Supplementary Material

**Solvated and Generalized Born calculations differences using GPU CUDA and multi-CPU
simulations of an antifreeze protein with AMBER**

Antonio Peramo*

Facultad de Ciencias, Escuela Politécnica Superior del Chimborazo, Riobamba, Ecuador

***Corresponding author,**

Escuela de Física y Matemáticas

Facultad de Ciencias

Escuela Politécnica Superior del Chimborazo

Riobamba, Ecuador

S1. Technical considerations to maximize performance for GPU based calculations

As suggested by Amber developers in the Amber manual (see page 247), we used the following values and parametrization for the GPU runs.

1. Values used for NTPR, NTWX, NTWV, NTWE and NTWR were not small. As suggested by Amber developers we did not set any of these values to less than 100.
2. Avoid setting `ntave /= 0`.
3. Avoid using the NPT ensemble (NTB=2) when it is not required. In some cases we used the NPT ensemble, because it was required for our simulations.
4. Avoid the use of GBSA in implicit solvent GB simulations unless required. We did not run GBSA simulations.
5. We used Langevin dynamics in most cases and the Berendsen Thermostat (*ntt=1*) in others.
6. Turn off ECC (C2050 and later). The GTX780 GPU does not have ECC control capabilities.
7. Turn on boost clocks if supported. This was turned off.

S2. Structure preparation

For the 1KDF molecule not generated with *tleap*, the initial structure was obtained from the protein data bank (rcsb.org). Using the 1KDF PDB file as input structure, protons were added according to the calculated ionization states of the chemical groups at pH 7.0. We used the H++ server with the output files format containing PDB and charge information. The corresponding coordinate and topology files were generated using the AMBER99SB force field. By default, the H++ server treats only GLU, ASP, ARG, CYS, LYS, HIP (AMBER name for the doubly protonated HIS), and TYR as titratable. Since protonation states may be affected by ϵ choice, we used a standard 80 value for the external (solvent) and 10 for the internal epsilon. This value was selected after comparing the desolvation penalty (labeled `delta_self`) produced by the H++ server, which generally showed a very small desolvation

penalty for all the residues, with the exception of TYR-63 which was around 2pK units. This selection for internal ϵ is reasonable in view of the fact that this type of AFP do not have many titrable sites and that most of them are not buried. The “flip” option in the H++ server was also selected to correct for the orientation of the amide N and O atoms in ASN and GLN.

S3. Input script for *tleap*

The utility *tleap* was used to generate coordinate and topology files for the extended conformation of the molecule for gas phase runs in GB calculations. The basic input script was as follow:

```
tleap -f - <<EOF
source leaprc.ff12SB
loadamberparams frcmod.ionsjc_tip3p
set default PBRadii mbondi2
1kdf = sequence { NASN GLN ALA SER VAL VAL ALA ASN GLN LEU ILE PRO ILE ASN THR ALA LEU THR LEU
VAL MET MET ARG SER GLU VAL VAL THR PRO VAL GLY ILE PRO ALA GLU ASP ILE PRO ARG LEU VAL SER
MET GLN VAL ASN ARG ALA VAL PRO LEU GLY THR THR LEU MET PRO ASP MET VAL LYS GLY TYR ALA
CALA }
savepdb 1kdf 1kdf_vac_normal.pdb
saveamberparm 1kdf 1kdf_vac_normal.top 1kdf_vac_normal.crd
addions 1kdf Na+ 0
addions 1kdf Cl- 0
solvatebox 1kdf TIP3PBOX 10.0
savepdb 1kdf 1kdf_wat_normal.pdb
saveamberparm 1kdf 1kdf_wat_normal.top 1kdf_wat_normal.crd
quit
EOF
```

For the variation in the Born calculations relative to the Born radii, the “set default PBRadii mbondi2” was excluded from the script resulting in the use of the default *mbondi* parameters.

S4. Input files for MD computations, except for aMD

Input files (mdin) are included in Table 1 below and described next. Temperatures shown are for 268K runs (except for the JAC benchmark, at 300K).

A. Input file for the Joint Amber-Charmm (JAC) benchmark for results shown in Table 4 and Table 6

of the main document. The benchmark uses dihydrofolate reductase, dhfr, in an explicit water bath with cubic periodic boundary conditions. Cubic periodic box, 62.23 Å dimension, 8Å nonbond cutoff with 2Å buffer; 2 fs timestep, 1000 steps. equilibration temperature 300K. Input file was the following, as included in the JAC benchmark, with the difference that we used ASCII output (ioutfm=0), instead of netCDF. :

```
&cntrl
ntx=5, irest=1, ntc=2, ntf=2,
nstlim=500000, ntp=1000, ntwx=5000,
ntwr=100000, dt=0.002, cut=8.,
ntt=1, taup=10.0,
temp0=300.0,
ntb=2, ntp=1,
ioutfm=0,
ntb=2, ntp=1, taup=10.0,
ioutfm=1,
/
```

B. Input files for the production steps of the AMBER GPU reproducibility tests and for pmemd and sander reproducibility runs for results shown in Table 5, Table 6, Table 7 and Figures 2, 3 and 4 for the solvated crystal structure of 1KDF. As general procedure, solvated or gas phase structures were first minimized for 1000 steps with the initial 500 steps using the steepest descent algorithm with the final 500 steps using the conjugate gradient energy minimization with constraints were applied to the protein residues. This was followed by a second minimization procedure of 5000 steps, with the last 2000 using the conjugate gradient without any restrains to the system. The system was heated to the specified temperatures for 100 ps and production trajectories of 0.9 ns were obtained in isothermal-isobaric ensemble (1atm), for a total run of 1ns. Except as noted for some GB simulations, where the Berendsen thermostat was implemented, all other simulations used Langevin dynamics; the SHAKE-bond-length constraints applied to all the bonds involving the H atom and the Particle Mesh Ewald method to calculate long-range electrostatic interactions as implemented in AMBER. Time step was 1.0 fs. A formatted restart file, and default values of all other inputs of were used for the program *pmemd*. Where applicable, simulations were performed with a 10Å ° cut-off for non-bonded interactions. 1-4

EEL scale factors were used with the default value of 1.2 and 1-4 VDW scale factors used the default value of 2.0. Implicit solvent radii for these simulations were modified Bondi radii (mbondi2).

C) Input file for the production step of the gas phase GB calculations for the crystal structure of 1KDF.

Same as above, for the GB simulations of the 1KDF crystal structure for results shown in Figures 4 and 5.

D) Input files for the production step of the GB calculations with the extended sequence generated with tleap for results shown in Figures 5, 6 and 7. These were run for 268K and 298K for 50ns in 10 steps of 5ns each, and some were extended up to 250ns. The input file in D) is for the Langevin dynamics run used in all calculations (*pmemd.cuda*, *pmemd.mpi*) in both the Exxact workstation and the HP clone (see Table 1 for description). For the runs using the Berendsen thermostat, parameters were set as follows: *ntt=1*; *gamma_ln* not used.

<p>A. &cntrl irest=1, ntc=2, ntf=2, nstlim=500000, ntp=1000, ntwx=5000, ntwr=100000, dt=0.002, cut=8., ntt=1, taup=10.0, temp0=300.0, ntb=2, ntp=1, ntx=5, taup=10.0, ioutfm=0</p>	<p>B. &cntrl imin = 0, irest = 1, ntb = 2, pres0 = 1.0, ntp = 1, , ntx = 7, taup = 2.0, cut = 10.0, ntr = 0, ntc = 2, ntf = 2, tempi = 268.0, temp0 = 268.0, ntt = 3, gamma_ln = 1.0, nstlim = 900000, dt = 0.001, ntp = 1000, ntwx = 1000, ntwr = 10000, / &ewald skinnb = 2.0</p>	<p>C. &cntrl imin = 0, irest = 1, ntx = 7, ntb = 0, ntp = 0, igb = 1, taup = 2.0, cut = 999., rgbmax = 999., ntr = 0, ntc = 2, ntf = 2, tempi = 268.0, temp0 = 268.0, ntt = 3, gamma_ln = 1.0, nstlim = 900000, dt = 0.001, ntp = 1000, ntwx = 1000, ntwr = 10000, ig = -1 / &ewald skinnb = 2.0</p>	<p>D. &cntrl imin=0, irest=1, ntx=5, nstlim=2500000, dt=0.002, ntc=2, ntf=2, ntt=3, ntp=0, tautp=0.5, gamma_ln=2.0, tempi=268.0, temp0=268.0, ntpr=1000,ntwx=1000, ntwr=20000, ntb=0, igb=1, cut=999.,rgbmax=999., ig=-1 / &ewald skinnb = 2.0</p>
---	--	---	---

S5. aMD input files and parameters, for results shown in Fig 8.

Here we followed the procedure reported in reference 9, in which the system was prepared by several steps of water minimization, relaxation, followed by whole system minimization, heating, reaching

stable density and equilibration of 5ns. and then for a final production run with the boosted potential for a total of 100ns. We note that although the system was prepared with many more steps than in the non-aMD runs, the results *before* implementing the boost to the potentials were the same for the energies, as shown in Fig 8 in the manuscript.

Table 2. Sequence of input scripts to run the aMD simulation for the solvated 1KDF crystal structure.

Final temperature shown is 248K; values changed accordingly for other temperatures.

<pre> water minimization: &cntrl imin=1, ntmin=1, nmropt=0, drms=0.1, maxcyc=2000, ncyc=1500, ntx=1, irest=0, ntp=100, ntwr=100, iwrap=0, ntf=1, ntb=1, cut=10.0, nsnb=20, igb=0, ibelly=0, ntr=1, restraintmask="!:WAT", restraint_wt=10.0 / </pre>	<pre> Water relaxation &cntrl timlim = 999999., nmropt = 0, imin = 0, ntx = 1, irest = 0, ntr = 1, ntmo = 1, ntp = 500, ntwx = 500, ntwv = 0, ntwe = 0, ntwr = 5000, ntf = 2, ntb = 2, cut = 10.0, nsnb = 20, nstlim = 10000, nscm = 2500, iwrap = 1, t = 0.0, dt = 0.002, temp0 = 248.0, tempi = 200.0, tautp=0.5, ntt = 1, ntp = 1, taup = 1.0, ntc = 2, tol = 0.00001, ibelly=0, ntr=1, restraintmask="!:WAT", restraint_wt=10.0 / &ewald skinnb = 2.0 </pre>	<pre> Whole system minimization: &cntrl imin=1, ntmin=1, nmropt=0, drms=0.1 maxcyc=2000, ncyc=1500, ntx=1, irest=0, ntpr=100, ntwr=100, iwrap=0, ntf=1, ntb=1, cut=10.0, nsnb=20, igb=0, ibelly=0, ntr=0, / </pre>
<pre> Heating the system &cntrl imin=0, nmropt=1, ntx=1, irest=0, ntpr=500, ntwr=500, ntwx=500, iwrap=1, ntf=2, ntb=1, cut=10.0, nsnb=20, igb=0, ibelly=0, ntr=1, nstlim=250000, nscm=500, dt=0.002, ntt=1, temp0=0.0, tempi=0.0, tautp=0.5, ntc=2, restraintmask=':1-65', restraint_wt=10.0, / &wt type='REST', istep1=0, istep2=0, value1=1.0, value2=1.0, &end &wt type='TEMP0', istep1=0, istep2=250000, value1=0.0, value2=248.0, &end &wt type='END' </pre>	<pre> Stabilize for density &cntrl imin=0, irest=1, ntx=5, nstlim=250000, dt=0.002, ntc=2, ntf=2, cut=10.0, ntb=2, ntp=1, taup=1.0, ntpr=500, ntwx=500, ntt=3, gamma_ln=2.0, temp0=248.0, iwrap=1, ntr=1, restraintmask=':1-65', restraint_wt=10.0, / </pre>	<pre> Equilibrate for 5ns &cntrl imin=0, irest=1, ntx=5, nstlim=250000, dt=0.002, ntc=2, ntf=2, ig=-1, cut=10.0, ntb=2, ntp=1, taup=2.0, ntpr=1000, ntwx=1000, ntt=3, gamma_ln=2.0, temp0=248.0, / </pre>

After equilibration, output files were analyzed for each temperature and the following NVT production run applied, where we used $iamd = 3$ to boost the whole potential with an extra boost to the torsions, where the EthreshD (Ed), alphaD (α D), EthreshP (Ep) and alphaP (α P) were calculated from the average potential and dihedrals obtained from the 5ns equilibration run. Parameter calculation followed

the standard values provided in the AMBER12 manual: 3.5 kcal/mol/residue and use of 0.2 as factor to calculate αD . The number of residues was 65 and the total number of atoms 12491.

```
&cntrl  
imin=0,irest=1,ntx=5,  
nstlim=50000000,dt=0.002,  
ntc=2,ntf=2,ig=-1,  
cut=10.0, ntb=1, ntp=0,  
ntpr=2000, ntwx=2000,  
ntt=3, gamma_ln=2.0,  
temp0=248.0,ioutfm=1,iwrap=1,  
iamd=3,  
ethreshd=887.75, alphad=45.5,  
ethreshp=-38743.71, alphap=2498.2,  
/
```

The previous mdin file was for 248K. Values for 268K were ethreshd=893.01, ethreshp=-37616.8; for 298K ethreshd=907.45, ethreshp=-35971.3; for 310K ethreshd=907.07, ethreshp=-35336.68; for 343K ethreshd=922.33, ethreshp=-33580.92 and for 373K ethreshd=933.72, ethreshp=-31993.11.

Supplementary Material Fig. 1 Backbone atoms root mean-square fluctuations (RMSFs) of 1KDF residues 5 to 60 with respect to the crystal structure at varied T for 1ns trajectories as obtained with the CUDA implementation; the CPU *pmemd* implementation and the *sander.mpi* implementation of AMBER12

Supplementary Material Fig. 2 Comparison of total energies (A), densities (B) and backbone atoms RMSD values for the GPU *pmemd.cuda* JAC benchmark 1ns runs at 300K using two different precision models, the DPDP (n=5) and the SPDP (n=20). Runs for each group showed bitwise results, as mentioned in Table 4 so only one line per precision model is shown . Note that the RMSD values for the DPDP precision arithmetics remain consistently above the SPDP precision for the entire simulation.