Water liquid-vapor equilibrium by molecular dynamics: Alternative equilibrium pressure estimation

Michal Ilčin, Martin Michalík, Klára Kováčiková, Lenka Káziková, Vladimír Lukeš

Department of Chemical Physics, Slovak University of Technology, SK-812 37 Bratislava, Slovakia
michal.ilcin@stuba.sk

Abstract: The molecular dynamics simulations of the liquid-vapor equilibrium of water including both water phases — liquid and vapor — in one simulation are presented. Such approach is preferred if equilibrium curve data are to be collected instead of the two distinct simulations for each phase separately. Then the liquid phase is not restricted, e.g. by insufficient volume resulting in too high pressures, and can spread into its natural volume ruled by chosen force field and by the contact with vapor phase as vaporized molecules are colliding with phase interface. Averaged strongly fluctuating virial pressure values gave untrustworthy or even unreal results, so need for an alternative method arisen. The idea was inspired with the presence of vapor phase and by previous experiences in gaseous phase simulations with small fluctuations of pressure, almost matching the ideal gas value. In presented simulations, the first idea how to calculate pressure only from the vapor phase part of simulation box were applied. This resulted into very simple method based only on averaging molecules count in the vapor phase subspace of known volume. Such simple approach provided more reliable pressure estimation than statistical output of the simulation program. Contrary, also drawbacks are present in longer initial thermostatization time or more laborious estimation of the vaporization heat. What more, such heat of vaporization suffers with border effect inaccuracy slowly decreasing with the thickness of liquid phase. For more efficient and more accurate vaporization heat estimation the two distinct simulations for each phase separately should be preferred.

Keywords: water, liquid-vapor equilibrium, molecular dynamics simulation, equilibrium pressure estimation, vaporization heat

Introduction

Many macroscopic properties including the existence of condensed (liquid or solid) matter are determined by molecular interactions. These interactions are responsible for the state of substance at given ambient conditions (for example, if the molecules ruled by the mentioned interactions "stay together" under given conditions the substance is in liquid or solid state). Molecular interactions can be considered as the forces between atoms the molecules are composed of. These forces may be well-described by overall potential function consisting of many specific contributions for smaller groups of several atoms. Considering only atoms from the same molecule the interactions are denoted as intra-molecular. These typically include bond, angular and torsion contributions to intra-molecular potential. Interactions between atoms of distinct molecules are covered under inter-molecular (van der Waals) potential. Summing up all inter- and intra-molecular potential contributions for all particles involved in simulation the overall potential function is obtained. This function presents total potential energy of the system under study depending on the positions of individual

atoms. Together with kinetic energy of all individual atoms corresponds with internal energy and is useful for the evaluation of thermodynamic properties. Except some specific properties, e.g. magnetic properties or NMR shifts related with nuclear magnetic moments, many others including thermodynamic and transport properties can be calculated from the knowledge of the potential function only. Analytical expressions involving potential function into evaluation have been developed theoretically for some properties, but their application is often limited to very simple molecules or atoms only (like expressions for transport properties in Hirschfelder *et al.* (1954) and Maitland *et al.* (1981)). Due to the lack of analytical expressions or due to their limited applicability, the molecular dynamics simulation must be performed with suitable starting composition and appropriate initial conditions in order to obtain demanded quantity. The equations of motion for each atom are solved during simulation. So besides the potential energy, the forces on each atom are also necessary to calculate (as a negative gradient of potential function in the coordinates of the atom of interest) in every time step of simulation. Afterwards, adequate evaluation method must be applied to necessary data

collected during the simulation (some quantities may be evaluated with different methods using different data from simulation, see e.g. diffusion coefficient evaluation in the work of Sládek *et al.* (2014)). The simulations use to involve as small molecules count as possible and also use to be as short as possible to reach for the shortest calculation times. However, sufficiently large size and length of simulation is important for averaging in order to achieve reasonable low statistical error. Therefore some authors are presenting their results from at least two different size simulations, like van der Spoel et al. in their study of water models (van der Spoel *et al.*, 1998) with 216 and 820 water molecules. Sufficient length play important role mainly for pressure averaging in liquid phase simulation as pressure value oscillations used to be very large (in the mentioned work of van der Spoel the rms pressure fluctuations were several hundreds of atmospheres what was in accordance with our experiences).

Although large number of molecular dynamics simulations is available in the literature (e.g. study of polarizable AMOEBA water model of Ren *et al.* (2004) or vapor-liquid coexistence study of Sakamaki *et al.* (2011)), many technical details connected mainly with setting of initial conditions depending on the purpose of simulation are still open. In this work, the liquid-vapor equilibrium curve of water was obtained from the phase interface simulation including both phases in one simulation. Also heat of vaporization has been estimated from such simulations. Force field of the first choice was the AMOEBA force field (Ren *et al.*, 2002; Ren *et al.*, *J. Phys. Chem. B* 2003; Ren *et al.*, *J. Am. Chem. Soc.* 2003) because of good accuracy of the heat of vaporization reported (Ren *et al.*, 2004). Results are compared with one simpler and less computationally-intensive water potential model with quality of vaporization heat comparable to used AMOEBA force field.

Simulation and calculation details

Typical way of obtaining the heat of vaporization by molecular dynamics (MD) simulation (Ren *et al.*, 2004) is to simulate liquid phase in one simulation and vapor phase as a separate one. Simulation is usually performed in quite small simulation box involving several hundreds of molecules with the size of the simulation box corresponding to density of the real substance at the simulation temperature. But such state is very likely not the "natural equilibrium state" for potential function used in simulation due to its deviations from interactions in real matter. Another approach use also barostat besides the thermostat to settle also the pressure average on desired value (van der Spoel *et al.*, 1998). Pres-

sure and temperature use to be adjusted according the real substance equilibrium parameters in such simulations. The only establishing quantity is the volume of the simulation box influencing the density. Also such established state cannot be in general considered as the "natural equilibrium state" of used force field.

Our goal was to obtain true information about the temperature dependence of liquid-vapor equilibrium pressure only from simulation. Therefore we left the liquid phase to relax with vapor phase in one simulation. This simulation is mimicking the liquid-vapor phase interface with both phases present. Typically, the simulation is performed in rectangular periodic box with one side much longer than the others (Matsumoto, 1998; Xie, 2012). Inside the box the "central stripe" is assembled from the molecules of studied substance (see Fig. 1). Density of the central stripe approximately correspond to liquid phase. Empty space is reserved for the vapor phase. Used dimension of rectangular periodic box was from 90 to 130 Å × 40 Å × 40 Å (longest dimension was increased to let more molecules to evaporate mainly in order to test the reliability of results). Dimension of initial central stripe was 20 Å × 40 Å × 40 Å and 1080 water molecules were placed inside.

For the initial distribution of molecules into central stripe the Packmol (Martinez *et al.*, 2003; Martinez *et al.*, 2009) program has been used. Minimal interatomic distances for atoms in different molecules were set to 2.0 Å (default in Packmol). This distance obviously corresponds to repulsive region of used potential function what can be seen from the temperature evolution during the initial thermostatization (see Fig. 2).
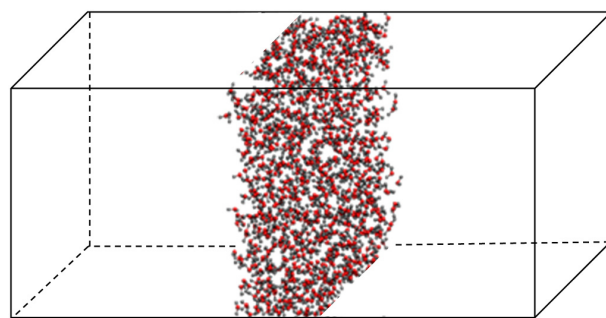

Fig. 1. Typical initial setup of molecules in the form of "central stripe" suitable for liquid-vapor equilibrium molecular dynamics study. Dimensions used in our simulation: smallest side of periodic box was 40 Å × 40 Å, longest edge was varying between 90 to 130 Å. Central stripe contained 1080 molecules with initial thickness of 20 Å.
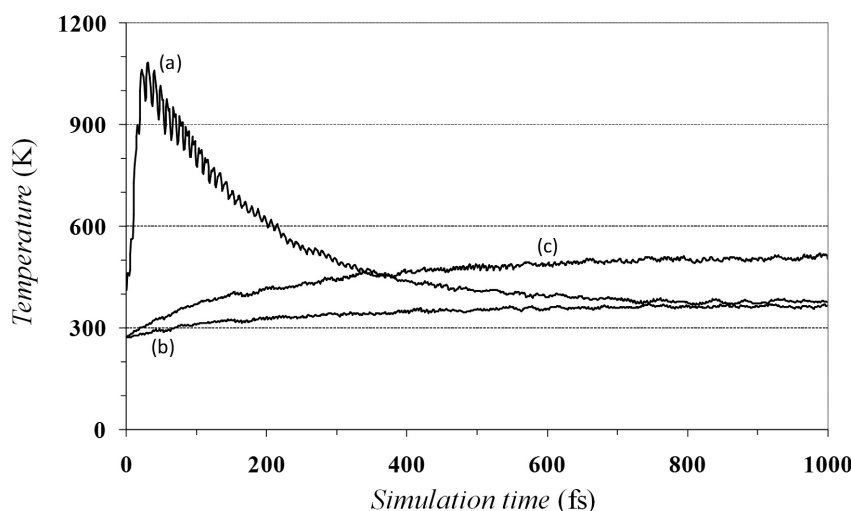
Fig. 2. Typical exponential response of Berendsen thermostat. Set temperature: 100 °C — cases (a) and (b), 250 °C — case (c). Case (a) — the simulation started as new one from initial distribution of molecules provided by Packmol. Case (b) and (c) — the simulation followed previous one thermostatized to 0 °C.

Simulations have been performed using TINKER program package (Ren *et al.*, 2002; Ren *et al.*, *J. Phys. Chem. B* 2003; Pappu *et al.*, 1998; Hodsdon *et al.*, 1996; Kundrot *et al.*, 1991; Ponder *et al.*, 1987). Simulation time step was 1 fs and default potential cut-offs were used. The AMOEBA (Ren *et al.*, 2002; Ren *et al.*, *J. Phys. Chem. B* 2003; Ren *et al.*, *J. Am. Chem. Soc.* 2003) polarizable atomic multipole force field was used for modeling water molecules interactions. The vaporization heats are quite close to the experimental one using AMOEBA force field (less than 4 % below experimental values in the temperature range of interest). On the other side, calculations involving this force field are more time consuming in comparison with other simpler models. For the sake of comparison also three-particle TIP3F model from OPLSAA force field (Jorgensen *et al.*, 1996; Maxwell *et al.*, 1995; Jorgensen *et al.*, 1998; McDonald *et al.*, 1998; Jorgensen *et al.*, 1999; Price *et al.*, 2001) was used. This model was selected due to the similar quality of heat of vaporization, around 5 % above the experimental values.

Introductory thermostatization has been done using Berendsen thermostat (Berendsen *et al.*, 1984) with time constant of 0.1 ps. From the Fig. 2 it seems the time of 1 ps is sufficient for thermostatization. It may be true for one phase simulation but was not for phase interface simulation with only liquid phase on the start of simulation. In such simulation the liquid phase needs to "correct" its volume and also the sufficient amount of molecules must evaporate to establish the evaporation-condensation equilibrium. This settlement of conditions took at most few picoseconds, thus the initial thermostatization time of 10 ps with Berendsen thermostat (Berendsen

et al., 1984) should be sufficient. The Berendsen thermostat in spite of Nose-Hoover thermostat is not producing Maxwell velocity distribution in general. Therefore the thermostat was then switched to Nose-Hoover one (Hoover, 1985; Nose, 2001). Time of 10 ps with new thermostat should assure Maxwell velocity distribution establishment. Histogram of molecular velocities compared with Maxwell distribution curve is presented on Fig. 3 for one such simulation and confirms this establishment.

The simulation time with Nose-Hoover thermostat was 50 ps. During this period the temperature was held in few degrees interval around the set point (standard deviation of temperature was less than 5 degrees from 100 time steps). Dedicating first 10 ps to the stabilization of new thermostat conditions the time of remaining 40 ps is suitable for demanded
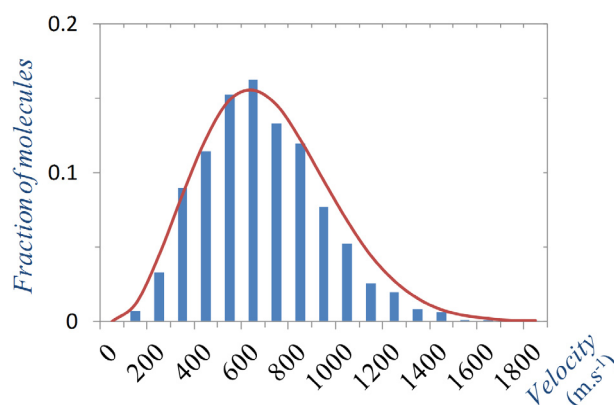

Fig. 3. Comparison of water molecules velocity distribution from simulation (histogram) and theoretical Maxwell velocity distribution (solid line) for temperature of 100 °C.

Ilčin M et al., Water liquid-vapor equilibrium by molecular dynamics...

quantity estimation. Furthermore, simulation was left to continue without thermostat for the next 50−100 ps. This part of the simulation was quite numerically stable with only small variations of total energy and almost the same temperature variations as during previous period with Nose-Hoover thermostat. Simulation period without thermostat was used for estimation of demanded quantities as well and with practically the same results.

Pressure in the simulation box is obtained from virial expression (Hansen *et al.*, 1986)

$$pV = NkT + \frac{1}{3}\sum_{i<j}\vec{r}_{ij}\cdot\vec{F}_{ij} \qquad (1)$$

and similarly as the temperature it needs to be averaged. But while in the simulation with few hundred molecules of water the temperature variations are within few degrees, the variations of pressure in liquid or both phases simulation are much larger than the average value (pure vapor phase pressure is quite reliable without extremely enormous variations). Therefore other approach has been used to get more reliable pressure information. The number of molecules has been averaged in the space around central liquid stripe corresponding to the vapor phase. Several different dimensions of vapor space has been tried with similar results. The increase of the simulation box dimension was done also in order to involve more molecules into averaging (mainly at lower temperature). From the averaged molecules count the pressure has been calculated using the equation for ideal gas.

The potential and kinetic energies evaluated at each time step of simulation become very useful in the calculation of heat (enthalpy) of vaporization. The sum of these contribution forming the total energy represents the internal energy. From the simulation of pure liquid phase with N molecules the internal energy $U_{(liquid)}$ and similarly from the simulation of pure vapor phase with N molecules the internal energy $U_{(vapor)}$ can be obtained. Dividing these internal energies by molecules count and multiplying by Avogadro constant N_A the molar internal energy $U_{m,(phase)}$ is calculated as

$$U_{m,(phase)} = \frac{U_{(phase)}}{N}N_A \qquad (2)$$

(the N need not to be same for liquid and vapor phase).

From molar internal energies of liquid and vapor phases the heat (enthalpy) of vaporization is

$$\Delta_{vap}H_m = U_{m,(vapor)} - U_{m,(liquid)} + RT \qquad (3)$$

approximating the product of pressure and molar change of volume during vaporization as RT. Furthermore, considering molecules have approxi-

mately the same kinetic and intramolecular potential energy in vapor phase as in liquid phase, the only difference is molar inter-molecular potential energy $V_{m,inter-mol.}$. Taking also very long distances between molecules in gas phase into account, virtually zero intermolecular potential energy is expected in the gas phase. Then the heat of vaporization can be approximated as

$$\Delta_{vap}H_m \approx -V_{m,inter-mol.}^{(liquid\ phase\ only)} + RT \qquad (4)$$

meaning only the simulation of liquid phase is necessary. But this approach cannot be easily done if Ewald summation of electrostatic interactions is applied (there is no intermolecular part of potential energy in simulation program output) what was also our case. What more, having both (liquid and vapor) phases in one simulation comes with some new complications. Part of the molecules are in vapor and the rest in liquid phase. This complicates the evaluation of molar energies − prior to this the molecules counts in liquid and vapor phase need to be estimated. For this reason, the molecule count averages in vapor phase have been evaluated. As the phase interface between liquid and vapor phase is not so easy to determine and also need not to be exact plane, the vapor molecule counts were evaluated for several planar borders closer and closer to phase interface. Until reaching the interface the number of molecules raises linearly with distance when approaching the interface. After crossing the phase interface the increase of molecules count started to be more rapid because the inclusion of molecules from liquid phase. This jump in molecules count provides quite easy way of finding the vapor and also liquid phase molecules counts estimated with quite small uncertainties. Using this information together with related internal energy and vapor molar internal energy from smaller vapor phase simulation the heat of vaporization has been estimated from phase interface simulation.

Results and discussion

The MD simulations have been performed in the temperature range between 373 K and 473 K (from 100 °C to 200 °C). In the Fig. 4 the screenshots of MD simulations of liquid-vapor phase interface for different temperatures are presented. First 3 pictures Fig. 4a−c represent screenshots from simulations using the AMOEBA force field for temperatures of 373 K, 423 K and 473 K. One can clearly observe the increase in the density of vapor phase. For the sake of comparison, also the screenshot from simulation using TIP3F water model for 473 K is depicted in Fig. 4d. Comparing with the AMOEBA picture for the same temperature, the vapor density is lower.
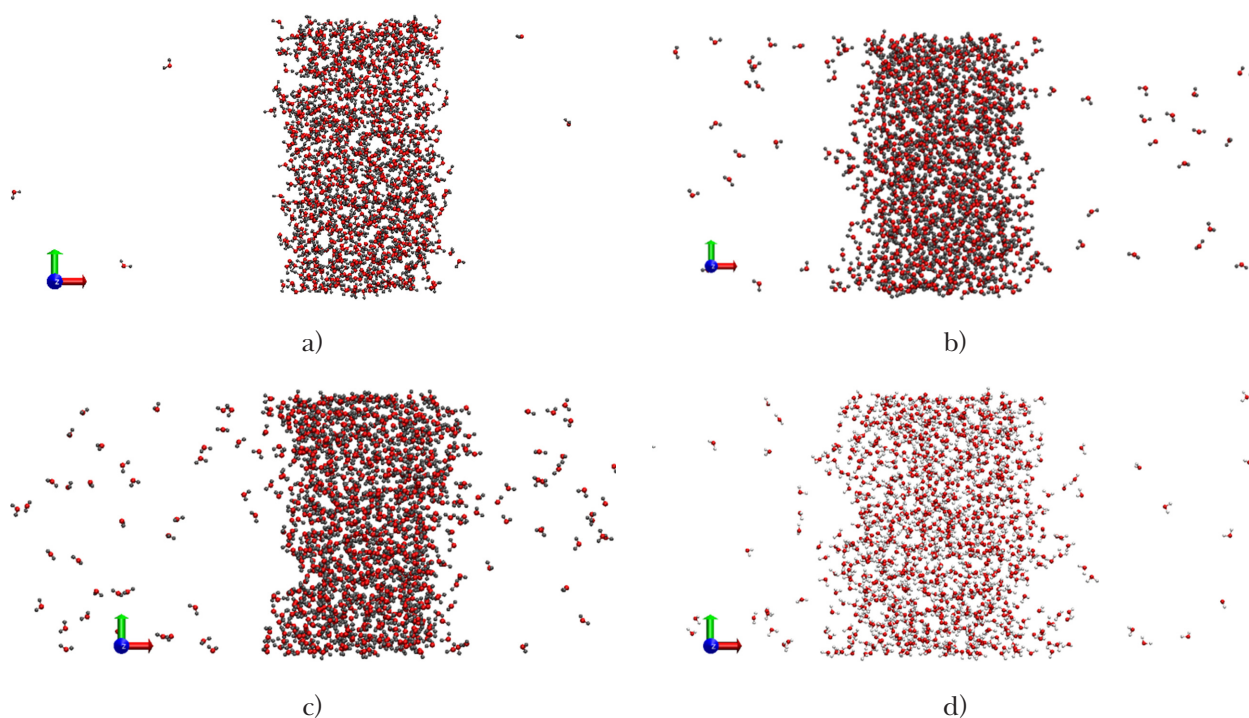
Fig. 4. Screenshots from simulations after thermostatization for AMOEBA force field at (a) 373 K, (b) 423 K, (a) 473 K and for TIP3F water model of OPLSAA force field at (d) 473 K.
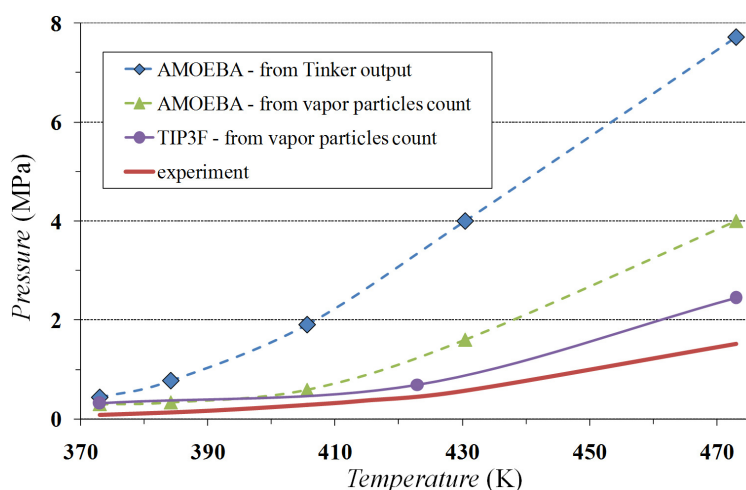

Fig. 5. Vapor-liquid equilibrium pressure dependence on the temperature — comparison of simulation results with experimental data (Dean, 1999).

Thus for TIP3F (OPLSAA force field) water model one can await lower equilibrium pressure than for AMOEBA force field.

In the Fig. 5 the liquid-vapor equilibrium pressure dependence on the temperature is presented. The pressure provided in TINKER output averaged for quite long period was overestimated in simulations using AMOEBA force field. For TIP3F water model the pressure averages from TINKER output were underestimated and even below zero. The results obtained from vapor phase molecules count are more real and closer to the real vapor-liquid equilibrium pressure dependence. TIP3F water model equilibrium pressures for higher temperatures are lower than for AMOEBA force field, what is in agreement with brief view of pictures in Fig. 4c and Fig. 4d. Among water models used in work of Sakamaki *et al.* (2011), the TIP3P model (similar to TIP3F model used in this work) provided also slightly overestimated equilibrium pressure.

Also several further water models used in the work of Sakamaki *et al.* (2011) provided overestimated pressures but the rest of water models provided underestimated results for equilibrium pressures. Due

Ilčin M et al., Water liquid-vapor equilibrium by molecular dynamics...

to large fluctuations the convergence of pressure values was slow and Sakamaki et al. needed to perform long simulations. Depending on used model, also obtained critical temperatures were different. Interesting conclusion of their work is the fact that despite of underestimating or overestimating the experimental equilibrium pressures, the pressure dependence on the reduced temperature (for different model temperature divided by different critical value) is underestimated for every model above the pressure of 0.1 MPa.

The estimated heats of vaporization from phase interface simulations are presented in Fig. 6. Values obtained from these simulations are underestimated comparing with experimental data and also comparing with "properly calculated" heats of vaporization. Properly calculated heats of vaporization are meant as evaluated from two simulations of purely liquid and purely vapor water phases. The difference lies in existence of "surface water" molecules in the simulation with phase interface comparing with pure liquid phase simulation containing no surface molecules. Comparing to the work of Ren *et al.* (2004), where they studied vaporization heats at the pressure of 1 atmosphere for temperatures up to almost 370 K, our "properly calculated" heats for higher temperatures were in good accordance with their results, satisfactorily matching at the ranges intersection.

The underestimation of the vaporization heats, observable in Fig. 6, has a clear reason in quite large fraction of molecules on the surface. While these molecules have neighbor molecules only from one side, their contribution to intermolecular potential energy is lower than ones from the middle of liquid. Therefore such values of the vaporization heat are underestimated comparing to the properly calculated heat of vaporization. Improving these underestimated vaporization heats is possible by increasing thickness of central stripe thereby lowering fraction of surface molecules. But with the increase of the thickness also much more molecules are involved into simulation, considerably prolonging simulation time with still not very accurate results. Properly calculated vaporization heats are reliable and more accurate if evaluated also from much smaller simulations of both phases separately. This makes the simulation of liquid-vapor phase interface an ineffective approach for vaporization heat estimation.

Conclusion

An attempt to obtain the temperature dependence of vapor-liquid equilibrium pressure from the MD simulation only has been made. Initial simulated system arrangement, with absolutely no restrictions (in one dimension) on condensed phase, provided enough freedom to molecules of liquid water, allowing them to occupy the space with appropriate volume. Neither volume of liquid nor the pressure was constrained. The volume occupied by liquid and pressure thus naturally results from potential function of chosen force field and the set temperature. Due to the space left for vapor phase, the temperature was the only quantity forced to
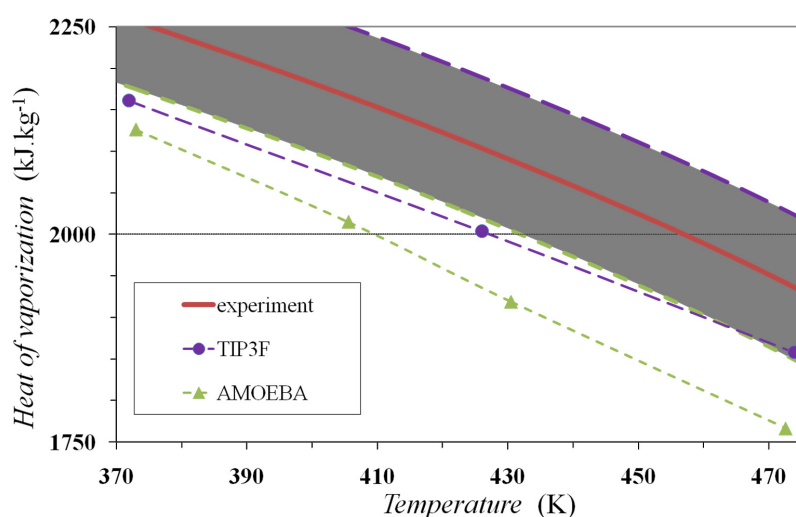

Fig. 6. The vaporization enthalpy obtained from the vapor-liquid phase interface simulation. Proper AMOEBA results agree with experimental curve (Dean, 1999) within 4 % deviation downwards and proper TIP3F results are over 5 % above this line (they are approximately on the edges of gray ±5 % error stripe). Underestimation of our two-phases simulation values is due to "border effect", i.e. existence of less strongly bounded "surface" molecules.

the system under study. Berendsen thermostat was applied in the initial phase and Nose-Hoover thermostat in the production phase of the simulation. After initial thermostatization, followed by subsequent production run with suitable thermostat permanently on, the simulation stepped into testing phase with thermostat switched off. Virtually the same averages of evaluated quantities were achieved during this test run without thermostat. But besides large standard deviations of the pressure also obtained values of pressure were not very reliable. Therefore the alternative method of pressure estimation were elaborated which works with vapor phase molecules count and ideal gas equation. Provided pressure averages were more reliable and with noticeably smaller deviations despite the simplicity of used method.

From equilibrated production and test runs, the heat of vaporization were also estimated. But not like at pressure estimation based on vapor phase molecules count, when it was sufficient to make the averages only in some subspace of overall vapor phase volume, now the total molecules count in vapor phase is necessary. Not counting for every vapor phase molecules causes that the not counted molecules remain regarded as the liquid phase ones. They artificially increase molecules count N of liquid phase in equation (2) leading in fictitiously decreasing absolute value of molar internal energy $U_{m,(liquid)}$. Considering $U_{m,(liquid)}$ is negative, the decrease in magnitude of $U_{m,(liquid)}$ will also shift the molar heat of vaporization $\Delta_{vap}H_m$ lower according to the equation (3). More clearly it can be seen from equation (4), in which the intermolecular potential is negative and thus positive value of

$$-V_{m,inter-mol.}^{(liquid\ phase\ only)},$$

affected similarly by artificial increase of N like molar internal energy, will decrease the value of $\Delta_{vap}H_m$ by fictitious decreasing of its own value. The need for correct total number of molecules in vapor phase makes the vaporization heat estimation more laborious comparing to the averaging of the energy values from the standard outputs of simulation program for separated liquid and vapor simulations. But also when the total number of molecules in vapor phase was correctly evaluated, the value of $\Delta_{vap}H_m$ still remained underestimated. Only with increasing the thickness of central liquid stripe the vaporization heat will raise its value closer to value from two distinct simulations of vapor and liquid phase. Therefore, because of the necessity to calculate $\Delta_{vap}H_m$ as a limit for infinitely increasing liquid stripe thickness and also due to more laborious estimation of $\Delta_{vap}H_m$ in every simulation for concrete liquid stripe thickness, calculation of vaporization heats from liquid-vapor equilibrium simulation is completely not recommended. Instead of this, standard method based on two distinct simulations of liquid and vapor phase is highly recommended for vaporization heats calculation.

Abbreviations

MD – molecular dynamics, AMOEBA – Atomic Multipole Optimized Energetics for Biomolecular Applications, OPLSAA – Optimized Potentials For Liquid Simulations – All Atoms

References

Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) *J. Chem. Phys.* 81: 3684.

Dean JA (1999) *Lange's Handbook of Chemistry*, 15th edition, McGraw-Hill, New York.

Hansen J-P, McDonald IR (1986) *Theory of Simple Liquids*, Academic Press, London.

Hirschfelder JO, Curtiss CF, Bird RB (1954) *Molecular Theory of Gases and Liquids*, Wiley: New York.

Hodsdon ME, Ponder JW, Cistola DP (1996) *J. Mol. Biol.* 264: 585.

Hoover WG (1985) *Phys. Rev. A* 31: 1695.

Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) *J. Am. Chem. Soc.* 117: 11225.

Jorgensen WL, McDonald NA (1998) *J. Mol. Struct. – THEOCHEM* 424: 145.

Jorgensen WL, Rizzo RC (1999) *J. Am. Chem. Soc.* 121: 4827.

Kundrot CE, Ponder JW, Richards FM (1991) *J. Comput. Chem.* 12: 402.

Maitland GC, Rigby M, Smith EB, Wakeham WA (1981) *Intermolecular Forces*, Clarendon Press: Oxford.

Martinez JM, Martinez L (2003) *J. Comput. Chem.* 24: 819.

Martinez L, Andrade R, Birgin EG, Martinez JM (2009) *J. Comput. Chem.* 30: 2157.

Matsumoto M (1998) *Fluid Phase Equilibria* 114: 307.

Nose S (2001) *J. Phys. Soc. Japan* 77: 70.

Maxwell DS, Tirado-Rives J, Jorgensen WL (1995) *J. Comput. Chem.* 16: 984.

McDonald NA, Jorgensen WL (1998) *J. Phys. Chem. B* 102: 8049.

Pappu RV, Hart RK, Ponder JW (1998) *J. Phys. Chem. B* 102: 9725.

Ponder JW, Richards FM (1987) *J. Comput. Chem.* 8: 1016.

Price MLP, Ostrovsky D, Jorgensen WL (2001) *J. Comput. Chem.* 22 1340.

Ren P, Ponder JW (2002) *J. Comput. Chem.* 23: 1497.

Ren P, Ponder JW (2003) *J. Phys. Chem. B* 107: 5933.

Ren P, Ponder JW (2003) *J. Am. Chem. Soc.* 125: 15671.

Ren P, Ponder JW (2004) *J. Phys. Chem. B* 108: 13427.

Sakamaki R, Sum AK, Narumi T, Yasuoka K (2011) J. Chem. Phys. 134: 124708.

Sládek V, Bučinský L, Matuška J, Ilčin M, Lukeš V, Laurinc V (2014) *Phys. Chem. Chem. Phys.* 16: 18519.

van der Spoel D, van Maaren PJ, Berendsen HJC (1998) *J. Chem. Phys.* 108: 10220.

Xie J-F, Sazhin SS, Cao B-Y (2012) *J. Therm. Sc. Tech.* 7: 1.