# Geometric Analysis of Shapes in Ion Mobility−Mass Spectrometry

Jean R. N. Haler,*,[||] Eric Béchet,[||] Christopher Kune, Johann Far, and Edwin De Pauw

ABSTRACT: Experimental ion mobility−mass spectrometry (IM−MS) results are often correlated to three-dimensional structures based on theoretical chemistry calculations. The bottleneck of this approach is the need for accurate values, both experimentally and theoretically predicted. Here, we continue the development of the trend-based analyses to extract structural information from experimental IM−MS data sets. The experimental collision cross-sections (CCSs) of synthetic systems such as homopolymers and small ionic clusters are investigated in terms of CCS trends as a function of the number of repetitive units (e.g., degree of polymerization (DP) for homopolymers) and for each detected charge state. Then, we computed the projected areas of expanding but perfectly defined geometric objects using an in-house software called MoShade. The shapes were modeled using computer-aided design software where we considered only geometric factors: no atoms, mass, chemical potentials, or interactions were taken into consideration to make the method orthogonal to classical methods for 3D shape assessments using time-consuming computational chemistry. Our modeled shape evolutions favorably compared to experimentally obtained CCS trends, meaning that the apparent volume or envelope of homogeneously distributed mass effectively modeled the ion−drift gas interactions as sampled by IM−MS. The CCSs of convex shapes could be directly related to their surface area. More importantly, this relationship seems to hold even for moderately concave shapes, such as those obtained by geometry-optimized structures of ions from conventional computational chemistry methods. Theoretical sets of expanding beads-on-a-string shapes allowed extracting accurate bead and string dimensions for two homopolymers, without modeling any chemical interactions.

KEYWORDS: ion mobility−mass spectrometry, synthetic polymers, structure interpretation, collision cross-section fitting, MoShade
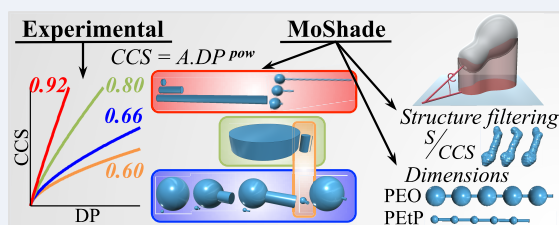
INTRODUCTION

Ion mobility−mass spectrometry (IM−MS) is being increasingly used for structural characterization of small molecules,[1−5] biomolecule folds,[6−10] or even synthetic polymer folds in the gas phase.[11−16] The two-dimensional shape information obtained from IM−MS as reduced ion mobility (K_0) or collision cross-section (CCS) does not offer atomic level structural characterizations. Therefore, theoretical chemistry calculations are usually coupled to the IM−MS experiments. Density functional theory (DFT), molecular mechanics (MM), molecular dynamics (MD), or different couplings of these structure calculation methods are applied to provide detailed three-dimensional structures from the experimentally determined CCS.[5,13−15,17−21] The best experiment−theory CCS value agreement[22] (often empirically fixed at ≤5% deviation) is then used as the criterion to choose the three-dimensional structure(s) representing the ion shape sampled by IM−MS. Structural interpretations can then be undertaken.

The interpretation of CCS trends over larger mass ranges or charge state ranges constitutes a different approach to interpreting structural and physicochemical information from IM−MS data.[9,16,23−26] If multiple charge states can be obtained, multiple CCS trends as a function of the mass or the charge state can be used for interpretation. One of the benefits of this methodology is that it can be independent of accurate CCS values, given that CCS trends represent changes in CCS rather than absolute CCS interpretations. Such interpretations can then for example also be performed on raw data.[16,23] This creates an interpretation strategy focused on experimental data, which allows obtaining physicochemical interpretations on multiple ions at once without having to calculate atomic fine structures for every single data point aligning on a CCS trend.

Given the structural homogeneity of synthetic (homo)-polymers, they constitute the ideal systems to explore these CCS trend-based methodologies over large mass and charge state ranges. They also allow building empirical interpretations based on polymer-to-polymer comparisons.[16,23] Similarly, artificial intelligence and deep learning approaches are gaining in popularity to build databases from experimental data trend analyses.[27−29]

Here, we continue the development of the CCS trend analysis, by adding the structural shape dimension, in addition
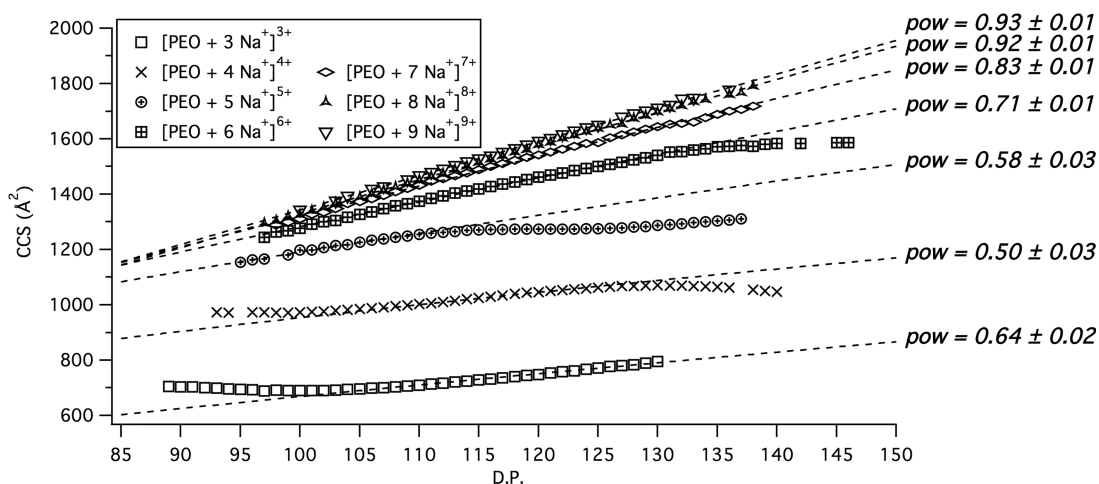
Figure 1. Collision cross-section (CCS) as a function of the degree of polymerization (DP) of poly(ethylene oxide) (PEO). The charge states range from 3+ to 9+ using sodium cation adducts. The dotted lines represent the CCS evolution fits (eq 1), and the *pow* fit parameter of each fit is given.

to the physicochemical dimension,[16,25,26] to the interpretation. For this purpose, we based our study on cation adducts of synthetic polymers where many data points can be generated (by increasing the polymer chain length, degree of polymerization, DP) without changing the essence of the interactions of the polymer−cation complexes. Aside from the chemically homogeneously growing structures of synthetic (homo)-polymers, a set of anionic tetrahydric clusters of iron(III), i.e., $[FeCl_3H(CH_2)_nCO_2]^-$, with a heterogeneous core (iron, chlorine, and linear carboxylate) and homogeneous ligand growth (length of the carboxylate alkyl chain) was also used as a model system. Their experimental trends of the CCS evolutions as a function of repetitive units are then interpreted through a fitting method.[9,16,23] Experimental fit parameters[16,23] were favorably compared to the fit parameters obtained from CCS trends of modeled geometric objects growing in different dimensions, which are perfectly defined and controlled. Modeled CCS evolutions are obtained through a new in-house CCS calculation program named MoShade. Our MoShade calculations are purely geometric shape-based analyses with no 3D structures from theoretical chemistry intervening in the process (i.e., not resolved at the atomic level and without mass considerations). Shape geometries defining the fit parameters (CCS evolutions) can be extracted and conclusions, which could also be applicable to other systems than the studied models, are found.

MATERIALS AND METHODS

Polymers, $[FeCl_3H(CH_2)_nCO_2]^-$ Clusters, and Ion Mobility−Mass Spectrometry. A Synapt G2 HDMS (Waters, UK) equipped with an electrospray ionization source (ESI) was used to perform IM−MS experiments. Poly-(ethylene oxide) PEO and poly(ethoxyphosphate) PEtP were measured in positive ion mode. $[FeCl_3H(CH_2)_nCO_2]^-$ clusters were measured in negative ion mode. Details can be found in the Supporting Information 1.

MoShade. MoShade is written in C++ and GPLv2-licensed. It is cross-platform compiled for Linux and Windows (64-bit). MoShade is a multithreaded, terminal-based (or command prompt) program. The shape inputs are stereo-lithography .stl or composite .gm files, allowing purely theoretical geometric objects or .stl-converted (e.g., using

VMD's QuickSurf representation[30]) ions or molecules to be computed by MoShade. MoShade computes the volume, the surface area, and the projected surface area (interpreted later as CCS in arbitrary units) of a shape. MoShade also calculates the ratio of the surface area and the CCS as well as the minimum and maximum CCS values from all projection angles, which can be printed out (optional). A concavity factor is also provided by computing the ratio between the sum of the reflex angles and the sum of all angles, taken at the edges of each contiguous and concatenated triangle. Its value ranges from 0 (perfectly convex shape) to 1. Our concavity factor depends on the number of triangles and cannot be compared between shapes having different numbers of triangles. However, this method does not require to define a shape factor by comparing the surface of the effective molecular shape to the surface of its corresponding convex envelope.[31]

MoShade can be compiled (on Linux) with or without a graphical visualization interface (using VTK or FLTK). The graphical visualization allows identifying unconnected triangles or even holes in the input structures, which could possibly yield incorrect results or increase the duration of the calculation through the loss of calculation optimizations.

The calculations were performed using optimized integrals and using $n = 15$ samples, representing $N = 900$ projection calculations ($N = (2n)^2$; for explanations, see section below). They were sufficient for yielding invariable MoShade results; verified by calculating $N = 3600$ projections ($n = 30$) for several structures.

A complete demonstration of the mathematics for (any) concave shapes can be found in the dedicated Supporting Information file (Supporting Information 2). Additional developments for convex shapes are also developed.

MoShade (and nutil library) can be downloaded at the following links: https://cadxfem.org/svn/cadxfem/moshade/, https://cadxfem.org/svn/cadxfem/nutil/, and https://github.com/JeanRNH/MoShade/releases.

MoShade Mathematical Background. In brief, MoSh-ade calculates the projected area of the shape by sampling its orientation. A short description of the software and its mathematical background can be found in the Supporting Information 1.

MoShade Input Preparation and MoShade Automator. The model structure input geometries (.stl files) were prepared using Autodesk 123D Design 14.2.2, free computer-aided design (CAD) software. Free software MoShade Automator (licensed GPLv2+), written in FreePascal (Object Pascal) using Lazarus, is a cross-platform (Linux and Windows 64-bit) graphical user interface for creating batch files of multiple structure inputs. It can also pilot MoShade by running, pausing, or stopping the execution of MoShade between inputs of multiple batch files.

MoShade Automator can be downloaded at the following link: https://github.com/JeanRNH/MoShadeAutomator/releases.

Theoretical PEO Structure Pool. The PEO calculations, based on conventional computational chemistry, were performed on PEO DP = 14 with 1 or 2 sodium cations. The structures were generated using both MM2 and MMFF94 force fields implemented in Chem3D Pro v.11.0. Structure optimizations after heat annealing to 300 K were performed. A pool of PEO structures was generated by intermediately stopping the calculation and saving the current structure. VMD[30] 1.9.2 yielded conform .stl files for MoShade calculations ("QuickSurf" shape representation, resolution: 1.0, radius scale: 1.0, density isovalue: 0.3, grid spacing: 1.0).

Theoretical Molecular Dynamics on Carboxylates. MD simulations were performed on butanoate, hexanoate, octanoate, decanoate, and dodecanoate carboxylates without adding the $FeCl_3$ core to the structure. A total of 4000 structures were recorded over 100 ns at 500 K using the AMBER94 force field[32] (implemented in Abalone v.1.8.94).

■ RESULTS AND DISCUSSIONS

Experimental IM−MS Evolution *Pow* Fitting Parameters for a Chemically Homogeneously Growing System: PEO Polymers. Figure 1 plots the experimental CCS as a function of the degree of polymerization (DP; converted from the mass-to-charge m/z ratio) of PEO (5000 g/mol). The CCS of highly charged polymer ions is generally increasing with increasing DP and charge state (e.g., $[PEO + 4Na^+]^{4+}$ to $[PEO + 9Na^+]^{9+}$). In different DP regions, the CCS decreases, while the DP still increases (e.g., $[PEO + 4Na^+]^{4+}$ 93 < DP < 100 and 130 < DP < 140). During these disruptions in the generally increasing CCS evolutions, the three-dimensional structures of the polymer complexes rearrange as a result of a decreased Coulomb repulsion induced by an increase in the charge solvation (i.e., increase in the monomeric unit/Na^+ ratio). After having undergone several structural rearrangements, the higher charge states merge with the most compact CCS evolution (i.e., the common trend line[12,13,16,23], here represented by the $[PEO + 3Na^+]^{3+}$ complexes above DP 100.

The CCS evolutions before and after the structural rearrangements can be fitted using eq 1.[23] In this study, we focused on the *pow* parameter. As we showed, it contains shape information on the complexes. Interpretations on the parameter A are the focus of a specifically dedicated paper.[16] When fitting the CCS evolutions,[16,23] the *pow* fit parameters of the polymer−cation complexes are found to roughly range from around 0.55 to 0.95.

$$\Omega = A \cdot DP^{pow} \qquad (1)$$

where Ω represents the CCS, DP represents the degree of polymerization, and A and *pow* are the fitting parameters (as developed in another paper from our group[23]). Since DP is directly related to the polymer mass (m), eq 2 derives from eq 1

$$\Omega = A' \cdot m^{pow} \qquad (2)$$

Application to Systems with a Chemically Heterogeneous Core: $[FeCl_3H(CH_2)_nCO_2]^-$ Clusters. The experimental *pow* fitting parameter has also been estimated for a set of iron(III) $[FeCl_3H(CH_2)_nCO_2]^-$ clusters with different alkyl chain lengths (i.e., n = 1, 2, 3, 5, 7, 11) corresponding to ethanoate, propanoate, butanoate, hexanoate, octanoate, and dodecanoate chains. Figure 2 plots the experimental CCS of

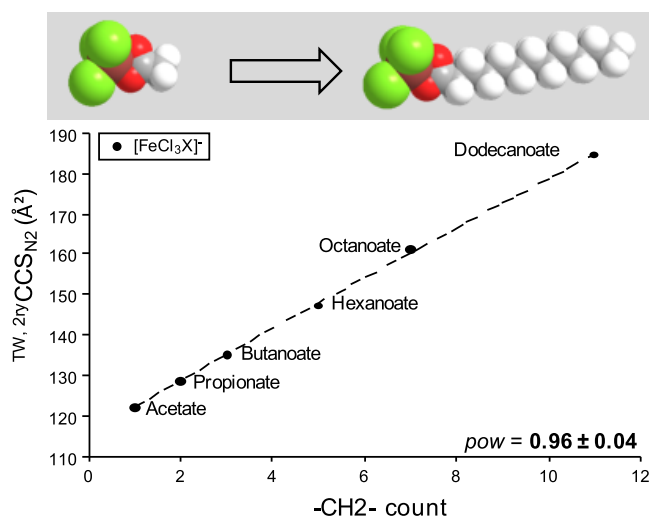**Figure 2.** Collision cross-section (CCS) as a function of the −CH2− count (i.e., n) of $[FeCl_3H(CH_2)_nCO_2]^-$ clusters, depicted $[FeCl_3X]^-$ (schematic representation on top). The dotted line represents the fit (eq 3) of the CCS evolution. The *pow* fitting parameter is given in the figure. The $^{TW,2ry}CCS_{N2}$ values have been obtained by CCS calibration of the Synapt G2 HDMS (TWIMS instrument) using published CCS values of polyalanine anions in N_2[33] (secondary calibration from TWIMS values[34]).

$[FeCl_3H(CH_2)_nCO_2]^-$ as a function of the number of −CH₂− units of the carboxylate ligands. This model system stands in opposition to synthetic polymers. Indeed, the synthetic (homo)polymers grow by adding chemically identical monomer units to their homogeneous monomer chain. The $[FeCl_3H(CH_2)_nCO_2]^-$ clusters, however, exhibit a very heterogeneous chemical composition with their iron and chlorine core and the alkyl ligand chain. The chosen cluster series nevertheless grows by chemically identical increments, i.e., −CH₂− entities. This model system can therefore expand the application of the CCS trend-based analysis to a larger range of molecules than polymers and biomolecules.[9,16,23−26]

In contrast to the PEO CCS trends (Figure 1), the $[FeCl_3X]^-$ clusters (Figure 2) evolve monotonically, exhibiting no structure or shape rearrangement within the screened carboxylate ligand sizes (or masses). Since the number of the monomeric units is small within the model of these clusters, the contribution of $FeCl_3$ (nonrepetitive ion core) on the CCS cannot be neglected. Therefore, the *pow* fitting parameter has been evaluated from eq 3, which is a generalization of eq 2. Eq 3 adds a CCS offset to eq 2

$$\Omega = A' \cdot m^{pow} + c \qquad (3)$$

In eq 3, the parameter c is added to take the heterogeneity of the $FeCl_3$ core compared to the carboxylate ligand into

Table 1. Shape Descriptions and Shape Deformations Used to Calculate CCS Values and Trends Using MoShade[e]

Shape description	Shape variation[a] & number of shapes (i)	Dimensions[a] (min.–max.)[b]	pow[c]	Scheme[d]	Figure
Sphere	Diameter (i = 14)	5 – 250	0.66688 ± 8.26×10⁻⁹		S2.a.
Cylinder	CL (i = 20)	CD: 20 CL: 5-300	0.91308 ± 0.0243 or Linear fit		S2.b.
Cylinder	CD (i = 20)	CD: 5-1000 CL: 200	0.79205 ± 0.0194		S2.c.
Cylinder	CD (i = 10)	CD: 5-125 CL: 200	0.57132 ± 0.0166		S2.d.
Ball-Cylinder	BD (i = 24)	BD: 3.8-45 CD: 3 CL: 5	0.65589 ± 0.00393		S3.a.
Ball-Cylinder	BD (i = 15)	BD: 3.8-9 CD: 3 CL: 5	0.60632 ± 0.00267		S3.b.
Ball-Cylinder	Uniform Scaling (i = 17)	BD: 1.6-60.9 CD: 0.6-22.8 CL: 0.7-26.6	0.66694 ± 0.000362		S3.c.
Ball-Cylinder	Uniform Scaling (i = 18)	BD: 1.6-65.8 CD: 0.6-24.7 CL: 1.4-57.5	0.66802 ± 0.00163		S3.c.
Ball-Cylinder	Uniform Scaling (i = 18)	BD: 1.6-65.8 CD: 0.6-24.7 CL: 2.1-86.3	0.66796 ± 0.00157		S3.c.
Ball-Cylinder	CL (i = 21)	BD: 20 CD: 3 CL: 3.5-125	Linear fit		S3.d.

[a]BD represents the ball diameter, CD is the cylinder diameter, and CL represents the cylinder length. [b]The dimensions are provided in arbitrary units. [c]The *pow* parameter is given with its 95% confidence interval (*pow* ± 95% confidence interval). [d]The schemes are provided from the starting shape (on the left) to the final shape (on the right). [e]The shape variations were sampled as described in the "Dimensions" column. The number of shapes sampled for each shape evolution is also provided by the index (*i*). The *pow* parameters were extracted from fitting (eq 1) the plots of the CCS as a function of the shape volume. The volume was calculated using the geometry dimensions, taking into account the volume overlaps between two overlapping shapes (e.g. for ball-cylinder shapes), and was controlled using MoShade volume calculations. Schemes of the sampled shape evolutions are given, and the CCS plots can be found in the Supporting Information 1 (Figures S3 and S4). Additional shapes are mathematically treated in the text.

account. The constant *c* represents the starting point of the CCS evolution of the cluster ions, i.e., a CCS offset due to the ion core. Note that the *c* parameter for the PEO model should tend to zero, since the PEO end-chains (i.e., CH_3O- and $-H$) have a negligible impact on the CCS evolution of large PEO ions. Therefore, the generalized eq 3 can be simplified to eq 2 for the PEO model. A *pow* fitting parameter of 0.96 ± 0.04 is observed for the $[FeCl_3H(CH_2)_nCO_2]^-$ cluster model. This *pow* parameter is almost equal to 1, meaning that eq 3 is quasilinear.

Predicted Theoretical *Pow* Fitting Parameters from Precise Geometric Structure Evolutions. The aim of the CCS modeling using MoShade is to investigate CCS evolutions of simple geometric shapes whose dimensions are perfectly tuned and controlled. From these shapes, controlled shape elongations and deformations are undertaken to study their effects on the CCS evolutions. In order to generate such model shapes, shapes were designed using computer-aided design software with no theoretical chemistry *a priori* intervening (e.g., no interaction potentials or atoms modeling), highlighting then the geometrical effects on CCS evolutions. The known shape dimensions allow calculating the shape volume, which relates to the experimental polymer DP values and enables our experiment−theory comparison. Additionally, MoShade calculates projections of the shapes without modeling any drift gas particles (i.e., no interactions with gas particles or any preferential orientations within an IMS cell), similarly to the projection approximation (PA) model[31,35−37] but for purely geometric objects. Indeed, our calculations take into account neither the mass nor the ion−drift gas particle interactions. Nevertheless, the obtained ᴹᵒˢʰᵃᵈᵉCCS values are related to an underlying apparent volume, i.e., the envelope of the ion−drift gas interactions as represented by the geometric object itself, as sensed by IM−MS.

Table 1 summarizes the different shapes that were analyzed: spheres, cylinders, and combinations of spheres and cylinders (ball-cylinder). The shape deformations, such as diameter increases, cylinder elongations, or a uniform (isotropic) scaling of the entire shape, describe each considered case.

Pow Parameter for Simple Shapes: Spheres and Cylinders. Three shapes have been considered as simple shapes: spheres, cylinders with spherical ends, and cylinders with plain level ends. For cylinders, the influence of the ratio between the diameter (*D*) and the length (*L*) on the *pow* parameter has been monitored.

First, we consider the case of a perfect sphere whose diameter increases. As expected in the literature,[23,38,39] the CCS evolution of spheres results in *pow* = 2/3 (or ~0.66, Table 1, Figure S2a). This can be demonstrated mathematically from a simplified equation for a solid of revolution where only one degree of freedom θ remains (eq 4; see the development of eq S1), the projected area is constant and is

given by eq 5. This then yields the average cross-section given by eq 6 (based on eq S1).

$$\overline{S}_p = \frac{\int_{-\pi/2}^{\pi/2} S_p(\theta) 2\pi \cos\theta \, d\theta}{\int_{-\pi/2}^{\pi/2} 2\pi \cos\theta \, d\theta} = \frac{1}{2} \int_{-\pi/2}^{\pi/2} S_p(\theta) \cos\theta \, d\theta$$
$$= \int_0^{\pi/2} S_p(\theta) \cos\theta \, d\theta \qquad (4)$$

$$S_p(\theta) = \pi \frac{D^2}{4} \qquad (5)$$

where S_p is the projected area of a sphere for a given orientation and D is its diameter.

$$\overline{S}_p = \int_0^{\pi/2} \pi \frac{D^2}{4} \cdot \cos\theta \, d\theta = \pi \frac{D^2}{4} \qquad (6)$$

where $\overline{S}_p$ is the average projected area (or average cross-section).

In addition, the volume of the sphere is given by $V = \pi \frac{D^3}{6}$. In our case, the volume is proportional to the mass $V \propto m$ so that $D \propto m^{1/3}$, and therefore, $\overline{S}_p \propto m^{2/3}$ (cf., eq 6). This relation is also found using MoShade calculations of perfect spheres with increasing diameter. By comparing the theoretical and experimental pow values, the growing spheres describe the envelope of the ion−drift gas particle interactions of the common trend line (Figure 1 [PEO + 3Na$^+$]$^{3+}$). This is in agreement with the expected globular shapes of PEO at high DP values and low Coulomb repulsion.

When considering a cylinder with spherical ends, the projected area again depends only on the angle θ, given that it is also a solid of revolution (eq 4 or eq S1). The contributions of the spherical and the cylinder parts can be decoupled for the determination of the projected area, since the cylindrical part does not shadow the spherical ends differently as if they were a simple sphere. The description of the two semispherical ends is identical with the case of a simple sphere (see above). For the cylindrical part, a term in $\cos\theta$ needs to be added in order to take into account the slope with respect to the projection plane (eq 7).

$$S_p(\theta) = \pi \frac{D^2}{4} + DL\cos\theta \qquad (7)$$

where $S_p(\theta)$ is the projected area dependent on the angle θ, and D and L are the cylinder diameter and length, respectively.

Given the volume $V = \pi \frac{D^3}{6} + \pi \frac{D^2}{4} L$ of the shape, the average projected area (average cross-section) yields eq 8 (based on eq 4 or eq S1).

$$\overline{S}_p = \int_0^{\pi/2} \left(\pi \frac{D^2}{4} + DL\cos\theta \right) \cdot \cos\theta \, d\theta$$
$$= \pi \frac{D^2}{4} + DL \int_0^{\pi/2} \frac{1 + \cos 2\theta}{2} \, d\theta = \frac{\pi}{4}(D^2 + DL) \qquad (8)$$

If we consider an isotropic growth of this cylinder (i.e., D/L ratio is constant), one can assume that $D \propto L \propto m^{1/3}$ since $V \propto m$ still holds, and therefore, $\overline{S}_p \propto m^{2/3}$. Identical to the case of simple spheres, $pow = 2/3$ is expected for an isotropic growth of cylindrical objects with spherical ends. This statement can in fact be extended to all convex shapes with isotropic growth.

However, different cases must be considered for cylinders subjected to anisotropic growth. First, if the cylinder with spherical ends increases in length without increasing in diameter (i.e., D/L ratio decreases when the object grows), one finds a linear relation between the average cross-section and the mass: D = constant and $L \propto m$; therefore, $\overline{S}_p \propto m$. This also holds for cylinders with plain level ends whose length increases at constant diameter (Table 1, Figure S2b). The elongating cylinder either yields $pow = 0.91$ or can be fitted with a linear function. Such structures are well represented by the highly charged ions ([PEO + 8Na$^+$]$^{8+}$ and [PEO + 9Na$^+$]$^{9+}$, cf. Figure 1). Second, if the diameter increases for a constant length (i.e., D/L ratio increases when the object grows), the behavior is more complex: L = constant and $D \propto m^{1/2}$, therefore yielding $\overline{S}_p \propto m^{1/2}$ if the diameter D is small with respect to L. Such shapes are represented by [PEO + 4Na$^+$]$^{4+}$ ions (cf. Figure 1). However, if L is small with respect to D, then $D \propto m^{1/3}$, and therefore, $\overline{S}_p \propto m^{2/3}$. This case resembles a distorted sphere. The pow parameter should therefore be comprised between $1/2$ and $2/3$ for these cases. For cylinders with plain level ends whose diameter D increases for a constant length L (i.e., D/L ratio increases when the object grows), an intermediate pow value of around 0.80 can be found if L is small with respect to D (Table 1, Figure S2c). This would correspond to the CCS trend of [PEO + 7Na$^+$]$^{7+}$ sampled in Figure 1. Nevertheless, the fit yielding 0.80 for cylinder diameter variations is highly influenced by CCS values at large volumes (L is small with respect to D). If the diameter D is small with respect to L, pow parameters that are smaller than 0.66 can be found ($pow = 0.57$; Table 1, Figure S2d). Such pow values would then correspond to [PEO + 4 Na$^+$]$^{4+}$ and [PEO + 5Na$^+$]$^{5+}$ complexes.

Finally, we can consider the effect of the thickness of a hollow sphere. If the increase in size is isotropic (increase in diameter and in thickness), it is the same procedure as for a simple sphere: only the volume V changes. However, if considering that the thickness of the sphere stays constant (e.g., fullerenes[40,41]), the volume becomes proportional to its surface and to the thickness e (considered very small). The average projected area (cross-section) of a simple sphere still holds $\overline{S}_p = \pi \frac{D^2}{4}$ (eq 8), with $V = \pi D^2 e$ so that $D \propto m^{1/2}$ and therefore $\overline{S}_p \propto m$. This is the same behavior as a cylinder of constant diameter that increases in length. It is therefore impossible to discriminate both cases by just comparing the exponent (of the "mass" variable) of the CCS evolution fit as a function of the mass.

Pow Parameter for Composite Shapes: Ball-Cylinders. Regarding composite shapes (sphere with cylinder), several size parameters can be varied. We considered the relative ball or cylinder diameters as well as only the cylinder length and a uniform scaling of the entire shape.

For increasing ball diameters in ball-cylinder shapes, a pow parameter of 0.655 is found (Table 1, Figure S3a). This result was expected, as the cylinder becomes increasingly negligible with the large growing sphere at high volumes, thus evolving like a sphere (0.66). However, the power fit is again highly influenced by the larger CCS values at large volume values. When fitting at small ball diameters, pow values of 0.60 can be reached (Table 1, Figure S3b). In this case, the cylinder and the sphere have almost identical diameters, and the shape

resembles a cylinder. Indeed, values smaller than 0.66 were found for cylinder diameter variations when D was small with respect to L (small CCS values; Table 1, Figure S2d). This ball-cylinder value can thus be explained solely on the basis of its spherical and cylindrical components and can be predicted.

When uniformly scaling a ball-cylinder shape (isotropic growth), the *pow* parameters reach 0.67. Identical *pow* values were found for all the ball-cylinder scaling series with different initial cylinder lengths, even for cylinder lengths surpassing the ball diameter (Table 1, Figure S3c). Given the experimental errors on the experimental fit parameters (Figure 1), this result indicates that IM−MS would be unable to differentiate a sphere with increasing diameter (Figure S2a), a ball-cylinder shape with a large and increasing sphere diameter (Figure S3a), or a ball-cylinder shape with a uniform and isotropic volume increase (Figure S3c). Similar to the conclusions drawn on convex shapes, IM−MS would then only be sensitive (i.e., different *pow* values) to anisotropic shape growths.

Finally, we analyzed the effect of a cylinder length increase of a ball-cylinder shape on the CCS evolution (Table 1, Figure S3d). Not surprisingly, this resulted in a linear CCS evolution as a function of the volume, similar to an elongating cylinder. The coupled sphere does only account for an offset in CCS or volume, given that its dimensions are kept unchanged. Such shapes could be related to the model $[\text{FeCl}_3\text{H}(\text{CH}_2)_n\text{CO}_2]^-$ cluster ions, where the FeCl_3 core corresponds to the ball segment of the object and the carboxylate ligands varying in length correspond to the elongating cylinder segment of the ball-cylinder object. As observed in Figure 2, the $[\text{FeCl}_3\text{X}]^-$ complexes lead to this expected linear CCS−ligand growth relation, with a CCS offset due to the FeCl_3 core of the ion (eq 3). Using conventional MD simulations, we modeled carboxylate ions (without the FeCl_3 core) and calculated their (median) distance variations with the increasing −CH$_2$− units (calculated as the difference between an xth CH$_2$ unit and the first C from the carboxylate group $C_x - C_1$). The cylindrical-like growth of FeCl_3−carboxylate suggested by the IM−MS data is supported by the theoretical $C_x - C_1$ distances (see Figure S5). The $C_x - C_1$ trend suggests that the increasing number of −CH$_2$− units follows a cylindrical growth, in good agreement with MoShade predictions, with a slight deviation for longer chains. This deviation might be an artifact of the MD calculation (a linear experimental trend of $1/K_0$ vs the number of −CH$_2$− units was observed in literature[42]), which also does not take into account the FeCl_3 core.

For polymers, a similar trend is expected for high charge states where several monomers bearing cation charges are subjected to strong Coulomb repulsions (cylinder segment of the ball-cylinder shape), and other monomers form a cation-solvating globule with little Coulomb repulsion (ball segment of the ball-cylinder shape).

Table 2 summarizes the results of the MoShade-obtained *pow* ranges for the evolutions of simple and composite shape deformations.

Tipping Points in the CCS Evolutions. Aside from the gradual increase in the CCS described by the CCS trends that can be fitted using eqs 1 to 3, the CCS vs DP plots also show deviations from these CCS trends (Figure 1). Such CCS deviations are often referred to as DP ranges where structural rearrangements of the ions occur. A nonexhaustive shape model is represented in Table S3. Here, we induced angle variations between composite ball-cylinder shapes (three balls

Table 2. Summary of the Shape Evolutions Correlated to Their *Pow* Range[a]

pow	shape evolutions
<0.66	−cylinders with spherical or plain level ends: diameter increase at constant length with CD < CL
	−ball-cylinders: ball diameter increase with BD < CL
~0.66	−spheres: diameter increase
	−cylinders with spherical ends: isotropic growth
	−cylinders with spherical ends: diameter increase at constant length with CL < CD
	−ball-cylinders: ball diameter increase with CL < BD
	−ball-cylinders: isotropic growth
	−isotropic growth
~0.80	−cylinders with plain level ends: diameter increase at constant length with CL < CD
≥0.90	−hollow sphere of constant thickness: diameter increase
	−cylinders with spherical or plain level ends: length increase at constant diameter
	−ball-cylinders: cylinder length increase

[a]See text for more detailed *pow* values.

and two cylinders), which were able to capture a decrease in CCS for constant volumes. Future modeling attempts focusing on the structural rearrangements could include reducing the length of one cylinder segment while increasing the diameter of one ball segment.

Relation between Surface Area and CCS. *Perfectly Convex Shapes and Theory.* Owing to the well-defined and known geometries, different relations between the CCS and geometry parameters can be tested. One recurring relation that stands out is the relation between the (geometric) surface area of the shape and its CCS. Indeed, the geometric surface area equals 4 × CCS (eq 9). This can be mathematically demonstrated for perfectly convex shapes (spheres, cylinders; see eqs S6 to S10 in the Supporting Information 2 and Cauchy's theorem[43,44]). Therefore, the ratio between surface area and CCS (S/CCS ratio) acts as a descriptor of the structure concavity.

$$\bar{S}_p = \frac{S}{4} \text{ or } S = 4 \cdot \text{CCS} \tag{9}$$

where S is the area of the shape and $\bar{S}_p$ or CCS is the average cross-section.

Even if this relation can only be mathematically demonstrated for perfectly convex structures[43,44] but not for concave structures, it seems nevertheless also valid for the above-analyzed composite ball-cylinder shapes (Table 1). These shapes begin to exhibit concavities but still yield surface-to-CCS ratios $S/\text{CCS} \leq 4.06$ (calculated from results obtained using Moshade).

Shapes with Concavities: Application to Theoretical Predicted Structures of PEO. The CCS evolutions of the spheres, cylinders, and ball-cylinder shapes, which are all adequate with the $S = 4 \cdot \text{CCS}$ relation, describe the "natural" CCS evolution of polymer ions when following one given charge state as a function of the DP (i.e., CCS trends[16,23]). This relation could then be used to facilitate the structure screening in conventional computational chemistry methods.

In order to check its applicability, we generated a pool of theoretical $[\text{PEO}_{\text{DP}=14} + 1\text{Na}^+]^{1+}$ and $[\text{PEO}_{\text{DP}=14} + 2\text{Na}^+]^{2+}$ structures (see Materials and Methods section for more information). Owing to the lack of Coulomb repulsion,[39] the shapes of the 1+ complexes (i.e., $[\text{PEO}_{\text{DP}=14} + 1\text{Na}^+]^{1+}$) are

globular, and the three-dimensional representation (volume envelope) should yield S/CCS ratios close to 4. We sampled two globular shapes (from optimized structures) of $[PEO_{DP=14} + 1Na^+]^{1+}$ and one shape considered as aberrant, which was manually distorted (elongated; i.e., a nonoptimized structure). The globular shapes yielded $S/CCS = 4.07$, whereas the elongated shape yielded $S/CCS = 4.13$ (Table S1). The globular shapes leading to S/CCS values closest to 4 can thus be correlated to the energy-optimized structure geometries.

Out of the 18 sampled structures for $[PEO_{DP=14} + 2Na^+]^{2+}$ (Table S1), the three most stable structures (shapes 1, 4, and 10 in Figure 3) exhibited the lowest S/CCS ratio, while ratios larger than 4 were observed for nonoptimized structures. Shapes 1 and 4 led to $S/CCS = 4.09$, and shape 10 led to $S/CCS = 4.08$.
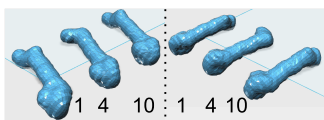
Figure 3. Shapes of $[PEO_{DP=14} + 2Na^+]^{2+}$ (represented in two different viewing angles) leading to the lowest S/CCS ratios out of all the 18 sampled shapes (see Table S1). Shapes 1 and 4 yield $S/CCS = 4.09$, and shape 10 yields $S/CCS = 4.08$. Their underlying structures correspond to optimized or nearly optimized structures using the MM2 and MMFF94 force fields.

Interestingly, even if the shapes extracted from conventional computational chemistry structures of PEO exhibit numerous concavities, the S/CCS ratios closest to 4 are found for their most stable structures. Therefore, it seems that the S/CCS ratio calculation could be a new tool for rapid structure screening in conventional computational chemistry methods as long as the shapes are not too concave. The limitations of structure filtering using the S/CCS ratio due to concavities nevertheless still need to be explored in more detail.

It should still be noted that the S/CCS relation is in fact independent of the CCS evolution and can therefore be used for single data point analysis. Indeed, one can calculate a S/CCS ratio for any structure, outgrowing the starting point of polymer ions and the need to sample CCS evolutions instead of single-point absolute CCS values. In this case, experimental CCS measurements must then be accurate[39,45] to be able to deduce their correct surface area. Structural elucidations could then be performed by matching experimentally deduced surface areas with a set of surface areas obtained from a theoretical predicted structure pool, which obliviates the calculation of theoretical CCS values. However, the theoretical inspected surface area should be accurately represented by the modeled 3D structures (e.g., by their electron density isosurfaces[46]) in order for them to pass the structure filter. Structure pools could then be "randomly" generated using computational chemistry methods[13,14,47−49] without the need for precise knowledge of experimental/simulation (ion or bath) temperatures. When finding suitable structures (using eq 9, $S = 4 \cdot CCS$), the effective experimental ion temperature could then be retraced.

CCS Evolution of the Highest-Charge-Density Ions: Beads on a String. Besides following the CCS evolutions of fixed charge states as a function of the DP, CCS evolutions when varying the charge states can also be analyzed. The literature predicts through theoretical chemistry that the highest charge states have a "beads-on-a-string" conformation.[13,14] The beads-on-a-string conformation is made out of monomer beads each solvating a cation and of monomer strings spacing the beads to reduce Coulomb repulsion. To test this structure hypothesis in our case, we created a CCS evolution, which is not based on a single fixed charge state by experimentally considering only the first detected ion of each charge state. In other words, we will consider only the first ion of each charge state series, which therefore bears the highest charge density. Their CCS values are used to generated a new CCS−DP trend, according to the charge state, for a given polymer. This CCS evolution can then be represented by a structure evolution of beads on a string where every new complex has an additional "bead-string" unit (or segment) for solvating the additional cation and Coulomb repulsion (see Figure 4b).

In order to experimentally sample the first complexes of multiple charge states, PEO polymers were analyzed at small DP values, covering DP 8 to ∼70 (750 and 2000 g/mol samples; Figure 4a). Fitting this CCS evolution yields a *pow* value of around 0.93. It can also be fitted using a linear fit function, which yields a slope of 13.9 (based on eq 2). Eq 3, i.e., eq 2 with a CCS offset, could also be used to take into consideration CCS contributions from the chain ends, especially for the smallest PEO polymer ions.

Modeled beads-on-a-string shapes, with different bead diameters and cylinder lengths or cylinder diameters, also yield linear CCS evolutions as a function of the volume. It should be noted that these structures do not follow the $S = 4 \cdot CCS$ relation described in eq 9 ($S/CCS \approx 4.10-4.40$), given their increased concavity. Shape variations that both followed eq 9 and exhibited linear CCS evolutions could be associated to elongating cylinders (convex shapes, see discussion above). Nevertheless, we still considered the beads-on-a-string structures as suggested by computational chemistry, given that our MoShade calculations can only yield "coarse grain" information on the volume envelope of the ion−drift gas particle interactions and may falsely discard beads-on-a-string fine-structures at the atomic level.

When comparing bead-string units with identical volume but different cylinder dimensions (Figure 4c, "BD 8 CD 3 CL 6" and "BD 8 CD 2.12 CL 12"), the CCS values change and the slope changes. Thus, only specific (nonaberrant) bead-string dimensions can lead to these specific slope values. Nevertheless, the cylinder diameter can be changed with only a negligible effect on the slope (Figure 4c, "BD 20 CD n" with n = 3 or 6 and CL = 3, 6, or 12). The cylinder length, which spaces each bead, and the bead diameter, leading to volume and CCS jumps, are the main parameters influencing the slopes. The beads-on-a-string linear evolution can thus be compared to the MoShade-obtained "coarse grain" cylinder elongations where the bead diameter defines the cylinder diameter and the string length defines the cylinder elongation.

When extracting the same CCS evolution (first complex of each charge state) for a different polymer, poly-(ethoxyphosphate) PEtP (Figure 4; PEtP scheme in Figure S1), a *pow* value of 0.98 can be attained. The slope of the linear fit yields a value of 35.3. The ratio of the PEtP and PEO slope values can then be calculated (PEtP/PEO slope ratio = 2.54; Table 3). Similar ratios can be calculated for all modeled beads-on-a-string CCS evolution slopes (Table S2), and matching experimental and theoretical ratios can be found. Two ratios of modeled beads-on-a-string shapes resulted in
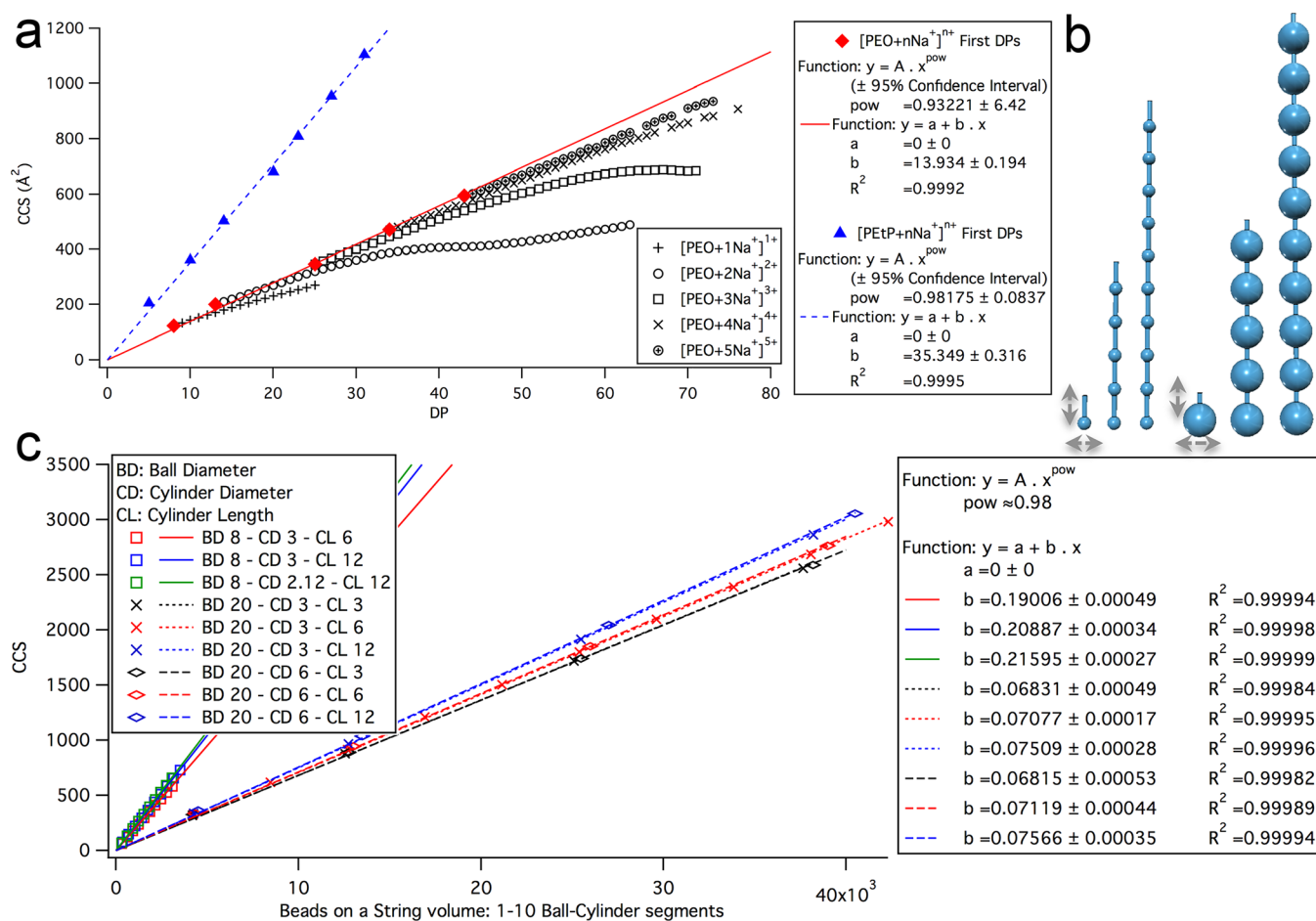
Figure 4. Comparison of the experimental CCS−DP plots considering only the lowest m/z (or DP/z) adduct of each charge state of sodiated PEO and PEtP ions with the theoretical CCS−volume plots from computer-aided design software-constructed shapes computed using MoShade. (a) Black markers depict the CCS evolution of PEO. The red markers represent the first PEO complex of each charge state, with the red plain line representing its linear fit function. This non-charge-state-based CCS evolution equates a series of the highest-charged complexes. The blue markers depict the same series for PEtP polymers (see Figure S1), with its linear fit function depicted as blue dotted line. (b) Describes the modeled shape evolutions of such complexes: each considered complex is made up of an additional bead-string unit (or segment). (c) Plot of the MoShade CCS (in arbitrary units) as a function of the volume of different beads-on-a-string shapes. The linear fit coefficients and the coefficient of determination are given for the different fits. BD represents the ball diameter, CD represents the cylinder diameter, and CL represents the cylinder length.

Table 3. Slopes of the Linear Fits (b) of Modeled Beads-on-a-String Shapes (See Figure 4c) and of Experimental PEtP and PEO Evolutions of Every First Complex of Each Charge State (See Figure 4a)[a]

bead-string unit (BD−CD−CL)	b (slope)	$b_{8-3-6}/b_{20-3-12}$; $b_{8-3-6}/b_{20-6-12}$
8−3−6	0.190	
20−3−12	0.075	2.531
20−6−12	0.076	2.512
polymer	b (slope)	b_{PEtP}/b_{PEO}
PEtP	35.349	
PEO	13.934	2.537

[a]The ratios of the modeled shapes are in close agreement with the experimental slope ratio of PEtP/PEO. Table S2 lists all the modeled shapes with all the calculated slope ratios. Bead diameter, cylinder diameter, and cylinder length are abbreviated as BD, CD, and CL, respectively.

close correlations with the experimental PEtP/PEO slope ratio (Table 3). The only difference in the two modeled shape evolutions is the cylinder diameter of the beads on a string, as expected due to its lesser influence on the slope. We are thus able to provide relative bead-string dimensions of PEO compared to the PEtP bead-string units (or vice versa). PEO should thus exhibit a bead diameter of 20 if PEtP exhibits a bead diameter of 8 (in arbitrary units). The cylinder length should be doubled when going from PEtP to PEO.

Knowing these relations between the envelope dimensions of PEO and PEtP, theoretical chemistry modeling methods (MM, MD, DFT···) could find enhanced hypotheses or interaction potential descriptions for structure calculations. The considered number of PEO monomer units needs to fill up the given volume (or envelope) and if volume corrections need to be performed, proportional corrections need also to be applied to the PEtP volume envelope. The missing dimension, i.e., the cylinder diameter, which only plays a minor role in the CCS evolution, will then be intrinsically defined through the width of the monomer units making up the string segment. We are hence able to give precise volume dimensions of the polymer ions without modeling any chemical interactions or atoms.

CONCLUSIONS AND PROSPECTS

The aim of this study was to link experimental CCS evolutions with the increase of repeating units to potential geometric shape evolutions of the considered ions, without the use of computational chemistry (for structural elucidation and CCS calculation) and without requiring accurate measured or absolute CCS values. We present a new orthogonal method for relating experimental IM−MS measurements to theoretical shape interpretations, without atomistic considerations. New software named MoShade allowed calculating CCS values of (mesh-represented) shapes through projection calculations, which yield the volume envelope of the ion−drift gas particle interactions. Our MoShade calculations are purely geometric shape-based analyses. No atoms or chemical interactions were modeled.

While we showed in a separate paper that the parameter A of the CCS fit equation $CCS = A \cdot X^{pow}$ (or $CCS = A \cdot X^{pow}+c$ with X being the number of repetitive units) contains physicochemical information of the analyzed systems,[16,23] we showed here that the *pow* parameter contains structural shape information accessible through IM−MS.

Basic shape geometries and shape variations were mathematically considered or computationally modeled. The modeled shape dimension variations yielded theoretical CCS trends, which were also fitted using power fit functions. Through comparisons between the experimental and mathematically obtained/modeled *pow* parameters, we were able to extract several shape evolutions which fitted with experimentally obtained CCS evolutions (summary in Table 2).

Generally, an isotropic growth of shapes should always exhibit a 2/3 *pow* value, as can be mathematically demonstrated. IM−MS CCS evolutions (*pow* values) for such shape evolutions are thus indistinguishable. Nonisotropically growing cylinders, cylinders with plain or spherical ends, and composite ball-cylinder shape evolutions were modeled, and theoretical *pow* parameter values were obtained. The models hold for both chemically homogeneously growing synthetic polymers as well as for chemically heterogeneous small cluster ions. A decrease in CCS, such as those observed during structural rearrangements, could be obtained by angle variations between composite ball-cylinder shapes.

Moreover, we suggested investigating the relation $S/CCS = 4$ as a potential structure filter in more detail. This relationship between the CCS and the geometric surface area holds for perfectly convex shapes. Validated first using the structures from structure evolutions of the CCS trend analysis, we found that energy-optimized structures of ions yielded a closer match to $S/CCS = 4$ than to nonoptimized structures. This relation could then potentially act as a structure filter for conventional computational chemistry methods, even for nonpolymeric systems as it is applicable to absolute CCS values[39,45] without the need for CCS evolutions. In other words, this relation can also be used for single data point analysis.

Finally, we investigated literature-advocated beads-on-a-string shapes.[13,14] We were able to extract precise shape dimensions of the bead-string units for two polymer systems (PEtP, PEO) without modeling any chemical interactions. Given that the number of monomer units making up the beads-on-a-string shape volume is known (MS identification), theoretical chemistry interaction potentials could be refined using the here-developed CCS trend and shape interpretation methodology.

The next step in the development of the CCS trend analysis could now focus on more heterogeneously growing systems, where the apparent densities[16,23] change during the shape growth.

ASSOCIATED CONTENT

Supporting Information file 1: Description and schemes of the polymers and anionic clusters used in the study; ion mobility−mass spectrometry experimental conditions; short mathematical background of MoShade and illustration of a shape projection; detailed plots and fits of the collision cross-sections of all the discussed shape evolutions calculated using MoShade; figure of the distance evolutions by −CH$_2$− unit increments for MD-modeled carboxylate ions; illustrations of the MM2- and MMFF94-derived MoShade shape inputs of PEO and table with their S/CCS ratios; table with the slopes of the linear fits obtained from modeled beads-on-a-string shape evolutions and their ratio calculations; illustration of angle variations for composite ball-cylinder shapes and table with their MoShade collision cross-sections (PDF)

Supporting Information file 2: Mathematical developments of convex shapes; detailed mathematical developments and projection illustrations of MoShade (PDF)

AUTHOR INFORMATION

Corresponding Author

Jean R. N. Haler − *Mass Spectrometry Laboratory, University of Liège, MolSys Research unit, B-4000 Liège, Belgium; Luxembourg Institute of Science and Technology − LIST, Materials Research & Technology MRT Department, L-4422 Belvaux, Luxembourg;* orcid.org/0000-0002-0315-1846; Email: jean.haler@uliege.be

Authors

Eric Béchet − *Aerospace & Mechanical Engineering Department, Computer-aided Geometric Design, University of Liège, B-4000 Liège, Belgium*

Christopher Kune − *Mass Spectrometry Laboratory, University of Liège, MolSys Research unit, B-4000 Liège, Belgium;* orcid.org/0000-0002-3010-8173

Johann Far − *Mass Spectrometry Laboratory, University of Liège, MolSys Research unit, B-4000 Liège, Belgium;* orcid.org/0000-0003-1208-6262

Edwin De Pauw − *Mass Spectrometry Laboratory, University of Liège, MolSys Research unit, B-4000 Liège, Belgium;* orcid.org/0000-0003-3475-1315

Complete contact information is available at:
https://pubs.acs.org/10.1021/jasms.1c00266

ACKNOWLEDGMENTS

■ REFERENCES

(1) Hines, K. M.; Ross, D. H.; Davidson, K. L.; Bush, M. F.; Xu, L. Large-Scale Structural Characterization of Drug and Drug-Like Compounds by High-Throughput Ion mobility−mass Spectrometry. *Anal. Chem.* **2017**, *89* (17), 9023−9030.

(2) Boschmans, J.; Lemière, F.; Sobott, F. Analyzing Complex Mixtures of Drug-like Molecules: Ion Mobility as an Adjunct to Existing Liquid Chromatography-(Tandem) Mass Spectrometry Methods. *J. Chromatogr. A* **2017**, *1490*, 80−88.

(3) Kune, C.; Far, J.; De Pauw, E. Accurate Drift Time Determination by Traveling Wave Ion Mobility Spectrometry: The Concept of the Diffusion Calibration. *Anal. Chem.* **2016**, *88* (23), 11639−11646.

(4) Czerwinska, I.; Far, J.; Kune, C.; Larriba-Andaluz, C.; Delaude, L.; De Pauw, E. Structural Analysis of Ruthenium−Arene Complexes Using Ion Mobility Mass Spectrometry, Collision-Induced Dissociation, and DFT. *Dalt. Trans.* **2016**, *45* (15), 6361−6370.

(5) Kune, C.; Haler, J. R. N.; Far, J.; De Pauw, E. Effectiveness and Limitations of Computational Chemistry and Mass Spectrometry in the Rational Design of Target-Specific Shift Reagents for Ion Mobility Spectrometry. *ChemPhysChem* **2018**, *19* (21), 2921−2930.

(6) Massonnet, P.; Haler, J. R. N.; Upert, G.; Degueldre, M.; Morsa, D.; Smargiasso, N.; Mourier, G.; Gilles, N.; Quinton, L.; De Pauw, E. Ion Mobility-Mass Spectrometry as a Tool for the Structural Characterization of Peptides Bearing Intramolecular Disulfide Bond-(S). *J. Am. Soc. Mass Spectrom.* **2016**, *27* (10), 1637−1646.

(7) López, A.; Tarragó, T.; Vilaseca, M.; Giralt, E. Applications and Future of Ion Mobility Mass Spectrometry in Structural Biology. *New J. Chem.* **2013**, *37* (5), 1283.

(8) Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T. Ion Mobilityg-Mass Spectrometry Reveals a Conformational Conversion from Random Assembly to β-Sheet in Amyloid Fibril Formation. *Nat. Chem.* **2011**, *3* (2), 172−177.

(9) Haler, J. R. N.; Massonnet, P.; Far, J.; Upert, G.; Gilles, N.; Mourier, G.; Quinton, L.; De Pauw, E. Can IM−MS Collision Cross-Sections of Biomolecules Be Rationalized Using Collision Cross-Section Trends of Polydisperse Synthetic Homopolymers? *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 990.

(10) Delvaux, C.; Massonnet, P.; Kune, C.; Haler, J. R. N.; Upert, G.; Mourier, G.; Gilles, N.; Quinton, L.; De Pauw, E.; Far, J. Combination of Capillary Zone Electrophoresis-Mass Spectrometry, Ion Mobility-Mass Spectrometry, and Theoretical Calculations for Cysteine Connectivity Identification in Peptides Bearing Two Intramolecular Disulfide Bonds. *Anal. Chem.* **2020**, *92*, 2425.

(11) Haler, J. R. N.; Far, J.; Aqil, A.; Claereboudt, J.; Tomczyk, N.; Giles, K.; Jérôme, C.; De Pauw, E. Multiple Gas-Phase Conformations of a Synthetic Linear Poly(Acrylamide) Polymer Observed Using Ion Mobility-Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (11), 2492−2499.

(12) Haler, J. R. N.; Massonnet, P.; Chirot, F.; Kune, C.; Comby-Zerbino, C.; Jordens, J.; Honing, M.; Mengerink, Y.; Far, J.; Dugourd, P.; De Pauw, E. Comparison of Different Ion Mobility Setups Using Poly (Ethylene Oxide) PEO Polymers: Drift Tube, TIMS, and T-Wave. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (1), 114−120.

(13) Larriba, C.; Fernandez De La Mora, J. The Gas Phase Structure of Coulombically Stretched Polyethylene Glycol Ions. *J. Phys. Chem. B* **2012**, *116* (1), 593−598.

(14) Trimpin, S.; Plasencia, M.; Isailovic, D.; Clemmer, D. E. Resolving Oligomers from Fully Grown Polymers with IMS-MS. *Anal. Chem.* **2007**, *79* (21), 7965−7974.

(15) De Winter, J.; Lemaur, V.; Ballivian, R.; Chirot, F.; Coulembier, O.; Antoine, R.; Lemoine, J.; Cornil, J.; Dubois, P.; Dugourd, P.; Gerbaux, P. Size Dependence of the Folding of Multiply Charged Sodium Cationized Polylactides Revealed by Ion Mobility Mass Spectrometry and Molecular Modelling. *Chem. - A Eur. J.* **2011**, *17* (35), 9738−9745.

(16) Haler, J. R. N.; Far, J.; De La Rosa, V. R.; Kune, C.; Hoogenboom, R.; De Pauw, E. Using Ion Mobility-Mass Spectrometry to Extract Physicochemical Enthalpic and Entropic Contributions from Synthetic Polymers. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (1), 330−339.

(17) Boschmans, J.; Jacobs, S.; Williams, J. P.; Palmer, M.; Richardson, K.; Giles, K.; Lapthorn, C.; Herrebout, W. A.; Lemière, F.; Sobott, F. Combining Density Functional Theory (DFT) and Collision Cross-Section (CCS) Calculations to Analyze the Gas-Phase Behaviour of Small Molecules and Their Protonation Site Isomers. *Analyst* **2016**, *141* (13), 4044−4054.

(18) Counterman, A. E.; Clemmer, D. E. Gas Phase Polyalanine: Assessment of i → i + 3 and i → i + 4 Helical Turns in [Alan + 4H]4+ (n = 29−49) Ion. *J. Phys. Chem. B* **2002**, *106* (46), 12045−12051.

(19) Kinnear, B. S.; Kaleta, D. T.; Kohtani, M.; Hudgins, R. R.; Jarrold, M. F. Conformations of Unsolvated Valine-Based Peptides. *J. Am. Chem. Soc.* **2000**, *122* (38), 9243−9256.

(20) Kune, C.; Delvaux, C.; Haler, J. R. N. N.; Quinton, L.; Eppe, G.; De Pauw, E.; Far, J. A Mechanistic Study of Protonated Aniline to Protonated Phenol Substitution Considering Tautomerization by Ion Mobility Mass Spectrometry and Tandem Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (11), 2238−2249.

(21) Haler, J. R. N.; Lemaur, V.; Far, J.; Kune, C.; Gerbaux, P.; Cornil, J.; De Pauw, E. Sodium Coordination and Protonation of Poly(Ethoxy Phosphate) Chains in the Gas Phase Probed by Ion Mobility-Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (3), 633−641.

(22) Lapthorn, C.; Pullen, F. S.; Chowdhry, B. Z.; Wright, P.; Perkins, G. L.; Heredia, Y. How Useful Is Molecular Modelling in Combination with Ion Mobility Mass Spectrometry for 'small Molecule" Ion Mobility Collision Cross-Sections?'. *Analyst* **2015**, *140* (20), 6814−6823.

(23) Haler, J. R. N.; Morsa, D.; Lecomte, P.; Jérôme, C.; Far, J.; De Pauw, E. Predicting Ion Mobility-Mass Spectrometry Trends of Polymers Using the Concept of Apparent Densities. *Methods* **2018**, *144*, 125.

(24) Haler, J. R. N.; de la Rosa, V. R.; Massonnet, P.; Far, J.; Hoogenboom, R.; De Pauw, E. Fundamental Studies on Poly(2-Oxazoline) Side Chain Isomers Using Tandem Mass Spectrometry and Ion Mobility-Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (7), 1220−1228.

(25) Kokubo, S.; Vana, P. Obtaining the Dielectric Constant of Polymers from Doubly Charged Species in Ion-Mobility Mass Spectrometry. *Macromol. Chem. Phys.* **2017**, *218* (17), 1700126.

(26) Kokubo, S.; Vana, P. Easy Access to the Characteristic Ratio of Polymers Using Ion-Mobility Mass Spectrometry. *Macromol. Chem. Phys.* **2017**, *218* (1), 1600373.

(27) Meier, F.; Köhler, N. D.; Brunner, A.-D.; Wanka, J.-M. H.; Voytik, E.; Strauss, M. T.; Theis, F. J.; Mann, M. Deep Learning the Collisional Cross Sections of the Peptide Universe from a Million Experimental Values. *Nat. Commun.* **2021**, *12* (1), 1185.

(28) Vega-Márquez, B.; Nepomuceno-Chamorro, I.; Jurado-Campos, N.; Rubio-Escudero, C. Deep Learning Techniques to Improve the Performance of Olive Oil Classification. *Front. Chem.* **2020**, *7*, 929.

(29) Plante, P.-L.; Francovic-Fontaine, É.; May, J. C.; McLean, J. A.; Baker, E. S.; Laviolette, F.; Marchand, M.; Corbeil, J. Predicting Ion Mobility Collision Cross-Sections Using a Deep Neural Network: DeepCCS. *Anal. Chem.* **2019**, *91* (8), 5191−5199.

(30) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33−38.

(31) 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.* **2011**, *308* (1), 1−10.

(32) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179−5197.

(33) Forsythe, J. G.; Petrov, A. S.; Walker, C. A.; Allen, S. J.; Pellissier, J. S.; Bush, M. F.; Hud, N. V.; Fernández, F. M. Collision Cross Section Calibrants for Negative Ion Mode Traveling Wave Ion Mobility-Mass Spectrometry. *Analyst* **2015**, *140* (20), 6853−6861.

(34) Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L.P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D.G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan, C. J.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinos, K.; Valentine, S. J.; Wyttenbach, T. Recommendations for Reporting Ion Mobility Mass Spectrometry Measurements. *Mass Spectrom. Rev.* **2019**, *38* (3), 291−320.

(35) 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.* **1996**, *100* (40), 16082−16086.

(36) Bleiholder, C.; Johnson, N. R.; Contreras, S.; Wyttenbach, T.; Bowers, M. T. Molecular Structures and Ion Mobility Cross Sections: Analysis of the Effects of He and N2 Buffer Gas. *Anal. Chem.* **2015**, *87* (14), 7196−7203.

(37) Shvartsburg, A. A.; Mashkevich, S. V.; Baker, E. S.; Smith, R. D. Optimization of Algorithms for Ion Mobility Calculations. *J. Phys. Chem. A* **2007**, *111* (10), 2002−2010.

(38) Ruotolo, B. T.; Benesch, J. L. P. P.; Sandercock, A. M.; Hyung, S.-J. J.; Robinson, C. V. Ion Mobility-Mass Spectrometry Analysis of Large Protein Complexes. *Nat. Protoc.* **2008**, *3* (7), 1139−1152.

(39) Haler, J. R. N.; Kune, C.; Massonnet, P.; Comby-Zerbino, C.; Jordens, J.; Honing, M.; Mengerink, Y.; Far, J.; De Pauw, E. Comprehensive Ion Mobility Calibration: Poly(Ethylene Oxide) Polymer Calibrants and General Strategies. *Anal. Chem.* **2017**, *89* (22), 12076−12086.

(40) Von Helden, G.; Hsu, M. T.; Gotts, N.; Bowers, M. T. Carbon Cluster Cations with up to 84 Atoms: Structures, Formation Mechanism, and Reactivity. *J. Phys. Chem.* **1993**, *97* (31), 8182−8192.

(41) Von Helden, G.; Hsu, M. T.; Kemper, P. R.; Bowers, M. T. Structures of Carbon Cluster Ions from 3 to 60 Atoms: Linears to Rings to Fullerenes. *J. Chem. Phys.* **1991**, *95* (1991), 3835−3837.

(42) Gao, H.; Niu, W.; Hong, Y.; Xu, B.; Shen, C.; Huang, C.; Jiang, H.; Chu, Y. Negative Photoionization Chloride Ion Attachment Ion Mobility Spectrometry for the Detection of Organic Acids. *RSC Adv.* **2014**, *4* (109), 63977−63984.

(43) Cauchy, A. Mémoire Sur La Rectification Des Courbes et La Quadrature Des Surfaces Courbes. *Mem. Acad. Sci. Paris* **1850**, *22*, 3.

(44) SANGWINE-YAGER, J. R. Mixed Volumes. *Handbook of Convex Geometry* **1993**, 43−71.

(45) Stow, S. M.; Causon, T. J.; Zheng, X.; Kurulugama, R. T.; Mairinger, T.; May, J. C.; Rennie, E. E.; Baker, E. S.; Smith, R. D.; McLean, J. A.; Hann, S.; Fjeldsted, J. C. An Interlaboratory Evaluation of Drift Tube Ion Mobility-Mass Spectrometry Collision Cross Section Measurements. *Anal. Chem.* **2017**, *89* (17), 9048−9055.

(46) Alexeev, Y.; Fedorov, D. G.; Shvartsburg, A. A. Effective Ion Mobility Calculations for Macromolecules by Scattering on Electron Clouds. *J. Phys. Chem. A* **2014**, *118* (34), 6763−6772.

(47) Schenk, E. R.; Nau, F.; Fernandez-Lima, F. Theoretical Predictor for Candidate Structure Assignment from IMS Data of Biomolecule-Related Conformational Space. *Int. J. Ion Mobil. Spectrom.* **2015**, *18* (1−2), 23−29.

(48) Molano-Arevalo, J. C.; Jeanne Dit Fouque, K.; Pham, K.; Miksovska, J.; Ridgeway, M. E.; Park, M. A.; Fernandez-Lima, F. Characterization of Intramolecular Interactions of Cytochrome c Using Hydrogen-Deuterium Exchange-Trapped Ion Mobility Spectrometry-Mass Spectrometry and Molecular Dynamics. *Anal. Chem.* **2017**, *89* (17), 8757−8765.

(49) Schenk, E. R.; Almeida, R.; Miksovska, J.; Ridgeway, M. E.; Park, M. A.; Fernandez-Lima, F. Kinetic Intermediates of Holo-and Apo-Myoglobin Studied Using Hdx-Tims-Ms and Molecular Dynamic Simulations. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (4), 555−563.

Geometric Analysis of Shapes in Ion Mobility-Mass Spectrometry

Jean R. N. Haler[1,2,†], Eric Béchet[3,†], Christopher Kune[1], Johann Far[1], Edwin De Pauw[1]*

[1]Mass Spectrometry Laboratory, University of Liège, MolSys Research unit, Quartier Agora, Allée du Six Aout 11, B-4000 Liège, Belgium

[2]Luxembourg Institute of Science and Technology – LIST, Materials Research & Technology MRT Department, L-4422 Belvaux, Luxembourg

[3]Aerospace & Mechanical Engineering Department, Computer-aided Geometric Design, University of Liège, B-4000 Liège, Belgium

[†]J.R.N. Haler and E. Béchet equally contributed to this work

*Corresponding author email address: jean.haler@uliege.be

Supporting Information 1

Materials and Methods

Polymers.

Poly(ethylene oxide) PEO (CH₃O-PEO-H) polymers (Figure S1) were bought from Sigma-Aldrich (St. Louis, USA). Average molar masses of the samples were 750, 2000 and 5000 g/mol. Poly(2-ethoxy-1,3,2-dioxapholane 2-oxide) PEtP (or poly(ethoxyphosphate), Figure S1) was synthesized according to literature.[1,2]
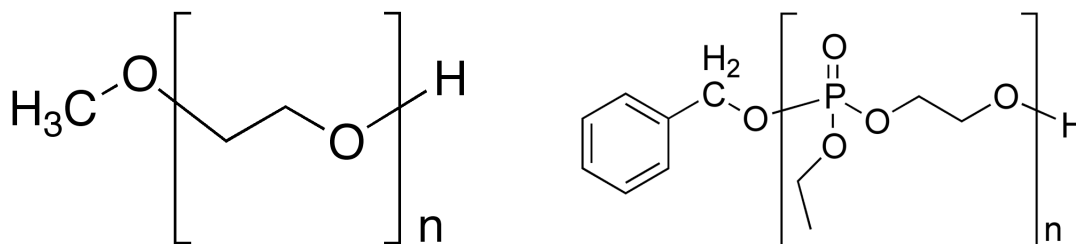
Figure S1: Poly(ethylene oxide) PEO (left) and poly(2-ethoxy-1,3,2-dioxaphospholane 2-oxide) or poly(ethoxyphosphate) PEtP polymer (right)

Polymer Ion Mobility-Mass Spectrometry.

The polymer samples were dissolved in pure methanol (Biosolve) spiked with sodium cations (NaCl salt) to obtain concentrations of around 10^{-6} to 5×10^{-6} M in polymer-sodium complexes.

The samples were infused into a traveling wave ion mobility mass spectrometer (Synapt G2 HDMS from Waters, UK) equipped with an electrospray ionization source (ESI) used in positive mode. The capillary voltage was set to 3 kV, the sampling cone voltage was set to 40 V and the extraction cone was 4 V. The desolvation gas flow was 500 L/h. The source and desolvation temperatures were set to 100 °C and 200 °C, respectively. The trap collision energy and the transfer collision energy (CE) were set to 4 V and 2 V,

respectively. The trap bias was 45 V. The IMS wave height was 40 V and the wave speed was set to 1200 m/s. The trap argon gas flow was 2 mL/min, the helium gas flow was 180 mL/min, and the nitrogen pressure in the IM cell was set to 2.6 mbar.

In order to convert drift times into CCS values, a CCS calibration procedure was followed[3], using reference values obtained in He to calibrate N_2 T-Wave measurements[4–9]. IM-MS data were interpreted using Waters' MassLynx 4.1 software. The arrival time distribution (ATD) peaks were fitted using PeakFit v.4.11 to extract accurate drift times. Data processing was performed using Excel 2011 and IgorPro 6.37.

$[FeCl_3H(CH_2)_nCO_2]^-$ clusters Ion Mobility-Mass Spectrometry.

Ferric(III) chloride, $FeCl_3$, were bought from Sigma-Aldrich (St. Louis, USA). Carboxylic acid used in this study (i.e., acetic acid, propanoic acid, butanoic acid, hexanoic acid, octanoic acid and dodecanoic acid were kindly provided by Professor L. Delaude and Professor A. Demonceau from the Laboratory of Catalysis (University of Liege). $FeCl_3$ were solubilized in methanol from Sigma-Aldrich (St. Louis, USA). Acetic acid, propanoic acid, butanoic acid were solubilized in water while hexanoic acid, octanoic acid and dodecanoic acid were solubilized in methanol. Solution of $FeCl_3$ and one carboxylic acid (one solution for each carboxylic acid) were prepared in 50% methanol solution to reach a final concentration of 5 and 50 µM, respectively.

The clusters were infused in the same ion mobility mass spectrometer as the polymer samples. The capillary voltage was set to -2.2 kV, the sampling cone voltage was set to -30 V and the extraction cone was -3 V. The desolvation gas flow was 500 L/h. The source and desolvation temperatures were set to 150 °C and 200 °C, respectively. The trap bias

was 35 V. The IMS wave height was 40 V and the wave speed was set to 2100 m/s. The nitrogen pressure in the IMS cell was set to 2.58 mbar. The helium window (before TWIMS cell) was filled with helium at a pressure of 5.4×10^{-1} mbar and the TRAP and TRANSFER cells, placed before and after the TWIMS cell respectively, were filled with argon at a pressure of 2.5×10^{-1} mbar. The TWIMS instrument was calibrated in negative mode using polyalanine anions as calibrating substances. The $^{TW, 1ry}CCS_{N2}$ values were reported by Forsythe and coworkers[10] (in N_2), and were considered for this calibration to obtain the $^{TW, 2ry}CCS_{N2}$ values (secondary calibration from TWIMS values[11]) of $[FeCl3H(CH2)nCO2]-$ clusters ions.

MoShade mathematical background.

In brief, MoShade calculates the projected area of the shape by sampling its orientation based on the angle θ and the rotation around the axis ϕ (Figure 1). Eq. S1 yields the average projected area $\bar{S}_p$. $\bar{S}_p$ is the integral of the cross-section (projected area) for each orientation (θ, ϕ), normalized (or weighted) by the probability to find the shape in one orientation or the other. This probability depends on the solid angle covered by an infinitesimal area on the unit sphere. This area is $2\pi \cos\theta \, d\theta$. The projected area is therefore integrated and divided by the integral of the probability density. Because of symmetries, we can restrict the computation of the average $\bar{S}_p$ to 1/4 of the total solid angle (sphere) covered by the variables (θ, ϕ). Eq. S1 can be rewritten as Eq. S2 because of symmetries, mainly because a shadow 'from below' has the same area as a shadow 'from above' for any given shape.

$$\overline{S}_p = \frac{\int\limits_{-\pi}^{\pi}\int\limits_{-\pi/2}^{\pi/2} S_p(\theta,\phi)\,2\pi\cos\theta\,d\theta\,d\phi}{\int\limits_{-\pi}^{\pi}\int\limits_{-\pi/2}^{\pi/2} 2\pi\cos\theta\,d\theta\,d\phi} = \frac{1}{4\pi}\int\limits_{-\pi}^{\pi}\int\limits_{-\pi/2}^{\pi/2} S_p(\theta,\phi)\cos\theta\,d\theta\,d\phi \qquad (S1)$$

$$\overline{S}_p = \frac{1}{\pi}\int\limits_{0}^{\pi}\int\limits_{0}^{\pi/2} S_p(\theta,\phi)\cos\theta\,d\theta\,d\phi \qquad (S2)$$
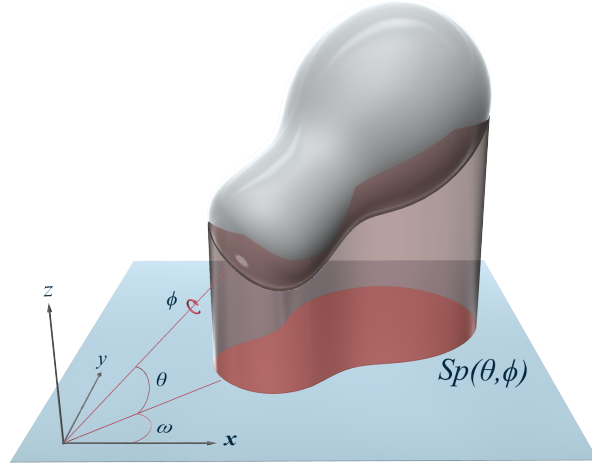
Figure S2: Illustration of a shape projection (in red below the white/gray shape) as performed by MoShade. The angles θ and rotation axis ϕ describe the shape's orientation and are sampled during the projection calculations. ω is the angle in the *(x,y)* plane. The shape input is given as a mesh and the projected area is calculated as the sum of the projections of each of the mesh's triangles (see text).

If one wishes to use numerical integration instead of trying to integrate analytically, one may use Eq. S3 as simple quadrature of Eq. S2, with a uniform repartition of the N samples on $\theta_i,\phi_i \in \left[0...\frac{\pi}{2}\right]\times[0...\pi]$. The n samples parameterizing a MoShade calculation hence lead to *N=(2n)²*.

$$\overline{S_p} = \frac{1}{\pi} \int_0^\pi \int_0^{\pi/2} \quad S_p(\theta, \phi) \cdot \cos\theta d\theta d\phi \approx \frac{1}{\pi} \sum_i \quad \frac{\pi^2}{2N} S_p(\theta_i, \phi_i) \cdot \cos\theta_i \tag{S3}$$

To evaluate $S_p(\theta_i, \phi_i)$, the surface of the shape must be decomposed into a mesh of simple geometrical elements (e.g. triangles or quads). Therefore MoShade uses mesh input files, where the shape is subdivided into small flat triangles (*.stl* files). The difficulty here is to compute the merged projected area for each triangle 'at once', i.e. by taking care of the shadowing of some triangles by others. MoShade therefore eliminates triangles that are not oriented adequately (see Supporting Information 2), and computes the intersection between the projection of each triangle and the current projected area for a given shape orientation (θ_i, ϕ_i), and adds only the 'not already covered' area, if it exists.

However, directly using the quadrature in Eq. S3 yields numerical issues. Indeed, the result is inaccurate when integrating a constant cross-section coming from a spherical shape for instance. This bad behavior originates from the term in $\cos\theta$. One could expect an exact result even with only one sample point in this case, as for any reasonably good quadrature such as Gaussian integration. To reach this goal (i.e. an exact result), the integration is optimized using a change of variables (see Supporting Information 2 for more details). The term $\cos\theta$ is cancelled in the integrand so that $\cos\theta \, d\theta$ is seen as constant. By changing $\theta = \sin^{-1}\psi$, and by substituting the boundaries of the integral in Eq. S2 to their sinus, one obtains Eq. S4, which constitutes the optimized integration used by MoShade. Once again, the *n* samples parameterizing a MoShade calculation lead to $N=(2n)^2$.

$$\overline{S}_p = \frac{1}{\pi} \int_0^\pi \int_0^1 S_p\left(\sin^{-1}\psi, \phi\right) d\psi \, d\phi \approx \frac{1}{N\pi} \sum_i S_p\left(\sin^{-1}\psi, \phi_i\right) \tag{S4}$$

It should be noted that the change of variable works because the bounds of the integral are such that the application $\theta = \sin^{-1}\psi$ is monotonous in the interval.

Results and Discussions

a

b

c

d

Figure S3: Plots of the Collision Cross-Section (CCS) calculated using MoShade as a function of the shape volume. The fit functions of the CCS evolutions (arbitrary units) are either power functions or linear functions. The structures accompanying each plot represent

the smallest considered structure, an intermediate size structure (except for S2.d) and the largest structure of each series. The plots depict: S2.a a sphere with a growing diameter; S2.b a cylinder with an increasing cylinder length; S2.c a cylinder with an increasing cylinder diameter and S2.d an extract of S2.c at small CCS an volume values with a more appropriate fit weighted at smaller values (represented in red). Table 1 sums up the shape variations with the minimum and maximum dimensions, as well as the number of samples and the fit parameter values.
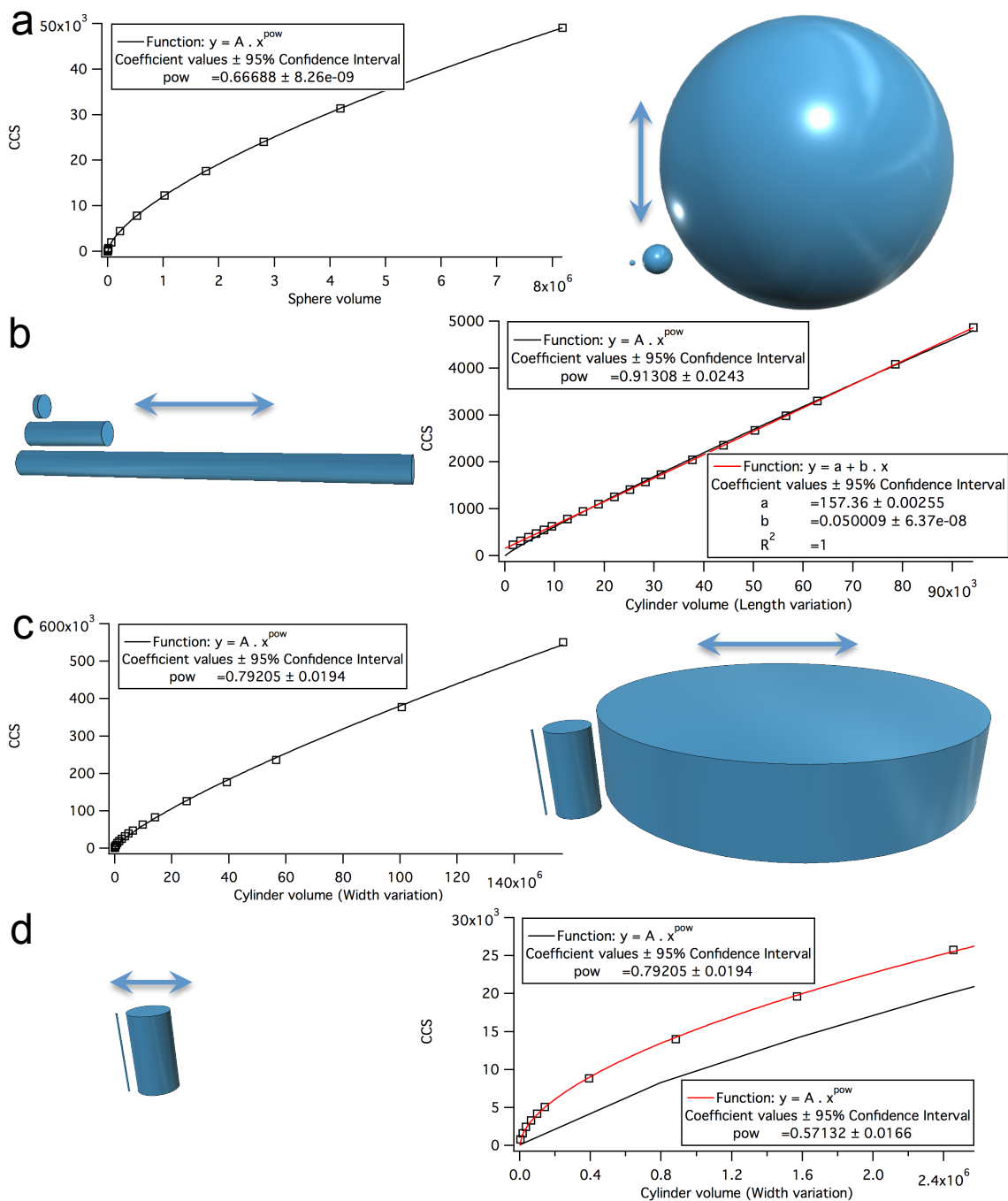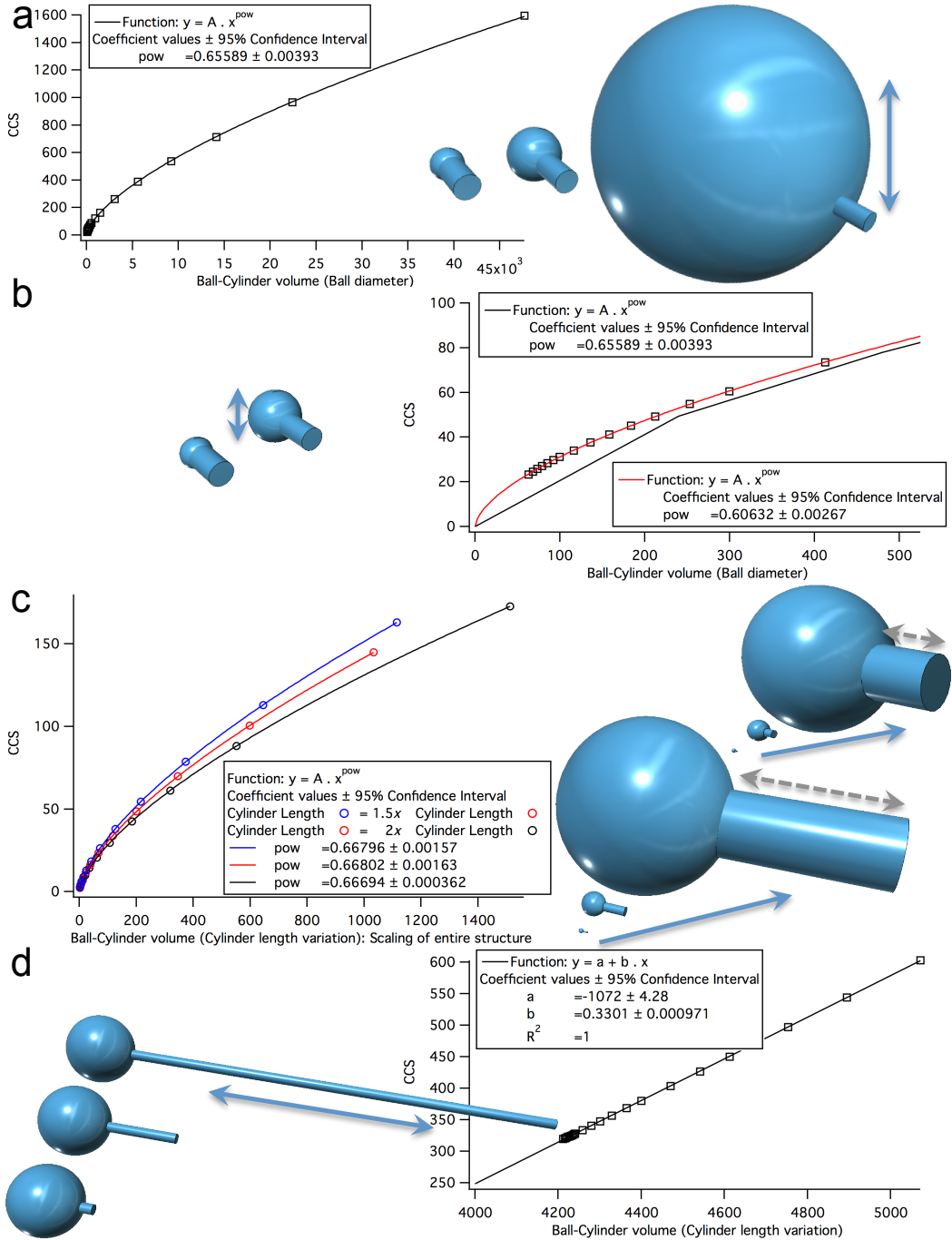
Figure S4: Plots of the Collision Cross-Section (CCS) calculated using MoShade as a function of the shape volume. The fit functions of the CCS evolutions (arbitrary units) are either power functions or linear functions. The structures accompanying each plot represent the smallest considered structure, an intermediate size structure (except for S3.b) and the

largest structure of each series. The plots depict composite ball-cylinder shapes: S3.a ball-cylinder with a growing ball diameter; S3.b an extract of S3.a at small CCS an volume values with a more appropriate fit weighted at smaller values (represented in red); S3.c ball-cylinders with differing cylinder lengths which undergo a uniform isotropic shape increase/scaling; S3.d a ball-cylinder with an increasing cylinder length. Table 1 sums up the shape variations with the minimum and maximum dimensions, as well as the number of samples and the fit parameter values.
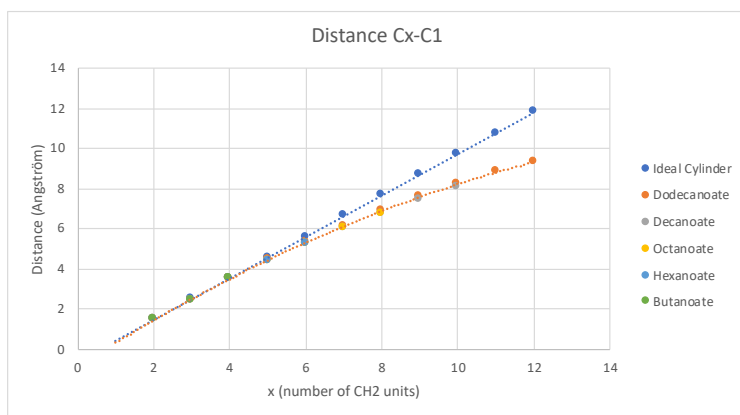
Figure S5: Distance evolutions by -CH_2- unit increments for the case of an ideal cylinder (in blue; linear evolution) and for MD-modeled carboxylate ions (without $FeCl_3$ core; median distance calculated from 4000 structures extracted over 100 nanoseconds at 500 K). The distance is calculated as the difference between an xth CH_2 unit in the chain and the first C from the carboxylate group "C_x-C_1".

Table S1: Structure pools generated by MM2 and MMFF94 force fields implemented in Chem3D Pro v.11.0 with and without structure optimization. a. $[\text{PEO}_{DP=14} + 1\text{Na}^+]^{1+}$ and b. $[\text{PEO}_{DP=14} + 2\text{Na}^+]^{2+}$ structure candidates (represented in 2 different viewing angles). The structures pool is obtained by sampling different simulation temperatures. The last column of the table gives the value of the S/CCS ratio (S as the geometric surface area) of each structure, as calculated by MoShade (arbitrary units). The bold values represent the structures providing the best agreement with Eq. 9 ($S = 4.CCS$). These structures can be correlated to optimized structure geometries.
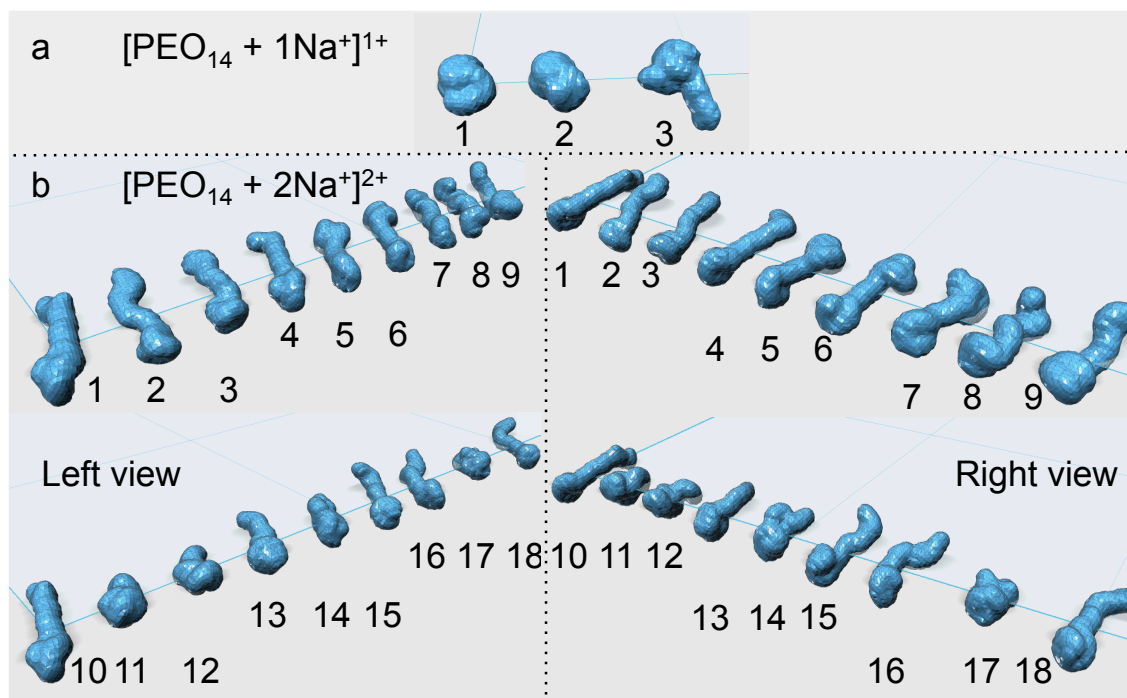
Complex description	Representation	Ratio S/CCS
	a 1	**4.07**
$[PEO_{14} + 1Na^+]^{1+}$	**a 2**	**4.07**
	a 3	4.13
	b 1	**4.09**
	b 2	4.14
	b 3	4.15
	b 4	**4.09**
	b 5	4.14
	b 6	4.16
	b 7	4.14
	b 8	4.27
	b 9	4.17
$[PEO_{14} + 2Na^+]^{2+}$	**b 10**	**4.08**
	b 11	4.10
	b 12	4.17
	b 13	4.12
	b 14	4.13
	b 15	4.14
	b 16	4.19
	b 17	4.15
	b 18	4.13

Table S2: Slopes of the linear fits (*b*) of the CCS plotted as a function of the volume of all the modeled beads on a string shapes (see Figure 4.c in the main article). The dimensions of the bead-string unit are given as descriptors. BD represents the ball diameter, CD represents the cylinder diameter and CL represents the cylinder length. The ratios leading to the best agreement with the experimental ratio between PEtP and PEO (see main text) are written in bold.

Description BD – CD – CL			8 – 3 – 6	8 – 3 – 12	8 – 2.12 – 12	20 – 3 – 3	20 – 3 – 6	20 – 3 – 12	20 – 6 – 3	20 – 6 – 6	20 – 6 – 12
	Slope b	b_2	0.190	0.209	0.216	0.068	0.071	0.075	0.068	0.071	0.076
	b_1	b_1/b_2									
8 – 3 – 6	0.190		1	0.910	0.880	2.782	2.686	**2.531**	2.789	2.670	**2.512**
8 – 3 – 12	0.209		1.099	1	0.967	3.058	2.951	2.782	3.065	2.934	2.761
8 – 2.12 – 12	0.216		1.136	1.034	1	3.161	3.051	2.876	3.169	3.033	2.854
20 – 3 – 3	0.068		0.359	0.327	0.316	1	0.965	0.910	1.002	0.960	0.903
20 – 3 – 6	0.071		0.372	0.339	0.328	1.036	1	0.942	1.038	0.994	0.935
20 – 3 – 12	0.075		0.395	0.360	0.348	1.099	1.061	1	1.102	1.055	0.992
20 – 6 – 3	0.068		0.359	0.326	0.316	0.998	0.963	0.908	1	0.957	0.901
20 – 6 – 6	0.071		0.375	0.341	0.330	1.042	1.006	0.948	1.045	1	0.941
20 – 6 – 12	0.076		0.398	0.362	0.350	1.108	1.069	1.008	1.110	1.063	1

Table S3: Composite Ball-Cylinder shapes constituted by 3 balls and 2 cylinders. The shape variation is made up of angle variations as shown in the Figure. The corresponding CCS values, calculated using MoShade, and their percentage variations are shown (calculated taking the shape with a 0° angle as reference).
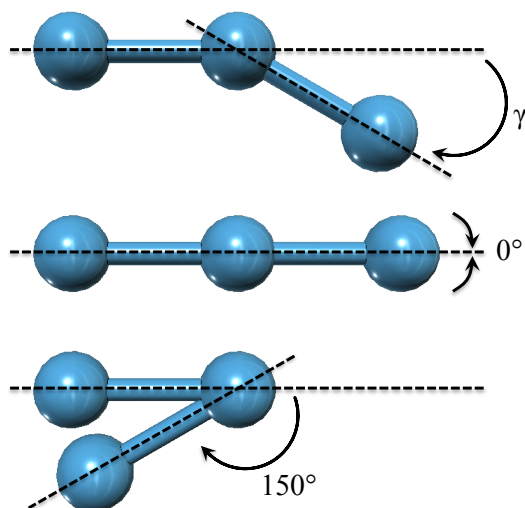
Angle γ	CCS (arbitrary units)	CCS difference in % (reference taken at 0°)
0°	1092	0.0
15°	1095	-0.2
30°	1092	0.1
45°	1087	0.5
60°	1089	0.3
75°	1087	0.5
90°	1086	0.5
105°	1081	1.0
120°	1073	1.8
135°	1061	2.9
150°	1031	5.7

References

(1)   Clément, B.; Grignard, B.; Koole, L.; Jérôme, C.; Lecomte, P. Metal-Free Strategies for the Synthesis of Functional and Well- De Fi Ned Polyphosphoesters. *Macromolecules* **2012**, *45*, 4476–4486. https://doi.org/10.1021/ma3004339.

(2)   Haler, J. R. N.; Morsa, D.; Lecomte, P.; Jérôme, C.; Far, J.; De Pauw, E. Predicting Ion Mobility-Mass Spectrometry Trends of Polymers Using the Concept of Apparent Densities. *Methods* **2018**. https://doi.org/10.1016/j.ymeth.2018.03.010.

(3)   Ruotolo, B. T.; Benesch, J. L. P. P.; Sandercock, A. M.; Hyung, S.-J. J.; Robinson, C. V. Ion Mobility-Mass Spectrometry Analysis of Large Protein Complexes. *Nat. Protoc.* **2008**, *3* (7), 1139–1152. https://doi.org/10.1038/nprot.2008.78.

(4)   Counterman, A. E.; Valentine, S. J.; Srebalus, C. A.; Henderson, S. C.; Hoaglund, C. S.; Clemmer, D. E. High-Order Structure and Dissociation of Gaseous Peptide Aggregates That Are Hidden in Mass Spectra. *J. Am. Soc. Mass Spectrom.* **1998**, *9* (8), 743–759. https://doi.org/10.1016/S1044-0305(98)00052-X.

(5)   Bush, M. F.; Campuzano, I. D. G. G.; Robinson, C. V. Ion Mobility Mass Spectrometry of Peptide Ions: Effects of Drift Gas and Calibration Strategies. *Anal. Chem.* **2012**, *84* (16), 7124–7130. https://doi.org/10.1021/ac3014498.

(6)   Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. Conformer-Dependent Proton-Transfer Reactions of Ubiquitin Ions. *J. Am. Soc. Mass Spectrom.* **1997**, *8* (9), 954–961. https://doi.org/10.1016/S1044-0305(97)00085-8.

(7)   Shelimov, K. B.; Jarrold, M. F. Vacuum : An Activation Barrier for Gas-Phase Protein Folding. *J. Am. Chem. Soc.* **1997**, *119* (13), 2987–2994.

(8)   Chen, Y. L.; Collings, B. A.; Douglas, D. J. Collision Cross Sections of Myoglobin

and Cytochrome c Ions with Ne, Ar, and Kr. *J. Am. Soc. Mass Spectrom.* **1997**, *8* (7), 681–687. https://doi.org/10.1016/S1044-0305(97)00033-0.

(9) Valentine, S. J.; Anderson, J. G.; Ellington, A. D.; Clemmer, D. E. Disulfide-Intact and -Reduced Lysozyme in the Gas Phase: Conformations and Pathways of Folding and Unfolding. *J. Phys. Chem. B* **1997**, *101* (19), 3891–3900. https://doi.org/10.1021/jp970217o.

(10) Forsythe, J. G.; Petrov, A. S.; Walker, C. A.; Allen, S. J.; Pellissier, J. S.; Bush, M. F.; Hud, N. V.; Fernández, F. M. Collision Cross Section Calibrants for Negative Ion Mode Traveling Wave Ion Mobility-Mass Spectrometry. *Analyst* **2015**, *140* (20), 6853–6861. https://doi.org/10.1039/c5an00946d.

(11) Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan, C. J.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinos, K.; Valentine, S. J.; Wyttenbach, T.; T. Bowers, M.; Bilbao, A.; F. Bush, M.; Campbell, J. L.; Campuzano, I. D. G.; J. Causon, T.; H. Clowers, B.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Wyttenbach, T. Recommendations for Reporting Ion Mobility Mass Spectrometry Measurements. *Mass Spectrom. Rev.* **2019**, *38* (3), 291–320. https://doi.org/10.1002/mas.21585.

Geometric Analysis of Shapes in Ion Mobility-Mass Spectrometry

Jean R. N. Haler[1,2,†], Eric Béchet[3,†], Christopher Kune[1], Johann Far[1], Edwin De Pauw[1]*

[1]Mass Spectrometry Laboratory, University of Liège, MolSys Research unit, Quartier Agora, Allée du Six Aout 11, B-4000 Liège, Belgium

[2]Luxembourg Institute of Science and Technology – LIST, Materials Research & Technology MRT Department, L-4422 Belvaux, Luxembourg

[3]Aerospace & Mechanical Engineering Department, Computer-aided Geometric Design, University of Liège, B-4000 Liège, Belgium

[†]J.R.N. Haler and E. Béchet equally contributed to this work

*Corresponding author email address: jean.haler@uliege.be

Supporting Information 2

Computation of an average cross-section of 3D shapes

The goal of this document is to set up formulae used to compute the average cross section $\overline{S_p}$ of various shapes. The "actual" cross section $S_p(\dots)$ is measured as the area of the projected shade onto the Oxy plane, when light comes from the z direction (from above), and obviously depends on the orientation of the shape with respect to the Oxy plane. The orientation depends on parameters (e.g. $\theta, \phi, \dots$). There are symmetries, though. For instance, the actual cross section does not change when one revolves the solid around the Oz direction, this helps in establishing a simplified numerical quadrature in cases where one cannot resort to explicit formulae. In what follows, we go from simple cases to the general case for concave solids. For convex solids, an explicit analytical result holds as shown below.

1 For a revolving solid

The choice of the orientation is generally made with only two degrees of freedom (as one rotation around the axis of symmetry of the solid does not change anything). In addition, as stated before, a revolution around a normal (Oz) of the projection plane does not change anything either. Remains only one degree of freedom. What is the probability to find the shape in one orientation or the other? It depends on the solid angle covered by an infinitesimal area on the unit sphere (which is the locus of tip the orientation vector). This in turn depends on the parametric representation used here. One may use θ as the elevation and ϕ as the azimuth:

$$p(\theta, \phi) = \begin{cases} cos\phi cos\theta \\ sin\phi cos\theta \\ sin\theta \end{cases} \tag{S0}$$

so that the infinitesimal surface related to a small variation in θ and ϕ is $d^2 s = cos\theta d\phi d\theta$.

As stated above, there is no change in the projection for a revolution around Oz , therefore one can integrate for ϕ going from 0 to 2π to get the infinitesimal area for a small variation in θ only : we get $ds = 2\pi cos\theta d\theta$. This is the "weight" or probability that the orientation of the solid is θ, for any value of ϕ. One should therefore integrate all the projected surfaces and divide the result by the integral of the probability density to get the average area. It yields (using symmetries to reduce the integration bounds to the minimal possible interval, that is, consider that a projection from above when θ is positive leads to the same contribution than a projection from below with a negative θ):

$$\bar{S}_p = \frac{\int_{-\pi/2}^{\pi/2} S_p(\theta) \cdot 2\pi cos\theta d\theta}{\int_{-\pi/2}^{\pi/2} 2\pi cos\theta d\theta} = \frac{1}{2}\int_{-\pi/2}^{\pi/2} S_p(\theta) \cdot cos\theta d\theta = \int_0^{\pi/2} S_p(\theta) \cdot cos\theta d\theta \tag{S1}$$

We now try to establish the average projection for two trivial cases.

1.1 Case of a sphere

In this case, the projected area (in absolute value) is constant and is given by $S_p(\theta) = \pi\frac{D^2}{4}$.

The average cross section is therefore given by:

$$\bar{S}_p = \int_0^{\pi/2} \pi\frac{D^2}{4} \cdot cos\theta \, d\theta = \pi\frac{D^2}{4} \tag{S2}$$

In addition; the volume is given by $V = \pi\frac{D^3}{6}$, and its area is $S = \pi D^2$. Therefore, the ratio between the area of the sphere and the average cross section is 4.

What is of interest here is the behaviour of the mean projected surface with respect to the molecular mass m. In our case, $V \propto m$ so $D \propto m^{1/3}$, and therefore $\bar{S}_p \propto m^{2/3}$.

1.2 Case of a cylinder with spherical ends

Here, the projected surface depends on the angle θ. For the spherical parts, same as previous paragraph because the two half spheres are present at the extremities, and are not "more" shadowed by the cylindrical part than if those were part of a simple sphere. For the cylindrical part, one should take a term in $cos\theta$ into account because of the slope with respect to the projection plane:

$$S_p(\theta) = \pi\frac{D^2}{4} + DL\cos\theta \qquad\qquad (S3)$$

The volume is equal to $V = \pi\frac{D^3}{6} + \pi\frac{D^2}{4}L$ in this case.

The area of the exterior surface is $S = \pi(D^2 + DL)$.

We have therefore:

$$\overline{S_p} = \int_0^{\pi/2}(\pi\frac{D^2}{4} + DL\cos\theta)\cdot\cos\theta d\theta = \pi\frac{D^2}{4} + DL\int_0^{\pi/2}\frac{1+\cos2\theta}{2}d\theta = \frac{\pi}{4}(D^2 + DL) \qquad (S4)$$

In this case, $V \propto m$ still holds, so $D \propto L \propto m^{1/3}$, and therefore $\bar{S}_p \propto m^{2/3}$ if one considers that the increase of size is isotropic. Nothing really changes here with respect to the simpler spherical case, and **it is so for every convex shape**. However, if the cylinder increases in length without increase in the diameter (that means the aspect ration changes), D is a constant and $L \propto m$, therefore one finds that $\bar{S}_p \propto m$!

The average cross section is in fact proportional to the molecular mass in this case.

If only the diameter increases, the behaviour is more complex: L is constant and $D \propto m^{1/2}$ therefore $\bar{S}_p \propto m^{1/2}$ if the diameter D is small with respect to L, however $D \propto m^{1/3}$ and therefore $\bar{S}_p \propto m^{2/3}$ if L is small with respect to D – a lesser common case.

1.3 Case of a hollow sphere

If the increase in size is isotropic, it is the same procedure than in 1.1; only the volume V changes. However, let us imagine that the thickness of the sphere is constant, so the volume becomes proportional to its surface and the thickness e (considered very small).

$$\bar{S}_p = \pi\frac{D^2}{4} \qquad\qquad (S5)$$

, with $V = \pi D^2 e$, so that $D \propto m^{1/2}$ and therefore $\bar{S}_p \propto m$, which is the same behaviour as a cylinder of constant diameter that increases its length. It is therefore impossible to discriminate both cases just by comparing how the cross sections evolve with an increase in molecular mass.

2   General case
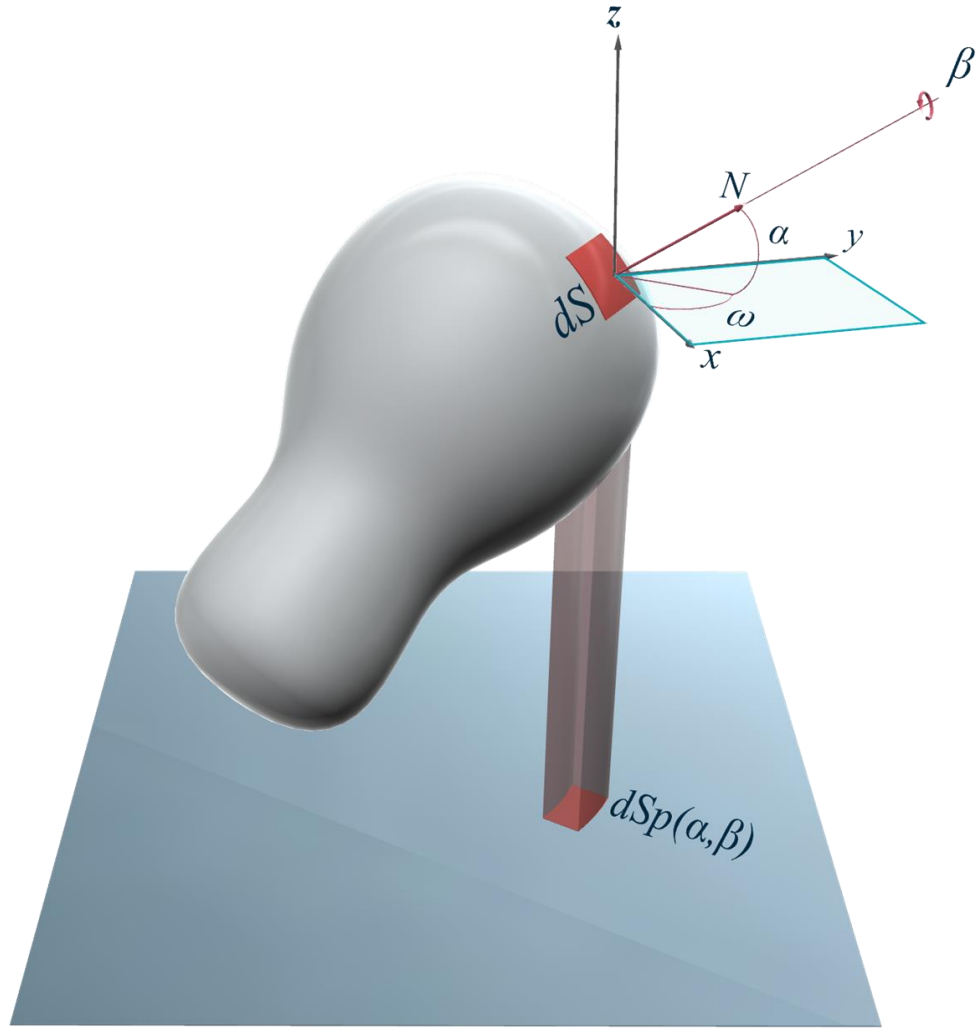
As a matter of fact, the case of a convex shape, therefore including the cases of the sphere and cylinder with hemispherical ends, has been explored in the past by Cauchy [1,2] and leads to the following remarkable identity:

$$\bar{S}_p = \frac{S}{4}$$

This allows to compute the average cross section at once, knowing the area of the solid, without having to compute any of the actual cross sections $S_p(\theta,\dots)$. However, this is of no help when the shape is concave (non-convex).

Let us consider an infinitesimally small patch δS located on the surface of the solid. Now, depending on the orientation (α, β) of this patch, its projected area will is noted $\delta S_p(\alpha, \beta)$. One should note that (α, β) is an absolute orientation, i.e. when $(\alpha, \beta) = (0,0)$ the normal to the patch is horizontal, and its projected area is simply zero. Now the shape to which the patch belongs still has its orientation parametrized globally by (θ, ϕ), and in this setting, the patch has an additional shift (a constant local orientation with respect to the shape reference frame) given by (ξ, η), so that in fact $(\alpha, \beta) = (\theta + \xi, \phi + \eta)$. If the expression $\delta S_p(\alpha, \beta)$ is known, then we shall compute the average projected area. The choice of orientations here has three degrees of freedom. Again, as with the first example above, a rotation around a normal to the projection plane Oz does not change anything. There remains two degrees of freedom: one angle α with respect to an axis in the projection plane, which defines an out-of-plane axis, and a rotation around the normal of the patch β. What is the probability to find the shape in one orientation or the other? It is the portion of solid angle covered by infinitesimals on the unit sphere, as before: $2\pi\cos\alpha d\alpha$. Obviously, the projected area does not depend on β, but we will keep the variable in the integration procedure in this case. We should therefore integrate all the projected surfaces and divide by the integral of the probability density to get the average cross section of the patch δS. It yields the following:

$$\delta\bar{S}_p = \frac{\int_{-\pi}^{\pi}\int_{-\pi/2}^{\pi/2}\delta S_p(\alpha,\beta)\cdot 2\pi\cos\alpha d\alpha d\beta}{\int_{-\pi}^{\pi}\int_{-\pi/2}^{\pi/2}2\pi\cos\alpha d\alpha d\beta} = \frac{1}{4\pi}\int_{-\pi}^{\pi}\int_{-\pi/2}^{\pi/2}\delta S_p(\alpha,\beta)\cdot\cos\alpha d\alpha d\beta \qquad (S6)$$

In fact, this patch will be shadowed at some point by another part of the shape. So, what is the expression of $\delta S_p(\alpha, \beta)$? This expression may be expanded as $\delta S_p(\alpha, \beta) = \delta S_{tot}(\alpha, \beta) \cdot Vis(\alpha, \beta)$, where $Vis(\alpha, \beta)$ is the visibility of the patch with respect to the projection plane. $Vis(\alpha, \beta)$ is either equal to 1 if the patch is visible, or 0 otherwise. There are no intermediate values because the patch is infinitely small and therefore considered absolutely flat. The expression $\delta S_{tot}(\alpha, \beta)$ refers to the projected area of the patch: $\delta S_{tot}(\alpha, \beta) = \delta S \sin\alpha$.

In the case of a convex shape, the expression simplifies because $Vis(\theta, \phi) = 1$ if $0 \leq \alpha \leq \pi/2$, and $Vis(\alpha, \beta) = 0$ for $\alpha < 0$. Therefore, the shading does actually occur on exactly one hemisphere, and the following simplification of Eq. S6 occurs:

$$\delta \bar{S}_p = \frac{1}{4\pi} \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} Vis(\alpha, \beta) \delta S \cdot \sin\alpha \cdot \cos\alpha d\alpha d\beta = \frac{1}{4\pi} \int_{-\pi}^{\pi} \int_{0}^{\pi/2} \delta S \cdot \sin\alpha \cdot \cos\alpha d\alpha d\beta = \frac{\delta S}{4} \quad (S7)$$

At the end, we have:

$$\delta \bar{S}_p = \frac{\delta S}{4} \quad (S8)$$

Now; if one takes the integral over the whole surface of the shape to get the total projected area, one gets

$$\bar{S}_p = \int_S \delta \bar{S}_p \, dS = \frac{1}{4\pi} \int_{-\pi}^{\pi} \int_{0}^{\pi/2} \int_S \delta S \cdot \sin\alpha \cdot \cos\alpha dS \, d\alpha d\beta = \frac{1}{4} \int_S \delta S \, dS = \frac{S}{4} \quad (S9)$$

, thus the remarkable identity pointed out by Cauchy [1,2]:

$$\overline{S_p} = \frac{S}{4} \quad (S10)$$

Here, S is the area of the exterior surface of the shape. This holds only because the shadowing is predictable: exactly one half of the orientation hemisphere is shadowed for every small patch located on the shape, therefore the contribution of each one is constant and equal to one fourth its area.

It should be noted that the global orientation of the shape (θ, ϕ) is not used here. Indeed, the expressions above holds for every patch on the surface, whatever its local orientation in the shape's frame. Orienting a patch using $(\alpha, \beta) = (\theta + \xi, \phi + \eta)$ or $(\theta, \phi) = (\alpha - \xi, \beta - \eta)$ leads to identical results since the integrals are written on the whole hemisphere (therefore a shift in the angles does not change anything), and the visibility in that case depends only on the absolute orientation of the patch, i.e. (α, β).

Now, what to do if the shape is not convex? The shadowing is not easily predictable anymore, and depends on the actual shape and its orientation. There is no way to get a simple expression of the visibility factor as in the previous case. In fact, one has to resort to numerical approximations to compute, in a global way, the average $\bar{S}_p$. The difficulty here is to compute $S_p(\theta, \phi)$ for any given orientation (θ, ϕ) of the shape.

$Sp(\theta,\phi)$
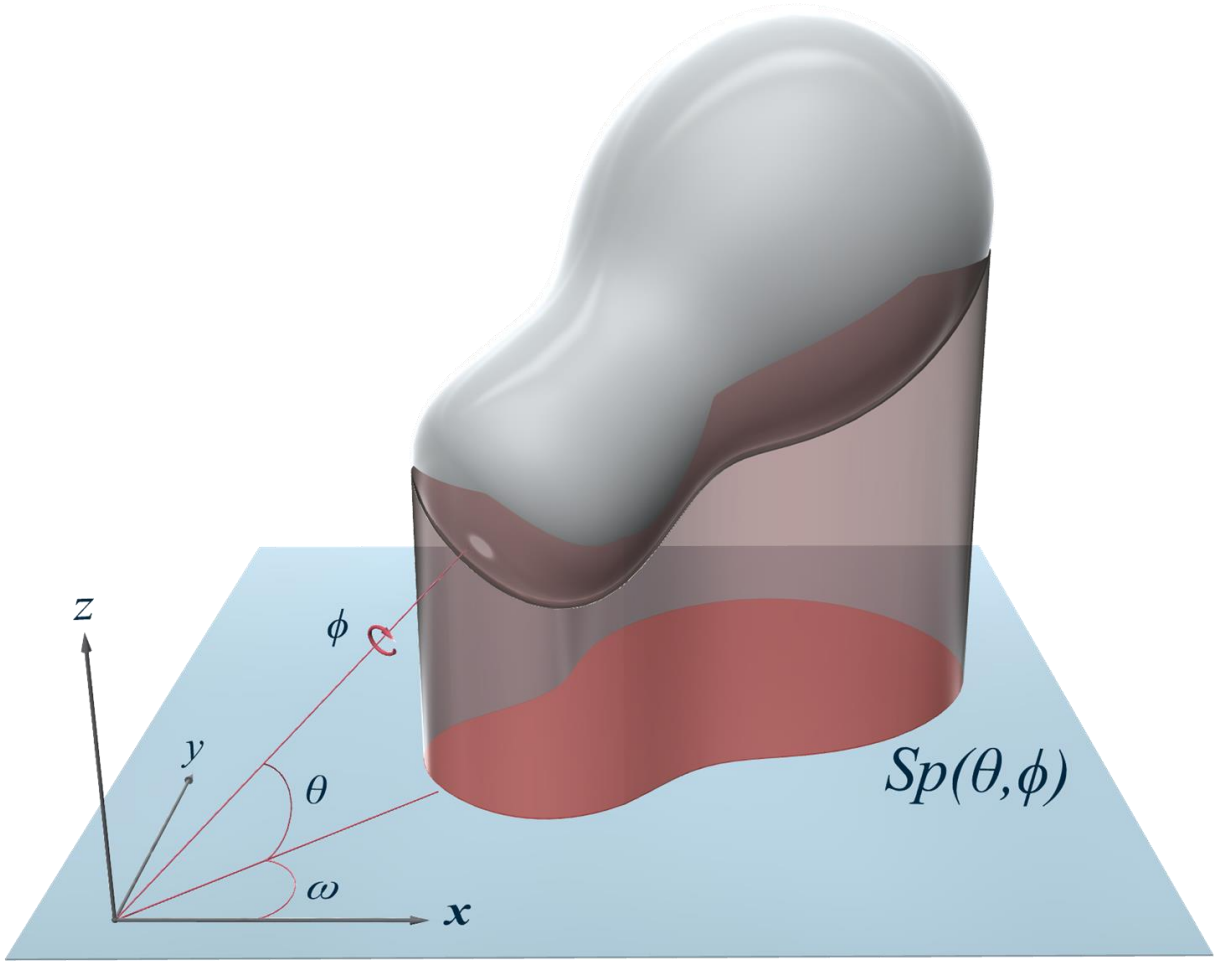
We are now dealing with the global orientation of the shape instead of individual patches, therefore the absolute orientation of a patch $(\alpha, \beta) = (\theta + \xi, \phi + \eta)$ will not be used here to avoid confusion.

$$\overline{S_p} = \frac{\int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} S_p(\theta,\phi) \cdot 2\pi cos\theta d\theta d\phi}{\int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} 2\pi cos\theta d\theta d\phi}$$

$$= \frac{1}{4\pi} \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} S_p(\theta, \phi) \cdot cos\theta d\theta d\phi$$

(S11)

If one wishes to use numerical integration instead of trying to integrate analytically as in the previous case, one may use the following simple quadrature that helps to approximate an integral as a discrete sum of evaluations of the integrand (e.g. the trapezoid rule):

$$\overline{S_p} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} S_p(\theta, \phi) \cdot cos\theta d\theta d\phi \approx \frac{1}{4\pi} \sum_i \frac{\pi^2}{2N} S_p(\theta_i, \phi_i) \cdot cos\theta_i$$

(S12)

, with an uniform distribution of the $4N$ samples on $\theta_i, \phi_i \in [-\frac{\pi}{2} \dots \frac{\pi}{2}] \times [-\pi \dots \pi]$.

In turn, to evaluate an approximation of $S_p(\theta_i, \phi_i)$, the surface of the shape must be decomposed into a mesh of simple geometrical elements (e.g. triangles or quads). This decomposition is usually not difficult to obtain: most 3D modelling software include an option to output an STL file that is perfectly suited to this purpose. The difficulty here is to compute the merged projected area for each triangle "at once", i.e. by taking care of the shadowing of some triangles by others. One should therefore eliminate triangles that are not oriented adequately (using e.g. back-face culling), and compute the intersection between the projection of each triangle and the current projected area, and add only the "not already covered" area, if it exists. It seems a simple task, but most of the code of MoShade lies there, although no theoretical complexity hides in that procedure. The implementation is however quite complex because of the need to keep track of already shadowed portions while computing the total area of the shadow. It is outside the scope and length of this document to explain it in a more detailed fashion.

3 Optimization of the numerical integration

3.1 Symmetries

Because of the symmetrical nature of the problem (a shadow "from below" has the same area as a shadow "from above" for any given shape as we consider "unsigned" or non-oriented measures of surfaces) , the following identity may be used to avoid computing twice the same terms in the discrete integral, by reducing the range of the variable θ to include only positive values :

$$\frac{1}{4\pi}\int_{-\pi}^{\pi}\int_{-\pi/2}^{\pi/2} S_p(\theta,\phi)\cdot\cos\theta d\theta d\phi = \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{0}^{\pi/2} S_p(\theta,\phi)\cdot\cos\theta d\theta d\phi \tag{S13}$$

which in turn leads to

$$\overline{S_p} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{0}^{\pi/2} S_p(\theta,\phi)\cdot\cos\theta d\theta d\phi \approx \frac{1}{2\pi}\sum_i \frac{\pi^2}{2N}S_p(\theta_i,\phi_i)\cdot\cos\theta_i \tag{S14}$$

, with an uniform distribution of the (now) $2N$ samples on $\theta_i,\phi_i \in [0\dots\frac{\pi}{2}]\times[-\pi\dots\pi]$.

There is another symmetry for the variable ϕ: the projection is same when the part is rotated by π radians around this axis. The expression can therefore be modified further by replacing the first integral along a full circle to only span π radians (either $-\pi/2$ to $\pi/2$ or 0 to π - it is strictly equivalent):

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{0}^{\pi/2} S_p(\theta,\phi)\cdot\cos\theta d\theta d\phi = \frac{1}{\pi}\int_{0}^{\pi}\int_{0}^{\pi/2} S_p(\theta,\phi)\cdot\cos\theta d\theta d\phi \tag{S15}$$

which then leads to

$$\overline{S_p} = \frac{1}{\pi}\int_{0}^{\pi}\int_{0}^{\pi/2} S_p(\theta,\phi)\cdot\cos\theta d\theta d\phi \approx \frac{1}{\pi}\sum_i \frac{\pi^2}{2N}S_p(\theta_i,\phi_i)\cdot\cos\theta_i \tag{S16}$$

, with an uniform distribution of the (now) N samples on $\theta_i,\phi_i \in [0\dots\frac{\pi}{2}]\times[0\dots\pi]$.

3.2 Harmonic term in the integrals

Now, the integrals used to compute the average cross section always contain a term in $\cos\theta d\theta$ as in the following expression:

$$\int_{a}^{b} F(\theta)\cdot\cos\theta d\theta \approx \sum_i \frac{b-a}{n}F(\theta_i)\cdot\cos\theta_i \tag{S17}$$

, with $0 \leq a < b \leq \pi/2$, and n is the number of samples used in the quadrature.

This term is such that when one uses a discrete integration quadrature to compute the average cross section for a sphere, the result is in fact depending on the number of samples, whereas for each sample, the cross section $F(\theta_i)$ is constant. This means that the integration quadrature is not even able to integrate a constant term exactly, which is not numerically sound.

In order to overcome this, and have an exact result for obvious cases (similarly to a Gauss integration of a constant function giving obviously the exact value for any number of samples), one has to use a different integration quadrature via a change of variable so that the integrand is constant when the cross section is constant. In fact, this kind of change of integration quadrature leads to better convergence results in the general case as well, allowing to use less samples for a given accuracy, because the samples will be uniformly distributed in the "physical" space instead of the parametric space associated with the orientation angle.

To achieve this, the term $\cos\theta$ must be cancelled in the integrand so that $\cos\theta d\theta$ is seen as a constant. Let us introduce the variable $\psi = \sin\theta$. Then, differentiating yields $d\psi = \cos\theta d\theta$ which does the trick. Therefore, one should use the following change of variable:$\theta = \sin^{-1}\psi$, and substitute the boundaries of the integral from a, b to $\sin a, \sin b$. Therefore, the integral from Eq. S17 becomes

$$\int_{a}^{b} F(\theta)\cdot\cos\theta d\theta = \int_{\sin a}^{\sin b} F(\sin^{-1}\psi)d\psi \approx \frac{\sin b - \sin a}{n}\sum_i F(\sin^{-1}\psi_i) \tag{S18}$$

, and this time, the uniform distribution of samples ψ_i is made on the variable ψ , over the modified interval [sin a,sin b], instead of θ.

In our case, we want to integrate the following expression (Eq. S16):

$$\overline{S_p} = \frac{1}{\pi}\int_{0}^{\pi}\int_{0}^{\pi/2} S_p(\theta,\phi)\cdot\cos\theta d\theta d\phi,$$ which can readily be transformed using the same change of variable:

$$\overline{S_p} = \frac{1}{\pi}\int_{0}^{\pi}\int_{0}^{1} S_p(\sin^{-1}\psi,\phi)\cdot d\psi d\phi \approx \frac{1}{\pi}\sum_i \frac{\pi}{N}S_p(\sin^{-1}\psi_i,\phi_i) \tag{S19}$$

Note that the change of variable works because the bounds of the integral are such that the application $\theta = \sin^{-1}\psi$is monotonous in that interval.

4 References

[1] A. Cauchy, Mémoire sur la rectification des courbes et la quadrature des surfaces courbes, Mem. Acad. Sci. Paris 22, 3ff, 1850 [Œuvres complètes (1), Vol 2, Gauthiers-Villars, Paris, 1908, pp. 167-177]

[2] J.R. Sangwine-Yager, Chapter 12 "Mixed Volumes", Handbook of Convex Geometry, P.M Gruber, J.M. Wills

Ed., North Holland, 1994.