Characterization of MCL1 inhibition via Fast Switching Double annihilation technology on the CRESCO3 cluster

Piero Procacci[1*]

Dipartimento di Chimica, Università di Firenze, Via della Lastruccia 3, I-50019 Sesto Fiorentino, Italy

Abstract. In this report we present the MPI performances of the ORAC molecular dynamics code on the CRESCO3 cluster for the Fast Switching Double Annihilation method in the determination of drug-receptor dissociation free energies. We also present some preliminary FS-DAM results on the inhibition of the Myeloid Cell Leukemia 1 (MCL1) protein using new potent specific inhibitors. MCL1 is a key apoptosis regulator and its selective inhibition can be of great importance in cancer therapy.

1 Introduction

In the recent years, thanks to the constant support of the CRESCO infrastructure, we developed a new massively parallel technology for the computation of the dissociation free energies in drug-receptor systems, described at the atomistic level. The method, termed Fast Switching Double Annihilation (FS-DAM) and thoroughly described in Refs[1, 2, 3, 4, 5], is based on a combination of Replica Exchange with Solute Tempering technique (REST) and non equilibrium (NE) alchemical approach. In contrast to the mainstream equilibrium alchemical protocols such as FEP/REST[6], in the NE alchemical approach, the equilibrium sampling is required only at the starting fully coupled bound and bulk states (attainable using conventional enhanced sampling simulation methods such as REST), thus eliminating altogether the need for the optimization of the so-called thermodynamic length along the alchemical coordinate as in equilibrium FEP or Thermodynamic Integration studies.[7] In FS-DAM, the dissociation free energies are estimated from the distributions of the NE works[8] obtained by launching fast and independent annihilation trajectories (originated from the configurations harvested in the REST stage) of the bound ligand and of the ligand in bulk solvent. Most importantly, unlike in FEP, the confidence level in FS-DAM predictions (assuming no error from the force field), a crucial quantity in a industrial approach, can be determined very reliably using standard bootstrap analysis on the collection of the NE works. FS-DAM, implemented in our in-house ORAC code[5], is particularly suited for non uniform memory access (NUMA) multi-core architectures such as the CRESCOx clusters, relying on a weak scaling parallelism for REST and NE technologies on the MPI level and enforcing an OpenMP layer for the strong scaling parallelization of the bonded and non bonded forces within a single shared memory access compute node.

In this report we examine the performances of FS-DAM on the MPI level as evaluated on the CRESCO3 cluster using the ORAC OpenMP/MPI hybrid code. We also present some important preliminary results on the inhibition of the apoptosis regulator Myeloid Cell Leukemia 1 protein (MCL1) by a potent recently discovered[9] indole-amide compound.

*Corresponding author. E-mail: procacci@unifi.it.

2 MPI perform aces of the FS-DAM on CRESCO3

OpenMP tests for the ORAC (intranode shared memory) have been thoroughly discussed in Ref. [5, 10]. Here we present the performance of the code for the MPI layer only in order to assess the internode communications overhead in the FS-DAM approach.

As described elsewhere,[5, 10] for each ligand-receptor pair, the computation of the alchemical annihilation free energy via FS-DAM involves two computational step:

i) the acquisition of a Boltzmann sampling of the fully coupled state (the complex and the ligand in bulk) using Replica Exchange Simulation with solute tempering (hereinafter **REST**). For such stage, a simulation time on the target state from few to few tens of ns is acquired by running concurrent multiple REST batteries in parallel, for a total simulation time of the order of the microsecond.

ii) the fast switching annihilation of the ligand in the complex and in bulk solvent in a swarm of independent non equilibrium trajectories (hereinafter **FS-NE**), typically lasting from few tens to few hundreds of picoseconds for a total simulation time of the order of the microsecond. The selected ligand-

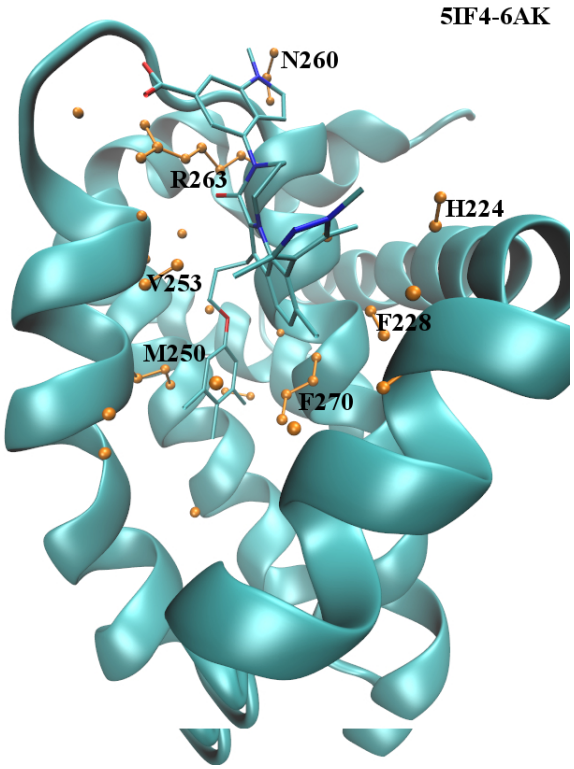

Figure 1: The MCL1-6AL complex [pdb id code 5if4] : The ribbon structure is the MCL1 protein. The 6AK indole ligand is depicted in the bond/stick representation. MCL1 atoms within 4.0 Å of the inhibitor are represented in ball and stick (orange color) along with residue specification (one letter code). Residue numbering is that of the pdb file 5if4.

receptor complex in this report, provided in the PDB file 5if4[9], is the Myeloid Cell Leukemia 1 (MCL1) protein in complex with a potent indole-amide inhibitor (indicated with 6AK). The complex is reported in Figure 2, where we show the main ligand-residue contacts that are involved in binding. The force field for the atomistic description of the ligand-receptor system is a combination of the AMBER99SB force field[11] for the protein, GAFF[6] for the 6AK ligand, and the TIP3 model[12] for the solvating water.

For the simulations of any kind, periodic boundary conditions are used while temperature and pressure are kept constant at 300 K and 1 Atm standardly using a Nosé thermostat and a Parrinello-Rahman Lagrangian, respectively.

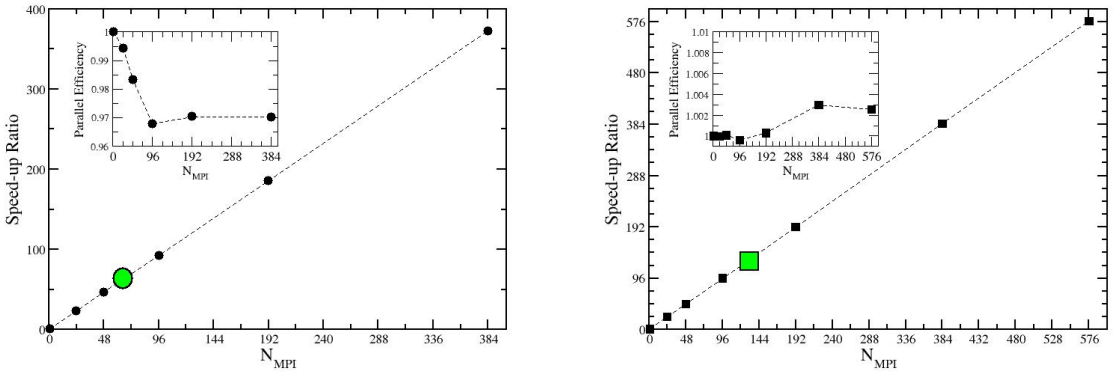

Figure 2: Left panel; Performance tests for the REST computational stage; the symbol in green corresponds to the interpolated speed up ratio for a typical value of 64 MPI instances (assuming 6-8 threads on the OpenMP layer). Right panel: Performance tests for the and FS-NE computational stage. The symbol in green corresponds to the interpolated speed up ratio for a typical value of 128 MPI instances (assuming 3-6 OpenMP threads).

In the Figures 1 and 2, we report the speed up ratio and parallel efficiency, *referred to the MPI layer only*, obtained for the MCL1-6AK protein in complex with an indole-amide inhibitor, a system of 11175 atoms) on the CRESCO3 HPC system. For the REST stage (see Figure 1), we run a REM simulation lasting 1.545 ps. Simulations were run (using in all cases one OpenMP thread per trajectories) with 1, 24,48,86,192 and 384 replicas (i.e. MPI processes) for a corresponding total simulation time of 1.545, 37.08, 74.16, 148.32, 296.64 and 593.28 picoseconds. All REST tests were done in an elapsed time ranging from 270.9 seconds (1 replica) to 279.2 seconds (384 replicas).

For FS-NE (see Figure 2), we launched NE trajectories lasting 0.6 ps. FS-NE parallel simulations were done producing 1, 24, 48, 96, 192, 384 and 576 independent non communicating trajectories, for a total simulation time of 0.6, 14.4, 28.8, 57.6, 115,2 , 230.4 and 345.6 picoseconds. All FS-NE tests were done in an elapsed time ranging from 119.8 to 120.2 seconds. The performances on the MPI level are practically ideal for both the REST and FS-NE computational stages, in a wide range including 48 (REST) and 128 (FS-NE) processors as required in the MCL1 project. While, expectedly, the embarrassingly parallel FS-NE computation exhibits no loss whatsoever of parallel efficiency (see the inset plot in Figure 2), in the REST computation a minor efficiency loss can be barely appreciated with the 384 replicas simulation running at 97% of parallel efficiency (see the inset plot in Figure 1). This moderate efficient loss in REST is due the communication overhead during replica exchange attempts. These are normally implemented each 15 fs (i.e. 103 attempts in total in the test) for all contiguous pairs of replicas and involve a `MPI_SEND / MPI_RECEIVE` couple of three double precision numbers (the REST scale factors) followed by an `MPI_BARRIER` on all MPI processes.

3 Dissociation Free energy of the MCL1-6AK complex via FS-DAM

The REST stage for the MCL1-6AK and the 6AK in bulk solvent, featuring about 11000 and 3000 atoms simulated for a total time of 20 ns in four eight-replicas batteries (according to the scaling protocol described in Ref [5, 10]) were computed on CRESCO3 (288 cores, 48 on the MPI layer and six on the OpenMP) in 48 and 13 wall clock hours, respectively, producing about 216/480 configurations sampled at 300 K and P=1 Atm every 10 ps. From these initial configurations of the bound ad bulk state, we started the FS-NE stage, by launching in parallel blocks of 48 NE ligand annihilation trajectories in the same thermodyanamic condition, each engaging 6 cores on the OpenMP layer for a total of 288 cores jobs on the CRESCO3 facility. The FS-NE stage on the whole collection of 216(complex) and 480 (bulk) starting states and using different annihilation times (from 90 ps to 720 ps), was completed in about two wall clock days.

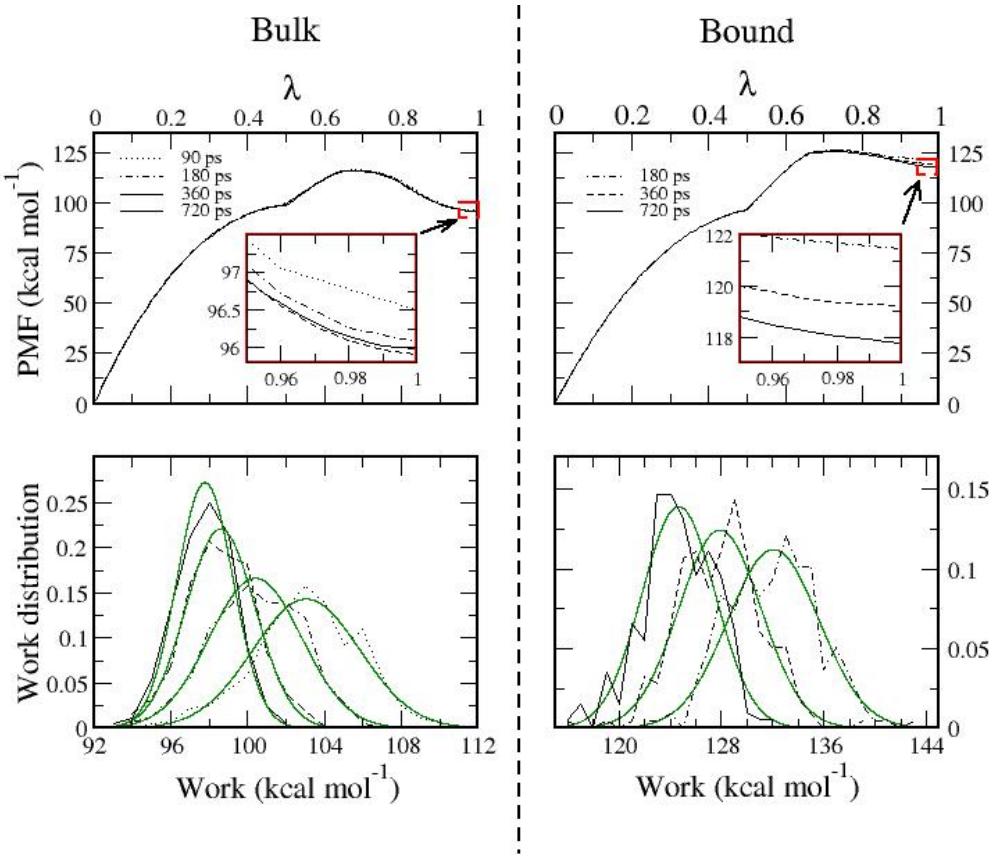

Figure 3: Upper panels: PMF for the annihilation of the ligand in bulk (left) and in the complex(right) obtained using different annihilation rates. The insets are a zoom of the PMF in the final stages $\lambda \simeq 1$ of the alchemical parameter λ Lower panel: Work distributions computed at various annihilation rate for the ligand in bulk (left) and in the complex (right). The corresponding Gaussian distributions, evaluated using the first two moments of the collection of 480 and 216 FS-NE works (unbound and bound state respectively), are shown in green color.

As discussed in Ref. [3, 4], the ligand annihilation free energy (in bulk and complex) along the alchemical path was computed using the first two moments (mean μ and variance σ) of the NE work distribution,

assumed to be normal, i.e.

$$\Delta G(\lambda) = \mu(\lambda) - \frac{\beta\sigma(\lambda)^2}{2} \qquad 0 \leq \lambda \leq 1. \tag{1}$$

where λ is the alchemical parameter with $\lambda = 0$ and $\lambda = 1$ indicating the fully coupled and fully decoupled ligand state, respectively. We stress that, in case of Gaussian distribution, Eq. 1 is a manifestation of the Crooks theorem[13] and as such is an *exact* equation.

In Figure 3 (bottom panels), we show the final annihilation work distributions evaluated at $\lambda = 1$ for the A6K decoupled ligand using various annihilation rate in bulk and in the complex. The corresponding superimposed normal distributions computed using the first two moments $\mu(1)$ and $\sigma(1)$ are also show in the Figure. In all cases, the Gaussian assumption appears to be fully justified. Further evidence comes from the calculation of higher moments of the distributions (kurthosis and skewness, data not shown) that are zero within statistical (bootstrap) error in all cases. In the same Figure 3 (upper panels), we show the annihilation free energy $\Delta G(\lambda)$ or Potential of Mean Force (PMF) for the unbound state (ligand in bulk) and for the bound state (ligand in the complex) for various annihilation times ranging from 90 to 720 ps (unbound state) and from 180 to 720 ps (bound state). Figure 3 represents a spectacular demonstration of the validity of the Crooks theorem, Eq. 1 and of the Gaussian assumption. As prescribed by the Crooks and Jarzynski theorems[13, 14], the invariance of the PMF to changes in the NE annihilation rate is observed at any λ within statistical (bootstrap) error for the annihilation of the ligand in bulk. For the annihilation of the ligand in the complex, the dissipation $W_{\rm diss} = \beta\sigma^2/2$, is significantly larger and a bias error[15] of $B = \beta\sigma^2/(2N_w)$ is observed for λ approaching to 1 (fully decoupled state). The λ annihilation protocol is, for all rates, such that for $0 \leq \lambda \leq 0.5$ the ligand-environment electrostatic interactions are switched off, while for $0.5 < \lambda \leq 1$, the ligand-environment Lennard-Jones interactions potential is brought to zero.

The differences between the two PMFs, for the the annihilation of the ligand in the complex and in bulk, provides important information on the nature of the binding in the 6AK-MCL1 system. Such difference is shown in Figure 4 (duration of the NE 720 ps for both bound and unbound systems) and shows that the electrostatic interactions slightly disfavors binding while a gain is observed in the dispersive repulsive (cavity) part. A similar trend was observed also in other cases.[4, 5]

Finally, the dissociation free energy of the 6AK ligand in the MCL1 system is given by the difference between the annihilation free energies at the alchemical end point $\lambda = 1$, i.e. $\Delta G_0 = \Delta G_b(1) - \Delta G_u(1)$. The actual value of the ΔG_0 must be corrected[3] for a standard state dependent volume term, $RT\ln(V_{\rm site}/V_0)$ with V_0 being the standard state volume, and possibly by a bias error.[15] The binding site volume, $V_{\rm site}$, can be estimated from the fluctuation of the center of mass distances between ligand and receptor in the REST equilibrium stage of the complex. In the Table 1, data are summarized.

Annihilation time (ps)	ΔG_b	ΔG_u	$RT\ln(V_{\rm site}/V_0)$	Bias[15]	ΔG_0
90	-	96.55 ± 0.54	-	-	-
180	121.34 ± 1.12	95.58 ± 0.44	-2.48 ± 0.1	-1.56	21.7 ± 1.6
360	119.19 ± 1.02	95.88 ± 0.23	-2.48 ± 0.1	-1.02	19.8 ± 1.3
720	117.66 ± 1.01	95.95 ± 0.16	-2.48 ± 0.1	-0.67	18.6 ± 1.2

Table 1: Annihilation free energies [ΔG_u (unbound ligand) and ΔG_b (bound ligand)], volume[3] and bias [15] corrections and standard dissociation free energies ΔG_0 for the 6AK-MCL1 complex. Reported errors have been evaluated by bootstrap analysis. All energy values are in kcal/mol. The experimental dissociation free energy for the 6AK-MCL1 complex[9] is below the theoretical limit in the given experimental condition[9] for competitive binding assay and is termed to be of "picomolar order" and is estimated to be $\Delta G_{\rm exp} \geq 13$ kcal/mol.

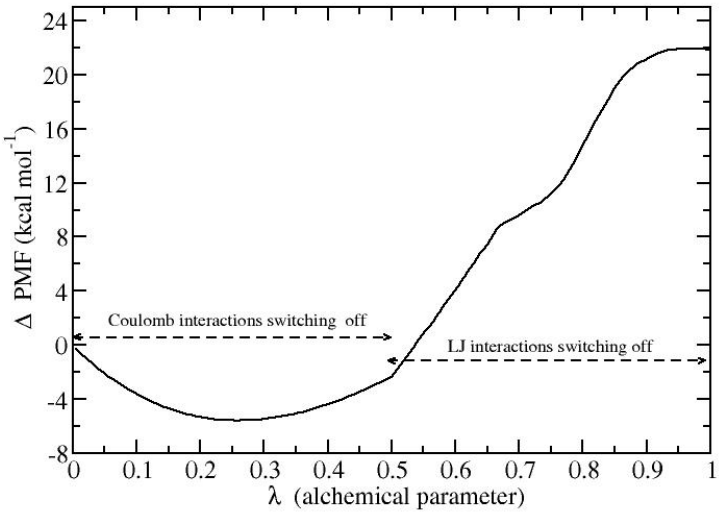

Figure 4: Difference between the PMF of the ligand in the complex and in bulk along the alchemical coordinate. The final value $\Delta\text{PMF}(\lambda = 1)$ represent the MCL1-6AK dissociation free energy (except for a standard state volume correction)

4 Conclusion

Using extensive tests on the CRESCO3 facility, we have shown that the ORAC molecular dynamics code exhibits essentially a linear scaling for thermodynamic parallelism (Replica Exchange and Fast Switching Non Equilibrium technologies) when implemented on the MPI layer. We have computed, using the FS-DAM technology, the dissociation constant of the MCL1-6AK constant, confirming the potency of the inhibitor versus the apoptosis regulator MCL1, with important implication in cancer therapy. A full paper on the MCL1 inhibition is in currently in preparation.

References

[1] Piero Procacci and Chiara Cardelli. Fast switching alchemical transformations in molecular dynamics simulations. *J. Chem. Theory Comput.*, 10:2813–2823, 2014.

[2] Robert B. Sandberg, Martina Banchelli, Carlo Guardiani, Stefano Menichetti, Gabriella Caminati, and Piero Procacci. Efficient nonequilibrium method for binding free energy calculations in molecular dynamics simulations. *J. Chem. Theory Comput.*, 11(2):423–435, 2015.

[3] Piero Procacci. I. dissociation free energies of drug-receptor systems via non-equilibrium alchemical simulations: a theoretical framework. *Phys. Chem. Chem. Phys.*, 18:14991–15004, 2016.

[4] Francesca Nerattini, Riccardo Chelli, and Piero Procacci. Ii. dissociation free energies in drug-receptor systems via nonequilibrium alchemical simulations: application to the fk506-related immunophilin ligands. *Phys. Chem. Chem. Phys.*, 18:15005–15018, 2016.

[5] Piero Procacci. Hybrid mpi/openmp implementation of the orac molecular dynamics program for

generalized ensemble and fast switching alchemical simulations. *J. Chem. Inf. Model.*, 56(6):1117–1121, 2016.

[6] Lingle Wang, Yujie Wu, Yuqing Deng, Byungchan Kim, Levi Pierce, Goran Krilov, Dmitry Lupyan, Shaughnessy Robinson, Markus K. Dahlgren, Jeremy Greenwood, Donna L. Romero, Craig Masse, Jennifer L. Knight, Thomas Steinbrecher, Thijs Beuming, Wolfgang Damm, Ed Harder, Woody Sherman, Mark Brewer, Ron Wester, Mark Murcko, Leah Frye, Ramy Farid, Teng Lin, David L. Mobley, William L. Jorgensen, Bruce J. Berne, Richard A. Friesner, and Robert Abel. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J. Am. Chem. Soc.*, 137(7):2695–2703, 2015.

[7] Levi N. Naden and Michael R. Shirts. Linear basis function approach to efficient alchemical free energy calculations. 2. inserting and deleting particles with coulombic interactions. *J. Chem. Theory Comput.*, 2015.

[8] Piero Procacci. Unbiased free energy estimates in fast nonequilibrium transformations using gaussian mixtures. *J. Chem. Phys.*, 142(15):154117, 2015.

[9] Taekyu Lee, Zhiguo Bian, Bin Zhao, Leah J. Hogdal, John L. Sensintaffar, Craig M. Goodwin, Johannes Belmar, Subrata Shaw, James C. Tarr, Nagarathanam Veerasamy, Shannon M. Matulis, Brian Koss, Melissa A. Fischer, Allison L. Arnold, DeMarco V. Camper, Carrie F. Browning, Olivia W. Rossanese, Amit Budhraja, Joseph Opferman, Lawrence H. Boise, Michael R. Savona, Anthony Letai, Edward T. Olejniczak, and Stephen W. Fesik. Discovery and biological characterization of potent myeloid cell leukemia-1 inhibitors. *FEBS Letters*, 591(1):240–251, 2017.

[10] Piero Procacci. *High Performance Computing on CRESCO infrastructure: research activities and results 2015*, chapter Fast switching alchemical simulations: a non equilibrium approach for drug discovery projects on parallel platforms. Giovanni Ponti, ENEA, DTE-ICT-HPC, CR Portici, 2016.

[11] Romelia Salomon-Ferrer, David A. Case, and Ross C. Walker. An overview of the amber biomolecular simulation package. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(2):198–210, 2013.

[12] W. L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79:926–935, 1983.

[13] G. E. Crooks. Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems. *J. Stat. Phys.*, 90:1481–1487, 1998.

[14] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, 1997.

[15] Jeff Gore, Felix Ritort, and Carlos Bustamante. Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements. *Proc. Natnl. Acad. Sci. USA*, 100(22):12564–12569, 2003.