

Extracting the electronic structure signal from x-ray and electron scattering in the gas phase

THOMAS NORTHEY,^{a*} ADAM KIRRANDER^b AND PETER M. WEBER^a

^a*Department of Chemistry, Brown University, Providence, RI 02912, United States,*

and ^bPhysical and Theoretical Chemistry Laboratory, Department of Chemistry,

University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom.

E-mail: peter-weber@brown.edu

Abstract

We examine x-ray and electron scattering from free gas-phase molecules using the independent atom model (IAM) and *ab initio* electronic structure calculations. The IAM describes the effect of the molecular geometry on the scattering, but does not account for the redistribution of valence electrons due to, for instance, chemical bonding. By examining the total, *i.e.* energy-integrated, scattering from three molecules, fluoroform (CHF₃), 1,3-cyclohexadiene (C₆H₈), and naphthalene (C₁₀H₈), we find that the effect of electron redistribution predominantly resides at small-to-medium values of the momentum transfer ($q \leq 8 \text{ \AA}^{-1}$) in the scattering signal, with a maximum percent difference contribution at $2 \leq q \leq 3 \text{ \AA}^{-1}$. We demonstrate a procedure to determine the molecular geometry from the large- q scattering, making it possible to more clearly identify the deviation of the scattering from the IAM approximation at small and intermediate q and to provide a measure of the effect of valence electronic structure on the scattering signal.

1. Introduction

1.1. Background on gas-phase x-ray and electron scattering

Scattering has provided an indispensable tool in advancing our understanding of the structure of matter (von Laue, 1915; Bragg & Bragg, 1915; Watson & Crick, 1953; Perutz *et al.*, 1960). Gas-phase scattering from molecules was a key component in early advances (Debye, 1915; Debye *et al.*, 1929; Debye, 1930; Pirenne, 1939; Pirenne, 1946) and the invention of x-ray free-electron lasers (XFELs) has sparked a renewed interest in gas-phase x-ray scattering (Küpper *et al.*, 2014; Minitti *et al.*, 2014), not the least in the domain of ultrafast x-ray scattering (Minitti *et al.*, 2015; Glowia *et al.*, 2016; Kirrander *et al.*, 2016; Stankus *et al.*, 2019; Ruddock *et al.*, 2019; Yong *et al.*, 2020; Yong *et al.*, 2021b; Gabalski *et al.*, 2022; Odate *et al.*, 2023), alongside advances in ultrafast electron diffraction (UED) experimental capabilities (King *et al.*, 2005; Sciaini & Miller, 2011; Weathersby *et al.*, 2015; Zandi *et al.*, 2017). We note that in the context of scattering, x-ray and electron scattering are close analogues (Stefanou *et al.*, 2017; Ma *et al.*, 2020) and in this paper we consider both.

A chemical binding effect has been observed in the electron scattering signal of molecules in the gas phase, observed mainly at small scattering angles (Iijima *et al.*, 1965; Fink & Kessler, 1967; Duguet & Jaegle, 1975; Hirota *et al.*, 1981; An-Ding & Xiao-Lei, 1995). This has been theoretically studied to quantify the effect for many molecules (Bonham & Iijima, 1963; Wang *et al.*, 1994; Hoffmeyer *et al.*, 1998; Shibata *et al.*, 1999; Shibata *et al.*, 2002). We discuss this effect in the total x-ray and electron scattering of gas-phase molecules; further adding quantification of the error that occurs in the structure determination, and proposing a way to separate out the molecular structure part from the electronic structure part in the signal.

1.2. Aim of the paper

The aim of this paper is to establish a procedure to extract the valence electronic structure contribution from the molecular scattering signal in the gas-phase. This contribution is the signature of redistribution of electrons away from spherical atom-centered distributions; predominantly the electrons in valence shells forming molecular orbitals as in chemical bonding. Although scattering is commonly viewed as a method to probe molecular geometry, x-rays scatter from the electrons in the target and thus the scattering relates to the electron density (Ben-Nun *et al.*, 1996; Kirrander, 2012; Suominen & Kirrander, 2014; Northey *et al.*, 2014; Northey *et al.*, 2016; Northey & Kirrander, 2019), and even to the pairwise correlation between electrons (Moreno Carrascosa *et al.*, 2019; Zotev *et al.*, 2020; Moreno Carrascosa *et al.*, 2022). This means that effects such as the redistribution of electrons due to chemical bonding, the delocalisation of electrons in aromatic rings, and the localisation of electrons in valence molecular orbitals show up in the scattering signal. Time-resolved experiments can thus detect the rearrangement of electrons due to photoexcitation (Yong *et al.*, 2020), dynamic charge transfer (Yong *et al.*, 2021c), or the breaking of chemical bonds (Ruddock *et al.*, 2019). Given sufficient time-resolution, it should become possible to track the dynamics of electrons in molecular system (Simmermacher *et al.*, 2019a; Simmermacher *et al.*, 2019b; Ziems *et al.*, 2023). The statements above remain true for electron scattering, however with additional terms due to the scattering of the incoming electrons from the nuclei (see Methods).

For gas-phase scattering, the absence of a crystalline lattice means that energy-integrating detectors pick up the total, rather than the elastic scattering. In this paper, we use the independent atom model (IAM) approximation (Debye *et al.*, 1929; Debye,

1930) of the scattering signal as a baseline that does not include any redistribution of electrons due to bonding, as the IAM assumes a non-interacting spherical distribution of electrons around each individual atom. These results are compared to accurate *ab initio* calculations of the total x-ray scattering that fully account for the redistribution of electrons (Moreno Carrascosa *et al.*, 2019; Zotev *et al.*, 2020; Carrascosa *et al.*, 2022). The difference between the IAM and the *ab initio* signal is identified as the valence electronic structure component. However, it is important to note that the exact molecular geometry is not necessarily known *a priori*. We therefore require a procedure to determine the molecular geometry as accurately as possible before the electronic component can be calculated. We show in this paper that the molecular geometry can be reliably determined from the large values of the momentum transfer q , while small and intermediate q values are most affected by electronic effects. In doing this, we use a recently developed simulated annealing algorithm to fit the molecular geometry to a target x-ray signal (Northey *et al.*, 2023) for various ranges of the momentum transfer vector q .

2. Methods

2.1. X-ray scattering

2.1.1. *Ab initio* calculation The *ab initio* x-ray scattering calculations were carried out using an in-house code from the Kirrander group (Northey *et al.*, 2014; Moreno Carrascosa *et al.*, 2019; Zotev *et al.*, 2020) that calculates the molecular x-ray scattering signal directly from the wavefunction expressed in a Gaussian basis and obtained via *ab initio* electronic structure methods, such as Hartree-Fock (HF) or multiconfigurational wavefunction methods (CASSCF, MRCI, MCCI etc). The code calculates the total x-ray scattering, *i.e.* both the elastic and inelastic components of the signal. In this paper, HF theory with the 6-31G* Pople basis set is used.

2.1.2. Independent atom model According to the independent atom model (IAM), for x-ray scattering from an isotropic ensemble of N_{at} -atomic molecules the total scattering intensity is proportional to,

$$I_{\text{IAM}}(q) = \sum_{i=1}^{N_{\text{at}}} f_i(q)^2 + \sum_{i \neq j}^{N_{\text{at}}} f_i(q) f_j(q) \frac{\sin(qR_{ij})}{qR_{ij}} + \sum_{i=1}^{N_{\text{at}}} S_i(q), \quad (1)$$

where the first sum constitutes the atomic scattering, $I_{\text{at}}(q)$, and runs over all atoms i with tabulated atomic x-ray scattering factors $f_i(q)$ (Prince, 2006). This sum contains no structural information about the molecule; instead, structure is contained in the second molecular scattering term, $I_{\text{mol}}(q)$, a double sum which runs over all pairs of atoms i and j (excluding $i = j$). This term involves the distance between atoms, $R_{ij} = |\mathbf{R}_i - \mathbf{R}_j|$, where $\mathbf{R}_i$ and $\mathbf{R}_j$ are the positions of atoms i and j , respectively. The final term accounts for inelastic scattering and is a sum of tabulated inelastic scattering factors, $S_i(q)$. Finally, the amplitude of the scattering vector is $q = |\mathbf{q}|$, defined as $\mathbf{q} = \mathbf{k}_1 - \mathbf{k}_0$, with $\mathbf{k}_1$ and $\mathbf{k}_0$ the wave vectors of the scattered and incident x-ray photons.

For electron scattering, the Mott-Bethe formula (Mott, 1930) can be used to transform the x-ray atomic factors to electron factors,

$$f_i^{\text{electron}}(s) \propto \frac{Z_i - f_i^{\text{xray}}(s)}{s^2}, \quad (2)$$

with proportionality constant $2m_e c^2 / \hbar$. This means, the IAM electron scattering is very similar to the x-ray scattering equation aside from the $1/s^2$ and Z_i terms, where Z_i is the atomic number of atom i , due to the additional scattering of the electrons by the positive charge of the nuclei. By convention, in electron scattering the scattering vector is denoted as s instead of q .

In this paper, comparison between *ab initio* and IAM scattering is often done in terms of the percent difference (PD) defined as,

$$\text{PD}(q) = 100 \times \frac{I_{\text{IAM}}(q) - I_{\text{abinitio}}(q)}{I_{\text{abinitio}}(q)}. \quad (3)$$

Note that this is a percentage, and that the reference signal $I_{\text{abinitio}}(q)$ is subtracted from $I_{\text{IAM}}(q)$ and also appears in the denominator.

2.2. Fitting procedure

A recently developed simulated annealing (SA) approach is used to fit the predicted IAM signal to the target data. This approach is described in detail in reference (Northey *et al.*, 2023). It minimises the target function,

$$\zeta_{\text{targ}} = \zeta_{\text{signal}} + \zeta_{\text{aux}}, \quad (4)$$

by moving the molecular geometry $\mathbf{R}'$ and recalculating ζ_{targ} iteratively, where $\zeta_{\text{signal}} = (\mathbf{x} - \mathbf{y})^2 / \mathbf{y}$ for predicted x-ray (or electron) IAM scattering signal $\mathbf{x} = \mathbf{x}(\mathbf{R}')$ which depends on $\mathbf{R}'$, and the target signal $\mathbf{y} = \mathbf{y}(\mathbf{R}_{\text{targ}})$ is calculated using the *ab initio* scattering code, where $\mathbf{R}_{\text{targ}}$ is the target geometry. This fitting procedure minimises a squared-difference function iteratively by randomly moving the molecular geometry along its normal modes; accepting an iteration if the fitting improves. It has robustness, in that it can escape local minima via the capacity to take uphill steps with non-zero probability. We note that other fitting methods would likely be adequate to show the results in this paper (Mitzel & Rankin, 2003).

In this work, the target geometry is the Hartree-Fock(HF)/6-31G* ground state optimised geometry, $\mathbf{R}_0$, calculated using MOLPRO (Werner *et al.*, 2012). Notably, in an experiment the target geometry would not be known *a priori*; the goal is to find it by sampling around a reasonable initial guess (via the SA algorithm). Auxiliary harmonic oscillator terms $\zeta_{\text{aux}} = \sum_k A_k (d'_k - d_k^{\text{start}})^2$ are included in ζ_{targ} for each first and second nearest neighbour inter-atomic distances d'_k and d_k^{start} for the distances in $\mathbf{R}'$ and the starting geometry $\mathbf{R}_{\text{start}}$ respectively; this restrains the molecular geometry from unrealistic regions of its conformational space. These terms have weighting factors A_k which are tuned such that the scattering term ζ_{signal} is the predominant driving

force in the minimisation, *i.e.* $\zeta_{\text{signal}}/\zeta_{\text{aux}} \sim 10$.

The starting geometry for the fitting procedure is initialised by slightly perturbing $\mathbf{R}_0$ (by moving it along all modes by a random small interval), and the final molecular geometry $\mathbf{R}_{\text{best}}$ is determined from the minimum of ζ_{targ} . Due to the stochastic nature of the SA algorithm, it is run 20 times for each fitting and the outcome with the lowest ζ_{targ} is selected. This avoids getting stuck in higher lying minima and likely provides a structure very close to the global minimum. A frequency calculation on the ground state $\mathbf{R}_0$ geometry is performed to obtain the normal mode unit vectors, $\hat{\mathbf{D}}_k$, which are used to sample all dimensions of the nuclear coordinate space when minimising the target function.

A metric used in this work to assess how close a given molecular geometry is to the reference geometry, is the mean absolute percent deviation (MAPD) (Yong *et al.*, 2021a) defined as,

$$\text{MAPD} = 100 \times \frac{2}{N_{\text{at}}^{\circ}(N_{\text{at}}^{\circ} - 1)} \sum_{i=1}^{N_{\text{at}}^{\circ}} \sum_{j>i}^{N_{\text{at}}^{\circ}} \frac{|d_{ij} - d'_{ij}|}{d'_{ij}}, \quad (5)$$

for total number of atoms of interest N_{at}° , where $N_{\text{at}}^{\circ} \leq N_{\text{at}}$. In this work, the non-hydrogen atoms are chosen in the calculation of the MAPD *i.e.* N_{at}° equals the number of non-hydrogen atoms. The distances between atom i and atom j are d_{ij} and d'_{ij} , and the prime (') denotes the reference structure, which is the ground state optimized structure $\mathbf{R}_0$ unless otherwise specified. In the following, we proceed to consider the three molecules: fluoroform (CHF_3), 1,3-cyclohexadiene (C_6H_{10} , CHD), and naphthalene (C_{10}H_8), which are shown in Fig. 1.

3. Results

3.1. IAM and *ab initio* scattering calculations

The top panels in Fig. 2 show a comparison between IAM and *ab initio* x-ray (top) and electron (bottom) scattering for the three molecules in Fig. 1. The $I(q)$ signal

is multiplied by q to enhance the visibility of the signal at large q (see the unscaled scattering intensity $I(q)$ plotted in the supplementary information; Fig. 9). The $I(s)$ signal is multiplied by s^4 for the same reason. Often in electron scattering experiments $M(s) = I_{\text{mol}}(s)/I_{\text{at}}(s)$ is plotted, which cancels out the s^{-4} Rutherford scattering term in $I(s)$; however, the *ab initio* $I(s)$ signal in this work cannot be decomposed into separate $I_{\text{mol}}(s)$ and $I_{\text{at}}(s)$ terms, so $s^4 I(s)$ is shown instead, cancelling the Rutherford scaling. The bottom panels in Fig. 2 show the percent difference (as in Eq. (3)) between IAM and *ab initio* scattering.

For x-ray scattering, all three molecules are similar in that the IAM underestimates signal in the approximate range $0 \leq q \leq 4.6 \text{ \AA}^{-1}$ ($0 \leq q \leq 5.9 \text{ \AA}^{-1}$ for CHF_3), albeit in CHF_3 IAM slightly overestimates for $q < 2.1 \text{ \AA}^{-1}$, and in all three IAM overestimates the scattering at larger q (approximately $4.6 \leq q \leq 8.4 \text{ \AA}^{-1}$ for CHD and naphthalene, and approximately $5.9 \leq q \leq 8.8 \text{ \AA}^{-1}$ for CHF_3) up until $q > \approx 8 \text{ \AA}^{-1}$ when IAM and *ab initio* become very similar; excellent agreement is seen here with $|\text{PD}(q > 8)| < 0.3\%$. The peak percent differences for CHD and naphthalene are relatively large, with $\text{PD}(q = 2.4) = -9.6\%$ and $\text{PD}(q = 2.7) = -8.6\%$ respectively. They are pure hydrocarbons which have been reported to contain larger chemical binding effect compared to molecules such as CCl_4 , N_2 , O_2 , and CS_2 (Shibata *et al.*, 2002). This is due to delocalisation of electrons via chemical bonding which the IAM does not take into account. For example, non-single bonds, aromatic rings, and hydrogens bonded to heavier atoms. Conversely, CHF_3 has a smaller peak percentage difference of $\text{PD}(q = 3.0) = -3.7\%$ due to the three C–F single bonds, which are quite well described by IAM, and has an absence of double bonds or delocalised electrons.

Similar to x-ray scattering, the electron scattering results show a substantial difference between IAM and *ab initio*, with the IAM performing the best for CHF_3 , whereas CHD and naphthalene have larger peaks in percentage difference $\text{PD}(s)$. The

maximum peak differences for each molecule are: CHF_3 has $\text{PD}(s = 3.2) = 6.2\%$, CHD has $\text{PD}(s = 2.3) = 17.0\%$, and naphthalene has $\text{PD}(s = 2.5) = 13.9\%$. These peak percent difference values are similar to the x-ray scattering result in terms of magnitude and molecule order, showing that the redistribution of electrons away from atom-centered spherical distributions is similarly visible. Finally, at $8 < s < 24$ the mean absolute percent difference is $\langle |\text{PD}(s)| \rangle < 0.3\%$ for each molecule, comparable to the x-ray scattering results at $8 < q < 12$.

3.2. Fitting to the target signal

Figs. 3 to 5 and Figs. 6 to 8 show the results of fitting the IAM signals to *ab initio* x-ray and electron scattering signals, respectively. The predicted data is $\mathbf{x} = I_{\text{IAM}}(q)$ for the x-ray scattering, as in Eq. (1) (not $qI(q)$ as shown in the figures), and $\mathbf{x} = s^4 I_{\text{IAM}}(s)$ for electron scattering. The corresponding target data is calculated by *ab initio* scattering theory at $\mathbf{R}_{\text{targ}} = \mathbf{R}_0$ for both x-ray and electron scattering, *i.e.* $\mathbf{y} = I_{\text{abinitio}}(q)$ or $\mathbf{y} = s^4 I_{\text{abinitio}}(s)$ respectively.

Various q (or s) ranges were fitted to assess the accuracy of IAM as a function of the values of the momentum transfer, to find the region most suitable to determine the molecular geometry with minimal ‘non-IAM’ contamination. This, correspondingly, is the region where the valence electronic structure component of the signal is strongest (that is, where the IAM and *ab initio* signals are significantly different).

At the top of each figure the best fit molecular geometries $\mathbf{R}_{\text{best}}$ (in solid) are shown together with the reference geometries $\mathbf{R}_{\text{targ}}$ (translucent) for the various q (or s) ranges. For x-ray scattering, the top graph shows the best-fit signal scaled by q , $q\mathbf{x}_{\text{best}} = qI_{\text{IAM}}(q; \mathbf{R}_{\text{best}})$, compared to $q\mathbf{y} = qI_{\text{abinitio}}(q; \mathbf{R}_{\text{targ}})$ for each signal range $q_{\text{min}} \leq q \leq q_{\text{max}}$. For electron scattering, similarly the best fits for each range are shown $\mathbf{x}_{\text{best}} = s^4 I_{\text{IAM}}(s; \mathbf{R}_{\text{best}})$ compared to $\mathbf{y} = s^4 I_{\text{abinitio}}(s; \mathbf{R}_{\text{targ}})$ for $s_{\text{min}} \leq s \leq s_{\text{max}}$. The

curves for each range of q or s used in the fitting are shifted vertically for visualisation purposes.

The bar charts at the bottom summarise the fitting for each signal range in terms of ζ_{signal} (notably, not ζ_{targ}) as in Eq. (4), and MAPD (Eq. (5)). Tables 1 and 2 show the values from these charts. It is clear that $8 \leq q \leq 12 \text{ \AA}^{-1}$ finds the molecular geometry closest to $\mathbf{R}_{\text{targ}}$; in fact, it is exceptionally close to the target, with mean absolute atom-atom distance deviations $\leq 0.9\%$ for all three molecules. Similarly, for electron scattering, the range $8 \leq s \leq 24 \text{ \AA}^{-1}$ gives the best geometry with $\text{MAPD} \leq 0.5\%$. This can be seen clearly from the overlap of the solid and translucent structures in (e) (or (d) for Figs. 6 to 8), whereas (a), (b), etc. have quite large visible deviations from $\mathbf{R}_{\text{targ}}$, such as stretching of C–F bonds, aromatic C–C bonds compacted/stretched, different hydrogen positions (and C–H distances). Importantly, MAPD and ζ_{signal} are generally correlated (aside from the smallest electron scattering range $0 \leq s \leq 4 \text{ \AA}^{-1}$ outliers), and the high- q (or s) ranges $8 \leq q \leq 12 \text{ \AA}^{-1}$ and $8 \leq s \leq 24 \text{ \AA}^{-1}$ have the lowest MAPD and ζ_{signal} in all cases except for one outlier in the CHF_3 electron scattering where the $0 \leq s \leq 24 \text{ \AA}^{-1}$ range has slightly lower (albeit very close) MAPD, despite much higher ζ_{signal} . The correlation between MAPD and ζ_{signal} is still promising however, as experimentally we don’t know the MAPD (because we don’t inherently know the molecular geometry) but do know the value of ζ_{signal} from comparison to theory. Therefore, fitting the x-ray (or electron) scattering curve via minimisation of ζ_{signal} (or equivalent) should give a structure close to the true structure.

Finally, the middle graph shows the valence electronic structure component as a percentage, $\text{PD}(q \text{ or } s)$, for the best fit geometry, $\mathbf{R}_{\text{best}}$, *i.e.* the structure with lowest ζ_{signal} for the ranges $8 \leq q \leq 12 \text{ \AA}^{-1}$ and $8 \leq s \leq 24 \text{ \AA}^{-1}$ for x-ray and electron scattering, respectively, using the *ab initio* signal calculated at $\mathbf{R}_{\text{targ}}$ as the reference. It is

compared to the percent differences between IAM and *ab initio* both at $\mathbf{R}_{\text{targ}}$ (dashed red line). The results show good agreement between $\text{PD}(q \text{ or } s; \mathbf{R}_{\text{best}})$ and $\text{PD}(q \text{ or } s; \mathbf{R}_{\text{targ}})$ for both x-ray and electron scattering. Notably, the electron results find near perfect agreement in this regard, revealing that $\mathbf{R}_{\text{best}}$ is closer to $\mathbf{R}_{\text{targ}}$ compared to the x-ray data fitting, hinting that electron scattering could be a better tool for molecular structure determination at larger scattering vector amplitudes compared to x-ray scattering. This is likely due to the dominance of elastic electron scattering at high values of s , whereas high- q x-ray scattering is dominated by unmodulated inelastic scattering. Despite this, x-ray scattering still performed well in this region, finding low MAPD structures for each molecule. It is striking that that the elastic scattering structural information persists here (see the Supplementary Information; Fig. 10).

Table 1. *Fitting results between IAM x-ray scattering and ab initio calculated at the target geometry $\mathbf{R}_{\text{targ}}$ at different q -ranges, q_{min} is shown in the first column and $q_{\text{max}} = 12 \text{ \AA}^{-1}$, and the data corresponds to the bar charts in Figs. 3 to 5. The target function, ζ_{signal} , is described in Eq. (4), and the mean absolute percent deviation (MAPD) is in Eq. (5).*

$q_{\text{min}} \text{ (\AA}^{-1}\text{)}$	ζ_{signal}	MAPD (%)
CHF ₃		
0	1.0×10^{-2}	4.5
2	9.5×10^{-3}	4.2
4	2.7×10^{-3}	1.3
6	1.9×10^{-3}	1.9
8	4.7×10^{-4}	0.2
CHD		
0	6.0×10^{-3}	5.3
2	7.0×10^{-3}	5.8
4	2.5×10^{-3}	4.3
6	1.3×10^{-4}	1.7
8	1.6×10^{-5}	0.8
Naphthalene		
0	9.1×10^{-3}	3.5
2	9.9×10^{-3}	3.6
4	4.2×10^{-3}	2.8
6	5.8×10^{-4}	1.6
8	2.4×10^{-5}	0.9

Table 2. *Fitting results between IAM electron scattering and ab initio calculated at the target geometry $\mathbf{R}_{\text{targ}}$ at different s -ranges, s_{min} is shown in the first column and s_{max} in the second, and the data corresponds to the bar charts in Figs. 6 to 8. The target function, ζ_{signal} , is described in Eq. (4), and the mean absolute percent deviation (MAPD) is in Eq. (5).*

s_{min} ($\text{\AA}^{-1}$)	s_{max} ($\text{\AA}^{-1}$)	ζ_{signal}	MAPD (%)
CHF ₃			
0	4	9.3×10^{-3}	5.8
0	8	2.5×10^{-2}	2.3
0	24	1.5×10^{-2}	0.3
8	24	5.1×10^{-4}	0.5
CHD			
0	4	2.8×10^{-1}	6.0
0	8	3.0×10^{-1}	3.0
0	24	1.3×10^{-1}	0.8
8	24	1.9×10^{-3}	0.4
Naphthalene			
0	4	4.2×10^{-1}	4.0
0	8	3.9×10^{-1}	2.1
0	24	1.7×10^{-1}	0.3
8	24	3.2×10^{-3}	0.2

4. Conclusion

The molecular geometry can be determined from large momentum transfer or equivalently large-angle scattering, with $q > 8 \text{ \AA}^{-1}$ (in the following, q also encompasses s). Although the elastic scattering component containing structural information drops off more quickly for x-ray than electron scattering, we find that sufficient elastic scattering persists to retrieve the correct molecular geometry using IAM theory in the range $8 < q < 12 \text{ \AA}^{-1}$ for both modalities of scattering. An important aspect to note is that in this high q regime, the contribution to the scattering due to chemical bonding is negligible, allowing the structure to be determined reliably using IAM theory. Conversely, if using IAM across the full available range of q , the resulting structure may be distorted from the correct $\mathbf{R}_0$ geometry. Another point to note is that the concept of structure may become ill-defined in excited state dynamics, where the dispersion of the nuclear wave packet leads to the coexistence of a range of structures that manifest as an effective damping of the high- q signal (Kirrander & Weber, 2017).

Nevertheless, for states of reasonably well-defined geometry, such as molecules in their ground electronic state, we can determine the molecular geometry using the large- q scattering, allowing us in the next step to extract the contribution to the scattering from the bonding valence electrons, at small and intermediate q . In this q -range, we find the deviation from IAM to be significant: $\sim 10\%$ for CHD and naphthalene, and $\sim 5\%$ for CHF_3 , for both x-ray and electron scattering.

We note that any robust inversion algorithm that can transform the one-dimensional scattering signal into reasonable molecular geometries would work (Yang *et al.*, 2014; Ishikawa *et al.*, 2015; Acheson & Kirrander, 2023). A practical challenge is that the large- q signal is small and is detected on a background of featureless inelastic scattering, which is demanding in terms of experimental signal-to-noise. The reliable detection of large- q signals might therefore be most appropriate at facilities such as the upgraded LCLS-II where high-repetition rates and photon energies upwards of 18 keV can help overcome such shortcomings. Nevertheless, given data of sufficient quality, the procedure outlined in this paper demonstrates that it should be possible to isolate the electronic contributions to the scattering signal, potentially opening the door for exciting new insights into electronic structure (Carrascosa *et al.*, 2022).

Appendix A

Appendix title

Acknowledgements

This research was supported by the National Science Foundation, Award #CHE-1953839. AK also acknowledges funding from EPSRC EP/V049240/2, EP/X026698/1, and EP/ X026973/1, and the Leverhulme Trust RPG-2020-208.

References

- Acheson, K. & Kirrander, A. (2023). *Journal of Chemical Theory and Computation*, **19**(10), 2721–2734.
- An-Ding, J. & Xiao-Lei, Z. (1995). *Acta Physico-Chimica Sinica*, **11**(7), 663–666.
- Ben-Nun, M., Martinez, T. J., Weber, P. M. & Wilson, K. R. (1996). *Chem. Phys. Lett.* **262**, 405.
- Bonham, R. & Iijima, T. (1963). *The Journal of Physical Chemistry*, **67**(11), 2266–2272.
- Bragg, W. H. & Bragg, W. L. (1915). *Nobel Lecture*.
- Carrascosa, A. M., Coe, J. P., Simmermacher, M., Paterson, M. J. & Kirrander, A. (2022). *Phys. Chem. Chem. Phys.* **24**(39), 24542–24552.
URL: <http://xlink.rsc.org/?DOI=D2CP02933B>
- Debye, P. (1915). *Ann. Phys.* **46**, 809.
- Debye, P. (1930). *Phys. Zeits.* **31**, 419–428.
- Debye, P., Bewilogua, L. & Ehrhardt, F. (1929). *Phys. Zeits.* **30**, 84.
- Duguet, A. & Jaegle, A. (1975). *Chemical Physics Letters*, **33**(3), 436–440.
- Fink, M. & Kessler, J. (1967). *The Journal of Chemical Physics*, **47**(5), 1780–1782.
- Gabalski, I., Sere, M., Acheson, K., Allum, F., Boutet, S., Dixit, G., Forbes, R., Glownia, J. M., Goff, N., Hegazy, K., Howard, A. J., Liang, M., Minitti, M. P., Minns, R. S., Natan, A., Peard, N., Rasmus, W. O., Sension, R. J., Ware, M. R., Weber, P. M., Werby, N., Wolf, T. J. A., Kirrander, A. & Bucksbaum, P. H. (2022). *J. Chem. Phys.* **157**(16), 164305.
URL: <https://aip.scitation.org/doi/10.1063/5.0113079>
- Glownia, J. M., Natan, A., Cryan, J. P., Hartsock, R., Kozina, M., Minitti, M. P., Nelson, S., Robinson, J., Sato, T., van Driel, T., Welch, G., Weninger, C., Zhu, D. & Bucksbaum, P. H. (2016). *Phys. Rev. Lett.* **117**, 153003.
- Hirota, F., Kakuta, N. & Shibata, S. (1981). *Journal of Physics B: Atomic and Molecular Physics*, **14**(17), 3299.
- Hoffmeyer, R. E., Bündgen, P. & Thakkar, A. J. (1998). *Journal of Physics B: Atomic, Molecular and Optical Physics*, **31**(16), 3675.
- Iijima, T., Bonham, R. A., Tavard, C., Roux, M. & Cornille, M. (1965). *Bulletin of the Chemical Society of Japan*, **38**(10), 1757–1760.
- Ishikawa, T., Hayes, S. A., Keskin, S., Corthey, G., Hada, M., Pichugin, K., Marx, A., Hirscht, J., Shionuma, K., Onda, K., Okimoto, Y., Koshihara, S.-y., Yamamoto, T., Cui, H., Nomura, M., Oshima, Y., Abdel-Jawad, M., Kato, R. & Miller, R. J. D. (2015). *Science*, **350**(6267), 1501–1505.
URL: <http://www.sciencemag.org/cgi/doi/10.1126/science.aab3480>
- King, W. E., Campbell, G. H., Frank, A., Reed, B., Schmerge, J. F., Siwick, B. J., Stuart, B. C. & Weber, P. M. (2005). *Journal of Applied Physics*, **97**(11).
- Kirrander, A. (2012). *J. Chem. Phys.* **137**, 154310.
- Kirrander, A., Saita, K. & Shalashilin, D. V. (2016). *J. Chem. Theory Comput.* **12**, 957–967.
URL: <http://dx.doi.org/10.1021/acs.jctc.5b01042>
- Kirrander, A. & Weber, P. M. (2017). *Appl. Science*, **7**(6), 534.
URL: <http://www.mdpi.com/2076-3417/7/6/534>
- Küpper, J., Stern, S., Holmegaard, L., Filsinger, F., Rouzée, A., Rudenko, A., Johnsson, P., Martin, A. V., Adolph, M., Aquila, A., Bajt, S., Barty, A., Bostedt, C., Bozek, J., Caleman, C., Coffee, R., Coppola, N., Delmas, T., Epp, S., Erk, B., Foucar, L., Gorkhover, T., Gumprecht, L., Hartmann, A., Hartmann, R., Hauser, G., Holl, P., Hömke, A., Kimmel, N., Krasniqi, F., Kühnel, K.-U., Maurer, J., Messerschmidt, M., Moshhammer, R., Reich, C., Rudek, B., Santra, R., Schlichting, I., Schmidt, C., Schorb, S., Schulz, J., Soltau, H., Spence, J. C. H., Starodub, D., Strüder, L., Thøgersen, J., Vrakking, M. J. J., Weidenspointner, G., White, T. A., Wunderer, C., Meijer, G., Ullrich, J., Stapelfeldt, H., Rolles, D. & Chapman, H. N. (2014). *Phys. Rev. Lett.* **112**, 083002.
- von Laue, M. (1915). *Nobel lecture*, **13**.
- Ma, L., Yong, H., Geiser, J. D., Carrascosa, A. M., Goff, N. & Weber, P. M. (2020). *Struct. Dyn.* **7**(3), 034102.
URL: <http://aca.scitation.org/doi/10.1063/4.0000010>

- Mai, S., Richter, M., Ruckebauer, M., Oppel, M., Marquetand, P. & González, L., (2014). Sharc: Surface hopping including arbitrary couplings - program package for non-adiabatic dynamics. share-md.org.
- Minitti, M. P., Budarz, J. M., Kirrander, A., Robinson, J., Lane, T. J., Ratner, D., Saita, K., Northey, T., Stankus, B., Cofer-Shabica, V., Hastings, J. & Weber, P. M. (2014). *Faraday Discuss.* **171**, 81.
URL: <http://dx.doi.org/10.1039/C4FD00030G>
- Minitti, M. P., Budarz, J. M., Kirrander, A., Robinson, J. S., Ratner, D., Lane, T. J., Zhu, D., Glowina, J. M., Kozina, M., Lemke, H. T., Sikorski, M., Feng, Y., Nelson, S., Saita, K., Stankus, B., Northey, T., Hastings, J. B. & Weber, P. M. (2015). *Phys. Rev. Lett.* **114**(25), 255501.
URL: <http://link.aps.org/doi/10.1103/PhysRevLett.114.255501>
- Mitzel, N. W. & Rankin, D. W. (2003). *Dalton Transactions*, (19), 3650–3662.
- Moreno Carrascosa, A., Coe, J. P., Simmermacher, M., Paterson, M. J. & Kirrander, A. (2022). *Phys. Chem. Chem. Phys.* **24**(39), 24542–24552.
URL: <http://xlink.rsc.org/?DOI=D2CP02933B>
- Moreno Carrascosa, A., Yong, H., Crittenden, D. L., Weber, P. M. & Kirrander, A. (2019). *J. Chem. Theory Comp.* **15**, 2836–2846.
URL: <https://doi.org/10.1021/acs.jctc.9b00056>
- Mott, N. F. (1930). *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, **127**(806), 658–665.
- Northey, T. & Kirrander, A. (2019). *J. Phys. Chem. A*, **123**(15), 3395–3406.
URL: <https://doi.org/10.1021/acs.jpca.9b00621>
- Northey, T., Moreno Carrascosa, A., Schäfer, S. & Kirrander, A. (2016). *J. Chem. Phys.* **145**, 154304.
URL: <http://dx.doi.org/10.1063/1.4962256>
- Northey, T., Weber, P. M. & Kirrander, A. (2023). (*manuscript in preparation*).
- Northey, T., Zotev, N. & Kirrander, A. (2014). *J. Chem. Theory Comput.* **10**(11), 4911.
URL: <http://dx.doi.org/10.1021/ct500096r>
- Odate, A., Kirrander, A., Weber, P. M. & Minitti, M. P. (2023). *Adv. Phys.: X*, **8**(1), 2126796.
URL: <https://www.tandfonline.com/doi/full/10.1080/23746149.2022.2126796>
- Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. & North, A. (1960). *Nature*, **185**(4711), 416–422.
- Pirenne, M. H. (1939). *J. Chem. Phys.* **7**(3), 144–155.
URL: <https://doi.org/10.1063/1.1750402>
- Pirenne, M. H. (1946). *The Diffraction of X-Rays and Electrons by Free Molecules*. London: Cambridge University Press.
- Prince, E. (ed.) (2006). *International Tables for Crystallography Volume C: Mathematical, physical and chemical tables*. No. ISBN 978-1-4020-1900-5. Wiley, 2006th ed.
- Ruddock, J. M., Zotev, N., Stankus, B., Yong, H.-W., Bellshaw, D., Boutet, S., Lane, T. J., Liang, M., Carbajo, S., Du, W., Kirrander, A., Minitti, M. P. & Weber, P. M. (2019). *Angew. Chem. Int. Ed.* **58**(19), 6371–6375.
URL: <http://dx.doi.org/10.1002/anie.201902228>
- Sciaini, G. & Miller, R. D. (2011). *Reports on Progress in Physics*, **74**(9), 096101.
- Shibata, S., Hirota, F. & Shioda, T. (1999). *Journal of Molecular Structure*, **485**, 1–11.
- Shibata, S., Sekiyama, H., Tachikawa, K. & Moribe, M. (2002). *Journal of Molecular Structure*, **641**(1), 1–6.
- Simmermacher, M., Henriksen, N. E., Møller, K. B., Moreno Carrascosa, A. & Kirrander, A. (2019a). *Phys. Rev. Lett.* **122**, 073003.
URL: <https://link.aps.org/doi/10.1103/PhysRevLett.122.073003>
- Simmermacher, M., Henriksen, N. E., Møller, K. B., Moreno Carrascosa, A. & Kirrander, A. (2019b). *J. Chem. Phys.* **151**(17), 174302.
URL: <https://doi.org/10.1063/1.5110040>

- Stankus, B., Yong, H., Zotev, N., Ruddock, J. M., Bellshaw, D., Lane, T. J., Liang, M., Boutet, S., Carbajo, S., Robinson, J. S., Du, W., Goff, N., Chang, Y., Koglin, J. E., Minitti, M. P., Kirrander, A. & Weber, P. M. (2019). *Nat. Chem.* **11**, 716–721.
URL: <http://www.nature.com/articles/s41557-019-0291-0>
- Stefanou, M., Saita, K., Shalashilin, D. V. & Kirrander, A. (2017). *Chem. Phys. Lett.* **683**, 300–305.
URL: <https://doi.org/10.1016/j.cplett.2017.03.007>
- Suominen, H. J. & Kirrander, A. (2014). *Phys. Rev. Lett.* **112**, 043002.
- Wang, J., Tripathi, A. N. & Smith Jr, V. H. (1994). *The Journal of Chemical Physics*, **101**(6), 4842–4854.
- Watson, J. D. & Crick, F. H. (1953). *Nature*, **171**(4356), 737–738.
- Weathersby, S., Brown, G., Centurion, M., Chase, T., Coffee, R., Corbett, J., Eichner, J., Frisch, J., Fry, A., Gühr, M. *et al.* (2015). *Review of Scientific Instruments*, **86**(7).
- Werner, H.-J., Knowles, P. J., Knizia, G., Manby, F. R. & Schütz, M. (2012). *WIREs Comput. Mol. Sci.* **2**, 242–253.
- Yang, J., Makhija, V., Kumarappan, V. & Centurion, M. (2014). *Struct. Dyn.* **1**(4), 044101.
URL: <http://aca.scitation.org/doi/10.1063/1.4889840>
- Yong, H., Carrascosa, A. M., Ma, L., Stankus, B., Minitti, M. P., Kirrander, A. & Weber, P. M. (2021a). *Faraday Disc.* **228**, 104–122.
URL: <http://xlink.rsc.org/?DOI=D0FD00118J>
- Yong, H., Cavaletto, S. M. & Mukamel, S. (2021b). *The Journal of Physical Chemistry Letters*, **12**(40), 9800–9806.
- Yong, H., Xu, X., Ruddock, J. M., Stankus, B., Carrascosa, A. M., Zotev, N., Bellshaw, D., Du, W., Goff, N., Chang, Y., Boutet, S., Carbajo, S., Koglin, J. E., Liang, M., Robinson, J. S., Kirrander, A., Minitti, M. P. & Weber, P. M. (2021c). *Proc. Nat. Acad. Sci.* **118**(19), e2021714118.
URL: <http://www.pnas.org/content/118/19/e2021714118.abstract>
- Yong, H., Zotev, N., Ruddock, J. M., Stankus, B., Simmermacher, M., Carrascosa, A. M., Du, W., Goff, N., Chang, Y., Bellshaw, D., Liang, M., Carbajo, S., Koglin, J. E., Robinson, J. S., Boutet, S., Minitti, M. P., Kirrander, A. & Weber, P. M. (2020). *Nat. Comm.* **11**(1), 2157.
URL: <http://www.nature.com/articles/s41467-020-15680-4>
- Zandi, O., Wilkin, K. J., Xiong, Y. & Centurion, M. (2017). *Structural Dynamics*, **4**(4).
- Ziems, K. M., Simmermacher, M., Gräfe, S. & Kirrander, A. (2023). *J. Chem. Phys.* **159**(4), 044108.
URL: <https://pubs.aip.org/jcp/article/159/4/044108/2904227/The-contribution-of-Compton-ionization-to>
- Zotев, N., Carrascosa, A. M., Simmermacher, M. & Kirrander, A. (2020). *J. Chem. Theory Comput.* **16**, 2594–2605.
URL: <https://dx.doi.org/10.1021/acs.jctc.9b00670>

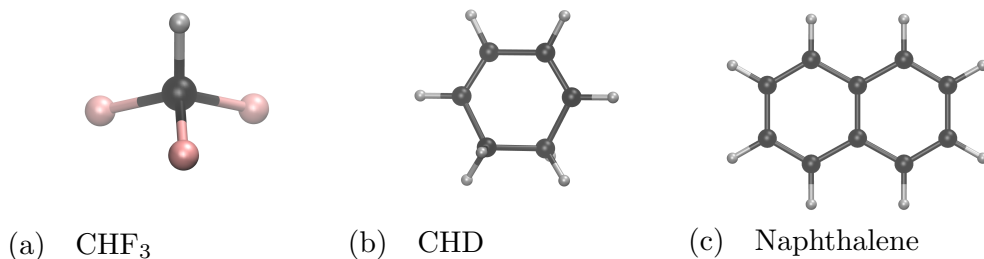


Fig. 1. The three molecules (a) fluoroform (CHF₃), (b) 1,3-cyclohexadiene (C₆H₈, CHD), and (c) naphthalene (C₁₀H₈).

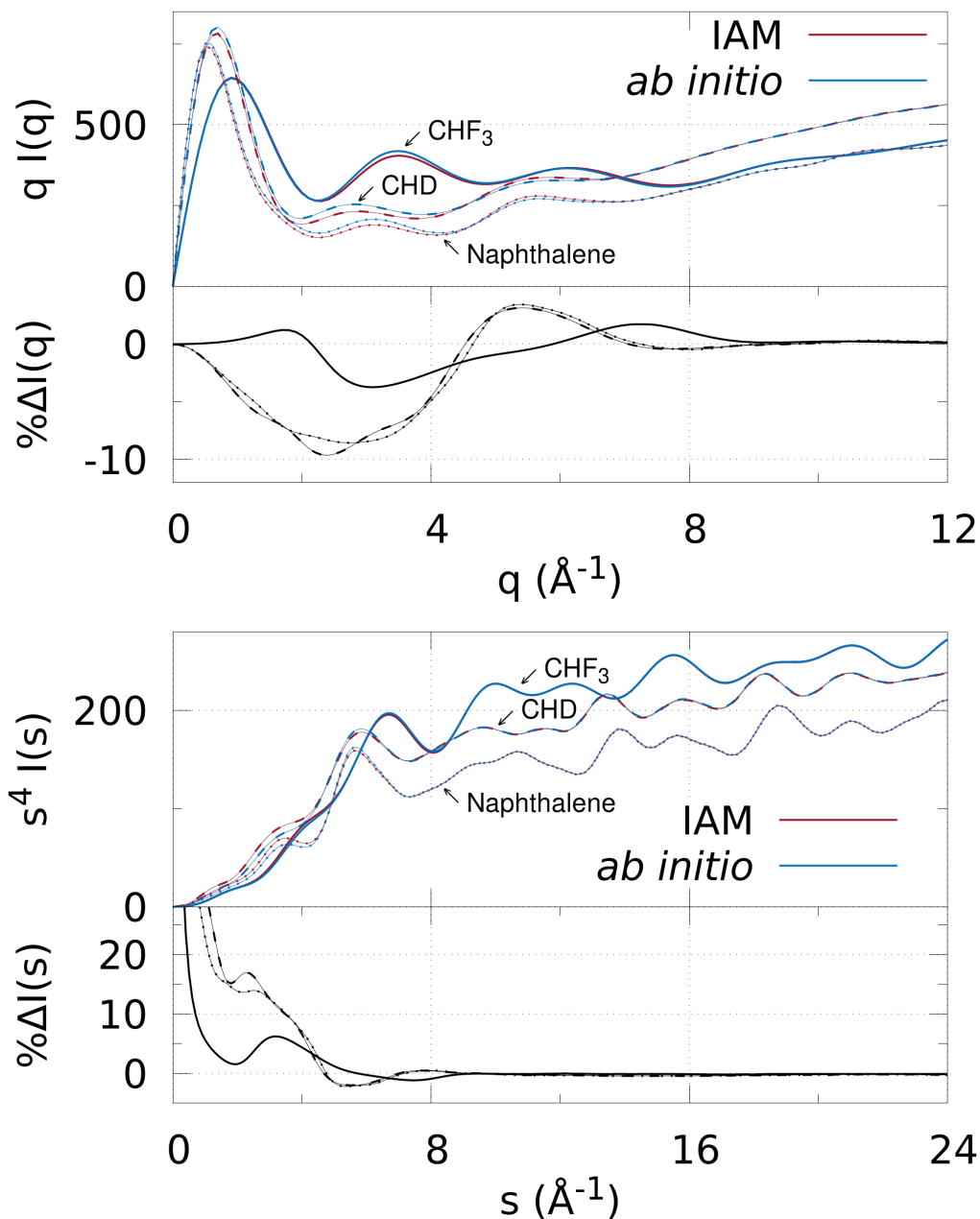


Fig. 2. Comparison between IAM and *ab initio* scattering (top figure: x-ray scattering, bottom figure: electron scattering) for the three molecules CHF₃, CHD, and naphthalene in their optimised ground state geometries $\mathbf{R}_0$. The scattering signals are shown in the top panels (IAM in red lines and *ab initio* in blue), with the signals scaled by q and s^4 , respectively, *i.e.* $qI(q)$ and $s^4I(s)$. In both top and bottom figures, the naphthalene signal has been multiplied by 0.5 for visualisation purposes. The bottom panels show percent differences between *ab initio* and IAM scattering, as defined in Eq. (3). Note that the electron scattering PD(s) becomes very large at small s due to division by small numbers, so the y -axis is truncated.

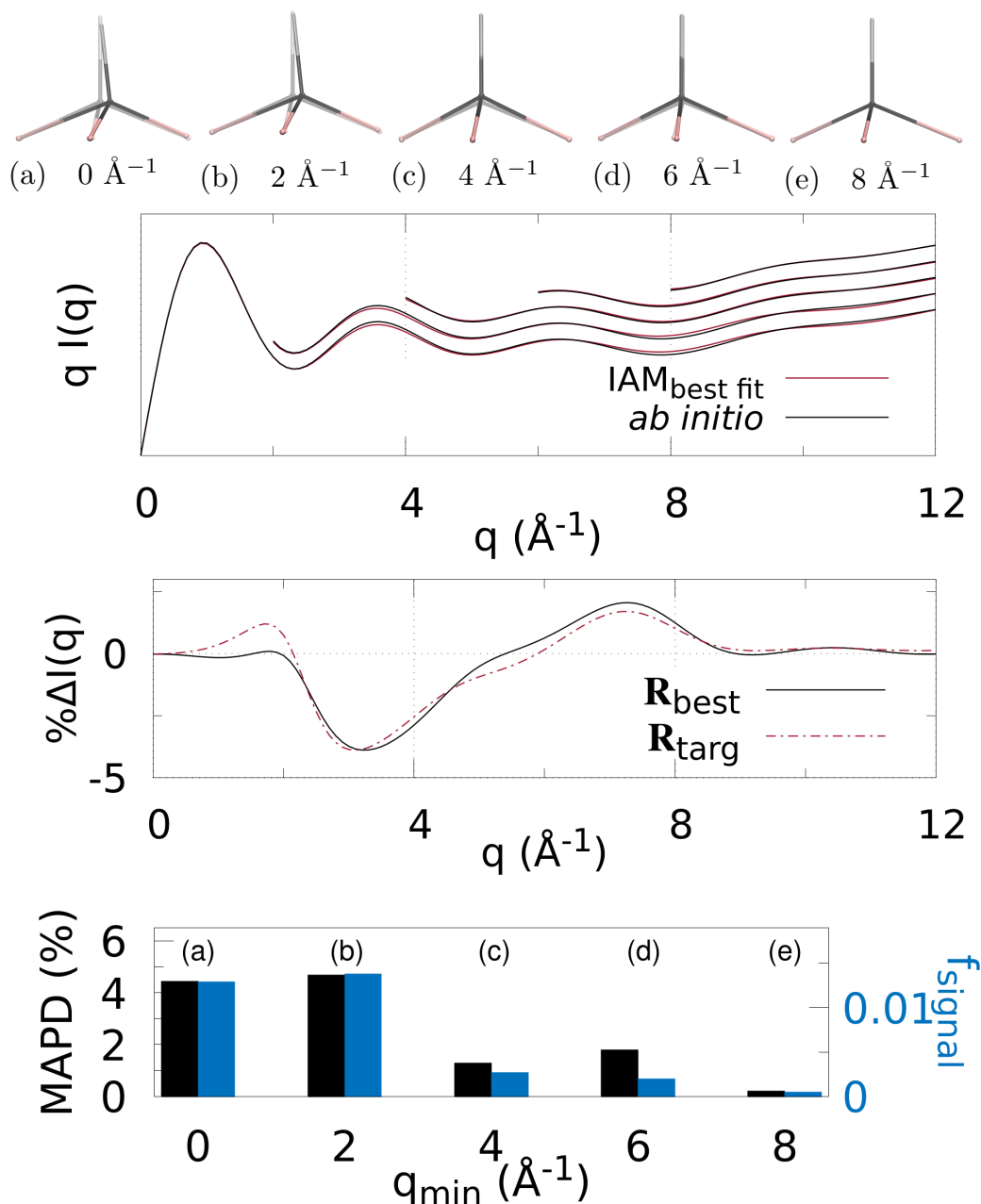


Fig. 3. X-ray scattering for CHF_3 using different q -ranges for structure determination: (a-e) The best fit geometry for each $q_{\min}$ (solid) versus the target structure $\mathbf{R}_{\text{targ}}$ (translucent), where $q \in [q_{\min}, q_{\max}]$, and $q_{\min} = [0, 2, 4, 6, 8] \text{ \AA}^{-1}$ and $q_{\max} = 12 \text{ \AA}^{-1}$. (Top) The IAM best fits to the *ab initio* x-ray scattering calculated at $\mathbf{R}_{\text{targ}}$ as a function of q -range. The $q_{\min} \geq 2 \text{ \AA}^{-1}$ curves are shifted vertically for visibility. (Middle) The percent difference $\text{PD}(q)$ (Eq. (3)) (solid black line) for the lowest ζ_{targ} structure. For comparison, $\text{PD}(q; \mathbf{R}_{\text{targ}})$ is also shown, using $I_{\text{IAM}}(q; \mathbf{R}_{\text{targ}})$ and $I_{\text{abinitio}}(q; \mathbf{R}_{\text{targ}})$. (Bottom) Bar chart showing the MAPD (Eq. (5)) and ζ_{signal} for each value of $q_{\min}$.

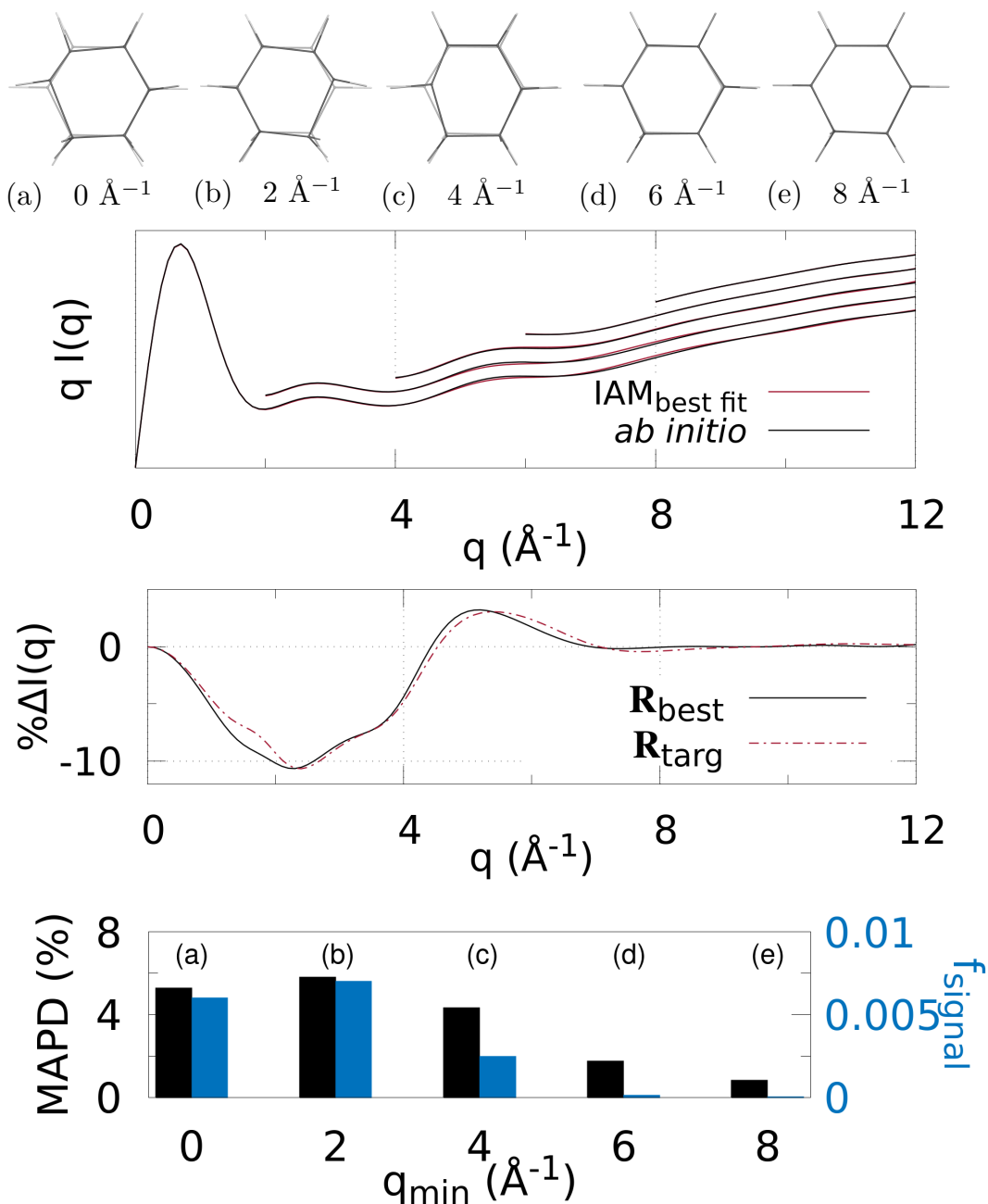


Fig. 4. X-ray scattering of CHD using different q -ranges for structure determination: (a-e) The best fit geometry for each $q_{\min}$ (solid) versus the target structure $\mathbf{R}_{\text{targ}}$ (translucent), where $q \in [q_{\min}, q_{\max}]$, and $q_{\min} = [0, 2, 4, 6, 8] \text{ \AA}^{-1}$ and $q_{\max} = 12 \text{ \AA}^{-1}$. (Top) The IAM best fits to the *ab initio* x-ray scattering calculated at $\mathbf{R}_{\text{targ}}$ as a function of q -range. The $q_{\min} \geq 2 \text{ \AA}^{-1}$ curves are shifted vertically for visibility. (Middle) The percent difference PD(q) (Eq. (3)) (solid black line) for the lowest ζ_{targ} structure. For comparison, PD($q; \mathbf{R}_{\text{targ}}$) is also shown, using $I_{\text{IAM}}(q; \mathbf{R}_{\text{targ}})$ and $I_{\text{abinitio}}(q; \mathbf{R}_{\text{targ}})$. (Bottom) Bar chart showing the MAPD (Eq. (5)) and ζ_{signal} for each value of $q_{\min}$.

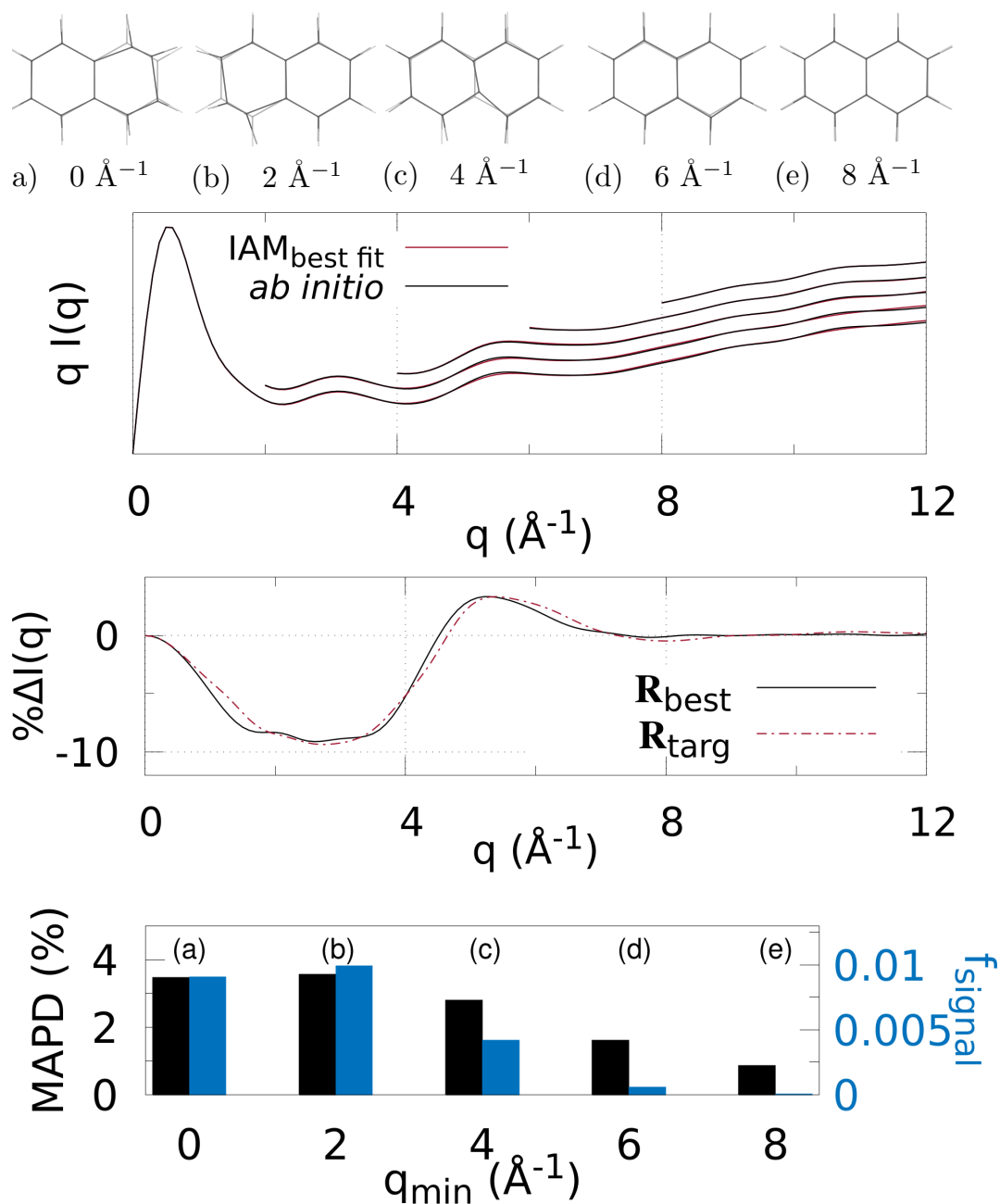


Fig. 5. X-ray scattering of naphthalene using different q -ranges for structure determination: (a-e) The best fit geometry for each $q_{\min}$ (solid) versus the target structure $\mathbf{R}_{\text{targ}}$ (translucent), where $q \in [q_{\min}, q_{\max}]$, and $q_{\min} = [0, 2, 4, 6, 8] \text{ \AA}^{-1}$ and $q_{\max} = 12 \text{ \AA}^{-1}$. (Top) The IAM best fits to the *ab initio* x-ray scattering calculated at $\mathbf{R}_{\text{targ}}$ as a function of q -range. The $q_{\min} \geq 2 \text{ \AA}^{-1}$ curves are shifted vertically for visibility. (Middle) The percent difference $\text{PD}(q)$ (Eq. (3)) (solid black line) for the lowest ζ_{targ} structure. For comparison, $\text{PD}(q; \mathbf{R}_{\text{targ}})$ is also shown, using $I_{\text{IAM}}(q; \mathbf{R}_{\text{targ}})$ and $I_{\text{abinitio}}(q; \mathbf{R}_{\text{targ}})$. (Bottom) Bar chart showing the MAPD (Eq. (5)) and ζ_{signal} for each value of $q_{\min}$.

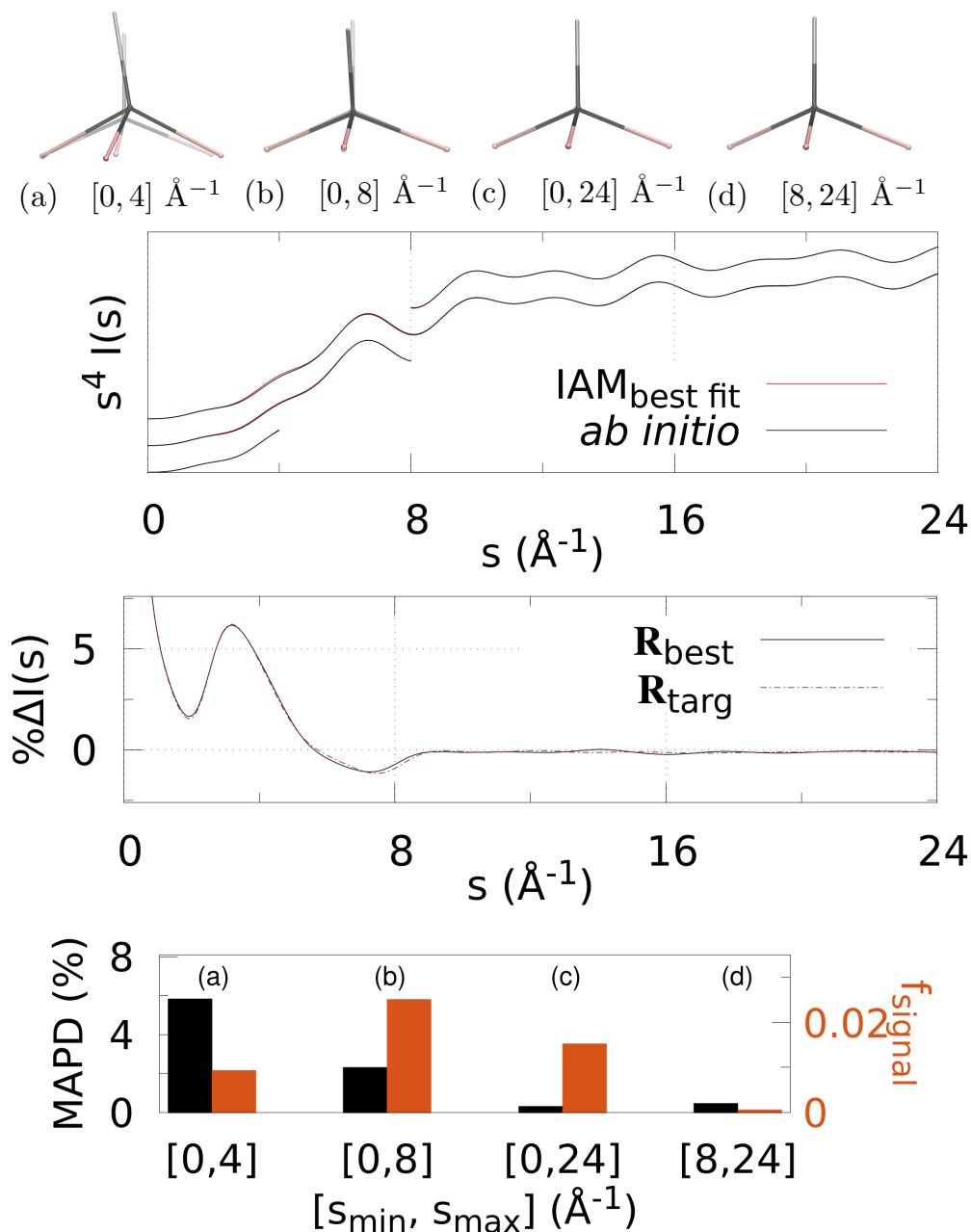


Fig. 6. Electron scattering of CHF₃ using different s -ranges for structure determination: (a-d) The best fit geometry for each s -range (solid) versus the reference structure $\mathbf{R}_{\text{targ}}$ (translucent), where $s \in [s_{\text{min}}, s_{\text{max}}] = [0, 4], [0, 8], [0, 24], [8, 24] \text{ \AA}^{-1}$. (Top) The IAM best fits to the *ab initio* electron scattering calculated at $\mathbf{R}_{\text{targ}}$ as a function of s -range. The $s_{\text{max}} > 4 \text{ \AA}^{-1}$ curves are shifted vertically for visibility. (Middle) The percent difference $\text{PD}(s)$ (Eq. (3)) (solid black line) for the lowest ζ_{targ} structure. For comparison, $\text{PD}(s; \mathbf{R}_{\text{targ}})$ is also shown, using $I_{\text{IAM}}(s; \mathbf{R}_{\text{targ}})$ and $I_{\text{abinitio}}(s; \mathbf{R}_{\text{targ}})$. (Bottom) Bar chart showing the MAPD (Eq. (5)) and ζ_{signal} for each value of $[s_{\text{min}}, s_{\text{max}}]$.

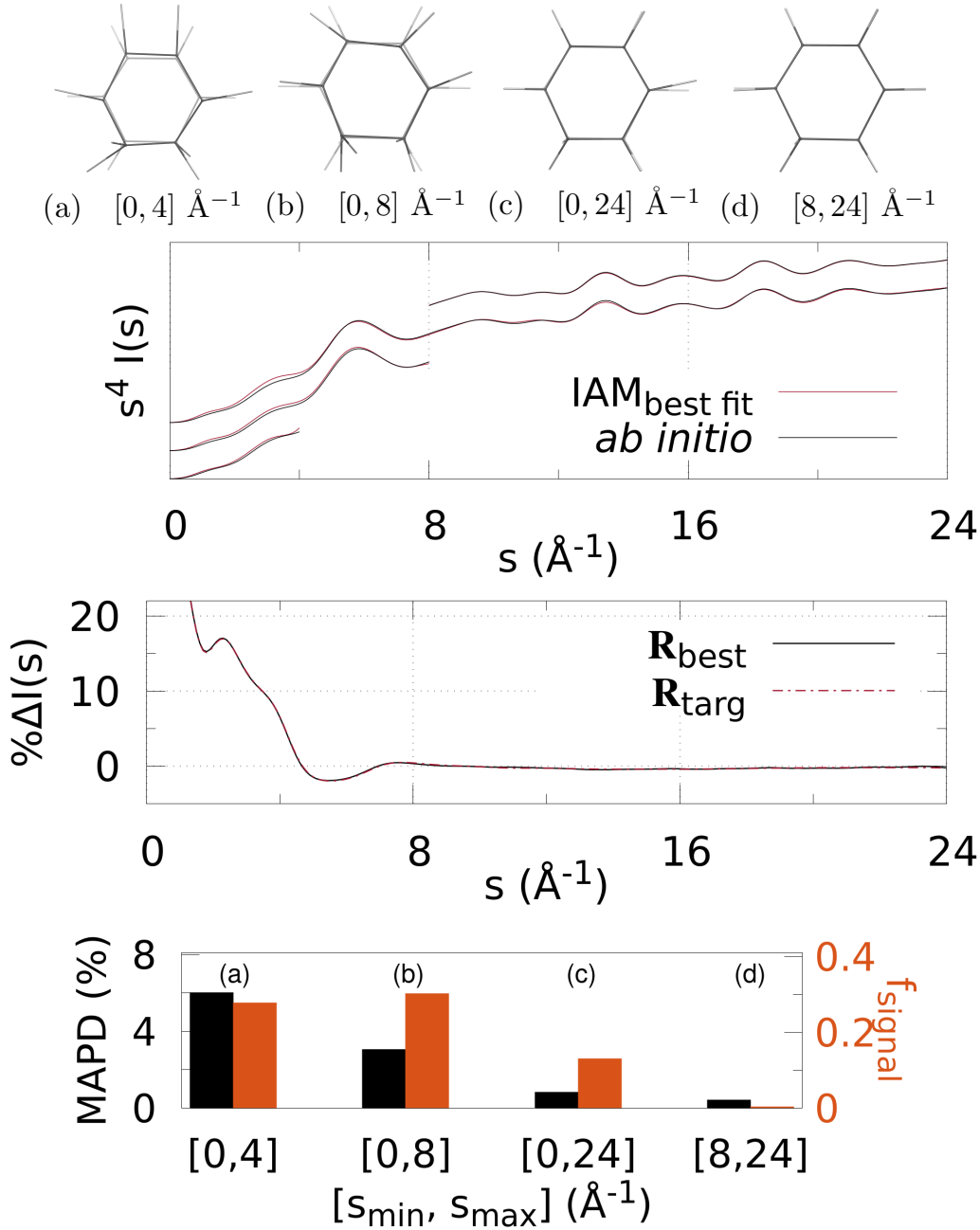


Fig. 7. Electron scattering of CHD using different s -ranges for structure determination: (a-d) The best fit geometry for each s -range (solid) versus the reference structure $\mathbf{R}_{\text{targ}}$ (translucent), where $s \in [s_{\text{min}}, s_{\text{max}}] = [0, 4], [0, 8], [0, 24], [8, 24] \text{ \AA}^{-1}$. (Top) The IAM best fits to the *ab initio* electron scattering calculated at $\mathbf{R}_{\text{targ}}$ as a function of s -range. The $s_{\text{max}} > 4 \text{ \AA}^{-1}$ curves are shifted vertically for visibility. (Middle) The percent difference PD(s) (Eq. (3)) (solid black line) for the lowest ζ_{targ} structure. For comparison, PD($s; \mathbf{R}_{\text{targ}}$) is also shown, using $I_{\text{IAM}}(s; \mathbf{R}_{\text{targ}})$ and $I_{\text{abinitio}}(s; \mathbf{R}_{\text{targ}})$. (Bottom) Bar chart showing the MAPD (Eq. (5)) and ζ_{signal} for each value of $[s_{\text{min}}, s_{\text{max}}]$.

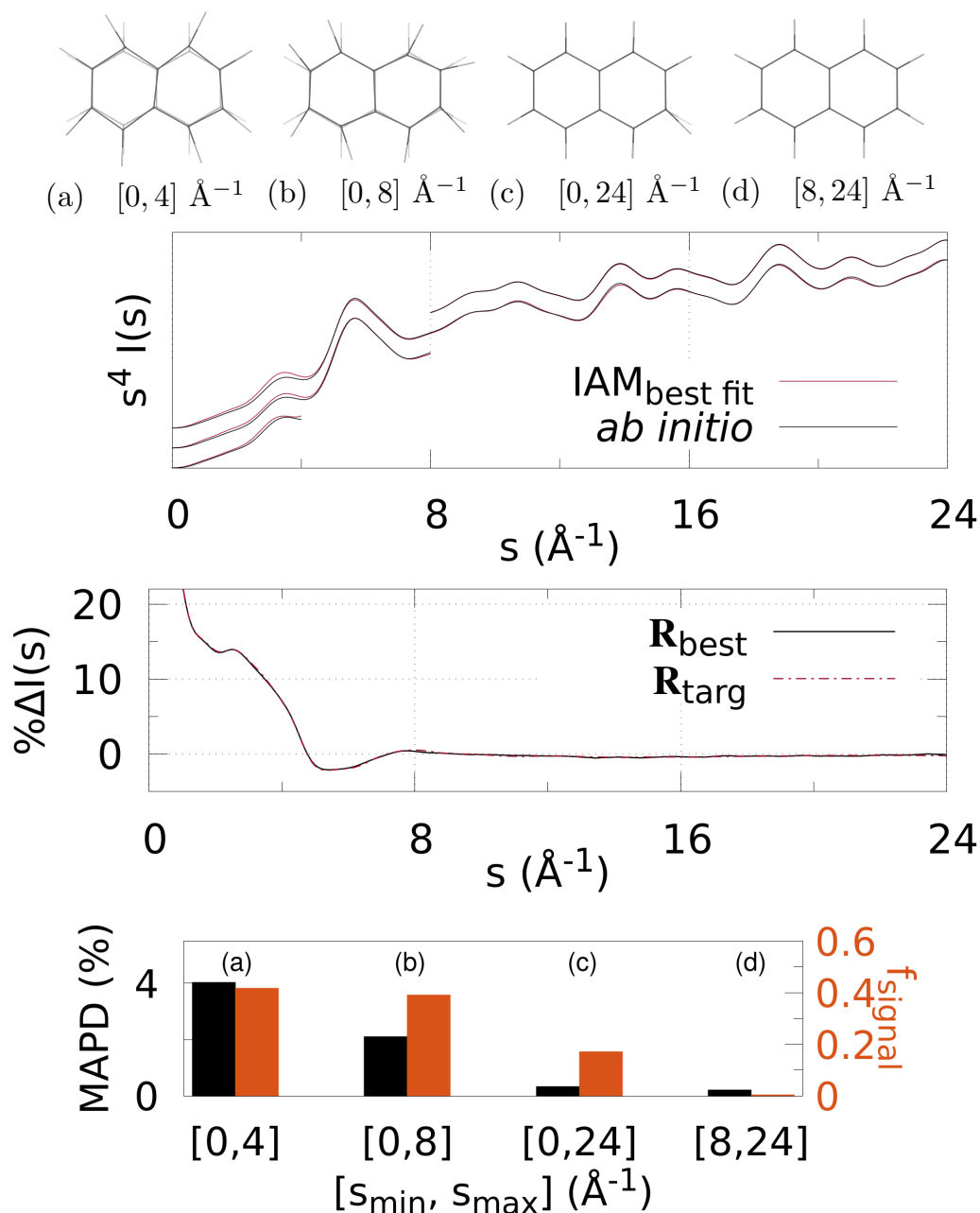


Fig. 8. Electron scattering of naphthalene using different s -ranges for structure determination: (a-d) The best fit geometry for each s -range (solid) versus the reference structure $\mathbf{R}_{\text{targ}}$ (translucent), where $s \in [s_{\text{min}}, s_{\text{max}}] = [0, 4], [0, 8], [0, 24], [8, 24] \text{ \AA}^{-1}$. (Top) The IAM best fits to the *ab initio* electron scattering calculated at $\mathbf{R}_{\text{targ}}$ as a function of s -range. The $s_{\text{max}} > 4 \text{ \AA}^{-1}$ curves are shifted vertically for visibility. (Middle) The percent difference $\text{PD}(s)$ (Eq. (3)) (solid black line) for the lowest ζ_{targ} structure. For comparison, $\text{PD}(s; \mathbf{R}_{\text{targ}})$ is also shown, using $I_{\text{IAM}}(s; \mathbf{R}_{\text{targ}})$ and $I_{\text{abinitio}}(s; \mathbf{R}_{\text{targ}})$. (Bottom) Bar chart showing the MAPD (Eq. (5)) and ζ_{signal} for each value of $[s_{\text{min}}, s_{\text{max}}]$.

Appendix B Supplementary

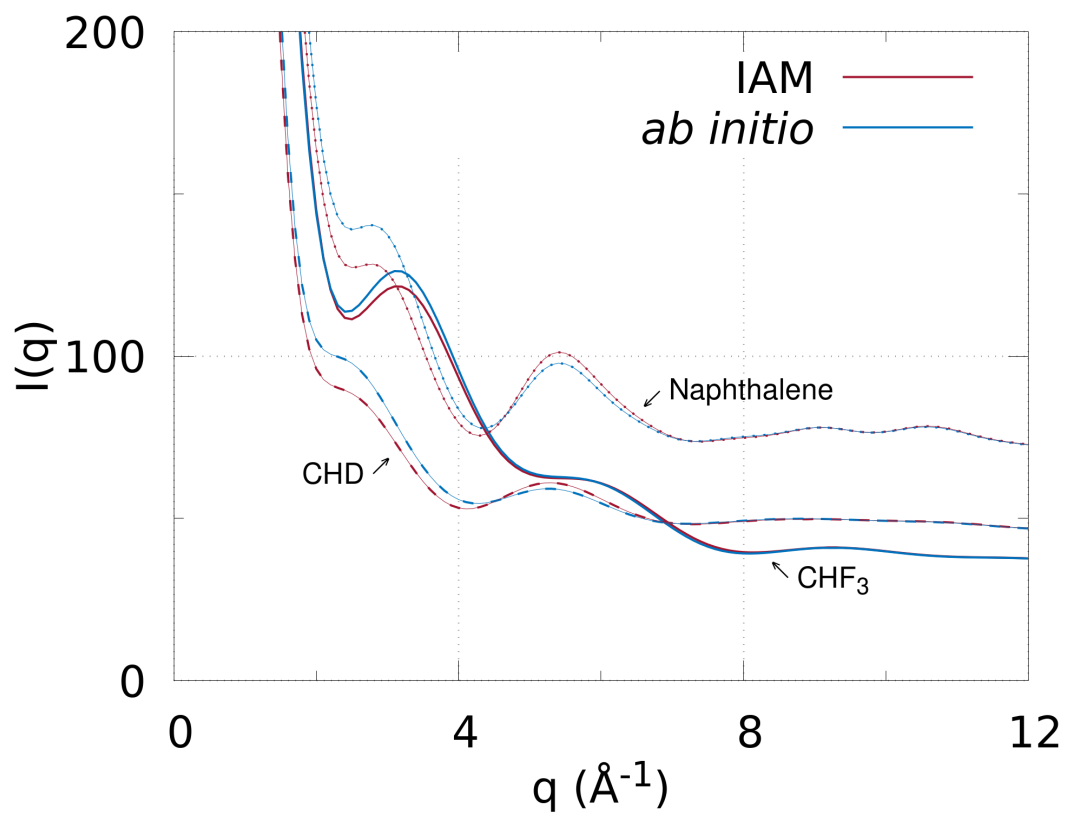


Fig. 9. A comparison between IAM and *ab initio* theory total x-ray scattering signals for the three molecules at their optimised ground state geometries ($\mathbf{R}_0$). The signal at $q = 0$ for each molecule is equal to its number of electrons squared, i.e. $I(q = 0) = N_{\text{el}}^2$ (off the scale of the figure), and at large q the signals tend to N_{el} , i.e. $I(q \rightarrow \infty) = N_{\text{el}}$.

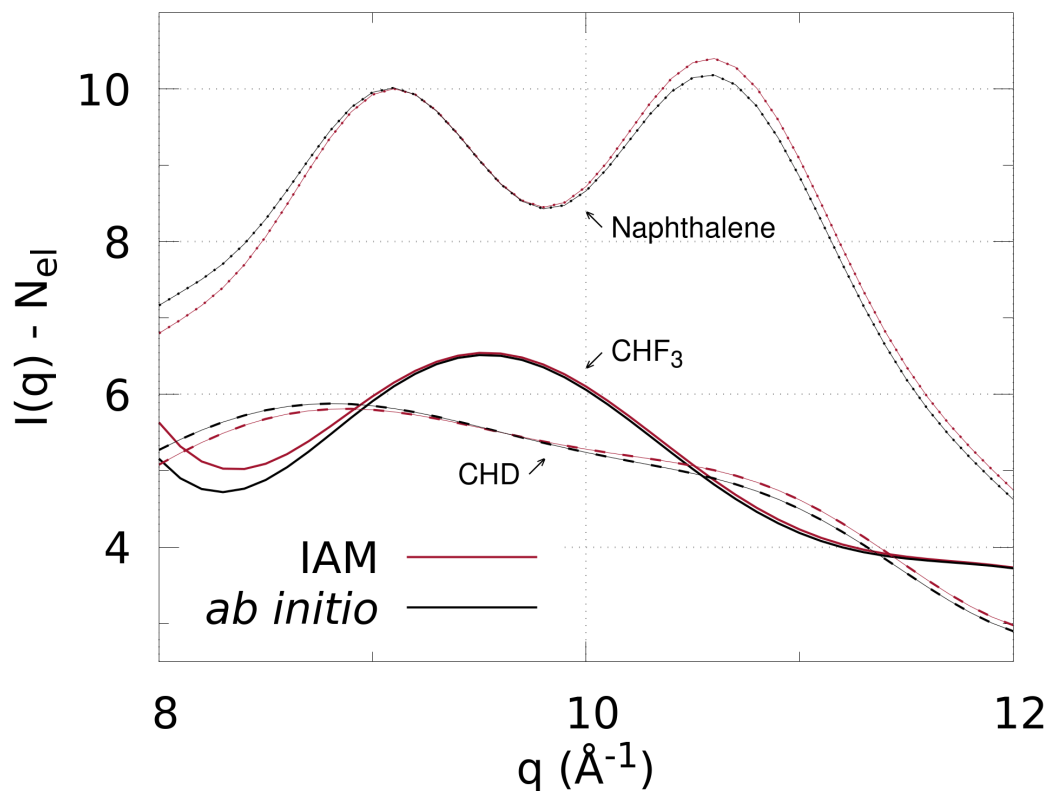


Fig. 10. High- q x-ray scattering signals ($q \in [8, 12] \text{ \AA}^{-1}$) calculated with IAM and *ab initio* scattering theory minus number of electrons N_{el} for each molecule.

B.1. Consideration of vibrations

The synthetic target data, calculated using *ab initio* theory did not account for vibrational effects since the focus of this study was on the electronic contribution. To justify this, Fig. 11 shows an average of $N = 1,000$ (which is converged) IAM signals for CHD obtained via geometries from a 300 K Wigner distribution centred at the ground state geometry. The Wigner distribution was obtained via the SHARC software (Mai *et al.*, 2014). The plot shows the percentage difference compared to the IAM of signal of the central geometry. The vibrational effects are $\sim 0.4\%$, which is much smaller than the 5 – 10% arising from the electronic structure effects. Notably,

the entire q -range is sensitive to the nuclear structure vibrations.

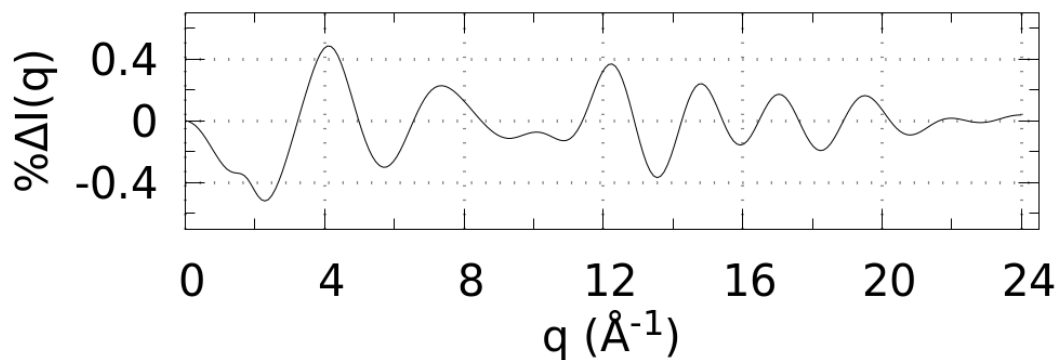


Fig. 11. Percentage difference between: the average of $N = 1,000$ x-ray scattering calculations (using IAM) for the CHD molecule, with geometries taken from a 300 K Wigner distribution centred at the ground state geometry; and the signal of the central geometry.

Synopsis

A new method to determine molecular structure from large-angle scattering for x-ray and electron scattering from free gas-phase molecules.
