RSC Advances

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2025, **15**, 39587

From volatile organic compounds to aerosols: quantum-level insights into humidity-induced cluster formation in atmospheres

Arnab Patla and Ranga Subramanian *

New particle formation (NPF) is a key factor in producing the aerosol populations in the atmosphere. This has enormous effects on controlling the climate, improving air quality, and protecting human health. Volatile organic compounds (VOCs) are a significant cause of NPF because they help prenucleation clusters form, which are the building blocks of aerosol particles. We employed density functional theory (DFT) to investigate the structure and thermodynamics of molecular clusters composed of acetic acid, acetone, acetaldehyde, and formaldehyde in the presence of ammonia and varying amounts of water molecules ($n = 0$–30). Our study examines binding energies, enthalpies, and free energies that are BSSE-corrected across a range of atmospheric temperatures (T) and pressures (P), reflecting the changes in the real-world environment. We also look at Rayleigh scattering levels ($\Re_n$), which shows a steady rise with cluster hydration. The results show that different VOC systems interact and cluster in distinct ways, providing more information about their roles in the atmosphere. This study may offer fresh insights into how different VOCs function in specific scenarios, which is necessary to comprehend new particle formation events in polluted areas.

Received 10th September 2025
Accepted 10th October 2025

DOI: 10.1039/d5ra06836c

rsc.li/rsc-advances

1. Introduction

An aerosol consists of liquid or solid particles dispersed in a gas, with sizes ranging from 1 nm to 0.1 mm. The lower threshold produces molecules and molecular aggregates, while the higher threshold leads to rapid sedimentation.[1,2] These aerosol droplets significantly influence atmospheric chemical processes, owing to their unique surface properties. Aerosols, mostly condensed water with a particle diameter of 10 μm, are prominently observed in the atmosphere and function as cloud condensation nuclei. Organic compounds (OCs) are a significant component of atmospheric aerosols.[3] In the atmosphere, organic aerosols may be primary particles, such as those from biomass and fossil fuel combustion, and secondary particles, such as those from volatile organic compound (VOC) oxidation. Secondary organic aerosols (SOA) are formed when oxidized VOCs condense onto pre-existing aerosols. This process occurs when VOCs are oxidized in the atmosphere.[3] One of the most prevalent oxygenated volatile organic compounds (OVOCs) found in the atmosphere is acetone (AC), with mixing ratios varying from parts per trillion (ppt) in remote areas to parts per billion (ppb) in polluted areas. Acetone is directly released into the atmosphere from various sources, such as natural and anthropogenic sources, biomass burning, vehicular exhaust, and industrial and combustion processes.[4–6] Acetaldehyde (AD) and formaldehyde (FD) are key components of atmospheric photochemical reactions. The majority of atmospheric acids contributing to the acidity of clouds and raindrops in remote areas are formic (HC(O)OH), abbreviated as FA, and acetic acid ($CH_3C(O)OH$), abbreviated as AA.[7,8]

Water is the primary condensed-phase element of the atmosphere. Various types of atmospheric liquid water, including rain, fog, cloud droplets, and water, are bonded to the aerosol particles.[9] The study on water-containing clusters has progressed using different experimental and computational methods.[10–13] Ammonia is essential in aerosol production since it prevents the evaporation of nucleating acids and rapidly interacts with acidic compounds to generate ammonium salts, hence reducing the atmospheric particle free energy *via* the formation of ion pairs.[14] The primary sources of NH_3 (A) in the atmosphere are emissions, NH_3-based fertilisers, and animal waste. The average amount of ammonia in the continental lower atmosphere ranges from 0.1 to 10 parts per billion (ppb).[15,16] Indeed, ammonia is a crucial precursor for atmospheric nucleation.[17]

AC, AD, FD, AA, W, and A can act as hydrogen bond donors and acceptors, respectively. VOCs use hydrogen bonds to form clusters with water and ammonia molecules. VOC-containing hydrogen-bonded clusters are of interest in this study. Numerous theoretical and experimental investigations have focused on VOC-water clusters in recent years.[17–22] Zhang *et al.* determined the interaction between water and acetic acid-sodium halide aerosol using molecular dynamics (MD)

Department of Chemistry, Indian Institute of Technology Patna, 801103, India. E-mail: ranga@iitp.ac.in

simulations.[23] Using DFT, Romero-Montalvo analysed the hydrogen bond interactions in water cluster-organic molecule complexes.[21] Wang *et al.* examined the theoretical aspects of gas-phase hydrolysis of formaldehyde to produce methanediol and its implications for forming new particles. They conclude that aldehydes were important precursor species in the NPF.[22] This study examines the atmospheric consequences of hydration during the production of organic acid–water, organic aldehyde–water, and organic ketone–water clusters in the gas phase.

In the present work, we address the following questions: (i) How do organic acids, aldehydes, and ketones form new particles with water in the atmosphere? (ii) What are the ambient atmospheric conditions for producing organic acids, aldehydes, and ketone-water clusters? (iii) How do hydrated organic acids, aldehydes, and ketone clusters scatter the incoming sunlight? (iv) What are the atmospheric implications of the hydrated organic acids, aldehydes, and ketone clusters? To address these questions, we first determined the lowest energy structures of VOC(A)(W)$_n$, (VOC = AA, AC, AD, FD) (n = 0–10, 15, 20, 25, and 30). We conducted DFT to compute the thermodynamic parameters of the examined clusters throughout a broad spectrum of atmospherically relevant temperatures (T) and pressures (P) (ranging from T and P = 298.15 K and 1013.25 atm to T and P = 226.65 K and 226.32 atm). Topological parameter analysis employs atoms in molecules (AIM) theory to enhance understanding of bonding characteristics. The scattering qualities were significantly influenced by the examined clusters at different incoming light wavelengths (static, 400, 500, 600, and 700 nm).

2. Computational details

2.1. Cluster sampling

The initial geometries of every combinational geometry VOC(A)(W)$_n$, (VOC = AA, AC, AD, FD) (n = 0–10, 15, 20, 25, 30) were sampled using the molecular mechanics (MM)[24] force field (UFF),[25] applying the steepest descent algorithm in the Avogadro package.[24] This procedure ensures that the clusters remain approximately spherical, mimicking the morphology of a forming droplet. Water molecules were arranged in various orientations around AA, AC, AD, and FD to construct these spherical assemblies. A sequential process to create a larger cluster by step-wise addition of a water molecule. Using the result of MM, each conformation of VOC(A)(W)$_n$, (VOC = AA, AC, AD, FD) (n = 0–10, 15, 20, 25, 30) clusters are further sampled with the Born Oppenheimer molecular dynamics (BOMD) method with simulated annealing at the M06-2X/6-31+G(d,p) level of theory. The molecular dynamics simulation involved heating each water molecule by approximately 5 kcal mol^{-1} and evenly removing it over a period of up to 5 ps at 300 K. In all instances, energy was completely dissipated within 0.5–1.0 ps and equilibrated until the end of the run. We employed the BOMD method, as implemented in the Gaussian 16 package[26] for thermal sampling. This method uses a fifth-order polynomial fitted to the energy, gradient, and Hessian at each time step. This method enables migration from one minimum to another, which is necessary when dealing with many water molecules in clusters due to a shallow potential energy surface. After the completion of the simulated annealing run, it gives several structures. We selected only one structure with the lowest potential energy from several candidates and optimised it using the DFT/6-311++G(2d,2p) theoretical level (DFT = M06-2X, ωB97X-D, and PW91PW91) to yield the final structure. We selected DFT with the functionals M06-2X, ωB97X-D, and PW91PW91 for our work due to their effective application to atmospheric molecular clusters in various prior investigations.[10,27–29] We have also employed this sampling procedure in our previous work.[16,30–33] The BOMD technique enables the identification of a potentially optimal estimate for the global minimum corresponding to each cluster.[16,30,32–34]

2.2. Structural and topological parameters

AIMALL software[35] was used to conduct Atoms in Molecules (AIM) analysis at the M06-2X/6-311++G(2d,2p) level to gain a better understanding of the nature of bonding present in VOC(A)(W)$_n$, (VOC = AA, AC, AD, FD) (n = 0–10, 15, 20, 25, 30) clusters. The M06-2X level is one of the best non-local exchange-correlation functionals assisted by AIMALL.[35,36] Topological variables such as the electron density (ρ), Laplacian ($\nabla^2\rho$), and total energy density (M) were employed to characterize bond critical points (BCP). The energetic descriptors at BCPs include M and its components, electronic kinetic energy density (N) and potential energy density (V), which are related by $M = V + N$ Intermolecular interactions were further examined using the non-covalent interaction (NCI) approach based on the reduced density gradient (RDG). The NCI index is derived from the reduced density gradient (s), which depends on electron density (ρ) and its gradient ($\nabla\rho$)

$$s = \frac{1}{2(3\pi^2)^{1/3}} \frac{|\nabla\rho|}{\rho^{4/3}} \quad (1)$$

The blue, green, and red codes in the NCI isosurfaces represent stabilisation of hydrogen bonding, weak van der Waals interactions, and destabilising steric interactions, respectively. NCI-RDG analysis was performed using high-quality grids with the MultiWfn application.[37] The Visual Molecular Dynamics (VMD)[38] application visualizes a gradient isosurface in real space.

2.3. Thermodynamics properties

We calculated the thermodynamics of the chosen isomers of all cluster systems. Binding energy, free energy, and enthalpy were calculated using the lowest energy-optimized structure from DFT/6-311++G(2d,2p) (DFT = M06-2X, ωB97X-D, and PW91PW91). We used the counterpoise (CP) correction method[39] to estimate BSSE corrected binding energies ($\Delta E_{\mathrm{bie}}^{\mathrm{cp}}$) and successive binding energy (ΔE_n^{cp}). The Gaussian 16 software[26] was used for all computations. We evaluate enthalpy (ΔH), free energy (ΔG), and successive enthalpy (ΔH_n) and successive free energy (ΔG_n) of studied clusters using without BSSE correction at M06-2X/6-311++G(2d,2p) level of theory. The

thermodynamic properties were calculated using the following equations.

$$\Delta E_{\text{bie}}^{\text{cp}} = E_{\text{VOC(A)(W)}_n}^{\text{cp}} - \sum_i E_{\text{Monomers},i} \quad (2)$$

$$\Delta E_n^{\text{cp}} = E_{\text{VOC(A)(W)}_n}^{\text{cp}} - E_{\text{VOC(A)(W)}_{n-1}}^{\text{cp}} - E_{\text{W}} \quad (3)$$

In eqn (2) and (3), $E_{\text{VOC(A)(W)}_n}^{\text{cp}}$ is the BSSE corrected binding energy of the cluster. $E_{\text{VOC(A)(W)}_{n-1}}^{\text{cp}}$ is the BSSE corrected binding energy of the $(n-1)$ cluster. $E_{\text{Monomers},i}$ is the binding energy of the monomer. E_{W} binding energy of a water molecule. ΔG and ΔG_n are calculated using eqn (4) and (5), respectively.

$$\Delta G = G_{\text{VOC(A)(W)}_n} - \sum_i G_{\text{Monomers},i} \quad (4)$$

$$\Delta G_n = G_{\text{VOC(A)(W)}_n} - G_{\text{VOC(A)(W)}_{n-1}} - G_{\text{W}} \quad (5)$$

In eqn (4) and (5), $G_{\text{VOC(A)(W)}_n}$is the free energy of VOC(A)$(W)_n$ cluster, $G_{\text{VOC(A)(W)}_{n-1}}$ is the free energy of VOC(A)$(W)_{n-1}$ cluster, $G_{\text{Monomers},i}$ is the free energy of monomers, G_{W} is the free energy of water.

ΔH and ΔH_n are calculated using eqn (6) and (7), respectively.

$$\Delta H = H_{\text{VOC(A)(W)}_n} - \sum_i H_{\text{Monomers},i} \quad (6)$$

$$\Delta H_n = H_{\text{VOC(A)(W)}_n} - H_{\text{VOC(A)(W)}_{n-1}} - H_{\text{W}} \quad (7)$$

In eqn (6) and (7) $H_{\text{VOC(A)(W)}_n}$ is the enthalpy of VOC(A)$(W)_n$ cluster, $\text{H}_{\text{VOC(A)(W)}_{n-1}}$ is the enthalpy of VOC(A)$(W)_{n-1}$ cluster, $H_{\text{Monomers},i}$ is the enthalpy of monomers, H_{W} is the enthalpy of water.

ΔG_n is the successive free energy of an n-hydrate cluster, T is the temperature in K, and R is the molar gas constant, $R = 1.987$ cal (mol K)$^{-1}$. For any particular n-hydrated cluster, the relative population x_n can be computed using the eqn (8).[17,40]

$$x_n = \left(\frac{p(\text{H}_2\text{O})}{p_{\text{ref}}}\right)^n x_0 \text{e}^{-\frac{\Delta G_n}{RT}} \quad (8)$$

p_{ref} is the reference pressure (1 atm) $p(\text{H}_2\text{O})$ is the water partial pressure, which is expressed as:

$$p(\text{H}_2\text{O}) = p(\text{H}_2\text{O})_{\text{eq}} \times \text{RH} \quad (9)$$

$p(\text{H}_2\text{O})_{\text{eq}}$ is the water saturation vapor pressure ($p(\text{H}_2\text{O})_{\text{eq}} = 0.0316$ atm at 298.15 K), RH is the relative humidity. The population of the dry cluster is x_0. x_0 is chosen so that,

$$\sum_0^{10} x_n = 1 \quad (10)$$

We calculated the hydrate distributions (x_n) at relative humidities of 20%, 40%, 60%, 80%, and 100%.

To evaluate the importance of the VOC(A) (VOC = AA, AC, AD, and FD) cluster in the atmosphere, the ratio of VOC(A)(W) to $(\text{A})_2(\text{W})$ and VOC(A) to $(\text{A})_2$ in the following equilibrium reactions, determine X (where X represents none or water):

$$\text{A}\cdot\text{W} + \text{VOC} = \text{VOC(A)W} \quad (\text{R1})$$

$$\text{A}\cdot\text{W} + \text{A} = (\text{A})_2\text{W} \quad (\text{R2})$$

The equilibrium constants K_1 and K_2 for reactions (R_1) and (R_2) are expressed as follows

$$K_1 = \frac{[\text{VOC(A)W}]}{[\text{A}\cdot\text{W}][\text{VOC}]} = \text{e}^{-\frac{\Delta G_1}{RT}} \quad (11)$$

$$K_2 = \frac{[(\text{A})_2\text{W}]}{[\text{A}\cdot\text{W}][\text{A}]} = \text{e}^{-\frac{\Delta G_2}{RT}} \quad (12)$$

From eqn (11) and (12), the ratio of the concentration of VOC·A·W to $(\text{A})_2\cdot\text{W}$ can be expressed as follows:

$$\frac{[\text{VOC(A)(W)}]}{[(\text{A})_2\text{W}]} = \frac{[\text{VOC}]}{[\text{A}]} \times \text{e}^{-\frac{\Delta(\Delta G)}{RT}} \quad (13a)$$

In the absence of water, the ratio is

$$\frac{[\text{VOC(A)}]}{[(\text{A})_2]} = \frac{[\text{VOC}]}{[\text{A}]} \times \text{e}^{-\frac{\Delta(\Delta G)}{RT}} \quad (13b)$$

$\Delta(\Delta G)$ is related to the difference in Gibbs free energy between (R_1) and (R_2) reactions.

The following equation was used to calculate the concentrations of VOC(A)$(W)_n$ ($n = 0$–3) clusters:

$$k_{\text{eq0}} = \frac{[(\text{VOC})(\text{A})]}{\{[\text{VOC}]\cdot[\text{A}]\}} \quad (14)$$

$$k_{\text{eq1}} = \frac{[(\text{VOC})(\text{A})(\text{W})]}{\{[(\text{VOC})(\text{A})]\ [\text{W}]\}} \quad (15)$$

$$k_{\text{eq2}} = \frac{[(\text{VOC})(\text{A})(\text{W})_2]}{\{[(\text{VOC})(\text{A})(\text{W})]\ [\text{W}]\}} \quad (16)$$

$$k_{\text{eq}n} = \frac{[(\text{VOC})(\text{A})(\text{W})_n]}{\{[(\text{VOC})(\text{A})(\text{W})_{n-1}]\ [\text{W}]\}} \quad (17)$$

In eqn (11)–(14), [VOC] is the concentration of VOC (VOC = AA, AC, AD, and FD), [A] is the ammonia concentration, and [W] is the concentration of water, [(VOC)·(A)] is the concentration of the VOC-ammonia cluster, and [(VOC)(A)(W)] is the concentration of the VOC-ammonia-water cluster. $k_{\text{eq}n}$ is the equilibrium constant.

The evaporation rate $\gamma_{i(i+1)}$, of an i-th molecule from a cluster $i+j$, which is formed through the association of the i-th and j-th molecules, can be described as a function of their Gibbs free energies of formation.[41,42]

$$\gamma_{i(i+1)} = \beta_{ij}\frac{P}{k_{\text{b}}Tn_0}\exp\left(\frac{\Delta G_{i+j} - \Delta G_i - \Delta G_j}{RT}\right) \quad (18)$$

In this expression, ΔG_{i+j}, ΔG_i, ΔG_j represent the Gibbs free energies of formation of the cluster $i+j$ and the monomers or clusters i and j, respectively, at the reference pressure P and temperature T. Here k_{b} is the Boltzmann constant, n_0 is the initial number of molecules in the cluster, and β_{ij} is the

collision rate under thermal equilibrium conditions. The collision rate, based on kinetic gas theory, is given by ref. 41–44

$$\beta_{ij} = (r_i + r_j)^2 \sqrt{8\pi k_b T \left(\frac{1}{m_i} + \frac{1}{m_j}\right)} \quad (19)$$

where m_i and m_j denote the molecular masses, r_i and r_j are the radii of the molecules i and j, respectively.

2.4. Optical properties

The calculated isotopic mean polarizability ($\bar{\alpha}$) of H_2O(W), NH_3(A), CH_3COOH(AA), CH_3COCH_3(AC), CH_3COH(AD), and HCOH(FD) at both M06-2X/6-311++G(2p,2d) and CAM-B3LYP/aug-cc-pVDZ level of theory and compare it with experimental value (see Table S24 in SI). After comparison with experimental value, we concluded that the findings indicate that CAM-B3LYP/aug-cc-pVDZ more accurately replicates experimental polarizabilities than M06-2X/6-311++G(2p,2d). The enhanced performance results from using a long-range corrected functional in conjunction with a correlation-consistent augmented basis set, which more accurately records the electronic response. Consequently, CAM-B3LYP/aug-cc-pVDZ is the most dependable method for forecasting molecular polarizabilities. The CAM-B3LYP/aug-cc-pVDZ level of theory was used to calculate the mean isotropic polarizabilities ($\bar{\alpha}$) and anisotropic polarizabilities ($\Delta\alpha$) of atmospheric VOC(A)(W)$_n$, (VOC = AA, AC, AD, FD) (n = 0–1, 15, 20, 25, 30) clusters to determine the Rayleigh scattering intensities. Elm J. *et al.* reported that CAM-B3LYP/aug-cc-pVDZ was a good balance between efficiency and accuracy based on an earlier benchmark study,[34,45] resulting in agreement with both experimental and CCSD(T) polarizability values. Static (∞ nm), 700, 600, 500, and 400 nm wavelengths were used to determine polarizabilities. All of the optical parameters are reported in atomic units, and the calculations were done with the Gaussian-16.[46] Rayleigh scattering happens when light interacts with particles, such as atmospheric gas molecules, that are much smaller than its wavelength. Using the CAM-B3LYP/aug-cc-pVDZ level of theory, the Rayleigh scattering intensities of the atmospheric molecular clusters were estimated by calculating the mean isotropic polarizabilities ($\bar{\alpha}$) and anisotropic polarizabilities ($\Delta\alpha$). For molecules that are randomly orientated in the gas phase, it is beneficial to establish scalar measurements of the polarizability tensor (α_{ij}). The isotropic mean polarizability ($\bar{\alpha}$) and the anisotropic polarizability ($\Delta\alpha$) are two instances of scalar values that are often used to describe eqn (20) and (21), respectively. The optical properties ($\bar{\alpha}$, $\Delta\alpha$, $\Re_n$, σ) for the gas phase clusters were determined using eqn (20)–(25).

$$\bar{\alpha} = \frac{1}{3}\sum_i \alpha_{ii} \quad (20)$$

$$(\Delta\alpha)^2 = \frac{1}{2}\left[(\alpha_{xx} - \alpha_{yy})^2 + (\alpha_{yy} - \alpha_{zz})^2 + (\alpha_{zz} - \alpha_{xx})^2\right] + 3\left[(\alpha_{xy})^2 + (\alpha_{xz})^2 + (\alpha_{yz})^2\right] \quad (21)$$

The polarizability tensor (α_{ij}), the indices $i,j = x, r, z$ are the Cartesian coordinate axes, α_{xx}, α_{yy}, α_{zz} are the diagonal elements of the tensor and α_{xy}, α_{yz}, α_{xz} are of diagonal elements of the tensor.

The Rayleigh scattering activities of the $\Re_{p\parallel}$, $\Re_{p\perp}$ components of linearly polarized light, as well as the $\Re_n$ The component of natural light is given by:

$$\Re_n = 45(\bar{\alpha})^2 + 13(\Delta\alpha)^2 \quad (22)$$

$$\Re_{p\parallel} = 6(\Delta\alpha)^2 \quad (23)$$

$$\Re_{p\perp} = 45(\bar{\alpha})^2 + 7(\Delta\alpha)^2 \quad (24)$$

The so-called depolarisation ratio (σ) is a crucial variable in scattered light measurements. It is described as the ratio of the parallel ($\Re_{p\parallel}$) perpendicular ($\Re_{p\perp}$) intensities to the scattering plane.

$$\sigma = \frac{\Re_{p\parallel}}{\Re_{p\perp}} = \frac{6(\Delta\alpha)^2}{45(\bar{\alpha})^2 + 7(\Delta\alpha)^2} \quad (25)$$

3. Results and discussions

3.1. Structures and topology

The M06-2X/6-311++G(2d,2p) was used to optimize the VOC(A)(W)$_n$, (VOC = AA, AC, AD, FD) (n = 0–10, 15, 20, 25, 30) cluster. Optimized structures are shown in Fig. 1 and 2 (large clusters are shown in Fig. S1 to S4 in the SI), and coordinates are presented in Table S1 (see SI). To form a circle, the AA(A) clusters were hydrated by between 0 and 3 water molecules. On the other hand, the AC(A), AD(A), and FD(A) clusters were hydrated by between 2 and 4 water molecules. The structure of AC(A)(W)$_n$, AD(A)(W)$_n$, FD(A)(W)$_n$ (n = 2 and 4) clusters were in agreement with the results obtained by Chen *et al.*[47] AA(A), AC(A), AD(A), and FD(A) clusters are more likely to form a cage when they are hydrated with 4–10, 15, 20, 25, and 30 water molecules. Clusters AA(A), AC(A), AD(A), and FD(A) hydrated by 4–10, 15, 20, 25, and 30 water molecules contain more hydrogen bonds (H-bonds) than those hydrated by 0–3 water molecules. Thus, we can conclude that clusters with three-dimensional cage structures are more stable than those with planar, linear, or circular structures. This conclusion is further corroborated by topological and thermal analysis (Fig. 3).

AIM analysis can determine the characteristics of hydrogen bonds, and NCI offers a graphical representation of the regions in which NCI occurs in real space.[16,33] The nature and strength of the bonds between molecules are revealed by electron density (ρ) and its Laplacian ($\nabla^2\rho$) measurement at BCP.[32,33,48] Tables S2–S5 and Fig. S5–S8 in the SI provide the AIM parameters of the studied clusters. A positive value of the Laplacian ($\nabla^2\rho$) indicates that the kinetic electron energy density (N) exceeds the potential electron energy density (V), which suggests that the interaction is non-covalent.[33] H_2N–H$\cdots$O=C, and HO–H$\cdots$O=C (O=C: carbonyl group of AA, AC, AD, and FD) hydrogen bonds

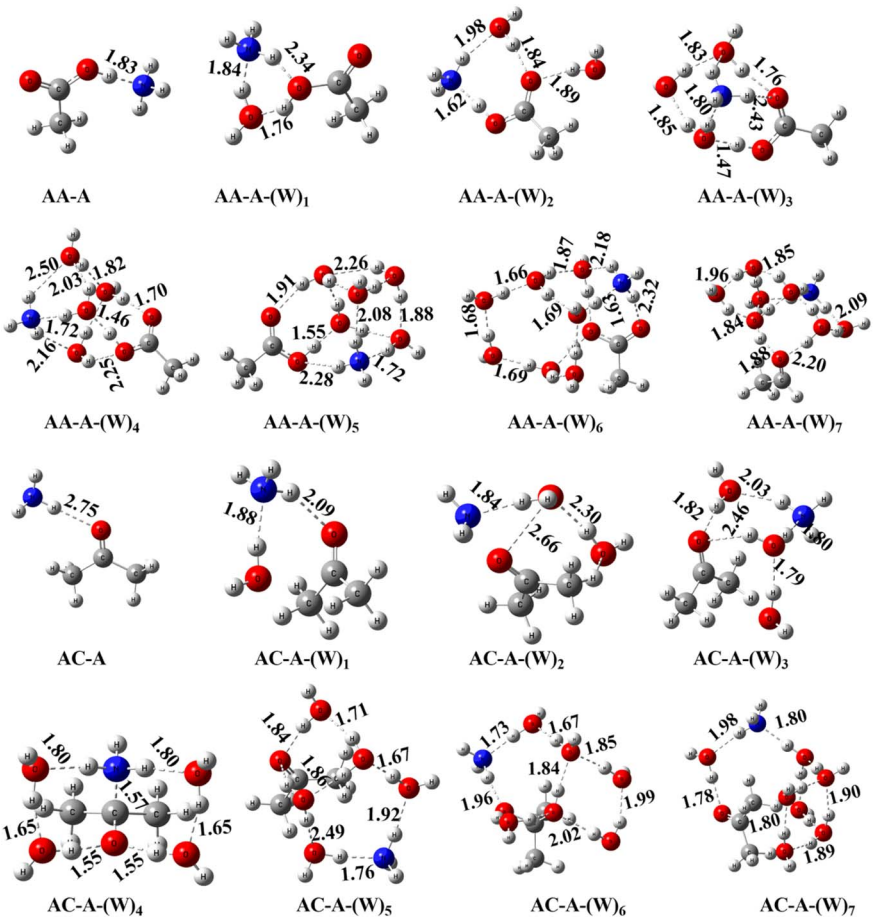

Fig. 1 The optimized structure of AA(A)(W)$_n$, AC(A)(W)$_n$ (n = 0–7), at M06-2X/6-311++G(2d,2p).

have a positive value of Laplacian ($\nabla^2\rho$) at BCP in VOC(A)(W)$_n$, (VOC = AA, AC, AD, FD) (n = 0–10, 15, 20, 25, 30) clusters. The nature of H_2N–H···O=C, and HO–H···O=C hydrogen bonds are non-covalent in AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters, because of the positive value of the Laplacian ($\nabla^2\rho$) of H_2N–H···O=C, and HO–H···O=C hydrogen bonds. In AA(A)(W)$_n$ C(O)HO···HNH_2, and C(O)O–H···NH_3, hydrogen bonds have a positive value of Laplacian ($\nabla^2\rho$) at BCP, which indicates that these bonds are non-covalent. We have plotted the NCI plots in Fig. 4 using the MultiWfn software, based on the optimisation structures calculated at the M06-2X/6-311++G(d,p) level of theory. The non-covalent interaction (NCI) analysis method has been demonstrated between repulsive steric contacts, hydrogen bonds, and van der Waals interactions. It also visually represents the spatial regions where non-covalent interactions occur.[49–51] NCI plots are produced using the RDG (a.u) *versus* (sign λ_2)ρ (a.u) plots, where the (sign λ_2)ρ is the electron density, which is multiplied by the sign of the second Hessian eigenvalue (λ_2). This investigation indicates the presence of weak interactions. When the requirement sign $\lambda_2\rho < 0$ is satisfied, it suggests the existence of strong non-covalent interactions due to hydrogen bonding. Conversely, instances where the value of sign $\lambda_2\rho \approx 0$, indicate weak interactions due to van der Waals forces. The presence of a repulsive force attributable to the steric effect may be confirmed when sign $\lambda_2\rho > 0$.[52–55] Fig. S9–S12 show the NCI plots of the studied clusters. The colored RDG scatterplot of VOC(A)(W)$_n$ (n = 1–2) cluster shows no red spike, indicating the absence of repulsive interactions in these clusters. Small pikes observed beyond two water molecules in VOC(A)(W)$_n$ (n = 3–10, 15, 20, 25, 30) clusters indicate some amount of steric repulsion is present in these clusters. In VOC(A)(W)$_n$ (n = 0, 1–10, 15, 20, 25, 30) clusters, RDG scatterplot, shows the spike in the green region at 0.01 a.u to −0.02 a.u, which indicates van der Waal's interaction present in VOC(A)(W)$_n$ clusters.

Strong hydrogen bonding interaction in VOC(A)(W)$_n$ clusters, as indicated by the spike in the blue region at −0.03 a.u. to −0.05 a.u. as shown by the RDG scatterplot. At the same time, the repulsive interaction between two oxygen atoms in water and a bulky methyl group is represented by a red isosurface. The AC(A)$(H_2O)_n$ clusters exhibit a stronger repulsive interaction than with the other clusters, as indicated by a more red spike in the RDG-scatter graph for the former, and may be attributed to the two bulky ($-CH_3$) groups in AC(A)$(H_2O)_n$ clusters.

3.2. Thermochemical analysis

We examined the thermodynamic properties of VOC(A)(W)$_n$, (VOC = AA, AC, AD, FD) (n = 0–10, 15, 20, 25, 30) to assess their

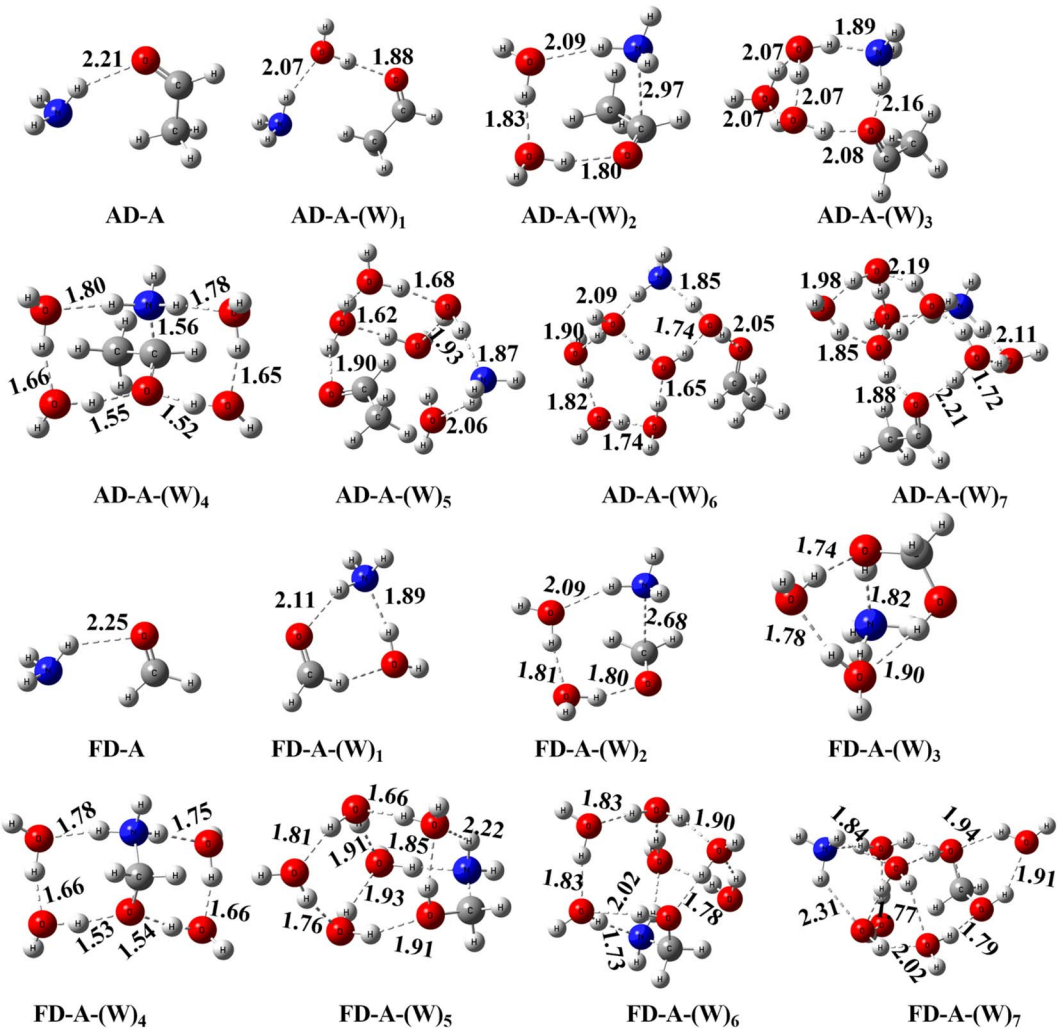

Fig. 2 The optimized structures of AD(A)(W)$_n$, FD(A)(W)$_n$ (n = 0–7), at M06-2X/6-311++G(2d,2p).

potential formation under atmospheric conditions. The lowest-energy structures and corresponding properties were obtained at the M06-2X/6-311++G(2d,2p) level of theory, without BSSE correction, in line with benchmark calculations.[27,56–58] This paper discusses binding energies, enthalpies, and free energies. Frequency calculations ascertain the zero-point correction, the thermal correction to enthalpy, and the thermal correction to free energy are calculated at M06-2X/6-311++G(2d,2p) level of theory. Total binding energies, total enthalpies, and total free energies for all isomers of VOC(A)(W)$_n$were computed. Additionally, we calculated successive binding energies (ΔE_n^{cp}), successive enthalpies (ΔH_n), and incremental successive free energies (ΔG_n) for the VOC(A)(W)$_n$clusters.

The BSSE-corrected total binding energies (ΔE_{bie}^{cp}) and successive binding energies (ΔE_n^{cp}) of VOC(A)(W)$_n$ clusters are calculated at 298.15 K. The values (see Tables S6–S9 in the SI) of total binding energies are plotted in Fig. 5. With the growth of these clusters, the total binding energy of clusters decreases, which implies that the VOC(A)(W)$_n$ clusters are energetically favorable. The successive binding energies (ΔE_n^{cp}) of VOC(A)(W)$_n$ clusters are given in Tables S6–S9. It is noted that the growth of ΔE_n^{cp} for all clusters during each hydration phase is negative. The ΔE_n^{cp} for AA(A)(W)$_n$ (n = 0–10) clusters are −5.60 to −14.28, −4.30 to −17.16, −6.10 to −17.18 kcal mol^{-1} at M06-2X, PW91PW91, ωB97X-D respectively. The ΔE_n^{cp} for AC(A)(W)$_n$ (n = 0–10) clusters are −4.64 to −19.09, −2.05 to −17.56, −3.86 to −17.67 kcal mol^{-1} at M06-2X, PW91PW91, ωB97X-D respectively. The ΔE_n^{cp} for AD(A)(W)$_n$ (n = 0–10) clusters are −4.13 to −15.98, −3.57 to −14.06, −4.21 to −14.61 kcal mol^{-1} at M06-2X, PW91PW91, ωB97X-D respectively. The ΔE_n^{cp} for FD(A)(W)$_n$ (n = 0–10) clusters are −4.19 to −25.53, −3.25 to −22.54, −3.78 to −24.38 kcal mol^{-1} at M06-2X, PW91PW91, ωB97X-D respectively. The ΔE_n^{cp} indicate that newly formed clusters are thermodynamically more favourable than the pre-existing clusters. The ΔE_n^{cp} estimation emphasises the gradual stabilisation of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters that happens when a new water molecule is added. This provides a more accurate picture of the progression of binding energy changes as cluster sizes change.

The enthalpies (ΔH) and successive enthalpies (ΔH_n) of VOC(A)(W)$_n$ clusters were calculated at M06-2X/6-311++G(2d,2p) level of theory using eqn (6) and (7), respectively, at 298.15 K.

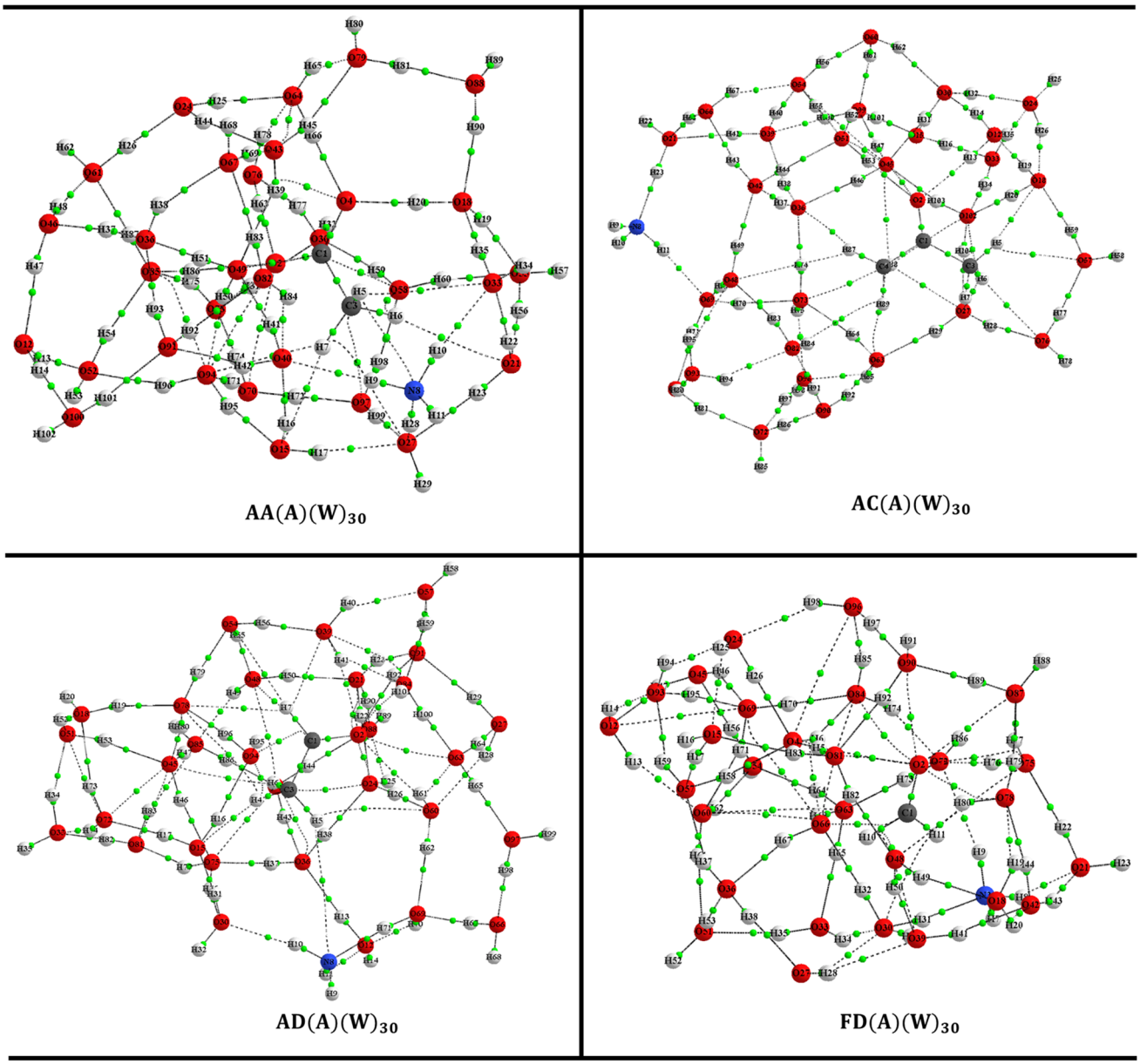

Fig. 3 Molecular topography of AA(A)$(W)_{30}$, AC(A)$(W)_{30}$, AD(A)$(W)_{30}$, and FD(A)$(W)_{30}$ clusters. It is obtained from the AIMALL software.[35]

Fig. 6 shows the enthalpies (ΔH) of VOC(A)$(W)_n$ clusters. With increasing numbers of water molecules, the intermolecular hydrogen bonding becomes stronger, making the addition of water molecules exothermic and thus thermodynamically favourable. A comparison across VOC(A)$(W)_n$ (VOC = AA, AC, AD, and FD) indicates that they form nearly the same number of intermolecular H-bonds, which directly influences the ΔH values.

The successive enthalpies (ΔH_n) of VOC(A)$(W)_n$ clusters are listed in Tables S10–S13 in SI at 298.15 K. For each hydration step, the consecutive enthalpies of AA(A)$(W)_n$, AC(A)$(W)_n$, AD(A)$(W)_n$ and AD(A)$(W)_n$ clusters are consistently negative, as shown in Tables S10–S13. The ΔH_n for AA(A)$(W)_n$ (n = 0–10) clusters are −3.88 to −12.26 kcal mol^{-1}. The ΔH_n for AC(A)$(W)_n$, AD(A)$(W)_n$ and FD(A)$(W)_n$ (n = 0–10) clusters are −3.31 to −17.01 kcal mol^{-1}, −2.65 to −12.74 kcal mol^{-1}, and −2.63 to −20.43 kcal mol^{-1} respectively. In terms of ΔH_n the formation of VOC(A)$(W)_n$ (VOC = AA, AC, AD, FD) clusters are thermodynamically favourable.

The free energy (ΔG) and successivefree energies (ΔG_n) were calculated throughout a range of atmospherically significant temperatures (T) and pressures (P) (from T and P = 298.15 K, and 1 atm to T and P = 216.65 K, and 0.2334 atm) to connect our results to the chemistry of the atmosphere, at the M06-2X/6-311++G(2d,2p). The International Standard Atmosphere (ISA) table was used to calculate these pressures and temperatures.[59] In Fig. 7, we have shown free energies (ΔG) with the cluster sizes for the AA(A)$(W)_n$, AC(A)$(W)_n$, AD(A)$(W)_n$, and FD(A)$(W)_n$ clusters at the T = 298.15 K to T = 216.65 K range. Concerning the hydrated and anhydrated clusters in Fig. 1 and 2, the

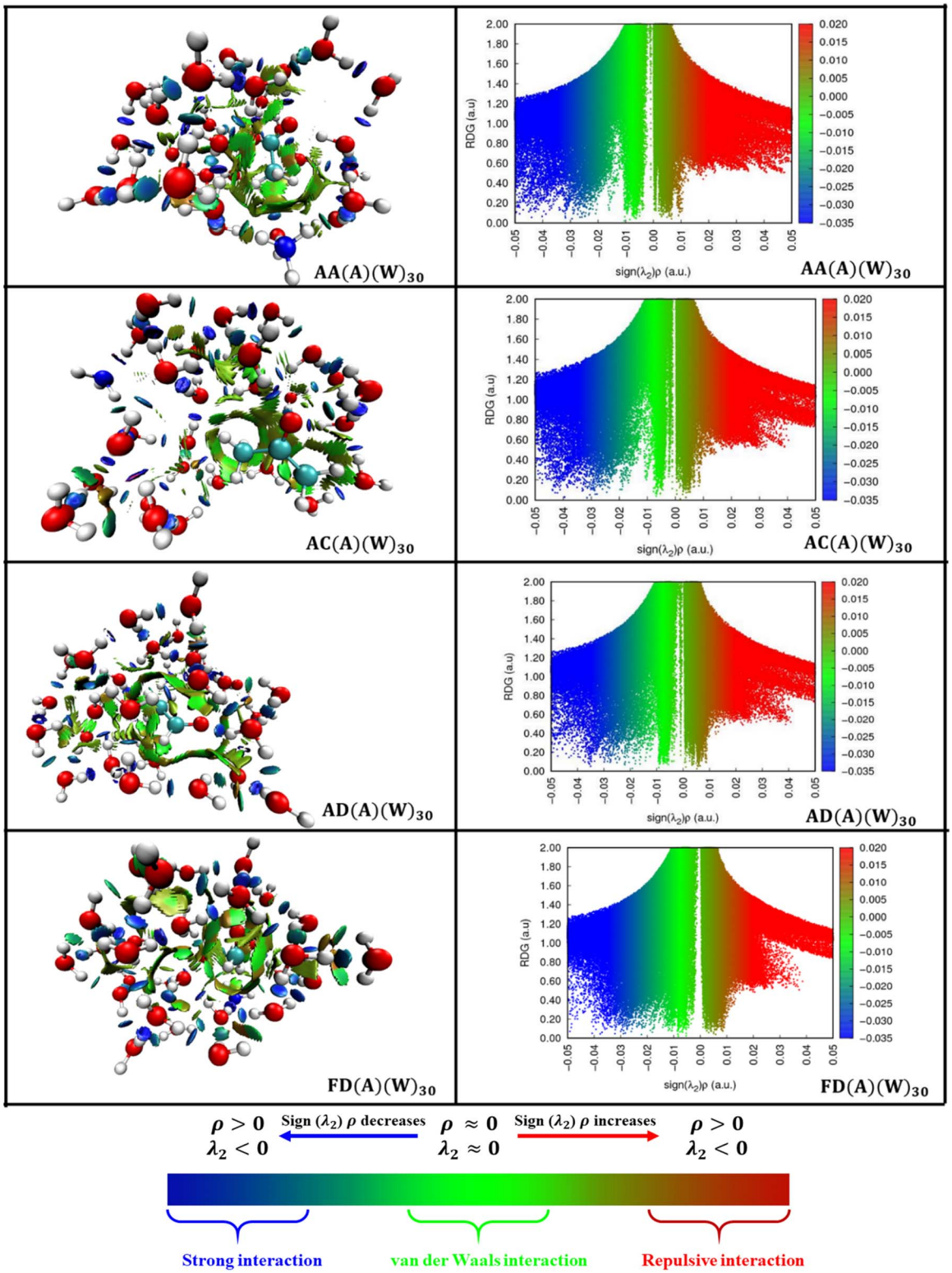

Fig. 4 NCI and RDG plot of AA(A)$(W)_{30}$, AC(A)$(W)_{30}$, AD(A)$(W)_{30}$, and FD(A)$(W)_{30}$ clusters. It is obtained by the MultiWfn application[37] and the Visual Molecular Dynamics (VMD).[38]

anhydrated AA(A) cluster has a lower binding free energy (ΔG) than the anhydrated AC(A), AD(A), and FD(A) clusters. It indicates that the formation of the AA(A) cluster is more favorable than AC(A), AD(A), and FD(A) clusters because of the strong interaction between the O–H group of $CH_3C(O)O$–H with NH_3 ($C(O)O$–H$\cdots NH_3$), such interaction is absent in AC(A), AD(A),

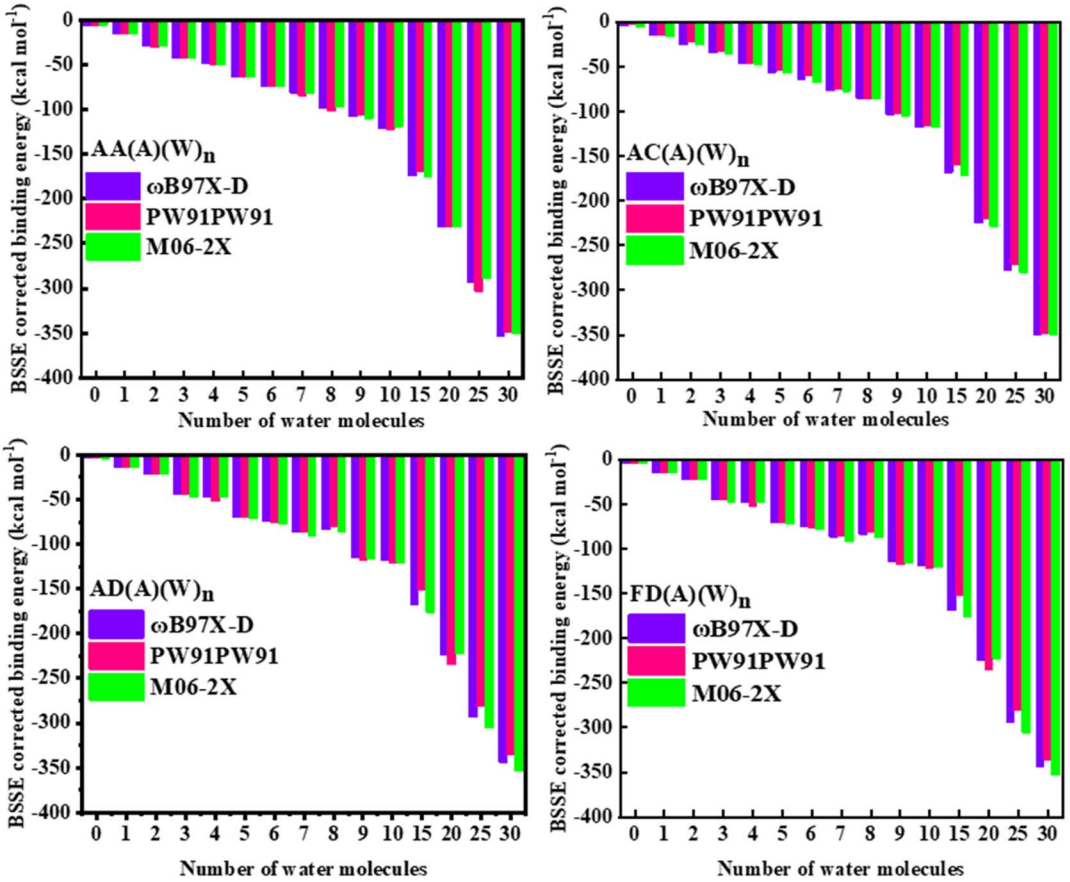

Fig. 5 The BSSE-corrected binding energies (kcal mol^{-1}) were calculated at the DFT/6-311++G(2d,2p) theoretical levels for the AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters (DFT: M06-2X, PW91PW91, and ωB97X-D).

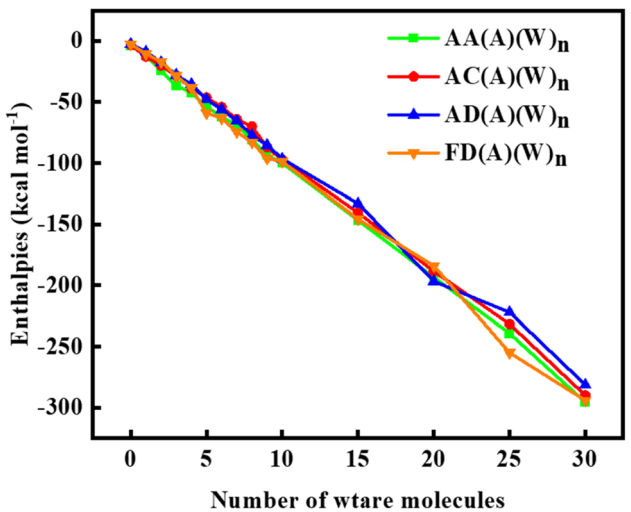

Fig. 6 The enthalpy as a function of water molecules for AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters (n = 0–10, 15, 20, 25, 30) at 298.15 K.

FD(A) clusters. In Tables S14–S17, the computed successive free energies (ΔG_n) are provided. Fig. 7, demonstrates that the ΔG becomes negative below a certain T and P for a specific cluster of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ (n = 0–10, 15, 20, 25, 30). The ΔG of AA(A)(W)$_n$ clusters are −3.19 to −51.78 kcal mol^{-1}. The free energies of AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters are 1.67 to −43.25 kcal mol^{-1}, 0.52 to −33.26 kcal mol^{-1}, 2.90 to −43.00 kcal mol^{-1} respectively, at T and P = 216.65 K, and 0.2234 atm,. Compared with AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters, AA(A)(W)$_n$ clusters have more negative ΔG because the carbonyl (O=C) and hydroxyl (–OH) groups of AA participate in the formation of hydrogen bonds with water and ammonia in AA(A)(W)$_n$ cluster. In AA(A)(W)$_n$ clusters, the carbonyl (O=C) group of AA acts as a hydrogen bond acceptor, and AA's hydroxyl (–OH) group of AA acts as a hydrogen bond acceptor and hydrogen bond donor. Since the AA(A)(W)$_n$ cluster has a more negative ΔG at 286.15 K to 216.65 K, the formation AA(A)(W)$_n$ cluster is thermodynamically more favorable, and it has a higher probability of surviving in the environment. Thus, AA is a suitable precursor for the NPF among AA, AC, AD, and FD. From Fig. 7, we can conclude that VOC(A) + nW = VOC(A)(W)$_n$ (VOC = AA, AC, AD, and FD) (n = 0–10, 15, 20, 25, and 30) the reaction is exothermic (negative value of binding enthalpies) in nature, T and P = 298.15 K, and 1 atm, to T and P = 216.65 K, and 0.2234 atm. Fig. 6 shows that intermolecular hydrogen bonding increases with exothermic enthalpy values as the number of water

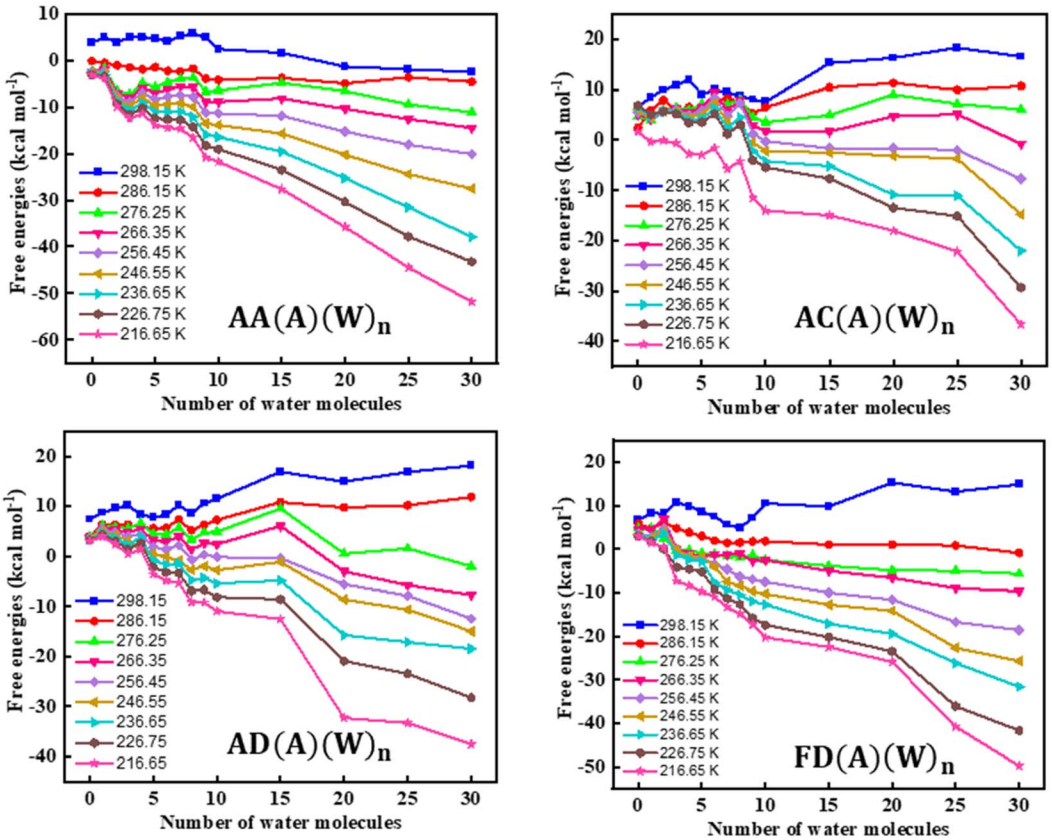

Fig. 7 ΔG as a function of water molecules for AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters (n = 0–10, 15, 20, 25, 30) at T and P = 298.15 K, and 1 atm to T and P = 216.65 K, and 0.2234 atm.

molecules increases. The free energies decrease in negativity as temperature and pressure reduce from T and P = 287.15 K and 1 atm to T and P = 216.65 K and 0.2234 atm. In light of this, the formation of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters become favorable concerning binding enthalpy and free energy as the temperature (T) and pressure (P) decrease. The equilibrium constants for the formation of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters were calculated using the successivefree energy (ΔG_n) at 298.15 K. The equilibrium constant of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters are given in Tables S10–S13.

Formation of a cluster is unfavourable at higher temperatures from the free energy ($\Delta G = \Delta H - T\Delta S$) point of view. Because monomers lose their translational and rotational flexibility when they aggregate, cluster formation is entropically unfavourable even if it is enthalpically favoured by hydrogen bonding. The equilibrium shifts towards free monomers at high temperatures when the $-T\Delta S$ term takes centre stage and the increased thermal motion further destabilises weak intermolecular interactions.

The estimated step-wise successive free energy (ΔG_n) of VOC(A)(W)$_n$ clusters are listed in Tables S14–S17. According to Tables S14–S17, the successive free energy of formation for AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$ and FD(A)(W)$_n$ clusters of a given size become more negative as temperature and pressure decrease. Under normal atmospheric circumstances, when high temperatures and pressures predominate, this thermodynamic tendency suggests that cluster formation is less favourable at higher temperatures and pressures, which may restrict their stability and chance of forming. These results also imply that lower temperatures or lower atmospheric layers would be more suited for the formation and expansion of such clusters. The findings demonstrate a substantial temperature and pressure dependency in the formation of VOC(A)(W)$_n$ (VOC = AA, AC, AD, FD).

3.3. Hydrated distribution and the impact of humidity

As demonstrated by prior research, the nucleation of amines and organic acids depends extensively on hydration.[60,61] Eqn (8) was used to determine the hydrated distributions of the studied cluster at 20%, 40%, 60%, 80%, and 100% relative humidities. The hydrate distributions of core VOC(A) (VOC = AA, AC, AD, and FD) at five different RHs values (20%, 40%, 60%, 80%, 100%) at a temperature of 298.15 K are presented to investigate the types of clusters that prevail at a given RH. Fig. 8 shows the hydrate distributions for the AA(A), AC(A), AD(A), and FD(A) clusters at five distinct relative humidity levels (20%, 40%, 60%, 80%, 100%). According to these results, the proportions of unhydrated AC(A) and FD(A) clusters were almost the same, but the proportion of hydrated clusters increased as RH increased. In the case of AA(A) and FD(A) clusters, with increasing RH, the

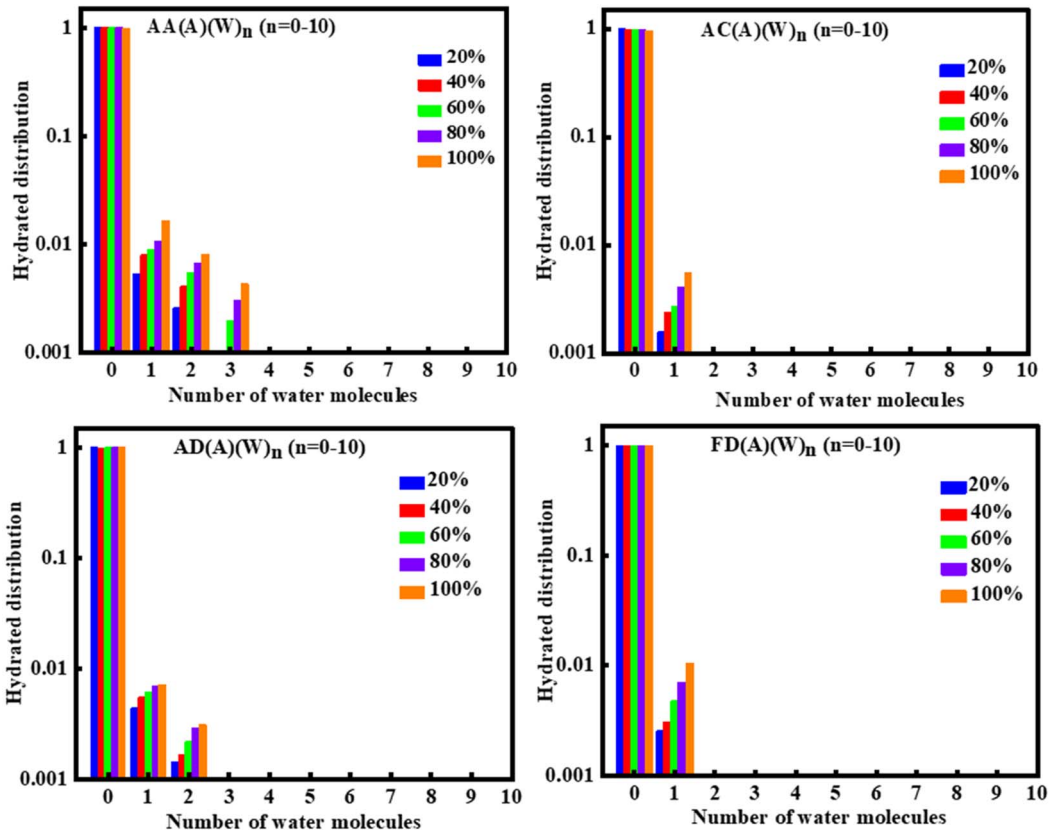

Fig. 8 The hydrated distributions (x_n) of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ (n = 0–10) clusters are determined at relative humidity of 20%, 40%, 60%, 80%, 100%. Note that the logarithmic scale is on the y-axis.

proportion of unhydrated AA(A) and FD(A) clusters decreases. In contrast, the proportions of hydrated AA(A) and AD(A) clusters increase. This result is consistent with the previously obtained result of the hydration of the acetic acid–dimethylamine complex.[62] The hydrated distributions are dominated by entirely unhydrated VOC(A) (VOC = AA, AC, AD, and FD) clusters. Simultaneously, a small portion is occupied by the mono-hydrate VOC(A) cluster, that is, (AA(A)(W)$_1$, AC(A)(W)$_1$, AD(A)(W)$_1$, and FD(A)(W)$_1$), as shown in Fig. 8. As the RH increased, the proportion of mono-hydrated clusters progressively increased from approximately 20% to 80%. The unhydrated cluster is always predominant; monohydrates and dihydrates comprise a small percentage, and trihydrates and tetrahydrates are essentially non-existent, as shown in Fig. 8. The results indicated that the unhydrated VOC(A) cluster was consistently predominant, and the hydrate distributions of the VOC(A)(VOC = AA, AC, AD, and FD) clusters were barely sensitive to relative humidity.

AC(A) and AD(A) are always observed to be unhydrated due to the extremely weak interaction between carbonyl (C=O) groups and water. AA(A) cluster mono, di, and tri hydrate are at high RH because of the strong interaction between the carboxylic acid (–COOH) group and water. AD(A) cluster mono, di, and tri hydrates are formed at high RH. Prior research has shown that the temperature sensitivity of hydrate distribution is relatively low at a constant relative humidity.[63] This might happen because temperature inversely affects the free energy of formation and water concentration. The ΔG decreases with a decrease in temperature, resulting in increased hydration.

3.4. Atmospheric abundances

Our study is related to atmospheric chemistry through the thermochemical analysis section, which calculates the successive free energies and equilibrium constants for the formation of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters at a pressure of 1 atm and temperature 298.15 K. At a temperature of 298.15 K and a pressure of 1 atm in the lower troposphere (100% relative humidity), the concentration of water vapor is approximately 7.80×10^{17} molecules per cm^3.[64] At different locations, the concentrations of acetic acid (AA), acetone (AC), acetaldehyde (AD), and formaldehyde (FD) varied at different locations. Determining the concentrations of the different hydrated AA(A), AC(A), FD(A), and clusters under a particular realistic atmospheric environment is an interesting task. We calculated the atmospheric concentration of VOC(A)(W)$_n$ (VOC = AA, AC, AD, and FD) (n = 0–3) clusters using eqn (11)–(14). To calculate the concentration of VOC(A)(W)$_n$ clusters, we used a water vapor concentration of 7.80×10^{17} molecules per cm^3, and the ammonia concentration was 12.3×10^{10} molecules per cm^3 (5 ppb)[64,65] atmospheric concentration of acetic acid is 0.05–1.9 ppbv over the marine site,[66] 0.2–17.8 ppbv over the urban site,[67] 0.9–1.4 ppbv over the semi-urban site.[68,69] Acetone

concentration range at the troposphere and tropopause is 0.2–3 ppb and 0.1–0.7 ppb, respectively.[4,70,71] In the oceanic boundary layer, simulated acetone concentrations usually are 0.3–0.5 ppb in the southern hemisphere and 0.5–1 ppb in the northern hemisphere.[72] The concentration of acetaldehyde in major urban areas (Brazil: São Paulo, Rio de Janeiro, and Salvador) was 35 ppb.[73] The acetaldehyde concentrations fluctuated between 3–18 and 2–7 ppbv.[74] The formaldehyde concentration in urban and rural areas is 0.065 μg m^{-3} (0.053 ppbv), 0.04 μg m^{-3} (0.033 ppbv).[75] The theoretically calculated atmospheric concentration of AA(A)$(W)_n$, AC(A)$(W)_n$, AD(A)$(W)_n$, and FD(A)$(W)_n$ clusters are provided in Tables S18–S21. Because AA is more abundant than AC, AD and, FD the atmospheric concentration of AA-containing water clusters, that is, AA(A)$(W)_n$ (n = 0–3) are higher than AC, AD, and FD-containing clusters, that is, AC(A)$(W)_n$, AD(A)$(W)_n$, and FD(A)$(W)_n$. This result suggests that AA-containing prenucleation clusters are at least 10^2–10^4 orders of magnitude higher than those of AC, AD and FD, indicating that AA plays an important role in atmospheric nucleation. Previously, Pal *et al.* calculated the atmospheric concentration of methanol–water and ethanol–water clusters; the concentration of methanol–water and water clusters varies from ~10^3 to ~10^8 molecules per cm^3, and ~10^1 to ~10^8 molecules per cm^3, respectively.[10] Comparing our computed values with previously studied results, it is apparent that our determined results are indeed reasonable.

The literature provides the concentration values for acetic acid, acetone, acetaldehyde, formaldehyde, and ammonia in the atmosphere, which are approximately 10^9–10^{11}, 10^9–10^{11}, 10^{10}–10^{11}, 10^8–10^{11}, and 10^9–10^{11} molecules per cm^3 respectively.[68,71,73,75,76] Therefore, the concentration ratios of [AA]/[A], [AC]/[A], [AD]/[A], [FD]/[A] is 10^{-2} to 10^1, 10^{-2} to 10^1, 10^{-1} to 10^2, and 10^{-3} to 10^0 respectively. Table S22 (SI) shows that the concentration ratio of [AA·A·X]/[$(A)_2$·X] was approximately 10^1, and 10^2 times higher than those of [AC·A·X]/[$(A)_2$·X] and [FD·A·X]/[$(A)_2$·X] clusters, respectively. Similarly, the concentration ratio of [AD·A·X]/[$(A)_2$·X] is approximately 10^1, and 10^2 times higher than those of [AC·A·X]/[$(A)_2$·X] and [FD·A·X]/[$(A)_2$·X] clusters. Compared to [AA·A·X]/[$(A)_2$·X] and [AD·A·X]/[$(A)_2$·X] clusters, respectively, the concentration of [AC·A·X]/[$(A)_2$·X] and [FD·A·X]/[$(A)_2$·X] clusters were lower. This outcome is merely an approximation that requires additional experimental confirmation.

3.5. Collision rate and evaporation rate

The collision rate (β_{ij}) and evaporation rate ($\gamma_{i(i+1)}$) of clusters is determined by the formation free energy obtained from quantum chemical calculations, and has been recognised as a crucial metric for examining the first phases of particle formation. The collision rate (β_{ij}) measures the number of molecules interacting with one another in a given time unit, while the evaporation rate measures the rate at which molecules escape from a cluster into the environment in the same time unit. Nucleation, the first stage in the formation of aerosols, may occur as a consequence of molecular collisions, making this amount crucial.[77,78] The collision rate (β_{ij}) and evaporation rate ($\gamma_{i(i+1)}$) for all cluster sizes (n = 0–10) were estimated using the eqn (18) and (19). The results for β_{ij} and $\gamma_{i(i+1)}$ present in Table S23 in SI. All the value of collision rates that have been computed are ~10^{-16} m^3 s^{-1}. This has the potential to start the process of aerosol nucleation and growth of new particles in the atmosphere. In the equilibrium state, the rate at which hydrates evaporate is strongly correlated with their ΔG. The collision and evaporation rates both fluctuate irregularly as hydration increases, as shown in Table S23 (see ESI). Evaporation rate varies 10^5 to 10^{12} s^{-1}, 10^8 to 10^{13} s^{-1}, 10^7 to 10^{15} s^{-1}, 10^6 to 10^{13} s^{-1} for AA(A)$(W)_n$, AC(A)$(W)_n$, AD(A)$(W)_n$, and FD(A)$(W)_n$ (n = 0–10) clusters respectively. A water molecule may readily escape from AC(A)$(W)_4$ cluster due to its higher evaporation rate than other clusters. AA(A)$(W)_2$ clusters have a lower evaporation rate than others, which is 1.10×10^5 s^{-1}. Previously, Ambe *et al.* and Chen *et al.* studied the evaporation rate of water containing organic volatile compounds, and they found that the evaporation rate varies from 10^2 to 10^{16} s^{-1}.[77,78] Therefore, our result is consistence with the previously studied result. The AA(A)$(W)_n$, AC(A)$(W)_n$, AD(A)$(W)_n$, and FD(A)$(W)_n$ (n = 0–10) clusters are short-lived in the atmosphere due to their higher evaporation rate, resulting in rapid removal *via* the release of water molecules into the atmosphere.

3.6. Scattering properties

The extinction properties of atmospheric aerosols influence air visibility, with Rayleigh scattering significantly contributing to these extinction characteristics. The impact of atmospheric prenucleation clusters on solar radiation remains poorly understood. Solar radiation has little influence on prenucleation clusters and contributes to the albedo effect, which entails the reflection of light into space. This facilitates the cooling of the Earth. The kind and distribution of particles in aerosols may either augment or diminish scattering, hence influencing cloud albedo, persistence, and atmospheric temperature.[79] Thus, we have critically investigated the Rayleigh-scattering properties of VOC(A)$(W)_n$ (VOC = AA, AC, AD, FD) (n = 0–10, 15, 20, 25, 30). The Rayleigh scattering ($\Re_n$) activity of the natural light depends on the isotropic ($\bar{\alpha}$) and anisotropic ($\Delta\alpha$) polarizabilities. The size of AA(A)$(W)_n$, AC(A)$(W)_n$, AD(A)$(W)_n$, and FD(A)$(W)_n$ (n = 0–10, 15, 20, 25, 30) clusters were less than 2 nm, so the scattering properties of those clusters were explained by $\Re_n$. Fig. 9–12 represent the $\Re_n$, $\bar{\alpha}$ and $\Delta\alpha$ of VOC(A)$(W)_n$ (VOC = AA, AC, AD, FD) (n = 0–10, 15, 20, 25, 30). Clusters. The isotropic polarizabilities ($\bar{\alpha}$) of AA(A)$(W)_n$, AC(A)$(W)_n$, AD(A)$(W)_n$, and FD(A)$(W)_n$ (n = 0–10, 15, 20, 25, 30) clusters are remarkably analogous. The $\bar{\alpha}$ had a linear correlation with cluster size as the quantity of water molecules increased. For all clusters, considerable isomer dependence and irregularity were observed in the anisotropic polarizabilities ($\Delta\alpha$). For all studied clusters, the $\Re_n$ increases with increasing water molecules in clusters (see Fig. 10). This conclusion agrees with some previous studies.[10,16,34] The $\bar{\alpha}$ of studied clusters increases with an increasing quantity of water molecules. The isotropic mean polarizability ($\bar{\alpha}$) of AA(A)$(W)_n$, AC(A)$(W)_n$, AD(A)$(W)_n$, and FD(A)$(W)_n$ varying from 48.14 to 361.79, 55.04 to

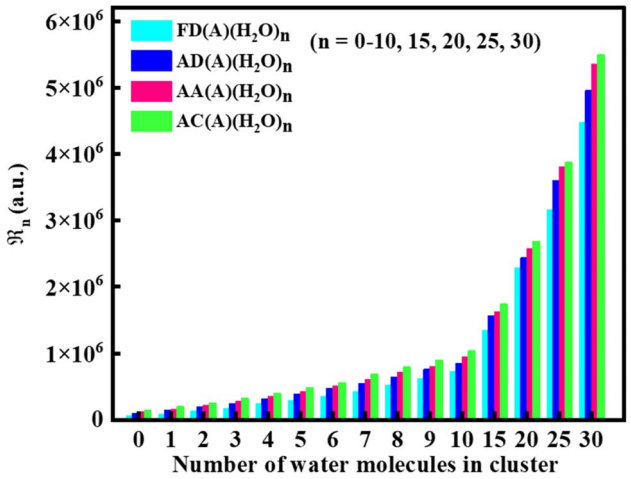

Fig. 9 The $\Re_n$ of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters at the static limit (∞) nm.

359.28, 55.04 to 348.04, and 31.34 to 329.88 a.u. at static limit (∞), 700, 600, 500, and 400 nm, respectively. The $\Re_n$ of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ at the static limit, as shown in Fig. 10, and we can conclude that AC(A)(W)$_n$ clusters have higher Rayleigh scattering intensity ($\Re_n$) than AA(A)(W)$_n$, AD(A)(W)$_n$ and FD(A)(W)$_n$ clusters; this is because of $\Re_n$ depends on isotropic mean polarizabilities ($\bar{\alpha}$); the acetone molecule has higher isotropic mean polarizabilities than acetic acid, acetaldehyde, and formaldehyde (see Table S24 in ESI), which is why acetone-containing clusters, *i.e.*, AC(A)(W)$_n$ have higher Rayleigh scattering intensity ($\Re_n$). The trend for isotropic mean polarizabilities ($\bar{\alpha}$) of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ is FD(A)(W)$_n$ < AD(A)(W)$_n$ < AA(A)(W)$_n$ < AC(A)(W)$_n$, because the effect of A, W, AA, AC, AD, and FD on the isotropic mean polarizability of clusters is as follows: W < A < FD < AD < AA < AC a similar trend is followed for Rayleigh scattering intensity ($\Re_n$) intensity, *i.e.*, FD(A)(W)$_n$ < AD(A)(W)$_n$ < AA(A)(W)$_n$ < AC(A)(W)$_n$. The Rayleigh scattering intensity ($\Re_n$), of AA(A)(W)$_n$ cluster increases 2.98–4.29% at 700 nm, 4.10–5.53% nm at 600 nm, 6.03–7.25% at 500 nm, and 9.90–12.20% at 400 nm. The Rayleigh scattering intensity ($\Re_n$), of AC(A)(W)$_n$ cluster increases 2.98–4.29%, 4.10–5.53% nm, 6.03–7.25% nm, and 9.90–12.20% at 700, 600, 500, and 400 nm, respectively. For AD(A)(W)$_n$, Rayleigh scattering intensity ($\Re_n$) increases by 2.96–3.79%, 4.08–5.27%, 6.01–7.84%, and 9.78–12.97% at 700, 600, 500, and 400 nm, respectively. In the case of FD(A)(W)$_n$, Rayleigh scattering intensity ($\Re_n$) increases by 2.99–4.04%, 4.11–5.62%, 6.02–8.37%, and 9.87–14.01% at 700, 600, 500, and 400 nm, respectively. As the cluster expands, it is found that the computed depolarisation ratios (σ) readily decrease. This is because the anisotropic polarizability is comparatively stable between 6 and 37 a.u., whereas the mean isotropic polarizability increases with the number of molecules for the studied clusters. The depolarisation ratios (σ) of AA(A)(W)$_n$, AC(A)(W)$_n$,

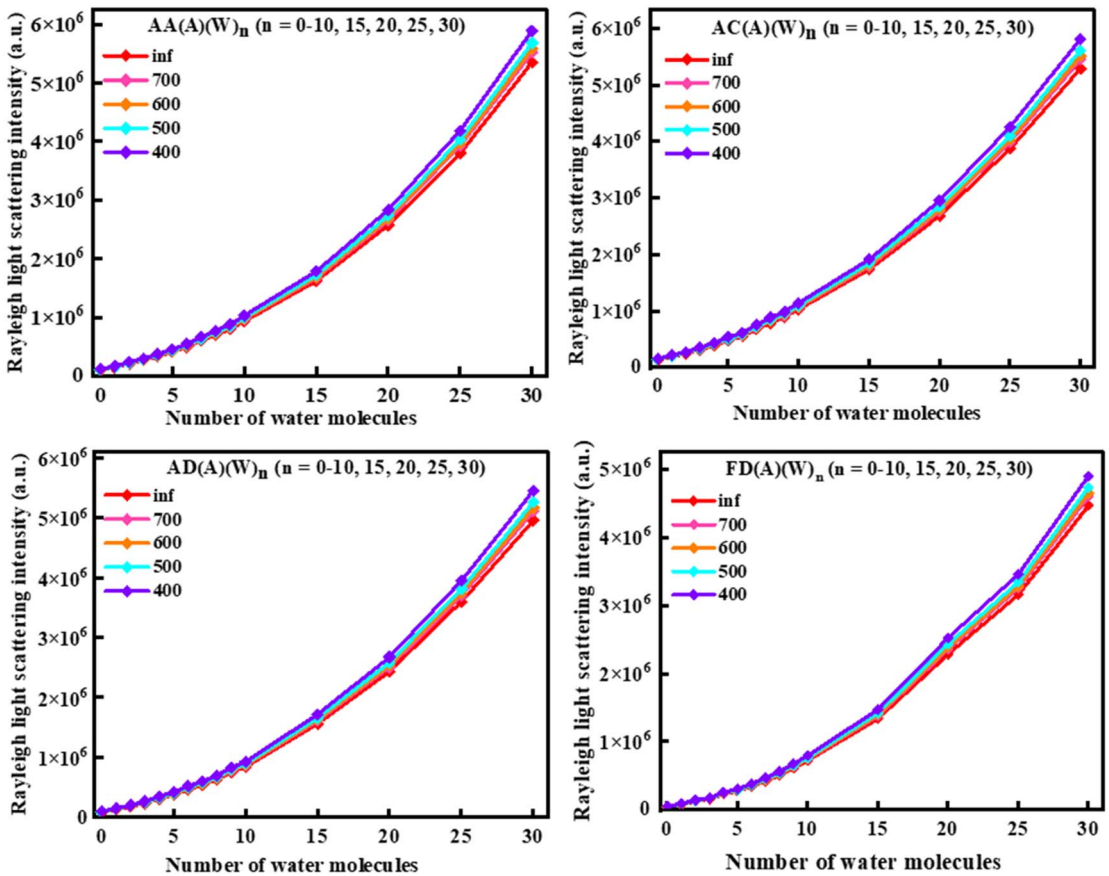

Fig. 10 The Rayleigh scattering intensity of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$, clusters at (∞), 700, 600, 500, 400 nm.

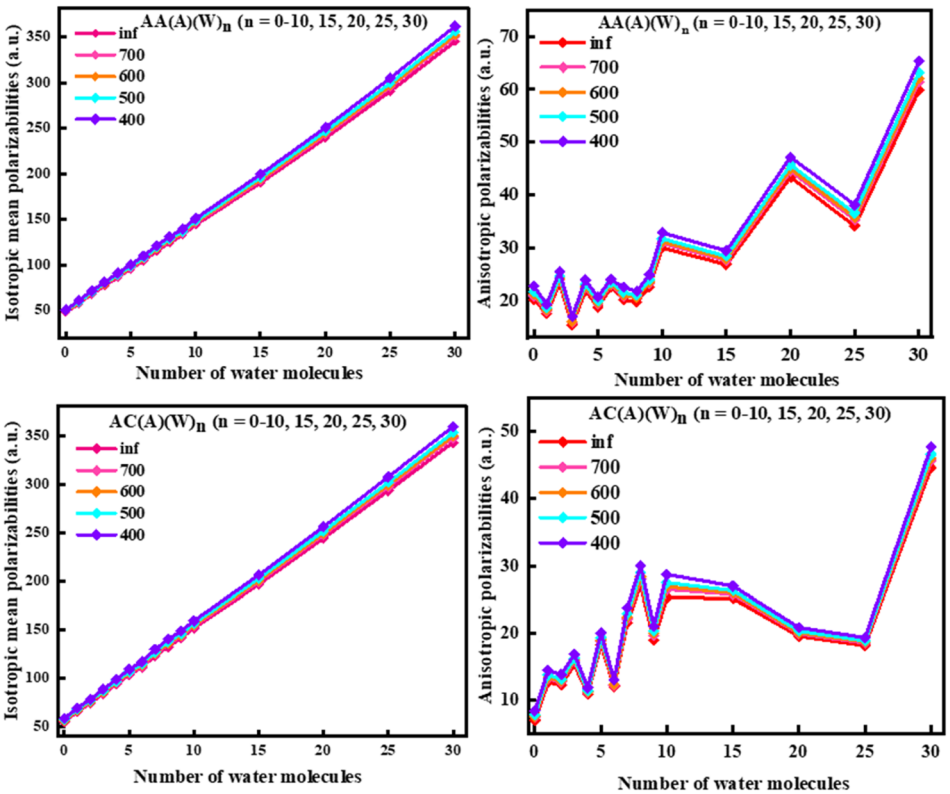

Fig. 11 The $\bar{\alpha}$ and $\Delta\alpha$ of AA(A)(W)$_n$, and AC(A)(W)$_n$ clusters at (∞), 700, 600, 500, 400 nm.

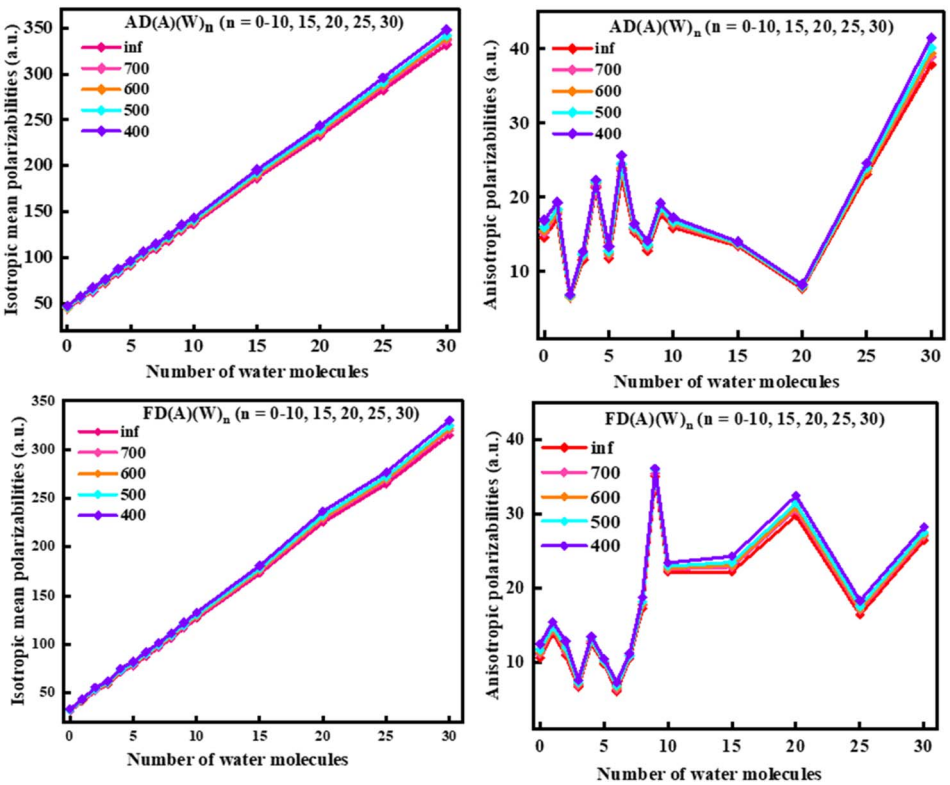

Fig. 12 $\bar{\alpha}$ and $\Delta\alpha$ of AD(A)(W)$_n$, FD(A)(W)$_n$ clusters at (∞), 700, 600, 500, 400 nm.

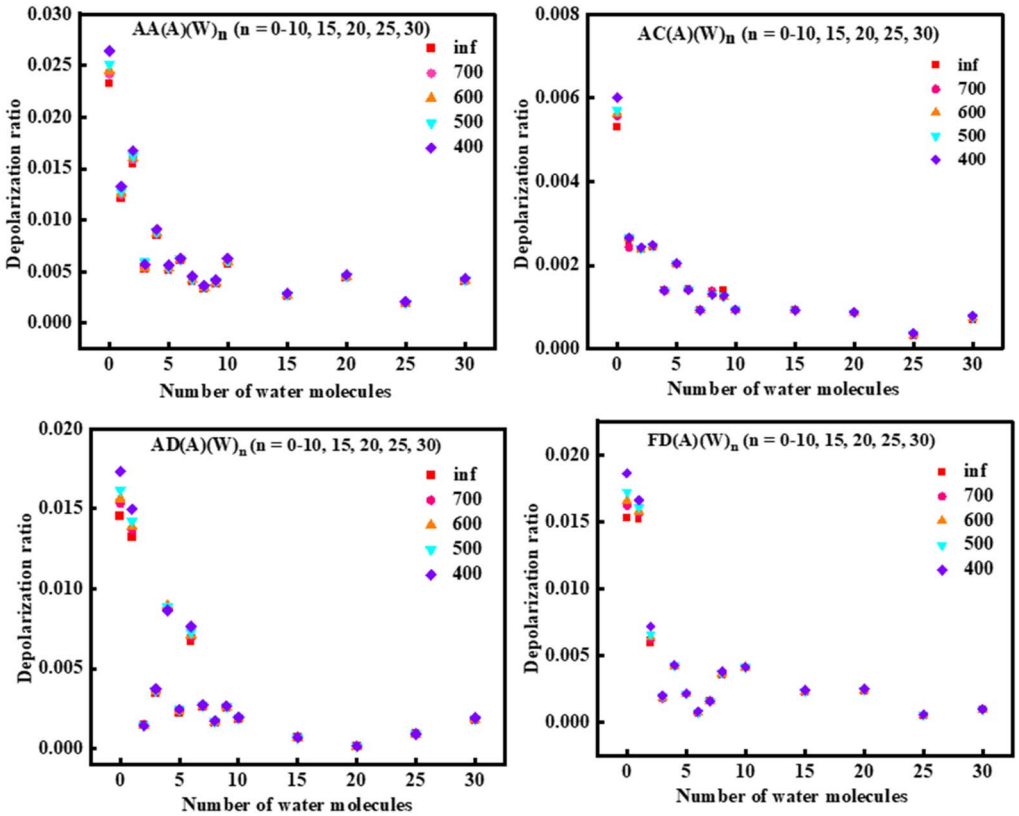

Fig. 13 σ of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters at (∞), 700, 600, 500, 400 nm.

AD(A)(W)$_n$, and FD(A)(W)$_n$ (n = 0–10) clusters fluctuate from 0.0001 to 0.029. For the large AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ (n = 15, 20, 25, 30) clusters, the depolarization ratios (σ) plummets to $\sim 10^{-4}$. The depolarisation ratios (σ) also exhibit a little change with the wavelength.

The quantity of water molecules influences the optical characteristics of the clusters, as the optical parameter values rise with an increasing amount of water molecules. The addition of water molecules or other substituents to the clusters increases in their size, therefore altering their visual properties. They have the potential to alter atmospheric visibility, which in turn affects human output and existence (Fig. 13).[80,81]

4. Conclusion

Density functional theory computations were conducted on a comprehensive array of hydrates of molecular clusters pertinent to the environment. Structural and thermodynamic analyses were conducted to assess the influence of 0–10, 15, 20, 25, and 30 water molecules on the AA(A), AC(A), AD(A), and FD(A) clusters. Five different relative humidity (RH) levels (20%, 40%, 60%, 80%, and 100%) were used to compute the hydrated distributions of the AA(A), AC(A), AD(A), and FD(A) clusters. From a structural perspective, hydrogen bonds are responsible for the formation of the AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters. Lower temperatures (T) and pressures (P) are preferred for the formation of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters, as the values of the free energies became more negative with decreasing temperature and pressure for a given cluster size in the studied temperature and pressure ranges. The percentages of monohydrates and dihydrates are negligible from the standpoint of hydrated distribution, while the percentages of trihydrates and tetrahydrates are almost non-existent. Because of the strong interaction between the carboxylic acid (COOH) group and water, AA(A) mono-, di-, and tri-hydrate clusters were somewhat populated under high relative humidity conditions. We determined that AA and AD had a greater impact on NPF than AC and FD based on the atmospheric concentration ratios of AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$. Collision and evaporation rates determine the stability and lifetime of atmospheric clusters. The stabilities of various clusters are assessed by determining their evaporation rates. It was found that both the thermodynamic and the optical properties of the studied clusters depend on the size of the clusters. Free energy becomes more negative with increasing cluster size, and the Rayleigh light scattering increases quadratically with cluster size. From the studies conducted, it appears that the cluster size is directly related to thermodynamic stability, which implies that a cluster will have a greater size if it has greater thermodynamic stability. And it has also been observed that the intensity of Rayleigh light scattering increases with the size of a cluster. Hence, for a cluster to have higher Rayleigh scattering, it must have higher thermodynamic stability. Acetone-containing clusters, such as AC(A)(W)$_n$, have greater $\Re_n$ than AA(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters because acetone molecules have higher

isotropic mean polarizabilities than acetic acid, acetaldehyde, and formaldehyde. The aerosols' extinction properties were more significantly impacted by the cluster's acetone content. The same growth pattern is observed at all investigated visible light wavelengths (700, 600, 500, and 400 nm). This work could provide fresh light on how different VOCs behave in certain situations for the formation of new particles in the atmosphere. The nature of the hydration profile and scattering characteristics of the AA(A)(W)$_n$, AC(A)(W)$_n$, AD(A)(W)$_n$, and FD(A)(W)$_n$ clusters were also elucidated by this research.

Author contributions

Arnab Patla: conceptualization, methodology, writing, original draft preparation, formal analysis, reviewing, and editing. Ranga Subramanian: conceptualization, methodology, formal analysis, supervision, reviewing, and editing.

Conflicts of interest

The authors report no potential conflict of interest.

Data availability

All relevant data that support this study are within the paper and its supplementary information (SI). Supplementary information: the optimized geometry (large clusters), colored RDG scattered plot and iso surfaces, scattering, and binding energy are in the SI. In addition, the energy information as binding energy, free energy, enthalpies, and atmospheric concentration are also available in the SI. See DOI: **https://doi.org/10.1039/d5ra06836c**.

Acknowledgements

This work was supported by the Council of Scientific and Industrial Research of India (Grant No. 01 (2890)/17/EMR-11) and the Prime Minister Research Fellowship (Id-2702447), Ministry of India. The author expresses gratitude to the Indian Institute of Technology Patna for offering research facilities at IIT Patna.

References

1 J. H. Seinfeld and S. N. Pandis, *Atmospheric Chemistry and Physics: from Air Pollution to Climate Change*, John Wiley & Sons, 2016.
2 W. C. Hinds and Y. Zhu, *Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles*, John Wiley & Sons, 2022.
3 H. Kokkola, P. Yli-Pirilä, M. Vesterinen, H. Korhonen, H. Keskinen, S. Romakkaniemi, L. Hao, A. Kortelainen, J. Joutsensaari and D. R. Worsnop, *Atmos. Chem. Phys.*, 2014, **14**, 1689–1700.
4 H. B. Singh, D. O'hara, D. Herlth, W. Sachse, D. R. Blake, J. D. Bradshaw, M. Kanakidou and P. J. Crutzen, *J. Geophys. Res. Atmos.*, 1994, **99**, 1805–1819.
5 A. J. Kean, E. Grosjean, D. Grosjean and R. A. Harley, *Environ. Sci. Technol.*, 2001, **35**, 4198–4204.
6 R. Holzinger, C. Warneke, A. Hansel, A. Jordan, W. Lindinger, D. H. Scharffe, G. Schade and P. J. Crutzen, *Geophys. Res. Lett.*, 1999, **26**, 1161–1164.
7 W. C. Keene, J. N. Galloway, G. E. Likens, F. A. Deviney, K. N. Mikkelsen, J. L. Moody and J. R. Maben, *J. Atmos. Sci.*, 2015, **72**, 2947–2978.
8 J. N. Galloway, G. E. Likens, W. C. Keene and J. M. Miller, *J. Geophys. Res. Ocean.*, 1982, **87**, 8771–8786.
9 V. F. McNeill, *Environ. Sci. Technol.*, 2015, **49**, 1237–1244.
10 J. Pal, A. Patla and R. Subramanian, *Chemosphere*, 2021, **272**, 129846–129858.
11 N. Goldman, C. Leforestier and R. J. Saykally, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 2005, **363**, 493–508.
12 M. Mahilang, M. K. Deb and S. Pervez, *Chemosphere*, 2021, **262**, 127771–127789.
13 S. K. Syamlal, C. B. S. Kumar, R. P. Reji, P. S. Roshal, Y. Sivalingam and V. J. Surya, *Chemosphere*, 2022, **307**, 135947–135956.
14 W. Xu and R. Zhang, *J. Phys. Chem. A*, 2012, **116**, 4539–4550.
15 F. M. Bowman, J. R. Odum, J. H. Seinfeld and S. N. Pandis, *Atmos. Environ.*, 1997, **31**, 3921–3931.
16 A. Patla and R. Subramanian, *Phys. Chem. Chem. Phys*, 2023, **25**, 7869–7880.
17 H. Wang, X. Wang, Y. Hu, Z. Su, X. Zhang, Q. Zhang, M. H. Hadizadeh, X. Zhao, F. Xu and Y. Sun, *J. Mol. Liq.*, 2023, **390**, 123001–123013.
18 Y. Yu, H. Xing, W. Cheng, W. Cui and R. Mu, *J. Ind. Eng. Chem.*, 2022, **111**, 289–299.
19 S. Trabelsi, M. Tlili, F. Hammami, S. Nasr, M.-C. Bellissent-Funel and J. Darpentigny, *J. Chem. Phys*, 2023, **159**, 224503–224513.
20 J. Gao, C. D. Carlson, J. Thomas, Y. Xu and W. Jaeger, *ChemPhysChem*, 2024, e202400573–e202400581.
21 E. Romero-Montalvo and G. A. DiLabio, *J. Phys. Chem. A*, 2021, **125**, 3369–3377.
22 C. Wang, X. Chen, Y. Liu, T. Huang and S. Jiang, *ACS omega*, 2023, **8**, 15467–15478.
23 C. Zhang, Y. Wang and H. Wang, *Powder Technol.*, 2017, **314**, 9–19.
24 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminform.*, 2012, **4**, 1–17.
25 A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.
26 M. J. ea Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, Ma. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson and H. Nakatsuji, *Gaussian 16, Revision A.03*, Gaussian, Inc., Wallingford, CT, 2016.
27 J. Elm, M. Bilde and K. V Mikkelsen, *J. Chem. Theory Comput.*, 2012, **8**, 2071–2077.
28 J. V. Kildgaard, K. V. Mikkelsen, M. Bilde and J. Elm, *J. Phys. Chem. A*, 2018, **122**, 5026–5036.
29 N. Myllys, J. Elm and T. Kurten, *Comput. Theor. Chem.*, 2016, **1098**, 1–12.
30 A. Patla, D. Kumari, A. Kumar, R. Jana and R. Subramanian, *Mol. Phys.*, 2025, e2478202–e2478222.

31 J. Pal, P. S. Teja and R. Subramanian, *Theor. Chem. Acc.*, 2020, **139**, 1–14.
32 A. Patla, J. Pal, K. Guleria and R. Subramanian, *Mol. Phys.*, 2024, e2325046–e2325057.
33 A. Patla, J. Pal and R. Subramanian, *Mol. Phys.*, 2024, e2372454–e2372472.
34 J. Elm, P. Norman, M. Bilde and K. V Mikkelsen, *Phys. Chem. Chem. Phys.*, 2014, **16**, 10883–10890.
35 T. A. Keith and T. K. Gristmill, Softw. Overl (Park, KS, USA), 2015, **http://aim.tkgristmill.com**.
36 M. Jabłoński, *ChemistryOpen*, 2019, **8**, 497–507.
37 T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580–592.
38 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.
39 S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553–566.
40 H. Henschel, J. C. A. Navarro, T. Yli-Juuti, O. Kupiainen-Maatta, T. Olenius, I. K. Ortega, S. L. Clegg, T. Kurtén, I. Riipinen and H. Vehkamaki, *J. Phys. Chem. A*, 2014, **118**, 2599–2611.
41 J. Elm, D. Ayoubi, M. Engsvang, A. B. Jensen, Y. Knattrup, J. Kubečka, C. J. Bready, V. R. Fowler, S. E. Harold and O. M. Longsworth, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2023, **13**, e1662–e1675.
42 I. K. Ortega, O. Kupiainen, T. Kurtén, T. Olenius, O. Wilkman, M. J. McGrath, V. Loukonen and H. Vehkamäki, *Atmos. Chem. Phys.*, 2012, **12**, 225–235.
43 H. Henschel, T. Kurtén and H. Vehkamaki, *J. Phys. Chem. A*, 2016, **120**, 1886–1896.
44 H. Wen, T. Huang, C.-Y. Wang, X.-Q. Peng, S. Jiang, Y.-R. Liu and W. Huang, *Atmos. Environ.*, 2018, **191**, 214–226.
45 X.-Q. Peng, Y.-R. Liu, T. Huang, S. Jiang and W. Huang, *Phys. Chem. Chem. Phys.*, 2015, **17**, 9552–9563.
46 R. Dennington, T. A. Keith and J. M. Millam, *GaussView 6.0.16*, Semichem Inc., Shawnee Mission, KS, USA, 2016, pp. 143–150.
47 L. Chen and D. E. Woon, *J. Phys. Chem. A*, 2011, **115**, 5166–5183.
48 J. Wang, M. Dou, X. Wang, B. Gao, T. Zhuang and Z. Ma, *Chemosphere*, 2022, **294**, 133741.
49 D. Chakraborty and P. K. Chattaraj, *J. Comput. Chem.*, 2018, **39**, 151–160.
50 C. L. Firme, N. K. V Monteiro and S. R. B. Silva, *Comput. Theor. Chem.*, 2017, **1111**, 40–49.
51 C. Lefebvre, G. Rubez, H. Khartabil, J.-C. Boisson, J. Contreras-García and E. Hénon, *Phys. Chem. Chem. Phys.*, 2017, **19**, 17928–17936.
52 M. D. Mohammadi, H. Y. Abdullah, H. Louis, E. E. Etim and H. O. Edet, *J. Mol. Liq.*, 2023, **387**, 122621–122630.
53 N. S. Venkataramanan and A. Suvitha, *J. Mol. Graph. Model.*, 2018, **81**, 50–59.
54 S. Sabud, M. Bera and J. Pal, *J. Mol. Model.*, 2023, **29**, 369–372.
55 S. Tiwari and R. Subramanian, *Theor. Chem. Acc.*, 2024, **143**, 1–15.
56 H. R. Leverentz, J. I. Siepmann, D. G. Truhlar, V. Loukonen and H. Vehkamaki, *J. Phys. Chem. A*, 2013, **117**, 3819–3825.
57 N. Bork, L. Du and H. G. Kjaergaard, *J. Phys. Chem. A*, 2014, **118**, 1384–1389.
58 J. Elm, M. Bilde and K. V Mikkelsen, *Phys. Chem. Chem. Phys.*, 2013, **15**, 16442–16445.
59 M. Cavcar, *Anadolu Univ. Turkey*, 2000, **30**, 1–6.
60 Y.-J. Han, Y.-J. Feng, S.-K. Miao, S. Jiang, Y.-R. Liu, C.-Y. Wang, J. Chen, Z.-Q. Wang, T. Huang and J. Li, *Phys. Chem. Chem. Phys.*, 2018, **20**, 25780–25791.
61 W. Xu and R. Zhang, *J. Chem. Phys*, 2018, **148**, 244103–244111.
62 J. Li, Y.-J. Feng, S. Jiang, C.-Y. Wang, Y.-J. Han, C.-X. Xu, H. Wen, T. Huang, Y.-R. Liu and W. Huang, *Atmos. Environ.*, 2019, **219**, 117005–117015.
63 S. Lv, F.-Y. Bai, X.-M. Pan and L. Zhao, *Environ. Chem.*, 2018, **15**, 372–385.
64 M. Torrent-Sucarrat, J. S. Francisco and J. M. Anglada, *J. Am. Chem. Soc.*, 2012, **134**, 20632–20644.
65 N.-T. Phan, K.-H. Kim, Z.-H. Shon, E.-C. Jeon, K. Jung and N.-J. Kim, *Atmos. Environ.*, 2013, **65**, 177–185.
66 D. W. Arlander, D. R. Cronn, J. C. Farmer, F. A. Menzia and H. H. Westberg, *J. Geophys. Res. Atmos.*, 1990, **95**, 16391–16403.
67 E. C. Tuazon, A. M. Winer and J. N. Pitts, *Environ. Sci. Technol.*, 1981, **15**, 1232–1237.
68 W. R. Hartmann, M. O. Andreae and G. Helas, *Physico-Chemical Behav. Atmos. Pollut. Air Pollut. Res. Reports*, 1990, 558–561.
69 N. Kumar, U. C. Kulshrestha, P. Khare, A. Saxena, K. M. Kumari and S. S. Srivastava, *Atmos. Environ.*, 1996, **30**, 3545–3550.
70 H. B. Singh, M. Kanakidou, P. J. Crutzen and D. J. Jacob, *Nature*, 1995, **378**, 50–54.
71 F. Arnold, V. Bürger, B. Droste-Fanke, F. Grimm, A. Krieger, J. Schneider and T. Stilp, *Geophys. Res. Lett.*, 1997, **24**, 3017–3020.
72 D. J. Jacob, B. D. Field, E. M. Jin, I. Bey, Q. Li, J. A. Logan, R. M. Yantosca and H. B. Singh, *J. Geophys. Res. Atmos.*, 2002, **107**, 1–17.
73 D. Grosjean, A. H. Miguel and T. M. Tavares, *Atmos. Environ. Part B. Urban Atmos.*, 1990, **24**, 101–106.
74 M. Possanzini, V. Di Palo and A. Cecinato, *Atmos. Environ.*, 2002, **36**, 3195–3201.
75 K. Toda, S. Yunoki, A. Yanaga, M. Takeuchi, S.-I. Ohira and P. K. Dasgupta, *Environ. Sci. Technol.*, 2014, **48**, 6636–6643.
76 R. Zhang, A. Khalizov, L. Wang, M. Hu and W. Xu, *Chem. Rev.*, 2012, **112**, 1957–2011.
77 N. L. Ambe, O. Holtomo, A. F. Tandong and D. Afungchui, *J. Aerosol Sci.*, 2025, **185**, 106524–106539.
78 J. Chen, S. Jiang, Y.-R. Liu, T. Huang, C.-Y. Wang, S.-K. Miao, Z.-Q. Wang, Y. Zhang and W. Huang, *RSC Adv.*, 2017, **7**, 6374–6388.
79 M. Wang, W. Chen, M. Shao, S. Lu, L. Zeng and M. Hu, *J. Environ. Sci.*, 2015, **28**, 128–136.
80 S. Ni, T.-T. Meng, G.-Q. Huang, Y.-Z. Tang, F.-Y. Bai and Z. Zhao, *J. Phys. Chem. A*, 2023, **127**, 5402–5413.
81 G.-C. An, *Chemosphere*, 2024, **368**, 143684–143693.