



Molecular dynamics simulations of supramolecular complexes under influence of an external force

Waleed N. AL-DARKAZALI^{a,*}, Omar HACHIM^b

^a Mustansiriyah University, College of Science, Chemistry Department, Baghdad, Iraq

^b Mustansiriyah University, College of Medicine, Microbiology Department, Baghdad, Iraq

ARTICLE INFO

Keywords:

External force
Weak interactions
Calix[n]arenes
Molecular dynamics
Gromacs

ABSTRACT

This work investigates the dynamics of supramolecular complexes Calix[n]arenes that are formed by weak forces. These interactions are important for the structure and function of biological molecules and for the design of synthetic host-guest systems. Molecular dynamics simulations are used to explore the reversible binding under of these complexes external force at the atomic level and to complement experimental methods.

1. Introduction

Computer simulations are becoming more indispensable for scientific research, as they expand the scope of application due to methodological developments and increasing computational power. For chemical problems, quantum chemical calculations and molecular dynamics (MD) simulations [1,2] can disclose the atomic-level events of molecular systems that are elusive to experimental methods. Therefore, computer simulations can aid in elucidating experiments or corroborating experimental data. Conversely, simulations can also prompt new experiments or address questions that are experimentally unfeasible.

In this work, we used MD simulations to explore the dynamics of reversibly bound supramolecular complexes. These complexes are sustained by weak, non-covalent interactions [3,4], which originate from electrostatic and dispersion forces. The most prevalent forms of non-covalent interactions are ion-pair interactions, hydrogen bonds and π -stacking interactions between aromatic groups.

Non-covalent interactions are crucial for the secondary structure and function of biological molecules such as proteins and nucleic acids [5,6]. However, supramolecular complexes are not only significant for natural chemistry. In host-guest chemistry, synthetic molecules are devised that can selectively bind complex molecules. For example, crown ethers can ensnare various cations [7]. By attaching dye molecules to the host molecules that change the wavelength of the emitted light when a guest molecule binds to the host, one can create a selective sensor that operates at the molecular level [8].

Non-covalent bonds have binding energies that are typically comparable to the thermal energy at room temperature [9,10]. Therefore, these bonds can fluctuate and have a limited lifetime. However, supramolec-

ular complexes are often stabilized by a large number of non-covalent bonds. For instance, in folded proteins there are on average about 0.7 hydrogen bonds per amino acid [8].

Non-covalent bonds have binding energies that are typically comparable to the thermal energy at room temperature [9,10]. Therefore, these bonds can fluctuate and have a limited lifetime. However, supramolecular complexes are often stabilized by a large number of non-covalent bonds. For instance, in folded proteins there are on average about 0.7 hydrogen bonds per amino acid [11]. If the two binding partners are close to each other in space, for example due to adjacent bonds, the broken bond can potentially be reformed. In this way, the stability of the system is not affected by small fluctuations in the number of bonds. The interaction of many reversible bonds significantly increases the lifetime of the complex compared to a single non-covalent bond.

A two-state model can provide the simplest description of a reversibly bound complex. Generally, the bound state corresponds to the global minimum and the unbound state corresponds to a local minimum on the potential surface of the system. The transition rates between these states are mainly determined by the energy barrier that separates them [12]. The rate for the transition from the bound to the unbound state is usually smaller than for the transition in the opposite direction [13]. For some special folding proteins [14], protein folding or unfolding cannot be observed in an equilibrium simulation, as the simulation time scale is much smaller than the time scale of these transitions. To overcome this time scale problem, there are several methods. They all share the common feature that they manipulate the system in such a way that the energy barriers between the different states can be more easily crossed. In the so-called metadynamics [15], artificial terms are added to the potential function, which energetically penalize conformations that have

* Corresponding author.

E-mail address: waleednabeel93727@uomustansiriyah.edu.iq (W.N. AL-DARKAZALI).

already occurred in the simulation. This way, conformations that are not normally observed in MD simulations are energetically favoured. In replica exchange MD simulations [16], several copies of the system are simulated in parallel at different temperatures. After a certain time, the coordinates of the system at different temperatures are randomly swapped. Through this swap, conformations that can only be reached at high temperatures also become accessible at low temperatures. Another method of overcoming this time scale problem is inspired by the experimental single-molecule force spectroscopy. By applying mechanical force, the complex can be stretched and even separated if the external force is large enough [17,18]. This also allows, for example, protein unfolding to be observed on the time scale of MD simulations. Unlike the other two methods, the dynamics of the system are specifically altered. However, it should be noted that the dissociation path may depend on the points where the external force is applied and the direction of pulling [19,20]. Such 'pull simulations' are called 'steered MD simulations' or 'force probe simulations'. In the two-state system model, the force lowers the energy barrier and increases the rate of transition from the bound to the unbound state [21,22]. In essence, the force shifts the equilibrium of the system from the bound to the unbound state. By determining the distribution of the detachment forces or the transition times, one can calculate the transition rates and the energy barrier. Since the force applied to the complex and the relative distance change between groups within the complex are both considered, one can also determine the elastic properties of the complex [23]. The conformational changes during dissociation also involve distance changes, which can be measured in experimental force spectroscopy. This allows one to obtain at least a coarse picture of the dissociation path [24,25]. For a more detailed examination, computer simulations are useful, as they can reveal the geometric dissociation path. By combining MD simulations and atomic force microscope experiments, the dissociation or folding paths of some receptor-ligand complexes and proteins could be determined [26,27].

The transitions induced by mechanical force are not deterministic but stochastic processes. Therefore, one observes, for example, a distribution of values for the detachment force when the same measurement is repeated several times. To determine the distributions, one needs a sufficiently large number of experiments or simulations.

2. Theoretical background

To simulate the temporal evolution of a molecular system on a computer, two basic ingredients are required. First, a potential function that describes how the particles interact with each other. These interactions involve the atomic nuclei and the electrons and are governed by quantum mechanics. Assuming the Born-Oppenheimer approximation [46], we can separate the motion of the nuclei and the electrons. If we neglect quantum effects such as tunneling, we can treat the nuclei as classical particles that follow the classical equations of motion. In this approximation, the interatomic interaction consists of an electrostatic term between the nuclei and an effective potential that arises from the interaction of the nuclei and electrons, and the electron-electron interaction. Ideally, we would compute this interatomic interaction potential using quantum chemical *ab initio* methods. These simulations are called *ab initio* molecular dynamics (AIMD) simulations. However, these methods are very computationally demanding, and they limit us to very small systems (~500 atoms) and short times (~50 ps) [47]. Alternatively, we can use classical molecular dynamics (MD) to calculate the potential without explicitly considering the electron-atom interaction. We can approximate it by an empirical function that only depends on the atomic coordinates. This function is known as the force field.

The second ingredient we need is a formalism that allows us to update the coordinates and momenta of the atoms from time t to time $t + \Delta t$. In molecular dynamics, we use Newton's equations of motion to integrate the temporal evolution of the system. The dynamics of the system in this case are deterministic, as they only depend on the initial conditions and the interaction potential.

To model the dynamics of an atomistic system, such as a molecule and its solvent, MD simulations are often used. The initial positions and velocities of all the particles in the system are specified. Then, the potential function is used to compute the forces on each particle based on their positions. These forces and the equations of motion are integrated to update the positions and velocities of the particles after a small-time interval Δt . This process of integration, force calculation, and position and velocity update is repeated until the simulation reaches the desired duration $t_{end} = N_{steps} * \Delta t$, where N_{steps} is the number of integration steps.

The particles in a simulation are confined to a simulation box. This means that only the particles in the middle of the box have other particles around them. The particles at the edge of the box, however, face a boundary with the vacuum. If there are N particles distributed evenly in the system, about $N^{2/3}$ of them are at the boundary. This means that with 106 particles, only 1% of them are affected by boundary effects. To reduce the errors caused by boundary effects, one could increase the size of the box and add more particles. But this would increase the computational cost of the simulation, which scales with N^2 . A better alternative is to use periodic boundary conditions. In this case, the box is replicated infinitely in all directions. The box acts like a unit cell of a periodic crystal structure. In the simulation, each particle and its periodic images move in the same way. This ensures that every particle has other particles around it and avoids the problem of a boundary with the vacuum.

A simulation that uses Newton's equations of motion to integrate the system at a fixed volume follows the microcanonical ensemble (NVE ensemble). However, this simulation cannot be directly compared with experiments that are done at constant temperature and/or pressure. Therefore, it is better to do the simulations in the canonical ensemble (NVT ensemble) or the isobaric-isothermal ensemble (NPT ensemble). The equidistribution theorem provides a link between the MD simulation and the thermodynamic temperature T . It states that the average kinetic energy of an ideal gas, which can be computed from the particle velocities v , is proportional to the temperature. By adjusting the particle velocities, the temperature can be controlled. In other words, the system is connected to a heat bath with the temperature T_B . If the system's temperature differs from T_B , it exchanges energy with the heat bath until they reach thermal equilibrium. The methods that control the temperature are called thermostats. For an ideal gas, the temperature determines the product of pressure P and volume V . By changing the volume of the system in each simulation step, the pressure can be controlled. The term barostat covers the algorithms used for this.

3. The importance of calix [n] arenes

Calix[n]arenes are cyclic oligomers that can be prepared by condensation of formaldehyde with *hp*-alkylphenol in a basic medium [28,29]. n indicates the number of monomer units (Fig. 1)

Calix4arenes are macrocyclic compounds that can be modified with various functional groups at their upper or lower rims (Fig. 2). One of the interesting modifications is the substitution of four urea groups at the upper rim, which gives rise to tetraurea calix4arenes. These compounds have the ability to form dimeric capsules in apolar solvents, which are stabilized by a network of intermolecular hydrogen bonds between the urea groups [30]. The dimerization can be fine-tuned by changing the size and shape of the loops or bulky residues attached to the urea groups [31]. The dimeric capsules can also act as hosts for various guests, such as ammonium salts, or as building blocks for rotaxanes and catenanes [32], which are mechanically interlocked molecules with potential applications in nanotechnology and molecular machines. Therefore, tetraurea calix4arenes are important synthetic targets for supramolecular chemistry and self-assembly studies [33].

The properties of calix[n]arenes can be controlled not only by the choice of the number of monomer units n , but also by the substituent R . Calix[n]arenes are relevant in complex chemistry, as they can complex

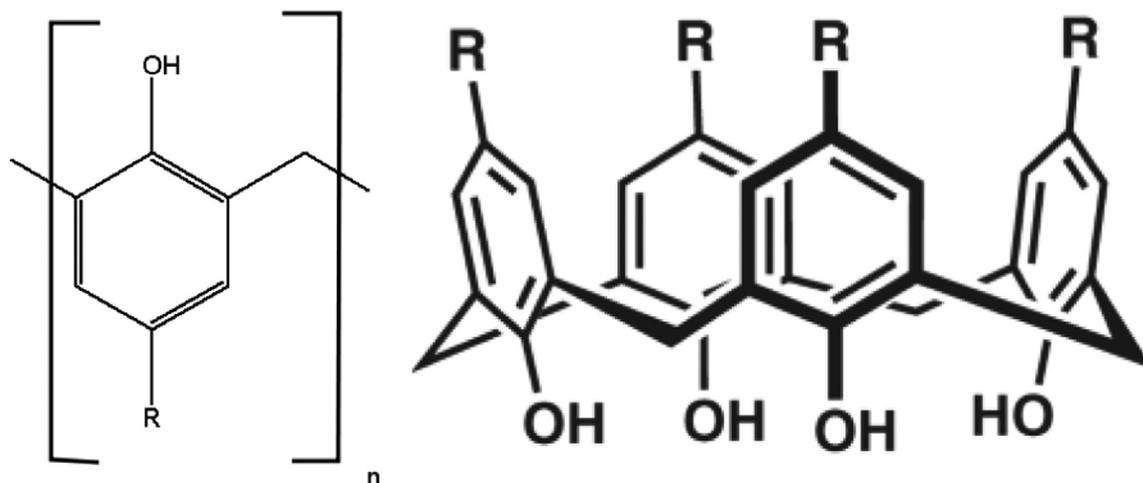


Fig. 1. Monomer unit of a calix[n]arene and Calix4arenes.

Different views of calix[4]arene

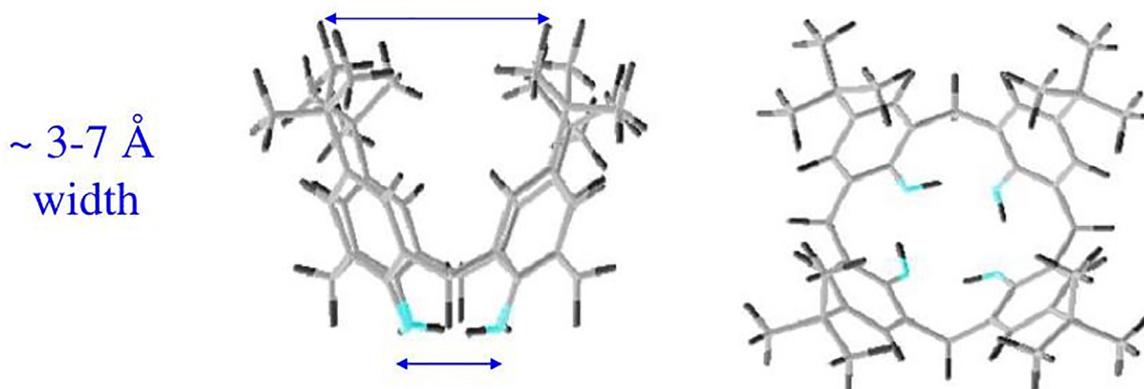


Fig. 2. Rod model of a dimer of two calix[4]arenes substituted with urea groups, left side view, right top view. For better distinguishability, the two monomers are shown differently.

a variety of molecules by suitable choice of size (number of monomer units) and substituents (R or by substitution of the hydroxyl group). For example, cesium ions can be bound with $R=CH_3$. If the calix[n]arene is substituted with sulfonamide groups, different anions can be complexed. Since the choice of substituents also determines the solubility behavior of calix[n]arenes in different solvents, calix[n]arenes are a very flexible ligand in complex chemistry [34]. Moreover, there are efforts to use substituted calix[n]arenes as selective sensors [35].

4. Calix[n]arenes conformers and derivatives

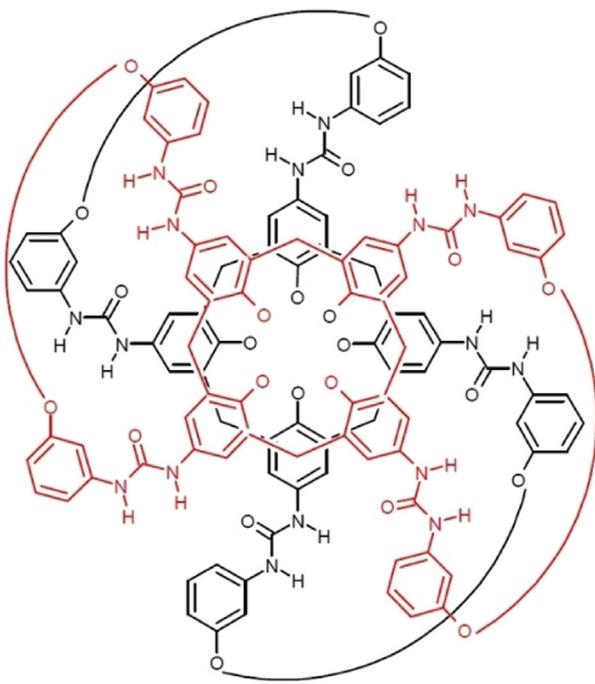
The calix[4]arenes substituted with urea groups were investigated. In the dimer, the two monomers arrange in such a way that the urea groups form a circular network of water bridge bonds. This results in a maximum number of 16 water bridge bonds, since two bonds can be formed between each of the neighbouring pairs of urea groups, of which there are eight.

The calix[4]arene dimer is therefore an ideal model system for an adhesion cluster under the influence of a mechanical pull. In experiments, an external force is applied to the system by means of a spring. The external force initially stretches the dimer system. If the force is sufficiently large, the water stoichiometry bonds can be opened and finally the dissociation of the dimer can be observed. If the stabilization by the

water, the spring can relax and the distance between the two monomers increases abruptly. The large distance between the two monomers prevents the water bonds from rebonding. The transition from the closed to the open structure is irreversible. However, if one also wants to observe rebonding events, the complete separation of the two monomers must be prevented. This can be achieved, for example, by introducing entangled alkyl chains into the system.

In the calix[4]arene catenane dimer these chains are attached to the oxygen atoms of the outer benzene rings, see Fig. 3, and are arranged in such a way that the alkyl chains of the different monomers are looped, similar to two intertwined rings. Generally, such systems, which consist of several monomers that are looped into each other and thus cannot be separated, are called catenanes [36]. In the closed structures, the alkyl chains are orthogonal to the connecting axis of the two monomers, which is why it can be assumed that they only influence the network of water stoichiometry bonds to a minor extent. If the dimer dissociates, the two monomers can separate at most until the alkyl chains are stretched. Further separation is prevented by the steric repulsion between the chains. Since the two monomers are now in relative proximity to each other, depending on the chain length, rebonding is in principle possible if the external force is reduced again. Such a system is called reversible because the transition from one structure (closed) to the other (stretched) can be reversed.

Bis-Loop-System



Tetra-Loop-System

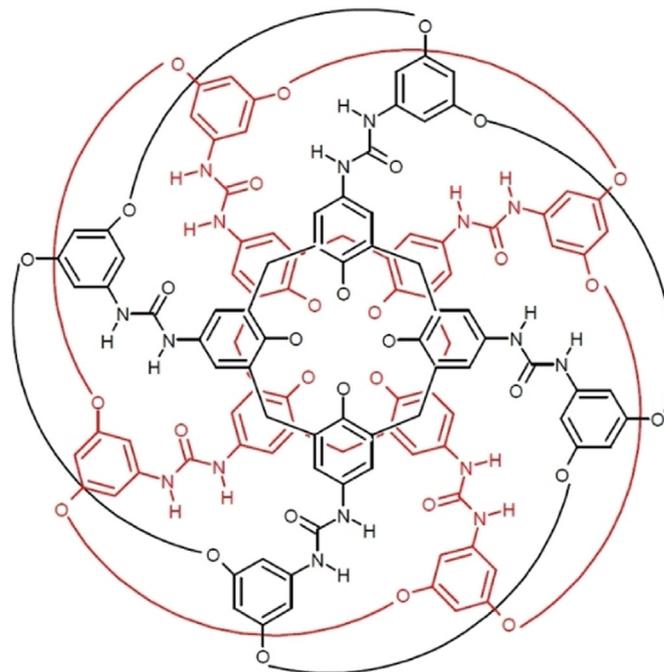


Fig. 3. Schematic representation of the structure of the two dimer systems. The alkyl chains are indicated by curved lines and consist of 20 (bis-loop) or 14 CH_2 units (tetra-loop) each. The single-bonded oxygen atoms in the inner ring each represent a methoxy group ($-\text{OCH}_3$).

In the following, the calix[4]arene-catenane will be abbreviated as calixarene. In an experimental work, Janke et al. [37] investigated two different calixarene systems using dynamic force spectroscopy. On the one hand, the so-called bis-loop system was studied. In this system, two alkyl chains with 20 CH_2 units each are attached to each monomer. The other system is called tetra-loop, as here four shorter (14 CH_2 units) alkyl chains are bound to each monomer. Both dimer systems are schematically shown in Fig. 3.

5. Molecular dynamics simulations details

We used the GROMACS version 2023.01 [38] program package, which is freely available, to perform all MD simulations. The Bis-Loop system was simulated with the GROMOS force field G53a5 [39], while the Tetra-Loop system was also simulated with the OPLS-AA [40] and the GAFF force field [41]. We generated the topologies for the calix[4]arene catenane using different methods depending on the force field. For the GROMOS G53a5 force field, we used the β -PRODRG2 server [42] to generate the topologies and then corrected the partial charges based on the charges of existing amino acids. For the OPLS-AA force field, we manually created the topologies. For the GAFF force field, we calculated the RESP charges [43] using the PyREDS server [44], which employs the GAUSSIAN09 [45] program package with the Hartree-Fock method and a 6-31G* basis for geometry optimisations and RESP charge calculations. We then used the 'Antechamber' module of the AMBER11 program package to generate the topologies and converted them into a GROMACS-readable format. We applied periodic boundary conditions to all simulations and used cut-offs of 1.4 nm for the GROMOS and OPLS force fields and 1.0 nm for the GAFF force field for non-bonding interactions. We updated the neighbourhood list every 10 fs (5 simulation steps) and used the PME method [47] for long-range Coulomb interactions and a dispersion correction for long-range van der Waals interactions. We fixed all bond lengths at their equilibrium value

Table 1
Atom type and charge of Calix [4] arene-catenane

Atom type	charge	Atom type	charge
CH_3	0.303	$\text{CH}_2(1)$	0.303
OA	-0.511	$\text{CH}_2(2)$	0.292
N	-0.310	C(O)	0.450
H	0.310	C(1)	0.208
O	-0.450	C(2)	-0.146
HC	0.146	C(3)	-0.200
CH_2	0.000	C(4)	0.108

with the LINCS algorithm [48], which enabled a time step of 2 fs. The simulation box size was about 5.4 nm \times 4.4 nm \times 4.4 nm for the Tetra-Loop system and 8.2 nm \times 4.6 nm \times 4.6 nm for the Bis-Loop system.

We present the force field parameters for calix[4]arene catenane and mesitylene molecules in Fig. 4. We show the atom types, partial charges, and van der Waals and bonding parameters for each molecule and force field. We use C4 symmetry to reduce the calix[4]arene catenane molecule to a quarter of its size and label the remaining parts as R* groups (R1 for the upper rim and R2 for the lower rim or alkyl chains; see figures 4). We simplify the molecular representations for all-atom force fields (OPLS and GAFF) by showing only one hydrogen atom per alkyl group.

The GROMOS G53a5 force field uses atom types to derive the parameters of the bonding and van der Waals interactions (Table 1). In some cases, atoms that belong to the same atom type have different charges, so an index '*' was added to these atom types. For example, the atom types C(1) and C(2) belong to the atom type C, but have different charges.

The system was then coupled to a heat bath with a temperature of $T = 300$ K. In a second equilibration phase, a barostat ($p = 1$ bar) was also coupled to control the pressure. The subsequent production or pulling

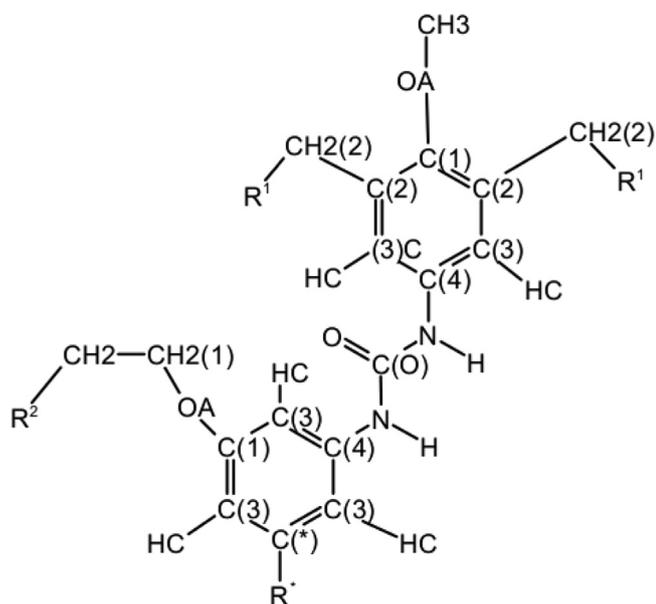


Fig. 4. Atomic types of the calix[4]arene catenane for the GROMOS force field, corresponding to R1C(2) the next quarter unit of the molecule. The missing alkyl groups of the alkyl chain are represented by R2 and consist only of CH₂ atomic types. Since the parameters of the tetra and bis-loop systems practically do not differ, both molecules are shown together. The only difference that arises between both systems enters the types of atoms C(*) and R*. In the case of the Tetra-Loop system, C(*)=C(2) and R*=OA-... or in the case of the Bis-Loop Systems C(*)=C(3) and R*=HC.

simulations, in which the data were collected, were performed under the same conditions and used the final structure of the second equilibration phase as the initial geometry.

Fig. 5. shows the groups relevant for the pulling simulation and illustrates how the external force is applied.

The groups relevant for the tensile simulation are the reference and the tensile group. Both are defined by the center of mass of the four methoxy carbon atoms of a monomer. The reference group is considered to be stationary in the tensile simulation and thus corresponds to the surface on which the molecule is bound in the experiment. A time-dependent harmonic tensile potential is applied to the tensile group. In the experiment, this corresponds, for example, to the cantilever of a force microscope. The force acting on the tensile group or the spring is given by the groups relevant for the tensile simulation which are the reference and the tensile group. Both are defined by the center of mass of the four methoxy carbon atoms of a monomer. The reference group is considered to be stationary in the tensile simulation and thus corresponds to the surface on which the molecule is bound in the experiment. A time-dependent harmonic tensile potential is applied to the tensile group. In the experiment, this corresponds, for example, to the cantilever of a force microscope. The force acting on the tensile group or the spring is given by

$$F = k_c(\nu t - z)$$

where $z = R(t) - R(t = 0)$ is the displacement of the pull group from its original position and thus indicates the change in end-to-end distance. Here k_c corresponds to the force constant of the potential and ν to the pulling speed.

In the simulations carried out here, the force constant $k_c = 830:5$ pN/nm and the tensile speeds $\nu = 10, 1$ and 0.1 m/s were used throughout. This results in loading rates in the order of magnitude of $\mu = k_c - \nu = 1011 - 1013$ pN/s. These are typical values for tensile simulations. For the lowest tensile speed ($\nu = 0:1$ m/s) 50 tensile simulations were carried out and for the other two speeds 100 simulations each. In order to observe all transitions was simulated

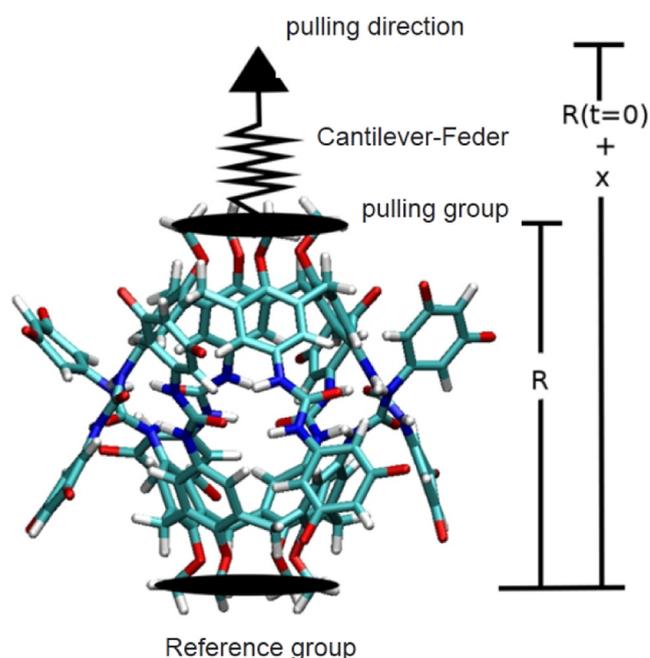


Fig. 5. Stick model of the calixarene (excluding the alkyl chains) and a schematic diagram of the pulling simulation. The center of mass of four methoxy carbon atoms (black ovals) represents either the reference or the pulling group. In the pulling simulation, an external potential is applied to the pulling group, resulting in the observed force. This potential is relative to the reference group, which is fixed during the pulling process.

with decreasing pulling speed for 0.5, 3 and 25 ns. For the relax-mode simulations, the end structure of a pull-mode simulation was chosen as the initial geometry, and the same simulation time as for the pull-mode simulations was used.

6. Results and discussion

If an external force is applied to the Tetra-Loop system, a defined transition to a stretched structure, this transition is accompanied by the opening of the H bonds between the urea groups of the two monomers (UU bonds). The new, stretched structure is stabilized by a network of H-bonds extending between the urea groups and the ether oxygens that connect the calixarene cups to the alkyl chains (UE bonds). The two states of the system can thus be clearly defined via the end-to-end distance R and the nature of the stabilizing H-bonds. In the following, the initially closed state shall be referred to as the C_T state and the stretched state referred to as the O_T condition.

In the following, the results of a pull-mode and a relax-mode simulation at a tensile speed of at the tensile speed $\nu = 0:1$ m/s will be presented. For the other pull- and relax-mode simulations, qualitatively similar results are found.

6.1. Results of a pull-mode simulation

Fig. 6 below shows the variation of the external force (force-distance curve; red) and the end-to-end distance (black) as a function of the control parameter $x = \nu t$ for an exemplary pull-mode simulation.

The trajectory can be divided into two regions, with the transition event at 1.42 nm separates these two regions. In the first region, the force initially increases linearly and the dimer is slowly stretched starting from $R = 1:45$ nm. In this region the system in the C_T state and the UU bonds hold the two monomers together.

Assuming that the system behaves harmonically, the force $F = k_{mol}z$ acts on the dimer, where k_{mol} describes the molecular force constant of the dimer along the tensile direction.

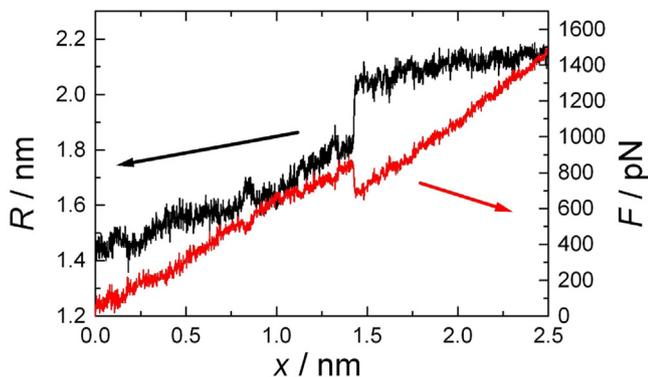


Fig. 6. Measured force F (red) and end-to-end distance R (black) of the Tetra-Loop system as a function of x with tensile speed $v = 0.1$ m/s.

If the system is in equilibrium, the same force given by Eq. (2) must act on the spring. Equating the two forces yields for the change of the end-to-end distance as a function of x :

$$z = (k_c/k_c + k_{mol}) \times x$$

Substituting this expression into $F = kmolz$, we obtain the force as a function of the control parameter:

$$F = (k_c \cdot k_{mol}/k_c + k_{mol})x = k_{eff}x$$

After the break-down event, R and F follow a linear course again, with the end-to-end distance increasing only slightly. For the change of the end-to-end distance as a function of the control parameter (derivation of equation (2) with respect to x):

$$\partial R/\partial x = \partial z/\partial x = K_c/K_c + k_{mol}$$

Fig. 7 shows on the left side the development of the UU bonds (green) and the UE bonds (blue) over time. In addition, the course of R (black) is plotted in order to be able to compare the curves more easily with Fig. 6. On the right side the dynamics of the individual H-bonds can be seen, where number of individual H-bonds (0, 1 or 2) is plotted. It can be clearly seen that in the C_T state, for small values of the control parameter x , 10–11 UU bonds are found on average. Since not the full number of the possible 16 UU bonds is formed, 1–2 UU bonds can already be formed in the C_T state. If the system is stretched, the number of UU bonds decreases almost linearly. From the detailed plot (cf. right-hand side of Fig. 7), it can be seen that initially one UU bond of each proton donor is formed.

The second bond opens at a much later point in time and one can assume that the two UU bonds have different stability. In crystal struc-

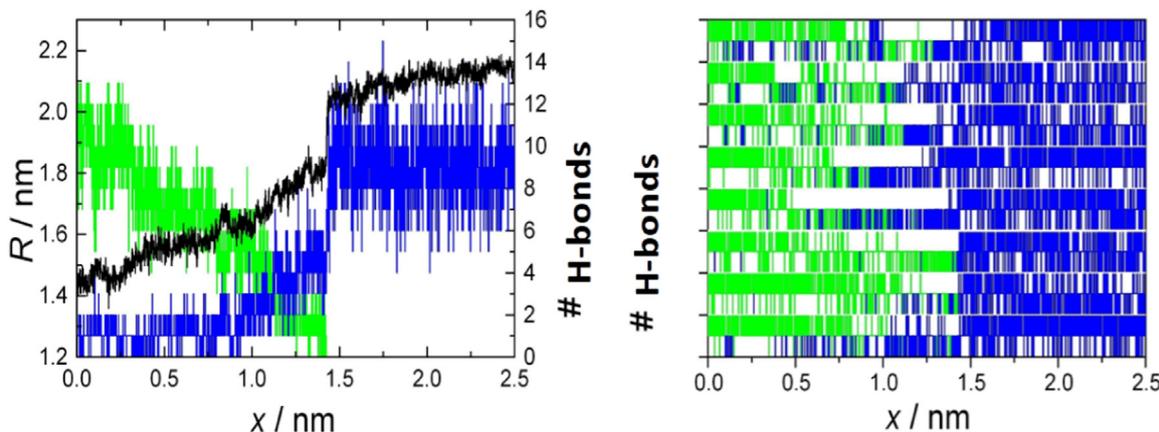


Fig. 7. Left: number of UU (green) and UE bonds (blue) as a function of x . In order to simplify the comparison with Fig. 6, R (black) is also drawn in. Right: Dynamics of the individual H bonds. Each track shows the course of the UU or UE bonds (zero, one or two) of an acceptor-donor pair.

tures of the Tetra-Loop system, for the N–O distances that are relevant for the UU bonds, measured two different distances [34]. Thus, the different stability of the both bonds are explained. From $x = 1.0$ nm there are almost exclusively the strong ones UU bindings and one observes an increase in the number of UE bindings. The rupture event at 1.42 nm is evident from the opening of the remaining UU bonds and recognized by a large increase in the number of UE bindings.

In O_T condition one observes an average of about 10–11 UE binding. Their number remains in the further course. However, the end-to-end distance changes only slightly when open changes.

The fact that the UU bonds have two different bond strengths, it is also evident from the dynamics of the total number of UU bonds (Fig. 7 left).

Erdman et al. investigated the time-dependent potential of rigid adhesion clusters using kinetic Monte Carlo simulations. They found that the bond number initially decreased slowly, but rapidly dropped near the transition event. Since the force was uniformly distributed among all bonds, each bond rupture increased the likelihood of the next one, leading to the results observed in the Monte Carlo simulations. The linear decrease of UU bonds in the tetra-loop systems observed in the MD simulations can be attributed to the different stability of the bonds. Initially, only the weaker bonds are broken, followed by the stronger ones.

6.2. Statistical analysis of the force probe simulations

The detachment and reattachment transitions are stochastic events, so a single trajectory of a pull experiment is not sufficient to analyze them. Instead, the statistical distribution of the observed quantities can be examined over multiple experiments. The most relevant quantity is the distribution of the transition forces. The mean force and the mean end-to-end distance are also interesting to analyze. The mean force can reveal the elastic properties of the system, and the mean end-to-end distance can indicate the reconnection behavior of the system.

Fig. 8 shows the distribution of the transition forces for the three pulling speeds on the left. No reattachment events were observed for the highest pulling speed $v = 10$ m/s, so the corresponding forces are absent from the distribution. On the right, the mean transition forces are plotted as a function of pulling speed.

The overall trend of the mean transition forces agrees with the theoretical expectation that the detachment forces increase with increasing pulling speed and the reattachment forces decrease [22]. The detachment forces measured in the simulations are an order of magnitude larger than the experimental ones for the bis-loop system. Moreover, the dependence of the detachment force on the loading rate, which is proportional to the pulling speed, is much steeper than in the experiments. Experimentally, a change in the loading rate by one decade results in an

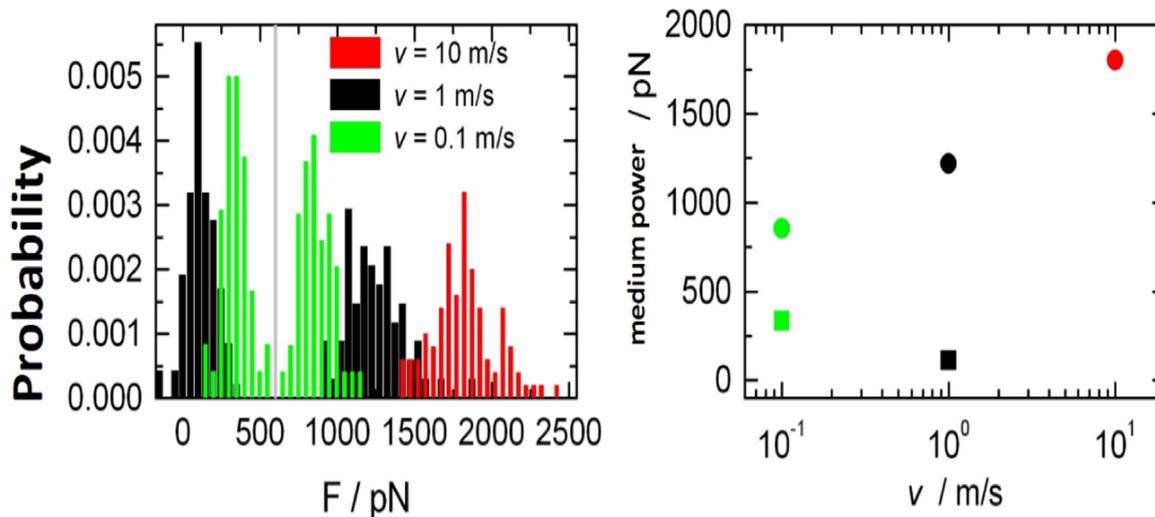


Fig. 8. Left panel: Distribution of the detachment (right of the gray line) and reattachment forces (left of the line) for different pulling speeds. Right panel: Plot of the mean transition forces versus the pulling speed v , with the detachment forces as circles and the reattachment forces as squares. The mean transition forces were obtained by fitting a Gaussian distribution to the respective data.

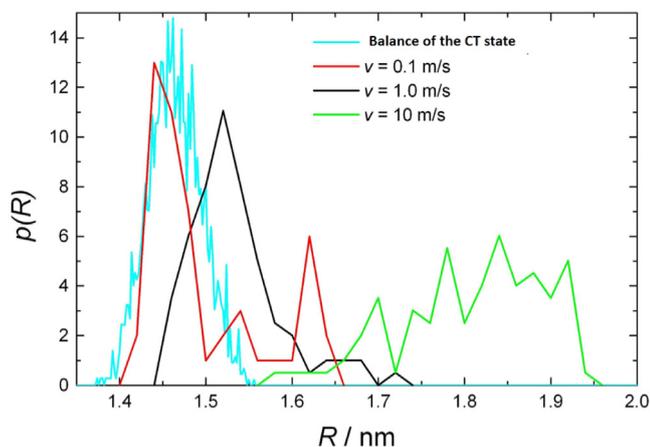


Fig. 9. Distribution $p(R)$ of end-to-end distances as found at the end of each relax mode simulation. The reference (cyan) is the distribution of the distances measured in an equilibrium simulation of the dimer system.

increase of the detachment force by 60 pN [27], while the simulations show an increase of 470 pN. This effect is also observed in other systems [66]. For $v = 1$ m/s, partially negative rebound forces are measured. Relaxes the dimer slower than the spring moves, the term $(z - vt)$ and thus the force can become negative.

In this case, the dimer is under compression. The reattachment events occur in this region, and negative reattachment forces are also measured. If the hysteresis is very large, the reattachment transitions can be shifted to forces that are not accessible within the timescale of the simulation and thus no reattachment events are observed. The timescale of a relax-mode simulation is determined by the simulation time (tend) of the preceding pull-mode simulation.

Fig. 9 shows the distribution of the end-to-end distances $p(R)$ found at the end of the relax-mode simulations. For reference (cyan), $p(R)$ is also shown for an equilibrium simulation, in which no external force was applied.

The largest end-to-end distance observed for the tetra-loop system in equilibrium is 1.55 nm. If a relax-mode simulation ends with a distance of $R < 1.55$ nm, then the system is in a conformation that is also found in equilibrium and the simulation can be considered fully reversible. Table 2 shows the fraction of simulations that satisfy this criterion. This

Table 2

Rebinding probability $P_{\text{rebinding}}$ of the three pulling speeds. $P_{\text{rebinding}}$ is defined as the fraction of the simulations that show an R at the end of the relax-mode simulation, that is also observed in the equilibrium simulations and thus corresponds to the overlap integral of these two distribution.

v (m/s)	$P_{\text{rebinding}}$
0.1	0.88
1	0.73
10	0.0

fraction can also be interpreted as the overlap integral of the distribution of the distances in the equilibrium simulation and the distribution of the distances at the end of the relax-mode simulations. For $v = 0.1$ m/s, most of the relax-mode simulations (88%) reach a conformation that is also observed in equilibrium. This fraction decreases as the pulling speed increases.

6.3. Average number of H-bonds

As mentioned earlier, the network of H-bonds is a characteristic quantity for distinguishing the C_T and O_T states besides the end-to-end distance. Fig. 10 shows the temporal evolution of the average number of H-bonds. For the UU bonds, the plot is truncated at $x = 3.5$ nm, since no UU bonds are observed beyond this value of x .

In the pull-mode simulations, the number of UU bonds decreases more slowly with increasing pulling speed than with lower pulling speed. This effect is due to the logarithmic dependence of the detachment force on the pulling speed. This relation holds not only for the detachment force of the whole system, but also for the individual bonds. Since the detachment time is proportional to the detachment force and $x = vt$, the mean value of the control parameter at which the rupture occurs has the same dependence. Therefore, with increasing pulling speed, the individual H-bonds are broken at a larger value of x than with lower pulling speed. When the number of UU bonds reaches 5, the slope of the curves changes ($x \approx 1, 1.5$ and 2.1 nm). At this point, all weak bonds are broken and the system is only stabilized by the strong bonds. As the system is further stretched, the number of UU bonds continues to decrease and the UE bonds are increasingly formed. For very large values

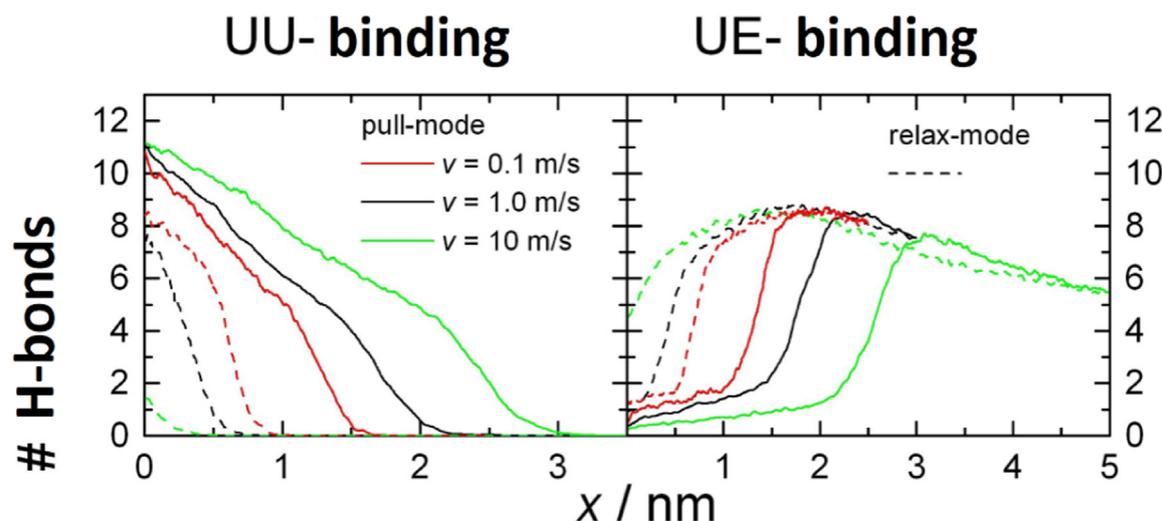


Fig. 10. Mean number of H-bonds as a function of the control parameter. left side the number of UU bonds and on the right side the number of UE bonds is plotted, solid lines indicate the pull-mode and broken lines the relax-mode simulations.

of x , the number of UE bonds decreases again. This suggests that the UE bonds stabilize the open state, but prevent the complete separation of the monomers by the alkyl chains. If these chains were longer, it should be possible to fully dissociate the network of UE bonds as well.

7. Summary

The tetra-loop system exhibits two distinct states, characterized by different end-to-end distances and water stub bond types (UU/UE). An external force can induce a transition from the closed C_T state to the stretched O_T state. The alkyl chains prevent further separation of the monomers. Unlike the experimental results, the system does not fully revert to the C_T state within the timescale of the MD simulations, as evidenced by the velocity-dependent hysteresis and the incomplete relaxation at different pulling speeds. At very high forces, some of the UE bonds that stabilize the O_T state are broken, suggesting that complete dissociation of the monomers is feasible in the bis-loop system with longer alkyl chains.

References

- [1] Cao Y, et al. Molecular dynamic simulations and quantum chemical calculations of adsorption process using amino-functionalized silica. *J Mol Liq* 2021;330:115544.
- [2] Ollitrault PJ, Miessen A, Tavernelli I. Molecular quantum dynamics: a quantum computing perspective. *Acc. Chem. Res* 2021;54(23):4229–38.
- [3] Abriata LA, Dal Peraro M. Assessing the potential of atomistic molecular dynamics simulations to probe reversible protein-protein recognition and binding. *Sci Rep* 2015;5(1):10549.
- [4] Bartocci A, et al. Molecular dynamics approach for capturing calixarene-protein interactions: the case of cytochrome C. *J Phys Chem B* 2020;124(50):11371–8.
- [5] Hobza P, Zahradník R, Müller-Dethlefs K. The world of non-covalent interactions: 2006. *Collect Czech Chem Commun* 2006;71(4):443–531.
- [6] Černý J, Hobza P. Non-covalent interactions in biomacromolecules. *Phys Chem Chem Phys* 2007;9(39):5291–303.
- [7] Gokel MR, et al. Crown ethers having side arms: a diverse and versatile supramolecular chemistry. *J Coord Chem* 2021;74(1–3):14–39.
- [8] Wang B, et al. A stable zirconium based metal-organic framework for specific recognition of representative polychlorinated dibenzo-p-dioxin molecules. *Nat Commun* 2019;10(1):3861.
- [9] Grandbois M, et al. How strong is a covalent bond? *Science* 1999;283(5408):1727–30.
- [10] Moy VT, Florin E-L, Gaub HE. Intermolecular forces and energies between ligands and receptors. *Science* 1994;266(5183):257–9.
- [11] Pace CN, et al. Contribution of hydrogen bonds to protein stability. *Protein Sci* 2014;23(5):652–61.
- [12] Serrano-Andrés L, Serrano-Pérez JJ, Leszczynski J, editor. Calculation of excited states: molecular photophysics and photochemistry on display, in handbook of computational chemistry 2016:1–88 Editor..
- [13] Buch I, Giorgino T, De Fabritiis G. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc Natl Acad Sci* 2011;108(25):10184–9.
- [14] Fersht AR. On the simulation of protein folding by short time scale molecular dynamics and distributed computing. *Proc Natl Acad Sci* 2002;99(22):14122–5.
- [15] Sutto L, Marsili S, Gervasio FL. New advances in metadynamics. *Wiley Interdiscipl Rev: Comput Molcul Sci* 2012;2(5):771–9.
- [16] Qi R, et al. Replica exchange molecular dynamics: a practical application protocol with solutions to common problems and a peptide aggregation and self-assembly example. *Peptid Self-Assemb: Method Protocol* 2018:101–19.
- [17] Wang Z, et al. On the Interpretation of Force-Induced unfolding studies of membrane proteins using fast simulations. *Biophys J* 2019;117(8):1429–41.
- [18] Marszałek PE, Dufrêne YF. Stretching single polysaccharides and proteins using atomic force microscopy. *Chem Soc Rev* 2012;41(9):3523–34.
- [19] Chen Y-F, et al. Direction-dependent force-induced dissociation dynamics of an entropic-driven lock-and-key assembly. *Phys Rev E* 2017;96(3):032610.
- [20] Izrailev S, et al. Steered molecular dynamics Computational Molecular Dynamics: Challenges, Methods. Deuffhard P, et al., editors. Berlin Heidelberg: Springer-Verlag; 1999.
- [21] Zwanzig R. Two-state models of protein folding kinetics. *Proc Natl Acad Sci* 1997;94(1):148–50.
- [22] Rhoades E, et al. Two-state folding observed in individual protein molecules. *J Am Chem Soc* 2004;126(45):14686–7.
- [23] Piana S, Klepeis JL, Shaw DE. Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations. *Curr Opin Struct Biol* 2014;24:98–105.
- [24] Freddolino PL, et al. Challenges in protein-folding simulations. *Nat Phys* 2010;6(10):751–8.
- [25] Bowman GR, et al. Progress and challenges in the automated construction of Markov state models for full protein systems. *J Chem Phys* 2009;131(12):124101.
- [26] González, L.A.M., Combining molecular dynamics simulations and atomic force microscopy experiments to rationalize the mechanical properties of double-stranded DNA and RNA. 2020.
- [27] Lee C-K, et al. Atomic force microscopy: determination of unbinding force, off rate and energy barrier for protein–ligand interaction. *Micron* 2007;38(5):446–61.
- [28] Hoskins C, Curtis ADM. Simple calix [n] arenes and calix [4]resorcinarenes as drug solubilizing agents. *J Nanomed Res* 2015;2(3).
- [29] Gorghiu LM, et al. Comparative Study of Some Calix [n] arenes and other Phenolic Compounds Influence on the thermal Stability of LDPE. *Rom Rep* 2004;56:466–72.
- [30] Braekers D, et al. Self-sorting dimerization of tetraurea calix [4]arenes. *J Org Chem* 2008;73(2):701–6.
- [31] Cho YL, et al. Hydrogen-bonding effects in Calix [4]arene capsules. *Chem–A Eur J* 2000;6(20):3788–96.
- [32] Vinicius Vieira Varejao E, Fatima Ade, Antonio Fernandes S. Calix [n] arenes as goldmines for the development of chemical entities of pharmaceutical interest. *Curr Pharm Des* 2013;19(36):6507–21.
- [33] Xu S, et al. Tetraurea calix [4]arenes with sulfur functions: synthesis, dimerization to capsules, and self-assembly on gold. *Org Biomol Chem* 2007;5(3):558–68.
- [34] Kajiwara T, Iki N, Yamashita M. Transition metal and lanthanide cluster complexes constructed with thiocalix [n] arene and its derivatives. *Coord Chem Rev* 2007;251(13–14):1734–46.
- [35] Akbarzadeh A, Zadmand R, Jalali MR. Synthesis of novel 6-piperidin-1-ylpyrimidine-2, 4-diamine 3-oxide substituted calix [4]arene as a highly selective and sensitive fluorescent sensor for Cu^{2+} in aqueous samples. *Tetrahedron Lett* 2020;61(13):151658.
- [36] Schill G. Catenanes, rotaxanes, and knots. Elsevier; 2017. Vol. 22.
- [37] Janke M, et al. Mechanically interlocked calix [4]arene dimers display reversible bond breakage under force. *Nat Nanotechnol* 2009;4(4):225–9.

- [38] Van Der Spoel D, et al. GROMACS: fast, flexible, and free. *J Comput Chem* 2005;26(16):1701–18.
- [39] Oostenbrink C, et al. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 2004;25(13):1656–76.
- [40] Kaminski GA, et al. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J Phys Chem B* 2001;105(28):6474–87.
- [41] Wang J, et al. Development and testing of a general amber force field. *J Comput Chem* 2004;25(9):1157–74.
- [42] Schlesier T, Diezemann G. Performance of different force fields in force probe simulations. *J Phys Chem B* 2013;117(6):1862–71.
- [43] Cornell WD, et al. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J Am Chem Soc* 2002;115(21):9620–31.
- [44] Wang F, et al. Object oriented programming for amber force fields. Université De Picardie-Jules Verne, Sanford| Burnham Medical Research Institute; Nov 2013.
- [45] Frisch M. gaussian09; 2009 <http://www.gaussian.com/>.
- [46] Case DA, et al. AMBER 11, San Francisco: University of California; 2010. Google Scholar There is no corresponding record for this reference.
- [47] Kratz EG, Duke RE, Cisneros GA. Long-range electrostatic corrections in multipolar/polarizable QM/MM simulations. *Theor Chem Acc* 2016;135:1–9.
- [48] Hess B, et al. LINCS: a linear conspulling solver for molecular simulations. *J Comput Chem* 1997;18(12):1463–72.