



# Quasi-harmonic approximation for the thermochemical stability of small proton bound clusters – a theoretical study



Physical & Theoretical Chemistry  
Wuppertal, Germany  
(Institute for Pure and Applied Mass Spectrometry)

Alexander Haack<sup>1</sup>; Walter Wissdorf<sup>1</sup>; Hendrik Kersten<sup>1</sup>; Thorsten Benter<sup>1</sup>

## Introduction

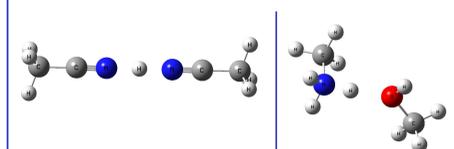
Small proton bound clusters are often observed in atmospheric pressure ionization (API) mass spectrometry. Especially clusters formed from water or other solvents, e.g., from a liquid sample injection, dominate the chemical environment of an API ion source and determine the degree of analyte protonation through the thermodynamic equilibrium of the system.<sup>[1]</sup>

For analytical purposes, it would be of great benefit, if the system at hand, i.e., the dominant cluster species and the analyte, and its thermodynamic equilibrium would be studied in greater detail, e.g. via theoretical methods, to prevent making poor choices regarding solvents / gas phase additives leading to reduced sensitivity.

So far, computational methods describing the thermochemical stability of such clusters and their protonation capabilities have been quite unsatisfactory. Standard harmonic oscillator approximations suffer from the strong anharmonicity of the loosely bound structures and more advanced methods (VPT2, VSCF) are often too demanding.

In this contribution we investigate the thermochemistry of proton bound clusters with the quasi-harmonic approximation (QHA), which can account for anharmonicity effects while still being computationally affordable.

As a case study we analyze two proton bound clusters, i.e., an acetonitrile dimer (ACN)<sub>2</sub>H<sup>+</sup> and a mixed amine methanol cluster [CH<sub>3</sub>NH<sub>3</sub> + CH<sub>3</sub>OH]<sup>+</sup>:



In both cases, a standard geometry optimization and frequency calculation of the cluster and its fragments led to unsatisfactory cluster stability data when compared to experimental data. This is mostly due to the harsh approximation of a harmonic potential to the proton movement potential and treating the internal rotations as vibrations. Additionally, coupling between modes is neglected in the harmonic treatment but is expected to have a significant contribution to the given examples.

## Computational Details

### Software package:

Gaussian16, Revision A.03<sup>[2]</sup> for DFT calculations

GaussView 6.0.16<sup>[3]</sup> for visualization

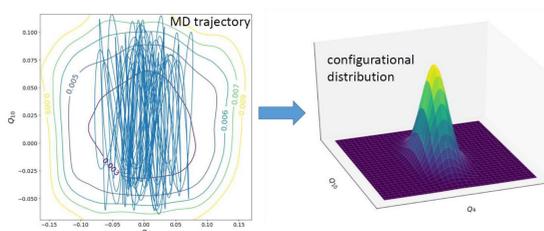
Python scripts for analysis and plotting

### Machine:

Linux-based computer cluster, 4×16-Core CPUs (6282SE AMD Opteron; Advanced Micro Devices GmbH, Dornach, Germany) and 32×16 GB memory

## QHA

- For thermochemical stability ( $\Delta_R G$  values) the partition function  $Z$  of reactants and products is needed
- Vibrational part of  $Z$  is difficult to describe outside the harmonic approximation to the potential energy surface (PES)
- The Quasi-harmonic approximation introduced by Karplus and Kushick<sup>[4]</sup> estimates classical  $Z_{vib}$  from configurational distribution determined by MD simulations:



- The integral over the configurational distribution (called the configuration integral) is proportional to the partition function
- If configurational distribution is fitted to a Gaussian:

$$Z_{vib} = \left( \frac{\sqrt{2\pi k_B T}}{h} \right)^{3N-6} \sqrt{\det(\mathbf{G}^{-1}|_{\mathbf{q}_0})} \sqrt{(2\pi)^{3N-6} \det(\boldsymbol{\sigma})}$$

- With  $\mathbf{G}^{-1}|_{\mathbf{q}_0}$  being the transformation matrix between Cartesian  $\mathbf{x}$  and internal coordinates  $\mathbf{q}$  at the equilibrium geometry  $\mathbf{q}_0$  and  $\boldsymbol{\sigma}$  being the covariance matrix of the Gaussian distribution

### Workflow:

- conduct MD simulation around  $\mathbf{q}_0$  at a certain temperature
- Transform trajectory into internal coordinates  $\mathbf{q}$
- Determine covariance matrix  $\boldsymbol{\sigma}$

## Implementation and special Considerations

### MD Simulation

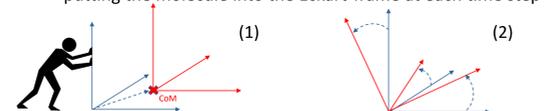
- Geometry optimization and normal mode analysis (both at B3LYP-GD3BJ/def2-TZVPP) for every compound
- Multiple ab initio MD simulations (B3LYP-GD3BJ/def2-SVP) around the respective equilibrium geometry  $\mathbf{q}_0$ , applying  $(3N - 6)k_B T$  kinetic energy to the nuclei
- export  $3N$  Cartesian coordinates and corresponding energies

### Internal Coordinates

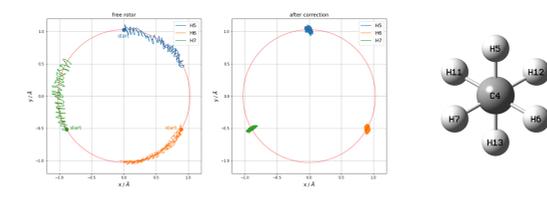
- In principle, all kinds of internal coordinates could be used
- We chose *normal* coordinates  $\mathbf{Q}$  because of their generality, the transformation matrix is obtainable from normal mode analysis, G-matrix is equal to unity matrix

### Eckart-Sayvetz conditions

- Before transformation from Cartesian to normal coordinates is performed, the trajectory is manipulated<sup>[5]</sup>
- Separate out overall translation (1) and rotation (2) by putting the molecule into the Eckart-frame at each time step



- Separate out internal rotations by rotating the moving atoms back around the internal rotation axis

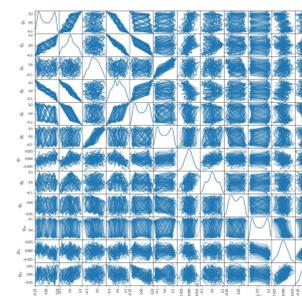


### Transformation and partition functions

- After Eckart treatment, transform (modified) Cartesian coord. to (mass weighted) normal coord.

$$Q_i(t) = \sum_{k=1}^{3N} \frac{\partial Q_i}{\partial x_k} \sqrt{m_k} (x_k(t) - x_k(0))$$

- Determine covariance  $\boldsymbol{\sigma}$  matrix (dimensionality of  $3N - 6 - N_{ir}$ )
- Partition functions for internal rotation (moment of inertia around int. rot. axis needed) and overall translation and rotation are calculated using standard formula
- right:** trajectory of ACN depicted as correlation matrix of the  $Q_i$

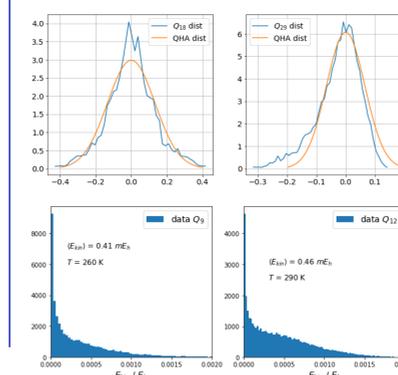


### Correction

- From the ab initio and QHA potentials ( $V$  and  $V_{QHA}$ ) a correction can be calculated:

$$\langle \dots \rangle \text{ is the average over the MD trajectory} \quad Z_{vib}^{QHA,corr} = Z_{vib}^{QHA} / \left\langle \exp \left( (V - V_{QHA}) / k_B T \right) \right\rangle$$

### Diagnostics



- Along most modes, the configurational distribution is well described by a Gaussian (top figures)
- Sometimes the distribution is asymmetric (due to an anharmonic potential); the Gaussian is thus broadened
- Sampling density must be high enough to ensure a converged distribution
- The kinetic energy in every mode changes over the MD trajectory; higher energies are less likely (bottom figures)
- This models the energy distribution of a canonical ensemble; corresponding Boltzmann plots show a near-linear correlation
- Mean kinetic energy in each mode is around  $\frac{1}{2} k_B T$

## Results

### (ACN)<sub>2</sub>H<sup>+</sup>

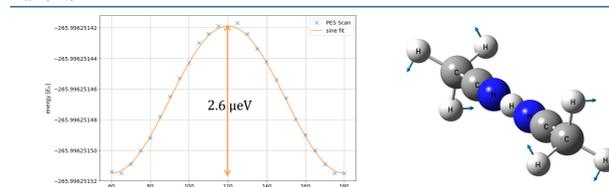
- $Z_{vib}$  differs largely from the harmonic values; this is due to failed description of high frequency modes ( $h\nu \gg k_B T$ ) which should be frozen out; in our treatment, everything is classically and even those modes have  $k_B T$  energy. At very high  $T$  the partition functions actually converge

- The correction factor is very different between the cluster and the fragments, indicating that the loosely bound cluster is more difficult to describe by this method; also indicates the importance of this correction

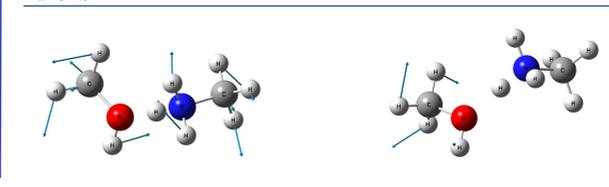
- The internal rotation is assumed to be a free rotation and its partition function is calculated

- The QHA  $\Delta_R G$  value is much closer to the experimental one of -92.2 kJ/mol<sup>[6a]</sup> than the one calculated with standard harmonic approximation (-108.6 kJ/mol)

	ACN	(ACN)H <sup>+</sup>	(ACN) <sub>2</sub> H <sup>+</sup>
$\epsilon_0 / E_h$	-132.816103564	-133.127065632	-265.996251514
$\det(\boldsymbol{\sigma}) / (\text{amu} \text{ \AA}^2)^M$	$1.1723 \times 10^{-32}$	$1.2505 \times 10^{-38}$	$1.4578 \times 10^{-68}$
$f_{corr}$	0.317831	0.600657	0.003921
$\ln(Z_{vib}^{QHA,corr})$	-27.042012	-30.567615	-54.663061
$\ln(Z_{vib}^{HA})$	-47.534292	-58.670193	-98.827696
$\ln(Z_{IR})$	-	-	0.963855
$\ln(Z_{rot})$	7.809736	7.887583	9.947041
$\ln(Z_{trans})$	16.150450	16.186853	17.208483
$G_{corr} / E_h$	0.002910	0.006131	0.025063
$\Delta_R G / \text{kJ/mol}$	-	-	-97.30



	CH <sub>3</sub> OH	CH <sub>3</sub> NH <sub>3</sub> <sup>+</sup>	[CH <sub>3</sub> NH <sub>3</sub> + CH <sub>3</sub> OH] <sup>+</sup>
$\epsilon_0 / E_h$	-115.781709438	-96.2659366250	-212.082861611
$\det(\boldsymbol{\sigma}) / (\text{amu} \text{ \AA}^2)^M$	$3.8346 \times 10^{-32}$	$1.5646 \times 10^{-50}$	$5.9329 \times 10^{-77}$
$f_{corr}$	0.331338	0.396865	0.000512
$\ln(Z_{vib}^{QHA,corr})$	-26.417544	-41.983030	-64.577338
$\ln(Z_{vib}^{HA})$	-53.822945	-83.555063	-136.467535
$\ln(Z_{IR})$	-	-	$\begin{pmatrix} 3.117 \\ 1.240 \end{pmatrix}$
$\ln(Z_{rot})$	8.053511	7.356496	11.499821
$\ln(Z_{trans})$	15.778953	15.780068	16.819231
$G_{corr} / E_h$	0.002441	0.017795	0.030121
$\Delta_R G / \text{kJ/mol}$	-	-	-66.50



### [CH<sub>3</sub>NH<sub>3</sub> + CH<sub>3</sub>OH]<sup>+</sup>

- Again,  $Z_{vib}$  differs largely from harmonic values
- $f_{corr}$  for the cluster is even smaller than in the ACN case!
- Two internal rotations (see pictures) were identified and separated out individually. Their partition functions are calculated assuming negligible barriers
- The QHA  $\Delta_R G$  value compares not so well to the experimental one of -49.31 kJ/mol<sup>[6b]</sup> and actually worse than the one calculated with standard harmonic approximation (-59.3 kJ/mol), although the latter is also unsatisfactory
- The molecule is very floppy which might hinder the description of the configurational distribution by normal coordinates; it is possible that internal ones (BAT coordinates) would perform better

## Conclusions

### Method:

- The QHA was applied to small molecular clusters for the determination of absolute classical vibrational partition functions and thermochemical data, such as free reaction enthalpies from MD simulations.
- Normal coordinates, obtained from standard ab initio NMA, were used for the description of the configurational distribution.
- It was important to separate out large amplitude motions such as overall translation, overall rotation and internal rotation by putting the molecule into the Eckart frame and separating out internal rotation by enforcing the Eckart-Sayvetz conditions at each time step of the MD trajectory.
- The covariance matrix of the normal coordinates, determined from the MD simulation, was tested for convergence, deviations from the multidimensional Gaussian and mean temperature in each mode.
- Although sometimes difficult, this analysis gave us confidence of the applicability of our treatment to small molecular clusters.

### Case study:

- Two clusters studied: (ACN)<sub>2</sub>H<sup>+</sup>, [CH<sub>3</sub>NH<sub>3</sub> + CH<sub>3</sub>OH]<sup>+</sup>
- While the one performed better, the second performed worse than the standard harmonic approximation
- While computationally more expensive than standard treatment, our method is faster than advanced techniques (VPT2, VSCF)

## Outlook

- Compare to results when internal (BAT) coordinates are used
- Study more molecular systems to test performance
- Study performance at different temperatures
- While the analysis is mostly automated, identification of the internal rotation still needs to be implemented in an automatic fashion

## Literature

- Klee et al. (2014): *J. Am. Soc. Mass Spectrom.*, 25(8), 1310–1321.
- Frisch et al. (2016): Gaussian16, Revision A.03, Gaussian, Inc., Wallingford CT.
- Dennington II, Keith, Millam (2016): GaussView 6.0.16, Semichem, Inc.
- (a) Karplus, Kushick (1981): *Macromolecules*, 14(2), 325–332 (b) Levy et al. (1984): *Macromolecules*, 17(7), 1370–1374 (c) Rojas et al. (1986): *J. Chem. Phys.*, 85(2), 1037–1043.
- (a) Eckart (1935): *Physical Review*, 47(7), 552–558. (b) Sayvetz (1939): *J. Chem. Phys.*, 7(6), 383–389.
- (a) Meot-Ner (Mautner), M. (2005): *Chem. Reviews*, 105(1), 213–284. (b) Meot-Ner (Mautner), M. (1984): *JACS*, 106(5), 1257–1264.

## Acknowledgement

Financial support by iGenTraX UG (Haan, Germany) is gratefully acknowledged.