ACCOUNTS of chemical research

pubs.acs.org/accounts Article

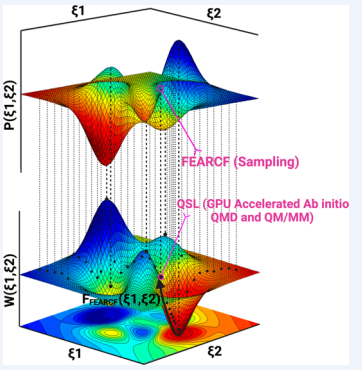
Multidimensional Free Energy and Accelerated Quantum Library Methods Provide a Gateway to Glycoenzyme Conformational, Electronic, and Reaction Mechanisms

Published as part of the Accounts of Chemical Research special issue "Direct Dynamics of Chemical Processes and Pathways".

Kevin J. Naidoo,* Tomás Bruce-Chwatt, Tharindu Senapathi, and Malcolm Hillebrand

Cite This: https://doi.org/10.1021/acs.accounts.1c00477 Read Online

CONSPECTUS: Enzyme reactions are complex to simulate accurately, and none more so than glycoenzymes (glycosyltransferase and glycosidases). A rigorous sampling of the protein frame and the conformationally plural carbohydrate substrate coupled with an unbiased treatment of the electron dynamics is needed to discover the true reaction landscapes. Here, we demonstrate the effectiveness of two computational methods ported in libraries that we have developed. The first is a flat histogram free energy method called FEARCF capable of multidimensional sampling and rapidly converging to a complete coverage of phase space. The second, the Quantum Supercharger Library (QSL), is a method that accelerates the computation of the ab initio electronic wave function as well as the integral derivatives on graphical processing units (GPUs). These QSL accelerated computations form the core components needed for direct quantum dynamics and QM/MM dynamics when coupled with legacy codes such as GAMESS and NWCHEM, making state of the art hyper-parallel electronic computations in chemistry and chemical biology possible. The combination of QSL (acceleration of ab initio QM computation) and FEARCF (multidimensional hyper-parallel reaction dynamics) makes the simulation of ab initio QM/MM reaction dynamics of enzyme catalysis feasible. Enzymes that process carbohydrates pose an added challenge as their pyranose ring substrates span multidimensional conformational space whose sampling is an intimate function of the catalytic mechanism. Here, we use the pairing of FEARCF and QSL to simulate the catalytic effect of the glycoenzyme *β*-*N*-acetylglucosamine transferase (OGT). The reaction mechanism is discovered from a variable three bond reaction surface using SCCDFTB. The role of the OGT in distorting the pyranose ring of *β*-*N*-acetylglucosamine (GlcNAc) away from the equilibrium $^4C_1$ chair conformation toward the E_3 envelope needed for the transition state is discovered from its pucker free energy hypersurfaces (or free energy volume, FEV). A complete GlcNAc ring pucker HF 6-31g FEV is constructed from ab initio QM dynamics in vacuum and ab initio QM/MM dynamics in the OGT catalytic domain. The OGT is shown to clearly lower the pathway toward the transition state E_3 ring conformer as well as stabilize it by 1.63 kcal/mol. Illustrated here is the use of QSL accelerated ab initio QM/MM dynamics that thoroughly explores carbohydrate catalyzed reactions through a FEARCF multidimensional sampling of the interdependence between reaction and conformational space. This demonstrates how experimentally inaccessible molecular and electronic mechanisms that underpin enzyme catalysis can be discovered by directly modeling the dynamics of these complex reactions.



KEY REFERENCES

- Naidoo, K. J. Multidimensional free energy volumes offer unique insights into reaction mechanisms, molecular conformation and association. *Phys. Chem. Chem. Phys.* **2012**, 14, 9026−9036.[1] *An overview of the FEARCF method applied to multidimensional simulations of reactions, conformation, and configurational sampling.*
- Barnett, C. B.; Naidoo, K. J. Ring puckering: A metric for evaluating the accuracy of AM1, PM3, PM3CARB-1 and SCC-DFTB carbohydrate QM/MM simulations. *J. Phys. Chem. B* **2010**, 114, 17142−17154.[2] *A detailed account of computing free energy pucker computation of furanose and pyranose rings.*
- Renison, C. A.; Fernandes, K. D.; Naidoo, K. J. Quantum supercharger library: Hyper-parallel integral derivatives algorithms for ab initio QM/MM dynamics. *J. Comp. Chem.* **2015**, 36, 1410−1419.[3] *Details of the QSL ab initio QM/MM development.*

Received: July 30, 2021

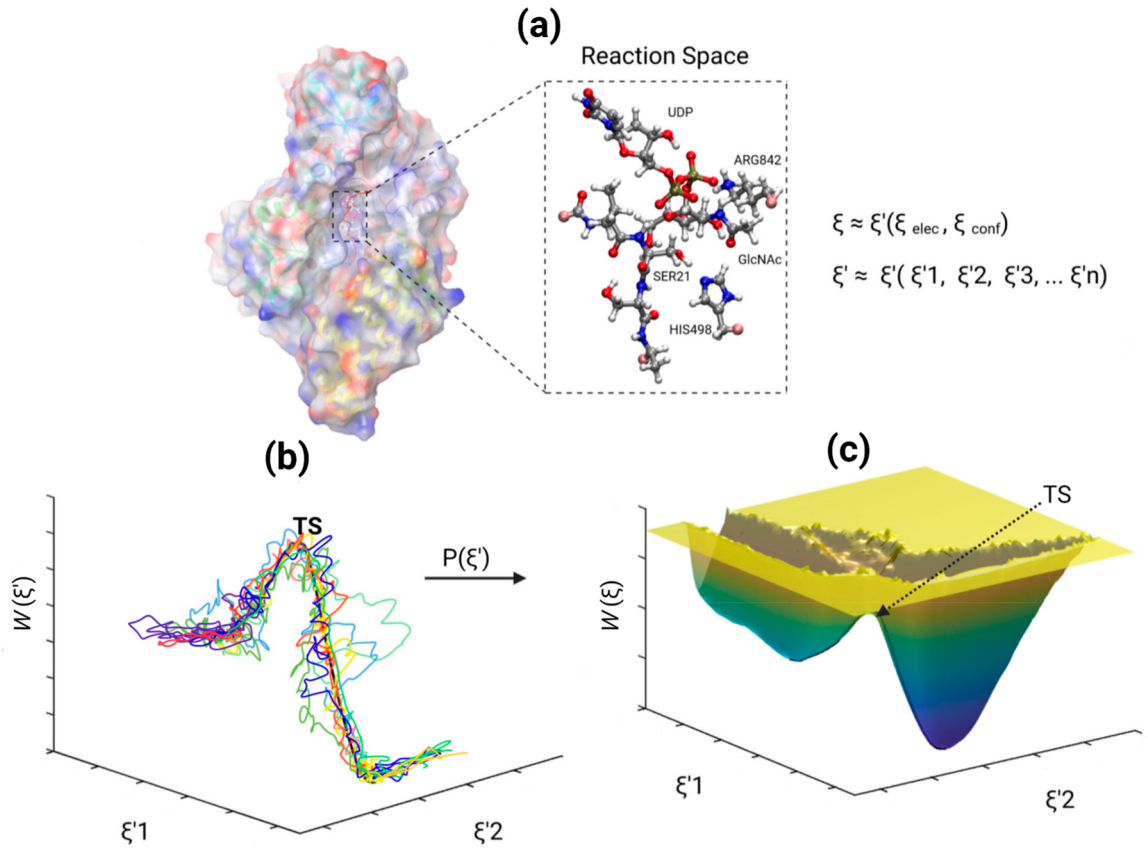


Figure 1. A model of a glycoenzyme mechanism showing (a) the conflated reaction coordinate, ξ, that includes conformational and electronic variables affecting the spatial and temporal distribution of electrons. (b) Computing the free energy by deconvoluting the reaction coordinate into progress variables, ξ' representing bond formation and cleavage. (c) The averaging along the progress variable to yield a reaction coordinate probability density $P(\xi')$ mapped to free energy surface $W(\xi')$.

- Rogers, I. L.; Naidoo, K. J. Producing DFT/MM enzyme reaction trajectories from SCC-DFTB/MM driving forces to probe the underlying electronics of a glycosyltransferase reaction. *J Comput Chem* **2017**, 38, 1789−1798.[4] *The FEARCF library implemented in NWChem to produce B3LYP/6-31G/MM glycosyltransferase reaction trajectories accompanied by the dynamic electronic evolution.*

INTRODUCTION

The vast and complex conformational space that make glycans important functional units in cellular biology presents a computational modeling challenge to sufficiently sample hemiacetal phase space.[1,2,4] While the cellular processing of glycans via glycoenzymes (glycosyltransferases and glycosidases) are known drug targets of a wide range of communicable and noncommunicable diseases, unpacking the mechanistic details and designing new drugs requires commensurate ab initio level QM/MM tools[3] to accurately simulate glycoenzyme catalytic action.[4] The epimeric choices leading to variation in cyclic pyranose monomers, the ability of the pyranose rings to pucker into 38 distinct conformations, and the 5 or more functional groups bonded to the ring carbons that can rotate into distinct orientations (cis, trans, gauche) makes carbohydrates ideal for data storage and information transfer. Glycosylation reactions (facilitated by glycosyltransferases), as well as deglycosylation reactions (facilitated by glycosidases), proceed via a transition state where the nonreducing sugar ring is puckered to enable electron redistribution.

Two major computational challenges compete when constructing an accurate simulation model of glycoenzyme action. These are to

(i) use a computational tool capable of expansive statistical sampling of carbohydrate conformational contributions to progress along reaction trajectory (Figure 1b)

(ii) build an ab initio dynamic model able to identify variables affecting the spatial and temporal distribution of electrons in the carbohydrate transition state (Figure 1c).

While the methods we have developed are universally applicable to computational enzymology and computational chemistry, the main objective driving the developments has been to overcome the special challenges posed by glycoenzyme modeling. A technical guiding principle has been to develop computational methods as functional libraries and modules able to link into legacy codes. This not only overcomes the challenges of (i) robust sampling of phase space and (ii) accessible ab initio dynamics but when linked to a legacy code can leverage the historical analytic and simulation methods built into those codes. For these reasons, we developed a QM/MM polar bond link atom (SLASH) method as a library,[5] a Free Energy and Reaction Dynamics (FEARCF) method[1,6] implemented as a library that when linked into MD packages comprehensively samples the phase space of pyranose puckering and glycoenzyme catalyzed, and a Quantum

https://doi.org/10.1021/acs.accounts.1c00477
Acc. Chem. Res. XXXX, XXX, XXX−XXX

supercharger library that accelerates ab initio Quantum and QM/MM dynamics.[3,7] Here, we describe the FEARCF and QSL and demonstrate their unique features on a Hartree−Fock QM/MM reaction dynamics simulation of *O*-linked *β*-*N*-acetylglucosamine transferase (*O*-GlcNAc or OGT).

A. FEARCF Efficient Sampling of Multidimensional Phase Space

Enzyme active sites are surrounded by an environment of thermally fluctuating protein conformational states. Computing chemical reaction rates must include this environmental diversity.[8] Moreover, in the case of glycoenzymes, the pyranose trajectories in the bond forming or breaking reaction path undergo necessary ring puckering to efficiently redistribute electrons about the anomeric center. The pyranose ring can pucker into chair (C), envelope (E), half chair (H), skew (S, also referred to as skew boat), and boat (B) conformations.

Investigating the role of a contorting sugar ring in a catalytic mechanism is central to accurate glycoenzyme modeling. The pyranose ring puckering is described using a triangular tessellation[9] definition. The triangular tessellation coordinates were specified for *β*-*N*-acetylglucosamine as $\xi_{\mathrm{GlcNac}} = (\theta_1, \theta_2, \theta_3)$, defining a three-dimensional phase space (Figure 2). $^4C_1$ is

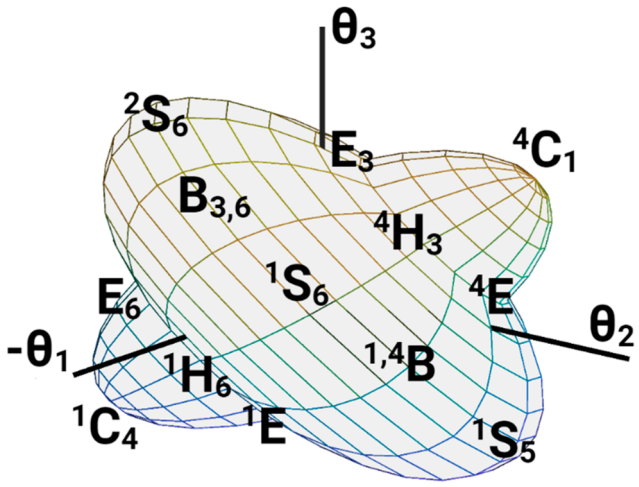


Figure 2. Distinct canonical configurations for a six-membered ring plotted onto a globe defined by the three triangular tessellated pucker angles.

at the north pole, with $^1C_4$ at the south pole. The equator contains the boat (B) and twist (S) conformations (six of each). The northern tropic contains six half-chairs (H) and six envelopes (E) that can be accessed from $^4C_1$ and the southern tropic contains six half-chairs and six envelopes that can be accessed from $^1C_4$. Theta refers to the angle between a defined reference plane (C4−O5−C2) and the triangular flaps defined by C1−C2−C3 (θ_1), C3−C4−C5 (θ_2), and C5−O5−C1 (θ_3).

The FEARCF Hamiltonian makes enhanced sampling of protein conformation possible, as well as enabling accurate sugar ring puckering in reaction simulations to explore phase space that beyond that found from equilibrium thermodynamics.[1] It is a generalized adaptive reaction coordinate force biasing method that has the capability of computing multidimensional free energies of reaction,[10] conformation,[11,12] association,[13] and configuration.[14] Recently, we have adapted FEARCF into a library format plugin for CHARMM,[15] NWChem,[16] and other molecular dynamics simulation legacy codes. Using the comprehensive sampling of reaction space, the competition between substitution and elimination pathways in the *Trypanosoma cruzi* trans-sialidase (TcTS) catalyzed reaction and TcTS's suppression of its side reactions to yield a single product were discovered through a CHARMM implementation of FEARCF to compute a SCCDFTB/MM free energy surface $W(\xi)$ or more specifically here, a reaction surface.[17] The underlying electronic dynamics were discovered using a NWChem implementation of FEARCF to compute the B3LYP/6-31G/MM dynamics employing driving forces from the SCC-DFTB/MM $W(\xi)$ for the TcTS glycosylation reaction.[4]

Ensemble averages are calculated from the frequency of occurrence of configurations that comply with the conditions of the reaction coordinate (ξ), and the probability density, $P(\xi)$. However, MD or Monte Carlo simulations at room temperature preferentially sample low energy configurations, making the computation of the average for $P(\xi)$ near impossible for chemical processes needing to overcome large energy barriers (Figure 1c). The large barriers are overcome by altering the potential of the system to artificially induce uniform sampling of the altered phase space (Figure 1b). A common strategy is umbrella sampling-type simulations that use a biasing potential chosen close to the quantity attempting to be determined. Flat histogram methods[18,19] attempt to automatically search for the optimal biasing potential on the fly without prior knowledge of the free energy profile. Non-Boltzmann sampling schemes have been developed for MD simulations where the Hamiltonian or Lagrangian is suitably modified to achieve equiprobable sampling for each state defined by ξ. FEARCF is such a flat histogram method[20] implemented as an embarrassingly parallelized algorithm building a biasing potential informed by previous sampling and designed to steer the system into unexplored areas (Figure 3). An early implementation of memory-dependence algorithms is local elevation,[21] which is central to the metadynamics algorithm,[22] both being examples of the flat histogram approaches. However, while these latter methods rely on a functional form to design the adaptive biasing potential, FEARCF employs a numerical potential built from appropriately weighted sums of histograms which in turn is related to the free energy (eq 1).

$$U_i(\xi) = -W_{i-1}(\xi) = k_B T \ln P_{i-1}(\xi) \tag{1}$$

Using the adaptive biasing potential, it is possible to access statistically rare events such as transition states that have relatively high free energies compared with the reactants and products. Multiple instances of the simulations are run in parallel, each one computed on a single core using the same biasing potential across all cores in each iteration. Simulations on each core are traveling along independent trajectories from which the statistics of sampling are then combined to update the biasing potential, as well as the free energy.[6]

The individual histograms from these parallel simulations are combined using a multidimensional weighted histogram analysis method (WHAM),[23] and n iterations of FEARCF produce a multidimensional free energy surface $W(\xi_i)$. To make the multidimensionality clear, we will use the phrase free energy volume (FEV) from here on.

The calculation of FEARCF driving forces, $F_i(\xi)$, as the gradient of $W_{i-1}(\xi)$ simplifies the extension to a multidimensional geometric reaction coordinate, $\xi(\xi_1, \xi_2, ...)$. The value of ξ is calculated for each step of the reaction dynamics

https://doi.org/10.1021/acs.accounts.1c00477
Acc. Chem. Res. XXXX, XXX, XXX−XXX

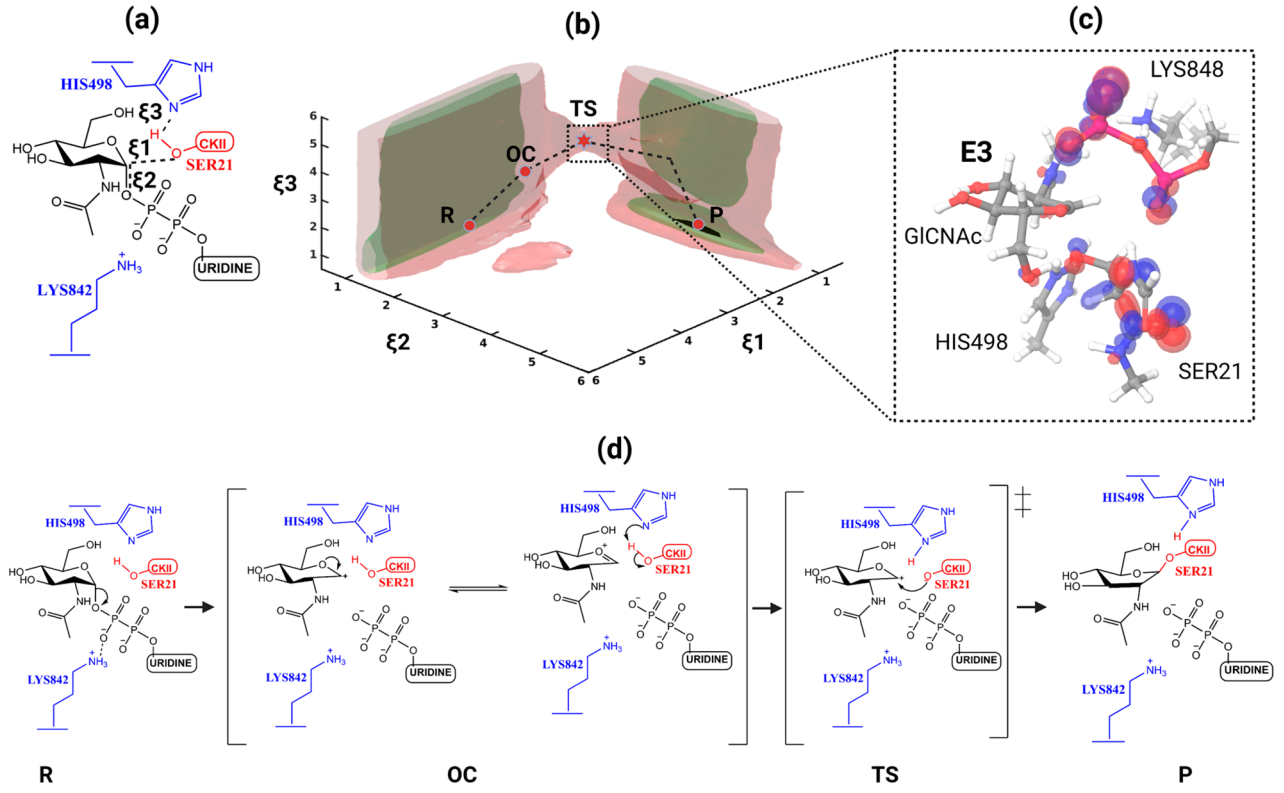


Figure 3. OGT FEARCF reaction simulation setup and the results summary. (a) The QM region used for the simulations and the three reaction coordinates labeled as $\xi 1$, $\xi 2$, and $\xi 3$. (b) FEV for OGT reactions. Isosurfaces are 0 kcal/mol, black; 8 kcal/mol, green; and 18 kcal/mol, red. Significant features along a crossing trajectory are R (reactants), OC (oxocarbenium ion), TS (transition state), P (product). (c) The transition state structure with HOMO and LUMO. The sugar ring is puckered to a E_3 conformation to facilitate the product formation. (d) The reaction scheme for the OGT catalytic itinerary.

trajectories and is used to compute driving forces $F(\xi_\alpha)$ from the negative partial derivative of the free energy surface (W), with respect to ξ_α (eq 2):

$$F_i(\xi_\alpha) = \frac{\partial W_{i-1}(\xi)}{\partial \xi_\alpha} \tag{2}$$

The chemical reasoning for carbohydrate active enzymes changing the conformation of substrates during catalysis is a unique example of how structure affects reactivity. It has been proposed that certain pucker conformations align the anomeric substituent, which is the leaving group in an axial orientation, ready for nucleophilic attack.[24,25] Alternatively, it has been advocated that delocalization of lone-pair electrons from the ring oxygen, which stabilizes the postulated cationic transition state and distorts the pyranose ring away from the equilibrium low energy chair conformation to one that favors such orbital overlap. The ring oxygen's axial lone pairs donate electrons into the σ^* antibonding orbital of axial C1−O1 bond, thereby stabilizing the oxo-carbenium ion transition state. This donation of electrons results in a partial double bond character at the C1−O5 bond, which suggests that at or extremely close to the transition state, the C-5-O-5-C-1-C-2 atoms must be coplanar.[24,26,27] Another FEARCF enabled discovery is the mechanism by which carbohydrate processing enzymes induce puckering of glucopyranose rings. This was first described in the context of cellobiohydrolase I (CBHI) catalysis of the hydrolysis of cellulose glycosidic bonds.[28]

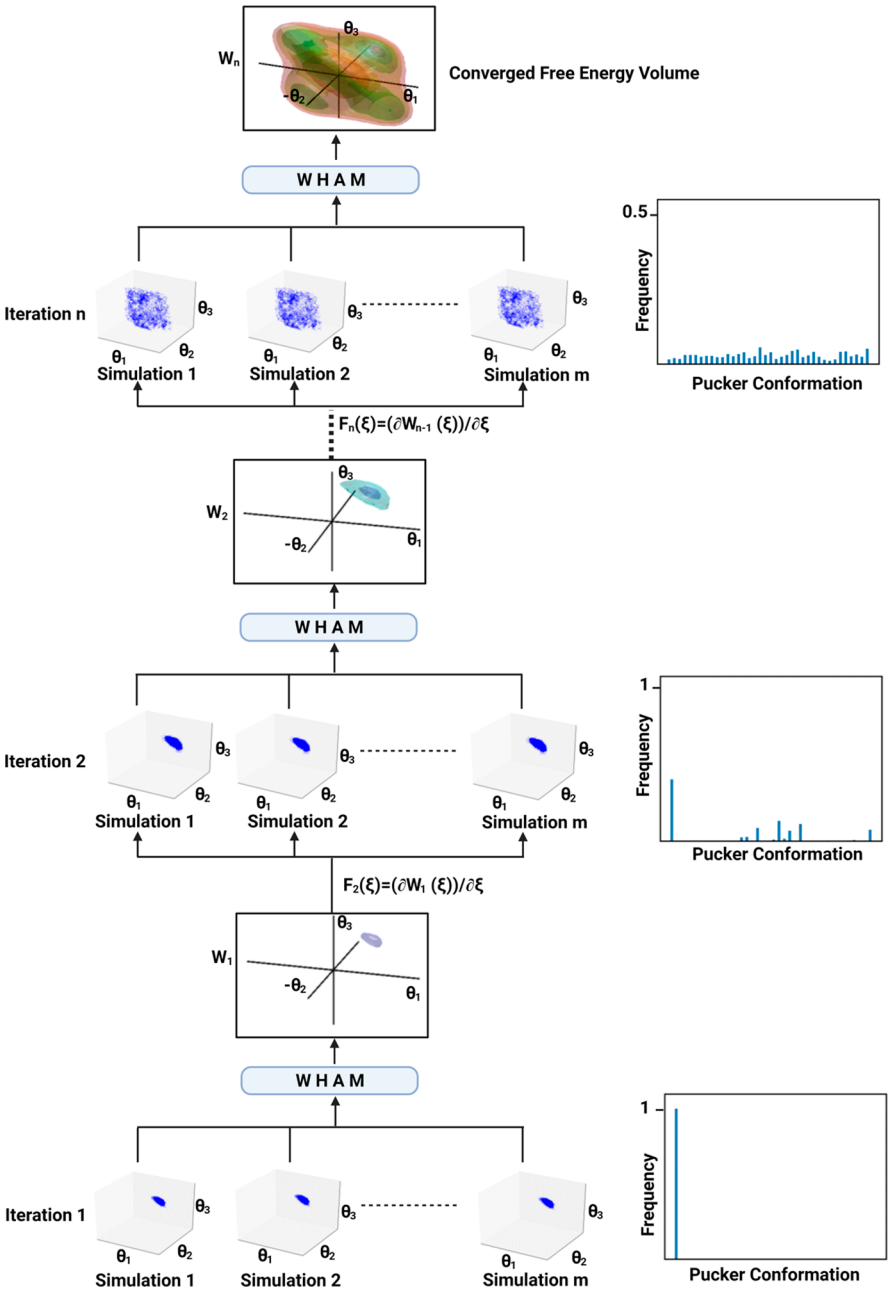
The convergence of a FEV can be illustrated in a bounded volume such as for the pucker free energy of the pyranose ring of β-*N*-acetylglucosamine (GlcNac). To generate the FEV for GlcNac in vacuum FEARCF was run for n = 50 iterations in CHARMM, with 120 parallel simulations for each iteration, with each iteration being 50 ps in length and run at 300 K using the Leapfrog integrator. Histograms for each reaction coordinate comprised 41 bins, over range of [−90,90] degrees for each pucker angle. All iterations were preceded by a 2 ps equilibrium step that did not include the biasing potential. An illustration of how the flat histogram was achieved is given in Scheme 1, where the first iteration (when the bias potential is set to zero) shows almost exclusive sampling in the $^4C_1$ configuration, which is the minimum energy configuration. In the second iteration, sampling has already improved in configurations like $^1S_5$ and $^1S_3$. By the 50th iteration, the ratio of least sampled conformations to most sampled has reached 1:5. To best show the sampling at each iteration, a single simulation was run using the same biasing potential used during the iteration with the same parameters but ran for a longer period of 2 ns.

B. Discovering the OGT Catalytic Reaction Mechanism

O-Linked β-*N*-acetylglucosamine transferase (*O*-GlcNAc transferase or OGT) is an essential enzyme in mammalian cells that plays a central role in cell recognition and signaling, as well as protein structure and function.[29] This nutrient dependent process is controlled by a pair of enzymes, OGT and *N*-acetyl-D-glucosaminidase (OGA). OGT acts by transferring *N*-acetylglucosamine (GlcNAc) from uridine-diphosphate-*N*-acetylglucosamine (UDP-GlcNAc) to the hydroxyl group of either the serine or threonine residues of proteins found in the

https://doi.org/10.1021/acs.accounts.1c00477
Acc. Chem. Res. XXXX, XXX, XXX−XXX

Scheme 1. FEARCF Convergence to Flat Histogram Sampling Illustrated with GlcNac Ring Pucker



cytoplasm, nucleus, and mitochondrion.[30] Following the OGT action, OGA hydrolyses the bound GlcNAc, removing it from the peptide residue, making the process dynamic and reversible.[31] Here, we discover the reaction mechanism of the *O*-GlcNAcylation of a well-characterized 14-residue CKII peptide (PGGSTPVSSANMM) acceptor from FEARCF reaction dynamics.

The human OGT crystal structure (PDB ID: 3PE4) was obtained from the RCSB protein data bank and used to prepare the Michaelis complex. Walker and coworkers crystallized the OGT CKII peptide: UDP-GlcNAc complex[32] capturing a completed reaction with the UDP cleaved and the *O*-GlcNAc bonded to a modified peptide serine residue adjacent to a valine residue in the peptide chain. The setup details are given in Supporting Information. The reaction free energy simulations were carried out using three reaction coordinates (ξ1, ξ2, and ξ3 defined in Figure 3a) to monitor the primary bond making and bond breaking trajectories. Histograms for each reaction coordinate comprised 55 bins over a sampling distance of 0.5−6 Å. A total of 15 FEARCF iterations were run, where each FEARCF iteration comprised 240 instances. Specifically, an iteration comprised of 240 cores, each computing distinct trajectories 30 ps in length. The SCCDFTB/MM level of theory was used to model the reaction that included a 2 ps equilibration run prior to the production runs in each instance for all FEARCF iterations. This is to ensure a unique starting structure for every trajectory.

Summing the reaction trajectories across all instances and FEARCF iterations, a total simulation time of 115.2 ns was used to construct the three-dimensional reaction FEV. The trajectories and histograms were collected after each iteration,

https://doi.org/10.1021/acs.accounts.1c00477
Acc. Chem. Res. XXXX, XXX, XXX−XXX

histograms were reweighted using WHAM, and FEVs were constructed for further analysis, as generally discussed above.

The crossing trajectories were isolated and compared with the minimum energy pathway (MEP) calculated from gradients of the free energy surfaces. A trajectory that progressed closest to the MEP was used to discover the relationship of the carbohydrate conformational dynamics to the reaction progress. The most significant of these are the substrate ring conformational pucker. There are several quantitative models where a reduced number of parameters is defined based on (i) perpendicular displacements from a mean plane of the ring,[33−36] (ii) the intracyclic torsion angles,[37−43] or (iii) the triangular tessellation (TT) of the ring.[9,33,44,45] The pucker conformations along the selected reaction trajectories were calculated using the TT or triangular decomposition (TD) method as it has been recently renamed.[9]

The reaction is initiated with a GlcNAc-UDP bond cleavage. The bond breaking mechanism is made possible through the strong interactions between the β-phosphate and the LYS848 residue that leads to a significant weakening of the GlcNAc-UDP bond. The decoupled GlcNAc forms a relatively stable oxocarbenium intermediate that leads to the GlcNAc ring reforming into an E_3 pucker conformation. This is a relatively unstable ring conformer in isolation. Strong interactions between GlcNAc and the leaving phosphate group, SER21, HIS498, as well as the primary alcohol aid to stabilize the GlcNAc in an E_3 pose. The complete proton transfer of the SER21 to HSD498 leads to the transition state formation, followed by the nucleophilic attack by the SER21 anion that leads to the product formation. The overall reaction followed a $D_N + A_{xh}D_H{}^{\ddagger}A_N$-type mechanism with 18.25 kcal/mol free energy barrier.

C. QSL ab Initio Quantum Dynamics and Hybrid QM/MM Dynamics

Stabilizing a positively charged oxocarbenium ion is central to the catalytic function of glycoenzymes, where some enzymes appear to act primarily by directly stabilizing the oxocarbenium ion.[46] Because ring puckering is a major driving force for chemical glycobiological reactions, we place a high priority on the proton affinities and electrostatic character (dipole moments) of the TS and other rings that are puckered away from the $^4C_1$ or $^1C_4$ chair conformers. Previously, we interrogated this by computing the complete free energy of pucker surfaces for ribose and the free energy of pucker volumes for glucose monosaccharides[2] using AM1,[47] PM3,[48] PM3CARB-1,[49] and SCC-DFTB[50] semiempirical (SE) methods. We then compared them with each other to selected DFT RB3LYP/6-311++G(d,p) stationary structures and to a free energy 6-31G Hartree−Fock surface.[2] Our conclusion that the SCCDFT method performed best but overall projected that computing accurate transition state properties for glycans in enzymatic reactions remained a universal failing of many SE methods. We address some of this with the development of a specialized NDDO method AM1/d-CB1 for modeling glycoenzymes.[51] The property prioritization of ring puckering dynamics, the dipole moments, and heats of formation of nonequilibrium ring conformers as well as the proton affinities, heats of formation, and dipole moments of amino acids were central to the development of AM1/d-CB1, and the performance is comparatively good for glycoenzyme modeling.[52] However, ab initio computations remain a superior albeit computationally inaccessible means of modeling glycoenzymatic mechanisms.

Computing ab initio electronic structures of molecules while limiting the computational time taken is rooted in the Hartree−Fock (HF) method of mutual electron repulsion and nuclear attraction.[53,54] The Schrödinger equation after the HF reduction is given by

$$f(r_1)\psi_i(r_1) = \varepsilon_i\psi_i(r_1) \tag{3}$$

where $\hat{f}$ is the Fock operator and ψ_i and ε_i are the wave function (expressible through Gaussian basis functions) and energy, respectively. HF computations use the concept of molecular orbitals to simplify the solution to the full Schrödinger equation for many-particle molecules while maintaining sufficient accuracy to provide meaningful, realistic outcomes.

In practice, HF self-consistent field (SCF) computations are primarily reduced to calculating single electron integrals for kinetic and nuclear attraction energies and two electron Coulomb and exchange energies that give the HF energy. In addition, for geometry optimizations and hybrid ab initio QM/MM simulations, derivatives of these integrals are required (Scheme 2). Particularly, the first derivative of the energy (or equivalently the Hamiltonian) with respect to the atomic coordinates is central to dynamics simulations. This derivative can be written in two terms as

$$\frac{\partial E}{\partial A_i} = \left\langle \Psi \left| \frac{\partial H}{\partial A_i} \right| \Psi \right\rangle + 2\left\langle \frac{\partial \Psi}{\partial A_i} | H | \Psi \right\rangle \tag{4}$$

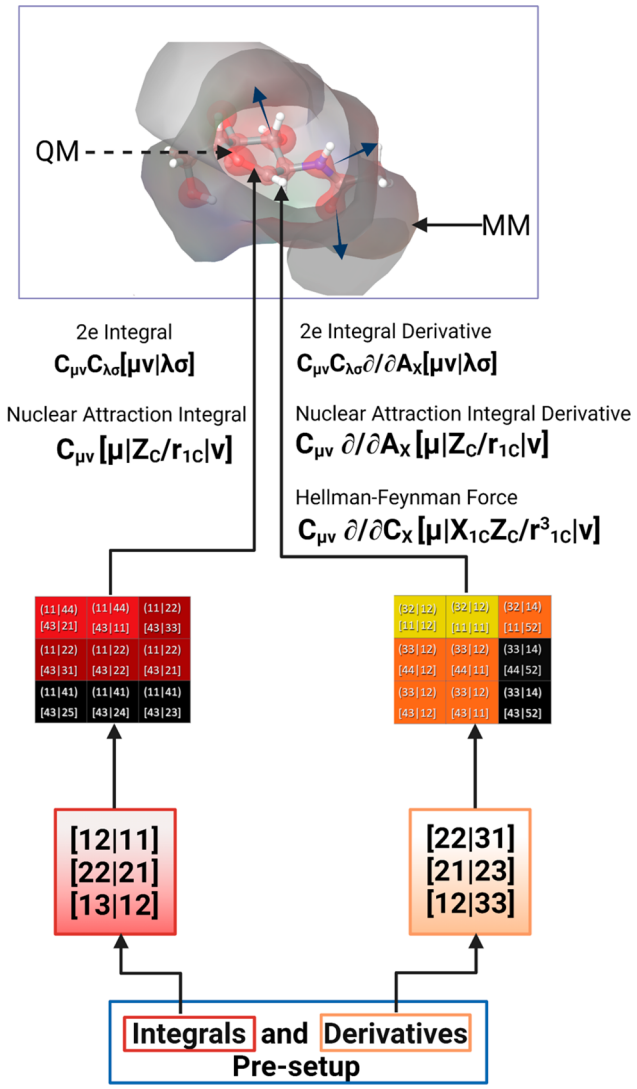
with the A_i representing the nuclear coordinate and Ψ representing the spatial wave functions. The first term here is referred to as the Hellmann−Feynman force and is calculated as an expectation value of the differential of the one-electron operator. The second term is a correction to the Hellmann−Feynman force. It is the Pulay force that is computed by differentiating the basis functions. This is done by evaluating both the one- and two-electron integral derivatives. Reducing their computational expense requires a focus of acceleration on these elements by performing efficient integral and integral derivative calculations.

There is a diverse array of methods proposed for computing these integrals and their derivatives. To fit the requirements of efficient hyper-parallelization, the scheme of choice for the QSL is the McMurchie−Davidson (MD) method.[55] The utility of this algorithm for GPU parallelism was outlined previously,[7] but the crux of the MD routine is the reduction of Gaussian integrals to the evaluation of Hermite polynomials which can be performed through recursion relations. This translates the integral contraction problem from an analytical nightmare to one that is computationally tractable (albeit intensive) strategy ideal for acceleration through hyper parallelization on hardware accelerators such as graphical processing units (GPUs).

GPU Parallelization Strategy. Parallel computation of integrals and derivatives of the integrals are ideally suited to GPUs compared with CPUs, as the computations have a single instruction, multiple data parallelization strategy (SIMD) architecture, and GPU are designed to address this problem. However, the implementation from traditional CPU algorithms to GPUs is not straightforward. It is critical to prepare code that makes optimal use of all GPU threads; while there are many cores on a GPU, the relatively slower performance of

https://doi.org/10.1021/acs.accounts.1c00477
Acc. Chem. Res. XXXX, XXX, XXX−XXX

Scheme 2. Ab Initio QM/MM Dynamics with One and Two Electron Integrals (Left) Used to Compute the HF and One and Two Electron Integral Derivatives (Right) Used for Force Computations in 1D Grids[a]



[a]Sized according to the number of significant contracted shell pairs. Darker colors represent larger quartet Schwarz values.

each core to a CPU means that speedup is contingent on effective parallelization. First, load balancing is a key consideration ensuring that a maximum number of threads are working at all times. Second, given the multiple threads and reuse of memory registers avoiding race conditions (i.e., where parallel threads become out of sync and interfere with each other) is central to the GPU algorithm design.

Finally, minimizing communication between a relatively slow host (i.e., CPU) and device (i.e., GPU) through memory transfers, if not carefully managed, can result in the gains from GPU calculation speedup eventually balanced out by inefficient data management.

The QSL takes all of these elements into account in constructing an efficient acceleration to the contracted integral bottleneck and performing many integrals and integral derivatives in parallel on a GPU (Scheme 2).[3,7] The strategy is different from our earlier work on accelerating the GAMESS-UK package on a GPU[56] and is similar in approach to that of Martínez et al.[57−60] on GPU acceleration but with differing implementations of the primitive contractions.[7]

Briefly, the QSL performs two electron integral calculations through grouping integrals by angular momentum (up to permutation symmetry), sorted and screened according to Schwarz pair values (left-hand flow in Scheme 2). These contracted shell pairs are assigned to blocks in the GPU grid, and their contribution to the final matrix is stored in memory. Coulomb and exchange contributions are computed in similar ways, with slightly different symmetry considerations.[7]

Equally, the computation of integral derivatives required a GPU-specific algorithm. As with the two electron integrals, symmetry is used in setting up classes of angular momentum shell pairs and screening via a prefactor test. These shell pairs are then computed by a block of threads on the GPU. A significant deviation between the CPU and GPU algorithms is the choice to ignore translational invariance for GPU memory efficiency. The ability to immediately map gradient contributions to the relevant atoms also circumvents the need to form a full Fock matrix from the derivatives. A full outline of the QSL implementation of one and two electron derivatives can be found in Renison et al.[3]

The essence of the QSL is to transfer a minimum amount of data from the host to device (e.g., Fock matrices, integral information), perform the full computation on GPU, and then transfer back a minimal data set to the host (e.g., returned Fock and density matrices). This allows for a data transfer and computationally efficient algorithm. A major advantage of the QSL is the library format making links to legacy code packages such as GAMESS-UK,[61] GAMESS-US,[62] NWChem,[16] etc. relatively simple. It functions as a true wave function, optimization, and QM/MM forces acceleration library, replacing the CPU routines for 1- and 2-electron integral Fock updates, integral derivative with calls to the QSL GPU kernels.

The FEARCF simulation of the GlcNAc ring pucker in the OGT enzyme as well as in vacuum was carried out using QSL accelerating GAMESS-UK and the CHARMM QM/MM routines. The setup was similar to that described for the vacuum SCCDFTB case above. The initial coordinates were chosen to be near the $^4C_1$ conformer. The dynamics simulations were run for 15 ps each, with a time step of 1 fs at 300 K. Eight V100 GPUs were used in parallel to run one simulation each, and 15 total iterations for the enzyme and 25 iterations for the vacuum were run to achieve a sufficiently flat histogram from the FEARCF library. Using a 20A sphere comprising water solvated OGT, we obtained 1.5 ps of QM/MM dynamics per hour on a V100, and 3.4 ps per hour of vacuum QMD dynamics. Further timing details are given in the Supporting Information.

The minimum energy pathway (MEP) between $^1C_4$ and E_3 in the pucker phase space was calculated for both the GlcNAc FEARCF simulations in OGT and in vacuum, using gradient descent (Figure 4a). The GlcNAc pucker FEV in vacuum shows the extent of the ring flexibility as it is able to access both $^4C_1$ (northern) and $^1C_4$(southern) poles via a contour band of 12 kcal/mol. The GlcNAc pucker FEV in OGT reveals how the constrained spatial and electrostatic reaction cavity confines GlcNAc to explore only the $^4C_1$ (northern) hemisphere. This limits the space to conformations about the TS E_3 conformation. To illustrate the difference between the 2 MEPs, the free energy value of each step of the path is plotted as a function of the distance traveled on the TD

https://doi.org/10.1021/acs.accounts.1c00477
Acc. Chem. Res. XXXX, XXX, XXX−XXX

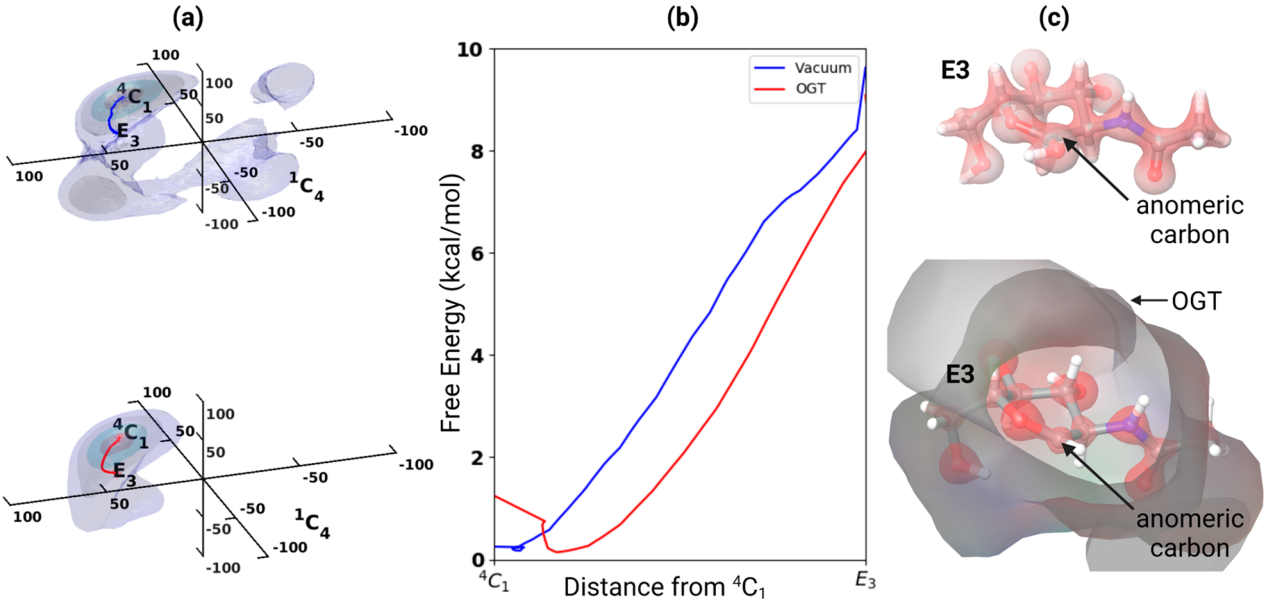


Figure 4. (a) The FEV of the beta-GlcNAc ring pucker including minimum energy path from $^4C_1$ to TS E_3 contoured at 1, 4, 8, and 12 kcal/mol with the top frame showing vacuum (blue) and bottom frame inside the OGT enzyme (red). (b) Vacuum and OGT MEPs shown as distance from $^4C_1$ chair to TS E_3 pucker. (c) TS E_3 vacuum electron density TS E_3 shown in the top frame and bottom frame inside the OGT enzyme.

spherical geometry from the $^4C_1$ chair to the TS E_3 puckered ring (Figure 4a). The OGT cavity disfavors the $^4C_1$ chair as is evident from the OGT MEP that is 1.02 kcal/mol higher than in vacuum. The OGT puckers the GlcNAc away from the usual equilibrium chair conformers, lowering this deviated path (red line Figure 4a) compared with the vacuum GlcNAc MEP pucker path (blue line Figure 4a). This is even despite the final difference in free energy at the E_3 conformation between OGT and FE being 1.63 kcal/mol. The detailed OGT mechanism that spells out the important contributing variables affecting the spatial and temporal distribution of electrons of this reaction will be reported elsewhere. However, the dynamics of the ring pucker and its role in lowering the reaction barrier is shown by comparing the HF/6-31g GlcNAc dynamic electron density (Figure 4c). By limiting the GlcNAc ring to the E_3 conformation in the reaction and electrostatically favoring this pose, the electron density is reduced on the anomeric carbon in comparison with the E_3 vacuum electron distribution.

CONCLUSION

Modeling catalytic mechanisms of glycoenzymes requires a method that can account for the essential conformational contributions of carbohydrate substrates that are being transformed. We have shown that multidimensional free energy methods can navigate not only conventional reaction space (bond formation and cleavage) but can measure the contributions made by the conformational deformation of carbohydrate rings to successful reactions. Combining this sampling methodology with a repurposed electronic Hamiltonian computed on GPU hardware accelerators that speeds up ab initio QM and QM/MM dynamics is central to building a gateway to directly discovering the reaction space and pathways of glycoenzymes.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.accounts.1c00477.

Additional simulation details and GPU timing benchmarks (PDF)

AUTHOR INFORMATION

Corresponding Author

Kevin J. Naidoo − *Scientific Computing Research Unit and Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa; Institute of Infectious Disease and Molecular Medicine, Faculty of Health Science, University of Cape Town, Rondebosch 7701, South Africa;* orcid.org/0000-0002-9898-3708; Email: kevin.naidoo@uct.ac.za

Authors

Tomás Bruce-Chwatt − *Scientific Computing Research Unit and Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa*

Tharindu Senapathi − *Scientific Computing Research Unit and Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa;* Present Address: Department of Chemistry, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda 10250, Sri Lanka; orcid.org/0000-0002-3277-4022

Malcolm Hillebrand − *Scientific Computing Research Unit and Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa; Nonlinear Dynamics and Chaos Group, Department of Mathematics, University of Cape Town, Rondebosch 7701, South Africa*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.accounts.1c00477

https://doi.org/10.1021/acs.accounts.1c00477
Acc. Chem. Res. XXXX, XXX, XXX−XXX

Notes

The authors declare no competing financial interest.

Biographies

Kevin J Naidoo is PhD graduate of the University of Michigan (United States) and trained as a Postdoctoral Fellow (1994−1995) at Cornell University (United States). He is a Professor in Physical Chemistry at the University of Cape Town (UCT). Since 2007, he has been the South African Research Chair in Scientific Computing. Over the last 10 years, his group's development of accelerated quantum code ported to GPUs has complimented the highly successful sampling FEARCF (Free Energies from Reaction Coordinate Forces) method developed over more than 20 years. His research objectives include producing ab initio classical mechanics (QM/MM) computational tools capable of nanosecond simulations of enzymatic reaction dynamics on GPU-based clusters.

Tomás Bruce-Chwatt obtained his BSc majoring in Physics and Biochemistry in 2017 from the University of Cape Town and is currently studying for his MSc in Computational Science at the SCRU under the supervision of Prof. Kevin J. Naidoo.

Tharindu Senapathi obtained his BSc in 2014 from the University of Sri Jayewardenepura. He obtained his MSc in computational science in 2017 and a PhD in chemistry in 2021 under the supervision of Professor Kevin Naidoo at the SCRU, University of Cape Town. He is currently a senior lecturer at the University of Sri Jayewardenepura as well as a research associate at the Scientific Computing Research Unit, University of Cape Town.

Malcolm Hillebrand obtained his BSc in 2017 and his PhD degree in Applied Mathematics from the University of Cape Town in 2021 under the supervision of Prof. Haris Skokos. He is currently pursuing postdoctoral research in the Nonlinear Dynamics and Chaos Group under Prof. Skokos as well as at the SCRU with Prof. Kevin Naidoo.

■ ACKNOWLEDGMENTS

This work is based on the research supported by the South African Research Chairs Initiative (SARChI Grant 449130) and the National Research Foundation of South Africa to K.J.N. T.B.-C. and T.S. thank the Scientific Computing Research Unit, the University of Cape Town SARChI, for graduate student funding. M.H. thanks the NRF of South Africa for Postdoctoral Funding (Grant 129630). The authors thank the Scientific Computer Research Unit and the Centre for High Performance Computing of South Africa for generous computing time.

■ ABBREVIATIONS

FEARCF, Free Energies From Adaptive Reaction Coordinate Forces; QSL, Quantum Supercharger Library; SLASH, Simple Link Atom Saccharide Hybrid; HF, Hartree−Fock; WHAM, weighted histogram analysis method

■ REFERENCES

(1) Naidoo, K. J. Multidimensional free energy volumes offer unique insights into reaction mechanisms, molecular conformation and association. *Phys. Chem. Chem. Phys.* **2012**, *14*, 9026−9036.
(2) Barnett, C. B.; Naidoo, K. J. Ring puckering: A metric for evaluating the accuracy of AM1, PM3, PM3CARB-1 and SCC-DFTB carbohydrate QM/MM simulations. *J. Phys. Chem. B* **2010**, *114*, 17142−17154.
(3) Renison, C. A.; Fernandes, K. D.; Naidoo, K. J. Quantum supercharger library: Hyper-parallel integral derivatives algorithms for ab initioQM/MM dynamics. *J. Comput. Chem.* **2015**, *36*, 1410−1419.
(4) Rogers, I. L.; Naidoo, K. J. Producing DFT/MM enzyme reaction trajectories from SCC-DFTB/MM driving forces to probe the underlying electronics of a glycosyltransferase reaction. *J. Comput. Chem.* **2017**, *38*, 1789−1798.
(5) Crous, W.; Field, M. J.; Naidoo, K. J. Simple Link Atom Saccharide Hybrid (SLASH) Treatment for Glycosidic Bonds at the QM/MM Boundary. *J. Chem. Theory Comput.* **2014**, *10*, 1727−1738.
(6) Naidoo, K. J. FEARCF a multidimensional free energy method for investigating conformational landscapes and chemical reaction mechanisms. *Sci. China: Chem.* **2011**, *54*, 1962−1973.
(7) Fernandes, K. D.; Renison, C. A.; Naidoo, K. J. Quantum supercharger library: Hyper-parallelism of the Hartree−Fock method. *J. Comput. Chem.* **2015**, *36*, 1399−1409.
(8) Masgrau, L.; Truhlar, D. G. The Importance of Ensemble Averaging in Enzyme Kinetics. *Acc. Chem. Res.* **2015**, *48*, 431−438.
(9) Hill, A. D.; Reilly, P. J. Puckering Coordinates of Monocyclic Rings by Triangular Decomposition. *J. Chem. Inf. Model.* **2007**, *47*, 1031−1035.
(10) Barnett, C. B.; Wilkinson, K. A.; Naidoo, K. J. Molecular Details from Computational Reaction Dynamics for the Cellobiohydrolase I Glycosylation Reaction. *J. Am. Chem. Soc.* **2011**, *133*, 19474−19482.
(11) Kuttel, M. M.; Naidoo, K. J. Free Energy Surfaces for the alpha(1 - 4)-Glycosidic Linkage: Implications for Polysaccharide Solution Structure and Dynamics. *J. Phys. Chem. B* **2005**, *109*, 7468−7474.
(12) Barnett, C. B.; Naidoo, K. J. Free Energies from Adaptive Reaction Coordinate Forces (FEARCF): An Application to Ring Puckering. *Mol. Phys.* **2009**, *107*, 1243−1250.
(13) Naidoo, K. J.; Lopis, A.; Westra, A. N.; Robinson, D. J.; Koch, K. R. Contact Ion Pair between Na+ and PtCl6 2- Favored in Methanol. *J. Am. Chem. Soc.* **2003**, *125*, 13330−13331.
(14) Gamieldien, M. R.; Strümpfer, J.; Naidoo, K. J. Hydration-Determined Orientational Preferences in Aromatic Association from Benzene Dimer Free Energy Volumes. *J. Phys. Chem. B* **2012**, *116*, 324−331.
(15) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(16) Aprà, E.; Bylaska, E. J.; Jong, W. A. d.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Valiev, M.; Dam, H. J. J. v.; Alexeev, Y.; Anchell, J.; Anisimov, V.; Aquino, F. W.; Atta-Fynn, R.; Autschbach, J.; Bauman, N. P.; Becca, J. C.; Bernholdt, D. E.; Bhaskaran-Nair, K.; Bogatko, S.; Borowski, P.; Boschen, J.; Brabec, J.; Bruner, A.; Cauët, E.; Chen, Y.; Chuev, G. N.; Cramer, C. J.; Daily, J.; Deegan, M. J. O.; Dunning, T. H., Jr.; Dupuis, M.; Dyall, K. G.; Fann, G. I.; Fischer, S. A.; Fonari, A.; Früchtl, H.; Gagliardi, L.; Garza, J.; Gawande, N.; Ghosh, S.; Glaesemann, K.; Götz, A. W.; Hammond, J.; Helms, V.; Hermes, E. D.; Hirao, K.; Hirata, S.; Jacquelin, M.; Jensen, L.; Johnson, B. G.; Jónsson, H.; Kendall, R. A.; Klemm, M.; Kobayashi, R.; Konkov, V.; Krishnamoorthy, S.; Krishnan, M.; Lin, Z.; Lins, R. D.; Littlefield, R. J.; Logsdail, A. J.; Lopata, K.; Ma, W.; Marenich, A. V.; Campo, J. M. d.; Mejia-Rodriguez, D.; Moore, J. E.; Mullin, J. M.; Nakajima, T.; Nascimento, D. R.; Nichols, J. A.; Nichols, P. J.; Nieplocha, J.; Otero-de-La-Roza, A.; Palmer, B.; Panyala, A.; Pirojsirikul, T.; Peng, B.; Peverati, R.; Pittner, J.; Pollack, L.; Richard, R. M.; Sadayappan, P.; Schatz, G. C.; Shelton, W. A.; Silverstein, D. W.; Smith, D. M. A.; Soares, T. A.; Song, D.; Swart, M.; Taylor, H. L.; Thomas, G. S.; Tipparaju, V.; Truhlar, D. G.; Tsemekhman, K.; Voorhis, T. V.; Vázquez-Mayagoitia, Á.; Verma, P.; Villa, O.; Vishnu, A. NWChem: Past, present, and future. *J. Chem. Phys.* **2020**, *152*, 184102.
(17) Rogers, I. L.; Naidoo, K. J. Multidimensional Reaction Dynamics Reveal How the Enzyme TcTS Suppresses Competing

https://doi.org/10.1021/acs.accounts.1c00477
Acc. Chem. Res. XXXX, XXX, XXX−XXX

Side Reactions and Their Side Products. *ACS Catal.* **2016**, *6*, 6384−6392.
(18) Wang, F. G.; Landau, D. P. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.* **2001**, *86*, 2050−2053.
(19) Wang, F. G.; Landau, D. P. Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2001** *64* DOI: 10.1103/PhysRevE.64.056101.
(20) Wang, F.; Landau, D. P. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.* **2001**, *86*, 2050.
(21) Huber, T.; Torda, A.; van Gunsteren, W. Local elevation: A method for improving the searching properties of molecular dynamics simulation. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695−708.
(22) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.
(23) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free□energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(24) Davies, G. J.; Planas, A.; Rovira, C. Conformational Analyses of the Reaction Coordinate of Glycosidases. *Acc. Chem. Res.* **2012**, *45*, 308−316.
(25) Walvoort, M. T. C.; van der Marel, G. A.; Overkleeft, H. S.; Codee, J. D. C. On the reactivity and selectivity of donor glycosides in glycochemistry and glycobiology: trapped covalent intermediates. *Chemical Science* **2013**, *4*, 897−906.
(26) Sinnott, M. L. Catalytic Mechanism of Enzymic Glycosyl Transfer. *Chem. Rev.* **1990**, *90*, 1171−1202.
(27) Rye, C. S.; Withers, S. G. Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* **2000**, *4*, 573−580.
(28) Barnett, C. B.; Wilkinson, K. A.; Naidoo, K. J. Pyranose Ring Transition State Is Derived from Cellobiohydrolase I Induced Conformational Stability and Glycosidic Bond Polarization. *J. Am. Chem. Soc.* **2010**, *132*, 12800−12803.
(29) Fardini, Y.; Dehennaut, V.; Lefebvre, T.; Issad, T. *O*-GlcNAcylation: A New Cancer Hallmark? *Front. Endocrinol.* **2013**, *4* DOI: 10.3389/fendo.2013.00099.
(30) Biwi, J.; Biot, C.; Guerardel, Y.; Vercoutter-Edouart, A.-S.; Lefebvre, T. The many ways by which *O*-GlcNAcylation may orchestrate the diversity of complex glycosylations. *Molecules* **2018**, *23*, 2858.
(31) Costa, A. F.; Campos, D.; Reis, C. A.; Gomes, C. Targeting Glycosylation: A New Road for Cancer Drug Discovery. *Trends in Cancer* **2020**, *6*, 757−766.
(32) Lazarus, M. B.; Nam, Y.; Jiang, J.; Sliz, P.; Walker, S. Structure of human *O*-GlcNAc transferase and its complex with a peptide substrate. *Nature* **2011**, *469*, 564−567.
(33) Bocian, D. F.; Pickett, H. M.; Rounds, T. C.; Strauss, H. L. Conformations of cycloheptane. *J. Am. Chem. Soc.* **1975**, *97*, 687−695.
(34) Cremer, D.; Pople, J. A. A general definition of Ring Puckering Coordinates. *J. Am. Chem. Soc.* **1975**, *96*, 1354−1358.
(35) Kilpatrick, J. E.; Pitzer, K. S.; Spitzer, R. The Thermodynamics and Molecular Structure of Cyclopentane1. *J. Am. Chem. Soc.* **1947**, *69*, 2483−2488.
(36) Pickett, H. M.; Strauss, H. L. Symmetry and Conformation of the Cycloalkanes. *J. Chem. Phys.* **1971**, *55*, 324−334.
(37) Altona, C.; Geise, H. J.; Romers, C. Conformation of non-aromatic ring Compounds—XXV: Geometry and conformation of ring D in some steroids from X-ray structure determinations. *Tetrahedron* **1968**, *24*, 13−32.
(38) Bérces, A.; Whitfield, D. M.; Nukada, T. Quantitative description of six-membered ring conformations following the IUPAC conformational nomenclature. *Tetrahedron* **2001**, *57*, 477−491.
(39) Diez, E.; Esteban, A. L.; Bermejo, F. J.; Rico, M. Ring-puckering geometrical models for five-membered rings. *J. Phys. Chem.* **1980**, *84*, 3191−3196.
(40) Geise, H. J.; Altona, C.; Romers, C. The relations between torsional and valency angles of cyclopentane. *Tetrahedron Lett.* **1967**, *8*, 1383−1386.
(41) Haasnoot, C. A. G. The conformation of six-membered rings described by puckering coordinates derived from endocyclic torsion angles. *J. Am. Chem. Soc.* **1992**, *114*, 882−887.
(42) Zefirov, N. S.; Palyulin, V. A. Quantitative Characteristic of Cycle Shape in Structural and Stereochemical Investigations. *Dokl. Akad. Nauk SSSR.* **1980**, *252*, 111−115.
(43) Zefirov, N. S.; Palyulin, V. A.; Dashevskaya, E. E. Stereochemical studies. XXXIV. Quantitative description of ring puckering via torsional angles. The case of six-membered rings. *J. Phys. Org. Chem.* **1990**, *3*, 147−158.
(44) Joshi, N. V.; Rao, V. S. R. Flexibility of the pyranose ring in α- and β-D-glucoses. *Biopolymers* **1979**, *18*, 2993−3004.
(45) Strauss, H. L.; Pickett, H. M. Conformational structure, energy, and inversion rates of cyclohexane and some related oxanes. *J. Am. Chem. Soc.* **1970**, *92*, 7281−7290.
(46) Berti, P. J.; McCann, J. A. B. Toward a Detailed Understanding of Base Excision Repair Enzymes: Transition State and Mechanistic Analyses of N-Glycoside Hydrolysis and N-Glycoside Transfer. *Chem. Rev.* **2006**, *106*, 506−555.
(47) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(48) Stewart, J. J. P. Optimization of parameters for semiempirical methods. I. Method. *J. Comput. Chem.* **1989**, *10*, 209−220.
(49) McNamara, J. P.; Muslim, A.; Abdel-Aal, H.; Wang, H.; Mohr, M.; Hillier, I. H.; Bryce, R. A. Towards a quantum mechanical force field for carbohydrates: a reparametrized semi-empirical MO approach. *Chem. Phys. Lett.* **2004**, *394*, 429−436.
(50) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM Implementation of the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) Method. *J. Phys. Chem. B* **2001**, *105*, 569−585.
(51) Govender, K.; Gao, J.; Naidoo, K. J. AM1/d-CB1: A Semiempirical Model for QM/MM Simulations of Chemical Glycobiology Systems. *J. Chem. Theory Comput.* **2014**, *10*, 4694−4707.
(52) Govender, K. K.; Naidoo, K. J. Evaluating AM1/d-CB1 for Chemical Glycobiology QM/MM Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 4708−4717.
(53) Fock, V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Eur. Phys. J. A* **1930**, *61*, 126−148.
(54) Hartree, D. R. Results of Calculations of Atomic Wave Functions. I. Survey, and Self-Consistent Fields for Cl^- and Cu^+. *Proceedings of the Royal Society of London. Series A* **1933**, *141*, 282−301.
(55) McMurchie, L. E.; Davidson, E. R. One and two electron integrals over Cartesian Gaussian functions. *J. Comput. Phys.* **1978**, *26*, 218−231.
(56) Wilkinson, K.; Sherwood, P.; Guest, M. F.; Naidoo, K. J. Acceleration of the GAMESS-UK Electronic Structure Package on Graphical Processing Units. *J. Comput. Chem.* **2011**, *32*, 2313−2318.
(57) Ufimtsev, I. S.; Martinez, T. J. Quantum chemistry on graphical processing units. 1. Strategies for two-electron integral evaluation. *J. Chem. Theory Comput.* **2008**, *4*, 222−231.
(58) 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.
(59) 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.
(60) Titov, A. V.; Ufimtsev, I. S.; Luehr, N.; Martínez, T. J. Generating Efficient Quantum Chemistry Codes for Novel Architectures. *J. Chem. Theory Comput.* **2013**, *9*, 213−221.

https://doi.org/10.1021/acs.accounts.1c00477
Acc. Chem. Res. XXXX, XXX, XXX−XXX

(61) Guest, M. F.; Bush, I. J.; van Dam, H. J. J.; Sherwood, P.; Thomas, J. M. H.; van Lenthe, J. H.; Havenith, R. W. A.; Kendrick, J. The GAMESS-UK electronic structure package: algorithms, developments and applications. *Mol. Phys.* **2005**, *103*, 719−747.
(62) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347−1363.

https://doi.org/10.1021/acs.accounts.1c00477
Acc. Chem. Res. XXXX, XXX, XXX−XXX