Supplementary material for

Dynamics of molecular collisions in air and its mean free path

Dimitrios G. Tsalikis,[1] Vlasis G. Mavrantzas,[1,2] and Sotiris E. Pratsinis[1,a)]

[1]Particle Technology Laboratory, Department of Mechanical and Process Engineering, Institute of Energy & Process Engineering, ETH Zürich, Zürich CH-8092, Switzerland

[2]Department of Chemical Engineering, University of Patras & FORTH-ICE/HT, 26504 Patras, Greece

[a)]Author to whom correspondence should be addressed: pratsinis@ethz.ch

S1. EXEMPLARY IMPACT OF MD ON AEROSOL NANOPARTICLE FORMATION AND GROWTH

In aerosol synthesis of materials Blaisten-Barojas and Zachariah[20] first showed by MD that upon coalescence of tiny molecular clusters of silicon, single Si atoms evaporate due to the heat generated by such fusion. Violi[21] first elucidated soot formation (inception) from first principles, consistent with experimental data, by combining MD with Monte Carlo simulations. In nanoparticle science, MD has helped elucidate how surface diffusion dominates the early stages of coalescence (fusion) of TiO_2[22] and Ag[23] or how Ag can occupy preferentially the surface of its nanoalloys with Au,[24] leading to biocompatible Ag-Au nanoparticles.[25] More recently, the method was used to explain such core-shell segregation of 45 bimetallic combinations[26] while reactive MD simulations elucidated, for the first time, soot inception and coalescence as a function of cluster size and temperature.[27] By uncovering the underlying mechanisms behind natural phenomena and processes utilizing atomic-level input, MD holds the merit of dramatically accelerating process scale-up and innovation.[28]

S2. SIMULATION SPECIFICS

Force Field Parameters

Table S1. Functional form and parameters specifying bonded and non-bonded interactions in the FA model (Ref. 32).

Bond	Potential	
	l_0 (Å)	
N-N	1.094	
O-O	1.208	
Lennard-Jones	$U_{LJ}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$	
	ε (kcal/mol)	σ (Å)
N-N	0.07232	3.32
O-O	0.10330	2.99
N-O	0.086433	3.155

Table S2. Characteristic threshold distance r_m for collisions by the hard-sphere (HS), Lennard-Jones (LJ) and fully atomistic (FA) models. The r_m refers to entire nitrogen or oxygen molecules for the HS and LJ models while for the FA, r_m refers to nitrogen or oxygen atoms.

	r_m (Å)		
	nitrogen – nitrogen	oxygen – oxygen	oxygen - nitrogen
HS model	3.798	3.467	3.6325
LJ model	4.263	3.892	4.077
FA model	3.726	3.356	3.541

Selection of Ensemble

Examination of selected atomistic trajectories from simulations in the NPT ensemble with the three MD models revealed that air molecules between two successive collisions do not move in a rectilinear path, but as Figure S1a shows their path exhibits many perturbations (zig-zags). This unexpected observation was traced to the dynamics of the barostat, since it is caused by the coupling between

volume fluctuations and equations of motion (i.e., the affine transformation of the atomistic positions during volume fluctuations to control pressure at its prescribed value) in the NPT ensemble.

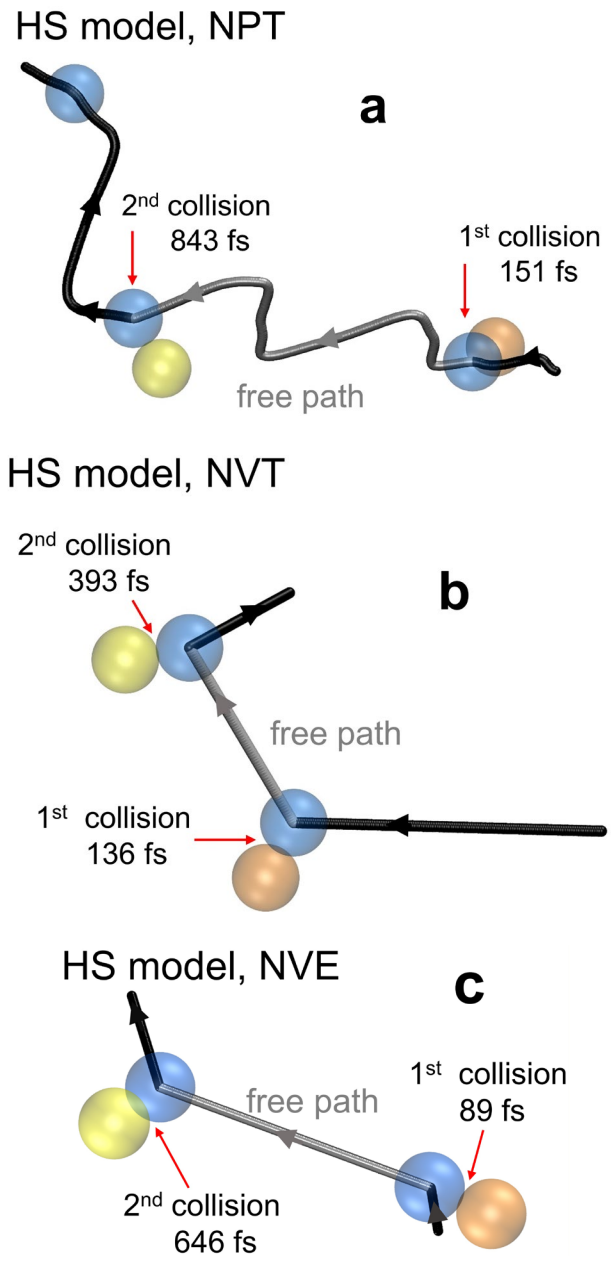

Figure S1. Typical trajectories of a molecule (blue sphere) before and after two successive collisions according to the HS model (equivalent to classic kinetic theory) in the course of an: a) NPT, b) NVT and c) NVE MD simulation. It illustrates the adverse impact (zig-zags instead of straight lines) of the barostat (NPT) on the trajectory of individual molecules between collisions.

In contrast, Figures S1b and S1c show that molecular trajectories between collisions from NVT and NVE simulations, respectively, are straight lines, which indirectly implies that, in contrast to the barostat,

the thermostat does not perturb the molecular motion. In turn, this implies that results for the mean free path from NVT and NVE MD simulations (at the same temperature and density conditions) should be identical, which is indeed the case (see Figure S8 in Sec. S9 of supplementary material). Nevertheless, and to avoid even the tiniest effect of the thermostat, all numerical results, schematic representations, and videos reported here have been extracted from microcanonical ensemble MD trajectories (NVE), at mass density of ρ = 1.177 kg/m^3 (corresponding to T= 300 K and P = 1 atm), unless otherwise noted.

S3. THEORETICAL MODELS FOR THE GAS MEAN FREE PATH

Hard-sphere model

From the kinetic theory (hard-sphere model), the distance of closest approach, r_m, for two spherical molecules is specified by solving[3]

$$\frac{1}{2}\mu_{ij}v_r^2 - u(r) - \frac{\mu_{ij}v_r^2 b^2}{2r^2} = 0 \tag{S1}$$

for r, where μ_{ij} denotes the reduced mass of the two colliding molecules ($\mu_{ij} = \frac{m_i m_j}{m_i + m_j}$ where m_i is the mass of molecule i and m_j that of j). The deflection angle is[3]

$$\chi(b, v_r) = \pi - 2b\int_{r_m}^{\infty} \frac{dr}{r^2\left[1 - \frac{2u(r)}{\mu v_r^2} - \frac{b^2}{r^2}\right]^{1/2}}\,. \tag{S2}$$

From that angle, the collision integrals

$$\Omega^{(l,s)} = \left(\frac{2k_B T}{\mu}\right)^{1/2} \int_0^{\infty}\int_0^{\infty} e^{-\gamma^2}\gamma^{2s+3}\left(1 - \cos^l\chi\right) b\,\mathrm{d}b\,\mathrm{d}\gamma \tag{S3}$$

with $\gamma = \frac{\frac{1}{2}\mu v_r^2}{k_B T}$ are computed to obtain analytical expressions for the transport coefficients. For example, the coefficient of viscosity to the first-order term in the Chapman-Enskog expansion involves the $\Omega^{(2,2)}$ collision integral and is given[3] by

$$\eta = \frac{5k_B T}{8\pi^{1/2}\int_0^\infty e^{-\gamma^2}\gamma^7\left[\int_0^\infty \sin^2 x\, b\, \mathrm{d}b\right]\mathrm{d}\gamma} \quad \text{(S4)}$$

For collisions between molecules of the same type *i*, the collision density from kinetic theory is[3,29,54]

$$Z_{ii} = \frac{1}{\sqrt{2}}\pi\sigma_i^2\,\bar{c}_i\,n_i^2 \quad \text{(S5)}$$

where n_i the number density and $\bar{c}_i$ the mean velocity of *i* molecules $\bar{c}_i = \sqrt{8k_\mathrm{B}T/\pi m_i}$ where k_B is the Boltzmann constant. For collisions between unlike molecules *i* and *j*, the collision density is[3]

$$Z_{ij} = \pi\sigma_{ij}^2\,\bar{c}\,n_i\,n_j \quad \text{(S6)}$$

where now $\bar{c} = \sqrt{8k_\mathrm{B}T/\pi\mu_{ij}}$.

The mean free path can be obtained from the mean flight time between collisions assuming an average velocity for the gas molecules:[29,30,54]

$$\lambda = \frac{\bar{c}}{Z_i} = \sqrt{\frac{8k_B T}{\pi m}}\cdot n\cdot\frac{1}{2Z_{ii}} \quad \text{(S7)}$$

where *m* denotes the mass and *n* the number density of gas molecules and Z_i is the total number of collisions of a single gas molecule per unit time (Z_{ii} is the total number of collisions of gas molecules per unit time per unit volume (the collision density)). By substituting into Equation (S7) the collision density Z_{ii} from Equation (S5), the following alternative expression for λ (assuming also an ideal gas behavior) is obtained:[3, 29]

$$\lambda = \frac{1}{\sqrt{2}}\cdot\frac{1}{\pi n\sigma^2} = \frac{1}{\sqrt{2}}\cdot\frac{k_B T}{\pi\sigma^2 P} \quad \text{(S8)}$$

or, more precisely for a binary mixture with nearly equal-mass molecules:[3,29]

$$\lambda_1 = \frac{1}{\sqrt{2}}\cdot\left[\frac{4}{4\pi n_1\sigma_1^2 + \pi n_2(\sigma_1+\sigma_2)^2}\right] \text{ and } \lambda_2 = \frac{1}{\sqrt{2}}\cdot\left[\frac{4}{4\pi n_2\sigma_2^2 + \pi n_1(\sigma_1+\sigma_2)^2}\right]. \quad \text{(S9)}$$

With σ_1 = 3.467 Å and σ_2 = 3.798 Å for oxygen and nitrogen (Table S2), and n_1 and n_2 corresponding to air mole fractions equal to 0.21 and 0.79, respectively, Equation (S9) yields λ_1 = 71.1 nm and λ_2 = 65.0

nm at 300 K and 1 atm, implying a value of λ equal to (0.21 x 71.1 + 0.79 x 65.0) nm = 66.3 nm. On the other hand, according to a widely adopted theoretical expression[69]

$$\lambda = \sqrt{\frac{\pi}{8}} \cdot \frac{\eta}{f} \cdot \frac{1}{\sqrt{\rho P}} \tag{S10}$$

where η is the viscosity of air, ρ the density of air and *f* a numerical factor equal to 0.4987445, resulting in λ = 67.3 nm at the same conditions (using η = 18.5 μPas and ρ = 1.177 kg/m^3).

Free path distributions

For a gas molecule that is in equilibrium, the distribution $p(\lambda')$ of free paths λ' is defined as:[29,30]

$$p(\lambda') = \lambda^{-1} \exp(-\lambda'/\lambda) \tag{S11}$$

From Eq. (S11), the mean free path λ is obtained as $\lambda = \int_0^\infty \lambda' \cdot p(\lambda') \cdot \mathrm{d}\lambda'$.

Variants of the hard-sphere model

Despite that the classic hard-sphere model provides significant insight into gas motion and helps extract a wealth of analytical expressions for most of transport properties of gases, the corresponding temperature scalings of these properties are not accurate. To overcome this, Bird[7] introduced the so-called Variable Hard-Sphere (VHS) model. This model follows the simple isotropic hard-sphere angular scattering law but considers the collision diameter not to be constant but to vary as a function of the relative energy in the collision. Thus, according to the VHS model, the effective diameter d_{ref} at a reference temperature T_{ref} is equal to[12]

$$d_{\mathrm{ref}} = \left(\frac{15\left(mk_{\mathrm{B}}T_{\mathrm{ref}}/\pi\right)^{1/2}}{2(5-2\omega)(7-2\omega)\eta_{\mathrm{ref}}} \right)^{1/2} \tag{S12}$$

where *m* denotes the molar mass, η_{ref} is the viscosity at the reference temperature T_{ref}, and ω is the scaling index for the viscosity. The corresponding hard-sphere expression is recovered for ω = 0.5 and reads[12]

$$d_{\mathrm{ref}} = \left(\frac{5}{16} \frac{\left(m k_{\mathrm{B}} T_{\mathrm{ref}} / \pi \right)^{1/2}}{\eta_{\mathrm{ref}}} \right)^{1/2} \tag{S13}$$

Typical values of VHS parameters for nitrogen and air are reported in Table S3.

Table S3. Values of the VHS parameters T_{ref}, η_{ref}, and ω (from Ref. 12).			
Gasses	T_{ref}	η_{ref}	ω
N_2	273.15	16.56	0.74
Air	273.15	17.19	0.77

Then, according to the VHS model, the mean free path of the gas at equilibrium is equal to[12]

$$\lambda = \left\{ 2^{1/2} \pi d_{\mathrm{ref}}^2 n \left(\frac{T_{\mathrm{ref}}}{T} \right)^{\omega - \frac{1}{2}} \right\}^{-1} . \tag{S14}$$

Almost a decade later, Koura and Matsumoto[8] introduced yet another variant of the hard-sphere model, the so-called Variable Soft-Sphere (VSS) model. The VSS model utilizes the same total cross-section model as the VHS model but includes a more realistic scattering model. In the same article, Koura and Matsumoto[8] provide two expressions of the VHS model for the gas mean free path, based on viscosity η and diffusivity D measurements, respectively:

$$\lambda_{\eta} = \left(8\eta/15\right)\left(3-\xi\right)\left(2-\xi\right)\left(2\pi m k_B T\right)^{1/2} \Big/ n \tag{S15}$$

$$\lambda_{D} = \left(4D/3\right)\left(2-\xi\right)\left(2\pi k_B T/m\right)^{1/2} \tag{S16}$$

with ξ being the energy exponent related to the viscosity exponent $\xi = \omega - 1/2$.[12]

The VSS model is an extension of the VHS model in the sense that the following expression is used to relate the impact parameter b to the collision angle χ:

$$b = d \cos^{\alpha}\left(\chi/2\right) \tag{S17}$$

with $a \geq 1$. According to the VSS model, the viscosity and the diffusivity are given through

$$\eta_{\mathrm{VSS}} = \eta_{\mathrm{VHS}} \big/ S_{\eta} \tag{S18}$$

$$D_{\mathrm{VSS}} = D_{\mathrm{VHS}} / S_D \tag{S19}$$

with the corresponding softness coefficients being equal to

$$S_\eta = 6a / \left[(a+1)(a+2) \right] \tag{S20}$$

$$S_D = 2 / (a+1) \tag{S21}$$

From the above transport coefficients, the following mean free path expressions are derived:

$$\lambda_{\eta,\mathrm{VSS}} = S_\eta \lambda_{\eta,\mathrm{VHS}} \tag{S22}$$

$$\lambda_{D,\mathrm{VSS}} = S_D \lambda_{D,\mathrm{VHS}} \tag{S23}$$

Also, according to the VSS model, the effective diameter of the gas molecules becomes

$$d_{\mathrm{ref}} = \left(\frac{5(a+1)(a+2)\left(m k_{\mathrm{B}} T_{\mathrm{ref}} / \pi\right)^{1/2}}{4a(5-2\omega)(7-2\omega)\eta_{\mathrm{ref}}} \right)^{1/2} \tag{S24}$$

For $a = 1$, Equation (S23) of the VSS model simplifies to Equation (S11) of the VHS model. Typical values of the parameters entering the VSS model for nitrogen (two data sets) and air are reported in Table S4.

Table S4. Values of the parameters ξ, a, S_η, and S_D entering the VSS model for nitrogen and air (from Ref. 8).

Gasses	ξ	α	S_η	S_D
N_2	0.238	1.5059	1.0285	0.7981
Air	0.268	1.5732	1.0266	0.7773

The generalized hard-sphere (GHS) model was introduced by Hassan and Hash[10,11] to enable a VHS-type of model to include the attractive portion of the interatomic potential energy surface (e.g., that appearing in the Lennard-Jones interaction) and thus a nonconstant power-law exponent for the transport properties. Unlike the VHS and VSS models, a reference diameter or reference cross-section

(i.e., reference relative velocity) cannot be defined for the GHS model, and values for all the GHS model parameters must be specified for each species pair.[13]

Table S5 summarizes the predictions of the above variants of the hard-sphere model for the mean free path of air, based on the parameters in Tables S3 and S4.

Table S5. Mean free path, λ, of nitrogen and air from theoretical expressions at T = 300 K and P = 1 atm.

	Nitrogen	Air
HS model, Chapman-Enskog, Equation (S12)	65	66.3
Jennings (1988) expression, Equation (S10)	66	67.3
VHS model, Equation (S11)	54	53.6
VHS model, from the dynamic viscosity, Equation (S14)	54.8	54.2
VHS model, from the diffusion coefficient, Equation (S15)	68.8	63.7
VSS model, Equations (S13) and (S23)	55.4	55

S4. MAXWELL-BOLTZMANN VELOCITY DISTRIBUTION

The distribution of velocities from the three models were calculated and compared (Figure S2) to the Maxwell-Boltzmann expression for the velocity distribution:[29]

$$f(u) = \left(\frac{m}{2\pi k_B T} \right)^{3/2} 4\pi u^2 \exp\left(-\frac{mu^2}{2k_B T} \right) \quad \text{(S25)}$$

where u is the molecular velocity and m the air molar mass. The distribution is normalized to unity. As expected, the results of all three MD models practically overlap with those from the above Maxwell-Boltzmann expression.

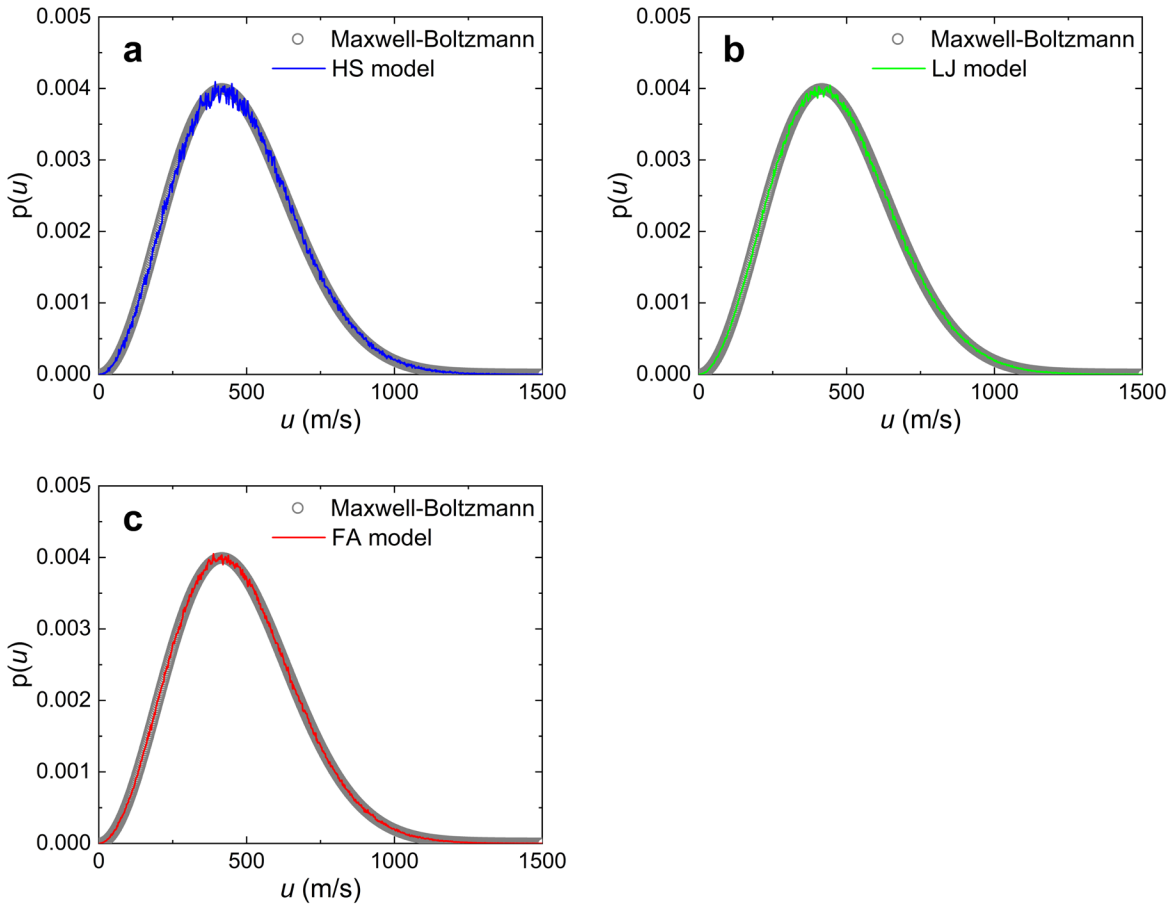

Figure S2. Molecular velocity distributions by the three MD models (HS, LJ, and FA) are in excellent agreement with those from the Maxwell-Boltzmann expression for the distribution of gas velocity.

S5. IMPACT OF SELECTED MD SYSTEM SIZE

The system size is selected by repeating simulations with larger and larger simulation cells till convergence of air properties is attained. Table S6 and Figure S3 below show the ρ, D and λ of air at T = 300 K and P = 1 atm with different system sizes from NVE simulations with the FA model. Clearly, when using simulation cells with, at least, 8'000 air molecules, the above air properties have reached their asymptotic values. In all of simulations, cells with 50'000 air molecules have been used.

Table S6. Air density, diffusivity and mean free path of air at 1 atm and 300 K by the FA model using progressively larger basic simulation cells (with increasing number of air molecules).				
Number of air molecules	Box length (nm)	Density (kg/m^3)	Diffusivity (cm^2/s)	Mean free path (nm)
2000	38.0	1.75 ± 0.01	0.107 ± 0.006	23.5 ± 0.6
4000	54.4	1.19 ± 0.01	0.202 ± 0.004	37.2 ± 1
8000	68.8	1.18 ± 0.01	0.205 ± 0.006	38.4 ± 1
10000	74.1	1.18 ± 0.01	0.206 ± 0.006	38.6 ± 1
50000	126.7	1.18 ± 0.01	0.203 ± 0.005	38.5 ± 1

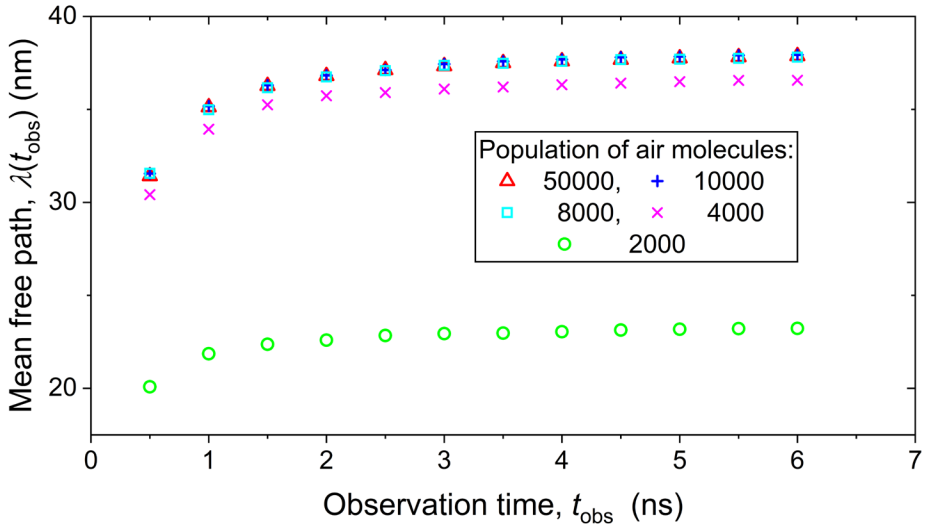

Figure S3. The evolution of the mean free path (by direct averaging) by the FA model with simulation cells of increasing size (number of air molecules). Clearly, at least, 8'000 molecules are needed for reliable results while 50'000 have been used in all simulations. The errors in the calculations are smaller than the symbol size.

S6. HAZARD PLOT ANALYSIS OF SPURIOUS COLLISIONS

Application of Hazard plot analysis[61,62] to the population of free paths generated by the three models (Figure S4) is shown in Figure S5. It reveals that the overwhelming majority of free paths that are larger than 0.2 Å occur at about the same rate (0.00239 and 0.00259 $Å^{-1}$ for the LJ and FA models, respectively). This analysis also reveals that a fraction of very short paths (shorter than 0.2 Å, Figure S5a,b insets) exists, occurring at a much faster rate (0.432 and 0.581 $Å^{-1}$ for the LJ and FA models, respectively).

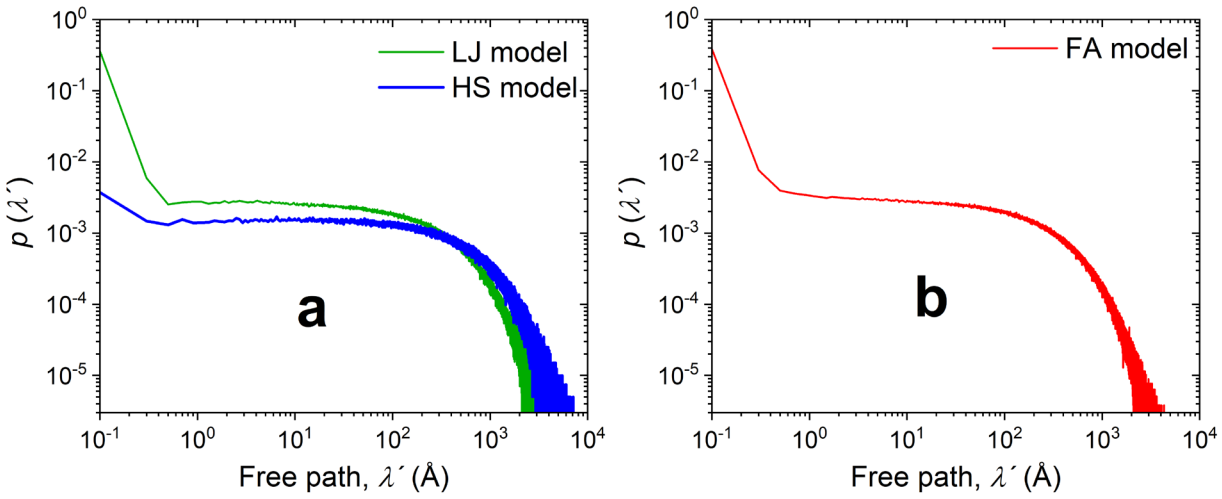

Figure S4. Probability density distribution of free paths (including spurious ones) for: a) the HS and LJ, and b) the FA models, by analyzing collisions from MD data generated over 6 ns of simulation time.

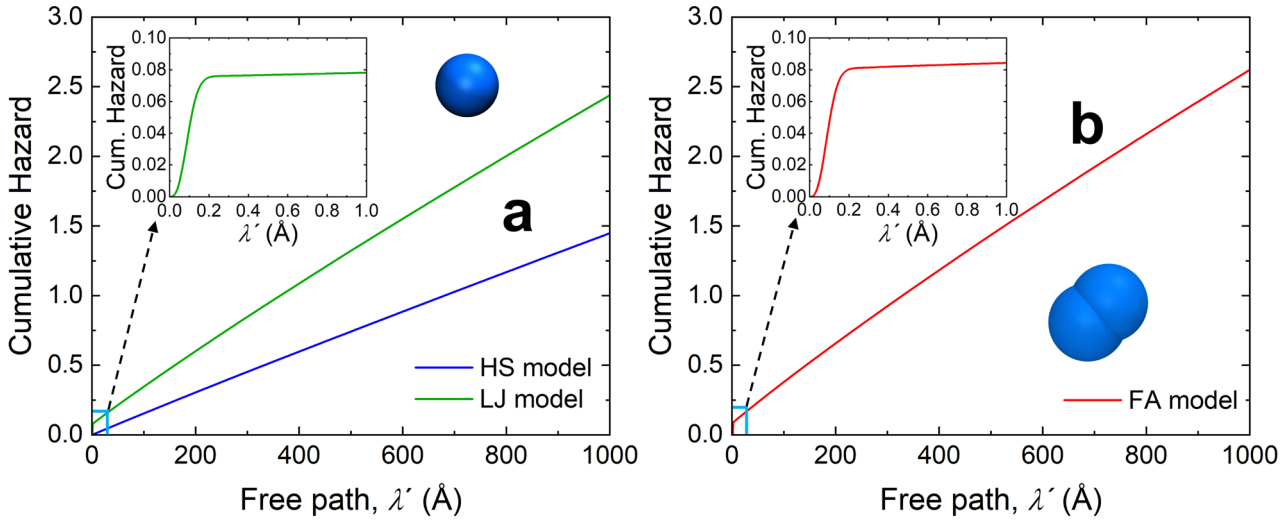

Figure S5. Hazard plots for identification of spurious collisions with the: a) HS and LJ, and b) FA models. Most of free paths follow an exponential distribution, a straight line of constant slope equal to 0.00145, 0.00239, and 0.00259 Å^{-1} for the HS, LJ, and FA models, respectively. The insets show the regime of spurious collisions corresponding to very short paths (< 0.2 Å) occurring at a much faster rate than the rest (i.e., normal collisions). For the HS model, no spurious collisions occur, so no inset is shown.

S7. STATISTICS OF MULTI-BODY COLLISIONS

Figure S6 shows the distribution of collision times from the LJ and FA models. For two-body collisions (Fig. S6a), the distributions are similar qualitatively, exhibiting a maximum at t = 230 ps and t = 290 ps for the FA and the LJ model, respectively. Only a tiny amount (less than ~1%) of two-body collisions

disengage instantly ($t \leq 20$ fs), in sharp contrast to the HS model. This behavior is attributed to the attractive part of the potential. On the overwhelming number of two-body collisions (~99%), the molecules stay together for some appreciable time that can be as large as 1000 fs for the FA model and 550 fs for the LJ model. On the other hand, the distribution of disengagement times in the FA model is broader than in the LJ model, which is mainly related to the non-spherical shape of the molecules accounted by the FA model. In particular, due to their diatomic representation in the FA model, the molecules influence more space, thus increasing the range of both their repulsive and attractive interactions, resulting in a significantly more expanded distribution of disengagement times.

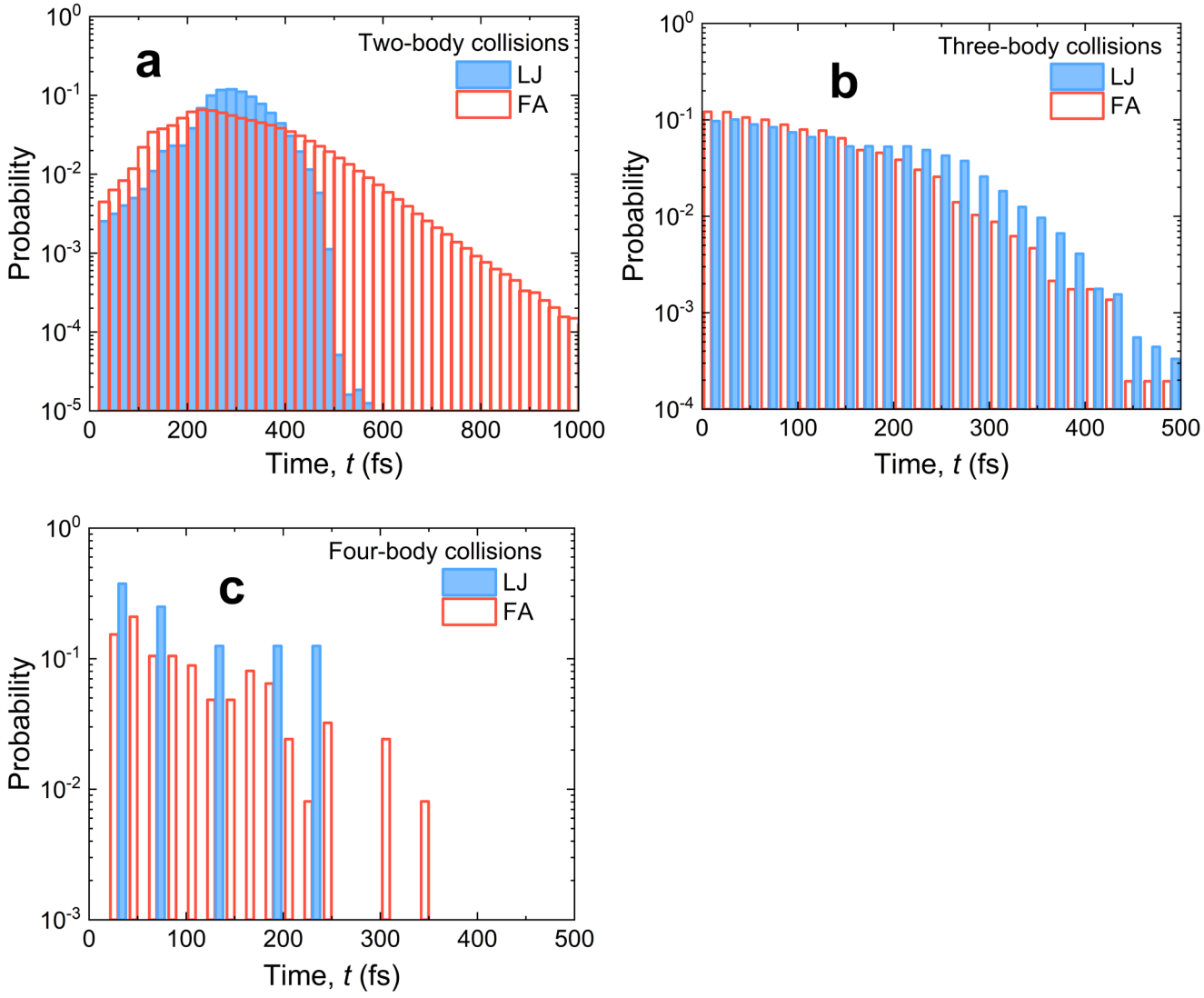

Figure S6. Probability distributions of collision times for: a) two-, b) three- and c) four-body collisions, from the LJ (blue filled bars) and FA model (red open bars) simulations, respectively.

For the three-body collisions (Figure S6b), the distributions are similar not only qualitatively but also quantitatively. Again, only a small population of colliding molecules disengage instantly (~9.7% for the LJ and ~12.1 % for the FA model); the rest tend to stay together for up to ~ 500 fs. The corresponding distributions of four-body collisions are shown in Figure S6c. They are quite noisy due to their relatively low occurrence probability.

Overall, the relatively simpler LJ model offers a pretty realistic description of the total number of collisions and their characteristic times despite its treating the molecules as perfect spheres, thus highlighting the critical role of accounting for the detailed molecular interactions by the force field. The distributions of disengagement times indicate that the impact of shape is less crucial than that of the force field but should not be neglected. Accounting for the detailed molecular shape, by making use of the FA model, the two-body collisions disentangle at longer times than in the LJ model, and the three-body collisions are 58% more than those in the LJ model (Table S7).

Table S7. Fractions (%) of two-, three- and four-body collisions as calculated by the detailed FA model and its two limiting cases that assume: a) fully elastic collisions between spherical air molecules (HS-model), and b) collisions between such molecules but with a Lennard-Jones potential between them (LJ-model).

Collision-type \ %	HS	LJ	FA
2-body	100	99.47	99.16
3-body	0	0.53	0.84
4-body	0	< 0.0005	< 0.005

In the HS model (results not shown), two-body collisions occur instantly due to the purely repulsive character of the potential, consistent with kinetic theory for rigid molecules that experience only elastic collisions; also, as already mentioned above, 3- and 4-body collisions never occurred with the HS model.

S8. COLLISION DENSITIES FOR SPURIOUS COLLISIONS

Accounting for all spurious collisions increases by an order of magnitude the collision densities (Table S8) by the LJ and FA models over those neglecting such collisions (Table I).

Table S8. Collision densities accounting for the spurious collisions by the Lennard-Jones (LJ) and fully atomistic (FA) models.				
	Collision densities (10^{33} $m^{-3}s^{-1}$)			
	oxygen - oxygen	nitrogen – nitrogen	nitrogen – oxygen	Total
LJ Spurious	6.3 ± 0.1	105.2 ± 2.8	51.6 ± 1.2	163.1 ± 4.2
FA Spurious	6.9 ± 0.2	114.1 ± 3.2	56.3 ± 1.4	177.3 ± 4.9

S9. DISTRIBUTION OF FREE PATHS

Figure S7 shows the evolution of the probability density distribution of free paths, λ', from the FA model for observation times t_{obs} from 0.5 to 6 ns. With increasing t_{obs}, the distribution shifts to the right as longer and longer paths are sampled. The shift rate slows down as the relative population of shorter λ' tends to stabilize while longer ones that are added to the distribution are characterized by significantly lower probabilities of occurrence. As a result, the λ' distribution tends to converge and stabilize asymptotically.

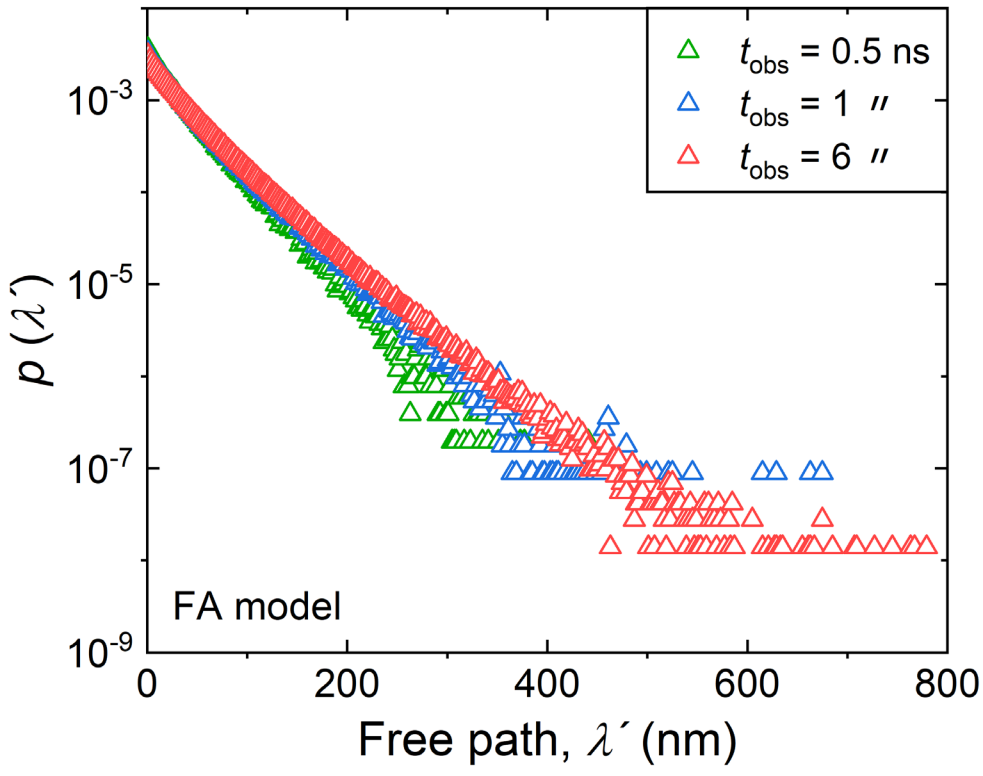

Figure S7. The probability density distribution $p(\lambda')$ of free paths λ' from the simulations with the FA model at observation times t_{obs} = 0.5, 1, and 6 ns.

To ensure that the λ' calculations do not depend on the initial configuration, we conducted a second MD simulation in the NVT ensemble, starting from an entirely different configuration of atomic positions and velocities for the three models (at the same temperature and density conditions with those in the NVE simulations). We again extracted the probability distributions of free paths and calculated their mean by directly averaging. Figure S8 shows the obtained λ as a function of observation time in two simulations for the HS, LJ, and FA models. The agreement of the results from the NVT and NVE independent simulations is remarkable; in all cases for t_{obs} > 1 ns, in particular, the NVT (crosses) and NVE sets of results are practically indistinguishable.

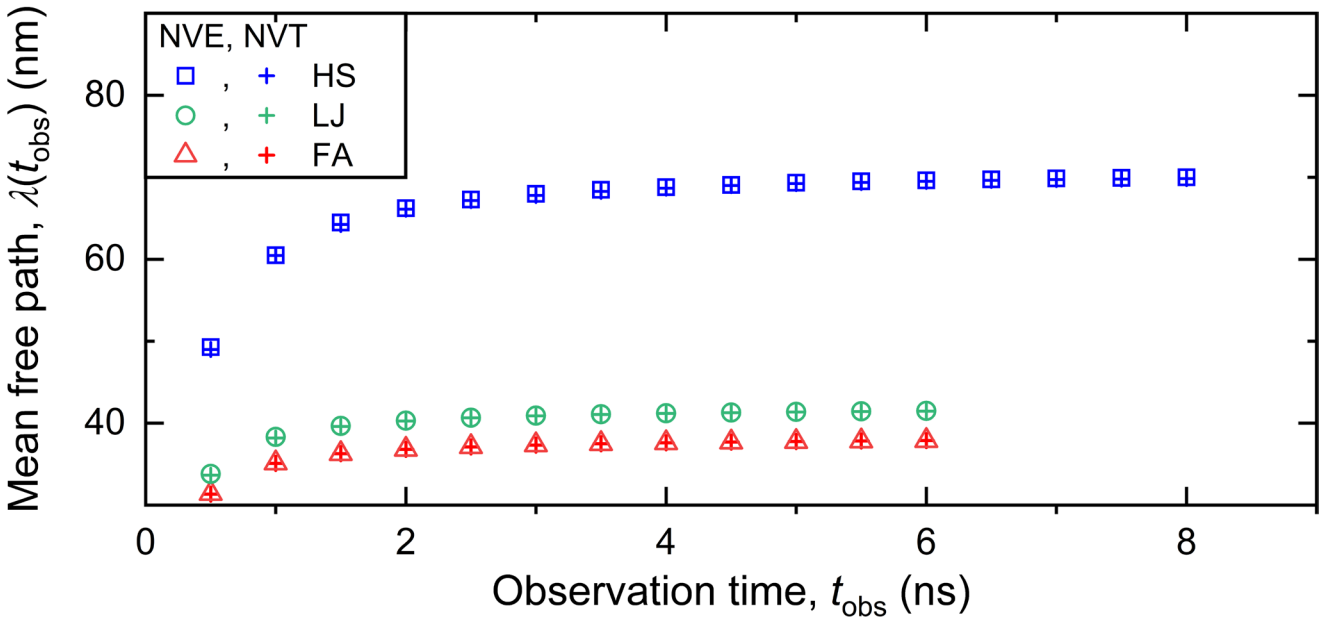

Figure S8. Comparison of the predictions of λ (by direct averaging) from MD simulations in NVT (crosses) and NVE (open symbols) by the three models HS (blue squares), LJ (green circles), and FA (red triangles). The errors in the calculations are smaller than the symbol size.

S10. IMPACT OF COLLISION DISTANCE DEFINITION ON MEAN FREE PATH

The λ values from the LJ and FA models in Table II have been computed for a collision distance equal to the characteristic LJ distance $2^{1/6}\sigma_{ij}$ at which the force between a pair of atoms (i, j) changes from attractive to repulsive as they approach each other. For the HS model, on the other hand, the collision distance is equal to the diameter σ_{ij} of the two molecules. Dongari et al.[71] calculated by MD the

λ of neon gas modeled as perfect spheres subject to a LJ potential (like the present LJ model) but using σ_{ij} as the collision distance. Their prediction for the λ of neon at T = 273.15 K and ρ = 0.9011 kg/m^3 was λ = 126.2 nm, which is in agreement with the corresponding kinetic theory result that λ = 125.5 nm. To check the sensitivity of λ on the choice of the collision distance but also to directly compare the two works, we performed simulations of the same system and thermodynamic conditions with exactly the same LJ parameters as in Dongari et al.,[71] using two values for r_m (Fig. S9), equal to $2^{1/6}\sigma$ and σ (filled and open circles, respectively, in Fig. S10), where σ is the neon atom LJ diameter.

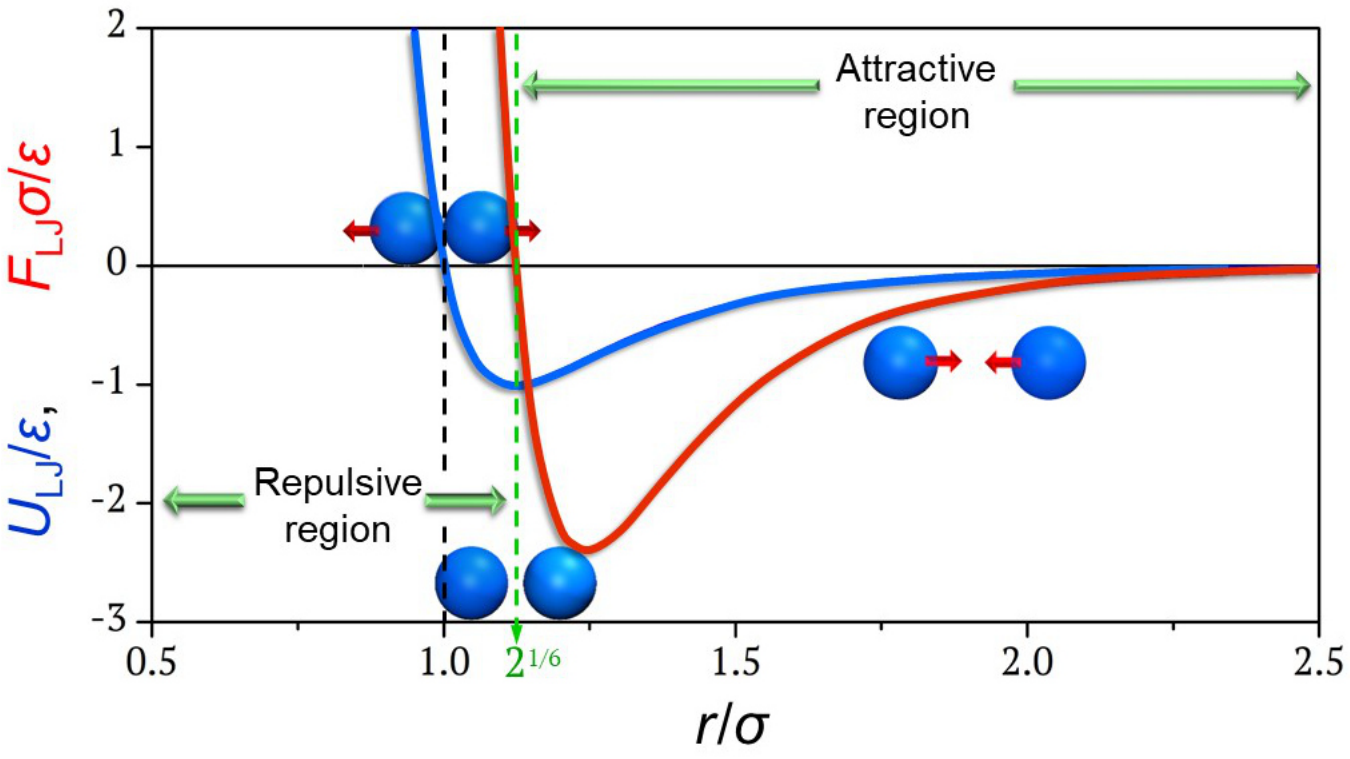

Figure S9. Comparison of the two choices for collision distance, r_m: $r_m = \sigma$ (i.e., $r_m / \sigma = 1$, vertical black dashed line), and $r_m = 2^{1/6}\sigma$ (i.e. $r_m / \sigma = 2^{1/6}$, vertical green dashed line). The former corresponds to the distance where the LJ potential energy becomes zero, while the latter to the distance where the force between the two becomes zero as it changes sign from negative (attractive) at $r_m / \sigma > 2^{1/6}$ to positive (repulsive) at $r_m / \sigma < 2^{1/6}$. All LJ and FA model results have been obtained with $r_m = 2^{1/6}\sigma$.

When $r_m = \sigma$, the λ from our LJ model is λ = 127.0 ± 1.5 nm, which matches well that of Dongari et al.[71] after about 3 ns of observation time, 126.2 nm. However, when $r_m = 2^{1/6}\sigma$, the results differ substantially, since λ = 85.2 ± 1.1 nm. This happens because when the collision distance is set to σ, all events corresponding to inter-atomic separations r in the interval $\sigma < r < 2^{1/6}\sigma$ are not counted as collisions. Then collision densities are underestimated, and consequently, the λ is over-estimated as the LJ potential is essentially ignored.

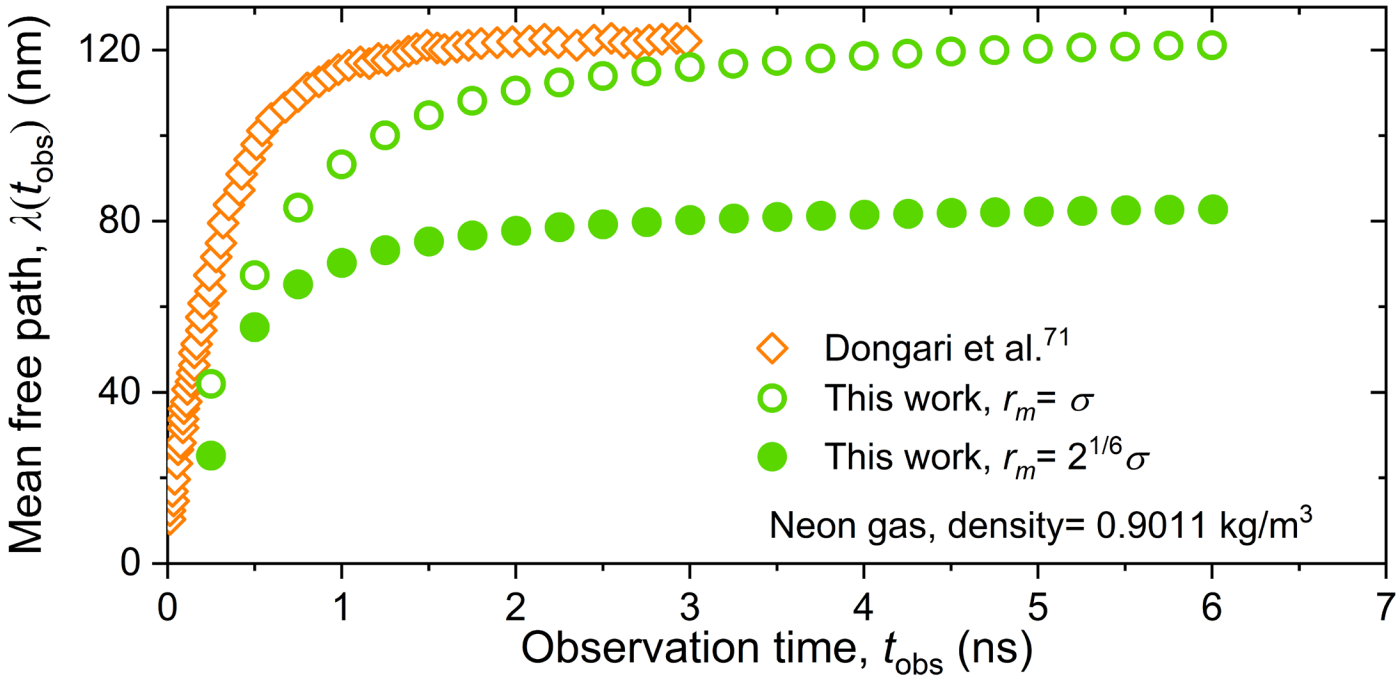

Figure S10. Comparison of computed mean free path as a function of observation time for neon gas molecules from Dongari et al.[71] (diamonds) where $r_m = \sigma$ was used, and from our work with $r_m = \sigma$ (open circles) and $r_m = 2^{1/6}\sigma$ (filled circles). The errors in the calculations are smaller than the symbol size. In all three cases. the MD simulations were performed with the same LJ potential, at exactly the same T = 273.15 K and ρ = 0.9011 kg/m^3. The difference between the two calculations indicates that one has to be careful in imposing the natural collision criterion (namely that $r_m = 2^{1/6}\sigma$, filled circles) and not that pertinent to hard spheres ($r_m = \sigma$, open symbols), as in the latter a significant number of collisions are neglected.

S11. IMPACT OF FORCE FIELD PARAMETERS ON MEAN FREE PATH

Table S9. Intermolecular parameter values for the different FA force fields for air.

Model	Lennard-Jones potential parameters			
	N-N		O-O	
	σ (Å)	ε (kcal/mole)	σ (Å)	ε (kcal/mole)
Zambrano et al.[32]	3.32	0.07232	2.99	0.10330
Wang et al.[35]	3.614	0.0797	3.297	0.1047
Bouanich[36]	3.29084	0.0739	3.0145	0.102823
Vrabec et al.[37]	3.3211	0.06935	3.1062	0.085813
Vacha et al.[38]	4.201	0.1973	2.955	0.2029
Zhang et al.[39]	3.31	0.07155	3.02	0.097373
Tokumasu and Matsumoto[41]	3.17	0.093857	-	-
Kosyanchuk and Yakunchikov[43]	3.4	0.043321	-	-

Table S10. The density, diffusivity, viscosity and mean free path of air at T = 300 K and P = 1 atm, using different force fields in the FA model.

	ρ (kg/m^3)	D (cm^2/s)	η (μPas)	λ (nm)
Experimental data	1.17	0.203	18.5	-
Zambrano et al.[32]	1.17	0.203	18.5 ± 0.1	38.5 ± 1
Wang et al.[35]	0.98	0.241	20.1 ± 0.1	39.8 ± 1
Bouanich[36]	1.17	0.213	18.4 ± 0.1	38.3 ± 1
Vrabec et al.[37]	1.16	0.214	18.1 ± 0.2	38.5 ± 1
Vacha et al.[38]	1.18	0.103	9.4 ± 0.2	20.5 ± 0.6
Zhang et al.[39]	1.17	0.212	18.2 ± 0.2	38.4 ± 1

Table S11. The density, diffusivity, viscosity and mean free path at T = 300 K and P = 1 atm of pure nitrogen with three different force fields (a purely classical one and two based on ab-initio parameterization) and comparison with experimental data.

	ρ (kg/m^3)	D (cm^2/s)	η (μPas)	λ (nm)
Experimental data[63,65,72,73]	1.126	0.219	17.8	-
Zambrano et al.[32]	1.14	0.213 ± 0.003	17.4 ± 0.3	37.6
Tokumasu and Matsumoto[41]	1.15	0.215± 0.003	18.6 ± 0.3	38.3
Kosyanchuk and Yakunchikov[43]	1.14	0.222± 0.003	19.5 ± 0.5	40.5

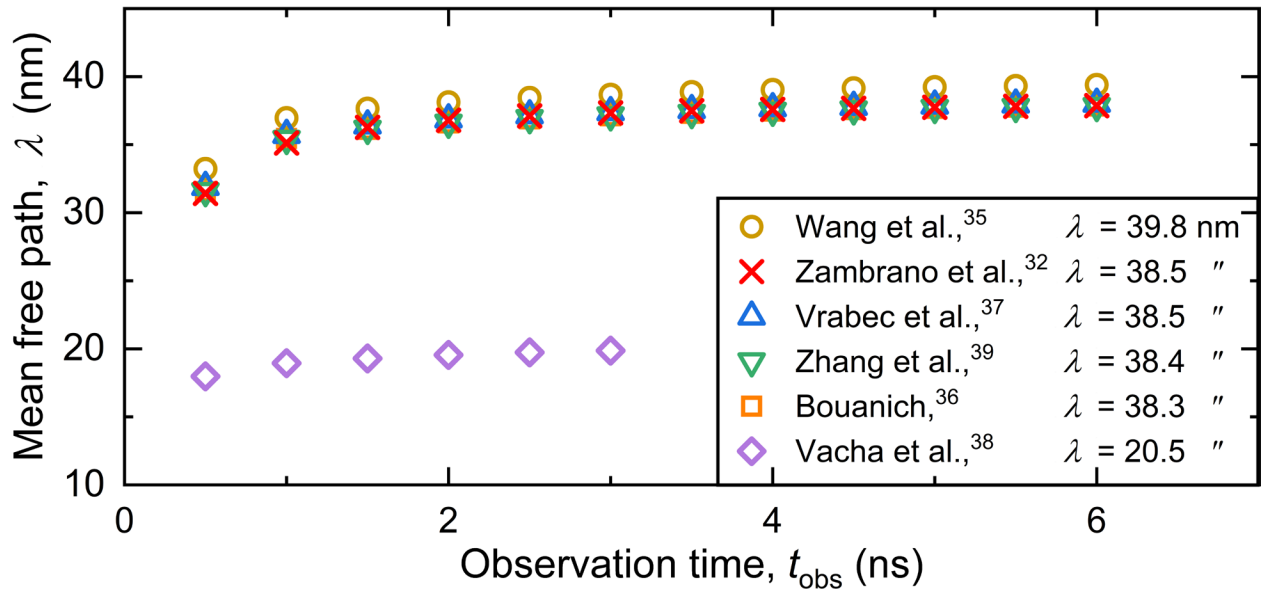

Figure S11. The mean free path of air as a function of observation time (symbols) from several fully atomistic MD models. The results have been obtained by directly averaging over the corresponding density distributions of free paths. The errors in the calculations are smaller than the symbol size. The λ values reported in the legend indicate the asymptotic values in the limit of long observation times, obtained by fitting the simulation data with a hyperbolic function and taking the limit $t_{obs} \to \infty$.

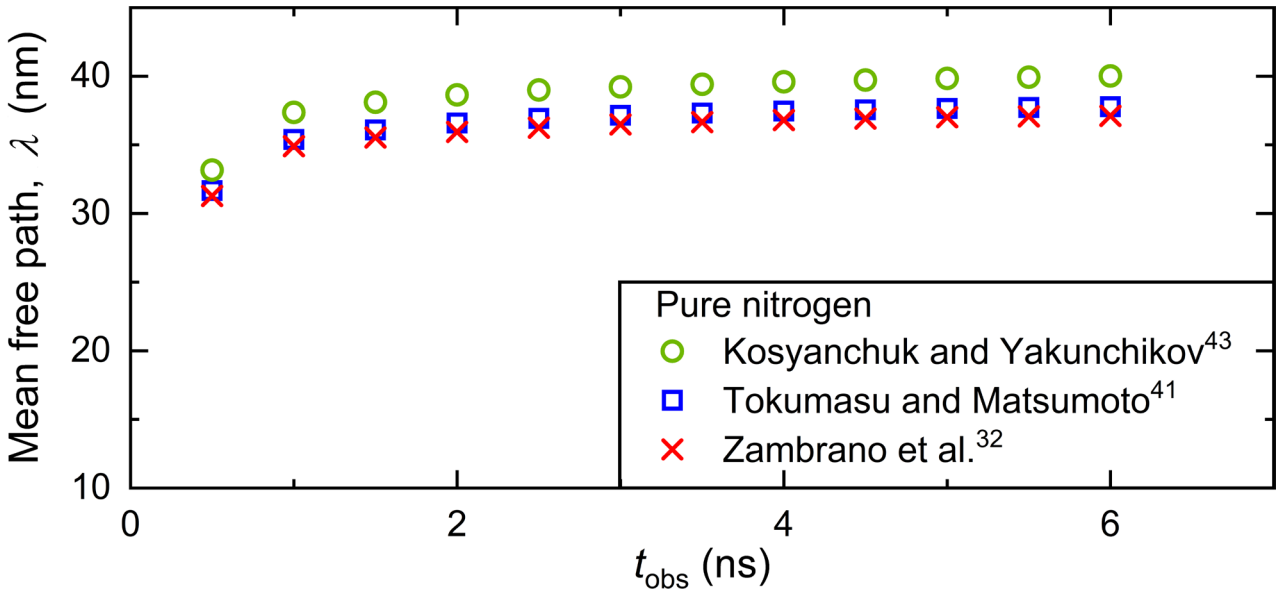

Figure S12. Mean free path of pure nitrogen as a function of observation time (symbols) from three fully atomistic models: a purely classical one[32] and two based on ab initio parametrization.[41,43] The results were obtained by directly averaging over the corresponding density distributions of free paths. The errors in the calculations are smaller than the symbol size.