Regular article

Computational structural determination and energy landscape analysis of the hepatic carcinogen 2-(acetylamino)fluorene

R. Q. Topper[1], K. Chung[1], C. M. Boelke[1], D. Louie[1], J. S. Kang[1], R. Hannan[1,2,3], T. Kiang[1,4], L. H. Chan[1,5]

[1]Department of Chemistry, Albert Nerken School of Engineering, The Cooper Union for the Advancement of Science and Art, 51 Astor Place, New York, NY 10003, USA
[2]Department of Chemistry, New York University, New York, NY 10003, USA
[3]Department of Pathology, Health Science Center at Brooklyn, State University of New York, 450 Clarkson Avenue, BSB 4–122A, Brooklyn, NY 11203, USA
[4]Department of Biomedical Engineering, School of Medicine, John Hopkins University, 720 Rutland Avenue, Baltimore, MD 21205, USA
[5]Novellus Systems Inc., 4000 N. First Street, San Jose, CA 95134, USA

Abstract. 2-(Acetylamino)fluorene (AAF), a potent mutagen and a prototypical example of the mutagenic aromatic amines, forms covalent adducts to DNA after metabolic activation in the liver. A benchmark study of AAF is presented using a number of the most widely used molecular mechanics and semiempirical computational methods and models. The results are compared to higher-level quantum calculations and to experimentally obtained crystal structures. Hydrogen bonding between AAF molecules in the crystal phase complicates the direct comparison of gas-phase theoretical calculations with experiment, so Hartree–Fock (HF) and Becke–Perdew (BP) density functional theory (DFT) calculations are used as benchmarks for the semiempirical and molecular mechanics results. Systematic conformer searches and dihedral energy landscapes were carried out for AAF using the SYBYL and MMFF94 molecular mechanics force fields; the AM1, PM3 and MNDO semiempirical quantum mechanics methods; HF using the 3-21G*and 6-31G* basis sets; and DFT using the nonlocal BP functional and double numerical polarization basis sets. MMFF94, AM1, HF and DFT calculations all predict the same planar structures, whereas SYBYL, MNDO and PM3 all predict various nonplanar geometries. The AM1 energy landscape is in substantial agreement with HF and DFT predictions; MMFF94 is qualitatively similar to HF and DFT; and

the MNDO, PM3 and SYBYL results are qualitatively different from the HF and DFT results and from each other. These results are attributed to deficiencies in MNDO, PM3 and SYBYL. The MNDO, PM3 and SYBYL models may be unreliable for compounds in which an amide group is immediately adjacent to an aromatic ring.

Keywords: 2-(Acetylamino)fluorene – Aromatic amines – Carcinogens and mutagens – Conformer searches

Correspondence to: R. Q. Topper
e-mail: topper@cooper.edu

Introduction

Many aromatic amines are known to be, or suspected to be, carcinogens [1, 2]. Aromatic amines are present in barbecued meats, tobacco smoke, automobile exhaust and dyes. A great deal of research over the past several decades has focused on the elucidation of the structural and biochemical properties of 2-acetamidofluorene or *N*-acetyl-2-aminofluorene, which is most commonly known in the literature as 2-(acetylamino)fluorene (AAF) (Fig. 1). AAF is a potent mutagen and a prototypical example of the mutagenic aromatic amines [3].

Although AAF is a synthetic compound and not an environmental carcinogen, more is known about its biological activity than any other aromatic amine and it serves as an important prototype in the literature. When activated metabolically in the liver, AAF adds to DNA at either C8 of guanine (forming the "major" adduct, 85% yield in vivo) or, less probably, at N2 of guanine (forming the "minor" adduct, 15% yield) [4, 5]. The major adduct to DNA in solution has been characterized experimentally by high-resolution NMR [6]. The crystal

Fig. 1. a Structure and atom numbering scheme for 2-(acetylamino) fluorene and **b** illustration of the (β, γ, δ) dihedral angles. The γ' dihedral angle is a rotation about the same bond as the γ dihedral angle

structure of AAF bound to C8 of guanine has been determined by X-ray crystallography [7] but an experimental structure of the minor adduct is not yet available. In addition, a number of X-ray crystallography studies have been carried out on AAF itself [8, 9]. The minor adduct is more persistent in vivo and is repaired more slowly than the major adduct, and so it has been speculated that the minor adduct may actually be the more important of the two (see Ref. [10] and references therein). A 1997 molecular dynamics study of the minor adduct predicted that the minor adduct lies along the minor groove of DNA [10], and this is the only structure determination of the minor adduct in the literature to date.

In the present work, a benchmark conformational study of AAF is presented using molecular mechanics and semiempirical methods. The structures thus obtained are compared to X-ray crystal structures [8, 9] and to structures obtained using Hartree Fock (HF) and nonlocal density functional theory (DFT) conformer searches. In addition, energy landscapes were generated for the various methods as a function of a subsequently identified critical dihedral angle. The need for the present study is clear. A number of molecular dynamics and molecular mechanics studies have been presented for AAF-modified oligomers, nucleosides and nucleotides [6, 10, 11, 12, 13, 14, 15] and one recent semiempirical study using MNDO has appeared for AAF-modified deoxyguanosine [16]. However, no theoretical studies of unmodified AAF have appeared in the literature to date. A survey of various aromatic amines was carried out using the MNDO, AM1 and PM3 semiempirical methods, but this survey did not include AAF or similar compounds [17]. As a result it is not possible to evaluate which computational methods

are appropriate for this class of systems without a study of the type presented here. A 1987 study of over 60 conjugated molecules by Fabian [18] showed that AM1 can be used successfully for many conformational problems whereas MNDO often fails, although rotational barriers and conformational energy differences are still often underestimated by AM1 [19, 20]. The goal of the present work is to find a reliable and practical conformer searching protocol which can be extended to reliable structure determinations of the major and minor nucleosides. In this spirit emphasis has been placed on the use of small, affordable basis sets in HF calculations as well as on the use of DFT, which includes the effect of explicit electron–electron correlation and also scales much more favorably with system size than post-HF methods [21].

Computational methods

All calculations were carried out using Spartan 5.1.3 [22], running on single- and dual-processor SGI Octane computers under versions 6.4 and 6.5 of the IRIX operating system. Data were generated for a total of nine computational methods, which include the AM1[20], PM3 [23] and MNDO [19] semiempirical quantum mechanics methods, the SYBYL [24] and MMFF94 [25] molecular mechanics force fields, HF theory using the 3-21G* and 6-31G* basis sets (HF/3-21G* and HF/6-31G*) [26], and BP DFT [27, 28] using the double-numerical DN, DN* and DN** basis sets.[29, 30, 31] The performance of many of these methods and basis sets for representative molecules and reactions has been extensively discussed elsewhere.[21] We note in particular that the DN basis set is similar in size to (but slightly larger than) the more well-known 6-31G basis set, DN* is slightly larger than the 6-31G* basis set, and DN** is slightly larger than the 6-31G** basis set. No extensions of the SYBYL and MMFF94 force fields were used. All dihedral angles were defined on the interval (0, 360°).

Results and discussion

Conformer searches of AAF

A series of systematic conformer searches was carried out on AAF. In a preliminary study, atomic coordinates from an X-ray crystallographic determination by Van Meerssche et al. [9] were used as the initial structure of AAF. The β dihedral angle (β = C1–C2–N C14 Fig. 1) was systematically and uniformly varied from its initial value (β = 44°) using 8-fold and 12-fold conformational sampling. Each resulting geometry was subjected to full (i.e. unconstrained) geometry optimization using either the MMFF94 or the SYBYL force fields. The results are shown in Table 1, which contains information about the relative conformer energies as well as the values of the β, γ and γ' dihedral angles (γ = C2–N C14–O and γ' = H–N–C14–O; if γ' = 180 the amide group is in the *trans* configuration).

The 8-fold and 12-fold searches were found to yield identical results to one another within each method, but the MMFF94 minimum-energy structures are planar (β = 1° and 179°) and the SYBYL structures are non-

Table 1. Relative energies (ΔE, kcal/mol) and dihedral angles (β, γ, γ', degrees) of 2-(acetylamino)fluorene (*AAF*) strain energy mimina located by 12-fold systematic conformer searches of the β dihedral angle of AAF using the MMFF94 and Sybyl force fields. $\Delta E = 0$ indicates the lowest-energy structure located by the search. Also shown are the dihedral angles obtained from X-ray crystallography [8,9]. Parameters whose values were not reported are indicated by n.r.

Source/method	ΔE	β	γ	γ'
MMFF94	0.0000	1	1	−178
	0.3821	179	−1	178
SYBYL	0.0000	−40	−2	176
	0.0679	140	−2	176
Ref. [8]	n/a	42	n.r.	n.r.
Ref. [9]	n/a	44	0	n.r.

Table 2. Relative energies (ΔE, kcal/mol), dihedral angles (β, γ, γ', degrees) and distance from N to the plane of its three covalent-adducts (R, Å) of AAF strain energy minima located by extensive systematic conformer searches of the (β, γ, δ) dihedral angles using the MMFF94 and SYBYL force fields. If N is sp^2 hybridized, $R = 0$

Method	ΔE	β	γ	γ'	R
MMFF94	0.0000	1	1	−178	0.00
	0.3821	179	−1	178	0.00
	4.1146	0	−180	0	0.00
	4.3489	178	179	−1	0.00
SYBYL	0.0000	−40	−2	176	0.01
	0.0679	140	−2	176	0.02
	1.4678	−58	177	−5	0.01
	1.5445	122	177	−5	0.01

planar ($\beta = -40°$ and 140°). The X-ray crystal structures yield a nonplanar structure, which seems to cast doubt on the MMFF94 result, but this is misleading. As discussed by Haisa and Kashino [8], in the solid phase the N–H hydrogen of each AAF molecule forms a hydrogen bond to the oxygen of its nearest neighbor, which forces the dihedral angle out of the plane of the fluorene ring system. However, the present calculations are in the gas phase and so crystal-phase hydrogen bonding between nearest neighbors is not taken into account in these calculations. The quality of these models is therefore most appropriately assessed by comparison with high-quality quantum-mechanical calculations (see later) rather than by direct comparison with the crystallographic data.

A second set of conformer searches was carried out using the lowest-energy geometry resulting from each of the two force fields as the initial structure. This procedure generated a second, completely different set of initial structures for geometry optimization in each case. Identical results were obtained once again for each force field. Finally, a series of conformer searches were carried out using the lowest-energy structure found for each force field in the previous studies as a starting configuration for geometry optimization. In addition to β, these conformer searches included rotation about the γ dihedral angle. The δ dihedral angle ($\delta =$ N–C14–C15–H) was also included in some searches although δ is basically a spectator angle as it involves the "spinning" of a methyl group, but its inclusion sometimes was observed to yield a small quantitative effect on the results. The three angles (β, γ, δ) were systematically and uniformly varied, sampling each angle as few as ten times (which generates 1,000 initial conformers) and as many as 50 times (125,000 conformers) for MMFF94 and 25 times (15,625 conformers) for SYBYL. These results are summarized in Table 2. In addition to the three angles (β, γ, γ'), the minimum distance between the nitrogen atom and the plane formed by the C2, H and C14 atoms bonded to it, R, is also given in Table 2. If $R = 0$, the three atoms are coplanar with N and N is trigonal planar. This is the case for both force fields.

It was observed that a tenfold search of 1,000 initial conformers was sufficient to completely resolve the accessible minimum-energy structures in both cases. It was also observed that the two lowest-energy structures were always the same as those discovered by only varying the β dihedral angle (Table 1). This helps to justify the use of only one torsion angle for sampling purposes in the subsequent quantum-mechanical calculations. This is an important point because subjecting thousands of conformers to direct geometry optimization using the high-level quantum mechanics methods is beyond the capacity of available computational resources.

An eightfold conformer search on β using each of the seven quantum-mechanical computational methods described in the previous section was also carried out. On the basis of the results of the more extensive molecular mechanics conformer searches, this seems to be sufficient for the resolution of the most important features of the energy landscape for the quantum methods. This was specifically confirmed for the case of the AM1 method by carrying out another series of extensive conformer searches in which each of the three dihedral angles was sampled as many as 20 times (generating 8,000 conformers for analysis). For all quantum calculations the initial Hessian was set equal to unity (HESS = UNIT) and the NOSYMTRY keyword was used as well. In the case of the DFT calculations, the optimized geometries obtained from an eightfold BP/DN conformer search were used as input to BP/DN* and BP/DN** geometry optimizations. The quantum-mechanical conformer search results are summarized in Table 3.

The AM1, HF/3-21G*, HF/6-31G* and BP/DN conformer searches presented in Table 3 are seen to have produced completely planar minimum-energy structures. The structures are observed to remain planar when the geometry is optimized at the BP/DN* and BP/DN** levels. In contrast, SYBYL, MNDO and PM3 conformer searches produced nonplanar structures. Moreover, these nonplanar structures are in substantial disagreement with one another. SYBYL predicts a minimum-energy structure with $\beta = -40°$; MNDO predicts $\beta = -80°$, and PM3 predicts $\beta = -56°$. Finally, the

Table 3. Relative energies (ΔE, kcal/mol), dihedral angles (β, γ, γ′, degrees) and distance from N to the plane of its three covalent adducts (R, Å) of AAF energy minima located by an eightfold conformer search of the β dihedral angle using the AM1, PM3, MNDO, HF/3-21G*, BP/DN, BP/DN* and BP/DN**methods. If N is sp^2 hybridized, $R = 0$

Method	ΔE	β	γ	γ′	R
AM1	0.0000	1	0	180	0.00
	0.0977	−179	0	180	0.00
PM3	0.0000	−56	18	158	0.26
	0.1406	−128	−18	−157	0.26
	0.7154	17	−19	−166	0.20
MNDO	0.0000	−80	13	161	0.20
	0.0211	−101	−13	−161	0.20
HF/3-21G*	0.0000	0	0	180	0.00
	0.4336	180	0	180	0.00
BP/DN	0.0000	0	−1	180	0.00
	0.1562	178	1	180	0.01
BP/DN*	0.0000	2	3	180	0.02
	0.5441	180	3	180	0.02
BP/DN**	0.0000	2	−3	180	0.02
	0.5830	178	3	180	0.02

PM3 and MNDO methods predict significant noncoplanarity of the amide group and pyramidalization of the nitrogen atom; PM3 predicts $\gamma = 18°$ and $R = 0.26$ Å, and MNDO predicts $\gamma = 13°$ and $R = 0.20$ Å. In contrast, BP/DN** calculations (the highest level of geometry optimization used in this study) predict a planar structure with $\beta = 2°$, with a $\beta = 178°$ planar structure lying 0.58 kcal/mole higher in energy. AM1 results are similar to those obtained with BP/DN**, but with a much smaller energy gap (0.10 kcal/mole).

The reason for the stabilization of planar structures in the HF and DFT calculations can be rationalized upon examination of the highest occupied molecular orbital (HOMO). For the purpose of simplified visualization we consider a smaller model compound,

Fig. 2. *Top*: top view of the HF/6-31G* HOMO of acetylaminobenzene. *Bottom*: side view

acetamidobenzene (AAB), which has virtually all of the same features as AAF. In Figure 2 we show the HOMO predicted by a HF/6-31G* calculation (again, the HOMO of AAF at the is qualitatively similar to that of AAB but is somewhat simpler to visualize). The HOMO simultaneously has strong π character in the ring and in the adjacent N–C bond of the amide group. The coplanarity of the two groups form a molecular orbital with strongly aromatic character, while tolerating a certain degree of repulsion between hydrogen atoms.

Dihedral angle coordinate driving studies of AAF: energy landscapes

In order to visually highlight some of the similarities and differences between the various models, a series of coordinate driving studies of AAF were carried out using various methods. β was systematically and uniformly varied from 0 to 360° by as few as 20 increments and as many as 40. A constrained geometry optimization was carried out with β fixed at each new dihedral angle. For the sake of consistency all coordinate driving calculations were carried out using the lowest-energy structure predicted by the MMFF94 force field as the initial geometry. Plots of the relative energy as a function of β were created for each method. From this data, a further comparison may then be made as to which methods are qualitatively similar to one another and which ones are qualitatively different from one another.

The resulting approximate energy landscapes for the MMFF94 and SYBYL force fields are shown in Fig. 3. Once again it is evident that the two force fields produce qualitatively different results, with minima and maxima in completely different locations. The energy landscapes for AM1, PM3 and MNDO calculations are presented in Fig. 4. Once again, the three semiempirical methods produce results that are completely at odds with one another, predicting minima and maxima in completely different locations.
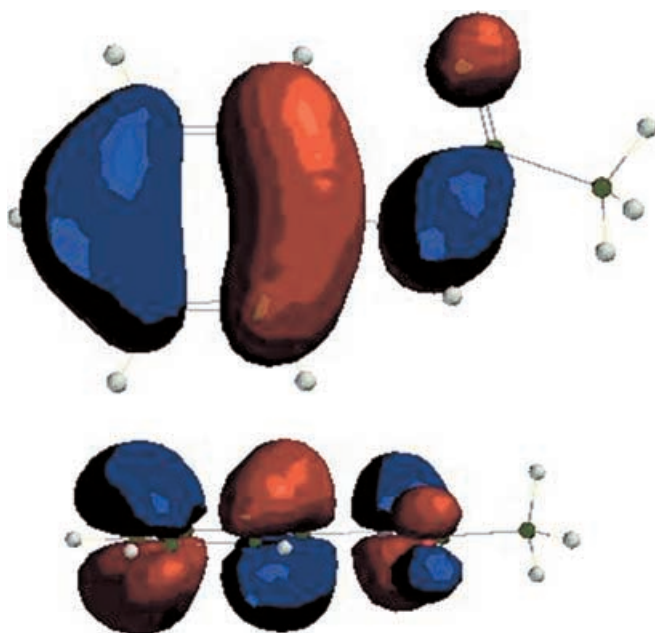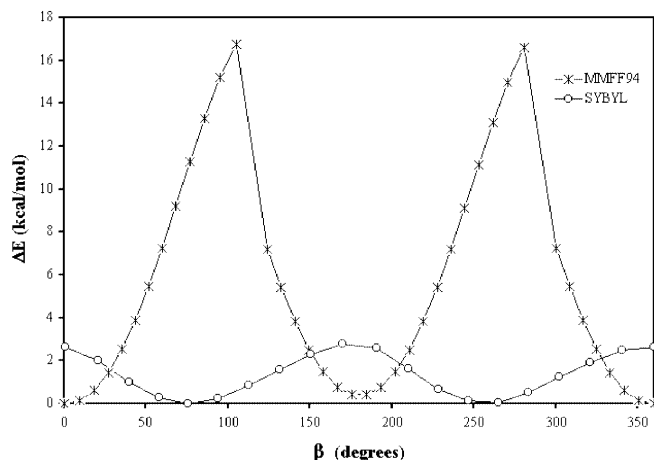
Fig. 3. Relative energy, $\triangle E$, as a function of the dihedral angle, β, using the MMFF94 and SYBYL force fields
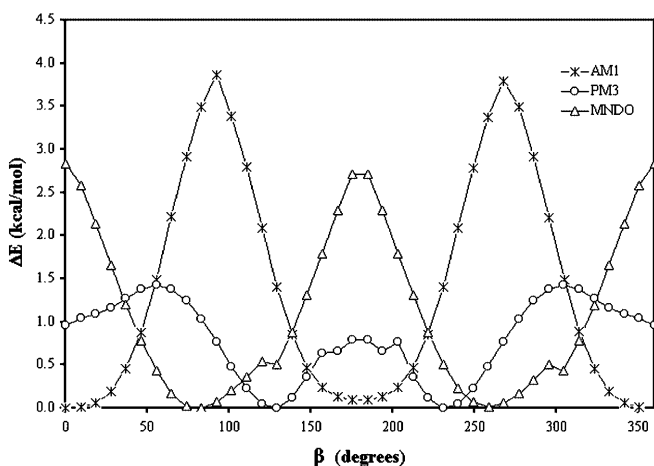
Fig. 4. $\triangle E$ as a function of β using the MNDO, AM1 and PM3 semiempirical models
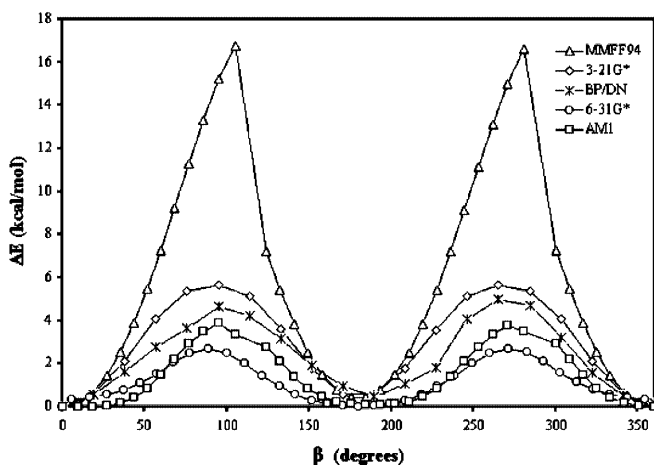

Fig. 5. $\triangle E$ as a function of β obtained from HF/3-21G*, HF/6-31G*, BP/DN, AM1 and MMFF94 calculations

Finally, the energy landscapes obtained from HF/3-21G*, HF/6-31G* and BP/DN calculations are presented together with the MMFF94 and AM1 landscapes in Fig. 5. All five of these methods produce qualitatively similar landscapes, with maxima and minima in the same locations. The MMFF94 landscape shows significantly larger barriers than the other models, but as MMFF94 is not specifically parameterized to predict rotational energy barriers this is not a surprising result. As is often the case, increasing the size of the basis set in the HF calculations substantially changes the barrier heights [21].

Conclusions

If a method does not correctly predict the structure of AAF, that same method is unlikely to be reliable for predictive studies of AAF adducts to guanine or to deoxyguanosine. In this spirit, it may be concluded that SYBYL, MNDO and PM3 are probably not reliable methods for calculations of AAF or its adducts [16]. In addition, they may not be appropriate models for any compound which contains an amide group adjacent to an aromatic ring system, [8, 32, 33, 34] although further investigation is warranted and is currently taking place in this laboratory. Unlike these three methods, the AM1, MMFF94, HF and DFT calculations all predict completely planar geometries for AAF. This appears to be due to the highly aromatic character of the HOMO, which extends across most of the nuclear framework and contains significant contributions from both the aromatic ring π system and the amide π bond. This has the net effect of preferentially stabilizing the planar geometry. The SYBYL, MNDO and PM3 models are not appropriately parameterized to accurately take this effect into account.

Although the failure of PM3 to perform properly for this molecule was not particularly obvious at the outset, it is no surprise that MNDO fails to describe this system properly. MNDO's inability to correctly describe the rotational energy landscape of nitobenzene and benzaldehyde was noted by Dewar et al. [20] in their original presentation of the AM1 method, and was attributed by them to an overestimation of the repulsive nonbonded interaction between the ring hydrogens and the oxygen atoms. Subsequent studies of many compounds confirmed the tendency of MNDO to overestimate repulsions between nonbonded atoms and to incorrectly predict nonplanar geometries for conjugated organic molecules (see Ref. [18] and references therein). AM1 was designed to overcome this limitation (and other limitations) of MNDO. In addition, the MNDO and PM3 models incorrectly predict nonplanarity of the amide group [35], which is treated in MOPAC and other codes by the use of the MMOK keyword. We have investigated the use of this keyword in PM3 geometry optimization of the lowest-energy PM3 conformer using SPARTAN '02 [36, 37] and found that the amide group is still nonplanar ($\gamma' = 158°$) and the amide and fluorene

groups are still not coincident ($\beta = -56°$). This study reinforces the importance of carefully benchmarking any semiempirical or molecular mechanics method as thoroughly as possible for small model compounds before tackling larger systems.

The MMFF94 and AM1 models are seen to perform reliably for aspects of AAF structure prediction in comparison to HF and DFT calculations; however, there is as yet no direct evidence as to how well or how poorly these methods will perform for adducts of AAF with guanine, deoxyguanosine or DNA. Current research in our group includes the use of the MMFF94 and AM1 methods as "prescreening" methods for extensive conformer searches of the major and minor AAF-guanine and AAF-deoxyguanosine adduct complexes, submitting the resulting family of conformers to subsequent DFT geometry optimizations.

References

1. Kriek E, Miller JA, Juhl U, Miller EC (1967) Biochemistry 6:177
2. Nesnow S, Argus M, Bergman H, Chu K, Frifth C, Helmes T, McGaughy R, Ray V, Slaga TJ, Tennant R, Weisburger W (1988) Mutat Res 185:1
3. Kriek E (1992) J Cancer Res Clin Oncol 118:481
4. Irving CC, Veazey RA (1969) Cancer Res 29:1799
5. Kriek E (1974) Biochim Biophys Acta 355:177
6. O'Handley SF, Sanford DG, Xu R, Lester CC, Hingerty BE, Broyde S, Krugh TR (1993) Biochemistry 32:2481
7. Neidle S, Kuroda R, Broyde S, Hingerty BE, Levine RA, Miller DW, Evans FE (1984) Nucleic Acids Res 12:8219
8. Haisa M, Kashino S, Ueno T, Shinozaki N, Matsuzaki Y (1980) Acta Crystallogr Sect B 36:2306
9. Van Meerssche M, Germain G, Decleroq JP, Touillaux R (1980) Crystallogr Struct Commun 9:515
10. Grad R, Shapiro R, Hingerty BE, Broyde S (1997) Chem Res Toxicol 10:1123
11. Hingerty B, Broyde S (1982) Biochemistry 21:3243
12. Broyde S, Hingerty B (1983) Biopolymers 22:2423
13. Hingerty B, Broyde S (1986) J Biomol Struct Dyn 4:365
14. Shapiro R, Hingerty BE, Broyde S (1989) J Biomol Struct Dyn 7:493
15. Fritsch V, Westhof E (1991) J Comput Chem 12:147
16. Besson M, Mihalek CL (2001) Mutat Res 473:211
17. Smith DA, Ulmer II CW, Gilbert MJ (1992) J Comput Chem 13:640
18. Fabian WMF (1988) J Comput Chem 9:369
19. Dewar MJS, Theil W (1977) J Am Chem Soc 99:4899
20. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107:3902
21. Hehre W J, Lou L (1997) A guide to density functional calculations in SPARTAN. Wavefunction, Irvine, CA
22. Wavefunction Inc SPARTAN 5.1.3. Wavefunction, Irvine, CA, http://www.wavefun.com/
23. Stewart JJP (1989) J Comput Chem 10:221, and references therein
24. Clark M, Cramer RD III, van Opdensch N (1989) J Comput Chem 10:982
25. Halgren TA (1996) J Comput Chem 17:490
26. Hehre WJ, Radom L, Schleyer PvR, Pople JA (1986) Ab initio molecular orbital theory. Wiley, New York
27. Becke AD (1988) Phys Rev A 38:3089
28. Perdew JP (1986) Phys Rev B 33:8822
29. Becke AD (1988) J Chem Phys 88:2547
30. Verslcis L, Ziegler T (1988) J Chem Phys 88:322
31. Delley B (1990) J Chem Phys 92:508
32. Evans FE, Miller DW (1983) J Am Chem Soc 105:4863
33. Huggett AC, Cone JL, Thorgeirsson SS, Roller PP (1987) J Org Chem 52:4933
34. Wulff G (2002) Chem Rev 102:1, and references therein
35. Neidle S, Subbiah A, Mason A, Islam SA (1981) Carcinogenesis 2:901
36. Wavefunction Inc SPARTAN '02. Wavefunction Inc, Irvine, CA, http://www.wavefun.com/
37. Kong J, White CA, Krylov AI, Sherill CD, Adamson RD, Furlani TR, Lee MS, Lee AM, Gwaltney SR, Adams TR, Ochsenfeld C, Gilbert ATB, Kedzoria GS, Rassolov VA, Maurice DR, Nair N, Shao Y, Besley NA, Maslen PE, Dombroski JP, Daschel H, Zhang W, Koramblath PP, Baker J, Byrd EFC, Van Voorhis T, Oumi M, Hirata S, Hsu C-P, Ishikawa N, Florian J, Warshel A, Johnson BG, Gill PMW, Head-Gordon M, Pople JA (2000) J Comput Chem 21:1532