

frontpage

assignment 1/2

assignment 2/2

ACKNOWLEDGEMENT

text....

SUMMARY

Text in English.

keywords:

text, text, text

SOUHRN

Text.

klíčová slova:

text, text, text

CONTENTS	1
1.1 Studied compounds	2
1.1.1 Polylactic acid	2
1.1.2 Active pharmaceutical ingredients	3
2 THEORETICAL PART	5
2.1 Force fields	5
2.1.1 OPLS force field	6
2.2 Periodic boundary conditions	9
2.2.1 Setting the cutoff	9
2.2.2 Ewald summation	10
2.2.3 Particle-Particle Particle-Mesh	10
2.3 Molecular dynamics	11
2.3.1 Verlet integration	11
2.3.2 Constraint dynamics	12
2.4 Measuring the properties	13
2.4.1 Statistical ensembles	13
2.4.2 Mean Square Displacement	13
2.4.3 Radial Distribution Function	14
2.4.4 Block averaging scheme	15
3 COMPUTATIONAL METHODS	16
3.1 Simulations of neat PLA	16
3.2 Simulations of neat APIs and mixtures with PLA	17
4 RESULTS AND DISCUSSION	18
4.1 Simulations of neat PLA	18
4.1.1 Structural properties	18
4.1.2 Polydispersity effect	21
4.1.3 Glass transition modeling	22
4.2 Simulations of neat API	25
4.2.1 Sulfathiazole FF parametrization	25
4.3 Simulations of mixtures of APIs and PLA	27
4.3.1 Excess properties	27
4.3.2 Radial distribution functions	29
4.3.3 Diffusion coefficients	38
4.3.4 Glass transition temperature	40
5 CONCLUSION	43

1 INTRODUCTION

In crystalline forms, most of the newly discovered active pharmaceutical ingredients (APIs) are poorly soluble in water, which limits their bioavailability, dissolution, and then their distribution through the organism. This fact limits potential wider use of numerous API as a solid drug in medical treatment. Today, combinatorial chemistry techniques and high-throughput screening have led to a sharp increase in the quantity of proposed nonsoluble API molecules, so the oral administration of poorly soluble drugs has become the biggest challenge for formulation scientists in the pharmaceutical industry. [1] There are different strategies to overcome this issue, such as cocrystal formation [2], conversion of an API to its salt [3] or using dispersions of API in various matrices. [4]

In 1961, Sekiguchi and Obi provided the earliest account of the so-called first-generation solid dispersion, when they discovered that the creation of eutectic mixtures enhances the rate of drug release and bioavailability. First-generation solid dispersions were built from crystalline carriers such as urea or sugars, forming crystalline solid dispersions. The second generation of solid dispersions is based on replacing crystalline carriers by amorphous carriers such as polymers, forming an amorphous product in which API is dissolved. There exist also a third generation of solid dispersions using a surfactant carrier or a combination of amorphous polymers and surfactants. [5]

The aim of researchers is to overcome the poor solubility of APIs by using amorphous solid phases of APIs and to avert the rearrangement of their molecules into a crystal lattice. However, crystalline forms of APIs are advantageous because of their better stability during long-term storage and more reliable predictions of material properties at the molecular level under defined conditions. [6] The better solubility of APIs in amorphous forms comes from a higher Gibbs energy of the amorphous form compared to the crystalline forms. During processing, storage, and after contact with water or humidity, the thermodynamically metastable amorphous forms tend to crystallise. This is a price we have to pay for the higher Gibbs energy resulted in better solubility. Solid mixtures of API and excipients (e.g. polymeric excipients) create amorphous solid dispersions (ASD) and offer a way to inhibit crystallisation of the API before and after oral administration of the dose. [7]

The creation of an amorphous dispersion of an API can generally have a twofold effect on the rate of solid-state crystallisation, affecting both thermodynamic and kinetic aspects. Thermodynamically, it reduces the Gibbs energy of the dispersion due to strong beneficial intermolecular interactions between API and its excipient, as well as it increases kinetic barriers to recrystallisation. On the atomic scale of individual interactions stabilizing such solid dispersions, hydrogen bonding makes the most significant contribution. [8]

Other suitable biocompatible and biodegradable polymers for ASD could be polyethylene glycol (PEG) and polyvinylpyrrolidone (PVP). [9] My aim in this thesis is to use computational chemistry methods to determine the suitability of mixing four model APIs with polylactic acid as a representative of a relevant biocompatible polymer. Firstly, the properties of the pure polymer such as glass transition temperature and density will be discussed. Then, PLA mixtures with different concentrations of APIs will be prepared and the effect of excipient will be studied in comparison with the pure compound.

1.1 Studied compounds

1.1.1 Polylactic acid

Polylactic acid (PLA) was chosen as a biocompatible polymer excipient. PLA is a biodegradable polymer formed by the polymerisation of lactic acid. The formula of the PLA monomer unit is shown in Figure 1. In this work, two condensed units of D-PLA were considered as the simplest building block for creating all of the other longer polymer chains. Polymer samples of a length up to 100 dimer units were created by replicating these dimer units. The molar weight of our dimer unit considered in the investigation is $M_w = 162.14 \text{ g mol}^{-1}$, which means that the longest polymer chain used in the simulations has a molar weight equal to $M_w = 14\,431 \text{ g mol}^{-1}$.

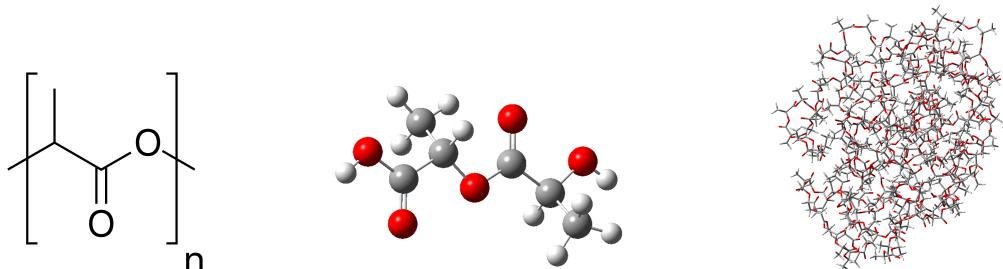


Figure 1: PLA formula on the left, PLA dimer block representing the chain unit used to build up polymer chain in the middle and a folded PLA chain containing 100 dimer block used to create mixtures with APIs on the right.

1.1.2 Active pharmaceutical ingredients

The first selected API is **ibuprofen**, systematically 2-(4-Isobutylphenyl)propanoic acid ($C_{13}H_{18}O_2$) as an example of a widely used analgesic, antipyretic, anti-inflammatory drug. The racemic mixture of ibuprofen is commonly used in medical treatment.

The S-enantiomer has a stronger pharmaceutical activity than the R-enantiomer, which is metabolically transformed to S-form in the organism. [10] In this work the S-form, which is visualised in Figure 2, is used. The molar weight of ibuprofen is $M_w = 206.28 \text{ g mol}^{-1}$ and the melting temperature of the enantiopure crystal is 324.4 K. [11]

The second selected API is **naproxen**, systematically 2-(6-Methoxynaphthalen-2-yl)propanoic acid ($C_{14}H_{14}O_3$), a non-steroidal anti-inflammatory drug, used as a painkiller. Naproxen contains three oxygen atoms (one carboxyl group and one ether bond), the structure is shown in Figure 2 in the upper right corner. On the basis of its structure, naproxen can donate one hydrogen bond and accept up to six hydrogen bonds, steric factors limits the actual coordination. Naproxen is a white crystalline powder, with a molar weight of $M_w = 230.263 \text{ g mol}^{-1}$ and melting temperature is 429.3 K. [11]

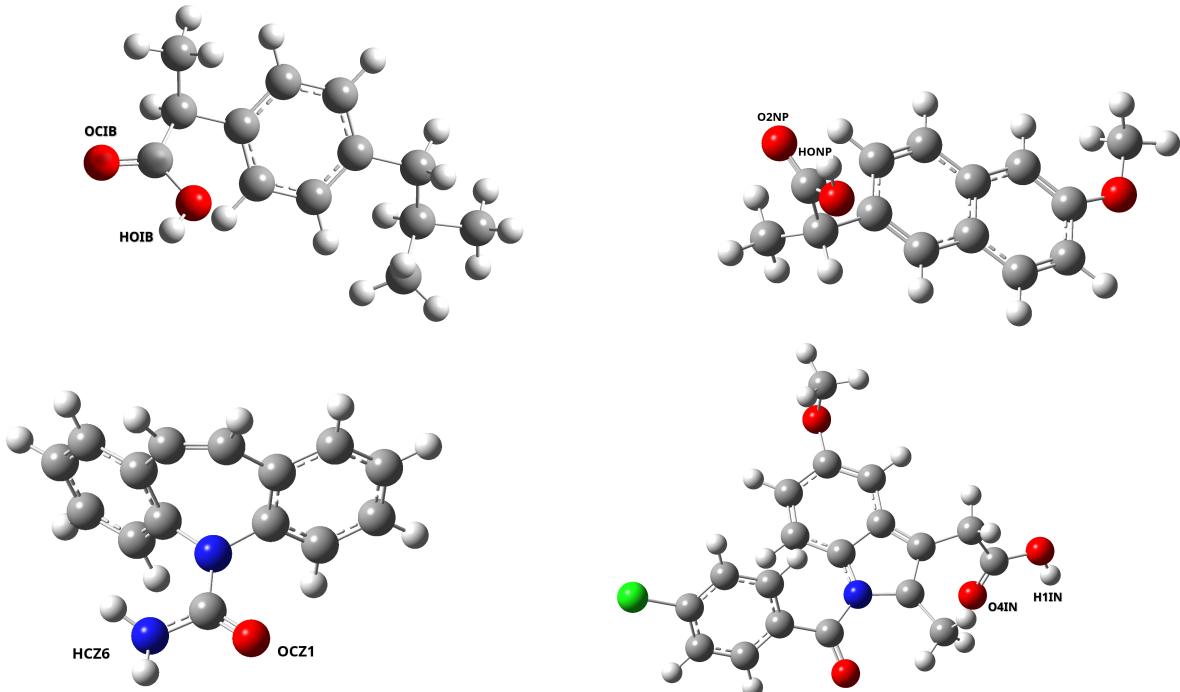


Figure 2: Molecular structures of ibuprofen (**top left**), naproxen (**top right**), carbamazepine (**bottom left**) and indomethacin (**bottom right**). Atom types contributing most to the hydrogen bonding are tagged for each molecule.

Carbamazepine, alternatively 5-Carbamoyl-5H-dibenzo(b,f)azepine ($C_{15}H_{12}N_2O$) is a representative anticonvulsant, which is used for the treatment of seizures and neuropathic pain. Carbamazepine contains two nitrogen atoms (amide group) and one oxygen in the carbonyl group; its structure is shown in Figure 2 on the left side. According to its structure, carbamazepine can theoretically accept up to two hydrogen bonds and donate up to four hydrogen bonds. Carbamazepine is a white crystalline powder, with a molar weight of $M_w = 236.273 \text{ g mol}^{-1}$ and melting temperature of 463.6 K. [11]

Indomethacin, 2-1-[(4-Chlorophenyl)carbonyl]-5-methoxy-2-methyl-1H-indol-3-ylacetic acid ($C_{19}H_{16}ClNO_4$), whith structure depicted in Figure 2, is used in the treatment of musculoskeletal and joint disorders. The molar weight is $M_w = 357.8 \text{ g mol}^{-1}$ and the melting temperature is 433.3 K. [11]

The last selected API was **sulfathiazole** with systematic name 4-amino-N-thiazol-2-ylidenebenzene-sulfonamide as a representative antibiotic drug from the sulfonamides group, which is used in the treatment of pyogenic cutaneous infections. Sulfathiazole is a white crystalline powder, with a molar weight (M_w) = 255.3 gmol^{-1} , which is highly polymorphic, five polymorphs have been discovered so far [6]. All known polymorphs of sulfathiazole crystallise in the $P2_1/c$ space group, but there are differences in intermolecular bonding and structural properties [12]. The II polymorph structure, shown in Figure 3, is used in this work. There are four molecules of sulfathiazole in the crystal monoclinic unit cell.

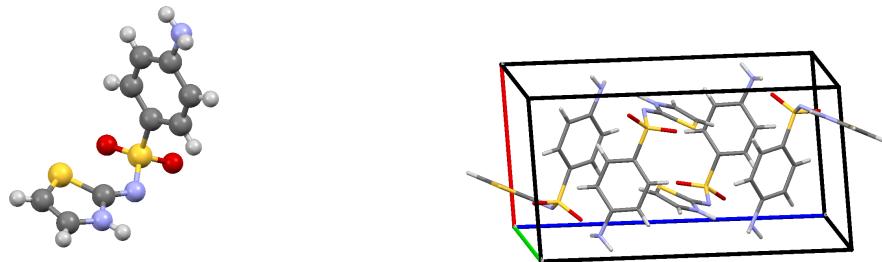


Figure 3: Sulfathiazole - molecular structure on the left and a unit cell of its II polymorph on the right.

2 THEORETICAL PART

In order to convert a real system consisting of individual molecules into a form that can be understood by a computer software, we must define individual parameters that are essential for describing mutual interactions of atoms. When describing real systems, it is usually necessary to consider a certain degree of approximation due to the possible computational complexity. In computational chemistry methods, we often encounter the so-called Born-Oppenheimer approximation, which is based on the decoupling of the motion of nuclei and electrons. The basis of the approximation is the orders-of-magnitude difference in mass between the electron and the nucleus. The latter thus move very slowly relative to the electrons and can thus be considered as a fixed point charge. This allows us to calculate the energy of a molecule as a function of the positions of the nuclei, in quantum chemistry we talk about the so-called Potential Energy Surface (PES), which is a function of $3N$ coordinates, where N is the number of nuclei. [13]

For small systems it is possible to calculate the PES based on quantum chemistry methods using reasonable resources, for very large systems this is not yet realistic. We therefore introduce a classical-mechanics set of analytic functions, yet empiric parameters called the Force Field (FF), which enable to evaluate the energy of simulated systems depending on the positions of the nuclei. Methods based on the use of such Force Fields are called molecular mechanics, which are applied especially when quantum phenomena are not of great importance and we can use classical mechanics approach or big amorphous systems where quantum-based calculations would be extremely expensive and resource taking. [14]

2.1 Force fields

The total energy calculated using the force field can be broken down into two contributions, a binding and a non-binding term, which are further expanded in Equations 2.1 and 2.2.

$$E_{\text{bonded}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}} \quad (2.1)$$

$$E_{\text{nonbonded}} = E_{\text{electrostatic}} + E_{\text{Van der Walls}} \quad (2.2)$$

By the level of functional description (number of terms) of the interactions in the force field, we distinguish force fields of three classes. Class 1 force fields contain the 5 terms mentioned in the two equations above (bond, angle, torsion, Lennard-Jones and electrostatic), examples of such FF are the DREIDING, AMBER [15], GAFF and OPLS. In addition, class 2 force fields include bond-bond and bond-angle coupling terms, anharmonic terms simultaneously with all class 1 terms, examples of such fields are PCFF or ReaxFF. The third class includes fluctuations of charge distribution in time (charge polarization effect), and they are called polarizable FF. [16]

During the parameterization of force fields (FF), we start from the assumption of transferability, similar chemical groups of different molecules interact in the same way. When constructing a force field for large molecules, we can use parameters obtained from data for small molecules, which are much more easily graspable and contain the same functional groups. [14] In model development, our aim is to achieve the most universal description of the system while still closely corresponding to its actual state. This can be facilitated by employing higher-order terms; however, incorporating anharmonic and cross terms introduces the need for a greater amount of FF parameters. We strive to avoid situations where we employ an overly adapted and detailed model that merely reproduces inserted information without providing any predictive capabilities. [16]

According to the level of parameterization, there are 3 basic types of force fields. In the first case, where the parameters are determined for each individual atom in the system, including hydrogens, we speak of an all-atom force field. A united atom force field is one where we parameterize the individual functional groups (interaction centers), such an interaction center could be for example a methyl group. The third type of force field is coarse grained, used mainly for protein and polymer simulations, offering higher computational efficiency for long simulations of large molecules by grouping them into "superatoms". [17]

2.1.1 OPLS force field

Optimized Potential for Liquid Simulations (OPLS) force field was developed by William L. Jorgensen [18] at Purdue University and later at Yale University based on previously released Assisted Model Building and Energy Refinement (AMBER) force field developed by Peter Kollman's group. [19] The OPLS force field consists of following terms written in Equation 2.3.

$$\begin{aligned}
U_{\text{OPLS}} = & \sum_{\text{bonds}} \frac{1}{2} k_b (r - r_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 \\
& + \sum_{\text{torsions}} \left(\frac{V_1}{2} [1 + \cos(\phi)] + \frac{V_2}{2} [1 - \cos(2\phi)] + \frac{V_3}{2} [1 + \cos(3\phi)] + \frac{V_4}{2} [1 - \cos(4\phi)] \right) \\
& + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j e^2}{r_{ij}} \right\} f_{ij}
\end{aligned} \tag{2.3}$$

i and j denote different atom types, N is the total number of atom pairs, ϵ_{ij} , σ_{ij} are LJ parameters, r_{ij} is the distance between atoms i and j , and f_{ij} is scaling factor equals 0.5 for 1-4 interactions ($i, j=1,4$) and 1 otherwise.

Bonds and angles are described as harmonic oscillators in OPLS FF. The equilibrium parameters are obtained by structural methods, such as x-ray diffraction NMR experiments. The values for force constants are then fitted to experimental data taken from vibrational spectroscopy. In Figure 4 is the visualization of the constants.

Proper **dihedral angle** ϕ between atoms i,j,k,l is represented in Figure 4. The torsional energy is described as a cosine expansion, where the first term corresponds to the rotation periodic after 360° , second term by 180° , the third term by 120° and the fourth term by 90° . Each of this term has the V_n constant, representing the barrier for rotation along the proper dihedral angle, which is dependent both on the torsional energy and non-bonded forces. Different approaches could be chosen, one of them is obtaining the parameters by QM calculations. [20] We first optimize the molecular geometry using QM methods and then we run scanning of dihedral angle of interest. At each step, we optimize geometry and calculate the change in potential energy. Then we compute potential energy of each optimized geometry using MD with dihedral parameters set to zero. Then we can fit the parameters by subtracting the results of MD from QM, this corresponds exactly to the influence of dihedrals. To do this procedure first all the other FF parameters should be known.

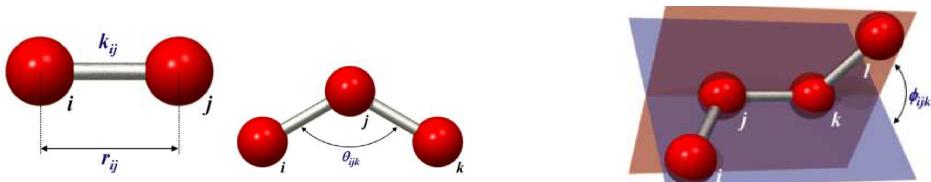


Figure 4: Pokud je v pořádku obrázek, vytvořím vlastní

Charges parametrization is also done by ab initio methods. When we are using the classical FF, the charges used to calculate the Coulomb potential remains the same during the simulations. Due to that, it is crucial to obtain the charges from equilibrium state of the molecules in order to avoid any errors from having the charges taken from structures with higher energy. When obtaining the charges, the first step is to optimize the geometry using the appropriate level of theory, meaning the basis set to be better or equal to 6-31G. Commonly used method to achieve the charges is CHELPG (CHarges from ELectrostatic Potentials using a Grid based method) [21], based on adjusting the partial charges at the centers of the nuclei in order to get the best representation of the electrostatic potential given by the wave functions. Calculation of the charges are often done using higher-level methods and basis sets such as B3LYP/cc-pVTZ or HF/6-31G**.

Van der Waals forces are most often represented by the Lennard-Jones (LJ) potential. The functional form with the illustration is in Figure 5. Lennard-Jones potential is a combination of two terms, repulsive term describes the Pauli repulsion at short distances and the attractive term describes the London dispersion force. The ϵ values are adjusted to experimental values of heats of vaporization and σ parameters are adjusted to experimental densities and structural data.

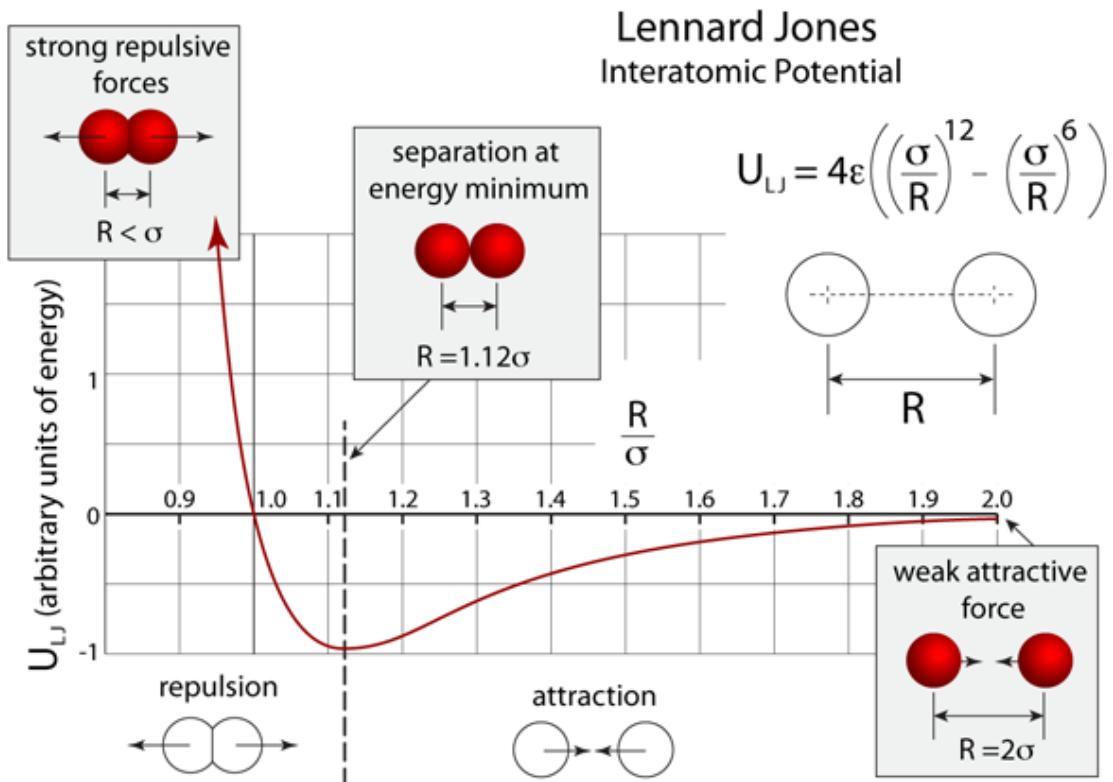


Figure 5: caption

2.2 Periodic boundary conditions

Due to the computational complexity, we are focused on only small region of a very complex real system. In order not to introduce errors caused by the boundaries of the system and interaction on them, we introduce periodic boundary conditions (PBC). This method is based on surrounding the simulated system with periodic images, thus achieving an approximation by a surface less system. We choose a shape of the simulation box that can be used to fill the space without problems, in our case a cubic box. For simulations in three dimensions, we usually introduce PBCs in the direction of all axes, which means that our simulated box is surrounded on all sides by a total of 26 replicas of the system. The behavior of the system is the same in all replicas, so we must include all the particles appearing in the replicas when calculating the pair interactions. [22]

2.2.1 Setting the cutoff

From the preceding paragraphs it is evident that the main problem in time complexity of the simulations is the evaluation of non-bonding energies. While the number of bonded terms increase linearly with the system size, the non-bonded terms show a quadratic increase in the number of contributions. A common strategy for reducing the computational time is to set specific cutoff distance, beyond which we neglect or estimate the non-bonded interaction energy contribution. We distinguish two type of interactions based on their decay with distance. First the so-called short-range interactions decreasing faster than $1/r^3$ such as van der Waals ($1/r^6$) that could be neglected on relatively short distance. By neglecting the van der Waals contributions at long distances, we introduce only a small numerical deviation for each pair, but the cumulative effect when all pairwise interactions are summed introduces larger deviations into the simulations. Theory says that van der Waals interactions are negligible beyond about 20 Å, we choose 12 Å as the optimal cutoff in this work because of the optimal computational complexity. [23]

For the long-range interactions, such as Coulombic interactions ($1/r$) the situation is not that easy and they require larger cutoff than van der Waals due to their long-range nature. The easiest option, as mentioned above, is to solve this issue by neglecting any contributions beyond the cutoff distance. The problem in that approach is that we introduce discontinuities in the potential and its derivatives. Better and more effective approach, especially for periodic systems is called Ewald summation.

2.2.2 Ewald summation

The idea of Ewald summation is based on a mathematical trick where we divide a three-dimensional very slowly and only relatively converging infinite series into two series that converge much faster. In practice, this trick means surrounding each of the point charges with a shielding diffusion cloud of opposite charge, which shape is described by Gaussian function. This shadow cloud is then compensated by a charge of identical shape but opposite sign, illustration is in Figure 6. Next, the contributions from the shadow cloud are summed in real space and the contributions of the compensation potential are summed in reciprocal space using Fourier transform. Due to the smoothness of the charge distribution using Gaussian functions this converge really fast in Fourier space. The width of the Gaussian peak determine the number of iterations to converge in reciprocal and direct space, the optimal width is then calculated in order to achieve even distribution of computational workload in both spaces. [24]

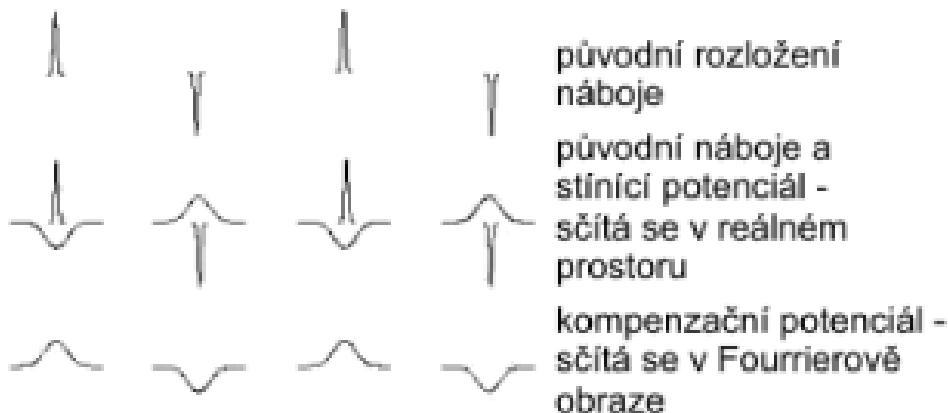


Figure 6: caption

2.2.3 Particle-Particle Particle-Mesh

The Particle-Particle Particle-Mesh (PPPM) method [25] is a development of the Ewald sum method. The PPPM method operates on the principle of decomposing the electrostatic potential into short-range and long-range components. The short-range interactions, which decay rapidly with distance, are typically calculated using direct methods like the Ewald summation or a cutoff-based approach. The long-range electrostatic potential is computed using a Fourier-based approach, where the charge distribution is mapped onto a grid using a fast Fourier transform (FFT). This grid-based representation allows for efficient calculation of the potential at grid points using convolution techniques. [24]

2.3 Molecular dynamics

The molecular dynamics method is based on solving the equations of motion of classical Newtonian mechanics for atoms. Let us choose the assumption that the interaction potential U is continuous and differentiable. The force acting on the i particle can thus be written as an equation 2.4

$$f_i = -\frac{\partial U(r^N)}{\partial r_i}, \quad i = 1, \dots, N. \quad (2.4)$$

In molecular dynamics, we are focused on the time development of the model. In other words, we are looking for the trajectory of the solution of the respective systems of differential equations. In Newtonian mechanics, acceleration is directly related to forces through the equations of motion. Formally, we can write the equation 2.5

$$\ddot{r}_i = \frac{f_i}{m_i}, \quad i = 1, \dots, N, \quad (2.5)$$

where the second time derivative of the positions appears on the left side. The equation 2.5 is a system of $3N$ ordinary differential equations for a set of N atoms. As initial conditions, we usually choose the knowledge of all atomic positions r_i and velocities \dot{r}_i at the initial time $t = t_0$.

We solve equation 2.5 using the finite difference method when we track the desired solution in the form of the function $r_i(t), i = 1, \dots, N$, in the time interval $[t_0, t_{max}]$ at discrete points of the form $t = t_0 + kh$, where h is the integration step and k is a non-negative integer.

To find a solution, it is necessary to calculate the forces acting on individual particles at each step of the simulation. One of the methods that is applied in this area is the Verlet integration method. It is a simple and very effective method that provides sufficiently accurate results in the physico-chemical context. Its great advantage is the time-reversibility and the conservation of the total energy of the system [23].

2.3.1 Verlet integration

Verlet integration method is a numerical method for integrating the equation 2.4. We express the second derivative using finite differences. From the second-order Taylor expansion $r_i(t \pm h)$ centred at t , we obtain the formula

$$\ddot{r}_i = \frac{r_i(t-h) - 2r_i(t) + r_i(t+h)}{h^2}, \quad (2.6)$$

binding atomic coordinates at three points in a time row ($t - h$, t and $t + h$). We will use this relation to calculate $r_i(t + h)$. By substituting 2.6 into 2.4 we get

$$r_i(t + h) = 2r_i(t) - r_i(t - h) + h^2 \frac{f_i(t)}{m_i}. \quad (2.7)$$

In this formulation, we are able to calculate the new positions at time $t + h$ from knowledge of the forces at time t , the positions of the particles at time t and the previous time $t - h$. The time reversibility of the method is clearly visible here. The advantage is that the force is calculated only once in each step of the simulation. For the position preceding the initial position ($t = -h$), we can use the expansion 2.8

$$r_i(-h) = r_i(t_0) - h\dot{r}_i(t_0) + h^2 \frac{f_i(t_0)}{2m_i}. \quad (2.8)$$

We can then calculate the position at $t = h$ from the known positions of $t = 0$ and $t = -h$. There are not explicit velocities in the equations, algorithm with velocities often used by computation software is called Velocity Verlet algorithm. [23]

2.3.2 Constraint dynamics

When integrating equations of motion, we often impose constraints on certain aspects of molecular geometries. The main reason is to enable using a longer simulation time step. If we simulate with a too large time step, we introduce large errors into the simulations, leading in extreme cases to a crash of the simulation. The calculated particle positions at time $t + h$ may lead to overlapping of particles, the calculated force acting on the particles may divert from physically reasonable configurations. Conversely, the use of inappropriately short simulation steps reduces the efficiency of the simulations (the most computationally and therefore time consuming element of the simulations is the calculation of the forces when integrating the equations of motion). [23] The criterion determining the optimal step length is the accuracy of the conservation of total energy. The step length can be determined by an Nyquist-Shannon [26] sampling theorem that says that the time step must be half or less of the period of the quickest dynamics exhibited in the system. Thus, for systems containing very light hydrogen atoms, we can either artificially increase the mass of the hydrogen atom while redistributing the masses of the other molecules to conserve the overall mass of the molecule, or fix the bond angles or bond lengths terminating in the hydrogen atoms. It is the fixation of hydrogen bond lengths that is most often implemented in the Verlet method, using an algorithm called SHAKE.

The SHAKE algorithm [27] is based on Verlet's integration method and is iterative. The first step is to initialize the initial velocities and positions of the atoms, then calculate the positions using Verlet's method without considering bond length constraint. We then create the λ correction of the atom positions to constrain the bond length. Fixing the bond length allows us to use a longer time step (we are no longer limited by the motion of very light particles) and, unlike fixing angles, does not introduce large deviations in the simulations. [23]

2.4 Measuring the properties

2.4.1 Statistical ensembles

MD gives us insight into the total energy of the system, which is naturally conserved in simulations. However, real systems are rarely thermodynamically closed and we often expose the system to external pressure or heat exchange. Thus we speak of different statistical ensembles depending on the conservation of quantities.

As already mentioned, the natural ensemble is the so-called microcanonical ensemble, *NVE*. Thermodynamically it is an adiabatically isolated and closed system, there is no heat exchange while maintaining the total number of particles, total energy and volume.

Another commonly used ensemble is the canonical ensemble, often referred to as *NVT*. Here, the system is maintained at constant temperature (T) and volume (V), allowing for heat exchange with a thermostat.

For many chemical systems, particularly those involving reactions in solution or under variable pressure conditions, the *NPT* ensemble is preferred. In this ensemble, both temperature and pressure are kept constant, ensuring that the system remains in equilibrium with its environment. The *NPT* ensemble is valuable for studying systems where pressure effects play a significant role.

2.4.2 Mean Square Displacement

The Mean Square Displacement (MSD) is a measure of the average displacement of particles in a system over time, used to analyze the diffusive behavior of particles. The average squared distance traveled by particle is measured, Equation 2.9

$$MSD(t) = \frac{1}{N} \sum_{i=1}^N \langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \rangle, \quad (2.9)$$

where $\mathbf{r}_i(0)$ is the initial position and $\mathbf{r}_i(t)$ is the particle position at time t . We can obtain diffusion coefficient by applying the following Equation 2.10 [28].

$$D = \frac{\text{MSD}(t)}{2dt} \quad (2.10)$$

where d is the dimensionality of the track, $d=3$ in three dimensions.

2.4.3 Radial Distribution Function

When we are studying the contacts in liquids, the basic structural quantity is radial distribution function (RDF). As distribution function, it describes the probability of finding a particle at distance r of some other particle. [23]

When we want to extract the RDF from MD simulations, the following strategy is applied. First we have to choose the particle, which surrounding we want to study. Then the space around this particle is divided into concentric spherical shells each having a specific radius. The number of particles inside each shell is counted and the histogram created is then normalized by the volume of each shell and total number of particles. For uncorrelated particles, ideal gas or liquid particles at long distances, the RDF is equal to one.

The first peak of RDF corresponds to the nearest neighbors of the particle. The position on x -axis is determined by the contact distance. The following minima correspond to the steric shielding of the first neighbours. The second peak has the meaning of the second shell of neighbors and so on for later peaks. The amplitude of the peaks corresponds to the relative occurrence in the mixture, the more contacts in the mixture the probability (height of the peak) is higher. [23]

When we want to extract more information from RDF and study how many particles are on average to the specific distance, we are talking about the cumulative RDF (CRDF). The CRDF is a function of coordination number on distance, that we can obtain by integrating the RDF from 0 to desired distance, usually to the first neighbor cutoff. The integration is shown in Equation ??, where $g(r')$ is the RDF. The number of particles within the radius r is then CDRF multiplied by density. [23]

$$\text{CRDF}(r) = \int_0^r 4\pi r'^2 g(r') dr' \quad (2.11)$$

2.4.4 Block averaging scheme

When we are extracting the properties such as volume, density or energy, we are handling with average value obtained from MD simulation. MD simulation data often exhibit correlations between consecutive frames due to the nature of the molecular dynamics algorithm. When we are estimating the uncertainties or errors associated with calculated properties, we have to consider that the simulated data might be correlated and use appropriate statistical methods.

The often used method to evaluate the standard errors is the block averaging scheme. The trajectory data generated from the MD simulation are divided into non-overlapping blocks of equal length. We are assuming, that data within those blocks are no more correlated. For each block, the average value of the property and its standard deviation is calculated from the simulation data. Finally, the averages and standard deviations from all blocks are combined to obtain an overall estimate of the uncertainty in the property of interest. The algorithm used in this work is described more into details by Allen [22].

3 COMPUTATIONAL METHODS

LAMMPS software [29] (version 5 May 2020) was used for all molecular dynamics calculations. The placement of molecular chains in the simulation boxes was done by Packmol [30], the chains were randomly distributed in the space of a cubic box preventing the overlap. The input files for the LAMMPS software were generated using the ffscript [31] script written in the Python programming language.

We also used periodic boundary conditions in the directions of all axes and the velocity Verlet integrator. Contributions of long-range charge interactions of distant atoms were calculated using the long-range solver using the particle-particle-particle mesh (PPPM) algorithm [32]. The bonds and angles were considered as harmonic oscillators, and for dihedral angles, OPLS (Optimised Potentials for Liquid Simulations) was used for every term. Coulombic point charges and the Lennard-Jones potential were used to model the particles interaction the cut-off distance for dispersion and Coulombic interactions was set to 12 Å. The SHAKE algorithm [27] was applied to constrain the lengths of covalent bonds that terminate in hydrogen atoms. The simulations were run under NpT conditions using the Nosé-Hoover thermostat and barostat [33], with relaxation times for temperature control as 100 fs and pressure control as 1000 fs. The simulations contained around 25 000 atoms in a simulation box. From previous research, this was considered to be a suitable setting. [9]

All-atom non-polarisable force fields were used during MD simulations, the parameterisation of the PLA force field was obtained from the literature [34], the parameterizations of APIs were also taken from the literature. [35]

3.1 Simulations of neat PLA

The PLA was first simulated separately from two initial conformers, a fibrilar and a globular chain. We try to verify that the resulting state of the system does not depend on the initial conformations that were used for the simulations. The Cartesian coordinates of the positions of atoms forming a globule were obtained from the last frame of a simulation of one linear polymer molecule in a large virtual empty box.

The initial simulation of the polymers led to the equilibration of the system. The simulation was carried out at 500 K and 1 bar in three blocks with a gradually increasing simulation step. First with a step 0.2 fs for 0.5 ns, followed by steps 0.4 fs and 0.7 fs each for 0.7 ns, then 1 ns of simulation with step a 1 fs.

3.2 Simulations of neat APIs and mixtures with PLA

We started all simulations with an equilibration simulation run from randomly packed simulation boxes under the temperature of 500 K and pressure of 1 bar in three blocks with a gradually increasing time-integration step. The simulation began with an equilibration procedure using a step of 0.25 fs, followed by steps 0.5 and 0.75 fs each for a simulation time of 0.5 ns, then 1 ns of simulation with a step of 1 fs. From this point, we cooled the system, down to 300 K over 2 ns with a step of 1 fs. After this cooling, we continued with a 10 ns long production run with a temperature of 300 K and pressure equal to 1 bar.

From these production runs, we evaluated the MSD (Mean Squared Displacement) of API and PLA molecules in the mixture and the RDF (radial distribution function) of atom interactions. We also performed a production run at the higher temperature of 500 K and the same pressure of 1 bar starting from conformations after the first equilibration under 500 K. We also sampled RDFs and MSDs. Those simulations were performed for neat APIs, PLA polymer, and mixtures with different concentrations of API. For each API, the concentration ratio API:PLA in terms of the number of molecules in a simulation box was 100:17, 200:17, and 300:17. The corresponding molar and mass fractions are available in the Table 1.

Table 1: The concentration of API in mixtures with PLA, expressed in molar and mass fractions.

N_{API}	N_{PLA}	x_{API}	w_{nap}	w_{cbz}	w_{ibu}	w_{indo}
100	17	0.85	0.086	0.088	0.078	0.13
200	17	0.92	0.16	0.16	0.14	0.23
300	17	0.95	0.22	0.22	0.20	0.30

To determine the glass transition temperature (T_g) of the mixtures, we performed simulated annealing simulations with a gradually decreasing temperature cooling rate (30 K ns^{-1}) starting at 800 K and ending at 200 K. Systems containing a mixture of API and PLA were first heated from 500 K to 800 K over 2 ns. To have statistically more reliable data, simulated annealing simulations were performed from 5 different initial conformations. To obtain those conformations a 5 ns long simulation at 800 K was done sampling atomic coordinates within the image of the box every 1 ns.

4 RESULTS AND DISCUSSION

The results section is divided into three logical blocks. First, simulations of the neat polymer were performed to verify the force field and to determine the key properties of the polymer. Subsequently, a series of neat API simulations were performed. Then, simulations of mixtures of different API concentrations in polymer were performed to see the behavior of the mixtures.

4.1 Simulations of neat PLA

4.1.1 Structural properties

First, the simulations from 2 initial states were done (fibrillar and globular). The Figure 7 shows an example of the shortest polymer in a fibrillar conformation state, whereas Figure 8 displays globular conformations of chains containing 20 and 200 monomer units on the left and right, respectively.

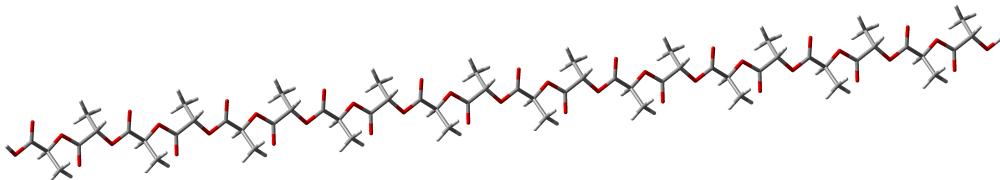


Figure 7: Fibrillar PLA polymer chain, 20 units



Figure 8: Globular PLA polymer chain, 20 and 200 units

To see the effects of initial conformations, molecular weight and thermal history of the polymer on bulk densities following simulations were performed. The design of the simulation boxes is in Table 2. The simulations were first performed at the temperature of 500 K (run 1), then the box was heated up to 1000 K followed by re-cooling and simulating at a temperature of 500 K (run 2). All subsequent simulations were observing the system for 10 ns with a step 1 fs at 1 bar. From

these simulations, the densities and the root mean square distance of the polymer chain termini with their standard deviations were evaluated.

Table 2: Design of the neat PLA simulation boxes

N_{units}	n_{chains}	N_{atoms}	$M, \text{ g mol}^{-1}$	$d, \text{ \AA}$
20	140	25620	1459.3	67.9
40	70	25410	2900.5	67.7
60	40	24978	4341.8	64.3
80	35	25305	5783.1	67.7
100	28	25284	7224.3	67.7
120	23	24909	8665.6	67.3
140	20	25260	10106.8	67.6
160	17	24531	11548.1	67.0
180	15	24345	12989.4	66.8
200	14	25242	14430.6	67.6

The following graphical representation of the results (Figure 9) shows a trend of an increasing density depending on the length of the chain (molecular weight), which is independent of the initial conformation of the molecule, the values differs only in . It is also visible that densities for longer chains converge to a constant value. From this finding, we can consider that a system above $M_w = 9\,000 \text{ g mol}^{-1}$ has reached the polymer limit within the simulation.

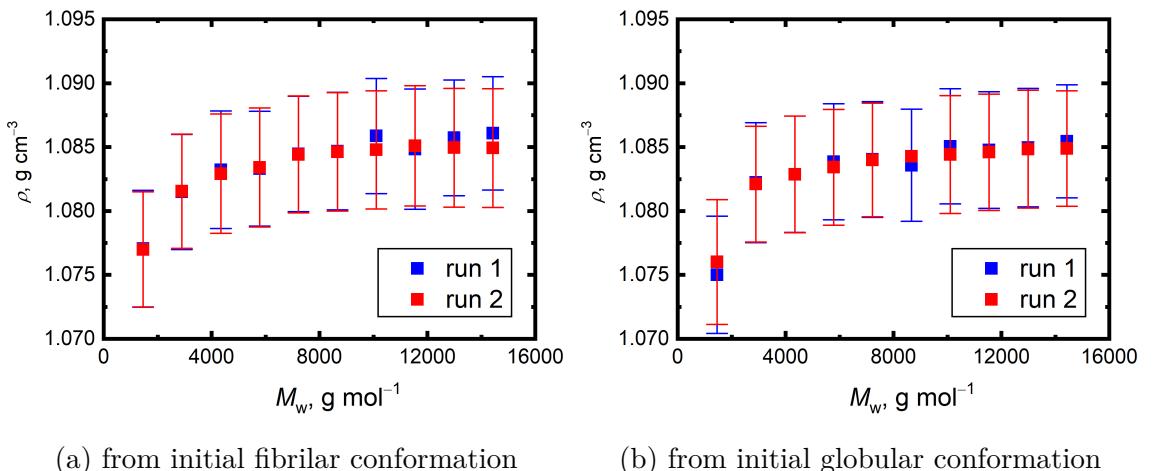


Figure 9: Average densities with their standard deviations for PLA polymer chains at 500 K and 1 bar as a function of the chain length before (run 1) heating and (run 2) after cooling.

From the density data, any impact of the conformation memory cannot be assessed, since the bulk density is too a crude point of view on the polymer structure. That is the reason why the distances of the polymer chain termini that are displayed in the

Figure 10 were calculated. The simulated time of 10 ns at the elevated temperature 1000 K was not enough to completely erase the polymer conformational memory, there is a noticeable deviation of the data sets obtained from the simulation initiated from the fibrilar and globular conformations.

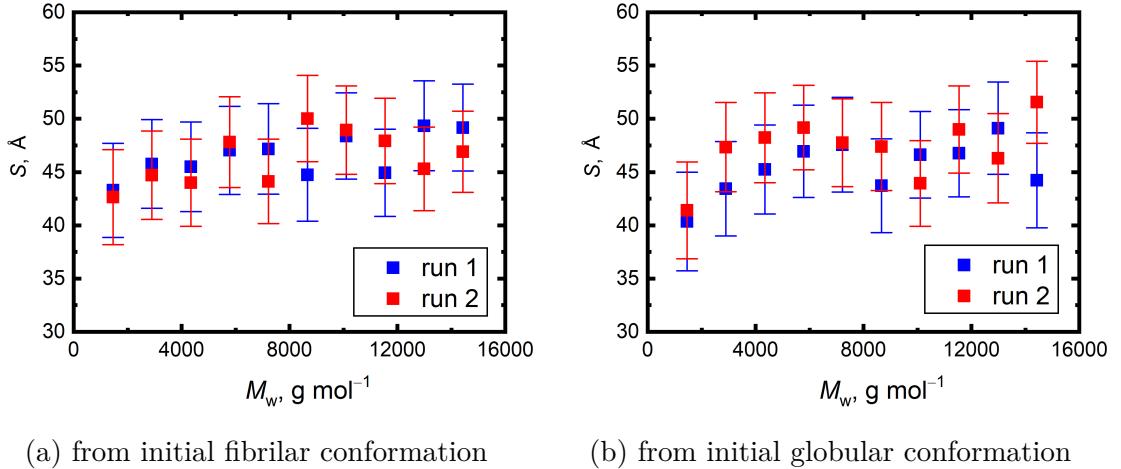


Figure 10: Root mean square distances of the polymer chain termini with their standard deviations for PLA polymer chains as a function of chain length before (run 1) heating and (run 2) after recooling.

The dynamics of the polymer is slow due to complex entanglement of individual chains even at elevated temperatures, for a complete loss of the conformational memory it would be necessary to simulate the system for a longer period of time, especially for longer chains.

To investigate the effect of the box size on the molecular simulations, the simulations with boxes containing 5-50 polymer chains inside, each having a molar mass $12\ 989\ \text{g mol}^{-1}$ were performed. The simulations were performed with a 3-block initial equilibration at 1000 K and 1 bar and subsequently cooled to 500 K and simulated for 10 ns. Mean densities obtained from these simulations are shown in the Figure 11a on the left. The simulation results prove that the size of the box has no significant effect on resulting average density. However, there is a visible effect of a lower uncertainty of standard deviations of densities for larger boxes. To have a better insight what is going on the structural level, we also analyzed the root-mean-squared end-to-end distance of the polymer chain termini shown in Figure 11b on the right. From the obtained data, there is visible growing trend in end-to-end distances. That means that the small boxes are not sufficient to represent correctly the distribution of the end-to-end distances in the polymer. We can say, that its value converges for boxes containing more than 40 polymer chains. As a conclusion, when we are focused only on macroscopic properties such as density, we can use smaller number of chains in

the simulation boxes in order to save the computational resources. However, when dealing with structural properties, we should be careful and set the number of chains in a simulation box more carefully.

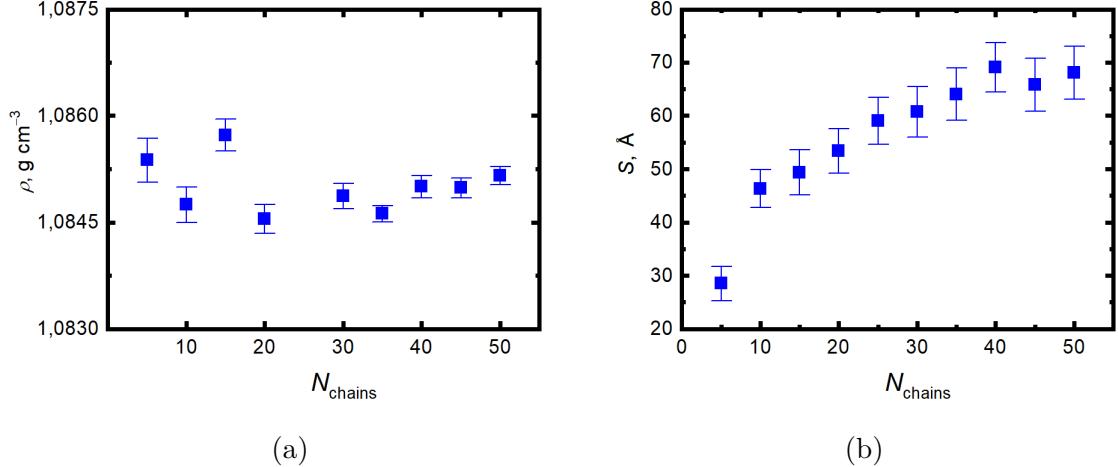


Figure 11: (a) dependence of density on the number of PLA polymer chains containing 90 dimer units in the box on the left and (b) dependence of root mean square distances of the polymer chain termini with their standard deviations on the right

To assess the accuracy of the calculated densities, the corresponding experimental data was obtained from literature. The comparison for two temperatures (300 and 500 K) in following Table 3 was taken from [9], where the methodology is also described in more details.

Table 3: Comparison of calculated and experimental PLA densities taken from [9].

$T, \text{ K}$	$\rho_{\text{exp}}, \text{g cm}^{-3}$	$\rho_{\text{MD}}, \text{g cm}^{-3}$	$100(\rho_{\text{MD}}/\rho_{\text{exp}} - 1)$
300	1.373 ± 0.003	1.193 ± 0.001	-13
500	1.181 ± 0.003	1.086 ± 0.001	-8.0

The experimental density value for 300 K is 13 % above the computed, for 500 K the value is 8 % above the computed density. In both cases, the value is higher than our limit value for the longest PLA polymer chain composed of 200 units.

4.1.2 Polydispersity effect

Under real conditions, it is hardly possible to experimentally prepare a monodisperse polymer containing only one selected chain length. For this reason, we simulated several polydisperse systems, each exhibiting a different distribution of molar masses of individual molecular chains, containing a total number of 50 chains. Their lengths come from an interval 8-244 units. Polydispersity index PDI = 1 corresponds to 50

chains of length 124 units, the other values were calculated using the equation 4.1. We than build composition of the other systems using the Gaussian distribution with the mean value of 124 units. By this procedure, we were able to obtain the PDI up to 1.3. To reach PDI 1.4 and 1.5, we had to increase the numbers of very short and long polymer chains present in the box to get the desired PDI. Designed compositions of systems based on PDI values are displayed in the Figure 12. The density was again evaluated from the simulations, in this case as a function of the PDI.

$$\text{PDI} = \frac{\sum_i N_i \cdot \sum_i N_i M_i^2}{(\sum_i N_i M_i)^2} \quad (4.1)$$

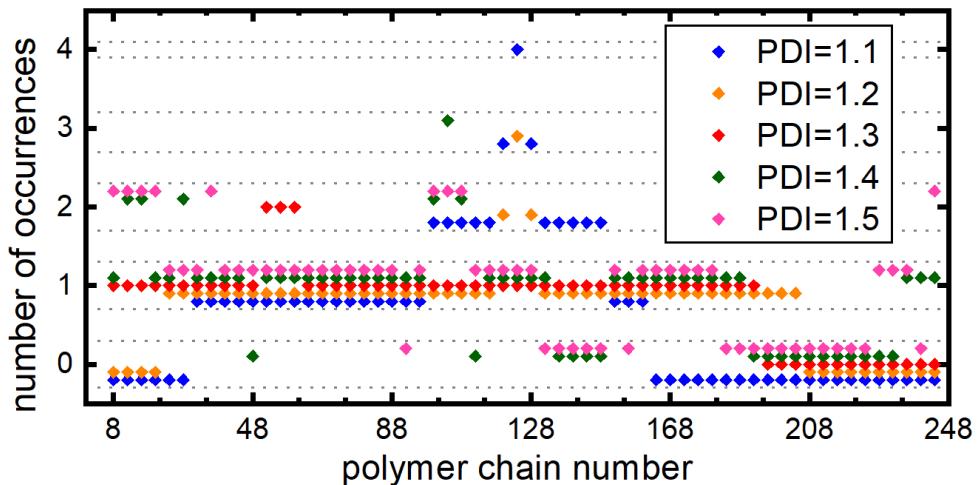


Figure 12: Number of occurrences of chains of a given length in the polydisperse systems with corresponding PDI.

Calculated densities depending on the PDIs are in the Figure 13. There is no significant difference among the densities obtained for different values of the polydispersity index. For the subsequent simulations we will use monodisperse systems as it will not introduce deviations from macroscopic behavior.

4.1.3 Glass transition modeling

Understanding the glass transition temperature of pharmaceutical materials is crucial for drug formulation, storage, and delivery. The T_g refers to the temperature at which an amorphous material transitions from a rigid, glassy state to a more flexible, rubbery state without melting. The knowledge of T_g helps us predict and control the stability of drugs during long-term storing.

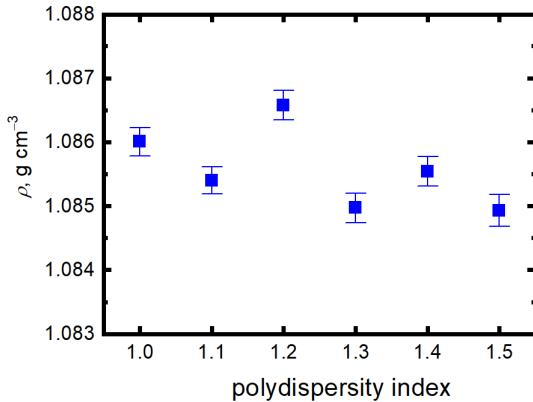


Figure 13: Density of PLA at 500 K and 1 bar extracted from MD simulations depending on the polydispersity coefficients.

To obtain the glass transition temperature of PLA (T_g), we started by simulating ten different polymer chain lengths for a set of temperatures in a range 140-485 K with a step of 15 K. For the system equilibration, each polymer chain simulation ran for 1 ns with a step of 1 fs. After the equilibration, the temperature interval for the following simulations was limited to the range of 200-485 K with a step of 15 K. Then, we run a NpT production simulation for each temperature with this limited interval for 10 ns with a step of 1 fs from which the equilibrium bulk phase densities were calculated and displayed as a function of temperature. To get the glass transition temperature (T_g) for PLA from volumetric data, the trend shift method was used [36].

The trend shift method is illustrated for two different lengths of the polymer chains in the Figure 14. When we plot the bulk phase densities as a function of temperature, there is a visible trend shift which has the meaning of the transition between the glassy state and the rubbery state of a polymer. The point of the transition divides the density dataset into two intervals. Coordinates of the breakpoint and one surrounding point from each side were excluded from the further processing and the two resulting intervals were interpolated by a linear function. The x -coordinate of the intersection of these two lines determines the glass transition temperature. For this purpose a script in Maple software solving the system of linear equations was created to get the T_g .

Available experimental T_g of L-PLA varies in the literature, the value of $T_g = 333 \pm 4$ K [37] was chosen as a reference value. The variation of experimental T_g data is caused by the stereo-configuration along the PLA chains, crystallinity of bulk PLA and the polydispersity. [9] The T_g value obtained in this work as an arithmetic

mean of the values for all polymer chain lengths is $T_g = 337 \pm 10$ K. The calculated T_g for each polymer chain are displayed as function of polymer chain weight in the Figure 15 with their arithmetic mean value drawn as a blue line. For each length of the polymer chain the regions for fitting were set manually. In the paper written by Klajmon et al[9], the data value was obtained from the same data sets using automated script. Their resulted value is $T_g = 351 \pm 11$ K. This corresponds with the fact that trend shift method is really sensitive to the intervals we determine, this is also illustrated in Figure 14. In this paper, there is also comparison with other methods used to determine T_g and also with results for other polymers.

The calculated T_g value is very close to the experimental one, there is no significant trend in T_g of polymer chain length of PLA from the calculated data. However, that small deviation between calculated and experimental value is more luck, when we compare the data for other polymers the difference is much larger.

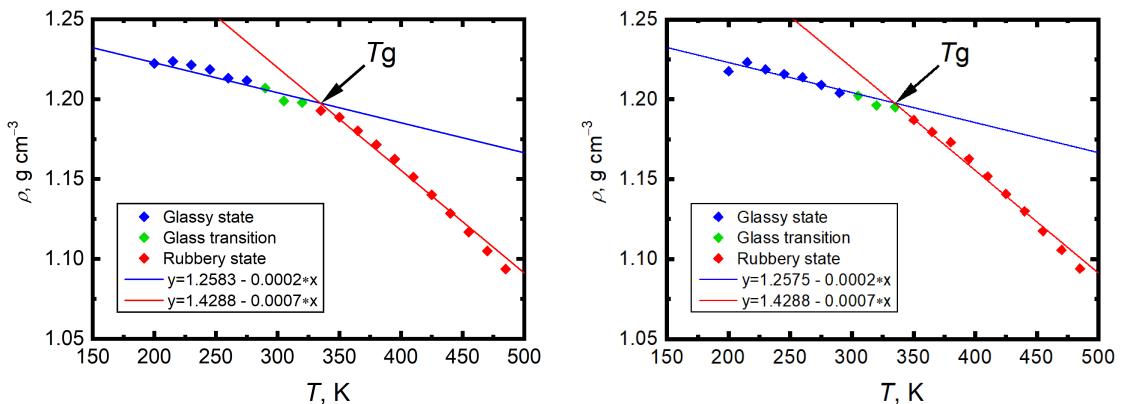


Figure 14: Illustration of the trend shift method for two polymer chains of different lengths, 20 units on the left and for 30 units on the right.

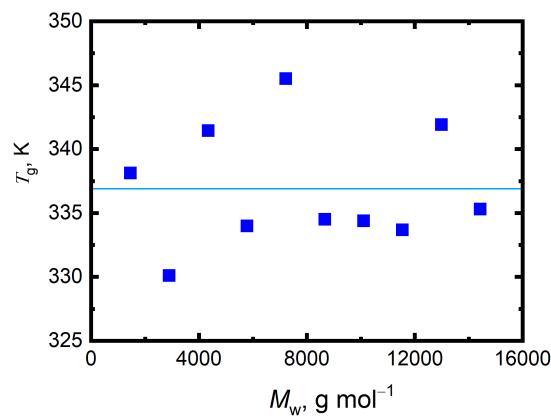


Figure 15: Calculated T_g for all lengths of polymer chains using trend shift method, where blue line represents mean value.

4.2 Simulations of neat API

The parameters of API simulation boxes after the equilibration run ($T=500$ K) are presented in Table 4, showing the number of API molecules (N_{API}), the total number of all atoms (N_{atoms}) and size of the cubic box (l_{box}).

Table 4: Properties of equilibrated simulations boxes for bulk API one-component systems, $T=500$ K.

API	N_{API}	N_{atoms}	$M, \text{ g mol}^{-1}$	$l_{\text{box}}, \text{\AA}$
carbamazepine	800	24 000	236.27	66
naproxen	800	24 800	230.26	66
ibuprofen	800	26 400	206.28	67
indomethacin	600	24 600	357.8	66

To validate the force fields, computed densities were compared with experimental literature values, the comparison is in the Table taken from literature. UPRAVÍM [35]

Table 2. Comparison of simulated and experimental bulk phase densities ρ ($\text{g}\cdot\text{cm}^{-3}$).

Compound	Phase	Temperature, K ^a	ρ_{MD}	ρ_{exp}	$100(\rho_{\text{MD}}/\rho_{\text{exp}} - 1)$
Carbamazepine	Crystal III	293	1.335	1.333 [41]	0.1
Ibuprofen	Crystal I	296	1.115	1.117 [40]	-0.2
	Liquid	350	1.006	0.966 [43]	4.1
		400	0.968	0.924 [43]	4.7
Indomethacin	Crystal α	203	1.408	1.420 [39]	-0.9
	Crystal γ	120	1.418	1.401 [38]	1.2
	Liquid	400	1.284	1.231 [43]	4.3
		450	1.264	1.183 [43]	6.9
Naproxen	Crystal	293	1.308	1.263 [37]	3.6
	Liquid	430	1.154	1.088 [43]	6.1
		480	1.116	1.048 [43]	6.5
Adenine	Crystal	293	1.506	1.494 [35]	0.8
Cytosine	Crystal	293	1.537	1.502 [36]	-1.6

^a Experimental density determination was performed at this temperature.

4.2.1 Sulfathiazole FF parametrization

A newly parametrized force field [38] was used for sulfathiazole. The charges on the atoms were calculated by quantum-chemical calculations using the Gaussian 16 code [39]. At first, the structure was optimized using the B3LYP functional with the aug-cc-pVTZ basis set with the dispersion correction GD3BJ [40]. After finding the optimal geometric structure of the molecule, the charges on the atoms were fitted through the CHELPG (Charges from EElectrostatic Potentials using a Grid-based method)[21] method and inserted into force field file.

The MD simulations of the sulfathiazole crystal (polymorph II) were performed to validate the force field file parameters. The triclinic simulation box and barostat setting was used to account for the anisotropic character of the molecular crystal, other settings are similar to that with the PLA instead of calculating long-range interactions. The sulfathiazole crystal was simulated for three temperatures of 100, 200 and 300 K at 1 bar. The simulation results and experimental data obtained from literature [12] are given in the Table 5. The MD simulations were performed with two different FF. First simulation was performed with FF field taken from literature (index MD1), second simulation with newly parametrized FF (CHELPG index).

Table 5: Comparison of simulated and experimental [12] crystallographic parameters for II polymorph of sulfathiazole. The index MD1 corresponds to force field taken from literature, index CHELPG to newly parametrized FF.

T , K	100	200	300
ρ_{exp} , g cm ⁻³	1.575(2)	1.560(2)	1.540(2)
ρ_{MD1} , g cm ⁻³	1.552(1)	1.534(2)	1.515(2)
$100(\rho_{\text{MD1}}/\rho_{\text{exp}} - 1)$	-1.5	-1.7	-1.6
ρ_{CHELPG} , g cm ⁻³	1.525(1)	1.538(2)	1.495(2)
$100(\rho_{\text{CHELPG}}/\rho_{\text{exp}} - 1)$	-3.2	-1.4	-2.9
β_{exp} , °	94.14(2)	93.905(7)	93.674(8)
β_{MD1} , °	91.56(7)	91.6(1)	91.7(1)
β_{CHELPG} , °	97.78(6)	96.79(8)	97.0(1)
a_{exp} , Å	8.1896(17)	8.2122(7)	8.2427(8)
a_{MD1} , Å	8.302(4)	8.325(5)	8.351(7)
a_{CHELPG} , Å	6.739(5)	6.648(9)	6.84(1)
b_{exp} , Å	8.532(2)	8.5604(9)	8.599(1)
b_{MD1} , Å	8.236(5)	8.275(7)	8.313(9)
b_{CHELPG} , Å	8.267(9)	8.00(1)	8.22(2)
c_{exp} , Å	15.447(4)	15.497(2)	15.563(2)
c_{MD1} , Å	15.99(1)	16.06(2)	16.13(2)
c_{CHELPG} , Å	20.15(2)	20.88(3)	20.35(5)

There is good agreement between the experimental and computed densities, the relative deviations low for both FF. Other parameters differs more between the two FF used. Generally, the MD1 FF describes better the β angle and the parameters of the unit cell within all simulated temperatures. It would be also useful to obtain the data for liquid state and compare them with experiments.

4.3 Simulations of mixtures of APIs and PLA

The simulations of mixtures of 4 APIs with PLA were performed following the scheme described in previous section. The images of the simulation boxes for mixtures with the least amount of API ($x_{\text{API}} = 0.85$) for temperature 500 K are in figure 16. The polymer was represented by quicksurf drawing method and diffused in the background, the API molecules are represented with balls and stick above the PLA molecules.

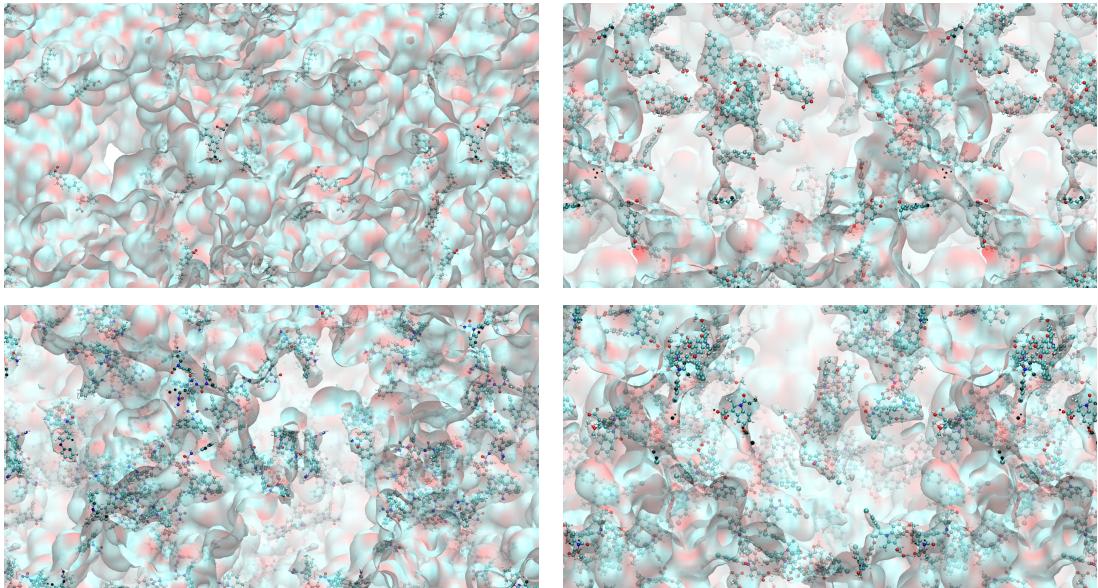


Figure 16: Image of the simulation box for $x_{\text{API}} = 0.85$ and $T = 500 \text{ K}$ for ibuprofen (**top left**), naproxen (**top right**), carbamazepine (**bottom left**) and indomethacin (**bottom right**).

4.3.1 Excess properties

Excess molar energy and excess molar volume are thermodynamic quantities that quantify deviations from ideal behavior in solutions. Ideal solutions exhibit no interactions between components, resulting in linear relationships between the properties of the mixture and its composition. However, real solutions often deviate from ideal behavior due to the interactions between molecules such as hydrogen bonding or their steric composition.

The excess molar energy represents the additional energy per mole of mixture compared to an ideal solution at the same temperature, pressure, and composition. It accounts for the energy associated with interactions between molecules in the mixture, including both attractive and repulsive forces. Positive excess molar energies indicate favorable interactions (hydrogen bonding), while negative values suggest repulsive interactions.

Similarly, the excess molar volume quantifies the deviation in volume per mole of mixture from that of an ideal solution. It reflects the changes in volume resulting from interactions between molecules, such as volume contraction due to strong molecular association or volume expansion due to repulsive interactions.

Excess molar volumes (V^E) of the mixtures were evaluated from the simulations for each concentration using the Equation 4.2. For this reason, the average molar volumes of the simulation boxes were evaluated from simulations of pure API (V_m^{API}) and pure polymer (V_m^{PLA}) and from a mixture (V_m^{MIX}) of API and PLA. To obtain the uncertainties block averaging scheme and error propagation law was used.

$$V^E = V_m^{\text{MIX}} - x_{\text{API}}V_m^{\text{API}} - x_{\text{PLA}}V_m^{\text{PLA}} \quad (4.2)$$

Excess molar energies of the mixtures were evaluated in a similar way by the Equation 4.3

$$E^E = E_m^{\text{MIX}} - x_{\text{API}}E_m^{\text{API}} - x_{\text{PLA}}E_m^{\text{PLA}}, \quad (4.3)$$

where E_m^{MIX} is the average molar energy of the simulation box of the mixture, E_m^{API} is the average molar energy of pure API, and E_m^{PLA} is the averaged molar energy of pure PLA obtained from simulation. Mixing energies and volumes were calculated for all systems and both temperatures (300, 500 K), and the comparison is shown in Table 6.

Table 6: Calculated excess energies (kJ mol⁻¹) and volumes (in cm³ mol⁻¹) for API mixtures of different concentrations from simulations under 300 K (V_{300}^E , E_{300}^E) and 500 K (V_{500}^E , E_{500}^E) with their standard uncertainties (k=1).

API	x_{API}	V_{300}^E	σ_{V^E}	V_{500}^E	σ_{V^E}	E_{300}^E	σ_{E^E}	E_{500}^E	σ_{E^E}
cbz	0.85	-5.47	0.27	0.92	0.50	0.63	0.30	4.74	0.46
	0.92	-3.45	0.14	0.64	0.27	0.54	0.15	3.20	0.21
	0.95	-2.34	0.11	0.22	0.17	-0.09	0.11	2.65	0.16
nap	0.85	-5.04	0.27	3.28	0.48	4.38	0.24	7.49	0.44
	0.92	-2.16	0.16	2.61	0.26	3.88	0.24	6.26	0.26
	0.95	-1.81	0.10	1.46	0.18	3.39	0.10	4.72	0.20
indo	0.85	-7.33	0.26	-1.24	0.54	5.20	0.25	8.55	0.65
	0.92	-3.85	0.13	-1.73	0.28	4.70	0.16	6.37	0.29
	0.95	-2.96	0.17	-2.00	0.22	2.04	0.15	3.53	0.31
ibu	0.85	-9.63	0.33	0.25	0.49	-1.27	0.35	3.28	0.42
	0.92	-6.45	0.21	0.11	0.29	-0.34	0.17	2.63	0.24
	0.95	-5.29	0.12	-0.14	0.17	0.23	0.16	2.12	0.15

For higher temperature of 500 K all excess energies are positive, this indicate that addition of the polymer can attenuate the API-API interactions, but those are not fully compensated by creating new intermolecular interactions between API and PLA in the mixtures. We can see, that for higher concentration of API in the mixture (x_{API}) the values of E^{E} are decreasing also with their uncertainties. The reason could be that more interactions of API-API are still presented resulting in less populated API-PLA interactions that do not fully compensate the API-API cohesion. The analysis of radial distribution functions of the most favorable interactions between API-API and API-PLA was performed in order to see the which specific interactions are forming. Same trend was observed for the height of the first peak for different concentrations. Complete results are given in following section.

The excess volumes are also decreasing with higher concentration of API in the mixture. For carbamazepine and naproxen the V^{E} values are positive, meaning less advantageous arrangement of molecules in space. For indomethacine all values are negative, more advantageous arrangement is observed. For ibuprofen the values are close to zero, for the highest API concentration the value is negative.

For the lower temperature of 300 K all excess volumes (V^{E}) are negative, creating more beneficial space conformation. The lowest value is for indomethacin indicates the strongest interaction, which is in agreement with the height of RDF first peak. For all APIs, there is visible trend, with higher concentration of PLA the V^{E} value is lower. The same trend is observed for excess energies for 3 of 4 APIs (cbz, nap, indo), for ibuprofen the trend of E^{E} is the opposite.

4.3.2 Radial distribution functions

Sampling the radial distribution function from the simulations was done to explore the interactions that are having the highest impact on material cohesion. The hydrogen bonds were mostly studied as the most important cohesive features. First the API-API contacts are discussed, then API-PLA interactions are studied.

The strongest API-API interaction was chosen and plotted to study the change between neat API and simulations of mixtures with different concentrations of PLA. For ibuprofen, naproxen and indomethacine the interaction between the oxygen and hydrogen atoms from carboxyl group was studied. However, carbamazepine does not contain complete carboxyl group, the hydrogen bonding was studied between the NH_2 and OC group. The selected atom types involved in interactions are visualised in the introduction section in Figure 2.

All RDFs of mixtures containing the API-API interaction were scaled onto the pure API RDF signal enabling us direct comparison of amplitudes of individual signals. The following Equation 4.4 was used.

$$\text{RDF}_{\text{scaled}} = \text{RDF} \cdot \frac{V_{\text{API}}}{N_{\text{API}}} \frac{N_{\text{mix}}}{V_{\text{mix}}}, \quad (4.4)$$

where V_{API} is the average volume of the pure API simulation box, N_{API} is the number of molecules in the pure API simulation box and analogously for mixtures.

API-API interactions

The RDF of the hydrogen bonding interaction in carboxyl group of indomethacin is shown in Figure 17. The dimer of the carboxyl groups is the strongest in indomethacine in comparison to other studied APIs. The closest contact distance between O-H \cdots O is 1.5 Å, with the first peak amplitude decreasing with less API in the mixture, for higher concentration of polymer in the mixture API-PLA interactions are more favorable. First coordination sphere contain one molecule, corresponding to the close dimer contact.

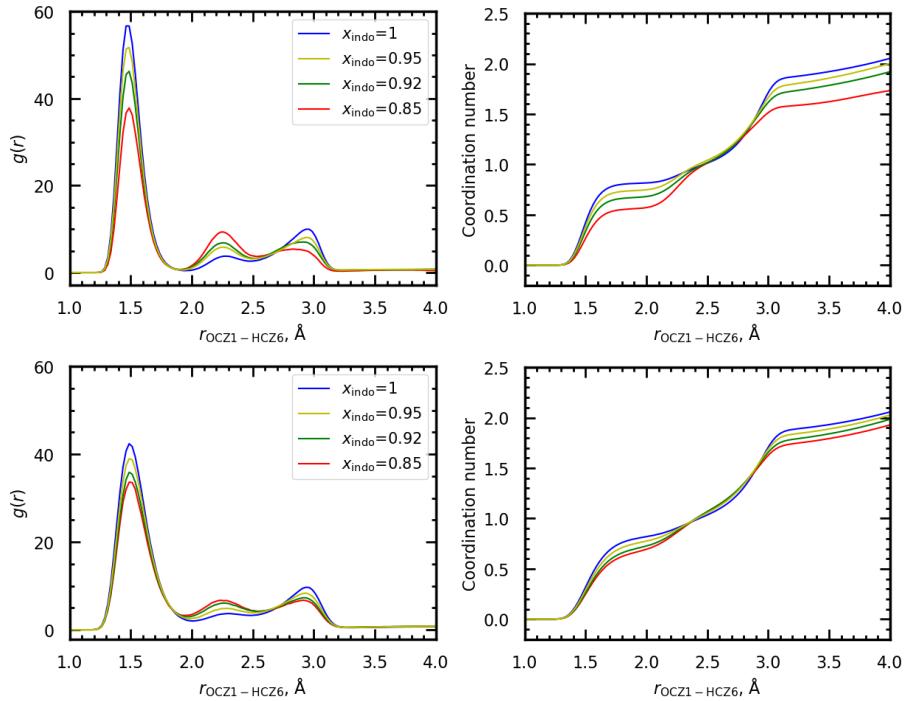


Figure 17: RDF of the API-API interaction between HCZ6 hydrogen atom and OCZ1 oxygen atom from COOH group in a mixture of indo and PLA for different concentration normalized on values for pure indo, temperature of 300 K in the left upper corner and 500 K bottom left, coordination numbers on the right.

The RDFs of the interaction in carboxyl groups of naproxen and ibuprofen are shown in Figure 18 and 19. The contact distance of O-H \cdots O is 1.75 Å in both cases. We can observe interesting trend for both API. In pure API, the dimer is linked by two hydrogen bonds, corresponding to two massive peaks of the function. With the higher concentration of PLA in the mixture, the shape of the RDF is changing. This change resulted in higher probability of the contact at longer distances, meaning that APIs prefer looser contact of their COOH dimers. The reason for that could be that first sphere is taken by interaction of hydrogen from API with oxygens from PLA which happen on distances 2 and 4 Å. This contact is described further in the PLA-API interaction part.

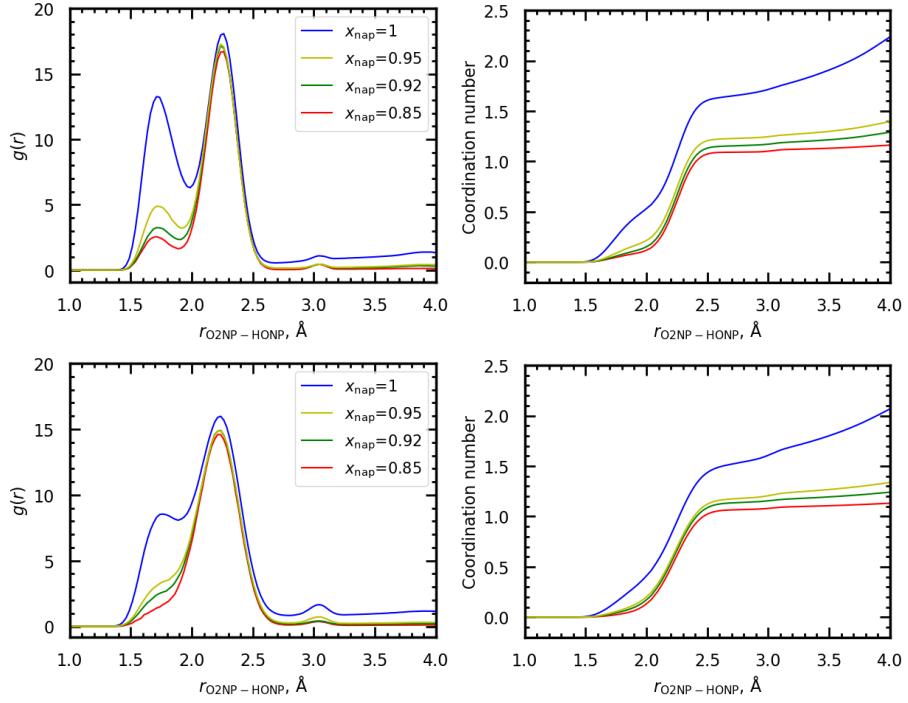


Figure 18: RDF of the API-API interaction between HONP hydrogen atom and O2NP oxygen atom from COOH group in a mixture of nap and PLA for different concentration normalized on values for pure nap, temperature of 300 K in the left upper corner and 500 K bottom left, coordination numbers on the right.

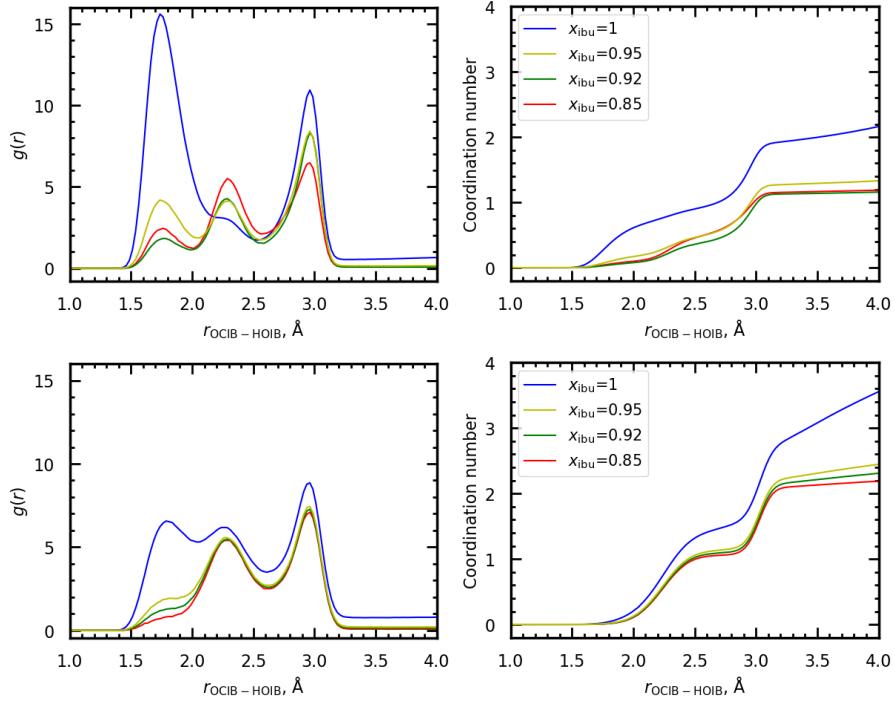


Figure 19: RDF of the API-API interaction between carboxyl HOIB hydrogen atom and OCIB oxygen atom in a mixture of ibu and PLA for different concentration normalized on values for pure ibu, temperature of 300 K in the left upper corner and 500 K bottom left, coordination numbers on the right.

The RDF of the hydrogen bonding interaction of carbamazepine is shown in Figure 20. There are two peaks presented, that correspond to interactions of 2 equivalent hydrogen atoms bonded on nitrogen, each participating the same linking the dimer with 2 hydrogen bonds. The contact distance $\text{O-H}\cdots\text{O}$ is 2.25 Å. The shape and signal response of the peaks is similar for neat API and mixtures for both temperatures. There is almost no change between different concentration of API, meaning that the impact of PLA on the cohesion of carbamazepine in the mixture is very low. The strength of API-API intermolecular interactions decreases in order of the APIs were discussed, indo < nap < ibu < cbz.

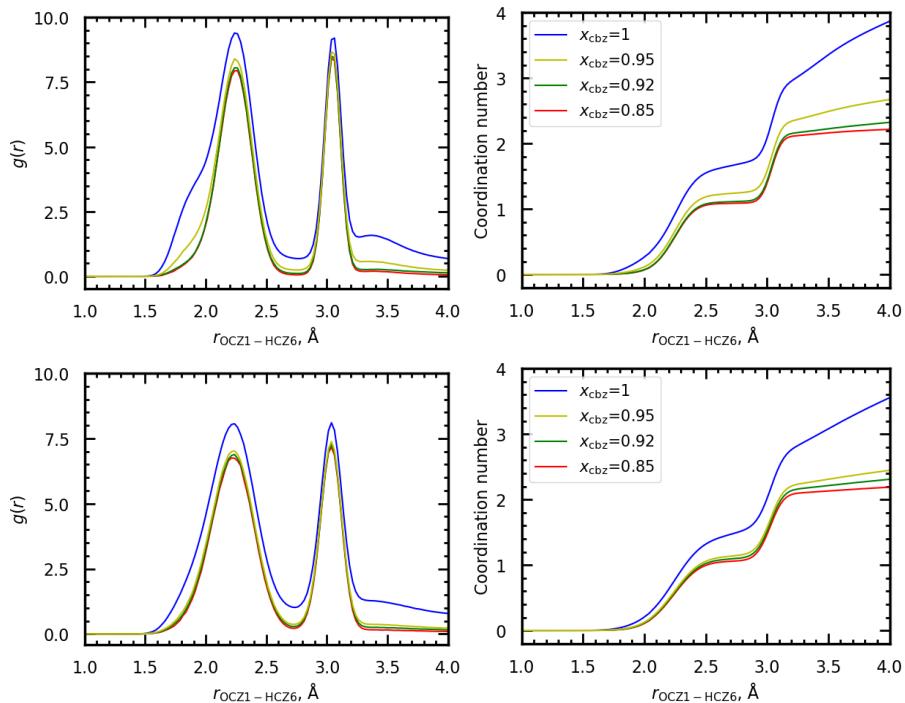


Figure 20: RDF of the API-API interaction between HCZ6 hydrogen atom bonded on nitrogen and OCZ1 oxygen atom in a mixture of cbz and PLA for different concentration normalized on values for pure cbz, temperature of 300 K in the left upper corner and 500 K bottom left, coordination numbers on the right.

API-PLA interactions

In this part we are focusing on the hydrogen bonding between API and PLA, where API acts as a donor of hydrogen and PLA as its acceptor. Optimization of the conformation was performed for each dimer composed of two units long PLA chain and one API molecule using quantum methods in Gaussian[39] software by B3LYP functional with the 6-31+g(d,p) basis set with the dispersion correction GD3BJ [40]. The optimized configuration with studied atom types marked is available in Figure 21.

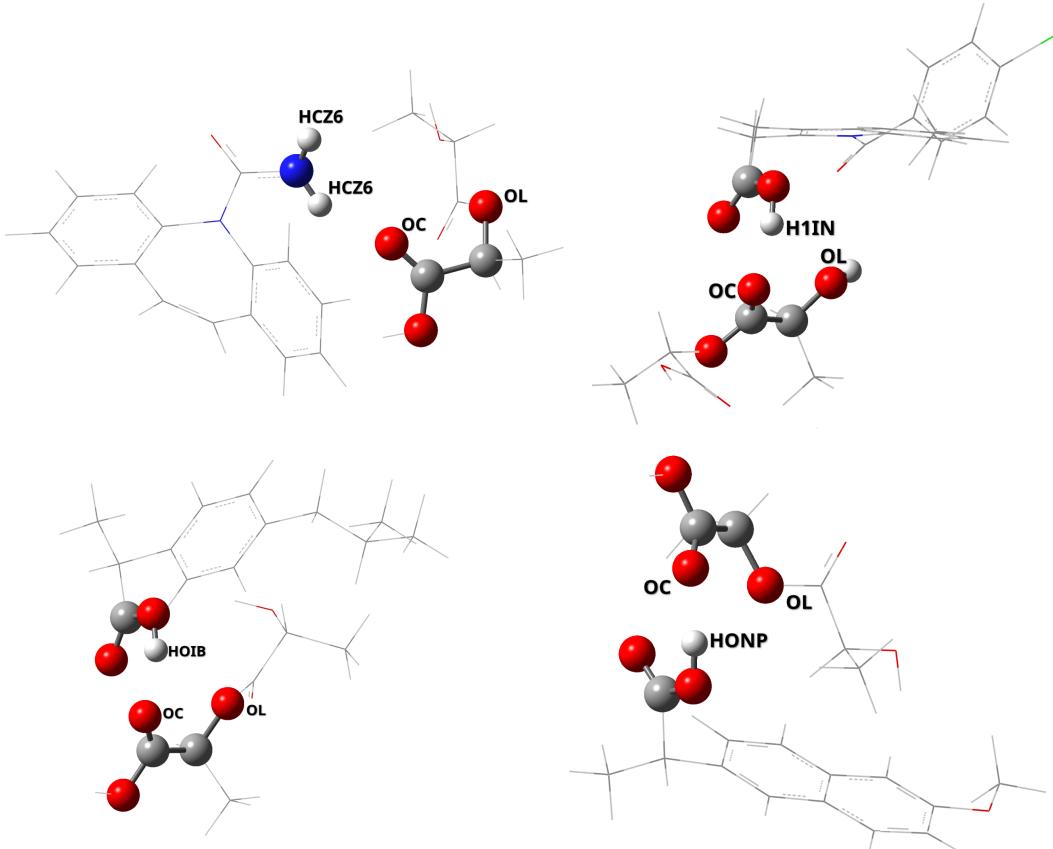


Figure 21: Visualization of the interaction between hydrogen atom from API and oxygen atom from PLA with marked atom types involved. Carbamazepine (**top left**), indomethacin (**top right**), ibuprofen (**bottom left**) and naproxen (**bottom right**).

The RDF of hydrogen bonding with the carbonyl group in PLA (OC) is shown in Figure 22. The RDF value converges to one in a long distance for all API-PLA interactions, which is in good agreement with theory. For indomethacine, the amplitude of the peak decrease with lower PLA molecules presented in the mixture. This correspond with the observation for trends in INDO-INDO interactions. The contact distance in dimer obtained from quantum optimization is 1.76 Å, from MD simulation we obtained the value of 1.75 Å.

In case of naproxen, the RDF data shows one peak at distance of $\text{? } \text{\AA}$, the value obtained by quantum simulation is 1.76 \AA . The peak amplitude is relatively high compared to other APIs. The RDF data for ibuprofen exhibit similar behavior, with slightly higher peak amplitude. For the temperature 300 K , there is different order of heights of the peaks. The peak corresponding to $x_{\text{API}} = 0.85$ is switched with the one corresponding to $x_{\text{API}} = 0.92$. The difference between those peaks is small, but it strange compared to trend observed in IBU-IBU RDF. The shift on x -axis is same as for naproxen, the calculated QM value is 1.77 \AA for IBU.

For carbamazepine, there are two peaks in the RDF function. The peaks corresponding to those interactions are weak, we can see that the intensity of the peaks is really low. Under temperature 300 K the intensity of the first peak is slightly above one, but for 500 K the first peak is below one, meaning that this interaction occurs less at a small distance than in the rest of the system. The impact of the PLA concentration change is not that visible, especially for higher temperature of 500 K . This also correspond to we saw in the CBZ-CBZ interactions. The contact distance value from quantum calculations is 2.01 \AA , which is the same as the distance of first RDF obtained by MD.

The hydrogen bonds with the oxygen bonded by ether bond in PLA structure are shown in Figure 23. All values also converge to 1 in a long distance for all APIs. Generally, those interactions are much weaker than those with oxygen from carbonyl group. For indomethacine and carbamazepine the value is below one at short distance, with the dimer contact distance obtained from QM 2.55 and 4.19 \AA respectively. For ibuprofen the order of peaks amplitude for 300 K is also different with the peak for $x_{\text{API}} = 0.92$ having the biggest value. For the temperature of 500 K the change in order is no longer visible. The QM distance between API hydrogen and ether-bonded oxygen in PLA for ibuprofen is 2.48 \AA . For naproxen the QM contact distance is 2.59 \AA .

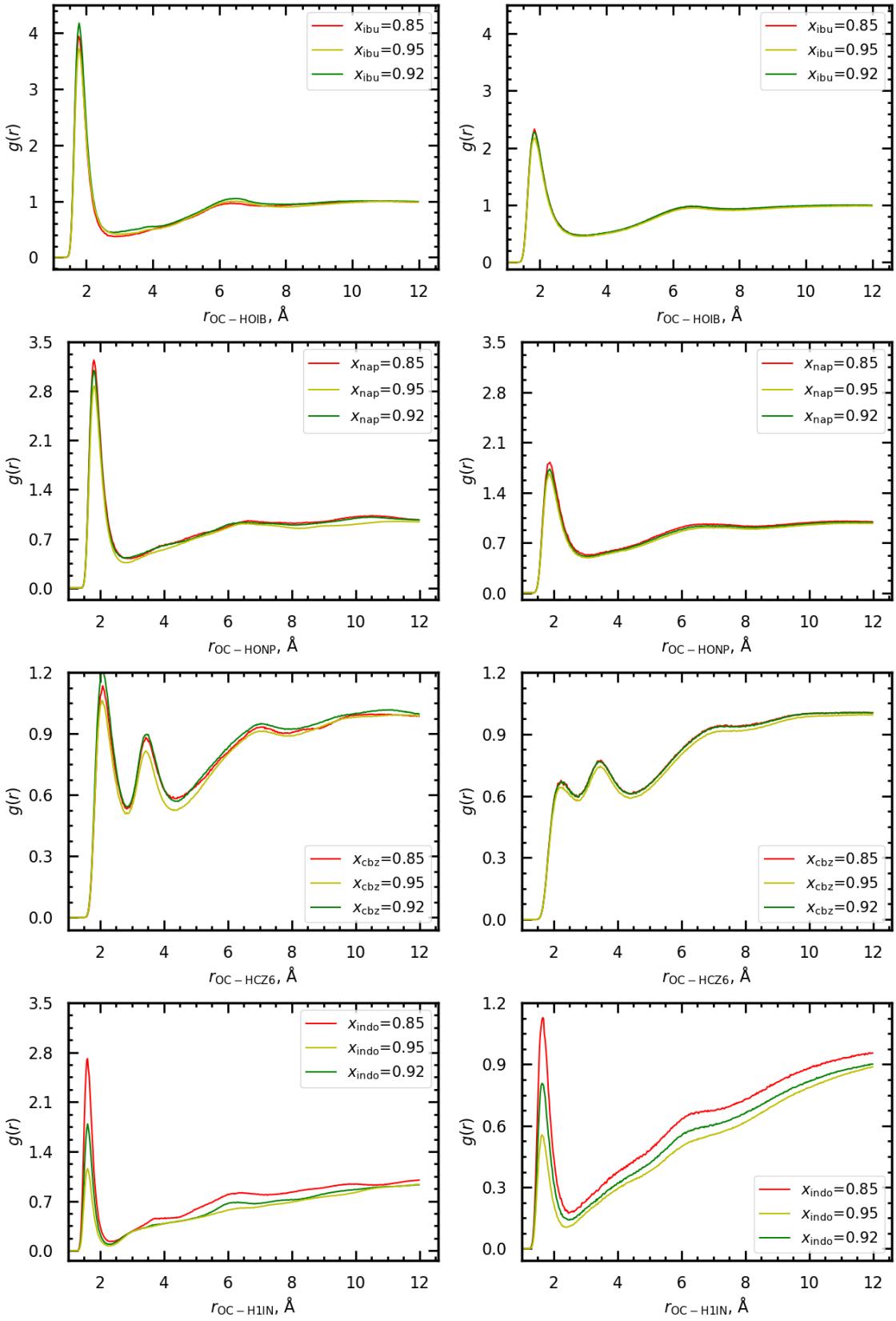


Figure 22: Radial distribution function of the interaction between hydrogen atoms and oxygen atom from carbonyl group in PLA, first ibuprofen on top, second naproxen, third carbamazepine and indomethacine in the bottom, temperature 300 K on the left and 500 K on the right.

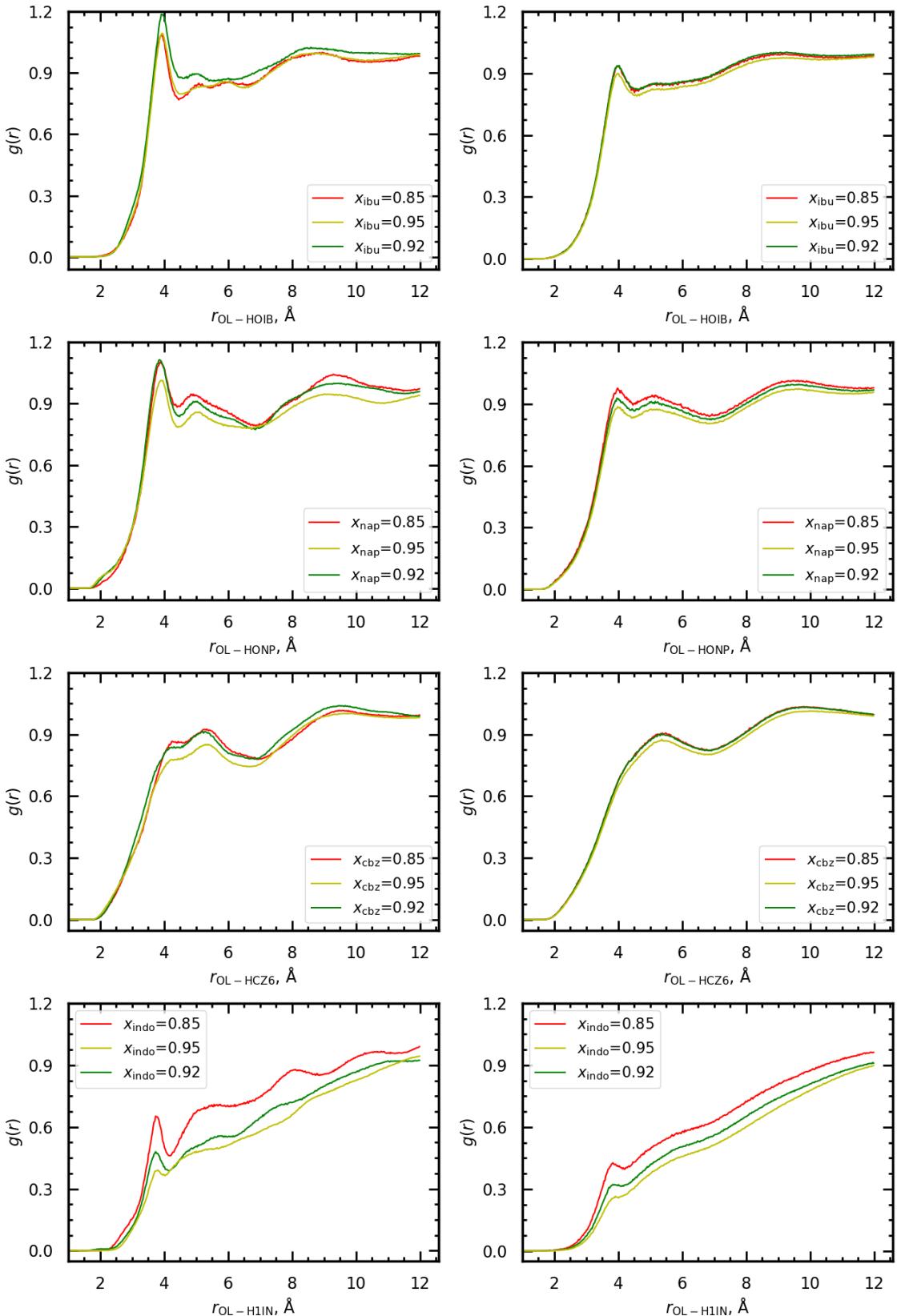


Figure 23: Radial distribution function of the interaction between hydrogen atoms and oxygen atom bonded by ether bond in PLA, first ibuprofen on top, second naproxen, third carbamazepine and indomethacine in the bottom, temperature 300 K on the left and 500 K on the right.

4.3.3 Diffusion coefficients

The MSDs were sampled from the 10 ns long run simulations every 1000 fs (integration step 1 fs). At each sampled time step, obtained MSD data were averaged over all API molecules and then plotted as a function of the simulation time. The MSD dependencies were then interpolated by linear functions and related self-diffusivities of the API in the mixtures were evaluated from the slope of the line using the following Equation 4.5 obtained by modifying the Einstein equation from the theoretical part.

$$D_{\text{API}} = \frac{a}{6}, \quad (4.5)$$

where D is the diffusion coefficient, a is the slope of the line.

The MSD data for APIs in mixtures with PLA for $T=500$ K are plotted in Figure 24. The data of carbamazepine shows that with increasing API concentration, the diffusivity also increases, the polymer is able to attenuate the diffusivity of neat Cbz. This behavior could stem from the intermolecular interactions between API-PLA, but this trend was not visible in analysis of RDFs. However, the dominating factor for diffusivity at higher temperatures is the shape of the molecules. This could be the case why PLA slows down relatively big carbamazepine molecules, even the strong interaction between API-PLA were not observed.

For naproxen, it seems that there is no change for different concentrations of mixtures, but the diffusivity is again slightly higher for neat API. Here we can describe it by forming API-PLA interaction and linking Nap to PLA molecule, which diffusivity is lower. This is in agreement with RDF analysis. We should also have a look at the data for lower temperature.

For ibuprofen, there is a significant difference between mobility of neat API and API mixed within the polymer, note the values on y -axis. This could be result of very strong API-PLA interaction forming in the mixtures. Analysis of RDF showed that ibuprofen had the strongest API-PLA interactions compared to other APIs.

The situation for indomethacine is completely different, the MSD values are much lower. For neat API the mobility is really low compared to mixtures with PLA. This behaviour seems strange in context of other API. The reason could be that in pure API, there are really strong API-API interactions that decrease the mobility. The number of those interactions is lower after adding the polymer. This was visible from RDFs, where indomethacine had the strongest intermolecular interactions decreasing with API concentration.

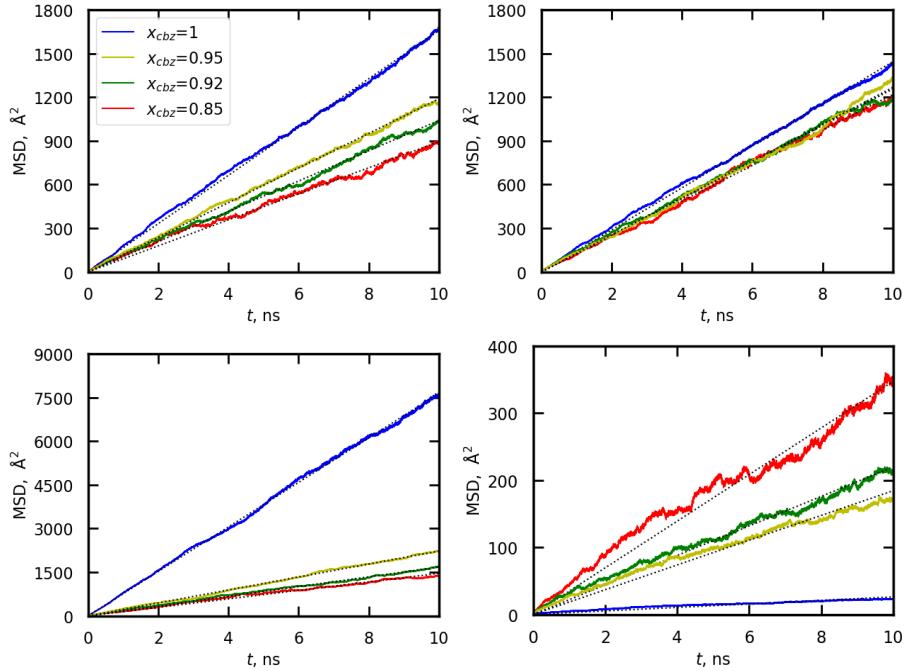


Figure 24: MSD from simulations under 500 K, carbamazepine (**top left**), naproxen (**top right**), ibuprofen (**bottom left**) and indomethacin (**bottom right**).

Self-diffusivities were evaluated from the above data and plotted in Figure 25.

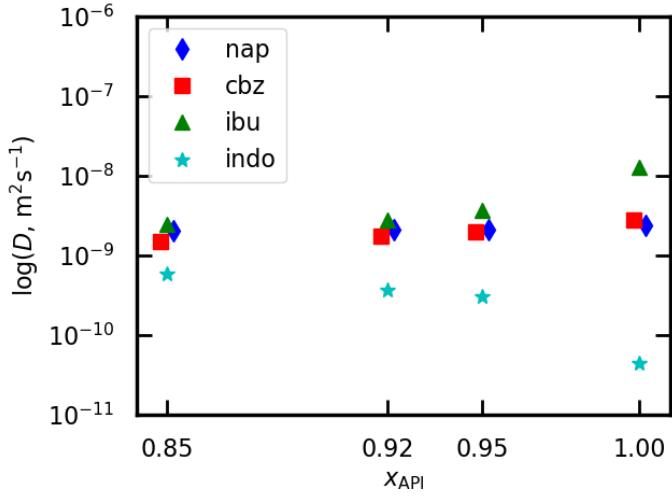


Figure 25: Self-diffusivities (D) for carbamazepine, naproxen, ibuprofen and indomethacine as a function of their concentration in the mixtures, temperature 500 K. Data shown in logarithmic scale

MSD was also evaluated at a lower temperature of 300 K, Figure 26. For carbamazepine there is no visible trend for change of the concentration and neat API has the signal lower than the mixtures. From that we can definitely say, that addition of PLA has minimal effect on the pure API.

The diffusivity of naproxen is always lower in mixtures, which is related to its stronger interactions with the polymer. The strength of the intermolecular interactions has a greater impact, meaning that paired molecules NAP-API slow the diffusion of other particles. The addition of polymer to ibuprofen results in decrease in diffusivity.

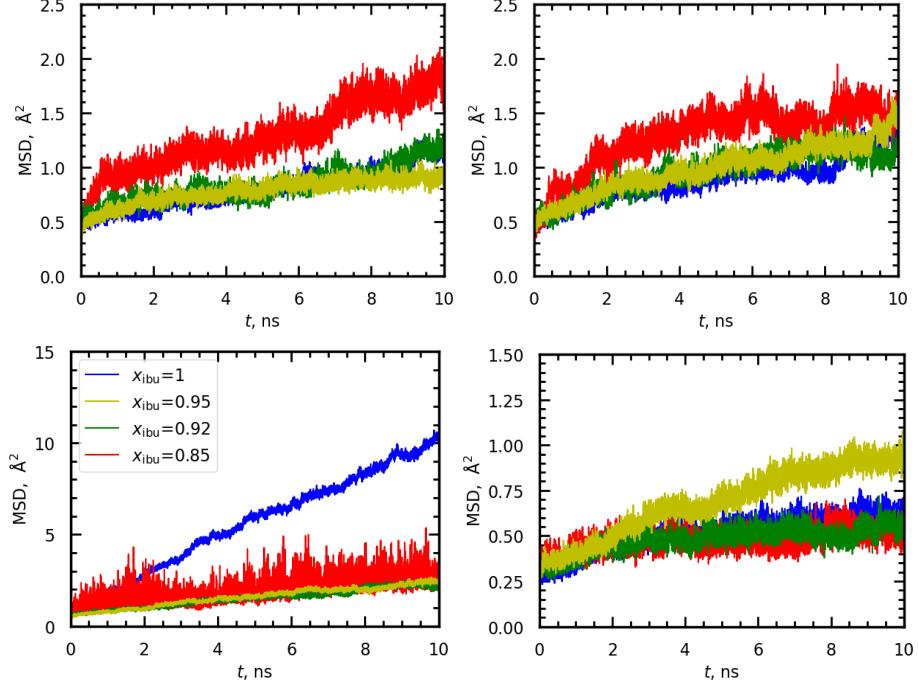


Figure 26: MSD from simulations under 300 K, carbamazepine (**top left**), naproxen (**top right**), ibuprofen (**bottom left**) and indomethacin (**bottom right**).

4.3.4 Glass transition temperature

The glass transition temperatures of the mixtures were evaluated from the simulated annealing runs by fitting a hyperbola to the temperature-density data. The whole methodology is described in the paper written by Alzate-Vargas et al.[41], the main equation of the fit is Equation 4.6

$$\rho(T) = \rho_0 - a(T - T_0) - b \left[\frac{1}{2} (T - T_0) + \sqrt{\frac{(T - T_0)^2}{4} + e^c} \right]. \quad (4.6)$$

Since this method is sensitive to the initial state of the simulated box, more simulated data starting from different conformations must be provided to evaluate T_g with the information about its uncertainty as it was described in methods section. We used

data from 5 simulations from different initial states. The average value and its uncertainty for each API is given in Table 7.

TO DO: DOPOČÍTAT PRO INDO A CBZ SIMULACE

Table 7: Calculated glass transition temperatures T_g of mixtures composed from API and PLA with concentration of $x_{\text{API}}=0.85$ and their standard uncertainties.

API	T_g	σ_{T_g}
carbamazepine	0	0
naproxen	370	6
ibuprofen	363	3
indomethacin	0	0

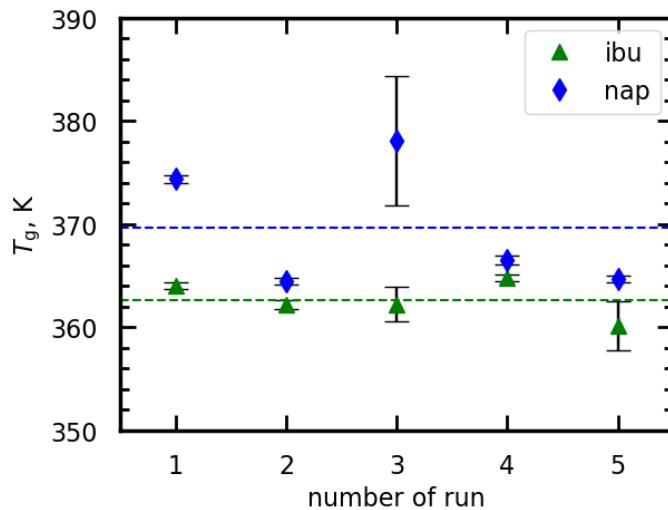


Figure 27: Calculated glass transition temperatures T_g of mixtures composed from API and PLA with concentration of $x_{\text{API}}=0.85$ starting from 5 different conformations, mean values plotted by dashed line for each API.

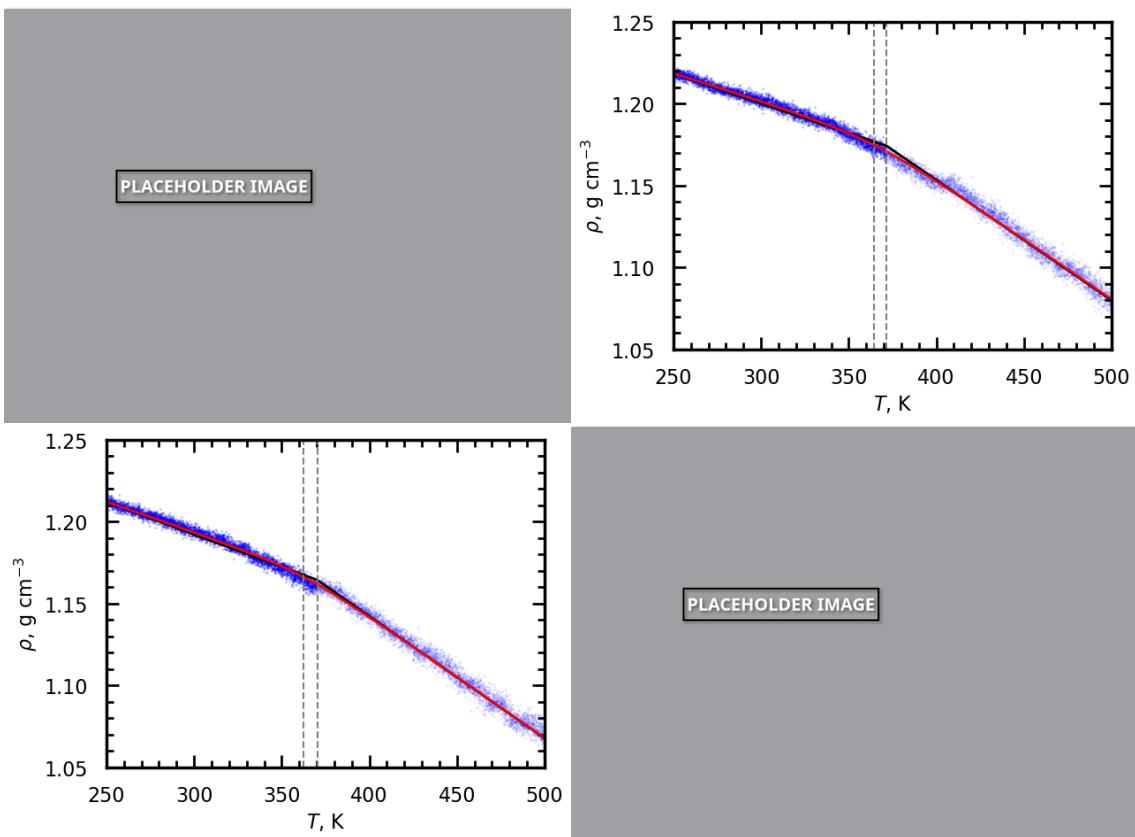


Figure 28: Determination of the glass transition temperature T_g of mixture with API concentration of $x_{\text{API}}=0.85$ by the non-equilibrium hyperbola fit method, second data set. Carbamazepine (**top left**), naproxen (**top right**), ibuprofen (**bottom left**) and indomethacin (**bottom right**).

5 CONCLUSION

text

References

1. Leuner, C. Improving drug solubility for oral delivery using solid dispersions. *European Journal of Pharmaceutics and Biopharmaceutics* **2000**, *50*, 47–60.
2. Batisai, E. Solubility Enhancement of Antidiabetic Drugs Using a Co-Crystallization Approach. *ChemistryOpen* **2021**, *10*, 1260–1268.
3. Huang, L. Impact of solid state properties on developability assessment of drug candidates. *Advanced Drug Delivery Reviews* **2004**, *56*, 321–334.
4. Srinarong, P.; De Waard, H.; Frijlink, H. W.; Hinrichs, W. L. Improved dissolution behavior of lipophilic drugs by solid dispersions: the production process as starting point for formulation considerations. *Expert Opinion on Drug Delivery* **2011**, *8*, 1121–1140.
5. Vasconcelos, T.; Sarmento, B.; Costa, P. Solid dispersions as strategy to improve oral bioavailability of poor water soluble drugs. *Drug Discovery Today* **2007**, *12*, 1068–1075.
6. Caron, V.; Tajber, L.; Corrigan, O. I.; Healy, A. M. A Comparison of Spray Drying and Milling in the Production of Amorphous Dispersions of Sulfathiazole/Polyvinylpyrrolidone and Sulfadimidine/Polyvinylpyrrolidone. *Molecular Pharmaceutics* **2011**, *8*, 532–542.
7. Prudic, A.; Ji, Y.; Sadowski, G. Thermodynamic Phase Behavior of API/Polymer Solid Dispersions. *Molecular Pharmaceutics* **2014**, *11*, 2294–2304.
8. Newman, A.; Zografi, G. What Are the Important Factors That Influence API Crystallization in Miscible Amorphous API-Excipient Mixtures during Long-Term Storage in the Glassy State? *Molecular pharmaceutics* **2022**, *19*, 378–391, Place: United States.
9. Klajmon, M.; Aulich, V.; Ludík, J.; Červinka, C. Glass Transition and Structure of Organic Polymers from All-Atom Molecular Simulations. *Industrial & Engineering Chemistry Research* **2023**, *62*, 21437–21448.
10. Rainsford, K. D. Ibuprofen: pharmacology, efficacy and safety. *Inflammopharmacology* **2009**, *17*, 275–342.
11. Štejfa, V.; Pokorný, V.; Mathers, A.; Růžička, K.; Fulem, M. Heat capacities of selected active pharmaceutical ingredients. *The Journal of Chemical Thermodynamics* **2021**, *163*, 106585.
12. Drebushchak, T. N.; Boldyreva, E. V.; Mikhailenko, M. A. Crystal structures of sulfathiazole polymorphs in the temperature range 100–295 K: A comparative analysis. *Journal of Structural Chemistry* **2008**, *49*, 84–94.
13. Leach, A. R. *Molecular modelling : principles and applications*, 2nd ed.; Prentice-Hall International: Hemel Hempstead, UK, 1997.

- tice Hall Harlow, England: Harlow, England, 2001; Section: xxiv, 744 pages, 16 unnumbered pages of plates : illustrations (some color) ; 24 cm.
- 14. Monticelli, L.; Tielemans, D. P. In *Biomolecular Simulations*; Monticelli, L., Salonen, E., Eds.; Humana Press: Totowa, NJ, 2013; Vol. 924; p 197–213, Series Title: Methods in Molecular Biology.
 - 15. Brooks, B. R. et al. CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry* **2009**, *30*, 1545–1614.
 - 16. Vanommeslaeghe, K.; Guvench, O.; MacKerell, A. D. Molecular Mechanics. *Current Pharmaceutical Design* **2014**, *20*, 3281–3292.
 - 17. Da Silva, G. C. Q.; Silva, G. M.; Tavares, F. W.; Fleming, F. P.; Horta, B. A. C. Are all-atom any better than united-atom force fields for the description of liquid properties of alkanes? *Journal of Molecular Modeling* **2020**, *26*, 296.
 - 18. Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society* **1988**, *110*, 1657–1666.
 - 19. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* **1995**, *117*, 5179–5197.
 - 20. Mackerell, A. D. Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry* **2004**, *25*, 1584–1604.
 - 21. Breneman, C. M.; Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry* **1990**, *11*, 361–373.
 - 22. Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*, 2nd ed.; Oxford University PressOxford, 2017.
 - 23. Nezbeda, I.; Kolafa, J.; Kotrla, M. *Úvod do počítačových simulací: metody Monte Carlo a molekulární dynamiky*; Karolinum, 1998.
 - 24. Šritterová, V. Molecular simulations of the structure of porous metal-organic liquids. 2022.
 - 25. Eastwood, J.; Hockney, R.; Lawrence, D. P3M3DP—The three-dimensional periodic particle-particle/ particle-mesh program. *Computer Physics Communications* **1980**, *19*, 215–261.
 - 26. Shannon, C. Communication in the Presence of Noise. *Proceedings of the IRE*

- 1949**, *37*, 10–21.
- 27. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **1977**, *23*, 327–341.
 - 28. Braun, E.; Gilmer, J.; Mayes, H. B.; Mobley, D. L.; Monroe, J. I.; Prasad, S.; Zuckerman, D. M. Best Practices for Foundations in Molecular Simulations [Article v1.0]. *Living Journal of Computational Molecular Science* **2019**, *1*.
 - 29. Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; In 'T Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications* **2022**, *271*, 108171.
 - 30. Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry* **2009**, *30*, 2157–2164.
 - 31. Padua, A. ENS Lyon, France. <https://github.com/paduagroup/fftool>.
 - 32. Hockney, R.; Eastwood, J. *Computer Simulation Using Particles*, 0th ed.; CRC Press, 2021; p 540.
 - 33. Tuckerman, M. E.; Alejandre, J.; López-Rendón, R.; Jochim, A. L.; Martyna, G. J. A Liouville-operator derived measure-preserving integrator for molecular dynamics simulations in the isothermal–isobaric ensemble. *Journal of Physics A: Mathematical and General* **2006**, *39*, 5629–5651.
 - 34. McAliley, J. H.; Bruce, D. A. Development of Force Field Parameters for Molecular Simulation of Polylactide. *Journal of Chemical Theory and Computation* **2011**, *7*, 3756–3767.
 - 35. Červinka, C.; Fulem, M. Structure and Glass Transition Temperature of Amorphous Dispersions of Model Pharmaceuticals with Nucleobases from Molecular Dynamics. *Pharmaceutics* **2021**, *13*, 1253.
 - 36. Klajmon, M.; Červinka, C. Does Explicit Polarizability Improve Molecular Dynamics Predictions of Glass Transition Temperatures of Ionic Liquids? *The Journal of Physical Chemistry B* **2022**, *126*, 2005–2013.
 - 37. Pyda, M.; Wunderlich, B. Reversing and Nonreversing Heat Capacity of Poly(lactic acid) in the Glass Transition Region by TMDSC. *Macromolecules* **2005**, *38*, 10472–10479.
 - 38. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Proper-

- ties of Organic Liquids. *Journal of the American Chemical Society* **1996**, *118*, 11225–11236.
- 39. Frisch, M. J. et al. GaussianTM16 Revision C.01. 2016.
 - 40. Smith, D. G. A.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. Revised Damping Parameters for the D3 Dispersion Correction to Density Functional Theory. *The Journal of Physical Chemistry Letters* **2016**, *7*, 2197–2203.
 - 41. Alzate-Vargas, L.; Fortunato, M. E.; Haley, B.; Li, C.; Colina, C. M.; Strachan, A. Uncertainties in the predictions of thermo-physical properties of thermoplastic polymers via molecular dynamics. *Modelling and Simulation in Materials Science and Engineering* **2018**, *26*, 065007.