsilon_{i,j}$.
- Van der Waals data in Amber force field files are given for each atom *i* as a single data pair: a radius $R_{min,i,i}/2$ ('van der Waals' radius of atom *i*, in Å) and the energy $\varepsilon_{i,i}$ (the minimum interaction energy between 2 atoms *i*, in kcal/mol) . These parameters are also called the self-interaction parameters.
- For Amber force fields, cross terms involving different atom types *i* and *j* are typically evaluated according to the Lorentz/Berthelot mixing rules:

$$R_{\min,i,j} = 0.5(R_{\min,i,i} + R_{\min,j,j}) \quad \text{(eq. 2)} \qquad \varepsilon_{i,j} = \sqrt{\varepsilon_{i,i} \cdot \varepsilon_{j,j}} \qquad \text{(eq. 3)}$$

- In the first part of this tutorial you will learn how to modify self-interaction LJ parameters, using as example the parameters for the oxygens of carboxylate groups in proteins. Notice that, by modifying the self-interaction parameters of these oxygens, you are in fact modifying the LJ interactions of *every atom type with these oxygens*, via the mixing rules given by eqs. 2,3 .
- The mixing rules have been shown to poorly represent the van der Waals interactions in certain cases. In the second part of this tutorial you will learn how to override the mixing rules for specific pairs of atoms *i* and *j,* and to use instead values of $R_{min,i,j}$ and/or ε_{ij} optimized for that interaction. Specifically, you will modify the LJ parameters for the interaction between Na^+ and the oxygens in the carboxylate groups of proteins.
- The new parameters are from Kashefolgheta, S. & Vila Verde, A. *PCCP,* **2017***, 19*, 20593-20607 , *doi: 10.1039/C7CP02557B* . They yield better agreement with experiment for the hydration free energy of acetate and the solution activity derivative of 0.5 m sodium acetate in TIP3P water. The same paper also reports optimized parameters for the NH_3^+ group of lysine, which we recommend using to obtain a better description of salt bridges in proteins.

- The original parameter files for the AMBER force field are in a directory which in our system can be found via environmental variable $AMBERHOME. The path to this directory is specific to each installation. If $AMBERHOME is not defined in your system and you don't know the path, you will need to ask your local IT support for help. You will leave the original files unchanged, and you will do modifications on local copies.

- Create a directory Tutorial/ at a location of your choice. Copy the necessary files from $AMBERHOME to Tutorial/ using the commands shown in fig. A.

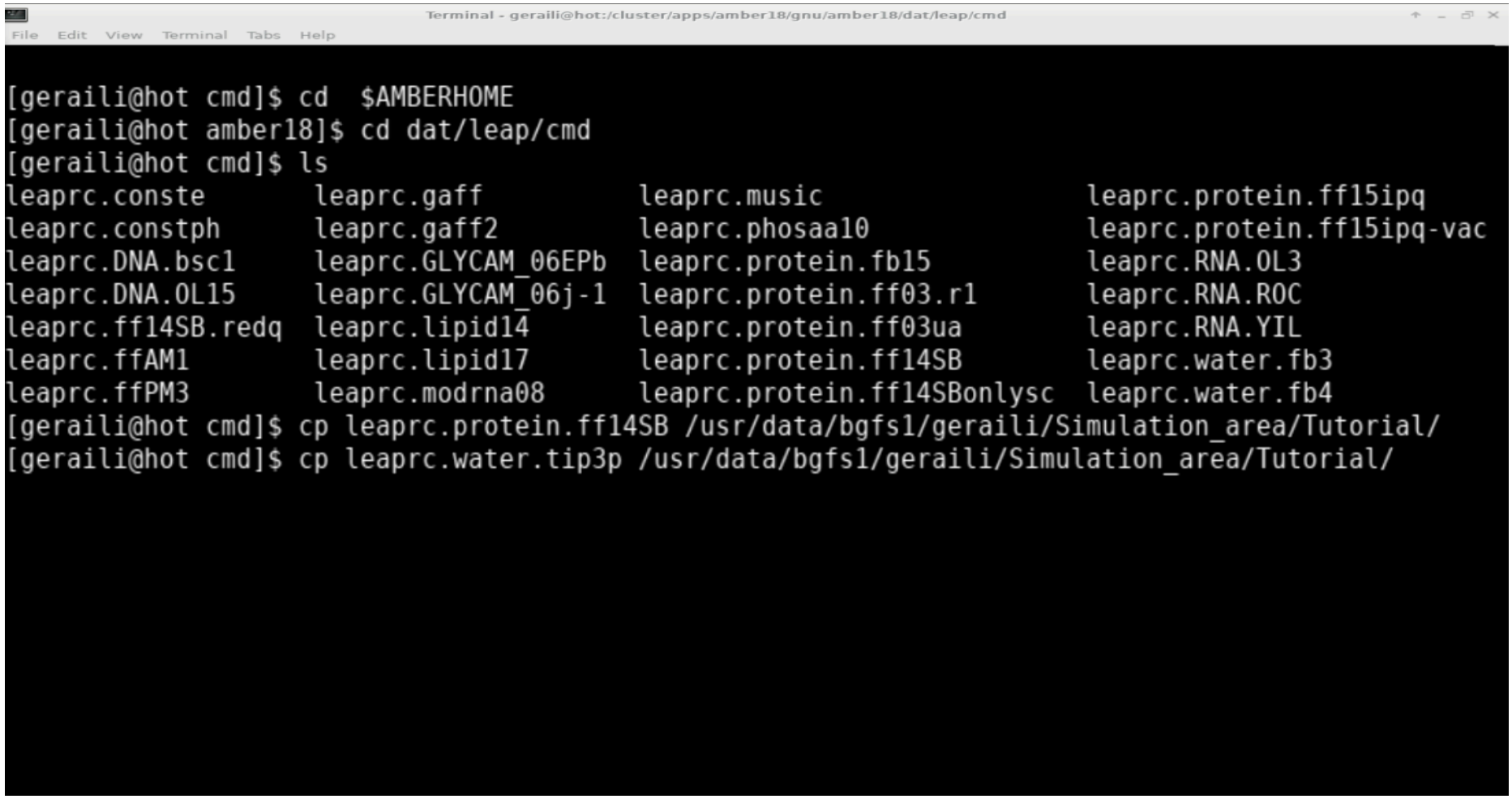

Fig. A

- Once this step is complete, your Tutorial/ directory should have the files shown in Fig. B.

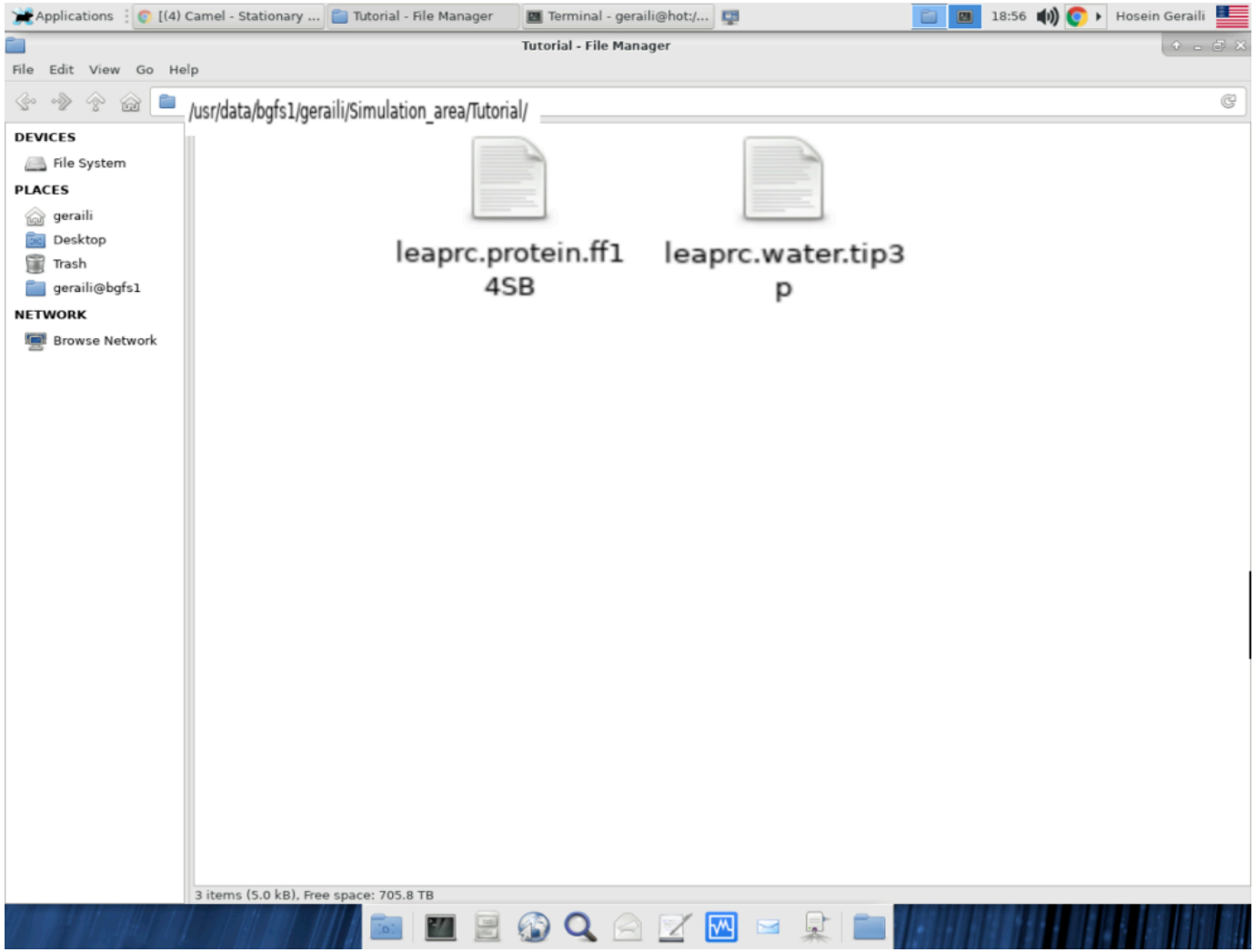

Fig. B: content of Tutorial/

- Copy the remaining necessary files following the commands in Fig. C.

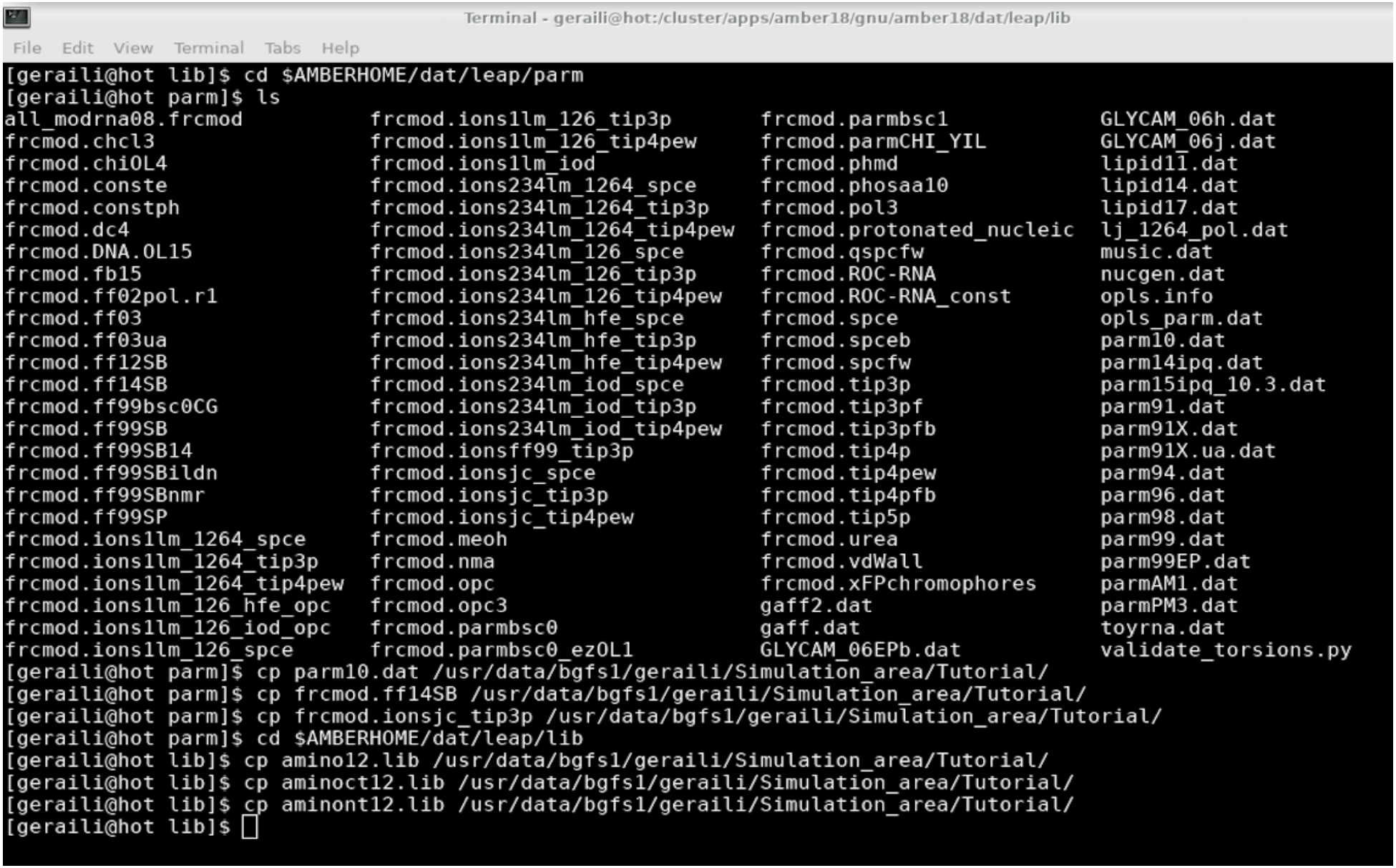

Fig. C

- Once you've copied all the files, Tutorial/ should have the content shown in Fig. D.
- The files *leaprc.water.tip3p* and *leaprc.protein.ff14SB* are loaded into "tleap" to build topology and coordination files for simulations of proteins in water. These files contain, among other things:
  - a list of atom types;
  - the path to the parameter files that will be loaded by "tleap":
    - *leaprc.protein.ff14SB* calls parameter files *parm10.dat* and *frcmod.ff14SB* and topology files *amino12.lib, aminoct12.lib, aminont12.lib;*
    - *leaprc.water.tip3p* calls several parameter and topology files for TIP3P water and TIP3P-compatible ions; one of the files called is *frcmod.ionsjc_tip3p*, which we have copied. We will not change anything in *leaprc.water.tip3p* or in *frcmod.ionsjc_tip3*. We copied these files because 1) we want to view the original Na+ parameters for the second part of the tutorial, and 2) it is convenient to have in a single directory all the files necessary to run a simulation in the AMBER MD software.
- Now that we have all the files necessary, we need to modify them one by one. **You will only modify files inside Tutorial/**.

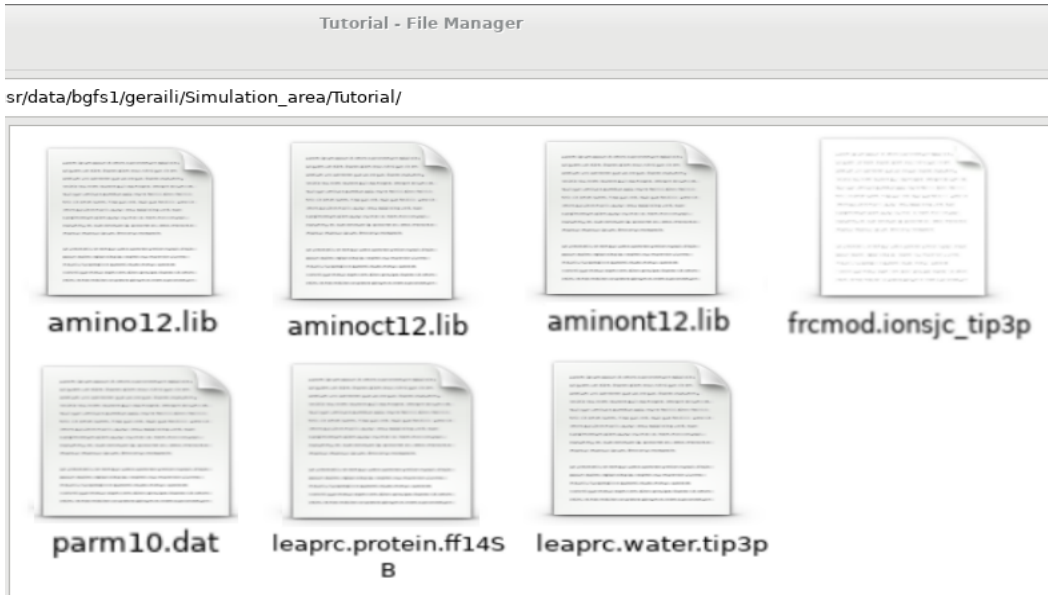

Fig. D: content of Tutorial/

- Our first, and most complex, task is to modify the self-interaction parameters of the oxygens in the carboxylate groups of proteins. Carboxylates exist in side chain of Asp (Asparate) and Glu (Glutamate) residues, and in any uncapped amino acid forming the C-terminus of the protein.

- In general, the difficulty in creating new atom types for a specific functional group is to make sure that we change the parameters of only the atom type on those functional groups, while leaving oxygens with the same original atom type but not belonging to the same functional group with the original parameters. In our specific example this problem does not arise because we only have these type of oxygens on the carboxylates in the side chain of Asp (aspartate) and Glu (glutamate) residues, and in any uncapped amino acid forming the C-terminus of the protein. We nevertheless follow a standard procedure to create a general tutorial for any other kind of optimization that might need to consider this important point.

- Open file *amino12.lib* with your favorite text editor and go to the entry for aspartate (“ASP”). The carboxylate oxygen has type “O2”. An example section of the unmodified version of this file is in Fig. E.; the red boxes show you the lines containing “O2” for ASP.

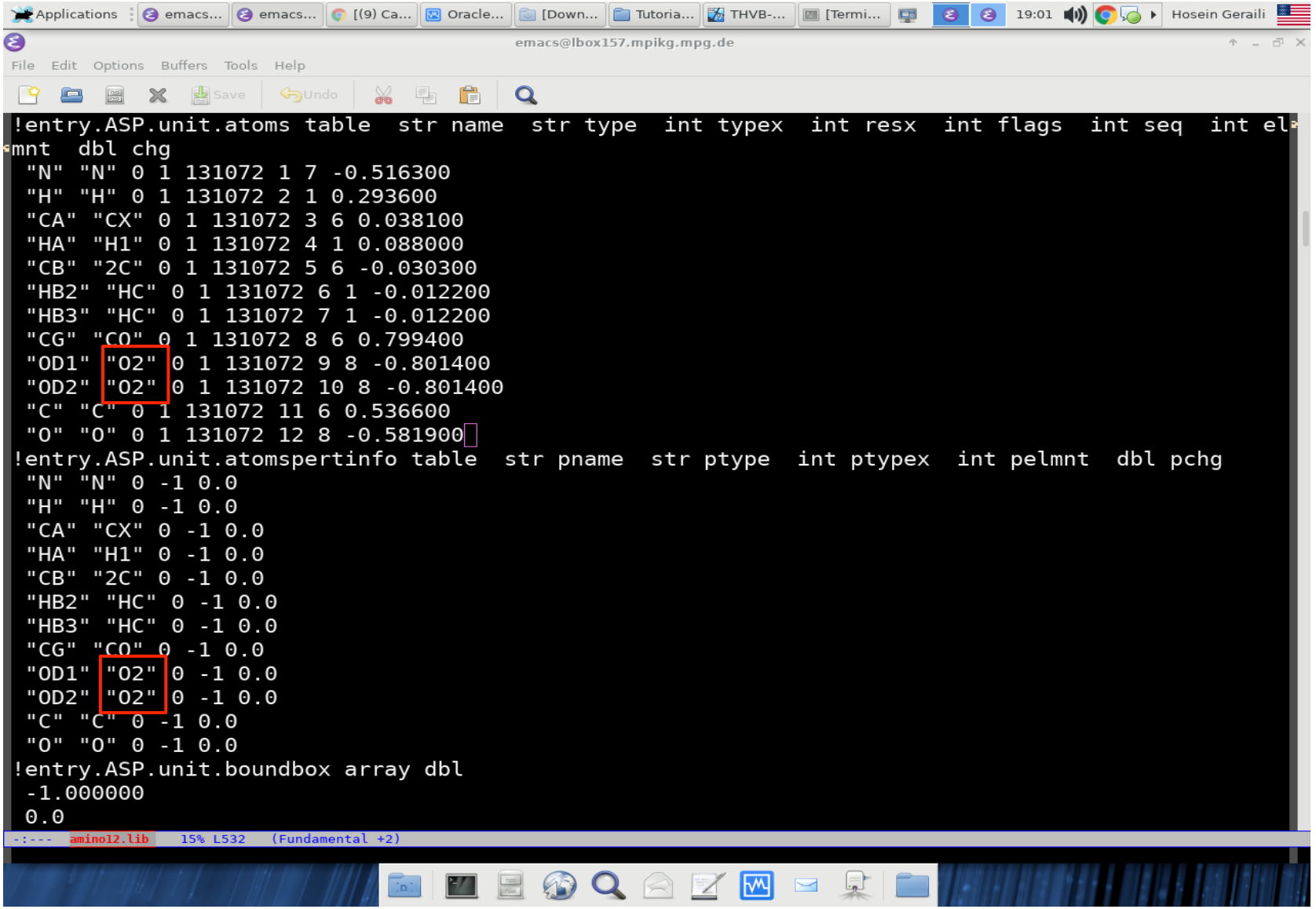

Fig. E: original *amino12.lib*

- In *amino12.lib*, change the atom type of the oxygens in the side chain of ASP from "O2" to a new type; we chose "9O". See Fig. F for the modified entries for ASP.
  - It is indispensable that the new atom type: 1) has only 2 letters; 2) is not already used for other atoms. You can check whether your new atom type is not being used by searching through this file.
- Search through this file and double check that atom type "O2" is changed to "9O" for every entry for ASP. Repeat the procedure for glutamate ("GLU"). Save and close the file when you are done.

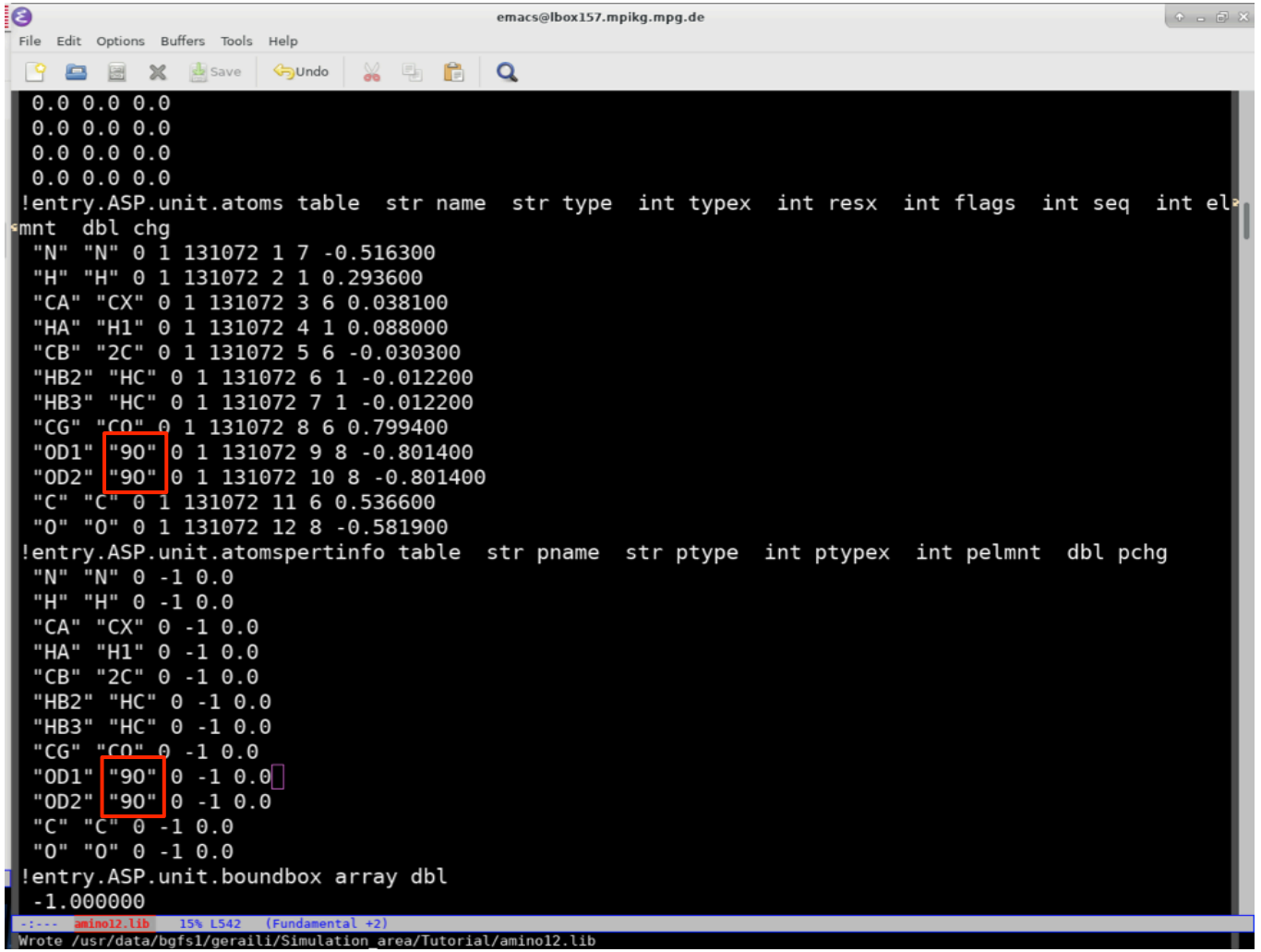

Fig. F: modified *amino12.lib*

Attention!
Always make sure you modify only the entries you want.

- Amino acids forming the C- or N-terminus of a protein have separate entries in the AMBER force field. The C-terminus entries are in file *aminoct12.lib*; the N-terminus entries are in file *aminont12.lib*. You will need to modify these files as well.
- An example section of the unmodified version of *aminoct12.lib* is shown in Fig. G. Notice that you now have 2 carboxylates for CGLU (red boxes): one in the side-chain and one forming the C-terminus.

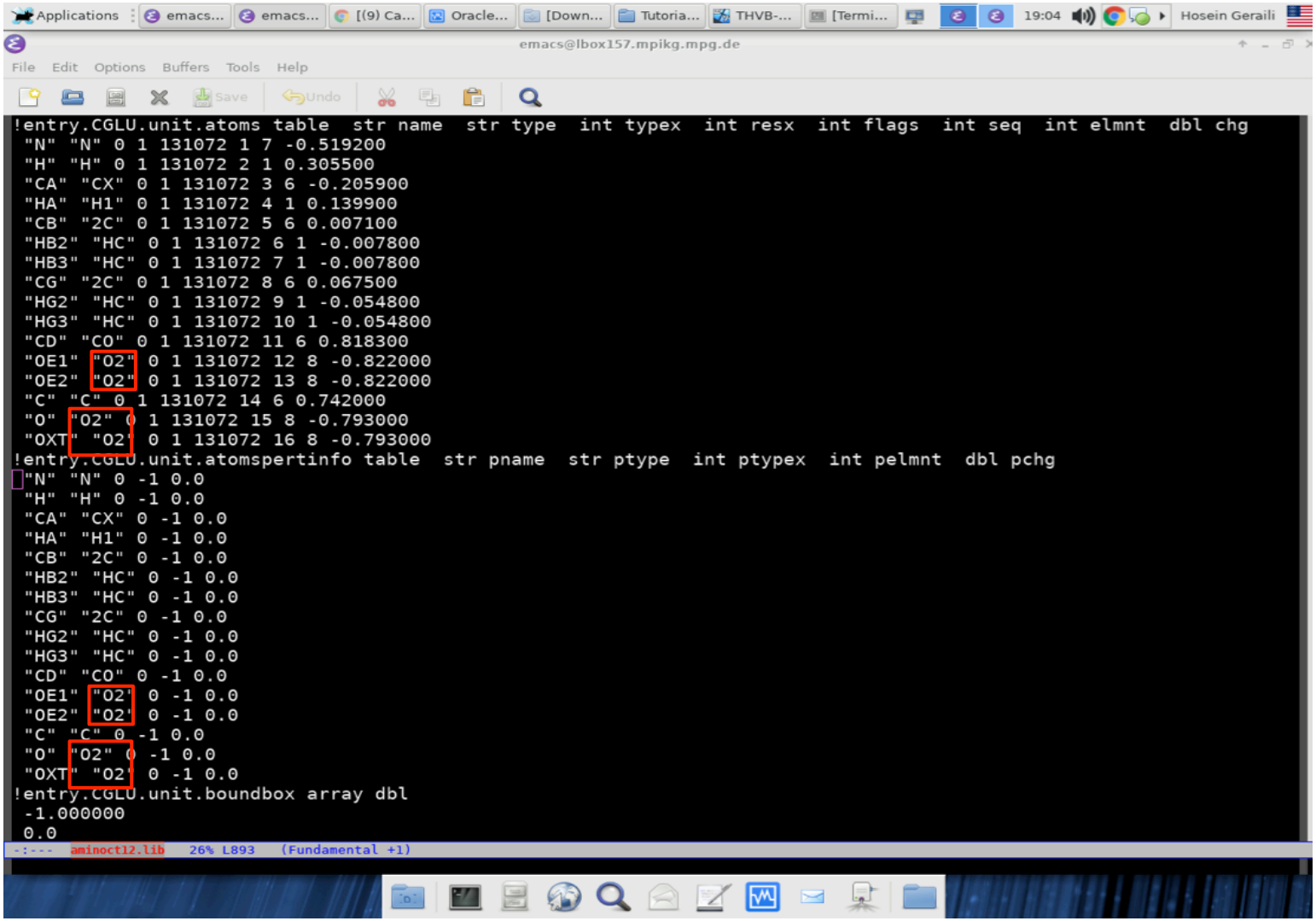

Fig. G: original *aminoct12.lib*

- For the C-terminus (*aminoct12.lib*):

  - Search for CGLU and change atom type “O2” to “9O”, similarly to what you did before. Fig H shows an example section of the modified *aminoct12.lib*.

  - Repeat for CASP.

  - Every other terminal amino acid has one carboxylate group; you will need to change “O2” to “9O” for all those carboxylates too (examples not shown).

  - Save and close the file when you are done.

- For the N-terminus (*aminont12.lib*):

  - Repeat the procedure to modify the side-chain oxygens in NASP and NGLU (images not shown). Only these amino acids need to be modified in this file, because N-terminus amino acids do not contain extra carboxylate groups.

  - Save and close the file when you are done.

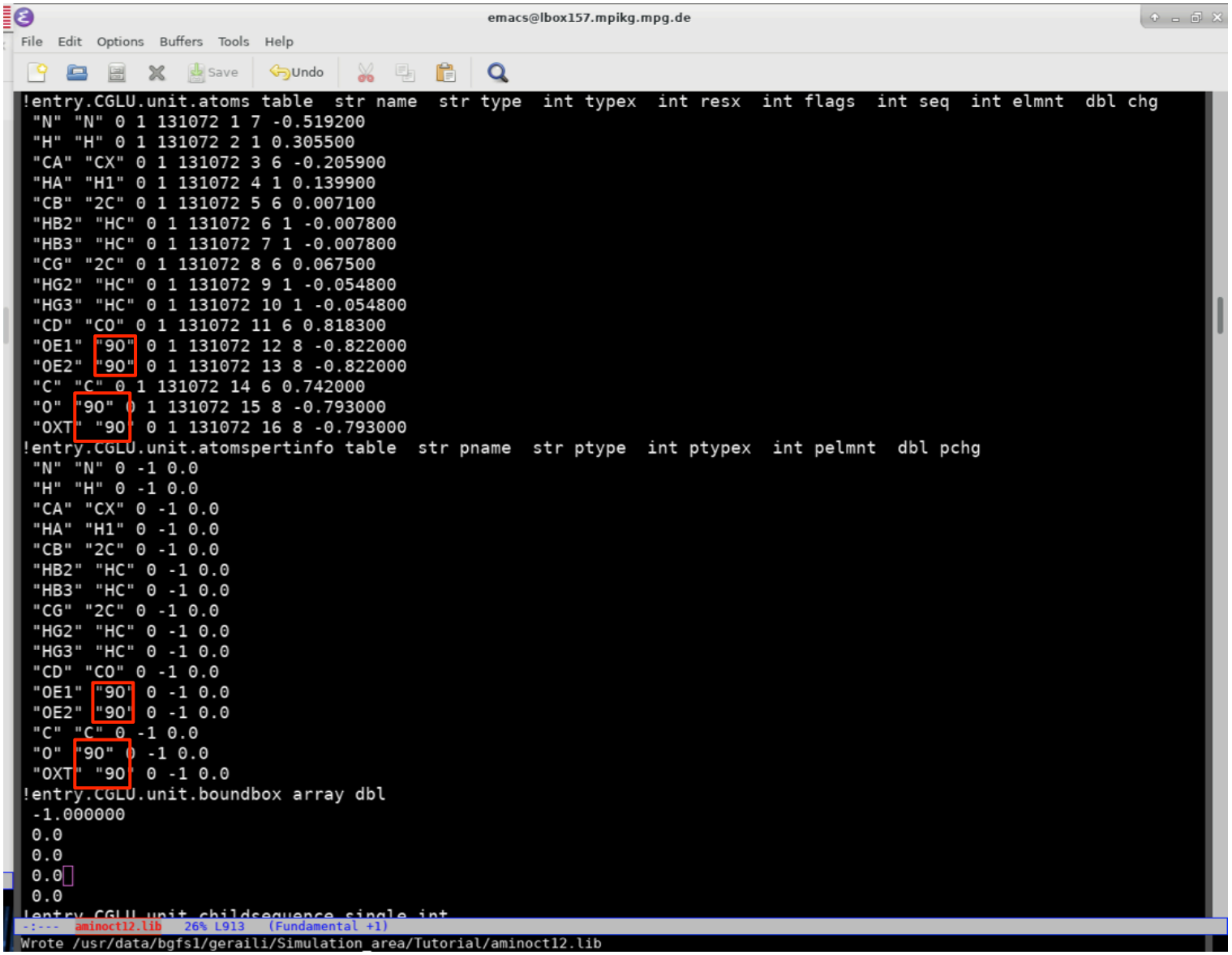

Fig. H: modified section of *aminoct12.lib*

- In the *leaprc.protein.ff14SB* file, copy the line where “O2” is defined (green box in Fig. I), and paste it directly below it as a new line.
- In the new line, change the “O2” to “YOUR NEW NAME” (in our example, the new name is “9O”; red box in Fig. I).
- Save and close the file.

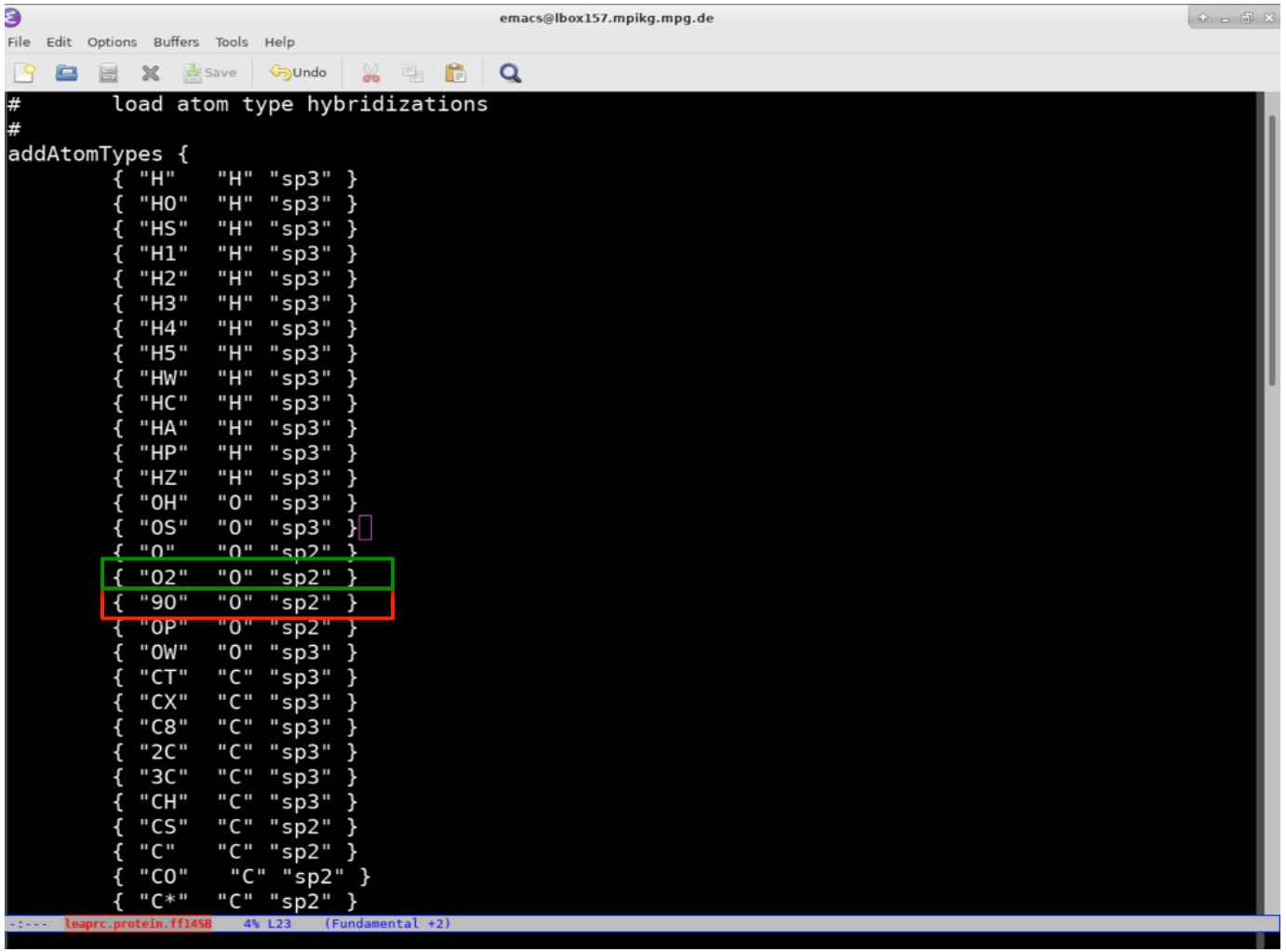

Fig. I: modified section of *leaprc.protein.ff14SB*

- The actual parameter values - atomic masses, charges, Lennard-Jones, bonds, angles, dihedral and improper potentials - for each interaction between 2 or more atoms are in files *parm10.dat* and *frcmod.ff14SB*.
  - *parm10.dat* is the main file; it contains the parameters for most interactions, and it is **never** modified.
  - *frcmod.ff14SB* supplements *parm10.dat* in two different ways: 1) it contains parameters for any new interaction; 2) it may also contain new parameters for any interaction that is already defined in *parm10.dat* but which we wish to override.
    - "tleap" reads *parm10.dat* first and *frcmod.ff14SB* afterwards; if it finds parameters for the same interaction in both files, it will use those from *frcmod.ff14SB* only.
  - In the next few slides we will illustrate the procedure to include in *frcmod.ff14SB* parameters for all interactions involving our new atom type "9O". In general, the steps in the procedure are (do not worry if you do not understand the procedure at this point; the examples will clarify it):
    - step 1: for each reference to carboxylate-"O2" in *parm10.dat*, check if the same interaction has been defined in *frcmod.ff14SB*.
      - If yes, duplicate the line in *frcmod.ff14SB* and past the duplicate directly below the original line; in the duplicate line, change "O2" to "9O".
      - if no, copy the line in *parm10.dat*, paste it at the end of the corresponding section in *frcmod.ff14SB*; in the copy, change "O2" to "9O".
      - by following these instructions, you make sure that all parameters in *parm10.dat* for carboxylate-"O2" are put into *frcmod.ff14SB*, without overwriting newer parameters for the same interaction if they exist in *frcmod.ff14SB*.
    - step 2: once you're done with step 1, you need to search for all references to carboxylate-"O2" in *frcmod.ff14SB*, and make sure that equivalent information is defined for "9O". This step is necessary because some interactions may only be defined in *frcmod.ff14SB*.

- Entering mass information for “9O”:
  - Look for the mass of “O2” in *parm10.dat* (Fig. J, red box).
  - Inspect the MASS section in *frcmod.ff14SB*. There is no entry for “O2” there.
    - Copy the line containing the mass information from *parm10.dat*, paste it at the end of the MASS section in *frcmod.ff14SB* and change the atom type from O2 to “9O” (Fig. K, red box).

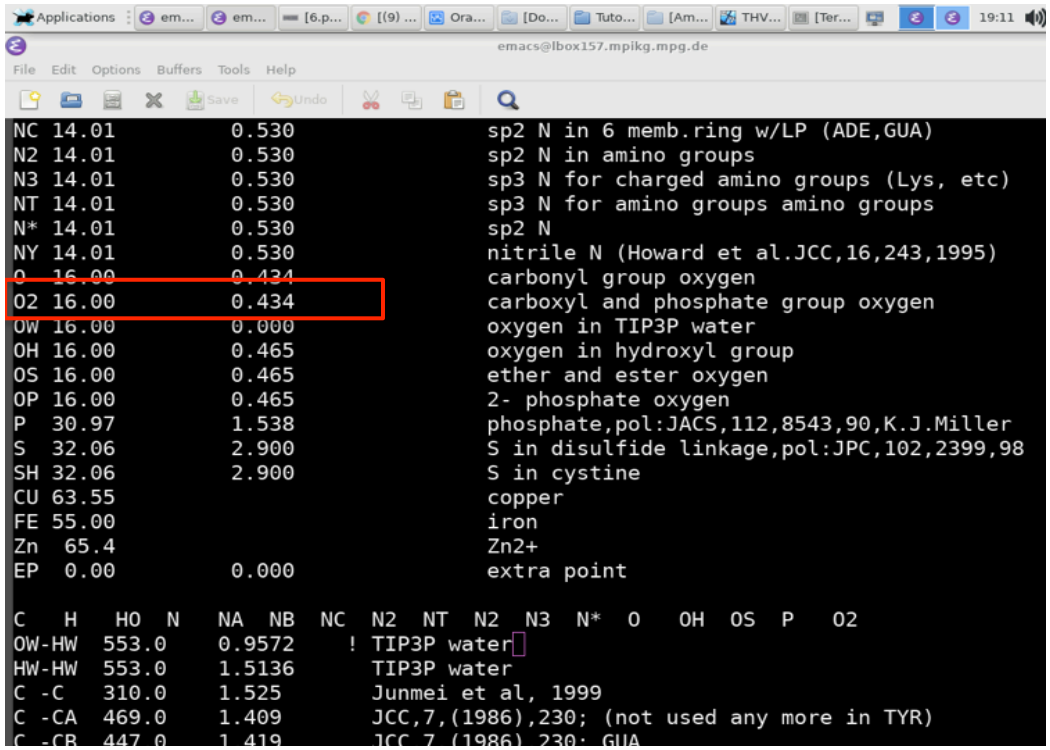

Fig. J: Original parm10.dat

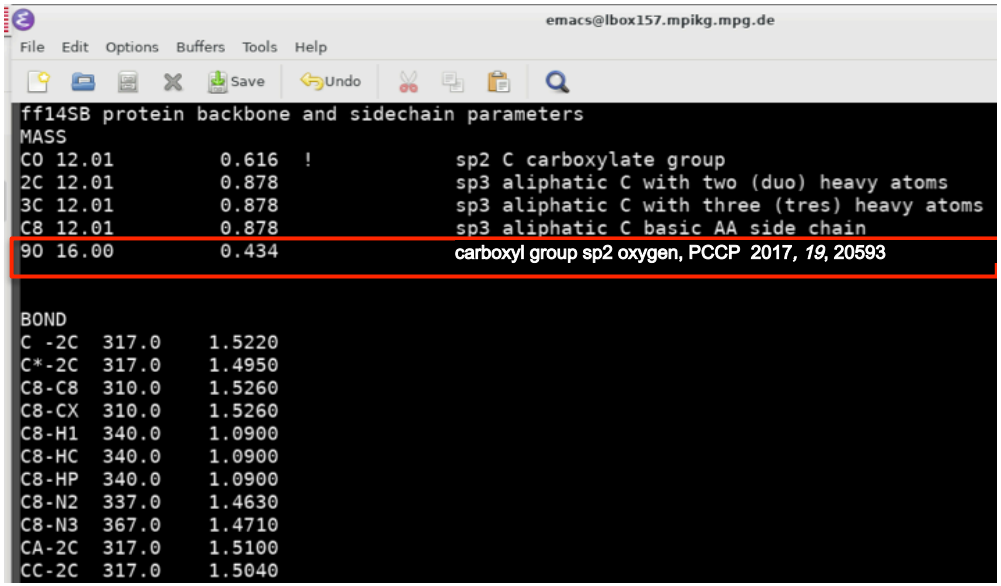

Fig. K: modified section of *frcmod.ff14SB*

- Entering parameters for bonds involving “9O”:
  - locate bond parameters involving “O2” in *parm10.dat*. Here is the relevant line:
    - C -O2 656.0 1.250 JCC,7,(1986),230; GLU,ASP (line 81)
    - O2-P 525.0 1.480 JCC,7,(1986),230; NA PHOSPHATES (line 173)
  - locate bond parameters involving “O2” in *frcmod.ff14SB*. There is only one line:
    - CO-O2 656.0 1.2500 (line 20)
  - Interactions “C –O2” and “O2-P” are only defined in *parm10.dat*. Copy the line for “C –O2” from *parm10.dat*, paste it at the end of the bond section in *frcmod.ff14SB* and change “O2” to “9O” (Fig. L, green box).
    - Note that you should **not** insert a line for the “9O-P” bond in *frcmod.ff14SB* . We developed these oxygen parameters specifically for carboxylates. Because of their high specificity, these parameters should not be used for anything other than the intended functional groups.
  - Interaction “CO-O2” only exists in *frcmod.ff14SB*. Duplicate this line, past it directly below the original and change “O2” to “9O” (Fig. L, red box).
    - Note: if “CO-O2” had simultaneously existed in *parm10.dat*, you would have ignored the *parm10.dat* data and would have done exactly the same.

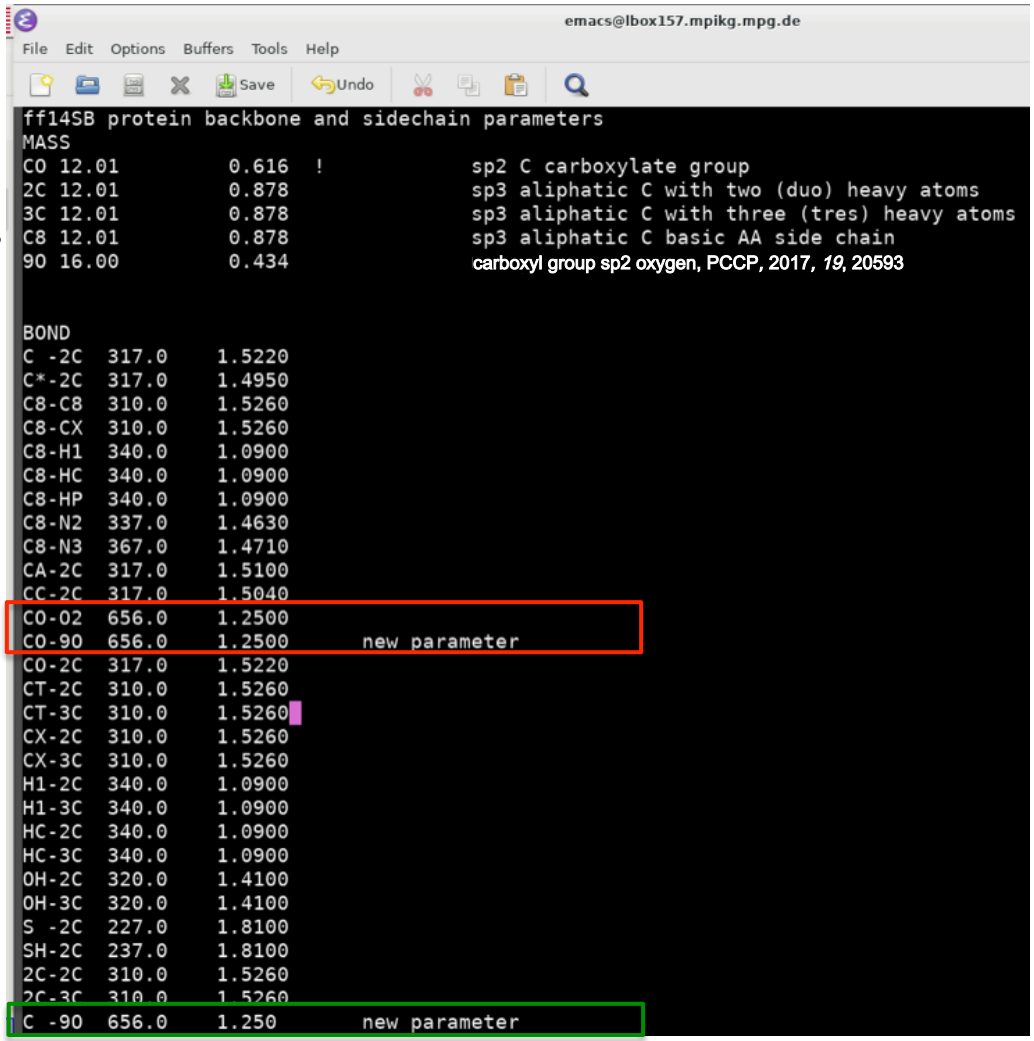

Fig. L: modified section of *frcmod.ff14SB*

- Entering parameters for angles, dihedrals or impropers involving carboxylate-“O2”:
  - To make these changes, you follow the same procedure we have exemplified for inserting parameters for bonds involving “9O”, with one difference: when you are changing the carboxylate-“O2” angle, dihedral, or improper, you must change them in a way to be able to consider any combination of interactions in the future.
    - search for carboxylate angles with “O2” in *parm10.dat*. There are multiple lines:
      - CT-C -O2 70.0 117.00 (line 237)
      - CX-C -O2 70.0 117.00 (was CT-C -O2) (line 238)
      - O2-C -O2 80.0 126.00 AA GLU (SCH JPC 79,2379) (line 254)
      - O2-P -OH 45.0 108.23 (line 539)
      - O2-P -O2 140.0 119.90 (line 540)
      - O2-P -OS 100.0 108.23 (line 543)
    - do the same thing in *frcmod.ff14SB*. Here are the lines:
      - O2-CO-O2 80.0 126.00
      - O2-CO-2C 70.0 117.00
    - Notice that each angle interaction is defined in only one file.
    - Duplicate the line with the “O2-CO-2C “ interaction in *frcmod.ff14SB* and paste it directly below the original. In the duplicate, change “O2” to “9O” (green box in Fig. M)
    - Duplicate the line with the “O2-CO-O2 “ interaction in *frcmod.ff14SB* and paste it twice directly below the original. In the duplicates, change “O2” to “9O” to allow all possible combinations of “O2” and “9O” (red box in Fig. M)
      - For our specific application, it would have been sufficient to insert only a line with “9O-CO-9O “ in *frcmod.ff14SB* because the *lib files we modified do not have an “O2-CO-O9” angle. We nevertheless suggest always including the combinations in *frcmod.ff14SB* to avoid unexpected problems if an angle “9O-CO-O2” ever becomes necessary.
    - Copy each of the lines from *parm10.dat* with angle interactions involving carboxlyate-“O2” and paste them one or more times (as necessary) at the end of the ANGLE section in *frcmod.ff14SB*. Make the necessary changes from “O2” to “9O” (images not shown).
    - Angle potentials involving the “O2-P” bond (in red above) should **not** be updated to “9O” in f*rcmod.ff14SB*, for the reasons explained in the previous slide.
    - Follow the same procedure to update dihedral and impropers (images not shown).

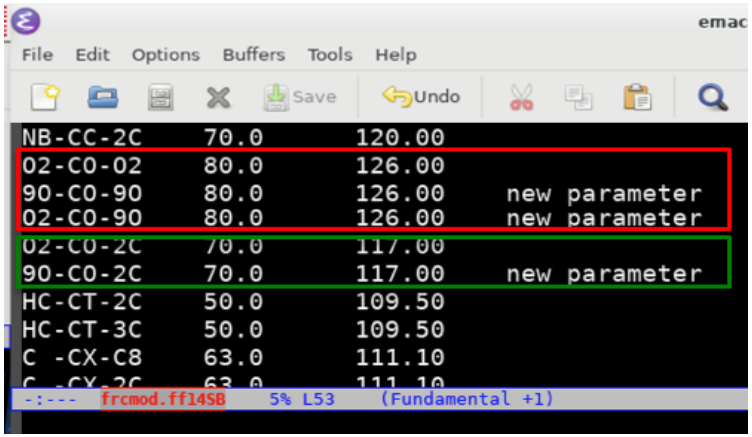

Fig. M: modified angle section of *frcmod.ff14SB*

- We will now input the new Lennard-Jones (LJ) parameters for atom type “9O”. LJ parameters exist in *parm10.dat* (Fig. N; red box shows parameters for “O2”), and in *frcmod.ff14SB*, in the “NONB” and “LJEDIT” sections (Fig. O, red boxes).
  - The “NONB” section will contain the new self-interaction parameters.
  - the “LJEDIT” will contain LJ parameters for those pairs of atoms for which we want to override the Lorentz-Berthelot combination rules.

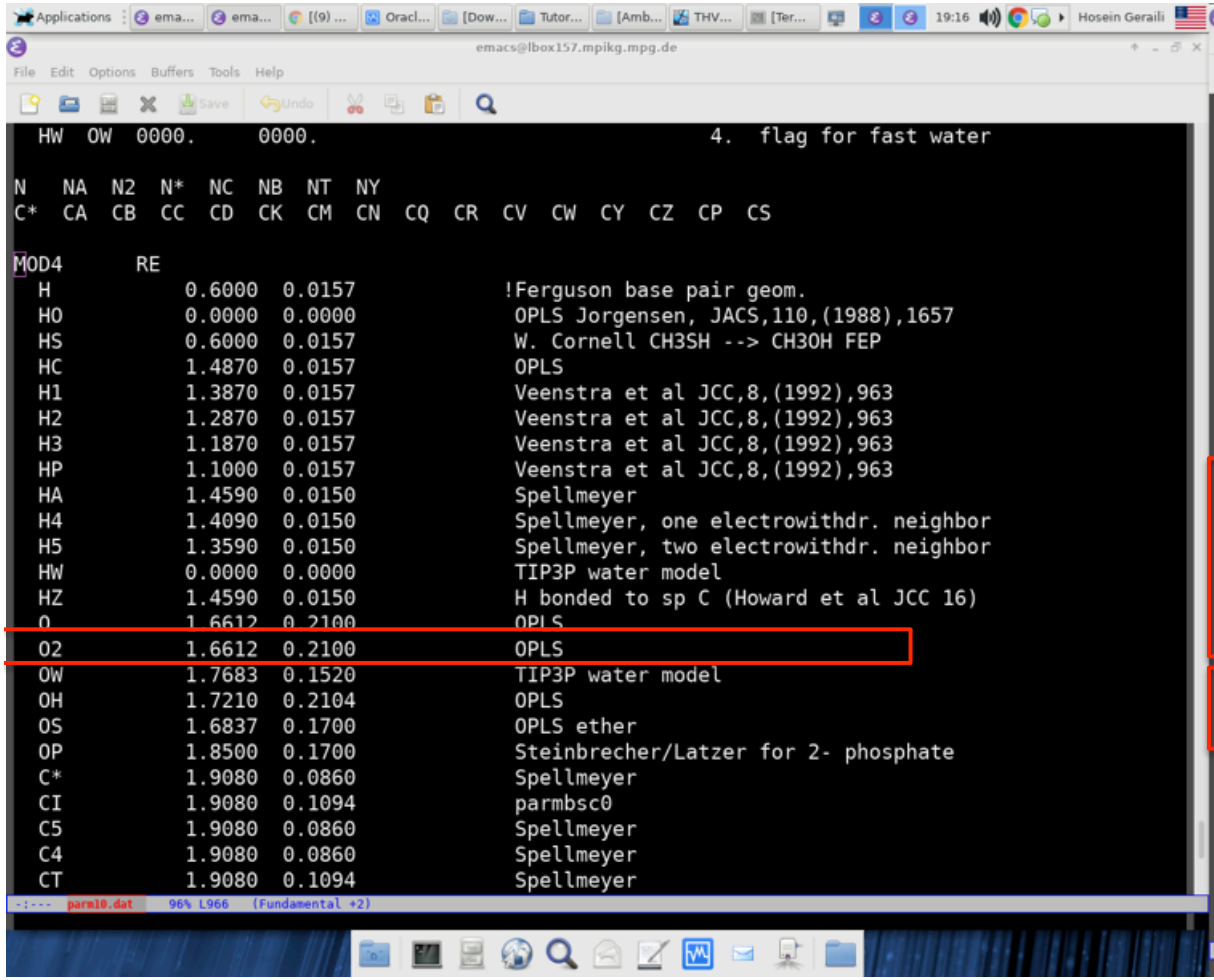

Fig. N: *parm10.dat*

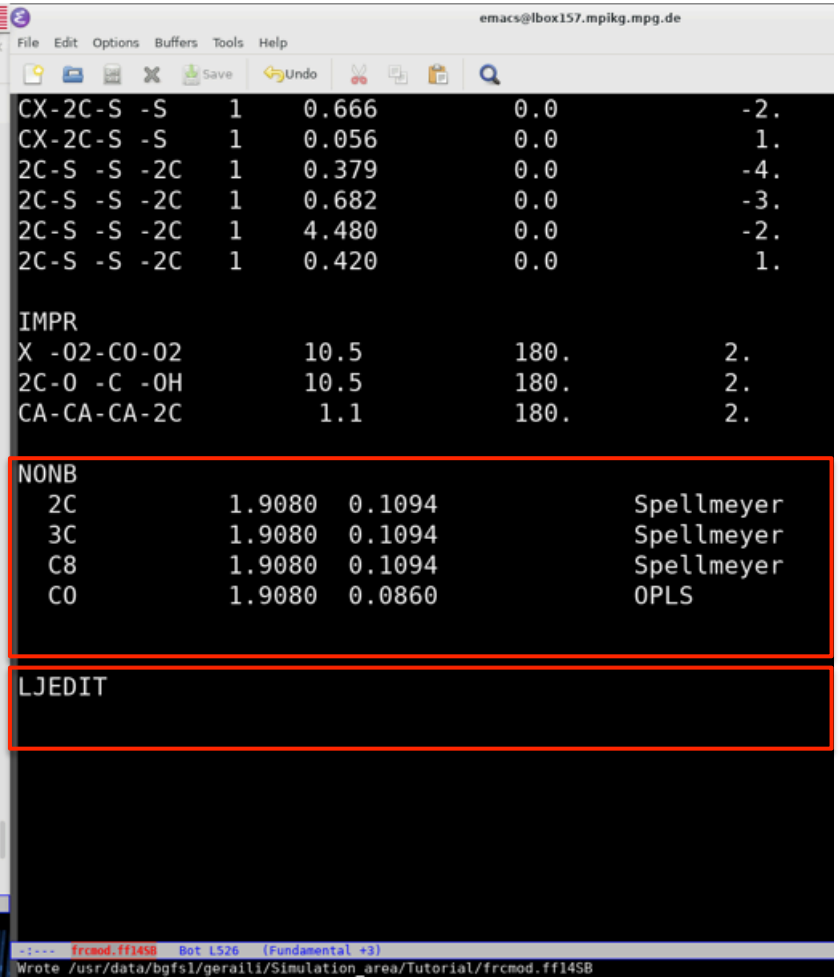

Fig. O: unmodified *frcmod.ff14SB*

- Download the *zip file that you'll find as supporting information of the article we mentioned in the introduction (doi: 10.1039/C7CP02557B). Open file Parameters/Acetate.top. This file contains the LJ parameter information in gromacs format; amber format parameters are given as comments. The relevant lines for this tutorial are:
  - ;name bond_type mass charge ptype sigma (nm) epsilon (kJ/mol) r(AMBER; A) epsilon(AMBER; kcal/mol)
  - OACE OACE 0.00000 0.00000 A 2.95992e-01 6.76553e-01 ; 1.6612 0.1617 ; 0.77x Original epsilon

  - ;i j func sigma epsilon Rij(AMBER; A) epsilon_ij(AMBER, kcal/mol)
  - OACE NA+ 1 2.75899e-1 4.97508e-1 ; 3.0969 0.1189 ; 1.022 times Original sigma (Rij in AMBER) ; OACE self iteraction correction for epsilon included

- "OACE" corresponds to our "9O". "r(AMBER)" and "epsilon (AMBER)" are the optimized self-interaction parameters. "r(AMBER)" is $R_{min,i,i}/2$ and "epsilon(AMBER)" is $\varepsilon_{i,i}$ in the notation of eq. 1 of this tutorial.
- "Rij(AMBER)" and "epsilon_ij(AMBER)" are the optimized LJ parameters for the interaction between "O9" and Na^+. Notice that "Rij(AMBER)" is $R_{min,i,j}$ in the notation of eq. 1.
- Notice that "r(AMBER)" is the same as in *parm10.dat* (Fig. N) but that "epsilon (AMBER)" is different: the optimized "r(AMBER)" is 0.77 times the original "r(AMBER)".
- Insert the new LJ self-interaction parameters ("r(AMBER)" and "epsilon (AMBER)") for "9O" into *frcmod.ff14SB* as shown in Fig. P.

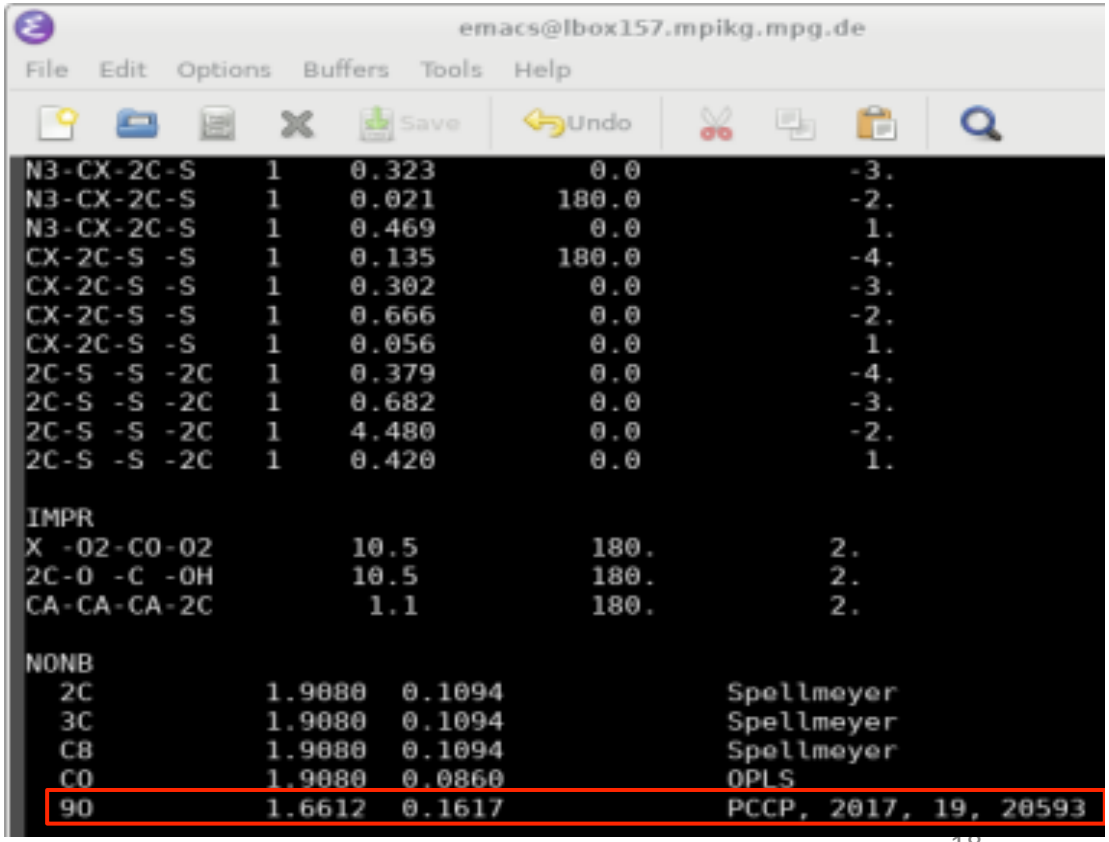

Fig. P: modified *frcmod.ff14SB*

- Now for the second (and shorter) part of this tutorial: modifying the LJ interactions between "9O" and Na^+:
- We will start by examining the original ion parameters (from JPCB 2008, 112, 9020-9041) typically used with TIP3P water. Open file Tutorial/*frcmod.ionsjc_tip3p*. The values under the NONBON section are $R_{min,i.i}/2$ (in A) and $\varepsilon_{i,i}$ (in kcal/mol). This file is loaded in "tleap" following the indications in *leaprc.water.tip3p*; to confirm this information, open *leaprc.water.tip3p* and search for the text string "frcmod.ionsjc_tip3p".

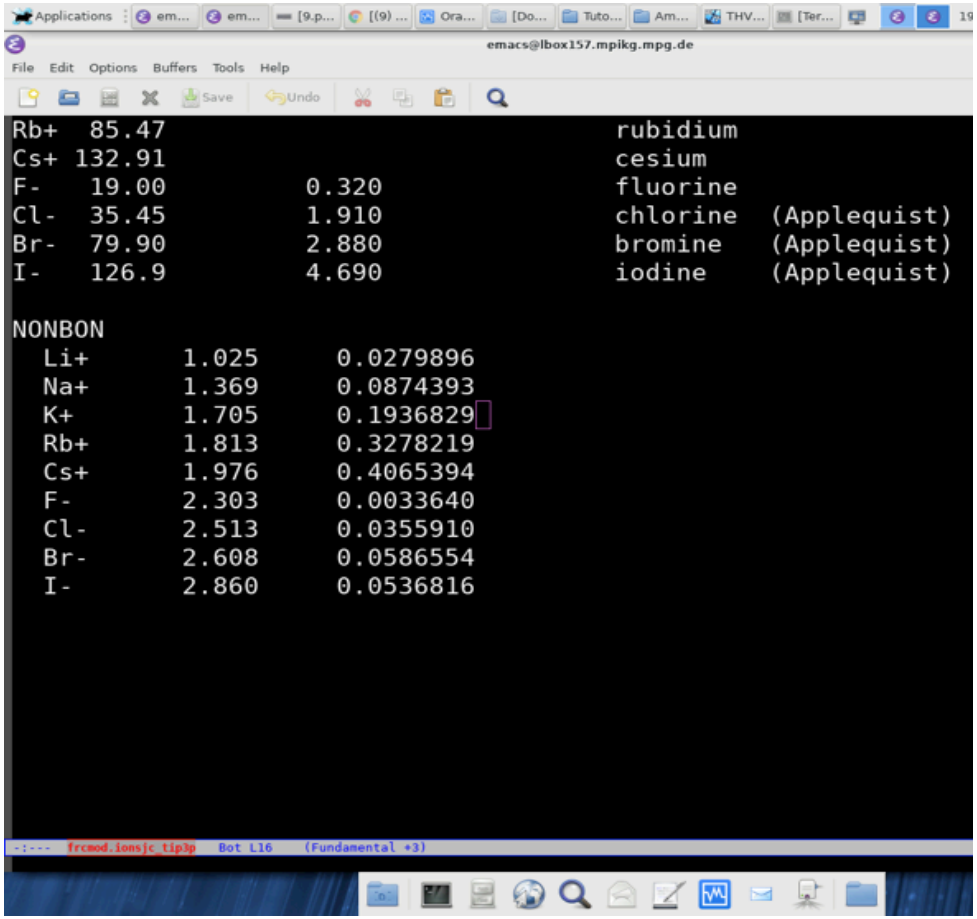

Fig. Q: original *frcmod.ionsjc_tip3p*

- Entries in the "LJEDIT" section in *frcmod.ff14SB* have the following (unexpected!) format:
  - atom_type_A atom_type_B d_A E_A d_B E_B ; "d" in angstrom; E in kcal/mol
  - these parameters are *exclusively* used to define the interactions between atom type A and atom type B as

$$R_{\min,A,B} = d_A + d_B \qquad \varepsilon_{A,B} = \sqrt{E_A E_B}$$

  - Notice that *any* combination of d_A and d_B, and of E_A and E_B, that yields the correct values $R_{min,A,B}$ and $\varepsilon_{A,B}$ is allowed. d_A,E_A and e_B,E_B are, by themselves, meaningless; only $R_{min,A,B}$ and $\varepsilon_{A,B}$ have meaning.
  - In the "LJEDIT" section of *frcmod.ff14SB*, add the new parameters for the "9O"…"Na+" interaction (red box, Fig. R).
  - Notice that you could have also written:

    9O Na+ 1.54845 0.1189 1.54845 0.1189

    or

    9O Na+ 0 1 3.0969 0.01414

    where 0.01414=0.1189*0.1189

  - Save and close *frcmod.ff14SB* when you are done.

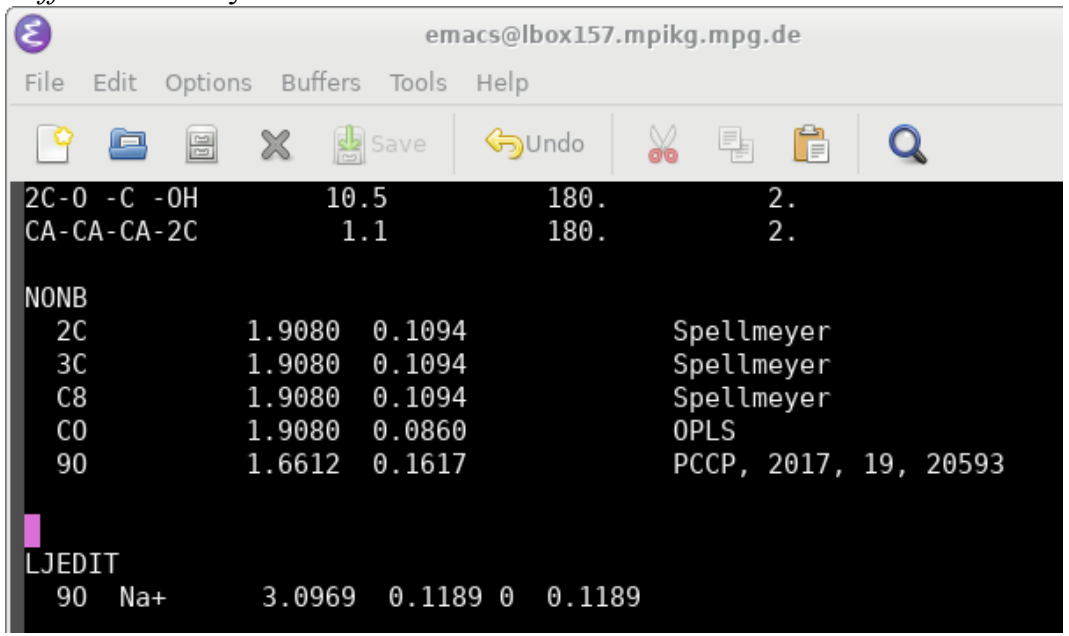

Fig. R

- The final step is to update the paths to your modified parameter and topology files so they are loaded by “tleap”.
  - Open Tutorial/*leaprc.protein.ff14SB* and update the paths of *parm10.dat*, *frcmod.ff14SB*, amino12.lib, *aminoct12.lib* and *aminont12.lib* to your Tutorial/ directory, as exemplified in Fig. S. We did not modify *parm10.dat* but it is convenient to call it also from Tutorial/ rather than from $AMBERHOME. Save and close *leaprc.protein.ff14SB* when you are done.
  - Create a new file (we called it *ff14_tleap_-f_ThisName.in*; Fig T) and add the following command:

    source /your/path/to/Tutorial/leaprc.protein.ff14SB

    - if you had made modifications in any of the water files, you could also add the command:

    source /your/path/to/Tutorial/leaprc.water.tip3p

    - save and close *ff14_tleap_-f_ThisName.in.*

- To build topology (.prmtop) and coordination (.inpcrd) files using the new parameters, open “tleap” with the following command:

  $ tleap –f “YOUR DIRECTORY”/ ff14_tleap_-f_ThisName.in

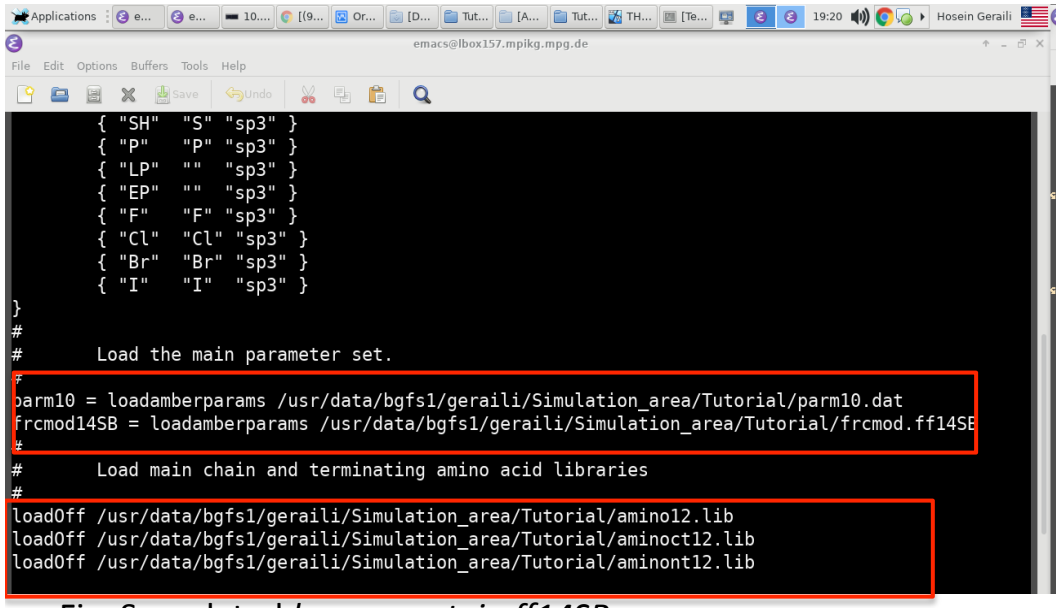

Fig. S: updated *leaprc.protein.ff14SB*

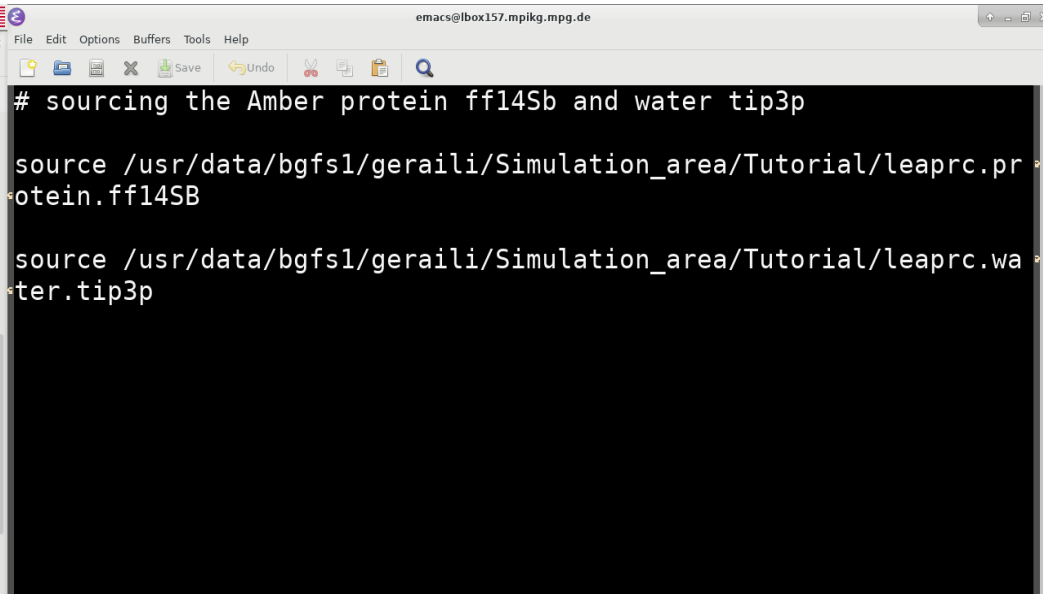

Fig. T: new file *ff14_tleap_-f_ThisName.in*

- It is always useful to verify if you've made the correct the modifications. To do so, download *wt1mg.pdb* from "ambermd.org" (http://ambermd.org/tutorials/advanced/tutorial8/files/wt1mg.pdb). This is a protein that has already been prepared for simulation. Put the pdb file in a "new_folder".
- Go to "new_folder", open a terminal there, and type all the commands in Fig. U (use your own path to Tutorial/, where you've saved all the modified topology and parameter files). After all the commands are executed, type "quit".

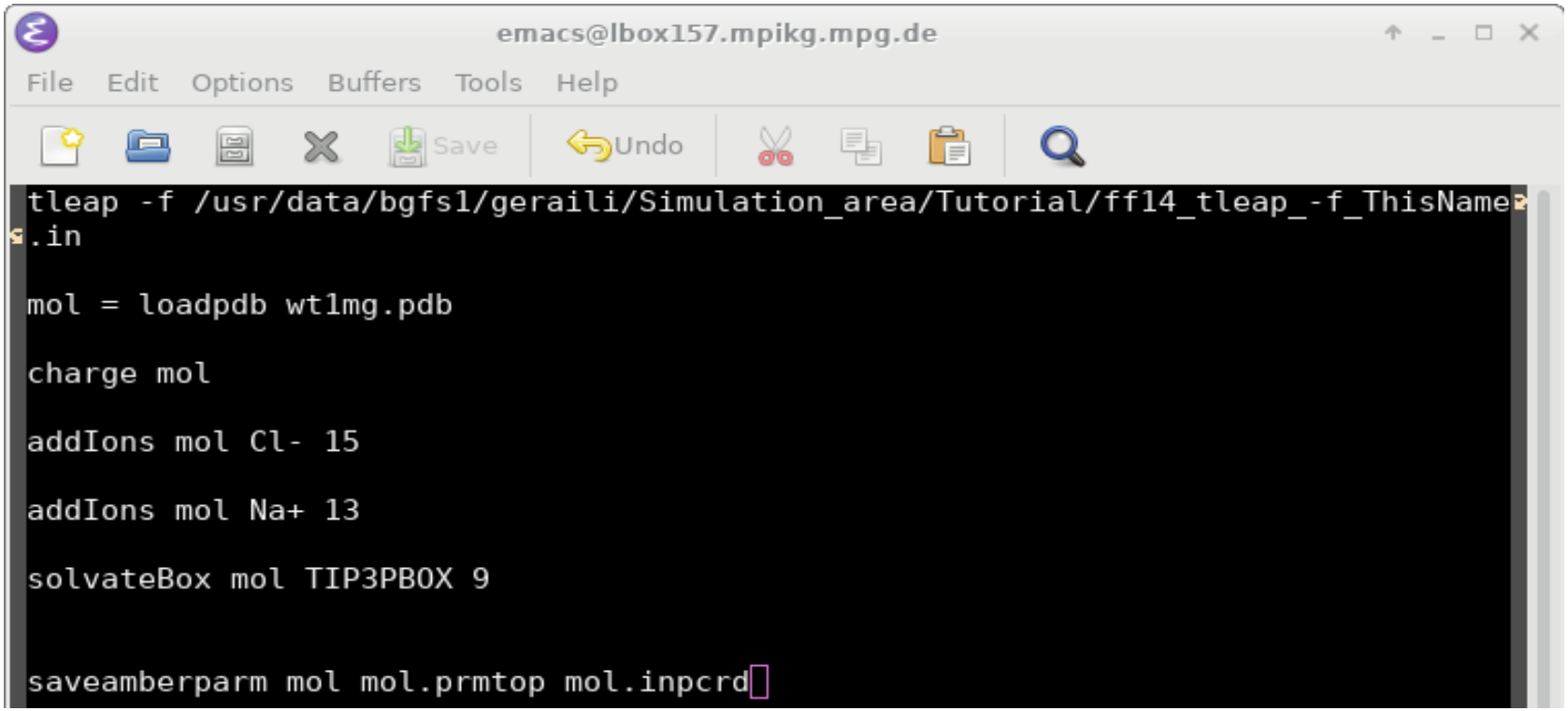

Fig. U

- Now we will use "Parmed" (https://parmed.github.io/ParmEd/html/index.html) to inspect the "mol.prmtop" topology file. Type the following commands in the terminal

  $ parmed –p mol.prmtop

  $ printLJMatrix :ASP

  $ printLJMatrix :Na+

  – The last column shows $\varepsilon_{i,j}$ in kcal/mol; the second-from-last column shows $R_{min,i,j}$ in angstrom.

- Take a look at the self-interaction parameters involving "9O" (Fig. V, red box) and at the interaction parameters between "9O" and "Na+" (Fig. V, green box). Compare them with those in file acetate.top, and with the parameters for "O2". Make sure there are no mistakes.

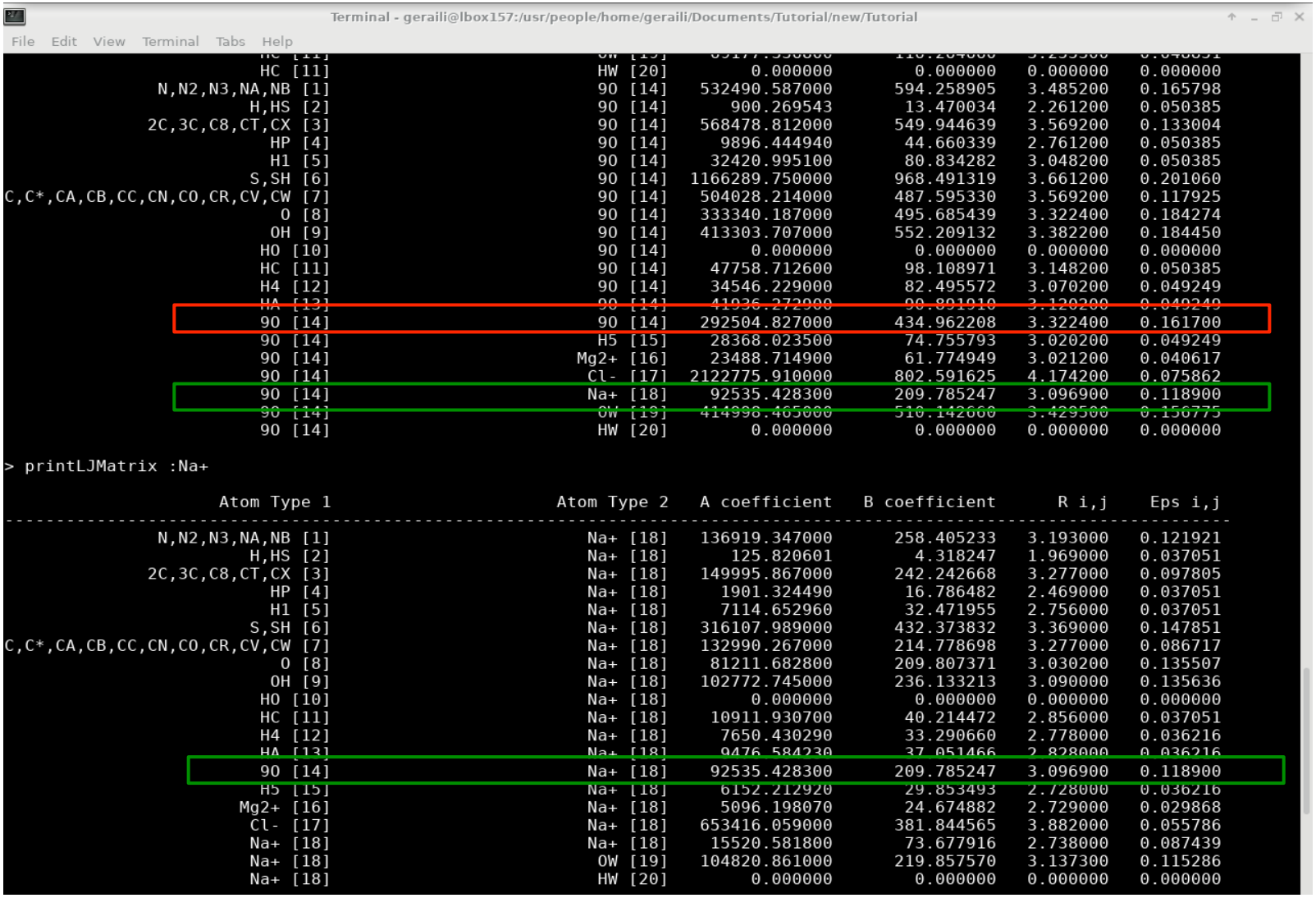

Fig. V

Well done; you’re good to go!