2017

Collidoscope: An Improved Tool for Computing Collisional Cross Sections with the Trajectory Method

Simon A. Ewing

Micah T. Donor

Jesse W. Wilson

James S. Prell

Collidoscope: An Improved Tool for Computing Collisional Cross Sections with the Trajectory Method

Simon A. Ewing[1], **Micah T. Donor**[1], **Jesse W. Wilson**[1], and **James S. Prell**[1,2,*]

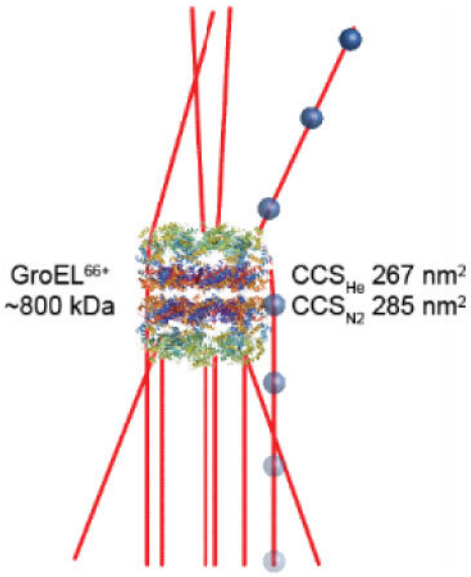
[1]Department of Chemistry and Biochemistry, University of Oregon, 1253 University of Oregon, Eugene, OR 97403-1253

[2]Materials Science Institute, University of Oregon, 1252 University of Oregon, Eugene, OR 97403-1252

Abstract

Ion Mobility-Mass Spectrometry (IM-MS) can be a powerful tool for determining structural information about ions in the gas phase, from small covalent analytes to large, unfolded, and/or denatured proteins and complexes. For large biomolecular ions, which may have a wide variety of possible gas-phase conformations and multiple charge sites, quantitative, physically explicit modeling of collisional cross sections (CCSs) for comparison to IMS data can be challenging and time-consuming. We present a “trajectory method” (TM) based CCS calculator, named “Collidoscope”, which utilizes parallel processing and optimized trajectory sampling, and implements both He and N_2 as collision gas options. Also included is a charge-placement algorithm for determining probable charge site configurations for protonated protein ions given an input geometry in pdb file format. Results from Collidoscope are compared to those from the current state-of-the-art CCS simulation suite, IMoS. Collidoscope CCSs are typically within 4% of IMoS values for ions with masses from ~18 Da to ~800 kDa. Collidoscope CCSs using x-ray crystal geometries are typically within a few percent of IM-MS experimental values for ions with mass up to ~3.5 kDa (melittin), and discrepancies for larger ions up to ~800 kDa (GroEL) are attributed in large part to changes in ion structure during and after the electrospray process. Due to its physically explicit modeling of scattering, computational efficiency, and accuracy, Collidoscope can be a valuable tool for IM-MS research, especially for large biomolecular ions.

Graphical Abstract



Introduction

Native Ion Mobility-Mass Spectrometry (IM-MS) can provide a wealth of information about the size, shape, charge state, mass, stoichiometry, and/or topology of ionized, native-like biomolecules and biomolecular complexes, in addition to other properties [1–30]. IM spectrometry separates ions upon passage through a buffer gas (typically helium, argon, or nitrogen) under the influence of an electric field as a result of the size- and conformation-dependent drag force induced by repeated collisions between the ion and buffer gas. In native IM-MS, ion mobility values in the low-field limit are typically measured using drift tube, traveling wave ion guide, or, more recently, trapped ion mobility instrumentation before subsequent mass spectrometry analysis [10, 31]. Structural information is typically obtained by converting measured ion mobility values to collisional cross sections (CCSs) and comparing these to other experimental data or CCSs computed for model structures. However, accurately computing CCSs with explicit computation of the scattering process can be very challenging and time-consuming due to contributions of long-range polarization interactions, orbiting, multiple scattering, and diffuse scattering, among other interactions. Due to the computational difficulty of modeling CCS in a physically explicit manner, ions with masses greater than ~10 kDa are commonly modelled using non-explicit methods (the “projection approximation”, PA [32–34]; the “projected superposition approximation”, PSA [35]; and the “local collision probability approximation”, LCPA [36]) or using hard-sphere collisions (the “exact” and “diffuse hard spheres scattering” methods, EHSS [33, 37] and DHSS [9, 38, 39]). Although physically explicit methods (the “trajectory method”, TM [33], and the “diffuse trajectory method”, DTM [9, 38, 39]) can be highly accurate, they can be prohibitively slow for large ions as currently implemented.

Generally, methods for approximating the CCS of an ion in the low-field limit balance computation time and the level of physical detail of the modeled collision events. Simple

methods (such as the PA and PSA) for calculating CCSs use projections of spherical model atoms to find the orientationally-averaged area of the ion's 2-dimensional silhouette. More complex methods model the ion in three dimensions but either do not explicitly model collision gas polarization effects (EHSS, DHSS) or 3-dimensional trajectories (LCPA). The most explicit currently-available method for calculating CCSs, the DTM, accounts for energy transfer between the collision particles and the ion using a hybrid of explicit classical scattering trajectories and stochastic modeling of inelastic scattering. Features of these various computational methods are summarized in Table 1.

The TM evaluates an ion's CCS by approximating the solution to the non-equilibrium Boltzmann Transport Equation [40, 41] as the momentum transfer integral $\Omega_{avg}^{(1,1)}$. A common choice of a potential energy surface for the TM is the sum of Lennard-Jones 6–12 (L-J 6–12) potential energy surface of each of the atoms in the ion for the buffer gas of interest, and, at added computational expense, an ion-induced dipole ("C_4") term [33, 39]. In principle, CCSs from the TM account for long-range scattering, multiple scattering, the ion's charge, and temperature effects. However, final computed CCSs for large biomolecules and assemblies using elastic collisions are typically about 5% higher than experimentally measured values [3]. This is believed to result in part from gas-phase compaction that occurs due to self-solvation of the biomolecular ion during the electrospray ionization process, in addition to any inherent errors in the approximation method, or in converting experimental drift times to CCSs [30].

Here, we introduce a new open-source program, Collidoscope, which approximates CCSs via the TM, using parallelized code to optimize performance. Collidoscope also contains an option to compute CCSs with N_2 as the buffer gas, using optimized L-J 6–12 parameters that result in CCSs closely matching experimental values. The essential features of the computational strategy and optimization procedure used for programming Collidoscope are described below, and results from Collidoscope are compared to experimental and other computational data. The utility of Collidoscope for interpreting structures of very large native-like biomolecular ions (with masses approaching ~1 MDa) is also discussed.

Theory

Physical Model

Collidoscope calculates an ion's CCS by approximating the temperature-dependent, orientationally-averaged momentum transfer term, $\Omega_{avg}(T)$, that arises in the solution to the Boltzmann Transport Equation. Collisions are simulated by integrating trajectories along the total ion-collision particle potential energy surface, approximated as the sum of L-J 6–12 potentials centered at each atom of the ion plus a "C_4" term to account for ion-induced dipole effects. According to classical scattering theory, $\Omega_{avg}(T)$ can be approximated in the low-field limit as the binary collision integral

$$\Omega_{avg}^{(1,1)}(T)=\frac{\int_0^\infty \int_0^{2\pi} \int_0^{\pi} \int_0^{2\pi} \int_0^\infty b\,(1-\cos(\chi_{b,\theta,\varphi,\gamma,g}))\,db\,d\gamma\,sin\varphi\,d\varphi\,d\theta * g^5 * e^{\frac{-\mu g^2}{2k_B T}} dg}{4\pi \int_0^\infty g^5 * e^{\frac{-\mu g^2}{2k_B T}} * dg} \quad (1)$$

where g is the initial speed of the collision particle relative to the ion; b is the impact parameter (defined as the initial distance of the approaching particle to the collision axis, see Figure 1); θ, φ, and γ are the angles which define the relative orientation of the ion and collisional particle; k_B is the Boltzmann constant; T is the absolute temperature; χ is the scattering angle; and μ is the reduced mass of the system [33, 40, 41]. (Note that the g^5 term reflects laboratory-frame orientational averaging of the particle and ion velocities, with probabilities given by Maxwell-Boltzmann distributions, as well as momentum transfer.)

In Equation 1, θ and φ define the direction of the collision axis, while γ describes the ion's relative rotation about that axis. Since the γ and b parameters together describe a plane perpendicular to the collision axis, the integral over $bdbd\gamma$ is equivalent to an integral over the differential area, dA. In addition, the double integral over $sin\varphi d\varphi d\theta$ can be written as an integral over differential solid angle, $d\omega$. Making the above substitutions results in the following equation:

$$\Omega_{avg}^{(1,1)}(T)=\frac{\int_0^\infty \iint \iint (1-\cos(\chi_{x,y,\theta,\varphi,g}))\;dA\,d\omega * g^5 * e^{\frac{-\mu g^2}{2RT}} dg}{4\pi \int_0^\infty g^5 * e^{\frac{-\mu g^2}{2RT}} * dg} \quad (2)$$

where dA is the differential area of a plane representing the origin of the collision particles before scattering (the "plane of origin", see Figure 1). In Equation 2, x and y describe the initial position of the collision particle in the plane of origin, while θ and φ describe the orientation of the plane of origin relative to the ion, and the denominator is a normalization factor. Collidoscope uses judicious choices of initial parameters to achieve uniform sampling of $d\omega$ and dA as well as a Maxwell-Boltzmann-weighted sampling of g. These sampling choices, as well as the choice of potential surface and trajectory integration method, are described in greater detail below.

Sampling of Relative Orientation

Orientational averaging is achieved using a user-defined set of "vantage points" that determine a set of collision axes and corresponding planes of origin for the scattering simulations. To accurately approximate the integral over $d\omega$ in Eq. 2, these vantage points should ideally be sampled uniformly over the sphere of possible ion orientations with high density, otherwise a poor estimate of $\Omega_{avg}(T)$ may be obtained. Vertices of regular or quasi-regular polyhedra are thus reasonable choices for the sets of vantage points.

Sampling of Impact Parameter

Within each plane of origin, trajectories are initiated from points arranged in a uniformly-spaced square grid. Thus, dA in Eq. 2 is approximated as the square of the distance between

neighboring grid points for each trajectory. Scattering trajectories are computed only for the grid points for which both plane-of-origin coordinates (*x* and *y* in Fig. 1) are smaller than the furthest distance of an atom in the ion to the center of mass, plus an additional "grid buffer" distance. Although the distance from the collision axis for which collision particles scatter strongly depends on the size and shape of the ion, use of a sufficiently large grid buffer distance ensures that significantly scattered trajectories are well-sampled at minimal computational expense.

For additional computational efficiency, Collidoscope performs a calculation to assess whether a trajectory within the region described above should be in included in the computation of $\Omega_{avg}^{(1,1)}$. The potential energy of a collision particle at its closest approach to any atom in the ion in the absence of any scattering is calculated. Scattering trajectories are then computed only if this potential energy is larger than a specified cut-off value, because they are likely to scatter significantly. Effectively, for each plane of origin, the trajectories included in the CCS calculation are initiated from a region that broadly resembles an expanded silhouette of the ion in that plane. This procedure for choosing trajectories to run results in more efficient computation. The computation time as a function of the potential energy cut-off is shown in Figure S-1.

Sampling of Relative Velocities

Initial collision speeds are sampled equally spaced, and the $\Omega_{avg}^{(1,1)}$ obtained for each collision particle kinetic energy is weighted by a probability factor to obtain $\Omega_{avg}^{(1,1)}(T)$ (see Eq. 2). Each initial kinetic energy state's contribution is weighted by a corresponding analytic integral of the probability density function from Eq. 2, $g^5 * \exp(\frac{-\mu g^2}{2k_B T})$, over a range of energies containing that initial kinetic energy state (see Figure 2). The number of initial kinetic energy states and the minimum and maximum initial kinetic energies used can be input by the user. For each sampled initial kinetic energy, the bounds of integration, which determine the weighting, are spaced equally in kinetic energy, with the exception that the lowest and highest bounds are 0 and infinity, respectively. This assures that $\Omega_{avg}^{(1,1)}(T)$, with this method of energy sampling, is always evaluated to the same value in the case that it is strictly independent of temperature. (Indeed, temperature dependence of CCSs appears to be very weak for ions with a fixed conformation containing at least a few dozen atoms over the range ~200–400 K; see Figure 2 and Supplementary Material, Figure S2). However, even the weak dependence of cross section on temperature necessitates more densely sampled parameters for temperatures far from standard temperature.

Model Potential Energy Surface

At present, modeling the polarization interactions quantum mechanically would require prohibitive computational expense even for small ions. Instead, we follow the general approach of MOBCAL [33], the IMoS suite [9, 38, 39], the PSA [35], and the LCPA [36] by approximating the potential energy surface as the sum of the L-J 6–12 potentials due to each atom in the ion, plus a "C_4" potential due to the polarizability of the collision particle and the localized charges on the ion:

$$U(r)=\sum_{atom}\varepsilon_i\left[\left(\frac{r_{min,i}}{r_i}\right)^{12}-2\left(\frac{r_{min,i}}{r_i}\right)^{6}\right]+\frac{\alpha e^2}{8\pi\varepsilon_0}\left(\left(\sum_{atom}z_i\frac{r_{x,i}}{{r_i}^3}\right)^2+\left(\sum_{atom}z_i\frac{r_{y,i}}{{r_i}^3}\right)^2+\left(\sum_{atom}z_i\frac{r_{y,i}}{{r_i}^3}\right)^2\right)$$

(3)

where ε_i and $r_{min,i}$ are the L-J 6–12 parameters for well-depth and radius at minimum potential between atom i and the collision particle, respectively, ε_0 is the permittivity of free space, α is the effective isotropic polarizability volume of the collision particle, e is the elementary charge, z_i is the charge state of each atom, and $\langle r_{x,i}, r_{y,i}, r_{z,i}\rangle$ is the vector between atom i and the collision particle (with length r_i). Collisions are modeled elastically, and the positions of the atoms and charges in the ions are held fixed throughout the scattering computations.

Ion-induced dipole interactions with He as the collision particle are modeled using its static electric dipolar polarizability [43]. N_2 is also modeled as a quasi-spherical collision particle, using an effective polarizability that is equal to the arithmetic mean of its principal polarizabilities determined by *ab initio* computations [44]. Based on its moment of inertia, over the duration of a typical simulated trajectory (tens of ps), a molecule of N_2 thermalized at room temperature would typically rotate about each axis a few dozen times. Thus, this quasi-spherical approximation assumes that the molecule is rotating rapidly enough that no significant rotational alignment effects occur due to collision particle polarization.

Integration of Trajectories

Integration of trajectories in time in Collidoscope is presently performed using Euler (i.e., first-order Runge-Kutta) integration. Therefore, in order to minimize calculation inaccuracy due to trajectory integration errors, multiple checks are performed to determine the numerical validity of a trajectory. The total energy of the system must be conserved to within a threshold over the course of a trajectory for it to be included in computing the CCS. To enforce this, the center-of-mass-frame kinetic and potential energy of the collision particle are calculated at the beginning and end of a trial trajectory. If the total energy changes by more than the allowed amount, the trajectory is recalculated with a shorter time step, and energy conservation is checked again. Using a smaller time step promotes stricter energy conservation, so re-calculated trajectories converge toward exact energy conservation as the time step is reduced.

If a particle loses energy due to integration error, it can occasionally become trapped near the ion by the attractive portion of the potential energy surface. Therefore, any trajectories exceeding a maximum number of steps are recalculated with a shorter time step. If a trajectory either does not conserve energy or exceeds the maximum number of time steps for five consecutive attempts, the trajectory is considered to have failed and is omitted from the CCS calculation.

Parallelization

Significantly reduced computation time is achieved in Collidoscope through parallelization of trajectory simulations. At present, many common computers have 12 parallel hardware threads on which calculations can be performed. Each thread can calculate independent trajectories, decreasing the wall time when using 12 threads to ~9% of the wall time needed when using one thread at a time. By default, Collidoscope automatically determines the number of hardware threads available on a computer and maximizes the number of trajectories that are run simultaneously. A version of Collidoscope which utilizes GPU parallelism will be released soon.

Ions Modeled

CCSs for molecules and ions ranging in mass from 12 Da to ~800 kDa with charge states between 0 and 73+ were calculated using Collidoscope and IMoS TM with He or N_2 as the collision gas, and the MOBCAL PA and EHSS methods. Collidoscope and MOBCAL calculations were performed using the Linux OS, parallelized over 12 threads on an Intel X5650 CPU. IMoS computations were performed using the Windows OS, parallelized over 7 threads on an Intel Core i7-4790 CPU. All computed values are reported using the notation CCS_{He} and CCS_{N2}, by analogy with the notation suggested by Barran and co-workers [45]. Sources for the atomic coordinates used in the computations as well as the locations of charge sites are described in further detail in Table S1 in the Supplementary Material. Charge for small, non-protein ions was equally distributed among all atoms for CCS computations. For protein ions, charge configurations were generated using the charge placement algorithm described below.

Charge Placement Algorithm for Protonated Protein Ions

Extending the charge placement algorithms of Williams [46, 47], Grandori [48], and Konermann [49], Collidoscope uses a Metropolis-Hastings-like [50, 51] charge placement algorithm for protonated protein ions based on the ion's input atomic coordinates, a user-defined charge state, the calculated point-charge electrostatic repulsion energy of a given charge configuration, and the total intrinsic proton affinity of the ion. The N-terminal amine group as well as each residue is considered to be a possible charge site throughout the charge placement computation, with one possible charge per residue. Initially, the user-defined number of charges are placed in a random configuration on the defined residues, and a total apparent proton affinity, PA_{app}, is calculated as the sum of the intrinsic proton affinities of all charged sites minus the electrostatic repulsion energy of the point-charge sites. This procedure is repeated 1000 times, and the charge configuration with the greatest PA_{app} is selected for Metropolis-Hastings optimization. One of the charges from this configuration is picked at random and moved to a new, random location not already occupied by a charge. The updated PA_{app} of the ion is calculated, and the new charge configuration is accepted if this value is greater than before the charge was moved. If the updated value of PA_{app} is instead lower, it is accepted with a probability $\exp\left(-\frac{\Delta PA_{app}}{k_B T}\right)$. This assures that the algorithm can escape local energy minima to find a low-energy charge configuration. This procedure is repeated until the standard deviation of PA_{app} is less than 6 $k_B T$ for the last 25% of the iterations and the average value of PA_{app} increases by less than 0.1 $k_B T$ between

the second-to-last and last 25% of the iterations. At this point, the optimization is considered to have converged, and the charge configuration with the highest PA_{app} among all the iterations is used for CCS calculations. Intrinsic proton affinities used in Collidoscope are identical to those used by Konermann and co-workers [49], and a relative permittivity of 2.5 is used in calculating electrostatic repulsion for all proteins other than GroEL, for which a value of 4 is used. For the proteins investigated here, the main effect of the charge placement algorithm is to spread the charges out among basic sites near the surface of the ion. (Examples of convergence of the charge placement for $GroEL^{73+}$ are illustrated in the Supplementary Information, Figure S3.)

Results

Default Parameter Optimization

Extensive tests were performed to optimize default values for parameters that minimize computation time while preserving accuracy of the resulting CCSs. Particularly important to the optimization process were the grid buffer distance and the impact parameter spacing. Tests were performed on ions spanning 4 orders of magnitude in mass (water, ondansetron, melittin, and LF_N), and results were compared to determine if the parameter should be scaled as a function of the mass or charge state of the ion. Calculated CCSs tend asymptotically toward a “dense-sampling limit” as parameters are sampled more densely (impact parameter spacing, initial kinetic energies, time step, and vantage points) or as the computations are more extensive (grid buffer distance, initial collision particle-ion separation, and potential energy threshold). Each parameter was optimized such that calculations with the optimized value generally produced CCSs within 1% of the dense-sampling limit for that parameter. These optimization tests are described in further detail below and in the Supplementary Material (Supplementary Figures S1 and S4–13).

Impact parameter spacing in Collidoscope is optimized to scale with the number of atoms in the ion, such that the CCS of an ion with 100 atoms is computed with an impact parameter spacing equal to the van der Waals radius of a carbon atom (1.7 Å; see Figure S8) [52]. The size of each square in the plane-of-origin grid is scaled by the number of atoms, so that the number of trajectories calculated scales approximately linearly with the number of atoms in the ion, for ions of similar density.

The grid buffer distance is optimized to be 15 Å. Smaller grid buffer distances result in significant deviations for the CCSs of ions with masses < ~1 kDa, and a grid buffer distance of 15 Å results in CCSs within 0.004% of their dense-sampling limit for all four ions (see Figure S9). This default grid buffer distance value ensures that some trajectories traveling through regions of low-magnitude potential energy are included in the calculation, but extraneous computation time and memory use are avoided by omitting many trajectories that would deflect negligibly. Note that the grid buffer affects the CCS less for large ions, so it can be significantly reduced or excluded (to reduce computation time) for large enough ions.

The initial distance from the plane of origin to the center of mass, the cut-off energy used to fine-grain omission of trajectories, the number and geometry of vantage points, the range and number of energy states computed, and the integration time step were all optimized

similarly, but these parameters were found to have less significant impact on computed CCSs than the parameters described above. Results used for optimization of these parameters, along with sample scattering trajectories simulated for C_{60}, melittin, and GroEL, are illustrated in the Supplementary Material (Figures S4, S5, and S10–S14).

Comparison of Collidoscope CCSs to Other Calculated and Experimental Values

CCS_{He} for species listed in Table S1 calculated using Collidoscope are compared to those obtained using IMoS in Figure 3. These Collidoscope CCS_{He} agree with those obtained via IMoS's TM to within 4% for all species larger than 1 kDa. Because Collidoscope and IMoS calculations were performed on different CPUs, their computation times are not directly comparable. However, Collidoscope calculations were completed with wall times typically ~2% of the equivalent TM calculations in MOBCAL, on the same CPU. This decrease in computation time is largely due to parallelism.

The CCS_{He} and the CCS_{N2} for $GroEL^{73+}$ were computed with Collidoscope, respectively, in 3.5 and 6 days (see Figure 4). (CCS_{He} were not computed for GroEL 14-mer ions using MOBCAL TM due to prohibitive computational expense, nor were they calculated using IMoS due to software restrictions on the size of molecule that can be used.) PA and EHSS CCS_{He} for this ion were computed with MOBCAL, and these values are 26.8% lower and 1.2% higher than the TM CCS computed with Collidoscope, respectively. These results are in agreement with previous observations that EHSS values tend to be very close to TM values for large, globular ions in low-field IMS, whereas PA values tend to be much lower. No noticeable trend is observed in the relative differences between CCS_{He} calculated with either MOBCAL or IMoS and with Collidoscope as a function of CCS. It is therefore expected that Collidoscope can be used to accurately calculate TM CCSs of ions with masses between ~10 kDa and several MDa that may require prohibitive computation time or memory use with other methods.

A log-log plot of computation time for CCS_{He} and CCS_{N2} versus ion mass is shown in Fig. 4. Interestingly, distinct scaling law behaviors are observed for small, dense ions and for larger biomolecular ions. In particular, computation time scales with ion mass, m, as $m^{0.97}$ (resp., $m^{0.92}$) for low-mass ions and as $m^{1.46}$ (resp., $m^{1.36}$) for larger biomolecular ions with He (resp., N_2) collision gas. This result is attributed to the different densities of these two types of ions, which in Collidoscope requires calculation of more trajectories per unit mass for large biomolecular ions than for smaller, denser ions. Furthermore, computations with N_2 buffer gas require roughly 2–3 times as much time as with He (see Fig. 4).

A comparison of Collidoscope CCS_{He} and CCS_{N2} to experimental results [18, 21, 53] is shown in Figure 5. Collidoscope CCS_{He} are typically ~1–3% higher than IMoS, and both values are typically higher than experimental CCSs. CCS_{He} calculated with Collidoscope are typically between 3% and 15% higher than experimental CCSs, with smaller relative errors for smaller ions (CCS < ~200 $Å^2$) than for larger ions, (Fig. 5). By comparison, results from IMoS TM for this set of ions are typically between 1% and 13% higher than experimental CCSs.

For Collidoscope CCS_{N2} calculations, optimized L-J 6–12 parameters for the quasi-spherical N_2 collision particle were determined by extensive parametrization tests for r_{min} and ε_i values for both carbon and hydrogen, using experimental data for an assortment of aromatic hydrocarbon ions (see Table S2). The r_{min} and ε_i parameters for carbon and hydrogen were optimized simultaneously, and default values were chosen that agree maximally with experimental results. We note that the L-J 6–12 parameters which we found to be optimal are in some cases somewhat different than the parameters used by IMoS. This result indicates that the TM is robust to variation of the L-J 6–12 parameters. CCS_{N2} obtained using these optimized parameters are typically 1% to 18% higher than experimental values for all ions studied, for both Collidoscope and IMoS (excluding GroEL, for which discrepant values have been reported in the literature [6, 29, 53]; see Fig. 5).

Discussion

The overall approach used to optimize Collidoscope is to reduce computational time without sacrificing accuracy and precision, and to provide a flexible, open-source, and parallelizable code base in anticipation of future needs for researchers in the field of native IM-MS. For example, if a researcher needs to calculate a CCS for a very large ion in Collidoscope with greater precision with limited computational time, s/he may choose to change the impact parameter grid spacing or the number of energy states used. Because the computation time varies as the inverse square of the grid spacing and only linearly with the number of energy states, it is less computationally expensive to increase the number of energy states used rather than to decrease the grid spacing. However, care must be taken when increasing the number of energy states, in order to ensure that the dg integral (see Eq. 2) is accurately estimated.

A major difference between Collidoscope and previous TM implementations is the type of sampling used for trajectory parameters. Instead of using Monte-Carlo sampling, Collidoscope samples all trajectory parameters uniformly and with sufficient density to obtain CCSs close to the dense-sampling limit. Computation times vary from a few seconds for small ions to a few days, even for ~MDa-sized complexes, such as the GroEL 14-mer. The above results indicate that Collidoscope computes TM CCS_{He} in close agreement with IMoS's TM implementation.

Because it accounts for long-range polarization effects, multiple scattering, and collision gas temperature, Collidoscope inherently models the physics of the scattering process at a greater level of detail than the PA/PA* or EHSS methods. Computing TM results for large ions is useful for validating empirical "geometry corrections" of values obtained from computationally inexpensive, but less physically detailed, methods such as PA and EHSS. Notably, the EHSS CCS_{He} of GroEL computed with MOBCAL is very close to the TM CCS_{He} for $GroEL^{66+}$ computed with Collidoscope, though both values are considerably higher than previously reported experimental results for this ion in N_2 buffer gas [6, 29, 53]. The MOBCAL PA CCS_{He} for GroEL is *higher* than the experimental CCS_{N2}, though for geometrical reasons the PA CCS generally provides a lower bound for the experimental low-field CCS of a dense, native-like ion. For each ion studied here, CCS_{N2} is strictly *greater* than CCS_{He}, due to both the larger size of the N_2 molecule and its greater polarizability.

These results all strongly suggest that the experimental GroEL CCSs previously reported pertain to conformations that are significantly more collapsed than the x-ray crystal structure. Therefore, calibration of IM-MS data using published experimental GroEL CCS_{He} or CCS_{N2} should be undertaken with careful consideration of the structures adopted under the experimental conditions used. Collidoscope can thus be a valuable tool for evaluating and developing calibration protocols for IM-MS data and for studying the conformational space of large ions.

Based on the comparison between Collidoscope CCSs with experimental values shown in Fig. 5, Collidoscope predicts CCSs for protein geometries derived from x-ray crystallography data that are consistently between 3% and 15% (respectively, 1% and 18%) higher than experimental IM-MS values for native-like ions using He (respectively, N_2) collision gas. This agrees with previous observations that many native-like biomolecular ions undergo some compaction during the electrospray ionization process due to removal of solvent and subsequent self-solvation [3]. Empirically, a scaling factor of 0.91 for Collidoscope CCS_{He} and CCS_{N2} obtained with default parameters using x-ray crystallography-derived geometries of proteins ions is recommended if the user desires to account for this compaction effect for native-like ions with masses between ~1 kDa and ~1 MDa. Based on the results reported here, after this correction, the method precision using both IMoS and Collidoscope is about 10% for protein ions.

For identical input structures, CCS_{He} computed using Collidoscope agree very closely with those computed using either the IMoS TM or MOBCAL EHSS (Fig. 3). Although N_2 buffer gas is not implemented in any MOBCAL method, CCS_{N2} are in close agreement between IMoS and Collidoscope. In addition, ratios between the CCS_{N2} and CCS_{He} computed using Collidoscope agree very well with experimental ratios for these values (Figure 6). Both sets of ratios converge toward an asymptotic value (~1.15) similar to those previously calculated for organic macro-ions using a variety of computational methods [9]. Together, these results indicate that Collidoscope's implementation of the TM for both He and N_2 gas provides accurate estimates of CCSs for the input structures used. Differences between Collidoscope CCSs and experimental values (see Fig. 5) for large protein ions are therefore attributed predominantly to discrepancies in the structures they represent, i.e., the experimental gas-phase structures studied are likely different from the crystal structures used here. A detailed analysis of possible structures adopted in these experiments is beyond the scope of this paper, but the ability to greatly reduce computational time for TM CCSs with Collidoscope (via GPU parallelism and other optimizations) will facilitate future work investigating the gas-phase conformational landscape for large ions.

Although we recommend using the extensively optimized default parameters presented here, customization of Collidoscope to suit the needs of the user is possible and relatively simple due to its object-oriented programming. Lennard-Jones parameters, collision gas properties, and the trajectory integration algorithm are coded separately and may be modified as needed.

Conclusions

Collidoscope is a computationally efficient tool for calculating CCSs of ions with a wide range of masses using the Trajectory Method that produces results in close agreement with measured low-field ion mobility values for both He and N_2 collision gas. Computation time is significantly decreased relative to MOBCAL TM calculations due to parallelized computing and optimized sampling of trajectory parameters. A GPU-parallelized version of Collidoscope is currently under development and will further reduce computational time. Because TM CCSs for megadalton-sized ions can be calculated using Collidoscope in a few days, the program makes detailed IM-MS analysis of the conformational space of large ions tractable at a high level of sophistication. CCSs computed with Collidoscope can also be used in combination with modeled structures to inform calibrations of IM-MS data for very large ions for which conformations are not precisely known or can vary with instrumental conditions, such as GroEL.

In the future, a more rigorous treatment of N_2 will be implemented in Collidoscope to account for its permanent quadrupole moment and full dipolar polarizability tensor. Higher-order Runge-Kutta integration methods will also be made available as an option to further reduce trajectory integration error. Finally, other buffer gases, including argon, will be implemented. Collidoscope is available for public use upon request from the authors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Computations on the University of Oregon ACISS Supercomputing Cluster were supported by the National Science Foundation (Grant OCI-0960354). The authors thank Prof. Bryan Krantz for the generous gift of anthrax toxin proteins and Elliott Ewing for helpful discussions.

References

1. Maisser A, Premnath V, Ghosh A, Nguyen TA, Attoui M, Hogan CJ. Determination of gas phase protein ion densities via ion mobility analysis with charge reduction. Phys Chem Chem Phys. 2011; 13:21630–21641. [PubMed: 22073403]
2. Alexeev Y, Fedorov DG, Shvartsburg AA. Effective Ion Mobility Calculations for Macromolecules by Scattering on Electron Clouds. J Phys Chem A. 2014; 118:6763–6772. [PubMed: 25137627]
3. Jurneczko E, Barran PE. How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. Analyst. 2011; 136:20–28. [PubMed: 20820495]
4. Ruotolo BT, Benesch JLP, Sandercock AM, Hyung SJ, Robinson CV. Ion mobility-mass spectrometry analysis of large protein complexes. Nat Protocols. 2008; 3:1139–1152. [PubMed: 18600219]
5. Bernstein SL, Dupuis NF, Lazo ND, Wyttenbach T, Condron MM, Bitan G, Teplow DB, Shea JE, Ruotolo BT, Robinson CV, Bowers MT. Amyloid-beta protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. Nat Chem. 2009; 1:326–331. [PubMed: 20703363]
6. Hogan CJ, Ruotolo BT, Robinson CV, de la Mora JF. Tandem Differential Mobility Analysis-Mass Spectrometry Reveals Partial Gas-Phase Collapse of the GroEL Complex. J Phys Chem B. 2011; 115:3614–3621. [PubMed: 21395304]

7. Lanucara F, Holman SW, Gray CJ, Eyers CE. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. Nat Chem. 2014; 6:281–294. [PubMed: 24651194]
8. Zhou M, Dagan S, Wysocki VH. Protein Subunits Released by Surface Collisions of Noncovalent Complexes: Nativelike Compact Structures Revealed by Ion Mobility Mass Spectrometry. Angew Chem-Int Edit. 2012; 51:4336–4339.
9. Larriba-Andaluz C, Fernandez-Garcia J, Ewing MA, Hogan CJ, Clemmer DE. Gas molecule scattering & ion mobility measurements for organic macro-ions in He versus N-2 environments. Phys Chem Chem Phys. 2015; 17:15019–15029. [PubMed: 25988389]
10. Ewing MA, Glover MS, Clemmer DE. Hybrid ion mobility and mass spectrometry as a separation tool. J Chromatogr A. 2016; 1439:3–25. [PubMed: 26592562]
11. Niu S, Rabuck JN, Ruotolo BT. Ion mobility-mass spectrometry of intact protein–ligand complexes for pharmaceutical drug discovery and development. Curr Opin Chem Biol. 2013; 17:809–817. [PubMed: 23856053]
12. Pacholarz KJ, Garlish RA, Taylor RJ, Barran PE. Mass spectrometry based tools to investigate protein-ligand interactions for drug discovery. Chem Soc Rev. 2012; 41:4335–4355. [PubMed: 22532017]
13. Beveridge R, Chappuis Q, Macphee C, Barran P. Mass spectrometry methods for intrinsically disordered proteins. Analyst. 2013; 138:32–42. [PubMed: 23108160]
14. Laganowsky A, Reading E, Allison TM, Ulmschneider MB, Degiacomi MT, Baldwin AJ, Robinson CV. Membrane proteins bind lipids selectively to modulate their structure and function. Nature. 2014; 510:172-+. [PubMed: 24899312]
15. Vahidi S, Stocks BB, Konermann L. Partially Disordered Proteins Studied by Ion Mobility-Mass Spectrometry: Implications for the Preservation of Solution Phase Structure in the Gas Phase. Anal Chem. 2013; 85:10471–10478. [PubMed: 24088086]
16. Morsa D, Defize T, Dehareng D, Jérôme C, De Pauw E. Polymer Topology Revealed by Ion Mobility Coupled with Mass Spectrometry. Anal Chem. 2014; 86:9693–9700. [PubMed: 25188877]
17. Freeke J, Robinson CV, Ruotolo BT. Residual counter ions can stabilise a large protein complex in the gas phase. Int J Mass Spectrom. 2010; 298:91–98.
18. Campuzano I, Bush MF, Robinson CV, Beaumont C, Richardson K, Kim H, Kim HI. Structural Characterization of Drug-like Compounds by Ion Mobility Mass Spectrometry: Comparison of Theoretical and Experimentally Derived Nitrogen Collision Cross Sections. Anal Chem. 2012; 84:1026–1033. [PubMed: 22141445]
19. Pukala TL, Ruotolo BT, Zhou M, Politis A, Stefanescu R, Leary JA, Robinson CV. Subunit Architecture of Multiprotein Assemblies Determined Using Restraints from Gas-Phase Measurements. Structure. 2009; 17:1235–1243. [PubMed: 19748344]
20. Quintyn RS, Yan J, Wysocki VH. Surface-Induced Dissociation of Homotetramers with D2 Symmetry Yields their Assembly Pathways and Characterizes the Effect of Ligand Binding. Chem Biol. 2015; 22:583–592. [PubMed: 25937312]
21. Salbo R, Bush MF, Naver H, Campuzano I, Robinson CV, Pettersson I, Jørgensen TJD, Haselmann KF. Traveling-wave ion mobility mass spectrometry of protein complexes: accurate calibrated collision cross-sections of human insulin oligomers. Rapid Commun Mass Spectrom. 2012; 26:1181–1193. [PubMed: 22499193]
22. Trimpin S, Tan B, Bohrer BC, O'Dell DK, Merenbloom SI, Pazos MX, Clemmer DE, Walker JM. Profiling of phospholipids and related lipid structures using multidimensional ion mobility spectrometry-mass spectrometry. Int J Mass Spectrom. 2009; 287:58–69.
23. Sterling HJ, Kintzer AF, Feld GK, Cassou CA, Krantz BA, Williams ER. Supercharging Protein Complexes from Aqueous Solution Disrupts their Native Conformations. J Am Soc Mass Spectrom. 2012; 23:191–200. [PubMed: 22161509]
24. Shepherd DA, Marty MT, Giles K, Baldwin AJ, Benesch JLP. Combining tandem mass spectrometry with ion mobility separation to determine the architecture of polydisperse proteins. Int J Mass Spectrom. 2015; 377:663–671.

25. Ferguson CN, Benchaar SA, Miao Z, Loo JA, Chen H. Direct Ionization of Large Proteins and Protein Complexes by Desorption Electrospray Ionization-Mass Spectrometry. Anal Chem. 2011; 83:6468–6473. [PubMed: 21774530]
26. Silveira JA, Fort KL, Kim D, Servage KA, Pierson NA, Clemmer DE, Russell DH. From Solution to the Gas Phase: Stepwise Dehydration and Kinetic Trapping of Substance P Reveals the Origin of Peptide Conformations. J Am Chem Soc. 2013; 135:19147–19153. [PubMed: 24313458]
27. Gabelica V, Baker ES, Teulade-Fichou MP, De Pauw E, Bowers MT. Stabilization and structure of telomeric and c-myc region intramolecular Gquadruplexes: The role of central cations and small planar ligands. J Am Chem Soc. 2007; 129:895–904. [PubMed: 17243826]
28. Uetrecht C, Rose RJ, van Duijn E, Lorenzen K, Heck AJR. Ion mobility mass spectrometry of proteins and protein assemblies. Chem Soc Rev. 2010; 39:1633–1655. [PubMed: 20419213]
29. van Duijn E, Barendregt A, Synowsky S, Versluis C, Heck AJR. Chaperonin Complexes Monitored by Ion Mobility Mass Spectrometry. J Am Chem Soc. 2009; 131:1452–1459. [PubMed: 19138114]
30. Allen SJ, Bush MF. Radio-Frequency (rf) Confinement in Ion Mobility Spectrometry: Apparent Mobilities and Effective Temperatures. J Am Soc Mass Spectrom. 2016; 27:2054–2063. [PubMed: 27582119]
31. Liu FC, Kirk SR, Bleiholder C. On the structural denaturation of biological analytes in trapped ion mobility spectrometry - mass spectrometry. Analyst. 2016; 141:3722–3730. [PubMed: 26998732]
32. Marklund, Erik G., Degiacomi, Matteo T., Robinson, Carol V., Baldwin, Andrew J., Benesch, Justin LP. Collision Cross Sections for Structural Proteomics. Structure. 2015; 23:791–799. [PubMed: 25800554]
33. Mesleh MF, Hunter JM, Shvartsburg AA, Schatz GC, Jarrold MF. Structural information from ion mobility measurements: Effects of the long-range potential. J Phys Chem. 1996; 100:16082–16086.
34. Paizs B. A divide-and-conquer approach to compute collision cross sections in the projection approximation method. Int J Mass Spectrom. 2015; 378:360–363.
35. Bleiholder C, Contreras S, Bowers MT. A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (IV). Application to polypeptides. Int J Mass Spectrom. 2013; 354–355:275–280.
36. Bleiholder C. A local collision probability approximation for predicting momentum transfer cross sections. Analyst. 2015; 140:6804–6813. [PubMed: 26178623]
37. Shvartsburg AA, Jarrold MF. An exact hard-spheres scattering model for the mobilities of polyatomic ions. Chem Phys Lett. 1996; 261:86–91.
38. Larriba C, Hogan CJ. Ion Mobilities in Diatomic Gases: Measurement versus Prediction with Non-Specular Scattering Models. J Phys Chem A. 2013; 117:3887–3901. [PubMed: 23488939]
39. Larriba C, Hogan CJ. Free molecular collision cross section calculation methods for nanoparticles and complex ions with energy accommodation. J Comput Phys. 2013; 251:344–363.
40. Kihara T. The Mathematical Theory of Electrical Discharges in Gases. B. Velocity-Distribution of Positive Ions in a Static Field. Rev Mod Phys. 1953; 25:844–852.
41. Mason EA, Schamp HW. Mobility of gaseous lons in weak electric fields. Ann Phys. 1958; 4:233–270.
42. Wyttenbach T, von Helden G, Batka JJ, Carlat D, Bowers MT. Effect of the long-range potential on ion mobility measurements. J Am Soc Mass Spectrom. 1997; 8:275–282.
43. Pachucki K, Sapirstein J. Relativistic and QED corrections to the polarizability of helium. Phys Rev A. 2000; 63:012504.
44. Spelsberg D, Meyer W. Static dipole polarizabilities of N2, O2, F2, and H2O. J Chem Phys. 1994; 101:1282–1288.
45. Surman AJ, Robbins PJ, Ujma J, Zheng Q, Barran PE, Cronin L. Sizing and Discovery of Nanosized Polyoxometalate Clusters by Mass Spectrometry. J Am Chem Soc. 2016; 138:3824–3830. [PubMed: 26906879]
46. Schnier PD, Gross DS, Williams ER. On the maximum charge state and proton transfer reactivity of peptide and protein ions formed by electrospray ionization. J Am Soc Mass Spectrom. 1995; 6:1086–1097. [PubMed: 24214055]

47. Schnier PD, Gross DS, Williams ER. Electrostatic Forces and Dielectric Polarizability of Multiply Protonated Gas-Phase Cytochrome c Ions Probed by Ion/Molecule Chemistry. J Am Chem Soc. 1995; 117:6747–6757.

48. Marchese R, Grandori R, Carloni P, Raugei S. A Computational Model for Protein Ionization by Electrospray Based on Gas-Phase Basicity. J Am Soc Mass Spectrom. 2012; 23:1903–1910. [PubMed: 22993040]

49. Popa V, Trecroce DA, McAllister RG, Konermann L. Collision-Induced Dissociation of Electrosprayed Protein Complexes: An All-Atom Molecular Dynamics Model with Mobile Protons. J Phys Chem B. 2016; 120:5114–5124. [PubMed: 27218677]

50. Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika. 1970; 57:97–109.

51. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of State Calculations by Fast Computing Machines. J Chem Phys. 1953; 21:1087–1092.

52. Bondi A. van der Waals Volumes and Radii. J Phys Chem. 1964; 68:441–451.

53. Bush MF, Hall Z, Giles K, Hoyes J, Robinson CV, Ruotolo BT. Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. Anal Chem. 2010; 82:9557–9565. [PubMed: 20979392]

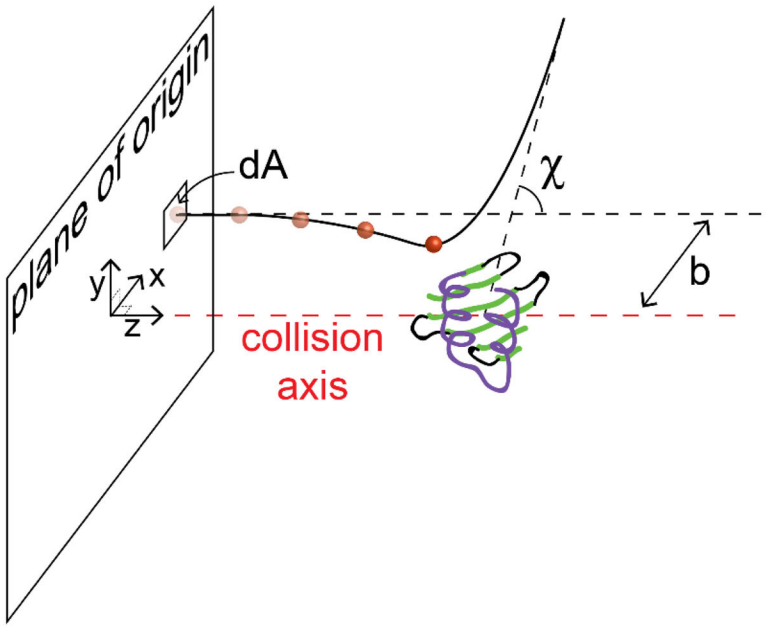


Figure 1.
Graphical depiction of the simulation system for a single trajectory. A collision particle incident from the plane of origin with impact parameter b scatters off the potential energy surface at angle χ.

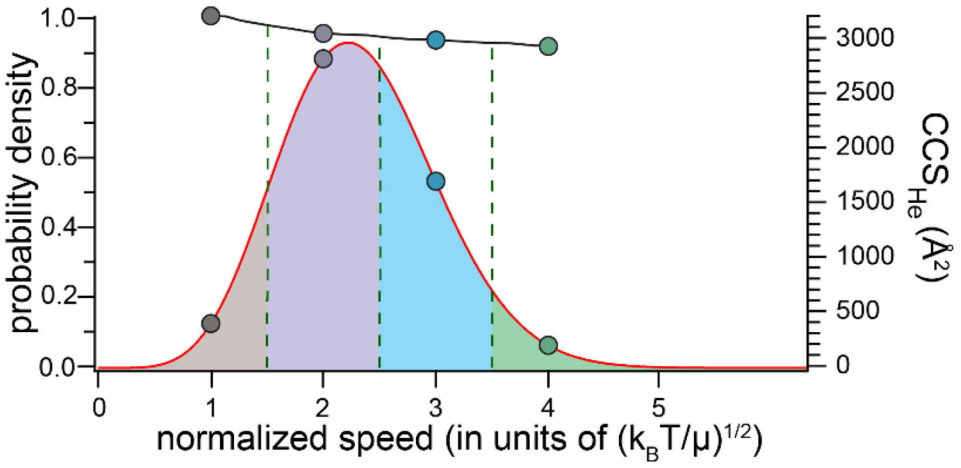


Figure 2.
Optimized default settings for initial collision speed sampling. The red curve represents the probability density function for the simulated trajectory's initial conditions as a function of normalized collision speed, dots represent sampled collision speed values, and dashed vertical lines represent limits of integration used for probability weightings equal to the areas of the corresponding colored regions (see text). The black curve is the calculated CCS_{He} for LF_N at 298 K, as a function of buffer gas speed.

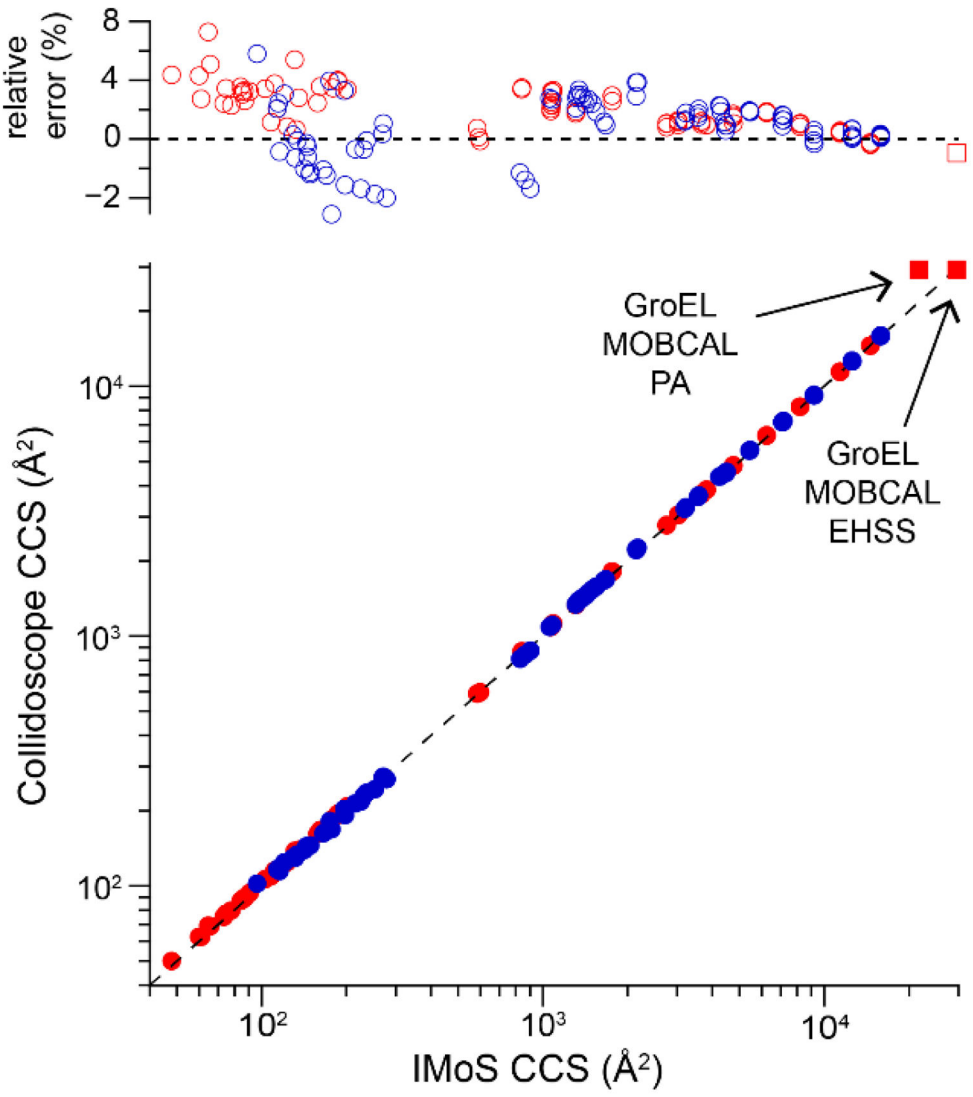


Figure 3.
Comparison of CCS (red, He; blue, N_2) obtained using Collidoscope and those obtained via IMoS's TM approximation. The diagonal line represents perfect agreement between the two methods. For GroEL, MOBCAL PA and EHSS values are used instead of IMoS's TM (see text).

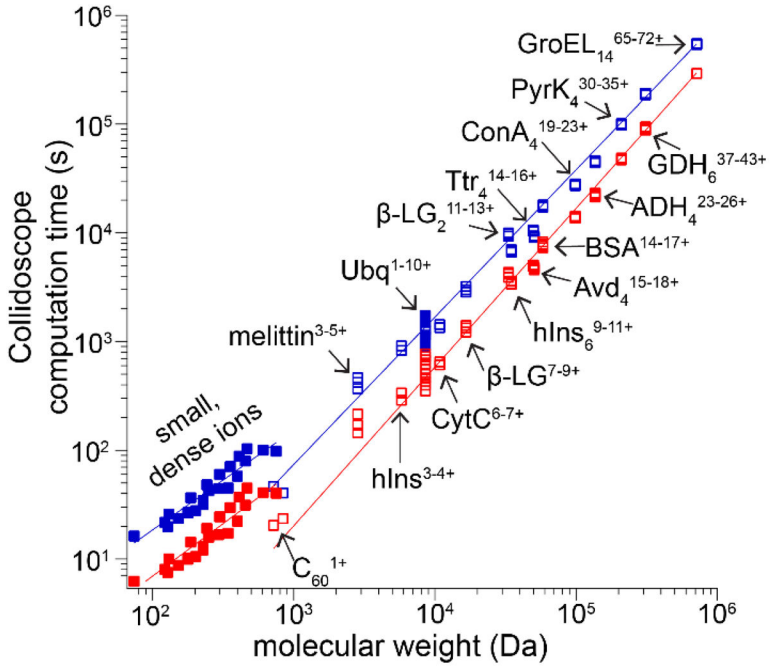


Figure 4.
Log-log plot of Collidoscope computation time versus molecular weight for small, dense ions and protonated proteins in He (red) and N_2 (blue) buffer gas. Fitted lines illustrate different scaling law behavior for ions with different densities.

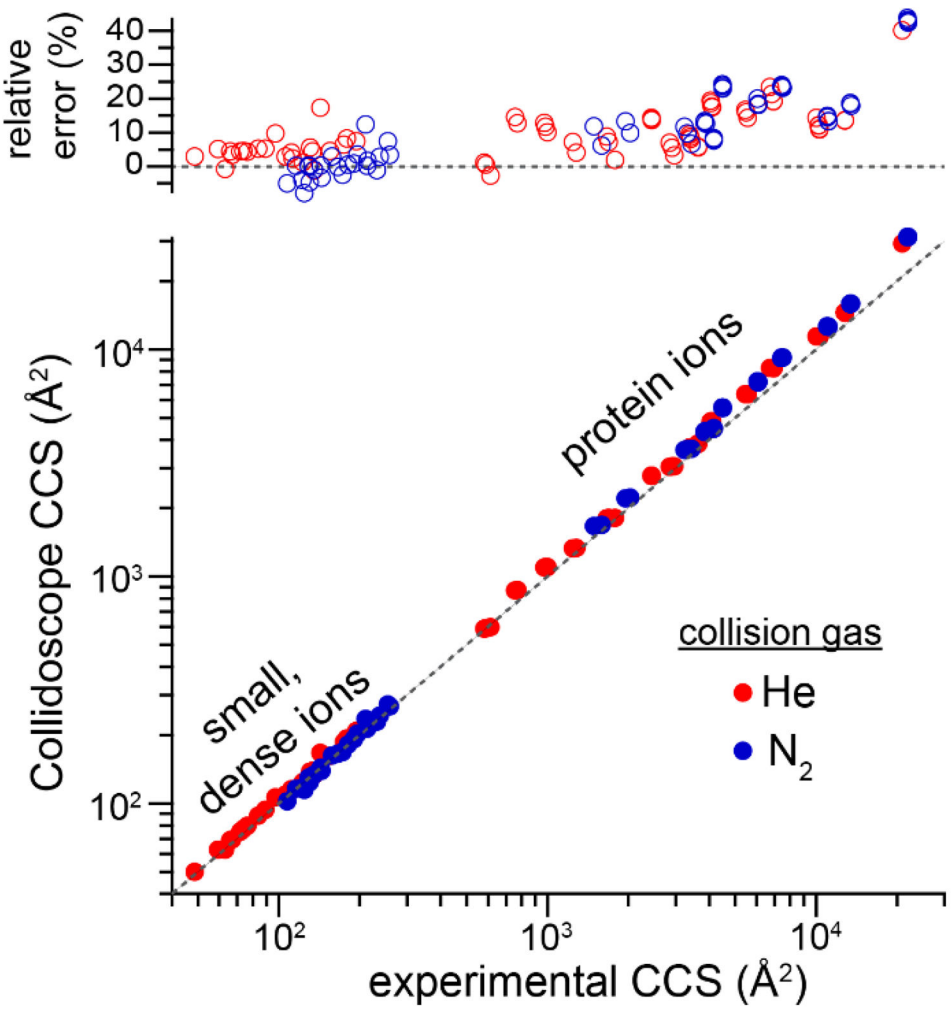


Figure 5.
Comparison of CCSs obtained using Collidoscope and those obtained from IMS experiments, using He (red) or N_2 (blue) buffer gas. The diagonal line represents perfect agreement between calculated and experimental values.

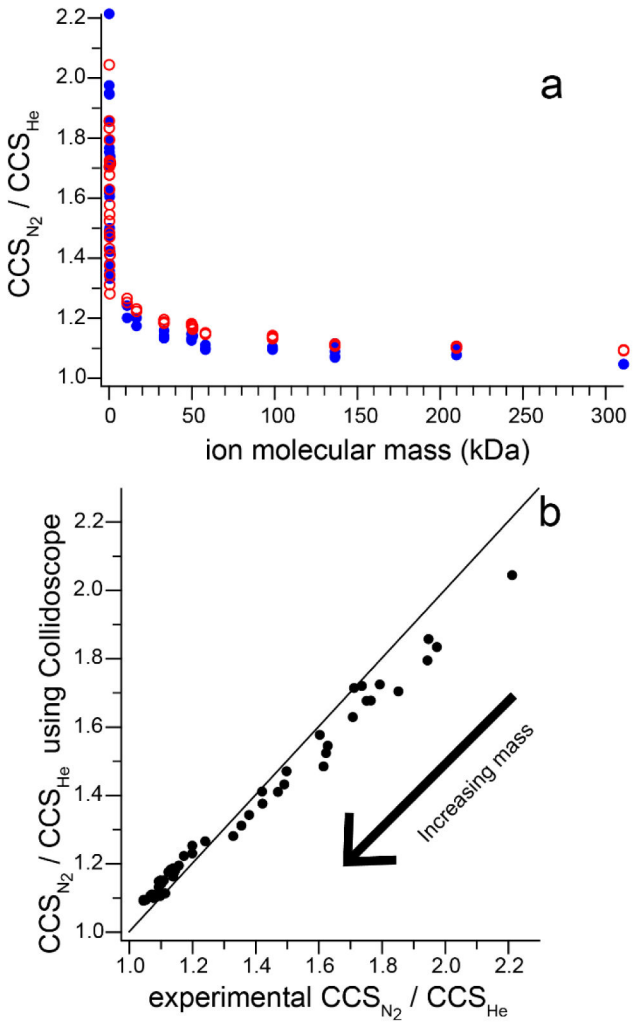


Figure 6.
a) Plot of the ratio of CCS_{N2} and CCS_{He} versus ion mass (experimental values, blue; Collidoscope values, red). b) Comparison of Collidoscope and experimental ratios (diagonal indicates perfect agreement).

Table 1

Features present (+) or absent (−) in various CCS calculation tools. Methods in bold type use 2-dimensional projections rather than explicit 3-dimensional scattering for CCS computations. Asterisks (*) indicate that a geometry-dependent correction factor is used to partially account for 3-dimensional scattering.

Computational Suite	Method	Momentum Transfer Type	Explicit Trajectories	Temperature/Charge State Dependent	3D Geometry Dependent	Collision particles
MOBCAL [33, 37]	**PA**	−	−	−/−	−	He
	PA*	−	−	−/−	+	He
	EHSS	elastic	+	−/−	+	He
	TM	elastic	+	+/+	+	He
WebPSA [35]	**PSA***	−	−	+/+	+	He/N_2
[36]	LCPA	−	−	+/+	+	He/N_2
Sigma [42], CCS [34], and IMPACT [32]	**PA***	−	−	−/−	+	He
IMoS [9, 38, 39]	DTM	elastic and inelastic	+	+/+	+	He/N_2
	DHSS	elastic and inelastic	+	+/−	+	He/N_2
Collidoscope	TM	elastic	+	+/+	+	He/N_2
	“ideal”	exact	+	+/+	+	all