Synthesis, ^{1}H, ^{13}C, ^{15}N, and ^{113}Cd NMR, ESI-TOF MS, Semiempirical MO (PM3), *ab initio*/HF and Cation/Anion Binding Studies of *N*-deoxycholyl-L-tryptophan

ELINA VIRTANEN[1],*,   JARI TAMMINEN[1], JUHA LINNANTO[1], PIA MÄNTTÄRI[1], PIRJO VAINIOTALO[2] and ERKKI KOLEHMAINEN[1]
[1]*Department of Chemistry, University of Jyväskylä, P.O.B. 35, FIN-40351 Jyväskylä, Finland;* [2]*Department of Chemistry, University of Joensuu, P.O.B. 111, FIN-80101 Joensuu, Finland*

Abstract

The synthesis, structural characterization, and cation/anion binding properties of a new bile acid-amino acid conjugate, *N*-deoxycholyl-L-tryptophan, are described. The structures of the ligand and its cadmium adduct at different pH conditions and various cadmium perchlorate concentrations were determined by modern multinuclear magnetic resonance spectroscopic as well as ESI-TOF MS techniques. Also semiempirical PM3 and *ab initio*/HF molecular modelling studies were performed. Based on ^{1}H,^{1}H NOESY measurements *N*-deoxycholyl-L-tryptophan in alkaline conditions was found to appear in a bent conformation which was clearly different from the conformations in neutral and acidic solutions. According to molecular modelling in its minimum energy structure the tryptophan backbone of the ligand was folded close to the deoxycholic acid skeleton and the structure was stabilized by an intramolecular hydrogen bond. The multinuclear magnetic resonance experiments indicated that Cd^{2+}-cation was bound with the ligand in neutral and alkaline conditions while in acidic conditions protons block the binding site. ESI-TOF MS revealed clearly a competition between sodium and cadmium ions, the ligand having a stronger affinity for sodium. Cadmium binding occurred only when excess of cadmium was used. Further, ESI-TOF MS spectra showed that various chlorine oxyanions originated from perchlorate anion formed together with cations different adducts with the ligand.

Introduction

The large, rigid, and curved steroidal skeleton, chemically different hydroxyl groups, enantiomeric purity, and unique amphiphilicity together with their availability and low cost make bile acids ideal building blocks for designing novel macrocyclic and also open-chain supramolecular hosts [1–8]. These host molecules can serve as model compounds for more complex biological systems and are important, for example, for molecular recognition of substrates in enzymatic processes.

Bile acids and their derivatives are important compounds also from the pharmacological point of view. They have been used in the treatment of bile acid deficiency diseases, liver diseases, and in dissolution of cholesterol gallstones [9]. The antiviral properties of bile acids and their sulfate esters [10] and the antifungal properties of some bile acid esters [11] have been studied with promising results. They have potential to act as carriers of liver-specific drugs, absorption enhancers and as cholesterol level lowering agents [12]. Other potential medicinal applications for bile acids and their conjugates have been their use as non-opiate analgesics [13], as sensitizers of Gram-negative bacteria to antibiotics [14–15], and as radiopharmaceuticals [16].

Glycine and taurine conjugates of bile acids are well-known ingredients of human bile. Synthetic procedures to prepare those natural bile acid-amino acid conjugates have been known for almost 100 years [17]. Later on many reports concerning syntheses of different kinds of natural and non-natural bile acid-amino acid conjugates have been published [18–22]. Roda *et al*. have studied the syntheses and physicochemical, biological, and pharmacological properties of bile acid-cyclic amino acid conjugates [23]. Albert and Feigel *et al*. have synthesized various 3α-aminolithocholic acid and phenylalanine based cyclopeptides and studied their conformations by several NMR-techniques, molecular dynamics calculations, and X-ray crystal studies [24–26]. However, as far as we are aware the syntheses and characteristics of bile acid-L-tryptophan conjugates are not reported. Tryptophan (Trp), as one of the naturally occurring amino acids, has been widely used as a fluorescence probe to monitor protein conformations and dynamics [27]. Based on the FTIR spectroscopic technique (DSD = dissolution-spray-deposition) and *ab initio* molecular orbital calculations it

* Author for correspondence: E-mail: elvirtan@cc.jyu.fi

has been proposed that for a zwitter ionic form of Trp in a solid KBr matrix there are several non-interconverting rotamers, one of which is predominant [28]. Also in solution the predominance of one conformer of Trp has been suggested based on ^{1}H NMR parameters, T_1-relaxation times, lanthanide induced chemical shift changes, and intramolecular nuclear Overhauser enhancements [29]. ^{1}H NMR and fluorescence spectroscopy have shown that terbium cation (having the same ionic radius as calcium) interacts with the amino group and the imido group of the indole ring of Trp [30]. De Wall *et al.* have presented experimental confirmation, based on NMR and X-ray studies, of cation-π interactions between alkali metal ions and the indole ring of Trp [31]. Recent mass spectrometric studies reveal that Trp can form complexes also with other cations such as aluminum and copper [32–34]. Further, Sharma *et al.* have reported synthetic, spectral, and thermal studies of cobalt(II), nickel(II), zinc(II), and cadmium(II) complexes with amino acid (e.g., Trp) derived Schiff bases [35].

A general objective of this study is to prepare bile acid-amino acid conjugates which have some potential to act as tissue targeting medicines. One should remember that L-tryptophan is a precursor of serotonin which is known as a neurotransmitter. A more chemically focused aim is to clarify how various solution conditions and cation concentrations are reflected in the multinuclear NMR spectroscopic, ESI-TOF MS spectrometric, and Cd^{2+}-complexing properties of *N*-deoxycholyl (or $3\alpha,12\alpha$-dihydroxy-5β-cholan-24-oyl)-L-tryptophan. The reasons for selecting deoxycholic acid for our first study are the physiological significance of the molecule and its ability to form choleic acids (crystalline inclusion compounds) [36]. The criterion for selecting Cd^{2+} as a complexing agent is that the ionic radius of the cation is close to the ionic radii of the biochemically important ions, Ca^{2+} and Zn^{2+}. However, the NMR characteristics of Cd^{2+} are much more favourable than those of Ca^{2+} and Zn^{2+} [37]. Further, in order to deepen the interpretation and discussion of the obtained spectroscopic data, some molecular orbital calculations have also been included. This study will serve our long-term project to prepare and characterize supramolecular hosts derived from bile acids and nitrogen-containing heterocycles as reviewed recently by us [8].

Experimental

Compounds

Deoxycholic acid ($3\alpha,12\alpha$-dihydroxy-5β-cholan-24-oic acid) was 98% reagent from Aldrich Chemical Co. L-tryptophan was purchased from Fluka Chemie AG as puriss. grade reagent. *N*-($3\alpha,12\alpha$-dihydroxy-5β-cholan-24-oyl)-L-tryptophan was synthesized by the mixed anhydride method following a previously reported procedure [23, 38].

The synthetic route to *N*-deoxycholyl-L-tryptophan is described in Scheme 1.

2.06 g (11 mmol) of tri-*n*-butylamine was added to a suspension of 3.93 g (10 mmol) of deoxycholic acid in 60

Scheme 1.
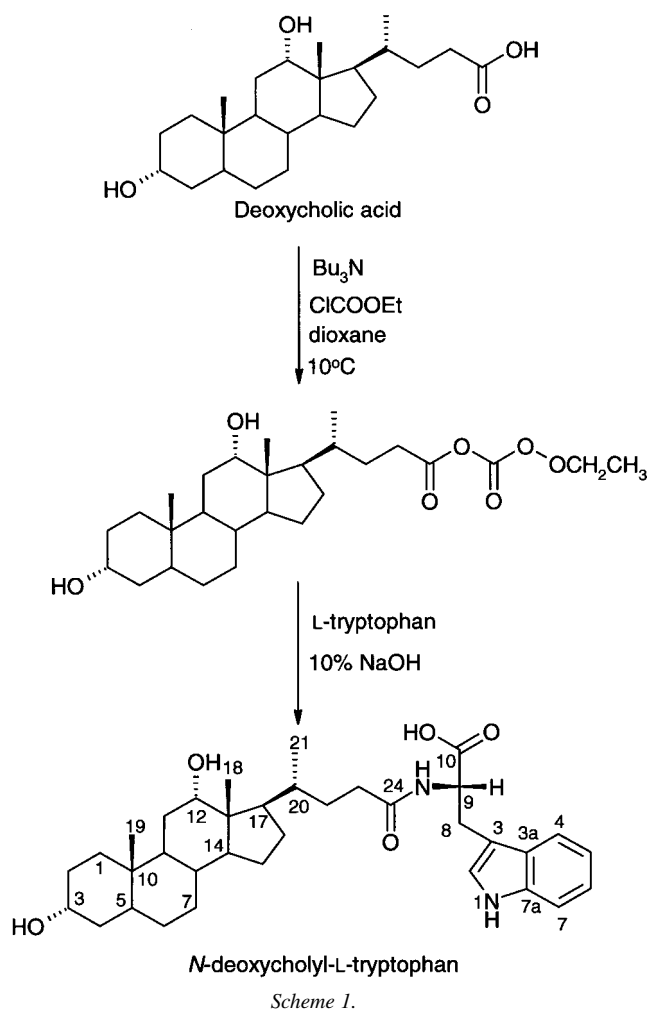
mL of sodium dried 1,4-dioxane at 10 °C. The temperature was kept constant while stirring the reaction mixture and a solution of 1.20 g (11 mmol) of ethyl chloroformate in 3 mL of 1,4-dioxane was added dropwise during 15 min. After the solution was stirred for an additional 30 min, a solution of 2.45 g (12 mmol) of L-tryptophan in 5 mL of 10% NaOH was added dropwise at the same temperature, and the resulting solution was stirred for 4 h. During the stirring the temperature was allowed to reach room temperature spontaneously. The reaction mixture was poured into 125 g of ice and the mixture was stirred and acidified (pH 2, conc. HCl). The stirring was continued for 30 min and then the mixture was allowed to stand for 15 min. The precipitate formed was filtered with suction and washed by chloroform to remove the tri-*n*-butylamine. The crude product obtained was finely ground and dried *in vacuo* at 50 °C for 3 h. The yield of the crude product was 4.20 g (73%), which was purified by column chromatography [39] (silica gel, chloroform:acetone:methanol, 70:50:10). The pure product was obtained in 45% yield. The purity and molecular structure of *N*-deoxycholyl-L-tryptophan was ascertained by ^{1}H and ^{13}C NMR and accurate ESI-MS measurements.

The Cd^{2+}-binding experiments were done by dissolving an accurate amount of $Cd(ClO_4)_2 \cdot 6H_2O$ in 1 mL of DMSO-d_6 containing 58 mg (0.1 mmol) of *N*-deoxycholyl-

L-tryptophan so that the cation:ligand molar ratios were 0, 1:2, 1:1, 2:1, 3:1, 4:1, and finally DMSO-d_6 solution saturated by cadmium perchlorate hexahydrate. In order to compare the ^{15}N and ^{113}Cd NMR chemical shifts of L-tryptophan and the ligand 83.5 mg of Cd(ClO$_4$)$_2$·6H$_2$O was added into 36 mg of L-tryptophan in DMSO-d_6 giving a cation:ligand molar ratio of 1.1:1. For the MS measurements the sample solutions were diluted to concentrations of 4 mg/mL by using DMSO (HPLC grade) and further to concentrations of 0.01 mg/mL by using acetonitrile (HPLC grade). For the accurate mass measurement 4 mg of the ligand was dissolved in 1 mL of DMSO (HPLC grade) and further diluted to a concentration of 0.01 mg/mL by using acetonitrile/HCOOH (0.2%).

Warning: Cd(ClO$_4$)$_2$·6H$_2$O can explode when heated or reacting with organic material. It is strictly forbidden to try to prepare an anhydrous salt by heating from the hexahydrate!

NMR and MS

All ^{1}H, ^{13}C, ^{13}C DEPT-135, PFG ^{1}H,^{13}C HMQC [40–41], PFG ^{1}H,^{13}C HMBC and PFG ^{1}H,^{15}N HMBC [42], DQF ^{1}H,^{1}H COSY [43 - 44], and ^{1}H,^{1}H NOESY [45–46] NMR measurements were performed in 0.1–0.2 M DMSO-d_6 solutions with a Bruker Avance DRX 500 spectrometer equipped with a 5 mm diameter broad band inverse observation (proton) probehead with a z-gradient accessory [47] working at 500.13 MHz in ^{1}H, 125.77 MHz in ^{13}C, and 50.70 MHz in ^{15}N experiments, respectively. Because of the insensitivity of the ^{113}Cd nucleus, ^{113}Cd NMR measurements were performed using a direct observation probehead without the z-gradient accessory at 110.94 MHz. The ^{1}H and ^{13}C NMR chemical shifts were referenced to the residual signal of partly deuteriated solvent: δ(C^{1}HD$_2$) = 2.50 ppm from the internal TMS and to the signal of solvent δ(^{13}CD$_3$) = 39.50 ppm from the internal TMS, respectively. The ^{15}N NMR chemical shifts were referenced to the signal of external nitromethane, δ(CH$_3$^{15}NO$_2$) = 0.0 ppm, in a 1 mm diameter capillary inserted coaxially inside the 5 mm NMR tube. The ^{113}Cd NMR chemical shifts were referenced to the signal of an external 0.1 M aqueous Cd(ClO$_4$)$_2$, δ(^{113}Cd) = 0.0 ppm, in a 1 mm diameter capillary inserted coaxially inside the 5 mm NMR tube. The complete lists of the acquisition and processing parameters are available from E.K. on request.

In the PFG ^{1}H,^{15}N HMBC experiment the pulse program, inv4gslplrnd, was selected from the manufacturer's library and a 50 ms evolution delay (corresponding to 10 Hz proton-nitrogen-15 spin-spin coupling) was incorporated in the pulse program. The PFG ^{1}H,^{15}N HMBC (heteronuclear multiple bond correlation) pulse program was selected because both in the reference and the studied compound in alkaline solutions the nitrogens are unprotonated. However, in the HMBC experiment all proton bearing nitrogens are also visible as doublets split by proton-nitrogen spin-spin couplings when the low-pass filter set at the start of the pulse program does not exactly match the evolution delay of ^{1}J(N,H). By this arrangement the information given by both

HMQC and HMBC is obtained by a single heteronuclear chemical shift correlation experiment.

Electrospray mass spectrometric measurements were performed using a LCT time of flight (TOF) mass spectrometer with electrospray ionization (ESI; Micromass LCT). The spectrometer utilized two hexapole RF lenses to transfer ions from the source to the orthogonal acceleration TOF mass analyzer. Ions emerging from the analyzer were detected by a dual microchannel plate detector and ion counting system. Controlling the LCT as well as acquiring and processing the data were performed with a MassLynx NT software system. In each experiment a flow rate of 20 μl/min was used for the sample solution and the sample droplets were dried with nitrogen gas. The potentials of 80 V and 5 V for the sample and extraction cones were applied. The RF lens was set at a potential of 500 V and the potential in the capillary at 3250 V. The desolvation temperature was set at 140 °C and the source temperature at 100 °C.

For the accurate mass measurement the calibration of the instrument was made with NaI. The nominally measured mass and the centroided spectrum were subjected to a mass scale correction obtained from the variation of the observed value of the reference mass from its theoretical monoisotopic mass. The necessary mass scale correction was applied to the spectrum relative to the fifth order calibration curve used for the acquisition of the raw data. An additional mass correction was applied from an algorithm which accounts for the observed statistical shift in the centroided mass with signal intensity resulting from the detector dead time. The reference was introduced by flow injection (20 μL/min) and the sample by loop injection. The potentials of 70 V and 3 V for the sample and extraction cones were applied. The RF lens was set at a potential of 250 V and the potential in the capillary at 3300 V. The desolvation temperature was set to 180 °C and the source temperature to 100 °C. Leucin-enkephalin (1 ng/μL) was used as a reference ion.

Molecular orbital calculations

The geometry of the ligand and its cadmium complexes (cation:ligand; 1:1 and 1:2, respectively) were fully optimized at the semiempirical PM3 [48] level on a Silicon Graphics O2 workstation by using SPARTAN (version 5.0) software [49]. After that a substructure of the ligand was optimized with and without Cd^{2+} cation at the *ab initio* Hartree-Fock (HF) level using the 3-21G(d) basis set on a Silicon Graphics Origin 200 workstation by Gaussian 98 software [50].

Results and discussion

NMR

^{1}H and ^{13}C NMR: The ^{1}H and ^{13}C NMR chemical shifts of *N*-deoxycholyl-L-tryptophan in neutral, acidic (HCl addition) and alkaline (NaOH addition) DMSO-d_6 solution without and with 1:1 Cd(ClO$_4$)$_2$·6H$_2$O additions are collected in Tables 1 and 2, respectively.

Table 1. ^{1}H NMR chemical shifts (ppm from DMSO-d_6, δ = 2.50 ppm from int. TMS) of the selected protons of *N*-deoxycholyl-L-tryptophan in neutral, acidic and alkaline DMSO-d_6 solutions without and with Cd(ClO$_4$)$_2\cdot$6H$_2$O additions (in a molar ratio of 1:1) measured at 30 °C

| | $\delta(^1\mathrm{H})$/ppm | | | | | |
| | Neutral | | Acidic | | Alkaline | |
Proton	ligand	ligand + Cd^{2+}	ligand	ligand + Cd^{2+}	ligand	ligand + Cd^{2+}
3β	3.37	3.37	3.35	3.35	3.35	3.35
12β	3.77	3.77	3.74	3.74	3.74	3.73
18 Me	0.57	0.56	0.51	0.50	0.55	0.50
19 Me	0.84	0.83	0.78	0.78	0.80	0.79
21 Me	0.88	0.87	0.82	0.82	0.82	0.81
23 α & β	1.93 & 2.06	1.94 & 2.08	1.96 & 2.07	1.95 & 2.07	1.86 & 2.01	1.88 & 2.02
Trp-2	7.06	7.12	7.09	7.08	7.03	7.08
Trp-4	7.49	7.51	7.47	7.47	7.42	7.47
Trp-5	6.91	6.95	6.94	6.94	6.81	6.91
Trp-6	7.00	7.04	7.03	7.03	6.90	7.00
Trp-7	7.29	7.31	7.31	7.30	7.25	7.28
Trp-8[a]	3.20	3.21	3.13	3.13	3.15	3.21
Trp-8[b]	2.99	2.98	2.97	2.97	2.93	2.97
Trp-9	4.25	4.48	4.41	4.41	4.18	4.40
Amide N*H*	7.39	7.71	8.04	7.98	Deprot.	7.40
Indole N*H*	10.80	10.71	10.75	10.61	Deprot.	10.55

[a]H *cis* to carbonyl oxygen of L-tryptophan.
[b]H *trans* to carbonyl oxygen of L-tryptophan.

For the deoxycholyl fragment only some proton chemical shifts are given in Table 1 because at 500 MHz unambiguous assignments of all protons are not possible. The ^{1}H and ^{13}C NMR chemical shift assignments are based on the literature spectra of deoxycholic acid [51] and L-tryptophan [52] as well as on the present ^{13}C DEPT-135, PFG ^{1}H,^{13}C HMQC, and PFG ^{1}H,^{13}C HMBC experiments.

Although the majority of the ^{1}H and ^{13}C NMR chemical shifts of *N*-deoxycholyl-L-tryptophan do not significantly depend on the pH or Cd^{2+}-cation content, some trends, however, can be observed. First of all, increasing the ionic strength by cadmium perchlorate causes opposite bulky effects on the ^{1}H and ^{13}C NMR chemical shifts of the atoms of the steroidal skeleton being deshielding in carbon-13 and shielding in proton, respectively. In a clearly acidic solution additions of cadmium perchlorate do not induce any significant changes due to the excess of protons which probably block the proper binding sites for cadmium. In neutral and alkaline conditions the Cd^{2+}-adduct formation is reflected in the ^{1}H NMR chemical shift changes, especially in the amide proton and H-9 of the Trp moiety, when the Cd^{2+}:ligand ratio is changed from 1:2 to 1:1. In the ^{13}C NMR the most significant change happens in the carbonyl carbon (C-10) of the Trp moiety, the signal shifting downfield by about 2 ppm when the cation:ligand molar ratio exceeds 1:1. In neutral conditions the signal additionally gets significantly broadened. A clear downfield (decreased shielding) change in the ^{13}C NMR chemical shift of C-24 of the deoxycholyl fragment is also observed. The changes in the ^{13}C NMR chemical shifts for a majority of the carbons of the Trp moiety also differ from the bulky effect. For example, an upfield

(increased shielding) change in the chemical shift of C-9 of the Trp moiety is detected.

In order to clarify whether the acidity *vs.* alkalinity of the solution has some influence on the three-dimensional structure of the *N*-deoxycholyl-L-tryptophan two-dimensional ^{1}H,^{1}H NOESY measurements have also been performed for the neutral, acidic, and alkaline DMSO-d_6 solutions of the ligand. In neutral and acidic conditions the NOESY maps are comparable revealing intra residual dipolar relaxations, such as the one between H-2 and H-8 in the tryptophan residue as well as the one between H-12β and the angular methyls 18 and 19 inside the steroidal system. In alkaline solution, however, some extra cross-peaks are visible which imply that there is close spatial proximity between the steroidal and amino acid moieties. For example, H-9 of the tryptophan part correlates with methyl 21 of the bile acid residue and a strong cross-peak is observed between the H-4 of tryptophan and H-12β of the steroid. Our molecular modelling studies reveal that in the minimum energy structure of the ligand the tryptophan moiety really has folded close to the steroid moiety and the structure is stabilized by an intramolecular hydrogen bond between 12α-*O*H and Trp-5-*H* as described later in this paper. Figure 1 shows a partial NOESY contour map of *N*-deoxycholyl-L-tryptophan in alkaline DMSO-d_6 solution showing the inter residual NOE cross-peaks.

It seems obvious that in alkaline solution a conformation, where the indole ring and the steroid skeleton have some kind of stacked arrangement explaining the observed inter residual NOE cross-peaks, is preferred.

^{15}N NMR: In order to characterize the possible binding sites for Cd^{2+}-cation, ^{15}N NMR chemical shifts of *N*-

Table 2. ^{13}C NMR chemical shifts (ppm from DMSO-d$_6$, δ = 39.50 ppm from int. TMS) of *N*-deoxycholyl-L-tryptophan in neutral, acidic and alkaline DMSO-d$_6$ solutions without and with Cd(ClO$_4$)$_2\cdot$6H$_2$O additions (in a molar ratio of 1:1) measured at 30 °C

| | δ(^{13}C)/ppm | | | | | |
| | Neutral | | Acidic | | Alkaline | |
Carbon	ligand	ligand + Cd^{2+}	ligand	ligand + Cd^{2+}	ligand	ligand + Cd^{2+}
1	35.11	35.18	35.57	35.65	35.36	35.34
2	30.20	30.27	30.48	30.56	30.19	30.20
3	69.92	70.04	70.77	70.99	70.34	70.44
4	36.27	36.33	36.50	36.57	36.22	36.21
5	41.61	41.68	42.15	42.26	41.89	41.93
6	26.97[a]	27.05[a]	27.49[a]	27.60[a]	27.22[a]	27.27[a]
7	26.08	26.17	26.61	26.73	26.33	26.39
8	35.64	35.72	36.22	36.34	35.94	36.00
9	32.90	33.00	33.51	33.66	33.23	33.30
10	33.78	33.86	34.31	34.43	34.03	34.08
11	28.56	28.66	29.03	29.13	28.72	28.76
12	71.00	71.12	71.93	72.18	71.48	71.59
13	45.94	46.02	46.52	46.64	46.20	46.24
14	47.40	47.51	47.99	48.14	47.66	47.74
15	23.48	23.56	24.05	24.16	23.76	23.81
16	27.13[a]	27.20[a]	27.58[a]	27.69[a]	27.39[a]	27.46[a]
17	46.19	46.26	46.74	46.89	46.38	46.49
18	12.40	12.49	12.93	13.05	12.65	12.70
19	23.04	23.12	23.54	23.64	23.27	23.30
20	35.14	35.18	35.54	35.65	35.36	35.43
21	17.08	17.15	17.56	17.67	17.29	17.33
22	31.49	31.57	32.06	32.21	31.64	31.73
23	32.74	32.53	32.79	32.95	33.09	32.96
24	171.71	172.42	174.18[b]	174.35[b]	172.60	173.25
Trp-2	123.19	123.39	124.12	124.25	125.65	123.56
Trp-3	111.26	110.85	110.42	110.54	110.77	111.53
Trp-3a	127.97	127.63	127.68	127.80	128.71	128.10
Trp-4	118.43	118.27	118.64	118.87	118.42	118.69
Trp-5	117.81	118.22	119.07	119.21	117.45	118.57
Trp-6	120.36	120.77	121.61	121.84	119.83	121.07
Trp-7	111.03	111.28	112.07	112.21	112.11	111.24
Trp-7a	135.91	136.06	136.64	136.77	137.71	136.09
Trp-8	27.51	27.84	27.71	27.84	28.04	28.06
Trp-9	54.56	53.49	53.66	53.80	55.71	54.63
Trp-10	174.70	176.94(br)	174.23[b]	174.57[b]	175.90	177.74

[a,b] Assignments may be interchanged.

deoxycholyl-L-tryptophan have been determined. Table 3 contains the ^{15}N NMR chemical shifts of L-tryptophan and its Cd^{2+}-adducts in neutral DMSO-d$_6$ solution as well as the shifts of *N*-deoxycholyl-L-tryptophan and its Cd^{2+}-adducts in neutral, acidic (HCl addition), and alkaline (NaOH addition) solutions.

As can be seen, in the neutral conditions the chemical shifts of the indole nitrogen of L-tryptophan, −248.9 ppm, and of *N*-deoxycholyl-L-tryptophan, −250.4 ppm, are quite close to the corresponding chemical shift of β-methylindole (−243.3 ppm) [53], which is a metabolite of tryptophan. The formation of an amide bond explains the big difference between the chemical shifts of the amino nitrogen of L-tryptophan, −339.6 ppm, and of *N*-deoxycholyl-L-tryptophan −255.7 ppm. Addition of Cd(ClO$_4$)$_2\cdot$6H$_2$O in

a molar ratio of 1.1:1 to the DMSO-d$_6$ solution of L-tryptophan giving a saturated solution led to only minor changes in the ^{15}N chemical shifts of the molecule, the resulting values being −249.1 ppm (indole) and −341.1 ppm (amide). Also, the ^{15}N chemical shift changes of the indole and amide nitrogens between the ligand and its cadmium adducts were considerably small (0.2 and 1.2 ppm, respectively) compared to the changes between the NH$_2$- and NHCO-groups of *N*-lithocholyl- and *N*-deoxycholyl-(2-aminoethyl)amides and their Cd^{2+}-complexes reported by us before [54].

In the alkaline solutions the indole and the amide protons were totally missing and the chemical shift of the indole nitrogen was clearly deshielded, resonating at −240.5 ppm, (−250.4 ppm in neutral conditions). The amide nitrogen

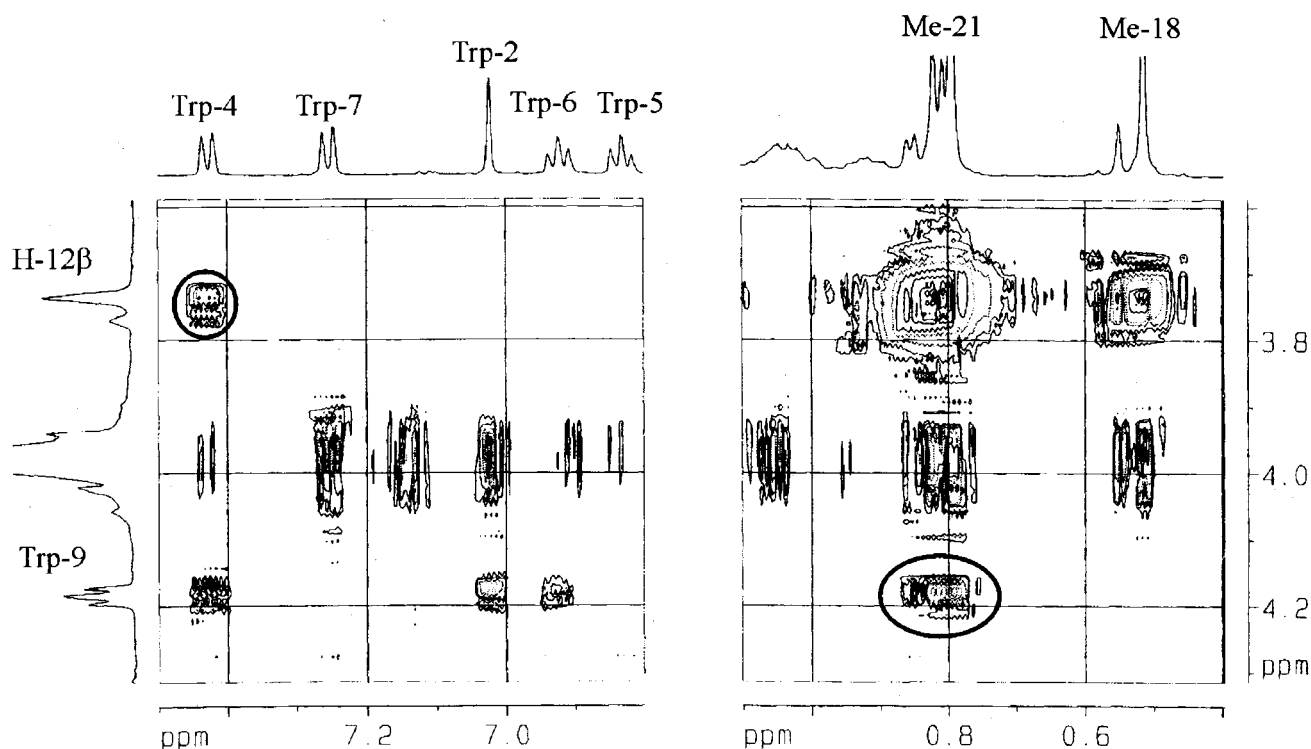*Figure 1.* A partial ^{1}H,^{1}H NOESY contour map of *N*-deoxycholyl-L-tryptophan in alkaline DMSO-d$_6$ solution showing (by circled cross-peaks) the correlation between H-9 of the tryptophan moiety and methyl 21 of the bile acid residue as well as the strong cross-peak between H-4 of tryptophan and H-12β of the steroid.

Table 3. ^{15}N NMR chemical shifts (ppm from ext. neat CH$_3$NO$_2$, δ = 0.0 ppm) of L-tryptophan and its Cd^{2+}-complex as well as of *N*-deoxycholyl-L-tryptophan and its Cd^{2+}-complexes in neutral, acidic and alkaline DMSO-d$_6$ solutions

	$\delta(^{15}$N/ppm)	
	Indole *N*	Amide *N* (amino *N* in L-tryptophan)
L-tryptophan	−248.9	−339.6
L-tryptophan + Cd^{2+}	−249.1	−341.1
ligand (neutral)	−250.4	−255.7
ligand + Cd^{2+} (neutral)	−250.6	−256.9
ligand (acidic)	−250.3	−258.4
ligand + Cd^{2+} (acidic)	−250.8	−258.4
ligand (alkaline)	−240.5	−254.2
ligand + Cd^{2+} (alkaline)	−251.6	−256.1

was also deshielded, resonating at −254.2 ppm, (−255.7 ppm in neutral solvent). Addition of Cd(ClO$_4$)$_2$·6H$_2$O in a molar ratio of 1:1 to the alkaline solution shifted the protolytic equilibrium back to the protonated indole form which showed a cross-peak at −251.6 ppm. Another interesting phenomenon in the ^{15}N NMR chemical shifts of *N*-deoxycholyl-L-tryptophan was that in the acidic conditions the chemical shift of the amide nitrogen was shielded, resonating at −258.4 ppm (−255.7 ppm in the neutral solvent).

^{113}Cd NMR: ^{113}Cd NMR opens a direct route to probe the environment of the bound cation itself. Only a weak signal for Cd^{2+}-cation bound with L-tryptophan was observed in neutral DMSO-d$_6$ at 299.5 ppm (from external 0.1 M aqueous Cd(ClO$_4$)$_2$) while the majority of the Cd^{2+}-cations was complexed with the solvent and resonated at −28.7 ppm (upfield from the reference). A 1:1 mixture of Cd^{2+}:*N*-deoxycholyl-L-tryptophan in neutral DMSO-d$_6$ resonated at 299.2 ppm (practically the same as in case of Trp). However, a sharp and strong peak revealed that the association constant was greater than in the case of Trp; especially because no signal of the Cd^{2+}:solvent complex was visible. The ^{113}Cd NMR chemical shifts of the cadmium:L-tryptophan and :*N*-deoxycholyl-L-tryptophan adducts measured in this work differ significantly from the chemical shifts of the Cd^{2+}-complexes of *N*-lithocholyl- and *N*-deoxycholyl-(2-aminoethyl)amides (92 and 84 ppm) measured by us previously [54]. The different binding sites could probably explain the big difference between the chemical shifts. In the PM3 optimized structure of the Cd^{2+}-complex of *N*-lithocholyl-(2-aminoethyl)amide, the Cd^{2+}-cation is coordinated to the carbonyl oxygen of the lithocholyl moiety and to both amino nitrogens of the 2-aminoethylamide moiety. In the present case coordination to the carbonyl oxygen and to the amide nitrogen of the Trp moiety is the most probable site based on steric reasons and also supported by molecular modelling. A disappearance of the ^{113}Cd NMR signal of the Cd^{2+}:*N*-deoxycholyl-L-tryptophan after saturating the solvent with Cd(ClO$_4$)$_2$·6H$_2$O is probably due to the decomposition of the Cd^{2+}-adduct. The buffer capacity

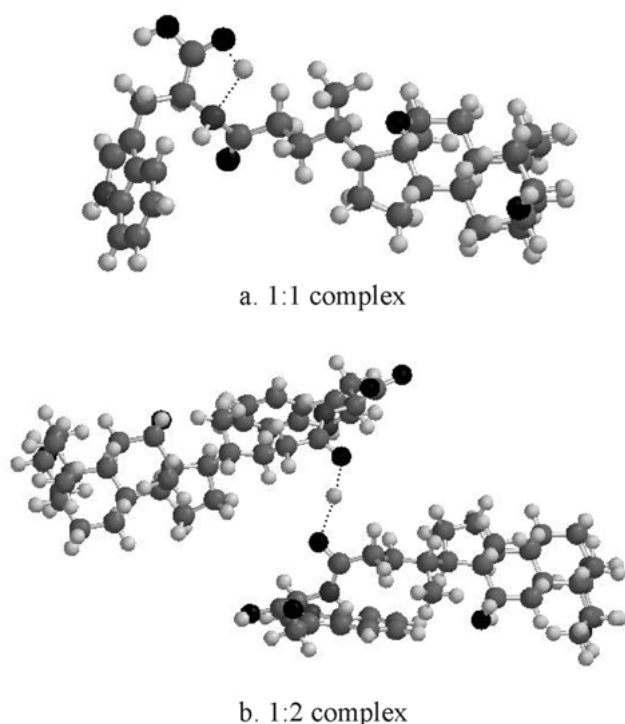a. 1:1 complex

b. 1:2 complex

Figure 2. The energetically most stable structures for cadmium:ligand 1:1 (a) and 1:2 (b) adducts optimized by semiempirical PM3 level of theory.

of the ligand is not able to withstand the strong acidifying effects caused by the increased concentration of cadmium cation. However, it is clear that the complexation tendency of *N*-deoxycholyl-L-tryptophan is stronger than that of L-tryptophan in similar conditions.

Although the ^{1}H, ^{13}C, and ^{113}Cd NMR data clearly show that Cd^{2+}-cation binds with the ligand in neutral and alkaline conditions, ^{15}N NMR shows surprisingly small changes concerning the cadmium binding. This could be due to the simultaneous interaction of the perchlorate counter anion that cannot be detected directly by NMR experiments. In order to determine if the counter anion interacts with the ligand as detected before for open-chain piperazine-containing ligands [55], careful ESI-TOF MS measurements were performed.

MS

In order to ensure the molecular weight of the ligand, an accurate mass measurement using leucin-enkephalin as a lock mass reference was performed. A centroided spectrum for the ligand resulted in MS *m/z*, ESI-TOF$^+$ found 579.3820 [M + H]$^+$, C$_{35}$H$_{51}$N$_2$O$_5$ requires [M + H]$^+$ = 579.3798. The measured value deviated 3.7 ppm from the theoretical one.

In order to find support for the results obtained from the different NMR experiments described before, ESI-TOF MS measurements at different solution conditions were performed. For ESI-TOF MS measurements the NMR samples were diluted to suitable concentrations which caused changes in the pH conditions compared to the circumstances prevailing during the NMR measurements. At neutral conditions singly charged [L + Na]$^+$ and [2L + Na]$^+$

ions were formed, even though alkali metal salts were not added to the sample solution. Fragmentation of the ligand did not occur under normal measurement conditions. At diluted acidic conditions signals of singly charged [L + H]$^+$, [L + K]$^+$, [L − H + 2Na]$^+$, [2L + H]$^+$, [2L + K]$^+$ and [2L − H + 2Na]$^+$ ions with low intensities in addition to the previously mentioned sodium adducts, which occurred as the main signals of the spectrum, were also observed. Lowering the pH with a small amount of acid was immediately detected by an increase in the intensity of the proton adducts. A higher proton concentration would probably have led to replacement of sodium and potassium entirely by protons. At alkaline conditions the singly charged signals of the potassium adducts [L + K]$^+$ and [2L + K]$^+$ were detected with higher intensities than in the previous cases. The [L + Na]$^+$ ion gave the most intense signal also at more alkaline conditions, but the signal of [2L + K]$^+$ was stronger than the signal of [2L + Na]$^+$. The other observed ions were the same as in the dilute acidic conditions described above.

In order to investigate the cadmium affinity of the ligand also by means of ESI-TOF MS the diluted NMR samples were measured. The salt used was cadmium perchlorate as mentioned before. The metal was at oxidation stage II and remained at this stage during complexation. Because of the numerous oxyanions of chlorine, however, the interpretation of the mass spectra was somewhat complicated. Signals of singly charged [ligand + oxyanion + cation(s)]$^+$ ions, typically [L + ClO + 2H]$^+$, [L + ClO + H + Na]$^+$, [L + ClO + H + K]$^+$, [L + ClO$_3$ + H + Na]$^+$, [L + ClO$_4$ + H + K]$^+$, and [L + 2ClO + H + 2Na]$^+$ as well as [2L + ClO + 2H]$^+$ and [2L + ClO$_3$ + 2H]$^+$ were detected at every cadmium concentration in addition to the proton, alkali metal, and cadmium adducts of the ligand. Interestingly, the oxyanion series was not observed at slightly alkaline pH conditions even though it was clearly detectable at acidic conditions.

Experiments at different cadmium:ligand ratios were performed in order to find out the amount of cadmium required for complexation reaction. These experiments were performed at neutral pH conditions. At a ratio of 1:2 (cadmium:ligand) the signal of singly charged [L + Na]$^+$ ion was the most intense of the signals of the spectrum. In addition to the previously mentioned oxyanion adduct series the [L + H]$^+$, [L − H + Cd]$^+$, [2L + H]$^+$, [2L + Na]$^+$, [2L − H + Cd]$^+$, and [2L + Cl − 2H + 2Cd]$^+$-ions were also detected. At a ratio of 1:1 the relative intensities of different ions remained almost the same, even though the percentage of the [2L + H]$^+$ ion was significantly higher than before. Excess cadmium caused disappearance of the signal of the [L + Na]$^+$ ion. The [L + H]$^+$ and [2L + H]$^+$ ions were still observed, but the intensities of the signals of the proton adducts were lower than those of the oxyanion and cadmium adducts. The most intense of the ligand-containing signals was due to the [L + 2ClO + H + 2Na]$^+$ ion where in fact two sodium hypochlorite molecules are bound non-covalently to the protonated ligand. The signal of the previously mentioned cadmium-containing monomeric adduct appeared almost as intense. Of the cad-

mium adducts the signal of the $[2L - H + Cd]^+$ ion was also clearly visible.

In order to examine the cadmium binding ability of the ligand at different pH conditions, measurements of the 1:1 complex in acidic and alkaline solutions were also performed. At slightly acidic conditions the most intense signal of the ligand-containing ions was given by $[L + H]^+$. Also signals given by the $[L + Na]^+$, $[L - H + Cd]^+$, $[2L + H]^+$, $[2L + K]^+$, $[2L - H + Cd]^+$, and $[2L + Cl - 2H + 2Cd]^+$ ions and the oxyanion adduct series were clearly detectable. At slightly alkaline conditions the signal of the $[L + Na]^+$ ion prevailed together with the signals of the $[L + H]^+$, $[2L + H]^+$ and $[2L + Na]^+$ ions. The cadmium adducts were not detected at all.

These findings clearly show that excess cadmium is required for its binding with the ligand. The result differs from the one obtained by NMR measurements in which the binding occurred between the molar ratios of 1:2 and 1:1 (cadmium:ligand), respectively. The difference between the results obtained by these two techniques depends most probably on the differences in the sensitivities and measurement conditions, a general feature of ESI-MS being alkali metal adduct formation. The mass spectrometric measurements clearly reveal that sodium and cadmium cations compete for the active site of the ligand. When the concentration of cadmium is low, complexes with sodium are formed more readily, that is to say, the sodium affinity of the ligand is higher than its cadmium affinity. Excess cadmium is required for displacing the sodium ions entirely with cadmium being the main complex-forming metal ion. The results also showed the ligand to have a relatively high proton and sodium affinity because protons and sodium ions were able to compete succesfully with cadmium ion.

MO calculations

As mentioned above the final optimization of the structure of the ligand and its cadmium complexes have been done at the semiempirical PM3 level of theory. In the energetically most favourable structure of the ligand with its rigid steroidal skeleton and flexible side chain, the tryptophan moiety has folded close to the deoxycholic acid moiety and the structure is stabilized by an intramolecular hydrogen bond between 12α-O*H* and Trp-5-*H* (2.70 Å). An additional hydrogen bond between Trp-CO*O*H and Trp-9-C*H* is also visible. The calculations suggest that if the concentration of cadmium is low, a 1:2 (cadmium:ligand) adduct is likely to form. When excess cadmium is added, the formation of the 1:1 adduct is favoured. The mass spectrometric measurements reveal that dimerization really occurs. However, the intensity of the signals of the monomeric $[L - H + Cd]^+$ and dimeric $[2L - H + Cd]^+$ cadmium adducts are almost the same at molar ratios of 1:2 and 1:1. This could be caused by solvent or water molecules which are also able to bind to the ligand lowering the total energy of the system so that dimerization is not necessary for the minimum energy structure. Further, the calculations do not take into account the other cations or anions present under the measurement conditions, that can interact with the ligand forming adducts containing two

or more ions. In the optimized 1:1 adduct the doubly coordinated Cd^{2+}-cation has bound to the C*O*OH (1.77 Å) and *N*H (2.20 Å) of the tryptophan moiety when in the case of the 1:2 adduct cadmium binds to the 24 C*O* oxygens of the bile acid moieties as can be seen in Figure 2. Calculations at the *ab initio* HF level using the 3-21G(d) basis set for the substructures of the Cd^{2+}-adducts support the semiempirical results giving similar model structures.

Conclusions

It has been shown that modern multinuclear magnetic resonance and mass spectrometric methods supported by sophisticated theoretical calculations can give a unique picture for the conformational and cation/anion binding characteristics of bile acid-amino acid conjugates. The ^{1}H,^{1}H NOESY measurements revealed that in alkaline solution the *N*-deoxycholyl-L-tryptophan conjugate had a folded conformation which differed clearly from the conformations prevailing in neutral and acidic solutions. According to molecular modeling studies in the minimum energy structure of the ligand the tryptophan moiety in fact had folded close to the deoxycholic acid moiety and the structure was stabilized by an intramolecular hydrogen bond. Further, multinuclear magnetic resonance data clearly showed that Cd^{2+}-ion was bound to the bile acid-amino acid conjugate in neutral and alkaline conditions. ESI-TOF MS measurements revealed that sodium and cadmium ions competed for the active site of the ligand and excess cadmium was needed for displacing sodium ion. Further, ESI-TOF MS spectra showed that various chlorine oxyanions originating from perchlorate anion formed together with cations different adducts with the ligand.

Acknowledgements

References

1. A.P. Davis: *Chem. Soc. Rev.* **22**, 243 (1993).
2. U. Maitra: *Curr. Sci.* **71**, 617 (1996).
3. A.P. Davis, R.P. Bonar-Law, and J.K.M. Sanders: *Receptors Based on Cholic Acid* (Comprehensive Supramolecular Chemistry v. 4, J.L. Atwood, J.E.D. Davies, D.D. MacNicol, and F. Vögtle eds., Elsevier (1996), p. 257.
4. M. Miyata and K. Sada: *Deoxycholic Acid and Related Hosts* (Comprehensive Supramolecular Chemistry v. 6, J.L. Atwood, J.E.D. Davies, D.D. MacNicol, and F. Vögtle (eds.)), Elsevier (1996), p. 147.
5. Y. Li and J.R. Dias: *Chem. Rev.* **97**, 283 (1997).
6. P. Wallimann, T. Marti, A. Fürer, and F. Diederich: *Chem. Rev.* **97**, 1567 (1997).
7. A.P. Davis and R.S. Wareham: *Angew. Chem., Int. Ed.* **38**, 2978 (1999).
8. J. Tamminen and E. Kolehmainen: *Molecules* **6**, 21 (2000).

9. A.F. Hofmann: *Ital. J. Gastroenterol.* **27**, 106 (1995).

10. F. Berlati, G. Ceschel, C. Clerici, R. Pellicciari, A. Roda, and C. Ronchi: WO 9,400,126 (1994).

11. B.A. Marples and R.J. Stretton: WO 9,013,298 (1990).

12. A. Enhsen, W. Kramer, and G. Wess: *Drug Discovery Today* **3**, 409 (1998).

13. M.R. Ruff, J.M. Hill, L.D. Kwart, and C.B. Pert: US 5,446,026 (1995).

14. C. Li, A.S. Peters, E.L. Meredith, G.W. Allman, and P.B. Savage: *J. Am. Chem. Soc.* **120**, 2961 (1998).

15. C. Li, L.P. Budge, C.D. Driscoll, B.M. Willardson, G.W. Allman, and P.S. Savage: *J. Am. Chem. Soc.* **121**, 931 (1999).

16. E. Campazzi, M. Cattabriga, L. Marvelli, A. Marchi, R. Rossi, M.R. Pieragnoli, and M. Fogagnolo: *Inorg. Chim. Acta* **286**, 46 (1999).

17. Bondi and Mueller: *Z. Physiol. Chem.* **47**, 499 (1906).

18. F. Cortese and L. Bauman: *J. Am. Chem. Soc.* **57**, 1393 (1935).

19. A. Norman: *Ark. Kemi* **8**, 331 (1955).

20. A.F. Hofmann: *Acta. Chem. Scand.* **17**, 173 (1963).

21. A.M. Ballatore, C.F. Beckner, R.M. Caprioli, N.E. Hoffman, and J.G. Liehr: *Steroids* **41**, 197 (1983).

22. S.M. Huijghebaert and A.F. Hofmann: *Gastroenterology* **90**, 306 (1986).

23. A. Roda, C. Cerrè, A.C. Manetta, G. Cainelli, A. Umani-Ronchi, and M. Panunzio: *J. Med. Chem.* **39**, 2270 (1996).

24. D. Albert and M. Feigel: *Tetrahedron Lett.* **35**, 565 (1994).

25. D. Albert and M. Feigel: *Helv. Chim. Acta* **80**, 2168 (1997).

26. D. Albert, M. Feigel, J. Benet-Buchholtz, and R. Boese: *Angew. Chem.*, *Int. Ed.* **37**, 2727 (1998).

27. K.J. Willis, A.G. Szabo, and D.T. Kracjarski: *Chem. Phys. Lett.* **182**, 614 (1991).

28. X. Cao and G. Fischer: *J. Phys. Chem. A* **103**, 9995 (1999).

29. B. Dezube, C.M. Dobson, and C.F. Teague: *J. Chem. Soc.*, *Perkin Trans. 2* 730 (1981).

30. K. Aizawa, S. Ohhata, H. Nishie, A. Ohsaka, K. Kato, K. Matsushita, and K. Hioka: *Biochem. Biophys. Res. Commun.* **146**, 791 (1987).

31. S.L. De Wall, E.S. Meadows, L.J. Barbour, and G.W. Gokel: *J. Am. Chem. Soc.* **121**, 5613 (1999).

32. M. Saraswathi and J.M. Miller: *Rapid Commun. Mass Spectrom.* **10**, 1706 (1996).

33. H. Lavanant and Y. Hoppilliard: *J. Mass Spectrom.* **32**, 1037 (1997).

34. C.L. Gatlin and F. Tureček: *J. Mass Spectrom.* **35**, 172 (2000).

35. P.K. Sharma, A.K. Sen, and S.N. Dubey: *Indian J. Chem.*, *Sect. A: Inorg.*, *Bio-inorg.*, *Phys.*, *Theor. Anal. Chem.* **33a**, 1031 (1994).

36. E. Giglio: *Inclusion Compounds* (Inclusion Compounds, Vol. 2: Structural Aspect of Inclusion Compounds formed by Organic Hosts Lattices v. 2, J.L. Atwood, J.E.D. Davies, and D.D. MacNicol (eds.)), Academic Press (1984), p. 207.

37. P. Granger: *Groups 11 and 12*, *Copper to Mercury* (Transition Metal Nuclear Magnetic Resonance: Studies in Inorganic Chemistry v. 13, P.S. Pregosin (ed.)), Elsevier (1991), p. 293.

38. A.M. Bellini, M.P. Quaglio, M. Guarneri, and G. Cavezzini: *Eur. J. Med. Chem.- Chim. Ther.* **18**, 191 (1983).

39. B. Dayal, J. Bhojawala, K.R. Rapole, B.N. Pramanik, N.H. Ertel, S. Shefer, and G. Salen: *Bioorg. Med. Chem.* **4**, 885 (1996).

40. A. Bax, R.H. Griffey, and B.L. Hawkins: *J. Magn. Reson.* **55**, 301 (1983).

41. A. Bax and S. Subramanian: *J. Magn. Reson.* **67**, 565 (1986).

42. A. Bax and M.F. Summers: *J. Am. Chem. Soc.* **108**, 2093 (1986).

43. M. Rance, O.W. Sørensen, G. Bodenhausen, G. Wagner, R.R. Ernst, and K. Wüthrich: *Biochem. Biophys. Res. Commun.* **117**, 479 (1984).

44. A. Derome and M. Williamson: *J. Magn. Reson.* **88**, 117 (1990).

45. J. Jeener, B.H. Meier, P. Bachmann, and R.R. Ernst: *J. Chem. Phys.* **69**, 4546 (1979).

46. G. Wagner and K. Wüthrich: *J. Mol. Biol.* **155**, 347 (1982).

47. R.E. Hurd and B.K. John: *J. Magn. Reson.* **91**, 648 (1991).

48. J.J.P. Stewart: *J. Comp. Chem.* **10**, 209 (1989).

49. SPARTAN, Version 5.0.2 (Wavefunction Inc., Irvine, CA, 1991–7).

50. M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, GAUSSIAN 98, Revision A.6, Gaussian Inc., Pittsburgh PA, 1998.

51. J.W. Blunt and J.B. Stothers: *Org. Magn. Reson.* **9**, 439 (1977).

52. M.S. Morales-Ríos, J. Espiñeira, and P. Joseph-Nathan: *Magn. Reson. Chem.* **25**, 377 (1987).

53. S. Berger, S. Braun, and H.-O. Kalinowski: *NMR Spectroscopy of the Non-Metallic Elements*, John Wiley & Sons Ltd. (1997), p. 158.

54. J. Tamminen, E. Kolehmainen, J. Linnanto, P. Vainiotalo, S. Vuorikoski, and R. Kauppinen: *J. Incl. Phenom. Macrocycl. Chem.* **37**, 121 (2000).

55. J.M.J. Nuutinen, M. Purmonen, J. Ratilainen, K. Rissanen, and P. Vainiotalo: *Rapid Commun. Mass Spectrom.* **15**, 1374 (2001).