PCCP

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2025, **27**, 22901

Received 2nd July 2025,
Accepted 8th October 2025

DOI: 10.1039/d5cp02528a

rsc.li/pccp

Combined machine learning and atomistic simulations reveal multi-state hydration in cationic brushes in the presence of halide counterions

Raashiq Ishraaq and Siddhartha Das *

In this communication, we employ a combination of all-atom molecular dynamics simulations and machine learning to establish the effect of different halide screening counterions (fluoride, chloride, bromide, and iodide ions) on the prevalence of two separate hydration states of the $\{N(CH_3)_3\}^+$ functional group of the PMETA (([poly(2-(methacryloyloxy)ethyl) trimethylammonium) cationic brushes.

Polyelectrolyte (PE) brushes,[1] which refer to charged PE chains in "brush"-like configurations, have been extensively used for functionalizing solid surfaces for applications such as protein adsorption,[2] sensing,[3] current rectification,[4] drug delivery,[5] water harvesting,[6] oil recovery,[7] *etc.* The overall behavior of the PE brushes and the brush-supported counterions and water molecules depends on the chemical nature of the species (PE and the counterion), chain architecture, configuration and external factors, and understanding this difference is crucial for designing novel brushes with desired properties. While there have been extensive investigations on PE brush systems *via* the use of coarse-grained molecular dynamics (MD) simulations, theory, and experiments (please see review papers in ref. 1 and 8 for some of these papers), the present group introduced all-atom MD simulations to capture the brush behaviour, where each brush atom and brush-supported water molecules and counterions are accounted for.[9–16] For example, in a recent paper, we studied the manner in which the different halide ions bind to the PMETAX (X = F^-,Cl^-,Br^-,I^-) PE brushes – it was revealed that due to their chaotropic nature, I^- and Br^- ions disrupted the water structure the most (enforcing the water molecules with least order parameter values around these ions) and accordingly, bound to the apolar $\{N(CH_3)_3\}^+$ functional group of the PMETAX brushes most strongly.[16] Additionally, in several recent papers, we employed unsupervised machine learning (ML) approaches (*e.g.*, clustering methods) to the all-atom MD simulation derived data to unravel several novel facets of the PE brushes and the brush-functionalized water and counterions.[17–20] For example, in one of these papers, we employed a ML framework (a combination of the Gaussian preprocessing and unsupervised clustering algorithm) on all-atom MD simulation data to establish the presence of two separate hydration states for the $\{N(CH_3)_3\}^+$ functional group of Cl^--ion-screened PMETAX brushes.[19] One of these hydration states was associated with the apolar hydration layer (with more structured water) formed around the $\{N(CH_3)_3\}^+$ group,[15] while the other hydration state was associated with the (unstructured) water molecules solvating the Cl^- ion that is attracted to the $\{N(CH_3)_3\}^+$ group.[16]

In this communication, we employ the same ML framework (see ref. 19) on all-atom MD data (MD simulations have been performed using LAMMPS[19]) to identify that for the PMETAX brushes screened with different halide counterions (F^-, Cl^-, Br^-, and I^- ions) there are always two hydration states of the $\{N(CH_3)_3\}^+$ group, as revealed by two separate clusters in the q–θ–CN tupule of water molecules. Here "q" and "θ" respectively denote the tetrahedral order parameter and dipole moment distribution water molecules around the $\{N(CH_3)_3\}^+$ group, while "CN" denotes the number of water molecules coordinated to the $\{N(CH_3)_3\}^+$ group. The "q" parameter depends on how much or how less the water structure is deviated from its ideal tetrahedral structure (for the case of the ideal tetrahedral structure, the value of the q parameter is 1). It depends on various factors such as the ion's size, its charge density, its polarizability, ionic structure, *etc.* The clusters defining the two hydration states differ in q-parameter values: the difference is largest (smallest) for the iodide-ion-screened (fluoride-ion-screened) PMETA brushes. Also, the relative size of the cluster representing the contribution of the solvation water (with reduced q) from the halide ion increases in the order $I^- > Br^- > Cl^- > F^-$. Both these results can be explained by noting that among the halide ions, water is the least (most) structured around the I^- (F^-) ion and I^- (F^-) binds most strongly (weakly) to the $\{N(CH_3)_3\}^+$ group. In summary, this communication represents an important discovery in the chemical sciences of

Department of Mechanical Engineering, University of Maryland, College Park, Maryland 20742, USA. E-mail: sidd@umd.edu

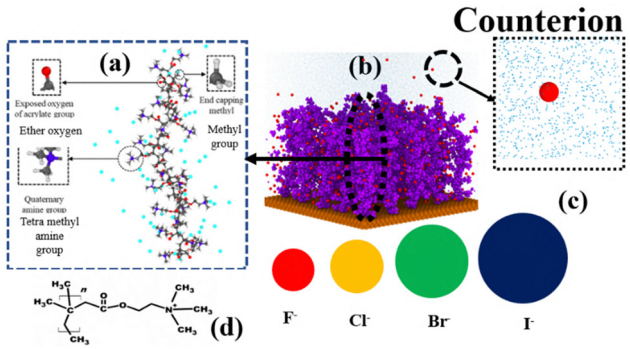

Fig. 1 MD simulation system. (b) MD simulation snapshot for the PMETA brushes with the chemical formula of the PMETA molecule shown in (d) and the detailed architecture of the different functional group of the brush shown in (a). The counterions in (b) are shown by red spheres with the counterion size variation (for different halide ion) being shown in (c).

the PE brushes *via* the combined use of all-atom MD simulations and unsupervised ML.

We first consider the all-atom MD simulations of the cationic PMETAX brushes for the separate cases where the brushes are screened by different halide counterions (*i.e.*, X= F^-, Cl^-, Br^-, I^-). Fig. 1 provides a representative MD simulation screenshot. These MD simulations explicitly account for the different atoms of the PE brushes, the counterions, and the water molecules. We have conducted such simulations in our previous paper;[16] we discuss the relevant details (including the calculations for "q", θ, and CN) in the SI. Additional information on the simulation procedure as well as methods to calculate different properties (*e.g.*, parameters like q, θ, CN) can be found in our previous papers.[16] These all-atom MD simulations provide different parameters associated with the PE brushes and the brush-supported water molecules and counterions. Given that we are interested to use ML for studying the microscopic hydration states of the $\{N(CH_3)_3\}^+$ group of the PMETAX brushes, following our previous study,[19] we shall first consider the three state variables (q, θ, CN) of the the water molecules around the $\{N(CH_3)_3\}^+$ group. This ML scheme comprises of two steps (see ref. 19 for details). First, we employ Gaussian pre-processing of the data (q, θ, CN of water molecules) and obtain the normalized distributions of these quantities in the q–θ–CN tupule space. Subsequently, we employ the unsupervised clustering to the q–θ–CN distributions to yield information about the hydration states of the $\{N(CH_3)_3\}^+$ group in the presence of different halide counterions.

We first discuss the Gaussian pre-processing of the data. Fig. 2(a, f and k) provide the all-atom MD simulation results on the probability distribution functions (PDFs) for the q, θ, and CN respectively associated with the water molecules around the $\{N(CH_3)_3\}^+$ group of the iodide-ion-screened PMETA brushes. These results represent the raw data for the PDFs, and it is important to identify the number of physical factors that govern the PDFs of q, θ and CN. For that purpose, one first needs to smoothen these PDFs by fitting several Gaussian functions to each PDF. In order to identify the number of Gaussians needed to smoothen the PDFs of q, θ and CN, we perform a goodness of fit calculations by quantifying the Akaike information criterion (AIC) and Bayesian information criterion (BIC) as functions of the number of Gaussians used individually for fitting q, θ and CN [see Fig. 2(b, g and l)]. These required number of Gaussians will be those for which AIC and BIC values got minimized. Fig. 2(c, h and m) provide these

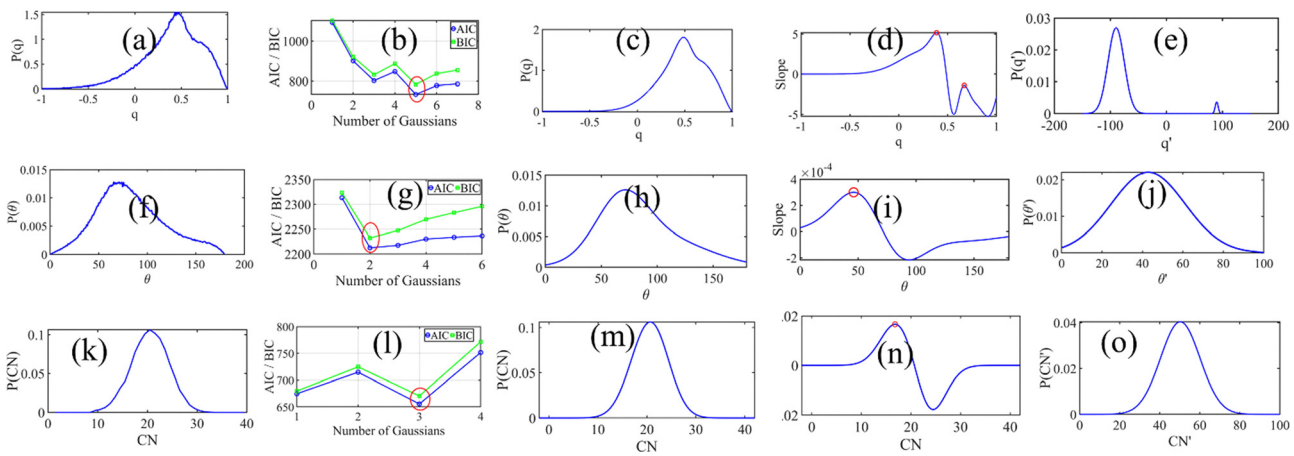

Fig. 2 Gaussian pre-processing of the all-atom MD data of the water properties around the $\{N(CH_3)_3\}^+$ group of the iodide ion screened PMETA brush layer (grafting density of σ_g = 0.20 chains nm^{-2}). (a), (f) and (k): Probability distribution functions (PDFs) corresponding to (a) tetrahedral order parameter (q) and (f) dipole orientation angle (θ) (expressed in degrees), and (k) co-ordination number (CN). (b), (g) and (l): Variation of AIC and BIC with the number of Gaussians used to fit (and smoothen) the PDFs corresponding to (b) q, (g) θ, and (l) CN. For each case, the number of Gaussians where AIC and BIC are minimized have been identified by circles. (c), (h) and (m): Smoothened PDFs (ensured by fitting with the identified number of Gaussians) corresponding to (c) q, (h) θ, and (m) CN. (d), (i) and (n): Variation of the slopes (whose peaks have been identified) of the PDFs corresponding to (d) q, (i) θ, and (n) CN. The number of peaks are equal to the number of Gaussians that will be used to fit the distributions shown in (c), (h) and (m). (e), (j) and (o): These Gaussians correspond to (e) q, (j) θ, and (o) CN; hence there are 2, 1, and 1 Gaussians used to fit q, θ, and CN PDFs, respectively, with the 2 Gaussians for the case of the q PDF being separated by a linear transformation. Also, these Gaussians represent scaled variables q', θ', and CN′, as these Gaussians are scaled to a different range and normalized.

smoothened PDFs (obtained by smoothening with the identified number of Gaussians; see the caption of Fig. 2) for q, θ, and CN. Once we have obtained these smoothened PDFs, we are in a position to identify the number of physical factors governing each PDF. As was explained in our previous paper,[19] the presence of multiple (single) driving factor can be manifested as the presence of multiple peaks or bulges (single peak) in the corresponding smoothened PDFs. The presence of such peak (or bulge) in the PDFs get manifested by a sudden change in the corresponding statistical properties (*e.g.*, peaks in the slope of the PDF). Fig. 2(d, i and n) demonstrate these peaks in the slope of the PDFs of the q, θ, and CN: one can clearly see two peaks in the slope of the q PDF, while one peak in the slopes of the θ and CN PDFs. These signify that there are two factors governing the q PDF, while 1 factor dictating each of θ and CN PDFs. This also implies that one can fit two Gaussians to the (smoothened) q PDF and one Gaussian to each of the (smoothened) θ and CN PDF. In Fig. 2(e, j and o), we plot these Gaussians (which represent these PDFs) in a manner such that they are linearly scaled and normalized (and hence denoted as q', θ', and CN′). Also, the 2 Gaussians used for fitting the q PDF are separated by a linear transformation. This procedure illustrated in Fig. 2 completes the Gaussian pre-processing of the all-atom-MD derived data on q, θ and CN of water molecules around the $\{N(CH_3)_3\}^+$ group of the iodide-ion-screened PMETA brushes. The results from the same Gaussian pre-processing of the data on q, θ and CN of water around $\{N(CH_3)_3\}^+$ group of the bromide-ion, chloride-ion, and fluoride-ion screened PMETA brushes have been provided in Fig. S3–S5 in the SI. As can be seen from Fig. S3–S5, 2, 1, and 1 Gaussians used to fit q, θ, and CN PDFs, respectively (with the 2 Gaussians for the case of the q PDF being separated by a linear transformation) for each of the cases of fluoride-ion, chloride-ion, and bromide-ion screened PMETA brushes.

The data points corresponding to the sampled data (represented by q', θ', and CN′ distributions) are further normalized (data-mean/standard deviation) and are denoted as q'_n, θ'_n and CN'_n. Fig. 3(a–d) provide the q'_n, θ'_n and CN'_n tupule distribution for the water molecules around the $\{N(CH_3)_3\}^+$ group of PMETAX (X = I^-, Br^-, Cl^-, and F^-) brushes. At this stage, we employ the ML algorithm to the q'_n, θ'_n and CN'_n tupule for each of the four cases. The ML method consists of first applying the mean shift algorithm that identifies the number of modes and their positions in the data set. Subsequently, Gaussian mixture model clustering was employed and the data points were separated. This process ensured that the number of modes manifested as the number of clusters (in the $q'_n - \theta'_n - CN'_n$ tupule space), which in turn represent the number of hydration states of the $\{N(CH_3)_3\}^+$ group. We can clearly see that for all the four cases, $\{N(CH_3)_3\}^+$ group demonstrates two separate hydration states varying significantly in the q parameter (or the degree of structuredness of water molecules) [see Fig. 4(a–d)]. The $\{N(CH_3)_3\}^+$ group itself, despite being positively charged, triggers an apolar hydration layer characterized by the presence of more structured water molecules (or water molecules having a greater q).[15] On the other hand, all the halide ions, being strong anions, trigger strongly-bound hydration layer that are characterized by the presence of less structured (lesser q) water molecules.[16] These halide ions get electrostatically attracted to the $\{N(CH_3)_3\}^+$ group and share their solvation water with the solvation water of the $\{N(CH_3)_3\}^+$ group, eventually resulting in generating these two separate hydration states of $\{N(CH_3)_3\}^+$ group [the state/cluster with smaller q is the contribution of the halide ion, while the state/cluster with greater q is contributed by the $\{N(CH_3)_3\}^+$ group itself].

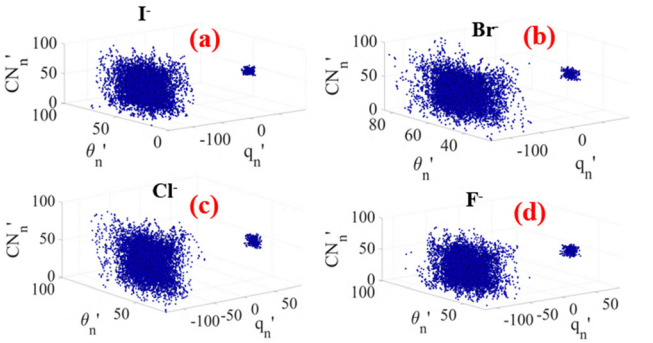

Fig. 3 Sample normalized data points represented as the $q'_n - \theta'_n - CN'_n$ tupule corresponding to the water molecules around the $\{N(CH_3)_3\}^+$ group of the PMETA brushes screened by for (a) iodide, (b) bromide, (c) chloride, and (d) fluoride ions.

Next we shall like to understand how the variation in the nature of the halide ion affects the two hydration states of the $\{N(CH_3)_3\}^+$ group. We characterize the prevalence of each of these hydration states by the number of points that constitute the cluster representing each of the hydration states. For all the four cases, the number of points corresponding to the hydration state with more structured water [or the state contributed directly by the $\{N(CH_3)_3\}^+$ group] is much lesser than the number of points corresponding to the hydration state with less structured water (or the state contributed by the halide ions).

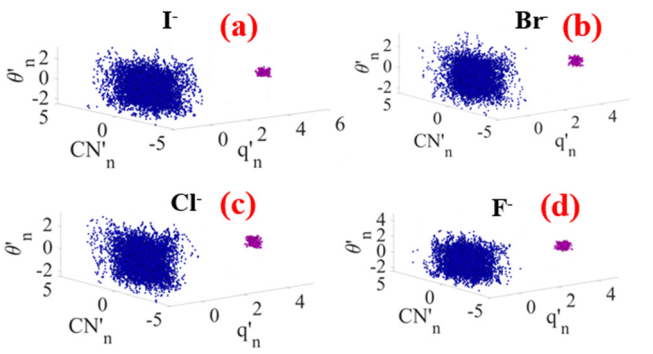

Fig. 4 Data clusters obtained by employing the ML method (consisting of the mean-shift algorithm and unsupervised clustering) to the normalized $q'_n - \theta'_n - CN'_n$ tupules shown in Fig. 3(a–d). The results are for the PMETA brushes screened by for (a) iodide, (b) bromide, (c) chloride, and (d) fluoride ions. For each case, there are two clusters representing two hydration states of the $\{N(CH_3)_3\}^+$ group: the one shown with blue (purple) points represent the hydration state with less (more) structured water, *i.e.*, water molecules with lesser (greater) q.

We can identify the ratio of points associated with these two states (state with lesser q to state with greater q) as r. Fig. 5(a) plots r values associated with the hydrations states of the $\{N(CH_3)_3\}^+$ group for four different cases of PMETAX brushes (X = I^-, Br^-, Cl^-, and F). We find that r progressively decreases as one goes from the cases corresponding to I^- to F^- ions. Fig. 5(b) provides the average distance between the two clusters (representing the two hydration states) for each of the four cases. Here too we find that the average cluster-to-cluster distances progressively decreases as one goes from I^- to F^- ions. In our previous paper, we have identified that due to the chaotropic nature of the I^- and Br^- ions and the kosmotropic nature of the Cl^- and F^- ions, the structure of the water molecules is severely disrupted around the I^- and Br^- ions, while the water is more structured around the Cl^- and F^- ions.[16] This ensured that q parameter distribution has greater values for water molecules around the Cl^- and F^- ions, as compared to that around the I^- and Br^- ions. Also, for the same reason, the $\{N(CH_3)_3\}^+$ group of the PMETA brushes interact much more strongly with the I^- and Br^- ions (as evident from the corresponding radial distribution function or RDF results; see ref. 16 for details), as compared to that with the Cl^- and F^- ions. This stronger interaction will imply that the contribution of the solvation-shell water of the I^- and Br^- ions in determining the (low-q) hydration state of the $\{N(CH_3)_3\}^+$ group will be much greater, and accordingly, the ratio r will be greater for the iodide and bromide ion screened PMETA brushes, as compared to the chloride and fluoride ion screened PMETA brushes [see Fig. 5(a)]. Also, the fact that the q values associated with the water molecules in the hydration shell of the iodide and bromide ions are smaller than that in the hydration shell of the chloride and fluoride ions, we can justify why the distance between the two clusters around the PMETA brush's $\{N(CH_3)_3\}^+$ group (one associated with high-q water molecules and another associated with low-q water molecules) increase as we go from fluoride-ion-screened PMETA to iodide-ion-screened PMETA brushes.

It is important to discuss some essential facets of this work. the presence of the two hydration states depends on the nature of the $\{N(CH_3)_3\}^+$ group, the nature of the confinement, and the type (namely, size and valence) of the counterion. Therefore, it is expected that similar behavior (where such two-state hydration is observed) can occur for other polyelectrolyte (PE) brush systems, containing quaternary functional groups (or other types of functional groups) that induce water structuring different from that induced by the corresponding screening counterions. Finally, it is worthwhile to mention that it might be possible to utilize some of the recently developed analysis techniques for experimentally observing these two separate hydration states. For example, Inbaraj *et al.* observed two separate states of water molecules (one strongly trapped and another less perturbed) in ionic liquids by utilizing Fourier Transform Infrared Absorption Spectroscopy.[21] Such methods might be useful for the present case as long as it is possible to access the narrow space inside the densely grafted brush layer.

To summarize, in this communication, we have combined atomistic MD simulations with ML methods to address a highly important question with regards to the counterion-dependent chemistry of the PE brushes: how the variation of the screening halide ions affects the hydration of a charged functional of cationic brushes? Considering the specific example of the PMETA brushes having the $\{N(CH_3)_3\}^+$ functional group, we show that for all the halide ions, the functional group demonstrates two separate hydrations states varying significantly in their relative abundance and the structural properties (tetrahedral order parameter or q) of the water molecules constituting these hydration states. The chaotropic nature of the iodide and bromide ions ensure that when they screen the PMETA brushes, there is a more dominant presence of the low-q hydration state for the water molecules around the $\{N(CH_3)_3\}^+$ group and the distance between the clusters representing the low-q and high-q hydration states are also enhanced.

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

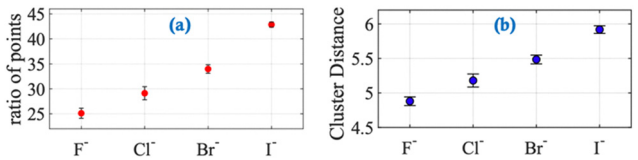

Fig. 5 (a) Variation of the ratio r, which is the ratio of the points constituting the cluster corresponding low-q water hydration state of $\{N(CH_3)_3\}^+$ to the points constituting the cluster corresponding high-q water hydration state of $\{N(CH_3)_3\}^+$, for the PMETA brushes screened with the different halide counterions. (b) Variation of the average distance between the clusters corresponding to low-q water and high-q water hydration states for the $\{N(CH_3)_3\}^+$ group of the PMETA brushes screened with different halide counterions. We ran four independent simulations for each counterion case and added error bars to the data points in both (a) and (b).

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included in the main paper as well as in the supplementary information (SI). Supplementary information is available. See DOI: **https://doi.org/10.1039/d5cp02528a**.

Acknowledgements

This work was supported by the Department of Energy Office of Science (Grant DE-SC0017741).

References

1 S. Das, M. Banik, G. Chen, S. Sinha and R. Mukherjee, *Soft Matter*, 2015, **11**, 8550.
2 E. Jumai'an, L. Zhang and M. A. Bevan, *ACS Nano*, 2023, **17**, 2378.

3 G. Xie, W. Tian, L. Wen, K. Xiao, Z. Zhang, Q. Liu, G. Hou, P. Li, Y. Tian and L. Jiang, *Chem. Commun.*, 2015, **51**, 3135.
4 B. Yameen, M. Ali, R. Neumann, W. Ensinger, W. Knoll and O. Azzaroni, *J. Am. Chem. Soc.*, 2009, **131**, 2070.
5 P. Matricardi, C. D. Meo, T. Coviello and F. Alhaique, *Expert Opin. Drug Delivery*, 2008, **5**, 417.
6 G. Liu, M. Cai, X. Wang, F. Zhou and W. Liu, *ACS Appl. Polym. Mater.*, 2014, **6**, 11625.
7 L. Qi, C. Song, T. Wang, Q. Li, G. J. Hirasaki and R. Verduzco, *Langmuir*, 2018, **34**, 6522.
8 R. Ishraaq and S. Das, *Chem. Commun.*, 2024, **60**, 6093.
9 H. S. Sachar, T. H. Pial, P. R. Desai, S. A. Etha, Y. Wang, P. W. Chung and S. Das, *Matter*, 2020, **2**, 1509.
10 H. S. Sachar, B. S. Chava, T. H. Pial and S. Das, *Macromolecules*, 2021, **54**, 2011.
11 T. H. Pial, H. S. Sachar, P. R. Desai and S. Das, *ACS Nano*, 2021, **15**, 6507.
12 T. H. Pial, H. S. Sachar and S. Das, *Macromolecules*, 2021, **54**, 4154.
13 H. S. Sachar, T. H. Pial, V. S. Sivasankar and S. Das, *ACS Nano*, 2021, **15**, 17337.
14 T. H. Pial, M. Prajapati, B. S. Chava, H. S. Sachar and S. Das, *Macromolecules*, 2022, **55**, 2413.
15 R. Ishraaq, T. S. Akash, A. Bera and S. Das, *J. Phys. Chem. B*, 2024, **128**, 381.
16 R. Ishraaq and S. Das, *Macromolecules*, 2024, **57**, 3037.
17 T. H. Pial and S. Das, *Soft Matter*, 2022, **18**, 8945.
18 A. Bera, T. S. Akash, R. Ishraaq, T. H. Pial and S. Das, *Macromolecules*, 2024, **57**, 1581.
19 R. Ishraaq, T. S. Akash and S. Das, *Macromolecules*, 2024, **57**, 5300.
20 R. Ishraaq and S. Das, *J. Phys. Chem. B*, 2025, **129**, 5854.
21 N. R. Inbaraj, S. Song, R. Chang, K. Fujita and T. Hayashi, *Langmuir*, 2023, **39**, 2558.