Enclosure 1b. Category 1 Application form 2015 – 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 Dutch on the website https://vscentrum.be/.

Title of the application:

pK_a calculation of Bromothymol Blue and modified derivatives

Name and first name of the applicant:

De Meyer Thierry

Institution:

Ghent University

Research group / department:

Center for Molecular Modeling

Title / position:

PhD Student (BOF)

email address:

thierry.demeyer@ugent.be

Total computing time that is needed, in node days:

3180

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

46,5

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

"The effect of dye-polymer interactions on the halochromic properties of azo dyes via a combined experimental and theoretical approach."

(BOF PhD grant - promotor: Prof. dr. ir. V. Van Speybroeck; copromotor: Prof. dr. ir. Karen De Clerck)

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):

This project is part of a collaboration between experimental and computational partners with the common goal of developing new sensor materials. More specifically, pH-sensitive (halochromic) dyes are applied onto polymeric materials, resulting in pH-sensitive polymers. One possible application is in smart wound bandages, such a bandage can warn the wearer of possible infection. Other applications, such as geotextiles, are also possible. In recent work, these dye molecules are modified to include a reactive group, which can later form a covalent bond with the parent textile material. This covalent bond provides the best attachment to the textile matrix.

The dye molecules under study undergo a colour change upon (de)protonation. Therefore, their most important parameter is the pK_a, which defines the pH-range of the colour change. The above mentioned modification however influences many properties of the dye molecule, including the pK_a. The azo dye Methyl Red was recently modified experimentally, causing the pK_a to drop from 5 to 2, making it unusable for wound bandages where the ideal pK_a would be around 6.5. The modification of molecules can be a long process and frustrating if the final result is unusable for the intended application. This can be avoided by theoretically predicting pK_a's, which is the goal of this project.

Molecular modeling is able to predict this change in pK_a, which can allow for screening of possible modifications. In a previous TIER1-project, advanced MD methods (insertion/deletion, I/D, and restrained MD, rMD) were tested on phenol derivatives and one dye molecule, phenol red, with excellent results compared to experiment. By the expertise gathered in the previous project, we are now ready to apply this methodology on Bromothymol Blue and two modified variants (up to 90 atoms). It has recently been shown experimentally that this modification has

a large effect on the pK_a and the recently tested methods will now be evaluated for the dye and its modified forms. This will illustrate the usefulness of predicting pK_a's of modified dye molecules.

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

The development of novel sensor materials, such as pH-sensitive textile materials, is a new and challenging research field. To develop such sensors, pH-sensitive dyes are applied onto textile materials. These dye molecules are chemically altered to allow for better interaction with the parent textile material. This modification however, also influences the pH-sensitivity itself, which is often difficult to predict. This project aims to develop a method using large-scale ab initio molecular dynamic simulations (MD) to accurately predict the pH-sensitive properties of these modified dyes, allowing for more thoroughly considered choices for dye modification in further experiments.

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 4 "EasyChair proposals submission procedure").

BOF PhD grant (grant code BOF14/DOC_V/343), see attachment.

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

Prof. dr. ir. Veronique Van Speybroeck

6. 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 of using HPC resources in the past (Tier0/Tier1/Tier2 infrastructure in Belgium and abroad)

Thierry De Meyer
Ghent University, Center for Molecular Modeling
BOF PhD fellow

4½ years of experience with TIER2 at Ghent University
½ year of experience with TIER1 at Ghent University

7. Explain why this project needs to run on a Tier1 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).

Performing all simulations proposed in this project will require about 3180 node days in computing time. To complete this project within a reasonable timeframe, the ability to run several multi-node jobs simultaneously on the TIER1 infrastructure will be indispensable. The long simulation times required for attaining statistically relevant data has long prohibited the application of molecular dynamics methods to the study of solvated dye systems, but nowadays the application of advanced molecular dynamics simulations for the accurate calculation of pK_a values is gaining interest internationally. Moreover, in the previous TIER1-project, these methods haven been shown to be successful in the calculation of pKa values for various systems, which is confirmed by literature. (A. Tummanapelli et al. J. Phys. Chem. B 118 (2014) 13651; B. Ensing et al. Acc. Chem. Res. 39 (2) (2006) 73) Also, the Center for Molecular Modeling has obtained a lot of experience employing advanced MD techniques. (S.L.C. Moors et al, *ACS Catal.*, 2013, 2556-2567; J. Van der Mynsbrugge, et al., *ChemCatChem*, 2014, 1906-1918). Access to the highly efficient TIER1 infrastructure with its fast nodes and inter-nodal communication will be indispensable to enable us to make further high-impact contributions to this highly competitive field.

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 Tier2 DATA/HOME partitions and the Tier1 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 Tier1 (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).

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

The computational requirements for this study have been carefully considered based on simulations previously done on the TIER1 infrastructure at Ghent University.

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

- In total, 3 dye molecules will be considered (including their modified variants). An initial **MD simulation** in the NVT ensemble will be performed at 300K, with a duration of 5 ps to allow the system to fully equilibrate. Given the total size of the system (one dye molecule, the biggest one about 90 atoms, and about 180 water molecules, 540 atoms) is about 630 atoms, each of these simulations requires about **10 node days**.

- For the I/D method, 3 simulations need to be run per dye molecule, each with a double cost of a normal MD simulation (due to the nature of this method). Depending on the characteristics of each system, about 45 ps (60000 steps) will be sampled. Per dye molecule about **600 node days** will therefore be required.

- For the rMD calculations, for each dye, 10 snapshots are taken during the deprotonation reaction. The system is kept fixed at each of these 10 points (10 simulations) during at least 20 ps (27000 steps). Per dye molecule **450 node days** will be required.

For both types of simulations, multiple nodes with all available memory (64GiB per node) will be used.

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

Computation-al task	# of such tasks	Wall clock time per task	# of nodes per task	Node days per simulation	Total node days	Total core days
MD (eq)	3	2 days	8	10	30	480
I/D	3	37 days	16	600	1800	28800
rMD	3x10	6 days	8	450	1350	21600
					3180	50880

Table 2 summarizes the requirements for scratch space and long-term storage (data) for each simulation type. The type of data stored is almost the same.

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]
MD (eq)	0.5	0.5	3*0.5 = 1.5	3*0.5 = 1.5
I/D	5	5	3*5 = 15	3*5 = 15
rMD	10*1 = 10	10	3*10 = 30	3*10 = 30

	46,5	46,5

Scaling of the CP2K software on the TIER1 has been thoroughly tested by user Andy Van Yperen–De Deyne on a water box, which is very similar to the system under study in this project and uses the same modules in CP2K. Similar tests have been performed for short MD runs on the Bromothymol Blue dye (the largest modified form) and the results are shown in Figure 1.

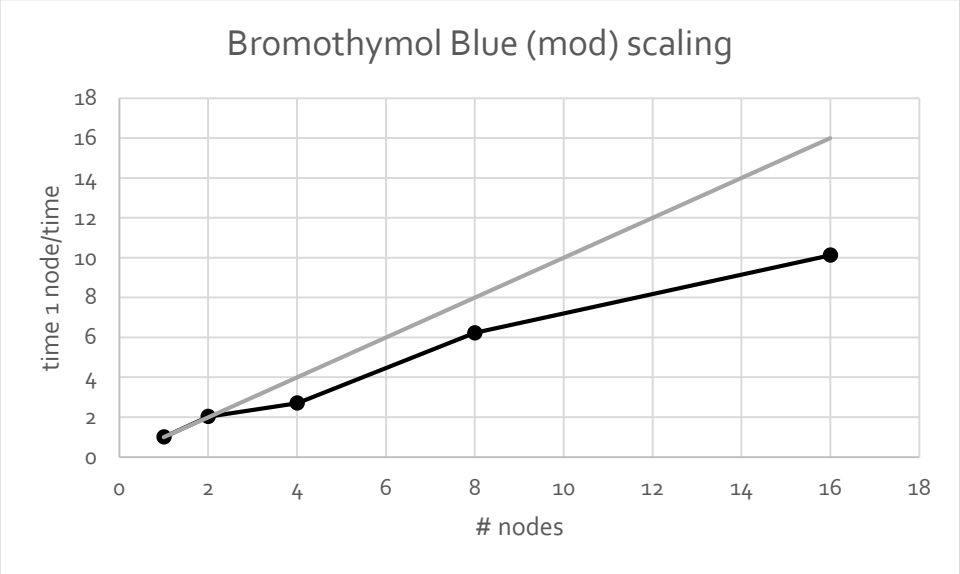

Figure 1. Efficiency vs. number of nodes for various the largest system with the GPW code of CP2K.

Figure 1 shows excellent scaling (more than 75%) up to 8 nodes, which will be used for these calculations.

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 Tier1 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 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)

Provide the results of scaling tests that were conducted with this software, preferably on Tier1 (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).

Molecular dynamics simulations will be performed using the CP2K software package (http://www.cp2k.org/), which is freely available under the GPL license. The required version of the programme (CP2K/20150904-intel-2015a-PLUMED-2.1.3) is already available on TIER1 and has already been evaluated for these systems.

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, …)?

Because intermediate analysis is required and the simulations themselves are very time consuming, a **period of 6 months** is requested for the completion of all the simulations, preferably **starting as soon as possible**. Because initial simulations have already been run, a continuous usage throughout the period is expected.

11. List the granted computing time allocations to the promoter(s) of this research project, on the Flemish Tier1 system, as well as other Tier1 and Tier0 systems. Also, describe the scientific output obtained within the framework of computing time that was granted during the past two years on the Flemish Tier1 or on other Tier1 or Tier0 supercomputers. DOI links are sufficient.

Below is a list of approved TIER1 projects from within our group, limited to those relevant to this proposal:

Accurate pKa calculations of pH-sensitive dye molecules
Period: 01/08/2015 to 31/01/2016
Node days: 3780
Users: T. De Meyer, O. De Vos, H. Goossens

Exploring the kinetics and selectivity of butene cracking using molecular dynamics simulations
Period: 01/01/2015 tot 30/06/2015
Node days: 4864 days
Users: J. Van der Mynsbrugge, P. Cnudde, K. De Wispelaere

Dynamics of poly(2-oxazoline)s
Period: 01/11/2014 tot 30/04/2015
Node days: 2900 days
Users: D. Hertsen, H. Goossens

Vibrational spectra of Mo-exchanged zeolite materials
Period: 13/11/2014 tot 30/04/2015
Node days: 3700 days
Users: K. Hemelsoet, A. Van Yperen-De Deyne

Dynamical kinetic study of zeolite-catalyzed reactions
Period: 07/07/2014 tot 31/12/2014
Node days: 4371 days
Users: K. De Wispelaere, J. Van der Mynsbrugge, S.L. Moors, V. Van Speybroeck

Molecular dynamics study of pentene in H-ZSM-5: toward a better estimate of adsorption enthalpies
Period: 07/03/2014
Node days: 1824 days
Users: J. Van der Mynsbrugge

List of all publications of our group where TIER1 calculations were used (DOI):

http://dx.doi.org/10.1021/cs400706e
http://dx.doi.org/10.1039/C3CP54132K
http://dx.doi.org/10.1063/1.4869937
http://dx.doi.org/10.1002/cctc.201402146
http://dx.doi.org/10.1039/C4MH00127C
http://dx.doi.org/10.3762/bjnano.5.184
http://dx.doi.org/10.1002/chem.201500473

+ “pKa calculations based on advanced MD simulations”, which is in preparation and based on the results obtained in the previous TIER1 project.

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 Hercules Foundation / Flemish Supercomputer Center.

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: rosette.vandenbroucke@vub.ac.be
For the other institutions: marc.luwel@herculesstichting.be

UNIVERSITEIT GENT

DIRECTIE ONDERZOEKSAANGELEGENHEDEN

Afdeling **Onderzoekscoördinatie**

Prof. Veronique Van Speybroeck
Vakgroep Toegepaste fysica
Technologiepark Zwijnaarde 903
9052 Zwijnaarde

Uw kenmerk	**Ons kenmerk**	**Datum**
	DOZA/IL/DDC/CWB/DL/186b-2012	22/08/2012
Contactpersoon	**E-mail**	**Contact**
David Lombart	BOF@UGent.be	tel. 09 264 31 23 fax 09 264 35 83

Betreft : BOF-oproep 2012 – Doctoraatsmandaten

Geachte collega,

Ik heb het genoegen u mee te delen dat na advies van de Onderzoeksraad dd. 21 augustus 2012 aan Thierry De Meyer een doctoraatsmandaat met als titel van het doctoraatsonderzoek *"The effect of dye-polymer interactions on the halochromic properties of azo dyes via a combined experimental and theoretical approach"* wordt toegekend. Dit mandaat heeft een duur van 24 maanden en start ten vroegste op 1 oktober 2012[1]. Aan dit mandaat wordt het dossiernummer 01D34312 toegekend.

De mandaathouder is verplicht zich in te schrijven voor het doctoraat bij de afdeling Studentenadministratie en studieprogramma's.
Indien de betrokken faculteit het doctoraatsonderwerp nog niet heeft goedgekeurd, dient de mandaathouder zich bij de Studentenadministratie aan te melden met het 'Voorstel tot Aanwerving' (zoals ingesloten bij de toekenningsbrief aan de mandaathouder) om zich voorlopig in te schrijven als kandidaat-doctoraatsbursaal.
Deze inschrijving voor het doctoraat of de voorlopige inschrijving als kandidaat-doctoraatsbursaal is een noodzakelijke stap om het dossier bij de directie Personeel en Organisatie in orde te brengen.

De afdeling Onderzoekscoördinatie heeft de directie Personeel en Organisatie (DPO) verzocht de aanwervingprocedure op te starten (middels het document 'Voorstel tot Aanwerving'). Op basis van de gegevens vermeld in het aanvraagdossier komt Thierry De Meyer in aanmerking voor een Dehoussebeurs. Mag ik u vragen om, indien het gaat om een eerste aanstelling aan de UGent, er op toe te zien dat zo snel mogelijk de overige vereiste documenten worden bezorgd aan de directie Personeel en Organisatie. Een overzicht van de in te dienen documenten is beschikbaar op: https://www.ugent.be/nl/werken/aanwerving/formulieren/wp/NW_document_db.htm/.

[1] Beslissing genomen bij bevoegdheidsdelegatie verleend aan de Rector door het bestuurscollege van 05/07/2012

Gekoppeld aan dit doctoraatsmandaat worden er ten behoeve van de promotor € 3.720 werkingsmiddelen per jaar toegekend (zijnde € 310/maand) op het WBS-element B/13193/01. Bij voortijdige stopzetting van het BOF doctoraatsmandaat wordt de werkingstoelage verhoudingsgewijs verminderd.

Het mandaat van Thierry De Meyer komt in aanmerking voor een verlenging. De afdeling Onderzoekscoördinatie zal de mandaathouder tijdig uitnodigen een door de promotor ondertekend wetenschappelijk vorderingsverslag in te dienen. Op basis van dit verslag adviseert de Onderzoeksraad het Bestuurscollege over een eventuele verlenging van het mandaat.

Indien u vragen heeft betreffende bovenstaande toekenning kan u contact opnemen met de afdeling Onderzoekscoördinatie (dhr. David Lombart, e-mail: BOF@UGent.be of tel.: 09/264.31.23).

Met collegiale groeten,

Prof. Paul Van Cauwenberge
Rector

c.c.:
- Prof. Karen De Clerck, Textielkunde, Technologiepark Zwijnaarde 907, 9052 Zwijnaarde

Directie Onderzoeksaangelegenheden – Afdeling Onderzoekscoördinatie
Sint-Pietersnieuwstraat 25, B-9000 Gent
www.UGent.be