Bulgarian Chemical Communications, Volume 49, Special Issue D (pp. 8-18) 2017

Hydrogen bonding reactivities of atomic sites in the nucleobases

V. Nikolova, D. Cheshmedzhieva, S. Ilieva, B. Galabov*

Department of Chemistry and Pharmacy, University of Sofia, James Bourchier Avenue, Sofia 1164, Bulgaria

Received March 27, 2017; Revised April 10, 2017

Dedicated to Acad. Ivan Juchnovski on the occasion of his 80th birthday

Theoretical parameters characterizing the electrostatic forces at individual atoms were applied in quantifying the hydrogen bonding reactivities of atomic sites in the five primary nucleobases. To avoid ambiguities associated with the usual formation of bidentate hydrogen bonds in the nucleobases, DFT M06-2X/6-311+G(2d,2p) computations were employed in deriving electrostatic potentials at nuclei (EPN), NBO, and Hirshfeld atomic charges for a set of 64 model molecules. For comparison, proton affinities for the proton-accepting sites and deprotonation energies for proton-donating sites also were evaluated. The functional groups in the selected model molecules are able to form single hydrogen bonds. The obtained correlation equations linking computed interaction energies and electrostatics-related parameters were applied in deriving reactivity descriptors for individual atomic sites in the nucleobases. Among the tested theoretical parameters, the EPN values provide the best correlations with hydrogen bonding energies. The derived relationships were employed in defining a hydrogen bond reactivity descriptor for the atomic sites in the nucleobases. The effects of base pairing on site reactivities also were assessed.

Key words: hydrogen bonding, nucleic acid bases, electrostatic potential, atomic charges, reactivity descriptors.

INTRODUCTION

The ability of different atomic sites in the purine and pyrimidine nucleic acid bases to form hydrogen bonds determines a number of key properties of DNA and RNA. These include the formation of double helix structure, the replication mechanisms, and the interactions with proteins and biologically active ligands [1-10]. In the present research we present an efficient computational method for evaluating hydrogen bonding reactivities of individual atomic sites in the nucleobases. The effects of base pairing on hydrogen bonding abilities are also considered.

Numerous theoretical studies have focused on quantifying the proton-accepting and proton-donating abilities of the purine and pyrimidine nucleobases [11-20]. Zeegers-Huiskens et al.[17-20] analysed using DFT computations the effectiveness of intrinsic basicities and acidities of the hydrogen bonding sites in the nucleobases in describing the interactions with a single water molecule. In a computational study Medhi et al. [21] characterized the reactivities of nucleobases by determining proton and metal ion affinities of the nucleobases accepting sites. In several theoretical studies Wetmore et al. [11-14] examined the hydrogen bonding abilities of the nucleobases in their interactions with small hydrogen bond forming molecules: HF, H_2O and NH_3. A barrier in quantifying the hydrogen bonding abilities of individual atomic sites arises from the particular structure of nucleobases. The interactions of molecules with proton accepting or donating sites in the bases involve in most cases the formation of cooperative bidentate hydrogen bonds with participation of neighbouring atoms. Thus, energy differentiations between these two simultaneous interactions cannot be evaluated from suitable experiments or from theoretical modelling. Information on site reactivities is particularly useful for QSAR and docking studies as well as in understanding fine mechanisms of interactions involving nucleic acid fragments.

Electrostatic forces have been shown to play a major role in defining the structure and properties of biopolymers [1-10,25,26]. Electrostatic interactions are also among the key factor in hydrogen bonding in general [1,8, 9, 27, 28]. It was of interest to follow how quantities, reflecting the electrostatics at different sites in the nucleobases, would be linked with their hydrogen bonding reactivities. In the present research, we evaluate three types of molecular parameters that are intrinsically related to site electrostatics: NBO atomic charges [29], Hirshfeld charges [30], and electrostatic potentials at nuclei (EPN) [22]. In addition, following the already established

---

* To whom all correspondence should be sent:
E-mail: *e-mail: galabov@chem.uni-sofia.bg*

methodology [17-20], we also determined the proton affinities (PAs) and deprotonation energies (DEs) for the proton accepting and donating sites in the primary nucleobases at the level of theory employed in the present investigation. It thus became possible to assess how well all five theoretical parameters considered correlate with hydrogen bonding energies.

Earlier studies from this laboratory [31-36] revealed the accuracy of electrostatic potential at nuclei (EPN) in characterizing the site reactivities in hydrogen bonding. In a recent investigation [37] on substituted benzenes, we showed that molecular electrostatic potential (MEP) values, evaluated for points in molecular space in close proximity to the ring carbon nuclei, describe quantitatively site reactivities of the respective positions. Using Morokuma energy decomposition analysis [38,39] for selected series of hydrogen bonded molecules we also showed [35] that the different terms contributing to interaction energies (polarization interaction, exchange repulsion, charge transfer) are *linearly correlated* with the electrostatic energy term. These results emphasize the importance of molecular parameters characterizing site electrostatics in quantifying hydrogen bonding reactivities.

To avoid the difficulty of multiple hydrogen bonding of model proton donor and acceptor molecules (such as HF, H_2O and NH_3) with neighbouring atoms in the nucleobases, we conducted computations for series of molecules containing functional groups in environments that resemble the structural arrangements in the nucleobases. The functional groups in the selected model molecules are able to form single hydrogen bonds. The established correlations between reactivity descriptors and interaction energies were than employed in characterizing quantitatively the reactivity of individual proton-accepting and proton-donating centers in the nucleic acid bases.

COMPUTATIONAL METHODS

DFT computations employing the MO6-2X functional [40] combined with the 6-311+G(2d,2p) basis set [41] were applied in evaluating hydrogen bonding energies and a set of molecular parameters characterizing local properties of atomic sites in the investigated systems. Both DFT method and basis set employed have been shown to provide reliable results for the energies of hydrogen bonding [42]. Harmonic vibrational frequency computations showed that the optimized structures are true minima in the potential energy surfaces. Zero point vibrational energies (ZPE) and bases set superposition error (BSSE) corrections [43] were applied in evaluating the energies of complex formation. The Gaussian09 program [44] was employed for all computations. Site electrostatics at the proton accepting and donating centers in the nucleobases was characterized by EPN values, NBO and Hirshfeld charges. The electrostatic potential at nuclei was first introduced by Wilson [45]. Politzer and Thruhlar [22] defined the electrostatic potential at nuclei Y (V_Y) by eqn (1):

$$V_Y \equiv V(R_Y) = \sum_{A \neq Y} \frac{Z_A}{|R_A - R_Y|} - \int \frac{\rho(r)}{|r - R_Y|} \qquad (1)$$

In this relationship, the singular term for nucleus Y is excluded. Z_A is the charge of nucleus A at position R_A, and $\rho(r)$ is the electron density function. As mentioned, following the original findings [31-34] that EPN values define quantitatively the ability of molecules to form hydrogen bonds, the EPN index was extensively applied in describing both hydrogen bonding and chemical reactivity of various molecular systems [35,36,44-51]. In contrast to other theoretical parameters that characterize site properties, the electrostatic potential at nuclei is a rigorously defined quantum mechanical quantity. The 1/r dependence of EPN (Eqn 1) determines considerably greater contributions to V_Y of negative and positive charges in close vicinity of nucleus Y.

To have a clearer physical interpretation of the variations of EPN (V_Y), we employed the shifts of EPN in the studied derivatives (X) with respect the values in selected small molecules, containing the same atoms in the respective hybridization state. For this purpose, we selected ammonia (sp^3-nitrogen), methanimine (sp^2-nitrogen), formaldehyde (sp^2-oxygen), and hydrogen molecule (for H atoms). The following relations were employed:

For N(sp3): $\Delta V_N = V_N(X) - V_N(NH_3)$ (2)

For N(sp2): $\Delta V_N = V_N(X) - V_N(H_2C{=}NH)$ (3)

For O(sp2): $\Delta V_O = V_O(X) - V_O(H_2C{=}O)$ (4)

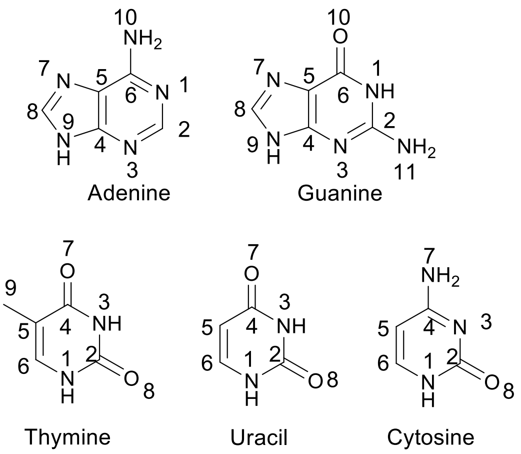
For H: $\Delta V_H = V_H(X) - V_H(H_2)$ (5)

In the tables, we present the shifts of EPN values as defined in Eqns. (2-5). The proton affinities (PA) of all nitrogen and oxygen atomic centers for the series of model nitrogen and oxygen containing compounds as well as for the nucleobases were evaluated as the difference of the ZPE corrected energies of protonated and neutral molecules. Deprotonation energies were also determined for the proton donating N-H sites.

METHODOLOGICAL APPROACH

Nitrogen and oxygen atomic centers are the principal proton accepting centers in the nucleobases able to form hydrogen bonds, while the N-H bonds may act as proton donors. As emphasized, it is not possible to evaluate theoretically the hydrogen bonding reactivities of individual atomic centers in the nucleobases by analyzing simply the energies of complex formation with model molecules. In most cases, proton donating or accepting molecules interact cooperatively with two atoms in the nucleobases [11-20]. These processes are illustrated in Fig. 1 for the interaction of thymine with HF, H_2O and C_2H_2. It is seen that two cooperative H- bonds are simultaneously formed. Even if a weaker proton donor such as C_2H_2 is employed in modeling the hydrogen bonding with a basic center, the neighboring proton donating groups contribute to the overall complexation energy by forming a π-hydrogen bond (Fig. 1C). Our approach involves four steps:

(1) The basic assumptions is that correlations between hydrogen bonding energies and molecular parameters established for model systems may be also applied for the respective functionalities in the nucleobases, following a verification. In general, similarities of the molecular environment of proton accepting and donating groups in the model compounds and in the nucleobases suggest that this is a well-based hypothesis. Nonetheless, to verify further its validity we analyzed the relationships between the employed molecular parameters (atomic charges, EPN values) with computed chemical properties of the respective atomic sites: proton affinities and deprotonation energies.



Scheme 1. Structure and numbering of atomic positions in the primary nucleobases.

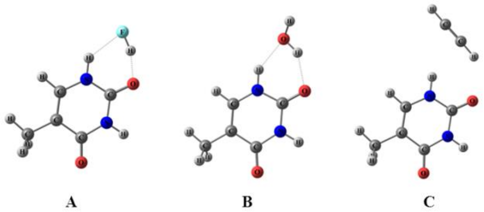
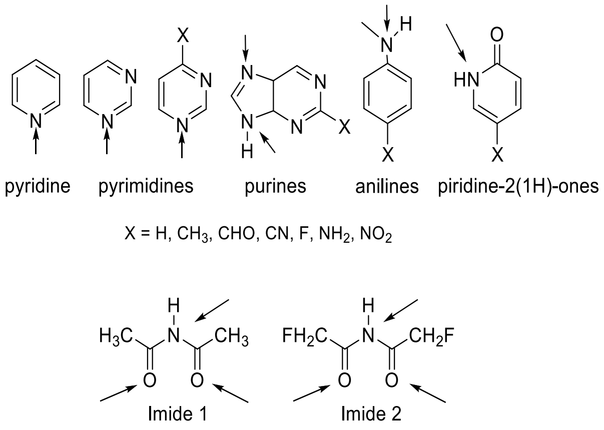


Fig. 1. Bidentate hydrogen bonding in thymine complexes with HF, H_2O and C_2H_2 from M06-2X/6-311+G(2d,2p) computations.

These correlations include both the model compounds and the nucleobases. Based on the obtained results, the most appropriate molecular parameter for the further analyses was selected. The criterion employed was how accurately the respective quantity (NBO and Hirshfeld charges, EPN) describes the local chemical reactivity of atomic sites.

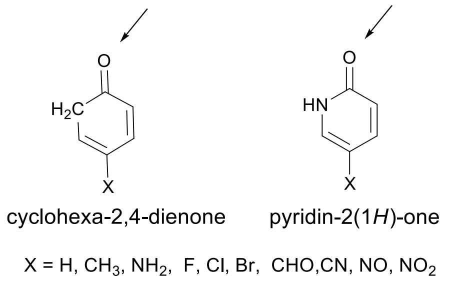
(2) In the second step, we studied in detail the relationships between energies of hydrogen bonding and molecular parameters for the sets of model compounds, containing the same functional groups as the nucleobases. We considered correlations of interaction energies with the already selected site-related parameter as well as with proton affinities and deprotonation energies. Series of substituted pyridine, pyrimidine, purine, aniline, pyridine-2(1H)-one, and imide derivatives (Scheme 2) were investigated computationally to characterize the hydrogen bonding abilities of nitrogen atoms in varying molecular environment.



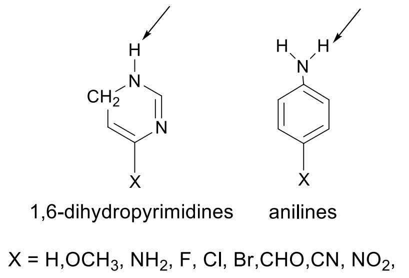
Scheme 2. Structures of model sets of molecules containing nitrogen atoms in varying molecular environment.

Computations for unsaturated cyclic amides and cyclodienones (Scheme 3) were conducted in characterizing the reactivities of carbonyl oxygen atoms. The proton affinities for the C=O oxygens were also evaluated. Finally, a suitable set of derivatives (Scheme 4) was selected for analogous computations, aimed at characterizing the proton-donating properties of N-H groups in hydrogen bonding. The arrows also indicate the chosen direction of hydrogen bond formation. We employed hydrogen fluoride as a model proton donor/proton acceptor in these computations. The established relationships between molecular quantities and interaction energies were than employed in characterizing the ability of individual atomic sites in the nucleic acid bases to form hydrogen bonds. The accuracy of theoretical predictions of hydrogen bonding energies by the different theoretical parameters was assessed.

(3) In the third step of our approach, we applied scaling procedures, aimed at generalizing the relationships between theoretical parameters and hydrogen bonding reactivities. The scaling was necessary because of different slopes and segments in the plots of interaction energies vs. theoretical parameters for the different types of functional groups considered in the present research. The generalized plots between hydrogen bonding energies and scaled parameters revealed the accuracy of the adopted approach.



Scheme 3. Series of model compounds containing oxygen proton accepting centers. The arrows point the atoms, for which electrostatic potential at nuclei (V_O), atomic charges (q_{NBO}, q_{Hirsh}), and proton affinities were calculated. The arrows also indicate the chosen direction of hydrogen bond formation.



Scheme 4. Model compounds containing nitrogen proton-donating centers. The arrows point the atoms, for which electrostatic potential at nuclei (V_H), atomic charges (q_{NBO}, q_{Hirsh}), and deprotonation energies were calculated. The arrows also indicate the chosen direction of hydrogen bond formation.

(4) In the fourth step, we define a suitable hydrogen bonding reactivity descriptor, based on the derived relationships between interactions energies and scaled molecular parameters.

RESULTS AND DISCUSSION

Correlations between site electrostatic parameters and proton affinities

Proton affinity (PA) and deprotonation energies (E_{dep}) are quantities that reflects well local properties of atomic sites in molecules. It was of special interest to examine how well molecular parameters associated with site electrostatics would correlate with quantities (PA, E_{dep}) reflecting purely chemical properties. Subsequent analyses of relationships between hydrogen bonding energies and molecular parameters may provide a clue for selecting the quantity describing *most accurately* the hydrogen bonding interactions. We evaluated by theoretical computations the proton affinities and electrostatics-related parameters (EPNs, NBO and Hirshfeld atomic charges), for sets of molecules containing nitrogen and oxygen proton accepting sites.

Tables S1 and S2 present these theoretically estimated quantities for the nitrogen and oxygen atomic sites in the selected model molecules and the five primary nucleobases. The last rows in Tables S1 and S2 show the correlation coefficients for the relationships between proton affinities and the three local parameters. A very good correlation between PAs and EPN (V_N) values for all considered 54 nitrogen proton accepting sites is established (correlation coefficient $r = 0.986$, $n= 54$). The plot is illustrated in Fig. S1. The correlation covers the model compounds (Scheme 2) and the respective nitrogen centers in the nucleobases. In contrast, the results reveal that the evaluated NBO and Hirshfeld charges for the nitrogen atoms do not correlate well with proton affinities (Table S1).

Table S2 presents the theoretically evaluated shifts of EPN values (ΔV_O) and atomic charges for the oxygen proton accepting centers for the series of model compounds (Scheme 3). A good correlation between PAs and EPN values (V_O) ($r = 0.985$, $n = 24$) characterizes the relationship between these two molecular quantities (Fig. S2). Again, the NBO and Hirshfeld charges for the oxygen atoms do not correlate satisfactorily with proton affinities.

Table 1. Hydrogen bonding energies and reactivity descriptors for nitrogen proton accepting centers in model molecules (Scheme 2) from M06-2X/6-311+G(2d,2p) computations.

Molecule	**ΔE_{corr} [kcal/mol l]**	**ΔV_N[a] [kcal/mol l]**	**PA [kcal/mol l]**
		sp^2	
Pyridine	-9.43	2.64	219.0
Pyrimidines			
H	-8.27	10.47	208.6
4-CH_3	-8.78	6.14	214.6
4-NH_2	-9.66	-2.18	224.0
4-F	-7.82	15.71	204.2
4-CHO	-7.33	20.17	203.4
4-CN	-6.91	27.08	197.1
4-NO_2	-6.57	30.09	193.9
Purines			
H	-8.31	10.31	210.9
2-CH_3	-8.57	7.71	214.3
2-NH_2	-8.99	4.16	219.2
2-F	-7.83	15.44	206.5
2-CHO	-7.52	17.42	206.0
2-CN	-7.14	22.69	201.1
2-NO_2	-7.16	24.78	199.1
Correlation coefficients[b]		**0.991**	**0.985**
		sp^3	
Anilines			
H	-6.83	24.20	207.3
4-CH_3	-6.90	18.96	210.0
4-OCH_3	-7.92	17.61	212.0
4-NH_2	-7.47	15.30	214.9
4-F	-6.77	23.90	204.6
4-Cl	-6.19	26.98	203.3
4-Br	-6.06	27.72	203.1
4-CHO	-4.93	34.54	199.0
4-CN	-4.84	37.74	195.0
5-Amino-pyrimidines			
H	-5.15	35.71	193.1
2-CH_3	-5.36	32.50	197.3
2-NH_2	-6.31	26.90	204.7
2-F	-4.90	39.23	189.9
2-CHO	-3.92	46.42	186.8
2-CN	-3.32	51.23	
2-NO_2	-3.16	53.60	
Correlation coefficients[b]		**0.987**	**0.958**

Table S3 present the theoretically evaluated deprotonation energies and molecular parameters for N-H proton donating sites in series of model compounds (Scheme 4) and in the nucleobases. The shifts in EPN values for the N-H hydrogen correlates with E_{dep}, though the correlation coefficients (r = 0.956) is lower compared to the plots between proton affinities and ΔV_Y values discussed so far. Still, the correlation between ΔV_H and E_{dep} is superior than the correlations of E_{dep} with q_H(NBO) and q_H(Hirsh) (Table S3).

Table 2. Hydrogen bonding energies and reactivity descriptors for oxygen proton accepting centers in model molecules (Scheme 3) from M06-2X/6-311+G(2d,2p) computations.

Molecule	**ΔE_{corr} [kcal/mol]**	**ΔV_O[b] [kcal/mol.]**	**PA [kcal/mol]**
		Cyclohexa-2,4-dienones	
H	-9.72	-22.41	209.0
4-CH_3	-10.11	-23.54	211.1
4-NH_2	-10.00	-23.10	211.8
4-F	-9.14	-15.26	202.5
4-Cl	-8.91	-14.69	203.6
4-Br	-9.21	-14.88	203.3
4-CHO	-8.48	-13.63	202.0
4-CN	-8.56	-8.36	196.8
4-NO	-8.81	-11.87	200.2
4-NO_2	-8.34	-6.91	195.6
		Benzoquinone	
	-6.92	1.31	188.6
		Pyridin-2(1H)-ones	
H	-11.38	-36.91	217.4
5-CH_3	-11.61	-39.42	220.9
5-NH_2	-12.18	-41.99	224.7
5-F	-10.80	-30.38	213.3
5-Cl	-11.15	-32.83	213.5
5-CN	-9.53	-18.46	203.9
5-NO_2	-9.22	-15.01	201.8
Correlation coefficients [a]		**0.991**	**0.979**

[a] Correlation coefficients for the relationships between ΔE_{cor} an molecular parameters. [b] ΔV_O is defined in eqn 4.

The relationships discussed above reveal good correspondence between chemical properties (proton affinities, deprotonation energies) and a theoretical parameter (EPN) describing the electrostatic forces acting at individual atomic sites. In the following discussion of hydrogen bonding interactions we will focus on the relationships employing EPN, PAs, and E_{dep} values. The computations showed that NBO and Hirshfeld atomic charges are less successful in describing small variations in site reactivities for the considered extended series of molecules. The results obtained also revealed the presence of outliers in the PA/EPN and E_{dep}/EPN relationships. Most critical for the purposes of the present investigation is the assessment of how well these molecular quantities would describe the abilities of atomic centers to form hydrogen bonds.

Zeegers et al. [17-20] have applied PAs and E_{dep} in characterizing the hydrogen bonding reactivities of the nucleobases. The derived by these authors relationships are based on analyzing the interaction energies of nucleobase functionalities with water molecule. As discussed, bidentate hydrogen bonds with neighboring polar groups in the bases form in almost all cases.

Table 3. Hydrogen bonding energies, electrostatic potential at nuclei, and deprotonation energies for N-H hydrogen in 1,6-dihydropyrimidine and aniline derivatives from M06-2X/6-311+G(2d,2p).

Derivative	**ΔE_{corr} [kcal/mol]**	**ΔV_H[b] [kcal/mol]**	**E_{dep} [kcal/mol]**
	1,6-Dihydro-pyrimidines		
H	-4.57	31.89	-363.50
4-OCH$_3$	-4.69	33.33	-361.05
4-NH$_2$	-4.72	31.32	-363.05
4-F	-5.25	41.67	-354.94
4-Cl	-5.33	41.99	-352.15
4-Br	-5.41	41.67	-350.29
4-CHO	-5.42	41.24	-353.23
4-CN	-5.77	41.57	-347.53
4-NO$_2$	-5.85	50.27	-345.29
	Anilines		
H	-3.64	17.14	-374.40
4-OCH$_3$	-3.54	13.31	-377.40
4-NH$_2$	-3.38	11.05	-377.89
4-F	-3.94	19.77	-372.49
4-Cl	-4.05	22.91	-367.67
4-Br	-4.04	27.37	-366.35
4-CHO	-4.49	30.57	-355.47
4-CN	-4.84	33.77	-354.91
4-NO$_2$	-5.18	36.91	-348.79
Correlation coefficient [a]		**0.979**	**0.964**

[a]Correlation coefficients for the relationships between ΔE_{corr} and ΔV_H. [b] ΔV_H is defined in eqn (5).

The specific structure of H_2O, however, may be a factor influencing the nature of such cooperative bonding. In the present work, we aim at establishing relations that are free of such effects

Relationships between hydrogen bonding energies and molecular parameters

Based on the above results it is possible to explore dependences between hydrogen bonding energies and molecular parameters for the sets of model molecules. As discussed, the selected model molecules (Schemes 2-4) possess atomic centers that are able to form single hydrogen bonds, without the complications of cooperative bonding. Our study focuses at analyzing how well proton affinities, deprotonation energies, and electrostatic potentials at atomic sites correlate (and predict) the hydrogen bonding energies for the sets of model molecules.

Interaction energies are, certainly, the most accurate descriptor of the ability of individual atomic centers to form hydrogen bonds. For the model derivatives, hydrogen bonding energies, associated with individual atomic sites, were theoretically evaluated. As mentioned, hydrogen fluoride was employed as a model proton donor or acceptor molecule. The established good correlations for the PA/EPN and E_{dep}/EPN relationships show that a molecular parameter, representing the electrostatic forces acting at particular atomic sites, may well predict site chemical properties for *both model systems and the nucleobases* (Table S1-S3). We explored the relationships between interaction energies and EPN values and proton affinities for the molecules shown in Schemes 2 and 3. Deprotonation energies were evaluated for proton-donating N-H bonds (Schemes 4). These model systems possess hydrogen bonding functional groups with sp^2 and sp^3 nitrogens, sp^2 oxygen, and N-H bonds in structural environments similar to those in the nucleobases. As shown, NBO and Hirshfeld charges were found to be less successful in describing site reactivities in the selected model systems. Nonetheless, in the Supporting Information we provide comparative tables (Tables S4-S6) that contain data for these two types of atomic charges.

Table 1 shows the computed hydrogen bonding energies (ΔE_{corr}) for the nitrogen centers (proton acceptors) for the series of model pyrimidine, purine, and aniline derivatives (Scheme 2). These values are juxtaposed to theoretically derived molecular parameters: ΔV_N and proton affinities. The results obtained reveal that no *common* relationship between ΔE_{corr} and any of these molecular parameters is found when molecules containing both sp^2 and sp^3 nitrogens are simultaneously considered. The data obtained demonstrate that the pyridine type nitrogens form distinctly stronger hydrogen bonds than the aniline type nitrogen atoms (Table 1). It was thus necessary to investigate separately these two sets of complexes. An excellent correlation between interaction energies and ΔV_N (sp^2 N) ($r = 0.991$, $n = 15$) for the entire set of substituted pyrimidines and purines was found (Table 1). A good correlation, though with *lower correlation coefficient* ($r = 0.985$, $n = 15$), was established between *proton affinities* and ΔE_{corr} (Table 1).

The correlation between ΔE_{corr} and V_N values for the series of aniline derivatives is also quite satisfactory with correlation coefficient $r = 0.987$. Again, this correlation is superior to the correlation between ΔE_{cor} and proton affinities for molecules containing sp^3 nitrogen centers ($r = 0.958$).

The equations reflecting the above relationships are as follows:

$\Delta E_{cor} = 0.097\Delta V_N\ (sp^2\ N)\ -9.390$ (6)
n=15, r = 0.991
$\Delta E_{corr} = 0.121\Delta V_N(sp^3\ N)\ -9.511$ (7)
n= 16, r = 0.987

The differences in slopes and intercepts in Eqns. 6 and 7 reflect the lower hydrogen bonding ability for the sp^3 nitrogen atoms in amino group environment. An appropriate scaling of $V_N(sp^3\ N)$ to reflect these difference results in the relation:

$\Delta V_N(sp^3\ N)_{scaled} = 1.24\ \Delta V_N(sp^3\ N)$ (8)

The segments in eqns. (6) and (7) are very close in value and, thus, the simplified relation 8 holds to a good approximation. Using the $\Delta V_N(sp^3\ N)_{scaled}$ the following relation covering all nitrogen proton accepting sites is derived:

$\Delta E_{corr} = 0.096\Delta V_N^{scaled} - 9.40$ (9)
r = 0.994, n = 31

ΔV_N^{scaled} comprises ΔV_N for the sp^2 nitrogens (scaling factor = 1) and $\Delta V_N(sp^3\ N)_{scaled}$ obtained from eqn (8). Using eqn (9) it becomes possible to prepare a reactivity scale for all nitrogen proton accepting centers in the model molecules.

To derive analogous relationships between structural variation in carbonyl compounds and molecular parameters (EPN, PAs), associated with oxygen proton accepting sites, a series of 18 carbonyl containing derivatives of cyclohexa-2,4-dienone and pyridin-2(1H)-one (Scheme 3) were considered.

To avoid cooperative bidentate hydrogen bonding of the model proton donor in the case of the aza heterocyclic molecules the hydrogen bond at the carbonyl oxygen was modeled at the side of the C=C double bond adjacent to the carbonyl moiety (Scheme 3). The computed corrected energies of hydrogen bonding (ΔE_{corr}), electrostatic potentials at nuclei (V_O), and proton affinities (PA) for the carbonyl oxygen sites are given in Table 2. The excellent correlation coefficient (r = 0.991) between the computed EPN values and energies of hydrogen bonding (ZPE and BSSE corrected) shows the near linear relationship between these quantities. A good correlation coefficient (r=0.979) characterizes the dependence between ΔE_{corr} and proton affinities. As in the case of nitrogen proton-accepting centers (Table 1), the electrostatic potential at nuclei values provide superior correlations with interaction energies compared to proton affinities for the series of carbonyl derivatives. This result reflects the strong dependence of hydrogen bonding on the electrostatic interaction between proton donor and acceptor molecules. The relationship between ΔV_O values and ΔE_{corr} results in the following expression:

$\Delta E_{corr}(C{=}O) = 0.112\Delta V_O - 7.382$ (10)
n = 18, r = 0.991

Equation 10 holds for two different series of cyclic compounds containing C=O functionality. This result shows that in spite of different molecular environment (ketone C=O group in the first set and amide C=O group in the second set) the same relationship between interaction energies and EPN values holds. Successful applications of EPN in quantifying hydrogen bonding reactivities of carbonyl derivatives belonging to different classes [25-27] are in harmony with the present results.

Considerations of the slopes and segments of eqns. (9) and (10) results in the following expression for the scaling of ΔV_O:

$\Delta V_O(sp^2\ O)_{scaled} = 1.167\Delta V_O(sp^2\ O) - 18.01$ (11)

Considering eqns (9-11) we arrive at a general relation quantifying the hydrogen bonding reactivities of all proton-accepting sites. Scaling of ΔV_O eqn. (11) is necessary to obtain such an equation. The derived general relation is:

$\Delta E_{corr} = -0.106\ \Delta V_Y^{scaled} - 9.22$ (Y=N,O) (12)
n = 49, r = 0.995

Eqn (12) provides a general scale for the hydrogen bonding abilities of all proton accepting site. It is based on the scaled values of the respective ΔV_Y values. ΔV_y^{scaled} are presented in Table 6. The plot between hydrogen bonding energies and ΔV_y^{scaled} is shown in Fig. 2.

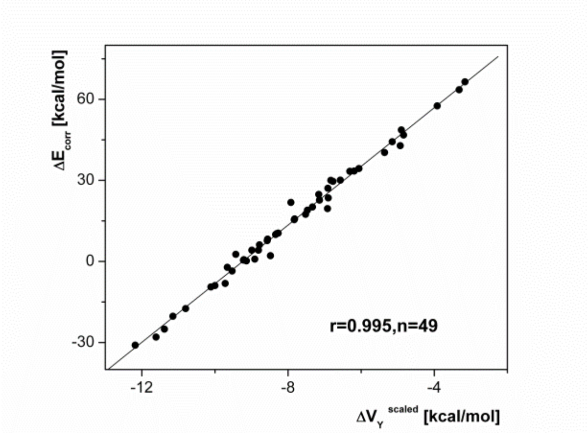


Fig. 2. Plot of hydrogen bonding energies vs. scaled electrostatic potentials at nuclei (V_Y^{scaled}) in in model pyrimidines, purines and carbonyl derivatives

Following the same approach we derived relationships between interaction energies and molecular parameters for the N-H proton-donating groups in the series of model 1,6-dihydropirimidines and anilines shown in Scheme 4. Table 3 presents the computed energies of hydrogen bonding and molecular parameters. A good correlation (r = 0.979) is established between ΔV_H and interactions energies (Fig. 3), while the respective relationship with E_{dep} is characterized with lower correlation coefficient (r = 0. 964). These results confirm that the shifts of EPN provide *superior correlations* with the hydrogen bonding energies compared to proton affinities/ deprotonation energies. . No scaling of EPN values is necessary in this case since the correlation involves only N-H group hydrogens. Eqn (13) shows that more negative ΔV_H values correspond to higher interaction energies.

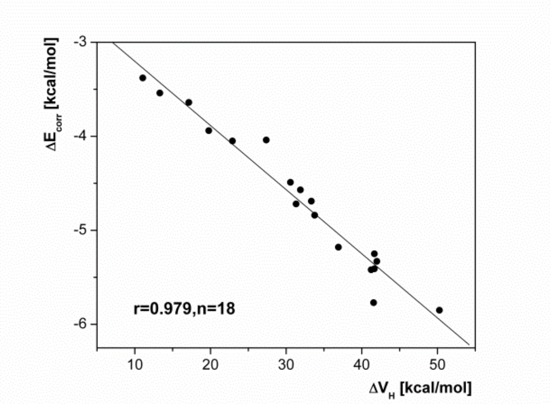


Fig. 3. Energies of hydrogen bonding (ΔE_{cor}) vs. shifts of electrostatic potentials at of proton-donating groups (ΔV_H) in substituted aniline and 1,6-dihydropyrimidine derivatives.

Thus, ΔV_H can directly be employed in deriving suitable hydrogen bonding reactivity descriptor for N-H proton-donating groups. The following equation applies:

$$\Delta E_{corr} = -0.068\Delta V_H - 2.52 \qquad (13)$$

n = 18, r = 0.979

Hydrogen bonding reactivity descriptors for nucleobases

ΔV_y may be employed in characterizing the hydrogen bonding reactivities in the nucleobases and their polymeric derivatives. It is only necessary to compute the respective scaled electrostatic potential values for the proton accepting centers (ΔV_y^{scaled}). Even easer assessment of bonding abilities comes by comparing the predicted hydrogen bonding energies using Eqn. 12.

It is appropriate to introduce a special hydrogen bond descriptor (HB_{descr}) for reactivities of atomic sites in nucleobases. We define this descriptor as the absolute values of the predicted energies of hydrogen bonding for particular sites in the bases using Eqns (12) and (13):

$$HB_{descr} = |\Delta E^{predicted}| \qquad (14)$$

The HB_{descr} values are given in the last columns of Tables 4 and 5. Table 4 shows the order of reactivities of all proton-accepting sites in the nucleobases. Inspection shows that the most reactive center for hydrogen bonding is the N3 nitrogen in uracil (HB_{descr}= 18.17).

Table 4. Scaled EPN values, predicted energies of hydrogen bonding (ΔEpredicted, in kcal/mol), and hydrogen bond descriptor (HBdescr) values for proton-accepting centers in the primary nucleobases.

Nucleo-bases	ΔV_Y	ΔV_Y scaled, a	$\Delta E^{pr,\ b}$	HB_{descr}
		Adenine		
N1(sp^2)	-1.84	-1.84	-9.02	9.02
N3(sp^2)	1.86	1.86	-9.42	9.42
N7(sp^2)	6.88	6.88	**-9.95**	9.95
N9(sp^3)	61.81	76.64	-17.34	17.34
N10(sp^3)	33.9	42.0	-13.68	13.68
		Guanine		
N1(sp^3)	56.85	70.49	-16.69	16.69
N3(sp^2)	9.76	9.76	-10.25	10.25
N7(sp^2)	0.81	0.81	-9.31	9.31
N9(sp^3)	60.94	75.56	-17.23	17.23
N11(sp^3)	45.07	55.89	-15.14	15.14
O10	-29.58	-34.52	-5.56	5.56
		Thymine		
N1(sp3)	54.49	67.57	-16.58	16.58
N3(sp3)	64.47	79.95	-17.69	17.69
O7	-20.16	-23.53	-6.73	6.73
O8	-19.94	-23.27	-6.75	6.75
		Uracil		
N1(sp3)	55.98	69.41	-16.58	16.58
N3(sp3)	68.12	84.47	-18.17	18.17
O7	-19.75	-23.04	-6.78	6.78
O8	-16.69	-19.48	-7.16	7.16
		Cytosine		
N1(sp3)	53.87	66.81	-16.30	16.30
N3(sp2)	-6.74	-6.74	-8.51	8.51
N7(sp3)	38.77	48.07	-14.32	14.32
O8	-41.87	-48.86	-4.04	4.04

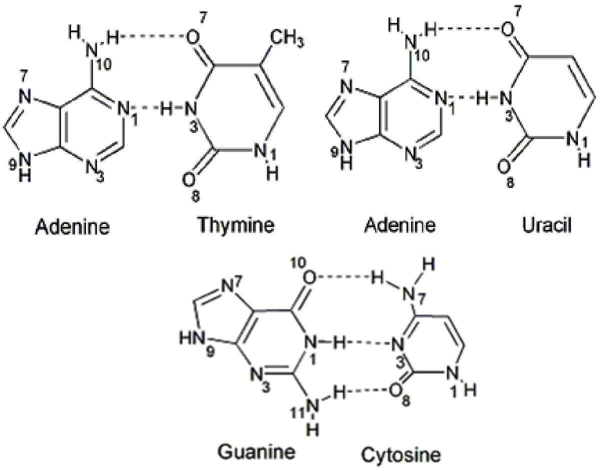
[a] Obtained from eqn (8) and eqn (11) [b] Predicted values using eqn (12).

Among the oxygen atomic sites, the O8 atom in uracil (HB_{descr}= 7.16) is predicted to possess highest reactivity. Table 5 compares the relative reactivities of proton-donating atomic sites in the nucleobases. The values of reactivity descriptors are obtained using relation 14. The effects of base pairing on hydrogen bond reactivities are illustrated in Table 6. The last column contains HB_{descr} values for the atomic sites in base pairs. The numbering of atomic sites is shown in Scheme 5. Comparison of Tables 4 and 6 reveals the effects of pairing on hydrogen bonding reactivities.

Table 5. Shifts of EPN values, predicted energies of hydrogen bonding ($\Delta E^{predicted}$, in kcal/mol), and hydrogen bond descriptor (HB_{descr}) values for proton-donating centers in the primary nucleobases.

Nucleo-bases	**ΔV_H[a]**	**$\Delta E^{pr.,}$[b]**	**HB_{descr}**
Adenine			
N9-H	57.56	-6.44	6.44
N10-H	31.44	-4.66	4.66
Guanine			
N1-H	50.95	-5.99	5.99
N9-H	56.85	-6.39	6.39
N11-H	40.91	-5.30	5.30
Tymine			
N1-H	56.43	-6.36	6.36
N3-H	47.82	-5.77	5.77
Uracil			
N1-H	59.86	-6.59	6.59
N3-H	49.32	-5.88	5.88
Cytosine			
N1-H	44.93	-5.58	5.58
N7-H	39.22	-5.19	5.19

[a] ΔV_H is defined in eqn (5). [b] Predicted values for interaction energies using eqn (13).



Scheme 5. Structure and numbering of atom position in base pairs Adenine-Thymine, Adenine-Uracil and Guanine-Cy tosine.

In the A-U base pair, the HB_{descr} for the N1 atom in adenine shifts from 9.02 to 9.59. For the N10 atom, the shift is from 13.68 to 12.90. Similar HB_{descr} shifts are also found for the A-T base pair. The HB_{descr} for the N1 atom in adenine changes from 9.02 to 9.54, while for the N10 from 13.68 to 12.86. The respective HB_{descr} values for guanine in the G-C pair change also upon base-paring. . The HB_{descr} value for the N1 atom in guanine shifts from 16.69 to 15.08, for the N11 atom from 15.14 to 12.91, and for O7 from 5.56 to 5.74. Distinct shifts are also found for the other base pairs (Tables 3 and 6).

Table 6. Scaled EPN values, predicted values of hydro gen bonding energies ($\Delta E^{predicted}$, in kcal/mol), and hydrogen bond descriptor (HB_{descr}) for proton-accepting centers in Adenine-Thymine, Adenine-Uracil, and Guanine-Citosine base pairs.

Base pair	**ΔV_Y**	**ΔV_Y^{scaled}**	**ΔE^{pr}**	**$HB_{descr.}$**
Adenine (A-U)				
N1(sp^2)	3.53	3.53	-9.59	9.59
N10(sp^3)	28.02	34.75	-12.90	12.90
Adenine (A-T)				
N1(sp2)	2.98	2.98	-9.54	9.54
N10(sp3)	27.73	34.38	-12.86	12.86
Guanine (G-C)				
N1(sp^3)	44.56	55.26	-15.08	15.08
N11(sp^3)	28.11	34.85	-12.91	12.91
O10	-28.16	-32.86	-5.74	5.74
Thymine (A-T)				
N3(sp3)	39.19	48.60	-14.37	14.37
O7	-21.99	-25.66	-6.50	6.5
Uracil (A-U)				
N3(sp3)	40.31	49.99	-14.52	14.52
O7	-21.88	-25.53	-6.51	6.51
Cytosine (G-C)				
N3(sp^2)	36.15	36.15	-13.05	13.05
N7(sp3)	45.46	56.37	-15.20	15.20
O8	-13.14	-15.33	-7.59	7.59

[a] Obtained from eqns (8) and (11). [b] Predicted values using eqn (12).

The shifts in HB_{descr} resulting from base pairing may be distinct, though not large. This is in harmony with the relatively small perturbations caused by hydrogen bonding between the nucleobases**.**

CONCLUSION

Hydrogen bonding reactivities of individual atomic sites in the primary nucleobases were characterized by applying theoretically derived electrostatic-related theoretical parameters. Several alternative quantities – electrostatic potentials at nuclei (EPN), NBO and Hirshfeld atomic charges as well as proton affinities and deprotonation energies – were employed in the analysis. For all studied systems, the EPN values provided the best correlations with interaction energies. A new reactivity descriptors (HB_{descr}) for the hydrogen bonding reactivities of atomic sites in the nucleobases is introduced based on predicted interaction energies. Using the derived equations, site reactivities for the primary nucleobases were evaluated. The effect of base pairing was also analyzed.

Acknowledgements: *This research was supported by the National Science Fund (Bulgaria), Grant DN 09/4.*

Electronic Supplementary Data available here.

REFERENCES

1. G. C. Pimental, A. L. McClellan, The Hydrogen Bond; Freeman: San Francisco, 1960.
2. G. A. Jeffrey, W. Saenger, Hydrogen Bonding in Biological Structures; Springer: Berlin, 1991.
3. J. T. Stivers,, Y. L. Jiang, *Chem. Rev.*, **103**, 2729 (2003).
4. S. S. David, S. D. Williams, *Chem. Rev*., **98**, 1221 (1998).
5. A. K. McCullough, M. L. Dodson, R. S. Lloyd, *Annu. Rev. Bioche.*, **68**, 255 (1999).
6. K. Muller-Dethlefs, P. Hobza, *Chem. Rev*., **100**, 143 (2000).
7. J. D. Watson, F. H. C. Crick, *Nature*, **171**, 737 (1953).
8. E. Martin, W. Saenger, Principles of Nucleic Acid Structure; Springer: New York, 1984.
9. D. Voet, J. G. Voet, *Biochemistry*, 3rd ed.; Hoboken, NJ: Wiley, New York, 2004.
10. L. Chan, M. Pineda, J. Heeres, P. Hergenrother, B. Cunningham, *ACS Chem. Biol*., **3**, 437 (2008).
11. M. Di Laudo, S. R. Whittleton, S. D. Wetmore, *J. Phys. Chem. A*, **107**, 10406 (2003).
12. S. R. Whittleton, K. C. Hunter, S. D. Wetmore, *J. Phys. Chem. A*, **108**, 7709 (2004).
13. K. C. Hunter, L. R. Rutledge, S. D. Wetmore, *J. Phys. Chem. A*, **109**, 9554 (2005).
14. T. L. McConnell, C. A. Wheaton, K. C. Hunter, S. D. Wetmore, *J. Phys. Chem. A*, **109**, 6351 (2005).
15. T. van Mourik, S. L. Price, D. C. Clary, *J. Phys. Chem. A*, **103**, 1611 (1999).
16. S. M. LaPointe, S. D. Wetmore, *Chem. Phys. Lett*., **408**, 322 (2005).
17. Л. Chandra, M. Nguyen, T. Uchimaru, Th. Zeegers- Huyskens, *J. Phys. Chem. A*, **103**, 8853 (1999).
18. ЛA. Chandra, M. Nguyen, T. Uchimaru, Th. Zeegers- Huyskens, *J. Phys. Chem. A*, **102**, 6010 (1998).
19. ЛA. Chandra, D. Michalska, R. Wysokinski, Th. Zeegers- Huyskens, *J. Chem. Phys A*, **108**, 9593 (2004).
20. R. Wysokinski, D. Bienko, D. Michalska, Th. Zeegers- Huyskens, *Chem. Phys.*, **315**, 17 (2005).
21. D. Talukdar, R. Parajuli, R. Kalita, C. Medhi, *Ind. J. Chem*., **45**, 1804 (2006).
22. P. Politzer, D. G. Truhlar, Chemical Applications of Atomic, Molecular Electrostatic Potentials; Eds.; Plenum: New York, 1981.
23. J. S. Murray, K. Sen, Molecular Electrostatic Potentials: Concepts, Applications; Eds.; Elsevier Science: Amsterdam, The Netherlands, 1996.
24. G. Naray-Szabo, G. G. Ferenczy, *Chem. Rev*., **95**, 829 (1995).
25. F. B. Sheinerman, B. Honig, *J. Mol. Biol*., **318**, 161 (2002).
26. H. Zheng, K. Comeforo, J. Gao, *J. Am. Chem. Soc*., **131**, 18 (2009).
27. G. A. Jeffrey, Introduction to Hydrogen Bonding; Oxford University Press: New York, 1997.
28. S. J. Grabowski, Hydrogen Bonding: New Insights; Springer: New York, 2006.
29. (a) A. Reed, R. Weinstock, F. Weinhold, *J. Chem. Phys.*, **83**, 735 (1985); (b) A. E. Reed, L. A. Curtiss, F. Weinhold, *Chem. Rev*., **88**, 899 (1988).
30. (a) F. L. Hirshfeld, *Theor. Chem. Acc*., **44**, 129 (1977); (b) J. P. Ritchie, *J. Am. Chem. Soc*., **107**, 1829 (1985); (c) J. P. Ritchie, S. M. Bachrach, *J. Comp. Chem*., **8**, 499 (1987).
31. P. Bobadova-Parvanova, B. Galabov, *J. Phys. Chem. A*, **102**, 1815 (1998).
32. B. Galabov, P. Bobadova-Parvanova, *J. Phys. Chem. A*, **103**, 6793 (1999).
33. B. Galabov, P. Bobadova-Parvanova, *J. Mol. Struct*., **550-551**, 93 (2000).
34. V. Dimitrova, S. Ilieva, B. Galabov, *J. Phys. Chem. A*, **106**, 11801 (2002).
35. B. Galabov, P. Bobadova-Parvanova, S. Ilieva, V. Dimitrova, *J. Mol. Struct. Theochem*, **630**, 101. (2003).
36. B. Galabov, S. Ilieva, G. Koleva, W. D. Allen, H. F. Schaefer, P. v. R. Schleyer, *WIREs Comput. Mol. Sci*., **3**, 37 (2013).
37. B. Galabov, V. Nikolova, S. Ilieva, *Chem.-Eur. J.*, **19**, 5149 (2013).
38. K. Morokuma, *J. Chem. Phys*., **55**, 1236 (1971).
39. K. Kitaura, K. Morokuma, *Int. J. Quantum Chem*., **10**, 325 (1976).
40. Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc*., **120**, 215 (2008).
41. a) A. D. McLean, G. S. Chandler, *J. Chem. Phys*., **72**, 5639 (1980); b) A. Krishnan, J. S. Binkley, R. Seeger, J. A. Pople, *J. Chem. Phys.*, **72**, 650 (1980); c) T. Clark, J. Chandrasekhar, G. W. Spitznagel, P. v. R Schleyer, *J. Comp. Chem.*, **4**, 294 (1983).
42. A. R. Neves, P. A. Fernandes, M. Ramos, *J. Chem. Theor. Comp.*, **7**, 2059 (2011).
43. S. F. Boys, F. Bernardi, *Mol. Phys*., **19**, 553 (1970).
44. M. J. Frisch, et al. *Gaussian 09 (Revision-A.02);* Gaussian, Inc., Wallingford CT, 2009.
45. E. B. Wilson, *J. Chem. Phys.*, 2232 (1962).
46. B. Galabov, S. Ilieva, H. F. Schaefer, *J. Org. Chem*., **71**, 6382 (2006).
47. B. Galabov, S. Ilieva, B. Hadjieva, Y. Atanasov, H. F. Schaefer, *J. Phys. Chem. A*, **112**, 6700 (2008).
48. B. Galabov, V. Nikolova, J. J. Wilke, H. F. Schaefer, W. D. Allen, *J. Am. Chem. Soc*., **130**, 9887 (2008).

49. N. Mohan, C. H. Suresh, *J. Phys. Chem. A*, **118**, 1697 (2014).
50. K. S. Sandya, C. H. Suresh, *Dalt. Trans.*, **43**, 12279 (2014).
51. F. B. Sayyed, C. H. Suresh, *J. Phys. Chem. A*, **116**, 5723 (2011).

РЕАКТИВОСПОСОБНОСТ КЪМ ВОДОРОДНО СВЪРЗВАНЕ НА АТОМНИТЕ ЦЕНТРОВЕ В АЗОТНИТЕ БАЗИ НА НУКЛЕИНОВИТЕ КИСЕЛИНИ

В. Николова, Д. Чешмеджиева, С. Илиева, Б. Гълъбов*

Факултет по химия и фармация, Софийски университет „Св. Климент Охридски”, бул. Дж. Баучер 1, 1164 София, България

Постъпила на 27 март 2017 г.; Коригирана на 10 април 2017 г

(Резюме)

Приложени са теоретични параметри, характеризиращи електростатичните сили при отделните атоми, за количествено описание на реакционна способност към водородно свързване на атомните центрове в петте основни бази на нуклеиновите киселини. С цел да се избегнат неяснотите, свързани с формирането на бидентатни вододродни връзки в нуклеиновите бази, бяха проведени DFT M06-2X/6-311+G(2d,2p) изчисления на атомния електростатичен потенциал (electrostatic potential at nuclei, EPN), NBO и Хиршфелд атомни заряди за серия от 64 моделни системи. За сравнение бяха определени протонните афинитети на протон-акцепторните групи и енергиите на депротониране на протон-донорните групи в базите. Моделните системи са така подбрани, че да могат да формират единични водородни връзки. Получените корелационни уравнения, свързващи изчислените енергии на взаимодействие с параметри, характеризиращи електростатичните сили при атомите, са приложени за извеждането на дескриптори на реактивоспособността за отделните атомни позиции в нуклеиновите бази. Сред изследваните теоретични параметри EPN дава най-добра корелация с енергиите на водородно свързване. Изведените зависимости бяха използвани за дефинирането на дескриптор на риакционната способност към водородно свързване за отделните атомни позиции в нуклеиновите бази. Ефектите на формиране на двойки бази върху реактивоспособността бяха също оценени количествено.