Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry. This journal is © The Royal Society of Chemistry 2020



Figure S2 Linear correlation plots between ΔG^{\ddagger} and DPE of the H1', BDFE of the C1' - H1'. The symbols \blacksquare , \blacktriangledown , \blacklozenge , and \bullet stand for duplex 5' A T* 3', 5' G T* 3',5' C_a T* 3', 5' T T* 3' and 5' U T* 3', respectively; their relative single-stranded sequences are represented by symbols \blacksquare , \blacktriangledown , \blacklozenge , and \bullet , respectively).



Figure S3 The linear correlation plots between ΔG^{\ddagger} and $\Delta Grxn$ of hydrogen abstraction by the peroxyl radical in the X-CH2-CH2-CH2-OO[•] models. (Calculated by M06-2X and CCSD(T) at 6-31+G(d,p) level, X represents CHCH₂, CCH, Ph, COH, COCH₃, COOCH₃, NO₂, F, COOH, H, NHCH₃, NH₂, NCH₃CH₃, NHCOH, OCH₃, OH, CH₂CH₃, CH₃).



Figure S4 The histogram plots of ΔG^{\ddagger} of hydrogen abstraction by the peroxyl radical in the X-CH₂-CH₂-CH₂-OO[•] models.



(a)





Figure S5 The model molecules used for assessment of the preparation energies. (a) and (b) used for C1'-H1' bond extension; (c) and (d) used for O_a - O_b bond extension.



Figure S6 All the transition structures of the H-abstraction reaction studied here.

Table S1 The activation energies and reaction energies (in kcal mol⁻¹) including zero-point correction energies (ΔE^{\ddagger} and ΔE), enthalpy (ΔH^{\ddagger} and ΔH) and Gibbs energies (ΔG^{\ddagger} and ΔG) of H1' abstraction by the peroxyl radical of duplex 5' X T* 3' and its opposite single-stranded 5' X T* 3'.

		5'AT*3'	5'GT*3'	5'C _a T*3'	5'CT*3'	5'TT*3'	5'UT*3'
	ΔE^{\ddagger}	16.3	17.5	16.0	24.6	15.6	15.1
	ΔE	1.5	1.4	0.9	5.3	1.8	1.2
đa	ΔH^{\ddagger}	16.1	17.0	15.4	24.3	15.0	15.0
ds	ΔH	1.6	1.2	0.8	5.5	1.4	1.4
	ΔG^{\ddagger}	16.8	19.9	17.2	24.9	16.6	15.9
	ΔG	0.9	2.7	1.2	5.3	2.8	1.4
	ΔE^{\ddagger}	19.5	25.5	19.7	19.7	21.9	21.9
	ΔE	6.3	10.1	3.8	3.8	6.5	6.5
~~	ΔH^{\ddagger}	19.1	24.8	19.2	19.2	21.5	21.6
88	ΔH	6.5	9.9	3.7	3.7	6.5	6.5
·	ΔG^{\ddagger}	20.8	27.1	21.3	21.3	22.6	22.7
	ΔG	6.3	11.4	4.4	4.4	7.5	7.1

The basis-set effect through using M06-2X/6-311++G(d,p) method.^a (in kcal mol⁻¹)

	ΔE^{\ddagger}	ΔΕ	ΔH^{\ddagger}	ΔH	ΔG^{\ddagger}	ΔG
ds-5'TT*3'	15.6	1.8	15.0	1.4	16.0	2.8
	(16.5)	(1.2)	(16.5)	(1.2)	(18.0)	(2.6)
ss-5'TT*3'	21.9	6.5	21.5	6.5	22.6	7.1
	(21.5)	(5.6)	(21.5)	(5.6)	(22.6)	(6.6)

^aValues in parentheses are calculated using M06-2X/6-311++G(d,p) method .

Table S2 The entropic effects, thermal effect, thermal correction, energies and imaginary frequencies.

		5'AT*3'	5'GT*3'	5'C _a T*3'	5'CT*3'	5'TT*3'	5'UT*3'
ds	298* ∆ S	-0.7	-2.9	-1.8	-0.6	-1.6	-0.9
SS	$298^* \Delta S$	-1.7	-2.3	-2.1	-2.1	-1.1	-1.1

The effects of the entropic on the reaction activation energies (in kcal mol⁻¹).

The effects of the thermal correction on activation energies and reaction energies (in kcal mol⁻¹) of H1' abstraction by the peroxyl radical of duplex 5' X T* 3' and its opposite single-stranded 5' X T* 3'.^a

		Thermal correction ^b	Thermal correction ^c	Thermal correction ^d
5! \ T*2!	A.E.	-3.9(-3.1)	-4.6(-4.8)	-4.6(-4.8)
5 AT '5	R.E.	-1.4(-0.6)	-0.6(-0.4)	-0.6(-0.4)
5'CT*2'	A.E.	-0.9(-1.6)	-3.8(-3.9)	-3.8(-3.9)
5'G1*3'	R.E.	1.3(1.3)	-0.3(-0.6)	-0.3(-0.6)
5'C T*3'	A.E.	-2.1(-2.4)	-3.8(-4.4)	-3.8(-4.4)
$5C_{a}1.5$	R.E.	0.1(0.3)	-0.3(-0.4)	-0.3(-0.4)
51CT*21	A.E.	-4.1(-2.4)	-4.7(-4.4)	-4.7(-4.4)
301*3	R.E.	-0.7(0.3)	-0.5(-0.4)	-0.5(-0.4)
5'TT*2'	A.E.	-2.3(-3.3)	-3.8(-4.4)	-3.8(-4.4)
511.5	R.E.	1.5(0.6)	0.9(-0.4)	0.9(-0.4)
5'I IT*2'	A.E.	-2.8(-3.3)	-4.2(-4.4)	-4.2(-4.4)
501*5	R.E.	-0.1(0)	-0.4(-0.6)	-0.4(-0.6)

^aValues for the single-stranded 5' X T* 3' in parentheses. ^bThermal correction to Gibbs free energy. ^cThermal correction to Enthalpy. ^dThermal correction to Energy. A.E. = Activation energy. R.E. = Reaction energy.

		5'AT*3'	5'GT*3'	5'C _a T*3'	5'CT*3'	5'TT*3'	5'UT*3'
T1	Popotont	0.933851	0.92751	0.949957	0.925581	0.931209	0.906372
	Reactant	(0.461792)	(0.466523)	(0.449293)	(0.449293)	(0.463874)	(0.437568)
to gibbo	тс	0.927694	0.919281	0.946658	0.918967	0.927588	0.901882
fraa	15	(0.45678)	(0.463961)	(0.445498)	(0.445498)	(0.458548)	(0.432318)
norgy	Droduct	0.931683	0.922756	0.950138	0.924402	0.933525	0.906189
energy	Product	(0.460861)	(0.468352)	(0.449813)	(0.449813)	(0.464781)	(0.437524)
	Reactant	1.114945	1.102834	1.132887	1.10398	1.11439	1.085066
Thermal		(0.570195)	(0.575918)	(0.555756)	(0.555756)	(0.573418)	(0.543672)
correction	TS	1.107629	1.096789	1.126766	1.096404	1.108264	1.078368
to		(0.562477)	(0.569727)	(0.548695)	(0.548695)	(0.566398)	(0.536652)
Enthalpy	D 1 (1.113922	1.102429	1.132424	1.103187	1.114526	1.084436
	Product	(0.56963)	(0.568767)	(0.555155)	(0.555155)	(0.572765)	(0.542723)
	Decetant	1.11400	1.101889	1.131942	1.103035	1.113446	1.04122
Thomas	Keactain	(0.56925)	(0.574973)	(0.554812)	(0.554812)	(0.572474)	(0.542728)
	ΤC	1.106685	1.095845	1.125822	1.09546	1.10732	1.077423
to Enormy	15	(0.561533)	(0.568783)	(0.547751)	(0.547751)	(0.565454)	(0.535708)
to Energy	Droduct	1.112978	1.101485	1.131479	1.102243	1.113582	1.083492
	Product	(0.568686)	(0.567823)	(0.554211)	(0.554211)	(0.571821)	(0.541779)

The thermal correction to gibbs free energy, thermal correction to enthalpy and thermal correction to energy of all computed structures.^a

^aValues for the single-stranded 5' X T* 3' in parentheses.

The energies (in Hartree Particle⁻¹) and imaginary frequencies (in cm^{-1}) for all computed structures.

			5'AT*3'	5'GT*3'	5'C _a T*3'	5'CT*3'	5'TT*3'	5'UT*3'
-		Reactant	-4734.223270	-4750.274796	-4789.558301	-4750.278642	-4734.216547	-4694.914696
	Energy	TS	-4734.190365	-4750.241661	-4789.527602	-4750.232345	-4734.186459	-4694.884862
ds		Product	-4734.219722	-4750.272479	-4789.556565	-4750.269084	-4734.214427	-4694.912245
-	Virtual		1935i	1863i	1821i	2258i	1767i	1786i
	frequency		(3976) ^a	(3898)	(4750)	(4725)	(4524)	(4408)
		Reactant	-2480.124031	-2555.338297	-2407.749890	-2407.749890	-2466.930804	-2427.629286
	Energy	TS	-2480.085927	-2555.292580	-2407.712211	-2407.712211	-2466.889470	-2427.587859
SS		Product	-2480.113097	-2555.322030	-2407.743384	-2407.743384	-2466.919777	-2427.617913
	Virtual		2070i	1885i	2008i	2008i	2120i	2131i
	frequency		(5335)	(3885)	(5131)	(5131)	(4038)	(4090)

^aValues in parentheses are the IR intensities (in km mol⁻¹).

	M06-2X		CCS	SD(T)
Substituent	$\Delta \mathrm{G}^{\ddagger}$	∆Grxn	$\Delta \mathrm{G}^{\ddagger}$	∆Grxn
CHCH ₂	20.8	1.9	23.8	2.9
ССН	22.0	5.3	25.0	6.6
Ph	22.4	5.9	25.3	9.0
СОН	21.9	6.5	25.4	8.7
COCH ₃	23.4	7.3	26.4	9.3
COOCH ₃	23.7	8.9	26.4	10.7
NO ₂	24.8	8.9	28.5	12.1
СООН	24.4	9.4	27.3	11.3
F	25.8	15.2	28.9	16.8
Н	27.3	19.4	30.0	20.2
NHNH ₃	17.7	8.7	21.4	10.5
NH ₂	18.6	9.1	20.7	11.3
NCH ₃ CH ₃	17.8	9.2	24.0	11.4
NHCOH	21.7	11.2	24.0	12.7
OCH ₃	21.4	11.3	24.8	12.8
ОН	22.0	11.4	24.1	13.1
CH ₂ CH ₃	23.5	15.2	25.7	16.5
CH ₃	24.3	15.4	26.7	16.7

Table S3 The ΔG^{\ddagger} and $\Delta Grxn$ of hydrogen abstraction by the peroxyl radical in the X-CH₂-CH₂-CH₂-OO[•] models. (in kcal mol⁻¹)

	5'AT*3'	5'GT*3'	5'C _a T*3'	5'CT*3'	5'TT*3'	5'UT*3'
$\mathbf{DD}(\mathbf{C1},\mathbf{C2})$	0.9	0.9	0.9	0.9	0.8	0.9
BD(CT-C2)	(0.9)	(0.8)	(0.9)	(0.9)	(1.0)	(0.9)
	3.1	3.0	3.0	2.7	3.1	3.0
BD(C2-H2)	(2.5)	(2.1)	(1.6)	(1.6)	(2.5)	(2.5)
$I P(\Omega 4')$	7.3	6.9	6.8	6.4	7.1	6.8
LF(04)	(6.6)	(7.6)	(7.0)	(7.0)	(6.7)	(6.7)
PD(0, 0)	0.5	0.6	0.2	0.4	0	0
$BD(O_a-O_b)$	(1.0)	(1.2)	(0.5)	(0.5)	(0.7)	(0.7)
	1.9	3.9	1.1	1.6	1.0	1.0
$LF(O_a)$	(1.6)	(1.0)	(0.3)	(0.3)	(0.9)	(0.9)
	2.8	0.4	3.8	4.5	2.8	3.5
$Lr(O_b)$	(0.8)	(1.6)	(0.5)	(0.5)	(1.8)	(1.4)
DD(N0, C8)	0.5	0.5	NI/A	NI/A	NI/A	N/A
BD(119-Co)	(1.0)	(0.7)	IN/A	IN/A	\mathbf{N}/\mathbf{A}	\mathbf{N}/\mathbf{A}
PD(N1,C)	NI/A	N/A	3.0	1.7	0	1.2
BD(NI-C)	\mathbf{N}/\mathbf{A}	1N/A	(2.4)	(2.4)	(2.1)	(1.9)
DD*(N1 C)	NI/A	NI/A	NI/A	NI/A	1.5	0
BD [*] (NI-C)	IN/A	IN/A	IN/A	IN/A	(0)	(1.3)
I D(N1)	N/A	N/A	N/A	N/A	2.0	1.7
	1N/A	1N/A	1N/A	1N/A	(0)	(0)
Total	17.0	16.2	18.8	18.2	18.3	18.1
Total	(14.4)	(15.0)	(13.1)	(13.1)	(15.7)	(16.4)

Table S4 The important second interaction energies ($E^{(2)}$, in kcal mol⁻¹) between $\sigma^*(C1' - H1')$ and electron donors.^a

^aValues for the single-stranded 5' X T* 3' in parentheses.

	5'AT*3'	5'GT*3'	5'C _a T*3'	5'CT*3'	5'TT*3'	5'UT*3'
DD(1)C11U11	1.978	1.976	1.974	1.977	1.975	1.974
BD(I)CI-HI	(1.979)	(1.979)	(1.979)	(1.979)	(1.979)	(1.979)
DD*(1)C1UUU	0.035	0.037	0.037	0.034	0.038	0.037
BD*(1)C1-H1	(0.034)	(0.035)	(0.035)	(0.035)	(0.034)	(0.034)
PD(1)Ob Oc	1.992	1.992	1.991	1.991	1.992	1.992
BD(1)00-0a	(1.992)	(1.993)	(1.993)	(1.993)	(1.992)	(1.992)
$\mathbf{DD}^{*}(1)\mathbf{Ob}\mathbf{O}_{2}$	0.010	0.013	0.015	0.012	0.011	0.011
BD* (1) 00-0a	(0.010)	(0.011)	(0.012)	(0.012)	(0.011)	(0.011)
PD(1)Ob Ob	0.987	0.987	0.988	0.989	0.988	0.988
BD(1)00-0a	(0.987)	(0.987)	(0.988)	(0.988)	(0.987)	(0.987)
PD*(1)Ob Ob	0.023	0.023	0.014	0.015	0.011	0.012
BD ⁺ (1)00-0a	(0.010)	(0.009)	(0.010)	(0.010)	(0.011)	(0.011)

Table S5 Electron Occupancy numbers of the selected bonds in the duplex and the single-stranded models.^a

 $^{\mathrm{a}}\mathrm{Values}$ for the single-stranded 5' X T* 3' in parentheses.

		5'AT*3'	5'GT*3'	5'C _a T*3'	5'CT*3'	5'TT*3'	5'UT*3'
		159.2	169.5	140.0	163.9	130.0	135.4
	ZCIHIOb	(171.7)	(147.7)	(154.7)	(154.7)	(175.0)	(113.2)
	/111/0_0	69.7	56.3	77.0	74.1	76.2	76.7
Depatant	$\angle HTO_bO_a$	(69.4)	(69.2)	(78.1)	(78.1)	(76.5)	(76.9)
Reactant	1	1.094	1.093	1.092	1.091	1.094	1.093
	u C1'-H1'	(1.093)	(1.092)	(1.092)	(1.092)	(1.092)	(1.092)
	$d_{\mathrm{Ob-Oa}}$	1.299	1.303	1.300	1.301	1.307	1.300
		(1.304)	(1.303)	(1.302)	(1.302)	(1.304)	(1.304)
		$\begin{array}{cccccccccccccccccccccccccccccccccccc$	173.6	173.0			
	ZCIHIOb	(169.2)	(169.9)	(159.8)	(159.8)	(154.2)	(155.0)
	/111/0_0	101.6	102.4	101.2	101.3	102.0	102.2
ΤS	∠HI U _b U _a	(100.9)	(101.8)	(102.3)	(102.3)	(99.2)	(99.2)
15	L	1.277	1.270	1.263	1.325	1.257	1.261
	d C1'-H1'	(1.295)	(1.279)	(1.294)	(1.294)	(1.310)	(1.311)
	d	1.382	1.381	1.382	1.382	1.381	1.382
	u _{Ob-Oa}	(1.386)	(1.381)	(1.384)	(1.384)	(1.387)	(1.387)

Table S6 The \angle C1'H1'O_b, \angle H1'O_bO_a (in °) and dC1'-H1', dO_b-O_a (in Å) in the reactants and transition species of duplex 5' X T* 3' and its opposite single-stranded models.^a

^aValues for the single-stranded 5' X T* 3' in parentheses.

		5'AT*3'	5'GT*3'	5'C _a T*3'	5'CT*3'	5'TT*3'	5'UT*3'
	$\pi^*(O_a-O_b)$	-9.7	-9.6	-9.7	-9.5	-9.6	-9.5
ds	σ(C1'-H1')	-11.1	-10.9	-11.3	-11.5	-11.1	-10.4
	ΔE	1.4	1.3	1.6	2.0	1.5	0.9
	$\pi^*(O_a-O_b)$	-9.7	-10.5	-9.0	-9.0	-9.7	-9.4
SS	σ(C1'-H1')	-11.6	-12.5	-11.0	-11.0	-11.8	-11.3
	ΔE	1.9	2.1	2.1	2.1	2.1	1.9

Table S7 The molecular orbital energy gap (ΔE , in ev) between π^* and σ (C1'-H1').

Table S8 The BDE, DPE and BDFE of natural deoxyribonucleosides (in kcal mol⁻¹)

	Method	dAMP	dGMP	dCMP	dTMP	dUMP
	BDE	95.4	92.1	95.0	95.4	95.4
ds	DPE	396.6	385.9	389.8	393.3	389.5
	BDFE	86.0	83.1	85.4	86.4	86.4
	BDE	94.4	94.4	94.5	95.8	95.7
SS	DPE	390.8	389.3	393.5	392.5	391.1
	BDFE	85.6	85.6	86.5	86.9	86.8

C1'-H1' BDEs of natural deoxyribonucleosides are calculated by M06-2X at 6-31+G(d,p) level. The values are shown in Table S6, it is should be noted that our values are excellent agreement with Guo's reports.¹ The acidity (represented by deprotonation enthalpy, DPE) and bond dissociation free energy (BDFE) of the C1'-H1' were calculated at the same level. For comparison, we also calculated the BDE, DPE and BDFE of duplex models.



 $X = CHCH_2$, CCH, Ph, COH, COCH₃, COOCH₃, NO₂, F, COOH, H, NHCH₃, NH₂, NCH₃CH₃, NHCOH, OCH₃, OH, CH₂CH₃, CH₃

Scheme S1 Schematic depiction of the intramolecular hydrogen abstraction by the peroxyl radical.

The reaction can be described as:

$$X - CH_2 - CH_2 - CH_2 - COO^{\bullet} \rightarrow X - {}^{\bullet}CH - CH_2 - CH_2 - COO^{\bullet} + H^{\bullet}$$
(1)

$$H^{\bullet} + X - {}^{\bullet}CH - CH_2 - CH_2 - COO^{\bullet} \rightarrow X - {}^{\bullet}CH - CH_2 - CH_2 - COOH$$
(2)

 ΔG rxn

$$= G(X - CH - CH_2 - CH_2 - COOH) - G(X - CH_2 - CH_2 - CH_2 - COO^{\bullet})$$

= $\Delta Grxn(1) + \Delta Grxn(2)$

= BDFE(C-H) - BDFE(O-H)

Definition of the torsion angle ϕ between the adjacent 5' and 3' bases

As the position of C2, N3, C4 on the modified thymine base is less changed, we use C2, N3 and O4 to determine the plane of pyrimidine base, C2, N1 and O6 to determine the plane of purine base. The ϕ is defined as the angle between 5' base plane and 3' base plane.





To reach a good computational efficiency without loss of accuracy, the ONIOM² approach was utilized in the calculation. As shown, the duplex 5' T T* 3' is divided into three regions. The most important reactive region for determining the properties of the molecular system includes the two stacking thymine nucleotides, which were treated at the M06-2X/6-31+G(d,p) level and represented using ball and bond. The two hydrogen-bonding supplementary adenine bases were dealt at the M06-2X/3-21G* level, which were set as the medium layer and represented as a tube. The remaining part of the whole molecule, including phosphate and sugar groups described as wireframe, was treated with a less rigorous but more computationally efficient PM6 theory.³ Hence, the ONIOM(M06-2X/ 6-31+G(d,p): M06-2X/3-21G*:PM6) combination method was used to optimize the model system. According to the ONIOM calculations, the activation barriers of the C1' - H1'and C5' - H5' bonds are 14.5 and 23.5 kcal mol⁻¹, respectively.

Notes and references

- 1. M. J. Li, L. Liu, Y. Fu and Q. X. Guo, J. Phys. Chem. B, 2005, 109, 13818-13826.
- 2. T. Vreven and K. Morokuma, *in Annual Reports in Computational Chemistry*, Ed. D. C. Spellmeyer, 2006, **2**, 35-51.
- 3. J. J. P. Stewart J. Mol. Model. 2007, 13, 1173-1213.