Enclosure 1. Tier-1 Application form – English version

APPLICATIONS ARE PREFERABLY DRAWN UP IN ENGLISH. AN ENGLISH TRANSLATION HAS TO BE ENCLOSED WITH APPLICATIONS SUBMITTED IN DUTCH.

The application form is available in English on the website

https://www.vscentrum.be/en/access-and-infrastructure/project-access-tier1

Title of the application:

Simulating oxygen transport through membranes at various temperatures

Name and first name of the applicant: De Vos Oriana

Institution: Ghent University

Research group / department: Center for Molecular Modeling

Title / position: Assistant

e-mail address: Oriana.DeVos@UGent.be

Total computing time that is needed, in node days: 660 nodedays

Total disk storage that is applied for (in GiB):

3 GiB on SCRATCH, 500 GiB on Tier-2 storage

The total number of pages in this application should not exceed 17, excluding possible appendices (confirmation letter of financing institution, software license,…) which may be taken into account by the Tier-1 Allocation Board.

1. Title of the research project (with IWETO or FRIS link if available) within the framework of which computing time is applied for:

Simulating oxygen transport through membranes at various temperatures

2. Describe your research project in short. Explicitly mention the scientific questions that you are planning to address and the overall scientific goals of the project. (max. 1 A4 in Arial 12):

Membranes are found in every cell. Oxygen must be transported through membranes in order to supply energy to the cell. The central question of this project is “how can oxygen diffuse through the membrane?” This project aims to investigate whether the presence of ordered rafts has an influence on oxygen transport. Specifically, the effect of temperature will be studied. Simulated data of oxygen diffusion is available at body temperature (310 K), whereas the experimental data on rafts is available at room temperature (298 K). This project aims at generating reliable reference simulations of ordered and disordered membranes at room temperature. This project will not only allow for a comparison between experiment and simulation, but will also give us insight in the effect of temperature on oxygen diffusion.

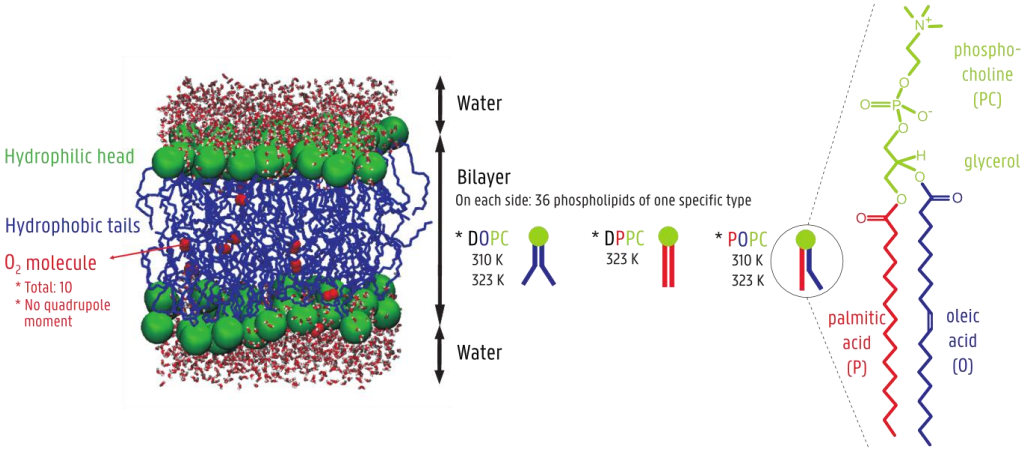


Figure 1. Oxygen molecules need to diffuse into and through phospholipid membranes in order to supply energy to the cell. Phospholipids can have various head groups and various tails.

Membranes consist of phospholipid bilayers. Phospholipids are composed of a hydrophilic head and two hydrophobic tails. The hydrophobic tails are facing each other in the membrane bilayer. Membranes can contain ordered domains rich in cholesterol, sphingomyelin (a group of phospholipids) and saturated phospholipids. These nanoscale domains, called rafts, are surrounded by disordered domains. The rafts function together with membrane proteins, both integral and peripheral proteins. There are indications that rafts play a role in signal transduction and intracellular transport. Cholesterol is known to have an influence on the flexibility of the cell membrane.

The long term goal is to investigate whether rafts have an influence on the oxygen transport. We expect that the transport might be easier in rafts than in disordered domains. In rafts the pathway of oxygen would be (almost) straight, but in disordered domains the pathway would be longer and would contain more curves. Oxygen mobility will be examined using computational modeling, since wet lab experiments cannot access the atomic scale information.

To study the oxygen diffusion in membranes, we will run molecular dynamics simulations of membranes surrounded by water. Three different phospholipid types that have been considered in previous simulations: DOPC, DPPC, and POPC. They have the same hydrophilic head, but they differ in the hydrophobic tails. DOPC contains two oleic acids, DPPC two palmitic acids and POPC an oleic and a palmitic acid. Oleic acid tails are unsaturated with a bent in the tail and are a model for disordered membranes. In contrast, palmitic acid tails are fully saturated and are a model for ordered membranes, such as the rafts. This project will focus on creating new reference simulations for these lipid types at a temperature of 298 K. The DPPC membrane will not be studied, as it is known to become a gel at room temperature. This will allow us to assess whether order or disorder in the membrane enhances the diffusion. In the future projects, this information may be linked to the study of rafts.

3. Provide an engaging abstract (10 lines) for scientific communication on the website in layman's terms. Should this application be bound by a confidentiality agreement (see also item 12 of this application form), provide more details about the specific nature of the confidentiality and indicate why an abstract may not be published.

Membranes are found in every cell and organelle. Oxygen must be transported through the membrane in order to supply energy to that cell/organelle. But how does oxygen diffuse through the membrane? Does the presence of ordered rafts in membranes influence this process? This project aims to investigate these questions through simulations of three model membranes with varying degree of lipid saturation, and hence varying degree of ordering. The effect of temperature will be investigated by comparing oxygen diffusion at room temperature and body temperature, which will help towards the understanding of oxygen transport in ordered rafts.

4. Financing institution or channel, financing the research project in full or in part (FWO, BOF, IWT, EU, ...): Please attach the confirmation letter as enclosure. In case the project has not gone through a scientific approval process attach a letter of approval of your own institute.

This round of TIER1-proposals is without charge, so we have not budgeted financing for this project.

The project is part of the PhD thesis of Oriana De Vos, who is employed as an assistant at Ghent University. Her PhD is focused on extracting dynamical information (diffusivities) from molecular simulations on biological systems, and more specifically on the permeation of biological membranes.

A letter can be found in attachment with the approval by Ghent University.

5. Name and email address of the promoter(s) of the research project:
   An Ghysels, an.ghysels@ugent.be

6. Persons mandated by the Applicant to compute on the Tier-1 within the framework of the present project: Please provide for every person:
   - name and first name
   - institution
   - research group / department
   - title / position
   - experience of using HPC resources in the past (Tier-0/Tier-1/Tier-2 infrastructure in Belgium and abroad)

   Oriana De Vos (vsc41441)
   Center for Molecular Modeling, Ghent University
   Assistant
   - Has started using the HPC infrastructure since April 2015 (TIER2).
   - Has developed experience with the program CHARMM after a successful application of a TIER1 Starting Grant (Oct 2015).
   - Has experience on VSC (UGent) with both TIER1 and TIER2
   - Has received a grant in Oct 2016 on the newest TIER1 infrastructure, a.o. to benchmark the scaling of CHARMM.
   - Has received several TIER1 grants for her work on oxygen diffusion.

   Prof. An Ghysels (vsc40051)
   Center for Molecular Modeling, Ghent University
   Has experience with a variety of computational codes, such as CHARMM in collaboration with foreign groups.
   Has experience since 2007 using HPC infrastructure.
   - on Biowulf (National Institutes of Health (NIH, USA), global cluster),
   - on LoBoS (Lots of Boxes on Shelves, Laboratory of Computational Biology, NIH, USA)
   - on VSC (UGent): early user since gengar/TIER2/TIER1

7. Explain why this project needs to run on a Tier-1 system, why the machine you have requested is suitable for the project and how the use of the system will enable the science proposed (max. ½ A4 in Arial 12).

   This project consists of one type of task: molecular dynamics simulations on systems of +10,000 atoms over longer timescales (at least 250 nanoseconds per membrane to get proper statistics). The membrane temperature will be varied to see its effect on the oxygen diffusion. The temperature setting does not change the length of the simulations.
   To make these many long simulations on large systems feasible in an acceptable amount of wall time, the TIER1 infrastructure is required. To make the comparison explicit: 250 ns of simulation time of 1 system is estimated to take *one year and three months* on 16 CPU on TIER2. Based on Table 2, the same simulation will take *41 days* on 8 nodes of TIER1. Indeed, our tests (see Table 2/Plot 1) show that the CHARMM software scales very well on TIER1, demonstrating that TIER1 is suitable hardware for the simulations.

8. Justify the number of node days requested. This should include information such as: number and nature of computing tasks, software used, and the sequence in which they will be performed.
   Indicate for each typical computing task the required resources:

   - wall clock time (note that 3 days is the maximal wall clock time for any job;)
   - memory (maximum 128 GiB/node; 256 GiB/node is available upon motivated request)
   - number of nodes
   - number of CPU cores
   - disk space (estimated volume in GiB and the total number of files); make a clear distinction between usage of Tier-2 DATA/HOME partitions and the Tier-1 SCRATCH partition
   - number of tasks, and an indication of how many such tasks would be submitted concurrently.

   This information should take the form of a table (an example is provided as Table 1 on the next page). Provide additional descriptions of the computing tasks and comments as needed and clearly relate the described tasks to the tasks in the table. Resource estimates should be preferably based on the results of actual calculations on Tier-1 (via, e.g., a Starting Grant) for system/problem sizes that are on par with those of the intended computing tasks (e.g., same mesh sizes, actual molecular system, …). If not, provide the

name, architecture, #cores, memory, etc. of the machine that was used to obtain these results and explain how you have calculated/rescaled the wall clock times, number of cores, etc.

(typically up to 2 A4 Arial 12).

+ see Table 1/Table 2

We have one particular type of simulations: running molecular dynamics simulations with the modelling package CHARMM. A previous TIER1 grant was used to test model systems of similar size.

To enable easy restarts and deal with the 3 day wall clock limit, we split up our 250 ns simulations per modelled membrane system in “runs” (or tasks) of 1 ns, which should be executed one after another. This allows for better management of the output (post-processing) and restarting. Restart files will be written every 1-3 hours. This gives a good trade-off between safety (restart) and performance (limiting i/o).

The membrane temperature will be varied in the simulations to 298 K. Two membranes composition types will be considered: (1) POPC, and (2) DOPC. At most 2 simulations would thus be submitted concurrently.

Task 1: one 1-ns run for one membrane composition takes 0.164 days on 8 TIER1 nodes according to our previous tests. To finalize the 250-ns simulations on 2 systems, it takes: 0.164 x 8 x 250 x 2 = **656 nodedays**. One run requires about 1 GiB of SCRATCH space. Since at most 2 simulations will be run concurrently, we need about 3 GiB SCRATCH space in total. We will regularly offload data to the Tier-2 storage, where 500 GiB is needed.

To perform a run, CHARMM only needs about 15 files (input files, restart files) on TIER1. Each run creates 4 new files. The 500 tasks will therefore generate 500*4 = 2000 files, which may be moved to the TIER2-DATA.

Table 1

	Node day calculation							Storage volume estimate	
Computational task	# of such tasks	Wall clock time (days) per task	# Tier-1 nodes per task	# total node days task	# CPU cores per task	Memory usage (GiB) / node per task	OpenMP / MPI / hybrid	Tier-2 DATA/HOME volume (GiB) + number of files	Tier-1 SCRATCH volume (GiB) + number of files
Task1: Molecular dynamics	500	0.164	8	656 (-> 660)	264	< 64 GB	MPI	500 GiB 2000 files	3 GiB 150 files

9. Describe the software required to perform the computing task(s). Please clearly provide the following per item in this regard:

   - a reference to the software's web page
   - the software license system (open source, GPL, etc.)
   - if there is no free academic use of the software, state which license makes the installation and the use valid on the Tier-1 by the Applicant (+ add a copy of the signed license)
   - if need be, which license server will be used (name + IP address)
   - whether the software is already available on the Tier-1 and, if this is not the case, compilation and installation instructions (possibly with reference to existing Tier-2 installation)

   CHARMM

   - http://www.charmm.org/
   - CHARMM has a commercial licence system. We have used the license previously on the TIER1. After purchase the license is valid forever for that software version.
   - The software is already available on the TIER1 infrastructure. We have also tested that CHARMM is running on the newest TIER1 infrastructure located in Leuven. The software is also available on TIER2.

   Provide the results of scaling tests that were conducted with this software, preferably on the current VSC Tier-1 (using, e.g., a Starting Grant) for system/problem sizes that are on par with those of the intended computing tasks (e.g., same mesh sizes, actual molecular system, …). If not run on the current VSC Tier-1, provide the name, architecture, #cores, memory, etc. of the machine that was used to obtain these results and how you think this compares to the current VSC Tier-1. If a different system/problem size is used provide some guidance how it relates to the problem size in the application.
   Provide both a table and scaling plot such as table 2 and plot 1 below (typically up to 3 A4 in Arial 12).

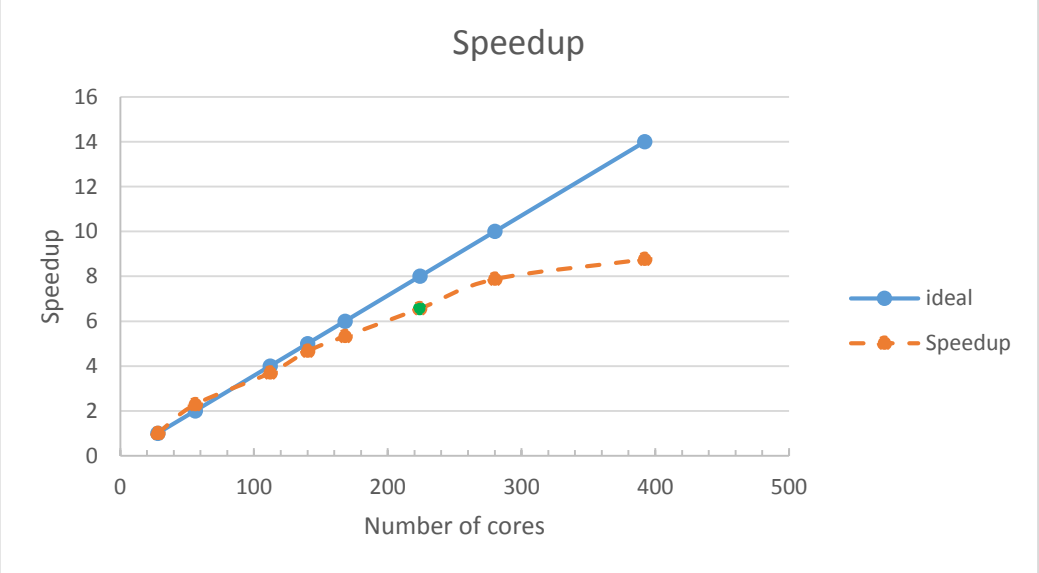
See Table 2 and Plot 1. The testing system had similar size as the membrane systems in this project as described in Task 1.

Table 2

# nodes	# cores	Wall clock time (s)	Speedup (with respect to 1 node)	Efficiency (FORMULE)
1	28	9286	1,0	1,00
2	56	4067	2,3	1,14
4	112	2515	3,7	0,92
5	140	1991	4,7	0,93
6	168	1744	5,3	0,89
8	224	1415	**6,6**	0,82
10	280	1179	7,9	0,79
14	392	1060	8,8	0,63

Table 2: Trajectory of 1e5 time integration steps. The size of the system is a large box (see Task 1). (The “runs” in the project have 1e6 time steps.)

Plot 1



Plot 1: The scalings are computed with the data of Table 2.

10. Describe how you will manage the resources requested in the period during which the task is to be performed. What usage pattern do you anticipate (similar usage on monthly basis, bursts, …)? Provide a data management plan (transfer of files to/from Tier1).

After setting up the simulations (which will take at most 1 week), we expect similar continuous usage per month throughout the project period. Intermediately, the simulations will be checked in order to verify if the simulations are proceeding correctly. We have experience with working on these specific molecular systems so we expect the setting up of the simulation to go smoothly. We will do one validation run on TIER2 in order to check the correctness of the simulation, before submitting to TIER1.

Data management

To enable easy restarts and deal with the 3 day wall clock limit, we split up our 250 ns simulations per modelled membrane system in 250 “runs” of 1 ns (250 tasks of type “Task 1”), which should be executed one after another. This allows for better management of the output (post-processing) and restarting. Restart files will be written every 1-3 hours. This gives a good trade-off between safety (restart) and performance (limiting i/o).

At most 3 simulations will be submitted concurrently.

Task 1:

To perform a run, CHARMM only needs about 15 files (input files, restart files, which takes about 10 MB in total) on TIER1. Each run creates 4 new files (which takes about 1 GB in total).

So, one run requires about 1 GiB of SCRATCH space for the output. Because at most 2 simulations will run concurrently, we need about 3 GiB SCRATCH space in total.

We will regularly offload the output data to the Tier-2 storage, where 500 GiB is needed (there are 250 runs).

11. List the granted computing time allocations to the promoter(s) of this research project, on the Flemish Tier-1 systems, as well as other Tier-1 and Tier-0 systems. Also, describe the scientific output obtained within the framework of computing time that was granted during the past two years on the Flemish Tier-1 systems or on other Tier-1 or Tier-0 supercomputers. DOI links are sufficient.

TIER1 grant (An Ghysels, Samuel Moors)
Diffusion of hydrocarbons in porous frameworks
4320 nodedays, 07/07/2014 – 31/12/2014

Output:
"Shape-selective diffusion of olefins in 8-ring solid acid microporous zeolites" (http://dx.doi.org/10.1021/acs.jpcc.5b06010)

TIER1 Starting Grant (Oriana De Vos, promotor An Ghysels)
100 nodedays, 02/10/2015-31/01/2016
Study of the oxygen diffusion through cell membranes

TIER1 Grant (Oriana De Vos, An Ghysels)
1250 nodedays, 2/2016-31/8/2016
Study of the oxygen diffusion through cell membranes
Output: investigation of first systems

TIER1 Grant (An Ghysels, Oriana De Vos)
700 nodedays, 1/11/2016-30/4/2017
Molecular dynamics simulations of oxygen transport through membranes
Output: testing scaling of newest TIER1 in Leuven

TIER1 Grant (An Ghysels, Oriana De Vos)
1260 nodedays, 1/3/2017-31/8/2017
Effect of periodic boundary conditions on simulations of oxygen transport through membranes
Output: investigation of box size effect

TIER1 Grant (An Ghysels, Oriana De Vos)
985 nodedays, 3/7/2017-3/1/2018
Simulating oxygen transport through membranes with various lipid compositions
Output: investigation of lipid type effect

Publication: "*Effect of (dis)order on oxygen diffusion through biological membranes from simulations*" (in preparation)

Publication: Simulations were performed on the LoBoS cluster of the National Institutes of Health (NIH, Maryland, USA).
Ghysels, Venable, Pastor, Hummer, J. Chem. Theory Comput. (**2017**), 13 (6), pp 2962-2976, **DOI:** 10.1021/acs.jctc.7b00039, *"Position-dependent diffusion tensors in anisotropic media from simulation: oxygen transport in and through membranes."*

12. Are the applicants of this application bound by a confidentiality agreement? If so, the abstract of this application will not be published on the website of the FWO / Flemish Supercomputer Center, only the title.

no

Should you have any questions or encounter any difficulties during the electronic submission of an Application, please contact by e-mail:
Associatie KU Leuven: hpcinfo@kuleuven.be
Associatie Universiteit Gent: hpc@ugent.be
Associatie Universiteit Hogescholen Antwerpen: hpc@uantwerpen.be
Associatie Universiteit Hogescholen Limburg: geertjan.bex@uhasselt.be
Universitaire Associatie Brussel: hpc@vub.ac.be
For the other institutions: caroline.volckaert@FWO.be

DIRECTIE
ONDERZOEKSAANGELEGENHEDEN
AFDELING ONDERZOEKSCOÖRDINATIE

Dirk De Craemer
Afdelingshoofd Onderzoekscoördinatie

E dirk.decraemer@ugent.be
T +32 9 264 30 33
M +32 473 852 426

Campus Ufo, Rectoraat
Sint-Pietersnieuwstraat 25
B-9000 Gent
België

www.ugent.be

Prof. dr. An Ghysels
Vakgroep Toegepaste Fysica.
Tecnologiepark 903,
9052 Zwijnaarde

DATUM	PAGINA	ONS KENMERK
2 oktober 2017	1/1	DOZA/AOC/DDC/17/0904

BETREFT: uw aanvraag voor het bekomen van Tier-1 computertijd

Waarde collega

Ik heb kennis genomen van en ga akkoord met uw aanvraag voor het bekomen van computertijd op de Tier-1 infrastructuur in het kader van uw project getiteld: *Simulating oxygen transport through membranes at various temperatures*, dat kadert in het onderzoeksproject van uw doctoraatsstudente Oriana De Vos.

Met collegiale groet

Prof. dr. Ignace Lemahieu
Voorzitter Onderzoeksraad

UNIVERSITEIT
GENT