Combined Machine Learning and Molecular Dynamics Reveal Two States of Hydration of a Single Functional Group of Cationic Polymeric Brushes

Raashiq Ishraaq,[1] Tanmay Sarkar Akash,[1] and Siddhartha Das[1]*

[1]Department of Mechanical Engineering, University of Maryland, College Park, MD 20742

*Email: sidd@umd.edu

Abstract:

The state of hydration of a macromolecular system regulates a plethora of different properties of such a system. In this article, we develop a novel machine learning (ML) approach, based on the unsupervised clustering algorithm, for probing the hydration behavior of the $\{N(CH_3)_3\}^+$ functional group of the PMETAC [Poly(2-(methacryloyloxy)ethyl trimethylammonium chloride] polyelectrolyte (PE) brush system. The PE brushes and the brush-supported water molecules and counterions (chloride ions) are first described using all-atom molecular dynamics (MD) simulations. The simulation data is subsequently used in our ML framework to identify that (1) the $\{N(CH_3)_3\}^+$ functional groups of the PMETAC brushes have two distinct hydration states with one state (state 1) being characterized by less structured water molecules and the other state (state 2) being characterized by more structured water molecules and (2) an enhancement in the brush grafting density leads to the progressive dissapparenace of state 2. An increase in the grafting density increases the number of chloride counterions in a given volume around the $\{N(CH_3)_3\}^+$ functional group and increases the number of shared water molecules between the $\{N(CH_3)_3\}^+$ and Cl^-. The chloride counterions are associated with a hydration layer with much less structured water molecules. Therefore, with an increase in the grafting density, an increase in the percentage of shared water molecules leads to the prevalence of the hydration state [of the $\{N(CH_3)_3\}^+$ moiety] with less structured water molecules. Finally, we explain how the present findings are commensurate with two key previous related results, namely a significantly large chloride ion mobility inside the PMETAC brush layer and the $\{N(CH_3)_3\}^+$-Cl^- average distance remaining independent of the PMETAC brush grafting density. We anticipate that the combined ML-MD-simulation approach proposed in this study can be adapted to probe other soft matter systems to reveal new insights of the underlying mechanisms of emergent phenomenon.

1. INTRODUCTION

Macromolecular systems are often accompanied by solvating molecules (mostly water) [1], which dictate solvent-specific properties of the system [2], such as the macromolecular configuration, ion distribution around the macromolecules, etc. Moreover, the solvent behavior near polymeric structures plays a vital role in polymer applications; for example, one proposed mechanism of the non-fouling activity of the polymer-coated surface is the hydration of certain "non-fouling moieties" which are responsible for creating a well-structured water layer above the surface. This water structuring then imparts an enthalpic penalty forbidding the unwanted proteins to penetrate the water and adsorb on the surface [3]. Therefore, developing a microscale understanding of the nature of hydration of such polymeric systems becomes critical in deciphering the properties and behaviors of such systems. Polyelectrolyte (PE) brushes, which refer to charged polymer chains grafted at close proximity to each other, enforcing the PE chains to attain "brush-like" configurations, is one such macromolecular system that has found extensive applications in colloid stabilization [4], developing controlled flow in nanochannels [5], gene delivery [6], targeted drug delivery[7], fabrication of nanofluidic diodes [8] and current rectifiers[8], [9], surface modification [10], [11], lubrication [12], and many more. Extensive experimental, theoretical, and simulation studies have been conducted to study the behavior of PE brushes. Such brushes are invariably accompanied by counterions (ions that screen the charges of the PE brushes) and solvent (often water) molecules, albeit the amount of attention that has been devoted to studying the properties and behaviors of such brush-supported water molecules and counterions has been very limited. For example, experimental studies using neutron scattering[13]–[17], X-ray reflectometry[18]–[20], and IR spectroscopy[21], [22] have shed light on the structural and dynamical properties of various PE brushes. However, oftentimes, these approaches lack the

accuracy and resolution to capture the dynamics and properties of the corresponding brush-supported solvent molecules and counterions. Theoretical and simulation methods, on the other hand, often try to model such systems using a combination of mean-field and free-energy approaches, where the free energy is a function of coarse-grained variables [23]–[27]. As a result, the atom-specific interactions get ignored or are treated as an average, which obfuscates the behavior of the solvent molecules in the vicinity of the PE brushes. Under such circumstances, it is difficult (or nearly impossible) to accurately study the hydration characteristics of the PE brushes and to predict the mechanisms of the associated emergent phenomena. Techniques such as the all-atom Molecular Dynamics (MD) simulations relieve some of these problems (to a certain extent) by modeling the system in a manner where each atom is treated as an interaction site (or coarse-graining some unimportant sites for the study at hand) and then by following the time evolution (in the case of MD simulations) or by solely relying on energetics (for the case of the Monte Carlo simulations) to shed light on the mechanisms dictating the system behavior [28]–[31].

All-atom MD simulations of the PE brush systems generate data on the trajectories and velocities of all the atoms constituting the system (PE brush, counterions, and water molecules) [32]. Often, it becomes challenging to unearth the key science of the probed phenomenon from such a large amount of raw data. This is especially true in the context of obtaining the behavior of the brush-supported (or in general macromolecular-system-supported) polar solvent molecules (like water), where the solvent molecules can make various kinds of information-rich configurations. Therefore, there is also a need for a scheme (or a protocol) for a phenomenon-specific (but system-agnostic) approach to decipher the relevant science and reveal novel insights into such complicated systems. Currently, there is an ongoing endeavor in the scientific community to apply various Machine learning (ML) methods to sift through such large amounts of data to

generate meaningful information. Recently, Gasparatto and Ceriotti developed a protocol to determine the condition of water-water hydrogen bonding from the values of distances between the coordinates of the atoms obtained from the MD simulations [33]. Their study helped to automatically detect relevant structural patterns from the data generated by the MD simulations. Offei-Danso *et al.* developed a protocol to understand the water-water hydrogen bond networks. They used chemical intuition-based parameters to encode the structural information, followed by dimensional reduction and clustering. They were able to use the protocol to detect high dimensional fluctuations of various length scales, which eventually helped to describe the water-water hydrogen bond networks [34]. Ansari *et al.* [35] used similar ML techniques to investigate the fluctuation-driven void formation in liquid water and methane and showed that the work associated with creating a void of a specific shape is the same as the work required to solvate a hydrophobic polymer of the same shape. The utilization of ML algorithms similar to (but not limited to) the above-mentioned studies has paved the way for opening up novel research avenues for investigating different phenomena related to soft matter systems (e.g., PE brushes). One example is our recent study [36], where we used the clustering algorithm of Gasparatto and Ceriotti [33] for identifying the water-water hydrogen bonding statistics inside a densely grafted poly-acrylic acid (PAA) brush layer: our findings revealed a significant difference in the water-water hydrogen bonding statistics inside the brush layer as compared to that outside the brush layer. In order to further such efforts, it is important to develop protocols that use simple ML algorithms (with minimum abstraction) and provide room for utilizing domain-specific knowledge of the researcher for their problem at each critical point in the analysis pipeline.

In this article, we make such an effort of developing a ML protocol that uses simple unsupervised ML clustering algorithms in order to probe the hydration behavior inside the densely

grafted PMETAC [Poly(2-(methacryloyloxy)ethyl trimethylammonium chloride] brush layer. Specifically, we investigate the microscopic hydration behavior of the $\{N(CH_3)_3\}^+$ functional group (moiety of interest) of the PMETAC brushes. Employing the ML protocol on all-atom MD-simulation-generated data, we identify that the $\{N(CH_3)_3\}^+$ functional group of the PMETAC brushes has two distinct microscopic hydration states. One state (state 1) is characterized by water molecules with a smaller order parameter value (q) (i.e., these water molecules are less structured), whereas the other state (state 2) is characterized by water molecules with a greater q value (i.e., these water molecules are more structured). We further identify that an increase in the grafting density leads to a progressive disappearance of state 2. An increase in the grafting density increases the number of chloride counterions per unit volume around a given $\{N(CH_3)_3\}^+$ group and causes a significant increase in the extent of sharing of water molecules between $\{N(CH_3)_3\}^+$ group and Cl^- ions. The chloride counterions have higher charge density and disrupt the tetrahedral order of the hydration shell water molecules much more than the $\{N(CH_3)_3\}^+$ group of the PMETAC chains [37]. Accordingly, the proportion of less structured water molecules around the $\{N(CH_3)_3\}^+$ group increases, enforcing a progressive disappearance of the $\{N(CH_3)_3\}^+$ hydration state 2 with an increase in the grafting density. These findings, we further establish, are commensurate with the existing knowledge of water and chloride counterion behavior around $\{N(CH_3)_3\}^+$ group of the PMETAC brushes.

The rest of the manuscript is organized as follows. First, we introduce the details of the system and MD simulation method in section 2.1. The ML protocol that has been developed in this paper is next presented in section 2.2. The choice of state variables (section 3.1); results obtained from ML protocol (section 3.2); further investigation using classical methods (section 3.3); and discussions for providing a more complete picture of the hydration of the $\{N(CH_3)_3\}^+$

moiety of the PMETAC brushes (section 3.4) have been provided section 3. Finally, we conclude (section 4) by summarizing the key findings of our study and the future applications of our proposed method.

2. METHODOLOGY

2.1. Molecular Dynamics Simulations

We performed molecular dynamics (MD) simulations on water-solvated PMETAC brush systems for three possible brush grafting density values, namely, σ_g= 0.15, 0.2, 0.25 chains/nm^2. These grafting density values were in the range reported in previous experimental studies[38], [39]. Each PMETAC chain consisted of 24 PMETAC monomers and had two methyl groups capping at the end. The chains were arranged in a 5 x 5 square array and were solvated in water [Fig. 1(a) provides a snapshot of the system]. OPLS-AA [40] forcefield was used to model the interactions between the brush atoms (which was also used to model PMETAC chains [36]). For water, we used SPC/E water model [29], which when used with the OPLS-AA forcefield, showed excellent reproducibility of polymer properties (that match closely with those observed experimentally). While geometric rules were applied to determine the Lennard Jones (LJ) interaction parameters for dissimilar atoms of the PMETAC brush, an exception was made for ion-ion and ion-water LJ interactions. For these cases, similar to the approach of Chetham *et al.*[41], the Lorentz-Berthelot mixing rule was used. In section S1 of the SI (Supporting Information), we have provided the forcefield parameters used for modeling the different interactions considered in this study. Periodic boundary conditions were applied in *x* and *y* directions, while a fixed boundary condition was applied along the *z* direction [42].

With regards to the simulation procedure, at first energy minimization was carried out to remove any overlap among the particles. Subsequently, the system pressure was equilibrated to 1 atmosphere using the NP_zT ensemble. Finally, the equilibration and production run (12 ns) were carried out using the NVT ensemble. A detailed description of the equilibration time as well as the equilibration criteria can be found in our previous article as we have used the same data in this

study [37]. All simulations were carried out using the LAMMPS software package [43]. For visualization and illustration purposes we used OVITO [44] and Avogadro software[45]. The machine learning algorithms were implemented in python using the scikit learn library[46].

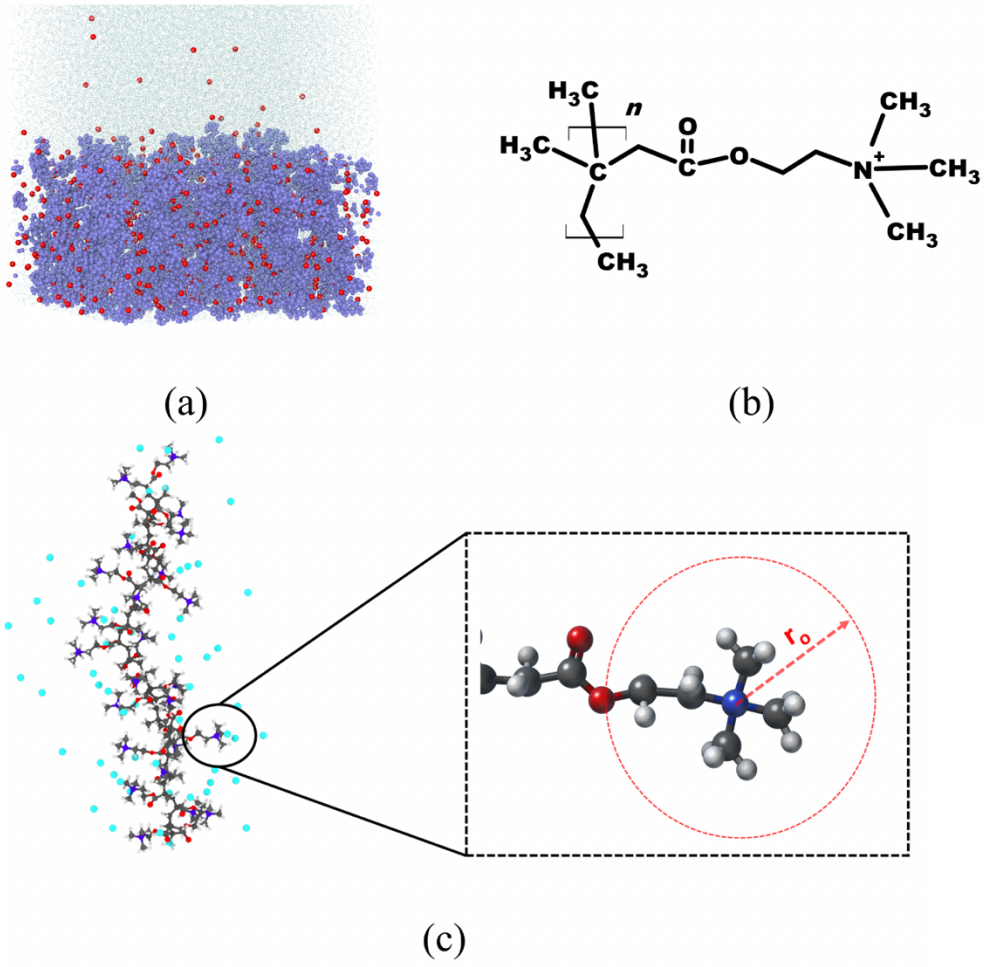

Figure 1. (a) A representative snapshot of the simulation system. Blue, red, and cyan colored particles respectively denote the brush atoms, the counterions, and the water molecules. (b) A scheme showing the chemical structure of a PMETAC monomer. Here *n* denotes the number of monomers (which is 24 in this case). (c) A snapshot showing an entire PMETAC polymer chain. The carbon, oxygen, nitrogen, hydrogen, and counterion particles are respectively colored in black, red, blue, white, and cyan. In (c), we have magnified the functional moiety of interest [namely, the $\{N(CH_3)_3\}^+$ moiety]. Water molecules inside a sphere of radius r_o (r_o =6.35 Å, i.e.,

the radius of the first solvation shell; r_0 has been identified in this magnified image) and centered around the N atom are considered when investigating the hydration state of the moiety.

2.2. Machine learning protocol

Any physical phenomenon can be described by relevant state variables. The interrelationship between the different state variables provides the necessary information about the underlying science [47]. One way to understand such interrelationships between the different state variables is to change some system parameters (i.e., apply some 'perturbations') and observe how the state variables change (i.e., identify the "effect" of such perturbations). In experiments, such "perturbations" are employed by changing some experimental conditions (or system parameters) repeatedly, plotting the variations of the different state variables (as functions of the imposed "perturbations"), and fitting appropriate functions that describe the interrelationships between these state variables. Any method or protocol that provides some information on the changes in the state variables (the "effect") with the changes in the system parameters (the "perturbations") is extremely valuable as it helps to elucidate the underlying governing mechanism dictating the system. In the present study, the scheme that we propose builds upon this principle: by detecting the changes in the state variables in the presence of the imposed changes in the system parameters for a microscopic system, we identify the presence of two separate hydration states around a single functional group [$\{N\text{-}(CH_3)_3\}^+$ group] of the cationic PMETAC chain inside the PMETAC PE brush layer.

In a macroscopic system, if the system is in equilibrium, each state variable has a single unique value. However, in a microscopic system (well below the thermodynamic limit), a state variable is described by a distribution (i.e., often by a probability distribution function or a PDF). Therefore, the overall state of the system is also described by a PDF, and accordingly, depending on the number of state variables, the number (and the dimensions) of the PDFs describing the overall state of the system will vary. The distribution of the state variables also provides information on

the appropriate physical factors driving the system (or the thermodynamic forces influencing the state variable). The presence of multiple peaks in the PDFs (or PDFs being represented by deformed Gaussians) can indicate that the state variable (described by the corresponding PDF) is affected by multiple physical factors, which gives rise to multiple modes or major states present in the system. Similar to macroscopic systems, detecting changes in the high dimensional probability distribution can reveal information about the underlying science (or the physical factors) governing the system observables. However, the detection of such changes is a tedious task and one often fails to identify such changes by employing classical methods, stemming from the fact that multiple state variables change simultaneously with the system parameters. The protocol that we propose in this paper addresses this issue by using an existing machine learning (ML) algorithm coupled with common data pre-processing techniques. A detailed description of the scheme has been provided below.

Fig. 2 shows a flow chart describing the entire scheme employed in this paper. The ***first step*** is to select the phenomenon that will be studied; for the present case, it will be the microscopic hydration state of the $\{N\text{-}(CH_3)_3\}^+$ moiety of the PMETAC brushes. In the ***second step,*** some appropriate state variables, relevant to describing this phenomenon, need to be selected. The criterion for selecting these state variables is that they should be weakly correlated or independent of each other so that the selected variables can together capture the multiple aspects (more information) of the selected phenomenon. In this study, as state variables, we select (1) the tetrahedral order parameter (q) of the water molecules solvating the $\{N\text{-}(CH_3)_3\}^+$ moiety, (2) the dipole orientation (θ) of the water molecules solvating the $\{N\text{-}(CH_3)_3\}^+$ moiety, and (3) the hydration number or the water coordination number (CN) of the $\{N\text{-}(CH_3)_3\}^+$ moiety. In Fig. 1(c), we have identified the water molecules that will be considered for this calculation, namely, the

water molecules inside a sphere that is of radius r_o and is centered around the N atom of the $\{N-(CH_3)_3\}^+$ moiety. Once these state variables have been selected, the ***third step*** involves sampling their individual distribution (or the probability distribution function or PDF) from a computational method of choice (in our case, all-atom MD simulations). Fig. 3(a) schematically represents such a PDF. The ***fourth step*** involves identifying the number of physical factors that are influencing the PDF of each state variable. For this purpose, first (***step* 4a**) the raw PDF function is smoothened by curve fitting to remove any noise or discontinuities. In our case, we smoothened the PDF [Fig. 3(b) shows such a smoothened pdf] by fitting multiple Gaussian functions. The number of Gaussians (***N***) that leads to the least RMSE (root mean square error) is selected to curve fit a particular PDF. At this point, as the next step (***step 4b***), one needs to identify the number of physical factors that affect the PDF (representing the state variable). For example, if only a single physical factor affects the state variable, one will have a single peak (in the smoothened PDF) and one can use a single Gaussian to model this smoothened PDF. On the other hand, if there are multiple physical factors affecting this state variable, there would be multiple peaks in the PDF, or there could be a sudden presence of a "bulge" or a localized peak in the PDF. Such a bulge or a peak often results in a sudden change in the statistical properties of the distribution [for example, the slope of the PDF; represented schematically in Fig. 3(c)] and the locations where such change occurs is known as the "change points" [see Fig. 3(b,c)]; there can be a host of methods for identifying such "change points" in a distribution [48]. For our case, we found these change points by detecting anomalous changes in the slope function (gradient) of the PDF. Multiple change points indicate the presence multiple physical factors influencing a given state variable; for such a case, one can fit more than one Gaussian (i.e., multiple secondary Gaussians, in addition to the main Gaussian) to model the PDF [see Fig. 3(d)]. Of course, these additional peaks or bulges

should occur at high probability values in order to ensure that the corresponding additional physical factors are significant and not merely "weak physical factors" that create irrelevant states (or the states that are attained by the system very few times). Under these circumstances, to conclusively identify the exact number of Gaussians that should be used to model the PDF, we identify the peaks of the slope function that occur at probability values of above "0.005" (or P(x)dx >0.005, here dx=$\frac{\text{range}}{200}$) (since such peaks refer to the relevant physical factors that significantly influence the particular state variable). For example, in the present case, we identify that the number of Gaussians that can be employed to fit the smoothened PDFs representing the variation of the q, θ, and CN are 2, 1, and 1, respectively (results are provided in Figs. 4-6). After identifying these appropriate number of Gaussians, as the ***fifth step***, a linear transformation is applied to separate the Gaussians for the case where a given PDF is fitted by multiple Gaussians (for example, the PDF for *q*, which is fitted by two Gaussians) [represented schematically in Fig. 3(e)]. This results in the generation of a modified high dimensional PDF and data points are sampled from that distribution. As the ***sixth step***, the sampled data points are then fed into a series of unsupervised machine learning (ML) algorithms. The separation of the Gaussians (see ***fifth step***) ensures that these machine learning (ML) algorithms find it easier to detect any changes in the state variables subject to the imposition of external "perturbations" (i.e., changes in the input conditions). For the present case, the ML algorithm consists of a Mode seeking and gaussian mixture model based clustering algorithm (discussed in detail in section S2 of the SI) [49]. The initial center and number of the clusters are obtained by using the mode-seeking algorithm (in this case, the mean shift [50]), while the points are grouped using the clustering algorithm (in this case, the Gaussian mixture model) [51] [schematically represented in Fig. 3(f)]. The number of clusters denotes the dominant states of the system. The changes in the cluster properties (e.g., changes in

center-to-center distance between the different clusters, number of points present in a given cluster, sphericity of the clusters, etc.) with the changes in the input conditions or system parameters ("perturbations") represent the "effect" triggered by these "perturbations" [see Fig. 3(g)]. These "effects" are generated by the modeled PDFs and can provide information about the underlying science and can also set the direction of the next step of analysis. If further analysis (through other classical methods) and data from existing literature support the generated information (or the "effect"), then the modeled high-dimensional PDF describing the system properties is accepted. This modeled PDF [see Fig. 3(h)] reveals major states (that are physically relevant) of the system and also the factors driving these states, which can further help to build a physical representation of the model by incorporating other information. If the effect cannot be verified or is not commensurate with existing knowledge, several steps need to be taken depending on the problem at hand. Such steps include changing the initial guess of the driving physical factors (by investigating "change points" of different statistical properties of the PDF), changing the state variables, changing the clustering algorithm, etc. Thus, iteratively, it is possible to reveal the underlying insights of the specific problem. An application of this protocol is used in this article, which will further help us understand how to use this protocol.

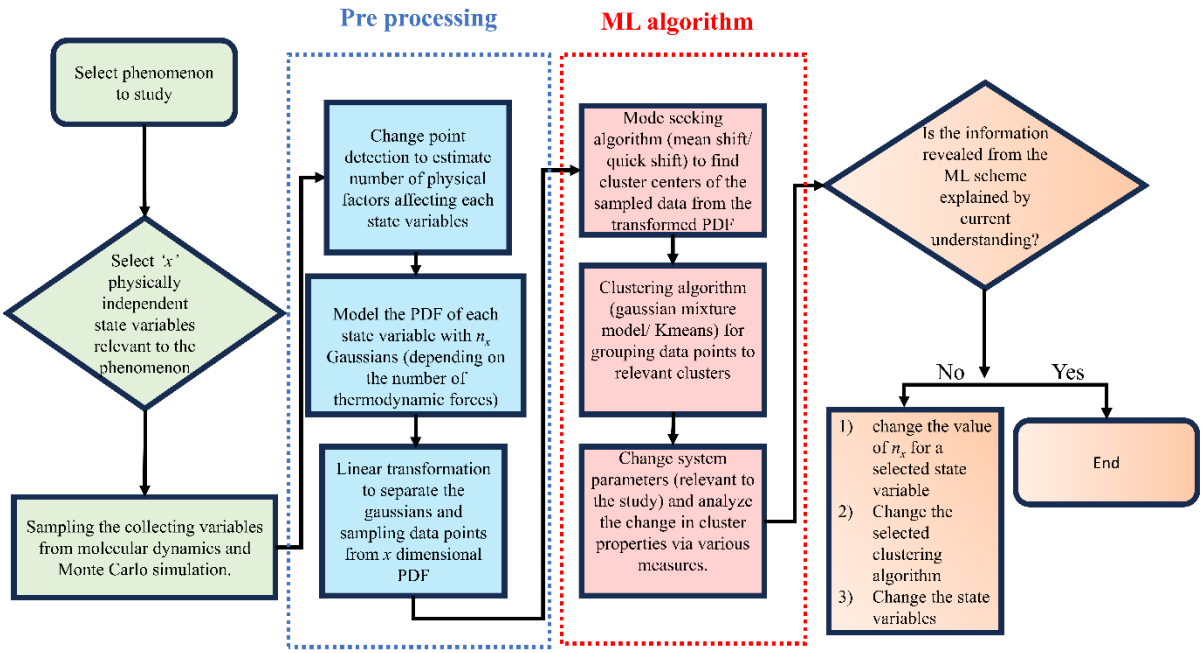

Figure 2. Flow chart of the scheme.

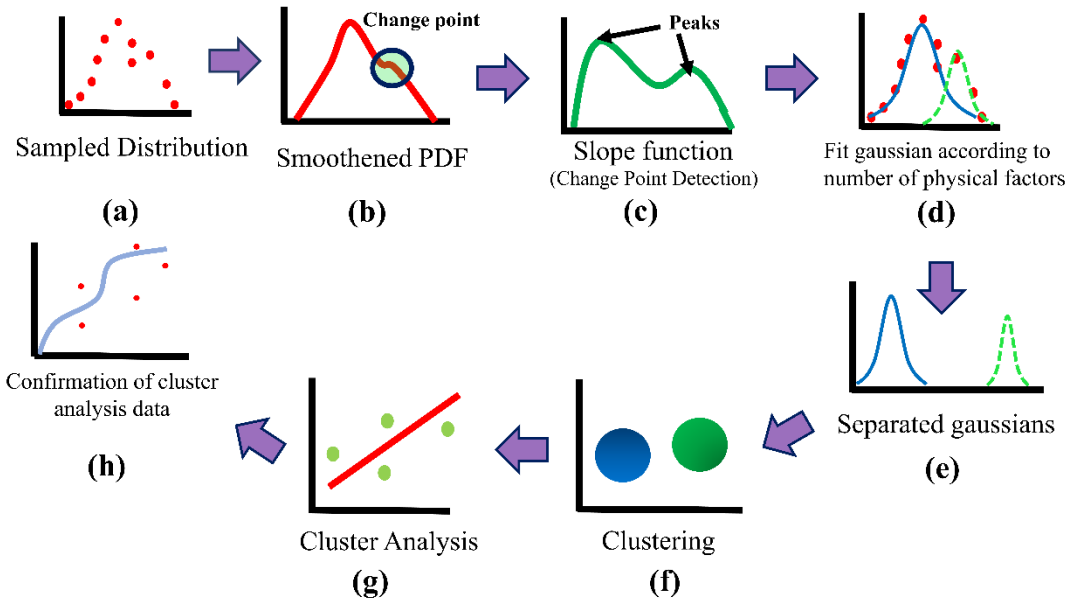

Figure 3. Schematic representation of the different steps to be employed in our proposed methodology.

3. RESULTS AND DISCUSSIONS

3.1. Choice of the state variables

As has been already pointed out above, the state variables that are selected to probe the hydration of the $\{N(CH_3)_3\}^+$ moiety of the grafted PMETAC chains are (1) the tetrahedral order parameter (q) of the water solvating the $\{N(CH_3)_3\}^+$ moiety, (2) the dipole orientation (θ) of the water solvating the $\{N(CH_3)_3\}^+$ moiety, and (3) the hydration number or the water coordination number (CN) of the $\{N(CH_3)_3\}^+$ moiety. These water molecules are those that are present within the sphere of radius r_0 around the N atom of the $\{N(CH_3)_3\}^+$ moiety [see Fig. 1(c)]. The CN denotes the number of water molecules in the 1st solvation shell of the selected moiety. It gives a measure of how well-hydrated the moiety is. The radius of the solvation shell is considered as the distance from the moiety center [N atom of the $\{N(CH_3)_3\}^+$] to the 1st minimum of the water-moiety RDF (Radial Distribution Function). Fig. S2 in section S3 in the SI provides the $\{N(CH_3)_3\}^+$-O_W (O_W is the oxygen atom of the water molecule) RDFs for different PMETAC brush grafting densities. The dipole orientation angle (θ) provides information on how the charge of the moiety affects the orientation of the water molecules surrounding the moiety. Finally, the tetrahedral order parameter (q) provides information on the structure of the solvating water molecules [37], [52]. All these state variables capture different aspects of the phenomenon to be studied. Please note, however, that this choice of the number and the types of state variables made by us for this study is not unique; other studies can choose different variables or more than three variables to study the same phenomenon. The distribution of these three state variables, q, θ, and CN, which are sampled from the all-atom MD simulation data for different grafting density values (σ_g), are shown in Figs. 4(a,f,k), 5(a,f,k), and 6(a,f,k). Following the procedure discussed in section 2.2., we first perform a curve fitting with Gaussian functions to remove any noise or discontinuities from the sampled

PDF using nonlinear regression method. The RMSE associated with this curve fitting depends on the number of Gaussians (N) chosen to fit a particular PDF; results showing the RMSE-*vs*-N variation have been provided in Figs. 4(b,g,l), Figs. 5(b,g,l), and Figs. 6(b,g,l). We choose that value for the number of Gaussians that yields the smallest value of the RMSE. The resultant PDF (resulting from fitting the original PDF with the selected number of Gaussians) is a significantly smoothened. These smoothened PDFs corresponding to the different state variables and different grafting density values have been shown in Figs. 4(c,h,m), Figs. 5(c,h,m), and Figs. 6(c,h,m). These smoothened PDFs enable us to identify the number of physical factors affecting the individual state variables (q, θ, and CN) for different grafting density values. For that purpose, we identify the peaks of the slope function (which is the gradient of these smoothened PDFs) that occur at probability values above "0.005" [or P(x)dx >0.005, here dx=$\frac{\text{range}}{200}$]. These gradient functions and the corresponding peaks have been identified in Figs. 4(d,i,n), Figs. 5(d,i,n), and Figs. 6(d,i,n): for all values of the grafting density, we observe two peaks in the slope function of the PDF for the case of the state variable q, but a single peak corresponding to the slope function of the PDF for the state variables θ and CN. Under these circumstances, we can infer that for all the grafting density values, the state variable q is influenced by two physical factors, i.e., one will require two Gaussians to fit the smoothened PDF representing the variation of q; these two Gaussians, after being separated by a linear transformation, have been shown in Figs. 4(e), 5(e), and 6(e). On the other hand, for all the grafting density values, the state variables θ and CN are dictated by one physical factor, and accordingly one will require a single Gaussian to fit and model the PDFs representing the variation of θ and CN; these Gaussians corresponding to θ are shown in Figs. 4(j), 5(j), and 6(j), while the Gaussians corresponding to CN are shown in Figs. 4(o), 5(o), and 6(o). The modeled PDFs of q, θ and CN are linearly scaled to a different range and normalized.

Accordingly, these variables are now denoted as q' [see Figs. 4(e), 5(e), and 6(e)], θ' [see Figs. 4(j), 5(j), and 6(j)], and CN' [see Figs. 4(o), 5(o), and 6(o)]. Linear scaling does not distort the information, and the normalization was performed such that the area under the curve integrates to 1. The range of the scaled state variables are: q': -150 to 150; CN': 0 to 100 and θ' : 0 to 100. The data sampled from these modeled PDFs are next fed to the ML protocol. Results from the ML protocol are discussed next.

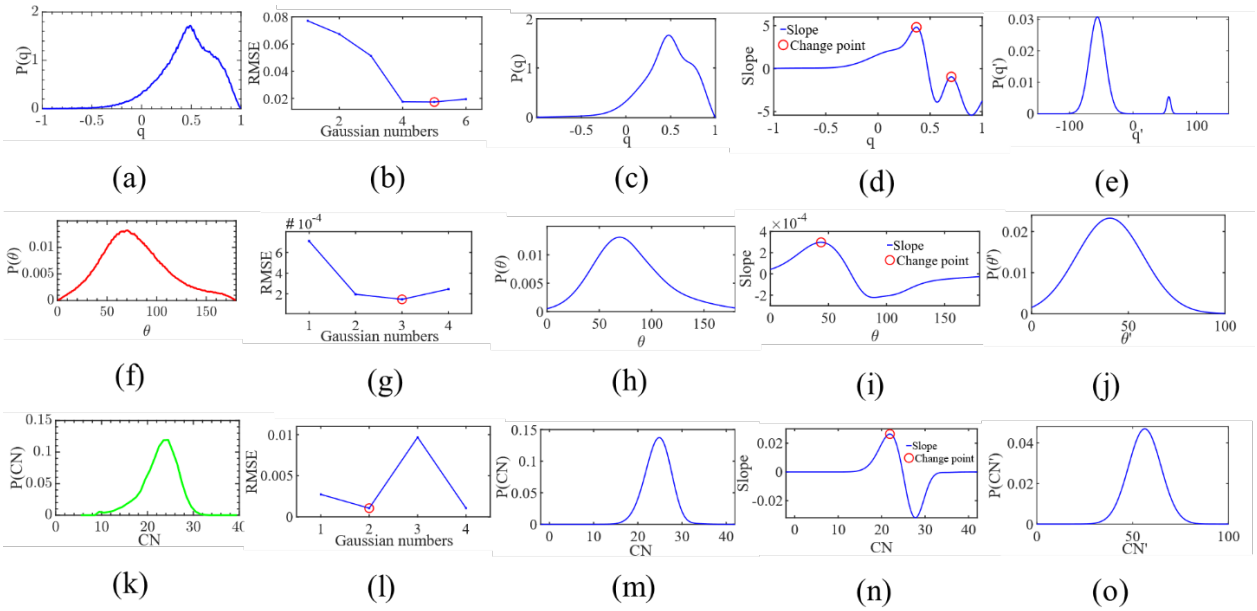

Figure 4. **(a,f,k):** Probability distribution functions (PDFs), sampled from the all-atom MD simulation data, corresponding to **(a)** tetrahedral order parameter (q), **(f)** dipole orientation angle (θ) (expressed in degrees), and **(k)** co-ordination number (CN). **(b,g,l):** RMSEs as functions of the number of Gaussians fitted to the PDFs (in order to smoothen the PDFs) corresponding to **(b)** tetrahedral order parameter (q), **(g)** dipole orientation angle (θ) (expressed in degrees), and **(l)** co-ordination number (CN). **(c,h,m):** Smoothened PDFs corresponding to **(c)** tetrahedral order parameter (q), **(h)** dipole orientation angle (θ) (expressed in degrees), and **(m)** co-ordination number (CN). **(d,i,n):** Variation of the function representing the slope of these smoothened PDFs corresponding to **(d)** tetrahedral order parameter (q), **(i)** dipole orientation angle (θ), and **(n)** co-ordination number (CN). The peaks of these slopes have been identified with red circles. **(e,j,o):** Gaussians used to fit the smoothened PDFs corresponding to **(e)** tetrahedral order parameter (q), **(j)** dipole orientation angle (θ) (expressed in degrees), and **(o)** co-ordination number (CN). The number of Gaussians used is equal to the number of peaks in the function representing the slope [see **(d,i,n)**]: hence for the case of the parameter (q), two Gaussians are used to fit the smoothened PDF [in (e), these two Gaussians, separated by a linear transformation, have been shown]. In

(e,j,o), these Gaussians fitted to the PDFs were linearly scaled to a different range and normalized (see the text for details); accordingly, these variable are denoted as $q^{/}$ [see (e)], $\theta^{/}$ [see (j)], and $CN^{/}$ [see (o)]. All the results provided here are for a PMETAC brush grafting density of σ_g = 0.15 chains/nm^2.

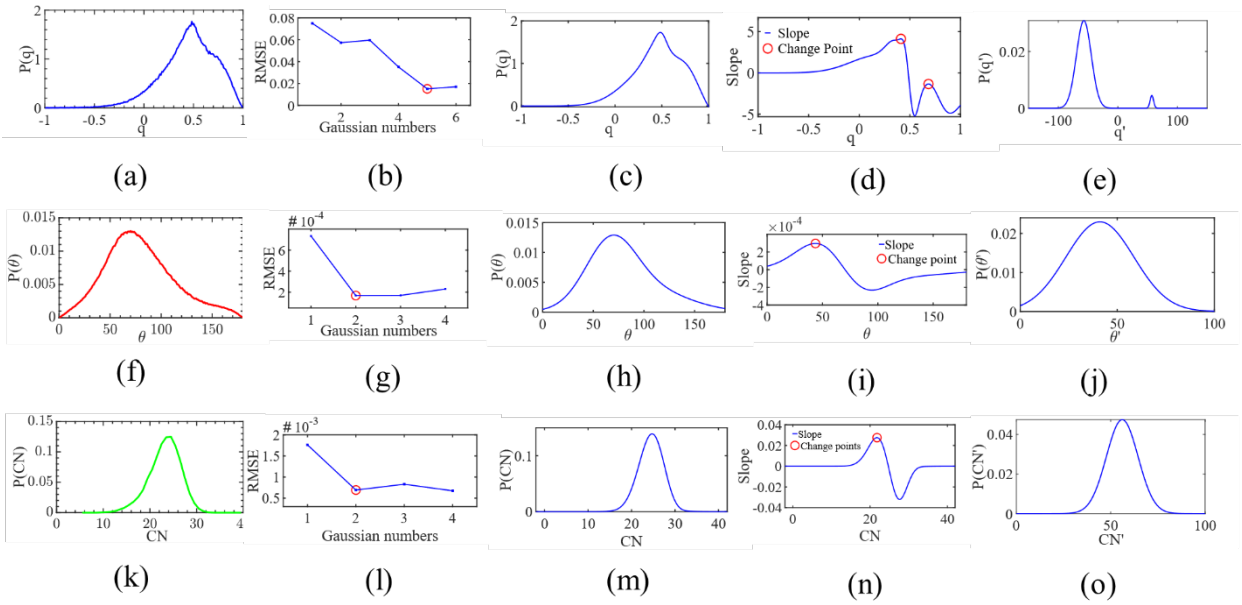

Figure 5. **(a,f,k):** Probability distribution functions (PDFs), sampled from the all-atom MD simulation data, corresponding to **(a)** tetrahedral order parameter (q), **(f)** dipole orientation angle (θ) (expressed in degrees), and **(k)** co-ordination number (CN). **(b,g,l):** RMSEs as functions of the number of Gaussians fitted to the PDFs (in order to smoothen the PDFs) corresponding to **(b)** tetrahedral order parameter (q), **(g)** dipole orientation angle (θ) (expressed in degrees), and **(l)** co-ordination number (CN). **(c,h,m):** Smoothened PDFs corresponding to **(c)** tetrahedral order parameter (q), **(h)** dipole orientation angle (θ) (expressed in degrees), and **(m)** co-ordination number (CN). **(d,i,n):** Variation of the function representing the slope of these smoothened PDFs corresponding to **(d)** tetrahedral order parameter (q), **(i)** dipole orientation angle (θ), and **(n)** co-ordination number (CN). The peaks of these slopes have been identified with red circles. **(e,j,o):** Gaussians used to fit the smoothened PDFs corresponding to **(e)** tetrahedral order parameter (q), **(j)** dipole orientation angle (θ) (expressed in degrees), and **(o)** co-ordination number (CN). The number of Gaussians used is equal to the number of peaks in the function representing the slope [see **(d,i,n)**]: hence for the case of the parameter (q), two

Gaussians are used to fit the smoothened PDF [in (e), these two Gaussians, separated by a linear transformation, have been shown]. In (e,j,o), these Gaussians fitted to the PDFs were linearly scaled to a different range and normalized (see the text for details); accordingly, these variable are denoted as q' [see (e)], θ' [see (j)], and CN' [see (o)]. All the results provided here are for a PMETAC brush grafting density of $\sigma_g = 0.20$ chains/nm^2.

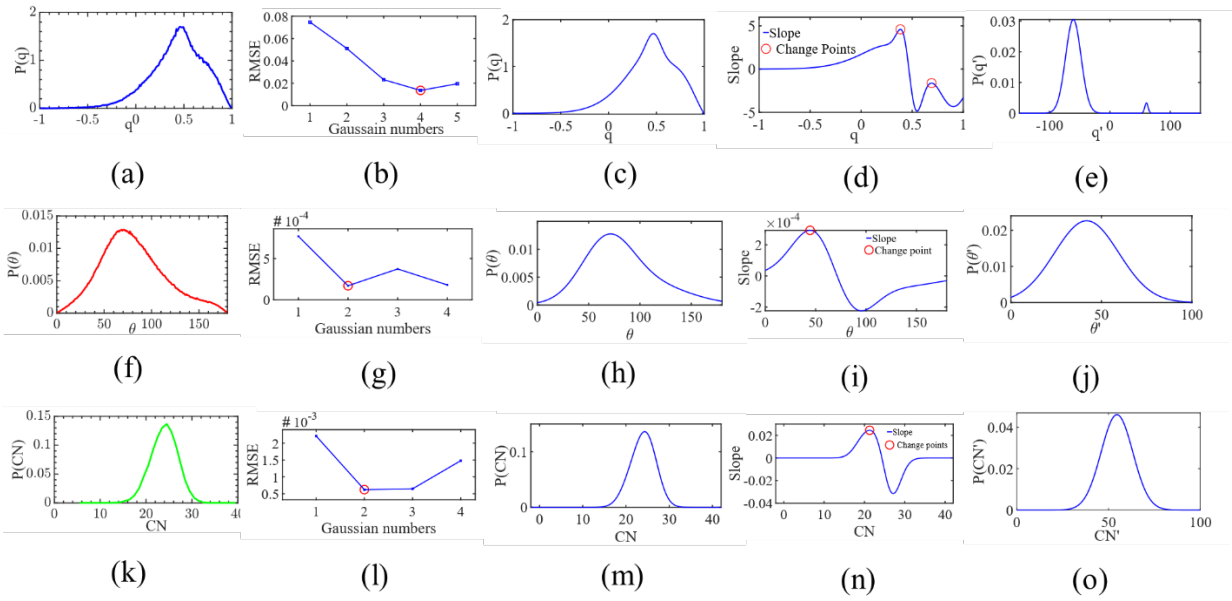

Figure 6. **(a,f,k):** Probability distribution functions (PDFs), sampled from the all-atom MD simulation data, corresponding to **(a)** tetrahedral order parameter (q), **(f)** dipole orientation angle (θ) (expressed in degrees), and **(k)** co-ordination number (CN). **(b,g,l):** RMSEs as functions of the number of Gaussians fitted to the PDFs (in order to smoothen the PDFs) corresponding to **(b)** tetrahedral order parameter (q), **(g)** dipole orientation angle (θ) (expressed in degrees), and **(l)** co-ordination number (CN). **(c,h,m):** Smoothened PDFs corresponding to **(c)** tetrahedral order parameter (q), **(h)** dipole orientation angle (θ) (expressed in degrees), and **(m)** co-ordination number (CN). **(d,i,n):** Variation of the function representing the slope of these smoothened PDFs corresponding to **(d)** tetrahedral order parameter (q), **(i)** dipole orientation angle (θ), and **(n)** co-ordination number (CN). The peaks of these slopes have been identified with red circles. **(e,j,o):** Gaussians used to fit the smoothened PDFs corresponding to **(e)** tetrahedral order parameter (q), **(j)** dipole orientation angle (θ) (expressed in degrees), and **(o)** co-ordination number (CN). The number of Gaussians used is equal to the number of peaks in the function representing the slope [see **(d,i,n)**]: hence for the case of the parameter (q), two Gaussians are used to fit the smoothened PDF [in (e), these two Gaussians, separated by a linear

transformation, have been shown]. In (e,j,o), these Gaussians fitted to the PDFs were linearly scaled to a different range and normalized (see the text for details); accordingly, these variable are denoted as q' [see (e)], θ' [see (j)], and CN' [see (o)]. All the results provided here are for a PMETAC brush grafting density of $\sigma_g = 0.25$ chains/nm^2.

3.2 Results of the Machine Learning protocol

We consider a representative of the sampled data ($q^{/}$, $\theta^{/}$, and $CN^{/}$) obtained from the modeled PDFs [see Figs. 4(e,j,o)], corresponding to a PMETAC brush grafting density of σ_g = 0.15 chains/nm^2. The data points are normalized $\left(\frac{data - mean}{standard\ deviation}\right)$ and are represented as $q_n^{/} - \theta_n^{/} - CN_n^{/}$ tupule (here subscript 'n' denotes normalized data) [Fig. 7(a)]. After that, we used ML algorithms to group the data points. At first, the mean shift algorithm was used to identify the number of modes and their location in the data set. Then Gaussian mixture model clustering was used to segregate the points. The number of modes was used as the number of clusters (or Gaussians) and the locations of these modes were used as the initial guesses for the mean values of the clustering Gaussians (see section S2 in the SI for greater details). Each cluster is a representation of a separate hydration state around the $\{N(CH_3)_3\}^+$ moiety. Therefore, corresponding to σ_g = 0.15 chains/nm^2, we find two distinct hydration states, represented by two distinctly colored clusters [see Fig. 7(b)].

We next aim to employ the clustering (ML) algorithm to obtain the effect of changing the grafting density (i.e., the effect of exerting a "perturbation") in altering these hydration states. For that purpose, we first consider the representative sampled data ($q^{/}$, $\theta^{/}$, and $CN^{/}$) corresponding to the PMETAC brush grafting densities of σ_g = 0.20 chains/nm^2 and σ_g = 0.25 chains/nm^2, and subsequently obtain the corresponding $q_n^{/} - \theta_n^{/} - CN_n^{/}$ tupule [see Figs. 7(c) and 7(e) for these tupules for σ_g = 0.20 chains/nm^2 and σ_g = 0.25 chains/nm^2] and the corresponding clustered data sets [see Figs. 7(d) and 7(f) for these clusters for σ_g = 0.20 chains/nm^2 and σ_g = 0.25 chains/nm^2]. To capture the effect of the PMETAC brush grafting density, we quantified, as functions of the PMETAC brush grafting density values, various distance metrics (e.g., radius of gyration of the

clusters and center-to-center distance between the clusters) that have direct correlations with various thermodynamic measures (see Fig. S3 in section S4 in the SI). The changes in the radii of gyration of the clusters denote the changes in fluctuations within the major states, while the changes in the center-to-center distances between the clusters point to the differences in the nature between the states. Fig. S3 in the SI confirms little variation in the cluster radius of gyration (within the major states) and the cluster center-to-center distance with the PMETAC brush grafting density. Hence, the study of these quantities (cluster radius of gyration and cluster center-to-center distance) cannot provide much information on the effect of the change in the grafting density on the hydration states of the $\{N(CH_3)_3\}^+$ moiety of the PMETAC brushes. However, we do see that with an increase in the grafting density, the number of points in state 1 [or clusters representing state 1, or clusters colored in blue in Fig. 7(b,d,f)] becomes significantly greater than the number of points in state 2 [or clusters representing state 2, or clusters colored in red in Fig. 7(b,d,f)]. In other words, the population of state 2 is decreasing significantly with an increase in the grafting density. Fig. 7(g) also confirms this observation. Such a behavior suggests that with an increase in the grafting density, state 1 becomes more dominant than state 2. Moreover, we can observe that the major difference between state 2 and state 1 is that state 2 has a higher q parameter (indicating more structured water) than state 1. These are two major pieces of information, as we can now direct the next step of the analysis to find out the physical factors that are causing state 2 (hydration state with more structured water molecules) to become less important with an increase in the grafting density. In summary, by using the ML algorithm we have (1) identified two distinct hydration states (one consisting of more structured water molecules and another consisting of less structured water molecules) of the $\{N(CH_3)_3\}^+$ moiety of the PMETAC brushes and (2) pinpointed the manner in which a change in the grafting density affects these distinct hydration states (an

increase in the grafting density makes the state with less structured water more favored). In the next section, further analysis will show that the generated effect is in harmony with the chemical understanding and results from previous studies.

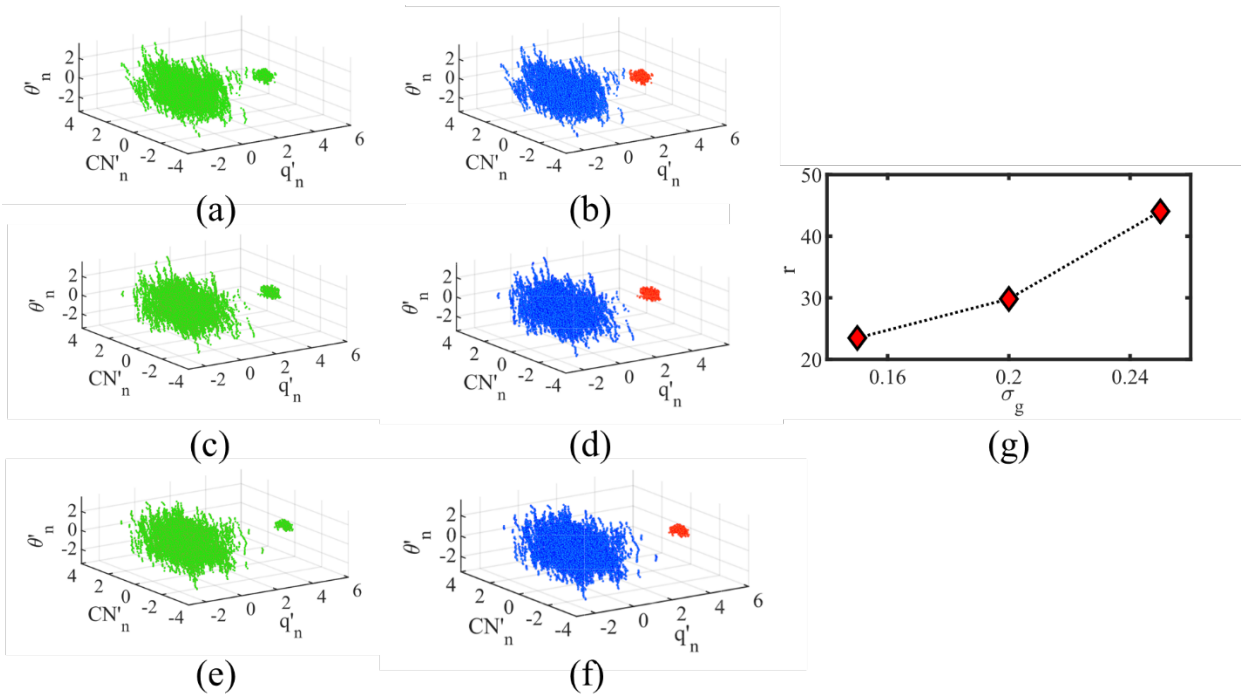

Figure 7. (a,c,e): Sampled normalized data points, represented as $q_n^{/} - \theta_n^{/} - CN_n^{/}$ tupule, obtained from the modeled PDF for **(a)** σ_g = 0.15 chains/nm^2 [corresponding Gaussians can be found in Figs. 4(e,j,o)], **(c)** σ_g = 0.20 chains/nm^2 [corresponding Gaussians can be found in Figs. 5(e,j,o)], and **(e)** σ_g = 0.25 chains/nm^2 [corresponding Gaussians can be found in Figs. 6(e,j,o)]. **(b,d,f)** Data clustered by employing the ML algorithm for **(b)** σ_g = 0.15 chains/nm^2, **(d)** σ_g = 0.20 chains/nm^2, and **(f)** σ_g = 0.25 chains/nm^2. The blue and red clusters [in (b,d,f)] respectively represent the two clusters (or two hydration states; blue: State 1 and red: State 2) detected by the ML scheme. **(g)** Variation of the ratio of points of the two clusters [in (b,d,f)], (represented by r = $\frac{points\ in\ cluster\ 1}{points\ in\ cluster\ 2}$), with the grafting density.

3.3 Two-state Hydration model as Explained By Existing Understanding

In our recent paper [36], we have found that the $\{N(CH_3)_3\}^+$ moiety of the PMETAC brushes induces apolar hydration due to low charge density. As a result, the water molecules present in the hydration layer of the $\{N(CH_3)_3\}^+$ group tend to interact strongly among themselves. Such enhanced interaction results in a higher *q* parameter value (or enhanced tetrahedral ordering) of the solvation shell water molecules, as compared to the water molecules solvating other functional groups (e.g., C=O) and counterions (Cl^- ions) present in the PMETAC brush system [37]. Moreover, experimental studies have also confirmed that water molecules solvating TMA (tri-methyl ammine) and similar salts have formed such apolar hydration layer and are characterized by a higher structural ordering [53]. Moreover, in one of our another previous papers [36], we found that at high grafting densities, i.e., under conditions when a greater percentage of water molecules inside the grafted PMETAC brush layer serve as the solvating water of the $\{N(CH_3)_3\}^+$ moiety, the "donor-oxygen-acceptor-oxygen-donor-hydrogen" angle of the solvating water has a lower standard deviation (less spread) than the bulk water. Such a situation indicates stronger hydrogen bonds (enhanced interaction) between the water molecules, which can be attributed to the presence of $\{N(CH_3)_3\}^+$ moiety.

As already pointed out, the machine learning protocol, acting upon the sampled data set obtained from the modelled PDF, confirms the presence of two hydration states (with the state with water with a greater *q* value disappearing with an increase in the grafting density) of the $\{N(CH_3)_3\}^+$ moiety of the PMETAC brushes. Previously, we have found that inside the PMETAC brush layer, the solvating water of the Cl^- ions, which serve as the counterions screening the charges of the $\{N(CH_3)_3\}^+$ moiety of the PMETAC brushes, have a *q* parameter that is smaller than the *q* parameter of the water molecules solvating the $\{N(CH_3)_3\}^+$ species [37]. The information

generated by the ML algorithm and the above-mentioned fact has motivated us to look into the water molecules around the region of the moiety of interest [in our case, $\{N(CH_3)_3\}^+$] and Cl^- counterion. Fig. 8(a) shows the percentage of hydration water of the $\{N(CH_3)_3\}^+$ moiety that is shared with other moieties inside the chloride-ion-screened PMETAC brush layer for a PMETAC brush grafting density of 0.15 chains/nm^2.

The water molecules inside the PMETAC brush can be divided into two types: free water which is not inside the solvation shell of an ionic species (thus not affected by the charged species) and bound water which is inside the solvation shell of an ionic species (and hence affected by the charged species). As the grafting density is increased, the % of bound water molecules increases for both $\{N(CH_3)_3\}^+$ [see Fig. 8(b)] and Cl^- counterion [see Fig. 8(c)]; such a scenario triggers the possibility of competition for solvation water (locally) between the charged species. More importantly, the percentage of the hydration shell water molecules of the $\{N(CH_3)_3\}^+$ moiety that is shared with that of the Cl^- counterion also increases [see Fig. 8(d)]. Given the fact that water molecules solvating the Cl^- ion have a smaller q (i.e., the water molecules are less structured) [37], such an increase in the share of the solvation water implies that an increase in the PMETAC brush grafting density causes a decrease in the number of the highly structured water molecules around the $\{N(CH_3)_3\}^+$ moiety. This explains the reason for the hydration state 2 (*i.e.*, the state where water molecules are more structured) vanishing with an increase in the grafting density [see Fig. 7(g)]. *Therefore, the findings of our combined ML-MD approach, which points to the possibility of the two-state hydration water model around the $\{N(CH_3)_3\}^+$ moiety of the PMETAC brushes, can be explained with our existing understanding [37] about the water and counterion structure around the PMETAC brushes.* We also calculated the change (if any) of the efficiency of screening of the $\{N(CH_3)_3\}^+$ moiety by the Cl^- counterions with an increase in the grafting density [see Fig.

8(e)] in order to pinpoint if the counterions remained inside the brush layer (for enhanced grafting density). This is important given the fact that a weakening in screening, which is tantamount to the expulsion of counterions from inside the PMETAC brush layer, with an increase in the grafting density, can indicate a different cause for the progressive vanishing of state 2 (See section S2.3 in the SI for the calculation details). We find that the screening efficiency remains high and unchanged with an increase in the grafting density, implying that for all values of the grafting density, the counterions remain inside the brush layer. To rule out other possibilities for the disappearance of state 2 with an increase in the grafting density, we also probed the effect of the increasing the grafting density on water-water hydrogen bonding inside the PMETAC brush layer. The results confirm that while the number of hydrogen bonds per water molecules decreased with an increase in the grafting density, the conditions (namely the critical value of donor-oxygen-acceptor-oxygen distance and the critical value of the donor-oxygen-acceptor-oxygen-donor-hydrogen angle) that dictated the formation of water-water hydrogen bonds remained unaffected by an increase in the grafting density (see Figs. S4(c,d) in section S5 in the SI). Such negligible impact of the change in the grafting density on these different quantities establish our inference that the disappearance of the hydration state 2 [of the $\{N(CH_3)_3\}^+$ moiety] due to an increase in the grating density can be solely attributed to the enhanced sharing of the solvation shell water molecules between the $\{N(CH_3)_3\}^+$ moiety and the Cl^-.

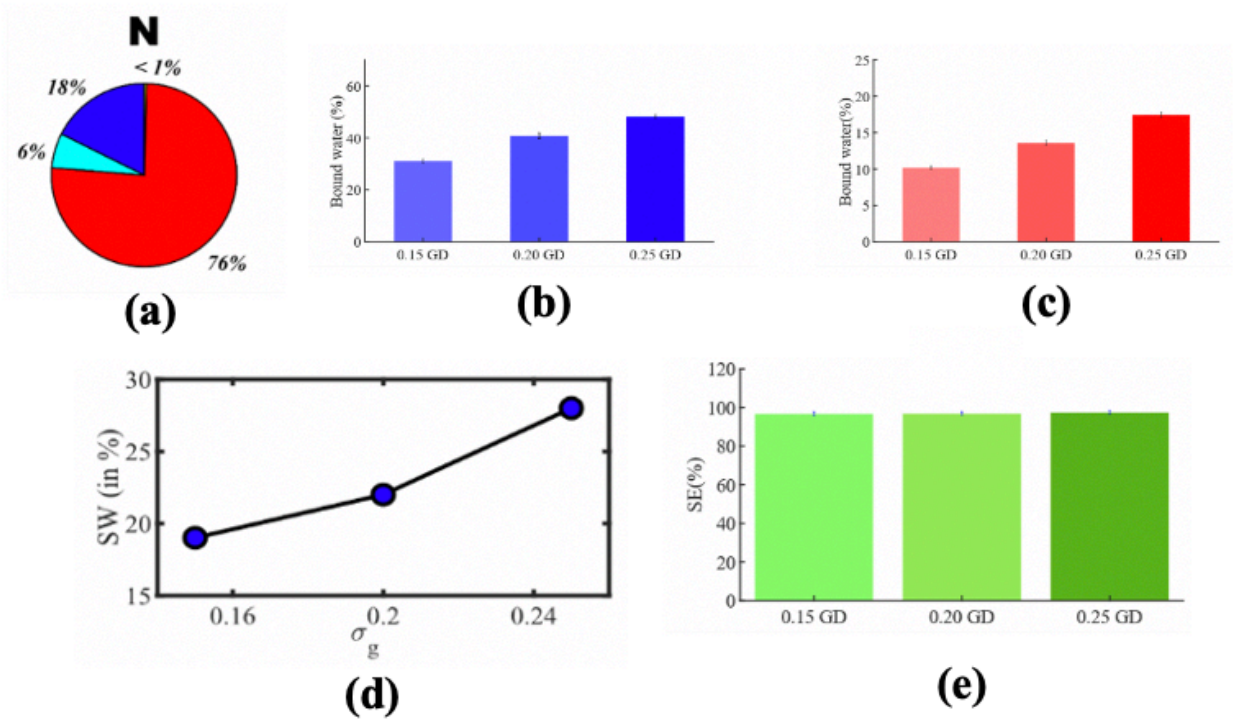

Figure 8: (a) A pie chart showing a representative case (σ_g=0.15 chains/nm^2) of the percentage of water molecules in the hydration shell of the $\{N(CH_3)_3\}^+$ moiety (of the PMETAC chains inside the PMETAC brush layer) shared with itself and other moieties as indicated with different colors [red: with itself; blue: with Cl^-; light blue: with C=O]. The results have been reproduced from Ref. [37]. **(b)** Percentage of the water molecules (for different grafting densities) inside the PMETAC brush that are bounded or inside the solvation shell of $\{N(CH_3)_3\}^+$ moiety. **(c)** Percentage of the water molecules (for different grafting densities) inside the PMETAC brush that are bounded or inside the solvation shell of Cl^- counterion. **(d)** Variation in the shared water (SW) molecules (in percentage) between the $\{N(CH_3)_3\}^+$ moiety and Cl^- ions as a function of the grafting density. (**e**) Screening efficiency of the counterions for different grafting densities.

3.4 A More Complete Picture of the Hydration of the $\{N(CH_3)_3\}^+$ moiety of the PMETAC Brushes In the Light Existing Results

Our combined MD-ML approach has unraveled a two-state hydration condition of the $\{N(CH_3)_3\}^+$ moiety of the PMETAC brushes. However, these findings must be coherent with the existing understanding of the water structure and counterion (Cl^- ion) distribution around the $\{N(CH_3)_3\}^+$ moiety, as revealed by existing studies (including our previous study [37]). In this section, we discuss the specific water and ion configuration that ensures such a coherence, thereby providing a more complete picture of the hydration of the $\{N(CH_3)_3\}^+$ moiety of the PMETAC brushes. Prior to that, let us first summarize the major information about the water molecules and counterions around the $\{N(CH_3)_3\}^+$ moiety inside the PMETAC brush layer that have been unraveled by the current combined MD-ML approach and by the existing previous studies.

1) There are two major states of hydration (one with more structured water molecules and another with less structured water molecules) around the $\{N(CH_3)_3\}^+$ moiety. With an increase in the grafting density, the state with more structured water molecules becomes less prevalent.
2) The percentage of the shared water molecules between the Cl^- ion and the $\{N(CH_3)_3\}^+$ moiety increases with an increase in the PMETAC brush grafting density [see Fig. 8(d)]. These findings are commensurate with the progressive disappearance of the hydration state [around the $\{N(CH_3)_3\}^+$ moiety] with more structured water molecules with an increase in the PMETAC brush grafting density.
3) The chloride counterions inside the PMETAC brush layer have very high mobility [see Fig. 7(d) and 7(e) of Ref. [37] and are also supported experimentally by [54]].

4) The average distance between the chloride ion and the $\{N(CH_3)_3\}^+$ moiety remains constant and does not change with the grafting density [see Fig. 4(e) of Ref. [37]].

The two-state hydration model of the $\{N(CH_3)_3\}^+$ moiety, as proposed in this study, is *commensurate* with the chloride counterions demonstrating very high mobility inside the PMETAC brush layer. This high mobility of the chloride ions has been attributed to the weak ion pairing between the $\{N(CH_3)_3\}^+$ moiety of the PMETAC chains and the chloride ions [37]. Moreover, we also identified that the $\{N(CH_3)_3\}^+$ and Cl^- interact as a solvent-separated ion pair, which means there is an intervening water layer present between $\{N(CH_3)_3\}^+$ and Cl^- and the attraction between the functional group and the counterion is not strong enough to expunge the intervening water layer [37]. As a result, water always remains closer to the $\{N(CH_3)_3\}^+$ functional group or there is a continuous hydration shell surrounding the moiety. Of course, the Cl^- ions, due to their large mobilities, can occupy various relative distances from the $\{N(CH_3)_3\}^+$ functional group for a given grafting density [see Fig. 7(b) in [37]]. When the Cl^- ion is significantly away from the $\{N(CH_3)_3\}^+$ functional group, the water structure around the $\{N(CH_3)_3\}^+$ is less disrupted, while when it is closer, the water structure is more disrupted. This explains how the 2-state hydration model is commensurate with high mobility of the Cl^- ions inside the brush layer.

For increasing grafting densities, the average distance between the $\{N(CH_3)_3\}^+$ and Cl^- ion does not change [see Fig. 4(e) of Ref. [37]], even though for greater grafting densities, the disruption in the apolar hydration structure of the solvation water becomes more pronounced. This information suggests that the disruption of the water structure around the $\{N(CH_3)_3\}^+$ moiety is not caused by the enhanced penetration (into the solvation layer) of the Cl^- ion. Rather, the increase in the propensity of the $\{N(CH_3)_3\}^+$ moiety to be in the hydration state 1 with an increase in the grafting density can be attributed to the corresponding increase in the number of counterions in a given

volume around a given $\{N(CH_3)_3\}^+$ (see Fig. 9) and the fact that the water molecules in the hydration layer of these counterions are always significantly less structured.

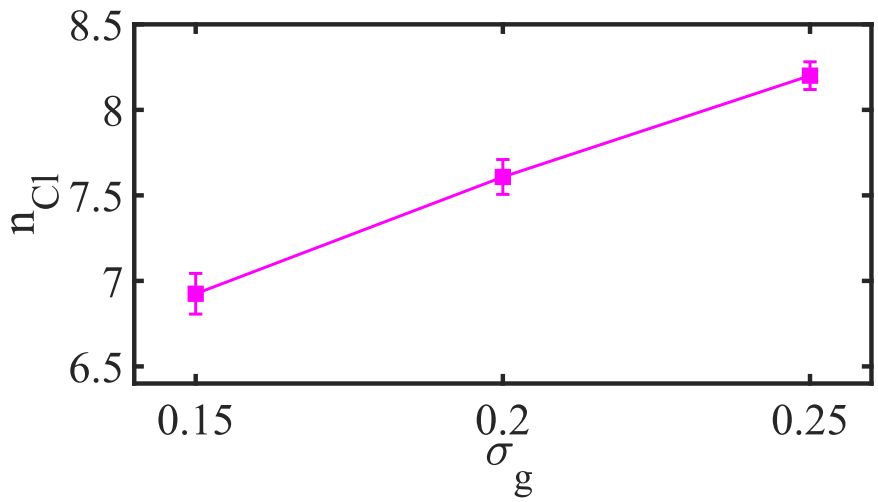

Figure 9: PMETAC-grafting-density-driven variation of the average number of Cl^- ion (n_{Cl^-}) inside a sphere (of radius 7Å) centered around the N atom of $\{N(CH_3)_3\}^+$ moiety.

CONCLUSIONS

The nature of hydration of macromolecules is extremely important given the often-unavoidable presence of water molecules in such macromolecular systems. Here, we used a combination of MD simulations and ML to probe into the science of hydration of the $\{N(CH_3)_3\}^+$ moiety of the PMETAC PE chains inside the densely grafted PMETAC brush layer. The combined approach revealed that there are two major hydration states of the $\{N(CH_3)_3\}^+$ functional group inside brush layer: one of the states consists of water molecules that are less structured, while the other state has more structured water molecules. At high grafting densities, due to the Cl^- crowding (instigated by enhanced confinement), the state with more structured water molecules become less prevalent. Finally, we argue that our findings are commensurate with existing results of enhanced counterion mobility (inside PMETAC brush layer) and invariance (with grafting density) of the $\{N(CH_3)_3\}^+$-Cl^- average distance.

On a different note, through this study, we have developed a combined MD-ML approach that can be easily modified and adapted to study several other phenomena of soft matter systems. Our approach will be useful to polymer community when trying to analyze a phenomenon that can be described by a high number of features. Furthermore, our results suggesting the dual hydration of $\{N(CH_3)_3\}^+$ opens up several questions related to such systems, such as the effect of such two-state hydration on several other properties (e.g., heat capacity, dielectric constant, etc.) of the PMETAC brush systems; the effect of counterion types and valences on such two-state hydration; etc. The answers to these questions should be sought in future research studies.

Supporting Information

MD simulation details; Machine Learning algorithm details; Calculation of Screening Efficiency; $N(CH_3)_3^+$-Cl^- and $N(CH_3)_3^+$-O_w radial distribution functions (RDFs) for different grafting densities; Change of radius of gyration of the clusters for different grafting density; Change of cluster-cluster centroid distance for different grafting density; Hydrogen bond analsis; Schematic depiction of the two state model; Schematic depiction of the vanishing of the second state due to Cl^- crowding.

Acknowledgement: This work has been supported by the Department of Energy Office of Science grant DE-SC0017741. The authors also gratefully acknowledge the Zaratan High-Performance Computing cluster at the University of Maryland for providing necessary computational resources and Dr. Turash Haque Pial for useful discussions.

References:

1. Funel, M. C. B. ; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.; Sterpone, F.; Spoel, D. V. D.; Yao, Xu; Garcia, A. E. Water Determines the Structure and Dynamics of Proteins. *Chem. Rev.* **2016**, *116 (13),* 7673–7697.
2. Cimatu, K. L. A.; Ambagaspitiya, T. D.; Premadasa, U. I.; Adhikari, N. M.; Kruse, A.; Robertson, E.; Guan, S.; Rong, L.; Advincula, R.; Bythell, B. J. Polymer-solvent interaction and Conformational Changes at a Molecular level: Implication to Solvent-Assisted Deformation and Aggregation at the Polymer Surface. *J. Colloid Interface Sci.* **2022**, *616*, 221-233.
3. White, A.; Jiang, S. Local and Bulk Hydration of Zwitterionic Glycine and Its Analogues through Molecular Simulations. *J. Phys. Chem. B*. **2011**, 115 (4), 660–667.
4. Klimkevicius, V.; Voronovic, E.; Jarockyte, G.; Skripka, A.; Vetrone, F.; Rotomskis, R.; Katelnikovas, A.; Karabanovas, V. Polymer Brush Coated Upconverting Nanoparticles with Improved Colloidal Stability and Cellular Labeling. *J. Mater. Chem. B.* **2022**, *10* (4), 625–636.
5. Li, F.; Jian, Y.; Xie, Z.; Liu, Y.; Liu, Q. Transient Alternating Current Electroosmotic Flow of a Jeffrey Fluid through a Polyelectrolyte-Grafted Nanochannel. *RSC Adv.* **2017**, *7* (2), 782–790.
6. Jiang, R.; Lu, X.; Yang, M.; Deng, W.; Fan, Q.; Huang, W. Monodispersed Brush-like Conjugated Polyelectrolyte Nanoparticles with Efficient and Visualized SiRNA Delivery for Gene Silencing. *Biomacromolecules* **2013**, *14* (10), 3643–3652.
7. Yang, Q.; Li, L.; Zhao, F.; Han, H.; Wang, W.; Tian, Y.; Wang, Y.; Ye, Z.; Guo, X. Hollow Silica–Polyelectrolyte Composite Nanoparticles for Controlled Drug Delivery. *J. Mater Sci.* **2019**, *54* (3), 2552–2565.
8. Ali, M.; Ramirez, P.; Mafé, S.; Neumann, R.; Ensinger, W. A PH-Tunable Nanofluidic Diode with a Broad Range of Rectifying Properties. *ACS Nano* **2009**, *3* (3), 603–608.

9. Yameen, B.; Ali, M.; Neumann, R.; Ensinger, W.; Knoll, W.; Azzaroni, O. Single Conical Nanopores Displaying PH-Tunable Rectifying Characteristics. Manipulating Ionic Transport with Zwitterionic Polymer Brushes. *J. Am. Chem. Soc.* **2009**, *131* (6), 2070–2071.
10. Chen, W. L.; Cordero, R.; Tran, H.; Ober, C. K. 50th Anniversary Perspective: Polymer Brushes: Novel Surfaces for Future Materials. *Macromolecules* **2017***, 50(11),* 4089–4113.
11. Li, M.; Pester, C. W. Mixed Polymer Brushes for "Smart" Surfaces. *Polymers.* **2020***, 12(7),* 1553.
12. Yang, W.; Zhou, F. Polymer Brushes for Antibiofouling and Lubrication. *Biosurf Biotribol* **2017**, *3* (3), 97–114.
13. Förster, S.; Wenz, E.; Lindner, P. Density Profile of Spherical Polymer Brushes. *Phys. Rev. Lett.* **1996**, *77*, 95.
14. Tran, Y.; Auroy, P.; Lee, L. T. Determination of the Structure of Polyelectrolyte Brushes. *Macromolecules* **1999**, *32* (26), 8952–8964.
15. Pesek, S. L.; Li, X.; Hammouda, B.; Hong, K.; Verduzco, R. Small-Angle Neutron Scattering Analysis of Bottlebrush Polymers Prepared via Grafting-through Polymerization. *Macromolecules* **2013**, *46* (17), 6998–7005.
16. Hore, M. J. A.; Ford, J.; Ohno, K.; Composto, R. J.; Hammouda, B. Direct Measurements of Polymer Brush Conformation Using Small-Angle Neutron Scattering (SANS) from Highly Grafted Iron Oxide Nanoparticles in Homopolymer Melts. *Macromolecules* **2013**, *46* (23), 9341–9348.
17. Hildebrandt, M.; Shin, E. young; Yang, S.; Ali, W.; Altinpinar, S.; Gutmann, J. S. Investigation of Roughness Correlation in Polymer Brushes via X-Ray Scattering. *Polymers (Basel)* **2020**, *12* (9).
18. Politsch, E.; Cevc, G.; Wurlitzer, A.; Lösche, M. Conformation of Polymer Brushes at Aqueous Surfaces Determined with X-Ray and Neutron Reflectometry. 1. Novel Data Evaluation Procedure for Polymers at Interfaces. *Macromolecules* **2001**, *34* (5), 1328–1333.
19. Kesal, D.; Christau, S.; Trapp, M.; Krause, P.; Von Klitzing, R. The Internal Structure of PMETAC Brush/Gold Nanoparticle Composites: A Neutron and X-Ray Reflectivity Study. *Phys. Chem. Chem. Phys.* **2017**, *19* (45), 30636–30646.

20. Delcea, M.; Helm, C. A. X-Ray and Neutron Reflectometry of Thin Films at Liquid Interfaces. *Langmuir* **2019**, *35* (26), 8519–8530.
21. Dronavajjala, K. D.; Rajagopalan, R.; Uppili, S.; Sen, A.; Allara, D. L.; Foley, H. C. A Simple Technique to Grow Polymer Brushes Using in Situ Surface Ligation of an Organometallic Initiator. *J. Am. Chem. Soc.* **2006**, *128* (40), 13040–13041.
22. Rahimi, M.; Nasiri, M. Polymer Brushes Prepared by Surface-Initiated Atom Transfer Radical Polymerization of Poly (N-Isopropyl Acrylamide) and Their Antifouling Properties. *Eur. Polym. J.* **2020**, *125*.
23. Zhulina, E. B.; Wolterink, J. K.; Borisov, O. V. Screening Effects in a Polyelectrolyte Brush: Self-Consistent-Field Theory. *Macromolecules* **2000**, *33* (13), 4945–4953.
24. Borisov, O. V.; Zhulina, E. B. Conformations of Polyelectrolyte Molecular Brushes: A Mean-Field Theory. *J. Chem. Phys.* **2018**, *149*, 184904.
25. Chen, L.; Merlitz, H.; He, S.; Wu, C. X.; Sommer, J. W. Polyelectrolyte Brushes: Debye Approximation and Mean-Field Theory. *Macromolecules* **2011**, *44*, 3109–3116.
26. Sachar, H. S.; Sivasankar, V. S.; Das, S. Revisiting the Strong Stretching Theory for Ph-Responsive Polyelectrolyte Brushes: Effects of Consideration of Excluded Volume Interactions and An Expanded Form of The Mass Action Law. *Soft Matter* **2019**, *15*, 559-574.
27. Okrugin, B. M.; Richter, R. P.; Leermakers, F. A. M.; Neelov, I. M.; Borisov O. V.; Zhulina, E. B. Structure and Properties of Polydisperse Polyelectrolyte Brushes Studied by Self-Consistent Field Theory. *Soft Matter* **2018**, *14*, 6230-6242.
28. Russano, D.; Carrillo, J. M. Y.; Dobrynin, A. V. Interaction between Brush Layers of Bottle-Brush Polyelectrolytes: Molecular Dynamics Simulations. *Langmuir* **2011**, *27* (17), 11044–11051.
29. Sachar, H. S.; Chava, B. S.; Pial, T. H.; Das, S. All-Atom Molecular Dynamics Simulations of the Temperature Response of Densely Grafted Polyelectrolyte Brushes. *Macromolecules* **2021**, *54*, 6342–6354.

30. Galuschko, A.; Sommer, J. U. Co-Nonsolvency Response of a Polymer Brush: A Molecular Dynamics Study. *Macromolecules* **2019**, *52* (11), 4120–4130.

31. Yagasaki, T.; Matubayasi, N. Molecular Dynamics Study of the Antifouling Mechanism of Hydrophilic Polymer Brushes. *Langmuir* **2023**, *39* (37), 13158–13168.

32. 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 J. Comput. Mol. Sci.* **2019**, *1* (1).

33. Gasparotto, P.; Ceriotti, M. Recognizing Molecular Patterns by Machine Learning: An Agnostic Structural Definition of the Hydrogen Bond. *J. Chem. Phys.* **2014**, *141* (17).

34. Offei-Danso, A.; Hassanali, A.; Rodriguez, A. High-Dimensional Fluctuations in Liquid Water: Combining Chemical Intuition with Unsupervised Learning. *J. Chem. Theory Comput.* ***2022****, 18(5),* 3136–3150.

35. Ansari, N.; Onat, B.; Sosso, G. C.; Hassanali, A. Insights into the Emerging Networks of Voids in Simulated Supercooled Water. *J. Phys. Chem. B.* **2020**, *124* (11), 2180–2190.

36. Pial, T. H.; Das, S. Machine Learning Enabled Quantification of the Hydrogen Bonds Inside the Polyelectrolyte Brush Layer Probed Using All-Atom Molecular Dynamics Simulations. *Soft Matter* **2022**, *18*, 8945−8951.

37. Ishraaq, R.; Akash, T. S.; Bera, A.; Das, S. Hydrophilic and Apolar Hydration in Densely Grafted Cationic Brushes and Counterions with Large Mobilities. https://doi.org/10.26434/chemrxiv-2023-b8mjm.

38. Wu, B.; Wang, X.; Yang, J.; Hua, Z.; Tian, K.; Kou, R.; Zhang, J.; Ye, S.; Luo, Y.; Craig, V. S. J.; Zhang, G.; Liu, G. Reorganization of Hydrogen Bond Network Makes Strong Polyelectrolyte Brushes Ph-Responsive. *Sci. Adv.* **2016**, *2*, e1600579.

39. Chu, X.; Yang, J.; Liu, G.; Zhao, J. Swelling Enhancement of Polyelectrolyte Brushes Induced By External Ions. *Soft Matter* **2014**, *10*, 5568-5578.

40. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

41. Joung, I. S.; Cheatham, T. E. III. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.

42. Sachar, H. S.; Pial, T. H.; Desai, P. R.; Etha, S. A.; Wang, Y.; Chung, P. W.; Das, S. Densely Grafted Polyelectrolyte Brushes Trigger "Water-in-Salt"-like Scenarios and Ultraconfinement Effect. *Matter* **2020**, *2*, 1509–1521.

43. 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. *Comput. Phys. Commun.* **2022**, *271*, 108171.

44. Stukowski, A. Visualization and Analysis of Atomistic Simulation Data with OVITO-the Open Visualization Tool. *Model Simul. Mat. Sci. Eng.* **2010**, *18* (1).

45. Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *SOFTWARE Open Access Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform*. *J. Cheminf.* **2012**, *4*, 17.

46. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Cournapeau, D.; Passos, A.; Brucher, M.; Perrot, M. and Duchesnay, E. Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**,*12*, 2825-2830.

47. Callen, H. B. *Thermodynamics and an introduction to thermostatistics*; Wiley (New York), 1985.

48. Truong, C.; Oudre, L.; Vayatis, N. Selective Review of Offline Change Point Detection Methods. *Signal Processing* **2020**, *167*, 107299.

49. Ceriotti, M. Unsupervised Machine Learning in Atomistic Simulations, between Predictions and Understanding. *Journal of Chemical Physics*. *J. Chem. Phys.* **2019**, *150*, 150901.

50. Comaniciu, D. and Meer, P. "Mean shift analysis and applications," *Proceedings of the Seventh IEEE International Conference on Computer Vision*, Kerkyra, Greece, **1999**, *2*, 1197-1203.

51. Adams, S.; Beling, P. A. A Survey of Feature Selection Methods for Gaussian Mixture Models and Hidden Markov Models. *Artif. Intel.l Rev.* **2019**, *52* (3), 1739–1779.

52. Duboué-Dijon, E.; Laage, D. Characterization of the Local Structure in Liquid Water by Various Order Parameters. *J. Phys. Chem. B* **2015**, *119* (26), 8406–8418.

53. Madan, B.; Sharp, K. Changes in Water Structure Induced by a Hydrophobic Solute Probed by Simulation of the Water Hydrogen Bond Angle and Radial Distribution Functions. Biophys. Chem. **1999**, *78(1-2)*, 33-41.

54. Zhang, C.; Chu, X.; Zheng, Z.; Jia, P.; Zhao, J. Diffusion of Ionic Fluorescent Probes atop Polyelectrolyte Brushes. *J. Phys. Chem. B* **2011**, *115* (51), 15167–15173.