

Does glycosyl transfer involve an oxacarbenium intermediate? Computational simulation of the lifetime of the methoxymethyl cation in water*

Ian H. Williams^{1,‡}, J. Javier Ruiz Pernía¹, and Iñaki Tuñón²

¹Department of Chemistry, University of Bath, Bath, BA2 7AY, UK; ²Department of Chemistry, University of Valencia, 46100 Burjassot, Spain

Abstract: 2D free-energy surfaces for transfer of the methoxymethyl cation between two water molecules are constructed from molecular dynamics (MD) simulations in which these atoms are treated quantum-mechanically within a box of 1030 classical solvent water molecules at 300 K. This provides a simple model for glycosyl transfer in water. The AM1/TIP3P surfaces with 2D-spline corrections at either MPWB1K/6-31+G(d,p) or MP2/6-31+G(d,p) contain a shallow free-energy well corresponding to an oxacarbenium ion intermediate in a $D_N^*A_N$ mechanism. MD analysis at three temperatures leads to a classical estimate of the lifetime of the methoxymethyl cation in water; when quantum corrections for vibrational zero-point energy are included, the lifetime is estimated to be about 1 ps, in agreement with the best experimental estimate. This suggests that computational simulation, with appropriate high-level correction, is a reliable tool to obtain detailed and reliable mechanistic descriptions for glycosidases. In view of the importance of developing improved anti-influenza drugs, simulations of sialidases that considered both sialyl oxacarbenium ion and covalent sialyl-enzyme as possible intermediates could provide particular insight.

Keywords: computational simulation; glycosyl transfer; oxacarbenium ion; lifetime; quantum mechanics/molecular mechanics (QM/MM); solvent dynamics.

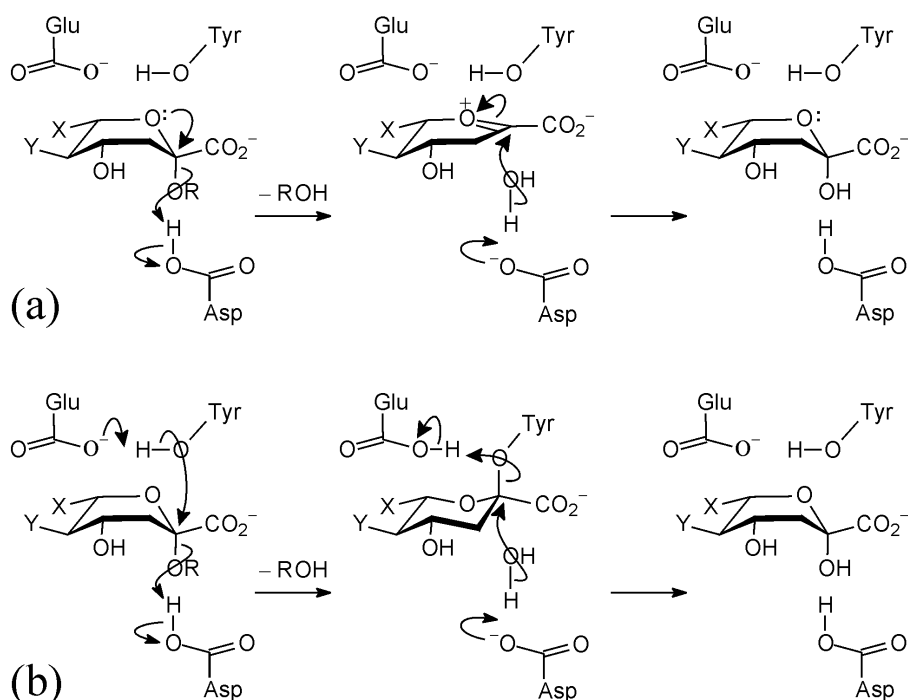
INTRODUCTION

The present frontline of defense against an influenza pandemic includes the antiviral drugs Tamiflu and Relenza, which were both designed as inhibitors of influenza sialidase, an enzyme found on the outer coating of the flu virus, which normally functions to cleave the terminal sialic acid residue from an oligosaccharide moiety of a glycoconjugate attached to a cell surface, thereby assisting the release of progeny virus particles from infected host cells.

The main features of the mechanism (Scheme 1a) originally proposed for influenza sialidase were: protonation by a general acid of the aglycone of the bound substrate with concomitant leaving group departure, stabilization of a sialyl oxacarbenium intermediate in the active site, and general base catalysis of attack by water on the intermediate yielding the sialic acid product [1,2]. All known sialidases catalyze hydrolysis of the anomeric sialyl ketal bond with overall retention of configuration and possess an active site containing a pair of acidic residues (Asp and Glu) and a conserved Tyr. However, recent studies of a *trans*-sialidase [3] and a sialidase [4] have demonstrated the involvement of Tyr as

*Paper based on a presentation made at the 20th International Conference on Physical Organic Chemistry (ICPOC-20), Busan, Korea, 22–27 August 2010. Other presentations are published in this issue, pp. 1499–1565.

‡Corresponding author

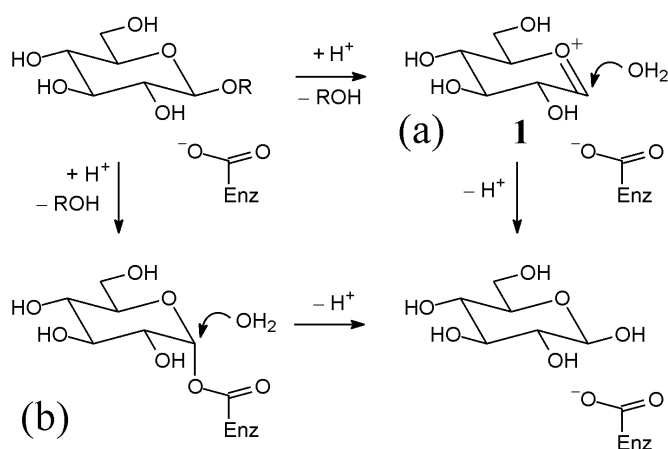


Scheme 1 Sialidase mechanisms involving (a) an oxocarbenium ion intermediate or (b) a covalent sialyl-enzyme intermediate.

the catalytic nucleophilic leading to a covalent sialyl intermediate [3], implying a mechanism (Scheme 1b) analogous to the standard double displacement reaction of retaining glycosidases that hydrolyze acetal linkages [5].

The question that now arises in regard to sialidases is the same as that posed many years ago for lysozyme: Is the key intermediate in the retaining mechanism a discrete oxocarbenium ion [6] or a covalently bound glycosyl-enzyme [7]? Following Phillips [6] and Vernon [8], it was long held that an oxocarbenium was involved, being stabilized electrostatically by the adjacent aspartate [9], but experimental evidence [10] has since unambiguously demonstrated the involvement of the covalent intermediate bound to Asp52 [11]. Nonetheless, it seems that an oxocarbenium intermediate is still considered by some to be involved in the mechanism of influenza sialidase [12]. The design of new and better anti-influenza drugs with, for example, less susceptibility to the development of drug resistance, might be aided by better understanding of the true nature of the intermediate involved in the mechanism, and it is this motive that has inspired the present study of the very simplest model for glycoside hydrolysis.

Glycoside hydrolysis mechanisms in aqueous solution teeter on the borderline between stepwise $D_N + A_N$ (S_N1) and concerted $A_N D_N$ (S_N2) mechanisms [13], with the lifetime of the oxocarbenium ion as a putative intermediate being the key point of issue. Extrapolations of kinetic measurements using the azide-clock method for a range of methoxy-substituted carbenium ions in water led Amyes and Jencks [14] to conclude that the stability of the glucosyl cation **1** (Scheme 2) was similar to that of the methoxymethyl cation **2**, and to estimate the lifetime in aqueous solution as ~ 1 ps. In the presence of a neutral nucleophile or leaving group, these cations exist for longer than a typical bond vibration (0.01–0.1 ps) but shorter than the typical time scales for reorganization within an ion pair (~ 10 ps) or diffusional separation (~ 60 ps) of an ion pair; thus neither has time to become equilibrated with dilute solutes or solvent molecules prior to their capture [15]. Consequently, substitutions involving these

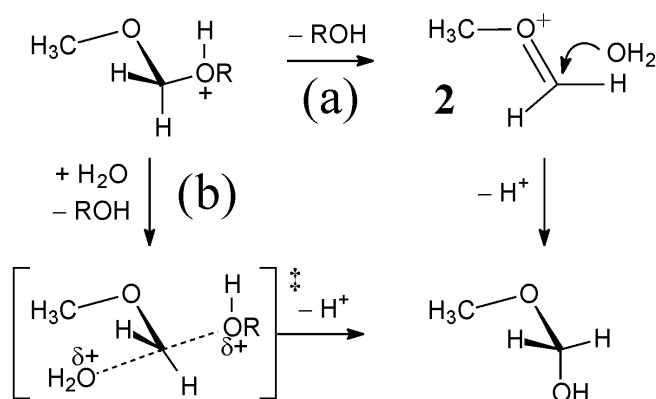


Scheme 2 Glucosidase mechanisms involving (a) an oxocarbenium ion intermediate or (b) a covalent glucosyl-enzyme intermediate.

cations with neutral nucleophiles require pre-association of the reacting species in a $D_N^*A_N$ mechanism or with anionic nucleophiles are enforced to be concerted [16,17].

Despite its simplicity, the methoxymethyl system holds an important place in the development of a physical organic understanding of borderline mechanisms. Kinetic studies of reactions of methoxymethyl substrates with dinitrophenolate [18] and *N,N*-dimethylaniline leaving groups [19] revealed very low sensitivity to the pK_a of the (conjugate acid of the) nucleophile, very high sensitivity to the pK_a of the nucleofuge, and α -deuterium kinetic isotope effects typical of stepwise $D_N + A_N$ reactions, despite being unambiguously bimolecular. This gave rise to the notion of the “exploded” transition state, and to the idea that the mechanism of a reaction could be enforced as concerted owing to the lifetime of the putative intermediate being too short for it to have an existence as a discrete species [16,17]. If a nucleofuge is not able to diffuse away from an electrophile in solution before nucleophilic attack occurs, the reaction is either concerted $A_N D_N$ or requires the nucleophile to be pre-associated in a $D_N^*A_N$ mechanism. The “exploded” transition state has significant oxocarbenium-ion character: there is little bond formation to the nucleophile and almost complete breaking of the bond to the leaving group [18,19].

Specific acid-catalyzed hydrolysis of methoxymethanol is, arguably, the simplest model for glycosyl transfer in water: Scheme 3 depicts the alternative $A_N D_N$ and $D_N^*A_N$ mechanisms for this system. By determining the lifetime of the methoxymethyl cation, we may gain some insight regarding the lifetime of the glycosyl cations such as glucopyranosylium **1** or sialyl, despite the more highly substituted character of these species. In any case, by performing a computational simulation of the lifetime of the methoxymethyl cation **2** in water, we are able to ascertain the level of theoretical treatment necessary in order to provide a realistic description of this type of species in a condensed medium as a prelude to further study of glycosyl transfer both in water and as catalyzed by glycosyl hydrolases or transferases.



Scheme 3 Acid-catalyzed hydrolysis of methoxymethanol ($R = \text{H}$) involving (a) a stepwise $\text{D}_\text{N}^*\text{A}_\text{N}$ mechanism via an intermediate oxocarbenium ion **2** or (b) a concerted $\text{A}_\text{N}\text{D}_\text{N}$ mechanism.

SIMULATION METHODS

A combination of quantum mechanics/molecular mechanics (QM/MM) for electronic structure calculations along with molecular dynamics (MD) and statistical mechanics techniques is employed, by means of the DYNAMO library [20]. The QM subsystem contains the methoxymethyl cation and two water molecules (nucleophile and leaving group) and is described initially by the semiempirical AM1 method. The MM subsystem comprises over 1000 surrounding solvent water molecules in a pre-equilibrated cubic box of side 31.5 Å centered on the central carbon (C_α) and described by the TIP3P potential. The system is first equilibrated by MD in the NVT ensemble subject to harmonic constraints to keep the nucleophilic water molecule close to the other QM atoms and the $\text{O}_{\text{lg}}-\text{C}_\alpha-\text{O}_{\text{nu}}$ angle close to linearity.

Transfer of the methoxymethyl group between the nucleofuge and nucleophile is described by a reaction coordinate RC defined as the difference between the distances of the central carbon atom to the oxygen atoms of these reacting water molecules, $RC = d(\text{C}_\alpha-\text{O}_{\text{lg}}) - d(\text{C}_\alpha-\text{O}_{\text{nu}})$; the sum of these two distances defines a tightness coordinate, $TC = d(\text{C}_\alpha-\text{O}_{\text{lg}}) + d(\text{C}_\alpha-\text{O}_{\text{nu}})$ describing the change from a “tight” pentacoordinate complex to a “loose” solvated oxocarbenium ion (Fig. 1a). A 2D potential of mean force (2D-PMF) is constructed with respect to RC and TC by averaging the atomic motions of the whole system at 300 K as governed by the QM/MM potential energy function and forces in each of 1156 simulation windows using double umbrella sampling with harmonic constraints applied to the two coordinates, as previously described [21]. The resulting surface provides information about the change in free energy both for transformation of reactants (negative RC) into products (positive RC) and for variation in the tight (small TC)/loose (large TC) character of intervening species.

In order to improve the quality of the QM treatment, interpolated corrections were applied in 2D by computing the gas-phase difference between single-point energies of the QM subsystem obtained with a “high-level” method and with “low-level” AM1, as a function of RC and TC , and with 2D-spline fitting. This procedure allows the corrected energy and gradient to be interpolated at each and every step of the MD trajectory, thereby leading to more reliable estimates of free-energy differences from the resulting 2D PMF [22].

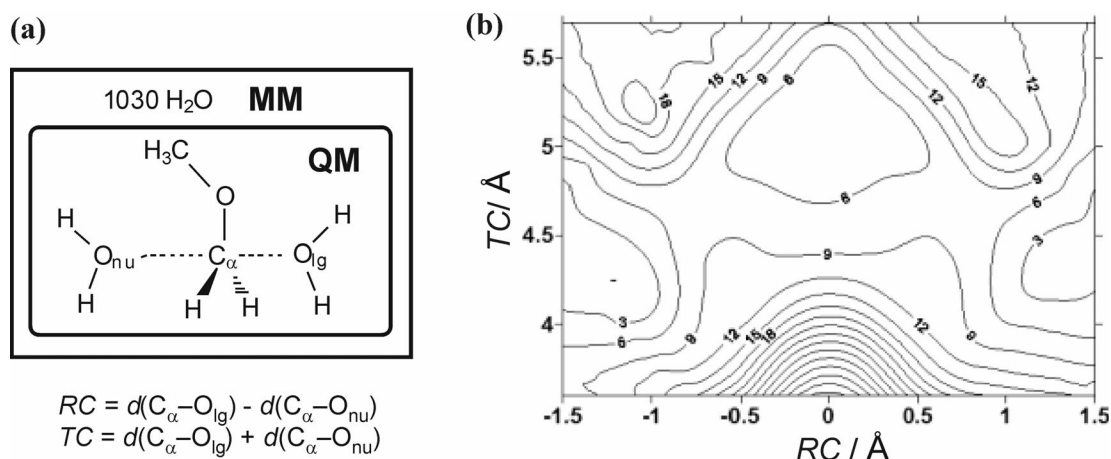


Fig. 1 (a) Atom labelling for QM and MM regions and (b) 2D-PMF contour map showing free energies (kcal mol⁻¹) computed at the MP2/6-31+G(d,p)-corrected AM1/TIP3P level for methoxymethyl cation transfer between water molecules in water.

RESULTS

The AM1/TIP3P 2D-PMF surface contains a single saddle point corresponding to the transition state for a concerted $A_N D_N$ mechanism, but the MPWB1K/6-31+G(d,p) density functional theory corrected 2D-PMF surface reveals a shallow free-energy minimum for a symmetrical intermediate between equivalent but asymmetrical transition states [21]. Figure 1b shows the contour map of the 2D-PMF with 2D-spline-corrections at the MP2/6-31+G(d,p) level of ab initio theory. The topographical features of this surface are very similar to those of the previously reported MPWB1K-corrected surface: the not-quite perfect symmetry about $RC = 0$ shows that the sampling of solvent-configurational space was adequate in construction of the 2D-PMF even though no symmetry constraints were imposed upon the simulations for this identity transfer of the methoxymethyl cation between water molecules in water. The instantaneous solvent environment is never symmetric, and symmetry is gained only by averaging over all accessible configurations. The nucleophile and leaving-group waters are each about 2.6 Å from C_α in the intermediate, but these distances are about 2.0 and 2.7 Å in the transition states. However, the intermediate well depth is deeper (1.5 kcal mol⁻¹) with MP2 than with MPWB1K (0.5 kcal mol⁻¹) and the free-energy barrier for methoxymethyl group transfer is also a bit smaller (7.5 vs. 8.5 kcal mol⁻¹). Both methods concur in predicting a stepwise $D_N^* A_N$ mechanism involving an intermediate that is appreciably looser than the symmetric transition state found at the uncorrected AM1/TIP3P level. The region of the free-energy surface around the intermediate is quite flat in both RC and TC ; slight changes in the system (e.g., structural modification in the reacting substrate or in the surrounding environment) could easily modify the shape of the surface and thereby the course of the most favorable reaction path. The mechanism inevitably involves pre-association of the nucleophile and electrophile for what is a hydrolysis in aqueous solution.

The lifetime of the intermediate MPWB1K-corrected surface was evaluated [21] by means of a series of MD trajectories initiated from the underlying potential-energy well. Simulations were performed at three temperatures with harmonic constraints imposed upon both $d(C_\alpha - O_{lg})$ and $d(C_\alpha - O_{nu})$ to equilibrate the system within this region. Thus, for each of 200 structures selected at 1 ps intervals along the constrained trajectory at each temperature, an unconstrained MD trajectory was initiated with random velocities taken from the Boltzmann distribution at that temperature; NVT trajectories were followed for 30 ps. Any configuration with both $d(C_\alpha - O_{lg})$ and $d(C_\alpha - O_{nu}) > 2.1$ Å was defined to be an intermediate; configurations with one of these distances < 2.1 Å were considered to belong to either the

reactant or the product basin. The time-evolution of the intermediate population $n(t)$ at each temperature was obtained by averaging over the set of 200 trajectories, and fitting to an exponential decay function of the type $n(t) = n_0 \exp(-t/\tau)$, where τ is the lifetime. This yielded classical lifetimes of 4.91, 3.56, and 3.53 ps at 300, 325, and 350 K, respectively, and an Arrhenius plot of $\ln(1/\tau)$ vs. $1/T$ gave a classical activation energy of 1.4 kcal mol⁻¹ for the intermediate to escape from the energy well. Application of quantum corrections to the vibrational degrees of freedom lowered this barrier by a further ~ 1 kcal mol⁻¹ leading to a best estimate of ~ 1 ps for the lifetime of the methoxymethyl cation in water at 300 K [21]. The slightly deeper well depth on the MP2-corrected surface might suggest a longer lifetime of about 5 ps.

DISCUSSION

Amyes and Jencks [14] estimated the lifetime of the methoxymethyl cation in three ways. First, they correlated $\log k$ for the hydration of oxacarbenium ions and for the addition of sulfite dianion [23] to the corresponding carbonyl compounds, from which a long extrapolation predicted a rate constant of approximately 5×10^{12} to 10^{15} s⁻¹ for the methoxymethyl cation. Second, they corrected the rate constant of $2\text{--}5 \times 10^{10}$ s⁻¹ for hydration of the oxacarbenium ions derived from propionaldehyde derivatives by a factor of 20–40, for stabilization by a methyl group, to give a rate constant of roughly 10^{12} s⁻¹ for the methoxymethyl cation. Thirdly, they correlated logarithms of rate constants for the hydration of oxacarbenium ions and logarithms of equilibrium constants for their formation from the corresponding α -azido ethers and extrapolated to obtain 10^{12} s⁻¹ as a lower limit for the methoxymethyl cation. They concluded that the lifetime of the methoxymethyl cation in water was “on the order of 10^{-12} s” [14]. Our results are clearly in agreement with theirs, but there are uncertainties with both the experimental and computational estimates: the former involved long extrapolations and the latter would seem to depend on what method is used to describe the QM region.

Pre-association of the nucleophile with the electrophile is not a diffusive process but instead presumably involves rotation of a solvent molecule to present a lone pair to the electrophilic carbon atom with accompanying loss of a hydrogen bond from a donor solvent molecule. The reaction coordinate RC for transfer of the methoxymethyl group is defined simply as the difference of distances of C_α to the nucleophile and leaving group but, by analogy with the mechanism for reorientation [24] of a hydrogen-bond donor water molecule (electrophile) between two hydrogen-bond acceptor water molecules (nucleophile and leaving group), and within the range of validity for $d(C_\alpha\text{--}O_{lg})$ and $d(C_\alpha\text{--}O_{nu}) < 4$ Å, progress along the reaction coordinate may be accompanied by changes in coordination numbers of the nucleophilic and leaving-groups waters and associated rotations. Whether or not the methoxymethyl cation exists as a discrete intermediate may depend on whether these motions within the hydrogen-bonded network of the solvent are uncoupled or coupled. Recent analysis of microwave and far-infrared dielectric spectroscopy data for liquid water suggests that the relaxation time of ~ 1 ps at 298 K corresponds to a “single molecule rotation” [25]. The “cation sandwich” ($H_2O \cdots CH_3OCH_2^+ \cdots OH_2$) corresponds to a minimum on the QM potential energy surfaces in vacuo, so the effect of bulk solvation in water is not such to cause the methoxymethyl cation to exist as a short-lived intermediate but rather, by virtue of solvent dynamics, to determine its lifetime.

Amyes and Jencks also argued [14] that the stability of the glucosyl cation in water was similar to or larger than that of the methoxymethyl cation, and consequently that its lifetime was also ~ 1 ps. It would be very desirable to perform computational simulations for the glucosyl cation in water analogous to those we have carried out for the methoxymethyl cation. Meanwhile, we would note the tendency of AM1/TIP3P to overestimate the tightness, or degree of concertedness, of methoxymethyl transfer in water. The same may also be true in an enzyme active site, but it would be expected that a glucosyl cation might anyway be too unstable to exist in the presence of an anionic (rather than a neutral) nucleophile, thus enforcing a concerted $A_N D_N$ mechanism. Note that such a mechanism was predicted by recent AM1/OPLS PMF simulations [26] for the glycosylation step of hydrolysis of *p*-nitro-

phenylxylobioside catalyzed by the retaining endo- β -1,4-xylanase from *B. circulans*. If high-level QM corrections can be applied to simulations for this and other enzyme-catalyzed glycoside hydrolysis and glycosyl transfer reactions, it might be expected that they would avoid any artificial bias toward concerted mechanisms and provide a reliable description of the preferred mechanism.

In view of the experimental difficulty of obtaining direct information regarding very short-lived intermediates in enzyme-catalyzed reactions, computational simulation has an important complementary role in the elucidation of mechanism. Our aim is perform reliable computational simulations of enzyme-catalyzed glycosyl transfer reactions; the present study offers encouragement that subtle distinctions between alternative mechanisms that do or do not involve short-lived glycosyl intermediates can be made, provided that an appropriate QM method is used to describe the reacting subsystem. As yet, very few computational modeling studies of glycosidase mechanisms include consideration of chemical bond making and breaking within the fully solvated protein environment [26–29], and fewer still examples of 2D free-energy surfaces [30–31]. The electrostatic environment within an enzyme active site is very different from that of water, and the issue of whether glycosyl cations exist as discrete intermediates in stepwise pre-association $D_N^*A_N$ mechanisms, or as oxacarbenium-like transition states in concerted A_ND_N mechanisms, is moot. The lessons learned from the methoxymethyl study should enable us in due course to apply appropriate high-level corrections in order to obtain detailed and reliable mechanistic descriptions for glycosidases; simulations of sialidases are currently in progress [32] with a view to the design of an improved anti-influenza drug.

REFERENCES

1. A. K. J. Chong, M. S. Pegg, N. R. Taylor, M. von Itzstein. *Eur. J. Biochem.* **207**, 335 (1992).
2. N. R. Taylor, M. von Itzstein. *J. Med. Chem.* **37**, 616 (1994).
3. A. G. Watts, I. Damager, M. L. Amaya, A. Buschiazzo, P. Alzari, A. C. Frasch, S. G. Withers. *J. Am. Chem. Soc.* **125**, 7532 (2003).
4. S. L. Newstead, J. A. Potter, J. C. Wilson, G. Xu, C-H. Chien, A. G. Watts, S. G. Withers, G. L. Taylor. *J. Biol. Chem.* **283**, 9080 (2008).
5. D. L. Zechel, S. G. Withers. *Acc. Chem. Res.* **33**, 11 (2000).
6. C. C. F. Blake, L. N. Johnson, G. A. Mair, A. C. T. North, D. C. Phillips, V. R. Sarma. *Proc. R. Soc. London B* **167**, 378 (1967).
7. D. E. Koshland. *Biol. Rev.* **28**, 416 (1953).
8. C. A. Vernon. *Proc. R. Soc. London B* **167**, 389 (1967).
9. A. Fersht. *Structure and Mechanism in Protein Science*, Freeman (1999).
10. D. J. Vocadlo, G. J. Davies, R. Raine, S. G. Withers. *Nature* **412**, 835 (2001).
11. A. J. Kirby. *Nat. Struct. Biol.* **8**, 737 (2001).
12. T. Islam, M. von Itzstein. *Adv. Carbohydr. Chem. Biochem.* **61**, 293 (2008).
13. M. L. Sinnott. *Carbohydrate Chemistry and Biochemistry: Structure and Mechanism*, RSC Publishing, Cambridge, UK (2007).
14. T. L. Amyes, W. P. Jencks. *J. Am. Chem. Soc.* **111**, 7888 (1989).
15. A. J. Bennet, T. E. Kitos. *J. Chem. Soc., Perkin Trans. 2* 1207 (2002)
16. W. P. Jencks. *Acc. Chem. Res.* **13**, 161 (1980).
17. W. P. Jencks. *Chem. Soc. Rev.* **10**, 345 (1981).
18. G. A. Craze, A. J. Kirby, R. Osborne. *J. Chem. Soc., Perkin Trans. 2* 357 (1978).
19. B. L. Knier, W. P. Jencks. *J. Am. Chem. Soc.* **102**, 6789 (1980).
20. M. J. Field, M. Albe, C. Bret, F. Proust-De Martin, A. Thomas. *J. Comput. Chem.* **21**, 1088 (2000).
21. J. J. Ruiz-Pernía, I. Tuñón, I. H. Williams. *J. Phys. Chem. B* **114**, 5769 (2010).
22. (a) J. J. Ruiz-Pernía, E. Silla, I. Tuñón, S. Martí, V. Moliner. *J. Phys. Chem. B* **108**, 8427 (2004);
(b) J. J. Ruiz-Pernía, E. Silla, I. Tuñón, S. Martí. *J. Phys. Chem. B* **110**, 17663 (2006).

23. P. R. Young, W. P. Jencks. *J. Am. Chem. Soc.* **99**, 8238 (1977).
24. D. Laage, J. T. Hynes. *J. Phys. Chem. B* **112**, 14230 (2008).
25. A. Beneduci. *J. Mol. Liq.* **138**, 55 (2008).
26. M. E. S. Soliman, J. J. Ruiz-Pernía, I. R. Greig, I. H. Williams. *Org. Biomol. Chem.* **7**, 5236 (2009).
27. S. Kozmon, I. Tvaroška. *J. Am. Chem. Soc.* **128**, 16921 (2006).
28. A. L. Bowman, I. M. Grant, A. J. Mulholland. *Chem. Commun.* 4425 (2008).
29. J. Liu, X. Wang, D. Xu. *J. Phys. Chem. B* **114**, 1462 (2010).
30. L. Petersen, A. Ardèvol, C. Rovira, P. J. Reilly. *J. Phys. Chem. B* **113**, 7331 (2009).
31. L. Petersen, A. Ardèvol, C. Rovira, P. J. Reilly. *J. Am. Chem. Soc.* **132**, 8291 (2010).
32. R. Castillo, G. D. Ruggiero, I. H. Williams. Paper in preparation.