

## 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:

**Modeling aromatics formation under MTO conditions**

Name and first name of the applicant:

**De Wispelaere Kristof**

Institution:

**Ghent University**

Research group / department:

**Center for Molecular Modeling**

Title / position:

**Postdoctoral researcher**

Email address:

**kristof.dewispelaere@ugent.be**

Total computing time that is needed, in node days:

**4858**

Total scratch disk storage that is applied for (in GiB) and number of files:

**107 GiB scratch space – ca. 1 TB long-term storage (provided by UGent) –  
18480 files**

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

If this is a resubmission or continuation from an earlier Tier-1 proposal, mention the EasyChair code (e.g. 2018-XX).

Selectivity tuning in the alcohol to hydrocarbons conversion by molecular engineering of active sites (IWETO code 3E004716)

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):

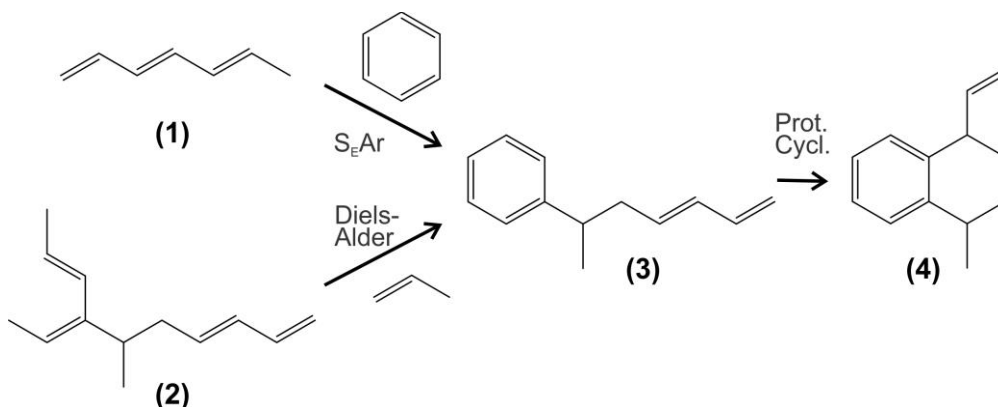
Conversion of methanol-to-olefins (MTO) provides an interesting alternative to crude oil based processes to produce fuels, plastics and chemicals. A crucial aspect in this zeolite-catalyzed process is obtaining optimal control on the product selectivity and catalyst lifetime. Decades of intensive research revealed that limiting the amount of aromatics formed in the pores of the zeolite maximizes the propene yield – the most wanted product in this process – and the catalyst lifetime.[1] **To enable the design of a next generation of highly efficient catalysts we need to obtain a deep understanding on how aromatics are formed in situ in the zeolite pores at operating conditions.**

Recent experimental work identified polyene compounds (molecules **1** and **2** in Figure 1) as precursors for aromatics formation in zeolite H-SSZ-13 through detailed Raman spectroscopy experiments. Via advanced molecular modeling tools we now want to obtain a profound understanding of how these polyenes behave in a CHA zeolite at operating conditions and how they can react to form (poly)aromatics (molecules **3** and **4** in Figure 1).

We will focus on the following 3 aspects

- a) How do polyenes and its derivatives behave at operating conditions in terms of mobility and nature of the adsorption intermediates (i.e. physisorbed or chemisorbed)? Particular focus will be on the presence of multiple active sites per unit cell.
- b) Which reaction paths are viable to form (poly)aromatics? Based on experimental data pathways via an electrophilic aromatic substitution ( $S_EAr$ ) or Diels-Alder type reaction are proposed (Figure 1). Via further protonation and cyclization precursors for polyaromatics can be formed.
- c) Do the calculated Raman spectra of the intermediates match with the experimentally measured spectra? Calculating Raman spectra will allow a one-on-one comparison with experimental results.

The search for new reaction paths will be guided by experimental spectroscopy results in a close collaboration with the group of Prof. Andy Beale (University College London).



**Figure 1.** Proposed pathways for the conversion of polyenes into (poly)aromatics in CHA.

[1] K. Hemelsoet, J. Van der Mynsbrugge, K. De Wispelaere, M. Waroquier, V. Van Speybroeck, *ChemPhysChem* **2013**, *14*, 1526.

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.

In the search for sustainable carbon resources, the use of elementary carbon building blocks such as  $CO_2$ , CO or methanol becomes more and more important. A typical example is the conversion of methanol-to-olefins. After decades of research it is still not clear how aromatics are formed in the zeolite pores during the early stages of the reaction. Based on experimental results we investigate potential reaction mechanisms for (poly)aromatics formation in a CHA catalyst under MTO conditions. To this end we will apply advanced molecular dynamics techniques to account for true operating conditions.

4. Financing institution or channel, financing the research project in full or in part (FWO, BOF, IWT, EU, etc.): 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.

FWO postdoctoral fellowship Kristof De Wispelaere FWO16-PDO-047

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 (Tier-0/Tier-1/Tier-2 infrastructure in Belgium and abroad) in the past (specify names and when)

Dr. ir. Kristof De Wispelaere – vsc40490

Ghent University, Center for Molecular Modeling

FWO Postdoctoral researcher

3 years experience with Tier-1 infrastructure Muk

1½ year experience with Tier-1 infrastructure BrENIAC

½ year of experience with Tier-1 and Tier-2 in The Netherlands (SurfSara)

7.5 years of experience with Tier-2 at Ghent University

1 year of experience with Tier-1 and Tier-2 at Stanford University/SLAC

Ir. Alexander Hoffman – vsc41949

Ghent University, Center for Molecular Modeling

PhD student

1½ year experience with HPC UGent Tier-2 clusters

1½ year experience with Tier-1 infrastructure BrENIAC

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).

Mechanistic investigations with respect to the early stages of the MTO reaction are a very hot topic nowadays. The CMM has ample experience with such studies, but **to maintain a leading position in the field it is vital to have access to the most performant machines** available. As we want to obtain insights on the behavior of intermediates and the kinetics of complex reaction paths at operating conditions we will apply first principle advanced molecular dynamics methods. Such simulations are

computationally very demanding and require access to Tier-1 to allow obtaining results within a reasonable time span. The computational expense for molecular dynamics based simulations is due to the long simulation times that are required to ensure a statistically relevant sampling. The CMM has a leading role in applying first principles advanced MD simulations to study complex conversions in nanoporous materials leading to high-level publications in top journals (Chemical Science, ACS Catalysis, Nature Materials ...). It is crucial to be able to perform such simulations as efficiently as possible. To obtain roughly 7 ps of useful information from an MD simulation, a wall time (on a single node) of 72h is required, depending on the system size. Therefore, the TIER1 infrastructure is crucial to achieve the goals of the project.

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: (typically, up to 2 A4 Arial 12).

As indicated in Figure 1, a variety of intermediates and reactions is going to be investigated in a CHA unit cell consisting of 110 atoms (unit cell size 13 Å x 13 Å x 14 Å). The total number of atoms in the system depends of course on the guest molecules included in the model. We distinguish 3 types of simulation jobs that will run on **2 nodes** at a time (based on a scaling test on BrENIAC for a representative test system, *vide infra*).

**MD** regular molecular dynamics runs to assess the mobility and adsorption behavior of the various intermediates. MD runs will be performed in the NPT ensemble and we need 100 ps of simulation time for each. As we chose our integration time step as 0.5 fs this means we need 200000 MD steps per simulation. We will perform such MD simulations for the following 7 systems:

- The empty zeolite
- Zeolite + methanol
- Zeolite + molecule **1**
- Zeolite + molecule **2**
- Zeolite + molecule **1** + benzene
- Zeolite + molecule **2** + propene
- Zeolite + molecule **3**

For each of these systems we will perform a simulation with a single Brønsted acid site (BAS) in the zeolite and 4 simulations with 2 BASs (we

will vary the relative positions of the two sites). This results in  $7 \times 5 = 35$  MD jobs.

**US** Umbrella sampling simulations are needed to study the reaction paths that connect the experimentally observed intermediates. The SEAr reaction is a multistep reaction and we estimate that 2 US simulations will be needed to capture the entire transformation. The Diels-Alder reaction consists of a cycloaddition step followed by 2 dehydrogenation steps, resulting in 3 US simulations. Finally, protonation and cyclization occur to result in the precursor for polyaromatics. During US simulations, 12 windows will be used that subdivide the entire simulation into 12 biased portions. Per window a sampling of 50 ps (100000 steps) is required and the computational expense for this is roughly identical to the one for regular MD simulations. A total amount of  $7 \times 12 = 84$  US simulations is needed.

**RM** To allow a one-on-one comparison with experimental results, the MD simulations will be used to calculate Raman vibrational spectra. The calculation of Raman spectra from molecular dynamics simulations requires the calculation of the polarizability tensor, which follows from the derivative of the dipole moment with respect to the electric field. To obtain this polarizability tensor at distinct time steps, we will calculate the dipole moment for four different situations, once without electric field and thrice with electric field applied along the three cartesian directions. Afterwards, the components of the polarizability tensor will be determined using a finite difference approach. To simulate the Raman spectra within the desired frequency range ( $0-4000 \text{ cm}^{-1}$ ), the polarizability tensor needs to be calculated every 2.5 fs (every 5 MD steps). The proposed methodology has proven to be very efficient and successful for the simulation of Raman spectra of nanoporous materials (A. E. J. Hoffman et al., *J. Phys. Chem. C* 122, 2734-2746, 2018). The computational expense is also similar to the one for MD simulations. We will calculate Raman spectra for the 7 systems listed above for a zeolite with 1 BAS and one with 2 BASs. This results in a total amount of  $7 \times 2 = 14$  RM simulations.

Table 1 summarizes the estimated number of tasks, node days and 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 job restarts (based on restart files generated by CP2K). As a result, the scratch volume per simulation is relatively low. We will also opt not to store the results of every single MD step as subsequent steps are statistically correlated. We will also compress completed runs as much as possible for archiving purposes.

**Table 1.** Estimated node days, memory requirements and number of files for each simulation type.

Computational task	Node day calculation				# CPU cores per task	Memory usage (GiB) / node per task	OpenMP / MPI / hybrid	Storage volume estimate	
	# of such tasks	Wall clock time (days) per task	# Tier-1 nodes per task	# total node days task				Tier-2 DATA/HOME volume (GiB) + number of files	Tier-1 SCRATCH volume (GiB) + number of files
MD	7 x 5 = 35	27	2	1890	56	25	MPI	122.5 GiB + 7700 files	17.5 GiB + 7700 files
US	7 x 12 = 84	14	2	2352	56	25	MPI	588 GiB + 10080 files	84 GiB + 10080 files
RM	7 x 2 = 14	22	2	616	56	25	MPI	39.2 GiB + 700 files	5.6 GiB + 700 files
<b>TOTAL</b>	<b>133</b>			<b>4858</b>				<b>749.7 GiB + 18480 files</b>	<b>107.1 GiB + 18480 files</b>

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)

Molecular dynamics simulations and umbrella sampling will be performed using the CP2K software package (<http://www.cp2k.org/>) with the PLUMED code (<http://www.plumed.org>), which are both freely available under the GPL license. The required version of the program (CP2K/3.0-intel-2016a-PLUMED-2.2.1-impi-5.1.2.150) is already available on Tier-1.

Provide the results of efficiency 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 efficiency plot such as Table 2 and Plot 1 below (log scale X-axis) (typically up to 3 A4 in Arial 12).

Important: the baseline configuration (first line in the table) should be the *minimal* configuration on which your computation could be carried out, where the number of cores for this baseline configuration could be 1, 2, ..., 28, 56, .... Please start the scaling of your code within one node (1, 2, ... cores), or explicitly mention why this is not possible. *Table and plot should be created w.r.t. the baseline configuration.*

Discuss the plot and table (explain any anomalies). Explain why you make a specific choice for a given number of nodes (in case of multi node jobs).

In case your tasks are single node jobs (HTC), describe how you will manage your jobs, e.g. will you make use of the Worker Framework<sup>1</sup> or something similar?

If the simulation contains a significant amount of I/O, then mention its impact on the calculations.

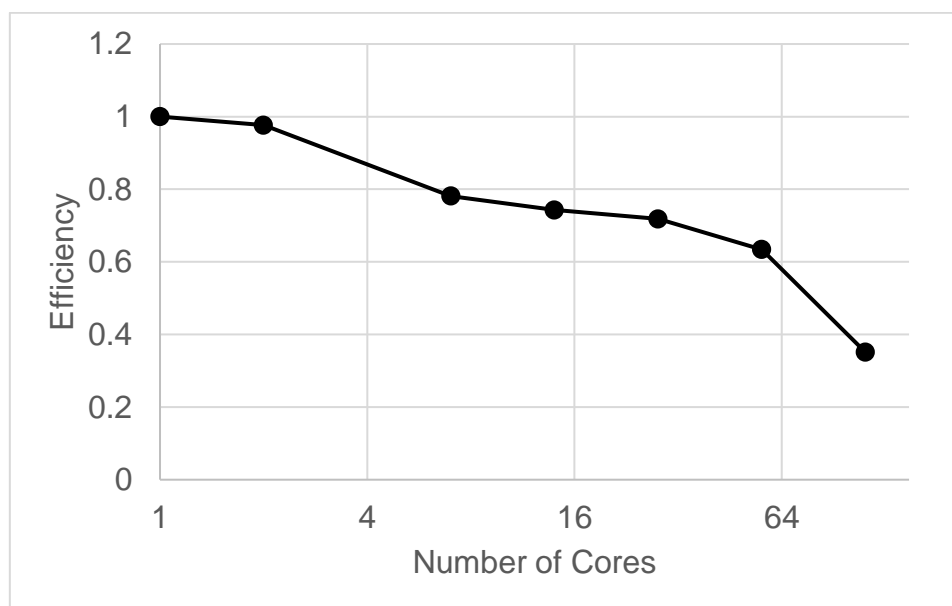
A scaling test for the CP2K software package (version CP2K/3.0-intel-2016a) on TIER1 (BrENIAC) has been conducted for a model system consisting of an adsorbed 1,3,5-heptatriene molecule (molecule **1** in Figure 1) in ZSM-5 containing one Brønsted acid site (i.e. 126 atoms per unit cell). Short NPT simulations (50 MD steps) have been performed on 1 - 112 cores of the Tier-1 cluster. It is expected that the scaling test holds for all systems and all types of calculations performed in this project as they are all based on MD. The results of the scaling test are summarized in Table 2. In Figure 2 the efficiency normalized to a simulation on a single core is plotted for the test MD simulations. Up to 56 cores (2 nodes) the efficiency is acceptable, however when using more cores it dramatically drops below 65%. This graph indicates motivates the choice to use a maximum of 2 nodes per job as this will be the optimal trade-off between efficiency and throughput time. Note that the efficiency drops dramatically between 1 and 7 cores. In that view, normalizing with respect to the 7 core test would yield slightly better efficiencies.

---

<sup>1</sup> <https://www.vscentrum.be/cluster-doc/running-jobs/worker-framework>

**Table 2.** Summary of the scaling test for a 50 MD steps simulation on 1,3,5-heptatriene in ZSM-5.

# nodes	# cores	Wall clock time (s)	Speedup	Efficiency
1	1	20632	1	1
1	2	10563	1.95	0.98
1	7	3773	5.47	0.78
1	14	1984	10.40	0.74
1	28	1025	20.11	0.72
2	56	581	35.50	0.63
4	112	524	39.34	0.35



**Figure 2.** Scaling efficiency normalized to 1 core for an MD simulation on 1,3,5-heptatriene in ZSM-5 using the CP2K software package.

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, ...)?

In case the disk storage or the number of files that is applied for exceed the defaults of 2 TiB and 200.000 files (cf. regulations §4, (5)), describe how the transfer of files to/from Tier-1 will be managed, if data reduction and/or compression of files will be performed, and if you can provide information about IOPS.

We estimate that the proposed simulations can be completed over a time period of approximately 8 months. Similar usage on a monthly basis is expected. For each of the 7 systems the 3 job types (MD, US, RM) will be performed in a sequential fashion. Per job type, a large number of simulations will run in parallel (for each different system). The estimated timeframe takes into account the actual runtime of the simulations, as well as intermediate data analysis tasks and frequent job restarts.

Since each simulation (MD, US and RM) cannot be completed within the wall time of 72 hours, several job restarts are required based on the automatically generated CP2K restart files. Restarting cannot be done automatically as intermediate result analysis is required. For each run, a maximum of 500MB of scratch space is required. After 72 hours, all data is automatically transferred to the Tier2 scratch storage and the Tier1 scratch space is cleared again before the job is restarted. Consequently, scratch volume per simulation is relatively low and the load on the Tier1 scratch will be minimal. Moreover, the data transfer is thus spread equally over the duration of the project. The data transfer of the input files at the start of the job is negligible (approximately 300KB). The amount of files generated per run of 72h varies between 25 and 100, resulting in a total amount of files in the project of 18480. We will also opt not to store the results of every single MD step as subsequent steps are statistically correlated. We will also compress completed runs as much as possible for archiving purposes

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.

List of granted research projects in the last 2 years on the Flemish Tier1 system:

- Dynamical first principle modelling of ethene oligomerization in Ni-SSZ-24 (project 2016-09, K. De Wispelaere, 3232 node days, 01/03/2016 - 30/06/2016)
- Dynamical first principle modelling of zeolite dealumination in H-SSZ-13 (project 2016-34, K. De Wispelaere, 3624 node days, 01/07/2016 - 31/10/2016)
- DFT study of reaction paths in zeolite-catalyzed 2-hexene cracking (project 2016-35, P. Cnudde, 4536 node days, 15/08/2016 - 31/12/2016)
- Benchmark study of ab initio molecular dynamics simulations for the methylation of HMB (project 2016-64, S. Bailleul, 4812 node days, 01/11/2016 - 30/04/2017)
- Mechanistic investigation of the early stages of the methanol-to-hydrocarbons conversion (project 2017-42, K. De Wispelaere, 3640 node days, 03/07/2017 - 03/01/2018)
- Ab initio study on the stability of cracking intermediates (project 2017-39, P. Cnudde, 4920 node days, 03/07/2017 - 03/01/2018)
- Enhanced sampling study of the methylation of ethene, propene and trans-2-butene (project 2017-66, S. Bailleul, 3300 node days, 01/11/2017 - 30/04/2018)
- Modeling the influence of zeolite acidity on MTO methylation reactions (project 2018-12, P. Cnudde, 4960 node days, 05/03/2018 - 05/11/2018)
- Investigating the low-frequency vibrational fingerprint of flexible metal-organic frameworks (project 2018-18, A. Hoffman, 1062 node days, 05/03/2018 - 05/11/2018)

List of scientific output within the framework of granted computing time of the last 2 years:

- <http://dx.doi.org/10.1016/j.cattod.2018.02.042>
- <http://dx.doi.org/10.1021/acscatal.7b01643>
- <http://dx.doi.org/10.1016/j.jcat.2017.03.007>
- <http://dx.doi.org/10.1016/j.jcat.2016.11.010>
- <http://dx.doi.org/10.1002/cctc.201600650>
- <http://dx.doi.org/10.1016/j.jcat.2016.05.018>

- <http://dx.doi.org/10.1039/C5CY02073E>
- <http://dx.doi.org/10.1021/acscatal.5b02139>
- <http://dx.doi.org/10.1021/acs.jpcc.7b11031>

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

De heer Kristof De Wispelaere  
Bargiekaai 27/302  
9000 GENT

**uw kenmerk**

**ons kenmerk**  
69566  
12N8617N

**contact**

wt@fwo.be  
02 550 15 88

**datum**

22 juni 2016

**betreft Toekenning postdoctoraal onderzoeker 2016-2019**

Geachte heer,

De raad van bestuur heeft u vandaag aangesteld als postdoctoraal onderzoeker van het FWO met ingang van 1 oktober 2016 tot en met 30 september 2019. Het reglement vindt u als bijlage.

Graag willen wij u feliciteren met deze aanstelling.

Voor de verdere administratieve opvolging zal het FWO u nog vóór 15 juli 2016 een arbeids-overeenkomst bezorgen.

Onderzoekers dienen, voor wat betreft de ethische problemen inzake de betrokkenheid van proefpersonen en/of proefdieren, het advies van de lokale ethische commissie in te winnen.

Voor bijkomende inlichtingen kan u via bovenvermeld e-mailadres steeds terecht bij de dossierbeheerder van uw wetenschapsgebied. Vóór 1 oktober 2016 zal u nog een gedetailleerde feedbackbrief ontvangen.

Tenslotte vragen wij u met aandring om op uw publicaties steeds uw titel "postdoctoraal onderzoeker van het FWO" te vermelden.

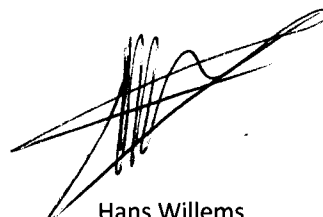
Wij wensen u veel succes toe in uw verdere onderzoeksloopbaan.

Met vriendelijke groeten,

Waarnemend secretariaat-generaal,



Danny Huysmans  
Directeur Intern Beheer



Hans Willems  
Directeur Steun aan Onderzoekers

**bijlagen: 1**