Fast switching alchemical simulations: a non equilibrium approach for drug discovery projects on parallel platforms

Piero Procacci[1*]

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

Abstract. We present a novel all-atoms molecular dynamics approach for computing ligand-receptor binding free energies in drug discovery projects. The technique is a non equilibrium variant of the alchemical method[1] and effectively combines generalized ensemble (GE) and fast switching double annihilation (FS-DAM) nonequilibrium (NE) technologies. The method is implemented in the ORAC program[2, 3]. The code was equipped with a hybrid Open Multi-Processing (OpenMP), Message Passing Interface (MPI) multilevel parallelism specifically tailored for non uniform memory access (NUMA) multi-core architectures. A demonstrative application of the FS-DAM approach on HPC platforms is presented.

1 Introduction

In the last decade, the exponential growth of the clock rate has finally ended and performance improvements in parallel applications are now coming from the increase in the number of cores on a processor unit (compute node).[4, 5] Cores per chipset doubles every 18 months instead of clock and 64-512 threads per node will become visible soon.[4, 6] The advent of non uniform memory access (NUMA) multi-core architectures substantiates therefore the need for an efficient strategical approach in multi-level parallelization on High Performance Computing (HPC) facilities putting us at a fundamental turning point in the development of software for the Molecular Dynamics (MD) simulation of complex biomolecular systems.

In modern MD codes[7, 8, 9, 10] running on HPC platforms, a huge effort has been invested in the parallelization of the strong scaling non bonded force computations that is handled in a distributed memory environment at the MPI level via the so-called neutral territory domain decomposition (NT-DD) approach.[11] The MPI-only DD parallel algorithm guarantees, at constant processor workload, an equal level of domain-domain communications independently of the size of the system. Like in any strong scaling computation, the scaling behavior of the DD scheme is limited by the mean number of particles assigned to a domain/processor.[12] When the the latter decreases, the mean workload per processor also decreases and the communication overhead fatally increases. In typically sized drug-receptor systems (from 10000 to 30000 atoms), the saturation level of the parallel performances of DD based MD codes are reached within few hundreds or even few tens of cores, depending on network speed, *de facto* limiting the productivity of the simulations to at most few hundreds of ns per day.

On the other hand, many important biophysical phenomena, such as conformational transitions in proteins or binding of a ligand to a receptor, are *rare events*, occurring in the time scale of the microseconds

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

or even milliseconds, whereby the need for smarter and inherently (weak scaling) parallel methods of configurational sampling such as Hamiltonian Replica Exchange (H-REM), Serial Generalized Ensemble (SGE) or multi-trajectories non equilibrium (NE) techniques. The present release of the hybrid OpenMP/MPI ORAC code[2, 3] for molecular dynamics of biomolecular systems is designed to optimally exploit the recent trend of NUMA HPC systems. The code enforces a weak scaling parallelism for Generalized Ensemble (GE) (H-REM, SGE) and NE technologies on the MPI level and an OpenMP layer to the parallelization of the bonded and non bonded forces within a single shared memory access compute node. As such, ORAC provides an effective and flexible tool for reliably determining the binding free energies in drug-receptor systems using the non equilibrium alchemical method in an hybrid parallel environment. In the so-called fast switching double annihilation method (FS-DAM), the efficiency of the GE approach can be effectively exploited to generate and accurate sampling of the fully coupled states (the bound ligand and the ligand in bulk). From these initial states, a swarm of many independent NE processes are launched, where an externally driven alchemical parameter, regulating the interaction of the ligand with the environment, is varied continuously and rapidly during the multi-step integration of the equations of motion according to a prescribed common time schedule.[13, 14, 15, 16] The drug-receptor absolute binding free energy can be recovered from the annihilation work histogram of the bound and unbound states by applying an unbiased unidirectional free energy estimate. The latter is computed on the assumption that any observed work distribution is the result of a linear combination of normal distributions, whose components are identical in either direction of the non-equilibrium process, with weights regulated by the Crooks theorem.[17, 15, 16] We refer to Refs. [15, 16] for a detailed description of the FS-DAM algorithms in drug design. In this contribution, we present a demonstrative example of an FS-DAM application on the CRESCO HPC platform to a live drug discovery project using the ORAC6.0. The code, along with documentation, testing and ancillary tools, is distributed under the provisions of the General Public License (GPL) and can be freely downloaded at the Internet address `www.chim.unifi.it/orac`.

2 Code performances for the OpenMP layer on a CRESCO5 compute node

The ORAC code, mostly written in FORTRAN, is highly portable and it compiles successfully with the most common compilers (Intel `ifort`, IBM `xlf90`, PGI `pgfortran` and GNU `gfortran`) on many unix platforms, from laptop computers to HPC facilities including BlueGeneQ IBM architectures. As outlined in the introduction, ORAC implements a weak scaling parallel algorithm via MPI calls in H-REM/SGE generalized ensemble simulations or driven simulations technologies based on NE theorems[18, 19] and a strong scaling parallel approach on the OpenMP layer based on particle decomposition for the computation of the forces. The code does not require any library to run except for the standard mathematical libraries. ORAC can use the efficient `FFTW` suite[20] for Fourier Transform in the evaluation of the Particle Mesh Ewald (PME) contribution to reciprocal lattice forces, but has its own built-in parallel FFT libraries. The nature of the executable (serial, MPI only and MPI/OpenMP with or without `FFTW` libraries) is controlled by the specification in the options of the `configure` script provided in the distribution. Parallel execution can therefore be done on either MPI or OpenMP level or on the two combined levels generating the appropriate target executable.

In the Figure 1 (left) we show the OpenMP speedups obtained on a single 16-cores compute node of the ENEA-CRESCO5 cluster, for a system including a penta-alanine mini-protein, solvated in 9300 TIP3P[1] water molecules for a total of 28000 atoms in a cubic MD simulation box with a side-length of $\simeq 64$ Å. The size of such benchmark system, is that typical for alchemically driven TACE-ligands FS-DAM simulations (*vide infra*). The solvated mini-protein was modeled standardly using the AMBER all atom amber force field[22], a direct lattice long range cutoff of 10 Å, constant pressure simulation with isotropic stress tensor,[23] and temperature control via a Nosé Hoover thermostat.[24] The equations of

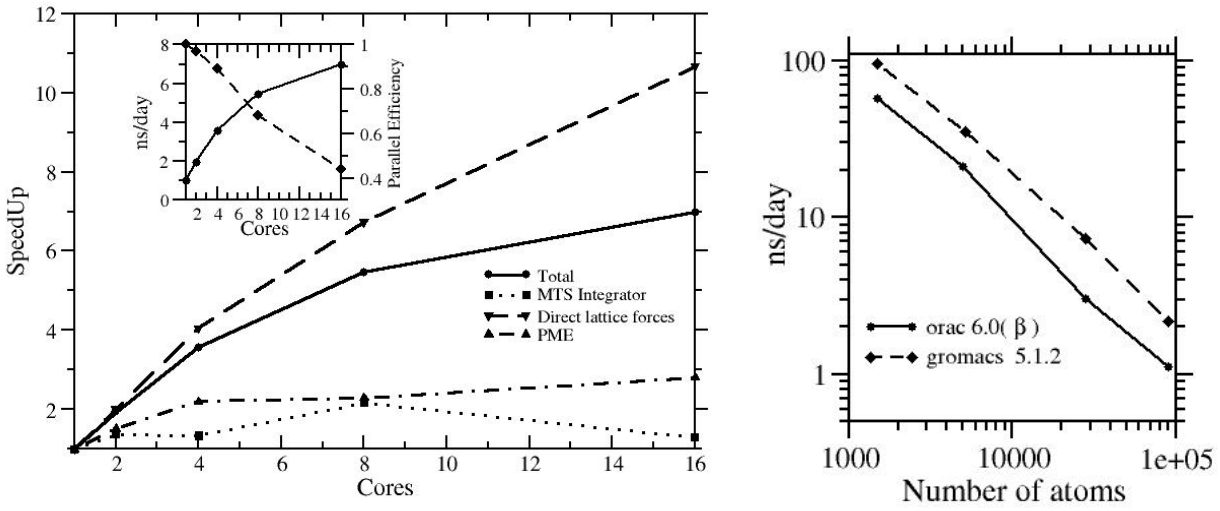

Figure 1: Left:Speedup measurements on the ENEA-CRESCO5[21] compute node (16-cores Intel(R) Xeon(R) CPU E5-2630 v3 2.40GHz) for a 28000 atoms systems (see text for details). Right: ORCA6.0 and gromacs5.1.2 performances for systems of various sizes on 8 cores on the CRESCO5 compute node (CPU E5-2630 v3).

motion were integrated using an efficient multiple time-step scheme specifically tuned for bio-molecular systems in the NPT ensemble.[23]. As the Intel `ifort` compiler generates optimized OpenMP code for the Intel `X_64` architectures, tests were run on a 16-cores genuine Intel CPU. The latter is the compute node of the CRESCO5 ENEA-grid facility and is made up of 2 sockets each mounting an 8-cores Intel(R) Xeon(R) CPU E5-2630 v3 with clock-rate of 2.40GHz and 64 GB/RAM. With eight processors (one socket), the overall parallel efficiency of the ORAC code is at about 70% with a nearly sixfold speedup factor. Performances are only marginally better when using all 16 cores on the node, with speed up factor at about 7 and with parallel efficiency dropping below 50%. Speedup factors are, however, disparate for the various type of calculations involved in the MD production. So, for example, whereas the scaling of the computation of the direct lattice non bonded forces is indeed satisfactory, with a parallel efficiency hitting 70% with 16 cores, the scaling of the out of place FFT using the parallel `FFTW` library is rather poor and in line with the recently reported benchmarks.[20] The computation labeled "MTS integrator" in Figure 1 includes constraints enforcing, fast bonded force calculations and asynchronous MTS velocity/coordinates advancement. All these combined tasks yield an overall contribution representing just 3% of the total MD workload for a serial job. However, due to the fine granularity, the MTS integration computation impacts on the overall parallel efficiency when increasing the number of processors. As stated previously, efficiency loss in a ORAC MPI/OpenMP job are exclusively attributable to the OpenMP strong scaling level, as the weak scaling H-REM/SGE or FS-DAM algorithms are implemented on the MPI layer with little of no communication overhead. Therefore, for a given amount of available core-hours, the assignment of too many cores on the OpenMP layer, while producing only moderate increase in the simulation speed, could limit the sampling efficiency in GE or FS-DAM simulations by reducing the number available MPI instances (each representing a GE-replica or a NE trajectory). On the overall, OpenMP speedups are satisfactory and efficiency loss is acceptable when running on a maximum of eight shared memory cores. In Figure 1(right), we compare the performances (measured in ns/day of simulation) of the OpenMP ORAC6.0β code to those obtained with the latest release of the popular GROMACS program. The latter MD code is admirably efficient due to algorithmic and processor-specific professional optimization, typically running 3-10 times faster

than many simulation programs.[8] Reported tests were done on 8-cores socket of the CRESCO5 CPU using solvated penta-alanine samples of various sizes. ORAC and GROMACS jobs were run using identical set up (constant pressure and temperature, bond constraints on X-H bonds, PME grid spacing of 1 Å, 10 Å cutoff). The log-log plot of Figure 1(c) shows that the performances of the two codes on a single 8-cores socket are comparable. For the largest system (90000 atoms), GROMACS runs at a speed of 2.2 ns/day compared to the 1.3 ns/day of ORAC6.0β. The efficiency gap between the two codes gets narrower with decreasing system size. For the smallest system (1500 atoms), GROMACS and ORAC produce about 95 and 60 ns of simulation in a day, respectively.

3 Dissociation free energy of TACE ligands via hybrid FS-DAM simulations on HPC platforms

In this section we provide a practical demonstration on how ORAC can be used to efficiently evaluate the binding free energy in drug-receptor systems on a HPC platform. To this end, we present some preliminary results obtained on the ENEA-CRESCO3 and Fermi BlueGeneQ HPC facilities using the recently developed FS-DAM technology for drug design on the Tumor necrosis factor converting enzyme (TACE)[25, 26, 27] in complex with the IK682 the tight binding inhibitor.[28] TACE is the membrane-anchored Zinc-protease that releases, from its membrane-bound precursor, the soluble form of the cytokine Tumor necrosis factor TNFa. The latter is a cell signaling protein involved in systemic inflammation and is one of the cytokines that make up the acute phase reaction. When produced in excess, TNFa is known to play a pivotal role in diseases provoking severe disabilities such a rheumatoid artritis, septic shock, Chron disease and Multiple Sclerosis. We should stress here that the results on the TACE-IK682 complex are presented here for demonstrative purposes. A thorough study on TACE inhibitors using FS-DAM will be the object of a forthcoming paper.

In Figure 2 we show a cartoon representation of the TACE-IK682 tight binding complex. The Zinc atom is the green sphere at the entrance of the binding site, bi-coordinated by the IK682 hydroxamate Zinc binding group (ZBG). The methylquinolin-methoxy-phenyl moiety of the inhibitor is deeply wedged in the hydrophobic environment of the so-called TACE "primed" pocket,[29] shown as surface representation. The TACE-IK682 presents various methodological challenges for the binding free energy determination via alchemical transformations. The TACE protein contains a doubly charged metal ion, interacting directly with the hydroxamate moiety of the ligand. Thus, the annihilation free energy of the ligand crucially depends on the modeling of the Zinc-hydroxamate electrostatic interactions. As a matter of fact, the Zn-ZBG mean electrostatic energy can be estimated to be[28] of the order of 300 kcal mol^{-1}. Hence, an error of few percent on this value may easily be transferred to the annihilation free energy of the binding ZBG, possibly producing discrepancies of several kcal mol^{-1} in the computed dissociation free energy.

The starting configuration of the bound state was obtained from the X-ray co-crystal structure of the TACE-IK682 complex (pdb code `2fv5`). Hydrogen atoms were automatically added by the code and the resulting all atoms experimental structure underwent a preliminary conjugate gradient minimization to adjust bond lengths and bending angles to the AMBER99SB-ILDN force field[31] corresponding equilibrium distances and angles. The complex was then arranged in an orthorhombic box in such a way that the shortest distance between any protein atom and the box walls was larger that 10 Å. The box was filled with TIP3P water at the density of 1g/cm^{-3}, discarding the solvent molecules that were found to overlap with the complex. The final simulation system consisted of a TACE-IK682 complex made of 4089 atoms, solvated in 6335 TIP3P water molecules for a total of 23094 atoms. On a local cluster, a preliminary equilibration in the NPT ensemble was carried on for 100 ps obtaining an equilibrated volume of about 225$\pm$1 nm^3. The starting configurations of the bound state for the subsequent FS-DAM annihilation were collected at regular time interval of 9 ps from a 5 ns 24 replicas H-REM simulation in the NPT ensemble with torsional tempering of the ligand and of the binding region. See Refs. [3, 16]

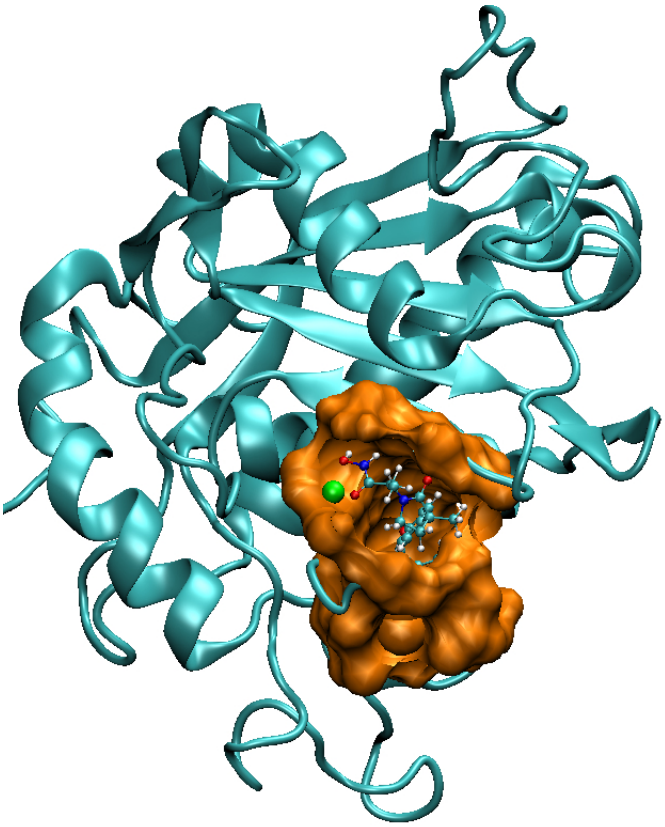

Figure 2: Tumor necrosis factor α converting enzyme in complex with the IK682 tight binding inhibitor. The model structure was built from the experimental structure[28] using the VMD software[30].

for further details on the H-REM torsional scaling scheme. The computation was done on 192 cores of the CRESCO3 cluster using a mixed OpenMP/MPI ORAC6.0 executable with eight OpenMP threads per MPI instance in less than three wall clock days. In order to collect the starting configurations of the unbound ligand, a 5 ns NPT H-REM simulation of torsionally tempered IK682 in bulk water using a cubic MD box of 30 Å side-length was performed in few wall-clock hours on 64 nodes of CRESCO3 using again an hybrid executable with 16 H-REM replicas and 6 OpenMP thread per MPI instance. Once collected, the final 512 starting configurations of the bound and unbound states were transferred to the Fermi/CINECA system. Fermi is made up of about 10,000 compute nodes, each featuring an IBM PowerA2 chip with 16 cores working at a frequency of 1.6 GHz, 16 GB of RAM. Two parallel OpenMP/MPI jobs on 2048 cores were finally submitted, producing the 512 FS-DAM annihilation 0.3 ns trajectories (each engaging 4 OpenMP threads) of the bound (23094 atoms) and unbound (2969 atoms) states, for a total combined 0.3 μs of simulation. Both jobs ended in a single wall clock day.

In Figure 3, we show the time record of the alchemical work in a subset of the FS-DAM annihilation trajectories of the IK682 ligand in the bound state and in bulk solvent along with the corresponding distribution of the final alchemical work. The nature of the work distribution was checked by computing higher moments of the distributions (skewness and kurtosis) and by using the standard Kolmogorov test as described in Refs. [14]. Both work distributions turn out be the superposition of more than one components. In particular, the work distribution referring to the bound state exhibits a significant left skewness, while that relative to the annihilation of the ligand in bulk showed a large kurtosis. By repre-

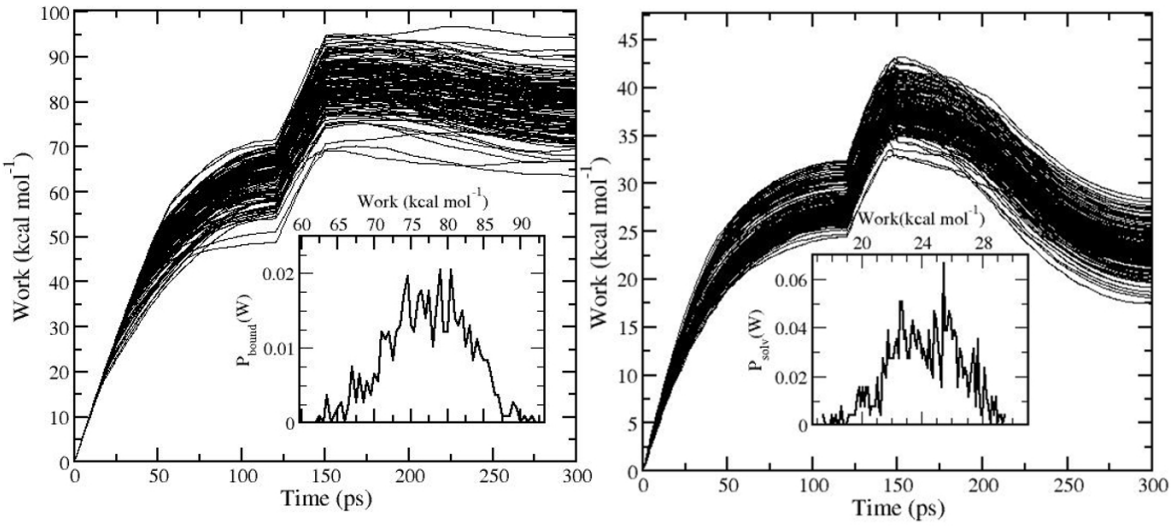

Figure 3: Time record of the annihilation work for a subset of the 512 FS-DAM trajectories of the bound state (left) and for the unbound state (right) for the IK682-TACE ligand-receptor pair. In the inset, the relative work distribution is reported.

senting the overall FS-DAM work distributions with an appropriate combination of normal components, the bound and unbound annihilation free energies (ΔG_b, ΔG_s,respectively) can be straightforwardly computed as described in Ref. [17], finding $\Delta G_b = 43.3 \pm 0.9$ kcal mol^{-1} and $\Delta G_s = 21.6 \pm 0.3$ kcal mol^{-1}. The errors on these values were computed by block-bootstrapping 20 random subsets containing 128 work samples. The dissociation free energy is given by the difference between these two values plus a standard state correction given by $k_B T \ln \frac{V_{\rm site}}{V_0}$, where $V_{\rm site}$ and V_0 are the binding site and standard state volumes. If we assume[27] that substrate/ligand binding occurs via the Zinc(II) cation, $V_{\rm site}$ can be taken as the influence volume of the Zinc(II) alone. The latter volume was computed by means of a Voronoi polihedra analysis[32] on the starting configurations of the complex, finding $V_{\rm site} = 4.9 \pm 0.7$ Å^3 kcal mol^{-1} and hence yielding a standard state correction of -3.5 ± 0.7 Å^3 kcal mol^{-1}. Based on these results the standard dissociation free energy of the TACE-IK682 complex can be estimated as $\Delta G_0 = 18.2 \pm 1.4$ kcal mol^{-1}.

Experimental measurements of the IK682 binding affinity *vs* TACE are very difficult due to the extreme strength of the binding. In any case, IK682 was found to exhibit picomolar activity[28] with an apparent K_i of the order 0.02 nm, yielding a standard dissociation free energy of $\simeq 15$ kcal mol^{-1} at pH=7.3. As previously stated, the TACE-IK682 is indeed challenging from a computational standpoint with the binding free energy crucially depending on many details (pH modeling, force fields, work distribution resolution). Nonetheless, in spite of the preliminary nature of this demonstrative Open-MP/MPI ORAC application, FS-DAM was able to correctly predict IK682 as a tight binding TACE inhibitor, with a semi-quantitative agreement between computed ($\Delta G_0^{(\rm FS-DAM)} = 18.2 \pm 1.4$ kcal mol^{-1}) and experimental ($\Delta G_0^{(\rm Exp.)} \simeq 15.0$ kcal mol^{-1}) standard dissociation free energies.

4 Conclusions and perspectives

In this paper we have presented an OpenMP/MPI implementation of the ORAC code[2, 3] for multi-core HPC platforms. The hybrid parallelization is designed in such a way that the distributed memory MPI level is reserved for thermodynamic parallelism only (H-REM, GE or FS-DAM low communication computations) while the strong scaling force decomposition algorithm is implemented in the OpenMP layer within the intra-node multi-cores shared memory environment. Such design is specifically tuned so as to maximize the sampling efficiency in complex heterogeneous systems that are characterized by a manifold of competing free energy attractors. Hence the need for an effective exploitation, in MD codes of biological system, of the so-called non Boltzmann methodologies that artificially enhance the probability of rare fluctuations on auxiliary thermodynamic states by rigorously re-weighting the statistics onto the target thermodynamic state. In some sense, H-REM and GE methodologies, as well as FS-DAM techniques are all aimed at expanding the statistics (i.e. the number of statistically *independent* simulated proteins) rather than the time scale of the system. These latter techniques are especially advantageous in a massively parallel environment where one can efficiently exploit the limited communication overhead that is needed for their implementation, easily producing, in a single wall clock day, an overall simulation time in the order of the microseconds. The OpenMP/MPI ORAC MD code is based on this simulation paradigm and is perfectly tailored for the current architectural trends in super-computing, limiting the strong scaling force computation to the intra-node shared memory environment with a moderate subtraction of resources to be assigned to the thermodynamic parallelism for enhanced sampling. In the Table below, based on tests done on the the 8-cores E5-2670 Intel CPU, we report the projected productivity on the full dedicated CRESCO4 cluster (4864 cores) in terms of processed ligand per day for a typical drug discovery project such as the TACE system. For each ligand-receptor pair, a 5 ns 16 H-REM simulation is assumed for the generation of the initial states, followed by the production of 256 0.3 ns FS-DAM trajectories for the computation of the work histograms. The gain in

OpenMP speed (ns/day)	N_{MPI}/socket	Ligands/day (overall)
1.9	4	16
3.0	2	20
4.9	1	25

Table 1: ORAC6.0β predicted throughput for a drug discovery project on the dedicated CRESCO4 cluster (4864 cores) using FS-DAM. Productivity is measured in the number of processed ligand-receptor pairs per wall-clock day based on full cluster occupancy

productivity as the number of MPI instance per socket increases, reflects the corresponding OpenMP efficiency loss as the the number of threads per MPI instance increases. A typical 100 ligand-receptor pairs drug discovery project could hence be completed in 4 wall clock days on the CRESCO4 cluster.

References

[1] W.L. Jorgensen and C. Ravimohan. Monte carlo simulation of differences in free energies of hydration. *J. Chem. Phys.*, 83:3050–3054, 1985.

[2] P. Procacci, E. Paci, T. Darden, and M. Marchi. Orac: A molecular dynamics program to simulate complex molecular systems with realistic electrostatic interactions. *J. Comput. Chem.*, 18:1848–1862, 1997.

[3] S. Marsili, G. F. Signorini, R. Chelli, M. Marchi, and P. Procacci. Orac: A molecular dynamics simulation program to explore free energy surfaces in biomolecular systems at the atomistic level. *J. Comput. Chem.*, 31:1106–1116, 2010.

[4] Herb Sutter. The free lunch is over: A fundamental turn toward concurrency in software. *Dr. Dobb's Journal (DDJ)*, 3:30, 2005.

[5] Christian Märtin. Multicore processors: Challenges, opportunities, emerging trends,. In *Proceedings Embedded World Conference 2014, 25-27 February, 2014, Nuremberg, Germany, Design & Elektronik*, pages 25–27, 25-27 February 2014.

[6] Giovanni Erbacci. Trends in HPC Architectures and Parallel Programmming, PRACE Winter school 2012: http://www.training.prace-ri.eu/uploads/tx_pracetmo/ParallelArchitectures-Erbacci.pdf (accessed 22 January 2016).

[7] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, 4(3):435–447, 2008.

[8] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilrd Pll, Jeremy C. Smith, Berk Hess, and Erik Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19–25, 2015.

[9] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, L. Skeel, and K. Schulten. Scalable molecular dynamics with namd. *J. Comput. Chem.*, 26:1781–1802, 2005.

[10] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.*, 117(1):1 – 19, 1995.

[11] David E. Shaw. A fast, scalable method for the parallel evaluation of distance-limited pairwise particle interactions. *J. Comput. Chem.*, 26(13):1318–1328, 2005.

[12] Kevin J Bowers, Ron O Dror, and David E Shaw. Overview of neutral territory methods for the parallel evaluation of pairwise particle interactions. *Journal of Physics: Conference Series*, 16(1):300, 2005.

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

[14] 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.

[15] 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.

[16] 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.

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

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

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

[20] Matteo Frigo and Steven G. Johnson. The design and implementation of FFTW3. *Proceedings of the IEEE*, 93(2):216–231, 2005. Special issue on "Program Generation, Optimization, and Platform Adaptation".

[21] G. Ponti, F. Palombi, D. Abate, F. Ambrosino, G. Aprea, T. Bastianelli, F. Beone, R. Bertini, G. Bracco, M. Caporicci, B. Calosso, M. Chinnici, A. Colavincenzo, A. Cucurullo, P. Dangelo, M. De Rosa, P. De Michele, A. Funel, G. Furini, D. Giammattei, S. Giusepponi, R. Guadagni, G. Guarnieri, A. Italiano, S. Magagnino, A. Mariano, G. Mencuccini, C. Mercuri, S. Migliori, P. Ornelli, S. Pecoraro, A. Perozziello, S. Pierattini, S. Podda, F. Poggi, A. Quintiliani, A. Rocchi, C. Scio, F. Simoni, and A. Vita. The role of medium size facilities in the hpc ecosystem: the case of the new cresco4 cluster integrated in the eneagrid infrastructure. In *Proceeding of the International Conference on High Performance Computing & Simulation*, pages 1030–1033. Institute of Electrical and Electronics Engineers (IEEE), 2014.

[22] Viktor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3):712–725, 2006.

[23] M. Marchi and P. Procacci. Coordinates scaling and multiple time step algorithms for simulation of solvated proteins in the npt ensemble. *J. Chem. Phys.*, 109:5194–5202, 1998.

[24] William G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, Mar 1985.

[25] G. Kassiotis and G. Kollias. Uncoupling the proinflammatory from the immunosuppressive properties of tumor necrosis factor (tnf) at the p55 tnf receptor level: Implications for pathogenesis and therapy of autoimmune demyelination. *J. Exp. Med.*, 193:427–434, 2001.

[26] C. Dai, D. Li, J. Popovici-Muller, L. Zhao, V.M. Girijavallabhan, and et al. 2-(2-aminothiazol-4-yl)pyrrolidine-based tartrate diamides as potent, selective and orally bioavailable tace inhibitors. *Bioorg. Med. Chem. Lett.*, 21:3172–3176, 2011.

[27] Carlo Guardiani and Piero Procacci. The conformational landscape of tartrate-based inhibitors of the tace enzyme as revealed by hamiltonian replica exchange simulation. *Phys. Chem. Chem. Phys.*, 15:9186–9196, 2013.

[28] X. Niu, S. Umland, R. Ingram, B.M. Beyer, Y. H. Liu, J. Sun, D. Lundell, and P. Orth. Ik682, a tight binding inhibitor of tace. *Arch. Biochem. Biophys.*, 451:43–50, 2006.

[29] Martina Banchelli, Carlo Guardiani, Eleonora Tenori, Stefano Menichetti, Gabriella Caminati, and Piero Procacci. Chemicalphysical analysis of a tartrate model compound for tace inhibition. *Phys. Chem. Chem. Phys.*, 15:18881–18893, 2013.

[30] W. Humphrey, A. Dalke, and K. Schulten. Vmd - visual molecular dynamics. *J. Mol. Graphics*, 14:33–38, 1996.

[31] Kresten Lindorff-Larsen, Stefano Piana, Kim Palmo, Paul Maragakis, John L. Klepeis, Ron O. Dror, and David E. Shaw. Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins*, 78(8):1950–1958, 2010.

[32] Piero Procacci and Riccardo Scateni. A general algorithm for computing voronoi volumes: Application to the hydrated crystal of myoglobin. *Int. J. Quantum Chem.*, 42(5):1515–1528, 1992.