DOI: 10.1002/ejoc.200701212

Structural Features of the (+)-Yatakemycin/d(GACTAATTGAC)-(GTCAATTAGTC) Complex – Quantum Mechanical Calculation of NMR Parameters as a Tool for the Characterization of Ligand/DNA Interactions

Simone Di Micco,[a] Dale L. Boger,[b] Raffaele Riccio,[a] and Giuseppe Bifulco*[a]

Dedicated to the memory of our dear colleague and friend Luigi Gomez Paloma

Keywords: Minor groove binders / NMR spectroscopy / Calculation of NMR parameters

(+)-Yatakemycin is a new and potent member of a class of natural antitumor compounds that derive their biological activities from specific alkylation of adenine residues in the minor grooves of AT-rich tracts. We have analyzed the covalent complex formed between (+)-yatakemycin and the d(GACTAATTGAC)-(GTCAATTAGTC) duplex, and have established that the ligand covalently binds to the A5 residue. For this purpose we used a hybrid approach based on 2D-NMR spectroscopy and quantum mechanical (QM) calculations of ^{1}H chemical shifts at the DFT/MPW1PW91 level. In this study we also show that the calculation of NMR parameters can be a useful tool for the structural characterization of ligand–receptor interactions.

Introduction

Small molecules that interact with DNA are of exceptional importance in drug discovery, because they can be employed in chemotherapy as antitumor, antiviral, and antifungal compounds. These molecules exert cytotoxic activity, interfering with factors that are necessary for replication, transcription, and repair mechanisms of damage. Minor groove binders can be potential drugs with high degree of target selectivity, because interactions with individual base pairs along the groove floor and the deoxyribose groove walls provide cooperative binding contacts and extended recognition specificity. Moreover, for a given dose, a minor groove binder with increased selectivity should produce a greater pharmacological response than a non-selective binder.

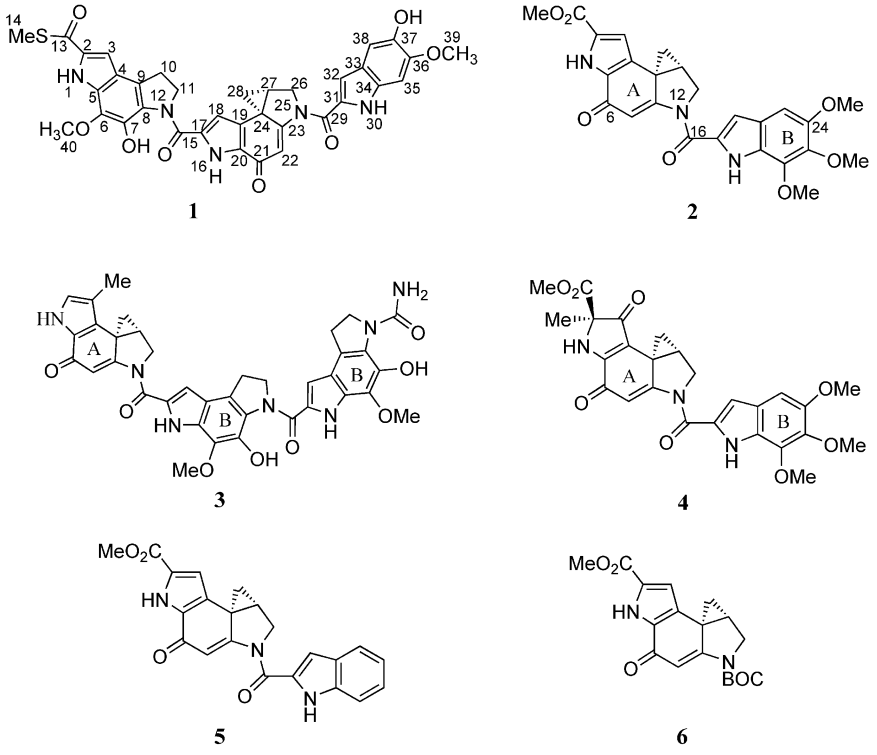
(+)-Yatakemycin[1–4] (**1**), isolated from the culture broth of *Streptomyces* sp. TP-A0356,[5] is the newest and most potent member of a class of natural antitumor compounds[1–4] that includes CC-1065 (**3**), and duocarmycins A (DA, **4**) and SA (DSA, **2**) shown in Scheme 1.[6]

These natural products derive their biological activities[1,7,8] from specific alkylation of AT-rich regions of DNA. These compounds display in vitro IC_{50} values against tumor cell lines at the pM level, and certain derivatives exhibit antitumor selectivity. Duocarmycin SA (DSA), for example, is able to cause apoptosis (programmed cell death) in tumor cells at a dose (<100 pM) well below that required for cell death by necrosis in non-sensitive cell lines.[9] Moreover, DSA displays better pharmacological properties[7,9–12] than the natural product (+)-CC-1065,[8] which shows high cytotoxicity in vitro and good antitumor activity in vivo, but induces a delayed fatal toxicity.[13] The unnatural enantiomer (–)-CC-1065 has the same activity as its natural enantiomer but it does not present the toxicity mentioned above associated with (+)-CC-1065 derivatives[14,15] and duocarmycin analogues.[16] In addition to their potential for direct use as anticancer agents, duocarmycins have also been proposed as conjugates for antibody-directed chemotherapy.[17,18]

All members of this class of compounds have a common structural portion (see, for instance, structure **2**): an alkylation subunit (indicated as A in Scheme 1) containing the reactive cyclopropyl ring that is attacked by adenine N3 upon binding to the DNA, and a binding subunit (indicated by B in Scheme 1) that contributes to the noncovalent binding to specific sequences of DNA duplex, establishing weak interactions (Van der Waals, hydrogen bonds) in the minor grooves of specific DNA sequences. Both subunits are rigid heteroaromatic polycycles, coplanar in the free state, pres-

[a] Dipartimento di Scienze Farmaceutiche, Università degli studi di Salerno,
Via Ponte don Melillo, 84084 Fisciano (SA), Italy
Fax: +39-089969602
E-mail: bifulco@unisa.it
[b] The Scripps Research Institute, Department of Chemistry, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Supporting information for this article is available on the WWW under http://www.eurjoc.org or from the author.



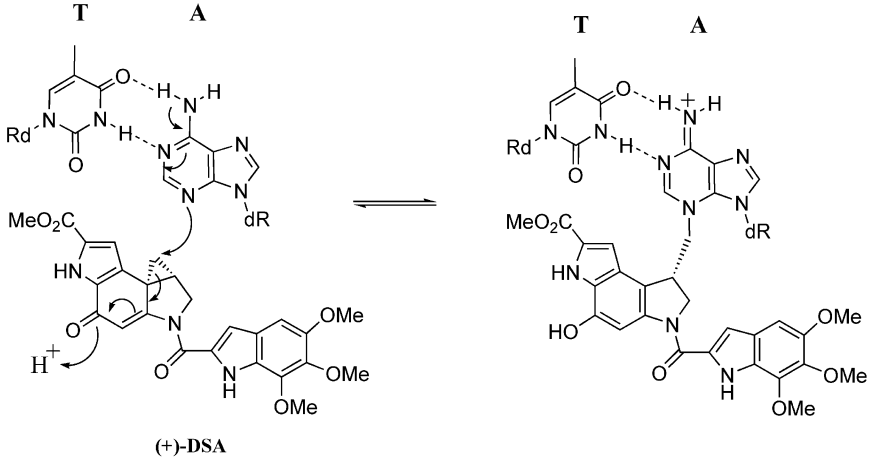
Scheme 1. Molecular structures of (+)-yatakemycin (**1**), DSA (**2**), CC-1065 (**3**), DA (**4**), (+)-DSI (**5**) and NBOC-DSA (**6**).

enting extended conjugation of π electrons of the A and B portions bound through a central carbonyl moiety, conjugated to the cyclohexadienone (left-hand side) as a function of the nitrogen lone pair of the vinylogous amide.

The DNA alkylation reaction proceeds through the nucleophilic attack of the nitrogen in the 3-position (N3) of an adenine residue on the cyclopropyl ring, leading to the transformation of the unsaturated ketone in the alkylating portion into a phenol (Scheme 2).

Extensive efforts have been devoted to improving understanding of the binding preferences of these ligands for the minor grooves of AT-rich tracts of DNA, the details of the alkylation reaction with duplex DNA, and the source of catalysis.[9,17–20] The alkylation reaction is a reversible, stereoelectronically controlled addition by N3 of adenine to the least substituted carbon of the cyclopropane ring.

The unusual stability of the cyclopropane ring when exposed to small nucleophiles[19,20] and its good reactivity with duplex-DNA have suggested a catalytic role of DNA in the reaction.[19–22] Structural analyses of the 1:1 complexes formed between (+)-DSA,[23] (+)-DA,[24] and (+)-DSI[25] and the DNA duplex d(GACTAATTGAC)–(GTCAATTAGTC) have supported the shape-dependent catalysis model, which involves in situ activation[26–30] in-



Scheme 2. Alkylation reaction mechanism. The nucleophilic attack of adenine N3 on the least substituted carbon of the cyclopropane moiety in (+)-DSA (**2**) is shown.

duced by a conformational change that the ligand undergoes in consequence of recognition and binding to DNA. The small molecule in the free state thus presents an extended conjugation, which is destroyed upon binding to the biological target.[31] The angle twisting destabilizes the small molecule, which as a result becomes more reactive to nucleophilic attack by adenine N3.

It has been shown that greater perturbations from the planar conformation cause higher alkylation efficiency.[31] For instance, DSA, presenting a larger $\chi 1$ value in the bound conformation than the other two analogues, shows a higher reaction efficiency than (+)-DA and (+)-DSI (Scheme 1).[24]

Subsequent studies on sandwiched analogues such as (+)- and (–)-CDPI-DSA-CDPI[1–4,32–33] have made great contributions to understanding of the shape-dependent catalysis, because these compounds show faster alkylation rates. The steric bulk of the sandwiched analogues in fact forces such a compound into greater perturbation from the free planar conformation to allow penetration into the minor groove.

The alkylation reaction has proven to be selective for AT-rich DNA regions. This selectivity is determined by noncovalent recognition of DNA by the binding subunit B, as witnessed by NMR structural analysis on the complexes.[23–25]

It is well known that most enzymes are proteins, while there are also many examples of catalytic roles of RNA. This in situ activation represents a rare case of catalysis by DNA; in fact the distortion of the ligand conformation induced by DNA recognition lowers the activation energy, and the initial noncovalent interactions put the reacting species in close proximity. The AT-rich region selectivity and the in situ activation shown by this class of compounds are very attractive properties that lay the foundations to achieving selective targeting in anticancer therapy.

(+)-Yatakemycin presents two binding subunits (B), flanking the alkylation subunit (B–A–B), thus defining a sandwich structure. Its alkylation subunit is identical to that of DSA. Its right-hand binding subunit is similar to the binding portion of DA (**4**) and DSA (**2**), whereas its left-hand binding subunit and central alkylation subunit are identical to the central and right-hand subunits, respectively, of CC-1065 (**3**).

(+)-Yatakemycin, like the preceding sandwiched derivatives,[4,32,33] shows improved alkylating properties,[1] displaying faster alkylation rate and higher alkylation selectivity, and consequently more potent cytotoxic activity than preceding analogues (Table 1).

These features of (+)-yatakemycin moved us to investigate the covalent complex formed by this sandwiched compound and the d(GACTAATTGAC)-(GTCAATTAGTC) duplex. In a recent study, Boger and co-workers presented data showing that (+)-yatakemycin and DSA are able to alkylate nucleosomal DNA[7] with an efficiency and selectivity nearly identical to those observed with protein-free DNA. In particular, these alkylating agents are able to recognize DNA tracts that face the proteins in the core particle of the nucleosome and that seem not to be accessible to drugs. These features are very important because, in addition to supporting a dynamic model of the nucleosome, it validates the use of free DNA as a model to predict the binding site and to analyze the structure–activity relationships of our ligand. We have focused our attention on determining which base in the DNA is involved in the covalent bond with the ligand. For these purposes we have analyzed the covalent complex by a hybrid approach based on 2D-NMR spectroscopy and quantum mechanical (QM) calculations of ^{1}H chemical shifts. In this study we report on how the calculation of NMR parameters could be a useful tool for structural characterizations of ligand–receptor interactions.

Table 1. Cytotoxic activities (L1210 mouse leukemia cell line) of compounds **1**–**4**, expressed as IC_{50} values.

Compound	IC_{50} (pM)
(+)-Yatakemycin	3
(+)-Duocarmycin SA	10
(+)-CC-1065	20
(+)-Duocarmycin A	200

Results and Discussion

^{1}H NMR Assignment and Characterization of Bound DNA Duplex

Following the lines of an established strategy,[34–40] we performed a sequence-specific assignment of the ^{1}H resonances of the duplex d(GACTAATTGAC)-(GTCAATTAGTC) in complexation with (+)-yatakemycin. Interproton distances are specified by Wüthrich's shorthand notation.[35] Sequential resonance assignments were made through the $d_i(6,8;1',2',2'')$, $d_s(6,8;1',2',2')$ connectivities in the NOESY spectra (t_{mix} = 200 ms), and the intraresidue 1′-2′-2′′ network was also confirmed by analysis of 2Q spectrum. The complete deoxyribose sugar spin system resonances were identified through scalar connectivities in the 2Q, PE-COSY, and TOCSY spectra, and were confirmed by inspection of the NOESY spectra (Table 2). Because of overlapping chemical shifts it was not possible to assign all 5′ and 5′′ proton resonances. All cytosine 5H–6H and thymine 5Me–6H resonances were identified by scalar connectivities in 2Q, PE-COSY and TOCSY spectra. All exchangeable protons and adenosine 2H were assigned by means of a NOESY (t_{mix} = 200 ms) spectrum acquired in H_2O. Table 2 reports the assignment of the ^{1}H resonances of the complex obtained as outlined above.

Identification of the (+)-Yatakemycin Binding Site

In order to detect which DNA bases made up the binding site in the complex formed from d(GACTAATTGAC)-(GTCAATTAGTC) (Figure 1) and (+)-yatakemycin (**1**), and also which adenine residue was involved in the covalent

Table 2. ^{1}H chemical shifts of the complex formed between d(GACTAATTGAC)-(GTCAATTAGTC) and (+)-yatakemycin (**1**) at pH 7.0, 300 K.

d(GACTAATTGAC)-(GTCAATTAGTC)										
Residue	N1–H, N3–H	NH_2	2-H, 5-H, 5-CH_3	6-H, 8-H	1′-H	2′-H	2′′-H	3′-H	4′-H	5′-H, 5′′-H
G1			–	7.79	5.54	2.41	2.67	4.78	4.13	
A2	–		8.05	8.24	6.32	2.73	2.92	5.01	4.46	
C3	–	6.81, 8.00	5.19	7.24	5.86	1.96	2.56	4.65	4.29	
T4	14.08	–	1.53	7.17	5.59	1.62	1.73	4.56	3.58	3.37
A5	–	–	6.68	8.49	5.03	2.59	2.89	4.82	3.29	
A6	–	–	7.22	8.15	5.32	2.33	2.44	4.07	3.59	
T7	12.88	–	1.15	6.97	5.60	1.84	2.14	4.57	3.40	
T8	13.51	–	1.44	6.84	5.70	1.56	1.82	4.54	3.36	
G9	12.51		–	7.82	5.44	2.67	2.67	4.96	4.33	
A10	–	–	7.86	8.08	6.19	2.59	2.80	4.94	4.36	4.10
C11	–	6.67, 8.09	5.28	7.25	5.97	1.98	2.04	4.39	3.92	
G12			–	7.88	5.96	2.60	2.72	4.73	3.91	
T13	13.66	–	1.24	7.35	6.02	2.08	2.41	4.81	4.19	
C14	–	6.60, 8.45	5.58	7.36	5.57	1.77	2.10	5.27	3.57	
A15	–		7.13	8.16	5.62	2.73	2.73	4.95		
A16	–		7.67	8.13	6.19	2.56	2.84	4.97	4.16	
T17	11.50	–	1.32	7.00	5.85	1.75	2.47	6.27	4.15	
T18	12.75	–	1.60	7.28	5.52	2.04	2.49	4.50	3.37	4.55
A19	–		7.27	8.15	5.84	2.45	2.54	4.33	4.15	
G20	12.83		–	7.40	5.66	2.13	2.40	4.57	5.13	
T21	14.00	–	0.85	6.81	5.53	1.76	2.27	4.40	3.82	
C22	–	8.25, 6.90	5.53	7.45	6.36	2.19	2.18	4.53	4.04	
(+)-Yatakemycin										
N1–H	3-H	N16–H	26-H	27-H	28-H	N30–H	32-H	35-H	38-H	39-H
11.93	7.47	11.36	4.26	4.65	3.83	10.99	7.71	7.14	7.25	3.71

bond, we first observed the variations of the NMR parameters of the receptor in the free and bound states, a strategy adopted for the structural analysis of covalent complexes involving the preceding analogues of **1**. The ^{1}H chemical shifts of DNA bound to **1** were therefore compared with the values relative to free nucleic acid, the resonances of which were already available.

Figure 1. Schematic representation of the DNA duplex. Sequence and numbering of the 11mer DNA duplex are indicated.

The comparison was performed by subtracting the values of the free duplex from the ^{1}H chemical shifts of bound DNA, to obtain the $\Delta(\delta_{\text{DNA bound}} - \delta_{\text{DNA free}})$ values depicted in Figure 2. As shown in Figure 2, the resonances of the A5 and A6 residues show remarkable shifts, of about 1 ppm, in their 1′-proton resonances. The 2′′-proton of T4 is shifted downfield by –0.69 ppm. The 3′-proton of the T18 residue shows a substantial perturbation, moving upfield by 1.5 ppm. The 3′-protons of T7 and G20 also vary substantially, by –0.69 ppm and –0.68 ppm, respectively. Notable perturbation can also be observed for the 4′-protons of T4, A6, T7, T8 and T18, which are shifted downfield in a range of 0.6–1.0 ppm, while G20 4′ is shifted upfield by ca. 0.8 ppm. The largest 4′-proton perturbation is shown by the A5 residue, which is shifted downfield by –1.1 ppm. Such perturbations are probably due to the ring current of (+)-yatakemycin.

The $\Delta\delta$ values for the 5′- and 5′′-protons were not determined, since these resonances could not be assigned in the free duplex due to overlap. The most dramatic change involved the imino proton resonance of T17 (–2.08 ppm), while other thymine NH resonances were also evidently shifted: T4 (δ = 0.57 ppm), T7 (–0.71 ppm), and T18 (–0.69 ppm). The observed $\Delta\delta$ values map out the binding site of the ligand to base pairs C3·G20 to T8·A15 (Figure 1).

Biological assays showed that (+)-yatakemycin selectively alkylated AT-rich DNA sequences (no guanine *N*3 or *N*7 addition)[1,8] and produced one adduct. Moreover, biological screenings showed that, statistically, the first base (in both directions) that flanked the alkylated base had to be an adenine or thymine, with this preference order: 5′-AAA-3′ > 5′-AAT-3′≈ 5′-TAA-3′ > 5′-TAT-3′.[1,8] In addition, there was a very strong preference, but not absolute requirement, for both the second 5′- and 3′-bases to be A or T.

The sandwiched compound would hence bind the adenine in the middle of the AT recognition sequence with the two binding subunits extending in both directions (3′ and 5′) from the alkylation point; this is unlike the cases of CC-1065 and duocarmycins A and SA, each of which bound an adenine at the end of the AT recognition sequence, with the binding subunit extending in the 3′→5′ direction.

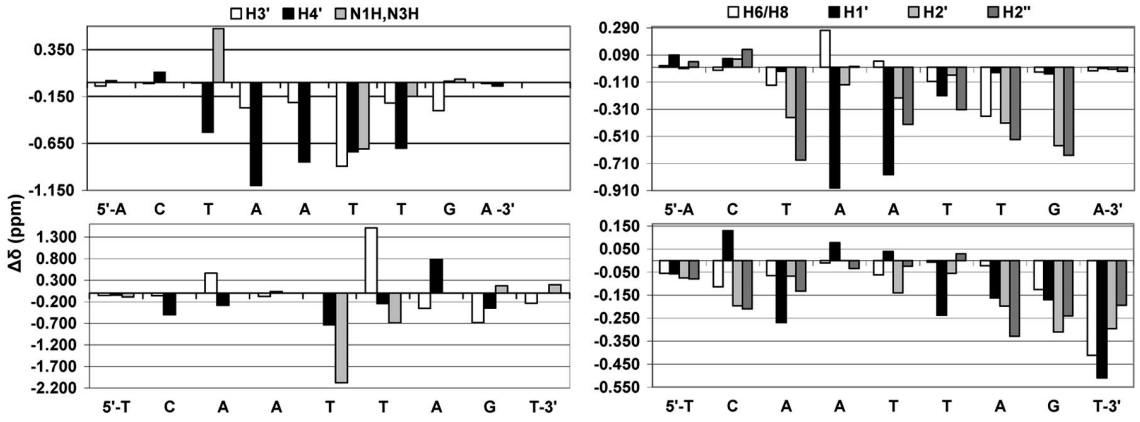


Figure 2. Chemical shift differences ($\Delta\delta = \delta_{\text{bound DNA}} - \delta_{\text{free DNA}}$) plots for the sugar and base protons of the DNA duplex. The base-pairs G1–C22 and C11–G12 are not considered because they do not show resonance shifts.

The results of the NMR inspection described above were in agreement with the results of the biological assays, because A5 and A6 were both in the middle of the AT sequence (Figure 1) responsible for the recognition of the ligand. Both A5 and A6 showed evident chemical shift perturbations, indicating the involvement of at least one of the two adenines in the alkylation reaction; nevertheless, the available experimental data were not sufficient to assess the alkylation point safely.

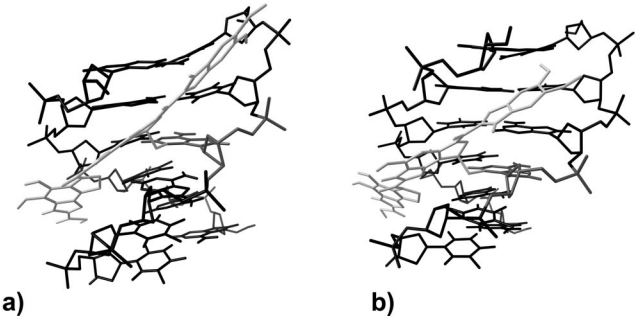


Figure 3. 3D models of (+)-yatakemycin covalently bound to A5 (a) and to A6 (b). For the nucleic acid the two models contain only the base-pairs C3·G20 to T8·A15. The DNA and ligand are represented in tube style. The DNA chains are depicted in black except for the adenine (in dark gray) covalently bound to ligand, while **1** is colored in light gray. The Figure highlights the covalent bond with adenine N3 and how the ligand assumes a curved conformation in the minor groove.

Identification of the Alkylated Adenine

In order to understand which adenine was responsible for the nucleophilic attack on the cyclopropane ring of the ligand, we integrated the NMR experimental data with the results of GIAO (gauge including atomic orbital)[41,42] calculations of NMR parameters, at DFT (density functional theory) theoretical level.[43,44] So far, the use of ^{1}H and ^{13}C chemical shifts calculations has been shown to be efficient in the elucidation of organic molecules,[45–50] in the interpretation of polymer spectra,[50] and in the conformational analysis of peptides,[51–53] oligosaccharides,[54] and calixarenes.[55] Here we propose a QM-NMR integrated approach for the characterization of a larger and more complex system, involving a ligand–receptor system and in particular a nucleic acid structure. We thus built two 3D models in which **1** was covalently bound to A5 (model A, Figure 3, a) and to A6 (model B, Figure 3, b), starting from the NMR structure of the DNA duplex and docking the ligand in the minor groove.

In order to perform quantum mechanical calculations with reasonable computation times and with a reliable model, we considered only six base pairs of the receptor (Figure 3). For these reduced models we calculated ^{1}H chemical shifts at the DFT/MPW1PW91 theoretical level with the 6-31G(d,p) basis set. Such a combination has been shown to provide satisfactory results with reasonable computational effort for calculations of NMR parameters of medium to large organic molecules.[48,56,57] The calculated values for A5 and A6 were compared with the experimentally determined ones, to verify which model reproduced the experimental data in the best way and to determine the point of alkylation. Besides the A5 and A6 resonances, we also considered the values relating to the imino proton of T17 and 3′H of T18, the chemical shift values of which showed dramatic variation (–2.08 ppm and 1.5 ppm, respectively; see Figure 2) upon binding of (+)-yatakemicin.

Table 3 reports the chemical shift differences ($\Delta\delta$), the Mean Absolute Errors (MAEs), and the Correlation coefficients (R) obtained for the significant protons under consideration for both models.

Analysis of these data suggests that, while a good agreement is observed between the calculated ^{1}H chemical shifts of Model A (alkylation at A5) and the experimental data, Model B (alkylation at A6) presents large differences between calculated and experimentally measured ^{1}H resonances on almost all protons considered. In particular, the resonances of the A6 residue (Model B), when bound covalently to structure **1**, mostly diverged from the experimen-

Table 3. Significant calculated ^{1}H NMR [MPW1PW91/6-31G(d,p) level] chemical shifts (δ values) for model A and model B, and the corresponding experimentally measured ^{1}H NMR data for DNA bound to **1**; $\Delta(\delta_{calcd} - \delta_{exp})$ is the difference between calculated and experimental chemical shifts.

Residue	H	Calcd. ^{1}H chemical shifts Model A	Model B	δ_{exp}	$\Delta(\delta_{calcd} - \delta_{exp})$ Model A	Model B	MAE[a] Model A	Model B	R[b] Model A	Model B
A5	8	8.68	8.06	8.49	0.19	–0.43	**0.31**	**0.58**	0.9932	0.9780
	2′	2.63	2.29	2.59	0.04	–0.30				
	2′′	2.43	1.93	2.89	–0.46	–0.95				
	1′	5.70	5.78	5.03	0.67	0.74				
A6	8	8.26	9.03	8.15	0.10	0.87				
	2′	1.89	1.79	2.33	–0.44	–0.55				
	2′′	2.83	2.66	2.44	0.39	0.22				
	1′	5.87	6.10	5.32	0.55	0.78				
T17	N3–H	11.43	10.80	11.50	–0.07	–0.70				
T18	3′	4.68	4.79	4.50	0.18	0.28				

[a] Mean absolute errors (MAEs) found for ^{1}H NMR chemical shifts of Models A and B vs. ^{1}H experimental values: $MAE = \sum[|(\delta_{exp} - \delta_{calcd})|]/n$. [b] Correlation coefficient obtained by linear fitting of calculated (δ_{calcd}) against experimentally measured (δ_{exp}) ^{1}H NMR chemical shifts.

tally ascertained values, with the exception of the 2′′-proton, suggesting that the ligand is not covalently linked to N3 of this base. In addition, the calculated resonances for the A5 residue in Model B did not reproduce the experimental data, indicating that A5 may be bound in the experimental complex. The calculated chemical shift values for model A, in contrast, reasonably reproduced the experimental values, showing good agreement with the experimentally measured resonances for both adenines A5 and A6. Model A differed from experimental data by 0.18 ppm (Table 3) for 3′H of the T18 residue and by –0.07 ppm for *N*3H of the T17 nucleotide, reproducing the magnetic environments of these two diagnostic protons well. In contrast, in the case of Model B, whereas a fair agreement is shown for T18 3′H ($\Delta\delta$ of 0.28 ppm), the imino proton of T17 displays a significant difference with the experimentally determined value ($\Delta\delta$ of –0.70 ppm; see Table 3).

From a quantitative point of view, the MAE of 0.31 calculated for Model A (Table 2), in comparison with 0.58 for model B, and the correlation coefficients of 0.9932 vs. 0.9780 were also consistent with alkylation at A5. Residual discrepancies between experimentally and theoretically determined ^{1}H chemical shifts for Model A could well be due to the absence of solvent in the calculations or to the low level of theory, chosen in view of the complexity of the system.

These results point towards a nucleophilic attack by *N*3 of adenine 5, and in the mean time suggest that a combined approach based on NMR and QM calculations can be a useful tool in the investigation of ligand–receptor interactions.

An analysis of key NOE contacts between the duplex and (+)-yatakemycin was necessary in order to confirm the reliability of our methodology (Figure S1, Supporting Information). In particular, the H-28 resonance of the ligand (δ = 3.83 ppm) shows dipolar effects with H-1′ (δ = 5.03 ppm) of the A5 residue and H-8 (δ = 8.15 ppm) of the A6 residue (Figure S2, Supporting Information). Such contacts are consistent with model A and are not compatible with model B.

The distances, in fact, of H-28 from H-1′ and H-8 are 1.96 and 3.46 Å, respectively (Figure 4), whereas those in model B are 6.15 and 5.02 Å.

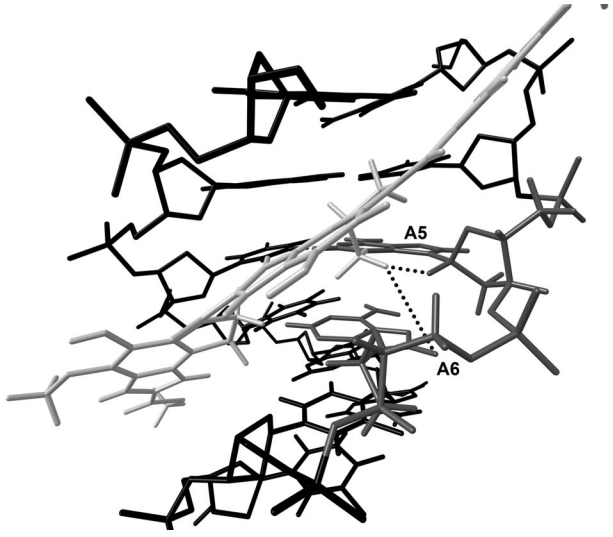


Figure 4. 3D model (duplex DNA base-pair region C3·G20 to T8·A15) of (+)-yatakemycin covalently bound to A5. The black broken lines show the NOE contacts observed experimentally and well reproduced by Model A.

Model A thus better reproduces the experimentally determined values, supporting the results obtained independently by our quantum chemical methodology.

Conclusions

We have analyzed the covalent complex formed between d(GACTAATTGAC)-(GTCAATTAGTC) and (+)-yatakemycin by an approach based on NMR spectroscopic data integrated with quantum mechanical calculations of the chemical shifts. The analysis of ^{1}H resonance shifts between the free and the bound state of DNA has allowed the bind-

ing site to be mapped out, confirming and assessing more accurately what has emerged from biological trials. The receptor-based NMR approach, very useful in previous structural studies on duocarmycins in covalent complexes, give us important information on the binding site but did not allow the adenine residue involved in the nucleophilic attack in cyclopropane ring to be identified, because A5 and A6 experienced comparable chemical shift perturbations between their free and bound DNA forms. Moreover biological experimentation suggested the formation of one covalent adduct, and the NMR spectra did not show a double pattern of signals attributable to the simultaneous presence of two complexes, with A5 and A6 both covalently bound to **1**. We developed our investigation by taking advantage of integration of the experimentally determined data with QM calculation of NMR parameters. Such methods, which have been widely applied to the structural characterization of small and medium-sized molecule, have here been used for a larger and complex system. The calculated ^{1}H chemical shifts have been used to integrate with the experimental NMR spectroscopic data, confirming the binding site of (+)-yatakemycin and suggesting A5 as the residue involved in the covalent bond with the ligand. The NOE contacts of H-28 (δ = 3.83 ppm) with H-1′ (δ = 5.03 ppm) in the A5 residue and with H-8 (δ = 8.15 ppm) in the A6 residue are in agreement with the QM calculation results, supporting the reliability of our methodology. In light of the results obtained, it has been shown that calculation of NMR parameters can be used not only for structure elucidation of small/medium-sized molecules, but also to obtain essential information about ligand–receptor interactions, and how the application of QM calculations of NMR parameters is not limited to systems with only modest numbers of atoms. The QM-based approach, moreover, could be useful in cases in which dipolar couplings between ligand and receptor are uncertain or unavailable; it could also be a useful addition to standard spectroscopic tools in the characterization of drug–DNA complexes. In this pioneering study, the choice of theoretical model was suggested by a compromise between required computation time and reliability of calculated values for reproduction of experimentally measured data. We suggest the use of higher theory levels when CPU time is not a limiting factor.

Experimental Section

Sample Preparation: The purified d(GTCTAATTGAC) and d(GTCAATTAGAC) DNA oligomers were purchased from Sigma Genosys. A typical analytical reaction was carried out by dissolving DNA duplex (200 nmol) in buffer solution (10 mM KCl, 10 mM KH_2PO_4, pH 7.0). The DNA solution was annealed at 353.15 K for five minutes, and then cooled slowly to room temperature. The agent was dissolved in [D_6]DMSO (30.5 μL), and a part of this solution (26.7 μL) was added to the annealed DNA solution. The reaction was conducted at 299.15 K in a Thermomixer (700 rpm). The progress of the reaction was monitored by reversed-phase HPLC over 24 h, with a 0.46 cm × 25 cm (Luna Phenomenex) C18 column. The HPLC conditions were as follows: eluent A, NH_4HCO_2 (pH 7, 50 mM); eluent B, acetonitrile (80%), 20% eluent A; flow rate, 1.0 mL min^{-1}; detection at 260 nm. The covalently bound strand separates from the complementary one and eluted from the HPLC column later. ^{1}H NMR spectroscopy showed that the alkylated duplex did not require any purification. The obtained sample was dissolved in buffered D_2O solution (10 mM KH_2PO_4 and 10 mM KCl, pH 7, 200 μL) and then lyophilized three times to remove residual H_2O. The sample was finally dissolved in buffered D_2O solution (pH 7, 10 mM KH_2PO_4 and 10 mM KCl, 200 μL) and placed in a 3 mm NMR tube (200 μM). The excess ligand precipitated and was left at the bottom of the tube. One NOESY experiment was carried out after the sample had been lyophilized and back-exchanged in H_2O/D_2O (9:1) for the chemical shift assignments of exchangeable protons.

NMR Experiments: NMR experiments were performed on a Bruker DRX 600 spectrometer fitted with a cryoprobe at T = 300 K. All spectra were acquired in the phase-sensitive mode, and the TPPI[58] method was used for quadrature detection in the ω_1 dimension. For the 2D spectra, the compounds were dissolved in D_2O (99.95%, 0.2 mL). The residual solvent signal in experiments acquired in ^2H_2O solution was suppressed by saturation of the resonance during the preparation period and/or mixing time. The 2Q[59] spectrum was executed with a number of 64 scans/t_1, a t_{1max} value of 71.09 ms.

The P.E.COSY[60] spectrum was executed with a number of 128 scans/t_1, a t_{1max} value of 100.36 ms. The TOCSY spectrum was acquired by use of the DIPSI-2 sequence[61] for spin locking with t_{mix} = 120 ms with a number of 128 scans/t_1, a t_{1max} value of 100.36 ms. The NOESY[62] spectrum in D_2O was acquired with a mixing time of 200 ms, a number of 192 scans/t_1, and a t_{1max} value of 39.20 ms. The NOESY spectrum in H_2O/D_2O (9:1) was acquired with a mixing time of 200 ms, a number of 320 scans/t_1, a t_{1max} value of 32.81 ms, with use of a 3–9–19[63,64] pulse sequence to suppress the water ^{1}H signal. All data were processed on a Silicon Graphics Indigo2 workstation with use of UXNMR software.

Computational Details: The NMR structure of the (+)-DSA-d(GACTAATTGAC)-(GTCAATTAGTC) covalent complex (PDB archive code 1DSA), was used as a template to build the homology model A (Figure 3, a) with (+)-yatakemycin bound to A5 (Figure 3, a). Structure **1** was preliminarily minimized at molecular mechanics level, with use of the Macromodel 8.5 software package[65] and the AMBER force field.[66–69] The structure was minimized by use of a Polak–Ribiere Conjugate Gradient (PRCG, 500 000 steps, convergence threshold 0.005 kJ mol^{-1} Å^{-1}). A GB/SA (Generalized Born/surface area)[70] solvent treatment was used, mimicking the presence of H_2O, in the calculations for reducing the artifacts derived from the absence of the solvent. In a second step the entire covalent complex was optimized with use of the same parameters (PRCG, 500 000 steps, convergence threshold 0.005 kJ mol^{-1} Å^{-1}), and with use of the GB/SA[70] solvent model in order to avoid the loss of B-DNA arrangement. Model B (Figure 3, b) was built by shifting the ligand (**1**) from $N3$ of A5 to the nucleophilic nitrogen of A6. On this model we performed an optimization of the whole complex by the same criteria as used for model A. Once built, we calculated the chemical shifts at quantum mechanical (QM) level for the two models. In order to reduce the CPU time of our calculations we cut the DNA including the base-pairs C3·G20 to T8·A15. Moreover, 5′-terminal phosphate groups were substituted with hydroxy groups. ^{1}H chemical shifts were calculated at the MPW1PW91 DFT level by use of the 6-31G(d,p) basis set (Gaussian 03 Software Package).[71]

Supporting Information (see footnote on the first page of this article): NOESY spectrum of the complex formed between (+)-yatakemycin and d(GACTAATTGAC)-(GTCAATTAGTC) complex acquired at 300 K (pH 7) in D_2O solution.

Acknowledgments

The authors are grateful to Maria Giovanna Chini for supporting us with the preparation of figures.

[1] J. P. Parrish, D. B. Kastrinsky, S. E. Wolkenberg, Y. Igarashi, D. L. Boger, *J. Am. Chem. Soc.* **2003**, *125*, 10971–10976.
[2] M. S. Tichenor, D. B. Kastrinsky, D. L. Boger, *J. Am. Chem. Soc.* **2004**, *126*, 8396–8398.
[3] M. S. Tichenor, J. D. Trzupek, D. B. Kastrinsky, F. Shiga, I. Hwang, D. L. Boger, *J. Am. Chem. Soc.* **2006**, *128*, 15683–15696.
[4] D. L. Boger, D. L. Hertzog, B. Bollinger, D. S. Johnson, H. Cai, J. Goldberg, P. Turnbull, *J. Am. Chem. Soc.* **1997**, *119*, 4977–4986.
[5] Y. Igarashi, K. Futamata, T. Fujita, A. Sekine, H. Senda, H. Naoki, T. Furumai, *J. Antibiot.* **2003**, *56*, 107–113.
[6] G. Bifulco, J. A. Smith, W. J. Chazin, L. Gomez-Paloma, *Advances in DNA Sequence-Specific Agents* **2002**, *4*, 47–73.
[7] J. D. Trzupek, J. M. Gottesfeld, D. L. Boger, *Nat. Chem. Biol.* **2006**, *2*, 79–82.
[8] D. L. Boger, D. S. Johnson, *Angew. Chem.* **1996**, *108*, 1542–1580; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1438–1474.
[9] D. L. Boger, D. S. Johnson, W. Wrasidlo, *Bioorg. Med. Chem. Lett.* **1994**, *4*, 631–636.
[10] T. Yasuzawa, Y. Saitoh, M. Ichimura, I. Takahashi, H. Sano, *J. Antibiot.* **1991**, *44*, 445–447.
[11] M. Ichimura, T. Ogawa, S. Katsumata, K. Takahashi, I. Takahashi, H. Nakano, *J. Antibiot.* **1991**, *44*, 1045–1053.
[12] M. Ichimura, T. Ogawa, K. Takahashi, A. Mihara, I. Takahashi, H. Nakano, *Oncol. Res.* **1993**, *5*, 165–171.
[13] V. L. Reynolds, J. P. McGrovren, L. H. Hurley, *J. Antibiot.* **1986**, *39*, 319–334.
[14] D. L. Boger, R. S. Coleman, B. Invergo, S. M. Sakya, T. Ishizachi, A. Munk, H. Zarrinmayeh, P. A. Kitos, S. C. Thompson, *J. Am. Chem. Soc.* **1990**, *112*, 4623–4632.
[15] J. P. McGrovren, G. L. Glarke, P. A. Pratt, T. F. DeKnig, *J. Antibiot.* **1984**, *37*, 63–70.
[16] D. L. Boger, *Advances in Heterocyclic Natural Product Synthesis* **1992**, *2*, 1–188.
[17] D. L. Boger, *Acc. Chem. Res.* **1995**, *28*, 20–29.
[18] D. L. Boger, H. Zarrinmayeh, S. A. Munk, P. A. Kitos, O. Suntornwat, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 1431–1435.
[19] D. L. Boger, R. S. Coleman, B. J. Invergo, S. M. Sakya, T. Ishizaki, S. A. Munk, H. Zarrinmayeh, P. A. Kitos, S. C. Thompson, *J. Am. Chem. Soc.* **1990**, *112*, 4623–4632.
[20] D. L. Boger, S. M. Sakya, M. Subas, *J. Org. Chem.* **1992**, *57*, 1277–1284.
[21] D. L. Boger, T. Ishizaki, P. A. Kitos, O. Suntornwat, *J. Org. Chem.* **1990**, *55*, 5823–5832.
[22] D. L. Boger, W. Yun, *J. Am. Chem. Soc.* **1994**, *116*, 5523–5524.
[23] P. S. Eis, J. A. Smith, J. M. Rydzewski, D. A. Case, D. L. Boger, W. J. Chazin, *J. Mol. Biol.* **1997**, *272*, 237–252.
[24] J. A. Smith, G. Bifulco, D. A. Case, D. L. Boger, L. Gomez-Paloma, W. J. Chazin, *J. Mol. Biol.* **2000**, *300*, 1195–1204.
[25] J. R. Schnell, R. R. Ketchem, D. L. Boger, W. J. Chazin, *J. Am. Chem. Soc.* **1999**, *121*, 5645–5652.
[26] Y. Kim, J. H. Prestegard, *J. Magn. Reson.* **1989**, *84*, 9–13.
[27] S. C. Wolkenberg, D. L. Boger, *Chem. Rew.* **2002**, *102*, 2477–2496.
[28] D. L. Boger, M. R. Garbaccio, *Bioorg. Med. Chem.* **1997**, *5*, 263–276.
[29] M. A. Warpehoski, D. E. Harper, *J. Am. Chem. Soc.* **1995**, *117*, 2951–2952.
[30] D. L. Boger, A. Santillan, M. Searcey, Q. Jin, *J. Am. Chem. Soc.* **1998**, *120*, 11554–11557.
[31] H. Sugiyama, M. Hosoda, I. Saito, A. Asai, H. Saito, *Tetrahedron Lett.* **1990**, *31*, 7197–7200.
[32] D. L. Boger, B. Bollinger, D. L. Hertzog, D. S. Johnson, H. Cai, P. Mesini, M. R. Garbaccio, Q. Jin, P. A. Kitos, *J. Am. Chem. Soc.* **1997**, *119*, 4987–4998.
[33] D. L. Boger, M. R. Garbaccio, *Acc. Chem. Res.* **1999**, *32*, 1043–1052.
[34] D. E. Wemmer, B. R. Reid, *Annu. Rev. Phys. Chem.* **1985**, *36*, 105–137.
[35] K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York **1986**, 304 pp.
[36] W. J. Chazin, K. Wüthrich, S. Hyberts, M. Rance, W. A. Denny, W. Leupin, *J. Mol. Biol.* **1986**, *190*, 439–453.
[37] D. J. Patel, L. Shapiro, D. Hare, *Quart. Rev. Biophys.* **1987**, *20*, 35–112.
[38] D. J. Patel, L. Shapiro, D. Hare, *Annu. Rev. Biophys. Biophys. Chem.* **1987**, *16*, 423–454.
[39] F. J. M. van de Ven, C. W. Hilbers, *Eur. J. Biochem.* **1988**, *178*, 1–38.
[40] S. S. Wijimenga, B. N. M. van Buuren, *Prog. Nucleic Magn. Reson. Spectrosc.* **1998**, *32*, 287–387.
[41] R. Ditchfield, *J. Chem. Phys.* **1972**, *56*, 5688–5691.
[42] K. Wolinski, J. F. Hinton, P. Pulay, *J. Am. Chem. Soc.* **1990**, *112*, 8251–8260.
[43] C. Adamo, V. Barone, *J. Chem. Phys.* **1998**, *108*, 664–675.
[44] C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158–6170.
[45] G. Barone, L. Gomez-Paloma, D. Duca, A. Silvestri, R. Riccio, G. Bifulco, *Chem. Eur. J.* **2002**, *8*, 3233–3239.
[46] P. Cimino, L. Gomez-Paloma, D. Duca, R. Riccio, G. Bifulco, *Magn. Reson. Chem.* **2004**, *42*, 26–33.
[47] M. Stahl, U. Schopfer, *J. Chem. Soc. Perkin Trans. 2* **1997**, *2*, 905–908.
[48] D. Duca, G. Bifulco, G. Barone, A. Casapullo, A. Fontana, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1024–1030.
[49] J. W. Bausch, R. C. Rizzo, L. G. Sneddon, A. E. Wille, R. E. Williams, *Inorg. Chem.* **1996**, *35*, 131–135.
[50] A. C. J. de Dios, *Prog. NMR Spectrosc.* **1996**, *29*, 229–278.
[51] P. S. Asirvatham, V. Subramanian, R. Balakrishnan, T. Ramasami, *Macromolecules* **2003**, *36*, 921–927.
[52] A. C. J. de Dios, E. Oldfield, *J. Am. Chem. Soc.* **1994**, *116*, 5307–5314.
[53] E. Oldfield, *J. Biomol. NMR.* **1995**, *5*, 217–225.
[54] C. W. Swalina, R. J. Zauhar, M. J. De Grazia, G. Moyna, *J. Biomol. NMR* **2001**, *21*, 49–61.
[55] G. Bifulco, L. Gomez-Paloma, R. Riccio, C. Gaeta, F. Troisi, P. Neri, *Org. Lett.* **2005**, *7*, 5757–5760.
[56] C. Bassarello, P. Cimino, L. Gomez-Paloma, R. Riccio, G. Bifulco, *Tetrahedron* **2003**, *59*, 9555–9562.
[57] G. Bifulco, P. Dambruoso, L. Gomez-Paloma, R. Riccio, *Chem. Rev.* **2007**, *107*, 3744–3779.
[58] D. Marion, K. Wüthrich, *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967–974.
[59] L. Braunschweiler, G. Bodenhausen, R. R. Ernst, *Mol. Phys.* **1983**, *48*, 535–560.
[60] L. P. Mueller, *J. Magn. Reson.* **1987**, *72*, 191–196.
[61] A. J. Shaka, C. J. Lee, A. Pines, *J. Magn. Reson.* **1988**, *77*, 274–293.
[62] S. Macura, R. R. Ernst, *Mol. Phys.* **1980**, *41*, 95–117.
[63] M. Piotto, V. Saudek, V. Sklenar, *J. Biomol. NMR* **1992**, *2*, 661–665.
[64] V. Sklenar, M. Piotto, R. Leppik, V. Saudek, *J. Magn. Reson.* **1993**, *102*, 241–245.
[65] *MacroModel*, version 8.5, Schrödinger LLC, New York, NY, **2003**.
[66] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz Jr, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

[67] P. K. Weiner, P. A. Kollman, *J. Comput. Chem.* **1981**, *2*, 287–303.
[68] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta Jr, P. K. Weiner, *J. Am. Chem. Soc.* **1984**, *106*, 765–784.
[69] S. J. Weiner, P. A. Kollman, D. T. Nguyen, D. A. Case, *J. Comput. Chem.* **1986**, *7*, 230–252.
[70] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
[71] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, *Gaussian 03*, revision B.05, Gaussian, Inc., Pittsburgh, PA, **2003**.

Received: December 20, 2007
Published Online: March 25, 2008

