

Cite this: *RSC Adv.*, 2019, **9**, 31343

Received 19th June 2019  
 Accepted 20th September 2019  
 DOI: 10.1039/c9ra04615a  
[rsc.li/rsc-advances](http://rsc.li/rsc-advances)

## Substituent effects on the stability of the four most stable tautomers of adenine and purine<sup>†</sup>

Halina Szatylowicz,<sup>id \*a</sup> Anna Jezuita,<sup>b</sup> Paulina H. Marek<sup>ac</sup>  
 and Tadeusz M. Krygowski<sup>c</sup>

Substituent effects at the C2-, C8- and N-positions of adenine and purine in their four the most stable tautomers are studied by means of B97D3/aug-cc-pvdz computation applying substituents of varying electronic properties: NO<sub>2</sub>, CN, CHO, Cl, F, H, Me, OMe, OH and NH<sub>2</sub>. The substituent effect is characterized by the substituent effect stabilization energy (SESE) and substituent Hammett constant  $\sigma$ . For adenine, SESE is obtained with purine as the reference system. Additionally, for both adenine and purine, SESE characteristics are estimated with benzene, imidazole and amino-pyrimidine as reference systems, when possible, taking into account substitution in topologically equivalent positions. The role of a C6-NH<sub>2</sub> group in adenine in modifying the substitution effect is observed and discussed. Additionally, the proximity effect for some asymmetric substituents (e.g. CHO, OMe) is recognized and meticulously analyzed.

### Introduction

Adenine is one of five building blocks of DNA and RNA nucleic acids<sup>1</sup> and is a simple amino derivative of purine. The low energetic prototropic abilities of adenine result from the presence of four endocyclic nitrogen atoms, one of which is bonded to a hydrogen atom. The proton may change its position, which leads to a formation of four low-energy tautomers. Among them, the 9H tautomer (Scheme 1) is the most stable, which has been confirmed both theoretically and experimentally; for details see Raczyńska *et al.*<sup>2</sup> and references therein. In contrast, there are also possibilities of tautomeric forms in which the exocyclic amino-group or carbon atoms are involved, but these tautomeric rearrangements are highly energetically unfavorable.<sup>3</sup> Proton transfers in tautomeric rearrangements are always accompanied by the migration of double bonds, that is, changes in  $\pi$ -electron structure.<sup>4</sup> As a result, it can be expected that the different sensitivity of a given position to the substituent effect depends on the type of tautomer considered. This is the reason why the impact of the substituent effect on various properties of purine or adenine systems has been studied. More than half a century ago, the effect of substituent on the hydrogen bonding of adenine derivatives was studied using IR technique.<sup>5</sup> This allowed to determine the association constants between variously substituted 1-cyclohexyluracil derivatives and 9-

ethyladenine and to compare them with the corresponding data for thymine derivatives. Almost at the same time, substituent effects on the optical activity of some purine nucleosides were studied.<sup>6</sup> The effect of exocyclic substituents of purine bases on the DNA curvature was also reported.<sup>7,8</sup> The substituent effects at the major and minor grooves of the double helix on the hydration and stability of Z- versus B-DNA were also investigated.<sup>9</sup> Subsequent studies concerned the influence of the substituent on process of pairing and polymerase recognition of simple unnatural base pairs.<sup>10</sup>

ADF computational package was used to study substituent effects on nanoswitches based on DNA base pairs<sup>11</sup> and on the NMR shielding constants of complexes based on AT and AU base pairs.<sup>12</sup> For AT and GC base pairs, systematic studies of the substituent effects on the individual hydrogen bond energies were carried out.<sup>13</sup> Recently, the effect of the substitution of halogen atoms at the position 8 in adenine 9H, 7H, 3H and 1H tautomers on their relative stability, vibrational frequencies and Raman spectra, were studied by means of DFT at the B3LYP/6-311+G(d,p) level of theory.<sup>14</sup> It was found that the stability of tautomers of substituted adenine molecules strongly depends on solvation. It was also found<sup>15</sup> that, for 20 different substituents in the position 8 of adenine, electrostatic potentials at the bonding atoms of monomers correlate nicely with interaction energies of adenine-thymine pairs.

All of the above-mentioned papers present research results indicating that many important properties of nucleic acid bases significantly depend on the substituent attached to their molecules. In addition, it should be mentioned that in most cases observed changes of the studied systems were not related to the electron properties of substituents. It is well known that

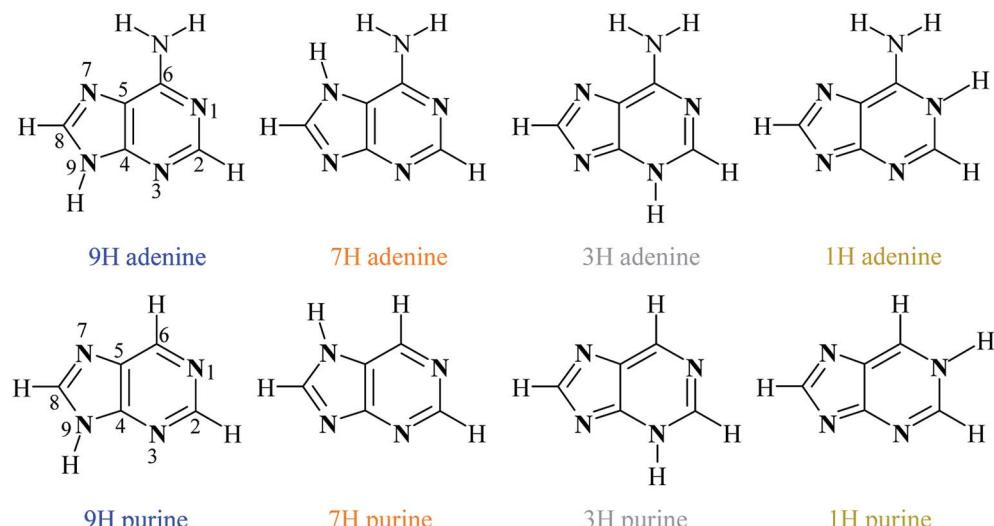
<sup>a</sup>Faculty of Chemistry, Warsaw University of Technology, Noakowskiego 3, 00-664 Warsaw, Poland. E-mail: halina@ch.pw.edu.pl

<sup>b</sup>Faculty of Chemistry, Opole University, Oleska 48, 45-052 Opole, Poland

<sup>c</sup>Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9ra04615a





Scheme 1 Structures and numbering scheme of four the most energetically favorable adenine and purine tautomers.

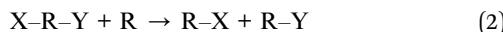
substituents can exert either electron withdrawing or electron donating properties. Thus, they may influence substituted systems in a different way and have a fundamental impact on the nature and consequences of the substituent effect.<sup>16</sup>

Classically, the quantitative characteristics of the substituent effect is described by the Hammett substituent constants  $\sigma$  (ref. 17 and 18) or their modifications.<sup>19</sup> Positive values of  $\sigma$  denote that the substituent is electron-attractive, while negative ones stand for electron-donating properties. Substituent constants have great utility in the so-called Hammett equation, where physicochemical properties of the substituted systems  $P(X)$  are linearly related to the substituent constants  $\sigma(X)$ , as shown in eqn (1)

$$P(X) = \rho\sigma(X) + \text{const} \quad (1)$$

where  $\rho$  is termed reaction constant and characterizes sensitivity of the property  $P(X)$  on influence of the substituent X. For exhaustive reviews see ref. 16 and 20–23.

Dynamic development of applications of quantum chemistry in chemical research<sup>24,25</sup> has resulted in the introduction of several models based on quantum chemistry describing substituents effects. One of them is the use of homodesmotic reactions<sup>26–28</sup> which for disubstituted species is presented by eqn (2)



The difference between energy of products and substrates, named SESE (substituent effect stabilization energy),<sup>29</sup> describes the substituent effect in energetic scale. A positive SESE value means that the intramolecular interactions between the substituents X and Y stabilize the X-R-Y system. This descriptor of the substituent effect takes into account all interactions in X-R-Y systems. Another quantum chemistry concept of the substituent effect is based on atomic charges. It is known that due to polarity of X-C bonds in substituted carbocyclic systems, atomic charges of substituents (in e.g.

benzene derivatives) do not correlate with substituent constants.<sup>30</sup> However, when instead of the charge of the X substituent,  $q(X)$ , the sum of charges in the substituent X and the *ipso* carbon atom are taken into account, the correlations with the constant substituents are very good.<sup>31,32</sup> This descriptor is named cSAR(X) (charge of the substituent active region) and independently of the atomic charge assessment method, it correlates well with substituent constants.<sup>33</sup> Recently, it has been shown that in the case of series of disubstituted benzene, cyclohexadiene and bicyclooctane derivatives, both descriptors, SESE and cSAR(X), modeled theoretically correlate well with traditional substituent constants.<sup>29,34–36</sup> It is important to note that cSAR(X) not only describes electron attracting/donating properties of the substituent, but also informs how these properties depend on its position in the substituted molecule. Another fruitful theoretical tool to study the substituent effect is the analysis of molecular electrostatic potential (MESP), which was recently published by Remya and Suresh<sup>37</sup> and references therein.

Adenine and purine are  $\pi$ -electron systems. As mentioned above, four endocyclic nitrogen atoms with the possibility of prototropic rearrangements give four tautomers. There are several questions that we should ask here: (i) how substituent effect descriptors based on quantum chemistry work in such complex molecules and whether they differ from benzene systems; (ii) how the substituent effect works in four different tautomers; (iii) whether the substituent can change the stability of the tautomer; (iv) how far the amino group in adenine may differentiate the substituent effect comparing to purine.

## Methodology

The research objects were C2-X, C8-X and N-X substituted derivatives (where X = NO<sub>2</sub>, CN, CHO, Cl, F, H, Me, OMe, OH and NH<sub>2</sub>) of the four most stable adenine and purine tautomers (Scheme 1). In the case of asymmetrical substituents (CHO, OMe and OH), both rotamers were taken into account, the one



with the lower energy was used for SESE calculations. For all studied systems, optimization without any symmetry constraints was performed using the Gaussian 09 package.<sup>38</sup> According to the results of our previous research,<sup>39</sup> the DFT-D methods were used, that is functionals B97D3 (ref. 40) and  $\omega$ B97XD,<sup>41</sup> both with Dunning's aug-cc-pvdz basis set.<sup>42</sup> To confirm that calculated structures correspond to the minima on the potential energy surface the vibrational frequencies were calculated at the same level of theory. No imaginary frequencies were found for the obtained systems. Since the results of calculations obtained on both calculation levels are very similar, in the further part of the work the first ones (B97D3/aug-cc-pvdz) are presented and discussed.

## Results and discussion

Adenine and purine are closely related to each other structurally, differing only in the amino group in the position 6 of purine. Their four most stable tautomers are realized by migration of the proton between the nitrogen atoms in positions 9, 7, 3 and 1 and are named 9H, 7H, 3H and 1H, respectively (Scheme 1).

The energy changes due to the presence of the most electron-accepting and electron-donating substituents ( $\text{NO}_2$  and  $\text{NH}_2$ , respectively) in purine and adenine are presented in Fig. 1; the data are shown in the order: 9H, 7H, 3H and 1H.

For all presented derivatives, it is clearly visible that the substitution of the C-X type causes a decrease in the energy of the system as compared to the substitution at the nitrogen atom

(N-X). Moreover, significant differences can be noticed when looking at the nitro and amino derivatives substituted at the C2 and C8 positions. In the case of the C2- and C8-nitro purine derivatives tautomer energies increase monotonically, while for adenine derivatives the lines are irregular and intersect between 7H and 3H tautomers. For amino derivatives of purine and adenine energy changes are similar. In both cases substitution at C2 leads to lower stability of 3H and 1H tautomers than in the case of C8 substitution. This means that the amino group affects energy of substituted tautomer derivatives strongly enough to modify the sequence of decreased stability. For the N-X systems shown in Fig. 1, the most significant changes in tautomer stability are observed for nitro-substituted purine derivatives, discussed below.

Relative energy values,  $E_{\text{rel}}$ , (with respect to the 9H tautomer 9H), for studied C8-X, C2-X and N-X substituted derivatives of adenine and purine tautomers can be found in Tables S1–S3.<sup>†</sup> Firstly, obtained results show that in all cases, the 9H tautomer is the most stable for both adenine and purine derivatives. The only exception is 7- $\text{NO}_2$ -purine, more stable than 9- $\text{NO}_2$ -purine by *ca.* 1.7 kcal mol<sup>-1</sup>, and the reason is a formation of a weak C–H $\cdots$ O intramolecular hydrogen bond in the former. The mean values of average relative stability in substituted adenines is  $\sim$ 10 kcal mol<sup>-1</sup>, whereas for purine series is only  $\sim$ 7.3 kcal mol<sup>-1</sup>. Undoubtedly, this difference results from the presence of the  $\text{NH}_2$  group at the 6-position in adenine, which additionally leads to greater variability of relative energies in adenine series. Secondly, we noted that for both adenine and

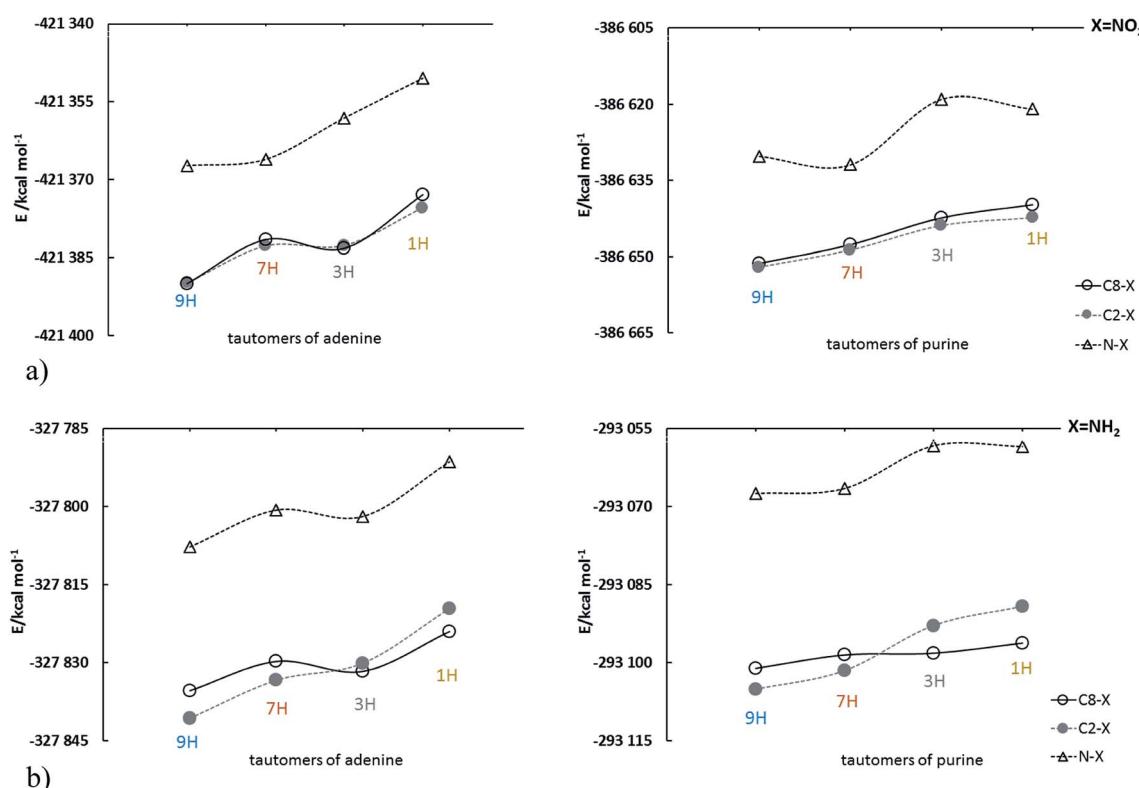


Fig. 1 Energy changes of adenine and purine tautomers in sequence: 9H, 7H, 3H, 1H for (a)  $\text{NO}_2$  and (b)  $\text{NH}_2$  substituted derivatives; for clarity, the points are connected by a solid or dashed line.

purine derivatives, in both C8 and C2 substituent positions, the  $E_{\text{rel}}$  values for tautomers 7H have the smallest range of variation (2.91 and 0.99 kcal mol<sup>-1</sup> for adenine, and 1.15 and 0.59 kcal mol<sup>-1</sup> for purine, respectively). The direct neighborhood of the C8 substituent in 9H and 7H tautomers is almost identical: one nitrogen atom with a lone pair in the molecular plane ( $2p_{xy}$ ) and one electron on the perpendicular  $2p_z$  orbital and second nitrogen (in N-H group) with  $2p_z$  electron pair. Thus, even if there are some proximity effects, they are similar, and the resulting  $E_{\text{rel}}$  differences are small. For C2-X substitution the ranges are the smallest, since there are no proximity effects that appear in the vicinity of C8. No such situations are found in other cases, with much larger  $E_{\text{rel}}$  ranges, generally between 8.07 kcal mol<sup>-1</sup> (3H, C8-X substitution in adenine) and 13.21 kcal mol<sup>-1</sup> (1H, C2-X substitution in adenine). In addition, it is worth noting that the substituents, except for the N-X cases with X = NO<sub>2</sub>, do not change the stability of purine tautomers (Tables S2 and S3<sup>†</sup>), which decreases in order of H9, H7, H3 and H1.

### Substituent effect (SE) descriptor based on relative stability of tautomer

The use of relative energy ( $E_{\text{rel}}$ ) differences with respect to the unsubstituted systems (X = H) allow for easier analysis of the SE. For 7H tautomers, this difference can be written as the energy of the homodesmotic reaction, named SESE<sub>rel</sub>:

$$\begin{aligned} 9\text{H} - \text{X} + 7\text{H} &\rightarrow 7\text{H} - \text{X} + 9\text{H} \\ \text{SESE}_{\text{rel}} &= E_{7\text{H-X}} + E_{9\text{H}} - (E_{9\text{H-X}} + E_{7\text{H}}) \\ &= E_{\text{rel 7H-X}} - E_{\text{rel 7H}} \end{aligned} \quad (3)$$

Analogous equations can be written for 3H and 1H tautomers. The obtained SESE<sub>rel</sub> values for 7H, 3H and 1H substituted adenine and purine tautomers are gathered in Tables 1 and 2, while the comparison of the substituent effect observed for adenine and purine derivatives is shown in Fig. 2; in addition, the statistics of obtained linear regressions are collected in Table S4.<sup>†</sup>

To reveal the influence of the presence of the amino group on the substituent effect the dependencies of SESE<sub>rel</sub>(adenine) on SESE<sub>rel</sub>(purine) are analyzed (Fig. 2 and Table S4<sup>†</sup>). Three substitution positions are shown: C8-X, C2-X and N-X. For the C-X series, a slope greater than 1.00 has been found for derivatives of 7H C8-X and 1H C2-X, its values are 2.45 ( $\pm 0.14$ ) and 1.176 ( $\pm 0.040$ ), respectively. In both cases NH is located near the amino group of adenine and a substituted carbon atom. In other cases, the slopes are lower than 1.0. In all C-X cases  $R^2 > 0.9$ , with the exception of 7H C2-X systems.

Relationships between SESE<sub>rel</sub> values (adenine vs. purine) of N-X substituted tautomers do not lead to unequivocal conclusions. However, when some of the most deviating points are omitted (empty circles in Fig. 2c), the reasonable correlations can be noticed. The slope is the highest for 7H ( $1.15 \pm 0.15$ , without X = CHO) then 0.71 ( $\pm 0.12$ ) for 1H (without X = NO<sub>2</sub> and OH) and finally 0.673 ( $\pm 0.031$ ,  $R^2 = 0.983$ ) for 3H substituted tautomers.

Thus, the comparison of the substituent effect, based on SESE<sub>rel</sub> model, in adenine with that present in purine, allows to conclude that more significant effect is observed in adenine when substituent is attached to the 5-membered ring.

Hammett  $\sigma$  constants are the most commonly used to describe the electron donating/accepting properties of the substituents. Their application to the data in Tables 1 and 2, this is SESE<sub>rel</sub> vs.  $\sigma$  regression (see Table S5<sup>†</sup>), leads to the observation that in the case of C8-X-substituted 7H tautomer, both adenine and purine, the determination coefficients are 0.878 and 0.810, respectively, with a large difference in the slope values: 2.206 ( $\pm 0.291$ ) and 0.855 ( $\pm 0.147$ ), respectively. This difference in the slopes can be related to a significant influence of the amino group in adenine on charge transfer from the substituent to the molecular moiety. It is important to note that in addition to these results, all other linear regressions with substituent constants always lead to  $R^2$  lower than 0.3 (for 3H C8-X purine series). These results provide important information that the model of interactions which describes well the substituent effect in benzene derivatives fails completely in some cases of adenine and purine derivatives. This problem will be discussed in more details later.

### Comparison of substituent effects from positions C8 and C2

Differences between the energies of pairs of the same molecule with different substitution positions allow to compare the substituent effect from these positions. For all adenine and purine tautomer series with substituents at the C8-X and C2-X positions

$$\Delta E_{8-2} = E_{\text{C8-X}} - E_{\text{C2-X}} \quad (4)$$

where  $E_{\text{C8-X}}$  and  $E_{\text{C2-X}}$  are the energies of the mono-substituted molecule at positions C8-X and C2-X, respectively.

It should be noticed that the close vicinity of C2 and C8 substitution positions in adenine and purine is topologically identical. The obtained  $\Delta E_{8-2}$  values for all adenine and purine tautomers are given in Tables 3 and 4, respectively.

Some interesting points resulting from the comparison of the data presented in Tables 3 and 4 can be discussed. In both series (adenine and purine), for 7H and 9H tautomers the differences are positive (except for CHO and Me derivatives), and for 1H and 3H are mostly negative, what can be seen in the averaged data for all substituents. The mean values of  $\Delta E_{8-2}$  averaged for all four tautomers of adenine and purine derivatives do not correlate with each other ( $R^2 = 0.421$ ), while the range values do very well ( $R^2 = 0.954$ ), as shown in Fig. S1.<sup>†</sup> This can be interpreted as follows: the nature of interactions caused by the substituent is different in purine and adenine, but the strength of the interactions is alike.

Fig. 3 presents a comparison of the obtained  $\Delta E_{8-2}$  values for adenine and purine – in other words – shows how the amino group in adenine influences interactions in the molecules. As mentioned above, the closest vicinity of positions C2 and C8 (Scheme 1) is similar. Either there are two nitrogen atoms with one  $\pi$ -electron and a lone pair in the plane of the molecule, or there is one N-atom of this kind and NH group, with  $2p_z$  electron pair. Obviously, in the first case, in N-C2-X-N fragment the nitrogen atoms can act as

**Table 1** The obtained  $SESE_{rel}$  (in  $\text{kcal mol}^{-1}$ ) values for 7H, 3H and 1H adenine and purine tautomers substituted at the C8 and C2 positions (Scheme 1)

	Adenine						Purine					
	C8-X			C2-X			C8-X			C2-X		
	7H	3H	1H									
NO <sub>2</sub>	1.32	-0.28	0.22	-0.09	0.08	-2.55	0.38	0.27	0.05	0.04	-0.53	-1.81
CN	1.05	-2.58	-2.17	0.36	2.69	1.49	0.33	-2.14	-2.32	0.09	2.25	1.78
CHO	0.84	3.45	4.22	-0.21	-4.42	-8.39	0.25	3.65	4.43	0.32	-5.02	-6.54
Cl	-0.19	-2.5	-2.95	0.17	2.48	2.12	-0.09	-2.56	-3.11	-0.01	2.02	1.76
F	-0.47	-4.62	-5.57	0.12	4.72	4.82	-0.18	-4.84	-5.80	-0.14	4.43	4.28
H	<b>0.00</b>											
Me	-0.26	0.45	0.10	-0.02	-0.57	-0.64	-0.04	-0.01	-0.09	0.12	-0.66	-0.55
OMe	-1.59	-1.33	-2.41	-0.63	0.46	-0.34	-0.77	-2.53	-2.78	-0.27	0.01	-0.27
OH	-1.58	-2.21	-3.90	-0.39	1.61	1.25	-0.74	-3.26	-4.03	0.30	1.46	1.42
NH <sub>2</sub>	-1.48	-3.32	-5.44	0.14	3.52	4.21	-0.67	-5.74	-6.56	0.30	3.55	4.42
Range	2.91	8.07	9.79	0.99	9.14	13.21	1.15	9.39	10.99	0.59	9.44	10.96
Average	-0.24	-1.29	-1.79	-0.05	1.06	0.20	-0.15	-1.72	-2.02	0.08	0.75	0.45
SD	1.08	2.29	3.00	0.29	2.57	3.75	0.44	2.76	3.24	0.19	2.66	3.17

hydrogen bond acceptors, whereas in NH-C8-X-N they may play a role of a donor and an acceptor, respectively.

Despite the similarity of the closest vicinity for substituents in the C2-X and C8-X positions and the fact that their energy difference is considered, differences appear in  $\Delta E_{8-2}$  for tautomers of adenine and purine. For 1H and 3H tautomers interactions are stronger in purine derivatives, and for 7H and 9H they are stronger in adenine series (Fig. 3). It is important to say that except for the 3H tautomer in all other cases the substituent effects are well linearly correlated with each other ( $R^2 > 0.945$ ) and slopes variety (between 0.868 and 1.256) could indicate dependence of the systems sensibility on the position of the substituent. Undoubtedly, the difference should be related to the presence of NH<sub>2</sub> group in adenine.

### SESE model based on adenine and purine reactions

Energies of substituted purine (PU) and adenine (AD) can be applied for homodesmotic reactions:

**Table 2** The obtained  $SESE_{rel}$  (in  $\text{kcal mol}^{-1}$ ) values for 7H, 3H, 1H adenine and purine tautomers substituted at the N-X positions

N-X	Adenine			Purine		
	7H	3H	1H	7H	3H	1H
NO <sub>2</sub>	-5.97	2.12	-0.13	-4.99	2.54	-2.20
CN	-3.20	3.90	0.87	-1.44	4.97	3.33
CHO	-4.87	2.20	0.18	-1.11	3.79	2.27
Cl	-1.84	-0.61	-1.02	-1.32	-1.79	-0.23
F	-2.88	-2.56	-4.10	-2.08	-4.22	-4.48
H	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
Me	0.67	-0.44	1.11	0.10	-0.89	1.51
OMe	-2.47	-1.95	-3.71	-0.89	-3.03	-2.46
OH	-1.12	-3.57	-2.28	-0.45	-5.59	0.77
NH <sub>2</sub>	-0.09	-1.23	-0.56	0.73	-1.60	1.52
Range	6.64	7.47	5.21	5.71	10.56	7.82
Average	-2.18	-0.21	-0.96	-1.15	-0.58	0.00
SD	2.04	2.34	1.73	1.59	3.45	2.41

$$9H\ C8-X-AD + 9H\ PU \rightarrow 9H\ AD + 9H\ C8-X-PU$$

$$SESE_{PU} = E_{9H\ AD} + E_{9H\ C8-X-PU} - (E_{9H\ C8-X-AD} + E_{9H\ PU}) \quad (5)$$

For C2-X substituted systems of the 9H tautomer we can write:

$$SESE_{PU} = E_{9H\ AD} + E_{9H\ C2-X-PU} - (E_{9H\ C2-X-AD} + E_{9H\ PU}) \quad (6)$$

Analogous equations can be written for the remaining substituted tautomers. The results of the  $SESE_{PU}$  approach are gathered in Table 5. Positive  $SESE_{PU}$  values indicate greater stability of substrates than reaction products.

Observed  $SESE_{PU}$  values are usually negative (*i.e.* destabilizing interactions) for electron-donating substituents and positive for the electron-attracting ones. In the case of strongly interacting substituents, for the C8-NH<sub>2</sub> and C2-NH<sub>2</sub> systems the average  $SESE_{PU}$  values (in  $\text{kcal mol}^{-1}$ ) are -2.47 and -0.46, whereas 2.37 and 1.72 for NO<sub>2</sub> derivatives, respectively. These differences in the energy values of the substituent effects in adenine, in comparison with purine, can be connected to the presence of the amine group at the 6-position in adenine.

It is interesting to compare the  $SESE_{PU}$  values with the traditional substituent effect descriptor, *i.e.* Hammett constants. Its mutual correlations for C8-X and C2-X adenine series are shown in Fig. 4, in addition, statistical details are presented in Table S6† (for both  $\sigma_p$  and  $\sigma_m$  relations).

A verification of the obtained interrelationships shows: (i) for C8-X systems  $SESE_{PU}$  correlates better with  $\sigma_p$  than with  $\sigma_m$ , (ii) the same applies to C2-X series of 3H and 1H tautomers, when X = CHO systems are excluded, whereas (iii) for C2-X derivatives of 9H and 7H tautomers, a better correlation of  $SESE_{PU}$  with  $\sigma_m$  than with  $\sigma_p$  can be found.



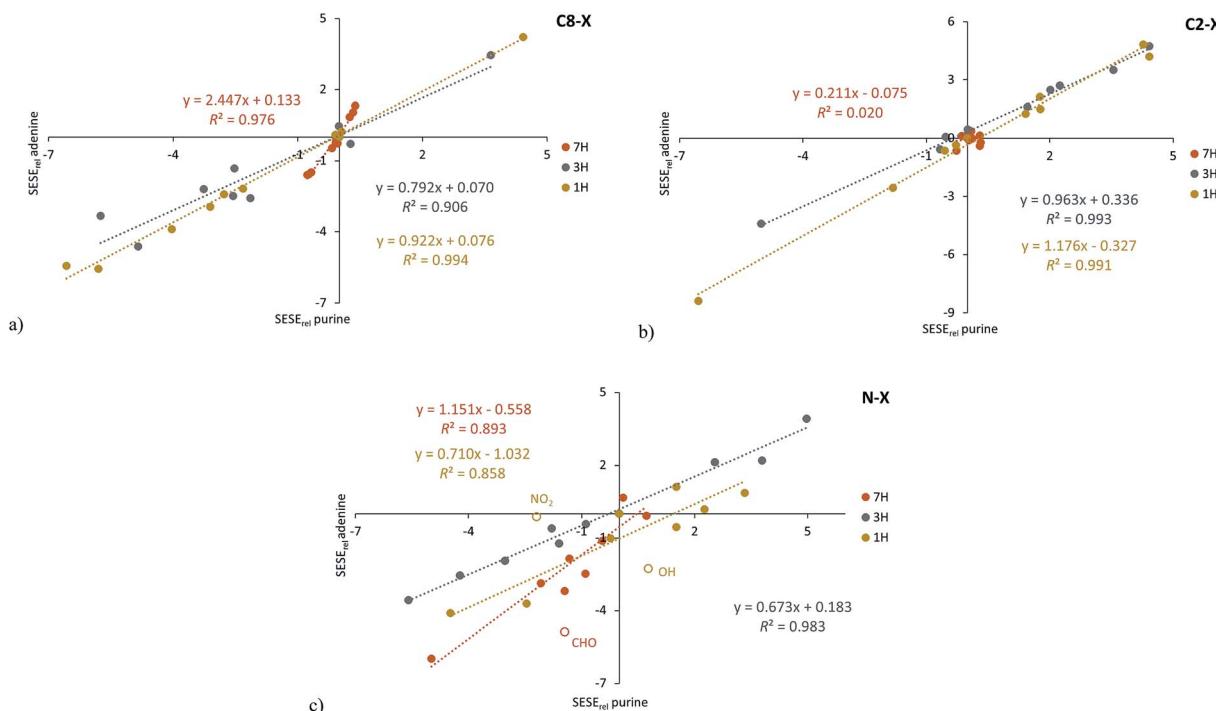


Fig. 2 Correlations between  $SESE_{rel}$  values (in  $\text{kcal mol}^{-1}$ ) of adenine and purine tautomers for C8-X (a), C2-X (b) and N-X (c) positions of substituents; empty circles indicate substituents not included in the regression.

Since  $SESE_{PU}$  values are based on differences in the energy values of the analogically substituted adenine and purine derivatives, obtained values inform about changes related to the amino group present in adenine tautomers. In the 7H tautomer the  $2p_z$  electron pair residing at N7 atom hinders charge transfer between electron-attracting substituents at the C8 position and the amino group at C6 (Scheme 1). This can be confirmed by the smallest value of the slope for this series. For 9H tautomers, where there is no such problem with the charge transfer, the interaction is stronger, hence the slope is steeper. The situation is more complex for 3H and 1H tautomers and

requires further study. However, it should be emphasized that for C8-X derivatives of 3H and 1H tautomers, the substituent effect is stronger than that observed for 9H and 7H, and the strongest occurring for the 3H C8-X adenine series (slope =  $5.08 \pm 0.40$ , Table S6†).

### The proximity effects

The proximity effect is associated with interactions of the functional group/substituent with the neighboring groups or even with the hydrogen atom, and in some cases with lone electron pairs. Due to this type of interaction, the electronic

Table 3 Energy difference values,  $\Delta E_{8-2}$  (eqn (4), in  $\text{kcal mol}^{-1}$ ), of the studied C8-X and C2-X substituted adenine (AD) tautomers

AD	$\Delta E_{8-2}$					
	9H	7H	3H	1H	Average	
$\text{NO}_2$	-0.16	1.25	-0.52	2.61	0.79	3.12
CN	0.01	0.71	-5.26	-3.65	-2.05	5.96
CHO	-5.70	-4.64	2.17	6.92	-0.31	12.62
Cl	2.95	2.59	-2.02	-2.12	0.35	5.07
F	7.45	6.86	-1.89	-2.94	2.37	10.39
H	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	
Me	-1.19	-1.43	-0.16	-0.45	-0.81	1.26
OMe	2.75	1.79	0.96	0.69	1.55	2.07
OH	4.45	3.26	0.82	-0.70	1.96	5.15
NH <sub>2</sub>	5.24	3.62	-1.60	-4.41	0.71	9.65
Range	13.15	11.51	7.43	11.33		
Average	1.58	1.40	-0.75	-0.41		
SD	3.77	3.11	2.08	3.33		

Table 4 Energy difference values,  $\Delta E_{8-2}$  (eqn (4), in  $\text{kcal mol}^{-1}$ ), of the studied C8-X and C2-X substituted purine (PU) tautomers

PU	$\Delta E_{8-2}$					
	9H	7H	3H	1H	Average	
$\text{NO}_2$	0.70	1.04	1.49	2.56	1.45	1.86
CN	0.69	0.93	-3.70	-3.41	-1.37	4.63
CHO	-4.36	-4.43	4.31	6.61	0.53	11.04
Cl	2.16	2.08	-2.42	-2.71	-0.22	4.87
F	6.19	6.15	-3.08	-3.89	1.34	10.08
H	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	
Me	-1.62	-1.78	-0.96	-1.15	-1.38	0.81
OMe	1.22	0.72	-1.32	-1.29	-0.17	2.54
OH	3.16	2.13	-1.56	-2.29	0.36	5.45
NH <sub>2</sub>	4.01	3.03	-5.28	-6.97	-1.30	10.98
Range	10.55	10.58	9.59	13.58		
Average	1.21	0.99	-1.25	-1.25		
SD	2.95	2.82	2.73	3.73		



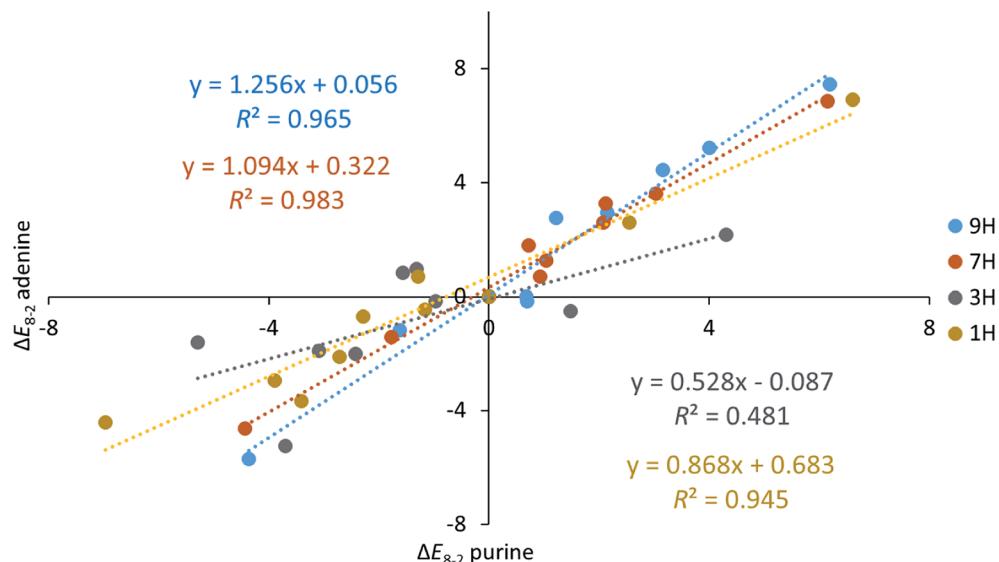


Fig. 3 Relations between the  $\Delta E_{8-2}$  values (in  $\text{kcal mol}^{-1}$ ) obtained for adenine and purine tautomers.

properties of the group/substituent can be significantly modified. Applications of the substituent effect descriptors estimated for *para*- and *meta*- substituted benzoic acids, such as the  $\sigma$  constants, to systems with proximity effects are usually not resulting with satisfactory correlations or even fail.<sup>43–45</sup> Sayyed and Suresh<sup>46</sup> showed that the use molecular electrostatic potential and a molecular fragment approach in conjunction with a rotation experiment on the COOH group of benzoic acid allowed quantifying the effects of proximity.

In the case of substituted adenine derivatives, the proximity effects appear frequently, and some examples are presented in detail below. As mentioned above, for C2- or C8- substituted adenines, depending on the type of tautomers, the substituent can be in direct neighborhood of the lone pair on the nitrogen atom in the plane of the molecule (or two such pairs) and the NH with the  $2p_z$  electron lone pair. Thus, some tautomers of adenine with asymmetric substituents, for example C8-CHO in 7H and 9H and

the C2-CHO in 1H and 3H systems, may exist in two conformations. One, with a low energy, when CH (from -CHO) interacts with a lone pair and the CO part with the NH, while in the second conformation, when CHO is rotated by  $180^\circ$ , strong repulsive interactions appear: CO is interacting with the lone pair and CH is interacting with NH. The observed difference in energy between these two rotamers is 6.59, 5.50, 9.53 and 5.92  $\text{kcal mol}^{-1}$ , respectively. Similar effects are observed for other nonsymmetric substituents, such as X = OH and OMe (Table S7†). These are examples of systems that can be obtained as energy molecule optimization results (local minima) and reveal possible interactions in the case of studied systems. Other substituents in the C2- and C8- position of adenine may have weaker proximity effects, but they still appear, and in many cases worse correlations with substituent constants or other substituent effect descriptors are observed. In this case, the proximity interactions CO···HN and CH···N (shown in Scheme 2) predominate the classical substituent effect based on resonance and inductive/field interactions. In other cases, this effect is not always significant, but it may contribute to lower values of determination coefficients, as shown in Table S6† and a few tables below.

Different interactions can be noticed in N-substituted adenine derivatives. Here, the proximity effect is mainly observed for N7- and N1- substituted species, where amino group in the 6-position is neighboring to substituents. Good examples of proximity effects are intramolecular hydrogen bonds presented in Scheme 2b.

#### SESE model based on reactions with benzene or aniline as references

So far, substituted adenine and purine series, differing only in the NH<sub>2</sub> group at C6, have been compared. One can ask how different are substituent effects observed in adenine and purine derivatives compared to those in benzene (BEN) and anilines (AN), which are considered classical objects in the SE research.

The corresponding equations are as follows:

$$\text{SESE}_{\text{BEN}} = E_{9\text{H AD}} + E_{\text{X-BEN}} - (E_{9\text{H C8-X-AD}} + E_{\text{BEN}}) \quad (7)$$



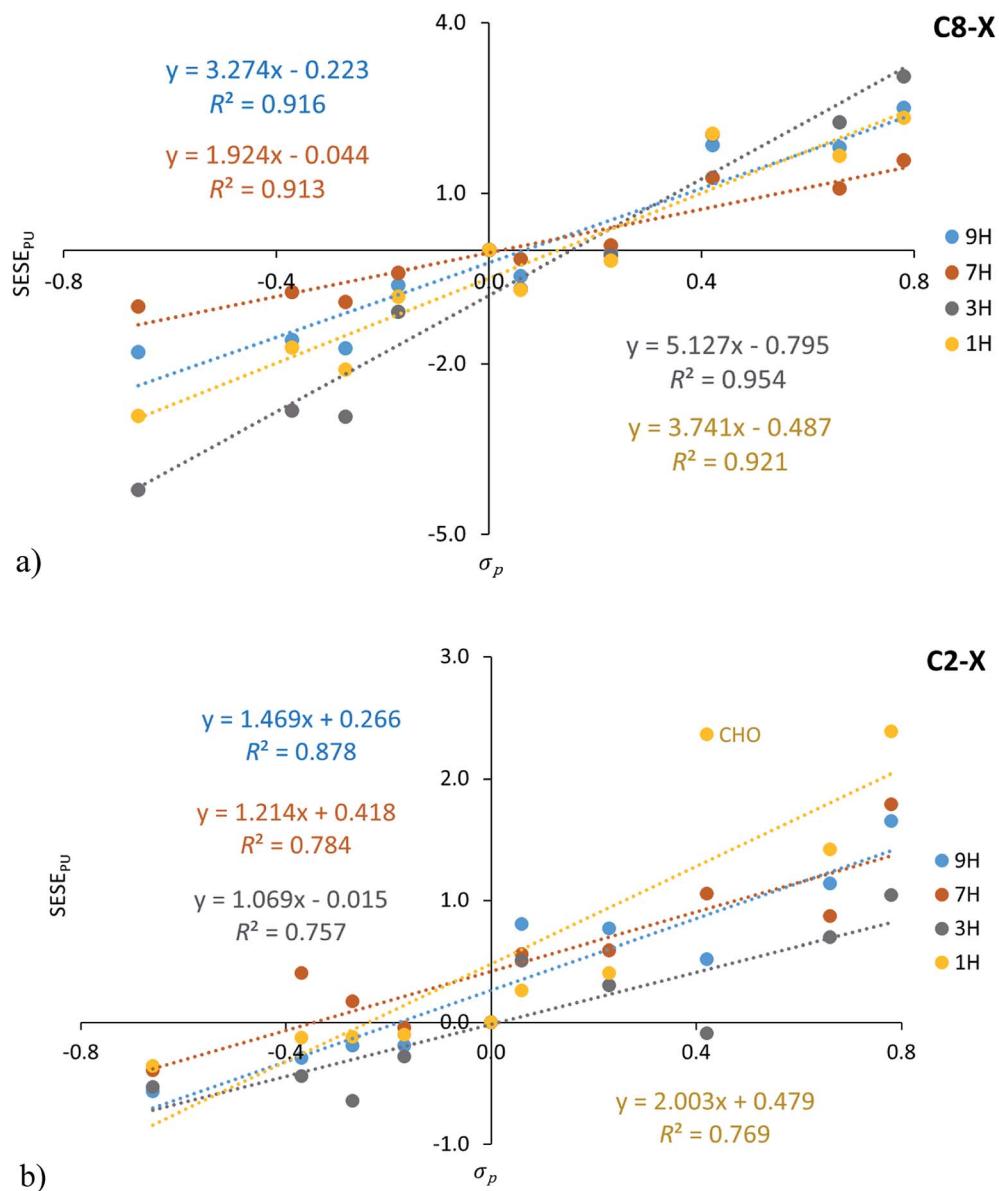
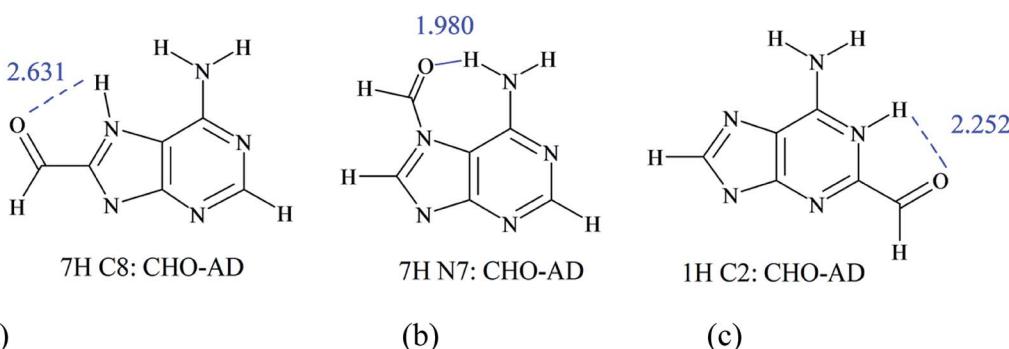


Fig. 4 Dependence of  $\text{SESE}_{\text{PU}}$  (in  $\text{kcal mol}^{-1}$ ) on substituent constants  $\sigma_p$  for C8-X (a) and C2-X (b) positions of substituents in adenine systems.

$$\text{SESE}_{\text{AN}} = E_{9\text{H AD}} + E_{4\text{-X-AN}} - (E_{9\text{H C8-X-AD}} + E_{\text{AN}}) \quad (8)$$

Analogous equations can be written for the remaining substituted tautomers of adenine and purines. The obtained

$\text{SESE}_{\text{BEN}}$  and  $\text{SESE}_{\text{AN}}$  values are presented in Tables 6 and S8,<sup>†</sup> respectively. It should be noted that the difference in values obtained for analogous derivatives of adenine and purines is  $\text{SESE}_{\text{PU}}$ .



Scheme 2 CO···HN interactions in selected formyl-adenine systems, with O···H distance marked in blue.

In almost all cases the obtained values for both  $\text{SESE}_{\text{BEN}}$  and  $\text{SESE}_{\text{AN}}$  are negative for electron-attracting substituents and positive for electron-donating ones. In addition, their absolute values are higher for purine systems than for adenine ones. This can be depicted by the slope of a linear relation of the SESE descriptors ( $\text{SESE}_{\text{BEN}}$  and  $\text{SESE}_{\text{AN}}$ ) between the values for adenine and purine systems as shown in Table 7. Thus, the amino group of adenine reduces the influence of the substituent effect. Moreover, it can be noticed that this reduction is more visible for substitution at C8-X than at C2-X. In all cases the discussed regressions are characterized by very high determination coefficients ( $R^2 > 0.95$ ).

Both  $\text{SESE}_{\text{BEN}}$  and  $\text{SESE}_{\text{AN}}$  are based on homodesmotic reactions with benzene derivatives as references. Thus, these descriptors, by definition, include a comparison of the substituent effect in heterocyclic systems (adenine or purine) with that realized in traditional benzene series (mono-substituted benzene or aniline derivatives). In the case of *para*-substituted aniline derivatives, we can characterize the SE using

Table 6  $\text{SESE}_{\text{BEN}}$  (in  $\text{kcal mol}^{-1}$ ) estimated for adenine (AD) (a) and purine (PU) (b) with benzene as reference

(a) $\text{SESE}_{\text{BEN}}$									
	9H		7H		3H		1H		
AD	C8-X	C2-X	C8-X	C2-X	C8-X	C2-X	C8-X	C2-X	
NO <sub>2</sub>	-4.71	-4.87	-6.03	-4.78	-4.44	-4.95	-4.93	-2.33	
CN	-4.07	-4.06	-5.12	-4.42	-1.49	-6.75	-1.90	-5.55	
CHO	1.90	-3.80	1.06	-3.59	-1.56	0.61	-2.33	4.59	
Cl	-2.82	0.13	-2.63	-0.04	-0.33	-2.35	0.13	-1.99	
F	-1.85	5.60	-1.38	5.48	2.76	0.87	3.71	0.77	
H	<b>0.00</b>								
Me	3.24	2.05	3.50	2.08	2.79	2.62	3.14	2.70	
OMe	4.41	7.17	6.01	7.80	5.74	6.71	6.82	7.51	
OH	4.06	8.51	5.64	8.90	6.26	6.90	7.95	7.26	
NH <sub>2</sub>	2.98	8.21	4.46	8.08	6.30	4.00	8.42	4.00	
Range	9.13	13.38	12.03	13.68	10.73	13.65	13.35	13.06	
Average	0.31	1.89	0.55	1.95	1.60	0.77	2.10	1.70	
SD	3.47	5.22	4.34	5.34	3.74	4.55	4.64	4.29	
(b) $\text{SESE}_{\text{BEN}}$									
	9H		7H		3H		1H		
PU	C8-X	C2-X	C8-X	C2-X	C8-X	C2-X	C8-X	C2-X	
NO <sub>2</sub>	-7.22	-6.53	-7.61	-6.57	-7.49	-6.00	-7.27	-4.72	
CN	-5.89	-5.20	-6.22	-5.29	-3.75	-7.45	-3.57	-6.98	
CHO	0.04	-4.32	-0.22	-4.64	-3.61	0.70	-4.39	2.22	
Cl	-2.79	-0.64	-2.71	-0.63	-0.24	-2.65	0.32	-2.40	
F	-1.40	4.79	-1.22	4.92	3.44	0.36	4.40	0.51	
H	<b>0.00</b>								
Me	3.86	2.24	3.90	2.12	3.87	2.90	3.95	2.80	
OMe	6.14	7.36	6.91	7.62	8.67	7.35	8.92	7.63	
OH	5.63	8.80	6.37	8.50	8.90	7.33	9.67	7.38	
NH <sub>2</sub>	4.77	8.78	5.44	8.47	10.51	5.22	11.33	4.36	
Range	13.36	15.32	14.51	15.06	18.00	14.80	18.60	14.60	
Average	0.31	1.53	0.46	1.45	2.03	0.78	2.34	1.08	
SD	4.75	5.79	5.12	5.80	6.09	5.13	6.38	4.83	

Table 7 Slope values,  $a$ , and determination coefficients,  $R^2$ , of  $\text{SESE}_{\text{BEN}}$  and  $\text{SESE}_{\text{AN}}$  relations between values obtained for adenine (AD) and purine (PU) tautomers with substituents at positions C2-X and C8-X

	C8-X		C2-X	
	$a$	$R^2$	$a$	$R^2$
<b><math>\text{SESE}_{\text{BEN}}(\text{AD})</math> vs. <math>\text{SESE}_{\text{BEN}}(\text{PU})</math></b>				
9H	0.712 $\pm$ 0.058	0.949	0.899 $\pm$ 0.026	0.994
7H	0.842 $\pm$ 0.031	0.989	0.918 $\pm$ 0.025	0.994
3H	0.612 $\pm$ 0.022	0.990	0.885 $\pm$ 0.023	0.990
1H	0.724 $\pm$ 0.019	0.995	0.871 $\pm$ 0.062	0.995
<b><math>\text{SESE}_{\text{AN}}(\text{AD})</math> vs. <math>\text{SESE}_{\text{AN}}(\text{PU})</math></b>				
9H	0.774 $\pm$ 0.026	0.991	0.927 $\pm$ 0.019	0.997
7H	0.875 $\pm$ 0.015	0.998	0.938 $\pm$ 0.018	0.997
3H	0.708 $\pm$ 0.014	0.997	0.927 $\pm$ 0.012	0.998
1H	0.792 $\pm$ 0.011	0.998	0.881 $\pm$ 0.037	0.986

their homodesmotic reaction with benzene, where energy,  $\text{SESE}_{\text{AN-BEN}}$ , can be calculated as:

$$\text{SESE}_{\text{AN-BEN}} = E_{\text{X-BEN}} + E_{\text{AN}} - (E_{\text{4-X-AN}} + E_{\text{BEN}}) \quad (9)$$

The obtained values are presented in Table S9.† In the same table, values from B3LYP/6-311++G(d,p) calculations from our previous paper,<sup>34</sup> are also given. These values are congruous with the results calculated at B97D3/aug-cc-pvdz level of theory.

The use of the SESE concept, with appropriately selected reactions, allows the comparison of the influence of a transmission moiety on the substituent effect. Purine is a reference for  $\text{SESE}_{\text{PU}}$ , whereas benzene for  $\text{SESE}_{\text{AN-BEN}}$ . The obtained results of linear relation  $\text{SESE}_{\text{PU}}(\text{adenine})$  vs.  $\text{SESE}_{\text{AN-BEN}}(-\text{aniline})$ , shown in Table 8, reveal: (i) a comparable transmission of the substituent effect only for C8-X 3H adenine (slope =  $1.062 \pm 0.104$ ); (ii) a weaker transmission of the substituent effect by the purine moiety than by benzene in all other cases; and (iii) the substituent effect acting from the adenine C8 position is stronger than from the C2 position.

In conclusion, a subsequent study confirmed that the more significant effect is observed in adenine when the substituent is attached to the 5-membered ring (substitution at C8-X position).

#### SESE model based on reactions with imidazole or pyrimidine as references

Purine (adenine) can be formed by the fusion of pyrimidine ( $\text{NH}_2$ -pyrimidine) and imidazole rings. Therefore, it can be of interest how the substituent effects observed in adenine and purine derivatives differ from those in imidazole (IM) and amino-pyrimidine (PY), substituted in a similar topological environment as in the purine (adenine) molecules. For this purpose, appropriate homodesmotic reactions, for the 9H tautomer shown in Scheme 3 (analogous reactions can be written for 7H tautomer), are used. Their energy allow to calculate  $\text{SESE}_{\text{IM}}$  and  $\text{SESE}_{\text{PY}}$ , both for adenine and purine C8-X and C2-X substituted series, respectively; for 9H adenine C8-X and C2-X derivatives, according to the following equations:



**Table 8** Slope values,  $a$ , and determination coefficients,  $R^2$ , of relation  $\text{SESE}_{\text{PU}}(\text{AD})$  vs.  $\text{SESE}_{\text{AN-BEN}}(\text{AN})$  for C2-X and C8-X substituted adenine tautomers

$\text{SESE}_{\text{PU}}(\text{AD})$ vs. $\text{SESE}_{\text{AN-BEN}}(\text{AN})$				
C8-X		C2-X		
$a$	$R^2$	$a$	$R^2$	
9H	$0.710 \pm 0.037$	0.978	$0.257 \pm 0.073$	0.609
7H	$0.417 \pm 0.024$	0.975	$0.236 \pm 0.058$	0.674
3H	$1.062 \pm 0.104$	0.929	$0.183 \pm 0.064$	0.502
1H	$0.804 \pm 0.054$	0.966	$0.455 \pm 0.054$	0.898

$$\text{SESE}_{\text{IM}} = E_{9\text{H AD}} + E_{\text{C2-X-IM}} - (E_{9\text{H C8-X-AD}} + E_{\text{IM}}) \quad (10)$$

$$\text{SESE}_{\text{PY}} = E_{9\text{H AD}} + E_{\text{C2-X-PY}} - (E_{9\text{H C2-X-AD}} + E_{\text{PY}}) \quad (11)$$

Analogous equations can be used for the remaining 9H and 7H tautomers of adenine and purine. The obtained values are presented in Table 9.

For particular substituents, calculated  $\text{SESE}_{\text{PY}}$  and  $\text{SESE}_{\text{IM}}$  values have opposite signs. As in the case of  $\text{SESE}_{\text{PU}}$ ,  $\text{SESE}_{\text{PY}}$  values are negative (*i.e.* destabilizing interactions) for electron-donating substituents and usually positive for the electron-attracting ones. The opposite changes apply to  $\text{SESE}_{\text{IM}}$ , that is for C8-X series both adenine and purine. Furthermore, the slopes of a linear relation of the SESE descriptors ( $\text{SESE}_{\text{IM}}$  and  $\text{SESE}_{\text{PY}}$ ) between the values for adenine and purine (both 9H and 7H) systems (Table 10) document that the amino group of adenine reduces the influence of the substituent effect. In addition, the substituent effect in the C8-X series is weaker than for the C2-X series and the discussed regressions are characterized by very high determination coefficients ( $R^2 > 0.92$ ).

#### Relations between $\text{SESE}_{\text{BEN}}$ , $\text{SESE}_{\text{AN}}$ , $\text{SESE}_{\text{IM}}$ and $\text{SESE}_{\text{PY}}$ descriptors and Hammett $\sigma_p$ constants

Statistical details of the mutual dependences of  $\text{SESE}_{\text{BEN}}$ ,  $\text{SESE}_{\text{AN}}$ ,  $\text{SESE}_{\text{IM}}$  and  $\text{SESE}_{\text{PY}}$  on  $\sigma_p$  for C8-X and C2-X adenine and purine derivatives are presented in Table S10.†

It should be added that  $\sigma_p$  should be understood as the generally accepted characteristic of the substituent effect in cases where contributions from inductive/field and resonance effects are equal.<sup>47–49</sup> The greater  $R^2$  of the regression  $\text{SESE}$  vs.  $\sigma_p$ , the interactions between substituents in adenine/purine series are more similar to those in benzene. This could mean that in the studied systems the classical resonance and inductive/field effects are realized nearly alike. For almost all cases in the adenine series  $R^2$  values are worse than the ones in purines (Table S10†). This can be explained by the proximity effects of the  $\text{NH}_2$  group in adenine and thus more complex nature of interactions. Therefore, benzene (and pyrimidine) as a reference for the

SE description in adenine systems is clearly worse than in purines. However, if aniline (*i.e.* 4-X-aniline) is a reference system, the determination coefficients improve, and  $\text{SESE}_{\text{AN}}$  becomes a good SE characteristic in both adenine and purine series. The same applies to  $\text{SESE}_{\text{IM}}$ , but this descriptor can only be used for the C8-X series. In addition, better correlations of  $\text{SESE}_{\text{AN}}$  with  $\sigma_p$  for C8-X series than for C2-X indicate a greater role of the resonance effect in the first case.

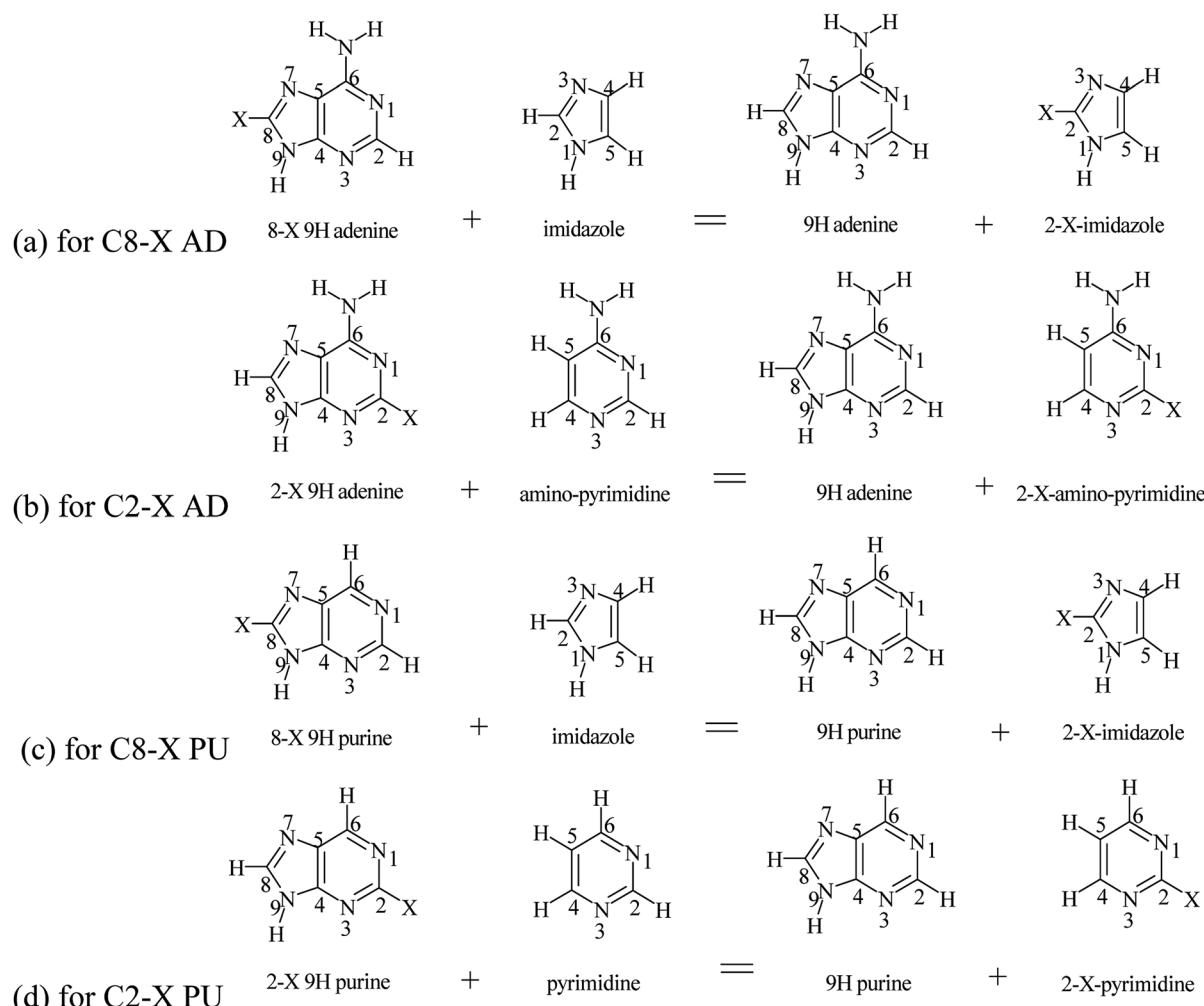
Considering the slope absolute values and benzene (aniline) as a reference, for the purine derivatives the strength of the SE increases with the stability of tautomers for C2-X series, while opposite trends can be seen for C8-X systems. In the case of adenine derivatives, the picture is not so clear. In addition, again, lower absolute slope values in the second series may be related to the presence of the amino group in adenine.

#### Classical substituent effect

The classical or traditional SE is the most common studied type of interaction – the properties of the “reaction site” Y (the fixed group in the series) are related to the properties of the X substituent in the disubstituted X-R-Y system. In the case of studied adenine derivatives, the C6-amino group is the fixed group (Y), purine moiety can be defined as R and substituents (X) can be introduced at the C2 or C8 positions (Scheme 1). The relationships between cSAR( $\text{NH}_2$ ) and  $\text{SESE}_{\text{PUR}}$  (as the SE descriptor) for all studied adenine tautomers are shown in Fig. 5 (the obtained cSAR( $\text{NH}_2$ ) values are gathered in Table S11†). In addition, statistics of these relations and others in which  $\text{SESE}_{\text{BEN}}$  and  $\text{SESE}_{\text{AN}}$  are used as the SE descriptor are presented in Table S12.†

Obtained dependences of cSAR( $\text{NH}_2$ ) on  $\text{SESE}_{\text{PUR}}$  indicate an interesting difference between regressions for the C8-X and C2-X series. It is important to note the high values of the determination coefficient (for C8-X  $R^2 \geq 0.987$ , for C2-X  $R^2 \geq 0.875$ ), which entitle us to a deeper analysis. The variability of cSAR( $\text{NH}_2$ ) sensitivity to the SE in the C8-X series is significantly greater (slopes varies from 0.012 to 0.027) than in the C2-X series (the slope ranges from 0.025 to 0.029). This discrepancy can be easily explained by resonance structures for these two systems, which resemble the interactions with electron-attracting groups observed in the case of *para* and *meta* positions in the substituted aniline<sup>34</sup> (or phenolate).<sup>50</sup> Adenine C8-X derivatives resemble *para*-substitution in benzene, where the intramolecular charge transfer between the electron donating and electron accepting substituents can be described by canonical forms with charge separation, shown in Scheme 4. In the case of C2-X systems, this transfer requires the use of canonical forms with double charge separation (as for *meta*-nitrophenolate). In general, when the number of bonds between carbon atoms to which electron attracting and electron donating substituents are attached is odd, then a charge transfer occurs (as in *para* substituted benzene), if the number is even, then interactions are like those for *meta* substituted benzene.<sup>51</sup>





Scheme 3 Homodesmotic reactions of adenine and purine with imidazole and aminopyrimidine (pyrimidine).

Table 9 SESE<sub>IM</sub> and SESE<sub>PY</sub> (in kcal mol<sup>-1</sup>) values for adenine and purine 9H and 7H tautomers with imidazole (IM) and aminopyrimidine (PY) as references

	Adenine				Purine			
	SESE <sub>IM</sub>		SESE <sub>PY</sub>		SESE <sub>IM</sub>		SESE <sub>PY</sub>	
	9H	7H	9H	7H	9H	7H	9H	7H
	C8-X	C8-X	C2-X	C2-X	C8-X	C8-X	C2-X	C2-X
NO <sub>2</sub>	-1.90	-3.21	0.54	0.63	-4.41	-4.79	0.71	0.67
CN	-1.67	-2.72	0.66	0.30	-3.49	-3.82	0.82	0.73
CHO	-1.77	-2.62	0.80	1.01	-3.63	-3.89	1.51	1.19
Cl	0.14	0.34	-0.73	-0.90	0.17	0.26	-0.72	-0.71
F	0.92	1.40	-1.16	-1.27	1.37	1.55	-1.18	-1.05
H	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
Me	1.13	1.39	-0.36	-0.33	1.74	1.78	-0.41	-0.53
OMe	2.44	4.03	-1.68	-1.05	4.17	4.93	-1.53	-1.26
OH	1.95	3.54	-1.88	-1.49	3.53	4.27	-1.20	-1.50
NH <sub>2</sub>	2.62	4.10	-1.57	-1.70	4.41	5.08	-1.82	-2.12
Range	4.52	7.32	2.68	2.72	8.82	9.87	3.34	3.31
Average	0.35	0.57	-0.49	-0.44	0.35	0.49	-0.35	-0.42
SD	1.64	2.66	0.98	0.90	3.12	3.52	1.06	1.04

An another equivalent way of counting the bonds is the number of bonds between the electron-donating atom (group D, nitrogen this time) and the electron-accepting atom (group A, usually oxygen atom in the NO or NO<sub>2</sub> groups or nitrogen atom in the CN group). In this case, if the number of all bonds between donating and accepting atoms in D and A groups, respectively, is even, then the charge transfer is privileged and well visible.<sup>52</sup>

Table 10 Slope values and determination coefficients of SESE<sub>IM</sub> and SESE<sub>PY</sub> relations between values obtained for 9H and 7H adenine (AD) and purine (PU) tautomers with substituents at positions C8-X and C2-X, respectively

SESE <sub>IM</sub> (AD) vs. SESE <sub>IM</sub> (PU)		SESE <sub>PY</sub> (AD) vs. SESE <sub>PY</sub> (PU)	
C8-X		C2-X	
<i>a</i>	<i>R</i> <sup>2</sup>	<i>a</i>	<i>R</i> <sup>2</sup>
9H	0.523 $\pm$ 0.013	0.995	0.877 $\pm$ 0.087
7H	0.753 $\pm$ 0.012	0.998	0.851 $\pm$ 0.062
			0.925
			0.959

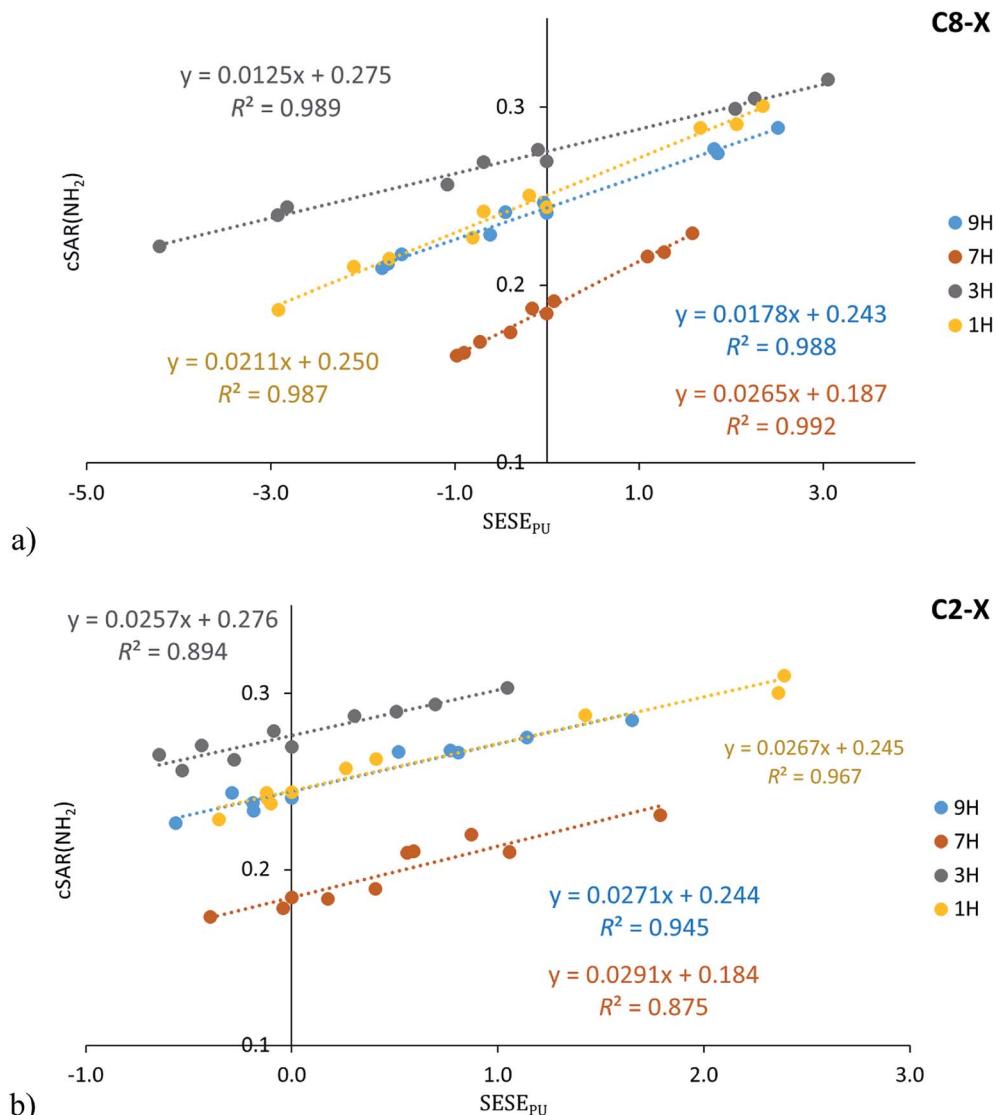


Fig. 5 Dependence of cSAR(NH<sub>2</sub>) on SESE<sub>PU</sub> (in kcal mol<sup>-1</sup>) for the C8-X (a) and C2-X (b) substituted adenine series.

The above observation is confirmed by determination coefficients of cSAR(NH<sub>2</sub>) relation to SESE<sub>BEN</sub> and SESE<sub>AN</sub> (obtained with benzene derivatives as a reference). In the case of the C8-X series, they are higher than 0.857 (excluding X = CHO, Table S12†), while in C2-X systems no correlation is observed,

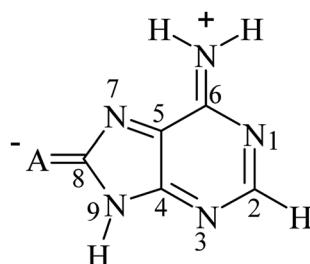
similarly as for substituted anilines in the *para* and *meta* positions, respectively.<sup>34</sup>

## Conclusions

Adenine, the amino derivative of purine, is a building block of DNA and RNA helices. Both, purine and adenine, can exist in various tautomeric forms. Therefore, impact of the substituent on their stability is a very important problem.

The strength of the substituent effect was analyzed depending both on the type of tautomers as well as on the substitution positions in the adenine and purine derivatives. The application of energetic descriptor, SESE, based on homodesmotic reactions with various reference systems, allowed to investigate different aspects of the substituent effect.

For all substituted adenine and purine series, the 9H tautomeric systems are the most stable (the only exception is the N-NO<sub>2</sub> 9H purine), while the 1H tautomer systems are the least



Scheme 4 Intramolecular charge transfer from electron donating NH<sub>2</sub> group to the electron attracting substituent.

stable. In the case of 7H and 3H tautomers, the order of energy changes for substituted adenine derivatives differs from the sequences in purines, which can be explained by the presence (interaction) of the  $\text{NH}_2$  group. The amino group in adenine induces about 1.6 times greater energy difference between the most and least stable tautomers comparing to purine. In addition, substituents, except for the  $\text{N-NO}_2$  systems, do not change stability of purine tautomers, which decreases in the order 9H, 7H, 3H and 1H.

The application of different reference systems in homodesmotic reactions and Hammett substituent constants reveals various similarity aspects in interaction due to the substituent effect.

The results of  $\text{SESE}_{\text{rel}}$  (with respect to 9H tautomer) analysis indicate the stronger substituent effect in adenine than purine derivatives for substituents attached adjacent to the  $\text{NH}$  group of the ring. In addition, a more significant effect is observed in adenine systems when the substituent is attached to the 5-membered ring. This observation is also confirmed by the results of analysis of the differences between the energies of pairs of the same molecule with the substituent at the C8-X and C2-X positions, for all tautomers.

The  $\text{SESE}_{\text{PU}}$  (purine as reference) approach informs about changes in the substituent effect due to the presence of the amino group in adenine tautomers. Like in the substituted aniline derivatives, the obtained values show stabilizing and destabilizing interactions of the amino group with electron-attracting and electron-donating substituents, respectively. Moreover, these interactions are more pronounced in C8-X than in the C2-X series.

The description of the substituent effect in adenines and purines based on their similarity to monocyclic reference systems (benzene, aniline, imidazole or pyrimidine) allows to conclude that the amino group of adenine reduces the effect of substituent when compared to purine. This weakening is more pronounced for the C8-X substitution, independently of the tautomer. However, it should be emphasized that these  $\text{SESE}$  values, with the exception of  $\text{SESE}_{\text{PY}}$ , suggest stabilizing interactions in the case of electron-donating substituents.

Dependences of  $\text{SESE}$  descriptors on the Hammett substituent constants allow to compare the similarity of the substituent effect in adenine (and purine) to that observed in benzene. The obtained relationships for  $\text{SESE}_{\text{PU}}$  reveal the similarity of the substituent effect in C8-X to the *para*, while in C2-X to the *meta* positions in benzene (documented by very good correlations with  $\sigma_p$  and  $\sigma_m$ , respectively). As in the case of aniline derivatives,<sup>34</sup> the sensitivity of the amino group to the substituent effect, along with much better correlations, exhibit C8-X systems when compared to C2-X ones. In addition, weaker transmission of the substituent effect by the purine moiety than by benzene is observed.

Furthermore, analysis of the classical substituent effect in adenine confirms the more pronounced substituent effect from the C8-position what can be explained by the stronger resonance effect in such arrangements.

Specific interactions between substituents and neighboring N or NH are the reason for deviations and lower

similarity of the substituent effect when compared with systems without these kind of interactions. Particularly strong deviations are observed for  $\text{CHO}$  group.

## Conflicts of interest

The authors declare no competing financial interest.

## Acknowledgements

We gratefully acknowledge the Interdisciplinary Center for Mathematical and Computational Modeling (Warsaw, Poland) and Wrocław Centre for Networking and Supercomputing (<http://wcss.pl>; grant No. 311) for providing computer time and facilities. Authors thank the National Science Centre and Ministry of Science and Higher Education of Poland for supporting this work under the grant no. UMO-2016/23/B/ST4/00082. P. H. M. would like to acknowledge Operational Project Knowledge Education Development 2014–2020 co-financed by European Social Fund.

## Notes and references

- 1 W. Saenger, *Principles of Nucleic Acids Structures*, Springer, New York, 1990.
- 2 E. D. Raczyńska, M. Makowski, K. Zientara-Rytter, K. Kolczyńska, T. M. Stępniewski and M. Hallmann, *J. Phys. Chem. A*, 2013, **117**, 1548–1559.
- 3 R. K. Singh, J. V. Ortiz and M. K. Mishra, *Int. J. Quantum Chem.*, 2010, **110**, 1901–1915.
- 4 E. D. Raczyńska, M. Makowski, M. Hallmann and B. Kaminska, *RSC Adv.*, 2015, **5**, 36587–36604.
- 5 Y. Kyogoku, R. C. Lord and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 1967, **57**, 250–257.
- 6 D. W. Miles, R. K. Robins and H. Eyring, *J. Phys. Chem.*, 1967, **71**, 3931–3941.
- 7 S. Diekman, E. von Kitzing, L. McLaughlin and J. Ott, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 8257–8261.
- 8 H. S. Koo and D. M. Krothers, *Biochemistry*, 1987, **26**, 3745–3748.
- 9 T. F. Kagawa, M. L. Howell, K. Tseng and P. S. Ho, *Nucleic Acids Res.*, 1993, **21**, 5978–5986.
- 10 G. T. Hwang and F. E. Romesberg, *Nucleic Acids Res.*, 2006, **34**, 2037–2045.
- 11 C. Fonseca Guerra, T. van der Wijst and F. M. Bickelhaupt, *ChemPhysChem*, 2006, **7**, 1971–1979.
- 12 A. C. Castro, M. Swart and C. Fonseca Guerra, *Phys. Chem. Chem. Phys.*, 2017, **19**, 13496–13502.
- 13 A. Ebrahimi, S. M. H. Khorassani, H. Delarmi and H. Esmaeeli, *J. Comput.-Aided Mol. Des.*, 2010, **24**, 409–416.
- 14 Y. L. Chen, D. Y. Wu and Z. Q. Tian, *J. Phys. Chem. A*, 2016, **120**, 4049–4058.
- 15 V. Nikolova and B. Galubov, *Maced. J. Chem. Chem. Eng.*, 2015, **34**, 159–167.
- 16 T. M. Krygowski and B. T. Stępień, *Chem. Rev.*, 2005, **105**, 3482–3512.
- 17 L. P. Hammett, *J. Am. Chem. Soc.*, 1937, **59**, 96–103.



18 L. P. Hammett, *Physical Organic Chemistry*, McGraw-Hill, New York, 1940, p. 196.

19 C. Hansch, A. Leo and R. W. Taft, *Chem. Rev.*, 1991, **91**, 165–195.

20 O. Exner, in *Advances in Linear Free Energy Relationships*, ed. N. B. Chapman and J. Shorter, Plenum Press, London, 1972, ch. 1, p. 1.

21 C. D. Johnson, *The Hammett Equation*, Cambridge University Press, 1973.

22 J. Shorter, in *Similarity Models in Organic Chemistry, Biochemistry and Related Fields*, ed. R. I. Zalewski, T. M. Krygowski, and J. Shorter, Elsevier, Amsterdam, 1991, ch. 2, p. 77.

23 O. Exner and S. Böhm, *Curr. Org. Chem.*, 2006, **10**, 763–778.

24 W. J. Hehre, L. Radom, P. v. R. Schleyer and J. A. Pople, *Ab initio Molecular Orbital Theory*, J. Wiley&Sons, N.Y., 1986.

25 S. M. Bachrach, *Computational Organic Chemistry*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2014.

26 P. George, M. Trachtman, C. W. Bock and A. M. Brett, *J. Chem. Soc., Perkin Trans. 2*, 1976, **11**, 1222–1227.

27 A. Pross, L. Radom and R. W. Taft, *J. Org. Chem.*, 1980, **45**, 818–826.

28 M. K. Cyrański, *Chem. Rev.*, 2005, **105**, 3773–3811.

29 T. Siodla, W. P. Oziminski, M. Hoffmann, H. Koroniak and T. M. Krygowski, *J. Org. Chem.*, 2014, **79**, 7321–7331.

30 T. M. Krygowski and N. Sadlej-Sosnowska, *Struct. Chem.*, 2011, **22**, 17–22.

31 N. Sadlej-Sosnowska, *Chem. Phys. Lett.*, 2007, **447**, 192–196.

32 N. Sadlej-Sosnowska and T. M. Krygowski, *Chem. Phys. Lett.*, 2009, **476**, 191–195.

33 O. A. Stasyuk, H. Szatylowicz, C. F. Guerra and T. M. Krygowski, *Struct. Chem.*, 2015, **26**, 905–913.

34 H. Szatylowicz, T. Siodla, O. A. Stasyuk and T. M. Krygowski, *Phys. Chem. Chem. Phys.*, 2016, **18**, 11711–11721.

35 T. Siodla, H. Szatylowicz, K. S. Varaksin and T. M. Krygowski, *RSC Adv.*, 2016, **6**, 96527–96530.

36 H. Szatylowicz, A. Jezuita, T. Siodla, K. S. Varaksin, M. A. Domanski, K. Ejsmont and T. M. Krygowski, *ACS Omega*, 2017, **2**, 7163–7171.

37 G. S. Remya and C. H. Suresh, *Phys. Chem. Chem. Phys.*, 2016, **18**, 20615–20626.

38 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci and G. A. Petersson, *et al.*, *Gaussian 09, Revision D.01*, Gaussian Inc., Wallingford, CT, 2013.

39 P. H. Marek, H. Szatylowicz and T. M. Krygowski, *Struct. Chem.*, 2019, **30**, 351–359.

40 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

41 J. D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.

42 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007–1023.

43 R. W. Taft, *J. Am. Chem. Soc.*, 1953, **75**, 4231–4238.

44 R. W. Taft, *Steric Effects in Organic Chemistry*, ed. M. S. Newman, Wiley, New York, 1956, p. 587.

45 M. Charton, *Prog. Phys. Org. Chem.*, 1971, **8**, 235–317.

46 F. B. Sayyed and C. H. Suresh, *New J. Chem.*, 2009, **33**, 2465–2471.

47 R. W. Taft and I. C. Lewis, *J. Am. Chem. Soc.*, 1958, **80**, 2436–2443.

48 R. W. Taft and I. C. Lewis, *J. Am. Chem. Soc.*, 1958, **80**, 5343–5352.

49 R. W. Taft, *J. Phys. Chem.*, 1960, **64**, 1805–1815.

50 M. Shahamirian, H. Szatylowicz and T. M. Krygowski, *Struct. Chem.*, 2017, **28**, 1563–1572.

51 T. M. Krygowski, M. Palusiak, A. Plonka and J. E. Zachara-Horeglad, *J. Phys. Org. Chem.*, 2007, **20**, 297–306.

52 H. Szatylowicz, T. M. Krygowski, M. Palusiak, J. Poater and M. Sola, *J. Org. Chem.*, 2011, **76**, 550–556.

