AARHUS UNIVERSITY

Dear Editor and Referee 1,

We sincerely thank referee 1 for evaluating our manuscript. However, we believe there is a fundamental misunderstanding of the manuscript's goal and methodology. In this cover letter, we outline the root of this disconnect, followed by detailed point-by-point responses to all specific comments below.

In the point-by-point responses, referee comments are given in ***bold italic***, while responses are given in roman (non-bold, non-italic). Excerpts from the revised manuscript to support our responses are highlighted in yellow. Line and page number are indicated by (L### P#).

We believe the referee's main criticisms stem from a fundamental misalignment regarding the manuscript's objective and its theoretical framework. First, this study is a methodological evaluation assessing how well machine learning models reproduce the long-range collision dynamics relevant to new particle formation. This evaluation is a mandatory prerequisite for accurately modeling large-scale nucleation simulations. Throughout the report, the referee critiques the study for evaluating systems below the critical cluster size, at a single temperature, and at an insufficiently high level of theory. However, these critiques apply to phenomenological kinetic studies rather than this methodological study. Because the underlying potential energy surface does not change with temperature, and the essential physics of monomer collisions must be validated before scaling to critical cluster sizes, exploring these additional avenues, while interesting, evaluates the manuscript against criteria outside its intended methodological scope.

Second, regarding the methodology, the referee approaches our work from the perspective of small-molecule gas-phase kinetics, where unstable excited complexes form and must be energetically evaluated for dissociation via a master equation. Atmospheric particle formation, on the other hand, standardly relies on a macroscopic cluster distribution dynamics formalism. This framework operates on the assumption that dissociation due to collisional excitation is negligible (Elm et al., 2020). This is physically motivated by the large number of internal degrees of freedom in the clusters studied here, which effectively accommodate excess collision energy until background gas thermalizes them. Therefore, the referee's disagreement lies fundamentally with the established theoretical framework of aerosol physics, rather than with our specific study. Given that macroscopic cluster dynamics simulations demonstrate good agreement with atmospheric experiments (e.g., Almeida et al., 2013), we do not believe this disagreement regarding standard field practices renders our study unfit for *Atmospheric Chemistry and Physics*.

This study represents a critical step toward utilizing machine learning interatomic potentials to run nucleation simulations at near-*ab initio* accuracy. Its focus on aerosol particle formation utilizing the research activity of machine learning aligns exactly with the Aims and Scope of *Atmospheric Chemistry and Physics*. While we have revised the manuscript to ensure its objectives and methodological framework is clear to future readers, the criticisms offered here do not accurately reflect the goals of this study or its situation within the broader atmospheric chemistry modeling community. As such, we respectfully request the Editor to carefully review our responses and consider our revised manuscript for publication in *Atmospheric Chemistry and Physics*.

Best regards,

Ivo Neefjes

Point-by-point response

Summary

Calculating the capture rate requires the long-range potential energy surface (PES) to be accurately defined. This paper uses molecular dynamics simulations (trajectory calculations) to explore the capture rate coefficient for three systems:

- $H_2SO_4 + H_2SO_4$***; binding energy*** $66\ kJ\,mol^{-1}$
- $H_2SO_4 + NH(CH_3)_2$***; binding energy*** $71\ kJ\,mol^{-1}$
- $H_2SO_4 + HSO_4{}^-$***: binding energy*** $125\ kJ\,mol^{-1}$

2 Machine learning architectures have been used to explore the PES for these systems based on data from molecular dynamics simulations at 2 levels of theory, semi-empirical and DFT.

Comparison is made between the Machine Learning methods and how things need to be improved to calculated accurate capture rates. All the capture rate coefficients are fast at 300 K.

We thank the referee for this summary of our study.

General criticism

This paper is mainly about the method of calculating the capture rate coefficient rather than the importance of the capture rate coefficients under consideration. If I wanted to calculate these capture rate coefficients, would it not be easier to do a high-level PES, much higher level than in this study, and then use capture rate theory to calculate the capture rate coefficient?

We thank the referee for their assessment and agree that this manuscript is primarily methodological. Its main objective is to investigate the extent to which machine learning algorithms can reproduce the attractive long-range interactions that govern sticking collisions and subsequent particle formation. We demonstrate that while all evaluated methods reproduce the free energy profile along the center-of-mass distance, only AIMNet2 successfully reproduces the collision rate coefficients due to its inclusion of long-range Coulombic and dispersion interactions.

While mapping a static, high-level potential energy surface (PES) might be feasible for a simple dimer collision, simulating the dynamics of spontaneous nucleation involves tens of flexible precursor molecules forming clusters of unpredictable sizes and configurations. Mapping a pre-defined PES for such a highly dimensional, multi-molecular system is computationally unfeasible. Molecular dynamics (MD) overcomes this bottleneck by sampling the PES on-the-fly, providing a cost-effective approach for unbiased cluster simulations.

Additionally, while methods like long-range transition state theory can provide collision rate coefficients at high levels of theory, they cannot account for explicit collision dynamics. Molecular dynamics, in contrast, allows us to track the time-evolution of the system, capturing phenomena such as proton transfers, preferred orientations, anharmonic vibrations, and interconversions between low-lying free energy minima.

Furthermore, as demonstrated in this study, machine learning interatomic potentials approach the accuracy of their underlying quantum chemical training data at a fraction of the computational cost. While we utilized ωB97X-3c due to its proven performance in recent benchmarks (e.g., Neefjes et al., 2026), the methodology is highly adaptable. Models can be readily trained at even higher levels of theory by performing nuclear gradient calculations on structures sampled at the GFN1-xTB level. By validating the long-range accuracy of these models, this manuscript lays the necessary groundwork for large-scale nucleation simulations, where particle formation rates and collision rate coefficients are derived directly from the MD trajectories using methods such as the mean first-passage time (MFPT; Wedekind et al. (2007)) or Yasuoka-Matsumoto (Yasuoka and Matsumoto, 1998) approaches.

We have revised the manuscript to highlight its methodological nature and objective:

In this methodological study, we assess the ability of the PaiNN and AIMNet2 architectures to describe collisions governed by long-range interactions. We sampled training configurations using GFN1-xTB dynamics, subsequently computing energies and forces at both the GFN1-xTB and ωB97X-3c levels. Additionally, we employed delta-learning to upscale GFN1-xTB simulations with PaiNN corrections to the ωB97X-3c level of theory. Since sulfuric acid is a key contributor to particle formation (Sipilä et al., 2010), we studied the sulfuric acid dimer, the sulfuric acid–dimethylamine system (to investigate stabilizing proton transfers), and the sulfuric acid–bisulfate system (to examine strong ionic long-range contributions). Following hyperparameter tuning, we evaluated model performance by comparing electronic energy and force predictions against independent test sets. Furthermore, we calculated the potential of mean force (PMF) through umbrella sampling to compare against GFN1-xTB reference data. Finally, we derived collision rate coefficients from MD collision trajectory simulations to evaluate the long-range dynamics of the models and examine how they vary across different levels of theory. By validating these ML models in the context of atmospheric particle formation, this study establishes the necessary groundwork for large-scale MD simulations in this domain.

L80 P3

An even more important point is that these capture rate coefficients are not the values for atmospheric chemistry. The calculated capture rate coefficient is for infinite pressure and I'm very confident that the systems explored here are nowhere near the high-pressure limit. The value required for atmospheric chemistry requires a Master Equation calculation using the derived PES for the system. It is also likely, the adducts are too weakly bound for them to not re-dissociate back to reagents at 300 K. Therefore, the equilibrium coefficient for these systems might also be important. You have not provided the atmospherically relevant rate coefficient and the equilibrium coefficient for the system.

The calculated collision rate coefficients are specifically intended for use in macroscopic cluster distribution models, such as those relying on birth-death equations (e.g., the Atmospheric Cluster Dynamics Code; McGrath et al. (2012)). In the birth-death framework, the collision process is explicitly separated from the evaporation process, operating under the assumption that nascent clusters live long enough to be thermalized by the background gas (Elm et al., 2020).

This assumption is physically justified by the size of the molecules likely involved in atmospheric new particle formation. The probability of fragmentation due to imperfect thermalization decreases rapidly as the number of vibrational modes ($3N-6$, where N is the number of atoms) increases. Because the relatively large, strongly bound dimers studied here possess a large number of internal degrees of freedom to distribute the collision energy over. Consequently, the assumption of thermalization without dissociation at pressures encountered in the lower atmosphere (0.1 to 1 atm) is reasonable and represents standard practice in aerosol physics.

While small gas-phase radicals that form unstable complexes may dissociate before thermalization, thus requiring a master equation to calculate a pressure-dependent effective rate, this is not the kinetic regime of the aerosol precursors modeled here. Furthermore, evaporation rate coefficients and equilibrium constants for these systems can be, and frequently are, obtained via detailed balance using binding free energies from static quantum chemistry calculations. However, providing survival probabilities for these specific dimers is not the objective of this methodological study, which focuses strictly on evaluating whether machine learning models can accurately capture the long-range collision kinetics required for large-scale molecular dynamics simulations. We therefore respectfully maintain that we have provided the exact rate coefficients relevant to our stated objectives.

We have revised the manuscript to highlight the cluster distribution dynamics framework used and its assumption of a pressure-independent collision rate coefficient:

> In cluster distribution dynamics models, such as the Atmospheric Cluster Dynamics Code (ACDC) (McGrath et al., 2012), cluster-forming collisions and cluster-removing evaporations are treated as independent processes, assuming that dissociation prior to thermalization from collisional excitation is negligible (Elm et al., 2020). This yields a pressure-independent collision rate coefficient, which represents the frequency of collisions per unit concentration.

L66 P3

The above two points raises the question of if ACP is the right journal for this paper. I do not think it is.

The referee's critique appears to stem from the framework of small-molecule fundamental gas kinetics, where master equation simulations are strictly required to track the dissociation of excited complexes. However, in the field of atmospheric new particle formation (NPF), this micro-kinetic framework is conventionally not applied, as large cluster precursors are physically understood to possess sufficient internal degrees of freedom to accommodate excitation energy.

By rejecting this assumption, the referee is inadvertently challenging the standard, widely accepted practice of modeling cluster distribution dynamics via macroscopic birth-death equations. Addressing the validity of foundational NPF kinetic frameworks falls far outside the scope of this study, which aims specifically to evaluate the long-range physical accuracy of machine learning potentials, with the collision rate coefficient serving as the primary metric of success. This evaluation is a necessary preparatory step for large-scale, machine-learning-driven nucleation studies currently being developed within our group and the broader computational atmospheric chemistry community.

Atmospheric Chemistry and Physics has a history of publishing impactful research on both cluster distribution dynamics modeling (e.g., McGrath et al., 2012; Ortega et al., 2012) and the calculation of fundamental collision rate coefficients (e.g., Halonen et al., 2019; Neefjes et al., 2022; Yang et al., 2023). Because our work directly provides validated computational tools designed specifically to advance atmospheric nucleation modeling, we strongly believe it is suitable for publication in *Atmospheric Chemistry and Physics*.

Additionally, can you really explore the long-range potential with such a low level of theory?

Yes, because the physics of long-range molecular interactions are fundamentally different from short-range chemical bonding. At long distances, the potential energy surface (PES) is dominated by asymptotic electrostatic forces (such as dipole–dipole and ion–dipole interactions) and dispersion. These macroscopic molecular properties, specifically dipole moments and polarizabilities, converge very rapidly with the level of theory. Methods like ωB97X-3c and even semi-empirical GFN1-xTB are well-parameterized to reproduce these asymptotic electrostatic properties.

Furthermore, the adequacy of describing long-range potentials using these fundamental properties is strongly supported by decades of experimental evidence (e.g., Tsikritea et al., 2022). Experimental mass spectrometry and selected ion flow tube measurements of ion-molecule reaction rates consistently match the theoretical collision rate limits predicted by classical capture theories (such as the Su-Chesnavich model; Su and Chesnavich (1982)). These theoretical models rely entirely on the exact same basic long-range electrostatic parameters (polarizability and dipole moment) captured by our underlying computational methods.

Is it easy to calculate the capture rate coefficient over a range of temperatures? Only 300 K is presented.

Calculating collision rate coefficients across a range of temperatures is computationally straightforward, provided the underlying machine learning interatomic potential is properly trained for those conditions.

As temperature increases, thermal motion allows the system to access higher-energy conformational states and shorter intermolecular collision distances. If a machine learning (ML) model is trained exclusively on lower-temperature data, simulating higher temperatures risks pushing the model into extrapolative configurations where it may predict inaccurate forces. Therefore, exploring a wide temperature range requires ensuring the training set is adequately augmented with data sampled at or above the maximum target temperature.

However, conducting a temperature-dependent kinetic analysis falls outside the methodological scope of this study. Our primary objective is to evaluate the fundamental ability of different ML architectures to capture long-range interactions, such as electrostatics and dispersion. The accuracy of the predicted potential energy surface (PES) at long distances is an intrinsic property of the ML architecture and its training and independent of the macroscopic simulation temperature. Because the underlying long-range PES does not change with temperature, benchmarking the models' physical accuracy at a single atmospherically relevant temperature (300 K) is sufficient to fulfill the objectives of this study.

Overall

For the reasons I outline above I do not think ACP is the right journal for this paper; these reactions are going to be pressure dependent and a long way from the capture rate coefficient at atmospheric pressure. To make this an ACP paper more work is required in order to calculate the atmospherically relevant rate coefficients, including the equilibrium coefficient.

As detailed in our previous responses, we respectfully maintain that the collisions of these aerosol precursors operate at or near the high-pressure limit at 1 atm, rendering the pressure-independent collision rate coefficient the relevant metric for atmospheric macroscopic cluster models.

Beyond the kinetic framework, this manuscript addresses a critical and highly specific bottleneck in atmospheric chemistry: the accurate simulation of new particle formation (NPF). Historically, large-scale molecular dynamics (MD) simulations of atmospheric nucleation have relied on classical force fields due to the prohibitive computational cost of *ab initio* MD. However, classical potentials fundamentally fail to model dynamic short-range quantum effects, such as the proton transfers that are essential for the thermodynamic stabilization of atmospheric acid–base clusters.

Recent advances in machine learning interatomic potentials (MLIPs) now allow us to run simulations that approach *ab initio* accuracy, explicitly capturing these atmospheric proton transfers, while maintaining accessible computational costs. However, before the atmospheric community can safely transition to machine-learning-driven nucleation simulations, it is essential to rigorously validate that these models do not sacrifice long-range physical accuracy (electrostatics and dispersion) to achieve short-range chemical accuracy. By evaluating this trade-off, this study represents an essential methodological step toward accurately simulating atmospheric nucleation systems.

Finally, regarding the manuscript's suitability, *Atmospheric Chemistry and Physics* explicitly lists machine learning as a research activity of interest within its Aims and Scope. Furthermore, our methodology directly provides the foundational kinetic inputs necessary for the macroscopic cluster distribution models (e.g., the Atmospheric Cluster Dynamics Code; (McGrath et al., 2012)) that have been extensively developed and published within *Atmospheric Chemistry and Physics* (e.g., Tuovinen et al., 2022; Shen et al., 2024). Because this study directly equips the atmospheric modeling community with validated computational tools to advance NPF research, we strongly believe it is suitable for publication in *Atmospheric Chemistry and Physics.*

Line-by-line

Are the collisions you are considering in this study important for NPF? I think it is later collisions that are rate determining, not the first step collisions considered in the present system.

The collision systems in this study were selected based on their fundamental importance to atmospheric new particle formation (NPF). Sulfuric acid is widely considered an essential driver of NPF due to its low volatility and its capacity to form strongly bound acid–base clusters. Among atmospheric amines, dimethylamine is recognized as the strongest candidate for such clustering, representing an optimal balance between atmospheric concentration, high basicity, and minimal steric hindrance. Furthermore, the bisulfate ion is a highly relevant component for modeling ion-induced particle formation pathways.

We acknowledge that, according to classical nucleation theory, the overall particle formation rate is governed by a critical cluster size, the threshold at which collisions begin to outweigh evaporation. Cluster distribution dynamic simulations (Olenius et al., 2013) and experiments (Kürten et al., 2014), however, show that atmospheric particle-forming systems, such as sulfuric acid–dimethylamine, do not have a critical cluster size but are collision controlled. Hence, the first collision forming the dimer is essentially the critical cluster.

Furthermore, reaching the critical cluster size is not the objective of this methodological study. Our aim is to evaluate, develop, and validate the computational tools necessary to accurately model the dynamics of nucleation using machine learning interatomic potentials. Because the entire macroscopic nucleation sequence is built upon a series of discrete collision events, accurately and efficiently describing the foundational monomer–monomer interactions is a mandatory first step. By benchmarking the machine learning models' ability to capture the long-range physics of these initial collisions, we are validating the tools required to eventually simulate the full pathway up to and beyond the critical cluster size.

The capture rate coefficient is appropriate at the high-pressure limit. A lot of associations—including the systems in this paper—are going to be pressure dependent at atmospheric pressure and will be much smaller than the capture limit. A Master Equation calculation is required.

As mentioned above, the framework of macroscopic cluster distribution dynamics assumes that the particle-forming precursors investigated in this study experience negligible dissociation due to collisional excitation. This is physically grounded in the structural complexity of the clusters. For instance, a sulfuric acid–dimethylamine dimer is relatively strongly bound and possesses 45 internal degrees of freedom across which the excess collision energy can be shared.

Consequently, the assumption of negligible dissociation (high-pressure limit) is not an approximation unique to our manuscript, but rather a foundational principle underpinning the established kinetic theory driving cluster dynamics modeling throughout the atmospheric sciences community.

What sort of errors. Do you just mean rate coefficients?

For the mentioned errors, we refer to a broader range of inaccuracies inherent to the semi-empirical GFN1-xTB method when applied to complex hydrogen-bonded systems, rather than errors in the collision rate coefficients themselves. Prominent examples of these limitations include both quantitative thermodynamic inaccuracies and qualitative structural misidentifications.

While GFN1-xTB offers exceptional computational efficiency, it can exhibit significant quantitative errors in binding energies. For instance, in our previous benchmarking study on hydrated clusters (Neefjes et al., 2026), GFN1-xTB yielded a mean electronic energy error of approximately 5 kcal/mol for various $(\text{acid/base})_{0-2}(\text{water})_{0-5}$ systems when compared to a highly accurate DLPNO-CCSD(T0)/aug-cc-pVTZ reference. For larger aerosol precursors, such as a cluster containing 6 sulfuric acid, 6 ammonia, and 10 water molecules, these quantitative electronic energy errors can exceed 40 kcal/mol against the same reference.

Furthermore, as an example of qualitative errors, GFN1-xTB frequently predicts a different lowest-energy conformational isomer (global minimum) compared to density functional theory methods, such as ωB97X-3c (Kubečka et al., 2019). Accurate evaporation rates depend fundamentally on correct thermodynamic stabilities and structural conformations.

We revised the manuscript to explicitly mention the type of errors we are alluding to:

While semi-empirical quantum chemistry methods such as GFN1-xTB (Grimme et al., 2017) can model bond-breaking, they can exhibit significant errors for complex hydrogen-bonded systems, including quantitative inaccuracies in binding energies (Neefjes et al., 2026) and qualitative misidentifications of lowest-energy cluster configurations (Kubečka et al., 2019).

L38 P2

Can MD simulations be run long enough to be relevant for more atmospheric reactions? Are you stuck at pico-second processes or shorter? For instance, adding H_2O to sulfuric acid dimer is important and the question is how many waters are added (together with further reactions) can MD tackle this problem, or is the timescale to long for MD? If you have an accurate PES from a much higher-level QM calculation, can this PES be used to calculate an accurate rate coefficient via capture rate theory?

We thank the referee for raising these practical considerations. With recent advances in machine learning interatomic potentials (MLIPs), the accessible timescale for *ab initio*-quality molecular dynamics (MD) has been significantly extended. While actual *ab initio* MD is often restricted to the picoseconds regime, MLIP-driven MD can simulate systems of tens to hundreds of molecules over durations of several to tens of nanoseconds.

Regarding the specific example of sequential water addition to a sulfuric acid dimer: due to the low concentration of particle-forming precursors compared to background gas in the actual atmosphere, simulating a realistic macroscopic slice of the atmosphere to observe these slow, sequential collisions in real-time is indeed impossible. However, MD studies routinely overcome this timescale limitation using well-established methodologies. Direct nucleation simulations employ highly supersaturated conditions to observe clustering events within nanosecond timescales. The resulting particle formation rates can then be rigorously extrapolated down to atmospherically relevant concentrations.

Beyond direct nucleation, targeted MD setups, such as the isolated collision trajectory simulations utilized in this study, allow us to extract precise collision rate coefficients and capture complex dynamical effects that are often neglected in static theoretical frameworks.

Finally, regarding the use of a high-level quantum mechanical potential energy surface (QM PES) with capture theory: as discussed in our previous response, while a static QM PES might be feasible for a simple monomer–monomer collision, mapping the multi-dimensional QM PES required for sequential additions (such as a dimer interacting with water molecules) quickly becomes computationally infeasible. MD bypasses this dimensionality explosion by evaluating the PES on-the-fly exactly where the dynamics visits.

The systems considered can be calculated at a much higher level via standard ab initio techniques. Would it not be better to do this as well, in order to really assign the accuracy of the present calculations?

The aim of this study is not to benchmark the accuracy of the employed quantum chemistry methods. In several benchmarks, ωB97X-3c has been found to provide accurate energetics and configurations compared to higher level ab initio techniques (Jensen and Elm, 2024; Neefjes et al., 2026). Our concern is if the chosen quantum chemistry method is accurately reproduced by the machine learning model.

We have, however, revised the manuscript to highlight that the same methodology can be used to train machine learning models at higher levels of theory:

> We note that while ωB97X-3c is used in this study, atomic forces can be calculated using any quantum chemistry method on the GFN1-xTB structures to obtain a training set at that level of theory.

L164 P6

This is not absolute chemical accuracy. What are you comparing against to state errors below 0.1 kcal mol^{-1}?

Defining “chemical accuracy” as an energy error of less than 1 kcal/mol is standard nomenclature within the computational chemistry community. This was likely established by John Pople. As stated in his Nobel lecture (Pople, 1999): “As the model becomes quantitative, the target should be that data is reproduced and predicted within experimental accuracy. For energies, such as heats of formation or ionization potentials, a global accuracy of 1 kcal/mole would be appropriate.”

Beyond its comparison with experimental uncertainty, this 1 kcal/mol threshold also has implications for kinetic and thermodynamic modeling. Free energies dictate equilibrium constants and evaporation rate coefficients via exponential relationships. At 300 K, an energy error of approximately 1.37 kcal/mol translates to a full order-of-magnitude error in these exponential quantities. Therefore, an energy error less than 1 kcal/mol ensures that rate and equilibrium errors are restricted to less than a factor of ten.

Regarding the specific values reported in our manuscript, the stated errors (below 0.1 kcal/mol) do not represent absolute accuracy compared to experimental "truth". Rather, they quantify how well the machine learning interatomic potential reproduces the potential energy surface of the specific quantum chemical level of theory (ωB97X-3c) it was trained on. An error below 0.1 kcal/mol confirms that the machine learning architecture introduces negligible additional error, effectively preserving the fundamental chemical accuracy of the underlying training data.

The highest observed RMSE is 0.20 $\mathrm{kcal\,mol^{-1}}$ for AIMNet2 applied to the H_2SO_4–$HSO_4{}^-$ system, which is five times lower than the standard threshold for chemical accuracy (1 $\mathrm{kcal\,mol^{-1}}$). In other words, the reproduction error introduced by the machine learning models is negligible compared to generally accepted error margins in computational chemistry.

L334 P14

References

Almeida, J., Schobesberger, S., Kürten, A., Ortega, I. K., Kupiainen-Määttä, O., Praplan, A. P., Adamov, A., Amorim, A., Bianchi, F., Breitenlechner, M., David, A., Dommen, J., Donahue, N. M., Downard, A., Dunne, E., Duplissy, J., Ehrhart, S., Flagan, R. C., Franchin, A., Guida, R., Hakala, J., Hansel, A., Heinritzi, M., Henschel, H., Jokinen, T., Junninen, H., Kajos, M., Kangasluoma, J., Keskinen, H., Kupc, A., Kurtén, T., Kvashin, A. N., Laaksonen, A., Lehtipalo, K., Leiminger, M., Leppä, J., Loukonen, V., Makhmutov, V., Mathot, S., McGrath, M. J., Nieminen, T., Olenius, T., Onnela, A., Petäjä, T., Riccobono, F., Riipinen, I., Rissanen, M., Rondo, L., Ruuskanen, T., Santos, F. D., Sarnela, N., Schallhart, S., Schnitzhofer, R., Seinfeld, J. H., Simon, M., Sipilä, M., Stozhkov, Y., Stratmann, F., Tomé, A., Tröstl, J., Tsagkogeorgas, G., Vaattovaara, P., Viisanen, Y., Virtanen, A., Vrtala, A., Wagner, P. E., Weingartner, E., Wex, H., Williamson, C., Wimmer, D., Ye, P., Yli-Juuti, T., Carslaw, K. S., Kulmala, M., Curtius, J., Baltensperger, U., Worsnop, D. R., Vehkamäki, H., and Kirkby, J.: Molecular understanding of sulphuric acid–amine particle nucleation in the atmosphere, Nature, 502, 359–363, https://doi.org/10.1038/nature12663, 2013.

Elm, J., Kubečka, J., Besel, V., Jääskeläinen, M. J., Halonen, R., Kurtén, T., and Vehkamäki, H.: Modeling the formation and growth of atmospheric molecular clusters: A review, J. Aerosol Sci., 149, 105 621, https://doi.org/10.1016/j.jaerosci.2020.105621, 2020.

Halonen, R., Zapadinsky, E., Kurtén, T., Vehkamäki, H., and Reischl, B.: Rate enhancement in collisions of sulfuric acid molecules due to long-range intermolecular forces, Atmos. Chem. Phys., 19, 13 355–13 366, https://doi.org/10.5194/acp-19-13355-2019, 2019.

Jensen, A. B. and Elm, J.: Massive Assessment of the Geometries of Atmospheric Molecular Clusters, J. Chem. Theory Comput., 20, 8549–8558, https://doi.org/10.1021/acs.jctc.4c01046, 2024.

Kubečka, J., Besel, V., Kurtén, T., Myllys, N., and Vehkamäki, H.: Configurational Sampling of Noncovalent (Atmospheric) Molecular Clusters: Sulfuric Acid and Guanidine, J. Phys. Chem. A, 123, 6022–6033, https://doi.org/10.1021/acs.jpca.9b03853, 2019.

Kürten, A., Jokinen, T., Simon, M., Sipilä, M., Sarnela, N., Junninen, H., Adamov, A., Almeida, J., Amorim, A., Bianchi, F., Breitenlechner, M., Dommen, J., Donahue, N. M., Duplissy, J., Ehrhart, S., Flagan, R. C., Franchin, A., Hakala, J., Hansel, A., Heinritzi, M., Hutterli, M., Kangasluoma, J., Kirkby, J., Laaksonen, A., Lehtipalo, K., Leiminger, M., Makhmutov, V., Mathot, S., Onnela, A., Petäjä, T., Praplan, A. P., Riccobono, F., Rissanen, M. P., Rondo, L., Schobesberger, S., Seinfeld, J. H., Steiner, G., Tomé, A., Tröstl, J., Winkler, P. M., Williamson, C., Wimmer, D., Ye, P., Baltensperger, U., Carslaw, K. S., Kulmala, M., Worsnop, D. R., and Curtius, J.: Neutral molecular cluster formation of sulfuric acid–dimethylamine observed in real time under atmospheric conditions, Proc. Natl. Acad. Sci. U.S.A., 111, 15 019–15 024, https://doi.org/10.1073/pnas.1404853111, 2014.

McGrath, M. J., Olenius, T., Ortega, I. K., Loukonen, V., Paasonen, P., Kurtén, T., Kulmala, M., and Vehkamäki, H.: Atmospheric Cluster Dynamics Code: a flexible method for solution of the birth-death equations, Atmos. Chem. Phys., 12, 2345–2355, https://doi.org/10.5194/acp-12-2345-2012, 2012.

Neefjes, I., Halonen, R., Vehkamäki, H., and Reischl, B.: Modeling approaches for atmospheric ion–dipole collisions: all-atom trajectory simulations and central field methods, Atmos. Chem. Phys., 22, 11 155–11 172, https://doi.org/10.5194/acp-22-11155-2022, 2022.

Neefjes, I., Knattrup, Y., Wu, H., Trolle, G. B., Elm, J., and Kubečka, J.: Thermodynamic benchmarking of hydrated atmospheric clusters in early particle formation, Aerosol Research, 4, 1–22, https://doi.org/10.5194/ar-4-1-2026, 2026.

Olenius, T., Kupiainen-Määttä, O., Ortega, I. K., Kurtén, T., and Vehkamäki, H.: Free energy barrier in the growth of sulfuric acid–ammonia and sulfuric acid–dimethylamine clusters, J. Chem. Phys., 139, 084 312, https://doi.org/10.1063/1.4819024, 2013.

Ortega, I. K., Kupiainen, O., Kurtén, T., Olenius, T., Wilkman, O., McGrath, M. J., Loukonen, V., and Vehkamäki, H.: From quantum chemical formation free energies to evaporation rates, Atmos. Chem. Phys., 12, 225–235, https://doi.org/10.5194/acp-12-225-2012, 2012.

Pople, J. A.: Nobel Lecture: Quantum chemical models, Rev. Mod. Phys., 71, 1267–1274, https://doi.org/10.1103/RevModPhys.71.1267, 1999.

Shen, J., Zhao, B., Wang, S., Ning, A., Li, Y., Cai, R., Gao, D., Chu, B., Gao, Y., Shrivastava, M., Jiang, J., Zhang, X., and He, H.: Cluster-dynamics-based parameterization for sulfuric acid–dimethylamine nucleation: comparison and selection through box and three-dimensional modeling, Atmos. Chem. Phys., 24, 10 261–10 278, https://doi.org/10.5194/acp-24-10261-2024, 2024.

Su, T. and Chesnavich, W. J.: Parametrization of the ion–polar molecule collision rate constant by trajectory calculations, J. Chem. Phys., 76, 5183–5185, https://doi.org/10.1063/1.442828, 1982.

Tsikritea, A., Diprose, J. A., Softley, T. P., and Heazlewood, B. R.: Capture theory models:

An overview of their development, experimental verification, and applications to ion–molecule reactions, J. Chem. Phys., 157, 060 901, https://doi.org/10.1063/5.0098552, 2022.

Tuovinen, S., Cai, R., Kerminen, V.-M., Jiang, J., Yan, C., Kulmala, M., and Kontkanen, J.: Survival probabilities of atmospheric particles: comparison based on theory, cluster population simulations, and observations in Beijing, Atmos. Chem. Phys., 22, 15 071–15 091, https://doi.org/10.5194/acp-22-15071-2022, 2022.

Wedekind, J., Strey, R., and Reguera, D.: New method to analyze simulations of activated processes, J. Chem. Phys., 126, 134 103, https://doi.org/10.1063/1.2713401, 2007.

Yang, H., Neefjes, I., Tikkanen, V., Kubečka, J., Kurtén, T., Vehkamäki, H., and Reischl, B.: Collision-sticking rates of acid–base clusters in the gas phase determined from atomistic simulation and a novel analytical interacting hard-sphere model, Atmos. Chem. Phys., 23, 5993–6009, https://doi.org/10.5194/acp-23-5993-2023, 2023.

Yasuoka, K. and Matsumoto, M.: Molecular dynamics of homogeneous nucleation in the vapor phase. I. Lennard-Jones fluid, J. Chem. Phys., 109, 8451–8462, https://doi.org/10.1063/1.477509, 1998.