

# Molecular dynamics simulation of vapour-liquid nucleation of water with constant energy

Michal Duška<sup>1,a</sup>, Tomáš Němec<sup>1</sup>, Jan Hrubý<sup>1</sup>, Václav Vinš<sup>1</sup> and Barbora Planková<sup>1</sup>

<sup>1</sup>Institute of Thermomechanics AS CR, v.v.i, Dolejškova 5, Prague 8, CZ-18200, Czech Republic

**Abstract.** The paper describes molecular dynamics study of nucleation of water in NVE ensemble. The numerical simulation was performed with the DL\_POLY. The metastable steam consisting of 10976 water molecules with TIP4P/2005 potential was driven on the desired energy level by a simulation at constant temperature, and then the nucleation at constant energy was studied for several tens of nanoseconds, which was sufficient for clusters to evolve at hundred molecules size. The results were compared with the previously published results and the classical nucleation theory predictions.

## 1 Introduction

Homogeneous nucleation is essentially a molecular-scale process therefore prone to be studied by molecular simulation. There are two main limitations for the method dealing with this transition process with an initial energy barrier. Macro scale time is needed, and in some cases a macro scale matter is needed as well, for the nucleation to have a fair chance to occur. This drawback can be overcome by a significant oversaturation of the micro scale sample and by fixing the temperature using an NVT (i.e., constant number of molecules, volume and temperature) simulation [1, 2]. In the present paper we tried to challenge this good practice and introduced an NVE (i.e., constant number of molecules, volume and energy) simulation to the realm of nucleation. The main motivation for this work is to move the mathematical experiment closer to reality with its self regulation by the release of the latent heat as the clusters grow bigger, and to see how it influences the growth rate.

Molecular dynamics (MD) is a quite new theoretical tool for our group. The presented paper illustrates the ability to enquire the necessary tools and develop the needed postprocessing capabilities for an effective study of the nucleation processes with MD and also the feasibility of the nucleation experiment in NVE ensemble. Although only a single nucleation experiment is described in the paper (which fully employed 64 CPU for roughly a month without postprocessing) the advantage of the NVE ensemble will be clearly visible.

Before starting the nucleation process a proper setting of the initial condition i.e. the density and the energy was essential. We used a setting which was recommended by Professor Jiří Kolafa through a personal communication [3].

<sup>a</sup> Corresponding author: duska@it.cas.cz

In this paper, first in Sec. 2 the theoretical apparatus of the classical nucleation theory is reviewed shortly to introduce the common definitions used within the nucleation theory. In Sec. 3, the details of our MD numerical experiment are given. And in Sec. 3, the nucleation process observed in the MD experiment is analyzed in detail and the results are compared with the predictions of the nucleation theory.

## 2 Classical nucleation theory

The classical nucleation theory (CNT) is used in this work as a theoretical complement to the nucleation parameters observed in the MD simulation. The CNT solves a non-equilibrium statistical-thermodynamic problem of the cluster size evolution in a thermodynamic system at a given temperature  $T$  and pressure  $p$ . In our simple case of a one-component system, the CNT equations can be solved analytically, resulting in the estimation of the steady-state nucleation parameters which are calculated as follows. The size of the critical cluster is given in terms of the critical radius  $r_c$  (m)

$$r_c = \frac{2\sigma v_l}{kT \ln \frac{p}{p_{sat}}}, \quad (1)$$

where  $\sigma$  is the surface tension,  $\text{N m}^{-1}$ ;  $v_l$  is the molecular volume of the liquid phase,  $\text{m}^3$ ;  $k$  is the Boltzmann constant,  $\text{J K}^{-1}$ ; and  $p_{sat}$  is the saturation pressure, Pa. The number of molecules in the critical cluster  $n_c$  is given as

$$n_c = \frac{4\pi r_c^3}{3v_l}. \quad (2)$$

The formation work  $W$  (J) of the critical cluster is calculated as

$$W = \frac{4}{3} \pi \sigma r_c^2. \quad (3)$$

Finally, the nucleation work  $J$  ( $\text{m}^{-3}\text{s}^{-1}$ ) expression takes the form

$$J = \sqrt{\frac{2\sigma}{\pi m}} v_l \frac{p^2}{(kT)^2} \exp\left(-\frac{W}{kT}\right), \quad (4)$$

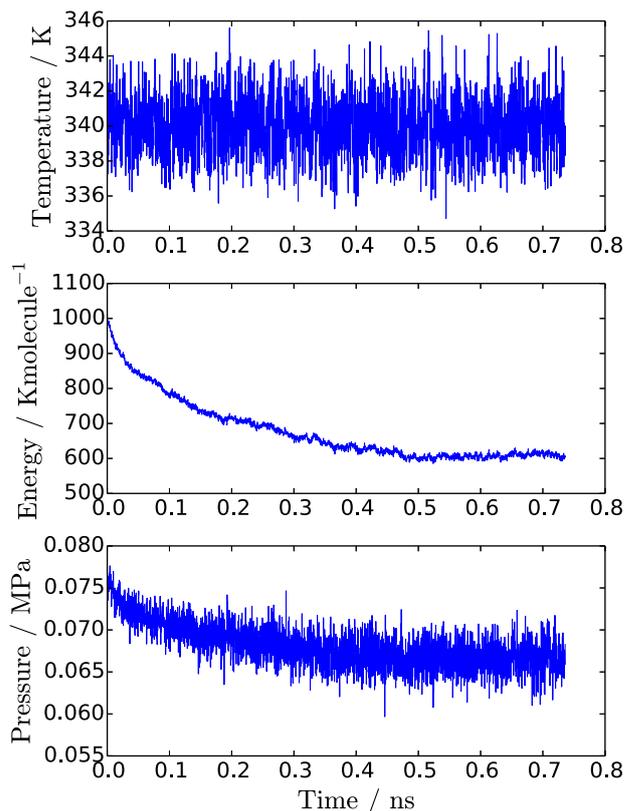
where  $m$  (kg) is the molecular mass. To evaluate Eqs. (1) to (4) several thermophysical properties of the system under investigation are required. The properties for TIP4P/2005 water are the liquid density (taken at the pressure of 0.1 MPa) [4], the saturation pressure [7] and the surface tension [8]. The respective thermophysical properties of real water are taken from the IAPWS formulation [9].

### 3 Numerical experiment

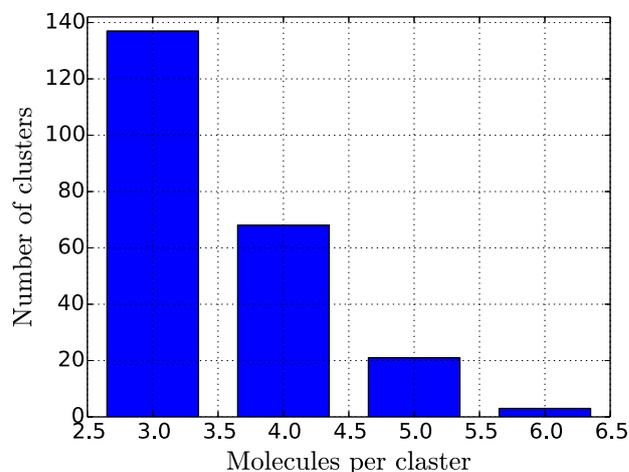
The intermolecular potential TIP4P/2005 [4] was chosen to be used in molecular dynamics experiment carried out with the DL\_POLY 4.05 numerical computational tool [5, 6]. Inhouse tools were developed in Python for the post processing and analysis.

The number of molecules was set to be 10 976 in order to distribute the water molecules evenly in the form of a cubic mesh within a cubic cell with one side length 869.446 Å (which is equivalent to a density of 0.5  $\text{kg m}^{-3}$ ) as a starting point to prepare the vapour. At this stage the vapour was equilibrated in an NVT ensemble at 1000 K for about 150 000 steps with the time step 1 fs. After that the system temperature was changed to 340 K. The temperature was controlled by the Andersen thermostat with the thermostat relaxation time of 0.5 ps and the softness of the thermostat of 0.5. The development of energy, temperature and pressure during this energy adjustment stage can be seen in figure 1.

As the starting point for the nucleation stage simulation the positions, velocities and forces of molecules were selected for the time step with an energy close to the desired value of 705  $\text{K molecule}^{-1}$  which was at 207 ps. At this time, the size of the largest cluster of the vapour is 6 molecules, as can be seen from figure 2. The molecules were marked to form a cluster if the distance between the hydrogen atom of one molecule and the oxygen atom of the other molecule was smaller than 2.5 Å or equal. It can be seen clearly from the figure 2 that largest clusters left in the vapour after the energy adjustment stage are well below the expected critical cluster size, around 20 molecules, so the metastable steam could appear in the next stage.



**Figure 1.** Development of thermodynamic properties during energy adjustment stage.

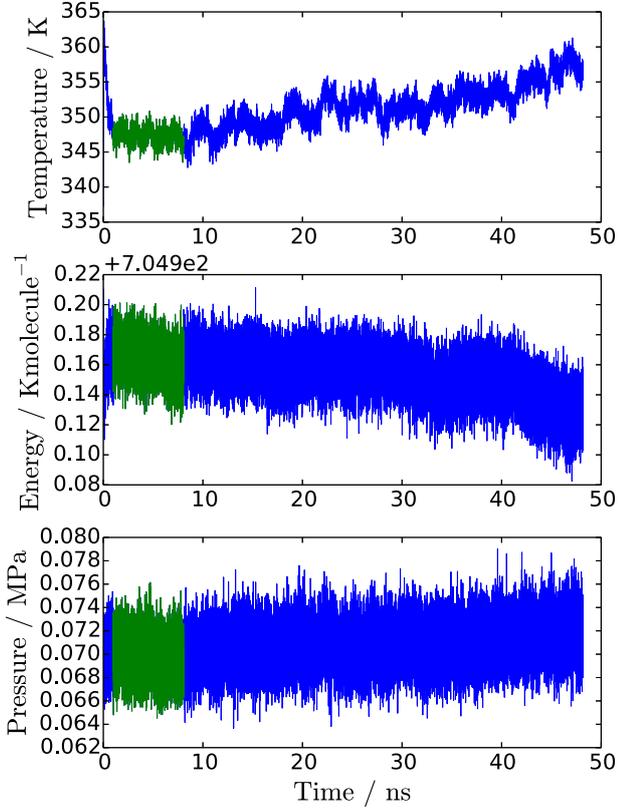


**Figure 2.** Histogram of trimmers and bigger clusters at 207 ps time step of energy adjustment stage.

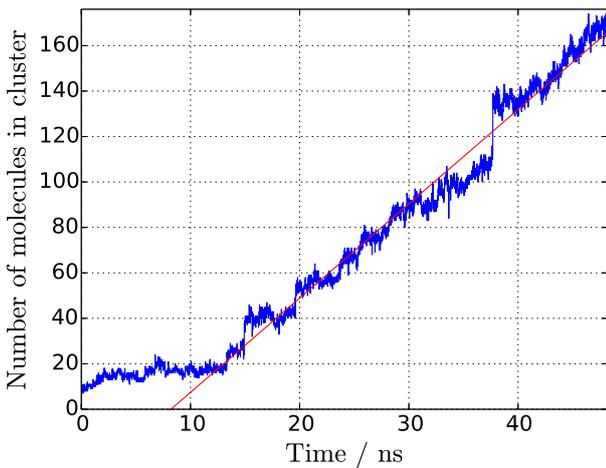
The nucleation stage was studied in an NVE ensemble with the time step of 2 fs. A cutoff of 50 Å was used for all interactions. The transition from the NVT to the NVE ensemble had a quite large influence on the temperature but the energy proved to be stable as can be seen from figure 3.

Figure 4 presents the size of the largest cluster evolving in the nucleation stage. The figure clearly shows three different phases of the nucleation process. The first phase is the temperature adjustment after the change of ensembles, which ended after around 1 ns. It was followed by induction phase, which ended at around 8 ns before the nucleation phase occurs. The precise beginning

of the nucleation phase was determined as an intersection of the linear regression curve of the cluster size higher than the expected critical cluster size and the x-axis. The critical cluster size can be spotted in figure 4; crossing cluster size around 20 molecules seems to be the turning point from which it is hard to downsize a cluster.



**Figure 3.** Development of thermodynamic properties during the nucleation stage. The color-coded data region represents the induction phase.

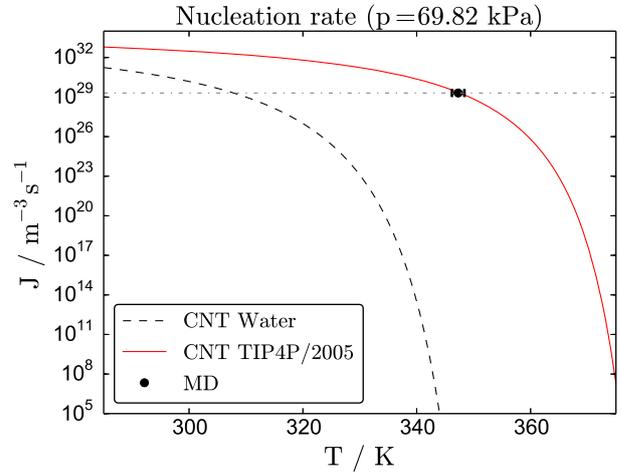


**Figure 4.** Number of molecules in the largest cluster during the nucleation stage. The solid line is a linear regression curve of the cluster size higher than expected critical cluster size.

## 4 Discussion

### 4.1 CNT nucleation rate for TIP4P/2005 water

The CNT nucleation rates of the real water and the TIP4P/2005 water calculated from Eq. (4) are shown in figure 5. The nucleation rates are plotted as functions of temperature for the pressure of 69.82 KPa, which is the mean value of pressure during the induction period, i.e. before the nucleation process started. A dramatic discrepancy can be observed between the two cases. The nucleation in TIP4P/2005 water is shifted to more than 30 °C higher temperatures than nucleation in real water. Therefore, the nucleation events simulated with the TIP4P/2005 water model can hardly be used as a complement to the nucleation experiments with real water.



**Figure 5.** Nucleation rate of real water (dashed line) and TIP4P/2005 water (full line) according to the CNT calculated as a function of temperature and compared to the nucleation rate observed in our MD experiment (red circle).

### 4.2 Kinetically defined critical cluster size

From the molecular kinetic point of view the critical cluster is characterized by an equal average loss and gain of molecules. To identify the critical cluster, an average size change was introduced as

$$\langle k \rangle (n) = \langle k \rangle_+ (n) + \langle k \rangle_- (n), \quad (5)$$

where  $\langle k \rangle_+ (n)$  is the growth rate for a cluster of  $n$  molecules growing by addition of monomers or any of the  $k$  sized clusters, and  $\langle k \rangle_- (n)$  is the decomposition rate

$$\langle k \rangle_+ (n) = \sum_{k=1}^{\infty} k \beta_{n,n+k}, \quad (6)$$

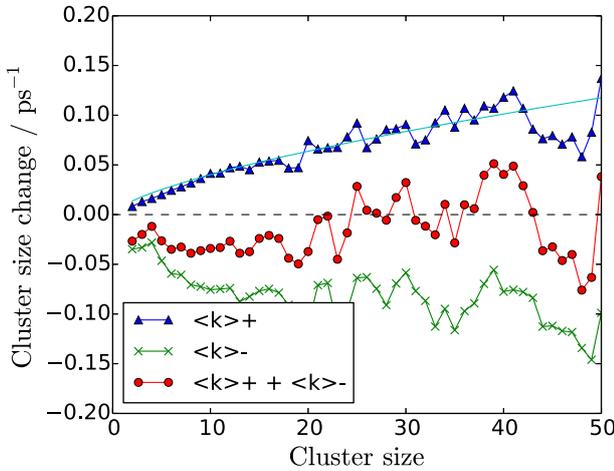
$$\langle k \rangle_-(n) = \sum_{k=-\infty}^{-1} k \beta_{n,n+k}. \quad (7)$$

The transition probability that a cluster of size  $n$  changes to the  $n+k$  size is

$$\beta_{n,n+k} = \frac{\langle M_{n,n+k}(t) \rangle_t}{\sum_{k=-\infty}^{\infty} \langle M_{n,n+k}(t) \rangle_t}, \quad (8)$$

where  $M_{n,n+k}(t)$  is the number of events when a cluster of  $n$  molecules changes its size to  $n+k$  molecules between the time  $t$  and  $t+\Delta t$ , for which the information about the molecules location is available, i.e. 5 ns. The operator  $\langle \rangle_t$  denotes the time averaging of a value during the nucleation phase.

The critical cluster size can be estimated from figure 6; the critical cluster appears when the average size change crosses the zero line before starting to oscillate around the zero line. Regular crossing of the zero line occurs for the cluster size of 25, but the clusters of 21 and 22 molecules are also very close to the zero line. Therefore, the critical cluster size based on the molecular kinetics can be located approximately between 21 and 25 molecules.



**Figure 6.** Average size change - circle, growth rate - triangle, and decomposition rate - cross. The solid line is the fitted curve of Eq. (9).

### 4.3 Comparison of growth rate between MD experiment and theory

The difference between the growth rate determined from our MD experiment and predicted by CNT is an important measure to compare the methods. We assume that both growth rates are proportional with respect to a constant parameter  $\xi$

$$\langle k \rangle_+(n) = \xi \sqrt{\frac{kT}{2\pi m}} \rho_v A_n. \quad (9)$$

In Eq. (9),  $\rho_v$  is the number density of vapour defined as a time average of the monomer density during the nucleation phase and  $A_n$  is the effective surface area of an  $n$ -mer given as

$$A_n = \left( \frac{6\sqrt{\pi}}{\rho_l} n \right)^{2/3}, \quad (10)$$

where  $\rho_l$  is the number density of the bulk liquid water.

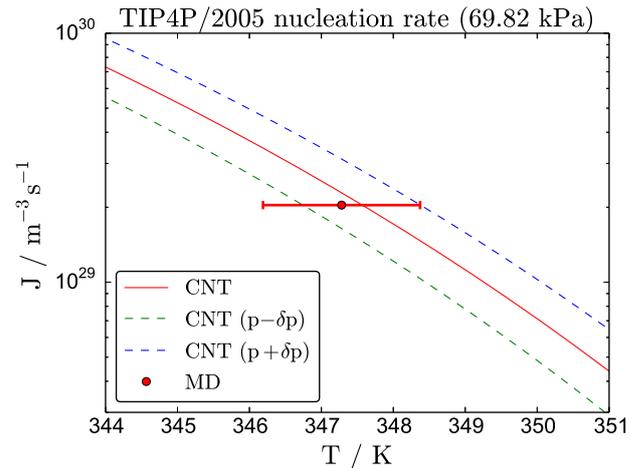
By fitting Eq. (9) to the MD data shown in figure 6 (the fit is presented as a light blue solid line) up to the cluster size of 30 molecules, we found that the MD growth rate is higher than the theoretical growth rate used by the CNT by the factor  $\xi = 8.13$ .

### 4.4 Nucleation rate estimated from MD experiment

During the nucleation phase four supercritical clusters emerged, but at the time of around 38 ns two supercritical clusters merged together. This can be seen in figure 4 as a sudden increase of the maximum cluster size. Therefore we included only limited period of time from  $t_0 = 8.06$  ns to  $t_1 = 38$  ns as a nucleation period with  $N = 4$  nucleation events. The volume of our simulation is  $V = (86.9)^3 \text{ nm}^3$ . Therefore, the average nucleation rate observed in our simulation is

$$J_{MD} = \frac{N}{V(t_1 - t_2)} = 2.04 \times 10^{29} \text{ m}^{-3} \text{ s}^{-1}. \quad (11)$$

The MD nucleation rate (11) is plotted in figure 7, and a good agreement between  $J_{MD}$  and the CNT prediction of the TIP4P/2005 nucleation rate can be observed.

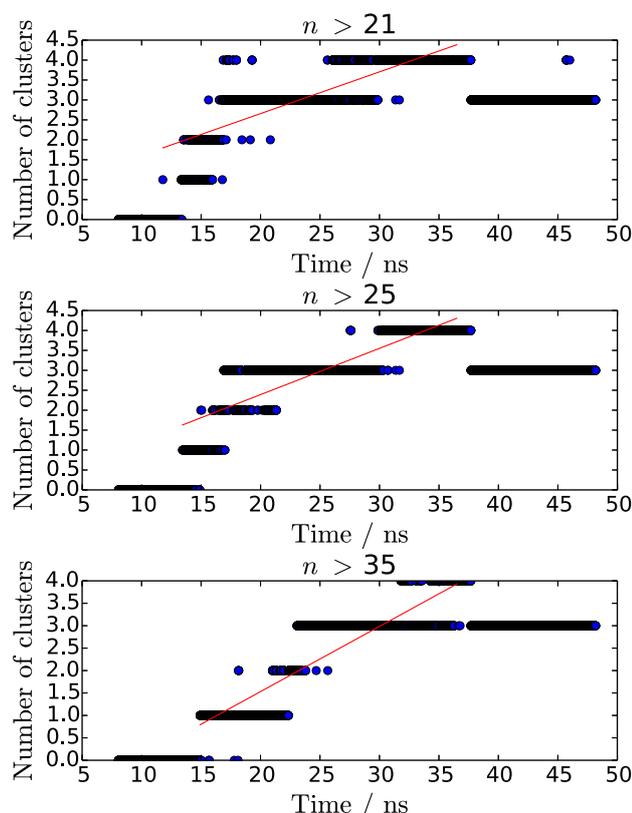


**Figure 7.** Nucleation rate of TIP4P/2005 water predicted by CNT (full line) compared to the nucleation rate observed in our MD experiment (red circle). The uncertainty in the CNT nucleation rate is given by the standard deviation of pressure in the induction period ( $\delta p = 1.67$  kPa). The uncertainty in the MD nucleation rate is given by the standard deviation of temperature in the induction period.

The temperature of our nucleation experiment is taken as the average temperature of the induction phase, i.e. 347.28 K with STD of 1.092 K, and the pressure is the average pressure during the induction phase, i.e. 69.82 kPa with STD of 1.67 kPa.

Apart from the above-presented rough estimate of the nucleation rate, the nucleation rate in MD simulations is usually calculated from the slope of the number of clusters above some defined supercritical size vs. time of the simulation [1,2]. Such a method can be used in simulations with considerably larger supersaturation, or considerably larger number of molecules, where enough supercritical clusters are observed. In our case, the use of the slope method is rather questionable, as the number of supercritical clusters reaches just 4. Nevertheless, the nucleation rate evaluated from the slope of the number of supercritical cluster is presented below for completeness, and is found to lie quite close to the rough estimate,

In figure 8, a linear fit to the number of clusters above three different supercritical sizes,  $n > 21$ , 25, 35, respectively, is plotted as a function of simulation time. Again, as in the previous case of the rough nucleation rate estimate, the fitting time was limited to 38 ns, i.e. before the coalescence of two supercritical clusters. The nucleation rate is now given by the slope of the fit in figure 8.



**Figure 8.** Time evolution of supercritical clusters, for the three thresholds in cluster size, 21, 25, and 35.

The slope should be independent on the threshold of the cluster size if it is larger than the critical cluster size. Therefore, the most robust estimate can be estimated by determining a mean slope for thresholds from cluster size

21 to 38 molecules which is  $2.03 \times 10^{29} \text{ m}^{-3} \text{ s}^{-1}$  with STD of different slopes is  $0.28 \times 10^{29} \text{ m}^{-3} \text{ s}^{-1}$  and aximum STD of slope estimate is  $0.04 \times 10^{29} \text{ m}^{-3} \text{ s}^{-1}$ . This very low uncertainty of slope estimate, despite of large variation of data (clearly seen from figure 8), is due to large number of data points of roughly 4400. The overall uncertainty, combining the two above mentioned uncertainties, is mainly influenced by the change of the slope for the different thresholds, so it is  $0.28 \times 10^{29} \text{ m}^{-3} \text{ s}^{-1}$ .

#### 4.5 Comparison of CNT predictions and results of numerical experiment

A comparison of the MD vs. CNT nucleation rate in the TIP4P/2005 system is shown in figure 7. The MD nucleation rate is plotted with an errorbar showing the uncertainty of the measurement of temperature during the induction phase given by the standard deviation of 1.092 K. To illustrate the uncertainty in the pressure measured in the induction period of the simulation, the CNT nucleation rate is calculated at pressures higher and lower by the standard deviation in pressure of 1.67 kPa.

The differences between the nucleation parameters predicted by the CNT and those observed in our MD experiment, as shown in figure 7, are summarized in Table 1. The CNT nucleation rate is slightly larger than the nucleation rate observed in our MD experiment. This is actually a very good agreement; within nucleation theory deviations in the nucleation rate of less than one order of magnitude are usually considered negligible.

The fact that CNT nucleation rate is higher than the MD nucleation rate is natural. The CNT was derived for an isothermal and isobaric system, or, in other words, for the NPT ensemble. Our MD simulation, on the contrary, was run in the NVE ensemble. In an NVE system, the temperature is not held constant by a thermostat, and as the nucleation process proceeds the temperature increases due to the evaporation heat released by the growing nuclei. Similarly, the pressure of the system increases, and both these changes can be observed in our simulation during the nucleation phase, see figure 3. The increasing temperature and pressure have a lowering effect on the instantaneous nucleation rate, and therefore the NVE nucleation process is inherently time-dependent, and it can never achieve the steady-state as in NPT. As a result, the number of nuclei produced in an NPT ensemble (as modelled by CNT) at a given starting temperature and pressure will be necessarily higher than the number of nuclei produced in an NVE ensemble (as seen in our MD simulation) at the same starting temperature and pressure.

Even the CNT nucleation rate is higher than the MD nucleation rate, the CNT prediction of the nucleation rate is still underestimated from the theoretical point of view, because it involves an ideal-gas, single-molecule approximation of the monomer attachment frequency for the evaluation of the nucleation rate. The CNT expression for the attachment frequency, as given on the right-hand side of Eq.(9), was found to be 8.13 smaller than the attachment frequency observed in our MD experiment, see Sec. 4.3. The single-molecule assumption in the growth process of clusters within the CNT is clearly not

satisfactory in our case of the high nucleation rate, where we observed cluster-cluster collisions even for two supercritical sizes. The CNT prediction of the nucleation rate at 347.28 K would be  $1.86 \times 10^{30} \text{ m}^{-3} \text{ s}^{-1}$  in case the real attachment frequency observed in our simulation would be utilized within CNT.

**Table 1.** Nucleation parameters of the TIP4P/2005 model system at  $T = 347.28 \text{ K}$  and  $p = 69.82 \text{ kPa}$  as observed in our MD simulation and predicted by the CNT.

	MD	CNT
$J / \text{m}^{-3} \text{s}^{-1}$	$2.03 \pm 0.28 \times 10^{29}$	$2.29 \times 10^{29}$
$n_c$	21 - 25	11.43
$r_c / \text{nm}$		0.438
$W / \text{zJ}$		49.83

Although the MD nucleation rate is satisfactorily predicted by CNT, there is a large discrepancy in the estimation of the critical cluster size  $n_c$ , as can be seen in Table 1. The kinetically defined critical cluster, Sec. 4.2, is roughly twice as large as the CNT critical cluster. In order to arrive at critical clusters larger than 21 molecules according to CNT, the temperature would need to be raised above 356 K and the corresponding CNT nucleation rate would drop below  $2.5 \times 10^{27} \text{ m}^{-3} \text{ s}^{-1}$ . To identify the reasons for the discrepancy in the estimation of the critical cluster size, a detailed analysis of cluster formation work in the MD simulation needs to be done. Such an analysis would enable to evaluate the formation work of the clusters in the simulation and estimate the maximum formation work. We will focus on the calculation of the formation work as one of our future activities.

## 5 Conclusions

DL\_POLY computational tool was used to carry out an MD experiment of water nucleation in NVE ensemble. A total of 10976 TIP4P/2005 water molecules were simulated in  $869^3 \text{ \AA}$  cubic cell for more than 45 ns, at an energy of  $705 \text{ K molecule}^{-1}$ , and the number density of vapour  $1.14 \times 10^{25} \text{ molecules m}^{-3}$ .

In-house computational tools were developed to analyse the cluster formation kinetics to determine the nucleation rate and the critical cluster size and to compare the MD nucleation parameters with CNT predictions. Despite excellent agreement between the simulated and theoretical nucleation rates, the critical cluster size determined from MD is roughly two times bigger than from CNT. There is also a disagreement in the growth rate between MD and CNT estimates; the MD growth rate is about eight times higher. We suggest that the main reason for the discrepancies between the simulation and theory lies in CNT simplification of cluster kinetics.

However, a more complex analysis is a task for future work involving the determination of the formation work of clusters in the MD simulation.

## Acknowledgement

We would like to express our gratitude to Professor Jiří Kolafa from The Institute of Chemical Technology, Prague for introducing us into the amazing world of molecular dynamics.

Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme "Projects of Large Infrastructure for Research, Development, and Innovations" (LM2010005), is greatly appreciated.

The project has been supported by Czech Science Foundation grant No. P101/11/1593 and with institutional support RVO:61388998.

## References

1. K. Yasuoka, M. Matsumoto, *J.Chem.Phys.*, **109** (1998)
2. H. Matsubara, T. Koishi, T. Ebisuzaki, K. Yasuoka, *J.Chem.Phys.*, **127** (2007)
3. J. Kolafa, personal communication, November 27, 2013
4. J.L.F. Abascal, C. Vega, *J.Chem.Phys.*, **123** (2005)
5. W. Smith, T.R. Forester, *J. Mol. Graph.* **14** (1996)
6. W. Smith, T.R. Forester, I.T. Todorov, THE DL\_POLY\_4 USER MANUAL, STFC Daresbury Laboratory Daresbury, Warrington WA4 4AD Cheshire, UK, Version 4.06, July 2014
7. C. Vega, J.L.F. Abascal, I. Nezbeda, *J. Chem. Phys.* **125** (2006)
8. C. Vega, E. de Miguel, *J. Chem. Phys.* **126** (2007)
9. Wagner and Pruss, *J. Phys. Chem. Ref. Data*, **31** (2002)