Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 338 (2003) 865–877
CARBOHYDRATE RESEARCH
www.elsevier.com/locate/carres

On the reaction pathways and determination of transition-state structures for retaining α-galactosyltransferases

Isabelle André,[a,*] Igor Tvaroška,[a,b,*] Jeremy P. Carver[a]

[a]*GlycoDesign Inc., 480 University Avenue, Suite 900, Toronto, ON, Canada M5G 1V2*
[b]*Institute of Chemistry, Slovak Academy of Sciences, Bratislava, Slovak Republic*

Received 8 April 2002; accepted 17 January 2003

Abstract

The catalytic mechanism of retaining glycosyltransferases is not yet completely understood, but one possible mechanism, by analogy with retaining glycosidases, is a double-displacement mechanism via a covalent glycosyl–enzyme intermediate (CGE). We have investigated various reaction pathways for this mechanism using non-empirical quantum mechanical methods. Because a double-displacement mechanism presumes a reaction happening in two steps, we have used predefined reaction coordinates to calculate the potential energy surface describing each step of the mechanism. By investigating several potential candidates to act as a catalytic base, this study attempts to shed some light on the unclear mechanism of the second step of the reaction. All intermediates and transition states on the reaction pathways were characterized using basis sets up to the DFT/B3LYP/6-311++G**//DFT/B3LYP/6-31G* level. Reaction pathways and structural changes were compared with the results previously obtained for inverting glycosyltransferases. The outcome of this study indicates, that among the reaction models investigated, the energetically favorable one is also the most plausible given the existing experimental data. This model requires the presence of only one catalytic acid in the active site with the UDP functioning as a general base in the second step of the reaction. This mechanism is in agreement with both kinetic data in the literature and the description of X-ray structures of retaining glycosyltransferases solved up to today.

Keywords: Glycosyltransferases; Catalytic mechanism; Reaction pathways; Transition states; Intermediates; Metal cofactor

1. Introduction

Glycosyltransferases of the Leloir pathway[1] are a large class of enzymes that catalyzse the transfer of a glycosyl residue from a specific activated nucleotide sugar donor to the hydroxyl group of a particular sugar acceptor. These enzymes usually reveal a defined specificity for both of their substrates, the sugar acceptor and the sugar donor. The glycosylation reaction can mechanistically proceed with either inversion or retention of stereochemistry at the anomeric carbon (C-1) of the donor sugar. To date, glycosyltransferases have been classified into about 60 families on the basis of sequence analysis and substrate/product stereochemistry,[2] of which about 18 lead to retention of the configuration at C-1. Reaction mechanisms of glycosidases and glycosyltransferases are commonly assumed to be rather similar, the difference being in the acceptor molecule to which the glycosyl residue is transferred, a water molecule in the case of glycosidases[3,4] and a free hydroxyl group of a specific sugar acceptor in the case of glycosyltransferases.[5–7] However, while the literature abounds of structural information on glycosidases, very little is yet known on the mechanism of glycosyltransferases. In the past few years, a large interest has been given to glycosyltransferases, and several crystal structures of glycosyltransferases have been reported in the literature.[8–21] The mechanism of inverting glycosyltransferases is now rather well documented, and it is thought to be a single nucleophilic displacement of the nucleotide by one of the hydroxyl groups of the acceptor.[22,23] The mechanism, however, of retaining glycosyltransferases remains unclear and raises a lot of interest. Indeed, only six of the X-ray structures of glycosyl-

* Corresponding author. Tel.: +416-593-6027; fax: +416-593-8998
E-mail addresses: isabelle@glycodesign.com (I. André), igor@glycodesign.com, chemitsa@savba.sk (I. Tvaroška).

0008-6215/03/$ - see front matter
doi:10.1016/S0008-6215(03)00050-8

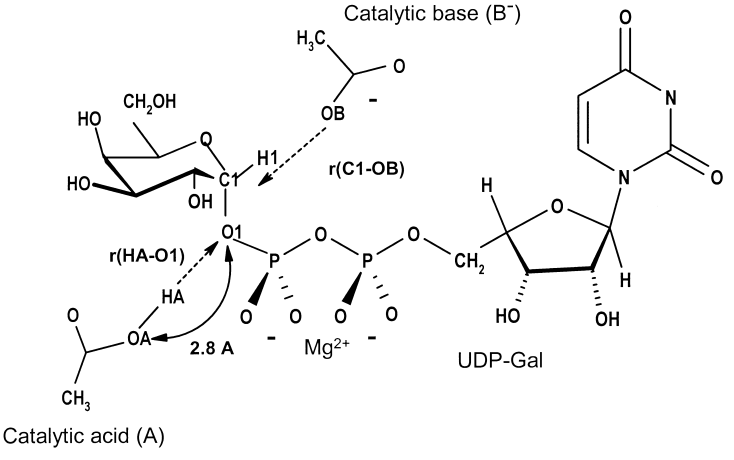
transferases available to date belong to retaining galactosyltransferases.[15,16,18–20] Like for glycosidases, the reaction catalyzed by retaining glycosyltransferases has long been thought to proceed via a double-displacement mechanism involving the participation of two catalytic residues. However, X-ray structures of retaining glycosyltransferases available to this date[15,16,18–20] suggest the possibility of alternative mechanisms. For example, some double-displacement mechanisms might involve only one catalytic residue as suggested for bovine α-(1→3)-GalT,[16] Glycogenin[19] and for GTA and GTB.[18] All the more intriguing are the recent results on the mechanism of α-(1→4)-GalT (LgtC)[15] pointing out to a mysterious mechanism that might simply not require any participation of a catalytic residue. The difficulty to identify a catalytic base in the active site, as well as the absence of any convincing evidence for the existence of a covalent intermediate, led to the suggestion that this enzyme might rather proceed through a single, front-side displacement reaction, also known as an S_Ni mechanism, which has previously been invoked for glycogen phosphorylases.[5] A similar mechanism has recently been suggested for OtsA glucosyltransferase.[20] However, the lack of experimental evidence does not yet permit one to clearly conclude on the actual type of mechanism involved.

This work is a continuation of our series of papers aiming to investigate the catalytic mechanism of glycosyltransferases using non-empirical quantum chemical methods. After describing the mechanism of inverting *N*-acetylglucosaminyltransferases,[24] we report herein high-level quantum chemical results on various possible pathways of a double-displacement mechanism for the retaining α-galactosyltransferases. Reaction pathways investigated herein involve one or two catalytic residues intervening in a double-displacement mechanism.

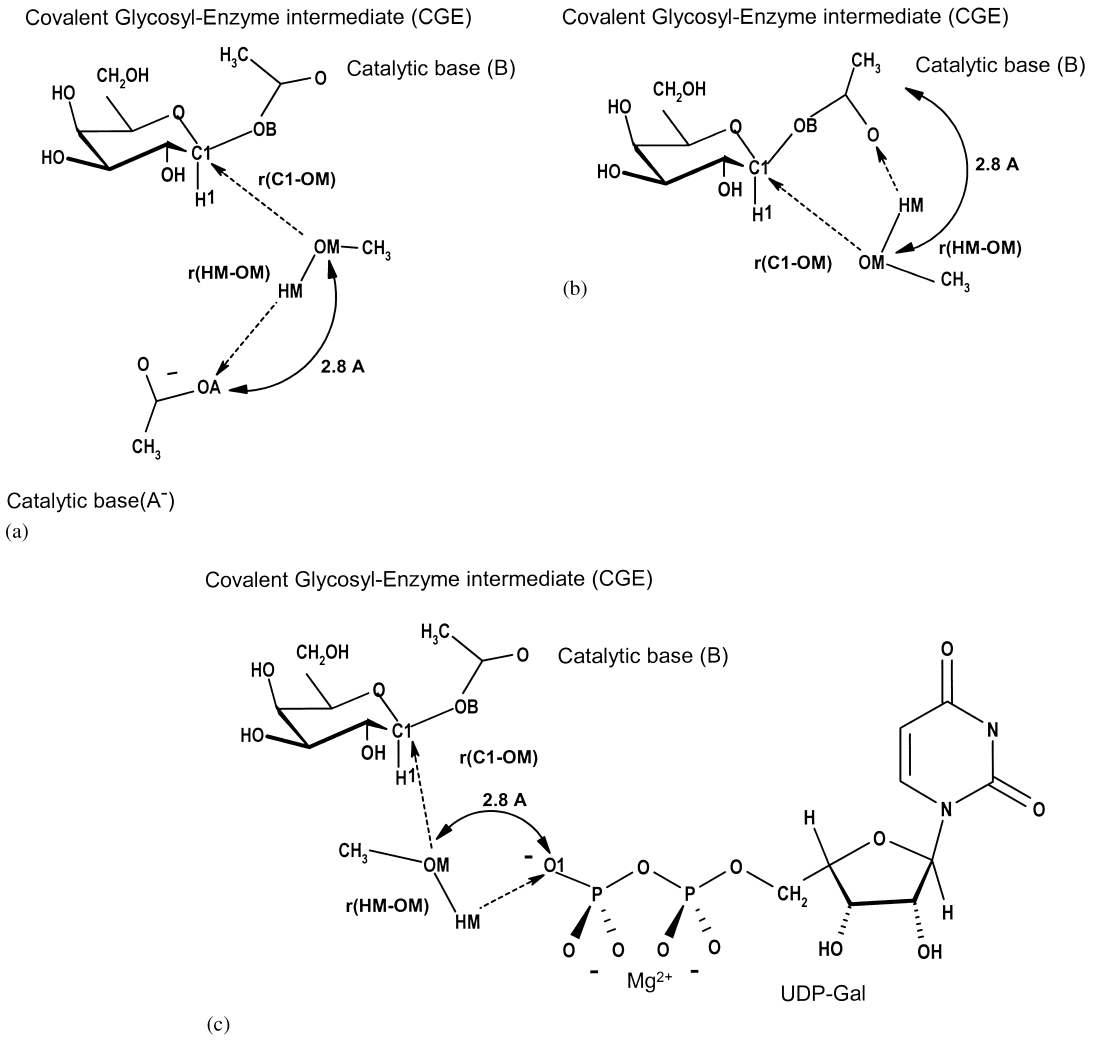
2. Model and computational procedures

By analogy with retaining glycosyl hydrolases, a double-displacement mechanism of retaining glycosyltransferases supposedly proceeds in two steps. First, the enzyme nucleophile (catalytic base B^-) is required for attack of the donor on the anomeric carbon (C-1) to form a covalent glycosyl–enzyme intermediate (CGE). A second amino acid (A) might be involved in this step to facilitate the departure of the leaving group (Scheme 1). In a covalent glycosyl–enzyme intermediate (CGE), the configuration at the C-1 atom is inverted to the equatorial orientation. In the second step (Scheme 2(a–c)), CGE is broken by a direct displacement of a new glycosidic linkage with the reactive hydroxyl group of the acceptor. The configuration at the C-1 is inverted from equatorial back to axial. In this step, a general base is required to deprotonate the hydroxyl group of the acceptor. Several scenarios involving distinct general catalytic bases can be considered for this deprotonation to proceed: (1) the presence and participation of a second catalytic residue (catalytic base A^- on Scheme 2(a)); (2) the participation of the catalytic base B covalently bound to the glycosyl residue in the glycosyl–enzyme intermediate (Scheme 2(b)); or (3) the proton transfer to the phosphate group of the leaving group UDP (Scheme 2(c)).

To analyze computationally a double-displacement mechanism of the transfer of a galactosyl residue by an α-galactosyltransferase, distinct structural models were used for each of the steps of the glycosyl-transfer reaction. All models are composed of all the essential molecules, or their fragments, assumed to be involved in a particular reaction step (Schemes 1 and 2(a–c)). Reaction pathways for each step of the reaction were calculated separately. The model describing the first step of the reaction is shown on Scheme 1. It contains the complete sugar–donor molecule, UDP-Gal, a diva-



Scheme 1. Schematic representation of the structural model used to describe the first step of the double-displacement mechanism of retaining α-galactosyltransferases.

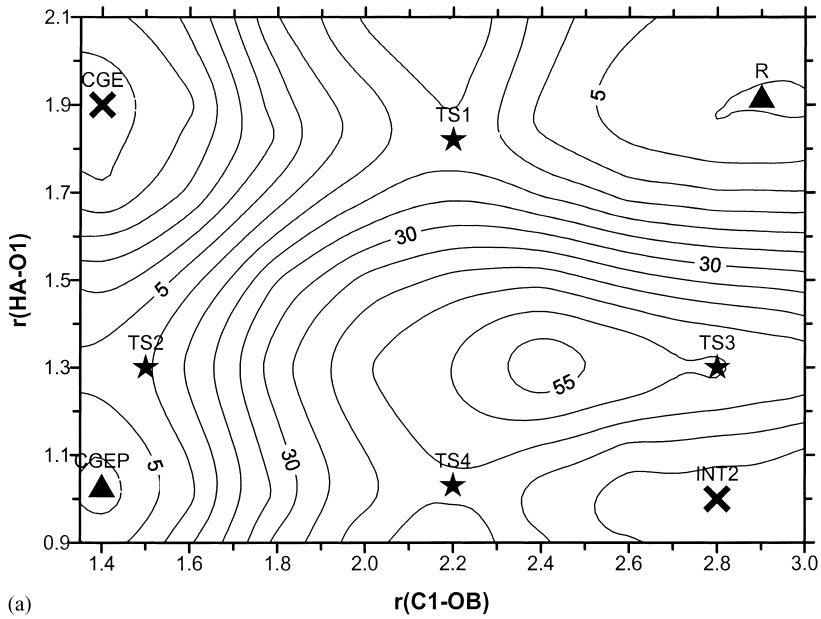


Scheme 2. Schematic representation of the structural models used to describe the second step of the double-displacement mechanism. (a) Assisted by a second catalytic base A^- in the active site. (b) Assisted by the catalytic acid B. (c) Assisted by UDP.

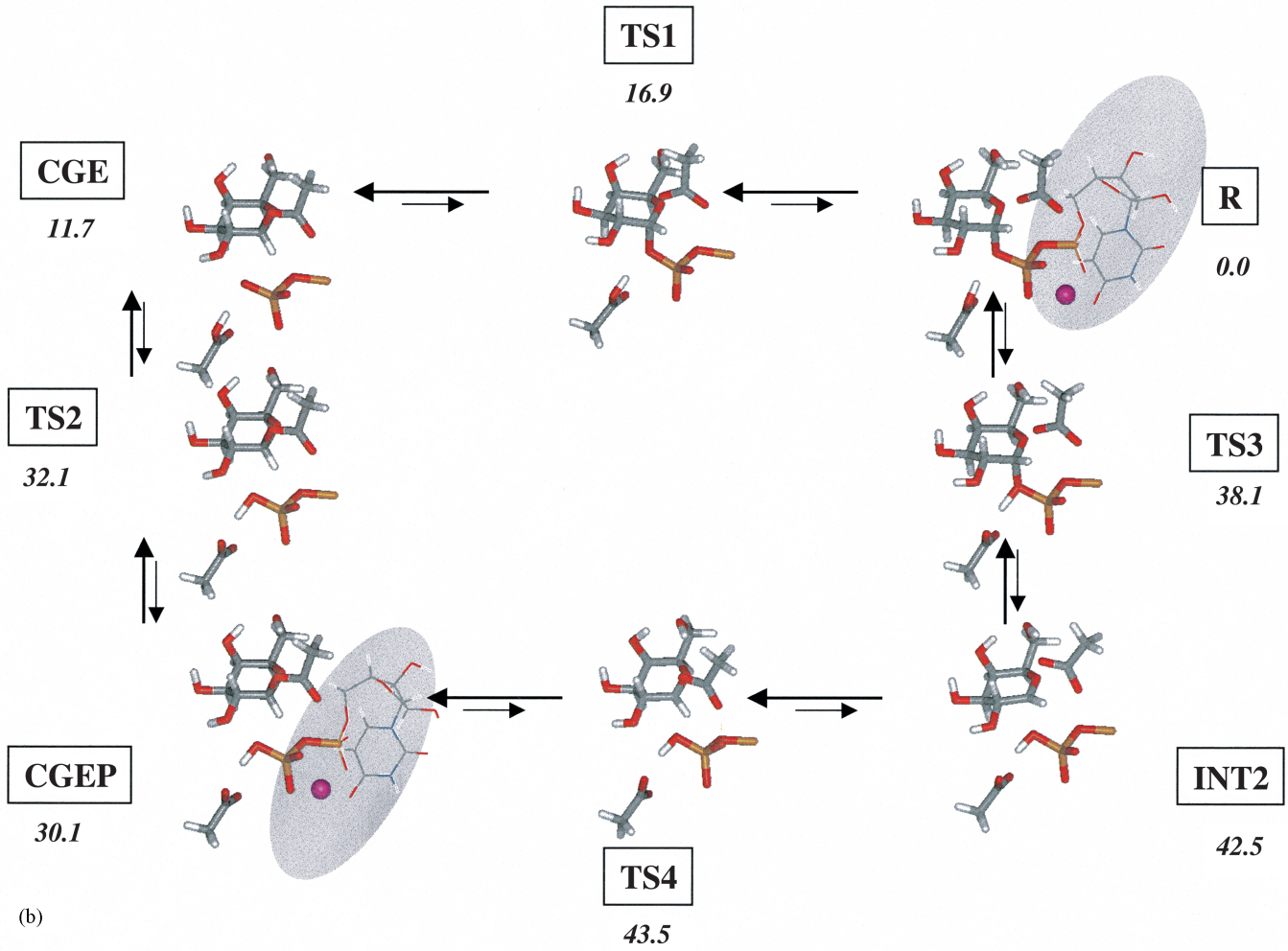
lent metal cofactor modeled by Mg^{2+}, as well as the essential parts of the catalytic acid A and catalytic base B^- represented by acetic acid and acetate molecules. This first model consists of 74 atoms and has an overall charge of -1. Various structural models have been used to investigate the second step of the reaction. Each of these models uses a different potential candidate to function as a general base in the proton transfer from the acceptor hydroxyl onto the base. These are described in Scheme 2(a–c) with the acceptor hydroxyl group modeled by methanol (HMOM-CH_3). The number of atoms in these distinct models varied between 36 and 74, and the number of basis functions from 291 to 1217, depending on the complexity of the model and the basis set used.

In the construction of our models, we have used a similar strategy to the one used in our earlier study on inverting *N*-acetylglucosaminyltransferases.[24] The two catalytic amino acids in the model were placed in an arrangement that emulates their orientation in the active site of retaining glycosyl hydrolases[3] and where the two carboxylates are located ~ 5.5 Å apart. In the first model (Scheme 1), the two amino acids are initially located about 2.8 Å away from the anomeric carbon (C-1). In the models used to describe the second step, the methanol oxygen atom OM, representing the reactive hydroxyl of the sugar–acceptor, was initially placed at 3.2 Å from the anomeric carbon (C-1) and at 2.8 Å from the oxygen of a general base (Scheme 2(a–c)).

Energy of the models described on Schemes 1 and 2(a and c) and calculated as a function of the reaction-coordinates is represented by means of three potential energy surfaces (PES's) (Fig. 1(a), Fig. 2(a) and Fig. 4(a), respectively). Energy of models represented in Scheme 2 was calculated as a function of the reaction coordinate representing the proton transfer and the nucleophilic attack, respectively. Each calculated point



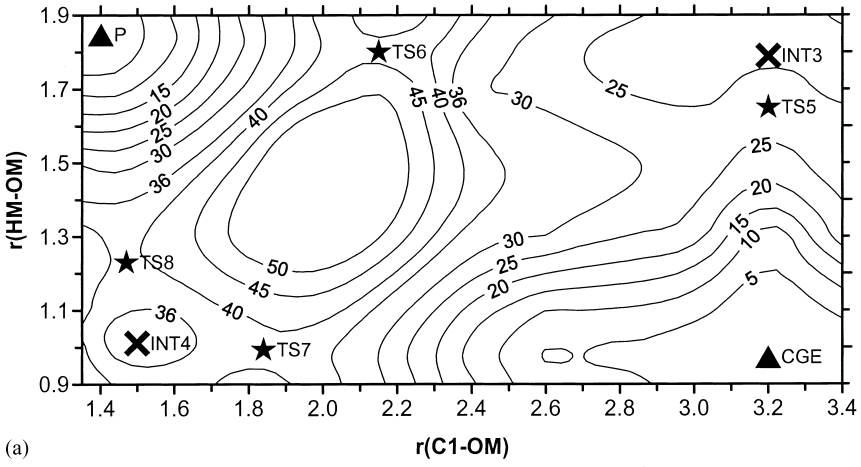
(a)



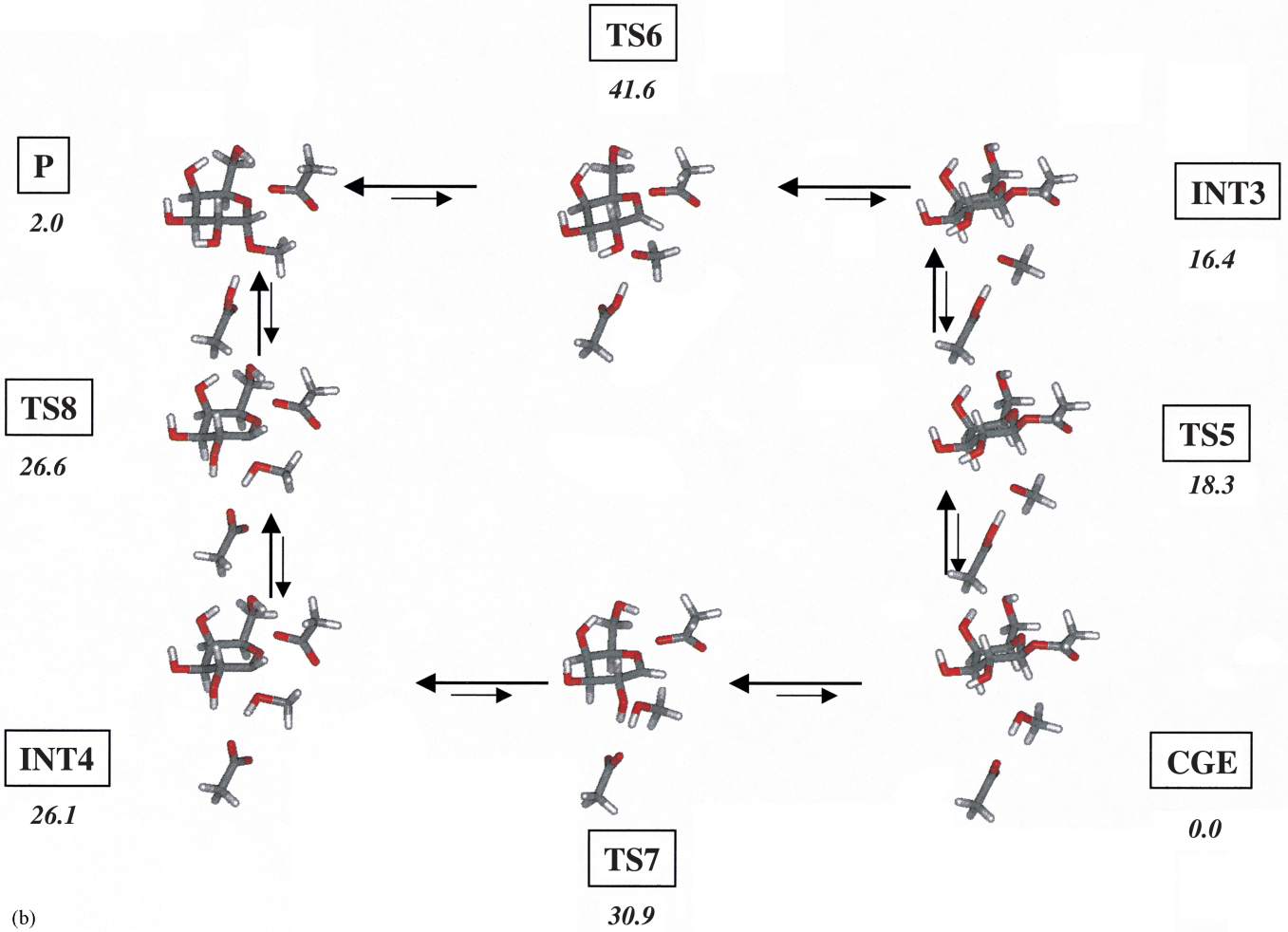
(b)

Fig. 1. (a) PES calculated at the HF/6-31G* level and corresponding to the first step of the mechanism involving a double-displacement (Scheme 1). (b) Geometrical representation of the different stationary points calculated at the DFT/B3LYP/6-31G* level. Numbers in italics represent relative energies (in kcal/mol) at the DFT/B3LYP/6-31++G**//DFT/B3LYP/6-31G* level. R, TS, INT, CGE, CGEP, and P represent the reactants, transition states, intermediates, glycosyl–enzyme intermediates, and products, respectively.

on the PES's of Fig. 1(a), Fig. 2(a), and Fig. 4(a) corresponds to the optimized structure of the model for the given distance used as reaction-coordinates. These distances were varied by 0.2-Å increments within the 3.0–1.35 Å range for r_{C1-OB}, 2.1–0.9 Å range for r_{HA-O1}, 0.9–1.9 Å for r_{HM-OM}, 1.0–1.8 Å for r_{HM-O1},



(a)



(b)

Fig. 2. (a) PES calculated at the HF/6-31G* level and corresponding to the second step (displacement step) of the mechanism involving a double-displacement and the assistance of a second catalytic residue in the active site (Scheme 2(a)). (b) Geometrical representation of the different stationary points calculated at the DFT/B3LYP/6-31G* level. Numbers in italics represent relative energies (in kcal/mol) at the DFT/B3LYP/6-31++G**//DFT/B3LYP/6-31G* level. R, TS, INT, CGE, and P represent the reactants, transition states, intermediates, glycosyl–enzyme intermediate, and products, respectively.

and within the 3.4–1.35 Å range for r_{C1-OM}. During the optimization, all geometrical parameters of the reactants were optimized with the exception of those defining the location and orientation of the two catalytic amino acids. As a result, each point on the PES represented by fixed values of the r_{C1-OB}, r_{HA-O1}, r_{C1-OM}, r_{HM-O1} and r_{HM-OM} distances have all their geometrical variables adjusted to their most stable values. The location of the local minima and transition barriers on the PES's is only approximate, and for that reason a further optimization of the stationary points with no constraints on the distances used as reaction-coordinates was required. These stationary points represent structures of the intermediates and transition states found on the different PES's and along the different reaction pathways. However, in order to avoid any confusion, we will use throughout the paper the same acronyms: TSi, and INTi, CGE, or CGEP, respectively, for the barriers located on PES's and for the stationary points later refined. R, P, and Prod are the reactants and products of the reaction, respectively.

The non-empirical calculations were carried out with the JAGUAR program.[25] The PES's were calculated at the SCF level with the 6-31G* basis set. The effects of electron correlation on the PES were estimated using the B3LYP density functional method.[26] Ultimately, selected geometries were used to estimate the effect of the basis set by calculating their single point energy with the 6-31++G** and 6-311++G** basis sets.

3. Results and discussion

Reaction pathways described throughout the manuscript are based on the assumption that retaining glycosyltransferases perform catalysis via a double-displacement mechanism involving the formation and the subsequent breakdown of a covalent glycosyl–enzyme intermediate (CGE). By analogy with retaining glycoside hydrolases, the presence of two catalytic residues in the vicinity of the reaction site would be required for this so-called classical mechanism (Scheme 1): a nucleophile (catalytic base B^-) operating a nucleophilic attack on the anomeric C-1 in order to form a covalent glycosyl–enzyme intermediate with inverted configuration, and a catalytic acid A protonating the glycosidic oxygen O-1 and thus allowing the departure of the leaving group (UDP). In the second step of this mechanism (Scheme 2(a)), the catalytic acid A of Scheme 1 has now become a proton acceptor, and it acts as a general base A^- deprotonating the reactive hydroxyl of the acceptor sugar (HM-OCH_3 in the model), which is performing a second inverting nucleophilic attack at C-1, resulting in a product with retained configuration. However, up to now such a classical double-displacement mechanism is not supported by any experimental evidence. Four of the available X-ray structures of retaining glycosyltransferases show the presence of only one catalytic residue in the active site,[16,18,19] which raises two questions to be answered. First, how can possibly the first step proceed without the contribution of a catalytic acid? Second, what is the most likely candidate that can act as a general base in the second step? We will attempt to answer these questions by evaluating various reaction pathways for each step separately.

The two steps of the reaction are described by means of three PES's represented in the form of two-dimensional reaction-coordinate contour diagrams in Fig. 1(a), Fig. 2(a) and Fig. 4(a). Distances along the x-axis determine the formation and scission of a new C-1–Ox linkage, C-1–OB bond in the first step of Fig. 1(a) and C-1–OM bond in the second step of Fig. 2(a) and Fig. 4(a). While such C-1–Ox distances along the x-axis represent the nucleophilic attack of either the catalytic base oxygen OB or the acceptor oxygen OM on the anomeric C-1, the distances along the y-axis characterize the proton-transfer processes from the catalytic acid A on the glycosidic oxygen O-1 in Fig. 1(a) and from the acceptor to the catalytic base A^- in Fig. 2(a), or from the acceptor to UDP in Fig. 4(a), respectively.

Different reaction pathways can generally be identified on these PESs. The reaction pathways parallel to the vertical and horizontal axes describe particular steps in a stepwise mechanism, while the reaction pathways following the diagonal across the PES represent a concerted mechanism. It is clear that the calculated two-dimensional PES's only represent a section of the potential energy hypersurface describing the entire catalytic reaction. Nevertheless, several qualitative conclusions can be formulated from the calculated two-dimensional PES's displayed on Fig. 1(a), Fig. 2(a) and Fig. 4(a), and they will be discussed below. Optimized structures of the different stationary points found along the reaction pathways of these PES's, and determined at the DFT/B3LYP/6-31G* level, are given in Fig. 1(b) and Fig. 2(b). The transition state determined on the PES of Fig. 4(a) is shown on Fig. 4(b). In our previous work,[27–29] we have shown that the inclusion of electron correlation results in slightly different magnitudes for the bond lengths and the increase of the basis set decreases the relative energy of the minima. Therefore, we will essentially base our discussions on the structures calculated at the DFT/B3LYP/6-31G* level and their energy estimated with the DFT/B3LYP/6-31++G**//DFT/B3LYP/6-31G* method. The energy for the pathway of the proposed mechanism has been calculated with the DFT/B3LYP/6-311++G**//DFT/B3LYP/6-31G* method. More detailed structural information on each stationary point is given for reference in Tables 1 and 2, while their relative energies (ΔE), determined at various levels, are listed in Tables

Table 1
Ab initio calculated geometrical parameters of the points observed on the PES described on Fig. 1(a) (first step) at the DFT/B3LYP/6-31G* level [a]

Conformer	Bond lengths				Bond angle	Torsion angle
	C-1–OB	HA–O-1	C-1–O-1	C-1–O-5	C-1–O-5–C-5	C-1–O-5–C-5–C-4
R	2.800	1.830	1.493	1.382	123.4	36.3
CGE	1.576	1.799	3.476	1.482	118.2	38.4
INT2	2.800	1.066	2.069	1.290	123.5	37.0
TS1	2.200	1.820	1.640	1.366	119.9	47.6
TS2	1.564	1.218	3.367	1.491	118.4	38.0
TS3	2.800	1.300	1.553	1.359	122.2	37.3
TS4	2.200	1.105	2.778	1.319	121.8	34.1
CGEP	1.563	1.095	3.422	1.490	118.5	38.4

[a] Lengths in Å, angles in °.

Table 2
Ab initio calculated geometrical parameters of the points observed for the second step of reaction at the DFT/B3LYP/6-31G* level [a]

Conformer	Bond lengths				Bond angle	Torsion angle
	C-1–OB	HM–OM	C-1–OM	C-1–O-5	C-1–O-5–C-5	C-1–O-5–C-5–C-4
Scheme 2(a)						
CGE	1.424	1.011	3.178	1.426	111.2	62.5
INT3	1.431	1.731	2.985	1.427	110.8	63.4
INT4	2.833	1.011	1.507	1.377	124.7	11.7
TS5	1.429	1.512	3.063	1.427	111.5	62.3
TS6	1.893	1.765	2.162	1.367	117.6	47.2
TS7	2.376	1.040	1.900	1.324	120.3	37.3
TS8	2.824	1.093	1.506	1.379	124.4	15.9
P	2.902	1.809	1.419	1.429	124.0	27.4
Scheme 2(c)						
CGE	1.574	0.997	2.800	1.486	120.2	37.1
TS9	2.279	1.022	2.190	1.322	121.3	48.7
Prod	2.993	3.847	1.418	1.417	119.9	44.4

[a] Lengths in Å, angles in °.

3 and 4. Fig. 3(a), Fig. 3(b) and Fig. 5 summarize the various possible reaction pathways found for the transfer of Gal as catalyzed by retaining α-galactosyltransferases via a double-displacement mechanism.

3.1. The formation of a covalent glycosyl–enzyme intermediate

The first step of the mechanism illustrated on Scheme 1, characterizes the attack of a nucleophile (catalytic base B^-) on the anomeric C-1 of UDP-Gal (sugar donor). In the case of a classical mechanism (two catalytic residues in the active site), the proton transfer from a catalytic acid A to the glycosidic oxygen O-1 precedes the nucleophilic attack. The PES describing the various reaction pathways for this step of the reaction mechanism, and calculated at the HF/6-31G* level, is given in Fig. 1(a). Distances plotted along both axes of the contour map describe, horizontally, the nucleophilic attack of the catalytic base B^- on the anomeric carbon C-1 of Gal, and vertically, the proton (HA) transfer from the catalytic acid A to the glycosidic oxygen O-1. Four energy barriers (TS1–TS4) located in valleys along the borders can be identified on the contour map. The reaction proceeds in a stepwise manner from reactants (R) to the covalent glycosyl–enzyme intermediates (CGE and CGEP). These two intermediates differ only by the proton transferred from the catalytic acid A to the glycosidic oxygen O-1, and thus distinguish between the mechanisms that involve either one or two

catalytic amino acids. The formation of the covalent glycosyl–enzyme intermediate CGE via the reaction pathway R → TS1 → CGE corresponds to the mechanism involving only one catalytic base (noted B^- on Scheme 1) in the reaction mechanism. In the case of a mechanism involving two catalytic residues, two distinct pathways lead to the covalent glycosyl–enzyme intermediate CGEP. They differ, however, in the sequence of the individual steps. In the first pathway (R → TS1 → CGE → TS2 → CGEP), the enzymatic reaction starts with the nucleophilic attack (along the horizontal axis) of the catalytic base oxygen OB on the anomeric carbon C-1 of the galactosyl residue, followed by the proton (HA) transfer (along the vertical axis) from the catalytic acid A to the glycosidic oxygen O-1. In the second pathway (R → TS3 → INT2 → TS4 → CGEP), the order of the steps is reversed with the proton transfer occurring before the nucleophilic attack. Comparison of the energy barriers required to proceed along these two pathways (Fig. 3(a)) reveals that the process starting with the nucleophilic attack is less energy demanding. As a matter of fact, the process starting with the proton transfer appears not only energetically unfavorable, but if no constraints are placed on the position of the proton HA, after transfer to O-1 (r_{HA-O1}), the proton goes back to its initial

Table 3
Comparison of the ab initio relative energies (kcal/mol) calculated by various methods for the points observed on the PES described on Fig. 1(a) (first step of the reaction)

Geometry energy	B3LYP/6-31G*6-31G*	B3LYP/6-31G*6-31++G**	B3LYP/6-31G*6-311++G**
R	0.00 [a]	0.00 [b]	0.00 [c]
CGE	9.02	11.66	9.48
INT2	46.38	42.52	41.96
TS1	16.89	16.85	16.52
TS2	31.88	32.11	31.80
TS3	39.38	38.07	37.88
TS4	46.81	43.54	43.87
CGEP	31.86	30.07	29.90

[a] $E = -2079498.13$ kcal/mol.
[b] $E = -2079615.52$ kcal/mol.
[c] $E = -2080032.47$ kcal/mol.

Table 4
Comparison of the ab initio relative energies (kcal/mol) calculated by various methods for the points observed in the second step of the reaction

Geometry energy	B3LYP/6-31G*6-31G*	B3LYP/6-31G*6-31++G**	B3LYP/6-31G*6-311++G**
Scheme 2(a)			
CGE	0.00 [a]	0.00 [b]	0.00 [c]
INT3	14.47	16.39	16.34
INT4	26.57	26.12	26.87
TS5	16.36	18.33	18.93
TS6	38.28	41.59	40.90
TS7	32.05	30.90	30.82
TS8	27.36	26.55	27.53
P	0.10	1.99	2.32
Scheme 2(c)			
CGE	0.00 [d]	0.00 [e]	0.00 [f]
TS9	10.09	9.39	8.45
Prod	−15.99	−17.54	−17.97

[a] $E = -743025.56$ kcal/mol.
[b] $E = -743105.99$ kcal/mol.
[c] $E = -743292.10$ kcal/mol.
[d] $E = -2008347.60$kcal/mol.
[e] $E = -2008463.96$ kcal/mol.
[f] $E = -2008860.42$ kcal/mol.

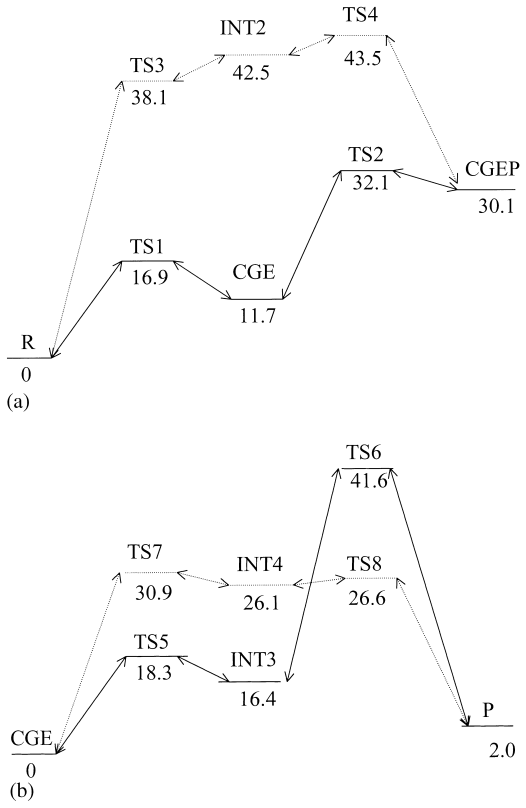


Fig. 3. Schematic energetic representations (in kcal/mol) of the possible reaction pathways observed in the different PES's for the transfer of the galactosyl residue catalyzed by retaining α-galactosyltransferases. (a) First step of the reaction. (b) Second step using a second catalytic base A^- to assist the reaction (Scheme 2(a)). Relative energies are calculated at the DFT/B3LYP/6-31++G**//DFT/B3LYP/6-31G* level.

position on the catalytic acid A. However, with the assistance of a stabilizing hydrogen bonding interaction between the hydroxyl group OH-2 of the galactosyl residue and the catalytic acid A, the transfer of HA to O-1 could be observed. The activation energy of 16.9 kcal/mol calculated for the nucleophilic attack in the first step is slightly higher compared to the value of 13.4 kcal/mol calculated for the inverting *N*-acetylglucosaminyltransferases.[24] This difference probably reflects a change in the nucleophilic character of the attacking groups. Geometrical changes occurring along the reaction process are shown on Fig. 1(b). Analysis of these changes revealed that along the reaction path: R → TS1 (16.9 kcal/mol) → CGE (11.7 kcal/mol) the C-1–O-1 bond length between the anomeric carbon and the leaving group gradually elongates from 1.493 to 3.422 Å as the distance between the anomeric oxygen and the attacking catalytic base r_{C1-OB} decreases (Table 1). The transition state for the nucleophilic attack occurring as the first step in the reaction, TS1, undergoes significant geometrical changes compared to the starting structure, R. As the C-1–OB reaction coordinate gets close to 2.2 Å as in TS1, the length of the C-1–O-1 scissile bond increases slightly by 0.15 Å going from 1.493 to 1.640 Å and the C-1–O-5 shortens from 1.382 to 1.366 Å. Geometrical changes are larger going from TS1 to CGE than from R to TS1. Indeed, in CGE the C-1–OB, C-1–O-1 and C-1–O-5 bonds have values of 1.576, 3.476, and 1.482 Å, respectively. Geometrical changes between the two covalent glycosyl–enzyme intermediates, CGE and CGEP, are minor. This indicates only a negligible effect of the proton transfer on these structures.

As for the previously calculated reaction mechanism of inverting *N*-acetylglucosaminyltransferases,[24] these results suggest that the proton transfer to the glycosidic oxygen is not required to the breaking of the C-1–O-1 bond. Therefore, in contrast to glycosyl hydrolases, it seems that glycosyltransferases do not need a second catalytic residue in their active site for the reaction to proceed. In fact, none of the solved crystal structure of glycosyltransferases, either inverting or retaining has so far revealed the presence of a second amino acid in the active site of the enzyme. One explanation resides probably in the significantly better leaving group character of a nucleotide compared to a glycosyl residue.

3.2. Displacement step

In the second step (Scheme 2(a–c)), the reactive hydroxyl of the acceptor sugar (HM–OM) attacks at the anomeric center of the glycosyl–enzyme intermediate (CGE) to displace the sugar off the enzyme, resulting in a product with the retained configuration. A general base is needed to deprotonate the reactive hydroxyl. In light of this, three possible bases can be considered; the catalytic base A^-; the oxygen of the catalytic acid B; and the oxygen of UDP. All three potential candidates were investigated and calculated separately. The classical double-displacement mechanism is represented on Scheme 2(a), where a second catalytic acid functioning as the catalytic base A^- is present in the model. Though this mechanism is supported by none of the available experimental data nor by our theoretical calculations, we cannot definitely rule out the presence of such a catalytic base in the active site of retaining glycosyltransferases not yet structurally solved. Indeed, only four out of many retaining glycosyltransferases have had their crystallographic structures determined. Therefore, we have also explored the possibility of having a second catalytic residue in the active site (Scheme 2(a)). In addition to making this study more complete, it will also shed some light on the proton-transfer mechanism.

The PES corresponding to the second step of the reaction mechanism employing a second catalytic base and calculated at the HF/6-31G* level is given in Fig. 2(a). The distances r_{C1-OM} and r_{HM-OM} along both axes of the contour map characterize horizontally the nucle-

ophilic attack of the methanol oxygen OM on the anomeric carbon C-1 of the galactosyl residue and, vertically, the proton (HM) transfer from the methanol to the catalytic base oxygen (OA). The PES shows two intermediates (INT3 and INT4) and four energy barriers (TS5–TS8) located along the borders of the contour map. Again, this map indicates that a concerted mechanism is improbable, leaving a stepwise mechanism as the most likely to occur. In this second step-reaction model, the starting molecule is the covalent glycosyl–enzyme intermediate CGE that has previously been formed in the first step of the reaction, and where UDP has left the reactive site to allow the acceptor molecule to attack the reactive center C-1 of the galactosyl moiety. Two distinct pathways can be seen on the PES in order to form the final product (P). In the first pathway (CGE → TS7 → INT4 → TS8 → P), the enzymatic reaction proceeds first with the nucleophilic attack (along the horizontal axis) of the methanol oxygen OM on the anomeric carbon C-1 of the galactosyl residue, followed by the proton (HM) transfer (along the vertical axis) from the methanol to the catalytic acid oxygen (OA). In the second possible pathway (CGE → TS5 → INT3 → TS6 → P), the order of the steps is reversed. Geometrical changes observed along the reaction process are displayed on Fig. 2(b). Like for the first step of the reaction (Fig. 1(b)), the main geometrical changes observed are connected with the nucleophilic attack occurring along the horizontal axis of the contour map. The C-1–OB bond length between the anomeric carbon and the catalytic base B in CGE gradually elongates from 1.424 to 2.902 Å, as the distance between the attacking methanol oxygen OM and the anomeric carbon C-1, r_{C1-OM}, decreases (Table 2). During the whole process, the C-1–O-5 bond in the galactopyranosyl ring varies sensibly, starting by shortening from 1.426 Å in CGE to 1.377 Å in INT4 after the nucleophilic attack, and it finally goes back to its initial value of 1.429 Å in P after the proton transfer has occurred.

The PES exhibits features similar to the map corresponding to the first step (Fig. 1(a)) or to the map calculated for inverting *N*-acetylglucosaminyltransferases.[24] However, their comparison revealed very relevant differences. Indeed, it appears for this mechanism that the reaction pathway with the proton transfer occurring as the first step is energetically favored over the reaction pathway where the nucleophilic attack at the anomeric carbon is the first step. The energy barrier for the proton transfer as the starting process in this reaction is in this case considerably lower (CGE → TS5 (18.3 kcal/mol) → INT3 (16.4 kcal/mol)) than the one beginning with the nucleophilic attack (CGE → TS7 (30.9 kcal/mol) → INT4 (26.1 kcal/mol)). This implies that the proton transfer is the key process in the second leg of a double-displacement mechanism.

For enzymes without a second catalytic residue in the active site, one can assume an alternative mechanism for the second step. Such an alternative mechanism would require another suitable general base to allow the proton transfer from the acceptor hydroxyl to the base. As potential candidates for this function, we have considered the oxygen of the catalytic acid B and the oxygen of UDP. The mechanism using the catalytic acid B as the potential nucleophile is quite interesting, as this would assist the breaking of the C-1–OB bond in CGE and thus facilitate the displacement of the catalytic base by a new glycosyl residue. This mechanism was evaluated by calculating a two-dimensional PES as a function of the C-1–OM and HM–O distances (Scheme 2(b)). The DFT/B3LYP/6-31G* calculated PES (data not shown) indicated only one valley along the diagonal of the contour map what clearly suggests a concerted mechanism. The energy barrier calculated for this pathway was over 65 kcal/mol. This very high-energy barrier denotes a very unfavorable mechanism. Furthermore, as we have investigated it, this mechanism would clearly lead to a glycoside product with inverted configuration at C-1, rather than the retained configuration expected by this type of enzymes.

Another potential mechanism, in which the oxygen of the phosphate group functions as the proton acceptor, has been proposed recently for α-(1 → 3)-GalT.[16,21] At start, this mechanism (Scheme 2(c)) did not seem very attractive, given that phosphates are very strong acids. However, this situation might be more favorable in a complex of UDP with the metal cofactor. Some kinetics measurements for retaining glycosyltransferases indicate that UDP is not released before the transfer of the sugar to the acceptor.[30–32] These data would be consistent with this type of mechanism. Interestingly, a proton transfer of this type from the hydroxyl of the acceptor to a phosphate group has also been proposed for the reverse reaction of maltodextrin phosphorylase.[33] The mechanism of the displacement step involving the proton transfer to the oxygen atoms of the diphosphate group was evaluated by calculating the reaction energy surface as a function of the reaction coordinates describing the proton transfer to the phosphate group r_{HM-O1} and the nucleophilic attack r_{C1-OM}, respectively. The calculated PES is given on Fig. 4(a), and barriers are given in Table 4. The horizontal axis represents the nucleophilic attack on the anomeric carbon of CGE. The vertical axis describes the proton transfer from the nucleophile hydroxyl to the UDP oxygen. The PES clearly shows the presence of a single transition state (TS9) in the central region of the map and on the diagonal going from reactants to products. The presence of only one transition state and a maximum at two corners of the map indicates the occurrence of a quasi-concerted mechanism, in which the formation of the C-1–OM bond and the proton transfer

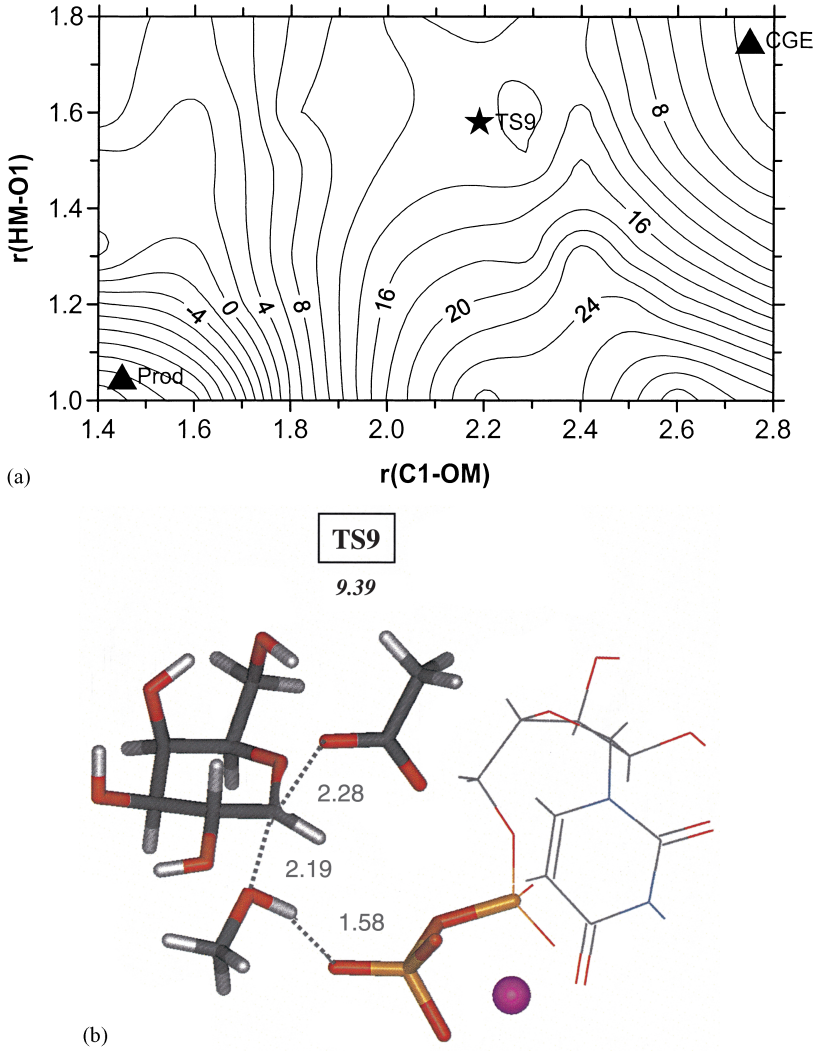


Fig. 4. (a) PES calculated at the HF/6-31G* level and corresponding to the second step (displacement step) and using UDP to assist the reaction (Scheme 2(c)). CGE, TS9, and Prod represent the glycosyl–enzyme intermediate, transition state, and products, respectively. (b) Geometrical representation of the transition state TS9 calculated at the DFT/B3LYP/6-31G* level. Numbers in italics represent relative energies (in kcal/mol) at the DFT/B3LYP/6-31++G**//DFT/B3LYP/6-31G* level.

from the acceptor to UDP occur simultaneously. In fact, the reaction pathway begins with the nucleophilic attack of the acceptor hydroxyl on the anomeric carbon of CGE. The nucleophilic hydroxyl group is being activated by interactions with the negatively charged UDP oxygen. As the hydroxyl approaches the anomeric carbon, the proton becomes more acidic and at the C-1–OM distance of 2.190 Å, the proton transfer to a general base, represented by the UDP oxygen, occurs. The proton transfer takes place on a very short range of C-1–OM distance as it is demonstrated by a steep PES in this region (Fig. 4(a)). Subsequently, the proton remains located at the UDP oxygen what might facilitate the release of UDP by decreasing its negative charge. The structure of the transition state TS9 (Fig. 4(b)) is represented by a distorted $^4C_1$ chair conformation of the galactopyranose ring; and C-1–OM and C-1–OB bond lengths of about 2.190 and 2.279 Å, respectively. The proton is located at an OM–HM distance of about 1.022 Å. After the proton is transferred to O-1, the pathway indicates the nucleophilic attack of the deprotonated hydroxyl until the CGE is cleaved off and a new glycosidic bond is formed. In the product (Prod), the anomeric carbon has the α configuration. A comparison of the transition state structures associated with the preferred reaction pathways for the two steps of the catalytic reaction revealed some interesting differences compared to inverting enzymes. In the first step of the retaining mechanism, R → TS1 → CGE, the ring geometry appears to be less affected along the reaction compared to inverting enzymes. The ring conformation proceeds through a distorted $^4C_1$ chair. This is illustrated by the values of C-1–O-5–C-5 and C-1–O-5–C-5–C-4 angles in Table 1. Along the second step of the mechanism, the alteration of the ring shape is clearly more pronounced. This is reflected by

the large variations in the ring geometry, C-1–O-5–C-5 and C-1–O-5–C-5–C-4 angles in Table 2. The conformation of the galactosyl–enzyme complex, CGE, resembles a standard conformation found for simple glycosides in accordance with the exo-anomeric effect.[34] It is noteworthy that the conformation about the axial anomeric bond does not represent the lowest energy conformation for this linkage. We believe that the sterical constraints used in the model to represent the enzyme active site are responsible for this unusual conformation. On the other hand, one can expect further stabilization of the product during its release from the active side by adopting the preferred conformation around C-1–O-1 in agreement with the exo-anomeric effect.[34] The calculated barrier for this reaction process at the DFT/B3LYP/6-31++G**//DFT/B3LYP 6-31G* level is about 9.4 kcal/mol, which is considerably lower than the energy barriers calculated for the process involving any of the other potential general bases considered in this study. All these results support the participation of UDP as a base in the displacement step of a double-displacement mechanism of retaining glycosyltransferases.

In the framework of a double-displacement mechanism, the following reaction mechanism whose reaction pathway is illustrated in Fig. 5 can be proposed based on our calculations. Relative energies of all stationary points along this pathway were calculated at the DFT/B3LYP/6-31++G**//DFT/B3LYP/6-31G* level in order to estimate the effect of the basis set (Tables 3 and 4, Fig. 5). The first step consists of a nucleophilic attack of the catalytic base on the anomeric carbon that formed the covalent glycosyl–enzyme complex (CGE) with inverted (β) configuration at C-1. The calculated barrier for this step is 16.5 kcal/mol. The hydroxyl group of the acceptor then attacks the CGE with the UDP functioning as a general base that deprotonates hydroxyl. This step proceeds in a quasi-concerted manner with the second inversion at the anomeric center with the resulting glycoside having the α configuration. The barrier for this step is about 8.5 kcal/mol. The reaction barrier for the proposed mechanism (18.0 kcal/mol on Fig. 5) lies in the range of experimentally observed barriers for glycosyltransferases of 15–25 kcal/mol.[24,35]

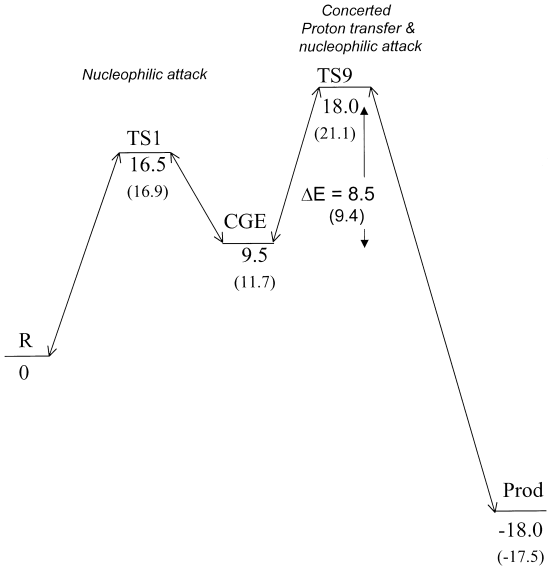


Fig. 5. Schematic energetic representation (in kcal/mol) of the possible reaction pathway observed in the different PES's for the transfer of the galactosyl residue catalyzed by retaining α-galactosyltransferases using UDP to assist the reaction (Scheme 2(c)). Relative energies are calculated at the DFT/B3LYP/6-311++G**//DFT/B3LYP/6-31G* and DFT/B3LYP/6-31++G**//DFT/B3LYP/6-31G* (in parenthesis) levels.

In conclusion, the first quantum mechanical calculations of the galactosyl residue transfer catalyzed by retaining α-galactosyltransferases provide an insight on the double-displacement mechanism of these enzymes. Structures of the covalent galactose–enzyme intermediate and transition states that we have determined can be employed as guides in rational design of inhibitors.[36] We have to emphasize that in the light of recent interesting results,[15,32] the double-displacement mechanism studied in this work might only represent one among several possible mechanisms followed by retaining glycosyltransferases. The existence of additional possible mechanisms to catalyze the glycosyl transfer is supported by the crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis*[15,32] where the key catalytic amino acids could not be identified within 5 Å of the reactive center C-1. This enzyme is believed to proceed through a single front-side displacement reaction, also known as an S_Ni mechanism, rather than via a double-displacement mechanism. Very recently published, the crystal structure of OtsA glucosyltransferase[20] seems also to suggest this type of mechanistic machinery, common to glycogen phosphorylases.[5] However, at this time the lack of experimental evidence does not permit one to yet conclude which of the two mechanisms are operative for these enzymes. Therefore, while we await further experimental data, alternative mechanisms of the retaining glycosyltransferases are currently being investigated in our laboratories and will be presented in future communications.

Acknowledgements

The authors wish to thank Professor S.G. Withers (Dept. of Chemistry, University of British Columbia, Vancouver, Canada) for his valuable discussion on the mechanism of retaining glycosyltransferases.

References

1. Leloir, L. F. *Science* **1971**, *172*, 1299–1303.
2. Coutinho, P.M.; Henrissat, B. Carbohydrate-Active Enzymes server at URL: http://afmb.cnrs-mrs.fr/CAZY.
3. Davies, G.; Henrissat, B. *Structure* **1995**, *3*, 853–859.
4. McCarter, J. D.; Withers, S. G. *Curr. Opin. Struct. Biol.* **1994**, *4*, 885–892.
5. Sinnott, M. L. *Chem. Rev.* **1990**, *90*, 1171–1202.
6. Davies, G.; Sinnott, M. L.; Withers, S. G. In *Comprehensive Biological Catalysis*; Sinott, M. L., Ed.; Academic Press Limited: New York, 1997; pp 119–208.
7. Zechel, D. L.; Withers, S. G. *Acc. Chem. Res.* **2000**, *33*, 11–18.
8. Vrielink, A.; Ruger, W.; Driessen, H. P. C.; Freemont, P. S. *EMBO J.* **1994**, *13*, 3413–3422.
9. Charnock, S. J.; Davies, G. J. *Biochemistry* **1999**, *38*, 6380–6385.
10. Gastinel, L. N.; Cambillau, C.; Bourne, Y. *EMBO J.* **1999**, *18*, 3546–3557.
11. Morera, S.; Imberty, A.; Aschke-Sonnenborn, U.; Ruger, W.; Freemont, P. S. *J. Mol. Biol.* **1999**, *292*, 717–730.
12. Ha, S.; Walker, D.; Shi, Y.; Walker, S. *Protein Sci.* **2000**, *9*, 1045–1052.
13. Pedersen, L. C.; Tsuchida, K.; Kitagawa, H.; Sugahara, K.; Darden, T. A.; Negishi, M. *J. Biol. Chem.* **2000**, *275*, 34580–34585.
14. Unligil, U. M.; Zhou, S.; Yuwaraj, S.; Sarkar, M.; Schachter, H.; Rini, J. M. *EMBO J.* **2000**, *19*, 5269–5280.
15. Persson, K.; Ly, H. D.; Dieckelmann, M.; Wakarchuk, W. W.; Whiters, S. G.; Strynadka, N. C. J. *Nat. Struct. Biol.* **2000**, *8*, 166–175.
16. Gastinel, L. N.; Bignon, C.; Misra, A. K.; Hindsgaul, O.; Shaper, J. H.; Joziasse, D. H. *EMBO J.* **2001**, *20*, 638–649.
17. Mulichak, A. M.; Losey, H. C.; Walsh, C. T.; Garavito, R. M. *Structure* **2001**, *9*, 547–557.
18. Patenaude, S. I.; Seto, N. O.; Borisova, S. N.; Szpacenko, A.; Marcus, S. L.; Palcic, M. M.; Evans, S. V. *Nat. Struct. Biol.* **2002**, *9*, 685–690.
19. Gibbons, B. J.; Roach, O. J.; Hurley, T. D. *J. Mol. Biol.* **2002**, *319*, 463–477.
20. Gibson, R. P.; Turkenburg, J. P.; Charnock, S. J.; Lloyd, R.; Davies, G. *J. Chem. Biol.* **2002**, *9*, 1337–1346.
21. Boix, E.; Zhang, Y.; Swaminathan, G. J.; Brew, K.; Acharya, K. R. *J. Biol. Chem.* **2002**, *277*, 28310–28318.
22. Unligil, U. M.; Rini, J. M. *Curr. Opin. Struct. Biol.* **2000**, *10*, 510–517.
23. Tarbouriech, N.; Charnock, S. J.; Davies, G. J. *J. Mol. Biol.* **2001**, *314*, 655–661.
24. Tvaroska, I.; André, I.; Carver, J. P. *J. Am. Chem. Soc.* **2000**, *122*, 8762–8776.
25. JAGUAR 3.5, S., Inc.: Portland, OR, 1998.
26. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
27. Tvaroska, I.; André, I.; Carver, J. P. *J. Mol. Struct. (Theochem.)* **1999**, *469*, 103–114.
28. Tvaroska, I.; André, I.; Carver, J. P. *J. Phys. Chem., Sect. B* **1999**, *103*, 2560–2569.
29. André, I.; Tvaroska, I.; Carver, J. P. *J. Phys. Chem., Sect. A* **2000**, *104*, 4609–4617.
30. Zhang, Y.; Wang, P. G.; Brew, K. *J. Biol. Chem.* **2001**, *276*, 11567–11574.
31. Ly, H. D.; Lougheed, B.; Wakarchuk, W. W.; Withers, S. G. *Biochemistry* **2002**, *41*, 5075–5085.
32. Lougheed, B.; Ly, H. D.; Wakarchuk, W. W.; Withers, S. G. *J. Biol. Chem.* **1999**, *274*, 37717–37722.
33. Geremia, S.; Campagnolo, M.; Schinzel, R.; Johnson, L. N. *J. Mol. Biol.* **2002**, *322*, 413–423.
34. Tvaroska, I.; Bleha, T. *Adv. Carbohydr. Chem. Biochem.* **1989**, *47*, 45–123.
35. Seto, N. O.; Compston, C. A.; Evans, S. V.; Bundle, D. R.; Narang, S. A.; Palcic, M. M. *Eur. J. Biochem.* **1999**, *259*, 770–775.
36. Compain, P.; Martin, O. R. *Bioorg. Med. Chem.* **2001**, *9*, 3077–3092.