Enclosure 1b. Category 1 Application form

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:

Dynamical kinetic study of zeolite-catalyzed reactions

Name and first name of the applicant:

De Wispelaere Kristof

Institution:

Ghent University

Research group / department:

Center for Molecular Modeling

Title / position:

FWO-PhD fellow

e-mail address:

Kristof.dewispelaere@ugent.be

Total computing time that is needed, in node days:

4371

Total disk storage that is applied for:

~ 2 TB SCRATCH, ~ 2 TB longer-term storage (provided by Ghent University)

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

Doctoral research project “Ab initio investigation on the activity and selectivity of single active sites in nanoporous materials” funded by FWO. (https://molmod.ugent.be/members/kristof-de-wispelaere).

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

In this project we will model three zeolite-catalyzed reaction types: methylations, methoxide formation and cracking reactions catalyzed by H-SAPO-5, H-SSZ-24, H-ZSM-5 and H-SAPO-34. When modeling these key reaction steps, taking into account the zeolite’s framework flexibility and the presence of several assisting molecules in the reaction environment requires a dynamic approach.

We will perform large-scale ab initio molecular dynamics (AIMD) free energy simulations in four steps. Firstly, MD runs will be performed to assess the behavior and reactivity of all reactants. Secondly, metadynamics (MTD) simulations will be performed to determine a free energy surface (FES) of the studied reaction (Figure 1). From this FES, a set of geometrical, thermodynamic and kinetic data can be inferred. Thirdly, a detailed committor analysis will shed light on the quality of the MTD simulations. This requires the generation of hundreds of MD paths on the obtained FES. Finally, the transition path sampling (TPS) approach will be applied based on an initial reactive trajectory obtained from the MTD simulations. By generating hundreds of trajectories by random ‘shooting moves’, one obtains an ensemble of transition paths connecting the reactants with the products free energy well. This ensemble of paths then serves to (a) test the completeness of the MTD sampling and (b) obtain more accurate kinetic data as one does no longer use the approximations of the transition state theory.

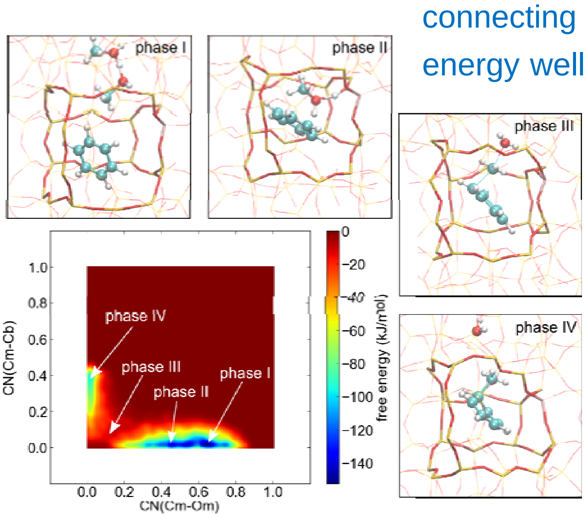

Figure 1. 2D free energy surface of the methylation of benzene with multiple methanols in H-ZSM-5 from metadynamics.

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).

Pre-doctoral FWO grant (see attachments)

4. Promoter of the research project:

Prof. Dr. ir. Veronique Van Speybroeck

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

Kristof De Wispelaere

Ghent University, Center for Molecular Modeling

FWO-PhD fellow

½ year experience with TIER1 at Ghent University

½ year experience with TIER1 and TIER2 in The Netherlands (Local infrastructure from UvA and Lisa cluster - SurfSARA)

3½ years experience with TIER2 at Ghent University

Samuel Moors

Ghent University, Center for Molecular Modeling

Post-doc

½ year experience with TIER1 at Ghent University

1½ years experience with TIER2 at Ghent University

5 years experience with TIER2 at KU Leuven

Jeroen Van der Mynsbrugge

Ghent University, Center for Molecular Modeling

Post-doc

5 years experience with TIER2 at Ghent University

½ year experience with TIER1 at Ghent University

Veronique Van Speybroeck

Ghent University, Center for Molecular Modeling

Full professor

5 years experience with TIER2 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

Large scale ab initio molecular dynamics (AIMD) simulations will be performed with the CP2K software package, using MPI. No vSMP system will be used.

Four types of AIMD simulations will be performed for each system – i.e. combination of catalyst material and reactant(s) – resulting in a total of approximately 6000 jobs. For approximately 50 jobs, multi-node jobs will be performed (the MD and MTD simulations, *vide infra*), meaning that the majority of the jobs will be single-node jobs.

(a) MD simulations in the NPT ensemble at high temperature, with a duration of 50 ps are required to assess the behaviour of the reactants and to obtain fully converged cell parameters for further simulations. Depending on the system, each of these simulations requires 10 to 23 node days. These estimates result from preliminary simulations on very similar systems, performed on the TIER2 (delcatty) infrastructure at Ghent University. The hardware from the delcatty cluster is comparable to the TIER1 system.
(b) On each system, a MTD simulation will be performed. Typical simulation times amount to 100 – 200 ps, depending on the height of the free energy barrier of the studied reaction. Depending on the size of the zeolite's unit cell, the amount of reactant molecules present in the system and the characteristics of the reactions under investigation, each simulation will take up to 35 node days.
(c) To assess the quality of the obtained results from the simulations performed in part (b), a committor analysis (CA) is performed by generating hundreds of AIMD paths based on the results from (b). From preliminary test simulations on delcatty, the time to generate 1 path is estimated to be 8h for the smallest unit cells (1 node, 16 cores on delcatty) and 18h for larger unit cells like for H-ZSM-5. To obtain sufficient sampling for statistical relevant results, it is estimated that at least 200 paths will have to be generated.
(d) A representative ensemble of transition paths will be generated with TPS. From a selection of points on a reactive pathway connecting reactants and products, a Monte Carlo step is performed – i.e. a perturbation of the momentum vector – to construct a number of new paths. By repeating this procedure for every obtained reactive pathway, a tree of hundreds of reactive pathways can be generated. To obtain an ensemble consisting of at least hundred successful paths – i.e. paths that successfully connect two stable states – a couple of hundred AIMD paths have to be generated. From preliminary test simulations on delcatty, the time to generate 1 path is estimated to be 8h (1 node, 16 cores on delcatty). To obtain sufficient sampling for statistical relevant results, it is estimated that at least 200 paths will have to be generated.

Scaling tests of CP2K on TIER1 performed by pilor user Andy Van Yperen – De Deyne has shown that excellent scaling is obtained up to 64 cores for a system with a similar number of atoms (Figure 2). Based on this test, good estimated of the computational requirements for this study could be made. Table 1 summarizes the required node and core days per type of simulation of all the systems under investigation. Similarly, Table 2 summarizes the estimated memory requirements. As the length of the simulations does not permit to complete a typical MD or MTD simulation in one run, the required SCRATCH storage per MD of MTD simulation – consisting of serial runs – is relatively low.

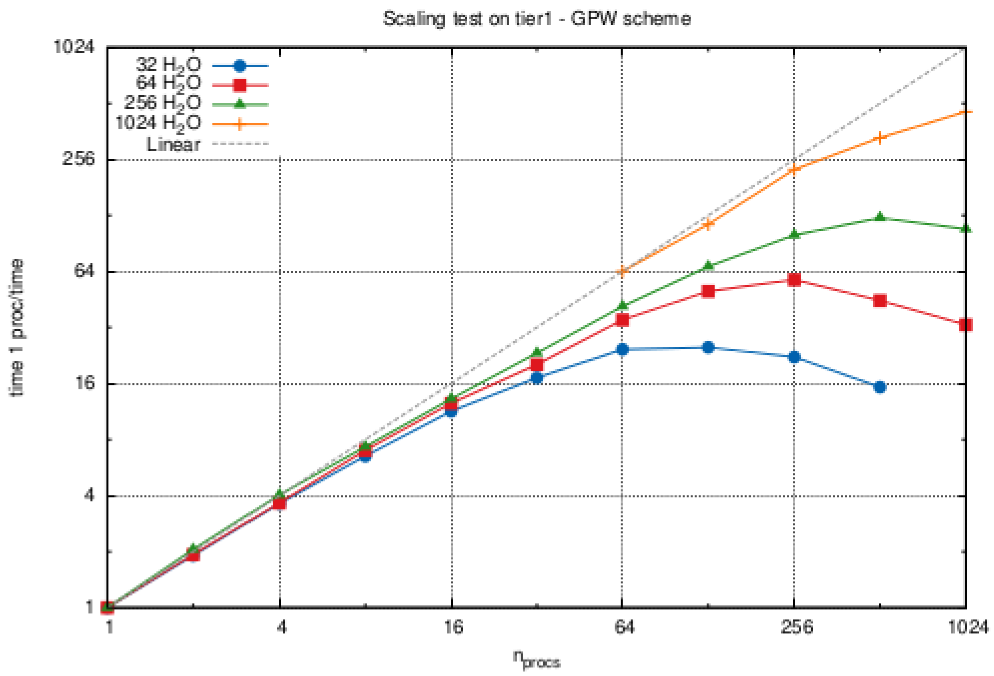

Figure. 2. Efficiency vs. number of CPUs for various system sizes with the GPW code of CP2K.

Table 1. Estimated core and node days for this project

Zeolite	Guest molecules	MD [node days]	MTD [node days]	CA [node days]	TPS [node days]	Total node days	Total core days
H-SAPO-5 (145 atoms / unit cell)	Benzene / propene + 1-3 methanols (2*3)	10	27	67	67	2*3*(10+27 +67+67) = 1026	1026*16
H-SSZ-24 (145 atoms / unit cell)	Benzene / propene + 1-3 methanols (2*3)	10	27	67	67	2*3*(10+27 +67+67) = 1026	1026*16
H-ZSM-5 (289 atoms / unit cell)	2 C4 and 4 C8 compounds (6)	23	35	150	150	6*(23+35+ 150+150) = 2148	2148*16
H-SAPO-34 (108 atoms / unit cell)	MeOH:H_2O in 1:0, 1:1 and 1:4 ratios (3)	10	27	67	67	3*(10+27+ 67+67) = 513	513*16
						4713	75408

Job type	SCRATCH [GB]	Permanent storage [GB]	Total SCRATCH [GB]	Total permanent storage [GB]
MD	1	5	(6+6+6+3)*1 = 21	(6+6+6+3)*5 = 105
MTD	1	5	(6+6+6+3)*1 = 21	(6+6+6+3)*5 = 105
CA	0.2	0.2	(6+6+6+3)*200*0.2 = 840	(6+6+6+3)*200*0.2 = 840
TPS	0.2	0.2	(6+6+6+3)*200*0.2 = 840	(6+6+6+3)*200*0.2 = 840
			1722	1890

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

Performing all the simulations as states in this proposal requires the execution of approximately 6000 jobs, corresponding with the requirement of more than 4000 node days. Being able to perform a large number of jobs on the highly efficient TIER1 infrastructure can significantly speed up this research. Moreover, the large amount of simulations required to perform a statistically relevant committor analysis and TPS simulation retains the use of this kind of simulations in the area of zeolite-catalyzed reactions where AIMD simulations are indispensable to accurately describe heterogeneously catalyzed reactions. Furthermore, the application of advanced molecular dynamics simulations in the field of heterogeneous catalysis is emerging. The Center for Molecular Modeling was one of the first research groups to succeed in applying metadynamics for the prediction of kinetics of zeolite-catalyzed reactions (S.L.C. Moors et al, *ACS Catal.*, 2013, 2556-2567). To be able to make some more substantial contributions with high impact to this highly competitive field, the involved researchers are in need of access to the TIER1 infrastructure to match the high MPI-standards for this work.

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)

The MD simulations will be performed using the CP2K software package: http://www.cp2k.org/

CP2K is freely available under the GPL license. The right version (CP2K/20130228-ictce-4.1.13) is already available on TIER1.

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

We estimate a simulation time period of approximately 6 months, starting a soon as possible. We estimate that next to long simulations times, also frequent intermediate analysis and restarting of simulations has to be performed.

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):

During the past year, highly valuable results were obtained with the dynamical approach as described in this project. The results were/will be reported in high impact journals; a list of relevant publications is listed below. Most of the simulations were performed on the TIER2 infrastructure and during, however to extend our methodology and obtain more accurate data and detailed insights, access to the TIER1 infrastructure will be indispensable.

- S.L.C. Moors, K. De Wispelaere, J. Van der Mynsbrugge, M. Waroquier, and V. Van Speybroeck, ACS Catal. (2013) 2556-2567.
- J. Van der Mynsbrugge, S. Moors, K. De Wispelaere, V. Van Speybroeck, ChemCatChem, 2014, in press.
- K. De Wispelaere et al., “Influence of water on the MTO reaction in H-SAPO-34”, 2014, in preparation.
- M. Westgard Erichsen, K. De Wispelaere et al., “How zeolitic acid strength and composition alter the reactivity of alkenes and aromatics towards methanol”, 2014, in preparation.