A genetic algorithm based determination of the ground and excited (^1L_b) state structure and the orientation of the transition dipole moment of benzimidazole.

Michael Schmitt*, Daniel Krügler†, Marcel Böhm, Christian Ratzer
Heinrich-Heine-Universität, Institut für Physikalische Chemie
D-40225 Düsseldorf, Germany.
Phone: +49 211 81 13691, Fax: +49 211 81 15195
E-mail: mschmitt@uni-duesseldorf.de

Violetta Bednarska‡, Ivo Kalkman and W. Leo Meerts
Molecular- and Biophysics Group,
Institute for Molecules and Materials, Radboud University Nijmegen,
P.O. Box 9010, NL-6500 GL Nijmegen, The Netherlands.
Phone: +33 24 365 3023. Fax: +33 24 356 3311
E-mail: Leo.Meerts@sci.kun.nl

October 13, 2005

Abstract

The structure of benzimidazole has been determined in the electronic ground and excited states using rotationally resolved electronic spectroscopy. The rovibronic spectra of four isotopomers and subsequently the structure of benzimidazole have been automatically assigned and fitted using a genetic algorithm based fitting strategy. The lifetimes of the deuterated isotopomers have been shown to depend on the position of deuteration. The angle of the transition dipole moment with the inertial axis could be determined to be $-30°$. Structures and transition dipole moment orientation have been calculated at various levels of theory and were compared to the experimental results.

1 Introduction

The thorough knowledge of excited state properties like structures, excitation energies, lifetimes, and transition dipole moments of indole and indole analogues is an important prerequisite for the

*corresponding author
†present address: Bruker Daltonik GmbH, D-28359 Bremen, Germany
‡present address: Politechnica Opolska, Technical University of Opole, Faculty of Electrical Engineering and automatic control, Ul. St. Mikolajczyka, 45-271 Oplole, Poland

interpretation of chromophore fluorescence in proteins. Many of the excited state properties of these molecules have been predicted at all levels of theory ranging from configuration interaction with single excitation over time-dependent density functional theory to complete active space self consistent field methods with second order perturbation corrections. The high level of approximations in some of the theoretical methods sometimes leads to contradictory results between the applied methods. Therefore, it is desirable to compare them to experimental results. Rotationally resolved fluorescence spectroscopy in combination with a genetic algorithm (GA) based fitting procedure is an ideal tool for determination of all excited state properties, mentioned above.

The automated assignment of very complex rotationally resolved electronic spectra using GA based fitting procedures has been shown to be very successfull if the cost function can be evaluated sufficiently fast. [1–3] The obtained molecular parameters can be used in the interpretation and evaluation of excited state molecular properties like the orientation of the transition dipole moment, the excited state lifetimes and the geometry in both electronic states connected by the transition. The problem of determining molecular structures from an insufficient number of inertial parameters is well known. If not enough information from singly substituted isotopomers is available to perform a complete Kraitchman analysis, a fit to a pseudo-Kraitchman structure can be used within a model geometry with reduced numbers of parameters to be fit. [4] Because of the nonlinear relation between internal coordinates and rotational constants, it is necessary to perform a nonlinear fit, which means that starting values for the geometry parameters are needed and an iterative procedure has to be applied. If there is an insufficient number of inertial parameters several assumptions on the geometry have to be made, regarding relations between the parameters. The more of these artificial constraints have to be imposed on the real geometry, the more unreliable the fit gets, regarding its ability to find the global minimum. Therefore the application of GA based strategies, which do not depend on the initial conditions, is also advantageous for the nonlinear fit of the geometry.

Benzimidazole easily forms several deuterated isotopomers upon coexpansion with D_2O which are spectrally well separated. Recording of isotopically substituted species always allows the determination of the orientation of the transition dipole moment (TDM), while rotationally resolved electronic spectra of a single isotopic species generally yields only the absolute value of the TDM angle. [3] Knowledge of the direction of the TDM contains the information about the electronic nature of the excited state, which can be compared to theoretical predictions. The use of the GA for automated fitting of the rotationally resolved spectra allows for a very accurate determination of the TDM angle, because all lines in the spectrum are considered in the calculation of the cost function.

The low resolution R2PI spectra of the four isotopomers investigated in the present study have been presented by Jacoby *et al.* [5] Furthermore, they performed *ab initio* calculations on the different tautomers of benzimidazole. The first vibronically resolved spectrum of benzimidazole in a molecular beam was published by Jalviste and Treshchalov. [6] Velino *et al.* reported the microwave spectrum of benzimidazole and the ND deuterated isotopomer. [7] They proved planarity in the electronic ground state on the basis of the rotational constants of both isotopomers. The inertial parameters in the first electronically excited were determined from a band contour analysis of benzimidazole in the vapour phase by Cané *et al.* [8] Later Berden *et al.* performed rotationally resolved LIF spectroscopy and proved as well the planarity of the molecule also in the first electronically excited state. [9] Serrano-Andrés and Borin presented a theoretical study of the absorption and emission spectra of benzimidazole. [10, 11] They determined the transition dipole moment orientation from a CASSCF calculation to be $+29°$ for the ^1L_b-state of benzimidazole.

This value is in contradiction to an experimentally determined one from polarization measurements in a stretched polymer of $-16°$. [12] A similar disagreement between the experimentally determined TDM and the CASSCF calculated value has been found for 7-azaindole. [3]

In the present study the structural parameters of benzimidazole in its ground and electronically excited states are determined via a genetic algorithm based fitting procedure both of the experimental spectrum as well as the geometry. Furthermore the absolute orientation of the TDM of the lowest excited state of benzimidazole will be determined and compared to the results of *ab initio* calculations.

2 Experimental

The experimental setup for the rotationally resolved LIF is described elsewhere [13]. Briefly, it consists of a ring dye laser (Coherent 899-21) operated with Rhodamine 110, pumped with 6 W of the 514 nm line of an Ar^+-ion laser. The light is coupled into an external folded ring cavity (Spectra Physics) for second harmonic generation (SHG). The molecular beam is formed by expanding benzimidazole, seeded in 600 mbar of argon, through a 70 μm hole into the vacuum. The molecular beam machine consists of three differentially pumped vacuum chambers that are linearly connected by skimmers (1 mm and 3 mm, respectively) in order to reduce the Doppler width. The molecular beam is crossed at right angles in the third chamber with the laser beam 360 mm downstream of the nozzle. The resulting fluorescence is collected perpendicular to the plane defined by laser and molecular beam by an imaging optics setup consisting of a concave mirror and two plano-convex lenses. The resulting Doppler width in this set-up is 25 MHz (FWHM). The integrated molecular fluorescence is detected by a photo-multiplier tube whose output is discriminated and digitized by a photon counter and transmitted to a PC for data recording and processing. The relative frequency is determined with a quasi confocal Fabry-Perot interferometer with a free spectral range (FSR) of 149.9434(56) MHz. The FSR has been calibrated using the combination differences of 111 transitions of indole for which the microwave transitions are known [14, 15]. The absolute frequency was determined by recording the iodine absorption spectrum and comparing the transitions to the tabulated lines [16].

3 Theoretical methods

3.1 *Ab initio calculations*

The structure of benzimidazole in the electronic ground state has been optimized at the HF/6-31G(d,p), B3LYP/6-31G(d,p), and MP2/6-31G(d,p) levels and at the CIS/6-31G(d,p) level for the electronically excited S_1-state with the Gaussian 98 program package (Revision a11). [17] The SCF convergence criterion used throughout the calculations was an energy change below 10^{-8} Hartree, while the convergence criterion for the gradient optimization of the molecular geometry was $\partial E/\partial r < 1.5 \cdot 10^{-5}$ Hartree/Bohr and $\partial E/\partial \varphi < 1.5 \cdot 10^{-5}$ Hartree/degrees, respectively. Additionally a CASSCF optimization of the ground and excited state structures has been performed with an active space of ten electrons in nine orbitals. The CAS orbitals are shown in Figure 1, both occupied and unoccupied ones sorted by increasing energy. The CASSCF calculations have been performed using the MOLCAS-5.4 program package. [18] Additionally the geometry of the electronically excited state has been optimized using time-dependent density functional the-

ory [19, 20] with the B3-LYP functional [21, 22] using the TZVP basis and Dunnings basis set of triple ζ quality (cc-pVTZ). [23] These calculations were performed with the program system TURBOMOLE [24, 25] and yield also the TDM direction of the electronically excited states.

3.2 The genetic algorithms

The genetic algorithm is basically a global optimizer, which uses concepts copied from reproduction and selection in nature. For a detailed description of the GA the reader is referred to the original literature [26–28]. The GA library PGAPack version 1.0, which can run on parallel processors, has been used [29]. We shortly introduce the elements of the GA, which will be used in the analysis of the spectra and the structures.

- Representation of the parameters: The molecular parameters are encoded binary, each parameter representing a gene. A vector of all genes, which contains all molecular parameters is called a chromosome. In an initial step the values for all parameters are set to random values between lower and upper limits which have to be chosen by the user.
- The solutions are evaluated by a fitness function, which is a measure for the agreement between simulation and experiment.
- One optimization cycle, including evaluation of the cost of all chromosomes, is called a generation.
- Pairs of chromosomes are selected for reproduction and their information is combined via a crossover process. Crossover combines information from the parent generations.
- The value of a small number of bits is changed randomly. This process is called mutation.

3.3 Fitting of the spectra

The cost function for evaluation of the quality of a given solution has been discussed in detail in refs. [1, 2]. It is expressed as:

$$C_{fg} = 100(1 - F_{fg}) = 100(1 - \frac{(\boldsymbol{f}, \boldsymbol{g})}{||\boldsymbol{f}||||\boldsymbol{g}||}), \tag{1}$$

where $\boldsymbol{f}$ and $\boldsymbol{g}$ represent the experimental and calculated spectra, respectively. The inner product $(\boldsymbol{f}, \boldsymbol{g})$ is defined with the metric $\boldsymbol{W}$

$$(\boldsymbol{f}, \boldsymbol{g}) = \boldsymbol{f}^T \boldsymbol{W} \boldsymbol{g} \tag{2}$$

and the norm of $\boldsymbol{f}$ defined as $\|\boldsymbol{f}\| = \sqrt{(\boldsymbol{f}, \boldsymbol{f})}$ (a similar definition holds for $\boldsymbol{g}$). $\boldsymbol{W}$ has the matrix elements $W_{ij} = w(|j - i|) = w(r)$. For $w(r)$ we used a triangle function [1] with a user controlled width of the base Δw

$$w(r) = \begin{cases} 1 - |r| \,/\, \left(\frac{1}{2}\Delta w\right) & \text{for } |r| \leqslant \frac{1}{2}\Delta w \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Since the GA performs a lineshape fit of the complete spectrum, much more information on the linewidth is gathered than from a lineshape fit to a few individual lines. In order to obtain the relevant parameters that determine the intensities in the spectrum, we performed a second GA fit with a reduced search range for the inertial parameters and the weight function width

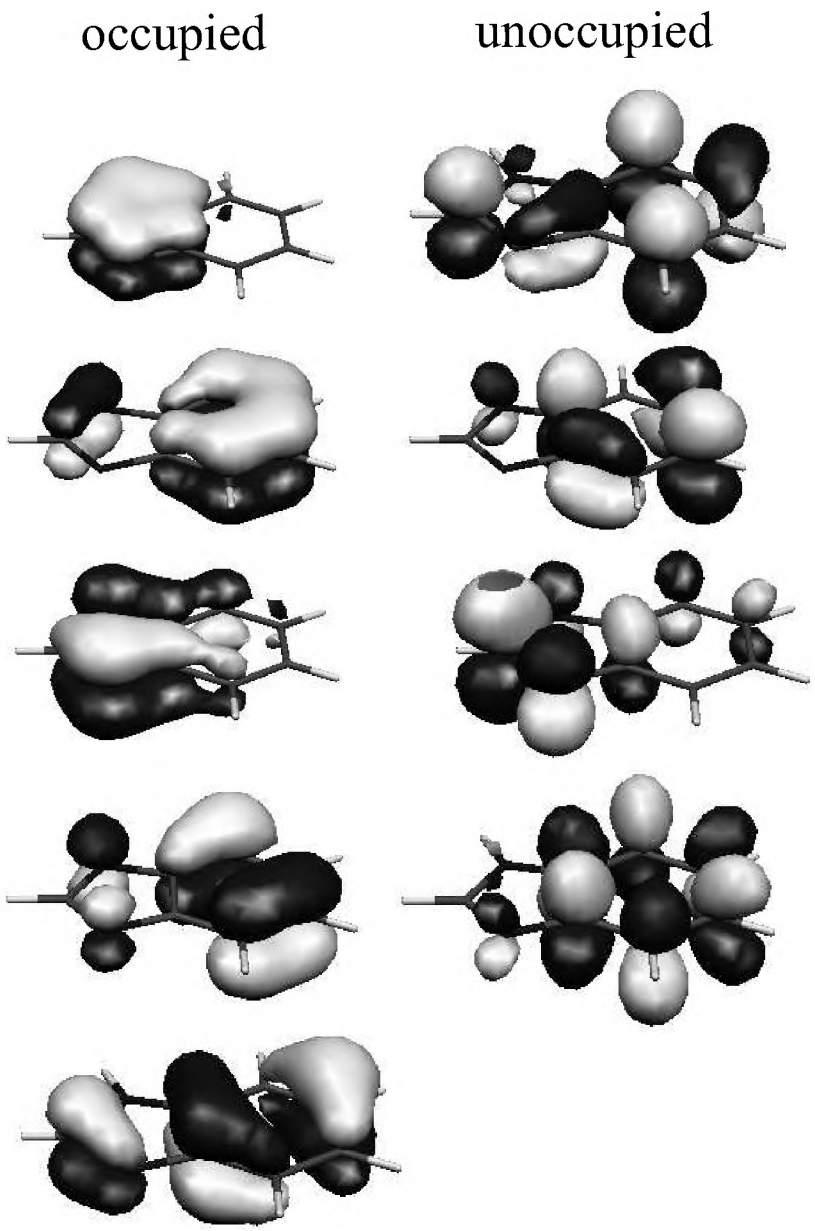

Figure 1: Occupied and unoccupied CASSCF orbitals sorted by increasing energy.

$\Delta w = 0$. This resulted in improved values for the angle θ that is connected to the components of the transition dipole moment by:

$$\lambda_a{}^2 = (\mu_a/\mu)^2 = \cos^2\theta \tag{4}$$

$$\lambda_b{}^2 = (\mu_b/\mu)^2 = \sin^2\theta \tag{5}$$

The determination of the Lorentzian component of the line width can be improved using the fit of all available intensities. The Gaussian width is fixed to the experimentally determined value of 25 MHz. The temperature dependence of the intensity is described by a two temperature model [30]:

$$n(T_1, T_2, w) = e^{-E/kT_1} + we^{-E/kT_2} \tag{6}$$

were E is the energy of the lower state, k is the Boltzmann constant, w is a weighting factor, T_1 and T_2 are the two temperatures.

3.4 Fitting of the structure

The program *pKrFit* [31] is used to determine the structure of benzimidazole in the S_0 and S_1-states. There are several new features of the program which are described in the following.

One of the most limiting problems in the process of the structure determination of larger molecules is the lack of independent data (rotational constants) compared to the number of model parameters. These parameters are the internal coordinates used to describe the geometry and a small number of parameters used to describe the vibrational contributions to the structure. Sometimes the number of model parameters to be determined can be reduced by using more symmetric models, where several parameters are interdependent by means of constraints (planarity, mirror planes, etc.). But often those symmetric models don't exist or they prove to be bad approximations to the real structure. Another well-known method to increase the number of independent data is to combine data of different sources. This method has been shown to work successful in combined fits of the intensities of vibronic spectra (Franck-Condon fit) and the changes of rotational constants upon electronic excitation (inertial parameter fit) in order to determine geometry changes upon excitation. [32] Another advantage of this method is, that it minimizes the correlations among data of given measurement series. This has been demonstrated by Lees who investigated the positive effect of mixing data of different sources (infrared and microwave data), although these had very different accuracies. [33]

To increase the number of data points in our structural determination fits we extended the functional capability of *pKrFit* to take both the moments of inertia as well as the angles (or the corresponding direction cosines) of the TDM of the respective isotopomer with the main inertial axes into account. Both experimental data sets are independent from each other and can be extracted from the spectra fits described in section 3.3. In the following we show how these two data sets can be used in a combined fit to determine the best parameter set of internal coordinates (i.e. distances, angles, and dihedral angles) describing a given molecular structure.

Using the nomenclature introduced in [31] the cost function to minimize in a pure rotational constants fit is given by the weighted sum of squared residuals: [34, 35]

$$\chi^2 = \chi^2_{rot} = \Delta\boldsymbol{y}^T \boldsymbol{W}_{rot} \Delta\boldsymbol{y}. \tag{7}$$

In this equation $\boldsymbol{y}^T = (B^0_{a,1}, B^0_{b,1}, B^0_{c,1}, ..., B^0_{a,k}, B^0_{b,k}, B^0_{c,k})$ is the vector of $n = 3k$ measured and calculated rotational constants $\boldsymbol{y}_{exp}$ and $\boldsymbol{y}_{calc}$ of k isotopomers, respectively, $\Delta\boldsymbol{y} = \boldsymbol{y}_{exp} -$

$\boldsymbol{y}_{calc}$ is their corresponding residual vector and $\boldsymbol{W}_{rot}$ denotes an n-by-n positive-definite weight matrix obtained from the covariance matrix of the n experimental rotational constants. Via the inverse relation to the moments of inertia, the rotational constants directly depend on the cartesian coordinates of the isotopomers, thus the mathematical relation between those n inertial values and the internal coordinates b_j, $j = 1..m$, of a hypothetical equilibrium structure is straightforward.

There exists no direct relation between the geometric TDM parameters, i.e. either polar angles (θ, ϕ) or squared direction cosines $(\lambda_a^2, \lambda_b^2)$, and the internal coordinates b_j of the molecule similar to that of the inertial parameters. But we can deduce an indirect relation by recognizing that the rotation of the inertial frame of a given reference isotopomer into any other isotopomer also implies a corresponding change of the TDM vector. In other words: We extend our former model of a molecule described by a set of internal coordinates by a molecular geometry and an additional TDM vector of unit length of an arbitrary reference isotopomer. It should be emphasized that although this ansatz adds two further fit parameters to the model, these are usually outweighed by the increased number of additionally available experimental TDM data. Mathematically, the TDM vector $\boldsymbol{\mu}_{ref} = (\mu_{ref,a}, \mu_{ref,b}, \mu_{ref,c})^T$ of the reference isotopomer is transformed into the TDM vector $\boldsymbol{\mu}_{other}$ of another isotopomer via

$$\boldsymbol{\mu}_{other} = \boldsymbol{R}_{other}^T \boldsymbol{R}_{ref} \boldsymbol{\mu}_{ref}. \tag{8}$$

In this equation $\boldsymbol{R}_i$ denotes the 3×3 orthogonal rotation matrix which transforms the initial geometry of the i^{th} isotopomer into its own inertial system. Our fit program accepts either reduced polar angles or squared direction cosines as input parameters. The polar angles are reduced, because the corresponding octant cannot be determined by the experiment and thus both angles are confined to the limited range 0–90° instead of their full range. The direction cosines are similarly constrained, because only their absolute value can be determined. Using the extended model the cost function is given by

$$\chi^2 = \chi_{rot}^2 + \chi_{TM}^2 = \Delta \boldsymbol{y}^T \boldsymbol{W}_{rot} \Delta \boldsymbol{y} + \Delta \boldsymbol{z}^T \boldsymbol{W}_{TDM} \Delta \boldsymbol{z} \tag{9}$$

where either $\boldsymbol{z}^T = (\theta_1, \phi_1, ..., \theta_r, \phi_r)$ or $\boldsymbol{z}^T = (\lambda_{a,1}^2, \lambda_{b,1}^2, ..., \lambda_{a,r}^2, \lambda_{b,r}^2)$ in the general *abc* case depending on the available r transition dipole moment data. In the special case of *ab* hybrid spectra, the vector $\boldsymbol{z}$ simplifies to $\boldsymbol{z}^T = (\theta_1, ..., \theta_r)$ or $\boldsymbol{z}^T = (\lambda_{a,1}^2, ..., \lambda_{a,r}^2)$, respectively. Similarly to the weight matrix $\boldsymbol{W}_{rot}$ of the rotational constants, $\boldsymbol{W}_{TDM}$ is the weight matrix obtained by the covariance matrix of the TDM data. Therefore the standard deviations of the TDM data were determined from the standard deviations of three independent runs of the GA using different (random) starting parameters. Additionally it was assumed that no correlations exist among the different polar angles or direction cosines, respectively.

The original optimization strategy of *pKrFit* used a gradient-based χ^2 minimizer. While the speed of this approach is appealing, the main disadvantage is the possibility of becoming trapped in local minima different from the absolute one. We decided to extend the program's capabilities by incorporating the same PGAPack global optimizer as used in the spectra fit. The GA library was used in minimization mode and thus directly used the correspondingly defined χ^2 value as cost function.

Furthermore our global minimizer has been extended in to allow intermediate local minimization steps rather than simple cost function evaluations. This technique was originally introduced by Li and Scheraga as part of a simulated annealing (SA) exploration of the potential hypersurface

of proteins. [36] Also known as *basin-hopping* [37] by Doyle and Wales, it was successfully combined with GA and SA to determine minimum-energy structures of fullerene and atomic clusters. [38–41]

4 Results

The atomic numbering for the designation of the isotopomers and the geometry parameters for the structure fit used throughout this publication is shown in Figure 2.

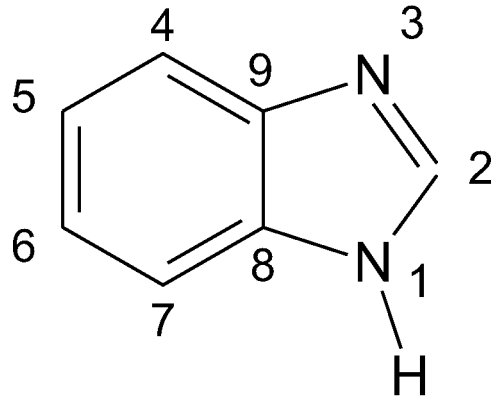

Figure 2: Atomic numbering used for definition of the isotopomers.

Figure 3 shows the rotationally resolved spectrum of the electronic origin ν_0 of 1H2H-benzimidazole at 36021.38 cm^{-1} along with the simulation using the best parameters given in Table 1, that are obtained by a GA fit as described in section 3.3. It is an *ab* hybrid type spectrum with an angle θ between the inertial *a*-axis and the transition dipole moment of 31.9(5)° (72.1 % *a*, 27.9 % *b*-type). The excited state lifetime of this isotopomer is determined to be 5.2(2) ns from a 30.8(10) MHz Lorentzian contribution to the Voigt profile.

Figure 4 shows the rovibronic spectrum of the electronic origin of 1D2D-benzimidazole. The electronic origins of the mixed isotopomers 1D2H-benzimidazole and 1H2D-benzimidazole are shown in the supplementary material. The results of the GA fits of all four spectra are summarized in Table 1. All bands are *ab* hybrids with similar angles θ between TDM and the inertial *a*-axis as the undeuterated isotopomer. Nevertheless, the small differences of the TDM angles of the different isotopomers are crucial in the determination of the absoute transition dipole moment orientation, as has been shown in Ref. [3].

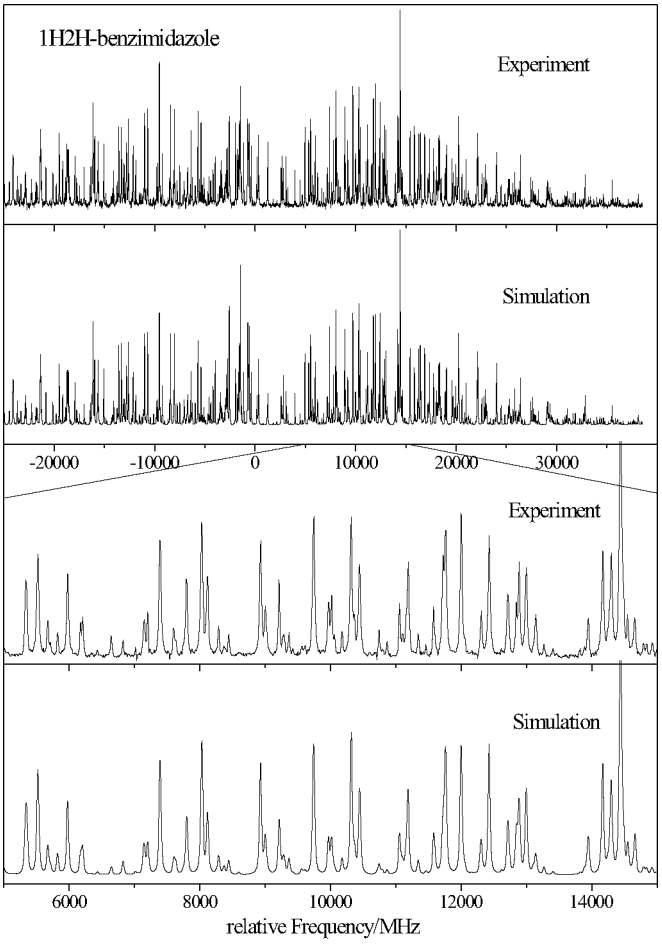

Figure 3: Rotationally resolved spectrum of the electronic origin ν_0 of 1H2H-benzimidazole at 36021.38 cm^{-1} along with the simulation using the best parameters given in Table 1.

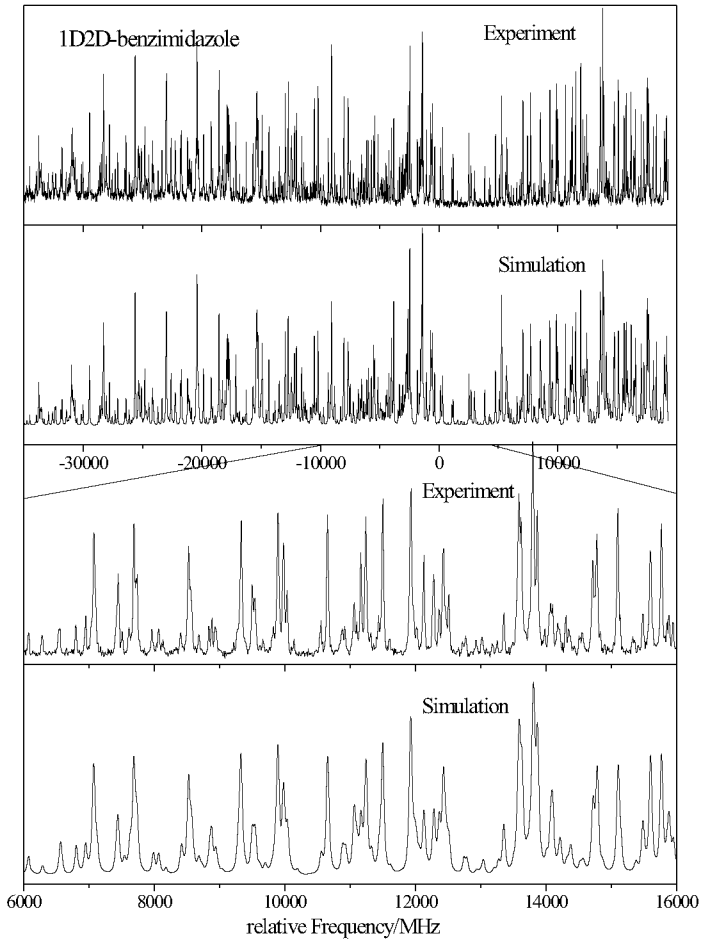

Figure 4: Rotationally resolved electronic spectrum of the origin ν_0 of 1D2D-benzimidazole at 36052.64 cm^{-1} along with the simulation using the best parameters given in Table 1.

Table 1: Molecular parameters of the electronic origin band of benzimidazole (B) isotopomers as obtained from the genetic algorithm fit. The numbering of the substitution position refers to Figure 2.

	1H2H-B[a]	1D2H-B[a]	1H2D-B	1D2D-B
A''/MHz	3929.720(7)	3811.29(19)	3930.61(69)	3810.69(57)
B''/MHz	1679.259(3)	1658.846(2)	1618.58(29)	1599.70(25)
C''/MHz	1176.747(1)	1156.052(3)	1146.71(28)	1126.94(24)
$\Delta\nu_0$/cm^{-1}[b]	0.00	7.08(1)	24.18(1)	31.26(1)
θ/°	−31.9(5)	−30.4(5)	−31.3(5)	−30.2(5)
$\Delta_{Lorentz}$/MHz	30.8(10)	22.8(10)	30.8(10)	26.5(10)
$t_{1/2}$/ns	5.2(2)	7.0(3)	5.2(2)	6.0(3)
ΔA/MHz	−155.72(22)	−149.98(11)	−156.27(14)	−150.26(13)
ΔB/MHz	−15.28(15)	−13.77(13)	−14.67(14)	−13.53(14)
ΔC/MHz	−21.40(11)	−20.47(11)	−20.65(13)	−19.85(13)

[a] fixed to the microwave values from Ref. [7].
[b] relative to the electronic origin of 1H2H-benzimidazole at 36021.38(1) cm^{-1}.

5 Discussion

5.1 Orientation of the transition dipole moment and excited state lifetimes

Off axis isotopic substitution results in a rotation of the respective inertial axis, but the orientation of the TDM remains the same. This procedure has successfully been applied to the determination of the absolute direction of the transition dipole moment in 7-azaindole. [3]

If the sign of the angle of the TDM with the inertial a-axis is positive (orientation I in Figure 5), the angle should increase for 1D2H-Benzimidazole, if the sign is negative (orientation II in Figure 5), the angle will decrease. As both 1D2H-benzimidazole and 1D2D-benzimidazole have a smaller angles θ than the respective 1H2H-benzimidazole and 1H2D-benzimidazole the sign of the angle θ is negative. Table 2 compares the angles of the TDM vector for the two lowest electronic ^1L_b and ^1L_a states from various experiments with the results of electronic structure calculations at different levels of theory. The CIS/6-31(d,p) calculations give an angle θ of $-15°$ in fair agreement with the experimental result of $-32°$. The determination of the TDM angle from the CASSCF(10,9) calculations yields a value of $+36°$ for the lowest state in contrast to the experimental findings. A similar result was found for 7-azaindole, where the experimentally determined orientation of the TDM has the opposite sign as obtained from CASSCF calculations. Further calculations on the transition dipole moment were performed with time-dependent B3-LYP calculations, using the TZVP basis set. Calculating the TDM in the geometry of the TDB3-LYP optimized S_1 geometry, a positive angle of $+36°$, close to the results of the CASSCF calculations resulted. Keeping the geometry fixed at the B3-LYP geometry obtained for the S_0 state, the TDM angle changes to 14°. While for time-dependent density functional theory a large difference between the TDM calculated in the ground state geometry and the one calculated in the excited state geometry is found, both CAS and CIS calculations show nearly no differences with respect to the reference geometry (S_0 or S_1). The large difference found at TDDFT level between the TDM calculated in the S_0 and

S_1 geometry, respectively can of course not be attributed merely to a rotation of the inertial axis system upon electronic excitation. The geometry changes are too small, which is immediately obvious from the Franck-Condon pattern of the vibronic spectrum, which shows the vibrationless origin as the by far most intense band.

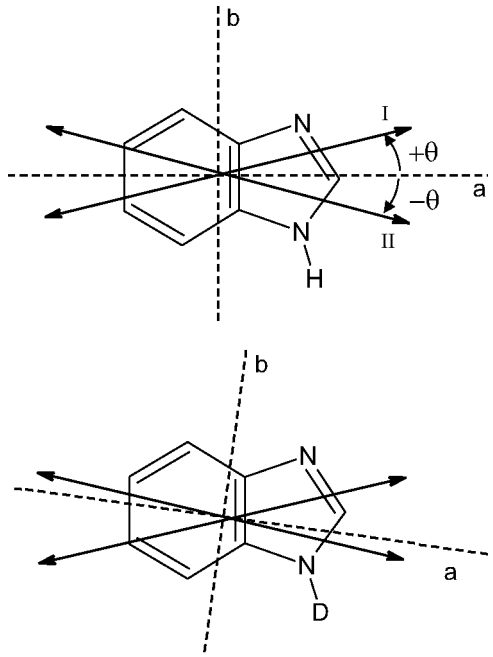

Figure 5: Angle θ between TDM and inertial a-axis for 1H2H-benzimidazole and 1D2H-benzimidazole. The rotation of the inertial axis system upon deuteration is not drawn to scale for reasons of clarity.

The lifetime of the four isotopomers has been determined from the Lorentzian contribution to the Voigt profile. The lineshape fit using the GA, results in very accurate values for the Lorentzian contribution, because in contrast to a conventional lineshape fit, all transitions which contribute to the band are taken into account, even if they are overlapping partially. Although the differences of the lifetimes, determined from the Lorentz width are small, they are well outside the uncertainties. The two isotopomers, that are deuterated at the N-H group (1D2H-benzimidazole and 1D2D-benzimidazole) have generally longer lifetimes (6.0 and 7.0 ns) than the ones undeuterated in this position (5.2 ns). A similar situation was found for phenol, where deuteration at the hydroxy group leads to a substantial increase of the lifetime. [31, 42, 43]. As Sobolewski and Domcke pointed out for the similar indole system, an increasing lifetime upon deuteration at the NH group, indicates a control of the lifetime by a tunneling process along the NH coordinate. [44] This tunneling takes place between the the excited $\pi - \pi^*$ and the $\pi - \sigma^*$ surfaces. The $\pi - \sigma^*$ surface has a conical intersection with the ground state surface which allows a very fast evolution of the excited state population to the ground state.

Table 2: Experimental and calculated TDM orientations of the two lowest excited states of benzimidazole. CASSCF alculations have been performed with the (10,9) space described in the text using the 6-31G(d,p) basis set. The TDDFT calculations have been performed using the B3-LYP functional and the TZVP basis set.

	Exp.[a]	Exp.[b]	CIS [c]	CIS[d]	CAS[c]	CAS[d]	TDDFT[c]	TDDFT[d]
S_1 (^1L_b)	−32	−16	−15	−14	+36	+36	+39	+14
S_2 (^1L_a)	–	−80	−71	−79	−52	−53	–[e]	−34

[a]This work
[b]From polarized absorption in a stretched polymer. [12]
[c]In the optimized geometry of the excited state
[d]In the optimized geometry of the ground state
[e]Due to state flipping the optimization did not converge for the S_2 state

5.2 Determination of the structure

5.2.1 Comparison to the *ab initio* calculations

The experimentally determined rotational constants of benzimidazole are compared to *ab initio* calculated rotational constants in Table 3. The MP2/6-31G(d,p) rotational constants agree within 1% with the experimental ones. Although the HF (for the ground state) and CIS calculations (for the excited state) with the same basis set show quite substantial deviations for the absolute values, the difference of both sets gives a very good approximation of the change of the rotational constants upon electronic excitation (cf. Table 3). The CASSCF(10/9) optimized structure shows both acceptable absolute values, slightly worse than at MP2 level for the ground state and also very accurate changes upon electronic excitation. Even better absolute rotational constants are obtained from density functional theory for the electronic ground state and time-dependent density functional theory for the electronically excited state. Both DFT and TDDFT have been performed using the B3-LYP functional with the TZVP basis set.

Table 3: Experimental and *ab initio* inertial constants of benzimidazole. All calculations have been performed using the 6-31G(d,p) basis set. All values are given in MHz. The DFT and TDDFT calculations have been performed using the B3-LYP functional and the TZVP basis set.

	Exp.	MP2	HF	CIS	CIS-HF	CAS(10/9)	B3-LYP/TDB3-LYP
A''	3929.7	3918	4016	-	-	3971[a]	3953
B''	1679.3	1681	1700	-	-	1684[a]	1682
C''	1176.7	1176	1194	-	-	1183[a]	1180
A'	3774.0	-	-	3876	-	3807[b]	3773[b]
B'	1664.0	-	-	1690	-	1657[b]	1672[b]
C'	1155.3	-	-	1177	-	1154[b]	1160[b]
ΔA	−155.7	-	-	-	−140	−164	−180
ΔB	−15.3	-	-	-	−10	−27	−10
ΔC	−21.4	-	-	-	−17	−29	−20

[a]Geometry optimized to the S_0-state.
[b]Geometry optimized to the S_1-state.

5.2.2 Fit of the structure

The r_0-structure of benzimidazole has been determined using the program *pKrFit*, whose new features are described in section 3.4. The bond lengths obtained from the fit to the rotational constants and to the changes of the TDM angle θ of four isotopomers are given in Table 4. The TDM of the undeuterated benzimidazole has been taken as reference and the change of the orientation of the three remaining isotopomers was used in the fit as additional parameter (Note that the method is generally invariant under the choice of the actual reference regardless of the quality of the initial TDM parameter value). Within the Born-Oppenheimer approximation being valid for this molecule, the different TDM orientations are only due to the geometry change induced rotations of the inertial axis system. The application of genetic algorithms for the variation of the geometry parameters minimized the risk to be trapped in a local minimum, which we found could easily happen if gradient based downhill optimizers were used. Furthermore, the possibility of defining physically meaningful parameter limits excludes the output of molecular structures which bear no relationship to reality.

Table 4: Comparison of experimentally determined geometry parameters of benzimidazole in the S_0 and in the S_1-state. All distances are given in pm.

	S_0	S_1	Δ
d(N_1C_8)	137.0(7)	133.8(18)	−3.2
d(C_8C_9)	140.7(0)	137.3(10)	−3.4
d(C_9N_3)	137.6(6)	132.5(17)	−5.1
d(N_3C_2)	129.6(10)	129.8(12)	+0.2
d(C_2N_1)	144.4[a]	143.3[a]	−1.1
d(C_9C_4)	141.3(11)	144.5(20)	+3.2
d(C_4C_5)	142.3(3)	139.9(8)	−2.3
d(C_5C_6)	144.6(1)	148.0(5)	+3.4
d(C_6C_7)	140.7(11)	146.0(22)	+5.3
d(C_7C_8)	139.0[a]	146.8[a]	+7.8

[a]this parameter is not needed in the z-matrix definition of the structure, since as last bond in a cyclic system, it is defined by the preceeding bond lengths. It is therefore not included in the fit, but calculated from the resulting structure and given here only for completeness.

Comparison with the results of the B3-LYP/TDB3-LYP optimized geometries shows a good agreement with the trends of the experimentally determined geometry changes. While a general expansion in the benzene ring is found upon excitation, the situation in the imidazole ring is different. We deduce a general decrease of bond lengths from the experiment. The difference of the B3-LYP bond lengths (for the ground state) and the time-dependent B3-LYP (for the excited state) in the imidazole ring shows alternating bond length changes.

Since the deuteration takes place only in the imidazole moiety of the molecule, the geometry changes in the benzene moiety are quite inaccurate and correlated and depend considerably on the chosen model. Therefore, the values of Table 4 present only a possible set of geometry changes. More isotopic substitutions are necessary for a more complete description of the geometry changes upon electronic excitation.

6 Conclusions

The geometry of benzimidazole has been determined experimentally in the S_0 and in the first electronically excited S_1 state. The structure in both states was fit to the rotational constants of four different isotopomers and the change of the TDM orientation between the isotopomers due to inertial axis rotation. As optimizer both a gradient based Levenberg-Marquart variant and a genetic algorithm have been used. The first allows very fast optimizations (in less than a second, on a fast desktop personal computer) while the latter has the advantage of being implicitly independent of the starting geometry. Convergence with the GA can be reached within 200 generations in about 2 minutes . The geometry changes upon electronic excitation can be divided into two major contributions: A general increase of the bond lengths in the benzene ring moiety and an overall decrease of bond lengths in the imidazole moiety. The main contributions to the $S_1(^1L_b) \leftarrow S_0$ transition in the one-electron approximation of the CIS calculations are HOMO $\rightarrow$ LUMO (0.6), HOMO-1 $\rightarrow$ LUMO+1 (0.3), and HOMO-1 $\rightarrow$ LUMO (0.1), cf Figure 1. The main (π) bonding contributions of HOMO and HOMO-1 are localized in the benzene ring, while the imidazole ring has non-bonding or anti-bonding characteristics. The LUMO and LUMO+1 are comprised of mainly anti-bonding interactions in the benzene ring and non-bonding and anit-bonding interactions in the imidazole ring. This fact might give a pictoral explanation of the origin for the observed geometry changes. Of course one has to keep in mind that other references than single excitations also contribute considerably to the excitation.

While the geometric structure of the electronically excited state can be calculated with sufficient accuracy even at time-dependent density functional theory, the determination of the transition dipole moment orientation poses serious problems to theory. First of all, although the rotational constants of the excited state from the CIS, TDDFT, and CASSCF calculations match closely the experimental values, even the quadrant of the TDM vector differs between the various calculations. Only the CIS calculated TDM orientation points to the negative quadrant (cf. Figure 5) as has been found experimentally. CASSCF calculations with the complete active π space result in a positive value. Obviously for an exact calculation of the TDM orientation the excitations in the σ space give important contributions. Although the CIS calculations includes only single excitations, the inclusion of σ excitations gives a better description of the TDM as a complete CI in a limited space (neglecting σ configurations) like in CASSCF. Nevertheless, the good agreement between experiment and CIS calculations has to be viewed as merely fortuitous, regarding the level of theory. Time-dependent DFT also considers only single excitations. Although the geometry change upon electronic excitation is small, the transition dipole moment orientation depends strongly on the reference geometry (S_0 or S_1). To compare with the vertical excitation scheme of our experiment we have to consider the value calculated in the optimized ground state geometry. This value is still positive, but much smaller than the respective CASSCF value. Future calculations, which include multiple excitations in a larger space (multi reference CI) together with time-dependent DFT should result in more accurate predictions for the transition dipole moment orientation.

Acknowledgment

The financial support of the Deutsche Forschungsgemeinschaft (SCHM 1043/9-4) is gratefully acknowledged. This work is part of the PhD thesis of Marcel Böhm. We thank Robert Brause for help with the *ab initio* calculations. The authors like to thank the National Computer Facilities of the Netherlands Organisation of Scientific Research (NWO) for a grant on the Dutch supercomputing

facility SARA.

References

[1] J. A. Hageman, R. Wehrens, R. de Gelder, W. L. Meerts, and L. M. C. Buydens, *J. Chem. Phys.*, 2000, **113**, 7955–7962.

[2] W. L. Meerts, M. Schmitt, and G. Groenenboom, *Can. J. Chem.*, 2004, **82**, 804–819.

[3] M. Schmitt, C. Ratzer, K. Kleinermanns, and W. L. Meerts, *Mol. Phys.*, 2004, **102**, 1605–1614.

[4] H. D. Rudolph, *Struc. Chem.*, 1991, **2**, 581–588.

[5] C. Jacoby, W. Roth, and M. Schmitt, *Appl. Phys. B*, 2000, **71**, 643.

[6] E. Jalviste and A. Treshchalov, *Chem. Phys.*, 1993, **172**, 325.

[7] B. Velino, A. Trombetti, and E. Cané, *J. Mol. Spec.*, 1992, **152**, 434–440.

[8] E. Cané, A. Trombetti, B. Velino, and W. Caminati, *J. Mol. Spec.*, 1991, **150**, 222.

[9] G. Berden, W. L. Meerts, and E. Jalviste, *J. Chem. Phys.*, 1995, **103**, 9596–9606.

[10] A. C. Borin and L. Serrano-Andrés, *Chem. Phys.*, 2000, **262**, 253–265.

[11] A. C. B. Luis Serrano-Andrés, *Chem. Phys.*, 2000, **262**, 267–283.

[12] I. Maki, K. Nishimoto, M.-A. Sugiyama, H. Hiratsuka, and Y.Tanizaki, *Bull. Chem. Soc. Jpn*, 1981, **54**, 8.

[13] M. Schmitt, J. Küpper, D. Spangenberg, and A. Westphal, *Chem. Phys.*, 2000, **254**, 349–361.

[14] R. D. Suenram, F. J. Lovas, and G. T. Fraser, *J. Mol. Spec.*, 1988, **127**, 472–480.

[15] W. Caminati and S. di Bernardo, *J. Mol. Struct.*, 1990, **240**, 253–262.

[16] S. Gerstenkorn and P. Luc, *Atlas du spectre d'absorption de la molécule d'iode*, CNRS, Paris, 1982.

[17] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian 98, revision a.11 Gaussian, Inc., Pittsburgh, PA, 2001.

[18] K. Andersson, M. Barysz, A. Bernhardsson, M. R. A. Blomberg, D. L. Cooper, M. P. Fülscher, C. de Graaf, B. A. Hess, G. Karlström, R. Lindh, P.-A. Malmqvist, T. Nakajima, P. Neogrády, J. Olsen, B. O. Roos, B. Schimmelpfennig, M. Schütz, L. Seijo, L. Serrano-Andrés, P. E. M. Siegbahn, J. Stalring, T. Thorsteinsson, V. Veryazov, and P.-O. Widmark, Molcas Version 5.4 Lund University, Sweden, 2002.

[19] R. Bauernschmitt and R. Ahlrichs, *Chem. Phys. Letters*, 1996, **256**, 454–464.

[20] R. Bauernschmitt, M. Häser, O. Treutler, and R. Ahlrichs, *Chem. Phys. Letters*, 1997, **264**, 573–578.

[21] C. Lee, W. Yang, and R. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

[22] A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.

[23] J. T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007.

[24] R. Ahlrichs, M. Bär, and H.-P. Baron, Turbomole (version 5.7) Universität Karlsruhe, Germany, 2002.

[25] R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, *Chem. Phys. Letters*, 1989, **162**, 165–169.

[26] J. H. Holland, *Adaption in Natural and Artificial Systems*, MI: The University of Michigan Press, Ann-Arbor, 1975.

[27] D. E. Goldberg, *Genetic Algorithms in search, optimisation and machine learning*, Addison-Wesley, Reading Massachusetts, 1989.

[28] I. Rechenberg, *Evolutionsstrategie - Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, Frommann-Holzboog, Stuttgart, 1973.

[29] D. Levine, PGAPack V1.0, PgaPack can be obtained via anonymous ftp from: ftp://ftp.mcs.anl.gov/pub/pgapack/pgapack.tar.z, 1996.

[30] Y. R. Wu and D. H. Levy, *J. Chem. Phys.*, 1989, **91**, 5278–5284.

[31] C. Ratzer, J. Küpper, D. Spangenberg, and M. Schmitt, *Chem. Phys.*, 2002, **283**, 153–169.

[32] D. Spangenberg, P. Imhof, and K. Kleinermanns, *Phys. Chem. Chem. Phys.*, 2003, **5**, 2505–2514.

[33] R. M. Lees, *J. Mol. Spec.*, 1970, **33**, 124–136.

[34] W. C. Hamilton, *Statistics in Physical Science*, Ronald Press, New York, 1964.

[35] S. Brandt, *Data Analysis*, Springer Verlag, Berlin, 3 ed., 1998.

[36] Z. Q. Li and H. A. Scheraga, *Proc. National Academy of Science*, 1987, **84**, 6611–6615.

[37] D. J. Wales and J. P. K. Doye, *J. Phys. Chem.A*, 1997, **101**, 5111–5116.

[38] D. M. Deaven and K. M. Ho, *Phys. Rev. Letters*, 1995, **75**, 288–291.

[39] S. K. Gregurick, M. H. Alexander, and B. Hartke, *J. Chem. Phys.*, 1996, **104**, 2684–2691.

[40] J. A. Niesse and H. R. Mayne, *J. Chem. Phys.*, 1996, **105**, 4700–4706.

[41] D. M. Deaven, N. Tit, J. R. Morris, and K. M. Ho, *Chem. Phys. Letters*, 1996, **256**, 155–200.

[42] A. Sur and P. M. Johnson, *J. Chem. Phys.*, 1986, **84**, 1206–1209.

[43] R. J. Lipert and S. D. Colson, *J. Phys. Chem.*, 1989, **93**, 135.

[44] A. L. Sobolewski and W. Domcke, *Phys. Chem. Chem. Phys.*, 2002, **4**, 1093–1100.

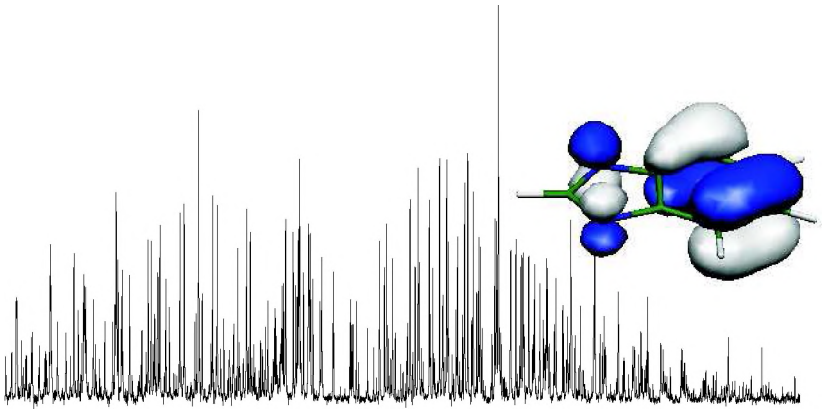

Figure 6: Graphical abstract