Biochemistry

pubs.acs.org/biochemistry Article

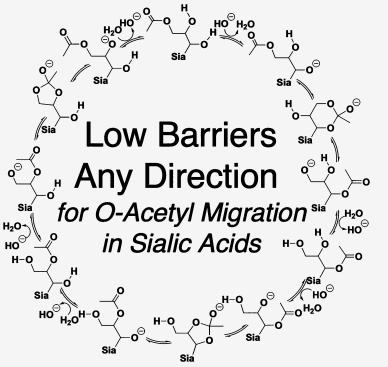
O-Acetyl Migration within the Sialic Acid Side Chain: A Mechanistic Study Using the *Ab Initio* Nanoreactor

Lisa Oh, Yang Ji, Wanqing Li, Ajit Varki, Xi Chen, and Lee-Ping Wang*

Cite This: https://doi.org/10.1021/acs.biochem.2c00343 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Many disease-causing viruses target sialic acids on the surface of host cells. Some viruses bind preferentially to sialic acids with *O*-acetyl modification at the hydroxyl group of C7, C8, or C9 on the glycerol-like side chain. Studies of proteins binding to sialosides containing *O*-acetylated sialic acids are crucial in understanding the related diseases but experimentally difficult due to the lability of the ester group. We recently showed that *O*-acetyl migration among hydroxyl groups of C7, C8, and C9 in sialic acids occurs in all directions in a pH-dependent manner. In the current study, we elucidate a full mechanistic pathway for the migration of *O*-acetyl among C7, C8, and C9. We used an *ab initio* nanoreactor to explore potential reaction pathways and density functional theory, pK_a calculations, and umbrella sampling to investigate elementary steps of interest. We found that when a base is present, migration is easy in any direction and involves three key steps: deprotonation of the hydroxyl group, cyclization between the two carbons, and the migration of the *O*-acetyl group. This dynamic equilibrium may play a defensive role against pathogens that evolve to gain entry to the cell by binding selectively to one acetylation state.



BACKGROUND

The sialic acid family includes more than 50 forms that are derived from the nine-carbon backbone monosaccharides neuraminic acid and 2-keto-3-deoxy-nononic acid (Kdn). Among these forms, *N*-acetylneuraminic acid (Neu5Ac) is most common.[1] Many disease-causing viruses target sialic acids (Sias) with specific *O*-acetyl (OAc) modifications. In one example, betacoronavirus OC43 and HKUI are known to bind to 9-*O*-acetylated Sias,[2,3] the most common Sia *O*-acetyl modification found in nature. In our recent article, we noted the preferential binding of SARS-CoV-2 S protein to an immobilized acetyl group at the C9−OH of a Sialyl Lewis x antigen by using an *N*-acetyl group analogue Neu5Ac9NAc which is resistant to migration over its *O*-acetyl counterpart Neu5,9Ac$_2$.[4] More recently, multivalent 9-*O*-acetylated-sialic acid glycoclusters were reported as potent inhibitors of SARS-CoV-2 infection.[5] Although it is important to understand the mechanisms of OAc-Sia binding that causes the diseases, many are unknown due to experimental difficulties from *O*-acetyl ester instability. *O*-Acetylated Sias are sensitive to pH, temperature, and esterases.[6−8] Additionally, the *O*-acetyl group is prone to spontaneously migration across the exocyclic glycerol-like Sia side-chain. This makes it very difficult to study glycan−protein binding using microarrays, as the expected *O*-acetyl group on the glycan may be cleaved or have migrated. In our recent study, we showed a pH-dependent Sia *O*-acetyl migration in all directions among the C9, C8, and C7 positions, to form Neu5,9Ac$_2$, Neu5,8Ac$_2$, and Neu5,7Ac$_2$, respectively (Figure 1).[7]

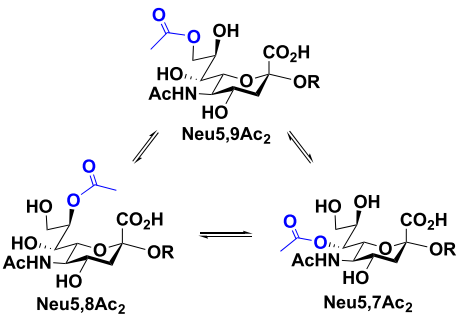


Figure 1. *O*-Acetyl migration is reversible between Neu5,9Ac$_2$, Neu5,8Ac$_2$, and Neu5,7Ac$_2$.

In this study, we probe the molecular details of this migration using density functional theory (DFT). There are many ways to computationally investigate mechanisms that involve finding critical points on the molecular potential energy surface, namely, energy-minimized reactants and products, intuitively driven intermediate structures, and the transition states connecting them *via* minimum energy pathways. Although conventional approaches to optimize transition

Received: June 14, 2022
Revised: August 16, 2022

states may connect the target reactant/product combination, success is highly dependent on user input and intuition. To overcome this, several approaches have been developed for automated reaction discovery.[9−18] Some methods include the use of semiempirical or fully *ab initio* molecular dynamics (AIMD) simulations, which can predict mechanisms independent of pre-existing information about the reaction or heuristic reactivity rules.[19] The nanoreactor is an AIMD-based method specialized to induce a large number of reactions in a small amount of simulation time by increasing reactivity through elevated temperature and spherically symmetric compressions that oscillate in time.[10,12,20] The nanoreactor trajectories contain many reaction events that are automatically refined to individual mechanistic steps, where a single transition state connects each reactant/product structure pair.

We used the nanoreactor to generate mechanistic guesses, followed by a combination of automated and manual energy refinement to obtain minimum energy paths of *O*-acetyl group migration.[10,12] We then tested the feasibility of our proposed mechanism under experimental conditions by computing pK_a values of ionizable groups on the Sia glycerol chain and free energy profiles of key bond dissociation steps using umbrella sampling simulations. We propose that an initial deprotonation step enables the *O*-acetyl group to form a cyclic intermediate, followed by migration with low activation barriers among the C9-, C8-, and C7-hydroxyl groups of Neu5Ac.

METHODS

The nanoreactor is a specially modified AIMD simulation with the primary goal of inducing reaction events on a short simulation time scale, achieved by means of a time-oscillating boundary potential that generates high-energy molecular collisions. The initial conditions consisted of one Sia molecule (either Neu5,7Ac$_2$ or Neu5,9Ac$_2$), α-anomer, surrounded by 12 water molecules randomly placed within a 5.8 and 5.4 Å sphere, using Packmol software,[21] followed by energy minimization. Although 12 explicit water molecules are not enough to fully solvate the Sia molecule, we found that it was sufficient to observe a range of solute−solvent reactivity, such as solvent-mediated proton transfer, without significantly increasing computational time. The level of theory for the nanoreactor simulations was B3LYP with a DFT-D3 dispersion correction, 3-21G Gaussian basis set, and SWIG-PCM polarizable continuum model of the aqueous reaction medium (ε = 78.4), denoted B3LYP-D3/3-21G/PCM. The implicit solvent provided an approximate description of bulk solvation beyond the explicit water molecules, which was needed to stabilize charged intermediate species. The relatively small 3-21G basis set was used in the nanoreactor simulations to mitigate the high computational cost of AIMD without unduly sacrificing the approximate accuracy needed for reaction discovery; even with the chosen small basis set, individual simulations were limited to a few ps in length. Calculations were performed using TeraChem quantum chemistry software[22−28] running on servers with an Intel Xeon CPU/Nvidia GPU (either E5-2630 v3 CPU/GTX 980 Ti GPU or E5-2637 v4 CPU/GTX 1080 Ti GPU).

Reaction events are induced by a time-dependent restraint potential in eq 1, where m is the atomic mass, R_1 and R_2 are radius parameters, and k_1 and k_2 are force constants divided by atomic mass. The restraint is a flat-bottomed harmonic potential with time-dependent parameters

$$V(r,t) = \frac{m}{2}k(t)\rho(t)^2;\ \rho = \begin{cases} r - R(t) & \text{if } r > R(t) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$k(t),\ R(t) = \begin{cases} k_1,\ R_1 & \text{if } t \bmod (t_1 + t_2) < t_1 \\ k_2,\ R_2 & \text{otherwise} \end{cases}$$

where the values of R and k oscillate in a square wave pattern and are equal to (R_1, k_1) for period t_1 and (R_2, k_2) for period t_2. Generally, $R_1 > R_2$, and during the period t_1, the molecules are free to move within a spherical volume of radius R_1 with no restraint force. When the radius decreases from R_1 to R_2, atoms beyond R_2 are accelerated inward, with the mass-proportional spring constant ensuring near-uniform acceleration of atoms within a molecule. Reaction events are initiated from high-velocity collisions of reactants, which generally occur within the smaller spherical volume of radius R_2.

Simulations used a time step of 0.5 fs with a Langevin thermostat set to 1000−2000 K and a friction coefficient of 3.3 ps^{-1}. Other simulation parameters were varied as we searched for the optimal combinations that yielded chemically reasonable reactivity within 2 ps of simulation time. At 2000 K, decarboxylation was observed to occur spontaneously, whereas temperatures between 1000 and 1500 K were optimal for relevant reaction exploration. The initial boundary radius (R_1) and second boundary radius (R_2) were chosen to be ∼1.5−2× and ∼0.8× that of the smallest initial radius used in Packmol. Reactivity can be further tuned by adjustment of the R_2 parameter, where smaller values of R_2 result in higher-energy collisions and a greater frequency of generated reactivity, although the reaction pathways found in this way tended to have higher activation energies after refinement. For example, some of our simulations used a combination of the following parameters: T = 1500 K, R_1 = 12.0 Å, k_1 = 1.0 kcal mol^{-1} Å^{-2} u^{-1}, t_1 = 1000 fs, R_2 = 4.4 Å, k_2 = 0.5 kcal mol^{-1} Å^{-2} u^{-1}, and t_2 = 500 fs.

The reaction events observed in the nanoreactor trajectory underwent an energy refinement procedure to produce optimized transition states and corresponding reactant/product structures. The refinement calculations used the ωB97X-D3/TZVP/PCM level of theory,[29−31] which is more accurate (and more costly) compared to what was used for the nanoreactor simulations. Initial transition state guesses were optimized using the nudged elastic band method[32,33] implemented in a development version of geomeTRIC optimization software[34] that calls TeraChem software[35,36] for energy and gradient evaluations. The transition state optimizations were performed in Q-Chem.[37] A frequency calculation was used to confirm the presence of a single imaginary mode in the optimized transition state, and an intrinsic reaction coordinate calculation was then performed to connect the optimized transition state with reactant to product structures.

Torsional energy profiles were computed for a few minor intermediate steps, for the purpose of matching each geometric endpoint between elementary steps. More specifically, for a few glycerol rotation steps that were unable to be optimized to single transition states, torsional energy profiles were computed by scanning relevant dihedrals by ∼3° per step using a development version of geomeTRIC optimization software[34] at the ωB97X-D3/TZVP level of theory. This was done to maintain maximal consistency between each intermediate within the reaction pathway.

https://doi.org/10.1021/acs.biochem.2c00343
Biochemistry XXXX, XXX, XXX−XXX

The pK_a values for two conformers of each glycerol ionizable group on Neu5,9Ac$_2$, Neu5,8Ac$_2$, or Neu5,7Ac$_2$, resulting in 12 acid/conjugate base pairs, were computed according to the following equation.[38,39]

$$pK_a = \frac{\Delta\Delta G}{2.303kT} = \frac{\Delta G_{base+H^+} - \Delta G_{acid}}{2.303kT} \quad (2)$$

Ten acid/conjugate base pairs were used to compute pK_a values for Neu5Ac, starting from the above-mentioned structures and replacing the acetyl group with hydroxyl. Additionally, to compare our results to experimental values, we used four acid/conjugate base pairs to compute the pK_as of the primary and secondary alcohols of glycerol, based on the Neu5Ac structures. In the pK_a calculations, the Sia (or glycerol) ionizable group was modeled by adding 7 explicit water molecules which forms a cubic hydrogen-bonded configuration together with the ionizable group, based on the precedence that inclusion of 3 waters can reduce the mean squared error between theory and experiment for a set of 72 organic compounds from around 8 to 1 pK_au, depending on the DFT functional.[39−41] In the conjugate base structure, three water molecules are positioned in a H-bond donor orientation, whereas for the acid, one of the H-bonding protons on water was rotated away, thereby maintaining a consistent number of H-bonds and maximal geometric consistency per acid/conjugate base pair.[39] The resulting structure was optimized at the ωB97X-D3/TZVP/PCM level of theory. Our choice of a cubic hydrogen-bonded solvent configuration maximizes the number of hydrogen bonds for each water molecule in the cluster and minimizes the formation of additional hydrogen bonds to the solute, which could detract from accuracy. Frequency analyses for vibrational contributions to free energies were calculated at the same level of theory at 298.15 K, while electronic energy of the optimized structure was calculated using the larger def2-TZVP-f basis set. In our pK_a calculations, we used a free energy value of 11.803 eV[42] for a proton at pH = 0. Figure 2a,b shows the optimized acid and conjugate base structures, and Figure 2c is an overlay of the water molecule structures to show the similarity in the H-bonding patterns.

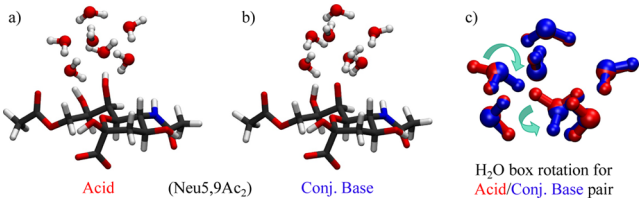


Figure 2. Diagram of Neu5,9Ac$_2$ surrounded by a water box of seven waters. Water box surrounding the ionizable C7−OH of Neu5,9Ac$_2$, maintaining three H-bonds, in the acid (a) and conjugate base (b) forms. (c) Illustration describing the proton bond rotations to maintain consistent H-bonding between the acid (red) and the conjugate base (blue) forms, highlighted in green arrows. Coloring key: H: white; C: gray; O: red; and N: blue.

Free energy profiles for cyclized bond-dissociation steps were computed from umbrella sampling simulations and using the plumed[43] plugin in TeraChem at the B3LYP/6-31G*/PCM level of theory with a harmonic spring constant of 0.50 hartree/bohr^2 and the Langevin thermostat set to 298.15 K. Seven explicit water molecules were placed around the oxide and near the dissociating bond to improve the description of solvation. Simulations of 5 ps in length were run with bond distances set from 2.60 to 4.50 bohr (1.38−2.38 Å) in increments of 0.10 bohr (0.05 Å), where larger distances resulted in unrealistic multi-bond cleavage within the Sia. Simulations within these distances were excluded when unusual reactivity and multi-bond cleavage occurred over the desired bond dissociation. Free energy profiles were generated from the last 4 ps of simulation data, using the Grossfield implementation of WHAM, version 2.0.10,[44] with 17−20 windows (depending on relevant simulations), from 2.6025 to 4.5975 bohr at 298.15 K, and with Monte Carlo bootstrap error analysis with 2000 iterations.[44,45]

RESULTS AND DISCUSSION

We initially looked for the migration from Neu5,9Ac$_2$ to Neu5,7Ac$_2$ through an intramolecular cyclization step, similar to a mechanism proposed for acetyl migration in lysophospholipids,[46] but found this to be energetically prohibitive with barriers of ∼50 kcal/mol (Figure S1). Following this, we carried out nanoreactor simulations to explore *O*-acetyl reactivity using Neu5,9Ac$_2$ and Neu5,7Ac$_2$ as key reactants and observed many proton transfer events between water and the glycerol hydroxyl groups (C7−OH, C8−OH, or C9−OH), de-*O*-acetylation, and HO$^-$ addition to the *O*-acetyl group. After energy refinement, we found that most reactions that started from deprotonated species had low free energy barriers of less than 15 kcal/mol (Figure S2). This is consistent with experimentally observed trends of increasing *O*-acetyl migration with higher pH.[47] Based on this reactivity, we explored *O*-acetyl migration steps involving deprotonation by HO$^-$.

Our calculations indicate that the migration of the acetyl group on Neu5,9Ac$_2$, Neu5,8Ac$_2$, or Neu5,7Ac$_2$ likely proceeds *via* a mechanism involving deprotonation and cyclization, followed by ring opening, subsequently resulting in migration, similar to a recently reported mechanism for acetyl migration within pyranosides (Figure 3).[48] When a base is present, deprotonation of a glycerol hydroxyl group is facile, with reaction barriers under 5 kcal/mol. This is hypothesized to be the first step in the acetyl migration pathway (Figure 4a−c). After deprotonation of Neu5,9Ac$_2$ C7OH (**1**), the resultant oxide (**2**) is a nucleophile that attacks the carbonyl carbon of the acetyl group and forms a cyclic intermediate with the carbonyl carbon adopting a tetrahedral geometry (**3**, **7**, **11**) (Figure 4d−f). This cyclic intermediate then ring-opens to result in an oxide at the previously *O*-acetyl-substituted carbon and the migrated acetyl group. Within this process, the oxygen atoms directly bonded to C7, C8, and C9 remain in place, and only the C(=O)CH$_3$ atoms migrate. Re-protonation by water results in Neu5,9Ac$_2$, Neu5,8Ac$_2$, or Neu5,7Ac$_2$ (**1**, **9**, **5**). Further migrations repeat the previous steps of deprotonation, cyclization, migration, and re-protonation. In addition to these chemical steps, some conformational differences were observed within an intermediate (*e.g.*, the product structure of 1 → 2 and the reactant structure of 2 → 3) which required further calculations to determine the free energy barrier of the conformational changes. Figure 2 indicates a single highest barrier for a given step (deprotonation, cyclization, or migration) in the direction of Neu5,9Ac$_2$ to Neu5,7Ac$_2$ to Neu5,8Ac$_2$ (see Figure S3 for the full pathway), including any conformational changes. For example, in order to cyclize **2** to **3**, three rotations of glycerol O−H groups and C−C torsions are required to prepare the acetyl group for nucleophilic attack. Our calculations require the presence of a HO$^-$ anion,

https://doi.org/10.1021/acs.biochem.2c00343
Biochemistry XXXX, XXX, XXX−XXX

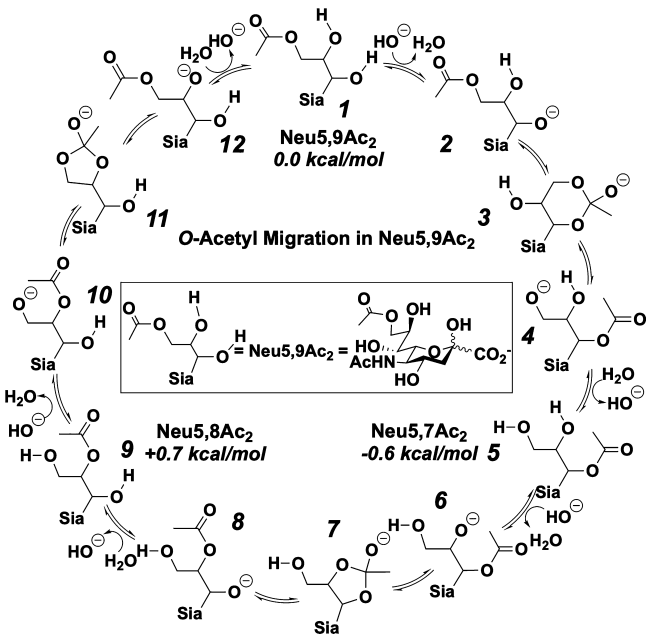


Figure 3. Mechanism of base-catalyzed *O*-acetyl migration among Neu5,9Ac$_2$, Neu5,8Ac$_2$, and Neu5,7Ac$_2$. Relative free energies calculated at the ωB97X-D3/TZVP level of theory as described in the Computational Details.

consistent with experiments that *O*-acetyl migration has a tendency to occur at higher pH values.

Although the base-catalyzed migration between Neu5,9Ac$_2$ and Neu5,7Ac$_2$ *via* the six-membered cyclic intermediate (**3**) has the lowest free energy barriers (<9 kcal/mol for each step), the transition state energies for the migrations between Neu5,9Ac$_2$ and Neu5,8Ac$_2$, or Neu5,7Ac$_2$ and Neu5,8Ac$_2$, are also low (<15 kcal/mol). The largest barriers in the migration route from Neu5,9Ac$_2$ to Neu5,7Ac$_2$ to Neu5,8Ac$_2$ are 12.7 kcal/mol for glycerol rotations of **10** (or **10′** after deprotonation of **9** to **10**) to geometrically pre-organize for the cyclization step in the *O*-acetyl migration from C8−OH to C9−OH (**11**). The reverse-direction maximum barrier is 14.2 kcal/mol, for the **8** to **7** glycerol rotations needed for the cyclization step to migrate the *O*-acetyl from C8−OH to C7−OH. Excepting these two glycerol rotation barriers, the maximum barrier in this migration cycle is ∼11 kcal/mol in either direction. Additionally, relative free energy differences of Neu5,8Ac$_2$ (+0.7 kcal/mol *vs* Neu5,9Ac$_2$) and Neu5,7Ac$_2$ (−0.6 kcal/mol *vs* Neu5,9Ac$_2$), calculated for structures **1**, **5**, and **9**, are small and within the margin of error of our calculations. Although Neu5,7Ac$_2$ is slightly lower in free energy than Neu5,9Ac$_2$, this may in part be due to the formation of an intramolecular H-bond between the amide proton of the *N*-acetyl on C5 and the carbonyl of *O*-acetyl on C7, with a distance of 2.0 Å. For the other Sias, the *N*-acetyl proton is closest to C7−OH with a distance of ∼2.5 Å. Because the barriers are low at each step of the base-catalyzed migration and the stabilities of Neu5,9Ac$_2$, Neu5,8Ac$_2$, and Neu5,7Ac$_2$ are similar, we think the acetyl group can migrate readily between positions (C7/8/9) at an increased pH.

This migration competes with de-*O*-acetylation, which is exergonic by −20.5 kcal/mol with the highest point on the reaction pathway ∼18 kcal/mol higher than the starting state. Given that the free energy barriers of *O*-acetyl migration in either direction from Neu5,9Ac$_2$ are <12 kcal/mol, the de-*O*-acetylation pathway, therefore, has a higher barrier, indicating that *O*-acetyl migration is kinetically more favorable. Thus, the de-*O*-acetylated Neu5Ac is the thermodynamic product, while the *O*-acetylated isomers are kinetic products.

In the cyclization and ring-opening calculations, we did not use the explicit solvent, which is needed to describe hydrogen bonding with solvent molecules. Given the mechanism's dependence on initial deprotonation and that these initial deprotonation steps have low barriers, we calculated pK_a values in the presence of explicit water to determine which hydroxyl may be deprotonated first. This could suggest a potential migration direction among Neu5,9Ac$_2$, Neu5,8Ac$_2$, and Neu5,7Ac$_2$ (see Methods for details). By taking the average of each ionizable group across the respective two Sias (ex. C9−OH results averaged from Neu5,8Ac$_2$ and Neu5,7Ac$_2$), the pK_a values of C9−OH and C7−OH are more similar at 14.5 and 14.1, respectively, compared to the slightly higher pK_a of C8−

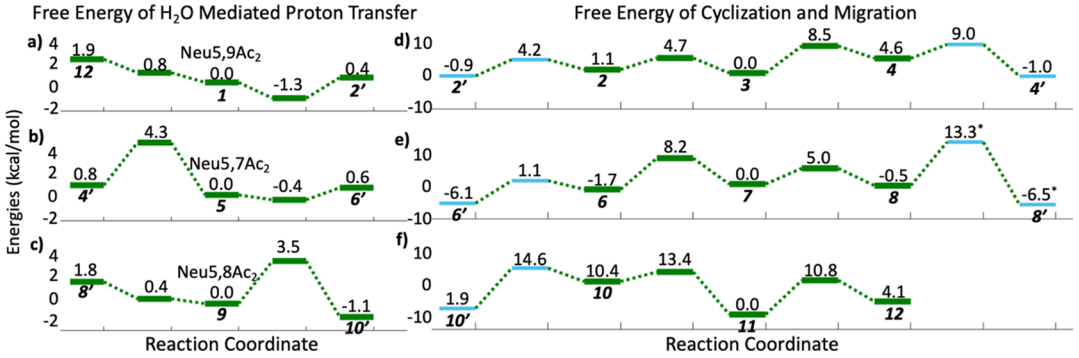


Figure 4. Reaction and activation free energies for base-catalyzed *O*-acetyl migration of Neu5,9Ac$_2$, Neu5,8Ac$_2$, and Neu5,7Ac$_2$. Small numbers in the bold font indicate structures depicted in Figure 3. (a−c) show deprotonation of C(7/8)-OH of Neu5,9Ac$_2$, C(9/7)-OH of Neu5,8Ac$_2$, and C(8/9)-OH of Neu5,7Ac$_2$, respectively, with free energies for each plot relative to the sialic acid and HO$^-$ (**1**, **5**, **9**). (d−f) show cyclization and ring-opening reactions for *O*-acetyl migration in C9−C7, C7−C8, and C8−C9, respectively, with free energies for each plot relative to the cyclic intermediate (**3**, **7**, **11**). Light-blue lines (*e.g.*, **2** ⟷ **2′**) indicate the free energies of conformational changes between the endpoints of the deprotonation and cyclization steps (*e.g.*, deprotonation of **1** results in **2′**, followed by conformational changes to **2** and then cyclization to **3**). *Energy barriers of conformational change steps were estimated by electronic energies from torsion scans (Figure S4). Free energies are calculated at the ωB97X-D3/TZVP level of theory as described in the Computational Details; negative activation free energies are an artifact of the quasi-harmonic approximation.

https://doi.org/10.1021/acs.biochem.2c00343
Biochemistry XXXX, XXX, XXX−XXX

OH at 15.2 (Figure 5). We performed additional pK_a calculations on the unmodified Neu5Ac for reference and

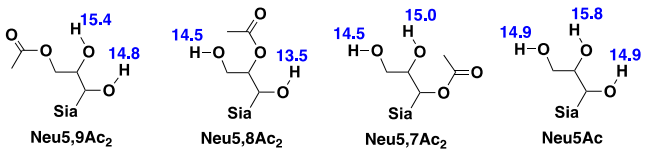


Figure 5. Computed pK_a values are highlighted in blue next to each ionizable group of Neu5,9Ac$_2$, Neu5,8Ac$_2$, Neu5,7Ac$_2$, and Neu5Ac. pK_a values were calculated at the ωB97X-D3/def2-TZVP-f level of theory for single-point energy contributions and ωB97X-D3/TZVP for vibrational contributions to free energies, as described in the Computational Details.

found the pK_as of C7−OH and C9−OH to be 14.9, lower than the pK_a for C8−OH at 15.8. As a second comparison, we computed the pK_a values for the primary and secondary alcohols of glycerol, resulting in 15.4 and 15.1. These values are close to the reported pK_a1 experimental value of 14.4 (within 1 pK_au).[49] Additionally, each computed pK_a for Sias is close to that of glycerol. As seen from our calculations on Sias C9/8/7-OH, as distance from the *O*-acetyl group increased, pK_a values tended to decrease by about 1. Even so, given the small differences between the pK_a values, this study further suggests that deprotonation of any hydroxyl group is comparatively easy among the Neu5,9Ac$_2$, Neu5,8Ac$_2$, and Neu5,7Ac$_2$ and that all deprotonated forms are likely present in an equilibrium mixture.

The conformational flexibility of the ring and glycerol chain across Sias was evident in our calculations of energy minima and transition states. To improve the sampling of particular reaction steps, we carried out umbrella sampling along the C−O distance coordinate for the cyclization and de-*O*-acetylation steps. Umbrella sampling may give a clearer picture of which ring is easier to break and in which direction the *O*-acetyl group is more likely to migrate. Figure 6 shows the free energy

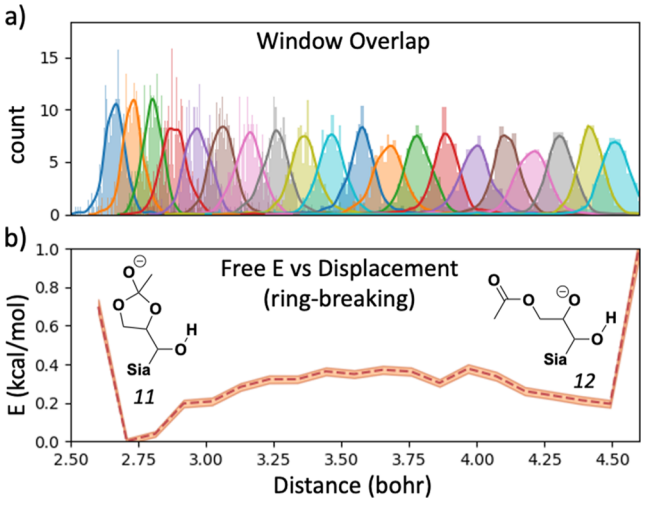


Figure 6. Example umbrella sampling of the cyclization/*O*-acetyl migration step between C9-/C8-hydroxyl groups. (a) Overlap of window sampling over bond length. (b) Free energy (red dashed line) of ring-break for *O*-acetyl migration from C8−OH to C9−OH with estimated error shaded in orange. Umbrella sampling carried out at the B3LYP/6-31G* level of theory as described in the Computational Details. See Figure S5 for all six steps.

of ring-opening from **11** to **12** in the *O*-acetyl migration from C8−OH to C9−OH, which involves a free energy barrier of <1.0 kcal/mol and a window overlap figure from our simulations. The figure indicates sufficient thermodynamic overlap between each window where the cyclized bond distance increases until the ring breaks, ensuring that the resultant free energy results are reliable. Each of the other ring-breaking steps has similar results, with free energy barriers of <1.0 kcal/mol and similar window overlaps (Figures 5 and S5). Umbrella sampling results with explicit water yielded lower activation free energies compared to those obtained from vibrational analysis without explicit water. This further supports the low-barriered, facile *O*-acetyl migration between the hydroxyl groups at C7, C8, and C9 of Sias.

To compare with the experiment, we calculated free energies from population ratios in our previous paper.[7] From the average population ratios (between SM1 and SM2) that reached an equilibrium at pH 8.0 and 37 °C, the free energies of Neu5,7Ac$_2$ and Neu5,8Ac$_2$ with respect to Neu5,9Ac$_2$ were +2.06 and +1.25 kcal/mol, respectively. Our values from a single structure of Neu5,7Ac$_2$ and Neu5,8Ac$_2$ (Figure 3) were −0.6 and +0.7 kcal/mol, respectively, with respect to Neu5,9Ac$_2$. The free energies for Neu5,8Ac$_2$ match within the margin of error, but the computed value for Neu5,7Ac$_2$ was slightly lower than expected. Our computational results are not precise enough to reproduce the experimentally measured population ratios of isomers; this could partly be due to the approximate level of theory used and/or our use of the quasi-harmonic approximation to evaluate these free energy differences. Although it may be possible to obtain more accurate free energy estimations using more advanced sampling methods,[50−52] it does not affect our main conclusion regarding the proposed migration mechanism, in which migration in any direction is facile provided that a base is present.

CONCLUSIONS

Our calculations indicate that when a base is present, *O*-acetyl migration occurs readily among the hydroxyl groups at C9, C7, and C8 of the Sias *via* deprotonation, migration, and cyclization steps. This conclusion is based on computed pK_a values, umbrella sampling, and DFT calculations on the migration steps and the competing de-*O*-acetylation reaction. When explicit water molecules were included in the calculations, their presence affected relative energies of intermediates but did not change the main conclusion that base-catalyzed acetyl migration is facile. *O*-Acetyl migration may allow host cells to evade viral binding. In our recent article, we noted the preferential binding of SARS-CoV-2 S protein to the immobilized acetyl group at the C9−OH of a Sialyl Lewis x antigen by using an *N*-acetyl group analogue Neu5Ac9NAc which is resistant to migration over its *O*-acetyl counterpart Neu5,9Ac$_2$.[4] Further work is necessary to probe the role of *O*-acetyl migration in viral recognition and pathogenesis.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.2c00343.

Video of the full computed migration mechanism (MP4)

https://doi.org/10.1021/acs.biochem.2c00343
Biochemistry XXXX, XXX, XXX−XXX

Reaction free energies and barriers for select migration steps without deprotonation, full migration pathway and free energies, torsional energy profile-selected glycerol rotation steps, and umbrella sampling results for each cyclization/*O*-acetyl migration step (PDF)

AUTHOR INFORMATION

Corresponding Author

Lee-Ping Wang − *Department of Chemistry, University of California, Davis, California 95616, United States;* orcid.org/0000-0003-3072-9946; Email: leeping@ucdavis.edu

Authors

Lisa Oh − *Department of Chemistry, University of California, Davis, California 95616, United States;* orcid.org/0000-0001-8572-9738

Yang Ji − *Glycobiology Research and Training Center, Departments of Medicine and Cellular and Molecular Medicine, University of California, San Diego, California 92093, United States*

Wanqing Li − *Department of Chemistry, University of California, Davis, California 95616, United States*

Ajit Varki − *Glycobiology Research and Training Center, Departments of Medicine and Cellular and Molecular Medicine, University of California, San Diego, California 92093, United States;* orcid.org/0000-0002-2206-975X

Xi Chen − *Department of Chemistry, University of California, Davis, California 95616, United States;* orcid.org/0000-0002-3160-614X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.biochem.2c00343

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We are grateful for the insightful discussions with Darón I. Freedberg and Audra A. Hargett at the Laboratory of Bacterial Polysaccharides, Food and Drug Administration (FDA). This work was supported by the United States National Institutes of Health grant no. R01AI130684 and R01GM032373.

REFERENCES

(1) Lewis, A. L.; Chen, X.; Schnaar, R. L.; Varki, A. Sialic Acids and Other Nonulosonic Acids. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2022.

(2) Hulswit, R. J. G.; Lang, Y.; Bakkers, M. J. G.; Li, W.; Li, Z.; Schouten, A.; Ophorst, B.; van Kuppeveld, F. J. M.; Boons, G.-J.; Bosch, B.-J.; Huizinga, E. G.; de Groot, R. J. Human coronaviruses OC43 and HKU1 bind to 9- O -acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 2681−2690.

(3) Tortorici, M. A.; Walls, A. C.; Lang, Y.; Wang, C.; Li, Z.; Koerhuis, D.; Boons, G.-J.; Bosch, B.-J.; Rey, F. A.; de Groot, R. J.; Veesler, D. Structural Basis for Human Coronavirus Attachment to Sialic Acid Receptors. *Nat. Struct. Mol. Biol.* **2019**, *26*, 481−489.

(4) Dhar, C.; Sasmal, A.; Diaz, S.; Verhagen, A.; Yu, H.; Li, W.; Chen, X.; Varki, A. Are Sialic Acids Involved in COVID-19 Pathogenesis? *Glycobiology* **2021**, *31*, 1068−1071.

(5) Petitjean, S. J. L.; Chen, W.; Koehler, M.; Jimmidi, R.; Yang, J.; Mohammed, D.; Juniku, B.; Stanifer, M. L.; Boulant, S.; Vincent, S. P.; Alsteens, D. Multivalent 9-O-Acetylated-Sialic Acid Glycoclusters as Potent Inhibitors for SARS-CoV-2 Infection. *Nat. Commun.* **2022**, *13*, 2564.

(6) Khedri, Z.; Xiao, A.; Yu, H.; Landig, C. S.; Li, W.; Diaz, S.; Wasik, B. R.; Parrish, C. R.; Wang, L.-P.; Varki, A.; Chen, X. A Chemical Biology Solution to Problems with Studying Biologically Important but Unstable 9-O-Acetyl Sialic Acids. *ACS Chem. Biol.* **2017**, *12*, 214−224.

(7) Ji, Y.; Sasmal, A.; Li, W.; Oh, L.; Srivastava, S.; Hargett, A. A.; Wasik, B. R.; Yu, H.; Diaz, S.; Choudhury, B.; Parrish, C. R.; Freedberg, D. I.; Wang, L.-P.; Varki, A.; Chen, X. Reversible O-Acetyl Migration within the Sialic Acid Side Chain and Its Influence on Protein Recognition. *ACS Chem. Biol.* **2021**, *16*, 1951−1960.

(8) Li, W.; Battistel, M. D.; Reeves, H.; Oh, L.; Yu, H.; Chen, X.; Wang, L.-P.; Freedberg, D. I. A Combined NMR, MD and DFT Conformational Analysis of 9-O-Acetyl Sialic Acid-Containing GM3 Ganglioside Glycan and Its 9-N-Acetyl Mimic. *Glycobiology* **2020**, *30*, 787−801.

(9) Pietrucci, F.; Andreoni, W. Graph Theory Meets Ab Initio Molecular Dynamics: Atomic Structures and Transformations at the Nanoscale. *Phys. Rev. Lett.* **2011**, *107*, 085504.

(10) Wang, L.-P.; Titov, A.; McGibbon, R.; Liu, F.; Pande, V. S.; Martínez, T. J. Discovering Chemistry with an Ab Initio Nanoreactor. *Nat. Chem.* **2014**, *6*, 1044−1048.

(11) Bergeler, M.; Simm, G. N.; Proppe, J.; Reiher, M. Heuristics-Guided Exploration of Reaction Mechanisms. *J. Chem. Theory Comput.* **2015**, *11*, 5712−5722.

(12) Wang, L.-P.; McGibbon, R. T.; Pande, V. S.; Martinez, T. J. Automated Discovery and Refinement of Reactive Molecular Dynamics Pathways. *J. Chem. Theory Comput.* **2016**, *12*, 638−649.

(13) Lapkin, A. A.; Heer, P. K.; Jacob, P.-M.; Hutchby, M.; Cunningham, W.; Bull, S. D.; Davidson, M. G. Automation of Route Identification and Optimisation Based on Data-Mining and Chemical Intuition. *Faraday Discuss.* **2017**, *202*, 483−496.

(14) Varela, J. A.; Vázquez, S. A.; Martínez-Núñez, E. An Automated Method to Find Reaction Mechanisms and Solve the Kinetics in Organometallic Catalysis. *Chem. Sci.* **2017**, *8*, 3843−3851.

(15) Jafari, M.; Zimmerman, P. M. Uncovering Reaction Sequences on Surfaces through Graphical Methods. *Phys. Chem. Chem. Phys.* **2018**, *20*, 7721−7729.

(16) Plehiers, P. P.; Marin, G. B.; Stevens, C. V.; Van Geem, K. M. Automated Reaction Database and Reaction Network Analysis: Extraction of Reaction Templates Using Cheminformatics. *J. Cheminf.* **2018**, *10*, 11.

(17) Unsleber, J. P.; Reiher, M. The Exploration of Chemical Reaction Networks. *Annu. Rev. Phys. Chem.* **2020**, *71*, 121−142.

(18) Dewyer, A. L.; Argüelles, A. J.; Zimmerman, P. M. Methods for Exploring Reaction Space in Molecular Systems. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, No. e1354.

(19) Iannuzzi, M.; Laio, A.; Parrinello, M. Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics. *Phys. Rev. Lett.* **2003**, *90*, 238302.

(20) Das, T.; Ghule, S.; Vanka, K. Insights Into the Origin of Life: Did It Begin from HCN and H2O? *ACS Cent. Sci.* **2019**, *5*, 1532−1540.

(21) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A Package for Building Initial Configurations for Molecular Dynamics Simulations. *J. Comput. Chem.* **2009**, *30*, 2157−2164.

(22) Ufimtsev, I. S.; Martínez, T. J. Quantum Chemistry on Graphical Processing Units. 1. Strategies for Two-Electron Integral Evaluation. *J. Chem. Theory Comput.* **2008**, *4*, 222−231.

(23) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 2. Direct Self-Consistent-Field Implementation. *J. Chem. Theory Comput.* **2009**, *5*, 1004−1015.

(24) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry

Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2619−2628.
(25) Liu, F.; Sanchez, D. M.; Kulik, H. J.; Martínez, T. J. Exploiting Graphical Processing Units to Enable Quantum Chemistry Calculation of Large Solvated Molecules with Conductor-like Polarizable Continuum Models. *Int. J. Quantum Chem.* **2019**, *119*, No. e25760.
(26) Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799−805.
(27) Liu, F.; Luehr, N.; Kulik, H. J.; Martínez, T. J. Quantum Chemistry for Solvated Molecules on Graphical Processing Units Using Polarizable Continuum Models. *J. Chem. Theory Comput.* **2015**, *11*, 3131−3144.
(28) Lange, A. W.; Herbert, J. M. A Smooth, Nonsingular, and Faithful Discretization Scheme for Polarizable Continuum Models: The Switching/Gaussian Approach. *J. Chem. Phys.* **2010**, *133*, 244111.
(29) Lin, Y.-S.; Li, G.-D.; Mao, S.-P.; Chai, J.-D. Long-Range Corrected Hybrid Density Functionals with Improved Dispersion Corrections. *J. Chem. Theory Comput.* **2013**, *9*, 263−272.
(30) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(31) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.
(32) Jonsson, H.; Mills, G.; Jacobsen, K. W. Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions. *Classical and Quantum Dynamics in Condensed Phase Simulations*; World Scientific; 1998; Chapter 16.
(33) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *J. Chem. Phys.* **2000**, *113*, 9901−9904.
(34) Wang, L.-P.; Song, C. Geometry Optimization Made Simple with Translation and Rotation Coordinates. *J. Chem. Phys.* **2016**, *144*, 214108.
(35) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2619−2628.
(36) Titov, A. V.; Ufimtsev, I. S.; Luehr, N.; Martinez, T. J. Generating Efficient Quantum Chemistry Codes for Novel Architectures. *J. Chem. Theory Comput.* **2013**, *9*, 213−221.
(37) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package. *Mol. Phys.* **2015**, *113*, 184−215.
(38) Rao, G.; Chen, N.; Marchiori, D. A.; Wang, L.-P.; Britt, R. D. Accumulation and Pulse Electron Paramagnetic Resonance Spectroscopic Investigation of the 4-Oxidobenzyl Radical Generated in the Radical S-Adenosyl-l-Methionine Enzyme HydG. *Biochemistry* **2022**, *61*, 107−116.
(39) Thapa, B.; Schlegel, H. B. Improved PKa Prediction of Substituted Alcohols, Phenols, and Hydroperoxides in Aqueous Medium Using Density Functional Theory and a Cluster-Continuum Solvation Model. *J. Phys. Chem. A* **2017**, *121*, 4698.
(40) Maheshwary, S.; Patel, N.; Sathyamurthy, N.; Kulkarni, A. D.; Gadre, S. R. Structure and Stability of Water Clusters (H2O)n, n = 8−20: An Ab Initio Investigation. *J. Phys. Chem. A* **2001**, *105*, 10525−10537.
(41) Qi, R.; Wang, L.-P.; Wang, Q.; Pande, V. S.; Ren, P. United Polarizable Multipole Water Model for Molecular Mechanics Simulation. *J. Chem. Phys.* **2015**, *143*, 014504.
(42) Wang, L.-P.; Wu, Q.; Van Voorhis, T. Acid−Base Mechanism for Ruthenium Water Oxidation Catalysts. *Inorg. Chem.* **2010**, *49*, 4543−4553.
(43) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.
(44) Grossfield, A. *WHAM: An Implementation of the Weighted Histogram Analysis Method*, Version 2.0.10. http://Membrane.Urmc.Rochester.Edu/Content/Wham/.
(45) Efron, B.; Tibshirani, R. *An Introduction to the Bootstrap*; Monographs on statistics and applied probability; Chapman & Hall/CRC: New York, 1994.
(46) Plückthun, A.; Dennis, E. A. Acyl and Phosphoryl Migration in Lysophospholipids: Importance in Phospholipid Synthesis and Phospholipase Specificity. *Biochemistry* **1982**, *21*, 1743−1750.
(47) Chokhawala, H. A.; Yu, H.; Chen, X. High-Throughput Substrate Specificity Studies of Sialidases by Using Chemoenzymatically Synthesized Sialoside Libraries. *ChemBioChem* **2007**, *8*, 194−201.
(48) Lassfolk, R.; Pedrón, M.; Tejero, T.; Merino, P.; Wärnå, J.; Leino, R. Acyl Group Migration in Pyranosides as Studied by Experimental and Computational Methods. *Chem.—Eur. J.* **2022**, *28*, No. e202200499.
(49) Sergeant, E. P.; Dempsey, B. *Ionisation Constants of Organic Acids in Aqueous Solution*, IUPAC Chem Data Ser No. 23; Pergamon Press; Oxford: New York, 1979
(50) Sherwood, P.; Brooks, B. R.; Sansom, M. S. Multiscale Methods for Macromolecular Simulations. *Curr. Opin. Struct. Biol.* **2008**, *18*, 630−640.
(51) Lee, J.; Miller, B. T.; Brooks, B. R. Computational Scheme for PH-Dependent Binding Free Energy Calculation with Explicit Solvent. *Protein Sci.* **2016**, *25*, 231−243.
(52) Jia, X.; Wang, M.; Shao, Y.; König, G.; Brooks, B. R.; Zhang, J. Z. H.; Mei, Y. Calculations of Solvation Free Energy through Energy Reweighting from Molecular Mechanics to Quantum Mechanics. *J. Chem. Theory Comput.* **2016**, *12*, 499−511.

https://doi.org/10.1021/acs.biochem.2c00343
Biochemistry XXXX, XXX, XXX−XXX