

---

# **Three Dimensional Structure Elucidation with the COSMOS-NMR Force Field**

---

## **Dissertation**

zur Erlangung des akademischen Grades doctor rerum  
naturalium (Dr. rer. nat.)

vorgelegt dem Rat der Physikalisch-Astronomischen Fakultät der  
Friedrich-Schiller-Universität Jena

von Diplom - Physiker Raiker Witter  
geboren am 05.11. 1971 in Gotha

Gutachter:

1. Prof. Dr. rer. nat. habil. Christian Jäger
2. Prof. Dr. rer. nat. habil. Eike Brunner
3. Prof. Dr. rer. nat. habil. Berthold Thomas

Tag der letzten Rigorosumsprüfung: 17.04. 2003

Tag der öffentlichen Verteidigung: 29.04. 2003

## Content

<b>1. Introduction</b>	<b>3</b>
<b>2. Structure Elucidation using Nuclear Spin</b>	<b>4</b>
2.1. From Wolfgang Pauli to Kurt Wüthrich	4
2.2. Structure Determination with NMR parameters	8
2.2.1. Basic NMR Theory	8
2.2.2. Basic NMR Hamiltonian	9
2.2.3. Magnetic Shielding and Chemical Shift	10
2.2.4. Direct Nuclear Dipole-Dipole Interaction	15
2.2.5. Indirect Nuclear Dipole-Dipole Interaction - J Coupling	16
2.2.6. Nuclear Electric Quadrupole Interaction	17
2.2.7. Nuclear Overhauser Effect (NOE)	18
2.2.8. Structure Determination with Force Fields	19
<b>3. Bond Polarization Theory (BPT)</b>	<b>21</b>
3.1. Introduction	21
3.2. Theoretical Background	21
3.3. Conclusion	25
<b>4. Efficient Charge Determination</b>	<b>26</b>
4.1. Introduction	26
4.2. Mulliken Population Analysis (MPA)	29
4.3. Electrostatic Charges (ESC)	32
4.4. Natural Population Analysis (NPA)	33
4.5. Conclusion	34
<b>5. <sup>13</sup>C Chemical Shift Molecular Dynamics Crystal Simulation</b>	<b>37</b>
5.1. Introduction	37
5.2. BPT <sup>13</sup> C NMR Chemical Shifts	38
5.3. Theoretical Background	39

5.4. Simulation of the $^{13}\text{C}$ Chemical Shift Spectra of Silk II	40
5.5. Results and Conclusion	43
<b>6. BPT Pseudo Forces</b>	<b>44</b>
6.1. Theory	44
6.2. Conclusion	51
<b>7. Proton Position Refinement of the X-ray Crystal Structure of D-Mannitol</b>	<b>53</b>
7.1. Introduction	53
7.2. Chemical Shift Driven Geometry Optimization	53
7.3. Conclusion	56
<b>8. 3D Solution Structure Determination of a Pseudopeptide Zinc Complex</b>	<b>57</b>
8.1. Introduction	57
8.2. Experimental	58
8.2.1. Modeling	58
8.2.2. Synthesis	59
8.2.3. NMR Experiments	59
8.3. NOE Pseudo Forces	61
8.4. Simulation and NMR Structure Refinement	63
8.5. Conclusion	67
<b>9. 3D Crystal Structure Refinement of Silk II</b>	<b>70</b>
<b>10. Results and Conclusion</b>	<b>71</b>
<b>11. Bibliography</b>	<b>73</b>
<b>Appendix</b>	
Integrals and Integral Derivatives	I
BPT Parameters of the Charge Parametrization	II
Deutsche Zusammenfassung	V
Thesen zur Dissertation	VII
Danksagung	IX

Ehrenwörtliche Erklärung

X

Lebenslauf

XII

## 1. Introduction

Nuclear Magnetic Resonance, NMR, is a precision measurement technique with a relative energy sensitivity of up to  $10^{-16}$  with respect to the Rydberg energy and can measure the life times of nuclear spin state in hours and coherence times in seconds. In most cases, the nuclear NMR Hamiltonian can be effectively separated from the electronic interactions which make theoretical considerations and simulations quite feasible. The NMR Hamiltonian and related relaxation processes are strongly coordinate dependent. For instance, direct nuclei coupling is proportional to the cube of their inter nuclear distance. The experimental set-up, which employs radio frequency methods, is one of the currently best developed technologies. Therefore, it is no surprise, that NMR and related techniques are employed in many other fields than physics, such as medicine, chemistry and biology. Several Nobel prizes are related to NMR techniques. As an example the latest prize, which was announced in October 2002, went to Kurth Wüthrich for his development of nuclear magnetic resonance spectroscopy for determining the 3D structure of biological macromolecules in solution.<sup>[1,2]</sup> Wüthrich contributed to this breakthrough by applying multidimensional NMR techniques such as proton Nuclear Overhauser Effect signals,<sup>[3]</sup> vicinal J coupling constants<sup>[4]</sup> and protein backbone  $^{13}\text{C}$  chemical shifts<sup>[5]</sup> for the extraction of structure dependent NMR parameters.

The dissertation presented here stands in direct relationship to this topic of structure determination with NMR parameters. In this work, the first general method for direct 3D-structure refinement using NMR magnetic shielding is introduced. This method is applicable for both solutions and solid states. The basic idea is the application of bond polarization theory, BPT<sup>[6]</sup> in the context of a force field approximation in order to calculate atomic charges and chemical shift tensors. Atomic polarization energy corrections, defined due to deviations between experimental and theoretical one-electron expectation values, are applied in a conventional force field. These atomic polarization energy corrections are the source for pseudo forces which drive the structure during force field simulations towards more reliable conformations. This method was then applied for structure elucidations employing  $^{13}\text{C}$  chemical shifts in solution and solid state. The structure of a pseudotriptide zinc complex in solution could be predicted, using this method. Furthermore, the proton positions of  $\beta$ -D-mannitol in the solid state were determined using crystal simulations, and the structure of a silk model was refined.

## 2. Structure Elucidation using Nuclear Spin

### 2.1. From Wolfgang Pauli to Kurt Wüthrich

The key phenomena which are responsible for making nuclear magnetic resonance possible are the Zeemann effect, which introduces a magnetic field as an essential element in spectroscopy, and the hyperfine structure of spectral lines, which lead to a realization that nuclear moments exist. In 1924, Wolfgang Pauli first succeeded in correctly interpreting hyperfine structures.<sup>[7]</sup> He proposed that atomic nuclei possess an intrinsic angular momentum, a spin and a magnetic moment parallel to its orientation. Its magnitude was assumed to be of the order of the nuclear magneton. It can be obtained by substituting the mass of a nuclei by the mass of an electron in the formula for the atomic Bohr magneton. In 1933, Otto Stern applied a method of molecular beams to the determination of the magnetic moments of the proton and deuteron in hydrogen molecules.<sup>[8,9,10]</sup> The year 1937 was the birth year of nuclear magnetic resonance. Gregory Breit and Isidor Isaac Rabi introduced their method for determining the magnetic properties of atomic nuclei,<sup>[11]</sup> for which they earned the Nobel Prize in 1944. They combined two STERN-GERLACH-Magnets of inhomogeneous fields, which were assembled reversed to each other. The exact magnitude of the field strengths had not to be known. By using nuclei with an odd atomic number, the only interaction between the magnetic field and these nuclei was due to the magnetic moments i.e. the intrinsic spins of the nuclei. The result of the field spin interaction and the selective effect of apertures was a spin state selective registration at the detector. A homogenous magnet with a transition coil was located between the two inhomogeneous fields. By fulfilling the resonance condition

$$\Delta E = \nu h = (M_1 - M_2) \mu B = \mu B , \quad (1)$$

the nuclei performed quantum jumps from one spin state to another ( $M_2 \longrightarrow M_1$ ), thereby landing in reversed positional directions to the homogenous magnetic field. This means that the atoms are not able to reach the detector due to deflection by the following inhomogeneous magnetic field. The effect of these quantum jumps is observable by the fact that the detector

registers a marked resonance minimum with precision of about one part per thousand, see Figure 1. Breit and Rabi could thus obtain a positive proton magnetic moment which differed by a

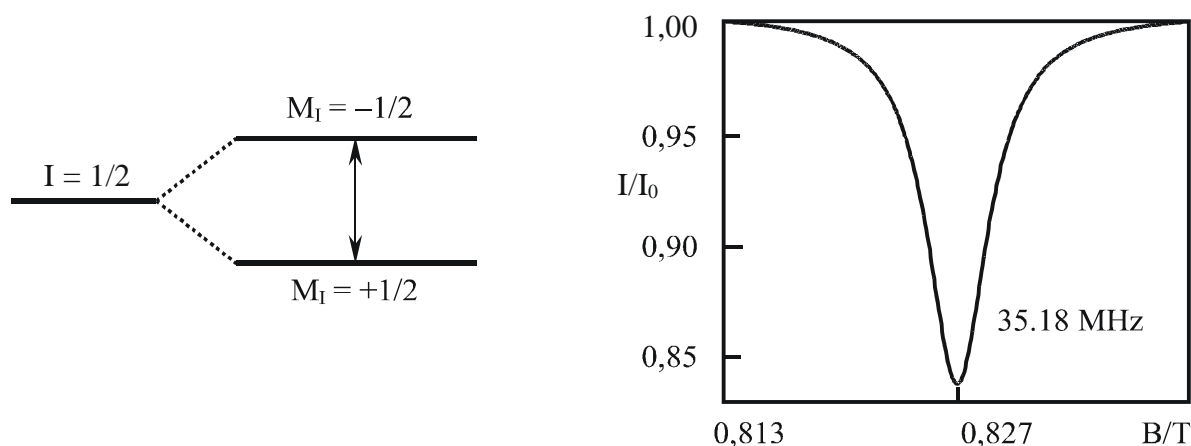


Figure 1  
left: two spin states of protons  
right: signal of the corresponding RABI-Experiment.

factor of roughly 2.85 ( $\pm 0.15$ ) from the theoretical prediction of Pauli:

$$\mu_p = 2.79 \mu_{Bohr} \frac{m_e}{m_p} = 2.79 \frac{e \hbar}{2 m_p} \quad (2)$$

With this accuracy, they also discovered the electric quadrupole moment of the deuterons in 1939.

Today, the principle of the RABI-Experiment is used in the Cesium atomic clock to define the length of a second. This technology for instance, has made the Global Positioning System (GPS) possible. A few years later, Felix Bloch suggested a much simpler way of detecting the reorientation of nuclear moments through the methods of radio reception. Electromagnetic induction caused by nuclear reorientation of a sample in a coil results in a detectable signal. It appears as a voltage difference between the terminals of an external electric circuit. Bloch described this phenomenon as "nuclear induction".<sup>[12,13]</sup> At the same time, Edward Mills Purcell introduced the concept of "nuclear magnetic resonance absorption".<sup>[14]</sup> It turned out that both principles are the same discovery, and is now termed the classical continuous wave NMR experiment which is simply a sample in a resonator located within a swept magnetic field. Both Purcell and Bloch became nobel laureates in 1952 for their development of this method.



In his famous paper,<sup>[13]</sup> Felix Bloch suggested pulsed NMR experiments. Following his suggestions, the first time-domain magnetic resonance experiments were performed 1949 by H.C. Torrey<sup>[15,16]</sup> and, in particular, by Erwin L. Hahn<sup>[17,18,19]</sup> who may be regarded as the true father of pulse spectroscopy. He invented the famous spin echo experiment. Since then, NMR has become an extraordinary method for structure investigation due to the fact that the nuclear sensors are, by nature, extremely well localized having a diameter of a few femtometers. These sensors interact highly sensitive with the local environment, although the corresponding energies are extremely small. The interaction frequency with the static magnetic field is about  $10^9$  Hz, but frequency differences are observable for instance in liquid state up to 0.1 Hz. That corresponds to an energy of  $4.1 \times 10^{-16}$  eV,  $4.0 \times 10^{-11}$  J/mol, or a temperature of  $1.6 \times 10^{-10}$  K. The macroscopic thermodynamic spin ensemble of the sample in the receiver coil can be described by a Boltzmann distribution at room temperature

$$\frac{N_1}{N_2} = e^{-\Delta E/kT} \approx 1 - \frac{h\nu}{kT}, \quad (3)$$

with an energy difference,  $\Delta E$ , about 0.4 J/mol ( $4.1 \times 10^{-6}$  eV, 1.6 K). As can be derived from the occupation number difference ( $N_1 - N_2$ ) at 293 K, it follows that only every  $1.6 \times 10^4$ th nuclei contributes to the NMR signal. Therefore, large numbers of spins are required to discriminate the weak signals from electronic noise.  $10^{14} - 10^{15}$  spins of one kind are needed in order to detect a signal within an experimental time of about one hour under optimum conditions with modern high field NMR spectrometers. Geometrical information can in principle be obtained from nuclear pair and nuclei-electron interactions, but the weakness of the nuclear spin interactions has led to severe detection problems. Four goals had to be achieved in order to promote NMR as a standard tool for molecular structure determination within the last half century:

- (1) optimization of the signal-to-noise ratio,
- (2) development of experiments to extract the enormous amount of molecular information,
- (3) geometric interpretations of the measured NMR data.

The revolution in NMR took place due to the introduction of the Pulse Fourier transform spectroscopy by Weston A. Anderson and Richard R. Ernst,<sup>[20,21,22]</sup> in 1964. It succeeded against numerous other methods as Stochastic Fourier transform spectroscopy,<sup>[23,24,25]</sup> Rapid Scan Fourier transform spectroscopy<sup>[26,27,28]</sup> and Synthesized Waveform Fourier transform spectroscopy.<sup>[29]</sup>

Pulse Fourier transform spectroscopy allows a rigorous analytical treatment, irrespective of the complexity of the spin system, and it is experimentally quite simple. It has unified NMR methodology across all fields, from liquid to solid state resonance, from relaxation measurements to high resolution NMR, and has entered into other fields such as ion cyclotron resonance,<sup>[30,22]</sup> microwave spectroscopy,<sup>[31]</sup> electron paramagnetic resonance<sup>[32]</sup> and laser spectroscopy.<sup>[33]</sup> Furthermore, the application of pulses brought the idea of introducing additional evolution periods. In 1971, Jean Jeener proposed at the Ampere International Summer School in Basko Polje, Yugoslavia, (unpublished) a two-dimensional Fourier transform experiment consisting of two  $\pi/2$  pulses with a variable time  $t_1$  between the pulses and an additional time variable  $t_2$  which measures the free induction decay after the second pulse. This experiment was named COSY, correlation spectroscopy. Later, in 1974, Richard Robert Ernst realized the first 2D-dimensional NMR experiment. He earned the Nobel Prize for his contributions to methodology development of high resolution nuclear magnetic resonance spectroscopy in 1991. Several kinds of experiments were introduced soon after the discovery of the multidimensional NMR spectroscopy:

- (1) correlation experiments,
- (2) exchange experiments,
- (3) separation experiments,
- (4) cross relaxation experiments.

The most important application has been the assignment of the NMR parameters to the atomic sites within a molecular system and its use for 3D structure determination. Most widely used are NOE distances, dipolar coupling distances, vicinal J coupling dihedral angles, and protein backbone/side chain chemical shift dihedral angles, especially in solution NMR. In October 2002, Kurth Wüthrich became a Nobel laureate for his development of nuclear magnetic resonance spectroscopy to determine the three-dimensional structure of biological macromolecules in solution.<sup>[1,2]</sup>

## 2.2. Structure Determination with NMR parameters

### 2.2.1. Basic NMR Theory

Given the quantum mechanical uncertainty relation:

$$\Delta x \Delta p \gtrsim \frac{\hbar}{2}, \quad (4)$$

it seems impossible, at first sight, to detect spatial dependencies within the Å-range with a typical NMR wavelength of  $\lambda \simeq 1$  m and a linear momentum  $p = \frac{h}{\lambda}$ . NMR works in a totally different manner compared to scattering experiments or microscopes with the geometric information being encoded in the spin Hamiltonian  $\hat{H}(q_1, \dots, q_f, t)$ , with  $q_k(t)$  being the nuclear coordinates. An accurate geometric measurement thus becomes an accurate energy measurement with a precision that is dependent on the observation time  $\Delta t$

$$\Delta E \Delta t \gtrsim \frac{\hbar}{2}. \quad (5)$$

The accuracy of NMR measurements is not restricted by the wavelength but rather by the dephasing of coherent nuclear energy states and relaxation-limited lifetimes,  $\tau = \Delta t$ . Coherence times are mostly in the range of milliseconds to seconds, and life times can be up to several hours. The underlying physical law is the quantum mechanical master equation of the spin density.<sup>[34]</sup> Using the superoperator notation for an operator  $\hat{O}$

$$\left[ \hat{O}, \hat{\sigma}(t) \right] = \hat{\tilde{O}} \hat{\sigma}(t), \quad (6)$$

the spin density can be represented as

$$\frac{d}{dt} \hat{\sigma}(t) = \left( -\frac{i}{\hbar} \hat{\tilde{H}} - \hat{\tilde{\Gamma}} \right) (\hat{\sigma}(t) - \hat{\sigma}_o). \quad (7)$$

$\hat{\tilde{\Gamma}}$  is the superoperator of the relaxation operator  $\hat{\Gamma}(q_1, \dots, q_f, t)$  and  $q_k(t)$  are the time dependent nuclear coordinates. The geometric information contained within  $\hat{\tilde{\Gamma}}$  is, in praxis, as important

as the Hamiltonian  $\hat{H}$  itself. The formal solution of (7) is given by

$$\hat{\sigma}(t) = \hat{\sigma}_o + \hat{T} \exp \left( - \int_0^t \left( \frac{i}{\hbar} \hat{H} + \hat{\Gamma} \right) dt \right) (\hat{\sigma}(0) - \hat{\sigma}_o) , \quad (8)$$

where  $\hat{T}$  is the time ordering operator. The information provided by the spin density can be measured by the free induction decay of the receiver coil and can be calculated for a standard NMR quadrature detector with the formula

$$s(t) \sim Tr \left[ \left( \hat{I}_x - i\hat{I}_y \right) \hat{\sigma}(t) \right] . \quad (9)$$

This concept can be easily generalized to include multidimensional NMR spectroscopy by

$$s(t_1, t_2, \dots, t_n) \sim Tr \left[ \left( \hat{I}_x - i\hat{I}_y \right) \hat{\sigma}(t_1, t_2, \dots, t_n) \right] . \quad (10)$$

Theoretical simulations of the multidimensional free induction decay can be carried out, for example, with the program packages GAMMA<sup>[35]</sup> and SIMMOL/SIMPSON.<sup>[36,37]</sup> Multidimensional experimental and theoretical spectra are obtained by Fourier transformation

$$S(\omega_1, \omega_2, \dots, \omega_n) \sim \int_0^\infty \int_0^\infty \dots \int_0^\infty s(t_1, t_2, \dots, t_n) e^{-i\omega_1 t_1} e^{-i\omega_2 t_2} \dots e^{-i\omega_n t_n} dt_n \dots dt_2 dt_1 . \quad (11)$$

This again, implies the uncertainty relation corresponding to (5)

$$\Delta\omega \Delta t \gtrsim \frac{1}{2} . \quad (12)$$

### 2.2.2. Basic NMR Hamiltonian

The basic NMR Hamilton operator consist of two parts

$$\hat{H} = \hat{H}_{ext} + \hat{H}_{int} . \quad (13)$$

$\hat{H}_{ext}$  is the external contribution arising from nuclei interactions with classical macroscopic

magnetic fields:

$$\widehat{H}_{ext} = \widehat{H}_0 + \widehat{H}_{RF}, \left\{ \begin{array}{l} \widehat{H}_0 = -\gamma B_z \hbar \widehat{I}_z \\ \widehat{H}_{RF} = -\gamma B_1 \hbar \left( \widehat{I}_x \cos(\omega_1 t + \xi) + \widehat{I}_y \sin(\omega_1 t + \xi) \right) \end{array} \right\}, \quad (14)$$

where  $\omega_1$  is the rotation frequency of the time dependent magnetic field  $\vec{B}_1$  of the transition coil. The second part of the NMR operator  $\widehat{H}_{int}$  results from atomic and molecular contributions:

$$\widehat{H}_{int} = \widehat{H}_{MS} + \widehat{H}_{DD} + \widehat{H}_J + \widehat{H}_Q, \quad (15)$$

where  $\widehat{H}_{MS}$  is magnetic shielding Hamiltonian,  $\widehat{H}_{DD}$  is the direct dipole-dipole interaction,  $\widehat{H}_J$  is the indirect dipole-dipole contribution, and  $\widehat{H}_Q$  is the quadrupole interaction. The informational content of the total internal nuclear Hamilton superoperator  $\widehat{\overline{H}}$  and the relaxation superoperator  $\widehat{\overline{\Gamma}}$  for large molecules, e.g. proteins with a molecular mass up to 100 kD, is immense. It is possible to determine the chemical shifts of thousands of atomic sites in a molecular system to an accuracy of  $10^{-16}$  with respect to the Rydberg energy. Internuclear distances of thousands of proton pairs can thus be measured to about a 0.1 Å accuracy and several hundred dihedral angles can, in principle, be determined with an uncertainty of less than  $10^\circ$ .

### 2.2.3. Magnetic Shielding and Chemical Shift

As early as in 1950, it was observed that NMR spectral frequencies of a particular species of nuclei in different chemical environments may be different. Furthermore, this difference depends upon the molecular structure. The electrons of the system interact with the external magnetic field and shield or deshield this field. The magnetic shielding Hamiltonian has the form

$$\widehat{H}_{MS} = \sigma \vec{B} \gamma \hbar \widehat{I}_z = \sigma \vec{B} \widehat{\mu}_z, \quad \vec{B} = (0, 0, B_z) \quad (16)$$

Applying the high field approximation,  $B_z \gg \left| \sigma \vec{B} \right|$ , the NMR chemical shielding Hamiltonian has the following form in the principal value system

$$\widehat{H}_{MS} = \omega_{MS} \hbar \widehat{I}_z, \quad (17)$$

## 2. Structure Elucidation using Nuclear Spin

with

$$\omega_{MS} = \gamma B_z \left( \sigma_{iso} + \frac{\Delta_\sigma}{3} [(3 \cos^2(\theta) - 1) - \eta_\sigma \sin^2(\theta) \cos(2\varphi)] \right). \quad (18)$$

$\theta$  and  $\varphi$  are the polar angles of the principal axis system with respect to  $\vec{e}_z$ . The three independent parameters  $\sigma_{iso}$ ,  $\Delta_\sigma$  and  $\eta_\sigma$  are obtained from the principal tensor values  $\sigma_{11}$ ,  $\sigma_{22}$  and  $\sigma_{33}$ , in order to fulfill the following condition:<sup>[38]</sup>

$$|\sigma_{33} - \sigma_{iso}| \geq |\sigma_{11} - \sigma_{iso}| \geq |\sigma_{22} - \sigma_{iso}|. \quad (19)$$

The isotropic value is given by

$$\sigma_{iso} = \frac{1}{3} (\sigma_{11} + \sigma_{22} + \sigma_{33}), \quad (20)$$

while the anisotropic value can be calculated as

$$\Delta_\sigma = \frac{3}{2} (\sigma_{33} - \sigma_{iso}). \quad (21)$$

The asymmetry is defined as:

$$\eta_\sigma = \frac{\sigma_{22} - \sigma_{11}}{\sigma_{33} - \sigma_{iso}}, \quad 0 \leq \eta \leq 1. \quad (22)$$

In liquid state, the spatial contributions  $(3 \cos^2(\theta) - 1)$  and  $\sin^2(\theta) \cos(2\varphi)$  average to zero due to Brownian motion thus, leaving only the isotropic value. A residual magnetic shielding tensor occurs only when small degrees of alignment are present such as in liquid crystals, solutions with slow molecular motions or molecular systems with notable magnetic susceptibility tensors.<sup>[39]</sup>

The frequency is given for an uniaxial phase by

$$\begin{aligned} \omega_{MS} = & \gamma B_0 \left( \sigma_{iso} + \frac{2}{3} (\sigma_{zz} - \frac{1}{2} (\sigma_{xx} + \sigma_{yy})) \overline{D_{0,0}^2(\theta, \varphi)} + \right. \\ & \left. \frac{1}{3} (\sigma_{xx} - \sigma_{yy}) \left( \overline{D_{0,2}^2(\theta, \varphi)} + \overline{D_{0,-2}^2(\theta, \varphi)} \right) \right) \hat{I}_z, \end{aligned} \quad (23)$$

$\overline{D_{m,n}^I(\theta, \varphi)}$  are the time-averaged Wigner rotation matrices,  $\theta$  and  $\varphi$  are the angles between the molecular frame (the principal value system of the order matrix) and the magnetic field. If the magnetic shielding tensor is given in the principal value system, it can be transformed into a

## 2. Structure Elucidation using Nuclear Spin

molecular frame of reference with the use of the direction cosines

$$\sigma_{\alpha,\beta} = \sum_{i,j} \sigma_{i,j} \cos(\theta_{\alpha,i}) \cos(\theta_{\beta,j}) \quad , \quad \begin{cases} \alpha, \beta = x, y, z \\ i, j = 1, 2, 3 \end{cases} . \quad (24)$$

In NMR spectroscopy, the chemical shifts  $\delta$  are extracted from the spectra with the frequency  $\nu$  instead of the magnetic shieldings  $\sigma$  :

$$\delta [ppm] = 10^6 \frac{\nu - \nu_{ref}}{\nu_{ref}} = \sigma_{ref} - \sigma . \quad (25)$$

These shifts  $\delta$  are calculated from the shieldings with respect to standardized nuclei specific references  $\sigma_{ref}$ . The magnetic shielding Hamiltonian can be derived in terms of the chemical shift and the frequency resembles:

$$\omega_{MS} = \gamma B_0 \left( \sigma_{ref} - \delta_{iso} - \frac{\Delta\delta}{3} [(3 \cos^2(\theta) - 1) - \eta_\delta \sin^2(\theta) \cos(2\varphi)] \right) . \quad (26)$$

With the help of equation (26) the magnetic shielding Hamiltonian can be expressed in terms of the chemical shift. From this the chemical shift Hamiltonian could be defined, although most NMR text books give a different definition.

So far, the Hamiltonian has been discussed in terms of the NMR experiment for extraction of the magnetic shielding or chemical shift. An outline for their theoretical calculation will now be given. The magnetic shielding  $\sigma$  is a tensor property and can be represented as an expectation value

$$\sigma = \langle \Psi_0 | \hat{\sigma}_N | \Psi_0 \rangle . \quad (27)$$

The wave function  $|\Psi_0\rangle$  is the solution of the Hamilton operator or Fock operator without external perturbations, and belongs to the ground state energy  $E_0$ . The shielding operator  $\hat{\sigma}_N$  yields the general form within the Coulomb gauge restriction introduced by Ramsey in 1950<sup>[40,41]</sup>

$$\hat{\sigma}_N = \hat{h}_N^{dia} - \frac{2}{m_e} \sum_{k \neq 0} \frac{\hat{h}_N^{orb} |\Psi_k^{(0)}\rangle \langle \Psi_k^{(0)}| (\hat{h}_N^{pso})^T}{E_k - E_0} . \quad (28)$$

This expression contains the excited states  $|\Psi_k^{(0)}\rangle$  belonging to the energy  $E_k$ , and the operators  $\hat{h}_N^{orb}$ ,  $\hat{h}_N^{pso}$  and  $\hat{h}_N^{dia}$ <sup>[42]</sup> which describe the involved electron interactions. The coupling of the

## 2. Structure Elucidation using Nuclear Spin

external field to orbital motion is expressed by a sum of orbital angular-momentum operators

$$\hat{h}_N^{orb} = \frac{1}{2a_o} \sum_k \hat{l}_{kO} . \quad (29)$$

The nuclear spin-orbit operator or orbital hyperfine operator describes the coupling of the nuclear magnetic moments to the orbital motion of the electrons

$$\hat{h}_N^{psO} = \alpha^2 a_o^2 \sum_k \frac{\hat{l}_{kN}}{r_{kN}^3} . \quad (30)$$

The diamagnetic electronic operator has the form

$$\hat{h}_N^{dia} = \frac{\alpha^2}{2} a_o \sum_k \frac{(\hat{r}_{kO} \hat{r}_{kN}) \hat{I} - (\hat{r}_{kN} \hat{r}_{kO}^T)}{r_{kN}^3} , \quad (31)$$

with Sommerfeld's fine structure constant  $\alpha = \frac{1}{4\pi\epsilon_0} \frac{e^2}{\hbar c} \approx \frac{1}{137}$ , and Bohr radius  $a_o = 0.53 \text{ \AA}$ . The shielding operator can be written as a sum of one-electron operators. The second order perturbation theory is included into the shielding operator itself (equation (28)). In principle, the calculation of this NMR parameter does not depend on the exact wave function model.

The Ramsey expression cannot be employed for most calculations, since it requires an explicit representation of the excited states. In practice, variational methods are used because all conventional *ab initio* methods can be cast in a variational form (HF, CI, MCSCF, CC, MP and DFT<sup>[43,44,45]</sup>). The magnetic shielding energy is a correction (equation (16)) to the total molecular energy. The magnetic shielding can be represented by a Taylor expansion of the total energy with respect to the magnetic field  $\vec{B}$  and the nuclear magnetic moment  $\vec{\mu}$ :

$$\begin{aligned} E(\vec{B}, \vec{\mu}) &= E(0) + \sum_i \left( \frac{dE}{dB_i} \right)_{B=0} B_i + \sum_i \left( \frac{dE}{d\mu_i} \right)_{\mu=0} \mu_i + \\ &\sum_{i,j} \left( \frac{d^2 E}{dB_j dB_i} \right)_{B=0} B_j B_i + \sum_{i,j} \left( \frac{d^2 E}{d\mu_j d\mu_i} \right)_{\mu=0} \mu_j \mu_i + \\ &\sum_{i,j} \left( \frac{d^2 E}{dB_j d\mu_i} \right)_{B,\mu=0} B_j \mu_i + \dots \end{aligned} \quad (32)$$

The second derivative of the energy with respect to the magnetic field and the nuclear magnetic



moment can be identified as the magnetic shielding, see equation (16)

$$\sigma_{ij} = \left. \frac{d^2 \langle \Psi | \hat{H} | \Psi \rangle}{dB_j d\mu_i} \right|_{B, \mu=0} . \quad (33)$$

The GIAO approach<sup>[46]</sup> will be discussed here, since it has become the standard for the calculation of magnetic properties.

The Hamiltonian of a molecular system can be constructed out of a sum of single electron (atomic orbital) operators  $\hat{h}_{ao}$ , and electron-electron interaction terms  $\hat{v}_{oa,ao'}$

$$\hat{H} = \sum_{ao}^{AOs} \hat{h}_{ao} + \sum_{ao}^{AOs} \sum_{ao'}^{AOs} \hat{v}_{oa,ao'} , \quad (34)$$

with

$$\hat{h}_{ao} = \frac{\hat{p}_{ao}^2}{2m_e} + \frac{e}{mc} \hat{A}_{ao} \hat{p}_{ao} + \frac{e^2}{2mc^2} \hat{A}_{ao}^2 + v(\vec{r}_{ao}) . \quad (35)$$

The vector potential  $\hat{A}_{ao}$  is localized at the atomic orbital position. It can be expressed by introducing a constant magnetic field  $\vec{B}$

$$\hat{A}_{ao} = \frac{1}{2} \vec{B} \times (\vec{r}_{ao} - \vec{R}_o) , \quad (36)$$

with  $\vec{r}_{ao}$  are the centers of the atomic orbitals and  $\vec{R}_o$  is the gauge origin - an arbitrary parameter. It describes a shift of the gauge origin from  $\vec{R}_o$  to  $\vec{r}_{ao}$ . Thus, the vector potential satisfies the (Coulomb gauge) condition

$$\vec{\nabla} \cdot \hat{A}_{ao} = 0 . \quad (37)$$

The atomic orbital wave functions are thus transformed

$$|\chi_{ao}(0)\rangle \mapsto |\chi_{ao}(B)\rangle = \exp\left(-\frac{ie}{2c\hbar} \vec{B} \times (\vec{r}_{ao} - \vec{R}_o)\right) |\chi_{ao}(0)\rangle . \quad (38)$$

The wave functions  $|\chi_{ao}(B)\rangle$  are called gauge including atomic orbitals (GIAOs) or London

orbitals. Introduction of these in the Hamiltonian results in

$$\begin{aligned} \hat{h}_{ao} = & -\frac{\hbar^2 \vec{\nabla}_{ao}^2}{2m_e} - \frac{ie\hbar}{2mc} \vec{B} \cdot \left( (\vec{r}_{ao} - \vec{R}_o) \times \vec{\nabla}_{ao} \right) + \\ & \frac{e^2}{8mc^2} \left( \vec{B} \times (\vec{r}_{ao} - \vec{R}_o) \right)^2 + v(\vec{r}_{ao}) . \end{aligned} \quad (39)$$

As an example the total molecular wave function  $|\Psi\rangle$  can be generated within the CC (coupled cluster) theory,<sup>[47,48,49]</sup> using an exponential excitation operator,  $\exp(\hat{T})$ , which is applied to a Slater determinant

$$|\Psi\rangle = \frac{1}{\sqrt{N!}} \exp(\hat{T}) \begin{vmatrix} |\Phi_1(1)\rangle & |\Phi_1(2)\rangle & \cdots & |\Phi_1(N)\rangle \\ \vdots & \vdots & \ddots & \vdots \\ |\Phi_N(1)\rangle & |\Phi_N(2)\rangle & \cdots & |\Phi_N(N)\rangle \end{vmatrix} \quad (40)$$

of atomic spin orbitals

$$|\Phi_j(k)\rangle = \exp\left(-\frac{ie}{2c\hbar} \vec{B} \times (\vec{r}_k - \vec{R}_o)\right) |\eta(\sigma_k)\rangle \sum_{ao} c_{j,ao} |\chi_{ao,k}(0)\rangle . \quad (41)$$

The GIAOs,  $|\chi_{ao,k}(B)\rangle$  imply the major effect of the magnetic field perturbation on the wave function. The remaining corrections of the CC and MO coefficients are rather small, and a rapid convergence within the energy minimization technique is ensured. Once the molecular wave function is known, the magnetic shielding can be calculated by the formalism given with equation (33). *Ab initio* calculations can be carried out with program packages as for instance GAUSSIAN98,<sup>[50]</sup> GAMESS,<sup>[51]</sup> DALTON<sup>[52]</sup> or ACES II.<sup>[53]</sup>

#### 2.2.4. Direct Nuclear Dipole-Dipole Interaction

The most obvious NMR parameter which contains geometric information is the nuclear spin interaction which can be represented in a high magnetic field approximation for homonuclear coupling as

$$\hat{H}_D^H = -\omega_D \hbar \frac{1}{2} \left( 3\hat{I}_{1z}\hat{I}_{2z} - \hat{I}_1\hat{I}_2 \right) . \quad (42)$$

The heteronuclear coupling is given by

$$\hat{H}_D^{IS} = -\omega_D \hbar \hat{I}_z \hat{S}_z . \quad (43)$$

For both equations,

$$\omega_D = \frac{\mu_0}{4\pi} \frac{\gamma_1 \gamma_2}{r_{12}^3} \hbar (3 \cos^2(\theta) - 1) . \quad (44)$$

$\theta$  is the angle between the internuclear vector and the direction of the magnetic field.<sup>[54]</sup> The spatial contribution averages to zero in liquid state, and the dipolar coupling vanishes. A long range residual dipolar coupling is left for small degrees of alignment.<sup>[55,56,57,39,58,59]</sup> The dipolar frequency of a spin pair is given for an uniaxial phase by

$$\omega_D = \frac{\mu_0}{4\pi} \frac{\gamma_1 \gamma_2}{r_{12}^3} \hbar f(\theta, \varphi, \alpha, \beta) , \quad (45)$$

with

$$f = \overline{D_{0,0}^2(\theta, \varphi)} (3 \cos^2(\theta) - 1) + \left( \overline{D_{0,2}^2(\theta, \varphi)} + \overline{D_{0,-2}^2(\theta, \varphi)} \right) \sqrt{\frac{3}{2}} \sin^2(\beta) \cos(2\alpha) . \quad (46)$$

The time averaged Wigner rotation matrices  $\overline{D_{m,n}^l(\theta, \varphi)}$  are the order parameters. They depend on the polar angles  $\theta$  and  $\varphi$  of the molecular frame with respect to the laboratory frame.  $\alpha$  and  $\beta$  are the polar angles of the inter nuclear dipole vector within the molecular frame. A reference frame, the alignment frame, is thus available and orientations are possible with respect to this frame irrespective of the number of intervening bonds. The use of the residual dipolar couplings is becoming more common as a supplement to NOEs as well as scalar couplings for the refinement of high resolution NMR solution structures. It could very well be that residual chemical shift anisotropies, time averaged quadrupolar couplings and possibly residual anisotropic J coupling contributions will be employed more and more in modern solution NMR refinement techniques.

### 2.2.5. Indirect Nuclear Dipole-Dipole Interaction - J Coupling

The total electron spin vanishes in diamagnetic substances, and there is no coupling between nuclear and electron spins in first order perturbation theory. However, nuclear spins interact among themselves via the local electron spin density. This indirect coupling exists in second order perturbation theory, which has a similar form as the dipolar coupling<sup>[60]</sup>

$$\omega_J = \frac{C_{12} \left( \frac{-}{r} \right)}{r_{12}^3} P(\cos(\theta)) , \quad (47)$$

where  $C_{12}$  is a complicated function depending on the spin separation and orientation of the molecule in the external field.<sup>[61]</sup> Therefore the J coupling is not averaged to zero in solution or by solid state magic angle spinning, it has an isotropic contribution. In his famous paper,<sup>[4]</sup> Karplus described the relation between the protein backbone dihedral angle  $\varphi$  and the three-bond isotropic coupling, the so-called vicinal coupling,

$${}^3\omega_J = A \cos^2(\varphi) + B \cos(\varphi) + C . \quad (48)$$

Several sets of parameters were established<sup>[62,63,64,65,66,67,68]</sup> for different kinds of vicinal couplings. Furthermore, an empirical correlation has been found between an one-bond coupling and two backbone dihedral angles.<sup>[69]</sup> An analytical derivation for this empirical observation has not yet been successful. In general, J couplings are rather small in comparison to other NMR interactions, and can often be neglected in solid state NMR.

## 2.2.6. Nuclear Electric Quadrupole Interaction

Nuclei with spin  $I > \frac{1}{2}$  possess a nuclear electric quadrupole moment  $eQ$  which becomes larger with increasing atomic number. This quadrupole moment interacts with non-spherical symmetric electric field gradients which results from an electric field at the nuclei position generated by the surrounding electronic charge distribution of the atom or molecule. Most common nuclei have small proton numbers, and their quadrupolar Hamiltonian is much smaller than the Zeeman contribution, i.e. it can be truncated. The first order quadrupole Hamiltonian can be written similar to the homonuclear dipole interaction in the principal axis system of the electric field gradient

$$\hat{H}_Q = \omega_Q \hbar \frac{1}{2} \left( 3\hat{I}_z^2 - \hat{I}^2 \right) , \quad (49)$$

where the quadrupolar frequency  $\omega_Q$  is defined as

$$\omega_Q = \frac{eQ V_{zz}}{2 I (2I - 1) \hbar} \left[ (3 \cos^2(\theta) - 1) - \eta_Q \sin^2(\theta) \cos(2\varphi) \right] , \quad (50)$$

and the asymmetry parameter  $\eta_Q$  is given as

$$\eta_Q = \frac{V_{yy} - V_{xx}}{V_{zz}} , \quad (51)$$

with the electric field gradient being

$$V_{ij} = \left. \frac{\partial^2 V}{\partial x_i \partial x_j} \right|_0 = \frac{1}{4\pi\epsilon_0} \int \frac{\partial^2}{\partial x_i \partial x_j} \left( \frac{e \langle \Psi_0(\vec{r}') | \Psi_0(\vec{r}') \rangle}{|\vec{r} - \vec{r}'|} \right) d\vec{r}'^3. \quad (52)$$

It is often employed

$$V_{zz} = eq. \quad (53)$$

The molecular ground state wave function  $|\Psi_0\rangle$  and the electric field gradient can be calculated, for instance with GAUSSIAN98,<sup>[50]</sup> GAMESS,<sup>[51]</sup> DALTON<sup>[52]</sup> or ACES II.<sup>[53]</sup>

A residual interaction in solution can be measured with small degrees of order for the first order quadrupolar coupling. It is similar to the dipolar coupling

$$\omega_Q = \frac{3 e^2 q Q}{8 I (2I - 1) \hbar} f(\theta, \varphi, \alpha, \beta), \quad (54)$$

with

$$f = \overline{D_{0,0}^2(\theta, \varphi)} (3 \cos^2(\theta) - 1) + \left( \overline{D_{0,2}^2(\theta, \varphi)} + \overline{D_{0,-2}^2(\theta, \varphi)} \right) \sqrt{\frac{3}{2}} \sin^2(\beta) \cos(2\alpha). \quad (55)$$

### 2.2.7. Nuclear Overhauser Effect (NOE)

In 1953, A. Overhauser predicted that a small alteration in the electron spin population would produce a large change in the nuclear spin polarization. This behavior was named the Overhauser effect and has been adapted to NMR. Nowadays, it is a very important tool for the determination of complex molecular structures. NMR NOE spectroscopy (NOESY)<sup>[70]</sup> is performed by introducing a third  $\pi/2$  pulse into the simple COSY experiment direct after the indirect time dimension  $t_1$ , and before the last  $\pi/2$  pulse. These two pulses are separated by a mixing time of about 0.2 s during which random processes occur. These incoherent, dissipative processes drive the system back to equilibrium. Inter nuclear cross relaxation takes place mainly due to magnetic dipolar interactions during this mixing time. The NOE signal is proportional to  $\langle r^{-6} \rangle$  for rigid molecules assuming the molecules are slowly tumbling.<sup>[3]</sup> The rotational correlation time  $\tau_c$  is in the order of ns for large proteins in solution. The general assignment procedure of protons based on COSY and NOESY spectra as well as the computer procedures to determine

molecular structures based on NOE signals and J couplings were established by Wüthrich and his research group.<sup>[1,71]</sup> In fact, the NOE NMR parameter was the breakthrough of NMR in 3D structure determination of macromolecules in solution.

### 2.2.8. Structure Determination with Force Fields

Beside X-ray analysis the most common strategy of structure determination is the application of a force field with NMR constraints. There are three major contributions to the energy,

$$U^{total} = U^{bond} + U^{non-bond} + U^{NMR}, \quad (56)$$

bonded energies, non-bonded energies and pseudo potentials (constraint energies) which result from the NMR restrictions. One of the simplest description of bonded interactions includes potential energies for bond stretching, angle bending, torsional rotation, and out-of-plane (improper) deformations. There exists a whole variety of methods for the description of non-bonded inter molecular as well as intra molecular interactions. As an example, a Lennard-Jones 9-6 or an exp-6 term can be used to describe van der Waals interactions. A Coulomb energy function can be employed to describe the electrostatic interaction. Force fields which disregard anharmonic and coupling interactions are classified by Dinur and Hagler as Class I force fields.<sup>[72]</sup> The most important examples for biomolecular/bioorganic applications<sup>[73]</sup> are CHARM,<sup>[74,75]</sup> Quanta/CHARMm,<sup>[76,77,78]</sup> AMBER,<sup>[79]</sup> OPLS/AMBER,<sup>[80]</sup> GROMOS,<sup>[81]</sup> and Tripos.<sup>[82]</sup> Some force fields, for instance CHARM, introduce an additional combined bond stretch/bond angle term, the Urey-Bradley term,<sup>[83]</sup> which is especially important when vibrational spectra have to be calculated. Only force fields containing anharmonic and coupling interactions between terms are classified as Class II force fields. Examples of these force fields are MM2,<sup>[84]</sup> MM3,<sup>[85]</sup> MM4,<sup>[86]</sup> MMFF,<sup>[87]</sup> and CFF<sup>[88]</sup> which are also most popular for organic applications. In principle, Class II force fields offer a better description of the potential energy surface, at the cost of more computer power. The energy of bond stretching is described with the computationally-expensive Morse function,<sup>[89]</sup> for instance

$$U^{Morse} = D \left( 1 - e^{\alpha (r-R)} \right), \quad (57)$$

where  $D$  is the dissociation energy and  $\alpha$  is a force constant which controls the width of the potential well. The Morse function can be approximated by a harmonic energy function only

when  $r$  is very close to the equilibrium bond length  $R$ . Most Class II force fields use a more accurate quartic bond potential, and MM4 even has a sixth power bond potential. In this thesis, the COSMOS force field<sup>[90]</sup> is applied, which consists of Class II force field contributions and employs a semi-empirical approach for calculating coordinate dependent charges for the electrostatic interactions.

The computational effort and hence CPU time of *ab initio* and density functional methods is proportional to the number of basis functions to the power of three to four,  $B^{3...4}$ . In contrast to this, force field calculation of bonded interactions is only proportional to the number of atoms  $N$ . The non-bonded interaction terms run over all atom pairs, and hence the computational cost is proportional to  $N^2$ . Therefore, the speed of force field simulations can be further increased when one increases the efficiency of the non-bonded calculations. The most simple and most widely used method is to apply a cut-off distance beyond which non-bonded interactions are not calculated. The cost of the search for all atom pairs within the cut-off distance is also proportional to  $N^2$ . Therefore, a non-bonded list is often maintained with all neighboring atoms within the cut-off distance. The electrostatic interaction is proportional to  $r^{-1}$ , and is thus long ranged. The use of a cut-off radius causes discontinuities in the electrostatic interactions at the cut-off distance.<sup>[91]</sup> The result is a noise source in molecular dynamics simulations which increases the kinetic energy and respectively the temperature. A common way to reduce the noise is to multiply the non-bonded interaction terms with smoothing functions.<sup>[92]</sup>

The NMR constraints force the molecule to adopt a structure that is compatible with the NMR data. Harmonic approximations are often used to accomplish this

$$U^{NMR} = k \left( \langle \hat{P} \rangle - P^E \right)^2, \quad (58)$$

where  $k$  is a force constant,  $P^E$  is the experimental and  $\langle \hat{P} \rangle = P(\vec{r})$  the theoretical value of a NMR parameter, that is coordinate dependent. There are several packages for structure determination in solution, based on conformational constraints from NMR data, i.e. X-PLOR,<sup>[93]</sup> CYANA,<sup>[71]</sup> CHARMM<sup>[94]</sup> and CONGEN.<sup>[95]</sup> The COSMOS force field<sup>[90,96]</sup> which applies analytically derived energetic corrections as constraints is used in this work.

### 3. Bond Polarization Theory (BPT)

#### 3.1. Introduction

Localized bond orbitals are a suitable form to represent the molecular wave function if a quantum chemical method is combined with a force field. A bond orbital can be constructed for every bond defined within the force field. Anti-bonds are introduced to account for polarizations. Delocalizations can be treated in the same manner. These bond orbitals can be used to construct a molecular wave function for the solution of the energy equation of the Hamilton operator by employing a perturbation series

$$|\Psi_0\rangle = |\Phi_0\rangle + \sum_{i \neq 0} \frac{\langle \Phi_i | \hat{H} | \Phi_0 \rangle}{E_0 - E_i} |\Phi_i\rangle + \dots, \quad (59)$$

where the molecular ground state wave function  $|\Phi_0\rangle$  of the energy  $E_0$  is represented by a Slater determinant constructed from ideal bond orbitals. The excited state configurations  $|\Phi_i\rangle$  of the energies  $E_i$  are obtained by polarization and delocalization of the ground state wave function. The main idea of the bond polarization theory (BPT)<sup>[6]</sup> is to reduce the total quantum mechanical problem to polarization effects only. Delocalizations are neglected, and the Hamiltonian  $\hat{H}$  is approximated by interacting point charges.

#### 3.2. Theoretical Background

Molecular systems are thought to consist of bonds between two atoms which can be described by localized two centered bond orbitals that are linear combinations of appropriate hybrid atomic orbitals. The polarity of the bond is determined by the polarity parameter  $d$ . This polarity parameter is not explicitly needed within the BPT framework. However, it is the only free parameter for the construction of the bonds

$$|i\rangle = \sqrt{\frac{1+d}{2}} |\chi_A\rangle + \sqrt{\frac{1-d}{2}} |\chi_B\rangle \quad (60)$$



as well as the anti-bonds

$$|i^*\rangle = \sqrt{\frac{1-d}{2}} |\chi_A\rangle - \sqrt{\frac{1+d}{2}} |\chi_B\rangle . \quad (61)$$

Bonds and anti-bonds are created from hybrid orbitals

$$\langle \vec{r} | \chi \rangle = \chi(\vec{r}) = \sum_k h_k \phi_k(\zeta, \vec{r}) , \quad (62)$$

and they are assumed to be orthogonal (zero overlap approximation<sup>[97]</sup>). The basis functions of the hybrids are Slater type atomic orbitals (STOs<sup>[98]</sup>)

$$\begin{aligned} \langle \vec{r} | \phi_{nlm} \rangle &= \phi_{nlm}(\zeta, \vec{r}) \\ &= \sqrt{\frac{(2\zeta)^{2n+1}}{(2n)!}} r^{n-1} \exp(-\zeta r) Y_{lm}(\theta, \varphi) , \end{aligned} \quad (63)$$

where the orbital exponents  $\zeta$  are taken from Burns.<sup>[99]</sup> The functions  $Y_{lm}(\theta, \varphi)$  are spherical harmonics and depend on the polar angles  $\theta$  and  $\varphi$ . The asymptotic behaviors ( $r \rightarrow 0$  and  $r \rightarrow \infty$ ) of the STOs are similar to the solutions of the hydrogen atom.

It is common to identify the Hamiltonian of equation (59) with the Fock operator  $F$ . The most important term of the configuration interaction is clearly the polarization. This is indicated by rewriting the excited wave functions as a polarization state

$$|\Phi_i\rangle = |\Phi(i^*)\rangle . \quad (64)$$

If we restrict the perturbation series to this first order polarization term, it is easy to obtain an expression for the expectation value of an one-electron operator  $\hat{O}$  using equation (59)

$$\langle \Psi_0 | \hat{O} | \Psi_0 \rangle = \langle \Phi_0 | \hat{O} | \Phi_0 \rangle + \sum_{i \neq 0} \frac{\langle \Phi(i^*) | \hat{O} | \Phi_0 \rangle}{E_0 - E_{i^*}} \langle \Phi_0 | \hat{F} | \Phi(i^*) \rangle + \dots \quad (65)$$

The Fock operator  $\hat{F}$  can be separated into two parts,  $\hat{1}$  and  $\hat{2}$

$$\hat{F} = \hat{F}_1 + \hat{F}_2 . \quad (66)$$

Part  $\hat{1}$  deals only with the bond being considered and  $\hat{2}$  describes the rest of the system. This splitting can be performed because delocalizations from  $\hat{1}$  to  $\hat{2}$  are neglected. By employing

the same ideas as found in the PCILO method of Malrieu,<sup>[100]</sup> it is assumed that the polarity parameters  $d$  can be adjusted in order that the following requirement

$$\langle k_{\hat{1}} | \hat{F}_{\hat{1}} | k_{\hat{1}}^* \rangle = 0 \quad (67)$$

is fulfilled. From equation (65) it follows

$$\langle \Psi_0 | \hat{O} | \Psi_0 \rangle = 2 \sum_k \left( \langle k_{\hat{1}} | \hat{O} | k_{\hat{1}} \rangle + 2 \frac{\langle k_{\hat{1}}^* | \hat{O} | k_{\hat{1}} \rangle}{E_0 - E_{i^*}} \langle k_{\hat{1}} | \hat{F}_{\hat{2}} | k_{\hat{1}}^* \rangle \right). \quad (68)$$

The sum runs over all bonds  $\hat{1}$  in the molecule which can be decomposed into two sum terms, the first runs over all atoms  $A$  and the second runs over all bonds belonging to the atom

$$\sum_k \rightarrow \sum_A \sum_{i \in A}. \quad (69)$$

Within the BPT approach, the Fock operator  $\hat{F}_{\hat{2}}$  is approximated by a point charge distribution  $\hat{V}_{\hat{2}}$  of system  $\hat{2}$

$$\hat{F}_{\hat{2}} \approx \hat{V}_{\hat{2}} = \frac{1}{4\pi\epsilon_0} \sum_{x \in \hat{2}} \frac{q_x}{|\vec{R}_x - \vec{r}|}. \quad (70)$$

$q_x$  are the net atomic charges located at the atomic nuclei positions. So far, double occupied bond orbitals are only suitable for the description of  $\sigma$ -bonds and ideal  $\pi$ -bonds. In order to describe molecular systems with delocalized  $\pi$ -bonds, the occupation number adopts a value between one and two. Double bonds are treated as superimposed resonance structures. There is no easy way to obtain the occupation number theoretically. The occupation number of conjugated  $\pi$ -bonds is estimated from an empirical valence formula<sup>[101]</sup>

$$\nu = e^{(R-r)/0.37\text{\AA}}. \quad (71)$$

The occupation number of a double bond is defined as  $n = 2\nu$ , the  $\pi$ -bond contribution to this number becomes  $n = (2\nu - 1)$ . It depends on the equilibrium length of an ideal single bond  $R$  and on the actual bond length  $r$  (bond length are given in Å). By taking expression (68) and introducing bond specific constants one can obtain equation (72) for an one-electron operator

at nucleus position  $N$

$$\begin{aligned} \langle \Psi_0 | \hat{O} | \Psi_0 \rangle &= \sum_A \sum_{i \in A} \left( n_i O_i + n_i^2 P(O)_i \langle i | \hat{V}_2 | i^* \rangle \right) \\ &= \sum_A \sum_{i \in A} \left( n_i O_i + n_i^2 P(O)_i \sum_x^{x \in \hat{2}} \left\langle i \left| \frac{1}{4\pi\epsilon_0} \frac{q_x}{|\vec{R}_x - \vec{r}|} \right| i^* \right\rangle \right). \end{aligned} \quad (72)$$

where

$$O_i = 2 \langle i | \hat{O} | i \rangle \quad (73)$$

and

$$P(O)_i = \frac{4 \langle i | \hat{O} | i^* \rangle}{E_0 - E_i}. \quad (74)$$

The polarization parameters  $O_i$  and  $P(O)_i$  have to be determined by calibration. Formula (72) holds for any one-electron expectation value with a strong polarization dependency. The atomic coordinate dependent charges enter the matrix elements of the bond polarization energy. If the polarity parameters  $d$  are considered to be much smaller than 1 and the overlap elements (such as  $\langle \chi_A | \hat{V}_2 | \chi_B \rangle$ ) are also assumed to be small, the matrix elements of the bond polarization energy have the form

$$\langle i | \hat{V}_2 | i^* \rangle = \frac{\sqrt{1-d^2}}{2} \left[ \langle \chi_A^i | \hat{V}_2 | \chi_A^i \rangle - \langle \chi_B^i | \hat{V}_2 | \chi_B^i \rangle \right]. \quad (75)$$

The determination of bond orbital matrix elements thus means solving a sum of integrals

$$\langle \chi_\lambda | \hat{V}_2 | \chi_\lambda \rangle = \frac{1}{4\pi\epsilon_0} \sum_x^{x \in \hat{2}} \sum_k h_k^2 \int d\vec{r}^3 \phi_k^\lambda(\vec{r}) \frac{q_x}{|\vec{R}_x - \vec{r}|} \phi_k^\lambda(\vec{r}), \quad (76)$$

where the factor  $0.5\sqrt{1-d^2}$  becomes part of the polarization parameter

$$P(O)_i \rightarrow P(O)_i = \frac{2\sqrt{1-d^2} \langle i_A | \hat{O} | i_A^* \rangle}{E_0 - E_i}, \quad (77)$$

and the final form of an one-electron expectation value becomes

$$\langle \Psi_0 | \hat{O} | \Psi_0 \rangle = \sum_A \sum_{i \in A} \left( n_i O_i + n_i^2 P(O)_i \left[ \langle \chi_A^i | \hat{V}_2 | \chi_A^i \rangle - \langle \chi_B^i | \hat{V}_2 | \chi_B^i \rangle \right] \right). \quad (78)$$

Up to this point, it has been assumed that the expectation value of the operator  $\widehat{O}$  is isotropic in nature. One could normally consider anisotropic expectation values, and the BPT approach can theoretically be extended in a tensorial context

$$\begin{aligned} \langle \Psi_0 | \widehat{O}_{\alpha\beta} | \Psi_0 \rangle &= \sum_A \sum_i^{i \in A} \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i \left( n_i O_i^{\alpha'\beta'} + \right. \\ &\quad \left. n_i^2 P(O)_i^{\alpha'\beta'} \left[ \langle \chi_A^i | \widehat{V}_2 | \chi_A^i \rangle - \langle \chi_B^i | \widehat{V}_2 | \chi_B^i \rangle \right] \right), \end{aligned} \quad (79)$$

where  $D_{\alpha\alpha'}$  are the matrix elements of the coordinate transformation from the bond orbital frame to the reference frame.

### 3.3. Conclusion

With equation (79) the general tensorial BPT formalism is given which holds for any tensorial one-electron expectation value with a strong polarization dependence. In practice, it appears useful to introduce a local one-electron operator. The sum over all subsystems  $\widehat{1}$  in the equations (78) and (79) breaks down into

$$\sum_A \sum_i^{i \in A} \Rightarrow \sum_i^{i \in A}. \quad (80)$$

The net atomic charges and the  $^{13}\text{C}$  chemical shifts are calibrated within this approximation. But the BPT could also be used for efficient calculation of other expectation values such as the quadrupole tensor, the indirect dipolar coupling or the dipole moment.

## 4. Efficient Charge Determination

### 4.1. Introduction

Charges are essential in order to calculate any one-electron expectation values within the BPT approach [see equation (70)]. The definition of a local charge operator ( $\hat{O} = \hat{q}_A$ ) would enable the calculation of atomic charges with the BPT method, see equation (78):

$$q_A = \langle \Psi_0 | \hat{q}_A | \Psi_0 \rangle = \sum_i^{i \in A} \left( n_i q_i + n_i^2 P(q)_i \left[ \langle \chi_A^i | \hat{V}_2 | \chi_A^i \rangle - \langle \chi_B^i | \hat{V}_2 | \chi_B^i \rangle \right] \right). \quad (81)$$

By analyzing these charge equations, it is obvious that BPT charges ( $q_A$ ) are estimated from atomic charges ( $q_x$ )

$$q_A = \sum_i^{i \in A} \left( n_i q_i + n_i^2 P(q)_i \sum_x^{x \in \hat{2}} \frac{1}{4\pi\epsilon_0} \left[ \langle \chi_A^i | \frac{q_x}{|\vec{R}_x - \vec{r}|} | \chi_A^i \rangle - \langle \chi_B^i | \frac{q_x}{|\vec{R}_x - \vec{r}|} | \chi_B^i \rangle \right] \right). \quad (82)$$

The computational time for setting up all  $N$  charge equations is proportional to  $N^2$ . Charge calculation means solving this set of linear equations for which the number of floating point operations is proportional to  $N^3$ . Atomic charges are not measurable directly and a meaningful charge operator does not exist since the atomic character of many observables is lost within molecules, clusters and solids due to electron interference between different atoms. Atomic charges can only be indirectly determined by measuring strongly charge dependent observables such as the dipole moment, the electric field gradient or the magnetic shielding. However, the knowledge of atomic charges is of substantial interest, and there exist several definitions of this property, for instance ESC (electrostatic charges),<sup>[102]</sup> MPA (Mulliken population analysis),<sup>[103]</sup> DI (density integration),<sup>[104]</sup> NPA (natural population analysis),<sup>[105]</sup> ENC (electronegativity charges)<sup>[106]</sup> or PACHA (partial atomic charges and hardness analysis).<sup>[107]</sup> These models: NPA, MPA, ESC, PACHA and ENC were applied in this dissertation on a pseudopeptide zinc complex consisting of 64 atoms (H, C, N, O, Zn) as a first test. The structures were optimized with GAUSSIAN98 applying density functional theory with a 6-31G(d,p) basis set. The charge calculations were done with best performance parameters (NPA with a 6-31G basis set, MPA and ESC with a

#### 4. Efficient Charge Determination

3–21G basis set), see below. PACHA charges were calculated by Marc Henry (private communication) and ENC with the COSMOS program. For the zinc complex employed here, the results of these models were correlated and compared with each another. The following correlation matrix could be obtained:

$$\begin{pmatrix} 1 & 0.963 & 0.871 & 0.941 & 0.853 \\ 0.963 & 1 & 0.845 & 0.845 & 0.755 \\ 0.871 & 0.845 & 1 & 0.815 & 0.796 \\ 0.941 & .845 & 0.815 & 1 & 0.934 \\ 0.853 & 0.755 & 0.796 & 0.934 & 1 \end{pmatrix}. \quad (83)$$

The comparison of NPA charges with PACHA charges yield a correlation coefficient of 0.941 for instance. The best correlation is observed between NPA and MPA; the correlation coefficient is 0.963. The data was further analyzed by a multivariate normal distribution with the density function

$$p(\vec{q}) = \frac{1}{2\pi\sqrt{\det[C^{-1}]}} \exp\left[-\frac{1}{2}(\vec{q} - \vec{Q})^T C^{-1}(\vec{q} - \vec{Q})\right], \quad (84)$$

with the covariance matrix elements

$$C_{ij} = \overline{q_i q_j} - \overline{q_i} \overline{q_j}, \quad (85)$$

and hence

$$C = e^2 \begin{pmatrix} 0.228 & 0.183 & 0.182 & 0.078 & 0.102 \\ 0.183 & 0.159 & 0.148 & 0.058 & 0.074 \\ 0.182 & 0.148 & 0.192 & 0.062 & 0.086 \\ 0.078 & 0.058 & 0.062 & 0.03 & 0.04 \\ 0.102 & 0.074 & 0.086 & 0.04 & 0.061 \end{pmatrix}. \quad (86)$$

The vector  $\vec{q} = (q_1, q_2, q_3, q_4, q_5) = (q_{NPA}, q_{MPA}, q_{ESC}, q_{PACHA}, q_{ENC})$  is the vector of charge models and  $\vec{Q}$  is the corresponding vector of means. The covariance matrix  $C$  is symmetric, and a transformation in the principal value system can be performed, the coordinate system is changed by a rigid rotation to remove any correlations between the variables in  $q_i$

$$(\vec{q} - \vec{Q})^T \hat{T}^T \begin{pmatrix} 0.609 & 0 & 0 & 0 & 0 \\ 0 & 0.034 & 0 & 0 & 0 \\ 0 & 0 & 0.023 & 0 & 0 \\ 0 & 0 & 0 & 0.004 & 0 \\ 0 & 0 & 0 & 0 & 0.001 \end{pmatrix} \hat{T} (\vec{q} - \vec{Q}), \quad (87)$$

with

$$\vec{q}' - \vec{Q}' = \hat{T} (\vec{q} - \vec{Q}), \quad (88)$$

and the transformation matrix

$$\hat{T} = \begin{pmatrix} -0.270 & -0.220 & -0.235 & -0.092 & -0.123 \\ 0.172 & 0.183 & -0.368 & 0.034 & -0.026 \\ 0.044 & -0.214 & -0.094 & 0.165 & 0.341 \\ 0.245 & -0.245 & 0.023 & 0.139 & -0.245 \\ 0.197 & -0.119 & -0.009 & -0.375 & 0.079 \end{pmatrix}. \quad (89)$$

The eigenvector corresponding to the largest eigenvalue, i.e. the smallest variance, has the form

$$q'_1 \sim (q_{NPA} + 0.87 q_{ESC} + 0.81 q_{MPA} + 0.46 q_{ENC} + 0.34 q_{PACHA}). \quad (90)$$

Due to the large eigenvalue of  $0.609 \text{ 1/e}^2$ , which corresponds to the smallest variance of  $1.64 \text{ e}^2$  or a standard deviation of  $1.28 \text{ e}$ , this coordinate dominates all other coordinates  $q'_i$ . It is obvious that NPA has a major contribution to the eigenvector. It follows, that the probability distribution can be approximated by

$$p(\vec{q}') \approx \frac{1}{2\pi\sqrt{1.64}} \exp\left[-\frac{1}{2} \frac{(q'_1 - Q'_1)^2}{1.64 \text{ e}^2}\right]. \quad (91)$$

NPA might be the charge calculation model of choice judged by the pure statistical investigation above. In this thesis the BPT was parameterized using the methods NPA, ESC and MPA. The models PACHA and ENC were not employed for three reasons: they showed a low contribution to the eigenvector  $q'_1$ , they cannot be calculated by *ab initio* techniques and they are less coordinate dependent than the BPT formalism.

A set of 175 model structures including 12 zinc compounds were optimized with density functional theory (DFT/B3LYP) using a 6-31G(d,p) basis set. The optimized atomic coordinates can be obtained from the author upon request. A subset of 163 molecules was then chosen consisting of H, C, N, O, F, Si, P, S and Cl for calibration using 11 different basis sets. In a second step, the best basis sets that are also applicable to Zn calculations were then applied for the calibration of the set of 175 structures.

## 4.2. Mulliken Population Analysis (MPA)

The simplest and most common approximate *ab initio* treatment is SCF-MO-LCAO-CGTO (Self Consistent Field approach of Molecular Orbitals that are approximated as Linear Combinations of Atomic Orbitals using Contracted Gauss Type Orbitals). The SCF closed shell molecular wave function of  $N$  electrons is approximated from the CI or CC model as

$$|\Psi\rangle = \sum_l C_l |\Psi_l\rangle \approx |\Psi_1\rangle, \quad (92)$$

with an antisymmetric sum of products of  $N$  molecular spin orbitals  $|\Phi_i(j)\rangle$

$$|\Psi\rangle \approx \frac{1}{\sqrt{N!}} \sum_P (-1)^P \hat{P} [|\Phi_1(1)\rangle |\Phi_1(2)\rangle \cdots |\Phi_1(N)\rangle], \quad (93)$$

with  $|\Phi_i(2k-1)\rangle = |\alpha_i(k)\rangle |\chi_i(k)\rangle$ ,  $|\Phi_i(2k)\rangle = |\beta_i(k)\rangle |\chi_i(k)\rangle$  and  $k = 1 \dots N/2$ .  $\hat{P}$  is the permutation operator. The molecular orbitals  $|\chi_i(k)\rangle$  are approximated as LCAOs (linear combination of atomic orbitals) of STOs, see equation (63)

$$\langle \vec{r} | \chi_i(k) \rangle = \chi_i(\vec{r}_k) = \sum_{\mu} c_{i\mu} \phi_{\mu}(\vec{r}_k). \quad (94)$$

The success of atomic quantum mechanical calculations is essentially based on the use of Gaussian functions, since two-center integrals can be transformed into one-center integrals. The STOs are approximated as a linear combination of PGTOs (primitive Gauss type orbitals), and the approximated STOs are called CGTOs (contracted Gauss type orbitals)

$$\phi(\vec{r}) = \sum_{\gamma} d_{\gamma} g_{\gamma}(\vec{r}), \quad (95)$$

with the PGTOs

$$g = N x^u y^v z^w \exp(-\alpha r^2), \quad u + v + w = n - 1. \quad (96)$$

The contraction coefficients  $d_{\gamma}$  and exponentials  $\alpha_{\gamma}$  have typically been optimized in order to reflect the behavior of STOs. The STO-3G nomenclature means that a STO is approximated by a CGTO consisting of three PGTOs.

In order to understand a population analysis, the local density operator at atom position  $A$  is



defined as

$$\hat{\rho}_A = \left| \vec{r}_A \right\rangle \left\langle \vec{r}_A \right| , \quad (97)$$

with  $\sum_A \hat{\rho}_A = 1$ . The local density is given by

$$\begin{aligned} \rho_A &= \left\langle \Psi \left| \vec{r}_A \right\rangle \left\langle \vec{r}_A \right| \Psi \right\rangle \\ &= \left( \sum_i^{N/2} (\langle \alpha_i(k) | + \langle \beta_i(k) |) \langle \chi_i(k) | \vec{r}_A \rangle \right) \left( \sum_j^{N/2} \langle \vec{r}_A | \chi_j(l) \rangle (|\alpha_j(l)\rangle + |\beta_j(l)\rangle) \right) \\ &= 2 \sum_i^{N/2} \langle \chi_i(k) | \vec{r}_A \rangle \langle \vec{r}_A | \chi_i(l) \rangle \\ &= 2 \sum_i^{N/2} \sum_{\mu, \nu} c_{i\mu}^* c_{i\nu} \phi_\mu^* \left( \vec{r}_{A(k)} \right) \phi_\nu \left( \vec{r}_{A(l)} \right) . \end{aligned} \quad (98)$$

Defining the bond-order matrix  $\hat{D}$  with its matrix elements as  $D_{\mu\nu} = 2 \sum_i^{N/2} c_{i\mu}^* c_{i\nu}$ , it follows that

$$\rho_A = \sum_{\mu, \nu} D_{\mu\nu} \phi_\mu^* \left( \vec{r}_{A(k)} \right) \phi_\nu \left( \vec{r}_{A(l)} \right) . \quad (99)$$

Thus, the population at atom  $A$  can be calculated by

$$\begin{aligned} N_A &= \int \rho_A dV \\ &= \sum_{\mu, \nu} D_{\mu\nu} S_{\mu\nu} \\ &= \sum_\mu^A \left( D_{\mu\mu} + \sum_{\nu, \nu \neq \mu}^A D_{\mu\nu} S_{\mu\nu} + \sum_{B, B \neq A} \sum_{\nu, \nu \neq \mu}^B D_{\mu\nu} S_{\mu\nu} \right) , \end{aligned} \quad (100)$$

and the atomic charge can be defined as

$$q = e(Z - N_A) . \quad (101)$$

There are two problems connected with the MPA.<sup>[103]</sup> First, the overlap contribution, the so-called interference density,  $\sum_B \sum_{\nu, \nu \neq \mu}^B D_{\mu\nu} S_{\mu\nu}$  of atom  $A$  to all other atoms  $B$  is equally distributed to  $A$  as well as  $B$ , which is a wrong assumption for strongly polarized heteroatomic bonds. Second, the MO might have delocalized contributions which describe the electron density of

#### 4. Efficient Charge Determination

another atom than  $A$ ,  $\sum_{\mu}^A D_{\mu\mu}$ . This effect could be strongly basis set dependent.

MPA	$R$	$SD$ [e]	$SD/\Delta Q$ [%]
	STO-2G	0.9912	0.0247
STO-3G	0.9911	0.0247	13.22
STO-6G	0.9944	0.0244	10.50
3-21G	0.9941	0.0448	10.77
6-21G	0.9941	0.0439	10.80
6-31G(d,p)	0.9917	0.0391	12.73
6-31+G(d,p)	0.9583	0.1036	27.38
6-31++G(d,p)	0.9325	0.1325	33.69
6-311G(d,p)	0.9850	0.0476	16.98
6-311+G(d,p)	0.8466	0.1322	45.06
6-311++G(d,p)	0.8182	0.1463	47.05

Table 1

MPA charge parametrization. In the first column the basis sets are given, in the second the correlation coefficients are listed. The third column shows the standard deviation and in the last column the ratios of the standard deviation and the absolute charge distribution deviation  $\Delta Q$  are given.

The extreme dependence of the Mulliken charges on the basis set is illustrated in the following example: the second carbon atom of  $\text{CH}_2\text{CFCl}$  has a positive charge (+0.05e) calculated with a 3-21G basis set. It is negatively charged (-0.11e) when one employs a 6-311++G(d,p) basis. A sum of one-electron operators for MPAs can be defined

$$\hat{O} = \hat{q}_A^{MPA} = e \left( Z_A - \sum_i^A |\phi_i, \sigma(i)\rangle \langle \phi_i, \sigma(i)| \right). \quad (102)$$

This is an important fact because BPT assumes one-electron operators, see equation (78).

A first parameterization applying STO-3G Mulliken charges were done by Koch et al.<sup>[108,109]</sup>

The results of the calibrations of this work are listed in Table 1. The best correlation (correlation coefficient  $R = 0.9944$ ) is achieved with a STO-6G basis set. This basis set is incomplete (at least in GAUSSIAN98) and not applicable to Zn. Therefore, the basis set with the second best correlation (3-21G) was applied for necessary charge calculations to parametrize the BPT method on 175 molecules including 12 zinc structures. The correlation coefficient is 0.9933, the standard deviation is 0.05 e. The BPT parameters  $q_i$  and  $P(q)_i$ , see equation (??), are given in the Appendix.

### 4.3. Electrostatic Charges (ESC)

A totally different approach is used to compute partial charges via the molecular electrostatic potential. The molecular electrostatic potential is evaluated at points in space around the molecule, and the data are fitted to a classical atomic point charge model.<sup>[102]</sup> The fitting parameters are the charges. The number of layers and the density of points per unit area are input parameters. In this work 10 layers and 10 points per unit area were used since these parameters gave good results within a reasonable computational time.

ESC	$R$	$SD [e]$	$SD/\Delta Q [\%]$
	STO-2G	0.9597	0.0824
STO-3G	0.9595	0.0825	27.04
STO-6G	0.9659	0.0832	25.03
3-21G	0.9768	0.0818	20.93
6-21G	0.9758	0.0820	21.34
6-31G(d,p)	0.9668	0.0813	24.71
6-31+G(d,p)	0.9414	0.1155	31.75
6-31++G(d,p)	0.9426	0.1137	31.46
6-311G(d,p)	0.9470	0.1081	30.43
6-311+G(d,p)	0.9427	0.1172	31.44
6-311++G(d,p)	0.9427	0.1171	31.47

Table 2

ESC charge parametrization. In the first column the basis sets are given, in the second the correlation coefficients are listed. The third column shows the standard deviation and in the last column the ratios of the standard deviation and the absolute charge distribution deviation  $\Delta Q$  are given.

There is no sum of one-electron operators for ESCs defined which makes its theoretical application within the BPT approach uncertain. Since BPT parameters cannot be correctly interpreted. Nevertheless, the parametrization was carried out in respect to different basis sets. The results are given in Table 2. It is evident, that the correlations are not sufficient enough. The best parametrization is obtained with a 3-21G basis set,  $R = 0.9768$ . Using the same conditions, the complete parametrization of the 175 structures was performed. The correlation turned out to be 0.9721 with a standard deviation of 0.08  $e$ . The obtained BPT parameters are given in the Appendix.

#### 4.4. Natural Population Analysis (NPA)

The natural population analysis method was formulated by Reed et al.<sup>[105]</sup> It is based on the fact that the nonorthogonal atomic orbitals  $\{\phi_k\}$  can be transformed to orthonormal NAOs (natural atomic orbitals)  $\{\psi_\mu\}$

$$\psi_\mu = \sum_k T_{k\mu} \phi_k . \quad (103)$$

Due to this property of the NAOs, the overlap matrix elements vanishes for different orbitals

$$S_{\mu\nu} = \int \psi_\mu^* \left( \vec{r}_{A(k)} \right) \psi_\nu \left( \vec{r}_{A(l)} \right) dV = \delta_{\mu\nu} , \quad (104)$$

and the population at atom  $A$  can be calculated from equation (100):

$$N_A = \sum_\mu^A D_{\mu\mu} . \quad (105)$$

NPA	NPA		
	$R$	$SD [e]$	$SD/\Delta Q [\%]$
STO-2G	0.9955	0.0449	9.41
STO-3G	0.9878	0.0308	15.35
STO-6G	0.9928	0.0290	11.91
3-21G	0.9960	0.0437	8.89
6-21G	0.9960	0.0435	8.91
6-31G(d,p)	0.9967	0.0437	8.12
6-31+G(d,p)	0.9893	0.0774	14.40
6-31++G(d,p)	0.9897	0.0748	14.14
6-311G(d,p)	0.9896	0.0710	14.24
6-311+G(d,p)	0.9795	0.0959	19.69
6-311++G(d,p)	0.9796	0.0958	19.68

Table 3

NPA charge parametrization. In the first column the basis sets are given, in the second the correlation coefficients are listed. The third column shows the standard deviation and in the last column the ratios of the standard deviation and the absolute charge distribution deviation  $\Delta Q$  are given.

The overlap matrix elements and interference terms vanish. These are big advantages in comparison to MPA. For less localized basis sets, an erroneous description could possibly appear because the wave functions might have delocalized contributions which describe the electron

density of a neighboring atom. Compact basis sets should therefore be employed. A sum of one-electron operators of NPAs can be defined from the population  $N_A$  using the natural spin atomic orbitals by

$$\hat{q}_A^{NPA} = e \left( Z_A - \sum_i^A |\psi_i, \sigma(i)\rangle \langle \psi_i, \sigma(i)| \right). \quad (106)$$

The parametrization results (of 163 structures) can be seen in Table 3. The 6-31G(d,p) basis gave the best correlation ( $R = 0.9967$ ). For the parametrization including the 12 Zn-structures the same basis set was applied which yielded a correlation coefficient of 0.9961 with a standard deviation of  $0.05 e$ . The BPT parameters  $q_i$  and  $P(q)_i$  for effective charge calculation are given in the Appendix.

## 4.5. Conclusion

The BPT method was parameterized with different atomic charge models: electrostatic charges, Mulliken population analysis (MPA) and natural population analysis (NPA). The basis set dependence of the charge were investigated using 11 different basis sets. The following basis sets

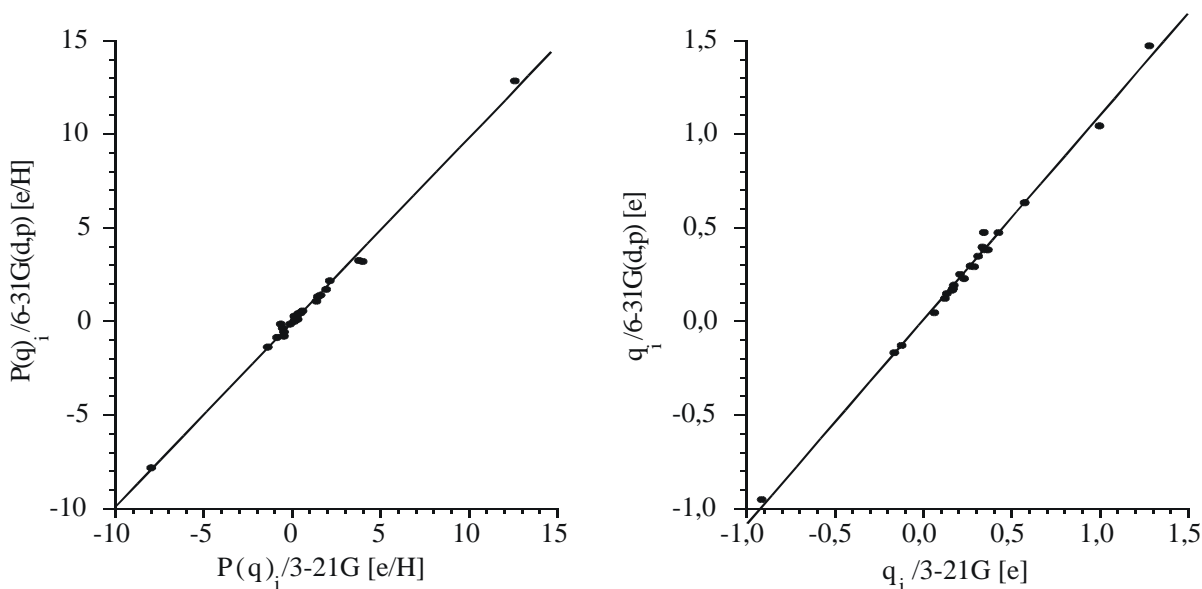


Figure 2

Correlation of the BPT/NPA parameters of different basis sets.

left: correlation of  $P(q)_i/3-21G$  and  $P(q)_i/6-31G(d,p)$  with  $R = 0.9969$  and  $SD = 0.26 e/H$

right: correlation of  $q_i/3-21G$  and  $q_i/6-31G(d,p)$  with  $R = 0.9973$  and  $SD = 0.03 e$

#### 4. Efficient Charge Determination

were chosen for parameterization of 175 structures consisting of H, C, N, O, F, Si, P, S, Cl and Zn atoms based on the test calculations of a set of 163 structures consisting of H, C, N, O, F, Si, P, S and Cl: 3-21G basis set for parameterization of the BPT charge calculation based on charges from Mulliken population analysis, 3-21G basis set for parameterization of the BPT charge calculation based on charges derived from electrostatic potentials, and 6-31G(d,p) basis set for parameterization of the BPT charge calculation based on charges from natural population analysis. In total,  $175+3\times 11\times 163+3\times 175=6079$  quantum chemical calculations and  $3\times 11+3=36$  parametrizations were performed. All BPT parameters are given in the Appendix. NPA yields the highest correlation coefficients as compared to all other charge models tested. A correlation coefficient of 0.9961 was obtained between BPT/NPA and DFT/NPA, and the standard deviation is 0.05  $e$ . The correlation of parameters of different compact basis sets, 3-21G and 6-31G(d,p) is shown in Figure 2. The correlation of non polarized parameters  $q_i$

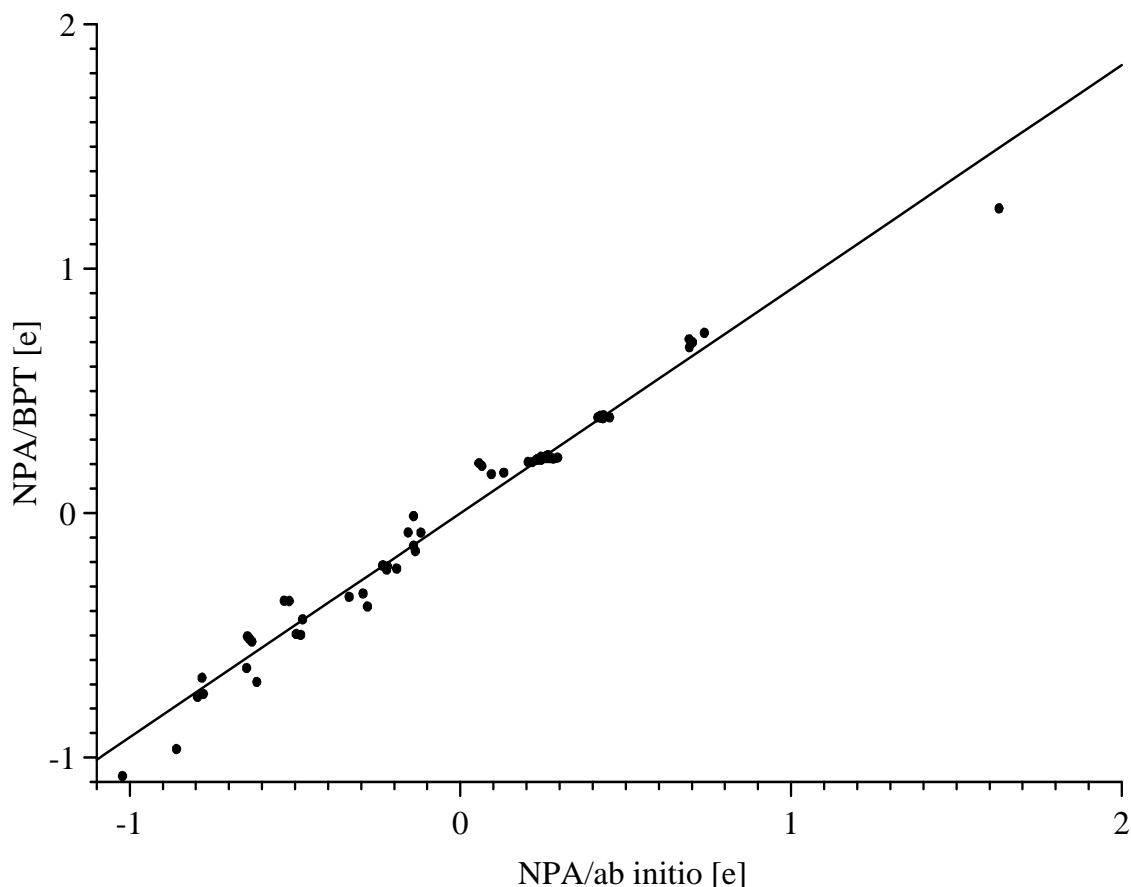


Figure 3

Testcorrelation of BPT/NPA and DFT/NPA on a pseudopeptide zinc complex with 64 atoms consisting of H, C, N, O and Zn ( $R = 0.9882$ ,  $SD = 0.07 e$ ).

#### 4. Efficient Charge Determination

yields a coefficient of 0.9973 ( $SD = 0.03 e$ ), while the polarization parameters  $P(q)_i$  show a  $R$ -value of 0.9969 and a standard deviation of  $0.26 e/H$ . These results bear the first test of the parameter stability of the BPT method as regarding different basis sets (see Figure 2). Thus far, all correlations found between DFT and BPT atomic charges are based on the pool of 175 structures used for parametrization in this work. The BPT/NPA method was therefore tested on DFT/NPA results of a pseudopeptide zinc complex consisting of 64 atoms (H, C, N, O, Zn), that was not included in the set of 175 parameterization structures, see Figure 3. The optimized coordinates are available from the author upon request. The correlation turned out to be 0.9882 with a standard deviation of  $0.07 e$ .

## 5. $^{13}\text{C}$ Chemical Shift Molecular Dynamics Crystal Simulation

### 5.1. Introduction

In 1957, P. Lauterbur and C. Holm independently recorded the first  $^{13}\text{C}$  NMR spectra, despite the low natural isotopic abundance of  $^{13}\text{C}$  (1.1 %). In 1972, Pines, Gibby and Waugh<sup>[110,111]</sup> introduced the concept of  $^{13}\text{C}$  polarization enhancement by a factor of  $\gamma_{1H}/\gamma_{13C} \approx 4$ , called the Bloembergen-Sorokin enhancement,<sup>[112]</sup> which was obtained by cross polarization by applying the Hartmann-Hahn energy level match  $\omega_{1H} = \omega_{13C}$ .<sup>[113]</sup> Since then,  $^{13}\text{C}$  NMR spectroscopy has gained more and more importance and has become a standard method for structural investigations of compounds containing carbon. The  $^{13}\text{C}$  magic angle spinning (MAS) spectrum of *Bombyx mori* silk is shown in Figure 4 and represents the high resolution isotropic  $^{13}\text{C}$  resonances. In addition, quantum mechanical computational procedures and computer hardware

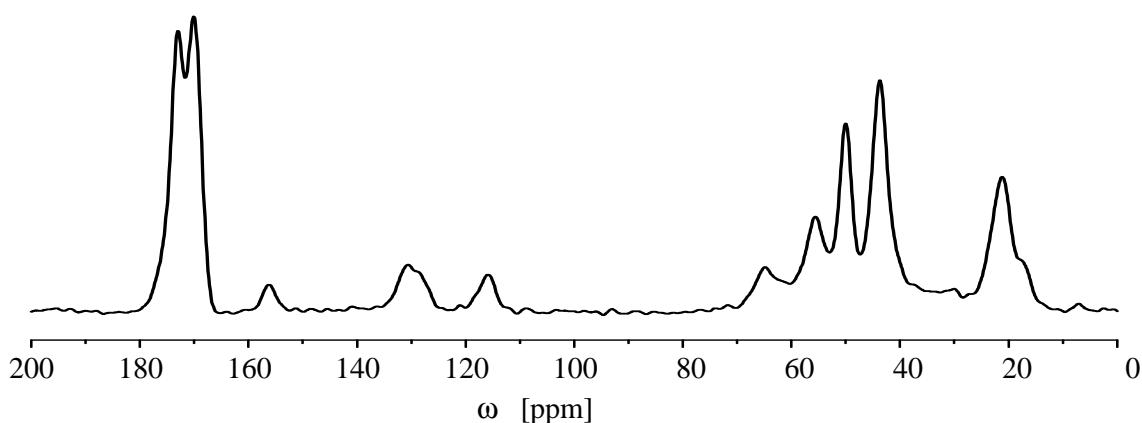


Figure 4

*Bombyx mori* silk 1D- $^{13}\text{C}$ -CP-MAS spectrum. Experimental parameters: CP contact time of 5 ms, experimental time of 1 s and 1024 scans.

have made it possible to perform  $^{13}\text{C}$  calculations for molecules or molecular fragments large enough to reflect the essential features of the local environment.<sup>[114]</sup> A minimum 3-21G basis set<sup>[115]</sup> is sufficient for systems of second row atoms (H-Ne). For more accurate  $^{13}\text{C}$  chemical shifts, the carbon sites can be treated with a 6-311+G\* basis set.<sup>[116]</sup> 59 contracted Gaussian functions are needed for N-formyl-L-alanine residues (10 atoms), if only the  $\text{C}_\alpha$  and  $\text{C}_\beta$  sites are of interest. The CPU time is proportional to  $59^4$  at the HF level, and it takes 24.2 minutes for a single chemical shift calculation on a 100 Mflop/s machine.<sup>[117]</sup> Current standard GHz



PCs run about 0.5 Gflop/s peak speed. Unfortunately, the computational effort for molecular dynamics simulations or geometry optimizations is still too large for a routine calculation. Only large parallel computer clusters with a peak speed of several Tflop/s are capable of performing such *ab initio* calculations within a sufficient time scale. Therefore, we are in need of fast semi-empirical quantum chemical methods. In this work, the bond polarization theory<sup>[6]</sup> is used to provide a vast method for the computation of chemical shifts and their related derivatives.

## 5.2. BPT <sup>13</sup>C NMR Chemical Shifts

The chemical shift can be considered as a sum of one-electron operators due to the definition<sup>[38]</sup>

$$\widehat{\delta} = \widehat{\sigma}_{ref} - \widehat{\sigma}_N . \quad (107)$$

The inner shell contributions of the shielding operator  $\widehat{\sigma}_N$  and the reference  $\widehat{\sigma}_{ref}$  are identical and cancel each other. The shift operator acts on all electrons in the bond orbitals and lone electron pairs. Hence, the sum of equation (68) runs over all bonds and lone pairs of the molecule being investigated. The expectation value of the shift is roughly proportional to the inverse distance of the electrons from the atom position  $A$ . In a first approximation, when only bonds that are directly connected to the nucleus of  $A$  are taken into account, the sum breaks down. In addition, carbon sites don't have lone electron pairs. It follows from equation (79) and (80) that

$$\begin{aligned} \langle \Psi_0 | \widehat{\delta}_{\alpha\beta} | \Psi_0 \rangle &= \sum_i^{i \in A} \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i \left( n_i \delta_i^{\alpha'\beta'} + \right. \\ &\quad \left. n_i^2 P(\delta)_i^{\alpha'\beta'} \left[ \langle \chi_A^i | \widehat{V}_2 | \chi_A^i \rangle - \langle \chi_B^i | \widehat{V}_2 | \chi_B^i \rangle \right] \right) . \end{aligned} \quad (108)$$

The bond increments  $\delta_i^{\alpha'\beta'}$  and polarization parameters  $P(\delta)_i^{\alpha'\beta'}$  are obtained by a calibration procedure.<sup>[118,119]</sup> A collection of crystal structures and single crystal chemical shift tensor measurements<sup>[120,121]</sup> is used to establish a set of linear equations in the form of equation (108). In addition, calculated *ab initio* data were used for calibration. The correlation coefficient is  $R = 0.994$  and the standard deviation  $SD = 7.2$  ppm (unpublished data by Wolfram Prieß). After the parametrization, the parameters  $\delta_i^{\alpha'\beta'}$  and  $P(\delta)_i^{\alpha'\beta'}$  are known. Thus, only the matrix elements of the bond polarization energy,  $\langle \chi_A^i | \widehat{V}_2 | \chi_A^i \rangle$  as well as  $\langle \chi_B^i | \widehat{V}_2 | \chi_B^i \rangle$  and the occupation numbers  $n_i$  have to be calculated. They can all be analytically expressed, thus making the BPT formalism very efficient computationally. In equation (108), there are two sums: the

first runs over all bond contributions of the atom in consideration, and the second runs over all polarizing charges of  $\widehat{V}_2$ . If the charges are known, the computational cost for a chemical shift calculation is proportional to  $N$ .

### 5.3. Theoretical Background

For conventional NMR spectra, the average sample volume is about  $10^2 \text{ mm}^3$ , and the measurements are on a time scale of  $\mu\text{s}$ . Therefore, statistical methods are crucial for NMR parameter simulations. In this thesis, the  $^{13}\text{C}$  chemical shift tensor averaging over time was achieved by molecular dynamics simulation by adapting a modified leap-frog algorithm to generate trajectories within the constant NVT ensemble (N...number of particles, V...volume, T...temperature).<sup>[122]</sup> The starting point of conventional molecular dynamics is the potential  $U(\vec{r}(t))$  which is minimal in an equilibrium state. The driving force to obtain a minimum state is the negative gradient of this potential. Any atom with mass  $m$  at the coordinate position  $\vec{r}(t)$  is subjected to the acceleration

$$\vec{a}(t) = -\frac{1}{m} \vec{\nabla} U(\vec{r}(t)) . \quad (109)$$

The particle velocity  $\vec{v}$  at the time  $t$  is initially calculated by

$$\vec{v}(t) = \vec{v}\left(t - \frac{\Delta t}{2}\right) + \vec{a}(t) \frac{\Delta t}{2} . \quad (110)$$

The overall translational velocity  $\vec{\sigma}$  and rotational moving  $\vec{\omega} \times (\vec{r} - \vec{R})$  around the center of gravity  $\vec{R}$  are then removed:

$$\vec{v}'(t) = \vec{v}(t) - \vec{\sigma} - \vec{\omega} \times (\vec{r} - \vec{R}) . \quad (111)$$

The temperature is then calculated by

$$T(t) = \frac{1}{3f k_B} \sum_{i=1}^N m_i \left(\vec{v}'_i(t)\right)^2 , \quad (112)$$

with the Boltzmann constant  $k_B$  and the number of degrees of freedom  $f = 3N - 7$ , where seven degrees of freedom are fixed: total momentum, total angular momentum and temperature. In

the next step, the tempering within the NVT ensemble is achieved by

$$\begin{aligned} \vec{v}'' \left( t + \frac{\Delta t}{2} \right) &= \left( 2\sqrt{1 + \frac{\Delta t}{\tau} \left( \frac{T_0}{T} - 1 \right)} - 1 \right) \vec{v}' \left( t - \frac{\Delta t}{2} \right) + \\ &\quad \sqrt{1 + \frac{\Delta t}{\tau} \left( \frac{T_0}{T} - 1 \right)} \vec{a} (t) \Delta t, \end{aligned} \quad (113)$$

where  $T_0$  is the destination temperature and  $\tau$  is the coupling time, which is in the range of ps.<sup>[122]</sup> The last cycle step of the leap-frog algorithm is the calculation of the new coordinates

$$\vec{r} (t + \Delta t) = \vec{r} (t) + \vec{v}'' \left( t + \frac{\Delta t}{2} \right) \Delta t, \quad (114)$$

which are then used for recalculating the potential  $U \left( \vec{r} (t + \Delta t) \right)$  and its gradient for the next step of the molecular dynamics simulation.

The potential is translational periodic within a lattice. Using the unit vectors of the unit cell  $\vec{a}$ ,  $\vec{b}$ ,  $\vec{c}$  and integers  $i$ ,  $j$ ,  $k$  it is

$$U \left( \vec{r} \right) = U \left( \vec{r} + i \vec{a} + j \vec{b} + k \vec{c} \right). \quad (115)$$

In this work  $\{i, j, k\} = 0, \pm 1$  was used. This condition does not influence the bonded contributions, but it cuts off the van der Waals interactions, the long range electrostatic potentials and the BPT calculation of the charges as well as chemical shift tensors. Nevertheless, a calculation of  $3 \times 3 \times 3 = 27$  unit cells can be considered as a good first approximation for crystal simulations because neighboring effects of unit cells in all 3 dimensions are considered.

#### 5.4. Simulation of the $^{13}\text{C}$ Chemical Shift Spectra of Silk II

Spider and silkworm silks are fibrous proteins with numerous actual and potential applications. These polymer materials are rather heterogeneous and can not be sufficiently characterized using diffraction methods. Silks have therefore been extensively studied using NMR methods (for review see for instance Zhao et al.<sup>[123]</sup>). In this work, degummed silk from *Bombyx mori* was studied which consists mainly of proteins forming  $\beta$ -sheets (Silk II<sup>[123]</sup>). The primary structure of this polymer is build up mainly from (–Gly–Ala–Gly–Ala–Gly–Ser–)–hexamers or –octamers with a lower content of tyrosine. Diffraction studies led to the conclusion (Takahashi et al.<sup>[124]</sup>) that the more crystalline silk Cp-fraction consists of two antipolar-antiparallel  $\beta$ -sheet structures in different orientations. The quest for the proper Silk II structure still stands, because

of the relative low resolution of fiber X-ray structural investigations. Alternative structure versions have been published by Fossey et al.<sup>[125]</sup> and Asakura et al.<sup>[126]</sup> The 1D- $^{13}\text{C}$ -CP-MAS NMR spectrum (shown in Figure 4) resembles very closely a spectrum that was published by Asakura and Kameda.<sup>[127]</sup> The assignment of the resonances was performed in accordance to this paper and to a paper of Asakura and Zhao.<sup>[123]</sup> A crystal molecular dynamics simulation was carried out with 20000 time steps  $\Delta t$  of 0.5 fs at a temperature of 393 K using the Silk II model of Takahashi et al. Coordinates were saved every 100th time step. 200 structures were thus obtained. These snapshots were geometry optimized and the chemical shift tensors were calculated and averaged. The resulting NMR Iso-Aniso spectrum is shown in Figure 5. The

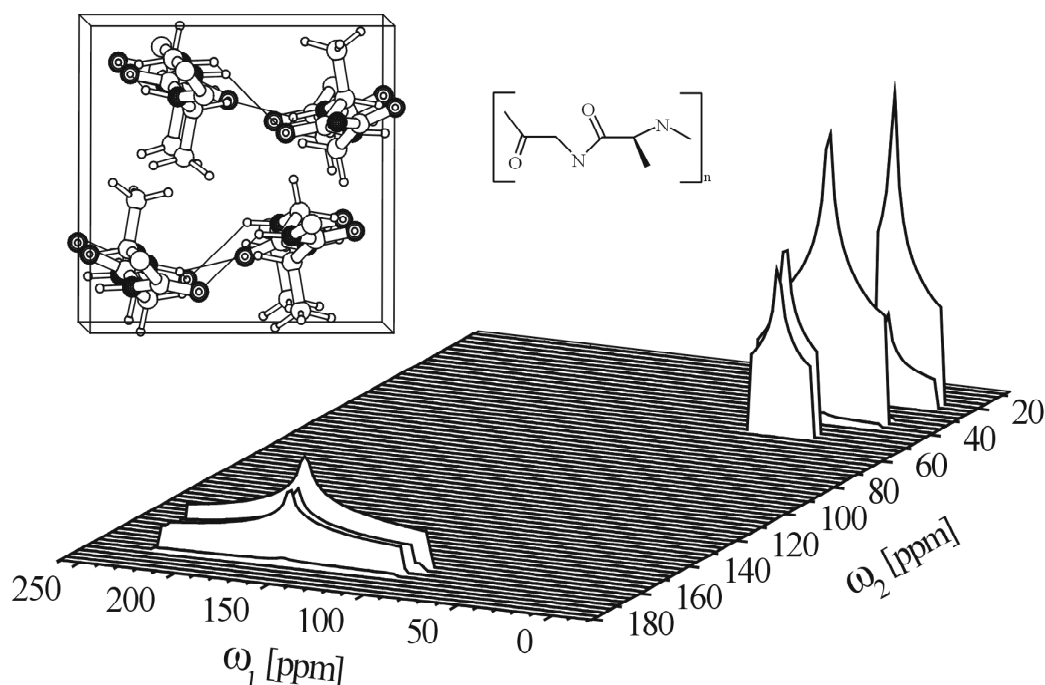


Figure 5

Simulated Iso-Aniso 2D spectrum of the Silk II model of Takahashi et al.,<sup>[124]</sup> which consists of two antiparallel-antiparallel  $\beta$ -sheet structures with different orientations. The view of the unit cell shows that the chains are oriented parallel to the  $\bar{c}$ -direction (perpendicular to the plane of the figure).

intensities  $I(\omega)$  of the spectra of the tensor powder patterns are calculated at the indirect  $\omega_1$

dimension according to<sup>[128]</sup>

$$I(\omega) = \begin{cases} \frac{\pi}{\sqrt{(\omega_{11}-\omega_{22})(\omega-\omega_{33})}} \int_0^1 \frac{dg}{(1-g^2)(1-m_1g^2)}, & \omega \in (\omega_{22}, \omega_{33}) \\ \frac{\pi}{\sqrt{(\omega_{11}-\omega)(\omega_{22}-\omega_{33})}} \int_0^1 \frac{dg}{(1-g^2)(1-m_2g^2)}, & \omega \in (\omega_{11}, \omega_{22}), \end{cases} \quad (116)$$

with

$$\begin{aligned} m_1 &= \frac{(\omega_{11}-\omega)(\omega_{22}-\omega_{33})}{(\omega-\omega_{33})(\omega_{11}-\omega_{33})} \\ m_2 &= \frac{(\omega_{11}-\omega_{22})(\omega-\omega_{33})}{(\omega_{11}-\omega)(\omega_{22}-\omega_{33})}. \end{aligned} \quad (117)$$

The chemical shift principal tensor values were reordered according to

$$\delta_{33} \geq \delta_{22} \geq \delta_{11}, \quad (118)$$

with  $\omega_{ii} = \gamma \hbar B_0 \delta_{ii}$ . A C++ routine in Numerical Recipes<sup>[129]</sup> was used for the calculation of the Legendre elliptic integrals of the 1st kind in equation (116).

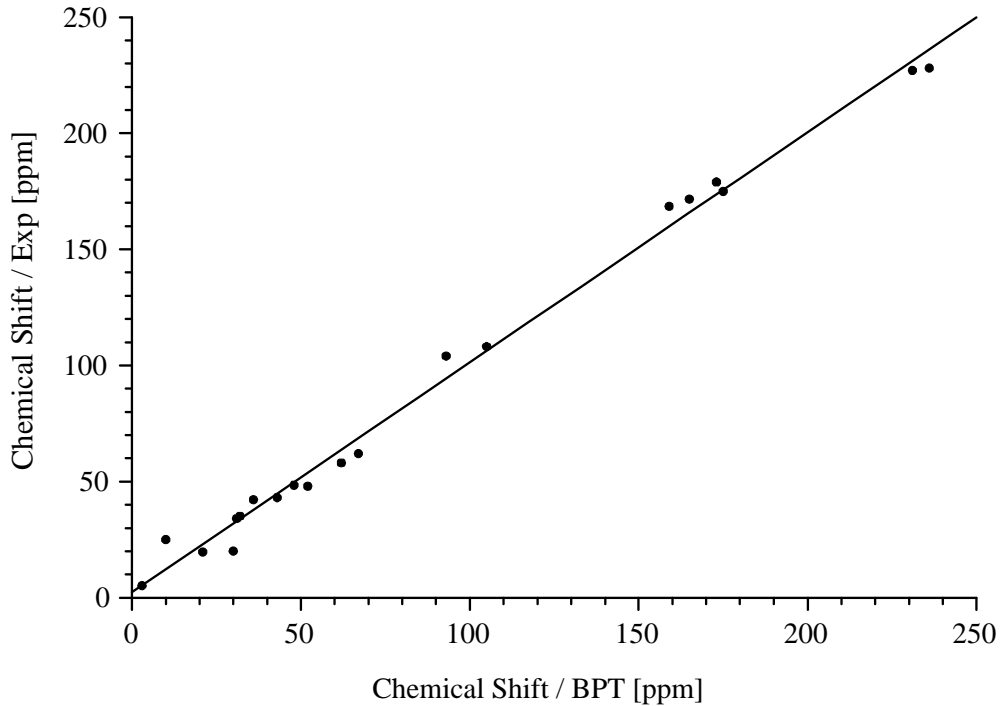


Figure 6  
Correlation between the averaged BPT chemical shift tensor values and experimental results of Witter et al.<sup>[130]</sup> ( $R = 0.996$ ,  $SD = 6.8$  ppm).

## 5.5. Results and Conclusion

The average chemical shift tensor values obtained by crystal molecular dynamics simulation of the Silk II model of Takahashi et al.<sup>[124]</sup> are listed in Table 4. A comparison to results from the literature is also given. There is a very good agreement between the averaged BPT <sup>13</sup>C chemical shift tensors and the experimental data obtained by Witter et al.<sup>[130]</sup> (see Figure 6). The correlation of 0.996 and the standard deviation of 6.8 ppm are remarkable as compared to the correlation of the original parametrization ( $R = 0.994$  and  $SD = 7.2$  ppm).

Carbon		Calculation		Experiment		
		This work	Zhou <sup>[131]</sup>	Witter <sup>[130]</sup>	Zhao <sup>[123]</sup>	Demura <sup>[132]</sup>
Ala C $\beta$	$\delta_{i,so}$	21(4)		19.6	20.2	
	$\delta_{11}$	31(8)		34(2)		
	$\delta_{22}$	30(8)		20(1)		
	$\delta_{33}$	3(8)		5(2)		
Gly C $\alpha$	$\delta_{i,so}$	36(4)		42.1	43.1	
	$\delta_{11}$	62(8)		58(2)		
	$\delta_{22}$	43(8)		43(1)		
	$\delta_{33}$	10(8)		25(2)		
Ala C $\alpha$	$\delta_{i,so}$	48(4)		48.4	49.4	
	$\delta_{11}$	67(8)		62(2)		
	$\delta_{22}$	52(8)		48(1)		
	$\delta_{33}$	32(8)		35(2)		
Gly CO	$\delta_{i,so}$	159(4)	167.4	168.5	169.5	174.3
	$\delta_{11}$	231(8)	246.8	227(2)		245(5)
	$\delta_{22}$	175(8)	164.3	175(1)		179(5)
	$\delta_{33}$	93(8)	91.0	104(2)		99(5)
Ala CO	$\delta_{i,so}$	165(4)	172.5	171.6	172.3	174.7
	$\delta_{11}$	236(8)	249.2	228(2)		242(5)
	$\delta_{22}$	173(8)	177.6	179(1)		186(5)
	$\delta_{33}$	105(8)	90.8	108(2)		96(5)

Table 4

<sup>13</sup>C chemical shift anisotropy data (ppm) of Silk II and results from the literature: Comparison of BPT calculated data, DFT chemical shift calculations of Zhou et al. based on a model of Fossey et al. (H-bonds are considered) and experimental data of Bombyx mori silk by Witter et al., <sup>13</sup>C CP/MAS NMR chemical shifts of the amino acid residues of B. mori fibroin combined of Silk II by Zhao et al. and experimental data of Demura et al. (solid state <sup>13</sup>C enriched powder patterns fitted on experimental spectra without considering <sup>15</sup>N-<sup>13</sup>C coupling).

## 6. BPT Pseudo Forces

### 6.1. Theory

The minimization of the Gibbs free enthalpy is the driving force leading to molecular equilibrium

$$G(p, T, N) = E_{kin} + U_{pot} - TS + pV . \quad (119)$$

It is a function of temperature and pressure and represents the energy of a molecular system at constant temperature  $T$  and pressure  $p$  when one excludes volume work  $-pV$  and heat exchange  $+TS$ . It includes entropic driving forces such as hydrophobic and hydrophilic effects. The gradient of the Gibbs potential can be calculated from a molecular force field and the gradient of the Hessian determinant by employing a harmonic approximation.<sup>[133,134]</sup> The potential energy ( $E_{pot}$ ) describes the interactions between residues, atoms or atomic groups, and molecules. The COSMOS force field<sup>[90]</sup> is used in this work. Force contributions to the free energy beyond the potential energy are entropic in nature. They involve the second and third derivative of the potential energy in the harmonic approximation and scale linearly with the temperature. They cannot be calculated efficiently and are therefore not used in practice. But, the molecular forces do not vanish in the equilibrium state at  $T > 0$ .

In the COSMOS force field, additional contributions are introduced, targeting calculated expectation values to experimental values, in our case  $^{13}\text{C}$  chemical shifts. In order to obtain energetic corrections, the polarization energy contribution of the bonds around the nucleus  $A$  has to be calculated. It follows from equations (68 and 71) and using the Fock operator  $\hat{O} = \hat{F}$  that

$$\begin{aligned} \langle \Psi_0 | \hat{F} | \Psi_0 \rangle &= \sum_A \sum_i^{i \in A} \left( 2n_i \langle i | \hat{F} | i \rangle + (2n_i)^2 \frac{\langle i^* | \hat{F} | i_A \rangle}{E_0 - E_{i^*}} \langle i | \hat{F}_2 | i^* \rangle \right) \\ &= \sum_A \sum_i^{i \in A} \left( 2n_i \langle i | \hat{F} | i \rangle + (2n_i)^2 \frac{|\langle i^* | \hat{F}_2 | i \rangle|^2}{E_0 - E_{i^*}} \right) . \end{aligned} \quad (120)$$

It can be deduced that

$$\begin{aligned} \langle \Psi_0 | \widehat{F} | \Psi_0 \rangle &= \sum_A \sum_i^{i \in A} \left( 2n_i E_i^o + (2n_i)^2 \frac{|\langle i^* | \widehat{F}_2 | i \rangle|^2}{E_0 - E_{i^*}} \right) \\ &= \sum_A \sum_i^{i \in A} \left( 2n_i E_i^o + (2n_i)^2 \frac{\frac{\sqrt{1-d^2}}{2} [\langle \chi_A^i | \widehat{F}_2 | \chi_A^i \rangle - \langle \chi_B^i | \widehat{F}_2 | \chi_B^i \rangle]^2}{E_0 - E_{i^*}} \right) \end{aligned} \quad (121)$$

$$E \approx \sum_A \sum_i^{i \in A} \left( 2n_i E_i^o + n_i^2 \frac{|\langle \chi_A^i | \widehat{F}_2 | \chi_A^i \rangle - \langle \chi_B^i | \widehat{F}_2 | \chi_B^i \rangle|^2}{E_0 - E_{i^*}} \right),$$

$$E \approx \sum_A E_A$$

with  $d^2 \ll 1$ . From this, we can define an atomic energy

$$E_A = E_A^o + E_A^P = \sum_i^{i \in A} \left( 2n_i E_i^o + n_i^2 \frac{V_i^2}{E_0 - E_{i^*}} \right), \quad (122)$$

with  $V_i = \langle \chi_A^i | \widehat{F}_2 | \chi_A^i \rangle - \langle \chi_B^i | \widehat{F}_2 | \chi_B^i \rangle$ . We are only interested in relative energies and disregard the constant contribution  $E_A^o$  because this concept has already been applied in the force field approach. The molecular polarization energy can be introduced

$$E^P = \sum_A E_A^P = \sum_A \sum_i^{i \in A} n_i^2 \frac{V_i^2}{E_0 - E_{i^*}}, \quad (123)$$

$E_A^P$  is the so-called atomic polarization energy. Any one-electron expectation value  $O_{\alpha\beta} = \langle \Psi_0 | \widehat{O}_{\alpha\beta} | \Psi_0 \rangle$  can be represented as a functional of  $E^P$  and vice versa using this approximation. Thus, it can be expanded in a Taylor series at a desired value  $O_{\alpha\beta}^0$

$$E^P(O_{\alpha\beta}^0) = \sum_A \left( E_A^P(O_{\alpha\beta}) + \frac{dE_A^P}{dO_{\alpha\beta}} \Delta O_{\alpha\beta} + \frac{d^2 E_A^P}{dO_{\alpha\beta}^2} \frac{\Delta O_{\alpha\beta}^2}{2} + o(\Delta O_{\alpha\beta}^3) \right), \quad (124)$$

with

$$\frac{dE_A^P}{dO_{\alpha\beta}} = \sum_{A'} \sum_i^{i \in A'} \left\{ \frac{\partial E_A^P}{\partial V_i} \left( \frac{dO_{\alpha\beta}}{dV_i} \right)^{-1} + \frac{\partial E_A^P}{\partial n_i} \left( \frac{dO_{\alpha\beta}}{dn_i} \right)^{-1} \right\}, \quad (125)$$



and

$$\begin{aligned} \frac{d^2 E_A^P}{dO_{\alpha\beta}^2} &= \sum_{A'} \sum_i^{i \in A'} \left\{ \frac{\partial^2 E_A^P}{\partial V_i^2} \left( \frac{dO_{\alpha\beta}}{dV_i} \right)^{-2} + \frac{\partial^2 E_A^P}{\partial V_i \partial n_i} \left( \frac{dO_{\alpha\beta}}{dV_i} \right)^{-1} \left( \frac{dO_{\alpha\beta}}{dn_i} \right)^{-1} + \right. \\ &\quad \left. \frac{\partial^2 E_A^P}{\partial n_i^2} \left( \frac{dO_{\alpha\beta}}{dn_i} \right)^{-2} + \frac{\partial^2 E_A^P}{\partial n_i \partial V_i} \left( \frac{dO_{\alpha\beta}}{dn_i} \right)^{-1} \left( \frac{dO_{\alpha\beta}}{dV_i} \right)^{-1} \right\}, \end{aligned} \quad (126)$$

with

$$\begin{aligned} \frac{\partial E_A^P}{\partial V_i} &= \frac{2n_i^2 V_i}{E_0 - E_{i^*}}, \\ \frac{\partial E_A^P}{\partial n_i} &= \frac{2n_i V_i^2}{E_0 - E_{i^*}}, \\ \frac{\partial^2 E_A^P}{\partial V_i^2} &= \frac{2n_i^2}{E_0 - E_{i^*}}, \\ \frac{\partial^2 E_A^P}{\partial n_i^2} &= \frac{2 V_i^2}{E_0 - E_{i^*}}, \\ \frac{\partial^2 E_A^P}{\partial V_i \partial n_i} &= \frac{4n_i V_i}{E_0 - E_{i^*}}, \\ \frac{dO_{\alpha\beta}}{dV_i} &= n_i^2 \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i P(O)_i^{\alpha'\beta'}, \\ \frac{dO_{\alpha\beta}}{dn_i} &= \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i \left( O_i^{\alpha'\beta'} + 2n_i P(O)_i^{\alpha'\beta'} V_i \right). \end{aligned} \quad (127)$$

We introduced a perturbation series for one-electron operators [see equation (79)] in order to obtain the following property:

$$O_i^{\alpha'\beta'} \gg 2n_i P(O)_i^{\alpha'\beta'} V_i. \quad (128)$$

Using this property, the energetic change due to chemical shift variation can be approximated

$$\begin{aligned} E^P(O_{\alpha\beta}^0) &= \sum_A \left( E^P(O_{\alpha\beta}) + \right. \\ &\quad \sum_i^{i \in A} \left( \frac{2 V_i}{\Delta E_i \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i P(O)_i^{\alpha'\beta'}} \right) \Delta O_{\alpha\beta} + \\ &\quad \sum_i^{i \in A} \left( \frac{2}{\Delta E_i n_i^2 \left( \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i P(O)_i^{\alpha'\beta'} \right)^2} \right) \frac{\Delta O_{\alpha\beta}^2}{2} + \\ &\quad \left. o(\Delta O_{\alpha\beta}^3) \right). \end{aligned} \quad (129)$$

All quantities of this formula are known or can be calculated, except  $\Delta E_i$ . We estimate it as follows:

The BPT charge polarization parameter  $P(q)_i$  is known

$$\begin{aligned}
 P(q)_i &= \frac{\langle i_A | \hat{q} | i_A^* \rangle}{\Delta E_i} \\
 &= \frac{2\sqrt{1-d^2} (\langle \chi_a^i | \hat{q} | \chi_a^i \rangle - \langle \chi_b^i | \hat{q} | \chi_b^i \rangle)}{\Delta E_i} \\
 &\approx \frac{2\sqrt{1-d^2} (-1e-0)}{\Delta E_i} \approx \frac{-2e}{\Delta E_i},
 \end{aligned} \tag{130}$$

and thus

$$\Delta E_i \approx \frac{-2e}{P(q)_i}. \tag{131}$$

Using this, we obtain the following expression of the pseudo potential

$$\begin{aligned}
 U^{O_{\alpha\beta}} &= E^P(O_{\alpha\beta}^0) - E^P(O_{\alpha\beta}) \\
 &= \sum_A \left( \sum_i^{i \in A} \frac{-P(q)_i V_i}{e \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i P(O)_i^{\alpha'\beta'}} \Delta O_{\alpha\beta} + \right. \\
 &\quad \left. \sum_i^{i \in A} \frac{-P(q)_i}{en_i^2 \left( \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i P(O)_i^{\alpha'\beta'} \right)^2} \frac{\Delta O_{\alpha\beta}^2}{2} + \right. \\
 &\quad \left. o(\Delta O_{\alpha\beta}^3) \right).
 \end{aligned} \tag{132}$$

The polarization energy change  $U^{O_{\alpha\beta}}$  is due to the deviation of an one-electron expectation value. It implies a new force contribution and is designated as a pseudo force. The experimental data are  $O_{\alpha\beta}^0 = O_{\alpha\beta}^E$  and the theoretical data are  $O_{\alpha\beta}^0 = O_{\alpha\beta}^T$ , such as  $\Delta O_{\alpha\beta} = O_{\alpha\beta}^E - O_{\alpha\beta}^T$ . We obtain the following expression by applying the harmonic approximation to the gradient of  $U^{O_{\alpha\beta}}$ ,

$$\vec{F}_j^{O_{\alpha\beta}} = -\frac{\partial U^{O_{\alpha\beta}}}{\partial x_j} \approx -k_{O_{\alpha\beta}} (O_{\alpha\beta}^E - O_{\alpha\beta}^T) \frac{\partial O_{\alpha\beta}^T}{\partial x_j}, \tag{133}$$

with the force constant

$$k_{O_{\alpha\beta}} = \sum_A \sum_i^{i \in A} \frac{2 P(q)_i}{en_i^2 \left( \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i P(O)_i^{\alpha'\beta'} \right)^2}. \tag{134}$$

The gradient of the one-electron expectation value is expressed as

$$\begin{aligned}
\frac{\partial O_{\alpha\beta}^T}{\partial x_j} &= \frac{\partial}{\partial x_j} \left\langle \Psi_0 \left| \widehat{O}_{\alpha\beta} \right| \Psi_0 \right\rangle \\
&= \sum_A \sum_i^{i \in A} \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i \left( \frac{\partial n_i}{\partial x_j} O_i^{\alpha'\beta'} + \right. \\
&\quad \left. 2n_i \frac{\partial n_i}{\partial x_j} P(O)_i^{\alpha'\beta'} \left[ \left\langle \chi_A^i \left| \widehat{V}_2 \right| \chi_A^i \right\rangle - \left\langle \chi_B^i \left| \widehat{V}_2 \right| \chi_B^i \right\rangle \right] + \right. \\
&\quad \left. n_i^2 P(O)_i^{\alpha'\beta'} \left[ \frac{\partial}{\partial x_j} \left\langle \chi_A^i \left| \widehat{V}_2 \right| \chi_A^i \right\rangle - \frac{\partial}{\partial x_j} \left\langle \chi_B^i \left| \widehat{V}_2 \right| \chi_B^i \right\rangle \right] \right). \tag{135}
\end{aligned}$$

The pseudo force (133) is proportional to the difference between the theoretical and the experimental data ( $O_{\alpha\beta}^T - O_{\alpha\beta}^E$ ) and is the result of the Taylor expansion of the atomic polarization energy (124). A scaling function is introduced in order to limit the pseudo force at large differences and to control the magnitude with respect to other force field interactions,

$$f_{scale} = S \frac{\exp\left(\frac{O_{\alpha\beta}^T - O_{\alpha\beta}^E}{\Delta_{\alpha\beta}}\right) - \exp\left(\frac{O_{\alpha\beta}^E - O_{\alpha\beta}^T}{\Delta_{\alpha\beta}}\right)}{\exp\left(\frac{O_{\alpha\beta}^T - O_{\alpha\beta}^E}{\Delta_{\alpha\beta}}\right) + \exp\left(\frac{O_{\alpha\beta}^E - O_{\alpha\beta}^T}{\Delta_{\alpha\beta}}\right)} = \frac{S}{\Delta_{\alpha\beta}} (O_{\alpha\beta}^T - O_{\alpha\beta}^E) + o^2, \tag{136}$$

$S$  controls the magnitude, and  $\Delta_{\alpha\beta}$  describes the relative sensitivity of the deviation ( $O_{\alpha\beta}^T - O_{\alpha\beta}^E$ ). The functional behavior is shown in Figure 7. In practise, it should be considered that  $S$  is in the order of  $\Delta_{\alpha\beta}$  ( $S \simeq \Delta_{\alpha\beta}$ ). The bond orbital matrix elements of the Fock operator

$$\left\langle \chi_\lambda \left| \widehat{F}_2 \right| \chi_\lambda \right\rangle \approx V_i = \left\langle \chi_a^i \left| \widehat{F}_2 \right| \chi_a^i \right\rangle - \left\langle \chi_b^i \left| \widehat{F}_2 \right| \chi_b^i \right\rangle \tag{137}$$

and their derivatives in respect to the coordinates of all atoms within the molecular system  $\frac{\partial V_i}{\partial x_j}$  have to be evaluated in order to determine the pseudo forces, see equation (133). Spherical coordinates are used for the computational routines

$$\frac{\partial V_i}{\partial x_j} = \frac{\partial V_i}{\partial R} \frac{\partial R}{\partial x_j} + \frac{\partial V_i}{\partial \cos(\theta)} \frac{\partial \cos(\theta)}{\partial x_j} + \frac{\partial V_i}{\partial \varphi} \frac{\partial \varphi}{\partial x_j}. \tag{138}$$

The integrals and their derivatives, in dependence on  $R$ ,  $\cos(\theta)$  and  $\varphi$  are given in the Appendix. They are necessary for  $sp^2$  and  $sp^3$  bond contributions. Figure 8 shows the bond coordinate system. The coordinate derivatives  $\frac{\partial R}{\partial x_j}$ ,  $\frac{\partial \cos(\theta)}{\partial x_j}$  and  $\frac{\partial \varphi}{\partial x_j}$  can be derived. The derivatives of the

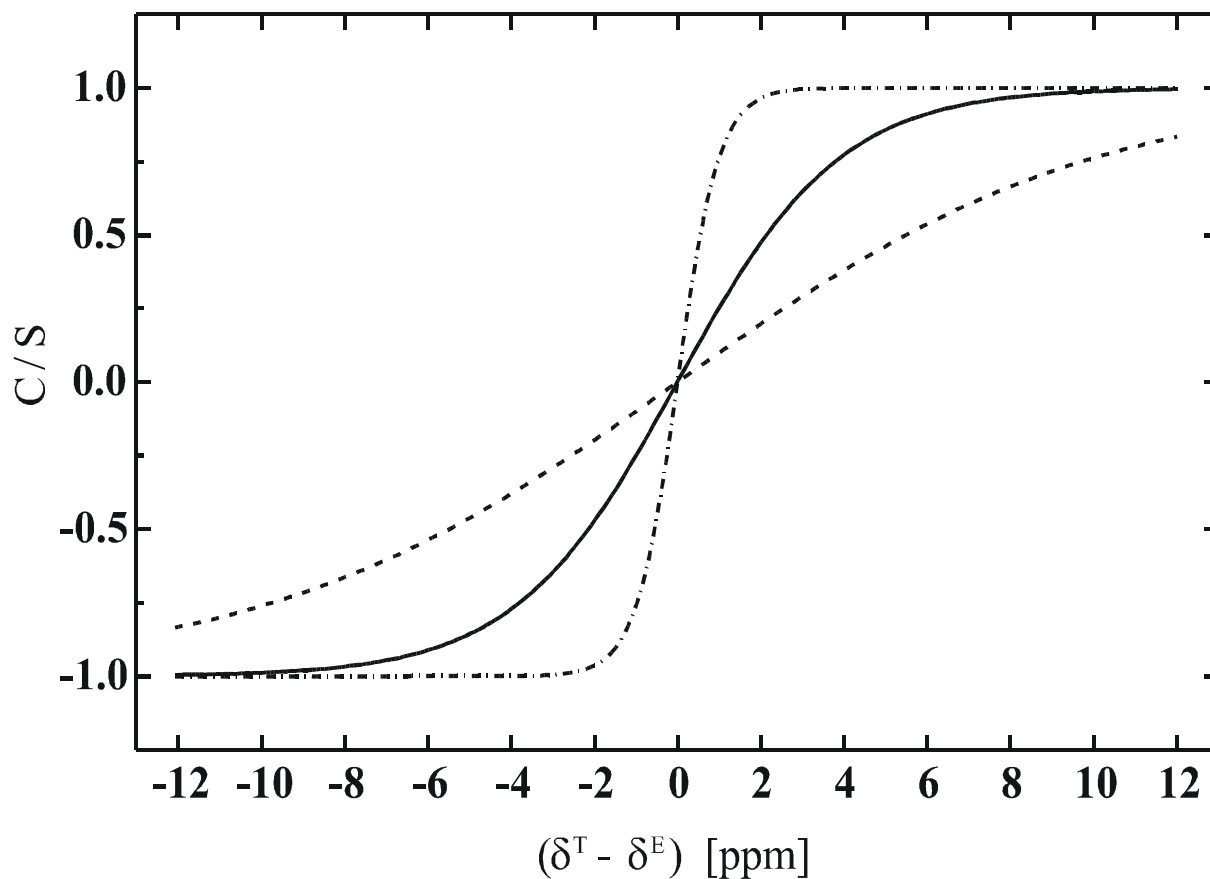


Figure 7

Functional dependence of the relative force scaling function  $f_{scale}/S$  on the difference between theoretical and experimental one-electron expectation value: (---)  $\Delta_{\alpha\beta}=1.0$ ; (—) with  $\Delta_{\alpha\beta}=3.9$  (resembles the standard deviation of the BPT chemical shifts in ppm) and (-·-)  $\Delta_{\alpha\beta}=10.0$ .

interconnecting vectors between a charge  $Q$  and the bonded atoms  $A$  or  $B$  are

$$\frac{\partial \left| \vec{AQ} \right|}{\partial A_i} = \frac{(\vec{Q}_i - \vec{A}_i)}{\sqrt{(Q_x - A_x)^2 + (Q_y - A_y)^2 + (Q_z - A_z)^2}} = +e_i^{AQ}, \quad (139)$$

$$\frac{\partial \left| \vec{AQ} \right|}{\partial Q_i} = -e_i^{AQ}, \quad (140)$$

$$\frac{\partial \left| \vec{BQ} \right|}{\partial B_i} = +e_i^{BQ}, \quad (141)$$

$$\frac{\partial \left| \vec{BQ} \right|}{\partial Q_i} = -e_i^{BQ}. \quad (142)$$

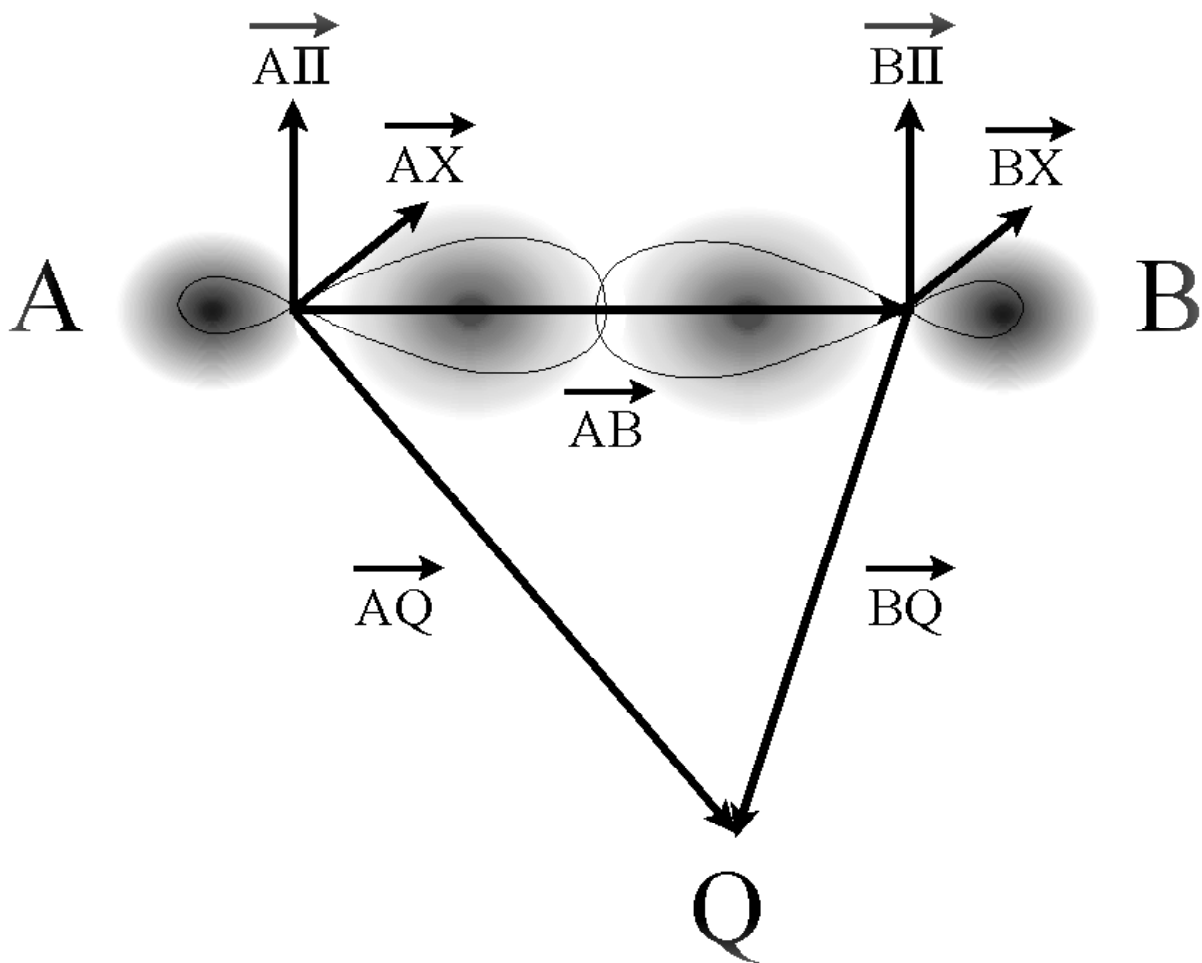


Figure 8  
Representation of the coordinates of the bond orbital matrix elements of the atoms  $A$  and  $B$ .  $Q$  is a polarizing point charge. The vectors  $\vec{A}\Pi$  and  $\vec{B}\Pi$  are the directions of the  $\pi$ -plane. Thus, a right handed coordinate system in  $A$   $\left( \vec{e}^{\vec{A}X}, \vec{e}^{\vec{A}\Pi}, \vec{e}^{\vec{A}B} \right)$  and a left handed system in  $B$   $\left( \vec{e}^{\vec{B}X}, \vec{e}^{\vec{B}\Pi}, \vec{e}^{\vec{B}A} \right)$  is defined.

The derivatives of the azimuthal angles are

$$\frac{\partial \cos(\angle BAQ)}{\partial A_i} = \frac{\partial \left( \vec{e}^{\vec{A}B} \vec{e}^{\vec{A}Q} \right)}{\partial A_i} = \frac{\partial \vec{e}^{\vec{A}B}}{\partial A_i} \vec{e}^{\vec{A}Q} + \frac{\partial \vec{e}^{\vec{A}Q}}{\partial A_i} \vec{e}^{\vec{A}B}, \quad (143)$$

$$\frac{\partial \cos(\angle BAQ)}{\partial Q_i} = \frac{\partial \vec{e}^{\vec{A}B}}{\partial Q_i} \vec{e}^{\vec{A}Q} + \frac{\partial \vec{e}^{\vec{A}Q}}{\partial Q_i} \vec{e}^{\vec{A}B}, \quad (144)$$

$$\frac{\partial \cos(\angle ABQ)}{\partial B_i} = \frac{\partial \vec{e}^{\vec{B}A}}{\partial B_i} \vec{e}^{\vec{B}Q} + \frac{\partial \vec{e}^{\vec{B}Q}}{\partial B_i} \vec{e}^{\vec{B}A}, \quad (145)$$

$$\frac{\partial \cos(\angle ABQ)}{\partial Q_i} = \frac{\partial \vec{e}^{\vec{B}A}}{\partial Q_i} \vec{e}^{\vec{B}Q} + \frac{\partial \vec{e}^{\vec{B}Q}}{\partial Q_i} \vec{e}^{\vec{B}A}. \quad (146)$$

The derivatives of the polar angles are

$$\frac{\partial \cos(\angle \Pi A Q)}{\partial A_i} = \frac{\partial}{\partial A_i} \left( \left[ \overleftarrow{e}^{AB} \times \overleftarrow{e}^{AQ} \right] \cdot \overleftarrow{e}^{A\Pi} \right), \quad (147)$$

$$= \left( \frac{\partial}{\partial A_i} \overleftarrow{e}^{AB} \times \overleftarrow{e}^{AQ} + \overleftarrow{e}^{AB} \times \frac{\partial}{\partial A_i} \overleftarrow{e}^{AQ} \right) \cdot \overleftarrow{e}^{A\Pi}, \quad (148)$$

$$\frac{\partial \cos(\angle \Pi A Q)}{\partial Q_i} = \left( \frac{\partial}{\partial Q_i} \overleftarrow{e}^{AB} \times \overleftarrow{e}^{AQ} + \overleftarrow{e}^{AB} \times \frac{\partial}{\partial Q_i} \overleftarrow{e}^{AQ} \right) \cdot \overleftarrow{e}^{A\Pi}, \quad (149)$$

$$\frac{\partial \cos(\angle \Pi B Q)}{\partial B_i} = \left( \frac{\partial}{\partial B_i} \overleftarrow{e}^{BA} \times \overleftarrow{e}^{BQ} + \overleftarrow{e}^{BA} \times \frac{\partial}{\partial B_i} \overleftarrow{e}^{BQ} \right) \cdot \overleftarrow{e}^{B\Pi}, \quad (150)$$

$$\frac{\partial \cos(\angle \Pi B Q)}{\partial Q_i} = \left( \frac{\partial}{\partial Q_i} \overleftarrow{e}^{BA} \times \overleftarrow{e}^{BQ} + \overleftarrow{e}^{BA} \times \frac{\partial}{\partial Q_i} \overleftarrow{e}^{BQ} \right) \cdot \overleftarrow{e}^{B\Pi}. \quad (151)$$

## 6.2. Conclusion

Energetic corrections of the molecular polarization energy provide additional forces within a force field that function to shift calculated expectation values in the direction of experimental data. The expectation value and its derivatives in dependence on coordinates have to be known in order to evaluate pseudo forces derived from one-electron expectation values. In a first step, the charges have to be calculated (see equation (??)). Afterwards, all other properties can be computed. All expressions can be evaluated very efficiently since they are analytically derived. In this work, the concept of BPT pseudo forces is applied for the  $^{13}\text{C}$  chemical shift, which is a localized expectation value. The sum over all atoms  $\sum_A$  breaks down, and the pseudo force becomes

$$\overleftarrow{F}_j^{CS_{\alpha\beta}} \approx -k_{CS_{\alpha\beta}} \left( \delta_{\alpha\beta}^E - \delta_{\alpha\beta}^T \right) \frac{\partial \delta_{\alpha\beta}^T}{\partial x_j}, \quad (152)$$

with

$$k_{CS_{\alpha\beta}} = \sum_i^{i \in A} \frac{2 P(q)_i}{en_i^2 \left( \sum_{\alpha'\beta'} D_{\alpha\alpha'}^i D_{\beta\beta'}^i P(\delta_{\alpha\beta})_i^{\alpha'\beta'} \right)^2}. \quad (153)$$

The computational cost depends, to the first degree, on the charge calculation which is proportional to the cube of the number of atoms,  $N^3$ . Determination of the charges, chemical shifts, pseudo energies and pseudo forces can be performed within seconds for systems of about 100 atoms on a Pentium II 350 MHz machine with an average performance of 200 Mflop/s. The

## 6. BPT Pseudo Forces

same calculation on a system of about  $10^4$  atoms takes roughly a day. This fact makes the method very feasible for molecular dynamics simulations and geometry optimization procedures, even for macromolecules or crystals.

## 7. Proton Position Refinement of the X-ray Crystal Structure of D-Mannitol

### 7.1. Introduction

The concept of  $^{13}\text{C}$  chemical shift pseudo forces can be applied to refine structures of large molecular systems. In this example, the proton positions of the mannitol crystal structure were refined (Witter et al.<sup>[135]</sup>). Because protons have no core electrons, their positions are not readily derived from X-ray investigations. Although, high resolution structures are possible for small molecules, it is necessary to improve the proton positions in large systems.  $\beta$ -D-mannitol is an example. The chiral and thus optical active acyclic sugar alcohol D-mannitol,  $\text{C}_6\text{H}_{14}\text{O}_6$ , is of medical and biological interest. D-mannitol crystallizes in a non-centrosymmetric  $\beta$  structure from a water solution in the orthorhombic space group  $\text{P}2_12_12_1$ . There is a X ray structure available, and its basic physical properties have been investigated.<sup>[136,137]</sup> In addition, cross-polarization, magic-angle spinning  $^{13}\text{C}$ -NMR spectra and the isotropic chemical shifts of D-mannitol have been published by Wasylishen et al.<sup>[138]</sup>

A comparison between the experimental and theoretical chemical shifts as derived by the BPT method, of the D-mannitol crystal structure shows a mean deviation of 1.7 ppm and a maximum difference of 2.7 ppm at the C1 carbon site. A proton position optimization procedure using crystallographic boundary conditions was applied during which the positions of the heavy atoms were held fixed. The central unit cell was replicated throughout space to simulate a periodic lattice with a total of 27 unit cells. Thus, there are no walls or surface molecules at the center unit cell. The positions of all other atoms in the neighborhood of the center unit cell were updated from the coordinates of the center. The forces act only on molecules belonging to the central cell. After 81 conventional conjugate gradient energy minimization steps and charge calculations, a much lower total energy was obtained. The mean chemical shift deviation increases up to 2.5 ppm.

### 7.2. Chemical Shift Driven Geometry Optimization

$^{13}\text{C}$  chemical shift pseudo forces were then switched on for the hydrogen sites. In the case of occurrence of differences between the target chemical shifts (the experimental values) and the theoretical chemical shifts, the local polarization energies of the carbon atoms will change (see



## 7. Proton Position Refinement of the X-ray Crystal Structure of D-Mannitol

Exp.	Theoretical $^{13}\text{C}$ Calculations							
	Scaling Constants:							
	$10^{-3}$	$10^{-2}$	$10^{-1}$	1	$10^1$	$10^2$	$10^3$	
Chemical Shift [ppm]								
C1	64.3	64.47	64.45	64.52	65.02	64.4	64.46	64.44
C2	71.7	73.17	73.22	73.06	72.58	71.77	71.74	71.71
C3	69.3	73.22	73.24	73.00	71.46	71.45	69.42	69.36
C4	67.4	71.47	71.46	71.32	70.57	70.02	67.81	67.74
C5	70.5	72.54	72.53	72.46	71.83	70.73	70.53	70.51
C6	62.8	63.24	63.25	64.52	63.14	62.83	62.72	62.71
$\text{SD}_\delta$	0	2.54	2.49	1.93	0.4	0.06	0.02	0.02
Standard Deviation of the Hydrogens in Respect to the X-ray Positions [ $\text{\AA}$ ]								
$\text{SD}_R$	0.195	0.199	0.198	0.189	0.188	0.265	0.128	0.132
Energy [kJ/mol]								
$E^{\text{pseudo}}$	-	-0.2	-1.9	-16.3	-101.8	-1077.1	-5043.7	-49314.5
$E^{\text{bonded}}$	-	-3.6	-2.0	69.6	1421.9	10600.4	9596.2	10632.1
$E^{\text{total}}$	-	-1776.0	-1760.0	-1003.0	13007.0	99008.0	50938.0	-381035.0

Table 5

Comparison of experimental and calculated data of D-mannitol Carbon atoms: experimental data (Column 2) and theoretical results of the NMR force field refined structure at various pseudo force scaling factors (Column 3 to 9).

equation (132)). These energy changes are derived in dependence on the coordinates and yield forces that reposition the hydrogens (see equation (133)). The carbon and oxygen sites were held stationary at their crystal positions. A range from  $10^{-3}$  to  $10^3$  was tested for the scaling constant  $S$ , see equation (136). The results are listed in Table 5.

The functional behavior of the chemical shift deviation in dependence on the logarithm of the scaling factor is shown in Figure 9. Little influence of the pseudo force is obtained with scaling from  $10^{-3}$  to  $10^{-2}$ . With scaling factors in this order of magnitude the forces are too weak to show any influence as regarding the repositioning of hydrogen atoms since the relative energy contribution (of the polarization energy) is too small as compared to the other energy terms in the force field. The total energy is about  $1.5 \cdot 10^5$  kJ/mol lower than the pure molecular force field result. Increasing the scale factor to  $10^{-1}$  or higher resulted in significant changes. The higher the scaling factor, the smaller the chemical shift deviation. At a scaling constant of  $10^2$ , the chemical shift deviation function reaches a plateau (see Figure 9) and further improvement beyond 0.02 ppm chemical shift deviation is not obtained by increasing the scale factor. An investigation of the average coordinate displacement of the protons with respect to the crystal

structure shows a decrease from 0.2 Å to 0.13 Å with respect to the scaling. At scaling constants higher than 100 no further improvement is obtained. There is an outlier at the value of 10. It can be understood by analyzing the total energy plot (Figure 9). In the COSMOS force field, the

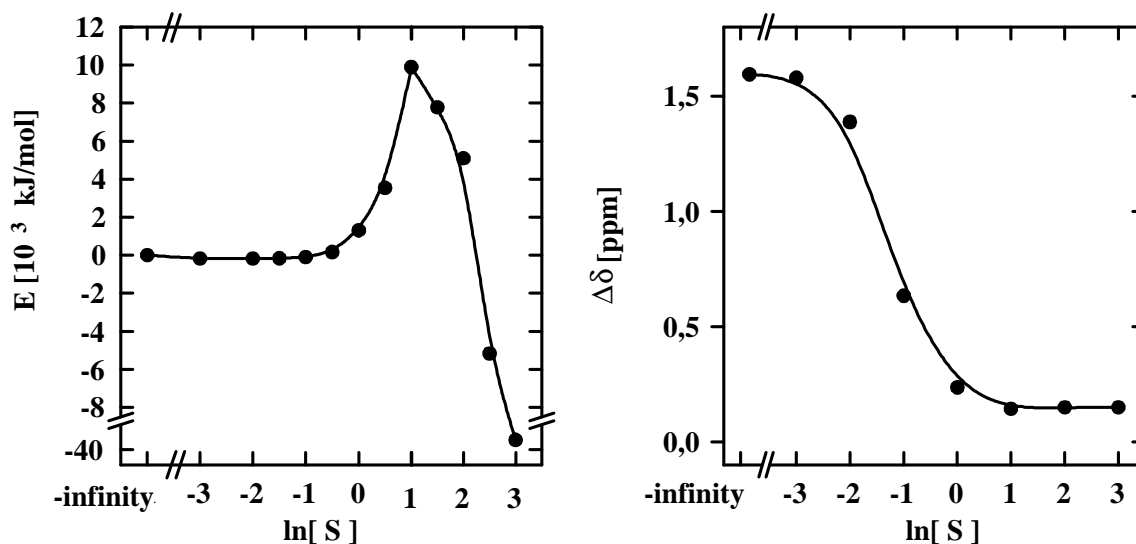


Figure 9

Left: total energy in dependence on the logarithm of the pseudo force scaling constant.

Right: theoretical chemical shift deviation from experiment in dependence on the logarithm of the pseudo force scaling constant.

non-bonded energy contributions dominate the total force field energy at small scaling factors. The pseudo forces increase with the scaling factor. Up to a scaling constant of  $S = 10$ , the harmonic bond length, bond angle and dihedral angle forces of the protons are comparably small. At scaling factors of 100 and higher, the bonded energy does not increase drastically, and the negative pseudo forces start to dominate all other contributions. Thus, the proton positions are controlled exclusively by the chemical shift pseudo forces at scaling factors of 100 and higher. The inaccuracies of the NMR experiment and the chemical shift calculations limit the precision of the structure determination. The refinement deviation is smaller than the uncertainties of the X-ray diffraction in this example. The given average proton displacement parameter derived from the temperature factor is about 0.2 Å. The standard deviation of the refined structure ( $S = 100$ ,  $S = 1000$ ) in respect to the X-ray structure is about 0.13 Å. Figure 10 shows the superposition of the X-ray and the proton refined structure ( $S = 1000$ ). The spheres at the proton positions are the isotropic 50% probability ellipsoids. The concept of the probability plots were taken from ORTEP-III.<sup>[139]</sup>

### 7.3. Conclusion

It is well known that hydrogen bond systems stabilize crystal structures. In the COSMOS force field,<sup>[90]</sup> hydrogen bonds are not treated with a separate energy term. However, a routine is included that recognizes hydrogen bonds. The total bond length contraction caused by H bond formation should be equal or greater to twice the van der Waals radius of the hydrogen atom (1.6 Å).<sup>[140]</sup> The H bond (X · · H) search is based on that rule: i.e. if the O · · H distance becomes lower than the mean covalent bond radii of H plus O plus 1.6 Å, a hydrogen bond is identified and the van der Waals interaction between these two atoms is switched off. One additional intra molecular hydrogen bond was detected after the chemical shift driven structure refinement (Witter et al.<sup>[135]</sup>), see Figure 10.

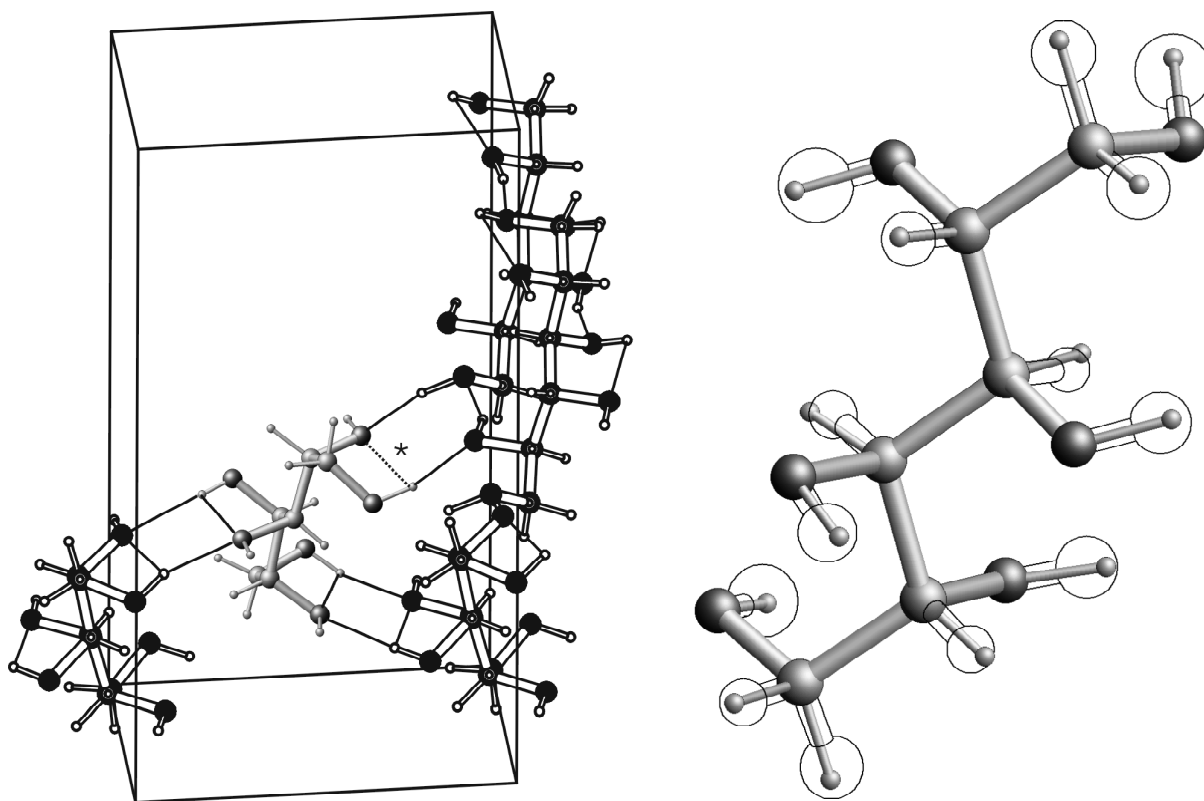


Figure 10

Left: refined  $\beta$ -D-mannitol crystal structure with sketched hydrogen bridges. The intramolecular hydrogen bond (\*) is newly detected after refinement.

Right: superposition of the X-ray and  $^{13}\text{C}$  refined structure (scaling constant  $S = 1000$ , solid structure). At the proton positions the X-ray 50% probability spheres are drawn.

## 8. 3D Solution Structure Determination of a Pseudopeptide Zinc Complex

### 8.1. Introduction

Bis(cysteinylyl) or bis(histidinyl) are the ligating units for zinc complexation in biological enzymes such as zinc fingers,<sup>[141,142]</sup> zinc twists, zinc clusters,<sup>[143,144,145]</sup> alcohol dehydrogenase,<sup>[146]</sup> metallothioneins<sup>[147]</sup> and carbonic anhydrase.<sup>[148]</sup> The protein sequences His–X–His or Cys–Y–Cys (X, Y are 1–4 amino acids) offer N, O or S atoms for coordination. The zinc ion is responsible for protein folding and catalytic binding of H<sub>2</sub>O or CO<sub>2</sub>. Tripeptides with bis(histidinyl) sequences have been investigated by Gockel et al.<sup>[149]</sup> Other authors have described tripodal histidine ligands,<sup>[150]</sup> pyrazolylborate ligands<sup>[151]</sup> as well as macrocyclic poly-amines.<sup>[152,153]</sup> In this work, as an attempt has been made to mimic the catalytic center of the carboanhydrase by using a His–X–His pseudotripeptide. A N–alkyl glycine derivative was used as residue X. Figure 11 shows the basic structure of this peptide.

Three compounds were investigated: the basic ligand Bz–His–Gly–His–NH<sub>2</sub> itself, the pseudotripeptide Bz–His–Ψ[CO–N(CH<sub>2</sub>)<sub>2</sub>–NH<sub>2</sub>]Gly–His–NH<sub>2</sub> and the zinc complex of the latter (Bz...benzoyl, Gly...glycine, His...histidine, Ψ...unusual structure motive). The basic peptide (His–Gly–His) and its dimeric complex has been previously investigated by Förster et al.<sup>[154]</sup> A N–alkyl chain was introduced in order to ensure monomeric complexation. Furthermore, the N–functionalized glycine residue leads to cis– and trans isomers of the peptide bond and is less flexible than the C<sub>α</sub>–substituents. This pseudotripeptide zinc complex was designed to bind H<sub>2</sub>O in aqueous solution. Since the complex did not dissolve very well in pure water, it was analyzed in a DMSO/H<sub>2</sub>O mixture.

The peptides were synthesized by solid state phase methods and purified with HPLC by Greiner et al.<sup>[155]</sup> The free ligand is quite soluble in H<sub>2</sub>O, in direct contrast to its zinc complex. Due to zinc coordination, the complex molecule loses four to five hydrophilic groups. This may also be a reason why we did not succeed in crystallizing the complex. Therefore, solution NMR techniques were used for structure determination. NMR methods generally are gaining in importance<sup>[156,157,158]</sup> in investigations of the metal-ligand-interactions in metallocomplexes. Modern methods use multidimensional NMR experiments combined with molecular mechanics simulations for structural determination.<sup>[159,1]</sup> Distance constraints are obtained from NOE intensities. <sup>13</sup>C chemical shifts provide additional information for the determination of the

conformation. The newly introduced method for structure refinement based directly on  $^{13}\text{C}$  chemical shifts<sup>[135]</sup> in connection with the COSMOS force field<sup>[90]</sup> was applied. Low energy conformations of the force calculations were optimized using a quantum chemical procedure and the chemical shifts were determined with a *ab initio* method.

## 8.2. Experimental

### 8.2.1. Modeling

An initial guess for the structure of the  $\text{Zn-Bz-His-}\Psi[\text{CO-N}(\text{CH}_2)_2\text{-NH}_2]\text{Gly-His-NH}_2$  complex was designed with the COSMOS program,<sup>[109]</sup> see Figure 11. The COSMOS force field<sup>[90]</sup> with coordinate dependent charges was applied. It was assumed, that the zinc ion coordinates with the imidazole rings. It is known that imidazole ligands in complexes are deprotonated.<sup>[160]</sup> They were therefore treated as negatively charged groups. A  $\text{H}_2\text{O}$  molecule was bound to the  $\text{Zn}^{2+}$  cation to account for the reaction step of the water activation of the carbonic anhydrase<sup>[161]</sup>. The N-alkyl chain was used to complete the most probable tetrahedral coordination sphere of zinc.

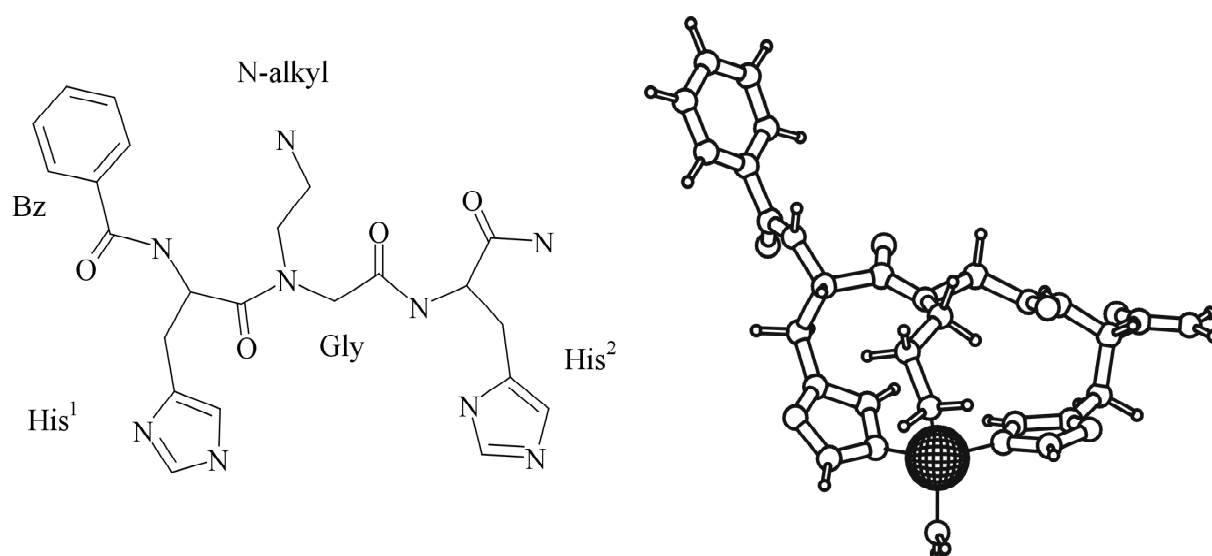


Figure 11

Left: basic structure of the peptide ligand.

Right: computer designed guess for zinc complexation of the pseudotriptide.

### 8.2.2. Synthesis

The syntheses of the peptides (Bz–His–Gly–His–NH<sub>2</sub> and Bz–His–Ψ[CO–N(CH<sub>2</sub>)<sub>2</sub>–NH<sub>2</sub>]Gly–His–NH<sub>2</sub>) were carried out with the solid phase methods on a semi-automated peptide synthesizer SP 650 (Bachem). The products were purified on a preparative HPLC LC-8A (Shimadzu) and checked by an analytical HPLC-System LC 10AT (Shimadzu). Zn(ClO<sub>4</sub>)<sub>2</sub> × 6 H<sub>2</sub>O was used for complex formation. MALDI mass spectra (Voyager-DETM RP Biospectrometry Workstation, PerSeptive Biosystems, Inc.) and high resolution ESI mass spectra were measured (MS MAT 95 XL Trap, Thermo Quest, Finnigan). A monomeric complexation was experimentally proven. For more details see Greiner et al.<sup>[155]</sup>

### 8.2.3. NMR Experiments

The <sup>13</sup>C spectra of the basic ligand and the pseudotriptide are illustrated in Figure 12. The

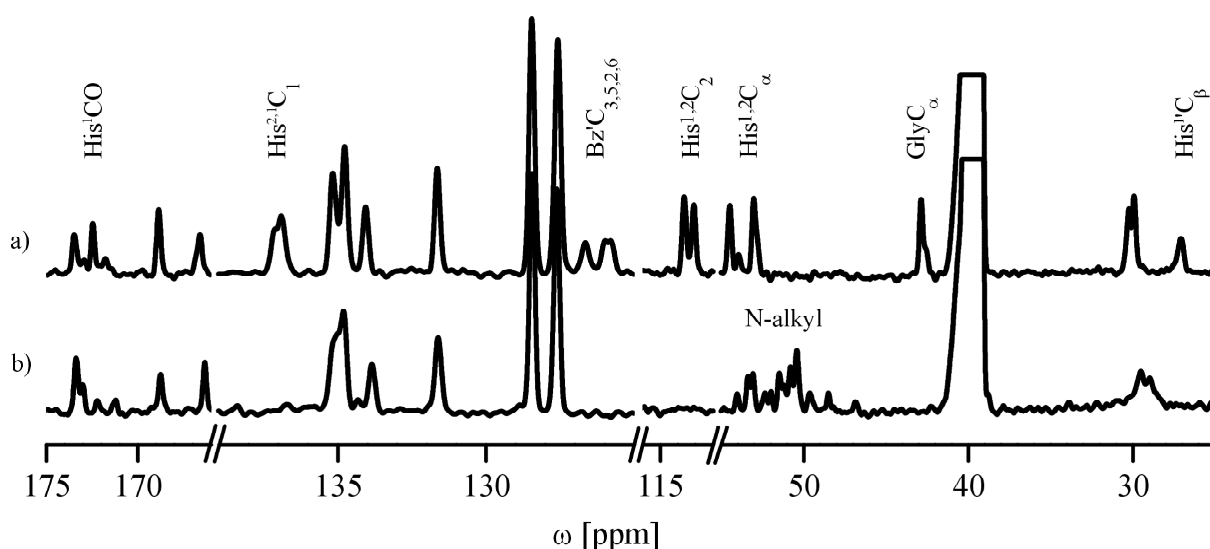


Figure 12

The <sup>13</sup>C spectra of a) the basic ligand Bz–His–Gly–His–NH<sub>2</sub> and b) the pseudotriptide Bz–His–[CO–N(CH<sub>2</sub>)<sub>2</sub>–NH<sub>2</sub>]Gly–His–NH<sub>2</sub>.

following solution NMR experiments were utilized for the full <sup>1</sup>H, <sup>13</sup>C and NOE assignment: 1D–<sup>1</sup>H, 1D–<sup>13</sup>C, 1D–<sup>13</sup>C Dept, 2D–COSY, 2D–TOCSY, 2D–NOESY, 2D–ROESY, 2D–HMBC and 2D–HSQC.<sup>[70]</sup> The experiments were carried out on a 350 MHz and a 500 MHz Bruker spectrometer. The resonances of His<sup>1</sup>CO, His<sup>2</sup>C<sub>1</sub>, His<sup>2</sup>C<sub>1</sub>, His<sup>1</sup>C<sub>2</sub> and His<sup>2</sup>C<sub>2</sub> are well-defined for the basic ligand, but are rather broad for the pseudotriptide. It is proved to be quite

difficult to extract data from 2D spectra due to this line broadening. The reason for this is the high flexibility of the peptide on the  $\mu\text{s}$  time scale. In addition, there are at least two stable conformations for the basic ligand which can be concluded from the appearance of second resonances for  $\text{Bz}'\text{C}_2$ ,  $\text{Bz}'\text{C}_3$ ,  $\text{Bz}'\text{C}_5$ ,  $\text{Bz}'\text{C}_6$ ,  $\text{His}^1\text{C}_\beta$  and  $\text{His}^1\text{CO}$ . The  $\text{GlyC}_\alpha$  resonance is shifted downfield for the pseudopeptide due to the additional N-alkyl chain. In Figure 12 b, it is obvious that the bulk of the  $\text{CH}_2$  resonances around 50 ppm contain information about the conformation of  $\text{GlyC}_\alpha$ ,  $\text{His}^1\text{C}_\alpha$ ,  $\text{His}^2\text{C}_\alpha$ ,  $\text{GlyNC}_1$  and  $\text{GlyNC}_2$ . In contrast to the two free ligands, the  $^{13}\text{C}$  spectra of the complex (Figure 13 and 14) display broader lines. Two extra  $^{13}\text{C}$

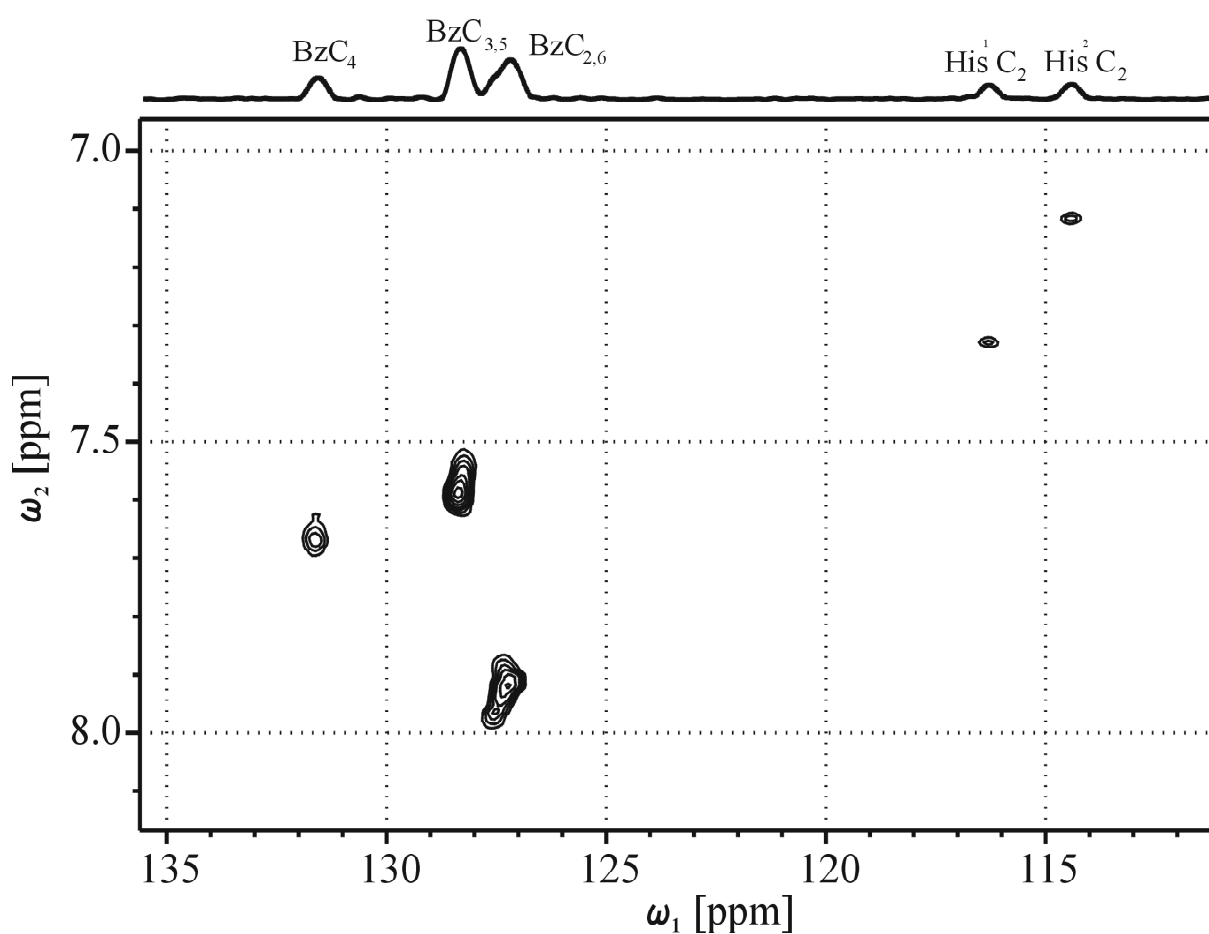


Figure 13

The part of aromatic resonances of the  $^1\text{H}$ - $^{13}\text{C}$  correlation spectrum (HSQC) of the pseudotriptide complex.

shieldings can be assigned to the  $\text{GlyNC}_2$  functionality and an additional one to the  $\text{BzCO}$  group. However, the intensities of these signals in the 1D spectrum are rather low which indicates a narrow conformational space. Surprisingly, the  $\text{His}^2\text{CO}$  functionality could not be assigned, a fact possibly due to fast rotation of the  $-\text{CONH}_2$  group at the chain end. The  $^{13}\text{C}$  NMR data is given in Table 6 and yields insights into the type of complexation of the pseudopeptide in

DMSO/H<sub>2</sub>O. The carbon chemical shifts were shifted downfield at certain positions possibly due to the vicinity of the positive zinc charge. The zinc ion has a large influence upon the positions of the His<sup>1</sup>C<sub>2</sub>, GlyCO, GlyC<sub>α</sub>, His<sup>2</sup>C<sub>4</sub> and His<sup>2</sup>C<sub>α</sub> shifts. The most important NOE connectivity was found between one proton from -NH<sub>2</sub> of the N-alkyl chain and His<sup>1</sup>C<sub>4</sub>H. This corresponds to the three complexes determined by the COSMOS force field refinement illustrated in Figure 15.

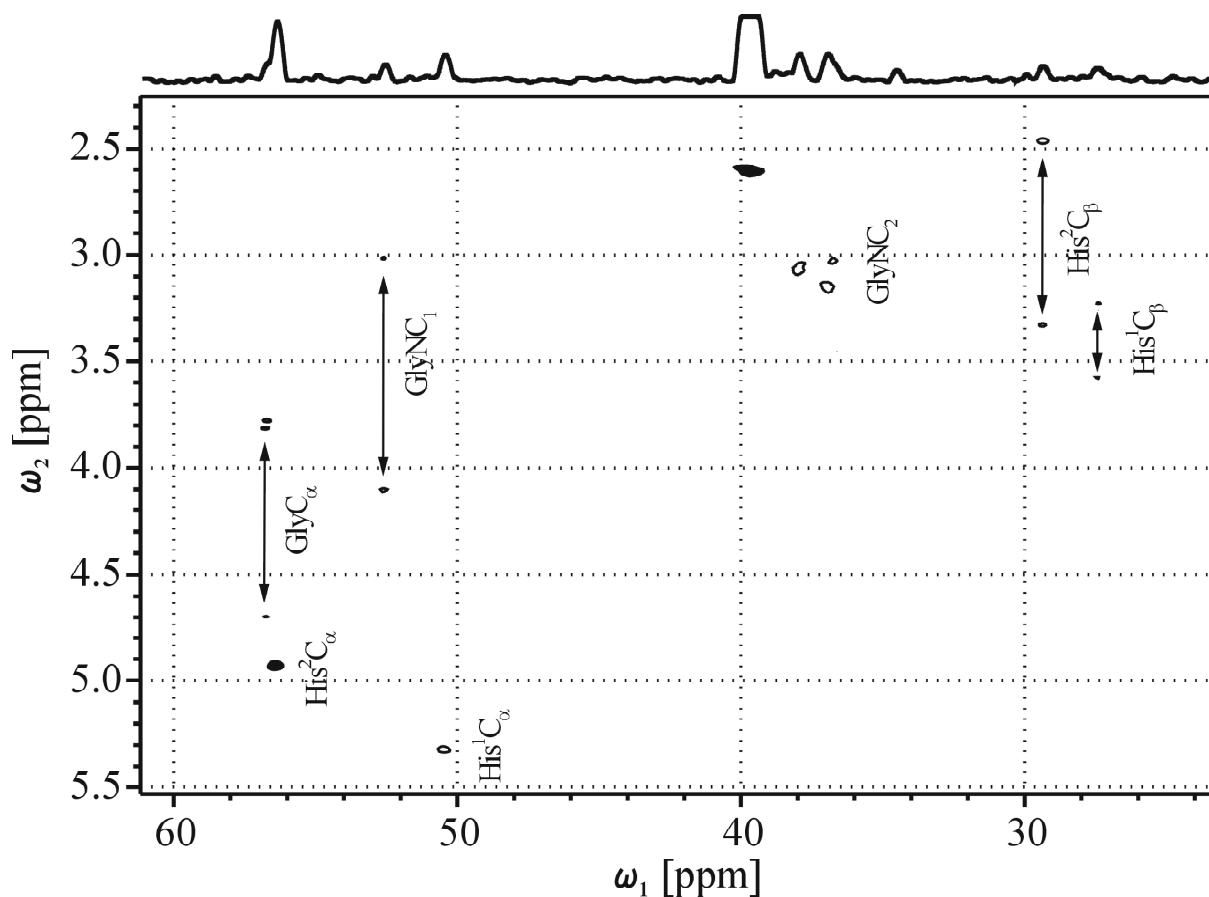


Figure 14  
<sup>13</sup>C-<sup>1</sup>H correlation (HSQC) spectrum section of the aliphatic <sup>13</sup>C chemical shifts of the zinc complex.

### 8.3. NOE Pseudo Forces

NMR parameters contain the coordinate information of the molecular structure averaged on the  $\mu$ s time scale. The NOE signals, J couplings and the chemical shifts are most notable in the liquid state. We therefore needed coordinate dependent theoretical models of the NMR parameters for our dynamic structure investigation that are applicable to metalloptides and



metaloproteins.  $^1\text{H}$  NOE signals depend largely on the cross relaxation between two protons. In contrast to the free ligand, the complex is rather rigid and the isotropic overall motion can be separated from the internal motions. The model free approach of Lipari & Szabo<sup>[3]</sup> leads to a functional behavior of roughly the inverse square of the third power of the effective NOE distance at first approximation. The Karplus equation<sup>[4]</sup> and similar relations<sup>[62,63,64,65,66,67,68]</sup> are well established models for the J couplings of the backbone or side chain dihedral angle dependencies of proteins and peptides. However, it is unclear whether these methods are sufficient for a complete description of pseudopeptide complexes. We therefore omitted these parameters in further considerations. Coordinate dependent  $^{13}\text{C}$  chemical shifts were calculated for the complex using the BPT approach.

The 3D structure elucidation was carried out with high temperature molecular dynamics followed by a simulated annealing procedure with NOE distance constraints. The structures were refined by geometry optimization with additional chemical shift restrictions. The COSMOS force field<sup>[90]</sup> was used since the computational effort required for molecular dynamics, simulated annealing and geometry optimizations with *ab initio* methods is still too large. Within the COSMOS force field approach, the coordinate dependent charge distribution was calculated using the semi-empirical bond polarization theory (BPT). Zinc had to be treated as a  $\text{Zn}^{2+}$  ion in the conformational search for the most stable complex. The van der Waals (VdW) radius of zinc had to be adjusted within the force field with respect to the BPT charge calculation.

$^1\text{H}$  NOE restrictions were considered first. They correspond to intermolecular interactions of proton pairs rather than valence forces. They were therefore treated as harmonic perturbations of the electrostatic and VdW proton pair energy

$$E^{el} + E^{VdW} = (E^{el} + E^{VdW})_{NOE} \left( 1 + S \frac{(R^T - R^{NOE})^2}{2 \Delta R^2} \right). \quad (154)$$

The variable  $S$  is the sign of the total non-bonded energy at the NOE distance. It ensures an energetic minimum with respect to the NOE restriction.  $\Delta R$  describes an acceptable deviation of the difference between the experimental and the theoretical NOE distance and should be chosen to be between 0.1 and 1 Å. A single NOE pseudo force added to the non-bonded forces of a proton becomes

$$F_{\alpha}^{NOE} = (E^{el} + E^{VdW})_{NOE} \frac{(R^T - R^{NOE})}{\Delta R^2} \frac{\partial R}{\partial x_{\alpha}}. \quad (155)$$

The VdW energy at the NOE distance can be calculated exactly. The hydrogen atomic charges for the Coulomb energy do not vary much with conformational changes and can be approximated with the starting values. Therefore, the energy minimum constant can be determined from the beginning.

#### 8.4. Simulation and NMR Structure Refinement

The modeled complex in Figure 11 (omitting the water molecule) was employed as initial structure for a 1 ns molecular dynamics calculation at a temperature of 2000 K in order to insure that the conformational energy barriers are surmountable. 1000 structures originated from this calculation were then selected and cooled to 0 K by applying simulated annealing methods with additional NOE pseudo forces. Due to the simplicity of our NOE distance model and the flexibility of the peptide complex, a half side harmonic NOE pseudo potential was employed; i.e. if any proton pair distance was equal to or closer than the desired NOE distance, the forces were set to zero. The functional behavior of the potential and its gradient were kept continuous by this assumption. Thus, only approximate proximities were obtained which resulted in rough structures. In order to obtain better structures, 1000 geometry optimizations with additional  $^{13}\text{C}$  chemical shift pseudo forces were performed. From the 22 available  $^{13}\text{C}$  chemical shifts, 17 were used.

The pseudo forces for the chemical shift prediction of  $\text{CH}_2$ -group carbon sites were omitted since the current BPT parameterization for these atom groups does not work satisfactorily. After the optimizations, the three energetic lowest conformations with the smallest NMR parameter violations (Table 7) were selected from the conformational space of all 1000 structures. These COSMOS-NMR force field structures are shown in Figure 15. The rms NOE distance deviation after refinement is about 0.5 and the rms difference between restricted  $^{13}\text{C}$  chemical shifts and experimental values is 0.1 ppm. If the  $-\text{CH}_2$  carbon sites are also considered, the deviation is 1.2 ppm (see Table 7).

The energetically most stable structure (of the BPT calculations) is illustrated in Figure 15a. The  $\text{NH}_2$ -group position of the N-alkyl chain is important since it is the most obvious structural difference between the first two conformations. The complex of Figure 15a possesses two additional stabilizing hydrogen bridges as compared to conformation 15b. In contrast to 15a and 15b, structure 15c shows two trans-cis alternations and an imidazole ring flip. In order to

## 8. 3D Solution Structure Determination of a Pseudopeptide Zinc Complex

C Atom	Base ligand	Ligand	Complex		
	A	B	C	C <sup>1</sup>	C <sup>2</sup>
BzC <sub>1</sub>	134.1	134.9	134.0	124.0	135.5
BzC <sub>2</sub> / BzC <sub>6</sub>	127.6 (125.8)	127.6	127.5	118.0	129.8
BzC <sub>3</sub> / BzC <sub>5</sub>	128.5 (126.7)	128.5	128.5	118.4	130.0
BzC <sub>4</sub>	131.7	131.6	131.8	121.6	133.8
BzCO	166.6	166.3	165.9 (166.9)	154.6	167.8
His <sup>1</sup> C <sub>1</sub>	136.9	135.1	133.6	131.1	145.2
His <sup>1</sup> C <sub>2</sub>	113.6	~113	116.4	117.8	128.0
His <sup>1</sup> C <sub>4</sub>	136.9	135.1	133.5	135.9	149.5
His <sup>1</sup> C <sub>β</sub>	29.9 (27.1)	29.0	27.7	28.7	32.1
His <sup>1</sup> C <sub>α</sub>	54.5 (54.0)	50.4	50.8	56.1	58.5
His <sup>1</sup> CO	172.5 (171.8)	173.0 (173.4)	174.7	167.9	184.4
GlyCO	168.9	168.8	172.3	164.4	182.1
GlyC <sub>α</sub>	42.9	50.8	56.6	48.9	52.1
GlyNC <sub>1</sub>	-	51.5 (48.3)	52.8	41.7	43.8
GlyNC <sub>2</sub>	-	39.7 (38.8, 38.6)	37.1 (38.0)	38.9	42.2
His <sup>2</sup> C <sub>1</sub>	137.1	135.1	135.6	130.0	142.7
His <sup>2</sup> C <sub>2</sub>	113.1	~113	114.2	114.5	122.8
His <sup>2</sup> C <sub>4</sub>	134.8	133.9	135.7	134.1	146.2
His <sup>2</sup> C <sub>β</sub>	30.3	29.5	29.1	28.1	28.9
His <sup>2</sup> C <sub>α</sub>	53.1	53.4 (53.1)	56.3	59.3	64.8
His <sup>2</sup> CO	173.5 (172.5)	172.2 (171.2)	-	155.8	171.1

Table 6

Experimental and theoretical <sup>13</sup>C NMR chemical shift data of the base ligand A (Bz-His-Gly-His-NH<sub>2</sub>), the ligand B (Bz-His-Ψ[CO-N(CH<sub>2</sub>)<sub>2</sub>-NH<sub>2</sub>]Gly-His-NH<sub>2</sub>) and the zinc complex C (Zn-Bz-His-Ψ[CO-N(CH<sub>2</sub>)<sub>2</sub>-NH<sub>2</sub>]Gly-His-NH<sub>2</sub>).

<sup>1</sup> SCF-GIAO B3LYP/6-31G(d,p) calculation

<sup>2</sup> SCF-GIAO B3LYP/6-311+G(d,p) calculation

Conformation	a)	b)	c)
$\Delta R$ [Å]	0.7	0.5	0.2
$\Delta R_{truncated}$ [Å]	0.3	0.3	0.2
$\Delta\delta$ [ppm]	1.2	1.2	1.2
$\Delta\delta_{truncated}$ [ppm]	0.1	0.1	0.1
$E_{COSMOS}$ [kJ/mol]	0	77.5	90.5
$E_{B3LYP}$ [kJ/mol]	6.3	0	42.5
$r$ [Å]	2.06 (0.15)	2.06 (0.14)	1.99 (0.07)
$\alpha$ [°]	109.2 (10.6)	09.1 (11.1)	120.0 (7.0)

Table 7

NMR refinement data of the three conformations a, b and c of Figure 15: rows 2 to 5 show the NOE distance deviation, the truncated NOE distance deviation (only distances are reported that are larger than the NOE distances), the <sup>13</sup>C chemical shift deviation from experiment and the shift deviation without CH<sub>2</sub>-group carbon sites. Rows 6 and 7 show the COSMOS force field energies and DFT minimum energies. The mean Zn-X bond lengths and the X-Zn-Y bond angles of the DFT optimized structures are given in rows 8 and 9. The deviation of the latter can be seen in parentheses.

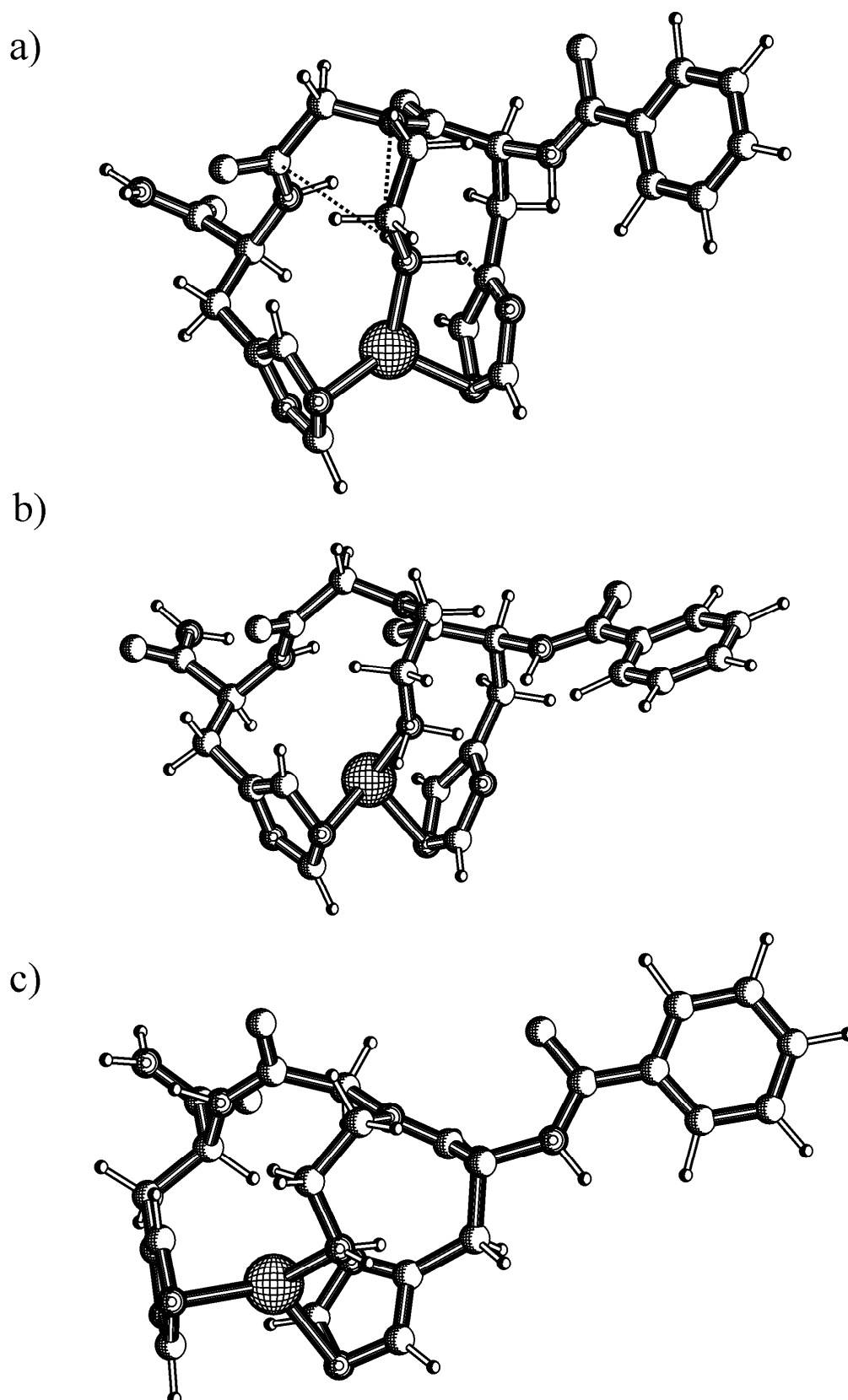


Figure 15

Lowest energy structures of the COSMOS-NMR force field results refined with NOE under inclusion of chemical shift pseudo forces. X-Zn-Y bonds are shown if the interatomic distance is shorter than 2.5 Å.

find energetic minima of the free complex nearest to the NMR optimized structures, B3LYP/6-31G(d,p)<sup>[162]</sup> full geometry optimizations were carried out with the GAUSSIAN98 program package<sup>[50]</sup> using the COSMOS-NMR structures as the starting point. The DFT optimized structure 15a is shown in Figure 16. A comparison of Figure 15a and Figure 16 shows that

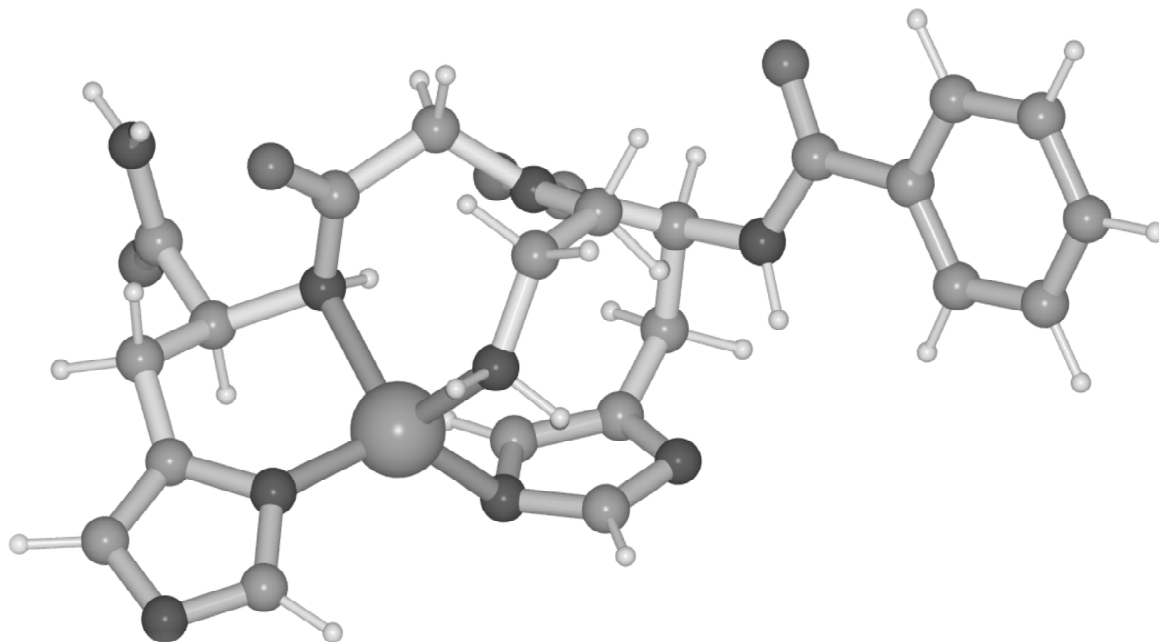


Figure 16  
B3LYP/6-31G(d,p) DFT optimized structure of the conformation in Figure 15a.

the position of the  $\text{NH}_2$ -group has changed and the hydrogen bridges were not detectable. This might be due to the fact that no diffuse functions were used. The B3LYP energy difference between 15a and 15b is rather small (6.3 kJ/mol). The corresponding energy difference in the COSMOS force field approach amounts to 77.5 kJ/mol and mainly originates from the three hydrogen bonds (see Figure 15).

The main difference between the DFT optimized structures 15a and 15b is the position of the  $-\text{CONH}_2$  group in His<sup>2</sup>. The average Zn-X bond length is ca.  $(2.0 \pm 0.1)$  Å. For the tetrahedral complexation of 15a and 15b, the bond angle is  $(109 \pm 12)^\circ$ . The theoretical angle would be  $109.3^\circ$ . The DFT optimized structure of 15c demonstrates a threefold coordination with a bond angle of  $(120 \pm 7)^\circ$ . The deviation from planarity is ca.  $4^\circ$ . This conformation could possibly be interpreted as a transition state structure for  $\text{H}_2\text{O}$  binding.

In order to confirm this possibility, a water molecule was added to structure 15c near the zinc ion and a full DFT geometry optimization was carried out. The minimum structure is shown

in Figure 17. The DFT relative energy difference (in respect to structure 15a together with a water molecule) is -45.8 kJ/mol and shows that the described complex can indeed be interpreted as a transition structure during water activation of the catalytic reaction cycle of the carbonic anhydrase. The total reaction has the form:  $\text{CO}_2 + 2\text{H}_2\text{O} \rightleftharpoons \text{HCO}_3^- + \text{H}_3\text{O}^+$ .<sup>[161]</sup>

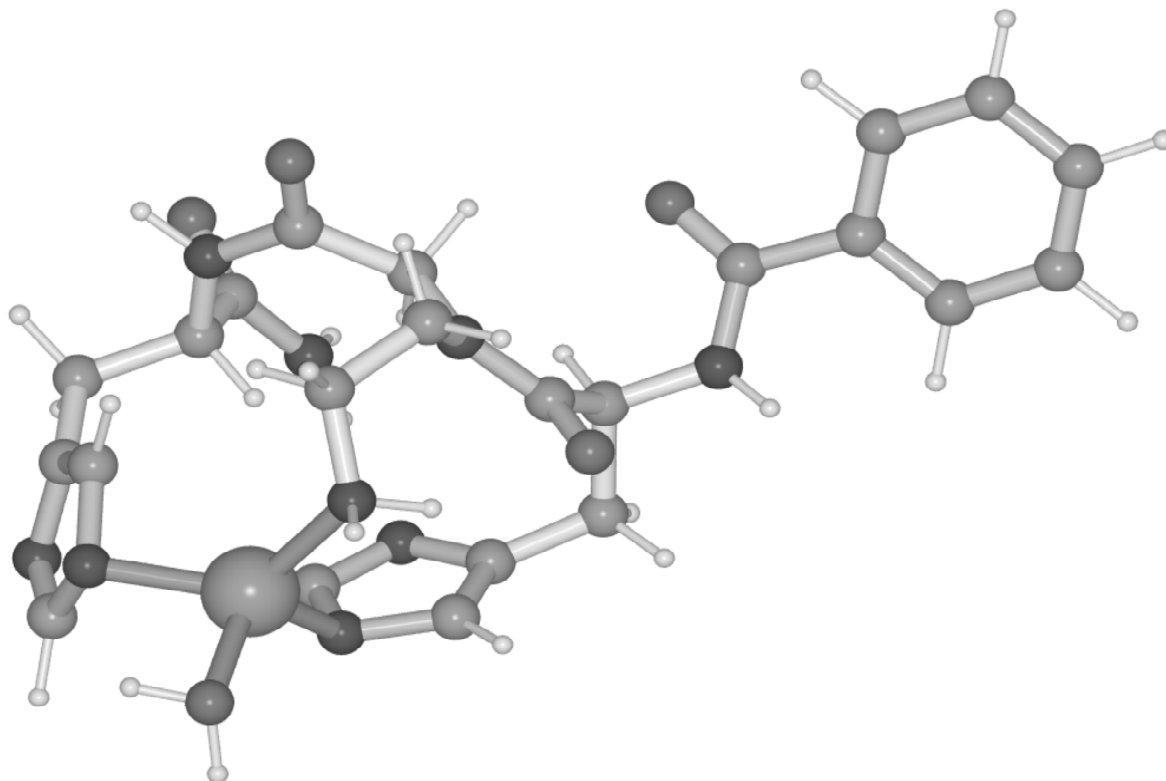


Figure 17

B3LYP/6-31G(d,p) DFT optimized structure of the complex Figure 15c with an additional  $\text{H}_2\text{O}$  molecule to demonstrate the possibility of a tetrahedral coordination.

## 8.5. Conclusion

The pseudotripeptide  $\text{Zn-Bz-His-}\Psi[\text{CO-N}(\text{CH}_2)_2\text{-NH}_2]\text{Gly-His-NH}_2$  was modeled based on a structure motive of the carbonic anhydrase reaction center. It was experimentally synthesized and its structure determined (Figure 15). NOE and  $^{13}\text{C}$  chemical shifts were used in the COSMOS-NMR force field as direct constraints. For the first time,  $^{13}\text{C}$  chemical shift pseudo forces have been applied to metal peptide complexes. This method could also be applied to metalloproteins. The force field structures were compared with the nearest *ab initio* minima and the energies were quantified. Finally the isotropic  $^{13}\text{C}$  chemical shifts were calculated

using the GIAO<sup>[46]</sup> B3LYP method. Two different basis sets were used: 6-31G(d,p) and 6-311G(d,p). The average over of all three structures of Figure 15a, b, c is shown in Table 6. As reference the TMS shieldings were computed (187.1 ppm and 178.9 ppm). The calculated average shifts correlate very well with the experimental data of the complex,  $R_c(6-31G)=0.9944$  as well as  $R_c(6-311G)=0.9945$  (Figure 18) and of the ligand itself with  $R_l(6-31G)=0.9938$  as well as  $R_l(6-311G)=0.9937$ . The standard deviations are  $SD_c(6-31G)=5.1$  ppm,  $SD_c(6-311G)=5.6$  ppm,  $SD_l(6-31G)=5.4$  ppm and  $SD_l(6-311G)=6.0$  ppm respectively (Witter et al.<sup>[163]</sup>). One cannot assume that  $R_c$  is much closer to unity than  $R_l$  or that  $SD_c$  is much smaller

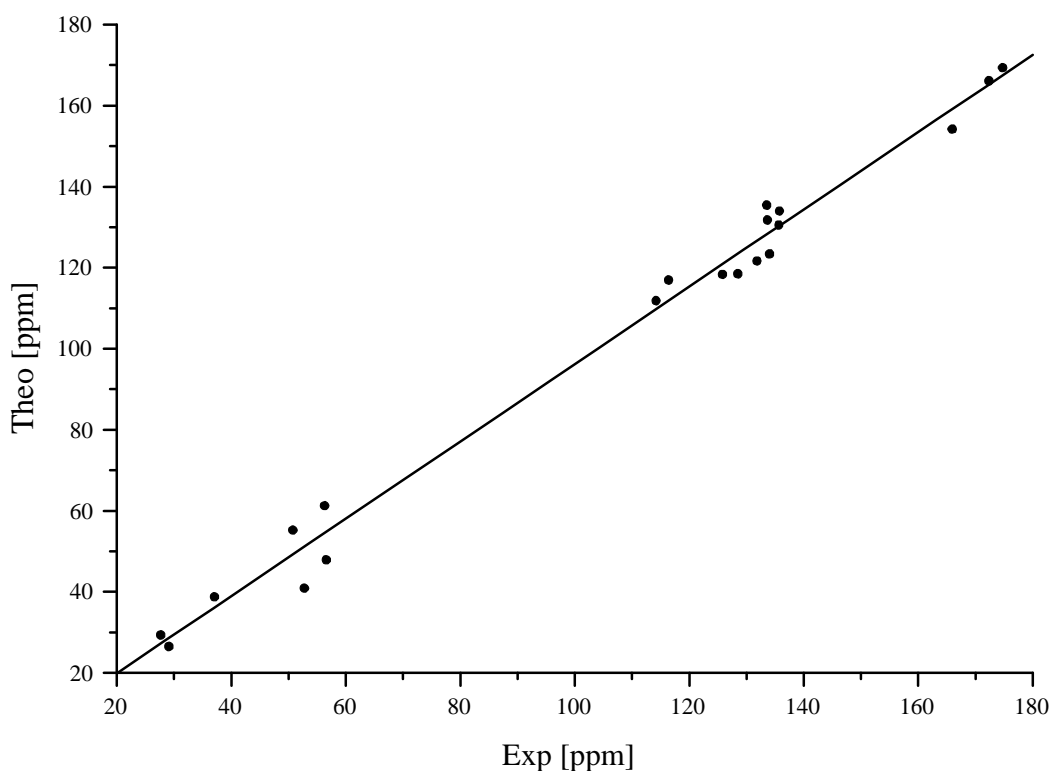


Figure 18

Correlation of the experimental and GIAO/6-31G(d,p) <sup>13</sup>C chemical shifts of the complex. The correlation coefficient is 0.9944 and the standard deviation is 5.1 ppm.

than  $SD_l$  since the correlation between experimental shifts of the pseudopeptide and the complex differs only by 0.001 from unity and the standard deviation is only 2.2 ppm. In addition, we used only DFT minimum structures and it is known that the GIAO method predicts chemical shifts at this level of theory with an average error of 5 ppm. Nevertheless, the GIAO chemical

## 8. 3D Solution Structure Determination of a Pseudopeptide Zinc Complex

shifts of the COSMOS–NMR force field complex structures, optimized by DFT, correlate best with the experimental data of the pseudotriptide zinc complex.

Last, but not least, it turned out that a structure analogous to the catalytic step of the water activation of the carboanhydrase is possible. A stable state of threefold coordination was found (Witter et al.<sup>[163]</sup>).



## 9. 3D Crystal Structure Refinement of Silk II

Crystal molecular dynamics simulations of the Silk II model of Takahashi et al.<sup>[124]</sup> were described in section 5.4. The resulting 200 minimum structures were optimized by applying additional chemical shift pseudo forces at the carbon sites within a unit cell. The experimental chemical shift standard deviation of the original crystallographic structure was 1.5 ppm. Geometry optimization yielded a decrease to about 1.0 ppm. Application of additional chemical shift pseudo forces produced a deviation of 0.5 ppm for the structure with the lowest total energy of 200 minimized structures. The original structure and the best NMR refined structure are shown in Figure 19. The energy of the hydrogen bridges were calculated with the COSMOS

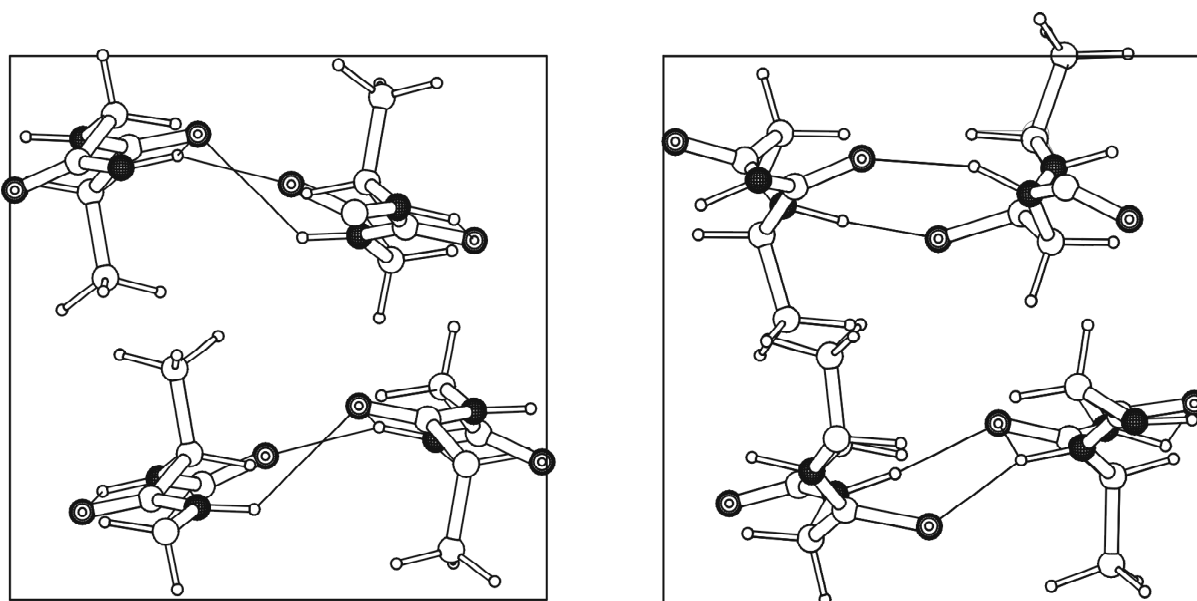


Figure 19

Left: original Takashi crystal structure<sup>[124]</sup>.

Right: best NMR refined structure. The Hydrogen bridges make the structure more stable. The energy gain per unit cell is 26 kJ/mol compared to the original crystal structure. The view of the unit cells show that the chains are oriented parallel to the  $\bar{c}$ -direction (perpendicular to the plane of the figure).

force field. The NMR refined structure showed a stabilization of 26 kJ/mol per unit cell which can be attributed to the newly formed hydrogen bridges. A microcrystalline polymer structure refinement was carried out for the first time with this procedure.

## 10. Results and Conclusion

The major theme of this dissertation is the improvement and application of the bond polarization theory (BPT) for 3D structure determinations. Within the BPT approach, ideal bonds of a molecular system are polarized by a surrounding charge distribution that can be approximated by atomic point charges. Delocalizations are neglected. The efficient calculation of one electron expectation values becomes very feasible within this concept. Contributions of ideal bonds are parameterized, and hence coordinates depend only on the location of the polarizing charges. This ensures the efficient application in a force field approach.

The BPT method was parameterized with different atomic charge models, electrostatic charges (ESC), Mulliken population analysis (MPA) and natural population analysis (NPA). The basis set dependence of these methods was tested for eleven basis sets (leading to 6079 quantum chemical calculations and 36 parametrizations). For every charge model, the best basis set was chosen to carry out the parametrization based on 175 structures consisting of H, C, N, O, F, Si, P, S, Cl and Zn atoms. The NPA model performed best. A correlation coefficient of 0.9961 was obtained between BPT/NPA and DFT/NPA with a standard deviation of 0.05  $e$ . For the first time, a parameter stability scale for compact basis sets was obtained. This stability was demonstrated by the correlation of BPT/NPA parameters derived from different compact basis sets, i.e. 3–21G and 6–31G(d,p). The parameters  $q_i$  have a  $R$ -value of 0.9973 ( $SD = 0.03 e$ ), while the parameters  $P(q)_i$  correlate with 0.9969 ( $SD = 0.26 e/H$ ).

The average  $^{13}\text{C}$  chemical shift tensor data were obtained by crystal molecular dynamics simulation of the Silk II model of Takahashi et al.<sup>[124]</sup> with the COSMOS force field including the application of the BPT/NPA charge model for the first time. There was a very good agreement between the averaged BPT chemical shift tensors and the experimental values obtained by Witter et al.<sup>[130]</sup> The correlation coefficient of the theoretical and experimental data is 0.996 and the standard deviation is 6.8 ppm. A complete 2D  $^{13}\text{C}$  NMR Iso-Aniso spectrum was simulated for this structure the first time.

In this work, the BPT has been formulated for non-local tensorial one-electron expectation values. Their coordinate derivatives were analytically derived. A molecular polarization energy was introduced that is a functional of such expectation values. Thus, polarization energy corrections respective additional contributions to the force field are derived. These pseudo forces shift calculated expectation values towards experimental data. The expectation value and its derivatives in dependence on the coordinates have to be calculated in order to determine the

one-electron expectation value pseudo forces. Once, the charges are computed, all other properties can be obtained. The computational efficiency of this method is obvious since all expressions are analytically derived. In this work, the concept of BPT pseudo forces was extended to  $^{13}\text{C}$  chemical shift calculations. The chemical shift pseudo forces are applied in geometry optimizations, molecular dynamics and simulated annealing simulations. They act on the 3D structure as long as there are differences between the theoretical and experimental expectation values.

The proton positions of the X-ray  $\beta$ -D-mannitol crystal structure were refined (Witter et al.<sup>[135]</sup>) based on the concept of  $^{13}\text{C}$  chemical shift pseudo forces. The average coordinate deviation of the NMR refined structure as compared to the crystal structure (0.13 Å) is smaller than the uncertainty of the X-ray proton position prediction (0.2 Å). One additional intramolecular hydrogen bond was obtained after the chemical shift driven structure refinement as compared to the original structure.

Furthermore, a three dimensional structure suggestion of a pseudotriptide Zn-Bz-His- $\Psi$ [CO-N(CH<sub>2</sub>)<sub>2</sub>-NH<sub>2</sub>]Gly-His-NH<sub>2</sub> in solution (DMSO/H<sub>2</sub>O) was derived. The resulting COSMOS-NMR force field structures were analyzed with density functional theory methods. A tetrahedral zinc complex formation is most probable. There is also a structure with threefold symmetry that is analogous to the water activation step of carboanhydrase. This fact is remarkable because the pseudotriptide was modeled according to the structure motive of the carbonic anhydrase reaction center (Witter et al.<sup>[163]</sup>).

Finally, a first crystal structure refinement of the Silk II model of Takahashi et al.<sup>[124]</sup> was carried out using experimental  $^{13}\text{C}$  chemical shifts (Witter et al.<sup>[130]</sup>). An additional stabilizing hydrogen bridge system is suggested for the crystal structure of this molecule.

In this dissertation, important contributions to the three dimensional structure determination of molecular systems in solution and solid state are presented based on NMR chemical shifts.

## 11. Bibliography

- [1] Wüthrich, K., *NMR of Proteins and Nucleic Acids*, Wiley Interscience, New York, 1986.
- [2] Jardetzky, O. and Finucane, M.D., *Dynamics, Structure and Function of Biological Macromolecules*, IOS Press, Amsterdam, 2001.
- [3] Lipari, G. and Szabo, A., *J. Am. Chem. Soc.*, 1982, **104**, 4546.
- [4] Karplus, M., *J. Chem. Phys.*, 1959, **30**, 11.
- [5] Spera, S. and Bax, A., *J. Am. Chem. Soc.*, 1991, **113 / 14**, 5490.
- [6] Sternberg, U., *J. Mol. Phys.*, 1988, **63**, 249.
- [7] Pauli, W., *Naturwiss.*, 1924, **12**, 741.
- [8] Frisch, R. and Stern, O., *Z. Physik*, 1933, **85**, 4.
- [9] Estermann, I. and Stern, O., *Z. Physik*, 1933, **85**, 17.
- [10] Estermann, I. and Stern, O., *Phys. Rev.*, 1934, **46**, 665.
- [11] Breit, G. and Rabi, I.I., *Phys. Rev.*, 1931, **38**, 2082.
- [12] Bloch, F., Hansen, W., and Packard, M.E., *Phys. Rev.*, 1946, **69**, 127.
- [13] Bloch, F., *Phys. Rev.*, 1946, **70**, 460.
- [14] Purcell, E.M., Torrey, H.G., and Pound, R.V., *Phys. Rev.*, 1946, **69**, 37.
- [15] Torrey, H.C., *Phys. Rev.*, 1949, **75**, 1326.
- [16] Torrey, H.C., *Phys. Rev.*, 1949, **76**, 1059.
- [17] Hahn, E.L., *Phys. Rev.*, 1949, **76**, 145.
- [18] Hahn, E.L., *Phys. Rev.*, 1950, **80**, 297.
- [19] Hahn, E.L., *Phys. Rev.*, 1950, **80**, 580.
- [20] Anderson, W.A. and Ernst, R.R., *US Patent No. 3.475.680*, 1965.
- [21] Ernst, R.R. and Anderson, W.A., *Rev. Sci. Instrum.*, 1966, **37**, 93.
- [22] Ernst, R.R., *Adv. Magn. Reson.*, 1966, **2**, 1.
- [23] Varian, R.H., *US Patent No 3.287.629*, 1956.
- [24] Primas, H., *Helv. Phys. Acta*, 1961, **34**, 36.
- [25] Ernst, R.R. and Primas, H., *Helv. Phys. Acta*, 1963, **36**, 583.
- [26] Dadok, J. and Specher, R.F., *J. Magn. Reson.*, 1974, **13**, 243.
- [27] Gupta, R.K., Ferretti, J.A., and Becker, E.D., *J. Magn. Reson.*, 1974, **13**, 275.
- [28] Ferretti, J.A. and Ernst, R.R., *J. Chem. Phys.*, 1976, **65**, 4283.
- [29] Tomlinson, B.L. and Hill, H.D.W., *J. Chem. Phys.*, 1973, **59**, 1775.
- [30] Comisarow, M.B. and Marshall, A.G., *Chem. Phys. Lett.*, 1974, **25**, 282.

- [31] McGurk, J.C., Mäder, H., Hofmann, R.T., Schmalz, T.G., and Flygare, W.H., *J. Chem. Phys.*, 1974, **61**, 3759.
- [32] Bowman, M.K. In ed. Kevan, L. and Bowman, M.K., *Modern Pulsed and Continuous-Wave Electron Spin Resonance*, p. 1, New York, 1990. J. Wiley.
- [33] Warren, W.S. and Zewail, A.H., *J. Chem. Phys.*, 1981, **75**, 5956.
- [34] Ernst, R.R., Bodenhausen, G., and Wokaun, A., *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford, 1989.
- [35] Smith, S., Levante, T., Meier, B., and Ernst, R., *J. Magn. Reson.*, 1994, **106**, 75.
- [36] Bak, M., Rasmussen, J.T., and Nielsen, N.C., *J. Magn. Reson.*, 2000, **26**, 1.
- [37] Bak, M., Schultz, R., Vosegaard, T., and Nielsen, N. C., *J. Magn. Reson.*, 2002, **154**, 28.
- [38] Mason, J., *Sol. Stat. Nucl. Magn. Reson.*, 1993, **2**, 285.
- [39] Dong, R.Y., *Nuclear Magnetic Resonance of Liquid Crystals*, Springer, New York, 1997.
- [40] Ramsey, N.F., *Phys. Rev.*, 1950, **91**, 303.
- [41] Ramsey, N.F., *Phys. Rev.*, 1950, **78**, 699.
- [42] Helgaker, T., Jaszunski, M., and Ruud, K., *Chem. Rev.*, 1999, **99**, 293.
- [43] Helgaker, T. and Jorgensen, P. In ed. Wilson, S. and Diercksen, G.H.F., *Methods in Computational Molecular Physics*, p. 353, New York, 1992. Plenum Press.
- [44] Kutzelnigg, W., *Theor. Chim. Acta*, 1992, **83**, 263.
- [45] Grotendorst, J., *Modern Methods and Algorithm of Quantum Chemistry*, NIC, Forschungszentrum Jülich, 2000.
- [46] Wolinski, K., Hilton, H.F., and Pulay, P., *J. Am. Chem. Soc.*, 1990, **112**, 8251.
- [47] Cizek, J., *J. Chem. Phys.*, 1966, **45**, 4256.
- [48] Cizek, J., *Adv. Chem. Phys.*, 1969, **14**, 35.
- [49] Cizek, J. and Paldus, J., *Int. J. Quantum Chem.*, 1971, **5**, 359.
- [50] *GAUSSIAN 98 Revision A.5*, Gaussian Inc., Pittsburgh, 1998.
- [51] Schmidt, M.W., Baldridge, K.K., Boatz, J.A., Elbert, S.T., Gordon, M.S., Jensen, J.H., Koseki, S., Matsunaga, N., Nguyen, K.A., Su, S., Windus, T.L., Dupuis, M., and Montgomery, J.A., *J. Comput. Chem.*, 1993, **14**, 1347.
- [52] Helgaker, T., Jensen, H.J.Aa., Jørgensen, P., Olsen, J., Ruud, K., Ågren, H., Auer, A.A., Bak, K.L., Bakken, V., Christiansen, O., Coriani, S., Dahle, P., Dalskov, E.K., Enevoldsen, T., Fernandez, B., Hättig, C., Hald, K., Halkier, A., Heiberg, H., Hettema, H., Jonsson, D., Kirpekar, S., Kobayashi, R., Koch, H., Mikkelsen, K.V., Norman, P., Packer, M.J., Pedersen, T.B., Ruden, T.A., Sanchez, A., Saue, T., Sauer, S.P.A., Schimmelpfennig, B., Sylvester-

- Hvid, K. O., Taylor, P.R., and Vahtras, O., *A Molecular Electronic Structure Program: Release 1.2*, 2001.
- [53] Stanton, J.F., Gauss, J., Watts, J.D., Nooijen, M., Oliphant, N., Perera, S.A., Szalay, P.G., Lauderdale, W.J., Kucharski, S.A., Gwaltney, S.R., Beck, S., Balková, A., Bernholdt, D.E., Baeck, K.K., Rozyczko, P., Sekino, H., Hober, C., and Bartlett, R.J., *A program product of the Quantum Theory Project*, University of Florida, Florida, 2002.
- [54] Schmidt-Rohr, K. and Spiess, W., *Multidimensional Solid-State NMR and Polymers*, Academic Press, London, 1994.
- [55] Gayathri, C., Bothner-By, A. A., van Zijl, P. C. M., and MacLean, C., *Chem. Phys. Lett.*, 1982, **87**, 192.
- [56] Bothner-By, A. A. In ed. Grant, D.M. and Harris, R.K., *Encyclopedia of Nuclear Magnetic Resonance*, p. 2932, Chichester, 1995. Wiley.
- [57] Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H., *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 9279.
- [58] Tjandra, N. and Bax, A., *Science*, 1997, **278**, 1111.
- [59] Brunner, E., *Concepts in Magn. Reson.*, 2001, **13**, 238.
- [60] Slichter, C.P., *Principles of Magnetic Resonance*, Springer-Verlag, Berlin, 1990.
- [61] Bloembergen, N. and Rowland, T.J., *Phys. Rev.*, 1955, **97**, 1679.
- [62] Pardi, A., Billeter, M., and Wüthrich, K., *J. Mol. Bio.*, 1984, **180**, 741.
- [63] Smith, L.J., Sutcliffe, M.J., Redfield, C., and Dobson, C.M., *Biochemistry*, 1991, **30**, 986.
- [64] Ludvigsen, S., Andersen, K.V., and Poulsen, F.M., *J. Mol. Bio.*, 1991, **217**, 731.
- [65] Vuister, G.W. and Bax, A., *J. Am. Chem. Soc.*, 1993, **115**, 7772.
- [66] DeMarco, A., Llinas, M., and Wüthrich, K., *Biopolymers*, 1978, **17**, 617.
- [67] Wang, A.C. and Bax, A., *J. Am. Chem. Soc.*, 1995, **117**, 1810.
- [68] Bystrov, V.F., *Progr NMR Spectr.*, 1976, **10**, 41.
- [69] Vuister, G.W., Delaglio, F., and Bax, A., *J. Am. Chem. Soc.*, 1992, **114**, 9674.
- [70] Braun, S., Kalinowski, H.O., and Berger, S., *150 and More Basic NMR-Experiments*, Wiley-VCH, New York, 1998.
- [71] Herrmann, T., Güntert, P., and Wüthrich, K., *J. Mol. Biol.*, 2002, **319**, 209.
- [72] Dinur, U. and Hagler, A.T. In ed. Lipkowitz, K.B. and Boyd, D.B., *Reviews in Computational Chemistry, Volume II*, p. 99, New York, 1991. VCH.
- [73] Boyd, D.B. In ed. Lipkowitz, K.B. and Boyd, D.B., *Reviews in Computational Chemistry, Volume VI*, p. 317, New York, 1995. VCH.

- [74] Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M., *J. Comp. Chem.*, 1983, **4**, 187.
- [75] MacKerell, A.D.Jr., Wiórkiewicz-Kuczera, J., and Karplus, M., *J. Am. Chem. Soc.*, 1995, **117**, 11946.
- [76] *QUANTA*, Molecular Simulations Inc., Waltham, MA, 2002.
- [77] Momany, F.A., Klimkowski, V.J., and Schäfer, L., *J. Com. Chem.*, 1990, **11**, 654.
- [78] Momany, F.A., Rone, R., Kunz, H., Frey, R.F., Newton, S.Q., and Schäfer, L., *J. Mol. Struct.: THEOCHEM*, 1993, **1**, 286.
- [79] Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M.Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A., *J. Am. Chem. Soc.*, 1995, **117**, 5179.
- [80] Kaminski, G., Duffy, E.M., Matsui, T., and Jorgensen, W.L., *J. Phys. Chem.*, 1994, **98**, 13077.
- [81] van Gunsteren, W.F. and Berendsen, H.J.C, *GROMOS: Groningen Molecular Simulation Library Manual*, Biomos, Groningen, 1987.
- [82] Clark, M., Cramer, R.D., and Opdenbosch, N.J., *J. Comp. Chem.*, 1989, **10**, 982.
- [83] Meier, R.J., *Vib. Spectrosc.*, 1996, **10**, 319.
- [84] Burkert, U. and Allinger, N.L., *MM2: Molecular Mechanics*, ACS Monograph, Washington DC, 1982.
- [85] Allinger, N.L., Yuh, Y., and Lii, J.H., *J. Am. Chem. Soc.*, 1989, **111**, 8551.
- [86] Allinger, N.L., Chen, K., Katzenellenbogen, J.A., Wilson, S.R., and Anstead, G.M., *J. Comp. Chem.*, 1996, **17**, 747.
- [87] Halgren, T.A., *J. Comp. Chem.*, 1996, **17**, 616.
- [88] Maple, J.R., Hwang, M.J., Stockfisch, T.P., Dinur, U., Waldman, M., Ewig, C.S., and Hagler, A.T., *J. Comp. Chem.*, 1994, **15**, 162.
- [89] Dinur, U. and Hagler, A.T., *J. Comp. Chem.*, 1994, **15**, 919.
- [90] Möllhoff, M. and Sternberg, U., *J. Mol. Mod.*, 2001, **7**, 90.
- [91] van Gunsteren, W.F. and Berendsen, H.J., *C. Angew. Chem. Int. Ed. Engl.*, 1990, **29**, 992.
- [92] Steinbach, P.J. and Brooks, B.R., *J. Comp. Chem.*, 1994, **15**, 667.
- [93] Brunger, A.T., *X-PLOR 3.1: A system for X-ray crystallography and NMR*, Yale University Press, New Haven, 1992.

- [94] MacKerell, A.D., Brooks, Jr.B., Brooks, C.L., Nilsson, L., Roux, B., Won, Y., and Karplus, M. In ed. Schleyer, P.v.R., *The Encyclopedia of Computational Chemistry*, p. 271, Chichester, 1998. John Wiley and Sons.
- [95] Bruccoleri, R.E., *Molecular Simulations*, 1993, **10**, 151.
- [96] Koch, F.T., Losso, P., and Sternberg, U., [www.comos-software.de](http://www.comos-software.de), 2003.
- [97] Re, G. Del, *J. Chem. Soc.*, 1958, p. 4031.
- [98] Slater, J.C., *Phys. Rev.*, 1930, **36**, 57.
- [99] Burns, G., *J. Chem. Phys.*, 1964, **41**, 1521.
- [100] Malrieu, J.P., *Mod. theor. Chem.*, 1977, **7**, 69.
- [101] O'Keefe, M. and Brese, N.E., *J. Am. Chem. Soc.*, 1991, **113**, 3226.
- [102] Singh, U.C. and Kollman, P.A., *J. Comp. Chem.*, 1984, **5**, 129.
- [103] Mulliken, R.S., *J. Chem. Phys.*, 1955, **23**, 1833, 1841, 2338, 2343.
- [104] Collins, J.B. and Streitwieser, A., *J. Comp. Chem.*, 1980, **1**, 81.
- [105] Reed, A.E., Weinstock, R.B., and Weinhold, F., *J. Chem. Phys.*, 1985, **83**, 735.
- [106] Uytterhoeven, L., Mortier, W.J., and Geerlings, P., *J. Phys. Chem. Solids*, 1989, **50**, 479.
- [107] Henry, M., *CHEMPHYSICHEM*, 2002, **3**, 561.
- [108] Koch, F.T., Möllhoff, M., and Sternberg, U., *J. Comp. Chem.*, 1994, **15**, 524.
- [109] Koch, F.T., Bräuer, M., Kunert, M., Sternberg, U., and Anders, E., *J. Mol. Mod.*, 2001, **7**, 54.
- [110] Pines, A., Gibbly, M.G., and Waugh, J.S., *J. Chem. Phys.*, 1972, **56**, 1776.
- [111] Pines, A., Gibby, M.G., and Waugh, J.S., *J. Chem. Phys.*, 1973, **59**, 569.
- [112] Bloembergen, N. and Sorokin, P., *Phys. Rev.*, 1958, **100**, 865.
- [113] Hartmann, S.R. and Hahn, E.L., *Phys. Rev.*, 1962, **128**, 2042.
- [114] Oldfield, E., *J. Bio. NMR*, 1995, **5**, 217.
- [115] Binkley, J.S., Pople, J.A., and Hehr, W.J., *J. Am. Chem. Soc.*, 1980, **102**, 939.
- [116] Hariharan, P.C. and Pople, J.A., *Theo. Chim. Acta*, 1973, **28**, 213.
- [117] Laws, D.D., Le, H., de Dios, A.C., Havlin, R.H., and Oldfield, E., *J. Am. Chem. Soc.*, 1995, **117**, 9542.
- [118] Sternberg, U. and Priess, W., *J. Magn. Reson.*, 1997, **125**, 8.
- [119] Priess, W. and Sternberg, U., *J. Mol. Struc.: THEOCHEM*, 2001, **544/1-3**, 181.
- [120] Veeman, W.S., *Progress in NMR Spectroscopy*, 1984, **20**, 193.
- [121] Sherwood, M.H., Alderman, D.W., and Grant, M.G., *J. Magn. Reson.*, 1989, **84**, 466.



- [122] Evans, D.J. and G.P.Morriss, *Statistical Mechanics of Nonequilibrium Liquids*, Academic Press, London, 1990.
- [123] Zhao, C. and Asakura, T., *Prog. Nuc. Magn. Res. Spec.*, 2002, **39**, 301.
- [124] Takahashi, Y., Gekoh, M., and Yuzuriha, K., *Int. J. Bio. Macromol.*, 1999, **24**, 127.
- [125] Fossey, S. A., Nemethy, G., Gibson, K. D., and Scheraga, H. A., *Biopolymers*, 1991, **31**, 1529.
- [126] Asakura, T., Demura, M., Date, T., and Miyahita, N., *Biopolymers*, 1997, **41**, 193.
- [127] Kameda, T. and Asakura, T., *Ann. Rep. NMR Spec.*, 2002, **46**, 101.
- [128] Grant, D.M. and Harris, R.K., *Encyclopedia of Nuclear Magnetic Resonance*, 1996, **2**, 1298.
- [129] Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P., *Numerical Recipes in C++: The Art of Scientific Computing*, Cambridge University Press, Cambridge, 2002.
- [130] Witter, R. and Sternberg, U., *J. Am. Chem. Soc.*, 2003, **submitted**.
- [131] Zhou, P., Li, G., Shao, Z., Pan, X., and Yu, T., *J. Phys. Chem.*, 2001, **105**, 12469.
- [132] Demura, M., Minami, M., Asakura, T., and Cross, T., *J. Am. Chem. Soc.*, 1998, **120**, 1300.
- [133] Go, N. and Scheraga, H.A., *J. Chem. Phys.*, 1969, **51**, 4751.
- [134] Karplus, M. and Kushick, J.N., *Macromolecules*, 1981, **14**, 325.
- [135] Witter, R., Prieß, W., and Sternberg, U., *J. Comp. Chem.*, 2002, **23**, 289.
- [136] H.M.Berman, Jeffrey, G.A., and Rosenstein, R.D., *Acta Crystallogr B*, 1968, **24**, 442.
- [137] Kaminsky, W. and Glazer, A.M., *Zeitschrift für Kristallographie*, 1997, **212**, 283.
- [138] Grindley, T.B., McKinnon, M.S., and Wasylshen, R.E., *Carbohydrate Research*, 1990, **197**, 41.
- [139] Burnett, M.H. and Johnson, C.K., *ORTEP-III*, OAK Ridge National Laboratory, Tennessee, 1996.
- [140] Vinogradov, S.N. and Linnell, R.H., *Hydrogen Bonding*, Van Nostrand Reinhold Company, New York, 1971.
- [141] Klug, A. and Rhodes, D., *Trends. Biochem. Sci.*, 1987, **12**, 464.
- [142] Krizek, B.A., Amann, B.T., and Kilfoil, V.J., *J. Am. Chem. Soc.*, 1991, **113**, 4518.
- [143] Marmorstein, R., Carey, M., Ptashne, M., and Harrison, S.C., *Nature*, 1992, **356**, 408.
- [144] Vallee, B.L., Coleman, J.E., and Auld, D.S., *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 999.
- [145] Luisis, B.F., Xu, W.X., Otwinowsky, Z., Freedman, L.P., Yamamoto, K.R., and Sigler, P.B., *Nature*, 1991, **352**, 497.
- [146] Magonet, E., Hayen, D., Delforge, D., Delaire, F., and Remacle, J., *Biochem. J.*, 1992, **287**, 361.

- [147] Stillman, M.J., Shaw, C.F., and Suzuki, K.T., *Methallothioneins*, VCH, Weinheim, 1992.
- [148] Zhang, X., Hubbard, C.D., and van Eldrick, R., *J. Phys. Chem.*, 1996, **100**, 9161.
- [149] Gockel, P., Gelinsky, M., Vogler, R., and Vahrenkamp, H., *Inorg. Chim. Acta.*, 1998, **272**, 115.
- [150] Herr, U., Spahl, W., Trojandt, G., Steglich, W., Thaler, F., and van Eldrick, R., *Bioorg. Med. Chem.*, 1999, **7**, 699.
- [151] Alsfasser, R., Ruf, M., Trofimenko, S., and Vahrenkamp, H., *Chem. Ber.*, 1993, **126**, 703.
- [152] Kimura, E. In ed. Karlin, K.D., *Progress in Inorganic Chemistry*, Vol. Vol. 41, p. 443, New York, 1994. Wiley.
- [153] van Eldrik, R., *Coord. Chem. Rev.*, 1999, **182**, 373.
- [154] Förster, M., Brasack, I., Duhme, A.K., Nolting, H.F., and Vahrenkamp, H., *Chem. Ber.*, 1996, **129**, 347.
- [155] Greiner, G., Seyfartz, L., Poppitz, W., Witter, R., Sternberg, U., and Reißmann, S., *Lett. Pept. Sci.*, 2000, **7**, 133.
- [156] Basosi, R., Gaggelli, E., Gaggelli, N., Pogni, R., and Valensin, G., *Inorg. Chim. Acta*, 1998, **275**, 274.
- [157] Magafa, V., Stavropoulos, G., and Tsiveriotis, P., *Inorg. Chim. Acta*, 1998, **272**, 7.
- [158] Tsiveriotis, P., Hadjiliadis, N., and Stavropoulos, G., *Inorg. Chim. Acta*, 1997, **261**, 83.
- [159] Williamson, M.P., *Natural Product Reports*, 1993, p. 207.
- [160] Alia, Matysik, J., Erkelens, C., Hulsbergen, F.B., Gast, P., Lugtenburg, J., and de Groot, H.J.M., *Chem. Phys. Lett.*, 2000, **330**, 325.
- [161] Mauksch, M., Bräuer, M., Weston, J., and Anders, E., *CHEMBIOCHEM*, 2001, **2**, 190.
- [162] Becke, A.D., *J. Chem. Phys.*, 1993, **98**, 5648.
- [163] Witter, R., Seyfart, L., Greiner, G., Reissmann, S., Weston, J., Anders, E., and Sternberg, U., *J. Biomol. NMR*, 2002, **24**, 277.

# Appendix

## Integrals and Integral Derivatives

Atomic Orbitals		
a)		$\left\langle \phi_{nlm} \left  \frac{1}{R-r} \right  \phi_{n'l'm'} \right\rangle$
$\langle ns  $	$ ns\rangle$	$\frac{1}{R}$
$\langle np_z  $	$ np_z\rangle$	$\frac{1}{R} + \frac{1}{R^3} \frac{(2n+1)(2n+2)}{5(2\zeta)^2} (3 \cos^2(\theta) - 1)$
$\langle ns  $	$ np_z\rangle$	$-\frac{2}{\sqrt{3}R^2} \frac{(2n+1)\sqrt{(2\zeta_1)^{2n+1}(2\zeta_2)^{2n+1}}}{(\zeta_1+\zeta_2)^{2n+2}} \cos(\theta)$
$\langle np_x  $	$ np_x\rangle$	$\frac{1}{R} + \frac{1}{R^3} \frac{(2n+1)(2n+2)}{2.5 \cdot (2\zeta)^2} [3(1 - \cos^2(\theta)) \cos(2\varphi) - (3 \cos^2(\theta) - 1)]$
$\langle np_y  $	$ np_y\rangle$	$\frac{1}{R} - \frac{1}{R^3} \frac{(2n+1)(2n+2)}{2.5 \cdot (2\zeta)^2} [3(1 - \cos^2(\theta)) \cos(2\varphi) + (3 \cos^2(\theta) - 1)]$
b)		$\frac{\partial}{\partial R} \left\langle \phi_{nlm} \left  \frac{1}{R-r} \right  \phi_{n'l'm'} \right\rangle$
$\langle ns  $	$ ns\rangle$	$-\frac{1}{R^2}$
$\langle np_z  $	$ np_z\rangle$	$-\frac{1}{R^2} - \frac{3}{R^4} \frac{(2n+1)(2n+2)}{5(2\zeta)^2} (3 \cos^2(\theta) - 1)$
$\langle ns  $	$ np_z\rangle$	$\frac{4}{\sqrt{3}R^3} \frac{(2n+1)\sqrt{(2\zeta_1)^{2n+1}(2\zeta_2)^{2n+1}}}{(\zeta_1+\zeta_2)^{2n+2}} \cos(\theta)$
$\langle np_x  $	$ np_x\rangle$	$-\frac{1}{R^2} - \frac{3}{R^4} \frac{(2n+1)(2n+2)}{2.5 \cdot (2\zeta)^2} [3(1 - \cos^2(\theta)) \cos(2\varphi) - (3 \cos^2(\theta) - 1)]$
$\langle np_y  $	$ np_y\rangle$	$-\frac{1}{R^2} + \frac{3}{R^4} \frac{(2n+1)(2n+2)}{2.5 \cdot (2\zeta)^2} [3(1 - \cos^2(\theta)) \cos(2\varphi) + (3 \cos^2(\theta) - 1)]$
c)		$\frac{\partial}{\partial \cos(\theta)} \left\langle \phi_{nlm} \left  \frac{1}{R-r} \right  \phi_{n'l'm'} \right\rangle$
$\langle ns  $	$ ns\rangle$	0
$\langle np_z  $	$ np_z\rangle$	$\frac{6}{R^3} \frac{(2n+1)(2n+2)}{5(2\zeta)^2} \cos(\theta)$
$\langle ns  $	$ np_z\rangle$	$-\frac{2}{\sqrt{3}R^2} \frac{(2n+1)\sqrt{(2\zeta_1)^{2n+1}(2\zeta_2)^{2n+1}}}{(\zeta_1+\zeta_2)^{2n+2}}$
$\langle np_x  $	$ np_x\rangle$	$-\frac{6}{R^3} \frac{(2n+1)(2n+2)}{2.5 \cdot (2\zeta)^2} [\cos(\theta) \cos(2\varphi) + \cos(\theta)]$
$\langle np_y  $	$ np_y\rangle$	$\frac{6}{R^3} \frac{(2n+1)(2n+2)}{2.5 \cdot (2\zeta)^2} [\cos(\theta) \cos(2\varphi) - \cos(\theta)]$
d)		$\frac{\partial}{\partial \varphi} \left\langle \phi_{nlm} \left  \frac{1}{R-r} \right  \phi_{n'l'm'} \right\rangle$
$\langle ns  $	$ ns\rangle$	0
$\langle np_z  $	$ np_z\rangle$	0
$\langle ns  $	$ np_z\rangle$	0
$\langle np_x  $	$ np_x\rangle$	$\frac{6}{R^3} \frac{(2n+1)(2n+2)}{2.5 \cdot (2\zeta)^2} [(1 - \cos^2(\theta)) \sin(2\varphi)]$
$\langle np_y  $	$ np_y\rangle$	$-\frac{6}{R^3} \frac{(2n+1)(2n+2)}{2.5 \cdot (2\zeta)^2} [(1 - \cos^2(\theta)) \sin(2\varphi)]$

Bond orbital matrix elements of the Fock operator (a), their derivatives with respect to  $R$  (b),  $\cos(\theta)$  (c) and  $\varphi$  (d).

**BPT Parameters of the Charge Parametrization**

<b>BPT parameters of MPA/3-21G</b>			
(q in e and P in e/H)			
q(C=O)	0.172112	q(H-N)	0.261307
P(C=O)	1.592837	P(H-N)	0.390599
P(C=C)	0.744594	q(C-F)	0.301326
q(C=N)	0.207822	P(C-F)	-0.06306
P(C=N)	2.147262	q(Cl-C)	0.232550
q(P=O)	0.192009	P(Cl-C)	0.011196
P(P=O)	0.158573	q(Si-H)	0.120357
q(S=O)	-0.29007	P(Si-H)	1.448965
P(S=O)	-0.19149	q(Si-C)	0.287954
q(S=C)	0.071907	P(Si-C)	0.894860
P(S=C)	0.990078	q(Si-O)	0.463323
q(P-O)	0.216148	P(Si-O)	-1.08260
P(P-O)	0.832157	q(Si-Cl)	0.257306
q(S-O)	0.759860	P(Si-Cl)	3.342545
P(S-O)	2.943327	q(S-H)	-0.10449
q(C-N)	0.145060	P(S-H)	-1.16316
P(C-N)	-0.77123	P(S-S)	2.035176
P(C-C)	0.362175	q(S-C)	0.264391
q(C-O)	0.257708	P(S-C)	-0.43828
P(C-O)	-1.38054	q(Zn-N)	0.244965
q(H-O)	0.336011	P(Zn-N)	4.333919
P(H-O)	0.269761	q(Zn-O)	0.397068
q(H-C)	0.198778	P(Zn-O)	-0.14277
P(H-C)	0.302874		

<b>BPT parameters of ESC/3-21G</b>			
(q in e and P in e/H)			
q(C=O)	0.190532	q(H-N)	0.197262
P(C=O)	2.408366	P(H-N)	3.139875
P(C=C)	3.130275	q(C-F)	0.188884
q(C=N)	-0.12967	P(C-F)	1.135402
P(C=N)	4.495816	q(Cl-C)	-0.00286
q(P=O)	0.552163	P(Cl-C)	1.707294
P(P=O)	0.111582	q(Si-H)	0.118685
q(S=O)	-1.20227	P(Si-H)	5.030602
P(S=O)	2.203628	q(Si-C)	0.096910
q(S=C)	-0.03686	P(Si-C)	2.701331
P(S=C)	1.792322	q(Si-O)	0.372915
q(P-O)	-0.15059	P(Si-O)	1.063902
P(P-O)	2.755085	q(Si-Cl)	0.212567
q(S-O)	1.112034	P(Si-Cl)	6.160564
P(S-O)	-0.11390	q(S-H)	-0.24591
q(C-N)	0.069866	P(S-H)	-2.06931
P(C-N)	2.832018	P(S-S)	7.401944
P(C-C)	2.526319	q(S-C)	-0.05875
q(C-O)	0.203412	P(S-C)	0.904944
P(C-O)	-0.64017	q(Zn-N)	0.113864
q(H-O)	0.403836	P(Zn-N)	4.483694
P(H-O)	0.324169	q(Zn-O)	0.697226
q(H-C)	0.101175	P(Zn-O)	2.234741
P(H-C)	1.899354		

<b>BPT parameters of NPA/6-31G(d,p)</b>			
(q in e and P in e/H)			
q(C=O)	0.247320	q(H-N)	0.381720
P(C=O)	1.489182	P(H-N)	0.273967
P(C=C)	0.599615	q(C-F)	0.397036
q(C=N)	0.137659	P(C-F)	-0.58294
P(C=N)	1.132873	q(Cl-C)	0.047769
q(P=O)	0.473276	P(Cl-C)	-0.00124
P(P=O)	0.145278	q(Si-H)	0.148906
q(S=O)	-0.98019	P(Si-H)	2.175241
P(S=O)	0.130090	q(Si-C)	0.384159
q(S=C)	-0.09903	P(Si-C)	1.081368
P(S=C)	1.290606	q(Si-O)	0.634816
q(P-O)	0.295536	P(Si-O)	-0.12469
P(P-O)	0.370479	q(Si-Cl)	0.348697
q(S-O)	1.494582	P(Si-Cl)	1.741996
P(S-O)	3.298970	q(S-H)	-0.16454
q(C-N)	0.120065	P(S-H)	-0.26373
P(C-N)	-0.97646	P(S-S)	3.210911
P(C-C)	0.538314	q(S-C)	0.159620
q(C-O)	0.296762	P(S-C)	-0.10032
P(C-O)	-1.39582	q(Zn-N)	1.045167
q(H-O)	0.476405	P(Zn-N)	-7.81040
P(H-O)	0.152492	q(Zn-O)	0.174562
q(H-C)	0.227809	P(Zn-O)	12.85987
P(H-C)	0.444273		

## Deutsche Zusammenfassung

Das Ziel der vorliegenden Dissertation ist die Weiterentwicklung und Anwendung der Bindungspolarisationstheorie (BPT) für molekulare Systeme. Man betrachtet ideale Bindungen in einer Umgebung aus Punktladungen, die die Ladungsverteilung im molekularen System repräsentieren. Diese Ladungen polarisieren die Bindungen zu mehr realistischen Bindungen. Mit diesem Konzept lassen sich Einelektronenerwartungswerte effizient berechnen. Ideale, konstante Bindungsanteile können parametrisiert werden, so dass eine Koordinatenabhängigkeit der Eigenschaften alleine auf Polarisierungen der Bindungen zurückgeführt werden kann. Dies macht den effizienten Einsatz in Verbindung mit einem Kraftfeld möglich. Als erstes wurde die Parametrisierung der BPT-Ladungsmethode anhand eines Satzes von 175 Eichmolekülen an Atomladungen von H, C, N, O, F, Si, P, S, Cl und Zn durchgeführt und optimiert. Dabei sind drei Ladungsmodelle (electrostatic charges: ESC, Mulliken population analysis: MPA, natural population analysis: NPA) unter Berücksichtigung von 11 verschiedenen Basissätzen in Anwendung gekommen. Erstmals wurden ESC- und NPA-Ladungen geeicht. Für jedes Modell wurde die beste Parametrisierung, Korrelationskoeffizient nahe 1, ermittelt. Die NPA-Eichung stellte sich bzgl. des theoretischen Ansatzes und der Parametrisierbarkeit als besonders geeignet heraus. Dies konnte aufgrund einer sehr guten Parameterstabilität bzgl. verschiedener kompakter Basissätze untermauert werden. Für die NPA-Ladungseichung wurde eine Korrelation zwischen *ab initio* und BPT Werten von  $R=0.996$  erreicht. Die absolute Standardabweichung beträgt 0.05 Einheiten der Elementarladung. Das Verhältnis der Standardabweichung zur Breite der absoluten Ladungsverteilung beträgt 9%.

Mittels des COSMOS-Kraftfeldes und unter Benutzung des neuen NPA Ladungsmodelles wurden mit Kristallmoleküldynamiksimulationen erstmals alle  $^{13}\text{C}$  chemischen Verschiebungstensorwerte von Seide II simuliert. Die erhaltenen gemittelten theoretischen Tensorwerte stimmen sehr gut mit experimentellen Daten überein. Die Korrelation beträgt 0.996 mit einer Standardabweichung von 6.8 ppm. Erstmals konnte auch ein komplettes 2D  $^{13}\text{C}$  NMR Iso-Aniso Spektrum simuliert werden.

In dieser Arbeit wurde erstmals die Bindungspolarisationstheorie für nicht lokale Einelektronenerwartungswerte formuliert. Desweiteren wurde eine allgemeine analytische Form für die Koordinatenableitung von tensoriellen stark polarisationsabhängigen Einelektronenerwartungswerten im Rahmen einer semiempirischen Theorie abgeleitet. Außerdem konnte im Rahmen der BPT eine molekulare Polarisationsenergie eingeführt werden, die Funktional der Einelek-



tronenerwartungswerte ist. Diesbezüglich wurden Energiekorrekturen und Kräfte abgeleitet, wobei von der Annahme gebrauch gemacht wurde, dass inkorrekte Strukturen Abweichungen von der Minimumpolarisationsenergie widerspiegeln und damit auch Abweichungen der Eielektronenerwartungswerte. Die abgeleiteten Kräfte bewirken innerhalb eines Kraftfeldes eine Anpassung des Strukturmodells an die experimentellen Daten. Dieses Konzept findet bei Geometrieoptimierungen, Moleküldynamik oder *simulated annealing* Rechnungen Anwendung, wobei die Pseudokräfte auf die 3D-Struktur von molekularen Systemen derart einwirken, dass theoretische und gemessene Erwartungswerte hinreichend genau in Übereinstimmung gebracht werden. In der vorliegenden Arbeit wurde dieses Verfahren mit der NMR chemischen Verschiebung realisiert. Für die isotropen Werte der  $^{13}\text{C}$  chemischen Verschiebung ist diese Methode implementiert worden, so dass es als Standardverfahren im COSMOS-Programm zur Strukturverfeinerung angewendet werden kann.

Mit Hilfe der BPT-Pseudokräfte wurden die Protonenpositionen der Röntgenstruktur von  $\beta$ -D-mannitol mittels  $^{13}\text{C}$  chemischen Verschiebungen verfeinert. Die mittlere Koordinatenabweichung der NMR Struktur von der Röntgenstruktur ( $0.13 \text{ \AA}$ ) ist kleiner als die Ungenauigkeit der Röntgenmessung ( $0.2 \text{ \AA}$ ). Außerdem wurde eine zusätzliche Wasserstoffbrücke gefunden. Die Methode der Pseudokräfte wurde desweiteren zur dreidimensionalen Strukturbestimmung eines Pseudotripeptidkomplexes des Zink ( $\text{Zn-Bz-His-}\Psi[\text{CO-N}(\text{CH}_2)_2\text{-NH}_2]\text{Gly-His-NH}_2$ ) in Flüssigkeit, einem Gemisch aus DMSO und  $\text{H}_2\text{O}$ , angewendet. Dafür wurden experimentell NOE-Abstände und  $^{13}\text{C}$  chemische Verschiebungen bestimmt. Eine Analyse der Ergebnisse mittels DFT (*density functional theory*) Verfahren ergab zwei verschiedene Strukturen: einen stabilen vierfach koordinierten Komplex und eine Struktur mit dreizähliger Symmetrie. Dies entspricht einer Übergangsstruktur zur Bindung von  $\text{H}_2\text{O}$ , einer Vorstufe zur Aktivierung von gelöstem  $\text{CO}_2$ . Die beschriebenen Geometrien weisen hohe Ähnlichkeit mit dem katalytischen Zentrum der Carboanhydrase, einem Enzym der Photosynthese zur Aktivierung von  $\text{CO}_2$ , auf. Schließlich konnte mit experimentellen  $^{13}\text{C}$  chemischen Verschiebungen die Kristallstruktur von Seide II (Takahashi Modell<sup>[124]</sup>) erstmalig verfeinert werden. Die energieniedrigste Struktur, mit den kleinsten Abweichungen der chemischen Verschiebungen vom Experiment, weist ein neues stabilisierendes Wasserstoffbrückensystem auf.

In dieser Dissertation konnten wesentliche Beiträge zur direkten dreidimensionalen Strukturbestimmung für Flüssigkeiten und Festkörper mittels der NMR chemischen Verschiebung erbracht werden.

## Thesen zur Dissertation

### ”Dreidimensionale Strukturaufklärung mit dem COSMOS-NMR Kraftfeld”

vorgelegt dem Rat der Physikalisch-Astronomischen Fakultät der  
Friedrich-Schiller-Universität Jena

von Diplom-Physiker Raiker Witter

- 1.) Die Bindungspolarisationstheorie läßt sich zur Berechnung von H-, C-, N-, O-, F-, Si-, P-, S-, Cl- und Zn-Atomladungen verschiedener Ladungsmodelle anwenden (*electrostatic charges, Mulliken population analysis, natural population analysis*).
- 2.) Für das Ladungsmodell *natural population analysis* lassen sich die Parameter der Bindungspolarisationstheorie am besten bestimmen. Die Korrelation zu *ab initio* Rechnungen beträgt 0.996. Eine Parameterstabilität bezüglich kompakter Basissätze kann gezeigt werden.
- 3.) Mittels Kristallmoleküldynamiksimulation kann man erstmals alle  $^{13}\text{C}$  chemischen Verschiebungstensorwerte von Seide II berechnen. Die gemittelten theoretischen Tensorwerte stimmen sehr gut mit experimentellen Daten überein. Der Korrelationskoeffizient beträgt 0.996 mit einer Standardabweichung von 6.8 ppm. Erstmals kann auch ein komplettes 2D  $^{13}\text{C}$  NMR Iso-Aniso Spektrum von Seide simuliert werden.
- 4.) Es ist möglich die Bindungspolarisationstheorie für nicht lokale Einelektronenerwartungswerte zu formulieren.
- 5.) Man kann eine allgemeine analytische Form für die Koordinatenableitungen von tensoriellen stark polarisationsabhängigen Einelektronenerwartungswerten im Rahmen der Bindungspolarisationstheorie angeben.
- 6.) Im Rahmen der BPT kann eine molekulare Polarisationsenergie eingeführt werden, die Funktional der Einelektronenerwartungswerte ist. Diesbezüglich können erstmals Energiekorrekturen und Kräfte abgeleitet werden, wobei von der Annahme gebrauch gemacht wird, dass inkorrekte Strukturen Abweichungen von der Minimumpolarisationsenergie widerspiegeln

und damit auch Abweichungen der Einelektronenerwartungswerte. Die abgeleiteten Kräfte bewirken innerhalb eines Kraftfeldes eine Anpassung des Strukturmodelles an die experimentellen Daten.

7.) Das Verfahren der Pseudokräfte ist mit der NMR chemischen Verschiebung umgesetzt und für die isotropen Werte der  $^{13}\text{C}$  chemischen Verschiebung im COSMOS-Programm implementiert, so dass es als Standardverfahren zur Strukturverfeinerung angewendet werden kann.

8.) Die Protonenpositionen der Röntgenstruktur von  $\beta$ -D-mannitol können mit  $^{13}\text{C}$  chemischen Verschiebungen verfeinert werden. Aufgrund der neuen Optimierungsmethode wird ein alternatives, an die experimentellen Daten besser angepasstes Wasserstoffbrückensystem, vorgeschlagen.

9.) Mit Hilfe experimentell ermittelter NOE Abständen und  $^{13}\text{C}$  chemischen Verschiebungen kann man die dreidimensionale Struktur von  $\text{Zn-Bz-His-}\Psi[\text{CO-N}(\text{CH}_2)_2\text{-NH}_2]\text{Gly-His-NH}_2$  in Flüssigkeit ermitteln. Die Analyse der Ergebnisse unter Verwendung von DFT (*density functional theory*) Verfahren ergibt einen stabilen vierfach koordinierten Komplex und eine Struktur mit dreizähliger Symmetrie. Diese Geometrien weisen hohe Ähnlichkeit mit dem katalytischen Zentrum der Carboanhydrase, einem Enzym der Photosynthese zur Aktivierung von  $\text{CO}_2$ , auf.

10.) Erstmals kann mit  $^{13}\text{C}$  chemischen Verschiebungen die Kristallstruktur von Seide II verfeinert werden. Die energieniedrigste Struktur, mit den kleinsten Abweichungen der chemischen Verschiebungen vom Experiment, weist auf ein alternatives Wasserstoffbrückensystem hin.

## **Danksagung**

Herzlich möchte ich mich bei meinem Doktorvater PD. Dr. Ulrich Sternberg für die fruchtbare Zusammenarbeit danken, und auch für seine Unterstützung in privat schwierigen Zeiten. Vielen Dank an Prof. Dr. Christian Jäger in dessen Gruppe ich arbeiten durfte. Ebenso ganzheitlich, auf wissenschaftlicher und privater Basis, möchte ich allen meinen Arbeitskollegen danken: Dagmar Bechstein, Katrin Redlich, Stephanie Hesse, Michaela Zeyer, Frank-Thomas Koch, PD. Dr. Peter Hartmann, Wolfram Prieß, Dr. Ole Hirsch, Dr. Wolfgang Heerdegen, Dr. Marcus Schulz und Frank Perner. Bedanken möchte ich mich für die erfolgreiche Zusammenarbeit innerhalb des SFB 436 "Metallvermittelte Reaktionen nach dem Vorbild der Natur" mit Prof. Dr. Ernst Anders, Prof. Dr. Siegmund Reißmann, Prof. Dr. Udo Gräfe, Dr. Georg Greiner, Dr. Lydia Seyfarth, Dr. Jenny Weston, Dr. Ute Böttger und Heike Heinecke.

Mein Dank geht an Dr. Margit Möllhoff für die Zusammenarbeit und ihre scharfen linguistischen Blicke.

Ich möchte mich bei Prof. Dr. Stefan Spange, Prof. Dr. Dieter Michel, Prof. Dr. Dieter Freude, Dr. Hardy Müller, Dr. Andre´ Pampel und Andreas Seifert für ihre Hilfe danken, dass ich neben meiner mehr theoretische orientierten Promotionsarbeit auch ein Reihe von experimentellen Erfolgen realisieren konnte.

Vielen Dank an Meister Chu Tan Cuong.

Bedanken möchte ich mich bei Prof. Dr. Richard Ernst, der durch sein öffentliches Auftreten meine ethisch und moralischen Grundsetze bestärkte.

Ich danke meinen liebsten Eltern und Steffi.

## Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, daß ich die vorliegende Arbeit selbständig, ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehenden aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich geholfen:

1. Dr. Margit Möllhoff und Frank-Thomas Koch: Hilfe bei der Auswahl des Satzes der Eichmoleküle für die Ladungsparametrisierung.
2. Prof. Marc Henry: Anwendung seiner PACHA-Ladungsberechnungsmethode an einem Pseudopeptidzinkkomplex.
3. Dr. Georg Greiner, Dr. Lydia Seyfart und Prof. Dr. Sigmund Reißmann: Synthese der in der Arbeit verwendeten Pseudopeptide und des Pseudopeptidzinkkomplexes.
4. Heike Heinecke und Prof. Dr. Udo Gräfe: Bereitstellung der Flüssigkeitsspektrometer.
5. Dr. Jennie Weston: Berechnung der chemischen Verschiebung des Pseudopeptidzinkkomplexes mit dem Basissatz 6-311+G(d,p)
6. Wolfram Priess: Bereitstellung der aktuellsten Parameter für die Berechnung der Tensoren der chemischen Verschiebung mittels der BPT-Methode.
7. PD. Dr. Ulrich Sternberg, Dr. Peter Losso und Frank-Thomas Koch: Programmierung verschiedener Funktionalitäten in der COSMOS-Software.
8. Sabine Irmer: Instandhaltung und Bereitstellung der Parallelrechner und des GAUSSIAN98-Programmes.

Weitere Personen waren an der inhaltlichen-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die geltende Promotionsordnung der Physikalisch-Astronomischen Fakultät ist mir bekannt.

Ich versichere ehrenwörtlich, daß ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Ort, Datum

Unterschrift

## Lebenslauf

**Name:**

Raiker Witter

**Geboren:**

5. November 1971 in Gotha

**Schulbildung:**

1978-1980: Polytechnische Oberschule "Theodor Neubauer", Gotha.

1980-1988: Polytechnische Oberschule "Anna Seghers", Gotha.

1988-1990: Erweiterte Oberschule "Arnoldi", Gotha.

**Wehrdienst:**

1990-1991: NVA der DDR und Bundeswehr der BRD.

**Studium:**

1991-1993: Friedrich-Schiller-Universität Jena (Physik Vordiplom).

1993-1994: Lancaster University, UK (Physik, Geschichte der Wissenschaft, Logik).

1994-1997: Friedrich-Schiller-Universität Jena (Physik Diplom).

1997 Diplomarbeit: "<sup>31</sup>P NMR Doppelquanten Untersuchungen an Phosphatgläsern".

**Tätigkeit:**

Seit 1997: wissenschaftlicher Angestellter an der Friedrich-Schiller-Universität Jena (SFB 436)

Seit 1998: Promotionsstudium an der Friedrich-Schiller-Universität Jena

Seit 1999: Nebentätigkeit in eigener Firma.

Ort, Datum

Unterschrift