Modeling Gas-Phase Unimolecular Dissociation for Bond Dissociation Energies: Comparison of Statistical Rate Models within RRKM Theory

Journal Article

Author(s):
Paenurk, Eno; Chen, Peter

Publication date:
2021-03-11

Permanent link:
https://doi.org/10.3929/ethz-b-000472010

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
The Journal of Physical Chemistry A 125(9), https://doi.org/10.1021/acs.jpca.1c00183
Modeling Gas-Phase Unimolecular Dissociation for Bond Dissociation Energies: Comparison of Statistical Rate Models within RRKM Theory

Eno Paenurk, Peter Chen*

Department of Chemistry and Applied Biosciences, ETH Zürich, 8093 Zürich, Switzerland
*E-mail: peter.chen@org.chem.ethz.ch

Abstract

The Rice-Ramsperger-Kassel-Marcus (RRKM) theory provides a simple yet powerful rate theory for calculating microcanonical rate constants. In particular, it has found widespread use in combination with gas-phase kinetic experiments of unimolecular dissociations to extract experimental bond dissociation energies (BDE). We have previously found several discrepancies between computed BDE values and the respective experimental ones, obtained with our empirical rate model, named L-CID. To investigate the reliability of our rate model, we conducted a theoretical analysis and comparison of the performance of conventional rate models and L-CID within the RRKM framework. Using previously published microcanonical rate data as well as reaction cross-section data, we show that the BDE values obtained with the L-CID model agree with the ones from the other rate models within the expected uncertainty bounds. Based on this agreement, we discuss the possible rationalization of the good performance of the L-CID model.

Introduction

Bond dissociation energy (BDE) is a central parameter in chemistry—knowing the bond strengths in a molecule assists chemists in designing reactions and estimating thermochemistry. The intrinsic BDEs, i.e., devoid of solvent influence, must be measured by gas-phase experiments. There exists a variety of suitable experimental methods,¹ thanks to which many experimental gas-phase BDE values are easily accessible in handbooks¹,² as well as in the Active Thermochemical Tables (ATcT).³,⁴ Besides the regular use by experimental chemists, such datasets are necessary for benchmarking computational methods; if the computations are sufficiently accurate and efficient, they can be further utilized to train efficient machine learning models for BDE prediction, such as ALFABET⁵ or BonDNet.⁶ Ideally, these advancements would enable chemists to effortlessly and accurately predict the strength of any bond.

Unfortunately, experimental data of BDEs in medium-to-large compounds of chemical relevance are scarce. To add to the problem, concerns were recently been raised about the reliability of computed interaction energies in large systems (~100 atoms), based on the disagreement between quantum diffusion Monte Carlo and coupled cluster approaches,⁷ both of which are generally considered highly
accurate ab initio methods. Such disagreements necessarily call the reliability of cheaper methods for systems of medium-to-large size into question and demand more experimental BDE data.

Gas-phase BDEs of large compounds cannot generally be measured directly, but are extracted from gas-phase rate data of unimolecular dissociation reactions instead. The rate data can be obtained by mass spectrometry (MS) based methods, which have been used extensively for measuring gas-phase thermochemistry of molecular ions; examples of such methods are threshold collision-induced dissociation (T-CID), threshold photoelectron photoion coincidence spectroscopy (TPEPICO), and blackbody infrared radiative dissociation (BIRD). The threshold energy ($E_0$) of a unimolecular bond dissociation reaction with no reverse barrier, such as most ion-molecule dissociation reactions, is equivalent to the respective BDE at 0 K (Figure 1a). In principle, when a compound has energy equivalent to $E_0$, it can undergo the dissociation reaction. However, the rate of dissociation at energy $E_0$ is generally slow for polyatomic species. The consequence is that $E_0$ is not equivalent to the energy at which the dissociation is detected in the experiment (termed appearance energy, $E_{app}$), because: 1) any rate slower than ~10$^2$ s$^{-1}$ is in competition with radiative decay, and 2) the reaction time in the experiment is limited. The second issue is generally more problematic, with the usual reaction time in CID experiments being on the order of 10$^{-5}$ s. This means that depending on the MS measurement sensitivity and the experimental time scale, the dissociation becomes detectable at a rate around 10$^2$–10$^5$ s$^{-1}$. This difference between $E_0$ and $E_{app}$, termed the kinetic shift, was first described by Chupka and is defined as the excess energy required for detectable dissociation (Figure 1b). The higher the $E_0$ is, and the more degrees of freedom the dissociating ion has, the larger is the kinetic shift. It is due to this kinetic shift that the $E_0$ value cannot be directly accessed and has to be extrapolated from the rate data above $E_{app}$.

![Figure 1. a) Potential energy surface (PES) of a dissociation reaction with no reverse barrier; b) Rate of the reaction as a function of internal energy.](image)

The extrapolation to $E_0$ requires modeling the rate curve $k(E)$ of the unimolecular dissociation reaction, which, under the assumption of statistical behavior of the reactant ions, can be done by statistical rate
theory. The most commonly used theory is one that was developed independently by Rosenstock et al.,\textsuperscript{24} who termed it quasi-equilibrium theory (QET), and Marcus,\textsuperscript{25,26} after whom the theory is now known as the Rice–Ramsperger–Kassel–Marcus (RRKM) theory.\textsuperscript{18} It is based on the following equation:

\[
k(E) = \frac{s N^\dagger(E - E_0)}{h \rho(E)}
\]  

(1)

where \(N^\dagger(E - E_0)\) is the (ro)vibrational sum of states of the transition state (TS), \(\rho(E)\) is the (ro)vibrational density of states of the reactant, \(s\) is the reaction symmetry factor/degeneracy, and \(h\) is Planck’s constant. If the functions \(N^\dagger(E)\) and \(\rho(E)\) are known, then \(E_0\) can be calculated by fitting the theoretical \(k(E)\) to the experimental \(k(E)\).

In 2007, we published a new rate model, intended for fitting T-CID data of large organometallic complexes with the RRKM equation (Equation 1), called L-CID (ligand collision-induced dissociation).\textsuperscript{27} Since then, we have determined BDEs in many compounds with the L-CID program, ranging from species relevant for catalysis,\textsuperscript{28–36} as well as for basic research.\textsuperscript{37} Such studies by the Chen group are generally accompanied by quantum chemical computations to find methods that can accurately reproduce experimental bond strengths. However, on several occasions, we have found large discrepancies of up to approximately 1 eV that we have yet to solve as no feasible computational method had provided agreement.\textsuperscript{28,29,36–40} The problem has also attracted the attention of the quantum chemistry community, specifically in trying to understand and solve the discrepancies in our WCCR10 benchmark set.\textsuperscript{41–43} In principle, the errors may reside in either the experiment or the computation, or both. On the one hand, as the experimental values depend on the physical model used in fitting the data, the reliability of the relatively novel L-CID model could be an issue. On the other hand, modern computational chemistry necessarily involves a set of approximations, and we have previously speculated about the possible sources of inaccuracy in the computational methods that one can apply to systems as large as those we routinely study.\textsuperscript{37,44,45} However, none of the speculated origins of the error have been unequivocally confirmed, given the difficulty, experimental and computational, in dealing with molecules with up to 200 atoms.

We recently investigated the reliability of the L-CID scheme with an experimental dataset of cation-linked pyridine dimer dissociations, designed to be structurally well-defined and well-behaved.\textsuperscript{45} Herein, we report a theoretical assessment of the performance of the L-CID rate model by comparison to well-established statistical rate models within the RRKM framework. We compare the performance of the models on our recently published dataset of reaction cross-sections as well as a previously published dataset of microcanonical rates. The latter dataset has been used in a comparison of rate models by Troe, Baer and coworkers,\textsuperscript{46} which we use as a benchmark for our implementation of the rate models.
Methods

Study design

Our main goal is to evaluate the performance of the rate models for fitting reaction cross-sections from T-CID measurements. However, as reaction cross-sections require a simulation of the experimental collision conditions (see Section S1.3 of the SI), they are not as sensitive to the quality of the rate model as pure microcanonical rates would be. Consequently, we decided to evaluate the performance of the rate models on reproducing two kinds of rate data: microcanonical rates from Troe and Baer and reaction cross-sections previously measured by us. The former allows for a direct comparison of the rate models, and the latter is pertinent to our experiments.

We chose the gas-phase dissociation rates of halobenzene radical cations (Scheme 1a) as our test set of microcanonical rates, measured by Troe, Baer and coworkers using TPEPICO. TPEPICO is a valuable method for testing the performance of statistical rate models, precisely because, unlike T-CID, it can measure gas-phase microcanonical rates. A further advantage of this data set is that Troe, Baer and coworkers themselves used the data to compare different rate models, providing a possibility to evaluate our implementation of the models and the fitting procedure. This data set also enabled the investigation of the effect of anharmonicity on these rates, as the halobenzene radical cations are simple and small enough for reliable vibrational perturbation theory calculations (i.e., they lack “floppy” modes or low-barrier rotors, which generally complicate such calculations).

Although TPEPICO is a superior experiment for accessing unimolecular kinetic data, it has several limitations, e.g., the size of compounds to which it can be applied. From that perspective, T-CID is a relatively simpler alternative that has much less stringent demands on the compounds; however, it does not directly measure the microcanonical rates, but reaction cross-sections instead. To compare the performance of the rate models with reaction cross-sections, we required a set of dissociation reactions with varying $E_0$ values and molecular sizes, to test different scenarios of the kinetic shift effect. In addition, we wanted well-behaving dissociation reactions to keep the comparisons free from complications, such as ambiguous reactant structures recently observed by us. We chose a dataset of four cation-linked pyridine dimer dissociations (Scheme 1b), recently reported by us. We selected the smallest and the largest well-behaving pyridine moieties from that study to sample different numbers of degrees of freedom, with both H+ and Ag+ linkers to sample two different $E_0$ regions; in other words, we selected cases with varying kinetic shifts.
a) Halobenzene radical cation dissociations for microcanonical rates

\[
\text{PhX}^+: X = \text{Cl, Br, I}
\]

b) Cation-linked pyridine dimer dissociations for reaction cross-sections

\[
\text{1-M: } R = \text{H, } M = \text{H, Ag} \\
\text{2-M: } R = \text{Ad, } M = \text{H, Ag}
\]

Scheme 1. The investigated dissociation reactions of: a) halobenzene radical cations, b) cation-linked pyridine dimers.

**Statistical rate models**

We simulated the microcanonical rates and reaction cross-sections with L-CID and four conventional statistical rate models: phase space limit (PSL),\(^{49}\) rigid activated complex RRKM (RAC-RRKM), simplified statistical adiabatic channel model (SSACM),\(^{50,51}\) and microcanonical variational transition state theory (μVTST).\(^{52-54}\) We note that our terminology for RAC-RRKM and SSACM is consistent with the common terminology in the PEPICO community,\(^{13}\) including the study by Troe, Baer and coworkers, to which we compare our results.\(^{46}\) We use a slightly modified nomenclature for the other two conventional models: μVTST instead of VTST to clarify that we use the microcanonical variant of the theory; PSL instead of PST (phase space theory) to emphasize that we use the limiting case model as described by Armentrout and coworkers,\(^{49}\) not PST, which explicitly conserves angular momentum.\(^{55-57}\) A more detailed description of the rate models and our implementation is provided in Section S1 of the SI.

The Python code used for the simulations (named *unikin*) is provided in the SI and is available on GitLab.\(^{58}\) The specific descriptions of the simulations for the systems in this study are provided in the respective Results sections. The plots and data from all the fitting results are provided in the SI.

**Structure optimizations, scans, and frequency calculations**

For testing the available computational methods for the purpose of our work, we optimized the structures and calculated the harmonic frequencies with density functional theory (DFT) as well as semi-empirical methods. Grimme’s extended tight-binding methods\(^{59}\) and several DFT functionals from different functional families were tested (see Section S2 of the SI). As the tested methods provided generally very similar results with the microcanonical rate data of halobenzene radical dissociations,
we decided to use BP86$^{60,61}$ for the majority of the calculations in this work, for its balance between efficiency and robustness. We also note that it has been suggested that BP86 harmonic frequencies are a good approximation to the “true” frequencies,$^{62}$ motivating its use for evaluating the vibrational densities of states.

The DFT calculations were performed with ORCA version 4.2.0.$^{63-65}$ The def2-TZVP basis set$^{66}$ was used together with the Stuttgart effective core potentials on the heavy atoms (Ag,$I$). The London dispersion interaction was included by adding Grimme’s D3 correction$^{69}$ with Becke-Johnson damping.$^{70}$ The resolution of identity approximation$^{71}$ was used for the Coulomb integrals with Weigend’s auxiliary basis set;$^{72}$ the integration grid was set with the Grid5 keyword. The optimizations were performed with ORCA’s tight optimization criteria (TightOpt). For open-shell systems (halobenzene radical cations), the unrestricted Kohn-Sham formalism was used. All other settings were kept at ORCA’s default.

The dissociation pathways for μVTST and RAC-RRKM models were calculated by relaxed surface scans. The dissociating bonds were scanned to 10 Å longer than their minimum-energy distance in ~0.5 Å steps. Due to convergence issues in the self-consistent field (SCF) procedure with the scans, the following measures were found to enable successful calculations: SlowConv keyword was invoked together with setting the maximum number of SCF iterations to 500; the rate of calculation of the full Fock matrix was set to every five iterations; and a temperature of 500 K was applied for electronic smearing. For consistency, these settings were also applied to the minimum energy structure (reactant) calculations used together with the scan data. We note that these convergence issues hint at the expected multi-reference problem in the electronic structure of a stretched covalent bond. However, because the stretching frequency of the stretched bond itself was not used in further calculations (it is a translational transitional mode), we expect that the most severe error was probably (at least partially) localized in the least relevant vibration. Therefore, we assumed the electronic smearing approach to be accurate enough, enabling us to calculate the relevant structural and vibrational parameters along the dissociation coordinate.

To calculate the anharmonic density of states, vibrational perturbation theory (VPT2)$^{73,74}$ calculations were performed with Gaussian 09 revision D 01.$^{75}$ We used BP86 with the same dispersion correction and basis set as above. As it has been suggested that VPT2 calculations give more accurate results with B3LYP$^{76-79}$ than BP86,$^{62}$ we also performed B3LYP VPT2 calculations for comparison. As the results were comparable, we only show BP86 results here; B3LYP results are provided in the Section S2.2 of the SI. These calculations were performed only on equilibrium structures as VPT2 is not applicable with imaginary frequencies. The integration grid was set to ultrafine and the structures were optimized with Gaussian’s VeryTight convergence criteria.
All the optimized structures and the respective frequencies are provided in the SI.

**Density and sum of states**

The density and sum of states were calculated with the Multiwell program suite. For the harmonic functions, the densum module was used on the harmonic frequencies obtained from the calculations as described above. For the anharmonic functions, the paradensum module, which utilizes the Wang-Landau algorithm, was used with one walker, 70% overlap and 95% flatness settings. The density and sum of states were calculated with an energy step of 10 cm⁻¹. The frequencies and the X-matrix for paradensum were obtained from the Gaussian VPT2 calculations. The mominert module of Multiwell was used to calculate the rotational constants from the atomic coordinates. The “hrb” hindered rotor model was used for treating hindered rotors; the rotational barriers were calculated with BP86-D3BJ/def2-TZVP by optimizing the TS structure of the rotation.

Representative input files for all the calculations can be found in Section S3 of the SI.

**Results**

**Fitting microcanonical rates**

To evaluate the performance of the rate models with microcanonical rates and to test our code, we performed our fitting procedure on to the gas-phase dissociation rates of halobenzene radical cations (Scheme 1a), measured by Troe, Baer and coworkers with TPEPICO over an impressive range of approximately 10³–10⁷ s⁻¹. In the original publication, the data were carefully compared to other existing experimental data as well as fitted with the four rate models of interest (RAC-RRKM, PSL, μVTST, and SSACM), providing a good data set for comparison. For PSL and SSACM, we closely followed the methodology of Troe, Baer and coworkers. For RAC-RRKM and μVTST, our approach differs from the one in the original publication, in which scaling factors and approximate analytical functions were used. We took a more explicit approach by calculating the frequencies at several points along the reaction coordinate. Details about the set-up of these calculations are provided in Section S1.5 of the SI. In addition to the four conventional models, we fit the data with the three-parameter L-CID rate model. We note that although in the conventional L-CID program the ν_eff fitting parameter has a maximal value of 1000 cm⁻¹, we increased it to 1500 cm⁻¹ for halobenzenes (see Section S2.3 of the SI for further details). The fitted rate curves are visualized in Figure 2a.

The effect of anharmonicity was investigated by fitting the PSL and SSACM rates with the harmonic and anharmonic density and sum of states, using either the Beyer-Swinehart or Wang-Landau algorithm, respectively (see Methods). Only the PSL and SSACM rates were calculated, because VPT2 cannot be used on imaginary frequencies, excluding RAC-RRKM and μVTST. The fitted rate curves are visualized in Figure 2b.
The effect of the external rotors was studied on the RAC-RRKM and $\mu$VTST rate models, in which the TS, and therefore the respective rotational constants, are structurally well-defined. Details about the set-up of these simulations are provided in Section S1.5 of the SI. The fitted rate curves for the $J$-rotor and the $K$-rotor look very similar to the ones without the rotors and are shown in Figure S2 of the SI.

All the fits are visualized in Figure 2, the $E_0$ values are given in Table 1, and the additional fit parameters for SSACM and L-CID in Table 2. To assess the statistical accuracy, we computed the population standard deviation, $\sigma$, of the calculated log($k$) values with respect to the reference. We additionally estimated the population standard deviation for the reference data itself by averaging the asymmetric error bounds of each log($k$) data point and computed the average of these averages. We assume these values to roughly correspond to the $2\sigma$ uncertainty. For comparison, we report our calculated $2\sigma$ values (Table 1), which for a normal distribution approximately correspond to a 95% confidence interval.

Figure 2. Comparison of: (a) all the “plain” rate models for halobenzene radical cation dissociation; (b) PSL and SSACM rates calculated by the harmonic and anharmonic (anh) models; (c) zoomed-in region of (a); (d) zoomed-in region of (b). All rates were calculated based on BP86-D3BJ/def2-TZVP data.
Table 1. Extrapolated $E_0$ values and the $2\sigma$ values of $\log(k)$ for the halobenzene radical cation dissociations with different rate models. 

<table>
<thead>
<tr>
<th>Rate model</th>
<th>PhCl$^+$</th>
<th>PhBr$^+$</th>
<th>PhI$^+$</th>
<th>PhCl$^+$</th>
<th>PhBr$^+$</th>
<th>PhI$^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAC-RRKM</td>
<td>3.253</td>
<td>2.664</td>
<td>2.245</td>
<td>0.17</td>
<td>0.21</td>
<td>0.40</td>
</tr>
<tr>
<td>RAC-RRKM ($J$)</td>
<td>3.241</td>
<td>2.662</td>
<td>2.238</td>
<td>0.25</td>
<td>0.27</td>
<td>0.30</td>
</tr>
<tr>
<td>RAC-RRKM ($K$)</td>
<td>3.226</td>
<td>2.643</td>
<td>2.227</td>
<td>0.16</td>
<td>0.21</td>
<td>0.42</td>
</tr>
<tr>
<td>$\mu$VTST</td>
<td>3.384</td>
<td>2.798</td>
<td>2.380</td>
<td>0.17</td>
<td>0.21</td>
<td>0.39</td>
</tr>
<tr>
<td>$\mu$VTST ($J$)</td>
<td>3.359</td>
<td>2.784</td>
<td>2.362</td>
<td>0.23</td>
<td>0.25</td>
<td>0.33</td>
</tr>
<tr>
<td>$\mu$VTST ($K$)</td>
<td>3.358</td>
<td>2.775</td>
<td>2.359</td>
<td>0.16</td>
<td>0.21</td>
<td>0.42</td>
</tr>
<tr>
<td>PSL</td>
<td>3.440</td>
<td>2.845</td>
<td>2.413</td>
<td>0.45</td>
<td>0.52</td>
<td>0.15</td>
</tr>
<tr>
<td>PSL (anh)$^b$</td>
<td>3.321</td>
<td>2.827</td>
<td>2.390</td>
<td>0.13</td>
<td>0.37</td>
<td>0.08</td>
</tr>
<tr>
<td>SSACM</td>
<td>3.337</td>
<td>2.780</td>
<td>2.399</td>
<td>0.10</td>
<td>0.15</td>
<td>0.07</td>
</tr>
<tr>
<td>SSACM (anh)$^b$</td>
<td>3.303</td>
<td>2.777</td>
<td>2.390</td>
<td>0.10</td>
<td>0.15</td>
<td>0.08</td>
</tr>
<tr>
<td>L-CID</td>
<td>3.500</td>
<td>2.849</td>
<td>2.365</td>
<td>0.14</td>
<td>0.12</td>
<td>0.10</td>
</tr>
</tbody>
</table>

| Reference$^c$    | 3.374    | 2.815    | 2.414   | 0.14     | 0.11     | 0.12    |

$^a$ Based on BP86-D3BJ/def2-TZVP data. $^b$ Density and sum of states from paradensum. $^c$ The $E_0$ values are from ATcT version 1.22p.$^{82}$ the $2\sigma$ values were calculated based on data in reference 46.

**Table 2.** Additional fit parameter values for the SSACM and L-CID models for the halobenzene radical cation dissociations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>PhCl$^+$</th>
<th>PhBr$^+$</th>
<th>PhI$^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSACM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c$ / meV</td>
<td>68</td>
<td>88</td>
<td>261</td>
</tr>
<tr>
<td>$c$ (anh)$^a$ / meV</td>
<td>577</td>
<td>144</td>
<td>$&gt;10^3$</td>
</tr>
<tr>
<td>$c$ (ref)$^b$ / meV</td>
<td>71</td>
<td>77</td>
<td>194</td>
</tr>
<tr>
<td>L-CID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$v_{eff}$ / cm$^{-1}$</td>
<td>$1332 \pm 2$</td>
<td>$1105 \pm 1$</td>
<td>$873 \pm 1$</td>
</tr>
<tr>
<td>$\alpha'$ / cm$^{-1}$</td>
<td>$6495 \pm 3$</td>
<td>$6499 \pm 1$</td>
<td>$6499 \pm 1$</td>
</tr>
</tbody>
</table>

$^a$ Density and sum of states from paradensum. $^b$ From reference 46.

**Fitting reaction cross-sections**

To compare the rate models with well-behaving dissociation reactions of larger ions, we used the reaction cross-section data from our recent set of cation-linked pyridine dimer dissociations, where the linking cation was either a proton or Ag(I) (Scheme 1b).$^{45}$ We chose the smallest and the largest pyridine moieties from that study as bookends to the number of degrees of freedom. In addition, we note that studying compounds with iso-propyl or tert-butyl groups, present in the other dimers of our previous study, would complicate the theoretical treatment of the internal rotors and could thereby introduce
additional model-dependencies and sources of error, as had been shown for the *tert*-butyl radical.\(^{83}\) This issue would be common to all the methods studied here and would not differentiate between them.

The fitting procedure was performed analogously to the halobenzene data fitting, with the exception that reaction cross-sections, not rates, were fitted. The reaction cross-sections were simulated with a Monte Carlo procedure (see Section S1.3 of the SI). The set-up of the rate models and the cross-section simulations is described in Section S1.6 of the SI.

The effect of the external *J*- and *K*-rotors was calculated analogously to the halobenzenes case (see Section S1.5 of the SI). To assess the influence of internal rotors, the rotation of the pyridines along the pyridine-cation-pyridine bond and the rotations of the adamantyl groups in family 2 were treated as harmonic vibrations as well as hindered rotors (HR), as implemented in the densum module of Multiwell. The fitted cross-sections with the hindered rotors as well as with the external rotors look very similar to the ones without the rotors and are shown in Section S4.2 of the SI.

Figure 3. Example reaction cross-section fits of all the “plain” rate models for the dissociation in: (a) compound family 1, and (b) compound family 2. The reference data were normalized based on the maximum value of smoothened data. All cross-sections were calculated based on BP86-D3BJ/def2-TZVP data.
Table 3. Extrapolated $E_0$ values (in eV) for the dissociation of cation-linked pyridine dimers with different rate models. $^a$

<table>
<thead>
<tr>
<th>Rate model</th>
<th>1-H</th>
<th>1-Ag</th>
<th>2-H</th>
<th>2-Ag</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAC-RRKM</td>
<td>0.98</td>
<td>1.78</td>
<td>1.20</td>
<td>1.95</td>
</tr>
<tr>
<td>RAC-RRKM ($J$)</td>
<td>0.96</td>
<td>1.78</td>
<td>1.21</td>
<td>1.98</td>
</tr>
<tr>
<td>RAC-RRKM ($K$)</td>
<td>0.97</td>
<td>1.77</td>
<td>1.19</td>
<td>1.96</td>
</tr>
<tr>
<td>$\mu$VTST</td>
<td>1.05</td>
<td>1.89</td>
<td>1.21</td>
<td>1.87</td>
</tr>
<tr>
<td>$\mu$VTST ($J$)</td>
<td>1.04</td>
<td>1.89</td>
<td>1.22</td>
<td>1.88</td>
</tr>
<tr>
<td>$\mu$VTST ($K$)</td>
<td>1.07</td>
<td>1.88</td>
<td>1.20</td>
<td>1.87</td>
</tr>
<tr>
<td>PSL</td>
<td>1.03</td>
<td>2.01</td>
<td>1.30</td>
<td>2.07</td>
</tr>
<tr>
<td>PSL (HR)</td>
<td>1.03</td>
<td>2.02</td>
<td>1.31</td>
<td>2.09</td>
</tr>
<tr>
<td>SSACM</td>
<td>1.03</td>
<td>2.00</td>
<td>1.30</td>
<td>2.06</td>
</tr>
<tr>
<td>SSACM (HR)</td>
<td>1.03</td>
<td>2.02</td>
<td>1.30</td>
<td>2.09</td>
</tr>
<tr>
<td>L-CID</td>
<td>1.18</td>
<td>1.91</td>
<td>1.40</td>
<td>1.84</td>
</tr>
<tr>
<td>Reference (L-CID)$^b$</td>
<td>1.13</td>
<td>1.93</td>
<td>1.46</td>
<td>1.93</td>
</tr>
</tbody>
</table>

$^a$ Based on BP86-D3BJ/def2-TZVP data. $^b$ From reference 45 with the production version of L-CID; the value for 2-Ag was updated. $^{54}$ The uncertainty in the values was estimated to be approximately 0.1 eV.

Table 4. Additional fit parameter values for the SSACM and L-CID models for the cation-linked dimer dissociations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>1-H</th>
<th>1-Ag</th>
<th>2-H</th>
<th>2-Ag</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSACM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c$ / meV</td>
<td>$&gt; 10^3$</td>
<td>$&gt; 10^3$</td>
<td>$&gt; 10^3$</td>
<td>$&gt; 10^3$</td>
</tr>
<tr>
<td>$c$ (HR) / meV</td>
<td>$&gt; 10^3$</td>
<td>$&gt; 10^3$</td>
<td>$&gt; 10^3$</td>
<td>$&gt; 10^3$</td>
</tr>
<tr>
<td>L-CID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$v_{eff}$ / cm$^{-1}$</td>
<td>515 ± 7</td>
<td>913 ± 41</td>
<td>506 ± 3</td>
<td>508 ± 5</td>
</tr>
<tr>
<td>$\alpha'$ / cm$^{-1}$</td>
<td>5750 ± 246</td>
<td>5887 ± 288</td>
<td>5280 ± 146</td>
<td>5567 ± 312</td>
</tr>
<tr>
<td>Reference$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$v_{eff}$ / cm$^{-1}$</td>
<td>581 ± 50</td>
<td>818 ± 72</td>
<td>621 ± 189</td>
<td>567 ± 33</td>
</tr>
<tr>
<td>$\alpha'$ / cm$^{-1}$</td>
<td>5805 ± 462</td>
<td>5688 ± 326</td>
<td>5823 ± 401</td>
<td>5867 ± 310</td>
</tr>
</tbody>
</table>

$^a$ From reference 45 with the production version of L-CID; the values for 2-Ag were updated. $^{54}$

**Discussion**

**Background**

Fitting the RRKM rate (Equation 1) to the experimental rate for obtaining the $E_0$ value requires the $N(E)$ and $\rho(E)$ functions to be known. In other words, obtaining accurate $E_0$ from experimental data is ultimately dependent on having reliable $N(E)$ and $\rho(E)$. As these functions are not generally
experimentally accessible, they have to be computed. This involves two choices: 1) how to treat the TS to evaluate $N^\ell(E)$, and 2) which method to use for the computation of $N^\ell(E)$ and $\rho(E)$.

The first choice relates to how the transitional modes, i.e., vibrations of the reactant that turn into rotations of the products, are treated at the TS. There are four commonly used methods (see Section S1.1 of the SI). PSL and RAC-RRKM are the extreme cases: the former treats the transitional modes as rotations of the products (a loose TS), the latter as vibrations of the TS (a tight TS). SSACM and $\mu$VTST both model the transitional modes as functions of energy, under the assumption that the TS becomes “tighter” at higher energy, and, therefore, the transitional modes behave less as free rotations. SSACM uses a rigidity factor with an additional fitting parameter ($c$) to dampen the effect of the transitional rotors to $N^\ell(E)$ at high energy; $\mu$VTST uses several (or even continuous) TS definitions with different transitional modes to calculate the lowest rate at each energy, effectively mimicking the change from a loose TS to a tighter one. Troe has analytically shown that $\mu$VTST rates for individual $J$-rotor levels are identical to the original SACM rates for model charge-dipole systems, and that in general $\mu$VTST can be considered a simplified version of SACM. Because the $N^\ell(E)$ of a rotor increases faster than that of an oscillator (Figure 4), the log($k$) curves for TS models with rotational transitional modes have a higher curvature at low energy. This curvature generally leads to higher $E_0$ values from fitting, and thereby models that treat the TS as loose return higher $E_0$ values, in general. Indeed, it has been shown by several studies that if applied on the same data, RAC-RRKM gives generally lower bound $E_0$ values, PSL higher bound ones, and $\mu$VTST and SSACM give values between the two extremes.

![Figure 4. Comparison of the sum of states function for a representative transitional vibrational and rotational degree of freedom computed with the Beyer-Swinehart algorithm and Equation S7, respectively.](image)

The second choice in the application of the RRKM theory is related to the method for calculating $N^\ell(E)$ and $\rho(E)$. The rotational and vibrational parts are generally calculated separately and then combined by
The rotational part of external as well as internal rotors are often calculated based on classical equations (e.g., Equations S7 and S8). The vibrational part can be treated on several levels of complexity. For a system of independent quantum harmonic oscillators, the most accurate and efficient method is the Beyer-Swinehart direct count algorithm. However, the harmonic approximation is non-ideal, because the calculated rates have been shown to be dependent on anharmonicities of the modes and specifically mode coupling. Unfortunately, there is, as yet, no one “gold standard” method for anharmonic \( \rho(E) \) calculation, but the two most common methods are: a) the Beyer-Swinehart-Stein-Rabinovitch algorithm, which calculates the density of states similarly to Beyer-Swinehart and can only treat uncoupled anharmonic oscillators; b) Monte Carlo based methods for coupled anharmonic oscillators, most notably the Wang-Landau method.

Alternative options that have been reported include a molecular dynamics based method, inverse Laplace transform of the anharmonic partition function, and empirical analytical formulas. As the calculation of anharmonic density of states is complicated as well as computationally expensive, the most common method for \( N^r(E) \) and \( \rho(E) \) calculation for routine applications is still the Beyer-Swinehart algorithm with harmonic frequencies from a quantum chemical calculation.

A further aspect to consider in RRKM rate calculations are the external rotors. The reactant and TS are usually treated as symmetric tops, with separable rotors. This means that they have an external 1D (K-)rotor and a 2D (J-)rotor (cf. Figure S1). The K-rotor is usually assumed to be active, i.e., it can exchange energy with other internal modes via, for example, Coriolis coupling. However, it has been suggested that for molecules with a large vibrational density of states—such as butene—the inclusion K-rotor does not have a significant impact on the rate, and it is therefore often ignored. The J-rotor is usually assumed to be adiabatic, such that the rotational quantum number (J) does not change during the dissociation. This leads to the so-called centrifugal effect, well presented in a review by Waage and Rabinovitch. In short, as the rotational energy is proportional to the moment of inertia and the moment of inertia changes during the dissociation, it follows that for different values of J, different amount of total energy \( E \) is “trapped” in the 2D rotation of the reactant and the TS. This tends to accelerate the rate at high energies, but decelerate at low energies, because the energy trapped in the rotor is not available for the dissociation.

Two commonly used programs that apply the RRKM theory to extract \( E_0 \) values from T-CID data are: CRUNCH by Armentrout and L-CID by Chen. CRUNCH calculates the rate by using the Beyer-Swinehart algorithm on calculated frequencies for the sum and density of states in the PSL framework. The L-CID fitting scheme, published in 2007 and built for deconvolution of the same kind of data as CRUNCH, made three innovative developments: (a) new, physically motivated, empirical density of states functions were derived (Equations S5 and S6), (b) the collision simulation was made more physically realistic by properly treating the centrifugal barrier with the model of Giomouisis and
Stevenson, and (c) Monte Carlo simulation for explicit energy and impact parameter distributions was introduced together with a genetic algorithm for the fitting procedure. The only structural information the user has to provide is the number of degrees of freedom, the number of internal rotors, and whether the TS is tight or loose; beyond that, the model is “structure-agnostic”. In other words, the two choices required for applying the RRKM equation, as mentioned in the introduction of this section, are handled by a “tight/loose switch” and the empirical equations of L-CID (point (a) above). This seemingly radical approximation offers a significantly simplified approach for large systems for which the structural complexity can become as severe as involving conformational and configurational isomers of the reactant and the TS. In such cases, rigorous application of the conventional methods that require explicit structure-dependent data can quickly become a complicated task; L-CID absorbs such structural complexity into the fitting parameters $v_{eff}$ and $\alpha'$ (cf. Equations S5 and S6). It is imperative to question whether this benefit of simplicity and extended applicability is worth the risk of potential decrease in accuracy. In the original publication, the L-CID program was shown to reproduce $E_0$ values calculated with CRUNCH for a selected set of reactions (of ions in a size range where both programs could be used) within the margin of uncertainty, implying that the approximations in L-CID do not adversely affect the accuracy. However, we have found several problematic disagreements in a range of about 0.3–1.1 eV between our subsequently measured BDEs for much larger molecular ions and quantum chemical calculations, with calculations generally returning higher BDEs, notably in the WCCR10 benchmark set, for which even the generally accurate $ab initio$ methods have not provided satisfactory agreement. In order to assess the validity of our experimentally derived BDEs, we decided to compare the L-CID rate model to other conventionally employed rate models; if the L-CID results agree with those of the conventional models, the experimental values should be valid, insofar as the experimental procedure itself is valid and RRKM theory is appropriate for these unimolecular dissociation reactions. In addition, we investigated the influence of anharmonicity and external rotors; as the L-CID rate model was derived and parametrized without explicitly considering these effects, it could be an indicator of possible issues in the parametrization if their magnitude were to be large.

**Comparison with microcanonical rates**

Our extrapolated $E_0$ values with the four conventional methods (RAC-RRKM, PSL, SSACM, and $\mu$VTST) agree with the values obtained by Troe, Baer and coworkers within the estimated uncertainty margins (Table 1; see Section S2.4 of the SI for an extended comparison of the results here and from reference 46). Consequently, the data shows the usual trend: RAC-RRKM returns the lowest $E_0$ values, PSL the highest, and SSACM and $\mu$VTST return values in between. This result is illustrated in Figure 5, which shows the deviation of the optimized $E_0$ value from the ATcT reference values, and the doubled standard deviation ($2\sigma$) of the log($k$) values, which describes how well the model fits the experimental data. As visible, both SSACM and $\mu$VTST return $E_0$ values close to the reference, but the fits are better
with SSACM. In fact, at least for PhCl\(^+\) and PhI\(^+\), the 2\(\sigma\) value from the SSACM fit is within the estimated average 2\(\sigma\) range of the reference data (0.12)—this is visible by the respective SSACM data points (green) being within the blue shaded area (estimated experimental 2\(\sigma\) range). Whereas PSL also returns values within or near the reference uncertainty range, the quality of the fits is considerably worse, specifically for PhCl\(^+\) and PhBr\(^+\) (2\(\sigma\) = 0.45 and 0.52, respectively), as was also observed by Troe, Baer and coworkers. RAC-RRKM consistently returns too low \(E_0\) values, although the fits at least for PhCl\(^+\) and PhBr\(^+\) are fine, as again was also noted by Troe and Baer. The reproduction of the qualitative trends and the good agreement between the results constitute an essential confirmation of our implementation of the different rate models, and gives confidence that our implementation of the rate models can be used for further analysis of the datasets with reaction cross-sections.

Although the L-CID rate model was developed for fitting T-CID data,\(^{27}\) it can be used equivalently on TPEPICO-derived rates, or any other type of unimolecular rate data. As can be seen in Figure 2, the L-CID rate curves follow those of the other rate models closely and give fits of similar quality to those of SSACM (Table 1 and Figure 5). Consequently, the extrapolated \(E_0\) values are also close to those from the other rate models (Table 1 and Figure 5); in other words, the L-CID rate equation returns rate curves and \(E_0\) values that are effectively equivalent, or at least comparable, to other popular rate models. Specifically, the values for PhCl\(^+\) and PhBr\(^+\) lie closest to the respective PSL values, whereas it is more similar to \(\mu\)VTST for PhI\(^+\). Therefore, on these data, L-CID predicts mid-to-high range \(E_0\) values, distinctly higher than those of RAC-RRKM. This implies that, as designed, L-CID treats the TS as loose, as is also reflected in the \(\alpha'\) values of 6500 cm\(^{-1}\) (Table 2; see reference 27 for a discussion on the meaning of the \(\alpha'\) parameter). The significance of this good agreement lies in that any of the other four investigated rate models require themselves at least some, if not substantial, computational effort before the fitting procedure, but L-CID fitting for reactions as simple as the halobenzene dissociation requires no more additional work beyond counting the number of atoms.
Figure 5. Deviation of $E_0$ from the reference value against the $2\sigma$ standard deviation in log($k$) for halobenzene radical cation dissociation. The vertical (gray) shaded area indicates the range where the deviation is below the average reference uncertainty ($\pm 0.10$ eV, Table 1); the horizontal (blue) shaded area indicates the range where the $2\sigma$ value is below the average estimated $2\sigma$ deviation in the reference log($k$) data (0.12).

The effects of anharmonicity and external rotors on microcanonical rates

We continue by evaluating the effect of anharmonicity on the rate curves and the extrapolated $E_0$ values. As shown in Table 1, using anharmonic densities and sums of states instead of the harmonic Beyer-Swinehart direct count ones leads to slightly lower $E_0$ values, most notably for PhCl$^+$ with PSL (0.12 eV). This is because the anharmonically calculated rate tends to be lower than the harmonic one—this effect has been comprehensively discussed and explained by Hase and coworkers.$^{18,91}$ The explanation relies on two features: a) the anharmonic density and sum of states are larger than the harmonic ones, and this discrepancy generally increases with energy; b) the internal energy of the reactant is higher than the internal energy of the TS. Thereby, in terms of Equation 1, anharmonicity increases the denominator more than the numerator, leading to the observed lowering of the rate. As indicated by the change in the $E_0$ values and visible in Figure 2b, the PSL rate is more affected by the change from harmonic to anharmonic treatment than SSACM. The two SSACM rate curves are probably more alike because the fitting procedure involves the additional fitting parameter $c$, which makes the rate curve more flexible. Indeed, the $c$ parameter values change going from the harmonic to the anharmonic model (Table 2). We note in passing that the drastically high $c$ value (effectively $\infty$) for PhI$^+$ with the anharmonic rate indicates that SSACM converges to the PSL rate, which was also recognized by Troe and Baer with harmonic calculations.$^{46}$ The key takeaway points are that a) the $E_0$ values calculated with the harmonic approximation provide the upper bound of the respective rate model, and b) for SSACM, the harmonic results are likely to be similar to the anharmonic results.
The inclusion of the external adiabatic rotor (J-rotor) induces a similar effect—the extrapolated $E_0$ values are smaller than without the external rotor, although only slightly so (Table 1). This means that the external adiabatic rotor decelerates dissociation. This deceleration has also been experimentally observed by Booze et al.\textsuperscript{103} as well as Kiermeier et al.\textsuperscript{107} However, the effect is small for halobenzenes, as was also assumed by Troe, Baer and coworkers.\textsuperscript{46} The incorporation of the external active rotor (K-rotor) leads to a greater lowering of the $E_0$ values than the J-rotor, but the effect is still only on the order of 1% (Table 1).

It appears that at least for the halobenzene dissociation reaction, both the anharmonic effects and the external rotor effects tend to reduce the $E_0$ value obtained from the fits. However, the changes are quantitatively not very significant—they are mostly in the same order of magnitude as the uncertainty bounds of the reference data (a few tens of meV). Even for the largest absolute change of 0.12 eV in the case of the anharmonic PSL model for PhCl\textsuperscript{+}, the relative change is still below 4% of the $E_0$ value. However, this uncertainty range should be kept in mind when applying the standard harmonic treatment for $E_0$ fitting.

**Comparison with reaction cross-sections**

As discussed in the Background section, the L-CID program introduced changes to: (a) the rate model by deriving empirical density of state functions, (b) the reaction cross-section simulation by properly treating the centrifugal barrier in collisions, and (c) the fitting procedure by using a genetic algorithm with a Monte Carlo simulation. In the previous sections, we showed that the L-CID rate model itself behaves analogously to the established rate models, with or without the inclusion of external rotors or anharmonic effects, thereby supporting the validity of change (a) for small systems. As the Giomousis-Stevenson model\textsuperscript{106} (change b) and the Monte Carlo procedure (change c) for collision simulations, while technically challenging, are not controversial changes from the perspective of principle, we decided to use the L-CID cross-section simulation procedure to compare all the rate models with larger systems. The original L-CID program was designed for the analytical L-CID rate model; in order to treat all the rate models on an equal footing, we rewrote the simulation algorithm such that it uses data arrays from the rate models instead of analytical functions and uses either a differential evolution algorithm or Brent’s method instead of a genetic algorithm for fitting (see details in Section S1.4 of the SI; the source code is provided in the SI). As the reaction cross-section data used in this paper has been previously fit with the production L-CID program,\textsuperscript{45} it is essential to compare the previous results with the ones from the current implementation. As visible in Table 3, the $E_0$ values from the reference L-CID calculations and the L-CID calculations from this work generally agree well, with the most notable deviation of 0.09 eV being in the $E_0$ value of 2-Ag. The similarity of the results is also echoed in the values of the other fit parameters, $v_{\text{eff}}$ and $\alpha'$ (Table 4). Therefore, the conclusions from the following analysis can be expected to apply also more generally to L-CID in its original implementation.
The fitted reaction cross-sections with all the models appear nearly indistinguishable (Figure 3), implying that all the investigated models can, in principle, be applied to CID reaction cross-section fitting with our algorithm. The numerical results from these fits are generally in line with the observations already made with microcanonical rates—the $E_0$ values vary maximally in a 0.2 eV range on average, with RAC-RRKM generally corresponding to the lowest values (except 2-Ag, discussed below) and PSL giving the highest values (Table 3). In other words, the $E_0$ values from different rate models are, with a certain generosity, within our usual estimated uncertainty bounds (approximately ±0.1 eV). We note that the high $c$ values for the SSACM model (Table 4) imply that SSACM does not improve on the PSL model for this data set, as also reflected in the similarity of their $E_0$ values (Table 3).

The effect of the $J$- and $K$-rotors on the $E_0$ values from RAC-RRKM and μVTST models is similar in magnitude to what was observed with microcanonical rate data (around 0.02 eV at most), but it occasionally differs in the direction of the change. Whereas the external rotors reduced the $E_0$ value with halobenzene radicals in each case, the change in the $E_0$ value is less systematic with the reaction cross-sections. For instance, in the case of 2-H, the $E_0$ value from RAC-RRKM with the $J$-rotor is larger by about 0.01 eV than the corresponding value without the $J$-rotor, but smaller by 0.01 eV in the case of the $K$-rotor (Table 3). Relatedly, the change of treating internal rotors (e.g., the rotation of the adamantyl substituents in family 2) as hindered rotors instead of harmonic oscillators in PSL and SSACM models only introduced modest changes in the resulting $E_0$ values, also on the order of 0.01–0.02 eV at most (Table 3). Based on these data, it thus appears that the quantitative effect of external as well as internal rotors is generally within the uncertainty bounds of the fits (approximately 0.1 eV), implying that the simplified treatment of rotors, i.e., neglecting the external ones and treating the internal ones as vibrations, is justified within the accuracy limit of reaction cross-section fitting.

The main objective and challenge of applying statistical rate theories for $E_0$ estimation is extrapolation of the rate data below the appearance energy ($E_{app}$, Figure 1b), or in other words, estimation of the kinetic shift. The larger the kinetic shift is, the more sensitive the final $E_0$ value is to the quality of the rate model. From that point of view, the most challenging system in our data set is 2-Ag; it has the largest expected kinetic shift, as it has the combination of a large number of degrees of freedom (207) as well as a high $E_0$ value (approximately 2 eV). However, based on the values in Table 3, the agreement between the different rate models does not significantly deteriorate, as compared to 1-Ag, which has a similar $E_0$ value, but a lower number of degrees of freedom (63); or to 2-H, which has the same number of degrees of freedom, but a lower $E_0$ value (1.2–1.4 eV)—in all these cases, the different rate models return $E_0$ values within a 0.2 eV range, and the specific values follow the generally known trends between the models (see Background). The abovementioned exception of RAC-RRKM for 2-Ag is likely due to the inherent arbitrariness of choosing a specific TS structure (see Section S6.2 of the SI
for further discussion). However, with or without accounting for the sensitive RAC-RRKM results, the rate models appear to estimate the $E_0$ values within a more or less uniform spread of 0.2 eV, implying a practically consistent treatment of the kinetic shift within that uncertainty range.

**Why does L-CID work?**

As mentioned above, the quality of the RRKM rate function is directly dependent on the quality of the density and sum of states functions. It follows that these functions must be carefully designed for a rate model to be accurate and applicable. Indeed, the density of states functions as defined for the reactant and the TS by Equations S5 and S6, respectively, are the most critical as well as inventive parts of L-CID. These functions were derived starting from an empirical equation, adapted from Lorquet:108

$$\rho(E) = \exp(\beta \cdot \sqrt{E})$$

(2)

One can recognize that the squared natural logarithm of Equation 2 is a linear function with respect to $E$, with a slope of $\beta^2$, i.e., $[\ln(\rho)]^2 = \beta^2 E$. However, for actual explicitly calculated $\rho(E)$ functions, this is not entirely true—although the functions are indeed linear at high energies, there exists a curved region at low energies (cf. Figure 6). The nonlinear part for the $[\ln(\rho)]^2$ vs $E$ plot covers the regions where $E$, as compared to the magnitude of a vibrational quantum, is not large, i.e., where the $\rho(E)$ function is “grainy.” To account for that, the energy $E$ was replaced with $\varepsilon = [E^2/(E + \alpha)]^{0.5}$ such that $\varepsilon = E$ at $E = 0$ and when $E$ is large. At intermediate energies, the value of the $\alpha$ parameter defines the curvature. The new equation therefore is:

$$\rho(E) = \exp\left(\beta \cdot \sqrt{\frac{E^2}{E + \alpha}}\right)$$

(3)

The $\alpha$ and $\beta$ parameters were parametrized against computational $\rho(E)$ functions calculated with the Beyer-Swinehart algorithm, using: 1) the linear region of $[\ln(\rho)]^2$ at high energies to parameterize $\beta$ with respect to the number of degrees of freedom and a single effective frequency, $\nu_{\text{eff}}$, and 2) the nonlinear low energy region to parameterize $\alpha$ with respect to the number of internal rotors and $\nu_{\text{eff}}$. This leads to Equation S5. More details on the derivation as well as the derivation of Equation S6 can be found in the supporting information of the original L-CID publication.27

One of the main assumptions of the density of states function of L-CID is that for the “true” (ro)vibrational density of states, the $[\ln(\rho)]^2$ function is linear at high energies. We investigated this by plotting the $[\ln(\rho)]^2$ function of the density of states used in our rate calculations (Figure 6). As also shown for a set of test cases in the original publication of L-CID,27 the $[\ln(\rho)]^2$ function is visibly linear at high energies over an impressive range of energy, for both the harmonic and anharmonic functions (calculated with the Beyer-Swinehart and Wang-Landau methods, respectively), reconfirming one of
the underlying assumptions of L-CID. A linear region in data can always be fit by a linear function with two parameters (slope and intercept). In terms of Equation 3, the two parameters are \( \alpha \) and \( \beta \), where \( \alpha \) controls the curvature of \( [\ln(\rho)]^2 \) at low energies, which effectively translates into the intercept of the linear part at high energies, and \( \beta \) controls the slope of the linear part. In L-CID, \( \alpha \) and \( \beta \) are further expressed by structure-related parameters (cf. Equations S5 and S6), of which a single effective frequency, \( \nu_{\text{eff}} \), and the TS “looseness” parameter, \( \alpha' \), are used as fitting parameters.

Although the shape of the computed \( [\ln(\rho)]^2 \) functions is amenable to accurate fitting by the L-CID equations, Figure 6 shows that the agreement of the \( [\ln(\rho)]^2 \) function with explicit calculations and from L-CID varies between the compounds, with deviations increasing in the series PhCl\(^+\) < PhBr\(^+\) < PhI\(^+\). Specifically, whereas the explicitly calculated functions change little between the compounds, the slope of the linear region increases significantly for the L-CID functions. At first glance, this may seem like a serious flaw, but the good correspondence between the actual rate data from L-CID and the conventional rate models (Table 1 and Figure 3) suggests that these errors do not propagate adversely in the L-CID rate model. In fact, whereas the error in the density and sum of states functions looks the largest for PhI\(^+\) and the smallest for PhCl\(^+\), the opposite trend exists in the extrapolated \( E_0 \) values. The probable reason is that the deviations in \( [\ln(\rho)]^2 \) and \( [\ln(N)]^2 \) are similar (Figure 6)—it is a known feature of the RRKM rate equation (Equation 1) that deviations in the denominator \( (\rho(E)) \) can be partially cancelled by deviations in the numerator \( (N^\dagger(E - E_0)) \). This implies that as long as the ratio between the \( \rho(E) \) and \( N^\dagger(E - E_0) \) is roughly correct, the absolute deviations do not matter much. It is this error cancellation effect that L-CID greatly benefits from—although the density and sum of state functions are parametrized in such a manner that they are usually in a reasonable range of values, they are not expected to be generally accurate with respect to the functions calculated from the molecular structure and vibrational frequencies. However, L-CID fitting can also be applied when the molecular structure and the vibrational frequencies are not known or well-defined, as can be the case for large fluxional compounds. Indeed, it is in such cases that the advantage of L-CID best emerges.
Figure 6. Plots of $[\ln(\rho)]^2$ (top) and $[\ln(N)]^2$ (bottom) for halobenzenes. The $[\ln(\rho)]^2$ plots also display zoomed insets of the low energy region, where the function is not linear. The vibrational part of the functions was obtained with the Beyer-Swinehart algorithm for PSL, and the Wang-Landau algorithm for PSL (anh); the rotational part for both cases was calculated with Equation S8. The L-CID data was calculated with Equations S5 and S6.

As explained, the RRKM equation itself benefits from error cancellation, which makes it less sensitive to absolute errors in the state count functions when fitting microcanonical rates. In the case of the reaction cross-section simulation, the density of states function also enters in the generation of the Boltzmann distribution of vibrational energies (Equation S9) and the kinetic energy release distribution (Equation S13). As both of these are probability distributions, they are necessarily normalized; this means that as long as the input density of states has the right shape, its absolute magnitude does not have an influence on the outcome, because of the normalization. As a consequence, the simulated reaction cross-sections are also insensitive to errors in the absolute magnitude of the density of states function.

Conclusions

We investigated the performance of various rate models within the RRKM framework, including our empirical L-CID model, on microcanonical dissociation rates of halobenzene radical cations and CID reaction cross-sections of cation-linked pyridine dimers. The former dataset provided grounds to compare the quality of purely the rate models, whereas the latter allowed us to estimate their merit in conjunction with cross-section simulations and with larger ions. We showed that in all cases, the resulting $E_0$ values agree within a couple of tenths of an eV. In addition, we observed that at least for the studied reactions, an explicit treatment of hindered rotors, external rotors, or anharmonicities, as opposed to the “plain harmonic” treatment, did not bring about significant changes in the resulting $E_0$ values, being generally less than the expected uncertainty of approximately 0.05 to 0.1 eV. Therefore,
the discrepancies between the computed BDE values and the experimental ones are unlikely to stem from the approximations L-CID makes on top of the RRKM theory.

We also briefly elaborated on the reasons for why L-CID works as well as it does, showing that the underlying assumption of the shape of the density of states function holds true for the systems in this study, also when compared with anharmonic functions. Because both the RRKM rate equation and the reaction cross-section simulation are forgiving for errors in the absolute values of the state counts, L-CID should provide reasonably accurate rates and cross-sections, and thereby $E_0$ values, as long as it gets the shapes of the functions right, which is what it was designed to do.

**Associated content**

**Supporting Information**

Description of the investigated rate models and the simulations; computational details of additional calculations; results from fitting with different computational methods for the structure optimization and frequency calculation with halobenzenes; comparison of the conventional rate models with data from Troe, Baer and co-workers; data from fitting the halobenzene rates with different $V_{\text{eff}}$ ranges; example input file templates for the calculations; rate and cross-section plots with accounting for the external and internal rotors; reference microcanonical rates used for fitting; new reference L-CID data for 2-Ag; discussion regarding RAC-RRKM with 2-Ag; data and plots of all the individual fits (PDF)

Rotational constants and frequencies for all the structures with all the used methods (ZIP)

Cartesian coordinates of all the optimized structures with the SCF energy and ZPE values (ZIP)

Reference cross-sections (ZIP)

The Python code used for the simulations and fitting (ZIP)

**Acknowledgments**

We thank Dr. Alexandra Tsybizova for helpful discussions and constructive feedback; Dr. Marek Bot for assistance with the original L-CID code and experimental data; and Felix Fleckenstein for assistance with the original L-CID code and for helpful comments on our Python code. We acknowledge financial support from the ETH Zürich, and the Swiss National Science Foundation.

**References**


(49) We also tested VPT2 calculations for the small cation-linked pyridines. The results appeared unreliable, judged by the large number (several hundred) of warnings regarding unreliable force constants in the log files and erratic predicted anharmonic intensities. We thereby decided not to use these calculations for further analyses.


(82) Ruscic, B.; Bross, D. H. Active Thermochemical Tables (ATcT) values based on ver. 1.122p of the Thermochemical Network (2020); available at ATcT.anl.gov.


(84) We found a minor bug in the original L-CID code that caused an intermittent error in the Boltzmann distribution sampling. We observed errors no larger than 0.1 eV in the final threshold energy values. The values used as the reference data here were obtained from refitting the data with the fixed code (see Section S6.1 of the SI for further details).


TOC Graphic