Enclosure 1b. Category 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:

Benchmark study of ab initio molecular dynamics simulations for the methylation of HMB

Name and first name of the applicant:

Bailleul Simon

Institution:

Ghent University

Research group / department:

EA17, Center for Molecular Modeling

Title / position:

PhD fellow

e-mail address:

simon.bailleul@ugent.be

Total computing time that is needed, in node days:

4812

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

50 GB scratch space - ~1.5 TB long-term storage (provided by UGent)

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

European Union's Horizon 2020 research and innovation programme (consolidator ERC grant agreement No 647755 – DYNPOR (2015-2020))

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):

In the methanol-to-olefin (MTO) process, which is an alternative for the depleting oil reserves, methanol is transformed into valuable base chemicals, catalyzed by zeolite or zeotype catalysts. In particular the chabazite structured H-SAPO-34 is of industrial interest due to its high selectivity to light olefins. [1] Studies on the MTO reaction mechanism led to the general acceptance of the hydrocarbon pool (HP) mechanism governing the MTO process, where an organic compound trapped in the catalyst acts as a co-catalyst. [2] Polymethylbenzenes are identified as dominant HP species in H-SAPO-34. [3] In this project, the methylation of hexamethylbenzene (HMB) is chosen to represent the MTO process, since it is an important reaction step in the hydrocarbon pool mechanism proposed for H-SAPO-34 (A joint experimental and theoretical study has emphasized the importance of this reaction. They found that the band at 400 nm of the in situ UV/Vis microscopy measurements could be assigned to polymethylbenzenes. Furthermore, the activation energies derived from the growth of this characteristic peak correlated well with the activation energy for methylation of the benzenic species. [4], [5]).

In an earlier project, the reaction profile of the methylation of HMB in H-SAPO-34 was calculated using static and dynamic ab initio calculations. For the dynamic simulations the metadynamics method (MTD) was used. Both reaction profiles are shown in Figure 1a, showing a clear difference in the product state. This difference is due to the insufficient description of the diffusional freedom of the formed water in the static simulation, as shown in Figure 1b. This emphasizes the importance of the use of advanced molecular dynamics simulations in tackling the problem of further elucidating the reaction mechanism of the MTO process. Therefore, the metadynamics simulations are benchmarked by umbrella sampling simulations, which will be performed using this Tier-1 project, to create more certainty on the already obtained results and further increase our insight into the reaction mechanism governing the MTO process.

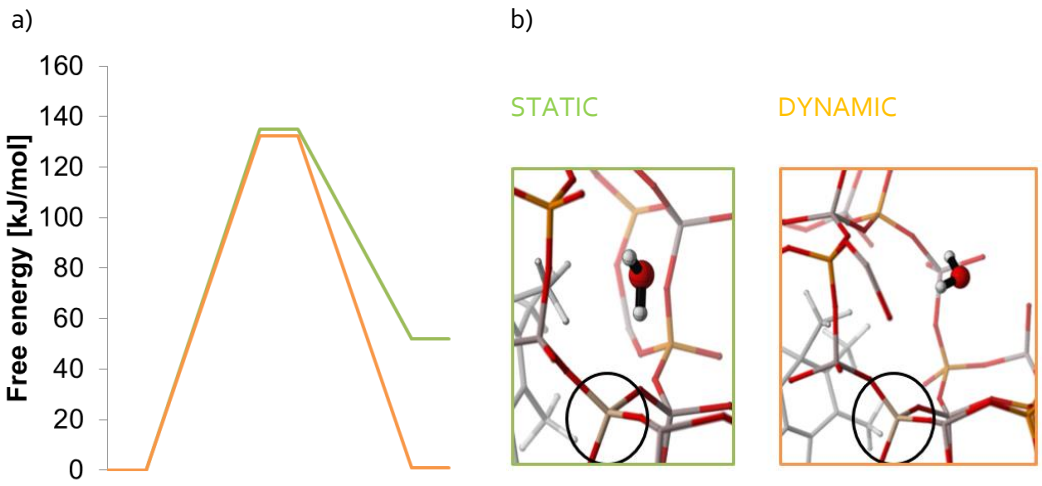

Figure 1: Reaction profile (a) and snapshots (b) of the methylation of HMB using static and dynamic ab initio simulations.

[1] J. Q. Chen et al. Catal. Today 106 (2005) 103.
[2] J. F. Haw and D. M. Marcus Top. Catal. 34 (2005) 41.
[3] B. P. C. Hereijgers et al. J. Catal., 264 (2009) 77.
[4] V. Van Speybroeck et al. ChemCatChem 5 (2013) 173.
[5] K. Hemelsoet et al. Chem. – Eur. J. 19 (2013) 16595.

3. Provide an abstract (10 lines) for scientific communication on the website in layman's terms. See also item 12 of this application form.

No abstract of this work may be published on the website (see item 12)

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 (see instructions in enclosure 3 "EasyChair proposals submission procedure").

European Research Council under the European Union's Horizon 2020 research and innovation programme (consolidator ERC grant agreement No 647755 – DYNPOR (2015-2020))

5. Name and email address of the promoter(s) of the research project:

Prof. Dr. ir. Veronique Van Speybroeck

veronique.vanspeybroeck@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)

Ir. Simon Bailleul

Ghent University, Center for Molecular Modeling

ERC PhD fellow

1 year of experience with TIER1 at Ghent University

2 year of experience with TIER2 at Ghent University

Ir. Pieter Cnudde

Ghent University, Center for Molecular Modeling

PhD fellow

1 ½ year of experience with TIER1 at Ghent University

3 year of experience with TIER2 at Ghent University

Prof. dr. ir. Veronique Van Speybroeck

Ghent University, Center for Molecular Modeling

Full Professor

2 ½ years of experience with TIER1 at Ghent University

7 year of experience with TIER2 at Ghent University

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).

To perform all simulations proposed in this study, a computing time of over 4000 node days will be needed to obtain accurate statistical averages from the umbrella sampling simulations. To complete this project in an acceptable timeframe, multi-node jobs are required, which makes the ability to run a large amount of jobs simultaneously on the Tier-1 infrastructure essential for the speedup of this research.

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; checkpointing should be used for longer run times)
   - memory (maximum 64 GiB/node)

- 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 2 in the appendix). Provide additional descriptions of the computing tasks and comments as needed. 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.

(max. 1 A4 Arial 12).

This type of ab initio molecular dynamics based simulations will be performed with the CP2K software package on a system consisting of 145 atoms. The computational requirements for this study have been carefully considered based on test simulations on the systems under study performed on the Tier-1 infrastructure at Leuven University.

During this project, two reactions will be analyzed, since the methylation of HMB can occur via a direct or a stepwise mechanism. The following calculation workflow will be used to model each of the two reactions:

1. **Creation of umbrellas:**
   Before the different windows in the umbrella sampling can be created, the system is equilibrated for 2 ps using an NVT ensemble at realistic reaction conditions, namely 623 K and 1 atm. Subsequently, to create the snapshots used in the different umbrellas of the simulations, a bias potential is moved along a reaction coordinate to drive the system from the reactant to the product state at the same conditions. This creation of the windows will take another 10 ps, which can be performed in **6 node days**. For each reaction, 40 snapshots are created using this procedure.
2. **Umbrella sampling simulations:**
   For each window created in step 1, 100 ps of biased molecular dynamics is needed to obtain converged results. These simulations are performed in the NVT ensemble at 623 K and 1 atm. Molecular dynamics simulations of similar systems on Tier-1 have shown that

5 ps simulations require approximately 3 node days. Therefore, each window will take **60 node days** to converge.

Table 1 summarizes the estimated node and core days required for each simulation type.

Table 1: Estimated core and node days required for the project.

Job type	# of simulations	Node days per simulation	Total node days	Total core days
Biased MD	2	6	12	336
US	80	60	4800	134400
			4812	**134736**

Table 2 summarizes the requirements for scratch space and long-term storage for each simulation type. Because of the long simulation times, a complete MD trajectory cannot be obtained within the wall time limit of 72 hours, but requires several restarts. As a result, the scratch volume per simulation is relatively low.

Table 2: Estimated scratch space and long-term storage requirements for the project.

Job type	Scratch [GB/run]	Long term storage [GB/run]	Total scratch [GB]	Total long term storage [GB]
Biased MD	0.5	10	2*0.5 = 1	2*10 = 20
US	0.5	20	80* 0.5 = 40	80*20 = 1600
			41	**1620**

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 (see https://www.vscentrum.be/cluster-doc/software/tier1-muk) and, if this is not the case, compilation and installation instructions (possibly with reference to existing Tier-2 installation)

Molecular dynamics simulations will be performed using the CP2K software package (http://www.cp2k.org/) with PLUMED code (http://www.plumed.org), which are freely available under the GPL license. The required version of the program (CP2K/20150904-intel-2015a-PLUMED-2.1.3) is not yet well compiled on Leuven Tier-1 but is available on the Gent Tier-1 machine.

Provide the results of scaling tests that were conducted with this software, preferably on 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, provide the name, architecture, #cores, memory, etc. of the machine that was used to obtain these results.
Provide both a table and scaling plot such as table 1 and plot 1 in the appendix (max. 2 A4 in Arial 12).

Scaling of the CP2K software (version CP2K/3.0-intel-2016a on the Tier-1 in Leuven) has been performed for a system similar to the system under study, i.e. H-ZSM-5. Short NPT runs (50 steps) have been performed on 28, 56, 112, 224 and 448 cores on the Tier-1 cluster for which the results are summarized in Table 3. Furthermore, Figure 2 shows the speedup normalized to the 28 core simulation for these simulations. This graph indicates that optimal scaling is achieved up to 2 nodes.

Table 3: Results of CP2K scaling test performed on Tier-1 for H-ZSM-5.

# nodes	# cores	absolute timing (s)	speedup	# cores x timing
1	28	499.2	1	13977.6
2	56	267.4	1.866866	14974.4
4	112	169.4	2.946871	18972.8
8	224	108.7	4.592456	24348.8
16	448	88.1	5.666288	39468.8

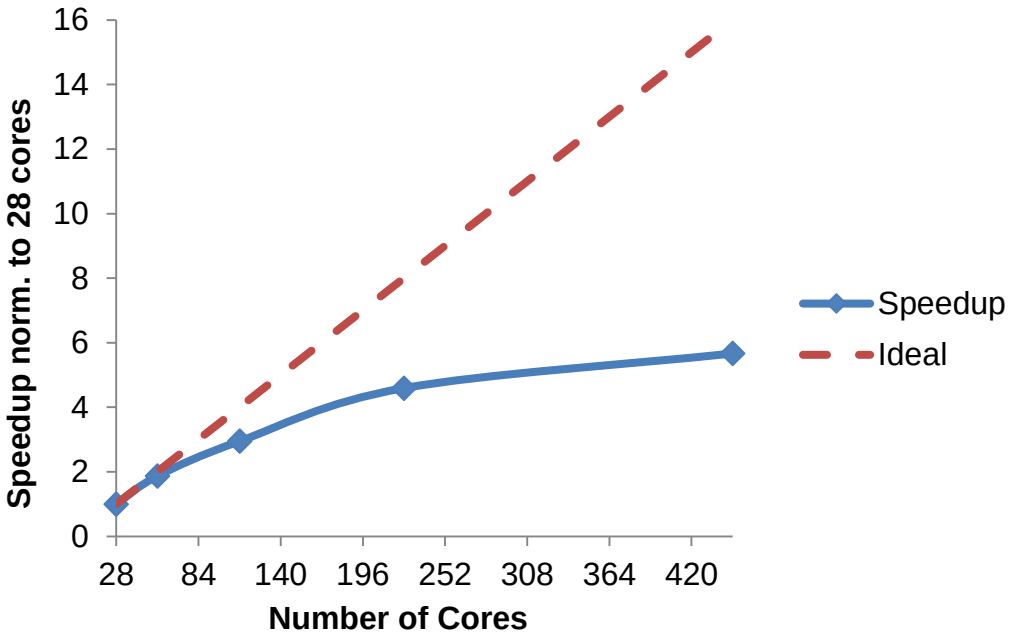

Figure 2: CP2K scaling tests performed on Tier-1 for H-ZSM-5.

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, …)?

We estimate the proposed simulations can be completed over a time period of approximately **6 months**, preferably **starting in January 2017**. Similar usage on a monthly basis is expected.

This estimated timeframe takes into account the actual runtime of the simulations, as well as intermediate data analysis tasks and frequent job restarts.

11. List the granted computing time allocations to the promoter(s) of this research project, on the Flemish Tier-1 system, 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 or on other Tier-1 or Tier-0 supercomputers. DOI links are sufficient.

List of granted research projects of the Flemish TIER1 system:

- Dynamical kinetic study of zeolite catalyzed reactions (K. De Wispelaere, 4371 node days, 07/07/2014 – 31/12/2014)

- Exploring the kinetics and selectivity of butene cracking using molecular dynamics simulations (J. Van der Mynsbrugge, 4864 node days, 01/01/2015 – 30/06/2015)
- Characterizing adsorption properties of C4 – C6 alkenes on H-ZSM-5 using molecular dynamics simulations (P. Cnudde, 4260 node days, 13/07/2015 – 31/12/2015)
- Dynamical first principle benchmark studies on alkene methylation in H-ZSM-5 (K. De Wispelaere, 1400 node days, 01/12/2015 – 30/06/2016)
- Ab initio molecular dynamics study on the role of water in the reaction mechanism during methanol conversion in H-SAPO-34 (S. Bailleul, 4880 node days, 01/03/2016 – 31/08/2016)
- DFT study of reaction paths in zeolite-catalyzed 2-hexene cracking (P. Cnudde, 4536 node days, 15/08/2016 – 31/12/2016)
- Dynamical first principle modelling of zeolite dealumination in H-SSZ-13 (K. De Wispelaere, 3624 node days, 01/07/2016 – 31/10/2016)

List of scientific output within the framework of granted computing time:

- http://dx.doi.org/10.1016/j.jcat.2016.05.018
- http://dx.doi.org/10.1002/cctc.201600650
- http://dx.doi.org/10.1039/C5CY02073E
- http://dx.doi.org/10.1021/acscatal.5b02139
- http://dx.doi.org/10.1002/chem.201500473
- http://dx.doi.org/10.1016/j.jcat.2015.01.013

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

Yes

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: rosette.vandenbroucke@vub.ac.be
For the other institutions: caroline.volckaert@FWO.be

Appendix: Conformation letter