

## Nucleotide insertion initiated by van der Waals interaction during polymerase beta DNA replication

ANDREW DAS ARULSAM Y

Condensed Matter Group, Division of Interdisciplinary Science, F-02-08 Ketumbar Hill, Jalan Ketumbar, 56100 Kuala-Lumpur, Malaysia  
e-mail: sadwerdna@gmail.com

MS received 18 January 2013; accepted 8 March 2013

**Abstract.** We present here an unambiguous theoretical analyses and to show that the exclusive biochemical reaction involved in a single nucleotide insertion into the DNA primer can be efficiently tracked using the renormalized van der Waals (vdW) interaction of a stronger type, the Hermansson blue-shifting hydrogen bond effect, and the Arunan composite hydrogen-vdW bond. We find that there are two biochemical steps involved to complete the insertion of a single base (cytosine) into the 3' end of a DNA primer. First, the O3' (from a DNA primer) initiates the nucleophilic attack on P $_{\alpha}$  (from an incoming dCTP), in response, O3 $_{\alpha}$  (bonded to P $_{\alpha}$ ) interacts with H' (bonded to O3'). These interactions are shown to be strongly interdependent and require the forming and breaking of P—O and H—O covalent bonds, which in turn imply that we do not need any external energy supply.

**Keywords.** Interacting biomolecules; peptide-bond formation; generalized van der Waals interaction; ionization energy theory.

### 1. Introduction

The reaction pathways involving polymerase(pol)- $\beta$  in eukaryotic cells, which have high fidelity and specificity are presently one of the most subtle scientific information needed to elucidate what constitute the origin of pol- $\beta$  specificity and fidelity.<sup>1,2</sup> We chose the eukaryotic enzyme pol- $\beta$  here for two obvious reasons—it is one of the simplest and smallest (39 kDa) known DNA polymerase enzymes, and it has been well-studied both experimentally and *via* simulations since the 1970s.<sup>3</sup> Radhakrishnan and Schlick,<sup>1</sup> and Alberts, Wang and Schlick<sup>2</sup> have carried out the constrained quantum/molecular mechanical calculations to obtain the relevant pol- $\beta$  DNA replication pathways with the lowest energies. Here, 'constrained' means certain variables are computationally made not-to-float, following the crystal structure data reported by Batra *et al.*,<sup>4</sup> and from their earlier studies.<sup>2</sup> The constrained variables include the position and the numbers of water molecules and ions (Mg<sup>2+</sup>, Na<sup>+</sup> and Cl<sup>-</sup>), the charge states and the conformational changes of the pol- $\beta$  residues (D190 (aspartate 190) and others) and the incoming dCTP (deoxyribocytosine 5'-triphosphate), and so on.<sup>2</sup> The crucial result<sup>1,2</sup> is that the initial deprotonation of the terminal DNA primer  $\delta^-$ -O3'H group to a water molecule occur concurrently with the nucleo-

philic attack of  $\delta^-$ -O3' on P $_{\alpha}^{\delta+}$  catalysed by Mg<sup>2+</sup> ions. Note that we have used  $\delta+$  and  $\delta-$  instead of 5+ and 2- for the phosphorous and oxygen ions, respectively, in order to accommodate the changes of these charges as a result of their changing polarizabilities in the presence of changing interaction strengths. In fact, we will implement this notation throughout the paper for any element in any molecule without caution. Any number that carries a prime (3' or 5') refers to a specific carbon atom in a given molecule.

All the computational analyses thus far focused on the 'pre- and post-chemistry avenues', which also correspond to the close- and open-conformations of pol- $\beta$  and DNA complex.<sup>1,2,5,6</sup> In these reports,<sup>1,2,5,6</sup> several responsible reaction-pathways were proposed entirely based on these avenues, disregarding the biochemical reaction route (with the highest specificity and fidelity) in the chemical process itself, sandwiched between those two avenues (for more information, please refer the section 3.3, Additional notes, given prior to conclusion). Tracking this route (within the chemical process) is far more efficient as there can only be a limited number of reaction routes to initiate the nucleophilic attack that could insert a single base to the growing DNA strand. Therefore, it is technically easier, accurate and reliable if we could exploit the information obtained from this chemical process to narrow down

the processes in the pre- and post-chemistry avenues. Here we make use of the ionization energy theory<sup>7</sup> and the energy-level spacing renormalization group method<sup>8</sup> to evaluate the van der Waals (vdW) interaction of a stronger type, down to the electronic level. We will show that such an evaluation will lead us to find the exclusive biochemical reaction (assisted by (i) the Hermansson blue-shifting hydrogen bond<sup>9–11</sup> and (ii) the Arunan composite hydrogen-vdW bond<sup>12–15</sup> effects) with the highest specificity and fidelity, in the presence of the correct Watson–Crick hydrogen bonds.<sup>16</sup>

## 2. Theoretical method

Here, we first introduce the ionization energy theory (IET) and its approximation, which are needed to renormalize the vdW interaction, and to determine the average atomic energy-level spacing. This is followed by the introduction to the renormalized vdW interaction, which will be shown to be responsible for the (bio)chemical reactions if this vdW is of a stronger type, much stronger than the standard vdW interaction that has a  $1/R^6$  dependence, where  $R$  is the separation between two nuclei.

### 2.1 Ionization energy approximation

Our analyses here use the ionization energy approximation, which can be understood in a straightforward manner if we start from the IET-Schrödinger equation,<sup>7</sup>

$$\begin{aligned} i\hbar(\partial\Psi(\mathbf{r}, t)/\partial t) &= [ -(\hbar^2/2m)\nabla^2 + V_{\text{IET}} ]\Psi(\mathbf{r}, t) \\ &= H_{\text{IET}}\Psi(\mathbf{r}, t) \\ &= (E_0 \pm \xi)\Psi(\mathbf{r}, t), \end{aligned} \quad (1)$$

where  $\Psi(\mathbf{r}, t)$  is the time-dependent many-body wave function,  $\hbar = h/2\pi$ ,  $h$  denotes the Planck constant and  $m$  is the mass of electron. The most important parameter here is the eigenvalue,  $E_0 \pm \xi$ , which is the real (true and unique) energy levels for a given quantum system (an atom or a molecule or a solid or anything in between). Equation (1) reduces to the standard Schrödinger equation by noting that  $E_0 \pm \xi = E$ , which is exact by definition, and it is not trivial.<sup>7</sup> Here,  $E_0$  is the energy levels at zero temperature and in the absence of any external disturbances. This means that  $E_0$  is a constant, while  $\pm\xi$  is known as the ionization energy (or the energy-level spacing) where  $\pm$  refers to electrons and holes, respectively. In this work,  $-\xi$  does not play any role for an obvious reason, but holes (positive

charge carriers) can play a vital role in biomolecules as proven experimentally by Fujitsuka and Majima<sup>17</sup> with respect to charge transfer in DNA. There are two formal proofs for the ionization energy approximation. The first one is indirect and reads,  $H_{\text{IET}}\Psi(\mathbf{r}, t) = (E_0 \pm \xi_{\text{matter}}^{\text{quantum}})\Psi(\mathbf{r}, t) \propto (E_0 \pm \xi_{\text{atom}}^{\text{constituent}})\Psi(\mathbf{r}, t)$ , available in ref. 7, while the second formal proof is direct (and is based on logic) and it is associated to the excitation probability of electrons and holes within the ionization energy based Fermi–Dirac statistics.<sup>18,19</sup>

The ionization energy approximation reads

$$\xi_{\text{matter}}^{\text{quantum}} \propto \xi_{\text{atom}}^{\text{constituent}}, \quad (2)$$

where  $\xi_{\text{matter}}^{\text{quantum}}$  is the real energy level spacings of a particular quantum system. In the most general sense,  $\xi$  is defined to be the real energy spacings between energy levels such that an electron from one energy level needs to overcome an energy cost when that electron tries to occupy another energy level. In molecules,  $\xi$  refers to the energy level spacing between an occupied level (in the highest occupied molecular orbital (HOMO)) and an empty level (in the lowest unoccupied molecular orbital (LUMO)), which is the molecular energy level spacing. Due to the above definition for  $\xi$ , the quantum matter discussed in this work refers to molecules. Now, we can readily predict the changes to  $\xi_{\text{matter}}^{\text{quantum}}$  if we know  $\xi_{\text{atom}}^{\text{constituent}}$ , where the values for  $\xi_{\text{atom}}^{\text{constituent}}$  can be obtained from the experimental atomic spectra available in refs. 20 and 21. For example,  $\xi_{\text{H}_2\text{O}} \propto [2\xi_{\text{H}^+} + \xi_{\text{O}^+}]$ . The ionization energy approximation becomes exact if the quantum matter are atoms or ions because  $\xi_{\text{matter}}^{\text{quantum}} = \xi_{\text{ion}}^{\text{atom}}$ .

In addition, we can also prove why an atomic hydrogen takes up the role of a cation in a water molecule, while an oxygen atom ends up as an anion. To prove this, we need to verify  $\xi_{2\text{H}^+} < \xi_{\text{O}^{2+}}$ , which can be confirmed from refs. 20 and 21 where  $\xi_{2\text{H}^+}$  ( $1312 \text{ kJmol}^{-1}$ )  $< \xi_{\text{O}^{2+}}$  ( $2351 \text{ kJmol}^{-1}$ ). Here, the averaging follows Eq. (3) (given below). This inequality also proves that the formation of  $\text{O}_2\text{H}$  is never possible. Therefore, an atomic hydrogen gives up (donates) its electron to an oxygen atom, and this process is never the other way round (having oxygen as a donor) because there is always a stronger Coulomb attraction between the electron from H and the oxygen nucleus. In other words, if  $\xi$  is larger for atomic A compared to atomic B, then the nucleus of atomic A will always attract the valence electrons from atomic B with a stronger Coulomb force, and not *vice versa*.<sup>7</sup> This reinforces the idea that an atomic hydrogen acts as a cation, whereas, an oxygen atom takes up the role of an anion, in a system consists of H and O. Consequently, one can use this (ionization

energy) approximation to evaluate the changing interaction strengths between two atoms or ions from different molecules. However, the bonds between atoms do not necessarily involve a single valence electron from each atom, and thus, we need to average the ionization energies for atoms that donate (or share) more than one electron. In this case, the ionization energy averaging follows

$$\xi_{\text{atoms}}^{\text{constituent}} = \sum_j \sum_i^z (1/z) \xi_{j,i} (X_j^{i+}), \quad (3)$$

where each subscript  $j$  represents one chemical element ( $X_j$ ) in a particular molecule, while  $i = 1, 2, \dots, z$ , where  $i$  counts the number of valence electrons coming from each chemical element. For example, if  $X_{j=1} = \text{P}$  such that P forms five bonds with its nearest neighbours, then  $\text{P} \rightarrow \text{P}^{5+}$ , and therefore,

$$\xi_{\text{P}^{5+}} = (1/5) \sum_i^5 \xi_{\text{P}^{i+}} = 3414 \text{ kJmol}^{-1}. \quad (4)$$

The first, second,  $\dots$ ,  $z$ th ionization energies (prior to averaging) were obtained from refs. 20 and 21. For a water molecule,  $X_{j=1} = \text{H}$  and  $X_{j=2} = \text{O}$ . It is to be noted here that the ionization energy theory was first formally proven to be a renormalized theory within the energy-level spacing renormalization group method.<sup>7,8</sup> This energy-level spacing renormalization procedure can be related exactly to the Shankar renormalization technique within the screened Coulomb potential derivation, and his renormalization method is documented here.<sup>22–24</sup> Further experimental supports for IET and its applications can be found here.<sup>25–33</sup>

Moreover, the importance of ionic valence states on the electronic transport properties of oxides and magnetic semiconductors was first experimentally proven by Mahendiran *et al.*,<sup>34</sup> the computational (using the Density Functional Theory) proof was given by Mahadevan and Zunger,<sup>35</sup> while the theoretical proofs (using IET) have been developed in ref. 36. In these reports,<sup>34–36</sup> the valence state of any cation is shown to vary depending on its surrounding. Therefore, our averaging procedure explained earlier, properly takes the valence-state effect into account. The ‘surrounding’ here refers to the nearest, and the next nearest neighbours only, such that these neighbours are directly bonded to this cation. For a given biochemical process however, the ‘surrounding’ also needs to consider the neighbours that are not directly bonded to a cation or an anion. This additional interaction originating from the non-bonding neighbours can be taken into account by means of changing interaction strength between two atoms or ions (see below for details).

## 2.2 Renormalized van der Waals interaction for biochemical reactions

Earlier, we have justified why the ionization energy approximation allows us to track the interaction strength between atoms from different molecules. Here, we need to capture that an increasing attractive strength between two atoms can lead to a (bio)chemical reaction. To do that, we need to develop a new vdW formalism that allows a stronger Coulomb attraction between two atoms (or ions) for all  $R \geq r_1 + r_2$ , where  $R$  is the separation between two nuclei,  $r_1$  and  $r_2$  denote the radii of two interacting atoms or ions.<sup>37</sup> This stronger Coulomb attraction has been captured using the renormalized vdW formalism, and the relevant formula is<sup>37</sup>

$$\begin{aligned} \tilde{V}'_{\text{Waals}}(\xi) &= V_{\text{Coulomb}}^{\text{e-ion}} + (1/2)\hbar\omega_0((1/\sqrt{2}) - 1) \\ &\quad \times \exp[(1/2)\lambda\xi], \end{aligned} \quad (5)$$

where,  $\hbar\omega_0$  denotes the ground state energy (of a given molecule) in the absence of external disturbances, and for  $T = 0\text{K}$ , and  $\tilde{V}'_{\text{Waals}}(\xi)$  is the stronger version of the attractive vdW interaction, after renormalization.<sup>37</sup> Moreover,  $\tilde{V}'_{\text{Waals}}(\xi)$  is negative guaranteed by the inequality  $1/\sqrt{2} - 1 < 0$  and the attractive Coulomb interaction between an electron and a nucleus is also negative by definition. Here,  $\hbar$  is the Planck constant divided by  $2\pi$ ,  $\lambda = (12\pi\epsilon_0/e^2)a_{\text{B}}$ , in which,  $a_{\text{B}}$  is the Bohr radius of atomic hydrogen,  $e$  and  $\epsilon_0$  are the electron charge and the permittivity of free space, respectively. It is interesting to note here that Arunan *et al.* were the first to point out the existence of a composite bond that consists of hydrogen bond and the vdW attraction,<sup>12–15</sup> which has been captured by Eq. (5) by means of the generalized vdW interaction.

The detailed derivation for  $\tilde{V}'_{\text{Waals}}(\xi)$ , its physical insights and why it is suitable for both long and short distances are given elsewhere.<sup>37</sup> Equation (5) has been proven to be stronger than the standard vdW prediction due to the existence of an asymmetric polarization between two atoms. The relevant equation that associates  $\xi$  to the atomic displacement polarization is given by<sup>38</sup>

$$\alpha_{\text{d}} = \frac{e^2}{M} \left[ \frac{\exp[\lambda(E_{\text{F}}^0 - \xi)]}{(\omega_{\text{ph}}^2 - \omega^2)} \right], \quad (6)$$

where  $\omega$  is the electric-field dependent frequency due to valence-electron displacement,  $\omega_{\text{ph}}$  denotes the phonon frequency of undeformable ions (mass,  $M$ ) and  $E_{\text{F}}^0$  is the Fermi energy at zero Kelvin. A stronger polarization exists if the two atoms have very different

ionization energy values such that  $\xi_A^{\text{atom}} < \xi_B^{\text{atom}}$  where atomic A will become a cation, while atomic B assumes the role of an anion-based on IET. Presently,<sup>1,2,5,6,39,40</sup> and usually, this asymmetric polarization are assumed to be static (by assigning fixed charges to atoms or ions) and therefore, the interaction between any two ions are electrostatic. For example, the red- and blue-shifted hydrogen bonds were determined entirely from the step-by-step changes to the distance between the supposedly hydrogen-bonded atoms.<sup>39,40</sup> It is physically alright to assume the charges assigned to the atoms of a particular molecule to be static where this assumption can be used to determine the electrostatic dipole–dipole, dipole-induced dipole and/or the London dispersion forces. For example, the above stated interactions can become stronger or weaker entirely from the step-by-step changes to the separation between two strongly interacting atoms ( $R$ ), however, this assumption (by definition) ignores the systematic changes to the atomic polarization when  $R$  changes.

The above mentioned atomic polarization and the vdW forces are often treated as adjustable parameters in the *ab initio* quantum mechanical calculations, and also in molecular dynamical simulations.<sup>41</sup> In particular, in any *ab initio* quantum mechanical calculations, polarization functions are introduced by hand to allow some variations to the atomic orbitals in order to evaluate the molecular integrals. For example, within the Gaussian basis set formalism, each basis function (indicated with a subscript ‘ $r$ ’) represents a normalized linear combination of a few Gaussians, namely, the contracted Gaussian<sup>42</sup>

$$\chi_r = \sum_{ijk} d_{ijk;r} g_{ijk}, \quad (7)$$

where  $d_{ijk;r}$  is the contraction coefficient, the primitive Gaussian based on the Cartesian coordinate (centred on atom  $b$ ),

$$g_{ijk} = N x_b^i y_b^j z_b^k \exp(-\alpha r_b^2), \quad (8)$$

and the Cartesian-Gaussian normalization constant,

$$N = (2\alpha/\pi)^{3/4} [(8\alpha)^{i+j+k} i!j!k!/(2i)!(2j)!(2k)!]^{1/2}. \quad (9)$$

Here, Eqs. (7), (8) and (9) are designed (not through an *ab initio* method) to allow polarizations where  $i$ ,  $j$  and  $k$  are nonnegative integers,  $\alpha > 0$  is an orbital exponent, and  $x_b$ ,  $y_b$ ,  $z_b$  are Cartesian coordinates such that the nucleus  $b$  is the origin. When  $i + j + k = l = 0$ , one has the  $s$ -type Gaussian, while the polarizable  $p$ - and  $d$ -type Gaussians can be obtained for  $l = 1$  and  $l = 2$ , respectively. Equa-

tion (7) represents the guessed linear combination of atomic orbitals (LCAO), which allow distortions such that an atomic orbital’s (AO) shape and its centre of charge can be shifted upon the formation of a molecule, by requiring  $l > 0$ . The point is, the above AO with  $l > 0$  are introduced at the outset by hand, and these AO-functions (see Eq. (7)) are not determined from any *ab initio* quantum mechanical calculation or derivation,<sup>42</sup> and therefore, we can now understand why the parameters associated to polarizations were treated as variationally adjustable parameters.<sup>41</sup>

As a consequence, we need to find an alternative way to understand how and why the change in the polarization strength can influence the blue- or red shifting hydrogen bond (or the generalized vdW interaction), in the presence of another ion. These information can be extracted from the *ab initio* calculations by imposing some guessed AO-functions (Eq. (7)) at the outset, and then treat the parameters related to polarizations as variationally adjustable. In our approach, the polarizations are determined directly without any inbuilt guessed functions or adjustable parameters. On the other hand, if one considers an electrostatic interaction, then an induced polarization can change the charge magnitude that was originally assigned to an ion, making such an assumption inadequate to study chemical reactions. Here, we will use the ionization energy approximation and the renormalized vdW theory to tackle the polarization effect that is responsible to initiate a chemical reaction between two strongly interacting atoms (or ions) from different molecules. In the following analyses, we will show that the changes in the polarization caused by different surrounding molecules give rise to the corresponding changes to the partial charges of the reactant atoms. These changes in the partial charges assigned to reactant atoms or ions (from different molecules) can lead to a lock-in configuration, which can proceed to a chemical reaction, if the configuration has the strongest Coulomb attraction.

### 2.3 Benchmark analyses: Peptide bond formation (condensation reaction)

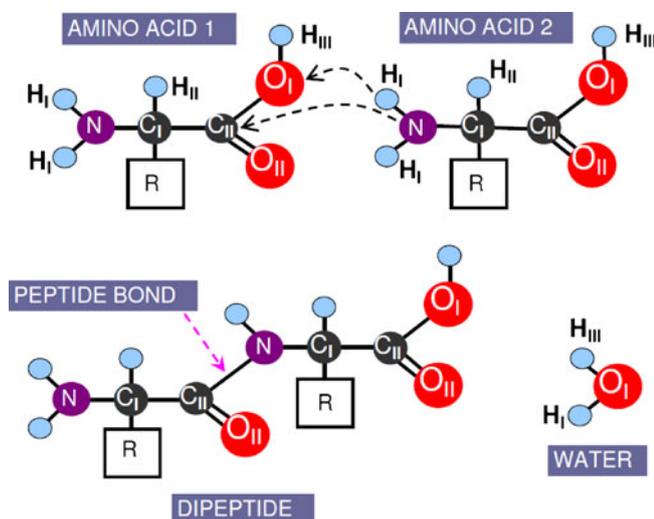
Here, we apply IET and the new vdW formula to evaluate the formation of a peptide bond between two  $\alpha$ -amino acids,



where R denotes an organic substituent, which is irrelevant here. The analyses presented here involve inequalities, but they are sufficient to achieve our objective to

unambiguously determine the correct biochemical reaction pathway, without the need to compute any numbers (apart from the ionization energy averaging).

The condensation reaction (given in Eq. (10)) is schematically depicted in figure 1 where the dashed arrows indicate the correct reaction pathways. The alternative reaction pathways can involve  $C_I$ ,  $O_{II}$ ,  $H_{II}$  and  $H_{III}$  (not shown). For a much simpler system given in ref. 25, in which the system consists of only one water molecule and a surface oxygen atom. Our strategy here is to track this condensation reaction pathway that has the strongest interaction (such that  $|\dot{V}'_{\text{Waals}}(\xi)|$  from Eq. (5) is maximum). This means that, a chemical reaction occurs continuously until completion or equilibrium (provided the external conditions to initiate this reaction are met), guided entirely by the changing instantaneous interaction strengths (mostly due to electron–electron and electron-ion types), not because the reactants are ‘consciously aware’ that their final destination or products (after the chemical reaction) will have a smaller free-energy, or maximum entropy, or



**Figure 1.** Two amino acids (the carboxylic-group from the amino acid 1 and the amine-group from the amino acid 2) react to form a dipeptide and a water molecule. We have identified three different hydrogen atoms,  $H_I$ ,  $H_{II}$  and  $H_{III}$ , two different carbon atoms ( $C_I$  and  $C_{II}$ ) and two different oxygen atoms, namely,  $O_I$  and  $O_{II}$ . The organic substituent,  $R$  is irrelevant in this reaction for obvious reasons. Dashed arrows indicate the correct reaction pathways ( $NH_I \rightarrow CO_I$ ), while the alternative reaction pathways are between these pairs of atoms,  $O_I H_{III}$  (from amino acid 2)  $\rightarrow O_I H_{III}$  (from amino acid 1),  $NH_I$  (from amino acid 2)  $\rightarrow O_I H_{III}$  (from amino acid 1),  $O_I H_{III}$  (from amino acid 2)  $\rightarrow CO_I$  (from amino acid 1) and  $CH_{II}$  (from amino acid 2)  $\rightarrow CO_I$  (from amino acid 1). The above notations imply that the first pair (of atoms) belongs to the acid amino 2, whereas the second pair (on the right-hand side of “ $\rightarrow$ ”) is from the acid amino 1. These alternatives will not be discussed—see text for details.

energetically more stable. There are five possibilities for two  $\alpha$ -amino acids to react (ignoring  $R$ ), but only one of the reaction pathways is responsible for the formation of a peptide bond. We have evaluated all the reaction pathways, but we only highlight the correct reaction pathway (indicated with dashed arrows in figure 1) as a benchmark analysis. Both the correct and the incorrect reaction pathways for dipeptide bond formation are similar to the reaction pathways discussed for pol- $\beta$  DNA replication (given in the following section).

The polarized charges ( $\delta+$  or  $\delta-$ ) for the relevant atoms in an isolated amino acid can be determined using Eqs. (2) and (3), while table 1 lists all the ionization-energy values required for our analyses. From these two inequalities,  $\xi_{N^{2+}}(2129 \text{ kJmol}^{-1}) > \xi_{2H^+}(1312 \text{ kJmol}^{-1})$  and  $\xi_{N^{4+}}(4078 \text{ kJmol}^{-1}) > \xi_{C^{4+}}(3571 \text{ kJmol}^{-1})$ , we can conclude that  $N$  is negatively charged,  $N^{\delta-}$ ,  $H_I$  is positively charged,  $H_I^{\delta+}$ , while  $C_I$  is slightly negatively

**Table 1.** Averaged atomic ionization energies ( $\xi$ ) for neutral and charged elements and their respective averaged valence states ordered with increasing atomic number  $Z$ . All the experimental ionization energy values were obtained from ref. 20. The notation 2nd denotes the 2nd ionization energy (without averaging), and so forth. However, the valence 1+ is also equivalent to the 1st ionization energy, while the valence state larger than 1+ has been averaged using Eq. (3).

Elements	Atomic number $Z$	Valence state / $\xi$ number	$\xi$ ( $\text{kJmol}^{-1}$ )
H	1	1+	1312
Carbon	6	4+	3571
Carbon	6	5+	10423
N	7	1+	1402
N	7	2 <sup>nd</sup>	2856
N	7	3 <sup>rd</sup>	4578
N	7	4 <sup>th</sup>	7475
N	7	2+	2129
N	7	3+	2946
N	7	4+	4078
O	8	1+	1314
O	8	2 <sup>nd</sup>	3388
O	8	3 <sup>rd</sup>	5301
O	8	4 <sup>th</sup>	7469
O	8	5 <sup>th</sup>	10990
O	8	2+	2351
O	8	4+	4368
O	8	5+	5692
P	15	1+	1012
P	15	2 <sup>nd</sup>	1907
P	15	3 <sup>rd</sup>	2914
P	15	4 <sup>th</sup>	4964
P	15	5 <sup>th</sup>	6274
P	15	6 <sup>th</sup>	21267
P	15	5+	3414

charged,  $C_I^{\delta-}$  because even though  $H_{II}$  contributes an electron to  $C_I$ , the negative charge due to this electron (from  $H_{II}$ ) is compensated by the electron (from  $C_I$ ) that is polarized toward N to create a covalent bond with N. This electron from  $C_I$  is actually asymmetrically polarized toward N due to  $\xi_{N^{++}}(4078 \text{ kJmol}^{-1}) > \xi_{C^{++}}(3571 \text{ kJmol}^{-1})$ . Therefore,  ${}^{H_I}\delta+ > {}^{H_{II}}\delta+$  and  ${}^{N}\delta- > {}^{C_I}\delta-$ . In addition, a C—C bond gives rise to an isotropic polarization (because  $\xi_C = \xi_C$ ), and thus, C does not gain any charge ( $\delta+$  or  $\delta-$ ) by making bonds with another C. However,  $C_{II}^{\delta+}$  is positively charged because  $\xi_{C^{++}}(3571 \text{ kJmol}^{-1}) < \xi_{O^{++}}(4368 \text{ kJmol}^{-1})$ , and no charge contribution from  $C_I$  to  $C_{II}$  due to isotropic polarization. Therefore, we have these inequalities,  ${}^{C_{II}}\delta+ > {}^{H_I}\delta+ > {}^{H_{II}}\delta+$  and  ${}^{O_I}\delta- > {}^{O_{II}}\delta-$ . The second inequality is due to  $\xi_{4H^+}(1312 \text{ kJmol}^{-1}) < \xi_{C^{++}}(3571 \text{ kJmol}^{-1})$ . Recall here that atoms with a lower averaged ionization energy easily gives away its electron, compared to an atom with a larger averaged ionization energy.

We can now proceed to evaluate the correct (experimentally proven) reaction pathway,  $NH_I$ (from amino acid 2) interacts with  $O_I H_{III}$ (from amino acid 1).  $H_I^{\delta+}$  first interacts with  $O_I^{\delta-}$ , which further polarizes the  $O_I^{\delta-}$  electrons (contributed by  $C_{II}^{\delta+}$  and  $H_{III}^{\delta+}$ ) much closer to  $O_I^{\delta-}$ . Moreover, the electron contributed by  $H_I^{\delta+}$  to  $N^{\delta-}$  is now more polarized closer to  $N^{\delta-}$  because  $H_I^{\delta+}$  now attracts an electron from  $O_I^{\delta-}$  originally contributed by  $C_{II}^{\delta+}$ . These induced polarizations imply that  ${}^{H_I \cdots O_I - C_{II}}\delta'+ > {}^{O_I - C_{II}}\delta+$  and  ${}^{O_I \cdots H_I - N}\delta''- > {}^{H_I - N}\delta-$ , where  $H_I \cdots O_I$  denotes the vdW attraction within the generalized vdW formalism. Within the hydrogen-bond formalism, it is known as the blue-shifting hydrogen bond effect between  $H_I$  and  $O_I$ . Here, the superscript,  $C_{II}-O_I$  represents the covalent bond between  $C_{II}$  and  $O_I$ . From these inequalities,  ${}^{H_I \cdots O_I - C_{II}}\delta'+ > {}^{O_I - C_{II}}\delta+$  and  ${}^{O_I \cdots H_I - N}\delta''- > {}^{H_I - N}\delta-$ , we can now deduce that  $N^{\delta-}$  interacts more strongly with  $C_{II}^{\delta+}$  thanks to the interaction between  $H_I^{\delta+}$  and  $O_I^{\delta-}$ . Next we show that the above charge inequalities (due to induced polarizations) will lead us to a lock-in configuration, which sets the stage for chemical reactions.

The lock-in configuration is obtained when  $H_I^{\delta+}$  interacts with  $O_I^{\delta-}$  and when  $N^{\delta-}$  interacts with  $C_{II}^{\delta+}$  where the strength of these interactions follow Eq. (5). We can now use the ionization energy approximation to calculate the relevant values, namely,  $\xi_{N \cdots C_{II}} \propto |\xi_{N^+} - \xi_{C^+}| = 317 \text{ kJmol}^{-1}$  and  $\xi_{H_I \cdots O_I} \propto |\xi_{H^+} - \xi_{O^+}| = 2 \text{ kJmol}^{-1}$ , and therefore, the  $\xi_{\text{total}}$  is  $317 + 2 = 319 \text{ kJmol}^{-1}$ , which is the highest compared to any other alternative combinations of atoms shown in figure 1. This highest value also gives the highest interaction strength when it is substituted into Eq. (5).

But this is only one-half of the requirement for the condensation reaction to proceed. The other half comes from the stronger bonds between  $C_{II}^{\delta+}$  and  $O_I^{\delta-}$  and between  $N^{\delta-}$  and  $H_I^{\delta+}$ . The bond between  $C_{II}^{\delta+}$  and  $O_I^{\delta-}$  is stronger because the electron from  $C_{II}^{\delta+}$  is polarized closer to  $O_I^{\delta-}$  in the presence of  $H_I^{\delta+}$ . Meanwhile, the  $N^{\delta-}-H_I^{\delta+}$  bond has also become stronger because the electron from  $H_I^{\delta+}$  is now polarized much closer to  $N^{\delta-}$  in the presence of  $O_I^{\delta-}$ . In other words, these two bonds have become stronger because they are in certain polarized modes, which are actually best suited to initiate selective chemical reactions compared to unpolarized bonds. For example, it is relatively easy to break a polarized bond via an electron and/or a proton transfer. On the contrary, unpolarized bonding electrons are more homogeneously distributed in a given molecule, and therefore, the charges for atoms within the molecule ( $\delta+$  and  $\delta-$ ) are either too small or negligible to initiate a proper and strong lock-in configuration. In the polarized modes however, atoms (in a molecule) have large polarized charges ( $\delta+$  or  $\delta-$ ), which can strongly interact with other atoms or ions from another molecule.

Note here that thermal energy is not useful to set up a lock-in configuration because thermal energy causes non-selective electron polarizations where it will excite all the bonding electrons homogeneously in a given molecule. Hence, thermal energy can be a useful source of energy required for continuous reactions and/or to keep the activation energy lower so that a given reaction (in certain polarized modes) can proceed.

We have identified these two bonds,  $C_{II}^{\delta+}-O_I^{\delta-}$  and  $N^{\delta-}-H_I^{\delta+}$  are stronger as a result of selective atomic polarizations. In other words, these bonds are more difficult to be broken by applying thermal energy, but they are more easily broken if we could lure these charged atoms with opposite charges (with an electron or a proton). In particular, the above stronger bonds will lead  $C_{II}^{\delta+}$  to strongly interact with  $N^{\delta-}$ , while  $O_I^{\delta-}$  is strongly interacting with  $H_I^{\delta+}$ . All we need now are some external energy supply (thermal energy for example) to initiate the simultaneous electron and proton transfers, namely an electron-transfer from  $N^{\delta-}$  to  $C_{II}^{\delta+}$  and a proton-transfer,  $H_I^{\delta+}$  to  $O_I^{\delta-}$ . These transfers are likely to proceed compared to other combinations of atoms, which have been predicted from the ionization energy approximation presented earlier. This lock-in configuration has the strongest attraction, which requires the minimal external energy supply (due to the lowest activation energy). The simultaneous electron and proton transfers will complete the condensation reaction, giving rise to a dipeptide and a water molecule as reaction products.

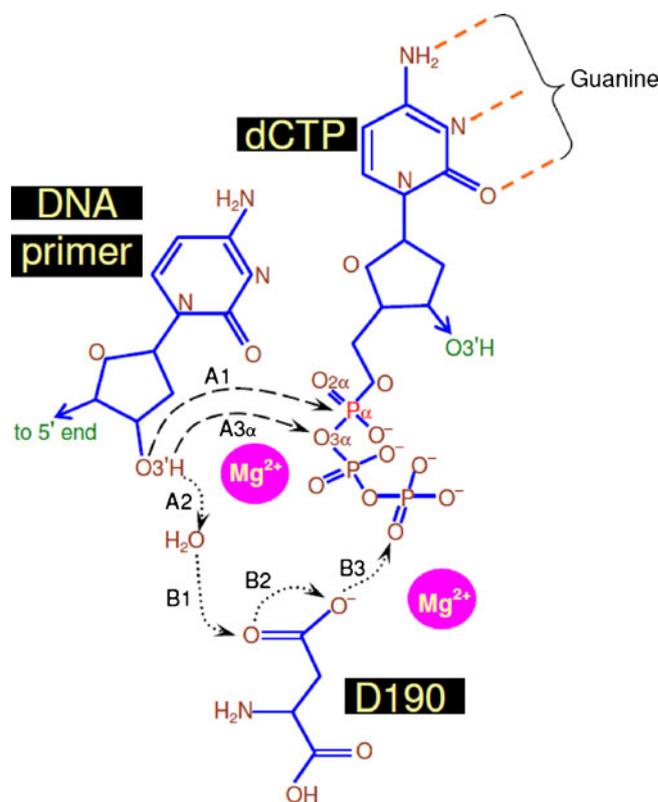
In summary, the charge ( $\delta+$  and  $\delta-$ ) redistribution (due to polarization) in atoms have given rise to a lock-in configuration, and subsequently have led to simultaneous electron and proton transfers in such a way that a new peptide bond is formed between  $N^{\delta'-}$  and  $C_{II}^{\delta'+}$ , while the  $N^{\delta'-}-H_1^{\delta'+}$  covalent bond is broken. On the other hand, a new  $O_1^{\delta-}-H_1^{\delta+}$  bond is formed, while the  $C_{II}^{\delta'+}-O_1^{\delta-}$  bond is broken. We can now move on and use the acquired knowledge to tackle the pol- $\beta$  DNA replication.

### 3. Results and discussion

Schlick *et al.*<sup>1,2</sup> found the nucleophilic attack, A1 runs simultaneously with A2 to B3 proton transfers. In contrast, we show here with unequivocal theoretical analyses that a direct deprotonation of  $H^+$  from  $O3'H$  to  $O3_\alpha$  ( $A3_\alpha$  pathway) and the A1 pathway are activated simultaneously. This concurrent process (with steps A1 and  $A3_\alpha$ ) is chemically driven (initiated by a stronger vdW interaction) with the highest specificity and fidelity (see figure 2). This direct deprotonation is also somewhat similar to Abashkin, Erickson and Burt,<sup>6</sup> in which, the difference being their<sup>6</sup> direct proton transfer is from  $O3'H$  to  $O2_\alpha$ , and then to  $O3_\alpha$  (after bond-rotation). We use the blue-shifting vdW interaction between electrons from different molecules,<sup>37</sup> or from different sections of a bigger molecule, as a determining factor for biochemical reactions, which completes one base-pair (CG) to the growing DNA strand. The letters, A, G, C and T denote the well-known DNA purine (adenine and guanine) and pyrimidine (cytosine and thymine) bases. To avoid any confusion, we will always write the name in full to denote the element carbon.

Of course, the steps, A1 and  $A3_\alpha$  can only be efficiently initiated in the presence of the correct (a) conformation of the incoming dCTP by interacting with the pol- $\beta$  residues and with the guanine (from the template DNA strand) via the Watson-Crick hydrogen bonds, (b) position of  $Mg^{2+}$  ions and water molecules, and (c) position of pol- $\beta$  residues and their interactions with cations, anions and water molecules. In fact, these points (a to c) are valid because numerous interactions (both quantum and classical types) will come into play in a biologically dynamical setting, which have been discussed explicitly by Spyraakis *et al.*<sup>43</sup> and Jayaram *et al.*<sup>44</sup> Apart from these, the readers can also referred to an excellent review by Joyce<sup>45</sup> for a list of the techniques available to study biochemical reactions with respect to DNA polymerases.

The fundamental information on the generalized vdW interaction leading to the insertion of a single base



**Figure 2.** There are two processes involved for a single base (C) insertion into the 3' end of the DNA primer, depicted in this schematic figure. The first process has been proposed in refs. 1 and 2 that consists of five steps, A1 to B3. Whereas, the second process proven here involves A1 and  $A3_\alpha$  occurring simultaneously (follow the dashed arrows). The first two major steps (discovered by Radhakrishnan and Schlick<sup>1</sup>) carry the labels, A1 and A2, while the other steps (from B1 to B3) are part of the subsequent proton( $H^+$ )-transfer steps assisted by D190, proposed by Alberts, Wang and Schlick<sup>2</sup> (follow the dotted arrows). The diagonal dashed lines on the top right corner denote the possible Watson-Crick hydrogen bonds with a guanine (from a template strand), while the solid diagonal arrow pointing to  $O3'H$  in an incoming nucleotide (dCTP) will be the new 3' end of the primer (growing) strand, after insertion. Another diagonal arrow pointing to the left (in the DNA primer) extends in the 5' end direction. All the sticks (including the isolated sticks inside the ring) and the diagonal solid arrows represent the covalent  $\sigma$  bonds. The conformations and the positions of the DNA primer, D190, dCTP and  $Mg^{2+}$  are not to scale. The number 190 refers to the position of an aspartate residue attached to a pol- $\beta$  enzyme.

(C) to the growing DNA strand are equally (if not more) important (i) to theoretically prove the possibility of replacing phosphorous with arsenic in the Californian bacterium, strain GFAJ-1, which was experimentally observed by Wolfe-Simon and her assemblage of scholars,<sup>46,47</sup> (ii) to evaluate the *in vitro* characterization of a DNA nanomachine that maps the changes in the spatial and temporal pH, as reported by Krishnan *et al.*,<sup>48</sup>

(iii) to understand how a female sex hormone (progesterone) chemically activates sperms into ‘hyperactivity’ by increasing the  $\text{Ca}^{2+}$  ions concentration (inside sperms),<sup>49,50</sup> (iv) to theoretically evaluate the possibility of exploiting G-quadruplexes (a four-stranded guanine structure) as a therapeutic target in oncology to develop anticancer drugs,<sup>51,52</sup> and (v) to study the initiation of human colon tumour that grows like a cancer due to pol- $\beta$  DNA replication error.<sup>53</sup>

### 3.1 Biochemical reactions in the chemical stage

In this work, we track the electron–electron and electron-ion interaction strengths via the changing energy-level spacing (presented earlier in the theoretical method section) during a particular biochemical reaction. Interestingly, the band structure and energy levels of proteins have also been suggested to be useful to understand proteins electronic interaction.<sup>54</sup> Here, we will track the above interaction strengths so as to understand why and how the nucleophilic attack proceeds with a unique path. In fact, there is only one approach that is well-defined and allows one to expose the chain of events leading to this exclusive path occurring at the critical point, which was first introduced<sup>25</sup> to capture the origin of a single water-molecule splitting on a MgO surface. The formal proofs that justify this approach are given elsewhere.<sup>55,56</sup> Added to that, these interaction strengths that we will be tracking are directly related to the ionization energy theory (IET) and the renormalization group method, which are also mathematically well-defined.<sup>7,8</sup>

Furthermore, our analyses can be made quantitative by invoking the ionization energy approximation, which gives us the license to claim that the molecular energy level spacings are directly proportional to the atomic energy level spacings of all the atom exist in a given molecule.<sup>19,57</sup> The atomic ionization energies of all the atoms in a molecule is of course different from the molecular ionization energies of that particular molecule. They are different because the atomic energy levels will have to transform (due to wave function transformation) during a chemical reaction to form the molecular energy levels.<sup>55,56</sup> However, the energy level spacings of a molecule containing certain atoms has been proven to be proportional (not equal) to the averaged atomic energy level spacings of all the isolated atoms present in that particular molecule.<sup>7</sup>

What we need to do now is to first understand the chain of events giving rise to the validity of the steps, A1 and A2 within IET. When the negatively charged  $\delta^- \text{O}3'$  comes closer to  $\text{P}_\alpha$  (see figure 2), the  $\delta^- \text{O}3' \text{—} \text{H}^{\delta+}$

bond is further reinforced, which is the effect we discovered in ref. 25. The reason is that the electrons contributed by H and carbon to create the  $\text{O}3' \text{—} \text{H}$  and  $\text{O}3' \text{—} \text{carbon}$  covalent bonds are now being strongly bound to  $\text{O}3'$  in the presence of  $\text{P}_\alpha^{\delta+}$ , which pulls these electrons further away from carbon and H (from now on, we denote this specific hydrogen as  $\text{H}'$ ). The  $\delta^- \text{O}3' \text{—} \delta^+ \text{H}'$  covalent bond is necessarily stronger due to induced electrostatic interaction between  $\delta^- \text{O}3'$  and  $\text{P}_\alpha$ . The reason is that the bonding electron (contributed by  $\text{H}'$ ) is now more polarized toward  $\delta^- \text{O}3'$  due to  $\text{P}_\alpha$ , and this will necessarily result in a stronger  $\delta^- \text{O}3' \text{—} \delta^+ \text{H}'$  covalent bond. Logical proof: the above covalent bond has to get stronger, or weaker, and definitely, its strength cannot remain the same in the presence of  $\text{P}_\alpha^{\delta+}$ . Here, all other conditions from the pre-chemistry avenue are relatively weak. They are weak because in the chemical stage, these external effects coming from the pre-chemistry avenue do not directly involve in the chemical reaction. Of course, the molecules, ions and the residues from the pre-chemistry avenue do set the stage for the chemical reaction to proceed, but they do not directly involve in the nucleophilic attack as reactants. Now, let us assume the said bond ( $\delta^- \text{O}3' \text{—} \delta^+ \text{H}'$ ) gets weaker in the presence of  $\text{P}_\alpha^{\delta+}$ , and this is only possible if the bonding electron from  $\text{H}'$  is weakly polarized toward  $\text{O}3'$ . Therefore, one can immediately deduce that this leads to a contradiction because the bonding electron contributed by  $\text{H}'$ , has to be more polarized toward  $\text{O}3'$  due to the attractive induced interaction between  $\text{P}_\alpha^{\delta+}$  and  $\delta^- \text{O}3'$ .

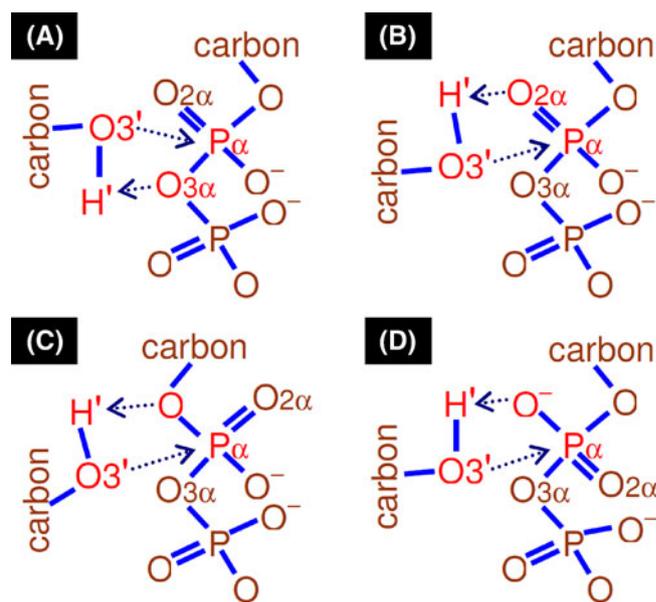
Consequently, it is energetically inconvenient to break this stronger  $\text{O}3' \text{—} \text{H}'$  covalent bond by brute force (by thermal energy), such as the electronic excitation of the ( $\text{O}3' \text{—} \text{H}$ ) bonding electron, which is now strongly polarized toward  $\text{P}_\alpha^{\delta+}$ . Therefore, it would be easier by allowing the oxygen ( $\text{O}_{\text{WAT}}^{\delta-}$ ) from a nearby water molecule to form the required hydrogen bond (between  $\text{O}_{\text{WAT}}$  and  $\text{H}'$ ) that can easily blue shift (*via* the Hermansson blue-shifting hydrogen bond effect) due to an attractive interaction between  $\text{O}3'$  and  $\text{P}_\alpha$ . Let us rephrase the last sentence due to its extreme importance and also because of its far-reaching effect in biological systems. The stronger covalent bond between  $\text{O}3'$  and  $\text{H}'$  is due to the attractive interaction between  $\text{O}3'$  and  $\text{P}_\alpha$ , and is actually more easily breakable, not by brute force, but when it is subjected to an induced electrostatic interaction provided by the  $\text{O}_{\text{WAT}}^{\delta-}$  from a water molecule. This means that, the original covalent bond between  $\text{O}3'$  and  $\text{H}'$  is not easily breakable by a water molecule in the absence of the attractive interaction between  $\text{O}3'$  and  $\text{P}_\alpha$  (note this point, or recall the theoretical method section).

Now, back to Hermansson, we know that her blue-shifting interaction can initiate the Radhakrishnan–Schlick deprotonation (step A2). Unfortunately, due to this deprotonation (losing  $H'$  to a water molecule), one now requires some external energy from an unknown source (this could be some additional heat or interaction) to overcome the activation energy and initiate the nucleophilic attack (A1). In other words, by having  $H'$  bonded to  $O3'$ , one has the strongest interaction between the bonding electron (coming from  $O3'—H'$ ) and the nucleus of  $P_\alpha^{\delta+}$  leading to smaller activation energy. To see why and how this activation energy is smaller, we just need to wait a bit longer because we need to introduce and build the arguments properly. Anyway, the other subsequent proton transfers will not be captured with IET because these latter paths, B1 to B3 are similar to the hydrogen-bonding mechanism stated above, and they are not unique. They are not unique because even though these steps, A1, A2 and B1 to B3 require minimal total energy compared to other pathways (that involve different residues and so on), one can always manipulate any number of pathways to further reduce the minimal total energy. For example, we can have a direct deprotonation from the  $H_2O - -H'$  to one of the oxygen in the outgoing pyrophosphate without involving D190, by manipulating the ‘constrained’ parameters. Here, the dashed lines, ‘- -’ means a hydrogen bond, ‘—’ denotes a covalent bond, and ‘=’ is a double covalent bond used in text.

In contrast, we prove here the existence of a unique reaction pathway that consists of two simultaneous steps, (A1)  $\delta^-O3'$  attacks  $P_\alpha^{\delta+}$  while (A3 $_\alpha$ )  $\delta^-O3_\alpha$  counterattacks  $\delta^+H'$  (from  $O3'$ ) that has the highest specificity and fidelity. We will also explain why and how these steps are strategically interdependent that make them exclusive. But before we dig deep, let us first properly introduce the ionization energy approximation relevant to this work. The origin of these charges,  $\delta^-$  and  $\delta^+$  are due to their respective polarization, which can be understood from the average atomic ionization energy ( $\xi$ ) values.<sup>26</sup> Here, large ionization energy value implies large energy needed to polarize or to excite the electron from its initial or ground state. For example, one can obtain the average ionization energy values from table 1<sup>20</sup> and write  $\xi_{2H^+}(1312 \text{ kJmol}^{-1}) < \xi_{O^{2+}}(2351 \text{ kJmol}^{-1})$ , and therefore, H and O are indeed positively and negatively charged, as they should be. This inequality proves that oxygen can never be a cation in water molecules. Using the same arguments, we can obtain the inequality,  $\xi_P < \xi_O$ , which is also true for all individual (1<sup>st</sup>, . . . , 5<sup>th</sup>) ionization energies (see table 1) to justify the type of charges ( $\delta^+$  or  $\delta^-$ ) carried by these elements,  $P_\alpha^{\delta+}$  and  $O^{\delta-}$  in the 5'-triphosphate group.

Up to this point, it is true that we have no constrained parameters to handle because we deliberately ignored them, not because ignorance is bliss, but to keep all the analyses exact and unambiguous within the chemical process, sandwiched between the pre- and post-chemistry avenues. The complex processes and their numerous possibilities involved in these avenues necessitate the introduction of constrained parameters. To avoid such parameters, the chemical process is isolated, which is theoretically valid because the interactions during this closing and opening stages, do not directly involve in the biochemical reactions that insert a base to the growing DNA strand. Thus, in the following paragraphs, our strategy is to first implement this isolation correctly, and then discuss the possibility of carrying out the pre- and post-chemistry investigations subject to the exclusive biochemical reaction pathway.

Apart from the well-known nucleophilic attack (A1 pathway), there are four other possible chemical reactions that correspond to four oxygen atoms bonded to  $P_\alpha$ , namely,  $P_\alpha—O3_\alpha—P$ ,  $P_\alpha=O2_\alpha$ ,  $P_\alpha—O—$ carbon, and  $P_\alpha—O^-$  (see figures 3A to D), in which only one of them should occur simultaneously (with A1). Apparently, we need to eliminate any three of them by



**Figure 3.** The above schematic diagrams are sketched to depict the nucleophilic-attack region, selected and magnified from figure 2 for different types of counterattack (counter to the nucleophilic attack) between oxygen (bonded to  $P_\alpha$ ) and  $H'$ . The dotted arrows indicate the induced electrostatic interactions between  $\delta^+$  and  $\delta^-$  charges for different oxygen atoms bonded to  $P_\alpha^{\delta+}$ , for instance, (A)  $\delta^-O3_\alpha$  (also bonded to another P), (B)  $\delta^-O2_\alpha$ , (C)  $O^{\delta-}$  (also bonded to carbon) and (D)  $O^-$ . Note here that the nucleophilic attack (between  $\delta^-O3'$  and  $P_\alpha^{\delta+}$ ) remains the same. The molecular conformations and their positions are not to scale.

considering each in turn, and then prove  $A3_\alpha$  has the highest specificity and fidelity. We have to inform the readers that the following analyses are entirely analytic and are inevitably subtle, plagued by different magnitudes of polarization for the element,  $P_\alpha$ , and the four oxygen atoms bonded to  $P_\alpha$  in the presence of  $O3'$ ,  $H'$ ,  $P$  (valid only for the oxygen in  $P_\alpha—O3_\alpha—P$ ) and carbon (valid only for the oxygen in  $P_\alpha—O—carbon$ ).

From the ionization energy approximation (given earlier), we can understand the validity of the assigned charges in  $P_\alpha^{\delta+}—O^{\delta-}—carbon^{\delta+}$ , and consequently the steps,  $A1$  and  $A3_\alpha$  are also correct from the induced Coulomb interaction between  $\delta+$  and  $\delta-$ . Of course, this induced interaction has a deep microscopic origin, which will activate the required chemical reactions to extend the DNA strand. But it is sufficient to consider only the partial charges now (without going into the details), to rule out the three out of the four steps listed above, and to find the crucial oxygen atom (bonded to  $P_\alpha$ ) that interact with the strongest induced interaction with  $H'$ . Once the correct oxygen is identified, we can then go on and exploit the existence of the renormalized vdW interaction of a stronger type (Eq. (5)) that is responsible for the simultaneous nucleophilic attack  $A1$ , and the counterattack  $A3_\alpha$ . The nucleophilic- and the counter-attack depicted in figures 3A to D are due to an induced attractive electrostatic interaction between  $\delta-O3'$  and  $P_\alpha^{\delta+}$ . This induced interaction is a prerequisite to form a new covalent bond ( $\delta-O3'—P_\alpha^{\delta+}$ ), which can be obtained from the stronger vdW interaction (see Eq. (5)). Using Eq. (5), we can go on to untangle each chemical step (sandwiched between the pre- and post-chemistry avenues) involved during the base,  $C$  insertion (from dCTP), and find the crucial oxygen (bonded to  $P_\alpha$ ) that strongly interact with  $H'$  during the nucleophilic attack.

### 3.2 $A1$ and $A3_\alpha$ steps

The moment  $\delta-O3'$  unleashes its polarizable electron on  $P_\alpha^{\delta+}$ ,  $\delta-O3_\alpha$  strikes back instantaneously, attacking  $\delta^+H'$ . As a matter of fact, these nucleophilic- and counter-attacks are simultaneous to the extent that one can also view it in the reverse attacking order. But this is incorrect. The above original order-of-attack (nucleophilic-attack invites the counterattack) can be understood by noting  $P_\alpha\delta+ > H'\delta+$  and  $O3'\delta- > O3_\alpha\delta-$ , which imply a stronger attractive interaction between  $\delta-O3'$  and  $P_\alpha^{\delta+}$ . The first inequality is obvious with four oxygen atoms (forming five covalent bonds) with  $P_\alpha$ . Whereas, the latter inequality is due to  $\frac{1}{2}[\xi(P_\alpha^{5+}) + \xi(P^{5+})] = 3414 \text{ kJmol}^{-1}$ ,  $\frac{1}{2}[\xi(carbon^{4+}) + \xi(H^+)] =$

$2442 \text{ kJmol}^{-1}$ , and therefore,  $\frac{1}{2}[\xi(P_\alpha^{5+}) + \xi(P^{5+})] > \frac{1}{2}[\xi(carbon^{4+}) + \xi(H^+)]$ . Here, the averaging follows Eq. (3). In other words, the electrons contributed by carbon and H are of lower ionization energies, and therefore more strongly polarized toward  $\delta-O3'$ . While the electrons responsible for the partial charge in  $\delta-O3_\alpha$  are coming from the less polarizable P (see figure 3A). Here, P is the least polarizable element (with five bonds) compared to carbon (with four bonds) and H (with fixed one bond). All bonds are predominantly covalent-type, unless stated otherwise.

The chemical steps depicted in figure 3A are also strategically interdependent. (i) The nucleophilic attack will first reinforce the  $O3'—H'$  bond, however, this stronger bond will be easily breakable only if  $H'$  is subjected to  $\delta-O3_\alpha$ . (ii) Conversely, the counterattack will only make the  $P_\alpha—O3_\alpha$  bond stronger, but again, this bond is also easily breakable if  $P_\alpha^{\delta+}$  is allowed to interact with  $\delta-O3'$ . As a consequence, we can observe that the two steps, (i) and (ii) nicely correspond to  $A1$  and  $A3_\alpha$ , and they are precisely interdependent. For example, both (i) and (ii) when proceed simultaneously, cause these bonds,  $P_\alpha—O3_\alpha$  and  $O3'—H'$  to break, which add C to the growing DNA strand due to the formation of  $P_\alpha—O3'$  bond. On the other hand, a  $PP_i$  molecule leaves the scene after the formation and breaking of  $O3_\alpha—H'$  and  $P_\alpha—O3_\alpha$  bonds, respectively. Recall here that we follow the path that likely to have the strongest electron–electron and electron–nucleus interactions in accordance with ref.<sup>37</sup> These strongest interactions will require minimal activation energy for a particular reaction to proceed (hence, require smaller external energy supply).

It is clear that the process consisting of  $A1$  and  $A3_\alpha$  steps is a type of lock-in process due to strong interdependence that can be made exclusive with the highest specificity and fidelity, if only we could systematically rule out the other oxygen atoms bonded to  $P_\alpha$ . This interdependence is actually initiated by a stronger vdW attraction between  $O3'$  and  $P_\alpha$ , and between  $O3_\alpha$  and  $H'$  with further assistance from the Hermansson blue-shifting hydrogen bond effect. The formation of these bonds,  $P_\alpha—O3'$  and  $H'—O3_\alpha$  are assisted by the Hermansson blue-shifting hydrogen bond effect, at least for the bond,  $H'—O3_\alpha$ . Whereas, the formation of  $P_\alpha—O3'$  bond is due to the Hermansson-like blue-shifting hydrogen bond effect. The readers are referred to ref. 25 for some straightforward details on the applications of the Hermansson-like bonds. Consequently, these steps do not require large external energy supply to break the above-stated bonds.

The process shown in figure 3B consists of the Abashkin–Erickson–Burt deprotonation steps.<sup>6</sup> In fact,

the  $O2_\alpha$  is the only close contender against the  $A3_\alpha$  step explained above. The other oxygen atoms can be ruled out quite easily. Recall that the nucleophilic attack (step (i)) introduced earlier remains the same here, and will remain so for all other oxygen atoms (see figures 3C and D). The instantaneous counterattack here is between  $O2_\alpha$  and  $H'$ , which will strengthen the  $O2_\alpha=P_\alpha$  double bond that should be easily breakable only if  $P_\alpha^{\delta+}$  is subjected to  $\delta^-O3'$ . There is no difference in the partial charge between  $\delta^-O2_\alpha$  and  $\delta^-O3_\alpha$  ( $^{O3_\alpha}\delta^- = ^{O2_\alpha}\delta^-$ ) because both electrons in each case are contributed by P. However,  $O2_\alpha=P_\alpha$  is a double bond, hence the single  $P_\alpha-O3_\alpha$  bond is in fact twice more easily breakable than the stated double bond. Unlike the breaking of  $P_\alpha-O3_\alpha$  bond in the  $A3_\alpha$  step, the interaction energy from the polarized  $O3'$  (during the nucleophilic attack given in figure 3B) is distributed equally to both bonds, until one of the double bond is broken. In contrast, for the  $A3_\alpha$  step, the same interaction energy is fully available to break the  $P_\alpha-O3_\alpha$  bond alone, reinforced by the counterattack between  $O3_\alpha$  and  $H'$ . Thus, the process indicated in figure 3B is inferior to A1 and  $A3_\alpha$  steps in terms of the highest specificity and fidelity.

In figure 3C, we require the O bonded to carbon and  $P_\alpha$  to counterattack  $H'$ . In this case, the partial charge in the  $\delta^-O$  atom are contributed by carbon and P, whereas two P contribute two of their electrons to  $\delta^-O3_\alpha$ , implying  $^{O\delta-} > ^{O3_\alpha}\delta-$  (recall the averaging presented earlier with carbon contributing a more polarizable electron, compared to P). Due to this inequality, the O-carbon bond is the one that tend to blue-shift more than the required O- $P_\alpha$  bond. Therefore, the O- $P_\alpha$  bond cannot take advantage of the counterattack between  $\delta^-O$  and  $\delta^+H'$  when  $P_\alpha^{\delta+}$  is subjected to  $\delta^-O3'$  (because the O-carbon bond will be fortified, instead of the O- $P_\alpha$  bond). As such, this third process can also be ruled out as it clearly lacks the required specificity and fidelity.

The logic required to exclude the final oxygen (see figure 3D) from the list is the easiest of them all. Here,  $O^-$  counterattacks  $H'$  without any disturbance to the  $P_\alpha^{\delta+}-O^-$  bond. Consequently, we can immediately see that the nucleophilic attack cannot proceed until the formation of the  $P_\alpha-O3'$  bond simply because this would require an additional core electron from  $P_\alpha$  with very large  $\xi$  (see table 1) to form the 6th bond between  $P_\alpha^{\delta+}$  and  $\delta^-O3'$ . This is definitely unlikely, and so is the process suggested in figure 3D.

In summary, it is an unequivocal claim that the steps, A1 and  $A3_\alpha$  have the highest specificity and fidelity within the biochemical reaction stage sandwiched between the pre- and post-chemistry avenues, which has been pointed out much earlier. Having said that, the next stage of research can be directed to fine-tune the

constrained parameters to include the conformational changes to the dCTP in such a way that the steps, A1 and  $A3_\alpha$  can be allowed to occur. Such steps may need some three-dimensional conformational changes to the dCTP, but without disturbing the Watson-Crick hydrogen bonds, where these hydrogen bonds will remain intact as one can confirm it from figure 2.

### 3.3 Additional notes

We have ignored the effects coming from the pre-chemistry avenue, for example, due to  $Mg^{2+}$  ions, pol- $\beta$  residues (D190 and probably others), and water molecules. They do contribute directly in the pre- and post-chemistry avenues, but they do not act as reactants in the chemical stage. Hence, in the chemical stage, we can indeed ignore the contribution made by these molecules and ions. In fact, we can logically rule out the effects coming from the pre-chemistry avenue—any ions or charged atoms (attached to nearby molecules or residues), which are not reactants, will affect all the nearby reactant atoms (the correct as well as the incorrect ones) equally *via* the electrostatic interaction. This means that, non-reactant molecules do not discriminate (not selective) when they interact *via* the Coulomb interaction (be it repulsive or attractive) with the reactant atoms (directly or indirectly). Interestingly, the effects coming from the pre-chemistry avenue have also been ruled out<sup>39</sup> using some molecular dynamical simulations. The results reported<sup>1,2</sup> emphasize the importance of pre-chemistry effects, but ignore the contribution of different oxygen atoms bonded to  $P_\alpha$  in the chemical stage because there is no counter-attack in their calculations. This means that, their calculations in the chemical stage do not consider the interactions between  $H'$  and other oxygen atoms bonded to  $P_\alpha$  because  $H'$  is lost to a water molecule.

There are two ways to include the effect from the environment—(i) the ions and molecules from the environment affect the relevant polarization of the atoms from the reactant molecules (such that there is no chemical reaction between them) before and/or during and/or after the DNA replication, or (ii) the ions or molecules from the environment chemically react with the reactant molecules before and/or during and/or after the DNA replication. Obviously, our work ignores point (ii), but satisfies point (i). For example, point (i) can be shown to reinforce the reaction pathway (that consists of two steps, A1 and  $A3_\alpha$ ) proposed here. In particular, the cations ( $Mg^{2+}$ ) enhance the magnitudes of the electron polarization of all anions in the vicinity (namely, the oxygen from  $O3'H$  and the oxygen

from  $O3_\alpha$ , see figure 2), and these non-discriminative polarization only enhance the lock-in configuration discussed above. Here,  $Mg^{2+}$  and water molecules do not selectively polarize the oxygen from  $O3'H$  and ignores the oxygen from  $O3_\alpha$ . This also means that, we did not constrain the position of  $Mg^{2+}$  ions in such a way to obtain stronger interaction strength. Point (ii) has been considered in refs. 1, 2 and 39. Of course, our reaction pathway becomes invalid if there is any chemical reaction between a reactant molecule and a molecule from the environment.

Activation energy is the minimum energy needed to initiate a chemical reaction, in which the activation energy exists due to small interaction strength between two reactant species. This means that, we need to supply external energy to overcome the activation energy or microscopically, we need to increase the interaction strength between two reactant species to initiate the reaction. Here, the reaction energy is the energy needed to complete a reaction (after overcoming the activation energy). The reaction pathway (see the steps, A1 and A3 $_\alpha$  in figure 2) responsible for the DNA replication requires breaking and forming of covalent bonds. To initiate the breaking and forming of these bonds, one needs to increase the interaction strength between H (from  $O3'H$ ) and  $O3_\alpha$ , and also between  $O3'$  and  $P_\alpha$ , which have been achieved by the strategic lock-in configuration between the atoms stated above. Now, the reaction energy is the total energy needed to complete the reaction, in which the reaction will last for as long as the interaction between the reactant species are strong enough to initiate the chemical reaction. This is why the reaction energy and the activation energy do not correlate, and any correlation between them (activation and reaction energies) is just a 'lucky coincidence'.

In any case, we have given unambiguous analyses as to why and how the A1 and A3 $_\alpha$  chemical steps have the highest specificity and fidelity, and are precisely interdependent (with lock-in steps), which can be made compatible with contributions from  $Mg^{2+}$  ions, pol- $\beta$  residues (D190 and probably others), and water molecules in the pre- and post-chemistry avenues. Our approach guarantees that the lock-in steps are only possible if H' stays bonded to  $O3'$  during the nucleophilic attack.

#### 4. Conclusions

In conclusion, we have actually attempted to tell an unambiguous and a self-consistent tale of how the electronic structure (with distinct energy-level spacings) may influence aging in humans using abstract analyses. These influences are found to be indirect

where the unique ionization energy values determine why and how a selected set of complex biochemical reactions are initiated and executed with negligible errors to maintain healthy eukaryotes. Here, we have considered only one type of enzyme known as the polymerase- $\beta$ , and a specific biochemical process involving the insertion of a single C base (from dCTP) into the growing DNA strand. Nonetheless, the above influence (due to ionization energy) is proven to be significant and far-reaching, such that the stronger renormalized vdW interaction and the Hermansson blue-shifting hydrogen bond effect are applicable to all biochemical processes. For example, these two interactions are responsible to set up the lock-in configuration that has the strongest Coulomb attraction.

We have proven that there is only one unique biochemical process, consisting of two steps with the highest specificity and fidelity. These steps are strongly interdependent to the extent that they can be regarded as lock-in steps, occurring simultaneously. In particular, the nucleophilic attack starts when the  $O3'$  uses its polarizable electron to attract the positively charged  $P_\alpha$  by means of a stronger vdW interaction. This step leads to a stronger  $O3'—H'$  bond, which becomes easily susceptible in the presence of  $O3_\alpha$ . At the same time, the counterattack between  $H'$  and  $O3_\alpha$  fortifies the  $P_\alpha—O3_\alpha$  bond, which can be easily weakened by subjecting  $P_\alpha$  to  $O3'$ , and this is nothing but the nucleophilic attack that is being unleashed as we speak. Obviously, these (nucleophilic- and counter-) attacks complement each other rather nicely in such a way that they lead to the formation of new  $O3'—P_\alpha$  and  $H'—O3_\alpha$  bonds, and the breaking of  $P_\alpha—O3_\alpha$  and  $O3'—H'$  bonds. In addition to this almost perfect matching, note that the bonds that are being formed are of the same types, as the ones that are being broken. This implies a negligible penalty in activation energy, or may not even require any external energy supply at all, to overcome the activation energy.

#### Acknowledgements

The author thanks Madam Sebastiammal Innasimuthu, Mr. Arulsamy Innasimuthu, Mr. Arokia Das Anthony, Madam Amelia Das Anthony, Mr. Malcolm Anandraj and Mr. Kingston Kisshenraj for their financial support. Thanks are also due to Mr. Alexander Jeffrey Hinde (The University of Sydney), Mr. Mir Massoud Aghili Yajadda (Commonwealth Scientific and Industrial Research Organisation, Lindfield), Mr. Ondřej Grulich (Tomas Bata University, Zlin) and Mr. Naresh Kumar Mani (École Normale Supérieure, Paris) for providing some of the listed references.

## References

- Radhakrishnan R and Schlick T 2006 *Biochem. Biophys. Res. Commun.* **350** 521
- Alberts I L, Wang Y and Schlick T 2007 *J. Am. Chem. Soc.* **129** 11100
- Hubscher U, Maga G and Spadari S 2002 *Annu. Rev. Biochem.* **71** 133
- Batra V K, Beard W A, Shock D D, Krahn J M, Pedersen L C and Wilson S H 2006 *Structure* **14** 757
- (a) Yang L, Beard W A, Wilson S H, Roux B, Broyde S and Schlick T 2002 *J. Mol. Biol.* **321** 459; (b) Arora K and Schlick T 2003 *Chem. Phys. Lett.* **378** 1; (c) Yang L, Beard W A, Wilson S H, Broyde S and Schlick T 2004 *Biophys. J.* **86** 3392
- Abashkin Y G, Erickson J W and Burt S K 2001 *J. Phys. Chem. B* **105** 287
- (a) Arulsamy A D 2010 *Pramana J. Phys.* **74** 615; (b) Arulsamy A D 2009 Many-body Hamiltonian based on ionization energy concept: a renormalized theory to study strongly correlated matter and nanostructures. PhD thesis, The University of Sydney, Australia
- Arulsamy A D 2011 *Ann. Phys. (N.Y.)* **326** 541
- Hermansson K 1991 *J. Chem. Phys.* **95** 3578
- Hermansson K 1993 *J. Chem. Phys.* **99** 861
- Hermansson K 2002 *J. Phys. Chem. A* **106** 4695
- Arunan E 1999 *Curr. Sci.* **77** 1233
- Goswami M, Mandal P K, Ramdass D J and Arunan E 2004 *Chem. Phys. Lett.* **393** 22
- Arunan E, Desiraju G R, Klein R A, Sadlej J, Scheiner S, Alkorta I, Clary D C, Crabtree R H, Dannenberg J J, Hobza P, Kjaergaard H G, Legon A C, Mennucci B and Nesbitt D J 2011 *Pure Appl. Chem.* **83** 1619
- Arunan E, Desiraju G R, Klein R A, Sadlej J, Scheiner S, Alkorta I, Clary D C, Crabtree R H, Dannenberg J J, Hobza P, Kjaergaard H G, Legon A C, Mennucci B and Nesbitt D J 2011 *Pure Appl. Chem.* **83** 1637
- Watson J D and Crick F H C 1953 *Nature* **171** 737
- Fujitsuka M and Majima M 2012 *Phys. Chem. Chem. Phys.* **14** 11234
- Arulsamy A D 2001 *Physica C* **356** 62
- Arulsamy A D 2002 *Phys. Lett. A* **300** 691
- Winter M J 2012 ([www.webelements.com](http://www.webelements.com)), University of Sheffield, UK
- Ralchenko Y, Kramida A E and Reader J 2009 NIST ASD Team, NIST Atomic Spectra Database version 3.1.4 (<http://physics.nist.gov/asd3>), National Institute of Standards and Technology, Gaithersburg, MD, USA
- Shankar R 1991 *Physica A* **177** 530
- Shankar R 1994 *Rev. Mod. Phys.* **66** 129
- Shankar R 2011 *Phil. Trans. R. Soc. A* **369** 2612
- Arulsamy A D, Kregar Z, Eleršič K, Modic M and Subramani U S 2011 *Phys. Chem. Chem. Phys.* **13** 15175
- Arulsamy A D, Eleršič K, Modic M, Cvelbar U and Mozetič M 2010 *Chem. Phys. Chem.* **11** 3704
- Shi F and Dong H 2011 *Dalton Trans.* **40** 6659
- Ratchford D, Dziatkowski K, Hartsfield T, Li X, Gao Y and Tang Z 2011 *J. Appl. Phys.* **109** 103509
- Dziatkowski K, Ratchford D, Hartsfield T, Li X, Gao Y and Tang Z 2011 *Acta Phys. Polon. A* **120** 870
- Amira A, Mosbah M F, Leblanc A, Molinie P and Corraze B 2004 *Phys. Status Solidi C* **1** 1944
- Mahtali M, Boudjema E H, Labbani R, Chamekh S, Bouabellou A, Toufik A and Simon C 2010 *Surf. Interf. Anal.* **42** 935
- Eleršič K, Pičman M, Hauptman N, Cvelbar U and Mozetič M 2011 *IEEE Trans. Plasma Sci.* **39** 2812
- Radhakrishnan A N, Rao P P, Mahesh S K, Thampi D S V and Koshy P 2012 *Inorg. Chem.* **51** 2409
- Mahendiran R, Tiwary S K, Raychaudhuri A K, Ramakrishnan T V, Mahesh R, Rangavittal N and Rao C N R 1996 *Phys. Rev. B* **53** 3348
- Mahadevan P and Zunger A 2003 *Phys. Rev. B* **68** 75202
- Arulsamy A D, Cui X Y, Stampfl C, Ratnavelu K 2009 *Phys. Status Solidi B* **246** 1060
- Arulsamy A D 2011 Retrieved from <http://arxiv.org/abs/1110.3412>
- Arulsamy A D and Ostrikov K 2009 *Phys. Lett. A* **373** 2267
- Prasad B R and Warshel A 2011 *Proteins* **79** 2900
- Zhou P P and Qiu W Y 2009 *Chem. Phys. Chem.* **10** 1847
- Gallicchio E and Levy R M 2004 *J. Comput. Chem.* **25** 479
- Levin I N 2000 *Quantum chemistry* (New Jersey, USA: Prentice-Hall)
- Spyrakakis F, Cozzini P, Bertoli C, Marabotti A, Kellogg G E and Mozzarelli A 2007 *BMC Struct. Biol.* doi:10.1186/1472-6807-7-4
- (a) Jayaram B and Beveridge D L 1996 *Annu. Rev. Biophys. Biomol. Struct.* **25** 367; (b) Reddy Ch K, Das A and Jayaram B 2001 *J. Mol. Biol.* **314** 619; (c) Jayaram B and Jain T 2004 *Annu. Rev. Biophys. Biomol. Struct.* **33** 343
- Joyce C M 2010 *Biochim. Biophys. Acta* **1804** 1032
- Wolfe-Simon F, Blum J S, Kulp T R, Gordon G W, Hoeft S E, Pett-Ridge J, Stolz J F, Webb S M, Weber P K, Davies P C W, Anbar A D and Oremland R S 2011 *Science* **332** 1163
- Wolfe-Simon F, Blum J S, Kulp T R, Gordon G W, Hoeft S E, Pett-Ridge J, Stolz J F, Webb S M, Weber P K, Davies P C W, Anbar A D and Oremland R S 2011 *Science* **332** 1149
- (a) Krishnan Y and Simmel F C 2011 *Angew. Chem.* **50** 3124; (b) Modi S, Bhatia D, Simmel F C and Krishnan Y 2010 *J. Phys. Chem. Lett.* **1** 1994; (c) Modi S, Swetha M G, Goswami D, Gupta G D, Mayor S and Krishnan Y 2009 *Nat. Nanotechnol.* **4** 325; (d) Saha S, Bhatia D and Krishnan Y 2010 *Small* **6** 1288
- Strünker T, Goodwin N, Brenker C, Kashikar N D, Weyand I, Seifert R and Kaupp U B 2011 *Nature* **471** 382
- Lishko P V, Botchkina I L and Kirichock Y 2011 *Nature* **471** 387
- Balasubramanian S 2011 *Chem. Commun.* **47** 7281
- Balasubramanian S, Hurley L H and Neidle S 2011 *Nat. Rev.* **10** 261
- Muniappan B P and Thilly W G 2002 *Cancer Res.* **62** 3271
- Evans M G and Gergely J 1949 *Biochim. Biophys. Acta* **3** 188
- Arulsamy A D 2011 Retrieved from <http://arxiv.org/abs/1105.5862>
- Arulsamy A D 2011 *Prog. Theor. Phys.* **126** 577
- Arulsamy A D 2005 *Phys. Lett. A* **334** 413