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://vscentrum.be/.

Title of the application: Diffusion of hydrocarbons in porous frameworks

Name and first name of the applicant: Ghysels An

Institution: Ghent University

Research group / department: Center for Molecular Modeling

Title / position: PhD, post-doctoral researcher

e-mail address: an.ghysels@ugent.be

Total computing time that is needed, in node days: 4320 nodedays

Total disk storage that is applied for: 720 GB

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

Multiscale modelling of diffusion in nanoporous materials

2. Short description of the research project within the framework of which computing time is applied for (max. 1 A4 in Arial 12):

Zeolites (SiO_2) are crystalline frameworks with nano-pores and nano-channels, in which guest molecules may adsorb and diffuse. By introducing Al and/or P atoms in the crystal, acid sites are created, onto which the guest molecules may undergo chemical reactions. As such, zeolites can act as catalysts in industrial

processes. An example is the metanol-to-olefin (MTO) process, where methanol conversion to industrially relevant olefins is achieved through a network of reactions. Unravelling this network is still a challenge, because experimentally only the *overall* reaction rate is measured, but not the *individual* reaction rates. Moreover, during the conversion process, coke formation clutters up the nanopores with bulky aromatic compounds, possibly hindering the mobility of the guest molecules.

Ultimately, the overall reaction rate is determined by the combination of individual reaction rates and this mobility of guest. The aim of this research is to investigate the mobility aspect with molecular dynamics (MD), by modeling diffusion of the compounds of the MTO process in several nanoporous materials, such as SAPO-34.

The methodology is well known: the trajectory of the guest molecule through the material is traced with MD, the mean square displacement Δr^2 is plotted as a function of time t, and the slope of this line gives the diffusivity D : $\Delta r^2 = 6\, D\, t$.

However, diffusion plays at a relatively large time scale. For example, ethene diffuses over a distance of approximately 3 nm in 100 ns time in the SAPO-34 material, illustrating that long simulation times are required. Moreover, the statistics of the mean-square displacement converge poorly, and long MD runs are needed to observe statistically significant differences in diffusivities. The time scale and the statistics motivate this application for TIER1 computer time.

We believe that our project is truly the next step in realistic modelling of the complexity playing at the atomic scale. Five years ago it would be computationally impossible. Now we can finally move forward and tackle the following research questions:
(1) What is the role of acid sites to diffusion? (2) How does coke formation affect the mobility? (3) Can we predict the role of framework flexibility?

These questions will be studied by varying the reaction conditions. The role of acid sites and coke formation will be tested by varying their distribution over the framework. Framework flexibility is known in literature to play an essential role, and various temperatures will be used to investigate the diffusion activation energy. Each variation in the reaction conditions (acid sited, cokes, temperature) requires a separate MD simulation to compute the diffusivity.

Finally, recent simulations in our group suggest that temperature creates geometrical distortions in the framework, in turn giving rise to both 'slow' and 'fast' diffusion channels. Therefore we foresee to not only compute diffusivities but also to perform a detailed analysis of structural properties.

3. 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).

FWO post-doctoral mandate of An Ghysels, see confirmation letter.

4. Promoter of the research project:

Prof. Veronique Van Speybroeck
Center for Molecular Modeling
Ghent University
veronique.vanspeybroeck@ugent.be

5. Persons mandated by the Applicant to compute on the TIER1 within the framework of the present project: Please provide for every person:

- name and first name
- institution
- research group / department
- title / position
- experience with TIER1/TIER2 infrastructure in Belgium and abroad

An Ghysels
Ghent University
Center for Molecular Modeling
PhD, post-doctoral researcher
5 years experience with TIER2 at Ghent University
5 years experience with LOBOS infrastructure at NIH
1 year experience with Beowulf infrastructure at NIH

Samuel Moors
Ghent University
Center for Molecular Modeling
PhD, post-doctoral researcher
½ year experience with TIER1 at Ghent University
2 years experience with TIER2 at Ghent University
5 years experience with TIER2 at KU Leuven

Veronique Van Speybroeck
Ghent University
Center for Molecular Modeling
Full professor
5 years experience with TIER2 at Ghent University
1 year experience with TIER1 at Ghent University

6. Description of the computing task, justification for the computing time, disk storage and memory that are applied for, and description of the software tools required (max. 3 A4 in Arial 12). Please clearly provide the following in this regard:

- the number of nodes/cores that are applied for per computing task, with a subdivision of the computing time in sub-tasks indicating the sequence of the sub-tasks
- whether these tasks use diversification (OpenMP, MPI, hybrid OpenMP/MPI …)
- the estimated memory use of a computing task (maximum 64GiB/node)
- whether a vSMP system will be used
- the requirements for disk storage (estimated volume in GiB and the total number of files), more specifically for:
    - required input files (data set, parameter files, etc.)
    - SCRATCH volume used during the performing of the computing tasks
    - result files

Jobs: 0.6 nodedays/job
MD simulations will be performed with DL_PLOY_CLASSIC. Each job is a single MD run of 10^7 time steps takes 0.6 nodedays on TIER1.

Tasks/Scaling: Figure 1 shows that the scaling on TIER1 is linear up to 16 cores (1 node). When using multiple nodes, the performance is poor. Therefore, the plan is to have **one job per node.**

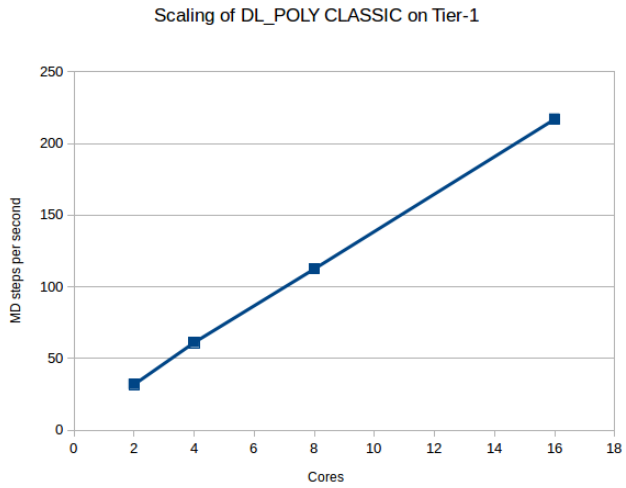

Figure 1. The progress of the DL_POLY simulation (number of MD steps) scales linearly with the number of cores.

Number of jobs:

We compute here how many jobs are needed for the project. The summary of the simulations may be found in Table 1: **12 materials** will be investigated at 30 different conditions.

Table 1 Overview of diffusivities to be calculated.

12 materials		30 conditions	
topology	chemical composition	temperature	loading
3x	4x	3x	10x
CHA AFX AEI	pure zeolite pure sapo zeolite with acid sites sapo with acid sites	300K 450K 600K	empty presence propene presence hmb presence pyrene
Total: 3 x 4 x 3 x 10 = 360 diffusivities			

- Three framework topologies will be tested: CHA, AFX, and AEI (3x).
- Each of these topologies can have various chemical compositions: pure zeolite, pure sapo, zeolite with acid sites, sapo with acid sites (4x).
- To assess the diffusion activation energy, these 12 materials will be modelled at temperatures of 300K, 450K, and 600K (3x).
- The 12 materials may be loaded with a very large variety of reactants and products concentrations. Here, we will focus on ethylene diffusion in the following circumstances:

a) in the empty material,
b) in presence of propylene,
c) in presence of the coke product hexamethylbenzene (hmb) - which may be distributed in several ways over the material's cages,
d) in presence of the coke product pyrene - which may be distributed in

several ways over the material's cages as well.
Instead of blindly trying out the several distributions of the propylene/hmb/pyrene over the cages, we will start with testing the most plausible distribution, such that we can learn from these tests to make a guided choice when trying out other distributions. A realistic estimate would be to try out 10 distributions of the loading in total (10x).

- ⇨ total number to be computed = 12 x 3 x 10 = **360 diffusivities**.

About **20 MD runs** (= 20 jobs) are required to reach the time scale of diffusion (100 to 200 ns) with acceptable statistics.

- ⇨ total number of MD runs = 360 x 20 jobs = **7200 jobs**
- ⇨ total nodedays = 360 jobs x 0.6 nodeday/job = **4320 nodedays**

Memory use:
maximum 100 MB per computing task
from tests: Maximum resident set size (kbytes): 46416

vSMP: No vSMP will be used.

Disk storage:
input files: 50 MB
scratch: 2 GB/job
output files: 100 MB/job
720 GB to be stored in total, may be transferred to another local storage location during the duration of the project, such that the post-analysis is not necessarily done on the TIER1.

Plan B
In case the diffusion is too slow for one or more of the systems in Table 1, we will rely on advanced ab initio free energy MD simulations such as metadynamics for these systems. These simulations will allow us to derive diffusion properties from free energy barriers of the small molecules travelling through the 8-rings. These free energy calculations, performed with the open source CP2K simulation package are well established by the applicants, and are known to run and scale well on the TIER1 computing infrastructure.

7. Please indicate why the TIER1 is the appropriate machine to perform the computing task (max. 1/2 A4 in Arial 12):

At the Center for Molecular Modeling (CMM), we have recently conducted our first successful study on the diffusion of small molecules in the framework of the MTO process, of which the report is currently in preparation. However, to take this research to the next level, much longer timescale force field MD simulations, as well advanced free energy ab intio MD simulations are needed.

Performing all the simulations summed in Table 1, requires the execution of approximately 7200 jobs, corresponding to 69120 core days (**4320 nodedays**). Being able to perform this large amount of jobs on the highly efficient TIER1 infrastructure can significantly speed up our research. The large amount of simulations proposed in this project is essential to obtain statistically relevant diffusion data. Moreover, we anticipate that the combination of large timescale MD simulations on the force field level with advanced ab initio free energy simulations will provide unique insight into the diffusion process of small molecules and thus provide significant contribution the highly competitive and fast moving field.

8. Summary of the software required to perform the computing task, and possible installation and compilation instructions (max. 2 A4 in Arial 12). Please clearly provide the following per item in this regard:

- a reference to the software's web page
- the software licence system (open source, GPL, etc.)
- if there is no free academic use of the software, state which licence makes the installation and the use valid on the TIER1 by the Applicant (+ add a copy of the signed licence)
- if need be, which licence server will be used (name + IP address)
- whether the software is already available on the TIER1 (see https://vscentrum.be/nl/tier1-rekenen) and, if this is not the case, compilation and installation instructions (possibly with reference to existing TIER2 installation)

Software: DL_POLY CLASSIC

Website: http://www.ccp5.ac.uk/DL_POLY_CLASSIC/

Licence: open source (BSD)

Already available: DL_POLY CLASSIC was already manually installed by the applicant on the TIER1 cluster.

9. Period during which the task is to be performed:

The estimated project time is 6 to 7 months, starting as soon as possible. Although the estimated total simulation time will be less, frequent analysis and restarting of simulations involves a significant overhead, which has to be accounted for.

10. Describe the results that were obtained within the framework of computing time that was attributed during the past two years on the TIER1 or on other TIER1 or TIER0 supercomputers (max. 2 A4 in Arial 12):

Using the TIER2 infrastructure of Ghent University, we have recently written two important articles that rely on advanced ab initio molecular dynamics simulations. We are convinced that these works move the field of zeolite-catalyzed chemical kinetics forward by explicitly taking into account the dynamics of the system:

- S. Moors, K. De Wispelaere, J. Van der Mynsbrugge, M. Waroquier, V. Van Speybroeck , Molecular dynamics kinetic study on the zeolite-catalyzed benzene methylation in ZSM-5 , ACS Catalysis, 2013 (3), 2556–2567.

- J. Van der Mynsbrugge, S. Moors, K. De Wispelaere, V. Van Speybroeck, Insight into the formation and reactivity of framework-bound methoxide species in H-ZSM-5 from static and dynamic molecular simulations, ChemCatChem, 2014, (in press).

However, the simulations that we performed for this research effectively required many months of simulation and analysis on the TIER2 infrastructure. We are currently performing follow-up simulations on the VSC TIER1 supercomputer, which enable us to move much faster, which is paramount to remain at the top of this highly competitive field.

As a direct result of these TIER1 simulations, we currently have one research article in preparation, and we expect at least two more articles to be submitted this year:

- M. Erichsen, K. De Wispelaere, K. Hemelsoet, S. Moors, T. Deconinck, S. Svelle, V. Van Speybroeck, U. Olsbye, How zeolitic acid strength and composition determines the reactivity of alkenes and aromatics during methanol conversion, 2014 (in preparation).

Should you have any questions or encounter any difficulties during the electronic submission of an Application, please contact by e-mail:

Associatie KU Leuven: leen.vanrentergem@kuleuven.be ; jan.ooghe@kuleuven.be

Associatie Universiteit Gent: hpc@ugent.be

Associatie Universiteit Hogescholen Antwerpen: stefan.becuwe@uantwerpen.be

Associatie Universiteit Hogescholen Limburg: geertjan.bex@uhasselt.be

Universitaire Associatie Brussel: rosette.vandenbroucke@vub.ac.be ; rosette.vandenbroucke@gmail.com

For the other institutions: dane.skow@herculesstichting.be