

Correlation between polar surface area and bioferroelectricity in DNA and RNA nucleobases*

See-Chuan Yam¹, Sharifuddin Md. Zain¹, Vannajan Sanghiran Lee^{1,a}, and Khian-Hooi Chew^{2,b}

¹ Department of Chemistry, Faculty of Science, University of Malaya, Kuala Lumpur, 50603, Malaysia

² Department of Physics, Centre for Theoretical and Computational Physics, Faculty of Science, University of Malaya, Kuala Lumpur, 50603, Malaysia

Received 2 February 2018 and Received in final form 27 June 2018

Published online: 18 July 2018

© EDP Sciences / Società Italiana di Fisica / Springer-Verlag GmbH Germany, part of Springer Nature, 2018

Abstract. We have performed computational molecular modelling to study the polarization switching and hysteresis loop behaviours of DNA and RNA nucleobases using the PM3 semi-empirical quantum mechanical approaches. All the nucleobases: adenine (A), thymine (T), guanine (G), cytosine (C), and uracil (U) were modelled. Our study indicates that all the nucleobases exhibit a zero-field polarization due to the presence of polar atoms or molecules such as amidogen and carbonyl. The shape of polarization P versus an applied electric field E hysteresis loop is square, implying typical ferroelectrics behaviour. The total energy U as a function of an applied electric field E exhibits a butterfly-like loop. The presence of zero-field polarization and ferroelectrics hysteresis loop behaviours in nucleobases may support the hypothesis of the existence of bioferroelectricity in DNA and RNA. We also found an interesting relationship between the minimum electric field required for switching E_C and the ratio of the topological polar surface area (TPSA) to the total surface area (TSA) of a nucleobase. In particular, the E_C of a nucleobase is inversely proportional to the TPSA/TSA ratio. This work may provide useful information for understanding the possible existence of ferroelectricity in biomaterials.

1 Introduction

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) have received wide interest due to their huge storage capacity, and highly conserved sequence [1–4]. It has long been speculated that DNA and RNA exhibit ferroelectric phenomena or more specifically, the so-called bioferroelectricity. The existence of the ferroelectricity in biological systems suggested that these physical properties may play an important role in biological processes and functions. For instance, Liu *et al.* [5–7] observed that the ferroelectric switching in elastin leads to a reduction in the distal shear stress by increasing the pulsatile flow and blood pressure in the heart arteries. Their results establish ferroelectricity as an important biophysical property of proteins. These feature properties could be extremely

useful for possible applications as a biosensor to detect DNA damage and mutation [8–11].

DNA and RNA have been reported experimentally to exhibit piezoelectric and ferroelectric properties [12–16]. Polonsky *et al.* [12] studied the DNA of cock erythrocytes in the form of sodium salt. Their study reveals that this DNA possesses ferroelectrics properties. As the temperature increases to 70 °C, the permittivity of the DNA increases. At the same time, the experimental work by Stanford *et al.* [13] also found a dielectric hysteresis loop behaviour in RNA, as well as a strong increase in permittivity with increasing temperature from 30 to 110 °C. However, Yarmarkin *et al.* [17] doubt the hypothesis that DNA and RNA support ferroelectricity. In order to validate the presence of ferroelectricity in either DNA or RNA, they conducted an investigation of the dielectric properties of DNA from calf thymus under heating and cooling in the air, humid, and vacuum environment within the temperature range from 20 °C to 60 °C. They argued that the observed increase in permittivity in the DNA samples used in theirs and Polonsky *et al.*'s [12] was induced by heating, and it is not related to ferroelectricity.

* Supplementary material in the form of a .pdf file available from the Journal web page at

<https://doi.org/10.1140/epje/i2018-11696-5>

^a e-mail: vannajan@um.edu.my

^b e-mail: khchew@um.edu.my

While extensive studies of ferroelectricity in DNA and RNA have been made, it is clear that the fundamental question on the origin and mechanism behind their ferroelectricity remain unclear. This motivates the present study in which the polarization switching and hysteresis loop behaviours of nucleobases are studied by quantum mechanics calculations. Since nucleobases are the building blocks of DNA and RNA, an understanding of the physical and electrical properties of nucleobases may shed light on the possible existence of bioferroelectricity in DNA and RNA. In this work, we performed a systematic study on the polarization switching and hysteresis loop behaviours of all the five nucleobases. For simplicity, the nucleobases are assumed to be surrounded by vacuum. All the computational calculations are performed by a quantum mechanics (QM) semi-empirical (PM3) method to obtain the minimum energy points. The molecular structure and electrical properties of all the nucleobases are calculated and studied. The topological polar surface area (TPSA) and total surface area (TSA) of each nucleobase are obtained. The electric field dependence of polarization and the total energy of all the nucleobases are also systematically studied. The correlation between the ratio of TPSA/TSA and the hysteresis loop behaviour of nucleobases is explored.

2 Methods

The three-dimensional (3D) molecular structures of the DNA and RNA nucleobases (A, T, G, C, U) were modelled and built using the commercial software package HyperChem7.5 [18]. The physicochemical properties of the nucleobases were performed using the QSAR module [19] of the HyperChem7.5. The molecular TPSA were obtained using the fragment-based method [20] by incorporating the molinspiration cheminformatics server [21]. In the present study, we have also attempted to calculate the optimized structures in water environment under the presence of an applied electric field. However, the calculation with water components is prohibitively time consuming with the available computing resources. Therefore, we mainly simulate and examine the effect of an applied electric field on the polar behaviours of DNA and RNA nucleobases in a gas phase.

In the calculations, the initial stage of the nucleobases is set based on the coordinate system, as shown in fig. 1. In the polarization reversal study, an electric field is applied along the *Y*-axis in the range between 0.00 to 0.011 atomic units (a.u.) (1 a.u. \sim 514 V/nm). Geometry optimizations have been computed using the quantum mechanics (QM) PM3 method with unrestricted Hartree-Fock (UHF) where the algorithm Polak-Ribiere (conjugate gradient) is employed to determine the minimum energy point [22,23]. The optimized structure with the lowest energy value was used to determine the electrical and polar properties of nucleobases. Calculation of the molecular volume based on the van der Waals surface was performed using the QSAR option. By calculating both the molecular dipole moments *D* and van der Waals volume *V*, we then determined the polarization *P* of the molecule using $P = 3.33564 * D/V$ [24].

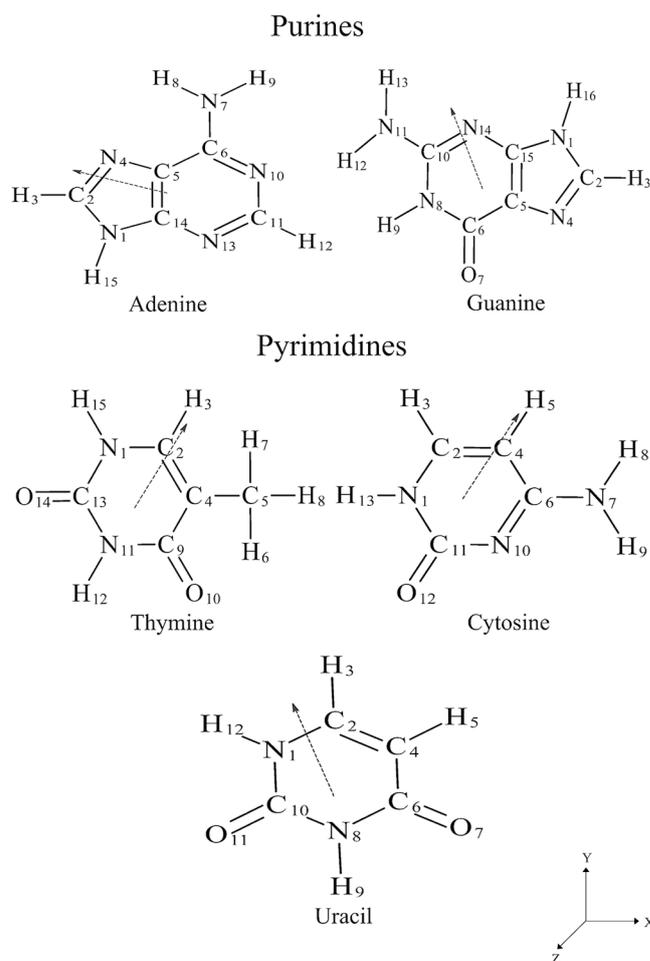


Fig. 1. Chemical structures and atom numbering for purines and pyrimidines of DNA and RNA nucleobases. Blue arrows indicate the direction of dipole moments. An electric field is applied along the *Y*-direction.

3 Results and discussion

3.1 Physical and structural properties

We first examined the molecular structures of the nucleobases under the absence of an applied electric field. In fig. 1, the chemical structures of the nucleobases (A, T, G, C, U) are shown. Four physicochemical properties and the chemical formula, number of atom, and heavy atom associated with the nucleobases are listed in table 1. It is observed that the polarizability is proportional to the molecular weight and volume (see table 1). The increasing order of polarizability for the nucleobases is $U > C > T > A > G$ correlated with the increasing molecular volume but it is slightly different with the molecular weight.

In table 2, we summarized the dipole moments (without an applied electric field) of the nucleobases (A, T, G, C, and U). In general, the calculated dipole moments of nucleobases based on the PM3 method agreed well to those obtained by Tasi *et al.* [25] as well as those reported by BdiKin *et al.* [26]. The values of dipole moments from

Table 1. Physicochemical properties of DNA and RNA nucleobases.

Nucleobases	Chemical formula	Number of atom	Heavy atom count	Molecular weight [g/mol]	Van der Waals surface volume [\AA^3]	Polarizability [\AA^3]
Adenine	$\text{C}_5\text{H}_5\text{N}_5$	15	10	143.19	110.42	13.71
Thymine	$\text{C}_5\text{H}_6\text{N}_2\text{O}_2$	15	9	132.16	105.10	11.86
Guanine	$\text{C}_5\text{H}_5\text{N}_5\text{O}$	16	11	159.19	118.77	14.22
Cytosine	$\text{C}_4\text{H}_5\text{N}_3\text{O}$	13	8	117.15	93.32	10.87
Uracil	$\text{C}_4\text{H}_4\text{N}_2\text{O}_2$	12	8	118.14	88.53	10.02

Table 2. Calculated dipole moments of DNA and RNA nucleobases.

Nucleobases	Dipole moment [Debye]		
	Present work	Tasi <i>et al.</i> [25]	Bdikin <i>et al.</i> [26]
Adenine	2.492	2.49	2.495
Thymine	3.989	3.88	4.51
Guanine	5.443	5.45	5.45
Cytosine	6.069	5.68	6.074
Uracil	3.990	3.90	–

the highest to lowest is $\text{C} > \text{G} > \text{U} > \text{T} > \text{A}$. Generally, this is due primarily to the difference in the electronegativity of the bonded atoms and other factors such as the distance between the charge separations. In other words, the greater the electronegativity difference, the greater the dipole moments. The TPSA and the TPSA/TSA ratio of the nucleobases are given in table 3. Among the 5 nucleobases, guanine has the largest TPSA and TSA, as well as the TPSA/TSA ratio. Both thymine and uracil possess the smallest TPSA, and the uracil has the smallest TSA of 131.02\AA^2 . The value of the TPSA/TSA ratio in the order from high to low is $\text{G} > \text{C} > \text{A} > \text{U} > \text{T}$.

3.2 Effect of an applied electric field on electrical, structural and physical properties

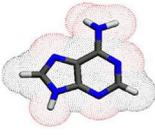
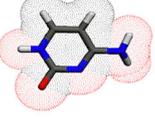
We first investigated and discussed the effects of an applied external electric field E on the polarization behaviours of the cytosine. The dependence of polarization P and total energy U on an applied electric field E of cytosine nucleobase are presented in fig. 2(a) and (b), respectively. The electric field is applied along the direction in the Y -axis of the cytosine molecule, which is illustrated as in fig. 2(a). The applied field is varied between -6 and $+6 \text{ V/nm}$. In general, the shape of the polarization P versus electric field E hysteresis loop is

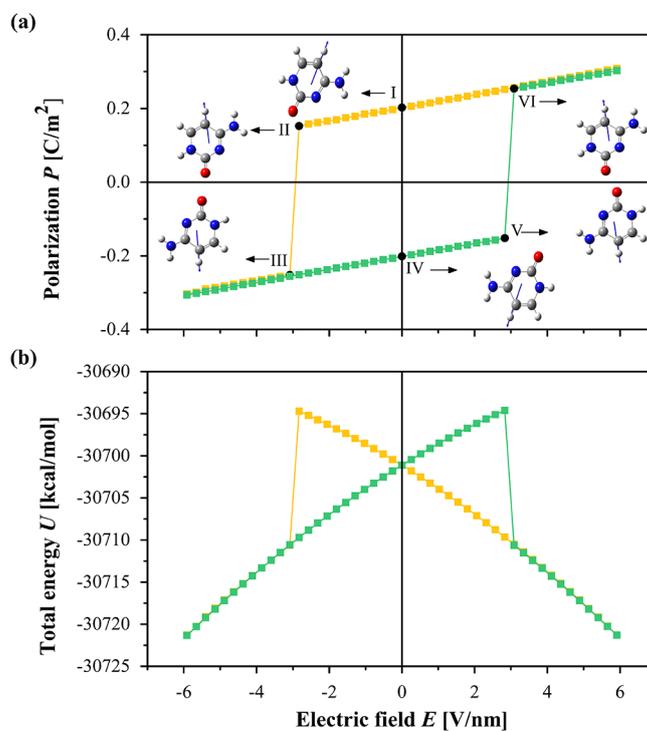
square and symmetry (see fig. 2(a)). If the applied electric field $E = 0$, the polarization P is $\sim \pm 0.200 \text{ C/m}^2$. Switching occurs at the coercive field $E_C = \pm 0.006 \text{ a.u.}$ which is equivalent to 3.085 V/nm ($1 \text{ a.u.} \sim 514 \text{ V/nm}$). Going from a low to a high negative electric field, the point I ($E = 0$) corresponds to a positive polarization state with $P = 0.202 \text{ C/m}^2$ and the corresponding total energy $U = -30701.113 \text{ kcal/mol}$. As the applied electric field increases (at point II), $E < E_{C-}$, the magnitude of the polarization P decreases while the total energy U increases. Upon further increasing $E > E_{C-}$ (at point III), the molecule switches from a positive to a negative polarization state. The reversal of polarization is accompanied by an abrupt decrease of U at $E \sim -3.085 = E_{C-}$. Similarly, the point IV ($E = 0$) corresponds to a negative polarization state with $P = -0.201 \text{ C/m}^2$ and $U = -30701.117 \text{ kcal/mol}$. The cytosine molecule reverses from a negative to a positive state (from V to VI) at $E \sim 3.085 = E_{C+}$.

Insets in fig. 2(a) show the molecular structure of cytosine under a particular electric field. It is seen that the cytosine molecule rotates from a positive (point II) to a negative (point III) state and from a negative (point V) to a positive (point VI) state at $E = E_C$. Hysteresis is not only found in the electric field dependence of polarization but also in the dependence of the total energy U on an applied electric field with a butterfly-like loop behaviour as illustrated in fig. 2(b). From the butterfly-like U - E loop, two small jumps can be found at $E = E_C$.

In fig. 3, a variation of the atomic charge distributions and geometrical parameters under the influence of the applied electric field for the cytosine nucleobase is shown. Note here that only atoms that are sensitive to an applied electric field are shown for the clarity of the discussion. In the following discussion, the label of the atom for the nucleobase is based on fig. 1 unless otherwise specified. The net atomic charges of N_1 , H_3 , H_5 , C_6 , N_7 , H_8 , H_9 , C_{11} , and H_{13} are positive. However, the atomic charge for C_2 , C_4 , N_{10} , and O_{12} are found to be negative. Among the atoms, the electron-rich oxygen O_{12} atom exhibits the highest negative charge whereas the carbon C_{11} atom carries the highest positive charge. These values are in good agreements with those reported by Santamaria *et al.* [27].

Table 3. Calculated topological polar surface area (TPSA), total surface area (TSA), and TPSA/TSA ratio for DNA and RNA nucleobases. Red and black colors represent the polar and non-polar surface areas, respectively.

Nucleobases	TPSA [\AA^2]	TSA [\AA^2]	TPSA/TSA $\times 100\%$	Visualization of TPSA
Adenine	80.50	165.64	48.60	
Thymine	58.20	161.75	35.98	
Guanine	96.20	174.74	55.05	
Cytosine	67.50	131.02	51.52	
Uracil	58.20	130.23	44.69	

**Fig. 2.** Electric field E -dependence of (a) polarization P and (b) total energy U of cytosine. Insets in (a) show the molecular structure of cytosine at certain electric fields. The initial state of molecule and the direction of an applied field is along the Y -axis, as shown in fig. 1.

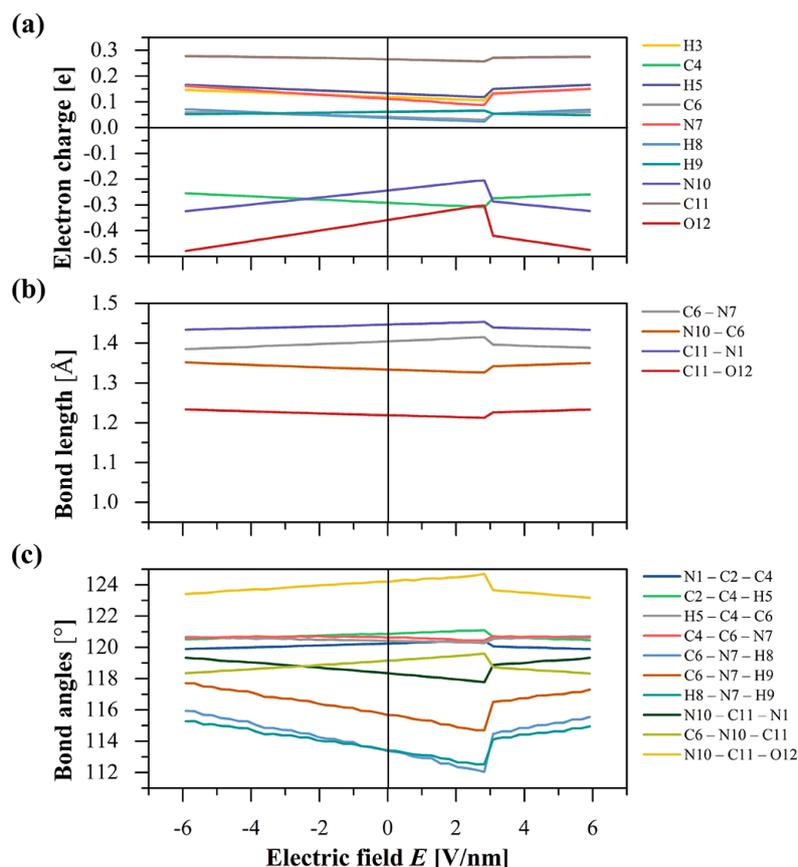


Fig. 3. Dependence of (a) electron charges, (b) bond lengths, and (c) bond angles of cytosine on an applied electric field.

As observed from the TPSA visualization, atoms N_1 , N_{10} , O_{12} , H_{13} and $-NH_2$ amidogen group (N_7 , H_8 , and H_9) can be attributed as the TPSA polar atom. Therefore, these atoms are particularly sensitive to the applied electric field. As shown in fig. 3(a), O_{12} is very sensitive to the electric field where an obvious change in the electron charge can be found at $E \sim E_C$. Similar trends can be seen in the electric-field dependence of bond lengths and angles in cytosine. Figure 3(b) shows the bond length of the cytosine as a function of the electric field. It is clearly seen that the bond length of the partial-polar TPSA C_6-N_7 , $N_{10}-C_6$, $C_{11}-N_1$, and $C_{11}-O_{12}$ bond length are sensitive to the applied electric field. The variation of bond length and angles are particularly obvious near $E \sim E_C$. On the other hand, the bond length or bond angle of the non-polar TPSA atoms is almost independent of an applied electric field (see fig. S1(b) in the supplementary material (SM)). Figure 3(c) depicts the predicted bond angles variation of cytosine under an applied electric field. The bond angles for non-polar atoms $H_5-C_4-C_6$, $H_3-C_2-C_4$, and $C_2-C_4-C_6$ are fairly insensitive to the applied electric field, as compared with $-NH_2$ amidogen group $H_8-N_7-H_9$. A complete picture for the dependence of atomic charge distributions and geometrical parameters on an applied electric field for the cytosine can be found in fig. S1 in the (SM). Figures S2–S5 in the SM show the results for adenine, thymine, guanine and uracil, respectively.

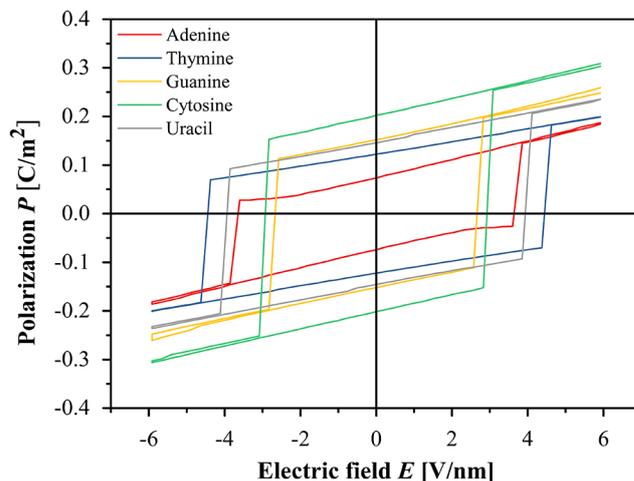


Fig. 4. Electric field E -dependence of polarization P for DNA and RNA nucleobases.

Figure 4 shows the P - E hysteresis loops for five nucleobases (A, T, G, C, and U). The corresponding total energy U - E loops for the adenine, thymine, guanine, and uracil nucleobases are shown in fig. 5. All the nucleobases are found to have butterfly-like loops, as expected.

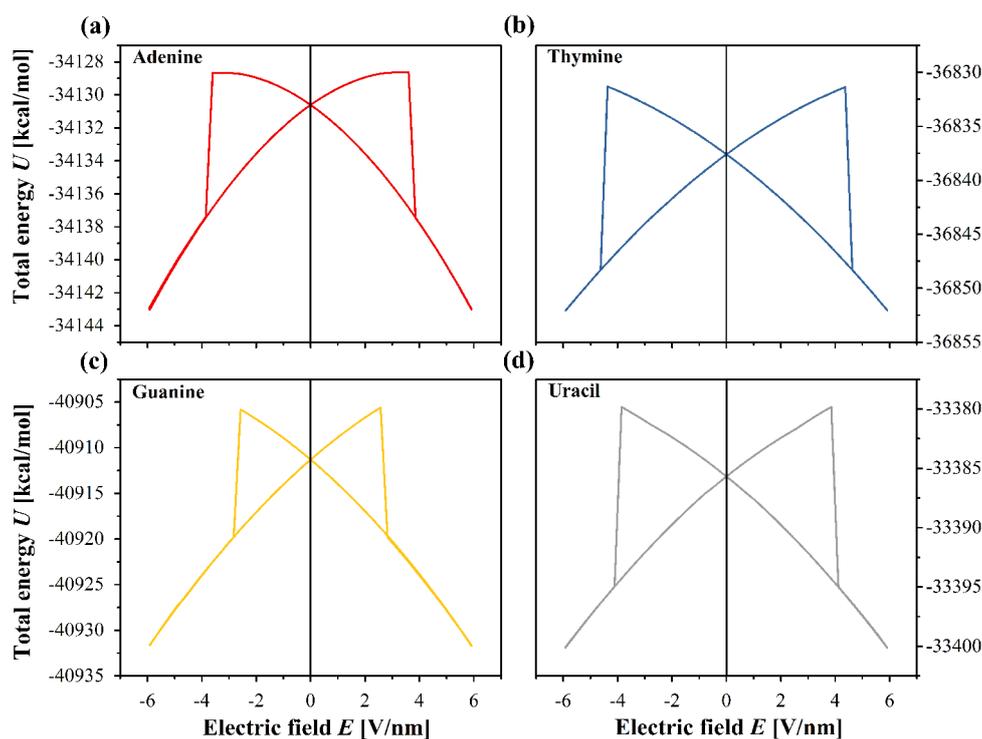


Fig. 5. Electric field E -dependence of total energy U for (a) adenine, (b) thymine, (c) guanine, and (d) uracil.

Table 4. Magnitude coercive field and zero-field polarization of DNA and RNA nucleobases. The values are obtained from the hysteresis loops of figs. 2(a) and 4.

Nucleobases	Magnitude coercive field [V/nm]	Zero-field polarization [C/m ²]
Adenine	3.857	0.073
Thymine	4.628	0.122
Guanine	2.828	0.152
Cytosine	3.085	0.202
Uracil	4.114	0.146

The coercive field and zero-field polarization of the nucleobases obtained from the hysteresis loop are summarized in table 4. Our calculated molecular polarization (at $E = 0$) and coercive field of the nucleobases are all in the same order of magnitude as reported by Bystrov *et al.* [28]. In general, thymine exhibits a broader hysteresis loop, and therefore, a higher coercive field. Among them, guanine and cytosine have large dipole moments. The magnitude of an average coercive field from high to low is $T > U > A > C > G$, whereas the average value of zero-field polarization from high to low is $C > G > U > T > A$. In order to further understand the hysteresis loop behaviour of the nucleobases, we analyse the relationship between the TPSA, polarization, and coercive field. TPSA is a descriptor defined as the sum of the surfaces of the polar atoms in a molecule, primarily oxygen, nitrogen and their attached hydrogen atoms bound to

these electronegative atoms [29]. An interesting result implies the inverse correlation between the TPSA/TSA ratio and the coercive field of the nucleobases. In particular, a nucleobase with a higher number of TPSA/TSA ratios has a smaller coercive field. For example, guanine has a higher value of TPSA/TSA ratio compared to other nucleobases and resulting in a narrow coercive field. On the contrary, thymine and uracil with a lower number of TPSA/TSA ratios require a higher electric field to switch the molecules. The inverse relationship between the TPSA/TSA ratio and the coercive field can be interpreted as follows. Amidogen $-NH_2$ group and oxygen on the carbonyl are polar atoms in the nucleobases of DNA and RNA. A higher value of TPSA/TSA ratio indicates that the total surface area of a nucleobase that is sensitive to an applied electric field is larger. Therefore, the minimum electric field required to switch the nucleobases is smaller. The analysis of other parameters such as molecular weight, charge, and size does not have an obvious relationship with the polarization and coercive field.

4 Conclusion

In this work, we have studied the polarization switching and hysteresis loop behaviour of the DNA and RNA nucleobases based on the PM3 method. Our studies revealed that all nucleobases exhibit zero-field polarization and square-shape hysteresis loops. The total energy as a function of electric field also showed a butterfly-like hysteresis loop feature. The results imply the possible existence of bioferroelectricity in these biomaterials.

The minimum field required for the polarization switching of a nucleobase is inversely proportional to the ratio of TPSA to TSA, though the zero-field polarization does not have an obvious correlation with the TPSA/TSA ratio. The observed switching mechanism and ferroelectrics properties of DNA and RNA nucleobases in this work may provide insight into the development of DNA- and RNA-based nanomaterials, electronic devices, and other possible potential applications.

This research was supported by the University of Malaya Research Grant (No. RP037C-17AFR).

Author contribution statement

KHC, VSL and SMZ conceived the research idea. SCY performed the simulations and analysis. SCY wrote and revised the manuscript. KHC improved the manuscript. All the authors have read and approved the final manuscript.

References

- G.M. Church, Y. Gao, S. Kosuri, *Science* **337**, 1628 (2012).
- Y. Erlich, D. Zielinski, *Science* **355**, 950 (2017).
- A. Extnance, *Nature* **537**, 22 (2016).
- R.K. Moyzis, J.M. Buckingham, L.S. Cram, M. Dani, L.L. Deaven, M.D. Jones, J. Meyne, R.L. Ratliff, J.R. Wu, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6622 (1988).
- Y. Liu, H.-L. Cai, M. Zelisko, Y. Wang, J. Sun, F. Yan, F. Ma, P. Wang, Q.N. Chen, H. Zheng, X. Meng, P. Sharma, Y. Zhang, J. Li, *Proc. Natl. Acad. Sci. U.S.A.* **111**, E2780 (2014).
- Y. Liu, Y. Wang, M.-J. Chow, N.Q. Chen, F. Ma, Y. Zhang, J. Li, *Phys. Rev. Lett.* **110**, 168101 (2013).
- T. Lenz, R. Hummel, I. Katsouras, W.A. Groen, M. Nijemeisland, R. Ruemmler, M.K.E. Schäfer, D.M.d. Leeuw, *Appl. Phys. Lett.* **111**, 133701 (2017).
- M.T. Hwang, P.B. Landon, J. Lee, D. Choi, A.H. Mo, G. Glinsky, R. Lal, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7088 (2016).
- Z. Altintas, I.E. Tothill, *Sens. Actuators B: Chem.* **169**, 188 (2012).
- X.C. Zhou, L.Q. Huang, S.F. Li, *Biosens. Bioelectron.* **16**, 85 (2001).
- J. Wang, *Anal. Chim. Acta.* **469**, 63 (2002).
- J. Polonsky, P. Douzou, C. Sadron, *C. R. Hebd. Seances Acad. Sci.* **250**, 3414 (1960).
- A.L. Stanford, R.A. Lorey, *Nature* **219**, 1250 (1968).
- Y. Ando, E. Fukada, *J. Polym. Sci.: Polym. Phys. Ed.* **14**, 63 (1976).
- E. Fukada, Y. Ando, *J. Polym. Sci. Part A-2: Polym. Phys.* **10**, 565 (1972).
- J. Duchesne, J. Depireux, A. Bertinchamps, N. Cornet, J.M. Van Der Kaa, *Nature* **188**, 405 (1960).
- V.K. Yarmarkin, S.G. Shul'man, V.V. Lemanov, *Phys. Solid State* **51**, 1881 (2009).
- HyperChem, *Tools for Molecular Modeling* (Hypercube, Inc., 2002).
- D. Hadjipavlou-Litina, *Curr. Med. Chem.* **7**, 375 (2000).
- P. Ertl, B. Rohde, P. Selzer, *J. Med. Chem.* **43**, 3714 (2000).
- <http://www.molinspiration.com>.
- C.S. Tsai, in *An Introduction to Computational Biochemistry* (John Wiley & Sons, Inc., 2003) p. 315.
- K.M. Khoda, Y. Liu, C. Storey, *J. Optimization Theory Appl.* **75**, 345 (1992).
- D. Klostermeier, M.G. Rudolph, *Biophysical Chemistry* (CRC Press, Taylor & Francis Group, 2017).
- G. Tasi, I. Palinko, L. Nyerges, P. Fejes, H. Foerster, *J. Chem. Inf. Comput. Sci.* **33**, 296 (1993).
- I. Bdikin, A. Heredia, S.M. Neumayer, V.S. Bystrov, J. Gracio, B.J. Rodriguez, A.L. Kholkin, *J. Appl. Phys.* **118**, 072007 (2015).
- R. Santamaria, A. Vázquez, *J. Comput. Chem.* **15**, 981 (1994).
- V.S. Bystrov, E. Seyedhosseini, I. Bdikin, S. Kopyl, S.M. Neumayer, J. Coutinho, A.L. Kholkin, *Ferroelectrics* **475**, 107 (2015).
- A. Pranitha, P. Lakshmi, *Iran. J. Pharm. Sci.* **10**, 47 (2014).