Supporting Information

Molecular Mechanism of the Glycosylation Step Catalyzed by Golgi α-Mannosidase II. A QM/MM Metadynamics Investigation

Luis Petersen,[†] Albert Ardèvol,[‡, §] Carme Rovira,[‡,§,¶] and Peter J. Reilly*[,†]

[†] *Department of Chemical and Biological Engineering, Iowa State University, Ames, Iowa 50011; Computer Simulation and Modeling Laboratory (CoSMoLab),* [‡] *Parc Científic de Barcelona, 08028 Barcelona, Spain;* [§] *Institut de Química Teòrica i Computacional (IQTCUB), Universitat de Barcelona, 08028 Barcelona, Spain;* [¶] *Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain*

CONTENTS

1. N-linked glycosylation pathway

The following scheme (Figure S1) illustrates the enzyme-catalyzed steps followed in the N-linked glycosylation pathway. First, three glucosyl units of the sugar complex are trimmed from $Glc_3Man_9GlcNAc_2$ by α-glucosidases. The product, $Man_9GlcNAc_2$, is then hydrolyzed to $Man_5GlcNAc_2$ by α-mannosidases in the endoplasmic reticulum and Golgi apparatus. Once the protein is inside the latter, a transferase (GlcNAc–TI) catalyzes the addition of a GlcNAc moiety to the carbohydrate complex. A second α-mannosidase (GMII) cleaves two more mannosyl residues to produce $GlcNAcMan_3GlcNAc_2$, which is further transformed by a second transferase (GlcNAc–TII) into $GlcNAc_2Man_3GlcNAc_2$, the precursor of more complex oligosaccharides.
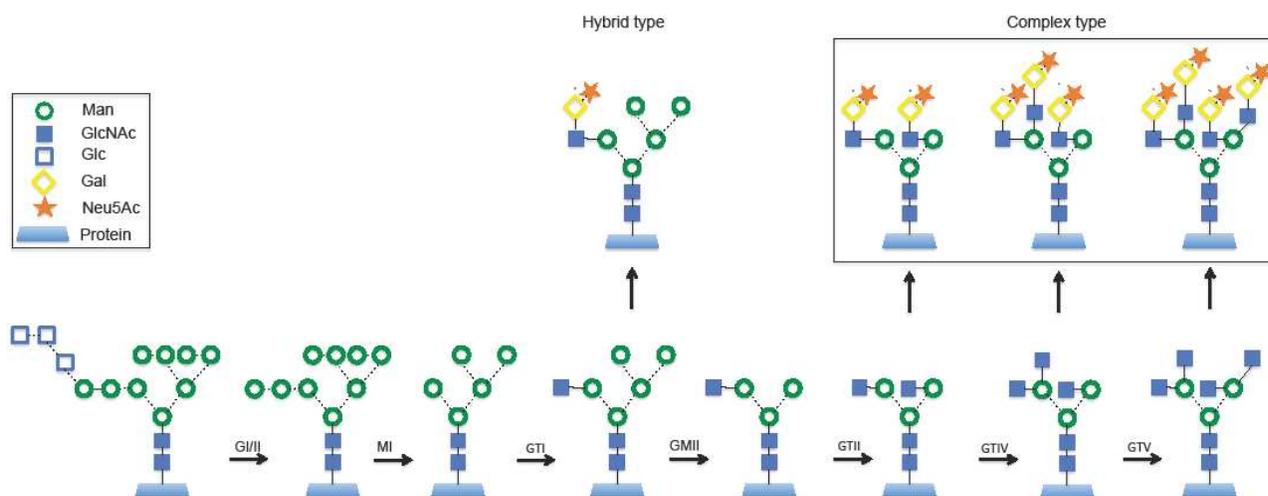

Figure S1. Golgi N-linked glycosylated structure modification pathway. See Figure 1 caption for explanation of bond angles and solid and dashed bonds.

2. Initial structure

To validate the quality of the initial x-ray crystal structure model, the B-factors of the active-site residues (PDB 3CZN)[1] are presented in Table S1. In addition, a superposition (Figure S2) between the QM/MM-minimized structure (after equilibration at natural conditions) and the x-ray crystal structure (PDB 3CZN) is presented.

Table S1. B-Factor Values of Active-Site
Residues of the PDB 3CZN Structure Used
in the Calculations

Residue	Atom	B-factor
His90	N	10.27
	CA	10.19
	C	10.45
	O	10.34
	CB	10.81
	CG	10.71
	CD2	11.08
	ND1	11.59
	CE1	10.93
	NE2	10.34
Asp92	N	11.45
	CA	11.10
	C	11.10
	O	11.45
	CB	10.99
	CG	12.64
	OD1	11.89
	OD2	16.68
Asp204*	N	13.52
	CA	13.45
	C	13.53
	O	13.90
	CB	13.61
	CG	15.49
	OD1	14.23
	OD2	14.35
Asp341	N	11.25
	CA	11.27
	C	11.44
	O	11.61
	CB	12.35
	CG	14.35
	OD1	15.09
	OD2	15.76
His471	N	9.03
	CA	9.39
	C	9.69
	O	10.61
	CB	9.79

	CG	9.90
	CD2	9.88
	ND1	10.80
	CE1	10.28
	NE2	10.05
Asp472	N	9.08
	CA	10.13
	C	10.32
	O	11.59
	CB	10.56
	CG	10.99
	OD1	11.38
	OD2	11.00
Zn1102	N	12.19
M5G1103	C11	32.80
(Ligand)	O11	34.02
	C12	24.29
	O12	27.60
	C13	17.46
	O13	18.24
	C14	13.73
	O14	13.30
	C15	16.92
	O15	17.72
	C16	15.64
	O16	15.23
	C17	16.05
	O17	17.52
	C21	31.72
	N21	31.24
	C22	21.99
	O22	24.71
	C23	16.15
	C24	13.46
	N24	16.27
	C25	17.93
	O25	18.29
	C26	14.75
	O26	16.03
	C27	15.75
	O27	18.10
	C31	30.91
	O31	29.27
	C32	20.42
	C33	17.15
	O33	17.21

C34	13.74
O34	12.79
C35	16.75
C36	14.49
O36	13.50
C37	13.71
O37	13.90
C41	30.35
C42	19.59
O42	21.04
C43	18.45
O43	20.41
C44	12.92
O44	12.85
C45	16.78
O45	18.05
C46	14.70
O46	14.07
C47	11.34
O47	12.25
C51	32.39
O51	33.37
C52	20.06
O52	21.05
C53	19.56
O53	17.94
C54	13.50
O54	14.58
C55	16.14
O55	17.03
C56	14.87
O56	14.60
C57	13.00
O57	14.50
C61	32.87
O61	35.98
C62	18.06
C63	21.47
O63	24.98
C64	15.72
O64	17.35
C65	17.28
C66	16.91
O66	22.21
C67	13.07
O67	12.73

C71	31.64
O71	31.67
C74	17.62
O74	19.87
C81	31.36
C84	19.15

* Modeled from PDB 3BUP
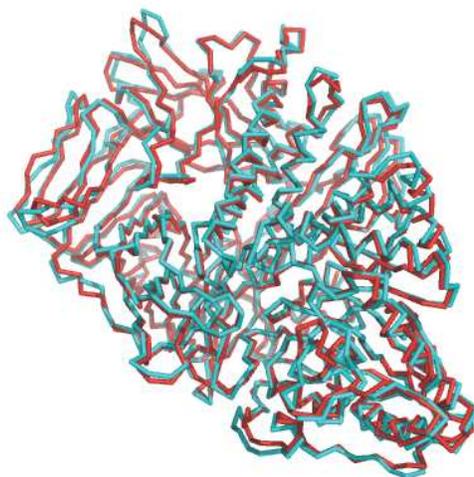
Figure S2. Structural alignment (RMSD 0.93 Å) between the backbone of the PDB 3CZN structure (red) and the one obtained after equilibration and QM/MM minimization (blue).

3. Further details of the CP QM/MM simulations

In this approach, the system is partitioned into quantum mechanics (QM) and molecular mechanics (MM) fragments. The dynamics of the atoms on the QM fragment depend on the electronic density, $\rho(r)$, computed with Density Functional Theory, while the dynamics of the atoms on the MM fragment are ruled by an empirical force field. The QM/MM interface is modeled by the use of a link-atom pseudopotential that saturates the QM region.[2] Seven link atoms were used, located at the Cα atoms of His90, Asp92, Asp204, Asp341, and His471. The electrostatic interactions between the QM and MM regions were handled via a fully Hamiltonian coupling scheme,[3] where the short-range electrostatic interactions between the QM and the MM

regions are explicitly taken into account for all atoms. An appropriately modified Coulomb potential was used to ensure that no unphysical escape of the electronic density from the QM to the MM region occurs. The electrostatic interactions with the more distant MM atoms were treated via a multipole expansion. Bonded and van der Waals interactions between the QM and the MM regions were treated with the standard AMBER force field.[4] Long-range electrostatic interactions between MM atoms were described with the P3M implementation[5] using a 64 x 64 x 64 mesh.

4. Further details of the metadynamics simulations

Metadynamics is a rare-events exploration method based on a dimension reduction: a set of collective variables that enclose the essential modes associated with the transitions in the analyzed process are defined.[6] Along the molecular dynamics simulation, small repulsive potential terms (Gaussian-like potentials) are added in the regions of the collective variable space that have already been explored. These added potentials make the system escape from already visited points to others, as soon as the biasing potential counterbalances the underlying free energy.[7] The method can be exploited for accelerating rare events, but also for mapping the free energy surface, which can be estimated as the negative of the sum of the Gaussian potential terms. This method has recently been applied to a variety of problems in the areas of biophysics, chemistry, and material science.[8]

The metadynamics simulations were done within the Car–Parrinello approach.[6b] In this scheme, the Car–Parrinello Lagrangian is extended by extra terms describing the fictitious dynamics of the collective variables. These additional fictitious particles are coupled (through a harmonic potential) to the value of the selected collective variables in the real system.[6] The mass for this fictitious particle and the force constant of the coupling potential were tested to ensure that the coupled particle naturally follows the value of the associated collective variable in the real system.

The collective variables used in the metadynamics simulations are a combination of coor-

dination numbers. We adopt the definition of coordination number of a species A with a species B, N_{coord} (A,B):[9,10]

$$N_{coord}(A,B) = \sum_{i=1}^{N_A} \frac{1}{N_A} \left[\sum_{i=1}^{N_B} \frac{1-(r_{ij}/d_{AB})^p}{1-(r_{ij}/d_{AB})^{(p+q)}} \right]$$

S1)

where N_A and N_B are the number of atoms of species A and B, respectively, r_{ij} are the interatomic distances, d_{AB} is a threshold distance for bonding, and p and q are exponents that determine the steepness of the decay of N_{coord} with respect to r_{ij}. This definition has been previously used in several studies of chemical and biochemical reactions using the metadynamics approach.[9–11]

The time evolution of the distances involved in the collective variables is presented in Figure S3. The forward and backward reactions were simulated to reconstruct the free energy surface. The evolution of the two collective variables CV1 and CV2 is found in Figure S4. To prevent product escape from the active site after the forward reaction occurred, a wall constraint was set on CV1 with a value of –0.93.
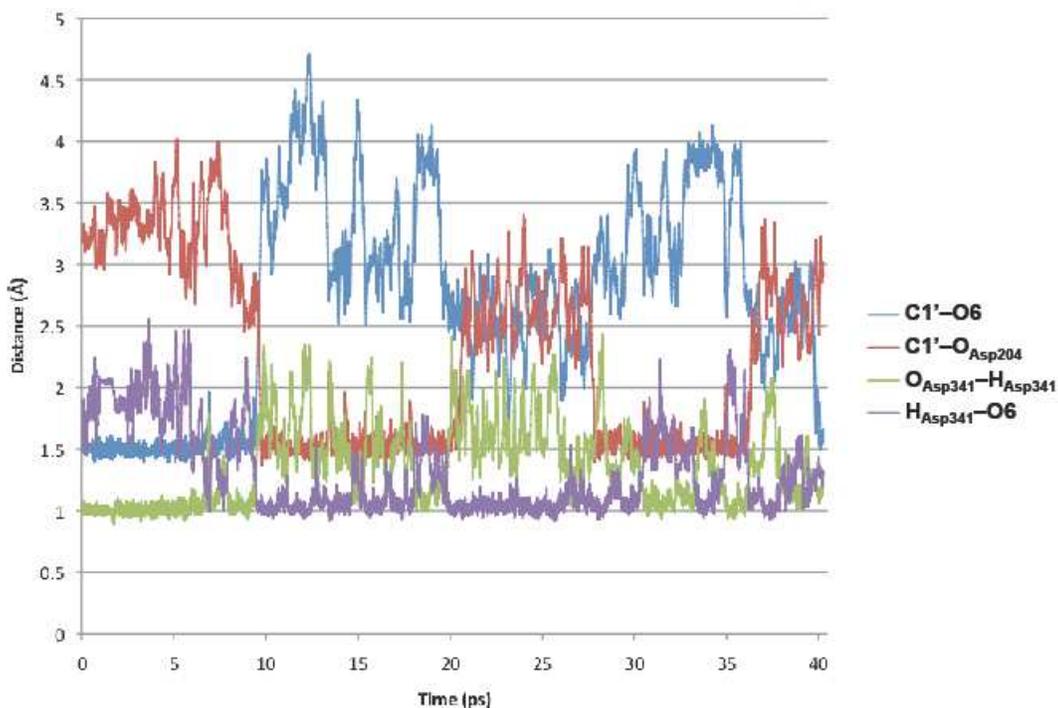
Figure S3. Time evolution of important distances along the metadynamics simulation. The metadynamics simulation started with the substrate–GMII complex in the reactant state, which was explored for ~4.8 ps until the forward reaction occurred. The scissile glycosidic bond was cleaved and the mannosyl–Asp204 covalent intermediate was formed. The system explored the product state (left side of the FES) for ~10.8 ps and then the middle region of the FES (corresponding to states with oxacarbenium ion (OCI)-like character) for a further 7.8 ps. Next, the system returned to the product state for another 8.4 ps. Finally, 4.1 ps later, the mannosyl glycosidic bond was reformed and the system reached the reactant state.
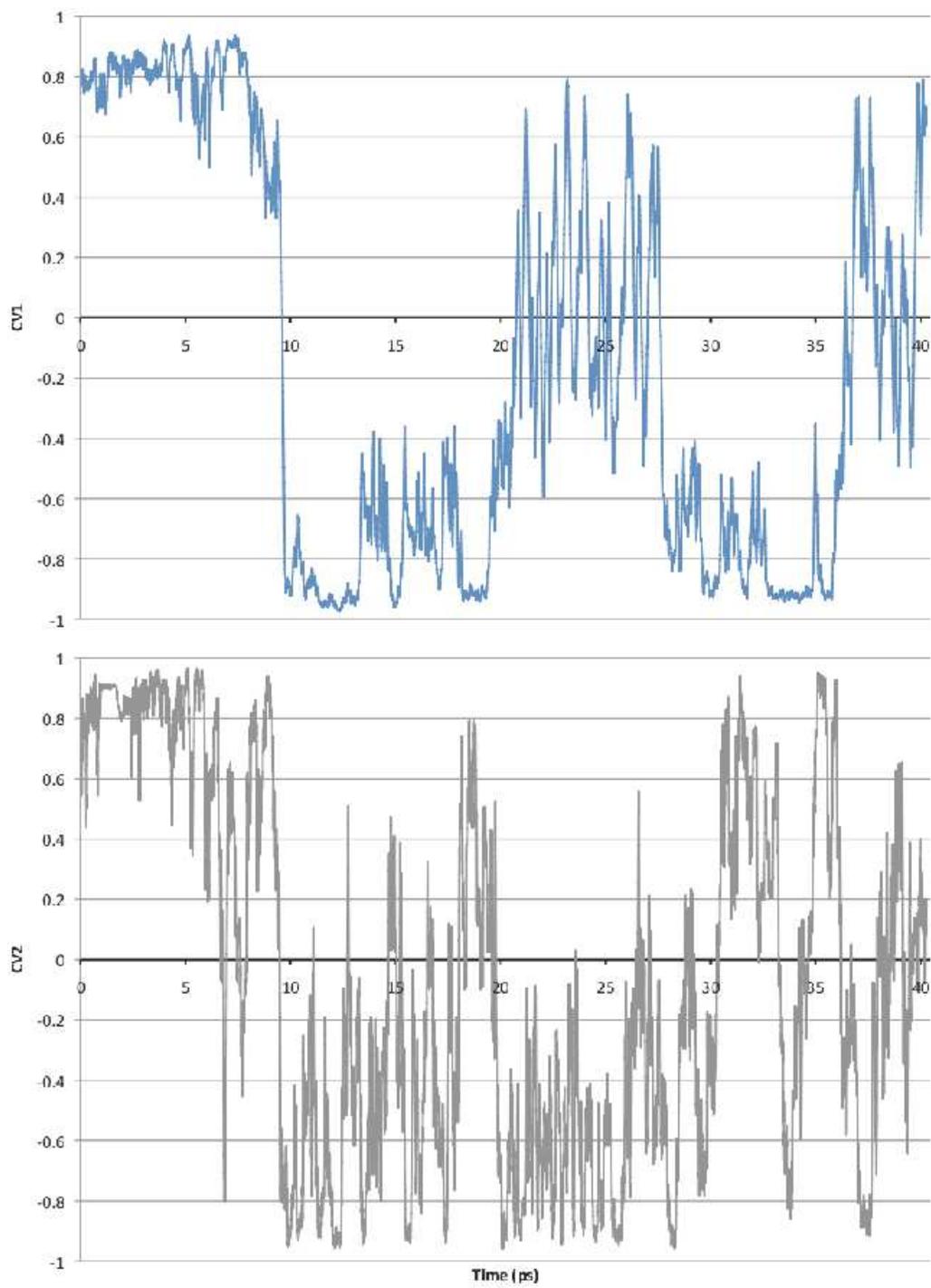
Figure S4. Time evolution of the collective variable along the metadynamics simulation.

The metadynamics simulation started with the substrate–GMII complex in the reactant state, which was explored for ~4.8 ps until the forward reaction occurred. The scissile glycosidic bond was cleaved and the mannosyl–Asp204 covalent intermediate was formed. The system explored the product state (left side of the FES) for ~10.8 ps and then the middle region of the FES (corresponding to states with oxacarbenium ion (OCI)-like character) for a further 7.8 ps. Next, the system returned to the product state for another 8.4 ps. Finally, 4.1 ps later, the mannosyl glycosidic bond was reformed and the system reached the reactant state.

Figure S5 shows the synchronous variation of the Zn and anomeric charges (δ^{an}) along the reaction pathway.
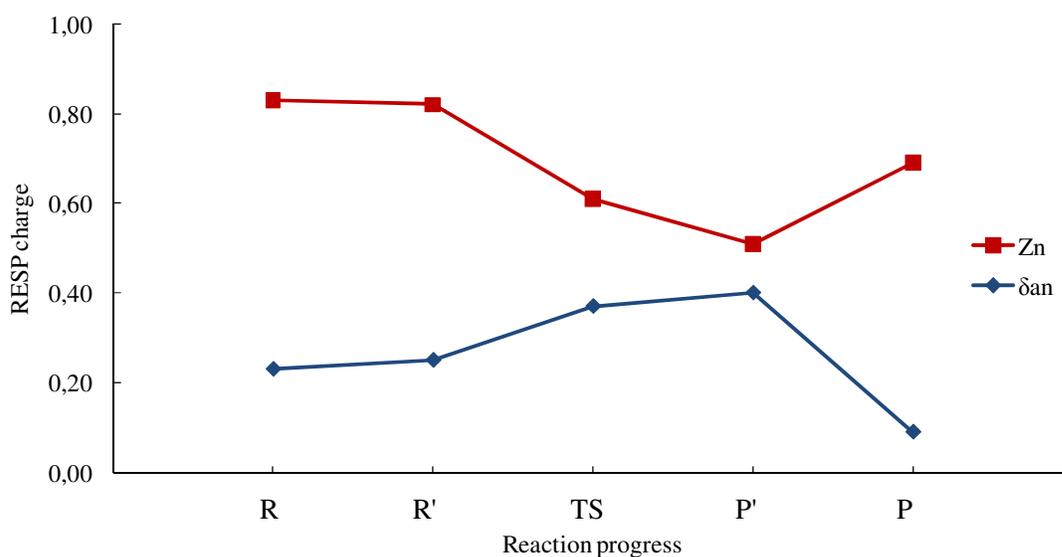

Figure S5. Variation of restrained electrostatic potential charges of the Zn ion and anomeric center along the reaction pathway.

5. Test calculations with different DFT functionals

To test the dependence of the results obtained with the exchange-correlation functional employed in the DFT calculations, we focused on a recent study[12] that compares the performance of a number of DFT functionals on the properties of small Zn model compounds, the largest being the tetracoordinated $Zn(NH_3)_4^{2+}$. This study showed that BLYP is a relatively poor functional, placed in positions 26 and 36 (over 39 functionals tested), depending on the property analyzed,

whereas the best performance is obtained with M05-2X.

To have a good measure of the error due to the functional employed, it is important to perform the calculations using the same model (complete enzyme) and methodology (CPMD QM/MM, with a 70-Ry PW basis set), varying only the DFT functional. Among the functionals that are available in the CPMD QM/MM package, only the BP86 functional is within the list of functionals tested by Amin et al.[12] and, interestingly, it is better ranked (in positions 13 and 35) than BLYP. Therefore we considered the BP86 functional in our GH38 test calculations. In addition, we considered the PBE functional (not included in the list of ref. 12), which has been used previously in GH mechanisms. A recent evaluation of basis sets for carbohydrates[13] concludes that PBE performs well among GGA functionals and better than the popular B3LYP. Therefore, test calculations of the optimized structures of the stable species (reactants and products) were performed with these two functionals. The results obtained (Table S2) show that the Zn-ligand distances differ on average by 0.025 Å for reactants and by 0.017 Å for products. The three functionals give similar values of the given reaction exothermicity (with differences of 1.88 and 1.22 kcal/mol for BP and PBE, respectively).

Table S2. Zn–Ligand Distances (Å) of the Optimized Structures

Bond	Reactant (R)	TS	Product (P)
BLYP			
$Zn–O_{Asp204}$	2.116	2.112	2.943
$Zn–O_{Asp92}$	2.219	2.243	2.052
$Zn–N_{His90}$	2.147	2.136	2.113
$Zn–N_{His471}$	2.152	2.137	2.113
Zn–O2'	2.339	2.331	2.196
Zn–O3'	2.276	2.312	2.279
PBE			
$Zn–O_{Asp204}$	2.105	2.089	2.930
$Zn–O_{Asp92}$	2.258	2.266	2.042
$Zn–N_{His90}$	2.164	2.141	2.124
$Zn–N_{His471}$	2.156	2.130	2.095
Zn–O2'	2.267	2.284	2.170
Zn–O3'	2.222	2.291	2.248
BP			
$Zn–O_{Asp204}$	2.096	2.088	2.930
$Zn–O_{Asp92}$	2.212	2.262	2.042
$Zn–N_{His90}$	2.140	2.122	2.115
$Zn–N_{His471}$	2.162	2.127	2.099
Zn–O2'	2.309	2.297	2.175
Zn–O3'	2.242	2.295	2.241

It is also important to verify that the trend of atomic charges obtained from our calculations (especially the synchronous variation of the Zn and anomeric charges (δ^{an}) along the reaction pathway) is reproduced when using different DFT functionals. Therefore we took several points along the reaction pathway (in particular points R, R', TS, P' and P in Figure 5 of the manuscript) and performed single-point calculations using the same functionals considered above (BLYP, BP86, and PBE) as well as M05-2X and B3LYP. The calculations with M05-2X and B3LYP were performed in a reduced model containing all atoms treated quantum-mechanically in the QM/MM calculations (98 atoms) and the NWCHEM program[12] was employed. Because of the system size, it was not possible to use a basis set for the Zn atom as large as the one considered by Amin et al. We used a 6-31G* basis set for all atoms except the Zn atom, which was described with an effective pseudopotential (LanL2DZ). The use of a small basis set is not expected to be a major problem here because we are only interested in the variation of atomic charges. The
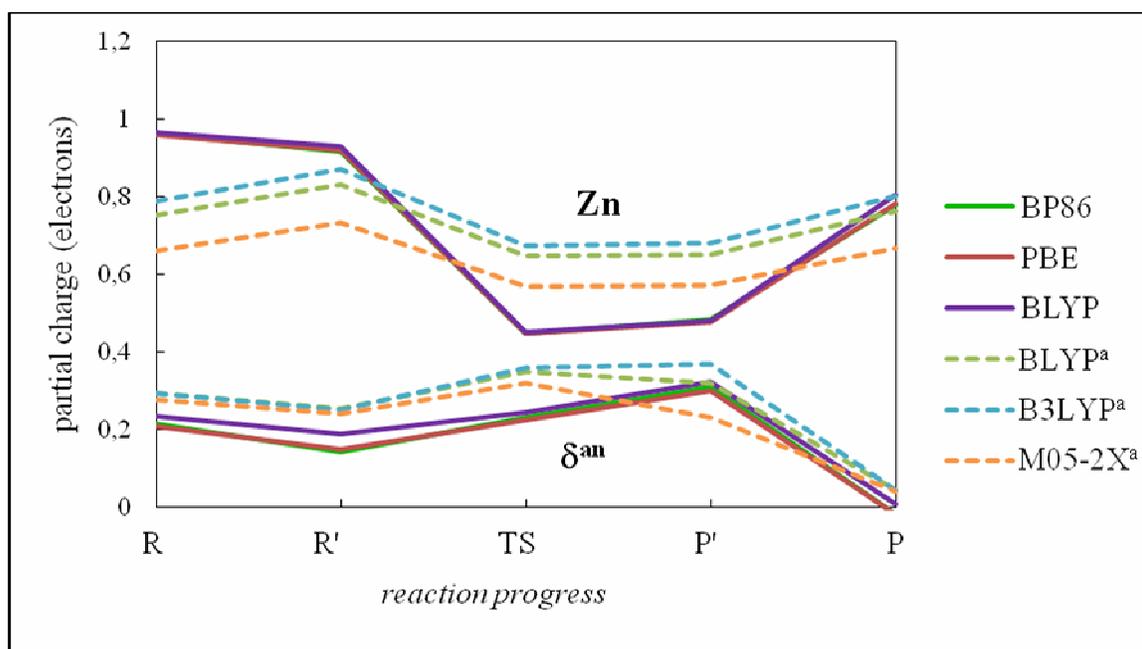

Figure S6. Variation of restrained electrostatic potential charges of the Zn ion and anomeric center along the reaction pathway, considering different functionals. The broken lines correspond to results with the reduced model and basis set (see text). This is also indicated with a superscript on the functional's name.

results obtained are shown in Figure S6. There is a synchronous variation of the Zn and anomeric charges (δ^{an}) along the reaction pathway independent of the functional.

References

(1) Shah, N.; Kuntz, D. A.; Rose, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9570–9575.

(2) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U. ; Sebastiani, D. *J. Chem. Phys.* **2005**, *122*, 013133 (1–6).

(3) Laio, A.; vandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941–6947.

(4) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1–41.

(5) Hünenberger, P. *J. Chem. Phys.*, **2000**, *113*, 10464−10476.

(6) (a) Laio, A.; Parrinello, M. *Proc. Nat. Acad. Sci. U.S.A.* **2002**, *99*, 12562–12566. (b) Iannuzzi, M.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2003**, *90*, 238302(4).

(7) (a) Ensing, B.; Laio, A.; Parrinello, M.; Klein, M. L. *J. Phys. Chem. B* **2005**, *109*, 6676–6687. (b) Laio, A.; Rodriguez-Fortea, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. *J. Phys. Chem. B* **2005**, *109*, 6714–6721. (c) Ensing, B.; De Vivo, M.; Liu, Z. W.; Moore, P.; Klein, M. L. *Acc. Chem. Res.* **2006**, *39*, 73–81.

(8) Recent applications include: (a) Oganov, A. R.; Martoňak, R.; Laio, A.; Raiteri, P.; Parrinello, M. *Nature* **2005**, *438*, 1142–1144. (b) Blumberger J.; Ensing B.; Klein M. L. *Angew. Chem. Int. Engl. Ed.* **2006**, *45*, 2893–2897 (c) Fiorin G.; Pastore A.; Carloni P.; Parrinello M. *Biophys. J.* **2006**, *91*, 2768–2777 (d) Martoňak R.; Donadio D.; Oganov, A. R.; Parrinello M. *Nature Materials* **2006**, *5*, 623–626 (e) Rodriguez-Fortea, A.; Iannuzzi, M.; Parrinello, M. *J. Phys. Chem. B* **2006**, *110*, 3477–3484. (f) Boero, M.; Ikeda, T. ; Ito, E. ; Terakura, K. *J. Am. Chem. Soc.* **2006**, *128*, 16798–16807. (g) Rodriguez–Fortea, A.; Iannuzzi, M.; Parrinello, M. *J. Phys. Chem. C* **2007**, *111*, 2251–2258. (h) Biarnés, X.; Ardèvol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2007**,

129, 10686–10693. (i) Gunaydin, H.; Houk, K. N. *J. Am. Chem. Soc.* **2008**, *130*, 10036–10037.

(9)   Iannuzzi, M.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.*, **2003**, *90*, 238302.

(10) Boero, M.; Ikeshoji, T.; Liew, C. C.; Terakura, K.; Parrinello, M. *J. Am. Chem. Soc.* **2004**, *126*, 6280–6286.

(11) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. *J. Phys. Chem. B* **2006**, *110*, 3533–3539.

(12) Amin, E. A.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 75–85.

(13) Csonka, G. I.; French, A. D.; Johnson, G. P.; Stortz, C. A. *J. Chem. Theory Comput.* **2009**, *5*, 679–692.

(14) E. J. Bylaska, W. A. de Jong, K. Kowalski, T. P. Straatsma, M. Valiev, D. Wang, E. Apra, T. L. Windus, S. Hirata, M. T. Hackler, Y. Zhao, P.-D. Fan, R. J. Harrison, M. Dupuis, D. M. A. Smith, J. Nieplocha, V. Tipparaju, M. Krishnan, A. A. Auer, M. Nooijen, E. Brown, G. Cisneros, G. I. Fann, H. Fruchtl, J. Garza, K. Hirao, R. Kendall, J. A. Nichols, K. Tsemekhman, K. Wolinski, J. Anchell, D. Bernholdt, P. Borowski, T. Clark, D. Clerc, H. Dachsel, M. Deegan, K. Dyall, D. Elwood, E. Glendening, M. Gutowski, A. Hess, J. Jaffe, B. Johnson, J. Ju, R. Kobayashi, R. Kutteh, Z. Lin, R. Littlefield, X. Long, B. Meng, T. Nakajima, S. Niu, L. Pollack, M. Rosing, G. Sandrone, M. Stave, H. Taylor, G. Thomas, J. van Lenthe, A. Wong, and Z. Zhang, "NWChem, A Computational Chemistry Package for Parallel Computers, Version 5.0" (2006), Pacific Northwest National Laboratory, Richland, WA, USA.