Electronic Supplementary Material (ESI) for Analyst.

A Local Collision Probability Approximation for Predicting Momentum Transfer Cross Sections.

Christian Bleiholder

Department of Chemistry and Biochemistry, Florida State University, Florida State University, Tallahassee, FL 32306-4390

Address reprint requests to

Christian Bleiholder, Department of Chemistry and Biochemistry, Florida State University, Florida State University, Tallahassee, FL 32306-4390, Phone +1-850-644-1290, Fax +1-850-644-8281

SUPPORTING INFORMATION

1 Theory

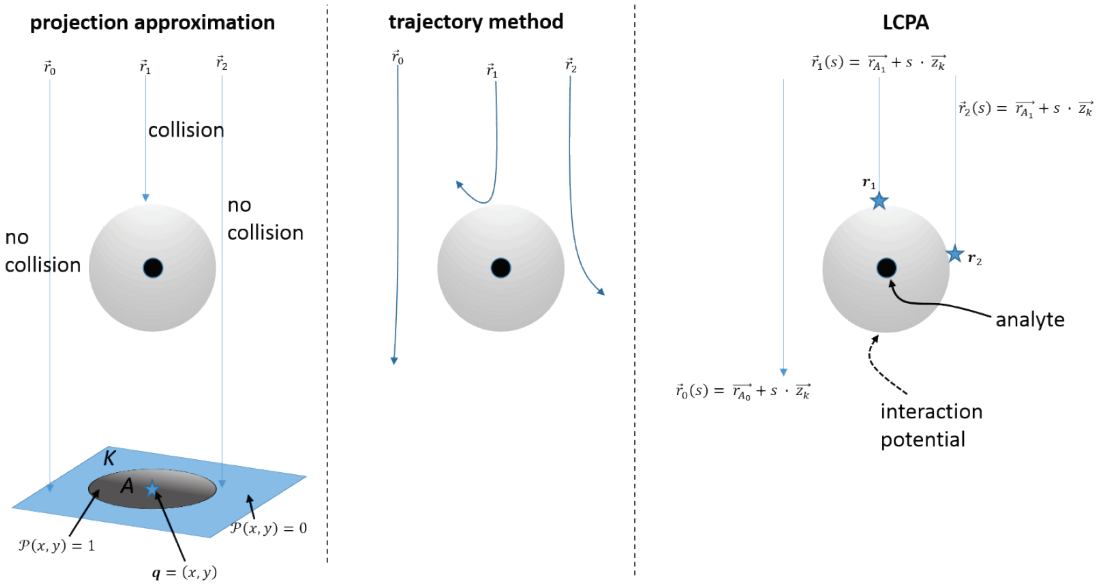

Figure S1. Comparison of the collision process in the projection approximation (PA), local collision probability approximation (LCPA), and trajectory methods. The center-of-mass of the analyte and the interaction potential are indicated by a black and a gray shaded circle, respectively. (left) The projection approximation determines points $\boldsymbol{q} = (x, y)$ on a projection plane *K* that represent collisions with the analyte ion. To this end, a probability function $\mathcal{P}(x, y)$ on the projection plane *K* is first defined. Then, any approaching buffer gas particle is described by a ray $\vec{r}$ that approaches the projection plane *K*. Any such ray that intersects plane *K* at a point $\boldsymbol{q} = (x, y)$ for which $\mathcal{P}(x, y)$=1 is considered 'a collision', or 'not a collision' otherwise. (center) The trajectory method computes the exact path that any approaching buffer gas particle takes during the encounter and records the corresponding deflection angle. Trajectories with large deflection angles (such as $\vec{r}_1, \vec{r}_2$) contribute strongly to the momentum transfer cross section while trajectories with a small deflection angle ($\vec{r}_0$) contribute only marginally. (right) The LCPA determines the points $\boldsymbol{r} = (x, y, z)$ that represent collisions with the buffer gas ($\boldsymbol{r}_1, \boldsymbol{r}_2$). To this end, the path of rays $\vec{r}(s) = \vec{r}_A + s \cdot \vec{z}_k$ is followed and the points $\vec{r}(s)$ at which a collision occurs according to the local collision probability $\tau(\varepsilon, \boldsymbol{r})$ is determined. The collection of these points $\vec{r}(s)$ defines the collision surface $\mathcal{S}$.

As depicted in Figure S1, the projection approximation (PA) determines points $\boldsymbol{q} = (x, y)$ on the projection plane *K* that fall within the 'effective area of the analyte' *A*. As discussed in the main text, a local probability measure $\mathcal{P}(x, y)$ is defined that decides whether or not $\boldsymbol{q} \in A$. The set of points $\boldsymbol{q}$ define the area of the analyte, projected onto *K*. In analogy to the PA, the LCPA determines points $\boldsymbol{r} = (x, y, z)$ *in three dimensions* at which collisions occur according to the probability measure $\tau(\varepsilon, \boldsymbol{r})$. Thus, the set of all

points $\boldsymbol{r}$ determined in the LCPA method defines a volumetric surface (in analogy of the area determined by points $\boldsymbol{q} \in A$ in the PA).

In order to construct an appropriate expression for $\tau(\varepsilon, \boldsymbol{r})$, two things must be accomplished. First, the (non-local) scattering process of the analyte ion and buffer gas particle that occurs on the interaction potential $U(\boldsymbol{r})$ must be decomposed into a superposition of (localized) potentials $V_\delta(\boldsymbol{r})$, each one only acting locally at position $\boldsymbol{r}$. Second, an appropriate expression for $\tau(\varepsilon, \boldsymbol{r})$ must be constructed for each of the locally acting potentials $V_\delta(\boldsymbol{r})$.

1.1 The interaction potential $\boldsymbol{U(r)}$ as a superposition of localized delta-potentials $\boldsymbol{V_\delta(r_i)}$

Any multivariate real function $f(x, y, \dots, z)$ can be expressed with Dirac-delta functions $\delta(x - \mu_i)$ as

$$f(x, y, \dots, z) = \int_{-\infty}^{-\infty} \cdots \int_{-\infty}^{+\infty} f(\mu_x, \mu_y, \dots, \mu_z)\delta(x - \mu_x)\delta(y - \mu_y) \dots \delta(z - \mu_x)\, d\mu_x d\mu_y \dots d\mu_z \tag{S1}$$

This decomposition applies in particular to the analyte-buffer gas particle interaction potential $U(\boldsymbol{r})$, with $\boldsymbol{r} = (x, y, z)$:

$$U(x, y, z) = \iiint_{-\infty}^{+\infty} dl\, dm\, dn\; U(l, m, n)\delta(x - l)\delta(y - m)\delta(z - n) \tag{S2}$$

For simplicity, this can be written in only one dimension x without loss of generality as

$$U(x) = \int_{-\infty}^{+\infty} dl\, U(l)\delta(x - l) \tag{S3}$$

Or, if the integral is replaced by the limit of a sum, as:

$$U(x) = \lim_{n\to\infty} \sum_{i=0}^{n} \Delta l_i \cdot U(l_i) \cdot \delta(x - l_i) \tag{S4}$$

The summand $\Delta l_i \cdot U(l_i) \cdot \delta(x - l_i)$ in equation S4 has the meaning of an interaction potential $V_\delta(l_i)$ with height $U(l_i)$ that acts only locally at position l_i (due to the $\delta(x - l_i)$ term). The term Δl_i corresponds to the length-scale of the local interaction potential $V_\delta(l_i)$ and is required because the term $\delta(x - l_i)$ really is only defined in the limit of the summation. (Additionally, the delta-measure $\delta(x - l_i)$ has the dimension of $[\mathrm{length}]^{-1}$ and so the term Δl_i has the meaning of the length-scale of the locally acting potential $V_\delta(l_i)$.) Thus, the interaction potential $U(x)$ can be decomposed into a superposition of locally acting potentials $V_\delta(l_i)$

$$U(x) = \lim_{n\to\infty} \sum_{i=0}^{n} V_\delta(l_i) \tag{S5}$$

These locally acting potentials $V_\delta(l_i)$ are from now on referred to as delta-potentials and written in three dimensions as $V_\delta(\boldsymbol{r_i})$, acting locally at position $\boldsymbol{r_i} = (x_i, y_i, z_i)$ with

$$V_\delta(\boldsymbol{r}_i) = l \cdot\ U(\boldsymbol{r}) \cdot \delta(\boldsymbol{r}_i) = g \cdot \delta(\boldsymbol{r}_i) \tag{S6}$$

Here, l is the length-scale, $U(\boldsymbol{r})$ the height, and g the area under $V_\delta(\boldsymbol{r}_i)$; the notation $\delta(\boldsymbol{r}_i)$ is used to abbreviate $\delta(x - x_i) \cdot \delta(y - y_i) \cdot \delta(z - z_i)$. An expression for the length-scale l will be provided in equation (S12). Thus, the potential $U(\boldsymbol{r})$ in three dimensions is written as a superposition over locally acting delta-potentials $V_\delta(\boldsymbol{r_i})$ according to

$$U(\boldsymbol{r}) = \lim_{n\to\infty} \sum_{i=0}^{n} V_\delta(\boldsymbol{r_i}) \tag{S7}$$

1.2 Local collision probability function $\boldsymbol{\tau(\varepsilon, r)}$

The meaning of $\tau(\varepsilon, \boldsymbol{r})$ is the probability for a collision to occur at position $\boldsymbol{r}$ when a ray $\boldsymbol{r}(s)$ is cast onto the molecule. In this section, an appropriate expression for $\tau(\varepsilon, \boldsymbol{r})$ will be constructed. To this end, the decomposition of the interaction potential $U(\boldsymbol{r})$ as a superposition over locally acting delta-potentials $V_\delta(\boldsymbol{r_i})$ is exploited (equation S7). Using equation S7, the (non-local) scattering process of the analyte – buffer gas interaction on the interaction potential $U(\boldsymbol{r})$ is treated as a superposition of localized, independent scattering events, each scattering event taking place at a delta-potential $V_\delta(\boldsymbol{r_i})$ at position $\boldsymbol{r_i}$.

As depicted in Figure S1, the LCPA treats the collision process by casting rays $\boldsymbol{r}(s) = \boldsymbol{r}_\infty + s \cdot \boldsymbol{z}$ onto the analyte (in analogy to the projection approximation). The scattering process *along the ray* $\boldsymbol{r}(s)$ can then be treated in one dimension quantum mechanically as the scattering of a matter wave with kinetic energy ε and reduced mass m. It is emphasized that this treatment amounts to a semi-classical picture of the scattering process and does not assume the form of a full quantum mechanical scattering treatment of a spherical wave.

There are two possible scenarios if a wave moving along ray $\boldsymbol{r}(s)$ interacts with the delta-potential $V_\delta(\boldsymbol{r}_i)$ at position $\boldsymbol{r}_i$. First, there is a probability for the wave being transmitted. This probability is given by the transmission coefficient T. Second, there is a chance for the wave being reflected. This probability is expressed by the reflection coefficient R with $T + R = 1$.

Since any encounter between the analyte ion and buffer gas particle during which their relative motion is altered counts as a 'collision', the LCPA defines the local collision probability $\boldsymbol{\tau}(\varepsilon, \boldsymbol{r})$ identical to the reflection coefficient R, i.e. $\boldsymbol{\tau}(\varepsilon, \boldsymbol{r}) \equiv R$. Thus, the LCPA expresses the local collision probability $\boldsymbol{\tau}(\varepsilon, \boldsymbol{r})$ as the probability R to observe a reflection when a particle wave with kinetic energy ε and reduced mass m interacts with the locally acting delta-potential $V_\delta(\boldsymbol{r})$:

$$\boldsymbol{\tau}(\varepsilon, \boldsymbol{r}) \equiv R = \frac{1}{1 + \frac{1}{\beta(\varepsilon, \boldsymbol{r})^2}} \tag{S8}$$

Here, $\beta(\varepsilon, \boldsymbol{r})$ is given by

$$\beta(\varepsilon, \boldsymbol{r}) = \frac{m}{\hbar^2} \frac{g}{k(\varepsilon, m)} \tag{S9}$$

where $k(\varepsilon, m)$ is the wave number of the scattered particle with kinetic energy ε and reduced mass m

$$k(\varepsilon, m) = \frac{1}{\hbar} \sqrt{2m\varepsilon} \tag{S10}$$

and g is the area under the delta-potential $V_\delta(\boldsymbol{r}) = g \cdot \delta(\boldsymbol{r})$ given by

$$g = l \cdot U(\boldsymbol{r}) \tag{S11}$$

As noted above, the interaction potential $U(\boldsymbol{r})$ has the meaning of the height and the 'length scale' l the meaning of the length of the delta-potential (*viz.* $\mathrm{area} = \mathrm{length} \cdot \mathrm{height}$). It is stressed that the length l of the delta-potential $V_\delta(\boldsymbol{r})$ is not a true distance in position space but rather corresponds to the length or duration of the interaction between the ion and the buffer particle (see section S1.1 above). That is, the length-scale l should decrease with increasing kinetic energy ε of the particle and

increase with increasing de-Broglie wavelength $\lambda(\varepsilon, m)$ of the particle. Thus, l can be written as

$$l = c \cdot \lambda(\varepsilon, m) \tag{S12}$$

where c is a proportionality constant.

For the limit of long-wavelength particles with $\lambda(\varepsilon, m) \gg l$ (i.e. $c \ll 1$, corresponding to *very cold particles*), the reflection coefficient R goes towards unity independent of the height $U(\boldsymbol{r})$ of the delta-potential $V_\delta(\boldsymbol{r})$. For the limit of asymptotically short-wavelength particles with $\lambda(\varepsilon, m) \to 0$ (i.e. $c \gg 1$, corresponding to *very hot particles*), the reflection coefficient R goes towards zero independent of the height $U(\boldsymbol{r})$ of the delta-potential $V_\delta(\boldsymbol{r})$. For finite, yet small $\lambda(\varepsilon, m) \ll l$, the scattering process becomes dominantly classical and the reflection coefficient R for scattering off a delta-potential $V_\delta(\boldsymbol{r})$ will be either zero (for $\varepsilon \leq U(\boldsymbol{r})$) or unity (for $\varepsilon > U(\boldsymbol{r})$). Consequently, the length scale l of the delta-potential is set equal to the de-Broglie wavelength $\lambda(\varepsilon, m)$ of the scattered particle (i.e. $c = 1$). Thus, the area g of the delta-potential $V_\delta(\boldsymbol{r}) = g \cdot \delta(\boldsymbol{r})$ is expressed in the LCPA as

$$g = g(\boldsymbol{r}, \varepsilon, m) = \lambda(\varepsilon, m) \cdot U(\boldsymbol{r}) \tag{S13}$$

Where $\lambda(\varepsilon, m)$ is the de-Broglie wavelength of the particle and $U(\boldsymbol{r})$ is the height of the interaction potential at position $\boldsymbol{r}$. Using the relationship $\lambda(\varepsilon, m) = {}^{2\pi}/_{k(\varepsilon, m)}$, the area g can then further be expressed as

$$g = g(\boldsymbol{r}, \varepsilon, m) = U(\boldsymbol{r}) \cdot \lambda(\varepsilon, m) = \frac{2\pi \cdot U(\boldsymbol{r})}{k(\varepsilon, m)} \tag{S14}$$

Thus, $\beta(\varepsilon, \boldsymbol{r})$ from equation S9 reduces to

$$\beta(\varepsilon, \boldsymbol{r}) = \frac{m}{\hbar^2}\frac{2\pi U(\boldsymbol{r})}{k(\varepsilon)^2} = \frac{m}{\hbar^2}\frac{2\pi U(\boldsymbol{r})}{1}\frac{\hbar^2}{2m\varepsilon} = \frac{\pi U(\boldsymbol{r})}{\varepsilon} \tag{S15}$$

And with equation (S15) inserted into equation (S8) it follows that $\tau(\varepsilon, \boldsymbol{r})$ is given by

$$\tau(\varepsilon, \boldsymbol{r}) = \frac{1}{1 + \left[\frac{\varepsilon}{\pi \cdot U(\boldsymbol{r})}\right]^2} \tag{S16}$$

This is the expression for $\tau(\varepsilon, \boldsymbol{r})$ used in the main text.

The treatment of the scattering process as a superposition of localized, independent scattering event occurring along a ray cast onto the analyte is certainly an approximation and assumes that the momentum transferred does not significantly depend on the exact path of the analyte – buffer gas trajectory. Thus, this treatment is appropriate for locally convex molecules.[1] The computed cross section must be corrected by a shape factor for non-convex molecules,[1,2] because here the analyte and buffer gas particle can interact multiple times during the collision process (see section S1.3 below).

1.3 The nature of the shape factor in the LCPA method

The momentum transfer cross section Ω of an analyte can be written as

$$\Omega = \sigma \cdot \rho \tag{S17}$$

Here, σ is the projection cross section of the analyte and the shape factor $\rho = {}^{\Omega}/_{\sigma}$ is really just the ratio between the momentum transfer and the projection cross sections. The momentum transfer cross section is identical to the projection cross section, i.e. ${}^{\Omega}/_{\sigma} = 1$, for any (locally) convex surface.[1] By contrast, for non-convex surfaces ${}^{\Omega}/_{\sigma} > 1$. As pointed out,[1] ${}^{\Omega}/_{\sigma} > 1$ for non-convex surfaces because here the analyte and

buffer gas particle can interact multiple times during the collision process. In other words, $\rho = \Omega/\sigma$ is a measure of how concave a given surface is. One intuitive way of expressing the degree of concaveness of a surface is to take the ratio of the areas of the surface and the convex hull of the surface,[2,3] *viz.*

$$\rho = \frac{\Omega}{\sigma} \approx \frac{\text{area of surface}}{\text{area of convex hull}} \tag{S18}$$

Estimating the shape factor through equation S18 has been successful for many molecular systems including carbon clusters, supramolecular coordination complexes, DNA, peptides, and proteins.[2–9] This pertains in particular to the molecules used for benchmark calculations in the current manuscript. However, the above approach fails, for example, for extended structures for high charge states of proteins under denaturing conditions in electrospray mass.[10] These shortcomings of equation S18 can be accounted for by replacing the area of the convex hull by the area of the fully enclosing surface with minimal surface area.[3]

The significance of the shape factor is that ρ can be fairly easily estimated for any given surface and thus allows to efficiently yet accurately estimate a momentum transfer cross section Ω for a non-convex surface by multiplication with the projection cross section σ.

1.4 Algorithmic details for a reference implementation to compute cross sections by the LCPA

In the following, a step-by-step description of a reference implementation of the LCPA method is provided. It is emphasized that these details are given here only for the sake of completeness and that the implementation of the LCPA method is not expected to follow these steps in the future.

1. A potential energy function $U(\boldsymbol{r})$ is defined for an input geometry which describes the interaction between the molecular analyte ion and the buffer gas particle at position $\boldsymbol{r}$. Currently, the potential $U(\boldsymbol{r})$ is constructed from atom-atom potentials as described:[11]

$$U(\boldsymbol{r}) = \sum_{l=1}^{atoms} u_l(\boldsymbol{r}) + V(\boldsymbol{r}) \quad \text{(S19)}$$

Specifically, the functions $u_l(\boldsymbol{r})$ can be identical to a Lennard-Jones 12,6 potential with parameters *E* and r_m centered at the origins of the atoms $\boldsymbol{R}_l$:

$$u_l(\boldsymbol{r}) = E\left[\left(\frac{r_m}{|\boldsymbol{r} - \boldsymbol{R}_l|}\right)^{12} - \left(\frac{r_m}{|\boldsymbol{r} - \boldsymbol{R}_l|}\right)^{6}\right] \quad \text{(S20)}$$

The function $V(\boldsymbol{r})$ can be used to described the charge-induced interaction potential for a buffer gas particle with polarizability α, as described.[11]

2. The potential $U(\boldsymbol{r})$ is evaluated in three dimensions on a grid $\mathcal{G}$ with elements $g_{x,y,z}$.
3. Given the temperature T of the buffer gas, a range of energies are selected for evaluation of the collision probability $\tau(\varepsilon, \boldsymbol{r})$ from the Boltzmann distribution.
4. The collision probability function $\tau(\varepsilon, \boldsymbol{r})$ is then computed for each grid element $g_{x,y,z}$ by Equation S16, for each energy ε.
5. The orientation-averaged cross section $\xi[\tau(\varepsilon, \boldsymbol{r})]$ is computed by casting rays $\vec{r}(s)$ onto the grid as follows:
    a) A plane *K* with normal vector $\overrightarrow{z_k}$ is randomly chosen.
        i. For each plane *K*, an area *A* is defined such that *A* fully contains the projection of the grid $\mathcal{G}$ on plane *K*.

ii. Rays $\vec{r}(s)$ with $\vec{r}(s) = \overrightarrow{r_A} + s \cdot \overrightarrow{z_k}$ originating from a point $\overrightarrow{r_A}$ within plane *K* with normal vector $\overrightarrow{z_k}$ are successively cast onto the grid $\mathcal{G}$.

iii. Points $\vec{r}(s)$ (i.e. grid elements $\mathcal{g}_{x,y,z}$) for which $p \leq \tau(\varepsilon, \overrightarrow{r_A} + s \cdot \overrightarrow{z_k})$, where *p* is a random number, is true are used to define the vertices of a collision surface $\mathcal{S}$ with $\mathcal{S} = \{\vec{r}(s) | p \leq \tau(\varepsilon, \vec{r}(s))\}$. A collision is said to occur at point $\vec{r}(s)$ if $\vec{r}(s) \in \mathcal{S}$. See Figure S2A for an example.

iv. Steps i) to iii) are repeated for every ray $\vec{r}(s)$ that can be define on the grid $\mathcal{G}$ starting from plane *K*.

b) For each plane *K*, an axis-aligned cross section $\sigma(\varepsilon, \overrightarrow{z_k})$ for kinetic energy **ε** is computed by:

$$\sigma(\varepsilon, \overrightarrow{z_k}) = A \frac{n_{coll}}{n_{coll} + n_{miss}} \tag{S21}$$

Here, *A* is the area of the bounding box in a plane *K* with normal vector $\overrightarrow{z_k}$ that fully contains the molecule. The quantities n_{coll} and n_{miss} denote the number of rays $\vec{r}(s) = \overrightarrow{r_A} + s \cdot \overrightarrow{z_k}$ that are considered a "collision" (n_{coll}) and "no collision" (n_{miss}), respectively, according to the Monte Carlo criterion above.

c) The orientation-averaged cross section $\xi[\tau(\varepsilon, \vec{r})]$ for kinetic energy **ε** is computed as the asymptotic average of orientation-aligned cross sections $\sigma(\varepsilon, \overrightarrow{z_k})$ for kinetic energy **ε** as:

$$\xi[\tau(\varepsilon, \vec{r})] = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \sigma(\varepsilon, \overrightarrow{z_k}) \tag{S22}$$

d) Steps a) to c) are repeated until the series converges to a defined threshold (typically 1%).

6. The shape-factor $\rho\left[\mathcal{S}[\tau(\varepsilon,\vec{r})]\right]$ is then computed from the collision surface constructed above according to:

$$\rho\left[\mathcal{S}[\tau(\varepsilon,\vec{r})]\right] = \frac{A(\mathcal{S}[\tau(\varepsilon,\vec{r})])}{C(\mathcal{S}[\tau(\varepsilon,\vec{r})])} \tag{S23}$$

Here, $C(\mathcal{S}[\tau(\varepsilon,\vec{r})])$ denotes the area of the convex hull of the collision surface $\mathcal{S}[\tau(\varepsilon,\vec{r})]$ and $A(\mathcal{S}[\tau(\varepsilon,\vec{r})])$ is computed according to:

$$A(\mathcal{S}[\tau(\varepsilon,\vec{r})]) = \lim_{\alpha\to 0} \mathrm{A}(\alpha, \mathcal{S}[\tau(\varepsilon,\vec{r})]) \tag{S24}$$

Here, $\mathrm{A}(\alpha, \mathcal{S}[\tau(\varepsilon,\vec{r})])$ is the exposed surface area of a Delaunay triangulation of momentum transfer surface $\mathcal{S}[\tau(\varepsilon,\vec{r})]$ in which all tetrahedrons with side lengths greater than α have been disregarded. See Figures S2B and S2C for examples. The shape factor can be improved by replacing the area of the convex hull $C(\mathcal{S}[\tau(\varepsilon,\vec{r})])$ by the minimum value of the exposed surface area $\mathrm{A}(\alpha, \mathcal{S}[\tau(\varepsilon,\vec{r})])$.[3]

7. The momentum transfer cross section $\Omega(T)$ can then be obtained according to:

$$\Omega(T) = \int_0^\infty f(\varepsilon,T) \cdot \xi[\tau(\varepsilon,\vec{r})] \cdot \rho\left[\mathcal{S}[\tau(\varepsilon,\vec{r})]\right] d\varepsilon \tag{S25}$$

Here, $f(\varepsilon,T)$ is the molecular Boltzmann factor; $\xi[\tau(\varepsilon,\vec{r})]$, $\mathcal{S}[\tau(\varepsilon,\vec{r})]$, and $\rho\left[\mathcal{S}[\tau(\varepsilon,\vec{r})]\right]$ are the orientation-averaged cross section, collision surface, and shape-factor for kinetic energy ε, respectively.

1.5 Examples of the collision surface

2 Comparison of LCPA, PSA, and trajectory method calculations on model compounds in helium

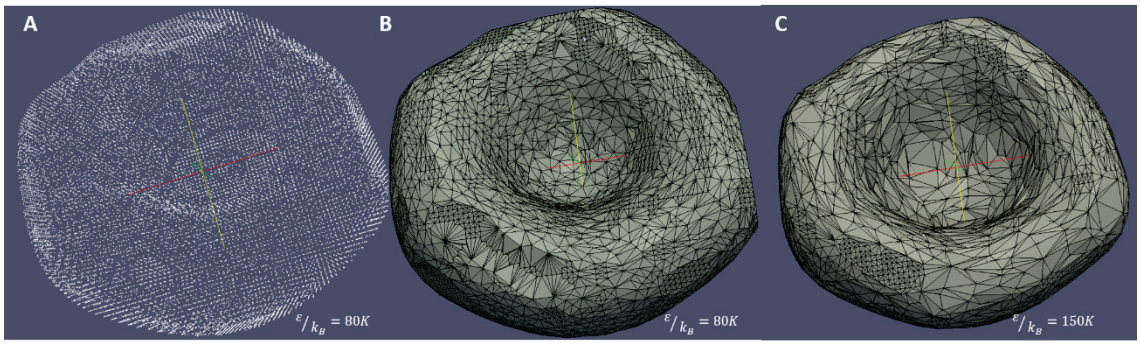

Figure S2. Generation of the collision surface $\mathcal{S}$ used to compute the microcanonical cross section $\Omega(\varepsilon)$ in the LCPA method. (A) Set of points comprising the vertices of $\mathcal{S}$ for the cup of C_{960}. These points have been generated by repeatedly casting rays on $\tau(\varepsilon, \boldsymbol{r})$, where $\tau(\varepsilon, \boldsymbol{r})$ was computed on a grid $\mathcal{G}$ with 0.15 Å spacing for an energy $\varepsilon/k_B = 80K$. (B) Collision surface $\mathcal{S}$ obtained from the set of points depicted in (A) after mesh generation. Surface reconstruction was achieved as described above. (C) Reconstructed collision surface for energy $\varepsilon/k_B = 150K$. The surface is significantly altered by superposition effects of the two-body Lennard-Jones interactions.

buffer gas

2.1 Cup and Sphere of C_{60}

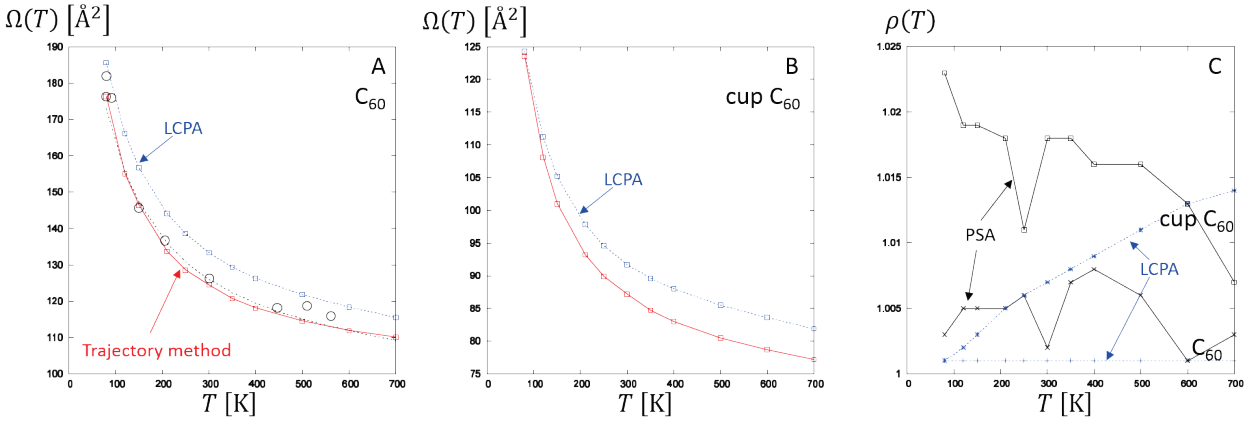

Figure S3. Predicted cross sections $\Omega(T)$ and shape factors $\rho(T)$ for C_{60} and the corresponding cup as a function of temperature. (A) Cross section $\Omega(T)$ of C_{60} predicted by the LCPA and trajectory methods using the same interaction potential.[11] LCPA cross sections using an optimized potential (dotted). Experimental cross sections[12] are plotted as circles. (B) Cross sections $\Omega(T)$ predicted for the cup of C_{60} by the LCPA and trajectory methods using the same interaction potential. (C) Shape factors $\rho(T)$ predicted by the LCPA and PSA methods.

2.2 Cup and Sphere of C_{180}

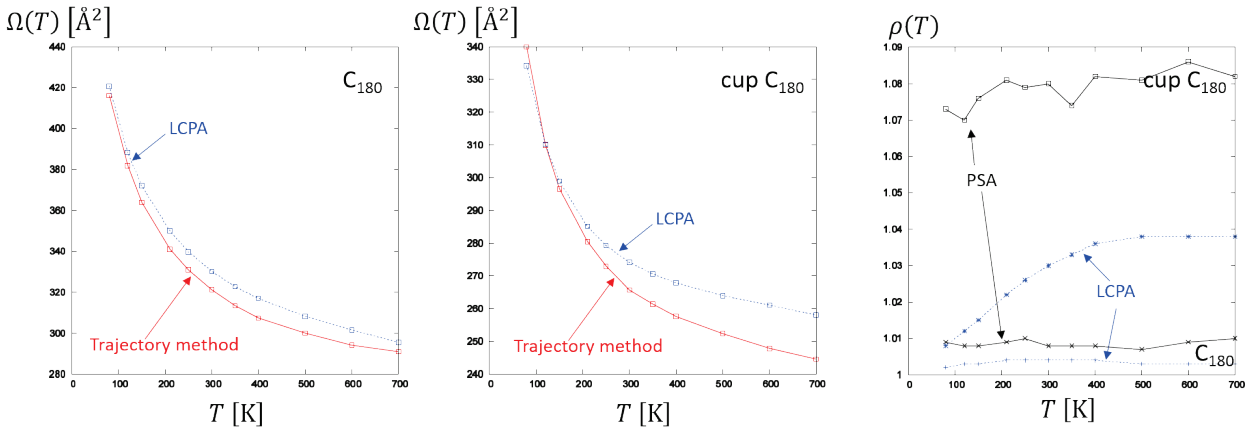

Figure S4. Predicted cross sections $\Omega(T)$ and shape factors $\rho(T)$ for C_{180} and the corresponding cup as a function of temperature. (A) Cross section $\Omega(T)$ of C_{180} predicted by the LCPA and trajectory methods using the same interaction potential. [11] (B) Cross sections $\Omega(T)$ predicted for the cup of C_{180} by the LCPA and trajectory methods using the same interaction potential. (C) Shape factors $\rho(T)$ predicted by the LCPA and PSA methods.

2.3 Cup and Sphere of C_{240}

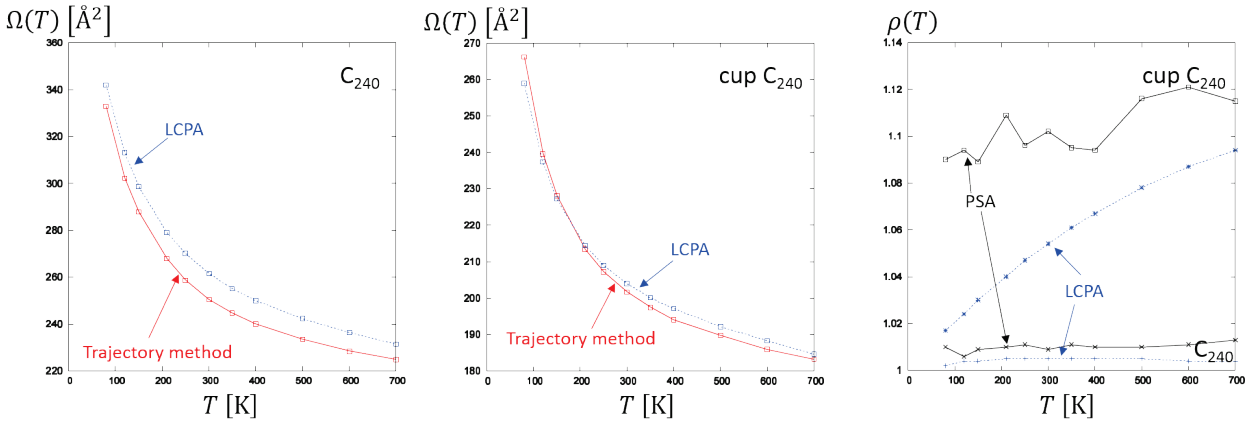

Figure S5. Predicted cross sections $\Omega(T)$ and shape factors $\rho(T)$ for C_{240} and the corresponding cup as a function of temperature. (A) Cross section $\Omega(T)$ of C_{240} predicted by the LCPA and trajectory methods using the same interaction potential.[11] (B) Cross sections $\Omega(T)$ predicted for the cup of C_{240} by the LCPA and trajectory methods using the same interaction potential. (C) Shape factors $\rho(T)$ predicted by the LCPA and PSA methods.

2.4 Statistical analysis

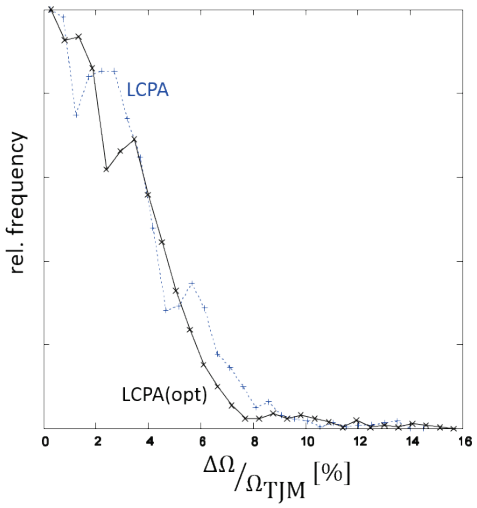

Figure S6. Relative difference in cross section $\Omega^{\mathrm{LCPA}} - \Omega^{\mathrm{TJM}}/\Omega^{\mathrm{TJM}}$ of LCPA and trajectory method (TJM) cross sections for helium for 208 benchmark systems in the temperature range from 80 to 700 K. Two interaction potentials were used to compute LCPA cross sections. The LCPA method achieves 90% and 95% quantiles of 6.3% and 7.1%, respectively, using the interaction potential fitted for use with the trajectory method ("LCPA", dotted/blue, E=0.0309 kcal/mol and r_{min}=3.405 Å). If the LCPA-optimized potential (E_{opt}=0.0347 kcal/mol and $r_{\mathrm{opt,min}}$=3.21 Å) is used, 90% and 95% quantiles of 4.9% and 6.3%, respectively, are obtained.

2.5 Comparison of Carbon-Helium Lennard-Jones potentials (graphite, trajectory, LCPA method)

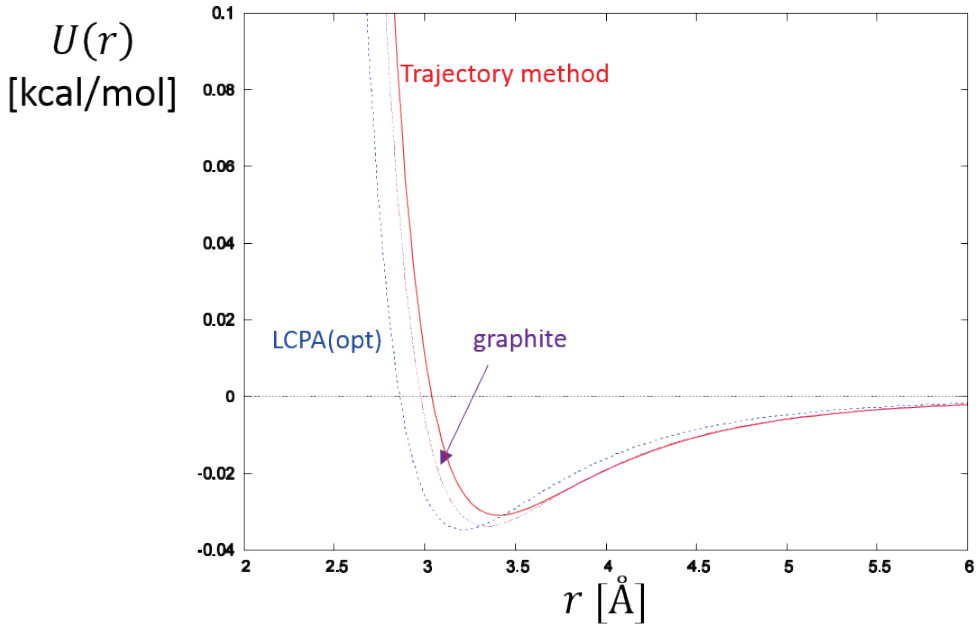

Figure S7. Lennard-Jones potentials for C_{60}-helium interactions. Plotted are the optimized potentials for the LCPA and trajectory[11] methods and also the potential deduced from the interlayer interactions in graphite.[11] The

carbon parameters optimized for the LCPA method are E_{opt}=0.0347 kcal/mol and $r_{\mathrm{opt,min}}$=3.21 Å, obtained from Monte-Carlo optimization for 100 iterations. Note the agreement between the different potentials.

3 Comparison of LCPA to the PSA and trajectory method calculations on model compounds for nitrogen buffer gas

In the following section, the performance of the LCPA method with respect to calculating cross sections in nitrogen buffer gas is assessed. To this end, nitrogen cross sections calculated by the LCPA, PSA, and trajectory methods are compared to each other. A few notes of caution are warranted at this point in order to appropriately interpret the results.

Trajectory method cross sections can be used as the reference standard when discussing the performance of the LCPA method for helium gas, mainly for four reasons. First, the trajectory method is a straight-forward implementation of the classical scattering process.[11] Second, the helium – analyte interaction potential has only a very shallow attractive well[13] and can therefore be described reasonably well by a Lennard-Jones potential. Third, the reported Lennard-Jones parameters can be taken as rather accurate because trajectory method calculations reproduce the temperature-dependency of the Buckminster-fullerene C_{60} cross sections reasonably well (see also Figures 2 in the main text and Figure S3).[2] Fourth, long-standing experience in using the trajectory method for ion mobility studies have shown that the trajectory method is often accurate when compared to the experiment.

For nitrogen, the situation is different. While the trajectory method itself is still a straight-forward implementation of the classical scattering process (#1 above), the validity of using a Lennard-Jones-like potential for nitrogen is in question due to the significant polarizability of nitrogen (#2 above).[13] Further, the exact representation of the analyte-

nitrogen interaction potential as implemented in the nitrogen-modification of the trajectory method[14,15] includes a number of modifications or additional parameters. One example is the use of a simple point-charge model to represent the quadrupole moment of nitrogen.[14] Another example is the use of scaling factors, which were fitted to the experiment on computed structures and atomic charges.[15] These changes were not rigorously assessed and tested as to their possible limitations (such as type, size, or charge state of the analyte). For example, distributed multipole methods are known to more accurately represent electric multipole moments than point charge models.

Further, accurate parameterization of an interaction potential for use with cross section algorithms must be based on temperature-dependent measurement of cross sections for analytes for which the structures are exactly known. This is so because only the temperature-dependency of the cross section can be used to invert back to the analyte-ion interaction potential,[16] and rigorously only if uncertainties in the structure are negligible. The helium parameters used in the trajectory method can be considered reliable, specifically because the temperature-dependency of the C_{60} cross sections are accurately reproduced (see Figure 3 in the main text and Figure S2).[2]

However, the nitrogen parameters for the trajectory method were obtained based on single-temperature measurements and calculated structures and atomic charges for calculations at 301 K.[15] Further, no study was reported on how well the nitrogen parameters describe the temperature-dependency of C_{60} cross sections (or of other rigid compounds whose structure is similarly well-characterized), see #3 above. Thus, it is not clear how broadly applicable the reported parameters really are with respect to differences in temperature or analyte size, charge, or shape. Further, since the nitrogen

implementation of the trajectory is a relatively recent development, comparison to experimental data is limited (#4 above).[8,15,17]

For these reasons, trajectory method cross sections computed for nitrogen buffer gas should be taken with some grain of salt and are not as well-suited as are their helium equivalents for use as reference cross sections to benchmark other algorithms. Consequently, an accurate and thorough comparison between the LCPA and trajectory methods for nitrogen cannot be accomplished here.

However, in order to assess *whether or not the LCPA method shows promise to compute cross sections in a number of electronically different buffer gases*, LCPA nitrogen cross sections are compared to PSA and trajectory method cross sections. Trajectory method cross sections were calculated using the recent modification for use with nitrogen buffer gas.[15] The comparison was restricted to temperatures ranging from 250 K to 400 K to reflect the parameterization of the trajectory method parameters at 301 K. Nitrogen parameters were estimated for the LCPA method by the combination rules described by Lorentz[18] and Berthelot[19] from the published helium and nitrogen parameters.[11,13] The resulting set of nitrogen Lennard-Jones parameters for the LCPA method are E=0.09153 kcal/mol and $r_{\min}$=4.01 Å (for details see Table S1). It is important to point out that the nitrogen-carbon interaction potentials used in the LCPA and trajectory methods differ from each other. Consequently, the LCPA and trajectory methods are not expected to predict identical cross sections, and nitrogen cross sections are thus discussed more in a qualitative than quantitative manner here. PSA cross sections were computed as described.[20]

Table S1. Nitrogen Lennard-Jones parameters for the LCPA method

	Lennard-Jones Parameter	
	E [eV]	σ [Å]
He-He[8]	0.00088	2.556
C-He[1]	0.00133	3.013
C-C[a]	0.00200	3.469
N_2-N_2[8]	0.00788	3.681
C-N_2[a]	0.00397	3.575
	E [kcal/mol]	$r_{\min}$ [Å]
C-N_2	**0.09153**	**4.013**

[a] from combination rules:[6,7]

$$\sigma_{X-Y} = \frac{\sigma_{X-X} + \sigma_{Y-Y}}{2}$$

$$E_{X-Y} = \sqrt{E_{X-X} \cdot E_{Y-Y}}$$

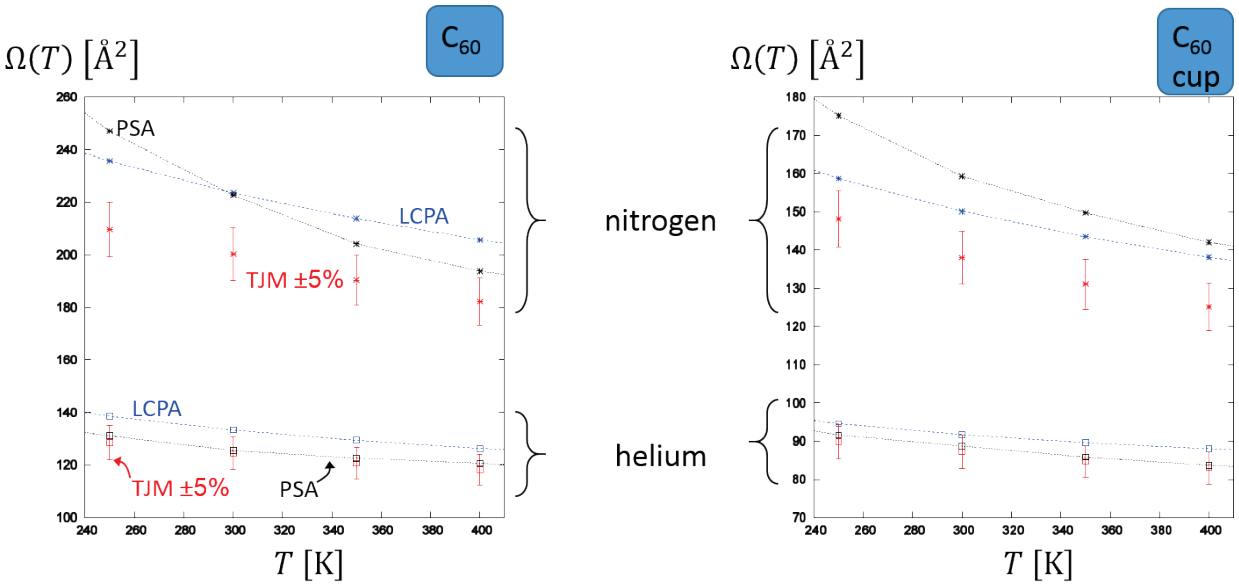

Figure S8. Comparison of cross sections predicted for C_{60} and its cup by the PSA (black), LCPA (blue), and trajectory (red) methods for nitrogen and helium buffer gas in the temperature range from 250 K to 400 K. A ±5% margin is indicated for the trajectory method. Comparison of the cross sections show that all methods agree with each other for helium (for details, see main text and reference 2). For nitrogen, LCPA cross sections are approximately 7 to 13% larger than those predicted by the trajectory method. PSA nitrogen cross sections are approximately 5 – 20 % larger than the corresponding trajectory values. Thus, although the LCPA and trajectory nitrogen cross sections do not agree with each other as well as for helium, the methods agree with each other qualitatively on the increase in cross section from helium to nitrogen buffer gas. Moreover, since a quantitative agreement between the different methods for nitrogen is not expected in the first place, it is more important to look at the trends visible in the Figure. For C_{60}, the slope $^{\Delta\Omega(T)}/_{\Delta T}$ of the nitrogen PSA cross sections is almost twice as large as the slope of the trajectory method values. This observation suggests that the use of the PSA nitrogen parameters may be restricted to charge states or compound classes that were used in the parametrization procedure. By contrast, LCPA cross sections are consistently 12 – 13 % larger than the trajectory method values for 250, 300, and 400 K, which indicates that the corresponding slopes $^{\Delta\Omega(T)}/_{\Delta T}$ are similar in that temperature range. The LCPA cross sections decrease by 30 Å^2 and 21 Å^2 for C_{60} and the cup, respectively, when increasing the temperature from 250 K to 400 K. This decrease is consistent with the decrease in cross section predicted by the trajectory method (28 Å^2 and 23 Å^2, respectively). This observation is significant, because it indicates that the LCPA method is capable to reproduce the differences between helium and nitrogen cross sections predicted by the trajectory method simply by estimating the nitrogen – carbon potential from the helium values using combination rules.

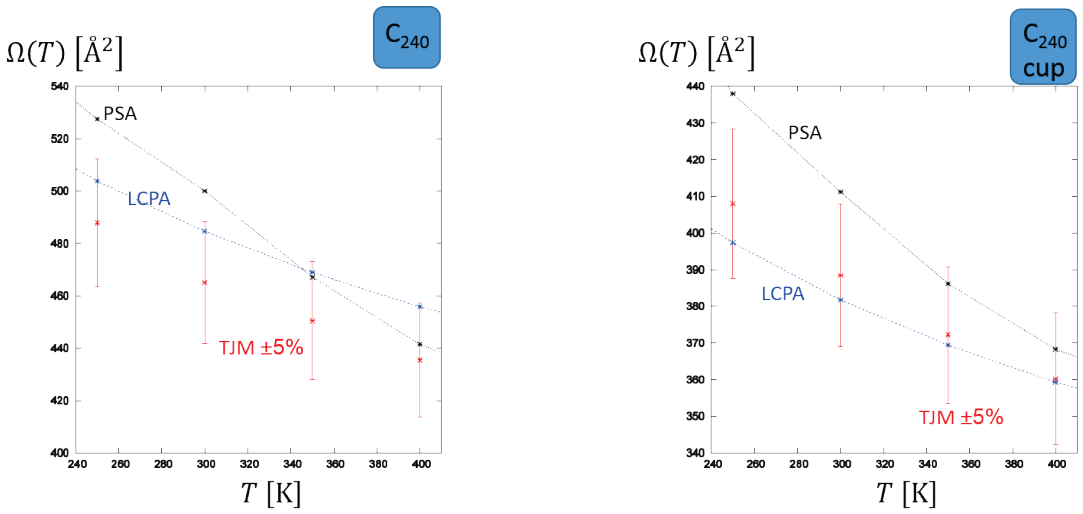

Figure S9. Comparison of cross sections predicted for C_{240} and its cup by the PSA (black), LCPA (blue), and trajectory (red) methods for nitrogen buffer gas in the temperature range from 250 K to 400 K. A ±5% margin is indicated for the trajectory method. The LCPA cross sections deviate from the trajectory method values between 3 and 5 % (C_{240}) and between 0 and 3 % (cup). The PSA cross sections are between 1 and 8 % (C_{240}) and 2 and 7 % (cup) larger than the trajectory method values. In the temperature range studied, the changes $\Delta\Omega(T)$ of the LCPA method (-48 and -38 Å^2 for C_{240} and the cup, respectively) are consistent with the changes of the trajectory method (-52 and -48 Å for C_{240} and the cup, respectively), while the changes of the PSA cross sections are much larger (-86 and -70 Å for C_{240} and the cup, respectively).

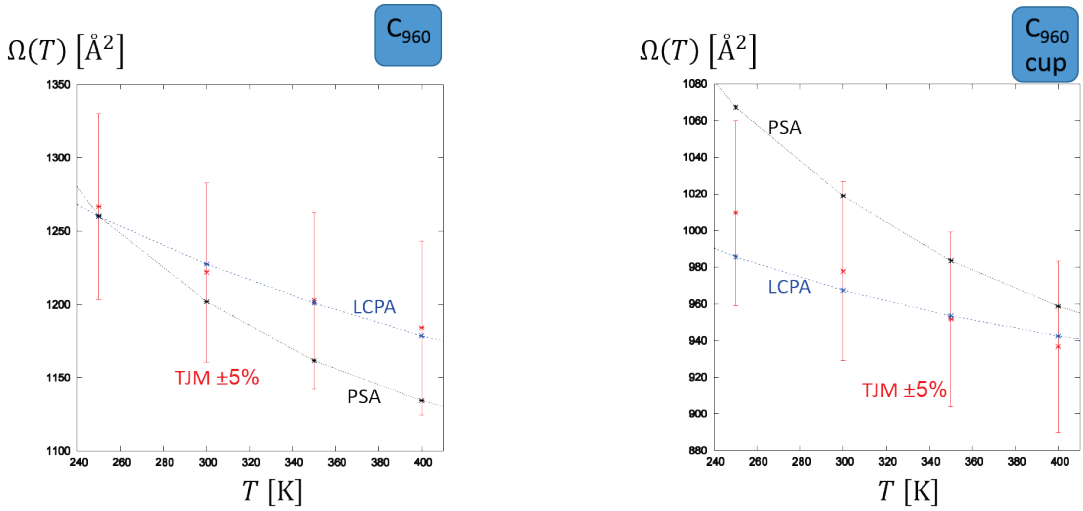

Figure S10. Comparison of cross sections predicted for C_{960} and its cup by the PSA (black), LCPA (blue), and trajectory (red) methods for nitrogen buffer gas in the temperature range from 250 K to 400 K. A ±5% margin is indicated for the trajectory method. The LCPA cross sections deviate from the trajectory method values between 0 and 1% (C_{960}) and between 1 and 2% (cup). The PSA cross sections are between 1 and 4% (C_{960}) and 2 and 6% (cup) larger than the trajectory method values. In the temperature range studied,

the changes in cross section $\Delta\Omega(T)$ predicted by the LCPA method are again consistent with the changes predicted by the trajectory method.

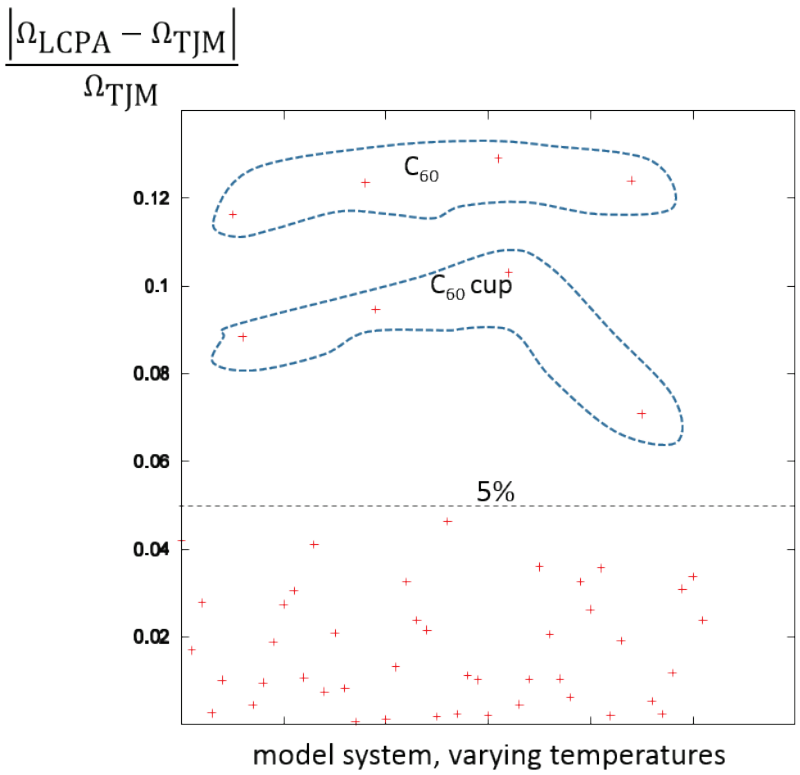

Figure 11. Relative deviation $\frac{|\Omega_{LCPA}-\Omega_{TJM}|}{\Omega_{TJM}}$ between the LCPA and trajectory method cross sections, Ω_{LCPA} and Ω_{TJM}, respectively, for temperatures ranging from 250 to 400 K for the fullerenes C_{60}, C_{160}, C_{240}, C_{540}, C_{960}, and their corresponding cups. The figure shows that the LCPA and trajectory method cross sections agree with one another within 5%. The two exceptions are the clusters C_{60} and its cup, where the two methods deviate from each other in the order of 10%.

4 CPU times used for benchmark calculations in helium and nitrogen buffer gas (LCPA vs. PSA)

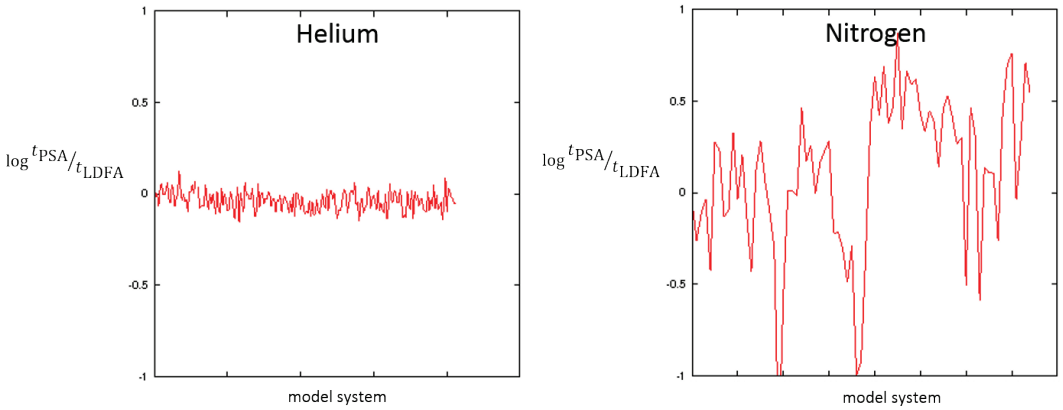

Figure S12. Ratio of CPU times used by the LCPA and PSA methods to compute the cross sections analyzed in this work for helium (left) and nitrogen (right) buffer gases. (left) The comparison of CPU times shows that the LCPA and PSA implementations are comparable in terms of the CPU time needed to calculate helium cross sections. Overall, the PSA method required approximately 71 hours of cumulative CPU time to finish the benchmark calculations. The LCPA method required approximately 77 hours to finish the calculations, exceeding the demand of the PSA by roughly 10%. (right) The comparison of CPU times shows that the PSA method is more demanding than the LCPA method in terms of the CPU time needed to calculate nitrogen cross sections. Overall, the PSA method required approximately 214 hours of cumulative CPU time to finish the benchmark calculations for nitrogen (see above). By contrast, the LCPA method required only approximately 94 hours to finish the calculations, only about 43% of the time required for the PSA calculations.

However, when interpreting these timings it should be kept in mind that the PSA method has matured over several years of development[2,7,5,6,3] while this is not the case for the LCPA method. Thus, overall, the data indicate that the LCPA promises to become a fast and accurate method to calculate collision cross sections. Trajectory method calculations are significantly more demanding than both LCPA and PSA calculations[2,8] and were omitted from the comparison for brevity. (For example, the nitrogen trajectory method took about 2664 hours to calculate the temperature-dependent cross sections for C_{960} whereas the LCPA method required 2.5 and the PSA method 5.5 hours.)

5 Examples of molecular surfaces used by the PSA method to compute shape factors

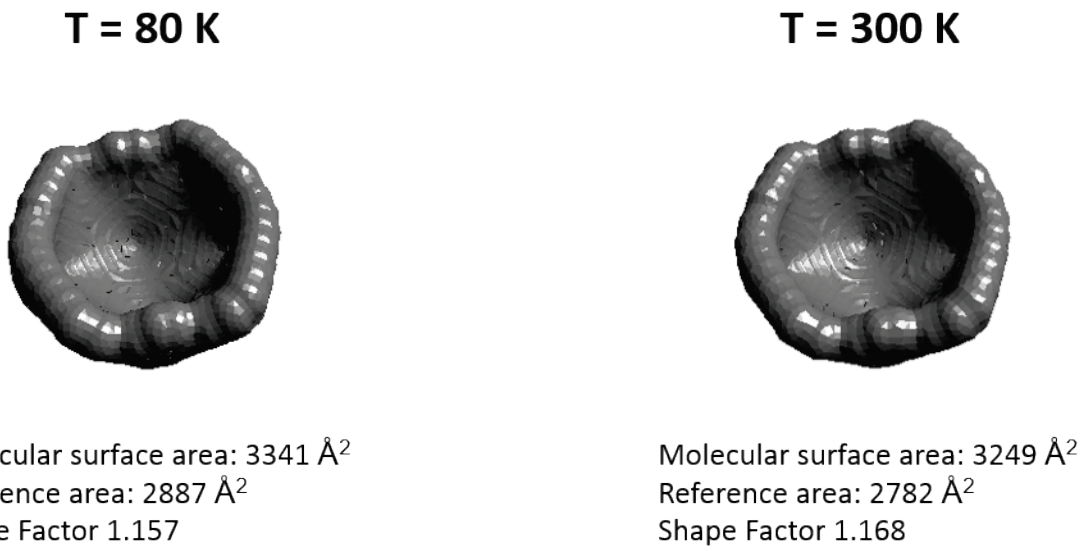

Figure S13. Molecular surfaces used in the PSA to compute the shape factor for the cup of C_{960}. (left) Temperature T=80 K. This surface corresponds to the center surface in Figure S2. (right) Temperature T=300 K.

References

1. Alexandre A. Shvartsburg and Martin F. Jarrold. An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.* **261,** 86–91 (1996).
2. Bleiholder, C., Wyttenbach, T. & Bowers, M. T. A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (I). Method. *Int. J. Mass Spectrom.* **308,** 1–10 (2011).
3. Wyttenbach, T., Bleiholder, C., Anderson, S. E. & Bowers, M. T. A new algorithm to characterise the degree of concaveness of a molecular surface relevant in ion mobility spectrometry. *Mol. Phys.* 1–6 (2015). doi:10.1080/00268976.2015.1042935

4. Wyttenbach, T., Bleiholder, C. & Bowers, M. T. Factors Contributing to the Collision Cross Section of Polyatomic Ions in the Kilodalton to Gigadalton Range: Application to Ion Mobility Measurements. *Anal. Chem.* **85,** 2191–2199 (2013).
5. Anderson, S. E., Bleiholder, C., Brocker, E. R., Stang, P. J. & Bowers, M. T. A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (III): Application to supramolecular coordination-driven assemblies with complex shapes. *Int. J. Mass Spectrom.* **330-332,** 78–84 (2012).
6. Bleiholder, C., Contreras, S. & Bowers, M. T. A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (IV). Application to polypeptides. *Int. J. Mass Spectrom.* **354-355,** 275–280 (2013).
7. Bleiholder, C., Contreras, S., Do, T. D. & Bowers, M. T. A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (II). Model parameterization and definition of empirical shape factors for proteins. *Int. J. Mass Spectrom.* **345-347,** 89–96 (2013).
8. Valentina D'Atri, Massimiliano Porrini, Frédéric Rosu, and Valérie Gabelica. Linking Molecular Models with Ion Mobility Experiments. Illustration with a rigid nucleic acid structure. *J. Mass Spectrom.* **50,** (2015).
9. Ganguly, P. *et al.* Tau Assembly: The Dominant Role of PHF6 (VQIVYK) in Microtubule Binding Region Repeat R3. *J. Phys. Chem. B* **119,** 4582–4593 (2015).
10. Goldstein, M., Zmiri, L., Segev, E., Wyttenbach, T. & Gerber, R. B. An atomistic structure of ubiquitin +13 relevant in mass spectrometry: Theoretical prediction and

comparison with experimental cross sections. *Int. J. Mass Spectrom.* **367,** 10–15 (2014).

11. Mesleh, M. F., Hunter, J. M., Shvartsburg, A. A., Schatz, G. C. & Jarrold, M. F. Structural information from ion mobility measurements: effects of the long-range potential. *J. Phys. Chem.* **100,** 16082–16086 (1996).
12. Wyttenbach, T., von Helden, G., Batka, J. J., Carlat, D. & Bowers, M. T. Effect of the long-range potential on ion mobility measurements. *J. Am. Soc. Mass Spectrom.* **8,** 275–282 (1997).
13. J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird. *Molecular Theory of Gases and Liquids*. (Wiley, 1954).
14. Kim, H. *et al.* Experimental and Theoretical Investigation into the Correlation between Mass and Ion Mobility for Choline and Other Ammonium Cations in N_2. *Anal. Chem.* **80,** 1928–1936 (2008).
15. Campuzano, I. *et al.* Structural Characterization of Drug-like Compounds by Ion Mobility Mass Spectrometry: Comparison of Theoretical and Experimentally Derived Nitrogen Collision Cross Sections. *Anal. Chem.* **84,** 1026–1033 (2012).
16. Mason, E. A. & McDaniel, E. W. *Transport Properties of Ions in Gases*. (Wiley, 1988).
17. Warnke, S. *et al.* Protomers of Benzocaine: Solvent and Permittivity Dependence. *J. Am. Chem. Soc.* **137,** 4236–4242 (2015).
18. H. A. Lorentz. Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. *Ann. Phys.* **248,** 127–136 (1881).

19. Daniel Berthelot. Sur le mélange des gaz. *Comptes Rendus Hebd. Séances L'Académie Sci.* **126,** 1703–1855 (1898).
20. Christian Bleiholder, Nicholas R. Johnson, Stephanie Contreras, Thomas Wyttenbach, and Michael T. Bowers. Molecular Structures and Ion Mobility Cross Sections: Analysis of the Effects of He and N2 Buffer Gas. *Anal. Chem.* **accepted,** (2015).