

# Reaction Mechanism of *N*-Acetylneuraminic Acid Lyase Revealed by a Combination of Crystallography, QM/MM Simulation, and Mutagenesis

Adam D. Daniels,<sup>†,||</sup> Ivan Campeotto,<sup>†,||,⊥</sup> Marc W. van der Kamp,<sup>‡</sup> Amanda H. Bolt,<sup>†</sup> Chi H. Trinh,<sup>†</sup> Simon E. V. Phillips,<sup>†,§</sup> Arwen R. Pearson,<sup>†</sup> Adam Nelson,<sup>§</sup> Adrian J. Mulholland,<sup>\*,‡</sup> and Alan Berry<sup>\*,†</sup>

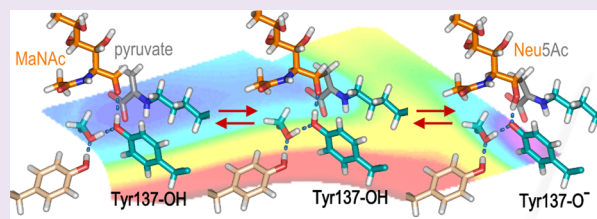
<sup>†</sup>Astbury Centre for Structural Molecular Biology and School of Molecular and Cellular Biology, University of Leeds, Leeds LS2 9JT, U.K.

<sup>‡</sup>Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, U.K.

<sup>§</sup>Astbury Centre for Structural Molecular Biology and School of Chemistry, University of Leeds, Leeds LS2 9JT, U.K.

## S Supporting Information

**ABSTRACT:** *N*-Acetylneuraminic acid lyase (NAL) is a Class I aldolase that catalyzes the reversible condensation of pyruvate with *N*-acetyl-*D*-mannosamine (ManNAc) to yield the sialic acid *N*-acetylneuraminic acid (Neu5Ac). Aldolases are finding increasing use as biocatalysts for the stereospecific synthesis of complex molecules. Incomplete understanding of the mechanism of catalysis in aldolases, however, can hamper development of new enzyme activities and specificities, including control over newly generated stereocenters. In the case of NAL, it is clear that the enzyme catalyzes a Bi-Uni ordered condensation reaction in which pyruvate binds first to the enzyme to form a catalytically important Schiff base. The identity of the residues required for catalysis of the condensation step and the nature of the transition state for this reaction, however, have been a matter of conjecture. In order to address this we crystallized a Y137A variant of the *E. coli* NAL in the presence of Neu5Ac. The three-dimensional structure shows a full length sialic acid bound in the active site of subunits A, B, and D, while in subunit C, discontinuous electron density reveals the positions of enzyme-bound pyruvate and ManNAc. These 'snapshot' structures, representative of intermediates in the enzyme catalytic cycle, provided an ideal starting point for QM/MM modeling of the enzymic reaction of carbon–carbon bond formation. This revealed that Tyr137 acts as the proton donor to the aldehyde oxygen of ManNAc during the reaction, the activation barrier is dominated by carbon–carbon bond formation, and proton transfer from Tyr137 is required to obtain a stable Neu5Ac-Lys165 Schiff base complex. The results also suggested that a triad of residues, Tyr137, Ser47, and Tyr110 from a neighboring subunit, are required to correctly position Tyr137 for its function, and this was confirmed by site-directed mutagenesis. This understanding of the mechanism and geometry of the transition states along the C–C bond-forming pathway will allow further development of these enzymes for stereospecific synthesis of new enzyme products.



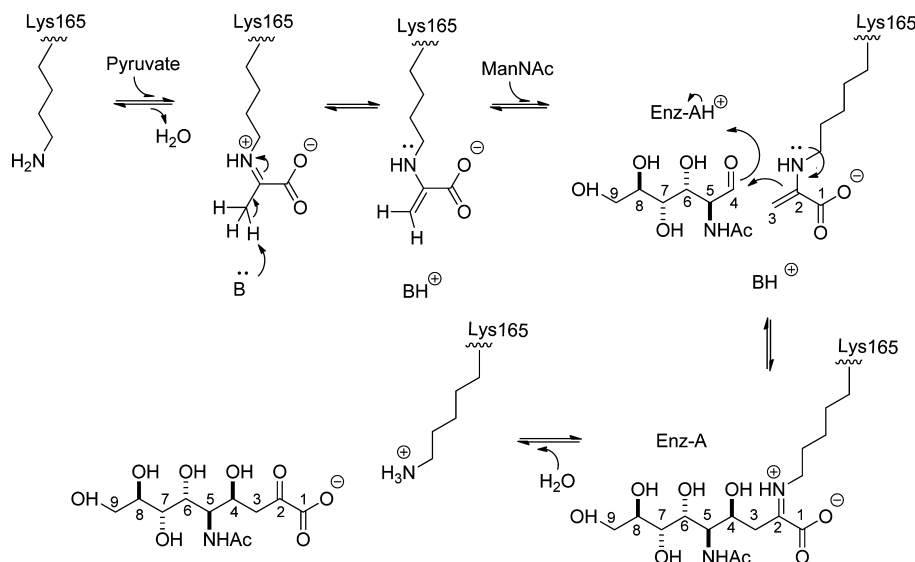
Aldolases catalyze a fundamental reaction in biology, namely, the formation or cleavage of carbon–carbon bonds. Examples are found in many pathways of both central and secondary metabolism. The basic aldol reaction involves the condensation of an aldehyde (the aldol acceptor) with a ketone (the aldol donor) to form the new carbon–carbon bond. Aldolases are generally highly stereoselective in their catalysis and form the new carbon–carbon bond while controlling the configuration of up to two new chiral centers. This ability to form new carbon–carbon bonds under mild conditions without the need for elaborate protecting group chemistry, while controlling the stereoselectivity of aldol condensation, has made the aldolase family an important group of enzymes for use in biocatalysis. Indeed, protein engineering and directed evolution studies have been directed against a number of aldolases to improve their utility in chemical synthesis.<sup>1</sup>

*N*-Acetyl-*D*-neuraminic acid lyase (NAL) (EC 4.1.3.3) is a pyruvate-dependent aldolase that catalyzes the reversible aldol condensation between *N*-acetyl-*D*-mannosamine (ManNAc) and pyruvate to yield *N*-acetyl-*D*-neuraminic acid (Neu5Ac), the most abundant of the sialic acids. The enzyme is a member of the Class I aldolase family in which the reaction proceeds through a Schiff base intermediate between a lysine and the substrate pyruvate.<sup>2</sup> NAL has been shown to follow an ordered Bi-Uni aldol condensation kinetic mechanism<sup>3</sup> (Scheme 1) in which pyruvate binds first as a Schiff base, followed by ManNAc. Aldol addition then produces Schiff base-bound Neu5Ac, which is then hydrolyzed to release Neu5Ac. A number of crystallographic studies<sup>4–9</sup> have determined three-

Received: August 16, 2013

Accepted: February 12, 2014

Published: February 12, 2014

Scheme 1. General Reaction Mechanism of NAL<sup>a</sup>

<sup>a</sup>The ordered sequential Bi-Uni reaction leading to the formation of NeuSAc is illustrated. After enamine formation between pyruvate and Lys-165, nucleophilic attack of the pyruvate-Schiff base onto the aldehyde carbon of ManNAc and protonation of the aldehyde oxygen generates the enzyme NeuSAc-Schiff base, which is hydrolyzed to yield NeuSAc as product. The reaction is reversible. The numbering scheme for NeuSAc is shown, and for ease of following the fate of the substrates these numbers are also used in the enzyme-Pyr-ManNAc complex.

dimensional NAL structures containing bound substrates and substrate analogues, including pyruvate,<sup>5,9</sup> a pyruvate analogue,<sup>7</sup> and a NeuSAc analogue.<sup>4</sup> These structures, while informative, provide only limited information on the enzyme mechanism since they do not reveal the transitions between these low energy, stable states. Obtaining atomic-resolution structural information for high energy transition states is impossible by experiment, but such information is available through computational modeling. Quantum mechanical/molecular mechanical (QM/MM) methods, in particular, have been very effective in modeling enzyme reactions.<sup>10</sup> Here we use a combined crystallographic and QM/MM approach to follow carbon-carbon bond formation catalyzed by NAL, with a view to revealing how substrates are bound and interact with the enzyme and determine the stereochemical outcome of the reaction.

The formation of the carbon-carbon bond between ManNAc and pyruvate depends on not only the proximity and orientation of the two bonding carbon atoms but also the transfer of a proton onto the aldehyde oxygen of ManNAc. The determination of several three-dimensional NAL structures revealed a lack of acidic or basic residues near the reaction center and led to a suggestion that the pyruvate carboxylate group could be involved in a "substrate-assisted" protonation step.<sup>11</sup> In addition, the three-dimensional structure of *Haemophilus influenzae* NAL with a substrate analogue, 4-oxo-NeuSAc, bound as a Schiff base with the catalytic lysine, revealed an interaction between the inhibitor C-4 hydroxyl group and the hydroxyl group of a tyrosine residue (Tyr136 in *H. influenzae* NAL).<sup>4</sup> This led to the hypothesis that this tyrosine mediates a proton transfer to and from the substrate carboxylate group.<sup>4,12</sup> The importance of this tyrosine was supported by the observation that its mutation to phenylalanine (Y130F) in the *Clostridium perfringens* NAL (equivalent to Tyr136 in *H. influenzae* NAL and Tyr137 in *E. coli* NAL (*E. coli* numbers used throughout the rest of this manuscript)) resulted in a significant decrease in activity.<sup>13</sup> Here we use

crystallography, QM/MM modeling, and mutagenesis to explore this proposed mechanism in further mechanistic detail.

## RESULTS AND DISCUSSION

**Crystal Structures.** The ideal starting point for QM/MM modeling of the formation or breakage of the carbon-carbon bond between ManNAc and pyruvate in NAL would be a crystal structure corresponding to either the ManNAc and pyruvate-bound structure or the NeuSAc-bound structure. No such structures had been reported until this study. The structure of the *H. influenzae* NAL containing the substrate analogue 4-oxo-NeuSAc<sup>4</sup> is a possible starting structure for QM/MM simulation but does not necessarily represent the true binding mode of NeuSAc because the C4 atom is sp<sup>2</sup> hybridized in this structure, rather than the sp<sup>3</sup> hybridization of the natural substrate, NeuSAc. Soaking crystals of wild-type *E. coli* NAL with NeuSAc always resulted in structures of the enzyme-pyruvate complex,<sup>5</sup> most likely due to rapid turnover of NeuSAc and unbinding of ManNAc.

We reasoned that inactivation of NAL by a Y137A mutation might produce an enzyme that could be crystallized with NeuSAc to act as a starting point for our modeling studies. Both Y137A and Y137F NAL variants were constructed by site-directed mutagenesis and expressed in *E. coli* as for the wild-type enzyme.<sup>14</sup> Steady-state enzyme kinetics showed that, as expected, the variants were severely impaired with only ca. 0.2% activity (as judged by  $k_{\text{cat}}/K_m$ ) for the Y137F variant and no detectable activity for the Y137A variant (Table 1). The Y137A variant was crystallized in the same conditions and space group as the wild-type enzyme. The crystals were then soaked with 75 mM NeuSAc, and the three-dimensional structure was solved (PDB code 4BWL; see Supplementary Table S1 for crystallographic data) by direct Fourier methods using the wild-type *E. coli* structure<sup>5</sup> as a starting model. In each subunit of the Y137A variant, additional electron density contiguous with the catalytic lysine residue, Lys165, was found, demonstrating that, despite the fact that no activity had been detected for this variant, the

**Table 1. Steady-State Kinetic Parameters of *E. coli* Tyr137 NAL Variants**

enzyme	$k_{\text{cat}}$ ( $\text{min}^{-1}$ )	$K_{\text{m}}$ (mM)	$k_{\text{cat}}/K_{\text{m}}$ ( $\text{min}^{-1} \text{mM}^{-1}$ )
wild-type	$250 \pm 5$	$2.2 \pm 0.1$	114
Y137F	$0.8 \pm 0.1$	$3.7 \pm 1.1$	0.2
Y137A	ND	ND	ND

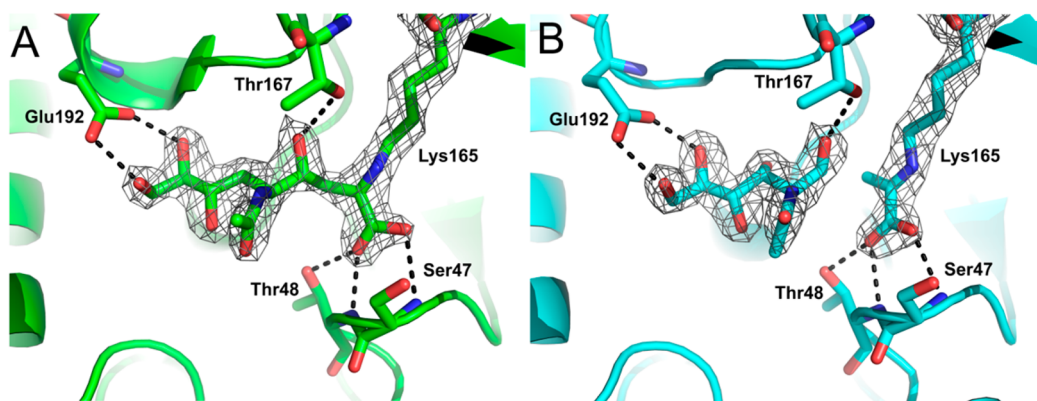
<sup>a</sup>Steady-state kinetic parameters for the cleavage of *N*-acetyl neuraminic acid (Neu5Ac) by wild-type and variant *E. coli* NAL determined using an LDH coupled enzyme assay.<sup>14,15</sup> Data were fitted to the Michaelis–Menten equation, and the fitted values  $\pm$  standard error of the fit are shown. ND = not detectable.

enzyme was still able to form a Schiff base with incoming substrates. This indicates that Tyr137 is not critical in forming the covalent enzyme–pyruvate complex. In subunits A, B, and D of the tetramer, the electron density observed in the active site was continuous and of the approximate size and shape to be fitted by Neu5Ac (Figure 1A). In subunit C, however, two distinct areas of electron density were visible, one similar to the NAL–pyruvate complex previously seen<sup>5</sup> and one of the approximate size and shape of ManNAc (Figure 1B). Extensive model building and testing (see Supporting Information) revealed that soaking with Neu5Ac in the Y137A variant serendipitously produced a structure of the Enz–ManNAc–pyruvate complex in subunit C. This result arises presumably because although we could not detect activity in our standard kinetic assays (Table 1), the Y137A mutant has an extremely low level of activity in the crystal sufficient for turnover to take place. Perhaps even more surprisingly, the continuous electron density found in subunits A, B, and D was consistent only with the 4-epimer of Neu5Ac (4-epi-Neu5Ac) rather than Neu5Ac (see Supporting Information for details of model fitting and testing). This result is unexpected because of the stereo-selectivity of the wild-type enzyme where the stereochemical outcome of the NAL condensation reaction depends on the structure of the substrate.<sup>25</sup> In the case of the wild-type enzyme acting on ManNAc and pyruvate, the substrate aldehyde group is attacked from the *si* face, generating a new stereocenter with 4*S*-configuration, whereas with other substrate pairs, *re*-face attack and subsequent 4*R*-configured products are found. This lack of overall stereochemical control together with <sup>1</sup>H NMR studies of product distribution during the course of the

reactions with different substrates<sup>26</sup> implies that NAL can bind substrate in different conformations. We suggest that in the crystals of the Y137A variant, cleavage of Neu5Ac and reformation of the carbon–carbon bond with the opposite stereochemical configuration at C4 must be possible, leading to 4-epi-Neu5Ac being trapped. In this respect it is important to note that the kinetics of various steps in the enzyme reaction in the cryo-cooled crystalline state may well be different from the kinetics measured in solution at RT,<sup>27</sup> leading to trapping of different intermediates. However, it is interesting to note that 4-epi-Neu5Ac can bind to NAL active site as it has been shown to be a competitive inhibitor of the wild-type *Clostridium perfringens* NAL.<sup>28</sup>

In all four subunits of the crystal structure, the carboxylate moiety of the substrate forms the same hydrogen bonding pattern with the backbone of Thr48 and Ser47 and with the hydroxyl group of Thr48 as found in the wild-type enzyme–pyruvate complex (PDB code 2WNN) (Figure 1A and B).<sup>5</sup> In subunits A, B, and D, the *N*-acetyl group of 4-epi-Neu5Ac at C5 lies close ( $\sim 3.4$  Å) to the side chain of Phe252 but does not appear to form any hydrogen bonds with the surrounding residues, with the exception of chain B, where it forms a hydrogen bond with an ordered water molecule. This is in agreement with the tolerance of NAL for substitutions at C5.<sup>25</sup> The hydroxyl groups of the C6–C9 moiety of 4-epi-Neu5Ac form slightly different hydrogen bonds with Ser208, Asp191, Glu192, and the backbone of Gly189 in the three subunits, consistent with a degree of flexibility of this portion of the substrate (see Supplementary Figure S2). The C4 hydroxyl group of 4-epi-Neu5Ac results from the protonation of the aldehyde oxygen of ManNAc, and it is interesting to note that this group forms a hydrogen bond with Thr167 in all of the subunits A, B, and D (Figure 1)

In subunit C, ManNAc and pyruvate are bound. The planarity of the electron density of the covalent Lys165–pyruvate is consistent with the pyruvate enamine complex previously seen in crystal structures of NAL.<sup>5</sup> The hydrogen bond network of ManNAc differs slightly from that of the C4–C9 portion of 4-epi-Neu5Ac in chains A, B, and D. The hydroxyl groups at carbons positions C4, C7, C8, and C9 of ManNAc (Neu5Ac numbering) form a similar hydrogen bond network as the corresponding hydroxyl groups of 4-epi-



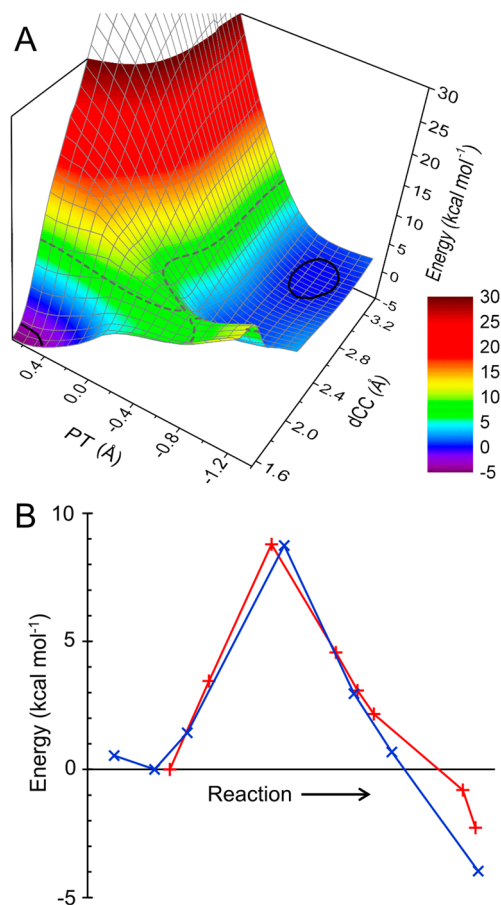
**Figure 1.** Ligand binding in the active site of the Y137A NAL crystal structure. (A) Continuous  $2F_o - F_c$  electron density, contoured at 2 RMSD around the Schiff base of Lys165 covalently linked to 4-epi-Neu5Ac in subunit A of the solved structure (PDB code 4BWL). (B)  $2F_o - F_c$  map (at the same contour level) for subunit C. The discontinuous electron density in subunit C reveals that cleavage of Neu5Ac has occurred, producing the Lys165–pyruvate Schiff base and ManNAc. Hydrogen bonds between the substrate and Glu192, Thr167, Thr48, and Ser47 are shown. For clarity other hydrogen bonds between the enzyme and substrate(s) are omitted.



Neu5Ac, but a shift of the hydrogen bond network is seen at C6 mainly due to the different orientation of ManNAc, which interacts with Asp191 and Ser208 instead of Ser208 alone.

**QM/MM Modeling.** After *in silico* mutation of Ala137 back to Tyr, the structure found in subunit C, with pyruvate and ManNAc bound, was used as a starting point for QM/MM modeling of the reaction mechanism. Initially, we modeled the carbon–carbon bond formation between ManNAc and pyruvate by performing QM/MM umbrella sampling molecular dynamics (MD) simulations along  $d(\text{C}–\text{C})$ , the distance between the carbons involved (the ManNAc aldehyde carbon atom and the carbon of the methyl group of pyruvate). ManNAc, the pyruvate-Schiff base (modeled in its enamine form) and the Tyr137 side chain were included in the QM region. Five series of umbrella sampling simulations (with different starting structures originating from unbiased QM/MM MD simulations) were performed. In all cases, the Tyr137 hydroxyl proton transferred spontaneously onto the ManNAc aldehyde oxygen, producing a Neu5Ac-Lys165 Schiff base in the final structures. Free energy profiles obtained from the umbrella sampling (Supplementary Figure S3) indicate that this intermediate is approximately 4–8 kcal mol<sup>−1</sup> lower in energy than the transition state. The spontaneity of the proton transfer and the apparently low free energy of the intermediate indicate that Tyr137 is the proton donor in the wild-type NAL reaction, consistent with previous speculations.<sup>4,12</sup> To test whether water might also be involved as a proton donor in the reaction, we repeated the simulations including an additional water molecule, in proximity to Tyr137 and the ManNAc aldehyde oxygen, in the QM region. In our simulations, this water molecule often moved away from this position, and proton transfer from Tyr137 onto the ManNAc aldehyde oxygen took place as before. It is therefore unlikely that water is directly involved in this reaction step. Water however is certainly required for the breakdown of the Schiff base between Lys165 and the substrate/product and may also serve to reprotonate the Tyr137 phenolate. In every umbrella sampling series along  $d(\text{C}–\text{C})$ , the ManNAc aldehyde underwent attack on the *si* face, producing the 4*S*-configured diastereoisomer. This is consistent with the previously reported high stereoselectivity of wild-type NAL with pyruvate and ManNAc.<sup>25,28</sup> Analysis of the MD trajectories revealed that the attacking pyruvate carbon is almost always positioned on the *si* face, and this preference becomes more pronounced as the carbon atoms approach each other. The simulations therefore strongly suggest that the ManNAc binding mode in wild-type NAL is the main contributory factor to the high stereoselectivity reported.<sup>25,28</sup>

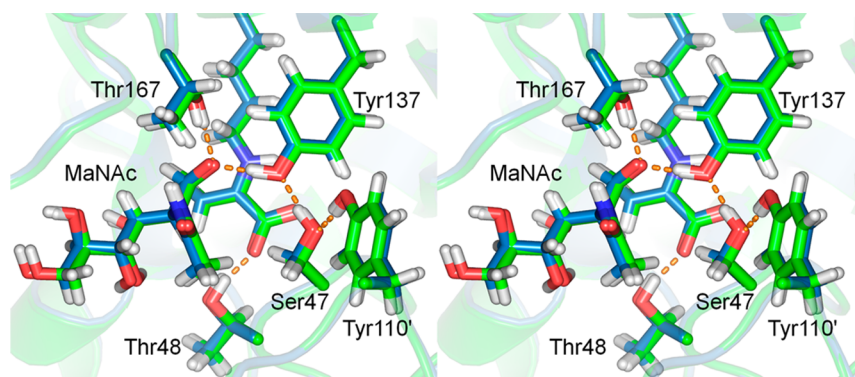
To model the mechanism of bond formation between ManNAc and pyruvate, QM/MM potential energy surfaces (PESs) of the reaction were obtained along two reaction coordinates:  $d(\text{C}–\text{C})$  and  $PT$ , the proton transfer between Tyr137 and the ManNAc aldehyde (defined as the difference between the proton-donor and proton-acceptor distances) (Figure 2). Seven PESs at the SCC-DFTB/CHARMM22 level were obtained from different enzyme–substrate conformations of the Neu5Ac-Lys165 Schiff base intermediate as sampled from QM/MM MD (Figure 2A, Supplementary Figure S4). These multiple PESs allowed for examination of conformational effects,<sup>23,29</sup> to test reproducibility and to provide starting points for reaction pathway optimization at higher levels of QM/MM theory (see below). Minimum energy paths along these PESs indicated three of the PESs had approximate activation energies (differences between the reactant minimum



**Figure 2.** QM/MM potential energy profiles for the carbon–carbon bond formation between ManNAc and the pyruvate-Lys165 Schiff base, with Tyr137 as proton donor. (A) SCC-DFTB/CHARMM22 potential energy surface along two reaction coordinates: the carbon–carbon distance,  $d_{\text{CC}}$ , and the proton transfer between Tyr137 and the ManNAc aldehyde,  $PT$ . Minima and approximate transition state energy are indicated by solid black and dashed gray lines, respectively. (B) Potential energies from two reaction path optimizations (blue and red, from ManNAc and pyruvate-Lys165 Schiff base on the left to the Neu5Ac-Lys165 Schiff base on the right). X-axis values are calculated as a combined reaction coordinate from optimized structures:  $((PT \text{ minus } d(\text{C}–\text{C})), \text{ see text})$ . Structures were optimized at the B3LYP/6-31+G(d)/CHARMM22 level with energies calculated at the SCS-MP2/aug-cc-pVDZ/CHARMM22 level. (See Supplementary Figure S5 for B3LYP/MM energies.)

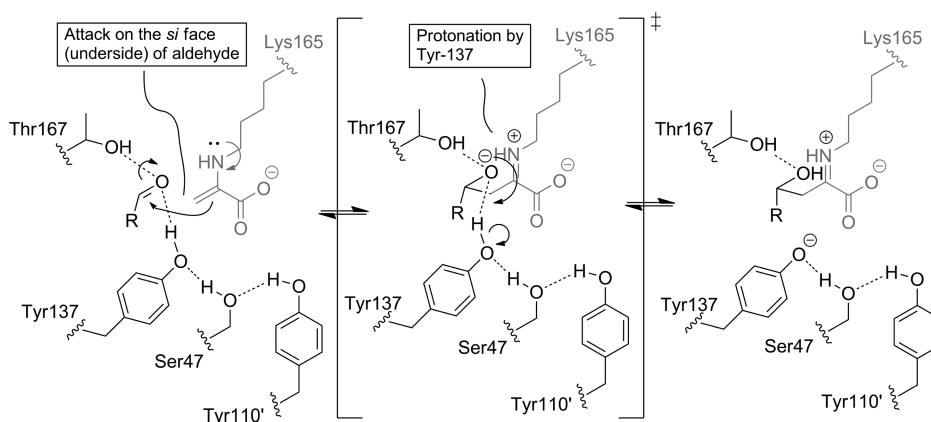
and the approximate transition state) of 10.1–12.0 kcal mol<sup>−1</sup> and four PESs with approximate activation energies of 6.2–7.7 kcal mol<sup>−1</sup>. The four PESs with lower activation energies all showed the same hydrogen bonding network (see also Figure 3): the Tyr137 hydroxyl oxygen accepts a hydrogen bond from Ser47 O $\gamma$ H and Ser47 O $\gamma$  accepts a hydrogen bond from the hydroxyl oxygen of a tyrosine residue in an adjacent subunit (Tyr110'). In the other three PESs, the hydrogen bond between Ser47 O $\gamma$ H and the Tyr137 hydroxyl oxygen was not formed. All seven PESs indicate that the carbon–carbon bond formation and proton transfer are asynchronous: proton transfer occurs only when  $d(\text{C}–\text{C})$  is 1.7 Å or smaller.

To confirm the mechanism and obtain more accurate reaction energetics and transition state structures, two reaction path optimizations were performed at the density functional theory (DFT) (B3LYP/6-31+G(d)/CHARMM22) QM/MM level, using starting conformations from two PESs with the



**Figure 3.** Stereo image showing the transition state structures (TSs) for the carbon–carbon bond formation between ManNAc and the pyruvate-Lys165 Schiff base, optimized at the B3LYP/6-31+G(d)/CHARMM22 level. The TS structure for the first profile is shown with green carbons (and hydrogen bonds in orange), and the TS structure for the second profile with light blue carbons. Side chains for the residues targeted in the mutagenesis studies are labeled.

**Scheme 2. Mechanism of Carbon–Carbon Bond Formation and Protonation during NAL Catalysis As Derived from the Crystal Structure of the Y137A Variant in the Presence of Pyruvate and ManNAc and QM/MM Modelling<sup>a</sup>**



<sup>a</sup>The pyruvate-Lys165 enamine is shown in grey as it lies below the plane of the ManNAc aldehyde. Hydrogen bonds are shown as dotted lines. The central structure is unstable and indicated to be a transition state by the modelling (see Figure 3).

Tyr137-Ser47-Tyr110' hydrogen bond network in place (Figure 2B). After geometry optimization, frequency calculations of the reactant, transition, and intermediate states verified that the structures were true minima (no imaginary frequency) and transition states (one imaginary frequency) and gave zero-point energy corrections (Supplementary Table S2). Energies of the optimized conformations were also calculated at the SCS-MP2/aug-cc-pVDZ/CHARMM22 QM/MM level; this *ab initio* method has been shown to give results in good agreement with 'gold standard' coupled-cluster theory and provides a high-level test of the DFT results.<sup>23,29</sup>

Zero-point corrected activation energies at the SCS-MP2/aug-cc-pVDZ/CHARMM22 level are 9.5 and 10.0 kcal mol<sup>-1</sup> for the two profiles, respectively. The activation free energy is likely to be higher due to entropy (as expected for a bond-forming reaction). The transition states of both pathways are very similar (Figure 3) and occur during the carbon–carbon bond formation ( $d(\text{C}–\text{C})$  at 2.13 and 2.04 Å, respectively), before proton transfer from Tyr137 takes place ( $d(\text{O}–\text{H})$  of the Tyr137 hydroxyl is 1.00 Å in both cases). The reaction paths obtained at the higher QM/MM level confirm that Tyr137 is the proton donor in the condensation reaction and that the deprotonated Tyr137 is stable in the context of the enzyme active site with a Neu5Ac-Lys165 Schiff base.

**Enzyme Mechanism and Mutagenesis.** Based on the crystal structure presented here and the QM/MM simulations, a detailed mechanism for formation of the carbon–carbon bond formation between ManNAc and pyruvate can be proposed (Scheme 2).

The enamine form of pyruvate, in Schiff-base complex with Lys165, specifically attacks the *si* face of the ManNAc aldehyde to form Neu5Ac. The increasing electron density around the (former) ManNAc aldehyde oxygen (due to the emerging tetrahedral conformation around the aldehyde carbon) is stabilized by hydrogen bonds from the Thr167 and Tyr137 hydroxyls. Once the carbon–carbon bond is (almost) formed, the (former) ManNAc aldehyde oxygen is protonated by the Tyr137 hydroxyl group. The resulting Tyr137 phenolate is stabilized by a hydrogen bond donated from the Ser47 hydroxyl, which may in turn be supported by a hydrogen donated to the Ser Oγ from the Tyr110' hydroxyl. The QM/MM energies indicate that this Neu5Ac Schiff base intermediate with a Tyr137 phenolate is quite stable. It is therefore possible that this phenolate exists until Neu5Ac release allows its reprotonation by solvent.

The triad of Tyr110', Ser47, and Tyr137 is similar to the Tyr107', Thr44, Tyr133 triad in the related ( $\alpha/\beta$ )<sub>8</sub> TIM barrel lyase, dihydrodipicolinate synthase (DHDPS)<sup>30,31</sup> and to

Tyr103', Thr43, and Tyr130 in another pyruvate-dependent aldolase, 2-keto-3-deoxygluconate aldolase.<sup>32</sup> In DHDPs these residues are thought to act as a proton shuttle to assist Tyr133 in its role as a proton donor during the DHDPs reaction. The presence of this catalytic triad in related enzymes provides strong evidence that the identified catalytic triad in wild-type *E. coli* NAL also plays a key role in its catalytic mechanism. In order to test this hypothesis further the active site residues identified here (Ser47, Tyr110, Tyr137, and Thr167) along with two other residues close to this area of the active site (Thr48, which interacts with the pyruvate carboxylate, and Phe252, which lies close to the *N*-acetyl of ManNAc) were targeted for site-directed mutagenesis. The two variants at Tyr137 (Y137A and Y137F) (Table 1) are extremely detrimental to activity, and it is now clear that this is because Tyr137 plays a direct, crucial role in the donation of a proton during the carbon–carbon bond formation between ManNAc and pyruvate. The importance of Thr167 in stabilizing the transition state by hydrogen bonding to the (former) aldehyde oxygen is confirmed by the mutagenesis results: removal of the possibility of H-bonding in the T167A mutant results in a 4-fold decrease in  $k_{\text{cat}}$ , whereas the replacement with serine (T167S) left the kinetic parameters unaffected (Table 2).

**Table 2. Steady-State Kinetic Parameters of *E. coli* NAL Variants<sup>a</sup>**

enzyme	$k_{\text{cat}}$ (min <sup>-1</sup> )	$K_{\text{m}}$ (mM)	$k_{\text{cat}}/K_{\text{m}}$ (min <sup>-1</sup> mM <sup>-1</sup> )
wild-type	250 ± 5	2.2 ± 0.1	114
S47A	18 ± 0.4	3.4 ± 0.2	5.3
S47C	13 ± 0.3	4.7 ± 0.3	2.8
S47T	340 ± 7.8	2.9 ± 0.2	120
T48A	510 ± 17	3.0 ± 0.3	170
T48S	680 ± 15	3.4 ± 0.2	200
Y110A	12 ± 0.8	4.3 ± 0.7	2.8
Y110F	170 ± 1.9	1.8 ± 0.1	94
T167A	71 ± 0.7	2.0 ± 0.1	36
T167S	200 ± 2.8	1.9 ± 0.1	110
F252A	260 ± 9.9	3.1 ± 0.3	84
F252Y	120 ± 2.3	1.3 ± 0.1	92

<sup>a</sup>Steady-state kinetic parameters for the cleavage of *N*-acetyl neuraminic acid (Neu5Ac) by wild-type and variant *E. coli* NAL determined using an LDH coupled enzyme assay.<sup>14,15</sup> Data were fitted to the Michaelis–Menten equation, and the fitted values ± standard error of the fit are shown.

Whereas mutations of active site residues postulated not to be involved in the catalytic mechanism (Thr48 and Phe252) had no or only minor effects on the kinetic behavior of the enzyme, replacement of Ser47 with alanine or cysteine (thus removing the potential to form the natural hydrogen bonding network) had a significant effect. However, in line with its postulated importance in a hydrogen bond network, the replacement of Ser47 with a threonine resulted in no significant change in kinetic parameters for the enzyme. In the case of Tyr110', the most distal of the triad from the active site, replacement with alanine results in a significant decrease (40-fold) in  $k_{\text{cat}}/K_{\text{m}}$ , whereas replacement with the bulkier phenylalanine results in no significant change in kinetic parameters. It seems that, in this case, the hydrogen bond is less important than the steric bulk of the residue in holding Ser47 in a correct position to hydrogen bond to Tyr137 and thereby stabilize its deprotonated, phenolate form once it has

donated its proton to the aldehyde oxygen of ManNAc during catalysis.

**Summary and Conclusions.** The current work exemplifies the mutual benefit of protein crystallography, QM/MM simulation, and mutagenesis in discerning enzyme mechanisms. The high-resolution structural information of the enzyme–substrate complex necessary for reliable QM/MM reaction modeling was obtained for NAL, allowing new insight into the mechanism of proton transfer in this enzyme in atomic detail. In turn, the modeled mechanism was then confirmed by kinetic analysis of enzyme mutants. On the basis of the X-ray crystallographic data, we have modeled the ManNAc–pyruvate–Schiff base complex with the pyruvate carboxylate group deprotonated (and it remained as such throughout the simulations). Our results indicate that a deprotonated Tyr137 in the vicinity of the Neu5Ac–Lys165 Schiff-base carboxylate is energetically stable, and there is thus no need to invoke a substrate-assisted mechanism as postulated for this enzyme<sup>11</sup> to allow reprotonation of Tyr137. This increased understanding of the catalytic cycle of NAL should provide the basis for future protein engineering to develop new aldolase-based biocatalysts.

## METHODS

**Expression and Purification of NAL.** Wild-type and variant *E. coli* NAL were expressed from the plasmid pK<sub>nan</sub>A-His<sub>6</sub> as previously described.<sup>14</sup>

**Site-Directed Mutagenesis.** Site-directed mutagenesis was performed using a QuikChange Lightning Site-Directed Mutagenesis Kit, (supplied by Agilent Technologies, Cheshire, U.K.) according to the manufacturer's guidelines.

**Kinetic Analyses.** Kinetic parameters of the aldol cleavage reaction were determined at 30 °C using a standard coupled assay<sup>14,15</sup> with lactate dehydrogenase (LDH) and NADH. The reaction (1 mL final volume) contained varying volumes (2–300 μL) of substrate (Neu5Ac) (100 mM, in Tris/HCl 50 mM, pH 7.4), 0.5 units LDH, 0.2 mM NADH, and Tris/HCl (50 mM; pH 7.4). NAL samples, in Tris/HCl (50 mM; pH 7.4), were added in volumes of between 10 and 200 μL. The decrease in absorbance at 340 nm was recorded on a Uvikon 930 spectrophotometer as the measure of enzyme activity. The rate of substrate cleavage was calculated using a molar extinction coefficient for NADH of 6220 M<sup>-1</sup> cm<sup>-1</sup>. Kinetic parameters were estimated by fitting the data to the Michaelis–Menten equation.

**Protein Crystallization.** *E. coli* NAL was crystallized using previously established conditions.<sup>5</sup> To form the Neu5Ac- and ManNAc/Pyr complexes NAL crystals were successively briefly soaked in the mother liquor containing 15% (w/v) PEG400, then 20% (w/v) PEG400, and subsequently for 5 min in 25% (w/v) PEG400 containing 75 mM Neu5Ac. Crystals were then flash cooled in liquid nitrogen prior to data collection.

**Data Collection and Refinement.** Diffraction data were collected from single crystals at the Diamond Light Source macromolecular crystallography beamline I02. The temperature for data collection was 100 K. Integration and scaling of data were carried out by MOSFLM<sup>16</sup> and SCALA.<sup>17</sup> REFMAC5<sup>18</sup> was used for refinement of the data, and after each refinement cycle model building was performed in COOT.<sup>19</sup> As the data were highly twinned (twin fraction 0.488, operator  $-h, -k, h+l$ ) the twin refinement option of REFMAC5 was used. Coordinates and restraint library files for the lysine residue covalently bound to a pyruvoyl moiety (HET code KPI) were as previously described.<sup>6</sup> Those for the Neu5Ac (HET code SI3) covalently bound to lysine and for ManNAc (HET code MN9) were generated using the PRODRG server. A number of models for the subunit C active site were tested (see Supporting Information for details) in order to identify that which best fit the observed electron density. The final models were validated using the PDB validation server.

**QM/MM Modeling.** A starting structure for QM/MM modeling was obtained by building the Tyr137 side chain onto Ala137 in subunit



C of the Y137A crystal structure (PDB code 4BWL), based on the conformation of Tyr137 in the previously solved wild-type enzyme with a pyruvate-Lys165 Schiff base (PDB code 2WNN).<sup>5</sup> ManNAc, pyruvate, Lys165 (from C $\gamma$ ), and Tyr137 (from C $\beta$ ) were treated quantum mechanically with self-consistent charge density functional tight-binding (SCC-DFTB), all other atoms with CHARMM22,<sup>20</sup> and all SCC-DFTB/CHARMM22 simulations were performed with CHARMM.<sup>21</sup> A 25 Å solvent sphere was added around the pyruvate-Lys165 Schiff base. Further details of the model and methods are described in the Supporting Information. In short, two 300 ps QM/MM MD trajectories (with different starting velocities) were run at 300 K. After at least 150 ps of simulation, 5 conformations were selected to start a series of short QM/MM MD simulations with a varying restraint on the carbon–carbon bond-forming distance d(C–C) ('umbrella sampling'). Subsequently, a series of energy minimizations ('adiabatic mapping') was applied along two reaction coordinates (see Results and Discussion), starting from the obtained intermediate. A set of structures along the thus identified SCC-DFTB/CHARMM22 minimum energy paths was selected to perform optimization of two reaction paths using B3LYP/6-31+G(d) for the QM region and applying the nudged-elastic band (NEB) and climbing-image NEB (CI-NEB) methods<sup>22,23</sup> in QoMMMa.<sup>24</sup> Transition states and minima were confirmed using QM/MM frequency calculations. Single-point spin-component scaled (SCS)-MP2/aug-cc-pVDZ/CHARMM22 calculations were performed on the optimized conformations.

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: [adrian.mulholland@bristol.ac.uk](mailto:adrian.mulholland@bristol.ac.uk).

\*E-mail: [a.berry@leeds.ac.uk](mailto:a.berry@leeds.ac.uk).

### Present Addresses

<sup>†</sup>Section of Microbiology and MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London SW7 2AZ, UK.

<sup>#</sup>Director, Research Complex at Harwell (RCaH), Rutherford Appleton Laboratory, Harwell (Oxford), Didcot, Oxon OX11 0FA, U.K.

### Author Contributions

<sup>||</sup>These authors contributed equally to this work.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank The Wellcome Trust and the BBSRC for supporting this work through Ph.D. studentships to I.C. (078110/Z/05/Z) and A.H.B. (075511/Z/04/Z), and A.D.D. (BB/D526502/1), respectively. A.J.M. is an EPSRC Leadership Fellow (grant number EP/G007705/1) and (with M.W.vdK.) thanks the EPSRC for support.

## ■ ABBREVIATIONS

NAL, N-acetylneuraminic acid lyase; Neu5Ac, N-acetylneuraminic acid; ManNAc, N-acetyl-D-mannosamine; QM/MM, quantum mechanical/molecular mechanical; LDH, lactate dehydrogenase; MD, molecular dynamics; DFT, density functional theory

## ■ REFERENCES

- (1) Bolt, A. H., Berry, A., and Nelson, A. (2008) Directed evolution of aldolases for exploitation in synthetic organic chemistry. *Arch. Biochem. Biophys.* 474, 318–330.
- (2) Uchida, Y., Tsukada, Y., and Sugimori, T. (1984) Purification and properties of N-acetylneuraminic lyase from *Escherichia coli*. *J. Biochem.* 96, 507–522.
- (3) Groher, A., and Hoelsch, K. (2012) Mechanistic model for the synthesis of N-acetylneuraminic acid using N-acetylneuraminic lyase from *Escherichia coli* K12. *J. Mol. Catal. B: Enzym.* 83, 1–7.
- (4) Barbosa, J. A. R. G., Smith, B. J., DeGori, R., Ooi, H. C., Marcuccio, S. M., Campi, E. M., Jackson, W. R., Brossmer, R., Sommer, M., and Lawrence, M. C. (2000) Active site modulation in the N-acetylneuraminic lyase subfamily as revealed by the structure of the inhibitor-complexed *Haemophilus influenzae* enzyme. *J. Mol. Biol.* 303, 405–421.
- (5) Campeotto, I., Bolt, A. H., Harman, T. A., Trinh, C. H., Phillips, S. E. V., Pearson, A. R., Nelson, A., and Berry, A. (2010) Structural insights into substrate specificity in variants of N-acetylneuraminic lyase produced by directed evolution. *J. Mol. Biol.* 404, 56–69.
- (6) Campeotto, I., Carr, S. B., Trinh, C. H., Nelson, A. S., Berry, A., Phillips, S. E. V., and Pearson, A. R. (2009) Structure of an *E. coli* N-acetylneuraminic acid lyase mutant, E192N, in complex with pyruvate at 1.45 Å resolution. *Acta Crystallogr., Sect. F* 65, 1088–1090.
- (7) Joerger, A. C., Mayer, S., and Fersht, A. R. (2003) Mimicking natural evolution *in vitro*: An N-acetylneuraminic lyase with an increased dihydropicolinate synthase activity. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5694–5699.
- (8) Izard, T., Lawrence, M., Malby, R., Lilley, G., and Colman, P. (1994) The three-dimensional structure of N-acetylneuraminic lyase from *Escherichia coli*. *Structure* 2, 361–369.
- (9) Lawrence, M. C., Barbosa, J. A., Smith, B. J., Hall, N. E., Pilling, P. A., Ooi, H. C., and Marcuccio, S. M. (1997) Structure and mechanism of a sub-family of enzymes related to N-acetylneuraminic lyase. *J. Mol. Biol.* 266, 381–399.
- (10) Van der Kamp, M. W., and Mulholland, A. J. (2013) Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* 52, 2708–2728.
- (11) Smith, B. J., Lawrence, M. C., and Barbosa, J. A. R. G. (1999) Substrate-assisted catalysis in sialic acid aldolase. *J. Org. Chem.* 64, 945–949.
- (12) Fessner, W. D., and Walter, C. (1997) Enzymatic C–C bond formation in asymmetric synthesis. *Top. Curr. Chem.* 184, 97–194.
- (13) Kruger, D., Schauer, R., and Traving, C. (2001) Characterisation and mutagenesis of the recombinant N-acetylneuraminic lyase from *Clostridium perfringens*: Insights into the reaction mechanism. *Eur. J. Biochem.* 268, 3831–3839.
- (14) Williams, G. J., Woodhall, T., Nelson, A., and Berry, A. (2005) Structure-guided saturation mutagenesis of N-acetylneuraminic acid lyase for the synthesis of sialic acid mimetics. *Protein Eng., Des. Sel.* 18, 239–246.
- (15) Blostein, R., and Rutter, W. J. (1963) Comparative studies of liver and muscle aldolase. II. Immunochemical and chromatographic differentiation. *J. Biol. Chem.* 238, 3280–3285.
- (16) Leslie, A. G. (2006) The integration of macromolecular diffraction data. *Acta Crystallogr., Sect. D* 62, 48–57.
- (17) Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr., Sect. D* 62, 72–82.
- (18) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D* 53, 240–255.
- (19) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) The features and development of Coot. *Acta Crystallogr., Sect. D* 66, 486–501.
- (20) MacKerell, A., Bashford, D., Bellott, M., Dunbrack, R., Evanseck, J., Field, M., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kucera, K., Lau, F., Mattos, C., Michnick, S., Ngo, T., Nguyen, D., Prodhom, B., Reiher, W., Roux, B., Schlenkerich, M.,

Smith, J., Stote, R., Straub, J., Watanabe, M., Wiorcikiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) *J. Phys. Chem. B* 102, 3586–3616.

(21) Brooks, B. R., Brucoleri, R. E., Olafson, B., States, D., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4, 187–217.

(22) Henkelman, G., Uberuaga, B. P., and Jonsson, H. (2000) A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* 113, 9901–9904.

(23) Van der Kamp, M. W., Żurek, J., Manby, F. R., Harvey, J. N., and Mulholland, A. J. (2010) Testing high-level QM/MM methods for modeling enzyme reactions: acetyl-CoA deprotonation in citrate synthase. *J. Phys. Chem. B* 114, 11303–11314.

(24) Harvey, J. N. (2004) Spin-forbidden CO ligand recombination in myoglobin. *Faraday Discuss.* 127, 165.

(25) Fitz, W., Schwark, J., and Wong, C. (1995) Aldotetroses and C(3)-modified aldohexoses as substrates for *N*-acetylneuraminic acid aldolase - a model for the explanation of the normal and the inversed stereoselectivity. *J. Org. Chem.*, 60.

(26) Lin, C. H., Sugai, T., Halcomb, R. L., Ichikawa, Y., and Wong, C. H. (1992) Unusual stereoselectivity in sialic-acid aldolase-catalyzed aldol condensations - synthesis of both enantiomers of high-carbon monosaccharides. *J. Am. Chem. Soc.* 114, 10138–10145.

(27) Makinen, M. W., and Fink, A. L. (1977) Reactivity and cryoenzymology of enzymes in the crystalline state. *Annu. Rev. Biophys. Bioeng.* 6, 301–343.

(28) Schauer, R., Stoll, S., Zbiral, E., Schreiner, E., Brandstetter, H. H., Vasella, A., and Baumberg, F. (1987) Interaction of *N*-acetyl-4-, 7-, 8- or 9-deoxyneuraminic acids and *N*-acetyl-4-, 7- or 8-mono-epi- and -7,8-di-epineuraminic acids with *N*-acetylneuraminic lyase. *Glycoconjugate J.* 4, 361–369.

(29) Mata, R. A., Werner, H. J., Thiel, S., and Thiel, W. (2008) Toward accurate barriers for enzymatic reactions: QM/MM case study on *p*-hydroxybenzoate hydroxylase. *J. Chem. Phys.*, 025104.

(30) Dobson, R. C. J., Perugini, M. A., Jameson, G. B., and Gerrard, J. A. (2009) Specificity versus catalytic potency: The role of threonine 44 in *Escherichia coli* dihydrodipicolinate synthase mediated catalysis. *Biochemie* 91, 1036–1044.

(31) Dobson, R. C. J., Valegard, K., and Gerrard, J. A. (2004) The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: Further evidence for a catalytic triad. *J. Mol. Biol.* 338, 329–339.

(32) Theodossis, A., Walden, H., Westwick, E. J., Connaris, H., Lamb, H. J., Hough, D. W., Danson, M. J., and Taylor, G. L. (2004) The structural basis for substrate promiscuity in 2-keto-3-deoxyglucuronate aldolase from the Entner-Doudoroff pathway in *Sulfolobus solfataricus*. *J. Biol. Chem.* 279, 43886–43892.