J Mol Model (2010) 16:1721–1729
DOI 10.1007/s00894-010-0722-8

ORIGINAL PAPER

Non-additive interactions of nucleobases in model dinucleotide steps occurring in B-DNA crystals

Piotr Cysewski

Received: 8 December 2009 / Accepted: 26 March 2010 / Published online: 8 May 2010

Abstract Non-additivity of base-base interactions in all ten possible model dinucleotide steps were analyzed on MP2/aug-cc-pvDZ quantum chemistry level. Conformations of four nucleobases exactly matched to ones occurring in B-DNA crystals. In most of thw 162 analyzed tetramers both three- and four-body contributions are negligible except for d(GpG) steps. However, in these dinucleotides both contributions are always of opposite signs and in all cases the sum of all non-additive part of intermolecular interactions do not exceed 2.6 kcal mol^{-1}. This stands for less than 5% of the overall binding energy of dinucleotide steps. Also replacements of guanine with 8-oxoguanine in d(GpG) systems introduces non-additivity of the same magnitude as for canonical dinucleotides. It is observed linear relationships between values of total binding energy obtained in the tetramer basis set and estimated energy exclusively in dimers basis sets with assumption of pairwise additivities. For all analyzed dinucleotides steps there are also linear correlations between amount of non-additive contributions and magnitude of pairs interactions. Based on differences in electrostatic contribution to the total binding energy of four nucleobases and polarity of dinucleotide steps three distinct classes of dinucleotide steps were identified.

Keywords B-DNA · Dinucleotide steps · Intermolecular interactions · Many-body · 8-oxoguanine

Electronic supplementary material The online version of this article (doi:10.1007/s00894-010-0722-8) contains supplementary material, which is available to authorized users.

P. Cysewski (✉)
Department of Physical Chemistry, Ludwik Rydygier Collegium Medicum in Bydgoszcz,
Nicolaus Copernicus University in Toruń,
Kurpińskiego 5,
85-950 Bydgoszcz, Poland
e-mail: piotr.cysewski@cm.umk.pl

P. Cysewski
Department of General Chemistry,
University of Technology and Life Sciences in Bydgoszcz,
ul. Seminaryjna 3,
85-326 Bydgoszcz, Poland

Introduction

Many-particle systems as immanent nature of physical world stand for emergence of new properties that cannot be derived just as simply adding up features of constituting elements. Non-negligible contributions of many-body terms to intermolecular interactions in molecular clusters are well known [1–4] and this aspect significantly complicates assessments of stabilization energies in multi-molecular systems. For example three-body terms have significant contribution in water clusters reaching even 30% to the interaction energy [2]. Fortunately, four-body and higher order terms are usually negligible [1–8]. This is of particular importance since almost all of the molecular modeling procedures ignore non-additive contributions to the total energy of molecular systems. A typical example is the contemporary molecular dynamics simulations for biomolecules. The size of these kinds of systems enable merely for application of force field parameterization inherently neglecting not only quantum effects but also all non-additivities of interacting systems. Implementing hybrid QM/MM methodology introduce improvement only to some extent. Furthermore, corrections for electron correlation are important contributions to the amount of individual n-body terms frequently imposing increase of their magnitudes. On the other hand many-body contributions can substantially

vary with the relative orientation of monomers [5, 6] and can originate from distinct types of interactions. For example in the case of water molecules energies of hydrogen bonding are two order stronger than for van der Waals interactions in noble gas clusters. Although for electrically neutral and non-polar systems non-additivity is typically small [7] for polar systems or condensed phases non-additivity can become substantial [8]. A special example of a system involving complicated interplay between many forces of different character [9–11] is found in the case of stabilization of polynucleotide chains. Many contributions are important for DNA involving contacts of its constituents as nucleobases, sugar moieties, phosphate backbones, surrounding water molecules, counter ions and diversity of ligands potentially present in solutions at physiological conditions. Even interactions between two heterocyclic nucleobases are non-homogeneous and typically are divided into three classes, namely hydrogen bonding, *intra*- and *inter*-strand stacking. It has been suggested that interaction between nucleic acids bases can in some extend exhibit such non-additive nature [12–16]. Although base-base interactions in dinucleotide steps of B-DNA are usually pair-additive but in the case of d(GpG) dinucleotide steps the absolute value of many-body contribution can be large and significantly structure-related [12–14]. As it was documented previously [12–16] high non-additive character of intermolecular interaction energy (IIE) has been found for many d(GpG) dinucleotide steps in B-DNA conformations. Although three- and four-body terms posses opposite signs they do not cancel each other and in most cases positive non-additivity is observed [15, 16]. The origin of such unique properties of d(GpG) steps is expected [12–16] to come from electrostatic repulsion between polar guanine-cytosine pairs which also was related to polymorphic changes of guanine-rich sequences of DNA double helices [16–19]. On the other hand much weaker electrostatics occurring in d(ApT) steps do not introduces significant non-additivity. However, small values of non-additive terms were also noticed [13, 14] for steps comprising two AT pairs. Many molecular clusters comprising for example water, ammonia or formamide molecules [2–4] exhibit similar sensitivity to polarization effects. Unfortunately, the level of theory used for characteristics of intermolecular interactions is a crucial factor determining magnitude of binding energies. This is related not only to level of electron correlation estimation [11–19], corrections of basis superposition errors (BSSE) [20] and quality of applied basis set [11–14] (or its extrapolation to complete one) but also to non-additive contributions to molecular cluster stabilizations [12–16]. As it is nowadays documented [16] accurate values for many-body effects require extensions of electron correlation beyond MP2 levels since this approximation can only encounter non-additivity *via* correlated exchange and induction components [21, 22]. However, due to fortunate compensation of errors such levels of theory as MP3, L-CCD or CCSD lead to substantially the same predictions as CCSD(T) one [16]. Due to the computer resources required for characteristics of nucleobases clusters there are only a few papers quantifying the non-additive nature of intermolecular interactions [13–15, 23–25]. Unfortunately the size of the analyzed systems prohibit application of advanced levels of theory as routine computations for a broad range of nucleobases conformations. Bearing in mind formal restrictions of MP2 level [21, 22] the qualitative estimation of non-additive nature of dinucleotide steps was characterized here for all ten unique dinucleotide steps adopting B-DNA conformations. Let's consider schematic representation of model dinucleotide step as it is presented in Fig. 1. The total binding of four nucleobases in such a system can be written in the following general form:

$$\Delta E^{total}_{d(XpY)} = E^{tetr}_{d(XpY)} - \sum_{i=1}^{4} E^{tetr}_{i} \tag{1}$$

The 5′-X and 3′-Y monomers present in strand I are numbered as 1 or 2, whereas corresponding complementary bases placed in strand II are labeled as 4 or 3, respectively. From the formal point of view all terms are to be estimated in the basis sets of the whole tetramer, what was indicated by *tetr* superscript. The decomposition of the total interaction energy into clusters of multi-body expansion can be defined as follows [26, 27]:

$$\Delta E^{total}_{d(XpY)} = \sum_{j>i=1}^{4} \varepsilon^{(2)}_{ij} + \sum_{k>j>i=1}^{4} \varepsilon^{(3)}_{ijk} + \varepsilon^{(4)}_{d(XpY)} \tag{2}$$

The analyzed dinucleotide steps are characterized by the following six two-body terms describing pairwise interactions between monomers:

$$\varepsilon^{(2)}_{ij} = E^{tetr}_{ij} - E^{tetr}_{i} - E^{tetr}_{j} \quad \{\mathrm{i,j} > \mathrm{i}\} = \{1,2,3,4\} \tag{3}$$

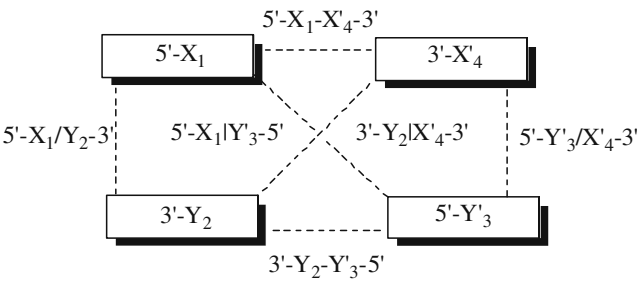

Fig. 1 The schematic representation of intermolecular interactions in model d(XpY) dinucleotide steps. There are three distinct types of contacts, namely *intra*-strand stacking (5'-X_1/Y_2-3', 5'-Y'_3/X'_4-3'), *inter*-strand stacking (5'-X_1|Y'_3-5', 3'-Y_2|X'_4-3') and hydrogen bonding (5'-X_1–X'_4-3', 3'-Y_2–Y'_3-5'), where X,Y denotes one of four nucleobases and X',Y' stands for corresponding complementary basex

The subscripts denote monomers according to notation used in Fig. 1. For the dinucleotide steps there are four tree-body terms defined as follows:

$$\varepsilon_{ijk}^{(3)} = E_{ijk}^{tetr} - E_i^{tetr} - E_j^{tetr} - E_k^{tetr} - \varepsilon_{ij}^{(2)} - \varepsilon_{ik}^{(2)} - \varepsilon_{jk}^{(2)} \quad \{i,j > i,k > j\} = \{1,2,3,4\} \tag{4}$$

Finally, the only four body term adopts one straightforward form:

$$\varepsilon_{d(XpY)}^{(4)} = E_{d(XpY)}^{tetr} - \sum_{i=1}^{4} E_i^{tetr} - \sum_{j>i=1}^{4} \varepsilon_{ij}^{(2)} - \sum_{k>j>i=1}^{4} \varepsilon_{ijk}^{(3)} \tag{5}$$

The additive part of the intermolecular interaction energy is usually defined as the sum of all two-body terms but all higher order components constitute the so-called non-additive part. The above scheme, although very simple and straightforward, is rather time-consuming and usually not-applicable for large number of dinucleotide conformations and advanced quantum chemistry levels. Thus, the most obvious simplification is neglecting all non-additive terms and estimation of all energies exclusively in the basis sets of dimers:

$$\Delta E_{d(XpY)}^{add,\dim} = \sum_{j>i=1}^{4} \left(E_{ij}^{dim} - E_i^{dim} - E_j^{dim} \right) \tag{6}$$

In such procedure monomers energies are recomputed separately for each pair but the overall CPU cost is about one order lower than full computations in tetramer basis set. The following difference quantifies the error of pairwise additivity of binding energy:

$$\Delta E_{d(XpY)}^{non-add} = \Delta E_{d(XpY)}^{total} - \Delta E_{d(XpY)}^{add,\dim} \tag{7}$$

The aim of this study is the quantitative analysis of magnitude of non-additive contributions to the total binding energy of dinucleotide steps and validation of assumption of pair-wise additivities of dimers stabilization energies. Besides, non-standard dinucleotide steps were considered that comprise one 8-oxoguanine molecule denoted by Z in two alternative sequences d(ZpG) or d(GpZ).

Methods

Model d(XpY) steps were prepared according to the procedure applied in our previous investigations [28–32] used for description of energetic heterogeneities in canonical and oxidized central guanine triad of B-DNA telomeric fragments [28, 29], polymorphism-related heterogeneities of guanine stacking in B- and A-DNA forms [30], characteristics of *inter*- and *intra*-strand stacking interactions in d(CpG) and d(GpC) steps found in B-DNA, A-DNA and Z-DNA crystals [31], quantification of all possible *intra*-strand stacking interactions between nucleobases [32] and assessment of many-body contributions to d(GpG) and d(CpC) dinucleotide steps [15]. The essential feature of applied approach is the characteristics of statistically significant number of conformations leading to distributions of intermolecular interaction energies (IIE) along with standard deviations and other statistical parameters. According to Standard Reference Frame [33] the complete description of three-dimensional arrangements of bases in nucleic acid structures is achieved by definition of values of base pair (*shear*, *stretch*, *stagger*, *buckle*, *propeller* and *opening*) and base pair step (*shift*, *slide*, *rise*, *tilt*, *roll* and *twist*) parameters. The values of these parameters corresponding to B-DNA crystals were taken from the Nucleic Acid Database [34] for all dinucleotide steps comprising canonical nucleobases. Preparation of model dinucleotide steps was done with an aid of X3DNA program [35], in which fragments library was modified by replacement of native geometries (without hydrogen atoms) with monomers pre-optimized on MP2/aug-cc-pvDZ (aDZ) level (imposing C_s symmetry). These monomers along with sets of 18 parameters (six for either hydrogen bonded pair and an additional six for stacked two-bases pairs) were used for preparation of tetramers in which orientations of nucleobases exactly correspond to ones found in the original B-DNA structures. In all cases the whole sugar-phosphate backbones were simplified just by hydrogen atoms. All ten unique model d(XpY) dinucleotide steps were used for single point energy estimation using density fitting MP2/aug-cc-pVDZ (aDZ) level of theory. The counterpoise correction for BSSE error [20] was included in all single point calculations of dimers, trimers, and tetramers. The MolPro package [36] was used in quantum chemistry calculations. Many body contributions were characterized according to formulas presented in the proceeding section.

Results and discussion

According to the procedure described above there were analyzed 162 model dinucleotide steps including 48 of d(GpG), 23 of d(ApA), 16 of d(CpG), 15 of d(TpA), 10 of either d(GpC), d(GpA), d(ApG), d(ApC), d(CpA) and d(ApT). All analyzed dinucleotide steps were described in Table S1 in supplementary material. As it was schematically presented in Fig. 1 there are three distinct types of intermolecular interactions that stabilize model dinucleotide steps. The most dominant contribution to the total binding comes from two-body interactions typically described as hydrogen bonding, *inter*- and *intra*-strand stacking. Three-body interactions consist of four distinct terms involving all combinations of *intra*-strand stacked

pair interacting with one nucleobase *via* hydrogen bonding. Finally, there is only one four-body term.

The magnitude of many body contributions

The calculated values of non-additive contributions to intermolecular interaction energy in ten possible model dinucleotide steps are presented in Figs. 2 and 3. First of all it is the visible distinct nature of d(GpG) steps since in most cases a significant non-additive part is obtained. For this step the distributions of both three- and four-body terms can vary significant as is presented in Fig. 2. The strongest attraction coming from three-body terms has been found for d(GpG) dinucleotide step in B-DNA coded as BD0085 and it is as high as -18.9 kcal mol^{-1}. The corresponding value of the four-body term is also unusually high and equals +20.2 kcal mol^{-1}. On the contrary in another utmost case (BD0015) an extremely high repulsion coming from three-body term (+24.2 kcal mol^{-1}) is noticed and an associated attraction originating from four-body term (-21.2 kcal mol^{-1}). However, these two examples are rare extremes since more than 90% of analyzed d(GpG) steps are characterized by three-body terms lower than +6 kcal mol^{-1}. The corresponding span of four-body contributions is from -5 kcal mol^{-1} to +2 kcal mol^{-1}. Closer inspection [15] reveals that not all three-body contributions are of the same magnitude for d (GpG) steps. All terms, except 134 one are usually quite small and typically do not exceed 2.0 kcal mol^{-1}. However, interactions stabilizing 134 trimer have a completely different nature. These contributions can have magnitude of several kcal mol^{-1} and are the main source of non-additivity in d(GpG) steps. This three-body cluster involves stacking of two cytosine molecules in strand II and hydrogen bonding of 5′-guanine present in strand I. Interestingly, although three-body contributions may have attractive or repulsive character, always an opposite sign of the four-body term is observed. However, this does not lead to cancelation of many body contributions and in case of d(GpG) step the total non-additivity is repulsive. As it was presented in Fig. 2 many-body contributions found for all other dinucleotide steps are much smaller and can be assumed as negligible since three-body terms are always within -1.8÷1.7 kcal mol^{-1} and four-body ones belong to interval -0.7÷0.2 kcal mol^{-1}. Since stabilization energy of model dinucleotide steps usually exceed -40 kcal mol^{-1} such small non-additivity affects at most modestly the overall biding of four nucleobases in d(XpY) steps. The median values of all non-additive contributions were presented in Fig. 3. Although all medians are positive the dinucleotide steps can be divided into three classes. As it was mentioned above the d(GpG) steps are characterized by highest repulsion coming from non-additive interactions. On the contrary for d (ApA), d(TpA) and d(ApT) dinucleotides nonadditivity although positive is close to zero. The third class comprises the rest of the dinucleotide steps, for which non-additivity is quite modest and repulsive. The statistical analysis applied to these three classes of dinucleotide steps proved that the differences in the median values are greater than would be expected by chance (P=<0.001). Hence all three dinucleotide distributions are statistically different. The normality test failed only for the first class comprising d(GpG) dinucleotides. The other two sets passed both normality (P=0.134) and equal variance tests (P=0.096).

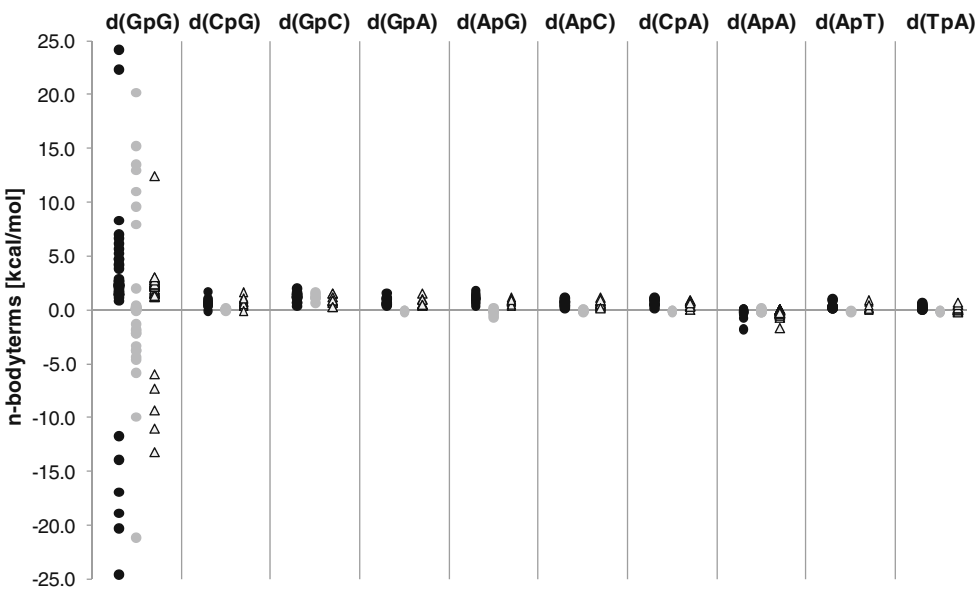

Fig. 2 The values of non-additive contributions to IIE energy of all ten possible model d(XpY) dinucleotide steps, where {(X,Y} = {A,G,C,T}. Black circles denote sum of all three-body contributions, gray circles correspond to four-body term and open triangles represent the sum of all non-additivities

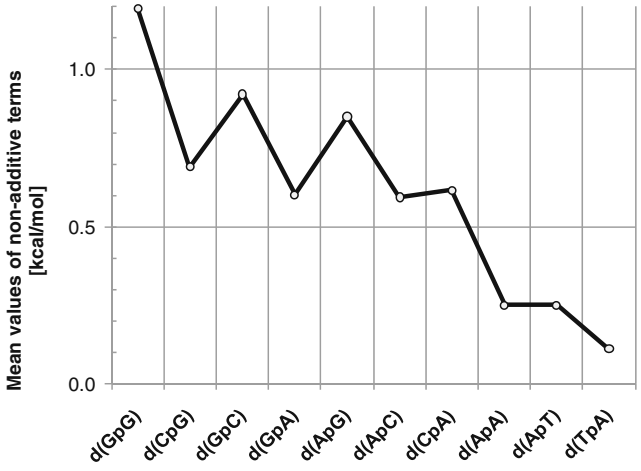

Fig. 3 The median values of all non additive terms estimated for ten possible dinucleotide steps

Linear relationships of intermolecular interactions

Since the intermolecular interactions between canonical and modified nucleobases are important aspect characterizing polynucleotide chains there is great temptation in assuming pairwise additivity of interactions stabilizing tetramers. This is justified not only by economy of computations but also by results presented in the previous section. However, in some cases, as for example for extreme conformations of d(GpG) steps, non-additive contributions can be significant. Thus, it is interesting to analyze the relationship between exact values of dinucleotide stabilization energies and simplified ones obtained after neglecting all non-additive terms in the binding energy of d(XpY) steps and utilization of exclusively dimer basis sets. Thus, the following formula is proposed for regression analysis:

$$\Delta E^{total,est}_{d(XpY)} = a \cdot \Delta E^{add,\dim}_{d(XpY)} + b \tag{8}$$

where a and b are regression coefficients and left-hand side estimates the total binding of nucleobases in d(XpY) step as the best fitted values to exact IIE of sum of pairwise additive terms estimated in dimer basis set. The applied regression analysis leads to plots provided in Fig. 4. As one could expect there is observed quite acceptable linear correlation between total binding energy and one estimated as simple sum of pairs energies (in the basis set of dimers). Even extreme cases for which dinucleotide steps are energetically unfavorable (+3.28 kcal mol^{-1} in case of d(ApC) found in BDL035 or with very small attraction -6.1 kcal mol^{-1} for d(ApA) (present in BDJ081) fulfill linear-dependence. At first glance one could recommend utilization of only one linear regression function in Fig. 4, but previously analyzed mean values of all non additive terms (see Fig. 3) suggested that three different classes of dinucleotide steps are to be considered. Since regressions presented in Fig. 4 allow for estimation of not only the total binding but also non-additive part it is interesting to see if linear relationships also hold for the latter quantity. Interestingly, due to different origin of non-additive parts in dinucleotide steps application of only one regression line would lead to very poor estimation of non-additive terms. On the contrary, as it was presented in Fig. 5, three relationships lead to a quite acceptable description of non-additive terms by exclusively pairs interactions. Hence, there are three

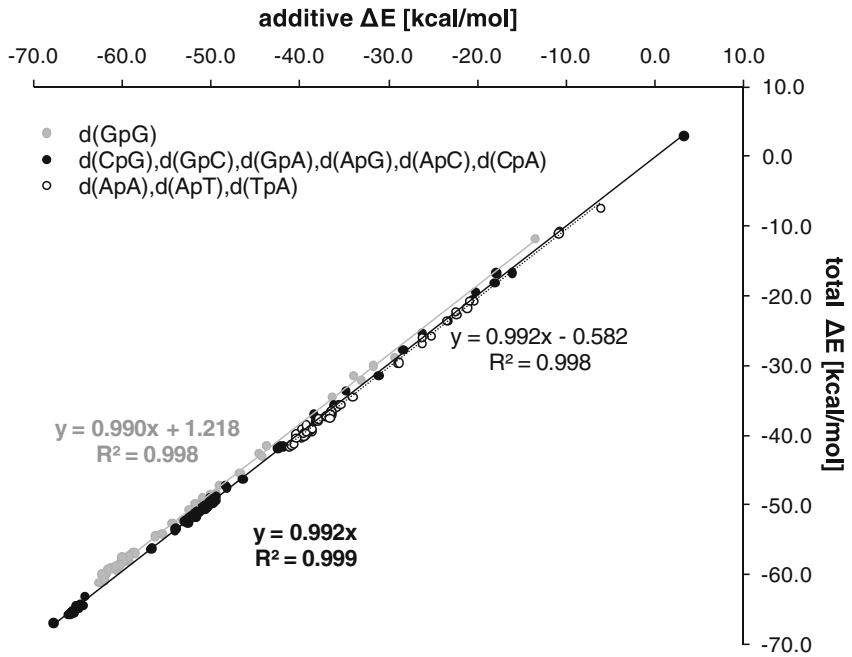

Fig. 4 The correlation between full estimation of total intermolecular interaction energies in dinucleotide steps including all non-additive terms in tetramer basis set $\left(\text{total}\Delta\text{E} = \Delta E^{total}_{d(XpY)} = E^{tetr}_{d(XpY)} - \sum_{i=1}^{4} E^{tetr}_{i}\right)$ and simplified assessment obtained after assumption of pairwise additivity of base-base interactions in dimer basis sets $\left(\text{additive}\Delta\text{E} = \Delta E^{add,\dim}_{d(XpY)} = \sum_{j>i=1}^{4}\left(E^{dim}_{ij} - E^{dim}_{i} - E^{dim}_{j}\right)\right)$

classes of dinucleotide steps with respect to many-body contributions to the total binding energy in dinucleotide steps. This can be further justified by values presented in Table 1, which collects the mean values of SCF and electron correlation contributions to the total values of IIE. Besides, the mean values of dipole moments of analyzed dinucleotide steps are also provided. Interestingly, although all SCF energies are negative indicating strong electrostatic contribution to the total binding energy, there are observed quite significant diversities between particular dinucleotide steps. First of all d(ApA), d(ApT) and d (TpA) systems are characterized by highest values of SCF term. Additionally, these tetramers belong to the less polar steps. On the contrary the d(GpG) is known [11, 12] as the most polar system among all ten possible dinucleotide steps, which also is confirmed by mean values of dipole moment collected in Table 1. Besides, much lower values of SCF contribution are noticed in this case if compared to previously described set of dinucleotides. The rest of the tetramers constitute a class of dinucleotide steps that are characterized by much stronger attraction coming from electron correlation energy and are of modest polarity. Thus, polarity expressed in terms of dipole moments along with electron correlation and SCF contribution to the total binding energy are quite satisfactory criterions of classification of dinucleotide steps into three distinct classes. Although the absolute values of non-additive terms seem to be negligible they are still an interesting source of distinguishing of dinucleotide steps.

Additivity of interactions in dinucleotide steps comprising 8-oxoguanine

Many products of nucleobases oxidation were identified in numerous irradiation and chemical experiments both *in vivo* and *in vitro* conditions of DNA and DNA-protein complexes (chromatin) [37, 38]. The use of gas chromatography–mass spectrometry with selected-ion-monitoring technique proved steady state levels of many free-radical derived analogues of canonical nucleobases [39]. They are believed to play a crucial role in ageing and different stages of cancer development [37, 38]. The structural and energetic consequences of B-DNA oxidation were not described till now. It is reasonable to expect that presence of 8-oxoguanine in B-DNA can significantly alter intermolecular interactions [40]. For example *intra*-strand stacking interactions of 8-oxoG with guanine as well as hydrogen bonding with cytosine are stronger compared to canonical guanine at least in telomere repeat units [28, 29]. It is interesting to see if modification of guanine leads also to any non-additivities in base-base interactions occurring in oxidized dinucleotide steps. Since a sufficient amount of experimental data comprising 8-oxoguanine is not available, alternatively molecular dynamics simulations were used for collecting structural parameters defining oxidized dinucleotide steps. Details are provided elsewhere [40]. The decomposition of dinucleotide stabilizing energies into many-body contributions can be found in Fig. 6. Presence of 8-oxoguanine in d (ZpG) or d(GpZ) dinucleotide steps introduces significant

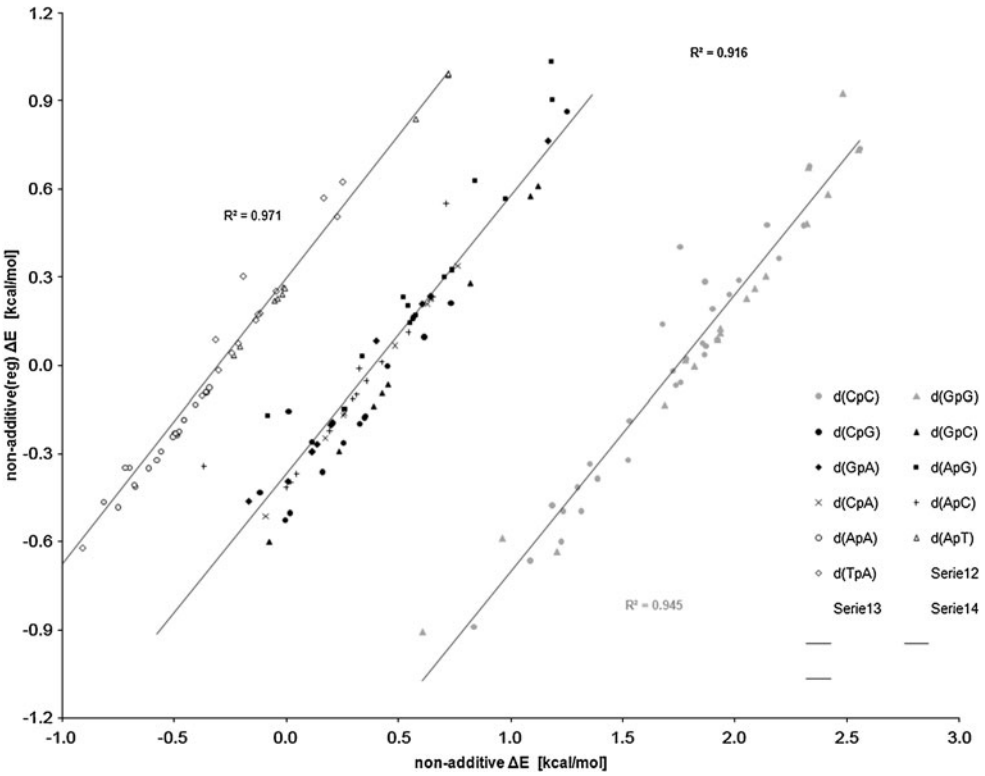

Fig. 5 The applicability of linear relationships obtained in Fig. 4 for estimation of non-additive contributions based on pair interactions estimated using dimer basis sets. As abscissa $\text{non}-\text{additive}\ \Delta E = \Delta E^{total}_{d(XpY)} - \Delta E^{add,\dim}_{d(XpY)}$, is used, while ordinate defines $\text{non}-\text{additive(reg)}\ \Delta E = \Delta E^{total}_{d(XpY)} - \left(a \cdot \Delta E^{add,\dim}_{d(XpY)} + b\right)$

Table 1 The mean values of dipole moments, SCF and electron correlation contributions to the total binding energy of model dinucleotide steps. Energies are expressed in kcal mol^{-1} and dipole moment are given in Debyes. Bold face was used for demonstration of differences between three distinct classes of dinucleotide steps

Class	Step	$\overline{\Delta E}^{total,SCF}_{d(XpY)}$	$\overline{\Delta E}^{total,correl}_{d(XpY)}$	Dipole moment
I	d(GpG)	-16.9±13.1	-51.0±10.7	11.8±1.3
II	d(CpG)	-24.7±16.0	-56.7±13.6	2.7±1.3
	d(GpC)	-29.2±1.9	-64.9±1.3	4.8±0.4
	d(GpA)	-13.2±6.9	-47.7±4.7	7.2±0.3
	d(ApG)	-11.3±13.5	-44.2±12.0	8.4±0.3
	d(ApC)	-6.2±23.5	-44.2±16.8	4.5±0.5
	d(CpA)	-23.3±2.2	-52.0±0.4	4.5±0.3
III	d(ApA)	-2.2±6.3	-38.3±4.7	3.9±0.4
	d(ApT)	-0.9±4.7	-37.6±5.6	1.6±1.3
	d(TpA)	0.3±12.2	-32.3±8.4	1.7±0.4

changes upon non-additivity. First of all, unlike the canonical d(GpG) sequence, both three- and four-body contributions are usually of the same sign. Thus, although actual values of non-additive terms are smaller if compared to canonical guanine dinucleotides it is expected that cumulative effect occurs. The mean values of all three-body terms are about +2.0 and +2.5 kcal mol^{-1} for d(ZpG) and d(GpZ) dinucleotide steps, respectively. The mean values of four body contributions are much smaller and do not exceed 0.13 kcal mol^{-1} in either case. Thus, many-body contributions are of repulsive character reducing total stabilization of the two stacked pairs. The mean values of non-additivities characterizing oxidized dinucleotide steps can reach 5% of total IIE. This is a similar

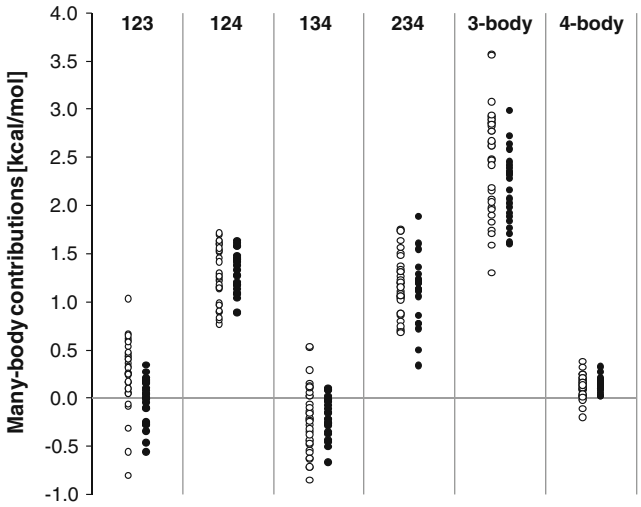

Fig. 6 Distribution of many body contributions to intermolecular interaction in oxidized d(GpG) dinucleotide steps. Open circles denote d(GpZ) steps, while black circles stand for d(ZpG) sequence, where Z denotes 8-oxoguanine

amount to one found in the case of d(GpG) steps [15]. No sequence dependence of many body contributions is observed for oxidized dinucleotides. There is another interesting features of three body terms found for modified dinucleotide steps. A majority of 123 and 134 contributions are systematically smaller than 124 and 234 ones. The common patterns in these terms are formed by purine-pyrimidine *inter*-strand stacking ($G_1|C_3$, $Z_1|C_3$, $C_2|G_4$ or $C_2|Z_4$). It seems that 13 interactions introduces attractions in analyzed steps, while 24 *inter*-strand interactions are responsible for repulsive contributions irrespectively of the analyzed sequences. In the case of d(GpG) this feature was not observed [15] since major contributions to the total non-additivity came from the 134 term. Finally, it is worth mentioning that there are linear relationships between full estimation of IIE values and ones utilizing pair-wise additive assumption. The applied linear regression analysis led to quite acceptable correlation with R^2=0.983 for oxidized dinucleotide steps and R^2=0.996 for canonical ones. This relationship is presented in subsequent study of this issue [40]. Finally, the total binding of four nucleobases in oxidized d(GpG) steps is about 2 kcal mol^{-1} stronger than found for canonical dinucleotide.

Conclusions

Results presented above qualitatively characterize non-additivity of base-base interactions stabilizing d(XpY) dinucleotide steps. Although the applied model excluded sugar-phosphate backbone, the most important features of *intra*-and *inter*-strand stacking along with hydrogen bonding were properly preserved. The most important aspects of the presented data is rational representation of four bases conformations and adequate level of computations. The first facet relies on the representation of meaningful conformations matching to ones occurring in B-DNA crystals. The main conclusion drawn from presented results is that pair-wise additivity simplification is rational since corresponding intermolecular interaction energy of d(XpY) steps suffers far less than 5% if compared to correct values. The four- and three- body terms are small and usually of the opposite signs, what further reduces its impact on the total binding energy. Dinucleotide steps comprising two stacked guanine molecules seem to be the only canonical systems for which non-additivity is non-negligible. However, as it was demonstrated in the previous report [15], in this case there are observed linear relationships which allow for corrections of IIE values without necessity of energy computations for the whole tetramer. Despite the fact that many body contributions are small, they can be found as an interesting source of classification of dinucleotide steps

into three distinct classes. First class comprised d(GpG) steps that were found as the most polar among all dinucleotide steps. To the second class one can include d(ApA), d(TpA and d(ApT) steps. They are characterized by the highest values of SCF contributions to the total IIE and lowest polarities. The rest of the dinucleotide steps constitute the third class with modest polarity and electrostatic contribution to the total interaction energy. Oxidation of guanine in d(GpG) dinucleotide steps introduces non-negligible non-additivity to IIE that is of the same order as observed for canonical systems. In the case of many potential structures contribution to analyzed quantity there is always the question related to arbitral selection of analyzed systems. It is impractical to analyze hundreds of d(XpY) conformations and misfortune sampling may significantly affect conclusions. However, in the previous paper [15] it was demonstrated that there is a simple origin of non-additivity in the case of d(GpG) dinucleotide steps. The interactions between hydrogen bonded nucleobases as well as *inter*-strand stacking interactions can vary significantly for different conformations of four bases. As the quantitative measure of such feature the *energy asymmetry coefficient* (δ) is proposed. It is defined as the sum of the two following terms. The first one denotes the difference in hydrogen bonding energies between two neighboring complementary nucleobases and the second one characterizes the difference between 5-G|C′-5′ and 3-G|C′-3′ *inter*-strand stacking interactions. In the case of d(GpG) dinucletode steps this parameters emphasizes the importance of hetero-bases contacts. There is a straightforward way of extension of δ coefficient for other tetramers and to characterize the symmetry/asymmetry diversities (in energy sense) of all dinucleotide steps. In any case close to zero values of δ coefficient indicate high symmetry of intermolecular interactions. On the contrary the positive values of this parameter suggest stronger interactions on 3′ side and vice versa negative values emphasizes stronger attraction on 5′-side. In Fig. 7 there is presented distribution of δ values estimated for analyzed dinucleotide steps. As is visible all analyzed systems may adopt a wide range of energetic asymmetry covering a broad range of structures. Hence, despite analogical distributions of energetic asymmetry the high non-additive contributions are observed only for canonical and oxidized d(GpG) steps. This suggests that the selection of dinucleotide conformations is adequate and presented conclusions are expected to be consistent also for other conformations of dinucleotide steps. The results presented in this paper are preliminary and further investigations are necessary. Several aspects were excluded from the analysis that probably can affect both stabilization energies of dinucleotide steps and many-body contributions. Among them the applied level of computations is the first way of improvement. More sophisticated methods should be applied for proper inclusion of electron correlations. Although, the presented results embrace only a small portion of this effect due to applied level of computations it is reasonable to expect that extension of the method will not change the general conclusion of the magnitude of many-body contributions to dinucleotide steps energetic. Besides there are other aspect worth additional considerations. For example it is commonly known that nucleobases comprising amino groups are not planar. These exocyclic substituent exhibit pyramidalization, which probably affects both IIE and many-body terms. Also relaxation of monomers geometry, especially hydrogen bonds are worthy of precise evaluation. The importance of all these aspects is to be validated in further investigations.

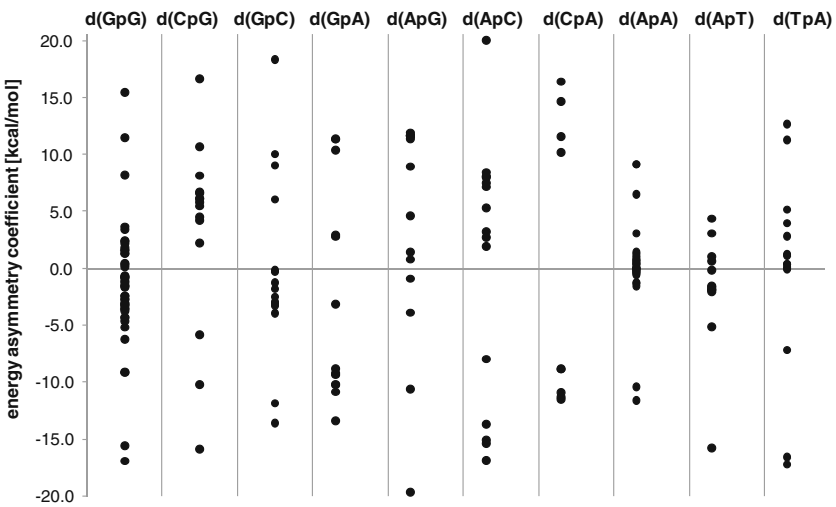

Fig. 7 The distribution of *energy asymmetry coefficient* (δ) estimated for all of analyzed dinucleotide steps

Acknowledgments Results were obtained as part of computational grant no 39 of Poznań Supercomputing and Networking Center (PSNC, Poznań, Poland). The allocation of computational facilities are greatly appreciated.

References

1. Elrod MJ, Saykally RJ (1994) Chem Rev 94:1975–1997
2. Xantheas SS (1994) J Chem Phys 100:7523–7534
3. Suhai SJ (1994) Chem Phys 103:7030–7039
4. Saykally R (1996) Science 271:929–933
5. Habitz P, Bagus P, Siegbabn P, Clementi E (1983) Int J Quantum Chem 23:1803–1806
6. Chalasinski G, Szczesniak MM, Cieplak P, Sheiner S (1991) J Chem Phys 94:2873–2883
7. Bukowski R, Szalewicz K (2001) J Chem Phys 114:9518–9531
8. Bukowski R, Szalewicz K, Groenenboom GC, van der Avoird A (2008) J Chem Phys 128:094314
9. Saenger W (1998) Principles of nucleic acid structures. Springer, Berlin
10. Bloomfield VA, Crothers DM, Tinoco I (1999) Nucleic acids: structures, properties and functions. University Science Books, Sausalito, CA
11. Hobza P, Šponer J (1999) Chem Rev 99:3247–3276
12. Šponer J, Riley KE, Hobza P (2008) Phys Chem Chem Phys 10:2595–2610
13. Šponer J, Gabb HA, Leszczynski J, Hobza P (1997) Biophys J 73:76–87
14. Pitoňák M, Neogrády P, Hobza P (2010) Phys Chem Chem Phys 12:1369–1378
15. Cysewski P (2010) Int J Quantum Chem. doi:10.1002/qua.22435
16. Svozil D, Hobza P, Šponer J (2010) J Phys Chem B 114:1191–1203
17. Lankas F, Cheatham TE, Špaĉkova N, Hobza P, Langowski J, Šponer J (2002) Biophys J 82:2592–2609
18. Ng HL, Dickerson RE (2002) Nucleic Acids Res 30:4061–4067
19. Šponer J, Florian J, Ng HL, Šponer J, Špaĉkova N (2000) Nucleic Acids Res 24:4893–4902
20. Boy SF, Bernardi F (1970) Mol Phys 19:553–556
21. Hesselmann A (2008) J Chem Phys 128:144112
22. Chalasinski G, Szczesniak MM (1994) Chem Rev 94:1723–1765
23. Peters M, Rozas I, Alkorta I, Elguero J (2003) J Phys Chem B 107:323–330
24. Gu J, Wang J, Leszczynski J (2004) J Phys Chem B 108:8017–8022
25. Biot C, Wintjens R, Rooman M (2004) J Am Chem Soc 126:6220–6221
26. Hankins D, Moskowitz JW, Stillinger FH (1979) J Chem Phys 53:4544–4554
27. Mierzwicki K, Latajka Z (2000) Chem Phys Lett 325:465–472
28. Cysewski P, Czeleń P (2007) J Mol Model 13:739–750
29. Cysewski P, Czeleń P (2009) J Mol Mod 15:607–613
30. Cysewski P (2008) J Mol Struct Theochem 865:36–43
31. Cysewski P (2009) J Mol Model 15:597–606
32. Cysewski P (2009) New J Chem 33:1909–1917
33. Olson WK, Bansal M, Burley SK, Dickerson RE, Gerstein M, Harvey SC, Heinemann U, Lu XJ, Neidle S, Shakked Z, Sklenar H, Suzuki M, Tung CS, Westhof E, Wolberger C, Berman HM (2001) J Mol Biol 313(1):229–237
34. Berman HM, Olson WK, Beveridge DL, Westbrook J, Gelbin A, Demeny T, Hsieh S-H, Srinivasan AR, Schneider B (1992) Biophys J 63:751–759
35. Lu XJ, Olson WK (2003) Nucleic Acids Res 31:5108–5121
36. Werner HJ, Knowles PJ, Lindh R, Manby FR, Schutz M, Celani P, Korona T, Rauhut G, Amos RD, Bernhardsson A, Berning A, Cooper DL, Deegan MJO, Dobbyn AJ, Eckert F, Hampel C, Hetzer G, Lloyd AW, McNicholas SJ, Meyer W, Mura ME, Nicklaß A, Palmieri P, Pitzer R, Schumann U, Stoll H, Stone AJ, Tarroni R, Thorsteinsson T (2006) MOLPRO, a package of ab initio programs designed by Werner HJand Knowles PJ. Version 2006.0 (Patch 2006.1) Cardiff, UK
37. von Sonntag C (2006) Free-radical-induced dna damage and its repair, a chemical perspective. Springer, Heidelberg
38. Rice-Evans C, Halliwell B, Lunt GG (1995) Free radicals and oxidative stress: environment Drugs and food additives. Portland, London
39. Halliwell B, Gutteridge JMC (1999) Free radicals in biology and medicine. Clarendon, Oxford
40. Cysewski P, Czeleń P (in press) J Mol Mod. doi:10.1007/s00894-010-0730-8