Journal of Biomolecular NMR, **24:** 277–289, 2002.
KLUWER/ESCOM

Structure determination of a pseudotripeptide zinc complex with the COSMOS-NMR force field and DFT methods

Raiker Witter[a,*], Lydia Seyfart[b], Georg Greiner[b], Siegmund Reissmann[b], Jennie Weston[c], Ernst Anders[c] & Ulrich Sternberg[a]

[a]*PAF / IOQ / HF, Friedrich-Schiller-Universität Jena, Max Wien Platz 1, 07743 Jena, Germany;* [b]*Institute of Biochemistry and Biophysics, Friedrich-Schiller-Universität Jena, Philosophenweg 12, D-07743 Jena, Germany;* [c]*Institute of Organic Chemistry and Macromolecular Chemistry, Friedrich-Schiller-Universität Jena, Humboldtstr. 10, D-07743 Jena, Germany*

Received 16 May 2002; Accepted 3 October 2002

Key words: carboanhydrase, carbonic anhydrase, geometry optimization, molecular dynamics, peptide, pseudo forces, restrictions, simulated annealing, structure determination, zinc complex

Abstract

A His-X-His pseudotripeptide zinc complex (X is a N-alkyl glycine derivative) similar to the catalytic center of the carbonic anhydrase was computer designed and experimentally synthesized. Using 2D-NMR techniques, all proton, carbon chemical shifts and nuclear overhauser effect signals were assigned. The three-dimensional structure of the complex was determined with the COSMOS (computer simulation of molecular structures) force field by applying ^{13}C bond polarization theory chemical shift pseudo forces and restrictions for NOE distances. From molecular dynamics, simulated annealing simulations and geometry optimizations, the three best force field structures were taken for a final investigation by density functional theory calculations.

Introduction

The ligating units for zinc complexation in biological enzymes such as zinc fingers (Klug and Rhodes, 1987; Krizek et al., 1991), zinc twists, zinc clusters (Marmorstein, 1992), alcohol dehydrogenase (Magonet et al., 1992), metallothioneins (Stillman et al., 1992) and carboanhydrase (Zhang et al., 1996) are bis(cysteinyl) or bis(histidinyl). These 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_2O or CO_2. Tripeptides with bis(histidinyl) sequences have been investigated by Gockel (1998). Other authors have described tripodal histidine ligands (Herr et al., 1999), pyrazolylborate ligands (Alsfasser et al., 1993) and macrocyclic polyamines (Kimura, 1994; van Eldrik, 1999). In this article, we have attempted to mimic the catalytic center of the carboanhydrase by a His-X-His pseudotripeptide. A N-alkyl glycine derivative was used for the residue X. Figure 1 shows the base structure of this peptide.

For consistency, we investigated three compounds: the base ligand Bz-His-Gly-His-NH_2 itself, the pseudotripeptide Bz-His-Ψ[CO-N$(CH_2)_2$-NH_2]Gly-His-NH_2 and the zinc complex of the latter. The base peptide of the His-Gly-His form and its dimeric complex has been previously investigated by Förster et al. (1996). We introduced a N-alkyl chain in order to ensure monomeric complexation. Furthermore, the N-functionalized glycine residue allows a cis-trans isomerisation and is less flexible than the C_α-substituents. This pseudotripeptide zinc complex was designed to bind H_2O in aqueous solution. It turned out that the complex did not dissolve very well in pure water so we analyzed it in a DMSO/H_2O mixture.

*To whom correspondence should be addressed. E-mail: witter@ioq.uni-jena.de

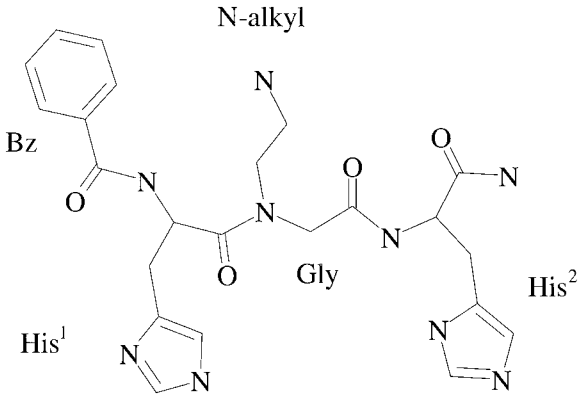

Figure 1. Base structure of the peptide ligand.

These peptides were synthesized by solid state phase methods and purified with HPLC. Compared to the zinc complex, the free ligand is very well soluble in H_2O. 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. We therefore used solution NMR techniques for structure determination. NMR methods of generally are gaining importance (Tsiveriotis et al., 1997; Basosi et al., 1998; Magafa et al., 1998) in investigations on the metal-ligand-interactions. Modern methods use multidimensional NMR experiments combined with molecular mechanics simulations for structure determination (Williamson, 1993). We obtained distance constrains from NOE intensities. The ^{13}C chemical shifts provided additional conformation information. We applied our newly introduced method for direct ^{13}C chemical shift refinement (Witter et al., 2002) in connection with the COSMOS force field (Möllhoff and Sternberg, 2001). The energetically most stable force field structures were finally optimized with quantum chemical procedures and the chemical shifts were theoretically determined.

Experimental

Modeling

An initial guess for the structure of the Zn-Bz-His-Ψ[CO-N$(CH_2)_2$-NH_2]Gly-His-NH_2 complex was designed with the COSMOS program (Koch et al., publ. www.cosmos-software.de), see Figure 2. The COSMOS force field (Möllhoff and Sternberg, 2001) with coordinate dependent charges was applied. We assumed that the zinc ion coordinates with the imidazole rings. It is known that imidazole ligands in complexes are deprotonated (Alia et al., 2000). They were therefore treated as negative charged groups. In order to account for the first reaction step of the carbonic anhydrase (Mauksch et al., 2001) a H_2O molecule was also bound to the Zn^{2+} cation. The N-alkyl chain was then used to complete the most probable tetrahedral coordination sphere on zinc.

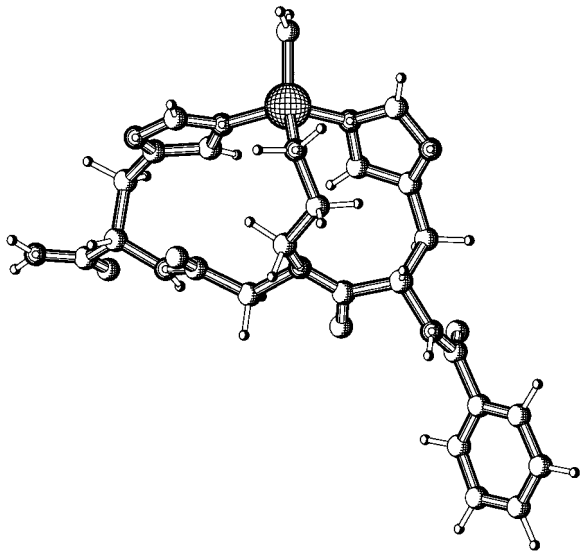

Figure 2. Computer designed guess for zinc complexation of the pseudotripeptide.

Synthesis

The syntheses of the peptides in consideration (Bz-His-Gly-His-NH_2 and Bz-His-Ψ[CO-N$(CH_2)_2$-NH_2]Gly-His-NH_2) 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). The complexation was carried out with $Zn(ClO_4)_2 \times 6\ H_2O$. MALDI mass (Voyager-DE^{TM} 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 could be experimentally shown. For more details see Greiner et al. (2000).

NMR

For the full ^1H, ^{13}C and NOE assignment we utilized the following solution NMR experiments: 1D ^1H, 1D ^{13}C, 1D ^{13}C Dept, 2D-COSY, 2D-TOCSY,

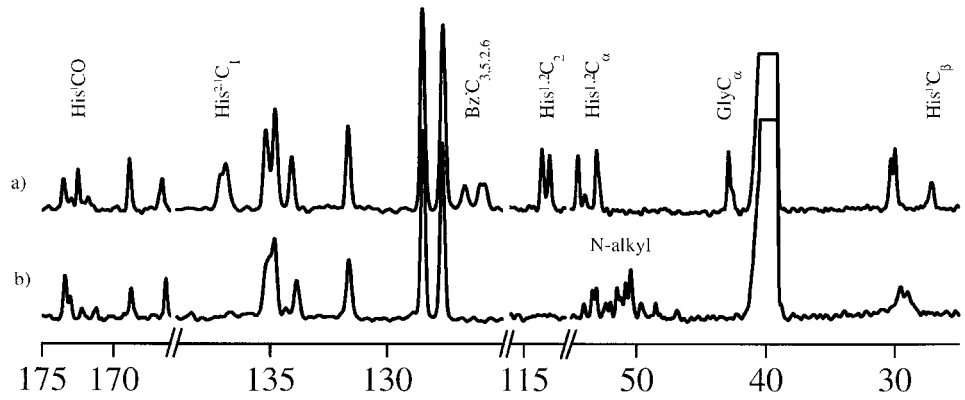

Figure 3. The ^{13}C spectra of a) the base ligand Bz-His-Gly-His-NH_2 and b) the pseudotripeptide Bz-His-Ψ[CO-N$(CH_2)_2$-NH_2]Gly-His-NH_2.

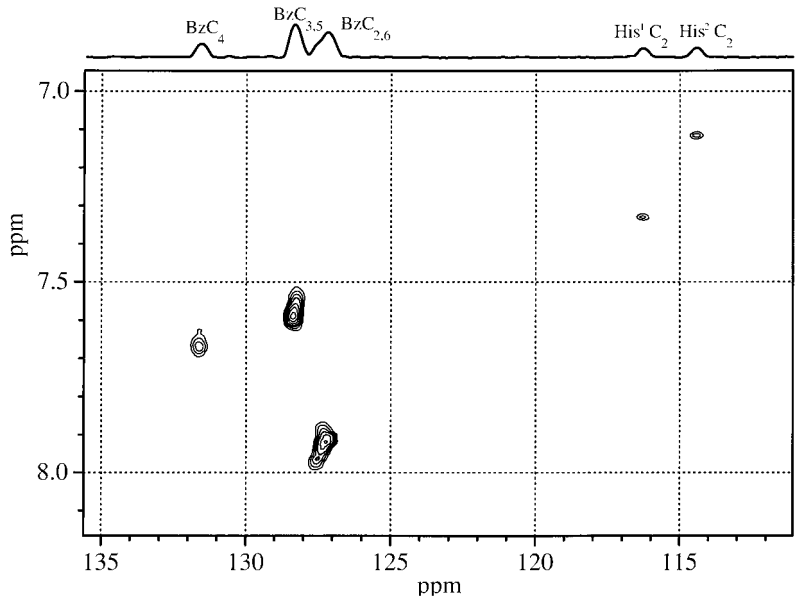

Figure 4. The aromatic resonances of the ^{1}H-^{13}C correlation (HSQC) spectrum of the pseudotripeptide complex.

2D-NOESY, 2D-ROESY, 2D-HMBC and 2D-HSQC (Braun et al., 1998). These measurements were carried out on a 350 MHz and a 500 MHz Bruker spectrometer.

The ^{13}C spectra of the base ligand and the pseudotripeptide are illustrated in Figure 3. The resonances of His^1CO, His^2C_1, His^2C_1, His^1C_2 and His^2C_2 are quite well-defined for the base ligand but rather broad for the pseudotripeptide. Due to this line broadening it is rather difficult to extract data from 2D-spectra. The reason for this is the high flexibility of the peptide on the μs time scale. Additional, there are at least two stable conformations for the base ligand which can be concluded from the appearance of second resonances for $Bz'C_2$, $Bz'C_3$, $Bz'C_5$, $Bz'C_6$, $His^{1'}C_\beta$ and $His^{1'}CO$.

The $GlyC_\alpha$ resonance is shifted downfield for the pseudopeptide duo to the additional N-alkyl chain. In Figure 3b) it is obvious that the bulk of CH_2 resonances contains conformations of $GlyC_\alpha$, His^1C_α, His^2C_α, $GlyNC_1$ and $GlyNC_2$ about 50 ppm.

In contrast to the two free ligands, the ^{13}C spectra of the complex (Figures 4 and 5) show broader lines. Two extra ^{13}C shieldings can be assigned to the $GlyNC_2$ functionality and an additional one to the BzCO group. However, the intensities of these signals in the 1D-spectrum are rather low which indicates a narrow conformational space. Surprisingly,

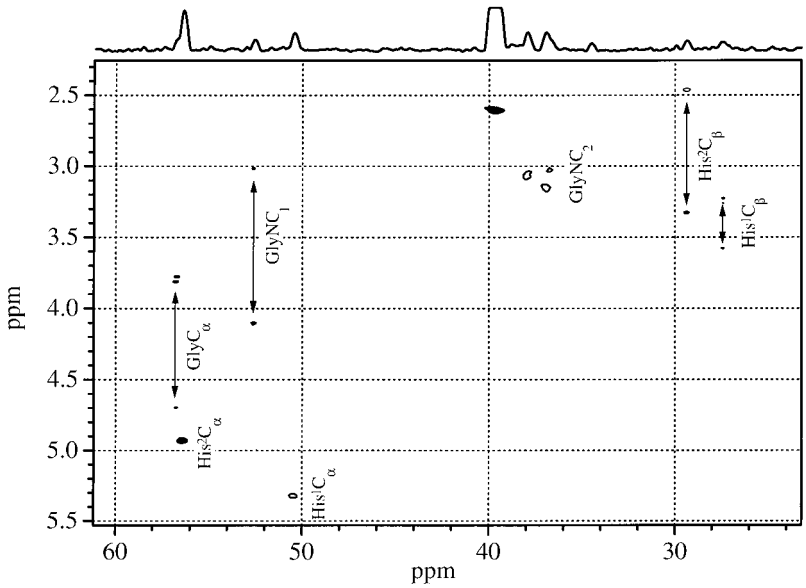

Figure 5. The ^{13}C-^{1}H correlation (HSQC) spectrum section of the aliphatic ^{13}C chemical shifts of the zinc complex.

Table 1. Experimental and theoretical ^{13}C NMR chemical shift data of the base ligand, the ligand and the zinc complex. The residue X is the N-alkyl glycine derivative Ψ[CO-N(CH_2)$_2$-NH_2]Gly

C atom	Bz-His-Gly-His^{Exp}	Bz-His-X-His^{Exp}	Zn-Bz-His-X-His^{Exp}	Zn-Bz-His-X-His^{Theo1}	Zn-Bz-His-X-His^{Theo2}
BzC_1	134.1	134.9	134.0	124.0	135.5
BzC_2/BzC_6	127.6 (125.8)	127.6	127.5	118.0	129.8
BzC_3/BzC_5	128.5 (126.7)	128.5	128.5	118.4	130.0
BzC_4	131.7	131.6	131.8	121.6	133.8
BzCO	166.6	166.3	165.9 (166.9)	154.6	167.8
His^{1}C$_1$	136.9	135.1	133.6	131.1	145.2
His^{1}C$_2$	113.6	~113	116.4	117.8	128.0
His^{1}C$_4$	136.9	135.1	133.5	135.9	149.5
His^{1}C$_\beta$	29.9 (27.1)	29.0	27.7	28.7	32.1
His^{1}C$_\alpha$	54.5 (54.0)	50.4	50.8	56.1	58.5
His^{1}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$_\alpha$	42.9	50.8	56.6	48.9	52.1
GlyNC$_1$	–	51.5 (48.3)	52.8	41.7	43.8
GlyNC$_2$	–	39.7 (38.8, 38.6)	37.1 (38.0)	38.9	42.2
His^{2}C$_1$	137.1	135.1	135.6	130.0	142.7
His^{2}C$_2$	113.1	~113	114.2	114.5	122.8
His^{2}C$_4$	134.8	133.9	135.7	134.1	146.2
His^{2}C$_\beta$	30.3	29.5	29.1	28.1	28.9
His^{2}C$_\alpha$	53.1	53.4 (53.1)	56.3	59.3	64.8
His^{2}CO	173.5 (172.5)	172.2 (171.2)	–	155.8	171.1

[1]SCF-GIAO B3LYP/6-31G(d,p) calculation.
[2]SCF-GIAO B3LYP/6-311+G(d,p) calculation.

the His^2CO functionality could not be assigned, a fact possibly due to fast rotation of the -$CONH_2$ group.

The ^{13}C NMR data is given in Table 1 and yield insights into the type of complexation of the pseudopeptide in DMSO. We found carbon downfield chemical shift changes at certain positions which are possibly due to the vicinity of the positive charge. The zinc ion has a large influence upon the positions of the His^1C_2, GlyCO, $GlyC_\alpha$, His^2C_4 and His^2C_α shifts. Otherwise the most important NOE connectivity was found between one proton from -NH_2 of the N-alkyl chain and His^1C_4H. This corresponds to the 3 complexes determined by the COSMOS-force field refinement illustrated in Figure 6.

Theory

NMR parameters contain the coordinate information of the molecular structure averaged on the μs time scale. Most notable in the liquid state are the NOE signals, J couplings and the chemical shifts. For our dynamic structure investigation we needed coordinate dependent theoretical models of the NMR parameters applicable for metalopeptides and metaloproteins. ^1H NOE signals depend largely on the cross relaxation between two protons. In contrast to the free ligand, the complex is rather rigid and the isotropical overall motion can be separated from the internal motions. The model free approach of Lipari & Szabo (1982) leads, in a first approximation, to a functional behavior of roughly the inverse square of the third power of the effective NOE distance (Spoel, 1996). For the J couplings, the Karplus equation (Karplus, 1959; Vuister et al., 1992) and similar relations (Spoel, 1996) are well established models for the backbone or side chain dihedral angles dependencies of proteins and peptides. It is however, unclear if these methods suffice for a complete description of pseudopeptide complexes. Thus we omitted these parameters in our considerations. The complex coordinate dependent ^{13}C chemical shifts were calculated using the BPT approach (Sternberg and Priess, 1997; Press and Sternberg, 2001) The main idea of this theory is that in first approximation any expectation value of an on-electron operator $\hat{O}$ can be expressed by

$$\left\langle \Psi_0 \left| \hat{O} \right| \Psi_0 \right\rangle = \left\langle \Phi_0 \left| \hat{O} \right| \Phi_0 \right\rangle + \sum_{i \neq 0} \left(\frac{\left\langle \Phi \binom{i^*}{i} \left| \hat{O} \right| \Phi_0 \right\rangle}{E_0 - E_i} \left\langle \Phi_0 \left| \hat{H} \right| \Phi \binom{i^*}{i} \right\rangle \right), \quad (1)$$

where the molecular ground state wave function $|\Phi_0\rangle$ of the energy E_0 is chosen to be a Slater determinant constructed from ideal bond orbitals. Because this formula is derived from a configuration interaction perturbation series of the total molecular wave function $|\Psi_0\rangle$, the excited configurations $|\Phi_i\rangle$ of the ground state wave function, with energies E_i, have to be taken into account. They main effect are polarizations $\left|\Phi \binom{i^*}{i}\right\rangle$, and we neglect electron de-localizations. From this, the basic idea of the BPT can be easily understood. It describes the polarization effect of the charge distributions on ideal bonds. The molecular system in consideration are thought to consist of simple bonds between connected neighbors. We describe these with localized two centred bond orbitals. They are linear combinations of appropriate hybrids. The polarity parameter d is the only free parameter for the construction of these bonds $|i\rangle$ and their anti-bonds $|i^*\rangle$. Bonds and anti-bonds are assumed to be orthogonal (zero overlap approximation (Del, 1958). The hybrids are built out of Slater atomic orbitals (Slater, 1930) which exponentials are taken from Burns (1964). Furthermore, we introduce the common identification of the Hamiltonian $\hat{H}$ with the Fock operator $\hat{F}$. Since the chemical shielding can be expressed as a sum of one-electron expectation values, so it is with the chemical shift, and it can be treated with the BPT approach (Equation 1). In this case, the Fock operator $\hat{F}$ can be thought to be decomposed into two subsystems A and B. Contribution $\hat{F}_A$ is designated to the bond in consideration and part $\hat{F}_B$ to the rest of the system. This splitting can be performed because delocalizations from system A to B are neglected. Using the ideas of the PCILO method, Malrieu et al. (1977), it can be assumed that the bond wave functions are adjusted in that way, that only subsystem $\hat{F}_B$ polarizes the bond contributions of subsystem A. Thus, it turns out that the chemical shift depends on the bond polarization energy matrix elements $F_i = \left\langle i \left| \hat{F}_B \right| i^* \right\rangle$. The shift operator acts on all electrons, hence the sum of Equation (1) runs over all bonds of subsystems A. It is known that the shift roughly depends on the inverse distance of the electrons from the nucleus N. Therefore, in first approximation, only bonds

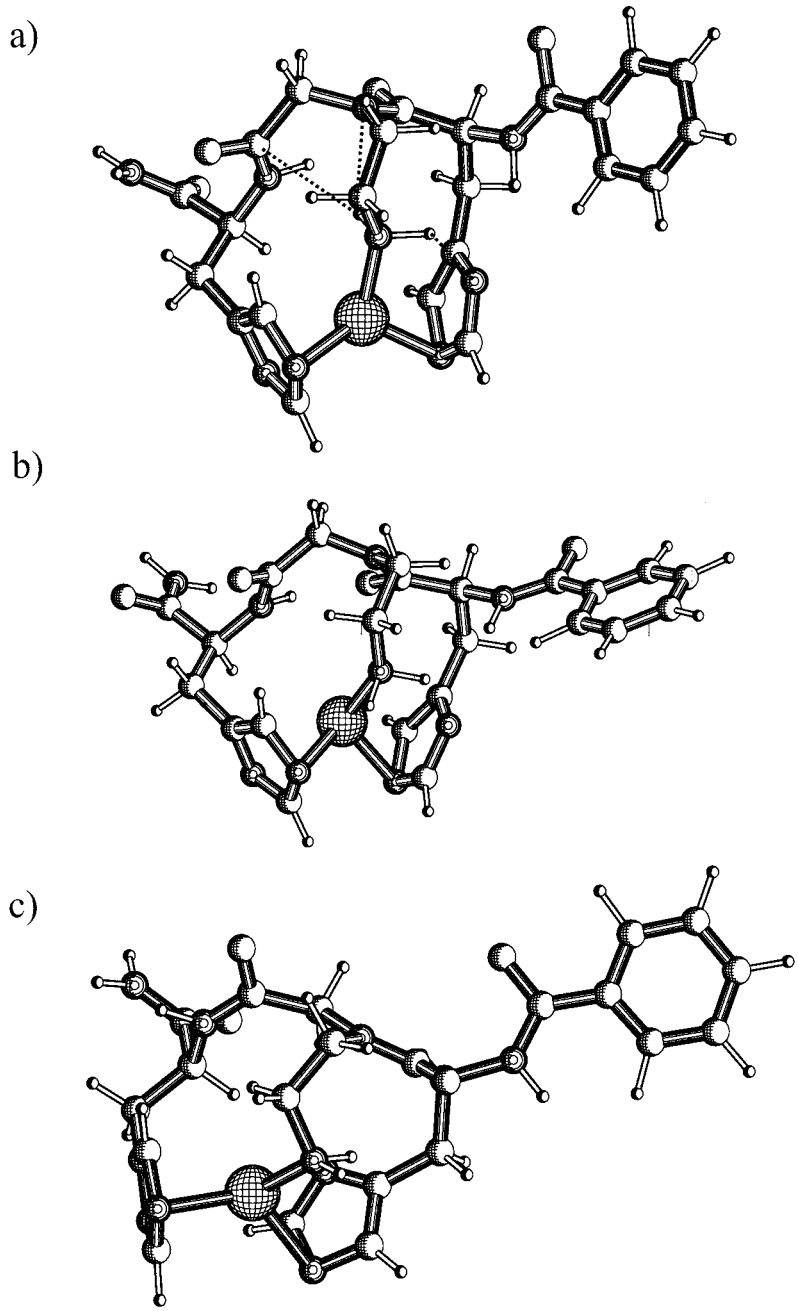

Figure 6. The three best COSMOS-NMR force field results with NOE as well as chemical shift pseudo forces. X-Zn-Y bonds are shown if the length is smaller than 2.5 Å.

directly connected to the nucleus N are taken into account. Additionally, we approximate the Fock operator $\hat{F}_B$ by a point charge distribution $\hat{V}_B$ of subsystem B ($F_i \approx V_i$). These are the net atomic charges located at the nuclei positions. And, we introduce an electron occupation number for conjugated π-bonds. They are calculated with the empirical valence formula (Malrieu, 1977) $n(r) = 2\exp((R-r)/0.37) - 2$. This occupation number is a function of the difference of the equilibrium single bond length R and the actual

bond length r. Using all these assumptions we end up with the BPT chemical shift formula (2)

$$\delta^T = \left\langle \Psi_0 \left| \hat{\delta}_N \right| \Psi_0 \right\rangle = \sum_{i}^{i\in N} \delta_i^{CS} n_i + A_i^{CS} n_i^2 \left\langle i \left| \hat{V}_B \right| i^* \right\rangle. \quad (2)$$

It is a sum over all bond contributions belonging to nucleus N. The functional behavior depends on the location of all atomic charges which polarize these bond contributions and it depends on the double bond length. The parameter $\delta_i^{CS} = 2\left\langle i \left| \hat{\delta} \right| i \right\rangle$ describes the ground state bond contribution, while $A_i^{CS} = 4\left\langle i \left| \hat{\delta} \right| i^* \right\rangle / \Delta E_i$ is the polarization parameter of bond contribution i. These bond polarization parameters δ_i^{CS} and A_i^{CS} are once obtained by calibrating on a collection of single crystal measurements (Veeman, 1984; Sherwood et al., 1989) and to some extend also to *ab initio* results (Sternberg and Priess, 1997; Priess and Sternberg, 2001). The integrals and their derivatives of the bond polarization energy matrix elements V_j are given in Witter et al. (2002). Now, only the atomic charges have to be known. It turns out, that the formula for the BPT atomic charges is similar to the chemical shift equation and has the form (Koch et al., 1994, 2001)

$$q^N = \sum_{j}^{N} q_j^Q n_j + A_j^Q n_j^2 V_j. \quad (3)$$

The parametrization is done with *ab initio* results. Because the bond polarization energy matrix elements depend linearly on the charges, calculation of the atomic charges means solving a set of linear equations. In addition, the Coulomb energy uses these coordinate dependent charges. Only the correct interplay of electrostatic forces, pseudo forces, VdW forces and valence forces ensures a successful structure determination.

The 3D structure elucidation was carried out with high temperature molecular dynamics followed by a simulated annealing procedures with NOE distance conditions. More precise structures were then finally determined by geometry optimization with additional chemical shift restrictions. We utilized our COSMOS force field (Mollhoff and Sternberg, 2001) since the computational effort required for molecular dynamics, simulated annealing and geometry optimizations with *ab initio* method (Grotendorst, 2000) is still too large. Within the COSMOS force field approach, the coordinate dependent charge distribution (Koch et al., 1994, 2001) is calculated using the semiempirical bond polarization theory (BPT) (Sternberg, 1988). In Koch et al. (2001) the COSMOS-force field was parametrized for calculations on organic zinc complexes. Since in this work we used the non-bonded approach for the zinc, it was treated as a Zn^{2+} ion and only its Van der Waals (VdW) radius had to be adjusted in respect to the BPT charge calculation. To the COSMOS force field suitable NMR pseudo forces were added.

We first considered the ^{1}H NOE restrictions. They correspond to intermolecular interactions of proton pairs rather than valence forces. We therefore treated them as harmonic perturbation of the electrostatic and VdW proton pair energy (Equation 4)

$$E^{el} + E^{VdW} = \left(E^{el} + E^{VdW}\right)\big|_{NOE} \left(1 + S\frac{\left(R^T - R^{NOE}\right)^2}{2\Delta R^2}\right). \quad (4)$$

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. ΔR 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 (Equation 5)

$$F_\alpha^{NOE} = \left|E^{el} + E^{VdW}\right|_{NOE} \frac{\left(R^T - R^{NOE}\right)}{\Delta R^2}\frac{\partial R}{\partial x_\alpha}. \quad (5)$$

The VdW energy at the NOE distance can be calculated exactly. The hydrogen atomic charges for the Coulomb energy don't vary much with conformational changes and can be approximated with the starting values. Therefore the energy minimum constant $\left|E^{el} + E^{VdW}\right|_{NOE}$ can be determined from the start.

We then introduced analytically derived chemical shift pseudo forces. The fundamental idea for the pseudo forces is that, in the bond polarization limit, the sum of the polarization energies of all the bond contributions belonging to an atom – the total atomic polarization energy E_P^N – is a functional of the corresponding chemical shift δ (Equation 6)

$$E_P^N = \sum_{i\neq 0}^{N} \left(\frac{\left\langle \Phi_0 \left| \hat{V}_B \right| \left(\Phi_i^{i*}\right) \right\rangle^2}{E_0 - E_i}\right). \quad (6)$$

Slight variations in the chemical shift cause perturbations in the polarization. In a first approximation, we derived the pseudo energy from the first order expansion with respect to the chemical shift (Witter et al., 2002). However, this limited the forces to polarization effects only. In order to account for strong π bond length dependencies of sp^2 coordination we had to add the second order energy term (Equation 7)

$$E_P^N[\delta] = E_P^N[\delta_0] + \left.\frac{\partial E_P^N}{\partial \delta}\right|_0 \Delta\delta + \left.\frac{\partial^2 E_P^N}{\partial \delta^2}\right|_0 \frac{(\Delta\delta)^2}{2} + o\left(\Delta\delta^3\right). \quad (7)$$

The constant is identified as the experimental value δ^E and the variable δ is the theoretical chemical shift δ^T. If we assume that we are near the minimum atomic polarization energy in the first approximation, the following expression (Equation 8) for the chemical shift pseudo force can be derived:

$$F_\alpha^{CS} \approx k_2 \left(\delta^T - \delta^E\right) \frac{\partial \delta^T}{\partial x_\alpha}. \quad (8)$$

For large differences in the chemical shift $\left(\delta^T - \delta^E\right)$, a modified scaling function is introduced (see Witter et al., 2002). The complexity and efficiency of the chemical shift pseudo force calculation depends on the expression of the chemical shift formula. From Equations (2) and (8) it might seem that the most CPU time is used for calculating the chemical shifts or its derivatives, but the determination of the charge distribution is actually the most time consuming component (Equation 3). With the knowledge of the atomic polarization energy (7), the chemical shift formula (2) as well as the charge equations (3) we are able to approximate the first and second derivative of the atomic polarization energy with respect to the chemical shift

$$k_1 = -\left.\frac{\partial E_P^N}{\partial \delta}\right|_0 = \sum_j^N \frac{\left|A_j^Q\right|}{e\, A_j^{CS}} V_j|_0$$
$$k_2 = -\left.\frac{\partial^2 E_P^N}{\partial \delta^2}\right|_0 = \sum_j^N \frac{\left|A_j^Q\right|}{e\, n_j^2 \left(A_j^{CS}\right)^2}. \quad (9)$$

Although we are able to calculate both derivatives, for simplicity we focus on the force constant k_2 and see, that it can be calculated from the bond polarization parameters of the charge and chemical shift of atom N (A_i^Q, A_i^{CS}). In SI units, the elementary charge e is used and the bond electron occupation numbers n_k have to be given. Our BPT chemical shift pseudo force takes the form

$$F_\alpha^{CS} \approx \left(\delta^T - \delta^E\right) \sum_i^N \frac{\left|A_i^Q\right|}{e\, n_i^2 \left(A_i^{CS}\right)^2} \frac{\partial \delta^T}{\partial x_\alpha}. \quad (10)$$

To evaluate this formula the chemical shift (Equation 2) and its derivative in respect to the coordinates have to be calculated. For this, of course, first the charges (Equation 3) have to be derived. Since all expressions are analytically known the efficiency is obvious. For small systems consisting of around 100 atoms these calculations are done within seconds on a current Pentium II machine with ∼200 Mflop/s average math performance.

Simulation and structure refinement

The initial complex in Figure 2 without the water molecule was used as initial structure for a 1 ns molecular dynamics calculation at a temperature of 2000 K to make conformative energy barriers surmountable. 1000 structures originating 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, we only used a half side harmonic NOE pseudo potential, i.e., if any proton pair was equal or closer than the desired NOE distance then the forces were set to zero. The functional behavior of the potential and its gradient were kept continuous by this assumption. Thus we only obtained approximate proximities which resulted in rough structures. In order to obtain better defined structures we performed 1000 geometry optimizations with additional ^{13}C chemical shift pseudo forces. From the 22 possible ^{13}C chemical shifts we used 17. The current BPT parametrization turned out to be erroneous for the chemical shift prediction of CH_2-group carbon sites. These pseudo forces were omitted. After the optimizations, the three energetic lowest conformations with the smallest NMR parameter violations (Table 2) were selected from the conformational space of all 1000 structures. These COSMOS-NMR force field structures are shown in Figure 6. The rms NOE distance deviation after refinement is about 0.5 Å and the rms difference between restricted ^{13}C chemical shifts and experimental values is 0.1 ppm. If the -CH_2 carbon sites are also considered, the deviation is 1.2 ppm (see Table 2).

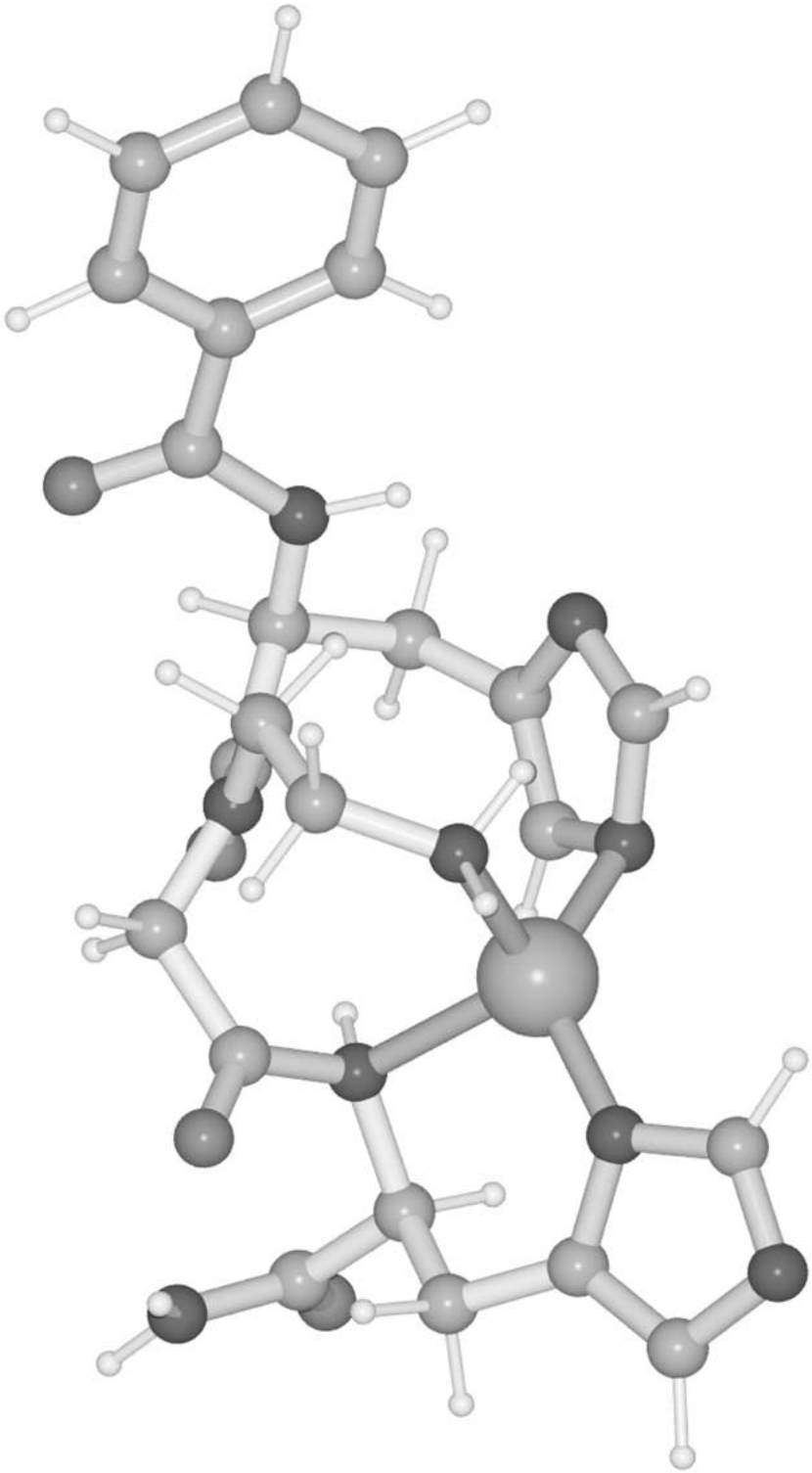

Figure 7. B3LYP/6-31G(d,p) DFT optimized structure of the conformation in Figure 6a.

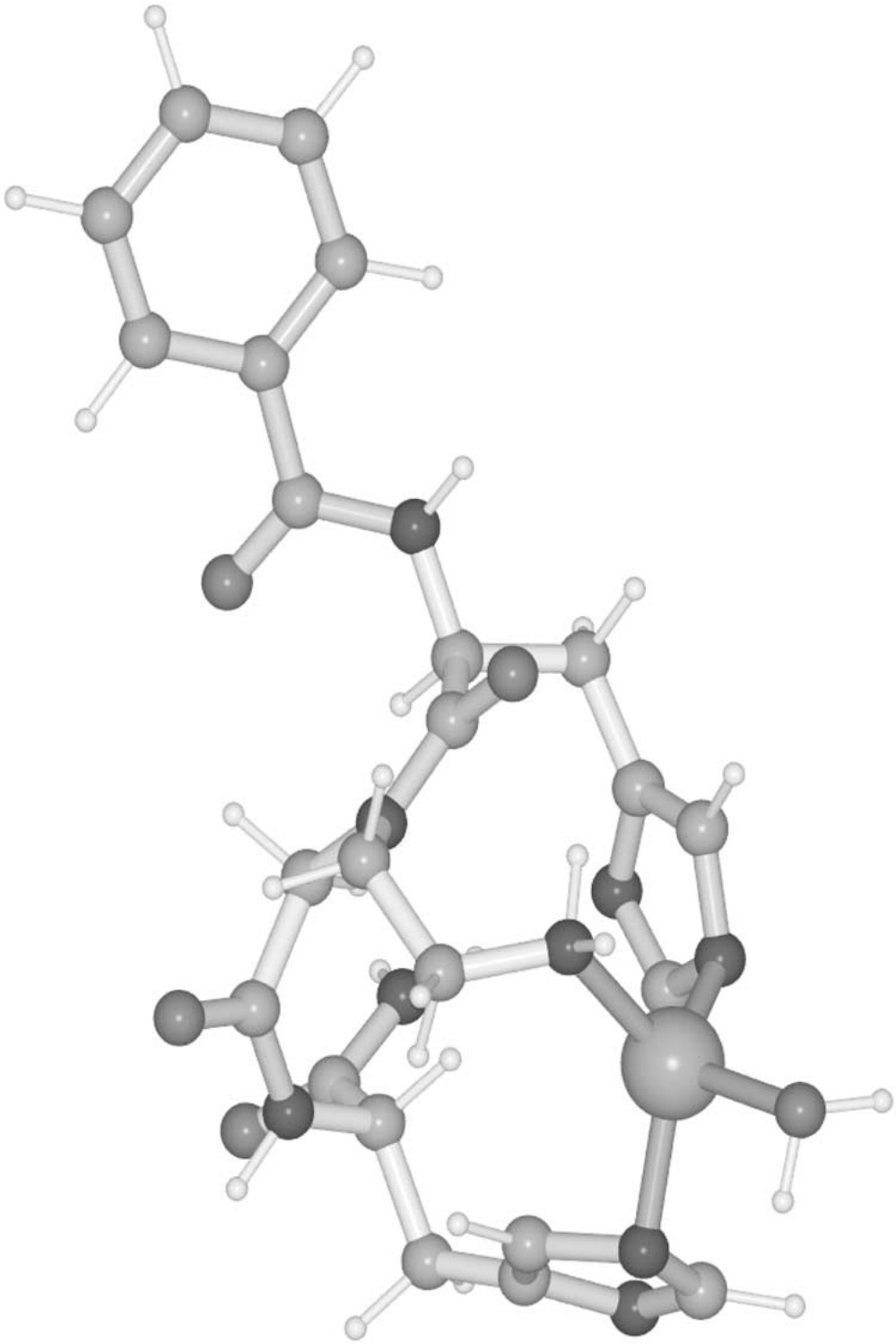

Figure 8. B3LYP/6-31G(d,p) DFT optimized structure of the complex Figure 6c) with an additional H_2O molecule to demonstrate the possibility of a tetrahedral coordination.

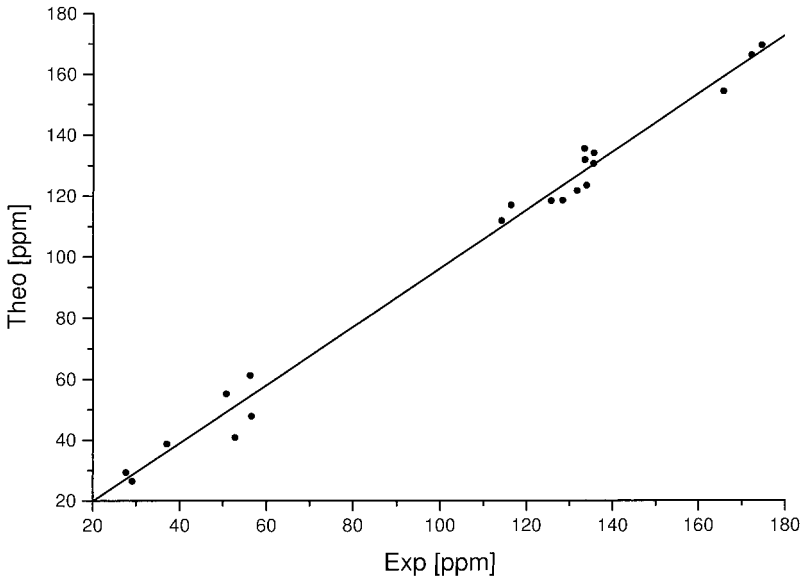

Figure 9. Correlation of the experimental and GIAO/6-31G(d,p) ^{13}C chemical shifts of the complex. The correlation coefficient is 0.9944 and the standard deviation is 5.1 ppm. In comparison to this, the correlation between the experimental values of the pure ligand and the theoretical complex results is slightly worse; R is 0.9938 and the deviation is 5.4 ppm. The experimental shifts of the free ligand and the complex correlate with 0.9991 and have an deviation of 2.2 pppm.

Table 2. NMR refinement data of three conformations a, b and c, Figure 6. In column 1 to 4 are the NOE distance deviation, the truncated NOE distance deviation (only distances are reported that larger than the NOE distances), the ^{13}C chemical shift deviation from experiment and the shift deviation without CH_2-group carbon sites

Conformation	a	b	c
ΔR [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

Table 3. Data of the conformations a, b and c found in Figure 6. The COSMOS force field energies and DFT minimum energies are given in columns 1 and 2. Columns 3 and 4 show the mean Zn-X bond lengths and the X-Zn-Y bond angles of the DFT optimized structures. The deviation of the latter is given in parentheses

Conformation	a	b	c
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)
α [°]	109.2 (10.6)	109.1 (11.1)	120.0 (7.0)

The energetically most stable structure is illustrated in Figure 6a. The NH_2-group position of the N-alkyl chain is important since it is the most obvious structural difference between the first two conformations. Complex 6a possesses two additional hydrogen bridges as compared to conformation b, thus stabilizing a.

In contrast to a and b structure c shows two trans-cis alternations and an imidazole ring flip. In order to find energetic minima of the free complex nearest to the NMR optimized structures B3LYP/6-31G(d,p) (Becke, 1993) full geometry optimizations were carried out with the GAUSSIAN98 program package (Gaussian 98, 1998) using the COSMOS-NMR structures as the starting point. The DFT optimized structure 6a is shown in Figure 7. It can be seen that the position of the NH_2-group has changed and the hydrogen bridges have disappeared. This might be duo to the fact that no diffuse functions were used. The B3LYP energy difference between 6a and 6b is rather small (6.3 kJ/mol). The corresponding energy difference in the COSMOS force field approach amounts to 77.5 kJ/mol, mainly originated from the three hydrogen bonds (see Figure 6).

The main difference between the DFT optimized structures 6a and 6b is the position of the -$CONH_2$ group in His^2. The average Zn-X bond length is ca.

(2.0 ± 0.1) Å. For the tetrahedral complexation of 6a and 6b, the bond angle is (109 ± 12)°. The theoretical angle would be 109.3°. The DFT optimized structure of 6c demonstrates a threefold coordination with an bond angle of (120 ± 7)°. The planarity varies around 4°. This conformation might be interpreted as being a transition state for binding H_2O. In order to confirm this possibility, a water molecule was added to structure 6c near the zinc ion and a full DFT geometry optimization was carried out. The minimum structure is shown in Figure 8. The relative energy in respect to structure 6a is −45.8 kJ/mol, taking the water molecule into account. This complex can be understood as the first step of the catalytic reaction cycle of the carbonic anhydrase. The reaction has the well known form: $CO_2 + 2H_2O \equiv HCO_3^- + H_3O^+$ (Mauksch et al., 2001).

Conclusion

The pseudotripeptide Zn-Bz-His-Ψ[CO-N$(CH_2)_2$-NH_2]Gly-His-NH_2 was modeled from a structure motive of the carbonic anhydrase reaction center. It was than experimentally synthesized and its structure determined (Figure 6). NOE and ^{13}C chemical shifts were used in a self developed COSMOS-NMR force field. For the first time, ^{13}C chemical shift pseudo forces have been applied to metal peptide complexes. This method could also be applied to metaloproteins. The force field structures were compared with the nearest *ab initio* minima and the energies were quantified. Finally we calculated the isotropic ^{13}C chemical shifts using the GIAO (Wolinski et al., 1990) B3LYP method. Two different basis sets were used: 6-31G(d,p) and 6-311G(d,p). The average over all three structures a, b, c is shown in Table 1. As reference the TMS shieldings were computed (187.1 ppm and 178.9 ppm). It turns out that the average shifts have a correlation with the experimental values of the complex with R_c(6-31G) = 0.9944 as well as R_c(6-311G) = 0.9945 (Figure 9) and of the free ligand 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, Sd_c(6-311G) = 5.6 ppm, Sd_l(6-31G) = 5.4 and Sd_l(6-311G) = 6.0 ppm, respectively.

One cannot assume that R_c is much closer to unity than R_l or that SD_c is much smaller than SD_l since the correlation between experimental shifts of the pseudopeptide and the complex differs only by 0.001 from unity and since the standard deviation is 2.2 ppm. Otherwise we used only free DFT minimum structures and it is known that the GIAO method predicts chemical shifts within an error of 5 ppm. Nevertheless the GIAO chemical shifts of the COSMOS-NMR force field complex structures, optimized by DFT, correlate best with the experimental data of the pseudotripeptide zinc complex.

Last, but not least, it turned out that a structure analogous to the first catalytic step of the carboanhydrase is possible. A stable transition state of threefold coordination was found.

Acknowledgements

Financial support by the Deutsche Forschungsgemeinschaft (Collaborative Research Centre 436, Jena, Germany) is gratefully acknowledged. We thank U. Graefe and H. Heinecke for the availability of the NMR equipment and their generous help; S. Irmer for the maintenance of the parallel clusters for the DFT calculations and the COSMOS GbR for the fruitful corporation.

References

Alia, Matysik, J., Erkelens, C., Hulsbergen, F.B., Gast, P., Lugtenburg, J. and de Groot, H.J.M. (2000) *Chem. Phys. Lett.*, **330**, 325–330.

Alsfasser, R., Ruf, M., Trofimenko, S. and Vahrenkamp, H. (1993) *Chem. Ber.*, **126**, 703.

Basosi, R., Gaggelli, E., Gaggelli, N., Pogni, R. and Valensin (1998) *Inorg. Chim. Acta*, **275**, 274–278.

Becke, A.D. (1993) *J. Chem. Phys.*, **98**, 5648.

Braun, S., Kalinowski, H.O. and Berger, S. (1998) *150 and More Basic NMR-Experiments*, Wiley-VCH, New York, NY.

Burns, G. (1964) *J. Chem. Phys.*, **41**, 1521.

Del R. (1958) *J. Chem. Soc.*, 4031.

Förster, M., Brasack, I., Duhme, A.-K., Nolting, H.-F. and Vahrenkam, H. (1996) *Chem. Ber.*, **129**, 347.

Gaussian 98 (1998) *Revision A.5*, Gaussian, Inc., Pittsburgh, PA.

Gockel, P., Gelinsky, M., Vogler, R. and Vahrenkamp, H. (1998) *Inorg. Chim. Acta*, **272**, 115.

Greiner, G., Seyfarth, L., Poppitz, W., Witter, R., Sternberg, U. and Reißmann, S. (2000) *Lett. Peptide Sci.*, **7**, 133–141.

Grotendorst, J. (2000) *Modern Methods and Algorithms of Quantum Chemistry*, NIC, 1.

Herr, U., Spahl, W., Trojandt, G. Steglich, W., Thaler, F. and van Eldrik, R. (1999) *Bioorg. Med. Chem.*, **7**, 699.

Karplus, M. (1959) *J. Chem. Phys.*, **30**, 11–15.

Kimura, E. (1994) In *Progress in Inorganic Chemistry*, Vol. 41, Karlin, K.D. (Ed.), Wiley, New York, NY, p. 443.

Klug, A. and Rhodes, D. (1987) *Trends. Biochem. Sci.*, **12**, 464.

Koch, F.T., Bräuer, M., Kunert, M., Sternberg, U. and Anders, E. (2001) *J. Mol. Model.*, **7**, 54–64.

Koch, F.-T., Losso, P. and Sternberg, U., *COSMOS: Computer Simulation of Molecular Structures*, www.cosmos-software.de

Koch, F.T., Möllhoff, M. and Sternberg, U. (1994) *J. Comp. Chem.*, **15**, 524.
Krizek, B.A., Amann, B.T., Kilfoil, V.J., Merkle, D.L. and Berg, J.M. (1991) *J. Am. Chem. Soc.*, **113**, 4518.
Lipari, G. and Szabo, A. (1982) *J. Am. Chem. Soc.*, **104**, 4546–4559.
Luisis, B.F., Xu, W.X., Otwinowsky, Z., Freedman, L.P., Yamamoto, K.R. and Sigler, P.B. (1991) *Nature*, **352**, 497.
Magafa, V., Stavropoulos, G. and Tsiveriotis, P. (1998) *Inorg. Chim. Acta*, **272**, 7–17.
Magonet E., Hayen, D., Delforge, D., Delaire, F. and Remacle, J. (1992) *Biochem J.*, **287**, 361 and references therein.
Malrieu, J.-P. (1977) *Mod. Theor. Chem.*, **7**, 69.
Marmorstein, R., Carey, M., Ptashne, M. and Harrison, S.C. (1992) *Nature*, **356**, 408.
Mauksch, M., Bräuer, M., Weston, J. and Anders, E. (2001) *Chem. Biochem.*, **2**, 190–198.
Möllhoff, M. and Sternberg, U. (2001) *J. Mol. Model.*, **7**, 90–120.
O'Keefe, M. and Brese, N.E. (1991) *J. Am. Chem. Soc.*, **113**, 3226–3229.
Priess, W. and Sternberg, U. (2001) *J. Mol. Struct: Theochem.*, **544** (1–3), 181–190.
Slater, J.C. (1930) *Phys. Rev.*, **36**, 57.
Spoel, D. (1996) *Structure and Dynamics of Peptides: Theoretical Aspects of Protein Folding*, Thesis, University Groningen, 23–28.
Sternberg, U. (1988) *J. Mol. Phys.*, **63**, 249.
Sternberg, U. and Priess, W. (1997) *J. Magn. Reson.*, **125**, 8–19.
Stillman, M.J., Shaw, C.F. and Suzuki, K.T. (1992) *Methallothioneins*, VCH, Weinheim.
Tsiveriotis, P., Hadjiliadis, N. and Stavropoulos, G. (1997) *Inorg. Chim. Acta*, **261**, 83–92.
Vallee, B.L., Coleman, J.E. and Auld, D.S. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 999.
van Eldrik, R. (1999) *Coord. Chem. Rev.*, **182**, 373.
Veeman, W.S. (1984) *Progr. NMR Spectrosc.*, **20**, 193–235.
Vuister, G.W., Delaglio, F. and Bax, A. (1992) *J. Am. Chem. Soc.*, **114**, 9674–9675.
Williamson, M.P. (1993) *Nat. Prod. Rep.*, 207–232.
Witter, R., Prieß, W. and Sternberg, U. (2002) *J. Comp. Chem.*, **23**, 298–305.
Wolinski, K., Hilton, J.F. and Pulay, P. (1990) *J. Am. Chem. Soc.*, **112**, 8251.
Zhang, X., Hubbard, C.D. and van Eldrik, R. (1996) *J. Phys. Chem.*, **100**, 9161.