Mechanism for dissociative hydrolysis of pyrimidine nucleoside d4N: inversion vs retention

Yi Zeng[1], Yuexin Zhao[1], Wei Zhang[1], and Yang Jiang[2]

[1]Xihua University
[2]Neijiang Normal University

October 10, 2022

Abstract

Employing the microhydration model that involves two to five explicit water molecules, two plausible dissociative hydrolysis pathways of 2',3'-didehydro-2',3'-dideoxyuridine (d4U), α-path with configuration-inversion and β-path with configuration-retention, have been investigated by M06-2X(CPCM)/6-31++G(d,p) method. Using this model, inclusion of three explicit water molecules (n=3) is shown to be the smallest system that gives the minimal activation free energy for α-path and β-path. Our results suggest that the glycoside cleavage is the RDS, and α-path is more favorable kinetically than β-path. Whereas β-path with exergonic formation of β-dihydrofuran-like sugar with keto pyrimidine complex possesses thermodynamic preference over α-path, where the formation of α-dihydrofuran-like sugar with enol pyrimidine complex is endergonic. The free energy barriers of RDSs for d4U (24.8 kcal mol^{-1}) and d4T (27.3 kcal mol^{-1}) suggest that the glycosidic bond in d4T is more stable than that in d4U. The relative lower stability of d4U is probably an important factor for less antiviral activity of d4U. The small free energy barrier differences of ˜1 kcal mol^{-1} for β-path over α-path, and the reaction free energy differences of ˜ -12 kcal mol^{-1} for β-path lower than α-path in d4T and d4U suggest a competitive β-path in pyrimidine d4Ns. The higher free energy barriers of RDSs in ddU (27.6 kcal mol^{-1}) and ddT (29.0 kcal mol^{-1}) indicate that the saturated sugar moiety increases the stability of glycosidic bond in contrast to the unsaturated counter parts in d4U and d4T. NBO analysis also shows the kinetic preference of α-path over β-path. Our results provide an exploration for the less antiviral activity of d4U and the influence of saturated ribose on the glycosidic bond stability of pyrimidine d4Ns.

Mechanism for dissociative hydrolysis of pyrimidine nucleoside d4N: inversion vs retention

Yi Zeng[a]*, Yuexin Zhao[a], Wei Zhang[a], Yang Jiang[b]*

[a]*School of Science, Key Laboratory of High Performance Scientific Computation, Xihua University, Chengdu, China 610039*

zengy@mail.xhu.edu.cn

[b]*College of Chemistry and Chemical Engineering, Neijiang Normal University, Neijiang, China 641100*

jiangyang@njtc.edu.cn

Abstract: Employing the microhydration model that involves two to five explicit water molecules, two plausible dissociative hydrolysis pathways of 2',3'-didehydro-2',3'-dideoxyuridine (d4U), α-path with configuration-inversion and β-path with configuration-retention, have been investigated by M06-2X(CPCM)/6-31++G(d,p) method. Using this model, inclusion of three explicit water molecules (n=3) is shown to be the smallest system that gives the minimal activation free energy for α-path and β-path. Our results suggest that the glycoside cleavage is the RDS, and α-path is more favorable kinetically than β-path. Whereas β-path with exergonic formation of β- dihydrofuran-like sugar with keto pyrimidine complex possesses thermodynamic preference over α-path, where the formation of α-dihydrofuran-like sugar with enol

pyrimidine complex is endergonic. The free energy barriers of RDSs for d4U (24.8 kcal mol^{-1}) and d4T (27.3 kcal mol^{-1}) suggest that the glycosidic bond in d4T is more stable than that in d4U. The relative lower stability of d4U is probably an important factor for less antiviral activity of d4U. The small free energy barrier differences of ˜1 kcal mol^{-1} for β-path over α-path, and the reaction free energy differences of ˜ -12 kcal mol^{-1} for β-path lower than α-path in d4T and d4U suggest a competitive β-path in pyrimidine d4Ns. The higher free energy barriers of RDSs in ddU (27.6 kcal mol^{-1}) and ddT (29.0 kcal mol^{-1}) indicate that the saturated sugar moiety increases the stability of glycosidic bond in contrast to the unsaturated counter parts in d4U and d4T. NBO analysis also shows the kinetic preference of α-path over β-path. Our results provide an exploration for the less antiviral activity of d4U and the influence of saturated ribose on the glycosidic bond stability of pyrimidine d4Ns.

KEYWORDS: hydrolysis, glycosidic bond, nucleoside, microhydration model, d4N

INTRODUCTION

2',3'-Didehydro-2',3'-dideoxynucleosides (d4Ns)[1-3] with a double bond in the furanose lacking the 2' and 3'-hydroxyl group exert their anti-human immunodeficiency virus (HIV) activity by the inhibition of reverse transcriptase (RT) and the termination of viral replication.[4-8] Particularly when the pyrimidine base is thymine, 2',3'-didehydro-2',3'-dideoxythymidine (d4T, Stavudine) as an approved drug constitutes an essential part of combined anti-retroviral therapy.[9,10] However, 2',3'-didehydro-2',3'-dideoxyuridine (d4U), different only in the absence of methyl group at the 5-position of d4T, pronounces less antiviral activity.[11,12] Consequently, earlier works were devoted to the relationship between conformational properties and the biological activity of d4T and d4U.

The important torsional angles χ(C2-N1-C1'-O4') around the glycosidic bond, γ(C3'-C4'-C5'-O5') and β(C4'-C5'-O5'-H5') orienting the hydroxyl on the 5'-end of the sugar residue, are labelled in Scheme 1. The X-ray structural studies of d4T and d4U revealed two conformations with similar χ-γ values: -118.0/60.6° and -174.0/53.8° for d4T,[13] versus -96.1(5)/53.2° and -178.0(4)/54.1(6)° for d4U.[14] A later high-resolution X-ray analysis of d4T determined the χ values to be -102.1(3)° and -117.2(4).[15]Hovorun group performed reliable quantum-chemical calculations on the stable conformers of d4T[16] and d4U[17]in gas phase. They reported that the most populated d4T and d4U conformers possessed very close torsion angles: χ(-103.4/-106.4°)[?]*anti* , γ(60.2/62.2°)[?]g^+ and β(68.8/66.2°)[?]g^+ . Both the experimental and theoretical investigations manifested that the d4T and d4U are very similar from a conformational point of view.

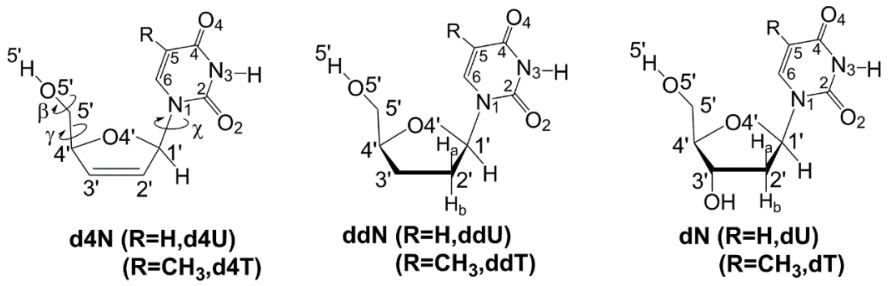
It is well known that the stability of glycosidic bond is another crucial factor to the biological activity of d4Ns.[18-20] Shi and coworkers found by reverse-phase high-performance liquid chromatography (HPLC) that the glycosidic bond of d4T would cleave with half-life ($t_{1/2}$) 35 days under acidic condition.[21] However, less information involving the transition states and energies for the glycosidic bond hydrolysis of d4T and other 2',3'-didehydro-2',3'-dideoxypyrimidines, are hitherto available from theoretical studies. Wetmore group engaged in the hydrolysis of natural 2'-deoxypyrimidines by employing microsolvated model with explicit water molecules (n = 1-3).[22,23] Their computational studies at PCM-B3LYP/6-31+G(d) level manifested the trend of the Gibbs activation energies for 2'-deoxythymidine (dT) and 2'-deoxyuridine (dU) as dT < dU, suggesting slightly weaker glycosidic bond of dT in comparison with dU.

Herein, we would examine the glycosidic bond hydrolysis of d4T and d4U to determine if the stability of glycosidic bond can be useful in understanding their different biological activity. This study parallels our earlier microsolvated calculation of the dissociative hydrolysis of 2',3'-didehydro-2',3'-dideoxyguanosine (d4G).[24] In this study, the microhydration model used in our previous studies for d4G and oxazolin-5-ones[25,26] is applied for d4U, which has the smallest pyrimidine base[27,28] in this group of molecules. With two to five explicit water molecules (n=2-5), the minimal microhydration model for the hydrolysis pathway of an extensive series of pyrimidine d4Ns and analogs has been found.

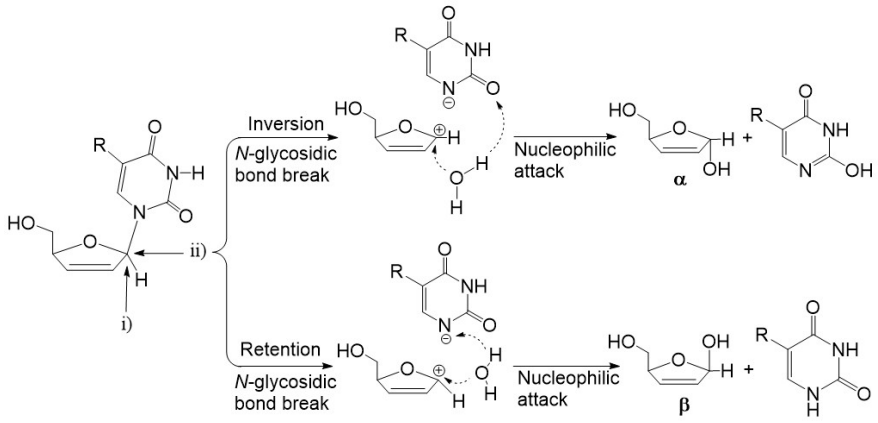
The two plausible hydrolysis paths: α-path and β-path, are indicated in Scheme 2. The dissociative hydrolysis paths experience the glycosidic bond breakage for the generation of oxacarbenium cation, and subsequent nucleophilic attack of water molecule on the oxacarbenium cation via positions (i) or (ii) corresponding to

inversion (α-path) or retention (β-path) of the C1' configuration, respectively. For nucleoside analogues, most of the published literatures focused on the inversion attack; in contrast, little attention was paid to the retention attack.[29]

In addition, to provide a better insight into the double bond in sugar moiety for the glycosidic bond stability of d4Ns, the saturated 2',3'-dideoxynucleosides (ddNs) including 2',3'-dideoxyuridine (ddU) and 2',3'-dideoxythymidine (ddT) among the potent anti-HIV agents,[30,31] are also examined as comparison using our simplistic microhydration model (Scheme 1).



Scheme 1. Structure and definition of the torsional angles in d4N and ddN.



Scheme 2. Orientations for inversion (**α** -path) and retention (**β** -path) hydrolysis.

COMPUTATIONAL DETAILS

Geometries of stationary points were obtained in water by a microhydrated clusters together with the conductor-like polarizable continuum model (CPCM)[32,33] at M06-2X[34,35]/6-31++G(d,p) computational level. These stationary points containing transition states (TS), reactant complexes (RC), product complexes (PC), and intermediates (IN) are validated by vibrational frequency calculations. This method has given excellent performance for thermochemistry and barrier height in our published hydrolysis of d4G system.[24] Intrinsic reaction coordinate (IRC)[36] calculations were used to confirm the connectivity of the stationary points in the potential energy profiles (PEP). Natural bond orbital (NBO)[37]calculation was employed for charge distributions and donor-acceptor interactions analysis. All calculations were based on the Gaussian09 program.[38]

RESULTS AND DISCUSSION

The nomenclature to identify each transition state and complex with three explicit water as model is presented in Figure 1. In α-path, the O_{w1} of the nucleophilic water attacks C1' from its backside, together with the

transfer of H_{w1} through proton shuttle: H_{w1} to O_{w2}, H_{w3} to O_{w3}, and eventually H_{w5} reaching to O2. In β-path, the nucleophilic O_{w1} attacks C1' from its frontside coupled with a direct shift of H_{w2} to N1, while the other two water molecules just act as explicit solvent to stabilize transition state. The geometrical parameters and charge distributions, as well as donor-acceptor interactions analysis of the key stationary points calculated in these ways, are listed in Table 1 and 2, respectively.

The computational models with one explicit water for α-path and one to two explicit water for β-path are not discussed for their inadequacy to describe a continuous reaction surface of d4U after extensive endeavors.

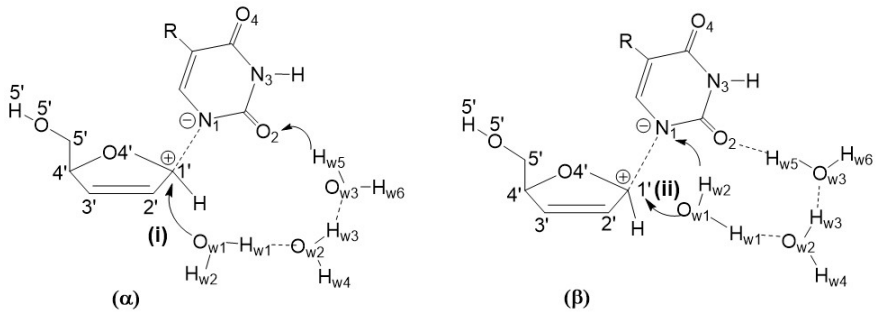


FIGURE 1 Nomenclature for the transition states in inversion (**α**) and retention (**β**) pathways with three explicit water.

3.1 α-Πατη οφ δ4Υ ωιτη τωο το φουρ εξπλιςιτ ωατερ μολεςυλες (ν = 2-4).

Two to three explicit water molecules (n = 2 and 3).

For α-path of d4U with two to three explicit water molecules (n = 2 and 3), the obtained free energy profiles with the optimized geometries are illustrated in Figure 2 and 3. In both cases, the dissociative hydrolysis of d4U manifests similar PEP with a rate-determining step (RDS) of glycosidic bond cleavage.

With two extra explicit water molecules, the oxacarbenium cation showed apparent higher stability compared to only one explicit water, and the activation free energy of RDS and the overall free energy barrier decreases from 27.7 and 29.9 kcal mol^{-1} (n=2) to 24.8 and 27.2 kcal mol^{-1} (n=3), respectively. For RDS from n=2 to 3 cases, the decrease of 2.9 kcal mol^{-1} in the activation free energy can be ascribed to the less twisted angle [?]N1C1'O_{w1} with 156.0° in**αΥ-TΣ1$_{3ω}$** relative to that 148.0° in **αΥ-TΣ1$_{2ω}$** . Meanwhile, the smaller distance change with 0.92 Å of the C1'-N9 (from 1.48 Å in **αΥ-P$^{\sim}_{3ω}$** to 2.40 Å in**αΥ-TΣ1$_{3ω}$**), compared to the corresponding value of 1.03 Å (from 1.47 Å in **αΥ-P$^{\sim}_{2ω}$**) to 2.50 Å in**αΥ-TΣ1$_{2ω}$**), also contributes to the activation free energy decrease. Thus, in this section, we would mainly focus on the case of n=3.

In contrast to the X-ray structural data of d4U,[14] the inclusion of water clusters around the C1' and uridine O2 makes d4U adopt a significant high-anti base orientation, with a glycosyl link χ value to be -76.6° for**α-P$^{\sim}_{3ω}$** . Similar effect of different water clusters on glycosyl torsional angle in d4T was reported by Palafox and co-workers.[39] The γ-β values, γ(59.9°)[?]g^+ and β(63.8°)[?]g^+ , are close to the experimental data.[14]

In **αΥ-TΣ1$_{3ω}$** , the breaking glycosidic bond C1'-N1 presents an striking second-order perturbation stabilization energy, $E^{(2)}$ value of 34.21 kcal mol^{-1} for n(N1)-π*(C1'-O4') (Table 2). The shorter C1'-O4' (1.28 Å) and C1'-C2' (1.44 Å) distances together with the positive charge +0.44 on C1' atom suggest an sp^2-hybridization character of C1', coupled with a great $E^{(2)}$ of 51.78 kcal mol^{-1} for π(C2'-C3')-π*(C1'-O4'). The generated oxacarbenium cation **αΥ-IN$_{3ω}$** bears a greater positive charge of +0.46, and is also energetically better stabilized, as shown sitting in a well with a depth of 1.5 kcal mol^{-1} on the free energy profile, in contrast to the corresponding well-depth of 0.8 kcal mol^{-1} for**αΥ-IN$_{2ω}$** . In the second stage of nucleophilic attack of water, **αΥ-TΣ2$_{3ω}$** displays a strong interaction of LP(O_{w1})-π*(C1'-O4') with an$E^{(2)}$ estimate of 70.36 kcal mol^{-1} and a decrease of the negative charge of O_{w1} from -1.06 to -0.95. The formation of**αΥ-**

$\Pi^{*}_{3\omega}$ by α- dihydrofuran-like sugar interacting with an enol tautomer of uridine is a thermodynamically disadvantageous pathway with a reaction free energy of 6.9 kcal mol^{-1}.

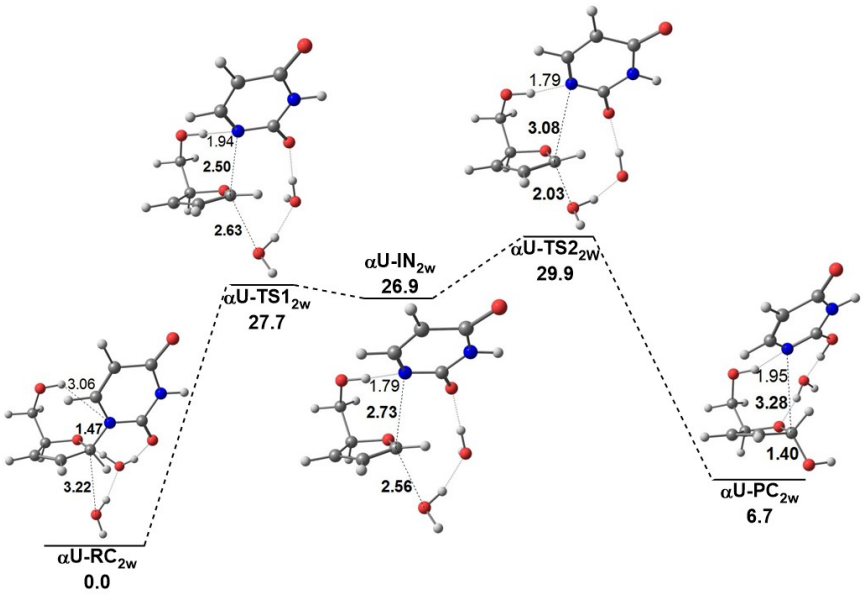


FIGURE 2 Free energy profile (kcal mol^{-1}) and optimized geometries of species in the inversion hydrolysis of d4U (α-path) with two explicit water molecules.

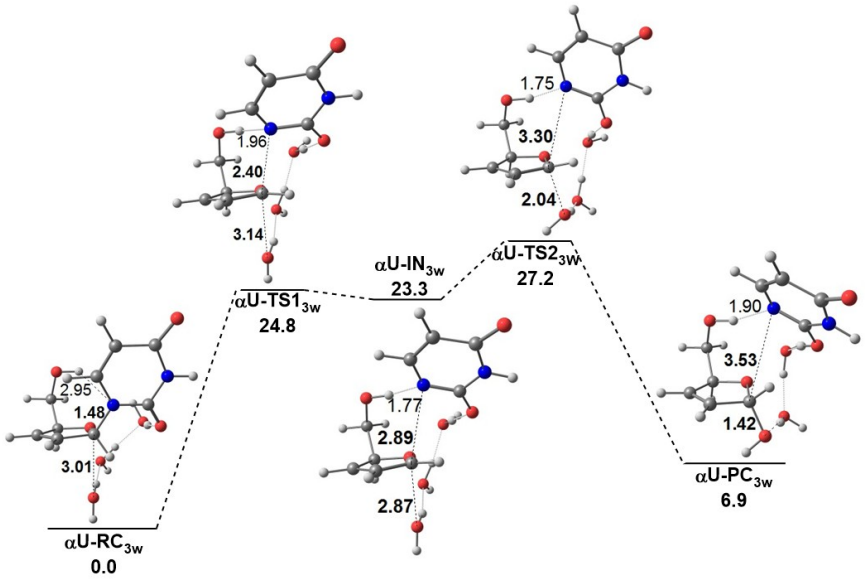


FIGURE 3 Free energy profile (kcal mol^{-1}) and optimized geometries of species in the inversion hydrolysis of d4U (α-path) with three explicit water molecules.

Four explicit water molecules (n = 4).

After considering the participation of the fourth water molecule, a more complicated PEP involving three

stages is predicted in Figure 4. Besides the two steps of glycosidic bond breakage (RDS) and nucleophilic attack of water, a third stage takes place where H_{w1} transfer experiences a hydrated proton process,**αϒ-IN2$_{4ω}$-αϒ-TΣ3$_{4ω}$** .

The activation free energy of RDS (24.9 kcal mol^{-1}) and overall free energy barrier (27.1 kcal mol^{-1}) in the former two stages are close to those in case of n=3. The**αϒ-IN1$_{4ω}$** lies in a shallow well of 0.3 kcal mol^{-1} on the free energy profile, suggesting the destabilizing effect of the extra water molecule for the oxacarbenium cation. In the third stage of H_{w1} transfer, the proton transfer from **αϒ-IN2$_{4ω}$** to**αϒ-TΣ3$_{4ω}$** is almost a barrier free process (9.1 vs. 9.4 kcal mol^{-1} in electronical energy barrier). In addition, the inclusion of the fourth water leads to an increase of 1.1 kcal mol^{-1} in the reaction free energy (8.0 in**αϒ-Π‵$_{4ω}$** vs. 6.9 kcal mol^{-1}**αϒ-Π‵$_{3ω}$**). Thereby, the microhydration model with three explicit water molecules is determined to be the smallest model for inversion hydrolysis mechanism of pyrimidine d4Ns. Such model was used for dU and dT, with the structures of transition states presented in Figure S1. At M06-2X(CPCM)/6-31++G(d,p) level, the activation free energies of glycosidic bond cleavage are 29.7 and 29.1 kcal mol^{-1} for dU and dT, respectively. The value of 29.7 kcal mol^{-1} for dU agrees well with the reported experimental result $\Delta G^{[?]}$=30.5 kcal mol^{-1} of glycosidic breakage at pH 6.8 and 25°C.[40] The stability order of glycosidic bond predicted here as dT < dU also coincides with the reported theoretical result.[23] These corroborate the consistency of our calculation.

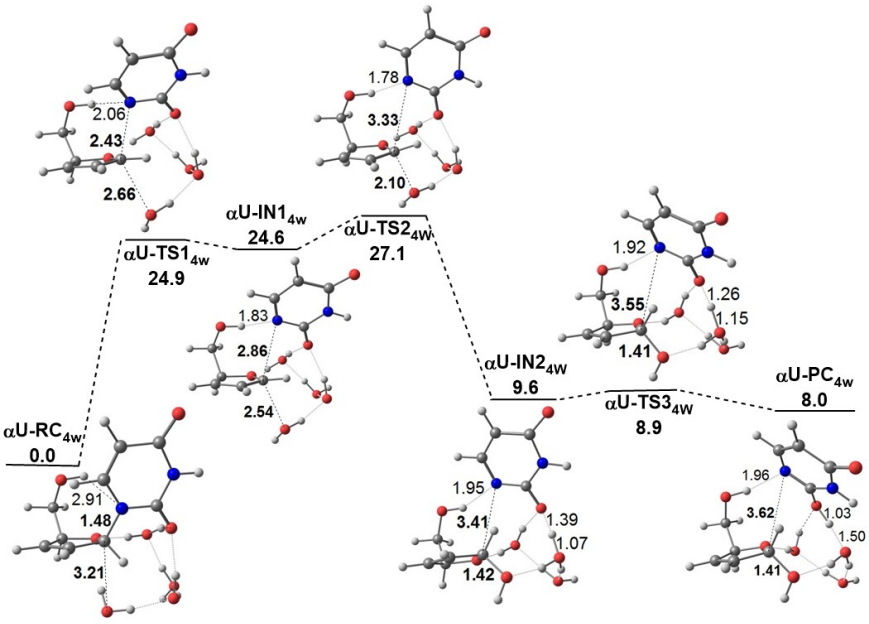


FIGURE 4 Free energy profile (kcal mol^{-1}) and optimized geometries of species sin the inversion hydrolysis of d4U (α-path) with four explicit water molecules.

3.2 β-Παθη οφ δ4ϒ ωιθη θηρεε το φιε εξπλιςιτ ωατερ μολεςυλες (ν = 3-5).

The microhydration model with three to five explicit water molecules for retention hydrolysis mechanism (β-path) is investigated in this section (Figure 5 to 7). Different from the one-step cleavage of glycosidic bond C1'-N1 with one transition state in α-path, the C1'-N1 cleavage in β-path puts through multiple steps, and more than one transition states are located in the PEPs. The rate-determining transition states for the C1'-N1 cleavage are **βϒ-TΣ1$_{3ω}$** ,**βϒ-TΣ1$_{4ω}$** and **βϒ-TΣ1$_{5ω}$**in the cases of n=3, 4 and 5.

3.2.1 Three Explicit Water Molecules (n = 3).

For n=3 case (Figure 5), the cleavage of C1'-N1 involves three transition states **βϒ-TΣ1$_{3ω}$** ,**βϒ-TΣ2$_{3ω}$**

and **βϒ-TΣ3$_{3ω}$**in which the activation free energy of rate-determining transition state**βϒ-TΣ1$_{3ω}$** is calculated to be 26.4 kcal mol^{-1}, higher than that 24.8 kcal mol^{-1} of **αϒ-TΣ1$_{3ω}$** . A smaller$E^{(2)}$ value of 32.39 kcal mol^{-1} for (N1)-π*(C1'-O4') in**βϒ-TΣ1$_{3ω}$** relative to the corresponding 34.21 kcal mol^{-1} in **αϒ-TΣ1$_{3ω}$**provides a theoretical support for the kinetical preference of α-path over β-path (Table 2). The other two $E^{(2)}$values for n(N1)-π*(C1'-O4') in **βϒ-TΣ2$_{3ω}$** and**βϒ-TΣ3$_{3ω}$** significantly decrease to 9.55 and 0.55 kcal mol^{-1}, with the lengthened C1'-N1 distances from 2.43 Å (**βϒ-TΣ1$_{3ω}$**) to 3.46 Å (**βϒ-TΣ3$_{3ω}$**). Simultaneously, the negative charges shift to N1 with an increment from -0.71 to -0.75 (Table 1).The departure of N1 causes an occurrence of double bond with sp^2 hybridization on C1', coupled with the great$E^{(2)}$ values for π(C2'-C3')-π*(C1'-O4') in the range of 51.32 to 61.88 kcal mol^{-1} for the above three transition states. Three intermediates**βϒ-IN1$_{3ω}$** , **βϒ-IN2$_{3ω}$** and**βϒ-IN3$_{3ω}$** with positive charge of about +0.45 at C1' are produced in the C1'-N1 cleavage process. For the final attack of water molecule, **βϒ-TΣ4$_{3ω}$** involves only 2.7 kcal mol^{-1} free energy barrier relative to**βϒ-IN3$_{3ω}$** . In **βϒ-TΣ4$_{3ω}$** , the movement of O_{w1} towards C1' corresponds an$E^{(2)}$ value for LP(O_{w1})-π*(C1'-O4') ascending to 54.43 kcal mol^{-1}, together with the decrease of the negative charge of O_{w1} to -0.98 and the C1' positive charge to +0.43. The conversion of sp^2 to sp^{3}hybridization on C1' leads to a significant diminution of$E^{(2)}$ value being 44.49 kcal mol^{-1} for π(C2'-C3')-π*(C1'-O4'). In addition, a keto uridine is formed after the shift of H_{w2} to N1. Different from the thermodynamical disadvantage of α- path, the formation of **βϒ-Π$_{3ω}$** is a thermodynamically favorable pathway with the reaction free energy of -5.7 kcal mol^{-1}.

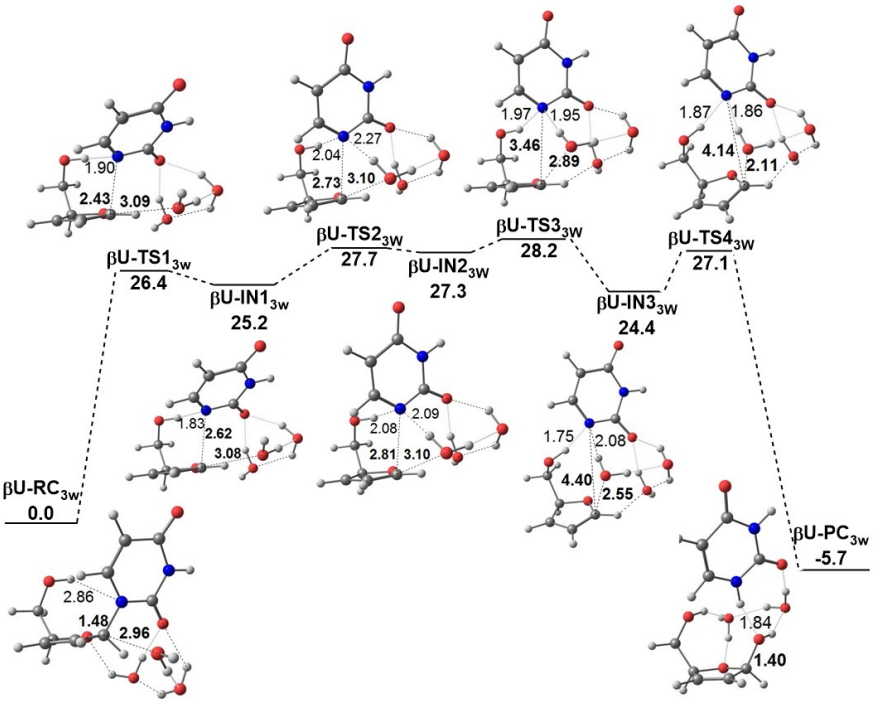


FIGURE 5 Free energy profile (kcal mol^{-1}) and optimized geometries of species in the inversion hydrolysis of d4U (β-path) with three explicit water molecules.

3.2.2 Four to five explicit water molecules (n = 4 and 5).

When the fourth and fifth extra water molecules are considered in the microhydration model, the C1'-N1 cleavages in β-path undergo two transition states, **βϒ-TΣ1$_{4ω}$** and**βϒ-TΣ2$_{4ω}$** vs. **βϒ-TΣ1$_{5ω}$**and **βϒ-TΣ2$_{5ω}$** .

For the n=4 case (Figure 6), the activation free energy of the first step of glycosidic bond dissociation

($\Delta G^{[?]}$= 26.3 kcal mol^{-1}) is almost unchanged, compared to that of n=3. The glycosidic bond cleavage takes place easier with a $\Delta G^{[?]}$=2.0 kcal mol^{-1}. The resultant intermediate $\beta\Upsilon$-$IN2_{4\omega}$ is located in a shallower well of 2.4 kcal mol^{-1} compared to corresponding$\beta\Upsilon$-$IN3_{3\omega}$ whose well is predicted to be 3.8 kcal mol^{-1}. It suggests the destabilizing of the fourth water to $\beta\Upsilon$-$IN2_{4\omega}$. Through the association of water to C1' with a small 2.5 kcal mol^{-1} free energy barrier relative to $\beta\Upsilon$-$IN2_{4\omega}$, the retention product $\beta\Upsilon$-$\Pi^{\curvearrowleft}_{4\omega}$ is produced and a similar destabilizing effect of the fourth water has also been found for$\beta\Upsilon$-$\Pi^{\curvearrowleft}_{4\omega}$ due to an increase of the reaction free energy of 2.8 kcal mol^{-1} (-2.9 in$\beta\Upsilon$-$\Pi^{\curvearrowleft}_{4\omega}$ vs. -5.7 kcal mol^{-1}$\beta\Upsilon$-$\Pi^{\curvearrowleft}_{3\omega}$).

For the n=5 case (Figure 7), the C1'-N1 cleavage and the nucleophilic attack of water stages resemble with those of n=4. Different from the above two cases where H_{w2} directly shifts to N1, a third stage here in which the H_{w1} transfers through a barrier free process of hydrated proton,$\beta\Upsilon$-$IN4_{5\omega}$-$\alpha\beta$-$T\Sigma5_{5\omega}$, produces an enol form of uridine with a reaction free energy of 4.9 kcal mol^{-1}. Besides, the fifth water causes a slight increase of activation free energy of rate-determining transition state$\beta\Upsilon$-$T\Sigma1_{5\omega}$ to 27.5 kcal mol^{-1}. Thus, from the kinetics and thermodynamics viewpoints, the smallest microhydration model for β-path is three explicit water molecules.

In general, both of the minimal microhydration model for α-path and β-path of d4U is with three explicit water molecules (n=3). It provides a better microhydration model for further investigation of α-path and β-path, which would be applied for the further investigation of these analogous.

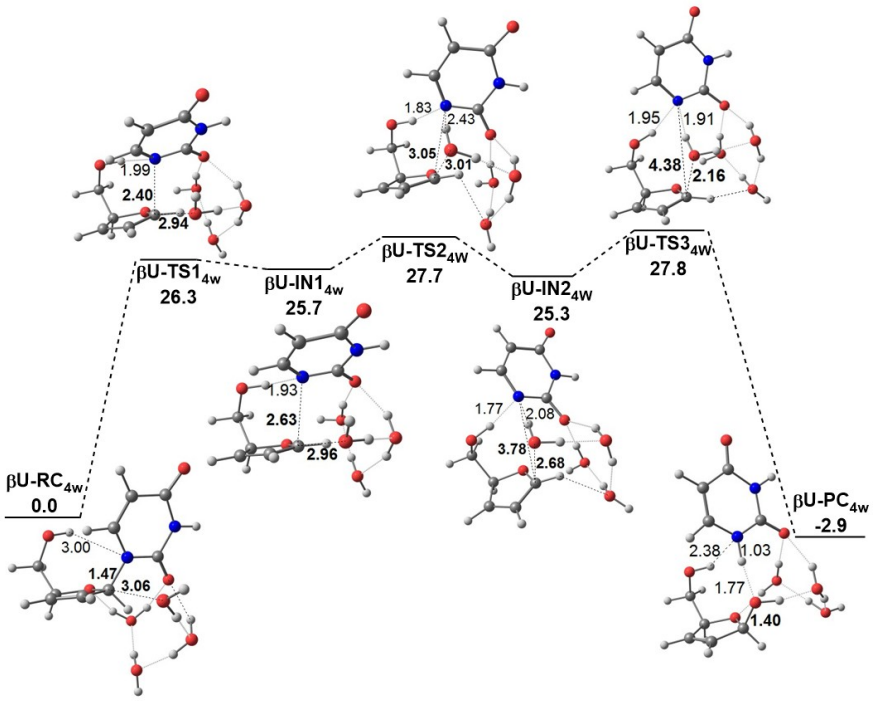


FIGURE 6 Free energy profile (kcal mol^{-1}) and optimized geometries of species in the inversion hydrolysis of d4U (β-path) with four explicit water molecules.

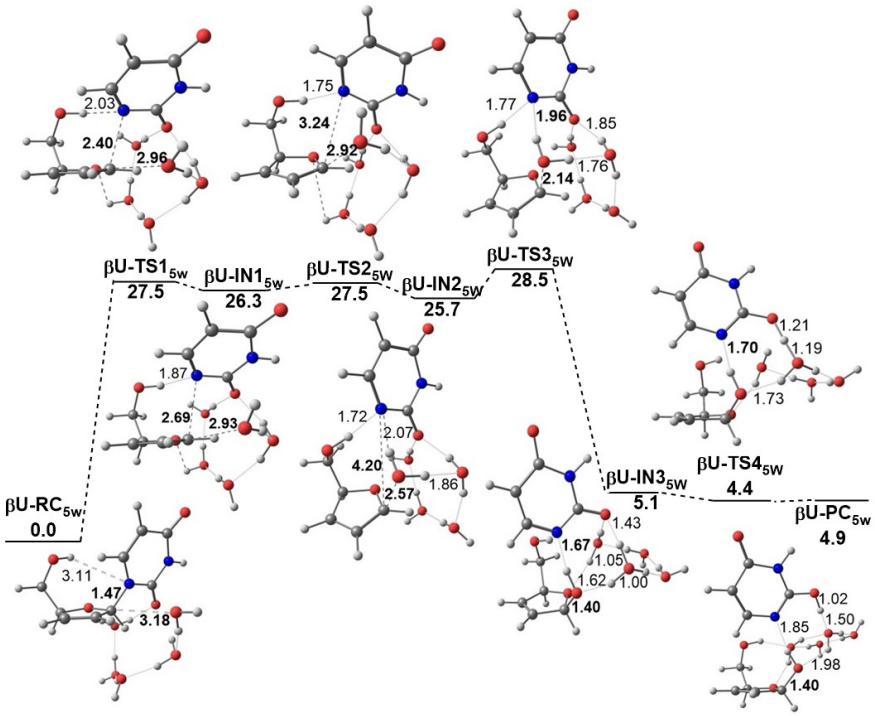


FIGURE 7 Free energy profile (kcal mol^{-1}) and optimized geometries of species in the inversion hydrolysis of d4U (β-path) with five explicit water molecules.

TABLE 1 Geometrical parameters and charge distributions of the stationary points in α-path and β-path (n=3).

	C1'-O4'	C1'-C2'	C2'-C3'	C1'	O4'	N1	O_{w1}
αϒ-P$^{\wedge}_{3\omega}$	1.41	1.50	1.33	0.25	-0.63	-0.50	-1.06
αϒ-TΣ1$_{3\omega}$	1.28	1.44	1.35	0.44	-0.47	-0.71	-1.06
αϒ-IN$_{3\omega}$	1.28	1.42	1.35	0.46	-0.45	-0.72	-1.06
αϒ-TΣ2$_{3\omega}$	1.29	1.45	1.34	0.43	-0.50	-0.71	-0.95
αϒ-Π$^{\wedge}_{3\omega}$	1.41	1.51	1.33	0.36	-0.65	-0.65	-0.80
βϒ-P$^{\wedge}_{3\omega}$	1.42	1.50	1.33	0.25	-0.63	-0.50	-1.05
βϒ-TΣ1$_{3\omega}$	1.29	1.44	1.35	0.43	-0.47	-0.71	-1.05
βϒ-IN1$_{3\omega}$	1.28	1.43	1.35	0.44	-0.45	-0.72	-1.05
βϒ-TΣ2$_{3\omega}$	1.28	1.43	1.35	0.45	-0.45	-0.74	-1.04
βϒ-IN2$_{3\omega}$	1.28	1.43	1.35	0.45	-0.44	-0.75	-1.05
βϒ-TΣ3$_{3\omega}$	1.28	1.43	1.35	0.44	-0.43	-0.75	-1.06
βϒ-IN3$_{3\omega}$	1.28	1.43	1.35	0.45	-0.44	-0.74	-1.04
βϒ-TΣ4$_{3\omega}$	1.30	1.45	1.34	0.43	-0.48	-0.75	-0.98
βϒ-Π$^{\wedge}_{3\omega}$	1.43	1.50	1.33	0.37	-0.65	-0.66	-0.83

TABLE 2 Selected donor-acceptor bond orbital interaction and second-order perturbation stabilization energy *E(2)* (kcal mol^{-1}) for key transition states in α-path and β-path of d4U, d4T, ddU and ddT (n=3).

	Donor	Acceptor	$E(2)$	$E(2)$	$E(2)$	$E(2)$
			$\boldsymbol{\alpha\Upsilon\text{-}T\Sigma1_{3\omega}}$	$\boldsymbol{\alpha\Upsilon\text{-}T\Sigma1_{3\omega}}$	$\boldsymbol{\alpha\Upsilon\text{-}T\Sigma2_{3\omega}}$	$\boldsymbol{\alpha\Upsilon\text{-}T\Sigma2_{3\omega}}$
d4U	LP(N1)	BD*(C1'-O4')	34.21	34.21	0.74	0.74
	LP(O_{w1})	BD*(C1'-O4')	0.44	0.44	70.36	70.36
	BD(C2'-C3')	BD*(C1'-O4')	51.78	51.78	44.25	44.25
			$\boldsymbol{\beta\Upsilon\text{-}T\Sigma1_{3\omega}}$	$\boldsymbol{\beta\Upsilon\text{-}T\Sigma2_{3\omega}}$	$\boldsymbol{\beta\Upsilon\text{-}T\Sigma3_{3\omega}}$	$\boldsymbol{\beta\Upsilon\text{-}T\Sigma4_{3\omega}}$
	LP(N1)	BD*(C1'-O4')	32.39	9.55	0.55	0.00
	LP(O_{w1})	BD*(C1'-O4')	0.32	0.45	1.65	54.43
	BD(C2'-C3')	BD*(C1'-O4')	51.32	56.85	61.88	44.49
			$\boldsymbol{\alpha T\text{-}T\Sigma1_{3\omega}}$	$\boldsymbol{\alpha T\text{-}T\Sigma1_{3\omega}}$	$\boldsymbol{\alpha T\text{-}T\Sigma2_{3\omega}}$	$\boldsymbol{\alpha T\text{-}T\Sigma2_{3\omega}}$
d4T	LP(N1)	BD*(C1'-O4')	32.24	32.24	0.11[a]	0.11[a]
	LP(O_{w1})	BD*(C1'-O4')	0.87	0.87	86.21[b]	86.21[b]
	BD(C2'-C3')	BD*(C1'-O4')	52.13	52.13	59.01[c]	59.01[c]
			$\boldsymbol{\beta T\text{-}T\Sigma1_{3\omega}}$	$\boldsymbol{\beta T\text{-}T\Sigma2_{3\omega}}$	$\boldsymbol{\beta T\text{-}T\Sigma3_{3\omega}}$	$\boldsymbol{\beta T\text{-}T\Sigma4_{3\omega}}$
	LP(N1)	BD*(C1'-O4')	29.22	9.55	0.55	0.00
	LP(O_{w1})	BD*(C1'-O4')	0.28	0.45	1.79	54.43
	BD(C2'-C3')	BD*(C1'-O4')	52.17	56.85	61.88	44.49
			$\boldsymbol{\alpha\Upsilon\text{-}T\Sigma1'_{3\omega}}$	$\boldsymbol{\alpha\Upsilon\text{-}T\Sigma2'_{3\omega}}$	$\boldsymbol{\beta\Upsilon\text{-}T\Sigma1'_{3\omega}}$	$\boldsymbol{\beta\Upsilon\text{- }T\Sigma2'_{3\omega}}$
ddU	LP(N1)	BD*(C1'-O4')	35.72	4.76	1.01	0.00
	LP(O_{w1})	BD*(C1'-O4')	6.75	41.27	1.17	15.12
	BD(C2'-H_a')	BD*(C1'-O4')	8.71	12.79	12.26	10.10
	BD(C2'-H_b')	BD*(C1'-O4')	8.90	6.51	9.05	10.09
			$\boldsymbol{\alpha T\text{-}T\Sigma1'_{3\omega}}$	$\boldsymbol{\alpha T\text{-}T\Sigma2'_{3\omega}}$	$\boldsymbol{\beta T\text{-}T\Sigma1'_{3\omega}}$	$\boldsymbol{\beta T\text{-}T\Sigma2'_{3\omega}}$
ddT	LP(N1)	BD*(C1'-O4')	36.23	5.48	0.84	0.00
	LP(O_{w1})	BD*(C1'-O4')	7.65	40.99	1.33	16.15
	BD(C2'-H_a')	BD*(C1'-O4')	8.65	12.69	12.46	10.01
	BD(C2'-H_b')	BD*(C1'-O4')	8.86	6.53	8.91	10.06

[a] LP(N1)-LP*(C1') [b]LP(Ow1)-LP*(C1')[c]LP(C2'-C3')-LP*(C1')

3.3 Comparison of d4U with d4T

The smallest microhydration model with three explicit water molecules for α-path and β-path was employed for d4T. The obtained PEPs of α-path and β-path, as well as transition states are depicted in Figure 8.

The calculation reveals that the PEPs of α-path and β-path in d4T are similar to those in d4U, with glycosidic bond dissociation step to be RDS. The presence of methyl group at the 5-position of d4U increases the activation free energies of RDS for α-path and β-path to be 27.3 ($\boldsymbol{\alpha T\text{-}T\Sigma1_{3\omega}}$) and 28.2 kcal mol^{-1} ($\boldsymbol{\beta T\text{-}T\Sigma1_{3\omega}}$), respectively, in contrast to the corresponding 24.8 ($\boldsymbol{\alpha\Upsilon\text{-}T\Sigma1_{3\omega}}$) and 26.4 kcal mol^{-1} ($\boldsymbol{\beta\Upsilon\text{-}T\Sigma1_{3\omega}}$) in d4U. Thus, different from the stability order of dT < dU, the order of glycosidic bond stability here is d4T > d4U, which can be understood by the decreasing $E^{(2)}$ value of n(N1)-π*(C1'-O4') as following: 32.24 kcal mol^{-1}($\boldsymbol{\alpha T\text{-}T\Sigma1_{3\omega}}$ in d4T) < 34.21 kcal mol^{-1} ($\boldsymbol{\alpha\Upsilon\text{-}T\Sigma1_{3\omega}}$ in d4U). The relative lower stability of d4U than d4T is probably an important factor for less antiviral activity of d4U.

In addition, the small free energy barrier differences for β-path over α-path in d4T (0.9 kcal mol^{-1}) and d4U (1.6 kcal mol^{-1}) suggest β-paths are competitive pathways, especially the β-path of d4T. And the reaction free energy differences of -11.6 (-3.0 vs 8.6 kcal mol^{-1} for$\boldsymbol{\beta T\text{-}\Pi^{*}_{3\omega}}$ and $\boldsymbol{\alpha T\text{-}\Pi^{*}_{3\omega}}$) and -12.6 kcal mol^{-1} (-5.7 vs 6.9 kcal mol^{-1} for $\boldsymbol{\beta\Upsilon\text{-}\Pi^{*}_{3\omega}}$ and$\boldsymbol{\alpha\Upsilon\text{-}\Pi^{*}_{3\omega}}$) indicate a striking thermodynamically favorable β-path in d4T and d4U.

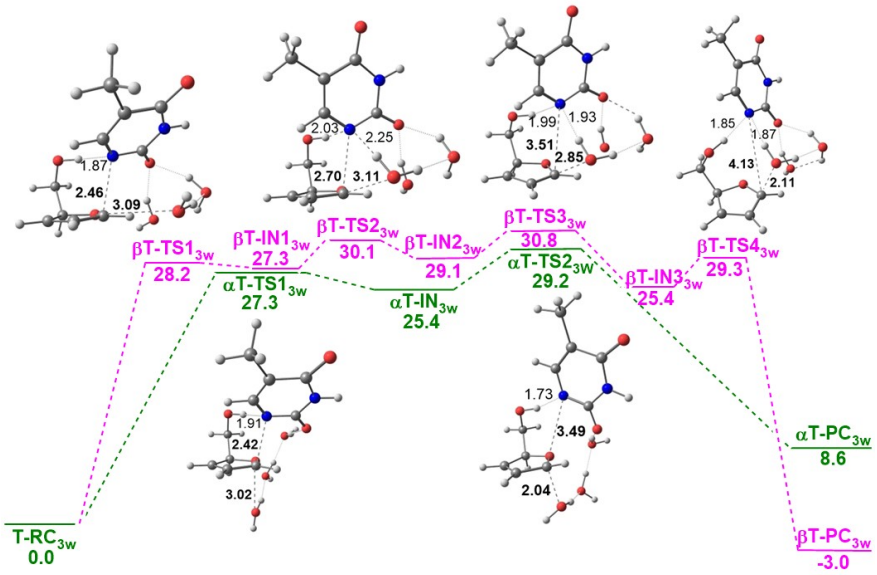


FIGURE 8 Comparison of Gibbs free energy profile (kcal mol^{-1}) for the inversion (green) and retention (pink) hydrolysis of d4T with three explicit water molecules in the cluster-continuum model.

3.4 Comparison of d4U/d4T with ddU/ddT

The saturated ddU and ddT are also examined with three explicit water molecules in microhydration model (Figure 9 and 10). Unlike the complicated PEPs involving multiple transition states for glycosidic bond cleavage in d4U and d4T, herein β-paths in ddU and ddT exhibit similar PEPs as α-paths, just with two transition states: glycosidic bond cleavage (RDS) ($\boldsymbol{\beta\Upsilon'\text{-}T\Sigma 1_{3\omega}}$ and$\boldsymbol{\beta T'\text{-}T\Sigma 1_{3\omega}}$) and the subsequent nucleophilic attack of water molecule ($\boldsymbol{\beta\Upsilon'\text{-}T\Sigma 2_{3\omega}}$ and$\boldsymbol{\beta T'\text{-}T\Sigma 2_{3\omega}}$).

The stability of glycosidic bond C1'-N1 in ddU and ddT has been reinforced with higher activation free energies of RDSs in contrast to those in d4U and d4T: 27.6 ($\boldsymbol{\alpha\Upsilon'\text{-}T\Sigma 1_{3\omega}}$) vs 24.8 kcal mol^{-1} ($\boldsymbol{\alpha\Upsilon\text{-}T\Sigma 1_{3\omega}}$), and 29.0 ($\boldsymbol{\alpha T'\text{-}T\Sigma 1_{3\omega}}$) vs 27.3 kcal mol^{-1} ($\boldsymbol{\alpha T\text{-}T\Sigma 1_{3\omega}}$). β-paths in these RDSs also show the same tendency. The increment of activation free energies in ddU and ddT can be understood by absence of strong interaction of π(C2'-C3')-π*(C1'-O4') by 52 kcal mol^{-1}in $\boldsymbol{\alpha(\beta)\Upsilon\text{-}T\Sigma 1_{3\omega}}$ and$\boldsymbol{\alpha(\beta)T\text{-}T\Sigma 1_{3\omega}}$, instead of the hyperconjugative effect by donation from C2'-H$_a$' and C2'-H$_b$' to the vacant antibonding orbital C1'-O4' in$\boldsymbol{\alpha(\beta)\Upsilon\text{-}T\Sigma 1'_{3\omega}}$ and$\boldsymbol{\alpha(\beta)T\text{-}T\Sigma 1'_{3\omega}}$, which is predicted to be just 9˜12 kcal mol^{-1}.

Furthermore, the saturation of sugar moiety enhances the kinetical difference of ˜5 kcal mol^{-1} for β-paths over α-paths: 27.6 ($\boldsymbol{\alpha\Upsilon'\text{-}T\Sigma 1_{3\omega}}$) vs 32.4 kcal mol^{-1} ($\boldsymbol{\beta\Upsilon'\text{-}T\Sigma 1_{3\omega}}$), and 29.0 ($\boldsymbol{\alpha T'\text{-}T\Sigma 1_{3\omega}}$) vs 34.3 kcal mol^{-1} ($\boldsymbol{\beta T'\text{-}T\Sigma 1_{3\omega}}$). The calculated $E^{(2)}$ energies of only ˜1 kcal mol^{-1} for n(N1)-π*(C1'-O4') in$\boldsymbol{\beta\Upsilon'\text{-}T\Sigma 1_{3\omega}}$ and$\boldsymbol{\beta T'\text{-}T\Sigma 1_{3\omega}}$ are much smaller than those of ˜36 kcal mol^{-1} in$\boldsymbol{\alpha\Upsilon'\text{-}T\Sigma 1_{3\omega}}$ and$\boldsymbol{\alpha T'\text{-}T\Sigma 1_{3\omega}}$, which provides an understanding to the greater kinetical favor for α-paths in ddNs. This suggests trivial possibility of β-path to take place compared with d4Ns.

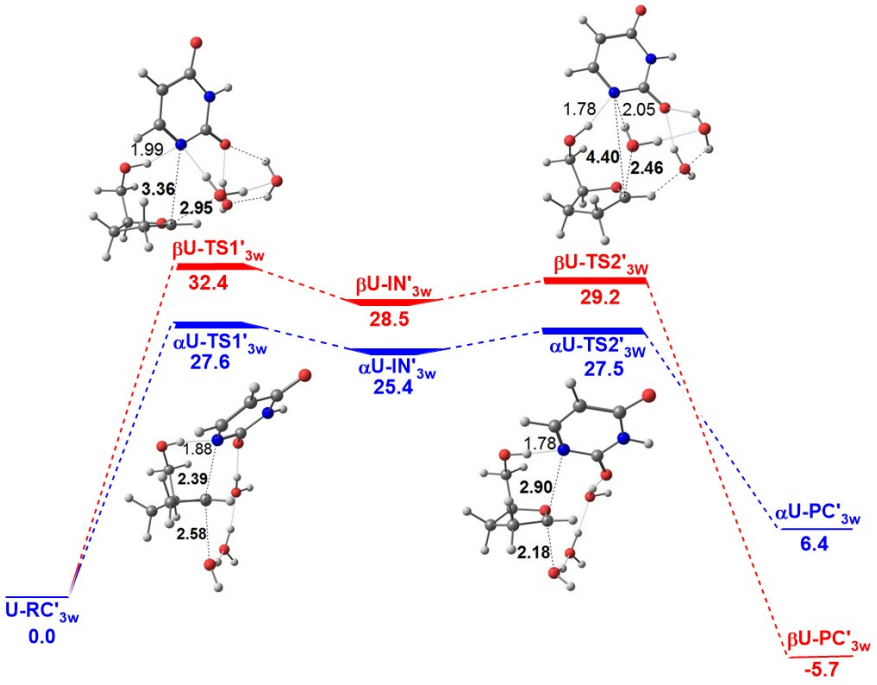


FIGURE 9 Comparison of Gibbs free energy profile (kcal mol^{-1}) for the inversion (blue) and retention (red) hydrolysis of ddU with three explicit water molecules.

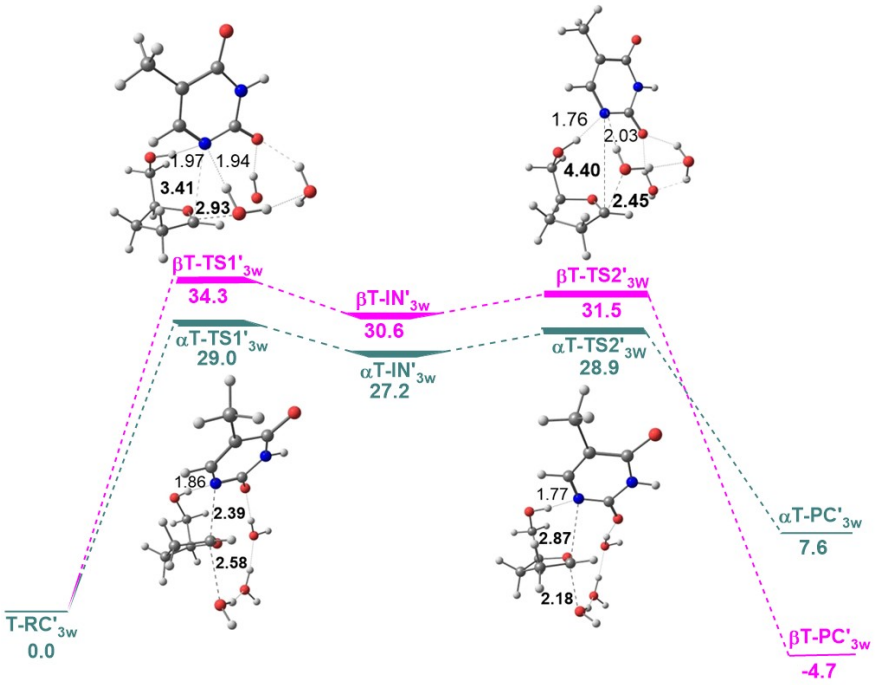


Figure 10. Comparison of Gibbs free energy profile (kcal mol^{-1}) for the inversion (green) and retention (pink) hydrolysis of ddT with three explicit water molecules.

CONCLUSION

Two plausible dissociative hydrolysis pathways of d4U, α-path with configuration-inversion and β-path with configuration-retention, have been examined using microhydration model with two to five explicit water molecules (n=2-5). The model with three explicit water molecules (n=3) is the smallest one for α-path and β-path. α-Path and β-path of d4T, as well as saturated ddU and ddT were also studied using microhydration model (n=3). Our results suggest that the glycosidic cleavage is the rate-determining step (RDS), and α-path is kinetically favorable than β-path, whereas β-path is thermodynamically favorable.

The free energy barriers of RDSs in α-path for d4U (24.8 kcal mol^{-1}) and d4T (27.3 kcal mol^{-1}) provide a kinetic support for the order of glycosidic bond stability: d4T > d4U, where the relative lower stability of d4U is probably an important factor for less antiviral activity of d4U. The small free energy barrier differences of ˜1 kcal mol^{-1} for β-path over α-path, as well as the reaction free energy differences of ˜ -12 kcal mol^{-1} for β-path lower than α-path in d4T and d4U suggest a competitive β-path in pyrimidine d4Ns. The saturated sugar moiety in ddU and ddT strengthens the stability of glycosidic bond in contrast to d4U and d4T. Our results provide an exploration for the less antiviral activity of d4U and the influence of saturated ribose on the glycosidic bond stability of pyrimidine d4Ns.

FUNDING INFORMATION This research was supported by Natural Science Foundation of China (21973062) and Yibin Key Laboratory of Chiral Small-Molecule Drug Research and Applications.

SUPPORTING INFORMATION Figure S1 showing the transition states involving in the inversion hydrolysis of dU and dT with three explicit water molecules, and Cartesian coordinates and related energies of all stationary points.

Reference

[1] 1 T. S. Lin, R. F. Schinazi, W. H. Prusoff, *Biochem. Pharmacol.* **1987** , *36(17)* , 2713.

[2] M. M. Mansuri, J. E. Jr. Starrett, I. Ghazzouli, M. J. M. Hitchcock, R. Z. Sterzycki, V. Brankovan, T. S. Lin, E. M. August, W. H. Prusoff, J.-P. Sommadossi, J. C. Martin, *J. Med. Chem.***1989** , *32* , 461.

[3] P. V. Roey, E. W. Taylor, C. K. Chu, R. F. Schinazi, *J. Am. Chem. Soc.* **1993** , *115* , 5365.

[4] E. De Clercq, *J. Clin. Virol.* **2001** , *22* , 73.

[5] E. De Clercq, *Biochim. Biophys. Acta.* **2002** ,*1587* , 258.

[6] D. J. Newman, G. M. Cragg, K. M. Snader, *Nat. Prod. Rep.***2000** , *17* , 215.

[7] P. Herdewijn, (Ed.) Modified nucleosides in biochemistry, biotechnology and medicine; Wiley: Weinheim, Germany, **2008** .

[8] E. De Clercq, *Trends Pharmacol. Sci.* **1987** ,*8(9)* , 339.

[9] M. J. M. Hitchcock, *Antiviral Chem. Chemother.***1991** , *2* , 125.

[10] C. K. Chu, R. F. Schinazi, B. H. Arnold, D. L. Cannon, B. Doboszewski, V. B. Bhadti, Z. Gu, *Biochem. Pharmacol.***1988** , *37* , 3543.

[11] J. Balzarini, G. J. Kang, M. Dalal, P. Herdewijn, E. De Clercq, S. Broder, D. G. Johns, *Mol. Pharmacol.* **1987** , *32* , 162.

[12] Y. Mehellou, J. Balzarini, C. McGuigan, *Org. Biomol. Chem.* **2009** , *7* , 2548.

[13] W. E. Harte Jr, J. E. Starrett Jr, J. C. Martin, M. M. Mansuri,*Biochem. Biophys. Res. Commun.* , **1991** , *175* , 298.

[14] P. Van Roey, E. W. Taylor, C. K. Chu, R. F. Schinazi, *J. Am. Chem. Soc.* **1993** , *115* , 5365.

[15] M. Mirmehrabi, S. Rohani, M. C. Jennings, *Acta Cryst. C.: Cryst. Struct. Commum.* **2005** , *C61* , o696.

[16] A. G. Ponomareva, Y. P. Yurenko, R. O. Zhurakivsky, T. van Mourik, D. M. Hovorun, *Curr. Phys. Chem.* **2013** , *3* , 83.

[17] A. G. Ponomareva, Y. P. Yurenko, R. O. Zhurakivsky, T. van Mourik, D. M. Hovorun. *Phys. Chem. Chem. Phys.* **2012** ,*14* , 6787.

[18] Z. J. Devereaux, C. C. He, Y. Zhu, H. A. Roy, N. A. Cunningham, L. A. Hamlow, G. Berden, J. Oomens, M. T. Rodgers, *J. Am. Soc. Mass Spectrom.* **2019** , *30* , 1521.

[19] A. C. Drohat, A. Maiti, *Org. Biomol. Chem.***2014** , *12(14)* , 8367.

[20] M. T. Bennett, M. T. Rodgers, A. S. Hebert, L. E. Ruslander, L. Eisele, A. C. Drohat, *J. Am. Chem. Soc.* **2006** ,*128* , 12510.

[21] J. Shi, A. S. Ray, J. S. Mathew, K. S. Anderson, C. K. Chu, R. F. Schinazi, *Bioorg. Med. Chem. Lett.* **2004** , *14* , 2159.

[22] J. L. Przybylski, S. D. Wetmore, *J. Phys. Chem. B.***2009** , *113* , 6533.

[23] J. L. Przybylski, S. D. Wetmore, *J. Phys. Chem. B.***2010** , *114* , 1104.

[24] Y. Jiang, Y. Xue, Y. Zeng, *J. Phys. Chem. B.***2018** , *122* , 1816.

[25] Y. Zeng, Y. Xue, G. Yan, *J. Phys. Chem. B* .**2008** , *112* , 10659.

[26] Y. Zeng, Y. Xue, G. Yan, *J. Phys. Chem. A.***2011** , *115* , 4995.

[27] A. L. Millen, L. A. B. Archibald, K. C. Hunter, S. D. Wetmore,*J. Phys. Chem. B.* **2007** , *111* , 3800.

[28] Y. E. R. Jeong, S. A. P. Lenz, S. D. Wetmore, *J. Phys. Chem. B.* **2020** , *124* , 2392.

[29] R. Rios-Font, L. Rodríguez-Santiago, J. Bertran, M. Sodupe,*J. Phys. Chem. B.* **2007** , *111* , 6071.

[30] D. H. Everaert, O. M. Peeters, C. J. De Ranter, N. M. Blaton, A. van Aerschot, P. Herdewijn, *Antiviral Chem. Chemother.***1993** , *4(5)* , 289.

[31] L. S. Jeong, Y. N. Choi, D. K. Tosh, W. J. Choi, H. O. Kim, J. Choi, *Bioorg. Med. Chem.* **2008** , *16* , 9891.

[32] V. Barone, M. Cossi, *J. Phys. Chem. A.* **1998** ,*102* , 1995.

[33] M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Comput. Chem.* **2003** , *24* , 669.

[34] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008** ,*120* , 215.

[35] Y. Zhao, D. G. Truhlar, **2008** , *41* , 157.

[36] K. Fukui, *Acc. Chem. Res.* **1981** , *14* , 363.

[37] A. E. Reed, L. A. Curtiss, F. Weinhold, *Chem. Rev.***1988** , *88* , 899.

[38] M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al. Gaussian 09, Revision A.02, Gaussian, Inc., Wallingford, CT, **2009** .

[39] M. A. Palafox, N. Iza, M. de la Fente, R. Navarro, *J. Phys. Chem B.* **2009** , *113* , 2458.

[40] G. K. Schroeder, R. Wolfenden, *Biochemistry.***2007** , *46* , 13638.