Understanding the interaction of nucleotides with UVC light an insight from quantum chemical calculation-based findings

Tan, Chunjian; Wang, Shaogang; Yang, Huiru; Huang, Qianming; Li, Shizhen; Liu, Xu; Ye, Huaiyu; Zhang, Guoqi

DOI
10.1039/d2cp05054d

Publication date
2023

Document Version
Final published version

Published in
Physical Chemistry Chemical Physics

PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d2cp05054d

Understanding the interaction of nucleotides with UVC light: an insight from quantum chemical calculation-based findings†

Chunjian Tan, [ab] Shaogang Wang,[ab] Huiru Yang,[b] Qianming Huang,[b] Shizhen Li,[b] Xu Liu,[ab] Huaiyu Ye*[b] and Guoqi Zhang*[a]

Received 28th October 2022,
Accepted 5th December 2022

DOI: 10.1039/d2cp05054d

rsc.li/pccp

Short-wave ultraviolet (also called UVC) irradiation is a well-adopted method of viral inactivation due to its ability to damage genetic material. A fundamental problem with the UVC inactivation method is that its mechanism of action on viruses is still unknown at the molecular level. To address this problem, herein we investigate the response mechanism of genome materials to UVC light by means of quantum chemical calculations. The spectral properties of four nucleotides, namely, adenine, cytosine, guanine, and uracil, are mainly focused on. Meanwhile, the transition state and reaction rate constant of uracil molecules are also considered to demonstrate the difficulty level of adjacent nucleotide reaction without and with UVC irradiation. The results show that the peak wavelengths are 248.7 nm, 226.1 nm (252.7 nm), 248.3 nm, and 205.8 nm (249.2 nm) for adenine, cytosine, guanine, and uracil nucleotides, respectively. Besides, the reaction rate constants of uracil molecules are 6.419×10^{-49} s^{-1} M^{-1} and 5.436×10^{11} s^{-1} M^{-1} for the ground state and excited state, respectively. Their corresponding half-life values are 1.56×10^{48} s and 1.84×10^{-12} s. This directly suggests that the molecular reaction between nucleotides is a photochemical process and the reaction without UVC irradiation almost cannot occur.

1 Introduction

Over the past few decades, several lethal pathogens including SARS-CoV,[1] MERS-CoV,[2] Ebola,[3] H1N1,[4] and SARS-CoV-2,[5] have hugely impacted public health and economic outlook worldwide. Meanwhile, the emergence of these viruses also continuously brings the available viral and bacterial inactivation approaches into the spotlight. For sterilization, a number of disinfection methods, *e.g.*, chemical,[6–8] thermal,[6] photochemical,[9,10] and advanced approaches,[11,12] have been explored to inactivate lethal pathogens from contaminated surfaces or in the air. Among these methods, UVC disinfection is regarded as the most promising technique for inactivating pathogens due to its manageable risk and low-cost.[13] UVC irradiation using photons of light from the 100–280 nm region of the electromagnetic spectrum, an effective and non-contact method for inactivating viruses and bacteria, has been widely used for disinfecting food, water supplies, public facilities, and medical instruments.[14–17] The UVC range shows the strongest antiviral and antibacterial activity among the three applied UV wavelength ranges, including UVA (315–400 nm) and UVB (280–315 nm) ranges.[18] The available scientific literature shows that three UVC range subsets (*i.e.*, 207–222 nm,[19,20] 254 nm,[21,22] and 260–280 nm[23,24]) are considered to be optimal for pathogen disinfection. A wavelength of 254 nm is the commonly used inactivation wavelength in most disinfection systems. Studies suggest that the 260–280 nm wavelength range is more efficient than 254 nm because these wavelengths are closer to the maximum absorption wavelength of genome materials.[25,26] Besides, there is increasing interest in 207–222 nm wavelengths because of their high germicidal effectiveness.[27,28] These literature studies intensely highlight the high feasibility of UVC irradiation in disinfection. Nevertheless, the mechanism of action of UVC light on pathogens is still debatable.

It was commonly understood that the inactivation mechanism of UVC irradiation against pathogens is DNA and RNA damage, primarily through the formation of thymine and pyrimidine dimers.[29–31] This genome damage would disrupt nucleic acid replications, thus leading to the death of various pathogens. To uncover the inactivation mechanism of UVC irradiation, considerable efforts have been made and substantial achievements have been made.[32–34] Four possible mechanisms have

[a] *Electronic Components, Technology and Materials, Delft University of Technology, 2628 CD Delft, The Netherlands. E-mail: G.Q.Zhang@tudelft.nl*
[b] *Engineering Research Center of Integrated Circuits for Next-Generation Communications, Ministry of Education, School of Microelectronics, Southern University of Science and Technology, Shenzhen 518055, P. R. China. E-mail: yehy@sustech.edu.cn*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d2cp05054d

been reported to be responsible for UVC light inactivation against pathogens. One is the viral protein oxidation, which is related to the reduction of viral infectivity in feline calicivirus and bacteriophage MS2.[35–37] Second is the destruction of the viral capsid protein, being found in the treatment of MNV-1 (murine norovirus 1) by using UVC irradiation.[38] Third is the viral protein–genome cross-linking, being observed in poliovirus treatment.[39] Fourth is the viral genome damage, being supported by the UVC irradiation of influenza viruses.[40] Briefly, the viral inactivation mechanism with UVC irradiation can be attributed to the genome damage and/or the protein destruction due to the absorption of UVC photons by nucleotides and/or proteins. Furthermore, one published study reported that UVC irradiation induces the genome damage of adenovirus, but has almost no effect on its proteins.[32] Until now, most of the previous studies focusing on UVC inactivation have only been at the experimental level, and theoretical research studies on interactions between UVC light and nucleotides still do not provide quantitative evidence to illustrate the effect of UVC light. Consequently, the necessity and scientific significance of the study of the inactivation mechanism of UVC irradiation against viruses by a molecular-level simulation is strongly emphasized.

In this article, by using quantum chemical calculations, we theoretically interpret the underlying mechanism of genome matter interactions with UVC light. Four nucleotides, namely, adenine, cytosine, guanine, and uracil molecules, are chosen as the objects for investigation. The absorption spectra clearly show that the absorption peaks of these four nucleotides are completely positioned in the range of 200–280 nm, which is close to the reported wavelengths.[41,42] Besides, the reaction rate constants of uracil molecules at the ground state and excited state turned out to be 6.419×10^{-49} s^{-1} M^{-1} and 5.436×10^{11} s^{-1} M^{-1}, respectively. The corresponding half-life values are 1.56×10^{48} s and 1.84×10^{-12} s, respectively. These results directly confirm the role of UVC light in inactivating viruses and the difficulty level of adjacent nucleotide reaction without and with UVC light.

2 Computational methodologies

All DFT and TDDFT calculations are performed using the Gaussian 16 program package.[43] As mentioned above, genome damage is one of the major causes of viral death. Therefore, we only consider the interaction of nucleotides composed of viral RNA with UVC light in this study. The geometrical structures, and UV-vis absorption spectrum of four single nucleotides, namely, adenine, cytosine, guanine, and uracil with the sugar phosphate backbone, are calculated by adopting the B3LYP hybrid functional with the 6-311+G(d,p) basis set for all atoms.[44] For determining the electron excitation feature, transition sate, and intrinsic reaction coordinate, we also invoke the wB97XD hybrid functional with the def2tvzp basis set in order to obtain better electron excitation description and more accurate energy parameters (*e.g.*, electronic energy). These computational methods are adequate and affordable for the four investigated nucleotides. The UV-vis absorption spectrum and hole–electron analysis are dealt with using the Multiwfn program that is powerful for implementing electronic wavefunction analysis.[45] The VMD package (version 1.9.3) is used to better visualize the investigated nucleotides and their hole–electron distribution.[46] In addition, the thermodynamic parameters of the reactant and transition state, required in the calculation of the reaction rate constant, are calculated by utilizing the Shermo code.[47]

The reaction rate constant predicted using the transition state theory (TST) can help theoretically characterize the speed of chemical reactions. Therefore, we can determine whether nucleotides are susceptible to reaction under S0 and S1 through this parameter. On the basis of illustration of the above section, we only focus on the reaction rate constant of uracil molecules to confirm the photochemical reaction between nucleotides. Theoretically, the conventional TST rate can be expressed as

$$k_{\mathrm{TST}}(T) = \sigma \frac{k_{\mathrm{B}}}{h} \frac{Q_{\mathrm{TS}}(T)}{\Phi_{\mathrm{R}}(T)} \exp\left[-V^{\ddagger}/k_{\mathrm{B}}T\right] \quad (1)$$

where k_{B}, h, $V^{\ddagger}$, and T are Boltzmann's constant, Planck's constant, the classical barrier height, and the thermodynamic temperature.[48] Besides, $Q_{\mathrm{TS}}(T)$ and $\Phi_{\mathrm{R}}(T)$ are the quantum mechanical transition state quasi-distribution function and reactant distribution function, respectively. According to the thermodynamic theory, the TST formula can be rewritten in an equivalent thermodynamic form:

$$k_{\mathrm{TST}}(T) = \sigma \frac{k_{\mathrm{B}}T}{h} \left(\frac{RT}{P_0}\right)^{\Delta n} \mathrm{e}^{-\Delta G^{0,\neq}/(k_{\mathrm{B}}T)} \quad (2)$$

where $G^{0,\neq}(T) = G^{0}_{\mathrm{TS}}(T) - G^{0}_{\mathrm{Reactant}}(T)$ is the standard Gibbs energy of activation[49,50] and $G^{0}_{\mathrm{TS}}(T)$ does not consider the contribution of imaginary frequency. Moreover, Δn is set as 1 for the gas phase bimolecular reaction, whereas it is zero for the single molecule reaction. RT/P_0 is the reverse of the molar concentration. It should be noted that σ would not be taken into account in $k_{\mathrm{TST}}(T)$ if the rotational symmetry is considered when calculating Gibbs energy. The reaction path degeneracy σ can be calculated using the formula

$$\sigma = \frac{\sigma_{\mathrm{rot,Reactant}}}{\sigma_{\mathrm{rot,TS}}} \quad (3)$$

where $\sigma_{\mathrm{rot,Reactant}}$ and $\sigma_{\mathrm{rot,TS}}$ are the rotational symmetry numbers of the reactant and transition state, respectively. In our calculation, T, P, R, and σ are set to 298.15 K, 1 bar, 8.31 J mol^{-1} K^{-1}, and 1 respectively.

3 Results and discussion

3.1 UV-vis absorption spectra of the four RNA nucleotides

Absorption spectra, also called absorption curves, refer to a spectrum resulting from the transition of electrons from a low-energy level to a high-energy level when substances absorb photons, which mainly depends on the atomic and molecular composition of samples.[51] An absorption spectrum can be employed to characterize the excited state of electrons, and

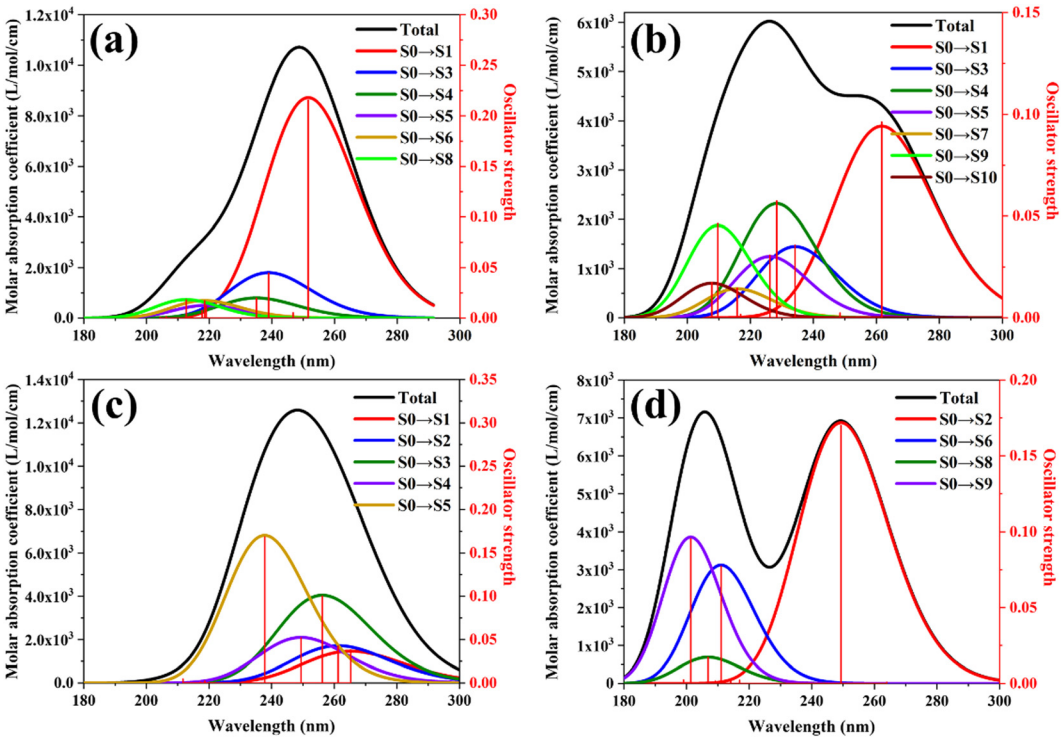

Fig. 1 UV-vis absorption spectra of adenine (a), cytosine (b), guanine (c), and uracil (d) nucleotides.

achieve a link between experimental and theoretical studies on the absorption properties of biological molecules. In addition, the knowledge of the absorption spectrum also plays an important role in the optical applications of biological molecules. Most importantly, some previous studies focusing on the viral inactivation through UV light have clearly revealed that UV photon-energy absorption is the leading cause of the production of intrastrand cyclobutyl-pyrimidine dimers, which directly results in genetic mutations or viral death.[52–55] Consequently, in this section, we discuss the absorption spectra of four single nucleotides in the gas phase. To clarify the contribution of each excited state, only the excited states with an oscillator strength greater than 0.01 are considered. Note that ten excited states are considered, namely, S1, S2, S3, S4, S5, S6, S7, S8, S9, and S10. In Fig. 1, we show the absorption spectra of adenine, cytosine, guanine, and uracil nucleotides in the gas phase. The total absorption spectra of four nucleotide molecules are consistent with the depicted spectra in the literature.[41,42] From the perspective of the total absorption curve, the strongest absorption peaks for adenine, cytosine, guanine, and uracil nucleotides are observed at 248.7 nm, 226.1 nm (second peak: 252.7 nm), 248.3 nm, and 205.8 nm (second peak: 249.2 nm) respectively, which are close to the reported wavelengths.[41,42] Meanwhile, these absorption peaks are also close to the reported values in the literature, in which the absorption peaks for adenine, cytosine, guanine, and uracil in the gas phase are at 255.84 nm, 262.27 nm, 259.18, and 255.70 nm, respectively.[56] As shown in Fig. 1(a), we can observe that the total absorption spectrum of adenine is dominantly composed of S1, S3, and S4, and the contribution magnitudes for the absorption peak are 0.80, 0.13, and 0.04, respectively. Therefore, the adsorption peak is led by S1. For cytosine, it can be found that S1, S3, S4, S5, S7, and S9 dominate the total absorption spectrum, as plotted in Fig. 1(b). Its strongest peak mainly consists of S3, S4, S5, S7, and S9, and their contribution rates are 0.19, 0.38, 0.21, 0.06, and 0.09, respectively. The second peak around 260 nm is almost dominated by S1 and the corresponding contribution rate is 0.72. Fig. 1(c) clearly displays that S1, S2, S3, S4, and S5 mainly form the total absorption spectrum of guanine. The order of contribution magnitude for the absorption peak is S5 > S3 > S4 > S2 > S1 and the specific values are 0.40, 0.28, 0.17, 0.09, and 0.06, respectively. In the case of uracil, as presented in Fig. 1(d), the total absorption spectrum is almost contributed by S2, S6, S8, and S9. Its first absorption peak around 200 nm is led by the S6, S8, and S9, and their contribution rates are 0.38, 0.10, and 0.48, respectively. The second absorption peak near 250 nm is almost completely dominated by S2 and the contribution magnitude is 0.99.

3.2 Electron excitation characteristics of four RNA nucleotides

Adenine molecule. Hole–electron analysis, as a very intuitive and versatile method for revealing the excitation characteristics of electrons in a system, has been widely used in many research studies.[57–60] Therefore, to interpret the excitation characteristics of electrons of adenine nucleotides, the isosurface diagrams of hole and electron distributions in different excited states are plotted in Fig. S1 and S2 (ESI†). The corresponding excitation characteristic parameters are listed in Table 1. The electron excitation involved in UV-vis spectroscopy can be

Table 1 Excitation characteristic parameters of various excited states for adenine. The D index is referred to the total magnitude of charge transfer length. Sr is defined as the overlapping extent of holes and electrons. The H index is an overall measure of the average degree of spatial extension of the hole and electron distribution in the $x/y/z$ direction. The t index is designed to measure the separation degree of holes and electrons in the charge transfer direction. The $\Delta\sigma$ index reflects the difference in the overall spatial distribution breadth of electrons and holes. HDI and EDI are hole and electron delocalization indexes, respectively, being pretty useful in quantifying the breadth of the spatial distribution. Ecoul represents the Coulomb attraction between holes and electrons. T type represents the transition type of excited states

Excited state	D (Å)	Sr	H (Å)	t (Å)	$\Delta\sigma$ (Å)	Ecoul (eV)	HDI	EDI	T type
S0 → S1	0.243	0.827	2.393	−1.240	0.308	5.90	8.50	8.78	π → π*
S0 → S2	0.449	0.885	2.254	−0.842	0.011	6.23	8.17	8.28	π → π*
S0 → S3	0.660	0.485	2.197	−0.637	0.293	6.35	15.38	9.69	n → π*
S0 → S4	0.650	0.524	2.118	−0.620	0.148	6.61	15.63	9.77	n → π*
S0 → S5	0.330	0.526	2.428	−1.104	0.235	5.79	14.33	8.80	n → π*
S0 → S6	0.670	0.778	2.611	−1.078	0.310	5.37	8.23	6.97	π → π*
S0 → S7	1.149	0.532	2.892	−0.618	1.079	4.65	9.45	2.72	π → 3s
S0 → S8	0.342	0.810	2.649	−1.400	0.092	5.37	8.31	7.70	π → π*
S0 → S9	0.632	0.821	2.481	−0.946	0.403	5.59	8.20	7.65	π → π*
S0 → S10	0.438	0.615	2.492	−1.214	−0.182	5.67	10.77	8.45	n → π*

divided into two categories, namely, valence and Rydberg excitations. The valence excitation refers to the excitation of electrons from the valence occupied orbital to the valence empty orbital. Rydberg excitation is described as the excitation of electrons from the valence occupied orbital to the Rydberg orbital. From the perspective of charge transfer, the valence excitation includes local excitation (LE) and charge-transfer (CT) excitation. According to natural transition orbital (NTO) analysis, we find that the excited state ranging from 1 to 10 can be classified into three transition categories, namely, π → π*, n → π*, and π → 3s. Among all the excited states, the excited state 7 (S7) is a π → 3s transition type that is a Rydberg excitation, whereas S3, S4, S5, and S10 belong to the n → π* category. The rest of the excited states including S1, S2, S6, S8, and S9 are determined as the transition type π → π*. Note that the isosurface maps of the NTOs of different excited states of adenine nucleotides are not presented here. On the basis of charge transfer, combining the results shown in Fig. S1 and S2 (ESI†) and Table 1, all the excited states of adenine nucleotides belong to LE except for S7. For instance, S1 is clearly an LE because it possesses a small D-index, a large Sr-index, an apparently negative t-index, and a modest $\Delta\sigma$-index. The isosurface map of the Chole–Celectron of S1 shown in Fig. S1 (ESI†) further verifies this analytical result. At the same time, the corresponding four indexes in S8 also exhibit the same feature as those in S1. Interestingly, the D, Sr, t, and $\Delta\sigma$ indexes of S3 and S4 are unlike those of the S1 where these indexes can be used to determine what type of transition does S1 belongs to. These four indexes of S3 and S4 are in an intermediate state with the determination of their excitation features being difficult. However, their HDI and EDI indexes are the largest among these excited states, which indicates a small uniformity in the hole and electron distribution, as shown in Fig. S1 (ESI†). For S6 and S9, although their D indexes are similar to those of S3 and S4, their t indexes are more negative than those of S3 and S4. Thus, the excitation features of S6 and S9 can be easily determined as LE. For S2 and S10, we find a similar behavior to that of S6 and S9.

Cytosine molecule. By means of NTO analysis, the excitation of cytosine nucleotide can be categorised into three categories, namely, n → π*, π → π*, and π → 3s. Surprisingly, all the excited states ranging from 1 to 10 belong to the n → π* transition type, except for S5 and S9. The NTO result suggests that S5 and S9 belong to the π → π* and π → 3s transition type, respectively. Hence, it is clear that S9 is the Rydberg excitation. Depending on hole–electron analysis, there are three excitation categories (*i.e.*, the LE, CT, and Rydberg excitations) of electrons for cytosine nucleotides within the investigated excited states. As listed in Table 2, all the D indexes are larger than 0.5 Å; thus it is necessary to utilize other indexes to better determine the excitation feature of electrons. It should be noted that there is no criterion for determining that the excitation of electrons is a CT excitation if the D index is greater than a certain value. Therefore, it is misleading to adopt a strict classification criterion for determining the excitation features.

Table 2 Excitation characteristic parameters of various excited states for cytosine

Excited state	D (Å)	Sr	H (Å)	t (Å)	$\Delta\sigma$ (Å)	Ecoul (eV)	HDI	EDI	T type
S0 → S1	0.750	0.727	2.137	−0.612	−0.048	6.47	11.09	10.00	n → π*
S0 → S2	1.244	0.457	1.879	0.114	0.448	6.82	19.31	10.44	n → π*
S0 → S3	0.959	0.712	2.264	−0.303	−0.082	6.04	11.98	9.36	n → π*
S0 → S4	1.294	0.553	2.083	−0.111	0.282	6.35	20.92	10.50	n → π*
S0 → S5	1.641	0.507	2.836	−0.130	1.093	4.59	10.30	3.53	π → π*
S0 → S6	1.534	0.643	2.514	−0.161	0.534	5.24	12.93	6.85	n → π*
S0 → S7	0.763	0.673	2.457	−0.833	0.417	5.70	10.48	8.42	n → π*
S0 → S8	1.247	0.536	2.082	0.072	0.416	6.26	17.39	9.63	n → π*
S0 → S9	1.699	0.409	3.102	−0.099	1.683	4.21	11.47	2.79	π → 3s
S0 → S10	0.925	0.610	2.428	−0.323	0.446	5.71	13.45	8.82	n → π*

For S1, we can observe a minimum D value, a maximum Sr value, and a most negative t; thus S1 is determined as LE. Even if the $\Delta\sigma$ index of S7 is much larger than that of S1, its D and Sr indexes are close to those of S1 and the corresponding t index is more negative than that of S1. Hence, it belongs to the same excitation category as S1, which is verified using the Chole–Celectron maps of S1 and S7. They have a similar Chole–Celectron distribution, as shown in Fig. S3 and S4 (ESI†). Importantly, it is found that the t indexes of S2 and S8 are positive, suggesting that these two excited states belong to CT excitation. The corresponding isosurface maps of Chole–Celectron of S2 and S8 provide powerful evidence, presenting that the separation of holes and electrons is sufficient. Besides, their D values also further confirm this discussion. In S3 and S10, it is obvious that their D, Sr, and t values are quite close, but there is a big difference between their $\Delta\sigma$ values. Meanwhile, these focused indexes help determine their excitation features with difficulty. Combining their isosurface diagrams of hole and electron distributions, S3 and 10 are determined as LE due to an insufficient hole–electron separation. S4, S5, and S6 possess a sufficiently large D value (see Table 2), which means that these three excited states belong to CT excitation. At the same time, their t values are slightly negative compared to those of S1 and S7. This provides another evidence for this determination. In addition, it should be noted that the Sr value of S6 is larger than that of S4 and S5, indicating that the hole–electron overlap of S6 is higher than that of other two excited states. As we can observe in the isosurface maps of Chole–Celectron of S4, S5, and S6, it is found that their hole–electron separation is not yet sufficient.

Guanine molecule. Differentiating with adenine and cytosine nucleotides, there are five transition categories for all the studied excited states in guanine nucleotides also based on NTO analysis, as listed in Table 3. Among these five categories, the valence excitation includes $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$, whereas $\pi \rightarrow$ 3s, $\pi \rightarrow$ 3py, and $\pi \rightarrow d_{x^2}$–d_{y^2} correspond to the Rydberg excitation. As discussed above, only one type ($\pi \rightarrow$ 3s) of Rydberg excitation is found in adenine and cytosine nucleotides. We can observe that S1, S2, and S7 belong to the $\pi \rightarrow \pi^*$ transition type, while S3, S6, S8, and S10 belong to the $n \rightarrow \pi^*$ transition type. Moreover, the NTO result for S4, S5, and S9 obviously illustrates that their isosurface maps of hole and electron distributions have significant Rydberg excitation characteristics; thus their transition features correspond to $\pi \rightarrow$ 3s, $\pi \rightarrow$ 3py, and $\pi \rightarrow d_{x^2}$–d_{y^2}, respectively. Similar to adenine and cytosine nucleotides, the LE and CT excitations are also determined as the transition features of the remaining excited states except for S4, S5, and S9. S8 possesses a minimum D value in the excited states ranging from 1 to 10, but its corresponding Sr value is modest. Fortunately, its t value is sufficiently negative and the $\Delta\sigma$ value is not large. Accordingly, we can conclude that S1 belongs to LE, and its Chole–Celectron distribution also confirms this conclusion. For S6 and S7 with similar D values, the magnitude of their D and t values is consistent with the numerical characteristics of LE. In particular, the Sr value of S7 is the maximum in the investigated excited states. Hence, it is easy to determine that the excitation type of electrons for S7 is LE. In contrast, it is difficult to make a decision for the excitation type of S6 because of its small Sr value. However, through observing the Chole–Celectron isosurface diagram of S6, it is clearly observed that its separation of holes and electrons is not significant. Above all, we can conclude that S6 belongs to LE. For S2, although its D value is not large or small, it is easy to conclude that S2 belongs to LE through combining the Sr and t values, as well as its Chole–Celectron map. For S3 and S10, the positive value of the t index directly displays that both the excited states are CT excitations due to their sufficient separation of holes and electrons, which is also further demonstrated by the corresponding Chole–Celectron distribution diagrams.

Uracil molecule. Similar to cytosine nucleotides, three transition categories, namely, $n \rightarrow \pi^*$, $\pi \rightarrow \pi^*$, and $\pi \rightarrow d_{x^2}$–d_{y^2}, are observed in uracil nucleotides by means of NTO analysis. Meanwhile, most of the excited states for both nucleotides belong to the $n \rightarrow \pi^*$ transition type. As listed in Table 4, it is found that S2 and S4 are determined as the $\pi \rightarrow \pi^*$ transition type, while the remaining excited states belong to the $n \rightarrow \pi^*$ transition type except for S7. Because of the $\pi \rightarrow d_{x^2}$–d_{y^2} transition feature, S7 is identified as the Rydberg excitation. For S2, a small D value, a large Sr value, a negative t value, and a modest $\Delta\sigma$ value directly provide powerful evidence that S2 belongs to LE. For S3 and S4, the D and t values of these two excited states are close, but their magnitude is not large or small. Thus, it is difficult to determine the excitation features of S3 and S4. Fortunately, the Sr value of S4 appears to be large enough. Therefore, S4 belongs to LE. Linking with the Chole–Celctron isosurface map of S3, we find that there is no obvious

Table 3 Excitation characteristic parameters of various excited states for guanine

Excited state	D (Å)	Sr	H (Å)	t (Å)	$\Delta\sigma$ (Å)	Ecoul (eV)	HDI	EDI	T type
S0 → S1	0.475	0.758	2.338	−0.874	0.034	5.99	9.34	9.05	$\pi \rightarrow \pi^*$
S0 → S2	0.933	0.731	2.466	−0.560	0.362	5.54	9.13	9.05	$\pi \rightarrow \pi^*$
S0 → S3	1.231	0.474	1.791	0.086	0.581	7.16	25.41	12.78	$n \rightarrow \pi^*$
S0 → S4	2.100	0.326	2.944	0.375	1.135	4.24	9.78	2.78	$\pi \rightarrow$ 3s
S0 → S5	2.132	0.375	3.134	0.335	1.600	3.98	9.06	3.12	$\pi \rightarrow$ 3py
S0 → S6	0.339	0.460	2.332	−1.020	0.341	5.98	15.53	9.39	$n \rightarrow \pi^*$
S0 → S7	0.331	0.769	2.696	−1.228	0.687	5.16	8.95	7.24	$\pi \rightarrow \pi^*$
S0 → S8	0.219	0.524	2.282	−1.025	0.283	6.17	16.16	9.64	$n \rightarrow \pi^*$
S0 → S9	2.015	0.409	3.339	−0.130	1.917	3.91	8.81	2.58	$\pi \rightarrow d_{x^2}$–d_{y^2}
S0 → S10	2.453	0.407	2.261	1.142	0.925	4.75	22.69	9.62	$n \rightarrow \pi^*$

Table 4 Excitation characteristic parameters of various excited states for uracil

Excited state	D (Å)	Sr	H (Å)	t (Å)	$\Delta\sigma$ (Å)	Ecoul (eV)	HDI	EDI	T type
S0 → S1	1.448	0.456	1.802	0.116	0.414	7.16	24.91	11.86	n → π*
S0 → S2	0.325	0.747	2.170	−1.114	−0.117	6.60	11.15	9.77	π → π*
S0 → S3	0.826	0.466	2.204	−0.351	−0.193	6.31	20.11	11.85	n → π*
S0 → S4	0.859	0.680	2.311	−0.292	0.026	6.05	10.89	11.39	π → π*
S0 → S5	1.633	0.562	2.044	0.654	0.081	5.90	15.90	10.18	n → π*
S0 → S6	0.735	0.484	2.476	−0.996	−0.463	5.85	16.15	9.03	n → π*
S0 → S7	0.411	0.455	2.949	−1.355	1.110	4.68	10.39	2.71	π → d_{x^2}–d_{y^2}
S0 → S8	2.309	0.456	2.436	0.604	−0.281	4.97	16.00	8.64	n → π*
S0 → S9	3.427	0.297	2.189	1.992	−0.170	4.16	17.28	9.59	n → π*
S0 → S10	2.175	0.471	2.270	0.717	−0.086	5.19	18.40	9.65	n → π*

hole–electron separation. As a result, S3 belongs to LE. Even if S6 possesses relatively small D and Sr values, an obviously negative t value leads to a hypothesis that S6 may belong to LE. Hereafter, this hypothesis is confirmed by an insufficient hole–electron separation, as plotted in the Chole–Celectron isosurface map of S6. Surprisingly, the remaining excited states, *i.e.*, S1, S5, S8, S9, and S10, are assigned to the CT excitation because their t indexes are positive. At the same time, the corresponding D values also further give another powerful evidence.

3.3 Transition state of dual uracil reaction

Biologically, nucleic acids have excellent photostability. However, photochemical damage has been observed in adjacent base monomers, oligonucleotides, and DNA.[61,62] Photo-dimerization induced by UV damage dominantly occurs in adjacent pyrimidines (in order YY = UU, CU, UC, and UU) bonded with cyclobutane-type linkages.[63] The cross-linking and photo-dimerization between uracil nucleotides are responsible for inactivation of RNA viruses.[64] The long RT-qPCR genome analysis has clearly confirmed that there are dimer sites in the SARS-CoV-2 genome after UVC treatment.[65] As plotted in Fig. S9 (ESI†), the most UV-reactive dinucleotides belong to the pyrimidine group, such as UU, UC, CU, and CC, with UU dinucleotides showing the highest UV-reactivity, with the most extreme data even beyond the interquartile range. Additionally, non-YY sequences may not result in cyclobutane-type dimers, which may be due to transient contacts of base pairs.[63] Accordingly, we only focus on the photochemical reaction process of uracil dimers to reveal the molecular mechanism of viral inactivation. It should be noted that the photo-dimerization process of nucleic acids mainly takes place in the base portion, not in the sugar phosphate portion; thus there is no sugar phosphate portion in uracil.[17,64,65]

The initial guess structures of the transition state of uracil base reaction in S0 and S1 are utilized to search for the actual

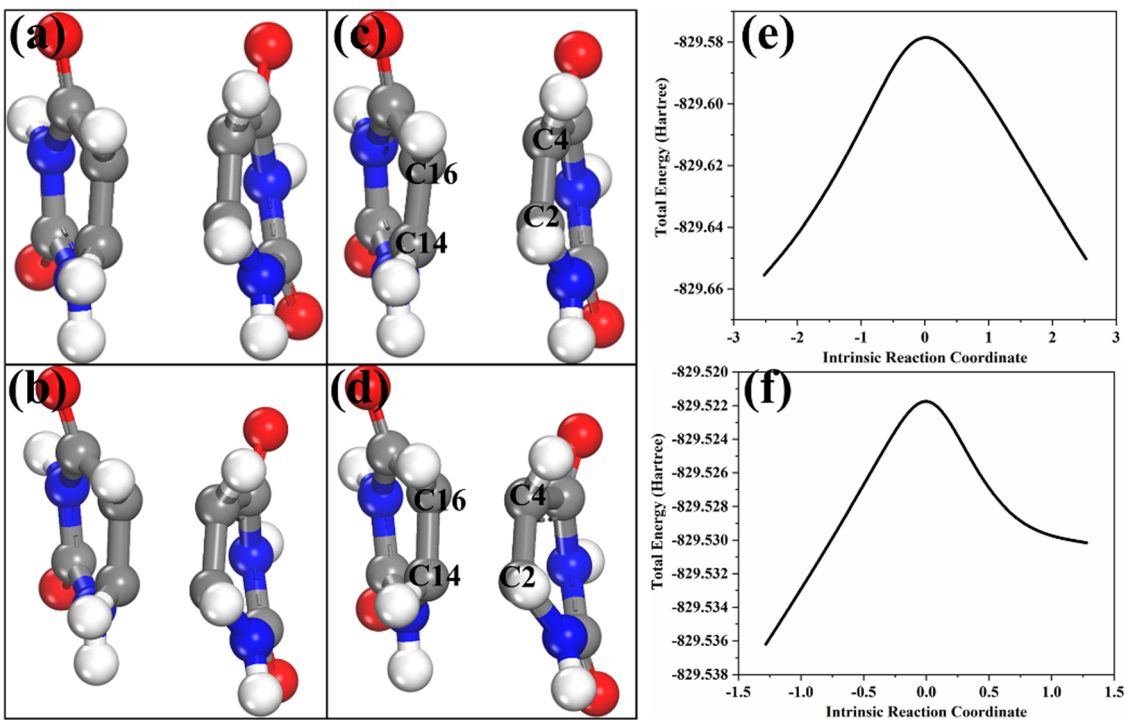

Fig. 2 (a and b) Initial guess structures and (c and d) actual structures of the transition-state of uracil photochemical reaction in the ground state and excited state and (e and f) their intrinsic reaction coordinate.

Table 5 Structural parameters of the transition state of uracil base reaction in S0 and S1

State	C2–C4 (Å)	C2–C14 (Å)	C4–C16 (Å)	C14–C16 (Å)	Energy (Hartree)
S0	1.462	2.145	2.116	1.398	−829.578
S1	1.529	1.687	1.726	1.502	−829.522

transition-state structures in these two states. As shown in Fig. 2(a)–(d), it can be observed that the actual transition-state structures in S0 and S1 are similar to their initial guess structures, which means that the initial guess structures are reasonable. The intrinsic reaction coordinate of the transition-state structures in S0 and S1 further confirms the rationality of the predicted transition-state structures. In Fig. 2(e) and (f), it can be observed that there is only one intermediate activated complex point during the photochemical reaction, which corresponds to the transition-state structure of dual-uracil with the highest total energy in the reaction process. Meanwhile, the transition state contributes to the decomposition of the activation complex, forming a cyclobutane uracil dimer. As listed in Table 5, we find that the C–C double bonds (C2═C4 and C14═C16) of uracil molecules are destroyed to form C–C single bonds (C2–C4 and C14–C16). Simultaneously, the unpaired electrons of C2 and C4 atoms prefer to bind to those of C14 and C16 atoms to form the cyclobutane ring of the uracil dimer. Furthermore, we also observe that the length of C2–C14 and C4–C16 bonds of the transition state in S1 is smaller than that in S0, and the total energy in S1 is higher than that in S0. Therefore, the reaction of uracil molecules in the excited state is easier than that in the ground state. The length of C2–C4 and C14–C16 bonds in the S0 and S1 states also further demonstrates this result because the length of C2–C4 and C14–C16 bonds in S0 is larger than that in S1. This indicates that the strength of C2–C4 and C14–C16 bonds in S0 is stronger than those in S1.

3.4 Reaction rate constant of dual-uracil molecules

In order to determine the reaction rate constant of uracil molecules in S0 and S1 states, the corresponding thermodynamic quantities with thermal correction are calculated through Gibbs free energy. Table 6 clearly presents the thermodynamic quantities with thermal correction of uracil molecules with S0 and S1, including internal energy (U), enthalpy (H), and Gibbs free energy (G). In particular, G is the most important parameter for calculating the reaction rate constant. As listed in Table 6, for S0, the G values of single and dual uracil are −414.788 a.u. and −829.391 a.u., respectively, whereas they are −414.771 a.u. and −829.536 a.u. for S1, respectively. Therefore, $G^{0,\neq}$ for S0 and S1 turn out to be 0.131 a.u. (344.651 kJ mol^{-1}) and 5.331 × 10^{-3} a.u. (13.997 kJ mol^{-1}), respectively. Eventually, the reaction rate constant of uracil molecules in S0 is 6.419 × 10^{-49} s^{-1} M^{-1}, while it is 5.436 × 10^{11} s^{-1} M^{-1} for S1.[66] For the chemical reaction at room tempearture, k_{TST} = 3.33 × 10^{-3} s^{-1} M^{-1} has been a criterion for determining the ease of between molecules, which is derived from $t_{1/2}$= 1/(k[A]$_0$). In this half-life formula, $t_{1/2}$ and [A]$_0$ are assumed as five minutes and 1 M, respectively. When the reaction rate constant is larger than 3.33 × 10^{-3} s^{-1} M^{-1}, the reactions between molecules are preferably easier and faster. Unsurprisingly, the reaction rate constant of uracil molecules in S1 is much larger than 3.33 × 10^{-3} s^{-1} M^{-1}, which indicates that uracil molecules react easily to produce uracil–uracil dimers in the excited states. By contrast, the reaction is extremely difficult in the ground state due to a very low reaction rate constant. The calculated half-life further demonstrates this result. Depending on the half-life formula, the half-life of uracil in S0 is 1.56 × 10^{48} s, which suggests that the molecular reaction could barely occur. However, the half-life of uracil in S1 is calculated to be 1.84 × 10^{-12} s (*i.e.*, 1.84 ps); thus transition of uracil in S1 is a transient reaction.[67] As a result, we can reach to a conclusion that photochemical reactions occur between nucleotides under UV light conditions. Besides, the absorption of photon energies by nucleotides changes their molecular structures from the ground state to the excited states, lowering the reaction potential barrier and thus allowing the molecular reactions to occur more easily.

Table 6 Thermodynamic quantities with thermal correction of single and dual uracil in S0 and S1

State	Thermodynamic quantity	Single-uracil (a.u.)	Dual-uracil TS (a.u.)
S0	U	−414.751	−829.391
	H	−414.750	−829.391
	G	−414.788	−829.444
S1	U	−414.732	−829.488
	H	−414.731	−829.487
	G	−414.771	−829.536

4 Conclusions

In summary, by means of applying a quantum chemical method, the underlying inactivation mechanism of UVC irradiation at the molecular level is theoretically revealed. Four main nucleotides, forming the genetic materials of viruses, are chosen as the main study objects. According to the calculated absorption spectra, it is found that the strongest absorption peaks are at 248.7 nm, 226.1 nm (252.7 nm), 248.3 nm, and 205.8 nm (249.2 nm) for adenine, cytosine, guanine, and uracil nucleotides, respectively. Eventually, depending on the calculation of the transition state and reaction rate constant, we find that the molecular reaction for adjacent nucleotides is a photochemical process under UVC light. The reaction rate constant at the ground state is 6.419 × 10^{-49} s^{-1} M^{-1}, whereas for the excited state it is 5.436 × 10^{11} s^{-1} M^{-1}. Their corresponding half-life values are 1.56 × 10^{48} s and 1.84 × 10^{-12} s, respectively. Therefore, the reaction under UVC irradiation easily occurs and that without UVC irradiation is extremely difficult.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the National Key R&D Program of China (2018YFE0204600) and the Shenzhen Fundamental Research Program (JCYJ20200109140822796).

References

1 J. M. Hughes, *Trans. Am. Clin. Climatol. Assoc.*, 2004, **115**, 361.
2 C. T. Bauch and T. Oraby, *Lancet*, 2013, **382**, 662–664.
3 D. Malvy, A. K. McElroy, H. de Clerck, S. Günther and J. van Griensven, *Lancet*, 2019, **393**, 936–948.
4 G. Neumann, T. Noda and Y. Kawaoka, *Nature*, 2009, **459**, 931–939.
5 S. F. Pedersen and Y.-C. Ho, *J. Clin. Invest.*, 2020, **130**, 2202–2205.
6 G. Xiling, C. Yin, W. Ling, W. Xiaosong, F. Jingjing, L. Fang, Z. Xiaoyan, G. Yiyue, C. Ying and C. Lunbiao, *Sci. Rep.*, 2021, **11**, 1–9.
7 I. Zucker, Y. Lester, J. Alter, M. Werbner, Y. Yecheskel, M. Gal-Tanamy and M. Dessau, *Environ. Chem. Lett.*, 2021, **19**, 1779–1785.
8 T. Pottage, C. Richardson, S. Parks, J. Walker and A. Bennett, *J. Hosp. Infect.*, 2010, **74**, 55–61.
9 C. S. Heilingloh, U. W. Aufderhorst, L. Schipper, U. Dittmer, O. Witzke, D. Yang, X. Zheng, K. Sutter, M. Trilling and M. Alt, *Am. J. Infect. Control*, 2020, **48**, 1273–1275.
10 C. Meyers, J. Milici and R. Robison, *PLoS One*, 2017, **12**, e0187377.
11 A. Sakudo, Y. Yagyu and T. Onodera, *Int. J. Mol. Sci.*, 2019, **20**, 5216.
12 T. Furukawa, T. Ueno, M. Matsumura, M. Amarasiri and K. Sei, *J. Hazard. Mater.*, 2022, **424**, 127382.
13 J. Hadi, M. Dunowska, S. Wu and G. Brightwell, *Pathogens*, 2020, **9**, 737.
14 M. M. Delorme, J. T. Guimarães, N. M. Coutinho, C. F. Balthazar, R. S. Rocha, R. Silva, L. P. Margalho, T. C. Pimentel, M. C. Silva and M. Q. Freitas, *Trends Food Sci. Technol.*, 2020, **102**, 146–154.
15 P. M. F. Maciel, N. d M. N. Fava, A. W. Lamon, P. Fernandez-Ibañez, J. A. Byrne and L. P. Sabogal-Paz, *J. Water Process Eng.*, 2021, **43**, 102203.
16 P.-Y. Lai, H. Liu, R. J. H. Ng, B. Wint Hnin Thet, H.-S. Chu, J. W. R. Teo, Q. Ong, Y. Liu and C. E. Png, *Sci. Rep.*, 2021, **11**, 1–10.
17 S. K. Bhardwaj, H. Singh, A. Deep, M. Khatri, J. Bhaumik, K.-H. Kim and N. Bhardwaj, *Sci. Total Environ*, 2021, **792**, 148548.
18 F. P. Sellera, C. P. Sabino, F. V. Cabral and M. S. Ribeiro, *J. Photochem. Photobiol.*, 2021, **8**, 100068.
19 M. Buonanno, D. Welch, I. Shuryak and D. J. Brenner, *Sci. Rep.*, 2020, **10**, 1–8.
20 K. Narita, K. Asano, Y. Morimoto, T. Igarashi and A. Nakane, *PLoS One*, 2018, **13**, e0201259.
21 B. M. Tande, T. A. Pringle, W. A. Rutala, M. F. Gergen and D. J. Weber, *Infect. Control Hosp. Epidemiol.*, 2018, **39**, 1122–1124.
22 S. Ali, S. Yui, M. Muzslay and A. Wilson, *J. Hosp. Infect.*, 2017, **97**, 180–184.
23 K. Song, F. Taghipour and M. Mohseni, *Sci. Total Environ.*, 2019, **665**, 1103–1110.
24 K. Oguma, K. Kanazawa, I. Kasuga and S. Takizawa, *J. Photochem. Photobiol.*, 2018, **94**, 570–576.
25 M. Heßling, K. Hönes, P. Vatter and C. Lingenfelder, *GMS hyg. Infect. Control*, 2020, **15**, Doc08.
26 S.-J. Kim, D.-K. Kim and D.-H. Kang, *Appl. Environ. Microbiol.*, 2016, **82**, 11–17.
27 K. Narita, K. Asano, K. Naito, H. Ohashi, M. Sasaki, Y. Morimoto, T. Igarashi and A. Nakane, *J. Hosp. Infect.*, 2020, **105**, 459–467.
28 K. Narita, K. Asano, Y. Morimoto, T. Igarashi, M. R. Hamblin, T. Dai and A. Nakane, *J. Photochem. Photobiol., B*, 2018, **178**, 10–18.
29 M. B. Lore, B. K. Heimbuch, T. L. Brown, J. D. Wander and S. H. Hinrichs, *Ann. Occup. Hyg.*, 2012, **56**, 92–101.
30 S. Pi, L. R. Seng-Libi and D. Xao-Ping, *Biomed. Environ. Sci.*, 2003, **16**, 246–255.
31 H. Inagaki, A. Saito, H. Sugiyama, T. Okabayashi and S. Fujimoto, *Emerging Microbes Infect.*, 2020, **9**, 1744–1747.
32 F. Bosshard, F. Armand, R. Hamelin and T. Kohn, *Appl. Environ. Microbiol.*, 2013, **79**, 1325–1332.
33 R. Nishisaka-Nonaka, K. Mawatari, T. Yamamoto, M. Kojima, T. Shimohata, T. Uebanso, M. Nakahashi, T. Emoto, M. Akutagawa and Y. Kinouchi, *J. Photochem. Photobiol., B*, 2018, **189**, 193–200.
34 E. Araud, M. Fuzawa, J. L. Shisler, J. Li and T. H. Nguyen, *Appl. Environ. Microbiol.*, 2020, **86**, e02436–02419.
35 T. Tanaka, O. Nogariya, N. Shionoiri, Y. Maeda and A. Arakaki, *J. Biosci. Bioeng.*, 2018, **126**, 63–68.
36 D. Sano, R. M. Pinto, T. Omura and A. Bosch, *Environ. Sci. Technol.*, 2010, **44**, 808–812.
37 K. Rule Wigginton, L. Menin, J. P. Montoya and T. Kohn, *Environ. Sci. Technol.*, 2010, **44**, 5437–5443.
38 D. Park, H. M. Shahbaz, S.-H. Kim, M. Lee, W. Lee, J.-W. Oh, D.-U. Lee and J. Park, *Int. J. Food Microbiol.*, 2016, **238**, 256–264.
39 K. Wetz and K.-O. Habermehl, *J. Gen. Virol.*, 1982, **59**, 397–401.
40 Y. Nakaya, T. Fukuda, H. Ashiba, M. Yasuura and M. Fujimaki, *BMC Infect. Dis.*, 2020, **20**, 1–10.
41 B. Milovanović, J. Novak, M. Etinski, W. Domcke and N. Došlić, *Phys. Chem. Chem. Phys.*, 2021, **23**, 2594–2604.
42 W. Kowalski, *Ultraviolet germicidal irradiation handbook: UVGI for air and surface disinfection*, Springer science & business media, 2010.
43 R. D. Tosso, M. N. C. Zarycz, A. Schiel, L. Goicoechea Moro, H. A. Baldoni, E. Angelina and R. D. Enriz, *J. Comput. Chem.*, 2022, **43**, 1298–1312.
44 X. Zou, X. Dai, K. Liu, H. Zhao, D. Song and H. Su, *J. Phys. Chem. B*, 2014, **118**, 5864–5872.
45 T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580–592.
46 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.

47 T. Lu and Q. Chen, *Comput. Theor. Chem. Acc.*, 2021, **1200**, 113249.
48 A. Fernández-Ramos, B. A. Ellingson, R. Meana-Pañeda, J. Marques and D. G. Truhlar, *Theor. Chem. Acc.*, 2007, **118**, 813–826.
49 A. Fernandez-Ramos, B. A. Ellingson, B. C. Garrett and D. G. Truhlar, *Rev. Comput. Chem.*, 2007, **23**, 125.
50 P. Ptáček, F. Šoukal and T. Opravil, Introducing the Effective Mass of Activated Complex and the Discussion on the Wave Function of this Instanton, *Chemical Kinetics*, IntechOpen, 2018, vol. 27, DOI: **10.5772/intechopen.70734**.
51 P. F. Bernath, *Spectra of atoms and molecules*, Oxford university press, 2020.
52 G. D. Harris, V. D. Adams, D. L. Sorensen and M. S. Curtis, *Water Res.*, 1987, **21**, 687–692.
53 T. D. Cutler and J. Zimmerman, *Anim. Health Res. Rev.*, 2011, **12**, 15–23.
54 D. Battigelli, M. Sobsey and D. Lobe, *Water Sci. Technol.*, 1993, **27**, 339–342.
55 R. B. Setlow, *Science*, 1966, **153**, 379–386.
56 P. K. Samanta, A. K. Manna and S. K. Pati, *J. Phys. Chem. B*, 2012, **116**, 7618–7626.
57 X. Wang, Z. Liu, X. Yan, T. Lu, H. Wang, W. Xiong and M. Zhao, *Phys. Chem. Chem. Phys.*, 2022, **24**, 7466–7473.
58 S. Chen, N. Ullah and R. Zhang, *J. Phys. Chem. Lett.*, 2018, **9**, 4857–4864.
59 X. Tang, L.-S. Cui, H.-C. Li, A. J. Gillett, F. Auras, Y.-K. Qu, C. Zhong, S. T. Jones, Z.-Q. Jiang and R. H. Friend, *Nat. Mater.*, 2020, **19**, 1332–1338.
60 Y. Jiao, J. Zhu, Y. Guo, W. He and Z. Guo, *J. Mater. Chem. C*, 2017, **5**, 5214–5222.
61 R. E. Rycyna and J. L. Alderfer, *Biochemistry*, 1988, **27**, 3142–3151.
62 R. Beukers, A. P. Eker and P. H. Lohman, *DNA Repair*, 2008, **7**, 530–543.
63 W. Kladwang, J. Hum and R. Das, *Sci. Rep.*, 2012, **2**, 1–7.
64 R. L. Miller and P. G. Plagemann, *J. Virol.*, 1974, **13**, 729–739.
65 C.-W. Lo, R. Matsuura, K. Iimura, S. Wada, A. Shinjo, Y. Benno, M. Nakagawa, M. Takei and Y. Aida, *Sci. Rep.*, 2021, **11**, 1–11.
66 T. Lu, **http://sobereva.com/310** (accessed 08/21/2021).
67 T. Lu, **http://sobereva.com/506** (accessed 08/21/2021).