Acta Chim. Slov. **1999**, *46*(4), pp. 493-500

AN AB INITIO MODEL STUDY OF THE CYTOSINE – METHYLBENZENE INTERACTION*

Jože Koller
Faculty of Chemistry and Chemical Technology, University of Ljubljana, Slovenija

(Received 25.5.1999)

Abstract

The unconventional hydrogen bond involving the amino group as proton donor and benzene ring as proton acceptor was studied on the B3LYP/6–31G level. Medium effects were simulated by the Polarisable Continuum Model of Tomasi et al. Obtained results (hydrogen bond energy –1.28 kcal/mol, nitrogen – benzene distance 3.85 Å) were discussed and compared with the data from the literature. To our knowledge this is the first attempt of treating the amino/benzene hydrogen bond with the inclusion of the medium.

Introduction

Hydrogen bonding plays an important role in the structure and function of biological molecules. Traditionally it involves an essentially electrostatic interaction between a proton attached to an electronegative atom (O, N or C) and another oxygen or nitrogen atom (acting as a proton acceptor). These interactions are usually termed *conventional hydrogen bonds*. The evidence for aromatic rings acting as hydrogen bond acceptors was first discovered by spectroscopy. Infrared spectra of N–ethylacetamide in various solvents revealed that the NH–stretching frequency was the lowest when benzene was used as a solvent [1]. The authors suggest that the π electrons of benzene ring act as a base. The term *unconventional hydrogen bond* could be used for this type of interaction. When the donating proton is attached to a nitrogen atom the interaction is described as amino/aromatic hydrogen bond. The existence of such unconventional

* Dedicated to Prof. Dr. D. Leskovšek on the occasion of his 80th birthday.

hydrogen bonding is well established at least for complexes of small molecules. Examples are the water/benzene [2] and ammonia/benzene complexes [3] which both show unconventional hydrogen bonding. Structures were determined experimentally by spectroscopic techniques and theoretically by ab initio calculations. However, it is not yet clear how important the unconventional hydrogen bonding is in the biochemical processes. Perutz [4] concludes at the end of his paper: "The chemical literature contains many more examples of interaction between proton donors and aromatic acceptors, but the ones quoted here are sufficient to show that such interactions are of biologically significant strength, that they play an important role in synaptic and cellular signal transduction and also contribute to the stability of certain protein structures." On the other hand Mitchell et al. [5] argued that although some gas phase calculations do in fact favour amino/aromatic hydrogen bonded structures such structures are remarkably rare. In their study they considered 55 high–resolution protein chain structures and evidenced 1108 nitrogen/aromatic ring interactions but only 30 (or 2.7%) of them are amino/aromatic hydrogen bonds according to the predefined criteria.

However, some papers appeared recently indicating that aromatic hydrogen bonds can substitute conventional hydrogen bond in some specific protein – DNA interactions. Parkinson et al. [6] studied the binding of *Escherichia coli* catabolite activation protein (CAP) to DNA [7,8]. CAP, also referred as cAMP (cyclic adenosine monophosphate) receptor protein CRP, is a transcription activator protein. The complex of CAP and cAMP stimulates transcription by binding to specific DNA sites at or near promoters. Authors substituted the carboxylate side chain of Glu by the aromatic chains of Phe, Tyr and Trp. Unexpectedly they found that this substitution did not reduce the DNA binding affinity, presumably due to the formation of aromatic hydrogen bonds. We decided to model this specific part of interaction between the DNA and protein by ab initio calculation.

Computational

Escherichia coli catabolic activation protein (CAP) binds to DNA as a dimer of two identical protomers. The atomic coordinates of the CAP–DNA complex were

obtained from the Brookhaven Protein Data Bank (accession number 1RUO). The complete arrangement of 4433 atoms can be seen on Figure 1. Thirty–six water molecules are included. The unconventional hydrogen bond that we are interested in is established between the DNA cytosine 7 (C 7) and the CAP phenylalanine 181 (Phe 181).

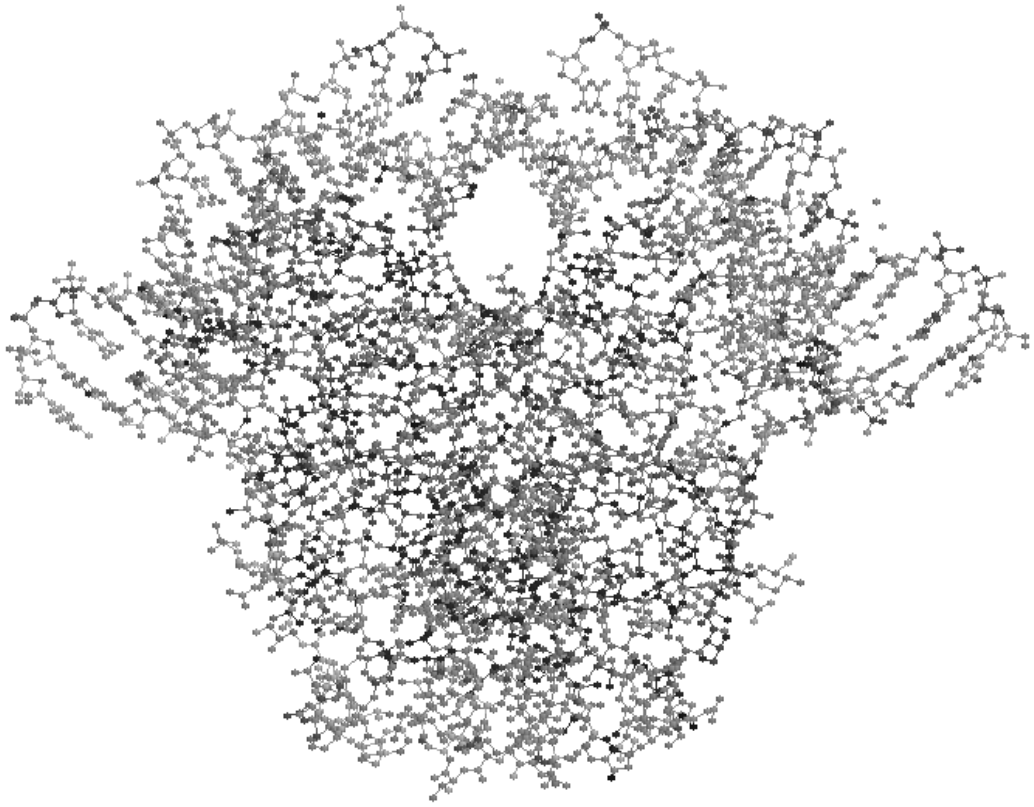

Figure 1: Structure of the CAP–DNA complex (Raswin picture using the coordinates from PDB)

The enlarged view on that part of complex in which the hydrogen bonded molecules can be seen is on the Figure 2. To make the calculations feasible and to retain all the essential interactions we substituted the phenylalanine molecule by methylbenzene. Figure 3 shows the molecules considered for calculations. The optimized geometry of the hydrogen bonded system was computed using the ab initio theory on the HF/6–31G level with the Gaussian 98 suite of programs [9]. To include some of the correlation energy the calculation was done with the density functional

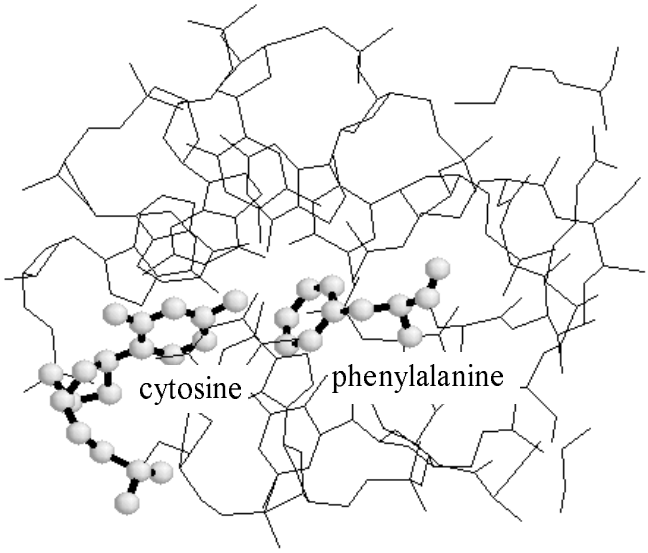

Figure 2: Enlarged view on the part of the CAP–DNA complex with the cytosine and phenylalanine molecules (hydrogen atoms are not shown)

theory (DFT) with the same basis set and Becke´s three parameter hybrid functional using the LYP correlation functional (B3LYP/6–31G). There are several possibilities how to include the interaction with the medium to model the real situation as good as possible. We choose the Polarized Continuum Model (PCM) of Tomasi and coworkers [10–12] which is included in the Gaussian 98 suite of programs and performed the total geometry optimization with the value ε=20 for the dielectric constant of the solvent. Different values for the dielectric permitivity ε of protein are used in the literature, ranging from 2 to 80. For there is no reliable experimental technique to determine ε, it is usually calculated theoretically. Because of its polar groups the dielectric constant of protein should be at least 10. Molecular dynamics simulations for two proteins [13] gave the values 30 and 36. The value ε=20 that we used could be taken as a compromise. To study the influence of the guanine molecule bonded with the three hydrogen bonds to cytosine we modeled also this interaction on the B3LYP/6–31G level. Treating the systems with several hydrogen bonds one must be aware of the nonadditivity of the hydrogen bond energies because of cooperative effects. However,

the overall geometry of the cytosine–methybenzene pair did not change significantly when the model guanine molecule was included in the calculation. This is the reason why we decided to consider only the cytosine–methybenzene complex. The results for the guanine–cytosine–methybenzene system are not reported here.

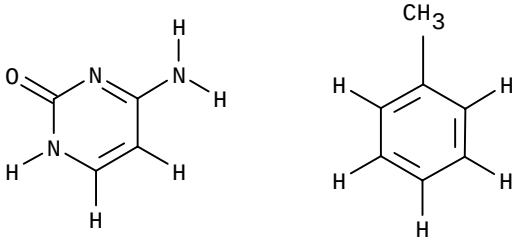

Figure 3: Cytosine and methybenzene molecules

Results and discussion

The numerical results including the hydrogen bond energies and geometric parameters are presented in Table 1. The B3LYP/6–31G calculation using the Polarized Continuum Model (PCM) of Tomasi is denoted as B3LYP/6–31G–Tom, for all other cases standard abbreviations are used. The approximate relative orientation of both interacting species is characterized by two parameters (Figure 4). R is the distance between the -NH_2 nitrogen of cytosine and the centre of the benzene ring of methylbenzene while α is the angle between R and the 6–fold symmetry axis of the benzene ring. Results show that rather small shortening of R takes place when the level of calculation is changed from HF to DFT. The same is true for the inclination angle α which changes from 16 to about 23 degrees. In all geometries the direction of the N–H bond (involved in the hydrogen bonding) points towards some of the carbon atoms which form the benzene ring because of the toroidal form of the π–electron density. As expected the hydrogen bond energy tends to become lower with inclusion of the solvent. Let us discuss the value –1.28 kcal/mol, obtained for the B3LYP/6–31G–Tom calculation. Conventional hydrogen bonds have the energies in the range from 2 to 20

kcal/mol [14]. The value obtained presently is less than the lowest end of the range commonly accepted for hydrogen bonding. Several theoretical studies of unconventional (aromatic) hydrogen bonds can be found in the literature. Suzuki et al. [2] studied the benzene – water complexes on the MP2/6–31G** level with basis set superposition error corrections. The theoretical aromatic hydrogen bond energy they obtained is –1.78 kcal/mol while the distance between the water oxygen atom and the

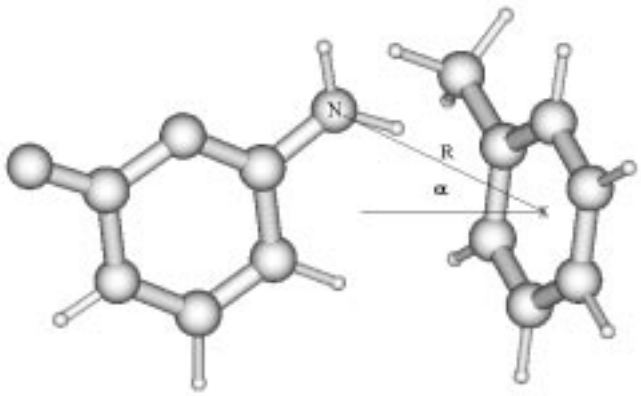

Figure 4: Geometry of the cytosine–methylbenzene complex (R – distance between the nitrogen atom and the centre of the benzene ring, α – angle between R and the 6–fold symmetry axis of the benzene ring)

Table 1: Geometric parameters and hydrogen bond energies ΔE_{hb} for the cytosine–methylbenzene complex calculated on the different levels of theory. For the definition of R and α see Figure 4.

theory	geometry optimization	R (10^{-10}m)	α (degrees)	ΔE_{hb} (kcal/mol)
6-31G	6-31G	3.90	16.0	–3.08
B3LYP/6-31G	B3LYP/6-31G	3.84	23.8	–3.51
B3LYP/6-31G-Tom[#]	B3LYP/6-31G	3.84	23.8	–1.78
B3LYP/6-31G-Tom[#]	B3LYP/6-31G-Tom[#]	3.85	22.8	–1.28

[#] B3LYP/6-31G calculation with the Tomasi PCM [10–12]

center of the benzene ring is 3.159 Å (the value deduced from experiment is 3.374 Å [2]). The benzene – ammonia dimer was studied by Rodham et al. [3] on the same level of theory (MP2/6–31G**). Values of –2.4 kcal/mol and 3.43 Å were obtained for the

bonding energy and the nitrogen – benzene distance, respectively. It is quoted in the paper concerning the binding of CAP to DNA [6] that the aromatic hydrogen bond Phe 181 – C 7 (DNA) contributes about 0.8 kcal/mol per CAP protomer to the binding free energy. According to the crystal structure of CAP – DNA complex the benzene ring is perpendicular to the cytosine >N–H bond and the distance of the nitrogen atom from the benzene ring is 3Å. We believe that our study of aromatic hydrogen bond supports the opinion that although its strength is relatively low it is important at least in some specific protein – DNA interactions.

Acknowledgement

Author wishes to thank Prof. D. Had□i for continuous support and helpful discussions.

References

[1] W. Klemperer, M.W. Cronyn, A.M. Maki, G.C. Pimentel, *J. Am. Chem. Soc.* **1954**, 76, 5846-5848.

[2] S. Suzuki, P.G. Green, R.E. Bumgarner, S. Dasgupta, W.A. Goddard, G.A. Blake, *Science* **1992**, 257, 942-945.

[3] D.A. Rodham, S. Suzuki, R.D. Suenram, F.J. Lovas, S. Dasgupta, W.A. Goddard, G.A. Blake, *Nature (London)* **1993**, 362, 735-737.

[4] M.F. Perutz, *Phil. Trans. Roy. Soc. A* **1993**, 345, 105-112.

[5] J.B.O. Mitchell, C.L. Nandi, I.K. McDonald, J.M. Thornton, S.L. Price, *J. Mol. Biol.* **1994**, 239, 315-331.

[6] G. Parkinson, A. Gunasekera, J. Vojtechovsky, X. Zhang, T.A. Kunkel, H. Berman, R.H. Ebright, *Nature Struct. Biol.* **1996**, 3, 837-841.

[7] A. Kolb, S. Busby, H. Buc, S. Garges, S. Adhya, *Ann. Rev. Biochem.* **1993**, 62, 749-795.

[8] G. Parkinson, C. Wilson, A. Gunasekera, Y.W. Ebright, R.E. Ebright, H.M. Berman, *J. Mol. Biol.* **1996**, 260, 395-408.

[9] Gaussian 98, Revision A.5, 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, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, **1998**.

[10] S. Miertus, E. Scroco, J. Tomasi, *Chem. Phys.* **1981**, 55, 117-129.

[11] S. Miertus, J. Tomasi, *Chem. Phys.* **1982**, 65, 239-252.

[12] V. Barone, M. Cosi, J. Tomasi, *J. Comp. Chem.* **1998**, 19, 404- 417.

[13] P.E. Smith, R.M. Brunne, A.E. Mark, W.F. van Gunsteren, *J. Phys. Chem.* **1993**, 97, 2009-2014.

[14] A.D. Buckingham, *The Hydrogen Bond: An Electrostatic Interaction?* in *Theoretical Treatments of Hydrogen Bonding*, D. Had□i ed., John Willley & Sons, **1997**, pp 1-12.

Povzetek

Z metodo B3LYP/6–31G smo študirali nekonvencionalno vodikovo vez, pri kateri je bila amino skupina donor, benzenov obroč pa akceptor protona v vezi. Vključili smo tudi vpliv okoliškega medija s pomočjo Tomasijevega modela polarizirnega kontinuuma. Kolikor vemo je to prvi poskus vključitve okolice pri teoretičnem obravnavanju tovrstne vodikove vezi. Dobljene rezultate (–1.28 kcal/mol za energijo vodikove vezi ter 3.85 Å za razdaljo med atomom dušika in težiščem benzenovega obroča) smo primerjali s podatki iz literature.

AN AB INITIO MODEL STUDY OF THE CYTOSINE – METHYLBENZENE INTERACTION

AB INITIO RAÈUNI INTERAKCIJE CITOZINA Z METILBENZENOM

avtor in naslov za korespondenco:

Prof. dr. Jo□e Koller

Fakulteta za kemijo in kemijsko tehnologijo

Aškerèeva 5, pp 537

1000 Ljubljana

Slovenija

telefon: 061 1760-614

fax: 061 1254-215

e-mail joze.koller@uni-lj.si