A Computational Study of the Molecular Recognition of Nucleic Acid Bases by Poly(vinyldiaminotriazine)

Jiří Czernek

Summary: The enthalpic gain upon formation of the complexes of the poly(vinyldiaminotriazine) model with purine, uracil and xanthine has been assessed by means of the RI-MP2 interaction energy calculations. Based on the underlying RI-DFT-D geometrical optimizations, the molecular-level insight into the recognition processes has been proposed.

Keywords: molecular modeling; nucleobases; poly(vinyldiaminotriazine); recognition

Introduction

Noncovalent interactions are being frequently used in the molecular self-assembly of well-defined supramolecular structures forming the basis of novel materials.[1] In polymer science, such interactions are important in determining the properties of supramolecular polymers[2] and can be employed in the design of new polymer architectures.[3] They also govern the molecular recognition of guest molecules by polymeric receptors.[4] A premier example of a newly developed supramolecular polymeric system is based on the ability of poly(2-vinyl-4,6-diamino-1,3,5-triazine) (PVDAT) to act as the artificial receptor for an efficient binding of nucleic acid bases and related compounds.[5–7] Thus, the recognition properties of the polymeric field of the PVDAT homopolymer and of some of the PVDAT-containing copolymers have found important technological applications.[8–9] However, due to the absence of direct structural information, the mechanism of the assembly of the PVDAT homo- and copolymers with guest molecules is not known precisely. Clearly, the knowledge of geometrical parameters involved in the formation of the complexes could be instrumental for the proper understanding and fine-tuning of the directionality of hydrogen bonding and other weak reversible interactions responsible for the binding.[10] Undoubtedly, modern techniques of computational chemistry can be usefully employed in the description of the recognition of the guests by polymers at the molecular level.[11] Here several quantum chemical methods were applied to investigate the geometries and energetics of model systems with the aim to describe the strength and directionality of the stabilizing structural elements in the PVDAT complexes. In particular, the differences were analyzed between the stacked and hydrogen bonded arrangements of purine, uracil, and xanthine bound to the PVDAT model with the surrounding uracil molecules, as they might help rationalize the outcome of the competitive adsorption experiments.[6] In addition, the issue of the additivity of the corresponding intermolecular interactions was addressed, which could also be important for an interpretation of the adsorption activity measurements.[6,7]

Computations

The quantum chemical modeling of the association of the PVDAT with nucleobases and related compounds was performed by applying the variational supermolecular approach to the interaction energy estima-

Institute of Macromolecular Chemistry, Academy of Sciences of the Czech Republic, Heyrovsky Square 2, Prague 162 06, The Czech Republic
E-mail: czernek@imc.cas.cz

tion for the optimized geometries of the reduced systems. Importantly, as the standard quantum chemical techniques[12] represented by the Hartree–Fock (HF) and the density-functional theory (DFT) methods are known (see reference[13] for the most recent discussion) to suffer from the improper description of the London dispersion contribution to the intermolecular interactions[14], more elaborate approaches, detailed below, were employed to investigate both the structure and energetics of the complexes. Thus, the PVDAT was approximated by the trimer structure capped with methyl groups (Figure 1). In the vicinity of the first and the last triazine unit, always the uracil molecule was placed, while in the central part of the trimer, one molecule from the set {purine; uracil; xanthine} was inserted. As a result, the structures consisted of 99, 98 and 101 atoms when dealing with the interactions in the system comprising the trimer and, accordingly, purine, uracil and xanthine in the presence of the two uracil molecules. For such models, a number of initial arrangements was prepared using the interactive computer graphics (program Insight II (2000), Accelrys Inc., San Diego, California), capturing various mutual orientations of the guests and conformations of trimer's backbone. They were subjected to the full geometrical optimization using the RI-BP-D approach (the resolution-of-the-identity integral approximation[15] to the DFT-based method with the BP86 exchange-correlation functional[16–18] and augmented with the empirical dispersion correction[19]) applied together with the SVP (single-valence plus polarization) atomic[20] and auxiliary[21] orbital basis sets (this technique will be abbreviated as RI-BP-D/SVP). The representative structures (see below) were verified to be minima of given potential energy hypersurface by means of the harmonic vibrational analysis.[12] The interaction energies, ΔE, in the resulting four-component systems *ABCD*, where *A* denotes the trimer and {*B*, *C*, *D*} the guest molecules as specified in Result and Discussion section, were established using the standard procedure[14] based on the counterpoise-corrected[22] RI-MP2 (the resolution-of-the-identity integral approximation[15] to the second-order Møller–Plesset perturbation theory) supermolecular calculations, which were performed with the cc-pVDZ (correlation-consistent polarized-valence double-zeta) atomic[23] and auxiliary[24] orbital basis sets. Both the HF, ΔE(HF), and correlation energy, ΔE(corr.), contributions to the total interaction energy, ΔE(tot.), $\Delta E(\text{tot.}) = \Delta E(\text{HF}) + \Delta E(\text{corr.})$, were extracted. All the calculations were carried out using the TURBOMOLE V5-10.[25] The geometries of all the structures can be obtained from the author upon request.

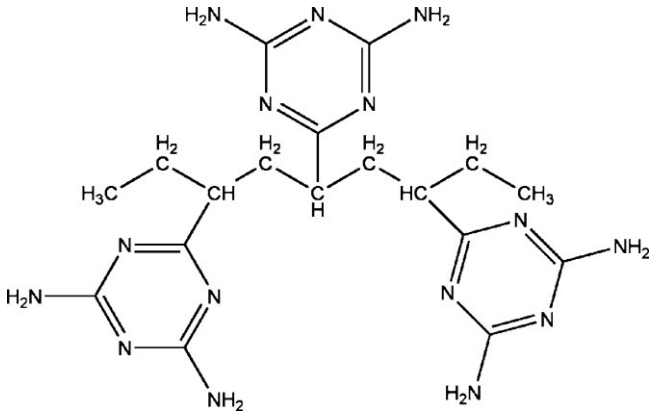

Figure 1.
The trimer of 2-vinyl-4,6-diamino-1,3,5-triazine used as the PVDAT model in the calculations.

Result and Discussion

Tables 1–3 summarize the results of the RI-MP2/cc-pVDZ calculations of the interaction energy components for the six RI-BP-D/SVP energy minima of the investigated complexes. Their geometries can be divided into two distinct groups. The first one, represented by the structures **1**, **3** and **5** (see Figures 2–4 accordingly), features the maximum number of hydrogen bonds between the PVDAT model and each of the guests. Thus, the complex **1** exhibits three hydrogen bonds between the host *A* and each of the uracils *B* and *D*, and also the hydrogen bonding contact of *A* with the purine molecule *C* (see Figure 2). The *AB* and *AD* interactions are of almost equal strength, and a little more than its one third amounts to the *AC* value (see Table 1). Interestingly, the mutual orientation of the uracil molecules is approximately T-shaped, and the resulting stabilization (ca. –10 kJ/mol) is due to the correlation energy (*cf.* the well-known case of an analogous arrangement in the benzene dimer[26]).

Table 1.
Details of the interactions in the systems comprising the PVDAT model, one purine and two uracil molecules (all values are in kJ/mol). See Figure 2 for the designation of the subsystems, and the text for details.

interaction type[a]	structure **1**			structure **2**		
	ΔE(HF)	ΔE(corr.)	ΔE(tot.)	ΔE(HF)	ΔE(corr.)	ΔE(tot.)
2b (AB)	−55.32	−17.72	−73.04	−50.91	−21.66	−72.58
2b (AC)	−17.76	−11.04	−28.80	29.48	−65.78	−36.30
2b (AD)	−59.38	−17.08	−76.46	−64.20	−16.49	−80.70
2b (BC)	1.45	−0.33	1.13	13.17	−18.65	−5.48
2b (BD)	0.28	−10.31	−10.03	1.11	−0.69	0.42
2b (CD)	1.16	−0.22	0.94	7.39	−15.11	−7.72
Σ *(2b)*	−129.56	−56.69	−186.25	−63.97	−138.38	−202.35
3b (ABC)	−0.63	0.09	−0.54	−0.72	0.62	−0.09
3b (ABD)	−1.08	0.40	−0.67	−0.34	0.05	−0.28
3b (ACD)	−0.33	0.03	−0.30	−2.19	0.94	−1.24
3b (BCD)	−1.08	0.13	−0.95	−0.20	0.04	−0.17
Σ *(3b)*	−3.11	0.66	−2.46	−3.44	1.66	−1.79
4b	0.98	−0.14	0.84	0.06	−0.01	0.05
all terms	−131.70	−56.17	−187.87	−67.35	−136.74	−204.09

[a] *2b*, *3b* and *4b* stand for two-body, three-body and four-body interaction energy term, respectively.

Table 2.
Details of the interactions in the systems comprising the PVDAT model and three uracil molecules (all values are in kJ/mol). See Figure 3 for the designation of the subsystems, and the text for details.

interaction type[a]	structure **3**			structure **4**		
	ΔE(HF)	ΔE(corr.)	ΔE(tot.)	ΔE(HF)	ΔE(corr.)	ΔE(tot.)
2b (AB)	−55.97	−17.62	−73.59	−45.93	−19.22	−65.16
2b (AC)	−60.67	−15.29	−75.97	7.34	−58.58	−51.24
2b (AD)	−59.38	−17.19	−76.57	−52.36	−17.33	−69.70
2b (BC)	0.55	−0.21	0.33	−8.61	−2.32	−10.92
2b (BD)	−1.30	−9.57	−10.86	0.56	−0.42	0.14
2b (CD)	0.47	−0.17	0.30	2.18	−9.20	−7.02
Σ *(2b)*	−176.29	−60.06	−236.34	−96.82	−107.08	−203.90
3b (ABC)	−0.04	0.02	−0.02	−2.34	0.42	−1.90
3b (ABD)	−0.07	0.29	0.22	0.12	−0.01	0.11
3b (ACD)	−0.02	0.01	−0.01	−1.19	0.51	−0.68
3b (BCD)	−0.04	0.03	−0.01	−0.11	−0.01	−0.12
Σ *(3b)*	−0.16	0.34	0.18	−3.52	0.93	−2.59
4b	−0.01	−0.02	−0.02	0.02	−0.01	0.01
all terms	−176.45	−59.73	−236.18	−100.32	−106.16	−206.48

[a] *2b*, *3b* and *4b* stand for two-body, three-body and four-body interaction energy term, respectively.

Table 3.
Details of the interactions in the systems comprising the PVDAT model, one xanthine and two uracil molecules (all values are in kJ/mol). See Figure 4 for the designation of the subsystems, and the text for details.

interaction type[a]	structure **5**			structure **6**		
	ΔE(HF)	ΔE(corr.)	ΔE(tot.)	ΔE(HF)	ΔE(corr.)	ΔE(tot.)
2b (AB)	−56.50	−17.42	−73.92	−59.23	−16.43	−75.67
2b (AC)	−58.86	−15.56	−74.43	−10.23	−63.81	−74.04
2b (AD)	−59.33	−17.22	−76.55	−58.87	−15.88	−74.75
2b (BC)	0.59	−0.26	0.33	11.44	−25.65	−14.20
2b (BD)	−0.96	−9.75	−10.71	1.01	−0.42	0.59
2b (CD)	0.37	−0.20	0.17	−0.14	−1.02	−1.16
Σ *(2b)*	−74.68	−60.42	−235.10	−116.01	−123.22	−239.23
3b (ABC)	−0.17	0.03	−0.14	−0.23	0.86	0.64
3b (ABD)	0.30	0.33	0.64	−0.13	0.02	−0.10
3b (ACD)	0.01	0.02	0.02	0.43	0.22	0.64
3b (BCD)	0.31	0.07	0.38	0.01	0.04	0.05
Σ *(3b)*	0.45	0.45	0.90	0.08	1.15	1.23
4b	−0.37	−0.06	−0.43	−0.02	−0.03	−0.05
all terms	−174.60	−60.02	−234.63	−115.95	−122.11	−238.05

[a] *2b*, *3b* and *4b* stand for two-body, three-body and four-body interaction energy term, respectively.

Clearly, all higher-order interaction energy terms can be considered as negligible (Table 1). As was the case for the complex **1**, the structures **3** and **5** also possess three hydrogen bonds between the PVDAT model and the terminal uracil molecules, which are also approximately perpendicular to each other. In addition, three such contacts are present in the mid part of the host, which binds uracil and xanthine, respectively. Expectedly, the values of all the *AB*, *AC* and *AD* terms are fairly similar (*cf.* Tables 2 and 3). All the three- and four-body ΔE components are close to zero, and

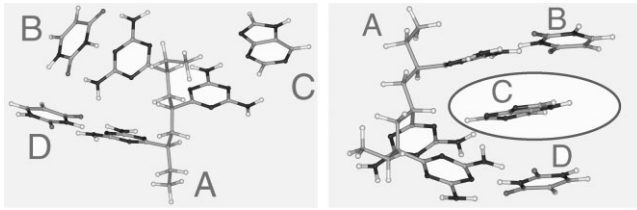

Figure 2.
The RI-BP-D/SVP optimized structures of the systems comprising the PVDAT model (*A*), one purine (*C*) and two uracil (*B* and *D*) molecules. Structure **1**, which maximizes the number of hydrogen bonds, is in the left panel. Structure **2**, which features a ''hydrophobic pocket'' (tentatively shown by the ellipse), is in the right panel.

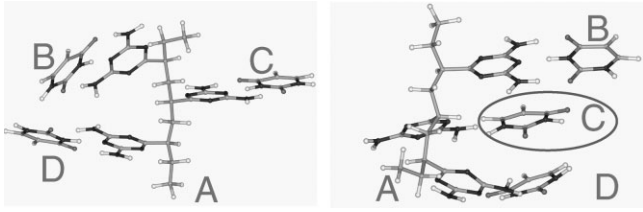

Figure 3.
The RI-BP-D/SVP optimized structures of the systems comprising the PVDAT model (*A*) and three uracil molecules (*B*, *C* and *D*). Structure **3**, which maximizes the number of hydrogen bonds, is in the left panel. Structure **4**, which features a ''hydrophobic pocket'' (tentatively shown by the ellipse), is in the right panel.

the same holds for the *BC* and *BD* terms (Tables 1–3). As a consequence, the total ΔE of the first group of structures (−188, −236 and −235 for **1**, **3** and **5**, respectively) can be considered as the additive sum of contributions of the hydrogen bonds and the *BD* term.

The RI-BP-D/SVP geometrical optimizations revealed also the second group of stabilizing structural arrangements. They contain hydrogen bonded uracils (*B* and *D*) on the flanks of the PVDAT model *A*, and a stacked guest molecule *C*, namely, purine, uracil and xanthine, respectively, in the complexes **2**, **4** and **6** (see Figures 2–4). The binding of *C* (*cf.* Tables 1–3) occurs predominantly via the contacts with the diaminotriazine units of *A* and by the stacking from the uracil molecules located above and below (the *BC* and *CD* contributions). This geometrical arrangement resembles a ''pocket'' in which the molecule of *C* is held predominantly by apolar interactions (*i.e.* the stacking of the nearby aromatic units). Such a binding motif, denoted as ''hydrophobic pocket'' in Figures 2–4, is highly stabilizing. Rather surprisingly, even the *AC* interaction alone can be as favorable as the formation of three hydrogen bonds. In particular, the *AB* and *AD* total interaction energies in the complexes **5** and **6** amount to, respectively, −76 and −75 kJ/mol, and −74 and −77 kJ/mol, while practically the same value (−74 kJ/mol) is obtained for the *AC* term, irrespective whether the xanthine molecule is hydrogen bonded to the PVDAT model (structure **5**) or located in the ''hydrophobic pocket'' of **6** (*cf.* Table 3 and Figure 4). Thus, for the models **2**, **4** and **6**, the *AC* term augmented with the *BC* and *CD* values (their sum exceeds 10 kJ/mol and is thus higher than the *BD* stabilization of the complexes **1**, **3** and **5** conferred by the T-shaped arrangement of the uracils) accounts for a significant portion of the overall stabilization. Namely, it equals to −49, −69 and −89 kJ/mol for **2**, **4** and **6**, respectively, while the contribution from the hydrogen bonding sites (the sum of *AB* and *AD*) is −154, −135 and −151 kJ/mol accordingly. As was the case of the first group of structures, the higher-order interaction energies of **2**, **4** and **6** are negligible, too. Their total ΔE (see Tables 1–3) can thus be approximated as being composed of the hydrogen bonding $(AB + AD)$ and the highly directional $(AC + BC + CD)$ components.

The computational results discussed above can be utilized in rationalizing, at a qualitative level, the outcome of the measurements. The adsorbing activity of PVDAT in water with respect to, accordingly, purine, uracil and theobromine was found[6] to increase as follows (arbitrary units): 0.05, 0.20 and 0.52, respectively. Thus, purine is adsorbed only to small extent. It possesses just one hydrogen bonding site, and the corresponding binding mode represented by the arrangement **1** confers much smaller total stabilization (−188 kJ/mol) than is the ΔE(tot.) of the stacked structure **2** (−204 kJ/mol, *cf.* Table 1 and Figure 2). This might indicate that, first, purine is adsorbed in water at ''type-II binding sites'' [6] (where, according to Komiyama et al.[6], the different number

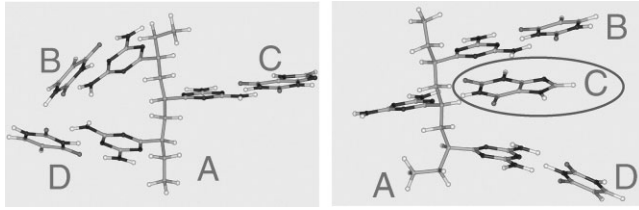

Figure 4.
The RI-BP-D/SVP optimized structures of the systems comprising the PVDAT model (*A*), one xanthine (*C*) and two uracil (*B* and *D*) molecules. Structure **5**, which maximizes the number of hydrogen bonds, is in the left panel. Structure **6**, which features a ''hydrophobic pocket'' (tentatively shown by the ellipse), is in the right panel.

of hydrogen bonding sites affects to a lesser extent the guest binding than for the ''type-I binding sites'', at which hydrogen bonding governs the adsorption). Second, this is in line with the experimental finding of only insignificant binding of purine by PVDAT in methanol,[7] where the number of the guest–diaminotriazine complementary hydrogen bonds is decisive for the binding activity and selectivity, and where the stacking interactions are minimized (see reference[9] for discussion). Uracil is predicted to be more strongly bound in the arrangement **3** than in **4** (the total ΔE of −236 and −206 kJ/mol, respectively). This agrees with the idea[6] of uracil molecules being adsorbed in the type-I binding sites in water (see above), and also with the precise recognition of uracil by PVDAT in methanol reported by Komiyama et al.[7] In order to reduce the computational cost, xanthine was used in the model instead of theobromine, whose adsorption was studied experimentally.[6] However, this replacement is expected to lead to only minor differences of the geometry of the hydrogen bonding and stacking sites. Indeed, theobromine can be bound to both type-I and type-II binding sites[6], which is in accord with quite similar ΔE(tot.) of −235 and −238 kJ/mol for the model structures **5** and **6**, respectively. In addition, because the difference of these values and the total stabilization of the arrangement **3** is small (less than one kcal/mol), one could anticipate the adsorption of uracil to be competitive to that of xanthine. This competition would likely be found experimentally, because the theobromine binding was measured[6] to be suppressed by the presence of thymine (which presumably occupies hydrogen bonding sites highly similar to those of uracil).

Conclusion

The enthalpic gain upon formation of the complexes of poly(vinyldiaminotriazine) model with purine, uracil and xanthine was assesses by means of the supermolecular RI-MP2/cc-pVDZ interaction energy calculations. Based on the underlying RI-BP-D/SVP geometrical optimizations, the major stabilizing structural factors were described. The results provide molecular-level insight into the recognition processes involving nucleic acid bases.

[1] S.-W. Kuo, *Polym. Int.* **2009**, *58*, 455.
[2] T. Haino, T. Fujii, A. Watanabe, U. Takayanagi, *Proc. Natl. Acad. Sci.* **2009**, *106*, 10477.
[3] W. H. Binder, R. Zirbs, *Adv. Polym. Sci.* **2007**, *107*, 1.
[4] H. M. Colquhoun, Z. Zhu, C. J. Cardin, Y. Gan, M. G. B. Drew, *J. Am. Chem. Soc.* **2007**, *129*, 16163.
[5] H. Asanuma, T. Ban, T. Hishiya, S. Gotoh, M. Komiyama, *Polym. J.* **1996**, *28*, 1024.
[6] H. Asanuma, T. Ban, S. Gotoh, T. Hishiya, M. Komiyama, *Macromolecules* **1998**, *31*, 371.
[7] H. Asanuma, T. Hishiya, T. Ban, S. Gotoh, M. Komiyama, *J. Chem. Soc., Perkin Trans. 2*, **1998**, 1915.
[8] O. Slinchenko, A. Rachkov, H. Miyachi, M. Ogiso, N. Minoura, *Biosens. Bioelectron.* **2004**, *20*, 1091.
[9] Z. Cao, W. Liu, D. Liang, G. Guo, J. Zhang, *Adv. Func. Mater.* **2007**, *17*, 246.
[10] R. P. Sijbesma, F. H. Beijer, L. Brunsveld, B. J. B. Folmer, J. H. K. K. Hirschberg, R. F. M. Lange, J. K. L. Lowe, E. W. Meijer, *Science* **1997**, *278*, 1601.
[11] L. Wu, Y. Li, *J. Molec. Recogn.* **2004**, *17*, 567.
[12] W. J. Hehre, L. Radom, P. v. R. Schleyer, J. A. Pople, ''*Ab initio Molecular Orbital Theory*'', Wiley, USA 1986.
[13] A. Krishtal, K. Vanommeslaeghe, A. Olasz, T. Veszprémi, C. V. Alsenoy, P. Geerlings, *J. Chem. Phys.* **2009**, *130*, Art. No. 174101.
[14] A. Stone, ''*Theory of Intermolecular Forces*'', Clarendon, Oxford 1996.
[15] O. Vahtras, J. Almlöf, M. W. Feyereisen, *Chem. Phys. Lett.* **1993**, *213*, 514.
[16] S. Vosko, L. Wilk, M. Nusair, *Can. J. Chem.* **1980**, *58*, 1200.
[17] A. D. Beck, *Phys. Rev. A* **1988**, *38*, 3098.
[18] J. P. Perdew, *Phys. Rev. B* **1986**, *33*, 8822.
[19] S. Grimme, *J. Comput. Chem.* **2004**, *25*, 1463.
[20] A. Schäfer, H. Horn, R. Ahlrichs, *J. Chem. Phys.* **1992**, *97*, 2571.
[21] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, *Chem. Phys. Lett.* **1995**, *242*, 652.
[22] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553.
[23] D. E. Woon, T. H. Dunning, Jr., *J. Chem. Phys.* **1995**, *103*, 4572.
[24] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, *Chem. Phys. Lett.* **1998**, *294*, 143.
[25] TURBOMOLE V5-10 2008, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmBH, 1989-2007, TURBOMOLE GmBH, since 2007; available from http://www.turbomole.com.
[26] P. Hobza, P. L. Selzle, E. W. Schlag, *J. Am. Chem. Soc.* **1993**, *116*, 3500.

