Insights into the mechanism of oncogenesis of the mutant protein PI3Kα from Molecular Dynamics simulations

Proposal ID: 2010PA0456

Paraskevi Gkeka and Zoe Cournia

Biomedical Research Foundation of the Academy of Athens,
4 Soranou Ephessiou, 11527 Athens, Greece

Abstract

Cancer is a leading cause of death worldwide, accounting for 7.6 million deaths in 2008 according to the World Health Organization. This number is expected to increase to 21 million by 2030. One of the signaling pathways which, when deregulated, becomes centrally involved in several types of cancers, like colon, mammary, and endometrial tumorigenesis, is served by phosphoinositide-3-kinase alpha (PI3Kα). The importance of PI3Kα in cancer is highlighted by the discovery that PIK3CA, the gene encoding the catalytic p110α subunit of PI3Kα, is frequently mutated in human malignancies. The goal of our project is to gain insights into the oncogenic mechanism of two commonly-expressed PI3Kα mutants by studying their conformational changes with Molecular Dynamics simulations, in comparison with the PI3Kα wild-type (normal, non-cancerous) protein. We performed Molecular Dynamics simulations using NAMD 2.8 and Gromacs 4.0 packages on CURIE and JUGENE. Our results show significant scaling up to 2,048 cores on CURIE and JUGENE for both NAMD and GROMACS. However, NAMD and GROMACS performance was significantly smaller on JUGENE. On CURIE the nodes are connected via an Infiniband QDR network, while on JUGENE the nodes are connected via a 10-Gbit Ethernet network. As NAMD requires frequent and high throughput communication between nodes, it performed better on CURIE than on JUGENE. GROMACS performed compared to NAMD. However, also GROMACS performance was reduced on JUGENE though to a smaller extent, probably due to the use of Ethernet interconnects on JUGENE versus infiniband on CURIE.

1 Background and significance

Every year 3.2 million Europeans are diagnosed with cancer, the second most common cause of death in Europe. Colorectal cancers account for about 13% of cancer deaths in both men and women. Breast cancer is the leading cause of female cancer deaths with a 30% incidence. The PIK3CA oncogene, which encodes the catalytic subunit of PI3Ka, is one of the two most frequently mutated oncogenes in human cancers and in particular in breast and colon cancers. Mutated proteins such as PI3Ka can drive tumor progression due to conformational changes that result in upregulated enzymatic activity. Understanding how these mutations drive tumorigenesis is central to developing new therapeutics for cancer.

Thus, the goal of this proposal is to gain insights into the oncogenic mechanism of two commonly-expressed PI3Kα mutants by studying their conformational changes with Molecular Dynamics simulations, in comparison with the PI3Ka wild-type (normal, non-cancerous) protein. Differences between these proteins will be identified by means of structural, thermodynamic, and kinetic properties derived from the simulations. Understanding the dynamic nature of the structural changes that occur in the mutated PI3Kα is of paramount importance, as such studies can be used to aid the design of novel, mutant-specific candidate drugs. Designing mutant-specific candidate drugs for PI3Kα is of striking

therapeutic potential as this strategy attempts to kill cancer cells selectively, while minimizing harmful side effects. Structure-based drug design is currently employed in our laboratory to develop inhibitors that act specifically for the H1047R PI3Ka mutant known to be present in a variety of cancers. Our drug design studies will be tremendously aided by understanding the dynamics of the mutant PI3Ka proteins. Moreover, the scientific community will benefit from understanding how these mutations lead to the increased PI3Kα activity and result to carcinogenesis.

2 Materials and methods

In this study, we performed Molecular Dynamics simulations using NAMD 2.8 and Gromacs 4.0 packages on CURIE and JUGENE. The CHARMM22 force field (includes phi, psi cross term map (CMAP) correction) was used to model all protein interactions, and the TIP3P model was used for water. The simulations were performed at constant pressure, temperature, and number of particles (NPT ensemble). The temperature was kept constant at 310 K, using the velocity rescaling with a relaxation time of 0.1 ps. The pressure of the system was isotropically coupled and maintained at 1 bar using the Berendsen algorithm with a time constant of 0.1 ps and a compressibility of $4.5x10^(-5)bar^(-1)$. The nonbonded potential energy functions were cut off and shifted at 14 Å, with forces smoothly decaying between 10 and 14 Å. The particle-mesh Ewald method (PME) was employed to calculate long-range electrostatic interactions with a grid spacing of 0.12 nm.The simulations were run using a 2-fs integration time step and a total number of 10,000 timesteps, which has been shown to be adequate for benchmark studies. The output coordinates, velocities and energies were saved every 100 steps. The system contained $\tilde{4}$00,000 atoms.

3 Benchmarks and scaling

Simulations were performed in 32, 64, 128, 256, 512, 1024, 2048 cores for NAMD and 32, 64, 128, 256, 512, 1024, 2048, 4096, 8192 cores for GROMACS. The results from our study are presented in the next section. Output/input files, etc, are available upon request.

Our results show significant scaling up to 2,048 cores on CURIE and JUGENE for both NAMD and GROMACS. However, NAMD and GROMACS performance was significantly smaller on JUGENE. On CURIE the nodes are connected via an Infiniband QDR network, while on JUGENE the nodes are connected via a 10-Gbit Ethernet network. As NAMD requires frequent and high throughput communication between nodes, it performed better on CURIE than on JUGENE. GROMACS performed compared to NAMD. However, also GROMACS performance was reduced on JUGENE though to a smaller extent, probably due to the use of Ethernet interconnects on JUGENE versus infiniband on CURIE.

#·cores	absolute·timing·(s)	speedup
256·	823·	1.000
512·	447	1.841
1024·	277·	2.971·
2048·	241·	3.415·

Figure 1: Scaling for NAMD on CURIE.

	sec/step			
	CURIE		JUGENE	
cores	NAMD	GROMACS	NAMD	GROMACS
32	0.46	0.096	0.411	0.910
64	0.26	0.05	0.405	0.440
128	0.133	0.027	0.406	0.400
256	0.072	0.014	0.215	0.270
512	0.041	0.008	0.115	0.190
1024	0.023	0.134	0.069	0.140
2048	0.017	0.004	0.047	0.020
4096	-	0.0047	0.045	0.016
8192	-	0.0038	0.042	0.020

Figure 2: Results from the benchmark analysis performed on the CURIE and JUGENE cluster.

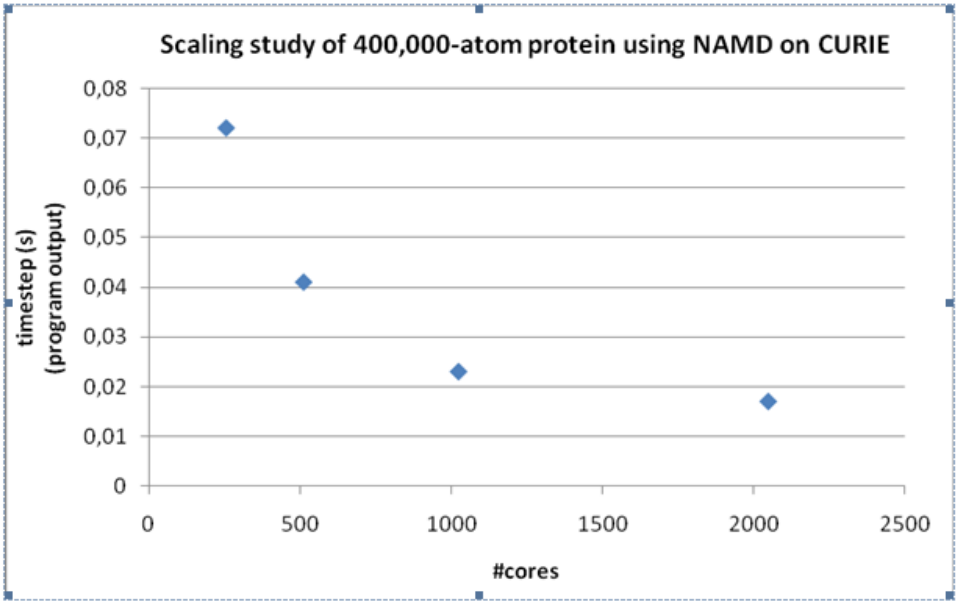

Figure 3: Timestep of the simulation in seconds, plotted against number of cores for the PI3Ka mutant protein.

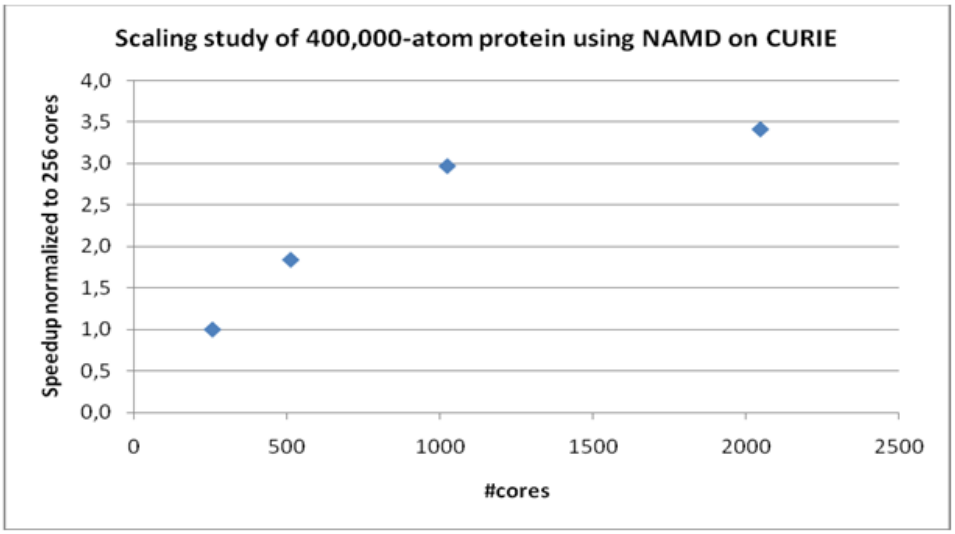

Figure 4: Speedup plotted against number of cores for the PI3Ka mutant protein over 10000 timesteps.

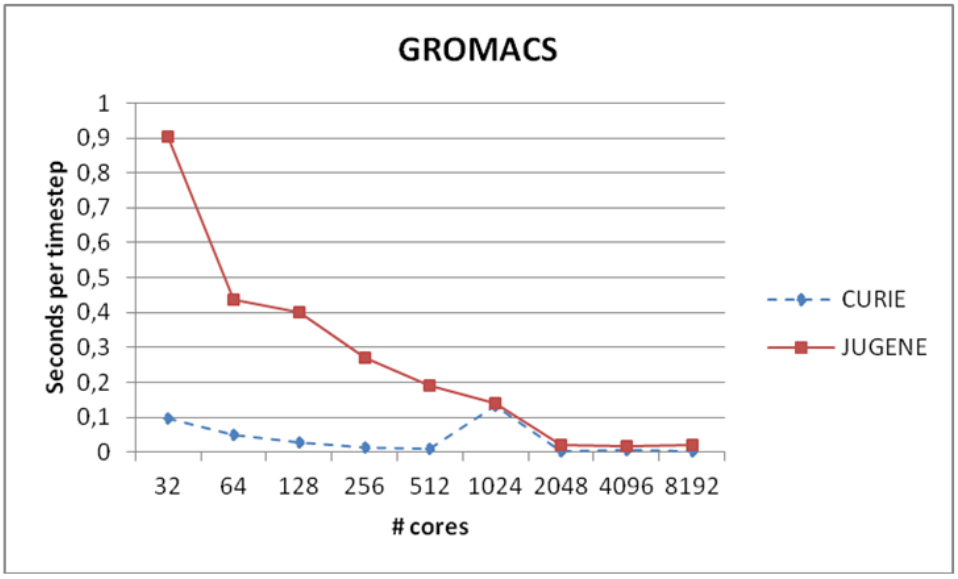

Figure 5: GROMACS performance on CURIE and JUGENE.

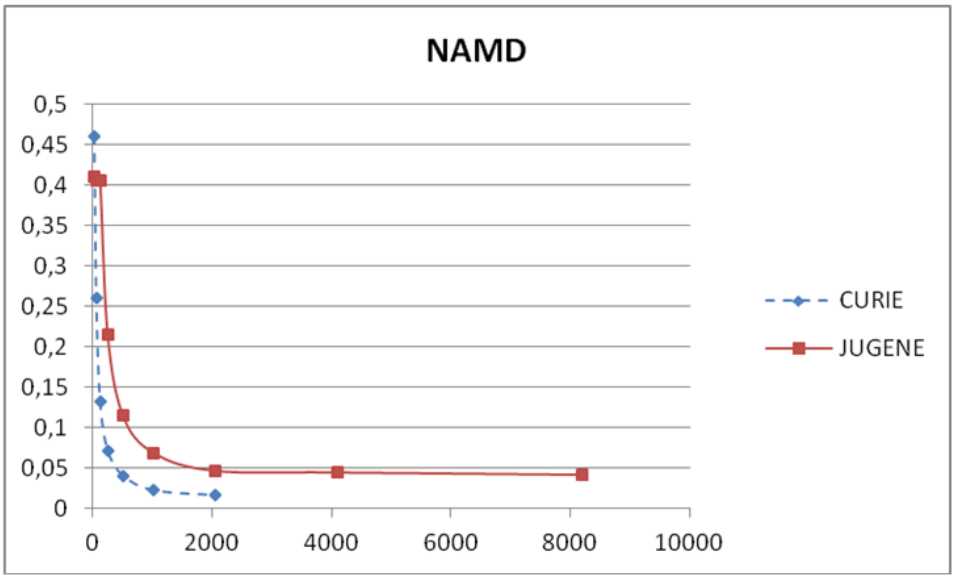

Figure 6: NAMD performance in CURIE and JUGENE.

Number·of· cores	Wall·clock·time	Speed-up·vs·the· first·one	Number·of· Nodes	Number·of· process
256·	823·	1.000	64	256
512·	447	1.841	128	512
1024·	277·	2.971·	256	1024
2048·	241·	3.415·	512	2048

Figure 7: NAMD on CURIE.

Number·of· cores	Wall·clock·time	Speed-up·vs·the· first·one	Number·of· Nodes	Number·of· process
256·	142	1.000	64	256
512·	79	1.798	128	512
1024·	1344*	--	256	1024
2048·	40	·3.550	512	2048
4096·	46	·3.087	1024	4096
8192·	38	·3.737	2048	8192

Figure 8: GROMACS on CURIE. [1]

Number·of· cores	Wall·clock·time	Speed-up·vs·the· first·one	Number·of· Nodes	Number·of· process
256·	·1723	1.000	64	256
512·	1151	1.497	128	512
1024·	·692	·2.500	256	1024
2048·	·467	·3.689	512	2048

Figure 9: NAMD on JUGENE.

Number·of· cores	Wall·clock·time	Speed-up·vs·the· first·one	Number·of· Nodes	Number·of· process
256·	2160	1.000	64	256
512·	1920	1.125	128	512
1024·	1260	·1.714	256	1024
2048·	196	·11.02	512	2048
4096·	155	·13.93	512	2048
8192·	203	·10.64	512	2048

Figure 10: GROMACS on JUGENE.

4 Conclusions

Our results show significant scaling up to 2,048 cores on CURIE and JUGENE for both NAMD and GROMACS. However, NAMD and GROMACS performance was significantly smaller on JUGENE. On CURIE the nodes are connected via an Infiniband QDR network, while on JUGENE the nodes are connected via a 10-Gbit Ethernet network. As NAMD requires frequent and high throughput communication between nodes, it performed better on CURIE than on JUGENE. GROMACS performed compared to NAMD. However, also GROMACS performance was reduced on JUGENE though to a smaller extent, probably due to the use of Ethernet interconnects on JUGENE versus infiniband on CURIE.

We were delighted from the use of both CURIE and JUGENE resources. For our purposes it seems that CURIE is the optimal resource. Our system size and programs scale up to 2,048 cores. With this scaling behavior we could achieve 10ns/day for our system.