This example is a large scale, production quality flexible docking and simultaneous loop search with LMOD on D-amino acid oxidase (4QFC). The ligand and 4 loops in contact with the ligand are sampled while the FAD and the rest of the protein structure are kept frozen. The results are displayed in three jpeg images, one showing the docked ligand poses where purple is the X-ray, brown is the starting pose for LMOD search, which is very far from the X-ray, and the low-energy LMOD poses are colored green. About half a dozen of the lowest energy poses overlap with the X-ray pose virtually perfectly. The second image is zoomed out also showing the variation in loop conformations. Note that the green loop above the ligand serves as a lid to cover the active site and the most variation is expected in this loop. The third image shows the results of a different LMOD search, which starts out of the X-ray structure. Unlike the other search where the starting ligand pose was set to a very un-X-ray-like orientation to demonstrate the flexible docking capabilities of LMOD search, in this experiment the starting pose of the ligand was left at the X-ray position. The low-energy ligand poses and associated loop variation in the green lid shown in the image draw attention to two properties of this protein system: i) the ligand is very tightly docked to its X-ray pose with virtually no room for variation (note that this LMOD search was done in the same way as the other, including numerous explicit translations and rotations applied to the ligand) and ii) even a very limited search of only 10 LMOD iterations found a partially closed lid orientation. In the bundle of the green loops above the ligand, purple shows the X-ray conformation of the open lid and the dark green loop shows the partially closed lid. Of course, this is just an example and a thorough search would require at least a 100 LMOD iterations. The job results can be found in 4QFC_run_lmod_dock.log and 4QFC_run_lmod_dock.log_X-ray, and the LMOD generated low-energy conformations are saved in individual PDB files as well as in a multi-PDB file.

To run this job, one simply needs to compile 4QFC_lmod_dock.nab (which will read 4QFCpatched_protein_FAD_ligand.pdb (where the ligand is in an altered pose) or 4QFCpatched_protein_FAD_ligand_X-ray.pdb and the associated 4QFCpatched_protein_FAD_ligand.prmtop). It is highly recommended to compile it in parallel using mpinab, and run it it on 10+ cores. 10 LMOD iterations with 100 total LMOD search steps took about 10 hours running on 8 cores of a 3.6 GHZ i7 processor. In general, 100+ LMOD iterations are recommended for seriously exploring a protein system, which definitely requires parallel execution with as many cores as there are available. LMOD scales quite well thanks to the efficient force parallelization engine in AmberTools.

The structure of a typical LMOD nab script and associated job logfile is fully explained in the main AmberTools documentation under NAB: Molecular Dynamics and Mechanics/Low-MODe (LMOD) optimization methods, here some specifics are pointed out.

1. The general LMOD workflow for proteins includes i) fully flexible minimization of the entire system, ii) applying LMOD search on the partially frozen system meaning that the low frequency modes are those associated with the flexible (specified as “moving”) loops in the presence of the external field (electrostatic, van der Waals, and solvation terms) exerted by the frozen atoms, and iii) re-minimizing the low-energy loop conformations after unfreezing the whole system. (In this example iii) was not actually performed, but it is straightforward to edit the nab script to include it in the calculation.)
2. With flexible docking one ore more ligands are explicitly translated and rotated on top of any low-mode displacement, see the documentation about defining the ligand(s) and setting the ro-translational parameters as shown in the following block in the nab script.

```
tr_min[1]    = 0.1;
tr_max[1]    = 1.0;
rot_min[1]   = 30.0;
rot_max[1]   = 180.0;
lig_start[1] = 5326;
lig_end[1]   = 5361;
lig_cent[1]  = 0;
```

3. It is good practice to use a hardware generated seed for the random number generator and save this seed value in the file names associated with an LMOD job. In the example script the number of seconds passed since zero hour, January first, 1970 is used for this purpose.
4. Noteworthy LMOD parameters: `lo.nmod=15` is the number of lowest frequency modes used, `lo.kmod=5` means that out of the 15, every new LMOD iteration explores 5 randomly selected modes (and follows them in both directions), and `lo.nrotran_dof=0` refers to the presence of frozen (or tethered) atoms. `lo.energy_window=50` is set intentionally very high to generate a variety of different ligand poses and associated loop conformations and not only focus on the lowest energy ones. `lo.conf_separation_rms=1.0` means that in the final set of LMOD generated loops every single pair of them will be at least 1.0 Å superposition RMSD apart. This RMSD calculation includes all moving atoms (LMOD has no knowledge of atom types, etc.). This parameter is quite useful in controlling the diversity of low-energy loops, or conformations in general. `lo.nof_lmod_steps=0` instructs LMOD to try determining a barrier passing event automatically (see documentation). For example, something like this in the logfile
```
15 /10 E = -2019.205 (0.082)  Rg = 4.332  rmsd= 0.962  p= 0.0715
15 / 5 E = -2009.059 (0.090)  Rg = 4.318  rmsd= 0.772  p= 0.0024
```
means that starting from a particular conformation, low-frequency mode #15 is explored in both directions. In one direction LMOD determined that after 10 LMOD steps (so-called zig-zag curvilinear perturbation, see documentation) a barrier was passed whereas the other direction required only 5 steps. After minimization, the resulting new ligand and flexible loop conformations are, respectively, displaced by 0.962 and 0.772 Å RMSD from the same starting conformation. Rg is the radius of gyration (of the moving atoms) and p is the Boltzmann probability of a conformation with respect to the starting conformation in that particular LMOD iteration. `lo.mc_option=2` sets the Monte Carlo search to “total quenching” which means that every new LMOD iteration continues the search with the lowest-energy conformation found in the previous iteration, irrespective of probability. This option is usually the best for finding the lowest energy structures. For a broad search, however, `lo.mc_option=1` is recommended, which is the traditional Metropolis criterion applied to the minimized energies. `lo.nlig=1` sets the number of ligands to one and `lo.apply_rigdock=1` means that every single LMOD iteration will include multiple ro-translational perturbations applied to the ligand in addition to the low-mode step on the entire moving system.
5. Note the following block taken from the logfile:
```
   6  *E = -2039.368 (0.091)  Rg = 4.135
2  / 5 E = -2012.577 (0.094)  Rg = 4.179  rmsd= 1.954  p= 0.0001
2  /14 E = -2018.969 (0.092)  Rg = 4.091  rmsd= 0.856  p= 0.0011
3  / 3 E = -2038.989 (0.087)  Rg = 4.150  rmsd= 0.157  p= 0.8813
3  / 4 E = -2039.734 (0.096)  Rg = 4.143  rmsd= 0.140  p= 1.0000
5  / 7 E = -2038.168 (0.084)  Rg = 4.125  rmsd= 0.372  p= 0.6703
```

```
5  / 2 E = -2038.895 (0.086)  Rg = 4.151  rmsd= 0.224  p= 0.8543
7  / 5 E = -2036.223 (0.061)  Rg = 4.166  rmsd= 0.686  p= 0.3505
7  /20 E = -2015.546 (0.048)  Rg = 4.446  rmsd= 1.905  p= 0.0004
10 / 4 E = -2035.521 (0.079)  Rg = 4.198  rmsd= 0.690  p= 0.2774
10 / 4 E = -2031.912 (0.089)  Rg = 4.135  rmsd= 0.540  p= 0.0833
-------------------------------------------------------------------
   7  *E = -2053.197 (0.091)  Rg = 4.159
```

In LMOD iteration #6 the starting conformation has an energy of E = -2039.368 [kcal/mol]. This conformation is perturbed via LMOD zig-zag displacements along 5 different low-frequency modes (in both directions) and will result in 10 new conformations. Note that so far this does not involve any docking perturbation. In a non-docking search, the selected new starting point in iteration #7 using total quench, would be the conformation with the lowest energy E= -2039.734 and p=1.0, however, as is shown above the energy of the new starting point in iteration #7 is in fact much lower, E= -2053.197. The reason is because in addition to the LMOD moves, `lo.nof_poses_to_try=20` explicit translations and rotations are applied to the ligand, which may or may not lower the energy (in this step it does) but in any case the next new starting point is the lowest energy pose out of the 20 tries. Also note that the * in front of E means that the low-frequency eigen-modes have been recalculated for the new starting point.

With any questions or comments please contact me at istvan@kolossvary.hu or ikolossv@bu.edu.